Towards a Bayesian framework for option pricing

Samuel W. Malone y
Enrique A. ter Horst

Estudio IESA N° 27

Versión presentada originalmente en lenguaje Latex

Hecho el depósito de ley
Depósito legal: lf23920063302632
ISBN: 980-217-311-8

Para ser publicado como *Estudio IESA* un texto tiene que ser aprobado por el Comité de Publicaciones. Las opiniones expresadas son del autor y no deben atribuirse al IESA, a sus directivos ni a Ediciones IESA. Para cualquier información sobre este estudio, favor dirigirse a Ediciones IESA, Apartado 1640, Caracas, Venezuela 1010-A. Teléfono: 58-212-555.44.52. Fax: 58-212-555-44-45. Dirección electrónica: ediesa@iesa.edu.ve.

Towards a Bayesian framework for option pricing

Samuel W. Malone* Enrique A. ter Horst†

February 15, 2006

Abstract

When using a model in continuous time finance for an asset, one can retrieve directly the likelihood function for the parameters of the asset process in order to perform statistical Bayesian analysis and estimate the posterior distributions via Gibbs sampling or other Bayesian methods. If the joint posterior distribution can be expressed in closed form, then the pricing formula for the option can be combined with the posterior to integrate out the relevant option parameters and obtain the posterior distribution for the option price. In this paper, we illustrate how to implement this procedure in a general framework, and further, we show how to introduce prior information about model parameters and use the example of geometric Brownian motion to derive the interesting result that in our framework, the martingale condition required for no-arbitrage option price is only obtained asymptotically as $t \to \infty$. when the drift has a posterior distribution and the risk-free interest rate is deterministic. We discuss that in the random case, both the drift and the risk-free interest rate have to be indistinguishable processes to rule out arbitrage.

JEL CLassification: C11, C13, G12
Keywords: Bayesian analysis; Lévy processes; Option pricing; Risk-Neutral measure; Model selection; Bayesian nonparametrics.

Acknowledgments

The authors would like to thank James O. Berger, Robert L. Wolpert , Goran Peskir, and German Molina for helpful conversations. All errors are our own.

1 Introduction

Since the seminal paper of Black and Scholes (1973), a large option pricing literature has developed around the problems implied by relaxing the simplifying assumptions of the original model, such as constant volatility, zero transactions costs, and a flat yield curve. This literature has yielded a plethora of alternative stochastic processes for the underlying, and has derived results showing how to

*Samuel W. Malone is a Ph.D. candidate at Oxford University, Oxford, United Kingdom.

†Enrique A. ter Horst is an assistant professor at the Instituto de Estudios Superiores de Administración IESA, in Caracas, DF, Venezuela.

price options on underlyings that follow such processes. Two popular alternatives to using geometric Brownian motion to model the stochastic process of the underlying are jump-diffusions and Lévy processes, which exhibit features such as skewness, kurtosis, and jumps that are also observed in price data in a wide variety of markets.

Alongside research that has focused on pricing derivatives for such processes, there has been innovation in the area of Bayesian econometrics on developing techniques for "integrating out" parameters from the risk-neutral pricing formula of the option after this formula has been derived in closed form. For an example of this technique see Eraker et al. (2000). However, no work has been done, to our knowledge, on integrating out the parameters during the transformation of the physical measure $\mathbb{P}$ to the risk neutral measure $\mathbb{Q}$, except when testing sequentially a precise hypothesis concerning the drift of a Brownian motion as in Paulo (2002), or computing Bayes factors between different models as in Polson and Roberts (1994). Using the Bayesian technique of integrating out parameters, Darsinos and Satchell (2001) derive the posterior distribution in closed-form for a European call option when the underlying follows a geometric Brownian motion, but they impose another likelihood in the computation of the price than that which is inherent in the stochastic process itself.

Thus the task of the present paper is to develop a methodology that is able to yield theoretically the posterior distribution (in closed-form as Darsinos and Satchell (2001) or numerically when that is not possible) of a call option price by integrating out the relevant parameters given their posterior distributions. These posterior distributions are constructed using the likelihood function that is implied by the underlying stochastic process used, and the prior distributions that are specified as the views of the market participant.

Although in practice, it is axiomatic that agents use their subjective beliefs as inputs into their valuation of assets, this runs counter to the famous and counterintuitive result of Black and Scholes (1973) that "risk aversion does not matter" in the determination of the fair price of a call option. This result, of course, is predicated on the crucial assumptions of continuous, zero-cost trading and an asset process (geometric Brownian motion) without jumps or other forms of market incompleteness. In contrast, when markets are incomplete, risk-aversion, and subjective beliefs, do matter.

When the underlying is modeled as a geometric Brownian motion, the equivalent [1] risk-neutral measure $\mathbb{Q}$ is unique, and the option has a unique price. When there are jumps, on the other hand, the risk-neutral measure $\mathbb{Q}$ is not unique, and thus a unique price for the option usually does not exist. Often, in the latter case, what we obtain is a range of admissible prices that still rule out arbitrage. There are ways of narrowing this range of prices; Cont and Tankov (2003) discuss how to circumvent the non-uniqueness of $\mathbb{Q}$, but their use of relative entropy methods also involves the problem we are trying to avoid, namely, the introduction of new information beyond that which is contained in the likelihood implied by the underlying stochastic process (and of course the

[1] A probability measure $\mathbb{P}$ is equivalent to another probability measure $\mathbb{Q}$ (or more compactly $\mathbb{P} \sim \mathbb{Q}$), if $\mathbb{P}$ and $\mathbb{Q}$ are mutually absolutely continuous.

prior beliefs of agents). The contribution of this paper is to show how one can link naturally Mathematical Finance with Bayesian statistics from a continuous time point of view. We show as well interesting problems concerning information asymmetry amongst agents through the use of prior distributions on the model parameters.

The outline of this paper is as follows. In section 2 we introduce Lévy processes and explain how they are used to model financial processes and price options by using the Esscher transform. The latter is but the Radon-Nikodym derivative of the risk-neutral measure $\mathbb{Q}$ with respect to the physical measure $\mathbb{P}$, which combined with a prior distribution yields a posterior. We show that integrating these parameters from the model preserves the martingale property of the Esscher transform. In section 3, we develop a Bayesian framework for option pricing together with examples from the Black & Scholes model, as well as in a diffusion case. Section 4 concludes.

2 Bayesian option pricing

The mainstream Bayesian literature has concerned itself with using state-space models as a way to get posterior distributions for derivatives perturbated around a Black & Scholes price of the following sort:

$$\log\left(\frac{S_t}{S_{t-1}}\right) = \mu + \sigma(W_t - W_{t-1}) \tag{1}$$

$$C_t = BS(\sigma, S_t) + \epsilon_t \tag{2}$$

where W_t is a Brownian motion, $\epsilon_t \sim N(0, \sigma^2)$, and $BS(\sigma, S_t)$ is the option price from the classical Black & Scholes model (see Johannes and Polson, 2002, p. 35-36 regarding this last result). Both Johannes and Polson (2002) and Darsinos and Satchell (2001) get the posterior distribution of the volatility σ from the discrete version of the continuous time process, which exists in discrete time, although is a degenerate point mass in continuous time as Polson and Roberts (1994) explain. As we saw, in the previous sections, we derived the likelihood for the Black & Scholes model that is already implied by a Geometric Brownian motion, which is a different road to that of. In order to get the posterior distribution of the theoretical Black & Scholes price, Polson and Roberts (1994) use a perturbation ϵ_t around the theoretical Black & Scholes price to construct a likelihood and proceed with a Bayesian analysis. Our framework consists of retrieving the likelihood function directly from the Radon-Nikodym process (Z_t^θ) used when performing a change of measure from the physical $\mathbb{P}$ to the risk-neutral $\mathbb{Q}$. Combining our likelihood Z_t^θ with a prior $\pi(\theta)$ enables us to derive a posterior distribution for $\pi(\theta|\{S_s : 0 \leq s \leq t\})$ given an observed price history $\{S_s : 0 \leq s \leq t\}$. Darsinos and Satchell (2001) are able to find a closed-form solution for the posterior of the call option, since their likelihood and priors exist with reference to Lebesgue measure. In general such densities do not always exist with respect to Lebesgue measure but (Z_t^θ) does. Our method can be carried out with the use of numerical simulations from the parameter's posterior distribution by standard Bayesian numerical methods.

The classsical framework of option pricing supposes a call option $C(t, S_t, \theta)$ whose payoff $h(S_t)$ depends[2] on our underlying S_t, and can be computed via the following integration:

$$C(t, S_t, \theta) \quad = \quad \int_S h(x) d\mathbb{Q}(x)$$

where integration is performed under the risk-neutral measure $\mathbb{Q}$, such that the discounted stock price $\exp(-rt) S_t$ is a $\mathbb{Q}$-martingale. General integration theory states that the following change of measure is also possible by invoking the Radon-Nikodym theorem:

$$C(t, S_t, \theta) \quad = \quad \int_S h(x) \frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)} d\mathbb{P}(x)$$

and if considering a prior $\pi(d\theta)$ that verifies some admissibility and integrability condition (see section 3.2), we can perform the following integration with respect to the prior $\pi(d\theta)$:

$$\begin{aligned}
C(t, S_t) &= \int_\Theta C(t, S_t, \theta) \pi(d\theta) \\
C(t, S_t) &= \int_\Theta \pi(d\theta) \int_S h(x) \frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)} d\mathbb{P}(x) \\
&= \int_\Theta \int_S h(x) \frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)} d\mathbb{P}(x) \pi(d\theta)
\end{aligned}$$

Although we do not always have a closed-form solution for this integral, we can compute them through classical Markov Chain Monte Carlo methods to get the marginal option price. In order to do this, we need a likelihood and a prior distribution on the model parameters in order to perform a Bayesian analysis. In the next section we show how to find this likelihood and what integrability conditions a prior must possess, as well as whether the martingale property of the Radon-Nikodym $\frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)}$ is preserved after integrating out the vector θ. The martingale property of $\frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)}$ is a crucial condition when performing a change of measure in stochastic analysis, and for option pricing as well. This is why in the next section we show that even after integrating out the parameter θ from $Z^\theta \equiv \frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)}$, the result is still a $\mathbb{P}$-martingale.

[2]For a European call option, the payoff function is equal to $\max(S_T - K, 0)$ where K is the strike price at termination date T.

3 Methodology

3.1 Change of measures and the likelihood function in option pricing

When modelling any underlying S_t of a derivative, it is common to use the following form:

$$S_t \;=\; \exp(-rt + X_t)$$

where X_t can either be a Lévy process or a diffusion, and thus S_t is the exponential[3] of either processes discounted by $-rt$. In order to price options on an underlying S_t, one needs to find an equivalent probability measure $\mathbb{Q} \sim \mathbb{P}$ such that $S_t \equiv \exp(-rt + X_t)$ is a martingale under $\mathbb{Q}$, i.e. the discounted stock price has to be a martingale under $\mathbb{Q}$. When performing a change of measure for a given stochastic process S_t under $\mathbb{P}$ to $\mathbb{Q}$, one can regard S as a random variable on the space $\Omega = \mathbb{D}[0, +\infty)$ of cádlág (Continue à droite limites à gauche) paths together with its associated filtration of measurable sets indexed by time $(\mathscr{F}_t)_{t \geq 0}$. This measure $\mathbb{P}$ is therefore defined on the space of sample paths of X, and so the Radon-Nikodym derivative $\frac{d\mathbb{Q}(x)}{d\mathbb{P}(x)}$ with respect to the reference measure $\mathbb{P}$ is the likelihood function after the process has been observed up to time t (see chapter X, section 2 Jacod and Shiryaev, 1987, on the equivalence between Radon-Nikodym derivatives and likelihood function).
Several popular methods are used to perform changes of measure, yielding likelihood functions in order to perform a Bayesian analysis. Among these methods we can cite the mininal entropy and the Esscher transform (see Cont and Tankov (2003)). The latter is named in honor of the Swedish actuary Frederik Esscher Bohman and Esscher (1963), and helps perform transformation of distribution functions, and thus of probability measures as well. This method is also known in the statistical literature as exponential tilting. For more information on the Esscher transform, see Cont and Tankov (2003).

The Esscher transform is just the Radon-Nikodym derivative of a measure with respect to another measure. In our framework, the dominating measure can either be $\mathbb{Q}$ or $\mathbb{P}$. This in turn allows us to work with the likelihood $\frac{d\mathbb{Q}}{d\mathbb{P}}$ in order to perform a Bayesian analysis, where $\mathbb{P}$ is the reference measure. (see Cont and Tankov, 2003, chapter 9, regarding the change of measure and the Esscher Transform).
The following theorem, due to Kallsen and Shiryaev (2002), provides an explicit method for computing the Esscher transform for exponential Levy processes such that the discounted stock price process S_t is a $\mathbb{Q}$-martingale.

Theorem 1 *Suppose T>0 and there exists $\theta_\star \in \mathbb{R}$ such that*

$$E_{\mathbb{P}}\{\exp(\theta_\star X_t)\} < +\infty \quad \text{and} \quad E_{\mathbb{P}}\{\exp((\theta_\star + 1)X_t)\} < +\infty \tag{3}$$

and the equation

[3] See Applebaum (2004) and Oksendal (2003) for the case when X_t is a Lévy process and a diffusion respectively.

$$k(\theta_\star + 1) - k(\theta_\star) = 0 \tag{4}$$

holds. Then

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\theta_\star X_T - k(\theta_\star)\right) \tag{5}$$

defines an equivalent martingale measure for $\{S_t\}_{0\leq t\leq T}$ under which it is a martingale. The process $\{X_t\}_{0\leq t\leq T}$ is a Lévy process under $\mathbb{Q}$ with characteristic triplet $(\mu_\star, \sigma_\star, \nu_\star)$, where

$$\mu_\star = \mu + \sigma\theta_\star + \int_{\mathbb{R}} (\exp\theta_\star x - 1)h(x)\nu(dx) \tag{6}$$

$$\sigma_\star = \sigma \tag{7}$$

$$\nu_\star = \exp\left(\theta_\star x\right)\nu(dx) \tag{8}$$

The Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is referred in the litterature as either the Esscher transform (see Gerber and Shiu (1994); Hubalek and Sgarra (2005)), exponential family in sequential analysis (see Jacod and Shiryaev (1987)), density process in probability, or likelihood function in statistics (once the process has been observed from time 0 to time t).

As we just explained, the Esscher transform is the initial building block to yield a likelihood with respect to $\mathbb{P}$ which combined with prior views through a prior distribution on θ, results in a posterior distribution for θ that can be used to incorporate the uncertainty of the model parameters in the option price, and to integrate it out.

It is important to point out that even when two market agents choose two different priors $\pi(\theta)_1$ and $\pi(\theta)_2$ on θ, their posteriors yield asymptotically the same results if both are consistent. The following two theorems are from Ghosh and Ramamoorthi (2003):

Theorem 2 *For each n, let $\pi(\theta|S_1, \ldots, S_n)$ be a posterior given $S_1, \ldots, S_n$. The sequence $\{\pi(\theta|S_1, \ldots, S_n)\}$ is said to be consistent at θ_0 if there is a $\Omega_0 \subset \Omega$ with $\mathbb{P}^\infty_{\theta_0}(\Omega_0) = 1$ such that if ω is in Ω_0 , then for every neighborhood U of θ_0,*

$$\pi(U|S_1, \ldots, S_n) \to 1 \tag{9}$$

Theorem 3 *Assume that the family $\{\mathbb{P}^\theta : \theta \in \mathbb{A}\}$ is dominated by a σ-finite measure μ and let p_θ denote the density of $\mathbb{P}^\theta$. Let θ_0 be an interior point of Ω_0, and π_1, π_2 be two priors densities with respect o a measure ν, which are positive and continuous at θ_0. Let $\pi(\theta|S_1, \ldots, S_n)_i$, $i = 1, 2$ denote the posterior densities of θ given $\{S_1, \ldots, S_n\}$. If $\pi(\theta|S_1, \ldots, S_n)_i$, $i = 1, 2$ are both consistent at θ_0 then:*

$$\lim_n \int |\pi_1(\theta|S_1, \ldots, S_n) - \pi_2(\theta|S_1, \ldots, S_n)| d\nu(\theta) = 0 \quad a.s.\ \mathbb{P}_{\theta_0} \tag{10}$$

Here $\mathbb{P}_{\theta_0}^{\infty}$ is the product probability measure defined on the space of infinite sequences $\Omega = (X^{\infty}, \mathbb{A}^{\infty})$. These last two theorems show that as time progresses, the importance of the prior distribution fades away (and thus the importance of the prior's hyperparameters as well), since we get to observe more and more data. In our framework, the dominating measure μ of Ghosh and Ramamoorthi (2003) is our $\mathbb{P}$ and their family $\mathbb{P}^{\theta}$ corresponds to our risk-neutral measure $\mathbb{Q}$.

3.2 The predictive risk-neutral measure

This section illustrates the use of the theorem by Kallsen and Shiryaev (2002) applied in a Bayesian framework. Our idea builds on the work of Paulo (2002) concerning the Bayesian treatment of integrating out all the parameters of S_t under $\mathbb{P}$ by introducing prior distributions for the drift μ. As Polson and Roberts (1994) point out, the posterior distribution for σ^2 converges to a point mass at the quadratic variation estimate as $\triangle t \to 0$. We shall thus only consider μ as our parameter and let $\theta = [\mu]$ the vector containing our parameter μ. We let $X_t \equiv (\mu - r - \frac{\sigma^2}{2})t + \sigma W_t$ and apply theorem (1).

$$
\begin{aligned}
E_{\mathbb{P}}\{\exp(\theta X_t)\} &= \exp\left(\theta(\mu - r)t + \frac{\sigma^2\theta^2 t}{2}\right) \\
E_{\mathbb{P}}\{\exp((\theta+1)X_t)\} &= \exp\left((\theta+1)(\mu - r)t + \frac{\sigma^2(\theta+1)^2 t}{2}\right)
\end{aligned}
$$

where all the assumptions of theorem (1) are met for we are dealing with Brownian motion. Solving for θ in $k(\theta_\star + 1) - k(\theta_\star) = 0$ we get:

$$
\begin{aligned}
\theta_\star &= \frac{(r + \frac{\sigma^2}{2} - \mu)}{\sigma^2} - \frac{1}{2} \\
\theta_\star &= \frac{(r - \mu)}{\sigma^2}
\end{aligned}
$$

we thus see how $\theta_\star$ is a function of $\theta \equiv [\mu,\sigma]$.
$\frac{d\mathbb{Q}|\mathscr{F}_t}{d\mathbb{P}|\mathscr{F}_t}$ is the restriction of $\frac{d\mathbb{Q}}{d\mathbb{P}}$ to $\mathscr{F}_t$, and in order to simplify notation, we shall write Z_t^{θ} for $\frac{d\mathbb{Q}|\mathscr{F}_t}{d\mathbb{P}|\mathscr{F}_t}$ interchangeably. It is worth noting that since Z_t^{θ} is a version of $\frac{d\mathbb{Q}|\mathscr{F}_t}{d\mathbb{P}|\mathscr{F}_t}$ and is a $\mathbb{P}$-martingale, we have that $\frac{1}{Z_t^{\theta}}$ is a version of $\frac{d\mathbb{P}|\mathscr{F}_t}{d\mathbb{Q}|\mathscr{F}_t}$ and is a $\mathbb{Q}$-martingale (see Protter, 1990, p. 101-102, for a proof and explanation).
Perform the following change of measure (via the Esscher transform) for a given $\theta_\star$ according to theorem 1:

$$
Z_t^{\theta_\star} = \exp(\theta_\star X_t - k(\theta_\star)) \tag{11}
$$

We can show that $Z_t^{\theta_\star}$ is a $\mathscr{F}_t$-martingale by Itô's lemma under $\mathbb{Q}$ and also that $E(Z_t^{\theta_\star}) = 1$. These last properties enable us to define $\forall t$ Z_t as:

$$Z_t = \int_{\mathbb{A}} Z_t^{\theta_\star} \pi(\theta_\star) d\theta_\star \tag{12}$$

where $\mathbb{A} \subset \Omega = \mathbb{D}[0, +\infty)$ is a measurable set. We define Z_t as *the marginal risk neutral density process*. We choose $\pi(\theta)$ such that $E_{\mathbb{P}}\{Z_t\} < +\infty$ as in Paulo (2002) and show that:

$$\begin{aligned} E_{\mathbb{P}}[Z_t|\mathscr{F}_s] &= E_{\mathbb{P}}[\int_{\mathbb{A}} Z_t^\theta \pi(\theta) d\theta|\mathscr{F}_s] && (13)\\ &= \int_{\mathbb{A}} E_{\mathbb{P}}[Z_t^\theta|\mathscr{F}_s]\pi(\theta)d\theta \qquad \text{(by Fubini's)} && (14)\\ &= \int_{\mathbb{A}} Z_s^\theta \pi(\theta) d\theta \qquad \text{(because } Z_s^\theta \text{ is a } \mathbb{P}\text{-martingale)} && (15)\\ &\equiv Z_s \qquad \text{(thus } Z_t \text{ is an } \mathscr{F}_t\text{-martingale)} && (16)\end{aligned}$$

Since Z_t is a $\mathbb{P}$-martingale and $E_{\mathbb{P}} Z_t = 1 \forall t$, we can then perform the following change of measure: $d\mathbb{Q} = Z_t d\mathbb{P}$.

4 Some illustrations

4.1 Posterior example under the Black & Scholes model

In the classic paper by Black and Scholes (1973), the stock price S_t is solution to the following SDE:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{17}$$

and the solution S_t is equal to:

$$S_t = S_0 \exp\left[(\mu - \frac{\sigma^2}{2})t + \sigma W_t\right] \tag{18}$$

Working with the discounted stock price $S_t \equiv \exp(-rt) S_t$:

$$S_t = S_0 \exp\left[(\mu - r - \frac{\sigma^2}{2})t + \sigma W_t\right] \tag{19}$$

enables us to determine the deterministic risk-neutral condition $\mu = r$ which is the one that determines the unique martingale measure. We then get that S_t is a martingale under $\mathbb{Q}$[4] and is equal to:

$$S_t = S_0 \exp\left[(r - \frac{\sigma^2}{2})t + \sigma W_t\right] \tag{20}$$

[4] S_t is a $\mathbb{Q}$-martingale, which is equivalent to showing that $Z_t^\theta S_t$ is a $\mathbb{P}$-martingale. See appendix for a proof.

As we shall see, this martingale condition holds asymptotically in our Bayesian framework as t tends to infinity. Doing the same algebra as the example from section 3.2, we find that the θ that solves the following equation:

$$k(\theta+1) - k(\theta) = 0 \tag{21}$$

yields $\theta = \frac{(r+\frac{\sigma^2}{2}-\mu)}{\sigma^2} - \frac{1}{2}$, where:

$$\begin{aligned} k(\theta+1) &= (\theta+1)[\mu - r - \frac{\sigma^2}{2}]t + \frac{\sigma^2(\theta+1)^2 t}{2} \\ k(\theta) &= \theta[\mu - r - \frac{\sigma^2}{2}]t + \frac{\sigma^2\theta^2 t}{2} \end{aligned}$$

we get the following posterior for μ

$$[\mu|\{S_s : 0 \leq s \leq t\}] \quad \sim \quad N\left[r - \frac{\sigma W_t}{t}, \frac{\sigma^2}{t}\right] \tag{22}$$

(please see the appendix for a derivation of the posterior). One important observation is that as $t \to +\infty$ we get $\frac{\sigma^2}{t} \to 0$ and thus $\mu \to \theta_0 = r$ since $\frac{\sigma W_t}{t} \to 0$ as $t \to +\infty$. By theorem 2, we confirm that the derived posterior for μ is consistent. We get the martingale condition $\mu = r$ asymptotically as Black and Scholes (1973). For a given t, μ is a normally distributed random variable and the probability that $\mu = r$ is zero. In the case where both μ and r can be stochastic, then the above equality has to be understood in an almost sure for every t, thus making both processes indistinguishable (see Protter, 1990, p. 4, for for a definition). The elicitation of a prior on μ imposes a probability distribution on r as well as the other way around. Defining a stochastic process for r imposes an objective prior on μ . Additionally, the risk neutral measure $\mathbb{Q}$ can be retrieved by using the following Esscher transform:

$$Z_t^\theta \quad = \quad \exp\left[X_t\left(\frac{r-\mu}{\sigma}\right) - \frac{(r^2-\mu^2)t}{2\sigma^2} + \frac{(r-\mu)t}{2}\right] \tag{23}$$

and

$$X_t \quad = \quad (\mu - \frac{\sigma^2}{2})t + \sigma W_t \text{ under } \mathbb{P} \tag{24}$$

thus

$$Z_t^\theta \quad = \quad \exp\left[W_t\left(\frac{r-\mu}{\sigma}\right) - \frac{t}{2}\left(\frac{r-\mu}{\sigma}\right)^2\right] \tag{25}$$

(see Cont and Tankov, 2003, chapter 9, regarding this last result). The use of a Bayesian framework for σ^2 is dealt by Polson and Roberts (1994) where it is shown that its posterior converges to a point mass at the quadratic variation estimate. This last result for σ^2 is not true when dealing with a jump-diffusion.

4.2 Posterior example under the diffusion case

When considering processes with time dependent drifts and volatilities ($\mu(t, X_t)$ and $\sigma(t, X_t)$), we are not considering anymore necessarily solutions of SDE's that are Lévy processes anymore. This is why, in order to derive the likelihood of such processes, we need to invoke Girsanov's theorem which gives us an *extended* Esscher transform which can incorporate time dependent drifts and volatilities. We shall consider a market that has two securities:

$$\begin{aligned} dX_t &= r(t,w)X_t dt && (26) \\ dS_t &= \mu(t,w)S_t dt + \sigma(t,w)S_t dW_t && (27) \\ & && (28) \end{aligned}$$

where $r(t,w)$, $\mu(t,w)$, and $\sigma(t,w)$ are stochastic processes that have to verify some conditions (see Oksendal (2003)). We construct the likelihood of the stochastic process S_t with respect to the dominating measure $\mathbb{Q}$ induced by:

$$dS_t = r(t,w)S_t dt + \sigma(t,w)S_t dW_t. \quad (29)$$

After observing the process from time 0 to t, the likelihood is then given by ($Z_t^\theta = \frac{d\mathbb{Q}}{d\mathbb{P}}$):

$$Z_t^\theta = \exp\left[\int_0^t \left(\frac{r(s,w) - \mu(s,w)}{\sigma(s,w)}\right)\frac{dS_s}{S_s} - \frac{1}{2}\int_0^t \left(\frac{r(s,w)^2 - \mu(s,w)^2}{\sigma(s,w)^2}\right) ds\right]$$

and if $\mu(t,w) = \mu$, $r(s,w) = r$ and $\sigma(s,w) = \sigma$ we get:

$$\begin{aligned} Z_t^\theta &= \exp\left[\left(\frac{r-\mu}{\sigma}\right)\int_0^t \frac{dS_s}{S_s} - \frac{1}{2}\left(\frac{r^2-\mu^2}{\sigma^2}\right)\int_0^t ds\right] \\ \int_0^t \frac{dS_s}{S_s} &= \log\frac{S_t}{S_0} + \frac{\sigma^2 t}{2} \end{aligned}$$

since we are looking for ($\frac{1}{Z_t^\theta} = \frac{d\mathbb{P}}{d\mathbb{Q}}$), we get after some simplifications:

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \exp\left[-\frac{\log\frac{S_t}{S_0}(r-\mu)}{\sigma^2} - \frac{(r-\mu)t}{2} - \frac{(\mu^2 - r^2)t}{2\sigma^2}\right]$$

and after combining squares we get the following posterior for μ: $[\mu \mid \{S_s : 0 \le s \le t\}] \sim N\left(\frac{\log\frac{S_t}{S_0}}{t} + \frac{\sigma^2}{2}, \frac{\sigma^2}{t}\right)$. In a same fashion as Polson and Roberts (1994), we can perform a Bayes factor analysis to perform model selection among competing models for the underlying using their methodology but with a different reference measure than theirs.

5 Discussion and Conclusions

When traders and market participants use pricing formulas for derivatives, the price is a function of some kind of parameter in some way. Ad-hoc methods are used such as historical estimates, or plain averages in order to plug those values back into the formula. By deriving the likelihood through various methods using the Esscher transform or Girsanov's theorem for diffusions, one is able to make inference on the drift of the underlying and other parameters depending on the mathematical nature of the latter.

When introducing a prior for the drift of a Geometric Brownian motion, reflecting the agent's beliefs/predictions, one gets that all agents have the same information about the drift since $\mu = r$ asymptotically as $t \to +\infty$. We have shown that this results holds only when μ is not random. Once we introduce a prior distribution on μ, we need to use the notion of indistinguishability between μ and r. The prior distribution for μ induces a probabilistical model for r, in the same way that a model for r induces a prior for μ through the martingale condition $\mu = r$.

An interesting point to consider would be to extend the the work of Polson and Roberts (1994) for computing Bayes factors for general Lévy processes in order to choose the best risk-neutral model that fits the data under the physical measure, as well as performing the same analysis but for Lévy processes with discontinuous sample paths. Also, as a final note, it is noteworthy to notice that when market participants introduce prior information concerning model parameters, the risk neutrality can only be achieved asymptotically, since μ becomes uncertain, while r is fixed in the BS model. An interesting feature is that in the case of jump-diffusion models, the posterior of the volatility is not degenerate at the quadratic variation of the process for it is equal to the diffusion component and the sum of its jumps (Protter (1990)). Future research concerning the posterior of the jump intensity measure (Lévy measure) for jump-diffusion models should be considered, but is still an ongoing research topic in the statistical literature (see Lancelot (2005)).

The framework developed in this paper shows how to get the marginal option price by finding the likelihood and using appropriate prior distributions on the model parameters. Promising results of implied volatility forecasts by Darsinos and Satchell (2001) might pave the way for exploring further the topic of applied Bayesian option pricing methods.

Appendix

Proof that S_t is a $\mathbb{Q}$-martingale when S_t is a Geometric Brownian motion

To show that S_t is an $\mathbb{Q}$-martingale is equivalent to show that $Z_t^\theta S_t$ is a $\mathbb{P}$-martingale and $\theta \equiv \frac{r-\mu}{\sigma^2} - \frac{1}{2}$. Recalling that:

$$Z_t^\theta S_t = \exp(Y_t)$$

where $Y_t \equiv (\theta+1)X_t - \theta(r-\mu)t - \frac{\theta^2\sigma^2 t}{2}$ and $X_t \equiv (\mu - r)t + \sigma W_t$ and W_t is a standard Brownian motion. By Itô's lemma we get:

$$\begin{aligned}
\exp(Y_t) - 1 &= \int_0^t \exp(Y_s)dY_s + \frac{1}{2}\int_0^t \exp(Y_s)d[Y_s, Y_s] \\
&= \int_0^t \exp(Y_s)dY_s + \frac{1}{2}\int_0^t \exp(Y_s)(\theta+1)^2\sigma^2 ds
\end{aligned}$$

and

$$dY_s = (\theta+1)(\mu-r)dt + (\theta+1)\sigma dW_t - \theta(\mu-r)ds - \frac{\theta^2\sigma^2 dt}{2}$$

and we get:

$$\exp(Y_t) - 1 = \int_0^t \exp(Y_s)(\theta+1)\sigma dW_s$$

which is a $\mathbb{P}$-martingale.

Posterior for μ using the Esscher transform

We found that the optimal θ making $S_t = \exp(X_t)$ into a $\mathbb{Q}$-martingale is equal to $\theta = \frac{(r+\frac{\sigma^2}{2}-\mu)}{\sigma^2} - \frac{1}{2}$ where:

$$X_t \equiv (\mu - r - \frac{\sigma^2}{2})t + \sigma W_t$$

and

$$
\begin{aligned}
\theta X_t &= -\frac{(\mu - r - \frac{\sigma^2}{2})^2 t}{\sigma^2} - \frac{(\mu - r - \frac{\sigma^2}{2})t}{2} + \sigma W_t \frac{(r - \mu + \frac{\sigma^2}{2})}{\sigma^2} - \frac{\sigma W_t}{2} \\
-\theta(\mu - r - \frac{\sigma^2}{2})t &= \frac{(\mu - r - \frac{\sigma^2}{2})^2 t}{\sigma^2} + \frac{(\mu - r - \frac{\sigma^2}{2})t}{2} \\
-\frac{\sigma^2\theta^2 t}{2} &= -\frac{(\mu - r - \frac{\sigma^2}{2})^2 t}{2\sigma^2} + \frac{(r + \frac{\sigma^2}{2} - \mu)t}{2} - \frac{\sigma^2 t}{8} \\
\theta^2 &= [\frac{(r + \frac{\sigma^2}{2} - \mu)}{\sigma^2} - \frac{1}{2}]^2 \\
&= \frac{(r + \frac{\sigma^2}{2} - \mu)^2}{\sigma^4} - \frac{(r + \frac{\sigma^2}{2} - \mu)}{\sigma^2} + \frac{1}{4}
\end{aligned}
$$

and the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ being equal to:

$$
\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\theta X_T - k(\theta)T\right)
$$

is thus proportional to (after subsitutions from above):

$$
\frac{d\mathbb{Q}}{d\mathbb{P}} \propto \exp\left(-\frac{(\mu - r - \frac{\sigma^2}{2})^2 t}{2\sigma^2} + \frac{(r + \frac{\sigma^2}{2} - \mu)t}{2} - \frac{\sigma^2 t}{8} + \sigma W_t \frac{(r - \mu + \frac{\sigma^2}{2})}{\sigma^2} - \frac{\sigma W_t}{2}\right)
$$

where we have:

$$
\exp\left(-\frac{(\mu - r - \frac{\sigma^2}{2})^2 t}{2\sigma^2} + \frac{(r + \frac{\sigma^2}{2} - \mu)t}{2} - \frac{\sigma^2 t}{8} + \sigma W_t \frac{(r - \mu + \frac{\sigma^2}{2})}{\sigma^2} - \frac{\sigma W_t}{2}\right) = \exp\left(W_t\theta - \frac{\sigma^2\theta^2 t}{2}\right)
$$

where $\theta = \left[\frac{(r-\mu+\frac{\sigma^2}{2})}{\sigma} - \frac{\sigma}{2}\right]$
Keeping only expressions depending on μ:

$$
\begin{aligned}
\frac{d\mathbb{Q}}{d\mathbb{P}} &\propto \exp\left(-\frac{(\mu - r - \frac{\sigma^2}{2})^2 t}{2\sigma^2} + \frac{(r + \frac{\sigma^2}{2} - \mu)t}{2} + \sigma W_t \frac{(r - \mu + \frac{\sigma^2}{2})}{\sigma^2}\right) \\
&\propto \exp\left(-\frac{t}{2\sigma^2}[(r + \frac{\sigma^2}{2} - \mu)^2 - 2(r + \frac{\sigma^2}{2} - \mu)\frac{1}{2}(\sigma^2 + \frac{2\sigma W_t}{t})]\right) \\
&\propto \exp\left(-\frac{t}{2\sigma^2}[r + \frac{\sigma^2}{2} - \mu - \frac{\sigma^2}{2} - \frac{\sigma W_t}{t}]^2\right) \\
&\propto \exp\left(-\frac{t}{2\sigma^2}[\mu - r + \frac{\sigma W_t}{t}]^2\right) \\
&\propto \exp\left(-\frac{t}{2\sigma^2}[\mu - (r - \frac{\sigma W_t}{t})]^2\right)
\end{aligned}
$$

and we conclude that the posterior for μ is:
$N\left[r - \frac{\sigma W_t}{t}, \frac{\sigma^2}{t}\right]$

References

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus.* Cambridge University Press.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* **81** 637–654.

Bohman, H. and Esscher, F. (1963). Studies in risk theory with numerical illustrations concerning distribution functions and stop loss premiums i,ii. *Skand. Aktuarietidskr.* 173–225.

Cont, R. and Tankov, P. (2003). *Financial Modelling with Jump Processes.* Chapman & Hall.

Darsinos, T. and Satchell, S. E. (2001). Bayesian Analysis of the Black-Scholes Option Price. Tech. rep., Cambridge Working Papers in Economics.

Eraker, B., Johannes, M. and Polson, N. (2000). The imact of jumps in volatility and returns. *Journal of Business* **75** 305–332.

Gerber, H. and Shiu, E. (1994). Option pricing by Esscher transforms. *Trans. Soc. Actuar.* **XLVI** 98–140.

Ghosh, J. and Ramamoorthi, R. (2003). *Bayesian Nonparametrics.* Springer Verlag.

Hubalek, F. and Sgarra, C. (2005). Esscher transforms and the minimal entropy martingale measure for exponential Lévy models. *Journal of Quantitative Finance (Forthcoming)* .

Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag.

Johannes, M. and Polson, N. (2002). MCMC methods for financial econometrics.

Kallsen, J. and Shiryaev, A. (2002). The cumulant process and the esscher's change of measure. *Finance Stoch.* **6** 397–428.

Lancelot, F. J. (2005). Bayesian Poisson Porcess Partition Calculus with an application to Bayesian Lévy Moving Averages. *Annals of Statistics* **33** 1771–1799.

Oksendal, B. (2003). *Stochastic Differential Equations : An Introduction with Applications.* Springer Verlag.

Paulo, R. (2002). *Problems on the Bayesian/Frequentist Interface.* Ph.D. thesis, Duke University, ISDS.

Polson, N. and Roberts, G. (1994). Bayes factors for discrete observations from diffusion processes. *Biometrika* **81** 11–26.

Protter, P. (1990). *Stochastic Integration and Differential Equations*. Springer-Verlag.