Incomplete markets and pricing techniques

Lamrabet Anas

Research paper directed by Philippe Spieser

Key words
Quantitative finance for economists, stochastic processes, incomplete markets, pricing of vanilla European options, stochastic volatility, jumps, smile of volatility, Black & Scholes, Merton, Heston, Bates, Girsanov, changing of probability, Itô formula for general semimartingales, jump diffusions, finite activity processes, Markov processes, Lévy processes, Poisson processes, Lévy measures, compensated martingales, square variation, Fokker Planck equation, Lévy-Khinchin formula, Lévy-Itô decomposition, Monte Carlo, Fast Fourier Transform

ESCP-EAP / DESS IM EVRY

2004-2005

Abstract

The empirical evidence of the smile volatility in the markets of options has shown the limit of Black & Scholes modeling with constant local volatility functions in the frame of complete markets, especially since the crash of 1987. Despite the effort made to improve this frame of modeling, in particular by Dupire[1] in 1992, it remains unable to give a satisfactory explanation to market smile volatility, and then to provide traders with an efficient and robust pricing technique of non at the money options. Hence, one has decided to introduce new classes of models with stochastic volatility and/or jumps and a theoretical frame which takes into account the smile effect for currency options and the skew effect for options written on stocks. While B&S modeling failed to explain this positive correlation between large downwards movements of the underlying asset and options prices, Heston modeling with stochastic volatility assumes a negative link between returns and volatility (called the Merton effect) and models the so-called leverage effect observed in the market. At the other hand, Merton modeling with jumps perfectly describes this effect as the fear of the participants from large downward jumps in the price of stocks which results in a more important remuneration of options out of the money.

But empirically, we observed that one has to implement too high values of the volatility's drift within a stochastic volatility model to meet the term structure of the volatility observed in the market without in compensation any real economic interpretation. In an equivalent way, modeling with finite number of jumps – so-called finite activity Lévy processes or jump diffusion processes – doesn't permit to fit in data with a respectable precision (only a modeling with infinite activity Lévy processes would do so). The mixture of these two models due to Bates, a volatility driven by a stochastic diffusion (Heston model) and a drifted Brownian component associated with jumps (Merton model) gives at the opposite a plenty flexibility in fitting the smile and the term structure of volatility. It also displays a robust theoretical frame, as shown with Bates and Bakchi, Cao, Chen for the particular case of options written on currencies and on stocks. This is the reason why this paper deals above all with the model of Bates, which is both efficient in fitting market data and quiet representative

[1] *Dupire provided a mathematical formula for the calibration of the local volatility function with the market data and proved its uniqueness under complete markets assumption*

of incomplete markets because it includes the two major sources of its incompleteness, jumps and stochastic volatility.

The market is said incomplete within our frame of modeling because jumps and volatility are not tangible assets and are not traded in the markets. This incomplete market modeling leads to some drawbacks. Firstly, the uniqueness of risk neutral probability under which assets are priced is no more guaranteed. So as contrary to the Black & Scholes model, modeling with jumps and stochastic volatility holds the inconvenient not to give a perfect hedge against the variations of the underlying. Secondly, the incomplete market models lead to some numerical approximation of the prices of vanilla options. In effect, a closed form of the probability density is rarely available and we only know closed form of the characteristic function of the log price process, which leads to a semi-closed formula of options price computable by numerical approximation. But as it seems illusory to consider the market to be complete in the real world, incomplete market modeling are considered as an approximated solution to the right problem. It is then naturally more efficient than the exact solution to the wrong problem provided by Black and Scholes modeling.

From a technical point of view, the incomplete market models seem to involve much more large mathematical concepts than the complete market ones. Besides, the technical nature of most papers on the subject gives sometimes the impression to business schools students or others not specialized in applied mathematics for finance that jump diffusion processes are beyond their reach. This paper shows that this is not so and has been written mostly for this reason. But our aim is of course not to give an exhaustive and sophisticated mathematical frame for jump diffusions and other Lévy processes, which would be a quite technical subject and is beyond our reach, but rather an intuitive comprehension of the tools involved with jump diffusions modeling.

We won't deal with stochastic interest rates like in Backchi, Cao and Chen as a generalization of Bates' model into SVJD models (stochastic volatility jump diffusions models), but give the tools to let these models affordable to the reader. More sophisticated modeling with infinite activity Lévy processes like Variance-Gamma, Normal Inverse Gaussian, tempered stable and generalized hyperbolic models is not developed in this paper because infinite activity Lévy processes require more sophisticated mathematical material.

The interested reader can report to the annexes for an introduction to such processes through the Lévy-Khinchin representation and to the bibliography

Throughout all the paper, we provide the reader with mathematical tools necessary to tackle the development of the model of Bates. Even if most of the concepts are recalled, the reader must be familiar with Itô processes and continuous semimartingales to get plenty of this paper.

We begin with a financial approach of stochastic processes in the first part to go smoothly through jump processes in the second one. Then we learn to work with discontinuous stochastic calculus in the part III, before tackling the models in the part IV, and give an application of those model to the pricing of vanilla options in the part V. We eventually explain in the part VI how to implement the model of Bates with Monte Carlo simulations and FFT algorithms.

We will systematically privilege intuitive results than exact formalized results and we will purely and simply ignore many technical difficulties by giving when possible only intuitive idea of the mathematical assumptions. Our aim is not to give mathematical proofs of the results used in the construction of the model of Bates, but to provide an intuitive understanding of the tools used to construct this model. We apologize in advance for the lack of technical details (in particular convergence arguments and (square) integrability) which would have made this paper much less clear, and for possible mistakes the author is responsible for.

We only hope that this paper will help the reader who is familiar with continuous theory of stochastic calculus to feel comfortable with those more and more common models which involve discontinuity and stochastic volatility, and will stimulate its interest in recent developments in quantitative finance.

Contents

I – Mathematical frame for incomplete markets modeling

A- Financial modeling frame with stochastic processes in incomplete markets

We consider a sequence of random variables X_t, $t \in [0,+\infty[$ where each X_t represents the random price of a financial asset at the time *t*.

Each X_t can be a priori defined as an application from Ω_t in $\mathbb{R}^+$ (as prices are positive variables), for each atom ω in the case where Ω_t can be decomposed in single atoms:

$$\begin{array}{ccc} \Omega_t & \to & \mathbb{R}^+ \\ \omega & \mapsto & X_t(\omega) = x \end{array}$$

where Ω_t is the set of possible paths of the process (X_t) until the time *t,* and each ω a path of the process (X_t) until the time *t*.

In fact, X_t is an application from the $\sigma - algebra$ F_t in B($\mathbb{R}^+$), the $\sigma - algebra$ of $\mathbb{R}^+$

$$\begin{array}{ccc} F_t & \to & \mathrm{B}(\mathbb{R}^+) \\ A & \mapsto & X_t(A) = x \end{array}$$

The $\sigma - algebra$ F_t can be intuitively represented as a set of combinations of atoms ω of Ω_t , which include Ω_t, and which is stable under reunion, finite intersection and by passing to the complementary event. The complementary of an element of F_t must be in F_t, just as the union and the intersection of two elements of F_t. The smallest $\sigma - algebra$ is $F_t = \{\emptyset, \Omega_t\}$ and the largest one is $F_t = \mathrm{P}(\Omega_t)$, the set of all subsets of Ω_t.

The $\sigma - algebra$ B($\mathbb{R}^+$) can be seen as a stable set of combinations of open intervals of the form $[a,b[$. It is the smallest stable set of combination of intervals which can generate all $\mathbb{R}^+$.

We use this $\sigma - algebra$ to make X_t measurable, i.e

$$\forall x \in B(\mathbb{R}^+) \quad \exists A \in F : \; x = X_t^{-1}(A)$$

Intuitively, X_t measurable means that every realization x of X_t has its origin in a stable combination of atoms A. Thus, every (point or interval) x of $B(\mathbb{R}^+)$ can be written $x = X_t(A)$, which means that every price or range of prices observed in the market are the realization of some event A. So if X_t is measurable with respect to F_t (or X_t is F_t-measurable) then all values observed in $\mathbb{R}^+$ are images of some events of F_t.

If f is an application from B($\mathbb{R}^+$) to B($\mathbb{R}^+$) and f is measurable with respect to the $\sigma-algebra$ B($\mathbb{R}^+$), then f is said to be *Borelian.* We use such functions in practice to make the random variable $f(X_t)$ measurable in respect with the $\sigma-algebra$ F_t of Ω.

$$\begin{array}{ccccc} F_t & \overset{X_t \; measurable}{\rightarrow} & B(\mathbb{R}^+) & \overset{f \; Borelian}{\rightarrow} & B(\mathbb{R}^+) \\ w & \rightarrow & X_t & \rightarrow & f(X_t)(w) = f(X_t(w)) \end{array}$$

In practice, we construct the $\sigma-algebra$ F_t using the past values of X_t. For each s, the realization of all random variables X_s - the path of assets until the time t - permits to define the $\sigma-algebra$ $F_t = \sigma(X_s, s \le t)$, called the $\sigma-algebra$ generated by all the random variables X_s with $s \le t$. F_t is mathematically the smallest set of stable combinations of antecedents of X_s, $s \le t$, that generate all the values of X_s, $s \le t$, a set of antecedents of X_s being of the form { $X_s^{-1}(x)$, $x \in B(\mathbb{R}^+)$ }.

F_t is then a stable set of all realizations of the process (X_s) before the time t. So it represents the information revealed by all the paths of X_s before the time t. By construction, each X_t is "naturally" F_t-measurable. The we say in this case that $(X_t)_{t\in[0,+\infty[}$ is adapted with respect to the filtration $(F_t)_{t\in\mathbb{R}^+}$.

Basically, each atom of F_s is included in F_t for $s \le t$, as $\omega \in F_s$ is a scenario for the evolution of the assets' prices until the time s, while $\omega \in F_t$ is a scenario until the time t.

Then we have

$$F_0 \subset F_1 \subset ... \subset F_s \subset ... \subset F_t \subset ... \subset F_\infty \subset F$$

The sequence of $\sigma - algebra$ $(F_t)_{t \in \mathbb{R}^+} = F_0,..F_s;..,F_t,...,F_\infty$ is called a filtration.

On the so-called measurable space (Ω, F), where Ω is the set of all Ω_t, $t \in \mathbb{R}^+$, we define a measure of probability p, an application from Ω to [0,1] stable by reunion of disjoint sets which holds the conditions p(Ω) = 1. This probability is said to be historical if Ω is the set of scenarios of the price paths in the real world.

Associated to the filtration $(F_t)_{t \in \mathbb{R}_+}$ and the probability p, the measurable space (Ω, F) becomes a stochastic basis (Ω, F, $(F_t)_{t \in \mathbb{R}_+}$, p) under which all prices' processes are constructed. (X_t) is then modeled as an adapted stochastic process (i.e each X_t is F_t-measurable) on the stochastic basis (Ω , F, $(F_t)_{t \in \mathbb{R}_+}$, p),

We will rather use the filtration $F_t = \sigma(X_s, s \leq t) \wedge N$ than $F_t = \sigma(X_s, s \leq t)$, where N is the set of negligible events[2] and $\wedge$ the union operator for σ-algebras. The stochastic basis holds then the famous "common conditions". When modeling càdlàg processes (see below), the condition of right-continuousity of F_t [3] is required. The reader can refer to [2] for more details on the crucial role of information in modeling price paths with stochastic processes.

For a fixed scenario (an event w in F_t), each application $t \to X_t(\omega)$ is a trajectory of the stochastic process $(X_t)_{t \in [0,+\infty[}$.

A stochastic process is said càdlàg – *continu à droite limite à gauche* - if each trajectory is left continuous with right limits, i.e

[2] N is precisely the subset of negligible events i.e the set of A $\subset$ B such as P(B) = 0. Adding it to F_0 means that we want impossible trajectories to be known in advance at time zero and the impossibility of this events is an available information in every F_t.

[3] $F_t = F_{t^+}$ with $F_{t^+} = \bigcap_{\varepsilon > 0} F_{t+\varepsilon}$

$$\forall t \in \mathbb{R}^+ \quad \lim_{\varepsilon \to 0} X_{t+\varepsilon} = X_t \qquad \text{almost surely under the probability p}$$

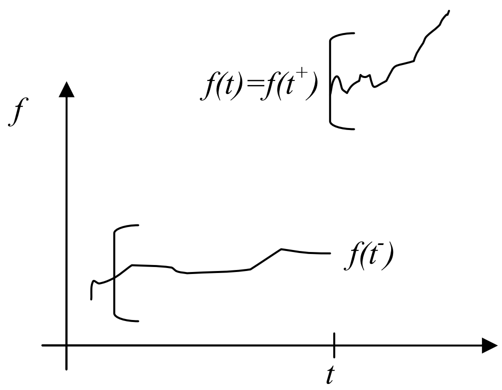
which means $\qquad X_{t^+} = X_t \qquad$ a.s. (almost surely under p)

Remark that such equalities which involve random variables are always written in the almost sure sense, without mentioning sometimes it is the case. The last equality means

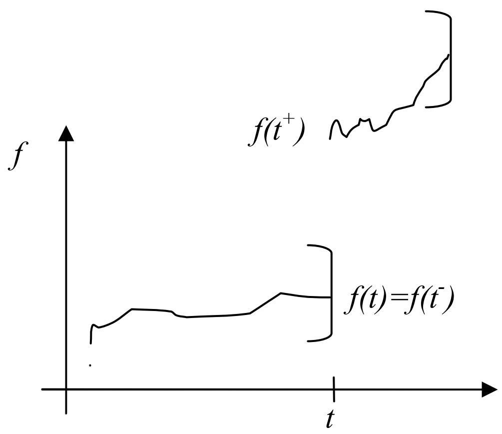
$$\forall w \in F_t \text{ such as } p(w) \neq 0, \ \forall t \in \mathbb{R}^+ \ \lim_{\varepsilon \to 0} X_{t+\varepsilon}(w) = X_t(w)$$

With càdlàg processes, jumps are unpredictable and are modeled as sudden events. In effect, the value X_t of the process at time *t* is always defined as the value after the jump $X_t = X_{t^+}$. As the jump is occurring after the time *t*, at t^+, the value X_t is not foreseeable by following the trajectory up to the time t. The unpredictability of jumps being an empirical fact in the market, we use these processes for modeling price dynamics.

At the opposite, càglàd processes (right continuous processes with left limits) are predictable processes. The value of the process X_t is the value before the jump, $X_t = X_{t^-}$, so we can foresee the value of X_t at time *t* by following the trajectory up to the time *t*. The càglàd processes are hence used to model dynamic strategies, where quantities of assets X_t to buy in period *[t, t+dt[* are known in advance.



càdlàg function
(unforeseeable jumps)



càglàd function
(previsible values)

A stopping time T is an application which gives the time of occurrence of an event, i.e

$$\begin{aligned} T: \quad F &\rightarrow \mathrm{B}(\Re^{+}) \\ \omega &\rightarrow T(\omega) = t \ \ for \ w \in F_t \end{aligned}$$

More precisely, T is said to be a stopping time[4] or a random time if the event $(T = t) \in F_t$, or equivalently $(T \leq t) \in F_t$. X_T is called a stopped process, and it takes values $X_{T(w)}(w)$ for every $w \in F$.

Given a σ-algebra G included in the σ-algebra F, we recall that the conditional expectation is the random variable defined by

$$\begin{aligned} \forall w \in F,\ E_p(X \mid G)(w) &= E_p(X \mid w) = E_{p_{|w}}(X) \quad \text{if } w \in G \\ &= E_p(X \mid \bar{w}) = E_{p_{|\bar{w}}}(X) \quad \text{if } w \notin G \end{aligned}$$

where the probability $p_{|w}$ is the conditional probability with respect to w, i.e

$$\forall A \in F,\ p_{|w}(A) = p(A \mid w) = \frac{p(A \cap w)}{p(w)}$$

and where the expectation of X under p is the average weighted value of X, i.e

$$E_p(X) = \int_{\Omega} X(w)dp$$

Remark that if we make the changing of variable $X(w) = x$, i.e $w = X^{-1}(x)$, we obtain

$$E_p(X) = \int_{\mathbb{R}} xdp(X^{-1}(x))$$

But

$$p(X^{-1}(x)) = p(\{w \mid X(w) = x) = p(X = x) = p_X(x)$$

[4] by reference to the definition of stopped processes used for the construction of local martingales.

$p_X(x)$ is called the distribution of X, and if $f_X(x) = \frac{dp_X(x)}{dx}$ is defined, it is called the density of X (or the Radon Nikodym derivative).

In this case,

$$E_p(X) = \int_{\mathbb{R}} x dp(X^{-1}(x)) = \int_{\mathbb{R}} x dp_X(x) = \int_{\mathbb{R}} x \frac{dp_X(x)}{dx} dx = \int_{\mathbb{R}} x f_X(x) dx$$

which is the usual definition of expectation with random variables which own a well defined density. With choosing any random variable $h(X)$ with h Borelian, and making the same reasoning, we show easily the transfer formula

$$E_p(h(X)) = \int_{\Omega} h(X)(w) dp = \int_{\mathbb{R}} h(x) f_X(x) dx$$

A Fourier transform of a function (or a measure) f is

$$\hat{f}(x) = \int_{\mathbb{R}} e^{i2\pi xy} f(y) dy$$

and the Inverse transform is

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i2\pi xy} \hat{f}(y) dy$$

The Fourier Transform of the density of a random variable is the characteristic function of the latest. With denoting f_X the density of the random variable X, its characteristic function is

$$\phi_X(u) = E(e^{iuX}) = \int_{\mathbb{R}} e^{iux} f_X(x) dx$$

Discounted prices are always modeled as martingales, as we can naturally think of the current price of a stock at time *t* as the sum of the futures market anticipations of the prices weighted objectively in a unique way (within complete markets) knowing the information revealed until this date. We remember that a process (X_t) is a martingale under the probability p (or a p-martingale) if its value at time *t* is its expectation at any future date *t+s* knowing the information revealed until the time *t,* i.e

$$E_P\ (X_{t+s} \mid F_t) = X_t$$

where E_P denotes the conditional expectation under the probability p [5].

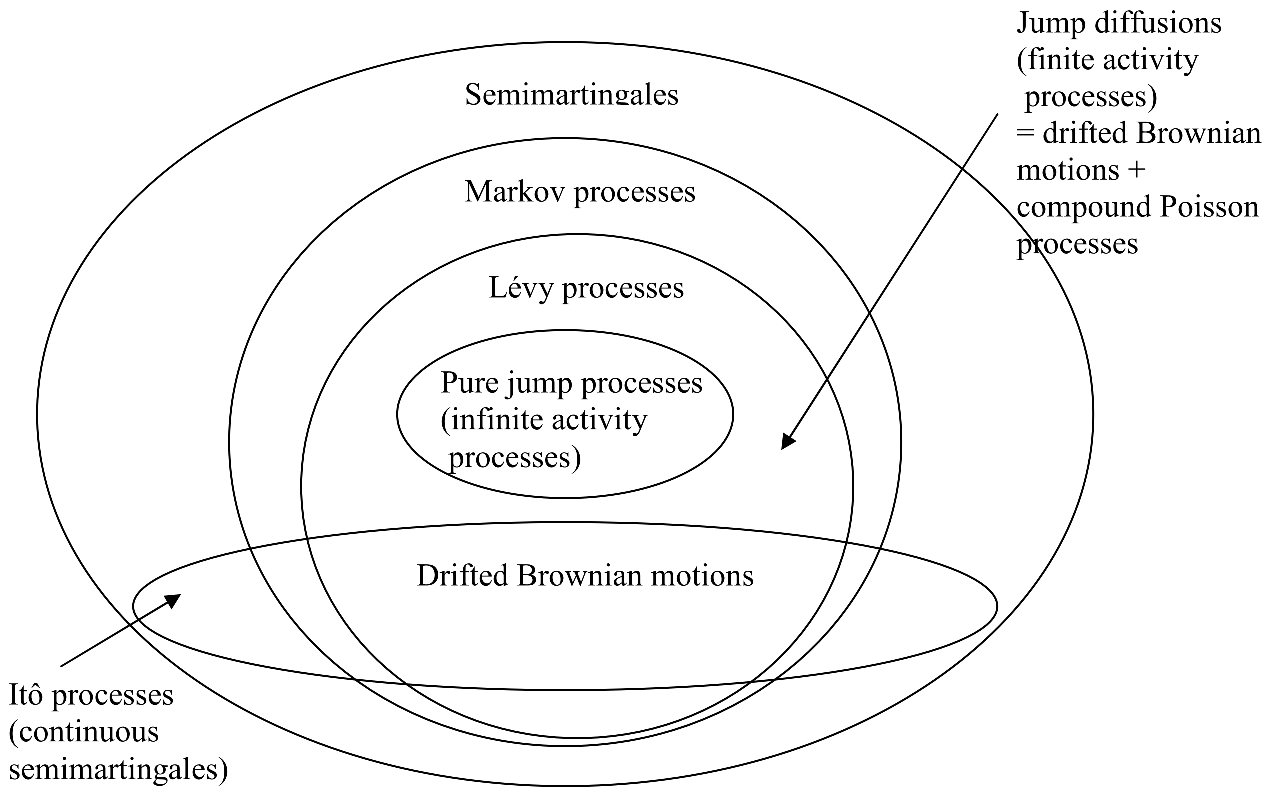
This is true only if the market actors agree for a unique perception of risk premium until they have to give the same weight for the different scenarios. The probability with such a distribution is called risk neutral probability. In almost all the cases, this probability differs from the historical probability. Under this probability, discounted prices of assets are martingales.

A complete market is a market where all sources of risks can be hedged perfectly, because of the possibility of replication of all assets. In this case, we can prove assuming there is no arbitrage opportunity that the risk neutral probability is unique (see [3]) . In an incomplete market, there is infinity of risk neutral probabilities under which discounted prices of assets are martingales. One has to use others criterions to discriminate between those probabilities. We'll develop this particular point in the part IV-C.

All stochastic processes used in the financial research are semimartingales. These are of two types : the continuous one, also called Itô processes and the discontinuous one, in which we find the family of pure Lévy processes (infinite activity process) and the jump diffusions (finite activity processes).

We'll see in the next part the reason why we choose this particular family of processes for financial modeling, before going trough the study of Lévy processes and jump diffusions.

[5] This expression must in fact have a sense so we suppose that X_t is integrable i.e $E_p(X) = \int_\Omega X(w)dp(w) < +\infty$


Jump diffusions
(finite activity
processes)
= drifted Brownian
motions +
compound Poisson
processes
Semimartingales
Markov processes
Lévy processes
Pure jump processes
(infinite activity
processes)
Drifted Brownian motions
Itô processes
(continuous
semimartingales)

B- Financial modeling with semimartingales

Let's suppose the market is composed of *n* assets, represented by the vector (S_t) = $(S_t^1, S_t^2, ..., S_t^n)$. Any strategy (Φ_t) is a sequence of vectors ($\Phi_t^1, \Phi_t^2, ..., \Phi_n^t$) which gives the composition of the portfolio V_t to which we apply the strategy at any time *t*, i.e

$$V_t(\Phi) = \sum_{k=1}^{n} \Phi_t^k S_t^k = \Phi_t . S_t \qquad \text{with the usual notation of scalar product}$$

A trading strategy consist to change ($\Phi_t^1, \Phi_t^2, ..., \Phi_n^t$) at any time *t* by selling and buying assets, without any external contribution of money. It means that the portfolio is self financed, i.e financed only by the growth or the decrease of the values of the assets which compose it.

Let's decompose the period of time [*0*,*T*] by discrete period of stopping times $t_0 = 0 \le t_1 \le ... \le t_{n-1} \le t_n = T$ to see what happens with the evolution of a self financed portfolio in the discrete case. The only information available at time t_i is F_{t_i}. We let then $t_0 = 0 \le t_1 \le ... \le t_{n-1} \le t_n = T$ be a partition of adapted random times which means that each t_i is F_{t_i}-measurable.

Reasonably, an investor should not know in advance at period of time $[t_{i-1}, t_i[$ what happens to assets prices during the period $[t_i, t_{i+1}[$, so he takes a decision to buy or sell assets for the next period $[t_i, t_{i+1}[$ (strategy Φ_{t_i}) with the information provided by the period $[t_{i-1}, t_i[$. At time of transaction t_i, the trader knows the new composition of the portfolio Φ_{t_i} without having any idea of the future evolution of assets until this moment. So he trades with the only information provided by $[t_{i-1}, t_i[$. If not, we can show very easily that strictly positive gains are almost sure with a zero initial investment, which would contradict hypothesis of non arbitrage opportunity.

Mathematically, it means that the trading strategy Φ_t at time t_i, i.e Φ_{t_i}, is $F_{t_{i-1}}$ - measurable. Thanks to this property, (Φ_t) is said to be a predictable process.

Thanks to self financing assumption, the portfolio value at time t is the sum of the plus-values made on the partition of random times plus the initial value i.e

$$V_{t_k}(\Phi) = V_0(\Phi) + \sum_{i=1}^{k} \Phi_{t_{i-1}} (S_{t_i} - S_{t_{i-1}})$$

We want now to generalize the stepwise trading strategy to a continuous one and give a sense to the expression

$$V_t(\Phi) = V_0(\Phi) + \int_0^t \Phi_{s^-} dS_s$$

When taking the partition to the limit[6], we see that the strategy Φ_t deployed at the period of time $[t \;, t^+[$ is defined at time t with the information provided by the period $[t^-, t[$.

By taking the limit at the right of the period of time $[t^-, t[$,

$$\lim_{s \to t^-} \Phi_s = \Phi_{t^-} ,$$

we define the strategy Φ_t which will be deployed in the period of time $[t \;, t^+[$. So the strategy Φ_{t^-} is equal to the strategy Φ_t. By analogy with the discrete case, Φ_t is $F_{t^-} - measurable$.

Φ_t is then predicted by Φ_{t^-}, i.e the value of Φ_t is defined at the time t^- by Φ_{t^-}, i.e

$$\Phi_t = \Phi_{t^-} .$$

(Φ_t) is then a càglàd process. So if we want to generalize the predictable trading strategy to the continuous case, we must deal with càglàd semimartingales.

Besides, the appropriate class of processes (Φ_t) and (S_t) we'll work with which must hold two fundamental properties of financial modeling :

[6] $\sup |t_k - t_{k-1}| \to 0$ a.s

- the martingale preserving property that guarantees that a portfolio of martingales is still a martingale
- the stability property which guarantees that a small change in the composition of the assets will lead only to a small change in the price of the portfolio

Martingale preserving property

From the discrete case, we can prove very easily the fact that if (Φ_t) is a predictable process and S_t is a martingale, then

$$V_t(\Phi) = V_0(\Phi) + \sum_{i=1}^{k} \Phi_{t_{i-1}} (S_{t_i \wedge t} - S_{t_{i-1} \wedge t}) \text{ is a martingale.}$$

This is a very important property, called the martingale preserving property. A portfolio of martingales must be a martingale and though the property of martingales must be stable by composition of portfolios with different assets.

Taken to the limits, $V_t(\Phi) = V_0(\Phi) + \sum_{i=1}^{k} \Phi_{t_{i-1}} (S_{t_i \wedge t} - S_{t_{i-1} \wedge t})$ might converge into

$$\int_0^t \Phi_{u^-} .dS_u = \int_0^t \Phi_u .dS_u$$

where (Φ_u) is a càglàd (semimartingale) strategy.

We use in fact the following result

If (Φ_t) *is a predictable process and* (S_t) *a martingale, then* $\int_0^t \Phi_t .dS_t$ *is a martingale.*

But we guess intuitively that any càdlàg semimartingale is the limit of a sequence of predictable processes. Then we can extend the property above to càdlàg semimartingales :

If (Φ_t) *is a càdlàg semimartingale and* (S_t) *a martingale, then* $\int_0^t \Phi_t .dS_t$ *is a martingale.*

This kind of generalization doesn't hold for this property only. It will be used implicitly in the sequel to generalize results holding with (Φ_t) predictable.

Stability property

Unfortunately, not all stochastic processes (Φ_t) and (S_t) guarantee this reasonable property. Small change of the composition of the portfolio (a change in Φ_t) which lead to a small change of

$$\int_0^t \Phi_t.dS_t$$

mathematically translates into :

If a sequence Φ_t^n of predictable strategies converges uniformly into Φ_t ,

then $\int_0^t \Phi_t^n.dS_t$ converges uniformly in probability into $\int_0^t \Phi_t.dS_t$.

The class of such processes that hold this property is the class of semimartingales. Even hard to identify with the property above, semimartingales can be characterized by the (non) unique decomposition into a finite variation process and a (local) martingale, a very technical result that we admit here. So it comes in particular that every finite variation process and every martingale are semimartingales. Besides, all Lévy processes are semimartingales : Poisson processes, Brownian motions etc… In fact, all processes encountered in financial modeling are semimartingales.

Let's point out that if (S_t) is a (semimartingale) Markov process, or more precisely a Lévy process, the stochastic integral with respect to (S_t) with a predictable (or more generally càdlàg semimartingale) process as integrand is not systematically a Markov process or a Lévy process anymore. But it is still a semimartingale. That is the reason why we work with this larger class of processes.

Semimartingales are the only framework for this kind of stability. They are more particularly stable with stochastic integration, i.e the following property is valid for càglàd semimartingales (Φ_t) :

If (S_t) is a semimartingale, then $\int_0^t \Phi_t.dS_t$ is a semimartingale.

Semimartingales provides a framework of stability under stochastic integration, linear transformations but also other smooth non linear transformations like changes of measures and changes of time. The product of two semimartingales is still a semimartingale (see [5]). Financial modeling is then naturally performed within the frame of semimartingale processes.

C- Lévy processes and Brownian motions

A Lévy process $(X_t)_{t\in[0,+\infty[}$ is a semimartingale with independent and stationary increments, left-continuous in probability, i.e. :

(i) $(X_{t+s} - X_t)$ is independent from F_t $\quad\quad$ ($F_t = \sigma(X_u, u \le t) \wedge N$)

(ii) $(X_{t+s} - X_t)$ has the same law than $X_s = (X_s - X_0)$ [7]

(iii) (X_t) is stochastically right continuous, i.e

$\forall\ \varepsilon > 0,\ \lim_{s\to t^+} \mathrm{p}(\mid X_s - X_t \mid > \varepsilon) = 0$ [8]

Property (i)

The first property means that the evolution of the price in the period [*t*, *t+s*] doesn't depend on the past values of the process before the time t. In equivalent way, for every increasing sequence of times $(t_1, t_2, ..., t_n)$ X_0, $X_{t_1} - X_0, X_{t_2} - X_{t_1}, \dots,$ $X_{t_n} - X_{t_{n-1}}$ are independent random variables. If a process holds this property, then it is said to be Markovian, or to be a Markov process. Intuitively, a Markovian price process depends on the past only by the present, i.e if a process is Markovian then

$$E(h(X_T) \mid F_s) = E(h(X_T) \mid \sigma(X_s))$$

If we know for instance at the moment *s* that $X_s = S$ on $A \in F$, i.e

$$A = (X = S) = \{w \in F, X(w) = S\}$$

then

$$E(h(X_T) \mid F_s)(A) = E(h(X_T) \mid X_s = S)$$

---

[7] By convention, $X_0 = 0$.

[8] The stochastic continuity is different and weaker that the continuity almost sure. The continuity almost sure is defined by p($\lim_{s\to t^-} X_s = X_t$) = 1.

There is many application of this property. The most direct one is the demonstration of the Black-Scholes formula for vanilla options (see annexes if necessary for the expression of X_T)

$$C_t = e^{-r(T-t)} E(h(X_T) \mid F_t) = e^{-r(T-t)} E(h(X_T) \mid X_t = S)$$

$$C_t = e^{-r(T-t)} E(h(X_t.e^{(\mu-\frac{1}{2}.\sigma^2)(T-t)+\sigma.W_{T-t}}) \mid X_t = S)$$

$$C_t = e^{-r(T-t)} E(h(S.e^{(\mu-\frac{1}{2}.\sigma^2)(T-t)+\sigma.W_{T-t}})$$

For a pay-off of a (borelian) type $h : x \to (x-K)^+$, the demonstration of the Black and Scholes formula is then possible with the manipulation of Gaussian variables.

Markov processes are used to model price paths because of an arbitrage argument in a perfect market. At the moment s, all information of the past is contained in the price X_s of the asset. If not, there will be an arbitrage opportunity because anyone can take a position on the asset and wait for the market to react to the new information.

Property (ii)

The second property deals with stability in distribution of the process. It equivalently means that for every Borelian function *f*, E($f\,(X_{t+s} - X_t)$) = E($f(X_s)$), where the symbol E denotes the expectation, by the definition of equality in distribution. The equality $X = Y$ in distribution is equivalent to $p_X(A) = p_Y(A)$ or to the equality of the respective Fourier transforms.

Property (iii)

With the process defined with property (i) and (ii) jumps can occur at both random times and deterministic times so we add the property of stochastic continuity to exclude processes with jumps times known in advance. The property of stochastic continuity merely means that jumps occur only at random times: given a deterministic time of jump *t*, the probability of the trajectory to be right-discontinuous in *t* converges to 0 and so there is no jump. Jumps known in advance can be regarded as calendar effects and do not match real evolution of the prices in the stock market.

We can show that there is a one to one correspondence between a Lévy process and its càdlàg representation. So we could have defined without loss of generality Lévy processes as càdlàg processes with property (i) and (ii), the property (iii) being a consequence of the almost sure right continuity[9].

We need the following theorem derived from the Itô-Lévy decomposition to make the link between Lévy processes and Brownian motions of the theory of continuous semimartingales.

Theorem

Let (X_t) be a Lévy process.

X_t is a continuous process if and only if it is a Gaussian process (or equivalently a drifted Brownian motion),

i.e $\exists a \in \Re$ *and* B_t *such as* $X_t = \delta t + B_t$, *with* B_t *a Brownian motion of variance t.*

The proof is simple starting from the Lévy-Itô decomposition (see annexes) and considering that the discontinuous part must be null. Consequently, the only continuous Lévy processes are the drifted Brownian motions. Though intuitive and very understandable, this theorem is based on the Lévy-Itô decomposition which is very technique and is at the core of the construction of all Lévy processes.

We will now study in deep Poisson processes to go further with jumps' modeling.

[9] The convergence almost sure implies the convergence in probability

II- Mathematical tools for financial modeling with jump processes

A- Poisson processes and Poisson measures

Let N_t a Poisson process with intensity λ.

$$
\begin{aligned}
N_t &= \sum_{k=1}^{Nt} 1 \\
&= \sum_{k=1}^{\infty} 1_{T_k \le t} \text{ [10]}
\end{aligned}
$$

where $(T_k)_{k\in[0,+\infty[}$ is an increasing sequence of exponential random variables such as $(T_{k+1}-T_k)$ are i.i.d. More precisely[11], we can show that in this case $T_k = \sum_{i=0}^{k} \tau_i$ where $(\tau_i)_{i\ge 0}$ is an i.i.d sequence of exponential variables with parameter λ.

$N_t = \sum_{k=1}^{\infty} 1_{T_k \le t}$ means that the Poisson process (N_t) jumps by the quantity 1 every time T_k until the time *t*.

Let's define the random Poisson measure M such as for $A \subset [0,T]$,

$$M(A) = \#\{\ k \ge 1 :\ T_k \in \mathrm{A}\}$$ [12]

where # means "the number of". *M(A)* is the number of jumps of the process N_t in the period of time *A*. *M* is defined as the jump measure of the Poisson process N_t. This measure is bounded (because of the finite number of jumps) and additive. We then have by definition

$$N_t = M([0, t])$$

[10] The indicating measure is the application from Ω to $\{0,1\}$ such as $1_A(w) = 1$ if $w \in A$ and *0* elsewhere.

[11] Thank you Willy for this remark !

[12] It of course means that $M(w,A) = \#\{\ k \ge 1 :\ T_k(w) \in \mathrm{A}\}$ for *w such as* $p(w) \neq 0$

We point out also that

$$N_t - N_s = M([0,t]) - M([0,s])$$
$$= \#\{\ k \geq 1 :\ T_k \in [0,t]\} - \#\{\ k \geq 1 :\ T_k \in [0,s]\}$$
$$= \#\{\ k \geq 1 :\ T_k \in [s,t]\}$$
$$= M([s,t])$$

which is a mere consequence of additivity.

By using this property on the subdivision $t_0 = 0 \leq t_1 \leq ... \leq t_{n-1} \leq t_n = t$,

$$N_t = \sum_{i=0}^{n-1} M(t_{i+1} - t_i)$$

By taking $n \to \infty$, i.e taking the subdivision to be infinitesimal,

$$N_t = \int_0^t M(ds) \qquad \text{(also denoted } \int_0^t dM(s)\text{)}$$

which is the random variable $w \to N_t(\mathrm{w}) = \int_0^t M(w, ds)$ to which we associate the random Poisson measure M :

$$\begin{array}{ccc} \Omega \times B([0,t]) & \to & \Re^+ \\ (w,A) & \mapsto & M(w,A) \end{array}$$

Poisson measure of compound Poisson processes

Let's consider a process J_t whose jumps' size are Y_k instead of 1 for simple Poisson processes.

$$J_t = \sum_{k=1}^{N_t} Y_k$$

J_t is called a compound Poisson process of jumps' size Y_k.

We want now the measure *M(B)* to be the total number of jumps of the process J_t whose size is equal to some element of *B*. Obviously,

$$M(B) = \sum_{k=1}^{Nt} 1_{Y_k \in B} = \sum_{k=1}^{Nt} 1_{Y_{T_k} \in B} \cdot 1_{T_k \le t} = \sum_{k=1}^{+\infty} 1_{Y_{T_k} \in B} \cdot 1_{T_k \le t}$$

M(*B*) can be considered as *M*([*0,t*],*B*), i.e the number of jumps of J_t whose size Y_k is equal to some element of *B* for the whole period of time [*0,t*].

Let's define the Poisson measure *M* such as for $A \subset [0,t]$, $B \in B(\mathbb{R}^*)$,

$$M(A, B) = \#\{\ k \ge 1 : \ T_k \in A, \ Y_k \in B\}$$

M is the number of jumps of the period of time *A*, which sizes belongs to the set *B*. It is the jump measure associated to the compound Poisson process J_t.

For example, $M([0,t], \{x\}) = \#\{\ k \ge 1 : \ T_k \in [0,t], \ Y_k = x\}$ is the number of jumps of the process J_t in the period of time [*0*,*t*] whose size is equal to *x*.

By the same reasoning than for simple Poisson processes,

$$\begin{aligned} M([0,t], \mathbb{R}^*) &= \#\{\ k \ge 1 : \ T_k \in [0,t], \ Y_k \in \mathbb{R}^* \} \\ &= \#\{\ k \ge 1 : \ T_k \in [0,t]\} \qquad \textit{as } Y_k \in \Re^* \text{ is always true} \end{aligned}$$

$$= \sum_{k=1}^{Nt} 1$$

$$= N_t$$

Then,

$$N_t = M([0,t], \mathbb{R}^*)$$

$$= \sum_{i=0}^{n} M(t_{i+1} - t_i, \mathbb{R}^*) \qquad \text{by left additivity}$$

$$= \sum_{j \in Z} \sum_{i=0}^{n} M(t_{i+1} - t_i, x_{j+1} - x_j,) \qquad \text{by right addivity}$$

$$= \int_{\mathbb{R}^*} \int_0^t M(ds, dx)$$

M(ds,dx) is the number of jumps of the process J_t during the period of time *ds* whose amplitude is *dx.*

Let's end this section with a remark. We can see that the intensity of a Poisson process its expectation per unit of time (per year).

$$E(N_t) = E(M(\mathbb{R}, [0,t]) = \lambda.t = \mu(\mathbb{R}).t$$

where $\mu(.)$ is the intensity of the process. The intensity λ of a Poisson process can be then interpreted as the average number of jumps per year.

We know $M([0,t], B)$ is a simple Poisson process for each $B \in B(\mathbb{R}^*)$, so we can define a Poisson process $M_t(B) = M([0,t], B)$ with variable intensity $E(M_t(B)) = \mu(B)t$.

This concept is useful for the construction of Poisson processes. The reader can report to the annexes for the construction of Poisson processes from a any Radon measure μ.

Marked Poisson processes (*)

Consider now the process $J = \sum_{n\geq 1, T_n \in [0,t]} f(T_n, Y_n) = \sum_{n\geq 1}^{N_t} f(T_n, Y_n)$

J is said to be the marked Poisson process of the sequence $(T_n, Y_n)_{n\geq 1}$ where

- (T_n) is an increasing sequence of stopping times with $T_n \underset{n\to\infty}{\overset{a.s}{\to}} \infty$ (the moments of jumps).

- Each value Y_n is revealed at T_n i.e each jump size Y_n is $F_{T_n} - measurable$.

The first condition is necessary to mean that for $T_n \leq t$, the number of jumps is a.s finite in the period of time [*0,t*]. The second means that the jump size is known only during the moment of jump (unpredictability of jumps).

We need then to prove the following property

If $\quad J = \sum_{n\geq 1}^{N_t} f(T_n, Y_n) \quad$ *then* $\quad J = \int_{\mathbb{R}^+} \int_0^t f(s,x) M(ds, dx)$

It is sufficient to prove this result for a stepwise function *f*, as every continuous function *f* can be approached by a convergent sequence of stepwise functions f_n.

Let's begin with a remark. Suppose that a function takes values $a_{i,j}$ on the set $[t_i, t_{i+1}[\times[x_j, x_{j+1}[$, i.e f is a stepwise function $f(s,x) = \sum_i \sum_j a_{i,j}.1_{(s,x)\in[t_i,t_{i+1}]\times[x_{j,},x_{j+1}]}$

Then by definition, as for the construction of the Stieltjes integrals,

$$\int_{\mathbb{R}^+} \int_0^t f(s,x) M(ds, dx) = \sum_i \sum_j a_{i,j} M([t_i, t_{i+1}[,[x_j, x_{j+1}[\ )$$

We assume now that *f* is piecewise and takes values $a_{i,j}$ on the set $[t_i, t_{i+1}[\times[x_j, x_{j+1}[$.

$$J = \sum_{n\geq 1} f(T_n, Y_n).1_{(T_n \in[0,t])}$$

$$J = \sum_{n\geq 1} f(T_n, Y_n).1_{T_n\in[0,t]}.1_{Y_n\in\mathbb{R}} \quad \text{(because } 1_{Y_n\in\mathbb{R}} \text{ is always equal to 1)}$$

$$J=\sum_{n\geq 1} f(T_n,Y_n).1_{(T_n,Y_n)\in[0,t]\times\mathbb{R}}$$ [13]

Considering the partitions ($[x_i,x_{i+1}[\ )_{i\in I}$ of $\mathbb{R}$, and ($[t_j,t_{j+1}[\ )_{j\in J}$ of $[0,t]$,

$$J=\sum_{n\geq 1}\ (\sum_{i\in I}\sum_{j\in J} f(T_n,Y_n).1_{(T_n,Y_n)\in[t_i,t_{i+1}[\times[x_j,x_{j+1}[})$$

And by changing the order of summing,

$$J=\sum_{i\in I}\sum_{j\in J}(\sum_{n\geq 1}\ f(T_n,Y_n).1_{(T_n,Y_n)\in[t_i,t_{i+1}[\times[x_j,x_{j+1}[})$$

$$J=\sum_{i\in I}\sum_{j\in J}(\sum_{n\geq 1}\ a_{i,j}.1_{(T_n,Y_n)\in[t_i,t_{i+1}[\times[x_j,x_{j+1}[})$$

$$J=\sum_{i\in I}\sum_{j\in J}\ a_{i,j}(\sum_{n\geq 1}1_{(T_n,Y_n)\in[t_i,t_{i+1}[\times[x_j,x_{j+1}[})$$

By definition, $\sum_{n\geq 1}1_{(T_n,Y_n)\in[t_i,t_{i+1}[\times[x_j,x_{j+1}[}=M([t_i,t_{i+1}[,[x_i,x_{i+1}[\ )$, because it's the number of jumps whose size belongs to $[x_j,x_{j+1}[$ during the period of time $[t_i,t_{i+1}[$.

So $$J=\sum_{i\in I}\sum_{j\in J}\ a_{i,j}.M([t_i,t_{i+1}[,[x_i,x_{i+1}[\ )$$

and finally, by assuming the convergence

$$J=\int_{\mathbb{R}^+}\int_0^t f(s,x)M(ds,dx)$$

[13] $1_{A\cap B}=1_A.1_B$

B- Construction and identification of jump diffusion processes

Let (X_t) a càdlàg Lévy process. We basically assume that the number of random times of jump is finite. If it were not, $(X_t)_{t\in[0,+\infty[}$ could not be a càdlàg process and be right-continuous at any point.[14]

Please note that from now on we denote $\lim_{s\to t^-} X_s = X_{t^-} = X_t^-$ and $\Delta X_t = X_t - X_t^-$.

ΔX_t is the amount of the discontinuous jump at time *t*. In continuous part of the trajectory, we have obviously $\Delta X_t = 0$. The jumps occurs on [*0*,*t*] at times T_k , k $\in$ [*0*,*n*], which means that $\forall k \in [|0,n|]$, $\Delta X_{T_k} = X_{T_k} - X_{T_k^-} \neq 0$.

Let $N_t = \#\{$ k with $\Delta X_{T_k} \neq 0$ and $T_k \le t \} = \#\{$ s $\in$ [0,t] such as $\Delta X_s \neq 0 \}$ where # denotes "the number of". N_t is the number of jumps occurring in period [0,*t*], and the random times T_k are the times of jump.

The random times are increasing, and identically independently distributed (i.i.d) of exponential law with some parameter λ, as we can prove that the time of occurrence of an event which probability of occurrence is uniform in a given interval follows an exponential distribution.

N_t, the total number of jumps is then naturally

$$\sum_{k=1}^{\infty} 1_{(T_k \le t)}$$

so it's a Poisson process of intensity λ

By obvious summing and taking a subdivision $t_0 = 0 \le t_1 \le ... \le t_{n-1} \le t_n = t$,

---

[14] In fact, we will see in the annexes that X_t can still be a Lévy process if there is an infinity of jumps whose size is around 0. The right condition is then : there is a finite number of jumps whose size is not close to 0, for example greater than 1. The reader can report to the annexes in the Lévy-Khinchine section for more details.

$$X_t = \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}}) + X_0$$

We separate the sum taken with n $\rightarrow \infty$ into two parts : the left continuous part where

$$X_{t_k} - X_{t_{k-1}} \xrightarrow[n \to \infty]{} 0$$

and the jumps occurring at random times $(T_k)_{k \in [1,n]}$ where

$$\Delta X_{T_k} \neq 0$$

By taking n $\rightarrow \infty$, i.e taking the subdivision to be infinitesimal, we obtain

- the continuous parts of X_t is equal to some expression denoted X_t^C
- and the discontinuous part to $\sum_{k=1}^{N_t} \Delta X_{T_k}$ where.

Thus we have $$X_t = X_t^C + \sum_{k=1}^{N_t} \Delta X_{T_k}$$

The continuous part of X_t is obviously a continuous Lévy process. It can be then expressed as a drifted Brownian motion, i.e

$$\exists \ (\gamma, \sigma) \in \mathbb{R} \times \mathbb{R}^+ \text{ such as } \ X_t^C = \gamma.t + \sigma.B_t .$$

The discontinuous part is then a compound Poisson process J_t

$$J_t = \sum_{k=1}^{N_t} \Delta X_{T_k}$$

Let's denote $Y_k = \Delta X_{T_k}$, jump size of the distribution X_t. Then J_t is a compound Poisson process with jumps' size distribution $(Y_k)_{k \in [0,n]}$, i.e

$$J_t = \sum_{k=1}^{N_t} Y_k$$

Hence, we have constructed a càdlàg Lévy process, in which a finite number of jumps occur at random times, called also as a jump diffusion

$$X_t = \gamma.t + \sigma.B_t. + J_t$$

Conversely, we often model under risk neutral probability the dynamics of the logarithm of prices of an asset with the jump diffusion

$$X_t = \gamma.t + \sigma.B_t. + J_t$$

where B_t is a Brownian motion and J_t a compound Poisson process of jumps Y_i. See section IV for a direct application with the model of Merton.

Unfortunately, we don't have the equivalence with the representation of all Lévy processes. The interested reader can report to the annexes to see that any Lévy process can be decomposed rather with subtracting additionally the expectation of the jumps smaller than 1 in absolute value

$$X_t = \gamma.t + \sigma.B_t + \sum_{k=1}^{N_t} \Delta X_{T_k} - E(\sum_{k=1}^{N_t} \Delta X_k.1_{|\Delta X_k|<1})$$

Marked point processes and jumps measure representation of X_t

Consider the random measure

$$M = \sum_{n\geq 1} 1_{(T_n, Y_n)}$$

i.e $$\forall \ (A,B) \subset [0,t]\times\mathbb{R}^*, \ \ M(A,B) = \sum_{n\geq 1} 1_{(T_n \in A, Y_n \in B)} \ ,$$

The process takes the value 1 if there exists a moment T_n such as the size of jumps is in the range *B* during the period *A*. So it merely counts the number of jumps in the period *A* whose amplitude belongs to *B*.

We denote J_X the jump measure associated to the jump diffusion

$$X_t = \gamma.t + \sigma.B_t + \sum_{k=1}^{N_t} \Delta X_k$$

i.e J_X is associated to the discontinuous part of X_t, which is the compound Poisson process

$$J_t = \sum_{k=1}^{N_t} \Delta X_k$$

Then *M(ds,dx)* defined in the precedent sections is precisely $J_X(ds,dx)$, the number of jumps of X_t during the period *ds* whose amplitude is *dx*, . The application J_X is then called the jump measure of X_t and is equal to the Poisson measure of the compound Poisson process J_t .

$$J_X(w,A,B) = \sum_{n\geq 1} 1_{(T_n(w)\in A, \Delta X_n(w)\in B)}$$

and in particular $$J_X(w,[0,t],x) = \sum_{n\geq 1} 1_{(T_n(w)\in[0,t], \Delta X_n(w)=x)}$$

and so $$J_X([0,t],x) = \sum_{s\in[0,t]}^{\Delta X_s \neq 0} 1_{(s, \Delta X_s = x)}$$

J_X is clearly the number of jumps between 0 and *t* whose amplitude is equal to *x*.

With keeping in mind that

$$J_X([0,t],B) = \int_0^t \int_B J_X(ds,dy)$$

we consider now the process of the form (marked point process)

$$J = \sum_{n\geq 1, T_n \in [0,t]} f(T_n, Y_n)$$

We showed

$$J = \int_0^t \int_{\mathbb{R}^*} f(s,y).M(ds,dy)$$

With setting $f(T_n, Y_n) = Y_n$, we can see that compound Poisson processes can be represented as follow

$$J_t = \sum_{k=1}^{N_t} \Delta X_k = \int_0^t \int_{\mathbb{R}^*} y.J_X(ds,dy)$$

So we can represent any jump diffusion as follow

$$X_t = \gamma.t + \sigma.B_t + \int_0^t \int_{\mathbb{R}^*} y.J_X(ds,dy)$$

More generally, any Lévy process can be decomposed as

$$X_t = \gamma.t + \sigma.B_t + \int_0^t \int_{\mathbb{R}^*} y.J_X(ds,dy) - \int_0^t \int_{|x|<1} y.v(dy)ds$$

where we used the Lévy measure v defined in the next section and the jump measure J_X of X_t.

C- Lévy measure and compensated martingales

We introduce in this section the definition of the Lévy measure which is at the core of the famous Lévy-Khinchin formula[15] and the construction of all jump processes.

Let (X_t) *be a Lévy process on* $\mathbb{R}$.
The measure ν *on* $\Re$ *defined by*
$\forall\ B \in B(\mathbb{R}),\qquad \nu(B) = E(\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in B\})$
is called the Lévy measure.

This measure is clearly linear with respect to the time.

$$E(\#\{t \in [0,2] : \Delta X_t \neq 0, \Delta X_t \in B\}) = E(\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in B\}) + E(\#\{t \in [1,2] : \Delta X_t \neq 0, \Delta X_t \in B\})$$

$$E(\#\{t \in [0,2] : \Delta X_t \neq 0, \Delta X_t \in B\}) = 2.E(\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in B\})$$

$\nu(B)$ is the expected number, per unit of time, of jumps whose size belongs to *B*. In practice, one unit of time is one year in finance modelling. So the average (the expectation of the) number of jumps whose size belong to *B* for a period of time *ds* is *ν(B)ds*. For instance, the average number of jumps whose size belongs to *B* for one month is

$$\frac{1}{12}\nu(B)$$

We see then clearly that by definition

$$E(M(ds,dx)) = E(\#\{t \in [0, ds] : \Delta X_t \neq 0, \Delta X_t \in dx\}$$
$$E(M(ds,dx)) = E(\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in dx\}.\ ds$$
$$E(M(ds,dx)) = \nu(dx).ds$$

[15] We give an intuitive demonstration of this formula in annexes with a pedagogical aim to help the reader to go further with modelling with jumps

Then $\int_0^t E(M(ds,dx)) = E(\int_0^t M(ds,dx)) = E(M([0,t],dx) = \int_0^t v(dx)ds = tv(dx)$

We obtain then $E(M([0,t],dx) = tv(dx)$

Let's work now only with jump diffusions of the form

$$X_t = \gamma.t + \sigma.B_t + J_t$$

and let *M* be the jump measure of X_t, i.e $M([0,t], dx)$ is the number of jumps ΔX_{T_k} of X_t whose size belongs to *dx* in the whole period [*0*,*t*] , i.e

$$M([0,t],dx) = \sum_{k=1}^{N_t} 1_{(\Delta X_{T_k} \in dx)} 1_{(T_k \le t)}$$

$$M([0,t],dx) = \sum_{k=1}^{N_t} 1_{(\Delta X_{T_k} \in dx)}$$ [16]

Let's denote $Y_k = \Delta X_{T_k}$, and *F* the distribution of the jumps Y_k , then

$$E(M([0,t],dx)) = E(\sum_{k=1}^{N_t} 1_{(\Delta X_{T_k} \in dx)})$$

$$E(M([0,t],dx)) = \sum_{n=0}^{\infty} E(\sum_{k=1}^{N_t} 1_{(\Delta X_{T_k} \in dx)} \Big| N_t = n) p(N_t = n)$$

$$E(M([0,t],dx)) = \sum_{n=0}^{\infty} E(\sum_{k=1}^{n} 1_{(\Delta X_{T_k} \in dx)}) p(N_t = n)$$

By independence of the jumps Y_k ,

$$E(M([0,t],dx)) = \sum_{n=0}^{\infty} \sum_{k=1}^{n} E(1_{(\Delta X_{T_k} \in dx)}) p(N_t = n)$$

And remembering they are i.i.d of distribution *F*,

[16] as $(k \le N_t)$ is equivalent to $(T_k \le t)$

$$E(M([0,t],dx)) = \sum_{n=0}^{\infty}\sum_{k=1}^{n} p(\Delta X_{T_k} \in dx)p(N_t = n)$$

$$E(M([0,t],dx)) = \sum_{n=0}^{\infty}\sum_{k=1}^{n} p(Y_k \in dx)p(N_t = n) = \sum_{n=0}^{\infty}(\sum_{k=1}^{n} F(dx))p(N_t = n)$$

$$E(M([0,t],dx)) = F(dx)\sum_{n=0}^{\infty} np(N_t = n) = F(dx)E(N_t)$$

$$E(M([0,t],dx)) = \lambda tF(dx)$$

So we obtain an important property thanks to

$$E(M([0,t],dx)) = tv(dx) = \lambda tF(dx)$$

The Lévy measure of a jump diffusion process is equal to the intensity of jumps by the distribution of the jumps i.e

$$v(dx) = \lambda F(dx)$$

Characteristic function of a compound Poisson process

This is a particular representation of Lévy processes when the activity is finite, i.e the number of big jumps is finite. The following formula holds in fact for the Lévy-Khinchin formula for compound Poisson processes.

Let X_t be a compound process of intensity λ and jumps' sizes Y_i of distribution F, i.e

$$p(Y_i \in A) = \int_A dF(x) = F(A)$$

then

$$E(e^{iuX_t}) = e^{t\int_{\mathbb{R}} (e^{iux}-1)\nu(dx)} = e^{t\lambda\int_{\mathbb{R}} (e^{iux}-1)F(dx)}$$

In effect, with denoting $\hat{F}$ the characteristic function of Y_1,

$$E(e^{iuX_t}) = \sum_n E(e^{iu\sum_{k=1}^{N_t} Y_k} \mid N_t = n)$$

$$E(e^{iuX_t}) = \sum_n E(e^{iu\sum_{k=1}^{n} Y_k})p(N_t = n) = \sum_n (\prod_{k=1}^{n} E(e^{iuY_k}))p(N_t = n)$$

$$E(e^{iuX_t}) = \sum_n (E(e^{iuY_1})^n p(N_t = n))$$

$$E(e^{iuX_t}) = e^{-\lambda t}\sum_n (\hat{F}(u)^n \frac{(\lambda t)^n}{n!}) = e^{-\lambda t} e^{\lambda t\hat{F}(u)}$$

With remarking that $\int_{\Re} F(du) = 1$,

$$E(e^{iuX_t}) = e^{\lambda t(\hat{F}(u)-1)} = e^{t\lambda(\int_{\mathbb{R}} e^{iux}F(dx)-1)}$$

$$E(e^{iuX_t}) = e^{t\lambda(\int_{\mathbb{R}} (e^{iux}-1)F(dx)}$$

The Lévy-measure of the compound process X_t being $\nu(dx) = \lambda F(dx)$, then

$$E(e^{iuX_t}) = e^{t(\int_{\mathbb{R}} (e^{iux}-1)\nu(dx)}$$

Compensated martingales

Let's begin with an example. Consider $N_t - \lambda t$ where N_t is a Poisson process with intensity λ, and let's take the filtration $F_s = \sigma(N_u, u \le s)$.

Then

$$E(N_t - \lambda t \mid F_s) = E(N_t \mid F_s) + \lambda t = E((N_t - N_s) + N_s \mid F_s) + \lambda t$$

$$E(N_t - \lambda t \mid F_s) = E(N_t - N_s \mid F_s) + E(N_s \mid F_s) + \lambda t$$

with considering that $N_t - N_s$ is independent from F_s and considering that N_s is F_s-measurable. So

$$E(N_t - \lambda t \mid F_s) = E(N_t - N_s) + N_s + \lambda t = \lambda(t-s) + \lambda t + N_s$$

$$= N_s - \lambda s$$

$N_t - \lambda t$ is then a martingale called the compensated martingale with respect to N_t.

Moreover, we can see trough this example that every Lévy process minus its expectation is a martingale, as we used only arguments of stationarity and independence of increments. In effect, λt is merely $E(N_t)$ and we have demonstrated implicitely that $N_t - E(N_t)$ is a martingale. We have then the following property.

For every Lévy process X_t *,*

$X_t - E(X_t)$ *is a martingale called the compensated martingale of the process* X_t.

Besides, $(X_t)^2 - E(X_t^2)$ *is a martingale.*

Consider now a marked point process of the form

$$J = \sum_{n \ge 1, T_n \in [0,t]} f(T_n, Y_n) \qquad = \int_0^t \int_{\mathbb{R}^*} f(s,y).M(ds,dy)$$

By linearity, we have

$$E(J)=\int_0^t\int_{\mathbb{R}^*} f(s,y)E(M(ds,dy))$$

We have seen $E(M(ds,dy)) = v(dy).ds$

Then $$E(J) = \int_0^t\int_{\mathbb{R}^*} f(s,y)v(dy)ds$$

So we obtain that

$$\int_0^t\int_{\mathbb{R}^*} f(s,y).M(ds,dy)-\int_0^t\int_{\mathbb{R}^*} f(s,y)v(dy)ds$$

is a martingale.

In particular for a compound Poisson process $J_t=\sum_{k=1}^{N_t} Y_k$,

$$J_t-E(J_t)=\int_0^t\int_{\mathbb{R}^*} y.J_X(ds,dy)-\lambda\int_0^t\int_{\mathbb{R}^*} y.F(dy)=\int_0^t\int_{\mathbb{R}^*} y.(J_X(ds,dy)-\lambda F(dy))$$

is a martingale which we can express in a simpler form. In effect, let's assume $E(Y_k)=E(Y_1)=\bar{k}$. Then $J_t-\lambda.\bar{k}.t$ is a martingale

In effect,

$$E(J_t)=\sum_{n=1}^{\infty} E(\sum_{k=1}^{n} Y_k)p(N_t=n)= \sum_{n=1}^{\infty}(\sum_{k=1}^{n} E(Y_k))p(N_t=n)$$

Y_k are i.i.d so

$$\sum_{k=1}^{n} E(Y_k) = n.E(Y_1)$$

Then

$$E(J_t)=E(Y_1)\sum_{n=1}^{\infty} n.p(N_t=n)$$

$$E(J_t)=E(Y_1)E(N_t)=E(Y_1).\lambda t=\bar{k}\lambda t$$

III – Stochastic calculus for semimartingales and jump diffusions

A- A brief review of Itô formula for continuous semimartingales

We denote obviously Itô processes

$$S_t = S_0 + \int_0^t \mu(s, S_s).ds + \int_0^t \sigma(s, S_s).dW_s$$

or, by obvious summing

$$\int_0^t dS_s = \int_0^t \mu(s, S_s).ds + \int_0^t \sigma(s, S_s).dW_s$$

with the simplified "differential notations"

$$dS_t = \mu(t, S_t).dt + \sigma(t, S_t).dW_t$$

In order to make the ideas clear, let's remember that a continuous semimartingale (also called an Itô process) S_t has a general form

$$dS_t = \mu(t, S_t).dt \ + \ \sigma(t, S_t).dW_t \qquad (1)$$

with μ integrable and σ square-integrable (i.e $\in L^2(\Omega)$ [17]), we have for any bounded function f C^1 in time and C^2 in space,

$$df(t, S_t) = \ f_t'(t, S_t).dt \ + \ f'_x(t, S_t).dS_t + \frac{1}{2}.f''_{xx}(t, S_t).d < S, S >_t$$

where we use the abbreviated notations

$$f_t^{'} = \frac{\partial f(t,x)}{\partial t}, \ f_x^{'} = \frac{\partial f(t,x)}{\partial x} \ \ and \ f_{xx}^{''} = \frac{\partial^2 f(t,x)}{\partial x^2}$$

The predictable variation $< X, Y >_t$ is the only increasing process such as $X_t Y_t - < X, Y >_t$ is a martingale. So the predictable variation $< S, S >_t$ is the only increasing process such as

[17] $L^2(\Omega)$ is the Hilbert space of square integrable random variable with the norm $\|X\|^2 = \int_\Omega X^2(w)dP(w) = E(X^2)$

$S_t^2 - < S,S >_t$ is a martingale. For Itô processes, the quadratic variation $d < S,S >_t$ is the process

$$d < S,S >_t = \sigma^2(t,S_t)dt .$$

The reader can check this result with manipulating

$$d < S,S >_t = < dS_t, dS_t >$$

as a bilinear form, and considering that the scalar product of two deterministic functions and the scalar product of a deterministic function and a Brownian motion is null.

$$
\begin{aligned}
& d < S,S >_t = < dS_t, dS_t > \\
& d < S,S >_t = < \mu(t,S_t).dt + \sigma(t,S_t).dW_t, \mu(t,S_t).dt + \sigma(t,S_t).dW_t > \\
& d < S,S >_t = \mu^2(t,S_t) < dt,dt > + \sigma^2(t,S_t). < dW_t, dW_t > \\
& \qquad + 2\sigma(t,S_t)\mu(t,S_t) < dt, dW_t > \\
& d < S,S >_t = \sigma^2(t,S_t). < dW_t, dW_t >
\end{aligned}
$$

Knowing $W_t^2 - t$ is a martingale, we conclude that $< W,W >_t = < W_t, W_t > = t$ and so $d < W,W >_t = < dW_t, dW_t > = dt$.

We obtain then, the famous Itô formula in time and space for continuous semimartingales of the form :

$$df(t,S_t) = f_t'(t,S_t).dt + f'_x(t,S_t).dS_t + \frac{1}{2}.f''_{xx}(t,S_t).\sigma^2(t,S_t).dt$$

and the Itô formula in space for bounded C^2 functions f :

$$df(S_t) = f'_x(S_t).dS_t + \frac{1}{2}.f''_{xx}(S_t).\sigma^2(t,S_t).dt$$

The readers who desire to recall how to use this formula can report to the annexes for a little application.

We'll use in this paper the Itô lemma in the case where S_t depends on two Brownian motions. For instance, we typically have for stochastic volatility models

$$\frac{dS_t}{S_t} = \mu(S_t,t)dt + \sigma(V_t,t)dW_t^1$$

$$dV_t = p(S,V,t)dt + q(S,V,t)dW_t^2$$

with $$< dW_t^1, dW_t^2 >= \rho.dt$$ [18]

S_t depends here on *t* and on V_t and any derivative written on S_t will depend on *t*, S_t and V_t.

So suppose now that a any vanilla derivative C written on S_t is a function of *t* and S_t like in the one dimensional case, and also V_t (which means implicitly that S_t depends itself on V_t). Let's denote it $C(t,S_t,V_t)$.We have then for any such function $C : (t,x,y) \rightarrow C(t,x,y)$ $C^{1,2,2}$ in time and in 2-dimensionnal space the 2-dimensionnal Itô lemma

$$\begin{aligned} dC(t,S_t,V_t) = & \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial x}dS_t + \frac{\partial C}{\partial y}dV_t \\ & + \frac{1}{2}\frac{\partial^2 C}{\partial x^2} < dS_t, dS_t > \\ & + \frac{\partial^2 C}{\partial x \partial y} < dS_t, dV_t > \\ & + \frac{1}{2}\frac{\partial^2 C}{\partial y^2} < dV_t, dV_t > \end{aligned}$$ [19]

---

[18] ρ is called the correlation coefficient between the two Brownian motions, because here the predictable variation is equal to the square variation, which definition is the covariation of the two processes. See III-B.

[19] The term in the third line of the equality is equal by symmetry to $2.\frac{1}{2}\frac{\partial^2 C}{\partial x \partial y} < dS_t, dV_t >$

B- Basis of discontinuous stochastic calculus

Let's begin this part with a very technical result, derived from the decomposition of Doob-Meyer. We know that in the continuous case, the decomposition of a semimartingale into a martingale and a process of finite variation is unique, but this result doesn't hold anymore for discontinuous semimartingales .

We only have the following result which allows us to introduce the concept of continuous part of a semimartingale.

If X_t is a semimartingale, then there exists processes such that

$$X_t = X_0 + X_t^C + M_t^d + A_t$$

where X_t^C is a continuous local martingale, M_t^d a purely discontinuous local martingale and A a process of finite variation. This decomposition is unique with respect to X_t^C, called the continuous part of X_t.

A local martingale is said purely discontinuous if and only if for every continuous local martingale M^C , M^C is orthogonal to M^d (i.e $M^C M^d$ is a local martingale).

We won't use in deep this technical theorem. We use this property only in the frame of jump diffusions. We know perfectly the form of the continuous part of (X_t) when X_t is a jump diffusion, and the continuous part of jumps diffusions is intuitive and comprehensive. In effect , if X_t is of the form

$$X_t = \gamma.t + \sigma.B_t + \sum_{k=1}^{N_t} \Delta X_k$$

then $$X_t^C = \gamma.t + \sigma.B_t$$

The quadratic variation is assimilated to predictable variation in the continuous case. We see often the definition

$$< X_t, Y_t > \; = \lim_{n \to \infty} \sum_{k=0}^{n} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k})$$ [20]

with the limit taken in the sense of L^2 convergence, but in fact it's rather the definition of the quadratic variation.

For instance, we often define in the continuous case the classic predictable variation of the Brownian motion as the square variation of the process $\sum_{k=0}^{n} (B_{t_{k+1}} - B_{t_k})^2$ which converges on L^2 in *t*.

Remark use the notation $< X >_t$ or $< X, X >_t$ to focus on the fact that it's a process, but it is of course can be considered as a bilinear form $(X_t, Y_t) \to < X_t, Y_t >$ where the scalar product between two finite variation processes and between a finite variation process and a Brownian motion is zero .

When we deal with (discontinuous) semimartingales, we must take into account this square variation ("crochet droit") as a different concept from the predictable variation ("crochet oblique").

$$[X, Y]_t = \lim_{n \to \infty} \sum_{k=0}^{n} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}) = \int_0^t [dX_s, dY_s]$$

We can remark that this definition is valid almost surely for a probability p , so it's also valid for another probability q which is equivalent to p. Hence, the square variation is independent from the probability q chosen , with the only condition that this probability must be equivalent to the initial probability. See the demonstration of the theorem of Girsanov in IV-B for a direct application.

[20] Remember that we take $t_n = t$ and that $n \to +\infty$ is equivalent to $\sigma \to 0$ where $\sigma = \sup | t_{i+1} - t_i |$

Let (X_t) and (Y_t) be two semimartingales processes. We denote from now on X_t^C and Y_t^C the continuous part of X_t and Y_t. By remarking that

$$(X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}) = X_{t_{k+1}} Y_{t_{k+1}} - X_{t_k} Y_{t_k} - Y_{t_k}(X_{t_{k+1}} - X_{t_k}) - X_{t_k}(Y_{t_{k+1}} - Y_{t_k})$$

We obtain by summing and taking the limit,

$$[X_t, Y_t] = X_t Y_t - X_0 Y_0 - \int_0^t X_{s^-} dY_s - \int_0^t Y_{s^-} dX_s \qquad (1)$$

In particular,

$$[X, X]_t = X_t^2 - X_0^2 - 2\int_0^t X_{s^-} dX_s$$

We have also demonstrated thanks to this equality the differential calculus rule

$$d(X_t Y_t) = X_s^- dY_s + Y_s^- dX_s + d[X, Y]_t$$

We have the relation between the quadratic variation and the predictable variation by splitting the jumps and the continuous part

$$[X_t, Y_t] = [X_t^C, Y_t^C] + \sum_{s \le t} \Delta X_s \Delta Y_s$$

As quadratic variation is equal to predictable variation for continuous semimartingales, we obtain :

$$[X_t, Y_t] = < X_t^C, Y_t^C > + \sum_{s \le t} \Delta X_s \Delta Y_s$$

We obtain then this very intuitive result: the total quadratic variation is equal to the predictable variation of the continuous part plus the jumps which are unpredictable.

Let's recall another important property of predictable variation to end this section

$$if \;\; X_t = \int_0^t \Phi_s^- dM_s \;\; then < X, X >_t = \int_0^t \Phi_s^2 d < M >_s$$ [21]

[21] M is here a continuous semimartingale. We use the notation $d < M >_s = d < M, M >_s$.

This property holds then naturally for càglàd processes in the form,

$$If\ \ X_t = \int_0^t \Phi_s dM_s \ \ then < X, X >_t \ = \int_0^t \Phi_s^2 d < M >_s$$

C- Itô Formula for semimartingales and applications for jump diffusions

Let's begin this part with an intuitive result. We consider paths (or functions) $t \to X_t(w)$ or equivalently a stochastic process (X_t) without Brownian component, where paths are a.s piecewise and continuous on the classical partition of [*0,t*], with $t_0 = 0$ and $t_n = t$.

Beside, let's define $X(t_{i+1}) = X^+_{t_{i+1}}$ [22]. (X_t) is then clearly a (discontinuous) càdlàg process to which we can apply classic (non stochastic) differential rules in its continuous part.

We assume the process is the sum of its continuous part and its discontinuous part to focus on what happens at points of discontinuity. Such process would be represented as

$$X_t = X_t^C + \sum_{i, t_i \le t} \Delta X_{t_i} \quad \text{where } \Delta X_{t_i} = X_{t_i} - X_{t_i}^-$$

We have for any smooth C^1 function *f*

$$f(X_t) - f(X_0) = \sum_{k=0}^{n-1} (f(X_{t_{i+1} \wedge t}) - f(X_{t_i \wedge t}))$$

$$f(X_t) - f(X_0) = \sum_{k=0}^{n-1} (f(X_{t_{i+1}}) - f(X_{t_i})) \text{ [23]}$$

$$f(X_t) - f(X_0) = \sum_{k=0}^{n-1} (f(X_{t_{i+1}}) - f(X_{t_{i+1}}^-) + f(X_{t_{i+1}}^-) - f(X_{t_i}))$$

$$f(X_t) - f(X_0) = \sum_{k=0}^{n-1} (f(X_{t_{i+1}}) - f(X_{t_{i+1}}^-)) + \sum_{k=0}^{n-1} (f(X_{t_{i+1}}^-) - f(X_{t_i}))$$

Replacing X_{t_i} by $X_{t_i}^- + \Delta X_{t_i}$, we obtain

$$f(X_t) - f(X_0) = \sum_{k=1}^{n} (f(X_{t_i}^- + \Delta X_{t_i}) - f(X_{t_i}^-)) + \sum_{k=0}^{n-1} (f(X_{t_{i+1}}^-) - f(X_{t_i}))$$

[22] We use this notation instead of $X_{t_{i+1}^+}$

[23] $t_0 = 0$ and $t_n = t$

The process (X_t) is continuous on $[t_i, t_{i+1}[$ or equivalently on $[t_i, t_{i+1}^-]$, [24] then $f(X_s)$ is continuous on this interval, and we can write then

$$f(X_{t_{i+1}}^-) - f(X_{t_i}) = \int_{t_i}^{t_{i+1}^-} f'(X_s) dX_s$$

thanks to the analogy with the classic differential property [25]

$$f(g(t)) - f(g(s)) = \int_s^t f^{'}(g(u)) g^{'}(u) du = \int_s^t f^{'}(g(u)) \frac{dg(u)}{du} du$$

$$f(g(t)) - f(g(s)) = \int_s^t f^{'}(g(u)) dg(u)$$

As this property is true when (X_t) is continuous on left points,

$$\sum_{k=0}^{n-1} (f(X_{t_{i+1}}^-) - f(X_{t_i})) = \sum_{k=0}^{n-1} \int_{t_i}^{t_{i+1}^-} f'(X_s) dX_s^C$$

Let's make now the changing of variable $X_s \to X_s^-$ so $dX_s^C \to dX_s^C$ by translation, $t_i \to t_i^+$ and $t_{i+1}^- \to t_{i+1}$, then

$$\sum_{k=0}^{n-1} (f(X_{t_{i+1}}^-) - f(X_{t_i})) = \sum_{k=0}^{n-1} \int_{t_i}^{t_{i+1}^+} f'(X_s^-) dX_s^C$$

With using the right continuity of the process,

$$\sum_{k=0}^{n-1} (f(X_{t_{i+1}}^-) - f(X_{t_i})) = \int_0^t f'(X_s^-) dX_s^C$$

we obtain

$$f(X_t) - f(X_0) = \sum_{k=0}^{n-1} (f(X_{t_{i+1}}) - f(X_{t_{i+1}}^-)) + \int_0^t f'(X_s^-) dX_s^C$$

---

[24] (X_t^-) is continuous on $[t_i^+, t_{i+1}]$ so is continuous on $[t_i, t_{i+1}]$ thanks to càdlàg property then $dX_s = dX_s^C$ on $[t_i, t_{i+1}]$

[25] Remember that X_t is non derivable so the derivative of X_t doesn't exist

or i.e $$f(X_t) - f(X_0) = \sum_{k=1}^{n} (f(X_{t_i}) - f(X_{t_i}^-)) + \int_0^t f'(X_s^-) dX_s^C$$

Replacing X_{t_i} by $X_{t_i}^- + \Delta X_{t_i}$, we obtain

$$f(X_t) - f(X_0) = \sum_{k=1}^{n} (f(X_{t_i}^- + \Delta X_{t_i}) - f(X_{t_i}^-)) + \int_0^t f'(X_s^-) dX_s^C$$

Let's remark that

$$\sum_{k=1}^{n} (f(X_{t_i}^- + \Delta X_{t_i}) - f(X_{t_i}^-)) = \sum_{s \le t}^{\Delta X_s \neq 0} (f(X_s^- + \Delta X_s) - f(X_s^-))$$

Then we obtain the formula

$$f(X_t) - f(X_0) = \sum_{s \le t}^{\Delta X_s \neq 0} (f(X_s^- + \Delta X_s) - f(X_s^-)) + \int_0^t f'(X_s^-) dX_s^C$$

which is in differential notations

$$df(X_t) = (f(X_s^- + \Delta X_s) - f(X_s^-)) + f'(X_t^-) dX_t^C$$

Starting from

$$f(t, X_t) - f(0, X_0) = \sum_{k=0}^{n-1} (f(t_{i+1}, X_{t_{i+1}}) - f(t_i, X_{t_i})) = \int_0^t df(s, X_s)$$

we can obtain with the same reasoning the formula in space and in time

$$df(t, X_t) = f_t^{'}(t, X_t) dt + f_x^{'}(t, X_t^-) dX_t^C + f(t, X_t^- + \Delta X_t) - f(t, X_t^-)$$

We can go further by remarking that

$$dX_s = dX_s^C + \Delta X_s$$

In effect, at a moment of jump, $dX_s = \Delta X_s$. At a moment where there is no jump, $\Delta X_s = 0$ so $dX_s = dX_s^C$. Then with replacing[26] dX_s^C with $dX_s - \Delta X_s$ when $\Delta X_s \neq 0$, we obtain the Itô formula in space for càdlàg processes without Brownian component in their continuous part :

$$f(X_t)\text{-}f(X_0)= \sum_{s \le t}^{\Delta X_s \neq 0} \left(f(X_s^- + \Delta X_s) - f(X_s^-) - \Delta X_s f'(X_s^-) \right) + \int_0^t f'(X_s^-) dX_s^C$$

In differential notations, we have a very intuitive result

$$df(X_t) = f'(X_t^-)dX_t + f(X_t^- + \Delta X_t) - f(X_t^-) - \Delta X_t f'(X_t^-)$$

It means that in point of continuity, i.e where $\Delta X_t = 0$, the increasing of $df(X_t)$ is $f'(X_t^-)dX_t$. In discontinuous points, where $\Delta X_t \neq 0$, dX_s^C=0. Then $dX_s - \Delta X_s$=0 and $f'(X_t^-)dX_t - \Delta X_t f'(X_t^-)$ = 0. Thus we have in points of discontinuity t $df(X_t)$ = $f(X_t^- + \Delta X_t) - f(X_t^-)$.

With the same reasoning, we can prove the Itô formula in space and time for any Lévy process without Brownian component , with any function $C^{1,1}$ in time and space

$$df(t, X_t) = f_t^{'}(t, X_t^-)dt + f_x^{'}(t, X_t^-)dX_t + f(t, X_t^- + \Delta X_t) - f(t, X_t^-) - \Delta X_t f'(t, X_t^-)$$

with merely using the differential rule in two dimensions

$$df(t, X_t) = f_t^{'}(t, X_t)dt + f_x^{'}(t, X_t)dX_t^C$$

on intervals where X_t is continuous.

[26] This is valid only if for jumps diffusions where the continuous has a trivial form.

General Itô formula for semimartingales

We tackle now the case of general semimartingales. In intervals of continuity, the increasing of $f(t,X_t)$ is no more $f_t^{'}(t,X_t)dt + f_x^{'}(t,X_t)dX_t$. We have to use Itô calculus for continuous semimartingales to determine $df(t,X_t)$ on intervals where X_t is continuous.

To do so, we consider again a stochastic process (X_t) where paths are a.s continuous on $]t_i,t_{i+1}[$ on the classical partition of [*0,t*], with $t_0=0$ and $t_n=t$ with the additional property

$$X(t_{i+1}) = X_{t_{i+1}}^{+}$$

to make it a (discontinuous) càdlàg process.

We have for any smooth C^2 function *f*, like in the precedent section,

$$f(X_t)-f(X_0) = \sum_{k=1}^{n}(f(X_{t_i}^{-}+\Delta X_{t_i})-f(X_{t_i}^{-}))+\sum_{k=0}^{n-1}(f(X_{t_{i+1}}^{-})-f(X_{t_i}))$$

Let's remark that

$$\sum_{k=0}^{n-1}(f(X_{t_{i+1}})-f(X_{t_{i+1}}^{-}))=\int_0^t df(X_t)=\int_0^t df(X_t^C)$$ [27]

having X_t^{-} left continuous by definition then continuous thanks to càdlàg property.

Then

$$f(X_t)-f(X_0)=\int_0^t df(X_t^C)+\sum_{s\le t}^{\Delta X_s\neq 0}(f(X_s^{-}+\Delta X_s)-f(X_s^{-}))$$

And in differential notations

$$df(X_t) = df(X_t^C) + f(X_t^{-}+\Delta X_t)-f(X_t^{-})$$

---

[27] Remark that X_t^{-} is left continuous by definition so it's a continuous process thanks to càdlàg property. Then we have $X_t^{-}=X_{t^{-}}^{C}$.

We have in the other hand on intervals where X_t is continuous

$$f(X_{t_{i+1}}^-) - f(X_{t_i}) = \int_{t_i}^{t_{i+1}^-} f_x^{'}(X_s)dX_s + \frac{1}{2}\int_{t_i}^{t_{i+1}^-} f_{xx}^{''}(X_s)d[X_s, X_s]$$

$$f(X_{t_{i+1}}^-) - f(X_{t_i}) = \int_{t_i}^{t_{i+1}^-} f_x^{'}(X_s)dX_s^C + \frac{1}{2}\int_{t_i}^{t_{i+1}^-} f_{xx}^{''}(X_s)d[X_s^C, X_s^C]$$

After changing variable $X_s \to X_s^-$ and summing, it gives

$$df(X_t^C) = f_x^{'}(X_t^-)dX_t^C + \frac{1}{2} f_{xx}^{''}(t, X_t^-)d[X_t^C, X_t^C]$$

So we obtain

$$df(X_t) = f_x^{'}(X_t^-)dX_t^C + \frac{1}{2} f_{xx}^{''}(t, X_t^-)d[X_t^C, X_t^C] + (f(X_t^- + \Delta X_t) - f(X_t^-))$$

With replacing dX_t^C by $dX_t - \Delta X_t$, we obtain the famous Itô formula for semimartingales in space,

$$df(X_t) = f_x^{'}(X_t^-)dX_t + \frac{1}{2} f_{xx}^{''}(t, X_t^-)d[X_t^C, X_t^C] + (f(X_t^- + \Delta X_t) - f(X_t^-) - \Delta X_t f'(X_t^-))$$

Starting from

$$f(t, X_t) - f(0, X_0) = \sum_{k=0}^{n-1} (f(t_{i+1}, X_{t_{i+1}}) - f(t_i, X_{t_i})) = \int_0^t df(s, X_s)$$

we can obtain with the same reasoning the formula in space and in time

$$df(t, X_t) = df(t, X_{t^-}^C) + f(t, X_t^- + \Delta X_t) - f(t, X_t^-)$$

As Itô calculus in 2 dimensions on continuous parts of X_t is typically

$$df(t, X_t) = f_t^{'}(t, X_t)dt + f_x^{'}(t, X_t)dX_t + \frac{1}{2} f_{xx}^{''}(t, X_t)d[X, X]_t$$

we obtain the famous Itô formula for semimartingales in time and space

$$\begin{aligned} df(t, X_t) \quad = \quad & f_t^{'}(t, X_t)dt + f_x^{'}(t, X_t^-)dX_t + \frac{1}{2} f_{xx}^{''}(t, X_t^-)d[X, X]_t^C \\ & + f(t, X_t^- + \Delta X_t) - f(t, X_t^-) - \Delta X_t f'(t, X_t^-) \end{aligned}$$

where $d[X,X]_t^C = d[X_t^C, X_t^C] = d < X^C >_t$ is the derivative of the quadratic variation of the continuous part of X_t.

In explicit notation, the Itô formula for semimartingales is

$$f(t, X_t) - f(0, X_0) = \int_0^t \frac{\partial f}{\partial s}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s^-)dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s^-)d[X,X]_s^C$$
$$+ \sum_{s \le t}^{\Delta X_s \ne 0} \left(f(s, X_s) - f(s, X_s^-) - \Delta X_s f'(s, X_s^-) \right)$$

Please note that this intuitive demonstration of Itô formula in the general case of semimartingales is rather a way to remind Itô formula quickly than a demonstration because of the divergence of the sum

$$\sum_{s \le t}^{\Delta X_s \ne 0} (f(s, X_s) - f(s, X_s^-))$$

in the case of an infinity of jumps. So the expression $f(t, X_t^- + \Delta X_t) - f(t, X_t^-)$ used in our demonstration is not defined in the case of infinite activity Lévy processes.

For a less intuitive but correct demonstration, we must rather show the convergence

$$f(t, X_t) - f(0, X_0) \to \int_0^t \frac{\partial f}{\partial s}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s^-)dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_s^-)d[X,X]_s$$
$$+ \sum_{s \le t}^{\Delta X_s \ne 0} \left(f(s, X_s) - f(s, X_s^-) - \Delta X_s \frac{\partial f}{\partial x}(s, X_s^-) - \frac{\partial^2 f}{\partial x^2}(s, X_s^-)\Delta X_s^2 \right)$$

thanks to a second order Taylor expansion and conclude with a particular property of the quadratic variation we showed in this chapter

$$[X_t, X_t] = [X_t^C, X_t^C] + \sum_{s \le t} \Delta X_s^2$$

Application to jump diffusion processes

We assume

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \sum_{i=1}^{N_t} \Delta X_i$$

with b_t and σ_t non anticipating [28] processes

With denoting $Z_t = \sum_{i=1}^{N_t} \Delta X_i$, we have with differential notations,

$$dX_t = b_t dt + \sigma_t dW_t + dZ_t$$

For any function $f\ C^{1,2}$ the process $f(t, X_t)$ can be represented as follow

$$f(t, X_t) - f(0, X_0) = \int_0^t \frac{\partial f}{\partial s}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s^-) b_s ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s^-) \sigma_s dW_s$$

$$+ \frac{1}{2} \int_0^t \sigma_s^2 \frac{\partial^2 f}{\partial x^2}(s, X_s) ds + \sum_{s \le t}^{\Delta X_s \ne 0} (f(X_s^- + \Delta X_s) - f(X_s^-))$$

In differential notations,

$$df(t, X_t) = f_t^{'}(t, X_t) dt + b_t f_x^{'}(t, X_t^-) dt + \sigma_t f_x^{'}(t, X_t^-) dW_t + \frac{1}{2} \sigma_t^2 f_{xx}^{''}(t, X_t) dt$$

$$+ f(t, X_t^- + \Delta X_t) - f(t, X_t^-)$$

In effect, from the general formula

$$df(t, X_t) = f_t^{'}(t, X_t) dt + f_x^{'}(t, X_t) dX_t + \frac{1}{2} f_{xx}^{''}(t, X_t) d[X, X]_t^C$$

$$+ f(t, X_t^- + \Delta X_t) - f(t, X_t^-) - \Delta X_t f'(t, X_t^-)$$

with replacing $dX_s - \Delta X_s$ with dX_s^C

[28] It is merely another formulation for adapted processes

$$df(t, X_t) \quad = \quad f_t^{'}(t, X_t)dt + f_x^{'}(t, X_t^-)dX_t^C + \frac{1}{2} f_{xx}^{''}(t, X_t^-)d[X, X]_t^C$$

$$+ \ f(t, X_t^- + \Delta X_t) - f(t, X_t^-)$$

We use the fact that

$$dX_t^C = b_t dt + \sigma_t dW_t$$

And the equality

$$d[X_t^C, X_t^C] = [dX_t^C, dX_t^C] =< b_t dt + \sigma_t dW_t, b_t dt + \sigma_t dW_t >$$

$$d[X_t^C, X_t^C] == \sigma_t^2 dt$$

The conclusion is then trivial.

IV- Incomplete Markets' Modeling

A- Principles of pricing

The markets are supposed incomplete, which means that there are sources of risks that can not be perfectly hedged. Assets' prices are supposed ruled by a dynamic based on a constant drift, jumps and stochastic volatility. Incomplete markets are due to the fact that jumps and stochastic volatility are not traded in markets so we cannot set a portfolio perfectly hedged with derivatives written on these stocks.

There is an equivalence between the existence of the risk neutral probability and the absence of arbitrage opportunities. We will suppose then that the market is free from arbitrage opportunity (no free lunch).

Since there is another equivalence between completeness of the market and uniqueness of risk neutral probability, there is an infinity of risk neutral probabilities (called also martingales measures) under which discounted prices of assets are martingales in the frame of incomplete markets. So we must provide any incomplete model with the chosen martingale measure we'll work with.

Once the martingale measure Q is determined, we use it to calculate the premium on derivatives written on the stock.

Typically, the premium of an option written on S_t with payoff $h(S_T)$ is :

$$C_t = C(t, S_t) = e^{-r(T-t)} E_Q \left(h(S_T) \mid F_t \right)$$

If the distribution of S_t is known in a closed form, then we can compute directly the expectation like in the case assets' prices follow Black Scholes diffusion, or like in the Merton model.

Since Black & Scholes, we know that the premium of a derivative is the price of the hedging portfolio of the latest. Then by deriving the diffusion of the semimartingale S_t with Itô formula for general semimartingales, and with using some classical arguments, like that the hedging portfolio must return a risk free rate, and using the fact that there is no arbitrage opportunity with the martingale property of discounted price of assets, we can deduce the risk neutral PDE (Partial Differential Equation) followed by the derivative under the risk neutral probability.

But unfortunately, this PDE is rarely directly solvable (so no closed formula is available) and we almost all the time need numerical methods to approximate the solution. These are provided by numerical simulations, but the time of calculus is often long and do not permit to price correctly all the derivatives with a reasonable precision.

Another choice is to solve the PDE followed by the characteristic function of the logarithm price of the underlying instead of solving from the PDE followed by the premium of the derivative. We can then obtain sometimes a closed form of the characteristic function of the asset price, and then a semi-closed form formula of the premium of the derivative by numerical approximation of Inverse Fourier transform.

Today, a model is considered efficient if it leads to closed form formula of the derivative, or at the limit to semi-closed form formulas. This is our particular topic with the jump diffusions.

In effect, with most of the jumps and stochastic volatility models, a closed formula of the Fourier transform of $\log S_t$ is available. We will see in the next chapter an application in the frame of pricing vanilla European options. Before tackling the models, additional sections on probability changing, risk free portfolios and martingales identification are fundamental to recall.

B- Mathematical tools for probabilities' changing

Equivalence of probabilities

Two probabilities Q and P are equivalent if they preserve the equalities in the almost sure sense i.e if they define the same set of negligible events,

$$P \sim Q \quad \Leftrightarrow \quad \forall\ A \in F,\ \ P(A)=0 \Leftrightarrow Q(A)=0$$

Changing of probability

This concept is very intuitive. Consider the calculation of the expectation of X under a new equivalent probability Q, that we want to express in function of the probability P.

$$E_Q(X) = \int_\Omega X(w)dQ(w) = \int_\Omega X(w)(\frac{dQ(w)}{dP(w)})dP(w)$$

If we denote the $L(w) = \frac{dQ(w)}{dP(w)} = \frac{dQ}{dP}(w)$ then

$$E_Q(X) = \int_\Omega X(w)L(w)dP(w) = E_P(XL)$$

L is then the so-called changing of probability $\frac{dQ}{dP}$ (or probability density of Q with respect to P).

Conversely we can choose any random variable to be a changing of probability with setting the new probability Q hold the condition $dQ = LdP$.

The only condition to respect is that

$$E_P(L) = \int_\Omega L(w)dP(w) = \int_\Omega dQ(w) = Q(\Omega) = 1$$

So we can define any changing of probability from a probability P to a new probability Q with merely choosing a random variable L with expectation 1 under Q.

Bayes Formula

For a given σ-algebra F, with defining $L = \dfrac{dQ}{dP}$ *we have* $E_Q(X \mid F) = \dfrac{E_P(LX \mid F)}{E_P(L \mid F)}$

In effect, $\forall A \in F$,

$$E_Q(X \mid F)(A) = E_Q(X \mid A) = E_{Q_{|A}}(X) = \int_\Omega X(w)dQ(w \mid A)$$

$$E_Q(X \mid F)(A) = \int_\Omega X(w)\frac{dQ(w \cap A)}{Q(A)} = \frac{\int_\Omega 1_{w \in A} X(w)dQ(w)}{Q(A)} = \frac{\int_A X(w)dQ(w)}{\int_A dQ(w)}$$

$$E_Q(X \mid F)(A) = \frac{\int_A X(w)L(w)dP(w)}{\int_A L(w)dP(w)} = \frac{\int_\Omega X(w)L(w)dP(w \cap A)}{\int_\Omega L(w)dP(w \cap A)}$$

$$E_Q(X \mid F)(A) = \frac{\int_\Omega X(w)L(w)\frac{dP(w \cap A)}{P(A)}}{\int_\Omega L(w)\frac{dP(w \cap A)}{P(A)}} = \frac{\int_\Omega X(w)L(w)dP(w \mid A)}{\int_\Omega L(w)dP(w \mid A)}$$

$$E_Q(X \mid F)(A) = \frac{E_P(XL \mid A)}{E_P(L \mid A)} = \frac{E_P(XL \mid F)(A)}{E_P(L \mid F)(A)}$$

The most direct application is the following. Suppose you define with the probabilities $Q(. \mid F_T)$ and $P(. \mid F_T)$ the changing of probability

$$L_T = \frac{dQ_{|F_T}}{dP_{|F_T}} \qquad \text{i.e} \qquad L_T(w) = \frac{dQ(w \mid F_T)}{dP(w \mid F_T)}$$

then $$E_Q(X_T \mid F_t) = \frac{E_P(L_T X_T \mid F_t)}{E_P(L_T \mid F_t)}$$

Let's recall the Girsanov theorem to end this section. We won't really use directly this theorem it in this paper but its importance in the reasoning makes it indispensable to recall.

Girsanov theorem (*)

If W_t^P is a Brownian motion under P, then with the changing of probability

$$L_T = \frac{dQ_{|F_T}}{dP_{|F_T}} \quad i.e \quad L_T(w) = \frac{dQ(w \mid F_T)}{dP(w \mid F_T)}$$

with defining [29] $$L_T = e^{\int_0^T \theta_s dW_s^P - \frac{1}{2}\int_0^T \theta_s^2 ds} < +\infty$$

and $$L_t = E_P(L_T \mid F_t) \quad for \quad t \le T$$

then

$$W_t^Q = W_t^P - \int_0^t \theta_s ds$$

is a Brownian motion under Q

We begin the demonstration by remarking that L_t is a martingale by definition. Then L_t is well defined changing of probability thanks to [30]

$$E(L_t) = E(L_0) = E(e^0) = 1$$

Thanks to the Bayes property

$$E_Q(W_t^Q \mid F_t) = \frac{E_P(L_T W_t^Q \mid F_t)}{E_P(L_T \mid F_t)}$$

$$E_Q(W_t^Q \mid F_t) = \frac{E_P(L_T W_t^Q \mid F_t)}{L_t}$$

[29] This expression must be finite. On can apply for instance Novikov condition $E_P(e^{\frac{1}{2}\int_0^T \theta_s^2 ds}) < +\infty$ which is sufficient but not necessary.

[30] If X_t is a martingale then $E(X_0) = E(E(X_t \mid F_0)) = E(E(X_t) \mid F_0) = E(X_t)$ for any $t \le T$

We need to prove now that $W_T^Q L_T$ is a P-martingale. We have first under the probability P

$$d(W_t^Q L_t) = L_t dW_t^Q + W_t^Q dL_t + [dW_t^Q, dL_t]$$

$$d(W_t^Q L_t) = L_t dW_t^P - \theta_t L_t dt + W_t^Q dL_t + [dW_t^P + \theta_t dt, dL_t]$$

$$d(W_t^Q L_t) = L_t dW_t^P - \theta_t L_t dt + W_t^Q dL_t + [dW_t^P, dL_t]$$

Besides, by Itô lemma for continuous semimartingales [31] applied to the Itô process $X_t = \ln(L_t)$ with $f : x \to e^x$,

$$df(X_t) = e^{X_t} dX_t + \frac{1}{2} e^{X_t} \theta_t^2 dt$$

then $$d(L_t) = L_t(\theta_t dW_t^P - \frac{1}{2}\theta_t^2 dt) + \frac{1}{2} L_t \theta_t^2 dt$$

and so, under P $$dL_t = L_t \theta_t dW_t^P$$

Then

$$d(W_t^Q L_t) = L_t dW_t^P - \theta_t L_t dt + W_t^Q L_t \theta_t dW_t^P + [dW_t^P, L_t \theta_t dW_t^P]$$

$$d(W_t^Q L_t) = L_t dW_t^P - \theta_t L_t dt + W_t^Q L_t \theta_t dW_t^P + \theta_t L_t dt$$

$$d(W_t^Q L_t) = L_t (1 + W_t^Q \theta_t) dW_t^P$$

So $W_t^Q L_t$ is a (Brownian) martingale under P and then

$$E_Q(W_t^Q \mid F_t) = \frac{L_t W_t^Q}{L_t} = W_t^Q$$

So W_t^Q is a *Q*-martingale. Besides, almost surely under P

$$[dW_t^Q, dW_t^Q] = [dW_t^P - \theta_t dt, dW_t^P - \theta_t dt] = [dW_t^P, dW_t^P] = dt$$

thanks to the fact that W_t^P is a P- Brownian motion.

[31] The informed reader would rather use directly Doléans-Dades exponential martingales property to obtain the result

As the square variation $[dW_t^Q, dW_t^Q]$ does not depend on the probability used [32], because it's the limit of the quadratic covariation in L^2 , we have the equality almost sure under Q

$$[dW_t^Q, dW_t^Q] = dt$$

So W_t^Q is a Q-martingale and its square variation is equal to *t*. Then W_t^Q is a Brownian motion under Q.

[32] Under the condition that the probabilities are of course equivalent (we must keep the equality almost sure)

C- Probabilities' changing and risk neutral diffusions

The martingale identification and changing of probabilities

We'll often use the following property. A continuous semimartingale asset price X_t - or by abuse of language dX_t - is a martingale [33] if its dt part is null.

In effect, by the theorem of representation of Brownian martingales,

$$\exists H \text{, such as } X_t = \int HdW \Leftrightarrow X_t \text{ is a martingale}$$

So if we want X_t to be a martingale, then it must be necessary of the form

$$dX_t = H_t dW_t$$

with a null dt part. Remark that this is natural since Itô processes are called continuous semimartingales because it's the sum of a martingale and a (deterministic) dt part.

If X_t is a jump diffusion, then we are in discontinuous frame and we don't have the equivalence anymore. X_t can be for instance the sum of a Brownian motion and a compensated Poisson martingale. So what we do is to find one martingale representation of X_t, and then remark that we have chosen the martingale measure by choosing this representation. The reader can find a direct application in the Merton model in chapter V.

We must not forget that a diffusion followed by X_t depends entirely on a chosen probability, since the diffusion is in the almost sure sense, and since the Brownian motion involved in this diffusion is not necessarily a Brownian motion under another probability.

So what find an equivalent probability, i.e under which equality a.s are preserved, in which the use of arbitrage arguments permits to prove that it's a martingale measure.

[33] In fact if the process holds this condition, it is then only a local martingale The regularity condition of square integrability is additionally required for the process to be a martingale. We assume this condition is fulfilled.

For instance, if under the historical probability p

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^p$$

where W_t^p is a Brownian motion under p, then by arbitrage argument – explained below - the drift under the risk neutral probability must be r, so

$$\frac{dS_t}{S_t} = rdt + \sigma(dW_t^p + \frac{(\mu - r)}{\sigma} dt)$$

All what we have to do is to find the probability Q under which

$$W_t^Q = W_t^p + \frac{(\mu - r)}{\sigma} t$$

is a Brownian motion. This is provided by the famous Girsanov theorem.

Then we obtain the so called risk-neutral diffusion

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t^Q \qquad \text{almost surely under } Q$$

where Q is well defined by the Girsanov theorem of changing of probability.

So the reasoning we'll apply systematically is the following.

We start from the diffusion under the historical probability p

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^p$$

and we assume there exists a risk neutral probability such as under Q

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t^Q$$

where W_t^Q is a Brownian motion under Q, then we search the value of μ under the arbitrage condition which is $\widetilde{S}_t = e^{-rt} S_t$ is a Q-martingale, as follow

$$d\widetilde{S}_t = -r.e^{-rt}.S_t.dt + e^{-rt} dS_t$$

$$d\widetilde{S}_t = e^{-rt}.S_t(-r.dt + \frac{dS_t}{S_t})$$

$$d\widetilde{S}_t = e^{-rt}.S_t((\mu - r).dt + \sigma.dW_t^Q)$$

$\widetilde{S}_t$ is then clearly a martingale if and only if the drift $\mu = r$.

Note that we obtain this equivalence by the theorem of representation of Brownian martingales, and then we have only an implication in the case of a jump diffusion within the frame of incomplete markets.

Then we can write that we have found the risk neutral probability Q (by making the Girsanov changing probability) under which we have the risk-adjusted or risk-neutral diffusion

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t^Q$$

In incomplete market models, we go with the same reasoning but there is no more equivalence. We find only one possible value of the drift under which the discounted price of the asset is a martingale. So we find one possible probability Q under which $\widetilde{S}_t$ is a martingale. The choice of this drift is then equivalent to the choice of the probability Q. We'll see a direct application of this reasoning when computing the model of Merton to vanilla options prices.

Let's recall the mathematical expression of a risk-free portfolio to end this section. Any risk free portfolio Π_t must hold the condition $d\Pi_t = r\Pi_t dt$. In effect, this condition is equivalent to

$$\frac{d\Pi_t}{\Pi_t} = r.dt$$

then with integrating this equality

$$\ln(\Pi_t) - \ln(\Pi_0) = rt$$

which is equivalent to

$$\Pi_t = \Pi_0 e^{rt}$$

D- The model of Merton

The model of Merton is a particular case of the model of Bates, as volatility is set local. Let S_t the asset prices observed in the market, and $X_t = \ln(S_t)$.

Merton's model assumes that $X_t = \gamma.t + \sigma.B_t + J_t$ with jumps' sizes Y_i i.i.d[34] with a Gaussian distribution, $Y_i \sim N(\mu, \sigma^2)$.

A closed form of the distribution of the process price

Let $A \in B(\mathbb{R}^+)$. We now try to give a closed expression of the distribution of X_t.

$$p(X_t \in A) = \sum_{k=0}^{\infty} p(X_t \in A \mid N_t = k).p(N_t = k)$$

The distribution of X_t knowing $N_t = k$ is a Gaussian distribution. In effect, we can see clearly that

$$\gamma.t + \sigma.B_t + \sum_{i=1}^{k} Y_i \text{ follows the Gaussian distribution } N(\gamma.t + k.\mu \ , \ \sigma^2.t + k.\delta^2)$$

as $B_t \sim N(0,t)$ so $\sigma B_t \sim N(0, \sigma^2 t)$ [35] and $\sum_{i=1}^{k} Y_i \sim N(k\mu, k\sigma^2)$

Then,

$$p(X_t = x \mid N_t = k) = p(\gamma.t + \sigma.B_t + \sum_{i=1}^{k} Y_i = x) = \frac{e^{(-\frac{(x-\gamma.t-k.\mu)^2}{2(\sigma^2.t+k.\delta^2)})}}{\sqrt{2\Pi(\sigma^2.t + k.\delta^2)}}$$

Let's denote $p_t(x)$ the distribution of X_t, i.e $p_t(x) = p(X_t = x)$

[34] identically and independently distributed

[35] Remember that $\mathrm{var}(aX) = a^2 \, \mathrm{var}(X)$

Using $p(N_t = k) = \dfrac{e^{-\lambda.t}(\lambda.t)^k}{k!}$, we obtain finally the law of X_t

$$p_t(x) = e^{-\lambda.t}\sum_{k=0}^{\infty}\frac{(\lambda.t)^k.e^{(-\frac{(x-\gamma.t-k.\mu)^2}{2(\sigma^2.t+k.\delta^2)})}}{k!.\sqrt{2\Pi(\sigma^2.t+k.\delta^2)}}$$

Risk neutral PDE of the model of Merton

We consider the same jump diffusion under the historical probability p,

$$S_t = S_0 e^{\mu t+\sigma W_t+\sum_{i=1}^{N_t} Y_i}$$

where N_t is a Poisson process with intensity λ, independent from W_t

and $Y_i \sim N(m,\delta^2)$ under the chosen risk neutral probability Q [36]

Let $X_t = \ln(S_t)$, $Z_t = \sum_{i=0}^{N_t} Y_i$

$$X_t = X_0 + \mu t + \sigma W_t + \sum_{i=0}^{N_t} Y_i$$

So applying Itô formula for semimartingales with $f : x \to e^x$

$$df(X_t) = f^{'}(X_t^-)dX_t^C + \frac{1}{2}f^{''}(X_t^-)d[X_t^C, X_t^C] + f(X_t^- + \Delta X_t) - f(X_t^-)$$

So

$$dS_t = S_t^-(\mu dt + \sigma W_t) + \frac{1}{2}\sigma^2 S_t^- dt + S_t^- e^{\Delta X_t} - S_t^-$$

$$\frac{dS_t}{S_t^-} = (\mu + \frac{1}{2}\sigma^2)dt + \sigma dW_t + (e^{\Delta X_t} - 1)$$

[36] which suppose that jumps are already risk adjusted, i.e jumps are not diversifiable. See below

$\Delta X_t = Y_t$ if t is a moment of jump and $\Delta X_t = 0$ elsewhere. So the process $e^{\Delta X_t} - 1$ worth $e^{Y_t} - 1$ if t is a moment of jump, and worth 0 elsewhere.

Consequently, $e^{\Delta X_t} - 1 = d\hat{Z}_t$ where $\hat{Z}_t$ is a compound Poisson process whose jumps are $e^{Y_t} - 1$, i.e

$$\hat{Z}_t = \sum_{i=0}^{N_t}(e^{Y_i} - 1) = \sum_{s \le t}(e^{\Delta X_s} - 1)$$

Let $\hat{Y}_i = e^{Y_i} - 1$ then $\ln(1 + Y_i) \sim N(m, \delta^2)$ so $\hat{Z}_t$ is a compound Poisson process with jumps size log-normally distributed. More precisely, like in the case of the Bates model, it is $1 + Y_i$ which is log-normal.

So under the historical probability P,

$$\frac{dS_t}{S_t^-} = (\mu + \frac{1}{2}\sigma^2)dt + \sigma dW_t + d\hat{Z}_t$$

Consider now a neutral probability Q under which $\widetilde{S}_t$ is a martingale. We shall have the diffusion

$$\frac{dS_t}{S_t^-} = (\mu_Q + \frac{1}{2}\sigma^2)dt + \sigma dW_t^Q + d\hat{Z}_t$$

or equivalently

$$X_t = X_0 + \mu_Q t + \sigma W_t^Q + \sum_{i=0}^{N_t} Y_i$$

wher W_t^Q is a Brownian motion under Q.

Let's find now the risk adjusted parameter μ under which $\widetilde{S}_t$ is a Q-martingale

$$d(\widetilde{S}_t) = d(e^{-rt}S_t) = -re^{-rt}S_t dt + e^{-rt}dS_t = -r\widetilde{S}_t dt + e^{-rt}dS_t = \widetilde{S}_t(-rdt + \frac{dS_t}{S_t})$$

$$d(\widetilde{S}_t) = \widetilde{S}_t((-r + \mu_Q + \frac{1}{2}\sigma^2)dt + \sigma dW_t^Q + d\hat{Z}_t)$$

We know $\hat{Z}_t - E_Q(\hat{Z}_t) = \hat{Z}_t - \lambda E_Q(\hat{Y}_i)t$ s is a martingale under Q so let it appear

$$d(\widetilde{S}_t) = \widetilde{S}_t((-r + \mu_Q + \frac{1}{2}\sigma^2 + \lambda E(\hat{Y}_i))dt + \sigma dW_t^Q + d\hat{Z}_t - \lambda E(\hat{Y}_i)dt)$$

$d\hat{Z}_t - \lambda E(\hat{Y}_i)dt$ is a compensated martingale under Q, and σW_t^Q is also a Q-martingale. So $d(\widetilde{S}_t)$ is a Q-martingale if the "*dt*" part is null, i.e

$$\mu_Q = r - \frac{1}{2}\sigma^2 - \lambda E(e^{Y_i} - 1) = r - \frac{1}{2}\sigma^2 - \lambda(e^{m+\frac{1}{2}\delta^2} - 1)$$

We have then the neutral diffusion under the probability Q,

$$S_t = S_0 e^{\mu_Q t + \sigma W_t^Q + \sum_{i=1}^{N_t} Y_i}$$

With a mere recurrence, we have the following equality

$$S_T = S_t e^{\mu_Q (T-t) + \sigma(W_T^Q - W_t^Q) + \sum_{i=N_t}^{N_T} Y_i}$$

and the following equality in distribution

$$S_T \overset{distribution}{=} S_t e^{\mu_Q (T-t) + \sigma W_{T-t}^Q + \sum_{i=1}^{N_{T-t}} Y_i}$$ [37]

which means that for every payoff h,

$$E_Q(h(S_T)) = E_Q(h(S_t e^{\mu_Q (T-t) + \sigma W_{T-t}^Q + \sum_{i=1}^{N_{T-t}} Y_i}))$$

As there is an infinity of choices of the probability Q, Merton choose this particular one because he modified only the drift of the Brownian motion (by modifying the *dt* part) and kept the jumps unchanged. He argued that the jump risk is diversifiable so no risk premium is attached to it. Mathematically, it means that changing the probability (and so changing the risk repartition) must not affect the jumps, and so that the jumps are already risk adjusted.

[37] If X=Y in distribution, then for any borelian function f, E(f(X))=E(f(Y)). Remark for the sequel this result : if X and Y are independent from another random variable A, then E(f(AX))=E(f(AY))

E- The model of Heston

We begin this part with general results on stochastic volatility models of the form

$$\frac{dS_t}{S_t} = \mu(t, S_t, V_t)dt + \sigma(t, S_t, V_t)dW_t^1$$

$$dV_t = p(t, S_t, V_t)dt + q(t, S_t, V_t)dW_t^2$$

where W^1 are W^2 are two Brownian motions with correlation ρ, i.e $< dW_t^1, dW_t^2 >= \rho.dt$. We can simplify the notation to make calculus clearer

$$\frac{dS}{S} = \mu.dt + \sigma dW^1$$

$$dV = p.dt + q.dW^2$$

The value of any derivative C written on S_t is a function of tree variables $C(t, S_t, V_t)$. The stochastic volatility V_t is a non tradable asset in markets. As the randomness feature of V depends entirely on W_2, the stochastic volatility renders the market incomplete because of the role of this second Brownian motion. In B&S model, one can hedge perfectly a derivative written on S_t because selling S_t is equivalent to selling the Brownian W_1, and then we can hedge from the only source of risk of S_t in C which is W_1.

With the presence of two sources of risk W_1 and W_2, we must hedge with :
- the underlying asset like in the Black and Scholes model to hedge dW_1
- another derivative to hedge the second source of risk dW_2.

Let's set a portfolio Π_t constituted of a derivative C_t, an amount Δ of underlying asset, and an amount Δ_1 of the second derivative used to hedge the risk of volatility.

$$\Pi(S_t, V_t, t) = C(S_t, V_t, t) - \Delta.S_t - \Delta_1 C_1(S_t, V_t, t)$$

i.e with simplified notations we have

$$\Pi = C - \Delta S - \Delta_1 C_1$$

The property of self financement allows to write

$$d\Pi = dC - \Delta dS - \Delta_1 dC_1$$

With Itô lemma for vectorial continuous semimartingales on C and C_1, we obtain easily using bilinear properties of predictable variation

$$dC = \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S}dS + \frac{\partial C}{\partial V}dV + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}d < S,S > + \frac{\partial^2 C}{\partial S \partial V}d < S,V > + \frac{1}{2}\frac{\partial^2 C}{\partial V^2}d < V,V >$$

$$dC = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial V^2}\right)dt + \frac{\partial C}{\partial S}dS + \frac{\partial C}{\partial V}dV$$

Then

$$\begin{aligned} d\Pi &= \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial V^2}\right)dt \\ &- \Delta_1\left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial V} + \frac{1}{2}q^2 \frac{\partial^2 C_1}{\partial V^2}\right)dt \\ &+ \left(\frac{\partial C}{\partial S} - \Delta_1 \frac{\partial C_1}{\partial S} - \Delta\right)dS + \left(\frac{\partial C}{\partial V} - \Delta_1 \frac{\partial C_1}{\partial V}\right)dV \end{aligned}$$

which is equivalent to

$$\begin{aligned} d\Pi &= \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2}q^2 \frac{\partial^2 C}{\partial V^2}\right)dt \\ &- \Delta_1\left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial V} + \frac{1}{2}q^2 \frac{\partial^2 C_1}{\partial V^2}\right)dt \\ &+ \left(\frac{\partial C}{\partial S} - \Delta_1 \frac{\partial C_1}{\partial S} - \Delta\right)S\mu.dt + \left(\frac{\partial C}{\partial S} - \Delta_1 \frac{\partial C_1}{\partial S} - \Delta\right)S.\sigma.dW_t^1 \\ &+ \left(\frac{\partial C}{\partial V} - \Delta_1 \frac{\partial C_1}{\partial V}\right)p.dt + \left(\frac{\partial C}{\partial V} - \Delta_1 \frac{\partial C_1}{\partial V}\right)q dW_t^2 \end{aligned}$$

The portfolio is risk free so we must have

$$d\Pi = r\Pi dt = r(C - \Delta S - \Delta_1 C_1)dt,$$

It implies in particular by identification that all stochastic terms (with W_t^1 and W_t^2) are equal to zero, then

$$\frac{\partial C}{\partial S} - \Delta_1 \frac{\partial C_1}{\partial S} - \Delta = 0,$$

and
$$\frac{\partial C}{\partial V} - \Delta_1 \frac{\partial C_1}{\partial V} = 0,$$

We obtain then

$$d\Pi = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C}{\partial V^2}\right)dt$$
$$-\Delta_1 \left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C_1}{\partial V^2}\right)dt$$

And with arranging terms after identification of parts with dt with $r(C - \Delta S - \Delta_1 C_1)dt$, we obtain

$$\frac{\left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C}{\partial V^2} + rS\frac{\partial C}{\partial S} - rC\right)}{\frac{\partial C}{\partial V}}$$
$$= \frac{\left(\frac{\partial C_1}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C_1}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C_1}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C_1}{\partial V^2} + rS\frac{\partial C_1}{\partial S} - rC_1\right)}{\frac{\partial C_1}{\partial V}}$$

The expression at the left is a function of C_1 only while the expression at the right is a function of C only.

We have with letting the function h

$$h(C(S,V,t) = \frac{\left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C}{\partial V^2} + rS\frac{\partial C}{\partial S} - rC\right)}{\frac{\partial C}{\partial V}}$$

we obtain the equality

$$h(C(S,V,t),S,V,t) = h(C_1(S,V,t),S,V,t)$$

Since this equality is true for any derivative C_1, no matter what is the characteristics of the derivative, it does not depend on C_1. In our purpose, C_1 and C are call premium and though the equality is true no matter what are the maturities and the strikes of the two options. Then the only way this equality can be always true is that both the expression at the left and the right do not depend on the first variable C of the function $(S,V,t) \to h(C(S,V,t),S,V,t)$.

In effect, consider that a function of two variables holds the condition

$h(x, y) = h(a, y)$ for every $a \in \Re$ and $y \in \Re$

Then if we freeze y we see that the function

$x \to h(x, y)$

is a constant equal to some real b_y. In effect, $\forall x \in \Re$ $h(x, y) = h(0, y)$ for instance. So if we set $b_y = h(0, y)$, then $\forall x \in \Re$, $h(x, y) = b_y$

So for every y, we see that the function $x \to h(x, y)$ is equal to the application $x \to b_y$. Then it comes that $h(x, y) = b_y = b(y)$, i.e h is a function of y only.

Then, for any arbitrary function g, we have consequently,

$$h(C(S,V,t),S,V,t) = g(S,V,t)$$

and so

$$\frac{\partial C}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}+\rho\sigma q S\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}q^2\frac{\partial^2 C}{\partial V^2}+rS\frac{\partial C}{\partial S}-rC=g.\frac{\partial C}{\partial V}$$

-*g* is called the risk neutral drift of the volatility, because it explains the variation of the volatility as we'll see in the sequel (we can already see that is the drift of the véga of the derivative)

Thus, without loss of generality, with letting $g = q\lambda - p$, for any function $\lambda(S,V,t)$ [38]

$$\frac{\partial C}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}+\rho\sigma Sq\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}q^2\frac{\partial^2 C}{\partial V^2}+rS\frac{\partial C}{\partial S}-rC=-(p-\lambda q)\frac{\partial C}{\partial V}$$

and equivalently, we obtain the partial differential equation to solve :

$$\frac{\partial C}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}+\rho\sigma Sq\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}q^2\frac{\partial^2 C}{\partial V^2}+rS\frac{\partial C}{\partial S}+(p-\lambda q)\frac{\partial C}{\partial V}-rC=0 \qquad \text{(IV-1)}$$

Let's explain now the choice of $g = \lambda q - p$, called the neutral drift of volatility.

Suppose $\Pi = C - \Delta S$, i.e that we want to hedge our portfolio in delta neutral. We want to see only how much the volatility *V* explains the risk in the variation of *C* with suppressing the risk coming from the variation of the underlying.

By self financing

$$d\Pi = dC - \Delta dS$$

So

$$d\Pi=\left(\frac{\partial C}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}+\rho\sigma q S\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}q^2\frac{\partial^2 C}{\partial V^2}\right)dt+\left(\frac{\partial C}{\partial S}-\Delta\right)dS+\frac{\partial C}{\partial V}dV$$

Thanks to the delta neutral hedging, the part with dS is zero, and we obtain :

[38] There is an equivalence because $h(S,V,t)\to qh(S,V,t)-p$ a bijection for $q\neq 0$

$$d\Pi = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C}{\partial V^2} \right) dt + \frac{\partial C}{\partial V} dV$$

$$d\Pi - r\Pi dt = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \rho\sigma q S \frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2} q^2 \frac{\partial^2 C}{\partial V^2} + rS\frac{\partial C}{\partial S} - rC \right) dt + \frac{\partial C}{\partial V} dV$$

$$= q\frac{\partial C}{\partial V}(\lambda dt + dW^2)$$

Then

$$d\Pi \quad = \quad r\Pi dt + q\frac{\partial C}{\partial V}(\lambda dt + dW^2) \quad \text{, i.e}$$

$$d\Pi \quad = \quad r\Pi dt + (q\frac{\partial C}{\partial V})\lambda dt + (q\frac{\partial C}{\partial V})dW_t^2$$

We observe then that for each risk unit generated in the period dt by the second Brownian motion W_2 there is λ additional units of risks added to the risk free rate in the same period. This is the reason why λ is called the risk neutral premium of volatility.

Besides, multiplying equation (IV-1) by *dt,* and recognizing *dt* terms of *dC* leads to

$$dC - \frac{\partial C}{\partial S}dS - \frac{\partial C}{\partial V}dV + rS\frac{\partial C}{\partial S}dt + (p - \lambda q)\frac{\partial C}{\partial V}dt - rCdt = 0$$

and by the way

$$dC - \frac{\partial C}{\partial S}(dS - rSdt) - \frac{\partial C}{\partial V}(dV - (p - \lambda q)dt) = rCdt$$

The drift μ of S_t becomes *r* and the drift of V_t becomes $p - \lambda q$ in the risk neutral diffusion of V_t. In effect, with denoting $\widetilde{C}$ the discounted price of the premium *C*, we have under a risk neutral probability *Q*,

$$dC - rCdt = \frac{\partial C}{\partial S}(dS - rSdt) - \frac{\partial C}{\partial V}(dV - (p - \lambda q)dt)$$

$$d\widetilde{C} = e^{-rt}(dC - rCdt) = e^{-rt}\ (\frac{\partial C}{\partial S}(dS - rSdt) - \frac{\partial C}{\partial V}(dV - (p - \lambda q)dt))$$

$$d\widetilde{C} = e^{-rt}\ (\frac{\partial C}{\partial S}(dS - rSdt) - \frac{\partial C}{\partial V}(dV - (p - \lambda q)dt))$$

If we want $\widetilde{C}$ to be a martingale under this probability then all parts with *dt* must be null. So the part with *dt* of *V* must be $(p - \lambda q)dt$ and the part with *dt* of *S* must be $rSdt$.

This the reason why we say that the risk neutral drift of $\widetilde{C}$ is *r* (as the part with *dt* of *dS/S* is *rdt* and the risk neutral drift of the volatility is $(p - \lambda q)$ (as the part with *dt* of *dV* is $(p - \lambda q)dt$.

As the probability under which we work is not entirely determined until λ is fixed, the market is rendered incomplete because of the infinity of possible risk premiums λ. Each choice of λ is equivalent to the choice of a martingale measure *Q* under which discounted prices are martingales.

Application to the model of Heston

We now set

- $\mu(t,S_t,V_t) = \mu(t)$ (μ depends only on t)
- $\sigma(t,S_t,V_t) = \sqrt{V_t}$
- $p(t,S_t,V_t) = \alpha(\overline{V} - V_t)$
- $q(t,S_t,V_t) = \eta \sqrt{V_t}$

which left us with the diffusion

$$\frac{dS_t}{S_t} = \mu(t)dt + \sqrt{V_t}\,dW_t^1$$

$$dV_t = \alpha(\overline{V} - V_t)dt + \sqrt{V_t}\,dW_t^2$$

with $$< dW_t^1, dW_t^2 >= \rho.dt$$

The first equation is similar to the one in extended Black Sholes models, but the volatility of the process S_t is not local anymore. The variance follows itself diffusion, which is a square root process due to Cox, Ingersoll and Ross.

The choice of the diffusion of V_t is at the source of plenty of stochastic volatility models, but the CIR process is empirically and theoretically the most relevant one for two major reasons :

- The CIR process gives a trend to the volatility with the property of return to the mean
- It displays always positive values of volatilities (as opposed to certain models that provide negative volatilities) [39]

[39] It displays also a hitting time of 0 which is infinite with appropriate parameters

For instance, if V_t - thanks an increasing trajectory of the Brownian motion- exceeds by too much the long volatility (long term mean of volatility) $\overline{V}$, then $\alpha(\overline{V}-V_t)dt$ will be an important negative term and the following variation dV_t would be roughly negative, which means that V_t will return progressively to the value $\overline{V}$. By symmetry, the reasoning is also true at the other side. The average speed of return to the mean is controlled by α. A high α will lead to a fast return to the mean. $\overline{V}$ is then naturally the parameter of long volatility

The correlation of the Brownian models the leverage effect, i.e the negative correlation between the price of stocks *S* and the volatility *V* (the parameter ρ is always set negative between -1 and 0). In this case, when the first Brownian is positive and contributes to give a positive trend to the variation *dS*, the second Brownian is negative and contributes to give a negative trend to the volatility.

The economic explanation is due to Merton : when shares fall down, the company becomes more indebted and then more risky, then the volatility of the shares increases. At the opposite, when shares rise, the company becomes less risky because it capital increases in proportion with respect to the debt, so the volatility decreases. Added to the fear of downward jumps of shares' prices since the crash of 1987, these are the most popular arguments to explain the form of the implicit volatility surface since then and the skew observed within stocks markets, and why in particular puts deep out of the money are overpriced and calls deep in the money are underpriced.

Intuitively, we can think about the risk premium of volatility λ as something proportional to the volatility $\sqrt{V_t}$ of the process S_t. Heston set then $\lambda(t,S_t,V_t) = \theta.\sqrt{V_t}$ with an arbitrary θ according to the works of Wiggins (1987) which demonstrated economically the existence of such a relation.

All this leads us to the equation of pricing of a vanilla derivative *C* written on *S* in the frame of the model of Heston

$$\frac{\partial C}{\partial t}+\frac{1}{2}V.S^2\frac{\partial^2 C}{\partial S^2}+\rho.\eta.V.S\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}\eta^2 V\frac{\partial^2 C}{\partial V^2}+rS\frac{\partial C}{\partial S}+(\alpha(\overline{V}-V_t)-\theta.\eta.V_t)\frac{\partial C}{\partial V}-rC=0$$

$$\frac{\partial C}{\partial t}+\frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}+\rho\sigma Sq\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}q^2\frac{\partial^2 C}{\partial V^2}+rS\frac{\partial C}{\partial S}+(p-\lambda q)\frac{\partial C}{\partial V}-rC=0$$

With setting $\alpha_r=\alpha+\theta.\eta$ and $\alpha_r.\overline{V}_r=\alpha.\overline{V}$, (so called risk adjusted parameters)

$$\alpha(\overline{V}-V_t)-\theta.V_t = \alpha_r(\overline{V}_r-V_t)$$

And we obtain the Heston's equation with risk adjusted mean reverting and long volatility α_r and $\overline{V}_r$, i.e we obtain the risk neutral partial differential equation of the price of a vanilla derivative,

$$\frac{\partial C}{\partial t}+\frac{1}{2}V.S^2\frac{\partial^2 C}{\partial S^2}+\rho V.S\frac{\partial^2 C}{\partial S\partial V}+\frac{1}{2}V\frac{\partial^2 C}{\partial V^2}+rS\frac{\partial C}{\partial S}+(\alpha_r(\overline{V}_r-V_t))\frac{\partial C}{\partial V}-rC=0$$

F- The model of Bates

The model of Bates is a generalization of the model of Heston and the model of Merton. It assumes that the asset's price is driven by both a stochastic volatility and jumps in the form

$$\frac{dS_t}{S_t} = \mu.dt + \sqrt{V_t}\,dW_t^1 + dZ_t$$

$$dV_t = \alpha(\overline{V} - V_t)dt + \eta\sqrt{V_t}\,dW_t^2$$

We use the same leverage effect than the one in Heston model by setting $< dW_t^1, dW_t^2 >= \rho.dt$ with ρ negative. The drift μ is set constant.

Z_t is a compound Poisson process with intensity λ [40] of jumps sizes Y_i, i.e

$$Z_t = \sum_{i=1}^{N_t} Y_i$$

where Y_i are i.i.d log-normally distributed. More precisely, with parameters k (mean of jumps' size) and σ^2 (variance of size of jumps) , we have

$$\ln(1+Y_i) \sim N(\ln(1+k) - \frac{1}{2}\sigma^2, \sigma^2)$$

This complex expression of the mean is set to lead to the equality $E(Y_i) = k$. In effect, thanks to Laplace transform [41]

$$E(e^{\ln(1+Y_i)}) = e^{\ln(1+k) - \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2} = 1\text{+k}$$

So when we use the model, we want to manipulate the parameter k which is the average size of jumps rather than any another complex parameter with obscure economic interpretation.

[40] This intensity of jumps λ has of course nothing to do with the risk premium of volatility λ of the model of Heston.

[41] If $X \sim N(m, \sigma^2)$ then $E(e^X) = e^{m + \frac{1}{2}\sigma^2}$ (Laplace transform)

Furthermore, this parameter must be risk adjusted in order not to be incompatible with the other risk adjusted parameters when passing to risk neutral diffusions. So we consider

$$k = E_Q(Y_i)$$

where Q is the chosen martingale measure.

Another presentation of the model

Log-normal distribution is used into this pseudo geometric Brownian motion of S_t to work with normal distributions when passing to log prices.

We can notice that $$Z_t = \sum_{i=1}^{N_t} Y_i = Z_t - Z_0 = \int_0^t dZ_s$$

$$Z_t = \sum_{i=1}^{n} Y_{T_i} = \sum_{s \leq t} Y_{s^-}(N_s - N_{s^-})$$

In effect, during a time of jump T_i such as $N_t = i$, $N_s - N_{s^-} = 1$. Elsewhere $N_s - N_{s^-} = 0$

Thus $$\int_0^t dZ_s = \int_0^t Y_{s^-} dN_s$$

and in differential notations $$dZ_s = Y_{s^-} dN_s$$

So $dZ_t = Y_{T_i}.1 = Y_{T_i}$ if $t = T_i$,

and $dZ_t = 0$ elsewhere.

Besides, we know that $E(dN_s) = \lambda dt$ [42]. The only two possible values for dN_s are trivially 0 and 1 so we can see that $E(dN_s) = 1.p(dN_s = 1) + 0.p(dN_s = 0) = \lambda dt$.

[42] by the fact that N_t is a Lévy process and so with stationary increments that $E(N_t - N_s) = \lambda(t - s)$

So we can conclude that $p(dN_s = 1) = \lambda dt$ and consequently $p(dN_s = 0) = 1 - \lambda dt$

Consequently, $p(dZ_s = Y_{s^-}) = \lambda dt$ and $p(dZ_s = 0) = 1 - \lambda dt$

That is the reason why we often see in the literature the presentation of the diffusion of the price within the model of Bates in this form

$$\frac{dS_t}{S_t} = \mu.dt + \sqrt{V_t}\,dW_t^1 + Y_t dN_t$$

where Y_i are i.i.d log-normal and N_t a Poisson process with intensity λ such as $p(dN_s = 1) = \lambda dt$ and $p(dN_s = 0) = 1 - \lambda dt$.

The risk neutral PDE of the model of Bates

The market is free from arbitrage opportunities if and only if discounted prices of assets are martingales. Let's find now the risk adjusted parameters under which $\widetilde{S}_t$ is a martingale. Under a risk neutral probability *Q*,

$$d(\widetilde{S}_t) = d(e^{-rt}S_t) = -re^{-rt}S_t dt + e^{-rt}dS_t = -r\widetilde{S}_t dt + e^{-rt}dS_t = \widetilde{S}_t(-rdt + \frac{dS_t}{S_t})$$

So $$d\widetilde{S}_t = \widetilde{S}_t(-rdt + \frac{dS_t}{S_t})$$

$$d\widetilde{S}_t = \widetilde{S}_t(-rdt + \mu dt + \sqrt{V_t}\,dW_t^1 + dZ_t)$$

We know that $dZ_t - E_Q(dZ_t) = dZ_t - \lambda E_Q(Y_i)dt$ is a martingale under *Q*, so we'll make it appear

$$d\widetilde{S}_t = \widetilde{S}_t(-rdt + \mu dt + \lambda k dt + \sqrt{V_t}\,dW_t^1 + dZ_t - \lambda k dt)$$

The part with *dt* must be zero, so $\mu - r + \lambda k = 0$. With arbitrage arguments, we proved that the risk adjusted drift is $(r - \lambda k)dt$, and we deal with risk adjusted diffusion

$$\frac{dS_t}{S_t} = (r - \lambda k).dt + \sqrt{V_t}\, dW_t^1 + dZ_t \qquad \text{(IV-3)}$$

$$dV_t = \alpha(\overline{V} - V_t)dt + \eta\sqrt{V_t}\, dW_t^2$$

Applying Itô formula for jump diffusions from (IV-3) to $X_t = \ln(S_t)$, we obtain

$$d(\ln(S_t)) = \frac{1}{S_t} dS_t^C - \frac{1}{2}\frac{1}{S_t^2} d[S_t^C, S_t^C] + \ln(S_{t^-} + \Delta S_t) - \ln(S_{t^-})$$

From (IV-3), we can see clearly that the jumps $\frac{\Delta S_t}{S_{t^-}}$ are equal to dZ_t, because the contribution of the continuous part to the jumps is null.

Besides,

$$dS_t^C = S_t(r - \lambda k)dt + S_t\sqrt{V_t}\, dW_t^1 \text{ so } d[S_t^C, S_t^C] = S_t^2 V_t dt.$$

Consequently,

$$dX_t = (r - \lambda k)dt + \sqrt{V_t}\, dW_t^1 - \frac{1}{2}V_t dt + \ln(1 + \frac{\Delta S_t}{S_{t^-}})$$

and so we obtain the equation for the log-price X_t

$$dX_t = (r - \lambda k - \frac{1}{2}V_t)dt + \sqrt{V_t}\, dW_t^1 + \ln(1 + dZ_t)$$

With denoting $d\hat{Z}_t = \ln(1 + dZ_t)$, we obtain the final form of this equation

$$dX_t = (r - \lambda k - \frac{1}{2}V_t)dt + \sqrt{V_t}\, dW_t^1 + d\hat{Z}_t \qquad \text{(IV-4)}$$

where $\hat{Z}_t$ is a compound Poisson process with jumps $\hat{Y}_i \sim N(\ln(1+k) - \frac{1}{2}\sigma^2, \sigma^2)$.

In effect, considering $dZ_t = Y_{T_i}$ if $t = T_i$ and 0 if not (i.e $dZ_t = Y_t$ if t is a time of jump and 0 elsewhere), then $d\hat{Z}_t = \ln(1 + dZ_t) = \ln(1 + Y_t) = \hat{Y}_t$ if t is a moment of jump and 0 elsewhere, where $\hat{Y}_t \sim N(\ln(1+k) - \frac{1}{2}\sigma^2, \sigma^2)$. $\hat{Z}_t$ is consequently a compound Poisson process with jumps $\hat{Y}_i \sim N(\ln(1+k) - \frac{1}{2}\sigma^2, \sigma^2)$.

We can use (IV-4) to compute easily derivative prices using Monte Carlo methods as we will see in the chapter VI.

V- Pricing European options in the frame of incomplete markets

A- The model of Merton

We start from the risk neutral PDE obtained in the previous section

$$S_T \overset{distribution}{=} S_t e^{\mu_Q (T-t)+\sigma W^Q_{T-t} + \sum_{i=1}^{N_{T-t}} Y_i}$$

Under *Q*, the European vanilla call price with payoff *h* (for instance $h: x \to (x-K)^+$ for a plain vanilla call or $h: x \to (K-x)^+$ for a plain vanilla put) is

$$C_t = e^{-r(T-t)} E_Q((S_T - K)^+ \mid F_t)$$

So S_t being a Lévy process, it is hence a Markov process

$$C_t = e^{-r(T-t)} E_Q(h(S_T) \mid S_t = S)$$

$$C_t = e^{-r(T-t)} E_Q(h(S_t e^{\mu_Q (T-t)+\sigma W^Q_{T-t} + \sum_{i=1}^{N_{T-t}} Y_i}) \mid S_t = S)$$

$$C_t = e^{-r(T-t)} E_Q(h(S e^{\mu_Q (T-t)+\sigma W^Q_{T-t} + \sum_{i=1}^{N_{T-t}} Y_i}))$$

$$C_t = e^{-r(T-t)} \sum_{n \geq 0} p(N_{T-t} = n) E_Q(h(S_t e^{\mu_Q (T-t)+\sigma W^Q_{T-t} + \sum_{i=1}^{N_{T-t}} Y_i}) \mid N_{T-t} = n)$$

So $$C_t = e^{-r(T-t)} \sum_{n \geq 0} p(N_{T-t} = n) E_Q(h(S e^{\mu_Q (T-t)+\sigma W^Q_{T-t} + \sum_{i=1}^{n} Y_i})$$

$$C_t = e^{-r(T-t)} \sum_{n \geq 0} p(N_{T-t} = n) E_Q(h(S e^{\mu_Q (T-t)} e^{\sigma W^Q_{T-t} + \sum_{i=1}^{n} Y_i})$$

We remark that $$\sum_{i=1}^{n} Y_i \sim N(nm, n\delta^2)$$

so $$\sum_{i=1}^{n} Y_i \overset{distribution}{=} nm + \frac{\delta\sqrt{n}}{\sqrt{T-t}}.W_{T-t}^{Q}$$

then $$\sum_{i=1}^{n} Y_i + \sigma W_{T-t}^{Q} \overset{distribution}{=} nm + \sqrt{\frac{\delta^2 n}{T-t} + \sigma^2}\ W_{T-t}^{Q}$$ [43]

And then if we set $$\sigma_n^2 = \sigma^2 + \frac{n\delta^2}{T-t},$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(Se^{\mu_Q(T-t)} e^{\sigma W_{T-t}^{Q} + \sum_{i=1}^{n} Y_i})$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(Se^{\mu_Q(T-t)} e^{nm + \sigma_n W_{T-t}^{Q}})$$

The idea here is to find an equivalent of the Black Scholes formula with a constant drift and a square variance depending on *n,* with choosing carefully an appropriate value of S_t.[44]

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(Se^{(r-\frac{1}{2}\sigma^2 - \lambda(e^{m+\frac{1}{2}\delta^2} - 1))(T-t) + nm + \sigma_n W_{T-t}^{Q}})$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(Se^{nm - \lambda(e^{m+\frac{1}{2}\delta^2} - 1)(T-t) + (r-\frac{1}{2}\sigma^2)(T-t) + \sigma_n W_{T-t}^{Q}})$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(Se^{nm - \lambda(e^{m+\frac{1}{2}\delta^2} - 1)(T-t) + \frac{n\delta^2}{2} + (r-\frac{1}{2}\sigma_n{}^2)(T-t) + \sigma_n W_{T-t}^{Q}})$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(Se^{nm - \lambda(e^{m+\frac{1}{2}\delta^2} - 1)(T-t) + \frac{n\delta^2}{2}} e^{(r-\frac{1}{2}\sigma_n{}^2)(T-t) + \sigma_n W_{T-t}^{Q}})$$

This leads us to

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) E_Q(h(S_n e^{(r-\frac{1}{2}\sigma_n{}^2)(T-t) + \sigma_n W_{T-t}^{Q}})$$

---

[43] if $X \sim N(m_1, \sigma_1^2)$ and $Y \sim N(m_2, \sigma_2^2)$ then $X + Y \sim N(m_1, \sigma_1^2 + \sigma_2^2)$

[44] In effect, we can show easily that Black Scholes formula is still available when the variance is depending on *t* or on another parameter if the drift is left constant. This is called extended Black and Scholes model.

$$C_t = \sum_{n\geq 0} p(N_{T-t} = n)\left(e^{-r(T-t)} E_Q(h(S_n e^{(r-\frac{1}{2}\sigma_n{}^2)(T-t)+\sigma_n W^Q_{T-t}})\right)$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} p(N_{T-t} = n) C_t^{BS}(S_n, \sigma_n)$$

$$C_t = e^{-r(T-t)} \sum_{n\geq 0} e^{-\lambda(T-t)} \frac{(\lambda(T-t))^n}{n!} C_t^{BS}(S_n, \sigma_n)$$

We obtain then the prices of options under Merton model with convergent series of well known Black Scholes prices

$$C_t = e^{-(r+\lambda)(T-t)} \sum_{n\geq 0} \frac{(\lambda(T-t))^n}{n!} C_t^{BS}(S_n, \sigma_n)$$

where $S_n = Se^{nm-\lambda(e^{m+\frac{1}{2}\delta^2}-1)(T-t)+\frac{n\delta^2}{2}}$ and $C_t^{BS}(S,\sigma)$ is the price of a plain vanilla call with maturity *T* and payoff *h* under Black Scholes model, given that S_t is equal to S at time *t*.

In effect, remember that with the Black Scholes Model, (the reader can see annexes for more details if necessary), we have

$$S_T = S_0.e^{(r-\frac{1}{2}.\sigma^2)T+\sigma.W_T}$$

$$S_T = S_t.e^{(r-\frac{1}{2}.\sigma_{BS}{}^2)(T-t)+\sigma_{BS}.W_{T-t}}$$

$$S_T = S.e^{(r-\frac{1}{2}.\sigma_{BS}{}^2)(T-t)+\sigma_{BS}.(W_T-W_t)}$$

And knowing $S_t = S_n$

$$S_T \overset{distribution}{=} S_n.e^{(r-\frac{1}{2}.\sigma^2)(T-t)+\sigma.W_{T-t}}$$

The price of a call with payoff *h* under Black Scholes modeling is then

$$C_t^{BS}(S_n,\sigma_n) = e^{-r(T-t)} E_Q(h(S_T) \mid F_t) = e^{-r(T-t)} E_Q(h(S_T) \mid S_t = S_n)$$

with $\quad S_T = S_0 . e^{(r-\frac{1}{2}.\sigma_n{}^2)T + \sigma_n . W_T}$

So

$$C_t^{BS}(S_n,\sigma_n) = e^{-r(T-t)} E_Q(h(S_n . e^{(r-\frac{1}{2}.\sigma_n{}^2)(T-t) + \sigma_n . W_{T-t}}))$$

B- Some remarks on the model of Heston

Let's start with an important theorem in differential calculus, systematically used when dealing with stochastic volatility models.

Theorem of Riccati *Given a differential equation such that*

$$f'(x) = a(x)f^2(x) + b(x)f(x) + c(x)$$

the following substitution

$$f(x) = -\frac{h'(x)}{a(x)h(x)}$$

transforms the first equation in a second order ODE (ordinary differential equation)

The demonstration is trivial. It is implicitly developed when using this property in the demonstration of the model of Bates.

Now we are interested in calculating the premium of a vanilla call. The premium is equal to the discounted expectation of the payoff of the call with respect to the risk probability *Q* which we have chosen during the construction of the model of Heston (when risk adjusting parameters). We remember that the market is incomplete and that the martingale measure *Q* under which discounted assets' prices are martingales is not unique.

$$C(S_t,t,T) = E^Q(e^{-r(T-t)}(S_T - K)^+ \mid F_t) = E^Q(e^{-r(T-t)}(S_T - K).1_{S_t \geq K} \mid F_t)$$

with the notation $(S_t - K)^+ = \max(S_T - K, 0)$

$$C(S_t,t,T) = E_Q(e^{-r(T-t)}S_T.1_{S_T \geq K} \mid F_t) - e^{-r(T-t)}K.E_Q(1_{S_T \geq K} \mid F_t)$$

$$C(S_t,t,T) = E_{Q_{|F_t}}(e^{-r(T-t)}S_T.1_{S_T \geq K}) - e^{-r(T-t)}K.E_{Q_{|F_t}}(1_{S_T \geq K})$$

To obtain plain vanilla options' prices we make the following famous changing of probability. For the first term we use the discounted stock price as numeraire and for the second we choose the risk free bonds.

$$dQ_{|F_t} = \frac{e^{-rT} S_T}{E_{Q_{|F_t}}(e^{-rT} S_T)} dP_1 = \frac{e^{-rT} S_T}{e^{-rt} S_t} dP_1 = e^{-r(T-t)} \frac{S_T}{S_t} dP_{1|F_t}$$

i.e $$\frac{dQ}{dP_1}|F_t = e^{-r(T-t)} \frac{S_T}{S_t}$$

and $$dQ_{|F_t} = \frac{e^{-\int_t^T r(s)ds}}{E_{Q_{|F_t}}(e^{-\int_t^T r(s)ds})} dP_2 = \frac{e^{-\int_t^T r(s)ds}}{B(t,T)} dP_2 = \frac{e^{-r(T-t)}}{e^{-r(T-t)}} dP_2 = dP_{2|F_t}$$

Since r is deterministic and non stochastic, the second changing of probability is in this particular case the identity. The reader would use the same technique with a stochastic interest like in Bakshi, Cao and Chen generalized SVJD models.

We obtain then

$$C(S_t,t,T) = E_{P_{1|F_t}}(S_t.1_{S_T \geq K}) - e^{-r(T-t)} K.E_{P_{2|F_t}}(1_{S_T \geq K})$$

Since S_t is $F_t - measurable$,

$$C(S_t,t,T) = S_t.E_{P_{1|F_t}}(1_{S_T \geq K}) - e^{-r(T-t)} K.E_{P_{2|F_t}}(1_{S_T \geq K})$$

$$C(S_t,t,T) = S_t.P_{1|F_t}(S_T \geq K) - e^{-r(T-t)} K.P_{2|F_t}(S_T \geq K)$$ [45]

We remark a very powerful and intuitive result which holds for all models

$$C(S_t,t,T) = (S_t - e^{-r(T-t)} K).P_{1|F_t}(S_T \geq K) + e^{-r(T-t)} K.P_{2|F_t}(S_T < K) \quad \text{(V-1)}$$

by merely setting $P_{2|F_t}(S_T \geq K) = 1 - P_{2|F_t}(S_T < K)$

The call price is equal to the probability of exercising the call multiplied by the payoff in case of exercising plus the present value of the strike multiplied by the probability of not exercising the option.

---

[45] $E_p(1_A) = 0.p(\overline{A}) + 1.p(A)$

The call premium depends now entirely on the probabilities P_1 and P_2 which are computed differently in each model.

One possible way to obtain semi-closed form formula for vanilla options is to put the expression (V-1) in the risk neutral PDE obtained in the section IV-E, and then to obtain a system of 2 PDEs where P_i, $i \in \{1,2\}$ are the unknowns. Then we take the Fourier transform $\widetilde{P}_i$, $i \in \{1,2\}$ of P_i, $i \in \{1,2\}$ and resolve the PDEs followed by $\widetilde{P}_i$, $i \in \{1,2\}$ thanks to the fact that we know the general form of solutions of such Fokker Planck equations. Starting from the substitution of the general solutions in the PDE, we use the Riccati theorem to obtain the expression of $\widetilde{P}_i$. Then we take the inverse Fourier transform to obtain the expression of P_i and the value of the call.

In our demonstration of the model of Bates, we put directly the characteristic function into the PDE followed by the price of derivative. These two approaches are equivalent and use strictly the same techniques of calculus. The Heston model is implicitly demonstrated in the next section by taking only the characteristic function of the continuous part of the log price.

C- The model of Bates

Remark that the form of the characteristic function of Heston model is included in the calculation of the characteristic function of X_t^C.

We develop the demonstration starting from the risk neutral diffusion obtained in IV-4

$$dX_t = (r - \lambda k - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t^1 + d\hat{Z}_t$$

$$dV_t = \alpha(\overline{V} - V_t)dt + \eta\sqrt{V_t}dW_t^2$$

With denoting X_t^C the continuous part of X_t,

$$dX_t = dX_t^C + d\hat{Z}_t$$

The jumps being independent of the Brownian motion, $\hat{Z}_t$ is independent from X_t^C. Then,

$$\phi_{X_t}(u) = E(e^{iu(X_t^C + \hat{Z}_t)}) = E(e^{iuX_t^C}e^{iu\hat{Z}_t}) = E(e^{iuX_t^C})E(e^{iu\hat{Z}_t}) = \phi_{X_t^C}(u)\ \phi_{\hat{Z}_t}(u)$$

We'll then calculate separately $\phi_{X_t^C}$ and $\phi_{\hat{Z}_t}$. For the case of Heston model without jumps, the demonstration is exactly the same with taking $\phi_{X_t} = \phi_{X_t^C}$. We follow by the way Heston demonstration for the calculus of the characteristic function of the continuous part of X_t.

We first freeze u, with denoting in the sequel $f_u = f$ and we let

$$f_u(x,v,t) = E_Q(e^{iuX_T^C} \mid X_t^C = x, V_t = v) \text{ [46]} \qquad \text{(V-2)}$$

and

$$M_t = f(X_t^C, V_t, t)$$

We see that

[46] $f(x,v,t) = E(e^{iuX_T^C} \mid (X_t^C = x) \cap (V_t = v))$

$$M_t = E_Q(e^{iuX_T^C} \mid X_t^C = X_t^C, V_t = V_t) = E_Q(e^{iuX_T^C} \mid \sigma(X_t^C, V_t))$$

In effect, let

$$F_{x,v} = \sigma((X_t = x) \cap (V_t = v)) = \{w \in F \mid X_t(w) = x \ \ and \ \ V_t(w) = v\}$$

(V-2) means that $\forall \ w \in F_{x,v}$,

$$f(X_t(w), V_t(w), t) = E_Q(e^{iuX_T^C} \mid \sigma(X_t, V_t) \ )(w)$$

which can be written

$$f(x, v, t) = E_Q(e^{iuX_T^C} \mid X_t^C = x, V_t = v)$$

But this equality holds in fact for every non negligible $w \in F$

$$f(X_t(.), V_t(.), t) = E_Q(e^{iuX_T^C} \mid \sigma(X_t, V_t) \ )(.,.)$$

i.e

$$f(X_t, V_t, t) = E_Q(e^{iuX_T^C} \mid \sigma(X_t, V_t) \ )$$

which is the targeted equality.

We remark also an important detail

$$f(x_0, v_0, 0) = E_Q(e^{iuX_T^C} \mid X_0^C = x_0, V_0 = v_0) = E_Q(e^{iuX_T^C} \mid F_0) = E_Q(e^{iuX_T^C}) = \phi_{X_T^C}(u)$$

Since every Lévy process is markovian,

$$M_t = E_Q(e^{iuX_T^C} \mid F_t)$$

But

$$M_T = E_Q(e^{iuX_T^C} \mid F_T) = e^{iuX_T^C} \quad [47]$$

[47] e^{iuX_T} is $F_T - measurable$, having X_T $F_T - measurable$ and $(x \to e^{iux})$ Borelian from $\Re$ to $\Re^2$ (by continuousity)

Then

$$M_t = E_Q(M_T \mid F_t) \quad \text{for every } t \le T \text{ and for every } T.$$

Applying Itô formula for continuous semimartingales to M_t leads to

$$df(X_t^C, V_t, t) = \frac{\partial f}{\partial x}(X_t^C, V_t, t) dX_t^C + \frac{\partial f}{\partial v}(X_t^C, V_t, t) dV_t + \frac{\partial f}{\partial t}(X_t^C, V_t, t) dt$$
$$+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(X_t^C, V_t, t) d < X_t^C > + \frac{\partial^2 f}{\partial v \partial x}(X_t^C, V_t, t) d < V_t, X_t^C > + \frac{1}{2}\frac{\partial^2 f}{\partial v^2}(X_t^C, V_t, t) d < V_t >$$

With simplifying the notations and making some trivial calculations,

$$dM_t = \left(\frac{\partial f}{\partial x}(r - \lambda k - \frac{1}{2}V_t) + \frac{\partial f}{\partial v}\alpha(\overline{V} - V_t) + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}V_t + \frac{1}{2}\frac{\partial^2 f}{\partial v^2}\eta^2 V_t + \frac{\partial^2 f}{\partial v \partial x}\eta\rho V_t\right)dt$$
$$+ \frac{\partial f}{\partial x}\sqrt{V_t}\, dW_t^1 + \frac{\partial f}{\partial v}\eta\sqrt{V_t}\, dW_t^2$$

M_t is a martingale so the part with *dt* must be null.

$$\frac{\partial f}{\partial x}(r - \lambda k - \frac{1}{2}V_t) + \frac{\partial f}{\partial v}\alpha(\overline{V} - V_t) + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}V_t + \frac{1}{2}\frac{\partial^2 f}{\partial v^2}\eta^2 V_t + \frac{\partial^2 f}{\partial v \partial x}\eta\rho V_t = 0$$

With remembering that $f(X_T^C, V_T, T) = M_T = e^{iuX_T^C}$ (terminal condition), f holds the so-called Fokker Planck Forward equation

$$\frac{1}{2}\frac{\partial^2 f(x,v,t)}{\partial x^2}v + \frac{1}{2}\frac{\partial^2 f(x,v,t)}{\partial v^2}\eta^2 v + \frac{\partial^2 f(x,v,t)}{\partial v \partial x}\eta\rho v$$
$$+ \frac{\partial f(x,v,t)}{\partial x}(r - \lambda k - \frac{1}{2}v) + \frac{\partial f(x,v,t)}{\partial v}\alpha(\overline{V} - v) + \frac{\partial f(x,v,t)}{\partial t} = 0$$

with the terminal condition

$$f(x,v,T) = e^{iux}$$

With abbreviated notations,

$$\frac{1}{2}f''_{xx}v+\frac{1}{2}f''_{vv}\eta^2 v+f''_{vx}\eta\rho v+f'_x(r-\lambda k-\frac{1}{2}v)+f'_v\alpha(\overline{V}-v)+f'_t=0 \qquad \text{(V-3)}$$

$$f(x,v,T)=e^{iux}$$

To solve (V-3) we guess a function *f* of the form

$$f_u(x,v,t)=e^{C(T-t)+vD(T-t)}e^{iux}$$

where C and D are functions of the time. This guess relies in fact on the linearity of the coefficients in the equation (V-3), provided that we have some techniques to solve such PDE's with affine coefficients. This choice permits to translate (V-3) in a further closed system of differential equations that we can solve by Ricatty in our purpose or more generally by some theorems provided by Duffie Pan Singleton in the general case of affine jump diffusions. In effect there is an infinity possible solutions to the backward Fokker-Planck equation, and we must target a precise form that permit closed calculus to resolve it.

With this substitution which reduces the range of solutions, the terminal condition becomes naturally

$$C(0)=0 \text{ and } D(0)=0$$

and the partial derivatives can be calculated

$$f'_t=(-C'(T-t)-vD'(T-t))e^{iux}e^{C(T-t)+vD(T-t)}$$

$$f'_x=iue^{iux}e^{C(T-t)+vD(T-t)}$$

$$f'_v=D(T-t)e^{iux}e^{C(T-t)+vD(T-t)}$$

$$f'_{vx}=iue^{iux}D(T-t)e^{C(T-t)+vD(T-t)}$$

$$f''_{vv}=(D(T-t))^2e^{iux}e^{C(T-t)+vD(T-t)}$$

$$f''_{xx}=-u^2e^{iux}e^{C(T-t)+vD(T-t)}$$

So the equation (V-3) becomes

$$-\frac{1}{2}u^2 v e^{iux+C(T-t)+vD(T-t)}+\frac{1}{2}\eta^2 vD^2(T-t)e^{iux+C(T-t)+vD(T-t)}$$
$$+iuD(T-t)e^{iux+C(T-t)+vD(T-t)}\eta\rho v+iue^{iux+C(T-t)+vD(T-t)}(r-\lambda k-\frac{1}{2}v)$$
$$+D(T-t)e^{iux+C(T-t)+vD(T-t)}\alpha(\overline{V}-v)-(C^{'}(T-t)+vD^{'}(T-t))e^{iux+C(T-t)+vD(T-t)}=0$$

By factorizing $e^{iux+C(T-t)+vD(T-t)}$ and remarking it always different from zero,

$$-\frac{1}{2}u^2 v+\frac{1}{2}\eta^2 vD^2(T-t)+iu\eta\rho vD(T-t)+iu(r-\lambda k-\frac{1}{2}v)+D(T-t)\alpha(\overline{V}-v)$$
$$-(C^{'}(T-t)+vD^{'}(T-t))=0$$

and by rearranging the terms

$$\left[-\frac{1}{2}u^2+\frac{1}{2}\eta^2 D^2(T-t)+iu\eta\rho D(T-t)-\frac{1}{2}iu-\alpha D(T-t)-D^{'}(T-t)\right]v+iu(r-\lambda k)+D(T-t)\alpha\overline{V}$$
$$-C^{'}(T-t)=0$$

Consequently [48], we obtain the two following ordinary differential equations

$$-\frac{1}{2}u^2+\frac{1}{2}\eta^2 D^2(T-t)+iu\eta\rho D(T-t)-\frac{1}{2}iu-\alpha D(T-t)-D^{'}(T-t)=0$$
$$iu(r-\lambda k)+D(T-t)\alpha\overline{V}-C^{'}(T-t)=0$$

Let $\tau=T-t$, then

$$D^{'}(\tau)=\frac{1}{2}\eta^2 D^2(\tau)-(\alpha-iu\eta\rho)D(\tau)-\frac{1}{2}u(u+i)$$
$$C^{'}(\tau)=\alpha\overline{V}D(\tau)+iu(r-\lambda k)$$

[48] $ax+b=0$ for all x if and only if a=0 and b=0 (we don't want $ax+b=0$ for $x=-\frac{b}{a}$ only)

We identify a structure *à la* Riccati in the first equation, that we might solve according to the theorem of Riccati with the substitution

$$D=-\frac{h^{'}}{\frac{\eta^{2}}{2}h}=-\frac{2}{\eta^{2}}\frac{h^{'}}{h}$$

So remarking

$$D^{'}=-\frac{2}{\eta^{2}}\frac{(h^{''}h-(h')^{2})}{h^{2}}$$

We obtain with little algebra the second order ordinary differential equation

$$\frac{2}{\eta^{2}}\frac{h^{''}}{h}-2\frac{(iu\eta\rho-\alpha)}{\eta^{2}}\frac{h^{'}}{h}-\frac{1}{2}u(u+i)=0$$

$$h^{''}+(\alpha-iu\eta\rho)h^{'}-\frac{\eta^{2}}{4}u(u+i)h=0 \qquad \text{(V-4)}$$

This equation is well known and its general solution is of the form

$$h(\tau)=Ae^{a_{1}\tau}+Be^{a_{2}\tau}$$

In effect, we target first a trivial particular solution of the form $h(\tau)=e^{a\tau}$ [49]

Then $h^{'}(\tau)=ae^{a\tau}$ and $h^{''}(\tau)=a^{2}e^{a\tau}$

Substituting it into (V-4), and suppressing the non zero factor $e^{\alpha\tau}$,

$$a^{2}+(\alpha-iu\eta\rho)a-\frac{\eta^{2}}{4}u(u+i)=0$$

The square delta of the polynomial function is

$$d=\sqrt{(\alpha-iu\eta\rho)^{2}+\eta^{2}u(u+i)}$$ [50]

[49] This is a well known method in classical differential calculus the informed reader might solve directly.

[50] The square is the complex one whatever is the sign of the amount under the square root. If $a<0$, $\sqrt{a}$ means $-i\sqrt{-a}$ in the classical sens of the square root.

and the solutions are

$$a_1 = \frac{-(\alpha - iu\eta\rho) + d}{2} \qquad\qquad a_2 = \frac{-(\alpha - iu\eta\rho) - d}{2}$$

Consequently, $h(\tau) = e^{a_1 \tau}$ and $h(\tau) = e^{a_2 \tau}$ are particular solutions of (V-4). By trivial linearity, any solution of the form $h(\tau) = Ae^{a_1\tau} + Be^{a_2\tau}$ is a solution of (V-4).

Finally, the general form of the solution of (V-4) is

$$h(\tau) = Ae^{\frac{-(\alpha - iu\eta\rho)+d}{2}\tau} + Be^{\frac{-(\alpha - iu\eta\rho)-d}{2}\tau} = Ae^{a_1\tau} + Be^{a_2\tau}$$

and we also have

$$h'(\tau) = a_1 Ae^{a_1\tau} + a_2 Be^{a_2\tau}$$

We now have to use terminal conditions to obtain the values of A and B

$$D(0) = 0 = -\frac{h'(0)}{\frac{\eta^2}{2}h(0)} \qquad \text{so} \qquad h'(0) = 0 \qquad \text{so} \qquad a_1 A + a_2 B = 0$$

In the other hand, we notice that $h(0) = A + B$

We obtain then a system of two unknowns, assuming we know $h(0)$

$$\begin{cases} a_1 A + a_2 B = 0 \\ A \quad + B \quad = h(0) \end{cases}$$

of solutions

$$A = \frac{a_2}{(a_2 - a_1)} h(0) \qquad\qquad B = \frac{a_1}{(a_1 - a_2)} h(0)$$

Now the only unknown left is the value of $h(0)$. We won't need to explicit its value because it disappears in the calculation of D.

$$D(\tau)=-\frac{2}{\eta^2}\frac{h'(\tau)}{h(\tau)}=-\frac{2}{\eta^2}\frac{a_1Ae^{a_1\tau}+a_2Be^{a_2\tau}}{Ae^{a_1\tau}+Be^{a_2\tau}}=-\frac{2}{\eta^2}\frac{a_1\frac{a_2}{(a_2-a_1)}h(0)e^{a_1\tau}-a_2\frac{a_1}{(a_2-a_1)}h(0)e^{a_2\tau}}{\frac{a_2}{(a_2-a_1)}h(0)e^{a_1\tau}-\frac{a_1}{(a_2-a_1)}h(0)e^{a_2\tau}}$$

$$D(\tau)=-\frac{2}{\eta^2}\frac{a_1a_2(e^{a_1\tau}-e^{a_2\tau})}{a_2e^{a_1\tau}-a_1e^{a_2\tau}}$$

The second equation $C^{'}(\tau)=\alpha\overline{V}D(\tau)+iu(r-\lambda k)$ is an ODE of first degree from which we can obtain the solution with a mere integration.

Using $D=-\frac{2}{\eta^2}\frac{h^{'}}{h}$, it gives for a τ_0 for which $\ln(h(\tau_0)$ is defined

$$\int_{\tau_0}^{\tau}C'(s)ds=-2\frac{\alpha\overline{V}}{\eta^2}\int_{\tau_0}^{\tau}\frac{h^{'}(s)}{h(s)}ds+iu(r-\lambda k)\int_{\tau_0}^{\tau}ds$$

$$C(\tau)-C(\tau_0)=-2\frac{\alpha\overline{V}}{\eta^2}[\ln(h(\tau))-\ln(h(\tau_0))]+iu(r-\lambda k)(\tau-\tau_0)$$

$$C(\tau)=-2\frac{\alpha\overline{V}}{\eta^2}\ln(h(\tau))+iu(r-\lambda k)\tau+C(\tau_0)+2\frac{\alpha\overline{V}}{\eta^2}\ln(h(\tau_0))-iu(r-\lambda k)\tau_0$$

With denoting $R=C(\tau_0)+2\frac{\alpha\overline{V}}{\eta^2}\ln(h(\tau_0)-iu(r-\lambda k)\tau_0$,

$$C(\tau)=-2\frac{\alpha\overline{V}}{\eta^2}\ln(h(\tau))+iu(r-\lambda k)\tau+R$$

To know the value of *R*, we use the terminal condition $C(0)=0$.

$$R=2\frac{\alpha\overline{V}}{\eta^2}\ln(h(0))$$

Although $h(0)$ is unknown, it disappears in the calculation of *C*

$$C(\tau) = -2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{a_2}{(a_2 - a_1)}h(0)e^{a_1\tau} + \frac{a_1}{(a_1 - a_2)}h(0)e^{a_2\tau}) + iu\tau(r - \lambda k) + 2\frac{\alpha\overline{V}}{\eta^2}\ln(h(0))$$

$$C(\tau) = -2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{a_2}{(a_2 - a_1)}e^{a_1\tau} + \frac{a_1}{(a_1 - a_2)}e^{a_2\tau}) + iu\tau(r - \lambda k)$$

$$C(\tau) = -2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{a_2 e^{a_1\tau} - a_1 e^{a_2\tau}}{a_2 - a_1}) + iu\tau(r - \lambda k)$$

We obtain finally the value of D and C

$$D(\tau) = -\frac{2}{\eta^2}\frac{a_1 a_2 (e^{a_1\tau} - e^{a_2\tau})}{a_2 e^{a_1\tau} - a_1 e^{a_2\tau}}$$

$$C(\tau) = iu(r - \lambda k)\tau - 2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{a_2 e^{a_1\tau} - a_1 e^{a_2\tau}}{a_2 - a_1})$$

with $\quad a_1 = \dfrac{-(\alpha - iu\eta\rho) + d}{2} \qquad a_2 = \dfrac{-(\alpha - iu\eta\rho) - d}{2}$

and $\quad d = \sqrt{(\alpha - iu\eta\rho)^2 + \eta^2 u(u+i)}$

We can go further with the form of the functions C and D, in order to raise the precision of the results obtained by numerical simulation.

For the calculus of D, we point out first that

$$a_1 a_2 = \frac{1}{4}(-(\alpha - iu\eta\rho) - d)(-(\alpha - iu\eta\rho) + d) = \frac{1}{4}((\alpha - iu\eta\rho)^2 - d^2) = -\frac{1}{4}\eta^2 u(u+i)$$

With factorizing $e^{\frac{(a_1 + a_2)t}{2}}$ in the numerator and the denominator, we obtain then with little algebra

$$D(\tau) = \frac{1}{2}u(u+i)\frac{e^{\frac{a_1 - a_2}{2}\tau} - e^{-\frac{a_1 - a_2}{2}\tau}}{a_2 e^{\frac{a_1 - a_2}{2}\tau} - a_1 e^{-\frac{a_1 - a_2}{2}\tau}}$$

and with remarking $d = a_1 - a_2$

$$D(\tau) = \frac{1}{2}u(u+i)\frac{e^{\frac{d\tau}{2}} - e^{-\frac{d\tau}{2}}}{a_2 e^{\frac{d\tau}{2}} - a_1 e^{-\frac{d\tau}{2}}}$$

$$D(\tau) = \frac{1}{2}u(u+i)\frac{e^{\frac{d\tau}{2}} - e^{-\frac{d\tau}{2}}}{-\frac{1}{2}d(e^{\frac{d\tau}{2}} + e^{-\frac{d\tau}{2}}) - \frac{1}{2}(\alpha - iu\eta\rho)(e^{\frac{d\tau}{2}} - e^{-\frac{d\tau}{2}})}$$

With using the hyperbolic cotangent function,

$$\coth(x) = \frac{\cosh(x)}{\sinh(x)} = \frac{\frac{1}{2}(e^x + e^{-x})}{\frac{1}{2}(e^x - e^{-x})} = \frac{(e^x + e^{-x})}{(e^x - e^{-x})}$$

we obtain

$$D(\tau) = -\frac{u(u+i)}{d\coth(\frac{d\tau}{2}) + (\alpha - iu\eta\rho)}$$

Using the same techniques of calculus

$$C(\tau) = iu\tau(r - \lambda k) - 2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{a_2 e^{a_1\tau} - a_1 e^{a_2\tau}}{a_2 - a_1})$$

$$C(\tau) = iu\tau(r - \lambda k) - 2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{e^{-\frac{(\alpha - iu\eta\rho)\tau}{2}}(a_1 e^{-\frac{d\tau}{2}} - a_2 e^{\frac{d\tau}{2}})}{d})$$

$$C(\tau) = iu\tau(r - \lambda k) - 2\frac{\alpha\overline{V}}{\eta^2}\left(\frac{-(\alpha - iu\eta\rho)\tau}{2} + \ln(\frac{a_1 e^{-\frac{d\tau}{2}} - a_2 e^{\frac{d\tau}{2}}}{d})\right)$$

$$C(\tau) = iu\tau(r - \lambda k) + \frac{\alpha\overline{V}\tau(\alpha - iu\eta\rho)}{\eta^2} - 2\frac{\alpha\overline{V}}{\eta^2}\ln(\frac{a_1 e^{-\frac{d\tau}{2}} - a_2 e^{\frac{d\tau}{2}}}{d})$$

But

$$\frac{a_1 e^{-\frac{d\tau}{2}} - a_2 e^{\frac{d\tau}{2}}}{d} = \frac{\frac{(a_1-a_2)}{2} e^{-\frac{d\tau}{2}} + \frac{(a_1-a_2)}{2} e^{\frac{d\tau}{2}} + \frac{(a_1+a_2)}{2} e^{\frac{-d\tau}{2}} - \frac{(a_1+a_2)}{2} e^{\frac{d\tau}{2}}}{d}$$

$$\frac{a_1 e^{-\frac{d\tau}{2}} - a_2 e^{\frac{d\tau}{2}}}{d} = \frac{\frac{(a_1-a_2)}{2} e^{-\frac{d\tau}{2}} + \frac{(a_1-a_2)}{2} e^{\frac{d\tau}{2}} + \frac{(a_1+a_2)}{2} e^{\frac{-d\tau}{2}} - \frac{(a_1+a_2)}{2} e^{\frac{d\tau}{2}}}{d}$$

$$\frac{a_1 e^{-\frac{d\tau}{2}} - a_2 e^{\frac{d\tau}{2}}}{d} = \frac{\frac{d}{2}(e^{-\frac{d\tau}{2}} + e^{\frac{d\tau}{2}})}{d} - \frac{\frac{(a_1+a_2)}{2}(e^{\frac{d\tau}{2}} - e^{-\frac{d\tau}{2}})}{d} = \cosh(\frac{d\tau}{2}) + \frac{(\alpha - iu\eta\rho)}{d}\sinh(\frac{d\tau}{2})$$

so we can conclude

$$C(\tau) = iu\tau(r-\lambda k) + \alpha\tau\overline{V}\,\frac{(\alpha - iu\eta\rho)}{\eta^2} - \frac{2\alpha\overline{V}}{\eta^2}\ln\left(\cosh(\frac{d\tau}{2}) + \frac{\alpha - iu\eta\rho}{d}\sinh(\frac{d\tau}{2})\right)$$

Then

$$\phi_{X_T^C}(u) = f(x_0, v_0, 0) = e^{C(T) + vD(T) + iux_0}$$

$$\phi_{X_T^C}(u) = e^{iuT(r-\lambda k) + \alpha T\overline{V}\frac{(\alpha - iu\eta\rho)}{\eta^2} - \frac{2\alpha\overline{V}}{\eta^2}\ln\left(\cosh(\frac{dT}{2}) + \frac{\alpha - iu\eta\rho}{d}\sinh(\frac{dT}{2})\right) - v_0\frac{u(u+i)}{d\coth(\frac{dT}{2}) + (\alpha - iu\eta\rho)} + iux_0}$$

$$\phi_{X_T^C}(u) = \frac{e^{iuT(r-\lambda k) + \alpha T\overline{V}\frac{(\alpha - iu\eta\rho)}{\eta^2} - v_0\frac{u(u+i)}{d\coth(\frac{dT}{2}) + (\alpha - iu\eta\rho)} + iux_0}}{\left(\cosh(\frac{dT}{2}) + \frac{\alpha - iu\eta\rho}{d}\sinh(\frac{dT}{2})\right)^{\frac{2\alpha\overline{V}}{\eta^2}}}$$

Let's now calculate the characteristic function of the jumps, $\Phi_{\hat{Z}_t}(u) = E(e^{iu\hat{Z}_t})$ where $\hat{Z}_t$ is a compound Poisson process of Gaussian distribution of jumps size $\hat{Y}_i \sim N(\ln(1+k) - \frac{1}{2}\sigma^2, \sigma^2)$.

By Fourier representation of compound Poisson processes,

$$\Phi_{\hat{Z}_t}(u) = e^{t\lambda \int_{\Re} (e^{iux} - 1) F(dx)}$$

where F is the distribution of $\hat{Y}_i$. So consequently,

$$\Phi_{\hat{Z}_t}(u) = e^{t\lambda \int_{\Re} e^{iux} F(dx) - t\lambda \int_{\Re} F(dx)} = e^{t\lambda \Phi_{\hat{Y}_i}(u) - t\lambda}$$

where $\Phi_{\hat{Y}_i}$ is the characteristic function of $\hat{Y}_i$. $\hat{Y}_i$ is Gaussian, so we can use Laplace transform,

$$\Phi_{\hat{Y}_i}(u) = e^{iu(\ln(1+k) - \frac{1}{2}\sigma^2) + \frac{1}{2}i^2 u^2 \sigma^2} = e^{iu(\ln(1+k) - \frac{1}{2}\sigma^2) - \frac{1}{2}u^2 \sigma^2}$$

So the characteristic function of the jumps is

$$\Phi_{\hat{Z}_t}(u) = e^{t\lambda (e^{-\frac{1}{2}u^2\sigma^2 + iu(\ln(1+k) - \frac{1}{2}\sigma^2)} - 1)}$$

We obtain after this long calculus the form of the characteristic function of the price process under the model of Bates

$$\Phi_{X_t}(u) = \frac{e^{iut(r - \lambda k) + \alpha t \bar{V} \frac{(\alpha - iu\eta\rho)}{\eta^2} - v_0 \frac{u(u+i)}{d \coth(\frac{dt}{2}) + (\alpha - iu\eta\rho)} + iux_0 + t\lambda (e^{-\frac{1}{2}u^2\sigma^2 + iu(\ln(1+k) - \frac{1}{2}\sigma^2)} - 1)}}{\left(\cosh(\frac{dt}{2}) + \frac{\alpha - iu\eta\rho}{d} \sinh(\frac{dt}{2}) \right)^{\frac{2\alpha\bar{V}}{\eta^2}}}$$

with $$d = \sqrt{(\alpha - iu\eta\rho)^2 + \eta^2 u(u+i)}$$

VI- Numerical Simulation of the model of Bates

A- Monte Carlo Methods

We use the equations

$$dX_t = (r - \lambda k - \frac{1}{2}V_t)dt + \sqrt{V_t}\,dW_t^1 + d\hat{Z}_t$$

$$dV_t = \alpha(\overline{V} - V_t)dt + \eta\sqrt{V_t}\,dW_t^2$$

On a grid $\{x_0 = 0, x_1, ..., x_n = T\}$ with a chosen precision *n*, we simulate

- a 2-dimensionnal Brownian motion vector (G^1, G^2) with correlation ρ
- a compound Poisson process Z with intensity λ and jumps' sizes Y_i

We computate for $j = 1$ to $j = n$

$$X_{x_j} = X_{x_{j-1}} + (r - \lambda k - \frac{1}{2}V_{x_{j-1}})(x_j - x_{j-1}) + \sqrt{V_{x_{j-1}}}\,G^1_{x_{j-1}} + Z_{x_{j-1}}$$

$$V_{x_j} = \alpha(\overline{V} - V_{x_{j-1}})(x_j - x_{j-1}) + \eta\sqrt{V_{x_{j-1}}}\,G^2_{x_{j-1}}$$

where X_0 and V_0 are a given parameters of the model.

Then we calculate $e^{-r(T-t)}h(e^{X_t})$ which is the price of the option. If the payoff is $h \to (e^{X_t} - K)^+$ (plain vanilla call) then $e^{-r(T-t)}(X_T - K)^+$ is one value of the premium of the call.

We repeat this operation 100 000 times, making hence 100 000 scenarios of evolution of the price, and we take the final average value of the premium.

In order to reduce the variance, we can use the method of the antithetic variable.

B- Semi-closed form pricing formula with FFT algorithm

We will see in this part the famous technique of Carr and Madan to compute numerically the price of an option knowing the characteristic function of the log price of the underlying X_t, or the characteristic function of the logarithm of the price. This technique is based on FFT transform as the pricing algorithm complexity for a whole range of strikes is comparable to the one of Black and Scholes model. This approach is systematically used when there is no closed form of the density of $X_t = \ln(S_t)$, but only a closed form of its characteristic function.

Let's consider the following premium of a plain vanilla call, with denoting f_{S_T} the density of S_T and f_{X_T} the density of X_T under Q,

$$C_T(K) = e^{-rT} E_Q((S_T - K)^+ \mid F_0)$$

$$C_T(K) = e^{-rT} E_Q((S_T - K)^+) = e^{-rT} E_Q((S_T - K)1_{S_T \geq K})$$

$$C_T(K) = \int_{-\infty}^{+\infty} e^{-rT}(s - K)1_{s \geq K} f_{S_T}(s) ds$$

$$C_T(K) = \int_{K}^{+\infty} e^{-rT}(s - K) f_{S_T}(s) ds$$

With a mere variable changing, we use the log price and we set $k = \ln K$,

$$C_T(k) = \int_{k}^{+\infty} e^{-rT}(e^s - e^k) f_{X_T}(s) ds$$

Now we would like to calculate the Fourier transform of the price of the call.

$$\phi_T(v) = \int_{\mathbb{R}} e^{ivk} C_T(k) dk$$

but this integral doesn't make sense since C_T is not square integrable. [51]

In effect, by Lebesgue theorem

[51] A necessary condition for C_T to be square integrable is that $\lim\limits_{k \to -\infty} C_T(k) = 0$

$$\lim_{k \to -\infty} C_T(k) = \int_{-\infty}^{+\infty} e^{-rT}(e^s - 0) f_{X_T}(s) ds$$

and considering $e^{-rt} S_t$ is a martingale

$$\lim_{k \to -\infty} C_T(k) = e^{-rT} E_Q(e^{X_T}) = E_Q(e^{-rT} S_T \mid F_0) = e^{-r.0} S_0 = S_0 > 0$$

Then

$$C_T(k) \geq S_0 \qquad \text{and} \qquad \int_{\mathbb{R}} C_T(k) dk \; \geq \; S_0 \int_{\mathbb{R}} dk = +\infty$$

Hence, C_T is not integrable and consequently not square integrable so its Fourier Transform doesn't exist.

We consider then a modified function $c_T(k) = e^{\alpha k} C_T(k)$ which is square integrable for a suitable α that must be calibrated with the others parameters of the model.

Now we get the Fourier Transform of the modified premium c_T

$$\phi_T(v) = \int_{\Re} e^{ivk} c_T(k) dk$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} e^{ivk} \left(e^{\alpha k} \int_{k}^{+\infty} (e^s - e^k) f_{X_T}(s) ds \right) dk$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} \int_{k}^{+\infty} e^{ivk} e^{\alpha k} (e^s - e^k) f_{X_T}(s) ds dk$$

We will now interchange the integrals with remarking that

$$-\infty \leq k \leq +\infty \quad \text{and} \quad k \leq s \leq +\infty$$

so the reverse integration area is given by

$$-\infty \leq s \leq +\infty \quad \text{and} \quad -\infty \leq k \leq s$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} \int_{-\infty}^{s} e^{ivk} e^{\alpha k} (e^s - e^k) f_{X_T}(s) dk ds$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} \int_{-\infty}^{s} e^{ivk} e^{\alpha k} (e^s - e^k) f_{X_T}(s) dk ds$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} f_{X_T}(s) \left(\int_{-\infty}^{s} (e^s e^{(\alpha+iv)k} - e^{(\alpha+1+iv)k}) dk \right) ds$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} f_{X_T}(s) \left[\frac{e^s}{\alpha+iv} e^{(\alpha+iv)k} - \frac{e^{(\alpha+1+iv)k}}{\alpha+1+iv} \right]_{-\infty}^{s} ds$$

$$\phi_T(v) = e^{-rT} \int_{-\infty}^{+\infty} \left(\frac{e^{(\alpha+iv+1)s}}{\alpha+iv} - \frac{e^{(\alpha+1+iv)s}}{\alpha+1+iv} \right) f_{X_T}(s) ds$$

$$\phi_T(v) = e^{-rT} E_Q\left(\frac{e^{(\alpha+iv+1)X_T}}{\alpha+iv} - \frac{e^{(\alpha+1+iv)X_T}}{\alpha+1+iv} \right)$$

$$\phi_T(v) = e^{-rT} E_Q\left(\frac{(\alpha+1+iv)e^{(\alpha+iv+1)X_T} - (\alpha+iv)e^{(\alpha+1+iv)X_T}}{(\alpha+iv)(\alpha+1+iv)} \right)$$

$$\phi_T(v) = e^{-rT} E_Q\left(\frac{(\alpha+iv+1)e^{X_T}}{\alpha^2+\alpha+(2\alpha+1)iv-v^2} \right)$$

$$\phi_T(v) = e^{-rT} E_Q\left(\frac{i(-i\alpha+v+-i)e^{X_T}}{\alpha^2+\alpha+(2\alpha+1)iv-v^2} \right.$$

$$\phi_T(v) = e^{-rT} E_Q\left(\frac{e^{i(v-i(\alpha+1))X_T}}{\alpha^2+\alpha+(2\alpha+1)iv-v^2} \right)$$

So we obtain

$$\phi_T(v) = e^{-rT} \frac{\Phi_{X_T}(v-i(\alpha+1))}{\alpha^2+\alpha+(2\alpha+1)iv-v^2}$$

where $\Phi_{X_T}(u) = E_Q(e^{iuX_T})$ is the characteristic function of X_T

By inverse Fourier Transform, we obtain a first form of the modified premium

$$c_T(k) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ivk} \phi_T(v) dv = \frac{e^{-rT}}{2\pi} \int_{\mathbb{R}} e^{-ivk} \frac{\Phi_{X_T}(v-i(\alpha+1))}{\alpha^2+\alpha+(2\alpha+1)iv-v^2} dv$$

We know that a function is real if and only if the real part of its Fourier transform is even and the imaginary part is odd. As $c_T(k) = e^{\alpha k} C_T(k)$ is a real function, we have

$$\mathrm{Re}(\phi_T(-v)) = \mathrm{Re}(\phi_T(v)) \qquad \text{and} \qquad \mathrm{Im}(\phi_T(-v)) = -\mathrm{Im}(\phi_T(v))$$

Then

$$c_T(k) = \mathrm{Re}(c_T(k)) = \frac{1}{2\pi}\int_{\Re} \mathrm{Re}(e^{-ivk}\phi_T(v))dv = \frac{1}{2\pi}\int_{\Re} \mathrm{Re}(e^{-ivk}\phi_T(v))dv$$

Set $\qquad h(v) = \mathrm{Re}(e^{-ivk}\phi_T(v))$

Then $\qquad h(v) = \mathrm{Re}(e^{-ivk})\mathrm{Re}(\phi_T(v)) - \mathrm{Im}(e^{-ivk})\mathrm{Im}(\phi_T(v))$ [52]

$$h(v) = \cos(-vk)\mathrm{Re}(\phi_T(v)) - \sin(-vk)\mathrm{Im}(\phi_T(v))$$

$$h(v) = \cos(vk)\mathrm{Re}(\phi_T(v)) + \sin(vk)\mathrm{Im}(\phi_T(v))$$

And so $\qquad h(-v) = \cos(-vk)\mathrm{Re}(\phi_T(-v)) + \sin(-vk)\mathrm{Im}(\phi_T(-v))$

$$h(-v) = \cos(vk)\mathrm{Re}(\phi_T(v)) + \sin(vk)\mathrm{Im}(\phi_T(v)) = h(v)$$

Using the fact that $h : v \to \mathrm{Re}(e^{-ivk}\phi_T(v))$ is even [53], we obtain the last form of the call price, expressed as a function of the characteristic function of X_T

$$C_T(k) = \frac{e^{-\alpha k}}{\pi}\int_0^{+\infty} e^{-ivk}\phi_T(v)dv = \frac{e^{-rT-\alpha k}}{\pi}.\int_0^{+\infty} e^{-ivk}\frac{\Phi_{X_T}(v - i(\alpha+1))}{\alpha^2 + \alpha + (2\alpha+1)iv - v^2}dv$$

This formula is semi-closed because all is known except the value of the integral that we simulate numerically. This expression will be computed using the FFT algorithm.

We can first numerically compute C_T on a uniform grid $\{x_0,..,x_j,..,x_N\}$ where $dx = x_{i+1} - x_i$

[52] For any complex numbers $u = a + ib$ and $v = c + id$
$\mathrm{Re}(uv) = \mathrm{Re}((a+ib)(c+id)) = ab - dc = \mathrm{Re}(u)\mathrm{Re}(v) - \mathrm{Im}(u)\mathrm{Im}(v)$

[53] If f is even then $f(x) = f(-x)$ so $\int_{-\infty}^{+\infty} f(x)dx = \int_0^{+\infty} f(x)dx + \int_{-\infty}^{0} f(x)dx = \int_0^{+\infty} f(x)dx - \int_{+\infty}^{0} f(-x)dx = 2\int_0^{+\infty} f(x)dx$

$$C_T(k)=\frac{e^{-\alpha k}}{\pi}\int_0^{+\infty} e^{-ivk}\phi_T(v)dv$$

as $$C_T(k)\approx\frac{e^{-\alpha k}}{\pi}\sum_{j=0}^{N-1} e^{-ix_j k}\phi_T(x_j)(x_{j+1}-x_j)=\frac{e^{-\alpha k}}{\pi}\sum_{j=0}^{N-1} e^{-ix_j k}\phi_T(x_j)\Delta x$$

where $O(N^2)$ operations are needed .

FFT algorithm provides a $O(n\log(n))$ number of operations for the computation Discrete Fourier integrals of the form

$$w_j=\sum_{u=0}^{N-1} e^{-i\frac{2\pi u}{N}j}x_u$$

To apply FFT algorithm to our premium C_T, we consider then a grid of log strikes distributed around the log spot price X_0 with a distance $\Delta_k(N)=\Delta_N=k_{j+1}-k_j$,

$$k_j=-\frac{1}{2}N\Delta_k+\Delta_k j+X_0 \qquad \text{with} \qquad j\in[0,N-1]$$

Then we approximate the premium of a call of strike k_j with the discrete Fourier transform

$$C_T(k_j)\approx\frac{e^{-\alpha k_j}}{\pi}\sum_{u=0}^{N-1} e^{-ix_u k_j}\phi_T(x_u)\Delta x$$

$$C_T(k_j)\approx\frac{e^{-\alpha k_j}}{\pi}\sum_{u=0}^{N-1} e^{-ix_u\Delta_k j}e^{ix_u(\frac{1}{2}N\Delta_k-X_0)}\phi_T(x_u)\Delta x$$

With choosing $\Delta_k=\frac{u}{x_u}\frac{2\pi}{N}$, i.e $x_u\Delta_k=u\frac{2\pi}{N}$

$$C_T(k_j)\approx=\frac{e^{-\alpha k_j}}{\pi}\sum_{u=0}^{N-1} e^{-i\frac{2\pi u}{N}j}e^{ix_u(\frac{1}{2}N\Delta_k-X_0)}\phi_T(x_u)\Delta x$$

So we can compute FFT with

$$w_j = C_T(k_j)$$

and

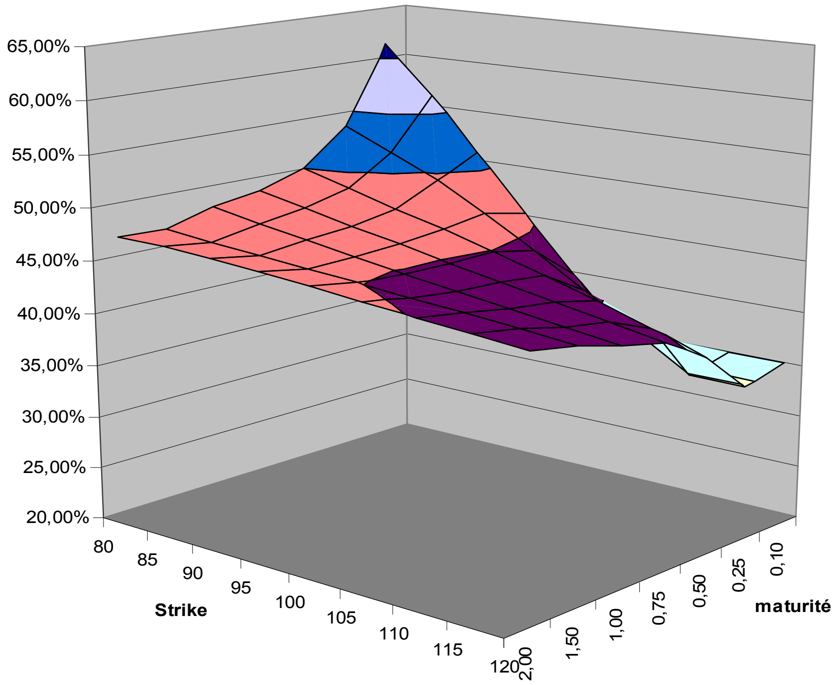
$$y_u = e^{ix_u(\frac{1}{2}N\Delta_k - X_0)} \phi_T(x_u)\Delta x .$$

With this method and the choice of

$$\Delta_k = \frac{u}{x_u}\frac{2\pi}{N}$$

we have to deal with the following trade-off : the choice of a high N will affect the time of execution of the routine while it while give an accurate range a strikes round the strike of S_0. We have to arbitrate between the density of the strikes and the computational time.

C- Numerical simulation of a surface of volatility



Implied volatility surface – Model of Bates (FFT)

Parameters

r = 0.05 (risk free rate)

λ=1 (average number of jumps per year)

k = -0.25 (average size of jumps in %)

σ=0.1 (volatility of jumps)

α = 4 (mean reverting)

V_0=0.09 (initial volatility)

$\bar{V}$ = 0.09 (long volatility)

η=1 (volatility of volatility)

ρ= -0.7 (leverage effect)

α_{FFT}= 0.75 (parameter of the FFT)

Bates' model :

$$dX_t = (r - \lambda k - \frac{1}{2}V_t)dt + \sqrt{V_t}\,dW_t^1 + d\hat{Z}_t$$

$$dV_t = \alpha(\bar{V} - V_t)dt + \eta\sqrt{V_t}\,dW_t^2$$

$$< dW_t^1, dW_t^2 >= \rho.dt$$

$$\hat{Z}_t = \sum_{i=1}^{N_t} \hat{Y}_i\,, \qquad \hat{Y}_i \sim N(\ln(1+k) - \frac{1}{2}\sigma^2, \sigma^2)\ ,\ N_t \sim P(\lambda)$$

Annexes

Illustration of classic Itô Calculus

The more famous application of such calculus is of course the Black-Scholes Formula, from which we derive closed-form expression of S_t from the equation of diffusion.

In this model, $\mu(t,S_t) = \mu.S_t$ and $\sigma(t,S_t) = \sigma.S_t$ where μ and σ are constant reals

$$dS_t = \mu.S_t.dt + \sigma.S_t.dW_t$$

which we denote it in the more usual form

$$\frac{dS_t}{S_t} = \mu.dt + \sigma.dW_t$$

to focus on the evolution of the return of the asset as a linear evolution over time (linear capitalization with the rate μ) plus a volatility multiplied by a normal distribution centered in 0 with variance *t*.

Applying Itô's formula with $f(x) = \ln(x)$, we obtain

$$d\ln(S_t) = \frac{1}{S_t}.dS_t + \frac{1}{2}(-\frac{1}{S_t^2}).\sigma^2.S_t^2.dt$$

$$\ln(S_t) - \ln(S_0) = \frac{1}{S_t}\mu.S_t.dt + \frac{1}{S_t}\sigma.S_t.dW_t - \frac{1}{2}.\sigma^2.dt$$

$$\ln(S_t) - \ln(S_0) = \mu.dt + \sigma.dW_t - \frac{1}{2}.\sigma^2.dt$$

But $$\int_0^t (\mu - \frac{1}{2}\sigma^2)dt = (\mu - \frac{1}{2}\sigma^2)t$$

and $$\int_0^t \sigma.dW_t = \sigma.\int_0^t dW_t = \sigma.(W_t - W_0) = \sigma.W_t$$ [54]

Then we obtain the closed-from formula of S_t,

$$S_t = S_0.e^{(\mu - \frac{1}{2}.\sigma^2)t + \sigma.W_t}$$

[54] $W_0 \sim N(0,0)$ so is a.s equal to 0.

Construction of a Poisson process from a Radon measure

For any Radon measure μ on $E \subset \Re$, there exists a Poisson measure *M* on *E* with intensity μ. In effect, it is sufficient to take

- a sequel of i.i.d random variables so that $p(X_i \in A) = \dfrac{\mu(A)}{\mu(E)}$
- $M(E)$ as a Poisson random variable with intensity $\mu(E)$, $M(E)$ independent of the X_i.

We construct then $M(A) = \sum\limits_{i=1}^{M(E)} 1_A(X_i)$

We obtain :

$$p(M(A)=k) = \sum_{n=1}^{\infty} p(M(A)=k \mid M(E)=n).p(M(E)=n)$$

$$p(M(A)=k) = \sum_{n=1}^{\infty} p(\sum_{i=1}^{M(E)} 1_A(X_i)=k \mid M(E)=n).p(M(E)=n)$$

$$p(M(A)=k) = \sum_{n=1}^{\infty} p(\sum_{i=1}^{n} 1_A(X_i)=k).p(M(E)=n)$$

We know that $$p(M(E)=n) = \frac{\mu(E)^n}{n!} e^{-\mu(E)}$$

In the other hand, every random variable $1_A(X_i)$ takes the value 1 with probability $\dfrac{\mu(A)}{\mu(E)}$ and the value 0 with probability (1-$\dfrac{\mu(A)}{\mu(E)}$). The random variables $1_A(X_i)$ being independent,

$$p(\sum_{i=1}^{n} 1_A(X_i)=k) = C_n^k (\frac{\mu(A)}{\mu(E)})^k (1-\frac{\mu(A)}{\mu(E)})^{n-k}$$

Note that if $k > n$ this expression is null. So we can count with $n \geq k$. Then,

$$p(M(A)=k) = \sum_{n=k}^{\infty} C_n^k (\frac{\mu(A)}{\mu(E)})^k (1-\frac{\mu(A)}{\mu(E)})^{n-k} \frac{\mu(E)^n}{n!} e^{-\mu(E)}$$

$$p(M(A)=k) = \sum_{n=k}^{\infty} C_n^k (\frac{\mu(A)}{\mu(E)})^k (1-\frac{\mu(A)}{\mu(E)})^{n-k} \frac{\mu(E)^{n-k}}{n!} \mu(E)^k e^{-\mu(E)}$$

$$p(M(A)=k)=\mu(A)^k\, e^{-\mu(E)} \sum_{n=k}^{\infty} \frac{C_n^k}{n!}(\mu(E)-\mu(A))^{n-k}$$

$$p(M(A)=k)=\frac{\mu(A)^k}{k!} e^{-\mu(E)} \sum_{n=k}^{\infty} \frac{(\mu(E)-\mu(A))^{n-k}}{(n-k)!}$$

$$p(M(A)=k)=\frac{\mu(A)^k}{k!} e^{-\mu(E)} \sum_{n=0}^{\infty} \frac{(\mu(E)-\mu(A))^{n}}{n!}$$

$$p(M(A)=k)=\frac{\mu(A)^k}{k!} e^{-\mu(E)} e^{\mu(E)-\mu(A)}$$

$$p(M(A)=k)=\frac{\mu(A)^k}{k!} e^{-\mu(A)}$$

M(.) is then a Poisson process of intensity $\mu(.)$

An introduction to Lévy Khinchin representation of Lévy processes (*)

Given a process X_t with finite number of jumps, we have seen that if X_t can be represented as follow

$$X_t = \gamma t + W_t + \sum_{s\in[0,t]} \Delta X_s = \gamma t + W_t + \int_{[0,t]\times\Re} x J_X(ds,dx)$$

where J_X is random Poisson measure with intensity $v(dx)dt$ then X_t is a Lévy process

Conversely, when we want to represent any Lévy process with this form we face a difficulty concerning the number of jumps. It is clear that any Lévy process that has an infinite number of jumps is not càdlàg nor a Lévy process anymore. But in reality, a Lévy process can have an infinity of jumps whose size is around 0, without being in contradiction with the càdlàg rule. The behaviour of a path with a an infinity of small jumps seems close to the behaviour of a Brownian in a macro level, while in reality we have on one hand a continuous process an in the other hand a pure discontinuous jump process.

The Lévy process is said then of infinite activity, as contrary to some finite activity processes like compound Poisson processes which have a bounded number of jumps of any size.

In this case, the measure $J_X(s,x)$ "blows up" when x is around 0, and the quantity $\int_{[0,t]\times\Re} x J_X(ds,dx)$ is non finite and does not make sense anymore.

Let's truncate the jumps by an arbitrary number, say 1 and see what happens in this kind of representation, when taking to the limit $\varepsilon \to 0$.

$$X_t = \gamma t + W_t + \int_{0\le s\le t,|x|\ge 1} x J_X(ds,dx) + \int_{0\le s\le t,\varepsilon\le|x|\le 1} x J_X(ds,dx)$$

The first term has a finite number of jumps and thus is a compound Poisson process. Remark that the truncating number could have been any strictly positive number.

The second one is still a non defined infinite quantity for $\varepsilon \to 0$. But the centered version of this integral

$$\int_{0\le s\le t,\varepsilon\le|x|\le 1} x\widetilde{J}_X(ds,dx) = \int_{0\le s\le t,\varepsilon\le|x|\le 1} x(J_X(ds,dx)) - E(\int_{0\le s\le t,\varepsilon\le|x|\le 1} x(J_X(ds,dx))$$

$$\int_{0\le s\le t,\varepsilon\le|x|\le 1} x\widetilde{J}_X(ds,dx) = \int_{0\le s\le t,\varepsilon\le|x|\le 1} x(J_X(ds,dx) - v(dx)dt)$$

is a martingale and we can show that it converges by central-limit arguments.

Consequently, the expression

$$X_t = \gamma t + W_t + \int_{0\le s\le t,|x|\ge 1} xJ_X(ds,dx) + \int_{0\le s\le t,|x|\le 1} x\widetilde{J}_X(ds,dx)$$

is defined and there is a quiet technical proof that every Lévy process can be decomposed this way. This property is known as the Lévy Itô decomposition. Here X_t can be as well a single process or a *d*-dimensional vector. Let *B* be the square matrix of variance and covariance of the Brownian vector W_t and

$$A = \frac{1}{t}B.$$

In the case where $d = 1$, *A* is a scalar which value is $\frac{\mathrm{var}(W_t)}{t} = 1$.

Under some regularity conditions of the Lévy measure explained below, the triplet (A, v, γ) is called the characteristic triplet of the Lévy process X_t because there is an equivalence between the definition of the triplet and the definition of the Lévy process.

Lévy-Itô decomposition

If X_t is a Lévy process on $\Re^d$ then there exist a triplet (A, v, γ) where

$$\int_{\Re} (|x|^2 \wedge 1)v(dx) < +\infty$$

such as

$$X_t = \gamma t + W_t + \int_{0\le s\le t,|x|\ge 1} xJ_X(ds,dx) + \int_{0\le s\le t,|x|\le 1} x\widetilde{J}_X(ds,dx)$$

The regularity conditions of the measure v are summarized in the condition $\int_{\Re}(|x|^2 \wedge 1)v(dx) < +\infty$ [55] which means both $\int_{|x|\geq 1} v(dx) < +\infty$ and $\int_{|x|<1} x^2 v(dx) < +\infty$

The measure v must be first finite for jumps which size is greater than one i.e $\int_{|x|\geq 1} v(dx) < +\infty$.

Besides the divergence around 0 must be dominated by x^2

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

which means that the variance of the part of the process where jumps size is smaller than one must be finite to make sure that the small jumps are regular around 0.

In effect, we have seen that $\qquad v(dx) = \lambda F(dx)$

thus $\qquad \int_{|x|<1} x^2 v(dx) = \lambda \int_{|x|<1} x^2 F(dx) = \lambda \operatorname{var}(X_t \mid X_t < 1) < +\infty$

i.e $\qquad \int_{|x|<1} x^2 v(dx) < +\infty \Leftrightarrow \operatorname{var}(X_t \mid X_t < 1) < +\infty$

Lévy-Khinchin representation

If X_t is a Lévy process on $\Re^d$ with characteristic triplet (A, v, γ) and

$$\int_{\Re}(|x|^2 \wedge 1)v(dx) < +\infty$$

then with using usual notation of scalar product

$$E(e^{iuX_t}) = e^{t\psi(u)}$$

$$\textit{with } \psi(u) = -\frac{1}{2}u.Au + i\gamma.u + \int_{\Re^d}(e^{iux} - 1 - iu.x1_{|x|\leq 1})v(dx)$$

Another formulation of this theorem is that if X_t is a Lévy process, then there exists a triplet (A, v, γ) such that the characteristic function of X_t can be written as above. This formulation

[55] $a \wedge b = \min(a,b)$

derives directly from the Lévy-Itô representation. Note that the drift γ depends directly on the truncature chosen, which is 1 in our paper.

The proof of this property starts from the Lévy-Itô decomposition, and uses both Laplace transform for Gaussian vectors and characteristic function form of compound Poisson processes. We demonstrate the proposition for the case d=1 with some notations of multidimensional notations to make it valid in the case where X_t is a multidimensional vector.

$$\phi_{X_t}(u) = E(e^{iuX_t}) = E(e^{iu(\gamma t + W_t + \int_{0\le s\le t, |x|\ge 1} xJ_X(ds,dx) + \int_{0\le s\le t, |x|\le 1} x\tilde{J}_X(ds,dx))})$$

$$= E(e^{iu(\gamma t + W_t + \int_{0\le s\le t, \Re} xJ_X(ds,dx) - \int_{0\le s\le t, |x|\le 1} xv(dx)ds)})$$

$$= e^{iu\gamma t} E(e^{iuW_t}) E(e^{iu\int_{0\le s\le t, \Re} xJ_X(ds,dx)}) E(e^{iu\int_{0\le s\le t, |x|\le 1} xv(dx)ds})$$

This calculus exploits the independence of the Brownian motion and the jumps, and the independence of the jumps distribution Y_i.

With Laplace transform for Gaussian vectors, *B* being the covariance matrix of the vector W_t

$$E(e^{iuW_t}) = e^{iu(E(W_t) + \frac{1}{2}(i^2 u.Bu)} = e^{-\frac{1}{2}tu.Au}$$

With characteristic function representation for compound Poisson processes

$$E(e^{iu\int_{0\le s\le t, \Re} xJ_X(ds,dx)}) = e^{t\int_{\Re}(e^{iux}-1)v(dx)}$$

Besides,

$$E(e^{iu\int_{0\le s\le t, 0\le x\le 1} xv(dx)ds}) = e^{iu\int_0^t\int_0^1 v(dx)ds} = e^{iut\int_{\Re} 1_{|x|\le 1} xv(dx)}$$

Consequently,

$$\phi_{X_t}(u) = e^{iu\gamma t - \frac{1}{2} tu.Au + t\int_{\Re}(e^{iux}-1)v(dx) + iut\int_{\Re} 1_{|x|\leq 1} xvd(x)}$$

$$\phi_{X_t}(u) = e^{t(iu\gamma - \frac{1}{2} u.Au + \int_{\Re}(e^{iux}-1-iu1_{|x|\leq 1} x)v(dx))}$$

The Lévy-Khinchin formula is very useful and powerful. We can know the characteristic function of every Lévy process and compute derivative prices through FFT transform (see part VI).

Bibliography

[1] "Financial Modelling With Jump Processes", Rama Cont & Peter Tankov, 2004, Chapman & Hall

[2] " Information Economique et Marchés Financiers", Philippe Spieser, 2000, Economica

[3] "Jumps and Stochastic Volatility : Exchange Rate Processes Implicit in Deutsche Mark Options", David S.Bates, 1996, Review of Financial Studies 9, 69-107

[4] "Empirical Performance of Alternative Option Pricing Models" G.Backshi, C. Cao and Z. Chen, 1997, Journal of Finance 52

[5] "Probabilités et potentiel – Théorie des martingales", Paul André Meyer and Claude Dellacherie, Hermann

[6] "Marchés financiers en Temps Continu", Rose-Anne Dana & Monique Jeanblanc-Picqué, 1998, Economica

[7] " Probabilités", Valérie Girardin and Nikolaos Limnios, 2001, Vuibert

[8] " Calcul stochastique discontinu et optimisation d'utilité dans le cadre d'un marché à sauts", 2004, Laurent Menard, from quantitative team of SGAM

[9] "Modèles à volatilité stochastique " , Rachid Lassoued, 2000, ESSEC research paper, www.finance-research.net

[10] "Lecture 1 : Stochastic Volatility and Local Volatility", Jim Gatheral, Merrill Lynch, Case Studies in Financial Modelling Course Notes, 2002

[11] "Option Valuation using the Fast Fourier Transform", P. Carr and D. Madan, 1999, Journal of Computational Finance 2

[12] "Option pricing when underlying stock returns are discontinuous", R. Merton, 1976, Journal of Finance Economics 3

[13] "Transform analysis and asset pricing for affine jump-diffusion", D. Duffie, J. Pan and K. Singleton, 2000, Econometrica

[14] "A closed-form solution for options with stochastic volatility with applications to bond and currency options", S. Heston, 1993, Review of Financial Studies 6