Sankhyā : The Indian Journal of Statistics
1989, Volume 51, Series A, Pt. 3, pp. 309–317.

UNIFORM INTEGRABILITY IN THE CESARO SENSE AND THE WEAK LAW OF LARGE NUMBERS

By TAPAS K. CHANDRA
Indian Statistical Institute

SUMMARY. It is shown here that L^1 convergence holds in large number of cases where the WLLN holds ; in fact, it is shown that the proof of the stronger fact is somewhat easier and more straightforward. In particular, several extensions and variations of the classical Khinchin WLLN are obtained.

The classical Khinchin's weak law of large numbers (WLLN) says that if $\{X_n\}$ is a sequence of independent and identically distributed random variables with finite $E(|X_1|)$, then $n^{-1}(X_1+\ldots+X_n)$ converges to $E(X_1)$ in probability ; actually 'mutual independence' can be replaced by 'pairwise independence' (see, e.g., Chung, 1974, Chapter 5). The usual proof of the above WLLN is due to Markov. Dharmadhikari (1976) gave an alternative (and somewhat simpler) proof of the same result ; in fact, he proved a slightly stronger result that $n^{-1}(X_1+\ldots+X_n)$ converges to $E(X_1)$ in L^1. It may be noted here that, under the above assumptions the strong law of large numbers also holds (see Etemadi, 1981).

The aim of the paper is to demonstrate that the L^1 convergence (and hence convergence in probability) of the sample mean holds under very general conditions. It is worth-mentioning that the proofs of Markov and Dharmadhikari use the truncation at levels $n\delta$ (see, e.g., Rao, 1973) and $n^{1/2}$ respectively ; this paper uses a different truncation.

Below the X_k are integrable random variables.

Definition. A sequence $\{X_n\}$ of random variables is said to be uniformly integrable in the Cesàro sense if

$$\lim_{a\to\infty} \sup_n \left\{ n^{-1} \sum_{k=1}^{n} \int_{|X_k|>a} |X_k|\, dP \right\} = 0.$$

Clearly, the above condition is implied by the uniform integrability of $\{X_n\}$ (for the definition of uniform integrability, see Chung, 1974),.

AMS (1985) *subject classification :* Primary 60F05 ; Secondary 60G42.

Key words and phrases : Uniform integrability, uniform integrability in the Cesaro sense, weak law of large numbers, martingale-difference sequence, φ- mixing sequence.

Remark 1. A sequence $\{X_n\}$ of integrable random variables is uniformly integrable in the Cesàro sense iff

$$\lim_{a\to\infty} \limsup_{n\to\infty} \left\{ n^{-1} \sum_{k=1}^{n} \int_{|X_k|>a} |X_k|\, dP \right\} = 0.$$

Theorem 1 : *Let $\{X_n\}$ be a sequence of pairwise independent random variables satisfying the uniform integrability condition in the Cesàro sense. If $E(X_n) = 0$ for each $n \geqslant 1$, then $n^{-1} \sum_{k=1}^{n} X_k$ converges to zero in L^1.*

To prove the above theorem, we shall use the following elementary result.

Lemma 1 : *If $\{X_n\}$ is a sequence of uniformly bounded pairwise independent random variables, then $n^{-1} \sum_{k=1}^{n} (X_k - E(X_k))$ converges to zero in L^1.*

Proof of *Lemma* 1 : Because of the Schwarz inequality it suffices to show that

$$n^{-2} \operatorname{var}\left(\sum_{k=1}^{n} (X_k - E(X_k)) \right) \to 0,$$

which is obvious because of the given assumptions.

Proof of Theorem 1 : Let N be an integer $\geqslant 1$ and put

$$Y_k = X_k \quad \text{if } |X_k| \leqslant N;$$
$$= 0 \quad \text{otherwise.}$$

Let
$$T_n = \sum_{k=1}^{n} Y_k \text{ and } S_n = \sum_{k=1}^{n} X_k.$$

Then
$$S_n = (T_n - E(T_n)) + \sum_{k=1}^{n} (X_k - Y_k) + E(T_n).$$

Hence
$$n^{-1} E(|S_n|) \leqslant n^{-1} E(|T_n - E(T_n)|) + n^{-1} \sum_{k=1}^{n} E(|X_k - Y_k|)$$
$$+ n^{-1} |E(S_n - T_n)| \quad (\text{since } E(S_n) = 0)$$
$$\leqslant n^{-1} E(|T_n - E(T_n)|) + 2n^{-1} \sum_{k=1}^{n} E(|X_k - Y_k|).$$

By Lemma 1, the first term of the right side goes to zero as $n \to \infty$ for each fixed $N \geqslant 1$. We, therefore, get for each $N \geqslant 1$

$$\limsup_{n\to\infty} n^{-1} E(|S_n|) \leqslant 2 \sup_n \left\{ n^{-1} \sum_{k=1}^{n} E(|X_k - Y_k|) \right\}.$$

Now letting $N \to \infty$ and noting that

$$E(|X_k - Y_k|) = \int_{|X_k| > N} |X_k| \, dP$$

we get the desired result.

The above theorem extends the first part of the Theorem in Landers and Rogge (1986) who prove, using a relatively more complicated argument, the WLLN under the assumption of the uniform integrability of $\{X_n\}$. It may also be noted here that the proof of the above extension of Khinchin's WLLN is much simpler than that, due to Markov, of Khinchin's WLLN.

A slight modification of Lemma 1 will yield the following useful generalisation of Theorem 1. See, in this connection, Jamison *et al.* (1965) who discuss the i.i.d. case.

Theorem 2 : *Let $\{X_n\}$ be a sequence of pairwise independent random variables with $E(X_n) = 0 \; \forall \; n \geqslant 1$. Let $\{a_n\}$ be a sequence of non-negative reals such that $\left(\sum_{k=1}^{n} a_k^2 \right) / b_n^2 \to 0$ as $n \to \infty$ where $b_n = \sum_{k=1}^{n} a_k$ which is assumed to be positive for all n. If*

$$\sup_n \left\{ \frac{1}{b_n} \sum_{k=1}^{n} a_k \int_{|X_k| > a} |X_k| \; dP \right\} \to 0 \text{ as } a \to \infty, \qquad \ldots \; (1)$$

then

$$E\left(\left| \sum_{k=1}^{n} a_k X_k \right| \right) / b_n \to 0 \text{ as } n \to \infty.$$

It may be noted that the condition $\left(\sum_{k=1}^{n} a_k^2 \right) / b_n^2 \to 0$ as $n \to \infty$ holds if $a_n = o_e(n^t)$ for some $t > 0$.

We now give an alternative description of the condition of the uniform integrability in the Cesàro sense ; for the corresponding description of the uniform integrability, see Chung (1974).

Theorem 3 : *A sequence $\{X_n\}$ of random variables satisfies the uniform integrability condition in the Cesàro sense if and only if the following two conditions are satisfied :*

(a) $\sup_n \left(n^{-1} \sum_{k=1}^{n} E(|X_k|) \right) < \infty$

(b) *for each $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever $\{A_k\}$ is a sequence of events satisfying the condition that*

$$\sup_n \left(n^{-1} \sum_{k=1}^{n} P(A_k) \right) < \delta, \qquad \ldots \; (2)$$

we have

$$\sup_n \left(n^{-1} \sum_{k=1}^{n} \int_{A_k} |X_k| \, dP \right) < \varepsilon. \quad \dots \quad (3)$$

Proof of Theorem 3 : We shall first prove the 'only if' part. Let $a_0 > 0$ be such that

$$\sup_n \left(n^{-1} \sum_{k=1}^{n} \int_{|X_k| > a_0} |X_k| \, dP \right) \leqslant 1.$$

Then

$$E(|X_k|) \leqslant a_0 + \int_{|X_k| > a_0} |X_k| \, dP$$

which implies that

$$n^{-1} \sum_{k=1}^{n} E(|X_k|) \leqslant a_0 + n^{-1} \sum_{k=1}^{n} \int_{|X_k| > a_0} |X_k| dP \leqslant a_0 + 1 ;$$

thus (a) holds.

Now fix an $\varepsilon > 0$. Let $a_0 > 0$ be such that

$$\sup_n \left(n^{-1} \sum_{k=1}^{n} \int_{|X_k| \geqslant a_0} |X_k| dP \right) < \frac{\varepsilon}{2}.$$

Put $\delta = \varepsilon/(2a_0)$. Then if (2) holds,

$$n^{-1} \sum_{k=1}^{n} \int_{A_k} |X_k| dP \leqslant n^{-1} \sum_{k=1}^{n} \left(a_0 P(A_k) + \int_{|X_k| \geqslant a_0} |X_k| dP \right)$$

$$= a_0 n^{-1} \sum_{k=1}^{n} P(A_k) + n^{-1} \sum_{k=1}^{n} \int_{|X_k| \geqslant a_0} |X_k| dP < a_0 \delta + \varepsilon/2 = \varepsilon.$$

Thus (b) holds.

For the 'if' part, put

$$K = \sup_n \left\{ n^{-1} \sum_{k=1}^{n} E(|X_k|) \right\}.$$

Then for each $a > 0$,

$$P(|X_k| \geqslant a) \leqslant a^{-1} E(|X_k|) \qquad \forall \, k \geqslant 1$$

and so

$$n^{-1} \sum_{k=1}^{n} P(|X_k| \geqslant a) \leqslant K/a \qquad \forall \, n \geqslant 1.$$

Fix an $\varepsilon > 0$. By (b), there exists a $\delta > 0$ such that (2) implies (3). Put $a_0 = K/\delta$. If $a > a_0$, then

$$\int_{|X_k| \geqslant a} |X_k| dP \leqslant \int_{|X_k| \geqslant a_0} |X_k| \, dP$$

which implies that

$$n^{-1}\sum_{k=1}^{n}\int_{|X_k|\geq a_0}|X_k|\,dP \leq n^{-1}\sum_{k=1}^{n}\int_{|X_k|\geq a_0}|X_k|\,dP < \varepsilon.$$

Hence the proof is complete.

We now show that for the L^1 convergence of the sample mean of *independent* random variables, the condition of uniform integrability in the Cesàro sense is not necessary.

Example 1. Let $\{X_n\}$ be a sequence of independent random variables where X_n follows $N(0\,;\ \sigma_n^2)\ \forall\ n \geq 1$. Put $\sigma_n = n^{1/4}$. Then $n^{-1}E(|X_1+\ldots+X_n|)\to 0$ iff $n^{-2}(\sigma_1^2+\ldots+\sigma_n^2)\to 0$ which is true, since

$$\sum_{k=1}^{n} k^{1/2} \leq \sum_{k=1}^{n}\int_{k}^{k+1} x^{1/2}\,dx = \int_{1}^{n+1} x^{1/2}\,dx.$$

Put $\alpha_n = 2E(N(0\,;1)\,I(N(0,1) > a/\sigma_n)),\ n \geq 1$. Then α_n increases with n and

$$\int_{|X_k|>a}|X_k|\,dP = \sigma_k\,\alpha_k.$$

We next show that uniform integrability in the Césàro sense fails for $\{X_n\}$; in fact, for every $a > 0$

$$\sup_n \left\{n^{-1}\sum_{k=1}^{n}\int_{|X_k|>a}|X_k|\,dP\right\} = \infty.$$

To see this, it suffices to note that

$$n^{-1}\sum_{k=1}^{n}\sigma_k\alpha_k \geq \alpha_1\left(n^{-1}\sum_{k=1}^{n}\sigma_k\right)$$
$$\geq \alpha_1\, n^{-1}\int_{0}^{n} x^{1/4}\,dx.$$

We next show that the uniform integrability in the Cesaro sense is strictly weaker than the uniform integrability.

Example 2. (due to B. V. Rao): Let $X_n = +1$ or -1 with probability $\frac{1}{2}$ each if n is not a perfect cube, and $X_n = +n^{1/3}$ or $-n^{1/3}$ with probability $\frac{1}{2}$ each if n is a perfect cube. Then $\sup_n E(|X_n|) = \infty$, so that $\{X_n\}$ is not uniformly integrable. But if $a \geq 1$, then

$$n^{-1}\sum_{k=1}^{n}\int_{|X_k|>a}|X_k|\,dP \leq n^{-1}\sum_{k=1}^{n}\int_{|X_k|>1}|X_k|\,dP \leq n^{-1}\sum_{\substack{k=j^3\\ k<n}} k^{1/3}$$
$$\leq ((n^{-1/3}+1)n^{2/3})/(2n)\to 0.$$

Now use Remark 1 to establish the uniform integrability in the Cesàro sense.

For the lemma below, note that the uniform integrability of $\{X_n\}$ implies Condition (1) of Theorem 2.

Lemma 2 : *Let $\{a_n\}$ and $\{b_n\}$ be as in Theorem 2. If whenever $a_n \geqslant 0$, $b_n \to \infty$ and $\left(\sum_{k=1}^{n} a_k^2\right)/b_n^2 \to 0$ as $n \to \infty$, Condition (1) of Theorem 2 holds, then $\{X_n\}$ is uniformly integrable.*

Proof : Suppose, by way of contradiction, that

$$\sup_{k \geqslant 1} \int_{|X_k| \geqslant a} |X_k| \, dP \not\to 0 \quad \text{as} \quad a \to \infty.$$

Then there exist an $\varepsilon > 0$ and a sequence $\{k_m\}$ of reals such that $1 < k_1 < k_2 < k_3 < \ldots$ and

$$\int_{|X_{k_m}| \geqslant m} |X_{k_m}| \, dP \geqslant \varepsilon \ \forall \ m \geqslant 1.$$

Define $a_1 = 1$, $a_{k_m} = 1 \ \forall \ m \geqslant 1$ and $a_n = 0$ for all other values of n. Then

$$\sum_{k=1}^{n} a_k^2 = 1 + \text{the number of } j \text{ such that } k_j \leqslant n;$$

$$= b_n,$$

so that $b_n \to \infty$ and $\left(\sum_{k=1}^{n} a_k^2\right)/b_n^2 \to 0$ as $n \to \infty$. Clearly we shall get a contradiction (to Condition (1)) if we show that

$$\limsup_{n \to \infty} \frac{1}{b_n} \left\{ \sum_{k=1}^{n} a_k \int_{|X_k| \geqslant m} |X_k| \, dP \right\} \geqslant \varepsilon$$

for each $m \geqslant 1$. Now fix an $m \geqslant 1$ and observe that for $j \geqslant m$,

$$\left\{ \sum_{i=1}^{k_j} a_i \int_{|X_i| \geqslant m} |X_i| \, dP \right\} / b_{k_j}$$

$$= \left\{ \sum_{i=1}^{j} \int_{|X_{k_i}| \geqslant m} |X_{k_i}| \, dP \right\} / j$$

$$\geqslant \frac{j-m+1}{j} \cdot \frac{1}{j-m+1} \left\{ \sum_{i=m}^{j} \int_{|X_{k_i}| \geqslant m} |X_{k_i}| \, dP \right\}$$

$$\geqslant \frac{j-m+1}{j} \varepsilon \quad \text{(by the choice of } \{k_m\}).$$

Thus

$$\limsup_{j\to\infty} \left\{\sum_{i=1}^{k_j} a_i \int_{|X_i|\geq m} |X_i|\,dP\right\}/b_{k_j} \geq \varepsilon.$$

This completes the proof.

We now replace the condition 'pairwise independence' in Theorem 1 by suitable 'other dependence conditions'.

Theorem 4 : *Let* $\{X_n\}$ *be a sequence of a martingale-difference random variables relative to* $\{\mathcal{B}_n\}$, *i.e.* $E(X_n \mid \mathcal{B}_{n-1}) = 0$ *for all* $n \geq 1$. *If* $\{X_n\}$ *is uniformly integrable in the Cesàro sense, then* $E(|n^{-1}S_n|) \to 0$ *as* $n \to \infty$ *where* $S_n = \sum_{i=1}^{n} X_i, n \geq 1$.

Proof : Let $\mathcal{B}_0$ be the trivial sigma-field, N an integer ≥ 1 and Y_k be as in the proof of Theorem 1 ($k \geq 1$) ; put

$$Z_n = \sum_{k=1}^{n} (Y_k - E(Y_k \mid \mathcal{B}_{k-1})),\ n \geq 1.$$

Then

$$S_n = Z_n + \sum_{k=1}^{n} E(Y_k \mid \mathcal{B}_{k-1}) + \sum_{k=1}^{n} (X_k - Y_k).$$

Hence

$$E(|n^{-1} S_n|) \leq n^{-1} E(|Z_n|) + n^{-1} E\left(\sum_{k=1}^{n} |E(X_k - Y_k \mid \mathcal{B}_{k-1})|\right)$$

$$+ n^{-1} \sum_{k=1}^{n} E(|X_k - Y_k|) \qquad \ldots (4)$$

since $E(X_k \mid \mathcal{B}_{k-1}) = 0$ for $k \geq 1$. Now the second term of Inequality (4) is

$$\leq n^{-1} E\left[\sum_{k=1}^{n} E(|X_k - Y_k| \mid \mathcal{B}_{k-1})\right]$$

$$= n^{-1} \sum_{k=1}^{n} E(|X_k - Y_k|).$$

Thus

$$E(|n^{-1} S_n|) \leq n^{-1} E(|Z_n|) + 2 \sup_m m^{-1} \sum_{k=1}^{m} E(|X_k - Y_k|)$$

Below we show that

$$\mathrm{var}(Z_n/n) \to 0 \text{ as } n \to \infty, \qquad \ldots (5)$$

which will imply that $n^{-1} E(|Z_n|) \to 0$ as $n \to \infty$ by the Schwarz inequality since $E(Z_n) = 0$. Letting $N \to \infty$ and using the uniform integrability of

$\{X_n\}$ in the Cesaro sense, the proof will be complete as in the proof of Theorem 1. Now

$$n^{-2}\operatorname{var}(Z_n) = n^{-2}\Big[\sum_{k=1}^{n}\operatorname{var}(Y_k - E(Y_k\,|\,\mathcal{B}_{k-1}))$$

$$+2\sum_{i<j} E(Y_i - E(Y_i\,|\,\mathcal{B}_{i-1}))(Y_j - E(Y_j\,|\,\mathcal{B}_{j-1}))\Big]$$

$$\leqslant n^{-2}(4nN^2+0)\to 0 \text{ as } n\to\infty$$

since for $1 \leqslant i < j \leqslant n$,

$$E(Y_i - E(Y_i\,|\,\mathcal{B}_{i-1}))(Y_j - E(Y_j\,|\,\mathcal{B}_{j-1}))$$

$$= E\{E(Y_i - E(Y_i\,|\,\mathcal{B}_{i-1}))(Y_j - E(Y_j\,|\,\mathcal{B}_{j-1}))\,|\,\mathcal{B}_{j-1}\}$$

$$= E\{(Y_i - E(Y_i\,|\,\mathcal{B}_{i-1}))(E(Y_j\,|\,\mathcal{B}_{j-1}) - E(Y_j\,|\,\mathcal{B}_{j-1}))\}$$

$$= 0.$$

For the next theorem, let $r > 1$ and recall the definition of φ-mixing sequence as given in Billingsley (1968, page 166).

Theorem 5 : *Let $\{X_n\}$ be a sequece of φ-mixing random variables such that*

$$n^{-1}\sum_{i=1}^{n-1}(\varphi(i))^r \to 0 \text{ as } n\to\infty$$

and $E(X_n) = 0$ for $n \geqslant 1$. If $\{X_n\}$ is uniformly integrable in the Césàro sense, then $E(|n^{-1}S_n|)\to 0$ as $n\to\infty$ where $S_n = \sum_{i=1}^{n} X_i$, $n \geqslant 1$.

Proof : Let N be an integer $\geqslant 1$ and define Y_k and T_n as in the proof of Theorem 1. Then, we get as before,

$$E(|n^{-1}S_n|) \leqslant n^{-1}E(|T_n - E(T_n)|) + 2\sup_m m^{-1}\sum_{k=1}^{m} E(|X_k - Y_k|).$$

It remains to show (as in the proof of Theorem 4) that for each $N \geqslant 1$,

$$\operatorname{var}(n^{-1}(T_n - E(T_n)))\to 0 \text{ as } n\to\infty.$$

Now $$n^{-2}\operatorname{var}(T_n) \leqslant n^{-2}(4n\,N^2 + \sum_{i<j}|\operatorname{cov}(Y_i, Y_j)|)$$

$$\leqslant n^{-2}(4n\,N^2 + 2N^2\sum_{i=1}^{n-1}(n-i)\varphi(i))$$

by Lemma 2 of Billingsley (1968, page 187). To complete the proof, we let $r^{-1}+s^{-1} = 1$ and note that

$$n^{-2} \sum_{i=1}^{n-1} (n-i)\varphi(i) \leqslant n^{-2} \left\{ \sum_{j=1}^{n-1} j^s \right\}^{1/s} \left\{ \sum_{i=1}^{n-1} (\varphi(i))^r \right\}^{1/r}$$

$$\leqslant n^{-2} \left\{ \sum_{j=1}^{n-1} \int_j^{j+1} x^s \, dx \right\}^{1/s} \left\{ \sum_{i=1}^{n-1} (\varphi(i))^r \right\}^{1/r}$$

$$= n^{-2} \left\{ \int_1^n x^s \, dx \right\}^{1/s} \left\{ \sum_{i=1}^{n-1} (\varphi(i))^r \right\}^{1/r}$$

$$\leqslant n^{-2} \left(\frac{n^{s+1}}{s+1} \right)^{1/s} \left\{ \sum_{i=1}^{n-1} (\varphi(i))^r \right\}^{1/r}$$

$$= (s+1)^{-1/s} \left\{ n^{-1} \sum_{i=1}^{n-1} (\varphi(i))^r \right\}^{1/r} \to 0 \text{ as } n \to \infty.$$

Remark 2. The assumption of 'pairwise independence' in Theorem 2 can similarly be relaxed to cover the above two notions of dependence. We omit the details.

Acknowledgement. The author is grateful to the referee for many helpful suggestions which improved the presentation of the paper.

REFERENCES

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*, John Wiley, New York.

CHUNG, K. L. (1974): *A Course in Probability Theory*, second ed., Academic Press, New York.

DHARMADHIKARI, S. W. (1976): A simple proof of mean convergence in the law of large numbers. *Amer. Math. Monthly*, **83**, 474-475.

ETEMADI, N. (1981): An elementary proof of the strong law of large numbers. *Z. Wahrs. Geb.*, **55**, 119-122.

JAMISON, B., OREY, S. and PRUITT, W. (1965): Convergence of weighted averages of independent random variables. *Z. Wahrs. Geb.*, **4**, 40-44.

LANDERS, D. and ROGGE, L. (1986): Laws of large numbers for piarwise independent uniformly integrable random variables. *Preprints in statistics* No. 98, University of Cologne.

RAO, C. R. (1973): *Linear Statistical Inference and Its Applications*, second ed., John Wiley and Sons.