LINEAR BOUNDS ON THE EMPIRICAL DISTRIBUTION FUNCTION

BY GALEN R. SHORACK* AND JON A. WELLNER

University of Washington and University of Rochester

Let $\Gamma_n$ denote the empirical df of a sample from the uniform (0, 1) df $I$. Let $\xi_{nk}$ denote the $k$th smallest observation. Let $\lambda_n > 1$. Let $A_n$ denote the event that $\Gamma_n$ intersects the line $\lambda_n I$ on [0, 1] and let $B_n$ denote the event that $\Gamma_n$ intersects the line $I/\lambda_n$ on $[\xi_{n1}, 1]$. Conditions on $\lambda_n$ are given that determine whether $P(A_n \text{ i.o.})$ and $P(B_n \text{ i.o.})$ equal 0 or 1. Results for $A_n$ (for $B_n$) are related to upper class sequences for $1/(n^{\xi_{n1}})$ (for $n^{\xi_{n2}}$).

Upper class sequences for $n^{\xi_{nk}}$, with $k > 1$, are characterized.

In the case of nonidentically distributed random variables, we present the result analogous to $P(A_n \text{ i.o.}) = 0$.

1. Introduction and statement of the theorems. Let $\xi_1, \ldots, \xi_n$ be independent uniform (0, 1) random variables having empirical df $\Gamma_n$ and whose ordered values are $0 \leq \xi_{n1} \leq \cdots \leq \xi_{nn} \leq 1$. The true df is the identity function on [0, 1], which we denote by $I$.

We let $\|f\|_{a,b} \equiv \sup_{a \leq t \leq b} |f(t)|$, and we simply write $\|f\|$ in case $a = 0$ and $b = 1$.

Note that $\Gamma_n$ lies entirely below the line $I\lambda$ if and only if $\|\Gamma_n/I\| \geq \lambda$ a.s. for each $n$. We can not make $\Gamma_n$ lie entirely above any line through the origin with positive slope since $\Gamma_n(t) = 0$ for $0 \leq t < \xi_{n1}$; however $\Gamma_n$ lies entirely above the line $I/\lambda$ on the interval $[\xi_{n1}, 1]$ if and only if $\|\Gamma_n/I\|_{\xi_{n1}} \leq \lambda$. Our main concern in this paper is bounding $\Gamma_n$ between straight lines through the origin. More precisely, we will characterize upper and lower class sequences for the random variables $\|\Gamma_n/I\|$ and $\|f/\Gamma_n\|_{\xi_{n1}}^k$.

"In probability upper and lower linear bounds" are well known (see Robbins (1954), Chang (1955) and Renyi (1973)); and see Shorack (1972) for applications. It is known that "a.s. linear bounds" do not exist (see Wellner (1977a)); also see Wellner (1977a) and (1977b) for applications of "a.s. nearly linear bounds."

Discussion of our theorems will be facilitated by contrasting them with the behavior of $\xi_{n1}$ and $\xi_{n2}$ that is set forth in Theorem 1.

THEOREM 1. Let $k \geq 1$ be a fixed integer.

(i) (Kiefer). If $c_n \searrow$, then

$$P(n^{\xi_{nk}} \leq c_n \text{ i.o.}) = 0$$

according as

$$\sum_{n=1}^{\infty} c_n^k n < \infty$$

$$= 1$$

$$= \infty.$$
(ii) (Robbins and Siegmund when \( k = 1 \)). Let \( c_n/n \downarrow \) and suppose either \( c_n \uparrow \) or \( \lim \inf_{n \to \infty} c_n / \log_2 n \geq 1 \). Then

\[
P(n^2 \xi_{nk} > c_n \text{ i.o.}) = 0 \quad \text{according as} \quad \sum_{n=1}^{\infty} \frac{c_n^k}{n} \exp(-c_n) < \infty
= 1 \quad = \infty.
\]

**Theorem 2.** Let \( n \lambda_n \uparrow \). Then

\[
P(||\Gamma_n/I|| \geq \lambda_n \text{ i.o.}) = 0 \quad \text{according as} \quad \sum_{n=1}^{\infty} \frac{1}{n \lambda_n} < \infty
= 1 \quad = \infty.
\]

Note that \( ||\Gamma_n/I|| = \max \{i/(n^2 \xi_{ni}) : 1 \leq i \leq n \} \) is \( \geq \lambda_n \) if \( n^2 \xi_{ni} \) is \( \leq c_n \equiv 1/\lambda_n \). Comparing the series criteria of Theorem 1(i) with Theorem 2, it is seen that small values of \( \xi_{ni} \) control large values of \( ||\Gamma_n/I|| \). Note however that \( ||\Gamma_n/I|| \) and \( (n^2 \xi_{ni})^{-1} \) have different limiting distributions.

Theorem 2 yields the known result \( \limsup_{n \to \infty} \log ||\Gamma_n/I||/\log_2 n = 1 \) a.s. In fact, \( \log \lambda_n = \sum_{i=2}^{n-1} \log n + \tau \log_2 n \), with \( p \geq 2 \), is upper class or lower class for \( \log ||\Gamma_n/I|| \) according as \( \tau > 1 \) or \( \tau \leq 1 \).

**Theorem 3.** Let \( \lambda_n/n \downarrow \) and suppose either \( \lambda_n \uparrow \) or \( \lim \inf_{n \to \infty} \lambda_n / \log_2 n \geq 1 \). Then

\[
P(||I/\Gamma_n||^2 \leq \lambda_n \text{ i.o.}) = 0 \quad \text{according as} \quad \sum_{n=1}^{\infty} \frac{\lambda_n^2}{n} \exp(-\lambda_n) < \infty
= 1 \quad = \infty.
\]

Note that \( ||I/\Gamma_n||^2 = \max \{n^2 \xi_{ni+1} / i : 1 \leq i \leq n \} \) is \( \geq \lambda_n \) if \( n \xi_{ni} \) is \( \geq c_n \equiv \lambda_n \). (Here, and in the following, \( \xi_{n+1,i} \equiv 1 \) for all \( n \).) Comparing the series criteria of Theorem 1(ii) with Theorem 3, it is seen that large values of \( \xi_{ni} \) control large values of \( ||I/\Gamma_n||^2 \). Note however (see Renyi (1973)) that \( ||I/\Gamma_n||^2 \) and \( n \xi_{ni} \) have different limiting distributions.

Theorem 3 yields the known result \( \limsup_{n \to \infty} ||I/\Gamma_n||^2 / \log_2 n = 1 \) a.s. In fact, \( \lambda_n = \log n + 3 \log_2 n + \sum_{i=4}^{n-1} \log n + \tau \log_2 n \), with \( p \geq 4 \), is upper class or lower class for \( ||I/\Gamma_n||^2 \) according as \( \tau > 1 \) or \( \tau \leq 1 \).

2. Proofs. Robbins (1954) showed that for any \( n \geq 1 \)

\[
P(||\Gamma_n/I|| \geq \lambda) = 1/\lambda \quad \text{for all} \quad \lambda > 1.
\]

**Proof of Theorem 2.** Suppose \( \sum (n \lambda_n)^{-1} < \infty \). Let \( n_k \equiv \text{int}(\alpha^k) \) where \( \alpha > 1 \) is fixed, and where \( \text{int}(\alpha^k) \) denotes that greatest integer function. Note that

\[
\sum_{k=2}^{\infty} \sum_{j=n_{k-1}+1}^{n_k} (n \lambda_n)^{-1} \geq \sum_{k=2}^{\infty} (n_k - n_{k-1})(n \lambda_n)^{-1} \geq \text{constant} \cdot \sum_{k=2}^{\infty} (n \lambda_n)^{-1}.
\]

Let \( A_k \equiv \max \{||\Gamma_n/I|| : n_k < n \leq n_{k+1} \} \geq \lambda_n \}; and note that monotonicity of \( n \Gamma_n \).
and \( n\lambda_n \) implies

\[
P(A_k) \leq P(n_{k+1} ||\Gamma_{n_{k+1}}/I|| \geq n_k \lambda_n) = n_{k+1}/(n_k \lambda_n) \quad \text{by} \quad (1)
\]

\[
\sim \alpha/\lambda_n
\]

so that (a) yields \( \sum_n^\infty P(A_n) < \infty \). Thus \( P(A_k \text{ i.o.)} = 0 \) by Borel–Cantelli; and hence \( P(||\Gamma_n/I|| \geq \lambda_n \text{ i.o.)} = 0. \)

Suppose \( \sum (n\lambda_n)^{-1} = \infty \). Now

\[
[||\Gamma_n/I|| \geq \lambda_n] = [\sup \{\sum_{t=1}^n I_{(0,t)}(\xi_t)/t : 0 < t \leq 1\} \geq n\lambda_n]
\]

\[
\sup \{I_{(0,t)}(\xi_t)/t : 0 < t \leq 1\} \geq n\lambda_n
\]

\[
= [\xi_n \leq (n\lambda_n)^{-1}]
\]

Now the events \([\xi_n \leq (n\lambda_n)^{-1}]\) are independent, and the sum of their probabilities equals \( \sum (n\lambda_n)^{-1} = \infty \). Thus Borel–Cantelli yields \( P(\xi_n \leq (n\lambda_n)^{-1} \text{ i.o.)} = 1 \); and hence \( P(||\Gamma_n/I|| \geq \lambda_n \text{ i.o.)} = 1. \)

Before proving Theorem 3, we need the following probability bound. For all \( n \geq 1 \) we have

\[
P(||I/I||_{\Gamma_n} \geq \lambda) \leq 16\lambda e^{-\lambda} \quad \text{for all} \quad \lambda > 1.
\]

The probability on the left-hand side of (2) is given in formula (17) on page 34 of Chang (1964); and for \( \lambda \geq 2 \) Chang’s next to the last formula on page 17 yields the bound \( 2(16e^{-\lambda})^k \), which when summed yields the right-hand side of (2). Note that (2) is trivial for \( \lambda \leq 1 \).

**Proof of Theorem 3.** Suppose \( \sum \lambda_i^{\alpha/n} [\exp (-\lambda_i) < \infty. \) Let \( n_j \equiv \text{int} (\exp (\alpha j/\log j)) \) for \( j \geq 2 \) with \( \alpha > 0 \) fixed. Let \( A_n \equiv [M_n \geq \lambda_n] \) where \( M_n \equiv ||I/I||_{\Gamma_n} \leq \max_{i \leq \lambda_n} (n\xi_{n+1,i}/i) \); and let \( B_j \equiv [M_{n_j} \geq (n_j/n_{j+1})\lambda_{n_{j+1}}] \). Note that

\[
M_n/n = \max_{i \leq \lambda_n} (\xi_{n+1,i}/i)
\]

is a \( \leq \) function of \( n \).

To see this, suppose \( \xi_{n+1,i} \) falls between \( \xi_{nk} \) and \( \xi_{nk+1} \); then

\[
0 \quad \xi_{n1} \quad \xi_{n2} \quad \cdots \quad \xi_{nk} \quad \uparrow \quad \xi_{n,k+1} \quad \cdots \quad \xi_{nn} \quad 1
\]

\[
\frac{\xi_{n+1,i}}{i} = \frac{\xi_{n+1,k+1}}{k} \leq \frac{\xi_{n+1,k+1}}{k} \quad \text{for} \quad 1 \leq i \leq k - 1
\]

\[
\frac{\xi_{n+1,i+1}}{i} = \frac{\xi_{n+1,k+1}}{k} \leq \frac{\xi_{n+1,k+1}}{k} \quad \text{for} \quad k + 1 \leq i \leq n
\]

so that (3) is established. From (3) and \( \lambda_n/n \leq \) we get

\[
\bigcup \{A_n : n_j \leq n < n_{j+1}\} \subseteq \bigcup \{[M_n/n \geq \lambda_{n_{j+1}}/n_{j+1}] : n_j \leq n < n_{j+1}\} = B_j.
\]

Thus to establish \( P(A_n \text{ i.o.)} = 0, \) it suffices to show \( \sum P(B_j) < \infty \). Let \( d_n \equiv \lambda_n \wedge 2 \log n \). Then

\[
d_n/n \searrow, \quad d_n \to \infty \quad \text{and} \quad \sum d_n (d_n^\alpha/n) \exp (-d_n) < \infty
\]
since \( d_n^3 \exp(-d_n) \leq \lambda_n^2 \exp(-\lambda_n) + (2 \log n)^2 \exp(-2 \log n) \). Since \( B_j \subset D_j \equiv \left\{ M_{n_j} \geq \frac{n_j}{n_{j+1}} \right\} \), it suffices to show that \( \sum_j P(D_j) < \infty \). Now

\[
\sum_{j=1}^\infty P(D_j) \leq \sum_{j=1}^\infty 16 \left( \frac{n_j}{n_{j+1}} \right) d_{n_j+1} \exp\left(-d_{n_j+1} \right) \exp\left( \left(1 - \frac{n_j}{n_{j+1}} \right) d_{n_j+1} \right)
\]

by (2)

\[
\leq (\text{Constant,}) \sum_{j=3}^\infty d_{n_j+1} \exp(-d_{n_j+1}) \quad \text{as in (2.45) of [6]}
\]

\(< \infty \quad \text{in complete analogy with Lemma 8 of [6] and using (b).}
\]

This completes the convergence half of the proof.

Suppose \( \sum n \exp(-\lambda_n) = \infty \). Note that \( \|I/\Gamma_n\|_{\nu_1} \geq n\xi_n \), and Theorem 1(ii) shows that \( P(n\xi_n \geq \lambda_n \text{ i.o.}) = 1 \). \( \square \)

**Remark.** Now \( \{n\xi_{n+i}/i : 1 \leq i \leq n\} \) is a reverse submartingale. This yields

\[
P\left( \|I/\Gamma_n\|_{\nu_1} \geq \lambda \right) \leq \inf_{r>0} E\left( \exp\left(rn\xi_n\right) \right)/\exp(r\lambda) \leq 14\lambda^2 \exp(-\lambda)
\]

for all \( \lambda > 1 \). This will only yield \( P\left( \|I/\Gamma_n\|_{\nu_1} \geq \lambda \text{ i.o.} \right) = 0 \) in Theorem 3 in case \( \sum n \exp(-\lambda_n) < \infty \).

**Proof of Theorem 1.** (i) See Kiefer (1972). (ii) See Robbins and Siegmund (1971) for the case \( k = 1 \). See Frankel (1976) for a statement of this result when \( k > 1 \) and \( \xi_n \nearrow \infty \); Frankel gives references to his 1972 thesis for a proof. It would appear that Frankel’s technique is similar to that of Wichura (1973); using diffusion processes and speed measure, Wichura establishes some results very closely related to the present ones.

The authors’ original version of this manuscript included a very long proof of Theorem 1(ii); it is available upon request. It uses only elementary techniques, and is a straightforward generalization of the proof of Robbins and Siegmund; the details are quite heavy. \( \square \)

3. **The case of arbitrary df’s.** Suppose \( X_{a1}, \ldots, X_{an} \) are independent with completely arbitrary df’s \( F_{a1}, \ldots, F_{an} \) on \( (-\infty, \infty) \). Let \( \bar{F}_n = n^{-1} \sum F_{at} \) denote the average df, and let \( F_n \) denote the empirical df of the observations.

**Theorem 4.** Let \( n\lambda_n \nearrow \infty \). Then \( \sum_{n=1}^\infty (n\lambda_n)^{-1} < \infty \) implies \( P\left( \|F_n/F_n\|_{\nu_1} \geq \lambda \text{ i.o.} \right) = 0 \).

**Proof.** By Theorem 1.1.1 and Corollary 1.3.1 of van Zuijlen (1976) we have

\[
P\left( \|F_n/F_n\|_{\nu_1} \geq \lambda \right) \leq 2\pi^2/3 \lambda \quad \text{for all } \lambda > 1.
\]

We can now recopy the proof of Theorem 2, except that an appeal to (4) replaces the appeal to (1). \( \square \)

We did not generalize Theorem 3 to the present case. It is possible to obtain an exponential bound in place of the bound in van Zuijlen’s equation (1.1.4) by applying a binomial exponential bound to the probability \( P(\sum_z, z > n - j + 1) \) of his proof. However, the resulting bound is not as strong as (2); and so we omit the resulting weak generalization of Theorem 3 that we can prove in the present case.
REFERENCES


DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195

DEPARTMENT OF STATISTICS
UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14627