PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 19, Fasc. 1 (1999), pp. 181–201

PSEUDO-MARTINGALES*

BY

R. JAJTE AND A. PASZKIEWICZ (ŁÓDŹ)

Abstract. For a probability space $(\Omega, \mathscr{F}, P)$ and a filtration $(\mathfrak{A}_n)$ in Ω, we consider the sequences (X_n) of random variables satisfying the condition

$$E(X_{n+1}-X_n \mid \mathfrak{A}_n) = 0, \quad n = 1, 2, \ldots$$

In general, the process (X_n) is not required to be $(\mathfrak{A}_n)$ adapted and it is called a *pseudo-martingale*. We indicate simple and natural conditions implying a good asymptotic behaviour of pseudo-martingales. For example: let $(X_n, \mathfrak{A}_n)$ be a uniformly integrable pseudo-martingale with $\mathfrak{A}_n \nearrow \mathscr{F}$. Then $X_n \to X$ weakly in L_1, where

$$X = \lim_{n\to\infty} E(X_n \mid \mathfrak{A}_n).$$

Some approximation results for σ-fields are obtained with implications to pseudo-martingales. A number of examples is collected.

1. The main goal of this paper is to enlarge the area of applications of martingale methods. Description of a 'fair' game is the most classical interpretation of martingale so let us begin with a gambling situation.

1.1. Let us assume that the game is described by a martingale $(Y_n, \mathfrak{A}_n)$, that is we think of Y_n as total winnings of a player after n successive trials and $\mathfrak{A}_n$ contains $\sigma(Y_1, \ldots, Y_n)$, the σ-field generated by the random variables $Y_1, \ldots, Y_n$. The winnings Y_n may be 'invested' (bank, stocks, inflation) and then the player receives the amount $X_n = \varphi_n Y_n$, according to a random interest rate φ_n. It is commonly adapted that all values are 'discounted' in a way that the simplest formulas are obtained, so we can require that $E^{\mathfrak{A}_n} \varphi_n = 1$ $(n = 1, 2, \ldots)$. In particular, φ_n may be independent of the 'gambling information' $\mathfrak{A}_n$ and normalized. The sequence $(X_n, \mathfrak{A}_n)$ satisfies the condition

$$(*) \qquad \int_A X_n = \int_A X_{n+1} \quad \text{for } A \in \mathfrak{A}_n,$$

but X_n may be not $\mathfrak{A}_n$-measurable.

---

* R. Jajte and A. Paszkiewicz, Department of Probability and Statistics, Łódź University.
Research supported by KBN grant 2 PO3A 023 15.

1.2. For a martingale $(X_n, \mathfrak{B}_n)$, we may consider a sequence $(X_n, \mathfrak{A}_n)$ with $\mathfrak{B}_n \supset \mathfrak{A}_n \nearrow$. Then, as a rule, (X_n) is not $(\mathfrak{A}_n)$ adapted, but formula $(*)$ still holds.

1.3. Let $(X_1, X_2, \ldots)$ be a sequence of random variables satisfying the condition

$$E(X_{n+2}-X_{n+1} \mid X_1, \ldots, X_n) = 0, \qquad n \geqslant 1.$$

Then, for $A \in \mathfrak{A}_n = \sigma(X_1, \ldots, X_n)$ (with $\mathfrak{A}_1 = \{\emptyset, \Omega\}$) formula $(*)$ holds. Obviously, (X_n) is not a martingale if, for example, $X_n - X_{n-1} = Y_{n-1} + Y_n$ for a non-trivial independent (Y_j) with $EY_j = 0$.

1.4. For a martingale $(Y_n, \mathfrak{A}_n)$, let us consider a sequence (Δ_n) of random variables. Assume that Δ's are $\mathfrak{A}_n$-centered, i.e. $E^{\mathfrak{A}_n} \Delta_n = 0$ $(n = 1, 2, \ldots)$. Then the random variables $X_n = Y_n + \Delta_n$ are not $\mathfrak{A}_n$-measurable in general, but formula $(*)$ still holds.

2. The above simple examples suggest the following definition.

Let $(\Omega, \mathscr{F}, P)$ be a probability space.

2.1. Definition. Let $(X_n) \subset L_1(\Omega, \mathscr{F}, P)$ and let $(\mathfrak{A}_n)$ be an increasing sequence of sub-σ-fields of $\mathscr{F}$. We say that $(X_n, \mathfrak{A}_n)$ is a *pseudo-martingale* (or that (X_n) is $(\mathfrak{A}_n)$ *pseudo-martingale*) if

$$(1) \qquad \int_A X_n = \int_A X_{n+1} \quad \text{for } A \in \mathfrak{A}_n,\ n = 1, 2, \ldots$$

It should be stressed here that X_n are not required to be $\mathfrak{A}_n$-measurable. Shortly, we shall often write 'p.m.' instead of 'pseudo-martingale'.

2.2. Remarks. (a) It is easy to show that $(X_n, \mathfrak{A}_n)$ is a p.m. if and only if $(E^{\mathfrak{A}_n} X_n, \mathfrak{A}_n)$ is a martingale. Indeed, for $A \in \mathfrak{A}_n$, we have

$$\int_A X_n = \int_A X_{n+1} \quad \text{iff} \quad \int_A E^{\mathfrak{A}_n} X_n = \int_A E^{\mathfrak{A}_{n+1}} X_{n+1}.$$

(b) Observe that the general form of a p.m. for a filtration $(\mathfrak{A}_n)$ is given by the formula

$$(2) \qquad X_n = (Y_n - E^{\mathfrak{A}_n} Y_n) + Z_n$$

with an arbitrary $(Y_n) \subset L_1$ and $(Z_n, \mathfrak{A}_n)$ being a martingale. In other words, any p.m. (X_n) is a perturbed martingale (Z_n) with a perturbation $\Delta_n = Y_n - E^{\mathfrak{A}_n} Y_n$.

In spite of generality of the notion of pseudo-martingale, for important types of convergence one can formulate natural sufficient conditions for p.m.'s to be convergent. Let us start with the weak convergence.

2.3. Theorem. *If $(X_n, \mathfrak{A}_n)$ is a uniformly integrable pseudo-martingale with $\mathfrak{A}_n \nearrow \mathscr{F}$, then $X_n \to Y$ weakly in L_1, where $Y = \lim_{n\to\infty} E^{\mathfrak{A}_n} X_n$.*

Proof. Let us remark that the limit Y exists almost everywhere and belongs to L_1 since $E^{\mathfrak{A}_n} X_n$ is an L_1-bounded martingale. Let us take an arbitrary $g \in L_\infty(\Omega, \mathscr{F}, P)$ and let $\varepsilon > 0$. We can find

$$\bar{g} = \sum_{k=1}^{N} \lambda_k 1_{A_k} \quad \text{with } \|g - \bar{g}\|_\infty < \frac{\varepsilon}{3M},$$

where $M = \sup_n \int_\Omega |X_n - Y| < \infty$. Put $\Lambda = \max_{1\leqslant k\leqslant N} |\lambda_n|$. Since $\mathfrak{A}_n \nearrow \mathscr{F}$, we find in the field $\bigcup_{s\geqslant 0} \mathfrak{A}_s$ and, consequently, in some $\mathfrak{A}_{n_0}$ the sets B_k such that $P(A_k \triangle B_k) < \delta$, δ being a number such that $P(Z) < \delta$ implies

$$\int_Z |X_n - Y| < \frac{\varepsilon}{3\Lambda N}.$$

Let us put $\bar{\bar{g}} = \sum_{k=1}^N \lambda_k 1_{B_k}$ and write

$$\int X_n g - \int Yg = \int (X_n - Y)(g - \bar{g}) + \int (X_n - Y)(\bar{g} - \bar{\bar{g}}) + \int (X_n - Y)\bar{\bar{g}} = A + B + C.$$

Then we have

$$|A| < \frac{\varepsilon}{3}, \quad |B| < \Lambda \sum_{k=1}^N \int_{A_k \triangle B_k} |X_n - Y| < \frac{\varepsilon}{3}.$$

Finally, since $\bar{\bar{g}}$ is $\mathfrak{A}_{n_0}$-measurable, for $n \geqslant n_0$ we get

$$\int X_n \bar{\bar{g}} = \int \bar{\bar{g}} E^{\mathfrak{A}_n} X_n \to \int \bar{\bar{g}} Y,$$

so we have $|C| < \varepsilon/3$ for n large enough. ■

In Theorem 2.3 we assume that $\mathfrak{A}_n \nearrow \mathscr{F}$. Obviously, if $\mathfrak{A}_n \nearrow \mathfrak{A}_\infty \neq \mathscr{F}$, then the limit $Y = \lim_{n\to\infty} E^{\mathfrak{A}_n} X_n$ is $\mathfrak{A}_\infty$-measurable, and instead of the weak convergence in L_1 we should deal with the integrals

$$\int X_n g \to \int Yg \quad \text{for } g \in L_\infty(\Omega, \mathfrak{A}_\infty, P).$$

Assuming that (X_n) is L_p-bounded $(p > 1)$ one can obtain the weak convergence in L_p, evidently stronger than the weak one in L_1.

2.4. Proposition. *Let $(X_n, \mathfrak{A}_n)$ be a pseudo-martingale with $\mathfrak{A}_n \nearrow \mathscr{F}$, $\|X_n\|_p \leqslant C < \infty$, $p > 1$. Then $X_n \to Y$ weakly in L_p.*

Proof. For $p > 1$, L_p-boundedness implies uniform integrability, so the argument used in the proof of Theorem 2.3 can be repeated with slight modifications. ■

3. One of the ways to obtain some pointwise and mean convergence results for p.m.'s is to estimate in some sense the degree of non-measurability of X_n's with respect to $\mathfrak{A}_n$'s. To this end we adopt the following elementary definitions.

3.1. For two σ-fields $\mathfrak{B}$ and $\mathfrak{A}$, let us introduce (non-symmetric) functions $\varrho(\mathfrak{B}, \mathfrak{A})$ and $\bar{\varrho}(\mathfrak{B}, \mathfrak{A})$ by putting

$$\varrho(\mathfrak{B}, \mathfrak{A}) = \sup_{B\in\mathfrak{B}} \inf_{A\in\mathfrak{A}} P(A \triangle B)$$

and

$$\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) = \sup_{B\in\mathfrak{B}} \inf_{\substack{A\in\mathfrak{A} \\ A\supset B}} P(A \setminus B).$$

We shall say that $\mathfrak{B}$ is *ε-approximated* by $\mathfrak{A}$ (or that $\mathfrak{A}$ *ε-approximates* $\mathfrak{B}$) if

$$\varrho(\mathfrak{B}, \mathfrak{A}) < \varepsilon.$$

We say that $\mathfrak{B}$ is *ε-surrounded* by $\mathfrak{A}$ (or that $\mathfrak{A}$ *ε-surrounds* $\mathfrak{B}$) if

$$\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) < \varepsilon.$$

3.2. It turns out that the functions ϱ and $\bar{\varrho}$ just introduced are very useful in the analysis leading us to some pointwise and mean convergence theorems for p.m.'s. It is obvious that adding a random perturbation Δ to a random variable Y we can completely change a σ-field $\sigma(Y)$ even when Δ is close to zero uniformly. Since we shall deal mostly with σ-fields, we would like to stress that, fortunately, some important and commonly appearing perturbations vanish outside the sets of small probability (for example, like errors in transmission of digital dates). Looking at a p.m. (X_n) as a perturbation of a martingale (Y_n), i.e. writing $X_n = Y_n + \Delta_n$ with $E^{\mathfrak{A}_n} \Delta_n = 0$ (which is always the case, compare Remarks 2.2), it is natural to assume that the perturbations Δ_n have small supports in probability, say $P(\Delta_n \neq 0) < \varepsilon_n$ with ε_n's small enough. The assumptions of such kind imply that for p.m. $(X_n, \mathfrak{A}_n)$ the conditions like $\varrho(\sigma(X_n), \mathfrak{A}_n) = O(\varepsilon_n)$ or $\bar{\varrho}(\sigma(X_n), \mathfrak{A}_n) = O(\varepsilon_n)$ are satisfied.

Thus the irregular situation coming from the lack of $\mathfrak{A}_n$-measurability for terms X_n can be still under control via ε-surrounding or ε-approximation of $\sigma(X_n)$ by $\mathfrak{A}_n$. Consequently, it makes it possible to prove several limit theorems for p.m.'s.

4. In this section we prove some auxiliary results before formulating limit theorems.

4.1. PROPOSITION. *For sub-σ-fields $\mathfrak{A}$ and $\mathfrak{B}$ of $\mathcal{F}$ and a set $Z \in \mathcal{F}$, assume that $\varrho(\mathfrak{B}, \mathfrak{A}) < \varepsilon$ and that $P(Z \triangle A) < \varepsilon$ for some $A \in \mathfrak{A}$. Then*

$$\varrho(\sigma(\mathfrak{B} \cup \{Z\}), \mathfrak{A}) < 4\varepsilon.$$

Proof. Note that $\sigma(\mathfrak{B} \cup \{Z\})$ is a family of all sets of the form $B_1 \cup B_2$, where $B_1 = C_1 Z$ and $B_2 = C_2 Z^c$ for some $C_j \in \mathfrak{B}$. Let $A, A_1, A_2 \in \mathfrak{A}$ satisfy the inequalities $P(A_i \triangle C_i) < \varepsilon$, $i = 1, 2$, and $P(A \triangle Z) < \varepsilon$. Then, clearly,

$$P[(B_1 \cup B_2) \triangle (AA_1 \cup A^c A_2)] < 4\varepsilon. \blacksquare$$

The following lemma will be crucial in the sequel.

4.2. LEMMA. *Let $\mathfrak{B}$ and $\mathfrak{A}$ be two σ-fields. Assume that $\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) < \varepsilon$. If $B_1, \ldots, B_n \in \mathfrak{B}$ and B_j are pairwise disjoint, then there exist $A_1, \ldots, A_n \in \mathfrak{A}$ such that*

$$B_i \subset A_i \quad \text{and} \quad P\Big(\bigcup_{i=1}^{n} (A_i \setminus B_i)\Big) < 7\varepsilon.$$

This means that the last estimation does not depend on the number n of sets B_j.

Proof. We can assume (and do it) that the number n is even, say $n = 2k$ (otherwise, we add the empty set $\emptyset = B_{2k}$). In the whole proof the letter k has always the same meaning: $n = 2k$. In the sequel, I always denotes a subset of $\{1, \ldots, 2k\}$. For every $I \subset \{1, \ldots, 2k\}$, let us fix $A_I \in \mathfrak{A}_n$ such that $A_I \supset \bigcup_{i \in I} B_i$ and $P(A_I \setminus \bigcup_{i \in I} B_i) < \varepsilon$. For $i = 1, \ldots, 2k$, let us put

$$A_i = \bigcap_{i \in I} A_I \in \mathfrak{A}.$$

Step I. We have the estimation

$$P\Big(\bigcup_{i=1}^{n}(A_i\backslash B_i)\cap\Big(\bigcup_{i=1}^{n}B_i\Big)^c\Big)\leqslant P\Big(A_{\{1,\ldots,2k\}}\cap\Big(\bigcup_{i=1}^{n}B_i\Big)^c\Big)<\varepsilon$$

by the definition of A_i.

Step II. Let us fix an arbitrary well-ordering $\prec$ in the set of pairs $\{(i,j);\ i,j=1,\ldots,2k\}$. Let us write

$$p_{ij}=P\Big((A_i\backslash B_i)B_j\backslash\bigcup_{(r,s)\prec(i,j)}(A_r\backslash B_r)B_s\Big).$$

It is enough to show that

(3) $$\sum_{i\neq j}p_{ij}<6\varepsilon.$$

Taking, if necessary, a permutation $\sigma\colon\{1,\ldots,2k\}\to\{1,\ldots,2k\}$, a sequence $(B_{\sigma(1)},A_{\sigma(1)}),\ldots,(B_{\sigma(2k)},A_{\sigma(2k)})$ instead of $(B_1,A_1),\ldots,(B_{2k},A_{2k})$, and a matrix $(p_{\sigma(i),\sigma(j)})$ instead of $(p_{i,j})$, we can assume that

(4) $$\sum_{i\leqslant k,j\geqslant k+1}p_{ij}\geqslant\sum_{i\in I,j\notin I}p_{ij}$$

for every I with k elements. Moreover, we have

(5) $$\sum_{i\leqslant k,j\geqslant k+1}p_{ij}\leqslant P\Big(\bigcup_{i\leqslant k}(A_i\backslash B_i)\cap\bigcup_{j\geqslant k+1}B_j\Big)\leqslant P\Big(A_{\{1,\ldots,k\}}\backslash\bigcup_{i\leqslant k}B_i\Big)<\varepsilon.$$

Now, we shall show that

(6) $$\sum_{i,j\leqslant k,i\neq j}p_{ij}\leqslant 2\varepsilon.$$

Assume the contrary, i.e.

(7) $$\sum_{i,j\leqslant k,i\neq j}p_{ij}>2\varepsilon.$$

Then, by (7) and (5), we would have

(8) $$\sum_{i\leqslant k}\Big(\sum_{l\leqslant k,l\neq i}p_{li}-\sum_{j\geqslant k+1}p_{ij}\Big)>\varepsilon.$$

Inequalities (8) and (5) imply that there exist indices $i_0\leqslant k$ and $j_0\geqslant k+1$ such that

$$\alpha=\Big(\sum_{l\neq i_0,l\leqslant k}p_{li_0}-\sum_{j\geqslant k+1}p_{i_0j}\Big)-\sum_{i\leqslant k}p_{ij_0}>0$$

(since the suitable sums have the same length). But then we would have, for $I=\{1,\ldots,i_0-1,j_0,i_0+1,\ldots,k\}$,

$$\begin{aligned}&\sum_{i\leqslant k,j\geqslant k+1}p_{ij}-\sum_{i\in I,j\notin I}p_{ij}\\&\quad=\sum_{j\geqslant k+1}p_{i_0j}+\sum_{i\leqslant k}(p_{ij_0}-p_{i_0j_0})-\sum_{j\geqslant k+1,j\neq j_0}p_{j_0j}-\sum_{l\leqslant k,l\neq i_0}(p_{li_0}-p_{j_0i_0})\\&\quad=-\alpha-p_{i_0j_0}-\sum_{j\geqslant k+1,j\neq j_0}(p_{j_0j}-p_{j_0i_0})<0,\end{aligned}$$

which contradicts (4), and (6) is proved.

In a similar way we prove that

(9) $$\sum_{i,j\geqslant k+1, i\neq j} p_{ij} \leqslant 2\varepsilon.$$

Finally, we have (3) by (5), (6) and (9). Namely,

$$\sum_{i,j\leqslant 2k, i\neq j} p_{ij} = \sum_{j\geqslant k+1, i\leqslant k} p_{ij} + \sum_{i\geqslant k+1, j\leqslant k} p_{ij} + \sum_{i,j\leqslant k, i\neq k} p_{ij} + \sum_{i,j\geqslant k+1, i\neq j} p_{ij} < 6\varepsilon.$$

The proof is completed. ∎

Clearly, the inequality $\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) < \varepsilon$ implies the following condition:

(10) for any $B\in\mathfrak{B}$ there exists a set $Z\in\mathfrak{A}$ such that

$$B\cup Z\in\mathfrak{A}, \quad B\backslash Z\in\mathfrak{A}, \quad P(Z) < 2\varepsilon.$$

Indeed, it is enough to put $Z = A_1\cap A_2$ for $A_1, A_2\in\mathfrak{A}$, $A_1\supset B$, $A_2\supset B^c$, $P(A_1\backslash B) < \varepsilon$, and $P(A_2\backslash B^c) < \varepsilon$.

Using Lemma 4.2 we can prove a result much more general than (10).

4.3. Lemma. *If $\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) < \varepsilon$, then for a sequence $B_1, B_2, \ldots$ of sets in $\mathfrak{B}$ there exists a set $Z\in\mathfrak{A}$ satisfying the conditions*

$$B_i\cup Z,\ B_i\backslash Z\in\mathfrak{A}, \quad P(Z) < 14\varepsilon.$$

Proof. Let $A_i^n\in\mathfrak{A}$, $i = 1, 2, \ldots, n$, be such that

$$B_i\subset A_i^n, \quad P\Big(\bigcup_{i\leqslant n} (A_i^n\backslash B_i)\Big) < 7\varepsilon$$

(cf. Lemma 4.2). Then $A_i^+ = \bigcap_{n\geqslant i} A_i^n\in\mathfrak{A}$ satisfy

$$B_i\subset A_i^+, \quad P\Big(\bigcup_{i\geqslant 1} (A_i^+\backslash B_i)\Big) < 7\varepsilon.$$

Similarly, there exist $A_i^-\in\mathfrak{A}$ such that

$$B_i^c\subset A_i^-, \quad P\Big(\bigcup_{i\geqslant 1} (A_i^-\backslash B_i^c)\Big) < 7\varepsilon.$$

It is enough to put $Z = \bigcup_{i\geqslant 1} (A_i^+\cap A_i^-)$. ∎

4.4. Theorem. *If $\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) < \varepsilon$, and $\bar{\mathfrak{A}}$ denotes the completion of the σ-field $\mathfrak{A}$, then there exists a set $Z\in\bar{\mathfrak{A}}$ such that*

$$P(Z) < 14\varepsilon, \quad \mathfrak{B}\cap Z^c\subset\bar{\mathfrak{A}}\cap Z^c.$$

Proof. Let us write $\mathfrak{B} = (B_i, i\in I)$. We put

$$\varepsilon_i = \inf\{P(C);\ C,\ B_i\cup C,\ B_i\backslash C\in\mathfrak{A}\}.$$

Taking $C_i = \bigcap_{k\geqslant 1} C_{ik}$, where $C_{ik}, B_i\cup C_{ik}, B_i\backslash C_{ik}\in\mathfrak{A}$, $P(C_{ik}) < \varepsilon_i + 1/k$, we have

(11) $$P(C_i) = \varepsilon_i, \quad C_i,\ B_i\cup C_i,\ B_i\backslash C_i\in\mathfrak{A}.$$

Let $\prec$ denote a well-ordering of I. We are going to define a family $(D_j, j\in I) \subset \bar{\mathfrak{A}}$ such that

$$B_j \cup D_j,\ B_j \backslash D_j \in \bar{\mathfrak{A}}, \quad P(D_j) = \varepsilon_j, \tag{12}$$

$$\bigcup_{k \prec j} D_k \in \bar{\mathfrak{A}}, \tag{13}$$

$$D_j \backslash \bigcup_{k \prec j} D_k = \varnothing \quad \text{or} \quad P\big(D_j \backslash \bigcup_{k \prec j} D_k\big) > 0 \tag{14}$$

hold true.

Assume that the conditions (12)–(14) are fulfilled for D_j with $j \prec i$ (for a fixed i). Let us remark that

$$\bigcup_{j \prec i} D_j = \bigcup_{j \prec i} \big(D_j \backslash \bigcup_{k \prec j} D_k\big) = \bigcup_{\substack{j \prec i \\ P(D_j \backslash \bigcup_{k \prec j} D_k) > 0}} \big(D_j \backslash \bigcup_{k \prec j} D_k\big) \in \bar{\mathfrak{A}}.$$

We set

$$\begin{aligned} D_i &= C_i \quad \text{when } P\big(C_i \backslash \bigcup_{j \prec i} D_j\big) > 0, \\ D_i &= C_i \cap \bigcup_{j \prec i} D_j \quad \text{when } P\big(C_i \backslash \bigcup_{j \prec i} D_j\big) = 0. \end{aligned} \tag{15}$$

Thus the whole family $(D_j, j\in I)$ satisfying (12)–(14) has been defined by the induction principle.

We are in a position to define Z by putting

$$Z = \bigcup_{j\in I} D_j = \bigcup_{\substack{j \in I \\ D_j \backslash \bigcup_{k \prec j} D_k \neq \varnothing}} D_j. \tag{16}$$

By Lemma 4.3 there exists a set $Z_0 \in \mathfrak{A}$ such that $B_j \cup Z_0 \in \mathfrak{A}$ and $B_j \backslash Z_0 \in \mathfrak{A}$ for a countable set of indices $j \in I$ satisfying $D_j \backslash \bigcup_{k \prec j} D_k \neq \varnothing$, $P(Z_0) < 14\varepsilon$.

For any $j \in I$ satisfying $D_j \backslash \bigcup_{k \prec j} D_k \neq \varnothing$, by the minimality of $P(C_j)$ (according to (11)), we have $P(C_j) = P(C_j \cap Z_0)$ and, by (15), $P(D_j) = P(D_j \cap Z_0)$. Consequently, by (16), $P(Z) = P(Z \cap Z_0) < 14\varepsilon$. ∎

Remark. In the last theorem the completion $\bar{\mathfrak{A}}$ of $\mathfrak{A}$ is necessary, in general. Indeed, let us consider

$$(\Omega, \mathcal{F}, P) = \big([0, 1] \times [0, 1],\ \text{Borel}([0, 1] \times [0, 1]),\ \lambda^2\big),$$

$$\mathfrak{A} = \{Z \times [0, 1];\ Z - \text{countable or } [0, 1] \backslash Z - \text{countable}\},$$

$$\mathfrak{B} = \{Z \subset [0, 1] \times [0, 1];\ Z - \text{countable or } [0, 1] \times [0, 1] \backslash Z - \text{countable}\}.$$

Then $\bar{\varrho}(\mathfrak{B}, \mathfrak{A}) = 0$ and for any $Z \in \mathfrak{A}$ satisfying $B \cup Z$, $B \backslash Z \in \mathfrak{A}$ for all $B \in \mathfrak{B}$ we have $Z = [0, 1] \times [0, 1]$.

5. Now, we pass to prove several pointwise and mean convergence theorems for p.m.'s.

We start with the following result which is interesting in itself and also important as a tool in the proofs of limit theorems for p.m.'s.

5.1. PROPOSITION. *Let $X \in L_\infty(\Omega, \mathscr{F}, P)$ with $\|X\|_\infty \leqslant 1$. Let $\mathfrak{A}$ be a sub-σ-field of $\mathscr{F}$. Then $\varrho(\sigma(X), \mathfrak{A}) < \varepsilon$ implies $\|X - E^{\mathfrak{A}} X\|_1 < 8\varepsilon$.*

Proof. Step I. Let $X = 1_B$. Then there exists an $A \in \mathfrak{A}$ such that $P(A \triangle B) < \varepsilon$ and we have the estimation

$$\int_\Omega |1_B - E^{\mathfrak{A}} 1_B| \leqslant \int_{A \cap B} (1 - E^{\mathfrak{A}} 1_B) + \int_{A \setminus B} E^{\mathfrak{A}} 1_B + \int_{A^c} (1_B + E^{\mathfrak{A}} 1_B)$$

$$\leqslant \int_A (1 - E^{\mathfrak{A}} 1_B) + P(A \setminus B) + 2 \int_{A^c} 1_B \leqslant P(A) - \int_A 1_B + 3\varepsilon < 4\varepsilon.$$

Step II. Let $0 \leqslant X \leqslant 1$. For $\delta > 0$ we write $X = \sum_{i=1}^n \lambda_i 1_{B_i} + Y$ with $\lambda_i > 0$, $\sum \lambda_i \leqslant 1$, and $\|Y\|_1 < \delta$. Then we have

$$\int |X - E^{\mathfrak{A}} X| < \sum_{i=1}^n \int |\lambda_i 1_{B_i} - E^{\mathfrak{A}} \lambda_i 1_{B_i}| + 2\delta$$

$$\leqslant \sum_{i=1}^n \lambda_i 4\varepsilon + 2\delta \leqslant 4\varepsilon + 2\delta \to 4\varepsilon \qquad (\delta \to 0).$$

Step III. For $|X| \leqslant 1$, we write $X = X^+ - X^-$. ∎

5.2. COROLLARY. (a) *If $(X_n, \mathfrak{A}_n)$ is a p.m., $\sup_n \|X_n\|_\infty \leqslant K < \infty$ and $\varrho(\sigma(X_n), \mathfrak{A}_n) \to 0$, then $X_n \to X_\infty$ in L_1, where $X_\infty = \lim_{n\to\infty} E^{\mathfrak{A}_n} X_n$.*

(b) *If, additionally, $\sum_{n=1}^\infty \varrho(\sigma(X_n), \mathfrak{A}_n) < \infty$, then $X_n \to X_\infty$ a.e.*

Proof. (a) is evident. For (b) it is enough to apply the Beppo–Levy theorem. ∎

5.3. THEOREM. *Let $(X_n, \mathfrak{A}_n)$ be a p.m. Assume that $|X_n| \leqslant Y \in L_1$ and that*

$$\sum_n \varrho(\sigma(X_n), \mathfrak{A}_n) < \infty.$$

Then $X_n \to X_\infty$ with probability one, where $X_\infty = \lim_{n\to\infty} E^{\mathfrak{A}_n} X_n$.

Proof. The proof of our theorem is simpler when we assume additionally that

$$\sigma\Big(\bigcup_{n \geqslant 1} \mathfrak{A}_n\Big) = \mathscr{F}.$$

That is why we present two independent arguments. The first one for $\mathfrak{A}_n \nearrow \mathscr{F}$, the second one for $\mathfrak{A}_n \nearrow \mathfrak{A}_\infty$ ($\neq \mathscr{F}$, in general).

Case $\mathfrak{A}_n \nearrow \mathscr{F}$. Let us remark that

$$\sum \varrho(\sigma(X_n 1_{(|X_n| < c)}), \mathfrak{A}_n) < \infty \qquad \text{for any } c > 0. \tag{17}$$

For a given $\varepsilon > 0$, let us take $Z = (Y < c)$, $Z_n = (|X_n| < c)$ with c large enough to have $P(Z) > 1-\varepsilon$.

By (17) and Proposition 5.1, we have

$$(18) \qquad X_n 1_{Z_n} - E^{\mathfrak{A}_n}(X_n 1_{Z_n}) \to 0$$

with probability one, so almost uniformly.

By the martingale convergence theorem,

$$E^{\mathfrak{A}_n}(Y 1_{Z^c}) \to Y 1_{Z^c} \text{ a.e.}$$

Consequently,

$$1_Z E^{\mathfrak{A}_n}(X_n^+ 1_{Z_n^c}) \to 0 \text{ a.e.}$$

as

$$E^{\mathfrak{A}_n}(X_n^+ 1_{Z_n^c}) \leqslant E^{\mathfrak{A}_n} Y 1_{Z^c}.$$

Similarly,

$$1_Z E^{\mathfrak{A}_n}(X_n^- 1_{Z_n^c}) \to 0 \text{ a.e.}$$

Finally, we have

$$(X_n - E^{\mathfrak{A}_n} X_n) 1_Z = [X_n 1_{Z_n} - E^{\mathfrak{A}_n}(X_n 1_{Z_n})] 1_Z - (E^{\mathfrak{A}_n} X_n 1_{Z_n^c}) 1_Z \to 0 \text{ a.e.}$$

Thus $X_n - E^{\mathfrak{A}_n} X_n \to 0$ almost uniformly by the arbitrariness of $\varepsilon > 0$, which concludes the proof.

Case $\mathfrak{A}_n \nearrow \mathfrak{A}_\infty$ ($\neq \mathscr{F}$, in general). Take once more $Z = (Y < c)$, $Z_n = (|X_n| < c)$, $P(Z) > 1-\varepsilon$, and assume that $\int_{Z^c} Y < \varepsilon^2$. Then the almost sure convergence in (18) holds. Moreover,

$$(19) \qquad E^{\mathfrak{A}_n}(Y 1_{Z^c}) \to E^{\mathfrak{A}_\infty}(Y 1_{Z^c}) \text{ a.e.}, \qquad \int E^{\mathfrak{A}_\infty}(Y 1_{Z^c}) = \int_{Z^c} Y < \varepsilon^2.$$

Then

$$P\big(E^{\mathfrak{A}_\infty}(Y 1_{Z^c}) > \varepsilon\big) < \varepsilon$$

and, for n_0 large enough, we have

$$P\big(\sup_{n > n_0} E^{\mathfrak{A}_n} |X_n 1_{Z_n^c}| \geqslant 2\varepsilon\big) \leqslant P\big(\sup_{n > n_0} E^{\mathfrak{A}_n}(Y 1_{Z^c}) \geqslant 2\varepsilon\big)$$
$$\leqslant P\big(E^{\mathfrak{A}_\infty}(Y 1_{Z^c}) \geqslant \varepsilon\big) + \varepsilon < 2\varepsilon.$$

That means that

$$(20) \qquad P\big(\sup_{n > n_0} |X_n 1_{Z_n^c} - E^{\mathfrak{A}_n}(X_n 1_{Z_n^c})| > 2\varepsilon\big) < 3\varepsilon.$$

Summing up, for $\varepsilon > 0$, we choose $Z = (Y > c)$ with $P(Z) > 1-\varepsilon$, and n_0 such that (20) holds. Then we fix an $n_1 \geqslant n_0$ in such a way that

$$(21) \qquad P\big(\sup_{n > n_1} |X_n 1_{Z_n} - E^{\mathfrak{A}_n} X_n 1_{Z_n}| > \varepsilon\big) < \varepsilon$$

by (18). Combining (20) and (21) we get the desired result. ∎

5.4. THEOREM. *If* $(X_n, \mathfrak{A}_n)$ *is a p.m.,* (X_n) *is uniformly integrable and* $\varrho(\sigma(X_n), \mathfrak{A}_n) \to 0$, *then* $X_n \to X_\infty$ *in* L_1, *where* $X_\infty = \lim_{n\to\infty} E^{\mathfrak{A}_n} X_n$.

Proof. For $\varepsilon > 0$ let us fix c such that

$$\int_{(|X_n|>c)} |X_n| < \varepsilon \quad \text{for all } n.$$

We have, by Proposition 5.1,

$$\|X_n 1_{(|X_n|\leqslant c)} - E^{\mathfrak{A}_n}(X_n 1_{(|X_n|\leqslant c)})\|_1 \to 0.$$

On the other hand, for all n,

$$\|X_n 1_{(|X_n|>c)} - E^{\mathfrak{A}_n}(X_n 1_{(|X_n|>c)})\|_1 < 2\varepsilon,$$

which completes the proof by the arbitrariness of ε. ∎

The theorems that have just been proved described completely the consequences of ε-approximation of $\sigma(X_n)$ by $\mathfrak{A}_n$ for pseudo-martingales $(X_n, \mathfrak{A}_n)$ in the context of limit theorems.

In the sequel the consequences of ε-surrounding being the subject of Lemmas 4.2, 4.3 and Theorem 4.4 will be discussed.

5.5. THEOREM. *Let* $(X_n) \subset L_1(\Omega, \mathscr{F}, P)$ *and let* $(\mathfrak{A}_n)$ *be an arbitrary sequence of* σ*-fields. Then*

$$\sum_n \bar{\varrho}(\sigma(X_n), \mathfrak{A}_n) < \infty \textit{ implies } X_n - E^{\mathfrak{A}_n} X_n \to 0 \textit{ a.e.}$$

In particular, as an immediate consequence of the above theorem we have the following result:

5.6. THEOREM. *Let* $(X_n, \mathfrak{A}_n)$ *be an* L_1*-bounded p.m. If* $\sum_n \bar{\varrho}(\sigma(X_n), \mathfrak{A}_n) < \infty$, *then* $X_n \to Y$ $(= \lim_n E^{\mathfrak{A}_n} X_n)$ *with probability one.*

We present two different proofs of Theorem 5.5. The first one, based only on Lemma 4.2, is in the spirit of discrete mathematics. The second proof is an immediate application of Theorem 4.4. However, it should be stressed here that Theorem 4.4 is a consequence of Lemma 4.2 via transfinite induction.

Elementary proof of Theorem 5.5. Let us put $\varepsilon_n = \bar{\varrho}(\sigma(X_n), \mathfrak{A}_n)$ and fix a sequence k_n with $P(|X_n| \geqslant k_n) < \varepsilon_n$ $(n = 1, 2, \ldots)$. For a fixed n we take a partition

$$-k_n < \lambda_0 < \ldots < \lambda_r = k_n \quad \text{with } \lambda_i - \lambda_{i-1} < 1/n.$$

Here and in the sequel, to avoid excessive accumulation of indices, we often omit n when the dependence on n is clear.

Let

$$B_i = \{\lambda_{i-1} \leqslant X_n < \lambda_i\},\ i = 1, \ldots, r, \qquad B_0 = \{X_n < -k_n \vee X_n \geqslant k_n\}.$$

By Lemma 4.2, there exist $A_i \in \mathfrak{A}_n$ such that $A_i \supset B_i$, $i = 0, \ldots, r$, and

$$P\Big(\bigcup_{i=0}^{r} (A_i \setminus B_i)\Big) < 7\varepsilon_n.$$

Let $C_i = B_i \setminus \bigcup_{j \neq i} A_j$. Since $\bigcup_{i=0}^{r} B_i = \Omega$, we have $C_i = A_i \setminus \bigcup_{j \neq i} A_j$, and, consequently, $C_i \in \mathfrak{A}_n$.

Put $D_n = \bigcup_{i \geqslant 1} C_i$. Since $C_i \subset B_i$, for any fixed $i_0 \geqslant 1$ we have

$$B_{i_0} D_n^c = B_{i_0} \setminus C_{i_0} = B_{i_0} \bigcup_{j \neq i_0} A_j = B_{i_0} A_{i_0} \bigcup_{j \neq i_0} A_j$$
$$= B_{i_0} \bigcup_{j \neq i_0} A_{i_0} A_j.$$

Moreover, for $i \neq j$,

$$A_i A_j = [(A_i \setminus B_i) \cup B_i] [(A_j \setminus B_j) \cup B_j] \subset (A_i \setminus B_i) \cup (A_j \setminus B_j).$$

Thus

$$\bigcup_{\substack{i \neq j \\ i,j=0,\ldots,r}} A_i A_j \subset \bigcup_{j=0,\ldots,r} (A_j \setminus B_j).$$

Consequently, we have

$$P(D_n^c) \leqslant P\big(D_n^c \bigcup_{i=1}^{r} B_i\big) + P(B_0) \leqslant P\big(\bigcup_{i \neq j} A_i A_j\big) + P(B_0)$$
$$\leqslant P\big(\bigcup_{i=0,\ldots,r} (A_i \setminus B_i)\big) + P(B_0) < 8\varepsilon_n.$$

Clearly, we have

$$|X_n - E^{\mathfrak{A}_n} X_n| < 1/n \text{ on } D_n \quad (n = 1, 2, \ldots),$$

so $X_n - E^{\mathfrak{A}_n} X_n$ tend uniformly to zero outside the set $\bigcup_{n \geqslant N} D_n^c$ for arbitrary $N \geqslant 1$.

Since

$$P\big(\bigcup_{n \geqslant N} D_n^c\big) < 8 \sum_{n \geqslant N} \varepsilon_n,$$

this means that $X_n - E^{\mathfrak{A}_n} X_n \to 0$ a.e. if $\sum_n \varepsilon_n < \infty$, which proves our theorem. ∎

S h o r t p r o o f o f T h e o r e m 5.5. Completing $\mathfrak{A}$ if necessary and using the notation of Theorem 4.4 we infer that X_n and $E^{\mathfrak{A}_n} X_n$ coincide on the sets Z_n with $\sum_n P(Z_n^c) < \infty$. ∎

Theorems 5.3, 5.4 and 5.6 imply immediately the following result:

5.7. Corollary. *Let $(X_n, \mathfrak{A}_n)$ be a pseudo-martingale with $\mathfrak{A}_n \nearrow \mathscr{F}$. Then:*
(a) *If (X_n) is uniformly integrable, then*

$$\varrho(\mathscr{F}, \mathfrak{A}_n) \to 0 \textit{ implies } X_n \to X_\infty \textit{ in } L_1.$$

(b) *If (X_n) is L_1-bounded, then*

$$\sum_n \bar{\varrho}(\mathscr{F}, \mathfrak{A}_n) < \infty \textit{ implies } X_n \to X_\infty \textit{ a.e.}$$

(c) *If* $|X_n| \leqslant Y$ *with* $Y \in L_1$, *then*

$$\sum_n \varrho(\mathscr{F}, \mathfrak{A}_n) < \infty \text{ implies } X_n \to X_\infty \text{ a.e.}$$

We always have $X_\infty = \lim_{n\to\infty} E^{\mathfrak{A}_n} X_n$ *a.e.*

We close this section discussing the role of "boundedness-type" conditions in limit theorems for pseudo-martingales.

Obviously, for the conditions

(i) $|X_n| \leqslant K$, $K \in \boldsymbol{R}$;

(ii) $|X_n| \leqslant Y$, $Y \in L_1$;

(iii) (X_n) is uniformly integrable;

(iv) (X_n) is bounded in L_1;

we have the implications

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(iv)}.$$

In our theorems one cannot use weaker assumptions.

We construct suitable elementary examples.

The probability space $(\Omega, \mathscr{F}, P)$ always equals $([0, 1), \text{Borel } [0, 1), \lambda)$, and for any $\omega \in [0, 1)$ we keep the notation $\omega = \varepsilon_1/2^1 + \varepsilon_2/2^2 + \ldots$, $\varepsilon_1, \varepsilon_2, \ldots$ are equal to 0 or 1 with infinite number of 0's.

5.8. Example. There exists a pseudo-martingale $(X_n, \mathfrak{A}_n)$ such that (X_n) is uniformly integrable, $\sum_n \varrho(\sigma(X_n), \mathfrak{A}_n) < \infty$ and, with probability one, X_n does not converge (cf. Theorem 5.3).

It is enough to put

$$(\mathfrak{A}_1, \mathfrak{A}_2, \ldots) = (\underbrace{\sigma(\varepsilon_1), \sigma(\varepsilon_1)}_{2 \text{ times}}, \ldots, (\underbrace{\sigma(\varepsilon_1, \ldots, \varepsilon_k), \ldots, \sigma(\varepsilon_1, \ldots, \varepsilon_k)}_{2^k \text{ times}}, \ldots),$$

$$(X_1, X_2, \ldots) = (\underbrace{2 \cdot 1_{[0, \frac{1}{2^2})} + 1_{[0, \frac{1}{2})^c}, 2 \cdot 1_{[\frac{1}{2}, \frac{1}{2} + \frac{1}{2^2})} + 1_{[\frac{1}{2}, 1)^c}}_{2 \text{ terms}}, \ldots,$$

$$\underbrace{2^k \cdot 1_{[0, \frac{1}{2^{2k}})} + 1_{[0, \frac{1}{2^k})^c}, 2^k \cdot 1_{[\frac{1}{2^k}, \frac{1}{2^k} + \frac{1}{2^{2k}})} + 1_{[\frac{1}{2^k}, \frac{2}{2^k})^c}, \ldots, 2^k \cdot 1_{[\frac{2^k-1}{2^k}, \frac{2^k-1}{2^k} + \frac{1}{2^{2k}})} + 1_{[\frac{2^k-1}{2^k}, 1)^c}}_{2^k \text{ terms}}, \ldots).$$

Then $E^{\mathfrak{A}_n} X_n = 1$, and $\limsup_{n\to\infty} X_n = 1$, $\liminf_{n\to\infty} X_n = 0$ for each non-dyadic ω. Moreover, for any $Z \in \mathscr{F}$ and for X_n from the k-th row, we have $\int_Z X_n \leqslant 2^{-k} + \lambda(Z)$, and (X_n) is uniformly integrable.

5.9. Example. Taking the same sequence $(\mathfrak{A}_n)$, we can define (X_n) in such a way that $\|X_n\|_1 \leqslant 2$, $E^{\mathfrak{A}_n} X_n = 0$, $\sum \varrho(\sigma(X_n), \mathfrak{A}_n) < \infty$, $\bar{\varrho}(\sigma(X_n), \mathfrak{A}_n) \to 0$, and $\lim_{n\to\infty} \|X_n - X_{n+1}\|_1 = 4$ (cf. Theorems 5.4 and 5.3).

Namely, define

$$(X_1, X_2, \ldots) = (\underbrace{2^2 \cdot 1_{[0,\frac{1}{2^2})} - 2 \cdot 1_{[0,\frac{1}{2})},\ 2^2 \cdot 1_{[\frac{1}{2},\frac{1}{2}+\frac{1}{2^2})} - 2 \cdot 1_{[\frac{1}{2},1)}}_{2\ \text{terms}}, \ldots,$$

$$\underbrace{2^{2k} \cdot 1_{[0,\frac{1}{2^{2k}})} - 2^k \cdot 1_{[0,\frac{1}{2^k})},\ 2^{2k} \cdot 1_{[\frac{1}{2^k},\frac{1}{2^k}+\frac{1}{2^{2k}})} - 2^k \cdot 1_{[\frac{1}{2^k},\frac{2}{2^k})}, \ldots, 2^{2k} \cdot 1_{[\frac{2^k-1}{2^k},\frac{2^k-1}{2^k}+\frac{1}{2^{2k}})} - 2^k \cdot 1_{[\frac{2^k-1}{2^k},1)}}_{2^k\ \text{terms}}, \ldots).$$

Then $\|X_n - X_{n+1}\|_1 = \|X_n\|_1 + \|X_{n+1}\|_1 \to 4$. The rest is obvious.

5.10. EXAMPLE. The martingale $X_n = n\, 1_{[0,1/n)}$ is an example of a well-known possibility $\|X_n\|_1 \leqslant 1$, $\|X_n - \lim_{n\to\infty} X_n\|_1 = 1$ (cf. Theorem 5.6).

In the next sections we discuss several examples of pseudo-martingales.

6. A large class of p.m.'s $(X_n, \mathfrak{A}_n)$ is given by the formula

$$X_n = \varphi_n Y_n, \quad n \geqslant 1, \tag{22}$$

where $(Y_n, \mathfrak{A}_n)$ is a martingale and (φ_n) a sequence of positive functions satisfying the condition

$$E^{\mathfrak{A}_n} \varphi_n = 1, \quad n \geqslant 1 \tag{23}$$

(as usual, $\mathfrak{A}_n \nearrow$).

Putting

$$\pi_n f = \varphi_n E^{\mathfrak{A}_n} f, \quad f \in L_1, \tag{24}$$

we have, for (X_n) in (22)

$$X_n = \pi_n X_{n+1}, \quad n \geqslant 1, \tag{25}$$

(π_n) being a sequence of positive contractive projections in L_1 satisfying the condition

$$\pi_n \pi_{n+1} = \pi_n, \quad n \geqslant 1. \tag{26}$$

Obviously, the projections π_n play here a similar role to conditional expectations in the classical theory of martingales. The sequence (X_n) in (25) (with $\mathfrak{A}_n \nearrow$, and φ_n satisfying (23)) will be called a *(π_n)-martingale*. It is worth noting here that, for fixed $(\mathfrak{A}_n)$ and (φ_n), the class of all (π_n)-martingales coincides with the class of sequences (X_n) satisfying (22). Indeed, it is enough to put $Y_n = E^{\mathfrak{A}_n} X_{n+1}$, $n \geqslant 1$.

Formula (24) turns out to be quite general. Using well known Ando's formula for contractive projections one can prove that any positive contractive projection in L_p $(1 \leqslant p \neq 2)$ is of the form (24) with $f \in L_p$ (cf. [1], [4], [5]). Consequently, for any sequence (π_n) of positive contractive projections in L_p

($1 \leqslant p \neq 2$), satisfying (26), and any $f \in L_p$, the sequence $(\pi_n f)$ is a p.m. with respect to σ-fields $\mathfrak{A}_n$ appearing in the description of π_n.

It should be stressed here that formula (26) means that $\ker \pi_{n+1} \subseteq \ker \pi_n$, but the sequence of projections (π_n) is not increasing in general, i.e. $\pi_{n+1}\pi_n \neq \pi_n$.

For monotone sequences of contractive projections (π_n) in L_p-spaces we have the following result analogous to the well-known theorem in the classical theory of martingales.

6.1. Theorem. *Let $p > 1$. If (π_n) is a monotone (decreasing or increasing) sequence of positive contractive projections in L_p, then $\pi_n f$ converges a.e. for all $f \in L_p$* (cf. [4]).

In contradistinction to the classical theory of martingales, there exists a decreasing sequence of positive contractive projections (P_n) in L_1 such that $P_n 1$ does not converge a.e. Indeed, we have the following example (cf. [4]).

6.2. Example. Denote by $\sigma(\chi)$ a σ-field generated by a Borel function χ on $[0, 1]$. We put

$$\Psi_i^n(x) = \begin{cases} \frac{1}{2} & \text{for } 0 \leqslant x < (n+i-1)/(2n), \\ (n+i+1)/2 & \text{for } (n+i-1)/(2n) \leqslant x < (n+i)/(2n), \\ 1 & \text{for } (n+i)/(2n) \leqslant x \leqslant 1, \end{cases}$$

$$\chi_i^n(x) = \max\big(x, (n+i)/(2n)\big).$$

Obviously, we have $\Psi_i^n \geqslant \frac{1}{2}$ on $[0, 1]$ and

$$E^{\sigma(\chi_j^n)} \Psi_i^n = 1 \quad \text{for } 1 \leqslant i \leqslant j \leqslant n, \tag{27}$$

$$\max(\Psi_1^n, \Psi_1^n \Psi_2^n, \ldots, \Psi_1^n \ldots \Psi_n^n) > n/2 \quad \text{on } [\tfrac{1}{2}, 1].$$

Let us write $c_0 = 1$, $n(1) = 4$ and, assuming that $c_0, \ldots, c_{k-1}$, $n(1), \ldots, n(k)$ have already been fixed, take $c_k = \min_{0 \leqslant x \leqslant 1}\big(\Psi_1^{n(k)} \ldots \Psi_{n(k)}^{n(k)}\big)(x)$ and $n(k+1)$ satisfying $\frac{1}{2} n(k+1) c_0 \cdot \ldots \cdot c_k > 2$.

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables uniformly distributed on $[0, 1]$. It is enough to take

$$(P_i) = (P_1^1, \ldots, P_{n(1)}^1, P_1^2, \ldots, P_{n(2)}^2, \ldots)$$

with

$$P_i^k = \prod_{l<k} \Big(\prod_{j \leqslant n(l)} \Psi_j^l\Big) \circ X_l \cdot (\Psi_1^k \ldots \Psi_i^k) \circ X_k\, E^{\sigma(\chi_i^{n(k)} \circ X_k, X_{k+1}, \ldots)}.$$

As usual, $\sigma(X_k, X_{k+1}, \ldots)$ denotes the σ-field generated by $X_k, X_{k+1}, \ldots$ It can be easily observed, by (27), that (P_i) is a decreasing sequence of projections (positive and contractive in L_1) but, for the function $\mathbb{1}(\omega) \equiv 1$,

$\max(P_1^k \mathbb{1}, \ldots, P_{n(k)}^k \mathbb{1}) > 2$ on the set $(\frac{1}{2} \leqslant X_k \leqslant 1)$. Thus $P_n \mathbb{1}$ do not converge on a set of probability 1, by the Borel–Cantelli lemma.

6.3. Remark. Let us observe that if (P_n) is an increasing sequence of positive, contractive projections in L_1, then $P_n f$ converges a.e. for all $f \in L_1$ (cf. [4]).

7. Indexing by stopping times is an important way of producing new martingales from a given one (cf. [6]). We have the following analogue for pseudo-martingales.

7.1. THEOREM. *Let $(X_n, \mathfrak{A}_n)$ be a pseudo-martingale and let $\tau_1, \tau_2, \ldots$ be an increasing sequence of finite stopping times relative to $(\mathfrak{A}_n)$ (i.e. $(\tau_j = k) \in \mathfrak{A}_k$ for all k and j). Let $Y_n = X_{\tau_n}$, $n \geqslant 1$. Assume that*

(a) $$Y_n \in L_1, \quad n \geqslant 1,$$

(b) $$\liminf_{k\to\infty} \int_{(\tau_n > k)} |X_k| = 0 \quad \textit{for } n = 1, 2, \ldots$$

Then $(Y_n, \mathfrak{B}_n)$ is a pseudo-martingale, where

$$\mathfrak{B}_n = \{A \in \mathscr{F}\colon\ A \cap (\tau_n = k) \in \mathfrak{A}_k \textit{ for all } k = 1, 2, \ldots\}.$$

Proof. Clearly, the sequence of σ-fields $(\mathfrak{B}_n)$ is increasing. We have to show that

(28) $$\int_B (Y_{n+1} - Y_n) = 0 \quad \text{for } B \in \mathfrak{B}_n.$$

Let $B \in \mathfrak{B}_n$. Since $B = \bigcup_{s=1}^{\infty} B(\tau_n = s)$, it is enough to show (28) for $C_s = B \cap (\tau_n = s) \in \mathfrak{A}_s$ (instead of B). Let us fix $k > s$. Since $(\tau_n = s) \subset (\tau_{n+1} \geqslant s)$, we have

$$\int_{C_s} Y_{n+1} = \sum_{i=s}^{k} \int_{C_s \cap (\tau_{n+1} = i)} Y_{n+1} + \int_{C_s \cap (\tau_{n+1} > k)} Y_{n+1}$$
$$= \sum_{i=s}^{k} \int_{C_s \cap (\tau_{n+1} = i)} X_i + \int_{C_s \cap (\tau_{n+1} > k)} X_k - \int_{C_s \cap (\tau_{n+1} > k)} X_k - Y_{n+1}.$$

But

$$\int_{C_s \cap (\tau_{n+1} = k)} X_k + \int_{C_s \cap (\tau_{n+1} > k)} X_k = \int_{C_s \cap (\tau_{n+1} \geqslant k)} X_k = \int_{C_s \cap (\tau_{n+1} \geqslant k)} X_{k-1}$$
$$= \int_{C_s \cap (\tau_{n+1} > k-1)} X_{k-1},$$

since $(\tau_{n+1} \geqslant k) \in \mathfrak{A}_{k-1}$, which together with $s < k$ gives $D_s \cap (\tau_{n+1} \geqslant k) \in \mathfrak{A}_{k-1}$.

Combining now the integrals

$$\int_{C_s \cap (\tau_{n+1} > k-1)} X_{k-1} \quad \text{and} \quad \int_{C_s \cap (\tau_{n+1} = k-1)} X_{k-1},$$

we get $\int_{C_s \cap (\tau_{n+1} > k-2)} X_{k-2}$. Continuing this procedure, we finally get

$$\int_{C_s} Y_{n+1} = \int_{C_s \cap (\tau_{n+1} \geqslant s)} X_s - \int_{C_s \cap (\tau_{n+1} > k)} (X_k - Y_{n+1}).$$

The integrals $\int_{C_s \cap (\tau_{n+1} > k)} X_k$ tend to zero as $k \to \infty$. Since $(\tau_{n+1} > k) \to \emptyset$, we have $\int_{C_s \cap (\tau_{n+1} > k)} Y_{n+1} \to 0$ as $k \to \infty$. Observing that $X_s = Y_n$ on $D_s = D_s \cap (\tau_{n+1} \geqslant s)$, we get (18) for $B = C_s$, which completes the proof. ■

8. Natural examples of pseudo-martingales appear when martingales are perturbed in some way. In Section 6 we discussed a large class of pseudo--martingales of such a type closely related to the theory of projections in L_p-spaces $(1 \leqslant p \neq 2)$.

8.1. The simplest but rather natural example is given by a random linear transformation of a martingale $(Y_n, \mathfrak{A}_n)$ of the form $(X_n, \mathfrak{A}_n)$ with

$$X_n = u_n Y_n + v_n, \tag{29}$$

where u_n, v_n, and Y_n are independent for each n, with $Eu_n = 1$ and $Ev_n = 0$ (mostly (u_n) and (v_n) are i.i.d. (noise) and independent of (Y_n)).

8.2. Disturbance of a martingale on some sets often leads to a p.m. Let $(X_n, \mathfrak{A}_n)$ be a martingale with $EX_n = 0$, and let $(\overline{X}_n)$ be a centered sequence in L_1. We fix a sequence (D_n) of events independent of (X_n) and $(\overline{X}_n)$, and set $B_n = D_n \cap D_{n+1} \cap \ldots$ Let us put

$$Z_n = 1_{D_n} X_n + 1_{D_n^c} \overline{X}_n \tag{30}$$

and

$$\mathscr{B}_n = \bigcup_{k=1}^{n} \{A \cap (B_k - B_{k-1});\ A \in \mathfrak{A}_n\} \cup \{B_n^c, \emptyset\}. \tag{31}$$

Then $(Z_n, \mathscr{B}_n)$ is a pseudo-martingale. Indeed, clearly, $(\mathscr{B}_n)$ is an increasing sequence of σ-fields. Moreover, for $A \in \mathfrak{A}_n$, $k \leqslant n$, we have

$$\int_{A \cap (B_k - B_{k-1})} (Z_{n+1} - Z_n) = P(B_k - B_{k-1}) \int_A (X_{n+1} - X_n) = 0.$$

Since $(X_n, \overline{X}_n, n \geqslant 1)$ are centered and independent of (D_n), $n \geqslant 1$, we have

$$\int_{B_n^c} Z_{n+1} = \int_{B_n^c \cap B_{n+1}} X_{n+1} + \int_{B_{n+1}^c} \overline{X}_{n+1} = 0.$$

Now, we specify the sets D_n by putting

$$D_n = \{f(Y_1) \ldots f(Y_n) \geqslant g(Y_1) \ldots g(Y_n)\}, \tag{32}$$

where $(Y_1, Y_2, \ldots)$ is a sequence of i.i.d. random variables, f is a density of distribution of Y_j, and g is another density of distribution on the real line.

Keeping the previous notation we assume that (Y_n) are independent of (X_n) and $(\overline{X}_n)$. Let us assume that the random variables $f(Y_j)$ and $g(Y_j)$ have finite variances. Then, for $\xi_i = \ln f(Y_i) - \ln g(Y_i)$, standard calculations lead us to the following estimations:

$$P(D_n^c) = P\big(f(Y_1)\dots f(Y_n) \leqslant g(Y_1)\dots g(Y_n)\big) = P\Big(\sum_{i=1}^{n} \xi_i \leqslant 0\Big)$$

$$= P\left(\frac{\sum_{i=1}^{n}(\xi_i - E\xi_i)}{\sigma\sqrt{n}} \leqslant \sqrt{n}\,\frac{E\xi_1}{\sigma}\right) = O(1)\,P\left(\xi < -\frac{|E\xi_1|}{\sigma}\sqrt{n}\right)$$

$$= O(e^{-\beta n}), \qquad \beta > 0,$$

where ξ is a Gaussian $N(0, 1)$ random variable. Let us note that $E\xi_1 < 0$ by the well-known inequality: $\int f(y)\ln f(y)\,dy < \int f(y)\ln g(y)\,dy$. It is clear that for the σ-fields $\mathfrak{A}_n$ and $\mathfrak{B}_n$ we have

$$\bar{\varrho}(\mathfrak{A}_n, \mathfrak{B}_n) < P(D_n^c) \leqslant \exp\{-\beta_0 n\} \qquad \text{for some } \beta_0 > 0.$$

The interpretation of the above example is the following. We observe the process (Y_j) (of independent measurements) and then (following the likelihood ratio test) we choose between two hypotheses: the density equals f or g. According to our decision we put $Z_n = X_n$ or $Z_n = \overline{X}_n$.

8.3. Let $(X_n, \mathfrak{A}_n)$ be a martingale. For $D_n \in \mathfrak{A}_n$, we put

$$Z_n = 1_{D_n} X_n + 1_{D_n^c} \overline{X}_n, \tag{33}$$

where $(\overline{X}_n)$ is an arbitrary sequence independent of $(\mathfrak{A}_n)$ with $E\overline{X}_n = 0$. Then $(Z_n, \mathfrak{A}_n)$ is a p.m.

8.4. Other examples of pseudo-martingales being simple transformations of martingales can be obtained as "moving averages" as follows.

For a martingale $(X_n, \mathfrak{A}_n)$, let us put

$$Z_n = \sum_{k=0}^{N} a_{n-k} X_{n-k}, \qquad n > N,$$

a_k being real numbers satisfying the condition: $a_{n+1} = a_{n-N}$ for $n > N$. Then $(Z_n, \mathfrak{A}_{n-N})$, $n > N$, is a p.m.

Similarly, putting

$$V_n = \sum_{k=0}^{N} b_{n+k} X_{n+k}, \qquad n \geqslant 1,$$

with the coefficients satisfying $b_{n+N+1} = b_n$, $n \geqslant 1$, we obtain a p.m. $(V_n, \mathfrak{A}_n)$.

More generally, let $(a_{n,k})_{k=0,1,\dots,k_n, n\geqslant 1}$ be a matrix satisfying the condition

$$\sum_{k=0}^{k_{n+1}} a_{n+1,k} = \sum_{k=0}^{k_n} a_{n,k}, \qquad n = 1, 2, \dots$$

(for example: $a_{n,k} = 1/(n+1)$, $0 \leqslant k \leqslant n$, or $a_{n,k} = \binom{n}{k} 2^{-n}$, $0 \leqslant k \leqslant n$, and zero elsewhere).

Let us put

$$Z_n = \sum_{k=0}^{n} a_{n,k} X_{n+k}.$$

Then $(Z_n, \mathfrak{A}_n)$ is a p.m.

9. In this section we discuss a class of pseudo-martingales which seems to be quite important. It is closely related to Markov chains and r-independent sequences. We adopt the following definition.

9.1. DEFINITION. A sequence $(X_n) \subset L_1$ is said to be a *pseudo-martingale of type* (r) if the following condition holds:

$$E(X_n - X_{n-1} \mid X_1, \ldots, X_{n-r}) = 0 \quad \text{for } n > r. \tag{34}$$

Obviously, a p.m. of type (1) is simply a martingale.

To indicate a close relation of the notion just defined with r-independence, we introduce even a little weaker condition than r-independence. Namely, we say that a sequence (X_n) of random variables is *successively r-independent* if, for any n, X_{n+r} is independent of $(X_1 \ldots X_n)$. Let $(X_n) \subset L_1$. Assume that $EX_n = 0$ and (X_n) is successively r-independent. Then, clearly, the sequence $S_n = \sum_{k=1}^{n} X_k$ is a p.m. of type (r).

For a sequence (X_n) of random variables, let $\Delta_k = X_k - X_{k-1}$. For $1 \leqslant s \leqslant r$, we define the sequences $X^{(s)}$ by putting

$$X_n^{(s)} = \Delta_s + \Delta_{s+r} + \ldots + \Delta_{s+nr}, \quad n \geqslant 1. \tag{35}$$

9.2. DEFINITION. We shall say that (X_n) is *r-uniformly integrable* (*r-bounded in* L_p, respectively) if all sequences $X^{(s)}$, $1 \leqslant s \leqslant r$, are uniformly integrable (L_p-bounded, respectively).

9.3. THEOREM. *Let (X_n) be a pseudo-martingale of type (r). If (X_n) is r-uniformly integrable, then $X_n \to X_\infty$ a.s. and in L_1, where*

$$X_\infty = \lim_{n\to\infty} E(X_n \mid X_1, \ldots, X_{n-r}).$$

Proof. The last limit exists since $E(X_n \mid X_1, \ldots, X_{n-r})$ is a uniformly integrable martingale. To prove that $X_n \to X_\infty$ we put $\Delta_k = X_k - X_{k-1}$ and consider the sequences $X^{(s)}$, $1 \leqslant s \leqslant r$, defined by formula (35). If $m = nr + k$ with $1 \leqslant k < r$, then

$$X_m = \sum_{s=1}^{k} X_n^{(s)} + \sum_{s=k+1}^{r} X_{n-1}^{(s)}. \tag{36}$$

Moreover, for any $s = 1, \ldots, r$, the sequence

$$(X_n^{(s)}, \mathfrak{A}_{(n-1)r+s})_{n=2,3,\ldots} \tag{37}$$

(with $\mathfrak{A}_k = \sigma(X_1, \ldots, X_k)$) is a uniformly integrable martingale, so $X_n^{(s)}$ converges (as $n \to \infty$) with probability one and in L_1 to some $Y^{(s)} \in L_1$ $(s = 1, \ldots, r)$.

By Theorem 2.3, $X_n \to X_\infty$ weakly, which implies that $\sum_{s=1}^{r} Y^{(s)} = X_\infty$ and, consequently, $X_n \to X_\infty$ a.e. and in L_1. ■

We have the following strong law of large numbers.

9.4. THEOREM. *Let $(X_n) \subset L_2$ be a zero-mean pseudo-martingale of type (r). Let us assume that, putting $\Delta_k = X_k - X_{k-1}$, we have*

$$\sum_{k=1}^{\infty} \frac{E\Delta_k^2}{k^2} < \infty.$$

Then $X_n/n \to 0$ a.s. and in L_2.

Proof. Keeping the notation of Sections 9.1–9.3, we have (36). For $s = 1, \ldots, r$, the sequence $(X_n^{(s)}, n \geq 1)$ is a martingale. Setting $Z_{k,s} = \Delta_{s+kr}$, we have

$$\sum_{k=1}^{\infty} \frac{EZ_{k,s}^2}{k^2} < \infty, \qquad s = 1, \ldots, r,$$

so the sequence $Y_{n,s} = \sum_{k=1}^{n} Z_{k,s}/k$, $n \geq 1$, is an L_2-bounded martingale. Consequently, $Y_{n,s} \to Y_s$ a.s. and in L_2 for $s = 1, \ldots, r$. By Kronecker's lemma,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} Z_{k,s} = 0,$$

which together with (36) gives $n^{-1} X_n \to 0$ a.s. and in L_2. ■

9.5. Natural examples of pseudo-martingales of type (r) can be provided by taking some functions of Markov chains. More exactly, let $Y = (Y_n)$ be a homogeneous Markov chain with states labelled by positive integers. Let $P = (P_{ij})$ be the transition probability matrix of Y. Let f be a function defined on the states of Y and satisfying the equality

$$(P^r - P^{r-1}) f = 0, \qquad r > 1.$$

Then the sequence $(X_j) = (f(Y_j))$ is a p.m. of type (r). Indeed, putting $i_{n-r} = i$, we have

$$\begin{aligned}(38)\quad E(X_n - X_{n-1} \mid Y_1 = i_1, \ldots, Y_{n-r} = i_{n-r}) &= E(X_n - X_{n-1} \mid Y_{n-r} = i)\\ &= \sum_j f(j) P_{ij}^{(r)} - \sum_j f(j) P_{ij}^{(r-1)} = \sum_j (P_{ij}^{(r)} - P_{ij}^{(r-1)}) f(j)\\ &= \big((P^r - P^{r-1}) f\big)(i) = 0\end{aligned}$$

for all states i.

Taking on both sides of equality (38) the conditional expectation $E(\cdot \mid X_1, \ldots, X_{n-r})$, we get

$$E(X_n - X_{n-1} \mid X_1 \ldots X_{n-r}) = 0,$$

which means that (X_n) is a p.m. of type (r).

Let us pass to some examples connected with a random walk. We shall confine our attention to a symmetric random walk $Y=(Y_n)$ on the lattice $\boldsymbol{Z}$ of integers. This means that a homogeneous Markov chain (Y_n) is governed by the transition probability matrix $\boldsymbol{P}=(P_{ij})$ with $P_{i,i+1}=P_{i,i-1}=1/2$ for $i\in\boldsymbol{Z}$, and $P_{ij}=0$ elsewhere. We want to describe all trajectories of a p.m. of type (r) related to the Markov chain $Y=(Y_n)$. We consider a process $X_n=f(X_n)$, where f is a function defined on $\boldsymbol{Z}$ and satisfying the equality

$$(\boldsymbol{P}^{r+1}-\boldsymbol{P}^r)f=0. \tag{39}$$

Writing (39) in the form $(\boldsymbol{P}-I)\boldsymbol{P}^r f=0$, we get

$$\sum_j (\boldsymbol{P}^r)_{ij} f(j)=a+bi, \quad i\in\boldsymbol{Z}, \tag{40}$$

for some $a, b\in\boldsymbol{R}$.

Consequently,

$$\frac{1}{2^r}\sum_{k=0}^{r}\binom{r}{k} f(i+r-2k)=a+bi, \quad i\in\boldsymbol{Z}.$$

Let us put

$$x(l)=f(2l)-a-b\cdot 2l, \quad y(l)=f(2l+1)-a-b(2l+1).$$

Then we obtain, after standard calculations,

$$\sum_{k=0}^{r}\binom{r}{k} x(l-k)=0, \quad \sum_{k=0}^{r}\binom{r}{k} y(l-k)=0. \tag{41}$$

The inverse of the generating function for (41) is of the form

$$\sum_{k=0}^{r}\binom{r}{k} x^k=(x+1)^r=0,$$

so it has one root (-1) of multiplicity r. Thus (cf. [3]),

$$\begin{aligned} x(l)&=(-1)^l(a_0+a_1 l+\ldots+a_{r-1} l^{r-1}),\\ y(l)&=(-1)^l(b_0+b_1 l+\ldots+b_{r-1} l^{r-1}). \end{aligned} \tag{42}$$

Formula (42) gives a complete description of trajectories of the p.m. of type (r) related to a symmetric random walk on $\boldsymbol{Z}$. This makes it possible to describe the asymptotic behaviour of trajectories of (X_n). Clearly, an essential exercise is to rewrite the law of the iterated logarithm for a process $(-1)^l(Y_{2l})^s$ (and $(-1)^l(Y_{2l+1})^s$), for some fixed exponent $s<r$.

Namely, we have the following proposition.

9.6. Proposition. *For a symmetric random walk* (Y_n) *on* $\boldsymbol{Z}$, *a process* $Z_l=(-1)^l(Y_{2l})^s$, *with a fixed exponent* $s\in\boldsymbol{Z}^+$, *satisfies the condition*

$$\limsup_{l\to\infty}\frac{\varepsilon Z_l}{(\sqrt{(4l)\ln\ln 2l})^s}=1$$

with probability one, for $\varepsilon=1$ *and* $\varepsilon=-1$ *as well.*

Proof. Random walk must go from one level to another passing through all intermediate levels. Thus the proposition is a consequence of the law of the iterated logarithm. ■

9.7. Corollary. *Assume that a pseudo-martingale* (X_n) *is given by the following conditions*:

(*) $X_n = f(Y_n)$, Y_n *being the symmetric random walk on* $\mathbf{Z}$, *and*

(**) $E(X_{n+r+1} - X_{n+r} \mid X_1, \ldots, X_n) = 0$.

Then, for some constants b, $b_1 \in \mathbf{R}$ *and some integer* $1 \leqslant s \leqslant r-1$, *we have*

$$\limsup_{m\to\infty} \frac{\varepsilon(X_m - mb)}{b_1(\sqrt{m \ln\ln m})^s} = 1$$

with probability one, for $\varepsilon = 1$ *and* $\varepsilon = -1$.

REFERENCES

[1] T. Ando, *Contractive projections in L^p-spaces*, Pacific J. Math. 17 (1966), pp. 391–405.

[2] R. B. Ash, *Real Analysis and Probability*, Academic Press, New York–San Francisco–London 1972.

[3] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics. A Foundation for Computer Science*, Addison–Wesley, Massachusetts, 1994.

[4] E. Hensz-Chądzyńska, R. Jajte and A. Paszkiewicz, *Almost sure convergence of projections in L_p-spaces*, Ann. Univ. Mariae Curie-Skłodowska 61 (1997), pp. 67–84.

[5] E. H. Lacey, *The Isometric Theory of Classical Banach Spaces*, Die Grundlehren der Math. Wissenschaften 208, Springer-Verlag, Berlin–Heidelberg–New York 1974.

[6] M. M. Rao, *Stochastic Processes: General Theory*, Kluwer Acad. Publishers, Dordrecht–Boston–London 1995.

Department of Probability and Statistics, Łódź University
Banacha 22, 90-238 Łódź, Poland
E-mail address: rjajte@math.uni.lodz.pl, adampasz@math.uni.lodz.pl

Received on 30.6.1998