

## ON SOME GLOBAL MEASURES OF THE DEVIATIONS OF DENSITY FUNCTION ESTIMATES

BY P. J. BICKEL<sup>1</sup> AND M. ROSENBLATT<sup>2</sup>

*University of California, Berkeley;*

*University of California, San Diego*

We consider density estimates of the usual type generated by a weight function. Limit theorems are obtained for the maximum of the normalized deviation of the estimate from its expected value, and for quadratic norms of the same quantity. Using these results we study the behavior of tests of goodness-of-fit and confidence regions based on these statistics. In particular, we obtain a procedure which uniformly improves the chi-square goodness-of-fit test when the number of observations and cells is large and yet remains insensitive to the estimation of nuisance parameters. A new limit theorem for the maximum absolute value of a type of nonstationary Gaussian process is also proved.

**1. Introduction.** Let  $X_1, X_2, \dots, X_n$  be independent and identically distributed random variables with continuous density function  $f(x)$ . By now there are a goodly number of papers on estimation of the density function (see [13] for a bibliography). The class of estimates  $f_n(x)$  that we consider are determined by a bounded integrable weight function  $w$ ,

$$(1.1) \quad \begin{aligned} f_n(x) &= \frac{1}{nb(n)} \sum_{j=1}^n w\left(\frac{x - X_j}{b(n)}\right) \\ &= \int \frac{1}{b(n)} w\left(\frac{x - s}{b(n)}\right) dF_n(s). \end{aligned}$$

In formula (1.1),  $F_n$  is the sample distribution function. Also  $b(n)$  is a bandwidth that tends to zero as  $n \rightarrow \infty$  but is such that  $n^{-1} = o(b(n))$ .

The local properties of such estimates have been discussed extensively. Our object will be to get global measures of how good  $f_n(x)$  is as an estimate of  $f(x)$ . In particular, the asymptotic distribution of the functionals

$$\max_{0 \leq x \leq 1} |f_n(x) - f(x)|/(f(x))^{1/2}$$

and

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

are evaluated under appropriate conditions as  $n \rightarrow \infty$ .

---

Received November 1971; revised March 1973.

<sup>1</sup> This work was performed while the author was a fellow of the John Simon Guggenheim Memorial Foundation and supported in part by the Office of Naval Research under Contract N00014-69-A0200-1038.

<sup>2</sup> This work was carried out while the author was a fellow of the John Simon Guggenheim Memorial Foundation and supported in part by the Office of Naval Research.

**Key words and phrases.** Density function estimates, weight function, bandwidth, global measure of deviation, asymptotic distribution, functionals of stochastic process, Gaussian process.

We shall state two results, one concerned with absolute deviation of the estimate  $f_n(x)$  from  $f(x)$  and the other with integrated quadratic deviation. They will give some insight into the type of result that is obtained. However, in order to give the result on absolute deviation it is convenient to introduce at this point certain convenient assumptions which we shall refer to as A1, A2, A3, and A4.

A1. The weight function  $w$  also assigns mass one to the line and either (a) vanishes outside an interval  $[-A, A]$  and is absolutely continuous on  $[-A, A]$  with derivative  $w'$  or (b) is absolutely continuous on  $(-\infty, \infty)$  with derivative  $w'$  such that  $\int |w'(t)|^k dt < \infty$ ,  $k = 1, 2$ .

A2. The density  $f$  is continuous, positive and bounded.

A3. The function  $f^{\frac{1}{2}}$  is absolutely continuous and its derivative  $\frac{1}{2}f'/f^{\frac{1}{2}}$  is bounded in absolute value. Moreover,

$$\int_{|z| \geq 3} |z|^{\frac{3}{2}} [\log \log |z|]^{\frac{1}{2}} [|w'(z)| + |w(z)|] dz < \infty.$$

A4. The second derivative  $f''$  of  $f$  exists and is bounded. Moreover  $w$  is symmetric (about 0) and  $z^2 w(z)$  is integrable.

We shall simply state a corollary of a main result on absolute deviations which is appealing because it is phrased in a form that is convenient if one wishes to set up a confidence band for the density function.

**COROLLARY.** *Let assumptions A1—A4 be satisfied with  $b(n) = n^{-\delta}$ ,  $\frac{1}{5} < \delta < \frac{1}{2}$ . Then*

$$(1.2) \quad \begin{aligned} \lim_{n \rightarrow \infty} P \left[ f_n(x) - \left( \frac{f_n(x)\lambda(w)}{nb(n)} \right)^{\frac{1}{2}} \left( \frac{z}{(2\delta \log n)^{\frac{1}{2}}} + d_n \right) \right. \\ \left. \leq f(x) \leq f_n(x) + \left( \frac{f_n(x)\lambda(w)}{nb(n)} \right)^{\frac{1}{2}} \left( \frac{z}{(2\delta \log n)^{\frac{1}{2}}} + d_n \right) \text{ for all } 0 \leq x \leq 1 \right] \\ = e^{-2e^{-z}} \end{aligned}$$

where

$$\lambda(w) = \int w^2(t) dt$$

and

$$d_n = (2\delta \log n)^{\frac{1}{2}} + \frac{1}{(2\delta \log n)^{\frac{1}{2}}} \left\{ \log \left( \frac{K_1(w)}{\pi^{\frac{1}{2}}} \right) + \frac{1}{2} [\log \delta + \log \log n] \right\}$$

if (a) of A1 holds and

$$K_1(w) = \frac{w^2(A) + w^2(-A)}{2} / \lambda(w) > 0,$$

and otherwise

$$d_n = (2\delta \log n)^{\frac{1}{2}} + \frac{1}{(2\delta \log n)^{\frac{1}{2}}} \left[ \log \frac{1}{\pi} \left( \frac{K_2(w)}{2} \right)^{\frac{1}{2}} \right]$$

with

$$K_2(w) = \frac{1}{2} \int_{-\infty}^{\infty} [w'(t)]^2 dt / \lambda(w).$$

The following result for a quadratic functional is also of some interest. The function  $a(x)$  used in the theorem is assumed to be a bounded piece-wise smooth integrable function.

**THEOREM.** *Let A1—A4 hold. Then if  $b(n) = o(n^{-\frac{1}{3}})$ ,  $n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{2}} = o(b(n))$  as  $n \rightarrow \infty$ ,*

$$b(n)^{-\frac{1}{2}}[nb(n) \int [f_n(x) - f(x)]^2 a(x) dx - \int f(x)a(x) dx \int w^2(z) dz]$$

*is asymptotically normally distributed with mean zero and variance*

$$2 \int (\int w(x+y)w(x) dx)^2 dy \int a^2(x)f^2(x) dx$$

*as  $n \rightarrow \infty$ .*

The basic technique in obtaining the results is that of approximating the normalized and centered sample distribution function by an appropriate Brownian motion process on a convenient probability space by using a Skorohod-like imbedding due to Brillinger and Breiman. The details of this approximation and remarks on approximation of other functionals are given in Section 2. The asymptotic theory of the maximal deviation and that of quadratic deviations are developed in Sections 3 and 4 respectively. Some computations on the power of these procedures are also carried out. In particular, we show that a goodness-of-fit test based on a quadratic functional is strictly better than the  $\chi^2$  test. There is also an appendix on the asymptotic distribution of the maximal deviation for a type of nonstationary Gaussian process.

**2. Approximations.** As has been indicated in the introduction our technique is to consider the statistics of interest as functionals of certain stochastic processes on the interval  $[0, 1]$  and then to substitute Gaussian processes with the same covariance structure for the latter where possible.

It is convenient to introduce  $Z_n^0(\cdot)$  given by

$$(2.1) \quad Z_n^0(t) = n^{\frac{1}{2}}(F_n^*(t) - t), \quad 0 \leq t \leq 1$$

where  $F_n^* = F_n(F^{-1})$  is the empirical distribution of  $F(X_1), \dots, F(X_n)$ . This will be approximated by  $Z^0(\cdot)$ , the Brownian bridge, given by

$$(2.2) \quad Z^0(t) = Z(t) - tZ(1)$$

where  $Z$  is a standard Wiener process on  $[0, 1]$ .

The process  $[nb(n)f^{-1}(t)]^{\frac{1}{2}}(f_n(\cdot) - E(f_n(\cdot)))$  is central to our discussion. It can be written as

$$(2.3) \quad Y_n(t) = b^{-\frac{1}{2}}(n)f^{-\frac{1}{2}}(t) \int_{-\infty}^{\infty} w\left(\frac{t-s}{b(n)}\right) dZ_n^0(F(s)).$$

Approximations  ${}_0Y_n$  and  ${}_1Y_n$  to this process are obtained by substituting  $Z^0(F(\cdot))$  and  $Z(F(\cdot))$  respectively for the random measure in (2.3). The resulting processes are well defined, at least if  $\int_{-\infty}^{\infty} w^2(t) dF(t) < \infty$ .

Two other processes which also arise naturally are given by

$$(2.4) \quad {}_2Y_n(t) = [b(n)f(t)]^{-\frac{1}{2}} \int w\left(\frac{t-s}{b(n)}\right) (f(s))^{\frac{1}{2}} dZ(s)$$

and

$$(2.5) \quad {}_3Y_n(t) = [b(n)]^{-\frac{1}{2}} \int w\left(\frac{t-s}{b(n)}\right) dZ(s)$$

where  $Z$  is a two-sided Wiener process on  $(-\infty, \infty)$  ( $dZ$  is Wiener measure). The process  ${}_3Y_n$  is well defined if  $\int w^2(t) dt < \infty$ , and all the integrals with respect to  $dZ^0(F(\cdot))$ ,  $dZ(F(\cdot))$ ,  $dZ(\cdot)$ ,  $dZ^0(\cdot)$  are taken in the  $L^2$  sense (cf. Doob [6] page 426). For convenience, suppose all our processes are realized as random elements taking their values in the space  $D[0, 1]$  (cf. [3]). For  $x \in D[0, 1]$  let  $\|x\| = \sup\{|x(t)| : 0 \leq t \leq 1\}$ . Our approximations rest on the following theorem of Brillinger (1969). (A similar argument appeared simultaneously in Breiman 1969.)

**THEOREM.** *There exists a probability space  $(\Omega, \mathcal{A}, P)$  on which one can construct versions of  $Z_n^0$  and  $Z$  such that*

$$(2.6) \quad \|Z_n^0 - Z^0\| = O_p(n^{-\frac{1}{4}}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{4}}).$$

From this we can derive

**PROPOSITION 2.1.** *If the processes  $Z_n^0$ ,  $Z^0$  are constructed as above and A1 and A2 hold, then*

$$(2.7) \quad \|Y_n - {}_0Y_n\| = O_p(b^{-\frac{1}{2}}(n)n^{-\frac{1}{4}}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{4}}).$$

**PROOF.** Write, using A1,

$$(2.8) \quad \begin{aligned} Y_n(q) &= [b(n)f(q)]^{-\frac{1}{2}}\{-w(A)Z_n^0(F(q - Ab(n))) \\ &\quad + w(-A)Z_n^0(F(q + Ab(n)))\} \\ &\quad + b^{-\frac{1}{2}}(n)f^{-\frac{1}{2}}(q) \int_{-\infty}^{\infty} Z_n^0(F(s))w'\left(\frac{q-s}{b(n)}\right) ds. \end{aligned}$$

(The first two terms inside the curly brackets are taken to be 0 in the event A1(b) holds but A1(a) does not.) A similar representation is valid for  ${}_0Y_n$  and (2.7) follows.

**PROPOSITION 2.2.** *If A2 holds then*

$$(2.9) \quad \|{}_0Y_n - {}_1Y_n\| = O_p(b^{\frac{1}{2}}(n)).$$

*If A2 and A3 hold then*

$$(2.10) \quad \|{}_2Y_n - {}_3Y_n\| = O_p(b^{\frac{1}{2}}(n)).$$

**PROOF.** From the representation (2.2),

$$(2.11) \quad \begin{aligned} |{}_0Y_n(q) - {}_1Y_n(q)| &= |Z(1)|[b(n)f(q)]^{-\frac{1}{2}} \\ &\quad \int w\left(\frac{q-s}{b(n)}\right)f(s) ds = |Z(1)|O(b^{\frac{1}{2}}(n)). \end{aligned}$$

Applying (2.8) and its analogues, if A1(a) holds,

$$\begin{aligned}
 & |{}_2 Y_n(q) - {}_3 Y_n(q)| \\
 & \leq b^{-\frac{1}{2}}(n) \{ [ |Z(Ab(n) + q)| |[f(Ab(n) + q)/f(q)]^{\frac{1}{2}} - 1| \\
 (2.12) \quad & + |Z(-Ab(n) + q)| |[f(-Ab(n) + q)/f(q)]^{\frac{1}{2}} - 1| ] \sup_t |w(t)| \\
 & + \int |Z(sb(n) + q)| |[f(q + sb(n))/f(q)]^{\frac{1}{2}} - 1| |w'(s)| ds \\
 & + \frac{1}{2}(b(n)) \int |Z(sb(n) + q)| |f'(q + sb(n))| [f(q)f(q + sb(n))]^{-\frac{1}{2}} |w(s)| ds \} \\
 & = O_p(b^{\frac{1}{2}}(n))
 \end{aligned}$$

by using A3 and the law of the iterated logarithm for the Wiener process. If A1(b) holds the first two terms vanish and the same argument applies.

To apply these propositions we make the elementary

REMARK. If  $\{g_n\}$  is a sequence of functionals on  $D[0, 1]$  satisfying Lipschitz conditions such that

$$(2.13) \quad |g_n(x) - g_n(y)| \leq M_n \|x - y\|$$

and  $A_n, B_n$  are stochastic processes realizable in  $D$  such that  $\|A_n - B_n\| = o_p(1/M_n)$ , then  $g_n(A_n)$  converges in law if and only if  $g_n(B_n)$  does, and to the same limit.

We shall apply this proposition in the next two sections to the functionals

$$\text{I} \quad (2|\log b(n)|)^{\frac{1}{2}} \left[ \max \left\{ \frac{|Y_n(t)|}{(\lambda(w))^{\frac{1}{2}}} : 0 \leq t \leq 1 \right\} - B([b(n)]^{-1}) \right]$$

where  $B$  is defined in Theorem A1 and,

$$\text{II} \quad b^{-\frac{1}{2}}(n) \left[ \int_{-\infty}^{\infty} Y_n^2(t) f(t) a(t) dt - \int_{-\infty}^{\infty} w^2(t) dt \right]$$

where  $a$  is an integrable weight function. Evidently, since  ${}_1 Y_n$  and  ${}_2 Y_n$  have the same joint laws, we can substitute  ${}_3 Y_n$  for  $Y_n$  in I if A1—A3 hold and

$$(2.14) \quad o\left(\frac{b(n)}{|\log b(n)|}\right) = n^{-\frac{1}{2}} \log n (\log \log n)^{\frac{1}{2}}$$

and  ${}_0 Y_n$  can be substituted for  $Y_n$  in II if A1 and A2 hold and,

$$(2.15) \quad o(b(n)) = n^{-\frac{1}{2}} (\log n)^{\frac{1}{2}} (\log \log n)^{\frac{1}{2}}.$$

Although we do not pursue this it is clear that the same technique can be applied to other functionals, e.g., a normalized version of the total time in  $[0, 1]$  spent by  $Y_n$  above a high level (cf. Berman (1971) [2]).

**3. The maximum absolute deviation.** The first measure of global deviation that we consider is  $\tilde{M}_n = \max \{|Y_n(t)| : 0 \leq t \leq 1\}$ . (There is no loss in considering  $[0, 1]$  rather than any other interval on which the density is bounded away from 0 and  $\infty$ .) The statistical interest of this functional is twofold as

- (i) A convenient way of getting a confidence band for  $f$ .
- (ii) A test statistic for the hypothesis  $H: f = f_0$ .

Under (ii) we shall also consider the possibility of testing composite hypotheses, for example, that  $f$  is Gaussian. The asymptotic theorem we need to discuss (i), and the behavior of (ii) under the null hypothesis is a consequence of our remarks in Section 2 and Theorem A1 of the appendix.

**THEOREM 3.1.** *Let  $w$  satisfy assumptions A1—A3 and*

$$b(n) = n^{-\delta}, \quad 0 < \delta < \frac{1}{2}.$$

*Then,*

$$(3.1) \quad P \left[ (2\delta \log n)^{\frac{1}{2}} \left( \frac{\tilde{M}_n}{(\lambda(w))^{\frac{1}{2}}} - d_n \right) < x \right] \rightarrow e^{-2e^{-x}},$$

*where*

$$(3.2) \quad \lambda(w) = \int w^2(t) dt,$$

*and*

$$(3.3) \quad d_n = (2\delta \log n)^{\frac{1}{2}} + \frac{1}{(2\delta \log n)^{\frac{1}{2}}} \left\{ \log \frac{K_1(w)}{\pi^{\frac{1}{2}}} - \frac{1}{2} [\log \delta + \log \log n] \right\}$$

*where*

$$K_1(w) = \frac{w^2(A) + w^2(-A)}{2} / \lambda(w),$$

*if  $K_1(w) > 0$ , and otherwise*

$$d_n = (2\delta \log n)^{\frac{1}{2}} + \frac{1}{(2\delta \log n)^{\frac{1}{2}}} \left[ \log \frac{1}{\pi} \frac{K_2(w)}{2} \right]$$

*where*

$$K_2(w) = \frac{1}{2} [\int [w'(t)]^2 dt] / \lambda(w).$$

**REMARK 1.** The natural weight function  $w(t) = \frac{1}{2}, |t| \leq 1, = 0$  otherwise, falls under the first case, while the “optimal” weight function of Epanechnikov (1969)  $w(t) = 3/(4(5)^{\frac{1}{2}})(1 - (v^2/5))$  if  $|v| \leq 5^{\frac{1}{2}}$ ,  $= 0$  otherwise, falls under the second.

**REMARK 2.** A similar result holds if one considers the maximum deviation (rather than absolute deviation) of a density function estimate as in Rosenblatt (1971). However, since one-sided deviations for density functions are unnatural the present result seems more interesting.

**REMARK 3.** The techniques of proof of this result may readily be adapted to prove limit theorems such as that of Woodroffe (1967) for the maximum deviation observed at an increasing finite number of points.

**PROOF.** It follows from Propositions 2.1 and 2.2 and the following remark that the limiting behavior of  $(2\delta \log n)^{\frac{1}{2}}[(\tilde{M}_n/(\lambda(w))^{\frac{1}{2}}) - d_n]$  is the same as that of  $(2 \log b(n))^{\frac{1}{2}}(\max \{ |Y_n(t)|/(\lambda(w))^{\frac{1}{2}} : 0 \leq t \leq 1 \} - d_n)$ . By the similarity transform for the Wiener process, the law

$$(3.4) \quad L(\int_0^1 Y_n(t) dt : 0 \leq t \leq 1) = L\left(\int_0^1 w\left(\frac{t}{b(n)} - s\right) dZ(s) : 0 \leq t \leq 1\right).$$

Since  $1/(\lambda(w))^{1/2} \int w(t-s) dZ(s)$  is a stationary Gaussian process with mean 0 and covariance

$$(3.5) \quad r(t) = \frac{\int w(s+t)w(s) ds}{\lambda(w)},$$

Theorem 3.1 follows from Corollary A.1 provided we show that  $r$  satisfies condition (v) and (vi) of Theorem A1 with  $\alpha = 1, 2$ . That (v) is satisfied is equivalent to Theorem B1. Moreover,

$$(3.6) \quad \int r^2(t) dt = \frac{1}{2\pi} \int |\hat{r}(t)|^2 dt = \frac{1}{2\pi\lambda^2(w)} \int |\hat{w}(t)|^4 dt$$

where  $\hat{\cdot}$  denotes Fourier transformation. Since  $w$  is integrable and bounded  $\hat{w}$  is square integrable and bounded and (vi) must hold.

APPLICATIONS. (i) To obtain a confidence band for  $f$  that is simple and explicit it is natural to consider  $\delta$  such that  $E(f_n)$  can be replaced by  $f$ . This is true if  $\delta > \frac{1}{5}$  and A4 holds. Then,

$$(3.7) \quad \frac{1}{b(n)} \int w\left(\frac{t-s}{b(n)}\right) f(s) ds = f(t) + O(b^2(n))$$

with 0 independent of  $t$ . If we now define  $Y_n^*$  by replacing  $E(f_n(t))$  with  $f(t)$  we conclude that

$$(3.8) \quad \|Y_n - Y_n^*\| = O([nb^5(n)]^{1/2}).$$

Using the usual approximations we conclude that  $\max\{|Y_n^*(t)| : 0 \leq t \leq 1\}$  behaves like  $\tilde{M}_n$  if A4 holds and  $\delta > \frac{1}{5}$ . In this case inverting as usual we obtain the confidence band

$$(3.9) \quad \begin{aligned} f &\leq f_n + \left(\frac{f_n}{nb(n)}\right)^{1/2} c(\alpha) \left(1 + \frac{c^2(\alpha)}{4nb(n)f_n}\right)^{1/2} + \frac{c^2(\alpha)}{2nb(n)} \\ f &\geq f_n - \left(\frac{f_n}{nb(n)}\right)^{1/2} c(\alpha) \left(1 + \frac{c^2(\alpha)}{4nb(n)f_n}\right)^{1/2} + \frac{c^2(\alpha)}{2nb(n)} \end{aligned}$$

where  $c(\alpha)$  is given by (3.11). A simpler band is obtained if we further substitute  $f_n$  for  $f$  in the denominator of  $Y_n^*$ . The resulting process  $Y_n^{**}$  (say) has

$$(3.10) \quad \begin{aligned} \|Y_n^* - Y_n^{**}\| &= O_p\left(\frac{\|Y_n^*\|^2}{(nb(n))^{1/2}} \|f_n^{-1/2}\|\right) \\ &= O_p\left(\frac{\log n}{(nb(n))^{1/2}}\right) \end{aligned}$$

if A1—A4 hold and  $\frac{1}{5} < \delta < \frac{1}{2}$ . The approximate confidence band obtained by looking at the maximum of  $|Y_n^{**}|$  is given in the introduction (1.2).

There is no choice of  $\delta$  which asymptotically makes this simple band as thin as possible, i.e. one should choose  $\delta$  as small as possible. This of course ignores the obvious—the speed with which bias disappears asymptotically depends on  $\delta$  as does the speed of convergence to the asymptote. However, for fixed  $n$  there

is an optimal  $\delta(n)$  (depending on  $\alpha$ )  $> 0$  which for moderate  $n$  and small  $\alpha$  may be the right thing to use if the choice of bandwidth is free.

(ii) To test  $H: f = f_0$  it is natural to compute  $\tilde{M}_n$  with  $f = f_0$  and reject for large values of the statistic. According to the theorem to obtain approximate level  $\alpha$  we should use as cutoff point,

$$(3.11) \quad c(\alpha) = -[\log |\log(1 - \alpha)| - \log 2] \frac{(\lambda(w))^{\frac{1}{2}}}{(2\delta \log n)^{\frac{1}{2}}} + d_n(\lambda(w))^{\frac{1}{2}}.$$

Under some assumptions the same cutoff point may be used for testing composite hypotheses of the form  $H: f = f_0(\cdot, \theta)$  where  $\theta$  is an unknown vector parameter by using  $\tilde{M}_n$  with an estimate  $\hat{\theta}$  substituted for the unknown parameter  $\theta$ . We need the following assumption.

A5. The estimate  $\hat{\theta}$  is such that if  $\theta = \theta_0$ , for every  $\theta_0$ ,

$$(3.12) \quad \sup \{|\int [f_0(t + sb(n), \hat{\theta}) - f_0(t + sb(n), \theta_0)]w(s) ds| : 0 \leq t \leq 1\} = o_p([nb(n) \log b(n)]^{-\frac{1}{2}})$$

and

$$\|f_0(\cdot, \theta_0) - f_0(\cdot, \hat{\theta})\| = o_p(|\log b(n)|^{-\frac{1}{2}}).$$

Typically for maximum likelihood and method of moments estimates

$$(3.13) \quad |\hat{\theta} - \theta_0| = O_p(n^{-\frac{1}{2}}).$$

If, moreover,  $\theta = (\theta^{(1)}, \dots, \theta^{(k)})$ ,  $\partial f_0 / \partial \theta^{(j)}$  is bounded for  $\theta$  in a neighborhood of  $\theta_0$ , all  $x$ , and  $1 \leq j \leq k$ , it is easy to see that A5 holds. To see that A5 is the needed assumption again introduce a process  $\bar{Y}_n$  with  $E_\theta(f_n)$  replaced by  $E_{\hat{\theta}}(f_n)$  and  $(f(\cdot, \theta))^{\frac{1}{2}}$  replaced in the denominator of  $Y_n$  by  $(f(\cdot, \hat{\theta}))^{\frac{1}{2}}$ . Then

$$(3.14) \quad \|Y_n - \bar{Y}_n\| = o_p(|\log b(n)|^{-\frac{1}{2}})$$

and the result follows.

To make local power calculations on the test of the simple hypothesis described above we need to consider the behavior of  $\tilde{M}_n$  (calculated under  $f_0$ ) for a sequence of alternatives of the form,

$$(3.15) \quad g_n(x) = f_0(x) + \gamma_n \eta(x) + o(\gamma_n)$$

where  $g_n$  satisfy A2—A3 uniformly in  $n$ ,  $\gamma_n \downarrow 0$  at a suitable rate, and  $o(\gamma_n)$  is uniform in  $x$  on  $[0, 1]$ . (Note that  $\eta$  must be continuous on  $[0, 1]$ .) Denote probabilities calculated under  $g_n$  by  $P_n$ . Our basic result is,

**THEOREM 3.2.** *Suppose that  $g_n$  are as above. Let  $w$  satisfy A1—A3 and define  $\tilde{M}_n$  in terms of  $f_0$ . Let*

$$\gamma_n = n^{-\frac{1}{2} + \delta/2} [2\delta \log n]^{-\frac{1}{2}}.$$

*Then,*

$$(3.16) \quad P_n \left[ (2\delta \log n)^{\frac{1}{2}} \left( \frac{\tilde{M}_n}{(\lambda(w))^{\frac{1}{2}}} - d_n \right) < x \right] \rightarrow \exp[-s(\eta)e^{-x}]$$

*where*

$$(3.17) \quad s(\eta) = \int_0^1 \{\exp[\eta(t)/(f_0(t)\lambda(w))^{\frac{1}{2}}] + \exp[-\eta(t)/(f_0(t)\lambda(w))^{\frac{1}{2}}]\} dt.$$

This result follows from Theorem A1 quite readily.

One interesting consequence of this formula is that our test is asymptotically strictly unbiased for such alternatives. The reason is that  $s(\eta) \geq 2$  with  $s(\eta) > 2$  unless  $\eta = 0$  and the family of distributions  $e^{-\theta e^{-x}}$  is an exponential family in  $\theta$ .

Unfortunately these tests are asymptotically inadmissible (have Pitman efficiency 0) when compared to the test based on the quadratic functional of the next section based on the same  $w$  and  $b(n)$ . The reason is that alternatives there may be permitted to come in to  $f_0$  at rate  $n^{-\frac{1}{2}+\delta/4}$  rather than  $n^{-\frac{1}{2}+\delta/2}$ . However, this test for moderate sample sizes and some alternatives may well be preferable.

In analogy to the confidence band situation it would appear that maximum power is achieved by taking  $\delta$  as small as possible. However, consideration of the approximation arguments suggests that  $s(\eta_n)$  is a better measure of the "true shift" than  $s(\eta)$  where,

$$(3.18) \quad \eta_n = (g_n - f_0)(2nb(n) \log b(n))^{\frac{1}{2}}.$$

Of course,  $s(\eta_n)$  may well be maximized for  $\delta > 0$ . In all of these questions it would be desirable to have some small sample Monte Carlo explorations.

**4. Quadratic functionals.** We are interested in the behavior of the functional,

$$(4.1) \quad T_n = nb(n) \int_{-\infty}^{\infty} [f_n(x) - E(f_n(x))]^2 a(x) dx = \int_{-\infty}^{\infty} L_n^2(x) a(x) dx,$$

where  $L_n = f^{\frac{1}{2}} Y_n$  and  $a$  is integrable. We have already remarked that if A1 and A2 hold and (say)  $b_n = n^{-\delta}$ ,  $\delta < \frac{1}{4}$ , then.

$$(4.2) \quad |T_n - \int_0 L_n^2(x) a(x) dx| = o(b^{\frac{1}{2}}(n)).$$

Moreover, if  $a$  is bounded as well as integrable and  $w$  and  $f$  are bounded, we can replace  $_0 L_n$  by  $_1 L_n = f^{\frac{1}{2}} {}_1 Y_n$  and hence by  ${}_2 L_n = f^{\frac{1}{2}} {}_2 Y_n$ . To see this note that,

$$(4.3) \quad \begin{aligned} & \left| \int ({}_1 L_n^2(x) - {}_0 L_n^2(x)) a(x) dx \right| \\ &= \left| \int \frac{1}{b(n)} \left\{ \left( Z(1) \int w \left( \frac{t-s}{b(n)} \right) f(s) ds \right)^2 \right. \right. \\ & \quad \left. \left. - 2Z(1) \int w \left( \frac{t-s}{b(n)} \right) dZ(F(s)) \int w \left( \frac{t-s}{b(n)} \right) f(s) ds \right\} a(t) dt \right| \\ &\leq Z(1)^2 b(n) \sup_x |f(x)| \int |a(t)| dt \\ & \quad + 2|Z(1)|b(n) \left| \int \left( \int w(y)c(s+b(n)y)a(s+b(n)y) dy \right) dZ(F(s)) \right| \end{aligned}$$

where

$$c(t) = \int w(y)f(t-b(n)y) dy.$$

But,

$$(4.4) \quad \begin{aligned} E(\int (\int w(y)c(s+b(n)y)a(s+b(n)y) dy) dZ(F(s)))^2 \\ = \int (\int w(y)c(s+b(n)y)a(s+b(n)y) dy)^2 dF(s) \end{aligned}$$

is bounded.

By (4.3) and (4.4),

$$(4.5) \quad |T_n - \int {}_2 L_n^2(x) a(x) dx| = O_p(b(n)).$$

(The infinite range poses no problem since we are approximating  $L_n$  rather than the normalized  $Y_n$ .)

The following lemma lets us determine the characteristic function of a quadratic functional

$$(4.6) \quad Z = \int Y(x)^2 a(x) dx$$

of a Gaussian process  $Y(x)$  under appropriate conditions.

LEMMA 4.1. *Let  $Y(x)$ ,  $EY(x) \equiv 0$ , be a Gaussian process with bounded, uniformly continuous covariance function  $r(x, y)$ . If  $a(x)$  is a piecewise smooth integrable function, the quadratic functional (4.6) has characteristic function formally given by*

$$(4.7) \quad E(e^{itZ}) = \exp \{ \sum_{k=1}^{\infty} 2^{k-1} (it)^k c_k / k \}$$

with

$$c_k = \int \cdots \int r(x_1, x_2) r(x_2, x_3) \cdots r(x_k, x_1) a(x_1) a(x_2) \cdots a(x_k) dx_1 \cdots dx_k.$$

The representation (4.6) is valid for  $|t| < 1/2M$  where  $M = ||r|| \int |a(t)| dt$ . The quantities  $(k-1)! 2^{k-1} c_k$  are of course the cumulants of (4.6).

The lemma is obtained by considering the form

$$(4.8) \quad \sum_{j=1}^n \bar{Y}_j^2 a_j$$

in jointly Gaussian random variables  $\bar{Y}_j$ ,  $E\bar{Y}_j \equiv 0$  with the  $a_j$ 's constants. Let  $R$  be the covariance matrix of the  $\bar{Y}_j$ 's with  $A$  the diagonal matrix with diagonal entries  $a_j$ . The characteristic function of (4.8) is then

$$|1 - 2itRA|^{-\frac{1}{2}} = \prod_{j=1}^n (1 - 2\lambda_j it)^{-\frac{1}{2}} = \exp \{ \sum_{k=1}^{\infty} 2^{k-1} (it)^k \text{tr}(RA)^k / k \},$$

at least if  $|t| < 1/2 \text{tr}(RA)$ .

Here  $\text{tr}(M)$  denotes the trace of  $M$ ,  $|M|$  its determinant and  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $RA$ . Lemma 4.1 is then obtained by going through an appropriate limiting operation.

The covariance function of the Gaussian process  ${}_2L_n(x)$  can be written

$$(4.9) \quad \begin{aligned} r(x, y) &= \int w(z) w(\alpha + z) f(x - b(n)z) dz \\ &= f(x) \int w(z) w(\alpha + z) dz + O(b(n)) \end{aligned}$$

where

$$\alpha = (y - x)/(b(n))$$

and  $O(b(n))$  is independent of  $x$  if  $f$  is bounded and has a uniformly bounded derivative and  $w^2(z)(1 + |z|)$  is integrable. Then

$$(4.10) \quad E(\int {}_2L_n(x)^2 a(x) dx) = \int f(x) a(x) dx \int w(z)^2 dz + O(b(n)).$$

Similarly if  $a$  is bounded as well as integrable and  $w$  is bounded and  $f$  is as above, the variance of  $\int {}_2L_n(x)a(x) dx$  is  $2b(n) \int [w * \bar{w}(u)]^2 du \int a^2(x)f^2(x) dx$  to first order as  $n \rightarrow \infty$ , where  $\bar{w}(t) = w(-t)$  and  $*$  denotes convolution. A similar argument shows that under the same conditions the  $k$ th cumulant of  $\int {}_2L_n(x)a(x) dx$  equals to first order  $(k-1)! 2^{k-1} b^{k-1}(n) [w * \bar{w}]^{(k)}(0) \int a^k(x)f^k(x) dx$  as  $n \rightarrow \infty$  where the

superscript  $(k)$  indicates that  $w * \bar{w}$  is convoluted with itself  $k$  times. As a result we have the following theorem which actually holds under the weaker assumptions indicated above.

**THEOREM 4.1.** *Let A1—A3 hold and suppose that  $a$  is integrable piecewise continuous and bounded. Suppose moreover that (2.16) holds. Then  $b^{-\frac{1}{2}}(n)(T_n - (\int f(x)a(x) dx) \int w^2(z) dz)$  is asymptotically normally distributed with mean 0 and variance  $2(w * \bar{w})^{(2)}(0) \int a^2(x)f^2(x) dx$  as  $n \rightarrow \infty$ .*

A particular case of interest for the application of the theorem is that in which as in Section 3,  $a(x)$  vanishes off an interval, say  $[0, 1]$ , and one sets  $a(x) = f(x)^{-1}$  on  $[0, 1]$ . In this case under A1—A3,  $T_n$  is asymptotically Gaussian with mean  $\int w^2(z) dz$  and variance  $2b(n)(w * \bar{w})^{(2)}(0)$ .

The statistic

$$(4.11) \quad \tilde{T}_n = nb(n) \int [f_n(x) - f(x)]^2 a(x) dx$$

is probably of greater interest than that considered in Theorem 4.1. However, let us expand  $\tilde{T}_n$  in the form

$$(4.12) \quad \begin{aligned} nb(n) \{ & \int [f_n(x) - Ef_n(x)]^2 a(x) dx \\ & + 2 \int [f_n(x) - Ef_n(x)][Ef_n(x) - f(x)]a(x) dx \\ & + \int [Ef_n(x) - f(x)]^2 a(x) dx \}. \end{aligned}$$

Let  $w$  be positive and symmetric about zero with

$$(4.13) \quad c = \int w(u)u^2 du < \infty.$$

Then if  $n^{-1} = O(b(n))$ ,  $b(n) \rightarrow 0$  as  $n \rightarrow \infty$ , A1 holds and  $f$  has a continuous bounded second derivative, the second term of (4.12) may, by the usual approximation arguments, be shown to be asymptotically normal with mean zero and variance

$$(4.14) \quad n^{-1}b(n)^4c^2 \int f''(x)^2 a(x)^2 f(x) dx$$

to the first order. Also, under the same conditions, the last term of (4.12) can be shown to be

$$(4.15) \quad b(n)^4c^2 \int f''(x)^2 a(x) dx$$

to the first order. Then  $[b(n)]^{-\frac{1}{2}}[\tilde{T}_n - T_n] = o_p(1)$  if and only if  $b(n) = o(n^{-\frac{1}{2}})$ . (The term (4.14) is then negligible.) The theorem quoted in the introduction follows.

**APPLICATIONS.** An explicit confidence band is hard to obtain from Theorem 4.1 and the theorem of the introduction. However we can test  $H: f = f_0$  at (approximate) level  $\alpha$  by calculating  $T_n$  for  $f = f_0$  and rejecting when  $T_n \geq d(\alpha)$  where by Theorem 4.1

$$(4.16) \quad \begin{aligned} d(\alpha) = & [\int f_0(x)a(x) dx][\int w^2(z) dz] \\ & + b^{\frac{1}{2}}(n)\Phi^{-1}(1 - \alpha)/[2(w * \bar{w})^{(2)}(0) \int a^2(x)f_0^2(x) dx]^{\frac{1}{2}}. \end{aligned}$$

As in Section 3 it is easy to see that in testing  $H: f = f_0(\cdot, \theta)$  where  $\theta$  is an unknown vector parameter we may use  $T_n$  with  $f$  replaced by  $f_0(\cdot, \hat{\theta})$  and  $d(\alpha)$  with  $f_0$  replaced by  $f_0(\cdot, \hat{\theta})$ , provided that A6 below holds.

A6. For each  $\theta_0$ ,  $(\partial^2 f(x, \theta)/\partial\theta^{(i)}\partial\theta^{(j)})$  is bounded in absolute value for all  $\theta$  in a neighborhood of  $\theta_0$  and all  $x, i, j$ . Moreover, if  $\theta_0$  is true,

$$(4.17) \quad |\hat{\theta} - \theta_0| = o_p([nb(n)]^{-\frac{1}{2}}).$$

To see this, taking  $k = 1$  for simplicity, expand as in (4.12) and note that it suffices to show that

$$(4.18) \quad \int [f_n(x) - E_{\theta_0}(f_n(x))] [E_{\theta_0}(f_n(x)) - E_{\hat{\theta}}(f_n(x))] a(x) dx = o_p([nb^{\frac{1}{2}}(n)]^{-1})$$

and

$$(4.19) \quad \int [E_{\theta_0}(f_n(x)) - E_{\hat{\theta}}(f_n(x))]^2 a(x) dx = o_p([nb^{\frac{1}{2}}(n)]^{-1}).$$

Taylor expanding the integral in (4.18) about  $\theta_0$  we obtain a first term

$$(\hat{\theta} - \theta_0) \int [f_n(x) - E_{\theta_0}(f_n(x))] \left[ \int \frac{\partial f(x + b(n)z, \theta)}{\partial\theta} \Big|_{\theta=\theta_0} w(z) dz \right] a(x) dx$$

which is  $O_p(|\hat{\theta} - \theta_0|n^{-\frac{1}{2}})$ , and a second term which is  $O_p([nb(n)]^{-\frac{1}{2}}(\hat{\theta} - \theta_0)^2)$ , and (4.18) follows. A similar argument yields (4.19).

To make local power calculations we again suppose  $g_n$  is as in (3.15) with  $g_n$  satisfying A2—A3 uniformly in  $n$  and  $o(\gamma_n)$  uniform in  $x$  and  $\eta$  is bounded.

**THEOREM 4.2.** *Let  $g_n$  be as above,  $w$  satisfy A1—A4,  $a$  be integrable piecewise continuous and bounded,  $b(n) = n^{-\delta}$ ,  $\delta < \frac{1}{4}$ ,  $\gamma_n = n^{-\frac{1}{2}+\delta/4}$ . Define  $T_n$  in terms of  $f_0$ . Then,*

$$(4.20) \quad b^{-\frac{1}{2}}(n)(T_n - [\int f_0(x)a(x) dx] \int w^2(z) dz)$$

*is asymptotically normally distributed with mean  $\int \eta_n^2(x)a(x) dx$  and variance*

$$2(w * \bar{w})^{(2)}(0) \int a^2(x)f_0^2(x) dx.$$

The proof is straightforward. As in Section 3 it follows that the test which rejects when  $T_n$  is  $\geq d(\alpha)$  is locally strictly unbiased if  $a(x) > 0$  for all  $x$ .

Also as before the asymptotics lead to choosing  $\delta$  as large as possible and again this conclusion is shaken if one uses the better approximation to the asymptotic mean,  $\int \eta_n^2(x)a(x) dx$  where

$$(4.21) \quad \eta_n(x) = \int w(z)[g_n(x + b(n)z) - f_0(x + b(n)z)] dz.$$

It is also clear that for fixed  $\delta$  we can let  $\lambda_n \rightarrow 0$  more quickly than for the sup functional and still get power. Thus the Pitman efficiency of the  $T_n$  test to the  $\tilde{M}_n$  test for the same  $\delta$  is  $\infty$ .

Suppose that  $f_0$  is the uniform density on  $[0, 1]$  an effect we can always achieve by applying the probability integral transformation to our observations before making the test. Let  $a(x) = 1$  on  $[0, 1]$  and 0 otherwise,  $w$  be the uniform density

on  $[-\frac{1}{2}, \frac{1}{2}]$ . Neglecting fringe effects we may then write

$$(4.22) \quad T_n = \int_0^1 \frac{(N[t - \frac{1}{2}b(n), t + \frac{1}{2}b(n)] - b(n))^2}{nb(n)} dt$$

where  $N[x, y]$  is the number of observations falling in the interval  $[x, y]$ . A related statistic for testing uniformity on the circle was considered by Watson in [15]. This is, of course, very similar to the  $\chi^2$  statistic for the problem based on the cells  $[0, b(n)]$ ,  $[b(n), 2b(n)]$ ,  $\dots$ ,  $[(K-1)\frac{1}{2}b(n), (K+1)\frac{1}{2}b(n)]$  given by,

$$(4.23) \quad \chi_n^2 = \sum_{k=1}^{*K} \frac{(N[\frac{1}{2}kb(n) - \frac{1}{2}b(n), \frac{1}{2}kb(n) + \frac{1}{2}b(n)] - nb(n))^2}{nb(n)}$$

where  $(K+1)\frac{1}{2}b(n) \leq 1 < (K+2)\frac{1}{2}b(n)$  and  $\sum^*$  is a sum over odd index.

Now we can write,

$$(4.24) \quad \chi_n^2/K = nb(n) \int_0^1 (f_n(t) - E(f_n(t)))^2 dA_n(t)$$

where  $A_n$  places mass  $1/K$  at each of the points  $\frac{1}{2}b(n), \dots, K\frac{1}{2}b(n)$ . It is easy to see that the arguments leading to Theorem 4.2 apply to functionals of this type also and that under the conditions of that theorem, if  $b(n) = n^{-\delta}$ ,  $\delta < \frac{1}{2}$ ,  $\chi_n^2/K$  is asymptotically normal with the natural parameters  $E(\chi_n^2/K)$  and  $\text{Var}(\chi_n^2/K)$ .

This result is, of course, known. A rigorous proof under milder conditions but using a different method may be found in Steck (1957). Now

$$(4.25) \quad E\left(\frac{\chi_n^2}{K}\right) = 1 + \frac{1}{K} \sum_{j=1}^{*K} nb(n) \left(1 - \frac{1}{b(n)} \int_{(j-1)\frac{1}{2}b(n)/2}^{(j+1)\frac{1}{2}b(n)/2} g_n(x) dx\right)^2 \\ = 1 + nb(n)\gamma_n^2 \frac{1}{K} \sum_{j=1}^{*K} \left[ \frac{1}{b(n)} \int_{(j-1)\frac{1}{2}b(n)/2}^{(j+1)\frac{1}{2}b(n)/2} \eta(x) dx \right]^2 + o(nb(n)\gamma_n^2)$$

$$(4.26) \quad \text{Var}\left(\frac{\chi_n^2}{K}\right) = \frac{1}{K} \text{Var}\left(\frac{N^2[0, b(n)]}{nb(n)}\right) + o\left(\frac{1}{K}\right) = \frac{2}{K} + o\left(\frac{1}{K}\right).$$

Thus if we take  $\gamma_n = n^{-\frac{1}{2}+\delta/4}$  as in Theorem 4.2, under  $g_n$  the statistics

$$(4.27) \quad W_n = b^{-\frac{1}{2}}(n) \left( \frac{\chi_n^2}{K} - 1 \right)$$

have a limiting Gaussian distribution with mean  $\int_0^1 \eta^2(x) dx$  and variance 2. Under the same circumstances the asymptotic mean of  $b^{-\frac{1}{2}}(n)(T_n - 1)$  with  $T_n$  given by (4.22) is also  $\int_0^1 \eta^2(x) dx$  while its asymptotic variance is,

$$(4.28) \quad 2w^{(4)}(0) = 2 \int_{-1}^1 (1 - |t|)^2 dt = \frac{4}{3}.$$

The Pitman efficiency of the tests based on  $T_n$  to those based on  $W_n$  is thus by the usual calculations,

$$(4.29) \quad e(T_n, W_n) = \left(\frac{3}{2}\right)^{\frac{1}{2}-\delta}$$

and thus at least  $(\frac{3}{2})^{\frac{1}{2}} = 1.217$  on the range  $\delta > 0$ . For the Mann-Wald (1942) prescription  $\delta = \frac{2}{5}$  we get an efficiency of 1.292.

Although as we have seen these asymptotic calculations are to be taken with a grain of salt we feel that the procedure  $T_n$  has promise as a competitor to the  $\chi^2$  test, at least for moderate sample sizes.

**Acknowledgment.** We are grateful to S. Berman and J. Pickands, III for providing us with preprints of some of the papers cited in the list of references.

**5. Appendix A. On the extrema of some nonstationary Gaussian processes.** Let  $Y_T(\cdot)$  be a sequence of separable Gaussian processes with mean  $\mu_T(\cdot)$  such that  $Y_T(\cdot) - \mu_T(\cdot)$  is stationary. Let  $r(\cdot)$  be the covariance function of  $Y_T$ ,

$$M_T = \max \{Y_T(t) : 0 \leq t \leq T\}, \quad m_T = \min \{Y_T(t) : 0 \leq t \leq T\}.$$

Let  $b_T(t) = \mu_T(t)(2 \log T)^{\frac{1}{2}}$ .

**THEOREM A1.** *Suppose that,*

- (i)  $b_T(t)$  is uniformly bounded in  $t$  and  $T$  on  $[0, T]$  as  $T \rightarrow \infty$ .
- (ii)  $b_T(t) \rightarrow b(t)$  uniformly on  $[0, T]$  as  $T \rightarrow \infty$ .
- (iii)  $T^{-1}\lambda[t: b(t) \leq x, 0 \leq t \leq T] \rightarrow \eta(x)$  the cdf of a probability measure as  $T \rightarrow \infty$ . ( $\lambda$  as usual denotes Lebesgue measure.)
- (iv)  $b(\cdot)$  is uniformly continuous on  $R$ .
- (v)  $r(t) = 1 - C|t|^\alpha + o(|t|^\alpha)$ ,  $0 < \alpha \leq 2$ , as  $t \rightarrow \infty$ .
- (vi)  $\int_0^\infty r^2(t) dt < \infty$ .

Let

$$\begin{aligned} B(t) &= (2 \log t)^{\frac{1}{2}} + \frac{1}{(2 \log t)^{\frac{1}{2}}} \\ &\times \left\{ \left( \frac{1}{\alpha} - \frac{1}{2} \right) \log \log t + \log(2\pi)^{-\frac{1}{2}} (C^{1/\alpha} H_\alpha 2^{(2-\alpha)/2\alpha}) \right\} \end{aligned}$$

where

$$H_\alpha = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^\infty e^s P[\sup_{0 \leq t \leq T} Y(t) > s] ds$$

and  $Y$  is a Gaussian process with,

$$(A.1) \quad E(Y(t)) = -|t|^\alpha, \quad \text{Cov}(Y(t_1), Y(t_2)) = |t_1|^\alpha + |t_2|^\alpha - |t_1 - t_2|^\alpha.$$

Then,

$$U_T = (2 \log T)^{\frac{1}{2}}(M_T - B(T)) \quad \text{and} \quad V_T = -(2 \log T)^{\frac{1}{2}}(m_T + B(T))$$

are asymptotically independent with,

$$(A.2) \quad P[U_T < z] \rightarrow e^{-\lambda_1 e^{-z}}, \quad P[V_T < z] \rightarrow e^{-\lambda_2 e^{-z}};$$

where,

$$(A.3) \quad \lambda_1 = \int e^z d\eta(z), \quad \lambda_2 = \int e^{-z} d\eta(z).$$

An immediate consequence of Theorem A1 is,

COROLLARY A1. *If  $\tilde{M}_T = \max \{|Y_T(t)| : 0 \leq t \leq T\}$  then under the conditions of the theorem,*

$$(A.4) \quad P[(2 \log T)^{\frac{1}{2}}(\tilde{M}_T - B(T)) < x] \rightarrow \exp[-(\lambda_1 + \lambda_2)e^{-x}].$$

*Note.*  $\lambda_1 + \lambda_2 \geq 2$  with strict inequality unless  $\eta$  concentrates at 0.

COROLLARY A2. *Let  $Y_0(t) - \mu(t)$  be a stationary mean 0 Gaussian process with covariance function  $r(t)$  satisfying the conditions of the theorem. Suppose that  $b(t) = (2 \log(t+2))^{\frac{1}{2}}\mu(t)$  is a bounded uniformly continuous function of  $t$  and that  $b(\cdot)$  satisfies condition (iii) of the theorem. Then,*

$$(A.5) \quad P[(2 \log T)^{\frac{1}{2}}(\max \{Y_0(s) : 0 \leq s \leq T\} - B(T)) < x] \rightarrow e^{-\lambda_1 e^{-x}}.$$

Similar assertions hold about the independence of maximum and minimum and the asymptotic distribution of the minimum.

This corollary may be viewed as complementing Theorem 4.1 of Qualls and Watanabe (1971) which deals with the extrema of a mean 0 process whose covariance function is asymptotically locally approximated by that of a stationary process while we deal with a process which is stationary when centered and asymptotically stationary.

The constants  $H_1$  and  $H_2$  are the only ones known explicitly. They are given by  $H_1 = 1$ ,  $H_2 = \pi^{-\frac{1}{2}}$  (cf. [11]).

PROOF OF COROLLARY A2. Define,

$$(A.6) \quad \begin{aligned} Y_T(t) &= Y_0(t) && \text{on } [\varepsilon(T), T] \\ &= Y_0(t) + ([\log(t+2)/\log(T+2)]^{\frac{1}{2}} - 1)\mu(t) && \text{otherwise} \end{aligned}$$

where  $\varepsilon(T) = o(T)$ ,  $\log \varepsilon(T) \sim \log T$ . Evidently,  $(2 \log T)^{\frac{1}{2}}E(Y_T(t)) \rightarrow b(t)$  uniformly and

$$(A.7) \quad \begin{aligned} & \left| P \left[ \max \{Y_T(s) : 0 \leq s \leq T\} < \frac{x}{(2 \log T)^{\frac{1}{2}}} + B(T) \right] \right. \\ & \quad \left. - P \left[ \max \{Y_0(s) : 0 \leq s \leq T\} < \frac{x}{(2 \log T)^{\frac{1}{2}}} + B(T) \right] \right| \\ & \leq 2P \left[ \max \{Y_0(s) - E(Y_0(s)) : 0 \leq s \leq \varepsilon(T)\} \geq \frac{x}{(2 \log T)^{\frac{1}{2}}} \right. \\ & \quad \left. - K + B(T) \right] \end{aligned}$$

where  $K = \max \{\mu(t) : 0 \leq t \leq \varepsilon(t)\}$ . Since  $B(\varepsilon(T)) - B(T) \rightarrow -\infty$  the term on the right of (A.7) tends to 0 by the theorem.  $\square$

PROOF OF THEOREM A1. The theorem is argued much as Theorem 3.1 of Pickands (1969). We refer the reader to this paper and Berman (1971) for the details of the argument.

LEMMA A1. Let  $\psi(x) = \phi(x)/x$  where  $\phi$  is the standard normal density. Let  $C = 1$ ,  $x = x(T) = B(T) + z_1/(2 \log T)^{\frac{1}{2}}$ . Then for  $a > 0$ ,

$$(A.8) \quad \begin{aligned} P[\max \{Y_T(t + akx^{-2/\alpha}), 0 \leq k \leq n\} > x] \\ &= \psi(x)e^{b(t)}H_\alpha(n, \alpha) + o(\psi(x)) \\ P[\min \{Y_T(t + kax^{-2/\alpha}), 0 \leq k \leq n\} < -x] \\ &= \psi(x)e^{-b(t)}H_\alpha(n, a) + o(\psi(x)) \end{aligned}$$

as  $T \rightarrow \infty$  uniformly in  $0 \leq t \leq T$  where

$$(A.9) \quad H_\alpha(n, a) = \int_{-\infty}^{\infty} e^s P[\max \{Y(ka) : 0 \leq k \leq n\} > s] ds.$$

Moreover, if  $y = y(T) = B(T) + z_2/(2 \log T)^{\frac{1}{2}}$  then

$$(A.10) \quad \begin{aligned} P[\max \{Y_T(t + kax^{-2/\alpha}) : 0 \leq k \leq n\} > x, \\ \min \{Y_T(t + kax^{-2/\alpha}) : 0 \leq k \leq n\} < -y] \\ &= o(\psi(x)) = o(\psi(y)), \end{aligned}$$

uniformly in  $0 \leq t \leq T$ . (Throughout,  $k$  may take on integer values only.)

PROOF. As in [11] consider the “local” process

$$(A.11) \quad \tilde{Y}_T(s) = x(Y_T(t + sx^{-2/\alpha}) - \mu_T(t)) - x.$$

$$(A.12) \quad \begin{aligned} P[\max \{Y_T(t + akx^{-2/\alpha}) : 0 \leq k \leq n\} < x] \\ &= \int_{-\infty}^{\infty} \gamma(z) P[\max \{\tilde{Y}_T(ka) : 0 \leq k \leq n\} > -x \mu_T(t) | \tilde{Y}_T(0) = z] dz \end{aligned}$$

where  $\gamma$  is the density of  $\tilde{Y}_T(0)$ ,

$$(A.13) \quad \gamma(z) = \frac{1}{x} \phi\left(x + \frac{z}{x}\right) = \psi(x) \exp[-z - z^2/2x^2].$$

It is easy to see using (ii) and (iv) that the finite dimensional conditional distributions of  $\tilde{Y}_T(s)$  given  $\tilde{Y}_T(0) = z$  converge uniformly in  $t$  to those of the process  $Y(s) + z$  where  $Y$  is given by (A.1). Arguing as in [11] the first part of (A.8) follows since  $x\mu_T(t) \rightarrow b(t)$  uniformly as required. By considering  $-\tilde{Y}_T$  we obtain the second part. To prove (A.10) let  $A$  be the event whose probability is being estimated. Then,

$$(A.14) \quad \begin{aligned} P\left(A, Y_T(t) > x - \frac{1}{x^{\frac{1}{2}}} + \mu_T(t)\right) \\ &\leq \int_{-x^{\frac{1}{2}}}^{\infty} \gamma(z) P[\min \{\tilde{Y}_T(ka) : 0 \leq k \leq n\} < z \\ &\leq -z - x(y + x + \mu_T(t)) | \tilde{Y}_T(0) = z] dz \\ &\leq \psi(x) \int_{-\infty}^{x^{\frac{1}{2}}} e^z P[\min \{\tilde{Y}_T(ka) + z : 0 \leq k \leq n\} \\ &< z - x(y + x + \mu_T(t)) | \tilde{Y}_T(0) = -z] dz \\ &\leq \psi(x) \sum_{k=0}^n \left\{ \int_{-\infty}^0 P[\tilde{Y}_T(ka) + z < z \right. \\ &\quad \left. - x(y + x + \mu_T(t)) | \tilde{Y}_T(0) = -z] dz \right. \\ &\quad \left. + x^{\frac{1}{2}} \exp x^{\frac{1}{2}} \max \{P[\tilde{Y}_T(ka) + z \right. \\ &\quad \left. < x^{\frac{1}{2}} - x(y + x + \mu_T(t)) | \tilde{Y}_T(0) = -z] : 0 \leq z \leq x^{\frac{1}{2}}\} \right\}. \end{aligned}$$

Applying the usual estimate  $\Phi(z) \leq \psi(|z|)$  for  $z \leq 0$  we conclude that the left-hand side of (A.14) is  $o(\psi(x))$ . Similarly,

$$(A.15) \quad P\left(A, Y_T(t) - \mu_T(t) < -y + \frac{1}{y^{\frac{1}{2}}}\right) = o(\psi(y)).$$

Finally,

$$(A.16) \quad \begin{aligned} & P\left(A, -y + \frac{1}{y^{\frac{1}{2}}} \leq Y_T(t) - \mu_T(t) \leq x - \frac{1}{x^{\frac{1}{2}}}\right) \\ & \leq \int_{-\infty}^{-z^{\frac{1}{2}}} \gamma(z) P[\max\{\tilde{Y}_T(ka) : 0 \leq k \leq n\} > -x\mu_T(t) \mid \tilde{Y}_T(0) = z] dz \\ & \leq \psi(x) \int_A^\infty e^z P[\max\{\tilde{Y}_T(ka) : 0 \leq k \leq n\} > z] dz \end{aligned}$$

for every  $A < \infty$ .

The final statement of the lemma follows.

**LEMMA A2.** *The assertion of Lemma A1 remains valid if  $a = 1$ ,  $k$  is permitted to range over all values in  $[0, n]$  and  $H_\alpha(n, a)$  is replaced by*

$$(A.17) \quad \bar{H}_\alpha(n) = \int_{-\infty}^\infty e^t P[\max\{Y(s) : 0 \leq s \leq n\} > t] dt$$

**PROOF.** We prove the analogue of (A.8); the other assertions follow similarly. We need to check that uniformly in  $T$ ,

(a) The conditional distributions of the continuous processes  $\tilde{Y}_T(t) - z$  given  $\tilde{Y}_T(0) = z$  converge weakly (in the sense of Prohorov) to that of  $Y(\cdot)$ ,

$$(b) \quad P[\max\{\tilde{Y}_T(k) : 0 \leq k \leq n\} > x\mu_T(t) \mid \tilde{Y}_T(0) = z] \leq g(z)$$

where  $\int e^{-z} g(z) dz < \infty$ .

To see that (a) holds it suffices to note that,

$$(A.18) \quad \text{Var}[(\tilde{Y}_T(s_1) - \tilde{Y}_T(s_2)) \mid \tilde{Y}_T(0) = z] \leq C |s_1 - s_2|^\alpha$$

and then apply Billingsley [3] page 95. To see that (b) is valid use the estimate of Fernique (1970) given below on the tails of  $\max\{|\tilde{Y}_T(k)| : 0 \leq k \leq n\}$ .

**LEMMA.** *Let  $Z(\cdot)$  be a Gaussian process on  $(0, 1)$ . Let  $a$  be such that  $P[\|Z\| \leq a] \geq \frac{3}{4}$ ,  $P[\|Z\| \geq a] \geq \frac{1}{4}$ . Then, for  $z \geq a$*

$$P[\|Z\| > z] \leq \exp\left\{-\frac{z^2}{24a^2} \log 3\right\}.$$

**LEMMA A3.** *Fix  $t > 0$  such that  $\inf\{s^{-\alpha}(1 - r(s)) : 0 \leq s \leq t\} \geq A(t) > 0$ . Define  $x$  and  $y$  as before. Let,*

$$(A.19) \quad H_\alpha(a) = \lim_{n \rightarrow \infty} \frac{H_\alpha(n, a)}{n}.$$

$$(A.20) \quad 0 < H_\alpha = \lim_{a \rightarrow 0} \frac{H_\alpha(a)}{a} = \lim_{n \rightarrow \infty} \frac{\bar{H}_\alpha(n)}{n}.$$

(See the note at the end of the lemma.)

Then,

$$(A.21) \quad P\left[\max\left\{Y_r(v + kax^{-2/\alpha}) : 0 \leq k \leq \left\lceil \frac{x^{2/\alpha}}{a} t \right\rceil\right\} > x\right] \\ = x^{2/\alpha} \psi(x) \frac{H_\alpha(a)}{a} \int_v^{v+t} \exp b(s) ds + o(x^{2/\alpha} \psi(x)),$$

$$(A.22) \quad P[\max\{Y_r(v + s) : 0 \leq s \leq t\} > x] \\ = x^{2/\alpha} \psi(x) [\int_v^{v+t} \exp b(s) ds] H_\alpha + o(x^{2/\alpha} \psi(x)),$$

uniformly in  $0 \leq v \leq T$ . Similar assertions hold for  $P[\min\{Y_r(v + s) : 0 \leq s \leq t\} < -x]$  with  $-b$  replacing  $b$ . Finally,

$$(A.23) \quad P[\max\{Y_r(v + s) : 0 \leq s \leq t\} > x, \min\{Y_r(v + s) : 0 \leq s \leq t\} < -y] \\ = o(x^{2/\alpha} \psi(x)).$$

*Note.* The existence of the limit in (A.19) was first proved in [11]. An incorrect proof of (A.20) was also given. Subsequently, a correct proof was communicated to the author by J. Pickands and another is included in [12]. We provide yet a third in Appendix B.

PROOF. We prove (A.22); (A.21) is argued similarly. Begin by bounding the left-hand side of (A.22) from above by,

$$(A.24) \quad \sum_{k=0}^M P[\max\{Y_r(v + knx^{-2/\alpha} + s) : 0 \leq s \leq nx^{-2/\alpha}\} > x]$$

where  $M = \lceil tx^{2/\alpha}/n \rceil$ . By Lemma A2 the expression above is asymptotic to

$$(A.25) \quad \frac{t\bar{H}_\alpha(n)}{n} x^{2/\alpha} \psi(x) \left[ \frac{1}{M+1} \sum_{k=0}^M \exp b(v + knx^{-2/\alpha}) \right] \\ = \frac{\bar{H}_\alpha(n)}{n} x^{2/\alpha} \psi(x) [\int_v^{v+t} \exp b(s) ds + o(1)]$$

since  $b$  is assumed uniformly continuous and bounded. On the other hand we can bound from below by the left-hand side of (A.21) which in turn is bounded from below by,

$$(A.26) \quad \sum_{r=0}^{M_a} P(A_r) - \sum_{0 \leq r \neq s \leq M_a} P(A_r A_s)$$

where  $A_r = [\{\max\{Y_r(v + kax^{-2/\alpha}), rn \leq k < (r+1)n\} > x\}], M_a = \lceil x^{-2/\alpha} t / na \rceil$ . If we apply Lemma A.1 to the first term on the right of (A.26) we obtain that,

$$(A.27) \quad \sum_{r=0}^{M_a} P(A_r) \sim \frac{H_\alpha(n, a)}{na} x^{2/\alpha} \psi(x) [\int_v^{v+t} e^{b(s)} ds].$$

Finally,

$$(A.28) \quad P(A_r A_s) \leq P(C_r C_s)$$

where

$$(A.29) \quad C_r = \left[ \max\{Y_r(kax^{-2/\alpha} + v) - \mu_r(kax^{-2/\alpha}), rn \leq k < (r+1)n\} \right. \\ \left. > x - \frac{K}{(2 \log T)^{1/2}} \right]$$

where  $K = \sup \{(2 \log T)^{\frac{1}{2}} |\mu_T(t)| : 0 \leq t \leq T\}$ . Now applying Lemma 2.3 of [11] and arguing as in Lemma 2.5 of the same paper we see that,

$$(A.30) \quad \sum P(C_r C_s) = o(x^{2/\alpha} \psi(x)).$$

Applying (A.20) we see that (A.22) follows. To prove (A.23) it suffices to show that,

$$(A.31) \quad \begin{aligned} & P[\max \{Y_r(v+s) : 0 \leq s \leq t\} > x] \\ & \quad \cup [\min \{Y_r(v+s) : 0 \leq s \leq t\} < -y] \\ & = x^{2/\alpha} \psi(x) H_\alpha \int_v^{v+t} [\exp b(\xi)] d\xi \\ & \quad + y^{2/\alpha} \psi(y) H_\alpha \int_v^{v+t} [\exp -b(\xi)] d\xi + o(x^{2/\alpha} \psi(x)). \end{aligned}$$

But we can bound the expression on the left of (A.31) from above by

$$P[\max \{Y_r(v+s) : 0 \leq s \leq t\} > x] + P[\min \{Y_r(v+s) : 0 \leq s \leq t\} < -y]$$

and from below as in (A.26) where we add  $A_{M_a+1}, \dots, A_{2M_a+1}$  with  $A_{M_a+j} = \{\min \{Y_r(v+kax^{-2/\alpha}) : (j-1)n \leq k < jn\} < -y\}$ . Now by (A.10)

$$(A.32) \quad \frac{1}{n} \sum_{j=0}^{M_a} P(A_j A_{M_a+j+1}) = o(x^{2/\alpha} \psi(x)).$$

Finally, again arguing as for the previous case,

$$(A.33) \quad \frac{1}{n} \sum_{0 \leq j \neq k \leq M_a} P(A_j A_k), \quad \frac{1}{n} \sum_{1 \leq j \neq k \leq M_a+1} P(A_{j+M_a} A_{k+M_a}) \quad \text{and} \\ \frac{1}{n} \sum_{0 \leq j \neq k \leq M_a} P(A_j A_{M_a+k+1}) \quad \text{are all } o(x^{2/\alpha} \psi(x)). \quad \square$$

The rest of the proof goes much as in Berman [1]. Neglecting fringe effects break the interval  $[0, T]$  up into  $2N$  intervals of which half,  $W_1, \dots, W_N$  are of length  $t$  and the others  $V_1, \dots, V_N$  of length  $\varepsilon$  so that  $V_i$  follows  $W_i$  which follows  $V_{i-1}$ ,  $i = 2, \dots, N$ . Of course,  $N \sim T/(t + \varepsilon)$ . Define  $x$  and  $y$  as in Lemma A1 and note that,

$$(A.34) \quad x^{2/\alpha} \psi(x) H_\alpha \sim \frac{1}{T} e^{-z_1}.$$

Then, by Lemma A3,

$$(A.35) \quad \begin{aligned} P[\max \{Y_r(\tau) : \tau \in \bigcup_{j=1}^N V_j\} \geq x] & \leq \sum_{j=1}^N P[\max \{Y_r(\tau) : \tau \in V_j\} \geq x] \\ & \sim [\sum_{j=1}^N \int_{V_j} \exp b(s) ds] \frac{e^{-z_1}}{T} \\ & = \varepsilon O\left(\frac{N}{T}\right) = \varepsilon O(1) \end{aligned}$$

where the  $O$  term is independent of  $\varepsilon$  and the  $V_j$ . A similar assertion holds for  $\min \{Y_r(\tau) : \tau \in \bigcup_{j=1}^N V_j\}$  and hence we need only show that,

$$(A.36) \quad \begin{aligned} \lim_{\varepsilon \rightarrow 0} \overline{\lim}_{T \rightarrow \infty} P[\max \{Y_r(\tau) : \tau \in \bigcup_{j=1}^N W_j\} \leq x, \\ \min \{Y_r(\tau) : \tau \in \bigcup_{j=1}^N W_j\} \geq -y] \\ = \exp -\{\lambda_1 e^{-z_1} + \lambda_2 e^{-z_2}\}, \end{aligned}$$

where  $\lambda_1, \lambda_2$  are defined in (A.3) and the bars above and below the limit sign indicate  $\limsup$  and  $\liminf$  respectively. Next choose  $a > 0$ . If  $W_j = [a_j, a_j + t]$ ,  $j = 1, \dots, N$ .

$$\begin{aligned}
 & \left| P[\max \{Y_r(\tau) : \tau \in \bigcup_{j=1}^N W_j\} \leq x] \right. \\
 & \quad \left. - P\left[Y_r(a_j + kax^{-2/\alpha}) \leq x : 0 \leq k \leq \left\lceil \frac{tx^{2/\alpha}}{a} \right\rceil, 1 \leq j \leq N\right] \right| \\
 (A.37) \quad & \leq \sum_{j=1}^N \left| P[\max \{Y_r(\tau) : \tau \in W_j\} \leq x] \right. \\
 & \quad \left. - P\left[\max \left\{Y_r(a_j + kax^{-2/\alpha}) : 0 \leq k \leq \left\lceil \frac{tx^{2/\alpha}}{a} \right\rceil\right\} \leq x\right] \right| \\
 & \sim \left[ \sum_{j=1}^N \int_{W_j} \exp b(s) ds \right] x^{2/\alpha} \psi(x) \left[ H_\alpha - \frac{H_\alpha(a)}{a} \right] e^{-z_1},
 \end{aligned}$$

by Lemma A3.

A similar argument holds for  $P[\min \{Y_r(\tau) : \tau \in \bigcup_{j=1}^N W_j\} \geq -y]$  and by simple probability manipulations it follows that to prove the theorem we need only show,

$$\begin{aligned}
 & \lim_{a \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{r \rightarrow \infty} P\left[-y \leq Y_r(a_j + kax^{-2/\alpha}) \leq x : 1 \leq j \leq N, \right. \\
 (A.38) \quad & \quad \left. 0 \leq k \leq \left\lceil \frac{tx^{2/\alpha}}{a} \right\rceil \right] \\
 & = \exp\{-[\lambda_1 e^{-z_1} + \lambda_2 e^{-z_2}]\}.
 \end{aligned}$$

Now in view of Lemma A3 it is easy to show that,

$$\begin{aligned}
 (A.39) \quad & \lim_{r \rightarrow \infty} \sum_{j=1}^N \left( 1 - P\left[-y \leq Y_r(a_j + kax^{-2/\alpha}) \leq x : 0 \leq k \leq \left\lceil \frac{tx^{2/\alpha}}{a} \right\rceil\right] \right) \\
 & = \frac{H_\alpha(a)}{a H_\alpha} \lim_{r \rightarrow \infty} \frac{1}{T} \sum_{j=1}^N \int_{W_j} \{\exp[b(s) - z_1] + \exp[-[b(s) + z_2]]\} ds.
 \end{aligned}$$

Since, by the boundedness of  $b$ ,  $T^{-1}[\sum_{j=1}^N \int_{W_j} \exp b(s) ds - \int_0^T \exp b(s) ds] = O(\epsilon)$  uniformly in  $T$  it follows from (A.39) and (A.20) that

$$\begin{aligned}
 (A.40) \quad & \lim_{a \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{r \rightarrow \infty} \sum_{j=1}^N \left( 1 - P\left[-y \leq Y_r(a_j + kax^{-2/\alpha}) \leq x : 0 \leq k \leq \left\lceil \frac{tx^{2/\alpha}}{a} \right\rceil\right] \right) \\
 & = \lambda_1 e^{-z_1} + \lambda_2 e^{-z_2}.
 \end{aligned}$$

Let  $E_j, j = 1, \dots, N$  be the events whose probabilities are being summed in (A.40). The assertion (A.38) corresponds to a limiting statement about  $P(E_1 \dots E_N)$ . If the  $E_j$  were independent assertion (A.38) would follow readily from (A.40). Let  $\tilde{P}$  be the measure which makes the vectors  $(Y_r(a_1), Y_r(a_1 + ax^{-2/\alpha}), \dots, Y_r(a_1 + ax^{-2/\alpha}[tx^{2/\alpha}/a]))$ ,  $(Y_r(a_2), \dots, Y_r(a_2 + ax^{-2/\alpha}[tx^{2/\alpha}/a])), \dots, (Y_r(a_N), \dots, Y_r(a_N + ax^{-2/\alpha}[tx^{2/\alpha}/a]))$  independent and otherwise agrees with  $P$ .

To conclude the proof of the theorem we need to show that.

$$(A.41) \quad \lim_{\epsilon \rightarrow 0} \overline{\lim}_T |(P - \tilde{P})(E_1 \cdots E_N)| = 0.$$

To do this apply the following modification of Lemma 4.1 of [1].

LEMMA A4. *Let*

$$(A.42) \quad \phi(x, y, p) = \frac{1}{2\pi(1-p^2)^{\frac{1}{2}}} \exp -\frac{(x^2 - 2pxy + y^2)}{2(1-p^2)}.$$

Let  $\Sigma_1 = |r_{ij}|$ ,  $\Sigma_2 = |s_{ij}|$  be  $k \times k$  nonnegative semi-definite matrices with  $r_{ii} = s_{ii} = 1$  for all  $i$ . Let  $\mathbf{X} = (X_1, \dots, X_k)$  be a mean 0 Gaussian vector with covariance matrix  $\Sigma_1$  or  $\Sigma_2$ . Let  $u_1, \dots, u_k$  be nonnegative numbers and  $u = \min_j u_j$ . Then,

$$(A.43) \quad |P_{\Sigma_1}[X_j \leq u_j, 1 \leq j \leq k] - P_{\Sigma_2}[X_j \leq u_j, 1 \leq j \leq k]| \leq 4 \sum_{i,j} |\int_{s_{ij}}^{r_{ij}} \phi(u_i, u_j; \lambda) d\lambda|.$$

PROOF. By the usual argument (see [1] page 931) the left-hand side of (A.43) is bounded by,  $4 \sum_{i,j} |\int_{s_{ij}}^{r_{ij}} \phi(u_i, u_j; \lambda) d\lambda|$ . But, by an elementary inequality

$$(A.44) \quad x^2 - 2pxy + y^2 \geq \frac{(1-p)}{2} (x+y)^2.$$

Thus,

$$(A.45) \quad \phi(u_i, u_j, \lambda) \leq \phi\left(\frac{u_i + u_j}{2}, \frac{u_i + u_j}{2}, \lambda\right) \leq \phi(u, u, \lambda). \quad \square$$

Take  $X_1 = Y_T(a_1) - \mu_T(a_1)$ ,  $X_2 = -Y_T(a_1) + \mu_T(a_1)$  etc.,  $k = 2N[tx^{2/\alpha}/a]$ ,  $|r_{ij}|$  corresponding to the distribution of  $\mathbf{X}$  under  $P$ ,  $|s_{ij}|$  corresponding to  $\tilde{P}$ ,  $u_1 = x - \mu_T(a_1)$ ,  $u_2 = y + \mu_T(a_1)$  etc. Evidently,

$$(A.46) \quad u = (2 \log T)^{\frac{1}{2}} + O((\log T)^{-\frac{1}{2}}).$$

It is clear now that we can apply to the bound of (A.43) exactly the same analysis as that given by Berman on pages 933–936 of [1] to arrive at the conclusion of the theorem.

Note. By applying the more refined analysis of Pickands [11] pages 64–72 we can show that the conclusion of the theorem also holds if (vi) is replaced by,

$$(A.47) \quad \lim_{t \rightarrow \infty} r(t) \log t = 0.$$

Unfortunately, the analysis of Berman appears to only yield the conclusion under the stronger

$$(A.48) \quad r(t)[\log t]^{2/\alpha} \rightarrow 0.$$

We do not enter into this further since (vi) is what we need for Theorems 1.1 and 1.2.

## 5. Appendix B. Miscellanea.

**THEOREM B1.** *Let  $w$  be an absolutely continuous square integrable function with*

a square integrable derivative  $w'$ . Let,

$$(B.1) \quad r(t) = \int w(t+s)w(s) ds.$$

Then  $r$  is twice differentiable and

$$(B.2) \quad r''(t) = - \int w'(t+s)w'(s) ds.$$

PROOF. We first show that

$$(B.3) \quad r'(t) = \int w'(t+s)w(s) ds = \int w(s-t)w'(s) ds.$$

Let  $\hat{w}$ ,  $\hat{w}'$  be the Fourier transforms of  $w$ ,  $w'$ . Then by Parseval,

$$(B.4) \quad \frac{r(t+h) - r(t)}{h} = \frac{1}{2\pi} \int \frac{(e^{-i(t+h)u} - e^{-itu})}{h} |\hat{w}(u)|^2 du.$$

Applying the dominated convergence theorem we obtain the existence of  $r'$  given by

$$r'(t) = -\frac{i}{2\pi} \int e^{-itu} u |\hat{w}(u)|^2 du = \int w'(t+s)w(s) ds.$$

Similarly

$$(B.5) \quad \begin{aligned} \frac{r'(t+h) - r'(t)}{h} &= \int \frac{w(s-t-h) - w(s-t)}{h} w'(s) ds \\ &= \frac{i}{2\pi} \int \left( \frac{e^{i(t+h)u} - e^{itu}}{h} \right) u |\hat{w}(u)|^2 du \\ &\rightarrow -\frac{1}{2\pi} \int e^{itu} u^2 |\hat{w}(u)|^2 du = - \int w'(s-t)w'(s) ds. \end{aligned}$$

The theorem follows. Note that  $r'(0) = 0$  from (B4) since  $|\hat{w}|$  is symmetric.

THEOREM B2. Let  $w$  be absolutely continuous on  $[-A, A]$  and 0 otherwise. Then  $r$  has left and right derivatives at 0 and

$$(B.6) \quad r_+'(0) = -r'_-(0) = -\frac{1}{2}(w^2(A) + w^2(-A)).$$

PROOF. Write, for  $h > 0$ ,

$$(B.7) \quad \begin{aligned} \int \frac{w(s+h) - w(s)}{h} w(s) ds \\ &= \int_{-A-h}^A \left[ \frac{1}{h} \int_s^{s+h} w'(z) dz \right] w(s) ds - \frac{1}{h} \int_{A-h}^A w^2(s) ds \\ &\rightarrow \int_{-A}^A w'(s)w(s) ds - w^2(A) = -\frac{1}{2}(w^2(A) + w^2(-A)) \end{aligned}$$

by arguing as in Theorem A1 and using Lebesgue's theorem. Since  $r(-t) = r(t)$  the result follows.

THEOREM B3. (Pickands) If  $H_\alpha(n, a)$ ,  $\bar{H}_\alpha(n)$  are defined as in (A.17), (A.19) then (A.20) holds.

PROOF. Suppose first that  $0 < \alpha < 2$ . Let for  $\gamma > 0$ ,

$$(B8) \quad \bar{H}_\alpha(n, \gamma) = \int_{-\infty}^{\infty} e^s [\max_{0 \leq t \leq n} Y(t) > s + \gamma] ds = e^{-\gamma} \bar{H}_\alpha(n).$$

Then

$$\begin{aligned}
 & \frac{1}{n} |H_\alpha(n, a) - \bar{H}_\alpha(na, \gamma)| \\
 & \leq \frac{1}{n} \left[ \int_{-\infty}^{\infty} e^s P[\max_{0 \leq t \leq na} Y(t) > s + \gamma, \max_{0 \leq k \leq n} Y(ka) \leq s] ds \right. \\
 (B.9) \quad & \quad \left. + \int_{-\infty}^{\infty} e^s P[s < \max_{0 \leq t \leq na} Y(t) \leq s + \gamma] ds \right] \\
 & \leq \frac{1}{n} \sum_{k=0}^{n-1} \int_{-\infty}^{\infty} e^s P[Y(ka) \leq s, \max_{ka \leq t \leq (k+1)a} Y(t) > s + \gamma] ds \\
 & \quad + \frac{1}{n} [\bar{H}_\alpha(na) - \bar{H}_\alpha(na, \gamma)].
 \end{aligned}$$

If the summands on the right of the first term of (B.9) are denoted by  $A(k, \gamma, a)$  then,

$$\begin{aligned}
 (B.10) \quad A(k, \gamma, a) &= \int_{-\infty}^{\infty} e^s \int_{-\infty}^s \tau(z, ka) \\
 &\quad \times P[\max_{0 \leq t \leq a} Y(t + ka) > s + \gamma \mid Y(ka) = z] dz ds
 \end{aligned}$$

where  $\tau(z, ka)$  is the density of  $Y(ka)$ . After some manipulation we obtain

$$\begin{aligned}
 (B.11) \quad A(k, \gamma, a) &= \int_{-\infty}^{\infty} \phi(w) \int_0^{\infty} e^s P[\max_{0 \leq t \leq a} (Y(t + ka) - Y(ka)) > s + \gamma \mid Y(ka) \\
 &= w + (ka)^\alpha] ds dw.
 \end{aligned}$$

As  $k \rightarrow \infty$ , the finite dimensional conditional distributions of  $Y(t + ka) - Y(ka)$  given  $Y(ka) = w + (ka)^\alpha$  tend for each  $w$  to those of  $Y(t)$ ,  $0 \leq t \leq a$ . Arguing as in Lemma A1 we conclude that,

$$(B.12) \quad \lim_k A(k, \gamma, a) = A(\gamma, a) = \int_0^{\infty} e^s P[\max_{0 \leq t \leq a} Y(t) > s + \gamma] ds.$$

Let  $Y^*(t) = Y(t) + |t|^\alpha$ . Then,

$$\begin{aligned}
 A(\gamma, a) &\leq \int_0^{\infty} e^s P[\max_{0 \leq t \leq a} Y^*(t) > s + \gamma] ds \\
 (B.13) \quad &= \int_0^{\infty} e^s P[\max_{0 \leq t \leq 1} Y^*(t) > (s + \gamma)a^{-\alpha/2}] ds \\
 &= a^{\alpha/2} e^{-\gamma} \int_{\gamma a^{-\alpha/2}}^{\infty} e^{-w a^{\alpha/2}} P[\max_{0 \leq t \leq 1} Y^*(t) > w] dw.
 \end{aligned}$$

Applying Fernique's estimate the right-hand side of (B.13) is  $O(\exp -a^{-\alpha/2})$  for every  $\gamma > 0$ . We conclude that,

$$\begin{aligned}
 (B.14) \quad & \limsup_a \limsup_n \frac{1}{na} |H_\alpha(n, a) - \bar{H}_\alpha(na, \gamma)| \\
 & \leq (1 - e^{-\gamma}) \limsup_a \limsup_n \frac{\bar{H}_\alpha(na)}{na}
 \end{aligned}$$

for every  $\gamma > 0$ . Since,

$$\begin{aligned}
 (B.15) \quad & P[\max_{0 \leq t \leq n} Y(t) > s] \\
 & \leq \sum_{k=0}^{n-1} P[\max_{k \leq t \leq k+1} Y(t) > s] \\
 & \leq \sum_{k=0}^{n-1} \{P[Y(k) \leq s, \max_{k \leq t \leq k+1} Y(t) > s] + P[Y(k) > s]\},
 \end{aligned}$$

it is easy to see that,

$$(B.16) \quad \sup_{x \geq 1} \frac{\bar{H}_\alpha(x)}{x} < \infty.$$

Hence,

$$(B.17) \quad \lim_a \limsup_n \frac{1}{na} |H_\alpha(n, a) - \bar{H}_\alpha(na)| = 0.$$

But from the argument of Lemma A3 it is clear that for every  $a > 0$ ,

$$(B.18) \quad \limsup_n \frac{H_\alpha(n, a)}{na} \leq \liminf_n \frac{\bar{H}_\alpha(na)}{na}.$$

The theorem follows for  $0 < \alpha < 2$ . For  $\alpha = 2$  we can use the representation  $Y(t) = 2^{\frac{1}{2}}tZ - t^2$  where  $Z$  is a standard normal deviate. Evidently,

$$(B.19) \quad \begin{aligned} \max_{0 \leq s \leq na/2^{\frac{1}{2}}} Y(s) &= \frac{Z^2}{2} \quad \text{if } 0 \leq Z < \frac{na}{2^{\frac{1}{2}}} \\ &= naZ - \frac{n^2a^2}{2} \quad \text{otherwise.} \end{aligned}$$

It follows that,

$$(B.20) \quad \begin{aligned} \frac{1}{na} \left| \bar{H}_2\left(\frac{na}{2^{\frac{1}{2}}}\right) - H_2\left(n, \frac{a}{2^{\frac{1}{2}}}\right) \right| \\ \leq \frac{1}{na} \int_0^{n^2a^2} e^{s/2} P[s^{\frac{1}{2}} < Z < (s + a^2)^{\frac{1}{2}}] ds \sim 2(1 - e^{-a^2/2}) \end{aligned}$$

by standard arguments. The theorem now follows generally.

#### REFERENCES

- [1] BERMAN, S. N. (1971). Asymptotic independence of the numbers of high and low level crossings of stationary Gaussian processes. *Ann. Math. Statist.* **42** 927–946.
- [2] BERMAN, S. N. (1971). Maxima and high level excursions of stationary Gaussian processes. *Trans. Amer. Math. Soc.* To appear.
- [3] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- [4] BREIMAN, L. (1969). *Probability*. Addison-Wesley, Reading.
- [5] BRILLINGER, D. (1969). An asymptotic representation of the sample distribution function. *Bull. Amer. Math. Soc.* **75** 545–547.
- [6] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.
- [7] EPANECHNIKOV, V. A. (1969). Nonparametric estimates of a multivariate probability density. *Theor. Probability Appl.* **14** 153–158.
- [8] FERNIQUE, X. (1970). Intégrabilité des vecteurs gaussiens. *C.R. Acad. Sci. Paris* **270** A 1698–99.
- [9] HAJEK, J. and SIDAK, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.
- [10] MANN, H. B. and WALD, A. (1942). On the choice of the number of class intervals in the application of the  $\chi^2$  test. *Ann. Math. Statist.* **13** 306–317.
- [11] PICKANDS, J. S. III (1960). Upcrossing probabilities for Gaussian processes. *Trans. Amer. Math. Soc.* **145** 51–73.
- [12] QUALLS, C. and WATANABE, H. (1971). Asymptotic properties of Gaussian processes. Tech. Report No. 736, Univ. of North Carolina, Chapel Hill.

- [13] ROSENBLATT M. (1971). Curve estimates. *Ann. Math. Statist.* **42** 1815-1842.
- [14] STECK, G. P. (1957). Limit theorems for conditional distributions. *Univ. California Publ. Statist.* **2** 237-284.
- [15] WATSON, G. S. (1967). Some problems in the statistics of directions. *Bull. I.S.I.* **17** 374-385.
- [16] WOODROOFE, M. (1967). On the maximum deviation of the sample density. *Ann. Math. Statist.* **38** 475-481.

DEPARTMENT OF STATISTICS  
UNIVERSITY OF CALIFORNIA  
BERKELEY, CALIFORNIA 94720

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF CALIFORNIA  
LA JOLLA, CALIFORNIA 92037