Parameter estimation of diffusion models*

Miljenko Huzak†

Abstract. *Parameter estimation problems of diffusion models are discussed. The problems of maximum likelihood estimation and model selections from continuous observations are illustrated through diffusion growth model which generalizes some classical ones.*

Key words: *diffusion growth model, diffusion process, stochastic differential equation, maximum likelihood estimation, model selection*

Sažetak. Procjena parametara difuzijskih modela rasta. *Razmatrani su problemi procjene parametara difuzijskih modela. Problemi procjene metodom maksimalne vjerodostojnosti i odabira modela na osnovi neprekidnih opservacija ilustrirani su na difuzijskom modelu rasta koji je poopćenje nekih klasičnih modela rasta.*

Ključne riječi: *difuzijski model rasta, difuzijski proces, stohastička diferencijalna jednadžba, procjenjivanje metodom maksimalne vjerodostojnosti, odabir modela*

1. Introduction

Diffusion processes have been used in modeling of various phenomena of noisy growth. For example, modeling of tumor growth (see e.g. [1, 12, 13, 24]) and interest rates (see e.g. [6, 7, 8, 9]).

A *parametric stochastic diffusion model of population growth* (briefly *diffusion growth model*) is introduced by a stochastic differential equation (SDE) (see [22]) of the form

$$dX_t = \mu(X_t, \vartheta)\, dt + \nu(X_t, \sigma)\, dW_t, \quad X_0 = x_0, \tag{1}$$

where $x_0 > 0$ is an initial size of the population; $\mu(\cdot, \vartheta)$ is a function from a deterministic growth model (see e.g. [2]) given by an ordinary differential equation $\dot{x}(t) = \mu(x(t), \vartheta)$, $t \geq 0$; $\nu(\cdot, \sigma)$ is a real function such that $\nu(0, \sigma) = 0$ and $\nu(x, \sigma) > 0$ for $x \neq 0$; $(W_t, t \geq 0)$ is *a linear standard Brownian motion* (see [22]); $X = (X_t, t \geq 0)$ is a *process of population growth* (briefly *growth process*); and $\vartheta = (\vartheta, \sigma)$ is a vector of the parameters of the model: ϑ is a vector of *drift parameters* and σ is a *diffusion parameter*. Usually, $\mu(\cdot, \vartheta)$ is called a *drift function* and $\nu(\cdot, \sigma)$ is called a *diffusion coefficient function* (see [22]).

---

*The lecture presented at the Mathematical Colloquium in Osijek organized by Croatian Mathematical Society - Division Osijek, January 9, 1998.

†Department of Mathematics, University of Zagreb, Bijenička cesta 30, HR-10 000 Zagreb, e-mail: huzak@cromath.math.hr

Let $\bar{\Theta}$ be *parameter space*, i.e. a set of all possible values of a parameter vector $\theta = (\vartheta, \sigma)$. Usually, it is supposed that $\bar{\Theta}$ is an open and connected set, a subset of Euclidean space $\mathbb{R}^{k+1}$ ($k \geq 1$).

The very first condition that any proposed diffusion growth model has to satisfy, is that SDE (1) has *a unique in law solution* with *almost sure continuous paths* (see [22]) for every $\theta \in \bar{\Theta}$. For example, this condition is satisfied if both drift and diffusion coefficient functions are locally Lipschitz ([23], Theorem V.12.1). In addition, if $\mu(\cdot, \vartheta)$ and $\nu(\cdot, \sigma)$ are of *bounded slope* (see again [23], Theorem V.12.1), then the growth process X is defined for all $t > 0$ (i.e. the solution of (1) does not explode in finite time). A diffusion model has to have *almost sure nonnegative paths* (see [22]) to be named as a *growth* model. Usually, this is a consequence of the forms of drift and diffusion coefficient functions.

Another property that any reasonable model has to have is *the stability on perturbations of parameters*, i.e. for any convergent sequence $(\theta_n; n \geq 1)$ of parameters in $\bar{\Theta}$, the sequence of growth processes $(X^{(n)}; n \geq 1)$ converges *in probability, uniformly in compacts* to the growth process $X^{(0)}$ (see [22]), where $X^{(n)}$ is a solution of the SDE (1) with parameter value $\theta = \theta_n$ for any $n \geq 0$ and $\theta_0 = \lim_n \theta_n$. For example, this condition is satisfied if both drift and diffusion coefficient functions are linear in parameters ϑ and σ ([21], Theorem V.15).

The problem we are going to discuss here is how to make parametric statistical inference from the observations of the growth process. Namely, we suppose that we observe the trajectories of the growth process satisfying SDE (1) of some specific diffusion growth model with an unknown vector of parameters θ which should be estimated.

An ideal way of observing the trajectories of a growth process is continuously through the time. Usually, this is not always possible, i.e. in many cases the trajectories can be observed only in discrete units of time. These imply two different estimation problems: parameter estimation problem from *continuous* observations and parameter estimation problem from *discrete* observations.

In this paper we will consider the estimation problem from continuous observations. The discussion of the estimation problem from discrete observations can be found, for example, in [3, 4, 10, 11, 16].

2. Maximum likelihood estimation

Let $\mathbb{P}_\theta$ be *the law* of the growth process X for the parameter vector $\theta \in \bar{\Theta}$ (see [22]). A solution of the estimation problem for some parametric diffusion model is possible if for any $\theta_1, \theta_2 \in \bar{\Theta}$, the assumption $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2}$ implies that $\theta_1 = \theta_2$. Let us assume that it holds.

Let $(X_t; 0 \leq t \leq T)$ be *a continuous observation up to the time* $T > 0$ of the growth process X. We assume that X is a solution of the SDE (1) with unknown true value $\theta_0 = (\vartheta_0, \sigma_0)$ of the model parameter vector.

The specific feature of the parametric estimation of diffusions based on continuous observations is that the value σ_0 of the diffusion parameter could be calculated through the following formula (see e.g. [5]) for *the quadratic variation process of* X

(see [22]):

$$\lim_n \sum_{j=1}^{2^n} (X_{jt2^{-n}} - X_{(j-1)t2^{-n}}) = \int_0^t \nu^2(X_s, \sigma_0)\, ds \ \text{ a.s.}, \quad 0 \le t \le T. \tag{2}$$

Hence, we will assume that the true value of the diffusion parameter σ is known and fixed. Let us denote $\nu(x, \sigma)$ briefly by $\nu(x)$. Let $\Theta = \Theta(\sigma)$ be a parameter space of unknown *drift* parameters ϑ. We assume that Θ is an open and connected set in Euclidean space $\mathbb{R}^k$ $(k \ge 1)$.

If drift and diffusion coefficient functions $\mu(\cdot, \vartheta)$ and ν are locally Lipschitz, then *the log-likelihood function* (LLF) of the diffusion growth model (up to some constant not depending on the parameter θ) based on the continuous observation $(X_t; 0 \le t \le T)$ is (see [14], Lemma 5.6. and definition 5.7.)

$$\ell_T(\vartheta) \equiv \ell(\vartheta | (X_t, 0 \le t \le T)) = \int_0^T \frac{\mu(X_t, \vartheta)}{\nu^2(X_t)}\, dX_t - \frac{1}{2} \int_0^T \frac{\mu^2(X_t, \vartheta)}{\nu^2(X_t)}\, dt. \tag{3}$$

Maximum likelihood estimator (MLE) of the unknown drift parameter ϑ is any value $\hat{\vartheta}_T \in \Theta$ such that

$$\ell_T(\hat{\vartheta}_T) = \max_{\vartheta \in \Theta} \ell_T(\vartheta). \tag{4}$$

The first problem is to find sufficient conditions for existence and measurability of MLE. If the drift function $\mu(x, \vartheta)$ is linear in ϑ, then LLF is a concave quadratic function. Hence, under some additional regularity conditions, for every $T > 0$, there exists a unique MLE $\hat{\vartheta}_T$ which has to be measurable (see [5]). If $\mu(x, \vartheta)$ is not linear in ϑ, then the problem of existence and measurability of MLE is a complex problem that usually has to be discussed for any particular model. For example, in the case of a one-dimensional nonlinear parameter ϑ, in [17] it has been shown that if $\mu(x, \vartheta)$ and $\nu(x)$ are sufficiently smooth, if X has *stationary distribution* (see e.g. [22]) and if some integrability conditions are satisfied, then there exists *a progressively measurable process* (see e.g. [22]) $(\hat{\theta}_T; T > 0)$ in $\Theta \cup \{+\infty\}$ such that

(i) there exists an a.s. finite random time τ such that for all $T \ge \tau$, $\hat{\vartheta}_T$ is in Θ and $\frac{d\ell_T}{d\vartheta}(\hat{\vartheta}_T) = 0$; and

(ii) a.s. $\lim_{T \to +\infty} \hat{\vartheta}_T = \vartheta_0$, where ϑ_0 is a true value of drift parameter in Θ.

The statement (i) tells us that a stationary point of LLF $\ell_T(\vartheta)$ exists for enough long observations of X. It may or may not be a MLE. The conditions on the model are too weak for stronger conclusions. But the same conditions are sufficient for the statement (ii) which tells us that the family of estimators $\hat{\vartheta}_T$, $T > 0$, is *strongly consistent* (see e.g. [19]). Moreover, under some additional conditions (see [17])

(iii) the process $(\sqrt{T}(\hat{\vartheta}_T - \vartheta_0); T > 0)$ has asymptotically normal distribution $\mathcal{N}(0, I_0^{-1})$, when $T \to +\infty$ and where

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T \left(\frac{\partial \mu}{\partial \vartheta}(X_t, \vartheta_0)\right)^2 \frac{dt}{\nu^2(X_t)} = I_0 \ \text{ a.s.}$$

The last statement tells us that the family of estimators $\hat{\vartheta}_T$, $T > 0$, is *asymptotically efficient* (see [15]). After all, our goal is to find estimators that are consistent and asymptotically efficient.

In [5] it has been shown that the MLE of a drift parameter of a diffusion model with a drift function $\mu(x, \vartheta)$ linear in ϑ has the same asymptotic properties. In addition, the asymptotic distribution of the test-statistic for testing the selection of (linear) model has been deduced.

It should be stressed that there exist some other methods of estimation which provide the estimators with the same or similar properties (see e.g. [17] or [20]).

3. Example

In [14] the diffusion growth model given by the SDE

$$dX_t = (\alpha - \beta h(\gamma, X_t))X_t\, dt + \sigma X_t\, dW_t, \quad X_0 = x_0 > 0, \tag{5}$$

has been introduced and discussed. The function h is defined by $h(\gamma, x) = (x^\gamma - 1)/\gamma$ if $\gamma \neq 0$ and $h(\gamma, x) = \log x$ if $\gamma = 0$ ($\gamma \in \mathbb{R}$, $x > 0$). The unknown drift parameter is $\vartheta = (\alpha, \beta, \gamma)$ and the diffusion parameter is σ. The parameter space is

$$\bar{\Theta} = \{(\alpha, \beta, \gamma, \sigma) \in \mathbb{R}^3 \times \langle 0, +\infty\rangle \;:\; \beta > 0,\ \gamma(\alpha - \frac{\sigma^2}{2}) + \beta > 0\}. \tag{6}$$

This model is a generalization of stochastic *Gompertz* model ($\gamma = 0$, see e.g. [13]), stochastic *logistic* model ($\gamma = 1$, see e.g. [1]) and stochastic *Bertalanffy* model ($\gamma = -1/3$, see [14]).

It turns out (see [14]) that the SDE (5) has a unique in law and a.s. continuous and positive solution. Moreover (see again [14]), the stationary distribution exists and the model is stable on perturbations of parameters.

The LLF for the drift parameter $\vartheta = (\alpha, \beta, \gamma)$, an element of drift parameter space Θ obtained from $\bar{\Theta}$ with fixed $\sigma > 0$, is (see [14])

$$\ell_T(\vartheta) = \frac{1}{\sigma^2}\int_0^T (\alpha - \beta h(\gamma, X_t))\frac{1}{X_t}\, dX_t - \frac{1}{2\sigma^2}\int_0^T (\alpha - \beta h(\gamma, X_t))^2\, dt. \tag{7}$$

$\ell_T : \Theta \to \mathbb{R}$ is of class C^3 on Θ and it is jointly measurable (see [14]). Let $D\ell_T(\vartheta)$ and $D^2\ell_T(\vartheta)$ denote the first and the second derivative of ℓ_T with respect to ϑ, respectively.

Let $\vartheta_0 = (\alpha_0, \beta_0, \gamma_0)$ be the true parameter value. The following two theorems are proved in [14].

Theorem 1. *There exists a progressively measurable process* $(\hat{\vartheta}_t; t \geq 0)$ *in* $\hat{\Theta}$ *(one-point compactification of* Θ*) such that*

(i) a.s. there exists $\tau > 0$ *such that for all* $T \geq \tau$, $\hat{\vartheta}_T \in \Theta$ *and* $D\ell_T(\hat{\vartheta}_T) = 0$*;*

(ii) $\lim_{T\to+\infty} \hat{\vartheta}_T = \vartheta_0$ *a.s.; and*

(iii) $\sqrt{T}(\hat{\vartheta}_T - \vartheta_0)$ *converges in law to the normal distribution* $\mathcal{N}(0, I_0^{-1})$ *when* $T \to +\infty$*, and where* $I_0 = \text{(a.s.)} \lim\limits_{T\to+\infty} \frac{1}{T}(-D^2\ell_T(\vartheta_0))$ *is a positively definite matrix.*

Let $c \in \mathbb{R}$ be given number and let

$$\Theta_c = \{ (\alpha, \beta) \in \mathbb{R}^2 \ : \ \beta > 0, \ c(\alpha - \frac{\sigma^2}{2}) + \beta > 0 \} \equiv \Theta \cap \{ \gamma = c \} \tag{8}$$

be the parameter space with the parameter γ being fixed and equal to c. Let $(\hat{\vartheta}_T^c; T \geq 0)$ be a sequence of MLEs of the parameter $\vartheta_0^c = (\alpha_0, \beta_0) \in \Theta_c$. These estimators exist (see [5]). The appropriate LLF is denoted by $\ell_T^c(\vartheta^c)$, $\vartheta^c \in \Theta_c$. Let $(\hat{\vartheta}_T; T \geq 0)$ be the sequence of the estimators of the parameter $\vartheta_0 = (\alpha_0, \beta_0, \gamma_0)$ from *Theorem 1*.

Theorem 2. *Under the assumption that $\gamma_0 = c$, $2(\ell_T(\hat{\vartheta}_T) - \ell_T^c(\hat{\vartheta}_T^c))$ converges in law to the χ^2-distribution with one degree of freedom, when $T \longrightarrow +\infty$.*

Theorem 2 provides the asymptotic distribution of the test-statistic $2(\ell_T(\hat{\vartheta}_T) - \ell_T^c(\hat{\vartheta}_T^c))$ that could be used for testing null-hypothesis H_0: $\gamma_0 = c$ as a part of the process of model selection.

References

[1] H. Aagaard-Hansen, G .F. Yeo, *A stochastic discrete generation birth, continuous death population growth model and its approximate solution*, J. Math. Biology **20**(1984), 69–90.

[2] Ž.Ḃajzer, S. Vuk-Pavlović, M. Huzak, *Mathematical modeling of tumor growth kinetics*, in: A Survey of Models for Tumor-Immune System Dynamics, (J. A. Adam and N. Bellomo, Eds.), Birkhäuser, Boston, 1997, 89–133.

[3] B. M. Bibby, M. Sørensen, *Martingale estimation functions for discretely observed diffusion processes*, Bernoulli, **1**(1/2)(1995), 17–39.

[4] B. M. Bibby, M. Sørensen, *On Estimation for Discretely Observed Diffusions: A Review*, Research Reports **334**, Department of Theoretical Statistics, University of Aarhus, 1995.

[5] B. M. Brown, J. I. Hewitt, *Asymptotic likelihood theory for diffusion processes*, J. Appl. Prob. **12**(1975), 228–238.

[6] K. C. Chan, G. A. Karolyi, F. A. Longstaff, A. B. Sanders, *An empirical comparison of alternative models of the short-term interest rates*, J. of Finance **47**(1992), 1209–1227.

[7] G. M. Constantinides, *A Theory of the nominal term structure of interest rates*, Review of Financial Studies **5**(1992), 531–552.

[8] G. Courtadon, *The Pricing of options on default-free bonds*, J. of Financial and Quantitative Analysis **17**(1982), 75–100.

[9] J. C. Cox, J. E. Ingersoll, S.A. Ross, *A Theory of the term structure of interest rates*, Econometrica **53**(1985), 385–407.

[10] D. Dacunha-Castelle, D. Florens-Zmirou, *Estimation of the coefficients of a diffusion from discrete observations*, Stochastics **19**(1986), 263–284.

[11] G. Dohnal, *On estimating the diffusion coefficient*, J. Appl. Prob. **24**(1987), 105–114.

[12] F. B. Hanson, C. Tier, *A stochastic model of tumor growth*, Math. Biosci. **61**(1982), 73–100.

[13] P. Holgate, *Varieties of stochastic model: a comparative study of the Gompertz effect*, J. Theor. Biol. **139**(1989), 369–378.

[14] M. Huzak, *Selection of diffusion growth process and parameter estimation from discrete observation*, Ph.D. thesis, University of Zagreb, 1997 (in Croatian).

[15] I. A. Ibragimov, R. Z. Has'minskii, *Statistical Estimation. Asymptotic Theory.*, Springer-Verlag, New York, 1981.

[16] M. Kessler, *Estimation of an ergodic diffusion from discrete observations*, Scandinavian Journal of Statistics **24**(1997), 211–229.

[17] V. Lanska, *Minimum contrast estimation in diffusion processes*, J. Appl. Prob. **16**(1979), 65–75.

[18] R. S. Liptser, A. N. Shiryayev, *Statistics of Random Processes I, General Theory*, Springer-Verlag, New York, 1977.

[19] R. S. Liptser, A. N. Shiryayev, *Statistics of Random Processes II, Applications*, Springer-Verlag, New York, 1977.

[20] A. A. Novikov, *Sequential estimation of the parameters of diffusion processes*, Math. Notes **12**(1972), 812-818.

[21] P. Protter, *Stochastic Integration and Differential Equations*, Springer-Verlag, Berlin, 1990.

[22] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, Berlin, 1991.

[23] L. C. G. Rogers, D. Williams, *Diffusion, Markov Processes, and Martingales*, Vol. 1-2, Wiley, Chichester, 1987.

[24] R. Wette, I. N. Katz, E. Y. Rodin, *Stochastic Processes for Solid Tumor Kinetics II. Diffusion Regulated Growth*, Math. Biosci. **21**(1974), 311–338.