The Annals of Statistics
1976, Vol. 4, No. 3, 532-541

MONOTONE PERCENTILE REGRESSION[1]

By Robert J. Casady and Jonathan D. Cryer

University of Iowa

Suppose that for each number t in [0, 1] there is a distribution with distribution function $F_t(\cdot)$ which has pth percentile $\xi(t)$. Consider the problem of estimating $\xi(\cdot)$ under the assumption that $\xi(\cdot)$ is monotone. An estimator which is analogous to the median regression estimator considered in Cryer, Robertson, Wright and Casady (1972), is studied. Asymptotic properties including consistency and law of the iterated logarithm results are obtained under various assumptions.

1. Introduction. Suppose that for each number t in [0, 1] we have a distribution with distribution function $F_t(\cdot)$. Let $\{t_j\}_{j=1}^{\infty}$ be a sequence of numbers in [0, 1], not necessarily distinct, to be called observation points, and let $\{Y_j\}_{j=1}^{\infty}$ be a sequence of independent random variables such that the distribution function of Y_j is $F_{t_j}(\cdot)$. For some fixed number p, such that $0 < p < 1$, and for each t in [0, 1] let $\xi(t)$ be the pth percentile of the distribution function $F_t(\cdot)$. In what follows we will adopt the definition

$$\xi(t) = \inf\{x \mid F_t(x) \geqq p\}$$

and the function $\xi(\cdot)$ will be referred to as the pth *percentile regression function*. Also, for the remainder of the paper, $f_t(\cdot)$ will denote the density of $F_t(\cdot)$ with respect to Lebesgue measure and $f_t'(\cdot)$ will denote $f_t'(x) = dF_t(x)/dx$.

In many experimental or survey situations, such as a study of growth in preschool children or a statewide survey of mathematical achievement for children in the primary grades, an estimator of the function $\xi(\cdot)$ is desired. The classical parametric approach is to assume that $\xi(\cdot)$ has some rather simple functional representation that depends on a few unknown parameters. For example, $\xi(t) = \alpha t + \beta$ where α and β are unknown parameters. These parameters are then estimated by some method such as least squares, minimum absolute deviations, or even maximum likelihood if further assumptions are made concerning $F_t(\cdot)$. We wish to pursue a nonparametric approach and assume only that $\xi(\cdot)$ is monotone nondecreasing.

Cryer, Robertson, Wright and Casady (1972) suggested an estimator for a monotone *median* regression function which is analogous to the estimator of a monotone *mean* regression function as given in Brunk (1969) and further investigated in Makowski (1973). Motivated by this estimator of the median

Received October 1974; revised November 1975.

[1] This research was supported by the National Science Foundation under Grant GP-30966X and forms a portion of the first author's Ph. D. dissertation at the University of Iowa.

AMS 1970 *subject classifications*. Primary 62G05; Secondary 60F15.

Key words and phrases. Monotone regression, law of the iterated logarithm, percentile estimation, consistency.

regression function the following estimator of the pth percentile regression function is proposed:

$$\hat{\xi}_n(t_i) = \max_{r \leqq t_i} \min_{t_i \leqq s} Z_p\{Y_j \mid j \leqq n, r \leqq t_j \leqq s\}$$
$$= \min_{t_i \leqq s} \max_{r \leqq t_i} Z_p\{Y_j \mid j \leqq n, r \leqq t_j \leqq s\}$$

where Z_p denotes the pth percentile of the empirical distribution function of the random variables described inside the braces. The equality of the two representations for $\hat{\xi}_n(t_i)$ follows from the work of Robertson and Waltman (1968). The value of $\hat{\xi}_n(t_i)$ at values of t between observation points can be specified in any manner so long as the estimator function remains monotone.

In Section 2, the general monotone percentile regression model is used to estimate the pth percentile for each of a finite number of populations when it is assumed that the pth percentiles are ordered. The proposed estimators are shown to be uniformly strongly consistent and a law of the iterated logarithm is shown to hold, thus establishing a rate of convergence for the estimators.

Section 2 also establishes that the proposed estimator of the monotone percentile regression function is uniformly strongly consistent on any interval $[a, b]$ where $0 < a < b < 1$. This result is an extension of a similar result for the estimator of a monotone median regression function as given in Cryer, Robertson, Wright and Casady (1972). It should also be mentioned that all of the other results in their paper are easily extended to the case of estimating monotone percentile regression functions. Also in this section a pointwise law of the iterated logarithm type result is proven under the assumption that $\xi(\cdot)$ is Lipschitz of order one.

Finally, in Section 3, a more general regression situation is considered in which the observation points are also considered to be random variables. The results previously mentioned for Section 2 are shown to hold for this more general model.

2. The main results. In this section we will first consider the problem of estimating the pth percentile, ξ_i, $i = 1, 2, \cdots, k$, for each of k populations when we assume that $\xi_1 \leqq \xi_2 \leqq \cdots \leqq \xi_k$. As an example, one might be interested in estimating the tenth percentile score on a standard achievement test for students in each for the six primary school grades.

In terms of our general monotone regression model, the preceding problem is equivalent to estimating $\xi(\cdot)$ at only a finite number of observation points, say $0 \leqq s_1 \leqq s_2 \leqq \cdots \leqq s_k \leqq 1$. The number of observations from population i, out of a total of n observations, will be denoted by $N_i(n)$.

The pointwise strong consistency of the estimators is established in the following theorem.

THEOREM 2.1. *If for i such that $1 \leqq i \leqq k$, $F_{s_i}(\xi(s_i)) = p$, $f_{s_i}(\xi(s_i)) > 0$ and $N_i(n) \to \infty$ as $n \to \infty$, then*

$$P\{\lim_{n\to\infty} |\hat{\xi}_n(s_i) - \xi(s_i)| = 0\} = 1 .$$

The proof of this theorem is implicit in Robertson and Waltman (1968) and depends primarily on well-known results concerning the almost sure convergence of sample percentiles to the population percentiles.

If the assumptions of Theorem 2.1 hold uniformly in i, then the following corollary establishing the uniform strong consistency of the estimators follows directly from Theorem 2.1.

COROLLARY 2.2. *If for each i such that $1 \leqq i \leqq k$, $F_{s_i}(\xi(s_i)) = p$, $f_{s_i}(\xi(s_i)) > 0$ and $N_i(n) \to \infty$ as $n \to \infty$, then*

$$P\{\lim_{n\to\infty} \max_{1\leqq i\leqq k} |\hat{\xi}_n(s_i) - \xi(s_i)| = 0\} = 1 .$$

The following theorem establishing a law of the iterated logarithm result for the monotone percentile regression estimators was motivated by the fact that a law of the iterated logarithm holds for the usual percentile estimators.

THEOREM 2.3. *If for i such that $1 \leqq i \leqq k$, $f_{s_i}(\xi(s_i)) > 0$, $f'_{s_i}(\cdot)$ is bounded in a neighborhood of $\xi(s_i)$ and $\liminf_n N_i(n)/n > 0$, then there exists a positive finite constant K such that*

$$P\{\limsup_n (N_i(n)/\log\log N_i(n))^{\frac{1}{2}} |\hat{\xi}_n(s_i) - \xi(s_i)| \leqq K\} = 1 .$$

The proof of this theorem is similar to the proof of Theorem 2.7 given later and hence will not be given here. The proof may be found in Casady (1972).

The following corollary is a direct consequence of Theorem 2.3.

COROLLARY 2.4. *Suppose that $\min_{1\leqq i\leqq k} f_{s_i}(\xi(s_i)) > 0$, there exist positive constants β and δ such that*

$$\max_{1\leqq i\leqq k} (\sup\{|f'_{s_i}(x)| : x \in [\xi(s_i) - \delta, \xi(s_i) + \delta]\}) < \infty$$

and

$$\min_{1\leqq i\leqq k} N_i(n)/n \geqq \beta \qquad \text{for all } n \text{ sufficiently large.}$$

Then there exists a finite constant K such that

$$P\{\limsup_n (n/\log\log n)^{\frac{1}{2}} \max_{1\leqq i\leqq k} |\hat{\xi}_n(s_i) - \xi(s_i)| \leqq K\} = 1 .$$

Thus far we have dealt with the convergence of $\hat{\xi}_n(\cdot)$ to $\xi(\cdot)$ at a finite number of distinct observation points when we assumed that the number of observations at each of these points becomes large. We now consider the same convergence under the general monotone percentile regression model. That is, we will consider the convergence of $\hat{\xi}_n(t)$ to $\xi(t)$ for each point t in (0, 1). It will be assumed, analogous to the previously mentioned assumption, that the number of observation points in each subinterval of (0, 1) becomes large. The following definition enables us to make this assumption more precise.

DEFINITION 2.5. For any number $t \in [0, 1]$ let $G_n(t) = n^{-1} \cdot \text{cardinality}\{j \mid 1 \leqq j \leqq n, t_j \leqq t\}$. That is, $G_n(\cdot)$ is the empirical distribution function of the first n observation points.

For the remainder of this section assume that for all t' and t'' such that $0 < t' < t'' < 1$

$$\liminf_n (G_n(t'') - G_n(t')) > 0 . \tag{2.1}$$

Using this assumption, Cryer, Robertson, Wright and Casady (1972) proved several theorems and corollaries dealing with consistency properties of the estimator of a monotone *median* regression function. This is a special case of monotone percentile regression and the proofs for the results in Cryer, Robertson, Wright and Casady (1972) can easily be altered for the more general case. For this reason the following theorem will be given without the proof.

THEOREM 2.6. *Suppose that the* pth *percentile function* $\xi(\cdot)$ *is continuous on* [0, 1], *for each positive number* ε

$$\inf_k F_{t_k}(\xi(t_k) + \varepsilon) - p > 0 , \tag{2.2}$$

$$p - \sup_k F_{t_k}(\xi(t_k) - \varepsilon) > 0 , \tag{2.3}$$

and (2.1) *holds. Then for* $0 < a < b < 1$

$$P\{\lim_n \sup_{a \leq t \leq b} |\hat{\xi}_n(t) - \xi(t)| = 0\} = 1 .$$

Our main interest is in establishing a rate of convergence by proving a law of the iterated logarithm-type result for the estimator $\hat{\xi}_n(\cdot)$. The proof of this result requires not only that the number of observation points in each interval become large but that this property hold uniformly on all intervals exceeding a certain length, say l_n, where $l_n \rightarrow 0$ as n becomes large. More precisely, it will be assumed that there exist positive constants β and c such that for n sufficiently large

$$(G_n(t'') - G_n(t'))/(t'' - t') \geqq \beta \tag{2.4}$$

for all t', t'' such that $|t'' - t'| > cn^{-\frac{1}{2}}$. Example 2.10 will show that this hypothesis is not vacuous.

THEOREM 2.7. *Suppose that* $\xi(\cdot)$ *is Lipschitz of order* 1 *on* [0, 1],

$$0 < \alpha_1 = \inf_k f_{t_k}(\xi(t_k)) \leqq \sup_k f_{t_k}(\xi(t_k)) = \alpha_2 < \infty , \tag{2.5}$$

and for some $\delta > 0$

$$\sup_k (\sup\{|f'_{t_k}(x)| : x \in [\xi(t_k) - \delta, \xi(t_k) + \delta]\}) = M < \infty . \tag{2.6}$$

Further, assume (2.4) *holds. Then for each t in* (0, 1) *there exists a* $K < \infty$ *such that*

$$P[\limsup_n (n/\log\log n)^{\frac{1}{2}} |\hat{\xi}_n(t) - \xi(t)| \leqq K\} = 1 .$$

PROOF. Let t be a fixed but arbitrary point in (0, 1), $\delta_n = (\log\log n/n)^{\frac{1}{2}}$, $S_n = t + \delta_n$ and $\varepsilon_n = (K_1 \log\log n/n)^{\frac{1}{2}}$ where K_1 is a positive constant to be chosen later. Now choose λ and γ such that $1 < \gamma < \lambda$, and note that

$$\begin{aligned} \hat{\xi}_n(t) - \xi(S_n) &= \max_{r \leq t} \min_{t \leq s} Z_p(Y_j \mid j \leqq n, r \leqq t_j \leqq s) - \xi(S_n) \\ &\leqq \min_{t \leq s \leq S_n} \max_{r \leq t} Z_p(Y_j \mid j \leqq n, r \leqq t_j \leqq s) - \xi(S_n) \\ &= \min_{t \leq s \leq S_n} \max_{r \leq t} Z_p(Y_j - \xi(S_n) \mid j \leqq n, r \leqq t_j \leqq s) \\ &\leqq \min_{t \leq s \leq S_n} \max_{r \leq t} Z_p(Y_j - \xi(t_j) \mid j \leqq n, r \leqq t_j \leqq s) \end{aligned}$$

so then we have

$$\{\hat{\xi}_n(t) - \xi(S_n) > \lambda\varepsilon_n\}$$
$$\subset \{\min_{t \leqq s \leqq S_n} \max_{r \leqq t} Z_p(Y_j - \xi(t_j) \mid j \leqq n, r \leqq t_j \leqq s) > \lambda\varepsilon_n\}$$
$$= \{\max_{r \leqq t} \min_{t \leqq s \leqq S_n} Z_p(Y_j - \xi(t_j) \mid j \leqq n, r \leqq t_j \leqq s) > \lambda\varepsilon_n\} .$$

For $j = 1, 2, \cdots$ and $n = 1, 2, \cdots$, define

$$X_{j,n} = 1 \quad \text{if} \quad Y_j \leqq \xi(t_j) + \lambda\varepsilon_n$$
$$= 0 \quad \text{otherwise,}$$

and for all real numbers a, b and x such that $a \leqq b$ let $I(a, b, x) = 1$ if $a \leqq x \leqq b$ and zero otherwise. Also, for $a \leqq b$ let $N(a, b, n) = \sum_{j=1}^n I(a, b, t_j)$ then for s' and r' such that $r' \leqq t \leqq s' \leqq S_n$

$$\{Z_p(Y_j - \xi(t_j) \mid j \leqq n, r' \leqq t_j \leqq s') > \lambda\varepsilon_n\}$$
$$\subset \{\sum_{j=1}^n I(r', s', t_j)X_{j,n} < N(r', s', n) \cdot p\}$$
$$= \{\sum_{j=1}^n I(r', s', t_j)(X_{j,n} - p) < 0\} .$$

Hence it follows that

$$\{\hat{\xi}_n(t) - \xi(S_n) > \lambda\varepsilon_n\} \subset \{\min_{r \leqq t} \max_{t \leqq s \leqq S_n} \sum_{j=1}^n I(r, s, t_j)(X_{j,n} - p) < 0\}$$
$$= \{\max_{r \leqq t} \min_{t \leqq s \leqq S_n} \sum_{j=1}^n I(r, s, t_j)(p - X_{j,n}) > 0\}$$
$$\equiv A_n .$$

Let $m_k = [\gamma^k]$ for $k = 1, 2, \cdots$, where $[x]$ denotes the integer part of x and let $r_{n,1} = t$ and $r_{n,2} > r_{n,3} > \cdots > r_{n,i_n} \geqq 0$ be the unique values of the t_j's such that $j \leqq n$ and $t_j < t$. Temporarily fix k and for n such that $m_k \leqq n \leqq m_{k+1}$ and l such that $1 \leqq l \leqq i_n$ define,

$$B_{n,l} = \{\max_{m_k \leqq m < n} \max_{r \leqq t} \min_{t \leqq s \leqq S_n} \sum_{j=1}^m I(r, s, t_j)(p - X_{j,n}) \leqq 0 ,$$
$$\max_{1 \leqq i < l} \min_{t \leqq s \leqq S_n} \sum_{j=1}^n I(r_{n,i}, s, t_j)(p - X_{j,n}) \leqq 0 ,$$
$$\min_{t \leqq s \leqq S_n} \sum_{j=1}^n I(r_{n,l}, s, t_j)(p - X_{j,n}) > 0\} .$$

It is easy to verify that $B_{n',l'} \cap B_{n'',l''} = \emptyset$ if either $n' \neq n''$ or $l' \neq l''$ and that $\bigcup_{n=m_k}^{m_{k+1}-1} A_n \subset \sum_{n=m_k}^{m_{k+1}} \sum_{l=1}^{i_n} B_{n,l}$. For n and l as before define

$$U_{n,l} = \{\sum_{j=n+1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - X_{j m_{k+1}}) \geqq 0\}$$

when $\sum_{j=1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j) \geqq 1$ and $n < m_{k+1}$. Otherwise, let $U_{n,l}$ be the sure event. On $U_{n,l} \cap B_{n,l}$ we have

$$\min_{t \leqq s \leqq S_n} \sum_{j=1}^n I(r_{n,l}, s, t_j)(p - X_{j,n}) > 0 ,$$

which implies

$$(2.7) \qquad \sum_{j=1}^n I(r_{n,l}, S_{m_{k+1}}, t_j)(p - X_{j,n}) > 0$$

because $S_{m_{k+1}} \leqq S_n$. Also, $X_{j,n} \geqq X_{j,m_{k+1}}$ by definition so that (2.7) implies

$$(2.8) \qquad \sum_{j=1}^n I(r_{n,l}, S_{m_{k+1}}, t_j)p \geqq \sum_{j=1}^n I(r_{n,l}, S_{m_{k+1}}, t_j)X_{j,m_{k+1}} .$$

But we also have on $U_{n,l} \cap B_{n,l}$

$$(2.9) \qquad \sum_{j=n+1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)EX_{j,m_{k+1}} \geqq \sum_{j=n+1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)X_{j,m_{k+1}},$$

so combining (2.8) and (2.9) we get

$$(2.10) \qquad \begin{aligned} \sum_{j=1}^{n} I(r_{n,l}, S_{m_{k+1}}, t_j)p + \sum_{j=n+1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)EX_{j,m_{k+1}} \\ = \sum_{j=1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)X_{j,m_{k+1}} . \end{aligned}$$

Finally, adding $\sum_{j=1}^{n} I(r_{n,l}, S_{m_{k+1}}, t_j)EX_{j,m_{k+1}}$ to both sides of (2.10) and rearranging terms we get

$$(2.11) \qquad \begin{aligned} \sum_{j=1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - X_{j,m_{k+1}}) \\ \geqq \sum_{j=1}^{n} I(r_{n,l}, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - p) . \end{aligned}$$

But $EX_{j,m_{k+1}} - p = F_{t_j}(\xi(t_j) + \lambda\varepsilon_{n_{k+1}}) - p \geqq 0$, $n \geqq m_k$ and $r_{n,l} \leqq t$ so that (2.11) implies

$$(2.12) \qquad \begin{aligned} \sum_{j=1}^{m_{k+1}} I(r_{n,l}, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - X_{j,m_{k+1}}) \\ \geqq \sum_{j=1}^{m_k} I(t, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - p) . \end{aligned}$$

Now let

$$\begin{aligned} C_k = \{\max_{r \leqq t} \sum_{j=1}^{m_{k+1}} I(r, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - X_{j,m_{k+1}}) \\ \geqq \sum_{j=1}^{m_k} I(t, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - p)\} \end{aligned}$$

and note that $r_{n,l} \in \{t_j \mid j \leqq m_{k+1}, t_j \leqq t\}$. Hence, (2.12) implies that $B_{n,l} \cap U_{n,l} \subset C_k$ for all n and l such that $m_k \leqq n \leqq m_{k+1}$, $1 \leqq l \leqq i_n$. Thus we have

$$(2.13) \qquad \begin{aligned} P\{C_k\} &\geqq P\{\sum_{n=m_k}^{m_{k+1}} \sum_{l=1}^{i_n} B_{n,l} \cap U_{n,l}\} \\ &= \sum_{n=m_k}^{m_{k+1}} \sum_{l=1}^{i_n} P\{B_{n,l} \cap U_{n,l}\} \\ &= \sum_{n=m_k}^{m_{k+1}} \sum_{l=1}^{i_n} P\{B_{n,l}\} \cdot P\{U_{n,l}\} , \end{aligned}$$

where the last equality holds since $B_{n,l}$ and $U_{n,l}$ are independent events. Also, for all j and n

$$\begin{aligned} EX_{j,n} &= F_{t_j}(\xi(t_j) + \lambda\varepsilon_n) \\ &= F_{t_j}(\xi(t_j)) + \lambda\varepsilon_n f_{t_j}(\xi(t_j)) + (\lambda\varepsilon_n)^2 f'_{t_j}(\theta_{j,n})/2 , \end{aligned}$$

where $t \leqq \theta_{j,n} \leqq t + \lambda\varepsilon_n$. Hence for n so large that $\lambda\varepsilon_n < \delta$ we have

$$|EX_{j,n} - p| \leqq \lambda\varepsilon_n \alpha_2 + (\lambda\varepsilon_n)^2 M/2 ,$$

which implies that $\lim_n \sup_j EX_{j,n} = p$. An application of the Berry–Esséen theorem shows there exists a Γ such that

$$0 < \Gamma \leqq \min\{P(U_{n,l}) : 1 \leqq l \leqq i_n, m_k \leqq n \leqq m_{k+1}\}$$

for k sufficiently large. Hence (2.13) implies that

$$(2.14) \qquad \begin{aligned} \Gamma^{-1}P\{C_k\} &\geqq \sum_{n=m_k}^{m_{k+1}} \sum_{l=1}^{i_n} P\{B_{n,l}\} \\ &\geqq P\{\bigcup_{n=m_k}^{m_{k+1}-1} A_n\} . \end{aligned}$$

Next it will be shown that $\sum_{k=1}^{\infty} P\{C_k\} < \infty$. Then by (2.14) and the Borel–Cantelli theorem we have $P\{\bigcup_{n=m_k}^{m_{k+1}-1} A_n \text{ i.o. } k\} = 0$ and hence $P\{A_n \text{ i.o. } n\} = 0$. Let

$$\lambda_k = (\textstyle\sum_{j=1}^{m_k} I(t, S_{m_{k+1}}, t_j)(EX_{j,m_{k+1}} - p))/N(0, S_{m_{k+1}}, m_{k+1}) ,$$

and then by an extension of Bernstein's inequality (see Hoeffding (1963)) we have

$$P\{C_k\} \leqq \exp(-N(0, S_{m_{k+1}}, m_{k+1}) \cdot \lambda_k^2/2(\sigma_k^2 + \lambda_k/3)) ,$$

where $\sigma_k^2 = \sum_{j=1}^{m_{k+1}} I(0, S_{m_{k+1}}, t_j) \operatorname{Var}(X_{j,m_{k+1}}/N(0, S_{m_{k+1}}, m_{k+1}))$. To show $\sum_{k=1}^{\infty} P\{C_k\} < \infty$ it is sufficient to show $\liminf_k N(0, S_{m_{k+1}}, m_{k+1})\lambda_k^2/(2(\sigma_k^2 + \lambda_k/3) \log k) > 1$, but using the uniform convergence of $EX_{j,n}$ to p it is easy to verify that $\lim_k \lambda_k = 0$ and $\lim_k \sigma_k^2 = pq$, hence we need only show

$$\liminf_k N(0, S_{m_{k+1}}, m_{k+1})\lambda_k^2/2pq \log k > 1 .$$

For k sufficiently large

$$\begin{aligned} EX_{j,m_{k+1}} - p &\geqq p + \lambda\varepsilon_{m_{k+1}}\alpha_1 - (\lambda\varepsilon_{m_{k+1}})^2 M/2 - p \\ &= \lambda\varepsilon_{m_{k+1}}\alpha_1 - (\lambda\varepsilon_{m_{k+1}})^2 M/2 . \end{aligned}$$

Hence for large k

$$\lambda_k \geqq N(t, S_{m_{k+1}}, m_k)(\lambda\varepsilon_{m_{k+1}}\alpha_1 - (\lambda\varepsilon_{m_{k+1}})^2 M/2)/N(0, S_{m_{k+1}}, m_{k+1}) .$$

Also,

$$\begin{aligned} S_{m_{k+1}} - t &= (\log\log[\gamma^{k+1}]/[\gamma^{k+1}])^{\frac{1}{2}} \\ &= ([\gamma^k]\log\log[\gamma^{k+1}]/[\gamma^{k+1}]^{\frac{1}{2}})m_k^{-\frac{1}{2}} . \end{aligned}$$

But $[\gamma^k]\log\log[\gamma^{k+1}]/[\gamma^{k+1}]$ diverges to $+\infty$. Hence for k large we have $S_{m_{k+1}} - t \geqq cm_k^{-\frac{1}{2}}$. Using assumption (2.4) we have

$$\begin{aligned} N(t, S_{m_{k+1}}, m_k) &\geqq (G_{m_k}(S_{m_{k+1}}) - G_{m_k}(t))m_k \\ &\geqq \beta(S_{m_{k+1}} - t)m_k . \end{aligned}$$

Thus $\lambda_k \geqq \beta(S_{m_{k+1}} - t)m_k(\lambda\varepsilon_{m_{k+1}}\alpha_1 - (\lambda\varepsilon_{m_{k+1}})^2 M/2)/N(0, S_{m_{k+1}}, m_{k+1})$ so that

$$\begin{aligned} &N(0, S_{m_{k+1}}, m_{k+1})\lambda_k^2/2pq \log k \\ &\quad\geqq (\beta(S_{m_{k+1}} - t)m_k(\lambda\varepsilon_{m_{k+1}}\alpha_1 - (\lambda\varepsilon_{m_{k+1}})^2 M/2))^2/(N(0, S_{m_{k+1}}, m_{k+1}) \cdot 2pq \log k) \\ &\quad\geqq (\beta(S_{m_{k+1}} - t)m_k(\lambda\varepsilon_{m_{k+1}}\alpha_1 - (\lambda\varepsilon_{m_{k+1}})^2 M/2))^2/2pqm_{k+1} \log k . \end{aligned}$$

It is easily verified that terms involving $\varepsilon_{m_{k+1}}$ to a power of three or greater vanish as $k \to \infty$ so that

$$\begin{aligned} &\liminf_k N(0, S_{m_{k+1}}, m_{k+1})\lambda_k^2/2pq \log k \\ &\qquad\geqq \liminf_k (\beta(S_{m_{k+1}} - t)m_k \lambda\varepsilon_{m_{k+1}}\alpha_1)^2/2pqm_{k+1} \log k \\ &\qquad= \liminf_k (\beta\lambda\alpha_1)^2 K_1^{\frac{1}{2}}[\gamma^k]^2 \log\log[\gamma^{k+1}]/2pq[\gamma^{k+1}]^2 \log k \\ &\qquad= ((\beta\lambda\alpha_1/\gamma)^2 K_1^{\frac{1}{2}}/2pq) \liminf_k (\log\log\gamma^{k+1}/\log k) \\ &\qquad= (\lambda/\gamma)^2((\beta\alpha_1)^2/2pq)K_1^{\frac{1}{2}} , \end{aligned}$$

where the first equality follows directly from the definitions. Let $K_1 = (2pq/(\beta\alpha_1)^2)^2$

and note that $(\lambda/\gamma) > 1$, then the desired result follows immediately. Thus, it has been shown that

$$P\{\hat{\xi}_n(t) - \xi(S_n) > \lambda\varepsilon_n \text{ i.o. } n\} = 0$$

where $\lambda > 1$ was arbitrary, so that

$$P\{\limsup_n (\hat{\xi}_n(t) - \xi(S_n))/\varepsilon_n \leqq 1\} = 1 .$$

By the Lipschitz condition there exists a $C_0 > 0$ such that $|\xi(t') - \xi(t'')| \leqq C_0|t' - t''|$ for all t', t'' in [0, 1] so that

$$\begin{aligned}\limsup_n (\hat{\xi}_n(t) - \xi(t))/\varepsilon_n &\leqq \limsup_n (\hat{\xi}_n(t) - \xi(S_n))/\varepsilon_n + \limsup_n (\xi(S_n) - \xi(t))/\varepsilon_n \\ &\leqq \limsup_n (\hat{\xi}_n(t) - \xi(S_n))/\varepsilon_n + C_0/K_1^{\frac{1}{4}} .\end{aligned}$$

Hence,

$$P\{\limsup_n (n/\log\log n)^{\frac{1}{4}}(\hat{\xi}_n(t) - \xi(t)) \leqq K_1^{\frac{1}{4}} + C_0\} = 1$$

and in a similar manner it can be shown that

$$P\{\limsup_n (n/\log\log n)^{\frac{1}{4}}(\xi(t) - \hat{\xi}_n(t)) \leqq K_1^{\frac{1}{4}} + C_0\} = 1$$

and we conclude that

$$P\{\limsup_n (m/\log\log n)^{\frac{1}{4}}|\hat{\xi}_n(t) - \xi(t)| \leqq K\} = 1$$

where $K = K_1^{\frac{1}{4}} + C_0$, and thus the theorem is proved.

The following example will illustrate a case in which assumptions (2.5) and (2.6) hold.

EXAMPLE 2.8. Suppose $F(\cdot)$ is a twice differential distribution function, $F(0) = p$, $F'(0) = \alpha > 0$ and for some $\delta > 0$, $\sup_{x\in[-\delta,\delta]} |F''(x)| = M < \infty$. Further, suppose that $\xi(\cdot)$ is a nondecreasing function on [0, 1] and $\sigma(\cdot)$ is a function on [0, 1] such that $0 < a = \inf_t \sigma(t) \leqq \sup_t \sigma(t) = b < \infty$. For each t in [0, 1] and for all real x let $F_t(x) = F((x - \xi(t))/\sigma(t))$, then $F_t(\xi(t)) = F(0) = p$,

$$f_t(x) = [F'((x - \xi(t))/\sigma(t))]/\sigma(t) \tag{2.15}$$

and

$$f_t'(x) = [F''((x - \xi(t))/\sigma(t))]/(\sigma(t))^2 . \tag{2.16}$$

From (2.15) we have

$$0 \leqq \alpha/b \leqq F_t'(\xi(t)) \leqq \alpha/a < \infty$$

which implies that (2.5) holds. If $\delta' = a\delta > 0$ then (2.16) implies that

$$\begin{aligned}\sup\{|f_t'(x)| : x \in [\xi(t) - \delta', \xi(t) + \delta']\} &\leqq \sup_{x\in[-\delta,\delta]} |F''(t)/(\sigma(t))^2| \\ &\leqq M/a^2 < \infty\end{aligned}$$

and hence assumption (2.6) holds.

It is not clear such sequences as assumed in (2.4) exist, however, the following theorem by Smirnov (1944) establishes the existence of such sequences.

THEOREM 2.9 (Smirnov). *Let $G_n(\cdot)$ be the empirical distribution function of a random sample $X_1, X_2, \cdots, X_n$ with distribution function $G(\cdot)$. If $G(\cdot)$ is continuous then*

$$P\{\limsup_n (n/\log\log n)^{\frac{1}{2}} \sup_x |G_n(x) - G(x)| \leqq (\tfrac{1}{2})^{\frac{1}{2}}\} = 1 .$$

EXAMPLE 2.10. Assume in addition to the hypotheses of Theorem 2.9 that $G(0) = 0$, $G(1) = 1$, $G'(\cdot)$ exists on (0, 1) and $0 < \beta^* = \inf_{x \in (0,1)} G'(x)$. Theorem 2.9 implies that for almost every observed sample sequence $x_1, x_2, \cdots$ we have for n sufficiently large,

$$\begin{aligned}(G(t'') - G(t'))/(t'' - t') - 2K(\log\log n/n)^{\frac{1}{2}}/(t'' - t') \\ \leqq (G_n(t'') - G_n(t'))/(t'' - t')\end{aligned}$$

where $(\frac{1}{2})^{\frac{1}{2}} < K < \infty$ and $0 < t' < t'' < 1$. There exists a θ such that $t' \leqq \theta \leqq t''$ and $(G(t'') - G(t'))/(t'' - t') = G'(\theta) \geqq \beta^*$. Hence

$$\beta^* - 2K(\log\log n)^{\frac{1}{2}}/(t'' - t') \leqq (G_n(t'') - G_n(t'))/(t'' - t') .$$

Now assuming $t'' - t' \geqq n^{-\frac{1}{4}}$ we have

$$\beta^* - 2K(\log\log n)^{\frac{1}{2}}/n^{\frac{1}{4}} < (G_n(t'') - G_n(t'))/(t'' - t')$$

and as $(\log\log n)^{\frac{1}{2}}/n^{\frac{1}{4}} \to 0$ as $n \to \infty$ we have for n sufficiently large

$$\beta < (G_n(t'') - G_n(t'))/(t'' - t')$$

where $\beta = \beta^*/2 > 0$.

Example 2.10 suggests that a more general type of regression situation should be considered in which the sequence of observation points, as well as $\{Y_j\}_{j=1}^{\infty}$, are random variables. This concept will be discussed in the next section.

3. An independent observations regression model. The above results can be generalized to the situation where the observation points $\{t_j\}_{j=1}^{\infty}$ are considered as observed values of random variables $\{T_j\}_{j=1}^{\infty}$ as done in Brunk (1969). Brunk's discussion formalizes the idea that $\{T_j\}_{j=1}^{\infty}$ is a discrete parameter stochastic process and conditional on a realization $\{t_j\}_{j=1}^{\infty}$ of that process the $\{Y_j\}$ are independent random variables such that $F_{t_j}(y) = P\{Y_j \leqq y \mid T_j = t_j\}$. We refer the reader to Brunk's paper for the detailed construction of this so-called *independent observations regression model*. Let $G_n(\cdot)$ denote the empirical distribution for $T_1, T_2, \cdots, T_n$.

COROLLARY 3.1. *Let $\{T_j, Y_j\}_{j=1}^{\infty}$ be an independent observations regression model such that, with probability one, $\{F_{T_j}(\cdot)\}_{j=1}$ has properties* (2.2) *and* (2.3) *and either*

(i) $\{G_n(\cdot)\}_{n=1}^{\infty}$ *has property* (2.1)

or

(ii) $\{T_j\}_{j=1}^{\infty}$ *are independent, identically distributed with distribution function* $G(\cdot)$ *strictly increasing on* (0, 1).

Then for $0 < a < b < 1$

$$P\{\lim_n \sup_{a \leq t \leq b} |\hat{\xi}_n(t) - \xi(t)| = 0\} = 1 .$$

PROOF. Part (i) follows from the hypotheses, Theorem 2.6 and the definition of independent observations regression model. Part (ii) will follow from part (i) if we can verify that the empirical distribution functions $\{G_n(\cdot)\}_{n=1}^{\infty}$ have property (2.1) with probability one. To verify the above, let $0 < t' < t'' < 1$ and choose $\varepsilon > 0$ such that $\varepsilon < (G(t'') - G(t'))/2$. By the Glivenko-Cantelli theorem we have with probability one

$$-\varepsilon < G_n(t'') - G(t'') \tag{3.1}$$

$$-\varepsilon < G(t') - G_n(t'') \tag{3.2}$$

for n sufficiently large. Combining (3.1) and (3.2) we get

$$0 < G(t'') - G(t') - 2\varepsilon < G_n(t'') - G_n(t') .$$

Hence, with probability one (2.1) holds and the result is proved.

COROLLARY 3.2. *In the independent observations regression model assume that, with probability one,* $\{F_{T_j}(\cdot)\}_{j=1}^{\infty}$ *has properties* (2.5) *and* (2.6) *and that* $\xi(\cdot)$ *is Lipschitz of order one. Then if either*

(i) *with probability one,* $\{G_n(\cdot)\}_{n=1}^{\infty}$ *has property* (2.4)

or

(ii) $\{T_j\}_{j=1}^{\infty}$ *are independent, identically distributed with distribution function* $G(\cdot)$ *as in Example* 2.10,

then there is a constant K such that for each t in (0, 1)

$$P\{\lim \sup_n (n/\log \log n)^{\frac{1}{2}} |\hat{\xi}(t) - \xi(t)| \leq K\} = 1 .$$

PROOF. The proof of (i) follows directly from Theorem 2.7 and that of part (ii) follows from the results of Example 2.10 applied to part (i).

REFERENCES

[1] BRUNK, H. D. (1969). Estimation of isotonic regression. In *Nonparametric Techniques in Statistical Inference*. Cambridge Univ. Press, 177-195.

[2] CASADY, ROBERT J. (1972). Monotone percentile regression estimation: large sample and other properties. Ph. D. dissertation, Univ. of Iowa.

[3] CRYER, J. D., ROBERTSON, TIM, WRIGHT, F. T. and CASADY, ROBERT J. (1972). Monotone median regression. *Ann. Math. Statist.* **43** 1459-1469.

[4] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13-20.

[5] MAKOWSKI, GARY G. (1973). Laws of the iterated logarithm for permuted random variables and regression applications. *Ann. Statist.* **1** 872-887.

[6] ROBERTSON, TIM and WALTMAN, PAUL (1968). On estimating monotone parameters. *Ann. Math. Statist.* **39** 1030-1039.

[7] SMIRNOV, N. V. (1944). Approximation of the laws of distribution of random variables by empirical statistics. *Advances in Math. Sci.* **10** 179-206.

NATIONAL CENTER FOR HEALTH STATISTICS
5600 FISHERS LANE
ROCKVILLE, MARYLAND 20852

DEPARTMENT OF STATISTICS
UNIVERSITY OF IOWA
IOWA CITY, IOWA 52242