PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 13, Fasc. 1 (1992), pp. 59–70

MATHEMATICAL EXPECTATION AND STRONG LAW OF LARGE NUMBERS FOR RANDOM VARIABLES WITH VALUES IN A METRIC SPACE OF NEGATIVE CURVATURE

BY

WOJCIECH HERER (WARSZAWA)

Abstract. Let f be a random variable with values in a metric space (X, d). For some class of metric spaces we define in terms of the metric d mathematical expectation of f as a closed bounded and non-empty subset of X. We then prove Kolmogorov's version of Strong Law of Large Numbers corresponding to that mathematical expectation.

0. Introduction. In this paper we introduce a concept of mathematical expectation of a random variable with values in a Polish metric space of negative curvature. This class of metric spaces contains complete simply connected Riemannian manifolds of non-positive (sectional) curvature with the geodesic metric. In [3, Chaper 5] Bussemann has studied a similar metric generalization of a non-positively curved Riemannian manifold (see also [6], Proposition 8.17 and Remark 8.18).

In Section 1 we introduce convex combinations of elements of a metric space. In Section 2 we define mathematical expectation of a random variable with values in a Polish space (X, d) of negative curvature. In Section 3 we prove Strong Law of Large Numbers for independent and identically distributed X-valued random variables together with its converse.

Results of this paper were partially announced in [8].

1. Convex combination. Let (X, d) be a metric space. By $(\boldsymbol{F}(X), h)$ we denote a metric space of closed bounded and non-empty subsets of X equipped with the Hausdorff metric defined as

$$h(F, F') = \max\{\sup_{x\in F} d(x, F'), \sup_{x'\in F'} d(x', F)\}.$$

We note the following identity:

$$h(\{x\}, F) = \sup_{y\in F} d(x, y) \quad \text{for } x\in X,\ F\in \boldsymbol{F}(X),$$

which will be used throughout this paper without reference.

DEFINITION 1.1. Let (X, d) be a metric space. For any system of non-negative reals $\{p_1, \ldots, p_n\}$ with $\sum_{i=1}^{n} p_i = 1$ and any subset $\{a_1, \ldots, a_n\}$ of X we define inductively a subset of X as follows:

1. If $n = 1$, we define: $1a_1 = \{a_1\}$.

2. Let $n > 1$, $\{a_1, \ldots, a_n\} \subset X$ and $\{p_1, \ldots, p_n\} \subset [0, 1]$ with $\sum_{i=1}^{n} p_i = 1$. Suppose the sets $\sum_{i=1}^{k} q_i b_i$ are already defined for all $k < n$ and any subsets

$$\{b_1, \ldots, b_k\} \subset X \quad \text{and} \quad \{q_1, \ldots, q_k\} \subset [0, 1] \text{ with } \sum_{i=1}^{k} q_i = 1.$$

We then define: $a \in \sum_{i=1}^{n} p_i a_i$ iff there exist non-empty disjoint and complementary subsets I_1 and I_2 of $\{1, \ldots, n\}$ and elements $a^1 \in \sum_{i \in I_1} p_i^1 a_i$, $a^2 \in \sum_{i \in I_2} p_i^2 a_i$, where $p_i^1 = p_i / \sum_{i \in I_1} p_i$ for $i \in I_1$ and $p_i^2 = p_i / \sum_{i \in I_2} p_i$ for $i \in I_2$ (with the convention $0/0 \neq 0$), such that

$$d(a, a^1) = \Big(\sum_{i \in I_2} p_i\Big) d(a^1, a^2), \qquad d(a, a^2) = \Big(\sum_{i \in I_1} p_i\Big) d(a^1, a^2).$$

We say that a metric space (X, d) is *convex* (*strictly convex*) if for any two elements a_1, a_2 of X the set $pa_1 + (1-p)a_2$ is non-empty (has exactly one element) for any $p \in [0, 1]$.

Remark 1.1. Given two elements $a, b \in X$ and a real $p \in [0, 1]$, Definition 1.1 reads as follows:

$$pa + (1-p)b = \{c \in X\colon\ d(c, a) = (1-p)d(a, b) \text{ and } d(c, b) = pd(a, b)\}.$$

If a metric space (X, d) is strictly convex, we identify the set $pa + (1-p)b$ with its unique element.

Remark 1.2. If a metric space (X, d) is complete, the above definition of convexity of a metric space agrees with the classical definition of Menger (see [2], Definition 14.1 and Theorem 14.1).

Remark 1.3. If a metric space (X, d) is convex (strictly convex), then the set $\sum_{i=1}^{n} p_i a_i$ is a closed (finite) non-empty subset of X for any $\{a_1, \ldots, a_n\} \subset X$ and $\{p_1, \ldots, p_n\} \subset [0, 1]$ with $\sum_{i=1}^{n} p_i = 1$.

This remark is a direct consequence of Definition 1.1 by the use of an inductive argument.

DEFINITION 1.2. We say that a strictly convex metric space (X, d) is of *negative curvature* iff for any four elements a_1, a_2, b_1, b_2 of X and any $p \in [0, 1]$ the following estimation holds:

$$d(pa_1 + (1-p)a_2,\ pb_1 + (1-p)b_2) \leq pd(a_1, b_1) + (1-p)d(a_2, b_2).$$

Remark 1.4. Let (X, g) be a complete simply connected Riemannian manifold and let d be a geodesic metric on X induced by g. Then the metric space (X, d) is of negative curvature if and only if the manifold (X, g) is of non-positive (sectional) curvature.

This property of sectional curvature of a Riemannian manifold was established by Bussemann in [3].

Remark 1.5. Let $(X, \| \ \|)$ be a *strictly convex* real Banach space (cf. [5]), i.e. such that the metric space (X, d) is strictly convex, where $d(x, y) = \|x-y\|$. Then one verifies easily the following:

(a) For any subsets $\{a_1, \ldots, a_n\} \subset X$ and $\{p_1, \ldots, p_n\} \subset [0, 1]$ with $\sum_{i=1}^n p_i = 1$ the set $\sum_{i=1}^n p_i a_i$ (in the sense of Definition 1.1) is a one-element set containing a linear combination of $a_1, \ldots, a_n$ with the coefficients $p_1, \ldots, p_n$.

(b) The metric space (X, d) is of negative curvature.

Remark 1.6. A metric space (X, d) is said to be *outer convex* iff for any two elements $a, b \in X$ and any $p \in [0, 1]$ there is an element $c \in X$ such that $b = pa+(1-p)c$.

Let (X, d) be a complete convex and outer convex metric space. In [1, Theorem 3.1] the authors have proved that the metric space (X, d) is isometric with a strictly convex real Banach space if and only if for any triplet a_1, a_2, a_3 of elements of X the set $\frac{1}{3}a_1+\frac{1}{3}a_2+\frac{1}{3}a_3$ is a one-element subset of X.

PROPOSITION 1.1. *Suppose (X, d) is a metric space of negative curvature. Then for any finite subsets $\{a_1, \ldots, a_n\}$, $\{b_1, \ldots, b_n\}$ of X and any subset $\{p_1, \ldots, p_n\}$ of $[0, 1]$ with $\sum_{i=1}^n p_i = 1$ the following estimation holds*:

$$h\Big(\sum_{i=1}^{n} p_i a_i, \sum_{i=1}^{n} p_i b_i\Big) \leqslant \sum_{i=1}^{n} p_i d(a_i, b_i). \tag{1.1}$$

Proof. We proceed by induction. For all one-element sets $\{a_1\}$, $\{b_1\}$ our proposition is true. Suppose it is true for all finite subsets of X with cardinality $k < n$, where $n > 1$.

Let $\{a_1, \ldots, a_n\}, \{b_1, \ldots, b_n\} \subset X$ and $\{p_1, \ldots, p_n\} \subset [0, 1]$ with $\sum_{i=1}^n p_i = 1$. It is sufficient to prove (by symmetry) that for each element $a \in \sum_{i=1}^n p_i a_i$ there is an element $b \in \sum_{i=1}^n p_i b_i$ such that

$$d(a, b) \leqslant \sum_{i=1}^{n} p_i d(a_i, b_i).$$

Suppose $a \in \sum_{i=1}^n p_i a_i$. Thus there are two non-empty disjoint and complementary subsets I_1, I_2 of $\{1, \ldots, n\}$ and two elements $a^1 \in \sum_{i\in I_1} p_i^1 a_i$ and $a^2 \in \sum_{i\in I_2} p_i^2 a_i$, where $p_i^1 = p_i/\sum_{i\in I_1} p_i$ for $i \in I_1$ and $p_i^2 = p_i/\sum_{i\in I_2} p_i$ for $i \in I_2$, such that

$$a = \Big(\sum_{i\in I_1} p_i\Big)a^1+\Big(\sum_{i\in I_2} p_i\Big)a^2$$

(see Definition 1.1 and Remark 1.1). It follows from the inductive hypothesis that there are elements $b^1 \in \sum_{i\in I_1} p_i^1 b_i$ and $b^2 \in \sum_{i\in I_2} p_i^2 b_i$ such that

$$d(a^1, b^1) \leqslant \sum_{i\in I_1} p_i^1 d(a_i, b_i) \quad \text{and} \quad d(a^2, b^2) \leqslant \sum_{i\in I_2} p_i^2 d(a_i, b_i).$$

Let $b = (\sum_{i\in I_1} p_i^1)b^1 + (\sum_{i\in I_2} p_i^2)b^2$. Since (X, d) is of negative curvature, we obtain

$$d(a, b) \leqslant (\sum_{i\in I_1} p_i^1)d(a^1, b^1) + (\sum_{i\in I_2} p_i^2)d(a^2, b^2)$$

$$\leqslant \sum_{i\in I_1} p_i d(a_i, b_i) + \sum_{i\in I_2} p_i d(a_i, b_i) = \sum_{i=1}^{n} p_i d(a_i, b_i),$$

which completes the induction and the proof of Proposition 1.1.

LEMMA 1.1. *Suppose (X, d) is a strictly convex metric space. Then for any $a, b \in X$ and $p, p' \in [0, 1]$ the following holds:*

$$d(pa+(1-p)b,\ p'a+(1-p')b) = |p-p'|\,d(a, b).$$

Proof. Suppose $p \geqslant p'$ and let

$$c = (p'/p)(pa+(1-p)b)+(1-p'/p)b.$$

We shall prove that $c = p'a+(1-p')b$. To prove this it is sufficient to show that $d(c, b) = p'd(a, b)$ and $d(c, a) = (1-p')d(a, b)$ (Remark 1.1).

We have

$$d(c, b) = (p'/p)d(pa+(1-p)b, b) \quad \text{and} \quad d(pa+(1-p)a, b) = pd(a, b)$$

(Remark 1.1), and hence $d(c, b) = p'd(a, b)$.

By the triangle inequality we have

$$d(c, a) \leqslant d(c, pa+(1-p)b)+d(pa+(1-p)b, a).$$

But (Remark 1.1)

$$d(c, pa+(1-p)b) = (1-p'/p)d(pa+(1-p)b, b) = (1-p'/p)pd(a, b)$$

and

$$d(pa+(1-p)b, a) = (1-p)d(a, b).$$

Hence we obtain $d(c, a) \leqslant (1-p')d(a, b)$. But this means (together with $d(c, b) = p'd(a, b)$) that $d(c, a) = (1-p')d(a, b)$, which shows finally that $c = p'a+(1-p')b$.

We thus obtain

$$d(pa+(1-p)b,\ p'a+(1-p')b) = d(pa+(1-p)b, c)$$
$$= (1-p'/p)d(pa+(1-p)b, b) = (1-p'/p)pd(a, b) = (p-p')d(a, b),$$

which completes the proof of Lemma 1.1.

LEMMA 1.2. *Let (X, d) be a metric space of negative curvature and let $(F_0(X), h)$ be a subspace of $(F(X), h)$ of non-empty finite subsets of X. Then the map φ of $[0, 1]\times F_0(X)\times F_0(X)$ into $F_0(X)$ defined as*

$$\varphi(p, F, G) = \{c \in X\colon c = pa+(1-p)b,\ a \in F,\ b \in G\}$$

is continuous.

Proof. We shall show that for any $x, x', y, y' \in X$ and any $p, p' \in [0, 1]$ the following inequality holds:

$$(1.2) \qquad d(px+(1-p)y,\ p'x'+(1-p')y') \leqslant pd(x, x')+(1-p)d(y, y') + |p-p'|d(x', y').$$

By the triangle inequality we have

$$d(px+(1-p)y,\ p'x'+(1-p')x') \leqslant d(px+(1-p)y,\ px'+(1-p)y') + d(px'+(1-p)y',\ p'x'+(1-p')y').$$

Since (X, d) is of negative curvature, we obtain

$$d(px+(1-p)y,\ px'+(1-p)y') \leqslant pd(x, x')+(1-p)d(y, y')$$

(Definition 1.2). From Lemma 1.1 we have

$$d(px'+(1-p)y',\ p'x'+(1-p')x') = |p-p'|d(x', y').$$

Thus we obtain (1.2).

The continuity of φ follows directly from the estimation (1.2) and the definition of the Hausdorff metric h.

PROPOSITION 1.2. *Suppose (X, d) is a metric space of negative curvature. Then for any finite subset $\{a_1, \ldots, a_n\}$ of X the application*

$$\psi(p_1, \ldots, p_n) = \sum_{i=1}^{n} p_i a_i$$

is a continuous map of a symplex

$$\Delta_n = \{(p_1, \ldots, p_n):\ \sum_{i=1}^{n} p_i = 1,\ p_i \geqslant 0,\ i = 1, \ldots, n\}$$

into $(F(X), h)$.

Proof. We proceed by induction. For all one-element sets $\{a_1\}$ our proposition is true. Suppose it is true for all finite subsets of X with cardinality $k < n$, where $n > 1$.

Let φ be a map defined in Lemma 1.2 and let $\mathscr{P}$ be a family of all non-empty sets $I \subset \{1, \ldots, n\}$ with non-empty complements I'. For each $I \in \mathscr{P}$ let ψ_I be an application of Δ_n into $(F(X), h)$ defined as

$$\psi_I(p_1, \ldots, p_n) = \varphi(\sum_{i\in I} p_i,\ \sum_{i\in I} p_i^1 a_i,\ \sum_{i\in I'} p_i^2 a_i),$$

where

$$p_i^1 = p_i/\sum_{i\in I} p_i \text{ for } i \in I, \qquad p_i^2 = p_i/\sum_{i\in I'} p_i \text{ for } i \in I'.$$

It follows from Lemma 1.2 and the inductive hypothesis that the application ψ_I is a continuous map from Δ_n into $(F(X), h)$ for any $I \in \mathscr{P}$. By the Definition 1.1

of a convex combination of n elements of X we have

$$\psi(p_1, \ldots, p_n) = \bigcup_{I\in\mathscr{P}} \psi_I(p_1, \ldots, p_n) \quad \text{for } (p_1, \ldots, p_n) \in \Delta_n.$$

This means that ψ is a continuous map of Δ_n into $(F(X), h)$ as a finite union of continuous maps ψ_I, which completes the induction and the proof of Proposition 1.2.

2. Mathematical expectation. Let (X, d) be a convex metric space and let $(\Omega, \mathscr{A}, P)$ be a probability space. By $\mathscr{S} = \mathscr{S}(\Omega, \mathscr{A}, P; X)$ we denote the set of *simple* random variables (r.v.) with values in X, i.e., Borel maps of Ω into X having a finite number of values.

We say that $\pi = \{A_1, \ldots, A_k\} \subset \mathscr{A}$ is a *partition* if $A_i \cap A_j = \emptyset$ for $i \neq j$, $i, j = 1, \ldots, k$, and $\bigcup_{i=1}^{k} A_i = \Omega$. If π and σ are partitions, we write $\pi \leqslant \sigma$ if each element of σ is included in some element of π.

Given a function $f \in \mathscr{S}$ we denote by $\Pi(f)$ the set of all partitions π such that $f(\omega) = f(\tilde{\omega})$ for any $A \in \pi$ and any $\omega, \tilde{\omega} \in A$.

DEFINITION 2.1. Given $f \in \mathscr{S}(\Omega, \mathscr{A}, P; X)$ and $\pi = \{A_1, \ldots, A_k\} \in \Pi(f)$ we define

$$E_\pi[f] = \sum_{i=1}^{k} P(A_i) a_i, \quad \text{where } a_i = f(\omega) \text{ for } \omega \in A_i,\ i = 1, \ldots, k,$$

and

$$E[f] = \mathrm{cl}\Big(\bigcup_{\pi\in\Pi(f)} E_\pi[f]\Big).$$

Remark 2.1. Suppose (X, d) is a complete convex and outer convex metric space. In [1, Theorem 2.1] the authors have proved that a metric space (X, d) is isometric with a real strictly convex Banach space if and only if for any triplet of elements a_1, a_2, a_3 of X the following equality holds:

$$\tfrac{1}{2}a_1 + \tfrac{1}{2}(\tfrac{1}{2}a_2 + \tfrac{1}{2}a_3) = \tfrac{1}{2}(\tfrac{1}{2}a_1 + \tfrac{1}{2}a_2) + \tfrac{1}{2}(\tfrac{1}{2}a_1 + \tfrac{1}{2}a_3).$$

Thus, in general, given $f \in \mathscr{S}$ the set $E_\pi[f]$ depends on the partition $\pi \in \Pi(f)$.

LEMMA 2.1. *Given $f \in \mathscr{S}$, the operator $E_\bullet[f]$ is increasing: If $\pi, \sigma \in \Pi(f)$ and $\pi \leqslant \sigma$, then $E_\pi[f] \subset E_\sigma[f]$.*

The proof results clearly from Definition 1.1 of a convex combination.

LEMMA 2.2. *Suppose a metric space (X, d) is of negative curvature. Then for any $f, g \in \mathscr{S}$ and any $\pi \in \Pi(f) \cap \Pi(g)$ the following inequality holds:*

$$h(E_\pi[f], E_\pi[g]) \leqslant \int_\Omega d(f(\omega), g(\omega))\, dP(\omega).$$

The proof results clearly from Proposition 1.1.

LEMMA 2.3. *If (X, d) is of negative curvature, then for any $f \in \mathscr{S}$ the set $E[f]$ is a bounded subset of X.*

Proof. Since the set $\Pi(f)$ is directed by the relation $\leqslant$ and the operator $E_{\bullet}[f]$ is increasing (Lemma 2.1), it is sufficient to prove that

$$\sup_{\pi\in\Pi(f)} \operatorname{diam} E_\pi[f] < \infty. \tag{2.1}$$

Let $\pi\in\Pi(f)$ and let $g(\omega) = a$, $\omega\in\Omega$, for some fixed element a of X. From Lemma 2.2 we have

$$h(E_\pi[f], \{a\}) \leqslant \int_\Omega d(f(\omega), a)\,dP(\omega),$$

which implies (2.1).

PROPOSITION 2.1. *Let (X, d) be a metric space of negative curvature. Then for any $f, g\in\mathscr{S}(\Omega, \mathscr{A}, P; X)$*

$$h(E[f], E[g]) \leqslant \int_\Omega d(f(\omega), g(\omega))\,dP(\omega). \tag{2.2}$$

Proof. To prove (2.2) it is sufficient to show that for any partition $\pi\in\Pi(f)$ and any element $a\in E_\pi[f]$ there is a partition $\sigma\in\Pi(g)$ and an element $b\in E_\sigma[g]$ such that $d(a, b)\leqslant\int_\Omega d(f, g)\,dP$.

For any partition $\pi\in\Pi(f)$ there is a partition $\sigma\in\Pi(f)\cap\Pi(g)$ such that $\pi\leqslant\sigma$. Since $h(E_\sigma[f], E_\sigma[g])\leqslant\int_\Omega d(f, g)\,dP$ (Lemma 2.2) and $E_\pi[f]\subset E_\sigma[g]$ (Lemma 2.1), for any $a\in E_\pi[f]$ there is $b\in E_\sigma[g]$ such that $d(a, b)\leqslant\int d(f, g)\,dP$.

Suppose (X, d) is a Polish metric space of negative curvature. We say that an X-valued random variable f is *integrable* iff $\int_\Omega d(x, f(\omega))\,dP(\omega) < \infty$ for $x\in X$. We denote by $\mathscr{L} = \mathscr{L}(\Omega, \mathscr{A}, P; X)$ the set of all integrable r.v.'s. and by $L = L(\Omega, \mathscr{A}, P; X)$ the set of all equivalence classes (for equality a.s.) of integrable r.v.'s. By $S = S(\Omega, \mathscr{A}, P; X)$ we denote a subset of L corresponding to the set of simple r.v.'s. Let d_1 be a metric on L given by

$$d_1(f, g) = \int_\Omega d(f(\omega), g(\omega))\,dP(\omega).$$

Since the metric space (X, d) is Polish, S is a dense subset of (L, d_1) (see Lemma 3.1). It is clear that $E[f] = E[g]$ if $f, g\in\mathscr{S}$ and $f = g$ a.s., i.e. an operator E acts on S. By Proposition 2.1, E is a uniformly continuous map of (S, d_1) into $(F(X), h)$. It is known that if a metric space (X, d) is complete, then a metric space $(F(X), h)$ is also complete [9, Vol. 1, § 33, IV]. Thus E admits a unique uniformly continuous extension to a map of (L, d_1) into $(F(X), h)$ which satisfies (2.2) for all $f, g\in L(\Omega, \mathscr{A}, P; X)$.

DEFINITION 2.2. Let (X, d) be a Polish metric space, $(\Omega, \mathscr{A}, P)$ a probability space, and $f\in\mathscr{L}(\Omega, \mathscr{A}, P; X)$. We say that a non-empty closed bounded subset $E[f]$ of X is a *mathematical expectation* of f. For any $f, g\in\mathscr{L}(\Omega, \mathscr{A}, P; X)$ an estimation (2.2) holds.

Remark 2.2. Suppose (X, d) is a Polish metric space of negative curvature. Then any $f\in\mathscr{L}(\Omega, \mathscr{A}, P; X)$ is integrable in the sense of Doss ([7], Definition 1). This is a consequence of the estimation (2.2) applied to f and $g(\omega) = x$, $\omega \in \Omega$.

Remark 2.3. Suppose $(X, \| \ \|)$ is a strictly convex real separable Banach space and $f \in \mathscr{L}(\Omega, \mathscr{A}, P; X)$, where X is equipped with the metric $d(x, y) = \|x-y\|$. Then $E[f]$ is a one-element set containing a Bochner integral of f (see Remark 1.5).

3. Strong Law of Large Numbers. We say that a metric space (X, d) is *finitely compact* iff each closed bounded subset of X is compact. Throughout this section we assume that (X, d) is a finitely compact metric space of negative curvature.

We put $\mathrm{Lim}_n F_n = F$ iff $\lim_n h(F_n, F) = 0$. We note the following known properties of the convergence in $(F(X), h)$ ([9], Vol. II, §42,1 and §42,2):

(α) $F_n \subset F'_n$ implies $\mathrm{Lim}_n F_n \subset \mathrm{Lim}_n F'_n$.

(β) If $\bigcup_n F_n$ is relatively compact in (X, d), then $\{F_n\}_{n=1}^{\infty}$ is relatively compact in $(F(X), h)$.

THEOREM 3.1. *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of independent and identically distributed (i.i.d.) integrable random variables with values in X and let*

$$F_n(\omega) = \sum_{i=1}^{n} \frac{1}{n} f_i(\omega) \quad \text{for } \omega \in \Omega,\ n = 1, 2, \ldots$$

Then

$$\mathrm{Lim}_n F_n(\omega) = E[f_1] \ \text{a.s.} \tag{3.1}$$

We shall precede the proof of Theorem 3.1 by three lemmas.

LEMMA 3.1. *Suppose $f_n \in \mathscr{L}(\Omega, \mathscr{A}, P; X)$ $(n = 1, 2, \ldots)$ are i.i.d. r.v.'s. Given $\varepsilon > 0$ there exists a sequence $g_n \in \mathscr{S}(\Omega, \mathscr{A}, P; X)$ $(n = 1, 2, \ldots)$ of i.i.d. r.v.'s such that $d(f_n, g_n)$ $(n = 1, 2, \ldots)$ are i.i.d. (real) r.v.'s and $\int_\Omega d(f_1, g_1)dP \leqslant \varepsilon$.*

Proof. Let $\varepsilon > 0$ be given and let x be a fixed element of X. Since the metric space (X, d) is Polish, there exists a compact subset K of X such that

$$\int_{\Omega \setminus f_1^{-1}(K)} d(x, f_1(\omega))dP(\omega) \leqslant \varepsilon/2.$$

Let $K = \bigcup_{i=1}^{k} B_i$, where B_i are non-empty pairwise disjoint Borel subsets of X with $\mathrm{diam}(B_i) \leqslant \varepsilon/2$ for $i = 1, \ldots, k$. Suppose $a_i \in B_i$ for $i = 1, \ldots, k$ and let us define

$$g_n(\omega) = \begin{cases} a_i & \text{if } \omega \in f_n^{-1}(B_i) \text{ for } i = 1, \ldots, k, \\ x & \text{if } \omega \notin f_n^{-1}(K). \end{cases}$$

It is clear that $\{g_n\}_{n=1}^{\infty}$ and $\{d(f_n, g_n)\}_{n=1}^{\infty}$ are i.i.d. sequences and

$$\int_\Omega d(f_1, g_1)dP = \int_{f_1^{-1}(K)} d(f_1, g_1)dP + \int_{\Omega \setminus f_1^{-1}(K)} d(f_1, x)dP \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

LEMMA 3.2. *Let $f \in \mathcal{S}(\Omega, \mathcal{A}, P; X)$. Then for each $\varepsilon > 0$ there exists a partition $\pi \in \Pi(f)$ such that $h(E_\pi[f], E[f]) \leqslant \varepsilon$.*

Proof. Given $f \in \mathcal{S}(\Omega, \mathcal{A}, P; X)$ the set $F = \bigcup_{\pi \in \Pi(f)} E_\pi[f]$ is bounded (Lemma 2.3), and thus relatively compact. Let $\{a_1, \ldots, a_k\}$ be an ε-net in F. Then $\{a_1, \ldots, a_k\} \subset E_\pi[f]$ for some $\pi \in \Pi(f)$, since $\Pi(f)$ is directed by $\leqslant$ and the operator $E_\bullet[f]$ is increasing (Lemma 2.1). It is clear that

$$h(E_\pi[f], E[f]) = h(E_\pi[f], \operatorname{cl} F) \leqslant \varepsilon.$$

LEMMA 3.3. *Suppose a probability space $(\Omega, \mathcal{A}, P)$ is non-atomic. Then for any partition $\pi = \{A_1, \ldots, A_k\}$ there exists a sequence $\{\pi_n\}_{n=1}^{\infty}$ of partitions*

$$\pi_n = \{A_{i,j}^n, B_l^n\colon 1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant m_n^i,\ 1 \leqslant l \leqslant r_n\} \quad (n = 1, 2, \ldots)$$

such that

$$A_{i,j}^n \subset A_i, \quad P\Big(A_i \setminus \bigcup_{j=1}^{m_n^i} A_{i,j}^n\Big) < 1/n, \quad P(A_{i,j}^n) = P(B_l^n) = 1/n$$

for $1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant m_n^i$, $1 \leqslant l \leqslant r_n$ $(n = 1, 2, \ldots)$.

Proof. Construction of the sequence $\{\pi_n\}_{n=1}^{\infty}$ is clear in view of the well-known property of a non-atomic measure: $\{P(B)\colon B \in \mathcal{A},\ B \subset A\} = [0, P(A)]$ for any $A \in \mathcal{A}$.

Proof of Theorem 3.1. If $(\Omega, \mathcal{A}, P)$ has an atom, then the existence of a sequence of i.i.d. r.v.'s defined on Ω implies that $\mathcal{A} = \{\emptyset, \Omega\}$ and our S.L.L.N. is trivially true. Thus we suppose that $(\Omega, \mathcal{A}, P)$ is a non-atomic probability space.

Assume first that f_1 is a simple random variable. Let $\varepsilon > 0$ be given. By Lemma 3.2 there exists $\pi \in \Pi(f_1)$ such that $h(E_\pi[f_1], E[f_1]) \leqslant \varepsilon$. Suppose $\pi = \{A_1, \ldots, A_k\}$ and $f(\omega) = a_i$ for $\omega \in A_i$, $i = 1, \ldots, k$. Let $\{\pi_n\}_{n=1}^{\infty}$ be a sequence of partitions constructed for that π in Lemma 3.3. Let $\{f_n'\}_{n=1}^{\infty}$ be a sequence of simple r.v.'s defined as

$$f_n'(\omega) = a_i \quad \text{for } \omega \in A_i^n \ (i = 1, \ldots, k),$$

where

$$A_1^n = \bigcup_{j=1}^{m_n^1} A_{1,j}^1 \cup \bigcup_{l=1}^{r_n} B_l^n \quad \text{and} \quad A_i^n = \bigcup_{j=1}^{m_n^i} A_{i,j} \text{ for } i = 2, \ldots, k.$$

We shall prove that

$$\text{(3.2)} \qquad \lim_n h(F_n(\omega), E_{\pi_n}[f_n']) = 0 \quad \text{for all } \omega \in \Omega \setminus N \text{ with } P(N) = 0.$$

Let us define for each $\omega \in \Omega$:

$$\nu_n^i(\omega) = \operatorname{card}\{s\colon 1 \leqslant s \leqslant n, f_s(\omega) = a_i\} \quad (i = 1, \ldots, k;\ n = 1, 2, \ldots).$$

We thus have (Definition 1.1)

$$F_n(\omega) = \overbrace{n^{-1}a_1 + \ldots + n^{-1}a_1}^{\nu_n^1(\omega)} + \overbrace{n^{-1}a_2 + \ldots + n^{-1}a_2}^{\nu_n^2(\omega)} + \ldots + \overbrace{n^{-1}a_k + \ldots + n^{-1}a_k}^{\nu_n^k(\omega)}$$

and

$$E_{\pi_n}[f_n'] = \overbrace{n^{-1}a_1 + \ldots + n^{-1}a_1}^{m_n^1 + r_n} + \overbrace{n^{-1}a_2 + \ldots + n^{-1}a_2}^{m_n^2} + \ldots + \overbrace{n^{-1}a_k + \ldots + n^{-1}a_k}^{m_n^k}.$$

Hence, by Proposition 1.1 we see that for each $\omega \in \Omega$ and $i = 1, 2, \ldots$

$$h(F_n(\omega), E_{\pi_n}[f_n']) \leqslant \sup_{1 \leqslant i,j \leqslant k} d(a_i, a_j) \left(\frac{|\nu_n^1(\omega) - m_n^1 - r_n|}{n} + \frac{|\nu_n^2(\omega) - m_n^2|}{n} + \ldots + \frac{|\nu_n^k(\omega) - m_n^k|}{n} \right).$$

For fixed $i = 1, \ldots, k$ the random variables ν_n^i $(n = 1, 2, \ldots)$ are the n-th partial sums of a sequence of i.i.d. r.v.'s with mean $P(A_i)$. By the construction of the partitions π_n, $P(A_i) - 1/n < m_n^i/n \leqslant P(A_i)$ and $r_n \leqslant k$ for $i = 1, \ldots, k$, $n = 1, 2, \ldots$ We thus infer, by the (real) S.L.L.N., that the right-hand side of the last inequality converges to zero for almost every $\omega \in \Omega$, which proves (3.2).

Let $x \in X$ and $\omega \in \Omega$ be fixed. An application of inequality (1.1) of Proposition 1.1 (for $a_i = x$, $b_i = f_i(\omega)$, $p_i = 1/n$ for $i = 1, 2, \ldots, n$) shows that

$$h(\{x\}, F_n(\omega)) \leqslant \sum_{i=1}^{n} n^{-1} d(x, f_i(\omega)) \quad \text{for all } \omega \in \Omega.$$

From the (real) S.L.L.N. applied to a sequence $\{d(x, f_n)\}_{n=1}^{\infty}$ of integrable i.i.d. r.v.'s we obtain

$$\limsup_n h(\{x\}, F_n(\omega)) \leqslant \int_\Omega d(x, f_1) dP \text{ a.s.}$$

This estimation means that for $\omega \in \Omega \setminus N'$ with $P(N') = 0$ the set $\bigcup_n F_n(\omega)$ is bounded, and thus relatively compact.

We shall prove that (3.1) holds for every $\omega \in \Omega \setminus (N \cup N')$. We may assume by (β), extracting a subsequence if necessary, that $\{F_n(\omega)\}_{n=1}^{\infty}$ is convergent in $(F(X), h)$. Thus for $\omega \in \Omega \setminus (N \cup N')$ we have, by (3.1),

$$\operatorname{Lim}_n F_n(\omega) = \operatorname{Lim}_n E_{\pi_n}[f_n'].$$

Since, by construction,

$$\lim_n \int_\Omega d(f_n', f_1) dP = 0,$$

we have $\operatorname{Lim}_n E[f_n'] = E[f_1]$. But $E_{\pi_n}[f_n'] \subset E[f_n']$ for $n = 1, 2, \ldots$, and thus by (α) we obtain $\operatorname{Lim}_n F_n(\omega) \subset E[f_1]$.

Let us consider the sequence of partitions $\sigma_n = \{A_1^n, \ldots, A_k^n\}$ $(n = 1, 2, \ldots)$. Since

$$\lim_n P(A_i^n \Delta A_i) = 0 \quad \text{for } i = 1, \ldots, k,$$

by Proposition 1.2 we obtain $\mathrm{Lim}_n E_{\sigma_n}[f_n'] = E_\pi[f_1]$. But $\sigma_n \leqslant \pi_n$ $(n = 1, 2, \ldots)$, and hence, by (α),

$$E_\pi[f_1] \subset \mathrm{Lim}_n E_{\pi_n}[f_n'] = \mathrm{Lim}_n F_n(\omega).$$

We thus finally obtain

$$E_\pi[f_1] \subset \lim_n F_n(\omega) \subset E[f_1] \quad \text{for all } \omega \in \Omega \backslash (N \cup N').$$

This implies, by the inequality $h(E_\pi[f_1], E[f_1]) \leqslant \varepsilon$, that $h(\mathrm{Lim}_n F_n(\omega), E[f_1]) \leqslant \varepsilon$. Since $\varepsilon > 0$ was chosen arbitrarily, the proof of (3.1) in the case of a simple r.v. f_1 is complete.

Let now $f_1 \in \mathscr{L}(\Omega, \mathscr{A}, P; X)$ be arbitrary and let $\varepsilon > 0$ be given. By Lemma 3.1 there is a sequence $\{g_n\}_{n=1}^\infty$ of simple X-valued i.i.d. r.v.'s such that $\{d(f_n, g_n)\}_{n=1}^\infty$ is an i.i.d. sequence and $\int_\Omega d(f_1, g_1)dP \leqslant \varepsilon$.

Let $G_n(\omega) = \sum_{i=1}^n n^{-1} g_i(\omega)$ for $\omega \in \Omega$, $n = 1, 2, \ldots$ From Proposition 1.1 we obtain

$$h(F_n(\omega), G_n(\omega)) \leqslant \sum_{i=1}^n n^{-1} d(f_i(\omega), g_i(\omega)) \quad \text{for } \omega \in \Omega.$$

Strong Law of Large Numbers applied to a sequence $\{d(f_n, g_n)\}_{n=1}^\infty$ implies that

$$\limsup_n h(F_n(\omega), G_n(\omega)) \leqslant \varepsilon \text{ a.s.}$$

By the triangle inequality we have

$$h(F_n(\omega), E[f_1]) \leqslant h(F_n(\omega), G_n(\omega)) + h(G_n(\omega), E[g_1]) + h(E[g_1], E[f_1]).$$

Since $\lim_n h(G_n(\omega), E[g_1]) = 0$ a.s. and

$$h(E[g_1], E[f_1]) \leqslant \int_\Omega d(f_1, g_1)dP \leqslant \varepsilon$$

(see Definition 2.2), we obtain

$$\limsup_n h(F_n(\omega), E[f_1]) \leqslant 2\varepsilon \text{ a.s.}$$

Since $\varepsilon > 0$ was chosen arbitrarily, this completes the proof of Theorem 3.1.

Strong Law of Large Numbers of Theorem 3.1 admits the following converse:

THEOREM 3.2. *Let $\{f_n\}_{n=1}^\infty$ be a sequence of i.i.d. X-valued r.v.'s and let $F_n(\omega) = \sum_{i=1}^n n^{-1} f_i(\omega)$ for $\omega \in \Omega$, $n = 1, 2, \ldots$ Suppose there exists $F \in F(X)$ such that $\lim_n h(F_n(\omega), F) = 0$ a.s. Then $f_1 \in \mathscr{L}(\Omega, \mathscr{A}, P; X)$ and $E[f_1] = F$.*

Proof. In view of Theorem 3.1 we need to prove only that $f_1 \in \mathscr{L}(\Omega, \mathscr{A}, P; X)$, that is $\int_\Omega d(x, f_1)dP < \infty$ for $x \in X$. To prove this it is sufficient to show that for a fixed $x \in X$ there is a constant M such that

(3.3) $$\limsup_n n^{-1} d(x, f_n(\omega)) \leqslant M$$

(for Kolmogorov's proof of the converse of the S.L.L.N. see, e.g., [10, Theorem 3.2.2]).

For each $\omega \in \Omega$ let $\{a_n(\omega)\}_{n=2}^{\infty}$ be an arbitrary sequence of elements $a_n(\omega) \in F_{n-1}(\omega)$ and let

$$b_n(\omega) = \tfrac{n-1}{n} a_n(\omega) + \tfrac{1}{n} f_n(\omega) \quad \text{for } n = 2, 3, \ldots$$

(Definition 1.1). By the triangle inequality we obtain

(3.4) $$n^{-1} d(x, f_n(\omega)) \leqslant n^{-1} d(x, a_n(\omega)) + d(a_n(\omega), b_n(\omega)), \quad \omega \in \Omega,\ n = 2, 3, \ldots$$

Since $\lim_n h(F_n(\omega), F) = 0$ a.s., we have

$$\lim_n h(\{x\}, F_n(\omega)) = h(\{x\}, F) = \tfrac{1}{2} M \text{ a.s.}$$

This implies (since $a_n(\omega) \in F_n(\omega)$, $b_n(\omega) \in F_n(\omega)$, $h(\{x\}, F_n(\omega)) = \sup_{y \in F_n(\omega)} d(x, y)$) that

$$\limsup_n d(x, a_n(\omega)) \leqslant \tfrac{1}{2} M \quad \text{and} \quad \limsup_n d(x, b_n(\omega)) \leqslant \tfrac{1}{2} M \quad \text{a.s.}$$

By the triangle inequality we obtain

$$\limsup_n d(a_n(\omega), b_n(\omega)) \leqslant M \text{ a.s.},$$

which implies (3.3) in view of (3.4).

REFERENCES

[1] E. Z. Andalafte, J. E. Valentine and S. G. Wayement, *Triangle median properties which characterize Banach spaces*, Houston J. Math. 5, No. 3 (1979), pp. 307–312.
[2] L. M. Blumenthal, *Theory and Applications of Distance Geometry*, Oxford 1953.
[3] H. Bussemann, *Spaces with non-positive curvature*, Acta Math. 80 (1948), pp. 259–310.
[4] – *The Theory of Geodesics*, Academic Press, New York 1955.
[5] J. Diestel, *Geometry of Banach Spaces*, Springer-Verlag, Berlin 1975.
[6] M. Gromov, *Structures métriques pour les variétés riemanniennes*, Cedic, Paris 1981.
[7] W. Herer, *Espérance mathématique au sens de Doss d'une variable aléatoire à valeurs dans un espace métrique*, C. R. Acad. Sci. Paris Sér. I 302 (1983), pp. 131–134.
[8] – *Espérance mathématique d'une variable aléatoire à valeurs dans un space métrique à courbure négative*, ibidem 306 (1988), pp. 681–684.
[9] K. Kuratowski, *Topology*, Vol. I, II, Academic Press, New York 1966, 1968.
[10] W. F. Stout, *Almost Sure Convergence*, Academic Press, New York 1974.

Institute of Mathematics, Technical University of Warsaw
pl. Politechniki 1, 00-661 Warsaw, Poland

Received on 17.2.1989;
revised version on 12.12.1991