The Annals of Statistics
1973, Vol. 1, No. 3, 401–421

ON CONSISTENCY IN MONOTONIC REGRESSION[1]

BY D. L. HANSON, GORDON PLEDGER[2] AND F. T. WRIGHT

University of Missouri, Columbia, Florida State University and University of Iowa

For each t in some subset T of N-dimensional Euclidean space let F_t be a distribution function with mean $m(t)$. Suppose $m(t)$ is non-decreasing in each of the coordinates of t. Let $t_1, t_2, \cdots$ be a sequence of points in T and let $Y_1, Y_2, \cdots$ be an independent sequence of random variables such that the distribution function of Y_k is F_{t_k}. Estimators $\hat{m}_n(t; Y_1, \cdots, Y_n)$ of $m(t)$ which are monotone in each coordinate of t and which minimize $\sum_{i=1}^n [\hat{m}_n(t_i; Y_1, \cdots, Y_n) - Y_i]^2$ are already known. Brunk has investigated their consistency when $N = 1$.

In this paper additional consistency results are obtained when $N = 1$ and some results are obtained in the case $N = 2$. In addition, we prove several lemmas about the law of large numbers which we believe to be of independent interest.

1. Summary. Let N be a positive integer; let $t = (t^1, \cdots, t^N)$ and $s = (s^1, \cdots, s^N)$ be "variable" ordered N-tuples of real numbers; let $t_k = (t_k^1, \cdots, t_k^N)$ for $k = 1, 2, \cdots$ be a fixed sequence of (not necessarily distinct) ordered N-tuples of real numbers; and let $m(t)$ be a real valued function of t which is non-decreasing in t^i for $i = 1, \cdots, N$.

For each t let F_t be a distribution function such that $\int x\, dF_t(x) = m(t)$. Let Y_k for $k = 1, 2, \cdots$ be an independent sequence of random variables and let Y_k have distribution function F_{t_k}. When y_k is used, it will be a real number and is to be thought of as an observed value of Y_k.

Let $\mathscr{L}_N$ be the collection of subsets L of R_N having the property:

$$(1)\qquad \left.\begin{cases} t \text{ in } L \\ t^i \leqq s^i \text{ for } i = 1, \cdots, N \end{cases}\right\} \text{ imply } s \text{ in } L.$$

Let $\mathscr{L}_N{}^c$ be the collection of complements (in R_N) of members of $\mathscr{L}_N$. Define

$$(2)\qquad \hat{m}_n(t_k; y_1, \cdots, y_n) = \max_{\{L: t_k \in L \in \mathscr{L}_N\}} \min_{\{K: t_k \in K \in \mathscr{L}_N{}^c\}} \frac{\sum_{\{\alpha: t_\alpha \in L \cap K, \alpha \leqq n\}} y_\alpha}{\#\{\alpha : t_\alpha \in L \cap K, \alpha \leqq n\}}$$

Received November 8, 1971; revised June 8, 1972.

[1] Research partially sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under AFOSR Grant No. 68–1394 and AFOSR Grant No. 71–2031. The United States Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation hereon.

[2] Part of this author's research was done while he held a National Science Foundation Traineeship at the University of Missouri-Columbia.

AMS 1970 *subject classifications.* Primary 62G05; Secondary 60G50.

Key words and phrases. Monotonic regression, isotonic regression, consistency, mean, law of large numbers.

where $\#(A)$ is the number of elements in the set A. (Most of the time we will write $\hat{m}_n(t) = \hat{m}_n(t; y_1, \cdots, y_n)$.) The functions $\hat{m}_n$ can be extended to all of R_N so that

(3) if $t^i \leqq s^i$ for $i = 1, \cdots, N$ then $\hat{m}_n(t) \leqq \hat{m}_n(s)$.

This extension is not unique.

The purpose of this paper is to present some results about the consistency of the sequence $\{\hat{m}_n(t; Y_1, \cdots, Y_n)\}$ of estimators of m in the cases $N = 1$ and $N = 2$.

Section 2 contains some more preliminaries. Sections 3 and 4 contain, respectively, our results for the cases $N = 1$ and $N = 2$; in each case an example and our theorems are presented first, followed by our proofs. Section 5 contains some remarks about some theoretical questions which have come up in the course of this investigation.

The first half of Theorem 1 ($N = 1$, Section 3) is a variant of Theorem 4.1 of Brunk [6] giving weak consistency instead of strong consistency but with weaker moment conditions and without Brunk's assumption about the asymptotic distribution in [0, 1] of the points at which observations are taken. Our Theorem 2 shows that if Brunk's assumption is just omitted then there exist cases in which one does not have strong consistency. The second half of our Theorem 1 gives Brunk's Theorem 4.1 under considerably weakened moment conditions. Our proofs are different from those given previously and depend heavily on some lemmas (which we believe to be of independent interest) about the law of large numbers. In particular, Lemma 3 shows that the strong law holds uniformly for all independent (not necessarily identically distributed) sequences of random variables satisfying certain moment conditions.

We believe our results in Section 4 ($N = 2$) to be the first explicitly dealing with a higher dimensional situation.

2. Preliminaries. The estimates $\hat{m}_n(t_i; y_1, \cdots, y_n)$ are known and are known to minimize

$$\sum_{k=1}^{n} (m_n(t_k) - y_k)^2 \tag{4}$$

when the minimum is taken over all functions $m_n(t)$ which are non-decreasing in each of the coordinates of t. To see the latter we let $s_1, \cdots, s_a$ be the distinct members of $(t_1, \cdots, t_n)$; let n_i for $i = 1, \cdots, a$ be the number of times s_i appears in $(t_1, \cdots, t_n)$; and let

$$\bar{y}_i = \sum_{\{\alpha: t_\alpha = s_i, 1 \leqq \alpha \leqq n\}} y_\alpha / n_i \,. \tag{5}$$

Let μ be the measure on the Lebesgue subsets of R_N which assigns measure n_i to $\{s_i\}$ for $i = 1, \cdots, a$ and which assigns measure zero to $R_N - \{s_1, \cdots, s_a\}$. Then (4) may be put in the form

$$\sum_{i=1}^{a} n_i(m_n(s_i) - \bar{y}_i)^2 + C \tag{6}$$

where C depends on $y_1, \cdots, y_n$ but not on m_n. Minimizing (4) then amounts to minimizing the sum in (6) which can be put in the form

$$(7) \qquad \int_{R_N} (m_n(s) - \bar{y}(s))^2 \, d\mu(s)$$

where

$$(8) \qquad \begin{aligned} \bar{y}(s) &= \bar{y}_i \quad \text{if } s = s_i \qquad i = 1, \cdots, a \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

We use the notation of Brunk, Ewing, and Utz [7] and set $\alpha = \bar{y}$, $m_n = \theta$, and $T(u) = u^2$ so that $F(u, v) = T(u) - T(v) - (u - v)T'(v) = (u - v)^2$. Then (7) is of the form $\int F[\alpha(s), \theta(s)] \, d\mu(s)$. By Theorem 3.3 (of [7]) a minimizing function $\hat{m}_n$ exists. By (4.15) and (4.16) of Theorem 4.3 it is unique at the points $s_1, \cdots, s_a$ and can be written either as

$$(9) \qquad \hat{m}_n(s_i) = \sup_{s_i \in L \in \mathscr{L}_N} \inf_{s_i \in K \in \mathscr{L}_N^c} \int_{L \cap K} \bar{y}(s) \, d\mu(s) / \mu(L \cap K)$$

or as

$$(10) \qquad \hat{m}_n(s_i) = \inf_{s_i \in K \in \mathscr{L}_N^c} \sup_{s_i \in L \in \mathscr{L}_N} \int_{L \cap K} \bar{y}(s) \, d\mu(s) / \mu(L \cap K) \, .$$

With a little manipulation (9) can be put in the form (2) and (10) can be put in the form

$$(11) \qquad \begin{aligned} &\hat{m}_n(t_i; y_1, \cdots, Y_n) \\ &\quad = \min_{\{K : t_i \in K \in \mathscr{L}_N^c\}} \max_{\{L : t_i \in L \in \mathscr{L}_N\}} \frac{\sum_{\{\alpha : t_\alpha \in L \cap K, \alpha \leqq n\}} y_\alpha}{\#\{\alpha : t_\alpha \in L \cap K, \alpha \leqq n\}} \, . \end{aligned}$$

The estimates being investigated are related to conditional expectation given a σ-lattice and can undoubtedly be obtained from the currently existing theory there or from slight modifications of it. The results of [7] seem to give the minimizing property of these estimates in a more straightforward manner.

If $N = 1$ and we order $s_1, \cdots, s_a$ so that $s_1 < s_2 < \cdots < s_a$ then an "obvious" extension of $\hat{m}_n$ to all of R_1 requires that it be continuous; be linear between s_i and s_{i+1} for $i = 1, \cdots, a - 1$; and be constant to the left of s_1 and to the right of s_a.

For all N define $t < s$ ($t \leqq s$) if $t^i < s^i$ ($t^i \leqq s^i$) for $i = 1, \cdots, N$. Define

$$(12) \qquad \begin{aligned} \underline{m_n}(t) &= \max \{\hat{m}_n(t_i) : t_i \leqq t, 1 \leqq i \leqq n\} \quad \text{if } t_i \leqq t \text{ for some } 1 \leqq i \leqq n \\ &= \min \{\hat{m}_n(t_i) : 1 \leqq i \leqq n\} \quad \text{otherwise} \end{aligned}$$

and

$$(13) \qquad \begin{aligned} \overline{m_n}(t) &= \min \{\hat{m}_n(t_i) : t \leqq t_i, 1 \leqq i \leqq n\} \quad \text{if } t \leqq t_i \text{ for some } 1 \leqq i \leqq n \\ &= \max \{\hat{m}_n(t_i) : 1 \leqq i \leqq n\} \quad \text{otherwise.} \end{aligned}$$

These are, respectively, the minimal and maximal possible definitions for $\hat{m}_n$ which are restricted to values in the interval $[\min \{\hat{m}_n(t_i) : 1 \leqq i \leqq n\}, \max \{\hat{m}_n(t_i) : 1 \leqq i \leqq n\}]$. They are not continuous. There are many continuous functions possible for $\hat{m}_n$ between $\underline{m_n}$ and $\overline{m_n}$ but the authors know of no obvious or "natural" choice except (as already given) in the case $n = 1$.

3. The case $N = 1$. Throughout this section $\hat{m}_n(t)$ will be the estimator $\hat{m}_n(t; Y_1, \cdots, Y_n)$ defined as in Section 2 so as to be continuous and piecewise linear for fixed values of $Y_1, \cdots Y_n$. We will restrict t to the interval $[0, 1]$.

One example of a situation in which one might wish to estimate the entire function $m(t)$ is as follows: t is to be thought of as dosage or concentration of a given drug and $m(t)$ is to be thought of as the expected proportion of the organisms of some particular type killed (under a fixed set of conditions) by that concentration of the drug. The experimenter decides in advance to run an experiment using dosages $t_1, \cdots, t_n$ and wants to estimate $m(t)$ in some interval based on the outcome of his experiment.

For $y \geqq 0$ define

$$F(y) = \sup_{k=1,2,\cdots} P\{|Y_k - m(t_k)| \geqq y\} \tag{14}$$

and for any set $J \in R$ define

$$N_n(J) = \#\{i\colon 1 \leqq i \leqq n \text{ and } t_i \in J\}\ . \tag{15}$$

THEOREM 1. *If m is continuous on* $[0, 1]$, $0 < a < b < 1$, $\{t_i\}$ *is dense in* $[0, 1]$, $F(y) \to 0$ *as* $y \to \infty$, *and* $\int_0^\infty y|dF(y)| < \infty$, *then for every* $\varepsilon > 0$

$$P\{\sup_{0 \leqq t \leqq b} [\hat{m}_n(t) - m(t)] < \varepsilon,\ \sup_{a \leqq t \leqq 1} [m(t) - \hat{m}_n(t)] < \varepsilon\} \to 1 \qquad \text{as} \quad n \to \infty\ .$$

If in addition $\limsup n/N_n(I)$ *is finite for every non-degenerate interval I contained in* $[0, 1]$ *then*

$$P\{\lim_{n\to\infty} \sup_{a \leqq t \leqq b} |\hat{m}_n(t) - m(t)| = 0\} = 1\ .$$

Theorem 1 is a modification of Theorem 4.1 of Brunk [6]. The following theorem shows that the condition "$\limsup n/N_n(I)$ is finite$\cdots$" cannot just be omitted in Theorem 1 if we are to have the conclusion of the second part of that theorem.

THEOREM 2. *Let m be continuous on* $[0, 1]$, *let* $0 < a < b < 1$, *and suppose there exists* $\delta > 0$ *such that* $\sup_t F_t[m(t) + \delta] < 1 - \delta$ *and* $\inf F_t[m(t) - \delta] > \delta$. *Then there exists a sequence* $\{t_i\}$ *dense in* $[0, 1]$ *and an* $\varepsilon > 0$ *such that*

$$P\{\inf_{a \leqq t \leqq 1} \limsup_n [\hat{m}_n(t) - m(t)] \geqq \varepsilon,\ \sup_{0 \leqq t \leqq b} \liminf_n [\hat{m}_n(t) - m(t)] \leqq -\varepsilon\} = 1\ .$$

The following two theorems give "rates of consistency."

THEOREM 3. *If m is continuous on* $[0, 1]$, $\varepsilon > 0$ *and* $0 < a < b < 1$, $\limsup n/N_n(I) < \infty$ *for every interval* $I \subset [0, 1]$, *and* $\lim_{y\to\infty} y^r F(y) = 0$ *with* $r > 1$, *then as* $n \to \infty$

$$n^{r-1} P\{\sup_{0 \leqq t \leqq b} [\hat{m}_n(t) - m(t)] \geqq \varepsilon \ \text{ or } \ \sup_{a \leqq t \leqq 1} [m(t) - \hat{m}_n(t)] \geqq \varepsilon\} \to 0\ .$$

THEOREM 4. *If m is continuous on* $[0, 1]$, $\limsup n/N_n(I) < \infty$ *for every interval* $I \subset [0, 1]$, $F(y) \to 0$ *as* $y \to \infty$, *there exisis* $\tau > 0$ *such that* $\int_0^\infty e^{\tau y}|dF(y)| < \infty$,

$\varepsilon > 0$, and $0 < a < b < 1$, then there exist a real number ρ in $(0, 1)$ and a real number C such that for all $n \geqq 1$

$$P\{\sup_{0 \leqq t \leqq b} [\hat{m}_n(t) - m(t)] \geqq \varepsilon \ \textit{or} \ \sup_{a \leqq t \leqq 1} [m(t) - \hat{m}_n(t)] \geqq \varepsilon\} \leqq C\rho^n .$$

PROOF INDICATION FOR THEOREM 1. We begin with three lemmas.

LEMMA 1. *Let $X_1, X_2, \cdots$ be independent random variables with $EX_k = 0$ and $EX_k^2 < \infty$ for all k. Suppose*

$$\sum_{k=1}^{\infty} EX_k^2/k^2 < \infty . \tag{16}$$

Corresponding to each positive integer $n \geqq 2$ let $i_{1,n}, \cdots, i_{n,n}$ be a permutation of the positive integers $1, 2, \cdots, n$ obtained by assigning a place to the integer n between some two successive integers of, or at the beginning of, or at the end of, the permutation corresponding to the integer $n - 1$. For $j = 1, \cdots, n$ define

$$S_{j,n} = \sum_{k=1}^{j} X_{i_{k,n}} .$$

Then

$$P\{\lim_{n\to\infty} \max_{j=1,\cdots,n} n^{-1}|S_{j,n}| = 0\} = 1 .$$

PROOF. This is the special case $r = 1$ of Theorem 6.1 of Brunk [3]. There are several typographical errors in the proof of Brunk's Theorem 6.1 and the hypothesis $r \geqq 1$ should be added to it.

LEMMA 2. *The preceding lemma is true if the hypothesis $EX_k^2 < \infty$ for all k and hypothesis (16) are replaced by the hypotheses*

$$F(y) \to 0 \quad \textit{as} \quad y \to \infty \quad \textit{and} \quad \int_0^\infty y|dF(y)| < \infty \tag{17}$$

where

$$F(y) = \sup_k P\{|X_k| \geqq y\} . \tag{18}$$

PROOF. The proof follows from Lemma 1 in the same way that Kolmogorov's Strong Law of Large Numbers ([11] page 239) follows from Theorem A ([11] page 238).

LEMMA 3. *Let F be a real valued function on $[0, \infty)$ satisfying (17). Then for every $\varepsilon > 0$ there exists a positive integer M such that if $\{X_i : i = 1, 2, \cdots\}$ is an independent sequence of random variables such that $EX_i = 0$ and $P\{|X_i| \geqq y\} \leqq F(y)$ for all i and all $y \geqq 0$, then*

$$P\{\sup_{M \leqq n} n^{-1}|X_1 + \cdots + X_n| > \varepsilon\} \leqq \varepsilon , \tag{19}$$

and

$$P\{\max_{1 \leqq k \leqq n} |S_k|/n > \varepsilon\} < \varepsilon \qquad \textit{for all} \ n \geqq M . \tag{20}$$

PROOF. Assume there exists no M such that (19) is true. Then there exist an $\varepsilon > 0$ and sequences $\{X_{k,M} : k = 1, 2, \cdots\}$ of independent random variables for $M = 1, 2, \cdots$ such that

$$P\{\sup_{M \leqq n} n^{-1}|X_{1,M} + \cdots + X_{n,M}| > \varepsilon\} > \varepsilon$$

for all M. We will assume that all the random variables $\{X_{i,j}: i, j = 1, 2, \cdots\}$ are defined on the same probability space and are independent; we can always define a product space on which are defined replicas of the original random variables having the desired properties. For each positive integer i let ν_i be a positive integer such that

$$P\{\max_{i \leq n \leq \nu_i} n^{-1}|X_{1,i} + \cdots + X_{n,i}| > \varepsilon\} > \varepsilon .$$

We now construct a sequence $\{X_i: i = 1, 2, \cdots\}$ of independent random variables such that $EX_i = 0$ and $P\{|X_i| \geqq y\} \leqq F(y)$ for all i and all $y \geqq 0$. The sequence $\{X_i\}$ will thus obey the strong law of large numbers.

Let $i_0 = a_0 = 0$. Suppose $X_1, \cdots, X_{a_k}$ have been chosen. If $k \neq 0$ let T_k be such that

$$P\{a_k^{-1}|X_1 + \cdots + X_{a_k}| \leqq T_k\} > \tfrac{1}{2} .$$

Let i_{k+1} be such that $i_k < i_{k+1}$, $a_k \leqq i_{k+1}$, and (when $k \neq 0$) $4a_k T_k/\varepsilon \leqq i_{k+1}$. Define $a_{k+1} = a_k + \nu_{i_{k+1}}$ and $X_{a_k+j} = X_{j,i_{k+1}}$ for $j = 1, \cdots, \nu_{i_{k+1}}$. The sequence $\{X_i\}$ has the property that for $k \geqq 1$

$$\begin{aligned} P\{\sup_{a_k<n} n^{-1}|X_1 + \cdots + X_n| > \varepsilon/4\} &\geqq P\{\max_{a_k+i_{k+1} \leq n \leq a_{k+1}} n^{-1}|X_1 + \cdots + X_n| > \varepsilon/4\} \\ &\geqq P\{\max_{i_{k+1} \leq j \leq \nu_{i_{k+1}}} (a_k + j)^{-1}|X_{1,i_{k+1}} + \cdots + X_{j,i_{k+1}}| > \varepsilon/2\} \\ &\quad \times P\{(a_k + i_{k+1})^{-1}|X_1 + \cdots + X_{a_k}| \leqq \varepsilon/4\} \\ &\geqq \varepsilon\tfrac{1}{2} . \end{aligned}$$

Thus $\{X_i\}$ does not obey the strong law of large numbers, giving a contradiction. This completes the proof of the first part of the lemma.

Now suppose there exists no M such that (20) is true. Then there exist an $\varepsilon > 0$, independent random variables $\{X_{i,j}: i, j = 1, 2, \cdots\}$ all defined on the same probability space and satisfying the distributional assumptions of the lemma, and a sequence $\{n_M\}$ of positive integers with $M \leqq n_M$ such that

$$P\{\max_{1 \leq k \leq n_M} |X_{1,M}, \cdots X_{k,M}|/n_M > \varepsilon\} \geqq \varepsilon$$

for $M = 1, 2, \cdots$. Let $M_0 = a_0 = 0$. Suppose $X_1, \cdots, X_{a_k}$ have been chosen. Let $T_0 = 1$ and if $k \geqq 1$ let T_k be such that

$$P\{|X_1 + \cdots + X_{a_k}|/T_k < \varepsilon/3\} > \tfrac{1}{2} .$$

Let M_{k+1} be such that $M_{k+1} > \max\{M_k, a_k, T_k\}$. Then define $a_{k+1} = a_k + n_{M_{k+1}}$ and define $X_{a_k+j} = X_{j,M_{k+1}}$ for $j = 1, \cdots, n_{M_{k+1}}$. The sequence $\{X_j\}$ defined inductively in this manner satisfies the hypotheses of Lemma 2 so that $\max_{1 \leq k \leq n} |S_k|/n \to 0$ a.e. and in probability. However

$$\begin{aligned} P\{\max_{1 \leq j \leq a_{k+1}} |S_j|/a_{k+1} \geqq \varepsilon/6\} &\geqq P\{\max_{a_k<j \leq a_{k+1}} |S_j|/a_{k+1} \geqq \varepsilon/6\} \\ &\geqq P\{|X_1 + \cdots + X_{a_k}|/a_{k+1} < \varepsilon/3\} \\ &\quad \times P\{\max_{a_k<j \leq a_{k+1}} |S_j - S_{a_k}|/a_{k+1} \geqq \varepsilon/2\} . \end{aligned}$$

Note that $a_{k+1} > n_{M_{k+1}} \geqq M_{k+1} > \max\{M_k, a_k, T_k\}$ so that the above is bounded below by

$$P\{|X_1 + \cdots + X_{a_k}|/T_k < \varepsilon/3\} \times P\{\max_{1\leqq j\leqq n_{M_{k+1}}} |X_{1,M_{k+1}} + \cdots + X_{j,M_{k+1}}|/n_{M_{k+1}} \geqq \varepsilon\} \geqq \tfrac{1}{2}\varepsilon .$$

This contradicts $\max_{1\leqq j\leqq n} |S_j|/n \to 0$ in probability and completes the proof of the lemma.

We return to the proof of Theorem 1. Let s be some fixed observation point (i.e. some t_i) in (0, 1). Choose u in the interval $(s, 1)$ such that $m(u) - m(s) < \varepsilon/2$. For n large we have from (2)

$$\begin{aligned}\hat{m}_n(s) &= \max_{\alpha\leqq s} \min_{s\leqq\beta} \textstyle\sum_{\{i:\alpha\leqq t_i\leqq\beta, i\leqq n\}} Y_i/N_n([\alpha, \beta)] \\ &< \max_{\alpha\leqq s} \textstyle\sum_{\{i:\alpha\leqq t_i\leqq u, i\leqq n\}} Y_i/N_n([\alpha, u]) \\ &< m(s) + \varepsilon/2 + \max_{\alpha\leqq s} \textstyle\sum_{\{i:\alpha\leqq t_i\leqq u, i\leqq n\}} [Y_i - m(t_i)]/N_n([\alpha, u]) .\end{aligned}$$

The sequence $\{Y_i - m(t_i)\}$ satisfies the conditions of Lemma 3 so when n is large enough that $N_n([s, u]) \geqq M = M_\delta$ then

$$P\{\max_{\alpha\leqq s} |\textstyle\sum_{\{i:\alpha\leqq t_i\leqq u, i\leqq n\}} [Y_i - m(t_i)]/N_n([\alpha, u])| > \delta\} \leqq \delta .$$

Thus,

$$P\{\hat{m}_n(s) \leqq m(s) + \varepsilon/2 + \delta\} > 1 - \delta$$

for n large enough, and as $n \to \infty$

$$P\{\hat{m}_n(s) - m(s) < \varepsilon\} \to 1 .$$

We choose a finite number of observation points $s_1, \cdots, s_k$ such that $0 < s_1 < \cdots < s_k - 1 < b < s_k < 1$, $m(s_1) - m(0) < \varepsilon/2$, and $m(s_i) - m(s_{i-1}) < \varepsilon/2$ for $i = 2, \cdots, k$. Then if $\hat{m}_n(s_i) - m(s_i) < \varepsilon/2$ for $i = 1, \cdots, k$ we see that $\hat{m}_n(t) - m(t) < \varepsilon$ for all t in $[0, b]$. It follows that for every $\varepsilon > 0$, as $n \to \infty$

$$P\{\sup_{0\leqq t\leqq b} [\hat{m}_n(t) - m(t)] < \varepsilon\} \to 1 .$$

The proof of the other half of the first part of Theorem 1 is "symmetric" and omitted.

The proof of the second part of Theorem 1 is essentially just Brunk's proof of Theorem 4.1 of [6] using Lemma 2 instead of Lemma 1.

INDICATION OF PROOF OF THEOREM 2. Let $\{t_1, \cdots, t_{n_0}\}$ be n_0 distinct "observation points" in (0, 1). Let γ be a positive integer with $1 \leqq \gamma \leqq n_0$. We give four ways of adding additional observation points to this sequence.

(A) Let $t^* = \min\{t_1, \cdots t_{n_0}\}$. Define $t_{n_0+k} = t^*/2^k$ for $k = 1, 2, \cdots$. If observations Y_i were to be taken at t_i for $i = 1, 2, \cdots$ then $\hat{m}_n(0) \leqq Y_n$ for $n > n_0$. Since m is continuous, $t_n \to 0$, and $\inf_t F_t[m(t) - \delta] > \delta$, we see that $P\{\liminf \hat{m}_n(0) \leqq m(0) - \delta\} = 1$. In particular, there exists $M = M(\nu)$ such that $P\{\min_{n_0<n\leqq n_0+M} \hat{m}_n(0) < m(0) - \delta/2\} \geqq 1 - 2^{-\nu}$. We fix ν and let our "lengthened" sequence be $t_1, \cdots, t_{n_0+M}$.

(B) Let $t^* = \max\{t_1, \cdots, t_{n_0}\}$. Define $t_{n_0+k} = 1 - (1 - t^*)/2^k$ for $k = 1, 2, \cdots$. An argument similar to that used in (A) would provide an M such that $P\{\max_{n_0 < n \leqq n_0+M} \hat{m}_n(1) > m(1) + \delta/2\} \geqq 1 - 2^{-\nu}$. Our "lengthened" sequence would be $t_1, \cdots, t_{n_0+M}$.

(C) Let s be the maximum member of the set $\{0, t_1, \cdots, t_{n_0}\}$ which is strictly less than t_γ. Define

$$t_{n_0+k} = t_\gamma - (t_\gamma - s)/2^k \qquad \text{for } k = 1, 2, \cdots .$$

Suppose observations Y_i were taken at t_i for $i = 1, 2, \cdots$. Note that if $k \geqq k_0$ then

$$\hat{m}_{n_0+k}(t_\gamma) \geqq \text{mim}\,\{[\textstyle\sum_{j=0}^{k_0-1} Y_{n_0+k-j} - \sum_{j=1}^{n_0} |Y_j|]/k_0\,, \quad [\sum_{j=0}^{k_0-1} Y_{n_0+k-j} - \sum_{j=1}^{n_0} |Y_j|]/(n_0 + k_0)\}\,.$$

Also note that if

$$A_{k_0} = \limsup_n \textstyle\bigcap_{k=1}^{k_0} \{Y_{n+k} \geqq m(t_{n+k}) + \delta\}$$

then $P(A_{k_0}) = 1$ so that (remember that m is continuous) $\omega \in A_{k_0}$ implies

$$\limsup_n [\hat{m}_n(t_\gamma)](\omega) \geqq \min\left\{m(t_\gamma) + \delta - \frac{1}{k_0} \textstyle\sum_{j=1}^{n_0} |Y_j(\omega)|,\ [m(t_\gamma) + \delta]\dfrac{k_0}{n_0 + k_0} - \dfrac{1}{n_0 + k_0} \sum_{j=1}^{n_0} |Y_j(\omega)|\right\}.$$

If $\omega \in \bigcap_{k_0=1}^{\infty} A_{k_0}$ then $\limsup_n [\hat{m}_n(t_\gamma)](\omega) \geqq m(t_\gamma) + \delta$ so we can find $M = M(\nu, t_1, \cdots, t_N)$ such that

$$P\{\max_{n_0 < n < n_0+M} \hat{m}_n(t_\gamma) > m(t_\gamma) + \delta/2\} \geqq 1 - 2^{-\nu}\,.$$

We "lengthen" our observation sequence to $t_1, \cdots, t_{n_0+M}$.

(D) Let s be the minimum member of the set $\{1, t_1, \cdots, t_{n_0}\}$ which is strictly greater than t_γ. Define

$$t_{n_0+k} = t_\gamma + (s - t_\gamma)/2^k \qquad \text{for } k = 1, 2, \cdots .$$

An argument similar to that used in (C) would provide an M such that $P\{\min_{n_0 < n < n_0+M} \hat{m}_n(t_\gamma) < m(t_\gamma) - \delta/2\} \geqq 1 - 2^{-\nu}$. Our "lengthened" sequence would be $t_1, \cdots, t_{n_0+M}$.

Now let $\varepsilon = \delta/5$ and choose $0 < t_1 < \cdots < t_{n_1} < 1$ so that $t_1 \leqq a$; $b \leqq t_{n_1}$; $m(t_1) - m(0) < \varepsilon$; $m(t_i) - m(t_{i-1}) < \varepsilon$ for $i = 2, \cdots, n_1$; and $m(1) - m(t_{n_1}) < \varepsilon$. If we can construct $\{t_i\}$ so that

$$P\{\liminf_n \hat{m}_n(t) < m(t) - \delta/2\} = 1$$

for $t = 0, t_1, \cdots, t_{n_1}$ and so that

$$P\{\limsup_n \hat{m}_n(t) > m(t) + \delta/2\} = 1$$

for $t = t_1, \cdots, t_{n_1}, 1$ we will be done.

It is sufficient to be able to guarantee that for each $t = 0, t_1, \cdots, t_{n_1}$ we can get $N_{t,\nu}$ and $M_{t,\nu}$ for $\nu = 1, 2, \cdots$ so that $N_{t,\nu} < M_{t,\nu}$, $N_{t,\nu} \to \infty$ as $\nu \to \infty$, and

$$P\{\min_{N_{t,\nu} < n \leqq M_{t,\nu}} \hat{m}_n(t) < m(t) - \delta/2\} \geqq 1 - 2^{-\nu}\,,$$

and to be able to guarantee that for each $t = t_1, \cdots, t_{n_1}$, 1 we can get $N^*_{t,\nu}$ and $M^*_{t,\nu}$ for $\nu = 1, 2, \cdots$ so that $N^*_{t,\nu} < M^*_{t,\nu}$, $N^*_{t,\nu} \to \infty$ as $\nu \to \infty$, and

$$P\{\max_{N^*_{t,\nu} < n \leqq M^*_{t,\nu}} \hat{m}_n(t) > m(t) + \delta/2\} \geqq 1 - 2^{-\nu} .$$

We do this as follows. Start with $\{t_1, \cdots, t_{n_1}\}$. Lengthen the sequence at $t = 0$ using (A) and $\nu = 1$.
Then lengthen the sequence using (D) and $\nu = 1$ at $t = t_1, \cdots, t = t_{n_1}$ successively. Now lengthen the sequence at $t = 1$ using (B) and $\nu = 1$. Next lengthen the sequence using (C) and $\nu = 1$ at $t = t_1, \cdots, t = t_{n_1}$ successively. Next lengthen the sequence by adding an observation point midway between each adjacent pair of observation points already defined. (The final sequence of observation points may not be dense unless this is done after completion of the other lengthening procedures for each value of ν.) Repeat this procedure using $\nu = 2$, $\nu = 3$, $\cdots$ to obtain $\{t_i; i = 1, 2, \cdots\}$ as needed for the validity of the theorem.

LEMMA 4. *Let $r > 1$ and let F be a real valued function on $[0, \infty)$ such that $y^r F(y) \to 0$ as $y \to \infty$. Then for every $\varepsilon > 0$ there exists a positive integer M such that if $\{X_i : i = 1, 2, \cdots\}$ is an independent sequence of random variables such that $EX_i = 0$ and $P\{|X_i| \geqq y\} \leqq F(y)$ for all i and all $y \geqq 0$, then for all $n \geqq M$*

$$n^{r-1} P\{\sup_{k \geqq n} k^{-1}|X_1 + \cdots + X_k| \geqq \varepsilon\} \leqq \varepsilon .$$

PROOF. If $\{X_i\}$ is a specific sequence satisfying the assumptions of the lemma, then an argument of Baum and Katz ([1] Theorem 4) shows that

$$n^{r-1} P\{\sup_{k \geqq n} k^{-1}|X_1 + \cdots + X_k| \geqq \varepsilon\} \to 0$$

as $n \to \infty$. The type of argument used in the proof of Lemma 3 completes the proof of this lemma. We omit it.

PROOF OF THEOREM 3. As in the proof of Theorem 1 we let s be a fixed observation point in (0, 1) and choose u in $(s, 1)$ so that $m(u) - m(s) < \varepsilon/2$. For n large enough we have

$$\hat{m}_n(s) \leqq m(s) + \varepsilon/2 + \max_{\alpha \leqq s} \sum_{\{i: \alpha \leqq t_i \leqq u, i \leqq n\}} [Y_i - m(t_i)]/N_n([\alpha, u]) .$$

The sequence $\{Y_i - m(t_i)\}$ satisfies the conditions of Lemma 4 so when n is large enough that $N_n([s, u]) \leqq M = M_\delta$ then

$$\{N_n([s, u])\}^{r-1} P\{\max_{\alpha \leqq s} |\sum_{\{i: \alpha \leqq t_i \leqq u, i \leqq n\}} [Y_i - m(t_i)]/N_n([\alpha, u])| \geqq \delta\} \leqq \delta .$$

Thus

$$\{N_n([s, u])\}^{r-1} P\{\hat{m}_n(s) \geqq m(s) + \varepsilon/2 + \delta\} \leqq \delta$$

for n large enough. Since $\limsup_n N_n([s, u]) < \infty$, as $n \to \infty$

$$n^{r-1} P\{\hat{m}_n(s) - m(s) \geqq \varepsilon\} \to 0 .$$

The remainder of the proof follows the arguments of Theorem 1.

The following lemma is well known. It follows easily from Theorem A of [9] but is essentially due to other author(s) in earlier work.

LEMMA 5. *Let F be a real valued function on $[0, \infty)$ such that $F(x) \to 0$ as $x \to \infty$, and suppose there exists $\tau > 0$ such that $\int_0^\infty e^{\tau x}|dF(x)| < \infty$. Then for every $\varepsilon > 0$ there exist a positive constant C and a constant ρ in $[0, 1)$ such that if $\{X_i : i = 1, 2, \cdots\}$ is an independent sequence of random variables such that $EX_i = 0$ and $P\{|X_i| \geqq y\} \leqq F(y)$ for all i and all $y \geqq 0$, then for $n = 1, 2, \cdots$*

$$P\{\sup_{k \geqq n} k^{-1}|X_1 + \cdots + X_k| \geqq \varepsilon\} \leqq C\rho^n .$$

PROOF OF THEOREM 4. This proof is essentially the same (given Lemma 5) as the proofs of Theorems 1 and 3. We omit it.

4. The case $N = 2$. Throughout this section $\hat{m}_n$ will be any monotone estimator taking on the correct values at observation points. We will worry about specific values of $\hat{m}_n$ only at observation points. Values of t will be restricted to a closed rectangle which we take to be the unit square $[0, 1] \times [0, 1]$.

One example of a situation in which one might wish to estimate $m(t) = m(t^1, t^2)$ is just the extension of the example of Section 3 to include time of exposure to the drug. t^1 is dosage or concentration of the drug (or insecticide, $\cdots$) under consideration and t^2 is minus the length of time the organism(s) are exposed to it. $m(t^1, t^2)$ could reasonably be assumed to be non-decreasing in each variable.

Another example could involve a study of the effect of a particular drug on the reaction time of the individual. t^1 might be dosage (or dosage per unit body weight) of the drug under consideration and t^2 the length of time elapsed since the drug was administered. $m(t^1, t^2)$ might be considered to be either the expected reaction time of an individual t^2 time units after a dosage t^1 of the drug was administered, or it might be the average of these for some group of people. In either case one would expect m to be monotone in t^1 and, provided t^2 were at least large enough for the body to have assimilated the drug, to be monotone in t^2 also. One could, of course, change sign on either or both variables to guarantee that m be non-decreasing in each variable.

We retain the notation (14) and (15). The definitions are modified very slightly in that $t_k = (t_k^1, t_k^2)$ is no longer just a real number and the sets J are now subsets of the plane.

THEOREM 5. *Suppose*

(21) *m is continuous on $[0, 1] \times [0,1]$;*

$0 < a < b < 1$; for some $r \geqq 1$

(22) *$s^r F(s) \to 0$ as $s \to \infty$;*

(23) *$\int_0^\infty s|dF(s)| < \infty$;*

(24) *$\liminf_{n\to\infty} N_n(J)/n > 0$ for every rectangle J of the form $(\alpha_1, \beta_1) \times (\alpha_2, \beta_2)$ with $0 \leqq \alpha_i < \beta_i \leqq 1$ for $i = 1, 2$; and*

(25) *there exist positive constants M, c, and d with $c < d$ such that if η is a positive integer and $cn^{\frac{1}{2}} \leqq \eta \leqq dn^{\frac{1}{2}}$ then*

$$N_n([i/\eta, (i+1)/\eta] \times [j/\eta, (j+1)/\eta]) \leqq M$$

for, $i, j = 0, \cdots, \eta - 1$.
Then for every $\varepsilon > 0$, as $n \to \infty$

(26) $n^{(r-1)/2}P\{\sup_{t\in B}[\hat{m}_n(t) - m(t)] \geqq \varepsilon$ *or* $\sup_{t\in A}[m(t) - \hat{m}_n(t)] \geqq \varepsilon\} \to 0$ *where* $A = [a, 1] \times [a, 1]$ *and* $B = [0, b] \times [0, b]$.

Note that (23) and "$F(s) \to 0$ as $s \to \infty$" imply (22) when $r = 1$, and that (22) implies (23) when $r > 1$.

THEOREM 6. *Suppose m is continuous on $[0, 1] \times [0, 1]$, $0 < a < b < 1$, there exist positive constants λ and D such that $F(y) \leqq De^{-\lambda y}$ for $y \geqq 0$, (24) holds, and*

(27) *there exists a constant M such that for every positive integer η there exists an $n(\eta)$ such that*

$$N_n([i/\eta, (i+1)/\eta] \times [j/\eta, (j+1)/\eta]) \leqq M\eta^{-2}n$$

for $i, j = 0, \cdots, \eta - 1$ and all $n \geqq n(\eta)$.
Then for every $\varepsilon > 0$ there exist positive constants $\rho < 1$ and C such that for all $n \geqq 1$

(28) $P\{\sup_{t\in B}[\hat{m}_n(t) - m(t)] \geqq \varepsilon$ *or* $\sup_{t\in A}[m(t) - \hat{m}_n(t)] \geqq \varepsilon\} \leqq C\rho^n$ *where* $A = [a, 1] \times [a, 1]$ *and* $B = [0, b] \times [0, b]$.

As in the case $N = 1$ (Theorem 2), when $N = 2$ we do not get consistency if assumptions (24) and either (25) or (27) are simply replaced by the assumption that the sequence of observation points (the sequence $\{t_i\}$) is dense in the unit square. An example similar to that of Theorem 2 can be constructed. A grid of points is properly chosen and the sequence $\{t_i\}$ chosen so as to approach the grid points closer and closer, one at a time, diagonally from the lower left and then diagonally from the upper right. In this way the sequence $\{\hat{m}_n(t)\}$ is "forced" to behave badly at each of the grid points t. We omit the details since the argument is like that given in the "proof" of Theorem 2, and since a great deal of space would be consumed if it were given in complete detail.

We will need the following lemmas in the proof of Theorem 5.

LEMMA 6. *Let F be a non-increasing function from $[0, \infty)$ into $[0, 1]$ such that $F(y) \to 0$ as $y \to \infty$ and such that $\int_0^\infty y|dF(y)| < \infty$. Then there exist a non-increasing sequence $\{a_n\}$ of real numbers converging to zero and a non-increasing real valued function g with $\lim_{T\to\infty} g(T) = 0$ such that whenever $X_1, X_2, \cdots$ is an independent sequence of random variables satisfying $EX_i = 0$ and $P\{|X_i| \geqq y\} \leqq F(y)$ for $i = 1, 2, \cdots$ and $0 \leqq y$, then if*

(29) $$X_n^* = \max_{0\leqq k\leqq n} S_k/n \qquad \textit{and} \qquad S_0 = 0$$

we have

$$E|X_n^*| \leqq a_n \quad \text{for} \quad n = 1, 2, \cdots, \qquad \text{and} \tag{30}$$

$$|E[(X_n^* - EX_n^*)I_{\{|X_n^* - EX_n^*| < T\}}]| \leqq g(T) \tag{31}$$

for $n = 1, 2, \cdots$ *and* $T \geqq 0$.

PROOF. Let $X_1, X_2, \cdots$ be a sequence satisfying the assumptions of the lemma.

We see that $E|X_n^*| \leqq \int_0^\infty x|dF(x)|$ for all n so we may assume $a_1 < \infty$, and in fact $a_1 \leqq \int_0^\infty x|dF(x)|$.

Suppose $\varepsilon > 0$. There exists $\delta > 0$ such that if $P(A) \leqq \delta$ and X is a random variable satisfying $P\{|X| \geqq x\} \leqq F(x)$ for all x then $\int_A |X|\, dP \leqq \varepsilon/2$. From Lemma 3 there exists M (independent of the particular sequence $\{X_i\}$ involved) such that $P\{|X_n^*| > \varepsilon/2\} < \delta$ for all $n \geqq M$. Then if $n \geqq M$

$$E|X_n^*| \leqq \int_{\{|X_n^*| \leqq \varepsilon/2\}} |X_n^*|\, dP + n^{-1} \sum_{k=1}^n \int_{\{|X_n^*| > \varepsilon/2\}} |X_k|\, dP \leqq \varepsilon .$$

This completes the proof of (30).

Now note that

$$\begin{aligned} &|E[(X_n^* - EX_n^*)I_{\{|X_n^* - EX_n^*| < T\}}]| \\ &\quad = |E[(X_n^* - EX_n^*)I_{\{|X_n^* - EX_n^*| \geqq T\}}]| \\ &\quad \leqq \int_{\{|X_n^* - EX_n^*| \geqq T\}} (|X_n^*| + a_1)\, dP \\ &\quad \leqq n^{-1} \sum_{k=1}^n \int_{\{|X_n^* - EX_n^*| \geqq T\}} |X_k|\, dP + a_1 P\{|X_n^* - EX_n^*| \geqq T\} . \end{aligned} \tag{32}$$

Using (30) of this lemma and using Lemma 3 choose M (independent of the sequence $\{X_i\}$) such that $n \geqq M$ implies $E|X_n^*| \leqq \frac{1}{2}$ and $P\{|X_n^*| \geqq \frac{1}{2}\} \leqq \min\{\delta, \varepsilon/2a_1\}$. Then, using (32), we see that if $T \geqq 1$ and $n \geqq M$ then

$$|E[(X_n^* - EX_n^*)I_{\{|X_n^* - EX_n^*| < T\}}]| \leqq \varepsilon .$$

Now notice that if $T > a_1$ and $n \leqq M$, then

$$\begin{aligned} \{|X_n^* - EX_n^*| \geqq T\} &\subset \{|X_n^*| \geqq T - a_1\} \\ &\subset \bigcup_{k=1}^n \left\{|X_k| \geqq \frac{T - a_1}{n}\right\} \subset \bigcup_{k=1}^M \left\{|X_k| \geqq \frac{T - a_1}{M}\right\} . \end{aligned}$$

There exists an x_0 such that $x_0 > 1$ and such that $\int_{x_0}^\infty x|dF(x)| < \min\{\delta/M, \varepsilon/2a_1 M\}$. Then if $T \geqq x_0 M + a_1$ we see that as long as $\{X_n\}$ satisfies the hypotheses of the lemma we have

$$P\{|X_n| \geqq (T - a_1)/M\} < \min\{\delta/M, \varepsilon/2a_1 M\}$$

for all n, and for $n \leqq M$ we have

$$P\{|X_n^* - EX_n^*| \geqq T\} < \min\{\delta, \varepsilon/2a_1\} .$$

Thus (32) is bounded by ε. Putting these together we see that if $T \geqq \max\{1, x_0 M + a_1\}$ then

$$|E[(X_n^* - EX_n^*)I_{\{|X_n^* - EX_n^*| < T\}}]| \leqq \varepsilon$$

for all n.

LEMMA 7. *Let F be as in Lemma 6, suppose $t \geqq 1$, and suppose $y^t F(y) \to 0$ as $y \to \infty$. Then there exists a real valued function $f(y)$ such that $f(y) \to 0$ as $y \to \infty$ and such that whenever $X_1, X_2, \cdots$ is an independent sequence of random variables satisfying $EX_i = 0$ and $P\{|X_i| \geqq y\} \leqq F(y)$ for $i = 1, 2, \cdots$ then*

$$y^t \sup_n P\{|X_n^* - EX_n^*| \geqq y\} \leqq f(y)$$

for all $y \geqq 0$.

PROOF. Since $EX_n^* \to 0$ uniformly in $\{X_i\}$ as $n \to \infty$ (from Lemma 6) it suffices to show the existence of $f(y)$ as above such that

$$y^t \sup_n P\{|X_n^*| \geqq y\} \leqq f(y)$$

for every sequence $\{X_i\}$ satisfying the assumptions of the lemma. Now if $X_1, X_2, \cdots$ is a sequence satisfying the assumptions of the lemma, and if

$$\begin{aligned} X_i^T &= X_i \quad \text{if} \quad |X_i| < ny \\ &= 0 \quad \text{otherwise,} \end{aligned}$$

then

$$\begin{aligned} P\{|X_n^*| \geqq y\} &\leqq P\{\max_{1\leqq k\leqq n} |S_k|/n \geqq y\} \\ &\leqq \textstyle\sum_{k=1}^n P\{|X_k| \geqq ny\} && (33) \\ &\quad + P\{|n^{-1} \max_{1\leqq k\leqq n} \textstyle\sum_{j=1}^k EX_j^T| \geqq y/2\} && (34) \\ &\quad + P\{|n^{-1} \max_{1\leqq k\leqq n} \textstyle\sum_{j=1}^k (X_j^T - EX_j^T)| \geqq y/2\}\,. && (35) \end{aligned}$$

In the following whenever $o(y^\alpha)$ is used for any real α (including $\alpha = 0$; i.e. $y^\alpha = 1$) it will denote some fixed function (independent of n and of $\{X_j\}$) such that $o(y^\alpha)/y^\alpha \to 0$ as $y \to \infty$.

Expression (33) is bounded by

$$\begin{aligned} nF(ny) &= y^{-t} n^{1-t}[(ny)^t F(ny)] \\ &\leqq y^{-t}[(ny)^t F(ny)] \\ &= o(y^{-t})\,. \end{aligned}$$

Note that the bound is uniform in n and in the sequences (X_i).

Expression (34) is bounded by

$$P\{n^{-1} \textstyle\sum_{j=1}^n E|X_j| \geqq y/2\} \leqq P\{\int_0^\infty x|dF(x)| \geqq y/2\}\,.$$

This expression is zero when $y > 2 \int_0^\infty x|dF(x)|$. We use the "Generalized Kolmogorov Inequality" (see problems 2 and 3 on page 263 of [11]) to bound (35). Let ν be a fixed even positive integer with $\nu > t$. Then (35) is bounded by

$$(2/ny)^\nu E[\textstyle\sum_{j=1}^n (X_j^T - EX_j^T)]^\nu\,. \tag{36}$$

Note that

$$\begin{aligned} |EX_j^T| &= |\textstyle\int_{|X_j|<ny} X_j\,dP| = |\textstyle\int_{|X_j|>ny} X_j\,dP| \\ &\leqq \textstyle\int_{|X_j|\geqq ny} |X_j|\,dP \leqq \int_{x\geqq ny} x|dF(x)| \end{aligned}$$

and that this converges to zero uniformly in n and in j as $y \to \infty$. If we let

$a_{n,j} = 1$ for $j = 1, \cdots, n$ and 0 otherwise then we can use the techniques and notation used on pages 354–355 of [8] to obtain

$$E[\sum_{j=1}^{n} (X_j^T - EX_j^T)]^\nu = \sum_1 \sum_2 C^{(\nu)}(m_1, \cdots, m_{a+b}) \times \prod_{k=1}^{a+b} a_{n,f(k)}^{m_k} E(X_{f(k)}^T - EX_{f(x)}^T)^{m_k} .$$

If t_0 is the largest positive integer less than t, then the fact that $y^t F(y) \to 0$ as $y \to \infty$ (along with the fact that the tails of the distributions of the X_j's are bounded by F uniformly in j) guarantees that $E|X_j^T|^{m_k}$ are uniformly bounded in j, n, and y for $m_k = 1, \cdots, t_0$. Combine this with the uniform bound on $|EX_j^T|$ to obtain a uniform bound on $|E(X_j^T - EX_j^T)^{m_k}|$ for $k = 1, \cdots, a$. As in [8], we use C to represent any constant whose exact numerical value does not matter. For $k = 1, \cdots, a$ we get

$$|a_{n,j}^{m_k} E(X_j^T - EX_j^T)^{m_k}| \leqq C|a_{n,j}| .$$

If $a + 1 \leqq k \leqq a + b$ then

$$\begin{aligned} |E(X_j^T - EX_j^T)^{m_k}| &\leqq 2^\nu |EX_j^T|^{m_k} + 2^\nu E|X_j^T|^{m_k} \leqq CE|X_j^T|^{m_k} \\ &\leqq C[\int_{x<ny} x^{m_k} |dF(x)| + (ny)^{m_k} F(ny-)] \\ &\leqq C[\int_0^{ny} x^{m_k - 1} F(x)\, dx + (ny)^{m_k - t} o(1)] \\ &\leqq C[1 + (ny)^{m_k - t} o(1) + \int_1^{ny} x^{m_k - t - 1} o(x^0)\, dx] . \end{aligned}$$

If $m_k = t$ this is bounded by $o(1) \log(ny)$. If $m_k > t$ it is bounded by $(ny)^{m_k - t} o(1)$. (We continue using the notation used on pages 354–355 of [8].) We want to bound

$$(37) \qquad (2/ny)^\nu \sum_2 |C^{(\nu)}(m_1, \cdots, m_{a+b}) \prod_{k=1}^{a+b} a_{n,f(k)}^{m_k} E(X_{f(k)}^T - EX_{f(k)}^T)^{m_k}|$$

for various of a and b. We first consider the case $a = 0$ with b fixed; we consider separately the two subcases $t = 1$ and $t > 1$. If $t = 1$ and $a = 0$ then $m_k \geqq 2 > t$ so (37) is bounded above by

$$\begin{aligned} C[ny]^{-\nu}[\sum_2 \prod_{k=1}^{b} a_{n,f(k)}][ny]^{\nu - bt} o(1) &\leqq Cn^{-b(t-1)} y^{-bt} o(1) \\ &\leqq Cy^{-b} o(1) . \end{aligned}$$

But this is $o(y^{-t})$ since $b \geqq 1$ and $t = 1$. If $t > 1$ and $b = 1$ then $m_1 = \nu > t$ so (37) is bounded above by

$$\begin{aligned} C[ny]^{-\nu}[\sum_2 a_{n,f(1)}][ny]^{\nu - t} o(1) &\leqq Cn^{-(t-1)} y^{-t} o(1) \\ &\leqq o(y^{-t}) . \end{aligned}$$

If $t > 1$ and $b > 1$ then (37) is bounded above by

$$C[ny]^{-\nu}[\sum_2 \prod_{k=1}^{b} a_{n,f(k)}][ny]^{\nu - bt}[\log(ny)]^b \leqq Cn^{-b(t-1)} y^{-bt}[\log(ny)]^b .$$

Since $t > 1$ and $b > 1$ this is $o(y^{-t})$. For the case $a > 0$ we see that (37) is bounded by

$$\begin{aligned} & C[ny]^{-\nu}[\sum_2 \prod_{k=1}^{a+b} a_{n,f(k)}][ny]^{\sum_{k=a+1}^{a+b} m_k - bt}[\log(ny)]^b[o(1)] \\ (38) \qquad & \leqq C[ny]^{-\sum_{k=1}^{a} m_k - bt}[n]^{a+b}[\log(ny)]^b[o(1)] . \end{aligned}$$

Note that $m_k \geqq 2$ and $t \geqq 1$ in (38). If $b = 0$ then $-\sum_{k=1}^{a} m_k - bt = -\sum_{k=1}^{a} m_k = -\nu$ so that (37) is bounded above by

$$C(ny)^{-\nu} n^a [o(1)] \leqq o(y^{-t}) .$$

If $b > 0$ then (38) is bounded above by

$$Cn^{-2a-b} y^{-2-t} n^{a+b} [\log (ny)]^b \leqq o(y^{-t}) .$$

When (36) is expanded as in [8] it is a sum of the form $\sum_1 (\cdots)$. We have shown that each of the terms of this finite sum is $o(y^{-t})$ so the whole sum is $o(y^{-t})$. This completes the proof of Lemma 7.

LEMMA 8. *Let F be a non-increasing function from* $[0, \infty)$ *into* $[0, 1]$ *such that* $F(y) = 0(y^{-1})$. *Let* $\{X_i : i = 1, 2, \cdots\}$ *be an independent sequence of random variables such that* $P\{|X_i| \geqq x\} \leqq F(x)$ *for* $i = 1, 2, \cdots$ *and* $x \geqq 0$ *and such that*

$$\lim_{T\to\infty} \sup_i |EX_i I_{\{|X_i| < T\}}| = 0 . \tag{39}$$

Let $\{a_{n,k} : n, k = 1, 2, \cdots\}$ *be real numbers such that*

$$\sum_k |a_{n,k}| \leqq C < \infty \tag{40}$$

for $n = 1, 2, \cdots$ *and such that*

$$\sup_k |a_{n,k}| \to 0 \qquad \textit{as} \quad n \to \infty . \tag{41}$$

Then

$$P\{|\sum_k a_{n,k} X_k| \geqq \varepsilon\} = o(\sum_k |a_{n,k}|)$$

for every $\varepsilon > 0$.

PROOF. This is almost Theorem 2b (case $t = 1$) of [10] and the proof is almost identical to the proof of Theorem 2b. The bounds (4.5), (4.6), and (4.7) on page 85 of [10] are used. In dealing with (4.5) and (4.7) in the proof of Theorem 2b, the condition $\sum_k |a_{n,k}| \to 0$ was used only to imply $\sup_k |a_{n,k}| \to 0$; we have assumed the latter in this lemma. In dealing with (4.6), the condition $\sum_k |a_{n,k}| \to 0$ was used to imply $|\sum_k a_{n,k} EY_{n,k}| \to 0$; in the proof of this lemma assumption (39) serves that function. We omit further details.

PROOF OF THEOREM 5. Choose $\varepsilon > 0$. It is sufficient to prove that

$$n^{(r-1)/2} P\{\sup_{t \in B} [\hat{m}_n(t) - m(t)] \geqq \varepsilon\} \to 0$$

and that

$$n^{(r-1)/2} P\{\sup_{t \in A} [m(t) - \hat{m}_n(t)] \geqq \varepsilon\} \to 0 .$$

The proofs are essentially the same. We will prove only the former.

Because of (24) and the uniform continuity of m on $[0, 1] \times [0, 1]$ there exists a k_0 such that for every $t \in B$ there exists a $t' \in \{t_1, \cdots, t_{k_0}\}$ such that $t < t' < ((1 + b)/2, (1 + b)/2)$ and such that $m(t) > m(t') - \varepsilon/3$. Let T^* be the collection of points $t' \in \{t_1, \cdots, t_{k_0}\}$ such that $t' < ((1 + b)/2, (1 + b)/2)$. For each $t \in B$ there exists a $t' \in T^*$ such that $t < t'$, such that $m(t') - m(t) \leqq \varepsilon/3$, and (from the monotonicity of $\hat{m}_n$) such that $\hat{m}_n(t) - \hat{m}_n(t') < 0$. It follows that

$$\{\hat{m}_n(t) - m(t) \geqq \varepsilon\} \subset \{\hat{m}_n(t') - m(t') \geqq 2\varepsilon/3\} ,$$

and thus it suffices to prove that

$$n^{(r-1)/2}P\{\max_{t \in T^*} [\hat{m}_n(t) - m(t)] \geqq 2\varepsilon/3\} \to 0$$

or that

$$n^{(r-1)/2}P\{\hat{m}_n(t) - m(t) \geqq 2\varepsilon/3\} \to 0$$

for every $t \in T^*$.

Now fix $t \in T^*$ and let $s = t + (\Delta t, \Delta t)$ where Δt is chosen so that $0 < \Delta t < (1 - b)/2$ and such that $m(s) < m(t) + \varepsilon/3$.

In that which follows almost everything depends on n. In order to simplify notation this dependence is not exhibited (by using subscripts or superscripts for example).

For n sufficiently large let $\eta = \eta(n)$ be a positive integer in the interval $[cn^{\frac{1}{2}}, dn^{\frac{1}{2}}]$ and divide $[0, 1) \times [0, 1)$ into η^2 "little squares" of the form

$$[i/\eta, (i + 1)/\eta) \times [j/\eta, (j + 1)/\eta) .$$

The "picture" (for $\eta = 16$) is as in Figure 1. Define

$$\Omega = \{x = (x^1, x^2) : (0, 0) \leqq x < s\}$$

and

$$D_0 = \{x : t \leqq x < s\} .$$

Let $\mathscr{S}$ be the collection of sets each of which is nonempty and is the intersection of Ω with one of the η^2 "little squares." Decompose $\mathscr{S}$ into "chains" so that all members of a given chain have the same main diagonal line running from upper right to lower left. One chain is shown in Fig. 1. Number these chains from one to λ. Note that $\lambda \leqq 2\eta - 1$. Suppose the ith chain consists of the sets (members of $\mathscr{S}$) $S_{i,1}, \cdots, S_{i,\mu_i}$ where the ordering starts with the member of the chain at the upper right and proceeds downward and to the left. Let

$$\mathscr{W}_{i,\tau} = \{Y_k : 1 \leqq k \leqq n \text{ and } t_k \in S_{i,\tau}\} .$$

Remove one Y_k from each $\mathscr{W}_{i,\tau}$ (i fixed) which is nonempty. There will be some number $\nu_{i,1}$ of these. If they are $Y_{k_1}, \cdots, Y_{k_{\nu_{i,1}}}$ taken at $t_{k_1} > t_{k_2} > \cdots > t_{k_{\nu_{i,1}}}$ define $Y_{i,1,j} = Y_{k_j}$. (Note that since two different t's are in two different members of the ith chain it follows that one is to the upper right of the other. Thus the $\nu_{i,1}$ observation points are linearly ordered.) Now remove one more Y_k from each $\mathscr{W}_{i,\tau}$ which is not now empty. There will be $\nu_{i,2}$ of these. If they are $Y_{\alpha_1}, \cdots, Y_{\alpha_{\nu_{i,2}}}$ taken at $t_{\alpha_1} > \cdots > t_{\alpha_{\nu_{i,2}}}$ define $Y_{i,2,j} = Y_{\alpha_j}$. Continue this procedure until all the collections $\mathscr{W}_{i,\tau}$ have been emptied. If

$$\gamma_i = \max_\tau \#(\mathscr{W}_{i,\tau})$$

and if the above procedure is carried out for $i = 1, \cdots, \lambda$ then $Y_{i,j,\nu}$ will have

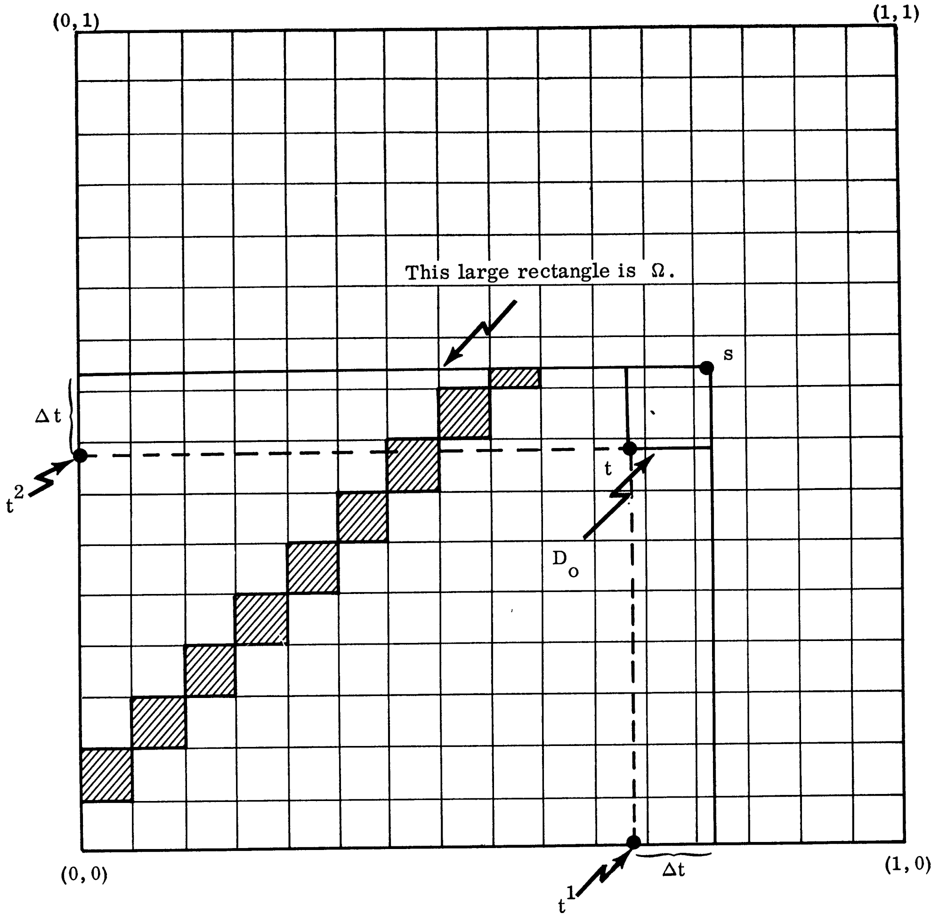

FIG. 1.

been defined for $i = 1, \cdots, \lambda$; $j = 1, \cdots, \gamma_i$; and $\nu = 1, \cdots, \nu_{i,j}$. From the definition

$$\hat{m}_n(t) = \max_{\{L: t \in L \in \mathscr{L}_2\}} \min_{\{K: t \in K \in \mathscr{L}_2^c\}} \frac{\sum_{\{\alpha: t_\alpha \in L \cap K, \alpha \leqq n\}} Y_\alpha}{N_n(L \cap K)}$$
$$\leqq \max_{\{L: t \in L \in \mathscr{L}_2\}} \frac{\sum_{\{\alpha: t_\alpha \in L \cap \Omega, \alpha \leqq n\}} Y_\alpha}{N_n(L \cap \Omega)} .$$

Since $EY_\alpha \leqq m(t) + \varepsilon/3$ for $t_\alpha \in \Omega$ we have

$$\hat{m}_n(t) \leqq \max_{\{L: t \in L \in \mathscr{L}_2\}} \frac{\sum_{\{\alpha: t_\alpha \in L \cap \Omega, \alpha \leqq n\}} (Y_\alpha - EY_\alpha)}{N_n(L \cap \Omega)} + m(t) + \varepsilon/3$$
$$\leqq \max_{\{L: t \in L \in \mathscr{L}_2\}} \frac{\sum_{\{\alpha: t_\alpha \in L \cap \Omega, \alpha \leqq n\}} (Y_\alpha - EY_\alpha)}{N_n(D_0)} + m(t) + \varepsilon/3 .$$

Since $\sum_{\{\alpha: t_\alpha \in L \cap \Omega, \alpha \leqq n\}} (Y_\alpha - EY_\alpha)$ can be written as a sum on i and j of partial sums of the sequence $Y_{i,j,1} - EY_{i,j,1}, \cdots, Y_{i,j,\nu_{i,j}} - EY_{i,j,\nu_{i,j}}$, it follows that if we define

$$Z_{i,j} = \max_{\nu=0,\cdots,\nu_{i,j}} \sum_{k=1}^{\nu} (Y_{i,j,k} - EY_{i,j,k})$$

where $\sum_{k=1}^{0} a_k$ is defined to be zero, then

$$\hat{m}_n(t) \leqq \sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} Z_{i,j}/N_n(D_0) + m(t) + \varepsilon/3$$

so that

$$P\{\hat{m}_n(t) - m(t) \geqq 2\varepsilon/3\} \leqq P\{\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} Z_{i,j} \geqq N_n(D_0)\varepsilon/3\} .$$

From (24), if n is large enough, then $N_n(D_0) \geqq C_1 n$ for some positive constant C_1 so

$$P\{\hat{m}_n(t) - m(t) \geqq 2\varepsilon/3\} \leqq P\left\{\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \frac{\nu_{i,j}}{n}\left(\frac{Z_{i,j}}{\nu_{i,j}}\right) \leqq 2\varepsilon'\right\}$$

where $\varepsilon' = C_1\varepsilon/6$. This is bounded by

$$P\left\{\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \frac{\nu_{i,j}}{n}\left[\frac{Z_{i,j}}{\nu_{i,j}} - E\left(\frac{Z_{i,j}}{\nu_{i,j}}\right)\right] \geqq \varepsilon'\right\} \tag{42}$$

$$+ P\left\{\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \frac{\nu_{i,j}}{n} E\left|\frac{Z_{i,j}}{\nu_{i,j}}\right| \geqq \varepsilon'\right\} . \tag{43}$$

From Lemma 6 there exists m^* such that $\nu_{i,j} \geqq m^*$ implies

$$E|Z_{i,j}/\nu_{i,j}| \leqq \sup_{m \leqq m^*} a_m < \varepsilon'/2 .$$

Then if $\sum_1$ is the sum over all pairs (i, j) such that $1 \leqq i \leqq \lambda$, $1 \leqq j \leqq \gamma_i$, and $\nu_{i,j} \leqq m^*$; and if $\sum_2$ is the sum over all pairs (i, j) such that $1 \leqq i \leqq \lambda$, $1 \leqq j \leqq \gamma_i$, and $\nu_{i,j} \geqq m^*$; we have

$$\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \frac{\nu_{i,j}}{n} E\left|\frac{Z_{i,j}}{\nu_{i,j}}\right| \leqq \sum\nolimits_1 m^* a_1/n + \sum\nolimits_2 (\nu_{i,j}/n)(\varepsilon'/2) . \tag{44}$$

Now $\lambda \leqq 2\eta - 1$ and (from assumption (25) and the definition of γ_j) $\gamma_i \leqq M$ so that the number of terms in $\sum_1$ is bounded by $2\eta M \leqq 2dMn^{\frac{1}{2}}$. Thus

$$\sum\nolimits_1 m^* a_1/n \leqq 2dMm^* a_1/n^{\frac{1}{2}} \rightarrow 0 .$$

We have

$$\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \nu_{i,j} = N_n(\Omega) \leqq n$$

so that $\sum_2 (\nu_{i,j}/n)(\varepsilon'/2) \leqq \varepsilon'/2$.

Thus lim sup of expression (44) is at most $\varepsilon'/2$ so that for n large enough (43) $= 0$. Note that

$$\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \frac{\nu_{i,j}}{n} \leqq 1 = n^0 ,$$

$$\max_{i,j} \left|\frac{\nu_{i,j}}{n}\right| \leqq \eta/n \leqq dn^{-\frac{1}{2}} , \qquad \text{and}$$

$$\sum_{i=1}^{\lambda} \sum_{j=1}^{\gamma_i} \left(\frac{\nu_{i,j}}{n}\right)^r \leqq \lambda M(\eta/n)^r$$

$$\leqq 2\eta M(\eta/n)^r \leqq Cn^{-(r-1)/2} .$$

Lemma 7 says that the tails of the distribution functions of the random variables $(Z_{i,j} - EZ_{i,j})/\nu_{i,j}$ (for all i, j, n) are uniformly $o(y^{-t})$. When $r > 1$ this fact and the inequalities directly above enable us to apply Theorem 2 of Franck and Hanson [8] with $\alpha = 0$, $\beta = \frac{1}{2}$, $t = r$, and $\rho = (r-1)/2$; this shows that (42) is $o(n^{-(r-1)/2})$ and completes the proof of Theorem 5 when $r > 1$. When $r = 1$ we use Lemma 8 (with Lemmas 6 and 7) to show that (42) is $o(1)$, completing the proof of Theorem 5.

PROOF OF THEOREM 6. Choose $\varepsilon > 0$. Defining T^* as in the proof of Theorem 5 we will show that for each $t \in T^*$ there exist positive constants C and $\rho < 1$ such that

$$P\{\hat{m}_n(t) - m(t) \geqq 2\varepsilon/3\} \leqq C\rho^n.$$

for all positive integers n. This is enough to guarantee that there exist positive constants C^* and $\rho^* < 1$ such that

$$P\{\sup_{t\in B}[\hat{m}_n(t) - m(t)] \geqq \varepsilon\} \leqq C^*(\rho^*)^n$$

for $n = 1, 2, \cdots$. The part of the proof involving

$$P\{\sup_{t\in A}[m(t) - \hat{m}_n(t)] \geqq \varepsilon\}$$

is similar and we omit it.

Now fix $t \in T^*$ and, as in the proof of Theorem 5, choose s, define Ω and D_0, and observe that (as in the proof of Theorem 5)

$$\hat{m}_n(t) - m(t) \leqq \max_{\{L: t\in L\in\mathscr{L}_2\}} \frac{\sum_{\{\alpha: t_\alpha\in L\cap\Omega, \alpha\leqq n\}}(Y_\alpha - EY_\alpha)}{N_n(D_0)} + \varepsilon/3\,. \tag{45}$$

We remind the reader that (24) implies that there exists a positive constant C_1 such that $N_n(D_0) \geqq C_1 n$ for n sufficiently large. Thus for n sufficiently large

$$P\{\hat{m}_n(t) - m(t) \geqq 2\varepsilon/3\}$$

is bounded above by

$$P\{\max_{\{L: t\in L\in\mathscr{L}_2\}} \sum_{\{\alpha: t_\alpha\in L\cap\Omega, \alpha\leqq n\}}(Y_\alpha - EY_\alpha) \geqq C_1 n\varepsilon/3\}\,. \tag{46}$$

Now for a fixed η (to be chosen later) divide $[0, 1) \times [0, 1)$ into the η^2 squares $[i/\eta, (i+1)/\eta) \times [j/\eta, (j+1)/\eta)$ for $i, j = 0, 1, \cdots, \eta - 1$. The reader is reminded of Fig. 1. Label the squares in the ith row (from the top) from right to left $I_{i1}, \cdots, I_{i\eta}$. Set $I_{i,0} = I_{i,\eta+1} = \phi$ for all i. Define $\mathscr{R}$ $(\subset \mathscr{L}_2)$ to be the collection of sets of the form $\bigcup_{i=1}^{n}\bigcup_{j=0}^{j_i} I_{ij}$ where $\eta \geqq j_1 \geqq j_2 \geqq \cdots \geqq j_\eta \geqq 0$, and for a fixed $R \in \mathscr{R}$ define $F_R = \bigcup_{i=1}^{\eta}\bigcup_{j=j_i+1}^{j_{i-1}+1} I_{ij}$ where $j_0 = \eta$. The number of sets I_{ij} in F_R is $\sum_{i=1}^{\eta}\sum_{j=j_i+1}^{j_{i-1}+1} 1 = \eta + \sum_{i=1}^{\eta}(j_{i-1} - j_i) \leqq 2\eta$. If $L \in \mathscr{L}_2$ then let $R(L)$ be the largest $R \in \mathscr{R}$ such that $R \subset L$. Note that $L \cap ([0, 1] \times [0, 1]) \subset R(L) \cup F_{R(L)}$. For convenience let $S_{A,n} = \sum_{\{\alpha:\alpha\leqq n, t_\alpha\in A\}}(Y_\alpha - EY_\alpha)$ where $A \subset [0, 1] \times [0, 1]$ and the sum is 0 if $\{\alpha : \alpha \leqq n,\ t_\alpha \in A\} = \phi$. Expression (46) is bounded above by

$$\sum_{\{R\in\mathscr{R}: t\in R\cup F_R\}} P\{S_{R\cap\Omega,n} + \max_{\{A: A\subset F_R\}} S_{A\cap\Omega,n} \geqq C_1 n\varepsilon/3\} \tag{47}$$

and for $R \in \mathscr{R}$ the summand in (47) is bounded by

$$P\{S_{R\cap\Omega,n} \geqq C_1 n\varepsilon/6\} + \sum_{\{A:A\subset F_R\}} P\{S_{A\cap\Omega,n} \geqq C_1 n\varepsilon/6\} . \tag{48}$$

We use Theorem A of [9] with $p = 1$, with weights $a_k = 1/n$ if $1 \leqq k \leqq n$ and $t_k \in A$ and $a_k = 0$ otherwise, and with ε replaced by $C_1\varepsilon/6$. This shows the existence of a positive constant C_2 depending only on D, λ, C_1, and ε (and in particular not on η) such that

$$P\{S_{A,n} \geqq C_1 n\varepsilon/6\} \leqq \exp(-C_2 n)$$

for every $A \subset [0, 1] \times [0, 1]$. Since

$$N_n(\overline{I_{ij}}) \leqq M\eta^{-2}n$$

for n sufficiently large, we see that, for n sufficiently large, there are at most $(2\eta)(M\eta^{-2}n)$ points in F_R and at most $2^{2M\eta^{-1}n}$ distinct subsets of F_R. Thus (48) is bounded above by

$$(2^{2M\eta^{-1}n} + 1) \exp(-C_2 n)$$

and (47) is bounded above by

$$2\eta^{\eta}(2^{2M\eta^{-1}n}) \exp(-C_2 n) .$$

If η is chosen large enough that

$$2^{2M\eta^{-1}} e^{-C_2} < 1$$

then

$$P\{\hat{m}_n(t) - m(t) \geqq 2\varepsilon/3\} \leqq C\rho^n \tag{49}$$

where $\rho = 2^{2M\eta^{-1}} e^{-C_2}$ and C is a positive constant at least as large as $2\eta^{\eta}$.

5. Concluding remarks. It would be interesting to know whether the rates of convergence presented herein are sharp (or reasonably so). It seems likely that the rates results presented in Section 3 are about as good as can be obtained, but then the method of proof used in that section would seem (at least at first) to involve inequalities which should be reasonably tight. It would be interesting to know whether the rates results presented in Section 4 can be improved significantly, and whether the bound consisting of expressions (42) plus (43) can be improved upon.

In Section 4 we ran into the following problem. Let Ω be a finite set, and for each i in Ω let X_i be a random variable. Let $\mathscr{L}$ be a lattice of subsets of Ω with $\Omega \in \mathscr{L}$. Define

$$S_T = \sum_{i\in T} X_i$$

for every subset T of Ω, and define $S_\phi = 0$. What can be said about probabilities of the form

$$P\{\max_{L\in\mathscr{L}} S_L \geqq \varepsilon\} ?$$

If $\Omega = \{1, \cdots, N\}$ and $\mathscr{L} = \{\phi, L_1, \cdots, L_n\}$ where $L_i = \{1, \cdots, i\}$ then the probabilities above occur standardly. If $\mathscr{L} = 2^\Omega$ then the probabilities above may be much bigger; how much bigger? Can Kolmogorov's inequality be modified so that it is of the form

$$P\{\max_{L \in \mathscr{L}} |S_L| \geqq \varepsilon\} \leqq C_{\mathscr{L}}(\sigma_1^2 + \cdots + \sigma_N^2)/\varepsilon^2$$

where $C_{\mathscr{L}}$ is a constant depending on the lattice $\mathscr{L}$? ($C_{\mathscr{L}} = N$ will always work but is it a good bound when $\mathscr{L} = 2^\Omega$, and how can it be improved upon when $\mathscr{L}$ is a smaller lattice?)

REFERENCES

[1] BAUM, L. E. and KATZ, MELVIN (1965). Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120** 108–123.

[2] BRUNK, H. D. (1955). Maximum likelihood estimates of monotone parameters. *Ann. Math. Statist.* **26** 607–616.

[3] BRUNK, H. D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Statist.* **29** 437–454.

[4] BRUNK, H. D. (1963). On an extension of the concept conditional expectation. *Proc. Amer. Math. Soc.* **14** 298–304.

[5] BRUNK, H. D. (1965). Conditional expectation given a σ-lattice and applications. *Ann. Math. Statist.* **36** 1339–1350.

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

[7] BRUNK, H. D., EWING, G. M. and UTZ, W. R. (1957). Minimizing integrals in certain classes of monotone functions. *Pacific J. Math.* **7** 833–844.

[8] FRANCK, W. E. and HANSON, D. L. (1966). Some results giving rates of convergence in the law of large numbers for weighted sums of independent random variables. *Trans. Amer. Math. Soc.* **124** 347–359.

[9] HANSON, D. L. (1967). Some results relating moment generating functions and convergence rates in the law of large numbers. *Ann. Math. Statist.* **38** 742–750.

[10] HANSON, D. L. and WRIGHT, F. T. (1971). Some convergence results for weighted sums of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **19** 81–89.

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

DEPARTMENT OF STATISTICS
UNIVERSITY OF MISSOURI
COLUMBIA, MISSOURI 65201