

## LAWS OF LARGE NUMBERS FOR L-STATISTICS

RIMAS NORVAIŠA

*Institute of Mathematics and Informatics  
Akademijos 4,  
Vilnius 2600, LITHUANIA*

(Received: February, 1992; revised: April, 1994)

### ABSTRACT

Consider  $L_n = n^{-1} \sum_{1 \leq i \leq n} c_{ni} g(X_{n:i})$  for order statistics  $X_{n:i}$  and let  $c_{ni} = n \int_{(i-1)/n}^{i/n} J d\lambda$  for some (Lebesgue)  $\lambda$ -summable over  $(0, 1)$  function  $J$ . Sufficient as well as necessary conditions for  $\lim_n L_n = \int_0^1 J g d\lambda$  to hold almost surely and in probability are given. Superposition (or Nemytskii) operators have been used to derive the laws of large numbers for  $L$ -statistics from the laws of large numbers in quasi-Banach function spaces for the empirical distribution functions based on  $X_1, \dots, X_n$ .

**Key words:** Order statistics, Law of Large Numbers,  $L$ -statistics, Central Limit Theorem.

**AMS (MOS) subject classifications:** 60F99, 60F17, 62G30.

### 1. Introduction

The strong law of large numbers (SLLN) for  $L$ -statistics has been considered in Wellner [19], Helmers [3], Sen [15], van Zwet [22], and Mason [12]. Most of the results concern sufficient conditions for the SLLN to hold for broad classes of  $L$ -statistics. Necessary and sufficient conditions are given by Mason [12] for the following special class of  $L$ -statistics. For each nonnegative and nonincreasing function  $g$  over  $T = (0, 1)$  and  $0 < p < \infty$ , set

$$L_n(p, g) = \frac{1}{n} \sum_{i=1}^n \left(\frac{i}{n}\right)^{p-1} g(U_{n:i}) \quad (1.1)$$

for Uniform  $(0, 1)$  order statistics  $U_{n:1} \leq \dots \leq U_{n:n}$ . By the first part of Theorem 1 from [12], we have that

$$\lim_{n \rightarrow \infty} L_n(p, g) = \int_T t^{p-1} g(t) dt =: M_p(g) \quad a.s.$$

if and only if

$$M_p(g) < \infty \quad \text{and} \quad E_p(g) = \int_T g^{1/p}(t) dt < \infty. \quad (1.2)$$

An analogous result for a nondecreasing function  $h$  and for the appropriately defined  $L$ -statistics  $L_n(p, h)$  also holds (cf. Theorem 2 from [12]). Then, invoking a boundedness condition on an arbitrary function  $J$  which appears in the weight constants given by  $c_{ni} = J(i/(n+1))$ , Mason's result in conjunction with those obtained by van Zwet [22] and Wellner [19] yield the best known results (cf. Corollary 1 in [12]). The proof of the main result in [12] relies on the asymptotic behavior results for order statistics.

Here another approach is used for the same purpose. To prove the law of large numbers for the  $L$ -statistics we firstly establish the law of large numbers for the corresponding empirical df in the space of measurable functions furnished with a function-norm depending on the weights of the  $L$ -statistics. Secondly, we use the fact that the  $L$ -statistic can be considered as an image of the empirical df by the superposition operator induced by the integral representation of the  $L$ -statistic. This approach leads to sufficient conditions for a broad class of  $L$ -statistics which appear to be necessary when specialized to a subclass closely related with (1.1). These conditions (the same as in (1.2)) also characterize the SLLN for the empirical df in an appropriate quasi-Banach function space, and in this way provides an interpretation of (1.2). Moreover, we complement and improve results obtained by van Zwet [22] and Wellner [19], respectively. The same approach in addition yields sufficient and necessary (in some sense) conditions for the weak law of large numbers (WLLN) to hold. To the best of our knowledge the convergence in probability of  $L$ -statistics was not treated in the literature previously. We hope that some of the results concerning laws of large numbers in quasi-Banach function spaces for empirical df's stated in a more general form than are presently needed are of interest in their own right. The use of the superposition operator for  $L$ -statistics seems to be new. The central limit theorem in function spaces for the empirical process in conjunction with differentiability properties of the superposition operators has been used in [14] to prove the asymptotic normality of  $L$ -statistics.

## 2. Results

Let  $U_1, \dots, U_n$  be iid Uniform  $(0, 1)$  rv's with order statistics  $U_{n:1} \leq \dots \leq U_{n:n}$  and let  $\lambda$  be Lebesgue measure on  $T = (0, 1)$ . We denote the *uniform empirical df* by

$$\mathbb{G}_n(t) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(0,t]}(U_i), \quad t \in T,$$

here and throughout the paper  $\mathbb{1}_A$  denotes the indicator function of the set  $A$ , and the *inverse uniform empirical df*  $\mathbb{G}_n^{-1}$  is the left-continuous function

$$\mathbb{G}_n^{-1}(t) = U_{n:i}, \quad \text{for } t \in ((i-1)/n, i/n], \quad i = 1, \dots, n.$$

We wish to consider  $L$ -statistics with the weight constants given by

$$c_{ni}(J) := \int_{(i-1)/n}^{i/n} J d\lambda, \quad i = 1, \dots, n,$$

for some  $\lambda$ -summable over  $T$  function  $J$  often called a *score function*. We will keep this assumption on the function  $J$  throughout the paper. More general versions of our results for the larger class of score functions considered in the previous works could have been given here using the standard approximation procedure. To simplify the presentation of the main ideas in this paper, such a generalization is not presented here. The choice of the score function is ruled by the following well-known integral representation of the  $L$ -statistic given by

$$L_n(J, g) := \frac{1}{n} \sum_{i=1}^n c_{ni}(J)g(U_{n:i}) = \int_T g(\mathbb{G}_n^{-1})Jd\lambda \quad (2.1)$$

$$= \int_T Jgd\lambda - \int_T \left\{ \int_t^{\mathbb{G}_n(t)} Jd\lambda \right\} dg(t) \quad (2.2)$$

which holds under appropriate assumptions on the existence of the integrals. We will refer to (2.1) and (2.2) as the first and the second representation of the  $L$ -statistic  $L_n(J, g)$  respectively. We will say that the  $L$ -statistic  $L_n(J, g)$  satisfies the SLLN (the WLLN) if the relation

$$\lim_{n \rightarrow \infty} L_n(J, g) = \int_T Jgd\lambda$$

holds almost surely (in probability, respectively).

Here we present the results for the case when the rv's are uniformly distributed and the function  $g$  satisfies the additional assumptions. But the proofs are given in a more general setting. A positive function  $g$  and  $T$  is  $u$ -shaped if  $g$  is nonincreasing on  $(0, 1/2]$  and symmetric about  $t = 1/2$ . The following statement is a consequence of Theorem 5.2 below.

**Corollary 2.1:** *Let  $g$  be a Borel measurable  $u$ -shaped function on  $T$ . The  $L$ -statistic (2.1) satisfies:*

(1) *the SLLN if*

$$\int_T g\left(\frac{t(1-t)}{2\log^\alpha(e/(t(1-t)))}\right) |J(t)| dt < \infty,$$

*for some  $\alpha > 1$ ;*

(2) *the WLLN if*

$$\int_T g(Mt(1-t)) |J(t)| dt < \infty,$$

*for every  $M \in (0, 1)$ .*

The first part of this statement extends Theorem 2.1 in [22] to nonintegrable functions  $g$ . For example, one can take  $g(t) = 1/t(1-t)$ ,  $t \in T$ , and the corresponding score function  $J$  to conclude the SLLN.

The main result may be considered to be a general statement in Theorem 6.1 below. The following consequence is a case of Corollary 6.2.

**Corollary 2.2:** *Let  $g$  be a left-continuous and nondecreasing function over  $T$  such that  $g(1/2) = 0$  and let  $0 < p < \infty$ . Suppose that there exists a finite constant  $C$  such that*

$$|J(t)| \leq C[t(1-t)]^{p-1}, \quad \text{for } \lambda\text{-a.a. } t \in T.$$

*Then the  $L$ -statistic (2.1) satisfies:*

(1) *the SLLN if*

$$\int_T |g(t)|^{1/p} dt < \infty \quad \text{and} \quad \int_T [t(1-t)]^{p-1} |g(t)| dt < \infty; \quad (2.3)$$

(2) the WLLN if only the second integral in (2.3) is finite.

The necessity of the conditions in (2.3) is treated in Proposition 6.3 under additional assumptions on the score function  $J$ . As noted in the introduction, conditions in (2.3) characterize the validity of the SLLN and the WLLN for the empirical df  $\mathbb{G}_n$  in the quasi-Banach function space  $L_p(T, dg)$  (cf. Theorems 4.2 and 4.3 below). In Theorem 6.4 below, one can find yet another sufficient condition for the SLLN for the  $L$ -statistic (2.2) to hold which complements the first part of Corollary 2.2 in the case when  $p > 1$ .

Now let us describe the scope of the paper. The next two sections are preparatory. Section 3 is devoted to some properties of superposition operators used here, and Section 4 deals with the characterization of the laws of large numbers for empirical df's with sample paths in function spaces. Another part of the paper is devoted to the laws of large numbers for  $L$ -statistics. The treatment is divided into Sections 5 and 6 depending on whether the first or the second representation of the  $L$ -statistic, respectively, is used.

### 3. Superposition Operators

The basic tool for what follows is a superposition operator acting between spaces of measurable functions. We recall here some properties of this operator, and refer to [1] for more information.

We start with some basic definitions and notations concerning quasi-Banach function spaces to be used throughout the paper. Let  $(T, \mu)$  be a complete  $\sigma$ -finite measurable space and let  $\mathbb{M}(T, \mu)$  be the linear space of all real valued  $\mu$ -measurable functions defined on  $T$ , where functions equal almost everywhere are identified. The mapping  $\|\cdot\|$  defined on  $\mathbb{M}(T, \mu)$  into  $[0, \infty]$  is called a *function quasi-norm* if  $\|\cdot\|$  has the following properties:

- (1)  $\|f\| = 0$  iff  $f = 0$  a.e.,  
 $\|cf\| = |c| \|f\|$ ,  $\forall c \in \mathbb{R}$ ,  
 $\exists \kappa < \infty$  such that  $\|f + g\| \leq \kappa[\|f\| + \|g\|]$ ,  $\forall f, g \in \mathbb{M}(T, \mu)$ ;
- (2) if  $|g| \leq |f|$  a.e., then  $\|g\| \leq \|f\|$ .

Given the function quasi-norm  $\|\cdot\|$ , define the set  $\mathbb{B}(T, \mu) = \{f \in \mathbb{M}(T, \mu) : \|f\| < \infty\}$ . Then  $\mathbb{B} = (\mathbb{B}(T, \mu), \|\cdot\|)$  is a quasi-normed linear space. If  $\mathbb{B}$  is complete, it is called a *quasi-Banach function space*. A  $\mu$ -measurable subset  $T_0 \subset T$  is called a *carrier* of  $\mathbb{B}$  if  $T \setminus T_0$  is the largest set in the sense of inclusion such that  $f \mathbb{1}_{T \setminus T_0} = 0$  for every  $f \in \mathbb{B}$ . Without repeating it every time we assume that the carrier of any  $\mathbb{B}(T, \mu)$  to appear below is equal to the whole of  $T$ . Another assumption to be assumed without further reference is that

$$\mathbb{1}_A \in \mathbb{B}(T, \mu), \quad \forall A \subseteq T \text{ such that } \mu(A) < \infty. \quad (3.1)$$

In passing, it may be remarked that in [11], the Banach function space  $\mathbb{B}$  with the condition (3.1) is said to be a Köthe function space.

Let  $\mathbb{B} = (\mathbb{B}(T, \mu), \|\cdot\|)$  be a quasi-Banach function space and let  $\{f_n : n \geq 1\}$  be a sequence in  $\mathbb{B}$ . We use the property that if there exists  $f \in \mathbb{B}$  such that  $\|f_n - f\| \rightarrow 0$  as  $n \rightarrow \infty$  then  $f_n \rightarrow f$  in  $\mu$ -measure. The proof of this fact for the function norm given in Theorem IV.3.1 from [5] can be extended to the function quasi-norm by invoking the existence of a  $p$ -homogeneous,  $0 < p < 1$ , norm, topologically equivalent to  $\|\cdot\|$ . A quasi-Banach function space  $\mathbb{B}$  is said to have an

order continuous function quasi-norm if it follows from  $0 \leq f_n \leq f$  in  $\mathbb{B}$  and  $f_n \rightarrow 0$   $\mu$ -a.e. ( in other words, from  $f_n \rightarrow 0$  in order) that  $\|f_n\| \rightarrow 0$ .

Turning to the definition of the superposition operator, let  $\phi(\cdot, \cdot)$  be a real valued function defined on  $T \times \mathbb{R}$ . A *superposition (or Nemytskii)* operator  $\Phi$  is defined by

$$\Phi f(t) = \phi(t, f(t)), \quad t \in T, \quad (3.2)$$

and we will refer to  $\Phi$  as to the superposition operator induced by the function  $\phi$ . A function  $\phi$  is said to be *superposition measurable* if the operator  $\Phi$  maps every measurable function into a measurable function, i.e., the superposition  $\phi(t, f(t))$  is measurable for every measurable function  $f$ . A stronger assumption often made on such operators is that  $\phi$  is a *Caratheodory function*, namely that  $\phi(\cdot, x)$  is measurable for each  $x \in \mathbb{R}$ , and  $\phi(t, \cdot)$  is continuous for  $\mu$ -almost all  $t \in T$ .  $\Phi$  is a nonlinear operator acting between subspaces of  $\mathbb{M}(T, \mu)$ . The basic fact to be used in this paper is due to Krasnosel'skii [7], (cf. also [8]) and, roughly speaking, says that  $\Phi$  is continuous from  $\mathbb{B}_1$  into  $\mathbb{B}_2$  if  $\Phi$  maps every function from  $\mathbb{B}_1$  into  $\mathbb{B}_2$ . The following statement is a variant of a more general fact proved in [9] (cf. also [6]), thus we omit the proof.

**Theorem 3.1:** *Let  $\phi$  be a Caratheodory function over  $T \times \mathbb{R}$  such that the superposition operator  $\Phi$  induced by  $\phi$  maps a quasi-Banach function space  $\mathbb{B}_1(T, \mu)$  into another quasi-Banach function space  $\mathbb{B}_2(T, \mu)$  with the order continuous function quasi-norm. Then  $\Phi$  is continuous.*

The assumption on  $\phi$  to be a Caratheodory function is comparatively strong in the present context. The advantage of the second representation (2.2) of the  $L$ -statistics is that the function  $\phi_J(t, \cdot) = \int_t^{\cdot} J d\lambda$  is continuous for all  $t \in T$  whenever the score function  $J$  is integrable. An analogous attempt to treat the  $L$ -statistics given by the first representation (2.1) requires continuity of  $g$ . Therefore, a weaker property of the superposition operator will be used in this case. Namely, if the function  $g$  is Borel measurable then an appropriately defined function  $\phi_g$  (cf. (5.2) below) is superposition measurable and the induced superposition operator  $\Phi_g$  preserves a uniform absolute continuity of the function norms. Recall that a subset  $W$  in a Banach function space  $\mathbb{B}$  with the finite measure space  $(T, \mu)$  is said to be of *uniformly absolutely continuous norm (UAC norm)* if given  $\epsilon > 0$ , there exists  $\delta > 0$  such that  $\|f\|_D < \epsilon$  holds for all  $f \in W$  and for all  $D \subset T$  such that  $\mu(D) < \delta$ .

**Theorem 3.2:** *Let  $(T, \mu)$  be a finite measure space and let  $\phi$  be superposition measurable function over  $T \times \mathbb{R}$  such that the superposition operator induced by  $\phi$  maps a Banach function space  $\mathbb{B}_1(T, \mu)$  into another Banach function space  $\mathbb{B}_2(T, \mu)$  with the order continuous function norm. Then  $\Phi$  maps every subsets  $F \subset \mathbb{B}_1$  of UAS norm into the set in  $\mathbb{B}_2$  of UAC norm also.*

The proof of this statement for  $\mathbb{L}_p$  spaces is to be found in [1] (Lemma 5.17.3). The proof in this more general setting follows the same lines, and we omit it.

## 4. Empirical Distribution Functions

We will be concerned further in this section with the question of when does the empirical df satisfy the law of large numbers considered as the normed partial sum of iid stochastic processes with almost all sample paths in a function space. This is a special case of a more general problem on the law of large numbers for vector valued rv's. The characterization of the SLLN for rv's with values in an arbitrary (possibly non-separable) Banach space is given by Talagrand [18]. Less is known about the convergence in probability and about the case of the non-Banach space valued rv's.

Let  $X_1, \dots, X_n$  be iid real rv's on a probability space  $(\Omega, Pr)$  with the df  $F$ . We assume that

$F$  is non-degenerate, i.e., that  $-\infty \leq a < b \leq +\infty$ , where  $a := \inf\{t \in \mathbb{R}: F(t) > 0\}$  and  $b := \sup\{t \in \mathbb{R}: F(t) < 1\}$ . Put  $T = T_F := (a, b)$ . We will consider the empirical df  $\mathbb{F}_n$  given by

$$\mathbb{F}_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(a,t]}(X_i), \quad t \in T_F,$$

as a stochastic process with almost all sample paths in a quasi-Banach function space  $\mathbb{B} = (\mathbb{B}(T_F, \mu), \|\cdot\|)$ , where the measure  $\mu$  is assumed to be Radon, i.e.,  $\mu([c, d]) < \infty$  for all compact subsets  $[c, d] \subset T_F$ . As usual, the asymptotic behavior of the empirical df  $\mathbb{F}_n$  near 0 and 1 at the endpoints  $a$  and  $b$  is of interest. The *symmetrization* of  $\mathbb{F}_n$  defined by

$$\mathbb{S}_n(t) := \begin{cases} -\mathbb{F}_n(t), & \text{if } t \in (a, c), \\ 1 - F(t), & \text{if } t \in [c, b). \end{cases} \quad (4.1)$$

for some point  $c \in T_F$ , provides a way to joint the analysis of both endpoints. for notational convenience, we assume  $c$  to be a continuity point of the df  $F$  such that  $F(c) = 1/2$  and put

$$m(t) := E\mathbb{S}_n(t) = \begin{cases} -F(t), & \text{if } t \in (a, c), \\ 1 - F(t), & \text{if } t \in [c, b). \end{cases}$$

The representation

$$\mathbb{S}_n(t) = \frac{1}{n} \sum_{i=1}^n \psi(X_i, t), \quad t \in T_F,$$

where the function  $\psi$  defined on  $T_F \times T_F$  is given by

$$\psi(x, \cdot) = \begin{cases} -\mathbb{1}_{[x, c)}, & \text{if } x \in (a, c), \\ \mathbb{1}_{[c, x)}, & \text{if } x \in [c, b), \end{cases}$$

is a simple consequence of the definition in (4.1).

Let  $\mathbb{B} = (\mathbb{B}(T, \mu), \|\cdot\|)$  be a quasi-Banach function space. For a given real rv  $X$ , the function  $\psi$  defines the map  $\psi \circ X := \psi(X, \cdot)$  with values in  $\mathbb{B}$ . By a standard terminology, the problem in question is when does the function  $n\psi$  satisfy the law of large numbers in  $\mathbb{B}$ . Since  $\mathbb{F}_n - F = m - \mathbb{S}_n$ , it is convenient to say that the *empirical df  $\mathbb{F}_n$  satisfies the SLLN (the WLLN) in  $\mathbb{B}$*  if there exists an element  $h \in \mathbb{B}$  such that  $\|\mathbb{S}_n - h\| \rightarrow 0$  almost surely (in probability, respectively). Since the convergence with respect to the function quasi-norm yields the convergence in  $\mu$ -measure, one can conclude that  $h = E\mathbb{S}_n = m$  whenever the law of large numbers holds for  $\mathbb{F}_n$ .

We start with the SLLN in a Banach function space  $\mathbb{B}$ . By Theorem 26 in [18], the empirical df  $\mathbb{F}_n$  satisfies the SLLN in  $\mathbb{B}$  if and only if  $\psi \circ X$  is properly measurable, and

$$\int^* \|\psi \circ X\| dPr < \infty. \quad (4.2)$$

The proper measurability notion is rather technical, and we refer to [18] for the definition. Since  $\psi \circ pX$  is a  $Pr \otimes \mu$ -measurable function on  $\Omega \times T$ , by Luxembour-Gribanov's theorem (cf. e.g. Theorem 99.2 from [21]), the map  $\Omega \ni \omega \mapsto \|\psi \circ X(\omega)\|$  is  $Pr$ -measurable. So, we can simply drop the asterisk in (4.2). The map  $\psi \circ X$  is scalarly measurable and need not have a separable valued range in general (cf. the proof of Theorem 1.5 from [2]). Thus, the well-known Bochner integrability condition for  $\psi \circ X$  is sufficient only (cf. [4]). It should be pointed out before

proceeding further, that the Banach function space  $\mathbb{B}$  is separable if and only if the function norm is order continuous and  $\mu$  is a separable measure. To verify Talagrand's proper measurability condition, we will assume that  $\psi \circ X$  is the map with values in an associate space rather than with values in an arbitrary Banach function space. For a given Banach function space  $\mathbb{B}$ , one can define the *associate space* (or Köthe dual space)  $\mathbb{B}'$  by

$$\mathbb{B}' = \{g \in \mathbb{M}(T, \mu) : \int_T |fg| d\mu < \infty, \quad \forall f \in \mathbb{B}\}.$$

Furnished with the associate norm

$$\|g\|_{\mathbb{B}'} = \sup\left\{ \left| \int fg d\mu \right| : \|f\| \leq 1 \right\}, \quad \forall g \in \mathbb{B}',$$

$\mathbb{B}'$  is a Banach function space also. This is a rather large class of function spaces including Lebesgue spaces  $L_p$  with  $1 \leq p < \infty$ .

Now we are ready to deduce the following statement:

**Proposition 4.1:** *Let  $X$  be a rv with the df  $F$  and let  $\mathbb{B} = (\mathbb{B}(T_F, \mu), \|\cdot\|_0)$  be a Banach function space. Then the empirical df  $F_n$  satisfies the SLLN in the associate space  $\mathbb{B}'$  if and only if  $E \|\psi \circ X\|_{\mathbb{B}'} < \infty$ .*

**Remark:** Also one can say that the Banach function space  $\mathbb{B}$  is a Glivenko-Cantelli class if and only if  $E \|\psi \circ X\|_{\mathbb{B}'} < \infty$ . Moreover, one can prove in a similar way, the SLLN in an arbitrary Banach function space  $\mathbb{B}$  if in addition  $\psi \circ X$  is a  $\mathbb{B}$ -valued rv with essentially separable valued range.

**Proof:** By Theorems 2.6(c) and 22(V) from [18], it is sufficient to prove that the set

$$Z_\psi = \left\{ \int_{T_F} f \psi \circ X d\mu : \|f\| \leq 1 \right\}$$

is order bounded, and the following relation

$$\overline{\lim}_{n \rightarrow \infty} E \|S_n - m\|_{\mathbb{B}'} = 0 \tag{4.4}$$

holds. It is obvious that the set  $Z_\psi$  is order bounded by  $\|\psi \circ X\|$ . Thus, we must only establish (4.4). Using the definition of the associate space and Fubini's theorem, one can deduce the following inequalities:

$$\begin{aligned} \|m\|_{\mathbb{B}'} &= \sup\left\{ \left| \int_F (E\psi \circ X) f d\mu \right| : \|f\| \leq 1 \right\} \\ &\leq \sup\left\{ E \left| \int_F (\psi \circ X) f d\mu \right| : \|f\| \leq 1 \right\} \leq E \|\psi \circ X\|_{\mathbb{B}'}. \end{aligned} \tag{4.5}$$

This is the only point where the definition of the associate space is used. Choose an arbitrary  $\epsilon > 0$ . By the integrability assumption, one can find finite numbers  $a_\epsilon$  and  $b_\epsilon$  such that

$$\int_a^{a_\epsilon} \|\psi(x, \cdot)\|_{\mathbb{B}'} F(dx) + \int_{b_\epsilon}^b \|\psi(x, \cdot)\|_{\mathbb{B}'} F(dx) \leq \epsilon. \tag{4.6}$$

Thus, by the triangle inequality and (4.5), we have

$$E \| S_n - m \|_{\mathbb{B}}' \leq 2E \| \psi \circ X \mathbb{1}_{(a, a_\epsilon)} \|_{\mathbb{B}}' + 2E \| \psi \circ X \mathbb{1}_{(b_\epsilon, b)} \|_{\mathbb{B}}' \\ + \| \mathbb{1}_{[a_\epsilon, b_\epsilon]} \|_{\mathbb{B}}' \mathop{Esup}_{t \in T_F} | F_n(t) - F(t) |.$$

We note that

$$\psi(x, \cdot) \mathbb{1}_{(a, a_\epsilon)} = 0 \quad \forall x > a_\epsilon$$

and

$$\psi(x, \cdot) \mathbb{1}_{(b_\epsilon, b)} = 0 \quad \forall x < b_\epsilon.$$

Hence, by (4.6) and by the Glivenko-Cantelli theorem in conjunction with the dominated convergence theorem, it follows that

$$\overline{\lim}_{n \rightarrow \infty} E \| S_n - m \|_{\mathbb{B}}' \leq 2\epsilon.$$

Since  $\epsilon$  is an arbitrary number, this gives us the desired relation (4.4). Proposition 4.1 is proved.  $\square$

**Remark:** The finiteness of the right side of (4.5) in conjunction with Theorem 6-1-2 from [17] yield that  $\psi \circ X$  is Pettis integrable and  $m$  is its Pettis integral.

Now we turn to the non-Banach function space case. In order to develop some feeling for what can be expected, we will compare the left and the right sides of (4.5) for  $L_p$ -space quasi-norms with  $0 < p < \infty$ . For all  $x, t \in T_F$ , we have (cf. Lemma 4.1 from [13])

$$| \psi(x, t) - m(t) | = | \mathbb{1}_{(a, t]}(x) - F(t) | \begin{cases} \leq | \psi(x, t) | + | m(t) |, \\ \geq \frac{1}{2} | \psi(x, t) | \wedge | m(t) |. \end{cases}$$

Thus, for any function quasi-norm  $\| \cdot \|$  we have

$$E \| \psi \circ X - m \| < \infty \quad \text{iff} \quad E \| \psi \circ X \| < +\infty \quad \text{and} \quad \| m \| < \infty. \quad (4.7)$$

The two conditions in the right side of (4.7) are different in general. For example, let  $U$  be a Uniform  $(0, 1)$  rv and let  $\mathbb{B} = (\mathbb{L}_p((0, 1), \mu), \| \cdot \|_p)$ , where  $\mu$  is the Lebesgue-Stieltjes measure induced by some nonnegative and nonincreasing function and  $\| \cdot \|_p$  is the usual  $\mathbb{L}_p$ -quasi-norm. Then it follows from the Proposition in the Appendix to [12] that the condition  $E \| \psi \circ U \|_p < \infty$  is stronger than, equivalent to, or weaker than the condition  $m \in \mathbb{L}_p$  depending on whether  $p > 1$ ,  $p = 1$ , or  $p < 1$ , respectively. This suggests that the characterization of the SLLN in the form of Proposition 4.1 ceases to hold for non-Banach function space valued rv's.

For given  $0 < p, q < \infty$ , define the quasi-Banach function space

$$\mathbb{L}_{p, q}(T, \mu) := \left\{ f \in \mathbb{M}(T, \mu) : \| f \|_{p, q} := \left( \int_a^c | f |^p d\mu \right)^{1/p} + \left( \int_c^b | f |^q d\mu \right)^{1/q} < \infty \right\}. \quad (4.8)$$

We shall be content here with the following.

**Theorem 4.2:** *Let  $X$  be a rv with the df  $F$  and let the quasi-Banach function space  $\mathbb{L}_{p, q}(T_F, \mu)$ ,  $0 < p, q < \infty$  be as above. Then the empirical df  $F_n$  satisfies the SLLN in  $\mathbb{L}_{p, q}$  if and only if*



$$E \|\psi \circ X - m\|_{p,q} < \infty. \quad (4.9)$$

**Proof:** Assume that the SLLN holds in  $\mathbb{L}_{p,q}$ . Hence  $\|\mathbb{S}_n - m\| \rightarrow 0$  almost surely as  $n \rightarrow \infty$ . Then the standard argument based on the Borel-Cantelli lemma implies (4.9).

Conversely, assume that (4.9) holds. Choose arbitrary numbers  $A, B$  such that  $a < A < B < b$  and put  $\Delta = \Delta(A, B) := (A, B)$ . Denote by  $\kappa_{p,q}$  the constant which appears in the triangle inequality for the quasi-norm  $\|\cdot\|_{p,q}$ . Then we have

$$\begin{aligned} \|\mathbb{S}_n - m\|_{p,q} &\leq \kappa_{p,q} \|\mathbb{I}_\Delta\|_{p,q} \sup_{t \in T_F} |\mathbb{F}_n(t) - F(t)| \\ &\quad + \kappa_{p,q}^2 [\|\mathbb{S}_n \mathbb{I}_{\Delta^c}\|_{p,q} + \|m \mathbb{I}_{\Delta^c}\|_{p,q}]. \end{aligned}$$

By the Glivenko-Cantelli theorem and (4.9) and due to the order continuity of the function quasi-norm  $\|\cdot\|_{p,q}$ , one can conclude that

$$\overline{\lim}_{n \rightarrow \infty} \|\mathbb{S}_n - m\|_{p,q} \leq \kappa_{p,q}^2 \overline{\lim}_{A \downarrow a, B \uparrow b} \overline{\lim}_{n \rightarrow \infty} \|\mathbb{S}_n \mathbb{I}_{\Delta^c(A,B)}\|_{p,q}, \quad a.s. \quad (4.10)$$

By the definition of  $\|\cdot\|_{p,q}$ , we have

$$\|\mathbb{S}_n \mathbb{I}_{\Delta^c(A,B)}\|_{p,q} = \|\mathbb{F}_n \mathbb{I}_{(a,A]}\|_p + \|(1 - \mathbb{F}_n) \mathbb{I}_{[B,b]}\|_q.$$

The two terms in the right side can be treated separately in an analogous manner. Thus, we will consider only the first of them. If  $p \geq 1$ , then by the SLLN in the Banach function space  $\mathbb{L}_p((a, A], \mu)$  (cf. Proposition 4.1) and by the order continuity of the function norm  $\|\cdot\|_p$ , we have

$$\overline{\lim}_{A \downarrow a} \overline{\lim}_{n \rightarrow \infty} \|\mathbb{F}_n \mathbb{I}_{(a,A]}\|_p = \overline{\lim}_{A \downarrow a} \|m \mathbb{I}_{(a,A]}\|_p = 0, \quad a.s.$$

Assume now  $p < 1$ . Since the function  $t^{p-1}$  is nonincreasing, by the mean value theorem, one can conclude that

$$\|\mathbb{F}_n \mathbb{I}_{(a,A]}\|_p^p \leq \int_a^A 2F^p(t) \mu(dt) + \int_a^A \mathbb{F}_n(t) F^{p-1}(t) \mu(dt).$$

Due to the SLLN in the space  $\mathbb{L}_1((a, A], \nu_p)$  with the measure  $\nu_p = F^{p-1} \mu$ , it follows that

$$\overline{\lim}_{A \downarrow a} \overline{\lim}_{n \rightarrow \infty} \int_a^A \mathbb{F}_n F^{p-1} d\mu = \overline{\lim}_{A \downarrow a} \|m \mathbb{I}_{(a,A]}\|_p^p = 0, \quad a.s.$$

These give us that the right side of (4.10) is zero almost surely and the proof of the theorem is completed.  $\square$

Now we proceed to the question of the convergence in probability of the empirical df in quasi-Banach function spaces. An application of Daniel's theorem in conjunction with the Glivenko-Cantelli theorem gives us necessary and sufficient conditions in this case. Only for the sake of simplicity we pose a condition on the function quasi-norm which implies measurability of  $\|\psi \circ X\|$ . Namely, for a given function quasi-norm  $\|\cdot\|$ , assume that for every subset  $\Delta \subset T$  of finite  $\mu$ -measure there exists a finite constant  $C_\Delta$  such that the inequality

$$\|f \mathbb{I}_\Delta\| \leq C_\Delta \int_\Delta |f| d\mu$$

holds for all  $f$ . Then, letting  $\Delta_n \uparrow T$  for any sequence of subsets with finite  $\mu$ -measure, we arrive to the desired measurability of the map  $\Omega \ni \omega \mapsto \|\psi \circ X(\omega)\|$  whenever the function quasi-norm is order continuous. It is clear that this condition is satisfied by the quasi-Banach function space  $L_{p,q}$ -spaces with  $0 < p, q < \infty$ .

**Theorem 4.3:** *Let  $X$  be a rv with the df  $F$  and let  $\|\cdot\|$  be an order continuous function quasi-norm on  $\mathbb{M}(T_F, \mu)$  as above. Then the empirical df  $\mathbb{F}_n$  satisfies the WLLN in  $\mathbb{B} = (\mathbb{B}(T_F, \mu), \|\cdot\|)$  if and only if  $m \in \mathbb{B}$ .*

**Proof:** Assume that the empirical df  $\mathbb{F}_n$  satisfies the WLLN in  $\mathbb{B}$ . Then there exists a subsequence  $\{n_k: k \in \mathbb{N}\}$  such that  $\|\mathbb{S}_{n_k} - h\| \rightarrow 0$  almost surely for some  $h \in \mathbb{B}$ . This implies the convergence  $|\mathbb{S}_{n_k} - h| \rightarrow 0$  in  $\mu$ -measure almost surely, and hence  $m = h \in \mathbb{B}$ .

Conversely, assume that  $m \in \mathbb{B}$ . choose an arbitrary number  $\epsilon > 0$ . By Daniel's theorem, there exists a positive number  $\lambda_\epsilon < \infty$  such that

$$Pr(\{sup_{t \in T} \frac{\mathbb{F}_n(t)}{F(t)} > \lambda_\epsilon\}) + Pr(\{sup_{t \in T} \frac{1 - \mathbb{F}_n(t)}{1 - F(t)} > \lambda_\epsilon\}) \leq \epsilon.$$

Due to the order continuity of the function quasi-norm, one can find real numbers  $A$  and  $B$  such that  $a < A < B < b$  and

$$\lambda_\epsilon \|m\|_{T_F \setminus [A, B]} < \epsilon/2\kappa.$$

This yields the inequality

$$Pr(\{\|\mathbb{S}_n - m\| > \epsilon\}) \leq \epsilon + Pr(\{\| \mathbb{1}_{[A, B]} \| sup_t |\mathbb{F}_n(t) - F(t)| > \epsilon/2\kappa\}).$$

An appeal to the Glivenko-Cantelli theorem completes the proof of the theorem.  $\square$

By the WLLN for Banach space valued rv's (cf. e.g. Theorem V.9.1 from [20]), one can conclude that for the empirical df  $\mathbb{F}_n$  to satisfy the WLLN in  $\mathbb{B}$ , the condition

$$\lim_{\lambda \rightarrow \infty} \lambda Pr(\{\|\psi \circ X - m\| > \lambda\}) = 0 \quad (4.11)$$

is sufficient if and only if the Banach space  $\mathbb{B}$  is of  $q$ -stable type and is necessary in an arbitrary Banach space. Thus, it would not be out of place to compare (4.11) with the condition  $m \in \mathbb{B}$  for the Banach function space  $\mathbb{B}$ . Since the following statement is not a principal result, the proof will be given for uniformly distributed rv's.

**Proposition 4.4:** *Let  $U$  be uniformly distributed on  $T = (0, 1)$  and let  $\mathbb{B}$  be a function Banach space with order continuous norm  $\|\cdot\|$ . Then the condition  $m \in \mathbb{B}$  implies (4.11).*

**Remark:** The converse implication is obvious. Moreover, in the separable Banach function space  $\mathbb{B} = \mathbb{L}_1$  which is not of 1-stable type, the condition  $m \in \mathbb{B}$  is equivalent to the Bochner integrability of  $\mathbb{B}$ -valued rv  $\psi \circ U$ . Thus, the WLLN and the SLLN for the uniform empirical df  $\mathbb{G}_n$  are equivalent in this case.

**Proof:** Define the family of functions  $\{f_u; u \in T\}$  on  $T$  by

$$f_u = \begin{cases} u\mathbb{1}_{[u, 1/2]}, & \text{if } u < 1/2, \\ (1-u)\mathbb{1}_{[1/2, u]}, & \text{if } u \geq 1/2. \end{cases}$$

Note that  $f_u \leq |m|$  for all  $u \in T$  and for every  $t \in T$  we have  $f_u(t) \rightarrow 0$  as  $u \downarrow 0$  and  $u \uparrow 1$ . The order continuity of the norm and the preceding remark yield that

$$\lim_{u \downarrow 0, u \uparrow 1} \|f_u\| = 0. \tag{4.12}$$

Choose an arbitrary number  $\epsilon > 0$ . By (4.12), there exists  $\delta \in (0, 1/2)$  such that

$$u \| \mathbb{1}_{[u, 1/2]} \| \leq \epsilon/4, \quad \forall u \in (0, \delta),$$

and

$$(1 - u) \| \mathbb{1}_{[1/2, u]} \| \leq \epsilon/4, \quad \forall u \in (1 - \delta, 1).$$

Then for all  $\lambda > 2 \| \mathbb{1}_{[\delta, 1 - \delta]} \| \vee 2 \| m \|$ , we have

$$\begin{aligned} \lambda Pr(\{ \| \psi \circ U - m \| > \lambda \}) &\leq \lambda Pr(\{ \| \psi \circ U - m \| > \lambda/2 \}) \\ &\leq \lambda Pr(\{ \| \mathbb{1}_{[u, 1/2]} \| > \lambda/2, U \leq \delta \}) + Pr(\{ \| \mathbb{1}_{[1/2, u]} \| > \lambda/2, U > 1 - \delta \}) \\ &\leq \lambda Pr(\{ U \leq \lambda\epsilon/2 \}) + \lambda Pr(\{ 1 - U \leq \lambda\epsilon/2 \}) = \epsilon. \end{aligned}$$

Since  $\epsilon$  is arbitrary, this completes the proof of the proposition. □

### 5. $L$ -Statistics Given by the First Representation

Assume in this section the function  $g$  is Borel measurable. Only in this section we will also confine ourselves to uniformly distributed rv's. By the first representation of the  $L$ -statistic (2.1), we have

$$\begin{aligned} L_n(J, g) &= \int_{1/n}^{1 - 1/n} g(\mathbb{G}_n^{-1}) J d\lambda + g(U_{n:1}) c_{n1}/n + g(U_{n:n}) c_{nn}/n \\ &=: T_n + \gamma_{n1} + \gamma_{nn}. \end{aligned} \tag{5.1}$$

By Theorem 2.1 from van Zwet [22], it is known that the  $L$ -statistic  $L_n(J, g)$  satisfies the SLLN if  $g \in \mathbb{L}_p(T, \lambda)$  and  $J \in \mathbb{L}_q(T, \lambda)$ , for some  $1 \leq p, q \leq \infty$  such that  $p^{-1} + q^{-1} = 1$  and  $T = (0, 1)$ . For the proof of this result, Vitali's theorem on convergence in the space  $\mathbb{L}_p(T, \lambda)$  has been applied. The proof of Theorem 5.2 below relies on Vitali's theorem on convergence in the space  $\mathbb{L}_1(T, |J| \lambda)$  in conjunction with Theorem 3.2 on superposition operators.

Turning to the definition of the superposition operator, put

$$h_{0,t}(t) := \begin{cases} \frac{1}{2} t / \log^\alpha(\frac{e}{t}), & \text{if } t \in (0, 1/2), \\ 1 - 2(1 - t) \log_2(\frac{e}{1-t}), & \text{if } t \in [1/2, 1), \end{cases}$$

and

$$h_{0,u}(t) := \begin{cases} 2t \log_2(\frac{e}{t}), & \text{if } t \in (0, 1/2), \\ 1 - \frac{1}{2}(1/t) \log^\alpha(\frac{e}{1-t}), & \text{if } t \in [1/2, 1), \end{cases}$$

for some  $\alpha > 1$ . Here, instead of the function  $w(t) = \log^\alpha(1/t)$  one can take any nondecreasing function  $w$  such that  $\int_0^1 [tw(t)]^{-1} dt < \infty$ . For a given  $M \in (0, 1)$ , define also

$$h_{M,l}(t) := \begin{cases} Mt, & \text{if } t \in (0, 1/2), \\ 1 - (1-t)/M, & \text{if } t \in [1/2, 1), \end{cases}$$

and

$$h_{M,u}(t) := \begin{cases} t/M, & \text{if } t \in (0, 1/2), \\ 1 - M(1-t), & \text{if } t \in [1/2, 1). \end{cases}$$

For every  $M \in [0, 1)$ , define the superposition operator  $\Phi = \Phi_M$  by (3.2) with  $\phi = \phi_M$  given by

$$\phi_M(t, x) := \begin{cases} g(h_{M,l}(t)), & \text{for } x < h_{M,l}(t), \\ g(x), & \text{for } x \in [h_{M,l}(t), h_{M,u}(t)], \\ g(h_{M,u}(t)), & \text{for } x > h_{M,u}(t), \end{cases} \quad (5.2)$$

for every  $t \in T$  and  $x \in \mathbb{R}$ ,

We begin by extending Lemma 2.1 from [22].

**Lemma 5.1:** *Let  $g$  be a Borel measurable function and let  $M \in [0, 1)$ . Then with probability 1,  $\phi_M(\cdot, \mathbb{G}_n^{-1})$  converges to  $g$  in Lebesgue measure, i.e., for every  $\epsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \lambda(\{t \in T: |\phi_M(t, \mathbb{G}_n^{-1}(t)) - g(t)| \geq \epsilon\}) = 0.$$

**Proof:** Suppose first that  $M = 0$ . By Corollary 1 in [16, p. 410] and by Theorem 2 in [16, p. 426], it follows that, with probability 1,  $\phi_0(\cdot, \mathbb{G}_n^{-1}) = g \circ \mathbb{G}_n^{-1}$  on the interval  $[1/n, 1 - 1/n]$  for sufficiently large  $n$ . Thus, the desired statement follows from Lemma 2.1 of [22] in this case. Suppose now that  $M \in (0, 1)$  and choose an arbitrary number  $\epsilon > 0$ . By Lusin's theorem there exist a Borel set  $B \subset T$  and a continuous function  $g_c$  defined on  $T$  such that  $\lambda(B) \leq \epsilon/2$  and  $g = g_c$  on  $T \cap B^c$ . Define  $B_n := \{t \in T: \mathbb{G}_n^{-1} \in B\}$ . Then it follows from the classical SLLN that

$$\lim_{n \rightarrow \infty} \lambda(B_n) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{1}_B(U_i) = \lambda(B) \text{ a.s.}$$

Define also  $A_n := \{t \in [\log n/n, 1 - \log n/n]: \mathbb{G}_n^{-1}(t) \notin [h_{M,l}, h_{M,u}]\}$ . Since  $M < 1$ , it follows from the estimate (10) in [16, p. 424] that with probability 1,  $A_n = \emptyset$  for sufficiently large  $n$ . Now, continuity of  $g_c$ , the Glivenko-Cantelli theorem and the preceding remarks yield that, with probability 1,

$$\begin{aligned} \overline{\lim}_n \lambda(\{t \in T: |\phi_M(t, \mathbb{G}_n^{-1}(t)) - g(t)| \geq \epsilon\}) &\leq \lambda(B) \\ &+ \overline{\lim}_n \left[ \lambda(B_n) + \lambda(A_n) + \frac{2 \log n}{n} \right] \\ &+ \overline{\lim}_n \lambda(\{t \in T: |g_c(\mathbb{G}_n^{-1}(t)) - g_c(t)| \geq \epsilon\}) \leq \epsilon \end{aligned}$$

establishing the desired statement in the case  $M \in (0, 1)$  also. Lemma 5.1 is thus proved.  $\square$

For a given function  $g$  and for every  $M \in [0, 1)$ , define the function  $g_M$  by

$$g_M(t) := \sup\{|g(s)| : s \in [h_{M,l}(t), h_{M,u}(t)]\}, \quad t \in T,$$

and assume  $g_M$  to be  $\lambda$ -measurable (take an upper  $\lambda$ -envelop if not). Now we are ready to state and prove the first result.

**Theorem 5.2:** *Let  $g$  be a Borel measurable function. The L-statistic  $L_n(J, g)$  represented by (5.1) satisfies:*

(1) *the SLLN if  $\gamma_{n1} + \gamma_{nn} \rightarrow 0$  a.s., as  $n \rightarrow \infty$ , and*

$$\int_T g_0 |J| d\lambda < \infty; \quad (5.3)$$

(2) *the WLLN if  $\gamma_{n1} + \gamma_{nn} \rightarrow 0$  in probability, as  $n \rightarrow \infty$ , and*

$$\int_T g_M |J| d\lambda < \infty, \quad \forall M \in (0, 1). \quad (5.4)$$

**Remark:** If  $g$  is unbounded and  $u$ -shaped then  $\gamma_{n1} + \gamma_{nn} \rightarrow 0$  almost surely (in probability) just by (5.3) (by (5.4), respectively). The first part of Theorem 5.2 also can be stated using an integrability condition on the function  $g_N$  from Theorem 3.2 of [22] which incorporates the convergence  $\gamma_{n1} + \gamma_{nn} \rightarrow 0$ .

**Proof:** (1) By upper and lower almost sure bounds on  $\mathbb{G}_n^{-1}$  (cf. Corollary 1 in [16, p. 426] and Theorem 2 in [16, p. 426]), it follows that with probability 1 we have  $\phi_0(\cdot, \mathbb{G}_n^{-1}) = g \circ \mathbb{G}_n^{-1}$  on the interval  $[1/n, 1 - 1/n]$  for sufficiently large  $n$ . Moreover,  $\phi_0(\cdot, I) = g$  on  $T$ , where  $I$  denotes the identity function  $I(t) = t$ . Hence, by the representation (5.1), it is sufficient to prove that, with probability 1,

$$\Phi_0 \mathbb{G}_n^{-1} \rightarrow \Phi_0 I \quad \text{in } \mathbb{L}_1(T, |J| \lambda) \quad \text{as } n \rightarrow \infty. \quad (5.5)$$

For this we check the assumptions of Theorem 3.2. Since the composition of a  $\lambda$ -measurable function and a Borel measurable function is  $\lambda$ -measurable, the function  $\phi_0$  in (5.2) is superposition measurable. Moreover, just by the definition, the inequality  $|\phi_0(\cdot, x)| \leq g_0$  holds for all  $x \in \mathbb{R}$ , and hence, by the assumption (5.3), the induced superposition operator  $\Phi_0$  maps  $\mathbb{L}_1(T, |J| \lambda)$  into  $\mathbb{L}_1(T, |J| \lambda)$ . Thus, we are in position to apply Theorem 3.2. Due to the finiteness of the measure  $|J| \lambda$  and by the Glivenko-Cantelli Theorem, with probability 1,  $\mathbb{G}_n^{-1}$  tends to  $I$  in the space  $\mathbb{L}_1(T, |J| \lambda)$ . Thus, an appeal to Theorem 3.2, assures us that with probability 1 the sequence  $\{\Phi_0 \mathbb{G}_n^{-1}; n \in \mathbb{N}\} \subset \mathbb{L}_1(T, |J| \lambda)$  is of UAC norm. Hence, (5.4) holds by Vitali's theorem whenever with probability 1 the sequence  $\phi_0(\cdot, \mathbb{G}_n^{-1})$  converges to  $g$  in Lebesgue measure. But this is exactly what Lemma 5.1 says in the case  $M = 0$ . This completes the first part of the proof.

(2) Choose an arbitrary number  $\epsilon > 0$ . By in-probability linear bounds on  $\mathbb{G}_n^{-1}$  (cf. Inequality 1 in [16, p. 419]), there exist  $M = M_\epsilon \in (0, 1)$  and a subset  $A_{n\epsilon}$  of  $\omega$  having  $Pr(A_{n\epsilon}) > 1 - \epsilon$  on which

$$g(\mathbb{G}_n^{-1}(t)) = \phi_{M_\epsilon}(t, \mathbb{G}_n^{-1}(t)), \quad \forall t \in [1/n, 1 - 1/n].$$

Thus, by the representation (5.1), we have

$$\begin{aligned} Pr(\{ |L_n(J, g) - \int_{1/n}^{1-1/n} g J d\lambda| \geq \epsilon \}) &\leq Pr(A_{n\epsilon}) + Pr(\{\gamma_{n1} + \gamma_{nn} \geq \epsilon/2\}) \\ &+ Pr(\{ \int_T |\phi_{M_\epsilon}(\cdot, \mathbb{G}_n^{-1}) - g| |J| d\lambda \geq \epsilon/2 \}). \end{aligned} \quad (5.6)$$

Now, as in the case (1), we conclude that (5.5) holds for the superposition operator  $\Phi_M$  instead of  $\Phi_0$ , thereby proving the convergence to zero of the right side of (5.6). This completes the proof of Theorem 5.2.  $\square$

## 6. $L$ -Statistics Given by the Second Representation

We shall consider in this section the  $L$ -statistic

$$L_n(J, g) = \frac{1}{n} \sum_{i=1}^n c_{ni}(J)g(X_{n:i}) \quad (6.1)$$

for iid rv's  $X_1, \dots, X_n$  with the df  $F$ . Concerning the function  $g$  on  $T_F = (a, b)$ , assume throughout this section that there exists a signed Lebesgue-Stieltjes (Radon) measure  $dg$  on  $T_F$  such that

$$g(t) = \int_{[c, t]} dg, \quad t \in T_F.$$

For such a  $g$ , its total variation function  $v(g)$  can be defined by

$$v(g)(t) = \begin{cases} \text{var}(g; [c, t]), & \text{for } t \in (c, b), \\ 0, & \text{for } t = c, \\ -\text{var}(g; [t, c]), & \text{for } t \in (a, c), \end{cases}$$

and the induced total variation measure denote by  $v(dg)$ . These can be done if and only if  $g$  is left-continuous and of locally bounded variation over  $T_F$ .

To justify the second representation of the  $L$ -statistic (6.1) assume that

$$\int_0^1 |Jv(g) \circ pF^{-1}| d\lambda < \infty, \quad (6.2)$$

where  $F^{-1}(t) = \inf\{x \in T_F: F(x) \geq t\}$ ,  $t \in (0, 1)$ . Note that (6.2) is equivalent to the finiteness of two integrals with the positive variation function  $g^+$  and the negative variation function  $g^-$  instead of the total variation function  $v(g)$  in (6.2), respectively. This in conjunction with the expression  $g = g^+ - g^-$  justifies the identity

$$\begin{aligned} L_n(J, g) &= \int_0^1 Jg \circ pF^{-1} d\lambda - \int_{T_F} \left\{ \int_F^n J d\lambda \right\} dg \\ &= : \theta + T_n. \end{aligned} \quad (6.3)$$

Turning to the superposition operator  $\Phi = \Phi_J$  defined by (3.2), choose an arbitrary number  $\gamma \in (1/2, 1)$  and put

$$\phi_J(t, x) = \frac{|x| \wedge \gamma}{|m(t)|} \int J_c(s, t) ds, \quad t \in T_F, x \in \mathbb{R}, \quad (6.4)$$

where

$$J_c(s, t) = J(s)\mathbb{1}_{(a, c)}(t) + J(1-s)\mathbb{1}_{[c, b)}(t), \quad s, t \in T_F, \quad (6.5)$$

and  $m = ES_n$  (cf. (4.1) above). The number  $\gamma$  in (6.4) is used to estimate the score function  $J$  when it is unbounded (cf. (6.10) below).

An advantage of the following general statement is that it is true in a more general setting than we use here. For example, the  $L$ -statistic and the corresponding empirical df  $\mathbb{F}_n$  can be based on nonidentically distributed rv's with an arbitrary dependence structure.

**Theorem 6.1:** *For a given df  $F$ , a score function  $J$  and the function  $g$  as above assume that (6.2) holds. Suppose that there exists a quasi-Banach function space  $\mathbb{B} = (\mathbb{B}(T_F, v(dg)), \|\cdot\|)$  such that the superposition operator  $\Phi_J$  induced by the function  $\phi_J$  given by (6.4) maps  $\mathbb{B}$  into  $\mathbb{L}_1(T_F, v(dg))$ . Then the  $L$ -statistic (6.1) satisfies the SLLN (the WLLN) whenever the empirical df  $\mathbb{F}_n$  based on the rv's with the df  $F$  satisfies the SLLN (the WLLN, respectively) in  $\mathbb{B}$ .*

**Proof:** It is clear that

$$\int_{F(t_0)}^{\mathbb{F}_n(t)} J(s)ds = \int_{|m(t)|}^{|\mathbb{S}_n(t)|} J_c(s, t)ds, \quad \forall t \in T_F.$$

By the Glivenko-Cantelli theorem, we have that, with probability 1,

$$|\mathbb{S}_n| \leq |\mathbb{F}_n - F| + 1/2 \leq \gamma \quad \text{on } T_F$$

for sufficiently large  $n$ . Hence, taking into account (6.3) and (6.4) we see that, with probability 1,

$$|L_n(J, g) - \theta| = |T_n| \leq \int_{T_F} |\Phi_J \mathbb{S}_n| v(dg), \quad (6.6)$$

for sufficiently large  $n$ . Due to the assumption, we have that  $\mathbb{S}_n \rightarrow m$  almost surely in  $\mathbb{B}$ . Hence, by Theorem 3.1, the right side of (6.6) tends to zero as  $n \rightarrow \infty$ . This establishes the first part of the theorem. Turning to the second part, choose an arbitrary number  $\epsilon > 0$ . By Theorem 3.1 once again, there exists a number  $\delta = \delta_\epsilon > 0$  such that

$$\int_{T_F} |\Phi_J f| v(dg) \leq \epsilon$$

for all  $f \in \mathbb{B}$  such that  $\|f - m\| \leq \delta$ . Therefore, we have

$$\begin{aligned} Pr(\{|L_n(J, g) - \theta| > \epsilon\}) &\leq Pr(\{\sup_{t \in T_F} |\mathbb{S}_n(t)| > \gamma\}) + Pr(\{\int_{T_F} |\Phi_J \mathbb{S}_n| v(dg) > \epsilon\}) \\ &\leq Pr(\{\sup_{t \in T_F} |\mathbb{F}_n(t) - F(t)| > \gamma - 1/2\}) + Pr(\{\|\mathbb{S}_n - m\| > \delta\}) \end{aligned}$$

and the second part of the theorem follows by the Glivenko-Cantelli theorem and the WLLN in  $\mathbb{B}$  for the empirical df  $\mathbb{F}_n$ .  $\square$

Invoking a score boundedness condition: there exists a finite constant  $C$  such that

$$|J(t)| \leq CT^{p-1}(1-t)^{q-1}, \quad \text{for } \lambda\text{-a.a. } t \in (0, 1); \quad (6.7)$$

for the quasi-Banach function space  $\mathbb{B}$  in Theorem 6.1, one can take the space  $\mathbb{L}_{p,q}$  defined by (4.8). In this way, Theorem 6.1 in conjunction with Theorems 4.2 and 4.3 yield:

**Corollary 6.2:** *Assume to be given the df  $F$ , the function  $g$  as above and the score function  $J$  such that (6.7) holds for some  $0 < p, q < \infty$ . Then the  $L$ -statistic (6.1) satisfies the SLLN if*

$$\int_a^c |v(g)|^{1/p} f F + \int_c^b |v(g)|^{1/q} dF < \infty \quad (6.8)$$

and

$$\int_a^c |v(g)| dF^p + \int_c^b |v(g)| d(1-F)^q < \infty. \quad (6.9)$$

Moreover, the  $L$ -statistic (6.1) satisfies the WLLN if only (6.9) holds.

**Proof:** The formula for integration by parts and the score boundedness condition (6.7) yield that (6.2) is a consequence of (6.9). By (6.7) once again, it follows the existence of a finite constant  $C = C(p, q, \gamma)$  such that for the function  $\phi_J$  defined by (6.4) we have

$$|\phi_J(\cdot, x)| \leq C\{[|x|^p + |m|^p]_{(a,c)} + [ |x|^q + |m|^q ]_{[c,b_0]}\} \quad (6.10)$$

for all  $x \in \mathbb{R}$ . By virtue of the formula for integration by parts, we infer that  $m \in \mathbb{B}$  if and only if (6.9) holds. This in conjunction with (6.10) yields that the superposition operator  $\Phi_J$  induced by  $\phi_J$  (see (6.4)) maps  $\mathbb{L}_{p,q}(T, v(dg))$  into  $\mathbb{L}_1(T, v(dg))$ , and for the quasi-Banach function space  $\mathbb{B}$  in Theorem 6.1 we can take the space  $\mathbb{L}_{p,q}(T, v(dg))$ . Turning to the law of large numbers for the empirical df in  $\mathbb{L}_{p,q}$ , by (4.9) one can conclude that (4.11) holds iff (6.8) and (6.9) are satisfied. Hence, by Theorem 4.2 the SLLN in  $\mathbb{L}_{p,q}$  holds, and the first part of the corollary follows from Theorem 6.1. The second part corresponding to the WLLN follows in a similar way from Theorem 6.1 in conjunction with Theorem 4.3. This completes the proof of Corollary 6.2.  $\square$

Note that conditions (6.8) and (6.9) reduce to (1.2) for the  $L$ -statistics (1.1). Next we show that these conditions also are necessary when specialized to a subclass of  $L$ -statistics.

**Proposition 6.3:** *I. Let  $F$  be a df,  $g$  be a left-continuous nonincreasing and nonnegative function on  $T_F$  and take the score function  $J_p(t) = t^{p-1}$ ,  $t \in (0, 1)$ , for some  $p > 0$ . If, for the  $L$ -statistic (6.1), we have*

$$\overline{\lim}_{n \rightarrow \infty} L_n(J_p, g) < \infty \quad \text{a.s.}, \quad (6.11)$$

then

$$\int_{T_F} g^{1/p} dF < \infty. \quad (6.12)$$

*II. Let  $g$  be a nonnegative Borel measurable function on  $T = (0, 1)$  and let the score function  $J$  be nonnegative. Consider the  $L$ -statistic (2.1) based on uniformly distributed rv's. If (6.11) holds with the score function  $J$  instead of  $J_p$ , then*

$$\int_T g(t) J(2t) dt < \infty.$$

**Remark:** The necessity question for the WLLN is more complicated. Recall that  $A$ -integrability (of  $g \circ X$ ) is a necessary and sufficient condition for the WLLN to hold in the classical setting ( $J \equiv 1$ ).



**Proof:** I. Put  $g^- := -g$ . By Abel's identity, one can conclude that

$$L_n(J_p, g) = \frac{1}{p} \int_{T_F} F_n^p dg^- + \frac{g(b)}{p}.$$

Note that the integral is equal to  $\|F_n\|_p^p$  where  $\|\cdot\|_p$  is the function quasi-norm on the space  $\mathbb{L}_p(T_F, dg^-)$ ,  $0 < p < \infty$ , and by the triangle inequality we have

$$n^{-1} \|\mathbb{1}_{[X_n, b)}\|_p \leq \kappa_p \left\{ \|F_n\|_p + \left(\frac{n-1}{n}\right) \|F_{n-1}\|_p \right\}.$$

Hence, due to the assumption (6.11), we have

$$\overline{\lim}_{n \rightarrow \infty} n^{-1} \|\mathbb{1}_{[X_n, b)}\|_p < \infty \text{ a.s.}$$

Thus, by independence, the Borel-Cantelli lemma and identical distributions, one can conclude that

$$\infty > E \|\mathbb{1}_{[X, b)}\|_p = \int_{T_F} (g(t) - g(b))^{1/p} F(dt) \geq 2^{-1/p} \int_{\{t: g(t) > 2g(b)\}} g^{1/p} dF.$$

This implies (6.12) as required.

II. Take a sequence  $\{a_n: n \in \mathbb{N}\}$  defined by  $a_n := \log n/n$ . By the third relation in (10) in [16, p. 424], we have that with probability 1

$$\sup_{a_n \leq t \leq 1} G_n^{-1}(t)/t \leq 2$$

for all sufficiently large  $n$ . Hence, due to the assumption, it follows that

$$\infty > \overline{\lim}_{n \rightarrow \infty} L_n(J, g) \geq \overline{\lim}_{n \rightarrow \infty} \int_{a_n}^{1/2} g \circ p G_n^{-1} J d\lambda \geq \frac{1}{2} \int_0^1 J(2t)g(t)dt.$$

This completes the proof of Proposition 6.3. □

By the aforementioned result from Mason [12], the condition (6.8) is stronger than (6.9) if  $p > 1$  and  $q > 1$ . The next statement compliments this case by providing still another sufficient condition. The proof given below differs very little from that used for the proof of Theorem 3 in Wellner [19]. Namely, the ILN in the space  $\mathbb{L}_1$  for the empirical df is used instead of the SLLN in a sup-norm. The improvement consists that one can take a logarithmic factor in the definition of the function  $h$  (cf. Assumption 1 in [19]) instead of the power function  $t^{-\delta}$  for some  $\delta > 0$ . For simplicity, we give the proof for uniformly distributed rv's and assume that  $p = q$  in the scores boundedness condition (6.7).

**Theorem 6.4:** *Let the function  $g$  be as above and let the score function  $J$  be such that for some  $1 \leq p < \infty$  there exists a finite constant  $C$  such that*

$$|J(t)| \leq C[t(1-t)]^{p-1}, \quad \text{for } \lambda\text{-a.a. } t \in (0, 1). \tag{6.13}$$

*Then the L-statistic (2.2) based on uniformly distributed rv's satisfies the SLLN if*

$$\int_0^1 |m|^p [\log(e/|m|)]^{\alpha(p-1)} \nu(dg) < \infty, \tag{6.14}$$

for some  $\alpha > 1$ .

**Remarks:** 1. Instead of the function  $w(t) = \log^\alpha(1/t)$  in (6.14), one can take any nondecreasing function  $w$  such that  $\int_0^1 [tw(t)]^{-1} dt < \infty$ .

2. An analogous proof in the case  $0 < p < 1$  gives a weaker result than that in Corollary 6.2. But the condition for the WLLN appears to be the same as in Corollary 6.2.

**Proof:** Due to assumption (6.13), it follows that

$$\left| \int_t^{\mathbf{G}_n(t)} J(s) ds \right| = \left| \int_{|m(t)|}^{|\mathbf{S}_n(t)|} J_c(s, t) ds \right| \leq |\mathbf{S}_n(t)|^{p-1} \vee |m(t)|^{p-1} |\mathbf{S}_n(t) - m(t)|,$$

for all  $t \in T = (0, 1)$ , where the function  $J_c$  is given by (6.5). By the SLLN for the empirical df in a sup-norm (cf. Theorem 1 in [16, p. 410]), it follows that, with probability 1, the inequality

$$|\mathbf{S}_n|^{p-1} \leq 2 |m|^{p-1} [\log(1/|m|)]^{\alpha(p-1)} \quad \text{on } T$$

holds for sufficiently large  $n$ . Hence, with probability 1, for all sufficiently large  $n$

$$\left| \int_G \int_t^{\mathbf{G}_n(t)} J d\lambda g(dt) \right| \leq \int_T [ |m| \log^\alpha(1/|m|) ]^{p-1} |\mathbf{S}_n - m| v(dg). \quad (6.15)$$

Thus, by Theorem 4.1 for the Banach function space  $\mathbb{L}_1(T, \mu)$  with

$$\mu = [ |m| \log^\alpha(1/|m|) ]^{p-1} v(dg),$$

the right side of (6.15) approaches zero almost surely as  $n \rightarrow \infty$ . Hence, due to the second representation (2.2), we arrive to the conclusion of Theorem 6.4.  $\square$

**Acknowledgement:** The author wishes to thank Andrew Rosalsky and the referee for their valuable remarks.

## References

- [1] Appell, J. and Zabrejko, P.P., *Nonlinear Superposition Operators*, Cambridge Univ. Press 1990.
- [2] Dobrič, V., The law of large numbers, examples and counterexamples, *Math. Scand.* **60** (1987), 273-291.
- [3] Helmers, R., A strong law of large numbers for linear combinations of order statistics, *Math Centrum, Amsterdam Report SW 50/77* (1977).
- [4] Hoffmann-Jørgensen, J., The law of large numbers for nonmeasurable and nonseparable elements, *Astérisque* **131** (1985), 299-356.
- [5] Kantorovič, L.V. and Akilov, G.P., *Functional Analysis*, 2nd ed., Nauka, Moscow 1977.

- [6] Kozłowski, W., Nonlinear operators in Banach function spaces, *Comm. Math. (Prace Mat.)* **22** (1980), 85-103.
- [7] Krasnosel'skii, M.A., On the continuity of the operator  $Fu(x) = f(x, u(x))$ , *Doklady Akad. Nauk SSR* **77** (1951), 185-188 (in Russian).
- [8] Krasnosel'skii, M.A., *Topological Methods in the Theory of Nonlinear Integral Equations*, Gostehizdat, Moscow 1956 (in Russian). [Engl. trans: MacMillan, New York 1964].
- [9] Krasnosel'skii, M.A., Rutitskij, Ya.B., and Sultanov, R.M., On a nonlinear operator which acts in a space of abstract functions, *Izvestija Akad. Nauk Azerbajdzh SSR* **3** (1959), 15-21 (in Russian).
- [10] Krasnosel'skii, M.A., Zabrejko, P.P., Pustyl'nik, Je.L., and Sobolevskij, P.Je., *Integral Operators in a Space of Abstract Functions*, Nauka, Moscow 1966 (in Russian). [Engl. trans: Noordhoff, Leiden 1976].
- [11] Lindenstrauss, J. and Tzafriri, L., *Classical Banach Spaces II. Function Spaces*, *Ergebnisse* **97**, Springer-Verlag, Berlin-Heidelberg-New York 1979.
- [12] Mason, D.M., Some characterizations of strong laws for linear functions of order statistics, *Ann. Probab.* **10** (1982), 1051-1057.
- [13] Norvaiša, R., The central limit theorem for empirical and quantile processes in some Banach spaces, *Stoch. Processes and Their Appl.* **46** (1993), 1-27.
- [14] Norvaiša, R., The central limit theorem for L-statistics, In: *Tech. Rep. Ser. Lab. Res. Stat. Prob.* **209** (1992), Carleton U.-U. of Ottawa.
- [15] Sen, P.K., An invariance principle for linear combinations of order statistics, *z. Wahrsch. verw. Gebiete* **42** (1978), 327-340.
- [16] Shorack, G.R. and Wellner, J.A., *Empirical Processes with Applications to Statistics*, John Wiley & Sons, New York 1986.
- [17] Talagrand, M., Pettis integral and measure theory, *Mem. Amer. Math. Soc.* **51** (1984), 307.
- [18] Talagrand, M., The Glivenko-Cantelli problem, *Ann. Probab.* **15** (1987), 837-870.
- [19] Wellner, J.A., A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics, *Ann. Statist.* **5** (1977), 473-480.
- [20] Woyczyński, W.A., Geometry and martingales in Banach spaces. Part II: independent increments, *Advances in Probability*, Dekker, New York **4** (1978), 267-527.
- [21] Zaanen, A.C., *Riesz Spaces II*, North-Holland Publishing Company, Amsterdam, New York, Oxford 1983.
- [22] van Zwet, W.R., A strong law for linear functions of order statistics, *Ann. Probab.* **8** (1980), 986-990.