On strong laws for generalized L-statistics with dependent data

David Gilat, Roelof Helmers

Abstract. It is pointed out that a strong law of large numbers for L-statistics established by van Zwet (1980) for i.i.d. sequences, remains valid for stationary ergodic data. When the underlying process is weakly Bernoulli, the result extends even to generalized L-statistics considered in Helmers et al. (1988).

Keywords: L-statistic, GL-statistic, strong law of large numbers, stationary ergodic process

Classification: 60F, 62G, 28D

Let $X_1, X_2, \ldots$ denote a (real-valued) ergodic stationary process (ESP) defined on a single probability space (Ω, A, P). The marginal distribution of the ESP is the common distribution function (df) F of the X_i's. To begin with define (ordinary) L-statistics by

$$L_n = \sum_{i=1}^{n} X_{i:n} \int_{(i-1)/n}^{i/n} J_n(s)\, ds$$

where $J_n : (0,1) \to \mathcal{R}$, $n = 1, 2, \ldots$ are Lebesgue-integrable functions and for $n = 1, 2, \ldots$, $X_{1:n} \leq \ldots \leq X_{n:n}$ denote the ordered $X_1, \ldots, X_n$. For a Lebesgue-integrable function $J : (0,1) \to \mathcal{R}$, define the parameter

$$\theta = \theta_J(F) = \int_0^1 J(s)\, F^{-1}(s)\, ds$$

where $F^{-1}(s) = \inf\{x : F(x) \geq s\}$, for $0 < s < 1$. Our first main result – Theorem 1 below – asserts that a strong law of large numbers for linear combinations of order statistics L_n obtained by van Zwet (1980) for the case of i.i.d. processes $X_1, X_2, \ldots$, remains valid (with essentially the same proof) if the i.i.d. assumption is replaced by the much weaker requirement that $X_1, X_2, \ldots$ is an ESP. Formally, we have the following SLLN for L-statistics with dependent data which complements Theorem L of Aaronson et al. (1996):

Theorem 1. *Let $\{X_n\}_{n\geq 1}$ be an ESP. Let $1 \leq p \leq \infty$, $p^{-1} + q^{-1} = 1$, and suppose that $J_n \in \mathcal{L}_p$ for $n = 1, 2, \ldots$ and $F^{-1} \in \mathcal{L}_q$. If there is a function $J \in \mathcal{L}_p$ such that*

$$\lim_{n\to\infty} \int_0^t J_n(s)\, ds = \int_0^t J(s)\, ds$$

for every $t \in (0,1)$ *(i.e.* $J_n \to J$ *weakly in* $\mathcal{L}_p$ *(for* $1 \le p < \infty$*) and weak* in* $\mathcal{L}_\infty$ *(for* $p = \infty$*)), and if either*

(i) $1 < p \le \infty$ *and*

$$\sup_n ||J_n||_p < \infty,$$

or

(ii) $p = 1$ *and* $\{J_n, n = 1, 2, \dots\}$ *is uniformly integrable.*

Then

$$\lim_{n\to\infty} L_n = \theta,$$

with probability 1.

PROOF: The proof follows exactly the argument given by van Zwet (1980) (cf. the proofs of his Lemma 2.1, Theorem 2.1 and Corollary 2.1) without any changes. We only have to recall the well-known fact that by the ergodic theorem the SLLN and the Glivenko-Cantelli theorem is not only true for i.i.d. sequences, but remains valid for ESP. □

To extend Theorem 1 to *generalized* L-statistics (GL-statistics), their definition, as in Helmers et al. (1988), will first be reviewed. For a positive integer m, let h (the *kernel*) be a measurable function from $\mathcal{R}^m$ to $\mathcal{R}$, and let $W_{1:n} \le \cdots \le W_{(n)_m:n}$ denote the ordered values of $h(X_{i_1}, \cdots, X_{i_m})$ taken over the $(n)_m = n(n-1)\cdots(n-m+1)$ m-tuples $(i_1, \cdots, i_m)$ of distinct indices from $\{1, \cdots, n\}$. Given a sequence $J_n : (0,1) \to \mathcal{R}$ $(n = 1, 2, \cdots)$, of Lebesgue-integrable functions, form the sequence of statistics

$$GL_n = \sum_{i=1}^{(n)_m} W_{i:n} \int_{(i-1)/(n)_m}^{i/(n)_m} J_n(s)\, ds.$$

Note that when $m = 1$ and $h(x) = x$, GL_n reduces to the ordinary L-statistic L_n.

Next, the form of the limiting value of GL_n (to be proved to exist a.s. under appropriate conditions) will be identified. To this end, let $Y_1, Y_2, \cdots$ be independent F-distributed random variables. For the kernel $h : \mathcal{R}^m \to \mathcal{R}$ consider the distribution function

$$H_F(y) = P_F\{h(Y_1, \cdots, Y_m) \le y\}, \quad y \in \mathcal{R}$$

and for a Lebesgue-integrable function $J : (0,1) \to \mathcal{R}$, form the parameter

$$\eta = \eta_{J,h}(F) = \int_0^1 J(s) H_F^{-1}(s)\, ds$$

where here, as before, $H_F^{-1}(s) = \inf\{y : H_F(y) \ge s\}$.

Note again that for $m = 1$ and $h(x) = x$, the parameter η reduces to the previous θ. For interesting parameters of form η, obtained by appropriate choices of J and h, see the examples in Helmers et al. (1988).

To imitate van Zwet's argument in the extension of Theorem 1 to GL-statistics, a strong law for the U-statistic

$$\frac{1}{(n)_m}\sum_{i=1}^{(n)_m} W_{i:n} = \frac{1}{(n)_m}\sum_{1\le i_1\ne i_2\ne\ldots\ne i_m\le n} h(X_{i_1},\ldots,X_{i_m})$$

is needed. Unfortunately no such strong law is available for general ESP's (see Example 4a in Aaronson et al. (1996)). Thus a more stringent mixing condition than mere ergodicity has to be imposed on the data $X_1, X_2, \cdots$. Recall that the stationary sequence $\{X_n\}_{n\ge1}$ is called *weakly Bernoulli* (WB) (also known as *absolutely regular*) if $d(m) = \sup\big\{d(m,k) : k \ge 1\big\} \to 0$ as $m \to \infty$, where $d(m,k)$ is the supremum of

$$\sum_{i=1}^{n}\big|P(A_i\cap B_i) - P(A_i)P(B_i)\big|$$

over all families of disjoint sets $A_i \cap B_i$, $i = 1, \cdots, n$, where $A_i \in \sigma(X_1, \cdots X_k)$ and $B_i \in \sigma(X_{k+m+1}, \cdots)$. The kernel h has to be also somewhat restricted: For data with marginal distribution F, it is required that the kernel $h : \mathcal{R}^m \to \mathcal{R}$ be *bounded by an F-integrable product*, i.e. that $|h(x_1,\ldots,x_m)| \le f(x_1)\ldots f(x_m)$ where $f : \mathcal{R} \to \mathcal{R}_+$ and $\int f(x)\,dF(x) < \infty$.

Proposition 1 (Theorem U(iii) in Aaronson et al. (1996))**.** *Let $\{X_n\}_{n\ge1}$ be a weakly Bernouilli ESP with marginal F and let $h : \mathcal{R}^m \to \mathcal{R}$ be measurable and bounded by an F-integrable product. Then:*

$$\lim_{n\to\infty}\frac{1}{n^m}\sum_{1\le i_1,\ldots,i_m\le n} h(X_{i_1},\ldots,X_{i_m}) = E_F h(Y_1,\ldots,Y_m)$$

with probability 1.

Corollary 1. *Under the conditions of the Proposition:*

$$\lim_{n\to\infty}\frac{1}{(n)_m}\sum_{1\le i_1\ne i_2\ne\ldots\ne i_m\le n} h(X_{i_1},\ldots,X_{i_m}) = E_F h(Y_1,\ldots,Y_m)$$

with probability 1.

Proof: Since $\frac{(n)_m}{n^m} \to 1$ as $n \to \infty$ (m fixed), it suffices to prove that $\lim_{n\to\infty}\frac{1}{n^m}\sum' h(X_{i_1},\ldots,X_{i_m}) = 0$ a.s., where $\sum'$ indicates summation over all

m-tuples $(i_1, \dots, i_m)$ with $i_j = i_k$ for some $j \neq k$.
By assumption $|h(x_1, \dots, x_m)| \leq f(x_1) \dots f(x_m)$, so it suffices to establish

$$(*) \qquad \lim_{n\to\infty} \frac{1}{n^m} \sum{}' Z_{i_1} \dots Z_{i_m} = 0 \quad \text{a.s.}$$

where $\{Z_n\}_{n\geq 1} = \{f(X_n)\}_{n\geq 1}$ is a nonnegative integrable ESP.
The key ingredient in the proof of (*) is a very special case of a result of Aaronson (1981). This special case is presented here (with a simplified proof) for the sake of completeness.

Lemma 1. *Suppose $\{Z_n\}_{n\geq 1}$ is a nonnegative integrable ESP and let $\alpha > 1$. Then: $\frac{1}{n^\alpha} \sum_{i=1}^n Z_i^\alpha \to 0$ with probability* 1.

PROOF: Given $\epsilon > 0$, choose $M > 0$ sufficiently large for $U_i = Z_i^\alpha I\{Z_i > M\}$ to satisfy $EU_i^{1/\alpha} < \epsilon$ (this is possible since $0 \leq Z_i$ and $EZ_i < \infty$).
Let $V_i = Z_i^\alpha - U_i = Z_i^\alpha I\{Z_i \leq M\}$. Then both $\{U_n\}_{n\geq 1}$ and $\{V_n\}_{n\geq 1}$ are nonnegative ESP's with $\{V_n\}_{n\geq 1}$ uniformly bounded by M. Consequently,

$$\frac{1}{n^\alpha} \sum_{i=1}^n Z_i^\alpha = \frac{1}{n^\alpha} \sum_{i=1}^n U_i + \frac{1}{n^\alpha} \sum_{i=1}^n V_i \leq \left(\frac{1}{n} \sum_{i=1}^n U_i^{1/\alpha} \right)^\alpha + \frac{1}{n^{\alpha-1}} \left(\frac{1}{n} \sum_{i=1}^n V_i \right)$$

because $\alpha > 1$. Now, by the ergodic theorem, the second term tends to zero a.s. and the first term has an almost sure limit smaller than ϵ^α (by the choice of M). Since $\epsilon > 0$ is arbitrary, the lemma follows. $\square$

To complete the proof of (*), hence of the Corollary, let $S_n = Z_1 + \dots + Z_n$, $S_{n,j} = S_n - Z_j$ $(j = 1, \dots, n)$. Since $Z_i \geq 0$, it is evident that

$$\sum{}' Z_{i_1} \dots Z_{i_m} \leq \sum_{k=2}^m \sum_{j=1}^n Z_j^k S_{n,j}^{m-k}.$$

Consequently,

$$\frac{1}{n^m} \sum{}' Z_{i_1} \dots Z_{i_m} \leq \sum_{k=2}^m \frac{1}{n^k} \sum_{j=1}^n Z_j^k \left(\frac{S_{n,j}}{n}\right)^{m-k}.$$

For each fixed $2 \leq k \leq m$,
(i) $\frac{1}{n^k} \sum_{j=1}^n Z_j^k \to 0$ a.s. by the Lemma; for each fixed $j \geq 1$ and $2 \leq k \leq m$
(ii) $(\frac{S_{n,j}}{n})^{m-k} \to (EZ_1)^{m-k}$ a.s. by the ergodic theorem.
It is now easily seen from (i) and (ii) that for each $k = 2, \dots, m$
$\frac{1}{n^m} \sum_{j=1}^n Z_j^k S_{n,j}^{m-k} \to 0$ a.s., hence so does the sum of $m - 1$ $(k = 2, \dots, m)$ such terms.

The extension of Theorem 1 to GL-statistics is now readily available.

Theorem 2. *Suppose $\{X_n\}_{n\geq 1}$ is a weakly Bernoulli ESP with marginal F and let h, H_F, J_n, J, GL_n and η be as defined above. Suppose h is bounded by an F-integrable product. If J_n and J satisfy the conditions of Theorem 1 and if $H_F^{-1} \in \mathcal{L}_q$, then*

$$\lim_{n\to\infty} GL_n = \eta$$

with probability 1.

Proof: The argument is completely analogous to the proof of Theorem 1. Note that the only probabilistic ingredient in the proof of Theorem 1 is the strong law and the Glivenko-Cantelli theorem for the sequence of empirical distributions based on observations from the stationary ergodic process. The rest is purely function-analytic. The function-analytic part of the proof of Theorem 2 is exactly the same as in van Zwet (1980). Since the appropriate strong law has already been established in the Corollary, the only missing link is an appropriate Glivenko-Cantelli type result. To state it, recall the distribution-function $H_F(y) = P_F\{h(Y_1, \cdots, Y_m) \leq y\}$ (here as before $Y_1, Y_2, \cdots$ are independent F-distributed r.v.'s) corresponding to the kernel h, and consider the associated empirical distribution-function

$$H_n(y) = \frac{1}{(n)_m} \sum_{1\leq i_1\neq i_2\neq\cdots\neq i_m\leq n} 1\{h(X_i, \cdots, X_{i_m}) \leq y\}.$$

For each fixed y, $H_n(y)$ is a U-statistic based on the indicator-kernel $h_y : \mathcal{R}^m \to \mathcal{R}$ defined by

$$h_y(x_1, \cdots, x_m) = \begin{cases} 1 & h(x_1, \cdots, x_m) \leq y \\ 0 & \text{otherwise.} \end{cases}$$

Since h_y is bounded and the underlying ESP is assumed to be WB, it follows by the Corollary that $H_n(y) \to H_F(y)$ a.s. as $n \to \infty$. That this almost sure convergence is uniform in y over $\mathcal{R}$, i.e. that

$$\lim_{n\to\infty} \sup_{y\in\mathcal{R}} \left|H_n(y) - H_F(y)\right| = 0 \quad \text{a.s.}$$

is now established by a purely analytic argument as in the standard proof of the classical Glivenko-Cantelli theorem. □

Remarks.

1. In view of Example 4a in Aaronson et al. (1996), Theorem 2 is false for general ESPs, even if the kernel h is bounded.
2. It is clear from van Zwet (1980) that the L-statistic L_n in Theorem 1 (similarly for GL_n in Theorem 2; cf. Corollary 4.1 of Helmers et al. (1988)) can be replaced by the more general statistic

$$\sum_{i=1}^{n} g(X_{i:n}) \int_{(i-1)/n}^{i/n} J_n(s)\, ds,$$

for any Borel measurable function $g : \mathcal{R} \to \mathcal{R}$, provided the assumption $F^{-1} \in \mathcal{L}_q$ is modified to $g \circ F^{-1} \in \mathcal{L}_q$, and the limiting parameter θ (η for Theorem 2) is adjusted accordingly. Note that for i.i.d. sequences $X_1, X_2 \ldots$ Theorem 2 can also be inferred from Corollary 3.1 of Helmers et al. (1988).

Added in proof: Proposition 1 quotes Theorem U(iii) in Aaronson et al. from a 1994 preprint. Between then and the publication of the paper in 1996, the manuscript has evidently been revised so that our Corollary 1 is stated as Theorem U(iii) in the published version of the paper. Since Aaronson et al. give only a hint as to why our Corollary 1 follows from our Proposition 1 (the version of Theorem U(iii) at our disposal when working on the present paper), we decided to publish a full proof.

Acknowledgment. The idea for this note was generated while D. Gilat was visiting the CWI, Amsterdam. This author would like to thank Professor M. Keane and his group at the CWI for their hospitality and support. Both authors wish to thank Jon Aaronson for pointing them to reference [1] which led to the simple proof of Lemma 1.

References

[1] Aaronson J., *An ergodic theorem with large normalising constants*, Israel J. Math. **38** (1981), 182–188.

[2] Aaronson J., Burton R., Dehling H., Gilat D., Hill T., Weiss B., *Strong laws for L- and U-statistics*, Trans. Am. Math. Society **348** (1996), 2845–2866.

[3] Helmers R., Janssen P., Serfling R., *Glivenko-Cantelli properties of some generalized empirical df's and strong convergence of generalized L-statistics*, Probab. Th. Rel. Fields **79** (1988), 75–93.

[4] van Zwet W.R., *A strong law for linear functions of order statistics*, Ann. Probab. **8** (1980), 986–990.

School of Mathematical Sciences, Tel Aviv University, Ramat Aviv 69978, Israel

E-mail: gilat@math.tau.ac.il

Centrum voor Wiskunde en Informatica (CWI), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

E-mail: helmers@cwi.nl

(*Received* April 18, 1996)