isid/ms/2002/21
August 27, 2002
http://www.isid.ac.in/ ~ statmath/eprints

On Berry-Esseen type bound for least squares estimator for diffusion processes based on discrete observations

M. N. Mishra
and
B. L. S. Prakasa Rao

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

ON BERRY-ESSEEN TYPE BOUND FOR LEAST SQUARES ESTIMATOR FOR DIFFUSION PROCESSES BASED ON DISCRETE OBSERVATIONS

M. N. Mishra
Institute of Mathematics
and Applications
Bhubaneswar

and

B. L. S. Prakasa Rao
Indian Statistical Institute
New Delhi

Abstract: The paper is concerned with the distribution of the least squares estimator (LSE) of the drift parameter in the stochastic differential equation (SDE) of small diffusion observed over discrete set of time points. Convergence of the distribution of the least squares estimator to the standard normal distribution with an error bound has been obtained when the discretization step decreases with noise intensity.

Key words : Ito stochastic differential equation, Berry-Esseen bound, Least squares estimator.

1 Introduction

One of the basic assumptions in the study of estimation for parameters of a diffusion process is that the process can be observed continuously in time. This assumption is too strong and impossible to meet in practice. In view of this, it is desirable to study if the parameter can be reasonably estimated from the discrete data $\mathcal{D}_d = \{t_0, t_1, t_2, \cdots t_n; X(t_0), X(t_1), \cdots, X(t_n)\}$ This problem was first studied by Le Breton (1976) using Ito approximation and later by Mishra and Bishwal (1995) using Stratonovich approximation. They considered the linear stochastic differential equation (SDE)

$$\begin{aligned} dX(t) &= \theta X(t)dt + dW(t), \quad 0 < t \leq T, \\ X(0) &= 0 \end{aligned} \tag{1.1}$$

where $\{W(t), t \geq 0\}$ is the standard wiener process and studied the discrete approximation of the maximum likelihood estimator (MLE) $\hat{\theta}_T$ of θ when T is fixed, based on the process $X(t)$ observed over $[0, T]$, by an estimator based on X_{t_i}, $0 \leq i \leq n$ and $\Delta_n = \max\{|t_i - t_{i+1}|, 0 \leq i \leq n-1\} \to 0$ as $n \to \infty$. The problem was studied by Mishra and Prakasa Rao (2002) for non-linear stochastic differential equation (SDE) by using Stratonovich approximation.

Asymptotic optimality in the minimax sense of an approximate MLE and an approximate Bayes estimator of a parameter in the drift coefficient of a non-linear SDE, when the observations are made at regularly spaced and dense time points was studied by Mishra and Prakasa Rao (2001).

Considering the stochastic differential equation

$$\begin{aligned} dX(t) &= f(\theta, X(t))dt + dW(t), \quad 0 < t \leq T, \\ X(0) &= X_0, \end{aligned} \tag{1.2}$$

Dorogovchev (1976) studied the least squares estimator (LSE) $\hat{\theta}_n$ and proved that $\hat{\theta}_n \to \theta_0$ (the true value of the parameter) in probability as $\Delta_n \to 0$ and $T \to \infty$. Kasonga (1988) proved under mild regularity conditions that, the LSE $\hat{\theta}_n$ of θ derived from equation (1.2) based on $\{\mathcal{D}_d\}$ is strongly consistent. In addition to this, Prakasa Rao and Rubin (1981) studied the strong consistency and asymptotic normality of an estimator related to LSE for parameter in the nonlinear stochastic differential equation by studying families of stochastic integrals using Fourier analytic methods. Prakasa Rao (1983) studied the asymptotic distribution of the LSE of parameter in the drift coefficient in a nonlinear stochastic differential equation.

A comprehensive discussion on parameteric and non-parametric inference for stochastic process from sampled data is given in Prakasa Rao (1988) and more recently in Prakasa Rao (1999).

The study of the asymptotic distribution of an estimator is not very useful for practical purposes unless the rate convergence is known. No result of this type is known for the distribution of the least squares estimator of the drift parameter of a diffusion process described by a nonlinear homogenous stochastic differential equation even though asymptotic properties of the estimator are known (cf. Prakasa Rao (1999)).

In this paper we obtain the Berry-Esseen bound for the LSE of the drift parameter in Ito type stochastic differential equation of small diffusion, when the discretization step decreases with noise intensity. Section 2 of the paper deals with some preliminaries and main results.

2 Main Result

Let $X_\varepsilon(t)$ be the solution of the one dimensional nonlinear stochastic differential equation

$$\begin{aligned} dX_\varepsilon(t) &= f(\theta, X_\varepsilon(t))dt + \varepsilon dW(t), \quad 0 < t < T_\varepsilon \\ X_\varepsilon(0) &= X_0, \end{aligned} \tag{2.1}$$

where $\{W(t), t \geq 0\}$ is a standard wiener process, $f(\theta, x)$ is a known real valued function $\theta \in \Theta \subset \mathbb{R}$ and X_0 is independent of $\{W(t), t \geq 0\}$.

In this paper we are concerned with the discretization of the process $\{X_\varepsilon(t), t \geq 0\}$ when its values are observed at equidistant time points in $[0, T_\varepsilon]$. Actually we take $T_\varepsilon = \sqrt{\varepsilon n_\varepsilon}$ and choose design points as $t_k = (k-1)\frac{T_\varepsilon}{n_\varepsilon}$, $k = 1, 2, \ldots, n_\varepsilon + 1$. Then of course $\Delta t_k = t_{k+1} - t_k = \sqrt{\frac{\varepsilon}{n_\varepsilon}}$ for $k = 1, 2, \ldots, n_\varepsilon$, when $\varepsilon n_\varepsilon$ tends to infinity and $\varepsilon \to 0$. Let

$$Z_{n_\varepsilon}(\theta) = \sum_{k=1}^{n_\varepsilon} \left[X_\varepsilon(t_{k+1}) - X_\varepsilon(t_k) - f(\theta, X_\varepsilon(t_k))\sqrt{\frac{\varepsilon}{n_\varepsilon}} \right]^2. \tag{2.2}$$

We define the least square estimator (LSE) $\hat{\theta}_{n_\varepsilon}$ of θ to be the value of θ at which $Z_{n_\varepsilon}(\theta)$ attains its minimum.

Let us denote by θ_0 the true value of the parameter θ. The following notations are used in this paper.

(1) $f_k^{(i)} = \left.\frac{d^i f(\theta, X_\varepsilon(t_k))}{d\theta^i}\right|_{\theta=\theta_0}$, $1 \leq k \leq n_\varepsilon$ and $i = 1, 2$; $f_k = f(\theta_0, X_\varepsilon(t_k))$, $1 \leq k \leq n_\varepsilon$ and $f_k^{(2)}(\theta)$ denotes the second derivative of f evaluated at, some θ between θ_0 and $\hat{\theta}_{n_\varepsilon}$, and $X_\varepsilon(t_k)$. Finally we define,

(2) $W_k = W(t_{k+1}) - W(t_k)$, $V_k(\theta) = f_k^{(2)}(\theta) - f_k^{(2)}$ and $U_k = X_\varepsilon(t_{k+1}) - X_\varepsilon(t_k) - f_k\sqrt{\frac{\varepsilon}{n_\varepsilon}}$, $1 \leq k \leq n_\varepsilon$.

In the sequel, the notation E_θ(respectively$-P_\theta$) is used when an expectation (respectively-probability) is computed when the true parameter in (1) is θ.

We assume that the following set of regularity conditions hold.

(A_1) The process $\{X_\varepsilon(t), t \geq 0\}$ is a stationary process satisfying $E_\theta(X_\varepsilon^2(0)) < \infty$. More over for any $g(.)$, $E_\theta g(X_\varepsilon(0)) < \infty$, and $\frac{1}{n_\varepsilon}\sum_{k=0}^{n_\varepsilon} g(X_\varepsilon(t_k)) - E_\theta g(X_\varepsilon(0))$ tends to zero in probability as $n_\varepsilon \to \infty$ and $\varepsilon \to 0$.

(A$_2$) $f(\theta,x)$ is differentiable thrice with respect to θ and for every $\theta\in\Theta$, there exists a compact neighbourhood V_θ of θ such that,

$$\sup_{\theta'\in V_\theta}\left[E_\theta|f^{(j)}(\theta',X_\varepsilon(0))||f^{(i)}(\theta',X_\varepsilon(0))|\right]<\infty$$

for $0\le i,\ \ j\le 2$, where $f^{(0)}=f$ and $f^{(i)}$ denotes the i th derivative of f with respect to θ.

(A$_3$) $f^{(i)}(\theta,x)$ is Lipschitz in x for every $\theta\in\Theta$ for $i=0,1,2,$ and $|f^{(i)}(\theta,x)|\le C(\theta)(1+|x|),x\in I\!R$ for some $C(\theta)>0,i=0,1,2.$

(A$_4$) $E_{\theta_0}|f(\theta,X_\varepsilon(0))-f(\theta_0,X_\varepsilon(0))|^2=0$ iff $\theta=\theta_0$.

Suppose further that there exists a non-increasing function $\delta_\varepsilon\to 0,\ T_\varepsilon\,\delta_\varepsilon^2\to\infty$ as $n_\varepsilon\varepsilon\to\infty$ and $\varepsilon\to 0$.

(A$_5$) $P_{\theta_0}\left\{|\hat\theta_{n_\varepsilon}-\theta_0|\ge\delta_\varepsilon\right\}<C\delta_\varepsilon^{\frac14}$ where C is a constant.

(A$_6$) $E_\theta\left[f^{(i)}(\theta,X_\varepsilon(0))\right]^2<\infty$ for $\theta\in\Theta$ and $i=0,1,2.$

(A$_7$) $f^{(2)}(\theta,x)$ is Lipschitz in θ in a compact neighbourhood V_{θ_0} of θ_0 with Lipschitzian function $\varphi(x)$ satisfying $E_\theta\left[\varphi(X_\varepsilon(0))\right]^2<\infty$.

(A$_8$) $\lim\limits_{T_\varepsilon\to\infty}\frac{1}{T_\varepsilon}\int_0^{T_\varepsilon}\left[f^{(1)}(\theta_0,X_\varepsilon(t))\right]^2dt=A(\theta_0)$ P_{θ_0} -a.s. where $A(\theta_0)$ is a positive number.

(A$_9$) $\sup\limits_{\theta\in\Theta}P_\theta\left\{\left|\frac{I_{T_\varepsilon}(\theta)}{A(\theta)T_\varepsilon}-1\right|\ge\delta_\varepsilon\right\}\le C\delta_\varepsilon^{\frac12}$
where $I_{T_\varepsilon}(\theta)=\int_0^{T_\varepsilon}\left[f^{(1)}(\theta,X_\varepsilon(t))\right]^2dt$ and C is a constant.

Under the conditions (A$_1$) and (A$_2$), Prakasa Rao (1983) has shown that,

$$\begin{aligned}(\varepsilon n_\varepsilon)^{\frac14}(\hat\theta_{n_\varepsilon}-\theta_0) &= (\varepsilon n_\varepsilon)^{-\frac14}\sum_{k=1}^{n_\varepsilon}U_kf_k^{(1)}\left[\frac{1}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}f_k^{(1)^2}-\frac{1}{(\varepsilon n_\varepsilon)^{\frac12}}\sum_{k=1}^{n}U_kf_k^{(2)}-\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\right.\\ &\quad\left.\sum_{k=1}^{n_\varepsilon}U_kV_k(\theta^\star_{n_\varepsilon})+(\hat\theta_{n_\varepsilon}-\theta_0)^2O_P(1)\right]^{-1}. \qquad (2.3)\end{aligned}$$

Multiplying both sides of (2.3) by $\sqrt{A(\theta_0)}$ we obtain,

$$\sqrt{A(\theta_0)}(\varepsilon n_\varepsilon)^{\frac{1}{4}}(\hat{\theta}_{n_\varepsilon} - \theta_0) = \left\{ \frac{1}{\sqrt{A(\theta_0)}}(\varepsilon n_\varepsilon)^{-\frac{1}{4}} \sum_{k=1}^{n_\varepsilon} U_k f_k^{(1)} \right\}$$
$$\left[\left\{ \frac{\frac{\varepsilon^{\frac{1}{2}} n_\varepsilon^{\frac{1}{2}}}{n_\varepsilon} \sum_{k=1}^{n_\varepsilon} f_k^{(1)^2}}{A(\theta_0)(n_\varepsilon \varepsilon)^{\frac{1}{2}}} \right\} - \frac{1}{A(\theta_0)(n_\varepsilon \varepsilon)^{\frac{1}{2}}} \sum_{k=1}^{n_\varepsilon} U_k f_k^{(2)} \right.$$
$$\left. - \frac{1}{A(\theta_0)(n_\varepsilon \varepsilon)^{\frac{1}{2}}} \sum_{k=1}^{n_\varepsilon} U_k V_k(\theta^\star_{n_\varepsilon}) + (\hat{\theta}_{n_\varepsilon} - \theta_0)^2 O_P(1) \right]^{-1}. \tag{2.4}$$

For the proof of our main result we need following lemmas.

Lemma 2.1: Let $(\Omega, \mathcal{F}, P)$ be a probability space and f and g be $\mathcal{F}$-measurable functions. Then, for every $\varepsilon > 0$,

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

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

The proof of the lemma is given in Michael and Pfanzagl (1971).

Lemma 2.2: Let $\{W(t), t \geq 0\}$ be a standard Wiener process and Z be a non-negative random variable. Then, for every $\varepsilon > 0$ and for all x,

$$|P\{W(Z) \leq x\} - \Phi(x)| \leq (2\varepsilon)^{\frac{1}{2}} + P(|Z - 1| > \varepsilon)$$

For the proof of this lemma, we refer to Hall and Heyde (1980). From relation (2.4), using Lemma 2.1, we have

$$
\begin{aligned}
&\sup_y |P_{\theta_0}\left\{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}(\hat{\theta}_{n_\varepsilon}-\theta_0)\le y\right\}-\Phi(y)| \\
\le\ & \sup_y |P_{\theta_0}\left\{\frac{1}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\sum_{k=1}^{n_\varepsilon}U_k f_k^{(1)}\le y\right\}-\Phi(y)| \\
&+P_{\theta_0}\left[|\left\{\frac{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}(f_k^{(1)})^2/A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\right\}-\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\sum_{k=1}^{n_\varepsilon}U_k f_k^{(2)}\right. \\
&\left.-\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\sum_{k=1}^{n_\varepsilon}U_kV_k(\theta^\star_{n_\varepsilon})+(\hat{\theta}_{n_\varepsilon}-\theta_0)^2O_p(1)-1|>\delta_\varepsilon\right]+\delta_\varepsilon \\
\le\ & \sup_y |P_{\theta_0}\left\{\frac{1}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\sum_{k=1}^{n_\varepsilon}U_k f_k^{(1)}< y\right\}-\Phi(y)| \\
&+P_{\theta_0}\left[|\left\{\frac{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}\frac{(f_k^{(1)})^2}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\right\}-1|\ge\frac{\delta_\varepsilon}{4}\right] \\
&+P_{\theta_0}\left\{|\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\sum_{k=1}^{n_\varepsilon}U_k f_k^{(2)}|\ge\frac{\delta_\varepsilon}{4}\right\} \\
&+P_{\theta_0}\left\{|\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\sum_{k=1}^{n_\varepsilon}U_kV_k(\theta^\star_{n_\varepsilon})|\ge\frac{\delta_\varepsilon}{4}\right\} \\
&+P_{\theta_0}\left\{|(\hat{\theta_{n_\varepsilon}}-\theta_0)^2O_P(1)|\ge\frac{\delta_\varepsilon}{4}\right\}+\delta_\varepsilon \\
=\ & I_1+I_2+I_3+I_4+I_5+\delta_\varepsilon \quad \text{(say).}
\end{aligned}
$$

Now

$$
\begin{aligned}
|I_1| =\ & \sup_y |P_{\theta_0}\left\{\frac{1}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\sum_{k=1}^{n_\varepsilon}\left(U_k f_k^{(1)}-\varepsilon W_k f_k^{(1)}\right)\right. \\
&\left.+\frac{\varepsilon}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\int_0^{T_\varepsilon}H^{(1)}(t)dW(t)+\frac{\varepsilon}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\int_0^{T_\varepsilon}f^{(1)}(\theta_0,X(t))dW(t)\right\} \\
&\le y\}-\Phi(y)| \\
\le\ & \sup_y |P_{\theta_0}\left\{\frac{\varepsilon}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\int_0^{T_\varepsilon}f^{(1)}(\theta_0,X(t))dW(t)\le y\right\}-\Phi(y)| \\
&+P_{\theta_0}\left\{|\frac{1}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\sum_{k=1}^{n_\varepsilon}(U_k f_k^{(1)}-\varepsilon W_k f_k^{(1)})|\ge\frac{\varepsilon^{\frac{1}{4}}}{2}\right\} \\
&+P_{\theta_0}\left\{|\frac{\varepsilon}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\int_0^{T_\varepsilon}H^{(1)}(t)dW(t)|\ge\varepsilon^{\frac{1}{4}}/2\right\}+(\sqrt{2\pi})^{-1}\varepsilon^{\frac{1}{4}} \\
&\text{(by G.J.Babu et al. (1978))}
\end{aligned}
$$

where $H^{(i)}(t) = f^{(i)}(\theta_0, X(t_k)) - f^{(i)}(\theta_0, X(t)))$ for $i = 0,1,2,$ whenever $t_k \le t < t_{k+1}, 1 \le k \le n_\varepsilon$. Observe that

$$\begin{aligned}
\sum_{k=1}^{n_\varepsilon} E_{\theta_0}|U_k - \varepsilon W_k||f_k^{(1)}| &= \sum_{k=1}^{n_\varepsilon} E_{\theta_0}\left\{|\left[\int_{t_k}^{t_{k+1}} f(\theta_0, X_\varepsilon(t_k)) - f(\theta_0, X_\varepsilon(t))\right] dt||f_k^{(1)}|\right\} \\
&= \sum_{k=1}^{n_\varepsilon} E_{\theta_0}\left\{|\int_{t_k}^{t_{k+1}} H^{(0)}(t)dt||f_k^{(1)}|\right\} \\
&\le \sum_{k=1}^{n_\varepsilon}\left\{E_{\theta_0}\left(\int_{t_k}^{t_{k+1}} H^{(0)}(t)dt\right)^2 E_{\theta_0}(f_k^{(1)^2})\right\}^{\frac{1}{2}} \\
&\le \sum_{k=1}^{n_\varepsilon}\left(\int_{t_k}^{t_{k+1}} H^{(0)}(f_k^{(1)^2})\right\}^{\frac{1}{2}} \\
&\le \sum_{k=1}^{n_\varepsilon}\left\{\frac{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}{n_\varepsilon}\int_{t_k}^{t_{k+1}} E_{\theta_0}(H^{(0)}(t))^2 dt\, E_{\theta_0}(f_k^{(1)}(\theta_0, X_\varepsilon(0))^2\right\}^{\frac{1}{2}}
\end{aligned} \tag{2.5}$$

by stationarity of $X_\varepsilon(t)$. But

$$\int_{t_k}^{t_{k+1}} E_{\theta_0}(H^0(t))^2 dt \le C\int_{t_k}^{t_{k+1}} E_{\theta_0}(X_\varepsilon(t) - X_\varepsilon(t_k))^2 dt \tag{2.6}$$

by condition (A_3) for some constant $C > 0$. Again for $t_{k+1} \ge t \ge t_k$,

$$E_{\theta_0}(X(t) - X(t_k))^2 \le C(t - t_k) \le C\sqrt{\frac{\varepsilon}{n_\varepsilon}}$$

for some constant C independent of t (cf. Gihman and Skorokhod (1969)). Combining (2.5) and (2.6) we obtain that,

$$\sum_{k=1}^{n_\varepsilon} E_{\theta_0}|U_k - \varepsilon W_k||f_k^{(1)}| \le C\frac{(n_\varepsilon\varepsilon)^{\frac{3}{4}}}{n_\varepsilon^{\frac{1}{2}}} \tag{2.7}$$

for some constant $C > 0$, by using (A_6).

Therefore

$$\frac{1}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\sum_{k=1}^{n_\varepsilon} E_{\theta_0}|U_k - \varepsilon W_k||f_k^{(1)}| \le \frac{C_1}{\sqrt{A(\theta_0)}}\frac{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}{n_\varepsilon^{\frac{1}{2}}}.$$

Again, using this we have,

$$P_{\theta_0}\left\{\frac{1}{\sqrt{A(\theta_0)}(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\sum_{k=1}^{n_\varepsilon}|U_k - \varepsilon W_k||f_k^{(1)}| \ge \frac{\varepsilon^{\frac{1}{4}}}{2}\right\} \le \frac{C_1}{\sqrt{A(\theta_0)}}\frac{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}{n_\varepsilon^{\frac{1}{2}}\varepsilon^{\frac{1}{4}}} \le C_2\varepsilon^{\frac{1}{4}}. \tag{2.8}$$

Next,

$$
\begin{aligned}
E_{\theta_0}|\varepsilon(n_\varepsilon\varepsilon)^{-\frac{1}{4}}\left\{\int_0^{T_\varepsilon} H^{(1)}(t)dW(t)\right\}| &\le \varepsilon(n_\varepsilon\varepsilon)^{-\frac{1}{4}}\left\{E_{\theta_0}|\int_0^{T_\varepsilon} H^{(1)}(t)dW(t)|^2\right\}^{\frac{1}{2}} \\
&= \frac{\varepsilon}{(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\left\{\int_0^{T_\varepsilon} E_{\theta_0}(H^{(1)}(t))^2dt\right\}^{\frac{1}{2}} \le \frac{C_3\varepsilon}{(n_\varepsilon\varepsilon)^{\frac{1}{4}}}\left((n_\varepsilon\varepsilon)^{\frac{1}{2}}\frac{\varepsilon^{\frac{1}{2}}}{n_\varepsilon^{\frac{1}{2}}}\right)^{\frac{1}{2}} \\
&= \frac{C_3\varepsilon^{\frac{3}{2}}}{(n_\varepsilon\varepsilon)^{\frac{1}{4}}} = C_3\left(\frac{\varepsilon^3}{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\right)^{\frac{1}{2}},
\end{aligned}
\tag{2.9}
$$

by using (A_3) and arguments similar to those given above. Hence,

$$
P_{\theta_0}\left\{|\varepsilon(n_\varepsilon\varepsilon)^{-\frac{1}{4}}\int_0^{T_\varepsilon} H^{(1)}(t)dW(t)| \ge \frac{\varepsilon^{\frac{1}{4}}}{2}\right\} \le \frac{2C_3\varepsilon^{\frac{5}{4}}}{(n_\varepsilon\varepsilon)^{\frac{1}{4}}}. \tag{2.10}
$$

Denote

$$
Z_{T_\varepsilon} = \int_0^{T_\varepsilon} f^{(1)}(\theta_0, X_\varepsilon(t))dW(t)
$$

which is a square integrable martingale with zero mean. Hence, by Theorem 2.3 of Feigin (1976), there exists a standard Wiener process W(.) adapted to $\{\mathcal{F}_t, t \ge 0\}$ such that

$$
\frac{Z_{T_\varepsilon}}{\sqrt{A(\theta_0)T_\varepsilon}} = W\left(\frac{I_{T_\varepsilon}(\theta_0)}{A(\theta_0)T_\varepsilon}\right),\ P_{\theta_0} - \text{a.s. for all } T_\varepsilon > 0.
$$

Hence, by Lemma 2.2,

$$
\begin{aligned}
&|P_{\theta_0}\left\{\frac{\varepsilon}{\sqrt{A(\theta_0)T_\varepsilon}}\int_0^{T_\varepsilon} f^{(1)}(\theta_0, X(t))dW(t)) \le y\right\} - \Phi(y)| \\
&= \left|P_{\theta_0}\left\{W\left(\frac{I_{T_\varepsilon}(\theta_0)}{A(\theta_0)T_\varepsilon}\right) \le \frac{y}{\varepsilon}\right\} - \Phi\left(\frac{y}{\varepsilon}\right)\right| + \left|\Phi\left(\frac{y}{\varepsilon}\right) - \Phi(y)\right| \ \text{(by assumption}(A_9)) \\
&\le (2\delta_\varepsilon)^{\frac{1}{2}} + P_{\theta_0}\left\{|\frac{I_{T_\varepsilon}(\theta_0)}{A(\theta_0)T_\varepsilon} - 1| \ge \delta_\varepsilon\right\} + \left|\Phi\left(\frac{y}{\varepsilon}\right) - \Phi(y)\right| \ \text{(by Lemma 2.2)} \\
&\le C_4\left\{(2\delta_\varepsilon)^{\frac{1}{2}} + (T_\varepsilon\delta_\varepsilon^2)^{-1}\right\} \le 2C_4\left\{\delta_\varepsilon^{\frac{1}{2}} + (n_\varepsilon^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\delta_\varepsilon^2)^{-1}\right\}.
\end{aligned}
\tag{2.11}
$$

Now, combining (2.8), (2.10) and (2.11), we get,

$$
|I_1| \le C_5\left\{\delta_\varepsilon^{\frac{1}{2}} + \varepsilon^{\frac{1}{4}} + \left(n_\varepsilon^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\delta_\varepsilon^2\right)^{-1} + \varepsilon^{\frac{5}{4}}(n_\varepsilon\varepsilon)^{-\frac{1}{4}}\right\}. \tag{2.12}
$$

Let us consider,

$$
\begin{aligned}
|I_2| = P_{\theta_0}&\left[|\left\{\frac{(n_\varepsilon\varepsilon)^{\frac{1}{2}}}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}\frac{(f_k^{(1)})^2}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\right\}-1|\geq\frac{\delta_\varepsilon}{4}\right]\\
\leq\ & P_{\theta_0}\left[|\left\{\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}(f_k^{(1)})^2dt-\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}(f^{(1)^2}(\theta_0,X_\varepsilon(t))dt\right\}/A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}|\geq\frac{\delta_\varepsilon}{8}\right]\\
&+P_{\theta_0}\left\{|\left((\int_0^{T_\varepsilon}f^{(1)^2}(\theta_0,X_\varepsilon(t))dt/A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\right)-1|\geq\frac{\delta_\varepsilon}{8}\right\}\\
\leq\ & P_{\theta_0}\left[\left\{\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}|f_k^{(1)^2}-f^{(1)^2}(\theta_0,X_\varepsilon(t))|dt\right\}/A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\geq\frac{\delta_\varepsilon}{8}\right]\\
&+C_6\delta_\varepsilon^{\frac{1}{2}}\text{ (by assumption(A}_9\text{))}\\
\leq\ & \frac{C_7}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}E_{\theta_0}\left\{\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}\left|{f_k}^{(1)^2}-f^{(1)^2}(\theta_0,X_\varepsilon(t))\right|dt\right\}+C_7\delta_\varepsilon^{\frac{1}{2}}\\
\leq\ & \frac{C_7}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}\sum_{k=1}^{n_\varepsilon}E_{\theta_0}\left\{\int_{t_k}^{t_{k+1}}\left|f_k^{(1)}-f^{(1)}(\theta_0,X_\varepsilon^{(t)})\right|\ \left|f_k^{(1)}+f^{(1)}(\theta_0,X_\varepsilon^{(t)})\right|dt\right\}+C_7\delta_\varepsilon^{\frac{1}{2}}\\
\leq\ & \frac{C_8}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}\sum_{k=1}^{n_\varepsilon}\sqrt{\frac{\varepsilon}{n_\varepsilon}}E_{\theta_0}\left\{\int_{t_k}^{t_{k+1}}\left|f_k^{(1)}+f^{(1)}(\theta_0,X_\varepsilon(t))\right|dt\right\}\\
&+C_7\delta_\varepsilon^{\frac{1}{2}}\text{ (by assumption(}A_3\text{))}\\
\leq\ & \frac{C_8}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}\sqrt{\frac{\varepsilon}{n_\varepsilon}}\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}\left(E_{\theta_0}\left|f_k^{(1)}+f^{(1)}(\theta_0,X(t))\right|^2\right)^{\frac{1}{2}}dt+C_7\delta_\varepsilon^{\frac{1}{2}}\\
\leq\ & \frac{C_9}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}\sqrt{\frac{\varepsilon}{n_\varepsilon}}\left\{\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}\left\{E_{\theta_0}f_k^{(1)^2}(\theta_0,X_\varepsilon(0))+E_{\theta_0}f^{(1)^2}(\theta_0,X_\varepsilon(0))\right\}^{\frac{1}{2}}dt\right.\\
&+C_7\delta_\varepsilon^{\frac{1}{2}}\text{ (by stationarity)}\\
\leq\ & \frac{C_{10}}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}\frac{\varepsilon}{n_\varepsilon}n_\varepsilon+C_7\delta_\varepsilon^{\frac{1}{2}}\text{ (by assumption(}A_6\text{))}\\
\leq\ & C_{11}\left(\frac{\varepsilon}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon}+\delta_\varepsilon^{\frac{1}{2}}\right).
\end{aligned}
\tag{2.13}
$$

We know that

$$
|\sum_{k=1}^{n_\varepsilon}U_kf_k^{(2)}|\leq\sum_{k=1}^{n_\varepsilon}|U_k-\varepsilon W_k||f_k^{(2)}|+|\varepsilon\int_0^{T_\varepsilon}H^{(2)}(t)dW(t)|+|\varepsilon\int_0^{T_\varepsilon}f^{(2)}(\theta_0,X(t))dW(t)|
$$

and

$$
\begin{aligned}
E_{\theta_0} \sum_{k=1}^{n_\varepsilon} |U_k - \varepsilon W_k||f_k^{(2)}| &\le \sum_{k=1}^{n_\varepsilon} \left\{ E_{\theta_0} \left(\int_{t_k}^{t_{k+1}} H^{(\circ)}(t)dt \right)^2 E_{\theta_0} \left(f_k^{(2)} \right)^2 \right\}^{\frac{1}{2}} \\
&\le \sum_{k=1}^{n_\varepsilon} \left\{ \left(\frac{\varepsilon}{n_\varepsilon} \right) \int_{t_k}^{t_{k+1}} E_{\theta_0} \left(H^{(\circ)}(t) \right)^2 dt \, E_{\theta_0} (f_k^{(2)})^2 (\theta, X_\varepsilon(0))) \right\}^{\frac{1}{2}}. \quad \text{(by stationarity)} \\
&\le C_{12} \sum_{k=1}^{n_\varepsilon} \left\{ \left(\frac{\varepsilon}{n_\varepsilon} \right)^{\frac{1}{2}} \int_{t_k}^{t_{k+1}} E_{\theta_0} \left(X_\varepsilon(t_k) - X_\varepsilon(t) \right)^2 dt \right\} \quad \text{(by using(A}_3\text{) and (A}_6\text{)).} \\
&\le C_{13} n_\varepsilon \left(\frac{\varepsilon}{n_\varepsilon} \right)^{\frac{3}{4}} \\
&\le C_{14} n_\varepsilon^{\frac{1}{4}} \varepsilon^{\frac{3}{4}}. \qquad (2.14)
\end{aligned}
$$

Again,

$$
\begin{aligned}
E_{\theta_0} \left| \varepsilon \int_0^{T_\varepsilon} H^{(2)}(t) dW(t) \right| &\le \varepsilon \left\{ E_{\theta_0} \left| \int_0^{T_\varepsilon} H^{(2)}(t) dW(t) \right|^2 \right\}^{\frac{1}{2}} \\
&\le \varepsilon \left\{ \int_0^{T_\varepsilon} E_{\theta_0} \left(H^{(2)}(t) \right)^2 dt \right\}^{\frac{1}{2}} \le C_{15} \varepsilon \left((n_\varepsilon \varepsilon)^{\frac{1}{2}} (\frac{\varepsilon}{n_\varepsilon})^{\frac{1}{2}} \right)^{\frac{1}{2}} \\
&\le C_{16} \varepsilon^{\frac{3}{2}} \qquad (2.15)
\end{aligned}
$$

by using assumption (A_3) and similar arguments as given earlier. Further more

$$
\begin{aligned}
& E_{\theta_0} \left| \varepsilon \int_0^{T_\varepsilon} f^{(2)}(\theta_0, X_\varepsilon(t)) dW(t) \right| \\
&\le \varepsilon \left\{ \int_0^{T_\varepsilon} E_{\theta_0} (f^{(2)}(\theta_0, X_\varepsilon(0))^2 dt \right\}^{\frac{1}{2}} \text{ (by stationarity)} \\
&\le C_{18} \varepsilon (n_\varepsilon \varepsilon)^{\frac{1}{4}} \text{ (by assumption } (A_3)). \qquad (2.16)
\end{aligned}
$$

Therefore,

$$\begin{aligned}
|I_3| &= P_{\theta_0}\left\{\left|\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\sum_{k=1}^{n_\varepsilon} U_k f_k^{(2)}\right| \geq \frac{\delta_\varepsilon}{4}\right\} \\
&\leq P_{\theta_0}\left\{\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\sum_{k=1}^{n_\varepsilon} |U_k - \varepsilon W_k| \ \left|f_k^{(2)}\right| \geq \frac{\delta_\varepsilon}{12}\right\} \\
&\quad + P_{\theta_0}\left\{\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\left|\varepsilon\int_0^{T_\varepsilon} H^{(2)}(t)dW(t)\right| \geq \frac{\delta_\varepsilon}{12}\right\} \\
&\quad + P_{\theta_0}\left\{\frac{1}{A(\theta_0)(n_\varepsilon\varepsilon)^{\frac{1}{2}}}\left|\varepsilon\int_0^{T_\varepsilon} f^{(2)}(\theta_0, X(t))dt\right| \geq \frac{\delta_\varepsilon}{12}\right\} \\
&\leq C_{19}\frac{\varepsilon^{\frac{1}{2}}}{\sqrt{(n_\varepsilon\varepsilon)^{\frac{1}{2}}\delta_\varepsilon^2}}, \text{ (by using (2.14), (2.15) and (2.16))} \qquad (2.17)
\end{aligned}$$

Observe that,

$$|\sum_{k=1}^{n_\varepsilon} U_k V_k(\theta^\star_{n_\varepsilon})| \leq \sum_{k=1}^{n_\varepsilon} |(U_k - \varepsilon W_k)V_k(\theta^\star_{n_\varepsilon})| + |\sum_{k=1}^{n_\varepsilon} \varepsilon \int_{t_k}^{t_{k+1}} V_k(\theta^\star_{n_\varepsilon})dW(t)|.$$

Following the procedure in the calculation of the upper bound of $|I_1|$, we obtain,

$$\begin{aligned}
& E_{\theta_0}\sum_{k=1}^{n_\varepsilon} |U_k - \varepsilon W_k| V_k(\theta^\star_{n_\varepsilon}) \\
&\leq \sum_{k=1}^{n_\varepsilon} E_{\theta_0}\left\{|U_k - \varepsilon W_k| \sup_{\theta\in V_{\theta_0}} |V_k(\theta)|\right\} \\
&\quad \text{(since } \theta^\star_n \text{ is in the compact neighbourhood } V_{\theta_0} \text{of } \theta_0 \\
&\quad \text{ with probability close to 1 as } (n_\varepsilon\varepsilon)^{\frac{1}{2}} \to \infty, \text{ and } \varepsilon \to 0) \\
&\leq C_{20}\frac{(n_\varepsilon\varepsilon)^{\frac{3}{4}}}{n_\varepsilon^{\frac{1}{2}}} \quad \text{(since } E_{\theta_0}|\sup_{\theta\in V_{\theta_0}} V_k(\theta)|^2 < \infty \text{ by using (A}_7\text{)).}
\end{aligned}$$

Further more

$$\begin{aligned}
& E_{\theta_0}\left|\sum_{k=1}^{n_\varepsilon} \varepsilon \int_{t_k}^{t_{k+1}} V_k(\theta^\star_{n_\varepsilon})dW(t)\right| \\
&\leq \varepsilon E_{\theta_0}\sup_{\theta\in V_{\theta_0}}\left|\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}} V_k(\theta)dW(t)\right| \\
&= \varepsilon E_{\theta_0}\sup_{\theta\in V_{\theta_0}}\left|\sum_{k=1}^{n_\varepsilon}\varepsilon\int_{t_k}^{t_{k+1}} J(\theta,t)dW(t)\right|
\end{aligned}$$

where $J(\theta,t) = \begin{cases} V_k(\theta) & \text{if} \quad t_k \le t \le t_{k+1}, \\ 0 & \text{otherwise.} \end{cases}$ Now using assumptions (A_3) and (A_7), we get from Prakasa Rao and Rubin (1981, p.181) that there exists a $\gamma > \frac{1}{2}$ such that

$$\begin{aligned} & E_{\theta_0}\left|\sum_{k=1}^{n_\varepsilon} \varepsilon \int_{t_k}^{t_{k+1}} V_k(\theta^\star_{n_\varepsilon}) dW(t)\right| \\ \le \;& E_{\theta_0} \sup_{\theta \in V_{\theta_0}} \varepsilon \left|\int_0^{T_\varepsilon} J(\theta,t) dW(t)\right| \\ \le \;& C_{21} \varepsilon T_\varepsilon^{\frac{1}{2}} (\log T_\varepsilon)^\gamma \\ \le \;& C_{22} (n_\varepsilon \varepsilon)^{\frac{1}{4}} (\log n_\varepsilon \varepsilon)^\gamma \end{aligned}$$

Hence

$$|I_4| \le \frac{(\log (n_\varepsilon \varepsilon))^\gamma}{\sqrt{(n_\varepsilon \varepsilon)^{\frac{1}{2}} \delta_\varepsilon^2}}. \tag{2.18}$$

In view of assumption (A_5),

$$|I_5| \le C_{23} \delta_\varepsilon^{\frac{1}{4}}. \tag{2.19}$$

Therefore using (2.12), (2.13), (2.17), (2.18) and (2.19), we obtain

$$\begin{aligned} & \sup_y \left|P_{\theta_0}\left\{\sqrt{A(\theta_0)} (n_\varepsilon \varepsilon)^{\frac{1}{4}} (\hat{\theta}_{n_\varepsilon} - \theta_0) \le y\right\} - \Phi(y)\right| \\ \le \;& C_{24} \left(\frac{(\log(n_\varepsilon \varepsilon))^\gamma}{\sqrt{(n_\varepsilon \varepsilon)^{\frac{1}{2}} \delta_\varepsilon^2}} + \varepsilon^{\frac{1}{4}} + \delta_\varepsilon^{\frac{1}{2}} \right). \end{aligned}$$

We now have the main result.

Theorem : Under the conditions (A_1)-(A_9),

$$\sup_{-\infty < y < \infty} \left|P_{\theta_0}\left\{\sqrt{A(\theta_0)} (n_\varepsilon \varepsilon)^{\frac{1}{4}} (\hat{\theta}_{n_\varepsilon} - \theta_0) \le y\right\} - \Phi(y)\right| \le C \left(\varepsilon^{\frac{1}{4}} + \delta_\varepsilon^{\frac{1}{2}} + \frac{(\log(n_\varepsilon \varepsilon))^\gamma}{\sqrt{(n_\varepsilon \varepsilon)^{\frac{1}{2}} \delta_\varepsilon^2}} \right)$$

where C is a positive constant and $\gamma > \frac{1}{2}$.

Example : Considering the Ornstein-Uhlenbeck process

$$dX_\varepsilon(t) = -\theta X_\varepsilon(t) + \varepsilon dW(t), X_\varepsilon(0) = X_0$$

where $\theta > 0$ and X_0 is independent of the standard Wiener process $\{W(t), t \ge 0\}$. We obtain from Prakasa Rao (1983)

$$\sum_{k=1}^{n_\varepsilon} \left[\Delta_k - \left\{ -\theta_0 X_\varepsilon(t_k) - (\hat{\theta_{n_\varepsilon}} - \theta_0) X_\varepsilon(t_k) \right\} \frac{T_\varepsilon}{n_\varepsilon} \right] (-X_\varepsilon(t_k)) = 0$$

or

$$\sum_{k=1}^{n_\varepsilon}\left[\Delta_k+\theta_0 X_\varepsilon(t_k)\frac{T_\varepsilon}{n_\varepsilon}\right]X_\varepsilon(t_k)+(\hat\theta_{n_\varepsilon}-\theta_0)\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)\frac{T_\varepsilon}{n_\varepsilon}=0$$

or

$$\begin{aligned}\hat\theta_{n_\varepsilon}-\theta_0 &= -\frac{\sum_{k=1}^{n_\varepsilon}\left[\Delta_k+\theta_0X_\varepsilon(t_k)\frac{T_\varepsilon}{n_\varepsilon}\right]X_\varepsilon(t_k)}{\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)\frac{T_\varepsilon}{n_\varepsilon}}\\ &= -\frac{\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}\left[\Delta_k+\theta_0X_\varepsilon(t_k)(\frac{\varepsilon}{n_\varepsilon})^{\frac12}\right]X_\varepsilon(t_k)}{\frac{1}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)}\\ &= -\frac{\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}[U_kX_\varepsilon(t_k)]}{\frac{1}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)}.\end{aligned}$$

Then

$$\begin{aligned}P_{\theta_0}\left\{|\hat\theta_{n_\varepsilon}-\theta_0|\ge\delta_\varepsilon\right\} &= P_{\theta_0}\left\{|\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}U_kX_\varepsilon(t_k)/\frac{1}{n_\varepsilon}\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)|\ge\delta_\varepsilon\right\}\\ &\le P_{\theta_0}\left\{|\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}U_kX_\varepsilon(t_k)|\ge A(\theta_0)\delta_\varepsilon(1-\delta_\varepsilon)\right\}\\ &\quad +P_{\theta_0}\left\{|\frac{1}{n_\varepsilon A(\theta_0)}\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)-1|\ge\delta_\varepsilon\right\}.\end{aligned}$$

From (2.13) we get

$$P_{\theta_0}\left\{|\frac{1}{n_\varepsilon A(\theta_0)}\sum_{k=1}^{n_\varepsilon}X_\varepsilon^2(t_k)-1|\ge\delta_\varepsilon\right\}\le C_{25}\frac{\varepsilon}{\sqrt{n_\varepsilon\varepsilon\;\delta_\varepsilon^2}}+\delta_\varepsilon^{\frac12}. \tag{2.20}$$

Now,

$$\begin{aligned}&P_{\theta_0}\left\{|\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}U_kX_\varepsilon(t_k)|\ge A(\theta_0)\delta_\varepsilon(1-\delta_\varepsilon)\right\}\\ &= P_{\theta_0}\left\{|\frac{1}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}(U_k-\varepsilon W_k)X_\varepsilon(t_k)|\ge A(\theta_0)\frac{\delta_\varepsilon}{3}(1-\delta_\varepsilon)\right\}\\ &\quad +P_{\theta_0}\left\{|\frac{\varepsilon}{(n_\varepsilon\varepsilon)^{\frac12}}\sum_{k=1}^{n_\varepsilon}\int_{t_k}^{t_{k+1}}(X_\varepsilon(t_k)-X_\varepsilon(t))\,dW(t)|\ge A(\theta_0)\frac{\delta_\varepsilon}{3}(1-\delta_\varepsilon)\right\}\\ &\quad +P_{\theta_0}\left\{|\frac{\varepsilon}{(n_\varepsilon\varepsilon)^{\frac12}}\int_0^{T_\varepsilon}X_\varepsilon(t)dW(t)|\ge A(\theta_0)\frac{\delta_\varepsilon}{3}(1-\delta_\varepsilon)\right\}\\ &\le C_{26}\left\{\frac{(n_\varepsilon\varepsilon)^{\frac14}}{n_\varepsilon\varepsilon}\,\delta_\varepsilon+\frac{\varepsilon^{\frac32}}{(n_\varepsilon\varepsilon)^{\frac12}\,\delta_\varepsilon}\right\}\\ &\le C_{27}\frac{1}{(n_\varepsilon\varepsilon)^{\frac{1}{20}}}\quad\text{(by using (2.7) and (2.9)).}\end{aligned}\tag{2.21}$$

Let $\delta_\varepsilon = (n_\varepsilon \varepsilon)^{\frac{1}{5}}$. Then from (2.20) and (2.21) we get,

$$P_{\theta_0}\left\{|\hat{\theta}_{n_\varepsilon} - \theta_0| \geq \delta_\varepsilon\right\} \leq C_{28}\delta_\varepsilon^{\frac{1}{4}}.$$

Condition (A_9) is satisfied for the Ornstein-Uhlenbeck Process (cf. Mishra and Prakasa Rao (1985)). It can be verified that other conditions are satisfied for Ornstein-Uhlenbeck Process.

References :

1. Babu, G.J., Ghosh, M. and Singh, K. (1978), *On rates of convergence to normality for ϕ -mixing process*, Sankhyā 40(A),278-293.

2. Dorogovcev, A. Ja. (1976), *The consistency of an estimator of a parameter of stochastic differential equation,* Theory of Probability and Mathematical Statistics, **10**, 73-82.

3. Feigin, P. D. (1976), *Maximum likelihood estimation*, Adv. Appl. Prob. **8**, 712-736.

4. Gikhman I. and Skorokhod A. V.(1969) *Introduction to the Theory of Random Processes,* W.B.Saunders, Philadelphia.

5. Hall P. and Heyde C.C. (1980) *Martingale Limit theory and Appl.*, Academic Press, New York.

6. Kasonga, R. (1988), *The consistency of a nonlinear least square estimator for diffusion processes,* Stochastic Processes and their Applications, **36**, 263-275.

7. Le Breton, A. (1976), *On continuous and discrete sampling for parameter estimation in diffusion type processes,* Math Programming Studies, **5**, 124-144.

8. Michael, R. and Pfanzagl J.(1971), *The accuracy of the normal approximation for minimum contrast estimate*, Z. Warh. Theor. verw Gebiete, **18**, 37-84.

9. Mishra M.N. and Prakasa Rao B.L.S. (1985) *On the Berry-Essen bound for maximum likelihood estimator for linear homogeneous diffusion processes,* Sankhyā , **47(A)**, 392-398.

10. Mishra, M.N. and Bishwal, J.P.N. (1995), *Approximate maximum likelihood estimation for diffusion processes from discrete observations,* Stochastic and Stochastics Reports, **52**, 1-13.

11. Mishra, M.N. and Prakasa Rao, B.L.S. (2001), *Asymptotic minimax estimation in nonlinear stochastic differential equation from discrete observations,* Statistics and Decisions, **19**, 121-136.

12. Mishra, M.N. and Prakasa Rao, B.L.S. (2002), *Approximation of maximum likelihood estimator for diffusion processes from discrete observations,* Stochastic Analysis and Applications, **20** (to appear).

13. Prakasa Rao B.L.S. and Rubin H. (1981) *Asymptotic theory of estimation in nonlinear stochastic differential equations,* Sankhyā , **43(A)**, 170-189.

14. Prakasa Rao, B.L.S. (1983), *Asymptotic theory for nonlinear least squares estimator for diffusion processes,* Mathematische Operations forschung Statistik Serie Statistik,**14**, 195-209.

15. Prakasa Rao, B.L.S. (1988), *Statistical Inference from sampled data for stochastic processes,* Contemporary Mathematics, Vol.**80**, American Math Society, Providence, Rhode Island, 249-284.

16. Prakasa Rao, B.L.S. (1999), *Statistical Inference for Diffusion type processes,* Arnold, London.