

## Research Article

# The Use of Statistical Tests to Calibrate the Black-Scholes Asset Dynamics Model Applied to Pricing Options with Uncertain Volatility

Lorella Fatone,<sup>1</sup> Francesca Mariani,<sup>2</sup>  
Maria Cristina Recchioni,<sup>3</sup> and Francesco Zirilli<sup>4</sup>

<sup>1</sup> Dipartimento di Matematica e Informatica, Università di Camerino, Via Madonna delle Carceri 9, 62032 Camerino, Italy

<sup>2</sup> CERI-Centro di Ricerca "Previsione, Prevenzione e Controllo dei Rischi Geologici", Università di Roma "La Sapienza", Palazzo Doria Pamphilj, Piazza Umberto Pilozi 9, Valmontone 00038 Roma, Italy

<sup>3</sup> Dipartimento di Scienze Sociali "D. Serrani", Università Politecnica delle Marche, Piazza Martelli 8, 60121 Ancona, Italy

<sup>4</sup> Dipartimento di Matematica "G. Castelnuovo", Università di Roma "La Sapienza", Piazzale Aldo Moro 2, 00185 Roma, Italy

Correspondence should be addressed to Francesco Zirilli, f.zirilli@caspur.it

Received 28 October 2011; Revised 28 February 2012; Accepted 13 March 2012

Academic Editor: A. Thavaneswaran

Copyright © 2012 Lorella Fatone et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new method for calibrating the Black-Scholes asset price dynamics model is proposed. The data used to test the calibration problem included observations of asset prices over a finite set of (known) equispaced discrete time values. Statistical tests were used to estimate the statistical significance of the two parameters of the Black-Scholes model: the volatility and the drift. The effects of these estimates on the option pricing problem were investigated. In particular, the pricing of an option with uncertain volatility in the Black-Scholes framework was revisited, and a statistical significance was associated with the price intervals determined using the Black-Scholes-Barenblatt equations. Numerical experiments involving synthetic and real data were presented. The real data considered were the daily closing values of the S&P500 index and the associated European call and put option prices in the year 2005. The method proposed here for calibrating the Black-Scholes dynamics model could be extended to other science and engineering models that may be expressed in terms of stochastic dynamical systems.

## 1. Introduction

The Black-Scholes formulae [1] used to price European call and put options are based on an asset price dynamics model. This model is a stochastic dynamical system that may

be written as a stochastic differential equation. The solution to this model takes the form of a stochastic process called geometric Brownian motion. The model contains two real parameters: volatility and drift. The volatility and drift parameter values are necessary to apply the model to asset and option price forecasting; therefore, in practice, the values of these parameters must be determined prior to using the Black-Scholes dynamics model and the option pricing formulae derived from it. The problem of estimating the asset price dynamics model parameters must be considered based on the available data. This problem is a calibration problem and is an inverse problem for a stochastic dynamical system defined by a stochastic differential equation. The estimated parameter values obtained by solving the calibration problem may be used to forecast asset prices at future time points and to evaluate the option pricing formulae. The “accuracy” and reliability of the estimated parameter values determine the accuracy and reliability of the forecasted asset prices and the computed and/or forecasted option prices.

Note that in recent years, the validity of the asset price dynamics model proposed by Black and Scholes in 1973 [1] has been disputed in the mathematical finance literature, and several other refined models, such as the Heston model [2], have been introduced to describe asset price dynamics. Nevertheless, the Black-Scholes asset dynamics model and, particularly, the option pricing formulae derived from it remain widely used in financial market practice. We will see that the solution to the Black-Scholes model calibration problem, using statistical tests, and an investigation of effects on the option pricing problem, are easily derived using elementary mathematics. For this reason, the Black-Scholes asset dynamics model is the natural choice to begin our study of solutions to the problems associated with calibrating stochastic dynamical systems using statistical tests. The ideas introduced in this paper are rather general and are not limited to the study of the Black-Scholes model.

The data used in the calibration problem include observations of the asset prices over a finite set of (known) equispaced discrete time values. We show how elementary statistical tests (i.e., the Student’s  $t$  and the  $\chi^2$  tests) may be used to estimate the drift and volatility parameters of the Black-Scholes model with statistical significance. Recall that the first step in formulating a hypothesis testing problem consists of defining the null hypothesis,  $H_0$ , as well as an alternative hypothesis,  $H_1$ . When the goal is to establish an assertion about a probability distribution parameter based on support from a data set, the assertion is usually taken to be the null hypothesis  $H_0$ , and the negation of the assertion itself is taken to be the alternative hypothesis  $H_1$  (or vice versa). The null hypothesis  $H_0$  considered later is the hypothesis that a parameter of a probability distribution belongs to a given interval. In a statistical test, we must additionally consider the statistical significance. The statistical significance  $\alpha$ ,  $0 < \alpha < 1$ , is the maximum probability of rejecting the null hypothesis  $H_0$  when the hypothesis is true. Defining a type I error as the error associated with rejecting the null hypothesis  $H_0$  when  $H_0$  is true, the statistical significance  $\alpha$ ,  $0 < \alpha < 1$ , is the maximum probability of making a type I error. The result of the test is a decision to accept or reject the null hypothesis  $H_0$  with a significance level  $\alpha$ ,  $0 < \alpha < 1$ .

Let us describe the content of the paper in greater detail. As mentioned, the solution to the Black-Scholes asset price dynamics model is a stochastic process called geometric Brownian motion, which depends on two parameters: the drift and the volatility. From this fact, it follows that the asset price at any given time may be modeled as a random variable with a log-normal distribution and, therefore, the log return of the asset price is normally distributed. It is easy to see that the asset price log return increments associated with observations of the asset price over a (finite) set of equispaced time values comprise a sequence of values sampled from a set of independent identically distributed Gaussian

random variables. The mean and variance of these Gaussian random variables can be expressed using elementary formulae as a function of the drift and volatility parameters of the Black-Scholes asset price dynamics model and of the time interval between observations. Starting from a sample of observations of the log return increments over a finite set of equispaced (known) time values, elementary statistical tests (i.e., the Student's  $t$  and  $\chi^2$  tests) may be used to estimate, with a given (statistical) significance, the mean and variance of the Gaussian random variables associated with the log return increments. In fact, the Student's  $t$ -test [3, 4] is the test used to establish, with statistical significance, whether the mean of a normally distributed population of independent samples has a certain value or belongs to an interval specified in a "null" hypothesis. Similarly, the  $\chi^2$  test [4, 5] is the test used to establish, with statistical significance, whether the variance of a normally distributed population of independent samples has a value or belongs to an interval specified in a "null" hypothesis. Knowledge of the Gaussian random variable parameters associated with the log return increments derived from the statistical tests permits recovery of the corresponding parameters of the Black-Scholes model and the associated statistical significances.

The use of statistical tests to solve the problem associated with calibrating stochastic dynamical systems, such as the Black-Scholes model, is an interesting approach to inverse problems used elsewhere in mathematical finance as well as in application contexts other than mathematical finance. The significance levels obtained for the parameter values (intervals) in the Black-Scholes model are relevant in many practical situations. In Section 3, we consider the problem associated with pricing an option with an uncertain volatility. This problem has been considered by several authors in the scientific literature. For simplicity, we refer the reader to [6, 7], in which the (uncertain) volatility in the Black-Scholes framework is assumed to belong to a known interval and the corresponding price intervals for the (European vanilla) option prices may be determined using the Black-Scholes-Barenblatt equations. Thanks to our methodology, statistical significance levels may be attributed to the option price intervals determined using the Black-Scholes-Barenblatt equations.

Finally, the approach used to calibrate the Black-Scholes model is applied to the study of synthetic and real data. The synthetic data considered were generated by numerically integrating the stochastic differential equation that defines the asset price dynamics in the Black-Scholes model, for several choices of the parameter values. The real data studied were the time series of the daily closing values of the S&P500 index and the associated European vanilla option prices during the year 2005. The numerical results obtained were computationally simple and statistically convincing.

In mathematical finance, the problem associated with estimating the volatility of asset prices, starting from a time series of observed data (asset and/or option prices), has received significant attention. The methods most commonly used in the literature include the implied and historical (or realized) volatility methods. The approach proposed here for solving this problem is distinct. Unlike the implied volatility method, we do not consider the volatility implied by the option prices observed in the market. That is, we do not estimate the volatility parameter of the asset price using the prices of various options to the asset with different strikes and expiration dates. By contrast, our method estimates the volatility based on the asset prices observed over a finite set of discrete time values, similar to the approach used in the historical volatility method. Our method improves on the historical volatility method by associating a significance level to the volatility estimate. In some sense, the transition from the historical volatility method to the method proposed here corresponds to a transition from the use of a volatility estimate for "sample volatility" to the use of a statistical test (i.e., the  $\chi^2$  test) tailored to the random variable which depends on the parameter to be estimated.

Roughly speaking, the transition is from a method of descriptive statistics to a method of mathematical statistics.

The work presented here has several valuable features that are worth noting. First, the idea proposed here is very simple. Although the approach is introduced in the context of the Black-Scholes model, it can be applied to other stochastic dynamical systems used in mathematical finance to describe asset price dynamics, for example, the Heston model (see [2, 8, 9]), or some of its variants, such as the models introduced in [10–13] to study specific problems. In [8–13], the models were calibrated using the “implied volatility method” or the maximum likelihood method. These studies did not implement statistical tests to solve the calibration problems or estimate the parameters; therefore, no statistical significance could be associated with the parameter estimates. Note that in calibrating these models or generic stochastic dynamical system models in general, no elementary statistical tests (such as the Student’s  $t$  or  $\chi^2$  tests) could be used. Extending the method suggested here to studies of more general dynamical systems will depend on the development of new ad hoc statistical tests. The application of these new statistical tests to a sample data will most likely require substantial use of numerical methods.

The website <http://www.econ.univpm.it/recchioni/finance/w11/> makes available auxiliary material, including animations, to assist the reader in understanding the discussion present here. References to the authors’ more general studies in the field of mathematical finance are available at the website <http://www.econ.univpm.it/recchioni/finance/>.

The remainder of the paper is organized as follows. Section 2 formulates and solves the calibration problem for the Black-Scholes model. Section 3 addresses the problem of pricing options with statistical significance when the volatility value is uncertain and lies in a specified range. Section 4 tests the proposed method on time series data, and the results obtained from studies of the synthetic and real data are discussed. Section 5 presents some concluding remarks and a few possible extensions of this work.

## 2. The Calibration Problem for the Black-Scholes Model

Let  $S_t > 0$  denote the asset price at time  $t \geq 0$ . The Black-Scholes model [1] assumes that  $S_t$ ,  $t > 0$ , is a stochastic process, the dynamics of which are governed by the following stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad t > 0, \quad (2.1)$$

with the initial condition:

$$S_0 = \widehat{S}_0, \quad (2.2)$$

where  $\mu$  and  $\sigma$  are real parameters,  $\mu$  is the drift,  $\sigma > 0$  is the volatility,  $W_t$ ,  $t > 0$ , is a standard Wiener process,  $W_0 = 0$ ,  $dW_t$ ,  $t > 0$  is its stochastic differential, and the initial condition  $\widehat{S}_0 > 0$  is a given random variable. For simplicity, we assume that  $\widehat{S}_0$  is a random variable concentrated at a point with probability one. Abusing the notation, we denote this point as  $\widehat{S}_0 > 0$ . The parameters  $\mu$  and  $\sigma$  are the unknowns of the calibration problem.

The stochastic differential (2.1) defines the so-called geometric Brownian motion. In fact, (2.1) can be rewritten as

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \quad t > 0. \quad (2.3)$$

Let  $\ln(\cdot)$  denote the logarithm of  $\cdot$  for  $t \geq 0$ . The quantity  $G_t = \ln(S_t/\hat{S}_0)$  is the log return at time  $t$  of the asset with a price  $S_t$ . Using (2.3) and Itô's Lemma (see [14]), it follows that the process  $G_t$ ,  $t > 0$  satisfies the following stochastic differential equation:

$$dG_t = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t, \quad t > 0, \quad (2.4)$$

with initial conditions that follow from (2.2), that is,  $G_0 = 0$ . Equation (2.4) implies that  $G_t = \ln(S_t/\hat{S}_0)$ ,  $t > 0$ , is a generalized Wiener process with a constant drift  $\mu - \sigma^2/2$  and a constant volatility  $\sigma > 0$ . Therefore, for  $t \geq 0$ ,  $\tau > 0$ , the increment in  $G_t = \ln(S_t/\hat{S}_0)$  occurring between time  $t$  and time  $t + \tau$  is a Gaussian random variable with mean  $(\mu - \sigma^2/2)\tau$  and variance  $\sigma^2\tau$ . That is

$$G_{t+\tau} - G_t = \ln S_{t+\tau} - \ln S_t \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)\tau, \sigma^2\tau\right), \quad t \geq 0, \tau > 0, \quad (2.5)$$

where, for  $M$  and  $V$  real constants,  $V \neq 0$ ,  $\mathcal{N}(M, V^2)$  denotes the Gaussian distribution with mean  $M$  and variance  $V^2$ .

From (2.5), it follows that the Black-Scholes asset price log return increments associated with a discrete set of equispaced time values form a sequence of independent identically distributed Gaussian random variables. Let  $\Delta t > 0$  be a time increment, and let  $t_i = i\Delta t$ ,  $i = 0, 1, \dots, n$ , be a discrete set of equispaced time values that later will be chosen as the set of observation times. We define  $X_{t_i}$ , the asset price log return increment as  $t$  increases from  $t_{i-1}$  to  $t_i$ , as

$$X_{t_i} = \ln\left(\frac{S_{t_i}}{S_{t_{i-1}}}\right), \quad i = 1, 2, \dots, n. \quad (2.6)$$

It is easy to see that the random variables  $X_{t_i}$ ,  $i = 1, 2, \dots, n$  are independent identically distributed Gaussian random variables with mean  $M$  and variance  $V^2$ , where

$$M = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t, \quad V^2 = \sigma^2\Delta t. \quad (2.7)$$

We then have

$$X_{t_i} \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)\Delta t, \sigma^2\Delta t\right), \quad i = 1, 2, \dots, n. \quad (2.8)$$

Consider the following calibration problem: given a time increment  $\Delta t > 0$ , a statistical significance level  $\alpha$ ,  $0 < \alpha < 1$ , and the asset price  $\widehat{S}_i$  observed at time  $t = t_i = i\Delta t$ ,  $i = 0, 1, \dots, n$ , determine two intervals in which the parameters  $\mu$  and  $\sigma$  of the model (2.1), respectively, fall with the given significance level  $\alpha$ .

Note that in order to apply the calibration problem solution to the problem of pricing options with uncertain volatility (see Section 3), point estimates of the tested parameters are not useful. Instead, an interval of variability for the tested parameter (i.e., with a given significance level  $\alpha$ ) is necessary. Studies of the problem presented in Section 3, in particular, require the identification of an interval of variability for the volatility  $\sigma$ . That is, we would like to determine the interval within which the parameter  $\sigma$  may be found, and to this interval we would like to associate a significance level  $\alpha$ . The option prices in the Black-Scholes model are independent of the drift parameter  $\mu$  and depend on the risk-free interest rate. Usually, the risk-free interest rate is known for a given problem; however, in the numerical experiments presented in Section 4, a risk-free interest rate was selected, depending on the interval of variability of  $\mu$  determined in the calibration problem. For more details see Section 4.

The observations  $\widehat{S}_i$  of the asset price at time  $t = t_i$ ,  $i = 0, 1, \dots, n$ , are nonnegative real numbers that are assumed to be unaffected by errors. These  $n + 1$  observations of the asset price and the corresponding time values are the data used in the calibration problem. The corresponding observed log return increments  $\widehat{x}_i = \ln(\widehat{S}_i/\widehat{S}_{i-1})$ ,  $i = 1, 2, \dots, n$ , are a sample of  $n$  observations taken, respectively, from the random variables  $X_{t_i}$ ,  $i = 1, 2, \dots, n$ , that is, taken from a set of independent identically distributed Gaussian random variables. Using this sample data, the Student's  $t$ -test and  $\chi^2$  test (see [3–5]), we identified two intervals over which the mean  $M$  and variance  $V^2$  of the Gaussian random variables were determined to within a given significance level  $\alpha$ . From knowledge of the intervals determined for  $M$  and  $V^2$ , the corresponding intervals for  $\mu$  and  $\sigma$  could be recovered by inverting the relations (2.7).

In greater detail, given a significance level  $\alpha$ ,  $0 < \alpha < 1$ , we can perform statistical tests on the variance  $V^2$  and on the mean  $M$  of the random variables  $X_{t_i} = \ln(S_{t_i}/S_{t_{i-1}})$ ,  $i = 1, 2, \dots, n$  starting from the sample data  $\widehat{x}_i = \ln(\widehat{S}_i/\widehat{S}_{i-1})$ ,  $i = 1, 2, \dots, n$ , using the  $\chi^2$  and Student's  $t$  tests, respectively. This implies that for a given  $\alpha$ ,  $0 < \alpha < 1$ , and the relations (2.7), we can accept or reject, with a significance level  $\alpha$ , the following hypotheses regarding the parameters of the Black-Scholes asset price dynamics model:

$$\sigma_1 \leq \sigma \leq \sigma_2, \quad (2.9)$$

and

$$\mu_1 \leq \mu \leq \mu_2, \quad (2.10)$$

where

$$\sigma_i = \frac{V_i}{\sqrt{\Delta t}}, \text{ and } \mu_i = \frac{M_i}{\Delta t} + \frac{V_i^2}{2\Delta t}, \quad i = 1, 2, \quad (2.11)$$

and  $M_1 < M_2$ ,  $0 < V_1 < V_2$  are the quantities that define the corresponding hypotheses on  $M$  and  $V$ , respectively. This is performed simply by translating the results on  $V^2$  and  $M$  obtained

using the statistical tests to  $\sigma$  and  $\mu$ . We can proceed as follows. Given a sample data,

$$\hat{x}_i \text{ sampled from } X_{t_i} \sim \mathcal{N}(M, V^2), \quad i = 1, 2, \dots, n, \quad (2.12)$$

with  $M$  and  $V^2$  defined in (2.7), it is possible first to estimate  $V^2$  (and, therefore,  $\sigma^2$ ) using the  $\chi^2$  test along with (2.11), and, subsequently, to estimate  $M$  (and, therefore,  $\mu$ , thanks to knowledge of  $\sigma^2$  acquired using the  $\chi^2$  test) using the Student's  $t$ -test along with (2.11).

It should be noted that in many circumstances, it is more practical to try to determine an interval of variability for the drift  $\mu$  and volatility  $\sigma$  parameters in the Black-Scholes model than to try to determine their "exact" values.

To estimate  $V^2$  from the log return increments (2.12), we must choose the hypotheses that we would like to test. This can be done in many ways. The analysis of data time series in Section 4 uses the following procedure.

*Procedure 1.* Given the sample data  $\hat{x}_i, i = 1, 2, \dots, n$ , fix a statistical significance level  $\alpha, 0 < \alpha < 1$ . Choose a sufficiently large interval  $I = I^{(0)} = [a^{(0)}, b^{(0)}], 0 < a^{(0)} < b^{(0)}$  such that  $V^2 \in I^{(0)}$ . Partition  $I^{(0)}$  into  $m$  subintervals (of equal length)  $I_i^{(0)} = [a_i^{(0)}, b_i^{(0)}], i = 1, 2, \dots, m$ , and apply the  $\chi^2$  test to test the hypothesis  $V^2 \in I_i^{(0)}, i = 1, 2, \dots, m$ . We restrict our attention to the subinterval(s)  $I_{i^*}^{(0)} = [a_{i^*}^{(0)}, b_{i^*}^{(0)}] \subset I^{(0)}$  over which the (composite) hypothesis:

$$H_0 : a_{i^*}^{(0)} \leq V^2 \leq b_{i^*}^{(0)} \quad (2.13)$$

may be accepted with a significance level  $\alpha, 0 < \alpha < 1$ . If the hypothesis (2.13) cannot be accepted over any subintervals  $I_{i^*}^{(0)}$ , the choice of  $I^{(0)}$  and/or  $m$  may be changed. If the subinterval  $I_{i^*}^{(0)}$  is unique, we set  $I^{(1)} = [a^{(1)}, b^{(1)}] = I_{i^*}^{(0)}$ , otherwise, we set  $I^{(1)} = [a^{(1)}, b^{(1)}]$  equal to the union of the intervals over which (2.13) is accepted with a significance level  $\alpha$ . In both cases, the procedure described above is repeated, in the first case with the division of  $I^{(1)}$ , and in the second case, with the shrinking of  $I^{(1)}$ . In this way, a sequence of subintervals  $I^{(k)} = [a^{(k)}, b^{(k)}], k = 1, 2, \dots$  is constructed such that the hypothesis:

$$H_0 : a^{(k)} \leq V^2 \leq b^{(k)}, \quad k = 1, 2, \dots, \quad (2.14)$$

is accepted with a significance level  $\alpha, 0 < \alpha < 1$ . This procedure stops when  $b^{(k)} - a^{(k)} < \text{tol}$ , where  $\text{tol}$  is a given tolerance. We take the last set constructed using this procedure, over which the formulated hypothesis may be accepted, as a final estimate of an interval within which  $V^2$  falls with a significance level  $\alpha$ .

*Procedure 2.* A similar procedure is used to estimate an interval within which  $M$  falls with a significance level  $\alpha$ , based on the data sample  $\hat{x}_i, i = 1, 2, \dots, n$ , using the Student's  $t$  test.

The synthetic and real data are analyzed in Section 4 using Procedures 1 and 2 to identify intervals within which the parameters  $\sigma$  and  $\mu$  fall with a significance level  $\alpha$ .

Let us briefly discuss the design of Procedures 1 and 2. We focus on the motivations for the design of Procedure 1, which are not substantially different from the motivations for the design of Procedure 2. In Procedure 1, the statistical tests performed during

the calibration process were chosen as hypotheses with the goal of balancing two needs. On the one hand, a careful and conservative procedure is desirable, which means that a sufficiently large interval surrounding the volatility  $\sigma$  must be specified in the null hypothesis. On the other hand, the interval surrounding the volatility  $\sigma$  should be small enough to provide a meaningful range of variability around the corresponding option prices. Too large an interval of variability around the volatility  $\sigma$  generates a large interval of variability around the corresponding option prices, making the information about the option prices useless. To satisfy these contrasting needs, we initially select a large interval for the volatility  $\sigma$ , such that we trust that the estimated parameter lies within this interval. All scenarios, even the most extreme cases, may be considered systematically based in this choice. We then use an ad hoc procedure (i.e., Procedure 1) to gradually reduce the size of the chosen interval based on the idea that if a given hypothesis has already been accepted, it should be possible to refine it. The suggested procedures are iterative. At each step, the null hypothesis  $H_0$  of the considered test is modified in light of what has been discovered in the preceding steps. This leads to an iterative hypothesis testing procedure. Note that the only purpose underlying this iterative approach is to identify a satisfactory formulation of the test's null hypothesis. The sample data used in the tests do not vary with the iterative procedure. Type I errors that occur during each test are calculated as if the other tests were not performed. This iterative procedure is conceptually different from a multiple testing procedure [15] and from an analytic induction procedure [16]. In fact, in a multiple testing procedure (see [15]), a set of statistical inferences is considered simultaneously (i.e., each test has his own sample data), and the type I errors increase as the number of comparisons increases, unless the tests are perfectly dependent. In an analytic induction procedure (see [16]), the type I errors decrease because the size of the sample data set considered increases during the induction procedure.

An alternative approach to determining the range within which the volatility,  $\sigma$ , varies involves using an elementary method based on elementary statistics. For example, the sample data is used to obtain a point estimate of the parameter  $\sigma$  (i.e., compute the volatility of the sample data), then a confidence interval is constructed around this point estimate to quantify the uncertainty. Similarly, elementary statistics may be used to determine the range of variability of the drift  $\mu$ .

For the purposes pursued here, the method described in Procedures 1 and 2 is preferable to methods based on elementary statistics. Elementary statistics methods risk yielding dubious results in the event that a low-quality point estimate is selected. Also, the construction of large intervals of variability for the volatility can spoil any predictions regarding the corresponding option prices.

Finally, we note that in this paper, the idea of calibrating stochastic dynamical systems using statistical tests is suggested and implemented in the context of the Black-Scholes model. Procedures 1 and 2 are only heuristic procedures for selecting test hypotheses. Many other methods are available for estimating the model parameters using statistical tests that are plausible, such as methods based on elementary statistics, as mentioned previously.

### 3. Option Prices with Uncertain Volatility and Statistical Significance

Consider the Black-Scholes asset price dynamics model (2.1) and the problem of pricing options with uncertain volatility in the Black-Scholes model with a given statistical significance  $\alpha$ ,  $0 < \alpha < 1$ . We assume that the volatility value  $\sigma$  is not known exactly, but it is known that the volatility lies within a specified range, say,  $\sigma_1 \leq \sigma \leq \sigma_2$ , where  $\sigma_1$  and  $\sigma_2$



are constants and  $0 < \sigma_1 < \sigma_2$ , with a given significance level  $\alpha$ . The  $\chi^2$  test may be applied to the data used in the calibration problem, as described in Section 2, to decide whether to accept or reject the statement:

$$\sigma_1 \leq \sigma \leq \sigma_2, \quad \text{with significance level } \alpha. \quad (3.1)$$

We limit our attention to European vanilla call and put options.

The question that we want to answer is given a significance level  $\alpha$ ,  $0 < \alpha < 1$ , and assuming that the hypothesis  $H_0 : \sigma_1 \leq \sigma \leq \sigma_2$  is accepted with a significance level  $\alpha$ , determine the corresponding range within which the value of a European vanilla option lies (with a significance level  $\alpha$ ).

The answer to this question follows from the work of Avellaneda et al. [6] and of Lyons [7]. These authors proposed a method for estimating price options in the Black-Scholes model if the volatility  $\sigma$  is not known exactly but it is known that:

$$\sigma_1 \leq \sigma \leq \sigma_2. \quad (3.2)$$

In [6, 7], significance levels were not considered.

We remark that the bounds  $\sigma_1$  and  $\sigma_2$  in (3.1) and (3.2) were determined in different ways. In (3.1), these bounds correspond to parameters that can be chosen together with a significance level in such a way that the  $\chi^2$  test will accept the hypothesis  $H_0 : \sigma_1 \leq \sigma \leq \sigma_2$  over the sample data considered with a significance level  $\alpha$ . Alternatively, these bounds are parameters determined from the sample data through an iterative procedure, such as Procedure 1 described in Section 2. In (3.2), these bounds are assigned using common sense assumptions, or they are determined either by looking at extreme values of the volatility implied by the observed option prices or by looking at the low and high peak values of the historical volatility.

Let  $r$  be the risk-free interest rate,  $t$  be the time variable,  $S$  be the asset price,  $\mathcal{T} > 0$ ,  $g(S)$ ,  $S > 0$ , be the expiration date, and the pay-off function of the option to be priced. References [6, 7] showed that in the Black-Scholes framework, when the volatility satisfies (3.2), there exists an interval  $[\mathcal{U}_1, \mathcal{U}_2]$ , depending on  $S$  and  $t$ , such that the price  $\mathcal{U} = \mathcal{U}(S, t)$ ,  $S > 0$ ,  $0 < t \leq \mathcal{T}$ , of the option lies within this interval. That is, (3.2) implies that

$$\mathcal{U}_1(S, t) \leq \mathcal{U}(S, t) \leq \mathcal{U}_2(S, t), \quad S > 0, \quad 0 < t \leq \mathcal{T}. \quad (3.3)$$

The worst case option value  $\mathcal{U}_1(S, t)$ ,  $S > 0$ ,  $0 < t \leq \mathcal{T}$  satisfies the following nonlinear partial differential equation:

$$\frac{\partial \mathcal{U}_1}{\partial t} + \frac{1}{2} a(\Gamma_1)^2 S^2 \frac{\partial^2 \mathcal{U}_1}{\partial S^2} + rS \frac{\partial \mathcal{U}_1}{\partial S} - r\mathcal{U}_1 = 0, \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.4)$$

with final conditions

$$\mathcal{U}_1(S, \mathcal{T}) = g(S), \quad S > 0, \quad (3.5)$$

where

$$\Gamma_1 = \frac{\partial^2 \mathcal{U}_1}{\partial S^2}, \quad (3.6)$$

and

$$a(\Gamma_1) = \begin{cases} \sigma_2, & \text{if } \Gamma_1 \leq 0, \\ \sigma_1, & \text{if } \Gamma_1 > 0. \end{cases} \quad (3.7)$$

Similarly, the best-case option value  $\mathcal{U}_2(S, t)$ ,  $S > 0$ ,  $0 < t \leq \mathcal{T}$  satisfies the following nonlinear partial differential equation:

$$\frac{\partial \mathcal{U}_2}{\partial t} + \frac{1}{2} b(\Gamma_2)^2 S^2 \frac{\partial^2 \mathcal{U}_2}{\partial S^2} + rS \frac{\partial \mathcal{U}_2}{\partial S} - r\mathcal{U}_2 = 0, \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.8)$$

with final conditions

$$\mathcal{U}_2(S, \mathcal{T}) = g(S), \quad S > 0, \quad (3.9)$$

where

$$\Gamma_2 = \frac{\partial^2 \mathcal{U}_2}{\partial S^2}, \quad (3.10)$$

and

$$b(\Gamma_2) = \begin{cases} \sigma_2, & \text{if } \Gamma_2 \geq 0, \\ \sigma_1, & \text{if } \Gamma_2 < 0. \end{cases} \quad (3.11)$$

For example, for a European vanilla call option, we have  $g(S) = \max(S - K, 0)$ ,  $S > 0$ , and for a European vanilla put option, we have  $g(S) = \max(K - S, 0)$ ,  $S > 0$ , where  $K$  is the strike price of the option. To ensure the existence of a unique solution for the European vanilla call option, the following boundary conditions must be added to (3.4) and (3.5) and to (3.8) and (3.9):

$$\mathcal{U}(S, t) \rightarrow 0 \quad \text{as } S \rightarrow 0, \quad 0 < t < \mathcal{T}, \quad (3.12)$$

$$\mathcal{U}(S, t) \sim S - Ke^{-r(\mathcal{T}-t)} \quad \text{as } S \rightarrow \infty, \quad 0 < t < \mathcal{T}. \quad (3.13)$$

Similar boundary conditions must be added to (3.4) and (3.5) and to (3.8) and (3.9) when considering a European vanilla put option.

Equations (3.4) and (3.8) are known as the Black-Scholes-Barenblatt equations, and they reduce to the usual Black-Scholes equation in the absence of uncertainty about the volatility value (i.e., in the case of  $\sigma_1 = \sigma_2$ ). For a general pay-off function, these equations do

not have a closed-form solution and must be solved numerically; however, the properties of the Black-Scholes-Barenblatt (3.4) and (3.8) or, more generally, of the maximum principle of parabolic partial differential equations, ensure that when a call or a put option is considered, the convexity of the corresponding payoff functions  $g(S)$ ,  $S > 0$  imply that the functions  $\partial^2 \mathcal{U}_1 / \partial S^2$  and  $\partial^2 \mathcal{U}_2 / \partial S^2$  do not change sign for  $S > 0$ ,  $0 < t < \mathcal{T}$ . That is, when  $S > 0$ ,  $0 < t < \mathcal{T}$ , the functions  $\partial^2 \mathcal{U}_1 / \partial S^2$  and  $\partial^2 \mathcal{U}_2 / \partial S^2$  retain the value of their sign at  $t = \mathcal{T}$ ; therefore, in the context of a call or a put option, (3.4) and (3.8) reduce to the Black-Scholes equation. Note that when  $t = \mathcal{T}$ , we have  $\mathcal{U}_i(S, \mathcal{T}) = g(S)$ ,  $S > 0$ ,  $i = 1, 2$ , and in the context of the payoff functions  $g$  of the call and put European vanilla options, the corresponding  $\partial^2 g / \partial S^2$  are Dirac delta functions, which requires that the sign of  $\partial^2 g / \partial S^2$  must be interpreted in the context of the distributions. The Black-Scholes equation is linear. Under simple final conditions, such as the conditions relevant to the call and put options, this equation can be solved explicitly to yield the Black-Scholes formulae.

For example, consider the situation in which the worst-case option value  $\mathcal{U}_1$  is the solution to the following problem relating to a European call option:

$$\frac{\partial \mathcal{U}_1}{\partial t} + \frac{1}{2} \sigma_1^2 S^2 \frac{\partial^2 \mathcal{U}_1}{\partial S^2} + rS \frac{\partial \mathcal{U}_1}{\partial S} - r\mathcal{U}_1 = 0, \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.14)$$

$$\mathcal{U}_1(S, \mathcal{T}) = \max(S - K, 0), \quad S > 0, \quad (3.15)$$

with boundary conditions (3.12) and (3.13). Similarly, the best-case call option value  $\mathcal{U}_2$  satisfies

$$\frac{\partial \mathcal{U}_2}{\partial t} + \frac{1}{2} \sigma_2^2 S^2 \frac{\partial^2 \mathcal{U}_2}{\partial S^2} + rS \frac{\partial \mathcal{U}_2}{\partial S} - r\mathcal{U}_2 = 0, \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.16)$$

$$\mathcal{U}_2(S, \mathcal{T}) = \max(S - K, 0), \quad S > 0, \quad (3.17)$$

with boundary conditions (3.12) and (3.13).

The explicit solutions of (3.14) and (3.15) and of (3.16) and (3.17), in the context of the boundary conditions (3.12) and (3.13) are respectively [1]:

$$\mathcal{U}_1(S, t) = SN(d_1) - Ke^{-r(\mathcal{T}-t)}N(e_1), \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.18)$$

$$\mathcal{U}_2(S, t) = SN(d_2) - Ke^{-r(\mathcal{T}-t)}N(e_2), \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.19)$$

where

$$d_1 = \frac{\log(S/K) + (r + (1/2)\sigma_1^2)(\mathcal{T} - t)}{\sigma_1 \sqrt{\mathcal{T} - t}}, \quad e_1 = d_1 - \sigma_1 \sqrt{\mathcal{T} - t}, \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.20)$$

$$d_2 = \frac{\log(S/K) + (r + (1/2)\sigma_2^2)(\mathcal{T} - t)}{\sigma_2\sqrt{\mathcal{T} - t}}, \quad e_2 = d_2 - \sigma_2\sqrt{\mathcal{T} - t}, \quad S > 0, \quad 0 < t < \mathcal{T}, \quad (3.21)$$

and

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-(1/2)y^2} dy, \quad -\infty < x < \infty. \quad (3.22)$$

A similar analysis can be used to determine the worst-value and best-value European vanilla put options. Note that the formulae (3.18) and (3.19) are the Black-Scholes formulae.

We are now in a position to address the question posed at the beginning of this Section. First, let us assume that a “true” value of the volatility  $\sigma$  exists, even if it is unknown. In the Black-Scholes model, the price  $\mathcal{U}$  of an option is a monotonically increasing function of the volatility  $\sigma$ . We can therefore conclude that when (3.1) holds, we have:

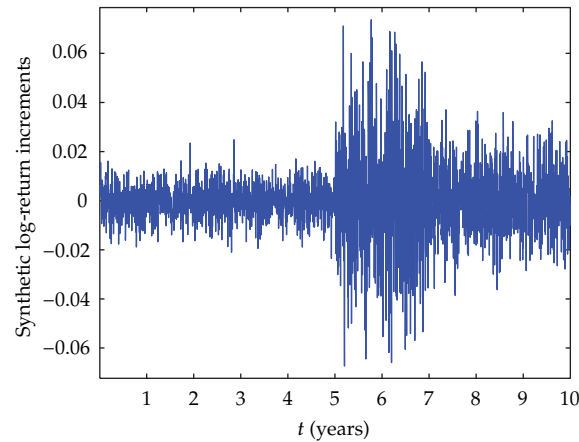
$$“\mathcal{U}_1(S, t) \leq \mathcal{U}(S, t) \leq \mathcal{U}_2(S, t), \quad S > 0, \quad 0 < t < \mathcal{T} \text{ with significance level } \alpha”, \quad (3.23)$$

where  $\mathcal{U}_1(S, t)$ ,  $\mathcal{U}_2(S, t)$ ,  $S > 0$ ,  $0 < t < \mathcal{T}$ , are the solutions to the appropriate Black-Scholes-Barenblatt equations. Note that if we consider European vanilla call and put options  $\mathcal{U}_1(S, t)$ ,  $\mathcal{U}_2(S, t)$ ,  $S > 0$ ,  $0 < t < \mathcal{T}$ , can be determined explicitly (see formulae (3.18) and (3.19), and similar formulae that can be deduced for put options).

The meaning of (3.23) can be restated as follows: if we assume that a “true” value of the volatility  $\sigma$  exists and that an analysis of sample data permits us to accept the hypothesis that this true value lies in the range  $[\sigma_1, \sigma_2]$  with a significance level  $\alpha$ , that is,  $\sigma_1 \leq \sigma \leq \sigma_2$  with probability  $1 - \alpha$ , it follows that the corresponding “true” value of the option price  $\mathcal{U}$  lies in the range  $[\mathcal{U}_1, \mathcal{U}_2]$  with probability  $1 - \alpha$ , that is, “ $\mathcal{U}_1(S, t) \leq \mathcal{U}(S, t) \leq \mathcal{U}_2(S, t)$ ,  $S > 0$ ,  $0 < t < \mathcal{T}$ , with significance level  $\alpha$ ”, where  $\mathcal{U}_1$  and  $\mathcal{U}_2$  are the solutions to the appropriate Black-Scholes-Barenblatt equations.

#### 4. Numerical Experiments

Several numerical experiments were conducted to demonstrate the principles described above. As a first example, we consider a numerical experiment that involves solving the calibration problem discussed in Section 2 using synthetic data. This experiment features an analysis of the time series corresponding to daily data of asset prices over a period of ten years. We assume that one year is composed of 253 trading days. The time series studied comprises  $253 \cdot 10 + 1 = 2531$  daily asset price data points, that is, it comprises the asset price  $\hat{S}_i$  observed at time  $t = t_i = i\Delta t$ ,  $i = 0, 1, \dots, 2530$ ,  $\Delta t = 1/253$ . The synthetic data were obtained by computing one trajectory using the stochastic differential equation (2.1) for several choices of the parameter values, then examining the computed trajectory at time  $t = t_i = i\Delta t$ ,  $i = 0, 1, \dots, 2530$ . We choose as the initial conditions, time  $t = t_0 = 0$ ,  $S_0 = \hat{S}_0 = 1200$ . We choose  $\mu = \bar{\mu}_1 = 0.01$  and  $\sigma = \bar{\sigma}_1 = 0.1$  in the first five years (i.e., for  $t = t_i$ ,  $i = 0, 1, \dots, 1264$ ),  $\mu = \bar{\mu}_2 = 0.06$  and  $\sigma = \bar{\sigma}_2 = 0.4$  in the sixth and seventh years (i.e., for  $t = t_i$ ,  $i = 1265, 1266, \dots, 1770$ ), and  $\mu = \bar{\mu}_3 = 0.03$  and  $\sigma = \bar{\sigma}_3 = 0.2$  in the last three years (i.e., for  $t = t_i$ ,  $i = 1771, 1772, \dots, 2530$ ). The synthetic data were generated such that the last data point of the fifth year was the initial data point of the sixth year. Similar statements hold



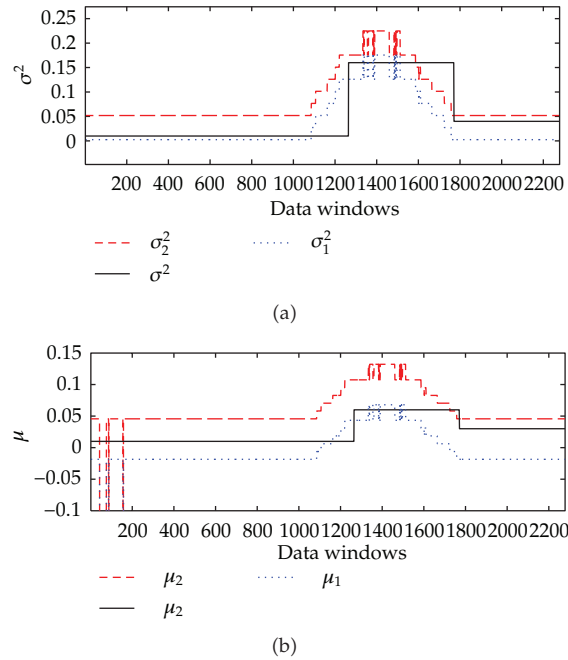
**Figure 1:** The daily log return increments (synthetic data).

for the initial data point values of the last three years. The daily log return increments of the synthetic asset prices generated in this way are shown in Figure 1. The fact that the data are generated using three different parameters can be readily seen by inspection of Figure 1.

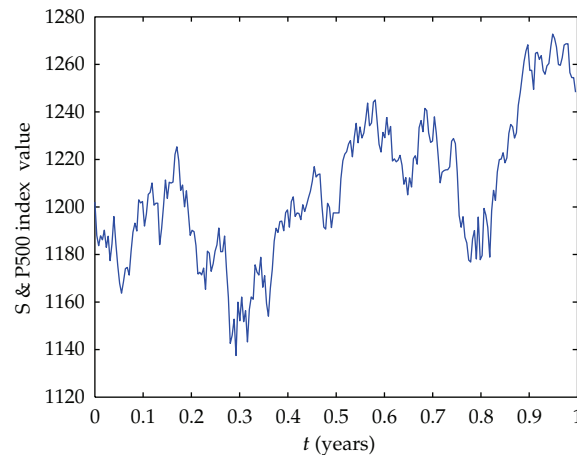
Consider the following calibration problem: given  $\Delta t = 1/253$ ,  $\alpha = 0.1$ , and the asset price observation  $\hat{S}_i$  at time  $t = t_i = i\Delta t$ ,  $i = 0, 1, \dots, 2530$ , determine intervals within which the parameters  $\mu$  and  $\sigma$  of model (2.1) fall with a significance level  $\alpha = 0.1$ .

We solve this calibration problem by applying Procedures 1 and 2, described in Section 2, to the data associated with a time window comprising 253 consecutive observations, that is, the observations corresponding to 253 consecutive trading days (one year). This window is moved stepwise across the ten years of data, discarding the datum corresponding to the first observation time within the window and inserting the datum corresponding to the next observation time after the window. The calibration problem is solved for each data time window by applying Procedures 1 and 2 described in Section 2. In other words, the problem is solved  $2530 - 253 + 1 = 2278$  times, identifying two intervals for each calibration problem solved within which the volatility and drift parameters fall with a significance level  $\alpha = 0.1$ . Referring to Procedures 1 and 2 described in Section 2, we choose  $m = 2$ ,  $\text{tol} = 2 \cdot 10^{-4}$  and appropriate initial intervals  $I = I^{(0)}$  to determine an interval within which the parameters  $\sigma$  and  $\mu$  fall. The parameter reconstructions obtained from moving the window along the data are shown in Figure 2. The abscissa in Figure 2 corresponds to the data window used to reconstruct the model parameters. The data windows are numbered in ascending order, beginning with one, according to the first day within the window being considered. Figure 2 shows that the intervals containing the parameters  $\mu$  and  $\sigma$ , chosen by Procedures 1 and 2, and the times at which the parameter values changed were reconstructed satisfactorily.

The second numerical experiment involved the use of real data. The real data studied were the 2005 daily values of the U.S. S&P500 index (see Figure 3) and of the prices of the European vanilla call and put options on this index. Recall that the U.S. S&P500 index is one of the leading indices of the New York Stock Exchange. Specifically, we considered the daily closing values of the S&P500 index and the bid prices of the vanilla European call and put options on the S&P500 index during a period of 12 months, beginning on January 3, 2005, and ending on December 30, 2005. Within this period are 253 trading days and more than 153 000

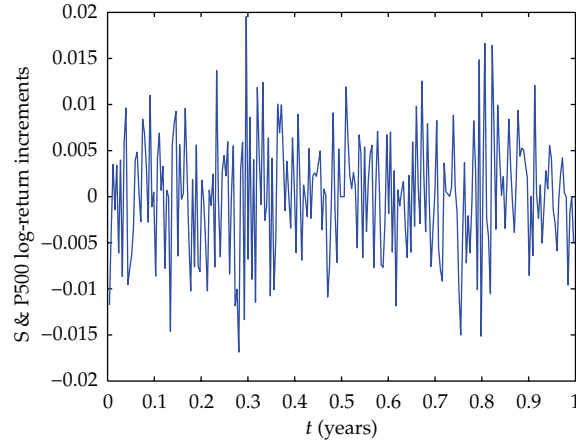


**Figure 2:** The parameters  $\sigma^2$  and  $\mu$ , reconstructed from the synthetic data.



**Figure 3:** The S&P500 index (year 2005).

option prices. We limit our study to the call and put prices corresponding to options with a positive volume (i.e., a positive number of contracts) and a positive bid price, traded on the day corresponding to the price considered. This data set included 46 823 options prices. Because there are 253 trading days in the year 2005, we define a “year” as 253 consecutively ordered trading days. The time  $t = t_0 = 0$  was assigned to the day of January 3, 2005. A total of 253 daily S&P500 index values  $\hat{S}_i$  were observed at time  $t = t_i = i\Delta t$ ,  $i = 0, 1, \dots, 252$ , with  $\Delta t = 1/253$  year. The S&P500 index and the corresponding (daily) log return increments in the year 2005 are shown in Figures 3 and 4, respectively.



**Figure 4:** The S&P500 daily log return increments (year 2005).

This data set were interpreted using the Black-Scholes model, as described in Section 2. We begin by studying the variance and drift in the Black-Scholes model applying the S&P500 index over the year 2005. The S&P500 daily log return increments  $\hat{x}_i = \ln(\hat{S}_i/\hat{S}_{i-1})$ ,  $i = 1, 2, \dots, 252$ , were analyzed using the Black-Scholes model. The data were considered as a sample (see Figure 4) of 252 observations taken from a set of independent identically distributed Gaussian random variables:

$$\hat{x}_i \text{ is sampled from } X_i \sim \mathcal{N}(M, V^2), \quad i = 1, 2, \dots, 252, \quad (4.1)$$

where the mean  $M$  and the variance  $V^2$  are defined in (2.7). Procedures 1 and 2, described in Section 2, were used to estimate  $V^2$  (and, therefore,  $\sigma^2$ ) using the  $\chi^2$  test. Subsequently,  $M$  (and, therefore,  $\mu$ ) was estimated using the Student's  $t$ -test.

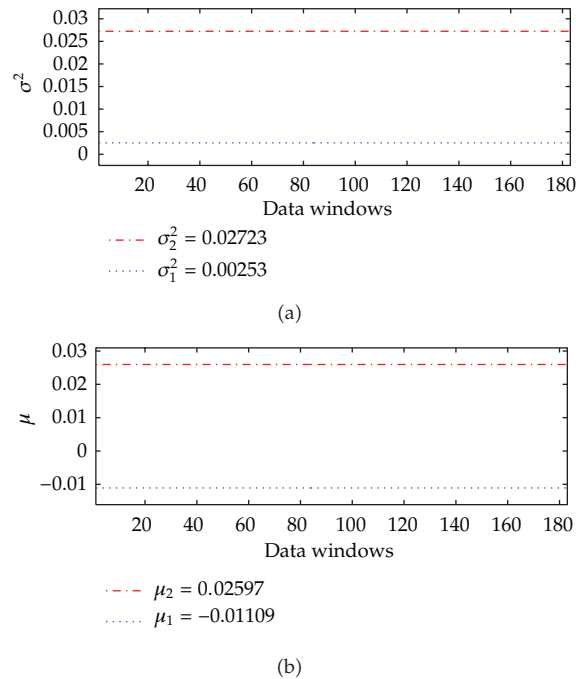
Given the sample data comprising the S&P500 daily log return increments  $\hat{x}_i$ ,  $i = 1, 2, \dots, 252$ , the significance level  $\alpha = 0.1$ ,  $m = 2$ ,  $\text{tol} = 10^{-4}$ , and the appropriate initial intervals  $I = I^{(0)}$ , we initiated Procedures 1 and 2 to find that the hypotheses:

$$2.5297 \cdot 10^{-3} = \sigma_1^2 \leq \sigma^2 \leq \sigma_2^2 = 2.7232 \cdot 10^{-2}, \quad (4.2)$$

$$-1.1087 \cdot 10^{-2} = \mu_1 \leq \mu \leq \mu_2 = 2.5968 \cdot 10^{-2}, \quad (4.3)$$

are accepted with a significance level  $\alpha = 0.1$ .

We next perform a type of stability analysis over the intervals (4.2) and (4.3) determined from the statistical tests. To do this, we fixed  $\alpha = 0.1$  and applied Procedures 1 and 2 to determine the intervals within which  $\sigma^2$  and  $\mu$  lay with the significance level  $\alpha$ , starting from a window of 70 consecutive observations corresponding to 70 consecutive observation times (i.e., 70 consecutive trading days). The data window was shifted stepwise over the data time series, discarding the datum corresponding to the first observation time of the window and inserting the datum corresponding to the next observation time after



**Figure 5:** The parameters  $\sigma^2$  and  $\mu$ , reconstructed from the S&P500 data (year 2005).

the window. This procedure generated  $252-70+1$  data windows over the data time series. For each window, the corresponding calibration problem was solved. We found  $252-70+1$  pairs of intervals within which the volatility and drift parameters lay with a significance level of  $\alpha = 0.1$ . Figure 5 shows that as the data window was shifted, the intervals determined through Procedures 1 and 2 remained stable. The abscissa in Figure 5 represents the data windows used to reconstruct the model parameters, numbered in ascending sequential order, beginning with one. Figure 5 shows the intervals determined by Procedures 1 and 2 of Section 2 corresponding to each data window.

Finally, we considered the 46.823 S&P500 European call and put option prices during the year 2005, and we attempted to interpret these values using the method developed in Section 3. The estimates (2.9), (2.10) established with a significance level  $\alpha = 0.1$  permitted us to determine the corresponding option price intervals using the Black-Scholes-Barenblatt equations. The worst-case option values  $\mathcal{U}_1(S, t)$ ,  $S > 0$ ,  $0 < t \leq \mathcal{T}$ , and the best-case option values  $\mathcal{U}_2(S, t)$ ,  $S > 0$ ,  $0 < t \leq \mathcal{T}$  could thereby be determined. Note that  $\mathcal{U}_1$  and  $\mathcal{U}_2$  could be determined using the Black-Scholes formula using (2.9). We selected  $r = (\mu_1 + \mu_2)/2$ , with  $\mu_1$  and  $\mu_2$ , as defined in (2.10). It should be noted that the value of  $r$  was not relevant in determining  $\mathcal{U}_1$  and  $\mathcal{U}_2$ , and varying  $r$  over a reasonable range did not substantially change  $\mathcal{U}_1$  or  $\mathcal{U}_2$ . Moreover, the time to maturity within the Black-Scholes framework may be computed by considering a year composed of 365 days. We compute the percentage of European call (% call) and put (% put) option prices on the S&P500, observed over the year 2005, that verified (3.23), assuming (2.9). The results obtained are shown in Tables 1–3. In these tables,  $N_{\text{call}}$  and  $N_{\text{put}}$  denote respectively the number of call prices and put prices corresponding to options with the characteristics described in the caption of the table. The quantities  $I_{\text{call}}$  and  $I_{\text{put}}$  denote respectively the average relative amplitude of the call price intervals and



**Table 1:** S&P500 option prices: in the money options (year 2005). These results were obtained using the estimates (4.2) and (4.3).

January–April 2005	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
	73.8% (172.22)	72.3% (313.01)	1571	1822	0.23	0.34	94.01	76.82
May–August 2005	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
	74.6% (335.25)	62.0% (202.34)	2005	1745	0.24	0.35	92.03	76.73
September–December 2005	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
	65.5% (401.78)	59.7% (557.22)	2174	2300	0.23	0.35	97.94	76.14

**Table 2:** S&P500 option prices: at the money options (year 2005). These results were obtained using the estimates (4.4) and (4.5).

January–April 2005	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
	51.1% (1358.18)	67.7% (1869.60)	852	902	0.27	0.28	28.29	24.76
May–August 2005	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
	56.5% (1899.99)	66.2% (1989.07)	1115	1154	0.28	0.28	27.40	22.52
September–December 2005	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
	41.5% (2377.19)	59.4% (2652.84)	1188	1208	0.27	0.28	30.29	23.45

put price intervals determined using the Black-Scholes-Barenblatt equations, and  $P_{\text{call}}$ ,  $P_{\text{put}}$  denote respectively the average bids of the call and put prices. It should be noted that the % call and % put columns may be expressed as the average number of contracts applied to the options that were considered to have been traded.

Recall that given the asset price  $S$  and the strike price  $K$  of an option, a call option (or a put option) is in the money if  $S > K$  (if  $S < K$ ) is out the money if  $S < K$  (if  $S > K$ ) and is at the money if  $S = K$ . In the numerical experiments, the condition  $S = K$  is substituted with  $|S - K| < \epsilon$ , where  $\epsilon$  is a (given) positive quantity. As a consequence, the conditions  $S > K$ ,  $S < K$  may be rewritten as  $S > K + \epsilon$  and  $S < K - \epsilon$ , respectively. Using these criteria, the 46 823 option prices considered above may be divided into three subsets corresponding to the prices of in the money, at the money, or out of the money options. Take  $\epsilon$  to be equal to one percent of the average strike price of the options considered.

Table 1 refers to the in the money S&P500 option prices, obtained by specifying (2.9) and (2.10) as (4.2) and (4.3), respectively.

A similar analysis of the S&P500 option prices corresponding to options out of and at the money reveals that the use of the intervals (4.2) and (4.3) leads to huge call and put price intervals, making the obtained results of dubious practical value. One way to overcome this drawback is to refine the estimates (4.2) and (4.3) by reducing the parameter  $\text{tol}$  in Procedures 1 and 2 of Section 2 until the option price intervals of “acceptable average relative amplitude” (i.e., average relative amplitude of some tens of percentage points) are obtained. Taking  $\text{tol}_1 = \text{tol}/4 = 10^{-4}/4$ , we find that the hypotheses:

$$8.7051 \cdot 10^{-3} = \sigma_1^2 \leq \sigma^2 \leq \sigma_2^2 = 1.4881 \cdot 10^{-2}, \quad (4.4)$$

$$1.2646 \cdot 10^{-3} = \mu_1 \leq \mu \leq \mu_2 = 1.0528 \cdot 10^{-2}, \quad (4.5)$$

**Table 3:** S&P500 option prices: out of the money options (year 2005). These results were obtained using the estimates (4.6) and (4.7).

	% call	% put	$N_{\text{call}}$	$N_{\text{put}}$	$I_{\text{call}}$	$I_{\text{put}}$	$P_{\text{call}}$	$P_{\text{put}}$
January–April 2005	20.1% (691.52)	2.41% (1061.76)	3412	4892	0.26	0.57	11.13	9.05
May–August 2005	13.4% (929.68)	3.55% (1337.39)	3644	6316	0.23	0.58	9.89	8.33
September–December 2005	12.6% (1598.65)	2.71% (1908.66)	4055	6468	0.24	0.59	12.73	8.58

are accepted with a significance level  $\alpha = 0.1$ . The choice of (4.4), (4.5) as intervals containing  $\sigma^2$  and  $\mu$ , respectively, leads to an average relative amplitude of some tens of percentage points for the option price intervals when the call and put options at the money are considered. Table 2 shows the results obtained on at the money option prices using (4.4) and (4.5).

In the context of the S&P500 option prices relative to options out of the money, the parameter  $\text{tol}$  must be further reduced to keep the average relative amplitudes of the option price intervals to reasonable values. For example, taking  $\text{tol}_2 = \text{tol}/10 = 10^{-4}/10$ , we find that the hypotheses:

$$1.1021 \cdot 10^{-2} = \sigma_1^2 \leq \sigma^2 \leq \sigma_2^2 = 1.2565 \cdot 10^{-2}, \quad (4.6)$$

$$4.7385 \cdot 10^{-3} = \mu_1 \leq \mu \leq \mu_2 = 7.0544 \cdot 10^{-3}, \quad (4.7)$$

are accepted with a significance level of  $\alpha = 0.1$ . Table 3 shows the results obtained from out of the money option prices using (4.6), (4.7).

Table 1 shows that the in the money S&P500 option prices were reasonably well interpreted by the model proposed here. Some 60%–80% of the prices of in the money S&P500 call and put options could be explained by the model, and 20%–40% of the average relative amplitudes of the option price intervals were considered to be of possible practical value. Table 1 further shows that in the case of in the money options, the call prices seemed to be better explained than the put prices. On the other hand, the numerical results shown in Table 2 indicated that the at the money S&P500 put option prices were better explained than the corresponding S&P500 call options prices. The results relative to at the money S&P500 option prices were satisfactory. Unfortunately, Table 3 shows that the out of the money S&P500 option prices and, above all, the S&P500 prices relative to out of the money put options, were not well interpreted using our methodology. This could result from the fact that the out of the money options usually had low prices. Tables 1–3 showed that the prices of the options out of the money were the smallest values.

## 5. Conclusions

A method for calibrating the Black-Scholes asset price dynamics model using the  $\chi^2$  and Student's  $t$  tests was developed to obtain estimates, with statistical significance, of the volatility and drift parameters of the model. The effects of these estimates on the option

pricing problem were considered. The pricing problem for European call and put options with uncertain volatility in the Black-Scholes framework was reinterpreted by associating a statistical significance to the option price intervals determined using the Black-Scholes-Barenblatt equations. The proposed method was tested on synthetic and real data. The real data considered were S&P500 values and the corresponding European call and put option prices over the year 2005. The calibration problem for the Black-Scholes model was solved based on the S&P500 data, and the S&P500 call and put option price data were interpreted in the framework of option prices with uncertain volatility and significance level. In the money and at the money S&P500 option prices could be modeled with a high degree of accuracy. The out of the money S&P500 option prices were not explained well by this model.

The methodology discussed in this paper was introduced in the context of the Black-Scholes asset dynamics model to take advantage of the model simplicity and to offer an alternative approach to the calibration techniques in widespread use by practitioners in the financial markets. The calibration problem may be solved using statistical tests that can affect the (call, put) option pricing, and these effects may be studied using elementary mathematics. Several extensions of our work may be considered. For example, the skew-normal models (see, e.g. [17]) are commonly used to interpret financial data. Many financial data sets exhibit asymmetric distributions, and several models, including the skew-normal models, have been proposed to capture this asymmetry and repair any inadequacies of the Black-Scholes model. In these models, a solution to the calibration problem that uses statistical tests should not differ too significantly from the solution proposed here. The extension of our methodology to the study of more general stochastic dynamical models, such as the Heston model or some of its variants, may require the development of new ad hoc statistical tests. Any application of new statistical tests to a sample data set should rely on the use of numerical methods. Another interesting idea to explore involves attempting to simplify the new statistical tests developed, solve the calibration problem, and approximate the relevant probability density functions using an asymptotic expansion within some meaningful limit. The intelligent use of these asymptotic expansions can provide significant computational advantages.

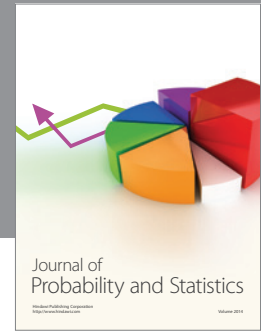
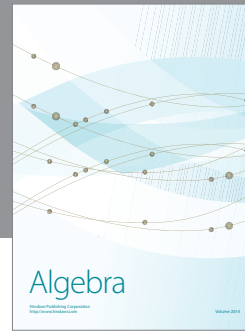
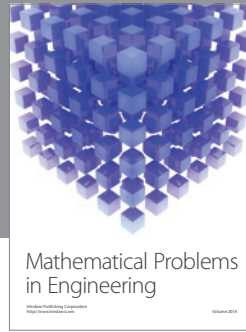
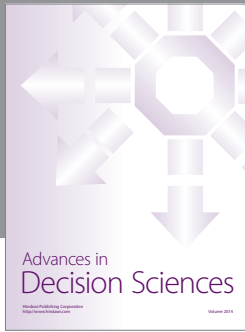
## Acknowledgments

The results reported in this paper were partially supported by MUR, the Ministero Università e Ricerca (Roma, Italy), 40%, 2007, under the grant: "The impact of population aging on financial markets, intermediaries and financial stability." The support and sponsorship of MUR are gratefully acknowledged. The numerical results reported in this paper were obtained using the computing grid of ENEA (Roma, Italy). The support and sponsorship of ENEA are gratefully acknowledged.

## References

- [1] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *Journal of Political Economy*, vol. 81, pp. 637–659, 1973.
- [2] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, vol. 6, no. 1, pp. 327–343, 1993.
- [3] W. S. Gosset, "The probable error of a mean," *Biometrika*, vol. 6, no. 1, pp. 1–25, 1908.
- [4] R. A. Johnson and G. K. Bhattacharyya, *Statistics: Principles and Methods*, John Wiley & Sons, New York, NY, USA, 5th edition, 2006.

- [5] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Philosophical Magazine*, vol. 50, pp. 157–175, 1900.
- [6] M. Avellaneda, A. Levy, and A. Parás, "Pricing and hedging derivative securities in markets with uncertain volatilities," *Applied Mathematical Finance*, vol. 2, pp. 73–88, 1995.
- [7] T. J. Lyons, "Uncertain volatility and the risk-free synthesis of derivatives," *Applied Mathematical Finance*, vol. 2, no. 2, pp. 117–133, 1995.
- [8] F. Mariani, G. Pacelli, and F. Zirilli, "Maximum likelihood estimation of the Heston stochastic volatility model using asset and option prices: an application of nonlinear filtering theory," *Optimization Letters*, vol. 2, no. 2, pp. 177–222, 2008.
- [9] L. Fatone, F. Mariani, M. C. Recchioni, and F. Zirilli, "The calibration of the Heston stochastic volatility model using filtering and maximum likelihood methods," in *Proceedings of the Dynamic Systems and Applications*, G. S. Ladde, N. G. Medhin, C. Peng, and M. Sambandham, Eds., vol. 5, pp. 170–181, Dynamic, Atlanta, Ga, USA, 2008.
- [10] L. Fatone, F. Mariani, M. C. Recchioni, and F. Zirilli, "Maximum likelihood estimation of the parameters of a system of stochastic differential equations that models the returns of the index of some classes of hedge funds," *Journal of Inverse and Ill-Posed Problems*, vol. 15, no. 5, pp. 493–526, 2007.
- [11] L. Fatone, F. Mariani, M. C. Recchioni, and F. Zirilli, "An explicitly solvable multi-scale stochastic volatility model: option pricing and calibration problems," *Journal of Futures Markets*, vol. 29, no. 9, pp. 862–893, 2009.
- [12] L. Fatone, F. Mariani, M. C. Recchioni, and F. Zirilli, "Calibration of a multiscale stochastic volatility model using European option prices," *Mathematical Methods in Economics and Finance*, vol. 3, no. 1, pp. 49–61, 2008.
- [13] L. Fatone, F. Mariani, M. C. Recchioni, and F. Zirilli, "The analysis of real data using a multiscale stochastic volatility model," *European Financial Management*. In press.
- [14] J. H. Hull, *Options, Futures, and Other Derivatives*, Pearson Prentice Hall, London, UK, 7th edition, 2008.
- [15] R. G. Miller Jr., *Simultaneous Statistical Inference*, Springer, New York, NY, USA, 2nd edition, 1981.
- [16] F. Znaniecki, *The Method of Sociology*, Farrar & Rinehart, New York, NY, USA, 1934.
- [17] A. Azzalini, "A class of distributions which includes the normal ones," *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 171–178, 1985.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

