The Annals of Mathematical Statistics
1971, Vol. 42, No. 6, 1815–1842

CURVE ESTIMATES[1]

BY MURRAY ROSENBLATT

University of California, San Diego

1. Introduction. There is a large class of problems in which the estimation of curves arises naturally (see [15], [34]). It is curious that one of the earliest extensive investigations of this type involves the estimation of the spectral density function when sampling from a stationary sequence ([1], [17], [27], [33]). Even though the simple histogram has been used for years, it was only later that the simpler question of estimating a probability density function was dealt with at some length ([26], [25], [9]). Because the final character of the usual results obtained in both problem areas is quite similar, and the arguments are much more transparent in the case of the probability density function, we shall develop the results for the probability density function first. Later some corresponding results for spectra will be given. The similarities and differences in the two areas will be noted. Since the literature is rather extensive by now, any presentation of theory as given can only be a selection of topics and cannot claim to be exhaustive or perhaps even representative. There are a number of attractive open problems that one can suggest solutions to on heuristic grounds. A few of these problems will be examined. In most cases it is clear that one will not use the techniques to be proposed in estimating a density function unless there is a good deal of data (many observations), little a priori information about the density function available, but a great need to get additional information about the density function, even if it is fairly crude.

2. Estimating the probability density function by independent observations. Consider a population with absolutely continuous distribution function $F(x)$ and probability density function $f(x) = F'(x)$. A simple sequence of estimates is determined by the choice of an *integrable bounded weight function* $w(u)$ with

$$\int w(u)\,du = 1 \tag{1}$$

and a *sequence of bandwidths* $b(n) \downarrow 0$ as $n \to \infty$. Notice that this implies that $\int w^2(u)\,du < \infty$. An estimate $f_n(x)$ of $f(x)$

$$f_n(x) = \frac{1}{nb(n)} \sum_{j=1}^{n} w\left(\frac{x - X_j}{b(n)}\right) \tag{2}$$

is determined by a sample $X_1, \cdots, X_n$ of independent observations from the population. If w is chosen to be nonnegative, the estimate $f_n(x)$ itself will be a probability

Received February 11, 1971.

[1] Invited Wald Lectures given at the annual meeting of the Institute of Mathematical Statistics held at Laramie, Wyoming, August 25–27, 1970; research supported by Office of Naval Research Contract N00014–69–A–0200–6016, Task NR–042–217. Reproduction in whole or in part is permitted for any purpose of the United States Government.

density function. Usually the weight function w will satisfy some additional regularity conditions.

First consider the mean and variance of the estimate $f_n(x)$. The mean

$$Ef_n(x) = \frac{1}{b(n)}\int w\left(\frac{x-u}{b(n)}\right)f(u)\,du = \int w(v)f(x-b(n)v)\,dv \tag{3}$$

while the variance

$$\sigma^2[f_n(x)] = \frac{1}{n}\left[\frac{1}{b(n)}\int w^2(v)f(x-b(n)v)\,dv - \left(\int w(v)f(x-b(n)v)\,dv\right)^2\right]. \tag{4}$$

The mean square error can then be simply written as

$$\begin{aligned} E|f_n(x)-f(x)|^2 &= \sigma^2[f_n(x)] + |Ef_n(x)-f(x)|^2 \\ &= \frac{1}{n}\left[\frac{1}{b(n)}\int w^2(v)f(x-b(n)v)\,dv - \left(\int w(v)f(x-b(n)v)\,dv\right)^2\right] \\ &\quad + \left|\int w(v)\{f(x-b(n)v)-f(x)\}\,dv\right|^2. \end{aligned} \tag{5}$$

A simple bound shows that *if* $nb(n) \to \infty$ *as* $n \to \infty$ *and* f *is a bounded function that is continuous at* x, *then there is consistency in mean square at* x *as* $n \to \infty$. The bound is now given. Let $\sup_x f(x) \leqq M$. Then

$$\sigma^2[f_n(x)] \leqq M[nb(n)]^{-1}\int w^2(v)\,dv. \tag{6}$$

For each $\varepsilon > 0$ let $\delta(\varepsilon) = \sup_{|y|\leqq\varepsilon}|f(x+y)-f(x)|$. Then

$$|Ef_n(x)-f(x)| \leqq 2M\int_{b(n)|v|\geqq\varepsilon}|w(v)|\,dv + \delta(\varepsilon). \tag{7}$$

The result follows by first letting $n \to \infty$ and then $\varepsilon \to 0$.

R. H. Farrel [14] showed that one cannot get a uniformly consistent sequence of estimators of the density function at a point (say $x = 0$) for a plausible class of density functions. He specifically showed that if C_M is the class of density functions f on $(-\infty, \infty)$ with

(a) f continuously differentiable everywhere

(b) $\sup_{x\in R} f(x) \leqq M$,

then even with a sequential estimator δ_N of $f(0)$, *the supremum of the mean square error*

$$\sup_{f\in C_M} E|\delta_N - f(0)|^2 \geqq \tfrac{1}{16}$$

if $M \geqq 3$ *and* $\sup_{f\in C_M} E_f N < \infty$. However, if one requires a Hölder condition for f itself in the neighborhood of zero, a uniformly consistent sequence of nonsequential estimators of $f(0)$ is easy to obtain. Let $C_{M,\varepsilon}$ be the class of continuously differentiable density functions satisfying conditions (a) and (b) and the additional requirement

(c) $|f(y)-f(0)| \leqq M|y|^{\varepsilon}$ for $|y| \leqq 1$

with $M, \varepsilon > 0$ *independent* of f. The types of bounds obtained in (6) and (7) indicate that *the sequence* $f_n(0)$ *is uniformly consistent in mean square if* $n \to \infty$ and $nb(n) \to \infty$, *for* $f \in C_{M,\varepsilon}$.

A number of people have suggested using the integrated mean square error

$$E\int |f_n(x)-f(x)|^2\, dx = \int E|f_n(x)-f(x)|^2\, dx \tag{8}$$

as a global measure of how good the estimator $f_n(x)$ is over R. Notice that

$$\begin{aligned}\int_{-\infty}^{\infty} \sigma^2[f_n(y)]\, dy &\leqq \frac{1}{nb(n)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w^2(v) f(x-b(n)v)\, dv\, dx \\ &= \frac{1}{nb(n)} \int_{-\infty}^{\infty} w^2(v)\, dv.\end{aligned} \tag{9}$$

Further

$$\begin{aligned}\int_{-\infty}^{\infty} (Ef_n(x)-f(x))^2\, dx &= \int_{-\infty}^{\infty} \{\int_{-\infty}^{\infty} w(u)[f(x-b(n)u)-f(x)]\, du\}^2\, dx \\ &\leqq A \int_{-\infty}^{\infty} |w(u)| \int_{-\infty}^{\infty} \{f(x-b(n)u)-f(x)\}^2\, dx\, du\end{aligned} \tag{10}$$

where

$$A = \int |w(u)|\, du \tag{11}$$

if $f \in L^2$. Let

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\lambda} \phi(\lambda)\, d\lambda. \tag{12}$$

Then

$$\int_{-\infty}^{\infty} \{f(x-b(n)u)-f(x)\}^2\, dx = \int_{-\infty}^{\infty} |\phi(\lambda)|^2 |\exp[-ib(n)u\lambda]-1|^2\, d\lambda. \tag{13}$$

Since $|\exp[-b(n)u\lambda]-1| \leqq b(n)|u\lambda|$ it follows that (10) is bounded above by

$$Ab(n)^2 \int_{-\infty}^{\infty} |w(u)|u^2\, du \int_{-\infty}^{\infty} |\phi(\lambda)|^2 \lambda^2\, d\lambda.$$

We make use of (9) and (10) and note that *if* $f_n(x)$ *is a sequence of estimators of* $f(x)$ *derived from a weight function* w *with*

$$\int |w(u)|u^2\, du < \infty, \tag{14}$$

then the integrated mean square error of f_n *tends to zero uniformly as* n, $nb(n) \to \infty$ *for all* f *with* $\int |f(x)|^2 dx$, $\int |f'(x)|^2 dx \leqq k < \infty$. The condition $\int |f'(x)|^2 dx \leqq k < \infty$ actually implies that $\int |f(x)|^2 dx \leqq k+1$.

For convenience *assume that* f *and its first two derivatives are continuous and bounded.* Then simple approximations can be given locally for the variance and bias of $f_n(x)$ as $n \to \infty$, $nb(n) \to \infty$, if $\int |w(u)| u^2 du < \infty$. It is then immediately clear from (3) and (4) that

$$\sigma^2[f_n(x)] \cong \frac{f(x)}{nb(n)} \int w^2(v)\, dv \tag{15}$$

if $f(x) > 0$ and (assuming $\int w(u) u du = 0$) that

$$Ef_n(x) - f(x) = \tfrac{1}{2} b(n)^2 f''(x) \int w(u) u^2 \, du + o(b(n))^2. \tag{16}$$

Notice that if $f(x) = 0$ with $f''(x) > 0$, then

$$\sigma^2[f_n(x)] \cong \frac{b(n)}{2n} f''(x) \int w^2(v)\, dv$$

as $n \to \infty$, $b(n) \to 0$.

With additional conditions on f (in terms of smoothness) and on w one can improve the rate at which the bias tends to zero. Such a discussion has been given in Bartlett [2] and we shall briefly indicate how this might be done. However, it will be clear that in order to bring this about weight functions which assume negative values must be used. In many situations such weight functions would not be plausible. *Suppose that f and its first four derivatives are continuous and bounded.* Further, let w be such that

$$\int w(u) u \, du = \int w(u) u^2 \, du = \int w(u) u^3 \, du = 0, \quad \int |w(u)| u^4 \, du < \infty. \tag{17}$$

Then

$$Ef_n(x) - f(x) = \frac{1}{4!} b(n)^4 f^{(4)}(x) \int w(u) u^4 \, du + o(b(n))^4. \tag{18}$$

A weight function satisfying the conditions (17) is given, for example, by

$$w(u) = \frac{3}{2}\left(1 - \frac{u^2}{3}\right) e^{-\frac{1}{2}u^2}.$$

There have been discussions in the past about the "optimal" shape of a window or weight function in spectral analysis (see [24]). The arguments usually given have been of an asymptotic character and it is a mistake to take them too literally from a finite sample point of view. But even asymptotic arguments if used and interpreted with care can yield meaningful ideas. Let us briefly consider a simple discussion due to Epanechnikov [13]. Under the assumptions leading to (17) and (18), the local estimate of mean square error obtained is

$$E|f_n(x) - f(x)|^2 = \frac{f(x)}{nb(n)} \int w^2(v)\, dv + \tfrac{1}{4}[b(n)]^4 [f''(x)]^2 \left(\int w(u) u^2 \, du\right)^2 + o\left(\frac{1}{nb(n)} + (b(n))^2\right). \tag{19}$$

The rate at which the mean square error tends to zero as $n \to \infty$ is maximized (if $\int w(u)u^2 du \neq 0$) if we set

$$b(n) = Kn^{-1/5} \tag{20}$$

with K the constant

$$K = \left[\frac{4f(x)\int w^2(v)\,dv}{(f''(x)\int w(v)v^2\,dv)^2}\right]^{1/5}. \tag{21}$$

The mean square error then becomes

$$E|f_n(x)-f(x)|^2 \cong 2^{3/5}[f(x)\int w^2(v)\,dv]^{4/5}|f''(x)\int w(v)v^2\,dv|^{2/5}n^{-4/5}+o(n^{-4/5}) \tag{22}$$

and this suggests that the integrated mean square error $\int E|f_n(x)-f(x)|^2 dx$ will be no smaller than

$$2^{3/5}\int f(x)^{4/5}|f''(x)|^{2/5}\,dx\,[\int w^2(v)\,dv]^{4/5}(\int w(v)v^2\,dv)^{2/5}n^{-4/5}$$

to the first order as $n \to \infty$. However, this would require locally scaling $b(n)$ with K depending on x when estimating $f(x)$. If the scaling were to be global with K independent of x, we would expect

$$\int E|f_n(x)-f(x)|^2\,dx = \frac{1}{nb(n)}\int w^2(v)\,dv+\tfrac{1}{4}b^4(n)\int_{-\infty}^{\infty}|f''(x)|^2\,dx\,(\int w(u)u^2\,du)^2 + o\left(\frac{1}{nb(n)}+b^4(n)\right).$$

This suggests taking $b(n) = Kn^{-1/5}$ with

$$K = \frac{2^{2/5}[\int w^2(v)\,dv]^{1/5}}{(\int (f''(x))^2\,dx)^{1/5}\{\int w(u)u^2\,du\}^{2/5}}$$

leading to

$$\int E|f_n(x)-f(x)|^2\,dx = 2^{3/5}[\int w^2(v)\,dv]^{4/5}[\int |f''(x)|^2\,dx\,(\int w(v)v^2\,dv)^2]^{1/5}n^{-4/5} + o(n^{-4/5}). \tag{23}$$

The estimate (23) can be obtained rigorously under the assumptions that

(i) *f is bounded, twice continuously differentiable with $f, f'' \in L^2$*

(ii) *w is bounded nonnegative and symmetric with $\int w(u)u^2du < \infty$.*

For then we can make use of the simple bound

$$\int [\int w(v)f(x-b(n)v)\,dv]^2\,dx = \int w(v)w(v')\{\int f(x-b(n)v)f(x-b(n)v')\,dx\}\,dv\,dv' \leqq [\int w(v)\,dv]^2 \int f(x)^2\,dx$$

and the estimate $f(x-b(n)u)-f(x) = -b(n)uf'(x)+\frac{1}{2}b(n)^2u^2f''(x-\theta b(n)u)$. The estimate implies that $\{\int w(u)[f(x-b(n)u)-f(x)]du\}^2 = \{\int w(u)\frac{1}{2}b(n)^2u^2f''(x-\theta b(n)u)\,du\}^2$. The expression for K and (23) can then be obtained with a small additional argument by using (4) and (10).

The type of discussion has been carried out several places some time ago. Epanechnikov's simple observation is that one simply ought to look for the weight function w minimizing

$$\int w^2(v)\,dv \tag{24}$$

subject to the restraints

$$\begin{aligned} &\text{(i)} \quad \int w(v)\,dv = 1 \\ &\text{(ii)} \quad w(v) = w(-v) \\ &\text{(iii)} \quad \int v^2 w(v)\,dv = 1. \end{aligned} \tag{25}$$

If δw represents a small deviation for an extremum subject to the restraints (i)–(iii), the variation of

$$\int_0^\infty w^2(v)\,dv + \lambda_1\{\int_0^\infty w(v)\,dv - \tfrac{1}{2}\} + \lambda_2\{\int_0^\infty w(v)v^2\,dv - \tfrac{1}{2}\}$$

should be zero (λ_1 and λ_2 are multipliers) and one is therefore led to

$$2\int_0^\infty w(v)\,\delta w(v)\,dv + \lambda_1 \int_0^\infty \delta w(v)\,dv + \lambda_2 \int_0^\infty \delta w(v) v^2\,dv = 0.$$

Thus

$$2w(v) + \lambda_1 + \lambda_2 v^2 = 0, \qquad w(v) = (-\lambda_1 - \lambda_2 v^2)/2.$$

The function $w(v)$ is zero at

$$v = \pm(-\lambda_1/\lambda_2)^{\frac{1}{2}}.$$

To have a function that is integrable and symmetric one ought to set

$$\begin{aligned} w(v) &= (-\lambda_1 - \lambda_2 v^2)/2 \qquad \text{if } |v| \leqq (-\lambda_1/\lambda_2)^{\frac{1}{2}} \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

The constants λ_1, λ_2 are determined by the conditions (i) and (iii). Then $\lambda_1 = -\frac{3}{4}5^{-\frac{1}{2}}$, $\lambda_2 = \frac{3}{4}t^{-\frac{3}{2}}$ so that the conjectured function is

$$\begin{aligned} w(v) &= \tfrac{3}{4}5^{-\frac{1}{2}}(1 - v^2/5) \qquad \text{if } |v| \leqq 5^{\frac{1}{2}} \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{26}$$

This is a nonnegative weight function. Consider a variation $\delta w(v)$ about (26) which is such that $\int \delta w(v)dv = 0$, $\int v^2 \delta w(v) dv = 0$ and $\delta w(-v) = \delta w(v)$ with $\delta w(v) \geqq 0$ for $|v| \geqq 5^{\frac{1}{2}}$. Then

$$\begin{aligned} \int (w(v) + \delta w(v))^2\,dv &= \int (w(v) + \delta w(v))^2\,dv - \tfrac{3}{2}5^{-\frac{1}{2}}\{\int (w(v) + \delta w(v))\,dv - 1\} \\ &\quad + \tfrac{3}{2}5^{-\frac{3}{2}}\{\int (w(v) + \delta w(v))v^2\,dv - 1\} \\ &= \int w(v)^2\,dv + \int \{2w(v) - \tfrac{3}{2}5^{-\frac{1}{2}} - \tfrac{3}{2}5^{-\frac{3}{2}}v^2\}\,\delta w(v)\,dv \\ &\quad + \int (\delta w(v))^2\,dv \\ &\geqq \int w^2(v)\,dv + \int (\delta w(v))^2\,dv \end{aligned} \tag{27}$$

because the integrand in the second integral of line (27) is zero for $|v| \leqq 5^{\frac{1}{2}}$ and nonnegative for $|v| \geqq 5^{\frac{1}{2}}$. *Thus the function $w(v)$ given by* (26) *gives an absolute minimum for* (24) *under the restraints* (i) *to* (iii) *in the class of nonnegative weight functions.* Exactly this variational problem was solved by Lehmann and Hodges in a nonparametric investigation [21]. Condition (ii) could be replaced by (ii)′ $\int uw(u)du = 0$.

However, one should note that if weight functions w with negative values are allowed, expression (24) can be made as small as is desired even though the restraints (i)–(iii) are satisfied. We reproduce a small part of a table of Epanechnikov (this is the point of greatest interest in his paper) to show how insensitive (24) is to the shape of w when dealing with nonnegative weight functions satisfying (i)–(iii). The minimizing weight function (26) is referred to as w_0.

TABLE 1

w		$L = \int w^2(u)\, du$	$r = \int w^2(u)\, du / \int w_0^2(u)\, du$
w_0		$3 \cdot 5^{-\frac{3}{2}}$	1
$1/6^{\frac{1}{2}} - \lvert y\rvert/6$	if $\lvert y\rvert \leqq 6$	$6^{\frac{1}{2}}/9$	1.015
0	if $\lvert y\rvert > 6$		
$(2\pi)^{-\frac{1}{2}}\, e^{-y^2/2}$		$2^{-1}\pi^{-\frac{1}{2}}$	1.051
$\frac{1}{2}\, 3^{-\frac{1}{2}}$	if $\lvert y\rvert \leqq 3^{\frac{1}{2}}$	$\frac{1}{2}\, 3^{-\frac{1}{2}}$	1.077
0	if $\lvert y\rvert > 3^{\frac{1}{2}}$		

From this table it is clear that one does almost as well with the Gaussian or even the rectangular weight function as one does with the optimal nonnegative weight function w_0. In computation, there would be clear advantages in dealing with a bandlimited weight function like w_0.

When $nb(n) \to \infty$ as $n \to \infty$ $(b(n) \to 0)$ the central limit theorem can be used to get a normal approximation to the probability distribution of $f_n(x)$. Some idea as to the error in the approximation is given by the Berry-Esseen theorem. Specifically, if the weight function w is bounded and nonnegative with f continuous, bounded and positive then

$$(28) \qquad |P[\sigma(f_n(x))^{-1}[f_n(x) - Ef_n(x)] \leqq x] - \Phi(x)| \leqq K/(nb(n))^{\frac{1}{2}}$$

where K for sufficiently large n is bounded above by

$$(29) \qquad 9 \int w(u)^3\, du\, \{\int w(u)^2\, du\}^{-\frac{3}{2}}[f(x)]^{-\frac{1}{2}}.$$

If f is twice continuously differentiable with $f(x) = 0$ and $f''(x) > 0$, one would require $nb(n)^3 \to \infty$ and $\int w^2(v)v^2 dv < \infty$ for asymptotic normality of the estimate $f_n(x)$. It is of some interest to see whether one can get better and more detailed approximations for the distribution of such estimates when using a simple band-limited weight function $w(u)$ (say triangular or even uniform). Steepest descent methods (see Daniels [12]) may be useful. An alternative procedure would be to condition the number of summands in (2) that are nonzero, if the weight function is bandlimited. Approximations suggested by a paper of Kolmogorov [23] may then be helpful.

It is clear that corresponding results can be obtained for estimates of multi-dimensional probability density functions. Just a few remarks will be made in the case of two dimensional estimates. *Let $w(u, v)$ be a bounded integrable weight function with*

$$\int w(u, v)\, du\, dv = 1. \tag{30}$$

Again let $b(n)$ be a linear bandwidth tending to zero as $n \to \infty$. The population is assumed to have an absolutely continuous two dimensional distribution function $F(x, y)$ with continuous probability density

$$f(x, y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y} F(x, y).$$

The estimate $f_n(x, y)$ of $f(x, y)$

$$f_n(x, y) = \frac{1}{nb(n)^2} \sum_{j=1}^{n} w\left(\frac{x - X_j}{b(n)}, \frac{y - Y_j}{b(n)}\right) \tag{31}$$

is computed from a sample $(X_1, Y_1), \cdots, (X_n, Y_n)$ of independent observations from the population. Assume that f and its partial derivatives up to second order are bounded and continuous. Further let

$$\int |w(u, v)| u^2\, du\, dv, \int |w(u, v)| v^2\, du\, dv < \infty \tag{32}$$

with

$$\int w(u, v) u\, du\, dv = \int w(u, v) v\, du\, dv = 0. \tag{33}$$

Then the variance and mean of $f_n(x, y)$ are given asymptotically by

$$\sigma^2[f_n(x, y)] \cong n^{-1} b(n)^{-2} f(x, y) \int w^2(u, v)\, du\, dv \tag{34}$$

and

$$Ef_n(x, y) = f(x, y) + \tfrac{1}{2} b(n)^2 [f_{xx}(x, y) m_{2,0} + 2 f_{xy}(x, y) m_{1,1} + f_{yy}(x, y) m_{0,2}] + o(b(n)^2) \tag{35}$$

as $n \to \infty$, $b(n) \to 0$. Here the $m_{\alpha, \beta}$ are the moments

$$\int u^\alpha v^\beta w(u, v)\, du\, dv \tag{36}$$

of the weight function w. The estimate is asymptotically normal if $nb(n)^2 \to \infty$ as $n \to \infty$. A more detailed discussion of multidimensional density estimates can be found in [7] and [13].

3. The process $[f_n(x) - Ef_n(x)]$. In this section the object is to study the process $[f_n(x) - Ef_n(x)]$, over an interval (say $x \in [0, 1]$), suitably normalized so as to obtain a nontrivial result as $n \to \infty$. The simple result obtained will be derived under unpleasant and completely impractical conditions. However, heuristically the result ought to hold under reasonable conditions. As we shall see, the type of result obtained suggests that limit theorems for sequences of random processes over finite intervals whose length is asymptotically unbounded may be of considerable interest. The asymptotic behavior of the process $[f_n(x) - Ef_n(x)]$ at a special set of points $x_j \in [0, 1]$, $j = 1, \cdots, n$, has been considered by Woodroofe [38]. To make it easier to follow the argument, the exposition is broken down into a number of stages. At the end the assumptions made in these stages are put together to get a coherent result. The argument is due to joint work with Peter Bickel.

(i) Consider the estimate $f_n(x)$ of $f(x)$ given by (2). Let us look at the covariance of $f_n(x), f_n(y)$

$$\text{(37)} \quad \operatorname{Cov}(f_n(x), f_n(y)) = \frac{1}{nb(n)^2} \int w\left(\frac{x-u}{b(n)}\right) w\left(\frac{y-u}{b(n)}\right) f(u)\, du - \frac{1}{n} \int w(z) f(x - b(n)z)\, dz \int w(z) f(y - b(n)z)\, dz.$$

From remarks made in Section 2, it is clear that to get a nontrivial limiting distribution at a fixed point x as $n \to \infty$, one ought to look at

$$\text{(38)} \quad (nb(n))^{\frac{1}{2}}[f_n(x) - Ef_n(x)].$$

We shall be interested in (38) as a process with parameter x varying over a finite range, say $0 \leqq x \leqq 1$, for convenience. To get a reasonable limiting covariance function, the scale in x has to be renormalized so that we look at

$$\text{(39)} \quad (nb(n))^{\frac{1}{2}}[f_n(b(n)x) - Ef_n(b(n)x)]$$

with $0 \leqq x \leqq b(n)^{-1}$. Thus

$$\text{(40)} \quad nb(n) \operatorname{Cov}[f_n(x), f_n(x + \alpha b(n))] = \int w(z) w(z + \alpha) f(x - b(n)z)\, dz - b(n) \int w(z) f(x - b(n)z)\, dz \int w(z) f(x + (\alpha - z) b(n))\, dz$$

and (40) tends to

$$\text{(41)} \quad f(x) \int w(z) w(z + \alpha)\, dz$$

as $n \to \infty$ and $b(n) \to 0$ if f is bounded and continuous at x while w is integrable and bounded. For asymptotic normality of (39) as $n \to \infty$, we require that

$nb(n) \to \infty$. Let us now assume that f is continuous and bounded away from zero on [0, 1]. Then we could look at

$$(nb(n))^{\frac{1}{2}}[f(x)]^{-\frac{1}{2}}[f_n(x) - Ef_n(x)] \tag{42}$$

so as to get the limiting covariance function

$$r(\alpha) = \int w(z)w(z+\alpha)\, dz, \tag{43}$$

with the proper change of scale.

(ii) We now compare the process (42) with the process obtained by looking at (42) at the points $x_j = jc(n)$ where $j = 1, 2, \cdots, [b(n)^{-1}c(n)^{-1}]$ ($[y]$ denotes the greatest integer less than or equal to y here) and interpolating linearly over the range $0 \leqq x \leqq b(n)^{-1}$. To estimate the difference between these two processes it will be enough to look at

$$\begin{aligned} \sup_{0 \leqq x \leqq b(n)^{-1}} (nb(n))^{\frac{1}{2}} \big| &f_n(b(n)x) - Ef_n(b(n)x) \\ &- f_n(b(n)[x/c(n)]c(n)) + Ef_n(b(n)[x/c(n)]c(n))\big|. \end{aligned} \tag{44}$$

The expression (44) can be conveniently rewritten as

$$\begin{aligned} \sup_{0 \leqq x \leqq b(n)^{-1}} (nb(n))^{\frac{1}{2}} \Big| \int &\{\exp[-itb(n)x] - \exp[-itb(n)[x/c(n)]c(n)]\} \\ &\cdot h(b(n)t)\{\varphi_n(t) - \varphi(t)\}\, dt \Big| \end{aligned} \tag{45}$$

where

$$h(t) = \int e^{itu} w(u)\, du \tag{46}$$

and

$$\varphi_n(t) = n^{-1} \textstyle\sum_{j=1}^n e^{itX_j}, \qquad \varphi(t) = E\exp(itX) \tag{47}$$

if $h \in L$. Notice that the second moment

$$\begin{aligned} &E(\sup_{0 \leqq x \leqq b(n)^{-1}} (nb(n))^{\frac{1}{2}} |\cdots|)^2 \\ &\quad \leqq nb(n) E\{\sup_{0 \leqq x \leqq b(n)^{-1}} \int |1 - \exp[itb(n)\{x - [x/c(n)]c(n)\}]|^2 |h(b(n)t)|\, dt \\ &\qquad \cdot \int |h(b(n)t)|\, |\varphi_n(t) - \varphi(t)|^2\, dt\} \end{aligned} \tag{48}$$

and that the right-hand side of inequality (48) is bounded by

$$nb(n)^3 c(n)^2 \int t^2 |h(b(n)t)|\, dt \int |h(b(n)t)| E|\varphi_n(t) - \varphi(t)|^2\, dt. \tag{49}$$

However,

$$E|\varphi_n(t) - \varphi(t)|^2 \leqq n^{-1}. \tag{50}$$

It follows from (49) and (50) that (44) tends to zero in probability as $n \to \infty$ if

$$c(n)^2/b(n) \to 0 \tag{51}$$

$(b(n) \to 0)$ and $h, t^2 h \in L$. Van Ryzin's paper [36] should be referred to for a related set of estimates.

(iii) Let ξ_i, $E\xi_i = 0$, $i = 1, \cdots, n$ be independent identically distributed random k vectors. Set

$$\rho_i = E|\xi_{1i}|^3 / E^{\frac{3}{2}} \xi_{1i}^2, \qquad i = 1, \cdots, k \tag{52}$$

where ξ_{1i} is the ith component of the random variable ξ_1. $F_n(x)$ is to denote the distribution function of $n^{-\frac{1}{2}} \sum_{i=1}^n \xi_i$ and $G(x)$ the k-variate normal distribution function with means zero and the same covariance structure as $F_n(x)$. Sazanov's Theorem 3 in [31] then states that

$$\sup_{x \in R^k} |F_n(x) - G(x)| \leqq C''(k) [\sum_{i=1}^k \rho_i n^{-\frac{1}{2}}]^{\frac{1}{3}} \tag{53}$$

where $C''(k) = Ck^2$ with C an absolute constant. In our context $k = [c(n)b(n)]^{-1}$ (the number of points $x_j = jc(n)$ in $0 \leqq x \leqq b(n)^{-1}$) and $\rho_i \leqq C' b(n)^{-\frac{1}{2}}$ for all i with C' an absolute constant. The estimate of the error on the right side of (53) becomes

$$C'' c(n)^{-7/3} b(n)^{-5/2} n^{-1/6}$$

with C'' an absolute constant and if we take $c(n), b(n) \to 0$ sufficiently slowly as $n \to \infty$, this estimate of the error will tend to zero.

(iv) The covariance function of the process (42) is

$$\{f(x)f(y)\}^{-\frac{1}{2}} b(n)^{-1} \Big[\int w\Big(\frac{x-u}{b(n)}\Big) w\Big(\frac{y-u}{b(n)}\Big) f(u)\, du - b(n)^2 \int w(z) f(x - b(n)z)\, dz \int w(z) f(y - b(n)z)\, dz \Big]. \tag{54}$$

Let ${}_0Y_n(x)$, $0 \leqq x \leqq 1$, be the Gaussian process with mean zero and covariance function (54). By using the Sazanov theorem, one can approximate the process (42) at the points $jc(n)b(n)$, $j = 1, \cdots, [c(n)b(n)]^{-1}$, in distribution by the Gaussian process ${}_0Y_n(x)$ at those same points. The type of argument used in (ii) shows that the supremum of the absolute difference between ${}_0Y_n(x)$ and the process obtained from ${}_0Y_n$ by linearly interpolating between the points $jc(n)b(n)$, $j = 1, \cdots, [c(n)b(n)]^{-1}$ tends to zero in probability as $n \to \infty$ if $c(n)^2/b(n) \to 0$ and $b(n) \to 0$. Let

$${}_1Y_n(x) = {}_0Y_n(x) + [(b(n)/f(x))^{\frac{1}{2}} \int w(z) f(x - b(n)z)\, dz] U \tag{55}$$

where U is a normal random variable with mean zero and variance one independent of the process ${}_0Y_n$. The process ${}_1Y_n(x)$ can be written as

$${}_1Y_n(x) = (b(n))^{-\frac{1}{2}} \int w\Big(\frac{x-u}{b(n)}\Big) \Big(\frac{f(u)}{f(x)}\Big)^{\frac{1}{2}} dB(u)$$

where $B(u)$ is a Brownian motion process with $-\infty < u < \infty$. Here $B(u)$ is a Gaussian process with increments over non-overlapping intervals independent, $B(0) \equiv 0$, and

$$E|B(u) - B(v)|^2 = |u - v|, EB(u) \equiv 0,$$

for $-\infty < u, v < \infty$. Further, the process ${}_1Y_n(x)$ can be approximated in the sup norm over $0 \leqq x \leqq 1$ by

$$ {}_2Y_n(x) = (b(n))^{-\frac{1}{2}} \int w\left(\frac{x-u}{b(n)}\right) dB(u). $$

This can be seen by making use of an idea from a paper of Garsia, Rodemich, and Rumsey [16]. Let

$$ U(x) = {}_1Y_n(x) - {}_2Y_n(x). \tag{56} $$

We are interested in estimating $\sup_{0 \leqq x \leqq 1} |U(x)|$. $U(x)$ is a Gaussian process with mean zero and covariance function

$$ R(x, y) = \frac{1}{b(n)} \int w\left(\frac{x-u}{b(n)}\right) w\left(\frac{y-u}{b(n)}\right) \left[\left(\frac{f(u)}{f(x)}\right)^{\frac{1}{2}} - 1\right]\left[\left(\frac{f(u)}{f(y)}\right)^{\frac{1}{2}} - 1\right] du. \tag{57} $$

Let

$$ \Delta R(x, y) = R(x, x) + R(y, y) - 2R(x, y). $$

The following Lemma is an immediate corollary of Theorem 2.2 of [16].

LEMMA. *Let $U(x)$, $0 \leqq x \leqq 1$, be a Gaussian process with mean zero and covariance function $R(x, y)$. Consider $p(u)$ a nonnegative even function that is continuous and nonnegative for $u \geqq 0$ and such that*

$$ \int_0^1 \frac{dp(u)}{u} < \infty. $$

If

$$ \int_0^1 \int_0^1 \frac{\Delta R(x, y)}{p^2(x-y)} dx\, dy < \infty $$

then

$$ \sup_{0 \leqq x \leqq 1} |U(x)| \leqq |U(0)| + B^{\frac{1}{2}} \int_0^1 \frac{dp(u)}{u} $$

where B is a nonnegative random variable with

$$ EB \leqq c \int_0^1 \int_0^1 \frac{\Delta R(x, y)}{p^2(x-y)} dx\, dy $$

and c an absolute constant.

We shall now apply this Lemma to estimate $E[\sup_{0 \leqq x \leqq 1} |U(x)|]^2$ where $U(x)$ is the Gaussian process (56) with covariance (57). In this case

$$ \Delta R(x, y) = \frac{1}{b(n)} \int [a(x, u) - a(y, u)]^2 \, du $$

with

$$ a(x, u) = w\left(\frac{x-u}{b(n)}\right)\left[\left(\frac{f(u)}{f(x)}\right)^{\frac{1}{2}} - 1\right]. $$

We clearly have

$$|\Delta R(x, y)| \leqq \frac{2}{b(n)} \int \{a(x, u)^2 + a(y, u)^2\} \, du.$$

Let f be continuously differentiable with f, f' bounded and f bounded away from zero on the interval [0, 1]. It then follows that

$$\frac{1}{b(n)} \int a(x, u)^2 \, du \leqq K_1 \int w(z)^2 [f(x - b(n)z) - f(x)]^2 \, dz$$
$$\leqq K_2 b(n)^2 \int w(z)^2 z^2 \, dz \leqq K_3 b(n)^2$$

if $\int w(z)^2 z^2 dz < \infty$ with the K_i constants. The uniform bound $|\Delta R(x, y)| \leqq 4K_3 b(n)^2$ for $\Delta R(x, y)$ is obtained. However, another bound will be required for $|\Delta R(x, y)|$ when x is close to y. Let $y - x = \alpha b(n)$ with $|\alpha| \leqq M < \infty$ where M is a constant. Then

$$|\Delta R(x, y)| \leqq 2\left[\int \left(w(z) - w\left(\frac{y-x}{b(n)} + z\right)\right)^2 \left[\left(\frac{f(x - b(n)z)}{f(x)}\right)^{\frac{1}{2}} - 1\right]^2 dz\right.$$
$$\left. + \int w\left(\frac{y-x}{b(n)} + z\right)^2 f(x - b(n)z)\{f(x)^{-\frac{1}{2}} - f(y)^{-\frac{1}{2}}\}^2 \, dz\right]$$
$$\leqq K_4 \int w'(z)^2 \, dz \, \alpha^2 + K_5 \alpha^2$$

if w' is continuous and bounded and in L^2. Thus $|\Delta R(x, y| \leqq K_6 \alpha^2$ if $y - x = \alpha b(n)$ with $|\alpha| \leqq M < \infty$. Let us set $p'(u) = u^{-\frac{1}{2}}$, $0 < u$, in applying the lemma since $\int_0^1 dp(u)/u = \int_0^1 u^{-\frac{1}{2}} du$ is finite. Also

$$\int_0^1 \int_0^1 \frac{\Delta R(x, y)}{p^2(x - y)} \, dx \, dy \leqq K \int_{b(n)^2 \leqq |x| < 2} \frac{b(n)^2}{x} \, dx + K' \int_{|x| \leqq b(n)^2} \frac{x^2}{b(n)^2 |x|} \, dx$$
$$\leqq K'' b(n)^2 |\log b(n)|.$$

Notice that the Gaussian process $Y(x) = {}_2Y_n(xb(n))$ has the desired covariance function

$$r(\alpha) = \int w(z) w(z + \alpha) \, dz.$$

We should now like to apply the result on the asymptotic distribution of the maximum of a Gaussian process as given, for example, on pages 271–272 of Cramér and Leadbetter [11] to the process $Y(x)$ on the range $0 \leqq x \leqq b(n)^{-1}$. Assume that the covariance function $r(\alpha)$ of the process $Y(x)$ as given by (43) is an even function that is continuously differentiable up to fourth order and such that

$$r(\alpha) = O(|\alpha|^{-\varepsilon}) \tag{58}$$

as $|\alpha| \to \infty$ for some $\varepsilon > 0$. Let

$$\lambda_0 = r(0) \tag{59}$$
$$\lambda_2 = r''(0).$$

Then the result cited in [11] indicates that

$$P\left[\max_{0\leq x\leq T}\lambda_0^{-\frac{1}{2}}Y(x)\leqq(2\log T)^{\frac{1}{2}}+\frac{A+z}{(2\log T)^{\frac{1}{2}}}\right]\to e^{-e^{-z}} \tag{60}$$

as $T\to\infty$ where

$$A=\log\{(-2\lambda_2/\lambda_0)^{\frac{1}{2}}/2\pi\}. \tag{61}$$

The estimates made in (ii), (iii) and (iv) (or obvious modifications of them) indicate that under appropriate conditions on f and w the distribution of the process (42) can be sufficiently well approximated by that of $Y(x)$ so as to obtain the following theorem.

THEOREM. *Let f be continuously differentiable and bounded away from zero on* [0, 1]. *Assume that the bounded, integrable weight function w used in the estimate f_n* (*see* (2)) *is twice continuously differentiable with $(1+v^4)|w''(v)|$ bounded and has a Fourier transform $h(t)$ with $(1+t^2)h(t)\in L$. Then if $n^{-1/24}=O(b(n))$ as $n\to\infty$, it follows that*

$$P\left[\max_{0\leq x\leq 1}\left(\frac{nb(n)}{\alpha f(x)}\right)^{\frac{1}{2}}\{f_n(x)-Ef_n(x)\}\leqq\{2\log b(n)^{-1}\}^{\frac{1}{2}}+\frac{A+z}{(2\log b(n)^{-1})^{\frac{1}{2}}}\right]\to e^{-e^{-z}} \tag{62}$$

as $n\to\infty$, where

$$\alpha=\int w^2(u)\,du \tag{63}$$

and

$$A=\log\frac{B^{\frac{1}{2}}}{2\pi},\qquad B=-\frac{2}{\alpha}\frac{d^2}{dt^2}\left(\int w(z)w(z+t)\,dz\right)\Big|_{t=0}. \tag{64}$$

The interest in the theorem given above is not in the detailed conditions. The result, in particular, must hold under much weaker requirements on $b(n)$. It is rather in its asymptotic nonparametric character and the ideas behind the proof. Notice that in the proof one approximates the process (42) in distribution with a corresponding Gaussian process.

The heuristics that suggest a result like that in the above Theorem also imply an asymptotic nonparametric character (under appropriate conditions) for

$$\int_0^1\frac{[f_n(x)-f(x)]^2}{f(x)}dx \tag{65}$$

when suitably normalized. There are the corresponding open questions for two-dimensional (and higher dimensional) density functions. For, example, let $f_n(x, y)$

be the estimate given in (31) of the density function $f(x, y)$. The marginal density estimates are

$$g_n(x) = \int f_n(x, y)\,dy \tag{66}$$
$$h_n(y) = \int f_n(x, y)\,dx.$$

A natural statistic to test independence over $[0, 1] \times [0, 1]$ is given by

$$\int_0^1 \int_0^1 \frac{|f_n(x, y) - g_n(x)h_n(y)|^2}{f_n(x, y)}\,dx\,dy \tag{67}$$

and its distribution should be investigated.

4. Probability density estimates for dependent sequences. It is also curious and pleasantly surprising that the asymptotic results on the behavior of probability density estimates obtained in Section 2 still hold even when sampling from a stationary process, if appropriate and rather reasonable conditions are satisfied. This is not true at all if the distribution function is estimated (see Billingsley [3, page 195] for a discussion) and the fact that it holds when estimating the probability density is due to the *local* character of the estimate.

Let $\{X_j, j = \cdots, -1, 0, 1, \cdots\}$ be a strictly stationary process. Assume that the instantaneous distribution function

$$F(x) = P[X_j \leqq x]$$

is absolutely continuous with continuous spectral density $f(x) = F'(x)$. We wish to estimate $f(x)$ by $f_n(x)$ as given in (2). The assumptions on the weight function w, the density function f, and the bandwidth $b(n)$ are the same as those in Section 2. The results on the bias of the estimate as given in (16) hold under the same conditions as in Section 2 since dependence has no effect on the bias. Let us now consider the covariance of the estimate. At this point the character of the dependence must be examined. Let the two dimensional distribution functions

$$F_j(x, y) = P[X_0 \leqq x, X_j \leqq y], \qquad j \neq 0, \tag{68}$$

be absolutely continuous with continuous density functions

$$f_j(x, y) = \frac{\partial}{\partial x}\frac{\partial}{\partial y} F_j(x, y). \tag{69}$$

The covariance of $f_n(x)$ and $f_n(y)$ is

$$\operatorname{Cov}[f_n(x), f_n(y)] = n^{-2} b(n)^{-2} \textstyle\sum_{j=-n}^{n} (n - |j|) \operatorname{Cov}[w(b(n)^{-1}(x - X_0)), w(b(n)^{-1}(y - X_j))] \tag{70}$$

with

$$\operatorname{Cov}[w(b(n)^{-1}(x - X_0)), w(b(n)^{-1}(y - X_j))] \tag{71}$$
$$= b(n)^2 \iint w(u)w(v)\{f_j(x - b(n)u, y - b(n)v) - f(x - b(n)u)f(y - b(n)v)\}\,du\,dv$$

if $j \neq 0$ and

(72) $\operatorname{Cov}[w(b(n)^{-1}(x-X_0)), w(b(n)^{-1}(y-X_0))]$

$$= b(n) \int w(u) w\left(u + \frac{y-x}{b(n)}\right) f(x-b(n)u)\, du$$
$$- b(n)^2 \int w(u) f(x-b(n)u)\, du \int w(v) f(x-b(n)v)\, dv.$$

Thus, if

$$\sum_{j \neq 0} |f_j(x,y) - f(x)f(y)| \leqq M < \infty \tag{73}$$

for all x and y, we have

$$nb(n)[\operatorname{Cov}(f_n(b(n)x), f_n(b(n)y))] = \int w(u) w(u+y-x) f(x-b(n)u)\, du + O(b(n)) \tag{74}$$

for all x and y as $n \to \infty$. The first term on the right of (74) is

$$f(x) \int w(u) w(u+y-x)\, du + O(b(n))$$

if f has a bounded continuous derivative. Additional conditions on the type of dependence are required to get a result on the asymptotic normality of the estimate $f_n(x)$. A standard result making use of the conditions will be described. The details of the proof can be found in [29]. Let $\mathscr{B}_n$ and $\mathscr{F}_m$ be the backward and forward Borel fields generated by the random variables $\{X_j; j \leqq n\}$ and $\{X_j; j \geqq m\}$ respectively. The stationary process $\{X_k\}$ satisfies the condition S if for every $\mathscr{F}_{n+k}$ measurable random variable $Y (k > 0)$ with $EY^2 < \infty$, $EY = 0$,

$$E|E(Y \mid \mathscr{B}_n)|^2 \leqq a(k) EY^2 \tag{75}$$

where

$$a(k) = O(k^{-4-\varepsilon}) \tag{76}$$

for some $\varepsilon > 0$ as $k \to \infty$. If $\{X_k\}$ satisfies condition S and the other requirements already specified, then

$$\{\int w^2(v)\, dv\}^{-\frac{1}{2}} (nb(n))^{\frac{1}{2}} [f_n(x) - Ef_n(x)]$$

is asymptotically normal with mean zero and variance one as $n \to \infty$, $nb(n) \to \infty$, $b(n) \to 0$. The paper of Roussas [30] discusses related questions.

The papers of Van Atta and Chen [34] and Frenkiel and Klebanoff [15] look at estimates of univariate and multivariate probability density functions without mentioning the estimation problem explicitly. They are interested in the probability density of the distribution of a component of the velocity in grid turbulence as well as the bivariate distribution of velocity components at the same point in space but with a time difference. They find that the univariate distribution appears to be normal but not the bivariate distribution.

A recent refinement in the "theory" of turbulence suggested by Kolmogorov is discussed in a paper of Yaglom [39]. The refinement suggests that a local spatial average of the energy dissipation might have a lognormal distribution. Currently, experimental results are being analyzed to estimate the probability density of such a local average of the energy dissipation to see whether it is reasonably approximated by a lognormal density.

5. Expansions in orthogonal functions. There are many ways of estimating a probability density function from a sample of independent observations on a population. Thus far in this paper, only estimates of the form (2) have been considered. Still another interesting class of estimates has been suggested by Van Ryzin [36]. An obvious class of estimates is also suggested by the notion of an orthogonal expansion. Just a few remarks will be made on this class but they will be enough to show that qualitatively results similar to those already obtained would be expected. However, it should be noted that there are still many open detailed questions.

Let $\{\varphi_j(x)\}$ be a complete orthonormal family of functions with respect to a given nonnegative weight function $w = w(x)$. Then if the probability density function f is in $L^2(w)$, the Fourier expansion

$$\sum_j c_j \varphi_j(x) \tag{77}$$

with

$$c_j = \int f(x)\overline{\varphi_j(x)}w(x)\,dx \tag{78}$$

will converge to f in $L^2(w)$. Let $F_n(x)$ be the sample distribution based on a sample of n independent observations $X_1, X_2, \cdots, X_n$ from the population with absolutely continuous distribution function $F(x)$ and continuous probability density $f(x) = F'(x)$. It is plausible to estimate c_j by

$$c_j{}^*(n) = \int \overline{\varphi_j(x)}w(x)\,dF_n(x) = n^{-1}\sum_{k=1}^{n} \overline{\varphi_j(X_k)}w(X_k). \tag{79}$$

Then

$$Ec_j{}^* = c_j \tag{80}$$

and

$$\begin{aligned} \operatorname{Cov}(c_j{}^*, c_k{}^*) &= n^{-1}\operatorname{Cov}(\varphi_j(X)w(X), \varphi_k(X)w(X)) \\ &= n^{-1}\{\int \varphi_j(x)\overline{\varphi_k(x)}w^2(x)f(x)\,dx - c_j\bar{c}_k\}. \end{aligned} \tag{81}$$

In all the cases we shall think of, the functions φ_j will be continuous in the domain in which they are defined. Consider a sequence of weights $\{\alpha_j(n)\}$ that are zero except for a finite number of values of j. Let the estimate of $f(x)$ be

$$f_n(x) = \sum_j c_j{}^*(n)\alpha_j(n)\varphi_j(x). \tag{82}$$

The bias

$$Ef_n(x) - f(x) = \sum c_j(\alpha_j(n) - 1)\varphi_j(x) \tag{83}$$

and the covariance

$$(84)\quad \mathrm{Cov}(f_n(x), f_n(y)) = n^{-1} \sum_{j,k} \varphi_j(x)\overline{\varphi_k(y)}\alpha_j(n)\alpha_k(n) \cdot \{\textstyle\int \varphi_j(u)\overline{\varphi_k(u)} w^2(u) f(u)\, du - c_j \bar{c}_k\}.$$

If the Fourier expansion of f converges to f absolutely, one would obviously require that

$$(85)\quad \alpha_j(n) \to 1$$

as $n \to \infty$ for each fixed j. The integrated weighted mean square error has mean value

$$(86)\quad \begin{aligned} E\textstyle\int |f_n(x) - f(x)|^2 w(x)\, dx &= \textstyle\int \sigma^2[f_n(x)] w(x)\, dx + \int |Ef_n(x) - f(x)|^2\, dx \\ &= n^{-1} \textstyle\sum |\alpha_j(n)|^2 \{\int |\alpha_j(u)|^2 w^2(u) f(u)\, du - |c_j|^2\} \\ &\quad + \textstyle\sum |c_j|^2 |\alpha_j(n) - 1|^2. \end{aligned}$$

It is sometimes convenient to rewrite $f_n(x)$ in the form

$$(87)\quad f_n(x) = n^{-1} \textstyle\sum_{j=1}^{n} w(X_j) k_n(x, X_j)$$

where

$$(88)\quad k_n(x, u) = \textstyle\sum_j \alpha_j(n) \varphi_j(x) \overline{\varphi_j(u)}$$

is a "generalized" kernel function (see G. S. Watson [37] for a related discussion). The variance of $f_n(x)$ can then also be written as

$$(89)\quad n^{-1}\{\textstyle\int w^2(u) |k_n(x, u)|^2 f(u)\, du - |\int w(u) k_n(x, u) f(u)\, du|^2\}.$$

The asymptotic behavior of the kernel function $k_n(x, u)$ is obviously of importance in determining the behavior of the variance (89) as $n \to \infty$ as well as the asymptotic distribution of $f_n(x)$.

We consider the case of the trigonometric functions

$$(90)\quad \varphi_j(x) = \frac{e^{ijx}}{(2\pi)^{\frac{1}{2}}}, \qquad -\pi \leqq x \leqq \pi,\ w(x) \equiv 1,$$

because it is simple and is amusingly like the case of spectral estimation. Here

$$(91)\quad k_n(x, u) = k_n(x - u) = \textstyle\sum_j \alpha_j(u)\, e^{ij(x-u)}.$$

For economy in computation one would require $\alpha_j(n) = 0$ for $|j| \geqq n$. In fact, the usual situation will be that in which $\alpha_j(n) = 0$ for $|j| \geqq m(n)$ with $m(n) = o(n)$. Assume that

$$(92)\quad \textstyle\int k_n(x - u)\, du = \alpha_0(n) = 1.$$

For convenience, assume that $k_n(u)$ is a nonnegative weight function that behaves asymptotically like a δ function, that is, given any $\varepsilon > 0$

$$(93)\quad \textstyle\int_{|u| \geqq \varepsilon} k_n(u)\, du \to 0$$

as $n \to \infty$. Asymptotic unbiasedness of the estimate $f_n(x)$ follows since

$$Ef_n(x) = \int k_n(x,u) f(u)\, du \to f(x) \tag{94}$$

as $n \to \infty$ if *f is bounded and continuous*. The variance $\sigma^2[f_n(x)]$ is given asymptotically by

$$\sigma^2[f_n(x)] \cong \frac{f(x)}{n} \int k_n(u)^2\, du. \tag{95}$$

A similar argument shows that

$$E\left|k_n(x,X) - Ek_n(x,X)\right|^4 \cong f(x) \int k_n(u)^4\, du. \tag{96}$$

A sufficient condition for (87) to be asymptotically normal with mean $Ef_n(x)$ and variance (95) as $n \to \infty$ is that

$$\frac{\int k_n(u)^4\, du}{n(\int k_n(u)^2\, du)^2} \to 0 \tag{97}$$

by the Liapounov form of the central limit theorem. Suppose we look at the case of simple truncation.

$$\begin{aligned} \alpha_j(n) &= 1 \qquad \text{if} \quad |j| \leqq m(n) \\ &= 0 \qquad \text{if} \quad |j| > m(n) \end{aligned} \tag{98}$$

or the Fejer weights

$$\begin{aligned} \alpha_j(n) &= 1 - |j|/m \qquad \text{if} \quad |j| \leqq m(n) \\ &= 0 \qquad\qquad\quad\; \text{if} \quad |j| > m(n). \end{aligned} \tag{99}$$

Then

$$\begin{aligned} \int k_n(u)^2\, du &\cong c_1 m(n) \\ \int k_n(u)^4\, du &\cong c_2 m(n)^3 \end{aligned} \tag{100}$$

with c_1, $c_2 > 0$ so that (97) is satisfied if $m(n) = o(n)$ as $n \to \infty$.

Similar computations will now also be carried out for orthonormal polynomials with respect to a weight function. The conditions assumed will be somewhat restrictive but the results obtained are of interest because they suggest what ought to hold under much weaker conditions. Let $w(x)$ be a weight function on $-1 \leqq x \leqq 1$ with $p_j(x)$, $j = 0, 1, \cdots$, the orthonormal polynomials with respect to this weight function. We wish to consider estimating a continuous density function $f(x)$ on $-1 \leqq x \leqq 1$ by (87) with

$$k_n(x,u) = \sum_j \alpha_j(n) p_j(x) p_j(u). \tag{101}$$

Let $x = \cos\theta$, $z = e^{i\theta}$ and

$$h(\theta) = w(\cos\theta)\left|\sin\theta\right|. \tag{102}$$

Assume that w is such that $h(\theta)$ *is a positive continuous function on the unit circle* $z = e^{i\theta}$, $-\theta < \theta \leqq \pi$ ($\theta = -\pi$ is identified with $\theta = \pi$), *and h satisfies a uniform Lipschitz condition.* Let $\phi_n(z)$, $n = 0, 1, 2, \cdots$, with $z = e^{i\theta}$, $-\pi < \theta \leqq \pi$, be the orthonormal trigonometric polynomials with weight function $h(\theta)$ generated from $1, z, z^2, \cdots$. The coefficient of z^n in $\phi_n(z)$ is κ_n. Then by [32, page 292]

$$p_n(x) = (2\pi)^{-\frac{1}{2}}\left\{1 + \frac{\phi_{2n}(0)}{\kappa_{2n}}\right\}^{\frac{1}{2}}\{z^{-n}\phi_{2n}(z) + z^n\phi_{2n}(z^{-1})\} \tag{103}$$

for $n \geqq 1$. Also, if

$$g(h; r\,e^{i\theta}) = g(r\,e^{i\theta}) = \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{i\lambda} + r\,e^{i\theta}}{e^{i\lambda} - r\,e^{i\theta}}\log h(\lambda)\,d\lambda\right\} \tag{104}$$

then by [18, page 51]

$$\lim_{n\to\infty}\kappa_n = [g(o)]^{-1} \tag{105}$$

$$\lim_{n\to\infty}\phi_n(o) = 0.$$

Further, by [18, page 53] one knows that

$$\phi_n(z) = z^n/\overline{g(h; z)} + \frac{\log n}{n}O(1) \tag{106}$$

with the $O(1)$ uniform in n and z. *Assume that* $|\alpha_j(n)| \leqq 1$ with $\alpha_j(n) = 0$ for $|j| \geqq m(n)$ *where* $m(n) = O(n)$ *and*

$$\lim_{n\to\infty}\alpha_j(n) = 1$$

for each fixed j. Then, if the expansion of f in terms of the orthonormal polynomials p_n is absolutely convergent, the estimate $f_n(x)$ will be asymptotically unbiased. To avoid difficulties arising from possible singularities at $x = \pm 1$, we shall *assume that*

$$f(x)(1 - x^2)^{-\frac{1}{2}} \tag{107}$$

is bounded on $-1 \leqq x \leqq 1$. The object is to again determine the asymptotic behavior of the variance of $f_n(x)$ and sufficient conditions for asymptotic normality of $f_n(x)$. From (103) and (106) it follows that

$$\begin{aligned} k_n(x, y) &= \sum \alpha_j(n)p_j(x)p_j(y) \\ &= \frac{1}{2\pi}\sum_j \alpha_j(n)\left(1 + \frac{\phi_{2j}(o)}{\kappa_{2j}}\right)^{-1}\left\{z^j/\overline{g(h; z)} + z^{-j}/\overline{g(h, z^{-1})}\right. \\ &\quad \left. + \frac{\log j}{j}O(1)\right\}\left\{\bar{z}'^j/g(h; z') + z'^{-j}/g(h; z'^{-1}) + \frac{\log j}{j}O(1)\right\} \end{aligned} \tag{108}$$

where $x = \cos\theta$, $z = e^{i\theta}$, $y = \cos\theta'$, $z' = e^{i\theta'}$. Let

$$c_n(\theta) = \sum_{j=1}^{n}\alpha_j(n)\,e^{ij\theta} \tag{109}$$

and assume that

(110) $$|c_n(\theta)| \leqq M(\varepsilon) < \infty \qquad \text{for } |\theta| \geqq \varepsilon.$$

Then

(111) $$\sigma^2[f_n(x)] = \frac{1}{n}\int w^2(u)|k_n(x,u)|^2 f(u)\,du \cong \frac{4}{n}\frac{f(x)}{(1-x^2)^{\frac{1}{2}}}\int_{-\pi}^{\pi}|c_n(\theta)|^2\,d\theta$$

as $n \to \infty$, if $|x \pm 1| \geqq \varepsilon > 0$ and

(112) $$(\log m(n))^2 = o(\textstyle\int |c_n(\theta)|^2\,d\theta)^{\frac{1}{2}}.$$

Further,

(113) $$\textstyle\int w^4(u)|k_n(x,u)|^4 f(u)\,du \leqq A\int_{-\pi}^{\pi}|c_n(\theta)|^4\,d\theta$$

with A a constant, as $n \to \infty$, if $|x \pm 1| \geqq \varepsilon > 0$ and

(114) $$(\log m(n))^2 = o(\textstyle\int |c_n(\theta)|^4 d)^{\frac{1}{4}}.$$

Thus, a sufficient condition for $f_n(x)$ to be asymptotically normally distributed with mean $Ef_n(x)$ and variance $\sigma^2[f_n(x)]$ as $n \to \infty$, given the requirements already specified, is that

(115) $$\frac{\int |c_n(\theta)|^4\,d\theta}{n(\int |c_n(\theta)|^2\,d\theta)^2} \to 0$$

as $n \to \infty$. The condition required on the weight function $w(u)$ is appropriate for the Chebyshev polynomials of the first kind but not for the other classical orthogonal polynomials. For just the Chebyshev polynomials of the first kind one would go through a direct and simple computation to get the results (111) and (115). The interest in deriving (111) and (115) under the conditions we have specified is that it suggests what ought to hold more generally for density estimates based on expansions in orthonormal polynomials on a finite interval. The case of classical orthonormal polynomials other than the Chebyshev polynomials of the first kind should be considered. Notice that asymptotic results like those obtained for expansions in the trigonometric functions can be derived for expansions in terms of the eigenfunctions of certain types of Sturm-Liouville problems by using the remarks made on page 117 of [18].

6. Spectra and stationary processes. There are many extended developments of the theory of stationary processes and discussions of spectral estimation [17], [33]. We shall briefly remark on some of the results that relate or are comparable to those already mentioned on estimation of the probability density.

Let $X(t)$, $t = \cdots, -1, 0, 1, \cdots$, be a discrete time parameter r-vector (column vector) valued stationary process with real-valued components. If second order moments exist, then

(116) $$X(t) = \textstyle\int_{-\pi}^{\pi} e^{it\lambda}\,dZ(\lambda)$$

where Z is an r-vector valued process (with complex-valued components) of orthogonal increments such that

$$E\,dZ(\lambda)\,dZ(\mu)' = \delta(\lambda+\mu)\,dF(\lambda), \qquad -\pi < \lambda, \mu < \pi. \tag{117}$$

In formula (117), δ is the Kronecker δ symbol

$$\begin{aligned} \delta(\lambda) &= 1 \quad \text{if} \quad \lambda = 0 \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

F is an $r \times r$ matrix-valued Hermitian non-decreasing function ($F(\lambda) - F(\mu)$ is positive semidefinite if $\lambda \geqq \mu$), and A' denotes the transpose of A. Since the components of $X(t)$ are real-valued

$$\overline{dZ(-\lambda)} = dZ(\lambda) \tag{118}$$

and $dF(\lambda) = dF(-\lambda)'$. In the case of a Gaussian process, the full probability structure is determined by the first and second order moments. It is convenient to assume that $EX(t) \equiv 0$.

The covariance matrices

$$r(t) = EX(\tau)X(t+\tau)' \tag{119}$$

are related to the second order spectral distribution function F by

$$r(t) = \int_{-\pi}^{\pi} e^{it\lambda}\, dF(\lambda). \tag{120}$$

If the existence of all moments is assumed, we have

$$m_{a_1,\cdots,a_k}(t_1,\cdots,t_k) = EX_{a_1}(t_1)\cdots X_{a_k}(t_k) = m_{a_1,\cdots a_k}(t_1+t,\cdots,t_k+t) \tag{121}$$

where $X_a(t)$ is the ath component of $X(t)$. Assume that the moments have Fourier representations.

$$\begin{aligned} &m_{a_1,\cdots,a_k}(t_1,\cdots,t_k) \\ &\quad = \int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi} \exp\left(i\textstyle\sum_{j=1}^{k} t_j w_j\right) dG_{a_1,\cdots a_k}(w_1,\cdots,w_k)\, dG_{a_1,\cdots,a_k}(w_1,\cdots,w_k) \\ &\quad = E\left(\textstyle\prod_1^k dZ_{a_j}(w_j)\right) \end{aligned} \tag{122}$$

with the functions G of bounded variation. The stationarity of the process $X(t)$ implies that

$$dG = 0 \quad \text{unless} \quad \textstyle\sum_1^k w_j = 0 \quad \text{modulo} \quad 2\pi. \tag{123}$$

The assumed representation (122) can also be written as

$$\begin{aligned} &c_{a_1,\cdots,a_k}(t_1,\cdots,t_k) \\ &\quad = \int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi} \exp\left(i\textstyle\sum_{j=1}^{k} t_j w_j\right) c\{dZ_{a_j}(w_j); j = 1,\cdots,k\}\, dF_{a_1,\cdots,a_k}(w_1,\cdots,w_k) \\ &\quad = c\{dZ_{a_j}(w_j); j = 1,\cdots,k\} \end{aligned} \tag{124}$$

where c is the corresponding cumulant function. It is convenient to assume that the cumulants $c_{a_1,\cdots,a_k}(t_1,\cdots,t_k)$ are in L_1 as functions of their $k-1$ (or k) t arguments (see Condition I in (137)). This can be regarded, if assumed up to order k, as a curious mixing condition of kth order. This then implies that

(125) $$dF_{a_1,\cdots,a_k}(w_1,\cdots,w_k)\eta(\textstyle\sum_1^k w_j) = f_{a_1,\cdots,a_k}(w_1,\cdots,w_k)\eta(\sum_1^k w_j)\,dw_1\cdots dw_k$$

($\eta(w) = 0$ if $w \neq 0$ modulo 2π and is equal to one otherwise) where the cumulant spectral density f is continuous and a function of only $k-1$ variables w since $\sum_1^k w_j = 0$ modulo 2π.

Let us first look at a real-valued stationary process $X(t)$ ($r = 1$) and consider estimating its spectral density of second order. Such an estimate is given by

(126) $$f^{(N)}(\lambda) = \frac{1}{2\pi}\sum_{v=-N}^{N} w_v^{(N)} R_v^{(N)} e^{-iv\lambda}$$

with

(127) $$w_v^{(N)} = \int_{-\pi}^{\pi} e^{iv\lambda} W_N(\lambda)\,d\lambda$$

and

(128) $$R_v^{(N)} = \frac{1}{N}\sum_{t,\tau=1,\,t-\tau=v}^{N} X(t)X(\tau).$$

Formula (126) could also be written as

(129) $$f^{(N)}(\lambda) = \int_{-\pi}^{\pi} W_N(\lambda-\alpha)I^{(N)}(\alpha)\,d\alpha$$

with

(130) $$I^{(N)}(\alpha) = \frac{1}{2\pi N}\left|\sum_{t=1}^{N} X(t)e^{-it\lambda}\right|^2,$$

the periodogram, and it is convenient to assume that

(131) $$W_N(u) = A_N B_N^{-1} W(B_N^{-1}u) \qquad \text{if } |u| \leqq \pi$$

with $B_N \downarrow 0$ as $N \to \infty$ and A_N a normalizing factor so that

(132) $$\int_{-\pi}^{\pi} W_N(u)\,du = 1.$$

Assume W_N is periodically extended with period 2π. Then if W has a finite second moment, is nonnegative and symmetric about zero and $NB_N \to \infty$, we can show that

(133) $$B_N^{-2}[Ef^{(N)}(\lambda) - f(\lambda)] = \int u^2 W(u)\,du \frac{f''(\lambda)}{2} + O(B_N^{-3})$$

and

(134) $$\begin{aligned}\operatorname{Cov}\{f^{(N)}(\lambda), f^{(N)}(\mu)\} &= 2\pi N^{-1}\{\textstyle\int_{-\pi}^{\pi} W_N(\lambda-\alpha)W_N(\mu+\alpha)f^2(\alpha)\,d\alpha \\ &\quad + \textstyle\int_{-\pi}^{\pi} W_N(\lambda-\alpha)W_N(\mu-\alpha)f^2(\alpha)\,d\alpha + O(N^{-1})\}.\end{aligned}$$

By (131) and (134), it follows that

$$\begin{aligned} &\lim_{N\to\infty} B_N N \operatorname{Cov}\{f^{(N)}(\lambda), f^{(N)}(\mu)\} \\ (135) \qquad &= 0 && \text{if } \lambda \neq \mu \\ &= 2\pi f^2(\lambda) \textstyle\int_{-\pi}^{\pi} W^2(u)\,du && \text{if } \lambda = \mu \neq 0, \pi \\ &= 4\pi f^2(\lambda) \textstyle\int_{-\pi}^{\pi} W^2(u)\,du && \text{if } \lambda = \mu = 0, \pi \end{aligned}$$

when $0 \leqq \lambda, \mu \leqq \pi$. If we look at the asymptotic mean square error of estimates, the analysis of Epanechnikov given in Section 2 still holds and one is led to the weight function $W(u)$ given in (26). The weight function $W(u)$ (26) is bandlimited but it leads to a sequence $w_\nu^{(N)}$ that is not bandlimited. But this need not be too frustrating since direct Fourier analysis via the fast Fourier transform (see [10]) may be preferable. One would directly compute the periodogram and then smooth it. The asymptotic results obtained in spectral estimation differ a bit from those obtained in estimating the probability density function. The basic range over which (135) holds is $0 < \lambda < \pi$ with a slightly different result at $\lambda = 0, \pi$. On the other hand, the results obtained in (15) for probability density estimation hold over the entire range of the independent variable x. Also, in the asymptotic formulas for the variance of the estimates, the constant $f(x)$ in (15) is replaced by $2\pi f^2(\lambda)$ in (135). Notice that if we rescale locally in the covariance result (134) for spectral estimation, let $\lambda = \mu + B_N a$ ($\lambda, \mu \neq 0, \pi$) and then normalize appropriately, as $N \to \infty$ one would expect a Gaussian process with covariance function

$$\int W(b+a)W(b)\,db. \tag{136}$$

This is again very much like what was obtained in Section 3 when estimating a probability density function. It would be interesting to look at the functionals of a global character such as (65) mentioned in Section 3, when estimating the spectral density. The applications of second order spectral techniques are legion and the development of theory is due to many people—Bartlett [1], Tukey [33], Grenander and Rosenblatt [17], Parzen [24] and others.

The success of these second order techniques suggested that it might be worthwhile developing ways of estimating higher order spectra for situations in which there are nonlinear or nonGaussian effects. Tukey [33], Van Ness [34], Brillinger and Rosenblatt [4] developed a certain theoretical background. There have been a limited number of applications of which a few are mentioned here—the work of Hasselman, Munk and MacDonald [19] on bispectra of ocean waves, Haubrich [20] on earth noise, Brillinger and Rosenblatt [5] on sunspots, D. Cartwright [8] with a 4th order analysis of tides and surges, and Huber, Kleiner, Gasser and Dumermuth [22] on electroencephalographs. We shall just briefly mention some asymptotic results due to Brillinger and Rosenblatt[4]. Let $c'_{a_1,\dots,a_k}(v_1,\cdots,v_{k-1})$ and $f'_{a_1,\dots,a_k}(\lambda_1,\cdots,\lambda_{k-1})$ be the cumulant function $c_{a_1,\cdots,a_k}$ and cumulant spectral density $f_{a_1,\cdots,a_k}$ looked at as functions of $k-1$ variables. First assume

$$\text{I} \qquad \sum_{v_1,\dots,v_{k-1}=-\infty}^{\infty} \left|v_j c'_{a_1,\dots,a_k}(v_1,\cdots,v_{k-1})\right| < \infty \tag{137}$$

for $j = 1, \cdots, k-1$ and any k-tuple $a_1, \cdots, a_k$ with $k = 2, 3, \cdots$. Condition I implies that $f'_{a_1,\ldots,a_k}(\lambda_1, \cdots, \lambda_{k-1})$ has a bounded uniformly continuous gradient. Let W be a weight function with

(138) $$W(-u_1, \cdots, -u_k) = W(u_1, u_2, \cdots, u_k)$$

and set

(139) $$W_N(u_1, \cdots, u_k) = B_N^{-k+1} W(B_N^{-1}u_1, \cdots, B_N^{-1}u_k).$$

If

(140) $$d_a^{(N)}(\lambda) = \sum_{t=0}^{N-1} X_a(t) \exp(-i\lambda t),$$

then

(141) $$I^{(N)}_{a_1,\ldots,a_k}(\lambda_1, \cdots, \lambda_k) = (2\pi)^{-k+1} N^{-1} \prod_{j=1}^{k} d^{(N)}_{a_j}(\lambda_j),$$

with $\sum_1^k \lambda_j = 0$ modulo 2π, is a kth order analogue of the periodogram. Consider the estimate

(142) $$\begin{aligned} f^{(N)}_{a_1,\ldots,a_k}&(\lambda_1, \cdots, \lambda_k) \\ &= (2\pi)^{k-1} N^{-k+1} \sum_{s_1=-\infty}^{\infty} \cdots \sum_{s_k=-\infty}^{\infty} W_N\left(\lambda_1 - \frac{2\pi s_1}{N}, \cdots, \lambda_k - \frac{2\pi s_k}{N}\right) \\ &\quad \cdot \Phi\left(\frac{2\pi s_1}{N}, \cdots, \frac{2\pi s_k}{N}\right) I^{(N)}_{a_1,\ldots,a_k}\left(\frac{2\pi s_1}{N}, \cdots, \frac{2\pi s_k}{N}\right) \end{aligned}$$

of $f_{a_1,\ldots,a_k}(\lambda_1, \cdots, \lambda_k)$ where it is understood that $\sum_1^k \lambda_j = 0$ modulo 2π. In (142) $\Phi(u_1, \cdots, u_k) = 1$ if $\sum_1^k u_k = 0$ modulo 2π but $\sum_{j \in J} u_j \neq 0$ on any nonvacuous proper subset J of $\{1, \cdots, k\}$. Make the additional assumption

(143) II $$\left|W\left(u_1, \cdots, u_{k-1}, -\sum_1^{k-1} u_j\right)\right|, \left|\frac{\partial}{\partial u_i} W\left(u_1, \cdots, u_{k-1}, -\sum_1^{k-1} u_j\right)\right| \leq A\left(1 + \left(\sum_1^{k-1} u_j^2\right)^{\frac{1}{2}}\right)^{-(k+\varepsilon-1)}$$

for $i = 1, \cdots, k-1$ with $A, \varepsilon > 0$.

Under assumptions I and II

(144) $$\begin{aligned} \operatorname{Cov}&\{f^{(N)}_{a_1,\ldots,a_k}(\lambda_1, \cdots, \lambda_k), f^{(N)}_{a_1',\ldots,a_k'}(\mu_1, \cdots, \mu_k)\} \\ &= 2\pi N^{-1} \sum_P \int_{-\infty}^{\infty} \cdots \int W_N(\lambda_1 - \alpha_1, \cdots, \lambda_k - \alpha_k) \\ &\quad \cdot W_N(\mu_1 + \beta_1, \cdots, \mu_k + \beta_k) \prod_1^k \eta(\alpha_j + \beta_{P(j)}) \eta(\textstyle\sum_1^k \alpha_j) \\ &\quad \cdot \prod_1^k f_{a_j, a'_{P(j)}}(\alpha_j)\, d\alpha_1 \cdots d\alpha_k\, d\beta_1 \cdots d\beta_k + O(B_N^{-k+2} N^{-1}) \end{aligned}$$

if $B_N^{k-1} N \to \infty$ as $B_N \to 0$ and $N \to \infty$. The summation in (144) is over all permutations P on the integers $1, \cdots, k$ and it is understood that

$$\sum_1^k \lambda_j = \sum_1^k \mu_j = 0 \quad \text{modulo} \quad 2\pi.$$

Let $X(t)$ be a stationary process satisfying assumption I. Assume that ${}_jf^{(N)}$, $j = 1, \cdots, b$, are spectral estimates of order $k_1 \leqq \cdots \leqq k_b$ whose weight functions satisfy assumption II. Let the bandwidths $B_N{}^{(j)}$ be such that

$$B_N{}^{(j)} \to 0, \qquad (B_N{}^{(j)})^{k_j-1} N \to \infty \tag{145}$$

as $N \to \infty$ with $B_N{}^{(1)} \leqq \cdots \leqq B_N{}^{(b)}$. Bandwidths of estimates of the same order are assumed equal. Then, the estimates are asymptotically jointly normal as $N \to \infty$, with estimates of different order asymptotically independent, and estimates of the same order having

$$\begin{aligned} \lim_{N\to\infty} B_N{}^{k-1} N \operatorname{Cov}\left[f^{(N)}_{a_1,\cdots,a_k}(\lambda_1,\cdots,\lambda_k), f^{(N)}_{a_1',\cdots,a_k'}(\mu_1,\cdots,\mu_k)\right] \\ = 2\pi \textstyle\sum_P \eta_{\lambda_1-\mu_{P(1)}} \cdots \eta_{\lambda_k-\mu_{P(k)}} f_{a_1,a'_{P(1)}}(\lambda_1) \cdots f_{a_k,a'_{P(k)}}(\lambda_k) \\ \cdot \textstyle\int_{-\infty}^{\infty} \cdots \int W(\tau_1,\cdots,\tau_k) W(\tau_{P(1)},\cdots,\tau_{P(k)}) \delta\left(\sum_1^k \tau_j\right) d\tau_1 \cdots d\tau_k. \end{aligned} \tag{146}$$

For a detailed discussion and interpretation of these results, [4] and [5] should be referred to. There are some marked differences between the case of spectral estimates and probability density estimates. In the higher order spectral case, asymptotic variances have products of 2nd order spectra in their principal term while in the case of probability density estimates, the multivariate density itself (see (34)) itself comes in. The stationarity and the real-valued character of the components of the time series imply that $\sum_1^k w_j = 0$ and $f(w_1, w_2, \cdots, w_k) = f(-w_1, -w_2, \cdots, -w_k)$ so that there are additional symmetries or restraints. This also leads to inhomogeneity on lower dimensional manifolds in w space unless one is careful. Related ideas are used by Brillinger [6] in a discussion of the spectral analysis of point processes.

Acknowledgment. I should like to thank colleagues who were kind enough to read the manuscript and send on their comments; and my thanks to Lillian Johnson for typing it.

REFERENCES

[1] BARTLETT, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika* **37** 1–16.

[2] BARTLETT, M. S. (1963). Statistical estimation of density functions. *Sankhyā, Ser. A* **25** 245–254.

[3] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[4] BRILLINGER, D. R. and ROSENBLATT, M. (1967). Asymptotic theory of estimates of kth order spectra in *Advanced Seminar on Spectral Analysis of Time Series*, ed. B. Harris. 153–188.

[5] BRILLINGER, D. R. and ROSENBLATT, M. (1967). Computation and interpretation of kth order spectra in *Advanced Seminar on Spectral Analysis of Time Series*, ed. B. Harris. 189–232.

[6] BRILLINGER, D. R. (1970). The spectral analysis of stationary interval functions. Proc. Sixth Berkeley Symp. *Math. Statist. Prob.* To appear.

[7] CACOULLOS, T. (1966). Estimation of a multivariate density. *Ann. Inst. Statist. Math.* **18** 179–189.

[8] CARTWRIGHT, D. E. (1968). A unified analysis of tides and surges round North and East Britain. *Philos. Trans. Roy. Soc. London Ser. A* **263** 1–55.

[9] CENCOV, N. N. (1962). Evaluation of an unknown distribution density from observations. *Soviet Math.* **3** 1559–1562.

[10] COOLEY, J. W. and TUKEY, J. W. (1965). An algorithm for the machine computation of complex Fourier Series. *Math. Comp.* **19** 297–301.

[11] CRAMÉR, H. and LEADBETTER, M. R. (1967). *Stationary and Related Processes.* Wiley, New York.

[12] DANIELS, H. (1964). Saddlepoint approximations in statistics. *Ann. Math. Statist.* **3** 631–650.

[13] EPANECHNIKOV, V.A. (1969). Nonparametric estimates of a multivariate probability density. *Theor. Probability Appl.* **14** 153–158.

[14] FARREL, R. H. (1967). On the lack of a uniformly consistent sequence of estimators of a density function. *Ann. Math. Statist.* **38** 471–475.

[15] FRENKIEL, F. N. and KLEBANOFF, P. S. (1965). Two-dimensional probability distribution in a turbulent field. *Phys. Fluids* **8** 2291–2293.

[16] GARSIA, A. M., RODEMICH, E. and RUMSEY, H. Jr. (1970). A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.* **20** 565–578.

[17] GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical Analysis of Stationary Time Series.* Wiley, New York.

[18] GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz Forms and Their Applications.* Univ. of California Press.

[19] HASSELMAN, K., MUNK, W. and MACDONALD, G. (1963). Bispectra of ocean waves in *Time Series Analysis*, ed. M. Rosenblatt. Wiley, New York.

[20] HAUBRICH, R A. (1965). Earth noise, 5 to 500 millicycles per second. *J. Geophys. Res.* **70** 1415–1427.

[21] HODGES, J. L. Jr. and LEHMANN, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

[22] HUBER, P. J., KLEINER, B., GASSER, TH. and DUMERMUTH, G. (1971). Statistical methods for investigating phase relations in stationary stochastic processes. To appear in the *IEEE Trans. Audio and Electroacoustics.*

[23] KOLMOGOROV, A. (1963). On the approximation of distributions of sums of independent summands by infinitely divisible distributions. *Sankhyā Ser. A* **25** 159–174.

[24] PARZEN, E. (1961). Mathematical considerations in the estimation of spectra. *Technometrics* **3** 167–190.

[25] PARZEN, E. (1962). On the estimation of a probability density and mode. *Ann. Math. Statist.* **33** 1065–1076.

[26] ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27** 832–835.

[27] ROSENBLATT, M. (1959). Statistical analysis of stochastic processes with stationary residuals in *H. Cramér Volume*, ed. U. Grenander. 246–275.

[28] ROSENBLATT, M. (1969). Conditional probability density and regression estimates, in *Multivariate Analysis*, ed. Krishnaiah. 25–31.

[29] ROSENBLATT, M. (1970). Density estimates and Markov sequences, in *Nonparametric Techniques in Statistical Inference*, ed. M. Puri. 199–210.

[30] ROUSSAS, G. (1969). Nonparametric estimation in Markov processes. *Ann. Inst. Statist. Math.* **21** 73–87.

[31] SAZANOV, V. V. (1968). On the multidimensional central limit theorem. *Sankhyā Ser. A* **30** 181–204.

[32] SZEGÖ, G.(1959). *Orthogonal Polynomials.* American Mathematical Society, Providence.

[33] TUKEY, J. W. (1959). An introduction to the measurement of spectra, in *H. Cramér Volume*, ed. U. Grenander. 300–330.

[34] VAN ATTA, C. W. and CHEN, W. Y. (1968). Correlation measurements in grid turbulence using digital harmonic analysis. *J. Fluid Mech.* **34** 497–515.

[35] VAN NESS, J. (1966). Asymptotic normality of bispectral estimates. *Ann. Math. Statist.* **37** 1257–1272.

[36] VAN RYZIN, J. (1970). On a histogram method of density estimation. Technical Report No. 226 Dept. of Statistics, Univ. of Wisconsin.

[37] WATSON, G. S. (1969). Density estimation by orthogonal series. *Ann. Math. Statist.* **40** 1496–1498.

[38] WOODROOFE, M. (1967). On the maximum deviation of the sample density. *Ann. Math. Statist.* **38** 475–481.

[39] YAGLOM, A. M. (1966). The influence of fluctuations in energy dissipation on the shape of turbulence characteristics in the inertial interval. *Soviet Physics Dokl.* **11** 26–29.