Probabilistic norms and convergence of random variables

Bernardo Lafuerza-Guillén and Carlo Sempi

a Departamento de Estadística y Matemática Aplicada, Universidad de Almería, 04120 Almería, Spain
b Dipartimento di Matematica “Ennio De Giorgi”, Università di Lecce, Lecce, Italy 73100

Received 5 March 2001
Submitted by U. Stadtmueller

Abstract

We prove that the probabilistic norms of suitable Probabilistic Normed spaces induce convergence in probability, $L^p$ convergence and almost sure convergence.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Convergence in probability; Almost sure convergence; Complete convergence; Probabilistic normed space; Probabilistic norm

1. History and introduction

The study of possible topologies for the various modes of convergence of sequences of random variables has a long history.

Let a probability space $(\Omega, \mathcal{A}, P)$ be given and let $L^0(\mathcal{A})$ be the linear space of (equivalence classes of) $E$-valued random variables, viz. $E$-valued measurable functions defined on it. Here $(E, | \cdot |)$ is a normed space.

For a sequence $\{f_n: n \in \mathbb{N}\}$ we shall consider the following modes of convergence: convergence in probability, convergence in $L^p$ and almost sure convergence.

Ky Fan [7] showed that the topology of convergence in probability can be metrized and introduced the metric that now bears his name. Fréchet [6] exhibited several metrics whose

---

* The research here reported was carried out as a part of the program “Processi Stocastici” of the Italian M.U.R.S.T.
* Corresponding author.
  E-mail addresses: blafuerz@ual.es (B. Lafuerza-Guillén), sempi@ilenic.unile.it (C. Sempi).

0022-247X/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0022-247X(02)00577-2
topology coincides with that of convergence in probability. Later, it was proved by Dugué [3], Marczewski [14], Thomasian [20,21] that, in general, no norm exists that generates this topology. Moreover, in general, there exists no metric whose convergence coincides with almost sure convergence [4]. These results are summarized in [13].

Finally, Fernique [5], using and generalizing a famous theorem of Skorohod’s [19] showed how one can build a probability space and a proper subset of $L^0(\mathcal{A})$ in which almost sure convergence can be defined through a metric topology.

As far as $L^p$ convergence is concerned, we shall make reference to the paper [10], but see also the previous paper [9].

We now plan to examine the whole question from the point of view of probabilistic normed spaces; we shall then recognize that there are probabilistic norms that imply either type of convergence.

A probabilistic normed space (PN space in the following) $(V, v, \tau, \tau^*)$ is a quadruple in which $V$ is a real linear space, the probabilistic norm $v$ is a mapping from $V$ into $\Delta^+$, the space of the distance distribution functions (i.e., those distribution functions $F$ that vanish at the origin, $F(0) = 0$), and $\tau$ and $\tau^*$, with $\tau \leq \tau^*$ are triangle functions; $\nu, \tau$ and $\tau^*$ are subject to the following conditions:

(N1) $v_p = \delta_0$ if, and only if, $p = \theta$ (the null vector of $V$);
(N2) $v_{-p} = v_p$ for every $p \in V$;
(N3) for all $p$ and $q$ in $V$, $v_{p+q} \geq \tau(v_p, v_q)$;
(N4) for every $p \in V$ and for every $\alpha \in [0, 1]$, $v_p \leq \tau^*(v_{\alpha p}, v_{(1-\alpha)p})$.

An important class of PN spaces is that of Šerstnev spaces: a PN space is called a Šerstnev space if (N1) and (N3) are satisfied along with the following condition:

\[ \forall \alpha \in \mathbb{R} \setminus \{0\} \forall \gamma > 0 \forall \nu_{\alpha p}(x) = v_p \left( \frac{x}{|\alpha|} \right), \quad (\tilde{S}) \]

which implies (N2) and (see [1]) (N4) in the strengthened form

\[ \forall \alpha \in [0, 1] \forall \nu \in V \nu_p = \tau_M(v_{\alpha p}, v_{(1-\alpha)p}). \quad (N4') \]

A Šerstnev space will be denoted by $(V, v, \tau)$, since the rôle of $\tau^*$ is played by a fixed triangle function, $\tau_M$, which then satisfies (N4').

For more information about PN spaces and for their properties see [1,9–12]. PN spaces were first introduced in a slightly less general form in [17]. The notation adopted in the present note is mainly from the book [16]. A PN space is endowed with the strong topology (see [16]) and this latter is metrizable. A sequence $\{p_n\}$ of elements of $V$ converges to $0_V$, the null element of $V$, in the strong topology if, and only if,

\[ d_S(v_{p_n}, \delta_0) \xrightarrow{n \to +\infty} 0, \]

where $d_S$ denotes the Lévy metric as modified by Sibley [15,18].
2. Convergence in probability

In the notation we have fixed in the previous section, let \( \nu : S \to \Delta^+ \) be defined, for every \( f \in L^0(A) \) and for every \( x \in \mathbb{R}_+ \), by
\[
\nu_f(x) := P\{ \omega \in \Omega : |f(\omega)| < x \},
\]
(1)
The couple \((L^0(A), \nu)\) is called an EN space. It was proved in [9] that the relationship, defined in the EN space \((L^0(A), \nu)\) by \( f \sim g \) if, and only if, \( \nu_f = \nu_g \) is an equivalence relationship. Moreover, if \( L^0(A) := L^0(A)/\sim \) is the quotient space and if the quotient mapping \( \tilde{\nu} : L^0(A) \to \Delta^+ \) is defined via
\[ \tilde{\nu}_f := \nu_f \]
for every \( f \) in the equivalence class \( \tilde{f} \), then \((L^0(A), \tilde{\nu})\) is a PN space under the triangle functions \( \tau_W \) and \( \tau_M \) [16]; in fact, more is true: \((L^0(A), \tilde{\nu}, \tau_W)\) is a Šerstnev space [8, Theorem 2.2.8].

As usual in probability theory, one writes \( f \) even when one means the equivalence class \( \tilde{f} \) of \( f \). The equivalence between convergence in probability and convergence with respect to the probabilistic norm \((1)\) is now almost obvious.

**Theorem 1.** For a sequence of (equivalence classes of) \( E \)-valued random variables \( \{f_n\} \), the following statements are equivalent:

(a) \( \{f_n\} \) converges in probability to \( \theta_S \), \( f_n \xrightarrow{P} \theta_S \);
(b) the corresponding sequence \( \{\nu_{f_n}\} \) of probabilistic norms converges weakly to \( \varepsilon_0 \), viz.
\[
d_S(\nu_{f_n}, \varepsilon_0) \xrightarrow{n \to +\infty} 0;
\]
(c) \( \{f_n\} \) converges to \( \theta_S \) in the strong topology of the Šerstnev space \((L^0, \nu, \tau_W)\).

**Proof.** Since (b) and (c) are equivalent by definition, it suffices to establish the equivalence of (a) and (b).

The sequence \( \{f_n\} \) converges to \( \theta_S \) in probability if, and only if, for every \( x > 0 \),
\[
\lim_{n \to +\infty} P(|f_n| < x) = 1,
\]
or, equivalently, on account of \((1)\), if, and only if, for every \( x > 0 \),
\[
\lim_{n \to +\infty} \nu_{f_n}(x) = 1.
\]
But this latter statement, in its turn, is equivalent (see [16]) to
\[
\lim_{n \to +\infty} d_S(\nu_{f_n}, \varepsilon_0) = 0,
\]
which proves the theorem. \( \Box \)

Of course, there is nothing special about \( \theta_S \) as a limit; if one wishes to consider the convergence in probability of the sequence \( \{f_n\} \) to the \((E\)-valued) random variable \( f \), then it suffices to consider the sequence \( \{f_n - f\} \) and its convergence to \( \theta_S \).
3. \( L^p \) convergence

In order to consider convergence in \( L^p \) with \( p \in [1, +\infty] \), the following result connecting the \( L^p \) norms with the probabilistic norm (1) will be needed (see [10]).

**Theorem 2.** For \( p \in [1, +\infty] \) the norm \( \| \cdot \|_p \) in the space

\[
L^p(A) := \left\{ f \in L^0(A) : \int |f|^p \, dP < +\infty \right\}
\]

is given, for \( f \) in \( L^p(A) \), by

\[
\| f \|_p = \left( \int_{\mathbb{R}^+} t^p \, d\nu_f(t) \right)^{1/p};
\]

for \( p = +\infty \), the norm \( \| \cdot \|_\infty \) in \( L^\infty(A) := \{ f \in L^0(A) : \text{ess sup} |f| < +\infty \} \) is given, for \( f \) in \( L^\infty(A) \), by

\[
\| f \|_\infty = \sup\{ t > 0 : \nu_f(t) < 1 \}.
\]

With the help of the previous result one can characterize \( L^p \) convergence. As in the previous section, there is no loss of generality in considering only convergence to \( \theta_S \), for, if one wishes to study the convergence of a sequence \( \{ f_n \} \) to \( f \neq \theta_S \), it suffices to replace \( \{ f_n \} \) by \( \{ f_n - f \} \).

**Theorem 3.** For a sequence of (equivalence classes of) \( E \)-valued random variables \( \{ f_n \} \) in \( L^p \), the following statements are equivalent:

if \( p \in [1, +\infty] \):

(a) \( \{ f_n \} \) converges to \( \theta_S \) in \( L^p \), \( f_n \xrightarrow{L^p} \theta_S \);

(b) the sequence of the \( p \)-th moments of the probabilistic norms \( \{ \nu_{f_n} \} \) tends to 0, viz.

\[
\int_{\mathbb{R}^+} t^p \, d\nu_{f_n}(t) \xrightarrow{n \to +\infty} 0;
\]

if \( p = +\infty \):

(c) \( \{ f_n \} \) converges to \( \theta_S \) in \( L^\infty \), \( f_n \xrightarrow{n \to +\infty} \theta_S \);

(d) for every \( t > 0 \), the sequence \( \{ \nu_{f_n}(t) \} \) is definitely equal to 1, viz., for all \( t > 0 \), there exists \( n_0 = n_0(t) \in \mathbb{N} \) such that \( \nu_{f_n}(t) = 1 \) if \( n \geq n_0 \).

**Proof.** Only the equivalence of (c) and (d) is not immediately obvious; thus we limit ourselves to proving it.
(c) ⇒ (d) Assume $\|f_n\|_\infty \xrightarrow{n \to \infty} 0$ and let $t$ be strictly positive; then, for every $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon)$ in $\mathbb{N}$ such that, for every $n > n_0$, one has

$$\sup \{ t > 0 : v_{f_n}(t) < 1 \} = \|f_n\|_\infty < \varepsilon,$$

so that $v_{f_n}(\varepsilon) = 1$; but then, for all $n > n_0$, $v_{f_n}(t) > v_{f_n}(\varepsilon) = 1$.

(d) ⇒ (c) For $t > 0$, let $n_0 = n_0(t) \in \mathbb{N}$ be such that $v_{f_n}(t) = 1$ for all $n > n_0$; therefore, if $n > n_0$, $\|f_n\|_\infty < t$, which yields $\|f_n\|_\infty \xrightarrow{n \to \infty} 0$. \(\square\)

In comparing Theorem 3 with Theorem 1, it should be noticed that, if the sequence $\{v_{f_n}(t)\}$ is definitely equal to 1 for every $t > 0$, then, a fortiori, one has $v_{f_n}(t) \xrightarrow{n \to \infty} 1$, or, equivalently $d_\delta(v_{f_n}, 0) \xrightarrow{n \to \infty} 0$; this is the translation in the language of PN spaces of the fact that convergence in $L^\infty$ implies convergence in probability. The converse is not true, as is well known: and, in fact, a sequence $\{v_{f_n}(t)\}$ may well converge to 1 without being definitely equal to 1.

4. Almost sure convergence

Consider the family $V := \{L^0(\mathcal{A})\}^\mathbb{N}$ of all the sequences of (equivalence classes of) $E$-valued random variables. The set $V$ is a real vector space with respect to the componentwise operations; specifically, if $s = \{f_n\}$ and $s' = \{g_n\}$ are two sequences in $V$ and if $\alpha$ is a real number, then the sum $s \oplus s'$ of $s$ and $s'$ and the scalar product $\alpha \odot s$ of $\alpha$ and $s$ are defined via

$$s \oplus s' = \{f_n\} \oplus \{g_n\} := \{f_n + g_n\}, \quad \alpha \odot s = \alpha \odot \{f_n\} := \{\alpha f_n\}.$$

A mapping $\phi : V \to \Delta^+$ will be defined on $V$ via

$$\phi_\delta(x) := P\left( \sup_{k \in \mathbb{N}} |f_k| < x \right) = P\left( \bigcap_{k \in \mathbb{N}} \{|f_k| < x\} \right),$$

where $x > 0$ and $s = \{f_k\}$. Then one can prove

**Theorem 4.** The triple $(V, \phi, \tau_W)$ is a Šerstnev space.

**Proof.** Let $O$ denote the null sequence, $O := \{o_n\}$, where, for every $n \in \mathbb{N}$, $o_n \equiv \theta_E$; here, $\theta_E$ is the null vector of $E$. Obviously, $O$ is the null element of the vector space $V$, $O = \theta_V$ and $\phi_O = \varepsilon_0$. In the other direction, assume $\phi_\delta = \varepsilon_0$; then one has, for every $x > 0$,

$$\phi_\delta(x) := P\left( \sup_{k \in \mathbb{N}} |f_k| < x \right) = 1,$$

so that, for every $x > 0$ and for every $k \in \mathbb{N}$,

$$P(|f_k| < x) = 1;$$

in other words, $v_{f_k} = \varepsilon_0$ for every $k \in \mathbb{N}$. As already noticed in the previous section, $(L^0, v, \tau_W)$ is a Šerstnev space; therefore, $f_k = \theta_E$ for all $k \in \mathbb{N}$, or, equivalently, $s = O$, which concludes the proof of (N1).
Property (N2) is obvious. Notice that, for $\alpha \in \mathbb{R}$, $\alpha \neq 0$, for every $s \in V$ and for all $x > 0$, one has
\[
\phi_{x,\alpha}(s)(x) = \mathbb{P}\left( \sup_{k \in \mathbb{N}} |\alpha f_k| < x \right) = \mathbb{P}\left( \sup_{k \in \mathbb{N}} |f_k| < \frac{x}{|\alpha|} \right) = \phi_x \left( \frac{x}{|\alpha|} \right).
\]
This proves (5).

For all sequences $s = \{f_k\}$ and $s' = \{g_k\}$, for all $x > 0$ and $t \in [0, x]$, one has
\[
P\left( \bigcup_{k \in \mathbb{N}} |f_k + g_k| \geq x \right) \leq P\left( \bigcup_{k \in \mathbb{N}} |f_k| \geq t \right) + P\left( \bigcup_{k \in \mathbb{N}} |g_k| \geq x - t \right)
\]
so that
\[
P\left( \bigcap_{k \in \mathbb{N}} |f_k + g_k| < x \right) \geq P\left( \bigcap_{k \in \mathbb{N}} |f_k| < t \right) + P\left( \bigcap_{k \in \mathbb{N}} |g_k| < x - t \right)
\]
\[
\geq 1 - P\left( \bigcup_{k \in \mathbb{N}} |f_k| \geq t \right) - P\left( \bigcup_{k \in \mathbb{N}} |g_k| \geq x - t \right)
\]
\[
+ P\left( \bigcap_{k \in \mathbb{N}} |f_k| < t \right) + P\left( \bigcap_{k \in \mathbb{N}} |g_k| < x - t \right)
\]
\[
\geq P\left( \bigcap_{k \in \mathbb{N}} |f_k| < t \right) + P\left( \bigcap_{k \in \mathbb{N}} |g_k| < x - t \right) - 1.
\]
Thus, one has, for every $t \in [0, x]$,\[
\phi_{x,\alpha'}(x) \geq \phi_x(t) + \phi_{x'}(x - t) - 1;
\]
therefore\[
\phi_{x,\alpha'} \geq \tau_W(\phi_x, \phi_{x'}),
\]
which proves (N3) and concludes the proof. $\square$
Given an element \( s \) of \( V \), viz. given a sequence \( s = \{ f_k : k \in \mathbb{N} \} \) of \( E \)-valued random variables, \( f_k \in L^0(A) \) for every \( k \in \mathbb{Z}_+ := \{0, 1, \ldots \} \), consider the \( n \)-shift \( s_n \) of \( s \), \( s_n := \{ f_{k+n} : k \in \mathbb{N} \} \), which again belongs to \( V \). We are now in a position to state our main result.

**Theorem 5.** A sequence \( s = \{ f_k : k \in \mathbb{N} \} \) of \( E \)-valued random variables converges almost surely to \( \theta_E \), the null vector of \( E \), if, and only if, the sequence \( \{ \phi_n : n \in \mathbb{Z}_+ \} \) of probabilistic norms of the \( n \)-shifts of \( s \) converges weakly to \( \varepsilon_0 \), or, equivalently, if, and only if, the sequence \( \{ s_n \} \) of the \( n \)-shifts of \( s \) converges to \( O := \{ \theta_E, \theta_E, \ldots \} \) in the strong topology of \( (V, \phi, \tau_W) \).

**Proof.** All the statements are equivalent to the assertion, which holds for every \( x > 0 \),

\[
1 = \lim_{n \to +\infty} \phi_{n_0}(x) = \lim_{n \to +\infty} P \left( \bigcap_{k \in \mathbb{N}} \{ |f_{k+n}| < x \} \right) = \lim_{n \to +\infty} P \left( \bigcap_{k \geq n} \{ |f_k| < x \} \right) = P \left( \liminf_{n \to +\infty} \{ |f_k| < x \} \right).
\]

This proves the result. \( \square \)

Given an element \( s \) in \( V \), define \( n : V \to \Delta^+ \) via

\[
n_{s}(x) := \sup_{n \in \mathbb{N}} \phi_{n_0}(x) = \lim_{n \to +\infty} P \left( \bigcap_{k \geq n} \{ |f_k| < x \} \right);
\]

then Theorem 5 says that \( s \) converges to \( \theta_E \) if, and only if, \( n_{s} = \varepsilon_0 \). An equivalence relationship can be defined on \( V \) by stipulating that \( s = \{ f_n \} \sim s' = \{ g_n \} \) if, and only if, \( \{ f_n - g_n \} \) converges almost surely to the null element \( O := \{ \theta_E, \theta_E, \ldots \} \) of \( V \), or, equivalently, if, and only if, \( n_{s} = \varepsilon_0 \). In the quotient space \( V_0 := V / \sim \), the constant sequence \( \{ f, f, \ldots \} \), where \( f \) is in \( L^0(A) \), is the family of the sequences of \( E \)-valued random variables that converge almost surely to \( f \).

As a final remark, notice that we have shown that the strong topology of the Šerstnev space \( (V, \phi, \tau_W) \) induces a convergence equivalent to almost sure convergence in \( L^0(A) \). This does not contradict the well-known fact that, in general, almost sure convergence does not derive from a topology on \( L^0(A) \).

**References**