J. Appl. Probab. **41A** (2004), 231-238 (C. C. Heyde Festscrift, ed. J. Gani & E. Seneta)

Summability Methods and Negatively Associated Random Variables

N. H. BINGHAM and H. R. NILI SANI

Abstract. The paper studies convergence of sequences of negatively associated (NA) random variables under various summability methods. The results extend previously known results for independence, and complement known results for ϕ-mixing.

Keywords. Negatively dependent r.v.'s, Negatively Associated r.v.'s, Summability Methods. [1]

§1. Introduction

We will need the following classical summability methods. We refer to e.g. Hardy (1949) for background and details.

Cesàro methods (C, α), $\alpha \geq 1$:

$s_n \to s \quad (C, \alpha)$ means $\binom{n+\alpha}{n}^{-1} \sum_{k=0}^{n} \binom{n-k+\alpha-1}{n-k} s_k \to s \qquad (n \to \infty)$.

Abel method, (A):

$s_n \to s \quad (A)$ means $(1-x) \sum_{n=0}^{\infty} x^k s_k \to s \qquad (x \uparrow \infty)$.

Euler methods (E_q), $\qquad 0 < q < 1$:

---

[1] AMS: 60F15

$s_n \to s \quad (E_q)$ means $\sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} s_k \to s \qquad (n \to \infty)$.

Borel method (B):

$s_n \to s \quad (B)$ means $e^{-x} \sum_{n=0}^{\infty} \frac{x^n}{n!} s_n \to s \qquad (x \to \infty)$.

For the Riesz (typical) means below, see e.g. Chandrasekharan & Minakshisundaram (1952):

Riesz method $R(\lambda_n, 1)$. For $\lambda_n \uparrow \infty, \qquad s_n := \sum_0^n a_k$,

$s_n \to s \quad R(\lambda_n, 1)$ means $\frac{1}{x} \int_0^x \{\sum_{\lambda_n \le y} a_n\} dy \to s \qquad (x \to \infty)$.

Delayed averages. The Riesz method with $\lambda_n = e^{\sqrt{n}}$ is particularly important for us, and we shall abbreviate it to (R). It is equivalent (Bingham and Tenenbaum (1986); Bingham and Goldie (1988)) to the *delayed average* method (D), where

$s_n \to s \quad (D)$ means $\frac{1}{u\sqrt{n}} \sum_{n}^{n+u[\sqrt{n}]} s_k \to s$ for all $u > 0$.

The methods of Borel (B) and Euler (E_q) play an important role in many areas of mathematics. For instance, in summability theory they are perhaps the most important methods other than the Cesàro (C_α) and Abel (A) methods, and two chapters of the classic book of Hardy (1949) are devoted to them. In probability, the distinction between methods of Cesàro-Abel and Euler-Borel type may be seen from the following two laws of large numbers, the first of which extends Kolmogorov's strong Law.

Theorem L (Lai 1974). For $X, X_1, X_2, ...$ independent and identically distributed, the following are equivalent:

(i) $\quad E \mid X \mid < \infty$ and $EX = \mu$,

(ii) $\quad X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (C_\alpha)$ for some (all) $\alpha \ge 1$,

(iii) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (A)$.

Lai's theorem makes precise a sense in which the summability methods C and A are tied to existence of *first moments*.

Theorem C (Chow 1973). For $X, X_1, X_2, ...$ independent and identically distributed, the following are equivalent:

(i) $EX^2 < \infty$ and $EX = \mu$,

(ii) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (E_p)$ for some (all) $p \in (0, 1)$,

(iii) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (B)$.

(iv) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (R)$ (equivalently, (D)).

Chow's theorem makes precise a sense in which the summability methods of Borel and Euler are linked to existence of *second moments*, or *variances*.

These results were extended to other summability methods, including the *random-walk* methods, the *circle* methods and the *Valiron* methods V_a $(a > 0)$, in a series of papers by the first author and collaborators in the 1980s. For details and references, see Bingham (1984a), (1984b), (1986a), (1986b), (1989), Bingham and Maejima (1985), Bingham and Tenebaum (1986), Bingham and Goldie (1988). A sample result is

Theorem BM (Bingham and Maejima 1985). For $X, X_0, X_1, ...$ independent and identically distributed, the following are equivalent:

(i) $Var X < \infty$ and $EX = \mu$,

(ii) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (P)$, for some (any) random -walk method,

(iii) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (V_\alpha)$, for some (all) $a > 0$,

(iv) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (C)$, for some (any) circle method,

(v) $X_n \to \mu \;\; a.s. \;\; (n \to \infty) \;\; (R)$.

For further extensions - to Jakimovski methods $[F, d_n]$ and Karamata-Stirling methods $KS(\lambda)$ - see Bingham (1988), Bingham & Stadtmüller (1990) (where analogues of the law of the iterated logarithm are also given).

These results, and indeed laws of large numbers generally, are basically cancellation phenomena. For such cancellation, independence is sufficient, but it is by no means necessary. A study of the extension to one type of weak dependence - ϕ-mixing - was begun by Peligrad (1985), (1989) and continued by Kiesel (1997), (1998). We focus here on extension to a different type of dependence - negative dependence (ND) and negative association (NA). The intuition is that such negative dependence or association between random variables *assists cancellation.*

§2. Negative Dependence and Negative Association

Definition 1 (Lehmann 1966). The random variables $X_1, \cdots, X_n$ are said to be *negatively dependent* (ND) if both

$$P(X_1 \leq x_1, \cdots, X_n \leq x_n) \leq \Pi_{i=1}^n P(X_i \leq x_i)$$

and

$$P(X_1 \geq x_1, \cdots, X_n \geq x_n) \leq \Pi_{i=1}^n P(X_i \geq x_i)$$

for all $x_1, \cdots, x_n \in R$.

The random variables $X_1, \cdots, X_n (n \geq 2)$ are said to be *pairwise negatively*

dependent (PND) if (X_i, X_j) is ND for every $i \neq j,\ i,j = 1, ..., n$. Events $\{E_n\}$ are said to be negatively dependent if their indicator functions are.
Definition 2 (Joag-Dev and Proschan 1983). The random variables $X_1, \cdots, X_n$ are said to be *negatively associated* (NA for short) if for every pair of disjoint nonempty subsets A_1, A_2 of $\{1, ..., n\}$,

$$Cov(f_1(X_i, i \in A_1), f_2(X_i, i \in A_2)) \leq 0,$$

whenever f_1 and f_2 are coordinatewise increasing and the covariances exist.
Definition 3 (Newman 1984). The random variables $X_1, \cdots, X_n (n \geq 2)$ are said to be *linearly negative dependent* (LIND for short) if for every pair of disjoint nonempty subsets A_1, A_2 of $\{1, ..., n\}$ and positive λ_j's, $\sum_{j \in A_1} \lambda_j X_j$ and $\sum_{j \in A_2} \lambda_j X_j$ are negatively dependent.
An infinite sequence is NA, etc., if every finite subsequence is. We will need the following result; see Joag-Dev and Proschan (1983), Matula (1992), Bozorgnia et al. (1996).

Proposition. Increasing functions defined on disjoint subsets of a set of negatively associated r.v.'s are negatively associated.
Let $\{X_n\}$ be a sequence of ND r.v's. Then
(i) $Cov(X_i, X_j) \leq 0 \quad i \neq j$,
(ii) If $\{f_n\}$ is a sequence of Borel functions all of which are monotone increasing (or all monotone decreasing) then $\{f_n(X_n)\}$ is a sequence of ND r.v's.
(iii) The Borel-Cantelli lemma holds for ND events.

For other related negative-dependence concepts, we refer to Lehmann(1966), Block et.al.(1982), Ebrahimi and Ghosh(1981), Jogdeo and Patil (1975), Chandra and Ghosal (1996a), (1996b), Karlin and Rinott (1980) and the monograph Joe (1997).

We write F_i^j for $\sigma(X_k : \ i \leq j)$, and recall the ϕ-mixing coefficients

$$\phi(n) := \sup\{\mid P(B|A) - P(B) \mid : \ A \in F_{-\infty}^0, \ B \in F_n^\infty, \ P(A) > 0\}.$$

§3. Results

The results below extend known results from independence to negative association. The first extends Lai's Th. 1, the second Th. 1 of Chow (1973) and the third Theorems C and BM (extended to include $KS(\lambda)$ and $[F, d_n]$).

Theorem 1: For $\{X, X_n, n\}$ negatively associated and identically distributed, $\alpha \geq 1$, μ given reals, then for the statements

(i) $E(X) = \mu, \ E(\mid X \mid) < \infty$,

(ii) $X_n \to \mu \quad (C, 1) \ a.s.$,

(iii) $X_n \to \mu \quad (C, \alpha) \ a.s.$,

(iv) $X_n \to \mu \quad (A) \ a.s.$,

we have $(i) \iff (ii) \implies (iii) \implies (iv)$. Moreover, if $\phi(1) < 1/4$ then the above conditions are equivalent.

Theorem 2 concerns *delayed sums* - sums of the form $S_{u,v}$, where $S_{u,v} := \sum_{n=[u]}^{[u]+[v]} X_n$ for $u, v \geq 0$ Write also $\bar{S}_{u,v} := \max\{S_{u,j} : 0 \leq j \leq v\}$, $S_{u,v}^* := \max\{|S_{u,j}| : 0 \leq j \leq v\}$; we abbreviate these to S_v, $\bar{S}_v$, S_v^* when $u = 0$.

Theorem 2: Let X, X_n be negatively associated and identically distributed.
(i) If $EX = 0,\ E \mid X \mid^r < \infty$ for some $1 \le r \le 2$, and $E(X^+)^p < \infty$, for some $p \ge r$, then for every $0 < \alpha < r/p$

$$\limsup n^{-1/p} \bar{S}_{n,n^\alpha} = 0 \quad a.s.$$

(ii) If $EX = 0$ and $E \mid X \mid^p < \infty$ for some $p \ge 1$, then for every $0 < \alpha < min(1, 2/p)$

$$n^{-1/p} S^*_{n,n^\alpha} \to 0 \quad a.s.$$

(iii) If $E \mid X \mid^p < \infty$ for some $1 \le p < 2$ and $EX = 0$, or if $E \mid X \mid^p < \infty$ for some $0 < p < 1$,

$$n^{-1/p} S^*_{n,n} \to 0 \quad a.s.$$

In Theorem 3, we take $d_n \ge d$ for some $d > 0$ and $d_n = O(n)$, as in Bingham (1988).

Theorem 3: For $\{X, X_n\}$ negatively associated and identically distributed, consider the following statements:
(i) $VarX < \infty$, $E(X) = \mu$,
(ii) $X_n \to \mu\ a.s.\ E_q$ for some (all) $q \in (0, 1)$,
(iii) $X_n \to \mu\ a.s.\ (B)$,
(iv) $X_n \to \mu\ a.s.\ (R)$,
(v) $X_n \to \mu\ a.s.\ [F, d_n]$.
(vi) $X_n \to \mu\ a.s.\ (KS(\lambda))$.
Then (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii), (i) $\Rightarrow$ (iv) and (ii) $\Rightarrow$ (v) $\Rightarrow$ (vi). If further $\phi(1) < 1/4$, then the above statements are equivalent.

§4. Proofs

We begin with two preliminary results.

Theorem S (Shao (2000)). Let $\{X_i, 1 \leq i \leq n\}$ be negatively associated and let $\{X_i^*, 1 \leq i \leq n\}$ be a sequence of independent r.v.'s such that X_i and X_i^* have the same distribution for each $i = 1, 2, .., n$. Then

(i) $Ef(\sum_{i=1}^n X_i) \leq Ef(\sum_{i=1}^n X_i^*)$ for any convex function f for which the expectation on the right hand side exists.

(ii) $Ef(\max_{1\leq k\leq n} \sum_{i=1}^n X_i) \leq Ef(\max_{1\leq k\leq n} \sum_{i=1}^n X_i^*)$ for any non-decreasing convex function f for which the expectation on the right hand side exists.

Theorem C* (Chow (1973), Th. 3): If for some $1/2 < \alpha < 2/3$,

$$\max_{0\leq j\leq n^\alpha} \mid A_n + A_{n+1} ... + A_{n+j} \mid /max(j, \sqrt{n}) \to 0,$$

then for every $q \in (0, 1)$

$$A_n \to 0 \quad E_q.$$

Proof of Theorem 1.

(i)$\Leftrightarrow$ (ii) is a result of Matula (1992) (or Chandra and Ghosal (1996b)).

(ii)$\Rightarrow$ (iii) $\Rightarrow$(iv) are results from analysis (that is, are true generally and do not depend on the probabilistic structure); see Theorem 55 of Hardy (1949) and Bingham (1989).

The proof that (iv)$\Rightarrow$ (i) follows as in Kiesel (1997), Proposition. One uses the Chow-Lai lemma (Lemma 1 there - this needs only $\phi(1) < 1$), the Lévy maximal inequality (Lemma 2 there - this uses $\phi(1) < 1/4$) and Lemma 3 there (using the Borel-Cantelli lemma for negative association rather than for ϕ-mixing).

Proof of Theorem 2.

Let $\{Z_n\}$ be independent random variables with Z_n identically distributed to X_n for each n. One now proceeds by following the truncation argument used in Chow (1973) (proof of Th. 1), using Shao's Th. S for $f(x) = e^x$ to majorise the relevant sum ($E \exp\{\sum_1^m Y_{n+i}\}$ in Chow's notation - the Y_n are obtained from the X_n by truncation, centring and scaling) by the corresponding sum obtained from the independent sequence $\{Z_n\}$, to which Chow's argument applies. The other main ingredients are the Borel-Cantelli lemma (Proposition, §2) and the maximal inequality for submartingales (the independent sequence is a martingale; its exponential gives a submartingale; see Chung (1974), Th. 9.3.1, 9.4.1, or Hall & Heyde (1980), Th. 2.1).

Proof of Theorem 3. By Theorem 2(ii) and Theorem C*, (ii) holds if (i) does. That (iii) follows from (ii) is analysis (see e.g. Theorem 128 of Hardy (1949)). That (i) implies (iv) follows from Theorem 2 and equivalence of convergence of delayed averages and Riesz means.

Now if $d_n \geq \delta > 0$ for all large n, as assumed, $E(1/\delta) \subset [F, d_n]$ by a result of Meir (1963) (see also Bingham (1988)). Thus (ii) implies (v) and (vi) (indeed (ii) $\Rightarrow$ (v) $\Rightarrow$ (vi) for suitable q, λ).

For the converse, that any of (ii) - (vi) implies (i), we again follow Lemma 3 of Kiesel (1997). This time, we obtain

$$X_n^s/\sqrt{n} \to 0 \;\; a.s.,$$

(in place of $X_n^s/n \to 0$ $a.s.$ in Theorem 1), and conclude $E(X^2) < \infty$ by a

Borel-Cantelli argument (in place of $E|X| < \infty$ in Th. 1).

§5. Remarks

1. Baum-Katz law. The Baum-Katz law (on 'complete convergence' - rates of convergence in the strong law) is proved for ϕ-mixing in Peligrad (1985), (1989) and Kiesel (1997), (1998), in varying degrees of generality. The Baum-Katz law for negative association is given in Shao (2000), Th. 5. Another approach is via the strong laws for summability methods proved here (extended to the generality of Kiesel (1997)), combined with the analysis results of Bingham & Goldie (1988). For NQD, see Liang et al. (2002).

2. Other limit theorems for negative association. Note in particular the Marcinkiewicz-Zygmund strong law (Matula (1992), Chandra & Ghosal (1996b)), the three-series theorem (Matula (1992)), the functional central limit theorem (Shao (2000)) and the law of the iterated logarithm (Shao & Su (1999)).

3. First mixing coefficient. The conditions $\phi(1) < 1$, $\phi(1) < 1/4$ are from Peligrad (1989), Kiesel (1997). It would be interesting to know whether the second is tight, or an artefact of the proof. For the NQD case, the first is discussed in a paper in Chinese by Wang, Su & Liu (1998), cited in Liang et al. (2002) (English translation available from the authors).

4. Convergence and integrability. The results here preserve the tight link between a.s. convergence and *integrability* found in the independent case. No such link holds in general. For a collection of dramatic counter-examples, see Bingham (1986b), §4.

5. Maximal inequalities. That one can pass from sums to maxima of sums at no extra cost - in, e.g., Th. 2 here and in the Baum-Katz law - may be compared with similar situations for ϕ-mixing (see e.g. Peligrad (1989), Th. 1.1 and Remark 1.3), and with maximal inequalities generally.

6. Large deviations. The range $1/2 < \alpha < 2/3$ in Chow's Th. C* is characteristic of the study in detail of summability methods of Euler-Borel type. See Hardy (1949), Th. 137, Meyer-König (1949), or for a modern account, Korevaar (2003+), Ch. 6. The probabilistic background concerns large deviations; see e.g. Feller (1971), XVI.7, Ibragimov & Linnik (1971), §9.1.

Acknowledgements. The second author acknowledges the support of an Iranian Government scholarship and the hospitality of the Department of Mathematical Sciences at Brunel University.

References

Bingham, N.H. (1984a). On Valiron and Circle convergence. *Math. Z.* **186**, 273-286.

Bingham, N. H. (1984b). Tauberian theorems for summability methods of random-walk type. *J. London Math. Soc.* **30**, 281-287.

Bingham, N.H. (1986a). Extensions of the strong law. *Adv. Appl. Probab.*, Supplement, 27-36.

Bingham, N.H. (1986b). Summability methods and dependent strong laws. *Dependence in Probability and Statistics*, Birkhäuser, 291-300.

Bingham, N. H. (1988). Tauberian theorems for Jakimovski and Karamata-Stirling methods. *Mathematika* **35**, 216-224.

Bingham, N.H. (1989). Moving Averages. *Almost Everywhere Convergence I*, Academic Press, 131-144.

Bingham, N.H. and Maejima, M. (1985). Summability methods and almost sure convergence. *Z.Wahrsch. Verw. Gebiete.* **68**, 383-392.

Bingham, N.H. and Tenenbaum, G. (1986). Riesz and Valiron means and fractional moments. *Math. Proc. Camb. Phil. Soc.* **99**, 143-149.

Bingham, N.H. and Goldie, C.M. (1988). Riesz means and self-neglecting functions. *Math. Z. 199*, 443-454.

Bingham, N.H. and Stadtmüller, U. (1990). Jakimovski methods and almost-sure convergence. *Disorder in Physical Systems* (J. M. Hammersley Festschrift, ed. G. R. Grimmett & D. J. A. Welsh) 5-18, Oxford University Press.

Block, H.W., Savits, T.H., and Shaked, M. (1985). A concept of negative dependence using stochastic ordering. *Statist. Probab. Lett.* **3**, 81-86.

Bozorgnia, A., Patterson, R.F. and Taylor, R.L. (1996). Limit theorems for negatively dependent random variables. *World Congress Nonlinear Analysis* **92**, 1639-1650.

Chandra, T.K. and Ghosal, S. (1996a). The strong law of large numbers for weighted average under dependence assumptions. *J. Theor. Prob.* **9**, 797-809.

Chandra, T.K. and Ghosal, S. (1996b). Extensions of the strong law of large numbers of Marcinkiewicz and Zygmund for dependent variables. *Acta Math. Hungar.* **71**, 327-336.

Chandrasekharan, K. and Minakshisundaram, S. (1952). *Typical means*, Oxford Univ. Press.

Chow, Y.S. (1973). Delayed sums and Borel summability of independent identically distributed random variables. *Bull. Inst. Math. Acad. Sinica* **1**, 207-220.

Chung, K.-L. (1974). *A course in probability theory*, 2nd ed., Academic Press.

Ebrahimi, N. and Ghosh, M. (1981). Multivariate negative dependence. *Comm. Statist. A* **10**, 307-337.

Feller, W. (1971). *An introduction to probability theory and its applications*, Vol 2, 2nd ed., Wiley.

Hall, P. and Heyde, C.C. (1980). *Martingale limit theory and its application.* Academic Press.

Hardy, G.H. (1949). *Divergent Series.* Clarendon Press, Oxford.

Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and stationary sequences of random variables.* Wolters-Noordhoff.

Jakimovski, A. (1959). A generalization of the Lototosky method. *Michigan Mathematical Journal.* **6**, 277-29.

Joe, H. (1997). *Multivariate models and dependence concepts.* Chapman and Hall.

Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables with applications. *Ann. Statist.* **11**, 286-295.

Jogdeo, K. and Patil, G. P. (1975). Probability inequalities for certain multivariate discrete distributions. *Sankhyā B* **37**, 158-164.

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. *J. Multivar. Anal.* **10**, 499-516.

Kiesel, R. (1997). Strong laws and summability for sequences of $\phi-$ mixing random variables in Banach spaces. *Elect. Comm. in Probab.* **2**, 27-41.

Kiesel, R. (1998). Strong laws and summability for ϕ-mixing sequences of random variables. *J. Theoretical Probability 11*, 209-234.

Korevaar, J. (2003+). *A century of Tauberian theory.* Book,to appear.

Lai, T.L. (1974). Summability methods for independent, identically distributed random variables. *Proc. Amer. Math. Soc.* **45**, 253-261.

Lehmann, E.L. (1966). Some concepts of dependence, *Ann. Math. Statist.* **43**, 1137-1153.

Liang, H.-Y., Chen, Z.-J. and Su, C. (2002). Convergence of Jamison-type weighted sums of pairwise negatively quadrant dependent random variables. *Acta Math. Applicatae Sinica* **18**, 161-168.

Matula, P. (1992). A note on the almost sure convergence of sums of negatively dependent random variables. *Statist. Probab. Lett. 15*, 209-213.

Meir, A. (1963). On the $[F, d_n]$- transformation of A. Jakimovski. *Bulletin of the Research Council of Israel* (F: Mathematics) **10**, 165-187.

Meyer-König, W. (1949). Untersuchungen über einige verwandte Limitierungsverfahren, *Math. Z.* **52**, 257-304.

Newman, C.M. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables. in: Tong, Y.L. (Ed.), *Inequalities in Statistics and probability.* IMS Lecture Notes-Monograph Series, Vol **5** 127-140, Hayward, CA.

Peligrad, M. (1985). Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrschein.* **70**, 307-314.

Peligrad, M. (1989). The r-quick version of the strong law for stationary ϕ-mixing sequences. *Almost Everywhere Convergence* (ed. G. A. Edgar & L. Sucheston) 335-348, Academic Press.

Shao, Q.-M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theoret. probab.* **13**, 343-356.

Shao, Q.-M. and Su, C. (1999). The law of the iterated logarithm for negatively associated random variables. *Stoch. Proc. Appl.* **83**, 139-148.

Department of Mathematical Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, UK
nick.bingham@brunel.ac.uk

Department of Mathematics, Birjand University, Birjand, Iran
Nili@math.um.ac.ir