The Annals of Statistics
1999, Vol. 27, No. 1, 24–35

THE LIMITING BEHAVIOR OF A MODIFIED MAXIMAL SYMMETRIC 2s-SPACING WITH APPLICATIONS[1]

BY JAN W. H. SWANEPOEL

Potchefstroom University for CHE

This paper presents a solution to an open problem in astrophysics, namely that of estimating nonparametrically the *strength* of the pulsed signal in a series of high-energy photon arrival times. The newly proposed estimator, based on a modified maximal symmetric $2s$-spacing, is shown to be strongly consistent and asymptotically normally distributed, and a Monte Carlo study shows that its small and moderate sample behavior is very satisfactory. Additionally, new results regarding the weak and strong limiting behavior of modified maximal $2s$-spacings are derived.

1. Introduction and main results. New results on the behavior of modified maximal symmetric $2s$-spacings will be derived and applied to solve an open problem in astrophysics, namely that of estimating the *strength* of the pulsed signal in a series of high-energy photon arrival times nonparametrically [see, e.g., Swanepoel, De Beer and Loots (1996)]. A typical data set consists of a sequence of arrival times, each arrival time representing either noise or pulsed radiation. The unknown periodic density f of the folded (modulo 1) arrival times (f is often called a *light curve* in the astrophysical literature) can be represented as

$$f(x) = 1 - p + pf_S(x), \qquad 0 < x < 1, \tag{1.1}$$

where p, $0 < p < 1$, is the unknown strength of the pulsed signal and f_S is the unknown *source function* that characterizes the radiation pattern of the source (pulsar).

It is usually assumed that

$$\min_{0<x<1} f_S(x) = 0,$$

so that the estimation of p can be reduced to the estimation of the minimum value of an unknown density function with bounded support, namely,

$$f(\theta) = \min_{0<x<1} f(x).$$

In (1.8) we propose a nonparametric estimator $\widehat{f(\theta)}$ for $f(\theta)$, and an estimator for p can therefore be defined by

$$\hat{p}_n = \max\left\{0, 1 - \widehat{f(\theta)}\right\}. \tag{1.2}$$

Received December 1995; revised August 1998.
[1] Supported by the South African Foundation for Research Development.
AMS 1991 *subject classifications*. 60F05, 60F15.
Key words and phrases. Limiting distributions, nonparametric estimators, order statistics, spacings, strong laws, weak convergence.

In order to define $\widehat{f(\theta)}$, we first introduce some general notation and discuss some known results regarding the behavior of maximal spacings. Let $X_1, X_2, \dots,$ be a sequence of independent and identically distributed random variables on some probability space $(\Omega, \mathscr{F}, P)$ with unknown univariate distribution function F, which is assumed to be absolutely continuous (with respect to Lebesgue measure) with density function f [not necessarily belonging to the class defined by (1.1)]. Throughout the discussion below we assume that the following two conditions hold:

(1.3) For some finite constants a and b, $a < b$, $f(x) > 0$ for all $x \in (a, b)$ and $f(x) = 0$ otherwise.

(1.4) There exists a constant $\theta \in (a, b)$ such that, for all $x \in (a, b)$, $x \neq \theta$, $f(x) > f(\theta) > 0$.

Denote the order statistics of $X_1, X_2, \dots, X_n$ by

$$Y_1 \le Y_2 \le \cdots \le Y_n.$$

Let $\{s_n\}$ be a nonrandom sequence of positive integers. The maximal s_n-spacing is defined for $1 \le s_n \le n - 1$ by

$$M_n = \max_{1 \le j \le n - s_n} \left(Y_{j+s_n} - Y_j\right). \tag{1.5}$$

A great deal is known about the behavior of M_n when $s_n \equiv 1$ and F is the uniform distribution function on (0, 1). For example, Devroye (1981, 1982) and Deheuvels (1982, 1983) derived laws of the iterated logarithm for M_n. Deheuvels and Devroye (1984) obtained analogous results if $s_n \to \infty$ at certain rates. However, few results are available when F is arbitrary. For $s_n \equiv 1$, Deheuvels (1984) derived strong limiting bounds for M_n. He pointed out, among others, that if F has a continuous density f, the major influence on the behavior of maximal spacings is exerted by the behavior of f in the neighborhood of its minimum. Under the assumption that Y_1 and Y_n belong to the domain of attraction of extreme-value distributions and that $s_n \equiv 1$, Deheuvels (1986) showed that the weak limiting behavior of Y_1 and Y_n characterizes completely the weak limiting behavior of M_n and he also obtained the corresponding limiting nonnormal distributions. Also, Barbe (1992) proved that M_n (appropriately standardized) converges in distribution to a Gumbel distribution if it is assumed, among others, that the density f has a positive minimum and $s_n \equiv 1$. The weak limiting behavior of M_n is related to the minimum of the density function and to the local behavior of the density function near its minimum, as is the case for the almost sure behavior of M_n [Barbe (1992)].

In this paper we investigate the strong and weak limiting behavior of a statistic V_n which is derived by modifying M_n as follows. Suppose that $2s_n \le n - 1$ for each $n = 1, 2, \dots,$ and $s_n \uparrow \infty$ as $n \to \infty$. For each n, let K_n be a positive integer-valued random variable defined by

$$Y_{K_n+s_n} - Y_{K_n-s_n} = \max_{s_n+1 \le j \le n-s_n} \left(Y_{j+s_n} - Y_{j-s_n}\right). \tag{1.6}$$

Since F is absolutely continuous, K_n is almost surely (a.s.) unique and the maximal symmetric $2s_n$-spacing $Y_{K_n+s_n} - Y_{K_n-s_n}$ is positive, a.s. Let $\{r_n\}$ be another nonrandom sequence of positive integers such that $r_n \le s_n$ for all $n = 1, 2, \dots,$ and $r_n \uparrow \infty$ as $n \to \infty$. Now, define

$$V_n = Y_{K_n+r_n} - Y_{K_n-r_n}. \tag{1.7}$$

A strong law of large numbers is proved for V_n in Theorem 1.1 below and it is shown to have an asymptotic normal distribution in Theorem 1.3. This contrasts sharply with the limiting nonnormal distributions derived in the literature for M_n.

We now suggest estimating $f(\theta) = \min_{a<x<b} f(x)$ by

$$\widehat{f(\theta)} = 2r_n(nV_n)^{-1}. \tag{1.8}$$

It follows immediately from (1.2) and Theorem 1.1 below that $\hat{p}_n \to p$ a.s. as $n \to \infty$. Also, from Theorem 1.3 and Slutsky's theorem we deduce that as $n \to \infty$,

$$(2r_n)^{1/2}\{\hat{p}_n - p\} \to_d N\left(-3^{-1}(k/2)^{1/2} f''(\theta)(f(\theta))^{-2}, (f(\theta))^2\right),$$

where k is defined by (1.18), and $\to_d$ denotes convergence in distribution.

The performance of $\hat{p}_n$ on real astrophysical data is extensively discussed by Swanepoel, De Beer and Loots (1996). It is pointed out, among other things, that an efficiency is achieved which is very close to a well-known norm set by astrophysicists. This norm was not nearly achieved by previous estimation procedures. An obvious alternative estimator of p, say $\tilde{p}_n$, is obtained if we replace $\widehat{f(\theta)}$ in (1.2) by the minimum value of a kernel density estimator. However, the Monte Carlo study of Section 5 shows that $\hat{p}_n$ performs substantially better than $\tilde{p}_n$ for small and moderate sample sizes.

Our first theorem shows that a strong law of large numbers holds for V_n defined by (1.7).

THEOREM 1.1. *Assume the following conditions hold*:

(1.9) *f is continuous in some neighborhood of θ*;
(1.10) $r_n/n \downarrow 0$, $s_n/n \downarrow 0$, $r_n/\log n \to \infty$ *as* $n \to \infty$.

Then, as $n \to \infty$,

$$f(\theta)n(2r_n)^{-1}V_n \to 1 \quad a.s.\ P.$$

In order to derive a limiting distribution for V_n, we need to know the limiting distribution of the discrete random variable K_n defined by (1.6). We state this as a theorem below, since the result is of independent interest. First, a stochastic process is defined which is needed for the formulation of the theorem.

Suppose that for each constant C, $0 < C \le \infty$, $\{Z_C(t), -\infty < t < \infty\}$ is a Gaussian process, originating from zero, with expectation zero and covariance function given by

$$\begin{aligned}(1.11)\qquad &\operatorname{Cov}\{Z_C(t), Z_C(t^*)\}\\ &\quad = B\{\min(|t|, 2C^{1/3}) + \min(|t^*|, 2C^{1/3}) - \min(|t - t^*|, 2C^{1/3})\},\end{aligned}$$

where

$$B = (f(\theta))^6/(f''(\theta))^2.$$

If $C = \infty$, $\{Z_\infty(t)\}$ is a two-sided Wiener–Levy process, which is defined as follows. Let $\{W_1(t), t \ge 0\}$ and $\{W_2(t), t \ge 0\}$ be two independent standard Wiener processes such that $W_1(0) = W_2(0) = 0$ a.s. Then,

$$(1.12)\qquad Z_\infty(t) = \begin{cases} W_1(2Bt), & \text{if } t \ge 0,\\ W_2(-2Bt), & \text{if } t < 0.\end{cases}$$

In this case the covariance function of $\{Z_\infty(t)\}$ is the limiting covariance function of $\{Z_C(t)\}$ as $C \to \infty$, namely,

$$\begin{aligned}(1.13)\qquad &\operatorname{Cov}\{Z_\infty(t), Z_\infty(t^*)\}\\ &\quad = 2B\min(|t|, |t^*|)\{I(t \ge 0, t^* \ge 0) + I(t < 0, t^* < 0)\},\end{aligned}$$

where $I(\cdot)$ denotes the indicator function.

THEOREM 1.2. *Suppose the following conditions hold*:

(1.14) *f has a bounded second derivative f'' in some neighborhood of θ, with $0 < f''(\theta) < \infty$, and f'' satisfies a Lipschitz condition of order α $(0 < \alpha \le 1)$ at θ*;
(1.15) $n^{-(3\alpha+8)} s_n^{(3\alpha+10)} \to 0$ *as* $n \to \infty$;
(1.16) $n^{-4} s_n^5 \to C$, *for some constant* $C, 0 < C \le \infty$.

Then, *as* $n \to \infty$,

$$n^{-1/3} s_n^{2/3}\big(n^{-1}K_n - F(\theta)\big) \to_d T,$$

where T is a random variable which maximizes the process

$$\{Z_C(t) - t^2, -\infty < t < \infty\}.$$

Contrary to the limiting nonnormal distribution derived up to now in the literature for M_n defined by (1.5) (usually some extreme-value distribution), the next theorem states that V_n has in fact an asymptotic normal distribution.

THEOREM 1.3. *Suppose* (1.14), (1.15) *and the following conditions hold*:

(1.17) $n^{-4} s_n^5 \to \infty$ *as* $n \to \infty$;
(1.18) $n^{-4} r_n^5 \to k$ *as* $n \to \infty$, *for some constant* k, $0 \le k < \infty$;
(1.19) $n^4 s_n^{-2} r_n^{-3} \to 0$ *as* $n \to \infty$.

Then, as $n \to \infty$, we have

$$(2r_n)^{1/2}\big\{f(\theta)n(2r_n)^{-1}V_n - 1\big\} \to_d N\big(-3^{-1}(k/2)^{1/2}f''(\theta)(f(\theta))^{-3}, 1\big).$$

REMARK. For inference purposes, a possible estimator for $f''(\theta)$ is $\hat{f}_h''(\hat{\theta}_n)$, where $\hat{f}_h''(\cdot)$ is the second derivative of the kernel estimator (with bandwidth h) defined in Section 5 and in view of (2.4) below, $\hat{\theta}_n$ can, for example, be chosen as

$$\hat{\theta}_n = \tfrac{1}{2}\big(Y_{K_n - s_n} + Y_{K_n + s_n}\big) \quad \text{or } \hat{\theta}_n = Y_{K_n}.$$

2. Proof of Theorem 1.1. Define the inverse of a distribution function H by $H^{-1}(t) = \inf\{x\colon H(x) \ge t\}$. Without loss of generality we assume that $X_i = F^{-1}(U_i)$, where $U_1, \dots, U_n$ is a sequence of independent uniform (0, 1) distributed random variables. Let $G_n(t) = n^{-1}\sum_{i=1}^n I(U_i \le t)$ and write $Q_n(t) = G_n^{-1}(t)$, the so-called *sample quantile function* based on $\{U_i\}$. For convenience, we use $\circ$ to denote composition; for example, $f \circ F^{-1}$ means $f(F^{-1})$. Then,

$$\begin{aligned}
&Y_{K_n+s_n} - Y_{K_n-s_n}\\
&\quad= \max_{(s_n+1)/n \le t \le 1-s_n/n} F^{-1}\circ Q_n(t+s_n/n) - F^{-1}\circ Q_n(t-s_n/n)\\
&\quad= \max_{(s_n+1)/n \le t \le 1-s_n/n} \left\{\frac{Q_n(t+s_n/n) - Q_n(t-s_n/n) - 2s_n/n}{f\circ F^{-1}(t+o(1))} + \frac{2s_n/n}{f\circ F^{-1}(t+o(1))}\right\},
\end{aligned} \tag{2.1}$$

where $o(1)$ converges to zero a.s. and uniformly in t [which follows from the fact that $\sup_t|Q_n(t) - t| = o(1)$ a.s.].

Under the conditions $s_n/\log n \to \infty$, $s_n/n \downarrow 0$ and $\log(n/s_n)/\log\log n \to \infty$ as $n \to \infty$, Mason (1984) proved that

$$\lim_{n\to\infty} \max_{(s_n+1)/n \le t \le 1-s_n/n} \frac{n|Q_n(t+s_n/n) - Q_n(t-s_n/n) - 2s_n/n|}{\sqrt{4s_n \log(n/2s_n)}} = 1 \quad \text{a.s.},$$

from which we deduce that

$$\lim_{n\to\infty} \max_{(s_n+1)/n \le t \le 1-s_n/n} \frac{n|Q_n(t+s_n/n) - Q_n(t-s_n/n) - 2s_n/n|}{2s_n} = 0 \quad \text{a.s.} \tag{2.2}$$

It is easy to check [using, e.g., the results of Komlós, Major and Tusnády (1976)] that (2.2) still holds if we drop the condition $\log(n/s_n)/\log\log n \to \infty$

assumed above. Hence, from assumption (1.9), (2.1) and (2.2), it follows that

$$\lim_{n\to\infty} \frac{n\left(Y_{K_n+s_n} - Y_{K_n-s_n}\right)}{2s_n} = \sup_{0<t<1} \frac{1}{f\circ F^{-1}(t)} = 1/f(\theta) \quad \text{a.s.} \tag{2.3}$$

Furthermore, we claim that

$$\lim_{n\to\infty} \frac{K_n}{n} = F(\theta) \quad \text{a.s.,} \tag{2.4}$$

which can be proved as follows. Suppose (2.4) does not hold. Then there is a subsequence $\{n(i)\}$ such that $\lim_{i\to\infty} K_{n(i)}/n(i) = F(\theta_1) \neq F(\theta)$, for some constant $\theta_1 \neq \theta$. By the same arguments as above, we now obtain

$$\begin{aligned}
&\lim_{i\to\infty} \frac{n(i)\left(Y_{K_{n(i)}+s_{n(i)}} - Y_{K_{n(i)}-s_{n(i)}}\right)}{2s_{n(i)}}\\
&\quad = \lim_{i\to\infty} \frac{n(i)\left\{F^{-1}\circ Q_{n(i)}\left(\left(K_{n(i)}+s_{n(i)}\right)/n(i)\right) - F^{-1}\circ Q_{n(i)}\left(\left(K_{n(i)}-s_{n(i)}\right)/n(i)\right)\right\}}{2s_{n(i)}}\\
&\quad = 1/f(\theta_1) \neq 1/f(\theta),
\end{aligned}$$

which is in contradiction with (2.3).

Finally, if $r_n \to \infty$,

$$n(2r_n)^{-1}V_n = \frac{n(2r_n)^{-1}\{Q_n((K_n + r_n)/n) - Q_n((K_n - r_n)/n)\}}{f\circ F^{-1}(F(\theta) + o(1))} \quad \text{a.s.,}$$

and the proof of the theorem now follows from (1.10) and by using (2.2) once more. □

3. Proof of Theorem 1.2. Notice that the first three derivatives of F^{-1} are given by

$$\begin{gathered}
(F^{-1})' = 1/f\circ F^{-1}, (F^{-1})'' = -(f'/f^3)\circ F^{-1},\\
(F^{-1})''' = -(f''/f^4)\circ F^{-1} + 3(f'/f^5)\circ F^{-1},
\end{gathered}$$

and in particular $(F^{-1})''(F(\theta)) = 0$, $(F^{-1})'''(F(\theta)) = -f''(\theta)/(f(\theta))^4$. Now, set

$$\begin{aligned}
\xi_n(t) = {} & F^{-1}(F(\theta) + t + s_n/n) - F^{-1}(F(\theta) + t - s_n/n)\\
& - F^{-1}(F(\theta) + s_n/n) + F^{-1}(F(\theta) - s_n/n)\\
& + \left(F^{-1}\circ Q_n - F^{-1}\right)(F(\theta) + t + s_n/n)\\
& - \left(F^{-1}\circ Q_n - F^{-1}\right)(F(\theta) + t - s_n/n)\\
& - \left(F^{-1}\circ Q_n - F^{-1}\right)(F(\theta) + s_n/n)\\
& + \left(F^{-1}\circ Q_n - F^{-1}\right)(F(\theta) - s_n/n).
\end{aligned}$$

Note that $|K_n/n - F(\theta) - \arg\max_t \xi_n(t)| \le 1/n$. Expanding F^{-1} up to order three and using (1.14), it readily follows that

$$
\begin{aligned}
&F^{-1}(F(\theta) + t + s_n/n) - F^{-1}(F(\theta) + t - s_n/n) \\
(3.1) \qquad &= (2s_n)/(nf(\theta)) - f''(\theta)(s_n t^2/n + s_n^3/(3n^3))/(f(\theta))^4 \\
&\quad + O(|t| + s_n/n)^{3+\alpha}.
\end{aligned}
$$

Since we are assuming that $f(\theta) > 0$ and $f''(x)$ exists in some neighborhood of θ, we deduce from Theorem 6 of Csőrgő and Révész (1978) that for each n there exists a Brownian bridge $\{B_n(t);\ 0 \le t \le 1\}$ such that

$$
\begin{aligned}
(3.2) \qquad &\sup_t |n^{1/2}(F^{-1} \circ Q_n - F^{-1})(t) - (B_n/f \circ F^{-1})(t)| \\
&\quad = O(n^{-1/2} \log n) \quad \text{a.s.},
\end{aligned}
$$

where the supremum is taken over t belonging to a fixed neighborhood of $F(\theta)$, with $0 < F(\theta) < 1$.

For brevity, write $B_n = B$. Recalling that $B(t) = W(t) - tW(1)$, where $\{W(t)\}$ is a standard Wiener process, we obtain from (3.1) and (3.2),

$$
\begin{aligned}
\xi_n(t) = {}& -f''(\theta) s_n t^2/(n(f(\theta))^4) \\
(3.3) \qquad &+ n^{-1/2}\{W(F(\theta) + t + s_n/n) - W(F(\theta) + t - s_n/n) \\
&\qquad - W(F(\theta) + s_n/n) + W(F(\theta) - s_n/n)\}/f(\theta) \\
&+ n^{-1/2} O_P(|t| + s_n/n)^2 + O_P(n^{-1} \log n) + O(|t| + s_n/n)^{3+\alpha}.
\end{aligned}
$$

Now, the asymptotic behavior of K_n/n can be easily analyzed. We have a quadratic drift $-f''(\theta)s_n t^2/n(f(\theta))^4$ of order $s_n t^2/n$ plus a Gaussian process which is of order $(t/n)^{1/2}$ [since $W(F(\theta) + t + s_n/n) - W(F(\theta) + s_n/n) =_d W(t)$ which is of order $t^{1/2}$]. To maximize $\xi_n(t)$, the quadratic drift should be of the same order as the Gaussian process [see also Kim and Pollard (1990)], which implies that t should be of order $n^{1/3}s_n^{-2/3}$.

Set $t = n^{1/3}s_n^{-2/3}\tau$, for some finite constant τ. From (1.16) we have $n^{-4}s_n^5 \to C$, for some constant C, $0 < C \le \infty$. First, consider the case when C is finite. Applying (1.15), (1.16) and the fact that $W(F(\theta) + \cdot\,) - W(F(\theta)) =_d W(\cdot)$ and $W(a\,\cdot\,) =_d (|a|)^{1/2}W(\cdot)$ (for any finite constant a), it easily follows from (3.3) that

$$
\begin{aligned}
Z_n(\tau) &:= (f(\theta))^4 n^{1/3} s_n^{1/3} \xi_n(n^{1/3}s_n^{-2/3}\tau)/f''(\theta) \\
&=_d -\tau^2 + ((f(\theta))^3/f''(\theta)) n^{-1/6} s_n^{1/3} \\
&\qquad \times \{W((C^{-2/15}\tau + C^{1/5})n^{-1/5}) - W((C^{-2/15}\tau - C^{1/5})n^{-1/5}) \\
&\qquad\qquad - W(C^{1/5}n^{-1/5}) + W(-C^{1/5}n^{-1/5})\} + o_P(1)
\end{aligned}
$$

$$
\begin{aligned}
&=_d -\tau^2 + \left(\left(f(\theta)\right)^3/f''(\theta)\right)\{W(\tau + C^{1/3}) - W(\tau - C^{1/3}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad -W(C^{1/3}) + W(-C^{1/3})\} + o_P(1) \\
&=: -\tau^2 + Z_C(\tau) + o_P(1) \\
&=: Z_C^0(\tau) + o_P(1),
\end{aligned}
$$

where $o_P(1) \to 0$ in probability, uniformly in $\tau \in [-\tau_0, \tau_0]$, for any finite constant $\tau_0 > 0$. Note that in the definition of $Z_C(\tau)$, $W(t)$ is defined as $W(t) = W_1(t)$ if $t \geq 0$, and $W(t) = W_2(-t)$ if $t < 0$, where $\{W_1(t), t \geq 0\}$ and $\{W_2(t), t \geq 0\}$ are two independent standard Wiener processes such that $W_1(0) = W_2(0) = 0$ a.s. It is easy to check that the covariance function of $\{Z_C(\tau)\}$ is given by (1.11). If $C = \infty$, we obtain as above, by using (1.16), that $Z_n(\tau) =_d -\tau^2 + Z_\infty(\tau) + o_P(1)$, where $Z_\infty(\tau)$ is now defined by (1.12) with the covariance function given by (1.13).

At this point it has been shown that, for any constant $0 < C \leq \infty$, $Z_n \to Z_C^0$ weakly on $D[-\tau_0, \tau_0]$ as $n \to \infty$, where $D[-\tau_0, \tau_0]$ is the space of real-valued functions defined on $[-\tau_0, \tau_0]$ that are right-continuous and have left-hand limits. Since τ_0 was arbitrary, it follows from Billingsley (1968) that the weak convergence result holds for all $\tau \in (-\infty, \infty)$. Now, for $x \in D(-\infty, \infty)$, let

$$h(x) = \min\left\{t\colon\ x(t) = \max_\tau x(\tau)\right\},$$

and notice that as $n \to \infty$,

$$
\begin{aligned}
n^{-1/3}s_n^{2/3}\left(n^{-1}K_n - F(\theta)\right) &= \arg\max_\tau \xi_n\left(n^{1/3}s_n^{-2/3}\tau\right) + O\left(n^{-4/3}s_n^{2/3}\right) \\
&= h(Z_n) + o(1).
\end{aligned}
$$

Hence, to complete the proof of the theorem, one uses the continuous mapping theorem to conclude that $h(Z_n) \to_d h(Z_C^0) = T$ as $n \to \infty$ [Chernoff (1964) and Groeneboom (1989) proved that $T < \infty$ a.s.]. □

4. Proof of Theorem 1.3. From Theorem 1.2 it follows that

$$
\begin{aligned}
n^{-1}K_n &= F(\theta) + O_P\left(n^{1/3}s_n^{-2/3}\right) \\
&=: F(\theta) + t_n.
\end{aligned}
\tag{4.1}
$$

Using the notation of Sections 1–3, it easily follows, by applying (1.14), (4.1) and the techniques required to derive (3.1) and (3.3), that

$$
\begin{aligned}
V_n &= Y_{K_n + r_n} - Y_{K_n - r_n} \\
&=_d F^{-1}\left(F(\theta) + t_n + r_n/n\right) - F^{-1}\left(F(\theta) + t_n - r_n/n\right) \\
&\quad + \left(F^{-1} \circ Q_n - F^{-1}\right)\left(F(\theta) + t_n + r_n/n\right) \\
&\quad - \left(F^{-1} \circ Q_n - F^{-1}\right)\left(F(\theta) + t_n - r_n/n\right)
\end{aligned}
$$

$$
\begin{aligned}
&= (2r_n)/(nf(\theta)) - f''(\theta)\big(r_n t_n^2/n + r_n^3/(3n^3)\big)/(f(\theta))^4 \\
&\quad + n^{-1/2}\{W(F(\theta) + t_n + r_n/n) - W(F(\theta) + t_n - r_n/n)\}/f(\theta) \\
&\quad + n^{-1/2}O_P(|t_n| + r_n/n)^2 + O_P(n^{-1}\log n) + O(|t_n| + r_n/n)^{3+\alpha}.
\end{aligned}
$$

Thus, by using (1.15) and (1.17)–(1.19), it can easily be seen that

$$
\begin{aligned}
&(2r_n)^{1/2}\{f(\theta)n(2r_n)^{-1}V_n - 1\} \\
&\quad =_d (2r_n)^{1/2}\Big\{-f''(\theta)\big(3t_n^2 + r_n^2/n^2\big)/\big(6(f(\theta))^3\big) \\
&\qquad\qquad + n^{1/2}\big(W(F(\theta) + t_n + r_n/n) \\
&\qquad\qquad\qquad - W(F(\theta) + t_n - r_n/n)\big)/(2r_n)\Big\} \\
&\qquad + (n/r_n)^{1/2}O_P(|t_n| + r_n/n)^2 + O_P\big(r_n^{-1/2}\log n\big) \\
&\qquad + \big(n/r_n^{1/2}\big)O(|t_n| + r_n/n)^{3+\alpha} \\
&\quad =_d -3^{-1}(k/2)^{1/2}f''(\theta)(f(\theta))^{-3} + \{W_1(1/2) - W_2(1/2)\} + o_P(1),
\end{aligned}
$$

and the proof of the theorem follows by applying Slutsky's theorem. □

5. Simulation study. In this section we present the results of a limited Monte Carlo study which compares the performance of $\widehat{f(\theta)}$, defined by (1.8), with that of an obvious alternative estimator of $f(\theta) = \min_{0<x<1} f(x)$, namely the minimum value of a kernel density estimator. The kernel method, introduced by Rosenblatt (1956), is probably the most commonly used density estimation technique and is certainly the best understood mathematically. The kernel estimator of f is defined by

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

where K is the kernel function, usually chosen to be a known density function symmetric around zero. The smoothing parameter h, which depends on the sample size n, is often referred to as the bandwidth. Hence, an alternative estimator of $f(\theta)$ is

$$\tilde{f}_h := \min_{0<x<1} \hat{f}_h(x).$$

Since the density f is assumed to have bounded support, we took $K(\cdot)$ throughout as the Epanechnikov kernel $K(x) = 0.75(1 - x^2)$ for $-1 \le x \le 1$. To implement $\tilde{f}_h$, we chose h to be the asymptotically optimal global bandwidth h_0 that minimizes the mean integrated squared error $E\int(\hat{f}_h(x) - f(x))^2\,dx$ asymptotically as $n \to \infty$. Other choices of h, for example data-driven bandwidths, are discussed in De Beer, Loots and Swanepoel (1999). It is well known [see, e.g., Wand and Jones (1995), page 22] that

$$h_0 = \left\{15\Big/\int_0^1 (f''(x))^2\,dx\right\}^{1/5} n^{-1/5}.$$

Furthermore, suppose $r_n^5 = kn^4$ for some constant $k > 0$. From Theorems 1.1 and 1.3 it follows, by using Slutsky's theorem, that as $n \to \infty$,

$$\widehat{f(\theta)} - f(\theta) \sim N\Big(k^{1/2}\gamma_1/(2r_n)^{1/2}, \gamma_2/(2r_n)\Big),$$

where

$$\gamma_1 = (1/18)^{1/2} f''(\theta)(f(\theta))^{-2}, \qquad \gamma_2 = (f(\theta))^2.$$

Hence, if for example the sequence $\{n^{4/5}(\widehat{f(\theta)} - f(\theta))^2\}$ is uniformly integrable (which will be the case under certain conditions), it follows from Theorem 5.4 of Billingsley (1968) that the mean-squared error of $\widehat{f(\theta)}$ can be approximated for large n by

$$\begin{aligned} E\Big\{\Big(\widehat{f(\theta)} - f(\theta)\Big)^2\Big\} &\sim \big(\gamma_2 + k\gamma_1^2\big)/(2r_n) \\ &= (1/2)n^{-4/5}\big(\gamma_2 k^{-1/5} + \gamma_1^2 k^{4/5}\big), \end{aligned}$$

which is minimized if k is chosen as

$$k_0 = \gamma_2/\big(4\gamma_1^2\big) = (9/2)(f(\theta))^6(f''(\theta))^{-2}.$$

This choice of k_0 yields

$$E\Big\{\Big(\widehat{f(\theta)} - f(\theta)\Big)^2\Big\} \sim 3^{-2/5}(2^{-4/5} + 2^{-14/5})(f(\theta))^{4/5}(f''(\theta))^{2/5}n^{-4/5}.$$

Note that the $n^{-4/5}$ rate of convergence is the same as that of the mean integrated squared error of the kernel estimator $\hat{f}_{h_0}$. However, it is an open question whether the mean-squared error of $\tilde{f}_{h_0}$ converges to zero at this rate. To implement $\widehat{f(\theta)} = (2r_n)/(nV_n)$, we therefore chose

$$\begin{aligned} r_n &= \min\big\{\big[k_0^{1/5}n^{4/5}\big], \quad [(n-1)/2]\big\}, \\ s_n &= \min\big\{\big[k_0^{1/5}n^{4/5+\nu}\big], \quad [(n-1)/2]\big\}, \end{aligned}$$

where $\nu = 0.0001$, and $[z]$ is the largest integer less than or equal to z. For small and moderate sample sizes, our simulation studies showed that $\widehat{f(\theta)}$ is most efficient if ν is chosen very small.

The estimators $\widehat{f(\theta)}$ and $\tilde{f}_{h_0}$ were evaluated by empirically calculating their biases and root mean-squared errors for sample sizes $n = 50$, 100 and 150. The following class of densities was considered:

$$f(x) = (1-\varepsilon)(a+1)x^a + \varepsilon(a+1)(1-x)^a, \qquad 0 < x < 1,$$

where $0 < \varepsilon < 1$ and $a > 0$. Values of θ, $f(\theta)$ and $f''(\theta)$ are displayed in Table 1. Table 2 contains Monte Carlo estimates of BIAS_1, BIAS_2 [bias of $\widehat{f(\theta)}$ and $\tilde{f}_{h_0}$, respectively], RMSE_1, RMSE_2 [root mean-squared error of $\widehat{f(\theta)}$ and $\tilde{f}_{h_0}$, respectively] and, to facilitate comparison,

$$\mathrm{RATIO} = 100 \times \frac{\mathrm{RMSE}_1}{\mathrm{RMSE}_2}\%.$$

TABLE 1
Some characteristics of f

ε	a	θ	$f(\theta)$	$f''(\theta)$
0.3	2.0	0.30	0.63	6.0
0.3	3.0	0.40	0.44	11.0
0.5	2.0	0.50	0.75	6.0
0.5	3.0	0.50	0.50	12.0

Each entry in Table 2 was based on 10,000 independent trials. The estimated standard errors of the averages were found to be negligibly small and were therefore omitted from the table.

From Table 2 it is clear that $\widehat{f(\theta)}$ has, in all cases considered, smaller bias than $\tilde{f}_{h_0}$. The latter estimator also suffers from large variability. This results in a large RMSE in comparison to the RMSE of $\widehat{f(\theta)}$, as is also evident from the RATIOs. Similar results were obtained for other choices of ε, a and the sample size n, but will not be reported here.

De Beer, Loots and Swanepoel (1999) performed an extensive Monte Carlo study to evaluate the performance of $\widehat{f(\theta)}$ when $\{r_n\}$ and $\{s_n\}$ are chosen *data-dependently*. They derived two data-based methods of choosing $\{r_n\}$ and $\{s_n\}$ by applying bootstrap techniques. The above findings regarding the superiority of $\widehat{f(\theta)}$ were also confirmed by these authors in studies where other classes of density functions [including the class defined by (1.1)] and sample sizes were used.

Computations were performed using Fortran programs together with IMSL (Version 2.1) routines on an IBM RS6000 43P PowerPC. Fortran code for the computation of $\widehat{f(\theta)}$ can be obtained by request from the author.

TABLE 2
Monte Carlo estimates of bias and root mean-squared error

n	ε	a	BIAS_1	BIAS_2	RMSE_1	RMSE_2	RATIO
50	0.3	2.0	0.002	−0.265	0.076	0.281	27.0
	0.3	3.0	−0.017	−0.048	0.075	0.106	70.8
	0.5	2.0	−0.020	−0.250	0.077	0.261	29.5
	0.5	3.0	−0.013	0.014	0.082	0.083	98.8
100	0.3	2.0	0.002	−0.259	0.064	0.269	23.8
	0.3	3.0	−0.018	−0.028	0.065	0.078	83.3
	0.5	2.0	−0.016	−0.214	0.064	0.222	28.8
	0.5	3.0	−0.020	0.030	0.069	0.080	86.3
150	0.3	2.0	−0.009	−0.256	0.058	0.264	22.0
	0.3	3.0	−0.014	−0.019	0.057	0.064	89.1
	0.5	2.0	−0.015	−0.195	0.059	0.202	29.2
	0.5	3.0	−0.018	0.036	0.063	0.080	78.8

Acknowledgments. The author would like to express deep appreciation to C. F. de Beer for his very generous help with the computer calculations and thanks H. Loots, who was coauthor in the first version of the manuscript but decided to withdraw from the project during the revision, for helpful discussions and critical comments. The author also thanks the Editor, Associate Editor and two referees for excellent and detailed suggestions which led to a substantial improvement of the paper.

REFERENCES

BARBE, PH. (1992). Limiting distribution of the maximal spacing when the density function admits a positive minimum. *Statist. Probab. Lett.* **14** 53–60.

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

CHERNOFF, H. (1964). Estimation of the mode. *Ann. Inst. Statist. Math.* **16** 31–41.

CSŐRGŐ, M. and RÉVÉSZ, P. (1978). Strong approximations of the quantile process. *Ann. Statist.* **6** 882–894.

DE BEER, C. F., LOOTS, H. and SWANEPOEL, J. W. H. (1999). Nonparametric estimation of the minimum value of a density function. *Comput. Statist. Data Anal.* To appear.

DEHEUVELS, P. (1982). Strong limiting bounds for maximal uniform spacings. *Ann. Probab.* **10** 1058–1065.

DEHEUVELS, P. (1983). Upper bounds for kth maximal spacings. *Z. Wahrsch. Verw. Gebiete* **62** 465–474.

DEHEUVELS, P. (1984). Strong limit theorems for maximal spacings from a general univariate distribution. *Ann. Probab.* **12** 1181–1193.

DEHEUVELS, P. (1986). On the influence of the extremes of an i.i.d. sequence on the maximal spacings. *Ann. Probab.* **14** 194–208.

DEHEUVELS, P. and DEVROYE, L. (1984). Strong laws for the maximal k-spacing when $k \leq c \log n$. *Z. Wahrsch. Verw. Gebiete* **66** 315–334.

DEVROYE, L. (1981). Laws of the iterated logarithm for order statistics of uniform spacings. *Ann. Probab.* **9** 860–867.

DEVROYE, L. (1982). A log log law for maximal uniform spacings. *Ann. Probab.* **10** 863–868.

GROENEBOOM, P. (1989). Brownian motion with a parabolic drift and airy functions. *Probab. Theory Related Fields* **81** 79–109.

KIM, J. and POLLARD, D. (1990). Cube root asymptotics. *Ann. Statist.* **18** 191–219.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent R.V.'s and the sample DF, II. *Z. Wahrsch. Verw. Gebiete* **34** 33–58.

MASON, D. M. (1984). A strong limit theorem for the oscillation modulus of the uniform empirical quantile process. *Stochastic Process. Appl.* **17** 127–136.

ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27** 832–837.

SWANEPOEL, J. W. H., DE BEER, C. F. and LOOTS, H. (1996). Estimation of the strength of a periodic signal from photon arrival times. *Astrophys. J.* **467** 261–264.

WAND, M. P. and JONES, M. C. (1995). *Kernel Smoothing*. Chapman and Hall, London.

DEPARTMENT OF STATISTICS
POTCHEFSTROOM UNIVERSITY FOR CHE
POTCHEFSTROOM 2520
SOUTH AFRICA
E-MAIL: sttcjs@puknet.puk.ac.za