PART C
LÉVY PROCESSES AND FINANCE

Matthias Winkel\textsuperscript{1}
University of Oxford

\textsuperscript{1}Departmental lecturer (Institute of Actuaries and Aon Lecturer in Statistics) at the Department of Statistics, University of Oxford
Prerequisites

Part A Probability is a prerequisite. BS3a/OBS3a Applied Probability or B10 Martingales and Financial Mathematics would be useful, but are by no means essential; some material from these courses will be reviewed without proof.

Aims

Lévy processes form a central class of stochastic processes, contain both Brownian motion and the Poisson process, and are prototypes of Markov processes and semimartingales. Like Brownian motion, they are used in a multitude of applications ranging from biology and physics to insurance and finance. Like the Poisson process, they allow to model abrupt moves by jumps, which is an important feature for many applications. In the last ten years Lévy processes have seen a hugely increased attention as is reflected on the academic side by a number of excellent graduate texts and on the industrial side realising that they provide versatile stochastic models of financial markets. This continues to stimulate further research in both theoretical and applied directions. This course will give a solid introduction to some of the theory of Lévy processes as needed for financial and other applications.

Synopsis


Special cases of increasing Lévy processes (subordinators) and processes with only positive jumps. Subordination. Examples and applications. Financial models driven by Lévy processes. Stochastic volatility. Level passage problems. Applications: option pricing, insurance ruin, dams.


Reading

- J.F.C. Kingman: *Poisson processes*. Oxford University Press (1993), Ch.1-5, 8
- A.E. Kyprianou: *Introductory lectures on fluctuations of Lévy processes with Applications*. Springer (2006), Ch. 1-3, 8-9

Further reading

- J. Bertoin: *Lévy processes*. Cambridge University Press (1996), Sect. 0.1-0.6, I.1, III.1-2, VII.1
- K. Sato: *Lévy processes and infinite divisibility*. Cambridge University Press (1999), Ch. 1-2, 4, 6, 9
Lecture 1

Introduction

Reading: Kyprianou Chapter 1
Further reading: Sato Chapter 1, Schoutens Sections 5.1 and 5.3

In this lecture we give the general definition of a Lévy process, study some examples of Lévy processes and indicate some of their applications. By doing so, we will review some results from BS3a Applied Probability and B10 Martingales and Financial Mathematics.

1.1 Definition of Lévy processes

Stochastic processes are collections of random variables $X_t$, $t \geq 0$ (meaning $t \in [0, \infty)$ as opposed to $n \geq 0$ by which means $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$). For us, all $X_t$, $t \geq 0$, take values in a common state space, which we will choose specifically as $\mathbb{R}$ (or $[0, \infty)$ or $\mathbb{R}^d$ for some $d \geq 2$). We can think of $X_t$ as the position of a particle at time $t$, changing as $t$ varies. It is natural to suppose that the particle moves continuously in the sense that $t \mapsto X_t$ is continuous (with probability 1), or that it has jumps for some $t \geq 0$:

$$\Delta X_t = X_{t+} - X_{t-} = \lim_{\varepsilon \downarrow 0} X_{t+\varepsilon} - \lim_{\varepsilon \downarrow 0} X_{t-\varepsilon}.$$

We will usually suppose that these limits exist for all $t \geq 0$ and that in fact $X_{t+} = X_t$, i.e. that $t \mapsto X_t$ is right-continuous with left limits $X_{t-}$ for all $t \geq 0$ almost surely. The path $t \mapsto X_t$ can then be viewed as a random right-continuous function.

**Definition 1 (Lévy process)** A real-valued (or $\mathbb{R}^d$-valued) stochastic process $X = (X_t)_{t \geq 0}$ is called a Lévy process if

(i) the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent for all $n \geq 1$ and $0 \leq t_0 < t_1 < \ldots < t_n$ (independent increments),

(ii) $X_{t+s} - X_t$ has the same distribution as $X_s$ for all $s, t \geq 0$ (stationary increments),

(iii) the paths $t \mapsto X_t$ are right-continuous with left limits (with probability 1).

It is implicit in (ii) that $\mathbb{P}(X_0 = 0) = 1$ (choose $s = 0$).
Figure 1.1: Variance Gamma process and a Lévy process with no positive jumps

Here the independence of $n$ random variables is understood in the following sense:

**Definition 2 (Independence)** Let $Y^{(j)}$ be an $\mathbb{R}^{d_j}$-valued random variable for $j = 1, \ldots, n$. The random variables $Y^{(1)}, \ldots, Y^{(n)}$ are called independent if, for all (Borel measurable) $C^{(j)} \subset \mathbb{R}^{d_j}$

$$
\mathbb{P}(Y^{(1)} \in C^{(1)}, \ldots, Y^{(n)} \in C^{(n)}) = \mathbb{P}(Y^{(1)} \in C^{(1)}) \ldots \mathbb{P}(Y^{(n)} \in C^{(n)}). 
$$

(1)

An infinite collection $(Y^{(j)})_{j \in J}$ is called independent if $Y^{(j_1)}, \ldots, Y^{(j_n)}$ are independent for every finite subcollection. Infinite-dimensional random variables $(Y^{(1)}_i)_{i \in I_1}, \ldots, (Y^{(n)}_i)_{i \in I_n}$ are called independent if $(Y^{(1)}_i)_{i \in F_1}, \ldots, (Y^{(n)}_i)_{i \in F_n}$ are independent for all finite $F_j \subset I_j$.

It is sufficient to check (1) for rectangles of the form $C^{(j)} = (a^{(j)}_1, b^{(j)}_1] \times \ldots \times (a^{(j)}_{d_j}, b^{(j)}_{d_j}]$.

### 1.2 First main example: Poisson process

Poisson processes are Lévy processes. We recall the definition as follows. An $\mathbb{N}(\subset \mathbb{R})$-valued stochastic process $X = (X_t)_{t \geq 0}$ is called a Poisson process with rate $\lambda \in (0, \infty)$ if $X$ satisfies (i)-(iii) and

(iv)$^{\text{Poi}} \quad \mathbb{P}(X_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \ k \geq 0, \ t \geq 0$ (Poisson distribution).

The Poisson process is a continuous-time Markov chain. We will see that all Lévy processes have a Markov property. Also recall that Poisson processes have jumps of size 1 (spaced by independent exponential random variables $Z_n = T_{n+1} - T_n, \ n \geq 0$, with parameter $\lambda$, i.e. with density $\lambda e^{-\lambda s}, \ s \geq 0$). In particular, $\{t \geq 0 : \Delta X_t \neq 0\} = \{T_n, n \geq 1\}$ and $\Delta X_{T_n} = 1$ almost surely (short a.s., i.e. with probability 1). We can define more general Lévy processes by putting

$$
C_t = \sum_{k=1}^{X_t} Y_k, \quad t \geq 0,
$$

for a Poisson process $(X_t)_{t \geq 0}$ and independent identically distributed $Y_k, \ k \geq 1$. Such processes are called compound Poisson processes. The term “compound” stems from the representation $C_t = S \circ X_t = S_{X_t}$ for the random walk $S_n = Y_1 + \ldots + Y_n$. You may think of $X_t$ as the number of claims up to time $t$ and of $Y_k$ as the size of the $k$th claim. Recall (from BS3a) that its moment generating function, if it exists, is given by

$$
\mathbb{E}(\exp\{\gamma C_t\}) = \exp\{-\lambda t(\mathbb{E}(e^{\gamma Y_1} - 1))\}.
$$

This will be an important building block of a general Lévy process.
1.3 Second main example: Brownian motion

Brownian motion is a Lévy process. We recall (from B10b) the definition as follows. An \( \mathbb{R} \)-valued stochastic process \( X = (X_t)_{t \geq 0} \) is called Brownian motion if \( X \) satisfies (i)-(ii) and

\[(iii)_{\text{BM}} \text{ the paths } t \mapsto X_t \text{ are continuous almost surely,}\]

\[(iv)_{\text{BM}} \mathbb{P}(X_t \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{y^2}{2t} \right\} dy, \quad x \in \mathbb{R}, \ t > 0. \ (\text{Normal distribution}).\]

The paths of Brownian motion are continuous, but turn out to be nowhere differentiable (we will not prove this). They exhibit erratic movements at all scales. This makes Brownian motion an appealing model for stock prices. Brownian motion has the scaling property \((\sqrt{c}X_t/c)_{t \geq 0} \sim X\) where "\(\sim\)" means "has the same distribution as".

Brownian motion will be the other important building block of a general Lévy process. The canonical space for Brownian paths is the space \( C([0, \infty), \mathbb{R}) \) of continuous real-valued functions \( f : [0, \infty) \rightarrow \mathbb{R} \) which can be equipped with the topology of locally uniform convergence, induced by the metric

\[d(f, g) = \sum_{k \geq 1} 2^{-k} \min\{d_k(f, g), 1\}, \text{ where } d_k(f, g) = \sup_{x \in [0, k]} |f(x) - g(x)|.\]

This metric topology is complete (Cauchy sequences converge) and separable (has a countable dense subset), two attributes important for the existence and properties of limits. The bigger space \( D([0, \infty), \mathbb{R}) \) of right-continuous real-valued functions with left limits can also be equipped with the topology of locally uniform convergence. The space is still complete, but not separable. There is a weaker metric topology, called Skorohod’s topology, that is complete and separable. In the present course we will not develop this and only occasionally use the familiar uniform convergence for (right-continuous) functions \( f, f_n : [0, k] \rightarrow \mathbb{R}, \ n \geq 1: \)

\[\sup_{x \in [0, k]} |f_n(x) - f(x)| \rightarrow 0, \quad \text{as } n \rightarrow \infty,\]

which for stochastic processes \( X, X^{(n)}, \ n \geq 1, \) with time range \( t \in [0, T] \) takes the form

\[\sup_{t \in [0, T]} |X_t^{(n)} - X_t| \rightarrow 0, \quad \text{as } n \rightarrow \infty,\]

and will be as a convergence in probability or as almost sure convergence (from BS3a or B10a) or as \( L^2 \)-convergence, where \( Z_n \rightarrow Z \) in the \( L^2 \)-sense means \( \mathbb{E}(|Z_n - Z|^2) \rightarrow 0. \)
1.4 Markov property

The Markov property is a consequence of the independent increments property (and the stationary increments property):

**Proposition 3 (Markov property)** Let $X$ be a Lévy process and $t \geq 0$ a fixed time, then the pre-$t$ process $(X_r)_{r \leq t}$ is independent of the post-$t$ process $(X_t+s-X_t)_{s \geq 0}$, and the post-$t$ process has the same distribution as $X$.

**Proof:** By Definition 2, we need to check the independence of $(X_{r_1}, \ldots, X_{r_n})$ and $(X_{t+s_1}-X_t, \ldots, X_{t+s_m}-X_t)$. By property (i) of the Lévy process, we have that increments over disjoint time intervals are independent, in particular the increments

$X_{r_1}, X_{r_2} - X_{r_1}, \ldots, X_{r_n} - X_{r_{n-1}}, X_{t+s_1} - X_t, X_{t+s_2} - X_{t+s_1}, \ldots, X_{t+s_m} - X_{t+s_{m-1}}$.

Since functions (here linear transformations from increments to marginals) of independent random variables are independent, the proof of independence is complete. Identical distribution follows first on the level of single increments from (ii), then by (i) and linear transformation also for finite-dimensional marginal distributions.

1.5 Some applications

**Example 4 (Insurance ruin)** A compound Poisson process $(Z_t)_{t \geq 0}$ with positive jump sizes $A_k$, $k \geq 1$, can be interpreted as a claim process recording the total claim amount incurred before time $t$. If there is linear premium income at rate $r > 0$, then also the gain process $rt - Z_t$, $t \geq 0$, is a Lévy process. For an initial reserve of $u > 0$, the reserve process $u + rt - Z_t$ is a shifted Lévy process starting from a non-zero initial value $u$.

**Example 5 (Financial stock prices)** Brownian motion $(B_t)_{t \geq 0}$ or linear Brownian motion $\sigma B_t + \mu t$, $t \geq 0$, was the first model of stock prices, introduced by Bachelier in 1900. Black, Scholes and Merton studied geometric Brownian motion $\exp(\sigma B_t + \mu t)$ in 1973, which is not itself a Lévy process but can be studied with similar methods. The Economics Nobel Prize 1997 was awarded for their work. Several deficiencies of the Black-Scholes model have been identified, e.g. the Gaussian density decreases too quickly, no variation of the volatility $\sigma$ over time, no macroscopic jumps in the price processes. These deficiencies can be addressed by models based on Lévy processes. The Variance gamma model is a time-changed Brownian motion $B_T$ by an independent increasing jump process, a so-called Gamma Lévy process with $T_s \sim \text{Gamma}(\alpha s, \beta)$. The process $B_T$ is then also a Lévy process itself.

**Example 6 (Population models)** Branching processes are generalisations of birth-and-death processes (see BS3a) where each individual in a population dies after an exponentially distributed lifetime with parameter $\mu$, but gives birth not to single children, but to twins, triplets, quadruplet etc. To simplify, it is assumed that children are only born at the end of a lifetime. The numbers of children are independent and identically distributed according to an offspring distribution $q$ on $\{0, 2, 3, \ldots\}$. The population size process $(Z_t)_{t \geq 0}$ can jump downwards by 1 or upwards by an integer. It is not a Lévy process but is closely related to Lévy processes and can be studied with similar methods. There are also analogues of processes in $[0, \infty)$, so-called continuous-state branching processes that are useful large-population approximations.
Lecture 2

Lévy processes and random walks

Reading: Kingman Section 1.1, Grimmett and Stirzaker Section 3.5(4)
Further reading: Sato Section 7, Durrett Sections 2.8 and 7.6, Kallenberg Chapter 15

Lévy processes are the continuous-time analogues of random walks. In this lecture we examine this analogy and indicate connections via scaling limits and other limiting results. We begin with a first look at infinite divisibility.

2.1 Increments of random walks and Lévy processes

Recall that a random walk is a stochastic process in discrete time

\[ S_0 = 0, \quad S_n = \sum_{j=1}^{n} A_j, \quad n \geq 1, \]

for a family \((A_j)_{j \geq 1}\) of independent and identically distributed real-valued (or \(\mathbb{R}^d\)-valued) random variables. Clearly, random walks have stationary and independent increments. Specifically, the \(A_j, j \geq 1\), themselves are the increments over single time units. We refer to \(S_{n+m} - S_n\) as an increment over \(m\) time units, \(m \geq 1\).

While every distribution may be chosen for \(A_j\), increments over \(m\) time units are sums of \(m\) independent and identically distributed random variables, and not every distribution has this property. This is not a deep observation, but it becomes important when moving to Lévy processes. In fact, the increment distribution of Lévy processes is restricted: any increment \(X_{t+s} - X_t\), or \(X_s\) for simplicity, can be decomposed, for every \(m \geq 1\),

\[ X_s = \sum_{j=1}^{m}(X_{js/m} - X_{(j-1)s/m}) \]

into a sum of \(m\) independent and identically distributed random variables.

Definition 7 (Infinite divisibility) A random variable \(Y\) is said to have an infinitely divisible distribution if for every \(m \geq 1\), we can write

\[ Y \sim Y_1^{(m)} + \ldots + Y_m^{(m)} \]

for some independent and identically distributed random variables \(Y_1^{(m)}, \ldots, Y_m^{(m)}\).

We stress that the distribution of \(Y_j^{(m)}\) may vary as \(m\) varies, but not as \(j\) varies.
The argument just before the definition shows that increments of Lévy processes are infinitely divisible. Many known distributions are infinitely divisible, some are not.

**Example 8** The Normal, Poisson, Gamma and geometric distributions are infinitely divisible. This often follows from the closure under convolutions of the type

\[ Y_1 \sim \text{Normal}(\mu, \sigma^2), Y_2 \sim \text{Normal}(\nu, \tau^2) \implies Y_1 + Y_2 \sim \text{Normal}(\mu + \nu, \sigma^2 + \tau^2) \]

for independent \( Y_1 \) and \( Y_2 \) since this implies by induction that for independent

\[ Y_1^{(m)}, \ldots, Y_m^{(m)} \sim \text{Normal}(\mu/m, \sigma^2/m) \implies Y_1^{(m)} + \ldots + Y_m^{(m)} \sim \text{Normal}(\mu, \sigma^2). \]

The analogous arguments (and calculations, if necessary) for the other distributions are left as an exercise. The geometric\((p)\) distribution here is \(\mathbb{P}(X = n) = p^n(1 - p), n \geq 0.\)

**Example 9** The Bernoulli\((p)\) distribution, for \( p \in (0, 1), \) is *not* infinitely divisible. Assume that you can represent a Bernoulli\((p)\) random variable \( X \) as \( Y_1 + Y_2 \) for independent identically distributed \( Y_1 \) and \( Y_2.\) Then

\[ \mathbb{P}(Y_1 > 1/2) > 0 \implies 0 = \mathbb{P}(X > 1) \geq \mathbb{P}(Y_1 > 1/2, Y_2 > 1/2) > 0 \]

is a contradiction, so we must have \( \mathbb{P}(Y_1 > 1/2) = 0, \) but then

\[ \mathbb{P}(Y_1 > 1/2) = 0 \implies p = \mathbb{P}(X = 1) = \mathbb{P}(Y_1 = 1/2)\mathbb{P}(Y_2 = 1/2) \implies \mathbb{P}(Y_1 = 1/2) = \sqrt{p}. \]

Similarly,

\[ \mathbb{P}(Y_1 < 0) > 0 \implies 0 = \mathbb{P}(X < 0) \geq \mathbb{P}(Y_1 < 0, Y_2 < 0) > 0 \]

is a contradiction, so we must have \( \mathbb{P}(Y_1 < 0) = 0 \) and then

\[ 1 - p = \mathbb{P}(X = 0) = \mathbb{P}(Y_1 = 0, Y_2 = 0) \implies \mathbb{P}(Y_1 = 0) = \sqrt{1 - p} > 0. \]

This is impossible for several reasons. Clearly, \( \sqrt{p} + \sqrt{1 - p} > 1, \) but also

\[ 0 = \mathbb{P}(X = 1/2) \geq \mathbb{P}(Y_1 = 0)\mathbb{P}(Y_2 = 1/2) > 0. \]

### 2.2 Central Limit Theorem and Donsker’s theorem

**Theorem 10 (Central Limit Theorem)** Let \((S_n)_{n \geq 0}\) be a random walk with \(\mathbb{E}(S_1^2) = \mathbb{E}(A_1^2) < \infty.\) Then, as \( n \to \infty, \)

\[ \frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - n\mathbb{E}(A_1)}{\sqrt{n\text{Var}(A_1)}} \to \text{Normal}(0, 1) \quad \text{in distribution.} \]

This result as a result for one time \( n \to \infty \) can be extended to a convergence of processes, a convergence of the discrete-time process \((S_n)_{n \geq 0}\) to a (continuous-time) Brownian motion, by scaling of both space and time. The processes

\[ \frac{S_{[nt]} - [nt]\mathbb{E}(A_1)}{\sqrt{\text{Var}(A_1)}}, \quad t \geq 0, \]

where \([nt] \in \mathbb{Z} \text{ with } [nt] \leq nt < [nt] + 1\) denotes the integer part of \( nt, \) are scaled versions of the random walk \((S_n)_{n \geq 0},\) now performing \( n \) steps per time unit (holding time \( 1/n), \) centred and each only a multiple \( 1/\sqrt{n\text{Var}(A_1)}\) of the original size. If \( \mathbb{E}(A_1) = 0, \) you may think that you look at \((S_n)_{n \geq 0}\) from further and further away, but note that space and time are scaled differently, in fact so as to yield a non-trivial limit.
Theorem 11 (Donsker) Let \((S_n)_{n \geq 0}\) be a random walk with \(\mathbb{E}(S_1^2) = \mathbb{E}(A_1^2) < \infty\). Then, as \(n \to \infty\),

\[
\frac{S_{[nt]} - [nt] \mathbb{E}(A_1)}{\sqrt{n \text{Var}(A_1)}} \to B_t \quad \text{locally uniformly in } t \geq 0,
\]

"in distribution", for a Brownian motion \((B_t)_{t \geq 0}\).

Proof: [only for \(A_1 \sim \text{Normal}(0,1)\)] This proof is a coupling proof. We are not going to work directly with the original random walk \((S_n)_{n \geq 0}\), but start from Brownian motion \((B_t)_{t \geq 0}\) and define a family of embedded random walks

\[
S_k^{(n)} := B_{k/n}, \quad k \geq 0, n \geq 1.
\]

Then note using in particular \(\mathbb{E}(A_1) = 0\) and \(\text{Var}(A_1) = 1\) that

\[
S_1^{(n)} \sim \text{Normal}(0, 1/n) \sim \frac{S_1 - \mathbb{E}(A_1)}{\sqrt{n \text{Var}(A_1)}},
\]

and indeed

\[
\left( S_{[nt]}^{(n)} \right)_{t \geq 0} \sim \left( \frac{S_{[nt]} - [nt] \mathbb{E}(A_1)}{\sqrt{n \text{Var}(A_1)}} \right)_{t \geq 0}.
\]

To show convergence in distribution for the processes on the right-hand side, it suffices to establish convergence in distribution for the processes on the left-hand side, as \(n \to \infty\).

To show locally uniform convergence we take an arbitrary \(T \geq 0\) and show uniform convergence on \([0, T]\). Since \((B_t)_{0 \leq t \leq T}\) is uniformly continuous (being continuous on a compact interval), we get a.s.

\[
sup_{0 \leq t \leq T} |S_{[nt]} - B_t| \leq \sup_{0 \leq s \leq T : |s-t| \leq 1/n} |B_s - B_t| \to 0
\]
as \(n \to \infty\). This establishes a.s. convergence, which "implies" convergence in distribution for the embedded random walks and for the original scaled random walk. This completes the proof for \(A_1 \sim \text{Normal}(0,1)\). \(\square\)

Note that the almost sure convergence only holds for the embedded random walks \((S_k^{(n)})_{k \geq 0}, n \geq 1\). Since the identity in distribution with the rescaled original random walk only holds for fixed \(n \geq 1\), not jointly, we cannot deduce almost sure convergence in the statement of the theorem. Indeed, it can be shown that almost sure convergence will fail. The proof for general increment distribution is much harder and will not be given in this course. If time permits, we will give a similar coupling proof for another important special case where \(\mathbb{P}(A_1 = 1) = \mathbb{P}(A_1 = -1) = 1/2\), the simple symmetric random walk.
2.3 Poisson limit theorem

The Central Limit Theorem for Bernoulli random variables $A_1, \ldots, A_n$ says that for large $n$, the number of 1s in the sequence is well-approximated by a Normal random variable. In practice, the approximation is good if $p$ is not too small. If $p$ is small, the Bernoulli random variables count rare events, and a different limit theorem is relevant:

**Theorem 12 (Poisson limit theorem)** Let $W_n$ be binomially distributed with parameters $n$ and $p_n = \lambda/n$ (or if $np_n \to \lambda$, as $n \to \infty$). Then we have
\[ W_n \to \text{Poi}(\lambda), \quad \text{in distribution, as } n \to \infty. \]

*Proof:* Just calculate that, as $n \to \infty$,
\[ \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{n(n-1) \cdots (n-k+1) (np_n)^k (1 - np_n)^n}{k!} \xrightarrow{n \to \infty} \frac{\lambda^k}{k!} e^{-\lambda}. \]

**Theorem 13** Suppose that $S_k^{(n)} = A_1^{(n)} + \cdots + A_k^{(n)}$, $k \geq 0$, is the sum of independent Bernoulli($p_n$) random variables for all $n \geq 1$, and that $np_n \to \lambda \in (0, \infty)$. Then
\[ S_k^{(n)} \to N_t \text{ \quad "in the Skorohod sense" as functions of } t \geq 0, \]
\[ S_k^{(n)} \to \text{Poisson } \lambda \text{ \quad "in distribution" as } n \to \infty, \text{ for a Poisson process } (N_t)_{t \geq 0} \text{ with rate } \lambda. \]

The proof of so-called finite-dimensional convergence for vectors $(S_{[nt]}^{(n)}, \ldots, S_{[ntm]}^{(n)})$ is not very hard but not included here. One can also show that the jump times $(T_m^{(n)})_{m \geq 1}$ of $(S_{[nt]}^{(n)})_{t \geq 0}$ converge to the jump times of a Poisson process. E.g.
\[ \mathbb{P}(T_1^{(n)} > t) = (1 - p_n)^{[nt]} = \left(1 - \frac{[nt]p_n}{[nt]}\right)^{[nt]} \to \exp\{-\lambda t\}, \]
for $[nt]/n \to t$ (since $(nt - 1)/n \to t$ and $nt/n = t$) and so $[nt]p_n \to t\lambda$. The general statement is hard to make precise and prove, certainly beyond the scope of this course.

2.4 Generalisations

Infinitely divisible distributions and Lévy processes are precisely the classes of limits that arise for random walks as in Theorems 10 and 12 (respectively 11 and 13) with different step distributions. Stable Lévy processes are ones with a scaling property $(\alpha X_t)_{t \geq 0} \sim X$ for some $\alpha \in \mathbb{R}$. These exist, in fact, for $\alpha \in (0, 2]$. Theorem 10 (and 11) for suitable distributions of $A_1$ (depending on $\alpha$ and where $\mathbb{E}(A_1^2) = \infty$ in particular) then yield convergence in distribution
\[ \frac{S_n - n\mathbb{E}(A_1)}{n^{1/\alpha}} \to \text{stable}(\alpha) \quad \text{for } \alpha \geq 1, \quad \text{or} \quad \frac{S_n}{n^{1/\alpha}} \to \text{stable}(\alpha) \quad \text{for } \alpha \leq 1. \]

**Example 14 (Brownian ladder times)** For a Brownian motion $B$ and a level $r > 0$, the distribution of $T_r = \inf\{t \geq 0 : B_t > r\}$ is $1/2$-stable, see later in the course.

**Example 15 (Cauchy process)** The Cauchy distribution with density $a/(\pi(x^2 + a^2))$, $x \in \mathbb{R}$, for some parameter $c \in \mathbb{R}$ is $1$-stable, see later in the course.
Lecture 3

Spatial Poisson processes

Reading: Kingman 1.1 and 2.1, Grimmett and Stirzaker 6.13, Kyprianou Section 2.2
Further reading: Sato Section 19

We will soon construct the most general nonnegative Lévy process (and then general real-valued ones). Even though we will not prove that they are the most general, we have already seen that only infinitely divisible distributions are admissible as increment distributions, so we know that there are restrictions; the part missing in our discussion will be to show that a given distribution is infinitely divisible only if there exists a Lévy process \( X \) of the type that we will construct such that \( X_1 \) has the given distribution. Today we prepare the construction by looking at spatial Poisson processes, objects of interest in their own right.

3.1 Motivation from the study of Lévy processes

Brownian motion \( (B_t)_{t \geq 0} \) has continuous sample paths. It turns out that \( (\sigma B_t + \mu t)_{t \geq 0} \) for \( \sigma \geq 0 \) and \( \mu \in \mathbb{R} \) is the only continuous Lévy process. To describe the full class of Lévy processes \( (X_t)_{t \geq 0} \), it is vital to study the process \( (\Delta X_t)_{t \geq 0} \) of jumps.

Take e.g. the Variance Gamma process. In Assignment 1.2.(b), we introduce this process as \( X_t = G_t - H_t, \ t \geq 0, \) for two independent Gamma Lévy processes \( G \) and \( H \).

But how do Gamma Lévy processes evolve? We could simulate discretisations (and will do!) and get some feeling for them, but we also want to understand them mathematically. Do they really exist? We have not shown this. Are they compound Poisson processes? Let us look at their moment generating function (cf. Assignment 2.4.):

\[
\mathbb{E}(\exp\{\gamma G_t\}) = \left(\frac{\beta}{\beta - \gamma}\right)^{\alpha t} \exp\left\{\alpha t \int_0^\infty (e^{\gamma x} - 1) \frac{1}{x} e^{-\beta x} dx\right\}.
\]

This is almost of the form of a compound Poisson process of rate \( \lambda \) with non-negative jump sizes \( Y_j, \ j \geq 1, \) that have a probability density function \( h(x) = h_{Y_1}(x), \ x > 0: \)

\[
\mathbb{E}(\exp\{\gamma C_t\}) = \exp\{\lambda t \int_0^\infty (e^{\gamma x} - 1) h(x) dx\}
\]

To match the two expressions, however, we would have to put

\[
\lambda h(x) = \lambda_0 h^{(0)}(x) = \frac{\alpha}{x} e^{-\beta x}, \quad x > 0,
\]
and \( h^{(0)} \) cannot be a probability density function, because \( \frac{\alpha}{x} e^{-\beta x} \) is not integrable at \( x \downarrow 0 \). What we can do is e.g. truncate at \( \varepsilon > 0 \) and specify
\[
\lambda \varepsilon h^{(\varepsilon)}(x) = \frac{\alpha}{x} e^{-\beta x}, \quad x > \varepsilon, \quad h^{(\varepsilon)}(x) = 0, \quad x \leq \varepsilon.
\]
In order for \( h^{(\varepsilon)} \) to be a probability density, we just put \( \lambda \varepsilon = \int_{\varepsilon}^{\infty} \frac{\alpha}{x} e^{-\beta x} dx \), and notice that \( \lambda \varepsilon \to \infty \) as \( \varepsilon \downarrow 0 \). But \( \lambda \varepsilon \) is the rate of the Poisson process driving the compound Poisson process, so jumps are more and more frequent as \( \varepsilon \downarrow 0 \). On the other hand, the average jump size, the mean of the distribution with density \( h^{(\varepsilon)} \) tends to zero, so most of these jumps are very small. In fact, we will see that
\[
G_t = \sum_{s \leq t} \Delta G_s,
\]
as an absolutely convergent series of infinitely (but clearly countably) many positive jump sizes, where \( \Delta G_s \) is a Poisson point process with intensity \( g(x) = \frac{\alpha}{x} e^{-\beta x}, x > 0 \), the collection of random variables
\[
N((a, b] \times (c, d]) = \# \{ t \in (a, b] : \Delta G_t \in (c, d] \}, \quad 0 \leq a < b, 0 < c < d
\]
a Poisson counting measure (evaluated on rectangles) with intensity function \( \lambda(t, x) = g(x), x > 0, t \geq 0 \); the random countable set \( \{ (t, \Delta G_t) : t \geq 0 \text{ and } \Delta C_t \neq 0 \} \) a spatial Poisson process with intensity \( \lambda(t, x) \). Let us now formally introduce these notions.

### 3.2 Poisson counting measures

The essence of one-dimensional Poisson processes \( (N_t)_{t \geq 0} \) is the set of arrival ("event") times \( \Pi = \{T_1, T_2, T_3, \ldots \} \), which is a random countable set. The increment \( N((s, t]) := N_t - N_s \) counts the number of points in \( \Pi \cap (s, t] \). We can generalise this concept to counting measures of random countable subsets on other spaces, say \( \mathbb{R}^d \). Saying directly what exactly (the distribution of) random countable sets is, is quite difficult in general. Random counting measures are a way to describe the random countable sets implicitly.

**Definition 16 (Spatial Poisson process)** A random countable subset \( \Pi \subset \mathbb{R}^d \) is called a spatial Poisson process with (constant) intensity \( \lambda \) if the random variables \( N(A) = \# \Pi \cap A, A \subset \mathbb{R}^d \) (Borel measurable, always, for the whole course, but we stop saying this all the time now), satisfy

(a) for all \( n \geq 1 \) and disjoint \( A_1, \ldots, A_n \subset \mathbb{R}^d \), the random variables \( N(A_1), \ldots, N(A_n) \) are independent,

(b) \( N(A) \sim \text{Poi}(\lambda |A|) \), where \( |A| \) denotes the volume (Lebesgue measure) of \( A \).

Here, we use the convention that \( X \sim \text{Poi}(0) \) means \( \mathbb{P}(X = 0) = 1 \) and \( X \sim \text{Poi}(\infty) \) means \( \mathbb{P}(X = \infty) = 1 \). This is consistent with \( \mathbb{E}(X) = \lambda \) for \( X \sim \text{Poi}(\lambda), \lambda \in (0, \infty) \). This convention captures that \( \Pi \) does not have points in a given set of zero volume a.s., and it has infinitely many points in given sets of infinite volume a.s.

In fact, the definition fully specifies the joint distributions of the random set function \( N \) on subsets of \( \mathbb{R}^d \), since for any non-disjoint \( B_1, \ldots, B_m \subset \mathbb{R}^d \) we can consider all
intersections of the form $A_k = B_1^* \cap \ldots \cap B_m^*$, where each $B_j^*$ is either $B_j^* = B_j$ or $B_j^* = B_j^c = \mathbb{R}^d \setminus B_j$. They form $n = 2^m$ disjoint sets $A_1, \ldots, A_n$ to which (a) of the definition applies. $(N(B_1), \ldots, N(B_m))$ is a just a linear transformation of $(N(A_1), \ldots, N(A_n))$.

Grimmett and Stirzaker collect a long list of applications including modelling stars in a galaxy, galaxies in the universe, weeds in the lawn, the incidence of thunderstorms and tornadoes. Sometimes the process in Definition 16 is not a perfect description of such a system, but useful as a first step. A second step is the following generalisation:

**Definition 16 (Spatial Poisson process, continued)** A random countable subset $\Pi \subset D \subset \mathbb{R}^d$ is called a spatial Poisson process with (locally integrable) intensity function $\lambda : D \to [0, \infty)$, if $N(A) = \#\Pi \cap A, A \subset D$, satisfy

(a) for all $n \geq 1$ and disjoint $A_1, \ldots, A_n \subset D$, the random variables $N(A_1), \ldots, N(A_n)$ are independent,

(b) $N(A) \sim \text{Poi} \left( \int_A \lambda(x) dx \right)$.

**Definition 17 (Poisson counting measure)** A set function $A \mapsto N(A)$ that satisfies (a) and (b) is referred to as a Poisson counting measure with intensity function $\lambda(x)$.

It is sufficient to check (a) and (b) for rectangles $A_j = (a_j^{(i)}, b_j^{(i)}) \times \ldots \times (a_d^{(i)}, b_d^{(i)})$. The set function $\Lambda(A) = \int_A \lambda(x) dx$ is called the intensity measure of $\Pi$. Definitions 16 and 17 can be extended to measures that are not integrals of intensity functions. Only if $\Lambda(\{x\}) \geq 0$, we would require $\mathbb{P}(N(\{x\}) \geq 2) > 0$ and this is incompatible with $N(\{x\}) = \#\Pi \cap \{x\}$ for a random countable set $\Pi$, so we prohibit such “atoms” of $\Lambda$.

**Example 18 (Compound Poisson process)** Let $(C_t)_{t \geq 0}$ be a compound Poisson process with independent jump sizes $Y_j, j \geq 1$ with common probability density $h(x), x > 0$, at the times of a Poisson process $(X_t)_{t \geq 0}$ with rate $\lambda > 0$. Let us show that

$$N((a, b] \times (c, d]) = \#\{t \in (a, b]: \Delta C_t \in (c, d]\}$$

defines a Poisson counting measure. First note $N((a, b] \times (0, \infty)) = X_b - X_a$. Now recall

**Thinning property of Poisson processes:** If each point of a Poisson process $(X_t)_{t \geq 0}$ of rate $\lambda$ is of type 1 with probability $p$ and of type 2 with probability $1 - p$, independently of one another, then the processes $X^{(1)}$ and $X^{(2)}$ counting points of type 1 and 2, respectively, are independent Poisson processes with rates $p\lambda$ and $(1 - p)\lambda$, respectively.

Consider the thinning mechanism, where the $j$th jump is of type 1 if $Y_j \in (c, d]$. Then, the process counting jumps in $(c, d]$ is a Poisson process with rate $\lambda \mathbb{P}(Y_1 \in (c, d])$, and so

$$N((a, b] \times (c, d]) = X_b^{(1)} - X_a^{(1)} \sim \text{Poi}((b - a)\lambda \mathbb{P}(Y_1 \in (c, d])).$$

We identify the intensity measure $\Lambda((a, b] \times (c, d]) = (b - a)\lambda \mathbb{P}(Y_1 \in (c, d])$.

For the independence of counts in disjoint rectangles $A_1, \ldots, A_n$, we cut them into smaller rectangles $B_i = (a_i, b_i] \times (c_i, d_i], 1 \leq i \leq m$ such that for any two $B_i$ and $B_j$ either $(c_i, d_i] = (c_j, d_j]$ or $(c_i, d_i] \cap (c_j, d_j] = \emptyset$. Denote by $k$ the number of different intervals $(c_i, d_i], w.l.o.g. (c_i, d_i]$ for $1 \leq i \leq k$. Now a straightforward generalisation of the thinning property to $k$ types splits $(X_t)_{t \geq 0}$ into $k$ independent Poisson processes $X^{(i)}$ with rates $\lambda \mathbb{P}(Y_1 \in (c_i, d_i]), 1 \leq i \leq k$. Now $N(B_1), \ldots, N(B_m)$ are independent as increments of independent Poisson processes or of the same Poisson process over disjoint time intervals.
3.3 Poisson point processes

In Example 18, the intensity measure is of the product form \( \Lambda((a, b] \times (c, d]) = (b - a)\nu((c, d]) \) for a measure \( \nu \) on \( D_0 = (0, \infty) \). Take \( D = [0, \infty) \times D_0 \) in Definition 16. This means, that the spatial Poisson process is homogeneous in the first component, the time component, like the Poisson process.

**Proposition 19** If \( \Lambda((a, b] \times A_0) = (b - a) \int_{A_0} g(x)dx \) for a locally integrable function \( g \) on \( D_0 \) (or \( = (b - a)\nu(A_0) \) for a locally finite measure \( \nu \) on \( D_0 \)), then no two points of \( \Pi \) share the same first coordinate.

**Proof:** If \( \nu \) is finite, this is clear, since then \( X_t = N([0, t] \times D_0), \ t \geq 0 \), is a Poisson process with rate \( \nu(D_0) \). Let us restrict attention to \( D_0 = \mathbb{R}^* = \mathbb{R} \setminus \{0\} \) for simplicity – this is the most relevant case for us. The local integrability condition means that we can find intervals \( (I_n)_{n \geq 1} \) such that \( \bigcup_{n \geq 1} I_n = D_0 \) and \( \nu(I_n) < \infty, \ n \geq 1 \). Then the independence of \( N((t_j, t_{j+1}] \times I_n), \ j = 1, \ldots, m, \ n \geq 1 \), implies that \( X^{(n)}_t = N([0, t] \times I_n) \), \( t \geq 0 \), are independent Poisson processes with rates \( \nu(I_n) \), \( n \geq 1 \). Therefore any two of the jump times \( (T^{(n)}_j, j \geq 1, n \geq 1) \) are jointly continuously distributed and take different values almost surely:

\[
\mathbb{P}(T^{(n)}_j = T^{(m)}_i) = \int_0^\infty \int_x f^{(n)}_{T^{(n)}_j}(x)f^{(m)}_{T^{(m)}_i}(y)dydx = 0 \quad \text{for all } n \neq m.
\]

[Alternatively, show that \( T^{(n)}_j - T^{(m)}_i \) has a continuous distribution and hence does not take a fixed value 0 almost surely.]

Finally, there are only countably many pairs of jump times, so almost surely no two jump times coincide. \qed

Let \( \Pi \) be a spatial Poisson process with intensity measure \( \Lambda((a, b] \times (c, d]) = (b - a) \int_c^d g(x)dx \) for a locally integrable function \( g \) on \( D_0 \) (or \( = (b - a)\nu((c, d]) \) for a locally finite measure \( \nu \) on \( D_0 \)), then the process \((\Delta_t)_{t \geq 0}\) given by

\[
\Delta_t = 0 \quad \text{if } \Pi \cap \{t\} \times D = \emptyset, \quad \Delta_t = x \quad \text{if } \Pi \cap \{t\} = \{(t, x)\}
\]

is a Poisson point process in \( D_0 \cup \{0\} \) with intensity function \( g \) on \( D_0 \) in the sense of the following definition.

**Definition 20 (Poisson point process)** Let \( g \) be locally integrable on \( D_0 \subset \mathbb{R}^{d-1} \setminus \{0\} \) (or \( \nu \) locally finite). A process \((\Delta_t)_{t \geq 0}\) in \( D_0 \cup \{0\} \) such that

\[
N((a, b] \times A_0) = \# \{t \in (a, b] : \Delta_t \in A_0\}, \quad 0 \leq a < b, A_0 \subset D_0 \text{ (measurable),}
\]

is a Poisson counting measure with intensity \( \Lambda((a, b] \times A_0) = (b - a) \int_{A_0} g(x)dx \) (or \( \Lambda((a, b] \times A_0) = (b - a)\nu(A_0) \)), is called a Poisson point process with intensity \( g \) (or intensity measure \( \nu \)).

Note that for every Poisson point process, the set \( \Pi = \{(t, \Delta_t) : t \geq 0, \Delta_t \neq 0\} \) is a spatial Poisson process. Poisson random measure and Poisson point process are representations of this spatial Poisson process. Poisson point processes as we have defined them always have a time coordinate and are homogeneous in time, but not in their spatial coordinates.

In the next lecture we will see how one can do computations with Poisson point processes, notably relating to \( \sum \Delta_t \).
Lecture 4

Spatial Poisson processes II

Reading: Kingman Sections 2.2, 2.5, 3.1; Further reading: Williams Chapters 9 and 10

In this lecture, we construct spatial Poisson processes and study sums \(\sum_{s \leq t} f(\Delta_s)\) over Poisson point processes \((\Delta_t)_{t \geq 0}\). We will identify \(\sum_{s \leq t} \Delta_s\) as Lévy process next lecture.

4.1 Series and increasing limits of random variables

Recall that for two independent Poisson random variables \(X \sim \text{Poi}(\lambda)\) and \(Y \sim \text{Poi}(\mu)\) we have \(X + Y \sim \text{Poi}(\lambda + \mu)\). Much more is true. A simple induction shows that \(X_j \sim \text{Poi}(\mu_j), 1 \leq j \leq m, \text{ independent} \Rightarrow X_1 + \ldots + X_m \sim \text{Poi}(\mu_1 + \ldots + \mu_m)\).

What about countably infinite families with \(\mu = \sum_{m \geq 1} \mu_m < \infty\)? Here is a general result, a bit stronger than the convergence theorem for moment generating functions.

**Lemma 21** Let \((Z_m)_{m \geq 1}\) be an increasing sequence of \([0, \infty)\)-valued random variables. Then \(Z = \lim_{m \to \infty} Z_m\) exists a.s. as a \([0, \infty]\)-valued random variable. In particular,

\[
\mathbb{E}(e^{\gamma Z_m}) \to \mathbb{E}(e^{\gamma Z}) = M(\gamma) \quad \text{for all } \gamma \neq 0.
\]

We have

\[
P(Z < \infty) = 1 \iff \lim_{\gamma \to 0} M(\gamma) = 1
\]

and

\[
P(Z = \infty) = 1 \iff M(\gamma) = 0 \text{ for all (one) } \gamma < 0.
\]

**Proof:** Limits of increasing sequences exist in \([0, \infty]\). Hence, if a random sequence \((Z_m)_{m \geq 1}\) is increasing a.s., its limit \(Z\) exists in \([0, \infty]\) a.s. Therefore, we also have \(e^{\gamma Z_m} \to e^{\gamma Z} \in [0, \infty]\) with the conventions \(e^{-\infty} = 0\) and \(e^{\infty} = \infty\). Then (by monotone convergence) \(\mathbb{E}(e^{\gamma Z_m}) \to \mathbb{E}(e^{\gamma Z})\).

If \(\gamma < 0\), then \(e^{\gamma Z} = 0 \iff Z = \infty\), but \(\mathbb{E}(e^{\gamma Z})\) is a mean (weighted average) of nonnegative numbers (write out the definition in the discrete case), so \(P(Z = \infty) = 1\) if and only if \(\mathbb{E}(e^{\gamma Z}) = 0\). As \(\gamma \uparrow 0\), we get \(e^{-\gamma Z} \uparrow 1\) if \(Z < \infty\) and \(e^{-\gamma Z} \to 0\) if \(Z = \infty\), so (by monotone convergence)

\[
\mathbb{E}(e^{\gamma Z}) \uparrow \mathbb{E}(1_{\{Z < \infty\}}) = P(Z < \infty)
\]

and the result follows. \(\square\)
Example 22 For independent \( X_j \sim \text{Poi}(\mu_j) \) and \( Z_m = X_1 + \ldots + X_m \), the random variable \( Z = \lim_{m \to \infty} Z_m \) exists in \([0, \infty)\) a.s. Now
\[
\mathbb{E}(e^{\gamma Z_m}) = \mathbb{E}((e^{\gamma})^{Z_m}) = e^{(e^{\gamma})^1(\mu_1 + \ldots + \mu_m)} \to e^{-(1-e^\gamma)\mu}
\]
shows that the limit is \( \text{Poi}(\mu) \) if \( \mu = \sum_{m \to \infty} \mu_m < \infty \). We do not need the lemma for this, since we can even directly identify the limiting moment generating function.

If \( \mu = \infty \), the limit of the moment generating function vanishes, and by the lemma, we obtain \( \mathbb{P}(Z = \infty) = 1 \). So we still get \( S \sim \text{Poi}(\mu) \) within the extended range \( 0 \leq \mu \leq \infty \).

4.2 Construction of spatial Poisson processes

The examples of compound Poisson processes are the key to constructing spatial Poisson processes with finite intensity measure. Infinite intensity measures can be decomposed.

Theorem 23 (Construction) Let \( \Lambda \) be an intensity measure on \( D \subset \mathbb{R}^d \) and suppose that there is a partition \((I_n)_{n \geq 1}\) of \( D \) into regions with \( \Lambda(I_n) < \infty \). Consider independently \( N_n \sim \text{Poi}(\Lambda(I_n)) \), \( Y_1^{(n)}, Y_2^{(n)}, \ldots \sim \frac{\Lambda(I_n \cap \cdot)}{\Lambda(I_n)} \), i.e. \( \mathbb{P}(Y_j^{(n)} \in A) = \frac{\Lambda(I_n \cap A)}{\Lambda(I_n)} \) and define \( \Pi_n = \{Y_j^{(n)} : 1 \leq j \leq N_n\} \). Then \( \Pi = \bigcup_{n \geq 1} \Pi_n \) is a spatial Poisson process with intensity measure \( \Lambda \).

Proof: First fix \( n \) and show that \( \Pi_n \) is a spatial Poisson process on \( I_n \)

**Thinning property of Poisson variables:** Consider a sequence of independent Bernoulli\((p)\) random variables \((B_j)_{j \geq 1}\) and independent \( X \sim \text{Poi}(\lambda) \).

Then the following two random variables are independent:

\[
X_1 = \sum_{j=1}^{X} B_j \sim \text{Poi}(p\lambda) \quad \text{and} \quad X_2 = \sum_{j=1}^{X} (1 - B_j) \sim \text{Poi}((1-p)\lambda).
\]

To prove this, calculate the joint probability generating function
\[
\mathbb{E}(r^{X_1}S^{X_2}) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)\mathbb{E}(r^{B_1 + \ldots + B_n}s^{n-B_1-\ldots-B_n})
= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}e^{-\lambda} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} r^k s^{n-k}
= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} (pr + (1-p)s)^n = e^{-\lambda p(1-r)}e^{-\lambda(1-p)(1-s)},
\]
so the probability generating function factorises giving independence and we recognise the Poisson distributions as claimed.

For \( A \subset I_n \), consider \( X = N_n \) and the thinning mechanism, where \( B_j = 1_{\{Y_j^{(n)} \in A\}} \sim \text{Bernoulli}(\mathbb{P}(Y_j^{(n)} \in A)) \), then we get property (b):

\( N_n(A) = X_1 \) is Poisson distributed with parameter \( \mathbb{P}(Y_j^{(n)} \in A)\Lambda(I_n) = \Lambda(A) \).
For property (a), disjoint sets $A_1, \ldots, A_m \subset I_n$, we apply the analogous thinning property for $m + 1$ types $Y_j^{(n)} \in A_i$, $i = 0, \ldots, m$, where $A_0 = I_n \setminus (A_1 \cup \ldots \cup A_m)$ to deduce the independence of $N_n(A_1), \ldots, N_n(A_m)$. Thus, $\Pi_n$ is a spatial Poisson process.

Now for $N(A) = \sum_{n \geq 1} N_n(A \cap I_n)$, we add up infinitely many Poisson variables and, by Example 22, obtain a $\text{Poi}(\mu)$ variable, where $\mu = \sum_{n \geq 1} \Lambda(A \cap I_n) = \Lambda(A)$, i.e. property (b). Property (a) also holds, since $N_n(A_j \cap I_n), n \geq 1, j = 1, \ldots, m$, are all independent, and $N(A_1), \ldots, N(A_m)$ are independent as functions of independent random variables.

\[ \square \]

### 4.3 Sums over Poisson point processes

Recall that a Poisson point process $(\Delta_t)_{t \geq 0}$ with intensity function $g : D_0 \to [0, \infty)$ first but this can then be generalised – is a process such that

\[ N((a, b] \times (c, d]) = \#\{a < t \leq b : \Delta_t \in (c, d]\} \sim \text{Poi}\left( (b - a) \int_c^d g(x)dx \right), \]

$0 \leq a < b$, $(c, d) \subset D_0$, defines a Poisson counting measure on $D = [0, \infty) \times D_0$. This means that

\[ \Pi = \{(t, \Delta_t) : t \geq 0 \text{ and } \Delta_t \neq 0\} \]

is a spatial Poisson process. Thinking of $\Delta_t$ as a jump size at time $s$, let us study $X_t = \sum_{0 \leq s \leq t} \Delta_s$, the process performing all these jumps. Note that this is the situation for compound Poisson processes $X$; in Example 18, $g : (0, \infty) \to [0, \infty)$ is integrable.

**Theorem 24 (Exponential formula)** Let $(\Delta_t)_{t \geq 0}$ be a Poisson point process with locally integrable intensity function $g : (0, \infty) \to [0, \infty)$. Then for all $\gamma \in \mathbb{R}$

\[ \mathbb{E}\left( \exp\left\{ \gamma \sum_{0 \leq s \leq t} \Delta_s \right\} \right) = \exp\left\{ t \int_0^\infty (e^{\gamma x} - 1)g(x)dx \right\}. \]

**Proof:** Local integrability of $g$ on $(0, \infty)$ means in particular that $g$ is integrable on $I_n = (2^n, 2^{n+1}]$, $n \in \mathbb{Z}$. The properties of the associated Poisson counting measure $N$ immediately imply that the random counting measures $N_n$ counting all points in $I_n$, $n \in \mathbb{Z}$, defined by

\[ N_n((a, b] \times (c, d]) = \{a < t \leq b : \Delta_t \in (c, d]\cap I_n\}, \]

$0 \leq a < b$, $(c, d) \subset I_n$, are independent. Furthermore, $N_n$ is the Poisson counting measure of jumps of a compound Poisson process with $(b - a)\int_c^d g(x)dx = (b - a)\lambda_n\mathbb{P}(Y_1^{(n)} \in (c, d])$ for $0 \leq a < b$ and $(c, d) \subset I_n$ (cf. Example 18), so $\lambda_n = \int_{I_n} g(x)dx$ and (if $\lambda_n > 0$) jump density $h_n = \lambda_n^{-1}g$ on $I_n$, zero elsewhere. Therefore, we obtain

\[ \mathbb{E}\left( \exp\left\{ \gamma \sum_{0 \leq s \leq t} \Delta_s^{(n)} \right\} \right) = \exp\left\{ t \int_{I_n} (e^{\gamma x} - 1)g(x)dx \right\}, \]

where $\Delta_s^{(n)} = \begin{cases} \Delta_s & \text{if } \Delta_s \in I_n \\ 0 & \text{otherwise} \end{cases}$
Now we have
\[ Z_m = \sum_{n=-m}^{m} \sum_{0 \leq s \leq t} \Delta_s^{(n)} \uparrow \sum_{0 \leq s \leq t} \Delta_s \]  
\[ \text{as } m \to \infty, \]

and (cf. Lemma 21 about finite or infinite limits), the associated moment generating functions (products of individual moment generating functions) converge as required:
\[ \prod_{n=-m}^{m} \exp \left\{ t \int_{2^n}^{2^{n+1}} (e^{\gamma x} - 1)g(x)dx \right\} \to \exp \left\{ t \int_{0}^{\infty} (e^{\gamma x} - 1)g(x)dx \right\}. \]

\[ \square \]

4.4 Martingales (from B10a)

A discrete-time stochastic process \((M_n)_{n \geq 0}\) in \(\mathbb{R}\) is called a **martingale** if for all \(n \geq 0\)
\[ \mathbb{E}(M_{n+1}|M_0, \ldots, M_n) = M_n, \quad \text{i.e. if } \mathbb{E}(M_{n+1}|M_0 = x_0, \ldots, M_n = x_n) = x_n \text{ for all } x_j. \]

This is the principle of a fair game. What can I expect from the future if my current state is \(M_n = x_n\)? No gain and no loss, on average, whatever the past. The following important rules for conditional expectations are crucial to establish the martingale property

- If \(X\) and \(Y\) are independent, then \(\mathbb{E}(X|Y) = \mathbb{E}(X)\).
- If \(X = f(Y)\), then \(\mathbb{E}(X|Y) = \mathbb{E}(f(Y)|Y) = f(Y)\) for functions \(f : \mathbb{R} \to \mathbb{R}\) for which the conditional expectations exist.
- Conditional expectation is linear \(\mathbb{E}(\alpha X_1 + X_2|Y) = \alpha \mathbb{E}(X_1|Y) + \mathbb{E}(X_2|Y)\).
- More generally: \(\mathbb{E}(g(Y)X|Y) = g(Y)\mathbb{E}(X|Y)\) for functions \(g : \mathbb{R} \to \mathbb{R}\) for which the conditional expectations exist.

These are all not hard to prove for discrete random variables. The full statements (continuous analogues) are harder. Martingales in continuous time can also be defined, but (formally) the conditioning needs to be placed on a more abstract footing. Denote by \(\mathcal{F}_s\) the “information available up to time \(s \geq 0\)”, for us just the process \((M_r)_{r \leq s}\) up to time \(s\) – this is often written \(\mathcal{F}_s = \sigma(M_r, r \leq s)\). Then the four bullet point rules still hold for \(Y = (M_r)_{r \leq s}\) or for \(Y\) replaced by \(\mathcal{F}_s\).

We call \((M_t)_{t \geq 0}\) a martingale if for all \(s \leq t\)
\[ \mathbb{E}(M_t|\mathcal{F}_s) = M_s. \]

**Example 25** Let \((N_s)_{s \geq 0}\) be a Poisson process with rate \(\lambda\). Then \(M_s = N_s - \lambda s\) is a martingale: by the first three bullet points and by the Markov property (Proposition 3)
\[ \mathbb{E}(N_t - \lambda t|\mathcal{F}_s) = \mathbb{E}(N_s + (N_t - N_s) - \lambda t|\mathcal{F}_s) = N_s + (t-s)\lambda - \lambda t = N_s - \lambda s. \]

Also \(E_s = \exp\{\gamma N_s - \lambda s(e^\gamma - 1)\}\) is a martingale since by the first and last bullet points above, and by the Markov property
\[ \mathbb{E}(E_t|\mathcal{F}_s) = \mathbb{E}(\exp\{\gamma N_s + \gamma (N_t - N_s) - \lambda t(e^\gamma - 1)\}|\mathcal{F}_s) \]
\[ = \exp\{\gamma N_s - \lambda t(e^\gamma - 1)\}\mathbb{E}(\exp\{\gamma (N_t - N_s)\}) \]
\[ = \exp\{\gamma N_s - \lambda t(e^\gamma - 1)\}\exp\{-\lambda(t-s)(e^\gamma - 1)\} = E_s. \]

We will review relevant martingale theory when this becomes relevant.
Lecture 5

The characteristics of subordinators

Reading: Kingman Section 8.4

We have done the leg-work. We can now harvest the fruit of our efforts and proceed to a number of important consequences. Our programme for the next couple of lectures is:

- We construct Lévy processes from their jumps, first the most general increasing Lévy process. As linear combinations of independent Lévy processes are Lévy processes (Assignment A.1.2.(a)), we can then construct Lévy processes such as Variance Gamma processes of the form $Z_t = X_t - Y_t$ for two increasing $X$ and $Y$.

- We have seen martingales associated with $N_t$ and $\exp\{N_t\}$ for a Poisson process $N$. Similar martingales exist for all Lévy processes (cf. Assignment A.2.3.). Martingales are important for finance applications, since they are the basis of arbitrage-free models (more precisely, we need equivalent martingale measures, but we will assume here a “risk-free” measure directly to avoid technicalities).

- Our rather restrictive first range of examples of Lévy processes was obtained from known infinitely divisible distributions. We can now model using the intensity function of the Poisson point process of jumps to get a wider range of examples.

- We can simulate these Lévy processes, either by approximating random walks based on the increment distribution, or by constructing the associated Poisson point process of jumps, as we have seen, from a collection of independent random variables.

5.1 Subordinators and the Lévy-Khintchine formula

We will call (weakly) increasing Lévy processes “subordinators”. Recall “$\nu(dx) \hat{=} q(x)dx$”.

Theorem 26 (Construction) Let $a \geq 0$, and let $(\Delta_t)_{t \geq 0}$ be a Poisson point process with intensity measure $\nu$ on $(0, \infty)$ such that

$$\int_{(0,\infty)} (1 \wedge x)\nu(dx) < \infty,$$

then the process $X_t = at + \sum_{s \leq t} \Delta_s$ is a subordinator with moment generating function $\mathbb{E}(\exp\{\gamma X_t\}) = \exp\{t\Psi(\gamma)\}$, where

$$\Psi(\gamma) = a\gamma + \int_{(0,\infty)} (e^{\gamma x} - 1)\nu(dx).$$
Proof: Clearly \((at)_{t \geq 0}\) is a deterministic subordinator and we may assume \(a = 0\) in the sequel. Now the Exponential formula gives the moment generating function of \(X_t = \sum_{s \leq t} \Delta_s\). We can now use Lemma 21 to check whether \(X_t < \infty\) for \(t > 0\):

\[
P(X_t < \infty) = 1 \iff \mathbb{E}(\exp \{\gamma X_t\}) = \exp \left\{ t \int_0^\infty (e^{\gamma x} - 1) \nu(dx) \right\} \to 1 \quad \text{as} \quad \gamma \uparrow 0.
\]

This happens, by monotone convergence, if and only if for some (equivalently all) \(\gamma < 0\)

\[
\int_0^\infty (1 - e^{\gamma x}) \nu(dx) < \infty \iff \int_0^\infty (1 - x) \nu(dx) < \infty.
\]

It remains to check that \((X_t)_{t \geq 0}\) is a Lévy process. Fix \(0 \leq t_0 < t_1 < \ldots < t_n\). Since \((\Delta_s)_{s \geq 0}\) is a Poisson point process, the processes \((\Delta_s)_{t_{j-1} \leq s < t_j}, j = 1, \ldots, n\), are independent (consider the restrictions to disjoint domains \([t_{j-1}, t_j) \times (0, \infty)\) of the Poisson counting measure

\[
N((a, b] \times (c, d]) = \{a \leq t < b : \Delta_t \in (c, d]\}, \quad 0 \leq a < b, 0 < c < d,
\]

and so are the sums \(\sum_{t_{j-1} \leq s < t_j} \Delta_s\) as functions of independent random variables. Fix \(s < t\). Then the process \((\Delta_{s+t})_{t \geq 0}\) has the same distribution as \((\Delta_s)_{s \geq 0}\). In particular, \(\sum_{0 \leq r \leq t} \Delta_s \sim \sum_{0 \leq r \leq t} \Delta_r\). The process \(t \to \sum_{s \leq t} \Delta_s\) is right-continuous with left limits, since it is a random increasing function where for each jump time \(T\), we have

\[
\lim_{t \uparrow T} \sum_{s \leq t} \Delta_s = \lim_{t \uparrow T} \sum_{s < T} \Delta_s 1_{s \leq t} = \sum_{s < T} \Delta_s \quad \text{and} \quad \lim_{t \uparrow T} \sum_{s \leq t} \Delta_s = \lim_{t \uparrow T} \sum_{s \leq T + 1} \Delta_s 1_{s \leq t} = \sum_{s \leq T} \Delta_s,
\]

by monotone convergence, because each of the terms \(\Delta_s 1_{s \leq t}\) in the sums converges. \(\square\)

Note also that, due to the Exponential formula, \(P(X_t < \infty) > 0\) already implies \(P(X_t < \infty) = 1\). We shall now state but not prove the Lévy-Khintchine formula for nonnegative random variables.

**Theorem 27 (Lévy-Khintchine)** A nonnegative random variable \(Y\) has an infinitely divisible distribution if and only if there is a pair \((a, \nu)\) such that for all \(\gamma \leq 0\)

\[
\mathbb{E}(\exp\{\gamma Y\}) = \exp \left\{ a \gamma + \int_{(0,\infty)} (e^{\gamma x} - 1) \nu(dx) \right\},
\]

where \(a \geq 0\) and \(\nu\) is such that \(\int_{(0,\infty)} (1 - x) \nu(dx) < \infty\).

**Corollary 28** Given a nonnegative random variable \(Y\) with infinitely divisible distribution, there exists a subordinator \((X_t)_{t \geq 0}\) with \(X_1 \sim Y\).

Proof: Let \(Y\) have an infinitely divisible distribution. By the Lévy-Khintchine theorem, its moment generating function is of the form (1) for parameters \((a, \nu)\). Theorem 26 constructs a subordinator \((X_t)_{t \geq 0}\) with \(X_1 \sim Y\). \(\square\)

This means that the class of subordinators can be parameterised by two parameters, the nonnegative “drift parameter” \(a \geq 0\), and the “Lévy measure” \(\nu\), or its density, the “Lévy density” \(g : (0, \infty) \to [0, \infty)\). The parameters \((a, \nu)\) are referred to as the “Lévy-Khintchine characteristics” of the subordinator (or of the infinitely divisible distribution). Using the Uniqueness theorem for moment generating functions, it can be shown that \(a\) and \(\nu\) are unique, i.e. that no two sets of characteristics refer to the same distribution.
5.2 Examples

Example 29 (Gamma process) The Gamma process, where $X_t \sim \text{Gamma}(\alpha t, \beta)$, is an increasing Lévy process. In Assignment A.2.4. we showed that

$$E(\exp \{ \gamma X_t \}) = \left( \frac{\beta}{\beta - \gamma} \right)^{\alpha t} \exp \left\{ t \int_0^\infty (e^{\gamma x} - 1) \alpha x^{-1} e^{-\beta x} dx \right\}, \quad \gamma < \beta.$$  

We read off the characteristics $a = 0$ and $g(x) = \alpha x^{-1} e^{-\beta x}$, $x > 0$.

Example 30 (Poisson process) The Poisson process, where $X_t \sim \text{Poi}(\lambda t)$, has

$$E(\exp \{ \gamma X_t \}) = \exp \{ t \lambda (e^\gamma - 1) \}.$$  

This corresponds to characteristics $a = 0$ and $\nu = \lambda \delta_1$, where $\delta_1$ is the discrete unit point mass in (jump size) 1.

Example 31 (Increasing compound Poisson process) The compound Poisson process $C_t = Y_1 + \ldots + Y_{X_t}$, for a Poisson process $X$ and independent identically distributed nonnegative $Y_1, Y_2, \ldots$ with probability density function $h(x), x > 0$, satisfies

$$E(\exp \{ \gamma C_t \}) = \exp \left\{ t \int_0^\infty (e^{\gamma x} - 1) \lambda h(x) dx \right\}.$$  

and we read off characteristics $a = 0$ and $g(x) = \lambda h(x), x > 0$. We can add a drift and consider $\tilde{C}_t = \tilde{\alpha} t + C_t$ for some $\tilde{\alpha} > 0$ to get a compound Poisson process with drift.

Example 32 (Stable subordinator) The stable subordinator is best defined in terms of its Lévy-Khintchine characteristics $a = 0$ and $g(x) = x^{-\alpha - 1}$. This gives for $\gamma \leq 0$

$$E(\exp \{ \gamma X_t \}) = \exp \left\{ t \int_0^\infty (e^{\gamma x} - 1) x^{-\alpha - 1} dx \right\} = \exp \left\{ t \frac{\Gamma(1 - \alpha)}{\alpha} (-\gamma)^\alpha \right\}.$$  

Note that $E(\exp \{ \gamma e^{1/\alpha} X_{t/c} \}) = E(\exp \{ \gamma X_t \})$, so that $(e^{1/\alpha} X_{t/c})_{t \geq 0} \sim X$. More generally, we can also consider e.g. tempered stable processes with $g(x) = x^{-\alpha - 1} \exp \{ -\rho x \}, \rho > 0$.

![Figure 5.1: Examples: Poisson process, Gamma process, stable subordinator](image)

5.3 Aside: nonnegative Lévy processes

It may seem obvious that a nonnegative Lévy process, i.e. one where $X_t \geq 0$ a.s. for all $t \geq 0$, is automatically increasing, since every increment $X_{s+t} - X_s$ has the same distribution $X_t$ and is hence also nonnegative. Let us be careful, however, and remember that there is a difference between something never happening at a fixed time and something never happening at any time. We have e.g. for a (one-dimensional) Poisson process $(N_t)_{t \geq 0}$

$$P(\Delta N_t \neq 0) = \sum_{n \geq 1} P(T_n = t) = 0 \quad \text{for all } t \geq 0, \text{ but } \quad P(\exists t : \Delta N_t \neq 0) = 1.$$
Here we can argue that if \( f(t) < f(s) \) for some \( s < t \) and a right-continuous function, then there are also two rational numbers \( s_0 < t_0 \) for which \( f(t_0) < f(s_0) \), so
\[
\mathbb{P}(\exists s, t \in (0, \infty), s < t : X_t - X_s < 0) > 0 \Rightarrow \mathbb{P}(\exists s_0, t_0 \in (0, \infty) \cap \mathbb{Q} : X_{t_0} - X_{s_0} < 0) > 0
\]
However, the latter can be bounded above (by subadditivity \( \mathbb{P}(\bigcup_n A_n) \leq \sum_n \mathbb{P}(A_n) \))
\[
\mathbb{P}(\exists s_0, t_0 \in (0, \infty) \cap \mathbb{Q} : X_{t_0} - X_{s_0} < 0) \leq \sum_{s_0, t_0 \in (0, \infty) \cap \mathbb{Q}} \mathbb{P}(X_{t_0 - s_0} < 0) = 0.
\]

Another instance of such delicate argument is the following: if \( X_t \geq 0 \) a.s. for one \( t > 0 \) and a subordinator \( X \), then \( X_t \geq 0 \) a.s. for all \( t \geq 0 \). It is true, but to say if \( \mathbb{P}(X_s \leq 0) > 0 \) for some \( s < t \) then \( \mathbb{P}(X_t \leq 0) \geq 0 \) may not be all that obvious. It is, however, easily justified for \( s = t/m \), since then \( \mathbb{P}(X_t \leq 0) \geq \mathbb{P}(X_{tj/m} - X_{t(j-1)/m} < 0 \text{ for all } j = 1, \ldots, m) > 0 \). We have to apply a similar argument to get \( \mathbb{P}(X_{tq} \leq 0) = 0 \) for all rational \( q > 0 \). Then we use again right-continuity to see that a function that is nonnegative at all rationals cannot take a negative value at an irrational either, so we get
\[
\mathbb{P}(\exists s \in [0, \infty) : X_s < 0) = \mathbb{P}(\exists s \in [0, \infty) \cap \mathbb{Q} : X_s < 0) \leq \sum_{s \in [0, \infty) \cap \mathbb{Q}} \mathbb{P}(X_s < 0) = 0.
\]

### 5.4 Applications

Subordinators have found a huge range of applications, but are not directly models for a lot of real world phenomena. We can now construct more general Lévy processes of the form \( Z_t = X_t - Y_t \) for two subordinators \( X \) and \( Y \). Let us here indicate some subordinators as they are used/arise in connection with other Lévy processes.

**Example 33 (Subordination)** For a Lévy process \( X \) and an independent subordinator \( T \), the process \( Y_s = X_{T_s}, s \geq 0 \), is also a Lévy process (we study this later in the course). The rough argument is that \( (X_{T_{s+u}} - X_{T_s})_{u \geq 0} \) is independent of \( (X_r)_{r \leq T_s} \), and distributed as \( X \), by the Markov property. Hence \( X_{T_{s+r}} - X_{T_s} \) is independent of \( X_{T_s} \) and distributed as \( X_{T_s} \). A rigorous argument can be based on calculations of joint moment generating functions. Hence, subordinators are a useful tool to construct Lévy processes, e.g. from Brownian motion \( X \). Many models of financial markets are of this type. The operation \( Y_s = X_{T_s} \) is called subordination – this is where subordinators got their name from.

**Example 34 (Level passage)** Let \( Z_t = a t - X_t \) where \( a = \mathbb{E}(X_1) \). It can be shown that \( \tau_s = \inf \{ t \geq 0 : Z_t > s \} < \infty \) a.s. for all \( s \geq 0 \) (from the analogous random walk result). It turns out (cf. later in the course) that \( \tau_s \) is a subordinator.

**Example 35 (Level set)** Look at the zero set \( \mathcal{Z} = \{ t \geq 0 : B_t = 0 \} \) for Brownian motion (or indeed any other centred Lévy process) \( B \). \( \mathcal{Z} \) is unbounded since \( B \) crosses zero at arbitrarily large times so as to pass beyond all \( s \) and \(-s\). Recall that \( (tB_{1/t})_{t \geq 0} \) is also a Brownian motion. Therefore, \( \mathcal{Z} \) also has an accumulation point at \( t = 0 \), i.e. crosses zero infinitely often at arbitrarily small times. In fact, it can be shown that \( \mathcal{Z} \) is the closed range \( \{ X_r, r \geq 0 \} \) of a subordinator \((X_r)_{r \geq 0} \). The Brownian scaling property \( (\sqrt{c}B_{1/c})_{t \geq 0} \sim B \) shows that \( \{ X_r/c, r \geq 0 \} \sim \mathcal{Z} \), and so \( X \) must have a scaling property. In fact, \( X \) is a stable subordinator of index \( 1/2 \). Similar results, with different subordinators, hold not just for all Lévy processes but even for most Markov processes.
Lecture 6

Lévy processes with no negative jumps

Reading: Kyprianou 2.1, 2.4, 2.6, Schoutens 2.2; Further reading: Williams 10-11

Subordinators $X$ are processes with no negative jumps. We get processes that can decrease by adding a negative drift $at$ for $a < 0$. Also, Brownian motion $B$ has no negative jumps. A guess might be that $X_t + at + \sigma B_t$ is the most general Lévy process with no negative jumps, but this is false. It turns out that even a non-summable amount of positive jumps can be incorporated, but we will have to look at this carefully.

6.1 Bounded and unbounded variation

The (total) variation of a right-continuous function $f : [0, t] \to \mathbb{R}$ with left limits is

$$||f||_{TV} := \sup \left\{ \sum_{j=1}^{n} |f(t_j) - f(t_{j-1})| : 0 = t_0 < t_1 < \ldots < t_n = t, n \in \mathbb{N} \right\}.$$

Clearly, for an increasing function with $f(0) = 0$ this is just $f(t)$ and for a difference $f = g - h$ of two increasing functions with $g(0) = h(0) = 0$ this is at most $g(t) + h(t) < \infty$, so all differences of increasing functions are of bounded variation. There are, however, functions of infinite variation, e.g. Brownian paths: they have finite quadratic variation

$$\sum_{j=1}^{2^n} |B_{ij2^{-n}} - B_{i(j-1)2^{-n}}|^2 \to t \quad \text{in the } L^2 \text{ sense}$$

since

$$\mathbb{E} \left( \sum_{j=1}^{2^n} |B_{ij2^{-n}} - B_{i(j-1)2^{-n}}|^2 \right) = 2^n \mathbb{E}(B_{12^{-n}}^2) = t$$

and

$$\mathbb{E} \left( \left( \sum_{j=1}^{2^n} |B_{ij2^{-n}} - B_{i(j-1)2^{-n}}|^2 - t \right)^2 \right) = \text{Var} \left( \sum_{j=1}^{2^n} |B_{ij2^{-n}} - B_{i(j-1)2^{-n}}|^2 \right) \leq 2^n (2^{-n}t)^2 \text{Var}(B_1^2) \to 0,$$
but then assuming finite total variation with positive probability, the uniform continuity of the Brownian path implies
\[
\sum_{j=1}^{2^n} |B_{t2^{-n}} - B_{t_{(j-1)2^{-n}}}|^2 \leq \left( \sup_{j=1,...,2^n} |B_{t2^{-n}} - B_{t_{(j-1)2^{-n}}}| \right) \sum_{j=1}^{2^n} |B_{t2^{-n}} - B_{t_{(j-1)2^{-n}}}| \to 0
\]
with positive probability, but this is incompatible with convergence to \( t \), so the assumption of finite total variation must have been wrong.

Here is how jumps influence total variation:

**Proposition 36** Let \( f \) be a right-continuous function with left limits and jumps \( \Delta f_s \) for \( 0 \leq s \leq t \). Then
\[
||f||_{TV} \geq \sum_{0 \leq s \leq t} |\Delta f_s|
\]

**Proof:** Enumerate the jumps in decreasing order of size by \( (T_n, \Delta f_{T_n})_{n \geq 0} \). Fix \( N \in \mathbb{N} \) and \( \delta > 0 \). Choose \( \varepsilon > 0 \) so small that \( \bigcup [T_n - \varepsilon, T_n] \) is a disjoint union and such that \( |f(T_n - \varepsilon) - f(T_n - \varepsilon)| < \delta/N \). Then for \( \{T_n - \varepsilon, T_n : n = 1, \ldots, N\} = \{t_1, \ldots, t_{2N+1}\} \) such that \( 0 = t_0 < t_1 < \cdots < t_{2N+1} < t_{2N+2} = t \), we have
\[
\sum_{j=1}^{2N+2} |f(t_j) - f(t_{j-1})| \geq \sum_{n=1}^{N} \Delta f(T_n) - \delta.
\]
Since \( N \) and \( \delta \) were arbitrary, this completes the proof, whether the right-hand side is finite or infinite. \( \square \)

### 6.2 Martingales (from B10a)

Three martingale theorems are of central importance. We will require in this lecture just the maximal inequality, but we formulate all three here for easier reference. They all come in several different forms. We present the \( L^2 \)-versions as they are most easily formulated and will suffice for us.

A stopping time is a random time \( T \) such that for every \( s \geq 0 \) the information \( \mathcal{F}_s \) allows to decide whether \( T \leq s \). More formally, if the event \( \{T \leq s\} \) can be expressed in terms of \( (M_r, r \leq s) \) (is measurable with respect to \( \mathcal{F}_s \)). The prime example of a stopping time is the first entrance time \( T_A = \inf\{t \geq 0 : M_t \in A\} \). Note that
\[
\{T \leq s\} = \{M_r \notin A \text{ for all } r \leq s\}
\]
(and at least for closed sets \( A \) we can drop the irrational \( r \leq s \) and see measurability, then approximate open sets.)

**Theorem 37 (Optional stopping)** Let \( (M_t)_{t \geq 0} \) be a martingale and \( T \) a stopping time. If \( \sup_{t \geq 0} \mathbb{E}(M_t^2) < \infty \), then \( \mathbb{E}(M_T) = \mathbb{E}(M_0) \).

**Theorem 38 (Convergence)** Let \( (M_t)_{t \geq 0} \) be a martingale such that \( \sup_{t \geq 0} \mathbb{E}(M_t^2) < \infty \), then \( M_t \to M_\infty \) almost surely.

**Theorem 39 (Maximal inequality)** Let \( (M_t)_{t \geq 0} \) be a martingale. Then \( \mathbb{E}(\sup\{M_s^2 : 0 \leq s \leq t\}) \leq 4\mathbb{E}(M_t^2) \).
6.3 Compensation

Let \( g : (0, \infty) \to [0, \infty) \) be the intensity function of a Poisson point process \((\Delta_t)_{t \geq 0}\). If \( g \) is not integrable at infinity, then \( \#\{0 \leq s \leq t : \Delta_s > 1\} \sim \text{Poi}(\int_0^\infty g(x)dx) = \text{Poi}(\infty) \), and it is impossible for a right-continuous function with left limits to have accumulation points in the set of such jumps (lower and upper points of a sequence of jumps will then have different limit points). If however \( g \) is not integrable at zero, we have to investigate this further.

**Proposition 40** Let \((\Delta_t)_{t \geq 0}\) be a Poisson point process with intensity measure \( \nu \) on \((0, \infty)\).

(i) If \( \int_0^\infty x\nu(dx) < \infty \), then

\[
\mathbb{E}\left(\sum_{s \leq t} \Delta_s\right) = t \int_0^\infty x\nu(dx).
\]

(ii) If \( \int_0^\infty x^2\nu(dx) < \infty \), then

\[
\text{Var}\left(\sum_{s \leq t} \Delta_s\right) = t \int_0^\infty x^2\nu(dx).
\]

**Proof:** These are the two leading terms in the expansion with respect to \( \gamma \) of the Exponential formula: the first moment can always be obtained from the moment generating function by taking \( \frac{\partial}{\partial \gamma}\big|_{\gamma=0} \), here

\[
\frac{\partial}{\partial \gamma} \exp\left\{ t \int_0^\infty (e^{\gamma x} - 1)\nu(dx)\right\} \bigg|_{\gamma=0} = t \int_0^\infty xe^{\gamma x}\nu(dx) \bigg|_{\gamma=0} = t \int_0^\infty \nu(dx),
\]

and the second moment follows from the second derivative in the same way. \( \square \)

Consider compound Poisson processes, with a drift that turns them into martingales

\[
Z^\varepsilon_t = \sum_{s \leq t} \Delta_s 1_{\{\varepsilon < \Delta_s \leq 1\}} - t \int_\varepsilon^1 x\nu(dx)
\]  
(1)

We have deliberately excluded jumps in \((1, \infty)\). These are easier to handle separately.

**Lemma 41** Let \((\Delta_t)_{t \geq 0}\) be a Poisson point process with intensity measure \( \nu \) on \((0, 1)\). With \( Z^\varepsilon \) defined in (1), \( Z^\varepsilon_t \) converges in \( L^2 \) if \( \int_0^1 x^2\nu(dx) < \infty \).

**Proof:** We only do this for \( \nu(dx) = g(x)dx \). Note that for \( 0 < \delta < \varepsilon < 1 \), by Proposition 40(ii) applied to \( g_{\delta, \varepsilon}(x) = g(x)1_{\{\delta \leq x < \varepsilon\}} \),

\[
\mathbb{E}(|Z^\varepsilon_t - Z^\varepsilon_t^\delta|^2) = t \int_\delta^\varepsilon x^2g(x)dx
\]

so that \((Z^\varepsilon_t)_{0 < \varepsilon < 1}\) is a Cauchy family as \( \varepsilon \downarrow 0 \), for the \( L^2 \)-distance \( d(X, Y) = \sqrt{\mathbb{E}((X - Y)^2)} \). By completeness of \( L^2 \)-space, there is a limiting random variable \( Z_t \) as required. \( \square \)

We can slightly tune this argument to establish a general existence theorem:

**Theorem 42 (Existence)** There exists a Lévy process whose jumps form a Poisson point process with intensity measure \( \nu \) on \((0, \infty)\) if and only if \( \int_{(0,\infty)} (1 \wedge x^2)\nu(dx) < \infty \).
Proof: The “only if” statement is a consequence of a Lévy-Khintchine type characterisation of infinitely divisible distributions on $\mathbb{R}$, cf. Theorem 44, which we will not prove. Let us prove the “if” part in the case where $\nu(dx) = g(x)dx$.

By Proposition 40(i), $\mathbb{E}(Z_{\varepsilon t}^e - Z_{\delta t}^\delta) = 0$. By Assignment A.2.3.(c), the process $Z_{\varepsilon t}^e - Z_{\delta t}^\delta$ is a martingale, and the maximal inequality (Theorem 39) shows that

$$
\mathbb{E}\left(\sup_{0 \leq s \leq t} |Z_{\varepsilon s}^e - Z_{\delta s}^\delta|\right) \leq 4\mathbb{E}(|Z_{\varepsilon t}^e - Z_{\delta t}^\delta|^2) = 4t \int_\delta^\varepsilon x^2 g(x)dx
$$

so that $(Z_{\varepsilon s}, 0 \leq s \leq t)_{0 < \varepsilon < 1}$ is a Cauchy family as $\varepsilon \downarrow 0$, for the uniform $L^2$-distance $d_{[0,t]}(X, Y) = \sqrt{\mathbb{E}(\sup_{0 \leq s \leq t} |X_s - Y_s|^2)}$. By completeness of $L^2$-space, there is a limiting process $(Z_{\varepsilon s})_{0 \leq s \leq t}$, which as the uniform limit (in $L^2$) of $(Z_{\varepsilon s})_{0 \leq s \leq t}$ is right-continuous with left limits. Also consider the independent compound Poisson process

$$
Z_{\varepsilon t}^{(2)} = \sum_{s \leq t} \Delta_s 1_{\{\Delta_s > 1\}}
$$

and set $Z = Z^{(1)} + Z^{(2)}$.

It is not hard to show that $Z$ is a Lévy process that incorporates all jumps $(\Delta_s)_{0 \leq s \leq t}$. □

**Example 43** Let us look at a Lévy density $g(x) = |x|^{-5/2}$, $x \in [-3, 0)$. Then the compensating drifts $\int_{-3}^x x g(x)dx$ take values 0.845, 2.496, 5.170 and 18.845 for $\varepsilon = 1$, $\varepsilon = 0.3$, $\varepsilon = 0.1$ and $\varepsilon = 0.01$. In the simulation, you see that the slope increases (to infinity, actually as $\varepsilon \downarrow 0$), but the picture begins to stabilise and converge to a limit.

![Figure 6.1: Approximation of a Lévy process with no positive jumps – compensating drift](image-url)
Lecture 7
General Lévy processes and simulation

Reading: Schoutens Sections 8.1, 8.2, 8.4

For processes with no negative jumps, we compensated jumps by a linear drift and incorporated more and more smaller jumps while letting the slope of the linear drift tend to negative infinity. We will now construct the most general real-valued Lévy process as the difference of two such processes (and a Brownian motion). For explicit marginal distributions, we can simulate Lévy processes by approximating random walks. In practice, we often only have explicit characteristics (drift coefficient, Brownian coefficient and Lévy measure). We will also simulate Lévy processes based on the characteristics.

7.1 Construction of Lévy processes

The analogue of Theorem 27 for real-valued random variables is as follows.

**Theorem 44 (Lévy-Khintchine)** A real-valued random variable $X$ has an infinitely divisible distribution if there are parameters $a \in \mathbb{R}$, $\sigma^2 \geq 0$ and a measure $\nu$ on $\mathbb{R} \setminus \{0\}$ with $\int_{-\infty}^{\infty} (1 \wedge x^2) \nu(dx) < \infty$ such that $E(e^{i\lambda X}) = e^{-\psi(\lambda)}$, where

$$\psi(\lambda) = -ia\lambda + \frac{1}{2} \sigma^2 \lambda^2 - \int_{-\infty}^{\infty} (e^{i\lambda x} - 1 - i\lambda x 1_{|x|\leq 1}) \nu(dx), \quad \lambda \in \mathbb{R}.$$

Lévy processes are parameterised by their Lévy-Khintchine characteristics $(a, \sigma^2, \nu)$, where we call $a$ the drift coefficient, $\sigma^2$ the Brownian coefficient and $\nu$ the Lévy measure or jump measure. $\nu(dx)$ will often be of the form $g(x)dx$, and we then refer to $g$ as the Lévy density or jump density.

**Theorem 45 (Existence)** Let $(a, \sigma^2, \nu)$ be Lévy-Khintchine characteristics, $(B_t)_{t \geq 0}$ a standard Brownian motion and $(\Delta_t)_{t \geq 0}$ an independent Poisson point process of jumps with intensity measure $\nu$. Then there is a Lévy process

$$Z_t = at + \sigma B_t + M_t + C_t, \quad \text{where } C_t = \sum_{s \leq t} \Delta_s 1_{\{|\Delta_s| > 1\}}.$$
is a compound Poisson process (of big jumps) and

\[ M_t = \lim_{\varepsilon \downarrow 0} \left( \sum_{s \leq t} \Delta_s 1_{\{\varepsilon < |\Delta_s| \leq 1\}} - t \int_{\{x \in \mathbb{R}: \varepsilon < |x| \leq 1\}} x \nu(dx) \right) \]

is a martingale (of small jumps – compensated by a linear drift).

**Proof:** The construction of \( M_t = P_t - N_t \) can be made from two independent processes \( P_t \) and \( N_t \) with no negative jumps as in Theorem 42. \( N_t \) will be built from a Poisson point process with intensity measure \( \mathcal{P}((c, d]) = \nu([-d, -c]), 0 < c < d \leq 1 \) (or \( \mathcal{P}(y) = g(-y), 0 < y < 1 \)).

We check that the characteristic function of \( Z_t = at + \sigma B_t + P_t - N_t + C_t \) is of Lévy-Khintchine type with parameters \((a, \sigma, \nu)\). We have five independent components. Evaluate at \( t = 1 \) to get

\[
\begin{align*}
\mathbb{E}(e^{\gamma a}) &= e^{\gamma a} \\
\mathbb{E}(e^{\gamma B_1}) &= \exp\left\{\frac{1}{2} \gamma^2 \sigma^2\right\} \\
\mathbb{E}(e^{\gamma P_1}) &= \exp\left\{ \int_0^1 (e^{\gamma x} - 1 - \gamma x) \nu(dx) \right\} \\
\mathbb{E}(e^{-\gamma N_1}) &= \exp\left\{ \int_0^1 (e^{-\gamma y} - 1 + \gamma y) \mathcal{P}(dy) \right\} = \exp\left\{ \int_{-1}^0 (e^{\gamma x} - 1 - \gamma x) \nu(dx) \right\} \\
\mathbb{E}(e^{i\lambda C_1}) &= \exp\left\{ \int_{|x|>1} (e^{i\lambda x} - 1) \nu(dx) \right\}.
\end{align*}
\]

The last formula is checked in analogy with the moment generating function computation of Assignment A.1.3 (in general, the moment generating function will not be well-defined for this component). For the others, now “replace” \( \gamma \) by \( i\lambda \). A formal justification can be obtained by analytic continuation, since the moment generating functions of these components are entire functions of \( \gamma \) as a complex parameter. Now the characteristic function of \( Z_1 \) is the product of characteristic functions of the independent components, and this yields the formula required. \( \square \)

We stress in particular, that every Lévy process is the difference of two processes with only positive jumps. In general, these processes are not subordinators, but of the form in Theorem 42 plus a Brownian motion component. They can then both take positive and negative values.

**Example 46 (Variance Gamma process)** We introduced the Variance Gamma process as difference \( X = G - H \) of two independent Gamma subordinators \( G \) and \( H \). We can generalise the setting of Exercise A.1.2.(b) and allow \( G_1 \sim \text{Gamma}(\alpha_+, \beta_+) \) and \( H_1 \sim \text{Gamma}(\alpha_-, \beta_-) \). The moment generating function of the Variance Gamma process is

\[
\begin{align*}
\mathbb{E}(e^{\gamma X_t}) &= \mathbb{E}(e^{\gamma G_t}) \mathbb{E}(e^{-\gamma H_t}) = \left( \frac{\beta_+}{\beta_+ - \gamma} \right)^{\alpha_+ t} \left( \frac{\beta_-}{\beta_- + \gamma} \right)^{\alpha_- t} \\
&= \exp\left\{ t \int_0^\infty (e^{\gamma x} - 1) \alpha_+ x^{-1} e^{-\beta_+ x^t} dx \right\} \exp\left\{ t \int_0^\infty (e^{-\gamma y} - 1) \alpha_- y^{-1} e^{-\beta_- y} dy \right\} \\
&= \exp\left\{ t \int_0^\infty (e^{\gamma x} - 1) \alpha_+ |x|^{-1} e^{-\beta_+ |x|^t} dx + t \int_0^{-\infty} (e^{\gamma x} - 1) \alpha_- |x|^{-1} e^{-\beta_- |x|^t} dx \right\}.
\end{align*}
\]
and this is in \( \text{Lévy-Khintchine form} \) with 
\[
\nu(dx) = g(x)dx
\]
with
\[
g(x) = \begin{cases} 
\alpha_+ |x|^{-1} e^{-\beta_+ |x|} & x > 0 \\
\alpha_- |x|^{-1} e^{-\beta_- |x|} & x < 0
\end{cases}
\]
The process \((\Delta X_t)_{t \geq 0}\) is a \text{Poisson point process} with intensity function \(g\).

**Example 47 (CGMY process)** Theorem 45 encourages to specify \text{Lévy processes} by their characteristics. As a natural generalisation of the \text{Variance Gamma process}, Carr, Geman, Madan and Yor (CGMY) suggested the following for financial price processes
\[
g(x) = \begin{cases} 
C_+ \exp\{-G|x|\}|x|^{-Y-1} & x > 0 \\
C_- \exp\{-M|x|\}|x|^{-Y-1} & x < 0
\end{cases}
\]
for parameters \(C_+ > 0, G > 0, M > 0, Y \in [0, 2)\). While the \text{Lévy density} is a nice function, the \text{probability density function} of an \text{associated Lévy process} \(X_t\) is not available in closed form, in general. The CGMY model contains the \text{Gamma model} for \(Y = 0\). When this model is fitted to financial data, there is usually significant evidence against \(Y = 0\), so the CGMY model is more appropriate than the \text{Variance Gamma model}.

We can construct \text{Lévy processes} from their \text{Lévy density} and will also simulate from \text{Lévy densities}. Note that this way of modelling is easier than searching directly for infinitely divisible \text{probability density functions}.

### 7.2 Simulation via embedded random walks

“Simulation” usually refers to the realisation of a random variable using a computer. Most mathematical and statistical packages provide functions, procedures or commands for the generation of sequences of pseudo-random numbers that, while not random, show features of independent and identically distributed random variables that are adequate for most purposes. We will not go into the details of the generation of such sequences, but assume that we have a sequence \((U_k)_{k \geq 1}\) of independent \text{Unif}(0, 1) random variables.

If the increment distribution is explicitly known, we simulate via time discretisation.

**Method 1 (Time discretisation)** Let \((X_t)_{t \geq 0}\) be a \text{Lévy process} so that \(X_t\) has probability density function \(f_t\). Fix a time lag \(\delta > 0\). Denote \(F_t(x) = \int_{-\infty}^{x} f_t(y)dy\) and 
\[
F_t^{-1}(u) = \inf\{x \in \mathbb{R} : F_t(x) > u\}
\]
Then the process
\[
X^{(1,\delta)}_t = S_{\lfloor t/\delta \rfloor}, \quad \text{where } S_n = \sum_{k=1}^{n} Y_k \text{ and } Y_k = F_{\delta}^{-1}(U_k),
\]
is called the \text{time discretisation} of \(X\) with time lag \(\delta\).

One usually requires numerical approximation for \(F_{\delta}^{-1}\), even if \(f_t\) is available in closed form. That the approximations converge, is shown in the following proposition.

**Proposition 48** As \(\delta \downarrow 0\), we have \(X^{(1,\delta)}_t \rightarrow X_t\) in distribution.

**Proof:** We can employ a coupling proof: \(t\) is a.s. not a jump time of \(X\), so we have \(X_{\lfloor t/\delta \rfloor \delta} \rightarrow X_t\) a.s., and so convergence in distribution for \(X^{(1,\delta)}_t \sim X_{\lfloor t/\delta \rfloor \delta}\). \(\Box\)
Example 49 (Gamma processes) For Gamma processes, $F_t$ is an incomplete Gamma function, which has no closed-form expression, and $F_{t-1}$ is also not explicit, but numerical evaluations have been implemented in many statistical packages. There are also Gamma generators based on more uniform random variables. We display a range of parameter choices. Since for a Gamma$(1,1)$ process $X$, the process $(\beta^{-1}X_{\alpha t})_{t \geq 0}$ is Gamma$(\alpha, \beta)$:

$$E(\exp\{\gamma \beta^{-1} X_{\alpha t}\}) = \left(\frac{1}{1 - \gamma \beta^{-1}}\right)^{\alpha t} = \left(\frac{\beta}{\beta - \gamma}\right)^{\alpha t},$$

we chose $\alpha = \beta$ (keeping mean 1 and comparable spatial scale) but a range of parameters $\alpha \in \{0.1, 1, 10, 100\}$ on a fixed time interval $[0, 10]$. We “see” convergence to a linear drift as $\alpha \to \infty$ (for fixed $t$ this is due to the laws of large numbers).

---

**Figure 7.1:** Simulation of Gamma processes from random walks with Gamma increments

---

**Figure 7.2:** Random walk approximation to a Lévy process, as in Proposition 48
Example 50 (Variance Gamma processes) We represent the Variance Gamma process as the difference of two independent Gamma processes and focus on the symmetric case, so achieve mean 0 and fix variance 1 by putting $\beta = \alpha^2/2$; we consider $\alpha \in \{1, 10, 100, 1000\}$. We “see” convergence to Brownian motion as $\alpha \to \infty$ (for fixed $t$ due to the Central Limit Theorem).

![Graphs of Variance Gamma processes with different shape parameters](image)

Figure 7.3: Simulation of Variance Gamma processes as differences of random walks

7.3 R code – not examinable

The following code is posted on the course website as `gammavgamma.R`.

```r
psum <- function(vector){
b=vector;
b[1]=vector[1];
for (j in 2:length(vector)) b[j]=b[j-1]+vector[j]; b
}
gammarw <- function(a,p){
unif=runif(10*p,0,1)
pos=qgamma(unif,a/p,a);
space=psum(pos);
time=(1/p)*1:(10*p);
plot(time,space,
pch=".",
sub=paste("Gamma process with shape parameter",a,"and scale parameter",a))
}
```
vgammarw <- function(a,p){
    unifpos=runif(10*p,0,1)
    unifneg=runif(10*p,0,1)
    pos=qgamma(unifpos,a*a/(2*p),a);
    neg=qgamma(unifneg,a*a/(2*p),a);
    space=psum(pos-neg);
    time=(1/p)*1:(10*p);
    plot(time,space,
         pch=".",
         sub=paste("Variance Gamma process with shape parameter",a*a/2,
                    "and scale parameter",a))
}

Now you can try various values of parameters $a > 0$ and steps per time unit $p = 1/\delta$ in `gammarw(a,p)`, e.g.

```r
 gammarw(10,100)
 vgammaw(10,1000)
```
In practice, the increment distribution is often not known, but the Lévy characteristics are, so we have to simulate Poisson point processes of jumps, by “throwing away the small jumps” and then analyse (and correct) the error committed.

8.1 Simulation via truncated Poisson point processes

Example 51 (Compound Poisson process) Let \((X_t)_{t \geq 0}\) be a compound Poisson process with Lévy density \(g(x) = \lambda h(x)\), where \(h\) is a probability density function. Denote \(H(x) = \int_{-\infty}^{x} h(y)\,dy\) and \(H^{-1}(u) = \inf\{x \in \mathbb{R} : H(x) > u\}\). Let \(Y_k = H^{-1}(U_{2k})\) and \(Z_k = -\lambda^{-1} \ln(U_{2k-1})\), \(k \geq 1\). Then the process \(X^{(2, \varepsilon)}_t = L_{N_t} - b_{\varepsilon}t\), where \(S_n = \sum_{k=1}^{n} Y_k\), \(T_n = \sum_{k=1}^{n} Z_k\), \(N_t = \#\{n \geq 1 : T_n \leq t\}\), has the same distribution as \(X\).

Method 2 (Throwing away the small jumps) Let \((X_t)_{t \geq 0}\) be a Lévy process with characteristics \((a, \sigma^2, g)\), where \(g\) is not the multiple of a probability density function. Fix a jump size threshold \(\varepsilon > 0\) so that \(\lambda \varepsilon = \int_{\{x \in \mathbb{R} : \varepsilon < |x| \leq 1\}} g(x)\,dx > 0\), and write \(g(x) = \lambda \varepsilon h_{\varepsilon}(x), \quad |x| > \varepsilon, \quad h_{\varepsilon}(x) = 0, \quad |x| \leq \varepsilon,\) for a probability density function \(h_{\varepsilon}\). Denote \(H(x) = \int_{-\infty}^{x} h(y)\,dy\) and \(H^{-1}(u) = \inf\{x \in \mathbb{R} : H(x) > u\}\). Let \(Y_k = H^{-1}(U_{2k})\) and \(Z_k = -\lambda^{-1} \ln(U_{2k-1})\), \(k \geq 1\). Then the process \(X^{(2, \varepsilon), \delta}_t = \sigma B_{t_{\varepsilon}} + X^{(2, \varepsilon)}_t\), where \(S_n = \sum_{k=1}^{n} Y_k\), \(T_n = \sum_{k=1}^{n} Z_k\), \(N_t = \#\{n \geq 1 : T_n \leq t\}\), and \(b_{\varepsilon} = a - \int_{\{x \in \mathbb{R} : \varepsilon < |x| \leq 1\}} x g(x)\,dx\), is called the process with small jumps thrown away.

For characteristics \((a, \sigma^2, g)\) we can now simulate \(L_t = \sigma B_t + X_t\) by \(\sigma B_{t_{\varepsilon}}^{(1, \delta)} + X^{(2, \varepsilon)}_t\). The following proposition says that such approximations converge as \(\varepsilon \downarrow 0\) (and \(\delta \downarrow 0\)). This is illustrated in Figure 6.3.
Proposition 52 As \( \varepsilon \downarrow 0 \), we have \( X_t^{(2,\varepsilon)} \to X_t \) in distribution.

**Proof:** For a process with no negative jumps and characteristics \((0, 0, g)\), this is a consequence of the stronger Lemma 41, which gives a coupling for which convergence holds in the \( L^2 \) sense. For a general Lévy process with characteristics \((a, 0, g)\) that argument can be adapted, or we write \( X_t = at + P_t - N_t \) and deduce the result:

\[
\mathbb{E}(\exp\{i\lambda X_t^{(2,\varepsilon)}\}) = e^{iat} \mathbb{E}(\exp\{i\lambda P_t^{(2,\varepsilon)}\}) \mathbb{E}(\exp\{-i\lambda N_t^{(2,\varepsilon)}\}) \to e^{iat} \mathbb{E}(\exp\{i\lambda P_t\}) \mathbb{E}(\exp\{-i\lambda N_t\}) = \mathbb{E}(e^{i\lambda X_t}).
\]

\( \square \)

**Example 53 (Symmetric stable processes)** Symmetric stable processes \( X \) are Lévy processes with characteristics \((0, 0, g)\), where \( g(x) = c|x|^{-\alpha - 1}, x \in \mathbb{R} \setminus \{0\} \) for some \( \alpha \in (0, 2) \) (cf. Assignment 3.2.). We decompose \( X = P - N \) for two independent processes with no negative jumps and simulate \( P \) and \( N \). By doing this, we have

\[
\lambda_\varepsilon = \int_\varepsilon^\infty g(x)dx = \frac{c}{\alpha} \varepsilon^{-\alpha}, \quad H_\varepsilon(x) = 1 - \left(\frac{\varepsilon}{x}\right)^\alpha \quad \text{and} \quad H_\varepsilon^{-1}(u) = \varepsilon(1 - u)^{-1/\alpha}.
\]

For the simulation we choose \( \varepsilon = 0.01 \). We compare \( \alpha \in \{0.5, 1, 1.5, 1.8\} \). All processes are centred with infinite variance. Big jumps dominate the plots for small \( \alpha \). Recall that \( \mathbb{E}(e^{i\lambda X_t}) = e^{-b|\lambda|^2} \to e^{-b\lambda^2} \) as \( \alpha \uparrow 2 \), and we get, in fact, convergence to Brownian motion, the stable process of index 2.

![Simulation of symmetric stable processes from their jumps](image-url)
Example 54 (Stable processes with no negative jumps) For stable processes with no negative jumps, we have \( g(x) = c_+ x^{-\alpha - 1}, \) \( x > 0. \) The subordinator case \( \alpha \in (0, 1) \) was discussed in Assignment A.3.1. \( X_t = \sum_{s \leq t} \Delta_s. \) The case \( \alpha \in [1, 2), \) where compensation is required, is such that \( \mathbb{E}(X_1) = 0, \) i.e. \( a = -\int_1^\infty xg(x)dx. \) We choose \( \epsilon = 0.1 \) for \( \alpha \in \{0.5, 0.8\} \) and \( \epsilon = 0.01 \) for \( \alpha \in \{1.5, 1.8\}. \)

Strictly speaking, we take as triplet \((a, \sigma^2, g)\) in Theorem 45 \( g \) as given, but for \( \alpha \in (0, 1) \) we take \( a = \int_0^1 xg(x)dx \) so that

\[
\mathbb{E}(e^{i\lambda X_t}) = \exp \left\{ t \int_0^\infty (e^{i\lambda x} - 1)g(x)dx \right\} \\
= \exp \left\{ -t \left( -i\lambda a - \int_0^\infty (e^{i\lambda x} - 1 - i\lambda x1_{\{|x|\leq 1\}})g(x)dx \right) \right\},
\]

since compensation of small jumps is not needed and we obtain a subordinator if we do not compensate, and for \( \alpha \in (1, 2), \) we take \( a = -\int_1^\infty xg(x)dx \) so that

\[
\mathbb{E}(e^{i\lambda X_t}) = \exp \left\{ t \int_0^\infty (e^{i\lambda x} - 1 - i\lambda x)g(x)dx \right\} \\
= \exp \left\{ -t \left( -i\lambda a - \int_0^\infty (e^{i\lambda x} - 1 - i\lambda x1_{\{|x|\leq 1\}})g(x)dx \right) \right\},
\]

since we can compensate all jump and achieve \( \mathbb{E}(X_t) = 0. \) Only with these choices we obtain the Lévy processes with no negative jumps that satisfy the scaling property.

This discussion shows that the representation in Theorem 45 is artificial and representations with different compensating drifts are often more natural.
8.2 Generating specific distributions

In this course, we will not go into the computational details of simulations. However, we do point out some principles here that lead to improved simulations, and we discuss some of the resulting modifications to the methods presented.

Often, it is not efficient to compute the inverse cumulative distribution function. For a number of standard distributions, other methods have been developed. We will here look at standard Normal generators. The Gamma distribution is discussed in Assignment A.4.2.

Example 55 (Box-Muller generator) Consider the following procedure

1. Generate two independent random numbers \( U \sim \text{Unif}(0,1) \) and \( V \sim \text{Unif}(0,1) \).
2. Set \( X = \sqrt{-2 \ln(U)} \cos(2\pi V) \) and \( Y = \sqrt{-2 \ln(U)} \sin(2\pi V) \).
3. Return the pair \((X,Y)\).

The claim is the \( X \) and \( Y \) are independent standard Normal random variables. The proof is an exercise on the transformation formula. First, the transformation is clearly bijective from \((0,1)^2\) to \(\mathbb{R}^2\). The inverse transformation can be worked out from \(X^2 + Y^2 = -2 \ln(U)\) and \(Y/X = \tan(2\pi V)\) as

\[
(U, V) = T^{-1}(X, Y) = (e^{-(X^2+Y^2)/2}, (2\pi)^{-1} \arctan(Y/X))
\]

(with an appropriate choice of the branch of \(\arctan\), which is not relevant here). The Jacobian of the inverse transformation is

\[
J = \begin{pmatrix}
    -xe^{-(x^2+y^2)/2} & -ye^{-(x^2+y^2)/2} \\
    \frac{1}{2\pi x^2 + y^2} & \frac{1}{2\pi x^2 + y^2}
\end{pmatrix}
\]

\[
\Rightarrow |\det(J)| = \frac{1}{2\pi} e^{-(x^2+y^2)/2}
\]

and so, as required,

\[
f_{X,Y}(x,y) = f_{U,V}(T^{-1}(x,y))|\det(J)| = \frac{1}{2\pi} e^{-(x^2+y^2)/2}.
\]

For a more efficient generation of standard Normal random variables, it turns out useful to first generate uniform random variables on the disk of radius 1:

Example 56 (Uniform distribution on the disk) For \( U_1 \sim \text{Unif}(0,1) \) and \( U_2 \sim \text{Unif}(0,1) \) independent, we have that \((V_1, V_2) = (2U_1 - 1, 2U_2 - 1)\) is uniformly distributed on the square \((-1,1)^2\) centered at \((0,0)\), which contains the disk \(D = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}\), and we have, in particular \(\mathbb{P}((V_1, V_2) \in D) = \pi/4\). Now, for all \(A \subset \mathbb{R}^2\), we have

\[
\mathbb{P}((V_1, V_2) \in A | (V_1, V_2) \in D) = \frac{\text{area}(A \cap D)}{\pi},
\]

so the conditional distribution of \((V_1, V_2)\) given \((V_1, V_2) \in D\) is uniform on \(D\). By the following lemma, this conditioning can be turned into an algorithm by repeated trials:

1. Generate two independent random numbers \( U_1 \sim \text{Unif}(0,1) \) and \( U_2 \sim \text{Unif}(0,1) \).
2. Set \((V_1, V_2) = (2U_1 - 1, 2U_2 - 1)\).

3. If \((V_1, V_2) \in D\), go to 4., else go to 1.

4. Return the numbers \((V_1, V_2)\).

The pair of numbers returned will be uniformly distributed on the disk \(D\).

**Lemma 57 (Conditioning by repeated trials)** Let \(X, X_1, X_2, \ldots\) be independent and identically distributed \(d\)-dimensional random vectors. Also let \(A \subset \mathbb{R}^d\) such that \(p = \mathbb{P}(X \in A) > 0\). Denote \(N = \inf\{n \geq 1 : X_i \in A\}\). Then \(N \sim \text{geom}(p)\) is independent of \(X_N\), and \(X_N\) has as its (unconditional) distribution the conditional distribution of \(X\) given \(X \in A\), i.e.

\[
\mathbb{P}(X_N \in B) = \mathbb{P}(X \in B | X \in A) \quad \text{for all } B \subset \mathbb{R}^d.
\]

**Proof:** We calculate the joint distribution

\[
\mathbb{P}(N = n, X_n \in B) = \mathbb{P}(X_1 \notin A, \ldots, X_{n-1} \notin A, X_n \in A \cap B) = (1 - p)^{n-1}\mathbb{P}(X_n \in A \cap B) = (1 - p)^{n-1}p\mathbb{P}(X_n \in B | X_n \in A).
\]

\[\square\]

We now get the following modification of the Normal generator:

**Example 58 (Polar method)** The following is a more efficient method to generate two independent standard Normal random variables:

1. Generate two independent random numbers \(U_1 \sim \text{Unif}(0, 1)\) and \(U_2 \sim \text{Unif}(0, 1)\).

2. Set \((V_1, V_2) = (2U_1 - 1, 2U_2 - 1)\) and \(S = V_1^2 + V_2^2\).

3. If \(S \leq 1\), go to 4., else go to 1.

4. Set \(P = \sqrt{-2(\ln(S))/S}\)

5. Return the pair \((X, Y) = (PV_1, PV_2)\).

The gain in efficiency mainly stems from the fact that no sine and cosine need to be computed. The method works because in polar coordinates \((V_1, V_2) = (R \cos(\Theta), R \sin(\Theta))\), we have independent \(S = R^2 \sim \text{Unif}(0, 1)\) and \(\Theta \sim \text{Unif}(0, 2\pi)\) (as is easily checked), so we can choose \(U = S\) and \(2\pi V = \Theta\) in the Box-Muller generator.
8.3 R code – not examinable

The following code is posted on the course website as \texttt{stable.R}.

\begin{verbatim}
stableonesided <- function(a,c,eps,p){
f=c*eps^(-a)/a;
 n=rpois(1,10*f);
t=runif(n,0,10);
y=(eps^(-a)-a*f*runif(n,0,1)/c)^(-1/a);
ytemp=1:n;res=(1:(10*p))/100;
 for (k in 1:(10*p)){for (j in 1:n){
 if(t[j]<=k/p)ytemp[j]<-y[j] else ytemp[j]<-0;
}}
 res
}

stable <- function(a,cp,cn,eps,p){
 pos=stableonesided(a,cp,eps,p);
 neg=stableonesided(a,cn,eps,p);
 space=pos-neg;time=(1/p)*1:(10*p);
 plot(time,space,
pch=".",
 sub=paste("Stable process with index",a,"and cplus=",cp,"and cminus=" ,cn))
}

stableonesidedcomp <- function(a,c,eps,p){
f=(c*eps^(-a))/a;
 n=rpois(1,10*f);
t=runif(n,0,10);
y=(eps^(-a)-a*f*runif(n,0,1)/c)^(-1/a);
ytemp=1:n;
 res=(1:(10*p))/100;
 for (k in 1:(10*p)){if (n!=0)for (j in 1:n){
 if(t[j]<=k/p)ytemp[j]<-y[j] else ytemp[j]<-0;
}}
 res
}
\end{verbatim}
Lecture 9
Simulation III

Reading: Ross 11.3, Schoutens Sections 8.1, 8.2, 8.4; Further reading: Kyprianou Section 3.3

9.1 Applications of the rejection method

Lemma 57 can be used in a variety of ways. A widely applicable simulation method is the rejection method. Suppose you have an explicit probability density function \( f \), but the inverse distribution function is not explicit. If \( h \geq cf \), for some \( c < 1 \) is a probability density function whose inverse distribution function is explicit (e.g. uniform or exponential) or from which we can simulate by other means, then the procedure

1. Generate a random variable \( X \) with density \( h \).
2. Generate an independent uniform variable \( U \).
3. If \( Uh(X) \leq cf(X) \), go to 4., else go to 1.
4. Return \( X \).

Proposition 59 The procedure returns a random variable with density \( f \).

Proof: Denote \( p = \mathbb{P}(Uh(X) \leq cf(X)) \). By Lemma 57 (applied to the vector \( (X, U) \)), the procedure returns a random variable with distribution

\[
\mathbb{P}(X \leq x | Uh(X) \leq cf(X)) = \frac{\mathbb{P}(X \leq x, Uh(X) \leq cf(X))}{p} = \frac{1}{p} \int_{-\infty}^{x} h(z) \mathbb{P}(U \leq cf(z)/h(z))dz = \frac{c}{p} \int_{-\infty}^{x} f(z)dz,
\]

and letting \( x \to \infty \) shows \( c = p \). \( \square \)

Example 60 (Gamma distribution) Note that the Gamma density for \( \alpha > 1 \) satisfies

\[
\frac{1}{\Gamma(\alpha)} \beta^{\alpha} x^{\alpha-1} e^{-\beta x} \leq \frac{\beta^{\alpha-1}}{\Gamma(\alpha)} \beta e^{-\beta x}
\]

so we can apply the procedure with \( f(x) = \beta e^{-\beta x} \) and \( c = \Gamma(\alpha)/\beta^{\alpha-1} \).
It is important that $c$ is not too small, since otherwise lots of iterations are needed until the random variable is returned. The number of iterations is geometrically distributed (first success in a sequence of independent Bernoulli trials) with success parameter $c$, so on average $1/c$ trials are required.

For the simulation via Poisson point processes (with truncation at $\varepsilon$, say), we can use properties of Poisson point processes to simulate separately jumps of sizes in intervals $I_n$, $n = 1, \ldots, n_0$ and can choose intervals $I_n$ so that the intensity function $g$ is almost constant.

**Example 61 (Distribution on $I_n$)** Suppose that we simulate a Poisson point process on a bounded spatial interval $I_n = (a, b]$, with some intensity function $g : I_n \to [0, \infty)$. Then we can take uniform

$$h(x) = 1/(b - a) \quad \text{and} \quad c = \frac{\int_a^b g(x) \, dx}{(b - a) \max\{g(x) : a < x \leq b\}}.$$

and simulate $\text{Exp}(\int_a^b g(x) \, dx)$-spaced times $T_n$ and spatial coordinates $\Delta T_n$ by the rejection method with $h$ and $c$ as given.

### 9.2 “Errors increase in sums of approximated terms.”

Methods 1 and 2 are based on sums of many, mostly small, independent identically distributed random variables. As $\delta \downarrow 0$ or $\varepsilon \downarrow 0$, these are more and more smaller and smaller random variables. If each is affected by a small error, then adding up these errors makes the approximations worse whereas the precision should increase.

For Method 1, this can often be prevented by suitable conditioning, e.g. on the terminal value:

**Example 62 (Poisson process)** A Poisson process with intensity $\lambda$ on the time interval $[0, 1]$ can be generated as follows:

1. Generate a Poisson random variable $N$ with parameter $\lambda$.
2. Given $N = n$, generate $n$ independent Unif(0, 1) random variables $U_1, \ldots, U_n$.
3. Return $X_t = \#\{1 \leq i \leq N : U_i \leq t\}$.

Clearly, this process is a Poisson process, since $X_n$ (and indeed $(X_{t_1}, \ldots, X_{t_n-1})$) is obtained from $X_1 \sim \text{Poi}(\lambda)$ by thinning as in Section 4.2.

This is not of much practical use since we would usually simulate a Poisson random variable by evaluating a unit rate Poisson process (simulated from standard exponential interarrival times) at $\lambda$. In the case of Brownian motion (and the Gamma process, see Assignment A.4.3.), however, such conditioning is very useful and can then be iterated, e.g. in a dyadic scheme:

**Example 63 (Brownian motion)** Consider the following method to generate Brownian motion on the time interval $[0, 1]$.
1. Set \( X_0 = 0 \) and generate \( X_1 \sim \text{Normal}(0,1) \) hence specifying \( X_{k2^{-n}} \) for \( n = 0, k = 0, \ldots, 2^n \).

2. For \( k = 1, \ldots, 2^n \), conditionally given \( X_{(k-1)2^{-n}} = x \) and \( X_{k2^{-n}} = z \), generate

\[
X_{(2k-1)2^{-n-1}} \sim \text{Normal} \left( \frac{x + z}{2}, 2^{-n-2} \right)
\]

3. If the required precision has been reached, stop, else increase \( n \) by 1 and go back to 2.

This process is Brownian motion, since the following lemma shows that Brownian motion has these conditional distributions. Specifically, the \( n = 0, k = 1 \) case of 2. is obtained directly for \( s = 1/2, t = 1 \). For \( n \geq 1, k = 1, \ldots, 2^n \), note that \( X_{(2k-1)2^{-n-1}} - X_{(k-1)2^{-n}} \) is independent of \( X_{(k-1)2^{-n}} \) and so, we are really saying that for Brownian motion

\[
X_{(2k-1)2^{-n-1}} - X_{(k-1)2^{-n}} \sim \text{Normal} \left( \frac{z - x}{2}, 2^{-n-2} \right),
\]

conditionally given \( Z_{k2^{-n}} - Z_{(k-1)2^{-n}} = z - x \), which is equivalent to the specification in 2.

A further advantage of this method is that \( \delta = 2^{-n} \) can be decreased without having to start afresh. Previous less precise simulations can be refined.

**Lemma 64** Let \( (X_t)_{t \geq 0} \) be Brownian motion and \( 0 < s < t \). Then, the conditional distribution of \( X_s \) given \( X_t = z \) is \( \text{Normal}(zs/t, s(t - s)/t) \).

**Proof:** Note that \( X_s \sim \text{Normal}(0, s) \) and \( X_t - X_s \sim \text{Normal}(0, t - s) \) are independent. By the transformation formula \( (X_s, X_t) \) has joint density

\[
f_{X_s,X_t}(x,z) = \frac{1}{2\pi \sqrt{s(t-s)}} \exp \left\{ -\frac{x^2}{2s} - \frac{(z-x)^2}{2(t-s)} \right\},
\]

and so the conditional density is

\[
f_{X_s|X_t=z}(x) = \frac{f_{X_s,X_t}(x,z)}{f_{X_t}(z)} = \frac{1}{\sqrt{2\pi \sqrt{s(t-s)/t}}} \exp \left\{ -\frac{x^2}{2s} - \frac{(z-x)^2}{2(t-s)} + \frac{z^2}{2t} \right\} = \frac{1}{\sqrt{2\pi \sqrt{s(t-s)/t}}} \exp \left\{ -\frac{(x-zs/t)^2}{2s(t-s)/t} \right\}
\]

For Method 2, we can achieve similar improvements by simulating \( (\Delta_t)_{t \geq 0} \) in stages. Choose a strictly decreasing sequence \( \infty = a_0 > a_1 > a_2 > \ldots > 0 \) of jump size thresholds with \( a_n \downarrow 0 \) as \( n \to \infty \)

\[
\Delta^{(k)}_t = \Delta_t 1_{\{a_k \leq \Delta_t < a_{k-1}\}}, \quad \Delta^{(-k)}_t = \Delta_t 1_{\{-a_{k} \geq \Delta_t > -a_{k-1}\}}, \quad k \geq 1, t \geq 0.
\]

Simulate the Poisson counting processes \( N^{(k)} \) associated with \( \Delta^{(k)} \) as in Example 62 and otherwise construct

\[
Z^{(k)}_t = \sum_{s \leq t} \Delta^{(k)}_s - t \int_{a_k}^{a_{k-1}} x 1_{\{0 < x < 1\}} g(x) dx
\]

as in Method 2 and include so many \( k = \pm 1, \pm 2, \ldots \) as precision requires.
9.3 Approximation of small jumps by Brownian motion

**Theorem 65 (Asmussen-Rosinski)** Let $(X_t)_{t \geq 0}$ be a Lévy process with characteristics $(a,0,g)$. Denote

$$\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} x^2 g(x) dx$$

If $\sigma(\varepsilon)/\varepsilon \to \infty$ as $\varepsilon \downarrow 0$, then

$$\frac{X_t - X_t^{(2,\varepsilon)}}{\sigma(\varepsilon)} \to B_t \quad \text{in distribution as } \varepsilon \downarrow 0$$

for an independent Brownian motion $(B_t)_{t \geq 0}$.

If $\sigma(\varepsilon)/\varepsilon \to \infty$, it is well-justified to adjust Method 2 to set

$$X_t^{(2+\varepsilon)} = X_t^{(2,\varepsilon)} + \sigma(\varepsilon) B_t$$

for an independent Brownian motion. In other words, we may approximate the small jumps by an independent Brownian motion.

**Example 66 (CGMY process)** The CGMY process is a popular process in Mathematical Finance. It is defined via its characteristics $(0,0,g)$, where

$$g(x) = C \exp\{-G|x|\} |x|^{-Y-1}, \quad x < 0, \quad g(x) = C \exp\{-M|x|\} |x|^{-Y-1}, \quad x > 0,$$

for some $C \geq 0$, $G > 0$, $M > 0$ and $Y < 2$. Let $(X_t)_{t \geq 0}$ be a CGMY process. We calculate

$$\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} x^2 g(x) dx \leq C \int_{-\varepsilon}^{\varepsilon} |x|^{1-Y} dx = \frac{2C}{2-Y} \varepsilon^{2-Y}$$

and for every given $\delta > 0$ and all $\varepsilon > 0$ small enough, the same quantity with $C$ replaced by $C - \delta$ is a lower bound, so that

$$\frac{\sigma(\varepsilon)}{\varepsilon} \sim \sqrt{\frac{2C}{2-Y}} \varepsilon^{-Y/2} \to \infty \iff Y > 0$$

Hence an approximation of the small jumps of size $(-\varepsilon, \varepsilon)$ thrown away by a Brownian motion $\sigma(\varepsilon) B_t$ is appropriate if and only if $Y > 0$. In fact, for $Y < 0$, the process has finite jump intensity, so all jumps can be simulated. Therefore, only the case $Y = 0$ is problematic. This is the Variance Gamma process (and its asymmetric companions).

Whether or not we can approximate small jumps by a Brownian motion, we have to decide what value of $\varepsilon$ to choose. By the independence properties of Poisson point processes, the remainder term that $X_t - X_t^{2,\varepsilon}$ is a (zero mean, for $\varepsilon < 1$) Lévy process with intensity function $g$ on $[-\varepsilon, \varepsilon]$ and variance

$$\sigma^2(\varepsilon) = \text{Var}(X_t - X_t^{(2,\varepsilon)})$$

(let $\delta \downarrow 0$ in the proof of Lemma 41). We can choose $\varepsilon$ so that the accuracy of $X_t^{(2,\varepsilon)}$ is within an agreed deviation $h$, i.e. e.g. $2\sigma(\varepsilon) = h$. In the setting of Theorem 65, this means that a deviation of $X_t$ from $X_t^{(2,\varepsilon)}$ by more than $h$ would happen with probability about 0.05.
9.4 Appendix: Consolidation on Poisson point processes

This section and the next should not be necessary this year, because the relevant material has been included in earlier lectures. They may still be useful as a reminder of key concepts.

We can consider Poisson point processes \((\Delta_t)_{t \geq 0}\) in very general spaces, e.g. (topological) spaces \((E, \mathcal{O})\) where we have a collection/notion of open sets \(O \in \mathcal{O}\) (and an associated Borel \(\sigma\)-algebra \(\mathcal{B} = \sigma(\mathcal{O})\), the smallest \(\sigma\)-algebra that contains all open sets, for which we also require that \(\{x\} \in \mathcal{B}\) for all \(x \in E\) and \(G = \{(x, x) : x \in E\} \in \mathcal{B} \otimes \mathcal{B}\).

We just require that \((\Delta_t)_{t \geq 0}\) is a family of \(\Delta_t \in E \cup \{0\}\) such that there is an intensity (Borel) measure \(\nu\) on \(E\) with \(\nu(\{x\}) = 0\) for all \(x \in E\),

(a) for disjoint \(A_1 = (a_1, b_1] \times O_1, \ldots, A_n = (a_n, b_n] \times O_n, O_i \in \mathcal{O}\), the counts
\[
N(A_i) = N((a_1, b_1] \times O_i) = \# \{t \in (a_i, b_i] : \Delta_t \in O_i\}, \quad i = 1, \ldots, n
\]
are independent random variables and

\(\text{jnhom}(b)\) \(N(A_i) \sim \text{Poi}((b_i - a_i)\nu(O_i)).\)

For us, \(E = \mathbb{R} \setminus \{0\}\) is the space of jump sizes, and \(\nu(O_i) = \nu((c_i, d_i)) = \int_{c_i}^{d_i} g(x) \, dx\) for an intensity function \(g : \mathbb{R} \setminus \{0\} \rightarrow [0, \infty)\). Property (a) for all open sets is then equivalent to property (a) for all measurable sets or all half-open intervals or all closed intervals etc. (all that matters is that the collection generates the Borel \(\sigma\)-algebra). It is an immediate consequence of the definition (and this discussion) that for (measurable) disjoint \(B_1, B_2, \ldots \subset \mathbb{R} \setminus \{0\}\), the “restricted” processes
\[
\Delta_t^{(i)} = \Delta_t 1_{\{\Delta_t \in B_i\}}, \quad t \geq 0,
\]
are also Poisson point processes with the restriction of \(g\) to \(B_i\) as intensity function, and they are independent. We used this fact crucially and repeatedly in two forms. Firstly, for \(B_1 = (0, \infty)\) and \(B_2 = (-\infty, 0)\) (and \(B_3 = B_4 = \ldots = \emptyset\)), we consider Poisson point processes of positive points (jump sizes) and of negative points (jump sizes). We constructed from them independent Lévy processes. Secondly, for a sequence \(\infty = a_0 > a_1 > a_2 > \ldots\), we considered \(B_i = [a_i, a_{i-1})\), \(i \geq 1\), so as to simulate separately independent Lévy processes (in fact compound Poisson processes with linear drift) with jump sizes only in \(B_i\).

9.5 Appendix: Consolidation on the compensation of jumps

The general Lévy process requires compensation of small jumps in its approximation by processes with no jumps in \((-\varepsilon, \varepsilon)\), as \(\varepsilon \downarrow 0\). This is reflected in its characteristic function of the form
\[
\mathbb{E}(e^{i\lambda X_t}) = e^{-t\psi(\lambda)}, \quad \psi(\lambda) = -ia_1 \lambda + \frac{1}{2} \sigma^2 \lambda^2 - \int_{-\infty}^{\infty} (e^{i\lambda x} - 1 - i\lambda x 1_{\{|x| \leq 1\}}) \nu(dx), \quad \lambda \in \mathbb{R}, \quad (1)
\]
where usually \( \nu(dx) = g(x)dx \). This is a parametrisation by \((a_1, \sigma^2, \nu)\) or \((a_1, \sigma^2, g)\), where we require the (weak) integrability condition \( \int_{-\infty}^{\infty} (1 \wedge x^2) \nu(dx) < \infty \).

The first class of Lévy processes that we constructed systematically were subordinators, where no compensation was necessary. We parametrised them by parameters \( a_2 \geq 0 \) and \( g : (0, \infty) \to [0, \infty) \) (or \( \nu \) measure on \((0, \infty))\) so that the moment generating function is of the form

\[
\mathbb{E}(e^{\gamma X_t}) = e^{\Psi(\gamma)}, \quad \Psi(\gamma) = a_2 \gamma + \int_0^\infty (e^{\gamma x} - 1) \nu(dx), \quad \gamma \leq 0. \tag{2}
\]

We required the stronger integrability condition \( \int_0^\infty (1 \wedge x) \nu(dx) < \infty \). Similarly, for differences of subordinators, we have a characteristic function

\[
\mathbb{E}(e^{i\lambda X_t}) = e^{-t\psi(\lambda)}, \quad \psi(\lambda) = -ia_2 \lambda + \frac{1}{2} \sigma^2 \lambda^2 - \int_{-\infty}^\infty (e^{i\lambda x} - 1) \nu(dx), \tag{3}
\]

under the stronger integrability condition \( \int_{-\infty}^\infty (1 \wedge |x|) \nu(dx) < \infty \). Compensation in (1) is only done for small jumps. This is, because, in general the indicator \( 1_{\{|x| \leq 1\}} \) cannot be omitted. However, if \( \int_{-\infty}^\infty (x \wedge x^2) \nu(dx) < \infty \), then we can also represent

\[
\mathbb{E}(e^{i\lambda X_t}) = e^{-t\psi(\lambda)}, \quad \psi(\lambda) = -ia_2 \lambda + \frac{1}{2} \sigma^2 \lambda^2 - \int_{-\infty}^\infty (e^{i\lambda x} - 1 - i\lambda x) \nu(dx). \tag{4}
\]

Equations (1), (3) and (4) are compatible whenever any two integrability conditions are fulfilled, since the linear (in \( \lambda \)) terms under the integral can be added to \( a_1 \) to give

\[
a_2 = a_1 + \int_{\{|x| \leq 1\}} x \nu(dx) \quad \text{and} \quad a_3 = a_1 - \int_{\{|x| > 1\}} x \nu(dx).
\]

Note that then (by differentiation at \( \lambda = 0 \)), we get \( a_3 = \mathbb{E}(X_1) \). If \( a_3 = 0 \), then \((X_t)_{t \geq 0}\) is a martingale. On the other hand, for processes with finite jump intensity, i.e. under the even stronger integrability condition \( \int_{-\infty}^\infty g(x)dx < \infty \), we get \( a_1 \) as the slope of the paths of \( X \) between the jumps. Both \( a_1 \) and \( a_3 \) are therefore natural parameterisations, but not available, in general. \( a_2 \) is available in general, but does not have such a natural interpretation.

We use characteristic functions for similar reasons: in general, moment generating functions do not exist. If they do, i.e. under a strong integrability condition \( \int_1^\infty e^{\gamma x} \nu(dx) < \infty \) for some \( \gamma > 0 \) or \( \int_{-\infty}^{-1} e^{\gamma x} \nu(dx) < \infty \) for some \( \gamma < 0 \), we get

\[
\mathbb{E}(e^{X_t}) = e^{\Psi(\gamma)}, \quad \Psi(\gamma) = a_1 \gamma + \frac{1}{2} \sigma^2 \gamma^2 + \int_{-\infty}^\infty (e^{\gamma x} - 1 - \gamma x 1_{\{|x| \leq 1\}}) g(x)dx. \tag{5}
\]

Moment generating functions are always defined on an interval \( I \) possibly including end points \( \gamma_- \in [-\infty, 0] \) and/or \( \gamma_+ \in [0, \infty) \), we always have \( 0 \in I \), but maybe \( \gamma_- = \gamma_+ = 0 \). If \( 1 \in I \) and \( a_1 \) is such that \( \Psi(1) = 0 \), then \((e^{X_t})_{t \geq 0}\) is a martingale.
Lecture 10

Lévy markets and incompleteness

Reading: Schoutens Chapters 3 and 6

10.1 Arbitrage-free pricing (from B10b)

By Donsker’s Theorem, Brownian motion is the scaling limit of most random walks and in particular of the simple symmetric random walk \( R_n = X_1 + \ldots + X_n \) where \( X_1, X_2, \ldots \) are independent with \( \mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2 \).

**Corollary 67** For simple symmetric random walk \((R_n)_{n \geq 0}\), we have \( e^{R_n/\sqrt{n}} \to e^{B_t} \), geometric Brownian motion, in distribution as \( n \to \infty \).

**Proof:** First note that \( E(X_1) = 0 \) and \( \text{Var}(X_1) = 1 \). Now for all \( x > 0 \)
\[
\mathbb{P}(e^{R_n/\sqrt{n}} \leq x) = \mathbb{P}(R_n/\sqrt{n} \leq \ln(x)) \to \mathbb{P}(B_t \leq \ln(x)) = \mathbb{P}(e^{B_t} \leq x),
\]
by the Central Limit Theorem. \( \square \)

This was convergence of for fixed \( t \). Stronger convergence, locally uniformly in \( t \) can also be shown. Note that \((R_n)_{n \geq 0}\) is a martingale, and so is \((B_t)_{t \geq 0}\). However,
\[
E(e^{R_n}) = \left(\frac{1}{2} e^{-1} + \frac{1}{2} e\right)^n \to \infty
\]

**Proposition 68** For non-symmetric simple random walk \((R_n)_{n \geq 0}\) with \( \mathbb{P}(X_i = 1) = p \), the process \((e^{R_n})_{n \geq 0}\) is a martingale if and only if \( p = 1/(1 + e) \).

**Proof:** By the fourth and first rules for conditional expectations, we have
\[
E(e^{R_{n+1}}|e^{R_0}, \ldots, e^{R_n}) = E(e^{R_n e^{X_{n+1}}}|e^{R_0}, \ldots, e^{R_n}) = e^{R_n} E(e^{X_{n+1}})
\]
and so, \((e^{R_n})_{n \geq 0}\) is a martingale if and only if
\[
1 = E(e^{X_{n+1}}) = pe + (1-p)e^{-1} \iff p(e-1/e) = 1 - 1/e \iff p = 1/(e+1). \quad \square
\]

The argument works just assuming that \( R_n = X_1 + \ldots + X_n \), \( n \geq 0 \), satisfies \( \mathbb{P}(|X_{n+1}| = 1|R_0 = r_0, \ldots, R_n = r_n) = 1 \). Among all joint distributions, the non-symmetric exponentiated random walk with \( p = 1/(1 + e) \) is the only martingale.

The concept of *arbitrage-free* pricing in binary models is leaving aside any randomness. We will approach the Black-Scholes model from discrete models.
Suppose we have a risky asset with random price process $S$ per unit and a risk-free asset with deterministic value function $A$ per unit. Consider portfolios $(U, V)$ of $U$ units of the risky asset and $V$ units of the risk-free asset. We allow that $(U_t, V_t)$ depends on the performance of $(S_s)_{0 \leq s \leq t}$ but not on $(S_s)_{s > t}$. We denote the value of the portfolio at time $t$ by $W_t = U_t S_t + V_t A_t$. The composition of the portfolio may change with time, but we consider only self-financing ones, for which any risky asset bought is paid for from the risk-free asset holdings and vice versa. We say that arbitrage opportunities exist if there is a self-financing portfolio process $(U, V)$ and a time $t$ so that $\mathbb{P}(W_t = 0) = 1$, $\mathbb{P}(W_t \geq 0) = 1$ and $\mathbb{P}(W_t > 0) > 0$. We will be interested in models where no arbitrage opportunities exist.

Example 69 (One-period model) There are two scenarios “up” and “down” (to which we may later assign probabilities $p \in (0, 1)$ and $1 - p$). The model consists of $(S_0, S_1)$ only, where $S_0$ changes to $S_1(\text{up})$ or $S_1(\text{down})$ after one time unit. The risk-free asset will evolve from $A_0$ to $A_1$. At time 0, we have $W_0 = U_0 S_0 + V_0 A_0$. At time 1, the value will change to either

$$W_1(\text{up}) = U_0 S_1(\text{up}) + V_0 A_1 \quad \text{or} \quad W_1(\text{down}) = U_0 S_1(\text{down}) + V_0 A_1. \quad (1)$$

It is easily seen that arbitrage opportunities occur if and only if $A_1/A_0 \geq S_1(\text{up})/S_0$ or $S_1(\text{down})/S_0 \geq A_1/A_0$, i.e. if one asset is uniformly preferable to the other.

A derivative security (or contingent claim) with maturity $t$ is a contract that provides the owner with a payoff $W_t$ dependent on the performance of $(S_s)_{0 \leq s \leq t}$. If there is a self-financing portfolio process $(U, V)$ with value $W_t$ at time $t$, then such a portfolio process is called a hedging portfolio process replicating the contingent claim. The value $W_0 = U_0 S_0 + V_0 A_0$ of the hedging portfolio at time 0 is called the arbitrage-free price of the derivative security. It is easily seen that there would be an arbitrage opportunity, if the derivative security was available to buy and sell at any other price (as an additional asset in the model). In general, not all contingent claims can be hedged.

Example 69 (One-period model, continued) Consider any contingent claim, i.e. a payoff of $W_1(\text{up})$ or $W_1(\text{down})$ according to whether scenario “up” or “down” take place. Equations (1) can now be used to set up a hedging portfolio $(U_0, V_0)$ and calculate the unique arbitrage-free price. Note that the arbitrage-free price is independent of probabilities $p \in (0, 1)$ and $1 - p$ that we may assign to the two scenarios as part of our model specification. Because of the linearity of (1), there is a unique $q \in (0, 1)$ such that for all contingent claims $W_1 : \{\text{up, down}\} \to \mathbb{R}$

$$W_0 = q W_1(\text{up}) + (1 - q) W_1(\text{down}).$$

If we refer to $q$ and $1 - q$ as probabilities of “up” and “down”, then $W_0$ is the expectation of $W_1$ under this distribution. If $A_0 = A_1 = 1$, $S_0 = 1$, $S_1(\text{up}) = e$, $S_1(\text{down}) = e^{-1}$, then we identify (for $W_1 = S_1$ and hence $(U_0, V_0) = (1, 0)$) that $q = 1/(1 + e)$.

The property that every contingent claim can be hedged by a self-financing portfolio process is called completeness of the market model.
Example 70 (n-period model) Each of n periods has two scenarios, “up” and “down”, as is the case for the model $S_k = e^{R_k}$ for a simple random walk $(R_k)_{0 \leq k \leq n}$. Then there are $2^n$ different combinations of “up” $(X_k = 1)$ and “down” $(X_k = -1)$. A contingent claim at time $n$ is now any function $W_n$ assigning a payoff to each of these combinations. By a backward recursion using the one-period model as induction step, we can work out hedging portfolios $(U_k, V_k)$ and the value $W_k$ of the derivative security at times $k = n - 1, n - 2, \ldots, 0$, where in each case, $(U_k, V_k)$ and $W_k$ will depend on previous “up”s and “down”s $X_1, \ldots, X_k$, so this is a specification of $2^n$ values each. $W_0$ will be the unique arbitrage-free price for the derivative security. The induction also shows that, for $A_0 = A_1 = \ldots = A_n = 1$, it can be worked out as

$$W_0 = \mathbb{E}(W_n(X_1, \ldots, X_n)), \quad \text{where } X_k \text{ independent with } \mathbb{P}(X_k = 1) = 1/(1 + e),$$

and that $(W_k)_{0 \leq k \leq n}$ is a martingale with $W_k = \mathbb{E}(W_n|X_1, \ldots, X_k)$, e.g. $(S_k)_{0 \leq k \leq n}$. If $A_k = (1 + i)^k = e^{\delta k}$, we get an arbitrage-free model if and only if $-1 < \delta < 1$, and then

$$W_0 = e^{-\delta n}\mathbb{E}(W_n), \quad \text{where } X_k \text{ independent with } \mathbb{P}(X_k = 1) = (e^{\delta n} - 1)/(e^{2n} - 1),$$

where now $(e^{-\delta k}W_k)_{0 \leq k \leq n}$ is a martingale. In particular, the n-period model is complete.

Example 71 (Black-Scholes model) Let $S_t = S_0 \exp(\sigma B_t + (\mu - \frac{1}{2} \sigma^2)t)$ for a Brownian motion $(B_t)_{t \geq 0}$ and two parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Also put $A_t = e^{\delta t}$. It can be shown that also in this model, every contingent claim can be hedged, i.e. the Black-Scholes model is complete. Moreover, the pricing of contingent claims $W_t$ can be carried out using the risk-neutral process

$$R_t = S_0 \exp\{\sigma B_t + (\delta - \frac{1}{2} \sigma^2)t\}, \quad t \geq 0,$$

where the drift parameter is $\delta$, not $\mu$. The discounted process $M_t = e^{-\delta t} R_t$ is a martingale and has analogous uniqueness properties to the martingale for the n-period model, but they are much more complicated to formulate here.

The arbitrage-free price of $W_t = G((S_s)_{0 \leq s \leq t})$ is now (for all $\mu \in \mathbb{R}$)

$$W_0 = e^{-\delta t}\mathbb{E}(G((S_s)_{0 \leq s \leq t})).$$

Examples are $G((S_s)_{0 \leq s \leq t}) = (S_t - K)^+$ for the European call option and $G((S_s)_{0 \leq s \leq t}) = (K - S_t)^+$ for the European put option. We will also consider path-dependent options such as Up-and-out barrier options with payoff $(S_t - K)^+1_{[S_t < H]}$, where $S_t = \sup_{0 \leq s \leq t} S_s$ and $H$ is the barrier. The option can only be exercised if the stock price does not exceed the barrier $H$ at any time before maturity.

10.2 Introduction to Lévy markets

The Black-Scholes model is widely used for option pricing in the finance industry, largely because many options can be priced explicitly and there are computationally efficient methods also for more complicated derivatives, that can be carried out frequently and for high numbers of options. However, its model fit is poor and any price that is obtained from the Black-Scholes model must be adjusted to be realistic. There are several models based on Lévy processes that offer better model fit, but the Black-Scholes methods for option pricing do not transfer one-to-one. Lévy processes give a very wide modelling freedom. For practical applications it is useful to work with parametric subfamilies. Several such families have been suggested.
Example 72 (CGMY process) Carr, Geman, Madan and Yor proposed a model with four parameters (the first letters of their names). It is defined via its intensity

\[ g(x) = C \exp\{-G|x|\} |x|^{Y-1}, \quad x < 0, \]
\[ g(x) = C \exp\{-M|x|\} |x|^{Y-1}, \quad x > 0. \]

Let \((X_t)_{t \geq 0}\) be a CGMY process. If \(M > 1\), then a risk-neutral price process can be modelled as \(R_t = R_0 \exp\{X_t - t(\phi(1) - \delta)\}\). Then the discounted process \(e^{-\delta t} R_t\) is a martingale, and it can be shown that arbitrage-free prices for contingent claims \(W_t = G((R_s)_{0 \leq s \leq t})\) can be calculated as

\[ W_0 = e^{-\delta t} E(G((R_s)_{0 \leq s \leq t})). \]

It can also be shown, however, that this is not the only way to obtain arbitrage-free prices, and other prices do not necessarily lead to arbitrage opportunities. Also, not every contingent claim can be hedged, the model is not complete.

### 10.3 Incomplete discrete financial markets

Essentially, arbitrage-free discrete models are complete only if the number of possible scenarios \(\omega_0, \ldots, \omega_m\) (for one period) is the same as the number of assets \(S_1^{(1)}, \ldots, S_1^{(m)}:\Omega = \{\omega_0, \ldots, \omega_m\} \to \mathbb{R}\) in the model, since this leads to a system of linear equations to relate a contingent claim \(W_1: \{\omega_0, \ldots, \omega_m\} \to \mathbb{R}\) to a portfolio \((U^{(1)}, \ldots, U^{(m)}, V)\)

\[ V_0 A_1 + \sum_{i=1}^{m} U_0^{(i)} S_1^{(i)}(\omega_j) = W_1(\omega_j), \quad j = 0, \ldots, m \]

that can usually be uniquely solved for \((U_0^{(1)}, \ldots, U_0^{(m)}, V_0)\), and we can read off

\[ W_0 = V_0 A_0 + \sum_{i=1}^{m} U_0^{(i)} S_0^{(i)}. \]

If the number of possible scenarios is higher, then the system does not have a solution, in general (and hedging portfolios will not exist, in general). If the number of possible scenarios is lower, there will usually be infinitely many solutions.

If the system has no solution in general, the model is incomplete, but this does not mean that there is no price. It means that there is not a unique price. We can, in general, get some lower and upper bounds for the price imposed by no-arbitrage. One way of approaching this is to add a derivative security to the market as a further asset with an initial price that keeps the no arbitrage property for the extended model. One can, in fact, add more and more assets until the model is complete. Then there exist unique probabilities \(q_j = P(\omega_j), 0 \leq j \leq m\), that make all discounted assets \((A_0/A_1) S^{(j)}, 1 \leq j \leq m\) (including the ones added to complete the market) martingales.

Example 73 (Ternary model) Suppose there are three scenarios, but only two assets. The model with \(S_0 = 1 = A_0 < A_1 = 2\), and \(1 = S_1(\omega_0) < 2 = S_1(\omega_1) < 3 = S_1(\omega_2)\) is easily seen to be arbitrage-free since \(S_1(\omega_0) < A_1 < S_1(\omega_2)\). The contingent claim \(0 = W_1(\omega_0) = W_1(\omega_1) < W_1(\omega_2) = 1\) can be hedged if and only if

\[ 2V_0 + U_0 = 0, \quad 2V_0 + 2U_0 = 0, \quad 2V_0 + 3U_0 = 1, \]
but the first two equations already imply \( U_0 = V_0 = 0 \) and then the third equation is false. Therefore the contingent claim \( W_1 \) cannot be hedged. The model is not complete.

For the model \((W, S, A)\) to be arbitrage-free we clearly require \( W_0 > 0 \) since otherwise we could make arbitrage with a portfolio \((1, 0, 0)\), just “buying” the security. Its cost at time zero is nonpositive and its value at time one is nonnegative and positive for scenario \( \omega_2 \). Now note that \((S, A)\) is arbitrage-free, so any arbitrage portfolio must be of the form \((-1/W_0, U_0, V_0)\) with zero value \(-1 + U_0 + V_0 = 0\) at time 0 and values at time 1

\[
U_0 + 2V_0 = 2 - U_0, \quad 2U_0 + 2V_0 = 2 > 0, \quad -1/W_0 + 3U_0 + 2V_0 = -1/W_0 + 2 + U_0,
\]

so that we need \( 2 \geq U_0 \geq 1/W_0 - 2 \), and this is possible if and only if \( W_0 \geq 1/4 \). Therefore, the range of arbitrage-free prices is \( W_0 \in (0, 1/4) \).

We can get these prices as expectations under martingale probabilities:

\[
1 = \frac{1}{2} E_q(S_1) = \frac{1}{2} q_0 + q_1 + \frac{3}{2} q_2, \quad W_0 = \frac{1}{2} E_q(W_1) = \frac{1}{2} q_2, \quad q_0 + q_1 + q_2 = 1.
\]

This is a linear system for \((q_0, q_1, q_2)\) that we solve to get

\[
q_0 = q_2 = 2W_0, \quad q_1 = 1 - 4W_0
\]

and this specifies a probability distribution on all three scenarios iff \( W_0 \in (0, 1/4) \). Since we can express every contingent claim as a linear combination of \( A_1, S_1, W_1 \), we can now price every contingent claim \( X_1 \) under the martingale probabilities as \( X_0 = \frac{1}{2} E_q(X_1) \).
Lecture 11

Lévy markets and time-changes

11.1 Incompleteness and martingale probabilities in Lévy markets

By a Lévy market we will understand a model \((S, A)\) of a risky asset \(S_t = \exp\{X_t\}\) for a Lévy process \(X = (X_t)_{t \geq 0}\) and a deterministic risk-free bank account process, usually \(A_t = e^{\delta t}, \ t \geq 0\). We exclude deterministic \(X_t = \mu t\) in the sequel.

**Theorem 74 (No arbitrage)** A Lévy market allows arbitrage if and only if either \(X_t - \delta t\) is a subordinator or \(\delta t - X_t\) is a subordinator.

**Proof:** We only prove that these cases lead to arbitrage opportunities. If \(X_t - \delta t\) is a subordinator, then the portfolio \((1, -1)\) is an arbitrage portfolio. \(\Box\)

The other direction of proof is difficult, since we would need technical definitions of admissible portfolio processes and related quantities.

No arbitrage is closely related (almost equivalent) to the existence of martingale probabilities. Formally, an equivalent martingale measure \(Q\) is a probability measure which has the same sets of zero probability as \(P\), i.e. under which the same things are possible/impossible as under \(P\), and under which \((e^{-\delta t}S_t)_{t \geq 0}\) is a martingale. For simplicity, we will not bother about this passage to a so-called risk-neutral world that is different from the physical world. Instead, we will consider models where \((e^{-\delta t}S_t)_{t \geq 0}\) is already a martingale. Prices of the form \(W_0 = e^{-\delta t}E_Q(W_t)\) are then arbitrage-free prices. The range of arbitrage-free prices is

\[
\{e^{-\delta t}E_Q(W_t) : Q \text{ martingale measure equivalent to } P\}.
\]

The proof of incompleteness is also difficult, but the result is not hard to state:

**Theorem 75 (Completeness)** A Lévy market is complete if and only if \((X_t)_{t \geq 0}\) is either a multiple of Brownian motion with drift, \(X_t = \mu t + \sigma B_t\) or a multiple of the Poisson process with drift, \(X_t = at + bN_t\) (with \((a - \delta)b < 0\) to get no arbitrage).

Completeness is closely related (almost equivalent) to the uniqueness of martingale probabilities. In an incomplete market, there are infinitely many choices for these martingale probabilities. This raises the question of how to make the right choice. While we can determine an arbitrage-free system of prices for all contingent claims, we cannot hedge the contingent claim, and this presents a risk to someone selling e.g. options.
11.2 Option pricing by simulation

If we are given a risk-neutral price process (martingale) \((S_t)_{t \geq 0}\), we can price contingent claims \(G((S_s)_{0 \leq s \leq t})\) as expectations

\[
P = e^{-\delta t} \mathbb{E}(G((S_s)_{0 \leq s \leq t})).
\]

Often, such expectations are difficult to work out theoretically or numerically, particularly for path-dependent options such as barrier options. Monte-Carlo simulation always works, by the strong law of large numbers:

\[
\frac{1}{n} \sum_{k=1}^{n} G((S_s^{(k)})_{0 \leq s \leq t}) \to \mathbb{E}(G((S_s)_{0 \leq s \leq t})) \quad \text{almost surely},
\]

as \(n \to \infty\), where \((S_s^{(k)})_{0 \leq s \leq t}\) are independent copies of \((S_s)_{0 \leq s \leq t}\). By simulating these copies, we can approximate the expectation on the right to get the price of the option.

Figure 11.1: Option pricing by simulation

11.3 Time changes

Lévy markets are one way to address shortcomings of the Black-Scholes model. Particularly quantities such as one-day return distributions can be fitted well. Other possibilities include modifications to the Black-Scholes model, where the speed of the market is modelled separately. The rationale behind this is to capture days with increased activity (and hence larger price movements) by notions of operational versus real time. In operational time, the price process follows a Brownian motion, but in real time, a busy day corresponds to several days in operational time, while a quiet day corresponds to a fraction of a day in operational time.

The passage from operational to real time is naturally modelled by a time-change \(y \mapsto \tau_y\), which we will eventually model by a stochastic process built from a Poisson point process. The price process is then \((B_{\tau_y})_{y \geq 0}\). This stochastic process cannot be observed directly in practice, but approximations of quadratic variation permit to estimate the time change.

The most elementary time change is for \(\tau_y = f(y)\), a deterministic continuous strictly increasing function \(f : [0, \infty) \to [0, \infty)\) with \(f(0) = 0\) and \(f(\infty) = \infty\). In this case, the
time-changed process $Z_y = X_{f(y)}$, $y \geq 0$, visits the same states as $X$ in the same order as $X$, performing the same jumps as $X$, but travelling at a different speed. Specifically, if $f(y) \ll y$, then, by time $y$, the process $X$ will have gone to $X_y$, but $Z$ only to $Z_y = X_{f(y)}$. We say that $Z$ has evolved more slowly than $X$, and faster if instead $f(y) \gg y$. If $f$ is differentiable, we can more appropriately make local speed statements according to whether $f'(y) < 1$ or $f'(y) > 1$. Note, however, that “speed” really is “relative speed” when comparing $X$ and $Z$, since $X$ is not “travelling at unit speed” in a sense of rate of spatial displacement; jumps and particularly unbounded variation make such notions useless. We easily calculate
\[
E(e^{i\lambda Z_y}) = E(e^{i\lambda X_{f(y)}}) = e^{-f(t)\psi(\lambda)}, \quad \text{if } E(e^{i\lambda X_t}) = e^{-t\psi(\lambda)}.
\]
and see that $Z$ is a stochastic process with independent increments and right-continuous paths with left limits, but will only have stationary increments if $f(y) = cy$ for all $y \geq 0$ and some $c \in (0, \infty)$.

**Example 76 (Foreign exchange rates)** Suppose that the EUR/USD-exchange rate today is $S_0$ and you wish to model the exchange rate $(S_t)_{t \geq 0}$ over the next couple of days. As a first model you might think of
\[
S_t = S_0 \exp\{\sigma B_t - t\sigma^2/2\},
\]
where $B$ is a standard Brownian motion $\sigma$ is a volatility parameter that measures the magnitude of variation. This magnitude is related to the amount of activity on the exchange markets and will be much higher for the EUR/USD-exchange rate than e.g. for the EUR/DKK-exchange rate (Danske Kroner, Danish crowns are not traded so frequently in such high volumes. Also, DKK is closely aligned with EUR due to strong economic ties between Denmark and the Euro countries).

However, in practice, trading activity is not constant during the day. When stock markets in the relevant countries are open, activity is much higher than when they are all closed and a periodic function $f' : [0, \infty) \to [0, \infty)$ can explain a good deal of this variability and provide a better model
\[
S_t = S_0 \exp\{\sigma B_{f(y)} - f(y)\sigma^2/2\} = S_0 \exp\{\tilde{B}_{f(y)} - \tilde{f}(y)/2\},
\]
where $\tilde{B}_s = \sigma B_{s/\sigma^2}$, $s \geq 0$, is also a standard Brownian motion and $\tilde{f}(y) = f(y)\sigma^2$ makes the parameter $\sigma$ redundant – the flexibility for $\tilde{f}$ retains all modelling freedom.

If we weaken the requirement of strict monotonicity to weak monotonicity and $f(y) = c$, $y \in [l, r)$, is constant on an interval, then $Z_y = X_c$, $y \in [l, r)$, during this interval. For a financial market model this can be interpreted as time intervals with no market activity, when the price will not change.

If we weaken the continuity of $f$ to allow (upward) jumps, then $Z_y = X_{f(y)}$, $y \geq 0$, does not evaluate $X$ everywhere. Specifically, if $\Delta f(y) > 0$ is the only jump of $f$, then $Z$ will visit the same points as $X$ in the same order until $X_{f(y-)}$, and then skip over $(X_{f(y-s)}+s)_{0 \leq s < \Delta f(y)}$ to directly jump to $X_{f(y)}$. In general, this is the behaviour at every jump of $f$.  

### Lecture 11: Lévy markets and time-changes
11.4 Quadratic variation of time-changed Brownian motion

In Section 6.1 we studied quadratic variation of Brownian motion in order to show that Brownian motion has infinite total variation (and is therefore not the difference of two increasing processes). Let us here look at quadratic variation of time-changed Brownian motion $Z_y = B_{f(y)}$ for an increasing function $f : [0, \infty) \to [0, \infty)$:

$$[Z]_t = p - \lim_{n \to \infty} [Z]_{t}^{(n)}, \quad \text{where} \quad [Z]_{t}^{(n)} = \sum_{j=1}^{[2^n t]} (Z_{j2^{-n}} - Z_{(j-1)2^{-n}})^2$$

and $p - \lim$ denotes a limit of random variables in probability. One may expect that $[B]_t = t$ implies that $[Z]_y = f(y)$, and this is true under suitable assumptions.

**Proposition 77** Let $B$ be Brownian motion and $f : [0, \infty) \to [0, \infty)$ continuous and increasing with $f(0) = 0$. Then $[Z]_y = f(y)$ for all $y \geq 0$.

**Proof:** The proof (for $Z$) is the same as for Brownian motion ($B$) itself, see Section 6.1 and Assignment 6.

Quadratic variation is accumulated locally. Under the continuity assumption of Brownian motion and its time change, it is the wiggly local behaviour of Brownian motion that generates quadratic variation. In Section 6.1 we showed that under the no-jumps assumption, positive quadratic variation implies infinite total variation. Hence, still under the no-jumps assumption, finite total variation implies zero quadratic variation. But what is the impact of jumps on quadratic variation? It can be shown as in Proposition 3 that

$$[f]_y \geq \sum_{s \leq y} |\Delta f_s|^2.$$

**Example 78** Consider a piecewise linear function $f : [0, \infty) \to [0, \infty)$ with slope 0.1 and jumps $\Delta f_{2k-1} = 1.8$, $k \geq 1$. Then $f(2k) = 2k$, but by Proposition 77

$$[f]_{2k} \geq \sum_{s \leq 2k} |\Delta f_s|^2 = k(1.8)^2 = 3.24k$$
and, in fact, this is an equality, since
\[
[f^{(n)}_{2k}] = k(1.8 + 2^{-n}0.1)^2 + (2^{n+1} - 1)k(2^{-n}0.1)^2 \rightarrow k(1.8)^2.
\]

Now define \(Z_y = B_{f(y)}\) and note that
\[
[Z^{(n)}_{2k}] = \sum_{i=1}^{k} \left( \sum_{j=1}^{2^n-1} (B_{(2i-2)+j2^{-n}0.1} - B_{(2i-2)+(j-1)2^{-n}0.1})^2 
+ (B_{2i-0.1} - B_{(2i-2)+0.1-2^{-n}})^2 
+ \sum_{j=1}^{2^n} (B_{2i-(j-1)2^{-n}0.1} - B_{2i-j2^{-n}0.1})^2 \right)
\rightarrow 2k0.1 + \sum_{i=1}^{k} (B_{2i-0.1} - B_{(2i-2)+0.1})^2,
\]
as \(n \rightarrow \infty\), which is actually \(0.1(2k) + \sum_{s \leq 2k} |\Delta Z_s|^2\). Note that this is a random quantity.

In general, quadratic variation consists of a continuous part due to Brownian fluctuations and the sum of squared jump sizes.
Lecture 12

Subordination and stochastic volatility

Subordination is the operation $X_{\tau_y}$, $y \geq 0$, for a Lévy (or more general Markov) process $(X_t)_{t \geq 0}$ and a subordinator $(\tau_y)_{y \geq 0}$. One distinguishes subordination in the sense of Bochner, where $X$ and $\tau$ are independent and subordination in the wide sense where $(\tau_y)_{y \geq 0}$ is a stopping time for all $y \geq 0$. These are both special cases of the more general concept of time change, where $(\tau_y)_{y \geq 0}$ does not have to be a subordinator.

12.1 Bochner’s subordination

Theorem 79 (Bochner) Let $(X_t)_{t \geq 0}$ be a Lévy process and $(\tau_y)_{y \geq 0}$ an independent subordinator. Then the process $Z_y = X_{\tau_y}$, $y \geq 0$, is a Lévy process, and we have

$$E(e^{i\lambda Z_y}) = e^{-\psi(\lambda)}$$

where $E(e^{i\lambda X_t}) = e^{-t\psi(\lambda)}$ and $E(e^{-q\tau_y}) = e^{-q\Phi(a)}$.

Proof: First calculate by conditioning on $\tau_y$ (assuming that $\tau_y$ is continuous with probability density function $f_{\tau_y}$)

$$E(\exp\{i\lambda Z_y\}) = E(\exp\{\gamma X_{\tau_y}\}) = \int_0^{\infty} f_{\tau_y}(t) E(\exp\{i\lambda X_t\}) dt = \int_0^{\infty} f_{\tau_y}(t) \exp\{-t\psi(\lambda)\} dt = e^{-\psi(-\lambda)}.$$

Now, for $r,s \geq 0$,

$$E(\exp\{i\lambda Z_y + i\mu(Z_{y+x} - Z_y)\}) = \int_0^{\infty} \int_0^{\infty} f_{\tau_y,\tau_{y+x}-\tau_y}(t,u) E(\exp\{i\lambda X_t + i\mu(X_{t+u} - X_t)\}) dt du = \int_0^{\infty} f_{\tau_y}(t) f_{\tau_x}(u) e^{-t\psi(\lambda)} e^{-u\psi(\mu)} dt du = e^{-\psi(\lambda)} e^{-\mu\Phi(\psi(\mu))},$$

so that we deduce that $Z_y$ and $Z_{y+x} - Z_y$ are independent, and that $Z_{y+x} - Z_y \sim Z_x$.

For the right-continuity of paths, note that

$$\lim_{\varepsilon \downarrow 0} Z_{y+\varepsilon} = \lim_{\varepsilon \downarrow 0} X_{\tau_y} = X_{\tau_y} = Z_y,$$
since \( \tau_y + \delta = \tau_{y+\varepsilon} \downarrow \tau_y \) and therefore \( X_{\tau_y+\delta} \to X_{\tau_y} \). For left limits, the same argument applies.

Note that \( \Delta Z_y = Z_y - Z_{y-} = X_{\tau_y} - X_{\tau_y-} \neq 0 \) can be non-zero if either \( \Delta \tau_y \neq 0 \), or \( \Delta X_{\tau_y} \neq 0 \), so \( Z \) inherits jumps from \( \tau \) and from \( X \). We have, with probability 1 for all \( y \geq 0 \) that

\[
\Delta Z_y = X_{\tau_y} - X_{\tau_y-} = \begin{cases} (\Delta X)_{\tau_y} & \text{if } (\Delta X)_{\tau_y} \neq 0, \\ X_{\tau_y} - X_{\tau_y-} & \text{if } \Delta \tau_y \neq 0. \end{cases}
\]

Note that we claim that \( X_{\tau_y-} = X_{\tau_y-} \), i.e. \( (\Delta X)_{\tau_y-} \neq 0 \) if \( \Delta \tau_y \neq 0 \), for all \( y \geq 0 \) with probability 1. This is due to the fact that the countable set of times \( \{\tau_y-, \tau_y : y \geq 0 \text{ and } \Delta \tau_y \neq 0\} \) is a.s. disjoint from \( \{t \geq 0 : \Delta X_t \neq 0\} \).

Note also that \( X_{\tau_y} = X_{\tau_y-} \) is possible with positive probability, certainly in the case of a compound Poisson process \( X \).

Heuristically, if \( X_t \) has density \( f_t \) and \( \tau \) Lévy density \( g_{\tau} \), then \( Z \) will have Lévy density

\[ g(z) := \int_0^\infty f_t(z) g_{\tau}(t) dt, \quad z \in \mathbb{R}, \tag{1} \]

since every jump of \( \tau \) of size \( \Delta \tau_y = t \) leads to a jump \( X_{\tau_y} - X_{\tau_y-} \sim X_t \), and the total intensity of jumps of size \( z \) receives contributions from \( \tau \)-jumps of all sizes \( t \in (0, \infty) \). We can make this precise as follows:

**Proposition 80** Let \( X \) be a Lévy process with probability density function \( f_t \) of \( X_t \), \( t \geq 0 \), \( \tau \) a subordinator with Lévy-Khintchine characteristics \((0, g_{\tau})\), then \( Z_y = X_{\tau_y} \) has Lévy-Khintchine characteristics \((0, 0, g)\), where \( g \) is given by (1).

**Proof:** Consider a Poisson point process \((\Delta_y)_{y \geq 0}\) with intensity function \( g \), then, by the Exponential Formula

\[
\mathbb{E} \left( \exp \left\{ i \lambda \sum_{s \leq y} \Delta_s 1_{\{|\Delta_s| \geq \varepsilon\}} \right\} \right)
= \exp \left\{ y \int_{-\infty}^\infty (e^{i \lambda z} - 1) g(z) 1_{\{|z| \geq \varepsilon\}} dz \right\}
= \exp \left\{ y \int_{-\infty}^0 \int_{-\infty}^\infty (e^{i \lambda z} - 1) f_t(z) 1_{\{|z| \geq \varepsilon\}} dz g_{\tau}(t) dt \right\}
\to \exp \left\{ -y \int_0^\infty (1 - e^{-i \lambda z}) g_{\tau}(t) dt \right\} = \exp \left\{ -y \Phi(\psi(\lambda)) \right\},
\]

as \( \varepsilon \downarrow 0 \), and this is the same distribution as we established for \( Z_y \).

Note that we had to prove convergence in distribution as \( \varepsilon \downarrow 0 \), since we have not studied integrability conditions for \( g \). This is done on Assignment sheet 6.

**Corollary 81** If \( X \) is Brownian motion and \( \tau \) has characteristics \((b, g_{\tau})\), then \( Z_y = X_{\tau_y} \) has characteristics \((0, b, g)\).
Proof: Denote \( \Phi_0(q) = \Phi(q) - bq \). Then the calculation in the proof of the proposition does not yield a characteristic exponent \( \Phi(\psi(\lambda)) \), but \( \Phi_0(\psi(\lambda)) \). Note that \( \Phi(\psi(\lambda)) = \Phi(\frac{1}{2}\lambda^2) = \frac{1}{2}b\lambda^2 + \Phi_0(\psi(\lambda)) \), so we consider
\[
b B_t + \sum_{s \leq t} \Delta_s
\]
for an independent Brownian motion \( B \), which has characteristic exponent as required.

Example 82 If we define the Variance Gamma process by subordination of Brownian motion \( B \) by a \( \Gamma(\alpha, \theta) \) subordinator \( \tau \) with Lévy density \( g_{\tau}(y) = \alpha y^{-1}e^{-\theta y} \), then we obtain a Lévy density
\[
g(z) = \int_0^\infty f_t(z)g_{\tau}(t)dt = \int_0^\infty \frac{1}{\sqrt{2\pi t}}e^{-z^2/(2t)}e^{-\theta t}dt
\]
and we can calculate this integral to get
\[
g(z) = \alpha |z|^{-1}e^{-\sqrt{2\theta} |z|}
\]
as Lévy density. The Variance Gamma process as subordinated Brownian motion has an interesting interpretation when modelling financial price processes. In fact, the stock price is then considered to evolve according to the Black-Scholes model, but not in real time, but in operational time \( \tau y, y \geq 0 \). Time evolves as a Gamma process, with infinitely many jumps in any bounded interval.

Note that all Lévy processes that we can construct as subordinated Brownian motions are symmetric. However, not all symmetric Lévy processes are subordinated Brownian motions.

12.2 Ornstein-Uhlenbeck processes

Example 83 (Gamma-OU process) Let \( (N_t)_{t \geq 0} \) be a Poisson process with intensity \( a\lambda \) and jump times \( (T_k)_{k \geq 1} \), \( (X_n)_{n \geq 1} \) a sequence of independent \( \Gamma(1, b) \) random variables, \( Y_0 \sim \Gamma(a, b) \), consider the stochastic process
\[
Y_t = Y_0 e^{-\lambda t} + \sum_{k=1}^{N_t} X_n e^{-\lambda(t-T_k)}
\]
We use this model for the speed of the market and think of an initial speed of \( Y_0 \) which slows down exponentially, but at times of a Poisson process, events occur that make the speed jump up at times \( T_k \), \( k \geq 0 \). Each of these also slow down exponentially. In fact, there is a strong equilibrium in that
\[
\mathbb{E}(e^{-q Y_t}) = \mathbb{E}(e^{-q e^{-\lambda t}Y_0}) \mathbb{E}\left( \exp\left\{ -q \sum_{k=1}^{N_t} X_n e^{-\lambda(t-T_k)} \right\} \right)
\]
\[
= \left( \frac{b}{b + q e^{-\lambda t}} \right)^a \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} \left( \int_0^t \frac{1}{b + q e^{-\lambda s}} ds \right)^n
\]
\[
= \left( \frac{b}{b + q e^{-\lambda t}} \right)^a \left( \frac{b + q e^{-\lambda t}}{b + q} \right)^a = \left( \frac{b}{b + q} \right)^a,
\]
so $Y_t$ has the same distribution as $Y_0$. In fact, $(Y_t)_{t \geq 0}$ is a stationary Markov process. The process $Y$ is called a Gamma-OU process, since it has the Gamma distribution as its stationary distribution. The time change process

$$
\tau_y = \int_0^y Y_s ds
$$

associated with speed $Y$ is called integrated Ornstein-Uhlenbeck process. Note that the stationarity of $Y$ implies that $\tau$ has stationary increments, but note that $\tau$ does not have independent increments. The associated stochastic volatility model is now the time-changed Brownian motion $(B_{\tau_y})_{y \geq 0}$.

In general, we can define Ornstein-Uhlenbeck processes associated with any subordinator $Z$ or rather its Poisson point process $(\Delta Z_t)_{t \geq 0}$ of jumps as

$$
Y_t = Y_0 e^{-\lambda t} + \sum_{s \leq t} \Delta Z_s e^{-\lambda (t-s)},
$$

for any initial distribution for $Y_0$, but a stationary distribution can also be found.

We can always associate a stochastic volatility model $(B_{\tau_y})_{y \geq 0}$ using the integrated volatility $\tau_y = \int_y Y_s ds$ as time change. Note that, by the discussion of the last section, we can actually infer the time change from sums of squared increments for a small time lag $2^{-n}$, even though the actual time change is not observed. In practice, the so-called market microstructure (piecewise constant prices) destroys model fit for small times, so we need to choose a moderately small $2^{-n}$. In practice, 5 minutes is a good choice.

### 12.3 Simulation by subordination

Note that we can simulate subordinators using simulation Method 1 (Time discretisation) or Method 2 (Throwing away the small jumps). The latter consisted essentially in simulating the Poisson point process of jumps of the subordinator. Clearly, we can apply this method also to simulate an Ornstein-Uhlenbeck process.

**Method 3 (Subordination)** Let $(\tau_y)_{y \geq 0}$ be an increasing process that we can simulate, and let $(X_t)_{t \geq 0}$ be a Lévy process with cumulative distribution function $F_t$ of $X_t$. Fix a time lag $\delta > 0$. Then the process

$$
Z^{(3,\delta)}_y = S_{[y/\delta]}, \quad \text{where } S_n = \sum_{k=1}^n A_k \text{ and } A_k = F^{-1}_{\tau_k - \tau_{k-1}}(U_k)
$$

is the time discretisation of the subordinated process $Z_y = X_{\tau_y}$.

**Example 84** We can use Method 3 to simulate the Variance Gamma process, since we can simulate the Gamma process $\tau$ and we can simulate the $A_k$. Actually, we can use the Box-Muller method to generate standard Normal random variables $N_k$ and then use

$$
\tilde{A}_k \sim \sqrt{\tau_{k\delta} - \tau_{(k-1)\delta}} N_k, \quad k \geq 1,
$$

instead of $A_k$, $k \geq 1$. 

13.1 The strong Markov property

Recall that a stopping time is a random time $T \in [0, \infty]$ such that for every $s \geq 0$ the information $\mathcal{F}_s$ up to time $s$ allows to decide whether $T \leq s$. More formally, if the event $\{T \leq s\}$ can be expressed in terms of $(X_r, r \leq s)$ (is measurable with respect to $\mathcal{F}_s = \sigma(X_r, r \leq s)$). The prime example of a stopping time is the first entrance time $T_I = \inf\{t \geq 0 : X_t \in I\}$ into a set $I \subset \mathbb{R}$. Note that

$$\{T \leq s\} = \{\text{there is } r \leq s \text{ such that } X_r \in I\}$$

(for open sets $I$ we can drop the irrational $r \leq s$ to show measurability.).

We also denote the information $\mathcal{F}_T$ up to time $T$. More formally,

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq s\} \in \mathcal{F}_s \text{ for all } s \geq 0\},$$

i.e. $\mathcal{F}_T$ contains those events that, if $T \leq s$, can be expressed in terms of $(X_r, r \leq s)$, for all $s \geq 0$.

Recall the simple Markov property which we can now state as follows. For a Lévy process $(X_t)_{t \geq 0}$ and a fixed time $t$, the post-$t$ process $(X_{t+s} - X_t)_{s \geq 0}$ has the same distribution as $X$ and is independent of the pre-$t$ information $\mathcal{F}_t$.

**Theorem 85 (Strong Markov property)** Let $(X_t)_{t \geq 0}$ be a Lévy process and $T$ a stopping time. Then given $T < \infty$, the post-$T$ process $(X_{T+s} - X_T)_{s \geq 0}$ has the same distribution as $X$ and is independent of the pre-$T$ information $\mathcal{F}_T$.

**Proof:** Let $0 < s_1 < \ldots < s_m$, $C_1, \ldots, C_m \subset \mathbb{R}$ closed, $A \in \mathcal{F}_T$. Then we need to show that

$$\mathbb{P}(A, T < \infty, X_{T+s_1} - X_T \leq c_1, \ldots, X_{T+s_m} - X_T \leq c_m)$$

$$= \mathbb{P}(A, T < \infty) \mathbb{P}(X_{s_1} \leq c_1, \ldots, X_{s_m} \leq c_m).$$

59
First define stopping times $T_n = 2^{-n}([2^n T] + 1)$, $n \geq 1$, that only take countably many values. These are the next dyadic rationals after time $T$. Note that $T_n \downarrow T$ as $n \to \infty$. Now note that $A \cap \{T_n = k2^{-n}\} \in \mathcal{F}_{k2^{-n}}$ and the simple Markov property yields

$$
\mathbb{P}(A, T_n < \infty, X_{T_n+s_1} - X_{T_n} \leq c_1, \ldots, X_{T_n+s_m} - X_{T_n} \leq c_m)
= \sum_{k=0}^{\infty} \mathbb{P}(A, T_n = k2^{-n}, X_{k2^{-n+s_1}} - X_{k2^{-n}} \leq c_1, \ldots, X_{k2^{-n+s_m}} - X_{k2^{-n}} \leq c_m)
= \sum_{k=0}^{\infty} \mathbb{P}(A, T_n = k2^{-n}) \mathbb{P}(X_{s_1} \leq c_1, \ldots, X_{s_m} \leq c_m)
= \mathbb{P}(A, T < \infty) \mathbb{P}(X_{s_1} \leq c_1, \ldots, X_{s_m} \leq c_m).
$$

Now the right-continuity of sample paths ensures $X_{T_n+s_j} \to X_{T+s_j}$ as $n \to \infty$ and we conclude

$$
\mathbb{P}(A, T < \infty, X_{T+s_1} - X_T \leq c_1, \ldots, X_{T+s_m} - X_T \leq c_m)
= \lim_{n \to \infty} \mathbb{P}(A, T_n < \infty, X_{T_n+s_1} - X_{T_n} \leq c_1, \ldots, X_{T_n+s_m} - X_{T_n} \leq c_m)
= \lim_{n \to \infty} \mathbb{P}(A, T_n < \infty) \mathbb{P}(X_{s_1} \leq c_1, \ldots, X_{s_m} \leq c_m)
= \mathbb{P}(A, T < \infty) \mathbb{P}(X_{s_1} \leq c_1, \ldots, X_{s_m} \leq c_m),
$$

for all $(c_1, \ldots, c_m)$ such that $\mathbb{P}(X_{T+s_j} - X_T = c_j) = 0$, $j = 1, \ldots, m$. Finally note that $(X_{T+s} - X_T)_{s \geq 0}$ clearly has right-continuous paths with left limits. 

\[\square\]

### 13.2 The supremum process

Let $X = (X_t)_{t \geq 0}$ be a Lévy process. We denote its supremum process by

$$
\overline{X}_t = \sup_{0 \leq s \leq t} X_s, \quad t \geq 0.
$$

We are interested in the joint distribution of $(X_t, \overline{X}_t)$, e.g. for the payoff of a barrier or lookback option. Moment generating functions are easier to calculate and can be numerically inverted. We can also take such a transform over the time variable, e.g.

$$
q \mapsto \int_0^\infty e^{-qt} \mathbb{E}(e^{\gamma X_t}) dt = \frac{1}{q - \Psi(\gamma)} \text{ uniquely identifies } \mathbb{E}(e^{\gamma X_t}),
$$

and the distribution of $X_t$. But $q \int_0^\infty e^{-qt} \mathbb{E}(e^{\gamma X_t}) dt = \mathbb{E}(e^{\gamma X_{\tau}})$ for $\tau \sim \text{Exp}(q)$.

**Proposition 86 (Independence)** Let $X$ be a Lévy process, $\tau \sim \text{Exp}(q)$ an independent random time. Then $\overline{X}_\tau$ is independent of $\overline{X}_\tau - X_\tau$.

**Proof:** We only prove the case where $G_1 = \inf\{t > 0 : X_t > 0\}$ satisfies $\mathbb{P}(G_1 > 0) = 1$. In this case we can define successive record times $G_n = \inf\{t > G_{n-1} : X_t > \overline{X}_{G_{n-1}}\}$, $n \geq 2$, and also set $G_0 = 0$. Note that, by the strong Markov property at the stopping times $G_n$ we have that $X_{G_n} > \overline{X}_{G_{n-1}}$ (otherwise the post-$G_{n-1}$ process $\tilde{X}_t = X_{G_{n-1}+t} - X_{G_{n-1}}$ would have the property $\tilde{G}_1 = 0$, but the strong Markov property yields $\mathbb{P}(\tilde{G}_1 > 0) =$...
\[ \mathbb{P}(G_1 > 0) = 1. \] So \( X \) can only reach new records by upward jumps, \( X_\tau \in \{X_{G_n}, n \geq 0\} \) and more specifically, we will have \( X_\tau = G_n \) if and only if \( G_n \leq \tau < G_{n+1} \) so that

\[
\mathbb{E}(e^{\beta X_\tau + \gamma (X_\tau - X_\tau)}) = \int_0^\infty q e^{-qt} \mathbb{E}(e^{\beta X_\tau + \gamma (X_\tau - X_\tau)}) dt
\]

\[
= q \mathbb{E} \left( \sum_{n=0}^{\infty} \int_{G_n}^{G_{n+1}} e^{-qt} e^{\beta X_\tau + \gamma (X_\tau - X_\tau)} dt \right)
\]

\[
= q \sum_{n=0}^{\infty} \mathbb{E} \left( e^{\beta X_{G_n}} e^{-qG_n} \int_{G_n}^{G_{n+1}} e^{-qs} e^{-\gamma (X_{G_n+z} - X_{G_n})} ds \right)
\]

\[
= q \sum_{n=0}^{\infty} \mathbb{E} \left( e^{-qG_n + \beta X_{G_n}} \right) \mathbb{E} \left( \int_0^{G_1} e^{-qs - \gamma X_s} ds \right)
\]

where we applied the strong Markov property at \( G_n \) to split the expectation in the last row – note that \( \int_{G_n}^{G_{n+1}} e^{-qs - \gamma (X_{G_n+z} - X_{G_n})} ds \) is a function of the post-\( G_n \) process, whereas \( e^{-qG_n + \beta X_{G_n}} \) is a function of the pre-\( G_n \) process, and the expectation of the product of independent random variables is the product of their expectations.

This completes the proof, since the last row is a product of a function of \( \beta \) and a function of \( \gamma \), which is enough to conclude. More explicitly, we can put \( \beta = 0, \gamma = 0 \) and \( \beta = \gamma = 0 \), respectively, to see that indeed the required identity holds:

\[
\mathbb{E}(e^{\beta X_\tau + \gamma (X_\tau - X_\tau)}) = \mathbb{E}(e^{\beta X_\tau}) \mathbb{E}(e^{\gamma (X_\tau - X_\tau)}).
\]

\[ \square \]

### 13.3 Lévy processes with no positive jumps

Consider stopping times \( T_x = \inf\{t \geq 0 : X_t \in (x, \infty)\} \), so-called first passage times. For Lévy processes with no positive jumps, we must have \( X_{T_x} = x \), provided that \( T_x < \infty \). This observation allows to calculate the moment generating function of \( T_x \). To prepare this result, recall that the distribution of \( X_t \) has moment generating function

\[
\mathbb{E}(e^{\gamma X_t}) = e^{\psi(\gamma)}, \quad \psi(\gamma) = a_1 \gamma + \frac{1}{2} \sigma^2 \gamma^2 + \int_{-\infty}^{0} (e^{\gamma x} - 1 - \gamma x 1_{|x| \leq 1}) g(x) dx.
\]

Let us exclude the case where \( -X \) is a subordinator, i.e. where \( \sigma^2 = 0 \) and \( a_1 - \int_{-1}^{0} x g(x) dx \leq 0 \), since in that case \( T_x = \infty \). Then note that

\[
\psi''(\gamma) = \sigma^2 + \int_{-\infty}^{0} x^2 e^{\gamma x} g(x) dx > 0,
\]

so that \( \psi \) is convex and hence has at most two zeros, one of which is \( \psi(0) = 0 \). There is a second zero \( \gamma_0 > 0 \) if and only if \( \psi'(0) = \mathbb{E}(X_1) < 0 \), since we excluded the case where \( -X \) is a subordinator, and \( \mathbb{P}(X_t > 0) > 0 \) implies that \( \psi(\infty) = \infty \).

**Theorem 87 (Level passage)** Let \( (X_t)_{t \geq 0} \) be a Lévy process with no positive jumps and \( T_x \) the first passage time across level \( x \). Then

\[
\mathbb{E}(e^{-qT_x 1_{\{T_x < \infty\}}}) = e^{-x \Phi(q)},
\]

where \( \Phi(q) \) is the largest \( \gamma \) for which \( \psi(\gamma) = q \).
Proof: We only prove this for the case where \( P(T_x < \infty) = 1 \), i.e. \( E(X_1) \geq 0 \). By Exercise A.2.3.(a) the processes \( M_t = e^{\tau X_t - t\Phi(\gamma)} \) are martingales. We will apply the Optional stopping theorem to \( T_x \). Note that \( E(M^2_t) = e^{t(\Psi(2\gamma) - 2\Psi(\gamma))} \) is not such that \( \text{sup}_{t \geq 0} E(M^2_t) < \infty \). However, if we put

\[
M_t^{(u)} = M_t \quad \text{if} \quad t \leq u \quad \text{and} \quad M_t^{(u)} = M_u \quad \text{if} \quad t \geq u,
\]
then \( (M^u_t)_{t \geq 0} \) is a martingale which satisfies \( \text{sup}_{t \geq 0} E((M^u_t)^2) < \infty \). Also, \( T_x \land u \) is a stopping time, so that for \( \gamma \geq \gamma_0 = 0 \) (so that \( \Psi(\gamma) \geq 0 \))

\[
1 = E(M_{T_x \land u}^{(u)}) = E(M_{T_x}) = E(e^{\gamma x - \Psi(\gamma) T_x}), \quad \text{as} \quad u \to \infty,
\]
by dominated convergence \( (M_{T_x \land u} \leq \exp\{\gamma x - \Psi(\gamma)T_x\} \leq \exp\{\gamma x\}) \). We now conclude that

\[
E(e^{-\Psi(\gamma) T_x}) = e^{-\gamma x}
\]
which for \( q = \Psi(\gamma) \) and \( \Phi(q) \) the unique \( \gamma \geq \gamma_0 = 0 \) with \( \Psi(\gamma) = q \). \( \square \)

Corollary 88 Let \( X \) be a Lévy process with no positive jumps and \( \tau \sim \text{Exp}(q) \) independent. Then \( \overline{X}_\tau \sim \text{Exp}(\Phi(q)) \).

Proof: \( F(\overline{X}_\tau > x) = P(T_x \leq \tau) = \int_0^\infty P(\tau \geq t)f_{T_x}(t)dt = E(e^{-qT_x}) = e^{-\Phi(q)x} \). \( \square \)

If we combine this with the Independence Theorem of the previous section we obtain.

Corollary 89 Let \( X \) be a Lévy process with no positive jumps and \( \tau \sim \text{Exp}(q) \) independent. Then

\[
E(e^{-\beta \overline{X}_\tau - X_\tau}) = \frac{q(\Phi(q) - \beta)}{\Phi(q)(\Phi(q) - \beta)}
\]

Proof: Note that we have from the Independence Theorem that

\[
E(e^{\beta \overline{X}_\tau})E(e^{-\beta \overline{X}_\tau - X_\tau}) = E(e^{\beta X_\tau}) = \int_0^\infty q e^{-qt}E(e^{\beta X_1})dt = \frac{q}{q - \Phi(\beta)}
\]
and from the preceding corollary

\[
E(e^{\beta \overline{X}_\tau}) = \frac{\Phi(q)}{\Phi(q) - \beta} \quad \text{and so} \quad E(e^{-\beta \overline{X}_\tau - X_\tau}) = \frac{q}{q - \Phi(\beta)} \frac{\Phi(q) - \beta}{\Phi(q)}.
\]

\( \square \)

13.4 Application: insurance ruin

Proposition 86 splits the Lévy process at its supremum into two increments. If you turn the picture of a Lévy process by 180°, this split occurs at the infimum, and it can be shown (Exercise A.7.1) that \( \overline{X}_\tau \sim X_\tau - \overline{X}_\tau \). Therefore, Corollary 89 gives \( E(e^{\beta \overline{X}_\tau}) \), also for \( q < 0 \) if \( E(X_1) > 0 \), since then

\[
E(e^{\beta \overline{X}_\tau}) = \lim_{q \downarrow 0} \frac{q(\Phi(q) - \beta)}{\Phi(q)(\Phi(q) - \beta)} = \frac{\beta E(X_1)}{\Psi(\beta)}
\]
since \( \Phi'(0) = 1/\Psi'(0) = 1/\text{E}(X_1) \) and note that for an insurance reserve process \( R_t = u + X_t \), the probability of ruin is \( r(u) = P(\overline{X}_\infty < -u) \), the distribution function of \( \overline{X}_\infty \) which is uniquely identified by \( E(e^{\beta \overline{X}_\tau}) \).
Lecture 14

Ladder times and storage models

Reading: Kyprianou Sections 1.3.2 and 3.3

14.1 Case 1: No positive jumps

In Theorem 87 we derived the moment generating function of $T_x = \inf\{t \geq 0 : X_t > x\}$ for any Lévy process with no positive jumps. We also indicated the complication that $T_x = \infty$ is a possibility, in general. Let us study this in more detail in our standard setting

$$
\mathbb{E}(e^{\gamma X_t}) = e^{\phi(\gamma)}, \quad \Psi(\gamma) = a_1 \gamma + \frac{1}{2} \sigma^2 \gamma^2 + \int_{-\infty}^{0} (e^{\gamma x} - 1 - \gamma x 1_{|x| \leq 1}) g(x) dx.
$$

The important quantity is

$$
\mathbb{E}(X_1) = \frac{\partial}{\partial \gamma} \mathbb{E}(e^{\gamma X_t}) \bigg|_{\gamma=0} = \Psi'(0) = a_1 - \int_{-1}^{0} x g(x) dx.
$$

The formula that we derived was

$$
\mathbb{E}(e^{-\Phi(T_x)} 1_{\{T_x < \infty\}}) = e^{-x \Phi(q)}
$$

where for $q > 0$, $\Phi(q) > 0$ is unique with $\Psi(\Phi(q)) = q$. Letting $q \downarrow 0$, we see that

$$
\mathbb{P}(T_x < \infty) = \lim_{q \downarrow 0} \mathbb{E}(e^{-q T_x} 1_{\{T_x < \infty\}}) = e^{-x \Phi(0+)}.
$$

Here the convexity of $\Psi$ that we derived last time implies that $\Phi(0+) = 0$ if and only if $\mathbb{E}(X_1) = \phi'(0) > 0$. Therefore, $\mathbb{P}(T_x < \infty) = 1$ if and only if $\mathbb{E}(X_1) \geq 0$.

Part of this could also be deduced by the strong (or weak) law of large numbers. Applied to increments $Y_k = X_{k\delta} - X_{(k-1)\delta}$ it implies that

$$
\frac{X_{n\delta}}{n\delta} = \frac{1}{n} \sum_{k=1}^{n} Y_k \to \frac{1}{\delta} \mathbb{E}(Y_1) = \frac{1}{\delta} \mathbb{E}(X_\delta) = \mathbb{E}(X_1),
$$

almost surely (or in probability) as $n \to \infty$. We can slightly improve this result to a convergence as $t \to \infty$ as follows

$$
\mathbb{E}(e^{\gamma X_t/t}) = e^{t \phi(\gamma/t)} \to e^{\gamma \phi'(0)} = e^{\gamma \mathbb{E}(X_1)} \Rightarrow \frac{X_t}{t} \to \mathbb{E}(X_1),
$$
in probability. We used here that \( Z_t \to a \) in distribution implies \( Z_t \to a \) in probability, which holds only because \( a \) is a constant, not a random variable. Note that indeed for all \( \varepsilon > 0 \), as \( t \to \infty \),

\[
\mathbb{P}(|Z_t - a| > \varepsilon) \leq \mathbb{P}(Z_t \leq a - \varepsilon) + 1 - \mathbb{P}(Z_t \leq a + \varepsilon) \to 0 + 1 - 1 = 0.
\]

From this, we easily deduce that \( X_t \to \pm \infty \) (in probability) if \( \mathbb{E}(X_t) \neq 0 \), but the case \( \mathbb{E}(X_t) = 0 \) is not so clear from this method. In fact, it can be shown that all convergences hold in the almost sure sense, here.

By an application of the Strong Markov property we can show the following.

**Proposition 90** The process \((T_x)_{x \geq 0}\) is a subordinator.

*Proof:* Let us here just prove that \( T_{x+y} - T_x \) is independent of \( T_x \) and has the same distribution as \( T_y \). The remainder is left as an exercise.

Note first that \( X_{T_x} = x \), since there are no positive jumps. The Strong Markov property at \( T_x \) can therefore be stated as \( \tilde{X} = (X_{T_x+s} - x)_{s \geq 0} \) is independent of \( \mathcal{F}_{T_x} \) and has the same distribution as \( X \). Now note that

\[
T_x + \tilde{T}_y = T_x + \inf\{s \geq 0 : \tilde{X}_s > y\} = T_x + \inf\{s \geq 0 : X_{T_x+s} > x + y\}
\]

so that \( T_{x+y} - T_x = \tilde{T}_y \), and we obtain

\[
\mathbb{P}(T_x \leq s, T_{x+y} - T_x \leq t) = \mathbb{P}(T_x \leq s, \tilde{T}_y \leq t) = \mathbb{P}(T_x \leq s)\mathbb{P}(\tilde{T}_y \leq t),
\]

since \( \{T_x \leq s\} \in \mathcal{F}_{T_x} \). Formally, \( \{T_x \leq s\} \cap \{T_x \leq r\} = \{T_x \leq s \wedge r\} \in \mathcal{F}_r \) for all \( r \geq 0 \) since \( T_x \) is a stopping time.

We can understand what the jumps of \((T_x)_{x \geq 0}\) are: in fact, \( X \) can be split into its supremum process \( \overline{X}_t = \sup_{0 \leq s \leq t} X_s \) and the bits of path below the supremum. Roughly, the times

\[
\{T_x, x \geq 0\} = \{t \geq 0 : X_t = \overline{X}_t\}
\]

are the times when the supremum increases. \( T_x - T_{x-} > 0 \) if the supremum process remains constant at height \( x \) for an amount of time \( T_x - T_{x-} \). The process \((T_x)_{x \geq 0}\) is called “ladder time process”. The process \((X_{T_x})_{x \geq 0}\) is called ladder height process. In this case, \( X_{T_x} = x \) is not very illuminating. Note that \((T_x, X_{T_x})_{x \geq 0}\) is a bivariate Lévy process.

**Example 91 (Storage models)** Consider a Lévy process of bounded variation, represented as \( A_t - B_t \) for two subordinators \( A \) and \( B \). We interpret \( A_t \) as the amount of work arriving in \([0,t]\) and \( B_t \) as the amount of work that can potentially exit from the system. Let us focus on the case where \( A \) is a compound Poisson process and \( B_t = t \) for a continuously working processor. The quantity of interest is the amount \( W_t \) of work waiting to be carried out and requiring storage, where \( W_0 = w \geq 0 \) is an initial amount of work stored.
Note that \( W_t \neq w + A_t - B_t \), in general, since \( w + A_t - B_t \) can become negative, whereas \( W_t \geq 0 \). In fact, we can describe as follows: if the storage is empty, then no work exits from the system. Can we see from \( A_t - B_t \) when the storage will be empty? We can express the first time it becomes empty and the first time it is refilled thereafter as

\[
L_1 = \inf \{ t \geq 0 : w + A_t - B_t = 0 \} \quad \text{and} \quad R_1 = \inf \{ t \geq L_1 : \Delta A_t > 0 \}.
\]

On \([L_1, R_1]\), \( X_t = B_t - A_t \) increases linearly at unit rate from \( w \) to \( w + (R_1 - L_1) \), whereas \( W \) remains constant equal to zero. In fact,

\[
W_t = w - X_t + \int_{L_1 \wedge t}^t 1ds = w - X_t + \int_0^t 1_{\{x=\overline{X}_s \geq w\}}ds = (w \vee \overline{X}_t) - X_t, \quad 0 \leq t \leq R_1.
\]

An induction now shows that the system is idle if and only if \( X_t = \overline{X}_t \geq w \), so that \( W_t = X_t + (w \vee \overline{X}_t) \) for all \( t \geq 0 \).

In this context, \((\overline{X}_t - w)^+\) is the amount of time the system was idle before time \( t \), and \( T_x = \inf \{ t \geq 0 : X_t > x \} \) is the time by which the system has accumulated time \( x - w \) in the idle state, \( x \geq w \), and we see that \((x - w)/T_x \sim x/T_x \to 1/\mathbb{E}(T_1) = 1/\Phi'(0) = \phi'(0) = \mathbb{E}(X_1) = 1 - \mathbb{E}(A_1)\) in probability, if \( \mathbb{E}(A_1) \leq 1 \).

**Example 92 (Dams)** Suppose that the storage model refers more particularly to a dam that releases a steady stream of water at a constant intensity \( a_2 \). Water is added according to a subordinator \((A_t)_{t \geq 0}\). The dam will have a maximal capacity of, say, \( M > 0 \). Given an initial water level of \( w \geq 0 \), the water level at time \( t \) is, as before

\[
W_t = (w \vee \overline{X}_t) - X_t, \quad \text{where} \quad X_t = a_2 t - A_t.
\]

The time \( F = \inf \{ t \geq 0 : W_t > M \} \), the first time when the dam overflows, is a quantity of interest. We do not pursue this any further theoretically, but note that this can be simulated, since we can simulate \( X \) and hence \( W \).

### 14.2 Case 2: Union of intervals as ladder time set

**Proposition 93** If \( X_t = a_2 t - C_t \) for a compound Poisson process \((C_t)_{t \geq 0}\) and a drift \( a_2 \geq 0 \vee \mathbb{E}(C_t) \), then the ladder time set is a collection of intervals. More precisely, \( \{ t \geq 0 : X_t = \overline{X}_t \} \) is the range \( \{ \sigma_y, y \geq 0 \} \) of a compound Poisson subordinator with positive drift coefficient.

**Proof:** Define \( L_0 = 0 \) and then for \( n \geq 0 \) stopping times

\[
R_n = \inf \{ t \geq L_n : \Delta C_t > 0 \}, \quad L_{n+1} = \inf \{ t \geq R_n : X_t = \overline{X}_t \}.
\]

The strong Markov property at these stopping times show that \((R_n - L_n)_{n \geq 0}\) is a sequence of \( \text{Exp}(\lambda) \) random variables where \( \lambda = \int_0^\infty g(x)dx \) is the intensity of positive jumps, and \((L_n - R_{n-1})_{n \geq 1}\) is a sequence of independent identically distributed random variables. Now define \( T_n = R_0 - L_0 + \ldots + R_{n-1} - L_{n-1} \) and \((\sigma_y)_{y \geq 0}\) to be the compound Poisson process with unit drift, jump times \( T_n, n \geq 1 \), and jump heights \( L_n - R_{n-1}, n \geq 1 \). □
The ladder height process \((X_{\sigma_y})_{y \geq 0}\) will then also have a positive drift coefficient and share some jump times with \(\sigma\) (whenever \(X\) jumps from below \(X_{t-}\) above \(X_{t-}\), but have extra jump times when \(X\) jumps from \(X_{t-}\) upwards, some jump times of \((\sigma_y)_{y \geq 0}\) are not jump times of \((X_{\sigma_y})_{y \geq 0}\) – if \(X\) reaches \(X_{t-}\) again without a jump.)

\[W_t = w - X_t + \int_0^t 1_{\{X_s = \overline{X}_s \geq w\}} ds \quad t \geq 0,\]

but note that the latter integral cannot be expressed in terms of \(\overline{X}_t\) so easily, but \(\sigma_y\) is still the time by which the system has accumulated time \(y - w\) in the idle state, for \(y \geq w\), so \(\overline{I}_t = \inf\{y \geq 0 : \sigma_y > t\}\) is the amount of idle time before \(t\).

### 14.3 Case 3: Discrete ladder time set

If \(X_t = a_2 t - C_t\) for a compound Poisson process (or indeed bounded variation pure jump process) \((C_t)_{t \geq 0}\) and a drift \(a_2 < 0\), then the ladder time set is discrete. We can still think of \(\{t \geq 0 : X_t = \overline{X}_t\}\) as the range \(\sigma_y, y \geq 0\) of a compound Poisson subordinator with zero drift coefficient. More naturally, we would define successive ladder times \(G_0 = 0\) and \(G_{n+1} = \inf\{t > G_n : X_t = \overline{X}_t\}\). By the strong Markov property, \(G_{n+1} - G_n, n \geq 0\), is a sequence of independent and identically distributed random variables, and for any intensity \(\lambda > 0\), we can specify \((\sigma_y)_{y \geq 0}\) to be a compound Poisson process with rate \(\lambda > 0\) and jump sizes \(G_{n+1} - G_n, n \geq 0\).

Note that \((\sigma_y)_{y \geq 0}\) is not unique since we have to choose \(\lambda\). In fact, once a choice has been made and \(q > 0\), we have \(\{\sigma_y : y \geq 0\} = \{\sigma_{qy}, y \geq 0\}\), not just here, but also in Cases 1 and 2. In Cases 1 and 2, however, we identified a natural choice \((q)\) in each case.

### 14.4 Case 4: non-discrete ladder time set and positive jumps

The general case is much harder. It turns out that we can still express

\[\{t \geq 0 : X_t = \overline{X}_t\} = \{\sigma_y : y \geq 0\}\]

for a subordinator \((\sigma_y)_{y \geq 0}\), but, as in Case 3, there is no natural way to choose this process. It can be shown that the bivariate process \((\sigma_y, X_{\sigma_y})_{y \geq 0}\) is a bivariate subordinator in this general setting, called the ladder process. There are descriptions of its distribution and relations between these processes of increasing ladder events and analogous processes of decreasing ladder events.
15.1 Galton-Watson processes

Let \( \xi = (\xi_k)_{k \geq 0} \) be (the probability mass function of) an offspring distribution. Consider a population model where each individual gives birth to independent and identically distributed numbers of children, starting from \( Z_0 = 1 \) individual, the common ancestor. Then the \((n+1)\)st generation \( Z_{n+1} \) consists of the sum of numbers of children \( N_{n,1}, \ldots, N_{n,Z_n} \) of the \( n \)th generation:

\[
Z_{n+1} = \sum_{i=1}^{Z_n} N_{n,i}, \quad \text{where } N_{n,i} \sim \xi \text{ independent, } i \geq 1, n \geq 0.
\]

**Proposition 95** Let \( \xi \) be an offspring distribution, \( g(s) = \sum_{k \geq 0} \xi_k s^k \) its generating function, then

\[
E(s^{Z_1}) = g(s), \ E(s^{Z_2}) = g(g(s)), \ldots, E(s^{Z_n}) = g^{(n)}(s),
\]

where \( g^{(0)}(s) = s \), \( g^{(n+1)}(s) = g^{(n)}(g(s)), n \geq 0 \).

**Proof:** The result is clearly true for \( n = 0 \) and \( n = 1 \). Now note that

\[
E(s^{X_{n+1}}) = E\left(s^{\sum_{i=1}^{Z_n} N_{n,i}}\right) = \sum_{j=0}^{\infty} P(Z_n = j) E\left(s^{\sum_{i=1}^{j} N_{n,i}}\right)
\]

\[
= \sum_{j=0}^{\infty} P(Z_n = j)(g(s))^j = E((g(s))^{Z_n}) = g^{(n)}(g(s)).
\]

**Proposition 96** \((Z_n)_{n \geq 0}\) is a Markov chain whose transition probabilities are given by

\[
p_{ij} = P(N_1 + \ldots + N_i = j), \quad \text{where } N_1, \ldots, N_i \sim \xi \text{ independent}.
\]

In particular, if \((Z^{(1)}_n)_{n \geq 0}\) and \((Z^{(2)}_n)_{n \geq 0}\) are two independent Markov chains with transition probabilities \((p_{ij})_{i,j \geq 0}\) starting from population sizes \( k \) and \( l \), respectively, then \( Z^{(1)}_n + Z^{(2)}_n, n \geq 0 \), is also a Markov chain with transition probabilities \((p_{ij})_{i,j \geq 0}\) starting from \( k + l \).
Proof: Just note that the independence of \((N_{n,i})_{i \geq 1}\) and \((N_{k,i})_{0 \leq k \leq n-1, i \geq 1}\) implies that

\[
\mathbb{P}(Z_{n+1} = j | Z_0 = i_0, \ldots, Z_{n-1} = i_{n-1}, Z_n = i_n) = \mathbb{P}(N_{n,1} + \ldots, N_{n,i_n} = j | Z_0 = i_0, \ldots, Z_n = i_n)
\]

\[
= \mathbb{P}(N_{n,1} + \ldots, N_{n,i_n} = j) = p_{i_n,j},
\]

as required. For the second assertion note that

\[
b_{(i_1,i_2),j} := \mathbb{P}(Z_{n+1}^{(1)} + Z_{n+1}^{(2)} = j | Z_n^{(1)} = i_1, Z_n^{(2)} = i_2)
\]

\[
= \mathbb{P}(N_{n,1}^{(1)} + \ldots + N_{n,i_1}^{(1)} + N_{n,1}^{(2)} + \ldots + N_{n,i_2}^{(2)} = j) = p_{i_1+i_2,j}
\]

only depends on \(i_1 + i_2\) (not \(i_1\) or \(i_2\) separately) and is of the form required to conclude that

\[
\mathbb{P}(Z_{n+1}^{(1)} + Z_{n+1}^{(2)} = j | Z_n^{(1)} + Z_n^{(2)} = i) = \frac{\sum_{i_1=0}^{i} \mathbb{P}(Z_n^{(1)} = i_1, Z_n^{(2)} = i - i_1)b_{(i_1,i-i_1),j}}{\mathbb{P}(Z_n^{(1)} + Z_n^{(2)} = i)} = p_{ij}.
\]

The second part of the proposition is called the branching property and expresses the property that the families of individuals in the same generation evolve completely independently of one another.

15.2 Continuous-time Galton-Watson processes

We can also model lifetimes of individuals by independent exponentially distributed random variables with parameter \(\lambda > 0\). We assume that births happen at the end of a lifetime. This breaks the generations. Since continuously distributed random variables are almost surely distinct, we will observe one death at a time, each leading to a jump of size \(k - 1\) with probability \(\xi_k, k \geq 0\). It is customary to only consider offspring distributions with \(\xi_1 = 0\), so that there is indeed a jump at every death time. Note that at any given time, if \(j\) individuals are present in the population, the next death occurs at a time

\[
H = \min\{L_1, \ldots, L_j\} \sim \text{Exp}(j\lambda), \quad \text{where } L_1, \ldots, L_j \sim \text{Exp}(\lambda).
\]

From these observations, one can construct (and simulate!) the associated population size process \((Y_t)_{t \geq 0}\) by induction on the jump times.

**Proposition 97** \((Y_t)_{t \geq 0}\) is a Markov process. If \(Y^{(1)}\) and \(Y^{(2)}\) are independent Markov processes with these transition probabilities starting from \(k\) and \(l\), then \(Y^{(1)} + Y^{(2)}\) is also a Markov process with the same transition probabilities starting from \(k + l\).

**Proof:** Based on BS3a Applied Probability, the proof is not difficult. We skip it here. \(\square\)

\((Y_t)_{t \geq 0}\) is called a continuous-time Galton-Watson process. In fact, these are the only Markov processes with the branching property (i.e. satisfying the second statement of the proposition for all \(k \geq 1, l \geq 1\)).
Example 98 (Simple birth-and-death processes) If individuals have lifetimes with parameter $\mu$ and give birth at rate $\beta$ to single offspring repeatedly during their lifetime, then we recover the case

$$\lambda = \mu + \beta \quad \text{and} \quad \xi_0 = \frac{\mu}{\mu + \beta}, \quad \xi_2 = \frac{\beta}{\mu + \beta}.$$ 

In fact, we have to reinterpret this model by saying each transition is a death, giving birth to either two or no offspring. These parameters arise since, if only one individual is present, the time to the next transition is the minimum of the exponential birth time and the exponential death time.

The fact that all jump sizes are independent and identically distributed is reminiscent of compound Poisson processes, but for high population sizes $j$ we have high parameters to the exponential times between two jumps – the process $Y$ moves faster than a compound Poisson process at rate $\lambda$. Note however that for $H \sim \text{Exp}(j\lambda)$ we have $jH \sim \text{Exp}(\lambda)$.

Let us use this observation to specify a time-change to slow down $Y$.

**Proposition 99** Let $(Y_t)_{t \geq 0}$ be a continuous-time Galton-Watson process with offspring distribution $\xi$ and lifetime distribution $\text{Exp}(\lambda)$. Then for the piecewise linear functions

$$J_t = \int_0^t Y_u du, \quad t \geq 0, \quad \varphi_s = \inf\{t \geq 0 : J_t > s\}, \quad 0 \leq s < J_\infty,$$

the process

$$X_s = Y_{\varphi_s}, \quad 0 \leq s < J_\infty,$$

is a compound Poisson process with jump distribution $(\xi_{k+1})_{k \geq 1}$ and rate $\lambda$, run until the first hitting time of 0.

**Proof:** Given $Y_0 = i$, the first jump time $T_1 = \inf\{t \geq 0 : Y_t \neq i\} \sim \text{Exp}(i\lambda)$, so

$$J_{T_1} = iT_1 \quad \text{and} \quad \varphi_s = s/i, \quad 0 \leq s \leq iT_1,$$

so we identify the first jump of $X_s = Y_{s/i}, 0 \leq s \leq iT_1$ at time $iT_1 \sim \text{Exp}(\lambda)$.

Now the strong Markov property (or the lack of memory property of all other lifetimes) implies that given $k$ offspring are produced at time $T_1$, the process $(Y_{T_1+i})_{t \geq 0}$ is a continuous-time Galton-Watson process starting from $j = i + k - 1$, independent of $(Y_r)_{0 \leq r \leq T_1}$. We repeat the above argument to see that $T_2 - T_1 \sim \text{Exp}(j\lambda)$, and for $j \geq 1$,

$$J_{T_2} = iT_1 + j(T_2 - T_1) \quad \text{and} \quad \varphi_{iT_1+s} = T_1/s + j, \quad 0 \leq s \leq j(T_2 - T_1),$$

and the second jump of $X_{iT_1+s} = Y_{iT_1+s/j}, 0 \leq s \leq j(T_2 - T_1)$, happens at time $iT_1 + j(T_2 - T_1)$, where $j(T_2 - T_1) \sim \text{Exp}(\lambda)$ is independent of $iT_1$. An induction as long as $Y_{T_m} > 0$ shows that $X$ is a compound Poisson process run until the first hitting time of 0. \qed
Corollary 100 Let $(X_s)_{s \geq 0}$ be a compound Poisson process starting from $l \geq 1$ with jump distribution $(\xi_{k-1})_{k\geq 1}$ and jump rate $\lambda > 0$. Then for the piecewise linear functions
\[ \varphi_s = \int_0^s \frac{1}{X_v} dv, \quad 0 \leq s < T_{\{0\}} , \quad \text{and} \quad J_t = \inf\{s \geq 0 : \varphi_s > t\}, \quad t \geq 0, \]
the process
\[ Y_t = X_{J_t}, \quad t \geq 0, \]
is a continuous-time Galton-Watson process with offspring distribution $\xi$ and lifetime distribution $\text{Exp}(\lambda)$.

15.3 Continuous-state branching processes

Population-size processes with state space $\mathbb{N}$ are natural, but for large populations, it is often convenient to use continuous approximations and use a state space $[0, \infty)$. In view of Corollary 100 it is convenient to define as follows.

Definition 101 (Continuous-state branching process) Let $(X_s)_{s \geq 0}$ be a Lévy process with no negative jumps starting from $x > 0$, with $\mathbb{E}(\exp\{-\gamma X_s\}) = \exp\{s\phi(\gamma)\}$. Then for the functions
\[ \varphi_s = \int_0^s \frac{1}{X_v} dv, \quad 0 \leq s < T_{\{0\}} , \quad \text{and} \quad J_t = \inf\{s \geq 0 : \varphi_s > t\}, \quad t \geq 0, \]
the process
\[ Y_t = X_{J_t}, \quad t \geq 0, \]
is called a continuous-state branching process with branching mechanism $\phi$.

We interpret upward jumps as birth events and continuous downward movement as (infinitesimal) deaths. The behaviour is accelerated at high population sizes, so fluctuations will be larger. The behaviour is slowed down at small population sizes, so fluctuations will be smaller.

Example 102 (Pure death process) For $X_s = x - cs$ we obtain
\[ \varphi_s = \int_0^s \frac{1}{x - cv} dv = -\frac{1}{c} \log(1 - cs/x), \quad \text{and} \quad J_t = \frac{x}{c} (1 - e^{-ct}), \]
and so $Y_t = xe^{-ct}$.

Example 103 (Feller diffusion) For $\phi(\gamma) = \gamma^2$ we obtain Feller’s diffusion. There are lots of parallels with Brownian motion. There is a Donsker-type result which says that rescaled Galton-Watson processes converge to Feller’s diffusion. It is the most popular model in applications. A lot of quantities can be calculated explicitly.

Proposition 104 $Y$ is a Markov process. Let $Y^{(1)}$ and $Y^{(2)}$ be two independent continuous-state branching processes with branching mechanism $\phi$ starting from $x > 0$ and $y > 0$. Then $Y^{(1)} + Y^{(2)}$ is a continuous-state branching process with branching mechanism $\phi$ starting from $x + y$. 
Lecture 16

The two-sided exit problem

Reading: Kyprianou Chapter 8, Bertoin Aarhus Notes, Durrett Sections 7.5-7.6

16.1 The two-sided exit problem for Lévy processes with no negative jumps

Let $X$ be a Lévy process with no negative jumps. As we have studied processes with no positive jumps (such as $-X$) before, it will be convenient to use compatible notation and write

$$
\mathbb{E}(e^{-\gamma X}) = e^{t\phi(\gamma)}, \quad \phi(\gamma) = a_X - \gamma + \frac{1}{2}\sigma^2 \gamma^2 + \int_{-\infty}^0 (e^{\gamma x} - 1 - \gamma x1_{|x|\leq 1}) g_X(x)dx
$$

$$
= -a_X - \gamma + \frac{1}{2}\sigma^2 \gamma^2 - \int_0^\infty (1 - e^{-\gamma x} - \gamma x1_{|x|\leq 1}) g_X(x)dx,
$$

where $a_X = -a_X$ and $g_X(x) = g_X(-x)$, $x > 0$. Then we deduce from Section 11.4 that, if $\mathbb{E}(X_1) < 0$,

$$
\mathbb{E}(e^{-\beta X}) = \frac{\beta \mathbb{E}(X_1)}{\phi(\beta)}, \quad \beta \geq 0. \tag{1}
$$

The two-sided exit problem is concerned with exit from an interval $[-a, b]$, notably the time

$$
T = T_{[-a,b]^c} = \inf\{t \geq 0 : X_t \in [-a, b]^c\}
$$

and the probability to exit at the bottom $\mathbb{P}(X_T = -a)$. Note that an exit from $[-a, b]$ at the bottom happens necessarily at $-a$, since there are no negative jumps, whereas an exit at the top may be due to a positive jump across the threshold $b$ leading to $X_T > b$.

Proposition 105 For any Lévy process $X$ with no negative jumps, all $a > 0$, $b > 0$ and $T = T_{[-a,b]^c}$, we have

$$
\mathbb{P}(X_T = -a) = \frac{W(b)}{W(a + b)}, \quad \text{where } W \text{ is such that } \int_0^\infty e^{-\beta x}W(x)dx = \frac{1}{\phi(\beta)}.
$$
Proof: We only prove the case $\mathbb{E}(X_1) < 0$. By (1), we can identify (the right-continuous function) $W$, since

$$-\frac{\beta \mathbb{E}(X_1)}{\phi(\beta)} = \mathbb{E}(e^{-\beta X_\infty}) = \int_0^\infty e^{-\beta x} f_{X_\infty}(x) dx$$

by partial integration, and so by the uniqueness of moment generating functions, we have $W(x) = c \mathbb{P}(X_\infty \leq x)$, where $c = -\mathbb{E}(X_1) > 0$.

Now define $\tau_a = \inf\{t \geq 0 : X_t < -a\}$ and apply the strong Markov property at $\tau_a$ to get a post-$\tau_a$ process $\tilde{X} = (X_{\tau_a + s} + a)_{s \geq 0}$ independent of $(X_r)_{r \leq \tau_a}$, in particular of $X_{\tau_a}$, so that

$$cW(b) = \mathbb{P}(X_\infty \leq b) = \mathbb{P}(X_{\tau_a} \leq b, X_\infty \leq a + b) = \mathbb{P}(X_{\tau_a} \leq b) \mathbb{P}(X_\infty \leq a + b) = \mathbb{P}(X_{\tau_a} \leq b) cW(a + b),$$

and the result follows.

Example 106 (Stable processes) Let $X$ be a stable process of index $\alpha \in (1, 2]$ with no negative jumps, then we have

$$\int_0^\infty e^{-\lambda x} W(x) dx = \lambda^{-\alpha} \Rightarrow \Gamma(\alpha) W(x) = x^{\alpha-1}.$$
Proof: We condition on $B_T$ and use the strong Markov property of $B$ at $T$ to obtain
\[
e^{-a\sqrt{2q}} = \mathbb{E}(e^{-qT(-a)})
\]
\[= \mathbb{P}(B_T = -a)\mathbb{E}(e^{-qT(-a)}|B_T = -a) + \mathbb{P}(B_T = b)\mathbb{E}(e^{-qT(-a)}|B_T = b)
\]
\[= \frac{b}{a+b}\mathbb{E}(e^{-qT}|B_T = -a) + \frac{a}{a+b}\mathbb{E}(e^{-qT}|B_T = b)
\]
\[= \frac{b}{a+b}\mathbb{E}(e^{-qT}|B_T = -a) + \frac{a}{a+b}\mathbb{E}(e^{-qT}|B_T = b)e^{-(a+b)\sqrt{2q}}
\]
and, by symmetry,
\[e^{-b\sqrt{2q}} = \frac{a}{a+b}\mathbb{E}(e^{-qT}|B_T = b) + \frac{b}{a+b}\mathbb{E}(e^{-qT}|B_T = -a)e^{-(a+b)\sqrt{2q}}.
\]
These can be written as
\[
\frac{b+a}{ab} = a^{-1}\mathbb{E}(e^{-qT}|B_T = -a)e^{a\sqrt{2q}} + b^{-1}\mathbb{E}(e^{-qT}|B_T = b)e^{-b\sqrt{2q}}
\]
\[
\frac{b+a}{ab} = a^{-1}\mathbb{E}(e^{-qT}|B_T = -a)e^{-a\sqrt{2q}} + b^{-1}\mathbb{E}(e^{-qT}|B_T = b)e^{b\sqrt{2q}}
\]
and suitable linear combinations give, as required,
\[
2\sinh(a\sqrt{2q})\frac{b+a}{ab} = 2\sinh((a+b)\sqrt{2q})b^{-1}\mathbb{E}(e^{-qT}|B_T = b)
\]
\[
2\sinh(b\sqrt{2q})\frac{b+a}{ab} = 2\sinh((a+b)\sqrt{2q})a^{-1}\mathbb{E}(e^{-qT}|B_T = -a).
\]
\[
\square
\]

**Corollary 108** For Brownian motion $B$, all $a > 0$ and $T = T_{[-a,a]}$, we have
\[
\mathbb{E}(e^{-qT}) = \frac{1}{\cosh(a\sqrt{2q})}
\]

Proof: Just calculate from the previous proposition
\[
\mathbb{E}(e^{-qT}) = \frac{V_q(a)}{V_q(2a)} = 2\frac{e^{\sqrt{2qa}} - e^{-\sqrt{2qa}}}{(e^{\sqrt{2qa}})^2 - (e^{-\sqrt{2qa}})^2} = \frac{1}{\cosh(a\sqrt{2q})}.
\]
\[
\square
\]

### 16.3 Appendix: Donsker’s Theorem revisited

We can now embed simple symmetric random walk (SSRW) into Brownian motion $B$ by putting
\[
T_0 = 0, \quad T_{k+1} = \inf\{t \geq T_k : |B_t - B_{T_k}| = 1\}, \quad S_k = B_{T_k}, \quad k \geq 0,
\]
and for step sizes $1/\sqrt{n}$ modify $T_k^{(n)} = \inf\{t \geq T_k : |B_t - B_{T_k^{(n)}}| = 1/\sqrt{n}\}.$
Theorem 109 (Donsker for SSRW) For a simple symmetric random walk $(S_n)_{n \geq 0}$ and Brownian motion $B$, we have

$$\frac{S_{[nt]}}{\sqrt{nt}} \to B_t, \quad \text{locally uniformly in } t \geq 0, \text{ in distribution as } n \to \infty.$$  

Proof: We use a coupling argument. We are not going to work directly with the original random walk $(S_n)_{n \geq 0}$, but start from Brownian motion $(B_t)_{t \geq 0}$ and define a family of embedded random walks

$$S^{(n)}_k := B_{T^{(n)}_k}, \quad k \geq 0, n \geq 1, \quad \Rightarrow \quad \left( S^{(n)}_{[nt]} \right)_{t \geq 0} \sim \left( \frac{S_{[nt]}}{\sqrt{nt}} \right)_{t \geq 0}.$$  

To show convergence in distribution for the processes on the right-hand side, it suffices to establish convergence in distribution for the processes on the left-hand side, as $n \to \infty$.

To show locally uniform convergence we take an arbitrary $T > 0$ and show uniform convergence on $[0, T]$. Since $(B_t)_{0 \leq t \leq T}$ is uniformly continuous (being continuous on a compact interval), we get in probability

$$\sup_{0 \leq t \leq T} \left| S^{(n)}_{[nt]} - B_t \right| = \sup_{0 \leq t \leq T} \left| B_{T^{(n)}_{[nt]}} - B_t \right| \leq \sup_{0 \leq s \leq T} \sup_{|s-t| \leq \varepsilon/n} \left| B_s - B_t \right| \to 0$$  

as $n \to \infty$, if we can show (as we do in the lemma below) that $\sup_{0 \leq t \leq T} |T^{(n)}_{[nt]} - t| \to 0$.

This establishes convergence in probability, which “implies” convergence in distribution for the embedded random walks and for the original scaled random walk. \square

Lemma 110 In the setting of the proof of the theorem, $\sup_{0 \leq t \leq T} |T^{(n)}_{[nt]} - t| \to 0$ in probability.

Proof: First for fixed $t$, we have

$$\mathbb{E}(e^{-qT^{(n)}_{[nt]}}) = \left( \mathbb{E}(e^{-qT^{(n)}_{[nt]}}) \right)^{[nt]} = \frac{1}{(\cosh(\sqrt{2q/n}))^{[nt]}} \to e^{-qt},$$

since $\cosh(\sqrt{2q/n}) \sim 1 + q/n + O(1/n)$. Therefore, in probability $T^{(n)}_{[nt]} \to t$. For uniformity, let $\varepsilon > 0$. Let $\delta > 0$. We find $n_0 \geq 0$ such that for all $n \geq n_0$ and all $t_k = k\varepsilon/2$, $1 \leq k \leq 2T/\varepsilon$ we have

$$\mathbb{P}(\left| T^{(n)}_{[nt_k]} - t_k \right| > \varepsilon/2) < \delta\varepsilon/2T,$$

then

$$\mathbb{P} \left( \sup_{0 \leq t \leq T} \left| T^{(n)}_{[nt]} - t \right| > \varepsilon \right) \leq \mathbb{P} \left( \sup_{1 \leq k \leq 2T/\varepsilon} \left| T^{(n)}_{[nt_k]} - t_k \right| > \varepsilon/2 \right) \leq \sum_{k=1}^{2T/\varepsilon} \mathbb{P} \left( \left| T^{(n)}_{[nt_k]} - t_k \right| > \varepsilon/2 \right) < \delta.$$  

\square

We can now describe the recipe for the full proof of Donsker’s Theorem. In fact, we can embed every standardized random walk $((S_k - k\mathbb{E}(S_1))/\sqrt{n\text{Var}(S_1)})_{k \geq 0}$ in Brownian motion $X$, by first exits from independent random intervals $[-A^{(n)}_{k}, B^{(n)}_{k}]$ so that $X_{T^{(n)}_k} \sim (S_k - k\mathbb{E}(S_1))/\sqrt{n\text{Var}(S_1)}$, and the embedding time change $(T^{(n)}_{[nt]})_{t \geq 0}$ can still be shown to converge uniformly to the identity.
Appendix A

Assignments

Assignment sheets are issued on Tuesdays of weeks 1-7. They are made available on the website of the course at

http://www.stats.ox.ac.uk/~winkel/ms3b.html.

Classes take place in weeks 2 to 8 at times and locations to be determined probably on Wednesdays 2.30pm, Thursdays 3.05pm and/or Fridays 8.55am. The class allocation can be accessed from the course website. Only undergraduates and MSc students in Mathematical and Computational Finance can sign up for classes. All others should talk to me after one of the first two lectures.

Scripts are to be handed in probably by Tuesdays 11am or Wednesdays 4pm in the Department of Statistics, but one of the classes will probably be in OCIAM.

Exercises on the problem sheets vary in style and difficulty. If you find an exercise difficult, please do not deduce that you cannot solve the following exercises, but aim at giving each exercise a serious try. Model solutions will be provided on the course website.

There are lecture notes available. Please print these so that we can make best use of the lecture time. I gradually replace last year’s notes by an updated version. The beginning has not changed much, but some typos and unclear passages have been improved.

Below are some comments on the recommended Reading and Further Reading literature.

Kyprianou: Introductory lectures on fluctuations of Lévy processes with applications. Springer 2006

This is the treatment that is closest to the course. It is based on a Masters course and has been written for Masters students. The book assumes a background in measure-theoretic probability, but is written in a friendly way suitable for a wide range of different backgrounds. The text contains some worked examples and exercises.
Kingman: Poisson processes. OUP 1993
This is a gentle introduction to (general, higher-dimensional) Poisson processes and contains a thorough discussion of the tools leading up to and including the study increasing Lévy processes (Section 8.4). Measure-theoretic arguments are isolated in a few proofs, and the reader can take the measure theory for granted. For consistency of terminology with other Oxford courses and most of the literature, in our course, we will reserve the term “Poisson process” to the one-dimensional process. What Kingman calls “Poisson process” is the associated counting measure, which we call “Poisson counting measure”.

This is not a textbook, but a monograph advertising Lévy process for finance applications. All models for financial stock prices that we study in our course, are discussed in detail, both their properties and how they can be calibrated to fit financial market data. There are also sections on simulation and option pricing.

Sato: Lévy processes and infinitely divisible distributions. CUP 1999
This is a graduate textbook on Lévy processes. The focus is on distributional properties and analytic methods.

Bertoin: Lévy processes. CUP 1996
This is a research monograph on Lévy processes. The approach is sample path based.

Grimmett and Stirzaker: Probability and Random Processes. OUP 2001
This is the standard probability reference book used in Oxford, overarching 1st, 2nd and 3rd year probability courses and more. It contains a section on spatial Poisson processes, a treatment of martingales.

Williams: Probability with Martingales. CUP 1991
This is the standard reference on measure theory and martingales used in Oxford.

We only use the simulation chapter.

This is a good graduate textbook on Probability. We essentially refer to Durrett only for a proof of Donsker’s theorem, but there is also a section on infinite divisibility.

This is a standard reference on Probability. The presentation is extremely concise and economical and the amount of material is encyclopaedic. We only refer to it for rigorous statements and proofs of convergence theorems of random walks to Lévy processes.
A.1 Infinite divisibility and limits of random walks

If in doubt, hand in scripts by Tuesday 22 January 2007, 11am, Department of Statistics

1. (a) Show that \( Y \sim \text{Gamma}(\alpha, \beta) \) with density
\[
g(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta x}, \quad x \geq 0,
\]
is an infinitely divisible distribution and that the independent and identically distributed “divisors” \( Y_{n,j} \) in \( Y_{n,1} + \ldots + Y_{n,n} \sim Y \) are also Gamma distributed.

(b) Show that the \( G \sim \text{geom}(p) \) distribution with probability mass function
\[
P(G = n) = p^n (1 - p), \quad n \geq 0,
\]
is infinitely divisible and that “divisors” are not geometrically distributed.

Hint: Study sums of geometric variables and guess the “divisor” distribution.

(c) Show that the uniform distribution on \([0, 1]\) is not infinitely divisible.

2. (a) Let \( X \) and \( Y \) be independent Lévy processes and \( a, b \in \mathbb{R} \). Show that \( aX + bY \) is also a Lévy process.

(b) Let \( C \) and \( D \) be two independent Gamma Lévy processes with \( C_1 \sim D_1 \sim \text{Gamma}(\alpha, \sqrt{2\mu}) \). Determine the moment generating function of \( C_s - D_s \), \( s \geq 0 \).

We will see later that the process \( C - D \) has in fact the same distribution as \( Z_s = B_{T_s} \), \( s \geq 0 \), for a Brownian motion \( B \) and a Gamma Lévy process \( T \). It is called Variance Gamma process, because \( \text{Var}(B_t) = t \) implies \( \text{Var}(B_{T_s}|T_s) = T_s \sim \text{Gamma}(\alpha s, \mu) \). It is a popular model for financial stock prices.

3. A large number \( N \) of policy holders in a given time period make claims independently of one another with small probability \( p_N \). Denote by \( S_N \) the total number of policy holders who make a claim in the time period. Assume that claim amounts \( A_1, A_2, \ldots \) are independent and identically distributed.

(a) State the Poisson limit theorem and use probability generating functions to prove it.

(b) Explain why \( S_N \) is approximately Poisson distributed and give its parameter.

(c) Calculate the moment generating function of the total amount \( T_N \) of claims.

(d) Show that the distribution of \( T_N \) is well-approximated by a compound Poisson distribution, by precisely formulating and proving a limit theorem of the form
\[
T_N = \sum_{n=1}^{S_N} A_n \rightarrow T_\infty = \sum_{n=1}^{\infty} A_n \quad \text{in distribution, as } N \rightarrow \infty.
\]

4. (a) Let \( A_1, A_2, \ldots \) be independent and identically distributed random variables with \( \mu = \mathbb{E}(A_1) \) and \( \sigma^2 = \text{Var}(A_1) \in (0, \infty) \). Define
\[
Y_{n,k} = \frac{A_k - \mu}{\sigma \sqrt{n}} \quad \text{and} \quad V_n = \sum_{k=1}^{n} Y_{n,k}.
\]
(i) Formulate the Central Limit Theorem for $A_1, A_2, \ldots$ in terms of $V_n$.

(ii) Let $x > 0$. Apply Tchebychev’s inequality to the random variable $B_n = |A_1 - \mu|1_{\{A_1 - \mu \geq \sigma x \sqrt{n}\}}$ to show that there is a sequence $\gamma_n(x) \to 0$ as $n \to \infty$ with

$$
\mathbb{P}(|A_1 - \mu| > \sigma x \sqrt{n}) \leq \frac{\gamma_n(x)}{n}.
$$

(iii) Define $M_n = \max\{|Y_{n,1}|, \ldots, |Y_{n,n}|\}$. Show that $\mathbb{P}(M_n \leq x) \to 1$ as $n \to \infty$, for all $x > 0$. Deduce that $M_n \to 0$ in probability.

(b) Consider an urn initially containing $r$ red and $s$ black balls, $r, s \geq 1$. One ball is drawn with replacement (stage 1). After this, a black ball is added to the urn and two balls are drawn, each with replacement (stage 2). After this, another black ball is added and three balls drawn with replacement (stage 3). Continue so that $n$ balls are drawn at stage $n$ followed by the addition of a single black ball. Let $Y_{n,k} = 1$ resp. $0$ if the $k$th ball of stage $n$ is red resp. black, $1 \leq k \leq n$ and $W_n = Y_{n,1} + \ldots + Y_{n,n}$.

(i) Show that $W_n \to \text{Poi}(r)$.

(ii) Show that $\mathbb{P}(Y_{n,k} = 0) \to 1$ as $n \to \infty$.

(iii) Define $M_n = \max\{|Y_{n,1}|, \ldots, |Y_{n,n}|\}$. Show that $\mathbb{P}(M_n = 0) \to e^{-r}$ as $n \to \infty$. Deduce that $M_n \not\to 0$ in probability.

(c) (i) Formulate Donsker’s theorem and the process version of the Poisson limit theorem in the settings of (a) and (b). Hint: Consider only $t \in [0,1]$ and evaluate the discrete processes $V$ and $W$ at $[nt]$, $t \in [0,1]$.

(ii) Show that in both cases $M_n$ converges in distribution to the size of the biggest jump of the limit process during the time interval $[0,1]$.

Hint for 3.(c)-(d): If $(X_n)_{n \geq 0}$ or $(X_t)_{t \geq 0}$ is a stochastic process and $N$ or $T$ an independent random time, then for real-valued functions $g$ for which the expectations exist, we have

$$
\mathbb{E}(g(X_N)) = \sum_{n=0}^{\infty} \mathbb{E}(g(X_n))\mathbb{P}(N = n) \quad \text{and} \quad \mathbb{E}(g(X_T)) = \int_0^{\infty} \mathbb{E}(g(X_t))f_T(t)dt.
$$

To prove the first formula for integer-valued $(X_n)_{n \geq 0}$ just note that

$$
\mathbb{E}(g(X_N)) = \sum_{k=-\infty}^{\infty} g(k)\mathbb{P}(X_N = k) = \sum_{k=-\infty}^{\infty} \sum_{n=0}^{\infty} g(k)\mathbb{P}(X_N = k, N = n) = \sum_{k=-\infty}^{\infty} \sum_{n=0}^{\infty} g(k)\mathbb{P}(X_n = k)\mathbb{P}(N = n) = \sum_{n=0}^{\infty} \mathbb{P}(N = n)\mathbb{E}(g(X_n)).
$$
A.2 Poisson counting measures

1. (a) Let \((X_t)_{t \geq 0}\) be a Poisson process with rate \(\lambda \in (0, \infty)\) and arrival times \(T_1, T_2, \ldots\). Show that \(N((c,d]) = \# \{ n \in \mathbb{N} : c < T_n \leq d \}\) is a Poisson counting measure on \([0, \infty)\) with constant intensity \(\lambda\).

(b) Let \(N\) be a Poisson counting measure on \([0, \infty)\) with time-varying intensity \(\lambda(t), \ t \geq 0\), continuous. Denote \(X_t = N([0,t])\) and \(T_j = \inf \{ t \geq 0 : X_t = j \}\), \(j \geq 1\).

(i) Show that \((X_t)_{t \geq 0}\) has independent increments.
(ii) Show that \((X_t)_{t \geq 0}\) has stationary increments if and only if the intensity function \(\lambda(t)\) is constant.
(iii) Show that \((X_t)_{t \geq 0}\) has right-continuous paths with left limits.
(iv) Calculate the distribution of \(X_t - X_s\).
(v) Calculate the survival function \(P(T_1 > s), s \geq 0\), of \(T_1\).
(vi) Show that \(T_2 - T_1\) is independent of \(T_1\) if and only if the intensity function \(\lambda\) is constant. Calculate the joint density of \((T_1, T_2 - T_1)\).

2. Let \(\Pi\) be a spatial Poisson process with constant intensity \(\lambda\) on the ball \(\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}\). Let \(P\) be the process given by the \((x, y)\) coordinate of the points (think of the points as being projected onto the \((x, y)\) plane passing through the centre of the ball). Show that \(P\) is a spatial Poisson process and find its intensity function. \(\text{Hint: For a rectangle } A \text{ in the } (x, y) \text{ plane, what points of the ball is } P\) counting?

3. Let \((X_t)_{t \geq 0}\) be a Lévy process with \(\mathbb{E}(X_t^2) < \infty\). Denote \(\mu = \mathbb{E}(X_1), \ \sigma^2 = \text{Var}(X_1)\) and \(e^{-\psi(\lambda)} = \mathbb{E}(e^{i\lambda X_1})\). If \(\mathbb{E}(e^{\gamma X_1}) < \infty\), denote \(e^{\Psi(\gamma)} = \mathbb{E}(e^{\gamma X_1})\). Show that the following processes are martingales.

(a) \(\exp\{\gamma X_t - t\Psi(\gamma)\}\), if \(\mathbb{E}(e^{\gamma X_1}) < \infty\). \(\text{Hint: First show that } \mathbb{E}(\exp\{\gamma X_t\}) = e^{t\Psi(\gamma)} \text{ for all } t = 1/m, \text{ then for all } t \in \mathbb{Q} \cap [0, \infty), \text{ and finally, using right-continuity, for all } t \in [0, \infty).}
(b) \(\exp\{i\lambda X_t + t\psi(\lambda)\}\).
(c) \(X_t - t\mu\).
(d) \((X_t - t\mu)^2 - t\sigma^2\).

4. (a) Show that for \(\beta > 0\) and \(\gamma < \beta\)

\[
\int_0^\infty (e^{\gamma x} - 1 - \frac{1}{x} e^{-\beta x}) dx = - \log \left( 1 - \frac{\gamma}{\beta} \right)
\]

\(\text{e.g. by suitable (and well-justified) differentiation under the integral sign.}\)
(b) Let \((\Delta_t)_{t \geq 0}\) be a Poisson point process with intensity function \(\alpha x^{-1} e^{-\beta x}\). Use the exponential formula for Poisson point processes to show that

\[C_t = \sum_{s \leq t} \Delta_s \text{ has a Gamma distribution, density } \frac{\beta^{\alpha t}}{\Gamma(\alpha t)} x^{\alpha t-1} e^{-\beta x}, \ x \geq 0.\]

(c) Show that \((C_t)_{t \geq 0}\) as defined in (b) is a Lévy process.
5. Let $X$ and $Y$ be two independent increasing compound Poisson processes. Denote the respective jump rates by $\lambda_X$ and $\lambda_Y$, assume that the jump size distributions are continuous with densities $h_X$ and $h_Y$. Denote $D = X - Y$.

(a) Show that $X$ and $Y$ have no jump times in common.

(b) Show that $D$ has jump times according to a Poisson process with rate $\lambda_X + \lambda_Y$.

(c) Calculate the distribution of the first jump size of $D$.

(d) Show that $(\Delta X_t)_{t \geq 0}$ is a Poisson point process and specify its intensity function.

(e) Show that $(\Delta D_t)_{t \geq 0}$ is a Poisson point process and specify its intensity function.

(f) Deduce from (d) and (e) that $D$ is also a compound Poisson process.

(g) Show that every real-valued compound Poisson process $C$ can be written uniquely as the difference of two independent increasing compound Poisson processes.

Note that the theory of Poisson point processes applied in (d)-(f) is neater than the conditioning in (c) that could also be developed and iterated to establish (f). Intensity functions just add, jump size distributions are mixtures/weighted averages.

Remark: Interchanging limits and expectation/integration/summation is not always permitted, and while we do not develop in this course the reasons why we may interchange, we add “by monotone convergence”, whenever we have increasing or decreasing limits of finite quantities. General measure-theoretic statements have been established in B10a, special cases for Lebesgue integrals have been established in Part A Integration, which is also not a prerequisite for this course. As in BS3a, it is enough for our purposes to formulate special cases, whose statements do not require any of the formal technical setup.

For convergence as $n \to \infty$ these are:

- $Z_n \uparrow Z$ and $\mathbb{E}(|Z_n|) < \infty$ for all $n \in \mathbb{N}$ implies $\mathbb{E}(Z_n) \uparrow \mathbb{E}(Z) \in \mathbb{R} \cup \{\infty\}$.

- $f_n \uparrow f$ and $\int_{\mathbb{R}} |f_n(x)| dx < \infty$ for all $n \in \mathbb{N}$ implies $\int_{\mathbb{R}} f_n(x) dx \uparrow \int_{\mathbb{R}} f(x) dx \in \mathbb{R} \cup \{\infty\}$.

- $a_m^{(n)} \uparrow a_m$ for all $m \in \mathbb{N}$ and $\sum_{m=0}^{\infty} |a_m^{(n)}| < \infty$ for all $n \in \mathbb{N}$ implies $\sum_{m=0}^{\infty} a_m^{(n)} \uparrow \sum_{m=0}^{\infty} a_m \in \mathbb{R} \cup \{\infty\}$.

- $\Delta_s^{(n)} \uparrow \Delta_s$ and $\sum_{s \leq t} |\Delta_s^{(n)}| < \infty$ for all $n \in \mathbb{N}$ implies $\sum_{s \leq t} \Delta_s^{(n)} \uparrow \sum_{s \leq t} \Delta_s \in \mathbb{R} \cup \{\infty\}$.

The last statement is useful to show right-continuity and the existence of left limits of sums of Poisson point processes, such as in 4.(c)
A.3 Construction of Lévy processes

1. Let $(\Delta_t)_{t \geq 0}$ be a Poisson point process with intensity function $g(x) = x^\kappa e^{-x}$ for a parameter $\kappa \in \mathbb{R}$.
   
   (a) Let $\kappa \in (-1, \infty)$. Show that
   
   $$ C_t = \sum_{s \leq t} \Delta_s $$
   
   is a compound Poisson process. Specify its jump rate $\lambda$ and jump density $h$.
   
   (b) Let $\kappa \leq -1$. Show that $\Delta(\frac{n}{t}) = \Delta t \mathbf{1}_{\{\Delta t > 1/n\}}$, $t \geq 0$, is a Poisson point process. Specify its intensity function. Show that
   
   $$ C^{(n)}_t = \sum_{s \leq t} \Delta^{(n)}_s $$
   
   is a compound Poisson process.
   
   (c) For $C^{(n)}_t$ as defined in (b), show that $C^{(n)}_t$ converges to a limit $\rightarrow C_t < \infty$ as $n \rightarrow \infty$ if and only if $\kappa > -2$. Specify the moment generating function of $C_t$.
   
   (d) Show that for $\kappa > -2$, we have
   
   $$ \sup_{s \leq t} |C^{(n)}_s - C_s| = |C^{(n)}_t - C_t|.$$ 
   
   Deduce that $C^{(n)}_t \rightarrow C_t$ a.s. (or in probability) locally uniformly.
   
   (e) Show that $C^{(n)}_t - \mathbb{E}(C^{(n)}_t)$ converges for $\kappa > -3$. Show that the limit is a Lévy process and that it has unbounded variation.

2. Let $(X_t)_{t \geq 0}$ be a stable subordinator in the sense that $(c^{1/\alpha}X_t/c)_{t \geq 0} \sim X$ for all $c > 0$ (scaling relation) for some $\alpha \in \mathbb{R}$.
   
   (a) Show that for all $\mu \geq 0$ and $t \geq 0$, we have $\mathbb{E}(e^{-\mu X_t}) \in (0, 1]$. Denote $\Phi_t(\mu) = -\ln(\mathbb{E}(e^{-\mu X_t}))$ and $\Phi = \Phi_1$.
   
   (b) Show that $\Phi_t(\mu) = t\Phi(\mu)$ for all $t \geq 0$. Deduce from the scaling relation that $\Phi(\mu) = \Phi(1)\mu^\alpha$ for all $\mu \geq 0$.
   
   (c) Show that $\frac{\partial}{\partial \mu} \mathbb{E}(e^{-\mu Y}) \leq 0$ and $\frac{\partial^2}{\partial \mu^2} \mathbb{E}(e^{-\mu Y}) \geq 0$ for any nonnegative random variable $Y$ and for all $\mu > 0$ with equality if and only if $\mathbb{P}(Y = 0) = 1$. Deduce that $\alpha \in (0, 1]$ or $\mathbb{P}(X = 0) = 1$.
   
   (d) By letting $\mu \downarrow 0$ in (c), show that $\mathbb{E}(X_t) = \infty$ for all $t > 0$ and $\alpha \in (0, 1)$.
   
   (e) For $\alpha \in (0, 1)$, calculate $g : (0, \infty) \rightarrow (0, \infty)$ such that
   
   $$ \Phi(\mu) = \int_0^\infty (1 - e^{-\mu x})g(x)dx.$$ 
   
   (f) For every $\alpha \in (0, 1]$ and $\Phi(1) = b > 0$, show that there exists a stable subordinator $(X_t)_{t \geq 0}$.
3. (a) Let \((X_t)_{t \geq 0}\) and \((Y_t)_{t \geq 0}\) be independent stable subordinators with common index \(\alpha\) and intensities \(b_X = \Phi_X(1)\) and \(b_Y = \Phi_Y(1)\). Show that \(Z = X - Y\) is also a stable process with index \(\alpha\) in the sense that \((c^{1/\alpha} X_t/c)_{t \geq 0} \sim X\) for all \(c > 0\).

(b) Let \(H\) be a real-valued random variable with symmetric distribution, i.e. \(H \sim -H\). Show that \(\mathbb{E}(e^{i\lambda H}) \in \mathbb{R}\) for all \(\lambda \in \mathbb{R}\). Hint: \(e^{ix} = \cos(x) + i \sin(x)\).

(c) In the setting of (a), for the special case \(b_X = b_Y\), show that \(\mathbb{E}(\exp\{i\lambda Z_t\}) = \exp\{-b_b|\lambda|^\alpha\}, \lambda \in \mathbb{R}\). Hint: Show that \(Z_t\) is symmetric and that all symmetric stable processes have a characteristic function of this form for some \(b > 0\). You may assume without proof that all characteristic functions are continuous in \(\lambda \in \mathbb{R}\) and that those of infinitely divisible distributions have no zeros.

(d) Fix \(\tilde{b} > 0\). Show that for \(\alpha \in (0, 2)\), the function
\[
\tilde{\psi}(\lambda) = \int_{-\infty}^{\infty} (\cos(\lambda x) - 1)\tilde{b} |x|^{\alpha-1} dx
\]
has the property \(\tilde{\psi}(\lambda c^{1/\alpha}) = c \tilde{\psi}(\lambda)\) for all \(c > 0\) and \(\lambda \in \mathbb{R}\). Deduce that
\[
\tilde{\psi}(\lambda) = b|\lambda|^\alpha
\]
for some \(b \in \mathbb{R}\).

(e) Using (d) or otherwise, show that a symmetric stable process \(R\) of index \(\alpha \in (0, 2]\) has bounded variation if and only if \(\alpha \in (0, 1]\), and deduce from the previous parts of the exercise that it can then be written as a difference of two stable subordinators. Show that \(\mathbb{E}(R_t)\) exists if and only if \(\alpha \in (1, 2]\), and that then \(\text{Var}(R_t) < \infty\) if and only if \(\alpha = 2\).

Warning: Densities of stable processes are only known in closed form for some special cases \(\alpha \in \{1/2, 1, 2\}\). It is known, however, that they all have smooth probability density functions.

You may use results from the lectures and previous assignment sheets without proof if you state them clearly, except that the Lévy density \(g\) of stable processes is to be derived here, and its form should not be assumed.

Question A.3.1 is the most relevant on this sheet for MSc MCF students – if pushed for time, please focus on this one.
A.4 Simulation

1. Let $U \sim \text{Unif}(0, 1)$ and $F : \mathbb{R} \rightarrow [0, 1]$ right-continuous (weakly) increasing with $F(-\infty) = 0$ and $F(\infty) = 1$. Define $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) > u\} \in [-\infty, \infty]$ for $u \in [0, 1]$. Show that $F^{-1}(U)$ is a random variable with cumulative distribution function $F$.

2. (a) Let $(X_t)_{t \geq 0}$ be a Gamma process with $X_t \sim \text{Gamma}(t, 1)$ for all $t > 0$. Consider $A = X_a$ and $B = X_{a+b} - X_a$ for some $a > 0$ and $b > 0$. Show that $R = A/(A+B)$ and $S = A+B$ are independent and that $R \sim \text{Beta}(a, b)$, where Gamma and Beta densities are recalled as follows:
\[
    f_S(s) = \frac{s^{a+b-1}e^{-s}}{\Gamma(a+b)}, \quad s \in (0, \infty),
\]
\[
    f_R(r) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} r^{a-1}(1-r)^{b-1}, \quad r \in (0, 1).
\]
Deduce, vice versa, the distribution of $(SR, S/(1-R))$ for independent $S \sim \text{Gamma}(1, c)$ and $R \sim \text{Beta}(cp, c(1-p))$, for some $c > 0$ and $p \in (0, 1)$.

(b) Let $U \sim \text{Unif}(0, 1)$ and $a > 0$. Show that $X = U^{1/a} \sim \text{Beta}(a, 1)$.

(c) Let $U \sim \text{Unif}(0, 1)$ and $V \sim \text{Unif}(0, 1)$ be independent and $a \in (0, 1)$. Calculate for $Y = U^{1/a}$ and $Z = V^{1/(1-a)}$
\[
\mathbb{P}\left(\frac{Y}{Y+Z} \leq t, Y+Z \leq 1\right)
\]
and deduce that the conditional distribution of $W = Y/(Y+Z)$ given $Y+Z \leq 1$ is $\text{Beta}(a, 1-a)$. Hint: Write both inequalities as constraints on $Z$ to find the bounds when writing the probability as a double integral.

(d) In the setting of (c), show that the conditional distribution of $TW$ given $Y+Z \leq 1$, for an independent $T \sim \text{Exp}(1) = \text{Gamma}(1, 1)$ random variable, is $\text{Gamma}(a, 1)$.

(e) Consider the following procedure due to Johnk. Let $a \in (0, 1)$.

1. Generate two independent random numbers $U \sim \text{Unif}(0, 1)$ and $V \sim \text{Unif}(0, 1)$.
2. Set $Y = U^{1/a}$ and $Z = V^{1/(1-a)}$.
3. If $Y + Z \leq 1$ go to 4., else go to 1.
4. Generate an independent $C \sim \text{Unif}(0, 1)$ and set $T = \ln(C)$.
5. Return the number $TY/(Y+Z)$.

What is this procedure doing? Explain its relevance for simulations.

3. (a) In the light of the previous exercise, explain how you can generate a Beta($a, b$) random variable from a sequence of Unif(0, 1) random variables, for any $a > 0$ and $b > 0$. Hint: Consider $a \in (0, 1)$ first and use the additivity of Gamma variables to generate Gamma($a, 1$) variables, from which the Beta variable can be constructed.
(b) Consider the following method to generate a Gamma process on the time interval $[0,1]$. Set $X_0 = 0$ and generate $X_1 \sim \text{Gamma}(1,1)$. For $n \geq 0$, if you have generated $X_{k2^{-n}}$, $k = 0, \ldots, 2^n$, generate $B_{k,n} \sim \text{Beta}(a_n, b_n)$ and set $X_{(2k-1)2^{-n-1}} = X_{(k-1)2^{-n}} + B_{k,n}(X_{(k+1)2^{-n}} - X_{k2^{-n}})$, $1 \leq k \leq 2^n$. For what choices of $a_n > 0$ and $b_n > 0$ does this procedure yield Gamma distributions for all $X_{k2^{-n}}$, and for what choice do you get stationary increments? Hint: $a_n = b_n = 2^{-n-1}$ works, but are there other choices?

(c) What are the advantages of this method when compared with the plain version of the time discretisation method (Method 1)?

4. Consider a variant of the Variance Gamma process of the form $V_t = at + G_t - H_t$ where $a \in \mathbb{R}$, $G_t \sim \text{Gamma}(\alpha_+, \beta_+)$ and $H_t \sim \text{Gamma}(\alpha_-, \beta_-)$

(a) For what values of $a, \alpha_+, \beta_+$ is $V$ a martingale?

(b) Write out the steps needed to simulate $V_t$

• by Method 1 (using a random walk with increment distribution $\sim V_0$)
• by Method 1 (applied to $G$ and $H$ separately)
• by the refinement of Method 1 given in A.4.3
• by Method 2 (simulating the Poisson point process of jumps truncated at $\varepsilon$)

(c) Carry out 9 simulations for a range of parameters $\alpha_+ \in \{1, 10, 100\}$ and $\alpha_- \in \{10, 100, 1000\}$, $\beta_+ = \alpha_+^2/2$ and $a$ such that $V$ is a martingale. This part of this question is optional.

Warning: The incomplete Gamma function $\Gamma_t(a) = \int_0^t x^{a-1}e^{-x}dx$ cannot be simplified into closed form (nor expressed in terms of the Gamma function), except for some special values of $a$ such as $a \in \mathbb{N}$. There are, however, numerical procedures to evaluate $\Gamma_t(a)$, which we will not address in this course.

If you have not used R, but would like to, you will find the “First steps with R” at http://www.stats.ox.ac.uk/~myers/stats_materials/R_intro/WA5_R.pdf useful. Following are brief explanations of the commands used in the sample file http://www.stats.ox.ac.uk/~winkel/gammadvgamma.R

This is a script file, which has to be run in the command window “R Console” to make the new commands available, e.g. select “Run all” in the drop-down menu “Edit”.

- `runif(n,a,b)` generates an $n$-vector of uniform variables on $[a,b]$.
- `qgamma(u,a,b)` evaluates $F^{-1}(u)$ for the Gamma($a,b$) inverse distribution function $F^{-1}$ at $u$. If $u$ is a vector, `qgamma` is applied to each component.
- `1:n` generates the vector $(1,2,\ldots,n)$. Multiplication of vectors $v$ by scalars $a$ can be written as $a*v$, similarly for addition and subtraction of vectors.
- `plot(x,y,pch=".",sub=paste("text"))` produces a scatter plot of pairs $(x_i,y_i)$ for vectors $x$ and $y$, with . marking the points, and text in the caption.
- `psum <- function(vector){\ldots}` defines a new command `psum` that takes a vector as an argument. When this line is executed, the command is just made available. To execute the command, type `psum(v)` for a vector $v$ to get the partial sums of $v$ displayed, or $s$=`psum(v)` to create a new vector $s$ containing the partial sums of $v$. 


A.5 Financial models

1. Consider a one-period model with three assets, a risk-free asset that increases from $A_0 = 1$ to $A_1 = e^\delta$ and two risky assets $B$ and $C$ that can each move up to or down from $B_0 = C_0 = 1$, so that there are four scenarios $\omega_1 = (\text{up, up}), \omega_2 = (\text{up, down}), \omega_3 = (\text{down, up})$ and $\omega_4 = (\text{down, down})$. Suppose that $B_1^{\text{up}} = B_1(\omega_1) = B_1(\omega_2) > B_1(\omega_3) = B_1(\omega_4) = B_1^{\text{down}}$ and $C_1^{\text{up}} = C_1(\omega_1) = C_1(\omega_3) > C_1(\omega_2) = C_1(\omega_4) = C_1^{\text{down}}$. Assume w.l.o.g. that $B_1^{\text{up}} < C_1^{\text{up}}$.

(a) For a portfolio $(T_0, U_0, V_0)$ of $T_0$ units of $A$, $U_0$ units of $B$ and $V_0$ units of $C$, specify the value $W_0$ and $W_1(\omega_i)$ of the portfolio at times 0 and 1, $i = 1, 2, 3, 4$.

(b) Show that this model is arbitrage-free if and only if $B_1(\omega_1) > A_1 > B_1(\omega_4) > C_1(\omega_4)$.

Consider the arbitrage-free case now.

(c) Give an example of a contingent claim $W_1(\omega_i)$ that cannot be hedged.

(d) Show that contingent claims of the form $W_1(\omega_1) = W_1(\omega_2), W_1(\omega_3) = W_1(\omega_4)$ can be hedged and priced as $e^{-\delta E(W_1)}$, where you should specify $q_B = \mathbb{P}(B_1 = B_1^{\text{up}})$ and $1 - q_B = \mathbb{P}(B_1 = B_1^{\text{down}})$.

In particular, $e^{-\delta t} B_t$, $t = 0, 1$, is then a martingale. Is $q_B$ unique?

(e) State the result analogous to (d) for contingent claims relating to $C$ only rather than $B$ only.

(f) Show that there are infinitely many possibilities to choose $p_1 = \mathbb{P}(\omega_1) = \mathbb{P}(B_1 = B_1^{\text{up}}, C_1 = C_1^{\text{up}}), p_2 = \mathbb{P}(\omega_2), p_3 = \mathbb{P}(\omega_3)$ and $p_4 = \mathbb{P}(\omega_4)$ so that $e^{-\delta t} B_t$ and $e^{-\delta t} C_t$, $t = 0, 1$ are martingales.

(g) Consider the contingent claim $W_1(\omega_1) = 1, W_1(\omega_2) = W_1(\omega_3) = W_1(\omega_4) = 0$. Using the range of possibilities for $p = (p_1, p_2, p_3, p_4)$, give the range of proposed (arbitrage-free) prices $W_0 = e^{-\delta E_p(W_1)}$.

2. Consider $X_t = N_t - \mu t$ for a Poisson process $N$ of rate $\lambda \in (0, \infty)$ and a drift coefficient $\mu \in (0, \infty)$. Let

$$S_n^{(\varepsilon)} = \sum_{i=1}^{n} X_i^{(\varepsilon)},$$

where $(X_i^{(\varepsilon)})_{i \geq 1}$ is a sequence of independent random variables with common probability mass function given by

$$\mathbb{P}(X_1^{(\varepsilon)} = 1 - \mu \varepsilon) = 1 - e^{-\lambda \varepsilon} =: p_{\varepsilon} \quad \text{and} \quad \mathbb{P}(X_1^{(\varepsilon)} = -\mu \varepsilon) = 1 - p_{\varepsilon}.$$

(a) Show that $S_{[t/\varepsilon]}^{(\varepsilon)} \rightarrow X_t$ in distribution as $\varepsilon \downarrow 0$. Hint: You can either prove this directly or consider $T_n^{(\varepsilon)} = S_n^{(\varepsilon)} + n \mu \varepsilon$ first.
(b) Show that the market model \((e^{\delta n}, e^{S_n^{(e)}})_{n \geq 0}\) is arbitrage-free if and only if 
\[-\mu < \delta < 1/\varepsilon - \mu\] 
and then also complete.

Consider the arbitrage-free case now.

(c) Show that the martingale probabilities are

\[q_{\varepsilon} = \mathbb{P}(\tilde{S}_1^{(e)} = 1 - \mu \varepsilon) = \frac{e^{\varepsilon (\delta + \mu)} - 1}{e - 1}.\]

(d) Show that under the martingale probabilities

\[\tilde{S}_{t/\varepsilon}^{(e)} \to \tilde{X}_t = \tilde{N}_t - \mu t,\]

where \((\tilde{N}_t)_{t \geq 0}\) is a Poisson process with rate \((\delta + \mu)/(e - 1)\).

(e) Show that in the notation of part (d), the discounted process \(e^{-\delta t}R_t\) associated with \(R_t = e^{\tilde{N}_t - \mu t}\) is a martingale. By conditioning on \(N_t\) and \(\tilde{N}_t\), respectively, explain briefly why the distribution of \(e^{\tilde{N}_t - \mu t}\) can be seen as providing martingale probabilities for \(e^{N_t - \mu t}\).

(f) Using the following subsets of the set of right-continuous paths with left limits

\[D_\mu = \{f \in D([0, 1], (0, \infty)) : \Delta \log f(t) = 1 \text{ or } (\log f)'(t) = -\mu \text{ for all } t \leq 1\},\]

show that \((e^{\tilde{N}_t - \mu t})_{0 \leq t \leq 1}\) is the only exponential Lévy process that has the same set of possible paths as \((e^{N_t - \mu t})_{0 \leq t \leq 1}\) and whose discounted process is a martingale. Remark: In fact, it is the only process that has the same set of possible paths as \((e^{N_t - \mu t})_{0 \leq t \leq 1}\) and whose discounted process is a martingale, so the market model is complete.
A.6 Time change and subordination

1. Consider Brownian motion \((B_t)_{t \geq 0}\) and a continuous increasing function \(f : [0, \infty) \to [0, \infty)\) with \(f(0) = 0\). Set \(Z_y = B_{f(y)},\ y \geq 0\).

(a) Show that \(Z\) has quadratic variation
\[
[Z]_y := \operatorname{p-lim}_{n \to \infty} \sum_{j=1}^{\lfloor 2^n y \rfloor} (Z_{j2^{-n}} - Z_{(j-1)2^{-n}})^2 = f(y),
\]
where \(\operatorname{p-lim}\) denotes a limit in probability of random variables.

(b) Assume that \(f\) is piecewise differentiable on \([0, \infty)\) with piecewise constant derivative \(\sigma^2(s) := f'(s)\), say taking values \(\sigma^2_j\) on intervals \([y_{j-1}, y_j)\) for some \(0 = y_0 < y_1 < \ldots < y_n < \ldots\). Let \((W_y)_{y \geq 0}\) be a Brownian motion. Show that the process
\[
\tilde{Z}_y = \int_0^y \sigma(r)dW_r := \sum_{i=1}^j \sigma_i(W_{y_i} - W_{y_{i-1}}) + \sigma_{j+1}(W_y - W_{y_j}),
\]
y\(_j \leq y < y_{j+1}\), has the same distribution as \(Z\).

This result holds in fact for a very wide class of stochastic processes \(\sigma\). This is why both time-change models and models where the Brownian motion coefficient varies are called stochastic volatility processes.

(c) Give an example of a Lévy process and a function \(f\) as in (b) for which \(X_{f(y)}\) does not have the same distribution as \(\int_0^y \sqrt{f'(s)}dX_s\).

In fact, Brownian motion is the only process for which this holds for all such functions \(f\).

2. Let \((X_t)_{t \geq 0}\) be a Lévy process with probability density function \(f_t\) and \((\tau_y)_{y \geq 0}\) a subordinator with characteristics \((0, g_\tau)\) (sum of jumps, no compensation!). Define
\[
g(z) = \int_0^\infty f_t(z)g_\tau(t)dt, \quad z \in \mathbb{R} \setminus \{0\}.
\]

(a) In the case \(\operatorname{Var}(X_1) < \infty\) and \(\operatorname{Var}(\tau_1) < \infty\), show that \(g\) satisfies the requirements of a Lévy density of a Lévy process.

(b) In the case where either \(\tau\) or \(X\) is compound Poisson, show that \(g\) also satisfies the requirements of a Lévy density of a Lévy process. More specifically, if \(X\) is a compound Poisson process with intensity \(\lambda\), then we have \(\mathbb{P}(X_t = 0) \geq e^{-\lambda t}\); assume that, in fact, \(\mathbb{P}(X_t = 0) = e^{-\lambda t}\) and that \(\mathbb{P}(X_t \in (a, b)) = \int_a^b f_t(x)dx\) for \((a, b) \neq 0\).

(c) If \(X\) is a stable process of index \(\alpha > 1\), show that \(g\) is the Lévy density of a bounded variation Lévy process if
\[
\int_0^\infty x^{1/\alpha}g_\tau(x)dx < \infty.
\]
3. Show that all Lévy processes \( X \) that can be obtained by subordination of Brownian motion with an independent subordinator are symmetric in the sense \( X \sim -X \), but that not all symmetric Lévy processes can be obtained in this way.

4. (a) Let \( C \) and \( D \) be two independent Gamma Lévy processes with parameters \( \alpha \) and \( \sqrt{2\lambda} \) for \( C_1 \sim D_1 \). Let \( T \) be a Gamma process with parameters \( \alpha \) and \( \lambda \), and let \( B \) be an independent Brownian motion. Show that \( X_s - Y_s \sim B_{T_s} \).

   This result was mentioned in Question A.1.2. as an explanation for the name Variance Gamma process.

   (b) Let \( B \) be Brownian motion, \( S \) an independent stable subordinator with index \( \alpha \in (0,1) \). Show that \( R_t = B_{S_t}, t \geq 0 \), is a stable process with index \( 2\alpha \).

   (c) Write down procedures to simulate the processes in (a) and (b) using Method 3 (Subordination).

5. Let \( X \) be a Lévy process with \( \mathbb{E}(X_1) = \mu \) and \( \mathbb{V} \text{ar}(X_1) = \sigma^2 \). Let \( S \) be an independent subordinator with \( \mathbb{E}(S_1) = m \) and \( \mathbb{V} \text{ar}(S_1) = q^2 \). Denote \( Z_t = X_{S_t}, t \geq 0 \).

   (a) Show that \( \mathbb{E}(Z_t) = m\mu t, t \geq 0 \).

   (b) Show that \( \mathbb{V} \text{ar}(Z_t) = (\sigma^2 m + \mu^2 q^2)t, t \geq 0 \). \( \text{Hint: Consider } \mathbb{E}(Z_t^2) \text{ first.} \)

   (c) Check this formula for the Variance Gamma process, using A.6.4.(a). For what values of \( \alpha \) and \( \lambda \) is \( \mathbb{E}(Z_1) = 0 \) and \( \mathbb{V} \text{ar}(Z_1) = 1 \)? Show that \( \mathbb{E}(Z_1^4) = 3\mathbb{E}(S_1^2) \) and deduce the range of \( \mathbb{E}(Z_1^4) \) for these values of \( \alpha \) and \( \lambda \). \( \text{Standardized fourth moments (curtosis) give an indication of heavy tails. They reflect why Lévy processes such as the Variance Gamma process can better fit financial price processes.} \)
A.7 Level passage events

1. Let \((X_s)_{s \geq 0}\) be a Lévy process and \(\underline{X} = \inf_{0 \leq s \leq t} X_s, \ t \geq 0.\)

   (a) For a fixed time \(t > 0\), show that the process \((\tilde{X}_s)_{0 \leq s \leq t}\) given by
   \[
   \tilde{X}_s^{(t)} = X_{t -} - X_{t - s}, \quad 0 \leq s \leq t,
   \]
   is a Lévy process with the same distribution as \((X_s)_{0 \leq s \leq t}\).

   (b) Show that this implies that for an independent random time \(\tau\) with density function \(f_\tau(x), x \in (0, \infty)\), we have \((\tilde{X}_s^{(\tau)})_{0 \leq s \leq \tau}\) \sim \((X_s)_{0 \leq s \leq \tau}\) in the sense that for all \(0 = s_0 < s_1 < \ldots < s_n < s_{n+1} = \infty\) and \(0 \leq m \leq n\) we have
   \[
   \mathbb{P}(X_{s_1} \in A_1, \ldots, X_{s_m} \in A_m, \tau \in [s_m, s_{m+1}) \cap B) = \mathbb{P}(\tilde{X}_{s_1}^{(\tau)} \in A_1, \ldots, \tilde{X}_{s_m}^{(\tau)} \in A_m, \tau \in [s_m, s_{m+1}) \cap B)
   \]
   for all intervals \(A_1, \ldots, A_n \subset \mathbb{R}\) and \(B \subset [0, \infty)\).

   (c) Using results and/or arguments from the lectures show that \(\underline{X}\) is independent of \((X_\tau - \underline{X}_\tau)\) for an independent \(\tau \sim \text{Exp}(\gamma)\).

   (d) Suppose now that \(X\) has no positive jumps. Calculate the distribution of \(\underline{X}\).

2. Let \((X_t)_{t \geq 0}\) be an \(\alpha\)-stable Lévy process with no positive jumps for some \(\alpha \in (1, 2]\), i.e. such that \(\mathbb{E}(e^{\gamma X_t}) = e^{t \gamma^\alpha}\). For \(\alpha = 2\) this is Brownian motion, for \(\alpha \in (1, 2)\), we have \(a_3 = 0\) and \(g(x) = c|x|^{-\alpha - 1}\). For \(x \geq 0\) denote \(T_x = \inf\{t \geq 0 : X_t > x\}\).

   (a) Using the strong Markov property of \((X_t)_{t \geq 0}\) at \(T_x\), show that \((T_x)_{x \geq 0}\) is a stable subordinator with index \(1/\alpha\).

   (b) Let \(Y\) have probability density function
   \[
   f_b(z) = \frac{b}{\sqrt{2\pi z^3}} e^{-b^2/(2z)}, \quad z > 0.
   \]
   Calculate the distribution of \(aY\) and deduce that \((f_b)_{b \geq 0}\) is the family of densities of stable distributions on \((0, \infty)\) of index \(1/2\).

   (c) Deduce that there is a constant \(c > 0\) such that
   \[
   \int_0^\infty e^{\gamma x} f_b(x)dx = e^{-cb\sqrt{\gamma}}.
   \]
   In fact, \(c = \sqrt{2}\).

3. (a) Let \(A_1, A_2, \ldots\) be identically distributed and \(S_n = A_1 + \ldots + A_n\) the associated random walk. Let \((N_m)_{m \geq 0}\) be an independent random walk. Denote the moment generating function of \(A_1\) by \(M(\gamma) = \mathbb{E}(\exp\{\gamma A_1\})\) and assume that it is finite for \(\gamma \in (-\varepsilon, \varepsilon)\). Denote the probability generating function of \(N_1\) by \(G(s) = \mathbb{E}(s^{N_1})\). Show that \(R_m = S_{N_m}, m \geq 0,\) is also a random walk (with independent and identically distributed increments).
(b) Let \((X_t)_{t \geq 0}\) be a Lévy process and \((T_s)_{s \geq 0}\) an independent increasing Lévy process. Show that \(Y_s = X_{T_s}, s \geq 0\), is also a Lévy process.

(c) (i) Let \((B_t)_{t \geq 0}\) be Brownian motion. For \(s \geq 0\) define \(T_s = \inf\{t \geq 0 : B_t + bt > s\}\), where \(b \geq 0\) is fixed. Using the strong Markov property at \(T_s\), show that \((T_s)_{s \geq 0}\) is an increasing Lévy process.

(ii) Show that \(\exp\{\gamma B_t - \frac{1}{2} \gamma^2 t\}\) is a martingale for all \(\gamma \in \mathbb{R}\). Use the Optional Stopping Theorem to show that

\[
\mathbb{E}(\exp\{\rho T_s\}) = \exp\{s(b - \sqrt{b^2 - 2\rho})\}.
\]

This distribution is called the inverse Gaussian distribution (note that \(B_{T_s} = s\) means that \(s \mapsto T_s\) is the right inverse of \(t \mapsto B_t\).) For an independent Brownian motion \((X_t)_{t \geq 0}\), the process \(Z_s = X_{T_s}, s \geq 0\), obtained as in (b) has the so-called Normal Inverse Gaussian (NIG) distribution. This is another popular process to model financial price processes.

The last question is optional:

4. Let \((X_t)_{t \geq 0}\) be standard Brownian motion with moment generating function \(\mathbb{E}(e^{\gamma X_t}) = e^{t\gamma^2/2}\). Denote \(\bar{X}_t = \sup_{0 \leq s \leq t} X_s\).

(a) For \(y > 0\) denote \(T_y = \inf\{t \geq 0 : X_t = y\}\). Use the strong Markov property to deduce that the process

\[
X^*_t = \begin{cases} X_t & t \leq T_y \\ 2y - X_t & t \geq T_y \end{cases}
\]

is a Brownian motion.

(b) Show that

\[
\mathbb{P}(X_t \leq x, \bar{X}_t > y) = \mathbb{P}(X^*_t > 2y - x), \quad y \in (0, \infty), x \in (-\infty, y).
\]

and deduce that

\[
f_{X_t, \bar{X}_t}(x, y) = \frac{2(2y - x)}{\sqrt{2\pi t^3}} \exp\left\{ - \frac{(2y - x)^2}{2t} \right\}, \quad y \in (0, \infty), x \in (-\infty, y).
\]

(c) Show that

\[
f_{T_x}(z) = \frac{x}{\sqrt{2\pi z^3}} e^{-x^2/(2z)}, \quad z > 0.
\]

Feedback on the various topics and how you perceived them given your background (no BS3a, no B10a, MScMCF etc.) will be most gratefully received: winkel@stats.ox.ac.uk.
Appendix B

Solutions

B.1 Infinite divisibility and limits of random walks

1. (a) Recall that for independent $A_1 \sim \text{Gamma}(\alpha_1, \beta)$ and $A_2 \sim \text{Gamma}(\alpha_2, \beta)$ we have $A_1 + A_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta)$. A quick proof can be given using moment generating functions. The Gamma distribution has moment generating function 

$$E(\exp\{\gamma A\}) = \int_0^\infty e^{\gamma x} \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta x} dx = \frac{\beta^\alpha}{(\beta - \gamma)^\alpha}, \quad \gamma < \beta.$$ 

We see that 

$$E(\exp\{\gamma (A_1 + A_2)\}) = E(\exp\{\gamma A_1\}) E(\exp\{\gamma A_2\}) = \frac{\beta^{\alpha_1 + \alpha_2}}{(\beta - \gamma)^{\alpha_1 + \alpha_2}}$$

and recognise the moment generating function of the Gamma$(\alpha_1 + \alpha_2, \beta)$ distribution. By the Uniqueness Theorem for moment generating functions, $A_1 + A_2 \sim \text{Gamma}(\alpha_1 + \alpha_2, \beta)$.

If we now choose $Y_{n,1}, \ldots, Y_{n,n} \sim \text{Gamma}(\alpha/n, \beta)$ independent, we obtain, by induction in $n$, that $Y_{n,1} + \ldots + Y_{n,n} \sim \text{Gamma}(\alpha, \beta)$. Since this holds for all $n \geq 1$, a random variable $Y \sim \text{Gamma}(\alpha, \beta)$ has an infinitely divisible distribution.

(b) First calculate for $B_1, B_2 \sim \text{geom}(p)$ independent that 

$$\mathbb{P}(B_1 + B_2 = n) = \sum_{k=0}^n \mathbb{P}(B_1 = k, B_2 = n-k) = \sum_{k=0}^n p^k (1-p) p^{n-k} (1-p)$$

$$= (n+1) p^n (1-p)^2,$$

and, e.g. by induction, for $A_m = B_1 + \ldots + B_m = A_{m-1} + B_m$ a negative binomial distribution. Alternatively, consider independent Bernoulli trials until the $m$th success, then $\{A_m = n\}$ means there have been $n$ failures and $m$ successes, the $m - 1$ first successes chosen from the first $n + m - 1$ trials, and we get 

$$\mathbb{P}(A_m = n) = \binom{n + m - 1}{m - 1} p^n (1-p)^m \frac{(n + m - 1)!}{(m - 1)! n!} p^n (1-p)^m$$

$$= \frac{\Gamma(n + m)}{\Gamma(m)n!} p^n (1-p)^m.$$
This formula makes sense for $m \in (0, \infty)$, and we refer to this probability mass function as $\text{NB}(m, p)$. Then we calculate the probability generating function for $A \sim \text{NB}(m, p)$

$$
\mathbb{E}(s^A) = \sum_{n \geq 0} \frac{\Gamma(n + m)}{\Gamma(m)n!} (sp)^n (1 - p)^m = \frac{(1 - p)^m}{(1 - sp)^m}, \quad s \in [0, 1],
$$

and if $B \sim \text{NB}(r, p)$ is independent, we obtain

$$
\mathbb{E}(s^{A+B}) = \frac{(1 - p)^{m+r}}{(1 - sp)^{m+r}},
$$

the probability generating function of the $\text{NB}(m + r, p)$ distribution, so we conclude by the Uniqueness Theorem for probability generating functions that $A + B \sim \text{NB}(m + r, p)$.

If we now choose $Y_{n,1}, \ldots, Y_{n,n} \sim \text{NB}(1/n, p)$ independent, we obtain, by induction in $n$, that $Y_{n,1} + \ldots + Y_{n,n} \sim \text{NB}(1, p) = \text{geom}(p)$. Since this holds for all $n \geq 1$, a random variable $Y \sim \text{geom}(p)$ has an infinitely divisible distribution.

(c) Assume that a random variable $U \sim \text{Unif}(0, 1)$ can be written as $U = Y_1 + Y_2$ for some independent and identically distributed $Y_1$ and $Y_2$. Then for $x \in [0, 1]$,

$$
1 - x = \mathbb{P}(U \geq x) \geq \mathbb{P}(Y_1 \geq x/2, Y_2 \geq x/2) \Rightarrow \mathbb{P}(Y_1 \geq x/2) \leq \sqrt{1 - x}
$$

and

$$
x = \mathbb{P}(U \leq x) \geq \mathbb{P}(Y_1 \leq x/2)^2 \Rightarrow \mathbb{P}(Y_1 \leq x/2) \leq \sqrt{x}.
$$

For $x = 1$ and $x = 0$, respectively, we deduce $\mathbb{P}(Y_1 \geq 1/2) = 0 = \mathbb{P}(Y_1 \leq 0)$. Now for $x \in (0, 1/2)$

$$
x = \mathbb{P}(U \leq x) \leq \mathbb{P}(Y_1 \leq x, Y_2 \leq x) \iff \mathbb{P}(Y_1 \leq x) \geq \sqrt{x}
$$

and the inequality on the left is an equality if and only if the inequality on the right is an equality. Similarly,

$$
x = \mathbb{P}(U \geq 1 - x) \leq \mathbb{P}(Y_1 \geq 1/2 - x)^2 \iff \mathbb{P}(Y_1 \geq 1/2 - x) \geq \sqrt{x}
$$

For $x = 1/4$, we get $\mathbb{P}(Y_1 \leq 1/4) \geq 1/2$ and $\mathbb{P}(Y_1 \geq 1/4) \geq 1/2$. If both inequalities were equalities, we would deduce from the left-hand equalities that $\mathbb{P}(Y_1 \in (1/8, 3/8)) = 0$ and this is incompatible with $\mathbb{P}(U \in (1/4, 3/8)) > 0$, so the assumption that $U = Y_1 + Y_2$ must have been wrong.

2. (a) (i) Independence of increments. By the independence of increments of $X$ and $Y$ and by the independence of $X$ and $Y$ we have for all $0 \leq t_0 < t_1 < \ldots < t_n$ that the following random variables are all independent:

$$
X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}} \quad \text{and} \quad Y_{t_0}, Y_{t_1} - Y_{t_0}, \ldots, Y_{t_n} - Y_{t_{n-1}}
$$

Since functions of independent random variables are independent, we can add take linear combinations and deduce independence of

$$
aX_{t_0} + bY_{t_0}, a(X_{t_1} - X_{t_0}) + b(Y_{t_1} - Y_{t_0}), \ldots, a(X_{t_n} - X_{t_{n-1}}) + b(Y_{t_n} - Y_{t_{n-1}}).
$$
(ii) Stationarity of increments. We have that \( X_{t+s} - X_t \) and \( Y_{t+s} - Y_t \) are independent, and also that \( X_s \) and \( Y_s \) are independent. By the stationarity of increments we have that \( X_{t+s} - X_t \sim X_s \) and \( Y_{t+s} - Y_t \sim Y_s \) and so the joint distributions of \((X_{t+s} - X_t, Y_{t+s} - Y_t)\) is the same as the joint distribution of \((X_s, Y_s)\). If we apply the same linear function to the random vectors, these will also have the same distribution, i.e.
\[
a(X_{t+s} - X_t) + b(Y_{t+s} - Y_t) \sim aX_s + bY_s.
\]

(iii) Right-continuity and left limits of paths. Linear combinations of such functions still have these properties.

(b) We calculated the moment generating function of the Gamma(\(\alpha, \beta\)) distribution in Exercise 1 as
\[
\mathbb{E}(\exp\{\gamma A\}) = \int_0^\infty e^{\gamma x} \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta x} \, dx = \frac{\beta^\alpha}{(\beta - \gamma)^\alpha}, \quad \gamma < \beta.
\]
If \( C_1 \sim D_1 \sim \text{Gamma}(\alpha, \sqrt{2\mu}) \), then \( C_s \sim D_s \sim \text{Gamma}(\alpha s, \sqrt{2\mu}) \). Hence
\[
\mathbb{E}(e^{\gamma(C_s-D_s)}) = \mathbb{E}(e^{\gamma C_s})\mathbb{E}(e^{-\gamma D_s}) = \frac{\sqrt{2\mu}^{\alpha s}}{(\sqrt{2\mu} - \gamma)^{\alpha s}(\sqrt{2\mu} + \gamma)^{\alpha s}} = \left(\frac{\mu}{\mu - \frac{1}{\gamma^2}}\right)^{\alpha s}
\]
for all \(-\sqrt{2\mu} < \gamma < \sqrt{2\mu}\).

3. (a) Let \( W_n \sim \text{Binomial}(n, p_n) \) with \( np_n \to \lambda \), then \( W_n \to \text{Poi}(\lambda) \) in distribution as \( n \to \infty \). To prove this, check
\[
\mathbb{E}(s^{W_n}) = \sum_{k=0}^n s^k \binom{n}{k} p_n^k (1-p_n)^{n-k} = \left(1 - \frac{np_n(1-s)}{n}\right)^n \to e^{-\lambda(1-s)},
\]
and this is the probability generating function of \( \text{Poi}(\lambda) \). By the Uniqueness Theorem and by the Continuity Theorem for probability generating functions, \( W_n \) converges in distribution to a \( \text{Poi}(\lambda) \) distribution.

(b) Since \( p_N \) is small, the Poisson limit theorem is appropriate, and since \( N \) is large, it will give a reasonably good approximation. As parameter of the Poisson distribution, \( Np_N \) is appropriate, since \( Np_N \to \lambda \) in the limit theorem for a \( \text{Poi}(\lambda) \) limit.

(c) Denote by \( B_1, \ldots, B_N \) the Bernoulli random variables so that \( B_j = 1 \) if policy holder \( j \) makes a claim. Then \( S_N = B_1 + \ldots + B_N \sim \text{Binomial}(N, p_N) \). We calculate the moment generating function
\[
\mathbb{E}(\exp\{\gamma T_N\}) = \mathbb{E}\left( \exp\left\{ \gamma \sum_{j=1}^{S_N} A_j \right\} \right)
\]
\[
= \sum_{k=0}^N \mathbb{E}\left( \exp\left\{ \gamma \sum_{j=1}^{k} A_j \right\} \right) \binom{N}{k} p_N^k (1-p_N)^{N-k}
\]
\[
= \sum_{k=0}^N \left( \mathbb{E}(e^{\gamma A_1}) \right)^k \binom{N}{k} p_N^k (1-p_N)^{N-k}
\]
\[
= \left( 1 - p_N + p_N \mathbb{E}(e^{\gamma A_1}) \right)^N,
\]
by the binomial theorem, for all \( \gamma \in \mathbb{R} \) for which \( \mathbb{E}(e^{\gamma A_1}) < \infty \).
(d) First, we can continue the argument in (b) to suggest that the actual random variable

\[ \sum_{n=1}^{S_N} A_n \]

has a distribution that is close to the distribution where \( S_N \) is replaced by a Poisson random variable.

Then we formulate the limit statement. Suppose that \( S_N \sim \text{Binomial}(N, p_N) \) for all \( N \in \mathbb{N} \), that \( S_\infty \sim \text{Poi}(\lambda) \) and that \( N p_N \to \lambda \in (0, \infty) \) as \( N \to \infty \). Suppose that \( A_1, A_2, \ldots \) are nonnegative, independent and identically distributed, independent of \( S_N, N \in \mathbb{N} \cup \{\infty\} \). Then

\[ \sum_{n=1}^{S_N} A_n \to \sum_{n=1}^{S_\infty} A_n \quad \text{in distribution, as } N \to \infty. \]

To prove this, consider the moment generating functions

\[ \mathbb{E}(\exp\{\gamma T_N\}) = \left(1 - \frac{N p_N(1 - \mathbb{E}(e^{\gamma A_1}))}{N}\right)^N \to \exp\{-\lambda(1 - \mathbb{E}(e^{\gamma A_1}))\}, \]

and this is the moment generating function of the compound Poisson distribution, which we calculate as follows

\[
\mathbb{E}\left(\exp\left\{\gamma \sum_{j=1}^{S_\infty} A_j\right\}\right) = \sum_{k=0}^{\infty} \mathbb{E}\left(\exp\left\{\gamma \sum_{j=1}^{k} A_j\right\}\right) \frac{\lambda^k}{k!} e^{-\lambda} \\
= \sum_{k=0}^{\infty} (\mathbb{E}(\exp\{\gamma A_1\}))^k \frac{\lambda^k}{k!} e^{-\lambda} \\
= e^{-\lambda} \exp\{\lambda \mathbb{E}(e^{\gamma A_1})\} = \exp\{-\lambda(1 - \mathbb{E}(e^{\gamma A_1}))\}.
\]

4. (a) (i) Note that

\[
\frac{\sum_{k=1}^{n} A_k - n \mathbb{E}(A_1)}{\sqrt{n \text{Var}(A_1)}} = \sum_{k=1}^{n} \frac{A_k - \mu}{\sigma \sqrt{n}} = \sum_{k=1}^{n} Y_{n,k} = V_n.
\]

Thus, the Central Limit Theorem in terms of \( V_n \) states \( V_n \to \text{Normal}(0, 1) \) in distribution as \( n \to \infty \).

(ii) Markov’s inequality \( \mathbb{P}(|X| > y) \leq \mathbb{E}(X^2)/y^2 \) yields

\[
\mathbb{P}(|A_1 - \mu| > \sigma x \sqrt{n}) = \mathbb{P}(|A_1 - \mu| 1_{|A_1 - \mu| > \sigma x \sqrt{n}}) > \sigma x \sqrt{n}) \\
\leq \frac{\mathbb{E}(|A_1 - \mu|^2 1_{|A_1 - \mu| > \sigma x \sqrt{n}})}{\sigma^2 x^2 n}.
\]

Now note that, as \( n \to \infty \),

\[
\mathbb{E}(|A_1 - \mu|^2 1_{|A_1 - \mu| < \sigma x \sqrt{n}}) \to \mathbb{E}(|A_1 - \mu|^2) = \sigma^2,
\]
(by monotone convergence) and so

\[ \gamma_n(x) := \frac{1}{\sigma^2 x^2} \mathbb{E}(|A_1 - \mu|^2 1_{(|A_1 - \mu| \geq \sigma x \sqrt{n})}) \]
\[ = \frac{1}{\sigma^2 x^2} (\sigma^2 - \mathbb{E}(|A_1 - \mu|^2 1_{(|A_1 - \mu| < \sigma x \sqrt{n})})) \rightarrow 0. \]

(iii) For all \( x > 0 \), calculate using (ii)

\[ \mathbb{P}(M_n \leq x) = \mathbb{P}(|Y_{n,1}| \leq x, \ldots, |Y_{n,n}| \leq x) = \mathbb{P}(|Y_{n,1}| \leq x)^n \]
\[ \geq \left( 1 - \frac{\gamma_n(x)}{n} \right)^n \rightarrow e^0 = 1. \]

This implies that \( \mathbb{P}(|M_n| > \varepsilon) = 1 - \mathbb{P}(|M_n| \leq \varepsilon) \rightarrow 0 \) for all \( \varepsilon > 0 \), so \( M_n \rightarrow 0 \) in probability.

(b) (i) At stage \( n \) there are \( r \) red balls and \( s + n - 1 \) black balls in the urn. So

\[ Y_{n,k} \sim \text{Bernoulli} \left( \frac{r}{r + s + n - 1} \right) \Rightarrow W_n \sim \text{Binomial} (n, p_n), \]

where \( p_n = r/(r + s + n - 1) \). Note that \( np_n \rightarrow r \), so that the Poisson limit theorem yields \( W_n \rightarrow \text{Po}(r) \).

(ii) Clearly \( \mathbb{P}(Y_{n,k} = 0) = 1 - p_n = 1 - r/(r + s + n - 1) \rightarrow 1 \), as \( n \rightarrow \infty \).

(iii) Now, as \( n \rightarrow \infty \),

\[ \mathbb{P}(M_n = 0) = \mathbb{P}(Y_{n,1} = 0, \ldots, Y_{n,n} = 0) = (1 - p_n)^n = \left( 1 - \frac{np_n}{n} \right)^n \rightarrow e^{-r}, \]

If \( M_n \rightarrow 0 \), then \( \mathbb{P}(|M_n| > \varepsilon) = 1 - \mathbb{P}(M_n = 0) \rightarrow 0 \) for all \( 0 < \varepsilon < 1 \), and this is incompatible with the limit above. So, \( M_n \not\rightarrow 0 \) in probability.

(c) (i) Define \( S_k^{(n)} = Y_{n,1} + \ldots + Y_{n,k}, k \geq 0, n \geq 1 \).

Donsker’s theorem says in the setting of (a), where \( V_n = S_n^{(n)} \), that \( S_t^{(n)} \rightarrow B_t \) locally uniformly in distribution for a Brownian motion \((B_t)_{t \geq 0}\).

The process version of the Poisson limit theorem says in the setting of (b), where \( W_n = S_n^{(n)} \), that \( S_{|nt|}^{(n)} \rightarrow N_t \) in the Skorohod sense in distribution for a Poisson process \((N_t)_{0 \leq t \leq 1}\) with rate \( r \).

(ii) Clearly, the size of the biggest jump of Brownian motion is 0, and we have \( M_n \rightarrow 0 \) in probability, hence also in distribution.

The number of jumps of \((N_t)_{0 \leq t \leq 1}\) is Poisson distributed with parameter \( r \). The size \( J \) of the biggest jump of \((N_t)_{0 \leq t \leq 1}\) is 1 if there is a jump, with probability \( \mathbb{P}(J = 1) = 1 - e^{-r} \), and \( \mathbb{P}(J = 0) = e^{-r} \) is the probability that there is no jump. This is the limit distribution that we wish to establish.

We have shown that

\[ \mathbb{P}(M_n = 0) \rightarrow e^{-r} = \mathbb{P}(J = 0) \]

and this implies \( \mathbb{P}(M_n = 1) = 1 - \mathbb{P}(M_n = 0) \rightarrow 1 - e^{-r} = \mathbb{P}(J = 1) \), as required.
B.2 Poisson counting measures

1. (a) The distribution of $\Pi$ is specified in terms of the associated counting measure

$$N((a,b]) = \#\Pi \cap (a,b] = \{j \geq 1 : a < T_j \leq b\} = X_b - X_a, \quad 0 \leq a < b.$$ 

Clearly, $N$ satisfies property $\text{hom}(b)$ of a Poisson counting measure: $N((a,b]) = X_b - X_a \sim \text{Poi}(\lambda(b - a))$ by the stationarity (ii) and Poisson (iv) properties for increments of $X$, and we identify the constant intensity function $\lambda(t) = \lambda$, $t \geq 0$.

$N$ also satisfies (a), since for disjoint intervals $(a_j, b_j]$, $j = 1, \ldots, n$, we have $N((a_j, b_j]) = X_{b_j} - X_{a_j}$ increments of $X$ over disjoint time intervals. By property (i) of the Poisson process, these are independent, as required.

(b) (i) Let $0 \leq t_0 < t_1 < \ldots < t_n$. Then $X_{t_j} - X_{t_{j-1}} = N((t_{j-1}, t_j])$. Since the sets $A_j = (t_{j-1}, t_j]$, $j = 1, \ldots, n$, are disjoint, property (a) of the Poisson counting measure yields the independence of the increments.

(ii) Fix $r \geq 0$. For an increment $X_{s+r} - X_s = N((s, s+r])$, property $\text{inhom}(b)$ of the Poisson counting measure yields a Poisson distribution with parameter $p_r(s) = \int_s^{s+r} \lambda(x)dx$. The differentiable function $s \mapsto p_r(s)$ is constant if and only if $0 = p'_r(s) = \lambda(s + r) - \lambda(s)$ for all $s \geq 0$.

Now $(X_t)_{t\geq0}$ has stationary increments if and only if $s \mapsto p_r(s)$ is constant for all $r \geq 0$ if and only if $\lambda(s) = \lambda(r + s)$ for all $r \geq 0$, $s \geq 0$. This is the case if and only if $x \mapsto \lambda(x)$ is constant.

(iii) Clearly $t \mapsto X_t$ is an increasing function, so all left and right limits exist. Denote by $\Pi$ the associated spatial Poisson process, then $\Pi = \{t \geq 0 : \Delta X_t > 0\} = \{t \geq 0 : \Delta X_t = 1\}$. The set $\Pi$ cannot have accumulation points since $\lambda$ is locally integrable, so $\Pi = \{T_j, j \geq 1\}$ and $X_t = N([0, t]) = j$ for $t \in [T_j, T_{j+1})$ is right-continuous at jump times, continuous elsewhere.

(iv) $X_t - X_s = N((s, t]) = \text{Poi}(\int_s^t \lambda(x)dx)$, by property $\text{inhom}(b)$ of the Poisson counting measure.

(v) $\mathbb{P}(T_1 > s) = \mathbb{P}(N([0, s]) = 0) = \exp\{-\int_0^s \lambda(x)dx\}$ for all $s \geq 0$.

(vi) The density of $T_1$ is obtained by differentiating the survival function:

$$f_{T_1}(s) = \lambda(s) \exp\left\{-\int_0^s \lambda(x)dx\right\}.$$ 

To calculate the joint distribution of $(T_1, T_2 - T_1)$, first calculate the joint distribution of $(T_1, T_2)$, from

$$\mathbb{P}(T_1 > s, T_2 > t) = \mathbb{P}(N([0, s]) = 0, N((s, t]) \leq 1)$$

$$= \exp\left\{-\int_0^s \lambda(x)dx\right\} \left(1 + \int_s^t \lambda(x)dx\right) \exp\left\{-\int_s^t \lambda(x)dx\right\}$$

and differentiation, first with respect to $s$ then with respect to $t$

$$f_{T_1,T_2}(s, t) = \lambda(s)\lambda(t) \exp\left\{-\int_0^s \lambda(x)dx\right\}.$$
and the transformation formula for \((T_1, T_2) \mapsto (T_1, T_2 - T_1)\) gives
\[
f_{T_1, T_2 - T_1}(s, r) = \lambda(s)\lambda(s + r) \exp \left\{ - \int_0^{s+r} \lambda(x)dx \right\}
\]
and then
\[
f_{T_2 - T_1 | T_1 = s}(r) = \lambda(s + r) \exp \left\{ - \int_s^{s+r} \lambda(x)dx \right\}
\]
\[
\Rightarrow \quad \mathbb{P}(T_2 - T_1 > r | T_1 = s) = \exp \left\{ - \int_s^{s+r} \lambda(x)dx \right\}
\]
is independent of \(s\) for all \(r \geq 0\) if and only \(x \mapsto \lambda(x)\) is constant, by the argument given in (ii).

2. For simplicity think of \(\Pi\) as a spatial Poisson process in \(\mathbb{R}^3\) with intensity function
\[
\lambda(x, y, z) = \lambda \text{ if } x^2 + y^2 + z^2 \leq 1, \lambda(x, y, z) = 0 \text{ otherwise.}
\]
We check the properties (a) and \(\text{inhom}(b)\) of a spatial Poisson process. Denote by \(N\) and \(N_P\) the associated counting measures of \(\Pi\) and \(P\). Then note that
\[
N_P((a, b] \times (c, d]) = N((a, b] \times (c, d] \times \mathbb{R}) \sim \text{Poi} \left( \int_a^b \int_c^d \int_{\mathbb{R}} \lambda(x, y, z)dzdydx \right)
\]
and we see that the intensity function of \(P\) will have to be
\[
\lambda_P(x, y) = \int_{\mathbb{R}} \lambda(x, y, z)dz = \int_{\sqrt{1 - x^2 - y^2}}^{\sqrt{1 - x^2 - y^2}} \lambda dz = 2\lambda \sqrt{1 - x^2 - y^2}
\]
for \((x, y) \in \mathbb{R}^2\) such that \(x^2 + y^2 \leq 1\), and \(\lambda_P(x, y) = 0\) otherwise.

Property (a) also holds since for disjoint \((a_j, b_j] \times (c_j, d_j], j = 1, \ldots, n,\) the sets \((a_j, b_j] \times (c_j, d_j] \times \mathbb{R}\) are also disjoint, and so independence of \(N_P((a_j, b_j] \times (c_j, d_j]), j = 1, \ldots, n,\) follows from the corresponding property \(N\).

3. (a) First note that \(e^{\Psi(\gamma)} = \mathbb{E}(e^{\gamma X_1})\) implies that \(\mathbb{E}(e^{\gamma X_1/m}) = e^{\Psi(\gamma)/m}\) since stationarity and independence of increments implies \(\mathbb{E}(e^{\gamma X_1/m}) = e^{\Psi(\gamma)}\), then \(\mathbb{E}(e^{\gamma X_q}) = e^{\psi(\gamma)}\), and then the right-continuity of sample paths implies that \(X_q \xrightarrow{a.s.} X_t\) almost surely and hence also in distribution, as \(q \downarrow t\). Therefore, characteristic functions converge and \(\mathbb{E}(e^{\gamma X_q}) = e^{\Phi(\gamma)} \rightarrow e^{t\Phi(\gamma)}\).

Now we use the independence and stationarity of increments to see
\[
\mathbb{E}(\exp\{\gamma X_t\} | \mathcal{F}_s) = \exp\{\gamma X_s\} \mathbb{E}(\exp\{\gamma (X_t - X_s)\}) = \exp\{\gamma X_s\} \exp\{(t - s)\Psi(\gamma)\}.
\]
(b) The argument in (a) applies, with \(\gamma = i\lambda\) and \(\psi\) instead of \(\Phi\) as appropriate. Recall that moment generating functions do not exist for all random variables, but characteristic functions always exist (because \(x \mapsto e^{i\lambda x}\) is bounded).
(c) The following argument can more easily be carried out for moment generating functions, but applies more generally if done for characteristic functions.

Differentiate $E(\exp\{i\lambda X_t\}) = e^{-t\psi(\lambda)}$ with respect to $\lambda$ at $\lambda = 0$ to get $iE(X_t) = -t\psi'(0)$ (see Grimmett-Stirzaker 5.7 for a statement and reference to the proof). The claim follows since $\mu = E(X_1)$ must now be the slope of this linear function.

Now, we use the independence and stationarity of increments to see

$$E(X_t - t\mu|\mathcal{F}_s) = E(X_s + (X_t - X_s) - t\mu|\mathcal{F}_s) = X_s + (t-s)\mu - t\mu = X_s - s\mu.$$ 

(d) Differentiate $E(\exp\{i\lambda X_t\}) = e^{-t\psi(\lambda)}$ twice with respect to $\lambda$ at $\lambda = 0$ to get $-E(X_t^2) = -t(\psi''(0) - t(\psi'(0))^2)$, so $\text{Var}(X_1) = \psi''(0)$.

Now we use the independence and stationarity of increments to see

$$E((X_t - t\mu)^2|\mathcal{F}_s) = E((X_s - s\mu)^2 + 2(X_s - s\mu)(X_t - X_s - (t-s)\mu) + (X_t - X_s - (t-s)\mu)^2|\mathcal{F}_s) = (X_s - s\mu)^2 + 2(X_s - s\mu)E(X_t - X_s - (t-s)\mu) + \text{Var}(X_t - X_s) = (X_s - s\mu)^2 + (t-s)\sigma^2.$$ 

4. (a) Fix $\beta > 0$. Note that the formula reduces to $0 = 0$ for $\gamma = 0$. It is therefore sufficient to show that the $\gamma$-derivatives of both sides coincide. To differentiate the left hand side, note that

$$\frac{\partial}{\partial \gamma}(e^{\gamma x} - 1)\frac{1}{x}e^{-\beta x} = e^{\gamma x}e^{-\beta x} \leq e^{-\beta x},$$

for $\gamma \leq 0$, where $x \mapsto e^{-\beta x}$ is integrable on $[0, \infty)$. Therefore, we may interchange $\gamma$-differentiation and $x$-integration and have to show that for all $\gamma < 0$

$$\int_0^\infty e^{\gamma x}e^{-\beta x}dx = \frac{1}{1 - \gamma/\beta},$$

which clearly is true.

The argument works for $\gamma \leq \gamma_0$ if we choose $e^{-(\beta-\gamma_0)}$ as integrable upper bound. Clearly, for every fixed $\gamma < \beta$, any $\gamma_0 \in (\gamma, \beta)$ will do.

(b) We apply the exponential formula for Poisson point processes and (a) to obtain

$$E\left(\exp\left\{\sum_{s \leq t} \Delta_s \right\}\right) = \exp\left\{t \int_0^\infty (e^{\gamma x} - 1)ax^{-1}e^{-\beta x}dx\right\} = \left(\frac{\beta}{\beta - \gamma}\right)^{\alpha t}.$$ 

We recognise the last expression as the moment generating function of the Gamma distribution with the required density. By the Uniqueness Theorem for moment generating functions, $\sum_{s \leq t} \Delta_s$ has this Gamma distribution.
5. (a) Denote by \(T\) the Poisson arrival processes of jumps. Denote by \(R\) the Poisson arrival processes of jumps. By the previous part, we have two Poisson point processes \(\Delta_s\) and \(\Delta_t\). The intensity function is \(\lambda_t\). The density is \(\lambda_t\). This is a mixture of the jump size distributions of \(X\) and \(Y\). We condition on whether \(X\) and \(Y\) are independent. The distributions are \(\text{Poi}(\lambda_j)\) and \(\text{Poi}(\lambda_j)\) as sum of two independent Poisson variables only depending on \(t_j - t_{j-1}\), (iii) paths are right-continuous with left limits, since it is a random increasing function where for each jump time \(T\), we have (by monotone convergence)

\[
\lim_{t \uparrow T} \sum_{s \leq t} \Delta_s = \sum_{s < t} \Delta_s \quad \text{and} \quad \lim_{t \uparrow T} \sum_{s \leq t} \Delta_s = \sum_{s \leq T} \Delta_s.
\]

5. (a) Denote by \(T_n \sim \text{Gamma}(n, \lambda_X)\) and \(T'_m \sim \text{Gamma}(m, \lambda_Y)\) the jump times of \(X\) and \(Y\). These are independent continuously distributed random variables and so \(\mathbb{P}(T_n = T'_m) = 0\). Therefore, (by subadditivity)

\[
\mathbb{P}(\{T_n, n \geq 1\} \cap \{T'_m, m \geq 1\} \neq \emptyset) \leq \sum_{m \geq 1} \sum_{n \geq 1} \mathbb{P}(T_n = T'_m) = 0.
\]

(b) Denote the Poisson arrival processes of jumps by \(R^X\) and \(R^Y\). Then \(R^X + R^Y\) satisfies the four properties of the Poisson process, since (i) \(R^X_{t_j} - R^X_{t_{j-1}} + R^Y_{t_j} - R^Y_{t_{j-1}}, j = 1, \ldots, n\), are independent as sums of independent random variables, (ii)/(iv) their distributions are \(\text{Poi}(\lambda_X(t_j - t_{j-1}) + \lambda_Y(t_j - t_{j-1}))\) as sum of two independent Poisson variables only depending on \(t_j - t_{j-1}\), (iii) paths are right-continuous with left limits as sums of two such paths.

(c) We condition on whether \(T_1 < T'_1\) or \(T'_1 < T_1\) and get for the first jump size \(J_1^D\) of \(D\)

\[
\mathbb{P}(J_1^D \in A) = \mathbb{P}(T_1 < T'_1)\mathbb{P}(J_1^X \in A | T_1 < T'_1) + \mathbb{P}(T_1 > T'_1)\mathbb{P}(J_1^Y \in A) = \frac{\lambda_X}{\lambda_X + \lambda_Y}\mathbb{P}(J_1^X \in A) + \frac{\lambda_Y}{\lambda_X + \lambda_Y}\mathbb{P}(J_1^Y \in A).
\]

This is a mixture of the jump size distributions of \(X\) and \(Y\). We deduce that the density is

\[
h_D(x) = \frac{\lambda_X}{\lambda_X + \lambda_Y} h_X(x) + \frac{\lambda_Y}{\lambda_X + \lambda_Y} h_Y(-x) = \begin{cases} 
\frac{\lambda_X}{\lambda_X + \lambda_Y} h_X(x) & x > 0 \\
\frac{\lambda_Y}{\lambda_X + \lambda_Y} h_Y(-x) & x < 0 
\end{cases}
\]

(d) This is bookwork, see Lecture 3, Example 18. The intensity function is \(\lambda_X h_X(x), x > 0\).

(e) By the previous part, we have two Poisson point process \(\Delta^X\) and \(\Delta^Y\) in \((0, \infty)\). It is easy to see that \(\Delta^{-Y} = -\Delta_Y\) is a Poisson point process in \((-\infty, 0)\) with intensity function \(\lambda_Y h_Y(-x), x < 0\). It is easy to see that the associated Poisson counting measures on \([0, \infty) \times (0, \infty)\) and \([0, \infty) \times (-\infty, 0)\) together form a Poisson counting measure on \([0, \infty) \times \mathbb{R} \setminus \{0\}\) via

\[
N(A \times B) = N_X(A \times (B \cap (0, \infty))) + N_Y(A \times (B \cap (-\infty, 0))).
\]

The intensity function is \(\lambda_X h_X(x), x > 0\) and \(\lambda_Y h_Y(-x), x < 0\).
(f) Since $(\Delta D_t)_{t \geq 0}$ is a Poisson point process and $D_t = \sum_{s \leq t} \Delta D_s$, $D$ is a compound Poisson process.

(g) For every real-valued compound Poisson process $C$ we can define the processes $X$ and $Y$ of positive and negative jumps. Since the associated processes $(\Delta X_t)_{t \geq 0}$ and $(\Delta Y_t)_{t \geq 0}$ inherit the properties of Poisson point processes (via their Poisson counting measures), this provides the required decomposition into two independent increasing compound Poisson processes. It is unique because any other decomposition must have more jumps, which must happen at the same time and cancel each other, but by (a), this is incompatible with independence.
B.3 Construction of Lévy processes

1. (a) If \( \kappa \in (-1, \infty) \), then

\[
\int_0^\infty g(x)dx = \int_0^\infty x^\kappa e^{-x}dx = \Gamma(\kappa + 1) < \infty.
\]

The Poisson point process is hence of the form of Example 18 and so \((C_t)_{t \geq 0}\) is a compound Poisson process with intensity \(\Gamma(\kappa + 1)\) and Gamma\((\kappa + 1, 1)\) jump distribution with density

\[
h(x) = \frac{1}{\Gamma(\kappa + 1)} x^\kappa e^{-x}, \quad x > 0.
\]

(b) The counting measures associated to \((\Delta_t)_{t \geq 0}\) and \((\Delta^{(n)}_t)_{t \geq 0}\) are

\[
N((a, b] \times (c, d]) = \#\{t \in (a, b] : \Delta_t \in (c, d]\}
\sim \text{Poi}\left((b - a) \int_c^d g(x)dx\right), \quad 0 \leq a < b, 0 < c < d,
\]

\[
N_n((a, b] \times (c, d]) = N((a, b] \times ((c, d] \cap (1/n, \infty)))
\sim \text{Poi}\left((b - a) \int_c^d g(x)1_{\{x > 1/n\}}dx\right), \quad 0 \leq a < b, 0 < c < d.
\]

\(N_n\) inherits the properties of a Poisson counting measure from \(N\). We read off the intensity function \(g_n(x) = g(x), x > 1/n, g_n(x) = 0, x \leq 1/n\). The argument of (a) shows that \(C_t^{(n)}\) is a compound Poisson process.

(c) \(C_t^{(n)}\) increases as \(n \to \infty\). We can study the limit of moment generating functions, whether or not the limit is finite. We get, as \(n \to \infty\),

\[
\mathbb{E}(e^{\gamma C_t^{(n)}}) = \exp\left\{ \int_{1/n}^\infty (e^{\gamma x} - 1)g(x)dx \right\} \downarrow \exp\left\{ \int_0^\infty (e^{\gamma x} - 1)g(x)dx \right\}
\]

and because for \(\gamma < 0\)

\[
\int_0^\infty (e^{\gamma x} - 1)g(x)dx < \infty \iff \int_0^\infty (1 \wedge x)g(x)dx < \infty,
\]

and by Lemma 21, we need to investigate the right hand condition. We check that

\[
\int_1^\infty g(x)dx < \infty, \quad \text{and} \quad \int_0^1 xg(x)dx < \infty \iff \kappa + 1 > -1,
\]

as required.

(d) We can write

\[
C_s - C_s^{(n)} = \sum_{r \leq s} \Delta_r 1_{\{\Delta_r \leq 1/n\}} \leq \sum_{r \leq t} \Delta_r 1_{\{\Delta_r \leq 1/n\}} = C_t - C_t^{(n)},
\]
and putting a supremum over \( s \leq t \) on the left hand side, we get the required estimate (as an equality because we can take \( s = t \) on the left. Now we showed in (c) that \( C_t^{(n)} \to C_t \) a.s., and so we deduce here that

\[
\sup_{s \leq t} |C_s^{(n)} - C_s| \to 0 \quad \text{as } n \to \infty,
\]

i.e. that the convergence is locally uniform.

(e) By Proposition 40(ii), we have for \( m \leq n \)

\[
\mathbb{E}(|C_t^{(n)} - \mathbb{E}(C_t^{(n)}) - (C_t^{(m)} - \mathbb{E}(C_t^{(m)}))|^2) = \text{Var}(C_t^{(n)} - C_t^{(m)})
= \int_{1/n}^{1/m} x^2 g(x) \, dx,
\]

and this decreases to zero as \( n \geq m \to \infty \) if and only if \( \int_0^1 x^2 g(x) \, dx < \infty \), i.e. \( \kappa > -3 \). In this case, \( (C_t^{(n)} - \mathbb{E}(C_t^{(n)}))_{n \geq 1} \) is a Cauchy sequence that converges by completeness of \( \mathbb{R} \) (and the associated \( L^2 \) space of \( \mathbb{R} \)-valued random variables).

The limiting process includes all jumps \((\Delta_s)_{s \leq t}\) (intuitively, a more formal argument uses uniform convergence that preserves the jumps, see Theorem 42), and by (c), these are not summable for \( \kappa \in (-3, -2] \). By Proposition 36, the limiting process has unbounded variation.

2. (a) Just note that for subordinators \( 0 \leq X_t < \infty \) a.s., and this implies that \( 1 \geq e^{-\mu X_t} > 0 \) a.s. and then also \( 1 \geq \mathbb{E}(e^{-\mu X_t}) > 0 \) as required. Therefore, \( \Phi_t \) is well-defined.

(b) This follows as in A.2.3, first for rational \( t \geq 0 \) and then, by right-continuity of paths and since a.s. convergence implies convergence in distribution, hence of moment generating functions. The scaling relation for fixed \( t \) translates to

\[
\Phi_{t/c}(e^{1/\alpha} \mu) = -\ln(\mathbb{E}(\exp(-\mu c^{1/\alpha} X_{t/c}))) = -\ln(\mathbb{E}(e^{-\mu X_t})) = \Phi_t(\mu).
\]

and therefore, for \( t = 1, c = \mu^{-\alpha} \), as required

\[
\mu^\alpha \Phi(1) = \Phi_{1/c}(1) = \Phi(\mu).
\]

(c) Clearly \( \mu \mapsto e^{-\mu X_t} \) is a.s. decreasing and so is hence \( \mu \mapsto \mathbb{E}(e^{-\mu X_t}) \), strictly decreasing if \( X_t > 0 \) with positive probability. Now, \( \Phi(\mu) = \Phi(1) \mu^\alpha \) is clearly differentiable for \( \mu > 0 \), and so

\[
\frac{\partial}{\partial \mu} \mathbb{E}(e^{-\mu X_t}) = \frac{\partial}{\partial \mu} e^{-t\Phi(1) \mu^\alpha} = -t \Phi(1) \alpha \mu^{\alpha-1} e^{-t\Phi(1) \mu^\alpha}
\]

and this is negative only for \( \alpha > 0 \) (or \( \alpha = 0 \) but then \( \mu \mapsto \mathbb{E}(e^{-\mu X_t}) \) is constant). To show that also \( \alpha \leq 1 \) note that \( \mu \mapsto e^{-\mu X_t} \) is also a.s. convex, and hence so is \( \mu \mapsto \mathbb{E}(e^{-\mu X_t}) \). Now, \( \Phi(\mu) \) is also twice differentiable so that

\[
\frac{\partial^2}{\partial \mu^2} \mathbb{E}(e^{-\mu X_t}) = t \Phi(1) \alpha \mu^{\alpha-2} e^{-t\Phi(1) \mu^\alpha} (t \Phi(1) \alpha \mu^\alpha - (\alpha - 1)),
\]

and this is nonnegative for all \( \mu > 0 \) if and only if \( \alpha \leq 1 \).
(d) Note that, (by monotone convergence), as $\mu \downarrow 0$, 

\[ t\Phi(1)\alpha^{\alpha-1}e^{-t\Phi(1)\mu^\alpha} = \mathbb{E}(X_t e^{-\mu X_t}) \uparrow \mathbb{E}(X_t), \]

where the left-hand side increases to $\infty$ for $\alpha \in (0, 1)$.

(e) Note that $\Phi(0) = 0$ implies that the equation holds for $\mu = 0$ no matter what $g$ is. Now differentiate both sides with respect to $\mu$ to get

\[ \Phi(1)\alpha^{\alpha-1} = \int_0^\infty e^{-\mu x} g(x) dx. \]

Remember that the density of the Gamma($1-\alpha, \mu$) distribution is $f(x) = (\Gamma(1-\alpha))^{-1} \mu^{1-\alpha} x^{-\alpha} e^{-\mu x}$. Therefore, we can (and have to, by the Uniqueness Theorem for moment generating functions) take

\[ g(x) = \frac{\Phi(1)\alpha}{\Gamma(1-\alpha)} x^{-\alpha-1}, \quad x > 0. \]

(f) For $\alpha \in (0, 1)$, the Construction Theorem for subordinators (Theorem 8) shows that we can construct the stable subordinator from a Poisson point process with intensity function $g$ as specified in (e). Note that $g$ satisfies the integrability condition

\[ \int_0^\infty (1 \wedge x) g(x) dx < \infty \]

since $x^{-\alpha-1}$ is integrable at $x = \infty$ and $x^{-\alpha}$ is integrable at $x = 0$.

For $\alpha = 1$ note that $\Phi_t(\mu) = \Phi(1)t\mu$. The associated subordinator is the deterministic drift $X_t = \Phi(1)t$.

3. (a) Just note that

\[ c^{1/\alpha}Z_{t/c} = c^{1/\alpha}X_{t/c} - c^{1/\alpha}Y_{t/c} \sim X_t - Y_t = Z_t \]

for fixed $t$, and that, as processes in $t \geq 0$, both the left-hand side and the right-hand side are Lévy processes. Therefore, the distributions as processes coincide.

(b) $H \sim -H$ implies

\[ \mathbb{E}(\cos(\lambda H)) + i\mathbb{E}(\sin(\lambda H)) = \mathbb{E}(e^{i\lambda H}) = \mathbb{E}(e^{-i\lambda H}) = \mathbb{E}(\cos(\lambda H)) - i\mathbb{E}(\sin(\lambda H)) \]

and so the imaginary part $\mathbb{E}(\sin(\lambda H))$ must vanish for all $\lambda \in \mathbb{R}$.

(c) Clearly $Z_t = X_t - Y_t \sim Y_t - X_t = -Z_t$, so $Z_t$ has a symmetric distribution. By (b), its characteristic function $\varphi_t(\lambda) = \mathbb{E}(e^{i\lambda Z_t})$ is real-valued. By the hint, we may assume that $\varphi_t$ is continuous, and since $Z_t$ is infinitely divisible, that $\varphi_t(\lambda) \neq 0$, so it must stay positive everywhere (note that $\varphi(0) = 1$). Define

\[ \psi_t(\lambda) = -\ln(\varphi_t(\lambda)), \quad \psi(\lambda) = \psi_1(\lambda), \quad \lambda \in \mathbb{R}. \]
By A.2.3.(b), we have \( \psi_t(\lambda) = t\psi(\lambda) \). The scaling relation implies

\[
\psi_t/c(\lambda) = -\ln(\mathbb{E}(e^{i\lambda c^{1/\alpha} Z_t})) = -\ln(\mathbb{E}(e^{i\lambda Z_t})) = \psi_t(\lambda),
\]

and as in 2.(b), this implies \( \psi(\lambda) = \psi(1)\lambda^\alpha \) for all \( \lambda \geq 0 \). For \( \lambda < 0 \) note that \( \psi(\lambda) = -\ln(\mathbb{E}(e^{i\lambda Z_t})) = -\ln(\mathbb{E}(e^{-i\lambda Z_t})) = \psi(-\lambda) \),

so we have \( \psi(\lambda) = \psi(1)|\lambda|^\alpha \).

(d) Before we start, note that the integral defining \( \tilde{\psi}(\lambda) \) converges for \( \alpha \in (0, 2) \) since the integrand behaves like \( x^{1-\alpha} \) at \( x = 0 \) and like \( x^{-\alpha-1} \) at \( |x| = \infty \). We then check, by change of variables \( y = c^{1/\alpha}x \) (hence \( x^{-1}dx = y^{-1}dy \), that

\[
\tilde{\psi}(\lambda c^{1/\alpha}) = \int_{-\infty}^{\infty} (\cos(\lambda c^{1/\alpha} x) - 1)b|x|^{-\alpha-1}dx
\]

\[
= \int_{-\infty}^{\infty} (\cos(\lambda y) - 1)bc|y|^{-\alpha-1}dy = c\tilde{\psi}(\lambda).
\]

The argument of (c) shows that this implies \( \tilde{\lambda} = b|\lambda|^\alpha \) for some \( b \geq 0 \) – the argument did not depend on \( \alpha \in (0, 1) \).

(e) Let \( R \) be a symmetric stable process of index \( \alpha \). In (d) we expressed the characteristic exponent in terms of the Lévy density \( g(x) = |x|^{-\alpha-1} \). Therefore, \( (R_t)_{t \geq 0} \) can be constructed from a Poisson point process of jumps with this density. We see from the criterion \( \int_{-\infty}^{\infty} (1 \wedge |x|)g(x)dx < \infty \) that jumps are absolutely summable if and only if \( \alpha \in (0, 1) \). In that case, we expressed \( R \) as the difference of two subordinators in (a), so \( R \) has indeed bounded variation. If \( \alpha \in [1, 2) \), jumps are not absolutely summable, so variation is unbounded, by Proposition 36. If \( \alpha = 2 \), then \( R \) is a multiple of Brownian motion, and it was shown in the lectures, that Brownian motion has also unbounded variation. We differentiate \( \psi(\lambda) \) in \( \lambda = 0 \) to see that \( \mathbb{E}(X_t) = t\psi'(0) < \infty \) if and only if \( \alpha \in (1, 2] \), and that \( \text{Var}(X_t) = t\psi''(0) < \infty \) if and only if \( \alpha = 2 \).
B.4 Simulation

1. Note that the definition $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) > u\}$ implies

$$F(t) > u \Rightarrow F^{-1}(u) \leq t \Rightarrow \forall \varepsilon > 0 F(t + \varepsilon > u) \Rightarrow F(t) \geq u$$

Therefore for all $t \in \mathbb{R}$

$$F(t) = \mathbb{P}(F(t) > U) \leq \mathbb{P}(F^{-1}(U) \leq t) \leq \mathbb{P}(F(t) \geq U) = F(t).$$

2. (a) Independent $A \sim \text{Gamma}(a,1)$ and $B \sim \text{Gamma}(b,1)$ have joint density

$$f_{A,B}(x,y) = \frac{x^{a-1}e^{-x} y^{b-1}e^{-y}}{\Gamma(a) \Gamma(b)}$$

The transformation $(R, S) = T(A, B) = (A/(A + B), A + B)$ is bijective $T : (0, \infty)^2 \to (0,1) \times (0, \infty)$ with inverse transformation $(A, B) = T^{-1}(R, S) = (SR, S(1 - R))$ that has the Jacobian

$$J(r, s) = \begin{pmatrix} s & r \\ -s & 1 - r \end{pmatrix} \Rightarrow |\det(J(r, s))| = s$$

and so the transformation formula yields

$$f_{R,S}(r, s) = |\det(J(r, s))| f_{A,B}(T^{-1}(r, s)) = s \frac{(sr)^{a-1}e^{-sr} (s(1 - r))^{b-1}e^{-s(1-r)}}{\Gamma(a) \Gamma(b)}$$

$$= \frac{\Gamma(a + b)}{\Gamma(a) \Gamma(b)} r^{a-1}(1 - r)^{b-1} s^{a+b-1}e^{-s},$$

as required.

Vice versa, for $c = a + b$ and $p = a/(a + b)$, we recognise $T^{-1}(R, S) = (A, B)$, which has joint distribution

$$f_{A,B}(x, y) = \frac{x^{a-1}e^{-x} y^{b-1}e^{-y}}{\Gamma(a) \Gamma(b)} = \frac{x^{cp-1}e^{-x} y^{(1-p)-1}e^{-y}}{\Gamma(cp) \Gamma(c(1-p))}.$$ 

and so any random variable $(\tilde{R}, \tilde{S})$ with joint distribution as $(R, S)$ will be such that $T^{-1}(\tilde{R}, \tilde{S}) \sim T^{-1}(R, S) = (A, B)$.

(b) $\mathbb{P}(X \leq x) = \mathbb{P}(U^{1/a} \leq x) = \mathbb{P}(U \leq x^a) = x^a, x \in (0,1)$ and so $f_X(x) = ax^{a-1}, x \in (0,1)$. We recognise $X \sim \text{Beta}(a,1)$.

(c) We calculate

$$\mathbb{P}\left(\frac{Y}{Y + Z} \leq t, Y + Z \leq 1\right) = \mathbb{P}(\frac{Y(1-t)}{t} \leq Z \leq 1-Y)$$

$$= \int_0^t \int_{y/(1-t)/t}^{1-y} ay^{a-1}(1-a)z^{-a}dzdy$$

$$= \int_0^t ay^{a-1} ((1-y)^{1-a} - y^{1-a}(1-t)^{1-a}t^{a-1}) dy.$$
We differentiate with respect to $t$ to get
\[
f_{W|Y+Z \leq 1}(t) = \frac{a t ((1-a)(1-t)^{-a} - (a-1)(1-t)^{-a-2})}{\mathbb{P}(X+Y \leq 1)}
\]
\[
= \frac{a(1-a)(1-t)^{-a}((1-t) + t)}{\mathbb{P}(X+Y \leq 1)}
\]
and we recognise the density of Beta($a, 1-a$), up to the normalisation constant, but we have calculated a conditional density which integrates to 1, so the normalisation constant must be the one of Beta($a, 1-a$).

(d) Given $Y + Z \leq 1$, $W$ is Beta($a, 1-a$)-distributed. Since $T$ is independent of $(Y, Z, W)$, its conditional distribution given $Y + Z \leq 1$ is still Gamma(1, 1) and it is conditionally independent of $W$ given $Y + Z \leq 1$. Therefore $\mathbb{P}(TW \leq h|Y + Z \leq 1) = \mathbb{P}(SR \leq h)$, and we can apply (a) for $c = 1$ and $p = a$ to deduce that, $SR \sim$ Gamma($a, 1$), i.e. the conditional distribution of $WT$ given $Y + Z \leq 1$ is Gamma($a, 1$).

(e) This procedure generates a Gamma($a, 1$) random variable. Specifically, the conditioning on $Y + Z \leq 1$ is realised by repeated trials until $Y + Z \leq 1$, see Lemma 57. The procedure is easily implemented and gives a more efficient way of simulating Gamma random variables from uniform random variables than inverting the distribution function of the Gamma distribution numerically.

3. (a) From 2.(a) we take that we obtain $A/(A+B) \sim$ Beta($a, b$) for independent $A \sim$ Gamma($a, 1$) and $B \sim$ Gamma($b, 1$). Johnk’s procedure works for $a \in (0, 1)$. To generate Gamma variables for higher parameters, we can write $a = [a] + \{a\}$ for integer part and fractional part and then represent
\[
A = \sum_{k=1}^{[a]} E_k + A_0
\]
where $(E_k)_{1 \leq k \leq [a]}$ is a sequence of independent Exp(1) random variables and $A_0 \sim$ Gamma($\{a\}, 1$). To summarise, the following procedure generates a Beta($a, b$) random variable:

1.-5. Run Johnk’s Gamma generator for parameter $\{a\}$. Set $A_0 = TY/(Y+Z)$.

6.-10. Independently of 1.-5., run Johnk’s Gamma generator for parameter $\{b\}$. Set $B_0 = TY/(Y+Z)$.

11. Generate independent $U_1, \ldots, U_{[a]+[b]} \sim$ Unif(0, 1) and set
\[
A = A_0 - \ln \left( \prod_{k=1}^{[a]} U_k \right) \quad \text{and} \quad B = B_0 - \ln \left( \prod_{k=[a]+1}^{[a]+[b]} U_k \right).
\]

12. Return the number $A/(A+B)$.

(b) The procedure generates a stochastic process successively at refining lattices of dyadic times. The key step (for $n = 0$ and then inductively for $n \geq 1$) is to take a $2^{-n}$-increment $Y_{k,n} = X_{k2^{-n}} - X_{(k-1)2^{-n}} \sim$ Gamma($2^{-n}, 1$) and a
\( B_{k,n} \sim \text{Beta}(a_n, b_n) \) random variable to split \( Y_{k,n} \) into two increments \( B_{k,n}Y_{k,n} \) and \( (1 - B_{k,n})Y_{k,n} \). For 2.(a) to apply, we need \( a_n + b_n = 2^{-n} \), i.e. \( a_n = 2^{-n}p \) and \( b_n = 2^{-n}(1 - p) \). In order to get identically Gamma\((2^{-n} - 1, 1)\) increments, further \( p = 1/2 \). This now yields stationary independent increments.

This is the only way to achieve Gamma distributions for all \( X_{k,2^{-n}} \) (at least in the framework of 2.(a), but in fact, in general) for Beta parameters not depending on \( k \). If we were to use \( B_{k,n} \sim \text{Beta}(a_{k,n}, b_{k,n}) \) for parameters that may depend on \( k \), then we get more general processes with independent Gamma increments that are not stationary, as soon as we have the consistency condition that \( a_{k,n} + b_{k,n} \) is the parameter of \( Y_{k,n} \).

(c) Johnk’s Gamma generator is more efficient than the inverse distribution function computation. The method is less liable to accumulating errors since time 1 is most accurate and errors only accumulate along the dyadic expansions, i.e. with local rather than global impact. Furthermore, we get an iterative procedure for which we do not have to fix the time lag \( \delta \) in advance, but can continue to fill in extra points until a satisfactory result is obtained.

4. (a) Since \( G_t \sim \text{Gamma}(\alpha_+ t, \beta_+) \) and \( H_t \sim \text{Gamma}(\alpha_- t, \beta_-) \), we have

\[
\mathbb{E}(at + G_t - H_t) = at + \frac{\alpha_+ t}{\beta_+} - \frac{\alpha_- t}{\beta_-} = 0 \iff a = \frac{\alpha_-}{\beta_-} - \frac{\alpha_+}{\beta_+}.
\]

(b) 

- Denote \( F_\delta(x) = \mathbb{P}(V_\delta \leq x) \).
  1. Set \( S_0 = 0 \) and \( n = 1 \).
  2. Generate \( U_n \sim \text{Unif}(0, 1) \).
  3. Set \( S_n = S_{n-1} + F_\delta^{-1}(U_n) \). If enough steps have been performed, go to 4., otherwise increase \( n \) by 1 and go to 2.
  4. Return \( (S_n)_{n \geq 0} \) as simulation of \( (V_\delta n)_{n \geq 0} \).

- Denote \( F(x; \alpha, \beta) = \mathbb{P}(G \leq x) \) for \( G \sim \text{Gamma}(\alpha, \beta) \).
  1. Set \( S_0 = 0 \) and \( n = 1 \).
  2. Generate two independent random numbers \( U_{2n-1} \sim \text{Unif}(0, 1) \) and \( U_{2n} \sim \text{Unif}(0, 1) \).
  3. Set \( S_n = S_{n-1} + a\delta + F^{-1}(U_{2n-1}; \alpha_+ \delta, \beta_+) - F^{-1}(U_{2n}; \alpha_- \delta, \beta_-) \). If enough steps have been performed, go to 4., otherwise increase \( n \) by 1 and go to 2.
  4. Return \( (S_n)_{n \geq 0} \) as simulation of \( (V_\delta n)_{n \geq 0} \).

- Fix \( t = 1 \), iterate for further time units if needed. Denote \( G(x; a, b) = \mathbb{P}(B \leq x) \) for \( B \sim \text{Beta}(a, b) \).
  1. Set \( V_0 = 0 \) and \( n = 0 \).
  2. Generate 2 independent random numbers \( U_1 \sim \text{Unif}(0, 1) \) and \( U_2 \sim \text{Unif}(0, 1) \).
  3. Set \( P_1 = F^{-1}(U_1; \alpha_+, \beta_+), N_1 = F^{-1}(U_2; \alpha_-, \beta_-) \) and \( V_1 = a + P_1 - N_1 \).
  4. Generate \( 2^n \) independent random numbers \( U_{2^n+k} \sim \text{Unif}(0, 1), k = 1, \ldots, 2^n \).
  5. Set \( B_{n,k} = G^{-1}(U_{2^n+k}; 2^{-n-1} \alpha_+, 2^{-n-1} \alpha_+) \), \( k = 1, \ldots, 2^{n-1} \) and \( C_{n,k} = G^{-1}(U_{2^n+k}; 2^{-n-1} \alpha_-, 2^{-n-1} \alpha_-), k = 1, \ldots, 2^{n-1} \).
6. Set \( P_{(2k-1)2^{-n}} = B_{n,k}P_{(2k-2)2^{-n}} + (1 - B_{n,k})P_{(2k)2^{-n}} \), \( N_{(2k-1)2^{-n}} = C_{n,k}N_{(2k-2)2^{-n}} + (1 - C_{n,k})N_{(2k)2^{-n}} \) and \( V_{(2k-1)2^{-n}} = (2k - 1)2^{-n}a + P_{(2k-1)2^{-n}} - N_{(2k-1)2^{-n}} \) for \( k = 1, \ldots, 2^{n-1} \). If the resolution is fine enough, go to 7., otherwise increase \( n \) by 1 and go to 4.

7. Return \( (V_{2^{-n}})_{k=1,\ldots,2^n} \).

Instead of \( F^{-1} \) and \( G^{-1} \), one can use Johnk’s Gamma generator of A.3.2. and the associated Beta generator of A.3.3.

- Denote \( H(x; \beta) = \int_{\varepsilon}^{x} y^{-1}e^{-\beta y}dy / \int_{\varepsilon}^{\infty} y^{-1}e^{-\beta y}dy \). Also denote

\[
\lambda = \alpha_{+} \int_{\varepsilon}^{\infty} y^{-1}e^{-\beta_{+} y}dy + \alpha_{-} \int_{\varepsilon}^{\infty} y^{-1}e^{-\beta_{-} y}dy
\]

and \( p = \lambda^{-1} \alpha_{+} \int_{\varepsilon}^{\infty} y^{-1}e^{-\beta_{+} y}dy \).

1. Set \( V_0 = 0, T_0 = 0 \) and \( n = 1 \).

2. Generate three independent random numbers \( U_{3n-2} \sim \text{Unif}(0, 1) \) and \( U_{3n-1} \sim \text{Unif}(0, 1) \) and \( U_{3n} \sim \text{Unif}(0, 1) \).

3. Set \( Z_n = -\ln(U_{3n})/\lambda \).

4. If \( U_{3n-1} > p \), let \( J_n = -H^{-1}(U_{3n}; \beta_{-}) \), otherwise let \( J_n = H^{-1}(U_{3n}; \beta_{+}) \).

5. Set \( T_n = T_{n-1} + Z_n \) and \( V_{T_n} = V_{T_{n-1}} + aZ_n + J_n \). If \( T_n \) is big enough, go to 6., otherwise increase \( n \) by 1 and go to 2.

6. Return \( (V_{T_n})_{n \geq 0} \).

(c) Below are 9 simulations for \( \alpha_{+} \in \{1, 10, 100\} \) (rows) and \( \alpha_{-} \in \{10, 100, 1000\} \) (columns). Note the big positive jumps for \( \alpha_{+} = 1 \), the cases \( \alpha_{+} = \alpha_{-} \) with \( a = 0 \) and convergence to Brownian motion from top left to bottom right. The code is a similar to the symmetric case and is available on the course website.
B.5 Financial models

1. (a) \( W_0 = T_0 + U_0 + V_0, W_1(\omega_1) = T_0 e^\delta + U_0 B_1^{\text{up}} + V_0 C_1^{\text{up}}, W_1(\omega_2) = T_0 e^\delta + U_0 B_1^{\text{down}} + V_0 C_1^{\text{down}}, W_1(\omega_3) = T_0 e^\delta + U_0 B_1^{\text{down}} + V_0 C_1^{\text{down}}, \) \( W_1(\omega_4) = T_0 e^\delta + U_0 B_1^{\text{up}} + V_0 C_1^{\text{up}}. \)

(b) By general reasoning, there is arbitrage if one asset is uniformly better than another asset. In particular:

- If \( B_1(\omega_1) \leq A_1, \) then \((1, -1, 0)\) is an arbitrage portfolio, since \( W_0 = 0 \) and \( W_1 \geq 0 \) with \( W_1(\omega_3) = W_1(\omega_4) > 0. \)
- If \( A_1 \leq B_1(\omega_4), \) then \((-1, 1, 0)\) is an arbitrage portfolio, since \( W_0 = 0 \) and \( W_1 \geq 0 \) with \( W_1(\omega_1) = W_1(\omega_2) > 0. \)
- Similarly \((1, 0, -1)\) or \((-1, 0, 1)\) are arbitrage portfolios if \( C_1(\omega_1) \leq A_1 \) or \( A_1 \leq C_1(\omega_4). \)

These can also be deduced from the standard two-asset binary model \((A, B)\) or \((A, C)\). Now let \( B_1^{\text{up}} > A_1 > B_1^{\text{down}} \) and \( C_1^{\text{up}} > A_1 > C_1^{\text{down}}. \) Since the model \((A, B)\) has no arbitrage, there is no arbitrage portfolio of the form \((T_0, U_0, 0). \)
Assume that \((T_0, U_0, 1)\) is an arbitrage portfolio. Then \(0 = W_0 = T_0 + U_0 + 1, W_1(\omega_1) > W_1(\omega_2) \geq 0 \) and \( W_1(\omega_3) > W_1(\omega_4) \geq 0. \)

- If \( U_0 \geq 0, \) then we have \(0 \leq W_1(\omega_4) = T_1 A_1 + U_0 B_1^{\text{down}} + C_1^{\text{down}} < (T_1 + U_0 + 1)A_1 = 0, \) which is a contradiction.
- If \( U_0 \leq 0, \) then we have \(0 \leq W_1(\omega_2) = T_1 A_1 + U_0 B_1^{\text{up}} + C_1^{\text{down}} < (T_1 + U_0 + 1)A_1 = 0, \) which is a contradiction.

Similarly, now assume that \((T_0, U_0, -1)\) is an arbitrage portfolio, then \(0 = W_0 = T_0 + U_0 - 1, W_1(\omega_2) > W_1(\omega_1) \geq 0 \) and \( W_1(\omega_3) > W_1(\omega_4) \geq 0. \)

- If \( U_0 \geq 0, \) then we have \(0 \leq W_1(\omega_3) = T_1 A_1 + U_0 B_1^{\text{down}} - C_1^{\text{up}} < (T_1 + U_0 - 1)A_1 = 0, \) which is a contradiction.
- If \( U_0 \leq 0, \) then we have \(0 \leq W_1(\omega_1) = T_1 A_1 + U_0 B_1^{\text{up}} - C_1^{\text{up}} < (T_1 + U_0 + 1)A_1 = 0, \) which is a contradiction.

So there is no arbitrage portfolio.

(c) The contingent claim \( W_1(\omega_1) = 1, W_1(\omega_2) = W_1(\omega_3) = W_1(\omega_4) = 0 \) cannot be hedged, since we would require

\[
0 = T_0 A_1 + U_0 B_1^{\text{up}} + V_0 C_1^{\text{down}} \\
= T_0 A_1 + U_0 B_1^{\text{down}} + V_0 C_1^{\text{down}} = T_0 A_1 + U_0 B_1^{\text{down}} + V_0 C_1^{\text{up}},
\]

for \( \omega_2, \omega_3, \omega_1, \) and these imply \( T_0 = U_0 = V_0 = 0, \) but then the fourth equation \(1 = T_0 A_1 + U_0 B_1^{\text{up}} + V_0 C_1^{\text{up}}\) fails.

(d) Since the contingent claim does not change as \( C_1 \) varies, we should consider portfolios of the form \((T_0, U_0, 0). \) Since the model \((A, B)\) with scenarios “up” and “down” is complete, the contingent claim \( \tilde{W}_1(\text{up}) = W_1(\omega_1), \tilde{W}_1(\text{up}) = W_1(\omega_4) \) can be hedged. Specifically,

\[
\tilde{W}_1(\text{up}) = T_0 A_1 + U_0 B_1^{\text{up}} \quad \text{and} \quad \tilde{W}_1(\text{down}) = T_0 A_1 + U_0 B_1^{\text{down}}
\]
has solution
\[
T_0 = \frac{\tilde{W}_1(\text{down})B_{1}^{\text{up}} - \tilde{W}_1(\text{up})B_{1}^{\text{down}}}{A_1(B_{1}^{\text{up}} - B_{1}^{\text{down}})} \quad \text{and} \quad U_0 = \frac{\tilde{W}_1(\text{up}) - \tilde{W}_1(\text{down})}{B_{1}^{\text{up}} - B_{1}^{\text{down}}},
\]
and so we read off from
\[
\tilde{W}_0 = T_0 + U_0 = \frac{A_1 - B_{1}^{\text{down}}}{A_1(B_{1}^{\text{up}} - B_{1}^{\text{down}})}\tilde{W}_1(\text{up}) + \frac{B_{1}^{\text{up}} - A_1}{A_1(B_{1}^{\text{up}} - B_{1}^{\text{down}})}\tilde{W}_1(\text{down}) \quad (1)
\]
that
\[
q_B = \frac{A_1 - B_{1}^{\text{down}}}{B_{1}^{\text{up}} - B_{1}^{\text{down}}} \in (0, 1).
\]

The martingale property is equation (1) for the contingent claim \(\tilde{W}_1(\text{down}) = B_{1}^{\text{down}}\) and \(\tilde{W}_1(\text{up}) = B_{1}^{\text{up}}\). The martingale probability \(q_B\) is unique and does not depend on \(\tilde{W}_1\).

(e) By symmetry, contingent claims of the form \(W_1(\omega_1) = W_1(\omega_3)\), \(W_1(\omega_2) = W_1(\omega_4)\) can be hedged and priced as \(e^{-\delta}\mathbb{E}(W_1)\), where
\[
q_C = \mathbb{P}(C_1 = C_1^{\text{up}}) = \frac{A_1 - C_1^{\text{down}}}{C_1^{\text{up}} - C_1^{\text{down}}} \in (0, 1).
\]

The process \(e^{-\delta}C_t\), \(t = 0, 1\), is a martingale under these probabilities.

(f) In order for both \(e^{-\delta}B_t\) and \(e^{-\delta}C_t\) to be martingales, we need
\[
q_B = \mathbb{P}(B_1 = B_{1}^{\text{up}}, C_1 = C_1^{\text{up}}) + \mathbb{P}(B_1 = B_{1}^{\text{up}}, C_1 = C_1^{\text{down}}) = p_1 + p_2
\]
and
\[
q_C = \mathbb{P}(B_1 = B_{1}^{\text{up}}, C_1 = C_1^{\text{up}}) + \mathbb{P}(B_1 = B_{1}^{\text{down}}, B_1 = B_{1}^{\text{down}}) = p_1 + p_3.
\]
Together with the normalisation condition \(p_1 + p_2 + p_3 + p_4\), we have three equations (of rank three) for four unknowns, so there is a one-dimensional solution space.

(g) The range of arbitrage-free prices \(W_0 = e^{-\delta}p_1\) depends on \(q_B\) and \(q_C\) as follows.

- If \(q_B + q_C \leq 1\), then \(p_1\) can be arbitrarily close to zero, and then \(W_0\) will be arbitrarily close to zero.
- If \(q_B + q_C > 1\), then \(q_B + q_C = 2p_1 + p_2 + p_3 < p_1 + 1\) and \(p_1 > q_B + q_C - 1\) and so \(W_0 > e^{-\delta}(q_B + q_C - 1)\).
- Clearly \(p_1 < \min\{q_B, q_C\}\) and so \(W_0 < e^{-\delta}\min\{q_B, q_C\}\).

So we get \(e^{\delta}W_0 \in (\max\{0, q_B + q_C - 1\}, \min\{q_B, q_C\})\). Note that this range is always non-empty.

2. (a) The direct proof is to calculate the moment generating function of \(X_t^{(\epsilon)}\)
\[
\mathbb{E}(e^{\gamma X_t^{(\epsilon)}}) = e^{-\gamma \mu - \lambda \epsilon} + e^{(1-\mu)\epsilon}(1 - e^{-\lambda \epsilon}) = e^{-\gamma \mu - (1 + (1 - e^{-\lambda \epsilon})(\epsilon^\gamma - 1))}
\]
and to see
\[
\mathbb{E}(e^{\gamma S_{[t/\epsilon]}^{(\epsilon)}}) = e^{-\gamma \mu [t/\epsilon]} \left(1 + \frac{[t/\epsilon](1 - e^{-\lambda \epsilon})(\epsilon^\gamma - 1)}{[t/\epsilon]}\right)^{[t/\epsilon]} \rightarrow e^{-\gamma \mu} e^{-\lambda (\epsilon^\gamma - 1)}
\]
which we recognise as being the moment generating function of \(X_t = N_t - \mu t\).
(b) This is a special case of the $n$-step generalisation of the two-asset model $(A, S)$ on two scenarios. Since $A_0 = S_0 = 1$, we have no arbitrage if and only if $S_1(\text{down}) < A_1 < S_1^{\text{up}}$. Here, this is
\[ e^{-\mu \varepsilon} < e^{\delta \varepsilon} < e^{1-\mu \varepsilon} \iff -\mu < \delta < 1/\varepsilon - \mu. \]

(c) We need
\[ 1 = S_0^{(e)} = e^{-\delta \varepsilon} \mathbb{E}_q(S_1^{(e)}) = e^{-\delta \varepsilon} (e^{-\mu \varepsilon}(1 - q_e) + e^{1-\mu \varepsilon} q_e) \]
and so
\[ q_e = \frac{e^{\delta \varepsilon} - e^{-\mu \varepsilon}}{e^{-\mu \varepsilon}(e - 1)} = \frac{e^{\mu \varepsilon} - e^{\delta \varepsilon} - 1}{e - 1}. \]

(d) This is in complete analogy to (a). We deduce this from the Poisson limit theorem considering $\tilde{T}_n^{(e)} = S_n^{(e)} - n \mu \varepsilon$, a Bernoulli random walk with success probability $q_e$. Noting that
\[ \frac{1}{\varepsilon} q_e = \frac{1}{e - 1} \frac{\varepsilon(\delta + \mu) - 1}{\varepsilon} \to \frac{\delta + \mu}{e - 1} \quad \text{as} \ \varepsilon \downarrow 0, \]
we obtain $\tilde{T}_{[t/\varepsilon]} \to \tilde{N}_t$ in distribution, as required. Now, clearly $[t/\varepsilon] \mu \varepsilon \to \mu t$, and taking differences in the two limit results completes the argument.

(e) Note from the moment generating function of the Poisson distribution that
\[ \mathbb{E}(e^{\tilde{N}_t}) = e^{\frac{\lambda + \mu \varepsilon}{\varepsilon}(e - 1)} = e^{\delta t + \mu t} \]
and so $M_t = e^{-\delta t} e^{\tilde{N}_t - \mu t}$ is a martingale, because for $s < t$
\[ \mathbb{E}(M_t | \mathcal{F}_s) = \mathbb{E}(e^{-\delta s} e^{\tilde{N}_s - \mu s} e^{-\delta(t-s)} e^{(\tilde{N}_t - \tilde{N}_s) - \mu(t-s)} | \mathcal{F}_s) = e^{-\delta s} e^{\tilde{N}_s - \mu s} e^{-\delta(t-s)} \mathbb{E}(e^{\tilde{N}_t - \tilde{N}_s}) = M_s. \]
Given $N_t = k$ or $\tilde{N}_t = k$, the two processes $(e^{N_s - \mu s})_{0 \leq s \leq t}$ and $(e^{\tilde{N}_s - \mu s})_{0 \leq s \leq t}$ have the same conditional distribution, since the $k$ jump times of $\tilde{N}$ and $N$ occur at independent uniform times on $[0, t]$. Since also $\mathbb{P}(N_t = k) > 0$ if and only if $\mathbb{P}(\tilde{N}_t = k) > 0$, the same paths are possible for the two processes. Since the discounted process $e^{-\delta t} e^{N_t - \mu t}$ is a martingale, it provides martingale probabilities for the equivalent process $e^{N_t - \mu t}$.

(f) $(N_t)_{t \geq 0}$ only has jumps of size 1, all other jumps are impossible, and the only Lévy processes with this property are Poisson processes with drift. If $(Y_t)_{t \geq 0}$ is a Poisson process with drift $-\nu t$, then we have
\[ \mathbb{P}((e^{Y_s})_{0 \leq s \leq t} \in D_\nu) = 1. \]
Since $D_\mu \cap D_\nu = \emptyset$ for $\mu \neq \nu$, we must have $\mu = \nu$ in order that $e^{Y_t}$ has the same possible paths as $e^{N_t - \mu t}$. We can now check that of all intensities $\lambda > 0$ of $Y$, only $\lambda = (\delta + \mu)/(e - 1)$ is such that $M_t = e^{-\delta t} e^{Y_t}$ is a martingale:
\[ \mathbb{E}(M_t | \mathcal{F}_s) = \mathbb{E}(e^{-\delta s} e^{Y_s} e^{-\delta(t-s)} e^{Y_t-Y_s} | \mathcal{F}_s) = e^{-\delta s} e^{Y_s} e^{-\delta(t-s)} \mathbb{E}(e^{Y_t-Y_s}) = M_s e^{-\delta(t-s)+\lambda(e-1)}. \]
B.6 Time change and subordination

1. (a) First note that

\[
\mathbb{E} \left( \sum_{j=1}^{[2^n y]} (Z_{j2^{-n}} - Z_{(j-1)2^{-n}})^2 \right) = \sum_{j=1}^{[2^n y]} \text{Var} \left( ( B_{f(j2^{-n})} - B_{f((j-1)2^{-n})} \right) \\
= \sum_{j=1}^{[2^n y]} (f(j2^{-n}) - f((j-1)2^{-n})) \\
= f([2^n y]2^{-n}) - f(0) = f([2^n y]2^{-n}) \to f(y),
\]

as \( n \to \infty \). For \( L^2 \)-convergence we then calculate

\[
\mathbb{E} \left( \left( \sum_{j=1}^{[2^n y]} (Z_{j2^{-n}} - Z_{(j-1)2^{-n}})^2 - f(y) \right)^2 \right) \\
\leq \text{Var} \left( \sum_{j=1}^{[2^n y]} (Z_{j2^{-n}} - Z_{(j-1)2^{-n}})^2 \right) + (f([2^n y]2^{-n}) - f(y))^2 \\
\leq \left( \sum_{j=1}^{[2^n y]} (f(j2^{-n}) - f((j-1)2^{-n}))^2 \right) \text{Var}(B_t^2) + (f([2^n y]2^{-n}) - f(y))^2 \\
\to [f_y^2 \text{Var}(B_t^2) = 0,
\]

provided that \( f \) is continuous (and increasing). Convergence in \( L^2 \) implies convergence in probability.

(b) Note first that both \( Z \) and \( \tilde{Z} \) are continuous. For marginal distributions, note that \( Z_y = B_{f(y)} \sim \text{Normal}(0, f(y)) \) and, for \( y_j \leq y < y_{j+1} \),

\[
\tilde{Z}_y = \sum_{i=1}^j \sigma_i (W_{y_i} - W_{y_{i-1}}) + \sigma_{j+1} (W_y - W_{y_j})
\]
is the sum of independent \( \sigma_i (W_{y_i} - W_{y_{i-1}}) \sim \text{Normal}(0, \tau_i^2) \), where

\[
\tau_i^2 = \sigma_i^2 (y_i - y_{i-1}) = \int_{y_{i-1}}^{y_i} f'(s)ds = f(y_i) - f(y_{i-1}),
\]

and these variances add up to \( f(y) \), as well. As for joint distributions, \( Z \) and \( \tilde{Z} \) have independent increments: for \( 0 = u_0 < u_1 < \ldots < u_n \)

\[
Z_{u_k} - Z_{u_{k-1}} = B_{f(u_k)} - B_{f(u_{k-1})} \sim \text{Normal}(0, f(u_k) - f(u_{k-1})),
\]

are independent as increments of \( B \); similarly, increments \( \tilde{Z}_{u_k} - \tilde{Z}_{u_{k-1}} \), for \( y_{k-1} < u_{k-1} \leq y_k \) and \( y_{r_k-1} < u_k < y_{r_k} \), are independent as linear combinations (for \( l_k < r_k \), just a multiple for \( l_k = r_k \)) of increments of \( W \)

\[
\tilde{Z}_{u_k} - \tilde{Z}_{u_{k-1}} = \sigma_{l_k} (W_{y_{l_k}} - W_{y_{l_k-1}}) + \sum_{i=l_k+1}^{r_k-1} \sigma_i (W_{y_i} - W_{y_{i-1}}) + \sigma_{r_k} (W_{y_{r_k}} - W_{y_{r_k-1}}) \\
\sim \text{Normal}(0, f(u_k) - f(u_{k-1})).
\]
(c) Take a Poisson process $X$ of rate $\lambda$. Then the process $Z = (X_f(y))_{y \geq 0}$ has jumps of size 1 only, for all continuous functions $f : [0, \infty) \to [0, \infty)$. However, if for a function as in (b), we have $\sigma_j \neq 1$ for some $j \geq 1$, then there is positive probability that $\tilde{Z}_y = \int_0^y \sqrt{f'(s)}dX_s$ has jumps of size $\sigma_j$, specifically, there will be a $\text{Poi}(\lambda(y_{j+1} - y_j))$ number of such jumps in the time interval $(y_j, y_{j+1}]$.

2. (a) If $\text{Var}(X_1) < \infty$ (and hence $\text{Var}(X_t) = t\text{Var}(X_1)$ and $\mathbb{E}(X_t) = t\mathbb{E}(X_1)$) and $\text{Var}(\tau_1) = \int_0^\infty t^2 g_r(t)dt < \infty$, we check the stronger integrability condition

$$
\int_{-\infty}^{\infty} z^2 g(z)dz = \int_{-\infty}^{\infty} z^2 \int_0^\infty f_l(z)g_r(t)dt dz
$$

$$
= \int_0^\infty \int_{-\infty}^{\infty} z^2 f_l(z)dz g_r(t)dt
$$

$$
= \int_0^\infty (\text{Var}(X_t) + (\mathbb{E}(X_t))^2)g_r(t)dt
$$

$$
= \int_0^\infty (t\text{Var}(X_1) + t^2(\mathbb{E}(X_1))^2)g_r(t)dt < \infty.
$$

(b) If $\tau$ is a compound Poisson process, i.e. $\int_0^\infty g_r(t)dt < \infty$, then

$$
\int_{-\infty}^{\infty} g(z)dz = \int_0^\infty \int_{-\infty}^{\infty} f_l(z)dz g_r(t)dt = \int_0^\infty g_r(t)dt < \infty.
$$

If $X$ is a compound Poisson process with intensity $\lambda$ and such that $\mathbb{P}(X_t \in (a, b)) = \int_a^b f_l(x)dx$ for all $(a, b) \notin 0$ and $\mathbb{P}(X_t \neq 0) = 1 - e^{-\lambda t}$, then

$$
\int_{\mathbb{R}\setminus\{0\}} g(z)dz = \int_0^\infty \int_{\mathbb{R}\setminus\{0\}} f_l(z)dz g_r(t)dt = \int_0^\infty (1 - e^{-\lambda t})g_r(t)dt < \infty.
$$

Note that (a) and (b) deal, respectively, with the integrability condition for small $z$ and large $z$. The general case, when neither the conditions of (a) nor of (b) are satisfied, we know that we still obtain a Lévy density, from the calculation of characteristic functions in the lectures, but the integrability condition for Lévy densities is difficult to check directly.

(c) A Lévy density is associated with a bounded variation Lévy process if and only if

$$
\int_{-\infty}^{\infty} (1 \wedge |x|)g(x)dx < \infty
$$

First note that the scaling relation $X_t \sim t^{1/\alpha}X_1$ implies, by the transformation formula, that

$$
f_l(y) = t^{-1/\alpha} f_l(t^{-1/\alpha}y), \quad y \in \mathbb{R}.
$$
Therefore, we can check that

\[
\int_{-\infty}^{\infty} |y| g(y) \, dy = \int_0^{\infty} \int_{-\infty}^{\infty} |y| f_t(y) dgy_r(t) \, dt = \int_0^{\infty} \int_{-\infty}^{\infty} |t^{1/\alpha} f_t(y) dxg_r(t) \, dt = \int_0^{\infty} t^{1/\alpha} g_r(t) dt \left( \int_{-\infty}^{\infty} \frac{|x|}{t^{1/\alpha}} f_1(x) \, dx \right).
\]

The last integral is \( E(|X_1|) \). To see that it is finite, split \( X = X^{(1)} + X^{(2)} \) where \( X^{(2)} \) is the compound Poisson process of big jumps \( \Delta X \{ |\Delta X| \geq 1 \} \) and \( X^{(1)} \) is the martingale of small jumps. Now \( E(|X_1|) \leq E(|X^{(1)}_1|) + E(|X^{(2)}_1|) < \infty \), since the moment generating function of \( X^{(1)} \) is an entire function, hence all moments are finite, and

\[
E(|X^{(2)}_1|) \leq E \left( \sum_{0 \leq s \leq 1} |\Delta s| \right) = \int_{|x| \geq 1} |x||x|^{-\alpha - 1} < \infty.
\]

3. Since \( B \sim -B \), we clearly have for \( X_y = B_{\tau_y} \)

\[
\mathbb{P}(X_y \leq a) = \int_0^{\infty} \mathbb{P}(B_t \leq a) f_{\tau_y}(t) \, dt = \int_0^{\infty} \mathbb{P}(B_t \geq -a) f_{\tau_y}(t) \, dt = \mathbb{P}(X_y \geq -a),
\]

and this is enough, since both \( X \) and \( -X \) have stationary, independent increments and right-continuous paths with left limits.

On the other hand, the symmetric process \( X = P - N \) for two independent Poisson processes of rate \( \lambda \) cannot be obtained as Brownian motion subordinated by an independent subordinator, since for every random time \( T \) with [or without] probability density function we have

\[
\mathbb{P}(B_T = 1) = \int_0^{\infty} \mathbb{P}(B_t = 1) f_T(t) \, dt = 0 \quad \text{[or]} \quad \int_{[0,\infty)} \mathbb{P}(B_t = 1) \mathbb{P}(T \in dt) = 0.
\]

4. (a) Calculate

\[
\mathbb{E}(\exp\{\gamma B_{T_s}\}) = \int_0^{\infty} \mathbb{E}(\exp\{\gamma B_t\}) \frac{\lambda^{as} t^{as-1}}{\Gamma(as)} e^{-\lambda t} \, dt
= \int_0^{\infty} \mathbb{E}(\exp\{\frac{1}{2} t \gamma^2\}) \frac{\lambda^{as} t^{as-1}}{\Gamma(as)} e^{-\lambda t} \, dt
= \mathbb{E}(\exp\{\frac{1}{2} \gamma^2 T_s\}) = \left( \frac{\lambda}{\lambda - \frac{1}{2} \gamma^2} \right)^{as}.
\]

and identify this with the formula for \( C_s - D_s \) in Exercise A.1.2.(b).
5. (a) \[
\mathbb{E}(e^{i\lambda R_t}) = \mathbb{E}(e^{i\lambda B_{S_t}}) = \int_0^\infty f_{S_t}(s)\mathbb{E}(e^{i\lambda B_s})ds
\]
\[
= \int_0^\infty f_{S_t}(s)e^{-\frac{1}{2}\lambda^2 s}ds
\]
\[
= \mathbb{E}(e^{-\frac{1}{2}\lambda^2 S_t}) = e^{-t\Phi\left(\frac{1}{2}\lambda^2\right)} = e^{-t\Phi(1)\left(\frac{1}{2}\right)^\alpha|\lambda|^{2\alpha}}.
\]
We identify the symmetric $2\alpha$-stable distribution.

(b) We condition on \( S_t \) to get

\[
\mathbb{E}(e^{i\lambda B_{S_t}}) = \mathbb{E}(e^{i\lambda B_{Z_y}}) = \int_0^\infty f_{Z_y}(s)\mathbb{E}(e^{i\lambda B_s})(s)ds
\]
\[
= \int_0^\infty f_{Z_y}(s)e^{-\frac{1}{2}\lambda^2 s}ds
\]
\[
= \mathbb{E}(e^{-\frac{1}{2}\lambda^2 Z_y}) = e^{-t\Phi\left(\frac{1}{2}\lambda^2\right)} = e^{-t\Phi(1)\left(\frac{1}{2}\right)^\alpha|\lambda|^{2\alpha}}.
\]

(c) For the Variance Gamma process, we have

Now recall Method 3: Let \((X_t)_{t\geq 0}\) be a Lévy process with cumulative distribution function \(F_t\). Fix a time lag \(\delta > 0\). Then the process

\[
Z^{(\alpha,\delta)}_y = S_{[y/\delta]}, \quad \text{where } S_n = \sum_{k=1}^n Y_k \text{ and } Y_k = F^{-1}_{\tau_k - \tau_{(k-1)}\delta}(U_k)
\]
is the time discretisation of the subordinated process \(Z_y = X_{\tau_y}\).

and Example 84: We can use Method 3 to simulate the Variance Gamma process, since we can simulate the Gamma process \(\tau\) and we can simulate the \(Y_k\). Actually, we can use the Box-Muller method to generate standard Normal random variables \(N_k\) and then use

\[
\tilde{Y}_k \sim \sqrt{\tau_{k\delta} - \tau_{(k-1)\delta}}N_k, \quad k \geq 1,
\]

instead of \(Y_k, k \geq 1\).

5. (a) \[
\mathbb{E}(Z_t) = \int_0^\infty \mathbb{E}(X_s) f_{S_t}(s)ds = \int_0^\infty \mu S f_{S_t}(ds) = \mu nt.
\]
(b) \[
\mathbb{E}(Z_t^2) = \int_0^\infty \mathbb{E}(X_s^2) f_{S_t}(s)ds = \int_0^\infty (\sigma^2s + \mu^2S^2) f_{S_t}(ds) = \sigma^2 mt + \mu^2 q^2 t + \mu^2 m^2 t^2
\]
and then \(\text{Var}(Z_t) = \sigma^2 mt + \mu^2 q^2 t + \mu^2 m^2 t^2 - \mu^2 m^2 t^2 = (\sigma^2 m + \mu^2 q^2) t\).

(c) For the Variance Gamma process, we have

\[
\text{Var}(C_t - D_t) = \text{Var}(C_t) + \text{Var}(D_t) = 2 - \frac{\alpha t}{\sqrt{2\lambda}} = \frac{\alpha t}{\lambda}.
\]

On the other hand, for \(\sigma^2 = 1, \mu = 0, m = \alpha/\lambda\)

\[
\text{Var}(Z_t) = \frac{\alpha}{\lambda}t.
\]

Clearly \(\alpha = \lambda\) corresponds to \(\text{Var}(Z_t) = 1\). Differentiate the moment generating function \(\mathbb{E}(e^{t B_t}) = e^{t^2/2}\) four times to get \(\mathbb{E}(B_t^4) = 3t^2\). Then

\[
\mathbb{E}(B_{S_t}^4) = \int_0^\infty \mathbb{E}(B_{S_t}^4) f_{S_t}(s)ds = \int_0^\infty 3s^2 f_{S_t}(s)ds = 3 \mathbb{E}(S_t^2)
\]

Since \(S_t \sim \text{Gamma}(\alpha, \lambda)\), we get

\[
\mathbb{E}(B_{S_t}^4) = 3 \left( \frac{\alpha}{\lambda^2} + \frac{\alpha^2}{\lambda^2} \right) = 3 \left( 1 + \frac{1}{\lambda} \right) \in (3, \infty).
\]
B.7 Level passage events

1. (a) The process \(\tilde{X}^{(t)}\) has independent increments: for all \(0 = s_0 < s_1 < \ldots < s_n = t\), we have
\[
\tilde{X}^{(t)}_{s_j} - \tilde{X}^{(t)}_{s_{j-1}} = X_{t-s_{j-1}} - X_{t-s_j} = X_{t_{n-j+1}} - X_{t_{n-j}}, \quad 1 \leq j \leq n,
\]
where \(t_{n-j+1} = t - s_j\), and furthermore \(\mathbb{P}(X_{t_k} = X_{t_l}, 1 \leq k \leq n) = 1\); now we conclude by the independence of increments of \(X\). The same argument yields the stationarity of increments and identity in distribution with \(X\), since any increment of \(\tilde{X}^{(t)}\) has been identified with an increment of \(X\) of the same length. Right-continuity and left limits are also deduced, using the following remarks. First, for a right-continuous function \(f\) with left limits the function \(g(s) = f(t-s)\) is left-continuous with right limits. Second, for a left-continuous function with right limits, the function \(h(s) = g(t-s)\) is right-continuous with left limits.

(b) We condition on \(\tau\) to get for \(B = [a, b]\)
\[
\mathbb{P}(\tilde{X}^{(\tau)}_{s_1} \in A_1, \ldots, \tilde{X}^{(\tau)}_{s_m} \in A_m, \tau \in [s_m, s_{m+1}) \cap B)
\]
\[
= \int_{s_m}^{s_{m+1}} \mathbb{P}(\tilde{X}^{(\tau)}_{s_1} \in A_1, \ldots, \tilde{X}^{(\tau)}_{s_m} \in A_m|\tau = t) f_r(t) dt
\]
\[
= \int_{s_m}^{s_{m+1}} \mathbb{P}(\tilde{X}^{(t)}_{s_1} \in A_1, \ldots, \tilde{X}^{(t)}_{s_m} \in A_m) f_r(t) dt
\]
\[
= \int_{s_m}^{s_{m+1}} \mathbb{P}(X_{s_1} \in A_1, \ldots, X_{s_m} \in A_m) f_r(t) dt
\]
\[
= \mathbb{P}(X_{s_1} \in A_1, \ldots, X_{s_m} \in A_m, \tau \in [s_m, s_{m+1}) \cap B),
\]
where we applied (a) to pass from \(\tilde{X}^{(t)}\) to \(X\).

(c) Since \((\tilde{X}^{(\tau)}_t)_{0 \leq \tau \leq \tau} \sim (X_t)_{0 \leq \tau \leq \tau}\), we deduce from Proposition 86 that the following two random variables are independent:
\[
\overline{X}^{(\tau)}_\tau = \sup_{0 \leq \tau \leq \tau} (X_{\tau} - X_{\tau-s}) = X_\tau - \inf_{0 \leq \tau \leq \tau} X_{\tau-s} = X_\tau - \underline{X}_\tau
\]
\[
\underline{X}^{(\tau)}_\tau - \overline{X}^{(\tau)}_\tau = X_\tau - \overline{X}_\tau - X_{\tau-s} = -X_\tau,
\]
where we also used that \(\mathbb{P}(X_{\tau-s} = X_{\tau}) = 1\).

(d) We calculate from Corollary 89
\[
\mathbb{E}(e^{\beta \overline{X}_\tau}) = \mathbb{E} \left( e^{-\beta (\overline{X}^{(\tau)}_\tau - \overline{X}^{(\tau)}_\tau)} \right) = \frac{q(\Phi(q) - \beta)}{\Phi(q)(q - \phi(\beta))}.
\]

2. (a) Let us first show that \((T_x)_{x \geq 0}\) is a subordinator. See the proof of Proposition 90 for the stationarity of increments and the independence of two consecutive increments. For \(0 = x_0 < x_1 < \ldots < x_n\) we now show that \(T_{x_k} - T_{x_{k-1}}, \quad \ldots, \quad T_{x_1} - T_{x_0}, \quad T_x - T_0\) are independent, and furthermore
\[
\mathbb{P}(T_{x_k} - T_{x_{k-1}} = \tau | \mathcal{F}_t) = \mathbb{P}(T_{x_k} - T_{x_{k-1}} = \tau, \mathcal{F}_t), \quad \tau \leq t,
\]
\[
\mathbb{P}(T_{x_k} - T_{x_{k-1}} = \tau | \mathcal{F}_t) = \mathbb{P}(T_{x_k} - T_{x_{k-1}} = \tau, \mathcal{F}_t), \quad \tau > t.
\]
k = 1, \ldots, n, are independent and proceed by induction on n. Suppose independence holds for n, then the strong Markov property at $T_{x_n}$ shows that $T_{x_{n+1}} - T_{x_n}$ is independent of $\mathcal{F}_{T_{x_n}}$, where

$$\{T_{x_n} \leq t\} \cap \{T_{x_1} \leq s_1, \ldots, T_{x_n} \leq s_n\} = \bigcap_{k=1}^{n-1} \{T_{x_k} \leq s_k \land t\} \in \mathcal{F}_t$$

and so $\{T_{x_n} \leq t\} \cap \{T_{x_1} \leq s_1, \ldots, T_{x_n} \leq s_n\} \in \mathcal{F}_{T_{x_n}}$. We deduce that

$$\mathbb{P}(T_{x_k} \leq s_k, 1 \leq k \leq n, T_{x_{n+1}} - T_{x_n} \leq s_{n+1})$$

$$= \mathbb{P}(T_{x_k} \leq s_k, 1 \leq k \leq n, \tilde{T}_{x_{n+1} - x_n} \leq s_{n+1})$$

$$= \mathbb{P}(T_{x_k} \leq s_k, 1 \leq k \leq n)\mathbb{P}(T_{x_{n+1} - x_n} \leq s_{n+1}),$$

so that $(T_{x_1}, \ldots, T_{x_n})$ (and by linear transformation $(T_{x_1}, \ldots, T_{x_n} - T_{x_{n-1}})$) is independent of $T_{x_{n+1}} - T_{x_n}$. By the induction hypothesis, we conclude that all variables $T_{x_1}, \ldots, T_{x_{n+1}} - T_{x_n}$ are independent.

For the right-continuity let $x_n = x + \delta_n \downarrow x$ and note that the definition $T_x = \inf\{t \geq 0 : X_t > x\}$ implies that there is a sequence $\varepsilon \downarrow 0$ such that $X_{T_x + \varepsilon_n} > x + \delta_n$, but then $T_x \leq T_{x+\delta_n} \leq T_x + \varepsilon_n$ implies that $T_{x+\delta_n} \to T_x$. The existence of left limits is trivial for increasing $x \mapsto T_x$.

Then we apply Theorem 87 to see that

$$\mathbb{E}(\exp\{-qT_x\mathbf{1}_{(T_x<\infty)}\}) = \exp\{-x\Phi(q)\},$$

where we calculate $\Phi(q)$ by inverting $q = \phi(\Phi(q)) = c(\Phi(q))^\alpha$ to get $\Phi(q) = c^{-1/\alpha} q^{1/\alpha}$, and we identify the distribution of a stable subordinator with index 1/\alpha.

(b) We apply the transformation formula to get

$$f_{\alpha V}(z) = \frac{b/a}{\sqrt{2\pi z^3/a^3}} e^{-b^2/(2z/a)} = \frac{\sqrt{ab}}{\sqrt{2\pi z^3}} e^{-(\sqrt{ab})^2/(2z)} = f_{\sqrt{ab}}(z).$$

From Exercise 2.(c) we know that $f_b$ is the distribution of $T_b$ in the case $\alpha = 1/2$, $c = 1/2$, so we have just shown that $aT_b \sim T_{\sqrt{ab}}$, or indeed, for $a = d^2$, $x = \sqrt{ab}$ that

$$T_x \sim d^2 T_{x/d}, \quad x \geq 0.$$ 

Since both processes $(T_x)_{x \geq 0}$ and $(d^2 T_{x/d})_{x \geq 0}$ are subordinators, we have identity of joint distributions and hence another derivation, independent of (a) of the fact that $(T_x)_{x \geq 0}$ is a stable subordinator of index 1/2.

(c) We now deduce from (b) and (a) respectively that for all $\gamma \geq 0$

$$\int_0^\infty e^{-\gamma x} f_b(x)dx = \mathbb{E}(e^{-\gamma T_b}) = e^{-cd\sqrt{\gamma}}$$

for some $c \in (0, \infty)$. 

\[\]
3. (a) Let us denote the increments of \((N_m)_{m \geq 0}\) and of \((R_m)_{m \geq 0}\) by \(B_j = N_j - N_{j-1}\) and \(C_j = R_j - R_{j-1}\), \(j \geq 1\). Then we calculate

\[
E(\exp\{\gamma C_1\}) = E(\exp\{\gamma S_{B_1}\}) = \sum_{n \in \mathbb{N}} P(B_1 = n)E(\exp\{\gamma S_n\})
\]

\[
= \sum_{n \in \mathbb{N}} P(B_1 = n)M(\gamma)^n = G(M(\gamma)).
\]

For \(m \geq 1\), the analogous calculation yields

\[
E(\exp\{\gamma_1 C_1 + \ldots + \gamma_m C_m\})
\]

\[
= \sum_{n_1, \ldots, n_m \in \mathbb{N}} P(B_1 = n_1, \ldots, B_m = n_m)
\]

\[
E(\exp\{\gamma S_1 + \ldots + \gamma_m (S_{n_1 + \ldots + n_m} - S_{n_1 + \ldots + n_{m-1}})\})
\]

\[
= \sum_{n_1, \ldots, n_m \in \mathbb{N}} P(B_1 = n_1) \ldots P(B_m = n_m)M(\gamma)^{n_1} \ldots M(\gamma)^{n_m}
\]

\[
= G(M(\gamma_1)) \ldots G(M(\gamma_m)),
\]

and we can deduce that \(C_1, \ldots, C_m\) are independent and identically distributed, as required.

(b) We apply the same argument as in (a) to first calculate

\[
E(\exp\{i\lambda Y_s\}) = E(\exp\{\gamma X_{T_s}\}) = \int_0^\infty f_{T_s}(t)E(\exp\{i\lambda X_t\})dt
\]

\[
= \int_0^\infty f_{T_s}(t)\exp\{-t\psi(\lambda)\}dt = M_{T_s}(-\psi(\lambda)),
\]

where we assumed that \(T_s\) has a density \(f_{T_s}\) and that \(E(e^{i\lambda X_1}) = e^{-\psi(\lambda)}\). Now, for \(r, s \geq 0\),

\[
E(\exp\{i\lambda Y_s + i\mu(Y_{s+r} - Y_s)\})
\]

\[
= \int_0^\infty \int_0^\infty f_{T_s, T_{s+r}}(t, u)E(\exp\{i\lambda X_t + i\mu(X_{t+u} - X_t)\})dtdu
\]

\[
= \int_0^\infty \int_0^\infty f_{T_s}(t)f_{T_r}(u)e^{-t\psi(\lambda)}e^{-u\psi(\mu)}dtdu = M_{T_s}(-\psi(\lambda))M_{T_r}(-\psi(\mu))
\]

so that we deduce that \(Y_s\) and \(Y_{s+r} - Y_s\) are independent, and that \(Y_{s+r} - Y_s \sim Y_r\). For the right-continuity of paths, note that

\[
\lim_{\varepsilon \downarrow 0} Y_{s+\varepsilon} = \lim_{\varepsilon \downarrow 0} X_{T_s+\varepsilon} = X_{T_s} = Y_s
\]

since \(T_s + \delta := T_{s+\varepsilon} \downarrow T_s\) and therefore \(X_{T_s+\delta} \to X_{T_s}\). For left limits, the same argument applies.

(c) (i) First note that \(X_t = B_t + bt\) is also a Lévy process, as a linear combination of two Lévy processes. Because of the path continuity of Brownian motion, we have

\[
X_{T_s} = s, \quad \text{where} \quad T_s = \inf\{t \geq 0 : X_t \in (s, \infty)\}.
\]
The strong Markov property of $X$ at the stopping time $T_s$ (first hitting time of $(s, \infty)$) yields that $X^{(s)} = (X_{T_s+u} - s)_{u \geq 0}$ is independent of $T_s$ and has the same distribution as $X$. Note further that

$$T_{s+r} = T_s + T^{(s)}_r,$$

where $T^{(s)}_r = \inf\{t \geq 0 : X^{(s)}_t > r\}$, since we can wait for $X$ to exceed level $s + r$ by first waiting for level $s$ and then waiting to increase by a further $r$. Now, we get from the strong Markov property that

$$T^{(s)}_r \sim T_r \quad \text{and} \quad T^{(s)}_{s+r} = T_{s+r} - T_s \text{ is independent of } T_s.$$

This yields the stationarity and independence of increments ($n = 2$ but can now be generalised by induction). As an increasing process, we automatically get the existence of left and right limits. Assume now that at $s$ the path $s \mapsto T_s$ is not right-continuous. Then there is $\varepsilon > 0$ so that for all $\delta > 0$

$$T_{s+\delta} - T_s > \varepsilon \quad \Rightarrow \quad X_{T_{s+\delta}} \leq s + \delta \quad \text{for all } 0 \leq u \leq \varepsilon$$

but then $X_{T_{s+\delta}} \leq s$ for all $0 \leq u \leq \varepsilon$, and this contradicts the definition of $T_s$. So, the path $s \mapsto T_s$ must be right-continuous at $s$.

(ii) We check for $E_t = \exp\{\gamma B_t - \frac{1}{2}\gamma^2 t\}$ that

$$\mathbb{E}(E_t|\mathcal{F}_s) = \mathbb{E}(\exp\{\gamma B_s + \gamma(B_t - B_s) - \frac{1}{2}\gamma^2 t\}|\mathcal{F}_s)$$

$$= \exp\{\gamma B_s - \frac{1}{2}\gamma^2 t\} \mathbb{E}(\exp\{\gamma(B_t - N_B)\})$$

$$= \exp\{\gamma B_s - \frac{1}{2}\gamma^2 t\} \exp\{-\frac{1}{2}\gamma^2(t-s)\} = E_s.$$ 

Now, applying the Optional Stopping Theorem at $T_s$ (which can be shown to be a stopping time for $(E_t)_{t \geq 0}$), we get from $B_{T_s} = X_{T_s} - bT_s = s - bT_s$ that

$$1 = \mathbb{E}(E_{T_s}) = \mathbb{E}(\exp\{\gamma B_{T_s} - \frac{1}{2}\gamma^2 T_s\}) = e^{\gamma s} \mathbb{E}(\exp\{-b\gamma + \frac{1}{2}\gamma^2 T_s\})$$

and this yields the claim by choosing $\rho = \rho(\gamma) = -b\gamma + \frac{1}{2}\gamma^2$, i.e. $\gamma = \sqrt{b^2 - 2\rho - b}$.

4. (a) By the strong Markov property, the post-$T_y$ process $\tilde{X}_s = X_{T_y+s} - y$, $s \geq 0$, is a Brownian motion independent of $\mathcal{F}_{T_y}$. Note that we can write

$$X_t = \begin{cases} X_t & t \leq T_y \\ y + \tilde{X}_{t-T_y} & t \geq T_y \end{cases}$$

and that $\tilde{X} \sim -\tilde{X}$ (by the symmetry of the centred multivariate Normal distribution), so it remains to replace $y + \tilde{X}_{t-T_y} \sim y - \tilde{X}_{t-T_y} = 2y - X_t$. Formally, we check for $0 = t_0 < t_1 < \ldots < t_n < t_{n+1} = \infty$ that

$$\mathbb{P}(X^{*}_{t_1} \in A_1, \ldots, X^{*}_{t_n} \in A_n)$$
\[
\sum_{k=0}^{n} \int_{t_k}^{t_{k+1}} \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k, \\
y - \tilde{X}_{t_{k+1}-t} \in A_{k+1}, \ldots, y - \tilde{X}_{t_n-t} \in A_n) f_{T_y}(t) \, dt
\]

\[
= \sum_{k=0}^{n} \int_{t_k}^{t_{k+1}} \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k) \\
\mathbb{P}(y - \tilde{X}_{t_{k+1}-t} \in A_{k+1}, \ldots, y - \tilde{X}_{t_n-t} \in A_n) f_{T_y}(t) \, dt
\]

\[
= \sum_{k=0}^{n} \int_{t_k}^{t_{k+1}} \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_k} \in A_k) \\
\mathbb{P}(y + \tilde{X}_{t_{k+1}-t} \in A_{k+1}, \ldots, y + \tilde{X}_{t_n-t} \in A_n) f_{T_y}(t) \, dt
\]

\[
= \mathbb{P}(X_{t_n} \in A_1, \ldots, X_{t_n} \in A_n).
\]

(b) Note that \( \bar{X}_t > y \) implies \( T_y < t \) and so, if also \( X_t \leq x \), then \( X_t^* = 2y - X_t \geq 2y - x \). Vice versa, if \( X_t^* > 2y - x \), then \( x < y \) implies that \( 2y - x > y \) and so \( T_y < t \) and hence \( \bar{X}_t > y \), but also \( X_t^* = 2y - X_t \), so \( X_t < x \). This means that

\[\{X_t < x, \bar{X}_t > y\} = \{X_t^* > 2y - x\} \Rightarrow \mathbb{P}(X_t < x, \bar{X}_t > y) = \mathbb{P}(X_t^* > 2y - x)\]

We apply (a) to get upon differentiation, first w.r.t. \( x \) then \( y \) that

\[f_{X_t,\bar{X}_t}(x, y) = \frac{\partial}{\partial x} \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{(2y - x)^2}{2t} \right\} = \frac{2(2y - x)}{\sqrt{2\pi t^3}} \exp \left\{ -\frac{(2y - x)^2}{2t} \right\}\]

for \( y \in (0, \infty), x \in (-\infty, y) \).

(c) Since \( T_y < t \) if and only if \( \bar{X}_t > y \), we first calculate

\[f_{X_t}(y) = \int_{-\infty}^{y} \frac{2(2y - x)}{\sqrt{2\pi t^3}} \exp \left\{ -\frac{(2y - x)^2}{2t} \right\} \, dx = \frac{2}{\sqrt{2\pi t}} \exp \left\{ -\frac{y^2}{2t} \right\}\]

and then differentiate the distribution function of \( T_y \)

\[f_{T_y}(t) = \frac{d}{dt} \mathbb{P}(T_y \leq t) = \frac{d}{dt} \mathbb{P}(\bar{X}_t > y) = \frac{d}{dt} 2\mathbb{P}(X_t > y) = \frac{d}{dt} 2\mathbb{P}(X_1 > y) = \frac{d}{dt} \frac{2}{\sqrt{2\pi}} \exp \left\{ -\frac{y^2}{2t} \right\} \cdot \frac{1}{y} \cdot \frac{1}{y^{3/2}} = \frac{y}{\sqrt{2\pi t^3}} \exp \left\{ -\frac{y^2}{2t} \right\},\]

for all \( t > 0 \).