Computational Finance

Course notes
as taught by prof. Daniele Marazzina

by Jacopo Stringara (@jstringara)

jstringara@gmail.com

September 28, 2023

The source code for this document (and many more) can be found at:
https://github.com/jstringara/Latex-projects/tree/master/CF

Contents

Chapter 1

Lecture 1:

1.1 Introduction

We will study Lévy processes and extend the Black and Scholes theory. We will depart from analytical formulas and use MonteCarlo techniques, partial differential equations, partial differential integral equations and Fast-Fourier transforms.

1.2 Black and Scholes

The Black and Scholes model is defined by the following stochastic differential equation:

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t \\ S_0 = s \end{cases}$$

Chapter 2

Lecture 2: 18/09

2.1 Poisson Process

In order to define the Poisson Processes we must first recall and explore the notions of:

- Exponential Random Variable
- Poisson Distribution

2.1.1 Exponential Random Variable

A random variable Y is said to be exponentially distributed with parameter $\lambda > 0$ if it has the following probability density function:

$$f_Y(y) = \lambda e^{-\lambda y} \mathbb{1}_{\mathbb{R}_+}(y)$$

or equivalently that it has the following cumulative distribution function:

$$F_Y(y) = 1 - e^{-\lambda y} \mathbb{1}_{\mathbb{R}_+}(y)$$

and we write $Y \sim \text{Exp}(\lambda)$.

Theorem 1 (Absence of Memory)**.** *Let $T \geq 0$ be a random variable such that:*

$$\mathbb{P}(T > t + s | T > s) = \Pr(T > t) \quad \forall t, s > 0$$

then T can only be exponentially distributed.

In other words, the exponential distribution is the only continuous distribution with absence of memory.

Proof.

- ($\Longleftarrow$) We start with $T \sim \text{Exp}(\lambda)$, then its pdf is:

  $$f_T(y) = \lambda e^{-\lambda y} \mathbb{1}_{\mathbb{R}_+}(y)$$

  Thus:

  $$\mathbb{P}(T > t + s | T > s) = \frac{\mathbb{P}(T > t + s)}{\mathbb{P}(T > s)} = \frac{\int_{t+s}^{\infty} \lambda e^{-\lambda y} dy}{\int_s^{\infty} \lambda e^{-\lambda y} dy} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}}$$
  $$= e^{-\lambda s} = \int_s^{\infty} \lambda e^{-\lambda y} dy = \mathbb{P}(T > s)$$

- ($\Longrightarrow$) We define $g(t) = \mathbb{P}(T > t)$, then thanks to the hypothesis we get:

  $$g(t + s) = \mathbb{P}(T > t + s) = \mathbb{P}(T > t + s | T > t)\mathbb{P}(T > t) = \mathbb{P}(T > s)\mathbb{P}(T > t) = g(s)g(t)$$

  Now, g is decreasing and right continuous thanks to the definition of a CDF.

  Theorem. *The function $g(t) = e^{-\lambda t}$ is the only function that satisfies the two requirements above.*

  Thus:

  $$\mathbb{P}(T \leq t) = 1 - \mathbb{P}(T > t) = 1 - g(t) = 1 - e^{-\lambda t} \implies T \sim \text{Exp}(\lambda)$$

□

2.1.2 Poisson Distribution

Let N be a random variable with integer values, then N is said to be Poisson distributed with parameter $\lambda > 0$ if it has the following probability mass function:

$$\mathbb{P}(N = n) = \frac{\lambda^n}{n!} e^{-\lambda} \mathbb{1}_{\mathbb{N}}(n)$$

and we write $N \sim \text{Pois}(\lambda)$. Its moment generating function is given by:

$$M_N(u) = \mathbb{E}[e^{uN}] = e^{\lambda(e^u - 1)}$$

We will now create a sort of "bridge" between the exponential distribution and the Poisson distribution.

Counting Process

Theorem 2 (Counting Process)**.** *Let $\{\tau_n\}$ be a sequence of iid random variables with exponential distribution with parameter $\lambda > 0$, then $\forall t > 0$, we define:*

$$N_t = \inf_t \left\{ n \geq 0 : \sum_{i=1}^{n+1} \tau_i > t \right\}$$

Then we can say that N_t is a Poisson distributed random variable with parameter λt. Hence:

$$\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \mathbb{1}_{\mathbb{N}}(n)$$

Remark 3. For us the sequence $\{\tau_n\}$ will be a sequence of inter-arrival times, i.e. the time between two consecutive jumps of the process. N_t is called the **counting process**. It counts the number of jumps of the process up to time t.

Proof. Let us define the process $T_n = \sum_{i=1}^n \tau_i$. Now, since $\tau_i \sim \text{Exp}(\lambda)$, then we have that the PDF of T_n is:

$$P_n(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \mathbb{1}_{\mathbb{R}_+}(t)$$

thus:

$$\begin{aligned}
\mathbb{P}(T_{n+1} > t) &= \int_t^\infty P_{n+1}(s)ds = 1 - \overbrace{\int_0^t P_{n+1}(s)ds}^{\mathbb{P}(T_{n+1} \leq t)} = 1 - \int_0^t \underbrace{\lambda e^{-\lambda s}}_{f} \underbrace{\frac{(\lambda s)^n}{n!}}_{g'} ds \\
&\Big| \\
\text{integrating by parts} &= 1 + \underbrace{e^{-\lambda t}}_{f} \underbrace{\frac{(\lambda s)^n}{n!}}_{g} - \int_0^t \underbrace{e^{-\lambda s}}_{f} \underbrace{\frac{(\lambda s)^n}{n!}}_{g'} n\, ds \\
&\Big| \\
&= 1 + e^{-\lambda t} \frac{(\lambda t)^n}{n!} - \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} ds \\
&\Big| \\
&= 1 + e^{-\lambda t} \frac{(\lambda t)^n}{n!} - \mathbb{P}(T_n < t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} + \mathbb{P}(T_n \geq t)
\end{aligned}$$

Hence we can write:

$$\mathbb{P}(N_t = t) = \mathbb{P}(T_{n+1} > t) - \mathbb{P}(T_n \geq t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

thus $N_t \sim \text{Pois}(\lambda t)$. □

Remark 4.

1. If $Y_1 \sim \text{Pois}(\lambda_1)$ and $Y_2 \sim \text{Pois}(\lambda_2)$ are independent, then $Y_1 + Y_2 \sim \text{Pois}(\lambda_1 + \lambda_2)$.
2. If $Y \sim \text{Pois}(\lambda)$, then we can divide it into:

$$Y = \sum_{i=1}^n Y_i \quad \text{where } Y_i \sim \text{Pois}\left(\frac{\lambda}{n}\right)$$

Definition 5 (Poisson Process (PP))**.** Let $\{T_n\}$ be a sequence of iid random variables with exponential distribution with parameter $\lambda > 0$, and $T_n = \sum_{i=1}^{n} \tau_i$, then:

$$N_t = \sum_{n=1}^{\infty} \mathbb{1}_{\{t > T_n\}}$$

is called a Poisson Process with intensity λ.

Ask if the definition is done with $\geq$ or $>$ in the indicator function.

Properties 6.

- $N_t \sim \text{Pois}(\lambda t)$ if $t > 0$.
- N_t is a Cadlag process. Hence it is right continuous and has left limits.
- its characteristic function can be computed as:

  $$\Phi_{N_t}(u) = \mathbb{E}[e^{iuN_t}] = e^{\lambda t(e^{iu}-1)}$$

  We also call $\psi_{N_t}(u) = \lambda t(e^{iu} - 1)$ the characteristic exponent of N_t.
- N_t has independent increments, i.e. $\forall t_1, \ldots, t_n$, we have that:

  $$N_{t_n} - N_{t_{n-1}} \perp\!\!\!\perp \ldots \perp\!\!\!\perp N_{t_2} - N_{t_1} \perp\!\!\!\perp N_{t_1} - \cancelto{0}{N_0}$$

We have so far defined the Weiner process, the Poisson process and we will at this point move onto the compound Poisson process and the compensated Poisson process. We will now see that the Poisson process is a special case of the compound Poisson process.
Let us take a family of iid random variables $\{Y_n\}$ and a sequence of iid random variables $\{\tau_n\}$ with exponential distribution of parameter $\lambda > 0$. Just like we did above we can define the process $T_n = \sum_{i=1}^{n} \tau_i$ and the counting process $N_t = \sum_{n=1}^{\infty} \mathbb{1}_{\{t \geq T_n\}}$. Now we move on to define the compound Poisson process.

Definition 7 (Compoind Poisson Process (CPP))**.** Let $\{Y_n\}$ be a sequence of iid random variables and $\{\tau_n\}$ be a sequence of iid random variables with exponential distribution of parameter $\lambda > 0$. Let also these two families be independent, such that $Y_i \perp\!\!\!\perp N_t = \sum_{n=1}^{\infty} \mathbb{1}_{\{t \geq T_n\}}$. Then we define the compound Poisson process as:

$$Y_t = \sum_{n=1}^{\infty} Y_n \mathbb{1}_{\{t \geq T_n\}}$$

Given a compound Poisson process Y_t, we can represent graphically its path as a step function, where the jump sizes are given by the values of Y_n and the time elapsed between two consecutive jumps is given by τ_n.

Add graph of a step function with the jumps and the time elapsed between them.

One unsatisfactory aspect of the Poisson process is that it is not a Martingale. Indeed, take any two times s and t such that $s < t$, then:

$$\mathbb{E}[N_t|N_s] = \mathbb{E}[N_t - N_s + N_s|N_s] = \mathbb{E}[N_t - N_s|N_s] + N_s = \mathbb{E}[N_t - N_s] + N_s = \lambda(t-s) + N_s \neq N_s$$

Definition 8. Let N_t be a Poisson process with intensity $\lambda > 0$, then the compensated Poisson process is defined as:

$$\tilde{N}_t = N_t - \lambda t$$

This is no longer a counting process since (due to the λt term) it is not strictly integer valued. However, it is a Martingale. Indeed by the same reasoning as above we have that:

$$\mathbb{E}[\tilde{N}_t|\tilde{N}_s] = \mathbb{E}[\tilde{N}_t - \tilde{N}_s|\tilde{N}_s] + \tilde{N}_s = \mathbb{E}[\tilde{N}_t - \tilde{N}_s] + \tilde{N}_s = 0 + \tilde{N}_s = \tilde{N}_s$$

Theorem 9 (Uniqueness of the Poisson Process)**.** *Let N_t be a counting process with intensity $\lambda > 0$, with integer values and independent and stationary increments then it can only be a Poisson process.*

Definition 10 (Stationary)**.** Let X_t be a stochastic process, then we say that it is stationary if $\forall t_1, \ldots, t_n$ and $\forall h > 0$ we have that:

$$(X_{t_1}, \ldots, X_{t_n}) \sim (X_{t_1+h}, \ldots, X_{t_n+h})$$

Check the definition of stationarity.

Recall. For the Weiner process we have that $W_t \sim \mathcal{N}(0,t) \overset{L}{=} \sqrt{t}\mathcal{N}(0,1)$. Thus simulating a Weiner process is rather simple as it only requires to simulate a normal random variable of mean 0 and variance 1. On the other hand, for the Poisson process we have:

- $\{\tau_n\}$ iid $\text{Exp}(\lambda)$
- $N_t \sim \text{Pois}(\lambda t)$

if we want to simulate a Poisson process we need to simulate either the inter-arrival times themselves or directly the counting process.

- The first method, that of simulating is called **Countdown Sampling**. And can be achieved in MATLAB with the following code:

```
% Countdown Simulation
lambda = 2;
t = 1;
tau = icdf('Exponential', rand, 1/lambda);
while (sum(tau) < t) %this is just the condition that we are before T
tau_temp = icdf('Exponential', rand, 1/lambda) #?????????
tau = [tau, tau_temp];
end
Nt = length(tau)-1;
T = cumsum(tau);
T = T(1:end-1); %the last time is beyond T
```

Finish up the coding part of the lecture.

Chapter 3

lecture 3: 19/09

3.1 Lèvy processes

In quite simple words a Lèvy process is a stochastic process with independent and stationary increments. We will see that this is a very strong property and useful property that allows us to model many phenomena. In particular we will try to model stocks prices of the form $S_t = S_0 e^{rt+X_t}$ where X_t is a Lèvy process.

Example.

1. Brownian motion W_t
2. $-\frac{\sigma}{2}t + \sigma W_t$

Definition 11 (Lèvy process)**.** Let X_t be a Cadlag stochastic process, such that $X_0 = 0$. We say that X_t is a Lèvy process if it satisfies the following properties:

i X_t has independent increments, i.e. for $0 \leq t_1 < t_2 < \cdots < t_n$ we have that:

$$X_{t_1} - \cancelto{0}{X_0} \perp\!\!\!\perp X_{t_2} - X_{t_1} \perp\!\!\!\perp \cdots \perp\!\!\!\perp X_{t_n} - X_{t_{n-1}}$$

ii X_t has stationary increments, i.e. $\forall t, s > 0$ we have that:

$$X_{t+s} - X_t \overset{L}{=} X_s$$

iii X_t has stochastic continuity, i.e. $\forall \varepsilon > 0$ we have that:

$$\lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| > \varepsilon) = 0$$

This last property ensures that the process is continuous in probability. In other words that there are no jumps in the distribution of the jump times. No single instant is more likely than all others to be a jump time. We have no way of knowing with any probability when the processes' next jump will occur.

Properties 12.

- **Infinite divisibility**: X_t is infinitely divisible, i.e. $\forall n \in \mathbb{N}$, $\forall t > 0$ we have that:

$$X_t = \sum_{i=1}^{n} \hat{X}_i \quad \hat{X}_i = X_{i \cdot t/N} - X_{(i-1) \cdot t/N} \text{ iid}$$

  Thus, we can break the process into an infinite (but countable) number of independent increments.

- **Characteristic function and exponent**:

$$\Phi_{X_t}(u) = \mathbb{E}[e^{iuX_t}] = e^{t\Psi(u)}$$

  where $\Psi(u)$ is the characteristic exponent of the process. The function $t \mapsto \Phi_{X_t}$ is the characteristic function of

$$\Phi_{X_{t+s}}(u) = \Phi_{X_t}(u) \cdot \Phi_{X_s}(u)$$

Indeed:

$$\begin{aligned}\Phi_{X_{t+s}}(u) &= \mathbb{E}[e^{iuX_{t+s}}] = \mathbb{E}[e^{iu(X_t - X_s)} \cdot e^{iuX_s}] \\ \text{thanks to independence} &= \mathbb{E}[e^{iu(X_t-X_s)}] \cdot \mathbb{E}[e^{iuX_s}] = \Phi_{X_{t+s}-X_s}(u) \cdot \Phi_{X_s}(u) \\ \text{thanks to stationarity} &= \Phi_{X_t}(u) \cdot \Phi_{X_s}(u)\end{aligned}$$

Remark 13. The PP is the only Counting process which is also a Lèvy process. On the other hand the CPP is the only Lèvy process with piecewise constant trajectories. Indeed we may recall that $X_t = \sum_{i=1}^{N} Y_i$ where Y_i are iid random variables. Thus, if we make the choice $Y_i = 1$ we get the PP. Furthermore the characteristic function of X_t depends entirely on the distribution of Y_i:

$$\begin{aligned}\Phi_{X_t}(u) = \mathbb{E}[e^{iuX_t}] &= e^{t\lambda \int_{\mathbb{R}^d}(e^{iuy}-1)\cdot f(dy)} \\ \implies \Psi_{X_t}(u) &= \lambda \int_{\mathbb{R}^d} (e^{iuy} - 1) \cdot f(dy)\end{aligned}$$

where f is the distribution of Y_i. Furthermore if Y_i is Lebesgue integrable then $f(dy) = k(y)dy$.

Proof.

$$\begin{aligned}\mathbb{E}[e^{iuX_t}] &= \mathbb{E}\left[\mathbb{E}[e^{iuX_t}|N_t]\right] = \mathbb{E}\left[\mathbb{E}\left[e^{iu\sum_{i=1}^{N_t} Y_i}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[e^{iuN_tY_i}|N_t\right]\right] \\ &= \mathbb{E}\left[\left(\mathbb{E}\left[e^{iuY_i}\right]\right)^{N_t}\right] = \mathbb{E}\left[\Phi_{Y_i}^{N_t}(u)\right] = \sum_{n=0}^{\infty} \Phi_{Y_i}^n(u) \cdot \mathbb{P}(N_t = n) \\ &= \sum_{n=0}^{\infty} \Phi_{Y_i}^n(u) \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t} = e^{-\lambda t} \cdot e^{\Phi_{Y_i}(u)\cdot\lambda t} = e^{\lambda t\left(\Phi_{Y_i}(u)-1\right)}\end{aligned}$$

Thanks to the indipendence of the Y_i and the Taylor expansion of the exponential function. Now:

$$\Phi_{Y_i}(u) = \mathbb{E}\left[e^{iuY_i}\right] = \int_{\mathbb{R}^d} e^{iuy} \cdot f(dy) = \int_{\mathbb{R}^d} e^{iuy} \cdot k(y)dy$$

Hence, by substituting in the previous equation we get:

$$\Phi_{X_t}(u) = e^{\lambda t\left(\int_{\mathbb{R}^d} e^{iuy}\cdot k(y)dy - 1\right)} = e^{\lambda t\left(\int_{\mathbb{R}^d} e^{iuy}\cdot k(y)dy - \int_{\mathbb{R}^d} 1\cdot k(y)dy\right)}$$

thus, thanks to the fact that $k(y)$ is a distribution function, we ultimately get:

$$\Phi_{X_t}(u) = e^{\lambda t \int_{\mathbb{R}^d}(e^{iuy}-1)\cdot k(y)dy}$$

□

Sometimes we encounter the notation $\nu(dy) = \lambda k(y)dy = \lambda f(dy)$. This is called the **Lèvy measure** of the process.

Definition 14 (Lèvy measure of a CPP)**.** Let A be a Borel set in $\mathbb{R}^d$. We define the following quantities:

$$\begin{aligned}\Delta X_t &= X_t - X_{t^-} \\ X_{t^-} &= \lim_{s \to t^-} X_s\end{aligned}$$

Where ΔX_t is the jump size of the process at time t. Then:

$$\nu(A) = \mathbb{E}\left[\#\left\{t \in [0,1] :\ \Delta X_t \neq 0,\ \Delta X_t \in A\right\}\right]$$

This is the average number of time in a unitary time interval in which a jump occurs in the set A.

Definition 15 (Random Measure)**.** While ν is a deterministic measure on the set of Borel sets of $\mathbb{R}^d$, we can define the following random measure:

$$J_X(B) = \#\left\{(t, \Delta X_t) \in B\right\}\ \forall B \text{ open set in } [0, +\infty) \times \mathbb{R}^d$$

This a random variable of the number of jumps times and jump sizes in the set B.

Application. Let X_t be a CPP.

$$X_t = \sum_{i=1}^{N_t} Y_i = \sum_{s \in [0,t]} \Delta X_s = \int_{[0,t]\times\mathbb{R}^d} x \cdot J_X(ds \times dx)$$

J_X is a Poisson Random Measure with intensity $\nu(ds)dt = \lambda f(x)dxdt$.

3.2 Exponential Lèvy processes

In our framework we will model stocks and other processes as exponential Lèvy processes of the form:

$$S_t = S_0 e^{rt+X_t}$$

where X_t is a Lèvy process. Indeed, different choices of X_t will lead to different models. For example:

- $X_t = -\sigma^2/2t + \sigma W_t$ is the Black-Scholes model under the risk-neutral measure $\mathbb{Q}$.
- $X_t = \sum_{i=1}^{N_t} Y_i$ yields us the Coumpound Poisson Process model. The problem is that since this model is piecewise constant, it cannot capture the continuous variation of the stock prices on the real market.
- $X_t = \delta t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$ yields us with a jump-diffusion model (either Merton or Kou). We will see later how δ shall be chosen in order to fit with the risk-neutral measure $\mathbb{Q}$.

Merton and Kou models

- **Merton model**: $Y_i \sim \mathcal{N}(\mu_j, \delta_j^2)$ hence we have a total of 4 parameters $\sigma, \lambda, \mu_j, \delta_j$. μ_j and δ_j are the mean and variance of the jumps size.
- **Kou model**:

$$Y_i = \begin{cases} \text{Exp}(\lambda^+) \text{ with probability p} \\ -\text{Exp}(\lambda^-) \text{ with probability 1-p} \end{cases}$$

  Hence we have a total of 5 parameters $\sigma, \lambda, p, \lambda^+, \lambda^-$.

3.3 Choice of δ

Let us recall that we are under the $\mathbb{Q}$ measure if we have that:

$$\begin{aligned} &\mathbb{E}^{\mathbb{Q}}\left[S_t \cdot e^{-rt}\right] = S_0 \\ &\implies S_0 \mathbb{E}[e^{X_t}] = S_0 \cdot \Phi_{X_t}(-i) \end{aligned}$$

Thus we must choose δ such that:

$$\Phi_{X_t}(-i) = 1 \iff \Psi_{X_t}(-i) = 0$$

in order to satisfy the risk-neutral measure condition. Thus δ is not a parameter of the model that we can freely choose but must calibrated from the market in order to adhere to the risk-neutral measure.

3.4 Lévy-Itô decomposition

So far we have encountered three types of levy processes:

- Continuous, the family of Brownian motions.
- Finite variation, the Compound Poisson Process, which is unique.
- Jump-diffusion processes.

But not all levy processes are these types. This dependes on whether ν can be well defined.

Add picture of how to divide up levy processes

For example:

$$\nu((2,3)) = +\infty \implies \mathbb{E}[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in (2,3)\}] = +\infty$$

This is clearly not a Cadlag process. We have infinite activity in the interval $(2,3)$. In other words we have an infinite number of terms in the summation $\sum \Delta X_t$.

Definition 16 (Lévy-Itô decomposition)**.** Let $(X_t)_{t\geq 0} \in \mathbb{R}^d$ be a Lévy process. Let ν be its Lévy measure. Then:

i) ν is a Radon measure in $\mathbb{R}^d \setminus \{0\}$. In other words any compact set in $\mathbb{R}^d \setminus \{0\}$ is measurable.

ii)
$$\int_{x\in\mathbb{R}^d:\,|x|\geq 1} \nu(dx) < +\infty$$

iii)
$$\int_{x\in\mathbb{R}^d:\,|x|<1} |x|^2\,\nu(dx) < +\infty$$

Missed a comment about this from the lecture

iv) We can define a triplet (γ, A, ν) such that:
$$X_t = \gamma t + B_t + X_t^l + \lim_{\varepsilon\to 0} \hat{X}_t^\varepsilon$$
where:

- B_t is a Brownian motion with varianca-covariance matrix A.
- $$X_t^l = \sum_{\substack{s\in[0,t]\\ |\Delta X_s|\geq 1}} = \int_{[0,t]\times\{x\geq 1\}} x\cdot J_X(ds\times dx)$$
- We have the Compound Poisson process with jumps in $(\varepsilon, 1)$:
$$X_t^\varepsilon = \sum_{\substack{s\in[0,t]\\ \varepsilon<|\Delta X_s|<1}} \Delta X_s$$
Thus we get the respective compensated process:
$$\begin{aligned}\hat{X}_t^\varepsilon &= \sum_{\substack{s\in[0,t]\\ \varepsilon<|\Delta X_s|<1}} \Delta X_s - t\int_{\varepsilon<|x|<1} x\,\nu(dx)\\ &= \int_{[0,t]\times\{\varepsilon<|x|<1\}} x\cdot(J_X(ds\times d) - \nu(dx)ds)\end{aligned}$$

Indeed we have that:
$$\int_{x\in\mathbb{R}^d:\,|x|<1} |x|^2\,\nu(dx) < +\infty \implies \lim_{\varepsilon\to 0^+} \hat{X}_t^\varepsilon < +\infty$$
$$\int_{x\in\mathbb{R}^d:\,|x|<1} |x|^2\,\nu(dx) < +\infty \not\Longleftarrow \lim_{\varepsilon\to 0^+} X_t^\varepsilon < +\infty$$

Thanks to Martingality and the Central Limiti Theorem.

Example (One-dimensional case, from Jump-diffusion to Levy decomposition)**.** Since we are starting from a jump-diffusion process, we have that:
$$X_t = \delta_t + \sigma W_t + \sum_{s\in[0,t]} \Delta X_s \longrightarrow \sum_{s\in[0,t]} \Delta X_s = \begin{cases} \int_{|x|\geq 1} \nu(dx) < +\infty \\ \int_{|x|<1} \nu(dx) < +\infty \end{cases}$$

Clearly we have that $A = \sigma^2$. The Levy-Itô decomposition is:
$$X_t = \delta t + \sigma W_t + \sum_{\substack{s\in[0,t]\\ |\Delta X_s|\geq 1}} \Delta X_s + \lim_{\varepsilon\to 0^+} \hat{X}_t^\varepsilon$$

Now, let X_t be a Jump-Diffusion model:

Missed something about it being Lévy + finite variation

$$X_t = \gamma t + \sigma W_t + \sum_{\substack{s\in[0,t]\\ |\Delta X_s|\geq 1}} \Delta X_s + \lim_{\varepsilon\to 0^+} \left[\sum_{\substack{s\in[0,t]\\ \varepsilon<|\Delta X_s|<1}} \Delta X_s - t\int_{\varepsilon<|x|<1} x\,\nu(dx) \right]$$

We can write since when $|\Delta X_s| < 1$ we have that:

$$\int_{|x|<1} \nu(dx) < +\infty \implies \int_{|x|<1} x\,\nu(dx) < +\infty$$

Thus we can write:

$$\begin{aligned}
&= \gamma t + \sigma W_t + \sum_{\substack{s\in[0,t]\\ |\Delta X_s|\geq 1}} \Delta X_s + \sum_{\substack{s\in[0,t]\\ |\Delta X_s|<1}} \Delta X_s - t\int_{|x|<1} x\,\nu(dx) \\
&= \underbrace{t\left(\gamma - \int_{|x|<1} x\,\nu(dx)\right)}_{\delta} + \sigma W_t + \sum_{s\in[0,t]} \Delta X_s
\end{aligned}$$

Definition 17 (Levy-Khincin formula)**.** Let X_t be a Lèvy process with Lèvy-Itô decomposition (γ, A, ν). Then:

$$\Phi_{X_t}(u) = \mathbb{E}[e^{iuX_t}] = e^{t\Psi_{X_t}(u)}$$

where:

$$\Psi_{X_t}(u) = iu\cdot\gamma - \frac{1}{2}u^T A u + \int_{\mathbb{R}^d} (e^{iux} - 1 - iu\cdot x\mathbb{1}_{|x|\leq 1})\,\nu(dx)$$

Proof. Thanks to the Lèvy-Itô decomposition we have that:

$$X_t = \underbrace{\gamma t + B_t}_{X_t^c,\text{ continuous}} + \underbrace{X_t^l + \lim_{\varepsilon\to 0^+} \hat{X}_t^\varepsilon}_{X_t^j,\text{ jump}}$$

Thus we can split the characteristic function into two parts thanks to the fact that the continuous and jumping parts are independent:

$$\Phi_{X_t}(u) = \mathbb{E}[e^{iuX_t^c}]\cdot\mathbb{E}[e^{iuX_t^j}]$$

We shall now analyze them separately. For X_t^c we have that:

$$\mathbb{E}[e^{iuX_t^c}] = \mathbb{E}[e^{iu\gamma t}\cdot e^{iuB_t}] = e^{iu\gamma t}\cdot\mathbb{E}[e^{iuB_t}] = e^{iu\gamma t}\cdot e^{-\frac{1}{2}tu^T Au}$$

Let us now analyze X_t^j, but first let us fix the value of ε:

$$\begin{aligned}
X_t^{\varepsilon,j} &= \sum_{|\Delta X_s|\geq 1} \Delta X_s + X_t^\varepsilon \\
X_t^j &= \lim_{\varepsilon\to 0^+} X_t^{\varepsilon,j}
\end{aligned}$$

With a fixed value for ε we can observe that $X_t^{\varepsilon,j}$ is a CPP. Thus:

$$\begin{aligned}
\mathbb{E}[e^{iuX_t^j}] &= \lim_{\varepsilon\to 0^+} \mathbb{E}[e^{iuX_t^{\varepsilon,j}}] = \lim_{\varepsilon\to 0^+} \mathbb{E}\left[e^{iu\cdot\left(\sum_{\varepsilon<|\Delta X_s|} - t\int_{\varepsilon<|x|<1} x\,\nu(dx)\right)}\right] \\
&= \lim_{\varepsilon\to 0^+} \mathbb{E}\left[e^{iu\sum_{|\Delta X_s|\geq 1}\Delta X_s}\right]\cdot\mathbb{E}\left[e^{iu\left(\sum_{\varepsilon<|\Delta X_s|<1}\Delta X_s - t\int_{\varepsilon<|x|<1} x\,\nu(dx)\right)}\right] \\
&\text{using the CPP properties } \lim_{\varepsilon\to 0^+} e^{t\int_{|x|\geq 1}(e^{iux}-1)\,\nu(dx)}\cdot e^{t\left(\int_{\varepsilon<|x|<1}(e^{iux}-1)\,\nu(dx) - \int_{\varepsilon<|x|<1} ux\,\nu(dx)\right)} \\
&= \lim_{\varepsilon\to 0^+} e^{t\int_{|x|\geq 1}(e^{iux}-1)\,\nu(dx)}\cdot e^{t\int_{\varepsilon<|x|<1}(e^{iux}-1-iux)\,\nu(dx)} \\
&= \lim_{\varepsilon\to 0^+} e^{t\int(e^{iux}-1-iux\mathbb{1}_{\varepsilon<|x|<1})\,\nu(dx)}
\end{aligned}$$

Now, if we let $\varepsilon\to 0^+$ we get:

$$\mathbb{E}[e^{iuX_t^j}] = e^{t\int(e^{iux}-1-iux\mathbb{1}_{|x|<1})\,\nu(dx)}$$

□

Property 18 (Splitting the integral in the Levy-Khincin formula)**.**

$$\lim_{\varepsilon\to 0^+} \hat{X}_t^\varepsilon = \lim_{\varepsilon\to 0^+} \int_{\{\varepsilon<|x|<1\}\times[0,t]} x\cdot(J_X(ds\times dx)-\nu(dx)ds)$$

For a general Levy process this is finite ($+\infty$). Under certain conditions we are able to split the integral into two parts:

$$\left.\begin{array}{l}\int_{|x|>1}\nu(dx)<+\infty \text{ finite activity}\\ \int_{|x|\le 1}|x|\,\nu(dx)<+\infty \text{ finite variation}\end{array}\right\} \implies \lim_{\varepsilon\to 0^+}\int_{\varepsilon<|x|<1} x\,\nu(dx)<+\infty$$

Check with professor about the integrals before the brackets.

Let us formally define what variation actually consists of:

Definition (Total variation)**.** Let $f:[a,b]\to\mathbb{R}^d$ be a function. We define the total variation of f as:

$$\sup_{a=t_0<\cdots<t_n=b}\sum_{i=1}^{n}|f(t_i)-f(t_{i-1})| \quad n\in\mathbb{N}$$

In plainer words, the total variation is the supremum of the sum of the "jumps" of the function over all possible partitions of its domain. Furthermore a function is said to be of bounded variation if its total variation is finite.

Theorem (Finite variation Levy process)**.** *Let X_t be a Levy process with Levy-Itô decomposition (γ,A,ν). Then X_t is of finite variation if and only if:*

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

This yields us the following hierarchy of Levy processes (with respect to their variation):

$$\begin{array}{l}\int_{|x|\le 1}\nu(dx)<+\infty \to \text{ CPP}\\ \int_{|x|\le 1}|x|\,\nu(dx)<+\infty \to \text{ Finite variation}\\ \int_{|x|\le 1}|x|^2\,\nu(dx)<+\infty \to \text{ General Levy process}\end{array}$$

Let us also observe that only the first two are of bounded variation, and thus can have their integral limit split into two parts.

Chapter 4

Lecture 4: 25/09

4.1 Subordinator process

In this lecture we will talk about the subordinator process. In very general terms a subordinator is a process that is increasing in time and that we use to sub-index another process. In other words we use it to change the "speed" of the flow information and of the process. We will see that the subordinator process is a very useful tool in the study of Lévy processes.

We will start by stating the formal definition of a subordinator process.

Definition 19 (Subordinator process)**.** Let X_t be a Lévy process. We say that X_t is a subordinator process if and only if it satisfies all of the following conditions:

1) $X_t \geq 0$ for all $t \geq 0$ almost surely.

2) $\exists \bar{t} > 0$ such that $X_{\bar{t}} \geq 0$ almost surely.

3) X_t is not decreasing.

4) Given that the Lévy decomposition of X_t is (γ, A, ν), then:

- $\nu((-\infty, 0)) = 0$. In other words, the Lévy measure of X_t is zero in the negative real line. This is equivalent to saying that the process X_t does not have negative jumps.
- $A = 0$. In other words, the process X_t does not have a continuous (Brownian Motion) component.
- $b = \gamma - \int_{|x|\leq 1} x\,\nu(dx) > 0$. In other words the drift of the process X_t is non-negative. Recall that $\int_{|x|\leq 1} x\,\nu(dx)$ is the compensator process.
- $\int_{|x|\leq 1} x\,\nu(dx) < +\infty$. In other words, the compensator process is finite.

Remark. Note that the fourth condition is equivalent to saying that X_t is a finite variation Lévy process with and non-negative drift ($b > 0$) and no negative jumps ($\nu((-\infty, 0)) = 0$).

Theorem 20. *The conditions above are equivalent to one another.*

Proof.

- 1) $\implies$ 2): This is trivial. If $X_t \geq 0$ for all $t \geq 0$ then $X_{\bar{t}} \geq 0$ for some $\bar{t} > 0$.
- 3) $\implies$ 1): This is also trivial. By assumption X_t is not decreasing and since it is Lévy $X_0 = 0$ almost surely, then $X_t \geq 0$ for all $t \geq 0$ almost surely.
- 2) $\implies$ 3): By hypothesis of 2) we know that $\exists \bar{t} > 0$ such that $X_{\bar{t}} \geq 0$ almost surely. Let us now take $n \in \mathbb{N}$, then we rewrite $X_{\bar{t}}$ as follows:

$$X_{\bar{t}} = \underbrace{X_{\bar{t}} - X_{\bar{t}\frac{(n-1)}{n}}}_{\geq 0} + \underbrace{X_{\bar{t}\frac{(n-1)}{n}} - X_{\bar{t}\frac{(n-2)}{n}}}_{\geq 0} + \cdots + \underbrace{X_{\frac{2\bar{t}}{n}} - X_{\frac{\bar{t}}{n}}}_{\geq 0} + \cancelto{0}{X_0} \geq 0$$

Thus since by the assumption on X_t being a Lévy process we know that the increments are iid, this becomes a sum of iid random variables, that is it is equal to n times the same random variable. Furthermore since we know that $X_{\bar{t}} \geq 0$ almost surely, then we know that each increment must be non-negative. Hence:

$$\implies \left(X_{\bar{t}\frac{i}{n}} - X_{\bar{t}\frac{i-1}{n}}\right) \geq 0 \quad \forall i \in \{1, \ldots, n\}$$

This can easily extended to be of the form:

$$\left(X_{\bar{t}p} - X_{\bar{t}q}\right) \geq 0 \quad \forall p > q \in \mathbb{Q}$$

Furthermore since the set $\mathbb{Q}$ is dense in the set of reals $\mathbb{R}$ this can be even further be extended to be of the form:

$$(X_t - X_s) \quad \forall t > s \in \mathbb{R}$$

So far we have thus proved that the first 3 conditions are equivalent among each other, now we will pass on to prove that the fourth condition is equivalent to the third.

- 4) $\implies$ 3): We can easily see that:

$$\int_{|x|\leq 1} x\,\nu(dx) < +\infty \implies X_t = bt + A \cdot B_t + \sum_{s\in[0,t]} \Delta X_s$$

Furthermore since we have that $\nu((-\infty, 0)) = 0$ then we know that $\Delta X_s \geq 0$ almost surely. Thus we can conclude that X_t is not decreasing since it is a sum of non-negative random variables.

- 3) $\implies$ 4): In order to prove this last result we need a further theorem.

Theorem. *Let f be a Cádlág function and not-decreasing. Then f is also of finite variation.*

Now, we also know that a Lévy process that is of finite variation also has that:

$$X_t \text{ FV} \implies A = 0, \int_{|x|\leq 1} |x|\,\nu(dx) < +\infty$$

These conditions thus lead us to conclude that:

$$X_t = bt + \sum_{s\in[0,t]} \Delta X_s \text{ is not decreasing} \iff \begin{cases} b > 0 \\ \Delta X_s \geq 0 \text{ a.s.} \iff \nu((-\infty, 0)) = 0 \end{cases}$$

□

4.2 How to construct a subordinator process

We will now go on to see how a subordinator process can be constructed at will.

Theorem 21. *Let X_t be a Lévy process in $\mathbb{R}^d$ (with $d \geq 1$) with Lévy-Itô decomposition (γ, A, ν). Let also $f : \mathbb{R}^d \mapsto \mathbb{R}_+$ such that $f(x) = O(|x|^2)$ in a neighborhood of 0. Then:*

$$S_t = \sum_{\substack{s\in[0,t] \\ \Delta X_s \neq 0}} f(\Delta X_s)$$

is a subordinator process.

Before prooving this theorem let us make a few remarks and observations:

Remark. Note that in this theorem we have only used ν but have not used A or γ at all.

Remark 22. Let us note that from the Lévy-Khintchine formula, we know that given (γ, A, ν) we can easily write the characteristic exponent:

$$\Psi_{X_s}(z) = i\gamma^T \cdot z - \frac{1}{2} z^T A z + \int_{\mathbb{R}^d} \left(e^{iz^T x} - 1 - iz^T x \mathbb{1}_{|x|\leq 1}\right) \nu(dx)$$

Where we divide small (i.e. $|x| \leq 1$) and large (i.e. $|x| > 1$) jumps. Furthermore if we know that $\int_{|x|\leq 1} |x|\,\nu(dx) < +\infty$ then:

$$\Psi_{X_s}(z) = i\gamma_c^T \cdot z - \frac{1}{2} z^T A z + \int_{\mathbb{R}^d} \left(e^{iz^T x} - 1\right) \nu(dx)$$

Where γ_c is:

$$\gamma_c = \gamma - \int_{|x|\leq 1} x\,\nu(dx)$$

Or we could also write:

$$\Psi_{X_s}(z) = i\gamma_{\hat{C}}^T \cdot z - \frac{1}{2} z^T A z + \int_{\mathbb{R}^d} \left(e^{iz^T x} - 1 - iz^T x \right) \nu(dx)$$

with the alternative choice of $\gamma_{\hat{C}}$:

$$\gamma_{\hat{C}} = \gamma + \int_{|x|>1} x \, \nu(dx)$$

Let us notice that the $\int_{|x|\leq 1} |x| \, \nu(dx)$ term only appears in the drift. It is rather easy to take a function $g : \mathbb{R}^d \mapsto \mathbb{R}$ such that:

$$g(x) = O(|x|) \text{ for } |x| \to 0$$
$$g(x) = O\left(\frac{1}{|x|}\right) \text{ for } |x| \to \infty$$

Then we can write the characteristic exponent as:

$$\Psi_{X_s}(z) = i\hat{\gamma}^T \cdot z - \frac{1}{2} z^T A z + \int_{\mathbb{R}^d} \left(e^{iz^T x} - 1 - iz^T x g(x) \right) \nu(dx)$$
$$\text{with } \hat{\gamma} = \gamma - \int_{\mathbb{R}^d} x \left(\mathbb{1}_{|x|\leq 1} + g(x) \right) \nu(dx)$$

Thus we can freely choose g as we wish and we will still have a valid characteristic exponent. This is a very useful property that we will use later on.

Let us also see a few examples of g that we can use:

- $g(x) = \mathbb{1}_{|x|\leq 1}$
- $g(x) = \mathbb{1}_{|x|\leq \varepsilon}$ for some $\varepsilon > 0$

Finally we move on to the proof of the theorem.

Proof. Since f is a non-negative function, then we know that S_t is non-decreasing. All we need to prove is that S_t is a Lévy process. This is equivalent to proving that S_t respects the following properties:

- S_t is right-continuous.
- S_t is left-limited.
- $S_0 = 0$.
- S_t has independent increments.
- S_t is stochastically continuous.

Recall that S_t is defined as:

$$S_t = \sum_{\substack{s\in[0,t] \\ \Delta X_s \neq 0}} f(\Delta X_s)$$

Where X_t is a Lévy process and f is a non-negative function such that $f(x) = O(|x|^2)$ in a neighborhood of 0.

From the definition, since t is included into the sum, we can quite trivially conclude that S_t is right-continuous. Furthermore, by the same logic on the definition, we can conclude that $S_0 = 0$. It also has independent increments since it is a sum of independent increments of X_t, and by the same logic it is also stochastically continuous. What we really need to prove is that the process is left-limited. This entails the two following properties:

$$\int_{|x|>1} \nu_S(dx) < +\infty$$
$$\int_{|x|\leq 1} |x|^2 \, \nu_S(dx) < +\infty$$

By the hypothesis that X_t be a Lévy process we know that:

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

And we want to pass to:

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

Which is also equivalent to asking that:

$$\int_{|x|\leq \varepsilon} |x|^2 \, \nu_S(dx) < +\infty \quad \forall \varepsilon > 0$$

Again, from the hypothesis that $f(x) = O(|x|^2)$ in a neighborhood of 0 we have that:

$$\forall C > 0 \, \exists \varepsilon > 0 \quad f(\Delta X_s) \leq C|\Delta X_s|^2 \quad \text{if } |\Delta X_s| \leq \varepsilon$$

This can be piecewise extended to write:

$$\sum_{\substack{s\leq t \\ 0<|\Delta X_s|\leq \varepsilon}} f(\Delta X_s) \leq C \cdot \sum_{\substack{s\leq t \\ 0<|\Delta X_s|\leq \varepsilon}} |\Delta X_s|^2$$

Now, we know that X_t is a Lévy process and thus that this second term is bounded. Thus we can conclude that:

$$S_t = \sum_{\substack{s\leq t \\ 0<|\Delta X_s|\leq \varepsilon}} f(\Delta X_s) + \sum_{\substack{s\leq t \\ |\Delta X_s|>\varepsilon}} f(\Delta X_s)$$

And here we have that the first term is bounded as seen above, and the second term is of finite cardinality, since we have a finite number of jumps that are larger than ε. Thus we can conclude that S_t is left-limited and therefore that it is a Lévy process. □

4.3 Kou Model: Matlab implementations

We will now move on to write a Matlab function to simulate the Kou model.

We will start off by writing a function that simulates directly the Lévy process X_t. Recall that in the Kou model we have the following definition:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$$

Where Y_i is a jump of the form:

$$Y_i = \begin{cases} \text{Exp}(\lambda^+) \text{ with probability } p \\ \text{Exp}(\lambda^-) \text{ with probability } 1-p \end{cases}$$

Hence the increase in the process is given by:

$$\Delta X_t = X_{(i+1)\Delta t} - X_{i\Delta t} = \gamma \Delta t + \sigma \underbrace{(W_{(i+1)\Delta t} - W_{i\Delta t})}_{\sim \mathcal{N}(0,\Delta t)} + \sum_{\substack{s\in[i\Delta t,(i+1)\Delta t] \\ Y_s \neq 0}} Y_s$$

So, in our function we will need to have the following inputs:

- N_{sim}: the number of simulations to run.
- T: the time horizon.
- M: the number of time steps.
- Model parameters: $\sigma, \lambda, \lambda^+, \lambda^-, p$.

Furthermore, recall that we are working under the risk-neutral measure and that X_t is modeling the exponent of the stock price $S_t = S_0 e^{rt+X_t}$. Thus we will have:

$$\mathbb{E}^{\mathbb{Q}}[S_t e^{-rT}] = S_0 \implies S_0 \mathbb{E}[e^{X_t}] = S_0 \implies \mathbb{E}[e^{X_t}] = 1 \iff \Phi_{X_t}(-i) = 1 \iff \Psi_{X_t}(-i) = 0$$

For the Kou process the characteristic exponent is given by:

$$\Psi_{X_t}(u) = -\frac{\sigma^2 u^2}{2} + i\gamma u + iu\lambda + iu\lambda\left(\frac{p}{\lambda^+ - iu} - \frac{1-p}{\lambda^- + iu}\right)$$
$$\implies \Psi_{X_t}(-i) = \frac{\sigma^2}{2} + \gamma - \lambda\left(\frac{p}{\lambda^+ + 1} - \frac{1-p}{\lambda^- + 1}\right) = 0$$
$$\implies \gamma = \lambda\left(\frac{p}{\lambda^+ + 1} - \frac{1-p}{\lambda^- + 1}\right) - \frac{\sigma^2}{2}$$

Thus after we can compute γ like so in order to obey the assumption of working under the risk-neutral measure.

For the actual simulation, we will sample the number of jumps for each simulation from a Poisson distribution with parameter λT. Then we will sample the continuous component and add the jumps. The code is as follows:

```
function X=Kou_Simulation(Nsim,T,params,M)

% get the fuction and model parameters
dt=T/M;
X=zeros(Nsim,M+1);
sigma=params(1);
lambda=params(2);
p=params(3);
lambdap=params(4);
lambdam=params(5);
% characteristic exponent (equivalent to the above)
Psi=@(u) -sigma^2*u^2/2+1i*u*lambda*...
    (p/(lambdap-1i*u)-(1-p)/(lambdam+1i*u));
drift=-Psi(-1i);

% Sample number of jumps for each simulation
NT = icdf('Poisson',rand(Nsim,1),lambda*T);
% Sample Continuous Component and Add the Jumps
Z=randn(Nsim,M);
for j=1:Nsim % simulation loop
    % sample NT(j) jump times from a uniform distribution from 0 to T
    Jump_Times=sort( rand(1,NT(j)) )*T;
    for i=1:M % time loop
        % continuous part (adjust the standard deviation and add the drift)
        X(j,i+1)=X(j,i)+drift*dt+sigma*sqrt(dt)*Z(j,i);
        % add the jump component
        for jj=1:NT(j)
            % if the jump time is in the interval [i*dt,(i+1)*dt]
            if Jump_Times(jj)>(i-1)*dt &&...
                    Jump_Times(jj)<=i*dt
                % sample the jump size
                if rand<p % positive jump
                    Y=icdf('Exponential',rand,1/lambdap);
                else % negative jump
                    Y=-icdf('Exponential',rand,1/lambdam);
                end
                % add jump
                X(j,i+1)=X(j,i+1)+Y;
            end
        end
    end
end
```

This is a slightly inefficient but very intuitive way of implementing the Kou model. We will now see a more efficient way of implementing the model by refactoring the loops. Indeed we can see:

$$\sum_{\substack{s\in[i\Delta t,(i+1)\Delta t]\\ Y_s\neq 0}} Y_s \sim \sum_{\substack{s\in[0,\Delta t]\\ Y_s\neq 0}} Y_s$$

```
function X=Kou_Simulation_v2(Nsim,T,params,M)

%% same code as above ...

% Sample number of Jumps in each time interval
Ndt=icdf('Poisson',rand(Nsim,M),lambda*dt);
% Sample Continuous Component and Add the Jumps
Z=randn(Nsim,M);
for i=1:M % time loop
    % continuous part (matrix multiplication)
    X(:,i+1)=X(:,i)+drift*dt+sigma*sqrt(dt)*Z(:,i);
    for j=1:Nsim % simulation loop
        for jj=1:Ndt(j,i)
            % sample the jump size
            if rand<p % positive jump
                Y=icdf('Exponential',rand,1/lambdap);
            else % negative jump
                Y=-icdf('Exponential',rand,1/lambdam);
            end
            % add jump
            X(j,i+1)=X(j,i+1)+Y;
        end
    end
end
```

Here we are directly sampling the number of jumps in each time interval rather than in the whole time horizon. This relieves us from having to check whether the jump time is in the interval $[i\Delta t, (i+1)\Delta t]$.

N.B. if $N_{\Delta t}(j,i) = 0$ then the loop will not be executed at all.

There was a third implementation of the Kou model that the professor did show but did not upload to Webeep.

Chapter 5

Lecture 5: 26/09

5.1 Merton Model: Matlab implementation

Just like we did for the Kou model, we will now implement a Matlab function to simulate the Merton model. Recall that in the Merton model the Lévy process X_t is given by:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$$

where W_t is a standard Brownian motion, N_t is a Poisson process with intensity λ, and Y_i are i.i.d. random variables with distribution $\mathcal{N}(\mu_J, \sigma_J^2)$ independent of W_t and N_t. Hence we have to take in the following model and simulation parameters:

- N_{sim}: the number of simulations to run.
- T: the time horizon.
- M: the number of time steps.
- Model parameters: $\sigma, \lambda, \mu_J, \sigma_J$.

Let us note that we have one less parameter than in the Kou model. The characteristic exponent of the Merton model is given by:

$$\Psi_{X_t}(u) = -\frac{\sigma^2 u^2}{2} + i\gamma u + \lambda\left(e^{\mu_J iu - \frac{\sigma_J^2 u^2}{2}} - 1\right)$$

Again, thanks to the requirement that the mean must be done under the risk neutral measure, we have that $\Psi_{X_t}(-i)$ must be equal to 0. Hence:

$$\gamma = -\frac{\sigma^2}{2} - \lambda\left(e^{\frac{\sigma^2}{2}+\mu_J} - 1\right)$$

After computing the drift we then pass on to the actual simulation in which we will sample the number of jumps between each time step from a Poisson distribution with intensity $\lambda\Delta t$ and the continuous part from a Standard Normal which we will then rescale by $\sigma\sqrt{\Delta t}$. For each simulation we will sample the jump sizes from a Normal distribution with mean μ_J and variance σ_J^2. Finally we will sum the jumps and the continuous part to get the simulated values of the Lévy process at each time step.

Such a procedure can be implemented into the following Matlab function:

```
function X=Merton_Simulation(Nsim,T,params,M)

% parameters and drift
dt=T/M;
X=zeros(Nsim,M+1);
sigma=params(1);
lambda=params(2);
muJ=params(3);
sigmaJ=params(4);
Psi=@(u) -sigma^2*u^2/2+lambda*...
    (exp(-sigmaJ^2*u^2/2+1i*muJ*u)-1);
drift=-Psi(-1i);
```

```
13
14  % Sample Number of Jumps in each time-bucket
15  Ndt=icdf('Poisson',rand(Nsim,M),lambda*dt);
16  % Sample Continuous Component Normal of size Nsim x M
17  Z=randn(Nsim,M);
18
19  % Simulation
20  for i=1:M % time loop
21     % continuous part for the time step (matrix multiplication)
22     X(:,i+1)=X(:,i)+drift*dt+sigma*sqrt(dt)*Z(:,i);
23     for j=1:Nsim % simulation loop
24         if Ndt(j,i)>0 % if there are jumps
25             % sample jumps
26             Y=sum( muJ+sigmaJ*randn(Ndt(j,i),1) );
27             % add jumps
28             X(j,i+1)=X(j,i+1)+Y;
29         end
30     end
31  end
```

If we would like to actually simulate the Stock Price process we would have to take the exponential of the simulated Lévy process. We can do it by adding the initial stock price to the function parameters and then returning the exponential of the simulated Lévy process. We modify the function like so:

```
1  function S=Merton_Simulation(Nsim,T,params,M,S0,r)
2
3  %% Same code as before ...
4
5  Psi=@(u) -sigma^2*u^2/2+lambda*...
6      (exp(-sigmaJ^2*u^2/2+1i*muJ*u)-1);
7  drift=r-Psi(-1i);
8
9  %% Same code as before ...
10
11 % get the stock price
12 S=S0*exp(X);
13
```

Why is there the r in the drift?

5.2 Monte Carlo Simulation techniques

Now that we have developed an implementation for both the Merton and Kou models to simulate stock prices, how can we use these in order to price options? The answer lies in Monte Carlo simulation methods. Hence in this lecture we will go over how to use Monte Carlo simulation methods to price different types of options, with a strong focus on exotic options. Monte Carlo methods entail doing multiple simulations of the possible paths of the underlying asset, and then averaging the results to get the mean of the payoff function conditional on the underlying asset's price at the start of the simulation.

Unfortunately because the number of simulation to achieve high accuracy is often fairly large, Monte Carlo methods suffer from a very high computational overhead.

5.2.1 Theoretical Framework

First of all, let us understand what are the tools at our disposal. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let X be a random variable on this space. Our aim is to estimate its mean $\vartheta = \mathbb{E}[X]$. The easiest unbasied estimator we can construct is the sample mean:

$$\hat{\vartheta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

where X_i are i.i.d. copies of X. Recall that such an estimator has variance:

$$\mathrm{Var}[\hat{\vartheta}_n] = \frac{1}{n}\mathrm{Var}[X]$$

Now we will see two crucial theorems relating to the convergence and the asymptotic distribution of this estimator.

Theorem (Law of Large Numbers)**.** *Let $\{X_n\}$ be a sequence of i.i.d. random variables with mean ϑ. If we define the random variable:*

$$Y(n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

then:

$$\lim_{n\to\infty} \mathbb{P}(|Y(n) - \vartheta| > \varepsilon) = 0 \quad \forall \varepsilon > 0$$

Theorem (Central Limit Theorem)**.** *Let $\{X_n\}$ be a sequence of i.i.d. random variables with mean ϑ and variance σ^2. Then:*

$$\frac{\sum_{i=0}^{n} X_i - n\vartheta}{\sigma \cdot \sqrt{n}} = \frac{\hat{\vartheta}_n - \vartheta}{\sigma/\sqrt{n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1)$$

In other words, for a sufficiently large n we have that the distribution of $\hat{\vartheta}_n - \vartheta$, i.e. the MonteCarlo error, is approximately normal with mean 0 and variance $\frac{\sigma^2}{n}$:

$$\hat{\vartheta}_n - \vartheta \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right) \text{ for } n \text{ large enough}$$

This an extremely useful result, because, no matter the distribution of X, we can always rely on the fact that the distribution of the MonteCarlo error will be normal. One problem that can arise from not knowing the distribution of X is that we have no knowledge of the actual variance, so we will use the sample variance instead:

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\vartheta}_n)^2}$$

This is an unbiased estimator of the variance, and it is also consistent. More in general, we can say the the Monte Carlo method has a $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$ convergence rate. In other words in order to cut our estimation error in half we need to quadruple the number of samples. This is quite inconvenient and computationally onerous.

This is why we often prefer to use confidence intervals instead of point estimates. In particular we take two values a and b and estimate the probability that ϑ lies in the interval $[a, b]$.

Recall that from the law of large numbers we have that for n large enough:

$$\mathbb{P}\left(\frac{\hat{\vartheta}_n - \vartheta}{\sigma/\sqrt{n}} \leq x\right) \approx \Phi(x)$$

where Φ is the cumulative distribution function of the standard normal. This means that we can construct a confidence interval for ϑ by inverting the cumulative distribution function:

$$\mathbb{P}(a < \vartheta < b) = \mathbb{P}\left(\frac{\hat{\vartheta}_n - b}{\sigma/\sqrt{n}} < \frac{\hat{\vartheta}_n - \vartheta}{\sigma/\sqrt{n}} < \frac{\hat{\vartheta}_n - a}{\sigma/\sqrt{n}}\right) = \Phi(b) - \Phi(-a)$$

If we take the choice $a = b = z_{\delta/2}$ where $z_{\delta/2}$ is the $1 - \delta/2$ quantile of the standard normal distribution, we get that:

Ask professor what the passages actually are.

$$\mathbb{P}\left(\hat{\vartheta}_n - z_{\delta/2}\frac{\sigma}{\sqrt{n}} < \vartheta < \hat{\vartheta}_n + z_{\delta/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \delta$$

This is a $1 - \delta$ confidence interval for ϑ. Again, we often do not know the actual variance σ^2, so we will use the sample variance instead:

$$\mathbb{P}\left(\hat{\vartheta}_n - z_{\delta/2}\frac{\hat{\sigma}}{\sqrt{n}} < \vartheta < \hat{\vartheta}_n + z_{\delta/2}\frac{\hat{\sigma}}{\sqrt{n}}\right) = 1 - \delta$$

and we say that the interval:

$$\left[\hat{\vartheta}_n - z_{\delta/2}\frac{\hat{\sigma}}{\sqrt{n}}, \hat{\vartheta}_n + z_{\delta/2}\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

is an asymptotically valid confidence interval for ϑ.

5.2.2 Variance Reduction Techniques

As we have said above, Monte Carlo method have a rather slow converge rate of $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$. We will now pass on to see a few techniques that can be used to reduce the variance of the Monte Carlo estimator and thus speed up the convergence rate.

Antithetic Variables

Let ϑ be the expected value of a monotone function g of a random variable X. We know that the sample mean $\hat{\vartheta}_n$ is an unbiased estimator of ϑ:

$$\hat{\vartheta}_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$$

whose variance is given by:

$$\text{Var}[\hat{\vartheta}_n] = \frac{1}{n}\text{Var}[g(X)]$$

or rather its unbasied estimator:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(g(X_i) - \hat{\vartheta}_n)^2$$

Now, let us instead take two samples X_1, X_2 from the same population. Clearly we have that:

$$\begin{aligned}\text{Var}\left[\frac{g(X_1)+g(X_2)}{2}\right] &= \frac{1}{4}(\text{Var}[g(X_1)] + \text{Var}[g(X_2)] + 2\text{Cov}[g(X_1), g(X_2)]) \\ &= \frac{1}{2}\text{Var}[g(X)] + \frac{1}{2}\text{Cov}[g(X_1), g(X_2)]\end{aligned}$$

Now, if the two sample are statistically independent this quantity will exactly coincide with the variance of the sample mean for 2 samples $\hat{\vartheta}_2$. On the other hand, if the two samples are negatively correlated, i.e. if $\text{Cov}[g(X_1), g(X_2)] < 0$, then the variance of this quantity will be smaller than that of the sample mean $\hat{\vartheta}_n$.

What we actually do in practice is to take a single sample X and then create from this first sample a second one X' that is negatively correlated with X. For example:

$$\begin{aligned}&U^1 \sim \mathcal{U}([0,1]) \to U^2 = 1 - U^1 \implies \text{Cov}(U^1, U^2) < 0 \\ &Z^1 \sim \mathcal{N}(0,1) \to Z^2 = -Z^1 \implies \text{Cov}(Z^1, Z^2) < 0\end{aligned}$$

then we can define the antithetic variable estimator:

$$\hat{\vartheta}_{AV} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\left(g(X_i) + g(X_i')\right)$$

where X' is the antithetic variable of X.

Remark. Note that the function g must be monotone for this to be valid.

5.2.3 Merton model: Antithetic Variables implementation

Let us now see how we can implement the antithetic variables technique in our Merton model simulation.

We can modify the Merton asset simulation function as follows:

```
function [S,SAV]=Asset_Merton_AV(Nsim,T,params,M,S0,r)

%% same code as in original merton simulation...

% create two samples
X=zeros(Nsim,M+1); XAV=zeros(Nsim,M+1);
for i=1:M % time loop
    % continuous part
    X(:,i+1)=X(:,i)+drift*dt+sigma*sqrt(dt)*Z(:,i);
    % antithetic variables is the same but with -sigma*sqrt(dt)*Z(:,i)
    % since BM is a normal variable
    XAV(:,i+1)=XAV(:,i)+drift*dt-sigma*sqrt(dt)*Z(:,i);
    for j=1:Nsim % simulation loop
```

```
        if Ndt(j,i)>0
            % sample jumps
            z=randn(Ndt(j,i),1);
            Y=sum( muJ+sigmaJ*z );
            % antithetic is a sum of -z
            YAV=sum( muJ-sigmaJ*z );
            % add jumps
            X(j,i+1)=X(j,i+1)+Y;
            XAV(j,i+1)=XAV(j,i+1)+YAV;
        end
    end
end
S=S0*exp(X);
SAV=S0*exp(XAV);
```

5.3 Exotic options: refresher

As we have seen Monte Carlo simulation offer us a very general and powerful toolbox for estimating means of random variables. But how can we translate this knowledge into a pricing method for options?

Let us recall that under the assumtpions of the fundamental theorem of asset pricing, we have that the price of an option is given by the discounted expected value of its payoff under the risk-neutral measure $\mathbb{Q}$:

$$V_0 = e^{-rT}\mathbb{E}^{\mathbb{Q}}\left[\Phi(S(T))\right]$$

where Φ is the payoff function of the option, and $S(T)$ is the price of the underlying asset at maturity.

Thus we can use Monte Carlo simulations to estimate the price of an option by simulating the possible paths of the underlying asset, and then averaging the payoff function over the simulated paths. Or more frequently by estimanting the $1-\delta$ confidence interval for the price of the option, where δ usually takes the value 0.05 or 0.01.

This methodology is often applied to exotic options, i.e. options whose payoff function is not a simple function of the underlying asset's price at maturity but often depends on the whole path of the underlying asset. This makes it impossible to price these options using analytical methods, and thus Monte Carlo simulation is often the only viable option.

Asian Options

Asian options are a class of exotic options whose payoff depends on the average price of the underlying asset over the whole life of the option. In particular the average can be calculated as follows:

- Arithmetic average:
$$A = \frac{\sum_{i+0}^{M} S(i\Delta t)}{M+1}$$
- Geometric average:
$$A = \Pi_{i=0}^{M} S(i\Delta t)$$

Where Δt is the time step of the simulation, and M is the number of steps. TODO: frequency might be linked to the market convention. Furthermore Asian options can be divided into two categories depending on the type of strike price:

- Fixed strike:
$$\Phi(S(T)) = (A-K)^+ \text{ for call options}$$
$$\Phi(S(T)) = (K-A)^+ \text{ for put options}$$
- Floating strike:
$$\Phi(S(T)) = (S(T)-A)^+ \text{ for call options}$$
$$\Phi(S(T)) = (A-S(T))^+ \text{ for put options}$$

Lookback Options

Loockback options are another class of exotic options whose payoff depends on the maximum or minimum price of the underlying asset over the whole life of the option. In particular, since some of the possible combinations would not be reasonable to trade, we only find a subset of the possible payoffs on the market:

- Fixed strike:

$$\Phi(S(T)) = (M - K)^+ \text{ for call options}$$
$$\Phi(S(T)) = (K - m)^+ \text{ for put options}$$

- Floating strike:

$$\Phi(S(T)) = (S(T) - m)^+ \text{ for call options}$$
$$\Phi(S(T)) = (M - S(T))^+ \text{ for put options}$$

where M is the maximum price of the underlying asset over the life of the option, and m is the minimum price of the underlying asset over the life of the option. Note that these two quantities assume following values in our computational framework:

$$M = \max_i S(i\Delta t)$$
$$m = \min_i S(i\Delta t)$$

Barrier Options

Barrier options are a class of exotic options whose payoff depends on whether the underlying asset prices has visited a certain region of the state space during the life of the option, in particular we have Up, Down and Knock options. Up options have a payoff that depends on whether the underlying asset price has ever been above a certain threshold, Down options have a payoff that depends on whether the underlying asset price has ever been below a certain threshold while Knock options have a payoff that depends on whether the underlying asset price has ever been above or below a certain threshold. Furthermore barrier options can be divided into In and Out options depending on whether the payoff is activated when the underlying asset price enters or leaves the region of the state space.

Let us summarize these combinations in a table:

	Up	Down	Knock
In	Up and In	Down and In	Knock and In
Out	Up and Out	Down and Out	Knock and Out

Just like for lookback options, we only find a subset of the possible payoffs on the market due to the reasonableness of the payoff. Let us see a few examples of barrier options:

- Down and Out Call:
$$\Phi(S(T)) = (S(T) - K)^+ \cdot \mathbb{1}_{\{\min_i S(i\Delta t) > b\}} \text{ with } S_0 > b$$

- Up and In Put:
$$\Phi(S(T)) = (K - S(T))^+ \cdot \mathbb{1}_{\{\max_i S(i\Delta t) \geq u\}} \text{ with } S_0 < u$$

- Knock and Out Call:
$$\Phi(S(T)) = (S(T) - K)^+ \cdot \mathbb{1}_{\{\min_i S(i\Delta t) > b \, and \, \max_i S(i\Delta t) \geq u\}}$$

- Knock and In Call:
$$\Phi(S(T)) = (S(T) - K)^+ \cdot \mathbb{1}_{\{\min_i S(i\Delta t) \leq b \, or \, \max_i S(i\Delta t) < u\}}$$

5.4 Monte Carlo Simulation for Exotic Options (some examples)

Now that we have had a refresher on exotic options and their respective payoffs, let us see how we can use Monte Carlo simulation to price them using the Matlab function that we have developed for the Merton model (we could achieve the same with the Kou model).

5.4.1 European Call

First of all let us see how we can use Monte Carlo simulation to price a European call option. Recall that the payoff function of a European call option is given by:

$$\Phi(S(T)) = (S(T) - K)^+$$

where K is the strike price of the option. Hence we can use Monte Carlo simulation to estimate the price of the option by simulating the possible paths of the underlying asset, and then averaging the payoff function over the simulated paths. We will also provide a confidence interval for the price of the option.

We can do this using the following matlab code:

```
%% Contract parameters (maturity and strike price)
T=1; K=100;
%% Market parameters (price and interest rate)
S0=98; r=0.0001;
%% Model parameters (sigma, lambda, muJ, sigmaJ)
params=[0.7, 10, -0.01, 0.4];
%% Numerical parameters
Nsim=1e7; M=1;

%% Standard MC
%% Sample S
S=Asset_Merton(Nsim,T,params,M,S0,r);
%% Compute the Discounted Payoff
DiscPayoff=exp(-r*T)*max( S(:,end)-K, 0);
%% Compute the price
% get the mean and the confidence interval using normfit
[Price,~,Price_CI]=normfit(DiscPayoff)

%% Antithetic Variable MC
[S,SAV]=Asset_Merton_AV(Nsim,T,params,M,S0,r);
DiscPayoff=exp(-r*T)*max( S(:,end)-K, 0);
DiscPayoffAV=exp(-r*T)*max( SAV(:,end)-K, 0);
[Price,~,Price_CI]=normfit( (DiscPayoff+DiscPayoffAV)/2 )

```

5.4.2 Asian Call with Floating Strike

Recall from above that an Asian call option with floating strike has payoff:

$$\Phi(S(T)) = (S(T) - A)^+$$

where A is the arithmetic average of the underlying asset's price over the life of the option. In our particular case we have that the arithmetic average is computed with observations taken every week, i.e. Δt is equal to $1/52$ of a year.

We can use Monte Carlo simulation to estimate the price of the option by using the following code in Matlab:

```
%% Price an Asian Call Option
%% ->Floating Strike Call
%% ->Weekly monitoring

%% Contract parameters (maturity, in years, and number of times steps)
T=1; M=round(52*T);

%% Same code as above ...

%% Sample S
S=Asset_Merton(Nsim,T,params,M,S0,r);
%% Compute the Discounted Payoff
% the mean is taken across the rows, i.e. across the simulations
DiscPayoff=exp(-r*T)*max( S(:,end)-mean(S,2), 0);
%% Compute the price
[Price,~,Price_CI]=normfit( DiscPayoff )
```

5.4.3 Lookback Call with Fixed Strike

For a lookback call option with fixed strike we have that the payoff function is given by:

$$\Phi(S(T)) = (M - K)^+$$

where M is the maximum price of the underlying asset over the life of the option. In our particular case we have that the maximum price is computed with observations taken every month, i.e. Δt is equal to 1/12 of a year.

```
%% Price a Lookback Call Option
%% ->Fixed Strike Call
%% ->Monthly monitoring

%% Contract parameters (time horizon, strike price, number of time steps)
T=1; K=100; M=round(12*T);

%% same code as above ...

%% Sample S
S=Asset_Merton(Nsim,T,params,M,S0,r);
%% Compute the Discounted Payoff
% to compute the max across the rows, i.e. across the simulations, we use the
% we use a dummy variable []
DiscPayoff=exp(-r*T)*max( max(S,[],2)-K, 0);
%% Compute the price
[Price,~,Price_CI]=normfit( DiscPayoff )
```

5.4.4 Down and Out Call

For a down and out call option we have that the payoff function is given by:

$$\Phi(S(T)) = (S(T) - K)^+ \cdot \mathbb{1}_{\{\min_i S(i\Delta t) > b\}}$$

where b is the barrier price of the option. In our particular case we have that the barrier price is equal to 95 and the barrier is checked on a Monthly basis, i.e. Δt is equal to 1/12 of a year.

```
%% Price a Down&Out Call Option
%% Monthly monitoring
%% Merton model

%% Contract parameters (time horizon, strike price, barrier price, number of
%% time steps)
T=1; K=100; b=95; M=round(12*T);

%% same code as above ...

%% Sample S
S=Asset_Merton(Nsim,T,params,M,S0,r);
%% Compute the Discounted Payoff
% just like for max we use a dummy variable [] to compute the min across the
% rows, i.e. across the simulations
% the expression (min(S,[],2)>b) is a logical vector that is 1 if the
% condition is true and 0 otherwise
DiscPayoff=(exp(-r*T)*max( S(:,end)-K, 0)).*...
    (min(S,[],2)>b);
%% Compute the price
[Price,~,Price_CI]=normfit( DiscPayoff )
```

Todo list