The Annals of Statistics
1981, Vol. 9, No. 2, 449–452

SUMS OF RANDOM VARIABLES INDEXED BY A PARTIALLY ORDERED SET AND THE ESTIMATION OF INTEGRAL REGRESSION FUNCTIONS[1]

BY F. T. WRIGHT

University of Missouri-Rolla

The estimator proposed by Brunk for the indefinite integral of a regression function defined on the unit cube in β dimensional Euclidean space is studied. It is shown to be strongly uniformly consistent if the errors satisfy a first moment type of condition and an almost sure rate of convergence of order $O((n/\log_2 n)^{-1/2})$ is obtained.

1. Introduction. Let $\{t_k\}$ be a sequence of points in Δ, a subset of R_β (the β-dimensional reals), and let $\{X_k\}$ be a sequence of independent random variables which are centered at their means. We think of X_k as associated with t_k for $k = 1, 2, \ldots$. For $A \subset R_\beta$, define $S_n(A) = \sum_{\{k\le n: t_k \in A\}} X_k$ with $S_n(A) \equiv 0$ if $A \cap \{t_1, t_2, \ldots, t_n\} = \varnothing$. Let $\ll$ denote the usual coordinate-wise partial order on R_β, let $\mathscr{W}$ denote the collection of subsets of the form $\{s: s \ll t\}$ with $t \in R_\beta$, and let $S_n(\mathscr{W}) = \max_{W\in\mathscr{W}} S_n(W)$.

We are interested in the almost sure convergence of $S_n(\mathscr{W})/n$ to zero, as well as rates of this convergence, because such results provide information concerning the strong consistency of an estimator proposed by Brunk (1970) for integral regression functions. For $\beta = 1$, Brunk (1970) has proved the almost sure uniform consistency of this estimator if the errors satisfy the r-order Kolmogorov condition. Lemma 2 of Hanson, Pledger and Wright (1973) combined with Brunk's work, shows that the r-order Kolmogorov condition can be replaced by a first moment type of assumption. (See the hypothesis of Theorem 1 of this paper.) Makowski (1976) demonstrated the strong consistency of the β-dimensional analogue of Brunk's estimator assuming the r-order Kolmogorov condition, and in Theorem 3 we show that the first moment condition mentioned above is also sufficient in the multivariate case. Makowski (1973, 1976) has also obtained an almost sure rate of convergence of order $O(n/\log_2 n)^{-1/(2\beta+2)})$, where $\log_2 x = \log\log x$. It is shown here that the exponent can be decreased to $-\frac{1}{2}$ and that the moment assumption can be weakened. (In the case $\beta > 1$, he has assumed that the errors have a moment generating function; but a second moment type of condition is sufficient.) The weak consistency properties of the estimator have been studied by Pledger (1976). He considered arbitrary β and triangular arrays of observation points.

The strong law needed to prove that the estimator is consistent was established by Smythe (1978). It is interesting to note that this result has no restrictions on the sequence $\{t_k\}$. In Wright (1978), a similar strong law was established for $S_n(\mathscr{L}) \equiv \max_{L\in\mathscr{L}} S_n(L)$ where $\mathscr{L}$ is the collection of lower layers. (A set $L \subset R_\beta$ is called a lower layer provided $t \in L$ and $s \ll t$ imply that $s \in L$.) However, because of the less restrictive nature of the elements of $\mathscr{L}$, assumptions on the placement of the t_k are needed to show that $S_n(\mathscr{L})/n$ converges to zero almost surely. (Smythe (1978) also considers $S_n(\mathscr{L})/n$ and gives a very elegant proof of some of the results in Wright (1978).) In this note we prove a law of the

Received June, 1978; revised February, 1980.

[1] This research was sponsored by the National Science Foundation under Grant Number MCS 78 01514.

AMS 1970 *subject classifications.* Primary 60F15; secondary 62E20.

Key words and phrases. Integral regression functions, laws of the iterated logarithm, partially ordered index sets.

iterated logarithm type of result for $S_n(\mathscr{W})$ and consider its implications in the estimation of integral regression functions.

2. Asymptotic behavior of $S_n(\mathscr{W})$. We suppose that the sequence $\{X_k\}$ is dominated in the following sense: there is a probability distribution function $G(x)$ for which

$$P[|X_k| \geq y] \leq \int_{|x|\geq y} dG(x) \qquad \text{for } k = 1, 2, \ldots \text{ and } y \geq 0. \tag{1}$$

The following result is a restatement of Theorem 2.1 of Smythe (1978).

THEOREM 1. *If* $\int_{-\infty}^{\infty} |x| dG(x) < \infty$, *then* $S_n(\mathscr{W})/n \to 0$ a.s.

We now prove a law of the iterated logarithm type of result for $S_n(\mathscr{W})$. Let σ_k^2 denote the variance of X_k and set $s_n^2 = \sigma_1^2 + \ldots + \sigma_n^2$ for $n = 1, 2, \ldots$.

THEOREM 2. *If* $\int_{-\infty}^{\infty} x^2 dG(x) < \infty$ *and if* s_n^2/n *is bounded away from zero, then*

$$P[\limsup_{n\to\infty} S_n(\mathscr{W})/(2s_n^2 \log_2 s_n^2)^{1/2} = 1] = 1$$

The following result is used in the proof of Theorem 2.

LEMMA. *Let j, k and m denote β-dimensional vectors of positive integers, let* $\{X_k: k \ll m\}$ *be a collection of independent random variables, let*

$$S_k = \textstyle\sum_{j\ll k} X_j$$

and let

$$T_m = \max_{k\ll m} S_k$$

If a median of a random variable Y is denoted med (Y), *then for any positive constant C*

$$P[T_m - \text{med}(T_m) \geq C] \leq 2^{\beta+1} P[S_m^* \geq C],$$

where S_k^* *is the symmetrized version of* S_k *for* $k \ll m$.

PROOF. If T_m' denotes $\max_{k\ll m} S_k'$ where the S_k' are independent replicas of the S_k then

$$P[\max_{k\ll m} S_k^* \geq C] \geq P[T_m - \text{med}(T_m) \geq C] P[T_m' - \text{med}(T_m') \leq 0]. \tag{2}$$

Paranjape and Park (1973) have shown that the first expression in (2) is bounded above by $2^\beta P[S_m^* \geq C]$ and so the desired result follows.

PROOF (THEOREM 2). Since $S_n(\mathscr{W}) \geq X_1 + \ldots + X_n$, the Hartman-Wintner law of the iterated logarithm shows that $\limsup_{n\to\infty} S_n(\mathscr{W})/(2s_n^2\log_2 s_n^2)^{1/2} \geq 1$ a.s. and so we only need to establish the reverse inequality. Hartman and Wintner (1941) have shown that there are sequences of random variables $\{Z_n\}$ and $\{Y_n\}$ and a sequence of real numbers $\{\alpha_n\}$ with $X_n = Z_n + Y_n + \alpha_n$ for all n; $\sum_{k=1}^n (|Y_k| + |\alpha_k|) = o((n \log_2 n)^{1/2})$a.s.; $\{Z_n\}$ is a sequence of independent random variables which are centered at their means and satisfy Kolmogorov's condition, that is ess sup $|Z_n| = o(d_n^2/\log_2 d_n^2)^{1/2})$ with $d_n^2 = \sum_{k=1}^n \text{Var}(Z_k)$ and $d_n^2 \log_2 d_n^2/(s_n^2 \log_2 s_n^2) \to 1$ as $n \to \infty$. Since s_n^2/n is bounded away from zero $S_n(\mathscr{W}) \leq S_n^{(Z)}(\mathscr{W}) + o((s_n^2\log_2 s_n^2)^{1/2})$, where $S_n^{(Z)}(\mathscr{W}) = \max_{W\in\mathscr{W}} S_n^{(Z)}(W)$ and $S_n^{(Z)}(W) = \sum_{\{k\leq n: t_k \in W\}} Z_k$. The proof now proceeds as the proof of Kolmogorov's law of the iterated logarithm. (See, for instance, Loève (1963).) So we must show that for any $\delta > 0$, $\sum_k P[S_{n_k}' > (1 + \delta)(2d_{n_k}^2 \log_2 d_{n_k}^2)^{1/2}] < \infty$ for a properly chosen subsequence n_k, with $S_{n_k}' =$

$\max_{n \le n_k} S_n^{(Z)}(\mathscr{W})$. For k fixed, a minimal β-dimensional grid can be chosen so that the points $t_1, t_2, \ldots, t_{n_k}$ are vertices of the grid. Label the vertices of the grid $(j_1, j_2, \ldots, j_\beta)$ for $j_i = 1, 2, \ldots, m_i$ where $m_i \le n_k$ for $i = 1, 2, \ldots, \beta$ so that the vertices have the same ordering with respect to $\ll$ as the associated labels, $(j_1, j_2, \ldots, j_\beta)$. Construct n_k replicas of this grid and label the vertices in the jth copy $(j_1, j_2, \ldots, j_\beta, j)$ for $j = 1, 2, \ldots, n_k$. So we have constructed a $(\beta + 1)$-dimensional grid. Define $X_{j_1 j_2 \cdots j_{\beta+1}} = Z_j$ if t_j is the vertex labeled $(j_1, j_2, \ldots, j_\beta)$ and $j_{\beta+1} = j$ and define $X_{j_1 j_2 \cdots j_{\beta+1}} \equiv 0$ otherwise. Now with $T_{m_1 m_2 \ldots m_{\beta+1}}$ defined as in the lemma and $m_{\beta+1} = n_k$,

$$S'_{n_k} = T_{m_1 m_2 \ldots m_{\beta+1}}.$$

Noting that $|\operatorname{med}(X)| \le |E(X)| + \sqrt{2V(X)} \le (1 + \sqrt{2})(E(X^2))^{1/2}$ and applying Theorem 6 of Gabriel (1977) to obtain $E(T^2_{m_1 m_2 \ldots m_{\beta+1}}) \le DE[(S_{n_k}^{(Z)}(R_\beta))^2]$ with D a positive constant, we see that $\operatorname{med}(T_{m_1 m_2 \ldots m_{\beta+1}}) = o((d^2_{n_k} \log_2 d^2_{n_k})^{1/2})$. So it suffices to consider

$$P[T_{m_1 m_2 \ldots m_{\beta+1}} - \operatorname{med}(T_{m_1 m_2 \ldots m_{\beta+1}}) \ge (1 + \delta')(2d^2_{n_k} \log_2 d^2_{n_k})^{1/2}] \tag{3}$$

with $\delta' > 0$. Next we apply the lemma to show that (3) is bounded above by $2^{\beta+1} P[\sum_{j=1}^{n_k} Z_j^* \ge (1 + \delta')(2d^2_{n_k} \log_2 d^2_{n_k})^{1/2}]$ where Z_j^* is the symmetrized version of Z_j. The remainder of the proof is like the proof of Kolmogorov's law of the iterated logarithm since ess sup $|Z_n^*| = o((d_n^2/\log_2 d_n^2)^{1/2})$.

Theorem 2 extends the first two results in Makowski (1973) to the case $\beta > 1$ and is "sharper" in that the exact value of the limit superior has been determined. Due to the modified inequality by Paranjape and Park the proof is also less complicated.

3. Integral regression functions. Let Δ be the closed unit cube in R_β, let $\mu(\cdot)$ be a real valued Borel-measurable function defined on Δ, let $F(\cdot)$ be a probability distribution function with support in Δ, and for $a, b, \in R_\beta$ with $a \ll b$ let $[a, b]$ denote $\{t \in R_\beta : a \ll t \ll b\}$. For $t \in \Delta$, the integral regression function has been defined by $M(t) = \int_{[0,t]} \mu(s)\, dF(s)$. Let $\{t_k\}$ be a sequence of observation points in Δ, let $\{Y_k\}$ be a sequence of independent random variables with $E(Y_k) = \mu(t_k)$ and let $F_n(\cdot)$ denote the empirical distribution function of $t_1, t_2, \ldots, t_n$. Brunk (1970) proposed $M_n(t) = \sum_{\{k \le n : t_k \ll t\}} Y_k/n$ as an estimator of $M(t)$. If the X_k of Section 2 are set equal to $Y_k - \mu(t_k)$, the theorems established there provide results concerning the consistency of $M_n(t)$. So we suppose that $G(\cdot)$ is a probability distribution function which dominates the sequence $\{Y_k - \mu(t_k)\}$ as in (1). As before $s_n^2 = \sum_{k=1}^n \operatorname{Var}(X_k)$.

Let $\Delta_{b,a}^{(i)} \mu = \mu(t_1, \ldots, t_{i-1}, b, t_{i+1}, \ldots, t_\beta) - \mu(t_1, \ldots, t_{i-1}, a, t_{i+1}, \ldots, t_\beta)$ be the ith coordinate difference operator and for $j = 1, 2, \ldots, \beta$ let $0 = u_{j,0} < \ldots < u_{j,n_j} = 1$ be a partition of $[0, 1]$. If for a fixed constant C

$$\sum_{i_1=0}^{n_1-1} \sum_{i_2=0}^{n_2-1} \cdots \sum_{i_\beta=0}^{n_\beta-1} \left| \Delta^{(1)}_{u_{1,i_1}, u_{1,i_1+1}} \cdots \Delta^{(\beta)}_{u_{\beta,i_\beta}, u_{\beta,i_\beta+1}} \mu(t_1, \ldots, t_\beta) \right| \le C \tag{4}$$

for all such partitions of $[0, 1]^\beta$ and if the k dimensional analogue $(1 \le k < \beta)$ of (4) holds for every function derived from μ by fixing $\beta - k$ of the variables $t_1, \ldots, t_\beta$, then μ is said to be of strictly bounded variation. Hobson (1927, Sections 253 and 254) discusses two definitions of bounded variation for functions of two real variables. The definition given here is an obvious generalization of the stronger of the two.

THEOREM 3. *If $\mu(\cdot)$ is continuous, $\int_{-\infty}^{\infty} |x|\, dG(x) < \infty$, and F_n converges uniformly to F, then* $\sup_{t \in \Delta} |M_n(t) - M(t)| \to 0$ a.s.

Furthermore, if $\int_{-\infty}^{\infty} x^2 dG(x) < \infty$, s_n^2/n is bounded away from zero, μ is of strictly bounded variation on $[0, 1]$, *and* $\limsup_{n\to\infty} (n/\log_2 n)^{1/2} \sup_{t \in \Delta} |F_n(t) - F(t)| \le M$, *for some real number M, then there is a real number M^* for which*

$$\limsup_{n\to\infty} (n/\log_2 n)^{1/2} \sup_{t \in \Delta} |M_n(t) - M(t)| \le M^* \text{ a.s.}$$

PROOF. Clearly, $\sup_{t\in\Delta} |M_n(t) - M(t)|$ is bounded above by

$$\sup_{t\in\Delta} \left| M_n(t) - \int_{[0,t]} \mu(s)dF_n(s) \right| + \sup_{t\in\Delta} \left| \int_{[0,t]} \mu(s)dF_n(s) - \int_{[0,t]} \mu(s)dF(s) \right| \tag{5}$$

and the first expression in (5) is $\max(S_n(\mathscr{W}), S_n^{(-)}(\mathscr{W}))/n$, where $S_n^{(-)}(\mathscr{W})$ is defined like $S_n(\mathscr{W})$ except X_k is replaced by $-X_k$. Hence, Theorems 1 and 2 show that it behaves as specified.

The proof is completed by showing that the second term converges to zero at the proper rate. While we only give the proof for $\beta = 2$, the proof for $\beta \neq 2$ is analogous. With $t = (t_1, t_2)$ fixed we consider a partition of $[0, t]$ determined by $0 = u_0 < u_1 < \ldots < u_k = t_1$ and $0 = v_0 < v_1 < \ldots < v_k = t_2$. The Lebesgue-Stieltjes integral $\int_{[0,t]} \mu(s)dF(s)$ can be shown to be the limit, as the norm of the partition converges to zero, of

$$\sum_{i=0}^{k-1} \sum_{j=0}^{k-1} F(u_i, v_j)[\mu(u_{i+1}, v_{j+1}) + \mu(u_i, v_j) - \mu(u_{i+1}, v_j) - \mu(u_i, v_{j+1})] + \sum_{i=0}^{k-1} F(u_i, t_2)[\mu(u_i, t_2) - \mu(u_{i+1}, t_2)] + \sum_{j=0}^{k-1} F(t_1, v_j)[\mu(t_1, v_j) - \mu(t_1, v_{j+1})] + F(t_1, t_2)\mu(t_1, t_2).$$

Since a similar statement holds for $\int_{[0,t]} \mu(s)dF_n(s)$, by using the conditions on $\mu(\cdot)$ and $|F_n(t) - F(t)|$ it can be shown that

$$(n/\log_2 n)^{1/2} \sup_{t\in\Delta} |\int_{[0,t]} \mu(s)dF(s) - \int_{[0,t]} \mu(s)dF_n(s)| < \infty,$$

and the proof is completed.

Using the techniques presented in Brunk (1970), one could obtain consistency results in the independent observations regression model by applying Theorem 3. We refer the interested reader to Brunk (1970), Makowski (1973, 1976) and Pledger (1976).

REFERENCES

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

[2] 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.

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

[4] HARTMAN, PHILIP AND WINTNER, AUREL (1941). On the law of the iterated logarithm. *Amer. J. Math* **63** 169–176.

[5] HOBSON, E. W. (1972). *The Theory of Functions of a Real Variable and The Theory of Fourier's Series*, Vol. I, 3rd Ed. Cambridge Univ.

[6] PARANJAPE, S. R. AND PARK, C. (1973). Laws of the iterated logarithm of multiparameter Wiener processes. *J. Multivariate Anal.* **3** 132–136.

[7] PLEDGER, GORDON (1976). Consistency in integral regression estimation with a triangular array of observation points. *Ann. Statist.* **4** 234–236.

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

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

[10] MAKOWSKI, G. G. (1976). Convergence rates for multidimensional integral regression function estimators. *Sankhyā*, Series A **38** 304–307.

[11] SMYTHE, R. T. (1978). Maxima of partial sums and a monotone regression estimator. *Ann. Probability* **8** 630–635.

[12] WRIGHT, F. T. (1979). A strong law for variables indexed by a partially ordered set with applications to isotone regression. *Ann. Probability* **7** 109–127.

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