Report n.165

Computing Maximum Likelihood Estimates Of a Class of One-Dimensional Stochastic Differential Equation Models from Discrete Data*

Eugene M. Cleur

Pisa, Marzo 2000

Computing Maximum Likelihood Estimates of a Class of One-Dimensional Stochastic Differential Equation Models from Discrete Data*

Eugene M. Cleur
Facoltà di Economia, Università di Pisa,
Via Ridolfi 10, 56124 Pisa, Italy.
email: cleur@ec.unipi.it

Abstract

The problem of computing the maximum likelihood estimate of the parameters of a specific class of stochastic differential equation (SDE) models with linear drift whose sample paths are observed at discrete time points is considered. This estimate is obtained as in Cleur and Manfredi (1999) by discretizing the explicit expressions for the estimates which maximize the likelihood function in continuous time, by discretizing the likelihood function through a quadrature approximation before maximizing it, and by maximizing the likelihood function of the Euler scheme approximation to the underlying continuous process. Simulation results indicate that for the constellation of parameter values considered all three approaches lead to very similar results.

Some key words: Discrete Maximum Likelihood Estimator, Euler scheme, Stochastic Differential Equation, Stochastic Integral.

*To appear in Journal of Time Series Analysis. This is a revision of the Report n. 146 in compliance with the suggestions made by the referee of JTSA

1 Introduction

Classical time series analysis has dedicated much of its attention to the modelling and estimation of data observed at discrete time points. The much used Box-Jenkins methodology is just one of many such examples. Of late, however, continuous time models expressed in the form of a Stochastic Differential Equation (SDE) have increasingly been considered and their estimation using data observed at equi-distant discrete time points is currently attracting much attention not only in theoretical, but also in applied research (see for instance Singer (1993), Polson and Roberts (1994), Broze et al (1995), Overbeck and Ryden (1997), Shoji and Ozaki (1997)).

This paper considers some computational aspects in the estimation, using maximum likelihood techniques, of the following class of one-dimensional SDE models:

$$dX_t = k(\vartheta - X_t)dt + \sigma X_t^{\beta} dW_t \tag{1}$$

when σ is unknown.

Model (1) was estimated, for instance, in Chan et al. (1992) using the generalized method of moments, and in Broze et al. (1995) and Calzolari et al. (1998) who both applied an indirect estimation procedure to correct for a so-called discretizaion bias.

Very often (1) is estimated for given values of β, such as $\beta = 0.00$ which defines the Ornstein-Uhlenbeck process (also known as the Vasicek model in financial economics), $\beta = 0.5$ which defines the Square Root process (Cox-Ingersoll-Ross model in financial economics) and $\beta = 1.00$ which defines a Geometric Brownian Motion process (Brennan-Schwarz in financial economics) thereby reducing the computational complexities in the estimation procedures.We will call such a model a "reduced model". Instead, equation (1), when β too is unknown and has to be estimated, will be labelled the "complete model". This paper will be concerned with the estimation of both the reduced and the complete model.

2 Maximum Likelihood Estimation

2.1 Reduced Model

For the following general one-dimensional SDE which includes equation (1)

$$dX_t = a(X_t, \Theta)dt + b(X_t, \sigma)dW_t \quad . \tag{2}$$

we may define the log likelihood ratio function in continuous time (see, for ex., Kloeden et al (1992)):

$$\log L(\Theta) = \int_0^T \frac{a(X_t, \Theta)}{\{b(X_t, \sigma)\}^2} dX_t - \frac{1}{2} \int_0^T \frac{\{a(X_t, \Theta)\}^2}{\{b(X_t, \sigma)\}^2} dt \tag{3}$$

For the SDE defined in (1), continuous time maximum likelihood estimates (CTMLE) of k and ϑ may be obtained conditional on given values of σ and β by maximising (3) (see Lipster and Shiryayev (1981) and Kloeden et al. (1992)) or by deriving explicit expressions for the estimators of k and ϑ as in Cleur and Manfredi (1999). For the SR and GBM models these expressions are given by (see Cleur and Manfredi (1999) for details)

$$\widehat{k} = \frac{I_1 I_5 - I_2 I_3}{I_3 I_4 - I_5^2} \quad \widehat{\theta} = \frac{I_1 I_4 - I_2 I_5}{I_1 I_5 - I_2 I_3} \tag{4}$$

where the I_j are either stochastic or non-stochastic integrals and are defined in Table I.

Table I.
Integrals in the CTMLE for the SR and GBM models

Model	I_1	I_2	I_3	I_4	I_5
SR	$\int_0^T \frac{dX_s}{X_s}$	$\int_0^T dX_s$	$\int_0^T \frac{ds}{X_s}$	$\int_0^T X_s ds$	$\int_0^T ds;$
GBM	$\int_0^T \frac{dX_s}{X_s^2}$	$\int_0^T \frac{dX_s}{X_s}$	$\int_0^T \frac{ds}{X_s^2}$	$\int_0^T ds$	$\int_0^T \frac{ds}{X_s}$

On the other hand, when the continuous time path is observed at equidistant time points, $0 = t_0$, $t_1 = t_0 + \delta$, $t_0 + 2\delta$, ... , T, where $\delta = t_i - t_{i-1}$ is known as the discretization step, various approaches are available for obtaining discrete time maximum likelihood estimates (DMLE henceforth) of k and ϑ for given values of σ and β. For instance, DMLE of k and ϑ, for given values of σ and β, may be obtained as in Cleur and Manfredi (1999) and which consists in discretizing the explicit expressions for the CTMLE of k and ϑ by means of a trapezoidal rule as reported in Tables II and III.

Table II.
Trapezoidal evaluation of the non-stochastic integrals

Model	I_3	I_4	I_5
SR	$\Delta \sum_{n=1}^{n_T} \frac{X_{n-1}^{-1}+X_n^{-1}}{2}$	$\Delta \sum_{n=1}^{n_T} \frac{X_{n-1}+X_n}{2}$	T
GBM	$\Delta \sum_{n=1}^{n_T} \frac{X_{n-1}^{-2}+X_n^{-2}}{2}$	T	$\Delta \sum_{n=1}^{n_T} \frac{X_{n-1}^{-1}+X_n^{-1}}{2}$

Table III.
Ito's formula evaluation of the stochastic integrals

Model	I_1	I_2
SR	$\log \frac{X_T}{X_0} + \frac{\sigma^2}{2} I_3$	$X_T - X_0$
GBM	$\sigma^2 I_5 - (\frac{1}{X_T} - \frac{1}{X_0})$	$\log \frac{X_T}{X_0} + \frac{\sigma^2}{2} T$

This estimator will be labelled "DMLE-Ito".

Alternatively, if the discretization step is sufficiently small, a quadrature approximation to (3) may be applied (see, for example, Florens-Zmirou (1989), Kloeden et al (1995), Pedersen (1995), Polson and Roberts (1994), Shoji (1997), and Shoji and Ozaki (1997) leading to

$$\begin{aligned} \log L(\Theta) &= \sum_{t_i \leq T} \frac{a(X_{t_{i-1}}, \Theta)}{\left\{b(X_{t_{i-1}}, \sigma)\right\}^2} (X_{t_i} - X_{t_{i-1}}) \\ &\quad -\frac{1}{2} \sum_{t_i \leq T} \frac{\left\{a(X_{t_{i-1}}, \Theta)\right\}^2}{\left\{b(X_{t_{i-1}}, \sigma)\right\}^2} (t_i - t_{i-1}) \end{aligned} \quad (5)$$

The estimator obtained by maximizing (5) is called the likelihood ratio maximum likelihood estimator in Shoji (1997) and Shoji and Ozaki (1997), but in this paper it will be labelled, for short, the "DMLE-Quad". It is, unfortunately, inconsistent as shown in Florens-Zmirou (1989); consistency is assured if $\delta \to 0$ and $T \to \infty$. Simulation results reported in this paper, however, indicate that the inconsistency is noticable only when the drift parameter k is relatively large, i.e. when there is an increase in the degree of convexity of the mean solution of the underlying continuous process.

In order to estimate the vector of parameters $\Theta = \{k, \vartheta\}$, both the "DMLE-Quad" and the "DMLE-Ito" require a preliminary knowledge of σ as well as of β which is assumed known in the reduced model. If the discretization step is sufficiently small, the following quadratic variation estimate, although inconsistent, has often been proposed (see, for ex., Polson and Roberts (1994), Shoji and Ozaki (1997))

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t_i \leq T} \frac{(X_{t_i} - X_{t_{i-1}})^2}{X_{t_{i-1}}^{2\beta}} \tag{6}$$

The inconsistency in this estimate, which is greatly dependent on the discretization step (in the limit, as the discretization step tends to zero the diffusion parameter σ is identified with probability one) as well as on the degree of convexity of the underlying process, could have negative repercussions on the properties of $\hat{k}$ as illustrated in Cleur (1999). In other words there is a vicious circle in act: when k is high, and only a finite number of data points are available, the quadratic variation estimate of σ substituted in (5) is more markedly biased and inconsistent, and this leads to a more markedly inconsistent estimate of k. The crucial role played by σ in the estimation of the drift parameter from discrete data is underlined in Prakasa Rao (1999), page 145.

The estimation procedure applied to the reduced model is as follows:

(a) the quadratic variation estimate defined in (6) is used as a preliminary value of σ,

(b) k and ϑ are computed by maximising (5) to obtain the DMLE-Quad and by applying the formulae in Tables 2 and 3 to obtain the DMLE-Ito,

(c) the final estimate of σ is obtained, for each of the two estimators in (b), by using the calculated residuals from the Euler scheme:

$$X_{t_i} = X_{t_{i-1}} + k(\vartheta - X_{t_{i-1}})\delta + \sigma X_{t_{i-1}}^{\beta} W_{t_i} \tag{7}$$

where $W_{t_i} \sim N(0, \sqrt{\delta}\,)$ and β is explicitly defined in the estimated reduced model.This final estimate of σ was also suggested in Shoji and Ozaki (1997), and Pedersen (1995), with reference to the Ornstein-Uhlenbeck process, considers it as an improvement over the quadratic variation estimate given by (6) in that it corrects for the influence of the drift.

The "DMLE-Ito" needs knowledge on solving Ito integrals and is, therefore, mathematically more demanding as against the "DMLE-Quad" which is much simpler to define and only needs a good computer program for optimizing a user supplied function. This paper also addresses the problem as to whether the disadvantage of the "DMLE-Ito" in terms of mathematics is offset by an advantage in terms of bias and efficiency.

2.2 Complete Model

In the reduced model above, β was known and this was motivated by the fact that often one has in mind the estimation of a model defined by a specific value of β. One might, however, decide to estimate β as well, for instance as a preliminary step in testing whether the observed time path was generated by a known model characterized by a specific value of β. In this case, an approximate Maximum Likelihood Estimator for the complete model may be derived as follows:

(7) can be rewritten as

$$X_{t_i} - X_{t_{i-1}} - k(\vartheta - X_{t_{i-1}})\delta = \xi_{t_i} \tag{8}$$

for which

$$\begin{aligned} \log L(x_T; k, \vartheta, \sigma, \beta) &= -T\log\sigma - \beta\sum \log X_{t_{i-1}} - \\ & \sum\left[\frac{X_{t_i} - X_{t_{i-1}} - k(\vartheta - X_{t_{i-1}})\delta}{\sigma\sqrt{\delta}X^{\beta}_{t_{i-1}}}\right]^2 \end{aligned} \tag{9}$$

Noting that the Euler scheme gives a reasonably good approximation to the time path of the underlying continuous process provided the discretization step is sufficiently small, the estimates of k, ϑ, σ, and β could be obtained by maximizing (9). This estimator, whose asymptotic properties have been considered in Shoji (1997), will be labelled "DMLE- Euler". The advantage with such an approach derives from the possibilty of estimating, simultaneously, all the parameters of the process.

An initial value for σ may be obtained from equation (6) as in the reduced model.

A simple procedure for obtaining an initial value for β could be based on the well known information criteria as follows:

Let $L_T(x_T; \hat{k}; \hat{\vartheta}; \hat{\sigma}; \beta^0)$ represent an estimate, using the data vector x_T , of the log likelihood function defined in (9) conditional on $\beta = \beta^0$. In the present context, the information criterion statistics AIC and BIC both reduce to finding that value of β^0 which maximizes $L_T(x_T; \hat{k}; \hat{\vartheta}; \hat{\sigma}; \beta^0)$. Hence, given a discrete approximation of the process, x_T, an initial estimate of β may be obtained by computing $L_T(x_T; \hat{k}; \hat{\vartheta}; \hat{\sigma}; \beta^0)$ over a wide range of values of β^0 and choosing that value which maximizes the likelihood. In simulations based on the values of $\beta = 0.3$, 0.5, 0.7 and 1.0, using values for β^0 in the range $\beta\pm$ 0.3, the AIC statistic always identified the true value.

As for k, starting values set casually in the interval (0.2, 1.0) all led to the same estimate upto the fourth decimal point, so this does not appear to pose any substantial computational problem.

The starting value for ϑ was the calculated mean of the observed process.

The DMLE-Euler may be calculated for the reduced model as well; in this case a comparison with the DMLE-Quad and DMLE-Ito defined above is made possible. Shoji (1997) (see also Prakasa Rao (1999), Section 3.4) has shown that the DMLE-Quad and the DMLE-Euler for the reduced model are asymptotically equivalent; they are consistent if $\delta \to 0$ and $T \to \infty$.The following Monte Carlo experiments should, therefore, enable us to evaluate the validity of this result for processes observed over a relatively short time interval.

3 Monte Carlo Experiments

3.1 Reduced Model

A series of computational experiments via simulation were undertaken to evaluate the performances of the three maximum likelihood estimators described above. Discrete time approximations of length T=100 and T=2000, with a fixed discretization step $\delta = 0.01$ and $X_0 = 0.10$, are simulated for the following constellation of parameter values using the order 1.5 strong Taylor scheme: $\vartheta = 0.1$; $\sigma = 0.06$; $k = 0.3$, 0.8 and 1.5 ; $\beta = 0.5$ and 1.0 in the reduced model. Each combination of these values is replicated 1000 times.

The standard errors of the estimates, reported in brackets, are the standard errors calculated from the 1000 estimates obtained for each parameter.

A preliminary control on the quadratic variation estimate of σ was carried out in the context defined by the constellation of parameter values used in this paper; the results are reported in Table IV. As can be seen, the estimator apparently provides very good starting values.

Table IV.
Quadratic Variation Estimate of σ in the SR process
$\delta = 0.01$(standard errors of estimates in brackets).

T	k=0.3	k=0.8	k=1.5
100	.0600 (.0004)	.0599 (.0004)	.0598 (.0004)
2000	.0600 (.0001)	.0599 (.0001)	.0598 (.0001)

In this and other studies, the estimation of ϑ has never posed a problem and hence any mention to this parameter will be avoided in order to save space, especially in the Tables. Tables V and VI allow us to compare the estimates for the SR process.

Table V.
Estimates of the Square Root process. T = 100
$\delta = 0.01$ (standard errors of estimates in brackets)

	$\hat{k}$			$\hat{\sigma}$		
True k	DMLE-Quad	DMLE-Ito	DMLE-Euler	DMLE-Quad	DMLE-Ito	DMLE-Euler
0.3	.3420	.3416	.3419	.0599	.0599	.0599
	(.0864)	(.0862)	(.0864)	(.0004)	(.0004)	(.0004)
0.8	.8368	.8368	.8367	.0598	.0598	.0598
	(.1275)	(.1275)	(.1275)	(.0004)	(.0004)	(.0004)
1.5	1.5268	1.5271	1.5267	.0595	.0595	.0595
	(.1703)	(.1703)	(.1703)	(.0004)	(.0004)	(.0004)

In this paper we are particularly interested in the comparative performances of the three estimators and not only in the entity of any bias present. Tables V and VI indicate a very similar behaviour of these three procedures of estimation which goes in favour of the DMLE-Euler and of the DMLE-Quad, both of which are very simple to define, over the DMLE-Ito which, as illustrated above, is more mathematically demanding.

Two principal conclusions may be drawn from these Tables: for a small T (T=100) the three estimates of the drift parameter are seriously biased and an increase in the degree of convexity of the process, i.e. in k, leads to an increase in the standard errors.

Table VI.
Estimates of the Square Root process. T = 2000
$\delta = 0.01$ (standard errors of estimates in brackets)

	$\hat{k}$			$\hat{\sigma}$		
True k	DMLE-Quad	DMLE-Ito	DMLE-Euler	DMLE-Quad	DMLE-Ito	DMLE-Euler
0.3	.3015	.3016	.3016	.0599	.0599	.0599
	(.0171)	(.0171)	(.0171)	(.0001)	(.0001)	(.0001)
0.8	.7988	.7990	.7988	.0598	.0598	.0598
	(.0276)	(.0276)	(.0276)	(.0001)	(.0001)	(.0001)
1.5	1.4906	1.4909	1.4906	.0596	.0595	.0596
	(.0375)	(.0375)	(.0375)	(.0001)	(.0001)	(.0001)

These assertions are backed by the contents of Tables VII and VIII, where the corresponding results for the GBM process are reported, as well as by the results for the well known Ornstein-Uhlenbeck process which, in order to save space, are not reported here.

Table VII.
Estimates of the GBM process. T = 100
$\delta = 0.01$ (standard errors of estimates in brackets)

	$\hat{k}$			$\hat{\sigma}$		
True k	DMLE-Quad	DMLE-Ito	DMLE-Euler	DMLE-Quad	DMLE-Ito	DMLE-Euler
0.3	.3420	.3419	.3419	.0599	.0599	.0599
	(.0867)	(.0867)	(.0867)	(.0004)	(.0004)	(.0004)
0.8	.8289	.8369	.8368	.0598	.0598	.0598
	(.1116)	(.1277)	(.1277)	(.0004)	(.0004)	(.0004)
1.5	1.5273	1.5270	1.5269	.0595	.0595	.0595
	(.1704)	(.1705)	(.1704)	(.0004)	(.0004)	(.0004)

Table VIII.
Estimates of the GBM process. T = 2000
$\delta = 0.01$ (standard errors of estimates in brackets)

	$\hat{k}$			$\hat{\sigma}$		
True k	DMLE-Quad	DMLE-Ito	DMLE-Euler	DMLE-Quad	DMLE-Ito	DMLE-Euler
0.3	.3016	.3016	.3016	.0599	.0599	.0599
	(.0173)	(.0172)	(.0172)	(.0001)	(.0001)	(.0001)
0.8	.7989	.7989	.7988	.0598	.0598	.0598
	(.0278)	(.0277)	(.0275)	(.0001)	(.0001)	(.0001)
1.5	1.4907	1.4907	1.4906	.0596	.0596	.0596
	(.0376)	(.0375)	(.0374)	(.0001)	(.0001)	(.0001)

In Tables VI and VIII, it is clear that as k increases, its estimate appears to converge towards a value less than the corresponding true value, i.e., an asymptotic bias is becoming increasingly evident. Infact, in an ad hoc simulation experiment conducted on the SR model where $k = 1.5$ and T = 20,000, all three procedures lead to $\hat{k} = 1.4880$ with a standard error of 0.1204E-1. Since the estimates of k are very sensitive to the initial value of σ, there are two possible solutions to this problem: the first consists in substituting the inconsistent quadratic variation estimate of σ with a consistent estimate if it exists (see Cleur (1999)), and the second involves a reduction of the discretization step which might not always be possible in practice, but whose positive effect through a better initial value for σ can be seen in Table IX where a partial result for the SR model is reported.

Table IX.
Estimates of the Square Root process. T = 2000
$\delta = 0.001$ (standard errors of estimates in brackets)

	$\hat{k}$			$\hat{\sigma}$		
True k	DMLE-Quad	DMLE-Ito	DMLE-Euler	DMLE-Quad	DMLE-Ito	DMLE-Euler
1.5	1.5006	1.5001	1.5006	.0600	.0600	.0600
	(.0389)	(.0374)	(.0389)	(.0000)	(.0000)	(.0000)

3.2 Complete Model

Consider now the estimation of (1) when the variance elasticity parameter, β, is unknown. Only the DMLE-Euler is applied.

Table X.
DMLE-Euler of $dX_t = k(\vartheta - X_t)dt + \sigma X_t^{\beta} dW_t$.
T = 100, $\delta = 0.01$ (standard errors in brackets)

True β	True k	$\hat{k}$	$\hat{\vartheta}$	$\hat{\sigma}$	$\hat{\beta}$
0.5	0.3	.3425 (.0864)	.0999 (.0064)	.0603 (.0043)	.5012 (.0308)
	0.8	.8368 (.1276)	.1000 (.0024)	.0601 (.0067)	.4997 (.0483)
	1.5	1.5267 (.1703)	.1000 (.0013)	.0600 (.0091)	.4982 (.0655)
1.0	0.3	.3420 (.0878)	.1000 (.0020)	.0616 (.0141)	1.0001 (.0991)
	0.8	.8369 (.1278)	.1000 (.0008)	.0638 (.0229)	1.0013 (.1538)
	1.5	1.5268 (.1704)	.1000 (.0004)	.0663 (.0322)	.9988 (.2049)

Table XI.
DMLE-Euler of $dX_t = k(\vartheta - X_t)dt + \sigma X_t^{\beta} dW_t$.
T = 2000, $\delta = 0.01$ (standard errors in brackets),

True β	True k	$\hat{k}$	$\hat{\vartheta}$	$\hat{\sigma}$	$\hat{\beta}$
0.5	0.3	.3020 (.0171)	.1000 (.0014)	.0601 (.0009)	.5004 (.0064)
	0.8	.8021 (.0277)	.1000 (.0005)	.0601 (.0015)	.5005 (.0105)
	1.5	1.5018 (.0376)	.1000 (.0003)	.0601 (.0020)	.5005 (.0143)
1.0	0.3	.3016 (.0172)	.1000 (.0005)	.0600 (.0028)	.9998 (.0206)
	0.8	.7989 (.0278)	.1000 (.0002)	.0596 (.0046)	.9973 (.0333)
	1.5	1.4906 (.0375)	.1000 (.0001)	.0591 (.0062)	.9939 (.0456)

The estimates of k in the complete model are almost identical to those in the reduced model, both in terms of bias as well as in terms of standard errors. The estimates of σ, however, undergo important changes when passing from the reduced model to the complete model; briefly, for a small T (T = 100) there is a notable increase in the standard errors which, when $\beta = 1.0$, is also accompanied by the presense of a large positive bias. This bias reduces considerably when T increases to 2000, but the standard error of σ continues to be between ten to sixty times the corresponding values in the reduced model.

4 Computer time

The simulation experiments carried out in this paper indicate that the three maximum likelihood estimators considered, which are asymptotically equivalent, have a very similar behaviour even for processes observed over a relatively short time interval. The above results also show that estimation of stochastic differential models from discrete data is not difficult, but estimation performance is a more important issue. Selection between the methods considered must therefore be made in terms of computation time and mathematical convenience.

The DMLE-Ito in terms of computer time is the most efficient, but it is mathematically the most demanding in that it is based on the derivation of the explicit expressions which also involve stochastic integrals for the estimates of the unknown parameter. The DMLE-Quad overcomes this difficulty, but is much slower since it requires the maximization of a likelihood function. The DMLE-Euler procedure which is also based on the maximization of a likelihood function is the least efficient procedure in terms of computer time, since convergence is reached after a significantly greater number of iterations, and therefore cannot be recommended for the estimation of the reduced model.

Overall, the above results tend to suggest that the DMLE-Quad might be a candidate for estimation in situations which are analytically difficult such as a nonlinear drift function.

5 Acknowledgements

The author especially thanks Piero Manfredi for his comments on an earlier draft of this paper. A word of thanks also goes to the referee for suggesting improvements over an earlier version of this paper.

This research is part of a project financed by the Italian Ministry for University and Research.

References

[1] Ait-Sahalia, Y. (1998), Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approach, Mimeo.

[2] Ball, C. A. and Torous, W. N. (1996), Unit roots and the estimation of interest rate dynamics, Journal of Empirical Finance, 3, 215-238.

[3] Broze L., Scaillet O.and Zakoian J.-M. (1995), Testing for continuous-time models of the short term interest rate, "Journal of Empirical Finance", 2, 199-223.

[4] Calzolari, G., Di Orio, F. and Fiorentini, G. (1998), Control variates for variance reduction in indirect inference: interest rate models in continuous time, Econometric Journal, Vol.1, C100-C112.

[5] Chan, K. C., Karolyi, G. A., Longstaff, F. A. and Sanders, A. B. (1992), An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, Journal of Finance, 47, 1209-1227.

[6] Cleur, E. M.and Manfredi, P. (1999), One dimensional SDE models, low order numerical methods and simulation based estimation: a comparison of alternative estimators, Computational Economics, 13, 477-497.

[7] Cleur, E. M. (1999), Estimating a class of diffusion models: An evaluation of the effects of sampled discrete observations, Working Paper 135, Dipartimento di Statistica e Matematica Applicata all'Economia, Università di Pisa, February, 1-16.

[8] Florens-Zmirou, D. (1989), Approximate discrete time schemes for statistics of diffusion processes, Statistics, 20, 547-557.

[9] Gard, T. (1987), Stochastic differential equations, Marcel Dekker, New York.

[10] Johnson, N. and Kotz, T. (1992), Continuous Univariate Distributions, Wiley, New York.

[11] Kloeden, P.E. and Platen, E. (1992), Numerical Solution of SDE, Springer Verlag, New York.

[12] Kloeden P.E., Platen, E. and Schurz, H. (1995), Numerical Solution of SDE through Computer Experiments, Springer Verlag, New York.

[13] Liptser, L.S. and Shiryayev, A. N. (1981), Statistics of random Processes, Springer Verlag, New York.

[14] Overbeck, L. and Rydén, T. (1997), Estimation in the Cox-Ingersoll-Ross Model, Econometric Theory, 13, 430-461.

[15] Pedersen, A.R. (1995), A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations, Scandinavian Journal of Statistics, Vol. 22, n. 1, 55-71.

[16] Polson, N. G. and Roberts, G. O. (1994), Bayes factors for discrete observations from diffusion processes, Biometrika, Vol. 81,1, 11-26.

[17] Prakasa Rao, B. L. S. (1999), Statistical Inference for Diffusion Type Processes, Kendall's Library of Statistics 8, Arnold, London.

[18] Shoji, I. (1997), A note on asymptotic properties of the estimator derived from the Euler method for diffusion processes at discrete times, Statistics and Probability Letters, 36, 153-159.

[19] Shoji, I. and Ozaki, T. (1997), Comparative Study of Estimation Methods for Continuous Time Stochastic Prosesses, Journal of Time Series Analysis, Vol.18, n.5, 485-506.

[20] Singer, H. (1993), Continuous-Time Dynamical Systems with Sampled Data, Errrors of Measurement and Unobserved Components, Journal of Time Series Analysis, 5, 527-546.