The Annals of Statistics
1976, Vol. 4, No. 1, 234–236

CONSISTENCY IN INTEGRAL REGRESSION ESTIMATION WITH A TRIANGULAR ARRAY OF OBSERVATION POINTS

BY GORDON PLEDGER

The University of Texas at Austin

Let μ be a continuous mean regression function defined on U, the unit cube in N-dimensional Euclidean space. Let F be a distribution function with support in U, and let M denote the indefinite integral of μ with respect to F. This paper provides consistency results, including rates of convergence, for a certain estimator of M in the case that the nth estimate is based on observations at points $\mathbf{t}_{n1}, \cdots, \mathbf{t}_{nn}$ of U. The estimator is the N-dimensional analogue of that considered by Brunk (1970).

Let N be a positive integer; let U denote the unit cube in N-dimensional Euclidean space; and, for $\mathbf{t}', \mathbf{t}''$ in U, let $[\mathbf{t}', \mathbf{t}''] = \{\mathbf{t} : \mathbf{t}' \leqq \mathbf{t} \leqq \mathbf{t}''\}$ where $\leqq$ is the usual partial ordering on U. Let μ be a real valued (Borel measurable) function defined on U, and for each positive integer n, let $\mathbf{t}_{n1}, \cdots, \mathbf{t}_{nn}$ be points in U. For each $\mathbf{t}$ in U let $Y(\mathbf{t})$ be a random variable with mean $\mu(\mathbf{t})$, and for each n let $Y_{n1}, \cdots, Y_{nn}$ be independent random variables such that Y_{ni} is distributed as $Y(\mathbf{t}_{ni})$. Here $\mathbf{t}_{ni}$ is to be thought of as an observation point and Y_{ni} as an observation at $\mathbf{t}_{ni}$.

Let $S_n(\mathbf{t})$ represent the sum of the observations made at points $\mathbf{t}_{ni} \leqq \mathbf{t}$; let F_n denote the empirical distribution function of $\mathbf{t}_{n1}, \cdots, \mathbf{t}_{nn}$; and let F be a distribution function with support in U. Define, for $\mathbf{t}$ in U,

$$M(\mathbf{t}) = \int_{[\mathbf{0},\mathbf{t}]} \mu(\mathbf{s})\, dF(\mathbf{s}) ,$$

the *integral regression function*, and

$$M_n(\mathbf{t}) = \int_{[\mathbf{0},\mathbf{t}]} \mu(\mathbf{s})\, dF_n(\mathbf{s}) = ES_n(\mathbf{t})/n .$$

Brunk (1970) proved that, under suitable conditions, $S_n(\mathbf{t})/n$ is uniformly, strongly consistent as an estimator of $M(\mathbf{t})$ in the case that $N = 1$ and $Y_{n+1,i} = Y_{ni}$ for $i = 1, \cdots, n$. For the same case, under stronger assumptions than Brunk's, Makowski (1973a) derived convergence rates of iterated logarithm type. Makowski (1973b) proved analogous results for $N > 1$. Both Brunk and Makowski considered permutations of a single sequence of random variables, a natural formulation arising, e.g., from the possibility that the kth observation may be made between two previous observation points. Commenting on Brunk's paper, Pyke (1970) suggested showing weak consistency when the observation points form a triangular array. It is shown below that the methods of Hanson,

Received December 1974; revised June 1975.
AMS 1970 *subject classifications*. Primary 62G05; Secondary 60G50.
Key words and phrases. Integral regression, multiple regression, consistency, triangular array.

Pledger and Wright (1973) yield such results, including rates of convergence, without explicitly dealing with the order in which the Y_{ij}'s are observed. Instead, a class of sequences of random variables is considered.

THEOREM. *Let G be a distribution function such that $G(0) = 0$ and $P(|Y(\mathbf{t}_{nk}) - \mu(\mathbf{t}_{nk})| \geqq x) \leqq 1 - G(x)$ for all $x \geqq 0$, $n = 1, 2, \cdots$; $k = 1, \cdots, n$. Let $\varepsilon > 0$.*

(a) *Assume either that $r = 1$ and $\int_0^\infty x\, dG(x) < \infty$, or that $r > 1$ and $x^r[1 - G(x)] \to 0$ as $x \to \infty$. Then as $n \to \infty$*

$$n^{r-1}P(\sup_{0 \leqq \mathbf{t} \leqq 1} |M_n(\mathbf{t}) - S_n(\mathbf{t})/n| \geqq \varepsilon) \to 0 .$$

(b) *If $\int_0^\infty e^{\theta x}\, dG(x) < \infty$ for some $\theta > 0$, then there exist $C > 0$, $0 < p < 1$, such that for $n = 1, 2, \cdots$,*

$$P(\sup_{0 \leqq \mathbf{t} \leqq 1} |M_n(\mathbf{t}) - S_n(\mathbf{t})/n| \geqq \varepsilon) \leqq Cp^n .$$

PROOF. Let $\{Z_i\}$ be a sequence of independent random variables, $A_n = Z_1 + \cdots + Z_n$, and denote a median of a random variable X by med (X). By Lemma 3 of Hanson, Pledger and Wright (1973) there exists J, depending only on G and ε, such that $\max_{k \leqq n} |\text{med}\,(A_k - A_n)/n| < \varepsilon/2$ for all $n \geqq J$ and all $\{Z_i\}$ satisfying

$$(1) \qquad EZ_i = 0 \quad \text{and} \quad P(|Z_i| \geqq x) \leqq 1 - G(x) \quad \text{for all} \quad x \geqq 0\,; \quad i = 1, 2, \cdots .$$

Then, for $n \geqq J$ and $\{Z_i\}$ satisfying (1), a Lévy inequality (see Loève (1963), page 247) implies that

$$\begin{aligned} P(\max_{k \leqq n} |A_k/n| \geqq \varepsilon) &\leqq P(\max_{k \leqq n} |A_k/n - \text{med}\,(A_k - A_n)/n| \geqq \varepsilon/2) \\ &\leqq 2P(|A_n/n| \geqq \varepsilon/2) . \end{aligned}$$

Note that $\sup_{0 \leqq \mathbf{t} \leqq 1} |M_n(\mathbf{t}) - S_n(\mathbf{t})/n| = \max_{k \leqq n} |M_n(\mathbf{t}_{nk}) - S_n(\mathbf{t}_{nk})/n|$ which has the form $\max_{k \leqq n} |A_k/n|$ where the sequence $\{Z_i\}$ changes with n. The conclusions of the theorem now follow from Lemmas 3, 4 and 5 of Hanson, Pledger and Wright (1973) which provide convergence rates *uniform* in sequences $\{Z_i\}$ satisfying (1).

Now suppose that, as in the *independent observations regression model* of Brunk (1970), the points $\mathbf{t}_{ni}$ are considered to be sample values of random vectors $\mathbf{T}_{ni}$, and let F_n be the empirical distribution function of $\mathbf{T}_{n1}, \cdots, \mathbf{T}_{nn}$. If F_n converges uniformly to F, and μ is continuous, then M_n can be replaced by M in part (a) of the theorem with $r = 1$. Suppose further that $\mathbf{T}_{n1}, \cdots, \mathbf{T}_{nn}$ is a random sample from F. Then theorem $1 - m$ of Kiefer (1961) implies that for $\delta > 0$, $\alpha > 0$,

$$P_F(\sup_{0 \leqq \mathbf{t} \leqq 1} |F(\mathbf{t}) - F_n(\mathbf{t})| \geqq \alpha) \leqq C(\delta, N)[e^{-(2-\delta)\alpha^2}]^n .$$

Hence the convergence rates in the theorem, with different values of C and p, carry over when M_n is replaced by M.

REFERENCES

[1] BRUNK, H. D. (1970). Estimation of isotonic regression. *Nonparametric Techniques in Statistical Inference*, 177–195. Cambridge Univ. Press.

[2] HANSON, D. L., PLEDGER, G. and WRIGHT, F. T. (1973). On consistency in monotonic regression. *Ann. Statist.* **1** 401–421.

[3] KIEFER, J. (1961). On large deviations of the empiric df of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649–660.

[4] LOÈVE, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, Princeton.

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

[6] MAKOWSKI, G. G. (1973 b). Convergence rates for multidimensional integral regression function estimators. Unpublished manuscript.

[7] PYKE, R. (1970). Discussion on Brunk's paper. *Nonparametric Techniques in Statistical Inference*, 196–197. Cambridge Univ. Press.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712