ALMOST SURE LIMIT THEOREMS
FOR DEPENDENT RANDOM VARIABLES

MICHAŁ SEWERYN

Faculty of Mathematics and Computer Science, University of Łódź
Banacha 22, 90-238 Łódź, Poland
E-mail: mseweryn@math.uni.lodz.pl

Abstract. For a sequence of dependent random variables $(X_k)_{k \in \mathbb{N}}$ we consider a large class of summability methods defined by R. Jajte in [11] as follows: For a pair of real-valued nonnegative functions $g, h : \mathbb{R}^+ \to \mathbb{R}^+$ we define a sequence of "weighted averages" $\frac{1}{g(n)} \sum_{k=1}^{n} \frac{X_k}{h(k)}$, where g and h satisfy some mild conditions. We investigate the almost sure behavior of such transformations. We also take a close look at the connection between the method of summation (that is the pair of functions (g, h)) and the coefficients that measure dependence between the random variables.

1. Introduction. Undeniably the classical Kolmogorov strong law of large numbers (for short SLLN) for iid random variables had a significant impact on the development of probability theory and mathematical statistics. This theorem was generalized in two different directions. The first path is the study of almost sure convergence of weighted averages, started by the work of Marcinkiewicz and Zygmund [12] in which one may find an extension of Kolmogorov's result to a wide class of means. On the other hand majority of researchers focused on extensions of the classical SLLN to the case of dependent random variables and so probably the most commonly known result in modern probability is the SLLN for pairwise independent identically distributed random variables due to Etemadi [8], which was lately generalized by Matula [16]. Our aim is to combine these two paths and state weighted strong laws of large numbers (for short WSLLN) for dependent random variables; we emphasize that our work is not new in this topic (see for instance [3], [15]).

First let us define the class of means we shall be dealing with. The following notation is due to Jajte [11].

NOTATION 1. We shall say that a pair of real valued, nonnegative functions (g, h) on $[0, \infty)$ satisfies the (F-J) condition if g is strictly increasing with range $(0, \infty)$ and moreover the function $\phi(x) := g(x)h(x)$ satisfies:

1. there exist $q \geq 0$ such that for any $x > y \geq q$ we have $\phi(x) > \phi(y)$,
2. there exist $r \geq 0, a \geq 0$ such that for any $y \geq r$ the following condition is satisfied $\phi^2(y) \sum_{n=y}^{\infty} \frac{1}{\phi^2(n)} \leq ay$.

For a given pair of functions (g, h) let us denote by $(\sigma_n)_{n \in \mathbb{N}}$ the sequence of weighted averages of a given sequence of random variables $(X_n)_{n \in \mathbb{N}}$, that is:

$$\sigma_n := \frac{1}{g(n)} \sum_{k=1}^{n} \frac{1}{h(k)}(X_k - m_k), \quad m_k := E(X_k \mathbf{I}\{|X_k| \leq \phi(k)\}).$$

The main part of our work is devoted to generalizations of the following theorem stated by Feller in 1945 [9] and revisited by Jajte in 2003 [11].

THEOREM 1.1. *Let* $(X_n)_{n \in \mathbb{N}}$ *be a sequence of iid random variables and* (g, h) *be a pair of functions that satisfies the condition* (F-J). *We have*

$$\lim_{n \to \infty} \sigma_n = 0 \ a.s. \iff E\phi^{-1}(|X_1|) < \infty.$$

In the present note we seek for optimal sufficient conditions for almost sure convergence of the sequence of weighted averages $(\sigma_n)_{n \in \mathbb{N}}$, thus for brevity we shall say that the sequence of dependent random variables $(X_n)_{n \in \mathbb{N}}$ satisfies the (g, h)-WSLLN if the corresponding sequence of weighted averages $(\sigma_n)_{n \in \mathbb{N}}$ converge to zero almost surely. For that purpose we need to extract a few smaller classes of weighted averages from the Feller-Jajte class.

It is worth mentioning that the Feller-Jajte WSLLN uncovers an interesting property of iid random variables, namely Theorem 1.1 implies:

REMARK 1. Let the function h be strictly increasing with range $(0, \infty)$ and (g, h) satisfy the condition (F-J), then:

$$\lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^{n} \frac{1}{h(k)}(X_k - EX_k) = 0 \ a.s. \iff \lim_{n \to \infty} \frac{1}{h(n)} \sum_{k=1}^{n} \frac{1}{g(k)}(X_k - EX_k) = 0 \ a.s,$$

where $(X_n)_{n \in \mathbb{N}}$ is as in the previous theorem.

We shall see that these two conditions are not always equivalent in the case of dependent random variables, which is not surprising as the pair (g, h) and (h, g) may define completely different summability methods. Let us also note that according to a theorem of Martikainen and Petrov [13] it remains an exercise to prove that no assumption on dependence of random variables is need in a special case of the previous theorem when there exist $r \geq 0, a \geq 0$, such that for any $y \geq r$ the following condition is satisfied $\phi(y) \sum_{n=y}^{\infty} \frac{1}{\phi(n)} \leq ay$.

In recent decades much attention was focused on investigating the asymptotic behavior of dependent random variables (see [4], [2], [7], [19], [1]). One can find a vast variety of conditions imposed on the dependence structure of the investigated sequence; we shall be interested in two classical ones: martingale difference sequences and α-mixing (i.e. strong

mixing) sequences. For the sake of completeness we recall the definitions of these two types of sequences below.

DEFINITION 1. Let $\mathcal{L}$, $\mathcal{F}$ be σ-fields on a probability space $(\Omega, \Im, P)$. We denote

$$\alpha(\mathcal{L}, \mathcal{F}) := \sup\{P(A \cap B) - P(A)P(B); A \in \mathcal{L}, B \in \mathcal{F}\}.$$

We shall say that a sequence of random variables $(X_k)_{k \in \mathbb{Z}}$ is α-*mixing* if

$$0 = \lim_{n \to \infty} \alpha_n := \lim_{n \to \infty} \left(\sup_{k \in \mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^k, \mathcal{F}_{k+n}^\infty) \right),$$

where

$$\mathcal{F}_{-\infty}^k = \sigma(X_l : l \le k), \mathcal{F}_{k+n}^\infty = \sigma(X_l : l \ge k + n).$$

DEFINITION 2. Let $(\mathcal{F}_n)_{n \in \mathbb{N}}$ be a filtration on a probability space. We shall say that a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ is a martingale difference sequence with respect to the filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ if the sequence of partial sums $(S_n)_{n \in \mathbb{N}}$ is a martingale with respect to the filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$, where $S_n := \sum_{k=1}^n X_k$ for $n = 1, 2, ...$

In the present paper we do not deal with associated (or negatively associated) sequences (for details see Matula [14]), for which the SLLN and WSLLN were studied in detail by Matula in [14], [15] (for further modifications see Chandra and Ghosal [5]). In view of Matula's generalization of the classical three series theorem (see Matula [14]) the proof of the remark below is an elementary exercise.

REMARK 2. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of identically distributed negatively associated random variables and (g, h) be a pair of functions that satisfies the condition (F-J). If $E\phi^{-1}(|X_1|) < \infty$, then $(X_n)_{n \in \mathbb{N}}$ satisfies the (g, h)-WSLLN.

Our main results generalize not only the Feller-Jajte SLLN, but also a few known WSLLN for dependent random variables (see Rio [18], Hall and Heyde [10] and references therein). Methods used in the proofs are modifications of those used by Rio [18] and Hall and Heyde [10], but our results cannot be derived from the above mentioned theorems.

2. WSLLN for α-mixing sequences. In the present section we state the main result for mixing sequences. For the sake of clarity and brevity we shall not give all the proofs here, instead we give reference to a forthcoming work, which contains more detailed discussion of this topic. Before stating our main result for α-mixing sequences we give a maximal inequality which is the main tool in the proof of the forthcoming theorem, but is also of interest on its own. First recall (what is known as the 'quantile transformation') that if X is a random variable with distribution function F then $X =^d F^{-1}(U)$, where $U \sim Unif([0,1])$ and F^{-1} denotes the generalized inverse of F. Moreover for any given nondecreasing sequence of nonnegative reals $(\alpha_n)_{n \in \mathbb{N} \cup \{0\}}$ and any random variable X let us make the following notation.

NOTATION 2. Let us denote by $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ a cadlag function given by

$$\alpha(t) := \alpha_{[t]}, t \in \mathbb{R}$$

and let

$$\alpha^{-1}(u) := \sup\{t \in \mathbb{R} : \alpha(t) > u\}, \ u \in \mathbb{R}^+.$$

Moreover in the same spirit let us denote by Q a generalized inverse function to the decay function, i.e. to $1 - F$, where F is the distribution function of X, defined as follows

$$Q(u) := \sup\{t \in \mathbb{R} : 1 - F(t) > u\}, \ u \in \mathbb{R}^+.$$

In the proof of the next lemma we shall use a covariance inequality due to E. Rio [17], which is as follows.

PROPOSITION 2.1. *Let X, Y be two square integrable random variables and let $\alpha := \alpha(\sigma(X), \sigma(Y))$ be as in Definition 1. We have*

$$|Cov(X, Y)| \leq 2 \int_0^{2\alpha} Q_X(t) Q_Y(t) dt,$$

where $Q_X(t)$ and $Q_Y(t)$ are generalized inverse functions of the decay functions of the random variables $|X|$ and $|Y|$ respectively and $\sigma(X), \sigma(Y)$ are σ-fields generated by X and Y respectively.

We are now ready to state the key lemma, a maximal inequality for α-mixing sequences, which is an adjustment of Theorem 4 of Rio [18] to our setting.

LEMMA 2.2. *Let $(X_n)_{n\in\mathbb{N}}$ be a α-mixing sequence and let $(\alpha_n)_{n\in\mathbb{N}}$ be the corresponding sequence of mixing coefficients as in Definition 1. Moreover let X_n be square integrable for each $n \in \mathbb{N}$ and let Q_n denote the inverse of the decay function of the random variable $|X_n|$. Let $h : \mathbb{R} \to \mathbb{R}$ be a nonnegative real function and $a_1 < a_2$ natural numbers. For any $x > 0$ and $r \in \mathbb{N}$ we have*

$$P\left(\max_{a_1 \leq p \leq a_2} \sum_{k=a_1}^{p} \frac{X_k}{h(k)} \geq 2x\right)$$

$$\leq \frac{2}{x^2} \sum_{k=a_1}^{a_2} \frac{1}{h^2(k)} \int_0^{2\alpha_0} Q_k^2(u) du + \frac{2}{x} \sum_{k=a_1}^{a_2} \frac{1}{h(k)} \int_0^{2\alpha_r} Q_k(u) du$$

$$+ \frac{8}{x^2} \sum_{k=a_1}^{a_2} \frac{1}{h(k)} \sum_{i=1}^{max(r-1, k-a_1)} \frac{1}{h(k-i)} \int_0^{2\alpha_i} Q_k^2(u) du$$

$$+ \frac{8}{x^2} \sum_{k=a_1}^{a_2} \frac{1}{h(k)} \sum_{i=1}^{max(r-1, k-a_1)} \frac{1}{h(k-i)} \int_0^{2\alpha_i} Q_{k-i}^2(u) du,$$

where $1/h(0) \equiv \mathbf{0}$.

Before stating a WSLLN for α-mixing sequences let us make some assumptions used in the sequel. Let us assume that there exist a nondecreasing nonnegative cadlag function ν such that for each $n, k \in \mathbb{N}, \ k < n$

$$\nu(k) := \sup_{k \leq n} \sum_{i=0}^{k-1} \frac{h(n)}{h(n-i)}, \nu(x) = \nu_{[x]}, x \in \mathbb{R}. \tag{1}$$

Moreover let us define a nonnegative sequence $(\theta_n)_{n\in\mathbb{N}}$ by the following condition

$$\phi(n) \sum_{k=1}^{n} \frac{1}{\phi(k)} = \theta_n, n \in \mathbb{N} \text{ and let } \theta(t) := \theta_{[t]}, \ t \in \mathbb{R}. \tag{2}$$

Let $(X_n)_{n\in\mathbb{N}}$ be an α-mixing sequence (as in Definition 1 and let α^{-1} be given in Notation 2), assume that there exist a nonnegative random variable X which dominates stochastically the sequence (X_n) i.e. such that for each $t \in \mathbb{R}$ $P(X > t) = \sup_{n\in\mathbb{N}} P(|X_n| > t)$ and $E\phi^{-1}(X) < \infty$. What is more let Q be the generalized inverse function to the decay function of the random variable X. Last let us assume that the pair of functions (g, h) satisfies the (F-J) condition.

THEOREM 2.3. *With the above notations and assumptions if*

(i) $\frac{\theta(n)}{n} = O(1)$ *and*

$$\int_0^1 \frac{\phi^{-1}(\nu(\alpha^{-1}(\frac{t}{2}))Q(t))}{\nu(\alpha^{-1}(\frac{t}{2}))}dt < \infty, \tag{3}$$

or if

(ii)

$$\int_0^1 \left|\frac{\theta(\phi^{-1}(\nu(\alpha^{-1}(\frac{t}{2}))Q(t))) - \nu(\alpha^{-1}(\frac{t}{2}))\theta(\phi^{-1}(Q(t)))}{\nu(\alpha^{-1}(\frac{t}{2}))}\right| dt < \infty, \tag{4}$$

(here $\frac{\theta(n)}{n}$ could be unbounded.),

then the sequence $(X_n)_{n\in\mathbb{N}}$ satisfies the (g, h)-WSLLN.

2.1. Optimality of the given conditions. Certainly a natural question arises on the optimality of the stated results, but the answer to this goes far beyond the scope of this short note. Nevertheless in the case of α-mixing sequences under the assumption that $\phi(n) := n^{\frac{1}{p}}$ and

$$ck \le \sum_{i=0}^{k-1} \frac{h(n)}{h(n-i)} \le dk \tag{5}$$

a construction similar to the one given in [6] may be performed, which in result shows the optimality of the stated results for mixing sequences. The counterexample is a well known (see [20]) result which indicates connections between the behavior of sums of independent random variables and Markov Chains. For convenience we shall be using the notion of β-mixing (for definition see below) instead of α-mixing in the sequel as it is well known (see [19]) that every stationary, ergodic Markov Chain is β-mixing.

DEFINITION 3. Let $\mathcal{L}, \mathcal{F}$ be σ-fields on a probability space $(\Omega, \Im, P)$. We denote

$$\beta(\mathcal{L}, \mathcal{F}) := \frac{1}{2}\sup \sum_{i\in I}\sum_{j\in J} P(A_i \cap B_j) - P(A_i)P(B_j),$$

where the supremum is taken over all possible finite partitions of the probability space $(A_i)_{i\in I}, (B_j)_{j\in J}$, which are $\mathcal{L}$ and $\mathcal{F}$ measurable respectively. We shall say that a sequence of random variables $(X_k)_{k\in\mathbb{Z}}$ is β-*mixing* if

$$0 = \lim_{n\to\infty}\beta_n := \lim_{n\to\infty}\left(\sup_{k\in\mathbb{Z}}\beta(\mathcal{F}_{-\infty}^k, \mathcal{F}_{k+n}^\infty)\right),$$

where

$$\mathcal{F}_{-\infty}^k = \sigma(X_l : l \le k), \quad \mathcal{F}_{k+n}^\infty = \sigma(X_l : l \ge k + n).$$

In the following lemma we postulate the existence of a Markov Chain with particularly "bad" properties from our point of view.

LEMMA 2.4. *For each $a > 1$ there exist a Markov Chain $(U_i)_{i \in \mathbb{Z}}$, where U_i is uniformly distributed over the interval $[0, 1]$ for each $i \in \mathbb{Z}$ and a sequence of numers $(\beta_n)_{n \in \mathbb{N}}$, such that $(U_i)_{i \in \mathbb{Z}}$ is β-mixing with mixing coefficients given by $(\beta_n)_{n \in \mathbb{N}}$ and what is more*

1. *$0 < \liminf_{n \to \infty} n^a \beta_n \le \limsup_{n \to \infty} n^a \beta_n < \infty$*
2. *For every integrable function $f : [0, 1] \to \mathbb{R}$ if*

$$\int_0^1 u^{\frac{1-p}{a}} (f(u))^p du = \infty, \tag{6}$$

   then

$$\lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^n \frac{1}{h(k)} (f(U_k) - E f(U_k)) = \infty \quad a.s.$$

It seems clear that by employing the so called 'quantile transformation' (see the beginning of the previous section) one may construct proper sequences that show optimality of the stated results in the case of power mixing rates. We refer the reader to a forthcoming work for the proof of this lemma and more details on the construction in the general case.

3. Main results for martingale difference sequences.
Now we turn our attention to martingale difference sequences and to this end we assume $(X_n)_{n \in \mathbb{N}}$ is a sequence of random variables and $(F_n)_{n=0}^{\infty}$ an increasing sequence of σ-fields such that X_n is measurable with respect to F_n for $n \in \mathbb{N}$. The next theorem is a generalization of a result due to Hall and Heyde [10].

THEOREM 3.1. *Let X be a random variable which dominates stochastically the sequence $(X_n)_{n \in \mathbb{N}}$ i.e. such that $P(|X_n| > x) \le P(|X| > x)$ for each $x \ge 0$ and $n \in \mathbb{N}$ and such that $E|X| < \infty$. Moreover let (g, h) be a pair of functions satisfying the (F-J) conditions. For θ given by (2) if $E(\theta(|X|)) < \infty$, then*

$$\lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^n \frac{X_k - E(X_k \mid F_{k-1})}{h(k)} = 0 \ a.s.$$

Proof. We set $\hat{X}_n := (X_n \wedge \phi(n)) \vee (-\phi(n))$ and $\hat{S}_n := \sum_{k=1}^n \frac{\hat{X}_k - E(\hat{X}_k \mid F_{k-1})}{h(k)}$, $n \in \mathbb{N}$, we have for $\epsilon > 0$

$$\sum_{k=1}^{\infty} P\left(\frac{\hat{X}_k - E(\hat{X}_k \mid F_{k-1})}{\phi^2(k)} > \epsilon \right) \le \sum_{k=1}^{\infty} \frac{E\hat{X}_n^2}{\phi^2(k)} \le EX^2 \sum_{k=1}^{\infty} \frac{1_{\{|X_k| \le \phi(k)\}}}{\phi^2(k)} \le E\phi^{-1}(|X|)$$

since the pair (g, h) satisfies the (F-J) condition and so due to Theorem 2.15 of Hall and Heyde [10] we have

$$\lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^n \frac{\hat{X}_k - E(\hat{X}_k \mid F_{k-1})}{h(k)} = 0 \text{ a.s.}$$

Since $\sum_{n=1}^{\infty} P(X_n \ne \hat{X}_n) < \infty$, then $\lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^n \frac{X_k - E(\hat{X}_k \mid F_{k-1})}{h(k)} = 0$ a.s. It re-

mains to prove that

$$0 = \lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^{n} \frac{|E(X_k \mid F_{k-1}) - E(\hat{X}_k \mid F_{k-1})|}{h(k)}$$

$$\leq \lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^{n} \frac{E(|X_k| 1_{\{|X_k| > \phi(k)\}} \mid F_{k-1})}{h(k)}$$

since

$$\sum_{k=1}^{\infty} \frac{E(|X_k| 1_{\{|X_k| > \phi(k)\}})}{\phi(k)} \leq E|X| \sum_{k=1}^{\infty} \frac{1_{\{|X_k| > \phi(k)\}}}{\phi(k)} \leq E\theta(|X|)$$

form the relation (2), then due to Kronecker's Lemma

$$\lim_{n \to \infty} \frac{1}{g(n)} \sum_{k=1}^{n} \frac{E(|X_k| 1_{\{|X_k| > \phi(k)\}} \mid F_{k-1})}{h(k)} = 0 \text{ a.s.}$$

(see Hall and Heyde [10] pages 31–39 for details). This completes the proof of the theorem. ■

REMARK 3. As far as martingale difference sequences are concerned a latest result of Sakhanenko [21] proves optimality of the classical SLLN (as stated in Hall and Heyde [10]) and it seems natural that a similar construction to the one in [21] should give counterexamples showing that the condition given in Theorem 3.1 (i.e. $E(\theta(|X|)) < \infty$) is optimal.

References

[1]   R. C. Bradley, *Basic properties of strong mixing conditions. A survey and some open problems*, Probab. Surveys 2 (2005), 107–144.

[2]   A. Bulinski and A. Shashkin, *Limit Theorems for Associated Random Fields and Related Systems*, World Scientific, 2007.

[3]   T. K. Chandra and S. Ghosal, *Some elementary strong laws of large numbers: a review*, Frontiers in Probability and Statistics (Calcutta 1994/95) Narosa, New Delhi, 1998, 61–81.

[4]   T. K. Chandra and S. Ghosal, *The strong law of large numbers for weighted averages under dependence assumptions*, J. Theor. Probab. 9 (1996), 797–809.

[5]   T. K. Chandra and S. Ghosal, *Extensions of the strong law of large numbers of Marcinkiewicz and Zygmund for dependent variables*, Acta Math. Hung. 71 (1996), 327–336.

[6]   P. Doukhan, P. Massart and E. Rio, *The functional central limit theorem for strongly mixing processes*, Ann. Inst. H. Poincaré Probab. Stat. 30 (1994), 63–82.

[7]   P. Doukhan, *Mixing: Properties and Examples*, Lecture Notes in Stat. 85, Springer-Verlag, 1994.

[8]   N. Etemadi, *An elementary proof of the strong law of large numbers*, Z. Wahrsch. Verw. Gebiete 55 (1981), 119–122.

[9]   W. Feller, *A limit theorem for random variables with infinite moments*, Amer. J. Math. 68 (1946), 257–262.

[10]   P. Hall and C. C. Heyde, *Martingale Limit Theory and its Applications*, Academic Press, 1980.

[11]   R. Jajte, *On the strong law of large numbers*, Ann. Probab. 31 (2003), 409–412.

[12] J. Marcinkiewicz and A. Zygmund, *Sur les fonctions indépendantes*, Fund. Math. 29 (1937), 60–90.

[13] A. I. Martikainen and V. V. Petrov, *On a theorem of Feller*, Theor. Probab. Appl. 25 (1980), 191–193.

[14] P. Matuła, *A note on the almost sure convergence of sums of negatively dependent random variables*, Stat. Probab. Letters 15 (1992), 209–213.

[15] P. Matuła, *Convergence of weighted averages of associated random variables*, Probab. Math. Stat. 16 (1996), 337–343.

[16] P. Matuła, *On some families of AQSI random variables and related strong law of large numbers*, Appl. Math. E-Notes 5 (2005), 31–35.

[17] E. Rio, *Covariance inequalities for strongly mixing processes*, Ann. Inst. H. Poincaré Probab. Stat. 29 (1993), 587–597.

[18] E. Rio, *A maximal inequality and dependent Marcinkiewicz-Zygmund strong laws*, Ann. Probab. 23 (1995), 918–937.

[19] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendantes*, Mathématiques & Applications 31, Springer, Berlin, 2000.

[20] M. Rosenblatt, *Random Processes*, Oxford University Press, 1962.

[21] A. I. Sakhanenko, *On coditions for SLLN for martingales with identically distributed increments*, Siberian Electronic Mathematical Reports 4 (2007), 547–552.