A Methodological Note on the Convergence of Sequences of Random Variables

Salvador Cruz Rambaud¹ & Antonio Luis Rodríguez López-Cañizares²

¹ Departamento de Dirección y Gestión de Empresas, University of Almería, La Cañada de San Urbano, Almería, Spain
² Departamento de Estadística e Investigación Operativa, University of Granada, Campus de Fuentenueva, Granada, Spain

Correspondence: Salvador Cruz Rambaud, Departamento de Dirección y Gestión de Empresas, University of Almería, La Cañada de San Urbano, s/n, Almería 04071, Spain. Tel: 34-950-015-184. E-mail: scruz@ual.es

Received: April 19, 2012 Accepted: May 7, 2012 Online Published: June 15, 2012
doi:10.5539/ijsp.v1n2p35 URL: http://dx.doi.org/10.5539/ijsp.v1n2p35

The research is financed by Proyectos de Excelencia de la Junta de Andalucía (FEDER). No. P09-SEJ-05404

Abstract

The aim of this note is to present a new point of view for introducing all well-known modes of convergence of sequences of random variables. In the one hand, we start from two noteworthy sets in convergence viz $T_{j,\varepsilon}$ and $S_{j,\varepsilon}$. The consideration of certain progressive assumptions on both $T_{j,\varepsilon}$ and $S_{j,\varepsilon}$ gives rise to a part of convergence concepts going from uniform convergence to convergence in probability. On the other hand, some key inequalities implies the rest of convergence concepts whose link with the former scheme lies in uniform and in probability convergence which end the circle of convergence modes. Moreover, all these steps are illustrated with their respective methodological charts.

Keywords: convergence, uniform, almost sure, complete, in probability, in mean

1. Introduction

Through time, there have been lot of textbooks describing the main modes of convergence of random variables sequences. Thus, the reader can find a complete survey of the different modes of convergence in Parzen (1954), Loève (1963), Neveu (1970), Rohatgi (1976), or Resnick (1999). One can say that all approaches coincide in introducing the different concepts and then deduce the relationships between them. In our opinion, one of the most systematized book is Karr (1993) where the modes of convergence have been classified in the following way:

- Convergence of random variables as functions.
- Convergence of distribution functions.
- Relationships among the modes
  - Implications always valid.
  - Counterexamples.
  - Implications of restricted validity.
  - Implications involving subsequences.
- Convergence of algebraic operations.
- Continuous mappings of random variables.
- Convergence of random vectors.

More recently, Kotz et al. (2006), based on Lukács, (1975), includes a complete survey of convergence modes in their Encyclopedia of Statistical Sciences, 2nd ed.

The organization of this paper is as follows. After this introduction, Section 2 presents all the well-known modes
of convergence through their respective theorem of characterization. The main novelty of our approach is that, before defining and characterizing all modes of convergence, we previously know the implications between them. This is because we introduce all known modes of convergence by progressive restrictions on both $T_{j,\epsilon}$ and $S_{j,\epsilon}$.

Finally, Section 3 summarizes and concludes.

2. Survey of Convergence of Sequences of Random Variables

When dealing with stochastic convergence of a sequence of random variables $\{X_t\}_{t=1}^{\infty}$ to $X$ (where $X_t, X: \Omega \rightarrow \mathbb{R}$), some noteworthy sets appear (we will follow Lukács, 1975, notation):

$$T_{j,\epsilon} = \{\omega \in \Omega : |X_j(\omega) - X(\omega)| \leq \epsilon\},$$

simply denoted by $[|X_j - X| \leq \epsilon]$, and

$$S_{t,\epsilon} = \bigcap_{j=t}^{\infty} T_{j,\epsilon}.$$  

With sets $S_{t,\epsilon}$ and $T_{j,\epsilon}$, we can write the following chain of statements each of them implying the following one:

1) For every $\epsilon > 0$, there exists an integer $t$, such that $S_{t,\epsilon} = \Omega$.

2) For every $\epsilon > 0$, there exists an integer $t$, such that $P(S_{t,\epsilon}) = 1$.

3) For every $\epsilon > 0$, $\lim_{t \to \infty} P(S_{t,\epsilon}) = 1$.

4) For every $\epsilon > 0$, $\lim_{t \to \infty} P(T_{j,\epsilon}) = 1$.

5) For every $\epsilon > 0$ and $\delta > 0$, there exists an integer $t$, such that, for any $x \in \mathbb{R}$, $|F_t(x) - F(x)| \leq [F(x + \epsilon) - F(x - \epsilon)] + \delta$, being $F_t$ and $F$ the distribution function of $X_t$ and $X$, respectively.

Observe that statement 3 is equivalent to:

3) For every $\epsilon > 0$, $\lim_{t \to \infty} P(S_{t,\epsilon}) = 0$.

and then $\lim_{t \to \infty} P\left(\bigcup_{j=t}^{\infty} T_{j,\epsilon}\right) = 0$. By the Countable Subadditivity Theorem (Billingsley, 1995, p. 25),

$$P\left(\bigcup_{j=t}^{\infty} T_{j,\epsilon}\right) \leq \sum_{j=t}^{\infty} P(T_{j,\epsilon}).$$

So three new statements stronger than 3 can be introduced:

2') For every $\epsilon > 0$, $\lim_{t \to \infty} \sum_{j=t}^{\infty} P(T_{j,\epsilon}) = 0$.

2'') There exists a sequence of positive numbers $\{\epsilon_t\}_{t=1}^{\infty}$ such that

$$\lim_{t \to \infty} \epsilon_t = 0 \quad \text{and} \quad \sum_{t=1}^{\infty} P(T_{t,\epsilon_t}) < \infty.$$

2'''') There exists a sequence of positive numbers $\{\epsilon_t\}_{t=1}^{\infty}$ such that

$$\sum_{j=1}^{\infty} \epsilon_t < \infty \quad \text{and} \quad \sum_{t=1}^{\infty} P(T_{t,\epsilon_t}) < \infty.$$

With respect to statement #1, we can enunciate

Theorem 1 The following two conditions are equivalent:

(i) For every $\epsilon > 0$, there exists an integer $t$, such that $S_{t,\epsilon} = \Omega$.

(ii) For every $\epsilon > 0$, there exists an integer $t$, such that $|X_j(\omega) - X(\omega)| \leq \epsilon$, for every $j \geq t$ and $\omega \in \Omega$.

Proof. It is obvious taking into account the definition of the set $S_{t,\epsilon}$.

Definition 1 A sequence of random variables $\{X_t\}_{t=1}^{\infty}$ is said to be uniformly convergent to $X$, denoted by $X_t \xrightarrow{u} X$, if it satisfies any of the two equivalent conditions in Theorem 1.
With respect to statement #2, we can enunciate

**Theorem 2** The following two conditions are equivalent:

(i) For every \( \epsilon > 0 \), there exists an integer \( t \), such that \( P(S_{t, \epsilon}) = 1 \).

(ii) There is a set \( A \subseteq \Omega \), \( P(A) = 1 \), such that, for every \( \epsilon > 0 \), there is an integer \( t_\epsilon \) such that, for \( t \geq t_\epsilon \), then

\[
|X_t(\omega) - X(\omega)| \leq \epsilon,
\]

for every \( \omega \in A \).

**Proof.** See Cruz (2011, p. 215-221).

**Definition 2** A sequence of random variables \( \{X_t\}_{t=1}^\infty \) is said to be *almost surely uniformly convergent* to \( X \), denoted by \( X_t \xrightarrow{a.s.u.} X \), if it satisfies any of the two equivalent conditions in Theorem 2.

With respect to statement #2’, we can enunciate

**Theorem 3** The following three conditions are equivalent:

(i) For every \( \epsilon > 0 \), \( \sum_{j=1}^{\infty} P(T_{j, \epsilon}) < \infty \).

(ii) For every \( \epsilon > 0 \), \( \lim_{t \to \infty} \sum_{j=1}^{t} P(T_{j, \epsilon}) = 0 \).

(iii) There exists a sequence of positive numbers \( \{\epsilon_t\}_{t=1}^{\infty} \) such that

\[
\lim_{t \to \infty} \epsilon_t = 0 \text{ and } \sum_{t=1}^{\infty} P(T_{t, \epsilon_t}) < \infty.
\]

**Proof.** (i) \( \Rightarrow \) (ii). It is obvious taking into account that

\[
\lim_{t \to \infty} \sum_{j=1}^{\infty} P(|X_j - X| > \epsilon) = \lim_{t \to \infty} \left( \sum_{j=1}^{\infty} P(|X_j - X| > \epsilon) - \sum_{j=1}^{t-1} P(|X_j - X| > \epsilon) \right) = 0.
\]

(ii) \( \Rightarrow \) (iii). By hypothesis, for every \( \epsilon > 0 \),

\[
\lim_{t \to \infty} \sum_{j=1}^{\infty} P(|X_j - X| > \epsilon) = 0.
\]

- For \( \epsilon = 1 \), there exists \( t_1 \), such that \( \sum_{j=1}^{\infty} P(|X_j - X| > 1) < \frac{1}{2} \), for all \( t \geq t_1 \).

- For \( \epsilon = \frac{1}{2} \), there exists \( t_2 \), such that \( \sum_{j=1}^{\infty} P(|X_j - X| > \frac{1}{2}) < \frac{1}{4} \), for all \( t \geq t_2 \).

- For \( \epsilon = \frac{1}{4} \), there exists \( t_3 \), such that \( \sum_{j=1}^{\infty} P(|X_j - X| > \frac{1}{4}) < \frac{1}{8} \), for all \( t \geq t_3 \).

- In general, for \( \epsilon = \frac{1}{r} \), there exists \( t_r \), such that \( \sum_{j=1}^{\infty} P(|X_j - X| > \frac{1}{r}) < \frac{1}{2r} \), for all \( t \geq t_r \).

Obviously, as \( \frac{1}{2} > \frac{1}{4} > \frac{1}{8} > \cdots > \frac{1}{2r} > \cdots \) and, as the corresponding probabilities are increasing, in order to reach smaller sums, it is necessary that \( t_1 < t_2 < t_3 < \cdots < t_r < \cdots \). Consider the following sequence of real numbers:

- \( \epsilon_1 = \epsilon_{t_1} = \cdots = \epsilon_{t_2-1} = 1 \).

- \( \epsilon_2 = \epsilon_{t_2} = \cdots = \epsilon_{t_3-1} = \frac{1}{2} \).

- \( \epsilon_3 = \epsilon_{t_3} = \cdots = \epsilon_{t_4-1} = \frac{1}{4} \).

- In general, \( \epsilon_r = \epsilon_{t_r} = \cdots = \epsilon_{t_{r+1}-1} = \frac{1}{2r} \).

Obviously, \( \{\epsilon_t\}_{t=0}^{\infty} \) converges to zero. On the other hand, for every \( \delta > 0 \), there exists a \( r \in \mathbb{N} \), such that \( \frac{1}{2r} < \delta \). In this case,

\[
\sum_{j=t_{r+1}}^{t_r-1} P(|X_j - X| > \epsilon_j = \epsilon_{t_{r+1}}) + \sum_{j=t_{r+2}}^{t_r-1} P(|X_j - X| > \epsilon_j = \epsilon_{t_{r+2}}) + \cdots < \\
< \sum_{j=t_{r+1}}^{\infty} P(|X_j - X| > \epsilon_j = \epsilon_{t_{r+1}}) + \sum_{j=t_{r+2}}^{\infty} P(|X_j - X| > \epsilon_j = \epsilon_{t_{r+2}}) + \cdots <
\]
Moreover, statement #2 (labelled \( \sum_{t=1}^{\infty} \epsilon_i \)-complete convergence and denoted by \( X_t \xrightarrow{\text{c}} X \)) implies condition (iii) in Theorem 3 and so complete convergence. It is obvious taking into account that \( \sum_{t=1}^{\infty} \epsilon_i < \infty \) implies \( \lim_{t \to \infty} \epsilon_i = 0 \). But, in general, the converse is not true, as the following example shows.

**Example** Let \( \Omega = [0, 1] \), for \( \mathcal{F} \) the Borel sets of this interval, and for \( P \) the Lebesgue measure. Consider the following sequence of random variables:

\[
X_t(\omega) = \begin{cases} 
0, & 0 \leq \omega \leq 1 - \frac{1}{t}, \quad t = 1, 2, \ldots \\
\frac{1}{t}, & 1 - \frac{1}{t} < \omega \leq 1 
\end{cases}
\]

For every \( \epsilon > 0 \), there exists a \( t_\epsilon \) such that \( \frac{1}{t} < \epsilon \), for all \( t \geq t_\epsilon \). Therefore,

\[
\sum_{t=1}^{\infty} P(|X_t| > \epsilon) < 1 + \frac{1}{2} + \cdots + \frac{1}{t_\epsilon - 1} < \infty.
\]

Thus, \( \{X_t\}_{t=1}^{\infty} \) converges completely to 0. On the other hand, let \( \{\epsilon_i\}_{i=1}^{\infty} \) be a sequence of positive real numbers such that \( \sum_{i=1}^{\infty} \epsilon_i < \infty \). So, the number of addends \( \epsilon_i \) lesser than their corresponding \( \frac{1}{t} \) is infinite, because, on the contrary, \( \sum_{i=1}^{\infty} \epsilon_i \) should be infinite. Thus,

\[
\sum_{i=1}^{\infty} P(|X_t| > \epsilon_i) > \sum_{i=1}^{\infty} \frac{1}{i} = \infty.
\]

Thus, \( \{X_t\}_{t=1}^{\infty} \) does not converge \((\sum_{i=1}^{\infty} \epsilon_i)\)-completely to 0.

With respect to statement #3, we can enunciate

**Theorem 4** The following three conditions are equivalent:

(i) For every \( \epsilon > 0 \), \( \lim_{t \to \infty} P(S_t, \epsilon) = 1 \).

(ii) \( P(\lim_{t \to \infty} X_t = X) = 1 \).

(iii) For every \( \epsilon > 0 \), there exists a set \( A \in \mathcal{F} \) such that \( P(\overline{A}) < \epsilon \) and \( X_t \xrightarrow{a.s.} X \) on \( A \).

**Proof.** See Theorem 2.1.2 in Lukács (1975, p. 30) and Egoroff’s Theorem in Aliprantis (2006).

**Definition 4** A sequence of random variables \( \{X_t\}_{t=1}^{\infty} \) is said to be almost surely convergent to \( X \), denoted by \( X_t \xrightarrow{a.s.} X \), if it satisfies any of the two first equivalent conditions in Theorem 4.

**Definition 5** A sequence of random variables \( \{X_t\}_{t=1}^{\infty} \) is said to be almost uniformly convergent to \( X \), denoted by \( X_t \xrightarrow{a.u.} X \), if it satisfies the third condition in Theorem 4.

With respect to statement #4, we can enunciate

**Theorem 5** The following two conditions are equivalent:

(i) For every \( \epsilon > 0 \), \( \lim_{t \to \infty} P(T_{t, \epsilon}) = 1 \).

(ii) For every \( \epsilon > 0 \), \( \lim_{t \to \infty} P(|X_t - X| > \epsilon) = 0 \).

**Proof.** It is obvious.
**Definition 6** A sequence of random variables \( \{X_t\}_{t=1}^{\infty} \) is said to be *convergent in probability* to \( X \), denoted by \( X_t \xrightarrow{P} X \), if it satisfies any of the two equivalent conditions in Theorem 5.

On the other hand, the implication \( \#4 \Rightarrow \#5 \) can be seen in Theorem 1.2.2. of (Lukács, 1975, p. 8) and, with respect to statement \( \#5 \), we can enunciate

**Theorem 6** *For every \( x \) where \( F \) is continuous, the following two conditions are equivalent:*

(i) For every \( \epsilon > 0 \) and \( \delta > 0 \), there exists an integer \( t \), such that \( |F_t(x) - F(x)| \leq [F(x + \epsilon) - F(x - \epsilon)] + \delta \).

(ii) For every \( \epsilon > 0 \), there exists an integer \( t \), such that \( |F_t(x) - F(x)| \leq \epsilon \).

**Proof.** (i) \( \Rightarrow \) (ii). It is obvious taking into account the continuity of the function \( F \).

(ii) \( \Rightarrow \) (i). It suffices to take the value of \( \epsilon \) as \( [F(x + \epsilon) - F(x - \epsilon)] + \delta \).

**Definition 7** A sequence of random variables \( \{X_t\}_{t=1}^{\infty} \) is said to be *convergent in distribution* to \( X \), denoted by \( X_t \xrightarrow{d} X \), if it satisfies any of the two equivalent conditions in Theorem 6.

Summarizing:

![Chart 1. Implications among general modes of convergence](image)

Next, we can write the following chain of statements each of them implied by the following one:

1) For every \( \epsilon > 0 \) and \( \delta > 0 \), there exists an integer \( t \), such that \( P(|X_t - X| \leq \epsilon) \geq 1 - \delta \).

2) For every \( s > 1 \), \( \lim_{t \to \infty} E_p|X_t - X|^s = 0 \).

3) For every \( r > s > 1 \), \( \lim_{t \to \infty} E_p|X_t - X|^r = 0 \).

Observe that

- Statement \( \#1 \) is equivalent to convergence in probability.
• Statement #2 (*convergence in the s-th mean*) implies convergence in probability, because the following relation (Chebyshev’s inequality):

\[
0 \leq P(|X_t - X| \geq \varepsilon) \leq \frac{E_P(|X_t - X|_s)}{\varepsilon^s},
\]

holds for any \( \varepsilon > 0 \).

• Statement #3 (*convergence in the r-th mean*) implies convergence in the s-th mean, because of the following relation:

\[
E_P(|X_t - X|_r) \leq [E_P(|X_t - X|_s)]^{s/r}.
\]

Finally, convergence in the r-th mean implies convergence of r-th moments:

\[
E_P|X_t|^r \rightarrow E_P|X|^r.
\]

Summarizing:

![Chart 2. Implications among modes of convergence of r.v. as functions](chart2.png)

*Chart 2. Implications among modes of convergence of r.v. as functions*

Chart 3 shows the implications always valid among the different modes of convergence. All these issues can be found in Karr (1993), Resnick (1999), and Rohatgi (1976). Observe that this new chart is the association of Charts 1 and 2 through the uniform convergence and convergence in probability.

Finally, in order to complete Chart 3, it is necessary to point out that the counterexamples for the implications that do not hold can be found in Karr (1993).
3. Conclusion

Convergence of sequences of random variables is an important issue in Mathematical Statistics. Its development can be justified by the need of defining a characteristic mode of convergence for statistical problems. In spite of the importance of almost sure convergence, there have been a proliferation of several other convergence concepts which have not had a common origin. In order to avoid this inconvenience, this paper uses two methodological resources. Firstly, we focus our attention on two noteworthy sets in convergence, $T_{j,\epsilon}$ and $S_{j,\epsilon}$, and secondly we use some key inequalities involving random variables. As a result, we can introduce all convergence concepts “in a natural way” with an added effect: this methodology allows us to see the proximity between two “consecutive concepts” and so to check the possibility of introduce a novel concept between them. Therefore, the methodological value of this paper can be located in both teaching and research processes.

Acknowledgment

We are very grateful for the comments and suggestions of two anonymous referees.

References


