Laws of large numbers for dependent heterogeneous processes

de Jong, R.M.

1991

Faculteit der Economische Wetenschappen en Econometrie

Serie Research Memoranda

LAWS OF LARGE NUMBERS FOR DEPENDENT HETEROGENEOUS PROCESSES

R.M. de Jong

Research Memorandum 1991-88

December 1991

vrije Universiteit amsterdam

LAWS OF LARGE NUMBERS FOR DEPENDENT HETEROGENEOUS PROCESSES

by

R. M. de Jong*
Free University Amsterdam
Department of Econometrics
De Boelelaan 1105
1081 HV Amsterdam, The Netherlands

September 25, 1991

Abstract:
This paper presents both a weak and a strong law of large numbers for weakly dependent heterogeneous random variables. The laws presented for near-epoch dependent random variables allow for relaxation of the dependence conditions that are necessary in nonlinear least squares theory for dependent processes in order to ensure strong and weak consistency of the nonlinear least squares estimator.

*) I thank dr. P. Spreij for his helpful comments.

1. Introduction

Laws of large numbers for mixingale sequences are an essential tool in the proof of consistency and asymptotic normality of many parametric and nonparametric estimators under data dependence. A mixingale sequence can be viewed upon as an asymptotic equivalent of a martingale difference sequence. In a recent paper, Andrews [1] extends the mixingale concept and establishes some (weak and L_1) laws of large numbers (LLN's) for mixingales. Andrews' work extends the results of McLeish [6] who introduced the mixingale condition. In section 2 of this paper his conditions for convergence in L_1 of a mixingale sequence will be relaxed by making use of an inequality of Azuma [2] for martingale differences. Recently, Hansen [5] proved some new strong LLN's for mixingales using Andrews' mixingale concept. Those results extend the results of McLeish. In section 3 of this paper the conditions of both McLeish and Hansen for a strong LLN to hold will be relaxed substantially for the important case of sample averages and bounded indices of magnitude of the mixingale sequence. The results obtained here are complementary to those obtained by Hansen and McLeish since our result will not be powerful in the case of slowly increasing indices of magnitude of the mixingale sequence, which is of limited interest in many contexts. In section 4 we apply our results to near-epoch dependent sequences. The results in this section allow for a substantial relaxation of the conditions for strong consistency of (non)linear least squares estimators in the dependent case as listed in Gallant and White [4].

2. A weak law of large numbers for mixingales

Let (Ω,F,P) denote a probability space. Let $\{X_i : i \geq 1\}$ be a sequence of random variables on (Ω,F,P). Let $\{F_i : i = \ldots, 0, 1, \ldots\}$ be any nondecreasing sequence of sub σ-fields of F. Often one will take $F_i = \sigma(X_1, \ldots X_i)$ for $i \geq 1$ and $F_i = \{\varnothing, \Omega\}$ for $i \leq 0$. $E(X_i|F_j)$ denotes the conditional expectation of X_i given F_j. Whenever $E(X_i|F_j)$ is used we assume $E|X_i|$ to be finite. Let $\|X\|_p$ denote $(E|X|^p)^{1/p}$. Our law of large numbers makes use of the following lemma of Azuma [2]:

***Lemma* 1 (*Azuma*).**

If $\{X_i, F_i\}$ is a zero mean martingale difference sequence and

$$|X_i| \leq B \qquad\qquad a.s.$$

then for all $\varepsilon > 0$

$$P(|\sum_{i=1}^{n} X_i| > \varepsilon) \leq 2\exp(-\varepsilon^2/(2nB^2)).$$

In order to make this note almost self-contained, a proof of this result will be given in section 5.

Andrews [1] defines a L_p-mixingale as follows:

***Definition* 1.**

The sequence $\{X_i, F_i\}$ is called an L_p-mixingale if there exist nonnegative constants $\{c_i : i \geq 1\}$ and $\{\psi(m) : m \geq 0\}$ such that $\psi(m) \to 0$ as $m \to \infty$ and for all $i \geq 0$ and $m \geq 0$, we have

(a) $\| E(X_i | F_{i-m}) \|_p \leq c_i \psi(m)$

(b) $\| X_i - E(X_i | F_{i+m}) \|_p \leq c_i \psi(m+1)$

Note that according to this definition L_p-mixingales are necessarily mean zero random variables. The laws of large numbers of Andrews [1] typically require

$$\limsup_{m\to\infty} \limsup_{n\to\infty} (1/n) \sum_{i=1}^{n} c_i\, \psi(m) = 0$$

for a weak L_1-law of large numbers to hold. The author points out that his law of large numbers does not require a rate of decay on the mixingale numbers $\psi(m)$ to be imposed. On the other hand, he does require $\limsup_n (1/n) \sum_{i=1}^{n} c_i$ to be finite for a L_1-LLN to hold. Since the mixingale magnitude indices c_i can in many cases be assumed bounded, e.g. by $\sup_i E|X_i|$ as in Andrews [1]' theorem 1b, this in many cases will be a reasonable assumption. It can however be shown that only a tradeoff condition between a rate of decay for the mixingale numbers and the rate of increase of the latter sum is required for a law of large numbers to hold:

Theorem 1.

Suppose the sequence $\{X_i, F_i\}$ *is a* L_p*-mixingale such that* $\sup_i E|X_i|^p < \infty$, *for some* $p \geq 1$. *If* $m_n = o(n^{1/2}\log(n)^{-1/2})$ *is a sequence such that*

$$\lim_{n\to\infty}(1/n)\sum_{i=1}^{n} c_i^p\,\psi(m_n)^p = 0$$

Then

$$E|(1/n)\sum_{i=1}^{n} X_i|^p \to 0 \quad \textit{as } n\to\infty \ (\textit{and therefore } |(1/n)\sum_{i=1}^{n} X_i| \to 0 \textit{ in prob.})$$

Proof: *See section* 5.

3. A strong law of large numbers for mixingale sequences

Strong laws of large numbers for mixingales are elaborated upon by McLeish [6] and Hansen [5]. Their approach is proving a.s. convergence of

$$\sum_{i=1}^{n} X_i$$

under some conditions, and they obtain an almost sure law by imposing those conditions on X_i/i and conclude that

$$\sum_{i=1}^{n} X_i/i$$

converges almost surely to some random variable. The almost sure law now follows by the Kronecker lemma, i.e. if a_i is a sequence of positive real numbers and $a_i\to\infty$ if $i\to\infty$,

$$\sum_{i=1}^{n} X_i/a_i < \infty \qquad \Rightarrow \qquad (1/a_n)\sum_{i=1}^{n} X_i \to 0.$$

See for example Chung [3]. This approach typically results in conditions on the c_i sequence of the type

$$\sum_{i=1}^{\infty}(c_i/i)^2 < \infty.$$

As we argued before, in many important cases c_i can be assumed to be bounded over all i by some constant C. The above condition on the other hand would for example allow for c_i sequences of order $i^{1/2-\alpha}$, for some $\alpha > 0$, leaving some room for improvement. Our approach only works for sample averages, which might also be a reason why we succeeded in improving the conditions that have to be imposed. McLeish [6] assumes the above condition on the c_i and requires the $\psi(m)$ sequence to be of *size* $-1/2$. A sequence $\psi(m)$ is said to be of size $-\beta$, $\beta>0$, if $\psi(m)=O(m^{-\beta-\varepsilon})$ for some $\varepsilon>0$. McLeish [6] restricts attention to L_2-mixingales. However, if we wish to restrict attention to the case of indices of magnitude that are bounded or increasing very slowly we can improve upon those results by means of the corollaries to the following theorem:

***Theorem* 2.**

Suppose the sequence $\{X_i,F_i\}$ is a uniformly L_p-integrable L_p-mixingale for some $p>1$. Suppose we can find nonnegative monotonously increasing sequences B_i and m_i such that the following conditions hold:

(*A*) $$\sum_{i=1}^{\infty} i^{-1} B_i^{1-p} < \infty$$

(*B*) $$\sum_{i=1}^{\infty} c_i\, i^{-1} \psi(m_i) < \infty$$

(*C*) *for all* $\delta>0$, $$\sum_{n=1}^{\infty} [m_n] \exp\left(-n\delta^2/([m_n]^2 B_n^2)\right) < \infty$$

Then

$$(1/n)\sum_{i=1}^{n} X_i \rightarrow 0 \qquad \textit{almost surely.}$$

Remark: Note the refinement of the strong LLN of Hansen [5] that has taken place. The following corollary is now easily established:

***Corollary* 1:**

Suppose the sequence $\{X_i,F_i\}$ is a L_p-mixingale such that $\sup_i E|X_i|^p<\infty$, for some $p>1$. Suppose the indices of magnitude c_i of the mixingale sequence are uniformly bounded. Suppose $\psi(m)$ is of order $O(1/\log^{1+\alpha}(m))$ as $m\rightarrow\infty$, for some $\alpha>0$. Then

$$(1/n)\sum_{i=1}^{n} X_i \rightarrow 0 \qquad \textit{almost surely.}$$

Proof: See section 5.

The following simple corollary to theorem 2 now shows that sample averages of mixingale sequences converge for sequences $\psi(m)$ of arbitrary size if the magnitude indices c_i are uniformly bounded:

Corollary 2:

Under the same conditions as corollary 1 except for the condition on the $\psi(m)$ sequence, $(1/n)\sum_{i=1}^{n} X_i \to 0$ almost surely if $\psi(m)$ converges to zero at a polynomial rate, i.e. $\psi(m)=O(m^{\alpha})$ for some $\alpha<0$.

4. Near epoch dependence

In the theory of consistency and asymptotic normality of (non)parametric estimators for dependent samples use is made of the near epoch dependence concept. The introduction of the concept of near epoch dependence is motivated by two problems that occur when merely mixing sequences are considered. Firstly, as is well-known, functions of mixing processes are again mixing, but this is not necessarily the case if a function of the entire history of the mixing process is considered. Further, even simple AR(1) processes can fail to be either φ- or α-mixing. Gallant and White [4] define the near epoch dependence concept as follows:

Definition 2.

Let $\{X_i : \Omega \to \mathrm{R}\}$ be a sequence of random variables with $EX_i^2<\infty$, $i=1,\ldots$. Then $\{X_i\}$ is near epoch dependent on $\{V_i\}$ of size $-a$ if and only if

$$\nu_m = \sup_i \| X_i - E(X_i \mid V_{i-m},\ldots,V_{i+m}) \|_2$$

is of size $-a$.

An inequality can now be used to show that near epoch dependent random variables on some mixing sequence are mixingales. An application of theorem 2 establishes the following theorem:

Theorem 3.

Suppose X_i is near epoch dependent on $\{V_i\}$, where $\{V_i\}$ is mixing with coefficients α_m in the strong mixing case and coefficients φ_m in the uniform mixing case. Suppose both the mixing numbers α_m or φ_m and the NED numbers ν_m decrease at a polynomial rate. Suppose $\sup_i E|X_i|^{2+\delta}<\infty$ for some $\delta>0$. Then

$$(1/n)\sum_{i=1}^{n} X_i \to 0 \qquad a.s.$$

Proof: *See section* 5.

Gallant and White [4], in order to establish that near epoch dependent sequences are mixingales of size -1/2, need to impose sizes on the mixing coefficients and on the near epoch dependence numbers ν_m. Our theorem allows us to simply assume a polynomial decay of both sequences. A theorem like Gallant and White's theorem 3.18 involving functions of near epoch dependent processes is very complicated since taking functions of near epoch dependent processes in general does not keep near epoch dependence numbers in tact. In order to establish that a particular function of the near epoch dependent process is of size -1/2 (and, as a consequence, satisfies a strong LLN) a lot of conditions are imposed. Our result allows the observation that such functions of near epoch dependent sequences keep the mixingale numbers at a polynomial rate of decay, and as a consequence, a strong law of large numbers will hold. The price that has to be paid for this result is a slight strengthening of the moment conditions on the X_i sequence. This observation allows for relaxation of all conditions involving strong consistency throughout the Gallant and White [4] book and, since no provisions for near epoch dependence numbers need to be made once they have been assumed to decay at polynomial rate, could lead to considerable simplification of the theory of nonlinear least squares estimation for the case of dependent datagenerating processes.

5. Proofs

This section contains the proofs of the various theorems and lemma's. We will start with a proof of Azuma's inequality for martingale difference sequences. The proof is nearly identical to the proof of Hoeffding's inequality given in Pollard [7].

***Proof of lemma* 1:**

Consider $E(\exp(tX_i)|F_{i-1})$. *By convexity,*

$$\exp(tX_i) \le \exp(-tB)(B-X_i)/(2B)+\exp(tB)(X_i+B)/(2B)$$

so

$$E(\exp(tX_i)|F_{i-1}) \le \exp(-tB)/2+\exp(tB)/2$$

since

$$E(X_i|F_{i-1})=0.$$

Analogously to [7, Appendix B], it can now be shown that

$$\log\left[E\left(\exp(tX_i)|F_{i-1}\right)\right] \le (1/2)t^2B^2.$$

We will use the successive conditioning strategy that is employed also in proofs of central limit theorems for martingale differences (e.g. Pollard [7]). Then it easily follows that

$$E\exp(t\sum_{i=1}^{n}X_i) \le \exp(nt^2B^2/2).$$

Since, by Chebishev's inequality, for all $t>0$,

$$P\left(\sum_{i=1}^{n}X_i \ge \varepsilon\right) \le \exp(-\varepsilon t)\,E\exp(t\sum_{i=1}^{n}X_i) \le \exp(-\varepsilon t + nt^2B^2/2)$$

the result now follows by setting $t=\varepsilon/nB^2$ *and applying the same result to* $\{-X_i\}$. ∎

Azuma's inequality is central to the proof of the following weak law of large numbers.

Proof of theorem 1:

We will demonstrate the proof for the case that X_i *is* F_i*-measurable, which implies* $X_i=E(X_i|F_{i+m})$ *a.s. .The proof of the theorem in its full generality does not pose any additional problems, but does mess up the proof substantially.*

Note that, for all $B>0$ *and all integer-valued* $m>0$,

$$(1/n)\sum_{i=1}^{n}X_i = (1/n)\sum_{i=1}^{n}E(X_i|F_{i-m}) + (1/n)\sum_{i=1}^{n}X_iI(|X_i|<B) - E(X_iI(|X_i|<B)|F_{i-m})$$

$$+ (1/n)\sum_{i=1}^{n}X_iI(|X_i|\ge B) - (1/n)\sum_{i=1}^{n}E(X_iI(|X_i|\ge B)|F_{i-m})$$

$$= T_1+T_2+T_3+T_4 \qquad (5.1)$$

We will take $m=m(n)=m_n$. *By uniform integrability of* $|X_i|^p$ *we can pick* B *so large that*

$$\sup_i\ E|X_i|^p I(|X_i|\ge B) < \varepsilon.$$

In that case $E|T_3|^p<\varepsilon$ *and* $E|T_4|^p<\varepsilon$. *Clearly by the mixingale definition*

$$E|T_1|^p \le (1/n)\sum_{i=1}^{n}c_i^p\psi(m_n)^p.$$

We have by Azuma's inequality, for all $\delta > 0$

$$P\left[\left|(1/n)\sum_{i=1}^{n} X_i I(|X_i| < B) - E(X_i I(|X_i| < B)|F_{i-[m_n]})\right| > \delta \right]$$

$$\leq \sum_{j=0}^{[m_n]-1} P\left[\left|(1/n)\sum_{i=1}^{n} E(X_i I(|X_i| < B)|F_{i-j}) - E(X_i I(|X_i| < B)|F_{i-j-1})\right| > \delta/[m_n] \right]$$

$$\leq \sum_{j=0}^{[m_n]-1} \exp(-2n\delta^2/([m_n]^2(2B)^2)) \leq [m_n]\exp(-2n\delta^2/([m_n]^2(2B)^2))$$

$$\rightarrow 0$$

if $m_n = o(n^{1/2}(\log(n))^{-1/2})$. Because T_2 is bounded (by $2B$), it also converges to zero in L_p. So if we are able to find some sequence m_n that is $o(n^{1/2}\log(n)^{-1/2})$ for which

$$(1/n)\sum_{i=1}^{n} c_i^p \psi(m_n)^p \rightarrow 0 \qquad \text{as } n \rightarrow \infty$$

the LLN for the sample average will hold since ε was arbitrary. ■

The following proof shows that Azuma's inequality can be useful also in proving a strong law of large numbers:

Proof of theorem 2:
Again, we will demonstrate the proof for the case that X_i is F_i-measurable. Once again, for all $B > 0$ and integer-valued $m > 0$, consider equation (5.1). For proving a strong LLN we will make both B and m depend on i. In order to obtain the result we prove the following three lemma's:

Lemma 5.1: *If*

$$\sum_{i=1}^{\infty} i^{-1} B_i^{1-p} < \infty$$

then $T_3 \rightarrow 0$ and $T_4 \rightarrow 0$ almost surely.

Proof:

Let $S_n = \sum_{i=1}^{n} X_i I(|X_i| \geq B_i)/i$. Then, for any $\delta > 0$,

$$P(\max_{j\le m}|S_{n+j}-S_n|>\delta)\le(1/\delta)\sum_{i=n+1}^{\infty} E|X_i|I(|X_i|\ge B_i)/i$$
$$\le\sum_{i=n+1}^{\infty}(1/i\delta)\,E|X_i|^p\;B_i^{1-p}\;\to\;0\;\; as\;\; n\to\infty$$

if $\sum_{i=1}^{\infty} i^{-1}B_i^{1-p}<\infty$ *which is imposed. So we conclude by the Cauchy criterion that* S_n *converges almost surely, so by the Kronecker lemma,* $T_3\to 0$ *almost surely. The same argument holds for* T_4. ■

***Lemma* 5.2:** *If*

$$\sum_{i=1}^{\infty} c_i i^{-1}\psi(m_i)<\infty$$

then $T_1\to 0$ *almost surely.*

Proof: Let $S'_n=\sum_{i=1}^{n} E(X_i|F_{i-m_i})/i$. *Then*

$$P(\max_{j\le m}|S'_{n+j}-S'_n|>\delta)\le(1/\delta)\sum_{i=n+1}^{\infty} E|E(X_i|F_{i-m_i})|/i$$

$$\le\;(1/\delta)\sum_{i=n+1}^{\infty} c_i\psi(m_i)/i\;\to\;0$$

as $n\to\infty$ *if* $\sum_{i=1}^{\infty} c_i\,i^{-1}\,\psi(m_i)$ *converges, which is imposed. So* S'_n *converges almost surely to some random variable, by the Cauchy criterion. So* $T_1\to 0$ *almost surely by the Kronecker lemma.* ■

***Lemma* 5.3.** *Let* m_i *and* B_i *be strictly positive monotonously increasing sequences such that, for all* $\delta>0$,

$$\sum_{n=1}^{\infty}[m_n]\exp\Big(-n\delta^2/([m_n]^2B_n^2)\Big)<\infty$$

Suppose $m(1)=1$. *Then* $T_2\to 0$ *almost surely.*

Proof: In this proof we use Azuma's lemma, together with the Borel-Cantelli lemma. Let

$$q(j)=\inf_{l\in\{1,2,..\}}\{l\ge m^{-1}(j)\}$$

Note that

$$P\Big(\;|(1/n)\sum_{i=1}^{n}X_iI(|X_i|<B_i)-E(X_iI(|X_i|<B_i)|F_{i-[m_i]})|>\delta\Big)\;=$$

$$P\left[\ |(1/n)\sum_{i=1}^{n}\ \sum_{j=1}^{[m_i]} E(X_iI(|X_i|<B_i)|F_{i-j+1})-E(X_iI(|X_i|<B_i)|F_{i-j})|>\delta\right] =$$

$$P\left[\ |(1/n)\sum_{j=1}^{[m_n]}\ \sum_{i=q(j)}^{n} E(X_iI(|X_i|<B_i)|F_{i-j})-E(X_iI(|X_i|<B_i)|F_{i-j-1})|>\delta\right] \le$$

$$\sum_{j=1}^{[m_n]} P\left[\ |(1/n)\ \sum_{i=q(j)}^{n} E(X_iI(|X_i|<B_i)|F_{i-j})-E(X_iI(|X_i|<B_i)|F_{i-j-1})|>\delta/[m_n]\right] \le$$

$$\sum_{j=1}^{[m_n]} P\left[\ |(1/n)\sum_{i=1}^{n} E(X_iI(|X_i|<B_i)|F_{i-j})-E(X_iI(|X_i|<B_i)|F_{i-j-1})|>\delta/2[m_n]\right] +$$

$$P\left[\ |(1/n)\ \sum_{i=1}^{q(j)}\ E(X_iI(|X_i|<B_i)|F_{i-j})-E(X_iI(|X_i|<B_i)|F_{i-j-1})|>\delta/2[m_n]\right] \le$$

$$[m_n]\exp(-n\delta^2/(32[m_n]^2B_n^2)) + \sum_{j=1}^{[m_n]}\exp(-n^2\delta^2/(32q(j)\,[m_n]^2B_n^2\ ) \le$$

$$O([m_n]\exp(-n\delta^2/(32[m_n]^2B_n^2))$$

This implies that, by virtue of the Borel-cantelli lemma, $T_2 \to 0$ *almost surely.*
Corollary 1 *now follows by taking* $m_i = O(i^{\varepsilon})$, *for some* $0<\varepsilon<1/2$ *and setting* $B_i^{1-p} = O(1/\log^2(i))$. *Without loss of generality we can assume* m_i *and* B_i *satisfy the restrictions of lemma* 5.3. *If the sequences are taken in the way described above, the condition of lemma* 5.3 *is satisfied too.* ■

Finally, we will prove our result regarding near epoch dependent sequences.

Proof of theorem 3:
Andrews(1988) *shows that near epoch dependent sequences are mixingales with coefficients* $c_i = 2+\|X_i\|_{2+\delta}$ *and* $\psi(m) = \nu([m/2]) + 6\alpha([m/2])^{1/2-1/(2+\delta)}$. *Noting that* c_i *is uniformly bounded by construction and that* $\psi(m)$ *decreases polynomially if both* ν *and* α *do, the theorem follows by an application of corollary* 2 *to theorem* 2. ■

References

1. Andrews, D. W. K. Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4 (1988): 458–467.
2. Azuma, K. Weighted sums of certain dependent random variables. *Tohoku Mathematical Journal* 19 (1967): 357–367.
3. Chung, K.L. *A course in probability theory*. New York: Academic Press, 1974.
4. Gallant, A.R. and H. White. *A unified theory of estimation and inference for nonlinear dynamic models*. New York: Basil Blackwell, 1988.
5. Hansen, B.E. Strong laws for dependent heterogeneous processes. *Econometric Theory* 7 (1991): 213–221.
6. McLeish, D. L. A maximal inequality and dependent strong laws. *Annals of probability* 3 (1975): 829–839.
7. Pollard, D. *Convergence of stochastic processes*. New York: Springer-Verlag.