Online teaching - Lecture notes "Financial derivatives and stochastic processes" M2 Probability et Finance

E. Gobet

October 14, 2021

Notations

We repeatedly use the same notations.

- Interest rate: $(r_u : u \geq 0)$
- Discount factor $D(t,T) = e^{-\int_t^T r_u \mathrm{d}u}$
- (riskless = non defaultable) Zero-coupon bond at time t for the maturity T: $B(t,T)$
- S for the cash-price
- $\mathtt{C}_t(\phi_T, T)$ for the cash-price (prix au comptant) at time t for the cashflow Φ_T delivered at T
- $\mathtt{F}_t(\phi_T, T)$ for the forward price (prix forward) at time t (and paid at T) for the cashflow Φ_T delivered at T
- $\mathtt{Fut}_t(\phi_T, T)$ for the Future price on Φ_T delivered at T

Week 3

2 Alternative modelling to pure log-normal models

2.1 Expected discounted cash flow

In the Black-scholes model and in the binomial model, we have seen that the price of a financial contract is given by the expected discounted cashflow under a probability measure such that the asset drift equals the no-risk asset drift. We would like to go beyond these two models, using some heuristic arguments, to get a pricing rule.

The question we ask is the following: given a uncertain cashflow Ψ_T paid at time T, what should be the cash-price $C_t(\Psi_T, T)$ at $t < T$ using some probabilistic arguments?

2.1.1 Heuristic/axiomatic

Such a pricing rule should be coherent with No Arbitrage principle. Therefore, it should obey to the following rules.

Linearity: $C_t(\Psi_T^1 + \Psi_T^2, T) = C_t(\Psi_T^1, T) + C_t(\Psi_T^2, T)$.

Homogeneity: $C_t(\lambda\Psi_T, T) = \lambda C_t(\Psi_T, T)$, $\forall \lambda \in \mathbb{R}$.

Monotonicity: $\Psi_T^1 \leq \Psi_T^2$ a.s. $\implies C_t(\Psi_T^1, T) \leq C_t(\Psi_T^2, T)$.

These suggest that short sales are possible. In addition the two first axioms suggest that there are infinite liquidity and no significant size effect.

2.1.2 Probabilistic modeling

Consider that the random variable Ψ_T is defined on a probability space $(\Omega, \mathcal{P}, \mathcal{F})$.

▷ First, one could set $C_0(\Psi_T, T)$ as the best predictor of Ψ_T from time $t = 0$. We recall that, for any scalar r.v. X,

$$\mathbb{E}(X) = \underset{c \in \mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\left((X - c)^2\right)$$

since $\mathbb{E}\left((X-c)^2\right) = \operatorname{Var}(X) + (\mathbb{E}(X) - c)^2$. It means that $\mathbb{E}(\Psi_T)$ is the best L_2-predictor of Ψ_T. However, taking that for a pricing rule does not match well: indeed, it does not account for the discounting effect of 1 € between 0 and T.

▷ Another candidate for being a predictor accounting for interest rate is to set

$$C_0(\Psi_T, T) := \mathbb{E}_{\mathcal{P}}\left(L_T e^{-\int_0^T r_s \mathrm{d}s} \Psi_T\right) \tag{2.1.1}$$

where L_T is a non-negative random variable. The properties of linearity and homogeneity of the cash price derive directly from the properties of expectation. The monotonicity condition is verified since $L_T \geq 0$.

2.1.3 Other constraints on L_t

The form (2.1.1) is quite general because L_T is so far; it is coherent with heuristics and it makes sense as a kind of best predictor. We now investigate whether some extra conditions on $(L_t)_{t\geq 0}$ are required. Here we assume that L_T does not depend on the cashflow Ψ_T, but depends on the maturity T.

When Ψ_T results from a cash position. It means that one invests 1 € and waits till T, and we get $e^{\int_0^T r_s \mathrm{d}s} = \Psi_T$ € . It gives that on the one hand, $C_0(\Psi_T, T) = 1$€ and on the other hand, our pricing rule gives

$$C_0(\Psi_T, T) = \mathbb{E}_{\mathcal{P}}\left(L_T e^{-\int_0^T r_s \mathrm{d}s}\Psi_T\right) = \mathbb{E}_{\mathcal{P}}\left(L_T\right) \text{€}\ .$$

Equating both gives

$$\mathbb{E}_{\mathcal{P}}\left(L_t\right) = 1, \forall t \geq 0.$$

Later, when dealing with a filtered probability space, we will prove that $(L_t)_{t\geq 0}$ must be a non-negative martingale, satisfying $\mathbb{E}_{\mathcal{P}}\left(L_T \mid \mathcal{F}_t\right) = L_t, \forall 0 \leq t \leq T$, and starting from $L_0 = 1$.

2.1.4 L_T defines a new probability measure $\mathbb{Q}_T$

For any measurable set $A \in \mathcal{F}$, set

$$\mathbb{Q}_T(A) := \mathbb{E}_{\mathcal{P}}\left(L_T \mathbf{1}_A\right).$$

Observe that this $\mathbb{Q}_T$ satisfies to σ-additivity properties, $\mathbb{Q}_T(\emptyset) = 0$, to non-negativity (since $L_T \geq 0$) and $\mathbb{Q}_T(\Omega) = \mathbb{E}_{\mathcal{P}}\left(L_T\right) = 1$: hence $\mathbb{Q}_T$ is a new probability measure. The rules for measure change yields that (2.1.1) becomes

$$C_0(\Psi_T, T) = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-\int_0^T r_s \mathrm{d}s}\Psi_T\right),$$

which writes as the expected discounted cashflow under the new measure $\mathbb{Q}_T$.

2.1.5 Return of a tradable asset under $\mathbb{Q}_T$

Assume that $(S_t)_{t\leq T}$ is the price process of a tradable asset that does not pay dividends. Then, buying the asset and holding it till T costs S_0 initially and worths S_T at T. Therefore, by the rule of unique price,

$$S_0 := C_0(S_T, T) = \mathbb{E}_{\mathbb{Q}_T}\left(e^{-\int_0^T r_s \mathrm{d}s} S_T\right).$$

Hence, $\mathbb{E}_{\mathbb{Q}_T}\left(e^{-\int_0^T r_s \mathrm{d}s} S_T/S_0\right) = 1$, which shows that S_T has the same return than the no-risk asset (cash account). For this reason, one may call $\mathbb{Q}_T$ a risk-neutral probability measure, since under this measure, the asset behaves <u>as if</u> its return is $(r_t)_{t\geq 0}$. The pricing measure does not pay for the risk of S, as if it were risk-indifferent or risk-neutral.

2.1.6 What about the dynamic aspects?

So far, we have clarified the form and the properties for a reasonable pricing rule $C_0(\Psi_T, T)$ at time $t = 0$. It is interesting to extend these arguments to $C_t(\Psi_T, T)$. It would require to deal with filtrations. Let us give the main results without details: assume that on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{P})$ the pricing rule is

$$C_t(\Psi_T, T) := \mathbb{E}_{\mathcal{P}}\left(L_T e^{-\int_t^T r_s \mathrm{d}s} \Psi_T \mid \mathcal{F}_t\right),$$

which is consistent with (2.1.1). Then

- $C_t(\Psi_T, T) := \mathbb{E}_{\mathbb{Q}_T}\left(e^{-\int_t^T r_s \mathrm{d}s} \Psi_T | \mathcal{F}_t\right)$, with the same probability measure $\mathbb{Q}_T$ as before,
- a discounted tradable asset $(e^{-\int_t^T r_s \mathrm{d}s} S_t)_{0 \leq t \leq T}$ without dividend should be martingale under $\mathbb{Q}_T$.

2.1.7 Price axiom

As a result of the previous discussion, we get a reasonable candidate for the cash price (and thus for the forward price) of a cashflow Ψ_T:

$$\begin{aligned} C_0(\Psi_T, T) &:= E_{\mathbb{Q}}\left[e^{-\int_0^T r_s \mathrm{d}s} \Psi_T\right] \text{(the cash price)}, \\ F_0(\Psi_T, T) &:= E_{\mathbb{Q}}\left[\frac{e^{-\int_0^T r_s \mathrm{d}s} \Psi_T}{B(0,T)}\right] \text{(the forward price)}. \end{aligned} \tag{2.1.2}$$

The probability measure is denoted here $\mathbb{Q}$ to simplify (this is our previous $\mathbb{Q}_T$) and is such that tradable asset without dividends have a return coinciding with the interest rate. In the rest of this section, we take this rule as granted. We will connect this rule to hedging portfolio later in Chapter III.

2.1.8 What probabilistic information can be found from market prices?

Imagine that we seek a probabilistic model consistent with market data, i.e.

$$C_0^{\texttt{Market}}((S_T - K)_+, T) = C_0^{\texttt{Model}}((S_T - K)_+, T).$$

Since the latter equals $\mathbb{E}_{\mathbb{Q}}\left(e^{-\int_0^T r_s \mathrm{d}s}(S_T - K)_+\right)$ using our axioms, we retrieve some information about the distribution of S_T under $\mathbb{Q}$.

Proposition 2.1 (Breeden-Litzenberger formula)**.** *The left and right derivatives of market call prices w.r.t. K are given as follows:*

$$\begin{aligned} \partial_K^- C_0^{\texttt{Market}}((S_T - K)_+, T) &= -\mathbb{E}_{\mathbb{Q}}\left(e^{-\int_0^T r_s \mathrm{d}s} \mathbf{1}_{S_T \geq K}\right), \\ \partial_K^+ C_0^{\texttt{Market}}((S_T - K)_+, T) &= -\mathbb{E}_{\mathbb{Q}}\left(e^{-\int_0^T r_s \mathrm{d}s} \mathbf{1}_{S_T > K}\right). \end{aligned}$$

Proof. We start from

$$C_0^{\texttt{Market}}((S_T-K)_+,T)=\mathbb{E}_{\mathbb{Q}}\left(e^{-\int_0^T r_s\mathrm{d}s}(S_T-K)_+\right).$$

We apply Lebesgue's derivation theorem to $(S_T-K)_+$, which derivative w.r.t. K is $-\mathbf{1}_{S_T\geq K}$ and $-\mathbf{1}_{S_T>K}$ according to the side of the derivative. In any case, this derivative times $e^{-\int_0^T r_s\mathrm{d}s}$ is bounded (uniformly in K) by $e^{-\int_0^T r_s\mathrm{d}s}$, and thus integrable. We are done. $\square$

This result shows that one has access to the survival distribution function of S_T under $\mathbb{Q}$ (up to the discounted factor). Of course, this holds in an ideal situation where market datas are available on a fine grid of strikes, in order to compute derivatives.

Another way to get this result is to apply the Carr formula (seen in Week 1 Lectures) to the smooth function $x\mapsto e^{-\lambda x}$ for any $\lambda>0$, which makes the link with the Laplace transform of the distribution of S_T under $\mathbb{Q}_T$. Since the Laplace transform characterizes positive random variables, we retrieve qualitatively the same result as before. The Carr formula states that Call/Put prices form a generating system for the price of smooth vanilla payoff, and in particular the Laplace transform of S_T under $\mathbb{Q}_T$.

Proposition 2.2 (Application of Carr formula)**.** *For any $\lambda>0$, we get (setting $f_0=\texttt{F}_0(S_T,T)$ for the forward price of S_T)*

$$\mathbb{E}\left(e^{-\int_0^T r_s\mathrm{d}s}e^{-\lambda S_T}\right)=e^{-\lambda f_0}B(0,T)+\int_{f_0}^{+\infty}\lambda^2e^{-\lambda K}\texttt{Call}_0^{\texttt{Market}}(T,K)\mathrm{d}K$$
$$+\int_0^{f_0}\lambda^2e^{-\lambda K}\texttt{Put}_0^{\texttt{Market}}(T,K)\mathrm{d}K.$$

Proof. The Carr formula writes for any h

$$\texttt{C}_0^{\texttt{Market}}(h(S_T,T))=h(f_0)B(0,T)+\int_{f_0}^{+\infty}h''(K)\texttt{Call}_0^{\texttt{Market}}(T,K)\mathrm{d}K+\int_0^{f_0}h''(K)\texttt{Put}_0^{\texttt{Market}}(T,K)\mathrm{d}K.$$

This holds also for `Model` price. In the particular case $h(x)=e^{-\lambda x}$ we get the announced formula. $\square$

2.1.9 Partial conclusion

The observation of the Call/Put market prices $\texttt{Call}_0^{\texttt{Market}}(T,K),\texttt{Put}_0^{\texttt{Market}}(T,K)$ gives a (almost) complete information on the distribution S_T seen from $t=0$.

Since T can be made arbitrary (at least in the set of market-available maturities), one has access to the marginal distribution of S (that is, t by t). However,

- it does not characterize at all the full distribution of the process S, described by the distribution of any vectors of coordinates $(S_{t_1},S_{t_2}...,S_T)$.

- this information is the one available at time $t = 0$ and of course, we do not have access to the dynamics of this distribution (how will it evolve from time $t = 0$ to t=1 day, and so on?).

Calibration (finding a model coherent with market data) is therefore a very difficult exercise: the model fit can be excellent today for call/put options, but completely wrong with the same model tomorrow. One needs to recalibrate everyday, which may be a non-sense (a model is a vision of risk over the next days/weeks, at least over $[0, T]$). A good calibration model is a model for which the calibrated parameters are stable over time.

2.2 Implied volatility

In the Black-Scholes formula for call/put options, prices depend mostly on the unknown parameter σ, while the other ones (S_0, T, K, r, q) are either defined in the contract or observed on the market. How to get the value of σ, which is not directly observable?

A statistical point of view. If the market data follows a geometric Brownian motion model, then the returns between observation times t_{k-1} and t_k must be

$$Y_k := \ln\left(\frac{S_{t_k}}{S_{t_{k-1}}}\right) = (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma \Delta W_{t_k}$$

where $\Delta t = t_k - t_{k-1}$ and $\Delta W_{t_k} = W_{t_k} - W_{t_{k-1}}$. Under the assumptions that the GBM assumption is satisfied, the above returns behave like independent Gaussian random variables $\mathcal{N}((\mu - \frac{1}{2}\sigma^2)\Delta t, \sigma^2 \Delta t) \stackrel{d}{=} \sqrt{\Delta t}\, \mathcal{N}((\mu - \frac{1}{2}\sigma^2)\sqrt{\Delta t}, \sigma^2)$. The mean term becomes negligible as $\Delta t \to 0$, thus the empirical variance of $Y_k/\sqrt{\Delta t}$ gives a consistent estimate of σ^2: a rigorous statement/proof could be written, we skip it. Anyway, all this requires the validity of the GBM over a long enough historical period, with constant parameters: these conditions are difficultly met in practice. Indeed, constancy of parameters are not observed in the market, at least not at the horizon length needed to ensure a reasonable estimate of σ^2. Besides, there is a conceptual problem about looking far in the past to estimate risk in the future (would you drive a fast-car while looking at a rear-view mirror?). The following point of view takes advantage of current available information, and therefore, it is seemingly more reliable.

Implied volatility. Since calls/puts are exchanged on the market, one may wonder how to extract information from these to get σ: this is an inverse problem, i.e. finding σ such that the BS prices (using that σ) coincide with the market prices. If the model follows a geometric Brownian motion, this inversion will be easily feasible and in addition, only a single market price will be enough to get the single parameter σ.

We start with a single call option with parameters (T, K), which market price at zero is denoted by $\mathtt{Call}_0^{\mathtt{Mkt}}(T, K)$.

Theorem 2.3 (definition and existence)**.** *The equation*

$$\texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma) = \texttt{Call}_0^{\texttt{Mkt}}(T, K)$$

has a unique solution in $\sigma \in (0, +\infty)$. The solution is called the Implied Volatility denoted by σ_I: it depends a priori on all other parameters (S_0, T, K).

Proof. We have :

$$\begin{aligned}
\texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma) &\xrightarrow{\sigma \to 0} \max\left(S_0 e^{-qT} - K e^{-rT}, 0\right), \\
\texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma) &\xrightarrow{\sigma \to \infty} S_0.
\end{aligned}$$

Besides, these limits squeeze the market price $\texttt{Call}_0^{\texttt{Mkt}}(T, K)$, in view of the no-arbitrage bounds on call price. Last, $\sigma \in (0, +\infty) \mapsto \texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma)$ is continuous and strictly increasing (since the Vega $\frac{\partial \texttt{Call}^{\texttt{BS}}}{\partial \sigma}(0, S_0, T, K, \sigma) = S_0 e^{-qT} \mathcal{N}(d_+) \sqrt{T - t}$ is strictly positive). Then, by the intermediate value theorem,

$$\texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma_I) = \texttt{Call}_0^{\texttt{Mkt}}(T, K)$$

for some real number σ_I. □

Proposition 2.4. *If σ_I is the volatility associated to the call option parameters (T, K), then one has also*

$$\texttt{Put}^{\texttt{BS}}(0, S_0, T, K, \sigma_I) = \texttt{Put}_0^{\texttt{Mkt}}(T, K).$$

In other words, the implied volatility is the same for call and put options. The above equality stems from the call/put relation which writes

$$\texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma) - \texttt{Put}^{\texttt{BS}}(0, S_0, T, K, \sigma) = \texttt{Call}_0^{\texttt{Mkt}}(T, K) - \texttt{Put}_0^{\texttt{Mkt}}(T, K), \qquad \forall \sigma.$$

Using Theorem 2.3, a single observation of market price is enough to get the unknown σ. If we take another market price (with other (T, K)), one should have the same value if the market follows a geometric Brownian motion. Unfortunately, computing $\sigma_I(T, K)$ for different strikes K gives a non-constant curve, that has a shape of

1. smile on FX market,
2. negative skew on equities/indices,
3. positive skew on commodities.

We usually use the term volatility smile for all these curves.

In addition, when varying T, the volatility smile has a term-structure and the surface $(T, K) \mapsto \sigma_I(T, K)$ is the volatility surface. For a clear exposure about volatility surface, see [Gat06].

Computation σ_I using Bisection Method. Computing σ_I can not be done using an explicit formula, it has to be done numerically.

The solution of the non-linear implied volatility equation can be calculated using the bisection algorithm for the function $\sigma \in (0, +\infty) \to \texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma)$. Actually, having an infinite interval $(0, +\infty)$ for the bisection is not convenient, it is better to use a closed interval. For this, observe that $x \in (0, 1) \mapsto \frac{x}{1-x} \in (0, +\infty)$ is a increasing bijection, hence it is equivalent to bisect on σ or x.

Then, the bisection method writes as follows. Set $x_0 = 0$, $y_0 = 1$ and $z_1 = \frac{x_0+y_0}{2}$.

- if $\texttt{Call}_0^{\texttt{Mkt}}(T, K) > \texttt{Call}^{\texttt{BS}}\left(0, S_0, T, K, \frac{z_i}{1-z_i}\right)$ then we set $x_i = z_i$, $y_i = y_{i-1}$, $z_{i+1} = \frac{x_i+y_i}{2}$.
- if $\texttt{Call}_0^{\texttt{Mkt}}(T, K) < \texttt{Call}^{\texttt{BS}}\left(0, S_0, T, K, \frac{z_i}{1-z_i}\right)$ then we set $x_i = x_{i-1}$, $y_i = z_i$, $z_{i+1} = \frac{x_i+y_i}{2}$ and we iterate...
- if equality, we stop.

At the end (when the variation of z_i to z_{i+1} is too small), put $\sigma_I = \frac{z_n}{1-z_n}$.
The choice of the number of iterations depends on the desired accuracy. The convergence in z_n is geometric.

Computation using Newton-Raphson Algorithm. The convergence using the Newton Algorithm is even faster. The principle of this algorithm is:

$$\sigma_{n+1} = \sigma_n - \frac{\texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma_n) - \texttt{Call}_0^{\texttt{Mkt}}(T, K)}{\frac{\partial \texttt{Call}^{\texttt{BS}}}{\partial \sigma}(0, S_0, T, K, \sigma_n)}.$$

It is known that this algorithm converges very quickly, if the starting point σ_0 is close to the required value. The choice of σ_0 is in general difficult, but here, a special value works systematically well, this has been observed by [MK82]. This is based in the following property.

Proposition 2.5. *Set*

$$\sigma_0 := \sqrt{\frac{2}{T}\left|\ln\left(\frac{Se^{(r-q)T}}{K}\right)\right|}.$$

For $\sigma > \sigma_0$, $\sigma \mapsto \texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma)$ is concave, and for $\sigma < \sigma_0$, it is convex.

It is easy to see (plot the iteration) that this situation is very benefitial for the Newton-Raphson Algorithm and ensures its convergence in a few iterations (the convergence is bi-quadratic, each time the number of accurate digits double).

Proof of Proposition 2.5. The Vega is $S_0 e^{-qT}\sqrt{T}\mathcal{N}'(d_+)$. Then, the vomma is

$$\frac{\partial^2 \texttt{Call}^{\texttt{BS}}}{\partial \sigma^2} = \frac{\partial\left(S_0 e^{-qT}\sqrt{T}\mathcal{N}'(d_+)\right)}{\partial \sigma}$$

$$
\begin{aligned}
&= S_0 e^{-qT}\sqrt{T}\left(\frac{-1}{\sigma^2\sqrt{T}}\ln\left(\frac{S_0e^{(r-q)T}}{K}\right)+\frac{\sqrt{T}}{2}\right)\mathcal{N}''(d_+)\\
&= S_0 e^{-qT}\sqrt{T}\frac{(-d_-)}{\sigma}(-d_+\mathcal{N}'(d_+))\\
&= \text{Vega}\frac{d_-d_+}{\sigma}.
\end{aligned}
$$

It comes

$$
\begin{aligned}
\frac{\partial^2 \texttt{Call}^{\texttt{BS}}}{\partial\sigma^2}>0 &\iff d_-d_+>0\\
&\iff \frac{1}{\sigma^2 T}\ln\left(\frac{S_0e^{(r-q)T}}{K}\right)^2>\frac{\sigma^2 T}{4}\\
&\iff \sigma<\sqrt{\frac{2}{T}\left|\ln\left(\frac{S_0e^{(r-q)T}}{K}\right)\right|}.
\end{aligned}
$$

□

2.3 Displaced Lognormal model

This model, introduced by [Rub83], is a variant of the Black-Scholes model, with an additional parameter a. It keeps the benefit of explicit call/put price formula. It leads to a non-constant implied volatility, which is interesting on its own.

2.3.1 Definition and first properties

To simplify the setting, we suppose zero interest rate $r = 0$.

Definition 2.6. *The dynamic of the displaced lognormal model (DLN) is given by:*

$$
\mathrm{d}S_t = (S_t + a)\sigma \mathrm{d}W_t, \tag{2.3.1}
$$

where $a \in \mathbb{R}$, $\sigma > 0$ and $S_0 + a > 0$. The asset price is explicitly given by

$$
S_t = (S_0 + a)\exp(\sigma W_t - \frac{1}{2}\sigma^2 t) - a, \qquad t \geq 0. \tag{2.3.2}
$$

To go from (2.3.2) to (2.3.1), apply the Ito equation to write

$$
\mathrm{d}S_t = (S_0+a)\mathrm{d}\left(\exp(\sigma W_t - \frac{1}{2}\sigma^2 t)\right) = (S_0+a)\exp(\sigma W_t - \frac{1}{2}\sigma^2 t)\sigma\mathrm{d}W_t = (S_t+a)\sigma\mathrm{d}W_t.
$$

Conversely, from (2.3.1), obverse that S_t+a satisfies to a geometric Brownian motion dynamics, which gives (2.3.2).

The dynamics of S is written without drift (it is equal to the zero interest rate): it means that one states here the dynamics under a risk-neutral probability (that does not change the shape of the noise terme, i.e. the $\mathrm{d}W_t$ term related to volatility, as we have seen in Black-Scholes and binomial tree modeling). The case $a = 0$ gets back to the usual Black-Scholes model.

Note that $S+a$ is a driftless geometric Brownian motion with volatility σ, and the interval of points attainable by S is $(-a, \infty)$. If modeling a nonnegative underlying asset such as a stock price, this model for $a > 0$ will misprice deep-out-of-the-money puts ($K \ll S_0$), due to the possibility of $S_T < 0$. For other strikes, this model still makes sense.

We list simple properties that follow from the definition.

Properties 2.7.

a) $(S_T + a)$ *is lognormal,*

b) $\mathbb{E}\,(S_T) = S_0$,

c) the instantaneous volatility defined by the process σ_t *such that* $\frac{\mathrm{d}S_t}{S_t} = \sigma_t \mathrm{d}W_t$ *is given by*

$$\sigma_t = \frac{S_t + a}{S_t}\sigma.$$

The instantaneous volatility (also called spot volatility) gives a simple intuition how much the asset price S changes as the market noise W changes

$$\text{relative return of } S = \frac{\Delta S_t}{S_t} \approx \sigma_t \Delta W_t.$$

Note that if $a > 0$, $\sigma_t \uparrow +\infty$, when $S_t \downarrow 0$ i.e. when the stock price decreases to 0, the volatility increases. This is somewhat intuitive with what is observed when the market drops.

2.3.2 Call price

Theorem 2.8. *For $K + a > 0$, a K-strike T-expiry European call option on a displaced lognormal S has a price explicitly given by*

$$\mathtt{Call}_0^{\mathtt{DLN}}((S_T - K)_+, T) = \mathtt{Call}^{\mathtt{BS}}(0, S_T + a, T, K + a).$$

Proof. According to the axiom (2.1.2), the call price in such a model writes

$$\mathbb{E}_{\mathbb{Q}}\left((S_T - K)_+\right) = \mathbb{E}_{\mathbb{Q}}\left([(S_0 + a)e^{\sigma W_T - \frac{1}{2}\sigma_T^2} - (K + a)]_+\right)$$

and we are back to computing a Black-Scholes formula. □

In general, payoffs invariant to parallel shifts of the S path and the contract parameters can be priced using Black-Scholes model valuation methods, but applied to displaced arguments. This applies to barrier options[1], for instance, which payoff writes

$$(S_T - K)_+ \mathbf{1}_{\min_{t\in[0,T]} S_t \le D} = (S_T + a - (K + a))_+ \mathbf{1}_{\min_{t\in[0,T]} (S_t + a) \le D + a}.$$

[1] which pricing methodolgy will be studied in next chapter

2.3.3 Implied volatility

Although the DLN call price is explicit, its implied volatility is not. However, it is possible to get its short-expiry behavior At The Money (ATM). We will derive the well-known $\frac{1}{2}$-rule of thumb, stating that

- the ATM implied volatility at short-expiry coincides with the spot volatility,
- the slope of ATM implied volatility at short-expiry coincides with $\frac{1}{2}$ of the slope of the spot volatility.

This rule of thumb is valid, not only in the DLN model as we will prove it, but also in any local volatility model [BBF02] (i.e. when the spot volatility is of the form $\sigma_t = \sigma(t, S_t)$).

2.3.4 Short-expiry behavior of implied volatility

The following theorem takes the short-expiry $T \downarrow 0$ limit of the implied volatility skew, and expresses the solution explicitly. This is done when $K = S_0$ (ATM).

Theorem 2.9. *The implied volatility $\sigma_I(T, S_0)$ of the Displaced Lognormal model with short maturity and at the money $K = S_0$ is such that:*

i) $\lim_{T\to 0} \sigma_I(T, S_0) = \sigma_0 = \frac{S_0+a}{S_0}\sigma$,

ii) $\lim_{T\to 0} \partial_K \sigma_I(T, S_0) = -\frac{a\sigma}{2S_0^2} = \frac{1}{2}\frac{\partial\sigma_0}{\partial S_0}$.

Proof. By definition of implied volatility $\sigma_I(T, K)$,

$$\mathtt{Call}_0^{\mathtt{DLN}}((S_T - K)_+, T) = \mathtt{Call}^{\mathtt{BS}}(0, S_0, T, K, \sigma_I(T, K))$$

and besides, by the DLN call price formula, recall $\mathtt{Call}_0^{\mathtt{DLN}}((S_T - K)_+, T) = \mathtt{Call}^{\mathtt{BS}}(0, S_0 + a, T, K + a)$. All in all,

$$\mathtt{Call}^{\mathtt{BS}}(0, S_0 + a, T, K + a) = \mathtt{Call}^{\mathtt{BS}}(0, S_0, T, K, \sigma_I(T, K)). \tag{2.3.3}$$

For any $T > 0$ and $\sigma > 0$, the left-hand side is C^∞ w.r.t. $(T, K) \in (0 + \infty)^2$, the function $(T, K, \sigma) \mapsto \mathtt{Call}^{\mathtt{BS}}(0, S_0, T, K, \sigma)$ is C^∞ on $(0 + \infty)^3$ and in addition the Vega $(= \partial_\sigma \mathtt{Call}^{\mathtt{BS}}(0, S_0, T, K, \sigma))$ is strictly positive. Therefore, by the implicit function theorem, $(T, K) \mapsto \sigma_I(T, K)$ is C^∞ on $(0 + \infty)^2$.

Moreover, observe $K = S_0 \implies \log(\frac{S_0}{K}) = 0$. This remark simplifies much the coefficients $d_\pm$ in the Black-Scholes formula. We get, from (2.3.3) and $K = S_0$,

$$(S_0+a)\mathcal{N}(\frac{1}{2}\sigma\sqrt{T}) - (S_0+a)\mathcal{N}(-\frac{1}{2}\sigma\sqrt{T}) = S_0\mathcal{N}\left(\frac{1}{2}\sigma_I(T, S_0)\sqrt{T}\right) - S_0\mathcal{N}\left(-\frac{1}{2}\sigma_I(T, S_0)\sqrt{T}\right).$$

As $T \to 0$, the left-hand side converges to 0. Hence, necessarily, $\sigma_I(T, S_0)\sqrt{T} \to 0$. Hence, a Taylor expansion at the third order (like in the Brenner-Subrahmanyan formula) gives

$$\begin{aligned}
&(S_0 + a)\left(\frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma\sqrt{T} - \frac{1}{2} + N'(0)\frac{1}{2}\sigma\sqrt{T} + \mathcal{O}(\sigma\sqrt{T})^3\right) \\
&= S_0\left(\frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma_I(T, S_0)\sqrt{T} - \frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma_I(T, S_0)\sqrt{T} + \mathcal{O}(\sigma_I(T, S_0)\sqrt{T})^3\right).
\end{aligned}$$

After simplifications, we get

$$(S_0 + a)\sigma\sqrt{T} + \mathcal{O}((\sigma\sqrt{T})^3) = S_0\sigma_I(T, S_0)\sqrt{T} + \mathcal{O}(\sigma_I(T, S_0)\sqrt{T})^3. \tag{2.3.4}$$

Looking at the main expansion term, this shows that the orders of magnitude of $\sigma_I(T, S_0)\sqrt{T}$ and $\sigma\sqrt{T}$ are the same, and thus

$$\mathcal{O}(\sigma_I(T, S_0)\sqrt{T})^3 = \mathcal{O}(\sigma\sqrt{T})^3 = \mathcal{O}(T^{3/2}),$$

the latter being valid as $T \to 0$. Hence, one can simplify the equality (2.3.4) and get

$$\sigma_I(T, S_0) = \frac{S_0 + a}{S_0}\sigma + \mathcal{O}(T).$$

We have proved the first result (i). To establish (ii), differentiate (2.3.3) w.r.t. K and take $K = S_0$: it gives

$$\begin{aligned}
\left.\frac{\partial(2.3.3)}{\partial K}\right|_{K=S_0} &= \frac{\partial \texttt{Call}}{\partial K}(0, S_0 + a, T, S_0 + a, \sigma) \\
&= \texttt{Vega}(0, S_0, T, S_0, \sigma_I(T, S_0))\frac{\partial \sigma_I}{\partial K}(T, S_0) + \frac{\partial \texttt{Call}}{\partial K}(0, S_0, T, S_0, \sigma_I(T, S_0)).
\end{aligned}$$

Invoke the explicit expression of Black-Scholes greeks, use the asymptotics $T \to 0$ with $K = S_0$:

$$\begin{aligned}
\frac{\partial \texttt{Call}}{\partial K}(0, S_0 + a, T, S_0 + a, \sigma) &= -\mathcal{N}(-\frac{1}{2}\sigma\sqrt{T}) = -\frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma\sqrt{T} + \mathcal{O}(T), \\
\frac{\partial \texttt{Call}}{\partial K}(0, S_0, T, S_0, \sigma_I(T, S_0)) &= -\frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma_I(T, S_0)\sqrt{T} + \mathcal{O}(T), \\
\texttt{Vega}(0, S_0, T, S_0, \sigma_I(T, S_0)) &= S_0\mathcal{N}'(d_+(0, S_0, T, S_0, \sigma_I(T, S_0)))\sqrt{T} = S_0\mathcal{N}'(0)\sqrt{T} + \mathcal{O}(T).
\end{aligned}$$

It readily follows that

$$\begin{aligned}
-\frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma\sqrt{T} &= S_0\mathcal{N}'(0)\sqrt{T}\frac{\partial \sigma_I}{\partial K}(T, S_0) - \frac{1}{2} + \mathcal{N}'(0)\frac{1}{2}\sigma_I(T, S_0)\sqrt{T} + \mathcal{O}(T), \\
S_0\frac{\partial \sigma_I}{\partial K}(T, S_0) &= \frac{1}{2}\sigma - \frac{1}{2}\underbrace{\sigma_I(T, S_0)}_{=\frac{S_0+a}{S_0}\sigma + \mathcal{O}(T)} + \mathcal{O}(\sqrt{T}) = -\frac{1}{2}\frac{a\sigma}{S_0} + \mathcal{O}(\sqrt{T}).
\end{aligned}$$

We get the announced formula. □

Consequences.

1. If $a > 0$, $\frac{\partial \sigma_I}{\partial K}(T, S_0) < 0$, i.e the ATM skew (of the implied volatility) is negative. This is coherent with what is observed on equity/index option price. Hence, for such markets, a calibrated DLN model would have $a > 0$.

2. Theorem 2.9 gives a simple way to calibrate a DLN model: compute on market data the level and slope of implied volatility ATM, which has to be fitted with DLN formulas. We have two equations and two unknowns (a, σ), this is feasible.

3. However, it turns out that the DLN model does exhibit a really linear smile shape, which is not much compatible with market data. Also the term structure (dependence w.r.t. T) of implied volatility of such DLN is very poor, which is again not much compatible with market data. This DLN model should be seen as a first step in volatility modeling, but it has to be improved.

2.4 Bachelier Model

2.4.1 Definition and first properties

Here again, for pedagogical reasons, we consider a simplified market, with zero interest rate $(r = 0)$ and a risky asset S_t with no dividends $(q = 0)$. The Bachelier model is a model in which the asset price is directly given by a Brownian motion

$$S_t = S_0 + \sigma W_t, \qquad t \geq 0,$$

where $S_0 > 0$, $\sigma > 0$. This model dates back to [Bac00].

Remark 2.10.

1. *S_t is a Gaussian variable, it can be negative with a non zero probability.*

2. *The difference between this model and the Black-Scholes model is somewhat analogous to the difference between linear and compound interests. The two models give similar results in the short run, as we'll see later (Theorem 2.12), but not in the long run.*

3. *This model although basic is sometimes used for deriving approximations in more sophisticated models. For instance, the stochastic volatitlity model called SABR [HKLW02] has an explicit (approximate) formula which is obtained through the use of Bachelier model as a zero-order proxy. See also the discussions in [BG12].*

2.4.2 Call price

Theorem 2.11. *Assume that the price of a Call Option on S with strike K and maturity T in the Bachelier model is computed as $\mathbb{E}_{\mathbb{Q}}\left((S_T - K)_+\right)$. Then the Call price is given by*

$$\mathtt{Call}_0^{\mathtt{Bach.}} := (S_0 - K)\mathcal{N}\left(\frac{S_0 - K}{\sigma\sqrt{T}}\right) + \sigma\sqrt{T}\mathcal{N}'\left(\frac{S_0 - K}{\sigma\sqrt{T}}\right).$$

Proof. Left as an exercise. □

In the following result, we compare the price of a Call Option in the Bachelier model with parameter σ^B and in the Black-Scholes model with parameter σ, provided that parameters are linked to each other.

Theorem 2.12. *Let $\sigma^B = \sigma^{BS} S_0$. Then, for an At The Money ($S_0 = K$) Call Option, the price of the Call Option in the Black-Scholes model ($\mathtt{Call}_0^{\mathtt{BS}}$) and the price in the Bachelier model ($\mathtt{Call}_0^{\mathtt{Bach.}}$) satisfy*

$$0 \leq \frac{\mathtt{Call}_0^{\mathtt{Bach.}} - \mathtt{Call}_0^{\mathtt{BS}}}{\mathtt{Call}_0^{\mathtt{Bach.}}} \leq \frac{T(\sigma^{BS})^2}{24}.$$

Proof. Left as an exercise. □

We can go a bit further. It is possible to prove [ST08] that, given the option price C_0 for a Call Option at the money, the difference between the Black-Scholes implied volatilities ($\sigma_I^{BS} := \sigma_I^{BS}(C_0)$) and the Bachelier implied volatility ($\sigma_I^{Bach} := \sigma_I^{Bach}(C_0)$) model satisfy

$$0 \leq \sigma_I^{BS} - \frac{\sigma_I^{Bach}}{S_0} \leq \frac{T(\sigma_I^{BS})^3}{24}.$$

Therefore, it is possible to use the simplified formula $\sigma_I^{BS} \approx \frac{\sigma_I^{Bach}}{S_0} = \frac{\sqrt{2\pi}}{\sqrt{T}}\frac{C_0}{S_0}$ for a Call Option at the money when T and σ_I^{BS} are small (Brenner-Subrahmanyan formula).

2.5 Merton model

This model [Mer76] is a variant of Black-Scholes-Samuelson model, where in addition to Geometric Brownian behaviour for the paths, we add some random jumps, random with respect to the times they occur and random w.r.t. their amplitudes. The jump times are given by those of a Poisson process, the jump sizes are log-normally distributed. All these random objects being independent. The jumps are aimed at modeling unpredictable shocks in the market, and their frequency is not large (a few per year). Interestingly, on the option pricing side,

- computing the call/put price can be made explicitly,
- this model is able to exhibit a smile for implied volatility surfaces, and thus this is a good candidate for model calibration on true market data.

We now enter into the details.

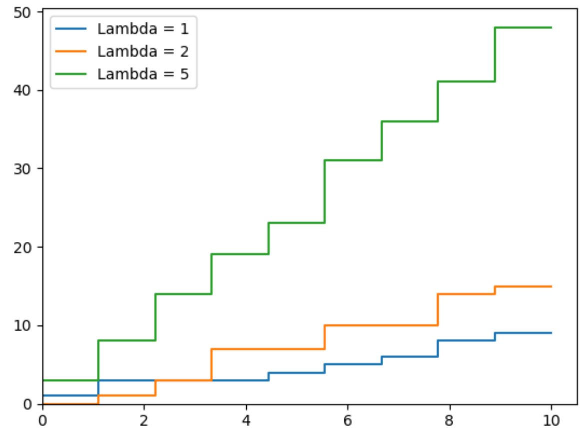

Figure 1: Sample path of Poisson processes with different intensities

2.5.1 Poisson process

Definition 2.13. *Given a sequence of independent random variables $(T_i : i \geq 1)$ following an exponential distribution with parameter $\lambda > 0$, we define the n-th jump times*

$$\tau_n := \sum_{i=1}^{n} T_i.$$

The Poisson process with intensity λ denoted by $N = (N_t : t \geq 0)$ is the counting process defined by :

$$N_t = \sum_{n=1}^{+\infty} \mathbf{1}_{\tau_n \leq t} = \#\{n : \tau_n \leq t\}.$$

For a reference on Poisson processes, see [Kin93]. We list some known properties.

Proposition 2.14. *If N is a Poisson process with intensity $\lambda > 0$, then the random variable N_t follows a Poisson distribution with parameter λt:*

$$\mathbb{P}\left[N_t = n\right] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n \in \mathbb{N}.$$

Proposition 2.15. *Given a Poisson process N with intensity λ, define its natural filtration $\mathcal{F}_t = \sigma(N_s, s \leq t)$. The process N has independent and stationary increments:*

- *for $s, t > 0, N_{t+s} - N_s$ is independent from $\mathcal{F}_s$,*
- *the distribution of $N_{t+s} - N_s$ is the same as $N_t - N_0 = N_t$.*

Compound Poisson process. N is a counting process, that counts one at each time a new jump time appears. A compound Poisson process J jumps at the same times, but with a random amplitude: in other words, a compound Poisson process with unit jump is a standard Poisson process. The process J can be defined by

$$J_t = \sum_{k=1}^{N_t} Y_k, \qquad t \geq 0,$$

where $(Y_k : k \geq 1)$ is a i.i.d. sequence of scalar random variables, modelling jump sizes.

We have the easy relations:

$$\Delta J_t := J_t - J_{t-} = Y_{N_t} \Delta N_t, \text{ where } \Delta N_t = N_t - N_{t-} \in \{0, 1\}.$$

2.5.2 Derivation of the Merton model

We define on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a Brownian motion W, a Poisson process N with intensity $\lambda > 0$, i.i.d. random variable $(J_k : k \geq 1)$ for jumps, all being independent. This serves to model the time evolution of an asset price S (typically a stock, or an index). The evolution of S_t at time t can now be described as follows:

- Between $[\tau_j, \tau_{j+1}[$, S follows a continuous Geometric Brownian motion dynamic, i.e. $\mathrm{d}S_t = S_t(\mu \mathrm{d}t + \sigma \mathrm{d}W_t)$,
- At time τ_j, S jumps and its jump is given by $\Delta S_{\tau_j} = S_{\tau_j} - S_{\tau_j^-} = S_{\tau_j^-} Y_j$ so that $S_{\tau_j} = S_{\tau_j^-}(1 + Y_j)$. Doing so, the process S has right-continuous path.

The formula for S can be derived by induction on each interval:

1. for $t \in [0, \tau_1[$,
$$S_t = S_0 e^{(\mu - \sigma^2/2)t + \sigma W_t},$$
2. the left limit at τ_1 is given by :
$$S_{\tau_1^-} = S_0 e^{(\mu - \sigma^2/2)\tau_1 + \sigma W_{\tau_1}},$$
3. at the first jumps,
$$S_{\tau_1} = S_{\tau_1^-}(1 + Y_1) = S_0(1 + Y_1) e^{(\mu - \sigma^2/2)\tau_1 + \sigma W_{\tau_1}},$$
4. for $t \in [\tau_1, \tau_2[$,
$$S_t = S_{\tau_1} e^{(\mu - \sigma^2/2)(t - \tau_1) + \sigma(W_t - W_{\tau_1})} = S_0(1 + Y_1) e^{(\mu - \sigma^2/2)t + \sigma W_t}$$
5. and so on.

Iterating the arguments, we obtain:

$$S_t = S_0 e^{(\mu-\sigma^2/2)t+\sigma W_t} \prod_{j=1}^{N_t}(1+Y_j) \qquad \text{(with the convention } \prod_{j=1}^{0} \cdots = 1).$$

Note that taking $\lambda = 0$ ($N \equiv 0$) gets back to the usual Geometric Brownian motion model.

In the Merton model, the jump size Y_i is distributed as a log-normal distribution:

$$\log(1+Y_i) \stackrel{d}{=} \mathcal{N}(\log(1+m) - \frac{\alpha^2}{2}, \alpha^2)$$

with $m > -1$ and $\alpha \geq 0$. Putting the parameters under this specific form is made for easier interpretation:

$$\mathbb{E}\,(1+Y_i) = \mathbb{E}\left(e^{\mathcal{N}(\log(1+m)-\frac{\alpha^2}{2},\alpha^2)}\right) = e^{\log(1+m)-\frac{\alpha^2}{2}}\,\mathbb{E}\left(e^{\mathcal{N}(0,\alpha^2)}\right) = e^{\log(1+m)-\frac{\alpha^2}{2}} e^{\frac{\alpha^2}{2}} = (1+m);$$

therefore m is the mean of a jump size: for instance $m = -0.1$ corresponds in average to a 10%-decrease of price when it jumps.

A first result states that conditionally to the number of jumps, S_t is log-normally distributed.

Proposition 2.16. *For any $k \in \mathbb{N}$,*

$$\log(S_t/S_0) \underset{|N_t=k}{\stackrel{d}{=}} \mathcal{N}\left((\mu - \frac{\sigma^2}{2})t + k(\log(1+m) - \frac{\alpha^2}{2}), \sigma^2 t + \alpha^2 k\right).$$

Proof. First

$$\log(S_t/S_0) = (\mu - \sigma^2/2)t + \sigma W_t + \sum_{j=1}^{N_t}(\log(1+m) - \frac{\alpha^2}{2} + \alpha G_i)$$

where G_i are i.id. standard Gaussian variables, independent from N and W. Then, conditionally to $N_t = k$, the above is a Gaussian r.v. with mean

$$(\mu - \sigma^2/2)t + \sum_{j=1}^{k}(\log(1+m) - \frac{\alpha^2}{2})$$

and variance

$$\sigma^2 t + \sum_{j=1}^{k} \alpha^2.$$

Here, we have used the independence between all random objects. The proof is complete. $\square$

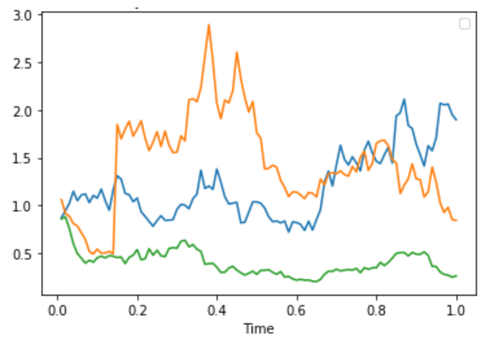

Figure 2: Sample paths of Merton model

2.5.3 Risk-neutral conditions

In view of our axioms (2.1.2), computing option price is made under a probability measure $\mathbb{Q}$ under with

$$\mathbb{E}_{\mathbb{Q}}\left(S_T\right) = S_0 e^{rT}. \tag{2.5.1}$$

Proposition 2.17. *Assume for S a Merton model under $\mathbb{Q}$. The condition (2.5.1) is satisfied if and only if*

$$\mu = r - \lambda m.$$

Proof. Decompose the expectation over the possible number of jumps and use Proposition 2.16:

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left(S_T\right) &= \sum_{k\geq 0} \mathbb{E}_{\mathbb{Q}}\left(S_T \mid N_T = k\right) \mathbb{P}\left[N_T = k\right] \\
&= \sum_{k\geq 0} \mathbb{E}_{\mathbb{Q}}\left(e^{\mathcal{N}\left((\mu-\frac{\sigma^2}{2})T+k(\log(1+m)-\frac{\alpha^2}{2}),\sigma^2 T+\alpha^2 k\right)}\right) e^{-\lambda T}\frac{(\lambda T)^k}{k!} \\
&= \sum_{k\geq 0} e^{(\mu-\frac{\sigma^2}{2})T+k(\log(1+m)-\frac{\alpha^2}{2})+\frac{1}{2}(\sigma^2 T+\alpha^2 k)} e^{-\lambda T}\frac{(\lambda T)^k}{k!} \\
&= e^{\mu T} e^{-\lambda T} \sum_{k\geq 0} \frac{((1+m)\lambda T)^k}{k!} \\
&= e^{\mu T} e^{-\lambda T} e^{(1+m)\lambda T} = e^{\mu T + \lambda m T}.
\end{aligned}$$

□

Exercise 2.18. *Replace the log-normal distribution by a double exponential distribution [KW04], i.e. when the distribution of $\log(1+Y_i)$ is given by*

$$\nu(\mathrm{d}y) = p\eta_1 e^{-\eta_1 y}\mathbf{1}_{y>0}\mathrm{d}y + q\eta_2 e^{-\eta_2 y}\mathbf{1}_{y<0}\mathrm{d}y$$

with $p \geq 0$, $q \geq 0$, $p+q=1$, $\eta_1 > 1$, $\eta_2 > 0$. Check that the risk-neutral condition writes $\mu + \lambda(\frac{p}{\eta_1 - 1} - \frac{q}{\eta_2+1}) = r$.

2.5.4 Merton's formula for Call option

We are now in a position to derive the price of call option in the Merton model.

Theorem 2.19. *Assume a Merton model under a risk-neutral pricing measure $\mathbb{Q}$ (with the condition $\mu + \lambda m = r$). Then*

$$\mathbb{E}\left(e^{-rT}(S_T - K)_+\right) = \sum_{k\geq 0} e^{-\lambda T}\frac{(\lambda T)^k}{k!}\texttt{Call}^{\texttt{BS}}(0, T, S_0(1+m)^k e^{-m\lambda T}, K, \sigma^2_{BS} = \sigma^2 + \alpha^2 k/T, r).$$

Proof. Decompose over the number of jumps:

$$\begin{aligned}\mathbb{E}\left(e^{-rT}(S_T - K)_+\right) &= \mathbb{E}\left(e^{-rT}(S_0 e^{(\mu-\sigma^2/2)T+\sigma W_T}\prod_{j=1}^{N_T}(1+Y_j) - K)_+\right)\\ &= \sum_k \mathbb{P}\left[N_T = k\right]\mathbb{E}\left(e^{-rT}(S_T - K)_+ \mid N_T = k\right).\end{aligned}$$

From Proposition 2.16 , we have $S_T \mid N_T = k \overset{d}{=} S_0 \exp(\mathcal{N}((\mu - \frac{\sigma^2}{2})T + k(\log(1+m) - \frac{\alpha^2}{2}), \sigma^2 T + \alpha^2 k))$. Thus, the computation of $\mathbb{E}\left(e^{-rT}(S_T - K)_+ \mid N_T = k\right)$ is similar to that for Black-Scholes formula, provided that we properly adjust parameters. This adjustment works as follows: the above distribution equals that of

$$S_T = \tilde{S}_0 \exp((r - \frac{\tilde{\sigma}^2}{2})T + \tilde{\sigma} W_T),$$

with

$$\tilde{\sigma}^2 T = \sigma^2 T + \alpha^2 k$$

and

$$S_0 \exp\left((\mu - \frac{\sigma^2}{2})T + k(\log(1+m) - \frac{\alpha^2}{2})\right) = \tilde{S}_0 \exp\left((r - \frac{\tilde{\sigma}^2}{2})T\right).$$

Simplifying the latter expression (using $\tilde{\sigma}$ and the risk-neutral condition) gives

$$\tilde{S}_0 = S_0 e^{k\log(1+m)+(\mu-r)T} = S_0(1+m)^k e^{-\lambda m T}.$$

□

Remark 2.20. *Computing numerically the Merton call formula is easy, although it writes as a series. Indeed, we can truncate the infinite sum keeping only the few first terms $k =$*

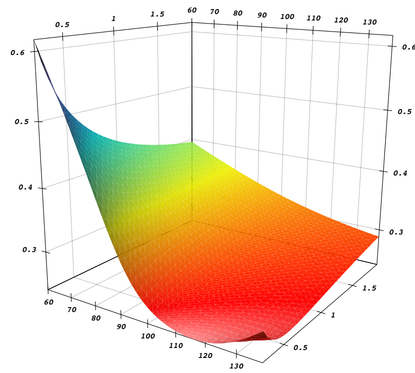

Figure 3: A sample of Implied volatility in Merton model

$0, \dots, k_0 - 1$, *the convergence is exponentially fast in* $k_0 \to +\infty$. *To be more precise, the No Arbitrage bound on call price gives*

$$\texttt{Call}^{\texttt{BS}}(0, T, S_0(1+m)^k e^{-m\lambda T}, K, \sigma_{BS}^2 = \sigma^2 + \alpha^2 k/T, r) \le S_0(1+m)^k e^{-m\lambda T}$$

and thus, truncating the sum after the term k_0 *yields an error at most equal to*

$$\begin{aligned} S_0 \sum_{k \ge k_0} e^{-\lambda T} \frac{(\lambda T)^k}{k!} (1+m)^k e^{-m\lambda T} &\le S_0 e^{-\lambda T} \frac{(\lambda T)^{k_0}}{k_0!} (1+m)^{k_0} e^{-m\lambda T} \sum_{k \ge k_0} \frac{(\lambda T)^{k-k_0}}{(k-k_0)!} (1+m)^{k-k_0} \\ &= S_0 e^{-\lambda T} \frac{(\lambda T)^{k_0}}{k_0!} (1+m)^{k_0} e^{-m\lambda T} e^{\lambda T(1+m)} \\ &= S_0 \frac{(\lambda T(1+m))^{k_0}}{k_0!}. \end{aligned}$$

2.5.5 Implied volatility surface

The implied volatility $\sigma_I(T, K)$ within the Merton model is defined as

$$\texttt{Call}^{\texttt{Merton}}(0, S_0, T, K) = \texttt{Call}^{\texttt{BS}}(0, S_0, T, K, \sigma_I(T, K)).$$

See Figure 3. It turns out that the presence of jumps helps to generate a smile, as observed in the market. Thus, fitting a Merton model to market data may a chance to work. However, hedging an option in this model is difficult since jumps lead to non-zero residual risks, see [EG11] for a discussion.

2.6 Numerical pricing using Fourier techniques

In this section, we present numerical techniques used to compute call/put prices when the distribution has some nice analytical properties. These techniques should be seen as tricks to compute very fastly prices, and as a result, it helps a lot for calibrating such a model in (quasi)-real-time. The "nice analytical properties" does not include all models, but some interesting ones such as the Heston model, the Bates model...

Let $T > 0$. We consider an asset, denote its price by S, and assume to know either the distribution of S_T or its characteristic function. We are interested in methods to price the European call following the axiomatic rules (2.1.2):

$$C_0^{\texttt{Model}}((S_T - K)_+, T) := \mathbb{E}\left(e^{-rT}(S_T - K)_+\right),$$

using a risk-neutral expectation. Here the risk-free rate is constant and equal to $r > 0$.

2.6.1 When the distribution of S_T is known

In the case where we know the distribution of S_T, the formula for call price can be rewritten as follows:

$$C(0, T, S, K) := e^{-rT}\mathbb{E}[(S_T - K)_+] = e^{-rT}\int_{\mathbb{R}}(y - K)_+ p(\mathrm{d}y),$$

where $p(\mathrm{d}y)$ is the distribution of S_T. This formula can either be used to obtain closed-form formula (see e.g. Black-Scholes formula) or to approximate the call price using usual integral approximation methods such as trapezoidal rule, etc. The distribution of S_T is known when the dynamic of S falls into the following cases: Black-Scholes model, Bachelier model, Merton model... otherwise it is usually not explicit and the above numerical integration can not be handled.

2.6.2 When the distribution of S_T is unknown

In some models, the distribution of S_T is unknown, but still, one may have a closed-form formula for the moment generating function of S_T defined as follows:

$$\Phi(z) := \mathbb{E}[e^{zX_T}], \qquad z \in \mathbb{C},$$

where $X_T := \log(S_T e^{-rT})$. Examples of such models include:

- The Heston model in which ν, the volatility of S, is stochastic and the process (ν, S) has the following dynamic:

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t \mathrm{d}t + \sqrt{\nu_t} S_t \mathrm{d}W_t^S, \\ \mathrm{d}\nu_t &= k(\theta - \nu_t)\mathrm{d}t + \xi\sqrt{\nu_t}\mathrm{d}W_t^\nu \ , \end{aligned}$$

  where W^S and W^ν are two correlated standard Brownian motion with correlation ρ.

  As shown in [DDS00], for example, the characteristic function has a log linear form; i.e. one can find A, B and C such that for all $(z_X, z_\nu) \in \mathbb{R}^2$,

$$\Phi(z) := \mathbb{E}[e^{z_X X_T + z_\nu \nu_T}] = e^{A(z_X, z_\nu, T) + B(z_X, z_\nu, T)X_0 + C(z_X, z_\nu, T)\nu_0}.$$

The next two sections present formula for the price of European call as summation of characteristic function.

2.6.3 Carr-Madan Formula

Let $k > 0$ be a log-strike. Write the price of an European call, using lo-variables:

$$C(k) := e^{-rT}\mathbb{E}\left[\left(e^{X_T+rT} - e^k\right)_+\right] ,$$

where $X_T := \log(S_T e^{-rT})$. It corresponds to $C(0, T, S, K)$ with strike $K := e^k$. The following Theorem is of great help to get fast approximation of the call price (compared to the approximation obtained by crude Monte Carlo for example) in the case where an explicit formula for $\Phi : z \mapsto \mathbb{E}[e^{zX_T}]$ is available in a tractable form.

Theorem 2.21 (Carr-Madan formula)**.** *Assume that there exist $\alpha > 0$ and $p > 1 + \alpha$, such that S_T admits a finite moment of order p, i.e. $\mathbb{E}[e^{pX_T}] < +\infty$. Then it holds:*

$$C(k) := e^{-\alpha k}\int_{\mathbb{R}} \frac{e^{-iuk}}{2\pi} e^{(iu+\alpha)rT} \frac{\Phi(iu + \alpha + 1)}{(iu + \alpha)(\alpha + 1 + iu)} \mathrm{d}u. \tag{2.6.1}$$

The result is presented for the first time in [CM98]. It relies mainly on the following classical result [Rud66, Theorem 9.11].

Lemma 2.22. *(Inversion Theorem) Let $f \in L^1(\mathbb{R}, \mathrm{d}x)$ and denote its Fourier transform by $\hat{f} : u \mapsto \int_{\mathbb{R}} f(x)e^{ixu}dx$. Assume $\hat{f} \in L^1(\mathbb{R}, \mathrm{d}x)$, then it holds a.e.:*

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(u)e^{-iux}\mathrm{d}u.$$

Proof of Theorem 2.21. Let $z(k) := e^{\alpha k}C(k)$ for $\alpha > 0$: z is continuous in $\mathbb{R}$, and moreover:

- By non-arbitrage argument, we have $C(k) \leq S_0$, thus $|z(k)| \leq e^{\alpha k}S_0$, and $k \mapsto e^{\alpha k}$ is integrable on $(-\infty, 0]$. So z is integrable on $(-\infty, 0]$.

- Observe that:

$$\begin{aligned}(e^{X_T+rT} - e^k)_+ &\leq e^{X_T+rT}\mathbf{1}_{X_T+rT\geq k}\\ &\leq e^{X_T+rT}e^{(X_T+rT-k)(p-1)}\mathbf{1}_{X_T+rT\geq k}\\ &\leq e^{(X_T+rT)p}e^{-(p-1)k}.\end{aligned}$$

  After taking expectation, we have:

$$|z(k)| = e^{\alpha k}e^{-rT}\mathbb{E}\left((e^{X_T+rT} - e^k)_+\right) \leq e^{-rT}\mathbb{E}\left(e^{(X_T+rT)p}\right)e^{(\alpha-(p-1))k}.$$

  In addition, $k \mapsto e^{(\alpha-(p-1))k}$ is integrable on $[0, \infty)$ since $\alpha-(p-1) < 0$; so z is integrable on $[0, \infty)$.

Hence we have shown that $z \in L^1(\mathbb{R}, \mathrm{d}x)$. Its Fourier transform is then well-defined, given by:

$$\begin{aligned}
\hat{z}(u) &= \int_{\mathbb{R}} e^{iuk} z(k) \mathrm{d}k \\
&= e^{-rT} \int_{\mathbb{R}} e^{iuk} e^{\alpha k} \mathbb{E}\left((e^{X_T + rT} - e^k) \mathbf{1}_{k \leq X_T + rT} \right) \mathrm{d}k \\
&= e^{-rT} \mathbb{E}\left(\int_{-\infty}^{X_T + rT} (e^{(iu+\alpha)k + X_T + rT} - e^{(\alpha+1+iu)k}) \mathrm{d}k \right) \\
&= e^{-rT} \mathbb{E}\left(\frac{e^{(iu+\alpha+1)(X_T+rT)}}{iu+\alpha} - \frac{e^{(\alpha+1+iu)(X_T+rT)}}{\alpha+1+iu} \right)
\end{aligned}$$

(since the contribution at -∞ is 0)

$$= e^{(iu+\alpha)rT} \frac{\Phi(iu+\alpha+1)}{(iu+\alpha)(\alpha+1+iu)}.$$

Observe that $\hat{z} \in L^1(\mathbb{R}, \mathrm{d}x)$, so we can apply Lemma 2.22 to get a.e.

$$z(k) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iuk} \hat{z}(u) \mathrm{d}u = \int_{\mathbb{R}} \frac{e^{-iuk}}{2\pi} e^{(iu+\alpha)rT} \frac{\Phi(iu+\alpha+1)}{(iu+\alpha)(\alpha+1+iu)} \mathrm{d}u.$$

Since $z(k) = e^{\alpha k} C(k)$, we have

$$C(k) = e^{-\alpha k} \int_{\mathbb{R}} \frac{e^{-iuk}}{2\pi} e^{(iu+\alpha)rT} \frac{\Phi(iu+\alpha+1)}{(iu+\alpha)(\alpha+1+iu)} \mathrm{d}u.$$

Since z is continuous, the result holds everywhere on $\mathbb{R}$, as stated in the theorem. □

Numerical aspects. We discuss below about the advantages of Carr-Madan formula, as well as its limitations. Theorem 2.21 expresses the price of a call as the inverse of a Fourier transform. This is a very nice result because Fast Fourier Transform algorithms can be used for fast computation of many calls at different strikes, which might be useful to compute the smile or Greeks, etc. More precisely, by setting $\nu_j := \eta(j-1)$ and discretizing (2.6.1) using the trapezoidal rule, for any strike $k > 0$, we obtain:

$$C(k) \approx \frac{e^{-\alpha k}}{\pi} \sum_{j=1}^{N} e^{-i\nu_j k} \Phi_T(\nu_j) \eta,$$

where we denoted $\Phi_T(u) := \frac{1}{2} e^{(iu+\alpha)rT} \frac{\Phi(iu+\alpha+1)}{(iu+\alpha)(\alpha+1+iu)}$. For any given value of k, by using N multiplications we obtain an estimate for $C(k)$ as above. Next, if N values of $C(k)$ were sought along a regular grid of k, one would need naively N^2 multiplications: however, Fast Fourier Transform algorithms (FFT) can be used to reduce the N^2 operations into $N \log(N)$ operations only.

The main drawback of Carr-Madan formula is that the call price given by the formula depends on a parameter α that remains to be chosen by the user. For example, after discretizing the integral by trapezoidal scheme, taking two different values of α will give two different approximations of the call[2]. Hence the user will certainly wonder which value of α he should take. Keep in mind that taking a rather small value requires a careful look at the integration close to 0, because the integrand is not integrable for $\alpha = 0$; and that taking α too large is not possible if S_T does not admit moments of all orders.

2.6.4 Lewis Formula

Lewis formula relies on the following classical result [Rud66, Theorem 9.13] and does not suffer of choosing a parameter α which competes with the number of finite moments of S_T. The result is presented for the first time in [Lew01].

Lemma 2.23 (Parseval-Plancherel theorem)**.** *Assume that f and $g \in L^2(\mathbb{R}, \mathrm{d}x)$. Then the following equality holds:*

$$\frac{1}{2\pi}\int_{\mathbb{R}} \overline{\hat{f}(\nu)}\hat{g}(\nu)d\nu = \int_{\mathbb{R}} f(x)g(x)\mathrm{d}x.$$

Theorem 2.24 (Lewis formula)**.** *It holds:*

$$C(k) = S_0 - \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{(1/2-iu)\log(Ke^{-rT})}}{1/4+u^2}\Phi(1/2+iu)\mathrm{d}u. \tag{2.6.2}$$

Proof of Theorem 2.24. We assume first that X_T admits a density w.r.t. $\mathrm{d}x$ and we denote it by $\rho(x)$.

• Write that

$$(e^{X_T+rT} - K)_+ = e^{X_T+rT} - \min(e^{X_T+rT}, K).$$

Taking expectation and multiplying by e^{-rT}, we have:

$$e^{-rT}\mathbb{E}\left((e^{X_T+rT} - K)_+\right) = \mathbb{E}\left(e^{-rT}e^{X_T+rT}\right) - e^{-rT}\mathbb{E}\left(\min(e^{X_T+rT}, K)\right). \tag{2.6.3}$$

• The first term in the r.h.s. of (2.6.3) can be computed as follows:

$$\mathbb{E}\left(e^{-rT}e^{X_T+rT}\right) = \mathbb{E}\left(e^{-rT}S_T\right) = S_0,$$

since we stand under the pricing measure, under which S has the same return as the risk-free rate.

• We now compute the second term in the r.h.s of (2.6.3):

$$\begin{aligned} &e^{-rT}\mathbb{E}[\min(e^{X_T+rT}, K)] \\ &= \int_{\mathbb{R}} \min(e^x, Ke^{-rT})\rho(x)\mathrm{d}x = \int_{\mathbb{R}} \min(e^x, Ke^{-rT})e^{-\frac{x}{2}}\rho(x)e^{\frac{x}{2}}\mathrm{d}x. \end{aligned} \tag{2.6.4}$$

For $x \in \mathbb{R}$, denote $f(x) := \min(e^x, Ke^{-rT})e^{-\frac{x}{2}}$, and $g(x) := \rho(x)e^{\frac{x}{2}}$.

[2] Of course, the two price match when the discretization step goes to 0.

Simpler case: Assume first that $\underline{g \in L^2(\mathbb{R}, \mathrm{d}x)}$.
▷ We easily check that f decays exponentially quickly as $x \to \pm\infty$, hence $f \in L^2$, so that we can apply Lemma 2.23 to the r.h.s. of (2.6.4). This gives

$$e^{-rT}\mathbb{E}[\min(e^{X_T+rT}, K)] = \frac{1}{2\pi}\int \overline{\hat{f}(u)}\hat{g}(u)du. \tag{2.6.5}$$

It remains to compute the two Fourier transforms to conclude:

$$\begin{aligned}
\hat{g}(u) &= \int_{\mathbb{R}} e^{iux}\rho(x)e^{\frac{x}{2}}\mathrm{d}x = \Phi(\frac{1}{2}+iu),\\
\hat{f}(u) &= \int_{\mathbb{R}} e^{iux}\min(e^x, Ke^{-rT})e^{-\frac{x}{2}}\mathrm{d}x\\
&= \int_{\log(Ke^{-rT})}^{+\infty} Ke^{-rT}e^{iux-\frac{x}{2}}\mathrm{d}x + \int_{-\infty}^{\log(Ke^{-rT})} e^{iux+\frac{x}{2}}\mathrm{d}x\\
&= -Ke^{-rT}\frac{e^{(iu-\frac{1}{2})\log(Ke^{-rT})}}{iu-\frac{1}{2}} + \frac{e^{(iu+\frac{1}{2})\log(Ke^{-rT})}}{iu+\frac{1}{2}}\\
&= e^{(iu+\frac{1}{2})\log(Ke^{-rT})}\frac{1}{u^2+\frac{1}{4}}.
\end{aligned}$$

▷ Plugging the Fourier transform expressions of $\overline{\hat{f}(u)}$ and $\hat{g}(u)$ into (2.6.5) gives the desired result and conclude the proof of the theorem in the simpler case.

General case: Now we $\underline{\text{drop the assumption that } g \in L^2(\mathbb{R}, \mathrm{d}x)}$. Take a small perturbation of X_T, i.e. $X_T^\varepsilon = X_T + G_\varepsilon$, where $G_\varepsilon \sim \mathcal{N}(0,\varepsilon)$ admits the density $g_\varepsilon(y) := \frac{e^{-\frac{y^2}{2\varepsilon}}}{\sqrt{2\pi\varepsilon}}$. It is an easy exercice to prove that the distribution of X_T^ε is absolutely continuous w.r.t. the Lebesgue measure in $\mathbb{R}$, with density $\rho_\varepsilon : z \in \mathbb{R} \mapsto \mathbb{E}\left(g_\varepsilon(z - X_T)\right)$. Indeed, for any bounded measurable function ϕ, we have:

$$\mathbb{E}\left(\phi(X_T^\varepsilon)\right) = \mathbb{E}\left(\int_{\mathbb{R}}\phi(X_T+y)g_\varepsilon(y)\mathrm{d}y\right) = \int_{\mathbb{R}}\phi(z)\mathbb{E}\left(g_\varepsilon(z-X_T)\right)\mathrm{d}z,$$

where we have changed variable and applied the Fubini theorem (on non-negative quantities). Observe that:

$$\begin{aligned}
\int_{\mathbb{R}}\rho_\varepsilon^2(z)e^z\mathrm{d}z &\le \int_{\mathbb{R}}\mathbb{E}\left(g_\varepsilon^2(z-X_T)\right)e^z\mathrm{d}z\\
&= \mathbb{E}\left(\int_{\mathbb{R}} e^z e^{-\frac{(z-X_T)^2}{\varepsilon}}\frac{\mathrm{d}z}{2\pi\varepsilon}\right)\\
&= \mathbb{E}\left(\int_{\mathbb{R}} e^{X_T+y}e^{-\frac{y^2}{\varepsilon}}\frac{\mathrm{d}y}{2\pi\varepsilon}\right) = S_0\int_{\mathbb{R}} e^y e^{-\frac{y^2}{\varepsilon}}\frac{\mathrm{d}y}{2\pi\varepsilon} < +\infty
\end{aligned}$$

where the first inequality holds by Jensen inequality. Hence, $x \mapsto \rho_\varepsilon(x)e^{\frac{x}{2}} \in L^2(\mathbb{R}, \mathrm{d}x)$, and we can apply (2.6.5) to get:

$$\begin{aligned}\mathbb{E}\left(e^{-rT}\min(e^{X_T^\varepsilon}, K)\right) &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{(\frac{1}{2}-iu)\log(Ke^{-rT})}}{\frac{1}{4}+u^2}\mathbb{E}\left(e^{(\frac{1}{2}+iu)X_T^\varepsilon}\right)\mathrm{d}u \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{(\frac{1}{2}-iu)\log(Ke^{-rT})}}{\frac{1}{4}+u^2}\mathbb{E}\left(e^{(\frac{1}{2}+iu)X_T}\right)\mathbb{E}\left(e^{(\frac{1}{2}+iu)G_\varepsilon}\right)\mathrm{d}u. \quad (2.6.6)\end{aligned}$$

It remains to let ε go to 0 on both sides of (2.6.6) to conclude. On the one hand, the right hand side converges to the quantity with $\varepsilon = 0$ by the dominated convergence theorem. The left hand side also converges, because $|\min(e^{X_T^\varepsilon}, K)|$ is bounded and continuous in ε. We are done with the Lewis formula. □

Numerical Aspects One may have to use numerical schemes to approximate the integral in (2.6.2). Once again this can be done using trapezoidal method or Gauss quadrature method. Note that the integrand is $\mathcal{O}(\frac{1}{x^2})$ at ∞ so that it converges fastly to 0. This is helpful for stability and numerical aspects.

References

[Bac00] L. Bachelier. Théorie de la spéculation. PhD thesis, Ann. Sci. École Norm. Sup., 1900.

[BBF02] H. Berestycki, J. Busca, and I. Florent. Asymptotics and calibration of local volatility models. Quantitative finance, 2(1):61–69, 2002.

[BG12] R. Bompis and E. Gobet. Asymptotic and non asymptotic approximations for option valuation. In T. Gerstner and P. Kloeden, editors, Recent Developments in Computational Finance: Foundations, Algorithms and Applications, chapter 4, pages 159–241. World Scientific Publishing Company, 2012.

[CM98] P. Carr and D.B. Madan. Option valuation using the fast Fourier transform. Journal of Computational Finance, 2(4):61–73, 1998.

[DDS00] J. Pan D. Duffie and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. Econometrica, 68(6):1343–1376, 2000.

[EG11] N. El Karoui and E. Gobet. Les outils stochastiques des marchés financiers: une visite guidée de Einstein à Black-Scholes. Editions de l'Ecole Polytechnique, 2011.

[Gat06] J. Gatheral. The Volatility Surface: A Practitioner's Guide. Wiley Finance, 2006.

[HKLW02] P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward. Managing smile risk. Willmott Magazine, pages 84–108, 2002.

[Kin93] J.F.C. Kingman. Poisson processes, volume 3 of Oxford Studies in Probability. The Clarendon Press Oxford University Press, New York, 1993. Oxford Science Publications.

[KW04] S.G. Kou and H. Wang. Option pricing under a double exponential jump diffusion model. Management Science, 50(9):1178–1192, 2004.

[Lew01] A. Lewis. A simple option formula for general jump-diffusion and other exponential Lévy processes. 2001.

[Mer76] R.C. Merton. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics, 3:125–144, 1976.

[MK82] S. Manaster and G. Koehler. The calculation of implied variances from the Black-Scholes model: A note. The Journal of Finance, 37(1):227–230, 1982.

[Rub83] M. Rubinstein. Displaced diffusion option pricing. Journal of Finance, 38(1):213–217, 1983.

[Rud66] W. Rudin. Real and complex analysis. McGraw-Hill, Toronto, 1966.

[ST08] W. Schachermayer and J. Teichmann. How close are the option pricing formulas of Bachelier and Black-Merton-Scholes? Math. Finance, 18(1):155–170, 2008.