The Annals of Applied Probability
1995, Vol. 5, No. 1, 158–170

PREDICTING INTEGRALS OF STOCHASTIC PROCESSES[1]

BY MICHAEL L. STEIN

University of Chicago

Consider predicting an integral of a stochastic process based on n observations of the stochastic process. Among all linear predictors, an optimal quadrature rule picks the n observation locations and the weights assigned to them to minimize the mean squared error of the prediction. While optimal quadrature rules are usually unattainable, it is possible to find rules that have good asymptotic properties as $n \to \infty$. Previous work has considered processes whose local behavior is like m-fold integrated Brownian motion for m a nonnegative integer. This paper obtains some asymptotic properties for quadrature rules based on median sampling for processes whose local behavior is not like m-fold integrated Brownian motion for any m.

1. Introduction. For a stochastic process $Z(\cdot)$ on $[0,1]$ with mean 0 and finite second moments, consider linear prediction of $\int_0^1 v(t)Z(t)\,dt$ for a fixed known function $v(\cdot)$ based on observing $Z(t)$ at $0 \le t_1 < \cdots < t_n \le 1$. For fixed n, an optimal quadrature rule minimizes the mean squared prediction error among all choices of $t_1, \dots, t_n$ and of the coefficients of the linear predictor. Except in some simple special cases, it is difficult to find optimal rules, so Sacks and Ylvisaker (1971) developed the notion of asymptotically optimal quadrature rules and found such rules for processes that behave locally like Brownian motion or integrated Brownian motion. Benhenni and Cambanis (1992a, b) extended these results in various directions. Considering the close relationship between predicting integrals and estimating regression coefficients for stochastic processes described by Sacks and Ylvisaker (1971), 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.

All of these works assume that the covariance function $K(s,t) = \text{cov}\big(Z(s), Z(t)\big)$ satisfies for some nonnegative integer m,

$$\alpha_m(t) = \lim_{s \uparrow t} \frac{\partial^{2m+1}}{\partial t^m\, \partial s^{m+1}} K(t,s) - \lim_{s \downarrow t} \frac{\partial^{2m+1}}{\partial t^m\, \partial s^{m+1}} K(t,s) \tag{1.1}$$

Received August 1993; revised April 1994.

[1]This research was supported in part by NSF Grant DMS-92-04504. This manuscript was prepared using computer facilities supported in part by NSF Grants DMS-89-052092, DMS-87-03942 and DMS-86-01732 awarded to the Department of Statistics at the University of Chicago, and by the University of Chicago Block Fund.

AMS 1991 *subject classifications*. Primary 62M20; secondary 41A55.

Key words and phrases. Optimal quadrature, design of time series experiments, median sampling, Riemann zeta function, fractional Brownian motion.

is finite and positive on $[0,1]$. For $Z(\cdot)$ m-fold integrated Brownian motion, (1.1) is satisfied with $\alpha_m(t)$ constant. More generally, since the smoothness of $K(s,t)$ for $s-t$ near 0 controls the local mean square behavior of $Z(\cdot)$, any process with covariance function satisfying (1.1) can be said to behave locally like m-fold integrated Brownian motion, at least in the mean square sense. As an example of when (1.1) does not hold, consider $K(s,t) = \exp\{-\theta|s-t|^\gamma\}$, which is a covariance function for $0 < \gamma \le 2$, $\theta > 0$, and is used by Sacks, Welch, Mitchell and Wynn (1989) to model the response surface from a computer experiment. For this K, (1.1) only holds for $\gamma = 1$, in which case $m = 1$ and $\alpha_1(t) \equiv 2\theta$. If $\gamma < 1$, then $\alpha_1(t) = \infty$ and if $1 < \gamma < 2$, $\alpha_1(t) \equiv 0$, but $\alpha_2(t) \equiv \infty$. When $\gamma = 2$, $\alpha_m(t) \equiv 0$ for all m. The results obtained here apply, for example, when $0 < \gamma < 2$; they do not apply when $\gamma = 2$.

To define the optimal quadrature problem formally, let $T_n = \{t_1, \dots, t_n\}$, with, for convenience, $0 \le t_1 < \cdots < t_n \le 1$ and $C_n = (c_1, \dots, c_n) \in \mathbb{R}^n$. Define for a function J from $[0,1]^2$ to $\mathbb{R}$,

$$e_J(T_n, C_n) = \int_0^1\int_0^1 v(s)v(t)J(s,t)\,dt\,ds - 2\sum_{i=1}^n c_i \int_0^1 v(t)J(t,t_i)\,dt + \sum_{i,j=1}^n c_i c_j J(t_i,t_j),$$

so that

$$e_K(T_n, C_n) = \operatorname{var}\left(\int_0^1 v(t)Z(t)\,dt - \sum_{i=1}^n c_i Z(t_i)\right).$$

Thus, the pair (T_n, C_n) defines a quadrature rule. Let D_n be the class of all n element subsets of $[0,1]$. The rule (T_n^*, C_n^*) is optimal if

$$e_K(T_n^*, C_n^*) = \inf_{\substack{T_n \in D_n,\\ C_n \in \mathbb{R}^n}} e_K(T_n, C_n).$$

A sequence of rules $\{(T_n^*, C_n^*)\}$ is asymptotically optimal if

$$\lim_{n\to\infty} \frac{e_K(T_n^*, C_n^*)}{\inf_{\substack{T_n \in D_n,\\ C_n \in \mathbb{R}^n}} e_K(T_n, C_n)} = 1.$$

Asymptotically optimal rules can sometimes be based on what Cambanis (1985) calls median sampling. Specifically, for a continuous density h on $[0,1]$, define $T_n(h) = \{t_{in}, \dots, t_{nn}\}$ by

$$\int_0^{t_{in}} h(t)\,dt = \frac{(i-(1/2))}{n}.$$

Sacks and Ylvisaker (1971) essentially show that for $m = 0$ or 1 in (1.1) with $\alpha_m(t)$ constant, by taking $h(t)$ proportional to $v(t)^{2/(2m+3)}$ and appropriately defining $C_n(h)$, $\{(T_n(h), C_n(h))\}$ is an asymptotically optimal sequence of quadrature rules.

Sacks and Ylvisaker are able to obtain this result because it is possible to put a sharp lower bound on the mean squared error of optimal quadrature rules when $m = 0$ or 1. Eubank, Smith and Smith (1981) extend this result to arbitrary m, but only for a very restricted class of covariance functions. For more general covariance functions satisfying (1.1) with α_m constant, the following is known [Benhenni and Cambanis (1992a)]: let $\mathscr{H}$ be the class of positive continuous densities on $[0,1]$ and $C_n(h)$ a sequence of weights that appropriately corrects for edge effects. Then if $h^*(t)$ is proportional to $v(t)^{2/(2m+3)}$,

$$(1.2) \qquad \lim_{n\to\infty} \frac{e_K(T_n(h^*),\, C_n(h^*))}{\inf_{h\in\mathscr{H}} e_K(T_n(h), C_n(h))} = 1.$$

Benhenni and Cambanis (1992a) also consider cases where α_m is not constant. Note that (1.2) falls short of proving the asymptotic optimality of $\{(T_n(h^*), C_n(h^*)\}$ in two regards. First, it is not known in general that the weights $C_n(h^*)$ as defined by Benhenni and Cambanis (1992a) do as well asymptotically as the optimal weights. Second, even if these weights are asymptotically optimal given the sequence of designs $\{T_n(h^*)\}$, the possibility that no regular sequence yields an asymptotically optimal quadrature rule has not been eliminated in general.

The goal here is, in a sense, to extend (1.2) to cases where m is not an integer. Specifically, letting $K(t) = \operatorname{cov}(Z(0), Z(t))$ for stationary $Z(\cdot)$, assume for some $a \neq 0$, $K(t) - a|t|^\alpha$ is smoother than $K(t)$ in a sense made precise in Section 2. Then the results of Section 2 show that (1.2) holds for $\alpha < 3$ but $\alpha \neq 2$. Making the identification $\alpha = 2m + 1$, we see that previous results only hold for α an odd integer. I expect that the results here can be extended to include $\alpha \geq 3$ (but not an even integer), but that would require a more sophisticated choice of weights [Benhenni and Cambanis (1992a)].

While this work only considers stationary processes, the extension to certain nonstationary processes is not difficult. For example, suppose $Z(\cdot)$ is a fractional Brownian motion [Yaglom (1987), page 406], which is a class of self-affine processes that has found application to fields such as music, geography and meteorology [Voss (1989)]. Then $\operatorname{var}(Z(s) - Z(t)) = 2a|s-t|^\alpha$ for some $\alpha \in (0,2)$, so that $\alpha = 1$ yields Brownian motion. It is not possible to make $\operatorname{var}(Z(t))$ independent of t for such a process, so the process is not stationary, although it does possess stationary increments. Theorem 1 below applies to fractional Brownian motion if we are willing to assume $\operatorname{var}(Z(0))$ is finite. Alternatively, if we normalize the c_i's to add up to 1, then again Theorem 1 applies without having to assume $\operatorname{var}(Z(0))$ is finite. One rough interpretation of Theorem 1 when $0 < \alpha < 2$ is that for processes whose local behavior is sufficiently similar to that of fractional Brownian motion, the problem of predicting integrals of the process is asymptotically the same as if the process were fractional Brownian motion.

As noted by Sacks and Ylvisaker (1971), there is a close relationship between the stochastic approach to optimal quadrature and the deterministic worst-case analysis commonly used in the numerical analysis literature. In Section 3, I compare the results here to some related work in numerical analysis.

Finally, in work done independently from mine, Pitt, Robeva and Wang (1995) obtain essentially all of the results in this work using rather different methods of proof.

2. Results. Suppose h is a positive density on $[0,1]$, H the corresponding distribution, $G = H^{-1}$ on $[0,1]$ and $g = G'$. Then $T_n(h) = \{G((j - 1/2)/n);\ j = 1,\dots,n\}$. Defining $u(t) = v(G(t))g(t)$, a reasonable predictor of $\int_0^1 v(t)Z(t)\,dt = \int_0^1 v(G(t))g(t)Z(G(t))\,dt$ based on observing $Z(\cdot)$ on $T_n(h)$ is $n^{-1}\sum_{j=1}^n u((j-1/2)/n)Z(G((j-1/2)/n))$. So define $C_n(h) = (c_{1n},\dots,c_{nn})$ by $c_{jn} = (1/n)u((j-1/2)/n)$.

THEOREM 1. *Suppose v and G, considered as functions on $[0,1]$, have two bounded derivatives with $g(t) > 0$ for $0 \le t \le 1$ and that for some $a > 0$, either*

$$K(t) = a_0 - a|t|^\alpha + Q(t) \quad \text{for } 0 < \alpha < 2$$

or

$$K(t) = a_0 - a_1 t^2 + a|t|^\alpha + Q(t) \quad \text{for } 2 < \alpha < 3,$$

where $Q(t)$, considered as a function on $[0,1]$, has $\lfloor\alpha\rfloor + 2$ derivatives,

$$Q(t) = o(t^\alpha) \quad \text{as } t \downarrow 0, \tag{2.1}$$

$$\frac{Q^{(\lfloor\alpha\rfloor+2)}(t)}{t^{\alpha-\lfloor\alpha\rfloor-2}} \text{ is bounded on } [0,1] \tag{2.2}$$

and

$$Q^{(\lfloor\alpha\rfloor+2)}(t) = o(t^{\alpha-\lfloor\alpha\rfloor-2}) \quad \text{as } t \downarrow 0. \tag{2.3}$$

Then as $n \to \infty$,

$$n^{\alpha+1}e_K(T_n(h), C_n(h)) \to 2a|\zeta(-\alpha)| \int_0^1 \frac{v(t)^2}{h(t)^{\alpha+1}}\,dt, \tag{2.4}$$

where $\zeta(\cdot)$ is the Riemann zeta function.

Before proving this result, first note that it does agree with Theorem 2 of Benhenni and Cambanis (1992a) when they both apply. If $\alpha = 2m+1$, then $\alpha_m(t)$ as defined in (1.1) is identically $(2m+1)!a$, so that their result implies

$$n^{2m+2}e_K(T_n(h), C_n(h), n) \to \frac{|B_{2m+2}|}{2m+2}a\int_0^1 \frac{v(t)^2}{h(t)^{2m+2}}\,dt,$$

where B_{2m+2} is a Bernoulli number [Abramowitz and Stegun (1965), page 807]. The two results coincide with $\alpha = 2m+1$ by noting $\zeta(-2m-1) = -B_{2m+2}/(2m+2)$. This correspondence provides support to the conjecture that Theorem 1 applies with $\alpha \geq 3$ as long as the definition of $C_n(h)$ is appropriately modified.

PROOF OF THEOREM 1. Define

$$e_J^{jk}(n) = \int_j \int_k \{u(t)u(s)\, J(G(t),G(s)) - u(t)u_k\, J(G(t),G_k) - u_j u(s)\, J(G_j,G(s)) + u_j u_k\, J(G_j,G_k)\}\, ds\, dt,$$

where $\int_j$ means $\int_{(j-1)/n}^{j/n}$, $G_j = G((j-1/2)/n)$ and $u_j = u((j-1/2)/n)$, so that

$$e_J(n) \overset{\text{def}}{=} e_J(T_n(h), C_n(h)) = \sum_{j,k=1}^{n} e_J^{jk}(n).$$

Letting $J_b(x,y) = |x-y|^b$, we then have, for $0 < \alpha < 2$, $e_K(n) = a_0 e_{J_0}(n) - a e_{J_\alpha}(n) + e_Q(n)$. The basic idea of the proof is to show that the term $e_{J_\alpha}(n)$ dominates the mean squared error asymptotically. By straightforward calculation, both $e_{J_0}(n)$ and $e_{J_2}(n)$ are $O(n^{-4})$, and so are $o(n^{-\alpha-1})$ for $\alpha < 3$. Next consider $e_Q(n)$. By (2.1),

$$\sum_{|j-k|\leq 1} e_Q^{jk}(n) = o(n^{-\alpha-1}). \tag{2.5}$$

Define

$$\mathrm{I}_J^{jk} = \int_j \int_k (u(t)-u_j)(u(s)-u_k)\, J(G(t),G(s))\, ds\, dt,$$

$$\mathrm{II}_J^{jk} = \int_j \int_k (u(s)-u_k)\{J(G(t),G(s)) - J(G_j,G(s))\}\, ds\, dt \quad \text{and}$$

$$\mathrm{III}_J^{jk} = \int_j \int_k \{J(G(t),G(s)) - J(G(t),G_u) - J(G_j,G(s)) + J(G_j,G_k)\}\, ds\, dt,$$

suppressing the dependence on n. Then

$$e_J^{jk}(n) = \mathrm{I}_J^{jk} + u_j \mathrm{II}_J^{jk} + u_k \mathrm{II}_J^{kj} + u_j u_k \mathrm{III}_J^{jk}.$$

In bounding these terms, we will need the following: by (2.2) and (2.3),

$$\begin{aligned}
\left|Q^{(\lfloor\alpha\rfloor+1)}(t)\right| &\leq \left|Q^{(\lfloor\alpha\rfloor+1)}(1)\right| + \left|\int_t^1 Q^{(\lfloor\alpha\rfloor+2)}(s)\, ds\right| \\
&= O(1) + o\left(\int_t^1 s^{\alpha-\lfloor\alpha\rfloor-2}\, ds\right) \\
&= o(t^{\alpha-\lfloor\alpha\rfloor-1}),
\end{aligned}$$

so that $t^{-\alpha+\lfloor\alpha\rfloor+1} Q^{(\lfloor\alpha\rfloor+1)}(t)$ is bounded on $[0,1]$ and tends to 0 as $t \downarrow 0$. Repeating this argument yields $t^{-\alpha+\lfloor\alpha\rfloor} Q^{(\lfloor\alpha\rfloor)}(t)$ is bounded on $[0,1]$ and tends to 0 as $t \downarrow 0$ unless $\alpha = 1$. For $\alpha = 1$, $Q'(t) \to 0$ as $t \downarrow 0$ by (2.1). For $2 < \alpha < 3$, $Q'(t)/t^{\alpha-1}$ is bounded and tends to 0 as $t \downarrow 0$ by (2.1). In all cases, for $1 \le r \le \lfloor \alpha \rfloor + 2$, $Q^{(r)}(t)/t^{\alpha - r}$ is bounded and tends to 0 as $t \downarrow 0$. Defining for $j > k$,

$$B_{jk}^{(r)} = \sup\{|Q^{(r)}(t)|: G_{j-1} - G_k \le t \le G_j - G_{k-1}\},$$

for $1 \le r \le \lfloor \alpha \rfloor + 2$ and $j > k + 1$, we thus have

$$\frac{B_{jk}^{(r)}}{((j-k)/n)^{\alpha - r}} \text{ is bounded} \tag{2.6}$$

and

$$\lim_{n\to\infty} \sup_{1 < j - k < \gamma(n)} \frac{B_{jk}^{(r)}}{((j-k)/n)^{\alpha - r}} = 0 \tag{2.7}$$

if $\gamma(n)/n \to 0$ as $n \to \infty$. Finally, the following bound will be repeatedly used: for a function f with second derivative bounded by c in a neighborhood of the origin,

$$\left| \int_{-1/(2n)}^{1/(2n)} \{f(t) - f(0)\}\, dt \right| \le \frac{c}{4n^3} \tag{2.8}$$

for all n sufficiently large.

Now consider I_Q^{jk}. For $j > k + 1$,

$$\begin{aligned}
|\mathrm{I}_Q^{jk}| \le & \left| \int_j \int_k (u(t) - u_j)(u(s) - u_k) \right. \\
& \left. \times \{Q(G_j - G_k) + (G(t) - G_j - G(s) + G_k) Q'(G_j - G_k)\}\, ds\, dt \right| \\
& + \tfrac{1}{2} B_{jk}^{(2)} \int_j \int_k |u(t) - u_j|\, |u(s) - u_k| (G(t) - G_j - G(s) + G_k)^2\, ds\, dt \\
= & \; O(n^{-6}) + O\left(n^{-6}\left(\frac{j-k}{n}\right)^{\alpha - 1}\right) + O\left(n^{-6}\left(\frac{j-k}{n}\right)^{\alpha - 2}\right)
\end{aligned}$$

using (2.6), (2.8) and the assumption that u has bounded second derivative on $[0, 1]$. Moreover,

$$\sum_{j-k>1} (j-k)^\beta \le n \sum_{j=1}^{n} n^\beta = O(n + n^{\beta + 2})$$

and $\mathrm{I}_Q^{jk} = \mathrm{I}_Q^{kj}$, so

$$\sum_{|j-k|>1} |\mathrm{I}_Q^{jk}| = O(n^{-4} + n^{-3-\alpha}) = o(n^{-\alpha - 1})$$

for $\alpha < 3$. Equation (2.5) allows us to handle the terms $|j - k| \le 1$.

Next consider II_Q^{jk}. For $\alpha < 2$ and $j > k+1$,

$$\begin{aligned}|\mathrm{II}_Q^{jk}| &\leq \left|Q'(G_j - G_k)\int_j\int_k (u(s)-u_k)(G(t)-G_j)\,ds\,dt\right| \\ &\quad + B_{jk}^{(2)}\int_j\int_k |u(s)-u_k|\left\{(G(t)-G_j)^2 + 2|G(t)-G_j||G(s)-G_k|\right\} ds\,dt \\ &= O\left(n^{-6}\left(\frac{j-k}{n}\right)^{\alpha-1}\right) + O\left(n^{-5}\left(\frac{j-k}{n}\right)^{\alpha-2}\right),\end{aligned}$$

yielding

$$\sum_{|j-k|>1} |u_j \mathrm{II}_Q^{jk}| = O(n^{-2-\alpha} + n^{-3}) = o(n^{-\alpha-1})$$

for $\alpha < 2$. For $2 < \alpha < 3$ and $j > k+1$,

$$\begin{aligned}|\mathrm{II}_Q^{jk}| &\leq \left|Q'(G_j - G_k)\int_j\int_k (u(s)-u_k)(G(t)-G_j)\,ds\,dt\right| \\ &\quad + \frac{1}{2}\left|Q''(G_j - G_k)\int_j\int_k (u(s)-u_k)\right. \\ &\qquad \times \left.\left\{(G(t)-G_j)^2 - 2(G(t)-G_j)(G(s)-G_k)\right\} ds\,dt\right| + B_{jk}^{(3)}O(n^{-6}) \\ &= O(n^{-6}) + O\left(n^{-6}\left(\frac{j-k}{n}\right)^{\alpha-3}\right),\end{aligned}$$

yielding

$$\sum_{|j-k|>1} |u_j \mathrm{II}_Q^{jk}| = O(n^{-4}) = o(n^{-\alpha-1}).$$

To bound III_Q^{jk}, first consider $\alpha < 1$ and $j > k+1$: $|\mathrm{III}_Q^{jk}| = O(B_{jk}^{(2)} n^{-4})$, so that

$$\sum_{|j-k|>1} |u_j u_k \mathrm{III}_Q^{jk}| = o(n^{-1-\alpha}) \tag{2.9}$$

follows from (2.7). For $1 \leq \alpha < 2$, (2.9) follows from (2.7) and

$$|\mathrm{III}_Q^{jk}| = \left|Q''(G_j - G_k)\int_j\int_k (G(t)-G_j)(G(s)-G_k)\,ds\,dt\right| + O(B_{jk}^{(3)} n^{-5}),$$

and for $2 < \alpha < 3$, (2.9) follows from

$$
\begin{aligned}
|\mathrm{III}_Q^{jk}| &= O(n^{-6}) \\
&\quad + \left| \frac{1}{6} Q^{(3)}(G_j - G_k) \int_j \int_k \{(G(t) - G_j)(G(s) - G_k)^2 \right. \\
&\quad \left. - (G(t) - G_j)^2 (G(s) - G_k)\} \, ds \, dt \right| + O(B_{jk}^{(4)} n^{-6}) \qquad (2.10) \\
&= O\left(n^{-6}\left(1 + \left(\frac{j-k}{n}\right)^{\alpha-3} + \left(\frac{j-k}{n}\right)^{\alpha-4}\right)\right).
\end{aligned}
$$

Finally, consider $e_{J_\alpha}(n)$. By following the preceding arguments, we have

$$\sum_{|j-k|>1} \{\mathrm{I}_{J_\alpha}^{jk} + u_j \mathrm{II}_{J_\alpha}^{jk} + u_k \mathrm{II}_{J_\alpha}^{kj}\} = o(n^{-1-\alpha}),$$

since this bound only requires (2.6), which is true for $Q(t) = |t|^\alpha$. Moreover, a straightforward argument shows

$$\sum_{|j-k|\le 1} \{\mathrm{I}_{J_\alpha}^{jk} + u_j \mathrm{II}_{J_\alpha}^{jk} + u_k \mathrm{II}_{J_\alpha}^{kj}\} = O(n^{-2-\alpha}).$$

Thus,

$$\sum_{j,k=1}^{n} \{\mathrm{I}_{J_\alpha}^{jk} + u_j \mathrm{II}_{J_\alpha}^{jk} + u_k \mathrm{II}_{J_\alpha}^{kj}\} = o(n^{-1-\alpha}).$$

For any function $\gamma(n)$ satisfying $\gamma(n) \to \infty$ and $\gamma(n)/n \to 0$ as $n \to \infty$,

$$
\begin{aligned}
&\sum_{|j-k|\le\gamma(n)} u_j u_k \mathrm{III}_{J_\alpha}^{jk} \\
&= \sum_{|j-k|\le\gamma(n)} u_j \left\{ u_j + O\left(\frac{\gamma(n)}{n}\right)\right\} \\
&\quad \times \int_j \int_k \left\{ \left| (t-s) g_j + O\left(\left(\frac{|j-k|+1}{n}\right)^2\right)\right|^\alpha \right. \\
&\qquad - \left| \left(t - \frac{k-1/2}{n}\right) g_j + O\left(\left(\frac{|j-k|+1}{n}\right)^2\right)\right|^\alpha \\
&\qquad - \left| \left(\frac{j-1/2}{n} - s\right) g_j + O\left(\left(\frac{|j-k|+1}{n}\right)^2\right)\right|^\alpha \\
&\qquad \left. + \left| \left(\frac{j-k}{n}\right) g_j + O\left(\left(\frac{|j-k|+1}{n}\right)^2\right)\right|^\alpha \right\} ds\, dt \\
&= \sum_{|j-k|\le\gamma(n)} \left\{ u_j^2 + O\left(\frac{\gamma(n)}{n}\right)\right\} g_j^\alpha
\end{aligned}
$$

$$\times \int_j \int_k \left\{ |t-s|^\alpha - \left| t - \frac{k-1/2}{n} \right|^\alpha - \left| \frac{j-1/2}{n} - s \right|^\alpha + \left| \frac{j-k}{n} \right|^\alpha + O\left(\left(\frac{|j-k|+1}{n} \right)^{1+\alpha} \right) \right\} ds\, dt$$

$$= \sum_{|j-k| \le \gamma(n)}' u_j^2 g_j^\alpha \int_j \int_k \left\{ |t-s|^\alpha - \left| t - \frac{k-1/2}{n} \right|^\alpha - \left| \frac{j-1/2}{n} - s \right|^\alpha + \left| \frac{j-k}{n} \right|^\alpha \right\} ds\, dt + O\left(\left(\frac{\gamma(n)}{n} \right)^{\alpha+2} \right),$$

so that the remainder is $o(n^{-\alpha-1})$ if, for example, $\gamma(n) = \log n$. Now

$$n^{\alpha+2} \int_j \int_k \left\{ |t-s|^\alpha - \left| t - \frac{k-1/2}{n} \right|^\alpha - \left| \frac{j-1/2}{n} - s \right|^\alpha + \left| \frac{j-k}{n} \right|^\alpha \right\} ds\, dt$$
$$= \frac{1}{(\alpha+1)(\alpha+2)} \left\{ |j-k+1|^{\alpha+2} - 2|j-k|^{\alpha+2} + |j-k-1|^{\alpha+2} \right\} - \frac{2}{\alpha+1} \left\{ \left| j-k+\frac{1}{2} \right|^{\alpha+1} - \left| j-k-\frac{1}{2} \right|^{\alpha+1} \right\} + |j-k|^\alpha$$

for $j \ne k$ and is

$$\frac{2}{(\alpha+1)(\alpha+2)} - \frac{1}{(\alpha+1)2^{\alpha-1}}$$

for $j = k$. Thus, for a positive integer r,

$$n^{\alpha+2} \sum_{k=j-r}^{j+r} \int_j \int_k \left\{ |t-s|^\alpha - \left| t - \frac{k-1/2}{n} \right|^\alpha - \left| \frac{j-1/2}{n} - s \right|^\alpha + \left| \frac{j-k}{n} \right|^\alpha \right\} ds\, dt$$
$$= \frac{2}{(\alpha+1)(\alpha+2)} \left\{ (r+1)^{\alpha+2} - r^{\alpha+2} \right\} - \frac{4}{\alpha+1} \left(r + \frac{1}{2} \right)^{\alpha+1} + 2 \sum_{l=1}^{r} l^\alpha.$$

For $0 < \alpha < 3$, as $r \to \infty$ [Berndt (1985), page 150],

$$\sum_{l=1}^{r} l^\alpha = \zeta(-\alpha) + \frac{r^{\alpha+1}}{\alpha+1} + \frac{r^\alpha}{2} + \frac{\alpha r^{\alpha-1}}{12} + o(1),$$

so that

$$\lim_{r\to\infty} \left[\frac{2}{(\alpha+1)(\alpha+2)} \left\{ (r+1)^{\alpha+2} - r^{\alpha+2} \right\} - \frac{4}{\alpha+1} \left(r + \frac{1}{2} \right)^{\alpha+1} + 2 \sum_{l=1}^{r} l^\alpha \right] = 2\zeta(-\alpha).$$

Using $\gamma(n) \to \infty$, a straightforward argument yields

$$\sum_{|j-k|\le\gamma(n)} u_j u_k \mathrm{III}_{J_\alpha}^{jk} \sim \frac{2\zeta(-\alpha)}{n^{\alpha+2}} \sum_{j=1}^{n} u_j^2 g_j^\alpha \sim \frac{2\zeta(-\alpha)}{n^{\alpha+1}} \int_0^1 u(t)^2 g(t)^\alpha \, dt$$
$$= \frac{2\zeta(-\alpha)}{n^{\alpha+1}} \int_0^1 \frac{v(t)^2}{h(t)^{\alpha+1}} \, dt.$$

For $j-k > \gamma(n)$, similar to (2.10),

$$|\mathrm{III}_{J_\alpha}^{jk}| = O\left(\left(\frac{j-k}{n}\right)^{\alpha-4} n^{-6}\right),$$

so that for $\alpha < 3$,

$$\sum_{|j-k|\le\gamma(n)} |u_j u_k \mathrm{III}_{J_\alpha}^{jk}| = O(n^{-\alpha-1}) \sum_{j=\gamma(n)}^{n} j^{\alpha-4} = o(n^{-\alpha-1}).$$

Theorem 1 follows. □

Using Hölder's inequality, it can be shown that $\int_0^1 v(t)^2 h(t)^{-\alpha-1}\,dt$ subject to h being a density on $[0,1]$ is minimized by taking $h(t)$ proportional to $v(t)^{2/(\alpha+2)}$. It follows that the unique continuous density minimizing the right-hand side of (2.4) is

$$h^*(t) = v(t)^{2/(\alpha+2)} \Big/ \int_0^1 v(s)^{2/(\alpha+2)}\, ds \tag{2.11}$$

and in this case,

$$\int_0^1 \frac{v(t)^2}{h^*(t)^{\alpha+1}}\, dt = \left\{ \int_0^1 v(t)^{2/(\alpha+2)}\, dt \right\}^{\alpha+2}. \tag{2.12}$$

For h^* as given in (2.11) to satisfy the conditions of Theorem 1, we must have $v(t) \neq 0$ on $[0,1]$. Since v is continuous, we may as well assume $v(t) > 0$ on $[0,1]$. Let $\mathscr{H}'$ be the class of all positive continuous densities on $[0,1]$ with bounded first derivatives on (0,1).

THEOREM 2. *Suppose v and K satisfy the conditions of Theorem 1 and $v(t) > 0$ on $[0,1]$. Then*

$$\lim_{n\to\infty} \frac{e_K(T_n(h^*), C_n(h^*))}{\inf_{h\in\mathscr{H}'} e_K(T_n(h), C_n(h))} = 1.$$

When $\alpha = 1$, so that both (1.2) and Theorem 2 apply, the same conclusion is reached except for the additional restriction here that densities in $\mathscr{H}'$ have bounded first derivative.

3. Discussion. There is an enormous literature on error analysis for quadrature formulae based on a worst-case approach. The basic idea [Motornyi (1990)] is for some class of functions $\mathscr{F}$, to define the largest error of a quadrature rule

$$R(\mathscr{F}, T, C) = \sup_{f\in\mathscr{F}}\left|\int_0^1 f(t)\,dt - \sum_{i=1}^{n} c_i f(t_i)\right|.$$

If, as is commonly done, $\mathscr{F}$ is all functions in some Sobolev space whose norm is at most 1, the worst-case approach and the probabilistic approach taken here are closely linked. For example, suppose $\mathscr{F}$ is all functions for which $f^{(m)}$ is absolutely continuous and $\left\{\int_0^1 f^{(m+1)}(t)^2\,dt\right\}^{1/2} \leq 1$. Then Sacks and Ylvisaker (1971) note that $R(\mathscr{F}, T, C) = e_K(T, C)$ if K is the covariance function for m-fold integrated Brownian motion. Another way to characterize this correspondence is that K is the reproducing kernel [Wahba (1990)] for the Hilbert space with norm given by the L^2 norm of $f^{(m+1)}$. Sacks and Ylvisaker (1970) and Eubank, Smith and Smith (1981), for example, exploit this correspondence to obtain their optimality results. While a covariance function satisfying the conditions of Theorem 1 is the reproducing kernel for some Hilbert space, unless $\alpha = 1$, it is not possible to express the norm on this space in terms of integrals of ordinary derivatives. Thus, it appears one would need to use Sobolev spaces of fractional order [Adams (1975)] to have any hope of handling covariance functions satisfying (2.1) for $\alpha \neq 1$.

A deterministic approach that yields similar looking results to those obtained here is to take $\mathscr{F}$ to be the class of all functions f that can be written as $f(t) = v(t)z(t)$ for a fixed known v satisfying some regularity conditions and $|z(s) - z(t)| \leq |s-t|^\beta$ for all s, t and some β satisfying $0 < \beta \leq 1$ [Motornyi (1990)]. Then [Motornyi (1990)]

$$\text{(2.13)}\qquad \lim_{n\to\infty} n^{2\beta} \inf_{\substack{T\in D_n,\\ C\in\mathbb{R}^n}} R(\mathscr{F}, T, C)^2 = \frac{1}{(\beta+1)^2 2^{2\beta}}\left(\int_0^1 v(t)^{1/(\beta+1)}\,dt\right)^{2(\beta+1)}.$$

Setting $\beta = \alpha/2$, we see the formal similarity between the right-hand sides of (2.13) and (2.12). However, it does not appear possible to apply (2.13) to the stochastic problem addressed here because of the lack of a direct relationship between taking suprema over $\mathscr{F}$ and the mean squared error under K as given in Theorem 1. In particular, the Riemann zeta function plays no role in (2.13).

However, the appearance of the Riemann zeta function in Theorem 1 is not entirely unexpected. Bucklew and Cambanis (1988) showed that for a class of processes covering the case $\alpha = \frac{1}{2}$ in Theorem 1, the Riemann zeta function does appear in the constant for the limiting variance when predicting integrals. Moreover, Sobolev (1974, 1992) showed how a multivariate generalization of the Riemann zeta function known as the Epstein zeta function occurs in the worst case error for integrating periodic functions on the unit cube based on observations on a regular lattice. Stein (1993a, b) gives somewhat analogous results for stationary processes on the unit cube based on observations on a square lattice. In particular, if $h(t) \equiv 1$, then Theorem 1 here is essentially a very special case of Proposition 3.1 of Stein (1993b), which can be seen by using the identity $\zeta(-\alpha) = -2(2\pi)^{-\alpha-1}\Gamma(\alpha+1)\zeta(\alpha+1)\sin(\pi\alpha/2)$.

Finally, it is worth noting that there is a considerable literature on probabilistic approaches to error analysis of quadrature formulae in the computational complexity literature [Traub, Wasilkowski and Woźniakowski (1988)], although there it is generally assumed that the stochastic process is m-fold integrated Brownian motion. To the extent that fractional Brownian motion and models with similar local behavior have found application, restricting attention to processes that behave locally like m-fold integrated Brownian motion appears to be a matter of mathematical convenience rather than any determination that other models are not useful. The results here and in Pitt, Robeva and Wang (1995) show that it is possible to obtain useful results about predicting integrals for processes that behave locally like fractional Brownian motion.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. (1965). *Handbook of Mathematical Functions*. Dover, New York.

ADAMS, R. A. (1975). *Sobolev Spaces*. Academic Press, New York.

BENHENNI, K. and CAMBANIS, S. (1992a). Sampling designs for estimating integrals of stochastic processes. *Ann. Statist.* **20** 161–194.

BENHENNI, K. and CAMBANIS, S. (1992b). Sampling designs for estimating integrals of stochastic processes using quadratic mean derivatives. In *Approximation Theory: Proceedings of the Sixth Southeastern Approximation Theorists Annual Conference* (G. A. Anastassiou, ed.) 93–123. Dekker, New York.

BERNDT, B. C. (1985). *Ramanujan's Notebooks, Part I.* Springer, New York.

BUCKLEW, J. A. and CAMBANIS, S. (1988). Estimating random integrals from noisy observations: sampling designs and their performance. *IEEE Trans. Inform. Theory* **34** 111–127.

CAMBANIS, S. (1985). Sampling designs for time series. In *Time Series in the Time Domain. Handbook of Statistics* (E. J. Hannan, P. R. Krishnaiah and M. M. Rao, eds.) **5** 337–362. North-Holland, Amsterdam.

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

MOTORNYI, V. P. (1990). Investigation of the optimization of quadrature formulas by Dnepropetrovsk mathematicians. *Ukrainian Math. J.* **42** 13–27.

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

SACKS, J. and YLVISAKER, D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist.* **37** 66–89.

SACKS, J. and YLVISAKER, D. (1968). Designs for regression problems with correlated errors: many parameters. *Ann. Math. Statist.* **39** 49–69.

SACKS, J. and YLVISAKER, D. (1970). Designs for regression problems with correlated errors III. *Ann. Math. Statist.* **41** 2057–2074.

SACKS, J. and YLVISAKER, D. (1971). Statistical designs and integral approximation. In *Proceedings of the Twelfth Biennial Canadian Mathematical Society Seminar* (R. Pyke, ed.) 115–136. Canadian Mathematical Congress, Montreal.

SACKS, J., WELCH, W. J., MITCHELL, T. J. and WYNN, H. P. (1989). Design and analysis of computer experiments (with discussion). *Statist. Sci.* **4** 409–435.

SOBOLEV, S. L. (1974). *Introduction to the Theory of Cubature Formulae.* Izdat Nauka, Moscow.

SOBOLEV, S. L. (1992). *Cubature Formulas and Modern Analysis.* Oxonian Press, New Delhi.

STEIN, M. L. (1993a). Asymptotic properties of centered systematic sampling for predicting integrals of spatial processes. *Ann. Appl. Probab.* **3** 874–880.

STEIN, M. L. (1993b). Predicting integrals of random fields using observations on a lattice. Technical Report 330, Dept. Statistics, Univ. Chicago.

TRAUB, J. F., WASILKOWSKI, G. W. and WOŹNIAKOWSKI, H. (1988). *Information-Based Complexity.* Academic Press, San Diego.

VOSS, R. F. (1989). Random fractals: self-affinity in noise, music, mountains and clouds. *Phys. D* **38** 362–371.

WAHBA, G. (1971). On the regression problem of Sacks and Ylvisaker. *Ann. Math. Statist.* **42** 1035–1053.

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

WAHBA, G. (1990). *Spline Models for Observational Data.* SIAM, Philadelphia.

YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions* **1**. Springer, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
5734 UNIVERSITY AVE.
CHICAGO, ILLINOIS 60637