

LOGNORMAL MOMENT MATCHING AND PRICING OF BASKET OPTIONS

PÅL NICOLAI HENRIKSEN

ABSTRACT. In this paper we will approximate the sum of the margins from a two dimensional lognormal variable by moment matching with a one dimensional lognormal variable. We will look at different cases of correlation and volatility in the two dimensional variable to analyse the accuracy of the approximation. Basket options will be valued based on the sum of two coupled lognormal assets and a univariate approximation to detect situations where the moment matching may be successfully applied in a financial market.

1. INTRODUCTION

Since there exists no closed form distribution function or characteristic function for the sum of two lognormal variables, extensive work has been done in the literature to approximate this. Examples are Zacks and Tsokos [8], Turnbull and Wakeman [7], Asmussen and Rojas-Nandayapa [2] and Hoedemakers [6]. One possible approximation is by moment matching, which will be the approach used in this document. Moment matching means that we will match the mean and variance of the bivariate lognormal variables $Y_1 + Y_2$ with the mean and variance of the univariate lognormal variable Y . An empirical analysis will be done based on the probability distributions of $Y_1 + Y_2$ and Y for different cases of volatility and correlation to assess the approximation.

We will see that moment matching is a good approximation in certain situations and quite bad in other. This makes an upper bound on the moment matching approximation error attractive. This error bound will be derived and displayed for the lognormal case based on work done by Akhiezer [1] and Berggren [3]. We will also use Taylor expansion to understand why the approximation fails in certain situations.

One important motivation for the moment matching approximation is valuation of options based on several underlying assets. Pricing options based on the approximated one dimensional asset can be done by closed form formulas, e.g. Black and Scholes, which can be calculated instantly. Pricing basket options based on two coupled underlying assets must as of date be done by time consuming simulations. We will look at situations where

Date: January 30, 2008.

Key words and phrases. Moment matching approximation, simulation, lognormal variables, error bound, basket option pricing.

the approximation works well and less well, and we will also have focus on whether the option pricing approximation is acceptable even if the distributional approximation is not.

Section 2 will give a short introduction to the univariate and bivariate lognormal models, and in Section 3 we will illustrate how one may simulate from the models. In Section 4 we will derive formulas for the moment matching, and look at the moment matching approximation through numerical examples and give a short analysis of the approximation error. In Section 5 we will price basket options of European call type based on simulations for the bivariate case and Black and Scholes formula for the approximated univariate case. In Section 6 we conclude.

2. THE LOGNORMAL MODEL

Most practical finance mathematics is based on the lognormal model given by

$$(2.1) \quad Y(t) = Y(0) \cdot e^{(r - \frac{1}{2}\sigma^2)t + \sigma W(t)},$$

where $Y(t)$ represents the risk neutral asset price, r the risk free interest rate and σ the asset volatility. $W(t)$ represents the Brownian motion stochastic term, which in this case is normally distributed $N(0, t)$. There are several advantages with this model; one is that the normal distribution gives a good fit to logreturn data, the lognormal model has given arise to closed form expressions of option prices through the Black and Scholes formula, the model is easily extended to higher dimensions.

Certain financial derivatives such as basket options and spread options are based on several underlying assets. Here we will study the case of two underlying correlated lognormal assets,

$$(2.2) \quad Y_1(t) = Y_1(0) \cdot e^{(r - \frac{1}{2}\sigma_1^2)t + \sigma_1 W_1(t)},$$

$$(2.3) \quad Y_2(t) = Y_2(0) \cdot e^{(r - \frac{1}{2}\sigma_2^2)t + \rho\sigma_2 W_1(t) + \sqrt{1 - \rho^2}\sigma_2 W_2(t)},$$

where ρ is the correlation coefficient between the logreturns of $Y_1(t)$ and $Y_2(t)$, while σ_1 and σ_2 are the volatility of the assets. $W_1(t)$ and $W_2(t)$ are two i.i.d. Brownian motions with probability distribution $N(0, t)$. From (2.2) and (2.3) it is easily seen that:

$$\begin{aligned} \ln Y_1(t) &\sim N(\ln Y_1(0) + (r - \frac{1}{2}\sigma_1^2)t, \sigma_1^2 t), \\ \ln Y_2(t) &\sim N(\ln Y_2(0) + (r - \frac{1}{2}\sigma_2^2)t, \sigma_2^2 t), \end{aligned}$$

which are the lognormal marginal distributions of the bivariate lognormal distribution for (Y_1, Y_2) . Using this model, all parameters are estimated through the covariance matrix of the asset logreturns.

3. SIMULATING LOGNORMAL VARIABLES

A straight forward Monte Carlo method is applied to simulate the univariate lognormal asset (2.1) at time t . The correlated bivariate lognormal assets (2.2) and (2.3) may be simulated in the same manner. The ease of simulating a bivariate lognormal variable is due to the fact that a bivariate normal distribution consists of two normally distributed

margins coupled by a correlation coefficient ρ . I.e. by fixing $Y_1(0)$, $Y_2(0)$, r , ρ , σ_1 and σ_2 and drawing i.i.d. ϵ_1 , ϵ_2 and ϵ_3 from a standard normal distribution, the lognormal variables are simulated by:

$$(3.1) \quad Y(t) = Y(0) \cdot e^{(r-\frac{1}{2}\sigma^2)t+\sigma\sqrt{t}\epsilon_1}$$

$$(3.2) \quad Y_1(t) = Y_1(0) \cdot e^{(r-\frac{1}{2}\sigma_1^2)t+\sigma_1\sqrt{t}\epsilon_1},$$

$$(3.3) \quad Y_2(t) = Y_2(0) \cdot e^{(r-\frac{1}{2}\sigma_2^2)t+\rho\sigma_2\sqrt{t}\epsilon_1+\sqrt{1-\rho^2}\sigma_2\sqrt{t}\epsilon_2}.$$

These formulae make the simulation of lognormal asset values at time t easy and efficient. If one would like to simulate a higher dimensional lognormal variable, this could be done just as easy by Cholesky decomposition of the covariance matrix of the logreturns.

4. MOMENT MATCHING

Similar work to that presented here for the lognormal case can be found in the literature, e.g. Zacks and Tsokos [8] or Borovkova, Permana and Weide [4].

The goal of the moment matching is to approximate the sum of the lognormal assets $Y_1(t) + Y_2(t)$ at time t :

$$(4.1) \quad Y(t) = (Y_1(0) + Y_2(0)) \cdot e^{a_Y t + \sigma_Y W(t)},$$

where $W(t) \sim N(0, t)$ is independent of $W_1(t)$ and $W_2(t)$. Expressions for a_Y and σ_Y are found by matching mean and variance of the variables. That is, by solving:

$$(4.2) \quad E(Y(t)) = E(Y_1(t) + Y_2(t)),$$

$$(4.3) \quad \text{Var}(Y(t)) = \text{Var}(Y_1(t) + Y_2(t)).$$

By using the facts that for $\ln X \sim N(\mu, \sigma)$, the mean and variance is given by:

$$E(X) = e^{\mu + \frac{1}{2}\sigma^2} \quad \text{Var}(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

and keeping in mind that $E(Y_1 + Y_2) = E(Y_1) + E(Y_2)$ and $\text{Var}(Y_1 + Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) + 2\text{Cov}(Y_1, Y_2)$, the parameters in (4.1) are found to be:

$$a_Y = r - \frac{1}{2}\sigma_Y^2$$

$$\sigma_Y^2 = \frac{1}{t} \cdot \ln \left(\frac{Y_1(0)^2 e^{\sigma_1^2 t} + Y_2(0)^2 e^{\sigma_2^2 t} + 2Y_1(0)Y_2(0)e^{\rho\sigma_1\sigma_2 t}}{(Y_1(0) + Y_2(0))^2} \right)$$

We immediately see that (4.1) is equivalent to (2.1) with σ replaced by σ_Y . This form on $Y(t)$ is risk-neutral, which is a basis for option pricing. However, (4.1) is not a risk neutral process in the sense that the moment matching is only valid at time t . Thus, we can not simulate risk neutral asset paths of $Y(t)$ based on a_Y and σ_Y , but only a moment matched lognormal variable at time t .

We have now established the framework where we can start simulating $Y_1(t) + Y_2(t)$ to compare the probability distribution of the sum with the probability distribution of $Y(t)$.

4.1. Numerical examples. A good way of comparing $Y_1 + Y_2$ with Y , is to simulate the probability density function of $Y_1 + Y_2$ and compare this with the analytic probability density function of Y . This comparison can be found in Figure 1 for different choices of ρ , σ_1 and σ_2 . All the simulations use $r = 0.04$, $t = 1$ and $Y_1(0) = Y_2(0) = 100$. The number of simulations in each case is 10^7 . This huge number of simulations has been used to minimize simulation error.

From Figure 1 it can be seen that for $\rho = -0.95$ the probability densities are nowhere close to each other. This result coincide with results from Zacks and Tsokos [8] where it is derived that the moment matching approximation works better for non-negative than negative correlations. Also for major differences in volatility between Y_1 and Y_2 the approximative distribution differs from the true one. The density of $Y_1 + Y_2$ is greater than Y at the peak of the distributions. Because of the variance matching, this means that the tail of $Y_1 + Y_2$ is somewhat heavier than the tail of Y . This is confirmed in Figure 2 where we have taken the logarithm of the densities to highlight the details in the distribution tails. It is clear that as σ_1 increases the probability densities become more and more similar for $\rho \geq 0$. For $\sigma_1 = \sigma_2$ the densities seem to be almost equal in our examples.

4.2. Analysis of approximation error. In Turnbull and Wakeman [7] the moment matching approximation error is studied through an Edgeworth expansion. In this section we will find an error bound of the moment matching approximation based on corollary 2.5.4 in Akhiezer [1] and Berggren [3]. In this report we will not discuss possible Monte Carlo simulation errors.

For the convenience of the reader, corollary 2.5.4 in Akhizer [1] is restated in the appendix. Based on the notation in the appendix, we let $F(x)$ be the cumulative distribution of $Y(t)$ and $G(x)$ the cumulative distribution of $Y_1(t) + Y_2(t)$. Further, we know from the moment matching that M_0 , M_1 and M_2 are equal for $F(x)$ and $G(x)$. The following expression for the error bound may be derived based on the corollary:

$$(4.4) \quad \left| \int_{-\infty}^{\kappa} dF(x) - \int_{-\infty}^{\kappa} dG(x) \right| \leq \left(\frac{1}{M_2 - M_1^2} (\kappa^2 - 2M_1\kappa + M_2) \right)$$

where M_1 and M_2 are given by:

$$\begin{aligned} M_1 &= e^{rt}(Y_1(0) + Y_2(0)) \\ M_2 &= e^{2rt}(Y_1(0)^2 e^{\sigma_1^2 t} + Y_2(0)^2 e^{\sigma_2^2 t} + 2Y_1(0)Y_2(0)e^{\sigma_1\sigma_2\rho t}) \end{aligned}$$

An example of the error bound is given in Figure 3. From the figure and expression (4.4) it is clear that the error bound is not very tight. As we see, around $E(Y_1 + Y_2)$, where Figure 3 peaks, the error bound is without information. However, in the tails expression (4.4) does indeed give us information about maximum possible error one makes when approximating the sum of two lognormal variables by matching the first two moments of one lognormal variable. This information is relevant for e.g. far out of the money options.

The error bound does not explicitly deal with the issue of different correlations and difference in volatility between Y_1 and Y_2 . As seen from Figure 1 and 2, negative correlation results in significantly worse approximations than non-negative correlations, and the same

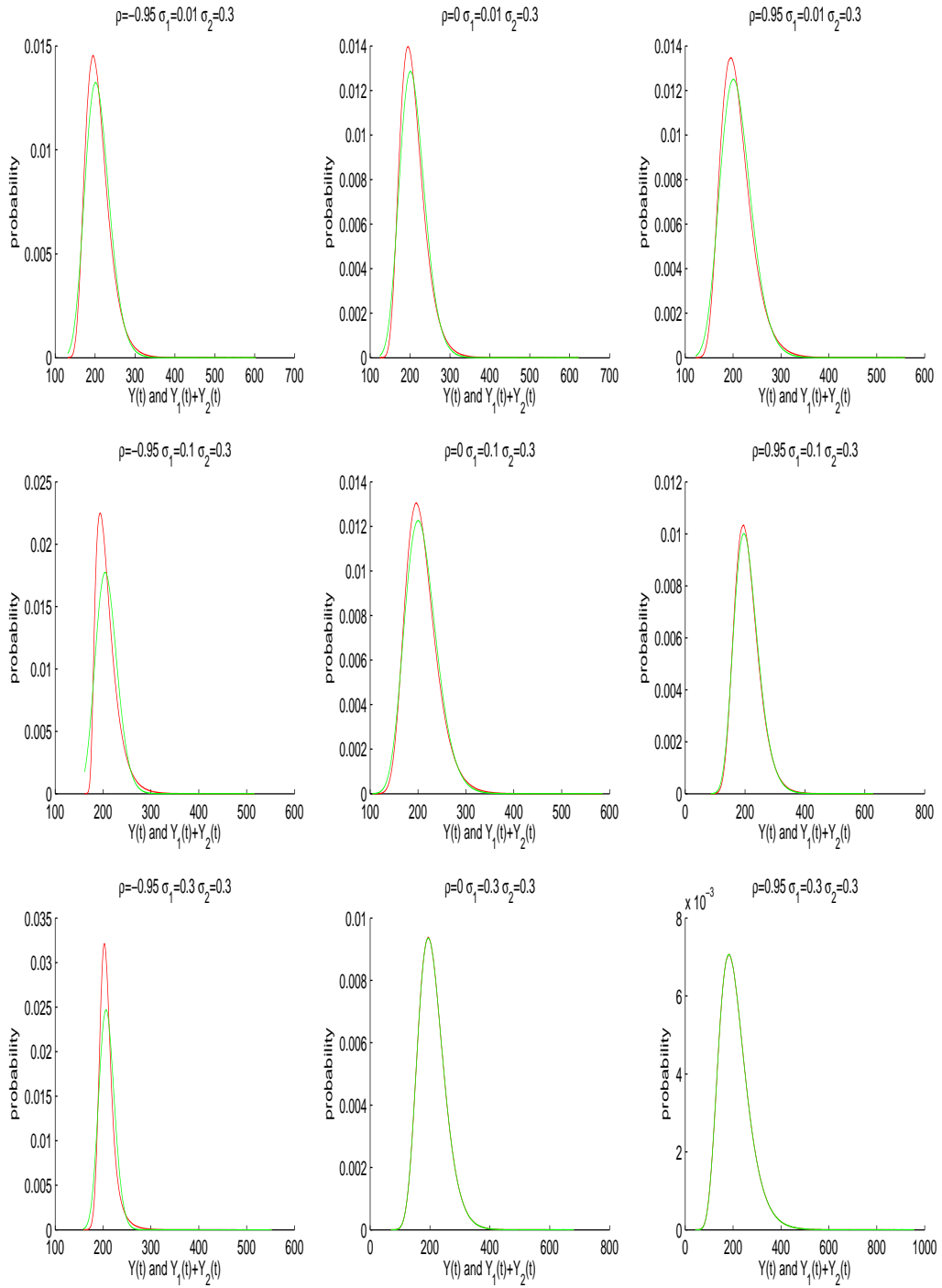


FIGURE 1. The distribution of $Y_1 + Y_2$ in red and Y in green for different choices of ρ , σ_1 and σ_2 .

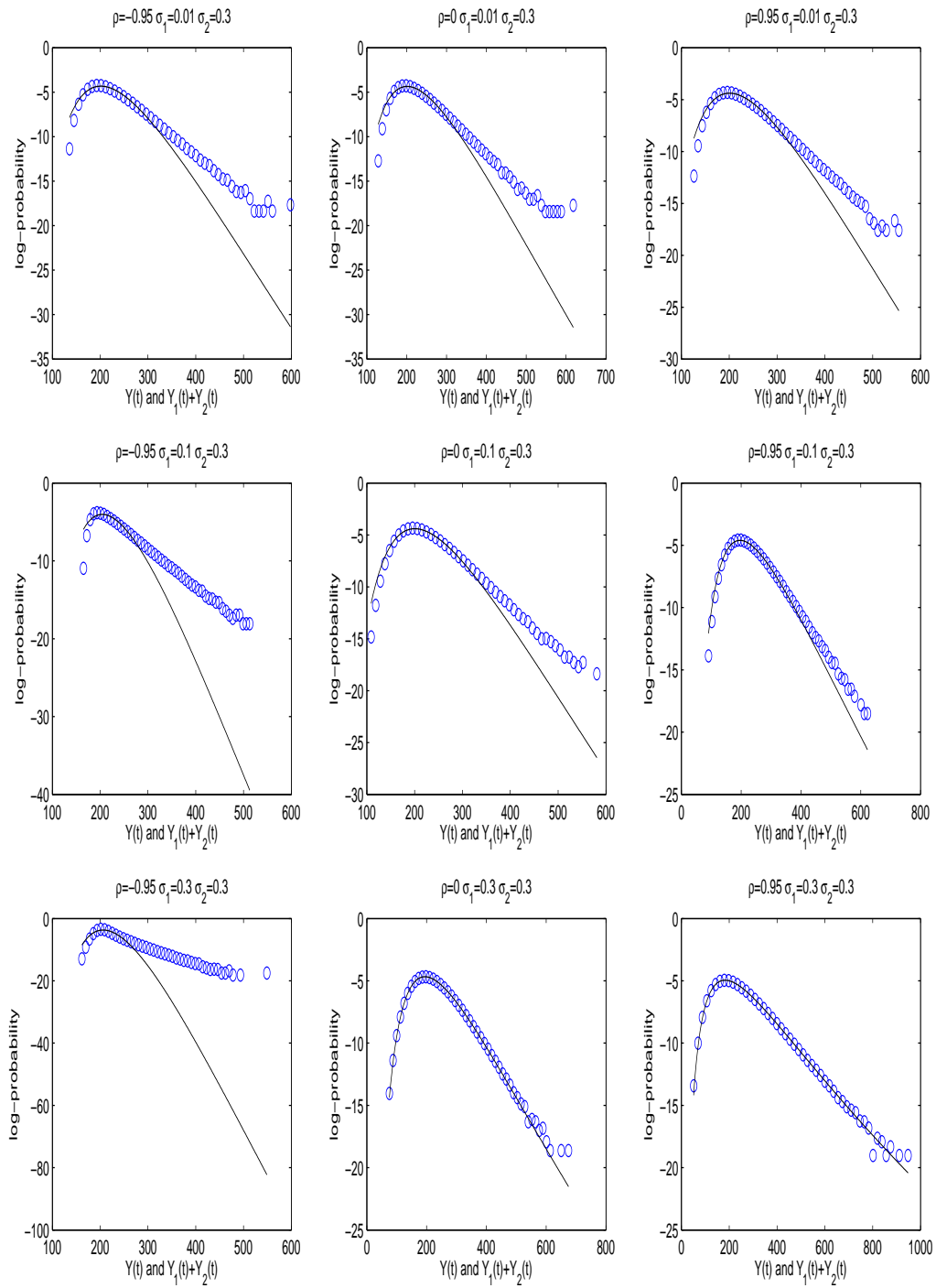


FIGURE 2. The logarithm of the $Y_1 + Y_2$ distribution as circles and Y as the whole line for different choices of ρ , σ_1 and σ_2 .

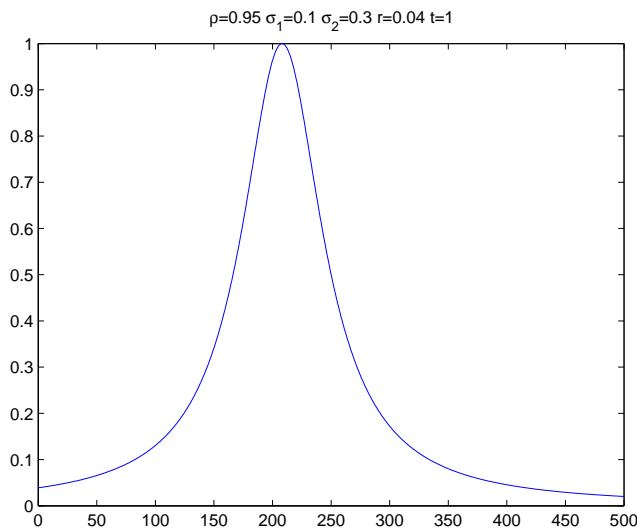


FIGURE 3. The error bound generated by expression (4.4).

is the case for a big difference in volatilities. For a better understanding of this, we will look at the following Taylor expansion:

$$\begin{aligned}
 \ln(e^X + e^Y) &= \frac{1}{2}X + \frac{1}{2}Y + \frac{1}{2}\ln(1 + e^{Y-X}) + \frac{1}{2}\ln(1 + e^{X-Y}) \\
 (4.5) \quad &\approx \frac{1}{2}X + \frac{1}{2}Y + \frac{1}{4}(\cosh(X - Y) - 1) + \ln(2) + O((X - Y)^4)
 \end{aligned}$$

Since we are studying the case where we are approximating $\exp(X) + \exp(Y)$, (X, Y) bivariate normally distributed, by $\exp(Z)$, Z normally distributed, the right hand side of (4.5) should be approximately normal. Since the margins of a bivariate normal distribution are normally distributed, it is clear that $X + Y$ is normal. Hence, one would want the rest of the terms on the right hand side to be as small as possible for the approximation to be good. Both the cosh-term and the rest-term of order 4 can be expected to be small when X and Y can be expected to be close to each other. This is clearly the case when the correlation is non-negative, and it is also the case, in a risk neutral setting, when the volatilities are similar.

5. PRICING OF BASKET OPTIONS

In this section we will price European call basket options based on two correlated assets to study the relationship between the moment matching approximation errors and the option prices. We wish to study if there is a smoothing in the pricing calculation that makes option prices similar even though the moment matching gives a not to good fit.

One great attribute when approximating $Y_1 + Y_2$ by Y is the simplification of option pricing. For European options this is particularly obvious as we can use the Black and

K	$\sigma_1, \sigma_2 \setminus \rho$	p_Y			$p_{Y_1+Y_2}$		
		-0.95	0	0.95	-0.95	0	0.95
150	0.01,0.3	55.999	56.028	56.061	55.883	55.887	55.929
	0.1,0.3	55.889	56.087	56.587	55.885	55.975	56.448
	0.3,0.3	55.882	56.879	59.266	55.886	56.869	59.271
200	0.01,0.3	15.907	16.255	16.594	15.412	15.789	16.181
	0.1,0.3	13.028	16.840	19.795	12.137	16.500	19.617
	0.3,0.3	10.810	20.958	27.218	9.974	20.947	27.221
250	0.01,0.3	1.681	1.873	2.068	2.133	2.292	2.472
	0.1,0.3	0.463	2.214	4.223	1.032	2.523	4.409
	0.3,0.3	0.055	5.117	10.567	0.387	5.132	10.566
300	0.01,0.3	0.078	0.101	0.128	0.231	0.256	0.290
	0.1,0.3	0.003	0.150	0.624	0.082	0.289	0.767
	0.3,0.3	$5.3 \cdot 10^{-6}$	0.929	3.694	0.026	0.944	3.689

TABLE 1. Option prices found from simulations of $Y_1(t) + Y_2(t)$ and from Black-Scholes formula based on the moment matched Y .

Scholes formula for the option price based on Y , whilst we need to use simulation for the option price based on $Y_1 + Y_2$. However, assuming no simulation error, we must keep in mind that the simulated option price based on $Y_1 + Y_2$ is the correct one.

In this document, we will compare European call option prices based on the fixed parameters $r = 0.04$, $Y_1(0) = Y_2(0) = 100$, exercise time $t = 1$ and varying ρ , σ_1 and σ_2 equivalent to the comparison of probability densities in Figure 1 and 2. Option prices are found at four different strikes; $K_1 = 150$, $K_2 = 200$, $K_3 = 250$ and $K_4 = 300$. Black and Scholes formula is applied for the option price p_Y based on Y :

$$p_Y = (Y_1(0) + Y_2(0))\Phi(u_1) - Ke^{-rt}\Phi(u_2)$$

where Φ represents the cumulative distribution function of a standard normal distribution, and

$$u_1 = \frac{\ln((Y_1(0) + Y_2(0))/K) + (r + \frac{1}{2}\sigma_Y^2)t}{\sigma_Y\sqrt{t}},$$

$$u_2 = \frac{\ln((Y_1(0) + Y_2(0))/K) + (r - \frac{1}{2}\sigma_Y^2)t}{\sigma_Y\sqrt{t}}.$$

We find the call price $p_{Y_1+Y_2}$ based on $Y_1 + Y_2$ by simulating Y_1 and Y_2 according to (3.2) and (3.3) and plugging this into:

$$(5.1) \quad p_{Y_1+Y_2} = e^{-rt}\mathbf{E}_Q[\max(Y_1(t) + Y_2(t) - K, 0)]$$

The Q represents the risk-neutral probability which corresponds to Y_1 and Y_2 being on the form (2.2) and (2.3). The options prices found are displayed in Table 1. From this table we see that for deep in the money options with strike $K = 150$ the moment matching approximation works well - even in the cases where the distributions in Figure 1 diverge.

This is not surprising when taking into account that very few of the simulation outcomes are below the strike, hence zero is hardly ever maximum in (5.1), and the first moment have been matched in equation (4.2). As the options go less and less deep into the money for strikes $K = 200$ and $K = 250$, the approximated option prices become less accurate. For $K = 200$ and non-negative correlation the approximation error is $< 3\%$. For negative correlation the approximation error is significant for this strike. An interesting point is that the approximative price, p_Y , overestimates the true price, $p_{Y_1+Y_2}$, for $K = 200$ while it underestimates the price for $K = 250$. Thus, for some strike $200 < K < 250$ the matching in the tails and the tail heaviness of the distributions start to dominate. This is clearly reflected in the option prices with strike $K = 300$. Here the option price approximation is very bad - except in the cases where the moment matched distribution tail coincides with the simulated distribution tail for $\sigma_1 = \sigma_2$ and $\rho \geq 0$. The accuracy of the approximative option price corresponds very well to the accuracy of the distribution in the tail plots of Figure 2 for the far out of the money options. One important result is that the worst approximation is for negative correlation and similar volatilities on Y_1 and Y_2 . This emphasizes the risk of moment matching approximations when the logreturns of the underlying assets are negatively correlated.

6. CONCLUSIONS AND FURTHER WORK

In this paper we have seen a method for approximating a sum of two coupled lognormal variables by matching the mean and the variance with a one dimensional lognormal variable. We have seen that this approximation may be done with ease by closed form formulas, and that the approximated variable preserves a risk neutral form, but is not in it self a risk neutral process. Through empirical analysis and analysis of approximation errors we have seen that the moment matching approximation works quite well as long as the correlation between logreturns of the two coupled lognormal variables is non-negative and the volatilities are relatively close to each other. For negative correlations the moment matching approximation is a less attractive method and should therefore be applied with care. In the pricing of European call basket options we saw that the moment matching approximation gave good results for in the money options - even in cases where the distribution of the approximated variable diverge significantly with the distribution based on the sum of the two coupled assets. Thus, there is some smoothing effect in the option pricing mechanism. On the other hand, the approximated prices had huge errors for out of the money options. The reason for this is that the approximated variable does not have a heavy enough tail. Improving this would be a good subject for further work, for instance by introducing variables that can pick up skewness and kurtosis in the financial market. In this paper, we have by construction moment matched *variables*. It could also be possible to restructure the moment matching to approximate a process, that is generalising it such that the approximation is valid for all times t . Other interesting further work could be to try to evaluate basket options by a totally different approach, e.g. by expanding the work done by Dempster and Hong [5] on transform based option pricing.

7. APPENDIX

Corollary 2.5.4 in Akhiezer says: If two nondecreasing functions $F(x)$ and $G(x)$ satisfy the relations

$$\begin{aligned} M_k &= \int_{-\infty}^{\infty} x^k dF(x) \\ &= \int_{-\infty}^{\infty} x^k dG(x) \quad k = 0, 1, 2, \dots, 2n - 2 \end{aligned}$$

where $\{M_k\}_0^{2n-1}$ is a positive sequence, then we have for arbitrary real κ

$$(7.1) \quad \left| \int_{-\infty}^{\kappa} dF(x) - \int_{-\infty}^{\kappa} dG(x) \right| \leq \Upsilon_{n-1}(\kappa)$$

Here $\Upsilon_n(\kappa)$ may be derived from the following:

For $n = 1, 2, \dots$ define the matrix

$$\mathbf{M}_n = \begin{bmatrix} M_0 & M_1 & \cdots & M_n \\ M_1 & M_2 & \cdots & M_{n+1} \\ \vdots & \vdots & \vdots & \vdots \\ M_n & M_{n+1} & \cdots & M_{2n} \end{bmatrix},$$

where M_n is the n th common moment of $F(x)$ and $G(x)$. Further, let

$$\mathbf{S}_n(\kappa) = \begin{bmatrix} M_0 & M_1 & \cdots & M_{n-1} & 1 \\ M_1 & M_2 & \cdots & M_n & \kappa \\ M_2 & M_3 & \cdots & m_{n+1} & \kappa^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ M_n & M_{n+1} & \cdots & M_{2n-1} & \kappa^n \end{bmatrix}.$$

From these matrices, define

$$\begin{aligned} P_0(\kappa) &= 1 \\ P_k(\kappa) &= \frac{\det \mathbf{S}_k(\kappa)}{\sqrt{\det \mathbf{M}_{k-1} \cdot \det \mathbf{M}_k}} \quad k \geq 1 \end{aligned}$$

from where we end up with

$$\Upsilon_n(\kappa) = \left(\sum_{k=0}^n |P_k(\kappa)|^2 \right)^{-1}$$

REFERENCES

- [1] Akhiezer N. I. (1965). The Classical Moment Problem and Some Related Questions in Analysis, Oliver Boyd: London
- [2] Asmussen, S., and Rojas-Nandayapa, L. (2006). Sums of dependent lognormal random variables: asymptotics and simulation. *Research Report*, Thiele Centre, University of Aarhus
- [3] Berggren, F. (2005). An error bound for moment matching methods of lognormal sum distributions. *Europ. Trans. Telecom.*, **16**, pp. 573-577
- [4] Borovkova, S., Permana, F. and Weide, H.(2006). A closed form approach to the valuation and hedging of basket and spread options. *Presentation at the Bachelier Finance Society Congress*.
- [5] Dempster, M. A. H., and Hong, S. S. G. (2000). Spread option valuation and the fast Fourier transform. *Working paper 26*, University of Cambridge.
- [6] Hoedemakers, T. (2005). Modern Reserving Techniques for the Insurance Business.*PhD Thesis*, Katholieke Universiteit Leuven.
- [7] Turnbull, S. M. and Wakeman, L. M. (1991). A Quick Algorithm for Pricing European Average Options *Journal of Financial and Quantitative Analysis*, **26**, pp. 377-389.
- [8] Zacks, S. and Tsokos, C. P. (1978). The distribution of sums of dependent log-normal variables. *Technical Report No 31*, Defense Technical Information Center.

(Pål Nicolai Henriksen), DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO, P.O. BOX 1053, BLINDERN, N-0316 OSLO, NORWAY

E-mail address: paalnhe@math.uio.no