The Annals of Probability
1992, Vol. 20, No. 3, 1213–1221

ONE-SIDED REFINEMENTS OF THE STRONG LAW OF LARGE NUMBERS AND THE GLIVENKO–CANTELLI THEOREM[1]

Dedicated to Lester Dubins on his seventieth birthday.

BY DAVID GILAT AND T. P. HILL

Tel Aviv University and Georgia Institute of Technology

A one-sided refinement of the strong law of large numbers is found for which the partial weighted sums not only converge almost surely to the expected value, but also the convergence is such that eventually the partial sums all exceed the expected value. The new weights are distribution-free, depending only on the relative ranks of the observations. A similar refinement of the Glivenko–Cantelli theorem is obtained, in which a new empirical distribution function not only has the usual uniformly almost-sure convergence property of the classical empirical distribution function, but also has the property that all its quantiles converge almost surely. A tool in the proofs is a strong law of large numbers for order statistics.

0. Introduction. The classical strong law of large numbers of Kolmogorov (1933) says that if $X_1, X_2, \dots$ are independent and identically distributed with finite mean μ, then the weighted partial sums $(n^{-1}X_1 + \cdots + n^{-1}X_n)$ converge almost surely to μ, and the recurrence of mean-zero random walks [e.g., Breiman (1968), Theorem 3.38] implies that for all nondegenerate distributions, these partial sums necessarily oscillate above and below μ infinitely often with probability 1.

One of the main objectives of this paper is to exhibit distribution-free weights $\alpha_{i,n}$, depending only on the relative ranks of the observations, so that the weighted partial sums $(\alpha_{1,n}X_1 + \cdots + \alpha_{n,n}X_n)$ not only converge to μ almost surely, but also are eventually always larger than or equal to μ almost surely. The proof uses (in addition to the classical strong law of large numbers, the Glivenko–Cantelli theorem and the law of the iterated logarithm) a limit theorem of Feller and a strong law of large numbers for order statistics.

Similarly, the classical empirical distribution function F_n converges uniformly to the underlying distribution function F almost surely, but in an oscillatory fashion, which implies that many parameters of F_n such as quantiles may not converge almost surely (e.g., medians for fair-coin tossing). A new empirical distribution function G_n is found which not only converges to F

Received June 1990; revised April 1991.

[1]Research partly supported by Israel–U.S. Binational Science Foundation Grant 88-00005 (for both authors) and NSF Grant DMS-89-01267 for Hill.

AMS 1980 *subject classifications.* Primary 60F15; secondary 62G30.

Key words and phrases. Strong law of large numbers, Glivenko–Cantelli theorem, order statistics, one-sided strong laws, convergence of medians.

uniformly almost surely, but also converges eventually from below with probability 1, so that quantiles and, in the case of measures with at least one atom, modes all converge almost surely.

This paper is organized as follows: Section 1 contains a strong law of large numbers for order statistics, Section 2 the one-sided refinement of the strong law of large numbers and Section 3 the one-sided refinement of the Glivenko–Cantelli theorem.

Throughout this paper $X, X_1, X_2, \dots$ will denote a sequence of independent and identically distributed random variables on a probability space $(\Omega, \mathscr{F}, P)$, F their common distribution function, and EX the expected value of X. For each positive integer n, $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ denote the *order statistics* associated with the first n observations $X_1, \dots, X_n$, that is, $X_{i:n} = X_{\pi_n(i)}$, where $\pi_n = \pi_{n,\omega}$ is any permutation of $\{1, \dots, n\}$ satisfying $X_{\pi_n(1)} \le X_{\pi_n(2)} \le \cdots \le X_{\pi_n(n)}$. Similarly, $R_n(X_i)$ is the *rank* of X_i among $X_1, \dots, X_n$ for $1 \le i \le n$, that is, $R_n(X_i) = \pi_n^{-1}(i)$, where again the dependence on ω is suppressed.

For a Borel probability measure on $\mathbb{R}$, the same symbol will be used to denote the (right-continuous) c.d.f., and the measure itself, so $G(A)$ is the G-probability of the set A, $G(x)$ the probability of the set $(-\infty, x]$ and $G(\{x\})$ the probability of the singleton $\{x\}$. For real numbers x and y, $\delta(x)$ denotes the Dirac measure or point-mass at x and $x \vee y$ denotes the maximum of x and y.

1. A Strong law of large numbers for order statistics.

The main result in this section, Theorem 1.1 below, is a strong law for order statistics which includes Kolmogorov's strong law of large numbers as a special case. Although Theorem 1.1 will be used here as a tool in the one-sided strong law in Section 2, it may be of some independent interest in itself; neither the version of Helmers (1977) nor its successor in van Zwet (1980) identifies the limit explicitly.

THEOREM 1.1. *If $X_1, X_2, \dots$ are i.i.d. with $E|X_1| < \infty$, then for each $k = 0, 1, 2, \dots$,*

$$\lim_{n\to\infty} \frac{k+1}{n^{k+1}} \sum_{i=1}^{n} i^k X_{i:n} = E(X_1 \vee \cdots \vee X_{k+1}) \quad a.s.$$

(Note that since $\sum_{i=1}^{n} X_{i:n} = \sum_{i=1}^{n} X_i$, the case $k = 0$ is precisely Kolmogorov's SLLN.)

PROOF. Fix k and observe that $E|X_1| < \infty$ implies $E|X_1 \vee \cdots \vee X_{k+1}| < \infty$. By Kolmogorov's SLLN,

$$(1) \quad \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i F^k(X_i) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_{i:n} F^k(X_{i:n}) = E\left[XF^k(X)\right] \quad \text{a.s;}$$

$$(2) \quad \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} |X_i| = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} |X_{i:n}| = E|X| \quad \text{a.s.}$$

Suppose first that F is absolutely continuous, in which case the Radon–Nikodym derivative $dF^{k+1}(x)/dx = (k+1)F^k(x)\,dF(x)/dx$ F-a.e. (where dx is Lebesgue measure), so

$$E\left[XF^k(X)\right] = \int_{\mathbb{R}} xF^k(x)\,dF(x) = \frac{1}{k+1}\int_{\mathbb{R}} x\,dF^{k+1}(x) = \frac{1}{k+1}E[X_1 \vee \cdots \vee X_{k+1}] \tag{3}$$

and, since $P(X_i = X_j \text{ for some } i \neq j) = 0$,

$$F_{n,\omega}(X_{i:n}(\omega)) = \frac{i}{n} \quad \text{a.s. for all } n \text{ and all } i = 1,\dots,n. \tag{4}$$

Let $\Delta_n = (1/n)\sum_{i=1}^n (i/n)^k X_{i:n} - (1/n)\sum_{i=1}^n X_{i:n}F^k(X_{i:n})$. By (1) and (3), it suffices to show that $\Delta_n \to 0$ a.s. By (4),

$$|\Delta_n| \le \frac{1}{n}\sum_{i=1}^n \left|\left(\frac{i}{n}\right)^k - F^k(X_{i:n})\right| |X_{i:n}| = \frac{1}{n}\sum_{i=1}^n \left|F_n^k(X_{i:n}) - F^k(X_{i:n})\right| |X_{i:n}| \quad \text{a.s.} \tag{5}$$

Now the uniform continuity of x^k on $[0,1]$, the finiteness of $E|X|$, (2) and the Glivenko–Cantelli theorem [e.g., Billingsley (1986), page 275],

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0 \quad \text{a.s.}, \tag{6}$$

together imply that the right-hand side of (5) goes to 0 almost surely, which completes the argument in the case F is absolutely continuous.

If F is not absolutely continuous [in which case both (3) and (4) may fail], the argument is to define new $\{\tilde{X}_i\}$ which are i.i.d. and continuous, which are uniformly close to $\{X_i\}$ and whose order statistics are uniformly close to those of the $\{X_i\}$ everywhere.

Fix $\varepsilon > 0$, and enlarging the probability space if necessary, let $\{Y_j\}_1^\infty$ be independent and independent of the $\{X_j\}$, with Y_j uniformly distributed on $[0, \varepsilon]$ for all j. Define $\{\tilde{X}_j\}_1^\infty$ by

$$\tilde{X}_j = X_j + Y_j.$$

Clearly the $\{\tilde{X}_j\}$ are i.i.d. and continuous and satisfy

$$|X_n - \tilde{X}_n| \le \varepsilon \quad \text{and} \quad |X_{i:n} - \tilde{X}_{i:n}| \le \varepsilon \quad \text{for all } \omega, \text{ all } n \text{ and all } i \le n. \tag{7}$$

By the continuous case result,

$$\frac{k+1}{n^{k+1}}\sum_{i=1}^n i^k \tilde{X}_{i:n} \to E\left[\tilde{X}_1 \vee \cdots \vee \tilde{X}_{k+1}\right] \quad \text{a.s.}, \tag{8}$$

so

$$\limsup_{n\to\infty}\left|\frac{k+1}{n^{k+1}}\sum_{i=1}^{n} i^k X_{i:n} - E[X_1 \vee \cdots \vee X_{k+1}]\right|$$

$$\le \limsup_{n\to\infty}\left[\frac{k+1}{n^{k+1}}\sum_{i=1}^{n} i^k |X_{i:n} - \tilde{X}_{i:n}| + \frac{k+1}{n^{k+1}}\left|\sum_{i=1}^{n} i^k \tilde{X}_{i:n} - E\left[\tilde{X}_1 \vee \cdots \vee \tilde{X}_{k+1}\right]\right| + \left|E[X_1 \vee \cdots \vee X_{k+1}] - E\left[\tilde{X}_1 \vee \cdots \vee \tilde{X}_{k+1}\right]\right|\right] \le 2\varepsilon \quad \text{a.s.},$$

where the last inequality follows by (7) and (8) and the fact that $((k+1)/n^{k+1})\sum_{i=1}^{n} i^k \to 1$ as $n \to \infty$. Since $\varepsilon > 0$ was arbitrary, this completes the proof. □

REMARKS. It follows easily from (3) and the Weierstrass approximation via a similar argument that the conclusion of Theorem 1.1 may be extended from $g(x) = x^k$ to general continuous functions g on [0, 1], in which case the result is $\lim_{n\to\infty}(1/n)\sum_{i=1}^{n} g(i/n)X_{i:n} = \int_0^1 F^{-1}(u)g(u)\,du$ (cf. Stigler (1974) in the weak-convergence setting). Note that if the underlying distribution F is nonatomic (continuous), then $\int_0^1 F^{-1}(u)g(u)\,du = E[Xg(F(X))]$. It can be seen that for each ω, $((k+1)/n^{k+1})\sum_{i=1}^{n} i^k X_{i:n}(\omega)$ is the bootstrap expected value of the maximum of $k+1$ independent observations from the empirical distribution $F_{n,\omega}$. Using this fact, an alternative proof of Theorem 1.1 can be given.

2. A one-sided refinement of the strong law of large numbers.

DEFINITION. Real numbers $r_1, r_2, \ldots$ *converge to r eventually from above* (written $r_n \to_+ r$) if (i) $\lim_{n\to\infty} r_n = r$; and (ii) there exists an integer N so $r_n \ge r$ for all $n \ge N$. Similarly, $r_n \to_{(-)} r$ if $-r_n \to_+ -r$.

REMARK. As pointed out by Gideon Schwartz, convergence eventually from above is equivalent to convergence in the Sorgenfrey or right-half-open-interval topology, which is the topology on the real line generated by half-open intervals of the form $[a, b)$ [cf. Steen and Seebach (1978)].

The following lemma, which is needed for part of the proof of the main result of this section, was communicated to us by Michael Klass.

LEMMA 2.1. *If* $E|X|^{1+\gamma} < \infty$ *for some* $0 < \gamma < 1$ *and* $\{S_n\}_{n=1}^{\infty}$ *is the sequence of partial sums of independent random variables distributed like* X, *then*

$$S_n - nEX = o(n^{1/(1+\gamma)}) \quad a.s.$$

PROOF. Since $\sum_{n=1}^{\infty} P(|X| \geq n^{1/(1+\gamma)}) < \infty$, the conclusion follows easily from Theorem 1 of Feller (1946) using the fact that if $E|X| < \infty$, then there is always a nonnegative increasing function ϕ on $[0, \infty)$ such that $\lim_{x \to \infty}(\phi(x)/x) = +\infty$ and $E(\phi(|X|)) < \infty$. □

The next theorem is the main result of this section.

THEOREM 2.2. *If $X, X_1, X_2, \ldots$ are i.i.d. with $E|X|^{1+\gamma} < \infty$ for some $\gamma > 0$, then for all α satisfying $2 < \alpha < \min\{2 + (\gamma/(1+\gamma)), \frac{5}{2}\}$,*

$$\sum_{i=1}^{n} \left(\frac{1}{n} - \frac{n+1}{2n^{\alpha}} + \frac{R_n(X_i)}{n^{\alpha}} \right) X_i \to_{+} EX \quad a.s., \tag{9}$$

or equivalently, in terms of order statistics,

$$\sum_{i=1}^{n} \left(\frac{1}{n} - \frac{(n+1)}{2n^{\alpha}} + \frac{i}{n^{\alpha}} \right) X_{i:n} \to_{+} EX \quad a.s. \tag{9'}$$

Note that for each n, the weights

$$\left\{ \frac{1}{n} - \frac{n+1}{2n^{\alpha}} + \frac{i}{n^{\alpha}} \right\}$$

form an arithmetic progression of strictly positive numbers with average value $1/n$ and sum 1, and assign the largest observations the most weight. Reversing the order of the weights will result in convergence from below.

PROOF. If X is degenerate, the theorem is trivial, so suppose X is nondegenerate. To simplify notation, set $\mu = EX$, $M = E(X_1 \vee X_2)$ and

$$\hat{S}_n = \sum_{i=1}^{n} \left(\frac{1}{n} - \frac{n+1}{2n^{\alpha}} + \frac{i}{n^{\alpha}} \right) X_{i:n}.$$

THE CASE $\gamma \geq 1$. Since $E|X|^{1+\gamma} < \infty \Rightarrow E|X|^2 < \infty$ for $\gamma \geq 1$, in this case $2 < \alpha < \frac{5}{2}$. Since X has finite variance, the classical law of the iterated logarithm [Khintchine (1924)] implies

$$\frac{1}{n} \sum_{i=1}^{n} X_{i:n} = \frac{1}{n} \sum_{i=1}^{n} X_i = \mu + o\left(\sqrt{\frac{\log n}{n}} \right) \quad \text{a.s.}, \tag{10}$$

so by Theorem 1.1 (for $k = 1$) and (10),

$$\begin{aligned} \hat{S}_n &= \mu + o\left(\sqrt{\frac{\log n}{n}} \right) + \frac{1}{2n^{\alpha-2}}(M - \mu + o(1)) \\ &= \mu + \frac{1}{2n^{\alpha-2}}(M - \mu) + o\left(\frac{1}{n^{\alpha-2}} \right) \quad \text{a.s.} \left(\text{since } 0 < \alpha - 2 < \frac{1}{2} \right). \end{aligned}$$

Now, X is nondegenerate if and only if $M - \mu > 0$, so for $0 < \alpha - 2 < 1/2$, $\hat{S}_n \to_+ \mu$ almost surely.

THE CASE $0 < \gamma < 1$. Recall that in this case, $2 < \alpha < 2 + (\gamma/(1+\gamma))$. By Lemma 2.1,

$$\frac{1}{n}\sum_{i=1}^{n} X_{i:n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \mu + o(n^{-\gamma/(1+\gamma)}),$$

so the conclusion follows essentially as in the case $\gamma \geq 1$, using the fact that $\alpha - 2 < \gamma/(1+\gamma)$. □

REMARKS. Other weights may also give one-sided convergence to the mean; this particular version was selected because it is a fair martingale-like arithmetic sequence, in the sense that, given the past, the weight assigned to the next observation will have expected value $1/n$ (so its expected *difference* from the classical weight is zero). As is seen in the proof, the interval for α for which the one-sided convergence holds is sharp. For all nondegenerate distributions, if the weights are too large ($\alpha \leq 2$), the term $\Sigma_{i=1}^{n}((n+1)/2n^{\alpha})X_i$ diverges a.s. for $\alpha < 2$, whereas for $\alpha = 2$, $\hat{S}_n$ converges a.s. to the wrong constant, namely $(M+\mu)/2$. On the other hand, for weights too close to $1/n$ (i.e., $\alpha \geq 5/2$), the desired one-sidedness is not attained because the adjusted weights are too small to override the (law of the iterated logarithm) oscillations.

3. A one-sided refinement of the Glivenko–Cantelli theorem. For a Borel probability measure G on $\mathbb{R}$, the (upper) β-quantile of G, or its generalized (upper) inverse, G^{-1}, $Q_\beta(G)$, is defined to be $Q_\beta(G) = G^{-1}(\beta) = \sup\{x \in \mathbb{R}: G(x) \leq \beta\}$ and this supremum is clearly attained for all $\beta \in (0,1)$. As was noted in the Introduction, the classical empirical distribution F_n for $X_1, X_2, \dots, X_n$, given by

$$F_n = \sum_{i=1}^{n} \frac{1}{n}\delta(X_i), \tag{11}$$

does converge as a c.d.f. uniformly to F almost surely [see equation (6)], but $Q_\beta(F_n)$ may not converge almost surely, even if F is continuous, as the next example shows.

EXAMPLE 3.1. Let $X_1, X_2, \dots$ be i.i.d. uniformly distributed on $[-2,-1] \cup [1,2]$. Then, clearly

$$\liminf_{n\to\infty} Q_{1/2}(F_n) = -1 < 1 = \limsup_{n\to\infty} Q_{1/2}(F_n) \quad \text{a.s.}$$

For exactly the same reason, the (upper) mode of the empirical c.d.f. for fair coin tossing also diverges almost surely.

The main purpose of this section is to exhibit a new empirical c.d.f. G_n which not only converges uniformly to F almost surely, but which also has all quantiles converging almost surely to the corresponding quantile of F and, if F has any atoms, has the modes of $\{F_n\}$ converging almost surely to the mode of F. In addition, the usual strong law of large numbers holds in that if F has finite mean, then the means of $\{G_n\}$ converge almost surely to the mean of F.

DEFINITION. If $X_1, X_2, \ldots, X_n$ are i.i.d., then for each $\alpha > 2$, the *refined* (upper) *empirical distribution* G_n^α is the random (depending on ω) probability measure

$$G_n^\alpha = \sum_{i=1}^{n} \left(\frac{1}{n} - \frac{(n+1)}{2n^\alpha} + \frac{i}{n^\alpha} \right) \delta(X_{i:n}). \tag{12}$$

Observe again that the weights attached to $\delta(X_i)$ are *distribution free*, as they depend only on the relative ranks of X_i and that G_n^α is a probability measure since the weights are strictly positive and sum to 1.

The next proposition records some immediate but useful comparisons between the classical empirical c.d.f. F_n (11) and the refined empirical c.d.f. G_n (12); here supp(G) is the support of G and $\|m\|$ is the total variation of the signed measure m.

PROPOSITION 3.2. *For all* $\alpha > 2$, $G_n = G_{n,\omega}^\alpha$ *satisfies*:

(i) $\mathrm{supp}(G_n) = \mathrm{supp}(F_n)$ *for all* ω;

(ii) $G_n(x) \le F_n(x)$ *for all* x *and all* ω (*i.e.*, G_n *stochastically dominates* F_n *everywhere*);

(iii)

$$\|G_n - F_n\| = \sum_{i=1}^{n} \left| \frac{n+1}{2n^\alpha} - \frac{i}{n^\alpha} \right| \le \frac{1}{4n^{\alpha-2}} \quad \textit{for all } \omega;$$

(iv) $Q_\beta(G_n) \ge Q_\beta(F_n)$ *for all* β *and all* ω.

PROOF. Conclusion (i) follows since the weights are strictly positive, (iii) follows immediately from the definitions of F_n and G_n and (iv) follows easily from (ii). To see (ii), note that by (11) and (12), for every x and ω,

$$\begin{aligned} G_n(x) &= \sum_{i=1}^{nF_n(x)} \left(\frac{1}{n} - \frac{n+1}{2n^\alpha} + \frac{i}{n^\alpha} \right) \\ &= F_n(x) - \frac{1}{2n^{\alpha-2}} F_n(x)(1 - F_n(x)) \le F_n(x). \end{aligned}$$ □

The next theorem is the main result in this section; recall that for a distribution G having at least one atom, the (upper) mode of G, mode(G), is $\max\{x \in \mathbb{R}: G(\{x\}) \ge G(\{y\}) \text{ for all } y \in \mathbb{R}\}$.

THEOREM 3.3. *Let* $X_1, X_2, \ldots$ *be i.i.d. with distribution* F. *Then for all* $\alpha \in (2, \frac{5}{2})$, *the refined empirical distribution functions* $\{G_n\} = \{G_{n,\omega}^\alpha\}$ *satisfy*:

(i) $\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |G_n(x) - F(x)| = 0$ *a.s.*;

(ii) $G_n(x) \to_{(-)} F(x)$ *a.s. for each* $x \in \mathbb{R}$;

(iii) $Q_\beta(G_n) \to Q_\beta(F)$ *a.s. for all* $\beta \in [0, 1]$;

(iv) *if* F *has at least one atom, then* $\mathrm{mode}(G_n) \to \mathrm{mode}(F)$ *a.s.*;

(v) *if* F *has finite mean, then* $\mathrm{mean}(G_n) \to \mathrm{mean}(F)$ *a.s.*

[Recall that by Theorem 2.2, if F has slightly more than finite mean, then for the appropriate α the convergence of means in (v) is even one-sided.]

PROOF OF THEOREM 3.3. (i) The conclusion follows from the Glivenko–Cantelli theorem and Proposition 3.2(iii), since $\alpha > 2$.

(ii) If $F(x) = 0$ or 1, the conclusion is trivial. If $0 < F(x) < 1$, the conclusion follows from Theorem 2.2 with $\{X_i\}$ replaced by $\{\hat{X}_i\}$, where $\hat{X}_i = I(X_i > x)$, $i = 1, 2, \dots,$ since the $\{\hat{X}_i\}$ are i.i.d. with finite mean and variance, so one-sided convergence holds a.s. for all $\alpha \in (2, \frac{5}{2})$.

(iii) If $\beta = 0$ or 1, the conclusion is easy. If $\beta \in (0, 1)$, the conclusion follows easily from (ii), since $Q_\beta(F)$ is finite. [Note that if $\beta < 1$, the convergence in (iii) is even eventually from above.]

(iv) Suppose F has at least one atom. If F has a unique mode, then by the Glivenko–Cantelli theorem, $\text{mode}(F_n) \to \text{mode}(F)$ a.s., so by Proposition 3.2(iii), $\text{mode}(G_n) \to \text{mode}(F)$ a.s. On the other hand, if the mode of F is not unique, it will suffice to show that if $F(\{x_1\}) = F(\{x_2\}) = p > 0$ for some $x_1 < x_2$, then

$$G_n(\{x_2\}) - G_n(\{x_1\}) \to_+ 0 \quad \text{a.s.} \tag{13}$$

To see (13), for each $i = 1, 2$, let $s_i = nF_n(\{x_i\})$ (= number of times value x_i occurs in $X_1, \dots, X_n$) and $N_i = nF_n(x_i)$ (= number of times values $\le x_i$ occur in $X_1, \dots, X_n$). Then

$$\begin{aligned} G_n(\{x_2\}) - G_n(\{x_1\}) &= \left(\frac{1}{n} - \frac{n+1}{2n^\alpha}\right)(s_2 - s_1) \\ &\quad + \frac{1}{n^\alpha}\left[\sum_{i=N_2-s_2+1}^{N_2} i - \sum_{i=N_1-s_1+1}^{N_1} i\right] \\ &\ge \left(\frac{1}{n} - \frac{n+1}{2n^\alpha}\right)\left(o\left(\sqrt{n \log n}\,\right)\right) \\ &\quad + \frac{1}{n^\alpha}\left[\sum_{i=N_1+1}^{N_1+s_2} i - \sum_{i=N_1-s_1+1}^{N_1} i\right] \\ &= o\left(\sqrt{\frac{\log n}{n}}\right) + \frac{1}{n^\alpha}\left[n^2p^2 + o\left(n\sqrt{n \log n}\,\right)\right] \\ &= \frac{p^2}{n^{\alpha-2}} + o\left(\frac{1}{n^{\alpha-2}}\right) \to_+ 0, \end{aligned}$$

where the first equality follows by the definitions of G_n, s_i and N_i, the inequality by (10) and the fact that $N_2 - s_2 \ge N_1$, the second equality by (10), the last equality since $\alpha \in (2, 5/2)$ and the one-sided convergence since $p > 0$. This establishes (13), and hence conclusion (iv).

(v) By the definition of G_n,

$$\text{mean}(G_n) = \sum_{i=1}^{n}\left(\frac{1}{n} - \frac{(n+1)}{2n^\alpha} + \frac{i}{n^\alpha}\right)X_{i:n}.$$

Since $\sum_{i=1}^{n} X_{i:n} = \sum_{i=1}^{n} X_i$, it follows from Kolmogorov's SLLN that $(1/n)\sum_{i=1}^{n} X_{i:n} \to EX = \text{mean}(F)$ a.s. and that $\sum_{i=1}^{n}((n+1)/2n^{\alpha})X_{i:n} \to 0$ a.s. since $\alpha > 2$. By Theorem 1.1 with $k = 1$,

$$\frac{2}{n^2}\sum_{i=1}^{n} iX_{i:n} \to E(X_1 \vee X_2) \quad \text{a.s.},$$

and recalling that the finiteness of EX_1 implies that of $E(X_1 \vee X_2)$, $\sum_{i=1}^{n}(i/n^{\alpha})X_{i:n} \to 0$ a.s. since $\alpha > 2$. □

In general, the one-sided convergence of G_n^{α} to F [conclusion (ii) in Theorem 3.3] does not hold uniformly, even for continuous F, as the next example shows, but does hold F-almost uniformly as can be seen from the proof.

EXAMPLE 3.4. Let $X_1, X_2, \ldots$ be i.i.d. uniform on (0, 1). Then for *all* $\omega \in \Omega$ and all $n \geq 1$, $\operatorname{ess\,sup} G_n^{\alpha} < 1$, so $G_n^{\alpha}(d) = 1 > F(d) = d$ for some d sufficiently close to 1, so $\{\omega: G_{n,\omega}(x) \to_{(-)} (F(x))$ for all $x\} = \varnothing$.

As the notion of mode for nonatomic distributions is less clear, and since the empirical c.d.f.'s (both classical and refined) are purely atomic, there does not appear to be a natural extension of the mode conclusion [(iv) in Theorem 3.3] to general distributions.

Acknowledgments. The authors are grateful to the referee for several helpful suggestions, to Professor P. Bickel for pointing out some of the order statistics references, to the Georgia Institute of Technology for the invitation to the first author to visit while this paper was written and especially to Professor Michael Klass for pointing them to Lemma 2.1.

REFERENCES

BILLINGSLEY, P. (1986). *Probability and Measure*, 2nd ed. Wiley, New York.

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, Mass.

FELLER, W. (1946). A limit theorem for random variables with finite moments. *Amer. J. Math.* **68** 257–262.

HELMERS, R. (1977). A strong law of large numbers for linear combinations of order statistics. Technical Report SW 50/77, Math. Centrum, Amsterdam.

KHINTCHINE, A. (1924). Über einen Satz der Wahrscheinlichkeitsrechnung. *Fund. Math.* **6** 9–20.

KOLMOGOROV, A. (1933). *Grundbegriffe der Wahrscheinlichheitsrechnung*. Springer, Berlin.

STEEN, L. and SEEBACH, J. (1978). *Counterexamples in Topology*. Springer, New York.

STIGLER, S. (1974). Linear functions of order statistics with smooth weight functions. *Ann. Statist.* **2** 676–693.

VAN ZWET, W. (1980). A strong law for linear functions of order statistics. *Ann. Probab.* **8** 986–990.

SCHOOL OF MATHEMATICAL SCIENCES
TEL AVIV UNIVERSITY
TEL AVIV 69978
ISRAEL

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332