EPIsciences

Communications in Mathematics **31** (2023), no. 1, 131–135
DOI: https://doi.org/10.46298/cm.10292
©2023 Yu-Lin Chou
This is an open access article licensed under the CC BY-SA 4.0

A general weak law of large numbers for sequences of L^p random variables

Yu-Lin Chou

Abstract. Without imposing any conditions on dependence structure, we give a seemingly overlooked simple sufficient condition for L^p random variables $X_1, X_2, \dots$ with given $1 \le p \le +\infty$ to satisfy

$$\frac{1}{a_n}\sum_{i=1}^{b_n}(X_i - \mathsf{E}X_i) \stackrel{L^p}{\to} 0 \quad \text{as } n \to \infty,$$

where $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ are pre-specified unbounded sequences of positive integers. Some unexpected convergences of sample means follow.

A law of large numbers is usually obtained by controlling both the dependence structure and the distributional homogeneity (including moment conditions here) of the underlying sequence of random variables. For classical treatments, one may refer to Etemadi [5] or Folland [6]; for more recent treatments, Chen and Sung [3] or Seneta [9]. The prototypical, most popular version of a weak law of large numbers is certainly the classical weak law asserting in-probability vanishing of sample means of independent identically distributed L^2 centered random variables.

In the related literature, there are works giving weak laws that are "non-typically" general in different directions. For instance, Loève [7] (p. 26) gives a necessary and sufficient condition for sample means of Bernoulli random variables, not necessarily independent, to obey a weak law; and Adler *et al.* [1] gives a weak law (in a suitable sense) for a class of independent random elements, whose moments need not exist, of a class of Banach spaces.

On the other hand, there are known laws of large numbers asserting L^p-vanishing of suitably scaled partial sums of centered random variables for special values of p. For instance, the classical Khintchine's theorem ensures L^1-vanishing of sample means of centered L^1 random variables under suitable conditions controlling both dependence structure and

MSC 2020: 60F05; 60F25
Keywords: convergence in probability, L^p-convergence, laws of large numbers
Affiliation:
Yu-Lin Chou – freelance author, Hsinchu County, Taiwan (R.O.C.)
E-mail: chou.y.l.edu@gmail.com

distributional homogeneity; the classical Markov's theorem asserts (under suitable conditions) the L^2-vanishing of the n^{-2}-scaled partial sums of centered L^2 random variables; and Lemma 1.5.1 in Chandra [2] asserts L^2-vanishing of sample means of uniformly bounded pairwise-independent centered L^2 random variables.

However, except possibly for the simple cases such as Bernoulli random variables, there seems not a general weak law for random variables, in the present evident sense, completely dropping control over dependence structure and at the same time offering a tractable sufficient condition. For instance, Theorem 1.2.2a in Révész *et al.* [8] asserts (in particular) in-probability vanishing of sample means of arbitrary random variables $X_1, X_2, \dots$ under the condition that the series $\sum_i i^{-1} X_i$ converges almost surely.

Independently of the related existing literature, we wish to give an overlooked law of large numbers suggested instead by the mathematical nature of the summation operators, which, without any dependence assumption, asserts in particular a generic weak law for random variables with finite mean (a condition being "negligible" in general) under precisely one simple distributional homogeneity condition in terms of the absolute first moments of the underlying random variables:

Theorem 1. *Given a probability space with* P *denoting the given probability measure, let* $1 \leq p \leq +\infty$*; let* $X_1, X_2, \dots \in L^p(\mathsf{P})$*; let* $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ *be unbounded sequences of positive integers. If* $a_n^{-1} \sum_{i=1}^{b_n} |X_i|_{L^p} \to 0$ *as* $n \to \infty$*, then*

$$\frac{1}{a_n} \sum_{i=1}^{b_n} (X_i - \mathsf{E}X_i) \xrightarrow{L^p} 0 \quad \text{as } n \to \infty.$$

Proof. By Minkowski's inequality we have

$$\left| \sum_{i=1}^{b_n} (X_i - \mathsf{E}X_i) \right|_{L^p} \leq \sum_{i=1}^{b_n} |X_i - \mathsf{E}X_i|_{L^p}$$

for all $n \in \mathbb{N}$.

Since $|f|_{L^r} \leq |f|_{L^\infty}$ for all $1 \leq r \leq +\infty$ and all $f \in L^r(\mathsf{P})$, Minkowski's and Jensen's inequalities (whenever suitable) jointly imply[1]

$$|X_i - \mathsf{E}X_i|_{L^p} \leq |X_i|_{L^p} + |\mathsf{E}X_i| \leq |X_i|_{L^p} + |X_i|_{L^1} \leq 2|X_i|_{L^p}$$

for each $i \in \mathbb{N}$. It follows that

$$\frac{1}{a_n} \sum_{i=1}^{b_n} |X_i - \mathsf{E}X_i|_{L^p} \leq \frac{2}{a_n} \sum_{i=1}^{b_n} |X_i|_{L^p}$$

for all n; but then the convergence assumption implies

[1] This observation appears in another preprint (Chou [4]) of the author for another purpose. At that time I did not observe the present observation, and it is evidently illogical to incorporate one of these works into the other. Mathematics happened to show itself in that way; I wrote it down.

$$\frac{1}{a_n}\sum_{i=1}^{b_n}(X_i - \mathsf{E}X_i) \overset{L^p}{\to} 0 \ \text{ as } n \to \infty.$$

This completes the proof. □

Remark 2. In Theorem 1, if $a_n = b_n = n$ for all n, then the sufficient condition may be replaced by the convergence $|X_i|_{L^p} \to 0$ as $i \to \infty$.

Moreover, Theorem 1 also holds for (a_n) an unbounded sequence of positive real numbers.

Corollary 3. *Given any probability space Ω with P denoting the given probability measure, let $X_1, X_2, \ldots$ be uniformly bounded random variables on Ω, i.e. such that $\sup_{i\in\mathbb{N}} |X_i| \le M$ on Ω for some (fixed) real M; let $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ be unbounded sequences of positive integers. If*

$$\frac{b_n}{a_n} \to 0 \quad \textit{as } n \to \infty,$$

then

$$\frac{1}{a_n}\sum_{i=1}^{b_n}(X_i - \mathsf{E}X_i) \overset{L^p}{\to} 0 \quad \textit{as } n \to \infty$$

for all $1 \le p \le +\infty$.

The potential utilities of Theorem 1 are further suggested in the following

Example 4. For each $x \in \mathbb{R}$, let δ_x be the Dirac measure $B \mapsto \mathbb{1}_B(x)$ on the Borel sigma-algebra of $\mathbb{R}$ concentrated at x. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of Rademacher-type random variables (on the same probability space) such that each X_i has $i^{-1}\delta_{-1} + (1 - i^{-1})\delta_{i^{-1}}$ as its distribution. Then $\mathsf{E}|X_i| = 2i^{-1} - i^{-2} \to 0$ as $i \to \infty$, and so

$$\frac{1}{n}\sum_{i=1}^{n}\mathsf{E}|X_i| \to 0 \ \text{ as } n \to \infty.$$

Since (X_i) is not necessarily independent and is by construction not identically distributed, no known law of large numbers seems to immediately assert a convergence of the sequence $(n^{-1}\sum_{i=1}^{n}(X_i - \mathsf{E}X_i))_{n\in\mathbb{N}}$, if not logically impossible. However, Theorem 1 asserts that

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathsf{E}X_i) \overset{L^1}{\to} 0 \ \text{ as } n \to \infty$$

and hence certainly

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathsf{E}X_i) \to 0 \text{ in probability}$$

as $n \to \infty$.

Example 5. Consider a sequence of normal random variables $(X_i)_{i\in\mathbb{N}}$ with mean zero such that each ξ_i has variance i^{-2}. Then $\mathsf{E}|X_i| = i^{-1}\sqrt{2/\pi}$ for all i, and so the sequence $X_1, X_2, \ldots$ of random variables satisfies the assumptions of Theorem 1.

The random variables $X_1, X_2, \ldots$ are not necessarily independent and are by construction not identically distributed, and hence the known laws of large numbers seem unable to assert a convergence of the sample mean of the centered random variables $X_i - \mathsf{E}X_i$. But its L^1-convergence and convergence in probability are ensured by Theorem 1.

Example 6. For a given sequence of L^p random variables $X_1, X_2, \ldots$ (with $1 \le p < +\infty$) that are identically distributed, Theorem 1 need not imply a convergence of the sample mean of the centered random variables $X_i - \mathsf{E}X_i$ (except for the trivial cases). However, since

$$\frac{1}{n^a}\sum_{i=1}^{n}|X_i|_{L^p} = \frac{1}{n^{a-1}}|X_1|_{L^p} \to 0 \ \text{ as } n \to \infty, \qquad \text{for all } a > 1,$$

Theorem 1 does assert the L^p-convergence of the sequence $(n^{-a}\sum_{i=1}^{n}(X_i - \mathsf{E}X_i))_n$ for all real $a > 1$. This covers some cases where the known laws of large numbers need not apply, e.g. where the dependence structure of (X_i) is unspecified.

Having given the above example, we construct another example for comparison. Let there be given some identically distributed sequence of non-negative L^1 random variables $\xi_1, \xi_2, \ldots$ with nonzero mean, and define $X_i := i\xi_i$ for all $i \in \mathbb{N}$. Then each X_i is L^1, and the sequence (X_i) is by construction non-identically distributed with an unspecified dependence structure; moreover, we have $\mathsf{E}|X_i| = i\mathsf{E}\xi_1 \to \infty$ as $i \to \infty$. Since

$$\frac{1}{n^{2+a}}\sum_{i=1}^{n}|X_i|_{L^1} \to 0 \ \text{ as } n \to \infty$$

for all real $a > 0$, Theorem 1 asserts for all real $a > 0$ the L^1-convergence of the random variables $n^{-2-a}\sum_{i=1}^{n}(X_i - \mathsf{E}X_i)$ as n goes beyond every bound. The existing laws of large numbers seem unable to assert this same conclusion.

For potential practical matters, we draw the following

Remark 7. The situations considered in the above examples would not be artificial. For instance, one may naturally consider certain types of observational (in contrast with "experimental") data as obtained from a time series of samples with finite mean for which it would be reasonable to assume that the absolute means of the samples vanish (at least in average) due to some systematic exogenous chronological structural factor such as continual technological advances over time. Thus Theorem 1 would also contribute to estimation or testing problems in the context of structural equation modeling.

For technical matters, we draw the following

Remark 8. One of the weak laws that are both technically friendly and application-friendly is the weak law for uncorrelated L^2 random variables, not necessarily identically distributed, whose n^{-2}-scaled partial sums of the variances of the first n random variables vanish, the conclusion being that the sample means of the centered random variables converge in probability. This weak law is certainly a special case of the Bernstein- Khintchine weak law (Theorem 1.5.1 in Chandra [2]).

The reader would then compare this common version of weak law with the implications of Theorem 1.

Acknowledgment

The author expresses gratitude for the helpful comments received.

References

[1] Adler A., Rosalsky A., and Taylor R. L.: A weak law for normed weighted sums of random elements in Rademacher type p Banach spaces. *Journal of Multivariate Analysis* 37 (2) (1991) 259–268.

[2] Chandra T. K.: *Laws of Large Numbers*. Narosa Publishing House (2012).

[3] Chen P. and Sung S. H.: A strong law of large numbers for nonnegative random variables and applications. *Statistics and Probability Letters* 118 (11) (2016) 80–86.

[4] Chou Y.-L.: Moderate laws of large numbers via weak laws. ResearchGate (2020).

[5] Etemadi N.: An elementary proof of the strong law of large numbers. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 55 (1981) 119–122.

[6] Folland G. B.: *Real Analysis: Modern Techniques and Their Applications*. Wiley (2007).

[7] Loève M.: *Probability Theory I*. Springer (1977).

[8] Révész P.: *The Laws of Large Numbers*. Academic Press (1968).

[9] Seneta E.: The weak law of large numbers for nonnegative summands. *Advances in Applied Probability* 50 (A) (2018) 241–252.

Received: February 20, 2021
Accepted for publication: April 02, 2021
Communicated by: Karl Dilcher