数理解析研究所講究録
第 777 巻 1992 年 68-83

Applications of the Malliavin calculus to higher order statistical inference *

Nakahiro YOSHIDA
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku,
Tokyo 106, Japan

November, 1991

Using the Malliavin calculus we can derive asymptotic expansions of mean values of statistics related to small diffusions. Applications to statistics and economics are mentioned.

1.Introduction

In this paper we review the applications of the Malliavin calculus to certain statistical problems. Consider a family of d-dimensional diffusion processes

$$\begin{aligned} dX_t^\epsilon &= V_0(X_t^\epsilon)dt + \epsilon V(X_t^\epsilon)dw_t, \\ X_0 &= x_0,\ t \in [0,T],\ \epsilon \in (0,1], \end{aligned} \tag{1.1}$$

where w is an r-dimensional standard Wiener process, V_0 and $V = (V_0, \cdots, V_r)$ are $\mathbf{R}^d$-valued and $\mathbf{R}^d \otimes \mathbf{R}^r$-valued smooth functions defined on $\mathbf{R}^d$ with bounded derivatives. T and x_0 are constants and ϵ is a parameter. Many statistical problems associated with this small diffusion are related to functionals of the form

$$F_T^\epsilon = F_0^\epsilon + \int_0^T f_0^\epsilon(X_t^\epsilon)dt + \epsilon \int_0^T f^\epsilon(X_t^\epsilon)dw_t, \tag{1.2}$$

where $F_0^\epsilon \in \mathbf{R}^k$, $f_0^\epsilon(x)$ is an $\mathbf{R}^k$-valued smooth function and $f^\epsilon(x)$ is an $\mathbf{R}^k \otimes \mathbf{R}^r$-valued smooth function.

*AMS 1980 subject classifications. 62E20, 62F12, 62M05
Key words and phrases. Small diffusion, Malliavin calculus, asymptotic expansion.

In the problem of pricing path dependent options in economics, the price X_t of underlying security is supposed to satisfy the one dimensional stochastic differential equation

$$(1.3)\qquad \begin{aligned} dX_t &= cX_t dt + \epsilon X_t dw_t,\\ X_0 &= x_0 \end{aligned}$$

where c and x_0 are constants. To price average options at time $t = 0$ we have to calculate the expectation

$$(1.4)\qquad E[Max\{Z_T^\epsilon - K, 0\}],$$

where

$$Z_T^\epsilon = \frac{1}{T}\int_0^T X_t dt$$

and K is a striking price, see Kunitomo-Takahashi (1990). The time-average Z_T^ϵ is in the form of (1.2). It is difficult to express this expectation explicitly, so several methods involving FFT or the numerical analysis for partial differential equations have been proposed. Kunitomo-Takahashi proposed that the log normal approximation to the distribution of Z_T^ϵ is applicable when ϵ is small. Here one problem is to derive the asymptotic expansion of (1.4) in ϵ-power.

A statistical problem occurs if the equation describing the observation process involves unknown parameters. Let X_t^ϵ satisfy the stochastic differential equation

$$(1.5)\qquad \begin{aligned} dX_t^\epsilon &= V_0(X_t^\epsilon, \theta)dt + \epsilon V(X_t^\epsilon)dw_t,\\ X_0 &= x_0,\ t \in [0,T],\ \epsilon \in (0,1], \end{aligned}$$

where V_0 and V are smooth functions and θ is an unknown parameter. The consistency and the first order properties of the maximum likelihood estimator and Bayes estimator when $\epsilon \downarrow 0$ are known, cf. Tsitovich (1977), Kutoyants (1978, 1984) and Genon-Catalot (1990). As for the higher order properties, the asymptotic expansions for distributions of the maximum likelihood estimator and the Bayes estimator were proved in Yoshida (1990a,b,1991a). When discussing the second order optimal property of these estimators we need asymptotic expansions of statistics such as log likelihood ratio statistics, Akahira-Takeuchi (1981). The log likelihood function and estimating equations of the M-estimator in parameter estimation problem are examples of functional (1.2). So it is necessary to show the asymptotic expansion for F_T^ϵ of (1.2).

The process X_t^ϵ defined by (1.1) converges to X_t^0 satisfying the differential equation

$$\frac{dX_t^0}{dt} = V_0(X_t^0),\ t \in [0,T],\ X_0^0 = x_0.$$

For instance, it is known that

$$\sup_{0\le t\le T} |X_t^\epsilon - X_t^0| \to 0$$

a.s. as $\epsilon \downarrow 0$. Hence, under some regularity conditions, we see that

$$F_T^\epsilon \to f_{-1}$$

a.s. as $\epsilon \downarrow 0$, where

$$f_{-1} = F_0^0 + \int_0^T f_0^0(X_t^0)dt.$$

Therefore it is more convenient to treat

$$\tilde{F}^\epsilon = \frac{F_T^\epsilon - f_{-1}}{\epsilon}$$

than F_T^ϵ. In view of the above two examples we consider the functional defined by

$$\varphi^\epsilon(\tilde{F}^\epsilon) I_A(\tilde{F}^\epsilon), \tag{1.6}$$

where $\varphi^\epsilon(x)$, $\epsilon \in [0,1]$, are smooth functions on $\mathbf{R}^k$, and give the asymptotic expansion of the mean value of this functional of F_T^ϵ

$$E[\varphi^\epsilon(\tilde{F}^\epsilon) I_A(\tilde{F}^\epsilon)], \tag{1.7}$$

which covers the asymptotic expansions of the distribution and moments of $\tilde{F}_T^\epsilon$ over some regions of the sample space. Then, it is a simple matter to obtain the asymptotic expansions of the mean value (1.4). Moreover, from our results we get the asymptotic expansions of the likelihood ratio statistic, which yield the bounds of the probability of concentration for second order asymptotically median unbiased estimators. These bounds serve to show the second order efficiency of the maximum likelihood estimator and the Bayes estimator.

Let us discuss from the mathematical view point. Assume that a family of the random variables F_ϵ has the asymptotic expansion

$$F_\epsilon \sim f_0 + \epsilon f_1 + \cdots$$

as $\epsilon \downarrow 0$ in some sense. Then, if function $T(x)$ satisfies certain regularity, we have the stochastic expansion

$$T(F_\epsilon) \sim \Phi_0 + \epsilon\Phi_1 + \cdots \tag{1.7}$$

as $\epsilon \downarrow 0$, where Φ_0, $\Phi_1, \cdots$ are determined by formal Taylor expansion and in particular $\Phi_0 = T(f_0)$ and $\Phi_1 = f_1 T'(f_0)$. Expectation of (1.7) yields the asymptotic expansion

$$E[T(F_\epsilon)] \sim E[\Phi_0] + \epsilon E[\Phi_1] + \cdots \tag{1.8}$$

as $\epsilon \downarrow 0$. If we are able to take $T(x) = I_A(x)$, the indicator function of the Borel set A, (1.7) gives formally

$$I_A(F_\epsilon) \sim I_A(f_0) + \epsilon f_1 I'_A(f_0) + \cdots \tag{1.9}$$

and hence

$$P(F_\epsilon \in A) \sim P(f_0 \in A) + \epsilon E[f_1 I'_A(f_0)] + \cdots \tag{1.10}$$

as $\epsilon \downarrow 0$. Thus we can obtain quite directly the asymptotic expansion of the distribution of F_ϵ. Unfortunately two difficulties arise. The second term of the right hand side of (1.9) is the composite function of the random variable f_0 and the Schwartz distribution I'_A. There is no usual meaning of a random variable as a measurable function on a probability space. So we have a problem of how to define or interpret such composite functionals. After removing this difficulty we still meet the question of how to justify the formal expansion (1.9). These difficulties have been solved by S. Watanabe (1983,1987) in the Malliavin calculus. He introduced the notion of the pull-back of the Wiener functional under the Schwartz distribution to justify the composite functionals in certain Sobolev spaces as generalized Wiener functionals (i.e. the Schwartz distribution on the probability space) and also exploited the method of the asymptotic expansion of generalized Wiener functionals to apply the heat kernels of the diffusion processes. This theory has been proved to be successfully applicable to the problems of the higher order statistical inference, Yoshida (1990a,b,1991a,b). To use this theory, the crucial step is to show the nondegeneracy of the Malliavin covariance of functionals. However, it dose not seem easy to do this even for a simple statistical estimator, whose Malliavin covariance is given by an integration of some non-adaptive process, cf., Yoshida (1990a). The Malliavin covariance corresponding to (1.2) is also written in a similar manner. Moreover, as for estimators appearing in parameter estimation, such as maximum likelihood estimators, we can not ensure their existence on the whole sample space, in general. So we need a modification of this theory with truncation. From these results we can derive asymptotic expansions of mean values of functionals of F_T^ϵ and asymptotic expansion of the distribution of the maximum likelihood estimator and the Bayes estimator.

The organization of this paper is as follows. In Section 2 we state the results for the asymptotic expansion of the mean value of the functional of $\tilde{F}_\epsilon$. Section 3 presents the illustrative examples of these results. Here we do not go into the technical details and all proofs of the results stated in this paper except Theorem 3.1 are omitted. For these points see Yoshida (1990a), (1990b), (1991a). Concluding this section we remark that our method is applicable to statistical models with discrete time parameter as well as models with continuous one. For example, we have similar results for AR(p) model with small noise, as this is the case where the model fitting does well, if they are realized on a Wiener space, Yoshida (1991b).

2.Asymptotic expansions of the mean values of functionals of F_T^ϵ

Let an $\mathbf{R}^d \otimes \mathbf{R}^d$-valued process $Y_t^\epsilon(w)$ be the solution of the stochastic differential equation

$$dY_t^\epsilon = \partial V_0(X_t^\epsilon)Y_t^\epsilon dt + \epsilon \textstyle\sum_{\alpha=1}^r \partial V_\alpha(X_t^\epsilon)Y_t^\epsilon dw_t^\alpha, \ t \in [0,T],$$
$$Y_0^\epsilon = I_d,$$

where $[\partial V_\alpha]^{ij} = \partial_j V_\alpha^i$, $\partial_j = \partial/\partial x^j$, $i,j = 1,\cdots,d$, $\alpha = 0,1,\cdots r$. Then, $Y_t := Y_t^0$ is a non-singular deterministic $\mathbf{R}^d \otimes \mathbf{R}^d$ -valued process. For function g^ϵ, $g^{(j)}$ denotes its j-th derivative in ϵ at $\epsilon = 0$. Let

$$x^0 = F_0^{(1)} + \int_0^T f_0^{(1)}(X_t^0)dt$$

and

$$a_t = \int_t^T \partial f_0^0(X_s^0)Y_s ds Y_t^{-1}V(X_t^0) + f^0(X_t^0).$$

We write $D_t = X_t^{(1)}$ and $E_t = X_t^{(2)}$. Let $\varphi^\epsilon(x)$, $\epsilon \in [0,1]$ be smooth functions on $\mathbf{R}^k$. In this paper we assume the following conditions for F_T^ϵ given in (1.2).

(1) The function F_0^ϵ is smooth in ϵ. The functions $\varphi^\epsilon(x), f_0^\epsilon(x)$ and $f^\epsilon(x)$ are smooth in (ϵ, x) on $[0,1] \times \mathbf{R}^d$, and any derivative is of polynomial growth order in x uniformly in ϵ.
(2) The matrix $\Sigma = \int_0^T a_t a_t' dt$ is positive definite.

We will use Einstein's rule for repeated indices. For matrix A, $[A]^{i,j}$ denotes its (i,j)-element. Moreover, $[A]^{i\cdot}$ and $[A]^{\cdot j}$ are the i-th row vector and the j-th column vector of A, respectively. For vector a, a^i is its i-th element. Put $\sigma^{ij} = [\Sigma]^{i,j}$ and $\sigma_{ij} = [\Sigma^{-1}]^{i,j}$. Let

$$\tilde{F}^\epsilon = \frac{F_T^\epsilon - f_{-1}}{\epsilon}.$$

Define as follows.

$$\begin{aligned}
&\lambda_{t,s} = Y_t Y_s^{-1} V(X_s^0),\\
&\lambda_{t,s}^i = [Y_t Y_s^{-1} V(X_s^0)]^{i\cdot},\ i = 1,\cdots,d,\\
&\mu_{i,t,s} = Y_t Y_s^{-1} \partial_i V(X_s^0),\ i = 1,\cdots,d,\\
&\mu_{i,t,s}^j = [Y_t Y_s^{-1} \partial_i V(X_s^0)]^{j\cdot},\ i,j = 1,\cdots,d,\\
&\nu_{i,j,t,s} = Y_t Y_s^{-1} \partial_i \partial_j V_0(X_s^0),\ i,j = 1,\cdots,d,\\
&\nu_{i,j,t,s}^l = [Y_t Y_s^{-1} \partial_i \partial_j V_0(X_s^0)]^l,\ i,j,l = 1,\cdots,d.
\end{aligned}$$

For $\mathbf{R}^1 \otimes \mathbf{R}^r$-valued function h_t,

$$C_0(h)_T = \int_0^T a_t h_t' dt,$$

For $\mathbf{R}$-valued function h_t,

$$C_1(h)_T = \int_0^T \int_0^t h_t \lambda_{t,s} a_s' ds dt.$$

For $\mathbf{R}^1 \otimes \mathbf{R}^r$-valued function h_t,

$$C_2^i(h)_T = \int_0^T \int_0^t a_t h_t' \lambda_{t,s}^i a_s' ds dt,\ i = 1,\cdots,d.$$

For $\mathbf{R}^1 \otimes \mathbf{R}^r$-valued functions b_t and c_t, put

$$C_2(b,c)_T = \frac{1}{2}\int_0^T \int_0^T a_t [b_t' c_s + c_t' b_s] a_s' ds dt.$$

Let

$$C_2^{i,j}(t) = C_2(\lambda_{t,\cdot}^i I_{\{\cdot \le t\}}, \lambda_{t,\cdot}^j I_{\{\cdot \le t\}})_T.$$

Define

$$\begin{aligned}
&A^{0,\alpha}\\
&= \tfrac{1}{2} F_0^{(2),\alpha} + \tfrac{1}{2}\int_0^T f_0^{(2),\alpha}(X_t^0) dt\\
&\quad + \tfrac{1}{2}\int_0^T \partial_\beta f_0^{0,\alpha}(X_t^0) \int_0^t \int_0^s \nu_{i,j,t,s}^\beta \lambda_{s,u}^i (\lambda_{s,u}^j)' du ds dt\\
&\quad + \tfrac{1}{2}\int_0^T \int_0^t \partial_i \partial_j f_0^{0,\alpha}(X_t^0) \lambda_{t,s}^i (\lambda_{t,s}^j)' ds dt,\\
&A_l^{1,\alpha} = [C_1(\partial_i f_0^{(1),\alpha}(X_\cdot^0))_T]^{i,j} \sigma_{jl} + [C_0([f^{(1)}(X_\cdot^0)]^{\alpha,\cdot})_T]^j \sigma_{jl},\\
&A_{p,q}^{2,\alpha}\\
&= \tfrac{1}{2}\int_0^T \partial_i \partial_j f_0^{0,\alpha}(X_t^0) [C_2^{i,j}(t)]^{m,n} dt\, \sigma_{pm} \sigma_{qn}\\
&\quad + \tfrac{1}{2}\int_0^T \partial_\beta f_0^{0,\alpha}(X_t^0) \int_0^t \nu_{i,j,t,s}^\beta [C_2^{i,j}(s)]^{m,n} ds dt\, \sigma_{pm} \sigma_{qn}\\
&\quad + \int_0^T \partial_\beta f_0^{0,\alpha}(X_t^0) [C_2^i(\mu_{i,t,\cdot}^\beta)_t]^{m,n} dt\, \sigma_{pm} \sigma_{qn}\\
&\quad + [C_2^i([\partial_i f^0(X_\cdot^0)]^{\alpha\cdot})_T]^{m,n} \sigma_{pm} \sigma_{qn}.
\end{aligned}$$

Now we can show the following theorem. $\mathbf{B}^k$ denotes the Borel σ-field of $\mathbf{R}^k$.

THEOREM 2.1. *The following asymptotic expansion holds.*

$$E[\varphi^\epsilon(\tilde{F}^\epsilon)I_A(\tilde{F}^\epsilon)] \\ \sim \int_A p_0(x)dx + \epsilon \int_A p_1(x)dx + \cdots,$$

as $\epsilon \downarrow 0$. *This expansion is uniform in* $A \in \mathbf{B}^k$. *In particular,*

$$\begin{aligned} p_0(x) &= \varphi^0(x)\phi(x; x^0, \Sigma), \\ p_1(x) &= \varphi^{(1)}(x)\phi(x; x^0, \Sigma) \\ &\quad + \varphi^0(x)\phi(x; x^0, \Sigma)[-A^{1,\alpha}_\alpha \\ &\quad + (A^{0,\alpha}\sigma_{\alpha p} - A^{2,\alpha}_{q,l}\sigma^{ql}\sigma_{\alpha p} - A^{2,\alpha}_{p,\alpha} - A^{2,\alpha}_{\alpha,p})[x - x^0]^p \\ &\quad + A^{1,\alpha}_p\sigma_{\alpha q}[x - x^0]^p[x - x^0]^q + A^{2,\alpha}_{p,q}\sigma_{\alpha l}[x - x^0]^p[x - x^0]^q[x - x^0]^l], \end{aligned}$$

where $\phi(x; x^0, \Sigma)$ *is the density of the normal distribution with mean* x^0 *and covariance matrix* Σ.

3.Examples

3.1. Applications to economics

In the problem of pricing path dependent options, the price X_t of underlying security is supposed to satisfy the one dimensional stochastic differential equation

$$\begin{aligned} dX_t &= cX_t dt + \epsilon X_t dw_t, \\ X_0 &= x_0 \end{aligned} \tag{3.2}$$

where c and x_0 are constants. To price average options at time $t = 0$ we have to calculate the expectation

$$E[Max\{Z^\epsilon_T - K, 0\}],$$

where

$$Z^\epsilon_T = \frac{1}{T}\int_0^T X_t dt$$

and K is a striking price, see Kunitomo-Takahashi (1990) and its references. It is difficult to express this expectation explicitly, so several methods involving FFT or the numerical analysis for partial differential equations have been proposed. Kunitomo-Takahashi proposed that the log normal approximation to the distribution of Z^ϵ_T is applicable when ϵ is small. We can derive asymptotic expansions for the diffusion defined by (1.1). Here we will only consider the asymptotic expansion of the distribution of Z^ϵ_T though we can treat transforms of Z^ϵ_T in the same manner.

Let X^ϵ satisfy the stochastic differential equation (1.1). For $F_T^\epsilon = Z_T^\epsilon$, the arithmetic mean on $\{X_t; 0 \leq t \leq T\}$, and $\varphi^\epsilon(x) = 1$, Theorem 2.1 gives the asymptotic expansion

$$\begin{aligned} &P[\epsilon^{-1}(Z_T^\epsilon - \tfrac{1}{T}\textstyle\int_0^T X_t^0 dt) \in A] \\ &\sim \textstyle\int_A p_0^Z(x)dx + \epsilon \int_A p_1^Z(x)dx + \cdots, \end{aligned}$$

as $\epsilon \downarrow 0$ uniformly in $A \in \mathbf{B}^d$. In particular,

$$\begin{aligned} p_0^Z(x) &= \phi(x; 0, \Sigma), \\ p_1^Z(x) &= \phi(x; 0, \Sigma)[(A^{0,\alpha}\sigma_{\alpha i} - A_{j,l}^{2,\alpha}\sigma^{jl}\sigma_{\alpha i} - A_{i,\alpha}^{2,\alpha} - A_{\alpha,i}^{2,\alpha})x^i \\ &\quad + A_{i,j}^{2,\alpha}\sigma_{\alpha,l}x^i x^j x^l], \end{aligned}$$

where

$$\begin{aligned} A^{0,\alpha} &= \tfrac{1}{2T}\textstyle\int_0^T \int_0^t \int_0^s \nu_{i,j,t,s}^\alpha \lambda_{s,u}^i (\lambda_{s,u}^j)' dudsdt \\ A_{p,q}^{2,\alpha} &= \tfrac{1}{2T}\textstyle\int_0^T \int_0^t \nu_{i,j,t,s}^\alpha [C_2^{i,j}(s)]^{m,n} dsdt\sigma_{pm}\sigma_{qn} \\ &\quad + \tfrac{1}{T}\textstyle\int_0^T [C_2^i(\mu_{i,t,\cdot}^\alpha)]^{m,n} dt\sigma_{pm}\sigma_{qn}. \end{aligned}$$

Here, $x^0 = 0$ and

$$a_t = \int_t^T \frac{1}{T} Y_s ds Y_t^{-1} V(X_t^0).$$

For example when X is the geometric Brownian motion (3.2),

$$p_1^Z(x) = \phi(x; 0, \Sigma) A_{1,1}^{2,1}(-3x + \Sigma^{-1}x^3),$$

where

$$\Sigma = x_0^2 (cT)^{-2} \int_0^T (e^{cT} - e^{ct})^2 dt$$

and

$$\begin{aligned} &A_{1,1}^{2,1} \\ &= \tfrac{1}{T}\textstyle\int_0^T \int_0^t \int_0^s \Sigma^{-2} x_0^3 (cT)^{-2} e^{ct}(e^{cT} - e^{cs})(e^{cT} - e^{cu}) dudsdt. \end{aligned} \tag{3.3}$$

We note that the asymptotic expansions given $\{X_s; 0 \leq s \leq t\}$, $t < T$ reduce to the unconditional case because of Markov property.

We may derive the asymptotic expansion of $E[Max\{Z_T^\epsilon - K, 0\}]$ directly in our context. Suppose that a one dimensional diffusion X satisfies (1.1) for $d = 1$. Let $F_T^\epsilon = Z_T^\epsilon$, $\varphi^\epsilon(x) = \epsilon x + f_{-1} - K$, $f_{-1} = \frac{1}{T}\int_0^T X_t^0 dt$ and $A = \{Z_T^\epsilon \geq K\}$. Then by Theorem 2.1 we see

$$\begin{aligned} &E[Max\{Z_T^\epsilon - K, 0\}] \\ &= E[(Z_T^\epsilon - K)I_A] \\ &\sim (f_{-1} - K)\textstyle\int_{x \geq \frac{K - f_{-1}}{\epsilon}} \phi(x; 0, \Sigma) dx \\ &\quad + \epsilon \textstyle\int_{x \geq \frac{K - f_{-1}}{\epsilon}} p_1(x) dx \\ &\quad + \epsilon^2 \textstyle\int_{x \geq \frac{K - f_{-1}}{\epsilon}} p_2(x) dx + \cdots, \end{aligned} \tag{3.4}$$

where

$$p_1(x) = x\phi(x;0,\Sigma) + (f_{-1} - K)\phi(x;0,\Sigma)[(A^{0,1}\Sigma^{-1} - 3A^{2,1}_{1,1})x \\ +A^{2,1}_{1,1}\Sigma^{-1}x^3]$$

and some smooth function $p_2(x)$. For (3.2)

$$p_1(x) = x\phi(x;0,\Sigma) + (f_{-1} - K)A^{2,1}_{1,1}(-3x + \Sigma^{-1}x^3)\phi(x;0,\Sigma)$$

with $A^{2,1}_{1,1}$ given in (3.3). If we want expansions in ϵ-power, we have from (3.4) that when $K - f_{-1} < 0$,

$$E[Max\{Z^{\epsilon}_T - K, 0\}] \\ \sim (f_{-1} - K) + \epsilon^2 \int p_2(x)dx + \cdots;$$

when $K - f_{-1} > 0$,

$$E[Max\{Z^{\epsilon}_T - K, 0\}] \sim O(\epsilon^n)$$

for $n = 1, 2, \cdots$; and when $K - f_{-1} = 0$,

$$E[Max\{Z^{\epsilon}_T - K, 0\}] \\ \sim \epsilon \int_{x\geq 0} x\phi(x;0,\Sigma)dx + \epsilon^2 \int_{x\geq 0} p'_2(x)dx + \cdots.$$

for some function $p'_2(x)$. Finally, we note that $p_2(x)$ in (3.4) can be specified by

$$\int_A p_2(x)dx \\ = E[f_1 I_A(f_0)] + E[f_0 f_1 \partial I_A(f_0)] \\ +(f_{-1} - K)E[f_2 \partial I_A(f_0) + \tfrac{1}{2}(f_1)^2 \partial^2 I_A(f_0)]$$

for $A \in \mathbf{B}^1$. It is possible to compute the right hand side from conditional expectations of fourfold Wiener integrals given a Wiener integral.

Using (3.4) we obtained similar numerical results for an example given in Kunitomo-Takahashi (1990).

3.2. Likelihood ratio statistic

Consider a parametric model of the d-dimensional small diffusions defined by

$$dX^{\epsilon}_t = V_0(X^{\epsilon}_t, \theta)dt + \epsilon V(X^{\epsilon}_t)dw_t, \\ X_0 = x_0,\ t \in [0,T],\ \epsilon \in (0,1], \tag{1.5}$$

where θ is an unknown parameter in $\bar{\Theta}$, the closure of a bounded convex domain Θ of $\mathbf{R}^k$. The likelihood function is given by the formula

$$\Lambda_\epsilon(\theta; X) = \exp\{\int_0^T \epsilon^{-2}V'_0(VV')^{+}(X_t,\theta)dX_t - \frac{1}{2}\int_0^T \epsilon^{-2}V'_0(VV')^{+}V_0(X_t,\theta)dt\},$$

where A^+ denotes the Moore-Penrose generalized inverse matrix of the matrix A, Liptser and Shiryayev (1977). Let $\theta_0 \in \Theta$ denote the true value of the unknown parameter θ. For $h \in \mathbf{R}^k$, the log likelihood ratio

$$\begin{aligned}
&l_{\epsilon,h}(w;\theta_0)\\
&= \log \Lambda_\epsilon(\theta_0 + \epsilon h; X) - \log \Lambda_\epsilon(\theta_0; X)\\
&= \textstyle\int_0^T \epsilon^{-1}[V_0(X_t^\epsilon, \theta_0 + \epsilon h) - V_0(X_t^\epsilon, \theta_0)]'(VV')^+ V(X_t^\epsilon)dw_t\\
&\quad -\tfrac{1}{2}\textstyle\int_0^T \epsilon^{-2}[V_0(X_t^\epsilon, \theta_0 + \epsilon h) - V_0(X_t^\epsilon, \theta_0)]'(VV')^+(X_t^\epsilon)[V_0(X_t^\epsilon, \theta_0 + \epsilon h) - V_0(X_t^\epsilon, \theta_0)]dt,
\end{aligned}$$

where X_t^ϵ is the solution of the above stochastic differential equation for $\theta = \theta_0$. Here we assume that $V_0(x,\theta) - V_0(x,\theta_0) \in M\{V(x)\}$: the linear manifold generated by column vectors of $V(x)$, for each x and θ. The asymptotic expansion of the distribution of the log likelihood ratio plays an important role in the higher order asymptotic theory to derive bounds of probability of concentration of statistics, see Akahira and Takeuchi (1981). We can obtain the asymptotic expansion from the result of Section 2.1.

Let $\delta_i = \partial/\partial\theta^i$. The Fisher information matrix $I(\theta_0) = (I_{ij}(\theta_0))$ is defined by

$$I_{ij}(\theta_0) = \int_0^T \delta_i V_0(X_t^0, \theta_0)'(VV')^+(X_t^0)\delta_j V_0(X_t^0, \theta_0)dt$$

for $i, j = 1, \cdots, k$. Let $I(\theta_0) = I = (I_{ij})$ and $I^{-1} = (I^{ij})$. We use the multi-index: $|\mathbf{n}| = n_1 + n_2 + \cdots + n_d$ and $\partial^{\mathbf{n}} = \partial_1^{n_1}\partial_2^{n_2}\cdots\partial_d^{n_d}$ for $\mathbf{n} = (n_1, n_2, \cdots, n_d)$; $|\nu| = \nu_1 + \nu_2 + \cdots + \nu_k$ and $\delta^\nu = \delta_1^{\nu_1}\delta_2^{\nu_2}\cdots\delta_k^{\nu_k}$ for $\nu = (\nu_1, \nu_2, \cdots, \nu_k)$.

In this subsection we assume the following conditions.

(1) V_0, V and $(VV')^+$ are smooth in $(x,\theta) \in \mathbf{R}^d \times \bar{\Theta}$.
(2) For $|\mathbf{n}| \geq 1$, $F = V_0, V, (VV')^+$, $\sup_{x,\theta} |\partial^{\mathbf{n}} F| < \infty$.
(3) For $|\nu| \geq 1$, a constant C_ν exists and

$$\sup_\theta |\delta^\nu V_0| \leq C_\nu(1 + |x|)^{C_\nu},$$

for all x.
(4) $I(\theta)$, $\theta \in \bar{\Theta}$ are positive definite.

Let

$$A_{i,j,n} = \frac{1}{2}\int_0^T \int_0^t [\partial_l\{\delta_i V_0'(VV')^+\delta_j V_0\}](X_t^0, \theta_0)[Y_t]^{l,m}[g_s]^{m,n} dsdt$$

and

$$B_{i,j,l} = \int_0^T \delta_i\delta_j V_0'(VV')^+\delta_l V_0(X_t^0, \theta_0)dt,$$

where
$$g_s = Y_s^{-1}VV'(VV')^+(X_s^0)\delta V_0(X_s^0,\theta_0)$$
$$= Y_s^{-1}\delta V_0(X_s^0,\theta_0).$$

Then, the following theorem is a corollary of Theorem 2.1.

THEOREM 3.1. *The probability distribution of the log likelihood ratio* $l_{\epsilon,h}(w;\theta_0)$, $h \neq 0$, *has the asymptotic expansion*

$$P[l_{\epsilon,h}(w;\theta_0) \in A] \sim \int_A p_0^L(x)dx + \epsilon \int_A p_1^L(x)dx + \cdots, \quad \textit{as } \epsilon \downarrow 0, \ A \in \mathbf{B}^1.$$

The expansion is uniform in $A \in \mathbf{B}^1$. *In particular,*

$$\begin{aligned} &p_0^L(x) = \phi(\overline{x};0,J),\\ &p_1^L(x) = [A_{i,j,l}h^ih^jh^l]J^{-3}[\overline{x}^3 - J\overline{x}^2 - 3J\overline{x} + J^2]\phi(\overline{x};0,J)\\ &+\tfrac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-2}[\overline{x}^2 - J\overline{x} - J]\phi(\overline{x};0,J), \end{aligned}$$

where $J = h'I(\theta_0)h$ *and* $\overline{x} = x + \frac{1}{2}J$. *The probability distribution function of* $l_{\epsilon,h}(w;\theta_0)$ *has the asymptotic expansion*

$$\begin{aligned} &P[l_{\epsilon,h}(w;\theta_0) \leq x] \sim \Phi(\overline{x};0,J)\\ &-\epsilon\{[A_{i,j,l}h^ih^jh^l]J^{-2}[\overline{x}^2 - J\overline{x} - J]\\ &+\tfrac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-1}[\overline{x} - J]\}\phi(\overline{x};0,J) + \cdots, \end{aligned}$$

where $\Phi(x;\mu,\sigma^2)$ *is the probability distribution function of the one-dimensional normal distribution with mean* μ *and variance* σ^2.

PROOF. We use Theorem 2.1 for $F_T^\epsilon = \epsilon l_{\epsilon,h}(w;\theta_0)$ and $\varphi^\epsilon(x) = 1$. Then, it is easy to show that

$$\begin{aligned} &x^0 = -\tfrac{J}{2},\\ &a_t = h^i\delta_i V_0'(X_t^0,\theta_0)(VV')^+V(X_t^0),\\ &p_{f_0}(x) = \phi(x;-\tfrac{J}{2},J)\\ &A^{0,1} = -\tfrac{1}{2}B_{i,j,l}h^ih^jh^l,\\ &A_1^{1,1} = -A_{i,j,l}J^{-1}h^ih^jh^l + \tfrac{1}{2}B_{i,j,l}J^{-1}h^ih^jh^l,\\ &A_{1,1}^{2,1} = A_{i,j,l}J^{-2}h^ih^jh^l. \end{aligned}$$

Thus we have the result. □

In this case, non-singularity of the Fisher information corresponds to the non-degeneracy of the Malliavin covariance. We also obtain the asymptotic expansion of the likelihood ratio statistic under the contiguous alternative $\theta_0 + \epsilon h$. By the argument of Yoshida (1990b), the following theorem can be proved.

THEOREM 3.2. *The probability distribution of the log likelihood ratio $l_{\epsilon,h}(w;\theta_0+\epsilon h)$ has the asymptotic expansion*

$$P[l_{\epsilon,h}(w;\theta_0+\epsilon h)\in A] \sim \int_A p_0^{Lc}(x)dx + \epsilon\int_A p_1^{Lc}(x)dx + \cdots, \quad as\ \epsilon\downarrow 0,\ A\in\mathbf{B}^1.$$

The expansion is uniform in $A\in\mathbf{B}^1$. In particular,

$$\begin{aligned} p_0^{Lc}(x) &= \phi(\underline{x};0,J),\\ p_1^{Lc}(x) &= [A_{i,j,l}h^ih^jh^l]J^{-3}[\underline{x}^3+2J\underline{x}^2-(3J-J^2)\underline{x}-2J^2]\phi(\underline{x};0,J)\\ &\quad +\tfrac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-2}[\underline{x}^2+J\underline{x}-J]\phi(\underline{x};0,J), \end{aligned}$$

where $\underline{x}=x-\frac{1}{2}J$. The probability distribution function of $l_{\epsilon,h}(w;\theta_0+\epsilon h)$ has the asymptotic expansion

$$\begin{aligned} &P[l_{\epsilon,h}(w;\theta_0+\epsilon h)\le x] \sim \Phi(\underline{x};0,J)\\ &+\epsilon\{[A_{i,j,l}h^ih^jh^l]J^{-2}[-\underline{x}^2-2J\underline{x}+J-J^2]\\ &+\tfrac{1}{2}[B_{i,j,l}h^ih^jh^l]J^{-1}[-\underline{x}-J]\}\phi(\underline{x};0,J)+\cdots. \end{aligned}$$

From now let us discuss the second order efficiency of estimators. Here we adopt the criterion by probability of concentration introduced by Takeuchi and Akahira. See Akahira-Takeuchi (1981), also see Taniguchi (1991) for time series. To avoid meaningless super-efficiency, an invariant condition is imposed on estimators in question. For simplicity we confine ourselves to considering the case $k=1$. An estimator T_ϵ is second order asymptotically median unbiased (second order AMU) if for any $\theta_0\in\Theta$ and any $c>0$

$$\lim_{\epsilon\downarrow 0}\sup_{\theta\in\Theta:|\theta-\theta_0|<\epsilon c}\epsilon^{-1}|P_\theta[\epsilon^{-1}(T_\epsilon-\theta)\le 0]-\frac{1}{2}|=0$$

and

$$\lim_{\epsilon\downarrow 0}\sup_{\theta\in\Theta:|\theta-\theta_0|<\epsilon c}\epsilon^{-1}|P_\theta[\epsilon^{-1}(T_\epsilon-\theta)\ge 0]-\frac{1}{2}|=0.$$

Then, we have the following theorem from Neymann-Pearson's fundamental lemma, Theorems 3.1 and 3.2., see Yoshida (1990b) for details.

THEOREM 3.3. *For any second order AMU estimator T_ϵ, for $h>0$*

$$\liminf_{\epsilon\downarrow 0}\epsilon^{-1}\{\Phi(J;0,J)+\epsilon[A_{1,1,1}h^3+\frac{1}{2}B_{1,1,1}h^3]\phi(J;0,J)-P_{\theta_0}[\epsilon^{-1}(T_\epsilon-\theta_0)\le h]\}\ge 0$$

and for $h < 0$

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-1}\{\Phi(-J;0,J)-\epsilon[A_{1,1,1}h^3+\frac{1}{2}B_{1,1,1}h^3]\phi(J;0,J)-P_{\theta_0}[\epsilon^{-1}(T_\epsilon-\theta_0) \leq h]\} \leq 0.$$

The second order distributions

$$\Phi(J;0,J) + \epsilon[A_{1,1,1}h^3 + \frac{1}{2}B_{1,1,1}h^3]\phi(J;0,J)$$

and

$$\Phi(-J;0,J) - \epsilon[A_{1,1,1}h^3 + \frac{1}{2}B_{1,1,1}h^3]\phi(J;0,J)$$

are called the bounds of second order distributions. An AMU estimator attaining these bounds for any $h > 0$ and $h < 0$ is called to be second order efficient.

Here we consider the maximum likelihood estimator and the Bayes estimators with respect to the quadratic loss functions and discuss their second order properties. When $\theta = \theta_0$ is true, the maximum likelihood estimator is denoted by $\hat{\theta}_\epsilon(w;\theta_0)$ and the Bayes estimator by $\tilde{\theta}_\epsilon(w;\theta_0)$. Let $\pi(\theta)$ be the Bayes prior which is a positive smooth function on Θ with all derivatives of polynomial growth order.

It is known that maximum likelihood estimators and Bayes estimators are asymptotically efficient for regular statistical experiments induced mainly from independent observations. As for the small diffusions, they have consistency and asymptotic normality and they are efficient in the first order, see, e.g., Kutoyants (1984). Here we are interested in their second order efficiency. For function $f(x,\theta)$ of x and θ, $f_t^\epsilon(\theta)$ denotes $f(X_t^\epsilon,\theta)$. We assume that

(5) for $\theta \in \bar{\Theta}$, $\theta \neq \theta_0$

$$\int_0^T [V_{0,t}^0(\theta) - V_{0,t}^0(\theta_0)]'(VV')_t^{+0}[V_{0,t}^0(\theta) - V_{0,t}^0(\theta_0)]dt > 0$$

as well as Conditions (1)-(4).

In the context of the higher order statistical asymptotic theory we need to modify the estimator to get an efficient estimator. We call an estimator $\hat{\theta}_\epsilon^*$ a bias corrected maximum likelihood estimator if

$$\hat{\theta}_\epsilon^* = \hat{\theta}_\epsilon - \epsilon^2 b(\hat{\theta}_\epsilon),$$

where $b(\theta)$ is a bounded smooth function with bounded derivatives. Similarly, we call an estimator $\tilde{\theta}_\epsilon^*$ a bias corrected Bayes estimator if

$$\tilde{\theta}_\epsilon^* = \tilde{\theta}_\epsilon - \epsilon^2 \tilde{b}(\tilde{\theta}_\epsilon),$$

where $\tilde{b}(\theta)$ is a bounded smooth function with bounded derivatives. The asymptotic expansion of the distributions of the maximum likelihood estimator and the Bayes estimator are known as in the following theorems (Yoshida 1990b,1991a).

Theorem 3.4. *The probability distribution of the bias corrected maximum likelihood estimator $\hat{\theta}^*_\epsilon(w;\theta_0+\epsilon h)$ under the contiguous alternative $P_{\theta_0+\epsilon h}$ has the asymptotic expansion*

$$
\begin{aligned}
&P[\tfrac{\hat{\theta}^*_\epsilon(w;\theta_0+\epsilon h)-(\theta_0+\epsilon h)}{\epsilon}\in A]\\
&\sim \textstyle\int_A p_0^c(y)dy+\epsilon\int_A p_1^c(y)dy+\cdots,\quad as\ \epsilon\downarrow 0,\ A\in\mathbf{B}^1,\ h\in\mathbf{R}^1,
\end{aligned}
$$

where $p_0^c, p_1^c, \cdots$ are smooth functions. The expansion is uniform in $A\in\mathbf{B}^1$ and $h\in K$, any compact set. In particular,

$$
\begin{aligned}
p_0^c(y) &= \phi(y;0,I^{-1}),\\
p_1^c(y) &= A_{1,1,1}[-y^3-hy^2+I^{-1}y+I^{-1}h]\phi(y;0,I^{-1})\\
&\quad +B_{1,1,1}[-\tfrac{1}{2}y^3-hy^2+I^{-1}y+I^{-1}h]\phi(y;0,I^{-1})\\
&\quad -b^1(\theta_0)Iy\phi(y;0,I^{-1}).
\end{aligned}
$$

THEOREM 3.5. *The probability distribution of the bias corrected Bayes estimator $\tilde{\theta}^*_\epsilon(w;\theta_0+\epsilon h)$ under the contiguous alternative $P_{\theta_0+\epsilon h}$ has the asymptotic expansion*

$$
\begin{aligned}
&P[\tfrac{\tilde{\theta}^*_\epsilon(w;\theta_0+\epsilon h)-(\theta_0+\epsilon h)}{\epsilon}\in A]\\
&\sim \textstyle\int_A \tilde{p}_0^c(y)dy+\epsilon\int_A \tilde{p}_1^c(y)dy+\cdots,\quad as\ \epsilon\downarrow 0,\ A\in\mathbf{B}^1,\ h\in\mathbf{R}^1,
\end{aligned}
$$

where $\tilde{p}_0^c, \tilde{p}_1^c, \cdots$ are smooth functions. The expansion is uniform in $A\in\mathbf{B}^1$ and $h\in K$, any compact set. In particular,

$$
\begin{aligned}
\tilde{p}_0^c(y) &= \phi(y;0,I^{-1}),\\
\tilde{p}_1^c(y) &= A_{1,1,1}[-y^3-hy^2+I^{-1}y+I^{-1}h]\phi(y;0,I^{-1})\\
&\quad +B_{1,1,1}[-\tfrac{1}{2}y^3-hy^2-\tfrac{1}{2}I^{-1}y+I^{-1}h]\phi(y;0,I^{-1})\\
&\quad -\tilde{b}(\theta_0)^1Iy\phi(y;0,I^{-1})+\pi(\theta_0)^{-1}\delta_l\pi(\theta_0)y\phi(y;0,I^{-1}).
\end{aligned}
$$

From Theorems 3.4 and 3.5 we see that the maximum likelihood estimator and the Bayes estimator are second order AMU if their biases are corrected by

$$
b(\theta_0) = -I^{-2}(\theta_0)A_{1,1,1}
$$

and

$$\tilde{b}(\theta_0) = -I^{-2}(\theta_0)A_{1,1,1} - \frac{3}{2}I^{-2}(\theta_0)B_{1,1,1} + I^{-1}\pi(\theta_0)^{-1}\delta\pi(\theta_0),$$

respectively. Then, we can prove that they are second order efficient comparing the bounds of Theorem 3.3 with the second order distributions of Theorems 3.4 and 3.5 for $h = 0$.

References

For the Malliavin calculus see Malliavin (1976,1978), Watanabe (1983,1984,1987), Ikeda and Watanabe (1984,1989), and Kusuoka and Strook (1984).

Akahira,M. and Takeuchi,K. (1981). Asymptotic efficiency of statistical estimators: concepts and higher order asymptotic efficiency, Lecture Notes in Statistics, 7, Springer, New York.

Genon-Catalot, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations, Statistics, 21, 99-116.

Ikeda,N. and Watanabe,S. (1984). An introduction to Malliavin's calculus, Stochastic Analysis, Proc. Taniguchi Intern. Symp. on Stochastic Analysis, Katada and Kyoto 1982, ed. by K. Itô, 1-52, Kinokuniya/North-Holland, Tokyo.

Ikeda,N. and Watanabe,S. (1989). Stochastic differential equations and diffusion processes, second edition, North-Holland/Kodansha, Tokyo.

Kunitomo,N. and Takahashi,A. (1990). Pricing average options, discussion paper, Faculty of Economics, University of Tokyo.(in Japanese)

Kusuoka,S. and Strook,D. (1984). Applications of the Malliavin calculus I, Stochastic Analysis, Proc. Taniguchi Internatl. Symp. Katata and Kyoto, 1982, K.Itô (ed.), 271-306, Kinokuniya, Tokyo.

Kutoyants,Yu.A. (1978). Estimation of a parameter of a diffusion process, Theory Probab. and Appl., 23, 641-649.

Kutoyants,Yu.A. (1984). Parameter estimation for stochastic processes, Translated and edited by B.L.S. Prakasa Rao, Heldermann.

Liptser,R.S. and Shiryayev,A.N. (1977).Statistics of random processes I, Springer.

Malliavin,P. (1976). Stochastic calculus of variation and hypoelliptic operators, Proc. Int. Symp. S.D.E. Kyoto, Kinokuniya, 195-263.

Malliavin,P. (1978). C^k-hypoellipticity with degeneracy, Stochastic Analysis, Academic Press, 199-214, 327-340.

Taniguchi,M. (1991). Higher order asymptotic theory for time series analysis, Lect. Notes in Statistics., 68, Springer, Berlin.

Tsitovich, I.I. (1977). On estimating the drift of a diffusion process from observations in a domain, Theor. Probab. and Appl., 22, 851-858.

Watanabe,S. (1983). Malliavin's calculus in terms of generalized Wiener functionals, Lecture Notes in Control and Inf. Sci., 49, Springer, New York.

Watanabe,S. (1984). Lectures on stochastic differential equations and Malliavin calculus, Tata Institute of Fundamental Research, Springer, Bombay.

Watanabe,S. (1987). Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels, Ann. Probab., 15, 1-39.

Yoshida,N. (1990a). Asymptotic expansion for small diffusion -an application of Malliavin-Watanabe theory, Research Memorandum, 383, Institute of Statistical Mathematics.

Yoshida,N. (1990b). Asymptotic expansion for small diffusion. An application of the theory of Malliavin-Watanabe. General case and second order efficiency, Research Memorandum, 392, Institute of Statistical Mathematics. To appear in Probab. Theory Related Fields.

Yoshida,N. (1991a). Asymptotic expansion for statistics related to small diffusion, Research Memorandum, 406, Institute of Statistical Mathematics.

Yoshida,N. (1991b). Asymptotic expansions for small noise system on Wiener space I: maximum likelihood estimator, Research Memorandum, 422 , Institute of Statistical Mathematics.

THE INSTITUTE OF STATISTICAL MATHEMATICS
4-6-7 MINAMI-AZABU, MINATO-KU,
TOKYO 106, JAPAN

EXPTR.TEX(910626)[~ 10625.TEX(910625)] EXPTR2.TEX(910801)
TENKAI.TEX(911224)SPELLC KOQROKU.TEX(920111)