AD-A100 916 MISSOURI UNIV-ROLLA DEPT OF MATHEMATICS F/G 12/1
THE ASYMPTOTIC BEHAVIOR OF MONOTONE PERCENTILE REGRESSION ESTIM--ETC(U)
MAY 81 F T WRIGHT N00014-80-C-0322
UNCLASSIFIED 4 NL

1 OF 1
AD A
100 916

END
DATE
FILMED
8-81
DTIC

LEVEL II

12

AD A100916

DTIC FILE COPY

# THE ASYMPTOTIC BEHAVIOR OF MONOTONE PERCENTILE REGRESSION ESTIMATES[1]

F. T. Wright

University of Missouri-Rolla

Abbreviated Title: Monotone Percentile Regression

American Mathematical Society 1970 subject classification:

Primary 62E20 Secondary 60F05

Key words and phrases: Monotone regression, percentile regression, asymptotic distributions, rate of convergence.

DTIC ELECTE JUL 0 6 1981 S D E

[1]This research was sponsored by the National Science Foundation under Grant Number MCS 78-01514 and by the Office of Naval Research under Contract ONR N00014-80-C-0322.

81 6 29 245

1. INTRODUCTION. The estimation of a regression function which is defined on an interval of the real line and is nondecreasing (nonincreasing) but is not assumed to be of a particular functional form has been considered in the literature. Typically, an estimate is chosen which minimizes a particular objective function subject to the appropriate montonicity constraints. Brunk (1958, 1970) considered the weighted least squares ($\ell_2$) estimate and, in the later reference, demonstrated its consistency and obtained its asymptotic distribution. The large sample distribution results have been extended by Leurgans (1979) and Wright (1981). Rates of convergence for this estimator have been studied by Makowski (1973) and Hanson, Pledger and Wright (1973).

The least absolute deviations ($\ell_1$) estimate was introduced by Robertson and Waltman (1968). Cryer, et al. (1972) also considered this estimator, showing its consistency, studying its rate of convergence and comparing it to the $\ell_2$ estimator by Monte Carlo techniques. Casady and Cryer (1976) have shown that based on $r$ observations, the almost sure rate of convergence for the $\ell_1$ estimator is of order no larger than $r^{-1/4}$, assuming the underlying regression function satisfies a first order Lipschitz condition. (This is the same rate obtained by Makowski (1973) for the $\ell_2$ estimator.) In this paper, the asymptotic distribution of the $\ell_1$ estimator is obtained and assuming the regression function has a positive slope at a point, the rate of convergence at that point is seen to be of order $r^{-1/3}$. The techniques presented here also apply to the weighted

$\ell_1$ estimate provided the weighting function is positive, continuous, bounded and bounded away from zero. As was seen in the least squares case (cf. Wright (1981)), such weights do not affect the limiting distribution and so we only consider the case of equal weights. (For a description of the weighted estimate see Robertson and Wright (1975).)

Leurgans (1979) has obtained the asymptotic distribution of another estimator which might be appropriate if the errors have heavier tails than the normal distribution. The estimate she has considered is defined to be the slope of the greatest convex minorant of a process determined by smoothly weighted linear combinations of order statistics. For Leurgans' estimator, the $\ell_1$ estimator and the $\ell_2$ estimator the order of the rate of convergence is the same and so the large sample relative efficiencies are determined by the multiplicative constants. These comparisons are not discussed here since they are the same as those for the ordinary one sample location problem.

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

2. ASYMPTOTIC DISTRIBUTION OF THE $\ell_1$ ESTIMATOR. Let $p \in (0,1)$ and for each $x \in I$, an interval of real numbers, let $D(x)$ be a probability distribution with pth quantile $\theta(x)$. For each positive integer r, let $x_{r1} \leq x_{r2} \leq \cdots \leq x_{rr}$ be points in I and let $Y_{r1}, Y_{r2}, \ldots, Y_{rr}$ be independent random variables with $Y_{rk}$ distributed as $D(x_{rk})$. The $x_{rk}$ are observation points and the $Y_{rk}$ are observations. (For the results given here the number of distinct observations must grow at least like some positive constant times r. If the number of distinct observation points is bounded, the work of Robertson and Waltman (1968) gives the asymptotic distribution.) The estimator proposed by Robertson and Waltman is defined at the observation points and any monotone extension to I might be appropriate. One such estimator is given by

$$(1) \qquad \hat{\theta}_r(x) = \max_{x_{rs} \leq x} \min_{x \leq x_{rt}} Q_r([x_{rs}, x_{rt}])$$

where $Q_r(A)$ is the pth sample quantile, that is the [rp]th order statistic, of the sample comprised of those $Y_{rk}$ for which $x_{rk} \in A$. This choice of $\hat{\theta}_r$ is constant on $(x_{rj-1}, x_{rj})$ for $j = 2, \ldots, r$, however, examining the proofs we see that the large sample results given here are valid for any nondecreasing estimator which coincides with $\hat{\theta}_r$ at the observation points. If $p = 1/2$ the monotone $\ell_1$ estimator is obtained.

The result that follows gives the asymptotic distribution of $\hat{\theta}_r(x_0)$ with $x_0$ in the interior of I. As in the $\ell_2$ case, the rate of convergence of the estimator, $\hat{\theta}_r(x_0)$, depends on the rate of growth of the regression function at $x_0$. We assume that for some $\alpha$ and $\beta$, both positive,

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

(Some implications of this assumption are discussed in Wright (1981).) Since the observation points may be the realization of a sequence of random variables each of which has support in I, we state the conditions on them in terms of their empirical distribution function,

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

We assume that there is a distribution function F, which is continuously differentiable in a neighborhood of $x_0$ with $F'(x_0) > 0$, for which

(3) $$\sup_x |F_r(x) - F(x)| = o(r^{-1/(2\alpha+1)}).$$

To apply the usual techniques for sample quantiles we make the following uniformity assumption:

(4) $$\sup\{|P\{Y_{rk}-\theta(x_{rk})\leq x\}-p-\eta x|: 1\leq k\leq r,\ r=1,2,\ldots,|x|\leq\rho\}=o(\rho) \text{ as } \rho\to 0.$$

Of course, (4) is satisfied if the $Y_{rk} - \theta(x_{rk})$ have a common distribution function G, with pth quantile 0 and $G'(0) = \eta$.

<u>THEOREM</u>. Suppose that $\theta$ is nondecreasing and satisfies (2); that the observation points satisfy (3); and that the observations, $\{Y_{rk}\}$, are independent for each r and satisfy (4). Then

(5) $$\{(\alpha+1)(rF'(x_0)\eta^2(p(1-p))^{-1})^{\alpha}/\beta\}^{1/(2\alpha+1)}(\hat{\theta}_r(x_0)-\theta(x_0))$$

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

<u>Proof</u>. The proof is similar to those given by Prakasa Rao (1969) and Brunk (1970) except that we must approximate the sample quantiles by averages before their techniques can be applied.

As in the $\ell_2$ case, we show that $\hat{\theta}_r(x_0)$ is asymptotically equivalent to an estimator based on observations with observation points in a sequence of intervals which converges to $x_0$. We make use of the modification of the arguments due to Prakasa Rao (1969) and Brunk (1970) given in Wright (1981). For an arbitrary c and r sufficiently large choose $\alpha_\ell(r)$, $\alpha_u(r)$, $\beta_\ell(r)$ and $\beta_u(r)$ so that

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

and

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

Set

$$\theta_r^* = \max_{x_{rs}\varepsilon(x_0-\alpha_\ell(r),x_0]}\min_{x_{rt}\varepsilon[x_0,x_0+\alpha_u(r))} Q_r([x_{rs},x_{rt}]).$$

The first step in the proof is to show that $\hat{\theta}_r(x_0)$ and $\theta_r^*$ are asymptotically equivalent by showing

$$\lim_{c\to\infty} \limsup_{r\to\infty} P\{\hat{\theta}_r(x_0)\neq\theta_r^*\} = 0. \tag{6}$$

Since the pth sample quantile is a Cauchy mean value function (cf. Robertson and Wright (1975)), it satisfies the averaging property used in the proof of the lemma in Wright (1981). That proof can be modified to obtain (6), but the following should be noted: with $\nu_r(A) = \text{card}\{k: x_{rk} \varepsilon A\}$, $I_A$ the indicator of the event A and $\sum_A^{(r)}$ the sum over those k for which $x_{rk} \varepsilon A$,

$$\{\min_{y\geq x_0} Q_r((x_0-\beta_\ell(r),y]) \leq \theta(x_0-\beta_\ell(r))\} \subset$$

$$\bigcup_{y\geq x_0} \{Av_r'((x_0-\beta_\ell(r),y])\geq p-(\nu_r((x_0-\beta_\ell(r),x_0])^{-1}$$

$$- \sum_{(x_0-\beta_\ell(r),y]}^{(r)} \mu_{rk}/\nu_r((x_0-\beta_\ell(r),y])\},$$

where $Av_r'(A) = \sum_A^{(r)}(I_{\{Y_{rk}\le\theta(x_0-\beta_\ell(r))\}}-\mu_{rk})/\nu_r(A)$ and

$\mu_{rk} = P\{Y_{rk}\le\theta(x_0-\beta_\ell(r))\}$. Furthermore, if $x_{rk} > x_0$,

$\mu_{rk} \le P\{Y_{rk}-\theta(x_{rk})\le\theta(x_0-\beta_\ell(r))-\theta(x_0)\}$, which using (4),

can be bounded by $p + \eta(\theta(x_0-\beta_\ell(r))-\theta(x_0))(1+o(1))$ where

$o(1)$ represents quantity that does not depend on k and

converges to zero as $r \to \infty$. If $x_{rk} \le x_0$, then

$$\mu_{rk}\le p+\eta(\theta(x_0-\beta_\ell/r))-\theta(x_{rk}))+(\theta(x_0-\beta_\ell(r))-\theta(x_0))o(1)$$

and so $\sum_{(x_0-\beta_\ell(r),y]}^{(r)}\mu_{rk}/\nu_r((x_0-\beta_\ell(r),y))\le p+(\theta(x_0-\beta_\ell(r))-\theta(x_0))o(1)+$

$$\eta\sum_{(x_0-\beta_\ell(r),x_0]}^{(r)}(\theta(x_0-\beta_\ell(r))-\theta(x_{rk}))/\nu_r((x_0-\beta_\ell(r),x_0]).$$

Next we approximate the pth quantiles by averages. Then, since the max-min operation on averages is the same as the slogcom (slope of the greatest convex minorant) of a cumulative sum process, we only need to make slight modifications on the arguments in the literature for such slogcom's. For arbitrary $\varepsilon$ with $0 < \varepsilon < c$, let $\gamma_\ell(r)$ and $\gamma_u(r)$ be defined by $\gamma_u(r) = \min\{x_{rk}: F(x_{rk})-F(x_0) \ge \varepsilon r^{-1/(2\alpha+1)}\}-x_0$ and $\gamma_\ell(r) = x_0 - \max\{x_{rk}: F(x_0)-F(x_{rk}) \ge \varepsilon r^{-1/(2\alpha+1)}\}$. Note that

$$(7)\quad \max_{x_{rs}\varepsilon(x_0-\alpha_\ell(r),x_0-\gamma_\ell(r)]}\min_{x_{rt}\varepsilon[x_0,x_0+\alpha_u(r))}Q_r([x_{rs},x_{rt}])\le\theta_r^*$$

$$\le\max_{x_{rs}\varepsilon(x_0-\alpha_\ell(r),x_0]}\min_{x_{rt}\varepsilon[x_0+\gamma_u(r),x_0+\alpha_u(r))}Q_r([x_{rs},x_{rt}]).$$

The large sample distribution for the lower and upper bounds in (7) depend on $\varepsilon$ and as $\varepsilon \to 0$ they approach the same limit.

The discussion for the upper and lower bounds are similar and so we only give the former. Denote the upper bound in (7) by $\theta_r'$. To approximate the quantiles by averages, we show that

$$r^{\alpha/(2\alpha+1)}\max_{x_{rs}\varepsilon(x_0-\alpha_\ell(r),x_0],x_{rt}\varepsilon[x_0+\gamma_u(r),x_0+\alpha_u(r))}|$$

$$\eta(Q([x_{rs},x_{rt}])-\theta(x_0)) - Av_r''([x_{rs},x_{rt}])| \overset{p}{\to} 0,$$

where $Av_r''(A) = \sum_A^r(p-I_{\{Y_{rk}\leq\theta(x_0)\}})/\nu_r(A)$. We appeal to the following lemma which is a generalization of a result given in Ghosh (1971):

<u>Lemma</u>. Consider arrays of random variables $\{A_{jk}^n: j=1,2,\ldots, J_n, k=1,2,\ldots,K_n, n=1,2,\ldots\}$ and $\{B_{jk}^n: j=1,2,\ldots,J_n, k=1,2,\ldots, K_n, n=1,2,\ldots\}$. If $\max_{jk}\{A_{jk}^n|$ is tight and for each $\delta > 0$ and real x

(8) $P\{A_{jk}^n\leq x \text{ and } B_{jk}^n\geq x+\delta \text{ for some } j \text{ and } k\} \to 0$

and

(9) $P\{A_{jk}^n\geq x+\delta \text{ and } B_{jk}^n\leq x \text{ for some } j \text{ and } k\} \to 0$

as $n \to \infty$, then $\max_{jk}|A_{jk}^n-B_{jk}^n| \overset{p}{\to} 0$.

<u>Proof</u>. Because of the tightness assumption it suffices to consider for arbitrary, positive $\tau$ and M

$$P\{|A_{jk}^n-B_{jk}| > 2\tau \text{ and } -M < A_{jk}^n < M \text{ for some } j \text{ and } k\}$$

$$\leq\sum_{\ell=1}^L P\{-M+(\ell-1)\tau<A_{jk}^n\leq-M+\ell\tau, B_{jk}^n\geq-M+(\ell+1)\tau \text{ for some } j \text{ and } k\}$$

$$+\sum_{\ell=1}^L P\{-M+(\ell-1)\tau<A_{jk}^n\leq-M+\ell\tau, B_{jk}^n\leq-M+(\ell-2)\tau \text{ for some } j \text{ and } k\}$$

where $L = [2M/\tau] + 1$. The above sums can be made as small as desired for fixed M and $\tau$.

Set $n_{1r} = \nu_r((-\infty, x_0-\alpha_\ell(r)]), n_{2r}=\nu_r((-\infty,x_0]), n_{3r}=$ $\nu_r((-\infty,x_0+\gamma_u(r)))$ and $n_{4r} = \nu_r((-\infty,x_0+\alpha_u(r)))$.

We apply the lemma with $n = r$,
$K_r = n_{4r} - n_{3r}$, $J_r = n_{2r} - n_{1r}$, $A^r_{jk} = r^{\alpha/(2\alpha+1)} Av''_r([x_{r,n_{2r}-j+1}, x_{r,n_{3r}+k}])$ and $B^r_{jk} = \eta r^{\alpha/(2\alpha+1)}(Q_r([x_{r,n_{2r}-j+1}, x_{r,n_{3r}+k}])-\theta(x_0))$. The proofs that (8) and (9) hold in this case are similar and we only give the latter. Fix x and note that $\{B^r_{jk} \le x\}$ is contained in

$$\{\textstyle\sum^r_{[x_{r,n_{2r}-j+1}, x_{r,n_{3r}+k}]} I\{Y_{r\ell} \le \theta(x_0)+x/(\eta r^{\alpha/(2\alpha+1)})\}$$

$$\ge [(j+k+n_{3r}-n_{2r})p]\}$$

and $\{B^r_{jk} \le x$ and $A^r_{jk} \ge x+\delta\}$ is contained in

$$(10) \quad \{\textstyle\sum^r_{[x_r,n_{2r}-j+1,x_r,n_{3r}+k]} U_{r\ell} \ge (x+\delta)(j+k+n_{3r}-n_{2r})r^{-\alpha/(2\alpha+1)}-1\}$$

where $U_{r\ell} = I_{\{Y_{r\ell}\le\theta(x_0)+xr^{-\alpha/(2\alpha+1)}/\eta\}} - I_{\{Y_{r\ell}\le\theta(x_0)\}}$. For the values of the index $\ell$ in the last sum $\theta(x_0-\alpha_\ell(r)) \le \theta(x_{r\ell}) \le \theta(x_0+\alpha_u(r))$ and applying (2), both $\theta(x_0)-\theta(x_0-\alpha_\ell(r))$ and $\theta(x_0+\alpha_u(r))-\theta(x_0)$ can be written as $\beta(2c/F'(x_0))^\alpha r^{-\alpha/(2\alpha+1)}(1+o(1))$. Hence, applying (4), $EU_{r\ell} = xr^{-\alpha/(2\alpha+1)} + o(r^{-\alpha/(2\alpha+1)})$ and since $n_{3r}-n_{2r} \ge \varepsilon r^{2\alpha/(2\alpha+1)}(1+o(1))$ for r sufficiently large (10) is contained in

$$\{\textstyle\sum^r_{[x_r,n_{2r}-j+1,x_r,n_{3r}+k]}(U_{r\ell}-EU_{r\ell}) \ge \varepsilon\delta r^{\alpha/(2\alpha+1)}/2\}.$$

With r fixed, define $X_{1,1} = \sum^r_{[x_{r,n_{2r}}, x_{r,n_{3r}}]}(U_{r\ell}-EU_{r\ell})$,

$X_{1,j+1} = U_{rn_{2r}-j+1} - EU_{rn_{2r}-j+1}$ for $j=1,2,\ldots,J_r$, $X_{k+1,1} = U_{rn_{3r}+k} - EU_{rn_{3r}+k}$ for $k = 1,2,\ldots,K_r$ and $X_{ij} \equiv 0$ otherwise. So for sufficiently larger r, $\{B^r_{jk} \le x$ and $A^r_{jk} \ge x+\delta$ for some j and k$\}$ is contained in $\{\max_{1\le j\le J_r+1, 1\le k\le K_r+1} \sum_{s=1}^{j}\sum_{t=1}^{k} X_{s,t} \ge \varepsilon\delta r^{\alpha/(2\alpha+1)}/2\}$ and applying Theorem 6 of Gabriel (1977) the probability of the latter event is bounded by

$$O(r^{-2\alpha/(2\alpha+1)}(J_r+K_r+n_{3r}-n_{2r})\max_{n_{1r}+1\le \ell\le n_{4r}} V(U_{r\ell})).$$

However, $J_r+K_r+n_{3r}-n_{2r} = 4cr^{2\alpha/(2\alpha+1)}(1+o(1))$, $V(U_{r\ell}) \le |EU_{r\ell}|$, and for $n_{1r} + 1 \le \ell \le n_{4r}$, $EU_{r\ell}$ was shown to be of the form $xr^{-\alpha/(2\alpha+1)} + o(r^{-\alpha/(2\alpha+1)})$. Hence, (8) holds. (The proof of (9) is similar.) Now we must show that $\max_{j\ k}|A^r_{jk}|$ is tight. Since $|EI_{\{Y_{r\ell}\le\theta(x_0)\}}|$ is uniformly bounded for $n_{1r} + 1 \le \ell \le n_{4r}$ and $r = 1,2,\ldots$, it suffices to show that

$$\max_{j,k}\left|\sum_{[x_{rn_{2r}-j+1}, x_{rn_{3r}+k}]}(I_{\{Y_{r\ell}\le\theta(x_0)\}}-EI_{\{Y_{r\ell}\le\theta(x_0)\}})/(j+k+n_{3r}-n_{2r})\right|$$

is tight. But this follows from an argument like the one above which uses the result due to Gabriel (1977).

Now we consider the large sample distribution of $\theta''_r = \eta^{-1}\max_{x_{rs}\varepsilon(x_0-\alpha_\ell(r),x_0]}\min_{x_{rt}\varepsilon[x_0+\gamma_u(r),x_0+\alpha_u(r))} Av''_r([x_{rs},x_{rt}])$ Since the proof is similar to that given in Wright (1981) we use the same notation. Let $y_{r1} < y_{r2} < \ldots < y_{r\lambda}$ be the distinct observation points in $(x_0-\alpha_\ell(r),\ x_0+\alpha_u(r))$; $n(r,k)$ the number of observations at $y_{rk}$; $\gamma=\sum_{\ell=1}^{\lambda} n(r,\ell)$; $\overline{X}_{rk} = Av''_r(\{y_{rk}\})$;

$t_{r0} = 0$, $t_{rk} = 2cD\sum_{\ell=1}^{k} n(r,\ell)/\gamma$ for $k=1,2,\ldots,\lambda$; $D=2(\sigma(x_0)B)^{-2}$; $\sigma^2(x_0) = p(1-p)/\eta^2$; and $B = \{(\alpha+1)(F'(x_0))^{\alpha}/\beta\sigma^{2\alpha}(x_0)\}^{1/(2\alpha+1)}$. Define a process on $[0,2cD]$ by $U_r(0) \equiv 0$,

$$U_r(t_{rk}) = 2cD\sum_{\ell=1}^{k} n(r,\ell)\overline{X}_{r\ell}/\gamma \text{ for } k = 1,2,\ldots,\lambda,$$

and linear interpolation between the points $t_{rk}$. Let $j(r)$ and $\ell(r)$ satisfy $y_{rj(r)} \le x_0 < y_{rj(r)+1}$ and $y_{r\ell(r)} = x_0+\gamma_u(r)$. In Chapter 1 of Barlow, et al. (1972) the relationship of the max-min operator to the left hand slope of the greatest convex minorant of the cumulative sum process is discussed. Using this relationship and denoting the slope from the left at x of the greatest convex minorant of the graph of $X(s)$ for $s \in S$ by slogcom $(x)\{(s,X(s)): s \in S\}$, we see that

$$\text{(11)} \quad \eta^{-1}\text{slogcom}(t_{rj(r)})\{(t,U_r(t)): 0\le t\le 2cD\}$$
$$\le \theta''_r \le \eta^{-1}\text{slogcom}(t_{r\ell(r)})\{(t,U_r(t)): 0\le t\le 2cD\}.$$

The argument starting on p. 446 of Wright (1981) shows that

$$r^{\alpha/(2\alpha+1)}B\eta^{-1}\text{slogcom}(t_{rj(r)})\{(t,U_r(t)): 0 \le t \le 2cD\}$$

converges weakly to

$$V = \text{slogcom}(0)\{(s,W(s) + |s|^{\alpha+1}): -cD \le s \le cD\}.$$

However, it should be noted that since $\sup_{1\le\ell\le\lambda}|EI_{\{Y_{r\ell}\le\theta(x_0)\}} - p - \eta(\theta(x_0) - \theta(x_{r,\ell}))| = o(r^{-\alpha/(2\alpha+1)})$, $r^{\alpha/(2\alpha+1)}B\eta^{-1}EU_r(t)$ can be written as the sum of $f_r(t)$ and a function which converges to zero uniformly in $t \in [0,2cD]$ as $r \to \infty$.

In considering the upper bound in (11), we first note that $\max_{1<\ell\le\lambda}(y_{r\ell}-y_{r\ell-1}) = o(r^{-1/(2\alpha+1)})$ which implies that

$\nu_r((x_0, x_0+\gamma_u(r)]) = \varepsilon r^{2\alpha/(2\alpha+1)}(1+o(1))$. Since in the arguments above we have seen that $t_{rj(r)} \to cD$, $t_{r\ell(r)} \to cD + \varepsilon D/2$. We appeal to the same weak convergence result as before to show that

$$r^{\alpha/(2\alpha+1)}B\eta^{-1}\text{slogcom}(t_{r\ell(r)})\{(t, U_r(t)): 0 \le t \le 2cD\}$$

converges weakly to

$$V(\varepsilon) = \text{slogcom}(\varepsilon D/2)\{(s, W(s)+|s|^{\alpha+1}): -cD \le s \le cD\},$$

but we must show that with probability one the convex minorant of $W(s) + |s|^{\alpha+1}$ on $[cD, cD]$ has unique slope at $\varepsilon D/2$. However, if the convex minorant has different left and right slopes at a fixed point t, then it agrees with $W(s) + |s|^{\alpha+1}$ at the point t. This implies that for $h > 0$, $(W(t+h) + |t+h|^{\alpha+1} - W(t) - |t|^{\alpha})/h$ is bounded below by the right hand slope of the convex minorant at t, but with probability one, $\liminf_{h\to 0}(W(t+h) - W(t))/h = -\infty$. So with probability one the convex minorant has unique slope at t. Again, using the same techniques $V(\varepsilon)$ converges weakly to V as $\varepsilon \to 0$.

We now show that $V_r = r^{\alpha/(2\alpha+1)}B(\theta_r^* - \theta(x_0))$ converges weakly to V. Let $r_k$ be a subsequence and V' a generalized random variable, that is one that may be infinite with positive probability, with $P\{V_{r_k} \le x\} \to \{V' \le x\}$ for each real x at which $P\{V' \le x\}$ is continuous. Since $A_r(\varepsilon) = r^{\alpha/(2\alpha+1)}B\theta_r''$ is bounded above and below by sequences which converge weakly, $A_r(\varepsilon)$ is tight for each $\varepsilon > 0$. So for each $\varepsilon > 0$, there is a further subsequence $r_{k(j)}$ with $A_{r_{k(j)}}(\varepsilon)$ converging weakly, say, to $A(\varepsilon)$. Let x be a real number in the continuity sets of $V, V', V(\varepsilon)$ and $A(\varepsilon)$ for a sequence of $\varepsilon$ converging to zero. Since $A_r(\varepsilon)$

and $B_r(\varepsilon) = r^{\alpha/(2\alpha+1)}B(\theta_r'-\theta(x_0))$ are asymptotically equivalent for each $\varepsilon > 0$, for real x and $\varepsilon$ in the sequence mentioned above we apply (7) and (11) to obtain

$$P\{V'\leq x\}\geq \lim_{j\to\infty}P\{B_{r_{k(j)}}\leq x\} = P\{A(\varepsilon)\leq x\} \geq P\{V(\varepsilon)\leq x\}.$$

Letting $\varepsilon \to 0$, we see that $P\{V'\leq x\} \geq P\{V\leq x\}$ for x in a dense set. The reverse inequality can be obtained by considering similar arguments for the lower bound on $V_r$. So $V = V'$ in distribution and this part of the proof is completed.

The only step remaining is to show that slogcom $(0)\{(s,W(s)+|s|^{\alpha+1}): -cD \leq s \leq cD\}$ converges to slogcom $(0)\{(s,W(s)+|s|^{\alpha+1}): -\infty < s < \infty\}$, but this is established in Wright (1981).

Casady and Cryer (1976) have shown that with probability one, $\limsup_{r\to\infty}(r/\log\log r)^{1/4}(\hat{\theta}_r(x_0)-\theta(x_0)) \leq k < \infty$ provided $\theta$ is Lipschitz of order 1. It would be interesting to determine if the exponent could be increased to 1/3 when $\theta'(x_0) > 0$. Of course, in this case, it can not be larger than 1/3.

## REFERENCES

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference Under Order Restrictions, New York: Wiley.

Brunk, H. D. (1958). On the estimation of parameters restricted by inequalities. Ann. Math. Statist. 29, 437-454.

Brunk, H. D. (1970). Estimation of isotonic regression. Nonparametric Techniques in Statistical Inference. 177-195, Cambridge University Press.

Cryer, J. D., Robertson, Tim, Wright, F. T. and Casady, R. J. (1972). Monotone median regression. Ann. Math. Statist. 43, 1459-1469.

Casady, R. J. and Cryer, J. D. (1976). Monotone percentile regression. Ann. Statist. 4, 532-541.

Gabriel, J. P. (1977). An inequality for sums of independent random variables indexed by finite dimensional filtering sets and its applications to the convergence of series. Ann. Probability 5, 779-786.

Ghosh, J. K. (1971). A new proof of the Bahadur representation of quantiles and an application. Ann. Math Statist. 42, 1957-1961.

Hanson, D. L., Pledger, Gordor, and Wright, F. T. (1973). On consistency in monotonic regression. Ann. Statist. 1, 401-421.

Leurgans, S. E. (1979). Asymptotic distribution of slope of greatest convex minorant estimators. Technical Report, Mathematics Research Center, University of Wisconsin-Madison.

Makowski, G. G. (1973). Laws of the iterated logarithm for permuted random variables and regression applications. Ann. Statist. 1, 872-887.

Prakasa Rao, B. L. S. (1969). Estimation of a unimodel density. Sankhyā (A) 31, 23-36.

Robertson, Tim and Waltman, Paul (1968). On estimating monotone parameters. Ann. Math. Statist. 39, 1030-1039.

Robertson, Tim and Wright, F. T. (1975). Consistency in generalized isotonic regression. Ann. Statist 3, 350-362.

Wright, F. T. (1981). The asymptotic behavior of monotone regression estimates. Ann. Statist. 9, 443-448.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>4 | 2. GOVT ACCESSION NO.<br>AD-A100 | 3. RECIPIENT'S CATALOG NUMBER<br>916 |
| 4. TITLE (and Subtitle)<br>The Asymptotic Behavior of Monotone Percentile Regression Estimates | | 5. TYPE OF REPORT & PERIOD COVERED |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>F. T. Wright | | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-80-C0322 new |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Department of Mathematics<br>University of Missouri-Rolla<br>Rolla, MO 65401 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research<br>Statistics and Probability Program Code 436<br>Arlington, Virginia 22217 | | 12. REPORT DATE<br>May 28, 1981 |
| | | 13. NUMBER OF PAGES<br>14 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

monotone regression, percentile regression, asymptotic distributions, rate of convergence

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The least absolute deviations estimate of a monotone regression function was derived by Robertson and Waltman (1968). Assuming the observation points become dense in the domain of the regression function, Casady and Cryer (1976) obtained an upper bound on the almost sure rate of convergence of the estimator. The asymptotic distribution of the estimator at a point is obtained here and a faster rate of convergence is obtained.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

F

8