The Annals of Statistics
1981, Vol. 9, No. 2, 443–448

THE ASYMPTOTIC BEHAVIOR OF MONOTONE REGRESSION ESTIMATES[1]

BY F. T. WRIGHT

University of Missouri-Rolla

An estimator for a monotone regression function was proposed by Brunk. He has shown that if the underlying regression function has positive slope at a point, then, based on r observations, the difference of the regression function and its estimate at that point has a nondegenerate limiting distribution if this difference is multiplied by $r^{1/3}$. To understand how the behavior of the regression function at a point influences the asymptotic properties of the estimator at that point, we have generalized Brunk's result to points at which the regression function does not have positive slope. If the first $\alpha - 1$ derivatives of the regression function are zero at a point and the αth derivative is positive there, then the norming constants are of order $r^{\alpha/(2\alpha+1)}$.

1. Introduction. The estimate proposed by Brunk (1958) for a nondecreasing regression function is obtained by a "max-min" operation on sample means (cf. equation (1)) and in some respects behaves like a sample mean. For instance, Hanson, Pledger and Wright (1973) have shown that the estimator is consistent if the errors satisfy a first moment type of condition. They have also shown that, based on r observations, the probability that the estimator differs from the underlying regression function by more than some fixed amount is $o(r^{-t+1})$ if the errors have finite absolute tth moments and that this probability converges to zero exponentially if the errors have moment generating functions. However, the actual rate of convergence of the estimator to the underlying regression function may be slower than that of a sample mean to a population mean. In particular, Brunk (1970) has shown that, with norming constants of order $r^{1/3}$, the difference between the estimator and the true regression function, at a fixed point, has a nondegenerate limiting distribution provided the true regression function has positive slope at that point. Parsons (1978) has shown that the norming constants are of order $r^{1/2}$ if the regression function is constant. The purpose of this note is to show how the rate of growth of the regression function at a point influences the rate of convergence of the estimator at that point. This will be accomplished by generalizing Brunk's result to points at which the regression function does not have a positive slope.

2. Asymptotic distribution of the estimator. For each $x \in I$, an interval of real numbers, let $D(x)$ be a probability distribution with mean $\theta(x)$. For each positive integer r, let $x_{r1} \le x_{r2} \le \cdots \le x_{rr}$ be points in I and let $Y_{r1}, Y_{r2}, \cdots, Y_{rr}$ be independent random variables with Y_{rk} distributed as $D(x_{rk})$ for $k = 1, 2, \cdots, r$. The x_{rk} are observation points and Y_{rk} is the observation at x_{rk}. (For the result given here the number of distinct observation points must grow at least like a positive constant times r.) Since it may be desirable to weight observations at different points differently, we consider a function w defined on I with $w(x) \ge w_0 > 0$ for all $x \in I$ and some w_0. We assume that θ is nondecreasing on I and consider an estimator proposed by Brunk (1958). Based on the rth set of observations, the estimator is defined by

Received January 1979; revised May 16, 1979.

[1] This research was sponsored by the National Science Foundation under Grant Numbers MCS 76-06949 and MCS 78-01514.

AMS 1970 *subject classifications.* Primary 62E20; secondary 62G05.

Key words and phrases. Isotone regression, limiting, distributions, rates of convergence, Wiener-Levy process.

(1) $$\hat{\theta}_r(x) = \max_{x_{rs} \le x} \min_{x \le x_{rt}} \mathrm{Av}_r([x_{rs}, x_{rt}]),$$

where $\mathrm{Av}_r(B) = \sum_{\{i: x_{ri} \in B\}} w(x_{ri}) Y_{ri} / \sum_{\{i: x_{ri} \in B\}} w(x_{ri})$.

This choice of $\hat{\theta}_r$ is constant on (x_{rj-1}, x_{rj}) for $j = 2, \cdots, r$. However, examining the proofs we see that the results given here are valid for any nondecreasing estimator which agrees with $\hat{\theta}_r$ at the observation points.

Let $\sigma^2(x)$ denote the variance of $D(x)$ and set $\phi(x) = w^2(x)\sigma^2(x)$. For the estimator to be consistent the observation points must satisfy certain regularity conditions. (See Brunk (1970) and Hanson et al. (1973).) In some situations, the observation points are the realization of a sequence (or more generally, a triangular array) of random variables and so we will state such regularity conditions in terms of the empirical distribution functions of the observation points,

$$F_r(x) = \mathrm{card}\{k: x_{rk} \le x\}/r.$$

To specify the rate of growth of the regression function at a point x_0, we assume that for some α and A, both positive,

(2) $$|\theta(x) - \theta(x_0)| = A|x - x_0|^\alpha(1 + o(1)) \qquad \text{as } x \to x_0.$$

(Of course, this and the fact that θ is nondecreasing imply that $\theta(x) < \theta(x_0) < \theta(y)$ for $x < x_0 < y$.) If $\alpha = 1$, then $\theta'(x_0) = A > 0$, which is the case considered by Brunk. If α is an integer greater than one and $\theta^{(\alpha)}(x_0)$ exists then $\theta^{(1)}(x_0) = \cdots = \theta^{(\alpha-1)}(x_0) = 0$ and $\theta^{(\alpha)}(x_0)/\alpha! = A$. (This can be seen by writing a Taylor series expansion for $\theta(x)$ (cf. Hardy (1952, page 278)).)

THEOREM 1. *Suppose that θ is nondecreasing and satisfies* (2); *that w and ϕ are continuous at $x_0 \in I$ and that ϕ is bounded on I; that for $r = 1, 2, \cdots, \{Y_{rl}\}_{l=1}^r$ are independent random variables with Y_{rl} distributed as $D(x_{rl})$ for $l = 1, 2, \cdots, r$ and that $\sum_{l=j(r)}^{k(r)} w(x_{rl})(Y_{rl} - \theta(x_{rl}))/(\sum_{l=j(r)}^{k(r)} \phi(x_{rl}))^{1/2} \xrightarrow{\mathscr{D}} \mathscr{N}(0, 1)$ for all $1 \le j(r) < k(r) \le r$ with $k(r) - j(r) \to \infty$; and that there is a distribution function F, which is continuously differentiable in a neighborhood of x_0 with $F'(x_0) > 0$, for which $\sup_x |F_r(x) - F(x)| = o(r^{-1/(2\alpha+1)})$. Then*

(3) $$\{(\alpha + 1)(F'(x_0)r)^\alpha(\sigma^{2\alpha}(x_0)A)^{-1}\}^{1/(2\alpha+1)}(\hat{\theta}_r(x_0) - \theta(x_0))$$

converges in distribution to the slope at zero of the greatest convex minorant of $W(t) + |t|^{\alpha+1}$, where W is the two-sided Wiener-Levy process with variance one per unit time.

PROOF. While the proof is a modification of the arguments given in Brunk (1970) and Prakasa Rao (1969), we give some of the details to show how the various quantities influence the rate of convergence of $\hat{\theta}_r(x_0)$ to $\theta(x_0)$. Since $F'(x) > 0$ on an open interval containing x_0, we may choose, for an arbitrary c and for r sufficiently large, positive numbers $\alpha_l(r)$ and $\alpha_u(r)$ for which

$$F(x_0) - F(x_0 - \alpha_l(r)) = F(x_0 + \alpha_u(r)) - F(x_0) = 2cr^{-1/(2\alpha+1)}.$$

Set

$$\theta_r^* = \max_{\{s: x_0 - \alpha_l(r) < x_{rs} \le x_0\}} \min_{\{t: x_0 \le x_{rt} < x_0 + \alpha_u(r)\}} \mathrm{Av}_r([x_{rs}, x_{rt}]).$$

LEMMA. *Assuming the hypotheses of the theorem,*

$$\lim_{c\to\infty} \limsup_{r\to\infty} P\{\hat{\theta}_r(x_0) \ne \theta_r^*\} = 0.$$

PROOF. For r sufficiently large, there exist positive numbers $\beta_l(r)$ and $\beta_u(r)$ for which $F(x_0) - F(x_0 - \beta_l(r)) = F(x_0 + \beta_u(r)) - F(x_0) = cr^{-1/(2\alpha+1)}$. We first argue that $P\{\hat{\theta}_r(x_0) \ne \theta_r^*\}$ is bounded by the sum of

(4) $$P\{\min_{y \ge x_0} \mathrm{Av}_r((x_0 - \beta_l(r), y]) < \max_{y \le x_0 - \alpha_l(r)} \mathrm{Av}_r([y, x_0 - \beta_l(r)])\}$$

and

(5) $$P\{\max_{y\le x_0}\mathrm{Av}_r([y, x_0+\beta_u(r))) > \min_{y\ge x_0+\alpha_u(r)}\mathrm{Av}_r([x_0+\beta_u(r), y])\}.$$

To see this, first take complements and note that for $x_{rs} \le x_0$ and $y_0 \ge x_0 + \alpha_u(r)$, $\mathrm{Av}_r([x_{rs}, x_0 + \beta_u(r))) \le \mathrm{Av}_r([x_0+\beta_u(r), y_0])$ implies (by the averaging property of means) that $\mathrm{Av}_r([x_{rs}, x_0+\beta_u(r))) \le \mathrm{Av}_r([x_{rs}, y_0])$. So

$$\hat{\theta}_r(x_0) = \max_{x_{rs}\le x_0}\min_{x_0\le x_{rt}<x_0+\alpha_u(r)}\mathrm{Av}_r([x_{rs}, x_{rt}]).$$

Using the fact that the maximum and minimum may be reversed in computing these estimates (cf. Brunk (1955)) and an argument like the one just completed, one can show that the intersection of the complements of the events in (4) and (5) is contained in $\{\hat{\theta}_r(x_0) = \theta_r^*\}$. So we now need to show that these two probabilities behave as specified. The proofs are similar and so we only give the argument for (4). Expression (4) is bounded above the sum of

(6) $$P\{\min_{y\ge x_0}\mathrm{Av}_r((x_0-\beta_l(r), y]) < \theta(x_0-\beta_l(r))\}$$

and

(7) $$P\{\max_{y\le x_0-\alpha_l(r)}\mathrm{Av}_r([y, x_0-\beta_l(r)]) > \theta(x_0-\beta_l(r))\}.$$

We first consider expression (6). Let $\delta_l(r) = \mathrm{card}\{j : x_0 - \beta_l(r) < x_{rj} \le x_0\}$ and observe that $\delta_l(r) = cr^{2\alpha/(2\alpha+1)}(1+o(1))$. (In the proofs given here, $o(1)$ denotes a sequence indexed by r which may depend on c, w and F but for each c it converges to zero as $r\to\infty$.) Since $\theta(\cdot)$ is nondecreasing, (6) is bounded by $P\{\max_{y\ge x_0}\mathrm{Av}_r^*((x_0-\beta_l(r), y]) > \varepsilon_r\}$ where Av_r^* is defined like Av_r except $w(x_{ri})Y_{ri}$ is replaced by $-w(x_{ri})(Y_{ri} - \theta(x_{ri}))$ and ε_r is

$$\left(\int_{(x_0-\beta_l(r),x_0]} w(x)\,dF_r(x)\right)^{-1}\int_{(x_0-\beta_l(r),x_0]} w(x)\{\theta(x)-\theta(x_0-\beta_l(r))\}\,dF_r(x).$$

Because w is continuous at x_0 and $\beta_l(r)\to 0$ as $r\to\infty$

$$\varepsilon_r = c^{-1}r^{1/(2\alpha+1)}(1+o(1))\int_{(x_0-\beta_l(r),x_0]}\{\theta(x)-\theta(x_0-\beta_l(r))\}\,dF_r(x).$$

Also $\beta_l(r) = c(F'(x_0))^{-1}r^{-1/(2\alpha+1)}(1+o(1))$ and

$$\int_{(x_0-\beta_l(r),x_0]}\{\theta(x)-\theta(x_0-\beta_l(r))\}dF(x)$$

$$= F'(x_0)(1+o(1))A(\beta_l(r))^{\alpha+1}(1+o(1)) - \int_{x_0-\beta_l(r)}^{x_0}(\theta(x_0)-\theta(x))\,dx\}$$

$$= F'(x_0)A(\alpha/(\alpha+1))(\beta_l(r))^{\alpha+1}(1+o(1)).$$

Integrating by parts,

$$\left|\int_{(x_0-\beta_l(r),x_0]}(\theta(x_0)-\theta(x))d(F_r(x)-F(x))\right| \le 2(\theta(x_0)-\theta(x_0-\beta_l(r)))\sup_x|F_r(x)-F(x)|$$

$$= 2Ac^{\alpha}(F'(x_0))^{-\alpha}r^{-(\alpha+1)/(2\alpha+1)}o(1)$$

and so for r sufficiently large, $\varepsilon_r = A_0c^{\alpha}r^{-\alpha/(2\alpha+1)}(1+o(1))$ where A_0 is a positive constant depending on A, α and $F'(x_0)$ but not c. Since $w(x)\ge w_0$, (6) is bounded by

$$P\{\max_{\delta_l(r)\le j\le r}S_{rj}/j > w_0A_0c^{\alpha}r^{-\alpha/(2\alpha+1)}(1+o(1))\}$$

where S_{rj} is, for each r, the sum of j independent random variables with zero means and bounded variances. Applying the Hájek-Rényi inequality (cf. Bauer (1972) page 178), one sees that (6) behaves as specified. A similar argument for (7) completes the proof of the lemma.

We now return to the proof of the theorem. Consider the rth set of observations. Let γ be the total number of these observations in $(x_0 - \alpha_l(r), x_0 + \alpha_u(r)$; let $y_{r1} < y_{r2} < \cdots < y_{r\lambda}$ be the distinct observation points in $(x_0 - \alpha_l(r), x_0 + \alpha_u(r))$ (of course, γ and λ depend on r); let $n(r, k)$ be the number of observations at y_{rk} (i.e., $n(r, k) = \text{card}\{j: x_{rj} = y_{rk}\}$); and let $\bar{X}_{rk}$ be the mean of the observations at y_{rk}. Set $t_{r0} = 0$ and, for $k = 1, 2, \cdots, \lambda$,

$$t_{rk} = 2cD \sum_{l=1}^{k} w(y_{rl})n(r, l)/(\gamma w_r)$$

where $w_r = \sum_{l=1}^{\lambda} w(y_{r,l})n(r, l)/\gamma = w(x_0)(1 + o(1))$, $D = 2(\sigma(x_0)B)^{-2}$ and $B = \{(\alpha + 1)(F'(x_0))^{\alpha}/(A\sigma^{2\alpha}(x_0))\}^{1/(2\alpha+1)}$. Define a process on $[0, 2cD]$ by $U_r(0) \equiv 0$,

$$U_r(t_{rk}) = 2cD \sum_{l=1}^{k} w(y_{rl})n(r, l)\bar{X}_{rl}/(\gamma w_r)$$

and $U_r(\cdot)$ is linear between the points t_{rk}. For some of the arguments that follow it is useful to note that

$$U_r(t) = 2cD \sum_{l=1}^{\lambda} h_{rl}(t)w(y_{rl})n(r, l)\bar{X}_{rl}/(\gamma w_r)$$

where $h_{rl}(t) = 0$ for $t < t_{rl-1}$, $h_{rl}(t) = (t - t_{rl-1})(t_{rl} - t_{rl-1})^{-1}$ for $t_{rl-1} \le t \le t_{rl}$ and $h_{rl}(t) = 1$ for $t > t_{rl}$. Using the algorithm based on the cumulative sum diagram discussed in Brunk (1956), θ_r^* is bounded above (below) by the slope from the left of the greatest convex minorant of the graph of $U_r(t)$ evaluated at the point $t = t_{rj(r)}(t = t_{rj(r)-1})$, where $y_{rj(r)-1} < x_0 \le y_{rj(r)}$. We denote the slope from the left at x of the greatest convex minorant of the graph of $X(s)$ for $s \in S$ by $\text{slogcom}(x)\{(s, X(s)) : s \in S\}$ and note that

$$\begin{aligned}(8)\qquad \text{slogcom}(t_{rj(r)-1})\{(t, U_r(t) - \theta(x_0)t) : t \in [0, 2cD]\} &\le \theta_r^* - \theta(x_0) \\ &\le \text{slogcom}(t_{rj(r)})\{(t, U_r(t) - \theta(x_0)t) : t \in [0, 2cD]\}.\end{aligned}$$

We obtain the limiting distribution of $r^{\alpha/(2\alpha+1)}B(\theta_r^* - \theta(x_0))$ by showing that $r^{\alpha/(2\alpha+1)}B$ times the lower bound in (8) and $r^{\alpha/(2\alpha+1)}B$ times the upper bound in (8) have the same limiting distribution. The two arguments are similar and so we only give the latter. With $M_r(t) = E(U_r(t))$ for $t \in [0, 2cD]$, we examine

$$W_r(t) = r^{\alpha/(2\alpha+1)}B(U_r(t) - M_r(t)).$$

This process is centered at its mean and since $\gamma = 4cr^{2\alpha/(2\alpha+1)}(1 + o(1))$, its covariance function is for $s \le t$ with $t_{rl(r)-1} < s \le t_{rl(r)}$,

$$t_{rl(r)-1}(1 + o(1)) + O(\max_k(t_{rk} - t_{rk-1})) = s + O(\max_k(t_{rk} - t_{rk-1})).$$

But because $F(\cdot)$ is continuous in a neighborhood of x_0 and because of the hypothesized rate of convergence of $F_r(\cdot)$ to $F(\cdot)$, $\max_k(t_{rk} - t_{rk-1}) = o(1)$. Clearly the finite-dimensional distributions converge to those of the Wiener process on $[0, 2cD]$. The sequence $\{W_r\}$ is shown to be tight by considering the modulus of continuity. The argument is an easy adaptation of those found on pages 59 and 70 of Billingsley (1968). (Note that $t_{rk(r)} - t_{rl(r)} \le \delta$ implies that, for r sufficiently large, $k(r) - l(r) - 1 \le (cD)^{-1}\delta\gamma$.)

Next we show that $f_r(t) = r^{\alpha/(2\alpha+1)}B(M_r(t) - \theta(x_0)t)$ converges uniformly to $|cD - t|^{\alpha+1} - (cD)^{\alpha+1}$. Since

$$t \equiv 2cD \sum_{l=1}^{\lambda} h_{rl}(t)w(y_{rl})n(r, l)/(\gamma w_r),$$

$f_r(t_{rk})$ can be rewritten as

$$(9)\qquad 2BcDr^{\alpha/(2\alpha+1)} \sum_{l=1}^{k} w(y_{rl})n(r, l)(\theta(y_{rl}) - \theta(x_0))/(\gamma w_r).$$

Furthermore, $f_r(t)$ is nonincreasing for $t \le t_{rj(r)-1}$ and nondecreasing for $t \ge t_{rj(r)-1}$. Since $r^{1/(2\alpha+1)}\alpha_l(r) \to 2c/F'(x_0)$, $t_{rj(r)-1}$ (and consequently $t_{rj(r)}$) converges to cD as $r \to \infty$. Fix $t \in (0, t_{rj(r)-1}]$ and set $k(r) = \max\{k : t_{rk} < t\}$. Then $f_r(t_{rk(r)}) \ge f_r(t) \ge f_r(t_{rk(r)+1})$ and both $t_{rk(r)}$

and $t_{rk(r)+1}$ converge to t as $r \to \infty$. So $\sum_{l=1}^{k(r)} n(r, l)/r^{2\alpha/(2\alpha+1)} \to 2t/D$ which implies that $F'(x_0)r^{1/(2\alpha+1)}(y_{rk(r)} - x_0 + \alpha_l(r)) \to 2t/D$. Also $f_r(t_{rk(r)})$ is

$$2^{-1}BDr^{(\alpha+1)/(2\alpha+1)}(1 + o(1)) \int_{(x_0-\alpha_l(r), y_{rk(r)}]} (\theta(x) - \theta(x_0))\, dF_r(x)$$

and

$$r^{(\alpha+1)/(2\alpha+1)} \int_{x_0-\alpha_l(r)}^{y_{rk(r)}} (\theta(x) - \theta(x_0))F'(x)\, dx$$

$$= AF'(x_0)(\alpha + 1)^{-1}(1 + o(1))r^{(\alpha+1)/(2\alpha+1)}\{(\alpha_l(r) - (y_{rk(r)} - x_0 + \alpha_l(r)))^{\alpha+1} - (\alpha_l(r))^{\alpha+1}\}$$
$$= (F'(x_0))^{-\alpha}(2/D)^{\alpha+1}A(\alpha + 1)^{-1}(1 + o(1))((cD - t)^{\alpha+1} - (cD)^{\alpha+1}).$$

(The sequence $o(1)$ in the last two expressions and the remainder of the proof may depend on t but for fixed c they converge uniformly for $t \in [0, 2cD]$.) Using (2) and the fact that $\max_l n(r, l)/r^{2\alpha/(2\alpha+1)} \to 0$, one can show that $\max_k |f_r(t_{rk}) - f_r(t_{rk+1})| \to 0$ and so both $f_r(t_{rk(r)})$ and $f_r(t_{rk(r)+1})$ are of the form $((cD - t)^{\alpha+1} - (cD)^{\alpha+1})(1 + o(1))$. The argument given above shows that $f_r(t_{rj(r)-1})$ and $f_r(t_{rj(r)})$ converge to $(-cD)^{\alpha+1}$. Considering separately the summands in expression (9) which have index $l < j(r)$ and $l \geq j(r)$, one can also show that $f_r(t_{rk(r)}) = (t - cD)^{\alpha+1}(1 + o(1)) - (cD)^{\alpha+1}(1 + o(1))$ for $t_{rk(r)} < t \leq t_{rk(r)+1}$ and $k(r) \geq j(r)$. The desired uniform convergence follows since the limit function is uniformly continuous and the $o(1)$ functions above converge uniformly.

Since constant functions do not influence the slope, making the change of variable $s = t - cD$, the expression $r^{\alpha/(2\alpha+1)}B$ slogcom$(t_{rj(r)})\{(t, U_r(t) - \theta(x_0)t) : t \in [0, 2cD]\}$ becomes

$$\text{slogcom}(t_{rj(r)} - cD)\{(s, X_r(s)) : s \in [-cD, cD]\} \tag{10}$$

where $X_r(s)$ converges weakly to $W_c(s) + |s|^{\alpha+1}$ and $W_c(s)$ is the two-sided Wiener-Levy process on $[-cD, cD]$ with variance 1 per unit time. Next we show that (10) converges weakly to slogcom$(0)\{(s, W_c(s) + |s|^{\alpha+1}) : s \in [-cD, cD]\}$. The proof is like that given at the beginning of Section 6 of Prakasa Rao (1969), but since the point at which the slope is being evaluated depends on r we apply Theorem 5.5 of Billingsley (1968). The set E, in the hypotheses of that theorem, is contained in the set of sample paths for which the convex minorant of $W_c(s) + |s|^{\alpha+1}$ does not have a unique slope at $s = 0$.

The proof is completed by showing that the probability that the slope of the convex minorant of $W_c(s) + |s|^{\alpha+1}$ on $[-cD, cD]$ is different from the slope of the convex minorant of $W(s) + |s|^{\alpha+1}$ on $(-\infty, \infty)$ converges to zero as $c \to \infty$. This is the analogue of Lemma 6.2 of Prakasa Rao (1969) and the proof is like his, except we must show that $P\{W(s) \geq |s|^{\alpha+1}$ for some $s > cD\} \to 0$ as $c \to \infty$. One could follow his proof or note that this follows from $W(s)/s \to 0$ a.s. as $s \to \infty$ (cf. Breiman (1968, page 265)).

3. Comments. There are a couple of comments that need to be made concerning the result of Theorem 1. In the case of random observation points, F_r will not converge to the appropriate F at the specified rate if $0 < \alpha \leq ½$. However, the observation points can be chosen deterministically so that $\sup_x |F_r(x) - F(x)| = O(r^{-1})$ which is $o(r^{-1/(2\alpha+1)})$. The assumption concerning the asymptotic normality of the sums of the variables $V_{rl} = W(x_{rl})(Y_{rl} - \theta(x_{rl}))$ requires, in general, some sort of uniformity condition. For instance, one could assume that ϕ is bounded away from zero and that the $E|V_{rl}|^{2+\delta}$ are uniformly bounded for some $\delta > 0$. Finally, this result has been obtained independently by Leurgans (1978) for uniformly spaced observations, a constant weight function w, and α an integer.

REFERENCES

[1] BAUER, HEINZ (1972). *Probability Theory and Elements of Measure Theory.* Holt, Rinehart and Winston, New York.

[2] BILLINGSLEY, PATRICK (1968). *Convergence of Probability Measures.* Wiley, New York.
[3] BREIMAN, LEO (1968). *Probability.* Addison-Wesley, Reading, Massachusetts.
[4] BRUNK, H. D. (1955). Maximum likelihood estimates of monotone parameters. *Ann. Math. Statist.* **26** 607–616.
[5] BRUNK, H. D. (1956). On an inequality for convex functions. *Proc. Amer. Math. Soc.* **7** 817–824.
[6] BRUNK, H. D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.* **29** 437–454.
[7] BRUNK, H. D. (1970). Estimation of isotonic regression. *Nonparametric Techniques in Statistical Inference.* 177–195. Cambridge Univ. Press.
[8] HANSON, D. L., PLEDGER, GORDON, and WRIGHT, F. T. (1973). On consistency in monotonic regression. *Ann. Statist.* **1** 401–421.
[9] HARDY, G. H. (1952). *A Course of Pure Mathematics*, Tenth Ed. Cambridge Univ.
[10] LEURGANS, S. E. (1978). Asymptotic distribution theory in generalized isotonic regression. Ph.D. dissertation, Stanford Univ.
[11] PARSONS, VAN L. (1978). The limiting distribution of the isotonic estimator at a point. Unpublished manuscript.
[12] PRAKASA RAO, B. L. S. (1969). Estimation of a unimodal density. *Sankhyā* (A) **31** 23–36.

DEPARTMENT OF MATHEMATICS
326 MATHEMATICS-COMPUTER SCIENCE BLDG.
UNIVERSITY OF MISSOURI-ROLLA
ROLLA, MISSOURI 65401