Fast Fourier Transform for Discrete Asian Options

Eric Benhamou*

First version: September 1999. This version: March 17, 2000.

JEL Classification: G13
MSC classification: 62P05
Keywords: Fast Fourier Transform, Asian options, Convolution, Fat-Tails.

Abstract

This paper presents an efficient methodology for the discrete Asian options consistent with different types of underlying densities, especially non-normal returns as suggested by the empirical literature (Mandelbrot (1963) and Fama (1965)). Based on Fast Fourier Transform, the method is an enhanced version of the algorithm of Caverhill and Clewlow (1992). The contribution of this paper is to improve their algorithm and to adapt it to non-lognormal densities. This enables us to examine the impact of fat-tailed distribution on price as well as on delta. We find evidence that fat tails lead to wider jumps in the delta.

Address: Financial Markets Group, Room G303, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom, Tel 0171-955-7895, Fax 0171-242-1006. E-mail: E.Ben-Hamou@lse.ac.uk. An electronic version is available at the author web-page: http://cep.lse.ac.uk/fmg/people/ben-hamou/index.html
I would like to thank Nicole El Karoui, Pierre Mella Barral, Alexander Müiermann, Stéphane Mallat, Stefano Galluccio, Grigoris Mamalis, Han Lee Aurore Lecanon and the participants of the Ph.D. seminar in Economics at the London School of Economics for interesting discussions and remarks. All errors remain mine.
1 Introduction

First introduced in Tokyo, Asian options are options based on any type of average of underlying equity prices, interest rates or indices. They are among the most popular path-dependent derivatives, since their characteristics capture partially the trajectory of the underlying, with often reduced exposure to volatility. In addition, Asian options are less sensitive to possible spot manipulations or extreme movements at settlement and offer flexibility in the way the average is settled. Consequently, they have become very attractive for investors since they provide a customized cheap way to hedge periodic cash-flows (see Longstaff (1995) for a discussion of the efficiency of Asian interest-rate options for corporations with reasonably predictable cash flows).

In the option pricing literature, the Black Scholes model often remains the standard model. This is particularly true for Asian option since the attention of many researchers has been kept by the puzzling fact that the distribution of an arithmetic average of lognormal distributions is not explicit. However, the empirical literature has rejected normality of returns and hence the geometric Brownian motion and rather suggested fat-tailed distributions (see Mandelbrot (1963) and Fama (1965) for the early ones).

The motivation of this paper is therefore to provide an efficient method for the pricing of Asian option consistent with various underlying densities especially non log-normal ones. Because of the challenge of a correct price for Asian option with the most accepted option pricing model, previous research has concentrated on the Black Scholes model, adopting different strategies. It has first focussed at geometric Asian option (Vorst (1992), Turnbull and Wakeman (1991), Zhang (1995)). It has as well simplified the question to the continuous time Asian options (Geman and Yor (1993), Rogers and Shi (1995), Alziary et al. (1997), He and Takahashi (1996), Forsyth et al. (1998), Nielsen and Sandmann (1998)). However, the average for traded Asian options is arithmetic and discrete, either daily, weekly or monthly. Approximating these option by their continuous time limit is inaccurate and misleading for options with a period of time between two prices longer than a day.

To take account for discrete averaging, it has been suggested to use different approximations of the density of the sum of lognormal variables leading to various closed form solutions: approximation via the geometric average (Vorst (1992)), via a log normal density (Turnbull and Wakeman (1991)), via an Edgeworth expansion (Levy (1992) and Jacques (1996)), via a Taylor expansion (Zhang (1998)) or via the reciprocal Gamma distribution (Milesvky and Posner (1997)).

It has also been advocated to use different numerical methods: Monte Carlo (Kemma and Vorst (1990)), tree methods (Hull and White (1997)) and Fast Fourier Transform techniques (Caverhill and Clewlow (1992)). However, all these works did not consider non-lognormal distributions.

When the underlying densities is not lognormal, the approximation methods
do not hold any more since they rely on the lognormal assumption. Numerical methods like PDE or lattice methods are as well not easy to adapt to the non-lognormal case, since we need to restrict ourselves to certain types of diffusion like stochastic volatility or deterministic volatility models which implies strong assumptions on the underlying diffusion. It is not very straightforward to derive an empirical density from market data, requiring very often a calibration stage. The two methods adaptable to an ad-hoc empirical non lognormal distribution without too much difficulty, are indeed the Monte Carlo and the Fast Fourier Transform method. However, these two methods perform poorly for non-lognormal case as well as for lognormal one. The Monte Carlo has the drawback to be slow. The algorithm of Caverhill and Clewlow (1992) requires large discretization grid and has slow convergence.

In this paper, we offer a solution to improve the method of Caverhill and Clewlow (1992) and to adapt it to the case of non-lognormal densities. To reduce the size of the grid and therefore the computational time, we recenter intermediate densities. We test this algorithm in the lognormal case since it is only in this particular situation that we have benchmarks in the literature. We then examine the impact of non-lognormal densities on the price as well as on the delta.

The remainder of this paper is organized as follows. In section 2, we describe our algorithm in detail. In section 3, we examine numerical results for the lognormal case, using it as a benchmark for the efficiency of our method. Section 4 deals with non-lognormal densities. It examines the impact of various densities on the price of the option as well as on the delta. We conclude briefly in section 5 suggesting further developments.

2 Description of the method

2.1 Framework

We consider a continuous time trading economy with infinite horizon. The uncertainty in the economy is classically modelized by a complete probability space \((\Omega, F, Q)\). The underlying is denoted by \((S_t)_{t \in \mathbb{R}^+}\). The information evolves according to the natural filtration \((F_t)_{t \in \mathbb{R}^+}\) implied by the underlying process. Following the traditional empirical literature, we assume that returns \((R_t)_{t \in \mathbb{N}}\), defined by \(R_t = \log(S_t/S_{t-1})\) for a given sequence of time \((t_i)_{i \in \mathbb{N}}\), are independently distributed and have a well known density \(f_i(\cdot)\), with a well known mean denoted by \(\mu_i\). In the case of the Black Scholes model, these densities are normal distribution with mean \(\left(r - \frac{\sigma^2}{2}\right)(t_i - t_i-1)\) and variance \(\sigma^2(t_i - t_i-1)\). The underlying price is then calculated as the initial price \(S_{t_0}\) increased by the different returns \(e^{R_{t_i}}\):

\[S_{t_i} = S_{t_0}e^{R_{t_1}+R_{t_2}+\ldots+R_{t_i}}\]
Assuming that we have \( n \) fixing dates for the average, denoted by \( t_1, t_2, \ldots, t_n \), the arithmetic average \( A \) is defined as the sum of the different fixings:

\[
A = \frac{1}{n} \sum_{i=1}^{n} S_{t_i}
\]  

(1)

In complete markets with no arbitrage opportunity, there is a unique risk neutral martingale measure denoted by \( Q \). In this framework, the price \( P \) of an Asian call, with strike \( K \), expiring at time \( T \), is defined as the expected value of the payoff discounted at the risk free rate \( r \):

\[
P = \mathbb{E}^Q \left[ e^{-rT} (A - K)^+ \right]
\]  

(2)

where \( X^+ \) stands for \( \max(X, 0) \). Since the discrete average process has no well known density, there is no closed formula. However, we show in this paper that we can compute numerically this density, giving a method which converges to the real densities as long as the size of the discretisation grid tends to the infinity.

2.2 Why Fast Fourier Transform?

Well known in signal theory, Fast Fourier Transform (FFT) is efficient for the computation of many numerical problems. Precisely, the FFT is an efficient algorithm for computing the sum:

\[
FFT (f(k)) = \frac{1}{\sqrt{2\pi}} \sum_{j=1}^{N-1} e^{-i\frac{2\pi}{N}(j-1)(k-1)} f(j) \quad \text{for} \ k = 1 \ldots N
\]

where \( N \) is typically a power of 2. This algorithm reduces the number of multiplications in the required \( N \) summations from an order of \( N^2 \) to that of \( N \log_2 (N) \). It suggests that for a \( 2^p \) grid, the complexity is \( p2^p \), which is typically the complexity of a binomial tree.

Recently, this technique has gained popularity in option valuations (Baskhi and Chen (1998), Scott (1997), Chen and Scott (1992), Carr and Madan (1999)). The property of the Fourier transform used here is its efficiency to calculate convolution products. The Fourier transform of a convolution product is simply the product of the Fourier transforms. This is helpful in getting the density of the sum of two variables since it is the convolution product of the individual densities as long as the variables are independent. In the case of the Asian option, the expression is not a straightforward sum of independent variables. In the algorithm section, we show how to use independent variables in a recursive scheme.

The interest of this method is its efficiency compared to a straightforward computation of the density. Instead of computing an \( n - 1 \) dimensional integral with a complexity of \( O(N^{n-1}) \), we reduce this complexity by means of Fast Fourier Transform to \( O(N^2 \log (N)) \).
The use of FFT method for Asian option was first suggested by Carverhill and Clewlow (1992). However, this is developed for lognormal densities and is not very efficient since it requires large grid and had slow convergence. To cope with smaller grid, we introduce a proxy for the mean of intermediate densities. This enables us to recenter the different variables. We extend as well the FFT method to non-lognormal densities. We concentrated on Student density as a well known example of fat-tailed distribution. Indeed, the method explained here is very general and can be applied to many other fat-tailed densities, like extreme value, Pareto and generalized Pareto distributions.

2.3 Algorithm

2.3.1 Inefficiency of the Carverhill and Clewlow method

To calculate simply a density of a sum of variables, it is worth obtaining independent variables. With our assumptions on the independence of returns, this comes naturally. Notice that when the underlying distribution is lognormal, returns are normal and their Fourier transform has a closed form solution equal to

\[
f(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(w-m)^2}{2\sigma^2}}
\]

where \( m \) stands for the mean and \( \sigma^2 \) the variance.

We introduce the following recursive sequence \((B_i)_{i=1}^{n}\) defined by its initial condition: \( B_1 = R_{t_n} \) and for \( i = 2...n \),

\[
B_i = R_{t_{n+1-i}} + \log (1 + \exp B_{i-1})
\]

The Steward and Hodges factorization expresses the sum variable \( A \) defined by (1) in terms of the last term of the recursive sequence \( B_{n-1} \):

Proposition 1  The sum variable \( A \) can be expressed in terms of the last term of the recursive sequence \( B_{n-1} \):

\[
A = \frac{S_{t_0}}{n} e^{B_n}
\]

Proof: we decompose the underlying price as a function of the difference of returns: \( S_{t_i} = S_{t_0} e^{R_{t_1} + R_{t_2} + ... + R_{t_i}} \). Factoring terms leads to a multiplicative expression of the sum variable:

\[
A = \frac{S_{t_0}}{n} [e^{R_{t_1}} \ast (1 + e^{R_{t_2}} \ast (1 + e^{R_{t_3}} (1 + ... (1 + e^{R_{t_n}})))))])
\]

When taking the logarithm of the above equation, we get an additive expression:

\[
A = \frac{S_{t_0}}{n} \ast \exp (R_{t_1} + \log (1 + \exp (R_{t_2} + \log (... + \log (1 + R_{t_n}))))))
\]

the term inside the exponential can be calculated recursively with the sequence \((B_i)_{i=1..n} \).
The proposition 1, with the recursive equation (3) was the starting point of the work of Carverhill and Clewlow (1992). In the recursive equation (3), at each step, we add the return \( R_{t_{n+1}} \) which is not centered and has often a positive mean. For high volatilities, however the mean is negative since it is equal to \( r - \frac{\sigma^2}{2} \). For positive mean, the distribution of \( B_{t_{n+1}} \) is consequently shifted to the right of the distribution of \( B_t \). If we discretise the distribution of \( B_{t_{n+1}} \) on the same grid as the one of \( B_t \), this implies that the discretisation grid be large enough to contain the two distributions. When we have \( n \) dates in our arithmetic average, this means that the grid contains \( n \) distributions more and more spaced as shown in figure 1. This is precisely why the algorithm of Carverhill and Clewlow requires a large grid.

![Figure 1: Evolution of the densities](image)

**2.3.2 Recentering intermediate densities**

To cope with a smaller grid and therefore reduce computational time, we can recenter densities at each step. The difficulty here is that we do not know the exact mean of the variable \( B_i \). Denoting by \( \mu_i \) the mean of the return \( R_{t_i} \) (\( \mu_i = \mathbb{E} [R_{t_i}] \)), which is supposed to be well known, we can approximate the mean of the variable \( B_i \) with the following sequence: \( (m_i)_{i=1...n} \) initialized with \( m_1 = \mu_n \) and for \( i = 2...n \)

\[
m_i = \mu_{n+1-i} + \log (1 + \exp m_{i-1}) \tag{4}
\]

The term \( m_i \) acts as a proxy for the mean of the variable \( B_i \). The approximation of the average is done by taking the lagged \( B_{i-1} \) equal to its mean \( m_{i-1} \) in the recursive equation (3). It is worth noticing that even if we do an approximation on the mean, it does not mean that we approximate the density of \( B_i \). It just means that we do not perfectly center this variable. However, there is no new error implied by the recentering. Indeed, since the function \( \log (1 + e^x) \) is convex, we are underestimating some convexity adjustment term as stated by the Jensen inequality for convex functions \( f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)) \).
The recentered sequence is defined as \( (A_i)_{i=1}^n \) with \( A_i = B_i - m_i \). Replacing \( B_{i-1} \) by its expression in terms of \( A_{i-1} \) and \( m_{i-1} \) leads to a recursive two dimensional sequence summarized by the following proposition:

**Proposition 2** The sum variable \( A \) can be expressed in terms of the last term of the recursive sequence \( A_n \) and \( m_n \):

\[
A = \frac{S_t}{n} e^{A_n + m_n}
\]

where the sequence \( (m_i)_{i=1}^n \) is defined as above (4) and the sequence \( (A_i)_{i=1}^n \) is given by the initial condition \( A_1 = R_t - m_1 \) and for \( i = 1..n \)

\[
A_i = R_{t_{n+1-i}} + \log (1 + \exp A_{i-1} \exp m_{i-1}) - m_i
\]

To get the density of \( A_i \) with respect to the one of \( A_{i-1} \), we use standard variable change theorem. It states that the density of a variable \( Y = g(X) \), denoted by \( dp_Y \), is given by the density of \( X \), denoted by \( dp_x \), expressed in terms of \( Y \) times the absolute value of the Jacobian of the function \( f^{-1}(Y) = X \)

\[
dp_Y = dp_x \left| J_{f^{-1}}(Y) \right| dy
\]

leading to the interpolation formula:

\[
dp_{\log(1+e^{m_{i-1}+x})} = \frac{e^{y+m_i}}{e^{y+m_i} - 1} p_X \left( \log \left( e^{y+m_i} - 1 \right) - m_{i-1} \right) 1\{y > m_i\} dy \quad (5)
\]

We can now describe the different steps of the first algorithm. The algorithm is initialized with the value of the two dimensional sequence \( m_1 = \mu_n \) and \( A_1 = R_t - m_1 \). It finishes when we get \( m_n \) and \( A_n \).

The recursive sequence is calculated as follows. Assume that we know the value of the bi-dimensional sequence at step \( i - 1 \), that is \( m_{i-1} \) and \( A_{i-1} \).

- We then interpolate the variable \( A_{i-1} \) by means of the remark (5) to get the density of the variable \( \log \left( 1 + e^{m_{i-1}+A_{i-1}} \right) - m_i \).
- We calculate the density of \( A_i \) as the sum of the two independent variable \( R_{t_{n+1-i}} \) and the variable \( \log \left( 1 + e^{m_{i-1}+A_{i-1}} \right) - m_i \) by calculating the convolution product via FFT.
- Having obtained the density of the average, we calculate the payoff of the option, defined as an expectation, by a numerical integration, using the Simpson rule.
2.3.3 Discussion of the numerical techniques

The FFT algorithm requires the density function to be represented at a sufficient number of equally spaced points. The grid for the discretisation of the different densities needs to be sufficiently thin as well as sufficiently large to avoid interference errors implied by the periodisation of the density function in the FFT algorithm. We use the FFT algorithm as described in Press et al. (1992).

Errors in the numerical integration by the Simpson rule (exact for the integration of polynom up to degree 3) are negligible compared to the ones produced by the discretisation of the distribution. The error in the Simpson rule for the integral of a function \( f \) infinitely differentiable \( \int_a^b f(x) \, dx \) can be shown to be an \( O \left( \left( \frac{b-a}{2} \right)^5 f^{(4)} \right) \).

3 Efficiency of the algorithm for the lognormal case

3.1 Black Scholes assumption

The information evolves according to the augmented filtration \( \{ F_t, t \in [0, T] \} \) generated by a standard one dimensional standard Brownian motion \( (W_t)_{t \in \mathbb{R}_+} \). We assume the underlying price process is a geometric Brownian motion, solution of the Black Scholes (1973) diffusion defined by equation (6) with initial condition \( S_{t=0} = S_0 \)

\[
\frac{dS_t}{S_t} = \mu dt + \sigma dW_t
\]  

In this case the returns \( R_{t_i} \) have a normal density with mean \( \left( r - \frac{\sigma^2}{2} \right) (t_i - t_{i-1}) \) and variance \( \sigma^2 (t_i - t_{i-1}) \).

3.2 Choice of the Grid

The choice of an efficient grid is not easy. The grid is determined by its range as well as its number of points. Choosing a range not correctly leads to interference errors. Taking a grid not dense enough leads as well to inaccurate Fourier Transform computation. We choose a centered grid with 4096 points, that is \( 2^{12} \) and with a width of \( 9n\sigma\sqrt{dt} \), where \( n \) stands for the number of fixing dates, \( \sigma \) the volatility, \( dt \) the time between two fixings. For a one year weekly Asian option, with fifty fixing, the number of fixing \( n \) is equal to 50 and the period of time between two following prices \( dt \) is equal to one week or \( 1/52 \) of a year.
3.2.1 Recentering the densities

The improvement of this paper is to recenter densities at each step. Since we approximate the mean, the recentering is imperfect as figure 2 shows. For low volatilities up to 20%, densities are perfectly recentered for a one-year weekly Asian option. For volatility higher than 20%, the approximation of the mean is not rigorously correct and leads to a shift of the different densities to the right. Indeed, since the function $\log(1 + e^x)$ is convex, we are underestimating a convexity adjustment term, as stated by the Jensen inequality for a convex function $f \cdot f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$. However, the bias in our estimation is quite small, since for high values of $x$, the function $\log(1 + \exp(x))$ is very little convex, roughly equal to $x$, justifying our method.

![Figure 2: Evolution of the density with recentering at each step. The two graphics concern a one year weekly Asian option. The figure on the right is with 30% of volatility whereas the one on the left is for 20% volatility.](image)

In the figure 2, we can see that for small volatility level like 20% (figure on the left), the recentering is perfect whereas for higher volatility like 30% (figure on the right), we are missing the convexity adjustment term. In the original algorithm of Caverhill and Clewlow, the grid size can be shown to be equal to $9n\sigma \sqrt{dt} + \mu_n$. The gain in our method can be measured by the grid width ratio $(9n\sigma \sqrt{dt} + m_n)/9n\sigma \sqrt{dt}$. For the case of a one year option with 10% volatility and a risk free rate of 20%, this gain is equal to 1.317. This means that with the old algorithm, we need 1317 points to get the same precision as 1000 points with the new one. This means that the equivalent of 4096 points with the new algorithm is about 5400 points with the old algorithm.

3.2.2 Interference on the FFT algorithm

When the grid is not large enough, interference alters the results quality as shown in figure 3, where we used only a grid width of $4n\sigma \sqrt{dt}$. This comes from the fact that the FFT algorithm assumes the periodicity of our function. It can cause interference terms when the grid size is too small.
3.3 Comparison of the different methods

Because of no well-known example, we arbitrary decided to use the same option example as in the work of Levy (1992) as a benchmark. We compute the price of a one year Asian option, with the underlying starting at 100 ($S = 100$), with a risk free interest rate of 10% ($r = 10\%$), and 50 fixings per year (weekly average with two weeks of holidays).

The results, given in the table 1, compare different methods and show that the convolution method is efficient for the pricing of Asian option. In the different column, MC stands for Monte Carlo with its standard error given in the next column SE. WE means Wilkinson Estimates, E Edgeworth method, RG the reciprocal Gamma approximation, CV the Convolution method of Caverhill and Clewlow, CVR the convolution method with recentering. The reference price is the one of the Monte Carlo simulation. The efficiency of a formula is given by its comparison with this reference price.

We found that recentering the density improves significantly the efficiency of the Fast Fourier Transform method for high volatilities since the estimation of the density becomes more important. Among traditional approximation methods, we tested Wilkinson estimates, Edgeworth expansion, and the reciprocal gamma approximation. We found that Wilkinson estimates was the most robust method. The Edgeworth expansion formula can blow up when the third and fourth moments are too different from the ones implied by a lognormal. We found as well poor results for the reciprocal gamma approximation. This comes from the small number of variables in our Asian options. The density of the average is therefore far from its asymptotic limit, which can be shown to be a reciprocal gamma density Milesvky and Posner (1997).
Table 1: Comparison of different methods for the Asian option $\sigma$ stands for the volatility, $K$ for the strike

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>K</th>
<th>MC</th>
<th>Std Err</th>
<th>WE</th>
<th>E</th>
<th>RG</th>
<th>CV</th>
<th>CVR</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>80</td>
<td>22.78</td>
<td>0.00</td>
<td>22.78</td>
<td>22.78</td>
<td>21.64</td>
<td>22.78</td>
<td>22.78</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>13.73</td>
<td>0.00</td>
<td>13.73</td>
<td>13.73</td>
<td>13.1</td>
<td>13.73</td>
<td>13.73</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>5.24</td>
<td>0.00</td>
<td>5.25</td>
<td>5.25</td>
<td>4.98</td>
<td>5.25</td>
<td>5.25</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>0.72</td>
<td>0.00</td>
<td>0.72</td>
<td>0.72</td>
<td>0.71</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>0.03</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>30%</td>
<td>80</td>
<td>23.07</td>
<td>0.01</td>
<td>23.14</td>
<td>23.07</td>
<td>21.92</td>
<td>23.09</td>
<td>23.08</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>15.22</td>
<td>0.01</td>
<td>15.30</td>
<td>15.16</td>
<td>14.46</td>
<td>15.29</td>
<td>15.26</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>9.01</td>
<td>0.01</td>
<td>9.08</td>
<td>9.00</td>
<td>8.56</td>
<td>9.08</td>
<td>9.05</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>4.83</td>
<td>0.01</td>
<td>4.84</td>
<td>4.85</td>
<td>4.58</td>
<td>4.86</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>2.35</td>
<td>0.01</td>
<td>2.33</td>
<td>2.40</td>
<td>2.23</td>
<td>2.40</td>
<td>2.33</td>
</tr>
<tr>
<td>50%</td>
<td>80</td>
<td>24.83</td>
<td>0.03</td>
<td>25.06</td>
<td>24.10</td>
<td>23.58</td>
<td>25.01</td>
<td>24.88</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>18.32</td>
<td>0.03</td>
<td>18.57</td>
<td>17.83</td>
<td>17.40</td>
<td>18.50</td>
<td>18.37</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>13.18</td>
<td>0.03</td>
<td>13.34</td>
<td>13.02</td>
<td>12.52</td>
<td>13.47</td>
<td>13.20</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>9.23</td>
<td>0.03</td>
<td>9.33</td>
<td>9.36</td>
<td>8.77</td>
<td>9.45</td>
<td>9.19</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>6.36</td>
<td>0.03</td>
<td>6.37</td>
<td>6.63</td>
<td>6.04</td>
<td>6.68</td>
<td>6.40</td>
</tr>
</tbody>
</table>

3.4 Density Comparison

Among standard approximation methods, we found that the most reliable one was the one based on the Wilkinson estimates. The reciprocal gamma approximation is inappropriate in our case since the density is far from its asymptotic limit. The Edgeworth expansion method could blow up when third and fourth moments were too different from the ones implied by a lognormal.

Our results confirm that the lognormal approximation slightly overprices Asian options (Levy and Turnbull (1992), Zhang (1998)). This is indicated by the skew to the right of the Wilkinson estimates (or lognormal approximation) density in figure 4. The efficiency of the FFT method is confirmed by the closed fit with the Monte Carlo sampling in figure 4. The Monte Carlo sampling was based on a simulation of a Sobol sequence with 30,000 draws.
It is worth noting that the precision of the method is heavily depending on the type of the options: in-at or out-the-money. One should expect little difference in price for options depending on a wide part of the distribution like in or at-the-money options. However, for out-the-money options, that are depending mainly on the tails of the distribution, there is a real advantage in terms of precision to use the Fast Fourier Transform method compared to Wilkinson estimates. Indeed, fat tails are the true motivation of this paper. It is already interesting to realize that even in the case of a lognormal underlying, the Fast Fourier Transform method takes better account for fat tails than most standard approximation methods with closed form.

4 Using non-lognormal densities

A major contradiction with normality of returns has been the instability of the Black Scholes implied volatility with different strikes and option maturities (see Dumas et al. (1995)). This has been referred as the volatility smile (Hull and White (1987), Chesney and Scott (1989) and Heston (1992)). This has indicated that fat-tailed distributions are more appropriate for modelling returns.

The interest of our methodology lies in its flexibility on the distributions of returns. We do not assume any specific distribution. The distribution is an input like other parameters. Therefore, we can use distribution derived from market data, like option prices. In this paper, we decided to illustrate the fat-tailed distribution with the specific case of a Student density. This is because this density is often used in the literature and converges as well to the normal density when the number of degree of freedom tends to infinity. Indeed, there are many other densities which could have been used, like Pareto, generalized Pareto, power-laws distributions and many more.
4.1 Densities for leptokurtic effect

To take account for leptokurtic returns, we assume that centered and pseudo normalized (with a parameter $\lambda > 1$) returns $\frac{R_i - (r - \frac{\mu^2}{2})(t_i - t_{i-1})}{\sqrt{\sigma^2(t_i - t_{i-1})}}$ have a density given by a Student distribution with a degree of freedom $n = \frac{2\lambda}{\lambda - 1}$ given by

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1 + \frac{\mu^2}{n}\right)^{-\frac{n+1}{2}}.$$  

The cumulative distribution is then given by

$$\Pr (X \leq t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \int_{-\infty}^{t} \left(1 + \frac{1}{n} u^2\right)^{-\frac{n+1}{2}} du$$

where $\Gamma (y) = \int_{0}^{\infty} e^{-x^2} x^{y-1} dx$ is the Gamma function at $y$. Since a student density has always a variance bigger than one we need to specify this variance by the parameter $\lambda$.

4.2 Numerical results

4.2.1 Effect on the price

As expected, fat-tailed distributions hereby illustrated by the Student density leads to a more expensive price of the Asian option. The Fast Fourier method is efficient as confirmed by a comparison with Monte Carlo simulations with 20,000 draws. To simulate Student density, we simulate uniform distribution and inverse the cumulative distribution by means of the approximation given in the appendix section.

Without any surprise, the discrepancy between the lognormal distribution and a distribution with fat tails increase with the volatility. It increases as well for distribution with fatter tails as shown by the increase of price between the Student density with 44 degree of freedom and the one with only 22. We have chosen the Student density since its asymptotic distribution is precisely the normal distribution when the degrees of freedom tend to infinity.

Interestingly, practitioners have kept on using the lognormal approximation for the Asian option. We have seen that this approximation is not valid when assuming lognormal underlying. It tends to overprice the Asian option. However, when assuming fat-tailed distribution for the underlying, we found as well a more expensive price. This explains the use by practitioners of the lognormal approximation since it includes the rise of price due to fat-tailed distributions.
<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>K</th>
<th>Lognormal 44 df</th>
<th>Student 44 df</th>
<th>MC Student 44 df</th>
<th>Student 22 df</th>
<th>MC Student 22 df</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>22.7838</td>
<td>22.7911</td>
<td>22.7914</td>
<td>22.8021</td>
<td>22.8028</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>100</td>
<td>5.2438</td>
<td>5.2843</td>
<td>5.2850</td>
<td>5.3278</td>
<td>5.3294</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>0.7211</td>
<td>0.7642</td>
<td>0.7649</td>
<td>0.8078</td>
<td>0.8094</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>0.0336</td>
<td>0.0358</td>
<td>0.0362</td>
<td>0.0423</td>
<td>0.0430</td>
</tr>
<tr>
<td>80</td>
<td>23.0733</td>
<td>23.2033</td>
<td>23.2050</td>
<td>23.3372</td>
<td>23.3406</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>15.2231</td>
<td>15.4014</td>
<td>15.4036</td>
<td>15.5808</td>
<td>15.5855</td>
<td></td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>4.8338</td>
<td>5.0345</td>
<td>5.0379</td>
<td>5.2355</td>
<td>5.2426</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>2.3545</td>
<td>2.5097</td>
<td>2.5117</td>
<td>2.6682</td>
<td>2.6728</td>
</tr>
<tr>
<td>80</td>
<td>24.8324</td>
<td>25.2213</td>
<td>25.2243</td>
<td>25.6099</td>
<td>25.6168</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>18.3207</td>
<td>18.7471</td>
<td>18.7510</td>
<td>19.1694</td>
<td>19.1779</td>
<td></td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>6.3615</td>
<td>6.7363</td>
<td>6.7412</td>
<td>7.1104</td>
<td>7.1209</td>
</tr>
</tbody>
</table>

**Table 2**: Price of Asian option with Fat-tails distributions. $\sigma$ stands for the volatility, K for the strike, df for degrees of freedom.

### 4.2.2 Delta hedging

The punch line of our numerical method is to examine the impact of fat-tailed distribution on the delta. Even if finite differences can lead to rough estimates of the Greeks (Benhamou (2000)), classical reduction techniques are relying on lognormal assumptions. Well known for discrete Asian option, delta jumps appear as soon as a fixing dates vanish.

The comparative study of the delta evolution with lognormal density and student density shows that fat-tailed distribution leads to higher jumps in the delta. This is logically due to the fact that fat-tailed distributions imply more expensive price and therefore larger drop of the price with the downfall of a fixing date. The difference in the delta is quite significant as shown by figure 5 and 6. The figure 5 show the evolution of the delta for a weekly Asian option far from the maturity of the option. The student density taken here is the one with 22 degrees of freedom.
There is no rule concerning the difference between the delta for lognormal densities and for Student densities. In the figure 5, the option is 50 to 44 weeks before the expiration. In this particular case, the delta implied by the Student density is on average more expensive. This is not the case when the option is close to the maturity as shown by figure 6 where there are only 10 to 1 week before the maturity of the option. However, it is worth noting that on average the delta is quite closed for the two different densities. This suggest that for a long-run delta hedging, assuming normal returns is not too much inaccurate. However, for short run delta hedging, the assumptions on the returns densities lead to very different hedging strategies.

5 Conclusion

In this paper, we have seen that Fast Fourier Transform is an efficient way of pricing discrete Asian options with non-lognormal densities. The systematic re-
centering of intermediate densities enables to reduce the size of the grid so as to fasten the convergence. We show that the price of the Asian option should be more expensive with fat-tailed distributions. This indicates that approximation methods overpricing the Asian option incorporate, in a way, fat-tailed. However, as far as the delta is concerned, fat-tailed distributions lead to very different hedging strategies, especially on the short run.

Our methodology appeals many remarks. First, the Fast Fourier Transform technique enables to take into account volatility smile since as an input, we can take returns distribution derived by market data incorporating the smile effect. Second, the same approach can be applied with minor changes to basket and multi-asset options. Third, this methodology rises the issue of the way of deriving the density from market data properly.

6 Appendix: Inverse of the cumulative distribution of the student density

The general algorithm for computing the inverse $t_p$ of the cumulative distribution of the Student density, with $n$ degree of freedom is given below with $0 < p < 1$ and with $x_p$ the inverse of the cumulative distribution of the normal density $N(0, 1)$:

$$
t_p = x_p + \frac{g_1(x_p)}{n} + \frac{g_2(x_p)}{n^2} + \frac{g_3(x_p)}{n^3} + \frac{g_4(x_p)}{n^4}
$$

\[
g_1(x) = \frac{1}{4} (x^3 + x)
g_2(x) = \frac{1}{96} (5x^5 + 16x^3 + 3x)
g_3(x) = \frac{1}{384} (3x^7 + 19x^5 + 17x^3 - 15x)
g_4(x) = \frac{1}{92160} (79x^9 + 776x^7 + 1482x^5 - 1920x^3 - 945x)
\]
References


