The Annals of Statistics
1978, Vol. 6, No. 1, 177-184

WEAK AND STRONG UNIFORM CONSISTENCY OF THE KERNEL ESTIMATE OF A DENSITY AND ITS DERIVATIVES

BY BERNARD W. SILVERMAN

University of Cambridge

The estimation of a density and its derivatives by the kernel method is considered. Uniform consistency properties over the whole real line are studied. For suitable kernels and uniformly continuous densities it is shown that the conditions $h \to 0$ and $(nh)^{-1} \log n \to 0$ are sufficient for strong uniform consistency of the density estimate, where n is the sample size and h is the "window width." Under certain conditions on the kernel, conditions are found on the density and on the behavior of the window width which are necessary and sufficient for weak and strong uniform consistency of the estimate of the density derivatives. Theorems on the rate of strong and weak consistency are also proved.

1. Introduction. Consider the kernel estimate f_n of a real univariate density f introduced by Rosenblatt (1956) and defined by

$$f_n(x) = \sum_{i=1}^{n} (nh)^{-1}\delta\{h^{-1}(x - X_i)\} \tag{1}$$

where $X_1, \cdots, X_n$ are independent observations from the density, δ is a kernel function, and $h(n)$ is a "window width." The explicit dependence of h on n will generally be suppressed.

The weak and strong uniform consistency properties of f_n have been considered by several authors, including Nadaraya (1965), Schuster (1969) and Van Ryzin (1969). In these papers the conditions placed on the window width for strong uniform consistency include $\Sigma \exp(-cnh^2) < \infty$ for all positive c. In this paper this condition is substantially weakened. The exact rate of weak and strong uniform convergence of $(f_n - \mathbb{E}f_n)$ to zero is demonstrated under mild conditions on δ, f and h, thus giving the best possible improvement of the result given by Schuster (1969). The conditions on h are again much weaker than those assumed by previous authors.

The final section deals with estimation of density derivatives and gives conditions under which the estimate of density derivatives due to Bhattacharya (1967) is uniformly consistent. The conditions on h are shown to be necessary as well as sufficient. Necessary and sufficient conditions in density estimation were also considered by Schuster (1969), who concentrated on the conditions placed on the density f and showed that the uniform continuity of f was necessary for uniform consistency, under the above condition on h.

Received June 1975; revised April 1977.

AMS 1970 *subject classifications.* Primary 62G05, 41A25, 60F15; Secondary 60G15, 60G17.

Key words and phrases. Density estimate, density derivative estimate, global consistency, kernel, Gaussian process, rates of convergence, modulus of continuity, supremum over real line.

The expressions sup, inf and $\int$, when unqualified, will be taken to be over the range $(-\infty, \infty)$. Let $F(x) = \int_{-\infty}^{x} f(t)\,dt$ throughout. Given any real function u of bounded variation, u induces a signed Lebesgue-Stieltjes measure μ_u on the Borel sets of the real line; this measure ascribes measure $u(b) - u(a)$ to the real interval $(a, b]$. The notation $\int g(x)|du(x)|$ will be used to denote an integral with respect to the absolute value, or total variation, of the measure μ_u, as defined on pages 117–118 of Rudin (1966). In the case where u is differentiable, $\int g(x)|du(x)|$ is equivalent to $\int g(x)|u'(x)|\,dx$.

2. Preliminary results. This section contains some results which will be useful in the main part of the paper. The first result demonstrates an elegant decomposition of the estimate.

PROPOSITION 1. *On a suitable probability space there is a version W^0 of the Brownian bridge such that*

$$f_n(x) = f(x) + b(x) + n^{-\frac{1}{2}}\rho(x) + \varepsilon(x)$$

where

$$b(x) = \mathbb{E}f_n(x) - f(x) = \int h^{-1}\delta\{h^{-1}(x-t)\}\,dF(t) - f(x)$$

$$\rho(x) = -\int h^{-1}W^0\{F(t)\}\,d_t\delta\{h^{-1}(x-t)\}$$

and

$$\varepsilon(x) = -(nh)^{-1}\log n \int Z_n(t)\,d_t\delta\{h^{-1}(x-t)\}$$

where, for some absolute constant C_0,

$$\limsup_n \sup_t |Z_n(t)| \leqq C_0 \quad \text{a.s.} \tag{2}$$

PROOF. Write $f_n(x)$ as $h^{-1}\int \delta\{h^{-1}(x-t)\}\,dF_n(t)$ and then use the probability integral transformation and Theorem 3 of Komlós, Major and Tusnády (1975) to decompose the empirical distribution function F_n. This technique, though not this notation, was used by Bickel and Rosenblatt (1973). To complete the proof, choose positive constants C, K and λ as in Theorem 3 of Komlós, Major and Tusnády (1975), which then gives, for any $\eta > \lambda^{-1}$,

$$\sum_n \Pr[\sup_t |Z_n(t)| > C + \eta] \leqq K \sum_n n^{-\lambda\eta} < \infty\,.$$

Setting $C_0 = C + \lambda^{-1}$, the result (2) now follows by applying the first Borel-Cantelli lemma to the sets

$$[\sup_t |Z_n(t)| > C + \eta] \qquad \text{for arbitrary} \quad \eta > \lambda^{-1}\,.$$

Notice that b and the probability structure of the Gaussian process ρ depend only on the window width h and not on the sample size n. The dependence of b, ρ and ε on n and h will not be explicitly expressed. The next proposition gives some properties of b and ε which follow easily from Proposition 1.

PROPOSITION 2. *Provided f is uniformly continuous, $b(x) \to 0$ uniformly over* $\mathbf{R}$ *as $h \to 0$, and, whatever the dependence of h on n,*

$$\limsup_{n\to\infty} nh(\log n)^{-1} \sup|\varepsilon(x)| \leqq C_0 V(\delta) \quad \text{a.s.}$$

where C_0 is the constant of Proposition 1 *and $V(\delta)$ is the variation of δ.*

PROOF. The first part follows immediately from elementary analysis. To prove the second part, note that, by standard properties of the integral,

$$nh\,(\log n)^{-1} \sup_x |\varepsilon(x)| \leqq V(\delta) \sup_t |Z_n(t)|$$

and apply Proposition 1.

The final result of this section is a lemma on the modulus of continuity of the Brownian bridge, an easy consequence of results of Garsia (1970) on the modulus of continuity of general Gaussian processes, combined with standard properties of the Brownian bridge.

PROPOSITION 3. *Let W^0 be a continuous version of the Brownian bridge, with modulus of continuity w_0. Let*

$$\begin{aligned} p(u) &= u^{\frac{1}{2}}(1-u)^{\frac{1}{2}} & 0 \leqq u \leqq \tfrac{1}{2} \\ &= \tfrac{1}{2} & u > \tfrac{1}{2} \end{aligned}$$

and let

$$q(u) = \textstyle\int_0^u \{\log(1/v)\}^{\frac{1}{2}}\, dp(v)\,.$$

Then, with probability one, w_0 is dominated by

$$16(\log B)^{\frac{1}{2}} p + 16(2)^{\frac{1}{2}} q$$

where B is a random variable with $B \geqq 1$ a.s. *and $\mathbb{E}B < 4(2)^{\frac{1}{2}}$.*

3. Estimation of the density. The first theorem of this section demonstrates the uniform consistency of the estimates under very mild conditions on h. The following conditions on δ are used in the theorems of this section.

(C1)
- (a) δ is uniformly continuous (with modulus of continuity w_δ) and of bounded variation $V(\delta)$
- (b) $\int |\delta(x)|\, dx < \infty$ and $\delta(x) \to 0$ as $|x| \to \infty$
- (c) $\int \delta(x)\, dx = 1$
- (d) $\int |x \log |x||^{\frac{1}{2}} |d\delta(x)| < \infty$

(C2)
- (e) setting $\gamma(u) = \{w_\delta(u)\}^{\frac{1}{2}}$, $\int_0^1 \{\log(1/u)\}^{\frac{1}{2}}\, d\gamma(u) < \infty$.

Conditions (C1) and (C2) are satisfied by a very wide variety of kernels, for example the normal density function, the Cauchy density function, and the spline kernel of Boneva, Kendall and Stefanov (1971).

THEOREM A. *Suppose δ satisfies conditions* (C1) *and f is uniformly continuous. Suppose $h \to 0$ and $(nh)^{-1} \log n \to 0$ as $n \to \infty$. Then, defining f_n as in* (1) *above,*

$$\sup |f_n - f| \to 0 \quad \text{a.s.} \qquad as \quad n \to \infty\,.$$

To prove Theorem A, decompose f_n as in Proposition 1 and apply Proposition 2 to dispose of the terms b and ε. It only remains to show that $\sup |n^{-\frac{1}{2}}\rho| \to 0$ to complete the proof. This is a consequence of the next proposition, giving asymptotic uniform bounds on the behavior of the process ρ *outside* a given interval I. To prove Theorem A, it will suffice to take an empty interval I, but the full result will be useful in the proof of Theorem B below. The notation p lim sup denotes the lim sup in probability. For notational convenience set

$$\alpha(h) = h^{\frac{1}{2}}\{\log(1/h)\}^{-\frac{1}{2}}$$

and

$$\beta(h) = h^{-\frac{1}{2}}\{\log(1/h)\}^{-\frac{1}{2}}.$$

PROPOSITION 4. *Suppose I is an interval (a, b), possibly empty. Suppose f is uniformly continuous and δ satisfies conditions* (C1) (a), (b) *and* (d). *Let* $M = \sup\{f(x) : x$ *outside* $I\}$. *Then, provided $h \to 0$ as $n \to \infty$,*

$$p \limsup_{n\to\infty} \alpha(h) \sup_{\mathbf{R}\setminus I} |\rho| \leqq 16 M^{\frac{1}{2}} 2^{\frac{1}{2}} \int |x|^{\frac{1}{2}} |d\delta(x)|.$$

If, in addition, $\Sigma h^{\Lambda} < \infty$ for some Λ,

$$\limsup_{n\to\infty} \alpha(h) \sup_{\mathbf{R}\setminus I} |\rho| \leqq 16 M^{\frac{1}{2}}(2^{\frac{1}{2}} + \Lambda^{\frac{1}{2}}) \int |x|^{\frac{1}{2}} |d\delta(x)| \quad \text{a.s.}$$

PROOF. Define $M_0 = \sup f$ and $M = \sup_{\mathbf{R}\setminus I} f$. Given any $M' > M$, by the continuity of f choose ε in $(0, 2M_0^{-1})$ such that

$$\sup\{f(x) : x \leqq a + \varepsilon \ \text{ or } \ x \geqq b - \varepsilon\} \leqq M'.$$

Given any x outside I, it follows from the definitions of ρ and w_0 and from Theorem 6.12 of Rudin (1966) that

$$\begin{aligned} |\rho(x)| &= |\textstyle\int h^{-1}[W^0\{F(x - h\xi)\} - W^0\{F(x)\}]\, d\delta(\xi)| \\ &\leqq h^{-1} \textstyle\int w_0\{|F(x - h\xi) - F(x)|\}|d\delta(\xi)|. \end{aligned}$$

Applying Taylor's theorem and the definition of ε,

$$\begin{aligned} \sup_{\mathbf{R}\setminus I} |\rho| \leqq\ & h^{-1} \textstyle\int_{|\xi|<\varepsilon/h} w_0(M'|\xi|h)|d\delta(\xi)| \\ & + h^{-1} \textstyle\int_{|\xi|\geqq\varepsilon/h} w_0(M_0|\xi|h)|d\delta(\xi)|. \end{aligned}$$

Now use Proposition 3 and the fact that $M_0|\xi|h \geqq \frac{1}{2}$ for $|\xi| \geqq \varepsilon/h$ to obtain, with probability one,

$$\begin{aligned} \alpha(h) \sup_{\mathbf{R}\setminus I} |\rho| \leqq\ & 16\beta(h)\{q(\tfrac{1}{2})2^{\frac{1}{2}} + p(\tfrac{1}{2})(\log B)^{\frac{1}{2}}\} \textstyle\int_{|\xi|\geqq\varepsilon/h} |d\delta(\xi)| \\ & + 16\beta(h)2^{\frac{1}{2}} \textstyle\int_{|\xi|<\varepsilon/h} q(M'|\xi|h)|d\delta(\xi)| \\ & + 16\beta(h)(\log B)^{\frac{1}{2}} \textstyle\int_{|\xi|<\varepsilon/h} p(M'|\xi|h)|d\delta(\xi)|. \end{aligned} \tag{3}$$

The terms of (3) are dealt with separately. Condition (d) of (C1) implies immediately that, for any $\varepsilon > 0$, as $h \to 0$,

$$\textstyle\int_{|\xi|\geqq\varepsilon/h} |d\delta(\xi)| = o\{\alpha(h)\}. \tag{4}$$

It follows from Proposition 3 that, if $h \to 0$ as $n \to \infty$,

$$(\log B)^{\frac{1}{2}}\{\log(1/h)\}^{-\frac{1}{2}} \to_P 0 \qquad \text{as} \quad n \to \infty \tag{5}$$

and that, if $\Sigma h^{\Lambda} < \infty$, by Borel-Cantelli,

$$\limsup_{n\to\infty} (\log B)^{\frac{1}{2}}\{\log(1/h)\}^{-\frac{1}{2}} \leqq \Lambda^{\frac{1}{2}} \quad \text{a.s.} \tag{6}$$

Combining (4) with (5) or (6) as appropriate shows that the first term of (3) can be neglected. Dealing with the second and third terms is a matter of investigation of the properties of the functions p and q.

For any $a > 0$ and sufficiently small h,

$$\beta(h)q(ah) = \int_0^a h^{\frac{1}{2}}p'(hv)\{1 + \log(1/v)/\log(1/h)\}^{\frac{1}{2}}\,dv \tag{7}$$

$$\rightarrow a^{\frac{1}{2}} \quad \text{as} \quad h \rightarrow 0 \tag{8}$$

by dominated convergence; the integrand in (7) is dominated for sufficiently small h by $(2v)^{-\frac{1}{2}}\{1 + |\log(1/v)|\}^{\frac{1}{2}}$, which is integrable on $[0, a]$, and the limit of the integrand is $\frac{1}{2}v^{-\frac{1}{2}}$ for each v.

For $M'|\xi|h \leqq \frac{1}{2}$ and sufficiently small h, the quantity $\beta(h)q(M'|\xi|h)$ is dominated by a constant multiple of $|\xi|^{\frac{1}{2}}(1 + \log|\xi|)^{\frac{1}{2}}$, which is integrable on $(-\infty, \infty)$ with respect to $|d\delta(\xi)|$ by condition (d) of (C1). Apply the dominated convergence theorem and (8) to obtain the result

$$\lim_{h\rightarrow 0}[\text{second term of } (3)] = 16(2M')^{\frac{1}{2}} \int |\xi|^{\frac{1}{2}}\,d\delta(\xi)\,.$$

The third term of (3) is now considered. Using the fact that $p(v)$ is dominated by $v^{\frac{1}{2}}$, the third term of (3) is dominated by

$$16M'^{\frac{1}{2}}(\log B)^{\frac{1}{2}}\{\log(1/h)\}^{-\frac{1}{2}} \int |\xi|^{\frac{1}{2}}|d\delta(\xi)|\,.$$

Apply (5) and (6) to give limiting results which complete the proof of Proposition 4.

Return to the proof of Theorem A. Given a sequence of window widths h which satisfy the hypotheses of Theorem A, let $N = \{n : h < n^{-\frac{1}{4}}\}$. Theorem 1 of Nadaraya (1965) implies that $\sup|f_n - f|$ tends to zero a.s. as n tends to infinity through N^c. [This is the only step which uses the uniform continuity of δ.] For n in N, Σh^8 is convergent, and so the second part of Proposition 4 implies that $\sup|n^{-\frac{1}{2}}\rho|$ is a.s. $O\{n^{-\frac{1}{2}}\alpha(h)\}$ and hence converges to zero as n tends to infinity through N. As remarked at the beginning of the proof, this completes the proof of Theorem A. The next theorem gives exact rates of weak and strong uniform consistency for the kernel density estimate.

THEOREM B. *Suppose that δ satisfies conditions* (C1) *and* (C2). *Suppose that f is uniformly continuous, and that, as $n \rightarrow \infty$, $h \rightarrow 0$ and*

$$(nh)^{-1}(\log n)^2\{\log(1/h)\}^{-1} \rightarrow 0\,.$$

Then

$$\{(nh)^{-1}\log(1/h)\}^{-\frac{1}{2}}\sup(f_n - \mathbb{E}f_n) \rightarrow_P C_1$$

and

$$\{(nh)^{-1}\log(1/h)\}^{-\frac{1}{2}}\inf(f_n - \mathbb{E}f_n) \rightarrow_P -C_1$$

where

$$C_1 = \{2\sup f \int \delta^2(x)\,dx\}^{\frac{1}{2}}\,.$$

If, in addition, $\Sigma h^{\Lambda} < \infty$ for some Λ,

$$1 \leqq C_1^{-1}\limsup_{n\rightarrow\infty}\{(nh)^{-1}\log(1/h)\}^{-\frac{1}{2}}\sup|f_n - \mathbb{E}f_n| \leqq (1+\Lambda)^{\frac{1}{2}} \quad \text{a.s.} \tag{9}$$

To prove Theorem B, decompose $f_n - \mathbb{E} f_n$ as $n^{-\frac{1}{2}}\rho + \varepsilon$ using Proposition 1, and use Proposition 2 to show that the contribution due to ε can be neglected. Notice that Proposition 4 applied to an empty interval I gives results which have the correct rate of convergence but the wrong constant C_1. The structure of the proof is to choose a suitable interval I and to use Proposition 4 to deal with $n^{-\frac{1}{2}}\rho$ outside I. An existing result of the author on the behavior of ρ within an interval is used to complete the proof.

Since f is uniformly continuous, $f(x) \to 0$ as $|x| \to \infty$. Choose a, $0 < a < \infty$, such that

$$\sup_{|x| \geqq a} f(x) < 512^{-1} C_1^2 / \{\int |x|^{\frac{1}{2}} |d\delta(x)|\}^2 = C_2\,, \quad \text{say.} \tag{10}$$

Proposition 4 then implies that

$$p \lim \sup_{n\to\infty} \alpha(h) \sup_{|x| \geqq a} |\rho(x)| < C_1\,. \tag{11}$$

By Theorem A of Silverman (1976), as n tends to infinity,

$$\begin{aligned} \alpha(h) \sup_{|x| \leqq a} \rho(x) &\to_P C_1 \qquad \text{and} \\ \alpha(h) \inf_{|x| \leqq a} \rho(x) &\to_P -C_1\,. \end{aligned} \tag{12}$$

Combining (11) and (12) completes the proof of the first part of Theorem B.

The first inequality of (9) can be deduced immediately. To demonstrate the second inequality, replace C_2 in (10) by $C_2(1 + 2^{-\frac{1}{2}}\Lambda^{\frac{1}{2}})^2$ and proceed similarly, using Theorem C of Silverman (1976) to give a strong result for the behavior of ρ within the interval $(-a, a)$.

4. Estimation of derivatives of the density. The results obtained in Section 3 can be extended to the case of estimating the rth derivative of a density. The estimator studied, suggested by Bhattacharya (1967), is obtained by taking the rth derivative of the Rosenblatt kernel estimate, using a kernel which is at least r times differentiable. In Theorem A sufficient conditions on h were given for uniformly consistent density estimates. In the case of estimates of density derivatives, however, the conditions on h will be shown to be sufficient *and necessary* for uniform consistency. Rates of consistency will not be considered here; corresponding results to Theorem B are easily obtained by observing that

$$h^r f_n^{(r)}(x) = \int h^{-1} \delta^{(r)}\{h^{-1}(x - t)\}\, dF(t)$$

and hence that $h^r f_n^{(r)}$ has the same structure as f_n with δ replaced by $\delta^{(r)}$. None of the work on $f_n - \mathbb{E} f_n$ uses condition (C1) (c), and so the fact that $\int \delta^{(r)}(x)\, dx$ is zero does not present any difficulty.

THEOREM C. *For some integer $r \geqq 1$, suppose δ is everywhere r times differentiable; that for $j = 0, \cdots, r$,*

$$\delta^{(j)}(x) \to 0 \quad \textit{as} \quad |x| \to \infty \qquad \textit{and} \qquad \int |\delta^{(j)}(x)|\, dx < \infty\,;$$

that $\int \delta(x)\, dx = 1$; that $\delta^{(r)}$ satisfies conditions (C1), (a), (b) *and* (d) *and condition* (C2); *and that the Fourier transform of δ is not identically one in any neighborhood* *of zero. Suppose f has uniformly continuous* rth *derivative. Then the conditions*

$$h \to 0 \quad \textit{and} \quad n^{-1}h^{-2r-1}\log(1/h) \to 0 \quad \textit{as} \quad n \to \infty$$

are necessary and sufficient for both

$$\sup|f_n^{(r)} - f^{(r)}| \to_P 0 \qquad \textit{as} \quad n \to \infty$$

and

$$\sup|f_n^{(r)} - f^{(r)}| \to 0 \quad \text{a.s.} \qquad \textit{as} \quad n \to \infty .$$

Notice that the condition on f is also necessary because the uniform continuity of $\delta^{(r)}$ implies that of $f_n^{(r)}$ and hence of any uniform limit of the $f_n^{(r)}$.

The proof of sufficiency follows almost exactly that of Theorem A and is omitted. The conditions are shown to be necessary for weak consistency; their necessity for strong consistency follows at once. Assume from now on that $\sup|f_n^{(r)} - f^{(r)}| \to 0$ in probability.

Suppose firstly that $h \nrightarrow 0$ but that (choosing a subsequence if necessary) $h \to h_0 \neq 0$. For each x, by the weak law of large numbers and the boundedness of $\delta^{(r)}$,

$$f_n^{(r)}(x) \to_P \int h_0^{-r-1}\delta^{(r)}\{h_0^{-1}(x-t)\}f(t)\,dt .$$

For weak uniform consistency it is therefore necessary that, for all x,

$$f^{(r)}(x) = \int h_0^{-r-1}\delta^{(r)}\{h_0^{-1}(x-t)\}f(t)\,dt$$

and hence, taking Fourier transforms and using Theorem 9.2 of Rudin (1966),

$$f^{(r)*}(s) = h_0^{-r}\delta^{(r)*}(h_0 s)f^*(s)$$

where * denotes Fourier transformation. By standard properties of Fourier transforms this becomes

$$(is)^r f^*(s) = (is)^r \delta^*(h_0 s)f^*(s) .$$

Since $f^*(s) \to 1$ as $s \to 0$, this implies that δ^* is identically one in a neighbourhood of zero, contradicting the assumption made in the theorem. Thus $h \to 0$ is necessary, since $h \to \infty$ would imply $\sup|f_n^{(s)}| \to 0$.

Now, letting $\tilde{f}_n$ be the estimate based on $X_2, \cdots, X_n$,

$$f_n^{(r)}(x) = n^{-1}h^{-r-1}\delta^{(r)}\{h^{-1}(x - X_1)\} + \{(n-1)/n\}\tilde{f}_n^{(r)}(x) .$$

Because $f_n^{(r)}$ is uniformly consistent, it is necessary that

$$\sup|n^{-1}h^{-r-1}\delta^{(r)}\{h^{-1}(x - X_1)\}| \to_P 0$$

and hence that $nh^{r+1} \to \infty$ as $n \to \infty$. For otherwise the asymptotic behaviour of $\tilde{f}_n$ would depend on X_1, contradicting the independence of $X_1, X_2, \cdots$.

By the analogous results to Propositions 1 and 2 and using Theorem A of Silverman (1976), using O_p to denote order in probability,

$$\begin{aligned}\frac{\sup|\varepsilon^{(r)}|}{\sup|n^{-\frac{1}{2}}\rho^{(r)}|} &= O_p[n^{\frac{1}{2}}h^{r+\frac{1}{2}}\{\log(1/h)\}^{-\frac{1}{2}}n^{-1}h^{-r-1}\log n] \\ &= o(1)\end{aligned}$$

since $h \to 0$ and $nh^{r+1} \to \infty$. [Notice that this conclusion cannot be drawn for $r = 0$.] It follows that $\sup |\varepsilon^{(r)}|$ is asymptotically negligible compared with $n^{-\frac{1}{2}} \sup |\rho^{(r)}|$ and hence that $n^{-\frac{1}{2}} \sup |\rho^{(r)}| \to_P 0$ is necessary for uniform consistency. By Theorem A of Silverman (1976), $\sup |\rho^{(r)}|$ is exactly $O_p[h^{-r-\frac{1}{2}} \{\log (1/h)\}^{\frac{1}{2}}]$ as $h \to 0$, and hence $n^{-1}h^{-2r-1} \log (1/h) \to 0$ is necessary for uniform consistency, completing the proof of Theorem C.

Acknowledgments. My sincere thanks are due to Dr. D. P. Kennedy and to the referee for comments on previous drafts of this paper.

REFERENCES

[1] BHATTACHARYA, P. K. (1967). Estimation of a probability density function and its derivatives. *Sankhyā Ser. A* **29** 373–382.

[2] BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095.

[3] BONEVA, L. I., KENDALL, D. G. and STEFANOV, I. (1971). Spline transformations. *J. Roy. Statist. Soc. B* **33** 1–70.

[4] GARSIA, A. M. (1970). Continuity properties of Gaussian processes with multidimensional time parameter. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 369–374. Univ. of California Press.

[5] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent random variables, and the sample distribution function. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** 111–131.

[6] NADARAYA, E. A. (1965). On non-parametric estimates of density functions and regression curves. *Theor. Probability Appl.* **10** 186–190.

[7] ROSENBLATT, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.* **27** 832–837.

[8] RUDIN, W. (1966). *Real and Complex Analysis.* McGraw-Hill, New York.

[9] SCHUSTER, E. F. (1969). Estimation of a probability density and its derivatives. *Ann. Math. Statist.* **40** 1187–1196.

[10] SILVERMAN, B. W. (1976). On a Gaussian process related to multivariate probability density estimation. *Math. Proc. Cambridge Philos. Soc.* **80** 135–144.

[11] VAN RYZIN, J. (1969). On strong consistency of density estimates. *Ann. Math. Statist.* **40** 1765–1772.

MATHEMATICAL INSTITUTE
24–29 ST. GILES
OXFORD OX1 3LB,
ENGLAND