Sankhyā : The Indian Journal of Statistics
1985, Volume 47, Series A, Pt. 3, pp. 392-398.

ON THE BERRY-ESSEEN BOUND FOR MAXIMUM LIKELIHOOD ESTIMATOR FOR LINEAR HOMOGENEOUS DIFFUSION PROCESSES

By M. N. MISHRA
Sambalpur University, India
and
B. L. S. PRAKASA RAO
Indian Statistical Institute

SUMMARY. The paper is concerned with the study of the rate of convergence of the distribution of a maximum likelihood estimator (MLE) of an unknown parameter in the drift coefficient of diffusion process described by a linear homogeneous stochastic differential equation. A bound on the rate of convergence of MLE to the true parameter is given.

1. Introduction

The study of the asymptotic distribution of an estimator is not very useful in general for practical purposes unless the rate of convergence is known. Michel and Pfanzagl (1971) and Pfanzagl (1971) studied the rate of convergence of the distribution of a minimum contrast estimator (MCE) to the normal distribution in the i.i.d. case. Prakasa Rao (1973) studied the rate of convergence of the distribution of the MLE to normal law for discrete time stationary Markov processes. No result of the Berry-Esseen type is known for the distribution of the maximum likelihood estimator of the drift parameter of a diffusion process described by a linear homogeneous stochastic differential equation even though asymptotic properties of the estimator are known (cf. Basawa and Prakasa Rao, 1980). Moreover, we do not know of any such results connected with central limit theorems for stochastic integrals.

In this paper, we shall study the rate of convergence of the distribution of MLE $\hat{\theta}_T$ of the parameter θ occurring in the drift coefficient of a linear homogeneous stochastic differential equation

$$dX_t = \theta\, a(X_t)dt + b(X_t)\, dW_t,\ t \geqslant 0;\ X_0 = x \in R \qquad \ldots \quad (1.1)$$

based on the realization $X_0^T = \{X_t\,;\, 0 \leqslant t \leqslant T\}$. We shall also obtain bounds on the difference $|\hat{\theta}_T - \theta|$. We assume that the equation (1.1) has a unique solution for any fixed θ.

AMS (1980) *subject classification :* 62M05.

Key words and phrases : Berry-Esseen bound ; Maximum likelihood estimator ; Stochastic differential equation ; Diffusion Process ; Rate of convergence.

2. Preliminaries

Let (Ω, F, P) be a probability space, $\{F_t, t \geq 0\}$ be an increasing family of sub σ-algebra of F such that F_0 contains all null sets of F and F_t is right continuous. Let $\{W_t, t \geq 0\}$ be the standard Wiener process adapted to $\{F_t, t \geq 0\}$. Denote by C_T, the space of continuous functions on $[0, T]$ with the supremum norm and B_T the associated Borel σ-algebra. Let $P_{\theta,T}$ be the measure induced by the process $\{X_t ; 0 \leq t \leq T\}$ satisfying (1.1) on (C_T, B_T) when θ is the parameter. Define by $P_{W,T}$ the measure induced by the Wiener process over $[0, T]$ on (C_T, B_T). Assume that $\theta \in \Theta \subset R$. Further suppose that

$$(A_1) \qquad 0 < E\left\{ \int_0^T \frac{a^2(X_t)}{b^2(X_t)}\, dt \right\} < \infty$$

for all $T > 0$. The log-likelihood function

$$L_T(\theta) = \log \frac{dP_{\theta,T}}{dP_{W,T}} \qquad \ldots \quad (2.1)$$

is a well-defined F_T-measurable function and it is given by

$$L_T(\theta) = \theta \int_0^T \frac{a(X_t)}{b^2(X_t)}\, dX_t - \frac{1}{2}\,\theta^2 \int_0^T \frac{a^2(X_t)}{b^2(X_t)}\, dt. \qquad \ldots \quad (2.2)$$

Hence the MLE $\hat{\theta}_T$ is given by

$$\hat{\theta}_T = \left\{ \int_0^T \frac{a(X_t)}{b^2(X_t)}\, dX_t \right\} \left\{ \int_0^T \frac{a^2(X_t)}{b^2(X_t)}\, dt \right\}^{-1}$$

$$= \theta + \left\{ \int_0^T \frac{a(X_t)}{b(X_t)}\, dW_t \right\} \left\{ \int_0^T \frac{a^2(X_t)}{b^2(X_t)}\, dt \right\}^{-1}. \qquad \ldots \quad (2.3)$$

Let

$$I_T = \int_0^T \frac{a^2(X_t)}{b^2(X_t)}\, dt \qquad \ldots \quad (2.4)$$

and suppose that there exists positive functions $Q(T) \uparrow \infty$, $\varepsilon(T) \downarrow 0$ such that

$$(A_2) \qquad Q(T)\varepsilon^2(T) \to \infty \text{ as } T \to \infty$$

and

$$\sup_{\theta \in \Theta} P_{\theta,T} \left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geq \varepsilon(T) \right\} = O(\varepsilon(T)^{1/2}). \qquad \ldots \quad (2.5)$$

Let

$$f(X_t) = a(X_t)[b(X_t)]^{-1} \text{ and } U_T = \int_0^T f(X_t)\, dW_t. \qquad \ldots \quad (2.6)$$

Then $\{U_t, F_t, t \geqslant 0\}$ is a square-integrable martingale with zero mean under (A_1). Hence, by Theorem 2.3 in Feigin (1976) due to Kunita-Watanabe, there exists a standard Wiener process $W(\cdot)$ adapted to $\{F_t, t \geqslant 0\}$ such that

$$\frac{U_T}{\sqrt{Q(T)}} = W\left(\frac{I_T}{Q(T)}\right) \; P\text{—a.s.} \qquad \dots \quad (2.7)$$

for all $T > 0$.

3. Main results

We shall use the following lemmas in the sequel.

Lemma 3.1 : *Let (Ω, F, P) be a probability space and f and g be F-measurable functions. Then, for any $\varepsilon > 0$,*

$$\sup_x \left| P\left\{w : \frac{f(w)}{g(w)} < x \right\} - \Phi(x) \right|$$
$$\leqslant \sup_y |P\{w : f(w) < y\} - \Phi(y)\}| + P\{w : |g(w) - 1| > \varepsilon\} + \varepsilon$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Proof : See Michel and Pfanzagl (1971).

Lemma 3.2 : *Let $\{W(t), t \geqslant 0\}$ be a standard Wiener process and Z be a non-negative random variable. Then, for every $x \in R$ and $\varepsilon > 0$,*

$$|P\{W(Z) \leqslant x\} - \Phi(x)| \leqslant (2\varepsilon)^{1/2} + P\{|Z - 1| > \varepsilon\}.$$

Proof : See Hall and Heyde (1980, p. 85).

We now state the main result of this paper.

Theorem 3.1 : *Under the assumption (A_1),*

$$\sup_{\theta \in \Theta} \sup_x |P_{\theta, T}\{\sqrt{Q(T)}(\hat{\theta}_T - \theta) \leqslant x\} - \Phi(x)|$$

$$< \sqrt{2\varepsilon(T)} + 2P_{\theta, T}\left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geqslant \varepsilon(T) \right\} + \varepsilon(T). \qquad \dots \quad (3.1)$$

If, further (A_2) holds, then the bound is of the older $O(\varepsilon(T)^{1/2})$.

Proof: Fix $\theta \in \Theta$. It is clear from (2.3) and the definition of U_T that $U_T = I_T(\hat{\theta}_T - \theta)$ and hence

$$|P_{\theta,T}\{\sqrt{Q(T)}(\hat{\theta}_T-\theta) \leq x\} - \Phi(x)|$$

$$= \left| P_{\theta,T}\left\{ \frac{U_T}{I_T} \sqrt{Q(T)} \leq x \right\} - \Phi(x) \right|$$

$$= \left| P_{\theta,T}\left\{ \frac{U_T/\sqrt{Q(T)}}{I_T/Q(T)} \leq x \right\} - \Phi(x) \right|$$

$$\leq \sup_y \left| P_{\theta,T}\left\{ \frac{U_T}{\sqrt{Q(T)}} \leq y \right\} - \Phi(y) \right| + P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geq \varepsilon(T) \right\} + \varepsilon(T)$$

(By Lemma 3.1)

$$= \sup_y \left| P\left\{ W\left(\frac{I_T}{Q(T)}\right) \leq y \right\} - \Phi(y) \right| + P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geq \varepsilon(T) \right\} + \varepsilon(T)$$

$$\leq (2\varepsilon(T))^{1/2} + 2P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geq \varepsilon(T) \right\} + \varepsilon(T)$$

(By Lemma 3.2).

This proves (3.1). Clearly the bound is of the order $O(\varepsilon(T)^{1/2})$ when (A_2) holds and the bound is also uniform in $\theta \in \Theta$. This completes the proof of the Theorem 3.1.

As a consequence of this theorem, we have the following result giving the rate of convergence of MLE $\hat{\theta}_T$.

Theorem 3.2 : *Suppose the condition (A_1) holds. Then for every $d > 0$ there exists a constant $c > 0$ such that*

$$\sup_{\theta \in \Theta} P_{\theta,T}\{|\hat{\theta}_T - \theta| \geq d\}$$

$$\leq c\, \varepsilon(T)^{1/2} + 2P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geq \varepsilon(T) \right\}. \quad \ldots \quad (3.2)$$

If, in addition (A_2) holds, i.e.,

$$\sup_{\theta \in \Theta} P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)} - 1 \right| \geq \varepsilon(T) \right\} = O(\varepsilon(T)^{1/2})$$

then the bound is $O(\varepsilon(T)^{1/2})$.

Proof: Let $\alpha_T(\theta, d) = P_{\theta,T}\{|\hat{\theta}_T - \theta| \geqslant d\}$. Then

$$\sup_{\theta \in \Theta} \alpha_T(\theta, d) \leqslant \sup_{\theta \in \Theta} |P_{\theta,T}\{|\hat{\theta}_T - \theta| \sqrt{Q(T)}$$

$$\geqslant d\sqrt{Q(T)}\} - 2\{1 - \Phi(\sqrt{Q(T)}d\} + 2\{1 - \Phi(\sqrt{Q(T)}d)\}$$

$$\leqslant (2\varepsilon(T))^{1/2} + 2 \sup_{\theta \in \Theta} P_{\theta,T}\left\{\left|\frac{I_T}{Q(T)} - 1\right| \geqslant \varepsilon(T)\right\} + \varepsilon(T)$$

$$+ \frac{2}{\sqrt{Q(T)}d} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Q(T)d^2}$$

by Theorem 3.1 and the inequality

$$1 - \Phi(x) < \frac{1}{x\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

for $x > 0$ (cf. Feller, 1968) p. 175). Since $Q(T)\varepsilon^2(T) \to \infty$ as $T \to \infty$, it follows that

$$\sup_{\theta \in \Theta} \alpha_T(\theta, d) \leqslant c\varepsilon(T)^{1/2} + 2 \sup_{\theta \in \Theta} P_{\theta,T}\left\{\left|\frac{I_T}{Q(T)} - 1\right| \geqslant \varepsilon(T)\right\}$$

for some $c > 0$ and the last term is $O(\varepsilon(T)^{1/2})$ under (A_2) completing the proof of Theorem 3.2.

4. Example

We now illustrate the above results by considering the stochastic differential equation

$$dX_t = -\theta X_t\, dt + dW_t,\ X_0 = 0,\ t \geqslant 0 \qquad \ldots \quad (4.1)$$

where $\theta \in (\alpha, \infty)$, $\alpha > 0$. Note that

$$X_t = e^{-\theta t} \int_0^t e^{\theta s}\, dW_s \qquad \ldots \quad (4.2)$$

and by Ito formula

$$X_t^2 = -2\theta \int_0^t X_s^2\, ds + t + 2\int_0^t X_s\, dW_s \qquad \ldots \quad (4.3)$$

and hence, if $I_T = \int_0^T X_s^2\, ds$, then

$$\frac{2\theta I_T}{T} - 1 = \frac{2}{T} \int_0^T X_t\, dW_t - \frac{1}{T} X_T^2.$$

It is easy to check that

$$E_\theta X_T^2 = \frac{1}{2\theta}(1-e^{-2\theta T}), \quad \ldots \quad (4.4)$$

and

$$E_\theta I_T = (T/2\theta)-(1/4\theta^2)(1-e^{-2\theta T}). \quad \ldots \quad (4.5)$$

Let $Q(T) = T/2\theta$ and $\varepsilon(T) \downarrow 0$ to be chosen later.

$$P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)}-1 \right| \geqslant \varepsilon(T) \right\}$$

$$= P_{\theta,T}\left\{ \left| \frac{2}{T}\int_0^T X_t\, dW_t - \frac{1}{T} X_T^2 \right| \geqslant \varepsilon(T) \right\}$$

$$\leqslant P_{\theta,T}\left\{ \frac{1}{T}\left| \int_0^T X_t\, dW_t \right| \geqslant \frac{\varepsilon(T)}{4} \right\} + P_{\theta,T}\left\{ \frac{1}{T} X_T^2 \geqslant \frac{\varepsilon(T)}{2} \right\}$$

$$\leqslant \frac{16 E_\theta I_T}{T^2\varepsilon^2(T)} + \frac{2E_\theta X_T^2}{T\varepsilon(T)} \leqslant \frac{16}{2\theta\varepsilon^2(T)} + \frac{2}{2\theta T\varepsilon(T)} \leqslant c(T\varepsilon^2(T))^{-1}$$

uniformly for $\theta \geqslant \alpha$ and T large. In particular, if we choose $\varepsilon(T) = T^{-2/5}$, it is easy to see that

$$\sup_\theta P_{\theta,T}\left\{ \left| \frac{I_T}{Q(T)}-1 \right| \geqslant \varepsilon(T) \right\} = O(\varepsilon(T)^{1/2})$$

and the conditions (A_1) and (A_2) hold. Hence

$$\sup_\theta \sup_x P_{\theta,T}\left| \left\{ \sqrt{\frac{T}{2\theta}}(\hat{\theta}_T-\theta) \leqslant x \right\} - \Phi(x) \right| = O(T^{-1/5})$$

and

$$\sup_\theta P_{\theta,T}\{|\hat{\theta}_T-\theta| \geqslant d\} = O(T^{-1/5})$$

from Theorems 3.1 and 3.2.

The rates obtained are not the best possibly due to the Skorokhod embedding technique applied by us. It is well known that this technique gives rates of order $O(n^{-1/4}\log n)$ in Berry-Esseen bounds for sums of martingale differences. It would be interesting to find whether rates of order $O(T^{-1/2})$ can be obtained as in Pfanzagl (1971) or Prakasa Rao (1973). The problem of obtaining the rates in the case of non-linear drift remains open. The embedding technique does not seem to be helpful in this case.

Acknowledgement. The authors thank Mr. Arup Bose for his detailed comments and suggestions which improved the results substantially from those in an earlier version of this paper.

REFERENCES

Basawa, I. W. and Prakasa Rao, B. L. S. (1980): *Statistical Inference for Stochastic Processes*, Academic Press, London.

Feigin, P. D. (1976): Maximum likelihood estimation for continuous time stochastic process. *Adv. Appl. Probability*, 8, 712-716.

Feller, W. (1968): *An introduction to Probability Theory and its Applications*, Wiley, New York.

Hall, P. and Heyde, C. C. (1980): *Martingale Limit theory and its Application*, Academic Press, New York.

Michel, R. and Pfanzagl, J. (1971): The accuracy of the normal approximation for minimum contrast estimate. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 18, 73-84.

Pfanzagl, J. (1971): The Berry-Esseen bound for minimum contrast estimates, *Metrika*, 17, 81-91.

Prakasa Rao, B. L. S. (1973): On the rate of convergence of estimators for Markov processes. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, 26, 141-152.

Paper received : January, 1984.

Revised : September, 1984.