Pseudo Maximum Likelihood and Moments Estimators for Some Ergodic Diffusions

Pedro Mota and Manuel L. Esquível

Abstract When $(X_t)_{t\geq 0}$ is an ergodic process, the density function of X_t converges to some invariant density as $t \to \infty$. We will compute and study some asymptotic properties of pseudo moments estimators obtained from this invariant density, for a specific class of ergodic processes. In this class of processes we can find the Cox-Ingersoll & Ross or Dixit & Pindyck processes, among others. A comparative study of the proposed estimators with the usual estimators obtained from discrete approximations of the likelihood function will be carried out.

1 Introduction

Ergodic diffusion processes like the Cox-Ingersoll & Ross [3], the geometric Ornstein-Uhlenbeck or Dixit & Pindyck [4] are widely used in the mathematical finance context, see [2] or [4].

Many times, for ergodic diffusions, the transition density is not known and the parameter estimation is made using approximations of the likelihood function based in some kind of discretization or using martingale estimating functions, see, for instance, [1, 5–7, 12]. In [10], a new parameter estimation technique was presented and applied to the stochastic processes satisfying the following stochastic differential equation,

$$dX_t = b(a - X_t)X_t^{\gamma}\,dt + \sigma\sqrt{X_t^{\gamma+1}}\,dB_t, \quad a, b, \sigma > 0, \gamma \geq 0, \tag{1}$$

and for the combination of parameters that makes this processes ergodic.

P. Mota (✉) · M. L. Esquível
Centro de Matemática e Aplicações (CMA), Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, Portugal
e-mail: pjpm@fct.unl.pt; mle@fct.unl.pt

T. A. Oliveira et al. (eds.), *Recent Studies on Risk Analysis and Statistical Modeling*, Contributions to Statistics,
https://doi.org/10.1007/978-3-319-76605-8_24

The idea was that if $(X_t)_{t\geq 0}$ is an ergodic process then as $t \to \infty$ the density function of X_t converges to the invariant density and then the process parameters can be estimated from the invariant density as if the observations, $X_1, \ldots, X_n$, of the process were independent and identically distributed (i.i.d.) random variables, all of them with the same invariant distribution.

Also in [10], pseudo maximum likelihood estimators were deduced from the invariant density and in the present work we will compute pseudo moments estimators and their asymptotic properties will be studied. In the final section a comparative study, through simulation, will be implemented to compare the pseudo moments estimators with the pseudo maximum likelihood estimators already mentioned and also with the usual estimators obtained from discrete approximations of the transition density.

2 Ergodicity

A continuous time diffusion process

$$dX_t = \mu(X_t; \boldsymbol{\theta})dt + \sigma(X_t; \boldsymbol{\theta})dB_t,$$

with state space $\mathbb{R}$, is said to be ergodic (see, for instance, [9]), if

$$S(x; \boldsymbol{\theta}) = \int_{x_0}^{x} \exp\left(-2\int_{x_0}^{y} \frac{\mu(v; \boldsymbol{\theta})}{\sigma^2(v; \boldsymbol{\theta})}dv\right) dy \to \pm\infty, \text{ as } x \to \pm\infty,$$

and

$$M(\boldsymbol{\theta}) = \int_{-\infty}^{+\infty} \frac{1}{\sigma^2(x; \boldsymbol{\theta})} \exp\left(2\int_{x_0}^{x} \frac{\mu(v; \boldsymbol{\theta})}{\sigma^2(v; \boldsymbol{\theta})}dv\right) dx < \infty,$$

with x_0 an interior point of the state space.
The invariant density is then given by

$$f_{\boldsymbol{\theta}}(x) = \frac{1}{M(\boldsymbol{\theta})\sigma^2(x; \boldsymbol{\theta})} \exp\left(2\int_{x_0}^{x} \frac{\mu(v; \boldsymbol{\theta})}{\sigma^2(v; \boldsymbol{\theta})}dv\right).$$

Theorem 1 *The processes satisfying the stochastic differential equation (1) are ergodic, when* $2ab > \sigma^2(\gamma + 1)$*, with invariant density,*

$$f_{(\alpha,\beta)}(x) = \frac{x^{\alpha-1}e^{-\beta x}\beta^{\alpha}}{\Gamma(\alpha)} \sim Gamma(\alpha, \beta), \text{ with } \alpha = \frac{2ab}{\sigma^2} - \gamma, \quad \beta = \frac{2b}{\sigma^2}.$$

Proof The processes have state space $]0, \infty[$ and with, $\boldsymbol{\theta} = (a, b, \gamma, \sigma)$, we have

$$S(x;\boldsymbol{\theta}) = \int_{x_0}^{x} s(y,\boldsymbol{\theta})dy = \int_{x_0}^{x} \exp\left(-2\int_{x_0}^{y} \frac{b(a-v)v^{\gamma}}{\sigma^2 v^{\gamma+1}}dv\right)dy$$
$$= x_0^{\frac{2ab}{\sigma^2}} e^{-\frac{2b}{\sigma^2}x_0} \int_{x_0}^{x} y^{-\frac{2ab}{\sigma^2}} e^{\frac{2b}{\sigma^2}y} dy \to \pm\infty, x \to +\infty, x \to 0,$$

and

$$M(\boldsymbol{\theta}) = \int_0^{+\infty} \frac{1}{\sigma^2 x^{\gamma+1}} \exp\left(2\int_{x_0}^{x} \frac{b(a-v)v^{\gamma}}{\sigma^2 v^{\gamma+1}}dv\right)dx$$
$$= \frac{x_0^{-\frac{2ab}{\sigma^2}} e^{\frac{2b}{\sigma^2}x_0}}{\sigma^2} \int_0^{\infty} x^{\frac{2ab}{\sigma^2}-\gamma-1} e^{-\frac{2b}{\sigma^2}x} dx < \infty, \text{ if } 2ab > \sigma^2(\gamma+1).$$

The invariant density is then given by

$$f_{\boldsymbol{\theta}}(x) = \frac{x_0^{-\frac{2ab}{\sigma^2}} e^{\frac{2b}{\sigma^2}x_0}}{\sigma^2} x^{\frac{2ab}{\sigma^2}-\gamma-1} e^{-\frac{2b}{\sigma^2}x} \left(\frac{x_0^{-\frac{2ab}{\sigma^2}} e^{\frac{2b}{\sigma^2}x_0}}{\sigma^2} \int_0^{\infty} x^{\frac{2ab}{\sigma^2}-\gamma-1} e^{-\frac{2b}{\sigma^2}x} dx\right)^{-1}$$
$$= \frac{x^{\alpha-1}e^{-\beta x}\beta^{\alpha}}{\Gamma(\alpha)} \sim Gamma(\alpha,\beta), \text{ with } \alpha = \frac{2ab}{\sigma^2} - \gamma,\ \beta = \frac{2b}{\sigma^2},$$

completing the proof.

3 Estimators and Consistency

If we are working with a strictly stationary ergodic process (for instance, if X_0 have already the invariant distribution), then for any $t > 0$ the random variable X_t will have the invariant distribution. In this framework we propose to deal with the observations of the process like if they were identically distributed with the invariant distribution and then use the invariant density for estimation purposes. From the previous section we know that the processes satisfying the stochastic differential equation (1) with $2ab > \sigma^2(\gamma+1)$ are ergodic with the invariant density $Gamma(\alpha, \beta)$, where $\alpha = \frac{2ab}{\sigma^2} - \gamma$, $\beta = \frac{2b}{\sigma^2}$.

In the following, let us suppose that we have observations $X_1, \ldots, X_n$ of the process, collected at equally spaced times $t_1 < \ldots < t_n$ and that γ and σ are known parameters, that is, the only parameters of interest for estimation purposes are a and b.

3.1 Pseudo Maximum Likelihood Estimators

We can compute pseudo maximum likelihood estimators, that is, defining the likelihood function

$$f_{X_1,\ldots,X_n}(\alpha,\beta;x_1,\ldots,x_n) := \prod_{i=1}^{n} f_{X_i}(\alpha,\beta;x_i)$$

just like in the case of i.i.d. observations and where f_{X_i} is the $Gamma(\alpha,\beta)$ density of Eq. (1).

Since

$$\forall i = 1,\ldots,n, \quad f_{X_i}(\alpha,\beta;x_i) = \frac{x_i^{\alpha-1}e^{-\beta x_i}\beta^{\alpha}}{\Gamma(\alpha)}$$

we get the likelihood function,

$$L(\alpha,\beta;X_1,\ldots,X_n) = \prod_{i=1}^{n}\frac{X_i^{\alpha-1}e^{-\beta X_i}\beta^{\alpha}}{\Gamma(\alpha)}$$

and the log-likelihood,

$$\log(L(\alpha,\beta;X_1,\ldots,X_n)) = (\alpha-1)\sum_{i=1}^{n}\log(X_i) - \beta\sum_{i=1}^{n}X_i + n\alpha\log(\beta) - n\log(\Gamma(\alpha)).$$

From differentiating the log-likelihood function and equating to zero, we get (with $\psi(.)$ the digamma function),

$$\frac{1}{n}\sum_{i=1}^{n}\log(X_i) + \log\left(\frac{2b}{\sigma^2}\right) - \psi\left(\frac{2\bar{X}_n b}{\sigma^2}\right) = 0 \tag{2}$$

with $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and

$$a = \bar{X}_n + \frac{\sigma^2\gamma}{2b}$$

getting the estimator $\hat{b}_n$ of b as the solution of the Eq. (2) and the estimator of a, as

$$\hat{a}_n = \bar{X}_n + \frac{\sigma^2\gamma}{2\hat{b}_n}.$$

Theorem 2 *If* $2ab > \sigma^2(\gamma+1)$, *the pseudo maximum likelihood estimators* $\hat{a}_n$ *and* $\hat{b}_n$ *are almost sure (a.s.) consistent estimators for a and b.*

Proof The proof of the theorem can be found in [10].

3.2 Pseudo Moments Estimators

We can obtain the moments estimators for a and b, using the invariant gamma density and by solving the equations,

$$\begin{cases} \bar{X}_n &= \frac{\alpha}{\beta} \\ M_{2,n} &= \frac{\alpha+\alpha^2}{\beta^2} \end{cases},$$

with $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ the sample mean and $M_{2,n} = \frac{1}{n}\sum_{i=1}^n X_i^2$ the empirical second moment.

Solving these equations we get the moments estimators for the parameters a and b,

$$\tilde{a}_n = \bar{X}_n + \frac{M_{2,n} - \bar{X}_n^2}{\bar{X}_n}\gamma \quad \wedge \quad \tilde{b}_n = \frac{\sigma^2 \bar{X}_n}{2(M_{2,n} - \bar{X}_n^2)}$$

or using the (non-central) sample variance $S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$,

$$\tilde{a}_n = \bar{X}_n + \frac{S_n^2}{\bar{X}_n}\gamma \quad \wedge \quad \tilde{b}_n = \frac{\sigma^2 \bar{X}_n}{2S_n^2}.$$

We have the following result about the consistency of the pseudo moments estimators.

Theorem 3 *If* $2ab > \sigma^2(\gamma+1)$, *the pseudo moments estimators* $\tilde{a}_n$ *and* $\tilde{b}_n$ *are a.s. consistent estimators for a and b.*

Proof Suppose that ξ is a random variable with the invariant gamma density $Gamma(\alpha, \beta)$, where $\alpha = \frac{2ab}{\sigma^2} - \gamma$, $\beta = \frac{2b}{\sigma^2}$. It is straightforward to prove the consistency of both estimators, since, using the ergodic theorem, we have that

$$\lim_{n\to\infty} \bar{X}_n = \mathbb{E}[\xi] = \frac{\alpha_0}{\beta_0}, \quad a.s.$$

and

$$\lim_{n\to\infty} S_n^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2 = \mathbb{V}[\xi] = \frac{\alpha_0}{\beta_0^2}, \quad a.s.$$

Then,

$$\lim_{n\to\infty}\left(\bar{X}_n + \frac{S_n^2}{\bar{X}_n}\gamma\right) = \frac{\alpha_0}{\beta_0} + \frac{\gamma}{\beta_0} = a_0 \quad a.s.$$

and

$$\lim_{n\to\infty}\frac{\sigma^2\bar{X}_n}{2S_n^2} = \frac{\beta_0\sigma^2}{2} = b_0 \quad a.s.$$

proving the consistency of the estimators.

Remark 1 We have assumed that σ is known, if σ is unknown the problem of estimating σ can be solved using the quadratic variation of the process, and following [11] we get, a estimator for σ^2,

$$\hat{\sigma}_{1,n}^2 = \frac{\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2}{\sum_{i=1}^{n-1} X_i^{\gamma+1}\Delta_n},$$

or following [12]

$$\hat{\sigma}_{2,n}^2 = \frac{1}{T}\sum_{i=1}^{n-1}\frac{(X_{i+1} - X_i)^2}{X_i^{\gamma+1}}.$$

4 Simulation and Data Analysis

In this section, we will compare through simulation the moments estimators with the approximate maximum likelihood estimators presented in [10] and the estimators based in discrete approximations of the log-likelihood function. We will suppose that the observations are equally spaced, that is, $t_{i+1} - t_i = \Delta, i = 1, \ldots, n$.

The estimators for a and b based on the discretized continuous-time likelihood function, see [1] or [7], $\check{a}_n$ and $\check{b}_n$, are given by:

$$\check{a}_n = \frac{\sum_{i=1}^{n-1}\frac{X_{i+1}-X_i}{X_i}\sum_{i=1}^{n-1} X_i^{\gamma+1} - (X_n - X_1)\sum_{i=1}^{n-1} X_i^{\gamma}}{\sum_{i=1}^{n-1}\frac{X_{i+1}-X_i}{X_i}\sum_{i=1}^{n-1} X_i^{\gamma} - (X_n - X_1)\sum_{i=1}^{n-1} X_i^{\gamma-1}}$$

and

$$\check{b}_n = \frac{1}{\Delta}\frac{\sum_{i=1}^{n-1}\frac{X_{i+1}-X_i}{X_i}\sum_{i=1}^{n-1} X_i^{\gamma} - (X_n - X_1)\sum_{i=1}^{n-1} X_i^{\gamma-1}}{\sum_{i=1}^{n-1} X_i^{\gamma-1}\sum_{i=1}^{n-1} X_i^{\gamma+1} - \left(\sum_{i=1}^{n-1} X_i^{\gamma}\right)^2}.$$

For simulation purposes, we will perform the generation of the trajectories of the processes using the approximation strong Taylor scheme of order 1.5, see [8].

The iterative scheme used is the following:

$$
\begin{aligned}
Y_{i+1} = & Y_i + b(a - Y_i)Y_i^{\gamma}\Delta + \sigma Y_i^{\frac{\gamma+1}{2}}\Delta B \\
& + \frac{\sigma^2(\gamma+1)}{4}Y_i^{\gamma}((\Delta B)^2 - \Delta) + \sigma b(\gamma(a - Y_i) - Y_i)Y_i^{\frac{3\gamma-1}{2}}\Delta Z \\
& + \frac{1}{2}\left(b^2(a - Y_i)(\gamma(a - Y_i) - Y_i) + \frac{1}{2}b\gamma\sigma^2(\gamma(a - Y_i) - a - Y_i)\right)Y_i^{2\gamma-1}\Delta^2 \\
& + \left(\frac{\sigma b}{2}(a - Y_i) + \frac{\sigma^3(\gamma-1)}{8}\right)(\gamma+1)Y_i^{\frac{3\gamma-1}{2}}(\Delta B\Delta - \Delta Z) \\
& + \frac{\sigma^3\gamma(\gamma+1)}{4}Y_i^{\frac{3\gamma-1}{2}}\left(\frac{1}{3}(\Delta B)^3 - \Delta\right)\Delta B,
\end{aligned}
$$

where $\Delta B = \sqrt{\Delta}U_1$, $\Delta Z = \frac{1}{2}\Delta^{3/2}(U_1 + U_2/\sqrt{3})$ and U_1 and U_2 are independent $N(0,1)$ random variables.

We simulated 500 trajectories and for the estimation of the parameter a we present the results when $n = 500$ in each trajectory, for the parameter b we considered $n = 250$, 500, and 1000 observations in each trajectory. We estimated a and b using the pseudo moments estimators $\tilde{a}_n$ and $\tilde{b}_n$ and we compared them with the pseudo maximum likelihood estimators $\hat{a}_n$ and $\hat{b}_n$ and the estimators $\check{a}_n$ and $\check{b}_n$ obtained from the discretized likelihood function.

We considered $\sigma = 0.1$, we present Table 1 for $\gamma = 0$ and Table 2 for $\gamma = 1$ (for other values of γ we get very similar results), the true value for a is always 1 and for b we considered the values 0.1, 0.5, 1, and 2.

Table 1 Mean and S.D. (standard deviation) for the estimators of a and b when $\gamma = 0$

		$b = 0.1$		$b = 0.5$		$b = 1$		$b = 2$	
Num. obs.	Estimator	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
500	$\tilde{a}_n$	0.9997	0.0447	1.0001	0.0091	1.0001	0.0047	1.0001	0.0026
	$\hat{a}_n$	0.9997	0.0447	1.0001	0.0091	1.0001	0.0047	1.0001	0.0026
	$\check{a}_n$	0.9997	0.0455	1.0001	0.0092	1.0001	0.0047	1.0001	0.0026
250	$\tilde{b}_n$	0.1207	0.0347	0.5147	0.0734	1.0136	0.1151	2.0182	0.2033
500		0.1099	0.0228	0.5051	0.0512	1.0015	0.0819	1.9986	0.1469
1000		0.1037	0.0150	0.4981	0.0351	0.9927	0.0565	1.9851	0.1029
250	$\hat{b}_n$	0.1215	0.0338	0.5228	0.0724	1.0293	0.1137	2.0480	0.2009
500		0.1111	0.0223	0.5125	0.0504	1.0157	0.0809	2.0253	0.1452
1000		0.1050	0.0147	0.5051	0.0347	1.0060	0.0558	2.0099	0.1017
250	$\check{b}_n$	0.1125	0.0308	0.4048	0.0495	0.6400	0.0573	0.8705	0.0611
500		0.1045	0.0211	0.4007	0.0354	0.6375	0.0406	0.8683	0.0428
1000		0.0995	0.0138	0.3967	0.0249	0.6342	0.0284	0.8657	0.0300

Table 2 Mean and S.D. (standard deviation) for the estimators of a and b when $\gamma = 1$

		$b = 0.1$		$b = 0.5$		$b = 1$		$b = 2$	
Num. obs.	Estimator	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
500	$\tilde{a}_n$	0.9988	0.0462	1.0001	0.0093	1.0001	0.0048	1.0001	0.0026
	$\hat{a}_n$	0.9980	0.0461	0.9999	0.0093	1.0000	0.0048	1.0000	0.0026
	$\check{a}_n$	0.9978	0.0469	1.0000	0.0093	1.0001	0.0048	1.0001	0.0027
250	$\tilde{b}_n$	0.1213	0.0350	0.5169	0.0722	1.0182	0.1143	2.0273	0.2038
500		0.1104	0.0227	0.5075	0.0505	1.0061	0.0816	2.0074	0.1474
1000		0.1041	0.0148	0.5007	0.0347	0.9976	0.0564	1.9943	0.1034
250	$\hat{b}_n$	0.1233	0.0349	0.5253	0.0717	1.0342	0.1131	2.0573	0.2013
500		0.1125	0.0230	0.5150	0.0500	1.0204	0.0807	2.0343	0.1457
1000		0.1059	0.0151	0.5079	0.0346	1.0110	0.0558	2.0192	0.1022
250	$\check{b}_n$	0.1141	0.0322	0.4035	0.0496	0.6356	0.0572	0.8615	0.0609
500		0.1051	0.0217	0.3994	0.0357	0.6330	0.0410	0.8588	0.0430
1000		0.0997	0.0141	0.3955	0.0250	0.6296	0.0285	0.8561	0.0300

In all the outputs, we can see that the proposed estimators for a and b give good results, very close to the approximate maximum likelihood estimators and we can also see that the estimator for b, $\check{b}_n$ based in the score function only produce good results when the true value of b is 0.1 (small).

5 Conclusion

In this paper we proposed moments estimators for some ergodic processes. The consistency proof of the proposed estimators and a simulation study to show the applicability of the estimators were provided. For future research, we have the open problem of proving the normality of the asymptotic distribution of the estimators.

Acknowledgements This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações).

References

1. Bibby, B.M., Sørensen, M.: Martingale estimation functions for discretely observed diffusion processes. Bernoulli **1**(1/2), 17–39 (1995)
2. Bjork, T.: Arbitrage Theory in Continuous Time. Oxford University Press, New York (1998)
3. Cox, J., Ingersoll, J., Ross, S.: A theory of the term structure of interest rates. Econometrica **53**(2), 385–408 (1985)
4. Dixit, A.K., Pindyck, R.S.: Investment Under Uncertainty. Princeton University Press, New Jersey (1994)

5. Florens-Zmirou, D.: Approximate discrete-time schemes for statistics of diffusion processes. Statistics **20**(4), 547–557 (1989)
6. Genon-Catalot, V.: Maximum contrast estimation for diffusion processes from discrete observations. Statistics **21**(1), 99–116 (1990)
7. Kessler, M.: Estimation of an ergodic diffusion from discrete observations. Scand. J. Stat. **24**(2), 211–229 (1997)
8. Kloeden, P., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin (1999)
9. Kutoyants, Y.A.: Statistical Inference for Ergodic Diffusion Processes. Springer, London (2004)
10. Mota, P., Esquível, M.L.: Estimation using the invariant density for some ergodic processes (2018, submitted)
11. Phillips, P.C.B., Yu, J.: A two-stage realized volatility approach to estimation of diffusion processes with discrete data. J. Econometrics **150**, 139–150 (2009)
12. Yoshida, N.: Estimation for diffusion processes from discrete observations. J. Multivar. Anal. **41**(2), 220–242 (1992)