Journal of Statistical Computation and Simulation
Vol. 00, No. 00, Month 2007, 1–14

Taylor & Francis
Taylor & Francis Group

Three-parameter stochastic lognormal diffusion model: statistical computation and simulating annealing – application to real case

R. GUTIÉRREZ, R. GUTIÉRREZ-SÁNCHEZ, A. NAFIDI and E. RAMOS*

Departamento de Estadística e Investigatión Operativa, Facultad de Ciencias, Universidad de Granada, Campus Fuentenueva, s/n, 18071 Granada, Spain

(Received January 2005)

In this paper, we propose a new study of a stochastic lognormal diffusion process (SLDP), with three parameters, which can be considered as an extension of the bi-parametric lognormal process with the addition of a threshold parameter. From the Kolmogorov equation, we obtain the probability density function and the moments of this process. The statistical inference of the parameter is studied by considering discrete sampling of the sample paths of the model and then using the maximum likelihood (ML) method. The estimation of the threshold parameter requires the solution of a nonlinear equation. To do so, we propose two methods: the classical Newton–Raphson (NR) method and one based on simulated annealing (SA). This methodology is applied to an example with simulated data corresponding to the process with known parameters. From this, we obtain the estimators of the parameters by both methods (NR and SA). Finally, the methodology studied is applied to a real case concerning the mean age of males in Spain at the date of their first wedding.

Keywords: Three-parameter lognormal diffusion process; Discrete sampling; Simulated annealing; Fits and forecast; Mean age of males in Spain at the date of their first wedding

1. Introduction

In the context of three-parameter lognormal distributions, we discuss the original study of the three-parameter lognormal diffusion process, which is one-dimensional in the endogenous variable and has three parameters.

We seek to obtain a three-parameter univariate lognormal model, which is flexible and broad enough to cover important fields of application that are currently not served by standard models. Moreover, taking possible applications into account, the maximum likelihood (ML) estimation is constructed on the basis of discrete sampling based on the conditioned likelihood derived from the transitions that are the solutions to the corresponding Kolmogorov equations.

*Corresponding author. Tel.: +34-958-246306; Fax: +34-958-243267; Email: ramosa@ugr.es

Journal of Statistical Computation and Simulation
ISSN 0094-9655 print/ISSN 1563-5163 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00949650701550622

The importance of three-parameter lognormal distribution as a probability model has been recognized in a wide range of scientific disciplines, including biology, geology, agricultural science, statistics and economics. For example, Wicksell [1] and Guerrieri [2] applied three-parameter distribution in a study of the age distribution of those marrying for the first time, using the direct estimation method for the parameters of Wicksell [1]. In 1957, Aitchison and Brown [3] applied all the methods available for 65 simulated samples. Royston [4] used this process in a study of the concentration of antibodies in blood, and Crawford [5] applied the same method to study the acidity of lakes. Given the importance of three-parameter lognormal distribution, many studies have been aimed at the problems involved in estimating the parameters of this distribution for data samples, particularly concerning the theoretical and computational difficulties that may arise when applying the ML method. In response to these difficulties, a considerable number of estimators and estimation methods have been proposed. Cohen [6] and Harter and Moore [7] obtained local ML estimates, while Hill [8] proposed Bayesian estimators. Calitz [9], on the other hand, used simulation procedures to compare likelihoods, percentiles and estimators of the moments. More recently, Giesbrecht and Kempthorne [10] obtained ML estimators for discrete models, considering grouped intervals, and Wingo [11] introduced *penalty and barrier functions* to increase convergence when calculating local ML estimators. Cohen and Witten [12], Cohen *et al.* [13] and Crow and Shimizu [14] proposed various modifications for ML estimators and for estimators of the moments. Lifson and Bhattacharyya [15] used a regression method to estimate the percentiles. Wingo [11] worked using a computational algorithm to maximize the likelihood function (as a function of the third parameter). Kappenman [16] estimated the three parameters by using iterative procedures and comparing the results with the likelihood estimators. Finally, Royston [4] estimated the *threshold parameter* using the properties of the median of the lognormal distribution, designing a test to compare hypotheses, based on the median.

An important theoretical difficulty that must be overcome is that the likelihood function is maximum where the parameters present unacceptable values. In this respect, Hill [8] obtained Bayesian arguments to justify the use of parametric estimators, which correspond to large local maxima. Moreover, Heyde [17] showed that the three-parameter lognormal distribution cannot be determined solely from its moments, as various problems arise concerning the estimation by the method of moments. From a non-Bayesian standpoint, Griffiths [18] showed that the ML method can be considered a reasonably reliable approach to the problem.

The main computational difficulty lies in the fact that the estimation of the parameters complicates the precautions that must be taken (in computational terms) when obtaining the estimators using iterative numerical methods. Many of these methods have been discussed, for example, by Cohen in refs. [6], [12] and [13] and Crow and Shimizu [14], Lambert [19], Harter and Moore [7] and Calitz [9]. If these methods are used without avoiding the region of attraction of infinite ML, then convergence difficulties will be encountered. To avoid such difficulties and to increase the convergence in the estimators of local ML estimators, Wingo [11] introduced frontier function methods with mobile cut-offs.

The main objective of this study is to add a threshold parameter to the biparameter stochastic lognormal diffusion process (SLDP), as studied by Gutiérrez *et al.* [20–22]. This work is structured as follows: in section 2, we describe the model using Kolmogorov equations and calculate the probability density function (pdf) and the moments of the model. In section 3, we estimate the parameters of the model by ML method using discrete sampling. In order to estimate the new parameter, we obtain a nonlinear equation, which we propose to solve using NR and SA, reducing the interval of the parameter search by means of Wingo's [11] reparametrization. In section 4, we simulate the trajectory of the process from the exact solution of Itô's stochastic differential equation (SDE) and estimate the parameters by the ML method.

In the last section, we apply the model to real data for the mean age of males in Spain at their first wedding.

2. The model and its basic probabilistic characteristics

2.1 *The model*

The one-dimensional lognormal diffusion process with three parameters can be introduced by means of the Kolmogorov backward and forward equation as a Markov process $\{X(t), t_0 \leq t \leq T\}$ with values in $]\gamma, +\infty[$, with almost-certainly continuous trajectories and with a distribution function for the process transition that is given by $P(y,t|x,s) = P[X(t) \leq y|X(s) = x]$, $x > \gamma$, $y > \gamma$ and $\gamma \in \mathbb{R}$.

And by assuming the following conditions:

- $\lim_{h\to 0} 1/h \int_{|y-x|>\epsilon} P(dy, t+h|x,t) = 0$,
- $\lim_{h\to 0} 1/h \int_{|y-x|\leq\epsilon} (y-x) P(dy, t+h|x,t) = A_1(x,t;\gamma) = \mu(x-\gamma)$,
- $\lim_{h\to 0} 1/h \int_{|y-x|\leq\epsilon} (y-x)^2 P(dy, t+h|x,t) = A_2(x,t;\gamma) = \sigma^2(x-\gamma)^2 > 0$,
- the higher-order infinitesimal moments are null,

where $\mu > 0$, $\sigma > 0$, $\mu, \sigma \in \mathbb{R}$, the infinitesimal moments of the process are

$$A_1(x,t;\gamma) = \mu(x-\gamma),$$
$$A_2(x,t;\gamma) = \sigma^2(x-\gamma)^2,$$

and the corresponding Kolmogorov backward and forward equation are

$$\frac{\partial p}{\partial s} + \frac{1}{2}\sigma^2(x-\gamma)^2\frac{\partial^2 p}{\partial x^2} + \mu(x-\gamma)\frac{\partial p}{\partial x} = 0,$$
$$-\frac{\partial p}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 (y-\gamma)^2 p}{\partial y^2} - \mu\frac{\partial (y-\mu) p}{\partial y} = 0,$$

where p denotes the transition pdf, $p(y,t|x,s)$, corresponding to the transition distribution function $P(y,t|x,s)$.

The common solution to these equations, which can be obtained using Ricciardi's theorem (see [23]), is

$$p(y,t|x,s) = \frac{1}{(y-\gamma)\sigma}\frac{1}{\sqrt{2\pi(t-s)}}\exp\left\{-\frac{1}{2\sigma^2(t-s)}(\ln(y-\gamma) - \ln(x-\gamma) - a(t-s))^2\right\}, \quad (1)$$

with the initial condition $p(y,s|x,s) = \delta(y-x)$, where $a = \mu - \sigma^2/2$.

The distribution of the random variable $X(t)|X(s) = x$ is the one-dimensional three-parameter lognormal distribution

$$\Lambda_1\left[\gamma + \ln(x-\gamma) + a(t-s); \sigma^2(t-s)\right],$$

$\Lambda_1[\alpha;\beta]$ denotes the one-dimensional three-parameter lognormal distribution with parameters α and β.

Alternatively, the above-defined process can be considered as the solution of the following Itô's SDE

$$\mathrm{d}X(t) = \mu(X(t) - \gamma)\mathrm{d}t + \sigma(X(t) - \gamma)\mathrm{d}W(t), \quad X(t_0) = x_0, \tag{2}$$

where $W(t)$ represents the Wiener process with independent increments $W(t) - W(s)$ distributed according to $\mathcal{N}(0, t - s)$ for $t > s$.

2.2 *Moments of the process*

The different moments of the process are obtained from the following expression,

$$E[(X(t) - \gamma)^k | X(s) = x_s] = (x_s - \gamma)^k \exp\left\{\left(ka + \frac{1}{2}k^2\sigma^2\right)(t - s)\right\}.$$

Then the conditioned trend function of the process is

$$E[X(t)|X(s) = x_s] = \gamma + (x_s - \gamma)\exp\{\mu(t - s)\}. \tag{3}$$

Taking into account the initial condition $P[X(t_0) = x_0] = 1$, the trend function is given by

$$E[X(t)] = \gamma + (x_0 - \gamma)\exp\{\mu(t - t_0)\}. \tag{4}$$

The covariance function has the following form

$$\mathrm{Cov}[X(t)X(s)] = (x_0 - \gamma)^2 \exp\{\mu((t - t_0) + (s - t_0))\}[\exp\{\sigma^2((t \wedge s) - t_0)\} - 1],$$

with $t \wedge s = \min(t, s)$.

3. Estimation of the parameters

3.1 *Maximum likelihood estimation*

We shall now estimate the parameters of the model using the ML method. Let us consider a discrete sampling of the process

$$\{X(t_1) = x_1, X(t_2) = x_2, \ldots, X(t_n) = x_n\}$$

for the instants $t_1, \ldots, t_n$, with the initial condition $P[X(t_1) = x_1] = 1$. The associated ML function is thus

$$\mathbb{L}(x_1, \ldots, x_n; a, \sigma^2, \gamma) = \prod_{i=2}^{n} P(x_i, t_i | x_{i-1}, t_{i-1}).$$

This function tends to infinity when γ tends to $x_{(1)}$, where $x_{(1)} = \inf_{0 \le j \le n}(x_j)$.

By using equation (1), the log-likelihood function is then

$$\begin{aligned}\ln[\mathbb{L}(x_1, \ldots, x_n; a, \sigma^2, \gamma)] = &-\frac{n-1}{2}\ln(2\pi\sigma^2) \sum_{i=2}^{n} \ln(x_i - \gamma) - \frac{1}{2}\sum_{i=2}^{n}\ln(t_i - t_{i-1}) \\ &- \frac{1}{2\sigma^2}\sum_{i=2}^{n}\frac{1}{(t_i - t_{i-1})}[\ln(x_i - \gamma) - \ln(x_{i-1} - \gamma) \\ &- a(t_i - t_{i-1})]^2.\end{aligned} \tag{5}$$

From equation (5), differentiating with respect to each of the parameters and setting the results equal to zero, we obtain the likelihood equations for a, σ^2 and γ. By performing some

calculations, we obtain the following expressions for the estimators $\hat{a}$, $\hat{\sigma}^2$ and $\hat{\gamma}$,

$$\hat{a} = \frac{[\ln(x_n - \hat{\gamma}) - \ln(x_1 - \hat{\gamma})]}{t_n - t_1}, \tag{6}$$

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^{n} \frac{[\ln(x_i - \hat{\gamma}) - \ln(x_{i-1} - \hat{\gamma}) - \hat{a}(t_i - t_{i-1})]^2}{t_i - t_{i-1}}, \tag{7}$$

$$\lambda(\hat{\gamma}) = \hat{\sigma}^2 \sum_{i=2}^{n} \frac{1}{(x_i - \hat{\gamma})} - \sum_{i=2}^{n} \frac{(x_i - x_{i-1})[\ln(x_i - \hat{\gamma}) - \ln(x_{i-1} - \hat{\gamma})]}{(x_i - \hat{\gamma})(x_{i-1} - \hat{\gamma})(t_i - t_{i-1})} + \hat{a} \frac{x_n - x_1}{(x_n - \hat{\gamma})(x_1 - \hat{\gamma})}, \tag{8}$$

where $\lambda(\hat{\gamma}) = 0$ and $\gamma < x_{(1)}$.

3.2 *Reparametrization of the process, according to Wingo*

Estimating the third parameter that is to be introduced into the three-parameter lognormal distribution is well known to be problematic.

The main theoretical difficulty is that the likelihood function reaches global maxima at points where the parameters present unacceptable values and where the global maximum value of the likelihood function is $+\infty$.

The main computational difficulty encountered is that the iterative numerical methods used for ML estimation must be employed with great care, as these methods do not avoid the region of attraction of infinite ML, and so problems of convergence arise. In other words, when we seek to find a solution to the equation (8) by means of a numerical approach, and as the initial estimate of γ is not close enough to the solution, this method will converge toward the degenerate solution $\gamma = -\infty$. Thus, we require an algorithm that is both computationally efficient and definitely convergent. In an attempt to resolve this problem, Wingo [11], for the case of three-parameter lognormal distribution, proposed a computational algorithm based on the reparametrization of the likelihood function, using a parametric transform, to reduce the interval of the real slope on which many finite local maxima of the log-likelihood function are located. This function is globally maximized by means of numerical methods within the reduced interval that has been found. The upper part of the search interval can be chosen depending on the precision of the data that are observed.

Let us now describe the algorithm used by Wingo and consider the same reparametrization for the case of the one-dimensional lognormal diffusion process with three parameters.

Consider the following transform

$$\gamma(\theta) = x_1 - \exp(-\theta), \quad]-\infty, +\infty[, \tag{9}$$

where x_1 is the minimum of the values in the sample.

It can be shown that $\gamma \longrightarrow x_1$ when $\theta \longrightarrow +\infty$ and that $\gamma \longrightarrow -\infty$ when $\theta \longrightarrow -\infty$, and substituting in equation (5), we obtain $\ln[\mathbb{L}(x_1, \ldots, x_n; a, \sigma^2, \gamma(\theta))]$, denoted by $\mathbb{L}^*(a, \sigma^2, \theta)$.

The computational algorithm is straightforward.

- Globally maximize $\mathbb{L}^*(a, \sigma^2, \theta)$ in a compact interval on the real slope to obtain a global maximum $\hat{\theta}$.
- The local ML estimator, $\hat{\gamma}$, is then calculated by means of

$$\hat{\gamma}(\theta) = x_1 - \exp(-\hat{\theta}). \tag{10}$$

- The remaining ML estimators are calculated by substituting the value $\hat{\gamma}$ in the equations (6) and (7).

The transform (9) can be seen, approximately, as a compression of the range of values over which the conditioned log-likelihood function must be maximized to find $\hat{\gamma}$. Obviously, this compression is computationally advantageous; it can be shown that maximizing equation (5) over γ may require up to 10 more evaluations of the objective function than when it is maximized over θ.

In the numerical data presented in ref. [11], the search interval chosen for θ was $[-10, 10]$. This choice of search range for θ has the advantage of being wide enough to cover the area in which the real value of the threshold parameter γ may be found, yet small enough as to exclude the infinite maximum of the log-likelihood function given in $\theta = +\infty$. This interval corresponds approximately to $[-20.000; x_1 - \epsilon]$ in γ-space, where $\epsilon \simeq 4 \times 10^{-5}$, a choice that may be useful for most of the samples likely to be found in practice.

Sometimes, it is useful to express the right-hand side of the search interval for θ as a function of the precision of the data. For example, if each of the observations x_i is fitted to three decimal places and the log-likelihood function has to be maximized over the interval $[-20.000, x_1 - \epsilon]$ of γ, then $\epsilon < 10^{-3}$ cannot be chosen, as the probability of the finite maximum of the log-likelihood function for γ belonging to the interval $x_1 - \epsilon < \gamma < x_1$ for $\epsilon < 10^{-3}$ is extremely remote. To reflect the precision of the data, we could choose $e < 10^{-d}$ or $\epsilon = |x_1| \times 10^{-10}$, where $d \geq 1$ is the number of decimal places of precision in the data. The corresponding upper limit of the search interval for θ could be $\theta = -\ln(10^{-d})$ or $\theta = -\ln(|x_1| \times 10^{-10})$.

By applying this transform to our process, the log-likelihood function would then be

$$\begin{aligned}\mathbb{L}^*(a, \sigma^2, \theta) = &-\frac{n-1}{2}\ln(2\pi\sigma^2) - \sum_{i=2}^{n}\ln(x_i - [x_1 - \exp(-\theta)]) \\ &- \frac{1}{2}\sum_{i=2}^{n}\ln(t_i - t_{i-1}) - \sum_{i=2}^{n}\frac{1}{2\sigma^2(t_i - t_{i-1})}[\ln(x_i - [x_1 - \exp(-\theta)]) \\ &- \ln(x_{i-1} - [x_1 - \exp(-\theta)]) - a(t_i - t_{i-1})]^2. \end{aligned} \tag{11}$$

By differentiating equation (11) with respect to θ and setting the result equal to zero, we obtain the following expression that provides an estimator for θ. By substituting this expression in the other likelihood equations, we obtain the estimators for the remaining parameters, as described above.

$$\begin{aligned}\hat{a}\left(\frac{1}{(x_n - \hat{\gamma}(\theta))} - \frac{1}{(x_1 - \hat{\gamma}(\theta))}\right) = &\sum_{i=2}^{n}\frac{\mathrm{e}^{-\theta}\hat{\sigma}^2}{(x_i - \hat{\gamma}(\theta))} \\ &- \sum_{i=2}^{n}\frac{(x_i - x_{i-1})[\ln(x_i - \hat{\gamma}(\theta))\ln(x_{i-1} - \hat{\gamma}(\theta))]}{(x_i - \hat{\gamma}(\theta))(x_{i-1} - \hat{\gamma}(\theta))(t_i - t_{i-1})}. \end{aligned} \tag{12}$$

3.3 *Simulated annealing optimization method*

The *simulated annealing* (SA) or *simulated overheating* method is a technique that has attracted considerable attention in large-scale optimization problems. It originated in the study by Metropolis *et al.* [24] aimed at minimizing a function on a very large finite set, although it could also be applied to optimization on a continuous set (see [25]). Various authors, such as

Kirkpatrick *et al.* [26], have demonstrated the utility of this method for finding global solutions to combinatory optimization problems.

The problem we address is

$$\max_{\theta\in\Theta} h(\theta) \tag{13}$$

or equivalently

$$\min_{\theta\in\Theta} -h(\theta). \tag{14}$$

The fundamental idea of the method is that a change of scale, called the temperature, allows faster movements over the surface of the function h to be maximized, the negative of which is termed energy (see [27, 28]). The optimum state, therefore, is achieved if T decreases slowly and well under control. The change in the partial scale avoids the attraction of local maxima.

To simulate the evolution of a physical system, we introduce the iterative method known as *the Metropolis acceptance rule*, which induces a change in the current state of the system in the following terms:

- If the system energy S decreases, the modification is accepted.
- If the energy increases by ΔS, the modification may be accepted with a probability of $\exp\{-\Delta S/T\}$, where T is the temperature and Δ the increase in h.

The implementation of the SA algorithm is remarkably easy. The following elements must be provided:

(i) a representation of possible solutions,
(ii) a generator of random changes in solutions,
(iii) a means of evaluating the problem functions,
(iv) an annealing schedule (an initial temperature an rules for lowering it as the search progresses).

In practice, the temperature parameter decreases by levels, generating a succession of states, enabling the system to approximate the equilibrium for each level. The algorithm stops for a small value of T, such that virtually no new generation is accepted. Thus, what is chosen as the solution to the problem is the state that optimizes the objective function among the states of the succession that is generated.

Various studies have shown that the asymptotic convergence of SA is guaranteed, but in practice it is necessary to implement the algorithm within a finite time. To do so, we define a finite succession of truncated Markov chains, associated with decreasing temperature values.

The basic structure of SA algorithm is presented in table 1. The following notations are used:

θ = the current solution,
θ^* = the best solution,
θ_n = neighboring solution,
$h(\theta)$ = the value of objective function at solution θ,
n = repetition counter,
T_0 = initial temperature,
T_f = final temperature,
L = number of repetition allowed at each temperature level,
p = probability of accepting θ_n when it is not better than θ.

The algorithm starts with an initial solution for the problem. As we can see in table 1, SA has two cycles, inner and outer. In the inner cycle, SA is repeated while $n < L$, a neighboring

Table 1. SA algorithm for minimization problem.

```
Initialize the SA control parameter (T0, L)
Select an initial solution, θ0
Set T = T0; Set θ = θ0; Set θ* = θ0; Calculate h(θ0);
While the stop cirterion is not reached do:
    Set n = 1
    While n < L do:
        Generate solution θn in the neighborhood of θ0; Calculate Δ = h(θn) − h(θ);
        If Δ ≤ 0
            θ = θn
        else
            generate a random number, r ∈ (0, 1)
            if (r ≤ p = e^{−Δ/T});
                θ = θn; n = n + 1;
            end
        end
        if (h(θ) < h(θ*))
            θ* = θn;
        end
    end
    reduce the temperature T;
end
```

solution θ_n of the current solution, θ is generated. If $\Delta \leq 0$ (θ_n is better than θ), then the generated solution replaces the current solution, otherwise the solution is accepted with the criterion probability $p = e^{-\Delta/T}$. The value of the temperature, T, decreases in each iteration of the outer cycle of the algorithm. The performance of SA depends on the definition of the several control parameters:

(i) The initial temperature should be high enough that, in the first iteration of the algorithm, the probability of accepting a worse solution is, at least, 80%.
(ii) We shall now determine all the initial values required by the algorithm and the temperature reduction rate. In the particular case in question, it is determined by means of the following expression:

$$T_i = \frac{T_0}{i+1},$$

where T_0 is the initial temperature fixed. This rate ensures that the cooling is slow enough for the optimum value to be reached.
(iii) The stopping criterion defines when the system has reached a desired energy level.

3.4 *Application of the simulated annealing optimization method*

The function to be minimized in the case in question is

$$\begin{aligned} -\ln[L(x_1,\ldots,x_n;a,\sigma^2,\gamma)] = {} & \frac{n-1}{2}\ln(2\pi\sigma^2) + \sum_{i=2}^{n}\ln(x_i-\gamma) + \frac{1}{2}\sum_{i=2}^{n}\ln(t_i-t_{i-1}) \\ & + \frac{1}{2\sigma^2}\sum_{i=2}^{n}\frac{1}{(t_i-t_{i-1})}[\ln(x_i-\gamma)-\ln(x_{i-1}-\gamma) \\ & - a(t_i-t_{i-1})]^2 \end{aligned} \tag{15}$$

By applying the reparametrization proposed by Wingo [11] to the above function, we obtain

$$
\begin{aligned}
-\ln[L(x_1,\ldots,x_n;a,\sigma^2,\theta)] = {} & \frac{n-1}{2}\ln(2\pi\sigma^2) + \sum_{i=2}^{n}\ln(x_i - [x_1 - \exp(-\theta)]) \\
& + \frac{1}{2}\sum_{i=2}^{n}\ln(t_i - t_{i-1}) + \sum_{i=2}^{n}\frac{1}{2\sigma^2(t_i - t_{i-1})} \\
& \times [\ln(x_i - [x_1 - \exp(-\theta)]) - \ln(x_{i-1} - [x_1 - \exp(-\theta)]) \\
& - a(t_i - t_{i-1})]^2,
\end{aligned}
$$

where x_1 is the minimum of the sample values.

There now arises the problem of deciding upon the new neighbor, *i.e.* how we are to choose a new value for the parameters. To do this, we generate the value of a distribution uniform, which is added to the previous value of the parameter, provided that always the new value is located within an appropriate range of possible values for the parameters. To ensure this, we must decide as to which range is acceptable for each parameter. In the present case, various graphic procedures are used, based on likelihood equations, so that we may bound the possible optimum values of the parameters.

4. Simulation studies

The stochastic differential equation (2) has a single continuous solution in the interval $[t_0, T]$, which corresponds to the three-parameter lognormal diffusion process, the explicit expression of which can be obtained by means of Itô's formula, applied to the transform $\ln(X(t) - \gamma)$, and which has the following form

$$
X(t) = \gamma + (x_0 - \gamma)\exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)(t - t_0) + \sigma(W(t) - W(t_0))\right\}. \tag{16}
$$

From this explicit solution, we can obtain the simulated trajectories of the process by discretizing the time interval $[t_0, T]$, with the initial condition $W(t_0) = 0$. The Wiener process is obtained as the sum of the distributions $\mathcal{N}(0, h)$, where $h = t_i - t_{i-1}$.

From this simulated process sample, we estimate the parameters by ML, first using the Newton–Raphson (NR) nonlinear approach to approximate the value of $\hat{\gamma}$. Secondly, we discuss the problems that occur in estimating the parameters of the three-parameter lognormal diffusion process. We propose SA optimization to the estimation of the paremeters, by which some of the above-discussed problems may be overcome.

The process is reparametrized, as described in section 3.2, in order to perform a compression of the range of values over which the conditioned log-likelihood function must be maximized to find $\hat{\gamma}$.

The parameters of the process are estimated by applying the method to the simulated data set described previously, which enable us to test the effectiveness of the method.

Table 2 shows the values used in the simulation and the results obtained by estimating the parameters, using the methods described above, implemented using the Mathematica packages by considering $h = 1$, $n = 25$ and an initial value $x_0 = 1.22139$.

These results clearly show that the SA algorithm is a good estimation method and that it enables us to eliminate many of the difficulties encountered with ML estimation.

Table 2. Starting values used in the simulation and estimation of the parameters.

	γ	μ	σ
Simulation	1	0.2	0.00010
Estimation NR	1.00006	0.200006	0.00008
Estimation SA	1.00483	0.214715	0.00091

5. Application to real data

5.1 *Description of the methodology*

We studied the evolution of a basic demographic indicator, that of the weddings taking place in a society and, specifically, the time-dependent stochastic variable:

$$Y(t) = \text{mean age of males in Spain at the date of their first wedding.}$$

The following steps were performed in the statistical methodology:

- Values were observed for the period 1982–2001 for the estimation of the parameters, reserving the values observed for the years 2002 and 2003 for comparison with the corresponding prediction by the model. The values observed correspond to observations over time intervals equal to 1 year. The source for the data was the Spanish National Institute of Statistics (INE).
- The estimations of the parameters were calculated by ML, using the expressions obtained in equations (6)–(8). The expression (8) was approximated numerically using the NR and SA methods.

Table 3. Observed values, MLCTF estimated trends for NR and SA and predicted value.

Years	Observed $Y(t)$	MLCTF-NR	MLCTF-SA
1982	26.70	26.7000	26.7000
1983	26.76	26.8303	26.7939
1984	26.90	26.8929	26.8556
1985	27.03	27.0389	26.9996
1986	27.15	27.1745	27.1333
1987	27.23	27.2997	27.2567
1988	27.39	27.3832	27.3389
1989	27.56	27.5501	27.5035
1990	27.81	27.7275	27.6783
1991	28.09	27.9883	27.9354
1992	28.35	28.2804	28.2233
1993	28.66	28.5517	28.4907
1994	28.92	28.8751	28.8095
1995	29.16	29.1464	29.0769
1996	29.46	29.3968	29.3237
1997	29.69	29.7097	29.6322
1998	29.85	29.9497	29.8687
1999	30.01	30.1166	30.0333
2000	30.16	30.2836	30.1978
2001	30.41	30.4400	30.3521
Prediction			
2002	30.63	30.7009	30.6092
2003	30.89	30.9304	30.8354

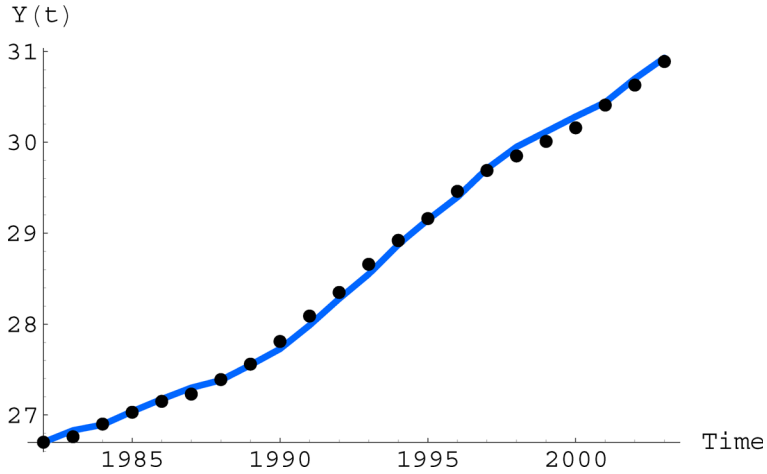

Figure 1. Fit and prediction based on MLCTF-NR.

- The ML conditional trend function of the process (MLCTF) can be obtained using Zehna's theorem (see [29]), replacing the parameter with the estimators in equation (3). Thus, we obtain two MLCTF, the first when $\hat{\gamma}$ is obtained by NR, which we denote as MLCTF-NR, and the second when $\hat{\gamma}$ is obtained by SA, which is denoted as MLCTF-SA. These two functions are used to fit and predict.

5.2 *Fitting the model using ML*

The expression (8) is approximated numerically using NR and we thus obtain the following ML estimates $\hat{\gamma} = 23.6911$, $\hat{\mu} = 0.042381$ and $\hat{\sigma} = 0.0140717$.

After estimating the parameters for the variable considered using SA, the results achieved are as follows: $\hat{\gamma} = 23.3911$, $\hat{\mu} = 0.027979$, $\hat{\sigma} = 0.0007018$.

Table 3 shows the observed values of the variable and those of the estimated ML conditioned trend functions using NR and SA.

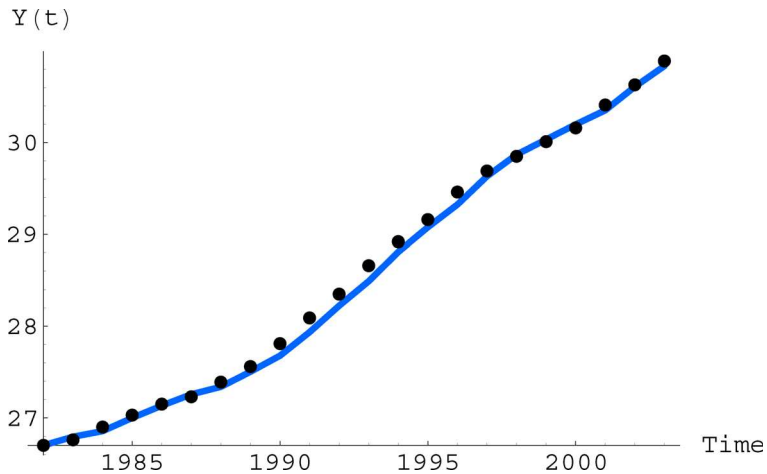

Figure 2. Fit and prediction based on MLCTF-SA.

Figure 1 shows the fit and the prediction for $Y(t)$ using the MLCTF-NR. Figure 2 shows the fit and the prediction for $Y(t)$ using the MLCTF-SA.

6. Conclusions

The fundamental goal of this study is to introduce and examine a stochastic three-parameter lognormal diffusion process. To do this, we determine the basic probabilistic results (section 2) and inferential results, in particular estimation based on discrete sampling of the process using ML (section 3). The problems that arise in applying ML estimation with a threshold parameter are addressed via an extension of Wingo's reparametrization technique in the case of a three-parameter lognormal distribution to the process proposed in this study. The main conclusion reached, from a computational standpoint is that when this reparametrization method is used, calculating the estimator of the theshold parameter of the process using numerical methods based on the reparametrized equation (12) and obtaining it by the alternative SA methodology produce similar results, regarding both simulated and real data.

The tables 4 and 5 show the results obtained using the two estimation methods described; table 4 shows the process parameters from the simulation of $X(t)$ and table 5 shows those estimated from the observed data of process $Y(t)$.

The small discrepancies observed between the two methods could be reduced by adjusting the precision and number of iterations when using the SA algorithm.

We calculated a global descriptive measure of the discrepancy between the adjusted values and those observed, using the sampling variance; in the case of NR estimation, the value was 0.003, and for estimation by the SA method it was 0.004, thus confirming the similarity of the two methodologies.

Although the proposed method is still being refined, the results obtained lead us to believe that it constitutes a good alternative to existing methods, as both its theoretical basis and its implementation or adaptation to any type of problem are much more straightforward. Moreover, most existing methods depend to a large extent on the initial solution. With the proposed method, however, if the Markov chain lengths produced are long enough, there is no such dependence, although this fact does mean that more time must be applied in searching for the solution.

Table 4. Estimation of the parameters for the set of simulated data.

	$\hat{\gamma}$	$\hat{\mu}$	$\hat{\sigma}$
Estimation NR	1.00006	0.200006	0.00008
Estimation SA	1.00483	0.214715	0.00091

Table 5. Estimation of the parameters for the observed $Y(t)$.

	$\hat{\gamma}$	$\hat{\mu}$	$\hat{\sigma}$
Estimation NR	23.6911	0.042381	0.0140717
Estimation SA	23.3911	0.027979	0.0007018

Acknowledgements

The authors are very grateful to the Editor and referees for constructive comments and suggestions. This work was supported partially by research project MTM 2005-09209, Ministerio de Educación y Ciencia, Spain.

References

[1] Wicksell, S.D., 1917, On logarithmic correlation with an application to the distribution of ages at first marriage. *Meddelande fran lunds Astronomiska Observatorium*, **84**.
[2] Guerrieri, G., 1969–1970, Sopra un nuovo metodo concernente la determinazione dei parametri della distribuzione lognormale e della distribuzioni pearsiane del III e del vee tipo. *J. Annali dell'instituto di statistica della universita di BARI*, 53–110.
[3] Aitchison, J. and Brown, J.A.C., 1980, *The Lognormal Distribution* (Cambridge University Press).
[4] Royston, P., 1992, Estimation, reference ranges and goodness of fit for the three-parameter lognormal distribution. *Statistics in Medicine*, **11**, 897–912 (John Wiley & Sons).
[5] Crawford, S.L., 1994, An application of the laplace method to finite mixture distributions. *Journal of American Statistical Association*, **89**(425), 259–267.
[6] Cohen, A.C., 1951, Estimating parameters of logaritmic-normal distributions by maximum likelihood. *Journal of American Statistical Association*, **46**, 206–212.
[7] Harter, H.L. and Moore, A.H., 1966, Local-maximum likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples. *Journal of American Statistical Association*, **61**, 842–851.
[8] Hill, B.M., 1963, The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic. *Journal of American Statistical Association*, **58**,72–84.
[9] Calitz, F., 1973, Maximum likelihood estimation of the three-parameter lognormal distribution a reconsideration. *The Australian Journal of Statistics*, **15**, 185–190.
[10] Giesbrecht, F. and Kempthorne, O., 1976, Maximum likelihood estimation in the three-parameter lognormal distribution. *Journal of the Royal Statistical Society, Series B*, **38**, 257–264.
[11] Wingo, D.R., 1984, Fitting three-parameter lognormal models by numerical global optimization–an improved algorithm. *Computational Statistics and Data Analysis*, **2**, 13–25.
[12] Cohen, A.C. and Witten, B.J., 1980, Estimation in the three-parameter lognormal distribution. *Journal of American Statistical Association*, **75**, 399–404.
[13] Cohen, A.C., Witten, B.J. and Ding, Y., 1985, Modified moment estimation for the three-parameter lognormal distribution. *Journal of Quality Technology*, **17**, 92–99.
[14] Crow, E.L. and Shimizu, K., 1988, *Lognormal Distribution Theory and Applications* (Marcel Dekker).
[15] Lifson, D.P. and Bhattacharyya, B.B., 1983, Quantile regression method and its application to estimate the parameters of lognormal and other distributions. *Contributions to Statistics*, 313–327.
[16] Kappenman, R.F., 1985, Estimation for the three-parameter Weibull, lognormal, and gamma distributions. *Computational Statistics and Data Analysis*, **3**, 11–23.
[17] Heyde, C.C., 1963, On a property of the lognormal distribution. *Journal of the Royal Statistical Society, Series B*, **25**, 392–393.
[18] Griffiths, D.A., 1980, Interval estimation for the three-parameter lognormal distribution via the likelihood function. *Applied Statistics*, **29**, 58–68.
[19] Lambert, J.A., 1964, Estimation of parameter in the three-parameter lognormal distribution. *Australian Journal of Statistics*, **6**, 29–32.
[20] Gutiérrez, R., Román, P. and Torres, F., 1999, Inference and first-passage-time for the lognormal diffusion process with exogenous factors. Application to modelling in economics. *Applied Stochastic Models in Business and Industry*, **15**(4), 325–332.
[21] Gutiérrez, R., Román, P. and Torres, F., 2001, Inference on some parametric functions in the univariate lognormal diffusion process with exogenous factors. *Test*, **10**(2), 357–373.
[22] Gutiérrez, R., González, A. and Torres, F., 1997, Estimation in multivariate lognormal diffusion process with exogenous factors. *Applied Statistics (JRSC)*, **46**(1), 140–146.
[23] Ricciardi, L.M., 1976, On the transformation of diffusion processes into the Wiener processes. *Journal of Mathematical Analysis and Application*, **54**, 185–99.
[24] Metropolis, N.A., Rosenbluth, M., Rosenbluth, A., Teller, A. and Teller, E., 1953, Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 1087–1092.
[25] Duflo, M., 1996, *Random Iterative Models. Applications of Mathematics. I*, Vol. 34 (Berlin: Springer-Verlag) .
[26] Kirkpatrick, S., Gelatt, D. and Vecchi, M.P., 1983, Optimization by simulated annealing. *Science*, **220**, 671–680.
[27] Casella, P.R. and Robert, C.P., 1999, *Monte Carlo Statistical Methods* (New York: Springer Verlag).
[28] Ramos, E., 2005, Multidimensional three-parameter lognormal diffusions with exogenous factors. PhD thesis, University of Granada, Spain.
[29] Anderson, T.W., 1984, *An introduction to Multivarite Statistical Analysis* (2nd edn) (Wiley), pp. 63–34.
[30] Gardiner, C.W., 1990, *Handbook of Stochastic Methods for Physics, Chemistry and Natural Sciences* (2nd edn) (Berlin, Germany: Springer Verlag).

Q4 [31] Gutiérrez, R., Gutiérrez-Sánchez, R., Ramos, E., Román, P. and Torres, F., 2003, Modelo de difusión lognormal para la descripción y predicción del precio de la vivienda. *Comunicación presentada en el* 27° *Congreso Nacional de Estadística e Investigación Operativa*, Lerida.

Q4 [32] Gutiérrez, R., Gutiérrez-Sánchez, R., Ramos, E., 2004, Inference in tri-parameter lognormal multidimensional diffusion processes with exogenous factors. Proceedings of the 17th European Meeting on Cybernetics and Systems Research, Viena, Vol. 1, pp. 157–162.

Q4 [33] Kloeden, P. and Platen, E., 1992, *The Numerical Solution of Stochastic Diferential Equations* (Berlin, Germany: Springers).

Q4 [34] Prakasa Rao, B.L.S., 1999, *Statistical Inference for Diffusion Type Process* (Arnold).