ASYMPTOTIC NORMALITY OF CONDITIONAL INTEGRALS OF DIFFUSION PROCESSES

MONTSERRAT FUENTES

NCSU MIMEO SERIES #2520

NOVEMBER 3, 1999

ASYMPTOTIC NORMALITY OF CONDITIONAL INTEGRALS OF DIFFUSION PROCESSES

MONTSERRAT FUENTES [1]

North Carolina State University

Consider predicting the integral of a diffusion process Z in a bounded interval A, based on the observations $Z(t_{1n}), \ldots, Z(t_{nn})$, where $t_{1n}, \ldots, t_{nn}$ is a dense triangular array of points (the step of discretization tends to zero as n increases) in the bounded interval. The best linear predictor is generally not asymptotically optimal. Instead, we predict $\int_A Z(t)dt$ using the conditional expectation of the integral of the diffusion process, the optimal predictor in terms of minimizing the mean squared error, given the observed values of the process. We obtain that, conditioning on the observed values, the order of convergence in probability to zero of the mean squared prediction error is $O_p(n^{-2})$. We prove that the standardized conditional prediction error is approximately Gaussian with mean zero and unit variance, even though the underlying diffusion is generally non-Gaussian. Because the optimal predictor is hard to calculate exactly for most diffusions, we present an easily computed approximation that is asymptotically optimal. This approximation is a function of the diffusion coefficient.

[1] M. Fuentes is an assistant professor at the Statistics Department, North Carolina State University, NC 27695. Email address: fuentes@stat.ncsu.edu.

AMS 1991 subject classifications: Primary 62M20; secondary 62M40, 41A25.

Key words and phrases. Diffusion process, fixed-domain asymptotics, infill asymptotics, numerical integration.

1. Introduction. Consider predicting $\int_A Z(t)dt$ for a diffusion process Z based on observations $Z_{(n)} = (Z(t_{0n}), \ldots, Z(t_{nn}))$, where A is a bounded interval $[a, b]$ and $a = t_{0n} < \ldots < t_{nn} = b$, $n = 1, 2, \ldots$. We assume the step of discretization tends to zero as $n \to \infty$: $\lim_{n\to\infty} \max_{2\leq i\leq n} (t_{in} - t_{i-1,n}) = 0$. Let $\widehat{Z}_n(t) = E\big[Z(t)|Z_{(n)}\big]$ and $c_n^Z(t_1, t_2) = E\big[Z(t_1)Z(t_2)|Z_{(n)}\big] - \widehat{Z}_n(t_1)\widehat{Z}_n(t_2)$ be the conditional mean and covariance functions, respectively, for Z. For Z Gaussian, $\widehat{Z}_n(t)$ is a linear function of $Z_{(n)}$ and $c_n^Z(t_1, t_2)$ is nonrandom. However, for diffusions, $\widehat{Z}_n(t)$ is generally not linear in $Z_{(n)}$ and $c_n^Z(t_1, t_2)$ is random.

Numerical integration over deterministic functions is studied, for example, by Davis and Rabinowitz (1984), Ghizzetti and Ossccini (1970) and Chakravarti (1970) by taking a linear function of $Z_{(n)}$. Numerical integration over random functions taking a linear function of $Z_{(n)}$ is done by Marnevskaya and Jarovich (1984), Stein (1993 and 1995), Pitt, Robeva and Wang (1995), and Ritter (1995). Stein (1987) presents a non-linear approximation to the integral of a transformation of a Brownian motion process, this is a special case of the central idea in this paper. Considering the close relationship between predicting integrals and estimating regression coefficients from stochastic processes described by Sacks and Ylvisaker (1966), the work by Sacks and Ylvisaker (1966, 1968, 1970), Eubank, Smith and Smith (1981) and Wahba (1971, 1974) on designs for estimating regression coefficients are also relevant for the prediction problem.

Matheron (1985) has developed a technique for approximating the unconditional distribution of $\int_A Z(t)dt$ when the process is non-Gaussian, but we are interested in making inferences about $\int_A Z(t)dt$ given $Z_{(n)}$, so the conditional distributions are relevant.

The conditional expectation of the integral of the diffusion process,

$$E\left[\int_A Z(t)dt \middle| Z_{(n)}\right] = \int_A \widehat{Z}_n(t)dt,$$

is the optimal predictor (in the sense of minimizing the mean squared error) of $\int_A Z(t)dt$ given $Z_{(n)}$. Since this predictor is hard to calculate exactly for most diffusions, we present an approximation that still yields better results than the best linear predictor (BLP) and is asymptotically optimal as $n \to \infty$.

The primary purpose of this paper is to show that, under certain conditions, the standardized conditional prediction error

$$\frac{\int_A Z(t)dt - \int_A \widehat{Z}_n(t)dt}{\left[\mathrm{var}\left\{\int_A Z(t)dt - \int_A \widehat{Z}_n(t)dt\right\}\right]^{1/2}}$$

is approximately $N(0,1)$, even if the underlying process is non-Gaussian. We also obtain an easily computed approximation for the general nonlinear predictor, $\int_A \widehat{Z}_n(t)dt$, that is a function of the diffusion coefficient.

First, in Section 2, we show that for large n, the conditional distribution of the prediction error is approximately Gaussian. Next, in Section 3, we show that the BLP is generally not asymptotically optimal and give an easily computed asymptotically optimal predictor. Then, in Section 4, we give an approximation for the conditional standard error of the prediction. Finally, in Section 5, we show a simulation study to assess the accuracy of this asymptotic result in finite samples, and in Section 6 we present some final remarks.

2. Asymptotic Normality of the Prediction Error.

We predict the integral of a diffusion over the interval $A = [0,1]$. It is straightforward calculation to generalize the results in this paper to any other bounded interval in $\mathbb{R}$. Thus, for a homogeneous diffusion process Z on $[0,1]$, consider the prediction of

$$\int_0^1 Z(t)dt$$

based on observing $Z(t)$ at $0 = t_{0n} < t_{1n} \cdots < t_{nn} = 1$. Specifically, let $Z_{(n)} = (Z(t_{0n}), ..., Z(t_{nn}))$ and predict the integral by the optimal predictor

$$\int_0^1 \widehat{Z}_n(t)dt.$$

We will often write t_i to denote t_{in} for $i = 0, \cdots, n$ to simplify the notation. In this section we will show that the standardized error in predicting the integral of a diffusion process with the optimal predictor (in the mean squared error sense), given the observed values of the process, is asymptotically $N(0,1)$.

We will make the following assumptions about the diffusion process $Z()$ throughout this paper,

$$(A.1) \qquad E\{Z(t+s) - Z(t)|Z(t) = x\} = s\mu(x) + o(s)$$

$$(A.2) \qquad E\{(Z(t+s) - Z(t))^2|Z(t) = x\} = s\sigma^2(x) + o(s)$$

$$(A.3) \qquad E\{|Z(t+s) - Z(t)|^3|Z(t) = x\} = s^{1+1/2}k(x) + o(s^{1+1/2})$$

where the remainder terms in $(A.1) - (A.3)$ are uniform in x. Note that conditions $A.1 - A.3$ require the existence of finite conditional moments of orders 1, 2 and 3. In most practical examples of diffusions, the indicated moments exist. It would be preferable to have assumption $(A.3)$ in terms of the diffusion parameters, $\mu(x)$ and $\sigma^2(x)$. The diffusion parameter $\mu(x)$ is frequently called the drift coefficient, and $\sigma^2(x)$ the diffusion coefficient. The following assumption gives explicit hypothesis on the diffusion parameters, and implies $(A.3)$:

$(A.4)$ The diffusion parameters, μ and $\sigma^2$are bounded, (see Karatzas and Shreve, p. 367), i.e.

$$\mu(x) + \sigma^2(x) \leq \rho; \quad 0 \leq t < \infty,$$

and the diffusion coefficient, σ^2, is a strictly positive and uniformly bounded function that has a continuous and uniformly bounded derivative.

We prove now that $(A.4)$ implies $(A.3)$. For a diffusion process $B(t)$ with constant diffusion coefficient σ^2, we have:

$$(2.1) \qquad E\{|B(t+s) - B(t)|^3|B(t) = x\} = \sqrt{2\pi}\; s^{1+1/2}\sigma^3$$

We apply the transformation function g to the diffusion process Z, where $g(x) = \int_{z_0}^{x} \sigma(y)^{-1} dy$. Then $g\left(Z(t)\right) = B(t)$ has constant diffusion coefficient and satisfies (2.1). By $(A.4)$ we can obtain the following Taylor expansion for $Z(t+s) = g^{-1}(B(t+s))$ centered at t, where g^{-1} is the inverse function of g:

$$[Z(t+s) - Z(t)] = \{g^{-1}(B(t))\}' [B(t+s) - B(t)] +$$
$$\{g^{-1}(B(t+\epsilon))\}'' [B(t+s) - B(t+\epsilon)]^2$$

where $\epsilon \in (0, s)$. Because $B(t)$ satisfies (2.1), and the first and second derivatives of g^{-1} are uniformly bounded (by the definition of g and (A.4)),

$$\{g(z)^{-1}\}' = \sigma(z) \quad \text{and} \quad \{g(z)^{-1}\}'' = \sigma(z)\sigma(z)' = \frac{\{\sigma^2(z)\}'}{2},$$

it follows that $Z(t)$ satisfies $(A.3)$.

We will sometimes require the following set of conditions for the dense triangular array of points in the bounded interval:

$(B.1)$ $\quad t_{in} = F^{-1}\left(\frac{i}{n}\right), t_{0n} = 0$ and $t_{nn} = 1$

$(B.2)$ $\quad F$ is a continuous strictly monotone cdf on $[0,1]$ with derivative f

$(B.3)$ $\quad f$ is a continuous function.

A sequence that satisfies $B.1 - B.3$ is called a regular sequence by Sacks and Ylvisaker (1966).

We define $\pi_i(x,t)$ to be the conditional density that from the state value x at time t the sample path of $Z()$ satisfies $Z(t_{i+1}) = z_{i+1}$ at time t_{i+1}, and make the following assumption:

$(C.1)$ $\quad \pi_i(x,t)$ has two continuous derivatives with respect to x

and it is a differentiable function with respect to t.

It would be preferable to have the previous assumption, $(C.1)$, in terms of the diffusion coefficients. The following assumption, $(C.2)$, has explicit hypotheses on μ and σ and implies $(C.1)$:

$(C.2)$ $\quad \sigma(x)$ and $\mu(x)$ are bounded and uniformly

Hölder-continuous functions

(see Karatzas and Shreve, 1991, p. 368).

In the following discussion o and O have the usual interpretation in terms of order of magnitude statements, while o_p and O_p mean that o and O hold respectively with a probability that can be chosen arbitrarily close to one.

We now give the following definitions that will be needed in the proof of the main theorem in this section:

(i) *Equivalent measures.* Let $\widetilde{Z}$ be a diffusion process of time parameter t, where $t \in T = [0, 1]$ with values $\widetilde{Z}(t) = \widetilde{Z}(\omega, t), \quad \omega \in \Omega$, on a probability space $(\Omega, \Psi, \widetilde{P})$. We assume that the σ-algebra Ψ is generated by $\widetilde{Z}(t) = \widetilde{Z}(\omega, t)$ on Ω as the parameter t runs through the set T. Let P_1 be another measure on the σ-algebra Ψ. It is said to be absolutely continuous with respect to $\widetilde{P}$ if $P_1(A) = 0$ whenever $\widetilde{P}(A) = 0$ for $A \in \Psi$. Measures P_1 and $\widetilde{P}$ are said to be *equivalent* if they are mutually absolutely continuous.

(ii) *Transition density.* Let $p(t, x, y)$ denote the transition density of Z at time t. That is,

$$p(t, x, y)dy = p\{y < Z(t) \leq y + dy | Z(0) = x\}.$$

Then $p(t, x, z_{i+1})$ is the solution of the following backward equation,

$$\frac{\partial}{\partial t} p(t, x, z_{i+1}) = \frac{\sigma^2(x)}{2} \frac{\partial^2}{\partial x^2} p(t, x, z_{i+1}) + \mu(x) \frac{\partial}{\partial x} p(t, x, z_{i+1}) \tag{2.2}$$

where the Radon-Nikodym derivatives are supplied by the Girsanov formula (see Karatzas and Shreve, 1991).

(iii) *Conditioned diffusion process.* Let Z_i^* be a process on $[t_i, t_{i+1}]$ such that,

$$\mathcal{L}(Z_i^*) = \mathcal{L}\big(Z | Z(t_i) = z_i, Z(t_{i+1}) = z_{i+1}\big) \tag{2.3}$$

where $\mathcal{L}(X)$ denotes the distribution law of a random variable X. The conditioned diffusion process is itself a diffusion with a time-varying drift $\mu^*(x, t)$ and same diffusion coefficient, $\sigma^2(x)$, as the original process. The transition probability density for the conditioned process is given by

$$p^*(s,t,x,y) = \frac{p(t-s,x,y)\,p(t_{i+1}-t,y,z_{i+1})}{p(t_{i+1}-s,x,z_{i+1})}.$$

(iv) *Prediction Error.* We define PE_n, the error in predicting $\int_0^1 Z(t)dt$ given $Z_{(n)}$ with the optimal predictor; that is,

$$PE_n = \int_0^1 \{Z(t) - \widehat{Z}_n(t)\}dt.$$

Furthermore, SPE_n is the standardized prediction error given $Z_{(n)}$,

$$SPE_n = \frac{PE_n}{\left[\text{var}\left\{PE_n | Z_{(n)}\right\}\right]^{1/2}}.$$

We will prove in Theorem 2.1 that SPE_n is asymptotically $N(0,1)$.

We present the following lemma that will be used in the proof of Theorem 2.1. We assume that the diffusion process Z satisfies assumptions $(A.1 - A.3)$, $(B1 - B.3)$ and $(C.1)$.

The transition density of Z_i^* for $t_i \leq s < t \leq t_{i+1}$, is given by

$$\begin{aligned} p_i^*(t,x;s,y)dy &= p(y < Z(s) \leq y + dy | Z(t_i) = z_i, Z(t) = x, Z(t_{i+1}) = z_{i+1}) \\ &= \frac{p(s-t,x,y)\pi_i(y,s)dy}{\pi_i(x,t)} \end{aligned} \tag{2.4}$$

for $t < s$. The drift coefficient for Z_i^*, for $t_i < t < t_{i+1}$, is

$$\mu_i^*(x,t) = \lim_{h \downarrow 0} \frac{1}{h} \int_{-\infty}^{+\infty} (y-x) p_i^*(t,x;t+h,y)dy.$$

LEMMA 2.1

The conditioned diffusion process Z_i^* has a non-homogeneous infinitesimal mean:

$$\mu_i^*(z_i,t_i) = \mu(z_i) + \frac{\partial p(t_{i+1}-t_i,z_i,z_{i+1})/\partial z_i}{p(t_{i+1}-t_i,z_i,z_{i+1})}\sigma^2(z_i). \tag{2.5.}$$

The diffusion coefficient of the process Z_i^* is σ^2, the diffusion coefficient of the process Z.

Proof of Lemma 2.1:

Lemma 2.1 is a known result (see Karlin and Taylor, 1981, p. 267-268). The proof is very straightforward once we have the following Taylor expansion for $\pi_i(x,t)$:

$$\pi_i(y, t+h) = \pi_i(x,t) + (y-x)\frac{\partial \pi_i}{\partial x}(x,t) + h\frac{\partial \pi_i}{\partial t}(x,t) + o(y-x) + o(h). \tag{2.6}$$

The Main Theorem

THEOREM 2.1

Consider predicting the integral of a diffusion process Z over $[0,1]$, based on the observations $Z_{(n)} = (Z(t_{1n}), \ldots, Z(t_{nn}))$, by $\int_0^1 \widehat{Z}_n(t)dt$. Suppose

(*i*) *the parameters of the diffusion process Z satisfy relations $(A.1 - A.3)$*

(*ii*) *conditions $(B.1 - B.3)$ are satisfied by the sequence of points in $[0,1]$*

(*iii*) *the conditional probabilities π_i satisfy condition $(C.1)$.*

Then conditional on $Z_{(n)}$

$$SPE_n \xrightarrow{\mathcal{L}} N(0,1) \tag{2.7}$$

with probability 1.

Proof of Theorem 2.1:

Given $Z_{(n)}$, PE_n is a sum of independent, mean zero random variables, so this is a triangular array situation. We study the conditional cumulants of orders 1 to 3 of the prediction error because, applying Lyapounov's condition (Billingsley, 1995) for $\delta = 1$, we can prove the weak convergence of the distributions. Suppose $Y_{1n}, \cdots, Y_{nn}$ are independent random variables, such that $\mathcal{L}(Y_{in}) = \mathcal{L}\big(X_i | Z(t_i) = z_i, Z(t_{i+1}) = z_{i+1}\big)$, where $X_i = \int_{t_{i-1}}^{t_i} Z(t)dt$ for $i = 1, \ldots, n$. Thus,

$$\text{(2.8)} \qquad \mathcal{L}(Y_{1n} + \cdots + Y_{nn}) = \mathcal{L}\big(X_{1n} + \cdots + X_{nn} | Z_{(n)} = z\big) = \mathcal{L}\Big(\int_0^1 Z(t)dt \Big| Z_{(n)} = z\Big).$$

Therefore the problem is reduced to studying just the cumulants of $\int_{t_{i-1}}^{t_i} Z_i^*(t)dt$ given Z_i^* at time t_i, where the process $Z_i^*()$ is defined in (2.3). The infinitesimal mean of Z_i^* satisfies

$$\text{(2.9)} \qquad \lim_{h\downarrow 0} \frac{1}{h} E\Big\{ Z_i^*(t+h) - Z_i^*(t) \Big| Z_i^*(t_i) = z_i \Big\} = \mu_i^*(z_i, t_i).$$

Recall that we obtained an expression for μ_i^* in Lemma 2.1.

In Lemma 2.1 we also showed that the variance of Z_i^* is

$$\text{(2.10)} \qquad \lim_{h\downarrow 0} \frac{1}{h} \text{var}\Big\{ Z_i^*(t+h) - Z_i^*(t) \Big| Z_i^*(t_i) = z_i \Big\} = \sigma^2(z_i).$$

In addition to the infinitesimal relations (2.8) and (2.9), the following higher-order infinitesimal moment relation is satisfied when $(A.3)$ holds:

$$\text{(2.11)} \qquad \lim_{h\downarrow 0} \frac{1}{h} E\Big\{ |Z_i^*(t+h) - Z_i^*(t)|^3 \Big| Z_i^*(t_i) = z_i \Big\} = 0.$$

All this means that

$$\text{(2.12)} \qquad \begin{aligned} &\text{var}\left\{ PE_n \middle| Z_{(n)} = z_{(n)} \right\} \\ &= \sum_{i=0}^{n-1} \text{var}\Big\{ \int_{t_i}^{t_{i+1}} Z_i^*(t)dt \Big| Z_i^*(t_i) = z_i \Big\} \\ &= \sum_{i=0}^{n-1} (t_{i+1} - t_i)^3 \{\sigma^2(z_i)/12\} + o(n^{-2}). \end{aligned}$$

We get the first equality by applying (2.8), and the last expression by integrating the infinitesimal variance of $Z_i^*(t)$ when t belongs to $[t_i, t_{i+1}]$, where by assumptions $(A.1) - (A.3)$, the remainder term in (2.12) is uniform in $z_{(n)}$. Therefore, the order of the conditional standard error of the prediction error is $O(n^{-1})$ as $n \to \infty$. It follows from Eq. (2.12) and conditions $(B.1 - B.3)$ that $n\big[\text{var}\{PE_n|Z_{(n)}\}\big]^{1/2}$ converges in probability to the random variable $L = \int_0^1 \sigma^2(Z(t))\omega(t)dt$ as $n \uparrow \infty$, where $\omega(t) = \frac{1}{12}f(t)^{-3}$.

Using the same argument as the one leading to Eq. (2.12) together with (2.11), we get that

$$\text{cum}_3\left\{PE_n \middle| Z_{(n)}\right\} = o_p(n^{-3}).$$

Let $s_n^2\left(Z_{(n)}\right)$ be the conditional variance of $S_n\left(Z_{(n)}\right) = (X_{1n}+\cdots+X_{nn}) - E\left(X_{1n}+\cdots+X_{nn}\middle|Z_{(n)}\right)$ given $Z_{(n)}$. By Eqs. (2.11) and (2.12), we get

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{s_n^3\left(Z_{(n)}\right)}E\left[\left|X_{kn} - E\left(X_{kn}\middle|Z_{(n)}\right)\right|^3\middle|Z_{(n)}\right] = 0 \text{ with probability } 1,$$

which is Lyapounov's condition for $\delta = 1$. Since Lyapounov's condition holds, $\frac{S_n\left(Z_{(n)}\right)}{s_n\left(Z_{(n)}\right)} \xrightarrow{\mathcal{L}} N(0,1)$. If we write the result in terms of the diffusion process Z, then relation (2.7) holds and the conditional prediction error is asymptotically normal.

In the following corollary whose proof is obvious, we reformulate the assumptions of Theorem 2.1 in terms of the diffusion parameters, $\mu(x)$ and $\sigma^2(x)$.

COROLLARY 2.1.

Theorem 2.1 still holds if (A.3) is replaced by (A.4) in (i) and (C.1) is replace by (C.2) in (iii).

3. Approximating the Optimal Predictor. The conditional expectation of the integral of the diffusion process is the optimal predictor of $\int_0^1 Z(t)dt$ given $Z_{(n)}$. But because this predictor is hard to calculate exactly for most diffusions, we present the following approximation to it that yields an asymptotically optimal predictor:

$$I_n(Z_{(n)}) = \sum_{i=1}^{n}(t_{in} - t_{i-1,n})Z_i + \frac{1}{2n}\sum_{i=1}^{n}(t_{in} - t_{i-1,n})\{\sigma^2(Z_i)\}'. \tag{3.1}$$

Here, we write Z_i to denote $Z(t_{in})$ to simplify the notation, and $\{\sigma^2(Z)\}'$ to denote the derivative of σ^2 with respect to Z. Approximation (3.1) is a linear function of $Z_{(n)}$, $\sum_{i=1}^{n}(t_{in} - t_{i-1,n})Z_i$, plus a nonlinear term in $Z_{(n)}$ that is a function of the derivative of the diffusion coefficient, and could be thought of as the adjustment for the conditional bias of the BLP.

We define $\widehat{PE}_n$, the error in predicting $\int_0^1 Z(t)dt$ with $I_n(Z_{(n)})$,

$$\widehat{PE}_n = \int_0^1 Z(t)dt - I_n(Z_{(n)}).$$

In this section we will show that the error $\widehat{PE}_n - PE_n$, in approximating $\int_0^1 \widehat{Z}_n(t)dt$ with $I_n(Z_{(n)})$, is negligible compared to the conditional standard deviation of the prediction error PE_n. Then, by Theorem 2.1, the prediction error $\widehat{PE}_n$ is asymptotically normal.

THEOREM 3.1

Under assumptions (A.1)-(A.3), (B.1)-(B.3), and (C.1), conditional on $Z_{(n)}$

$$\frac{\widehat{PE}_n}{\left[\mathrm{var}\{PE_n | Z_{(n)}\}\right]^{1/2}} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$$

with probability 1.

We now present a proposition that will be used in the proof of Theorem 3.1.

PROPOSITION 3.1

Under conditions (A.1-A.3), (B.1-B.3) and (C.1),

$$\int_0^1 \widehat{Z}_n(t)dt - I_n(Z_{(n)}) = O_p(n^{-(1+\delta)}).$$

Proof of proposition 3.1:

We now prove that the remainder term when we approximate

$$\int_0^1 \widehat{Z}_n(t)dt$$

with $I_n(Z_{(n)})$ is of order $\sum_{i=0}^{n-1}\{F^{-1}(\frac{i+1}{n}) - F^{-1}(\frac{i}{n})\}^{2+\delta}$ for some $\delta > 0$. By conditions $(B.1-B.3)$, we get that

$$\sum_{i=0}^{n-1}\left\{F^{-1}\left(\frac{i+1}{n}\right) - F^{-1}\left(\frac{i}{n}\right)\right\}^{2+\delta} = O\left(n^{-(1+\delta)}\right).$$

Thus, the error in the approximation, $\widehat{PE}_n - PE_n$, is negligible compared to the conditional standard deviation of PE_n, which we will prove in proposition 4.1 is $O_p\left(n^{-1}\right)$.

By the Markov property of the diffusion process Z,

$$\int_0^1 \widehat{Z}_n(t)dt = \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} E(Z(t)|Z(t_i), Z(t_{i+1}))dt.$$

Letting $z = (z_0, \cdots, z_n)$ be the observed values of $Z()$ at times $t_0, \cdots, t_n$, by straightforward calculation

$$(3.2) \qquad \begin{aligned}\int_0^1 E\big(Z(t)|Z_{(n)} = z\big)dt = &\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} E\big(Z(t) - Z(t_i)|Z(t_i) = z_i, Z(t_{i+1}) = z_{i+1}\big)dt\\ &+\sum_{i=0}^{n-1} z_i(t_{i+1} - t_i).\end{aligned}$$

In (2.4) we showed that μ_i^*, the transition density for the conditioned process, was a function of π_i (that satisfies $(C.1)$) and p, the transition density of Z. So, we need an explicit expression for p. We transform the coordinates applying the transformation function g, where $g(x) = \int_{z_0}^{x} \sigma(y)^{-1} dy$. Then $g(Z)$ is a process with constant diffusion coefficient.

Now, if P is the probability measure of the diffusion process Z, let $\overline{P}$ be a probability measure such that P and $\overline{P}$ are equivalent, and under $\overline{P}$, Z has drift coefficient

$$\overline{\mu}(x) = -\sigma^2(x) g''(x) \{2g'(x)\}^{-1}. \tag{3.3}$$

It is routine to verify that the process $B = g(Z)$ is a Brownian motion under $\overline{P}$. Therefore, by a change of variables we get an equation equivalent to (2.2),

$$\frac{\partial p(t, b, b_{i+1})}{\partial t} = \frac{1}{2} \frac{\partial^2 p(t, b, b_{i+1})}{\partial b^2}. \tag{3.4}$$

The solution is a Gaussian density. By a change of variables again, we can express the solution in terms of the process Z (under $\overline{P}$) and obtain

$$\frac{\partial p(t_{i+1} - t_i, z_i, z_{i+1})/\partial z_i}{p(t_{i+1} - t_i, z_i, z_{i+1})} = -\frac{3}{2}(t_{i+1} - t_i)^{-1} \sigma(z_i) \int_{z_i}^{z_{i+1}} \frac{1}{\sigma(y)} dy,$$

and

$$\overline{\mu}(z_i) = \frac{\sigma'(z_i)\sigma(z_i)}{2}.$$

Assumption $(C.1)$ postulates sufficient regularity for $\pi_i(x,t)$ to permit the use of the following Taylor expansion:

$$\begin{aligned} \pi_i(y, t+h) = \pi_i(x,t) + (y-x)\frac{\partial \pi_i}{\partial x}(x,t) + \int_t^{t+h} (t+h-\tau)\frac{\partial \pi_i}{\partial \tau}(x,\tau) d\tau \\ + \int_x^y \frac{(y-\epsilon)^2}{2} \frac{\partial^2 \pi_i}{\partial \epsilon^2}(\epsilon, t) d\epsilon \end{aligned} \tag{3.5}$$

where $\epsilon \in (x,y)$.

We define R_i the residual term when we approximate $\int_{t_i}^{t_{i+1}} Z(t)dt$ with the conditionally optimal predictor restricted to the interval $[t_{i+1}, t_i] \in [0,1]$. By the definition of $p^*(t,x;s,y)$ in (2.4) and the Taylor expansion (3.5), we obtain

$$R_i = \int_{t_i}^{t_{i+1}} \frac{\partial^2 \pi_i(x,t)/\partial x^2}{\pi_i(x,t)} \int_{-\infty}^{+\infty} (y-x)^3 p(t,x,y)dydt + o(n^{-2}).$$

By a change of variables again, we can get an explicit expression for the second derivative of $\pi_i(x,t)$ with respect of of x, using the same argument as in (3.4). By assumption A.3, R_i is of order $O_p\left(n^{-(2+\delta)}\right)$, where δ is the same as in (A.3). Thus, we obtain

$$\frac{\sum_{i=1}^{n-1} R_i}{\left[\mathrm{var}\left\{PE_n \middle| Z_{(n)}\right\}\right]^{1/2}} = O_p\left(n^{-\delta}\right).$$

Then we get that (under $\overline{P}$)

$$\begin{aligned}\int_0^1 \widehat{Z}_n(t)dt &= \sum_{i=1}^{n-1} Z_i(t_{i+1}-t_i) + \sum_{i=1}^{n-1}(t_{i+1}-t_i)(Z_{i+1}-Z_i)\\ &\quad -\sum_{i=1}^{n-1} \frac{\sigma'(Z_i)\sigma(Z_i)}{2}(t_{i+1}-t_i)^2 + O_p\left(n^{-(1+\delta)}\right)\\ &= I_n(Z_{(n)}) + O_p\left(n^{-(1+\delta)}\right).\end{aligned} \tag{3.6}$$

Thus, under the probability measure $\overline{P}$,

$$\begin{aligned} n^{(1+\delta/2)}\Bigg(\widehat{Z}_{(n)}(t)dt &- \sum_{i=1}^{n-1} Z_i(t_{i+1}-t_i) + \sum_{i=1}^{n-1}(t_{i+1}-t_i)(Z_{i+1}-Z_i)\\ &-\sum_{i=1}^{n-1}\frac{\sigma'(Z_i)\sigma(Z_i)}{2}(t_{i+1}-t_i)^2\Bigg) = n^{(1+\delta/2)}\left(\widehat{PE}_n - PE_n\right) \xrightarrow{\overline{P}} 0.\end{aligned} \tag{3.7}$$

It is straightforward consequence of the equivalence of P and $\overline{P}$ that (3.7) holds under P as well. In other words,

$$n^{(1+\delta/2)}\left(\widehat{PE}_n - PE_n\right) \xrightarrow{P} 0.$$

Thus, we approximate the optimal predictor of the integral $\int_0^1 \widehat{Z}_n(t)dt$ with $I_n(Z_{(n)})$ and the error in this approximation, $\widehat{PE}_n - PE_n$ is negligible compared to the standard deviation of PE_n.

Proof of Theorem 3.1:

We proved in Theorem 2.1 that the standardized prediction error SPE_n is, conditional on $Z_{(n)}$, asymptotically $N(0,1)$. By Propositions 3.1 and 4.1, $\widehat{PE}_n - PE_n$ is negligible compared to the conditional standard deviation of PE_n, so Theorem 3.1 follows.

A symmetric approximation for the optimal predictor.

We present a symmetric approximation for the predictor that is invariant if the order of time for the diffusion on the interval $[0,1]$ is reversed. We get this approximation as the average of two approximations for the optimal predictor. The first approximation is obtained by assuming that we observe the diffusion process Z on $[0,1]$ starting at time 0, that is, $t_0 = 0$ and $t_n = 1$. For the second approximation, we assume the observations start at time 1, that is, $t_0^* = 1$, and $t_n^* = 0$. We obtain

$$\begin{aligned} \widetilde{I}_n\left(Z_{(n)}\right) = {} & \sum_{i=2}^{n-1} Z_i(t_{i+1}-t_{i-1}) + \sum_{i=2}^{n-1}(t_{i+1}-t_{i-1})(Z_{i+1}-Z_{i-1}) \\ & - \sum_{i=2}^{n-1} \frac{\sigma'(Z_i)\sigma(Z_i)}{2}(t_{i+1}-t_{i-1})^2 \\ & + \frac{1}{2}\left[Z_1(t_2-t_1) + Z_n(t_n-t_{n-1})\right] \\ & + \frac{1}{2}\left[(t_2-t_1)(Z_2-Z_1) + (t_n-t_{n-1})(Z_n-Z_{n-1})\right] \\ & - \frac{\sigma'(Z_1)\sigma(Z_1)}{4}(t_2-t_1)^2 - \frac{\sigma'(Z_n)\sigma(Z_n)}{4}(t_n-t_{n-1})^2. \end{aligned} \tag{3.8}$$

Thus,

$$\int_0^1 \widehat{Z}_n(t)dt = \widetilde{I}_n\left(Z_{(n)}\right) + O_p\left(n^{-1-\delta}\right).$$

4. The Variance of the Prediction Error. In the following proposition we prove that the order of convergence to zero of the conditional variance of the prediction error is $O_p(n^{-2})$. Then the order of convergence to zero of the error in approximating $\int_0^1 \widehat{Z}_n(t)dt$ by $I_n(Z_{(n)})$ is faster than the order of the conditional standard deviation of PE_n.

PROPOSITION 4.1

As $n \to \infty$,

$$\text{var}\left\{ PE_n \middle| Z_{(n)} \right\} = O_p(n^{-2}). \tag{4.1}$$

Proof of proposition 4.1:

By the Markov property of the diffusion process and using the same argument as in (3.2), the variance of the prediction error, when conditions $(A.1 - A.3)$, $(B.1 - B.3)$ and $(C.1)$ are satisfied, can be written as

$$\begin{aligned} &\sum_{i=0}^{n-1} \text{var}\left\{ \int_{t_i}^{t_{i+1}} Z(t)dt \middle| Z(t_i), Z(t_{i+1}) \right\} \\ &= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \int_{t_i}^{t_{i+1}} \text{cov}\{Z_i^*(x) Z_i^*(y) | Z_i^*(t_i)\} dxdy \\ &= \sum_{i=0}^{n-1} (t_{i+1} - t_i)^3 \{\sigma^2(Z(t_i))/12\} + o_p(n^{-2}). \end{aligned} \tag{4.2}$$

Where the process Z_i^* is defined in (2.3). Then the order of the standard error of the prediction error is $O_p(n^{-1})$.

A symmetric approximation for the variance of the prediction error.

Eq. (4.2) gives us an approximation for the conditional variance which can also be expressed in a symmetric manner

$$\sum_{i=2}^{n-1} \frac{(t_{i+1}-t_{i-1})^3\sigma^2(Z_i)}{12} + \left[\frac{(t_2-t_1)^3\sigma^2(Z_1)}{24} + \frac{(t_n-t_{n-1})^3\sigma^2(Z_n)}{24}\right]. \tag{4.3}$$

5. Simulation Results with Known Diffusion Parameters. To assess the accuracy of the proposed approximations to the integral of a diffusion process Z and to the variance of the prediction error in finite samples, we performed several simulation experiments, assuming that the diffusion coefficient was a known function.

We simulated n values of a diffusion process Z at times $t_0, \dots, t_{n-1}$ where $t_i = \frac{i}{n-1}$, for $i = 0, 1, \dots, n-1$, $n = 100$ and Z is a transformation of a Brownian motion B,

$$Z(t) = \frac{\exp\{B(t)\}}{[1+\exp\{B(t)\}]} \tag{5.1}$$

This diffusion process is a uniformly bounded process and, hence, trivially satisfies the assumptions $(A.4)$ and $(C.2)$.

The parameters of the diffusion Z, as defined in (5.1), can be written as a function of the Brownian motion process. The infinitesimal variance of Z is

$$\sigma^2(Z(t)) = \frac{\exp\{2B(t)\}}{[1+\exp\{B(t)\}]^4}.$$

Neither $I_n(Z_{(n)})$, nor the approximation for the standard error, in the prediction error depend on the drift coefficient of Z, which can be an unknown function. So, for the asymptotics, we do not need an expression for the infinitesimal mean of Z. Clearly, $I_n(Z_{(n)})$ is a function of the infinitesimal variance of Z and also of the derivative of that parameter, given by

$$\frac{\partial\sigma^2(Z(t))}{\partial Z(t)} = \frac{2\,[1-\exp\{B(t)\}]\exp\{B(t)\}}{[1+\exp\{B(t)\}]^3}.$$

We conducted the following simulation experiment to assess the asymptotic normality of the conditional prediction error:

(i) We simulated n values of a Brownian motion and applied the transformation (5.1) to get n observations of the diffusion process Z. We show in Figure 5.1 $Z_{(n)}$, the n simulated values of Z, for a single realization of Z. Using Eq. (3.8) and (2.12) we get the value of the predictor and the conditional variance of the prediction error, because the diffusion coefficient is known.

(ii) In order to simulate the distribution of the prediction error, we need the value of the predictor obtained in (i), the conditional variance of the prediction error obtained also in (i), and $\int_0^1 Z(t)dt$, which is unknown. Then we have to simulate $\int_0^1 Z(t)dt$ conditional on $Z_{(n)}$. In order to obtain values of $\int_0^1 Z(t)dt$, we simulated m values of $Z(t)$ at m equally spaced locations in time in every interval $[t_i, t_{i+1}]$, conditioning on the observed values $Z_{(n)}$. We conducted several simulation experiments with different values for m, and $m = 10$ seemed to be sufficiently large in the sense that increasing m has only a negligible effect on the results. Figure 5.2 shows a single realization of the $10n$ simulated values of Z conditional on the n observed values of $Z_{(n)}$. We considered the mean of the $10n$ simulated values of Z conditioning on $Z_{(n)}$ to be the unknown quantity that we want to predict, $\int_0^1 Z(t)dt$.

(iii) We repeated step (ii) k times to obtain k simulated values for $\int_0^1 Z(t)dt$. We also obtained k simulated values of the prediction error $\widehat{PE}_n$ and the variance of $\widehat{PE}_n$, always conditioning on the n observed values $Z_{(n)}$.

(iv) We plotted the histogram of the k simulated values of the prediction error $\widehat{PE}_n$ (see Figure 5.3 (a)). Here $k = 100$.

Then we repeated steps (i)-(iv) 100 times to study the conditionally asymptotic normality, but conditioning now on other simulated values of $Z_{(n)}$.

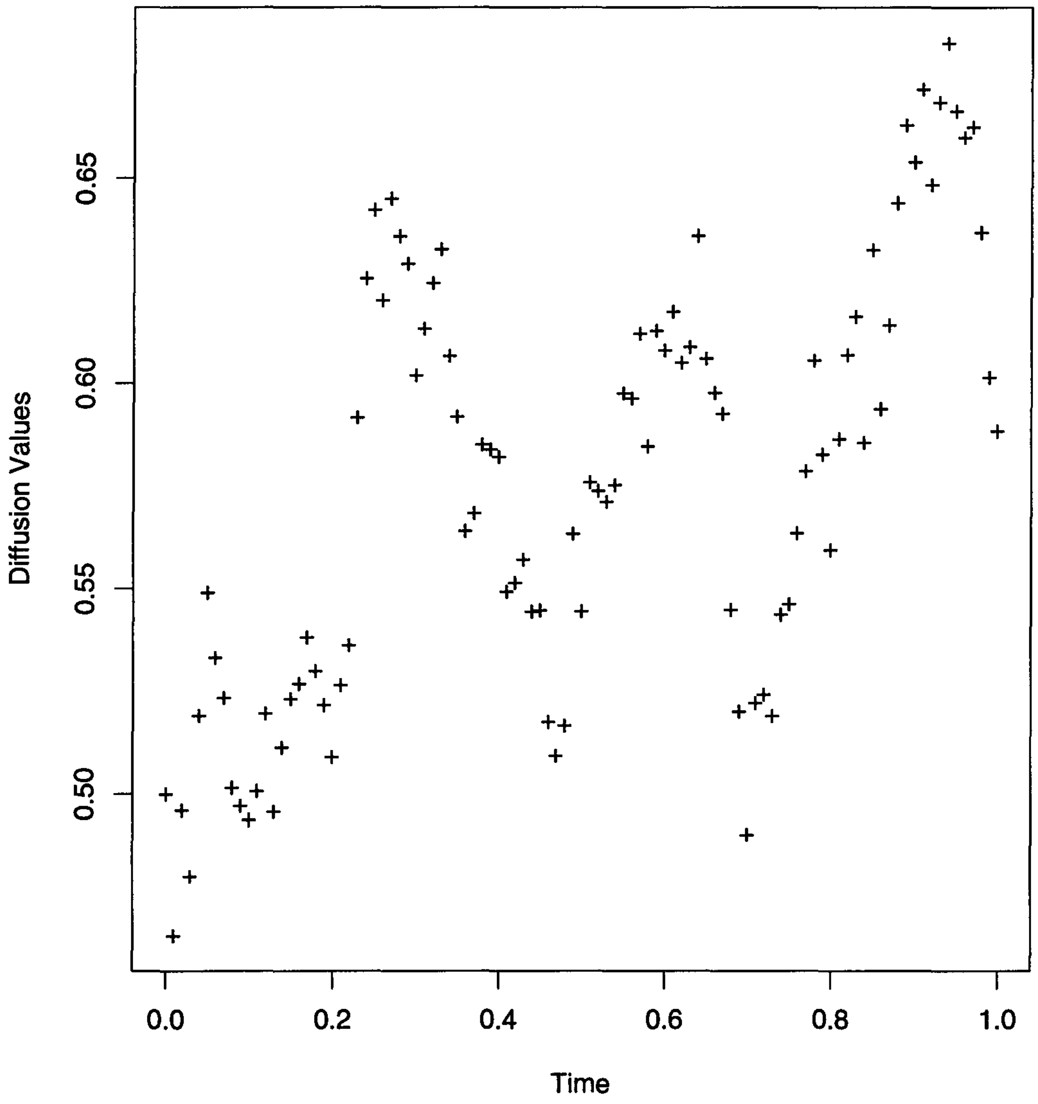


Figure 5.1. Simulation of $Z_{(n)}$, the observed values of the diffusion process Z, with $n = 100$.

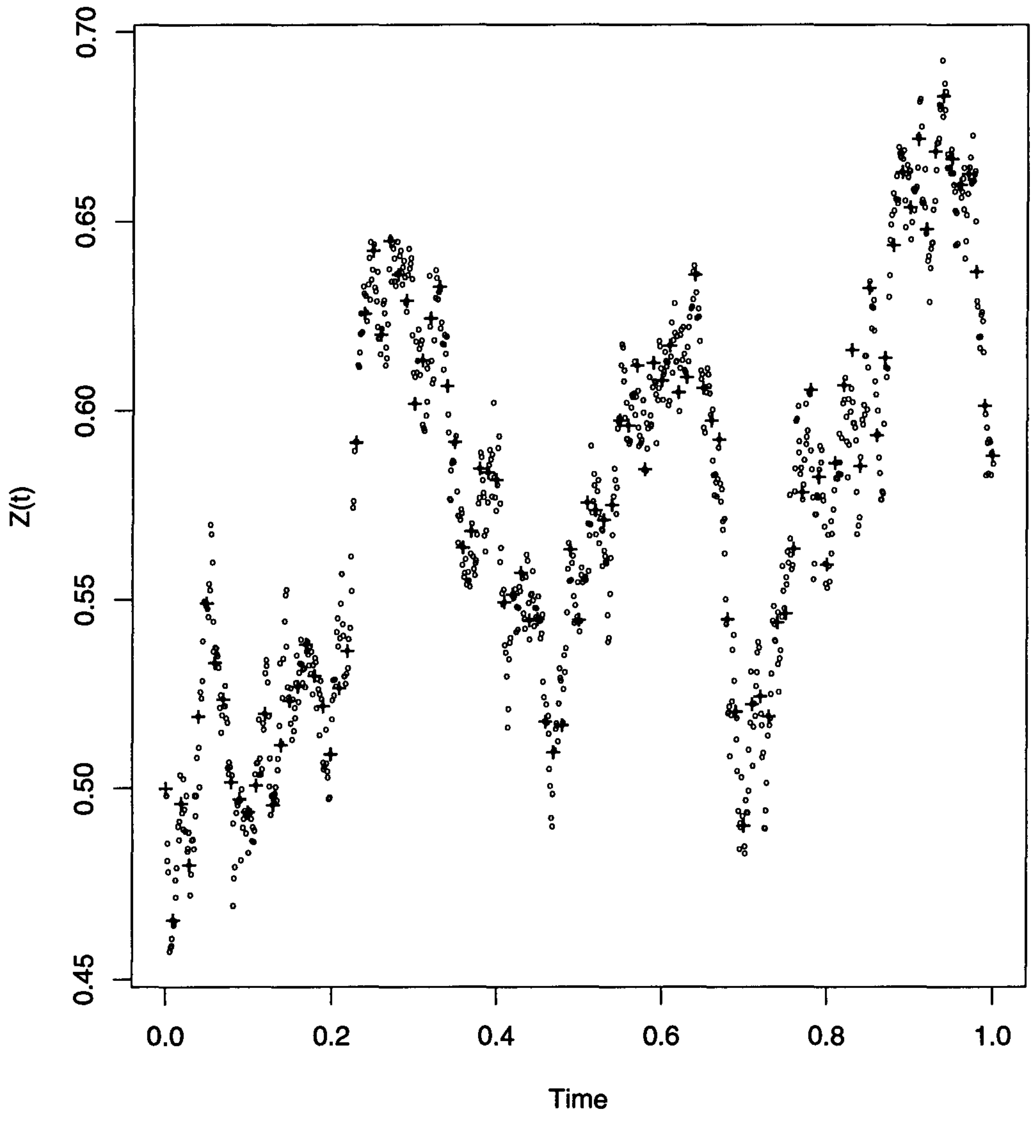
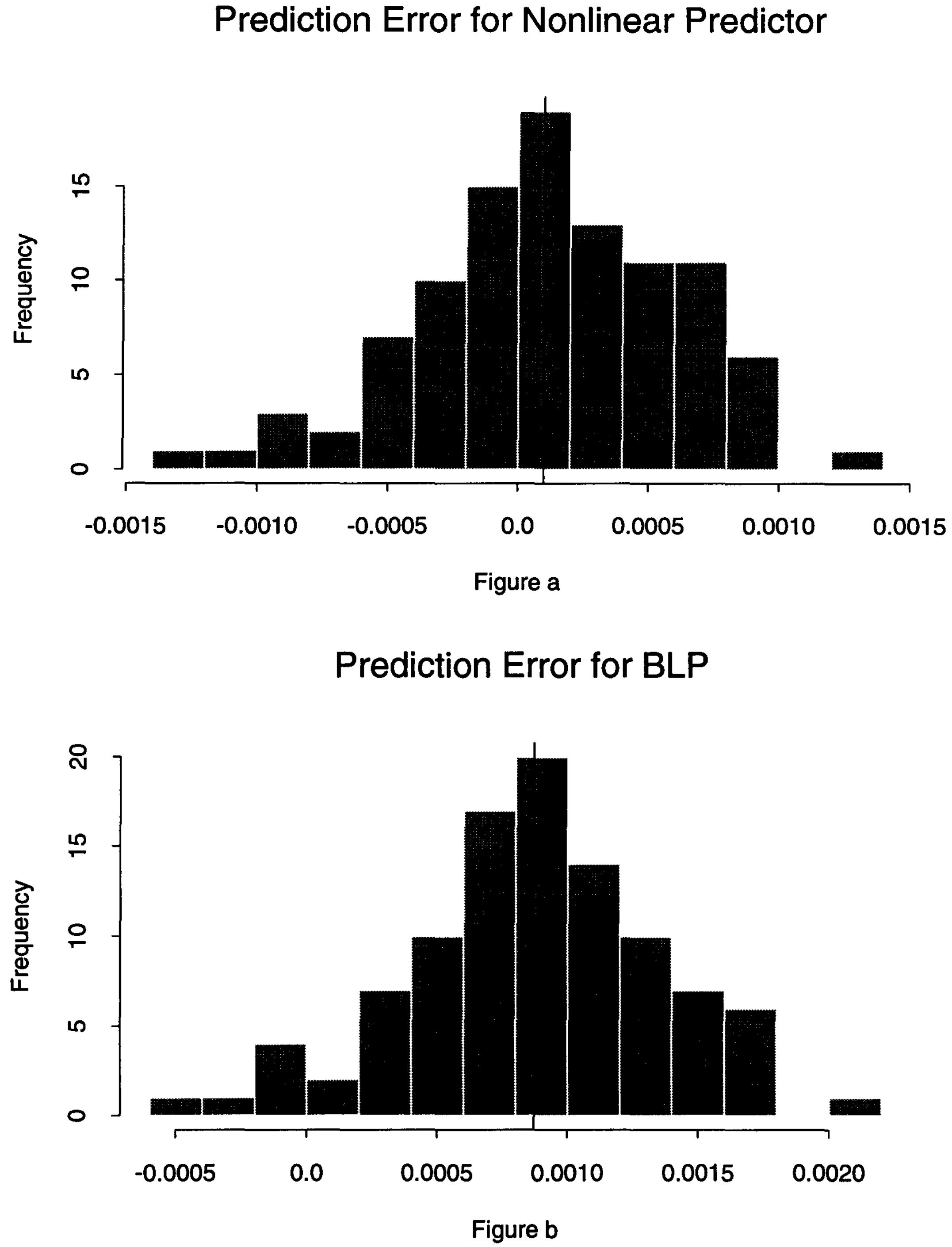


Figure 5.2. Simulation of $\int_0^1 Z(t)dt$ conditional on Z_n. The vertical axis shows 1000 simulated values of Z conditional on the n=100 observed values $Z_{(n)}$. The average of the 1000 simulated values approximates $\int_0^1 Z(t)dt$. The symbol + is used for the values of $Z_{(n)}$.

Figure 5.3 (a) is a histogram of the distribution of the prediction error using the proposed approximation for the optimal predictor, $I_n(Z_{(n)})$, for the particular $Z_{(n)} = z_{(n)}$ showed in Figure 5.1. Because the histogram is centered at zero and has a normal shape, this provides some evidence that the proposed predictor is conditionally unbiased and asymptotically normal. In this simulation study $n = 100$, $m = 10$ and $k = 100$. In the normality plot, Figure 5.4, we can see evidence of this asymptotic normality.

Figure 5.3 (b) shows the distribution of the prediction error when the predictor is the classic linear combination of the observed values. The vertical line in the graph is located at the mean value of the sampling distribution. The histogram is not centered at zero, which reflects the fact that this naive predictor, a linear combination of the observed values, has nontrivial conditional bias in this case (the value of the conditional bias is .00087), which is approximately $\frac{1}{2n}\int_0^1 \sigma(Z_t)\sigma'(Z_t)dt$.

Now we compare the mse of the linear and nonlinear predictor, for the particular $Z_{(n)} = z_{(n)}$ pictured in Figure 5.1, to show that the nonlinear predictor is better. The mse of the nonlinear predictor is 4.13×10^{-7} and the mse of the linear predictor is 8.01×10^{-7}, almost twice as large. In order to show that this result is not just due to a favorable value of $Z_{(n)}$, Table 5.1 compares the mse for more simulations conditioning on other values of $Z_{(n)}$. We compared the mse of the nonlinear predictor from the simulation conditional on the $Z_{(n)}$ shown in Figure 5.1 with the one obtained using the asymptotic approximation presented in Theorem 3.1, and the asymptotic mse is 3.97×10^{-7}, very close to the mse of the nonlinear predictor and roughly half the mse of the linear predictor.



Prediction Error for Nonlinear Predictor

Figure a

Prediction Error for BLP

Figure b

Figure 5.3. Distribution of the approximated prediction error, σ^2 known. (a) Distribution of the conditional prediction error when predicting $\int_0^1 Z(t)dt$ given $Z_{(n)} = z_{(n)}$ with the optimal predictor. The vertical axis represents the frequencies out of 100 simulations. (b) Distribution of the prediction error given $Z_{(n)} = z_{(n)}$ using the BLP. The conditional bias is .00087.

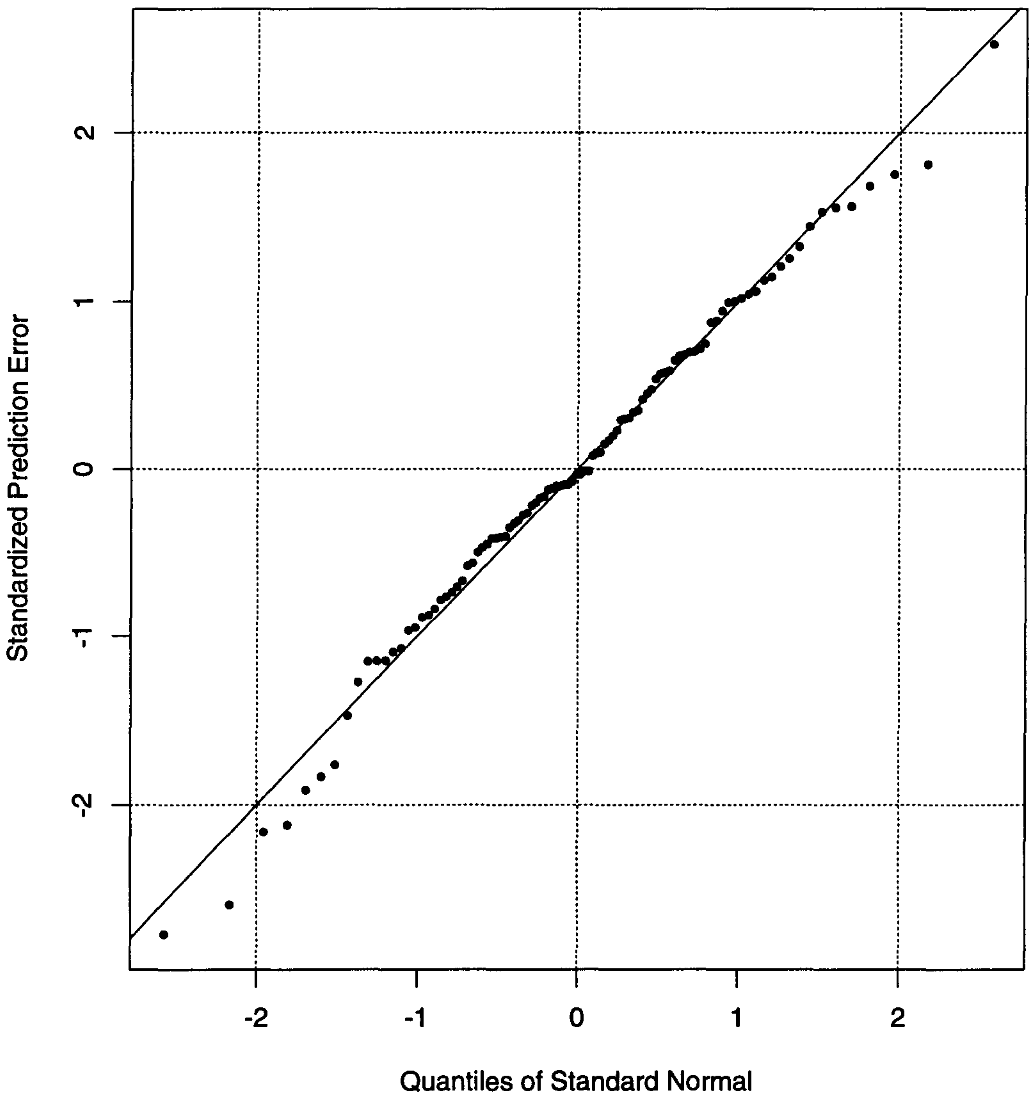


Figure 5.4. Normality plot for the approximated prediction error, σ^2 known. Normality plot for the standardized conditional prediction error when we predict $\int_0^1 Z(t)dt$ given $Z_{(n)} = z_n$ with the approximately optimal predictor, $I_n(Z_{(n)})$.

We also obtained naive 95% prediction intervals. For the nonlinear predictor in the simulation shown in Figure 5.3, conditioning on the particular value of $Z_{(n)} = z_{(n)}$ pictured in Figure 5.1, the 95% prediction interval, using the approximation for the conditional variance (4.3) presented in Section 4, is (0.572687, 0.575567). We calculated the 95% prediction interval for the linear predictor acting as if we have a Brownian motion, and estimated σ^2 with

$$\widetilde{\sigma}^2 = \sum_{i=1}^{n} (Z(t_i) - Z(t_{i-1}))^2 . \tag{5.2}$$

This estimator is uniformly minimum variance unbiased for σ^2 under the assumption that Z is a Brownian motion. Thus, we could use $n^{-2}\widetilde{\sigma}^2$ to build the prediction interval. For this simulation, using the value of $Z_{(n)} = z_{(n)}$ plotted in Figure 5.1, the value of $n^{-2}\widetilde{\sigma}^2$ is 4.52×10^{-7}. The prediction interval for the linear predictor using $n^{-2}\widetilde{\sigma}^2$ is (0.573507, 0.576137). Thus, the interval for the linear predictor is approximately the interval for the nonlinear predictor shifted to the right by around .0006, or by approximately the conditional bias for the linear predictor.

We repeatedly simulated $\int_0^1 Z(t)dt$, conditioning on the same $Z_{(n)}$ as in Figure 5.1, to see how well the obtained prediction intervals work. For this simulation we used the technique presented in part (ii) of this section. In 450 out of 500 simulations (90% of the time), the prediction interval for the nonlinear predictor contains $\int_0^1 Z(t)dt$; whereas, in 390 out 500 simulations (78% of the time), the prediction interval for the linear predictor contains $\int_0^1 Z(t)dt$. The median value of the simulated $\int_0^1 Z(t)dt$ is 0.574027, approximately the mid point of the interval for the nonlinear predictor. All these results suggest that the prediction interval, obtained when we use the nonlinear predictor and the approximate conditional standard prediction error, has somewhat better coverage than the one obtained for the linear predictor (see Figure 5.5).

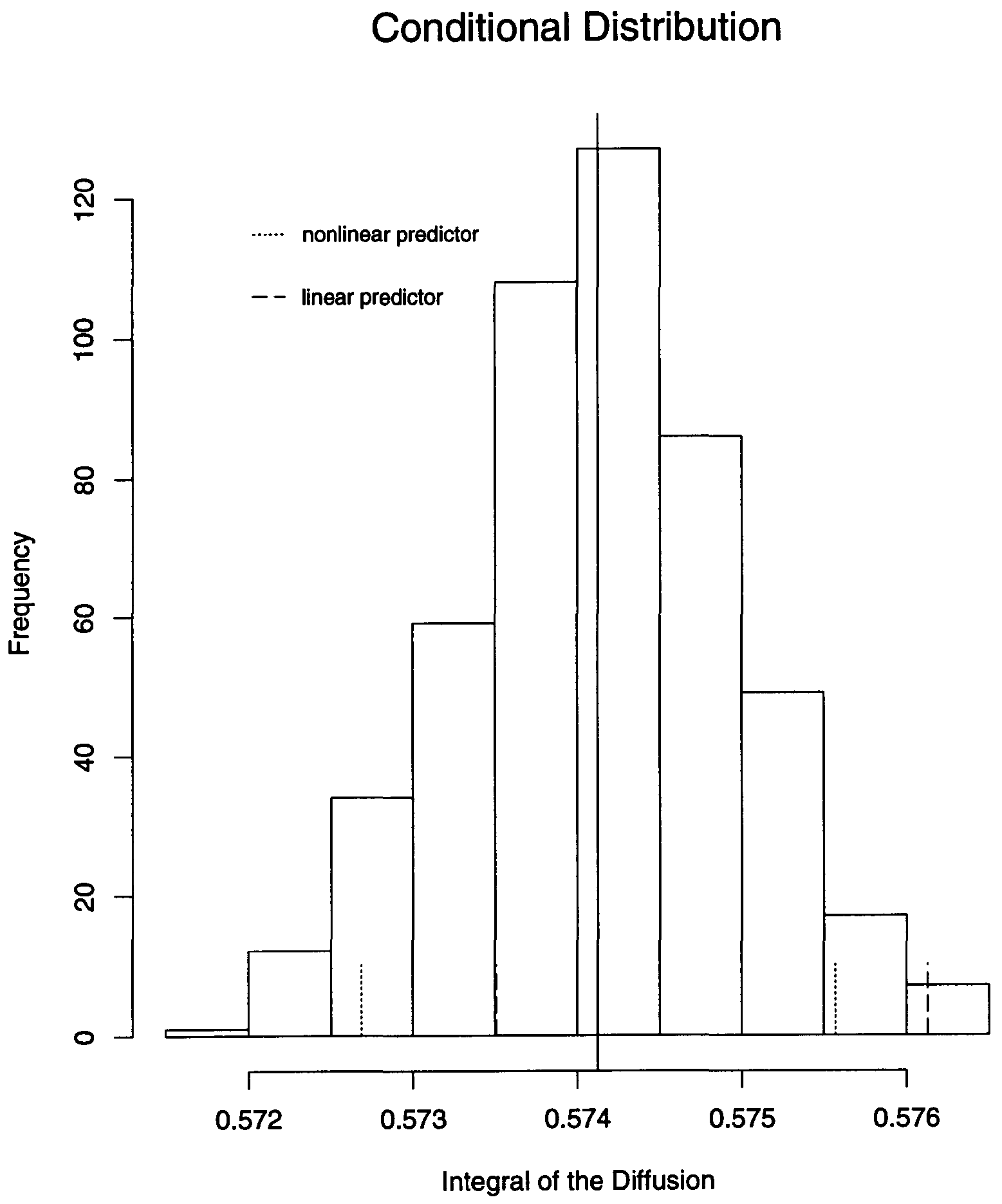


Figure 5.5. Prediction intervals for $\int_0^1 Z(t)dt$. Simulated distribution of the integral $\int_0^1 Z(t)dt$ given a fixed $Z_{(n)} = z_{(n)}$. The solid vertical line is located at the sample mean of the simulated values of $\int_0^1 Z(t)dt$. The marks on the horizontal axis are the prediction intervals for the linear and nonlinear predictors. The vertical axis shows the frequencies out of 500 simulations.

Because the results obtained could be just due to a favorable realization of $Z_{(n)}$, we ran many other simulations. We present the mse of the linear predictor obtained by using two different approaches (using the conditional variance for the linear predictor and using $n^{-2}\tilde{\sigma}^2$) and the mse of the nonlinear predictor using (4.3), for 10 simulated $Z_{(n)}$, with $n = 500$, assuming the diffusion coefficient is a known function.

Conditional mse for the linear and nonlinear predictors:

Linear	**Nonlinear**
1.18	1.04
1.09	0.31
1.86	0.61
2.07	0.52
1.70	1.09
2.23	1.42
2.02	0.04
1.99	0.61
1.98	0.29
1.92	0.79

Table 5.1 *Conditional mse for the linear and nonlinear predictors multiplied by* 10^8.

mse for the best linear predictor using the conditional variance using $n^{-2}\tilde{\sigma}^2$:

Conditional mse	**mse using $n^{-2}\tilde{\sigma}^2$**
1.18	1.14
2.09	2.10
1.86	2.02
2.07	2.14
1.70	1.64
2.23	2.26
2.02	1.97
1.99	2.02
1.98	1.98
1.92	2.03

Table 5.2 *mse values multiplied by* 10^8 *for the linear predictor.*

Table 5.1 shows in the first column the mse for the linear predictor using the conditional variance for the linear predictor. The second column is the mse for the nonlinear predictor using the approximation to the conditional variance presented in Section 4 (4.3), assuming σ^2 is known. The first column have mse values that are larger than the mse values for the nonlinear predictor.

Table 5.2 shows in the first column the mse, conditioning on the same values for $Z_{(n)}$ as in Table 5.1, for the linear predictor using the conditional variance for the linear predictor, and the second column shows the mse for the linear predictor acting as if it is a Brownian motion and using $n^{-2}\tilde{\sigma}^2$. The two columns have similar mse values, which indicates that using $n^{-2}\tilde{\sigma}^2$ to compute the mse for the linear predictor is an appropriate approach.

6. Final Remarks. In this paper we have presented the asymptotic normality of the prediction error for predicting the integral of a diffusion process with the integral of the conditional expectation of the process, the optimal predictor. We have also obtained a simple asymptotically optimal approximation for the predictor and a simple asymptotically valid approximation for the variance of the prediction error, but in both cases we assumed that the diffusion coefficient was a known function.

We could generalize Theorem 2.1 to other additive functionals, for instance $\int_0^1 g(Z_t)dt$, because, by the Ito transformation formula we easily obtain the diffusion parameters of $Y_t = g(Z_t)$ as a function of Z_t through g. If g is smooth enough so that assumptions $(A.4)$ and $(C.2)$ are satisfied by the new diffusion parameters, then Theorem 2.1 holds for the diffusion process $g(Z_t)$, and we obtain the asymptotic normality of $\int_0^1 \{g(Z_t) - E\left[g(Z_t)|g\left(Z_{(n)}\right)\right]\} dt$.

Acknowledgments. The author is grateful to Michael L. Stein for his many valuable comments, insights and guidance in obtaining the results in this paper.

REFERENCES

BILLINGSLEY, P. (1995). *Probability and Measure,* third edition. Wiley, New York.

CHAKRAVARTI, P. C. (1970). *Integrals and Sums: Some New Formulae for Their Numerical Evaluation.* Oxford University Press, New York.

DAVIS, P. J., and RABINOWITH, P. (1984). *Numerical Integration,* second edition. Academic Press, Orlando.

EUBANK, R. L., SMITH, P. L., and SMITH, P. W. (1981). Uniqueness and eventual uniqueness of optimal designs in some time series models. *Annals of Statistics,* **9**, 486–493.

FUENTES, M. (1999). Asymptotic Normality of Conditional Integrals of Diffusion Processes. Tecnical Report Statistics Department North Carolina State University.

GHIZZETTI, A., and OSSCCINI, A. (1970). *Quadrature Formulae.* Academic Press, New York.

KARATZAS, I., and SHREVE S. E. (1991). *Brownian Motion and Stochastic Calculus,* second edition. Springer–Verlag, New York.

MARNEVSKAYA, L. A. and JAROVICH, L. A. (1984). Laplace method for Riemann integrals of random fields. *Vesci Akademii Nauk Byelorussia SSR, Seryja Fizika-Matematycnyh Nauk.* Leninskii prospekt, 68, Minsk, USSR, **4**, 9–12.

MATHERON, G. (1985). Change of support for diffusion–type random functions. *Journal of the International Association for Mathematical Geology,* **17**, 137–165.

PITT, L. D., ROBEVA, R., and WANG, D. Y. (1995). An error analysis for the numerical calculation of certain random integrals: Part 1. *Annals of Applied Probability,* **5**, 171–197.

RITTER, K. (1995). *Average Case Analysis of Numerical Problems.* University of Erlangen.

SACKS, J. and YLVISAKER, D. (1966). Designs for regression problems with correlated error. *Annals of Mathematical Statistics,* **37**, 66–89.

SACKS, J. and YLVISAKER, D. (1968). Designs for regression problems with correlated error: many parameters. *Annals of Mathematical Statistics,* **39**, 46–69.

SACKS, J. and YLVISAKER, D. (1970). Designs for regression problems with correlated error III. *Annals of Mathematical Statistics,* **41**, 2057–2074.

STEIN, M. L. (1987). Gaussian approximations to conditional distributions for multi–Gaussian processes. *Mathematical Geology,* **19**, 387–405.

STEIN, M. L. (1993). Asymptotic properties of centered systematic sampling for predicting integrals of spatial processes. *Annals of Applied Probability,* **3**, 874–880.

STEIN, M. L. (1995). Predicting integrals of stochastic processes. *Annals of Applied Probability,* **5**, 158–170.

WAHBA, G. (1971). On the regression problem of Sacks and Ylvisaker. *Annals of Mathematical Statistics,* **42**, 1035–1053.

WAHBA, G. (1974). Regression design for some equivalence classes of kernels. *Annals of Statistics*, **2**, 925–934.