Sankhyā : The Indian Journal of Statistics
1993, Volume 55, Series A, Pt. 1, pp. 12-28.

CESÀRO UNIFORM INTEGRABILITY AND L_p-CONVERGENCE

By ARUP BOSE
Purdue University
and
TAPAS K. CHANDRA
Indian Statistical Institute

SUMMARY. We show how a new condition, called Cesàro uniform integrability, introduced by Chandra (1989) can be used in many cases to prove L_p-convergence of $n^{-1/p} S_n$ where $S_n = \sum_{i=1}^{n} X_i$.

1. Introduction

Let $(X_n : n \geqslant 1)$ be a sequence of random variables and let $S_n = X_1 + \ldots + X_n$. Pyke and Root (1968) proved that if $(X_n : n \geqslant 1)$ is an independent and identically distributed (i.i.d.) sequence and $E(|X_1|^p) < \infty$ for some $0 < p < 2$, then $n^{-1}E(|S_n - a_n|^p) \to 0$ as $n \to \infty$ where $a_n = 0$ if $0 < p < 1$ and $a_n = nE(X_1)$ if $1 \leqslant p < 2$. Chatterjee (1969) extended this result by assuming only that $(X_n : n \geqslant 1)$ is dominated in distribution by a random variable X such that $E(|X|^p) < \infty$ and taking $a_n = \sum_{k=1}^{n} E(X_k | X_1, \ldots, X_{k-1})$ if $1 \leqslant p < 2$. Chow (1971) strengthened this result by replacing the domination condition by the condition of uniform integrability (UI) of $(|X_n|^p : n \geqslant 1)$.

In a recent paper Chandra (1989), a new condition called "Cesàro uniform integrability" (CUI) was introduced. This condition is weaker than the usual UI condition and yet was shown to be strong enough to derive L_1-convergence in the weak law of large numbers (WLLN). In this paper we establish L_p-convergence, $0 < p < 2$, for several types of independent and dependent sequences under CUI. The dependent sequences include pairwise independent sequences, martingale differences and L_p-mixingale differences.

Paper received. August 1990 ; revised March 1991.

AMS (1985) *subject classifications.* Primary 60F25 ; secondary 60G48, 60G50.

Key words and phrases. Cesàro uniform integrability, L_p-convergence, martingale differences, L_p-mixingale differences, autoregressive process, moving average process.

It appears that this CUI condition will be useful in deriving strong law of large numbers (SLLN), more general than those known in the literature. See Chandra and Goswami (1992) for an account of the progress made in this direction.

2. Preliminaries

In this section we give the definition and basic properties of CUI sequences and introduce the concept of L_p-mixingales. The latter generalizes the concepts of mixingales introduced by McLeish (1975) and its extension given by Andrews (1989).

Definition 2.1. A sequence of real valued random variables $(X_n : n \geqslant 1)$ on $(\Omega, \mathscr{A}, P)$ is said to be Cesàro uniformly integrable (CUI) if

$$\lim_{a \to \infty} \limsup_{n \to \infty} \left(n^{-1} \sum_{k=1}^{n} E[|X_k| I(|X_k| > a)]\right) = 0.$$

Remark 2.2. In order that WLLN (or SLLN) holds for $(X_n : n \geqslant 1)$, it should be possible to allow a few of the X_n's to take large values. The CUI condition is capable (at least to a certain extent) of allowing such sequences. In this connection see Chandra (1989).

In the following lemma, we collect the basic facts we will require about the above condition.

Lemma 2.3. *Let $(X_n : n \geqslant 1)$ and $(Y_n : n \geqslant 1)$ be two sequences of random variables on $(\Omega, \mathscr{A}, P)$.*

(i) *$(X_n : n \geqslant 1)$ is CUI if and only if*

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

and

(b) *given $\epsilon > 0$, there esists a $\delta > 0$ such that for any sequence of measurable sets*

$$(A_n : n \geqslant 1) \text{ with } \limsup_{n} n^{-1} \sum_{k=1}^{n} P(A_k) < \delta, \ \limsup_{n} n^{-1} \sum_{k=1}^{n} E[|X_k| I(A_k)] < \epsilon$$

(ii) *If $(|X_n| : n \geqslant 1)$ is CUI and $|Y_n| \leqslant |X_n|$ a.s., then $(|Y_n| : n \geqslant 1)$ is CUI.*

(iii) *If for some $p > 0$, $(|X_n|^p : n \geqslant 1)$ and $(|Y_n|^p : n \geqslant 1)$ are CUI, then so is $(|X_n + Y_n|^p : n \geqslant 1)$.*

(iv) *Let $(\mathscr{F}_n : n > 1)$ be a sequence of sub-sigma fields of $\mathscr{A}$ and $p > 0$. If $(|X_n|^p : n \geqslant 1)$ is CUI, then so is $(Y_n = E(|X_n|^p \mid \mathscr{F}_n), n > 1)$.*

Proof. (i) is proved in Chandra (1989). (ii) is trivial. (iii) follows from the observation,

$$E(|X_k+Y_k|^p I(|X_k+Y_k| > a))$$

$$\leqslant 2^p E\left[[|X_k|^p I(|X_k| > \frac{a}{2}\right]+2^p E\left[|Y_k|^p I\left(|Y_k| > \frac{a}{2}\right)\right].$$

To prove (iv), note that since $I(|Y_k| > a)$ is $\mathscr{F}_k$-measurable,

$$\limsup_n n^{-1} \sum_{k=1}^n E[|Y_k| I(|Y_k|> a)] = \limsup_n n^{-1} \sum_{k=1}^n E[|X_k|^p I(|Y_k|> a)]. \quad \dots \quad (2.1)$$

Note that as $a \to \infty$,

$$\limsup_n n^{-1} \sum_{k=1}^n P(|Y_k| > a) \leqslant \left[\limsup_n n^{-1} \sum_{k=1}^n E(|X_k|^p)\right] a^{-p} \to 0.$$

Thus using the alternative criteria of CUI established in (i), the term in (2.1) $\to 0$ as $a \to \infty$ since $(|X_k|^p : k \geqslant 1)$ is CUI □

Remark 2.4. The following implications relate the concepts of UI and CUI.

$$(X_k)\ UI \Rightarrow (X_k)\ CUI \Rightarrow \left(Y_k = k^{-1} \sum_{i=1}^k X_i\right) UI.$$

The proof of this is easy using the criterion of Lemma 2.1 for for CUI and the similar criterion for UI. However, none of the reverse implications are true in general. To see this, let

$$X_{2k} = -X_{2k-1} \sim N(0,(2k-1)^{3/2}),\ k = 1, 2, \dots$$

Then it is easy to see that $(Y_k : k \geqslant 1)$ is UI. However $(2n)^{-1} \sum_{k=1}^{2n} E|X_k| \approx n^{3/4}$ as $n \to \infty$. So $(X_k : k \geqslant 1)$ is not CUI. For an example where CUI $\not\Rightarrow$ UI see Chandra (1989).

The concept of asymptotic martingales was introduced by McLeish (1975), who called them mixingales. Andrews (1989) extended this concept to what he called L_1-mixingales. We extend these concepts below through the following definitions.

Let $(X_n : n \geqslant 1)$ be a sequence of random variables on $(\Omega, \mathscr{A}, P)$ such that $E(|X_n|^p) < \infty$ for some $p \geqslant 1$ and for each $n \geqslant 1$. Let $(\mathscr{F}_n : n = 0, \pm 1, \pm 2, 2, \dots)$ be an increasing sequence of sub-sigma fields of $\mathscr{A}$. Let $\|\cdot\|_p$ denote the L_p-norm.

Definition 2.5. The pair $\{(X_n : n \geqslant 1), (\mathscr{F}_n : n = 0, \pm 1, \ldots)\}$ is called an L_p-mixingale difference sequence if there exist sequences of constants $(c_n : n \geqslant 1)$ and $(\psi_m : m \geqslant 0)$ such that $\psi_m \to 0$ as $m \to \infty$ and

(a) $\|E(X_n \mid \mathscr{F}_{n-m})\|_p \leqslant c_n \psi_m$

and

(b) $\|X_n - E(X_n \mid \mathscr{F}_{n+m})\|_p \leqslant c_n \psi_{m+1}$.

For some illuminating examples of L_1 and L_2 mixingale difference sequences in the above sense, see Hall and Heyde (1980) and Andrews (1989).

In the next sections C stands for a generic constant and S_n will denote $\sum_{i=1}^{n} X_i$, $\sum_{k=1}^{kn}$ or $\sum_{k=1}^{kn} a_{nk} X_{nk}$ as the case may be.

3. The main results

In this section we prove various L_p-convergence results. Our first result is an extension of a Theorem of Chow (1971) who proves the following result with the assumption of UI of the sequence $(X_n : n \geqslant 1)$ and deals with the case $0 < p < 1$.

Theorem 3.1. *Let* $0 < p < 1$ *and* $(|X_n|^p : n \geqslant 1)$ *be CUI. Then* $n^{-1}E(|S_n|^p) \to 0$.

Proof. For $a > 0$, defire

$$Y_n = X_n I(|X_n| \leqslant a), n \geqslant 1$$
$$Z_n = X_n - Y_n, \qquad n \geqslant 1.$$

Then

$$n^{-1}E(|S_n|^p) = n^{-1} E\left(\left|\sum_{k=1}^{n} Z_k + \sum_{k=1}^{n} Y_k\right|^p\right)$$
$$\leqslant n^{-1} E\left(\left|\sum_{k=1}^{n} Z_k\right|^p\right) + n^{-1} E\left(\left|\sum_{k=1}^{n} Y_k\right|^p\right)$$
$$\leqslant n^{-1} \sum_{k=1}^{n} E|Z_k|^p + n^{-1+p} a^p.$$

So

$$\limsup_{n\to\infty} n^{-1}E(|S_n|^p \leqslant \limsup_{n\to\infty} n^{-1} \sum_{k=1}^{n} E|Z_k|^p$$
$$= \limsup_{n\to\infty} n^{-1} \sum_{k=1}^{n} E|(X_k|^p I(|X_k| > a)).$$

Now letting $a \to \infty$ and using the fact that $(|X_n|^p : n \geqslant 1)$ is CUI, the result follows. □

The following theorem deals with the case $1 \leqslant p < 2$ and extends Theorem 2.22 of Hall and Heyde (1980) and Theorem 4 of Chandra (1989).

Theorem 3.2. *Let $(X_n : n \geqslant 1)$ be a martingale difference sequence such that $(|X_n|^p : n \geqslant 1)$ is CUI for some $1 \leqslant p < 2$. Then* $n^{-1}E(|S_n|^p) \to 0$.

Proof. Let Y_n, Z_n be defined as in Theorem 3.1. The case $p = 1$ is proved in Chandra (1989). We give below a simpler proof for this case.

$$n^{-1}E\left(\left|\sum_{k=2}^{n} X_k\right|\right) \leqslant n^{-1} E\left|\sum_{k=2}^{n} (Y_k - E(Y_k|X_1, \ldots, X_{k-1}))\right|$$

$$+n^{-1}E\left|\sum_{k=2}^{n} Z_k\right| + n^{-1} E\left|\sum_{k=2}^{n} E(Z_k|X_1, \ldots, X_{k-1})\right|$$

Since $(Y_k - E(Y_k|X_1, \ldots, X_{k-1}), k \geqslant 2)$ is a bounded martingale difference sequence (with respect $\mathscr{F}_k = \sigma(X_1, \ldots, X_{k-1})$) the first term above $\to 0$ as $n \to \infty$ (see for example Theorem 2.22 of Hall and Heyde (1980)). The last two terms are dominated by $2n^{-1}E(\sum_{k=1}^{n} |Z_k|)$. First letting $n \to \infty$ and then $a \to \infty$, this converges to 0 by CUI.

We now look at the case $1 < p < 2$. Let C denote a generic constant. By Burkholder's inequality (1966) (see Theorem 2.10 of Hall and Heyde, 1980),

$$E(|S_n|^p) \leqslant CE(|\sum_{k=1}^{n} X_k^2|^{p/2}) = CE(|\sum_{k=1}^{n} (Z_k^2 + Y_k^2)|^{p/2})$$

$$\leqslant CE(|\sum_{k=1}^{n} Z_k^2|^{p/2}) + CE(|\sum_{k=1}^{n} Y_k^2|^{p/2}) \leqslant C\sum_{k=1}^{n} E(|Z_k|^p) + C(na^2)^{p/2}.$$

Thus $\limsup_{n\to\infty} n^{-1}E(|S_n|^p) \leqslant C \limsup_{n\to\infty} n^{-1}\sum_{k=1}^{n} E(|Z_k|^p)$. Now the result follows as in Theoreem 3.1.□

Corollary 3.3. *If $(|X_n|^p : n \geqslant 1)$ is CUI for some $1 \leqslant p < 2$ then* $n^{-1} E(|\sum_{k=2}^{n} (X_k - E(X_k|X_1, \ldots, X_{k-1}))|^p) \to 0$.

Proof. Note that $(Y_k = X_k - E(X_k|X_1, \ldots, X_{k-1}), k \geqslant 2)$ is a martingale difference sequence. Further $|E(X_k|X_1, \ldots, X_{k-1})|^p) \leqslant E(|X_k|^p| X_1, \ldots, X_{k-1})$. So by applying Lemma 1 (ii), (iii) and (iv), $(Y_k : k \geqslant 2)$ is CUI and the corollary follows from Theorem 3.2. □

We now turn to L_p-mixingales and generalize Theorem 1 of Andrews (1989) in two directions. First, we prove L_p convergence for $p \geqslant 1$, whereas Andrews works with $p = 1$. Second, we reduce the assumption of UI to CUI.

Theorem 3.4. *Let* $\{(X_n : n \geqslant 1), (\mathcal{F}_i, i = 0, \pm 1, \pm 2, \dots)\}$ *be an* L_p*-mixingale difference sequence and* $(|X_n|^p : n \geqslant 1)$ *be CUI for some* $1 \leqslant p < 2$. *Further assume that* $\limsup_n n^{-1}\left(\sum_{i=1}^{n} c_i\right)^p < \infty$. Then $E(n^{-1}|S_n|^p) \to 0$.

Proof. For $n \geqslant 1$ and $i = 0, \pm 1, \pm 2, \dots$ define

$$Y_{ni} = E(X_i \mid \mathcal{F}_{n+i}) - E(X_i \mid \mathcal{F}_{n+i-1}).$$

For each i, $(Y_{ni}, \mathcal{F}_{n+i}, n \geqslant 1)$ is a martingale difference sequence. Further $(|Y_{ni}|^p : n \geqslant 1)$ is CUI by Lemma 1.

Define $S_{ni} = \sum_{k=1}^{n} Y_{ki}$. By Theorem 3.2, $n^{-1/p}\|S_{ni}\|_p \to 0$ as $n \to \infty$ for each i. Further

$$S_n = \sum_{k=1}^{n} (X_k - E(X_k \mid \mathcal{F}_{k+m})) + \sum_{k=1}^{n} E(X_k \mid \mathcal{F}_{k-m}) + \sum_{i=-m+1}^{m} S_{ni}.$$

Thus

$$\|S_n\|_p \leqslant \sum_{k=1}^{n} \|X_k - E(X_k \mid \mathcal{F}_{k+m})\|_p + \sum_{k=1}^{n} \|E(X_k \mid \mathcal{F}_{k-m})\|_p + \sum_{i=-m+1}^{m} \|S_{ni}\|_p$$

Thus

$$\limsup_n n^{-1/p}\|S_n\|_p \leqslant \limsup_{n \to \infty}\left(n^{-1/p} \sum_{k=1}^{n} c_k\right) \psi_{m+1} + \limsup_{n \to \infty}\left(n^{-1/p} \sum_{k=1}^{n} c_k\right) \psi_m$$

Now using the condition on c_i and the fact that $\psi_m \to 0$ as $m \to \infty$, the result follows. □

Remark. It is clear from the above proof that Theorem 3.4 continues to hold if the conditions (a), (b) in Definition 2.5 are replaced by

$$\text{(a)}' \quad \limsup_{m \to \infty} \; \limsup_{n \to \infty} \left(n^{-1/p} \sum_{k=1}^{n} \|E(X_k \mid \mathcal{F}_{k-m})\|_p\right) = 0$$

$$\text{(b)}' \quad \limsup_{m \to \infty} \; \limsup_{n \to \infty} \left(n^{-1/p} \sum_{k=1}^{n} \|X_k - E(X_k \mid \mathcal{F}_{k+m})\|_p\right) = 0.$$

In the following theorems, we show how normalization other than n may be used for S_n.

Theorem 3.5. *Let* $(X_n : n \geqslant 1)$ *be a sequence of identically distributed random variables with* $E(X_n) = 0$ *which is either pairwise independent or is a martingale difference. Suppose that* f *is a function such that* $f(x) > 0$ *for* $x > 0$, $x^{-1}f(x)$ *is nonincreasing as* $x \to \infty$ *and* $x^{-1}f^2(x) \uparrow \infty$ *and* $E[f^{-1}(|X_1|)] < \infty$. *Then* $(f(n))^{-1}E(|S_n|) \to 0$.

Proof. First assume that $(X_n : n \geqslant 1)$ is pairwise independent. Let

$$Y_{nj} = X_j I(|X_j| \leqslant f(n)), \quad j = 1, \ldots, n$$
$$Z_{nj} = X_j - Y_{nj}, \quad j = 1, \ldots, n$$
$$T_n = \sum_{j=1}^{n} Y_{nj}$$

$$\begin{aligned} E(|S_n|) &\leqslant E(|T_n|) + E(|S_n - T_n|) \leqslant [E(T_n^2)]^{1/2} + nE(|Z_{n1}|) \\ &= [V(T_n) + (E(T_n))^2]^{1/2} + nE(|Z_{n1}|) \\ &\leqslant \left[\sum_{j=1}^{n} V(Y_{nj}) + n^2(E(|Z_{n1}|))^2 \right]^{1/2} + nE(|Z_{n1}|) \\ &\leqslant [nE(Y_{n1}^2) + n^2(E(|Z_{n1}|))^2]^{1/2} + nE(|Z_{n1}|). \end{aligned}$$

Let $V = f^{-1}(|X_1|)$. For large n,

$$\begin{aligned} E(|Z_{n1}|) &= E[|X_1| I(|X_1| > f(n))] \\ &= E[f(V)V^{-1}\, VI(f(V) > f(n))] \leqslant n^{-1}f(n)E[VI(V > n)] \\ &= o(n^{-1}f(n)). \qquad \ldots \quad (3.1) \end{aligned}$$

As $x^{-1}f^2(x) \to \infty$, there exists integers $N_n \uparrow \infty$ such that $N_n^{-1}f^2(N_n) = o(n^{-1}f^2(n))$. To see this, define $g(x) = x^{-1}f^2(x)$, $n_0 = 1$, and given $n_0 < n_1 < \ldots < n_{k-1}$ define n_k to be such that $\dfrac{g(k)}{g(n)} \leqslant k^{-1}\ \forall n \geqslant n_k > n_{k-1}$. Now define

$$\begin{aligned} N_n &= 1 \text{ if } n_0 \leqslant n \leqslant n_1 - 1 \\ &= k \text{ if } n_k \leqslant n < n_{k+1} - 1. \end{aligned}$$

Thus

$$\begin{aligned} E(Y_{n1}^2) &= E[X_1^2 I(|X_1| \leqslant f(n))] \\ &= E[V^{-1}f^2(V)VI(V \leqslant n)] \\ &= E[V^{-1}f^2(V)VI(V \leqslant N_n)] \\ &\quad + E[V^{-1}f^2(V)VI(N_n < V \leqslant n)] \\ &\leqslant N^{-1}f^2(N_n)E(V) + n^{-1}f^2(n)E[VI(V > N_n)] \\ &= o(n^{-1}f^2(n)). \qquad \ldots \quad (3.2) \end{aligned}$$

Combining (3.1) and (3.2) we have the result when $(X_n : n \geqslant 1)$ is pairwise independent. When $(X_n : n \geqslant 1)$ is a martingale difference, the same proof works by replacing Z_{nj} and T_{nj} by $Z'_{nj} = Z_{nj} - E(Z_{nj} \mid X_1, \ldots, X_{j-1})$ and $T'_{nj} = T_n - E(T_{nj} \mid X_1, \ldots, X_{j-1})$. □

Choosing $f(x) = x^{1/p}$, we have the following Corollary.

Corollary 3.6. *If $(X_n : n \geqslant 1)$ is a sequence of pairwise independent identically distributed r.v.'s such that $E(|X_1|^p) < \infty$ for some $1 \leqslant p < 2$, and $E(X_1) = 0$, then $E(|S_n|) = o(n^{1/p})$.*

Theorem 3.7. *Let $(X_n : n \geqslant 1)$ be a martingale difference sequence. Let f be a function and let $1 \leqslant p < 2$ be such that $f(x)$ is nondecreasing, $x^{-p} f^2(x) \to \infty$ and*

$$\lim_{a \to \infty} \limsup_{n \to \infty} (f(n))^{-1} \sum_{j=1}^{n} E[|X_j|^p I(|X_j| > a)] = 0.$$

Then $(f(n))^{-1} E(|S_n|^p) \to 0$.

This is a more general version of Theorem 3.2 and the proof is omitted. □

The CUI condition can be adapted to prove L_p-convergence of weighted sums. Below, we give two such theorems for the case $1 \leqslant p < 2$. As was seen in Theorem 3.1, the case $0 < p < 1$ is much easier to deal with.

Theorem 3.8. *Let $(X_{nk} : 1 \leqslant k \leqslant k_n,\ n \geqslant 1)$ be a triangular array of random variables such that $(X_{nk} : 1 \leqslant k \leqslant k_n)$ is pairwise independent for each $n \geqslant 1$ and $EX_{nk} = 0$ and let $(a_{nk} : 1 \leqslant k \leqslant k_n, n \geqslant 1)$ be an array of real numbers such that*

$$\lim_{n \to \infty} \sum_{k=1}^{k_n} a_{nk}^2 = 0$$

and

$$\lim_{a \to \infty} \limsup_{n \to \infty} \sum_{k=1}^{k_n} |a_{nk}| E(|X_{nk}| I(|X_{nk}| > a)) = 0.$$

Then $E(|\sum_{k=1}^{k_n} a_{nk} X_{nk}|) \to 0$ as $n \to \infty$.

To prove the theorem, we need the following lemma.

Lemma 3.9. *Let (X_{nk}) (a_{nk}) be as in Theorem 3.8. Assume in addition that $\sup_{n,k} |X_{nk}| \leqslant A < \infty$ for some constant A. Then the conclusions of Theorem 3.8 hold.*

Proof. $E(|\Sigma a_{ni} X_{ni}|) \leqslant [E(\Sigma a_{ni} X_{ni})^2]^{1/2} \leqslant [\Sigma a_{ni}^2 A^2]^{1/2} \to 0$. □

Proof of Theorem 3.8. Fix $a > 0$ and define

$$S_n = \sum_{k=1}^{k_n} a_{nk} X_{nk},\ n \geqslant 1$$

$$T_n = \sum_{k=1}^{k_n} a_{nk} X_{nk} I(|X_{nh}| \leqslant a)$$

$$Y_n = S_n - T_n.$$

Then we have

$$S_n = T_n - ET_n + Y_n + E(T_n - S_n).$$

Hence

$$\|S_n\|_1 \leqslant \|T_n - ET_n\|_1 + \|Y_n\|_1 + \|E(T_n - S_n)\|_1.$$

By Lemma 3.9

$$\|T_n - ET_n\|_1 \to 0. \qquad \text{... (3.3)}$$

Further,

$$\|Y_n\|_1 \leqslant \sum_{i=1}^{k_n} \|a_{nk} X_{nk} I(|X_{nk}| > a\|_1$$

$$= \sum_{i=1}^{k_n} |a_{nk}|\{E[|X_{nk}| I(|X_{nk}| < a)]\} \qquad \text{... (3.4)}$$

$$|E(T_n - S_n)| \leqslant \sum_{k=1}^{k_n} |a_{nk}| E(|X_{nk}| I(|X_{nk}| < a) \qquad \text{... (3.5)}$$

Now the result follows by first letting $n \to \infty$ and then $a \to \infty$ and using relations (3.3), (3.4) and (3.5). □

Theorem 3.10. *Let $(X_{nk} : 1 \leqslant k \leqslant k_n,\ n \geqslant 1)$ be a triangular array of random variables which is a martingale difference sequence for each n on the probability space $(\Omega_n, \mathcal{A}_n, P_{n,\theta})$ for each $\theta \in K$ and let $(a_{nk} : 1 \leqslant k \leqslant k_n, n \geqslant 1)$ be an array of real numbers such that*

$$\sum_{k=1}^{k_n} a_{nk}^2 \to 0 \text{ and for some } 1 \leqslant p < 2,$$

$$\lim_{a \to \infty} \limsup_{n \to \infty} \sup_{\theta \in K} \sum_{k=1}^{k_n} |a_{nk}|^p E_{n,\theta}(|X_{nk}|^p I(|X_{nk}| > a)) = 0.$$

Then $\sup_{\theta \in K} E_{n,\theta}(|S_n|^p) \to 0.$

Proof. Note that $(a_{nk}X_{nk} : 1 \leqslant k \leqslant k_n,\ n \geqslant 1)$ is also a martingale difference sequence. By Burkholder's inequality, as in the proof of Theorem 3.2, defining $Y_{nk} = X_{nk}I(|X_{nk}| \leqslant a)$, $Z_{nk} = X_{nk}-Y_{nk}$,

$$
\begin{aligned}
E_{n,\theta}|S_n|^p &\leqslant CE_{n,\theta}\left(\sum_{k=1}^{k_n} a_{nk}^2X_{nk}^2\right)^{p/2} \\
&= CE_{n,\theta}\left[\sum_{k=1}^{k_n} a_{nk}^2(Y_{nk}^2+Z_{nk}^2)\right]^{p/2} \\
&\leqslant CE_{n,\theta}\left(\sum_{k=1}^{k_n} a_{nk}^2Y_{nk}^2\right)^{p/2} + C\sum_{k=1}^{k_n}|a_{nk}|^pE_{n,\theta}(|Z_{nk}|^p) \\
&\leqslant Ca^p\sum_{k=1}^{k_n} a_{nk}^2 + C\sum_{k=1}^{k_n}|a_{nk}|^p\,E_{n,\theta}[|X_{nk}|^pI(|X_{nk}| > a)].
\end{aligned}
$$

First letting $n \to \infty$ and then letting $a \to \infty$, the result follows. □

Remark 3.11. As is evident the above theorems are more general than Theorem 3.2 and 3.5. It is also clear that a version of Theorem 3.10 for L_p-mixingales can be proved along the same lines as the proof of Theorem 3.4.

In the following theorem, we generalize the classical WLLN of Markov (see e.g. Loève, 1977, p. 287 and Parzen, 1960, p. 418) to martingale differences and pairwise independent random variables.

Theorem 3.12. *Let $(X_k : k \geqslant 1)$ be a martingale difference sequence or a sequence of pairwise independent random variables satisfying Markov's δ-condition, $n^{-(1+\delta)}\sum_{k=1}^{n} E(|X_k|^{1+\delta}) \to 0$ for some $0 < \delta \leqslant 1$. Then $n^{-1}E(|S_n|) \to 0$.*

Proof. First assume that $(X_k : k \geqslant 1)$ is a martingale difference sequence. Define

$$
\begin{aligned}
X_{nk} &= X_kI(|X_k| \leqslant n), && k = 1, \ldots, n \\
X'_{nk} &= X_k - X_{nk}, && k = 1, \ldots, n \\
Y_{nk} &= E(X_{nk}|X_1, \ldots, X_{k-1}), && k = 1, \ldots, n \\
Z_{nk} &= E(X'_{nk}|X_1 \ldots, X_{k-1}) && k = 1, \ldots, n
\end{aligned}
$$

$$
n^{-1}\sum_{k=1}^{n} X_k = n^{-1}\sum_{k=1}^{n}(X_{nk}-Y_{nk}) + n^{-1}\sum_{k=1}^{n}(X'_{nk}-Z_{nk}) \qquad \ldots (3.6)
$$

$$
E\left|n^{-1}\sum_{k=1}^{n}(X'_{nk}-Z_{nk})\right| \leqslant 2n^{-(1+\delta)}\sum_{k=1}^{n} E|X_k|^{1+\delta} \to 0. \qquad \ldots (3.7)
$$

Note that $(X_{nk}-Y_{nk} : k = 1, \ldots, n)$ is a martingale difference sequence.

Further

$$E(X_{nk}Y_{nk}) = EY_{nk}^2.$$

Thus

$$V\left(n^{-1}\sum_{k=1}^{n}(X_{nk}-Y_{nk})\right) = n^{-2}\sum_{k=1}^{n}V(X_{nk}-Y_{nk})$$

$$= n^{-2}\sum_{k=1}^{n}E(X_{nk}-Y_{nk})^2$$

$$= n^{-2}\sum_{k=1}^{n}E(X_{nk}^2+Y_{nk}^2-2X_{nk}Y_{nk})$$

$$= n^{-2}\sum_{k=1}^{n}E(X_{nk}^2-Y_{nk}^2) \leqslant n^{-2}\sum_{k=1}^{n}E(X_{nk}^2)$$

$$\leqslant n^{-(1+\delta)}\sum_{k=1}^{n}E|X_k|^{1+\delta} \to 0 \text{ as } n\to\infty.$$

Further,

$$E\left[\left|n^{-1}\sum_{k=1}^{n}(X_{nk}-Y_{nk})\right|\right] \leqslant \left[E\left(n^{-1}\sum_{k=1}^{n}(X_{nk}-Y_{nk})^2\right)\right]^{1/2}. \quad \ldots \quad (3.8)$$

Combining (3.6), (3.7), and (3.8), the result follows for martingale differences.

The same proof works when $(X_k : k \geqslant 1)$ is pairwise independent by replacing Y_{nk} and Z_{nk} by the unconditional expectations of X_{nk} and X'_{nk} respectively. □

Remark 3.13. A version of Theorem 3.12 is true for mixingales and can be proved by using arguments given in the proof of Theorem 3.4.

Remark 3.14. Even though we have stated most of our definitions and results for sequences, it is easy to see that with appropriate changes everything extends to triangular array of variables.

4. Examples and counter-examples

In this section we give examples to show that L_p-convergence need not hold under weaker conditions. We also give some examples where our results can be applied.

Example 4.1. Theorem 3.2 need not hold if CUI of $(|X_k|^p : k \geqslant 1)$ is replaced by UI of $(n^{-1}|S_n|^p : n \geqslant 1)$. To see this let $(X_n : n \geqslant 1)$ be independent $N(0, n^{2/p-1})$ variables. Then $\sup_{n \geqslant 1} E(n^{-1}|S_n|^p)^2 < \infty$ and thus $(n^{-1}|S_n|^p : n \geqslant 1)$ is uniformly integrable. However

$$E(n^{-1}|S_1|^p) \rightarrow c > 0.$$

In fact, Theorem 3.2 is not even true under UI of $\left(n^{-1} \sum_{i=1}^{n} |X_i|^p : n \geqslant 1\right)$ as the following example shows.

Example 4.2. Let $(X_n : n \geqslant 1)$ be independent such that $X_n \sim N(0, \sigma_n^2)$ where $\sigma_n = (1+n)^{1/p}$ if $n = 2^m$ for some $m = 0, 1, \ldots$and $\sigma_n = 1$ otherwise. Note that if $X \sim N(0, \sigma^2)$ then $E|X|^p = c_p\sigma^p$. Now

$$E\left(n^{-1} \sum_{i=1}^{n} |X_i|^p\right) = n^{-2}\left[c_{2p} \sum_{i=1}^{n} \sigma_i^{2p} + 2 \sum_{i<j} c_p \sigma_i^p \sigma_j^p\right]$$

$$\leqslant \max(c_p^2, c_{2p}) \left(n^{-1} \sum_{i=1}^{n} \sigma_i^p\right)^2 < \infty.$$

Thus $\left(n^{-1} \sum_{i=1}^{n} |X_i|^p : n \geqslant 1\right)$ is L_2 bounded and hence UI. However, note that if $2^m \leqslant n < 2^{m+1}$, we have

$$\frac{\sigma_1^2+\ldots+\sigma_n^2}{n^{2/p}} \geqslant n^{-2/p} \sum_{j=1}^{m} (4^{1/p})^j \geqslant \left(\frac{(4^{1/p})^{m+1}-1}{4^{1/p}-1}\right)\frac{1}{(4^{1/p})^m}$$

Thus $\liminf_{n \to \infty} n^{-1} E(|S_n|^p) = \liminf_{n \to \infty} n^{-1} c_p(\sigma_1^2+\ldots+\sigma_n^2) > 0.$

The following example shows that Markov's weak law is false if Markov's condition is assumed to hold with $\delta > 1$.

Example 4.3. Let $(X_n : n \geqslant 1)$ be independent $N(0, \sigma_n^2)$ where $\sigma_n = n^\alpha$ and $\frac{1}{2} \leqslant \alpha < \frac{\delta}{(1+\delta)}$, $\delta > 1$. Then

$$n^{-(1+\delta)} \sum_{k=1}^{n} E(|X_k|^{1+\delta}) = c_\delta n^{-(1+\delta)} \sum_{k=1}^{n} n^{\alpha(1+\delta)}$$

$$\leqslant cn^{-(1+\delta)+\alpha(1+\delta)-1} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Thus $(X_n : n \geqslant 1)$ satisfies Markov's δ-condition.

Note that

$$V(n^{-1} S_n) = (\sigma_1^2+\ldots+\sigma_n^2)/n^2$$

$$= n^{-2} \sum_{k=1}^{n} k^{2\alpha} \approx cn^{2\alpha-1} \not\rightarrow 0.$$

Since $n^{-1} S_n$ is a mean zero normal variable this implies that $n^{-1} S_n \overset{P}{\not\to} 0$.

The next example shows that Markov's δ condition does not imply the CUI condition.

Example 4.4. Let $X_k \sim N(0, \sigma_k^2)$ where $\sigma_n = n^\alpha$ and $0 < \alpha < \delta/(1+\delta)$. Then it is easily seen that $(X_k : k \geqslant 1)$ satisfies Markov's δ-condition. Let X be a $N(0, 1)$ variable. Then for any $a > 0$,

$$n^{-1} \sum_{k=1}^{n} E(|X_k| I(|X_k| \geqslant a))$$

$$= n^{-1} \sum_{k=1}^{n} \sigma_k E[|X| I(|X| \geqslant a/\sigma_k)]$$

$$\approx n^{-1} \sum_{k=1}^{n} \sigma_k e^{-a/2\sigma_k^2} \to \infty \text{ as } n \to \infty.$$

Example 4.5. Let $(X_n : n \geqslant 1)$ be a martingale difference sequence, such that $(|X_n|^p : n \geqslant 1)$ is uniformly integrable for some $1 \leqslant p < 2$. Let $(b_n : n \geqslant 1)$ be a sequence of real numbers such that

$$\lim_{a \to \infty} \lim_{n} n^{-1} \sum_{i=1}^{n} |b_i|^p I(|b_i| > a) = 0.$$

Then $n^{-1} E\left(\left| \sum_{i=1}^{n} b_i X_i \right|^p \right) \to 0$.

To see this, first note that there is a $K > 0$ such that

$$\sup_n \left[n^{-1} \sum_{i=1}^{n} |b_i|^p + E|X_n|^p \right] \leqslant K < \infty.$$

Given $\epsilon > 0$, choose M such that $\sup_n E[|X_n|^p I(|X_n| > M) < \epsilon$.

$$n^{-1} \sum_{i=1}^{n} E[|b_i X_i|^p I(|X_i b_i| > a)]$$

$$\leqslant n^{-1} \sum_{i=1}^{n} |b_i|^p E[|X_i|^p 1(|X_i| > M) I\left(|b_i| \leqslant \frac{a}{M}\right)$$

$$+ n^{-1} \sum_{i=1}^{n} |b_i|^p E[|X_i|^p] I(|b_i| > a/M)$$

$$\leqslant \epsilon K + K n^{-1} \sum_{i=1}^{n} |b_i|^p I\left(|b_i| > \frac{a}{M}\right) \to 0.$$

Thus $(|b_n X_n|^p : n \geqslant 1)$ is Cesàro uniformly integrable and the result follows from Theorem 2.2.

Example 4.6. Let $(X_n : 1 \leqslant i \leqslant k_n,\ n \geqslant 1)$ be a triangular array of M-dependent random variables. Then this array is an L_p-mixingale with $\psi_m = 0$ for $m > M$ and $c_{ni} = \|X_{ni}\|_p$. If $(|X_{ni}|^p : 1 \leqslant i \leqslant k_n,\ n \geqslant 1)$ is CUI, then $k_n^{-1} E(|S_n|^p) \to 0$. To see this, note that CUI of the array implies $\sup_n \frac{1}{k_n} \sum_{i=1}^{k_n} E|X_{ni}|^p < \infty$, which in turn implies $\sup k_n^{-1}\left(\sum_{i=1}^{k_n} c_{n,i}\right)^p < \infty$ and Theorem 3.4 applies.

Example 4.7. (McLeish's mixingales) If $(X_n : n \geqslant 1)$ is a mixingale in the sense of McLeish (1975) and $\lim\sup n^{-1}\left(\sum_{i=1}^{n} c_i\right)^p < \infty$ and $(|X_n|^p : n \geqslant 1)$ is CUI then $n^{-1} E(|S_n|^p) \to 0$ for $1 \leqslant p < 2$.

Example 4.8. Let $X_i = \sum_{j=1}^{i} a_{ij}\epsilon_{i-j}$, where (a_{ij}) are real numbers and (ϵ_i) are random variables and $Y_0 = 0$ and $X_0 = 0$. Define

$$b_{nk} = \sum_{i=k}^{n} a_{i,(i-k)}$$

$$Y_{nk} = \epsilon_k b_{nk},\ 1 \leqslant k \leqslant n,\ n \geqslant 1.$$

If $(Y_{nk} : 1 \leqslant k \leqslant n,\ n \geqslant 1)$ is an L_p-mixingale for some $1 \leqslant p < 2$, $E\epsilon_k = 0$ and $\lim_{n\to\infty} n^{-1}\left(\sum_{k=1}^{n} c_{nk}\right)^p < \infty$ then $n^{-1}E(|S_n|^p) \to 0$ by an application of the triangular version of Theorem 3.4. The above conditions are satisfied if (ϵ_i) itself is an L_p-mixingale with $\sup_n n^{-1}\left(\sum_{i=1}^{n} c_i\right)^p < \infty$ and $(b_{nk} : 1 \leqslant k \leqslant n,\ n \geqslant 1)$ is uniformly bounded. In particular this condition on (b_{nk}) is satisfied for stationary ARMA processes of any finite order.

Example 4.9. If we allow "infinite past" in Example 4.8, we have

$$X_i = \sum_{j=0}^{\infty} a_{ij}\,\epsilon_{i-j} = \sum_{k=1}^{n} a_k b_{nk} + \sum_{k=-\infty}^{0} \epsilon_k b_{nk}$$

where b_{nk} is as defined in Example 4.8. So provided $n^{-1} E\left(\left|\sum_{k=-\infty}^{0} \epsilon_k b_{nk}\right|\right)^p \to 0$, and conditions of Example 4.8 are satisfied, we have $n^{-1} E\left(\left|\sum_{i=1}^{n} X_i\right|^p\right) \to 0$.

The above extra condition is satisfied when $\sup_{k \leqslant 0} E|\epsilon_k|^p < \infty$ and $n^{-1} \sum_{k=-\infty}^{0} |b_{nk}|^p \to 0$.

Example 4.10. Let $X_t = \epsilon_t + \alpha\epsilon_{t-1}$, $t \geqslant 1$, $\epsilon_0 = 0$ be a moving average process. The method of moment estimator of α based on $X_1, \ldots, X_n$ is given by $\alpha_n = n^{-1} \sum_{t=-1}^{n} X_t X_{t-1}$, which yields,

$$\alpha_n - \alpha = \frac{\alpha}{n} \sum_{t=1}^{n} (\epsilon_t^2 - 1) - \frac{\alpha}{n} + n^{-1} \sum_{t=1}^{n} \epsilon_t \epsilon_{t-1}$$

$$+ \frac{\alpha}{n} \sum_{t=1}^{n} \epsilon_t \epsilon_{t-2} + \frac{\alpha^2}{n} \sum_{t=1}^{n} \epsilon_{t-1} \epsilon_{t-2}.$$

Note that $(|\epsilon_t \epsilon_{t-1}|^p : t \geqslant 1)$ is CUI if $(\epsilon_t^{2p} : t \geqslant 1)$ is CUI. Thus with suitable mixingale conditions on ϵ_t^2, $\epsilon_t \epsilon_{t-1}$ and with $E\epsilon_t^2 = 1$, $E\epsilon_{t-1} \epsilon_{t-1} = 0$, we have $E|\alpha_n - \alpha|^p \to 0$.

Example 4.11. The usual (least squares) estimator in an autoregressive process is much harder to deal with. Let $X_t = \theta X_{t-1} + \epsilon_t$, $t \geqslant 1$ where $|\theta| < 1$, $X_0 = 0$. The least squares estimate of θ is given by $\theta_n = \Sigma X_t X_{t-1} / \Sigma X_{t-1}^2$. This yields $\theta_n - \theta = \sum_{t=1}^{n} X_{t-1} \epsilon_t / \sum_{t=1}^{n} X_{t-1}^2$. Hence $\frac{\Sigma X_{t-1}^2}{n} (\theta_n - \theta) = n^{-1} \sum_{t=1}^{n} X_{t-1} \epsilon_t$. Thus if $(X_{t-1} \epsilon_t : t \geqslant 1)$ satisfies the conditions of Theorem 3.4 we have

$$n^{-1} E\left[\left(\sum_{t=1}^{n} X_{t-1}^2 |\theta_n - \theta|\right)^p\right] \to 0 \text{ as } n \to \infty.$$

In particular if $(\epsilon_t : t \geqslant 1)$ is a martingale difference sequence then so is $(X_{t-1} \epsilon_t : t \geqslant 1)$ and one needs to check only the CUI condition.

To conclude the convergence of $E|\theta_n - \theta|$, we proceed as follows:

Define $Y_n = (\theta_n - \theta)$, $Z_n = \dfrac{\Sigma X_{t-1}^2}{n}$. Then

$$E|Y_n| = E[|Y_n| I(|Z_n| \leqslant a_n)] + E[|Y_n| I(|Z_n| > a_n)]$$

$$\leqslant [E(|Y_n^2|)]^{1/2} [P(|Z_n| \leqslant a_n)]^{1/2} + E[|Y_n Z_n|]/a_n.$$

If $(X_{t-1}\epsilon_t : t \geqslant 1)$ satisfies conditions of Theorem 3.4 with $p = 1$ then for some sequence $a_n \to 0$, the second term above$\to 0$. To show that the first term$\to 0$, it is enough to show that $\sup\limits_n E(Y_n^2) < \infty$ and $P(Z_n \leqslant a_n) \to 0$

$$EY_n^2 = E\left[\left(\sum_{t=1}^{n} X_{t-1}\epsilon_t)^2 \Big/ \sum_{t=1}^{n} X_{t-1}^2\right)^2\right]$$

$$\leqslant 2E\left[\frac{\left(\sum\limits_{t=1}^{n-1} X_{t-1}\epsilon_t\right)^2}{(\Sigma X_{t-1}^2)^2}\right] + 2E\left[\frac{X_{n-1}^2\epsilon_n^2}{\left(\sum\limits_{l=1}^{n} X_{t-1}^2\right)^2}\right]$$

If $(\epsilon_n : n \geqslant 1)$ is such that $\sup\limits_n E(\epsilon_n^2 \mid \mathscr{F}_{n-1}) \leqslant K < \infty$ then

$$EY_n^2 \leqslant 2E\left[\left(\sum_{t=1}^{n-1} X_{t-1}\epsilon_t\right)^2 \Big/ \left(\sum_{t=1}^{n} X_{t-1}^2\right)^2\right].$$

However

$$\left(\sum_{t=1}^{n-1} X_{t-1}\epsilon_t\right)^2 \leqslant \left[\sum_{n=1}^{n-1} |X_{t-1}|(|\theta|\,|X_{t-1}| + |X_t|)\right]^2$$

$$\leqslant 2|\theta|^2\left(\sum_{t=1}^{n-1} X_{t-1}|^2\right)^2 + 2\left(\sum_{t=1}^{n-1} |X_{t-1}X_t|\right)$$

$$\leqslant (2|\theta|^2 + 32)\left(\sum_{t=1}^{n-1} X_t^2\right)$$

Hence $\sup EY_n^2 \leqslant K < \infty$. Now for large n,

$$P(Z_n \leqslant a_n) = P\left(\frac{\Sigma(\epsilon_t^2 - 1)}{n} + 2\theta\frac{\Sigma X_{t-1}\epsilon_t}{n} - \frac{X_n^2}{n} \leqslant a_n(1-\theta^2) - 1\right)$$

$$\leqslant P\left(\left|\sum_{t=1}^{n}\frac{(\epsilon_t^2-1)}{n}\right| \geqslant \left|\frac{a_n(1-\theta^2)-1}{3}\right|\right)$$

$$+P\left(\left|\sum_{t=1}^{n}\frac{X_{t-1}\epsilon_t}{n}\right| \geqslant \left|\frac{a_n(1-\theta^2)-1}{6}\right|\right)$$

$$+P\left(\frac{X_n^2}{n} \geqslant \left|\frac{a_n(1-\theta^2)-1}{3}\right|\right).$$

The second term$\to 0$ since $E\left|\dfrac{\Sigma X_{t-1}\epsilon_t}{n}\right| \to 0$. It is easily seen that the third term $\to 0$. Thus, if further $n^{-1} \sum_{t=1}^{n} \epsilon_t^2 \xrightarrow{P} 1$, the first term also $\to 0$. To summarize, if $(\epsilon_n, \mathcal{F}_n : n \geqslant 1)$ is a martingale difference sequence such that

$$\sup_n E(\epsilon_n^2 \mid \mathcal{F}_{n-1}) \leqslant K < \infty \text{ and } n^{-1} \sum_{t=1}^{n} \epsilon_t^2 \xrightarrow{P} 1,$$

then $E(|\theta_n - \theta|) \to 0$. Note that the above conditions are satisfied if $(\epsilon_t : t \geqslant 1)$ is i.i.d., $E\epsilon_t = 0$ and $E\epsilon_t^2 = 1$.

References

Andrews, D. W. K. (1989). Laws of large numbers for dependent non-identically distributed random variables. *Cowles Foundation paper No.* 717, Cowles Foundation, Yale University.

Chandra, T. K. (1989). Uniform integrability in the Cesàro sense and the weak law of large numbers. *Sankhyā*, Ser. A. **51**, 309–317.

Chandra, T. K. and Goswami, A. K. (1992). Cesàro uniform integrability and the strong law of large numbers. *Sankhyā*, Ser A, **54**, 215–231.

Chatterjee, S. D. (1969). An L_p-convergence theorem. *Ann. Math. Statist.*, **40**, 1068-1070.

Chow, Y. S. (1971). On the L_p-convergence for $n^{-1/p} S_n$, $0 < p < 2$. *Ann. Math. Statist.*, **42**, 393-394.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application*, New York, Academic Press.

Loève, M. (1977). *Probability Theory*, Vol. I, Springer-Verlag, New York.

McLeish, D. L. (1975). A maximal inequality and dependent strong laws. *Ann. Probab.*, **3**, 829-839.

Parzen, E. (1960). *Modern Probability Theory and its Applications*, John Wiley and Inc.

Pyke, R. and Root, D. (1968). On convergence in r-mean of normalized partial sums. *Ann. Math. Statist.*, **39**, 379-381.

Devision of Theoretical Statistics and Mathematics
Indian Statistical Institute
203 B. T. Road
Calcutta 7000035
India.