A GLIVENKO-CANTELLI THEOREM AND STRONG LAWS FOR L-STATISTICS

by

Arup Bose
Indian Statistical Institute and Purdue University

Technical Report #96-4

Department of Statistics
Purdue University
West Lafayette, IN USA

March 1996

A GLIVENKO-CANTELLI THEOREM AND STRONG LAWS FOR L-STATISTICS

by

Arup Bose
Indian Statistical Institute and Purdue University

Abstract

We prove a weighted Glivenko-Cantelli theorem and apply it to study the rate of convergence in the strong law for L-statistics.

Key Words and Phrases: Glivenko-Cantelli theorem, strong law, law of iterated logarithm, L-statistics.

AMS 1991 Subject Classification: 60F15, 62G30.

1. Introduction.

Let $(X_i, i \geq 1)$ be a sequence of independent and identically distributed (i.i.d.) random variables with $E|X_1|^p < \infty$ for some $0 < p < 2$. The Marcinkiewicz-Zygmund strong law (MZSLLN) states that as $n \to \infty$,

$$n^{-1} \sum_{i=1}^{n} X_i = c + o(n^{(1-p)/p}) \text{ almost surely } \ldots \ldots \ldots \quad (1)$$

Here $c = EX_1$ if $p \geq 1$ and c may be taken to be 0 if $p < 1$. For a proof of this result, see Chow and Teicher (1978, page 122). The main motivation for this work was to establish this law for L-statistics.

Wellner (1977a,b) proved a SLLN and LIL for L-statistics by first establishing appropriate convergence results for empirical processes. When specialized to the sample mean, his results yield the SLLN when $E|X_1|^{1+\varepsilon} < \infty$ for some $\varepsilon > 0$ and the LIL when $E|X_1|^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$, thereby coming close to the minimal possible condition, namely $\varepsilon = 0$.

We take a similar approach and first establish a suitable Glivenko-Cantelli type theorem. This is then used to obtain the MZSLLN for L-statistics. When specialized to the sample mean for $1 < p < 2$, (1) holds under $E|X_1|^{p+\varepsilon} < \infty$ for some $\varepsilon > 0$, thereby coming close to the minimal condition.

When certain smoothness conditions are allowed on the weight functions, we take a more direct approach of using projections and establish the MZSLLN for L-statistics.

2. A Glivenko-Cantelli Theorem.

Let $(\xi_i, i \geq 1)$ be i.i.d. uniform (0,1) random variables. Let I denote the identity function and let Γ_n denote the empirical distribution of $\xi_1, \ldots, \xi_n$ so that $\Gamma_n(t) = n^{-1} \sum_{i=1}^{n} I_{[0,t]}(\xi_i)$. Let $\mathcal{H}^+$ be the set of all nonnegative, nondecreasing continuous functions on [0,1]. Let $\mathcal{H}$ denote the set of all functions h such that $h(t) = h(1-t) = \bar{h}(t)$ for $0 \leq t \leq 1/2$ and $\bar{h} \in \mathcal{H}^+$. For $h \in \mathcal{H}^+ \cup \mathcal{H}$, define

$$\rho_h(\Gamma_n, I) = \sup_{0 \leq t \leq 1} \frac{|\Gamma_n - I|}{h}$$

THEOREM 1. Suppose $h \in \mathcal{H}^+ \cup \mathcal{H}$ and $(\theta_i, i \geq 1)$ is a decreasing sequence of numbers between 0 and 1 such that

$$\text{(A1)} \qquad \sup_{0 \leq t \leq \theta_n} [t(h(t))^{-1}] = o(n^{(1-p)/p})$$

$$\text{(A2)} \qquad n^{-1/2}(\log\log\ n)^{1/2} \sup_{\theta_n \leq t} [w(t)h(t)]^{-1} = o(n^{(1-p)/p}) \text{ for some } \delta > 0.$$

where w is a positive real-valued function on [0,1] such that for some $0 < \delta \leq 1/2$, $t^{1/2}w(t)$ is monotone increasing on $(0, \delta]$, $(1-t)^{1/2}w(t)$ is monotone decreasing on $[1-\delta, 1)$, is bounded on $[\delta, 1-\delta]$ and

$$\int_0^1 w^2(t) \log\ \log(t(1-t))dt < \infty \ldots\ \ldots \qquad (2)$$

$$\text{(A3)} \qquad \sum_{i=1}^{\infty} i^{-1/p} \int_0^{\theta_i} (h(t))^{-1} dt < \infty \ldots\ \ldots\ \ldots \qquad (3)$$

or

$$\sum_{i=1}^{\infty} i^{-q_i/p} \int_0^{\theta_i} (h(t))^{-q_i} dt < \infty$$

for some $0 < q_i < 1$, $i \geq 1$. Then as $n \to \infty$,

$$\rho_h(\Gamma_n, I) = o(n^{(1-p)/p}) \text{ almost surely } \ldots\ \ldots \qquad (4)$$

□

Proof. By "symmetry", we may assume $h \in \mathcal{H}^+$.

$$\begin{aligned} \rho_h(\Gamma_n, I) &\leq \sup_{0 \leq t \leq \theta_n} \Gamma_n/h + \sup_{0 \leq t \leq \theta_n} t/h \\ &\quad + \sup_{\theta_n \leq t} (w(t)|\Gamma_n(t) - t|)\,[w(t)h(t)]^{-1} \\ &= T_1 + T_2 + T_3 \text{ say.} \end{aligned}$$

By Assumption (A1), $T_2 = o(n^{(1-p)/p})$. By a LIL for empirical processes (see for example James (1975, Theorem, page 770), the first factor of T_3 is $O(n^{-1/2}(\log\log\ n)^{1/2})$ almost surely. Hence by Assumption (A2), $T_3 = o(n^{(1-p)/p})$ almost surely. To tackle T_1, observe that

$$\begin{aligned} T_1 &\leq n^{-1} \sum_{i=1}^{n} I_{[0,\theta_n]}(\xi_i)/h(\xi_i) \\ &\leq n^{-1} \sum_{i=1}^{n} I_{[0,\theta_i]}(\xi_i)/h(\xi_i). \end{aligned}$$

Define $Y_i = i^{-1/p} I_{[0,\theta_i)}(\xi_i)/h(\xi_i)$. Observe that $(Y_i, i \geq 1)$ are independent and by (A3) $\Sigma E Y_i < \infty$ or $\Sigma E Y_i^{q_i} < \infty$. In either case, ΣY_i converges. (See for example Teicher (1978, page 114, Corollary 3)). By Kronecker's Lemma, $n^{-1/p} \sum_{i=1}^{n} I_{[0,\theta_i]} (\xi_i)/h(\xi_i) = o(1)$ almost surely and hence $T_1 = o(n^{(1-p)/p})$ almost surely, proving the theorem. □

Example 1. Theorem 1 holds with $h(t) = t^\alpha$, where $0 \leq \alpha < \min(1/p, 1)$. To see this assume that $1/2 < \alpha < \min(1/p, 1)$, let $\theta_i = i^{-\beta}$ and choose $w(t) = t^{-(1/2-\delta)}$, δ very small. Then

$$\begin{aligned} &\sup_{0 \leq t \leq \theta_n} [t(h(t))^{-1}] = i^{-\beta(1-\alpha)} = o(n^{(1-p)/p}) \text{ if} \\ &-\beta(1-\alpha) < (1-p)/p \\ &\text{or } \beta > (p-1)/p(1-\alpha) \ldots \ \ldots \ \ldots \end{aligned} \tag{5}$$

(A2) is satisfied if

$$-1/2 - \beta(1/2 - \alpha) < (1-p)/p$$

or

$$\beta < (p-2)/2p(1-2\alpha) \ldots \ \ldots \tag{6}$$

Note that $(p-1)/p(1-\alpha) < (p-2)/2p(1-2\alpha)$, hence a choice of $\beta > 0$ satisfying (5) and (6) is possible. With such a choice of β,

$$\begin{aligned} &\sum_{i=1}^{\infty} i^{-1/p} \int_0^{\theta_i} (h(t))^{-1} dt \\ &\leq c \sum_{i=1}^{\infty} i^{-1/p} i^{-\beta(1-\alpha)} < \infty \end{aligned}$$

since $-1/p - \beta(1-\alpha) < -1/p + (1-p)/p = -1$. Thus conditions (A1), (A2) and (A3) hold. If $0 \leq \alpha \leq 1/2$, then $\rho_h(\Gamma_n, I) = O(n^{-1/2}(\log\ \log\ n)^{1/2})$ almost surely by LIL of James (1975). So Theorem 1 holds trivially. □

3. Strong Laws for L-Statistics.

Let $\mathcal{G}$ be the set of left continuous functions on (0, 1) which are of bounded variation on $(\theta, 1-\theta)$ for any $\theta > 0$. Let $c_{n1}, \ldots, c_{nn}, (n \geq 1)$ be constants and $g_n \in \mathcal{G}$, $n \geq 1$. Define the sequence of L-statistics as

$$L_n = n^{-1} \sum_{i=1}^{n} c_{ni} g_n(\xi_{ni}) \ldots \ldots \tag{7}$$

Note that if $g_n(x) = g(x) = T(F^{-1}(x))$ for some distribution function F, then $(L_n, n \geq 1)$ has the same distribution as $(M_n, n \geq 1)$ where

$$M_n = n^{-1} \sum_{i=1}^{n} T(X_{ni}) \ldots \ldots \ldots \tag{8}$$

Here $X_{n1} \leq \cdots \leq X_{nn}$ is the ordered statistics of a sample of size n from F. Define $J_n(t) = c_{ni}$ if $(i-1)/n < t \leq i/n$, $1 \leq i \leq n$ and $J_n(0) = c_{n1}$. Let

$$\mu_n = \int_0^1 g_n J_n dt \ldots \ldots \ldots \tag{9}$$

For real numbers b_1, b_2, δ, define

$$B_{b_1,b_2}(t) = t^{-b_1}(1-t)^{-b_2}, \quad 0 < t < 1 \ldots \ldots \tag{10}$$

We impose the following conditions on $(g_n, n \geq 1)$ and $(J_n, n \geq 1)$. Below C denotes some generic constant.

(C1) $$|J_n| \leq C B_{\gamma,\varepsilon} \text{ for some } \gamma \text{ and } \varepsilon \ldots \ldots \tag{11}$$

(C2) For some (r, s) with $-1 < r + \gamma < 1/p$ and $-1 < s + \varepsilon < 1/p$,

$$|g_n| \leq ct^{-r}(1-t)^{-s}$$

(C3) $$\sup_{n \geq 1} \int_0^1 B_{\gamma,\varepsilon} B_{\delta,\delta} h d|g_n| < \infty$$

for some $\delta > 0$ and some h satisfying Theorem 1.

THEOREM 2. If $(g_n, J_n, n \geq 1)$ satisfy the conditions (C1), (C2) and (C3), then

$$L_n - \mu_n = o(n^{(1-p)/p}) \text{ almost surely } \dots \dots \qquad (12)$$

□

Proof. We may write

$$L_n - \mu_n = -(S_n + R_{n_1} + R_{n_2} + R_{n_3}) \dots \dots \dots \qquad (13)$$

where

$$\begin{aligned} S_n &= \int_{\xi_{n1}}^{\xi_{nn}} A_n(\Gamma_n - I) dg_n \\ &= \int_0^1 A_n^*(\Gamma_n - I) dg_n \\ R_{n_1} &= g_n(\xi_{n1})(\psi_n(0) - \psi_n(\xi_{n1})) \\ R_{n_2} &= g_n(\xi_{nn})\psi_n(\xi_{nn}) \\ R_{n_3} &= \int_{[\xi_{n1}, \xi_{nn})^c} g_n \; J_n \; d \; I \\ \psi_n(t) &= -\int_t^1 J_n dI \\ A_n &= \psi_n(\Gamma_n) - \psi_n(I)/(\Gamma_n - I) \\ A_n^*(t) &= A_n I_{[\xi_{n1}, \xi_{nn}]}(t) \end{aligned}$$

By Assumptions (C1) and (C2),

$$\begin{aligned} |R_{n1}| &\leq C(\xi_{n1})^{-r+1-\gamma} \\ |R_{n2}| &\leq C(1 - \xi_{nn})^{s+1-\varepsilon} \\ \left| \int_0^{\xi_{n1}} |g_n||J_n| dt \right| &\leq C \int_0^{\xi_{n1}} t^{-(r+\gamma)} dt \\ &\leq C(\xi_{n1}^{1-(r+\gamma)} \\ \int_{\xi_{nn}}^1 |g_n||J_n| dt &\leq C \int_{\xi_{nn}}^1 (1-t)^{-(s+\varepsilon)} dt \\ &\leq C(1 - \xi_{nn})^{1-(s+\varepsilon)} \end{aligned}$$

From Galambos (1978, page 261), $\xi_{n1} + (1 - \xi_{nn}) = O(n^{-1}(\log \; \log \; n))$ almost surely. Hence by using the conditions on r, v, s and ε,

$$\sum_{i=1}^{3} |R_{n_i}| = o(n^{(1-p)/p}) \text{ almost surely } \dots \dots \qquad (13)$$

By arguments given in Wellner (1977a, page 478) or Wellner (1977b, page 488), almost surely,

$$|A_n| \leq C B_{\gamma,\varepsilon} B_{\delta,\delta} \ldots \ldots \tag{14}$$

Using this, almost surely

$$|S_n| \leq C \rho_h(\Gamma_n, I) \int_0^1 B_{\gamma,\varepsilon} B_{\delta,\delta} h d|g_n| \ldots \ldots \tag{15}$$

The Theorem follows by using Assumptions (C3), (13) and (15). □

Example 2. Suppose $g_n(t) \equiv g(t) = F^{-1}(t)$, $J_n(t) \equiv 1$. Then $L_n = n^{-1} \sum_{i=1}^n X_i$, $\mu_n = \int_0^1 F^{-1}(t)dt = EX_1$. Suppose that $1 \leq p < 2$ and $E|X_1|^{p_1} < \infty$ for some $p < p_1$. Then it is easy to see that $g(t) = o(t^{-1/p_1})$ as $t \to 0$. Thus we may choose $\gamma = \varepsilon = 0$ in (C1) and $r = s = 1/p_1$ in (C2). Choose $h(t) = t^\alpha$ and $\delta > 0$ arbitrarily small such that $\alpha = \frac{1}{p_1} + 2\delta < 1/p$. Then the integral in (C3) equals $\int_0^1 (t(1-t))^{\delta + 1/p_1} d|F^{-1}(t)| < \infty$. Further $h(t)$ defined above satisfies conditions of Example 1. Hence Theorem 2 applies and $T_n - E(X_1) = o(n^{(1-p)/p})$ almost surely if $E|X_1|^{p_1} < \infty$ for some $p_1 > p$. Note that this remains true for $0 < p < 1$ if EX_1 is replaced by 0. To see this, modify the proof of Theorem 2 by replacing μ_n by 0. This falls just short of the usual MZSLLN for T_n given in (1). □

Example 3. The L-statistic often used in the $Q-Q$ plot is

$$L_n = n^{-1} \sum_{i=1}^n C_{n_i} X_{n_i} \ldots \ldots \tag{16}$$

where $C_{n_i} = \Phi^{-1}(i/(n+1))$ or $= \Phi^{-1}((i-1/2)/n)$ or some minor variations of these. It can be verified that in this case (C1) is satisfied with $\gamma = \varepsilon$, where $\varepsilon > 0$ is arbitrary small. As in Example 2, $g_n(t) \equiv g(t) = F^{-1}(t)$. By arguments given in Example 1, it again follows that for $0 < p < 2$, $T_n - \mu_n = o(n^{(1-p)/p})$ almost surely if $E|X_1|^r < \infty$ for some $r > p$. It may be noted that Wellner (1977b, Example 1b, page 492) proves the LIL for T_n under the additional assumption that F is normal. □

The above methods, being general cannot exploit any further smoothness or other conditions which may have been imposed on the weights. We now show how such assumptions may be exploited using first and/or second order differentials of statistics.

Define

$$T(F) = \int_0^1 F^{-1}(t)J(t)dt \tag{17}$$

We list some conditions which different researchers have used for the weight function J. See for example Shorack (1972) or Stigler (1974).

(A) J is bounded and continuous a.e. Lebesgue and a.e. F^{-1}.

(B) $J(u) = J(u') = 0$ for all $0 < u < \alpha < \beta < u' < 1$, ($\alpha, \beta$ fixed).

(C) J is bounded and continuous.

(D) J' exists everywhere on (0,1) and is Lip (δ).

Let F_n denote the empirical distribution of $X_1, \ldots, X_n$. Define (the leading term of $T(F_n) - T(F)$)

$$\alpha(x) = -\int_{-\infty}^{\infty} [I(x \leq y) - F(y)]J[F(y)]dy. \tag{18}$$

The following integration by parts result will serve as an important tool. For proof, see Serfling (1980, page 265). Let $K(t) = \int_0^t J(u)du$, $0 \leq t < 1$. Then

$$T(F_n) - T(F) = -\int_{-\infty}^{\infty} [K(F_n(x)) - K(F(x))]dx \tag{19}$$

Proposition 1. *Assume conditions (A) and (B) and for some* $0 < p < 2$ $E|\alpha(X_1)|^p < \infty$. *Then as* $n \to \infty$, $T(F_n) - T(F) = o(n^{(1-p)/p})$ *almost surely.* □

Proof. Define

$$W_{G,F}(x) = \begin{cases} \frac{K(G(x)) - K(F(x))}{G(x) - F(x)} - J(F(x)) & \text{if } G(x) \neq F(x) \\ 0 & \text{if } G(x) = F(x) \end{cases} \tag{20}$$

By using equations (19) and (20),

$$T(F_n) - T(F) = n^{-1} \sum_{i=1}^{n} \alpha(X_i) + R_{1n} \tag{21}$$

where

$$R_{1n} = -\int_{-\infty}^{\infty} W_{F_n,F}(x)[F_n(x) - F(x)]dx \tag{22}$$

Thus

$$|R_{1n}| \le ||W_{F_n,F}||_{L_1} \cdot ||F_n - F||_\infty$$

From Serfling (1980, page 281), the conditions (A) and (B) ensure that $||W_{F_n,F}||_{L_1} \to 0$ almost surely. Since $||F_n - F||_\infty = O(n^{-1/2} \log \log n)$ almost surely, this proves the result by an application of (1). □

J in Proposition 1 is a trimmed function (Assumption B). Our first result for untrimmed J function needs another assumption.

(E) There exists a function q such that $\int_{-\infty}^{\infty} q(F(x))dx < \infty$ and

$$||\frac{F_n - F}{qoF}||_\infty = O(n^{(1-p)/p}) \ a.s. \tag{22}$$

Proposition 2. *Let (A) and (E) hold and E* $|\alpha(X_1)|^p < \infty$ *for* $0 < p < 2$. *Then as* $n \to \infty$,

$$T(F_n) - T(F) = o(n^{(1-p)/p}) \text{ almost surely}$$ □

Proof. Recall (21). Now we use a different estimate of R_{1n}.

$$|R_{1n}| \le ||(qoF)W_{F_n,F}||_{L_1} \cdot ||\frac{F_n - F}{qoF}||_\infty.$$

By Lemma B of Serfling (1980, page 282), under Assumption (A) and (E),

$$||(qoF)W_{F_n,F}||_{L_1} \to \infty \text{ almost surely.}$$

This proves the proposition. □

We now use a two-term expansion. Define

$$Y_i(x) = I(X_i \le x) \tag{23}$$

$$\beta(x,y) = -\int_{-\infty}^{\infty} [I(x \le t) - F(t)][I(y \le t) - F(t)]J(F(t))dt \tag{24}$$

$$h(x,y) = [\alpha(x) + \alpha(y) + \beta(x,y)]/2 \tag{25}$$

Proposition 3. Assume conditions (C) and (D). Further assume that for some $1 < p < 2$, and $\delta > 0$, $1 < 1/p_0 < 1 + 1/p + \delta/2$ and $\frac{1}{p_1} < \min(1, 1/p + \delta/2)$,

(i) $E[\int_{-\infty}^{\infty}(Y_1(x)-F(x))^2]^{p_0} < \infty$

(ii) $E|\int_{-\infty}^{\infty}(Y_1(x)-F(x))(Y_2(x)-F(x))dx|^{p_1} < \infty$

(iii) $E[|h(X_1,X_2)|^p + |h(X_1,X_1)|^{p/(1+p)}] < \infty$

Then $T(F_n) - T(F) = o(n^{(1-p)/p})$ almost surely. □

Proof. Express $T(F_n) - T(F)$ as $V_{2n} + R_{2n}$ where

$$
\begin{aligned}
V_{2n} &= d_1T(F, F_n - F) + \frac{1}{2}d_2T(F, F_n - F) \\
&= n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n} h(X_i, X_j, F) \\
&= 2n^{-2}\sum_{1\le i<j\le n} h(X_i, X_j, F) + n^{-2}\sum_{i=1}^{n} h(X_1, X_1, F)
\end{aligned}
$$

with

$$
\begin{aligned}
R_{2n} = -\int_{-\infty}^{\infty} &\{K(F_n(t)) - K(F(t)) - J(F(t))[F_n(t) - F(t)] \\
&- \frac{1}{2}J'(F(t))[F_n(t) - F(t)]^2\}dt
\end{aligned}
$$

Since the first term of V_{2n} is (approximately) a U-statistic, by Remark 2.1 (i) of Bose and DasGupta (1994), this term is $o(n^{(1-p)/p})$ almost surely. By using Assumption (iii) and (1), the second term is $o(n^{-1+\frac{1-p/(1+p)}{p/(1+p)}}) = o(n^{(1-p)/p})$. For R_{2n}, from Serfling (1980, page 289)),

$$
\begin{aligned}
|R_{2n}| &\le c||F_n - F||_{L_2}^2 \cdot ||F_n - F||_{\infty}^{\delta} \\
&\le c||F_n - F||_{L_2}^2 0(n^{-\delta/2}) \ \textit{almost surely.}
\end{aligned}
$$

Note that

$$
\begin{aligned}
||F_n - F||_{L_2}^2 &= \int_{-\infty}^{\infty}[n^{-1}\sum_{i=1}^{n} Y_i(x) - F(x)]^2 dx \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}(Y_i(x) - F(x))^2 dx \\
&+ \frac{2}{n^2}\sum_{1\le i<j\le n}\int_{-\infty}^{\infty}(Y_i(x) - F(x))(Y_j(x) - F(x))dx
\end{aligned}
$$

By Assumption (i) and (1), almost surely

$$n^{-2}\sum_{i=1}^{n}\int_{-\infty}^{\infty}(Y_i(x)-F(x))^2\,dx = o(n^{-1+(1-p_0)/p_0})$$
$$= o(n^{-2+1/p_0})$$

By using Bose and DasGupta (1994) again, the second term is $o(n^{(1-p_1)/p_1})$ almost surely. Thus $|R_{2_n}| = o(n^{(1-p)/p})$ almost surely by using the restrictions on p_0 and p_1. This proves the Proposition. □

Remark 1. It can be verified that

$$\int_{-\infty}^{\infty}|Y_1(x)-F(x)|dx \le 2[|X_1| + E|X_1|]$$

and

$$[\int_{-\infty}^{\infty}|Y_1(x)-F(x)||Y_2(x)-F(x)|dx] \le 4[|X_1| + |X_2| + E|X_1|]$$

So $E|X_1|^{p_1} < \infty$ implies condition (ii) of Proposition 3. □

If higher moments of X_1 exist one may avoid checking all the conditions of Proposition 3 and the following variant of Proposition 3 holds.

Proposition 4. *If $1 < p < 2$, (C), (D) and (iii) of Proposition 3 hold and $E|X_1|^2 < \infty$ then $T(F_n) - T(F) = o(n^{(1-p)/p})$ almost surely.*

Proof. As before, write $T(F_n) - T(F) = V_{2n} + R_{2n}$. By Lemma B of Serfling (1980, page 288), for any sequence $(a_n \ge 0)$

$$\sum_{n=1}^{\infty}P(||F_n - F||_{L_2}^2 \ge a_n) \le \sum_{n=1}^{\infty}a_n^{-k}E[||F_n - F||_{L_2}^{2k}]$$
$$\le c\sum_{n=1}^{\infty}a_n^{-k}n^{-k} < \infty$$

if we choose $k = 2$ and $a_n \approx n^{\frac{1}{k}-1}(\log\ n)^{1+\varepsilon_0}$ for some $\varepsilon_0 > 0$. Hence almost surely,

$$R_{2n} = 0(n^{-\delta/2+1/2-1}(\log\ n)^{1+\varepsilon}) \text{ for } \varepsilon < \varepsilon_0.$$

Choosing δ small, this is $o(n^{(1-p)/p})$ almost surely, proving the Proposition. □

Remark 2. Often alternative forms of L statistics are used. For example Helmers (1977) considers

$$T_{1n} = n^{-1} \sum_{i=1}^{n} J(\frac{i}{n+1}) X_{(ni)}$$

where $X_{(n1)} \leq \ldots \leq X_{(nn)}$ is the ordered statistics of $(X_1, \ldots, X_n)$.

Note that

$$T(F_n) - T_{1n} = n^{-1} \sum_{i=1}^{n} X_{(ni)} \{ \int_{(i-1)/n}^{i/n} (J(t) - J(i/n)) dt \} + \frac{1}{n} \sum X_{(ni)} \{ J(i/n) - J(i/n+1) \}$$

Under Assumptions (C) and (D), it then follows that if $E|X_1|^{p/(1+p)} < \infty$, almost surely

$$|T(F_n) - T_{1n}| \leq cn^{-2} \sum_{i=1}^{n} |X_i| = o(n^{(1-p)/p}) \text{ almost surely}$$

Hence assuming the conditions of Propositions 3 or 4, $T_{1n} - T(F) = o(n^{(1-p)/p})$ almost surely.

Example 4. For $k \geq 0$, consider the sequence of L-statistics

$$L_n = n^{-1} \sum_{i=1}^{n} (\frac{i}{n})^k X_{(ni)}$$

and its smooth version

$$T(F_n) = \int_0^1 F_n^{-1}(t) J(t) dt$$

where

$$J(x) = x^k, \quad 0 \leq x \leq 1.$$

Note that $T(F) = \int_0^1 F^{-1}(t) J(t) dt$ is finite if $EX_1 < \infty$. Further, for any integer k,

$$T(F) = E(\max(X_1, \ldots, X_{k+1}))/(k+1).$$

Bose and DasGupta (1994, Remark 2.2) proved that if $E|X_1|^p < \infty$ for $1 \leq p < 2$, then for any **integer** $k \geq 1$, almost surely

$$L_n = T(F) = o(n^{(1-p)/p}) \ldots \ldots \ldots \tag{26}$$

By Lemma A of Serfling (1980, page 288), for any $p \geq 1$, $E|\alpha(X_1)|^p < \infty$ whenever $E|X_1|^p < \infty$. By Theorem 1, there exists a function q satisfying (22). Thus Proposition (2) applies and for any $k \geq 0$, if $E|X_1|^p < \infty$

$$T(F_n) - T(F) = o(n^{(1-p)/p}) \text{ almost surely } \dots \dots \dots \qquad (27)$$

As in Remark 2, since $T(F_n)$ and L_n are close, (26) also holds for any $k \geq 0$ provided $E|X_1|^p < \infty$. It is interesting to note that if we had applied Theorem 2 directly then (26) and (27) would require $E|X|^{p+\varepsilon} < \infty$ for some $\varepsilon > 0$. □

References

Bose, A. and DasGupta, R. (1994). On some asymptotic properties of U-statistics and one sided estimates. *Ann. Prob.* **23**, 4, 1715–1724.

Chow, Y. S. and H. Teicher (1978). *Probability Theory.* Springer Verlag, New York.

Galambos, J. (1978). *The asymptotic theory of extreme order statistics.* Krieger, Florida.

Helmers, R. (1977). The order of the normal approximation for linear combinations of order statistics with smooth weight functions. *Ann. Prob.* **5**, 940–953.

James, B. R. (1975). A functional law of the iterated logarithm for weighted empirical distributions. *Ann. Prob.* **3**, 762–772.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* J. Wiley, New York.

Shorack, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43**, 412–427.

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

Wellner, Jon A. (1977a). A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. *Ann. Statist.* **5, 3**, 473–480.

Wellner, Jon A. (1977b). A law of the iterated logarithm for functions of order statistics. *Ann. Statist.*, **5, 3**, 481–494.