International Journal of Pure and Applied Mathematics
Volume 68 No. 4 2011, 403-414

MILSTEIN APPROXIMATION OF POSTERIOR DENSITY FOR DIFFUSIONS

Jaya P.N. Bishwal
Department of Mathematics and Statistics
University of North Carolina at Charlotte
Charlotte, NC 28223-0001, USA
e-mail: J.Bishwal@uncc.edu

Abstract: We adopt the second order weak Milstein simulation scheme for estimating the true posterior density of diffusions and obtain the rate of weak convergence. We also obtain the Bernstein-von Mises Theorem concerning the convergence of the simulated posterior distribution to normal distribution and asymptotic normality of the Bayes estimator for smooth prior and loss functions.

AMS Subject Classification: 60F05, 60F10, 60H05, 62M05, 62F12
Key Words: diffusion processes, simulation, Milstein scheme, posterior distribution, Bernstein-von Mises Theorem

1. Introduction

Continuous time Markov models such as diffusion processes have been widely used in biology and finance.

For instance, Gompertz diffusion model describes the *in vivo* tumor growth. The drift parameter describes the intrinsic growth rate (mitosis rate) of the tumor. Consider the stochastic differential equation

$$dY_t = (\alpha Y_t - \beta Y_t \ln Y_t)\ dt + \sigma Y_t\ dW_t,\ t \geq 0,\quad Y_0 = y_0,$$

where $\{W_t\}$ is a standard Wiener process with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and $\alpha,\ \beta$ and σ are unknown parameter to be estimated on the basis of discrete observations of the process $\{Y_t\}$ at times $0 = t_0 < t_1 < \cdots t_n = T$.

Received: January 28, 2011

Here Y_t is the tumor volume which is measured at discrete time, α is the intrinsic growth rate of the tumor, β is the tumor growth acceleration factor, and σ is the diffusion coefficient.

The knowledge of distribution of the estimator may be applied to evaluate the distribution of other important growing parameters used to assess the tumor treatment modalities. Some of these parameters are the plateau of deterministic Gomperzian model, $Y_\infty = \exp(\alpha/\beta)$, tumor growth delay, and the first time the growth curve of the Gompertz diffusion reaches Y_∞.

In term structure of interest rates in finance, Cox-Ingersoll-Ross process is widely used as a short rate model and is given by the SDE

$$dY_t = \kappa(\alpha - Y_t)dt \ + \sigma\sqrt{Y}_t dW_t,$$

where $\{W_t\}$ is standard Brownian motion. The transition density is non-central chi-square. Here κ denotes the speed of mean reversion and $\kappa\alpha$ denotes the level of mean reversion of the interest rate and σ is the volatility.

Thus the data from these models are always discrete. Based on discrete observations, the main difficulty with diffusion processes inference is likelihood is intractable since the transition density is unknown except in some very special cases. Hence alternative strategies are necessary. Approximate Bayesian inference for stochastic differential equations based on discrete observations of the diffusion process has received recent attention, see Bishwal (2008). See also Bishwal (2000, 2001, 2006, 2007) and Bishwal and Bose (2001) about approximate likelihood and approximate posterior based inference for finite and infinite dimensional stochastic differential equations.

Consider the Itô diffusion

$$dY_t = a(\theta, Y_t)dt + b(\theta, Y_t)dW_t,$$

where $\{W_t\}$ is standard Brownian motion, a and b are smooth functions satisfying existence and uniqueness of a strong solution to the Itô stochastic differential equation, $\theta \in \mathbb{R}$ is the unknown parameter to be estimated based on discrete data $\{Y_1, Y_2, \cdots, Y_n\}$.

The true posterior distribution is given by

$$\pi(\theta|Y_1, Y_2, \cdots, Y_n) = \frac{\Pi_{i=1}^n P(Y_{i-1}, Y_i|\theta)\pi(\theta)}{\int \Pi_{i=1}^n P(Y_{i-1}, Y_i|\theta)\pi(\theta)d\theta},$$

where $y :\to P(x, y|\theta)$ is the intractable density of the chain $(Y_i)_{1\le i\le n}$.

In Section 2, we discuss the Milstein schemes. In Section 3, we obtain the rate of weak convergence of the approximation scheme to the posterior expectation. In Section 4, we obtain the Bernstein-von Mises Theorem concerning the convergence of the simulated posterior distribution to normal distribution and asymptotic normality of the Bayes estimator for smooth prior and loss functions.

2. Milstein Schemes

The first order Euler scheme has been extensively used for simulating the paths of diffusions. The scheme is given by

$$Y(x,m)_{i+1} \quad = \quad Y(x,m)_i + \tfrac{1}{m}a(\theta, Y(x,m)_i) + \tfrac{1}{\sqrt{m}}b(\theta, Y(x,m)_i)Z_i,$$

where $Z_i, i = 1, 2, \cdots, m$ are i.i.d. $\mathcal{N}(0,1)$ variables. The transition density of this scheme is Gaussian.

A natural idea is to approximate $P(x, \cdot|\theta)$ by the density $P^m(x, \cdot|\theta)$ of the terminal variable of a m-step Euler scheme on the interval $[0,1]$ and starting from x as in Pedersen (1995a). In order to improve the accuracy of approximation, we use Milstein scheme. First we consider the *second order mean square Milstein scheme.* More precisely, we consider the sequence

$$\begin{aligned} Y(x,m)_{i+1} = Y(x,m)_i + \frac{1}{m}a(\theta, Y(x,m)_i) + \frac{1}{\sqrt{m}}b(\theta, Y(x,m)_i)Z_i \\ + \frac{1}{2}b(\theta, Y(x,m)_i)b'(\theta, Y(x,m)_i)(Z_i^2 - 1), \end{aligned}$$

where $Z_i, i = 1, 2, \cdots, m$ are i.i.d. $N(0,1)$ variables. The transition density of this scheme is non-central chi square. The density $P^m(x, \cdot|\theta)$ of the terminal random variable $Y(x,m)_m$ approximates $P(x, \cdot|\theta)$. In particular, P^m tends to P as $m \to \infty$. Thus the approximate posterior density is

$$\pi^m(\theta|Y_1, Y_2, \cdots, Y_n) = \frac{\Pi_{i=1}^n P^m(Y_{i-1}, Y_i|\theta)\pi(\theta)}{\int \Pi_{i=1}^n P^m(Y_{i-1}, Y_i|\theta)\pi(\theta)d\theta},$$

in which m is chosen to be large enough, should be close to the exact posterior density. The main difficulty remains since m is larger than 1, there is no tractable expression for P^m. In this paper we study the rate of weak convergence of π^m to π as $m \to \infty$.

In order to improve the accuracy of weak approximation, we use the Second Order Weak Milstein Scheme, see Kloeden and Platen (1999). For $j = 0, 1, \cdots, m-1$,

$$Y_{t_i+(j+1)h} = Y_{t_i+jh} + a(\theta, Y_{t_i+jh})h + b(\theta, Y_{t_i+jh})(W_{t_{i+1}} - W_{t_i})$$

$$
\begin{aligned}
&+\frac{1}{2}b(\theta, Y_{t_i})b'(\theta, Y_{t_i+mh})\{(W_{t_i+(j+1)h} - W_{t_i+jh})^2 - h\} \\
&+\left(\frac{1}{2}a(\theta, Y_{t_i+jh})b'(\theta, Y_{t_i+jh}) + \frac{1}{2}a'(\theta, Y_{t_i+jh})b(\theta, Y_{t_i+jh})\right. \\
&\left.+\frac{1}{4}b^2(\theta, Y_{t_i+jh})b''(\theta, Y_{t_i+jh})\right)(W_{t_i+(j+1)h} - W_{t_i+jh})h \\
&+\left(\frac{1}{2}a(\theta, Y_{t_i+jh})a'(\theta, Y_{t_i+jh}) + \frac{1}{4}a''(\theta, Y_{t_i+jh})b^2(\theta, Y_{t_i+jh})\right)h^2.
\end{aligned}
$$

Here prime denotes derivative with respect to y.

3. Approximation of the Posterior Expectations

Cano, Kessler and Salmeron (2006) used the Euler scheme to approximate the posterior. We improve their approximation by using the second order weak Milstein scheme.

Let f be a bounded measurable function on Θ. In this section we obtain a bound on the error $|E_n f - E_n^m f|$, where E_n and E_n^m denote the expectation w.r.t. the exact posterior density $\pi(\theta|Y_1, Y_2, \cdots, Y_n)$ and approximate posterior density $\pi^m(\theta|Y_1, Y_2, \cdots, Y_n)$ respectively. The main tool is the control of the difference $P(x, \cdot|\theta)$ and $P^m(x, \cdot|\theta)$.

Assumptions. (A1) $P(x, y|\theta) + P^m(x, y|\theta) \le C(\theta)e^{-C'(\theta)|x-y|^2}$.
(A2) $|P(x, y|\theta) - P^m(x, y|\theta)| \le \frac{C(\theta)}{m^2}e^{-C'(\theta)|x-y|^2}$.
The following is the main result of this section.

Theorem 3.1. *Let $\epsilon = \int Q_n(\theta)\pi(\theta)d\theta$ denote the value of the predictive distribution after $Y_1, Y_2, \cdots, Y_n$ have been observed. Under assumptions (A1) and (A2), if $|Y_{j-1} - Y_j| > 2/m$ for every $j = 1, 2, \cdots, n$, then*

$$|E_n f - E_n^m f| \le 2||f|| \frac{\int C(\theta)^n e^{-nC'(\theta)\sum_{i=1}^n |Y_{i-1}-Y_i|^2}\pi(\theta)d\theta}{(m^2/n)\epsilon - \int C(\theta)^n e^{-nC'(\theta)\sum_{i=1}^n |Y_{i-1}-Y_i|^2}\pi(\theta)d\theta}$$

provided $\int C(\theta)^n e^{-nC'(\theta)\sum_{i=1}^n |Y_{i-1}-Y_i|^2}\pi(\theta)d\theta$ is finite. Here $||f||$ denotes the supnorm of f.

Proof. Let $Q_n(\theta) := \Pi_{i=1}^n P(Y_{i-1}, Y_i|\theta)$, $Q_n^m(\theta) := \Pi_{i=1}^n P^m(Y_{i-1}, Y_i|\theta)$. Define $h_n^m(\theta) := Q_n^m(\theta) - Q_n(\theta)$. We obtain

$$|E_n f - E_n^m f| \quad = \quad \frac{\int f(\theta)Q_n(\theta)\pi(\theta)d\theta}{\int Q_n(\theta)\pi(\theta)d\theta} - \frac{\int f(\theta)Q_n^m(\theta)\pi(\theta)d\theta}{\int Q_n^m(\theta)\pi(\theta)d\theta}$$

$$= \frac{\int f(\theta)Q_n(\theta)\pi(\theta)d\theta}{\int Q_n(\theta)\pi(\theta)d\theta} \frac{\int h_n^m(\theta)\pi(\theta)d\theta}{\int Q_n^m(\theta)\pi(\theta)d\theta} - \frac{\int f(\theta)h_n^m(\theta)\pi(\theta)d\theta}{\int Q_n^m(\theta)\pi(\theta)d\theta}.$$

Therefore,

$$|E_n f - E_n^m f| \leq 2||f|| \left| \frac{h_n^m(\theta)\pi(\theta)d\theta}{\epsilon + h_n^m(\theta)\pi(\theta)d\theta} \right|. \tag{3.1}$$

Using the fact that we can write the difference between the products of n terms as

$$\Pi_{i=1}^n a_i^n - \Pi_{i=1}^m a_i = (a_1^m - a_1)a_2 \dots a_n + \dots + a_1^m \dots n_{n-1}^m (a_n^m - a_n),$$

we obtain using assumption (A1) and (A2),

$$|h_n^m(\theta)| \leq nC(\theta)^n/m^2)e^{-nC'(\theta)\sum_{i=1}^n |Y_{i-1}-Y_i|^2}.$$

We also deduce that for m large enough,

$$-\epsilon < -\int n(c(\theta)^n/m^2)e^{-nC'(\theta)\sum_{i=1}^n |Y_{i-1}-Y_i|^2}$$

which together with (3.1) proves the theorem. $\square$.

Thus the approximate posterior converges weakly to the exact posterior as $m \to \infty$ for fixed n.

4. Bernstein-von Mises Theorem

In this section we study the convergence of the approximate posterior distribution to normal distribution.

Suppose that Π is a prior probability measure on $(\Theta, \mathcal{D})$, where $\mathcal{D}$ is the σ-algebra of Borel subsets of the Θ. Assume that Π has a density $\pi(\cdot)$ with respect to the Lebesgue measure and the density is continuous and positive in an open neighborhood of true value the unknown parameter θ_0.

Denote $Y^n := \{Y_1, Y_2, \cdots, Y_n\}$. The maximum likelihood estimator (MLE) based on Y^n is defined as

$$\theta^n := \arg\max L_n(\theta),$$

where

$$L_n(\theta) = \Pi_{i=1}^n P(Y_{i-1}, Y_i|\theta).$$

Between time points t_i to t_{i+1}, we simulate m data points by second order weak Milstein scheme. We refer to the collection of simulated data and the observations as the augmented data.

The transition function is given by $p(s,x,t,A;\theta) = P_\theta(X_t \in A|X_s = x)$. The approximate transition density uses the second order weak Milstein scheme. Define

$$p_m(s,x,t,y;\theta) = E_{P_{\theta,x,s}}(p_1(\tau_{m-1}, X_{\tau_{m-1},m}, t, y;\theta)).$$

We omit the details of the proof of the following proposition which is similar to the proof in Pedersen (1995a, b).

Proposition 4.1. *(i)* $p_m(s,x,t,y;\theta) \to p(s,x,t,y;\theta)$ *in* L^1 *as* $m \to \infty$ *for all* $0 \le s < t,\ x \in \mathbb{R}$ *and for all* $\theta \in \Theta$,

(ii) $L_{n,m}(\theta) \to L_n(\theta)$ *in probability under* P_{θ_0} *as* $m \to \infty$ *for all* $\theta \in \Theta$ *and* $n \in \mathbb{N}$ *where* θ_0 *denotes the true parameter value.*

The simulated maximum likelihood estimator (SMLE) based on the augmented data $Y^{n,m}$ is defined as

$$\hat{\theta}^{n,m} := \arg\max L_{n,m}(\theta).$$

We assume Θ to be compact. The consistency and the asymptotic normality of the SMLE as $m \to \infty$ and $n \to \infty$ can be proved by the methods of Pedersen (1995b). The asymptotic setting is that there exists a sequence of number of augmentations $m(n) \to \infty$ such that the number of observations $n \to \infty$. The simulated posterior density of θ given in $Y^{n,m}$ is given by

$$p(\theta|Y^{n,m}) := \frac{L_{n,m}(\theta)\pi(\theta)}{\int_\Theta L_{n,m}(\theta)\pi(\theta)d\theta}.$$

Let $\tau = (\frac{n}{m^2})^{1/2}(\theta - \hat{\theta}^n)$. Then the posterior density of $(\frac{n}{m^2})^{1/2}(\theta - \hat{\theta}^{n,m})$ is given by

$$p^*(\tau|Y^{n,m}) := (\frac{n}{m^2})^{-1/2} p(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau|Y^{n,m}).$$

The simulated posterior density of $(\frac{n}{m^2})^{1/2}(\theta - \hat{\theta}^{n,m})$ is given by

$$p^*(\tau|Y^{n,m}) := (\frac{n}{m^2})^{-1/2} p(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau|Y^{n,m}).$$

Let $\mathbb{P}_\theta^n$ be the measure generated by the simulated process, θ_0 be the true value of the parameter and

$$\nu_{n,m}(\tau) := \frac{d\mathbb{P}^n_{\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau}/d\mathbb{P}^n_{\theta_0}}{d\mathbb{P}^n_{\hat{\theta}^{n,m}}/d\mathbb{P}^n_{\theta_0}} = \frac{d\mathbb{P}^n_{\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau}}{d\mathbb{P}^n_{\hat{\theta}^{n,m}}},$$

$$C_{n,m} := \int_{-\infty}^{\infty} \nu_n(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau)d\tau.$$

Clearly,

$$p^*(\tau|Y^{n,m}) = C_{n,m}^{-1}\nu_{n,m}(\tau)\pi(\hat{\theta}^n + (\frac{n}{m^2})^{-1/2}\tau).$$

We study the asymptotic behavior of the posterior when the degree of augmentation m and the number of observations n increases.

Let the Fisher information be defined as $I(\theta_0) = E_{\theta_0}\left[\frac{\partial}{\partial\theta}\left(\frac{a(\theta,Y_0)}{b(\theta,Y_0)}\right)^2\right]$.

Let $K(\cdot)$ be a non-negative measurable function satisfying the following two conditions:

(K1) There exists a number $\eta, 0 < \eta < I(\theta_0)$, for which

$$\int_{-\infty}^{\infty} K(\tau)\exp\{-\frac{1}{2}\tau^2(I(\theta_0) - \eta)\}d\tau < \infty.$$

(K2) For every $\epsilon > 0$ and $\delta > 0$

$$e^{-\epsilon(\frac{n}{m^2})}\int_{|\tau|>\delta} K(\tau(\frac{n}{m^2})^{1/2})\pi(\hat{\theta}^{n,m} + \tau)d\tau \to 0 \text{ a.s. } [P_{\theta_0}]$$

as $n \to \infty,\ m \to \infty$ such that $m^2/n \to 0$.

We need the following Lemma to prove the Bernstein-von Mises Theorem.

Lemma 4.1. *Under the assumptions (K1)-(K2),*
(i) There exists a $\delta_0 > 0$ *such that*

$$\lim_{n\to\infty,\ m\to\infty}\int_{|\tau|\le\delta_0(\frac{n}{m^2})^{1/2}} K(\tau)\left|\nu_{n,m}(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau)\right.$$
$$\left. -\pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau = 0 \text{ a.s. } [P_{\theta_0}].$$

(ii) For every $\delta > 0$,

$$\lim_{n\to\infty,\ m\to\infty}\int_{|\tau|\ge\delta(\frac{n}{m^2})^{1/2}} K(\tau)\left|\nu_{n,m}(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau)\right.$$
$$\left. -\pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau = 0 \text{ a.s. } [P_{\theta_0}].$$

Proof. (i) follows by an application of dominated convergence theorem.

For every $\delta > 0$, there exists $\epsilon > 0$ depending on δ and β such that

$$\begin{aligned}
&\int_{|\tau|\geq\delta(\frac{n}{m^2})^{1/2}} K(\tau) \left|\nu_{n,m}(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau) \; -\pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau \\
\leq &\int_{|\tau|\geq\delta(\frac{n}{m^2})^{1/2}} K(\tau)\nu_{n,m}(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau)d\tau \\
&+ \int_{|\tau|\geq\delta(\frac{n}{m^2})^{1/2}} \pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2)d\tau \\
\leq &e^{-\epsilon(\frac{n}{m^2})} \int_{|\tau|\geq\delta(\frac{n}{m^2})^{1/2}} K(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau)d\tau \\
&+ \pi(\theta_0) \int_{|\tau|\geq\delta(\frac{n}{m^2})^{1/2}} \exp(-\frac{I(\theta_0)}{2}\tau^2)d\tau \\
=: &F_{n,m} + G_{n,m}.
\end{aligned}$$

By condition (K2), it follows that $F_{n,m} \to 0$ a.s. $[P_{\theta_0}]$ as $n \to \infty,\ m \to \infty$ for every $\delta > 0$. Condition K(1) implies that $G_{n,m} \to 0$ as $n \to \infty$. This completes the proof of the lemma. □

Now we are ready to prove the generalized version of the Bernstein-von Mises Theorem for discretely sampled simulated diffusions.

Theorem 4.1. *Under the assumptions (K1)-(K2), we have*

$$\lim_{n\to\infty,\ m\to\infty} \int_{-\infty}^{\infty} K(\tau) \left|p^*(\tau|Y^{n,m}) - (\frac{I(\theta_0)}{2\pi})^{1/2}\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau = 0$$

a.s. $[P_{\theta_0}]$.

Proof. From Lemma 4.1, we have

$$\begin{aligned}
\lim_{n\to\infty, m\to\infty} \int_{-\infty}^{\infty} K(\tau) &\left|\nu_n(\tau)\pi(\hat{\theta}^n + (\frac{n}{m^2})^{-1/2}\tau)\right. \\
&\left. -\pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau = 0 \quad \text{a.s. } [P_{\theta_0}].
\end{aligned} \tag{4.1}$$

Putting $K(\tau) = 1$ which trivially satisfies (K1) and (K2), we have

$$\begin{aligned}
C_{n,m} = &\int_{-\infty}^{\infty} \nu_{n,m}(\tau)\pi(\hat{\theta}^n + (\frac{n}{m^2})^{-1/2}\tau)d\tau \\
\to &\pi(\theta_0) \int_{-\infty}^{\infty} \exp(-\frac{I(\theta_0)}{2}\tau^2)d\tau \quad \text{a.s. } [P_{\theta_0}].
\end{aligned} \tag{4.2}$$

Therefore, by (4.1) and (4.2), we have

$$
\begin{aligned}
&\int_{-\infty}^{\infty} K(\tau)\left|p^*(\tau|Y^{n,m}) - (\frac{I(\theta_0)}{2\pi})^{1/2}\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau \\
&\le \int_{-\infty}^{\infty} K(\tau)\Big|C_{n,m}^{-1}\nu_n(\tau)\pi(\hat{\theta}^{n,m} + (\frac{n}{m^2})^{-1/2}\tau) \\
&\quad -C_{n,m}^{-1}\pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2)\Big| d\tau \\
&\quad + \int_{-\infty}^{\infty} K(\tau)\left|C_{n,m}^{-1}\pi(\theta_0)\exp(-\frac{I(\theta_0)}{2}\tau^2) - (\frac{I(\theta_0)}{2\pi})^{1/2}\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau \\
&\longrightarrow 0 \ \text{ a.s. } [P_{\theta_0}] \text{ as } n\to\infty \ \text{ and } m\to\infty. \qquad \square
\end{aligned}
$$

Theorem 4.2. *Suppose* $\int_{-\infty}^{\infty}|\theta|^r\pi(\theta)d\theta < \infty$ *for some non-negative integer* r *holds. Then as* $m^2/n \to 0$

$$
\lim_{n\to\infty,\ m\to\infty,}\int_{-\infty}^{\infty}|\tau|^r\left|p^*(\tau|Y^{n,m}) - (\frac{I(\theta_0)}{2\pi})^{1/2}\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau = 0 \ \text{ a.s. } [P_{\theta_0}].
$$

Proof. For $r = 0$, the verification of (K1) and (K2) is easy and the theorem follows from Theorem 4.1. Suppose $r \ge 1$. Let $K(\tau) = |\tau|^r, \delta > 0$ and $\epsilon > 0$. Using $|a+b|^r \le 2^{r-1}(|a|^r + |b|^r)$, we have

$$
\begin{aligned}
&e^{-\epsilon(\frac{n}{m^2})}\int_{|\tau|>\delta} K(\tau(\frac{n}{m^2})^{1/2})\pi(\hat{\theta}^{n,m}+\tau)d\tau \\
&\qquad \le (\frac{n}{m^2})^{r/2}e^{-\epsilon(\frac{n}{m^2})}\int_{|\tau-\hat{\theta}^{n,m}|>\delta}\pi(\tau)|\tau-\hat{\theta}^{n,m}|^r d\tau \\
&\qquad \le 2^{r-1}(\frac{n}{m^2})^{r/2}e^{-\epsilon(\frac{n}{m^2})}\Big[\int_{|\tau-\hat{\theta}^{n,m}|>\delta}\pi(\tau)|\tau|^r d\tau \\
&\qquad\quad + \int_{|\tau-\hat{\theta}^{n,m}|>\delta}\pi(\tau)|\hat{\theta}^{n,m}|^r d\tau\Big] \\
&\qquad \le 2^{r-1}(\frac{n}{m^2})^{r/2}e^{-\epsilon(\frac{n}{m^2})}\Big[\int_{-\infty}^{\infty}\pi(\tau)|\tau|^r d\tau + |\hat{\theta}^{n,m}|^r\Big] \\
&\qquad \longrightarrow 0 \text{ a.s. } [P_{\theta_0}] \text{ as } n\to\infty \text{ and } m\to\infty.
\end{aligned}
$$

from the consistency of $\hat{\theta}^{n,m}$ and hypothesis of the theorem. Thus the theorem follows from Theorem 4.1. $\square$

Remark 4.1. For $r = 0$ in Theorem 4.2, we have

$$
\lim_{n\to\infty,m\to\infty}\int_{-\infty}^{\infty}\left|p^*(\tau|Y^{n,m}) - (\frac{I(\theta_0)}{2\pi})^{1/2}\exp(-\frac{I(\theta_0)}{2}\tau^2)\right| d\tau = 0
$$

a.s. $[P_{\theta_0}]$.

This is the classical form of Bernstein-von Mises Theorem for discretely sampled simulated diffusions in its simplest form.

As a special case of Theorem 4.2, we obtain

$$E_{\theta_0}[(\frac{n}{m^2})^{1/2}(\hat{\theta}^{n,m} - \theta_0)]^r \to E[\xi^r]$$

as $n \to \infty$, $m \to \infty$ and $m^2/n \to 0$ where $\xi \sim \mathcal{N}(0, I^{-1}(\theta_0))$.

As an application of Theorem 4.1, we obtain the asymptotic properties of a regular Bayes estimator of θ. Suppose $l(\theta, \phi)$ is a loss function defined on $\Theta \times \Theta$. Assume that $l(\theta, \phi) = l(|\theta - \phi|) \geq 0$ and $l(\cdot)$ is non decreasing. Suppose that J is a non-negative function on $\mathbb{N} \times \mathbb{M}$ and $K(\cdot)$ and $G(\cdot)$ are functions on $\mathbb{R}$ such that:

(B1) $J(n,m)l(\tau(\frac{n}{m^2})^{-1/2}) \leq G(\tau)$ for all n and m.

(B2) $J(n,m)l(\tau(\frac{n}{m^2})^{-1/2}) \to K(\tau)$ as $n \to \infty$ and $m \to \infty$ uniformly on bounded subsets of $\mathbb{R}$.

(B3) $\int_{-\infty}^{\infty} K(\tau + s)\exp\{-\frac{1}{2}\tau^2\}d\tau$ has a strict minimum at $s = 0$.

(B4) $G(\cdot)$ satisfies (K1) and (K2).

Define the simulated posterior risk

$$B_{n,m}(\phi) := \int_\theta l(\theta, \phi)p(\theta|Y^{n,m})d\theta.$$

A regular simulated Bayes estimator $\tilde{\theta}^{n,m}$ based on $Y^{n,m}$ is defined as

$$\tilde{\theta}^{n,m} := \arg \inf_{\phi \in \Theta} B_{n,m}(\phi).$$

Assume that such an estimator exists.

The following theorem shows that SMLE and simulated Bayes estimators are asymptotically equivalent as $n \to \infty$ and $m \to \infty$.

Theorem 4.3. *Assume that (K1)-(K2) and (B1)-(B4) hold. Then we have:*
(i) $(\frac{n}{m^2})^{1/2}(\tilde{\theta}^{n,m} - \hat{\theta}^{n,m}) \to 0$ *a.s.-*$[P_{\theta_0}]$ *as* $n \to \infty$ *and* $m \to \infty$.
(ii) $\lim_{n\to\infty,\ m\to\infty} J(n,m)B_{n,m}(\tilde{\theta}^{n,m}) = \lim_{n\to\infty,\ m\to\infty} J(n,m)B_{n,m}(\hat{\theta}^{n,m})$

$$= (\frac{I(\theta_0)}{2\pi})^{1/2}\int_{-\infty}^{\infty} K(\tau)\exp(-\frac{I(\theta_0)}{2}\tau^2)d\tau \quad \text{a.s. } [P_{\theta_0}].$$

Proof. The proof is analogous to Theorem 4.1 in Borwanker et al (1972). We omit the details. □

Corollary 4.3. *Under the assumptions of Theorem 4.1, we have:*
(i) $\tilde{\theta}^{n,m} \to \theta_0$ *a.s.* $[P_{\theta_0}]$ *as* $n \to \infty$ *and* $m \to \infty$.
(ii) $(\frac{n}{m^2})^{1/2}(\tilde{\theta}^{n,m} - \theta_0)\xrightarrow{\mathcal{L}}\mathcal{N}(0, I^{-1}(\theta_0))$ *as* $n \to \infty$ *and* $m \to \infty$.

Proof. (i) and (ii) follow easily by combining Theorem 4.1 and strong consistency and asymptotic normality results of the SMLE. □

5. Example

For the Ornstein-Uhlenbeck process

$$dY_t = \theta Y_t dt + dW_t,$$

with $\theta < 0$, the simulated transition density $P^m(x, \cdot|\theta)$ is normal with mean $x(1+\theta/m)^m$ and variance $[(1+\theta/m)^{2m} - 1]/[\theta(\theta/m+2)]$ if $\theta \neq 0,\ -2m$ which converges to the true transition density $P(x, \cdot|\theta)$ which is normal density with mean xe^{θ} and variance $[e^{2\theta} - 1]/[2\theta]$ for all θ. The SMLE is

$$\hat{\theta}^{n,m} = m\Big\{\Big(\frac{\sum_{i=1}^{n} Y_i Y_{i-1}}{\sum_{i=1}^{n} Y_{i-1}^2}\Big)^{1/m} - 1\Big\}$$

which converges in probability to θ_0. If we choose the loss function to be $l(\theta, \phi) = l(|\theta - \phi|) = (\theta - \phi)^2$, then the simulated Bayes estimator $\tilde{\theta}^{n,m}$ would be the posterior mean.

References

[1] J.P.N. Bishwal, Rates of convergence of the posterior distributions and the Bayes estimators in the Ornstein-Uhlenbeck process, *Random Operators and Stochastic Equations*, **8** (2000), 51-70.

[2] J.P.N. Bishwal, The Bernstein-von Mises Theorem and spectral asymptotics of Bayes estimators for parabolic SPDEs, *J. Australian Math. Society*, **72**, No. 2 (2001), 289-300.

[3] J.P.N. Bishwal, Rates of weak convergence of the approximate minimum contrast estimator for the discretely observed Ornstein-Uhlenbeck process, *Statistics and Probability Letters*, **76**, No. 13 (2006), 1397-1409.

[4] J.P.N. Bishwal, A new estimating function for discretely sampled diffusions, *Random Operators and Stochastic Equations*, **15**, No. 1 (2007), 65-88.

[5] J.P.N. Bishwal, *Parameter Estimation in Stochastic Differential Equations*, Lecture notes in Mathematics, **1923**, Springer-Verlag (2008).

[6] J.P.N. Bishwal, A. Bose, Rates of convergenc of approximate maximum likelihood estimators in the Ornstein-Uhlenbeck process, *Comput. Math. Appl.*, **42** (2001), 23-38.

[7] J.D. Borwanker, G. Kallianpur, B.L.S. Prakasa Rao, The Bernstein-von Mises Theorem for Markov processes, *Ann. Math. Statist.*, **42** (1971), 1241-1253.

[8] J. Cano, M. Kessler, D. Salmeron, Approximation of the posterior density for diffusion processes, *Statist. Probab. Letters*, **76** (2006), 39-44.

[9] E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag (1999).

[10] A.R. Pederson, A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations, *Scand. J. Statist.*, **22** (1995), 55-71.

[11] A.R. Pederson, Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes, *Bernoulli*, **3** (1995), 257-279.