PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 22, Fasc. 1 (2002), pp. 1–11

WEAK CONVERGENCE OF RANDOM VECTORS AND DISTRIBUTIONS IN BANACH SPACES*

BY

RYSZARD JAJTE AND ADAM PASZKIEWICZ (ŁÓDŹ)

Abstract. Let (ξ_n) be a sequence of random vectors with values in a Banach space X with distributions p_{ξ_n} weakly converging to a given distribution p. We characterize a general form of a distribution of a weak limit of ξ_n in Banach space $L_1(X)$ of Bochner integrable vectors. We show that the weak convergence of random vectors (ξ_n) in $L_1(X)$ implies that $\|\xi_n(\omega)-\xi(\omega)\| \to 0$ stochastically. Moreover, the conditions $\|\xi_n(\omega)-\xi(\omega)\| \to 0$ stochastically and $\langle \xi_n(\omega)-\xi(\omega), x^* \rangle \to 0$ stochastically for any $x^* \in X^*$ are equivalent.

1991 AMS Subject Classification: 60B11, 46B09.

Key words and phrases: Distributions in Banach spaces, weak limits of random vectors, weak limits of distributions.

1. INTRODUCTION

The main goal of the paper is to investigate some general relations between weak convergence of Banach space valued random variables and their distributions. The advanced theory of probability in a Banach space often depends on its geometry. We formulate rather elementary results not depending on additional geometric conditions.

For a Banach space $(X, \|\cdot\|)$ the notion of weak convergence of X-valued random vectors has several aspects. For example, the following situations seem to be quite meaningful. For a sequence (ξ_n) of random vectors in X it may happen that

(i) for any $x^* \in X^*$, $\langle \xi_n(\omega)-\xi(\omega), x^* \rangle \to 0$ for almost all ω;

(ii) for any $x^* \in X^*$, $\langle \xi_n-\xi, x^* \rangle \to 0$ stochastically;

(iii) assuming that ξ, ξ_n are Bochner integrable, for any $x^* \in X^*$ and real-valued bounded random variable f, we have the convergence of expectations

$$E\langle \xi_n-\xi, x^* \rangle f \to 0;$$

* Research supported by KBN grant 2 P03A 023 15.

(iv) assuming that ξ, $\xi_n \in L_1(X)$, where $L_1(X)$ is the Banach space of Bochner integrable vectors, ξ_n tends to ξ in the weak topology in $L_1(X)$.

Interesting relations appear when we assume that distributions p_{ξ_n} (of the vectors ξ_n) converge weakly to a distribution p in X. Under this assumption we characterize all possible distributions of weak limits of ξ_n (Theorem 4.1). We show that the weak stochastic convergence (condition (ii)) implies the norm stochastic convergence (Theorem 3.3). Moreover, the weak convergence in $L_1(X)$ (condition (iii) or (iv)) implies the stochastic norm convergence (Theorem 3.1).

2. NOTATION

Throughout, X is a real separable Banach space (with the norm $\|\cdot\|$) and X^* is its dual. For a probability space, say $(\Omega, \mathscr{F}, P)$, we use the following notation. $L_0(\Omega, \mathscr{F}, P; X)$ denotes the space of all X-valued random vectors ξ on $(\Omega, \mathscr{F}, P)$, i.e. the maps $\xi\colon \Omega \to X$ such that $\xi^{-1}A \in \mathscr{F}$ for $A \in \operatorname{Borel} X$ (X is always equipped with the norm topology). p_ξ denotes the probability distribution of ξ, i.e. $p_\xi(A) = P(\xi^{-1}A)$ for $A \in \operatorname{Borel} X$. $L_1(\Omega, \mathscr{F}, P; X)$ stands for the space of those vectors in $L_0(\Omega, \mathscr{F}, P; X)$ which are P-integrable in the sense of Bochner. We shall also write shortly $L_1(X)$ or $L_1(\Omega, X)$. The notation $L_\infty(X^*)$ or $L_\infty(\Omega, X^*)$ or $L_\infty(\Omega, \mathscr{F}, P; X^*)$ will be used for the completion of the space of linear combinations

$$\{g(\omega) = \sum_k x_k^* f_k(\omega);\ x_k^* \in X^*, f_k \in L_\infty(\Omega, \mathscr{F}, P)\}$$

under the norm $\|g\|_\infty = \operatorname{sup\,ess}\{\|g(\omega)\|_{X^*};\ \omega \in \Omega\}$. We do not refer here to tensor products of spaces keeping all considerations in a rather elementary language.

For the norm in $L_1(\Omega, X)$ we set $\|\xi\|_1 = E\|\xi\| = \int_\Omega \|\xi(\omega)\|\, dP$.

The spaces $L_1(X)$ and $L_\infty(X^*)$ are in natural duality given by the formula

$$(\xi, \eta) = \int_\Omega \langle \xi(\omega), \eta(\omega)\rangle P(d\omega) \quad \text{for } \xi \in L_1,\ \eta \in L_\infty.$$

For ξ, $\xi_n \in L_1(\Omega, X)$, we say that ξ_n *tends weakly* to ξ if $(\xi_n, \eta) \to (\xi, \eta)$ for any $\eta \in L_\infty(\Omega, X^*)$. In other words, $\xi_n \to \xi$ in the $\sigma(L_1(X), L_\infty(X^*))$-topology.

For $\xi \in L_1(\Omega, \mathscr{F}, P; X)$ and any σ-field $\mathscr{G} \subset \mathscr{F}$, there exists a conditional expectation $E(\xi|\mathscr{G})$ uniquely defined, like in the classical case, by the conditions

$$E(\xi|\mathscr{G}) \in L_1(\Omega, \mathscr{G}, P; X),$$

$$\int_A \xi\, dP = \int_A E(\xi|\mathscr{G})\, dP \quad \text{for } A \in \mathscr{G}.$$

The conditional expectation is a bounded linear operator

$$E(\cdot|\mathscr{G}): L_1(\Omega, \mathscr{F}, P; X) \to L_1(\Omega, \mathscr{G}, P; X)$$

enjoing nice properties analogous to those which are well known in the classical case.

3. CONVERGENCE THEOREMS

THEOREM 3.1. *Let $\xi_n, \xi \in L_1(\Omega, \mathscr{F}, P; X)$. Let us assume that $p_{\xi_n} \Rightarrow p_\xi$ (weakly) and that $\xi_n \to \xi$ in the $\sigma(L_1(X), L_\infty(X^*))$-topology. Then $\|\xi_n(\omega) - \xi(\omega)\| \to 0$ stochastically.*

In the proof we shall use the following result due to Pratelli.

THEOREM 3.2 ([5], see also [2]). *Let ξ_n, ξ be real-valued integrable random variables. If $p_{\xi_n} \Rightarrow p_\xi$ (weakly) and $\xi_n \to \xi$ weakly in $L_1(R)$, then $\|\xi_n - \xi\|_1 \to 0$.*

Proof of Theorem 3.1. For simplicity, we assume that $P_\xi(\{\theta\}) = 0$. We leave to the reader an easy modification of the proof when θ is the atom of p_ξ. From the assumptions it follows in particular that, for any $x^* \in X^*$, we have $p_{\langle \xi_n, x^* \rangle} \Rightarrow p_{\langle \xi, x^* \rangle}$ and $\langle \xi_n, x^* \rangle \to \langle \xi, x^* \rangle$ weakly in $L_1(R)$. By Theorem 3.2 we have

$$\int_\Omega |\langle \xi_n, x^* \rangle - \langle \xi, x^* \rangle| \, dP \to 0 \quad \text{for any } x^* \in X^*.$$

Then also

$$\int_\Omega |\langle \xi_n 1_B, x^* \rangle - \langle \xi 1_B, x^* \rangle| \, dP \to 0 \quad \text{for } B \in \mathscr{F}$$

and, consequently, for any $x^* \in X^*$, we have

$$\Big|\int_\Omega \exp(i\langle \xi_n 1_B, x^* \rangle) \, dP - \int_\Omega \exp(i\langle \xi 1_B, x^* \rangle) \, dP\Big| \leqslant \int_\Omega |\langle \xi_n 1_B, x^* \rangle - \langle \xi 1_B, x^* \rangle| \, dP \to 0.$$

Thus the characteristic functionals of random vectors $\xi_n 1_B$ converge to the characteristic functional of $\xi 1_B$, which, together with the tightness of measures $p_{\xi_n 1_B}$, $n = 1, 2, \ldots,$ gives the weak convergence of measures $p_{\xi_n 1_B} \Rightarrow p_{\xi 1_B}$ for any $B \in \mathscr{F}$. Since $p_{\xi_n} \Rightarrow p_\xi$ and $p_\xi(\{\theta\}) = 0$, for any $\varepsilon > 0$ there exists a compact set $Z_\varepsilon \subset X$ such that $\theta \notin Z_\varepsilon$, $p(\xi_n \in Z_\varepsilon) > 1 - \varepsilon$ for n large enough, and $p(\xi \in Z_\varepsilon) > 1 - \varepsilon$. By a rather standard procedure we can find, for any $\delta > 0$, a partition $(A_1, \ldots, A_N)$ of Z_ε such that A_j are continuity sets of p_ξ (i.e. $p_\xi(\partial A_j) = 0$) and the points $z_1, \ldots, z_N \in X$ satisfying the conditions $\sup_{x \in A_i} \|x - z_i\| < \delta$ $(i = 1, \ldots, N)$.

Let us put

$$\bar{\xi}_n = \sum_{j=1}^{N} z_j 1_{\xi_n^{-1}(A_j)}, \quad \bar{\xi} = \sum_{j=1}^{N} z_j 1_{\xi^{-1}(A_j)}.$$

Let $\varepsilon > 0$ be fixed, and put $\delta = \varepsilon/4$. We shall show that

$$P(\|\xi_n - \xi\| > \varepsilon) < 3\varepsilon \quad \text{for } n \text{ large enough.}$$

Indeed, we have

$$P(\|\xi_n - \bar{\xi}_n\| > \delta) < \varepsilon \quad \text{and} \quad P(\|\xi - \bar{\xi}\| > \delta) < \varepsilon$$

(since $P(\xi_n \notin Z_\varepsilon) < \varepsilon$ and $P(\xi \notin Z_\varepsilon) < \varepsilon$). Let us note that $\theta \notin A$ implies

$$(\xi_n 1_B)^{-1}(A) = \{\omega\colon \xi_n(\omega) 1_B(\omega) \in A\} = (\xi_n^{-1} A) \cap B.$$

Moreover, if $\theta \notin \bar{A}$ and A is a continuity set of p_ξ, then also A is a continuity set of $p_{\xi 1_B}$ for any $B \in \mathcal{F}$. Thus, by the choice of the sets A_j, we have

$$p_{\xi_n 1_B}(A_j) \to p_{\xi 1_B}(A_j),$$

which means that

$$\int_\Omega 1_{(\xi_n \in A_j)} 1_B \, dP \to \int_\Omega 1_{(\xi \in A_j)} 1_B \, dP \quad \text{for any } B \in \mathcal{F}.$$

But this implies that

$$1_{(\xi_n \in A_j)} \to 1_{(\xi \in A_j)} \text{ weakly in } L_1(\boldsymbol{R}),$$

which, together with $P(\xi_n \in A_j) \to P(\xi \in A_j)$, gives

$$\int_\Omega |1_{\xi_n^{-1} A_j} - 1_{\xi^{-1} A_j}| \, dP \to 0 \quad \text{for } j = 1, 2, \ldots, N,$$

by Theorem 3.2. Thus, we have the estimation

$$P(\|\bar{\xi}_n - \bar{\xi}\| > \varepsilon/3) \leqslant \frac{3}{\varepsilon} E\|\bar{\xi}_n - \bar{\xi}\| = \frac{3}{\varepsilon} \int_\Omega \Big\| \sum_{j=1}^{N} z_j (1_{\xi_n^{-1} A_j} - 1_{\xi^{-1} A_j}) \Big\| \, dP$$

$$\leqslant \frac{3}{\varepsilon} \sum_{j=1}^{N} \|z_j\| \int_\Omega |1_{\xi_n^{-1} A_j} - 1_{\xi^{-1} A_j}| \, dP < \varepsilon \quad \text{for } n > n_0.$$

Since, clearly,

$$P(\|\xi_n - \bar{\xi}_n\| > \varepsilon/3) < \varepsilon \quad \text{and} \quad P(\|\xi - \bar{\xi}\| > \varepsilon/3) < \varepsilon,$$

we get

$$P(\|\xi_n - \xi\| > \varepsilon) < 3\varepsilon \quad \text{for } n > n_0. \ \blacksquare$$

The other conditions implying the stochastic convergence in norm are formulated in the following theorem:

THEOREM 3.3. *Let X be a Banach space and let $\xi, \xi_n \in L_0(\Omega, \mathcal{F}, P; X)$ be such that the distributions (p_{ξ_n}) are tight. Then the weak stochastic convergence of ξ_n to ξ implies the stochastic convergence in norm, i.e.*

$$\langle \xi_n - \xi, x^* \rangle \to 0 \text{ stochastically} \quad \text{for each } x^* \in X^*,$$

implies $\|\xi_n - \xi\| \to 0$ stochastically.

Proof. The algebraic sum of two compact sets in X is compact, so the sequence $(p_{\xi_n - \xi})$ is tight. Moreover, the characteristic functionals of $\xi_n - \xi$ tend to 1. This completes the proof. ∎

4. CHARACTERIZATION OF DISTRIBUTIONS OF WEAK LIMITS

In this section we show that some classes of probability distributions on a Banach space X coincide. For the convenience and clarity of formulations the random X-valued vectors appearing in different conditions will be defined on possibly different probability spaces.

Let p be a given probability distribution on X. We show the coincidence of the following four classes each of which is naturally related to the distribution p. The first one consists of all possible distributions of weak limits of random vectors ξ_n with $p_{\xi_n} \Rightarrow p$. The second one consists of all possible distributions of conditional expectations of random vectors with distribution p.

The third one is the class of all possible distributions of almost sure limits of arithmetic means $n^{-1}(\xi_1 + \dots + \xi_n)$ for some ξ_n with $p_{\xi_n} \equiv p$.

The fourth class consists of all possible distributions of almost sure limits of the ergodic averages $n^{-1}(\xi + \xi \circ T + \dots + \xi \circ T^{n-1})$ for some ξ with $p_\xi \equiv p$, $T: \Omega \to \Omega$ being an arbitrary measure-preserving bijection of Ω.

Formally, we have the following theorem:

THEOREM 4.1. *Let p be a probability distribution on a Banach space X with $\int_X \|x\| \, p(dx) < \infty$. Then the following classes of distributions on X coincide:*

$$\mathcal{W}_p = \{q;\ p_{\xi_n} \Rightarrow p,\ p_\xi = q$$
$$\text{for some } \xi_n, \xi \text{ with } \xi_n \to \xi \text{ in the } \sigma(L_1(X), L_\infty(X^*))\text{-topology}\};$$

$$\mathcal{C}_p = \{q;\ p = p_\xi,\ q = p_{E(\xi|\eta)} \text{ for some random vectors } \xi, \eta\};$$

$$\mathcal{M}_p = \{q;\ p = p_{\xi_n},\ q = p_\xi,\ n^{-1}(\xi_1 + \dots + \xi_n) \to \xi \text{ a.s.}$$
$$\text{for some random vectors } \xi_n, \xi\};$$

$$\mathcal{E}_p = \{q;\ p = p_\xi,\ q = p_\eta,\ n^{-1}(\xi + \xi \circ T + \dots + \xi \circ T^{n-1}) \to \eta \text{ a.s.}$$
$$\text{for some } \xi \in L_1(\Omega, \mathcal{F}, P; X) \text{ and measure-preserving bijection } T: \Omega \to \Omega\}.$$

Proof. It is enough to show the following inclusions: $\mathscr{C}_p \subset \mathscr{W}_p$, $\mathscr{W}_p \subset \mathscr{C}_p$, $\mathscr{C}_p \subset \mathscr{E}_p$, $\mathscr{M}_p \subset \mathscr{W}_p$. We do it in separate steps.

Step 1. $\mathscr{C}_p \subset \mathscr{W}_p$.

Let $X \in L_1(\Omega, \mathscr{F}, P; X)$ and let $Y \in L_0(\Omega, \mathscr{F}, P; X)$. It is enough to prove that there exist a probability space $(M, \mathscr{M}, \mu)$ and random vectors $\xi_n, \xi \in L_1(M, \mathscr{M}, \mu; X)$ such that $\xi_n \to \xi$ in the $\sigma(L_1(M, X), L_\infty(M, X^*))$-topology, with $p_{\xi_n} = p_X$ $(n = 1, 2, \ldots)$ and $p_\xi = p_{E(X|Y)}$.

Let us put $M = [0, 1] \times X$, $\mathscr{M} = \operatorname{Borel} M$, $\mu = \lambda \times p_Y$, where λ is the Lebesgue measure on $[0, 1]$. For $y \in X$, $B \in \operatorname{Borel} X$, let $q(y, B)$ be a regular conditional probability distribution for X, given $Y = y$, i.e. for any $y \in X$, $q(y, \cdot)$ is a probability measure on $\operatorname{Borel} X$, and for any $B \in \operatorname{Borel} X$, $q(y, B) = P(X \in B \mid Y = y)$ p_Y-a.e. In particular, we have $E(X \mid Y = y) = \int_X z q(y, dz)$. We shall also write $P_y(\cdot)$ instead of $P(\cdot \mid Y = y)$. Let a map j from $[0, 1]$ onto X be a fixed Borel isomorphism (see [3], p. 227). By the same letter j we denote the isomorphism of the Borel field of $[0, 1]$ onto the Borel field of X. Then, for every $y \in X$, $0 \leqslant \alpha \leqslant 1$, $F(y, \alpha) = P_y(j[0, \alpha))$ is a distribution function of a measure concentrated on the interval $[0, 1]$. For any $y \in X$, we denote by $g(x, y)$ the rearrangement ("inverse") of the distribution function $F(y, \alpha)$, i.e. for $0 \leqslant x \leqslant 1$ we set

$$g(x, y) = \inf\{\alpha\colon\ F(y, \alpha) \geqslant x\}, \qquad y \in X.$$

Let us put

$$h(x, y) = j \circ g(x, y) \quad \text{and} \quad \xi_n(x, y) = h(\tau^n x, y),$$

where τ is a fixed measure-preserving mixing transformation of $([0, 1], \lambda)$. Then we have

$$p_{\xi_n}(A) = p_X(A) \quad \text{for } A \in \operatorname{Borel} X.$$

Indeed, since j is the isomorphism of Borel structures, it is enough to show the above formula for $A = j([0, \alpha))$, $0 \leqslant \alpha \leqslant 1$. In this case we have

$$\begin{aligned} p_{\xi_n}(j[0, \alpha)) &= \mu\{(x, y)\colon\ \xi_n(x, y) \in j([0, \alpha))\} \\ &= \mu\{(x, y)\colon\ g(\tau^n x, y) < \alpha\} = \mu\{(x, y)\colon\ g(x, y) < \alpha\} \\ &= \int_X P_y(j[0, \alpha)) p_Y(dy) = p_X(j[0, \alpha)). \end{aligned}$$

Let us put

$$\xi(x, y) = E(X \mid Y = y)$$

(so ξ does not depend on x). It is easy to check that

$$p_\xi = p_{E(X|Y)}.$$

Indeed, let us define $\varphi\colon X \to X$ by putting $\varphi(y) = E(X \mid Y = y)$. Then we have

$E(X \mid Y) = \varphi(Y)$. Consequently, for $A \in \mathrm{Borel}\, X$,

$$p_{E(X|Y)}(A) = P\{\omega \in \Omega: \ \varphi(Y(\omega)) \in A\} = p_Y\{y \in X: \ \varphi(y) \in A\}$$
$$= \mu\{((x, y): \ E(X \mid Y = y) \in A) = p_\xi(A)\}.$$

Let us remark that

$$\int_0^1 h(x, y)\, dx = E(X \mid Y = y) \ \ p_Y\text{-a.e.} \tag{1}$$

Indeed, for P_Y-almost every y, we have $q(\cdot, y) = P(X \in \cdot \mid Y = y)$. Fixing such a y we can look at $g(x, y)$ as a random variable with distribution function

$$\lambda(x \in [0, 1]: g(x, y) < \alpha) = F_y(\alpha) = P_y(X \in j([0, \alpha))) = P_y(j^{-1} \circ X < \alpha),$$

so the random variables g over $([0, 1], \lambda)$ and $j^{-1} \circ X$ (over (Ω, P_y)) have the same distribution function. Consequently, the random variables $j \circ g$ and X have the same expectation, which means that the formula (1) holds. Let us remark that, by the mixing property of τ, for $\kappa \in L_1([0, 1], \lambda; X)$ and $\varphi \in L_\infty([0, 1], \lambda; X^*)$, we have

$$\int_0^1 \langle \kappa(\tau^n x), \varphi(x) \rangle\, dx \to \left\langle \int_0^1 \kappa(x)\, dx, \int_0^1 \varphi(x)\, dx \right\rangle.$$

Indeed, it is enough to check it for $\varphi = f \cdot x^*$ with $x^* \in X^*$ and $f \in L_\infty([0, 1], \boldsymbol{R})$, and then to pass to the limit with suitable linear combinations.

To complete the proof, let us take a $\psi \in L_\infty(M, \mathscr{M}, \mu; X^*)$. Then we have

$$\int_M \langle \xi_n(x, y), \psi(x, y) \rangle\, d\mu = \int_M \langle h(\tau^n x, y), \psi(x, y) \rangle\, d\mu$$
$$= \int_X \left(\int_0^1 \langle h(\tau^n x, y), \psi(x, y) \rangle\, dx \right) p_Y(dy) \to \int_X \left\langle \int_0^1 h(x, y)\, dx, \int_0^1 \psi(x, y)\, dx \right\rangle p_Y(dy)$$
$$= \int_X \int_0^1 \langle \xi(x, y), \psi(x, y) \rangle\, dx\, p_Y(dy) = \int_M \langle \xi(x, y), \psi(x, y) \rangle\, d\mu.$$

Step 2. $\mathscr{W}_p \subset \mathscr{C}_p$.

Let $\xi, \xi_n \in L_1(M, \mathscr{M}, \mu; X)$ and let

$$\xi_n \to \xi \ \text{ in the } \sigma(L_1(M, \mu; X), L_\infty(M, \mu; X^*))\text{-topology.} \tag{2}$$

Assume that $p_{\xi_n} \Rightarrow p$ (weakly). It is enough to prove that there exist a probability space $(\Omega, \mathscr{F}, P)$ and random vectors $X, Y \in L_0(\Omega, \mathscr{F}, P; X)$ such that $p = p_X$ and $p_\xi = p_{E(X|Y)}$.

Let us note first that $\int_X \|x\|\, p(dx) < \infty$. Indeed, by the weak convergence of ξ_n to ξ, we have $\sup_n \|\xi_n\|_{L_1(M,X)} = K < \infty$. Thus, for any $c > 0$, by the weak convergence of p_{ξ_n} to p we obtain

$$\int_{\|x\| \leqslant c} \|x\| \, dp(x) = \lim_{n \to \infty} \int_{\|x\| \leqslant c} \|x\| \, dp_{\xi_n}(x) \leqslant K,$$

so

$$\int_{X} \|x\| \, dp(x) \leqslant K < \infty.$$

Put $(\Omega, \mathscr{F}) = \big(X \times X, \operatorname{Borel}(X \times X)\big)$. Let P_n be a probability distribution of the vector (ξ_n, ξ), $n = 1, 2, \ldots$ Taking a subsequence if necessary, we can assume that $P_n \Rightarrow P$ for some probability measure P, so a probability space $(\Omega, \mathscr{F}, P)$ is defined. For $X(x, y) = x$ and $Y(x, y) = y$, we have obviously $p = p_X$. In particular, $E\|X\| < \infty$. The equality $p_\xi = p_{E(X|Y)}$ is a consequence of the following lemma.

LEMMA. *Let $\xi_n, \xi \in L_1(M, \mathscr{M}, \mu; X)$ satisfy* (2). *Assume that $p_n = p_{(\xi_n, \xi)} \Rightarrow p$ (weakly). Then, for the coordinates $X(x, y) = x$, $Y(x, y) = y$ on $(X \times X$, $\operatorname{Borel} X \times X$, $p)$, the equality $E(X \mid Y) = Y$ holds.*

Proof. By assumption, $p_n \Rightarrow p$ (weakly) and for any $x^* \in X^*$ the sequence $\big(\langle \xi_n(\omega), x^* \rangle\big)$ weakly converges in $L_1(R)$. This implies that for any $A \in \operatorname{Borel} X$ we have

$$(3)\quad \Big\langle \int_{(Y \in A)} X dp_n, x^* \Big\rangle = \int_{(Y \in A)} \langle X, x^* \rangle \, dp_n \to \int_{(Y \in A)} \langle X, x^* \rangle \, dp = \Big\langle \int_{(Y \in A)} X dp, x^* \Big\rangle.$$

Indeed, let us fix an $x^* \in X^*$. Then, for $c > 0$, we get

$$\int_{(Y \in A)} \langle X, x^* \rangle \, 1_{(|\langle X, x^* \rangle| \leqslant c)} \, dp_n \to \int_{(Y \in A)} \langle X, x^* \rangle \, 1_{(|\langle X, x^* \rangle| \leqslant c)} \, dp.$$

Moreover, by the uniform integrability of $\langle \xi_n, x^* \rangle$, we obtain

$$\int_{(Y \in A)} \langle X, x^* \rangle \, 1_{(|\langle X, x^* \rangle| > c)} \, dp_n = \int_{(\xi \in A)} \langle \xi_n, x^* \rangle \, 1_{(|\langle \xi_n, x^* \rangle| > c)} \, d\mu$$

$$\leqslant \int_{(|\langle \xi_n, x^* \rangle| > c)} |\langle \xi_n, x^* \rangle| \, d\mu \to 0 \quad \text{as } c \to \infty, \text{ uniformly in } n.$$

Also,

$$\int_{(Y \in A)} |\langle X, x^* \rangle| \, 1_{(|\langle X, x^* \rangle| > c)} \, dp \to 0 \quad \text{as } c \to \infty.$$

All the above estimations easily imply (3). On the other hand, for any $x^* \in X^*$,

$$\Big\langle \int_{(Y \in A)} X dp_n, x^* \Big\rangle = \int_{(\xi \in A)} \langle \xi_n, x^* \rangle \, d\mu = \int_{M} \langle \xi_n, x^* \rangle \, 1_{(\xi \in A)} \, d\mu \to \Big\langle \int_{(Y \in A)} Y dp, x^* \Big\rangle,$$

which together with (3) gives

$$\int_{(Y\in A)} X dp = \int_{(Y\in A)} Y dp \qquad \text{for } A \in \operatorname{Borel} X.$$

It means that $E(X \mid Y) = Y$. ■

The proof of Step 2 is completed.

Step 3. $\mathscr{C}_p \subset \mathscr{E}_p$.

Let $p = p_\xi$ and $q = p_{E(\xi|\eta)}$ for ξ, η defined on some probability space $(\Omega, \mathscr{F}, P)$. Obviously, one can assume that $E(\xi \mid \eta) = \eta$. Let us take a new probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ with $\tilde{\Omega} = \Omega \times X^{\mathbf{Z}}$,

$$\tilde{P}(B \times \cdots \times X \times A_{-m} \times \cdots \times A_m \times X \times \cdots) = \int_B \prod_{-m \leqslant i \leqslant m} P(\xi \in A_i \mid \eta) P(d\omega)$$

for any $B \in \mathscr{F}$, $A_i \in \operatorname{Borel} X$. One can prove by a rather standard argument that the probability measure $\tilde{P}$ is uniquely defined on the product σ-field $\tilde{\mathscr{F}} = \mathscr{F} \otimes \mathscr{G}$, $\mathscr{G}$ being the σ-field generated by the cylinders in $X^{\mathbf{Z}}$. The demanded random variables can be defined as $\tilde{\eta}(\omega, (x_i)) = \eta(\omega)$ and $\tilde{\xi}(\omega, (x_i)) = x_0$ for any sequence $(x_i)_{i\in\mathbf{Z}} \in X^{\mathbf{Z}}$. The bijection T can be defined as a shift $(\omega, (a_i)_{i\in\mathbf{Z}}) \to (\omega, (a_{i+1})_{i\in\mathbf{Z}})$. By the strong law of large numbers for random vectors in Banach space (used conditionally), we have

$$n^{-1}(\tilde{\xi} + \tilde{\xi} \circ T + \cdots + \tilde{\xi} \circ T^{n-1}) \to \tilde{\eta} \text{ a.s.}$$

The relation $\mathscr{C}_p \subset \mathscr{E}_p$ is shown.

Step 4. $\mathscr{M}_p \subset \mathscr{W}_p$.

Now, let $p_{\xi_n} = p$, $p_\eta = q$, $n^{-1}(\xi_1 + \cdots + \xi_n) \to \eta$ a.s. for some random variables ξ, ξ_n on a probability space $(\Omega, \mathscr{F}, P)$. Let us define $\tilde{\Omega} = \Omega \times [0, 1)$, $\tilde{\mathscr{F}} = \mathscr{F} \otimes \operatorname{Borel}[0, 1)$, $\tilde{P} = P \otimes \lambda$, and $\tilde{\xi}_n(\omega, t) = \tilde{\eta}_n(\omega, 2^n t - [2^n t])$ with

$$\tilde{\eta}_n(\omega, t) = \begin{cases} \xi_1(\omega) & \text{for } t \in [0, 1/n), \\ \dots\dots\dots\dots\dots\dots & \\ \xi_n(\omega) & \text{for } t \in [(n-1)/n, 1). \end{cases}$$

By a rather standard argument, the convergence $\tilde{\xi}_n \to \tilde{\xi}$ in the $\sigma(L_1(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}; X), L_\infty(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P}; X^*))$-topology is equivalent to the convergence

$$\int (\tilde{\xi}_n - \tilde{\xi}) \mathbf{1}_{A \times B} \, d\tilde{P} \to 0 \tag{4}$$

for any sets $A \in \mathscr{F}$, $B \in \operatorname{Borel}[0, 1)$.

For $\tilde{\xi}(\omega, t) = \eta(\omega)$, $(\omega, t) \in \tilde{\Omega}$, the convergence (4) can be obtained by an approximation of a Borel set B by a sum of dyadic intervals. The relation $\mathscr{M}_p \subset \mathscr{W}_p$ is shown.

5. FINAL REMARKS

Remark 5.1. Our Theorem 3.1 is closely related to the results of L. Pratelli. One of them is quoted as Theorem 3.2. In fact, it is a special case of Theorem 3.1 for $X = \boldsymbol{R}^1$. The second result of L. Pratelli [5] is the following.

Let $\xi_n, \xi \in L_1(\Omega, \mathscr{F}, P; X)$ satisfy the condition:
(α) for any f — real, bounded and Borel on X, and any $x^* \in X^*$,

$$\langle \int_\Omega (f \circ \xi)(\xi_n - \xi)\, dP, x^* \rangle \to 0.$$

Then the convergence $\xi_n \to \xi$ in L_1 is equivalent to the weak convergence of distributions $p_{\xi_n} \Rightarrow p_\xi$ and the uniform integrability of $(\|\xi_n\|)$.

In Theorem 3.1 we show that the convergence of distributions $p_{\xi_n} \Rightarrow p_\xi$ implies the stochastic convergence of $\|\xi_n - \xi\|$ to 0 under the assumption of the convergence of ξ_n to ξ in the $\sigma(L_1(X), L_\infty(X^*))$-topology (condition (iii)).

Pratelli's condition (α) is even less restrictive than the $\sigma(L_1(X), L_\infty(X^*))$-convergence but it is used together with the uniform integrability of $\|\xi_n\|$ to show the $L_1(X)$-convergence.

Remark 5.2. The characterization of distributions of weak limits given in Theorem 4.1 in the case of $X = \boldsymbol{R}$ was obtained in [2].

Remark 5.3. Concerning the conditions formulated in the Introduction it is worth noting that for a separable space X the equivalence of (iii) and (iv) can be shown by a rather standard argument. Clearly, the conditions (i) or (ii) for a sequence (ξ_n) with uniformly integrable norms $\|\xi_n\|$ imply the conditions (iii) and (iv).

Remark 5.4. The class of measures described in Theorem 4.1 coincides with the class of measures q subordinated to the distribution p (cf. Theorem T.53 in [4], Chapter XI). Namely, $q \prec p$ if and only if $\int f dq \geqslant \int f dp$ for any concave positive function f.

REFERENCES

[1] P. Billingsley, *Convergence of Probability Measures*, Willey, New York–Toronto 1968.
[2] R. Jajte and A. Paszkiewicz, *Conditioning and weak convergence*, Probab. Math. Statist. 19 (1999), pp. 453–461.
[3] K. Kuratowski, *Topologie. I*, Monografie Matematyczne, Vol. 3, Warsaw 1933.
[4] P. A. Meyer, *Probability and Potentials*, Blaisdell Publish. Comp., Toronto–London 1966.

[5] L. Pratelli, *Un caractérisation de la convergence dans L^1. Application aux quasimartingales*, Sémin. Probab. XXVI, Lecture Notes in Math. 1526, Springer, 1992, pp. 61–69.

[6] L. Pratelli, *Sur la convergence en moyenne pour des vecteurs aléatoires intégrables au sens de Bochner*, Rend. Mat. Accad. Lincei 3 (1992), pp. 299–306.

R. Jajte, A. Paszkiewicz
Faculty of Mathematics
University of Łódź
Banacha 22
PL-90-238 Łódź, Poland
E-mail: rjajte@math.uni.lodz.pl
adampasz@math.uni.lodz.pl

A. Paszkiewicz
Institute of Mathematics
Polish Academy of Sciences
Śniadeckich 8
PL-00-950 Warszawa, P.O. Box 17, Poland

Received on 10.8.2000