L-RANDOM ELEMENTS AND L^*-RANDOM ELEMENTS IN BANACH SPACES

EDITH MOURIER
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, PARIS

1. Introduction

It is well known that many applications of the theory of probability require the consideration of general random elements. We are mainly interested in generalizing laws of large numbers and central limit theorems, so we must consider random elements in topological vector spaces. As a first approach we study random elements in Banach spaces and in the first part of this paper give the definitions and results we have obtained (see [1] and [9]). The case of Banach spaces may seem very limited; however, we show in the second part that it allows many applications. Finally, in the last part, we indicate a new point of view and the results we have established in this way.

2. Definition of L-random elements in Banach spaces and their mathematical expectation

Let $(U, \mathcal{Q}, m)$ be a fundamental probability space of elements u, let $\mathfrak{X}$ be a Banach space of elements x, and $x(u)$ a function on U to $\mathfrak{X}$. We call $X = x(u)$ a *random element* in $\mathfrak{X}$, that is, the "value" of the random element X is $x(u) \in \mathfrak{X}$ if the outcome of the experiment is $u \in U$.

Let $\mathfrak{X}^*$ be the dual space of $\mathfrak{X}$, that is, the space of all continuous linear functionals x^* on $\mathfrak{X}$. We shall write $\langle x^*, x\rangle$ for the number obtained by applying the linear functional $x^* \in \mathfrak{X}^*$ to $x \in \mathfrak{X}$. The *mathematical expectation* of X is the element $E(X)$ of $\mathfrak{X}$, if one exists, such that, for all $x^* \in \mathfrak{X}^*$,

$$\langle x^*, E(X)\rangle = E(\langle x^*, X\rangle). \tag{2.1}$$

If $E(X)$ exists, it is unique [6], and $E(X)$ is the Pettis integral [11], $\int_U x(u)dm$, of $x(u)$ with respect to the measure m.

Such a definition of mathematical expectation implies that, for all $x^* \in \mathfrak{X}^*$, $\langle x^*, x(u)\rangle$ is a measurable function of u. We call *L-random elements* those random elements for which this property is fulfilled, and in this section we shall consider only L-random elements.

It is easy to prove (see [9]) the following:

(a) If α is a given number and if $E(X)$ exists, then $E(\alpha X)$ exists and $E(\alpha X) = \alpha E(X)$.

(b) If X is almost surely (a.s.) equal to a given element x, then $E(X)$ exists and $E(X) = x$.

(c) If X is almost surely (a.s.) equal to a given element x, and if A is a random variable such that $E(A)$ exists, then $E(AX)$ exists and $E(AX) = xE(A)$.

This paper was prepared with partial support of the Office of Naval Research.

(d) If $E(X_1)$ and $E(X_2)$ exist, then $E(X_1 + X_2)$ exists and $E(X_1 + X_2) = E(X_1) + E(X_2)$.

(e) If F is a linear function on $\mathfrak{X}$ to a Banach space $\mathfrak{Y}$ and if $E(X)$ exists, then $E[F(X)]$ exists and $E[F(X)] = F[E(X)]$.

(f) If $\|X\|$ is measurable and if $E(X)$ exists, then $\|E(X)\| \leqq E(\|X\|)$.

If $\mathfrak{X}$ is a *separable* Banach space and if $X = x(u)$ is an L-random element, then $\|x(u)\|$ is measurable. More generally, any continuous numerical function of x is measurable. If $\mathfrak{X}$ is a *reflexive, separable* Banach space and if $E(\|X\|) = m < +\infty$, then $E(X)$ exists. We shall see, as a consequence of the strong law of large numbers, that this property is true even if $\mathfrak{X}$ is not a reflexive space.

3. Laws of large numbers

3.1. *Strong law of large numbers. If $\mathfrak{X}$ is a separable Banach space and $\{X_j\}$ is a strictly stationary sequence of L-random elements in $\mathfrak{X}$ such that $E(\|X_j\|) < +\infty$, then almost surely $(1/n) \sum_{j=1}^{n} X_j$ tends, as $n \to +\infty$, to a limit Y which is an L-random element such that $E(\|Y\|) < +\infty$.* As a particular case let us suppose that $\{X_j\}$ is a sequence of independent L-random elements with the same probability law and with $E(\|X_j\|) < +\infty$. If a sequence converges strongly to a limit it converges weakly; therefore, for all $x^* \in \mathfrak{X}^*$,

$$\left\langle x^*, \left[\frac{1}{n}\sum_{j=1}^{n} X_j\right]\right\rangle = \frac{1}{n}\sum_{j=1}^{n} \langle x^*, X_j\rangle \to \langle x^*, Y\rangle, \text{ a.s.} \tag{3.1}$$

Hence, by Kolmogoroff's theorem, $(1/n) \sum_{j=1}^{n} \langle x^*, X_j\rangle \to E(\langle x^*, X_j\rangle)$ a.s. Thus, $Y = E(X_j)$, and this proves that if $\mathfrak{X}$ is a separable Banach space, $E(\|X\|) < +\infty$ implies that $E(X)$ exists.

3.2. *Law of large numbers in mean of order α. If $\mathfrak{X}$ is a separable Banach space and $\{X_j\}$ is a strictly stationary sequence of L-random elements in $\mathfrak{X}$ such that $E(\|X_j\|^\alpha) < +\infty$, $1 \leqq \alpha < +\infty$, there exists in $\mathfrak{X}$ an L-random element Y such that $E(\|Y\|^\alpha) < +\infty$, $E(Y) = E(X_j)$ and $\lim_{n\to\infty} E\left[\left\|(1/n) \sum_{j=1}^{n} X_j - Y\right\|^\alpha\right] = 0$.* The proof is based on an ergodic theorem of Yosida and Kakutani [12] and on the study [9], [1] of an auxiliary Banach space. This theorem reduces to an important classical theorem when $\mathfrak{X}$ is the real line and $\alpha = 2$. However, the theorem for any $\alpha \geqq 1$ is new. It is obvious that the theorem includes the case of multidimensional random variables.

3.3. *Rapidity of convergence in the mean.* It is very useful to know with what rapidity $E\left[\left\|(1/n) \sum_{j=1}^{n} X_j - Y\right\|^\alpha\right]$ tends to zero as n tends to infinity.

Let us suppose in addition that the X_j are *independent* L-random elements; then we can suppose without loss of generality that $EX_j = Y = \theta$ (the symbol θ denotes the null element of the space $\mathfrak{X}$). If $\mathfrak{X}^*$ is *separable* and if $\alpha \geqq 2$, there exists a positive number ρ such that, for all n,

$$E\left(\left\|\frac{1}{n}\sum_{j=1}^{n} X_j\right\|^\alpha\right) \geqq \rho_n^{-\alpha/2}. \tag{3.2}$$

Obviously it would be particularly interesting to have an inequality in the other direction. To get such a result it seems necessary to make some assumptions on $\mathfrak{X}$. Let us say that a Banach space $\mathfrak{X}$ is a $\mathfrak{G}$-space if there exists a positive number K and a function g on $\mathfrak{X}$ to $\mathfrak{X}^*$ such that for all $x, y \in \mathfrak{X}$

$$\|g(x)\| = \|x\|, \tag{3.3}$$

$$\langle g(x), x \rangle = \|x\|^2, \tag{3.4}$$

$$\|g(x) - g(y)\| \leqq k\|x - y\|. \tag{3.5}$$

In particular, spaces L_α, $\alpha \geqq 2$, are $\mathfrak{G}$-spaces.

If $\mathfrak{X}$ is a *separable* $\mathfrak{G}$-space and if $X_1, \cdots, X_n$ are independent L-random elements in $\mathfrak{X}$, with the same probability law *or not*, such that, for all j, $E(X_j) = \theta$ and $E(\|X_j\|^2) < +\infty$, then, for all n,

$$E\left(\left\|\sum_{j=1}^{n} X_j\right\|^2\right) \leqq K \sum_{j=1}^{n} E(\|X_j\|^2). \tag{3.6}$$

If we suppose that $n \to +\infty$ and that $\sum_{j=1}^{n} E(\|X_j\|^2) = O(n^\beta)$ with $\beta < 2$, it follows from (3.6) that, as n tends to infinity, $(1/n) \sum_{j=1}^{n} X_j$ tends to θ in mean of order two and also a.s. tends strongly to θ. Thus, not only do we have information about the rapidity of convergence, but also we get laws of large numbers under different assumptions: $\mathfrak{X}$ is a separable $\mathfrak{G}$-space, the X_j are independent but *not* of the same law. Then the inequality (3.6) is very important and we shall show now how it may be proved and why we have to make assumptions on $\mathfrak{X}$. Let $X_1, \cdots, X_n$ be independent L-random elements such that $E(X_j) = \theta$, $E(\|X_j\|^2) < +\infty$, $j = 1, \cdots, n$, and let $\mathfrak{X}$ be a separable $\mathfrak{G}$-space. We have

$$\begin{aligned} \|X_1 + X_2 + \cdots + X_n\|^2 &= \langle g(X_1 + X_2 + \cdots + X_n), X_1 + X_2 + \cdots + X_n \rangle \\ &= \sum_{j}^{n} \langle g(X_1 + X_2 + \cdots + X_n), X_j \rangle. \end{aligned} \tag{3.7}$$

Let us write $T_j = X_1 + X_2 + \cdots + X_{j-1} + X_{j+1} + \cdots + X_n$ so that $X_1 + X_2 + \cdots + X_n = T_j + X_j$. Let us put $g(X_1 + X_2 + \cdots + X_n) = g(T_j) + X^{*j}$ and, since $\mathfrak{X}$ is a $\mathfrak{G}$-space, we have $\|X^{*j}\| \leqq K\|X_j\|$. Then

$$\|X_1 + X_2 + \cdots + X_n\|^2 = \sum_{j=1}^{n} \langle g(T_j), X_j \rangle + \sum_{j=1}^{n} \langle X^{*j}, X_j \rangle, \tag{3.8}$$

and to evaluate the mathematical expectation of $\left\|\sum_{j=1}^{n} X_j\right\|^2$ we have to evaluate the mathematical expectation of $\langle g(T_j), X_j \rangle$ and of $\langle X^{*j}, X_j \rangle$. We shall prove below that, if $\mathfrak{X}$ is *separable*, then $E(\langle g(T_j), X_j \rangle) = 0$, so that

$$E\left(\left\|\sum_{j=1}^{n} X_j\right\|^2\right) = E\left(\sum_{j=1}^{n} \langle X^{*j}, X_j \rangle\right) \leqq K \sum_{j=1}^{n} E\left(\left\|X_j\right\|^2\right). \tag{3.9}$$

3.4. *Mathematical expectation of* $\langle X^*, X\rangle$. Let us suppose that X and X^* are independent random elements in a Banach space $\mathfrak{X}$ and in its dual space $\mathfrak{X}^*$ respectively. Let μ and μ^* be the probability measures of X and X^* respectively. The product measure $\mu \times \mu^*$ defines a measure in $\mathfrak{X} \times \mathfrak{X}^*$. Let us consider the random variable

$$A = \langle X^*, X\rangle . \tag{3.10}$$

THEOREM 3.1. *Let X be an L-random element in $\mathfrak{X}$ such that $E(X) = \theta$ and let X^* be any random element in $\mathfrak{X}^*$. If A, as defined by* (3.10), *is $\mu \times \mu^*$-measurable and if $E(|A|) < +\infty$, then $E(A) = 0$.*

This follows immediately from Fubini's theorem.

Sufficient conditions for A to be $\mu \times \mu^*$-measurable and $E(|A|) < +\infty$ are

1) $\mathfrak{X}^*$ is separable,
2) spheres of $\mathfrak{X}^*$ are μ^*-measurable,
3) $E(\|X^*\|) < +\infty$ and $E(\|X\|) < +\infty$.

It is sufficient to note that $\langle x^*, x\rangle$ is a continuous function (with the strong topology) on $\mathfrak{X} \times \mathfrak{X}^*$. Obviously, if there exists a complete vectorial subspace $\mathfrak{X}'^*$ of $\mathfrak{X}^*$ such that almost surely $X^* \in \mathfrak{X}'^*$, it is sufficient that the above conditions hold on $\mathfrak{X}'^*$.

THEOREM 3.2. *If X^* is an L-random element in $\mathfrak{X}^*$ such that $E(X^*) = \theta^*$ (the null element of $\mathfrak{X}^*$), and X is any random element in $\mathfrak{X}$, and if A is $\mu \times \mu^*$-measurable and $E(|A|) < +\infty$, then $E(A) = 0$.*

This again follows immediately from Fubini's theorem.

Sufficient conditions for A to be $\mu \times \mu^*$-measurable and $E(|A|) < +\infty$ are

1) $\mathfrak{X}^*$ is separable (then $\|X^*\|$ is μ^*-measurable and $\mathfrak{X}$ is separable),
2) spheres of $\mathfrak{X}$ are μ-measurable (then $\|X\|$ is μ-measurable),
3) $E(\|X\|) < +\infty$ and $E(\|X^*\|) < +\infty$.

Theorems 3.1 and 3.2 have been given in [9], but the conditions of validity—conditions which allow us to apply Fubini's theorem—were not specified there. In fact, in [9], theorems 3.1 and 3.2 were used in a separable Hilbert space and in this case all the necessary conditions are fulfilled.

Let us now come back to the computation of $E(\langle g(T_j), X_j\rangle)$. With the previous notation and under the conditions stated above, $g(T_j)$ and X_j are independent random elements in $\mathfrak{X}^*$ and $\mathfrak{X}$ respectively, and X_j is an L-random element in $\mathfrak{X}$ such that $E(X_j) = \theta$ and $E(\|X_j\|^2) < +\infty$. Further, $\mathfrak{X}$ is a separable $\mathfrak{S}$-space and $\langle g(T_j), X_j\rangle$ is a continuous function of the X_j's and is therefore measurable. Moreover,

$$|\langle g(T_j), X_j\rangle| \leqq \|g(T_j)\| \cdot \|X_j\| = \|T_j\| \cdot \|X_j\|, \tag{3.11}$$

and, therefore,

$$E(|\langle g(T_j), X_j\rangle|) < +\infty . \tag{3.12}$$

Consequently, from theorem 3.1 we get

$$E(\langle g(T_j), X_j\rangle) = 0 . \tag{3.13}$$

In the same way we see that $E(\langle X^{*i}, X_j\rangle)$ exists, and

$$E(\langle X^{*i}, X_j\rangle) \leqq KE(\|X_j\|^2) . \tag{3.14}$$

The inequality (3.6) follows immediately.

4. Characteristic functions

It is well known that the characteristic function is a very useful tool in the case of real-valued random variables, particularly to study sums of independent random variables and to get central limit theorems. Kolmogoroff [8] gave the definition and some properties of the characteristic function of an L-random element in a Banach space, but at that time he did not carry the study of such elements very far. The definition of the characteristic function is as follows. Let X be an L-random element in a Banach space $\mathfrak{X}$ and x^* any real, continuous, linear functional on $\mathfrak{X}$. We call the *characteristic function* of X the function

$$(4.1) \qquad \varphi(x^*) = E(e^{i\langle x^*, X\rangle})$$

defined for all real $x^* \in \mathfrak{X}^*$. We now list the properties of this function.

Property I. With the strong topology in $\mathfrak{X}^*$, $\varphi(x^*)$ is a uniformly continuous function of x^*, and it is a continuous function of x^* with the weak topology in $\mathfrak{X}^*$.

Property II. If $E(X)$ and $E(\|x\|^2)$ exist, then

$$(4.2) \qquad \varphi(x^*) = 1 + i\langle x^*, E(X)\rangle - \tfrac{1}{2}E(|\langle x^*, X\rangle|^2) + \|x^*\|^2\omega(x^*),$$

where $\omega(x^*) \to 0$ as $\|x^*\| \to 0$.

Given any positive integer n, any $x_1^*, \cdots, x_n^* \in \mathfrak{X}^*$, and any B-measurable set E_n in n-dimensional Euclidean space we call a cylinder set, $\mathcal{E}_n$, the set of all $x \in \mathfrak{X}$ such that $(\langle x_1^*, x\rangle, \cdots, \langle x_n^*, x\rangle) \in E_n$. Let $\mathfrak{B}$ denote the smallest Borel field which contains the cylinder sets. If $\mathfrak{X}$ is separable, $\mathfrak{B}$ contains the spheres and all the open sets in $\mathfrak{X}$.

Property III. The characteristic function determines the L-measure on $\mathfrak{B}$.

Property IV. Any characteristic function $\varphi(x^*)$ is a positive definite function with $\varphi(\theta^*) = 1$. In the present context a positive definite function $\varphi(x^*)$ is a numerical-valued function such that

(a) it is continuous with the strong topology in $\mathfrak{X}^*$,

(b) $\sum_{j,k}^{n} \varphi(x_j^* - x_k^*)\alpha_j\bar{\alpha}_k$ is real and greater than or equal to 0 for any positive integer n, any $x_1^*, \cdots, x_n^* \in \mathfrak{X}^*$, and any complex numbers $\alpha_1, \cdots, \alpha_n$.

Property V (Generalization of Bochner's theorem). It is well known that, in the case of an ordinary random variable, the characteristic function is a positive definite function and that, conversely, *any* positive definite function such that $\varphi(0) = 1$ is the characteristic function of a random variable. This holds also for an n-dimensional random variable, for any positive integer n. If X is an L-random element in a Banach space $\mathfrak{X}$, property IV shows that the characteristic function of X is a positive definite function. But in this case it is *not* true that any positive definite function φ, with $\varphi(\theta^*) = 1$, is the characteristic function of an L-random element in $\mathfrak{X}$. And this is the case even if $\mathfrak{X}$ is a separable Hilbert space. If $\mathfrak{X}$ is a reflexive, separable Banach space, we have obtained a necessary and sufficient condition—condition C in [9]—for a positive definite function φ, with $\varphi(\theta^*) = 1$, to be the characteristic function of an L-random element. From condition C we can deduce the following one, which is more useful in practice. If $\mathfrak{X}$ is a reflexive, separable Banach space, and if $\varphi_1(x^*), \varphi_2(x^*), \cdots, \varphi_n(x^*), \cdots$ is an infinite sequence of characteristic functions such that

(a) there exist $a > 0$ and $s > 0$ such that for every positive integer n

$$E_n\,(\|X\|^a) < s^a \tag{4.3}$$

(b) $\varphi(x^*) = \lim_{n \to +\infty} \varphi_n(x^*)$ exists,

(c) there exists an A such that $|\varphi_n(x^*) - \varphi(x^*)| \to 0$ uniformly for all $x^* \in \mathfrak{X}^*$ for which $\|x^*\| \leq A$,

then $\varphi(x^*)$ is the characteristic function of an L-random element in $\mathfrak{X}$.

Property VI (Addition of independent L-random elements). If $\varphi_X(x^*)$; $\varphi_Y(x^*)$ are the characteristic functions of two independent L-random elements in $\mathfrak{X}$, the characteristic function of $X + Y$ is

$$\varphi_{X+Y}\,(x^*) = \varphi_X\,(x^*) \cdot \varphi_Y\,(x^*)\,. \tag{4.4}$$

Obviously the property extends to any given number of independent L-random elements.

A detailed proof of the above properties is given in [9].

5. Laplacian L-random elements; central limit theorems

5.1. *Laplacian L-random elements.* An L-random element X in a Banach space $\mathfrak{X}$ is called by M. Fréchet [4] (see also [9]) a *Laplacian* L-random element if $\langle x^*, X\rangle$ is a Laplacian random variable for all $x^* \in \mathfrak{X}^*$.

If $X_1, \cdots, X_n$ are independent Laplacian L-random elements in $\mathfrak{X}$, if x_0 is a given element in $\mathfrak{X}$, and if $a_0, a_1, \cdots, a_n$ are given numbers, then $Z = a_0x_0 + a_1X_1 + \cdots + a_nX_n$ is a Laplacian L-random element. Conversely, if $Z = a_1X_1 + a_2X_2$ is a Laplacian L-random element and if X_1 and X_2 are independent L-random elements, then X_1 and X_2 are Laplacian L-random elements. A necessary and sufficient condition that X be a Laplacian L-random element in $\mathfrak{X}$ is that there exists in $\mathfrak{X}$ an L-random element Y, independent of X, such that $X + Y$ and $X - Y$ are independent L-random elements.

The characteristic function of a Laplacian L-random element X is

$$\varphi\,(x^*) = \exp\{iE\,(\langle x^*, X\rangle) - \tfrac{1}{2}E\,[\langle x^*, X\rangle - E\,(\langle x^*, X\rangle)\,]^2\}. \tag{5.1}$$

And, conversely, if X is an L-random element whose characteristic function is (5.1), then it is Laplacian.

Different problems arise. If X is a Laplacian L-random element, then $E(\langle x^*, X\rangle)$ exists for all $x^* \in \mathfrak{X}^*$; this is a necessary but not a sufficient condition for the existence of $E(X)$. Then what about $E(X)$? And $E(\|X\|^2)$? If X is a Laplacian L-random element in a separable Hilbert space, then $E(X)$ and $E(\|X\|^2)$ do exist.

Replacing X by Y in (5.1) we see that it is possible to say that the characteristic function of a Laplacian L-random element X is a function of the form

$$f\,(x^*) = \exp\{iE\,(\langle x^*, Y\rangle) - \tfrac{1}{2}E\,[\langle x^*, Y\rangle - E\,(\langle x^*, Y\rangle)\,]^2\}, \tag{5.2}$$

where Y is an L-random element in $\mathfrak{X}$ such that (5.2) has a meaning. Conversely, for any L-random element Y in $\mathfrak{X}$ such that (5.2) has a meaning, the function $f(x^*)$ defined by (5.2) is obviously a positive definite function, and if it is a characteristic function it is the characteristic function of a Laplacian L-random element X, which, in general, is different from Y. A criterion is needed for knowing whether or not (5.2) is a characteristic function. We have obtained the following.

THEOREM 5.1. *If $\mathfrak{X}$ is a reflexive, separable $\mathfrak{S}$-space and if $E(\|Y\|^2) < +\infty$, then (5.2) is the characteristic function of a Laplacian L-random element X such that $E(\|X\|^2) < +\infty$. Since $\mathfrak{X}$ is separable, this implies furthermore that $E(X)$ exists.*

This theorem leads us to distinguish among the functions (5.2) those that are of a particular kind: for $\mathfrak{X}$ any Banach space, Y is an L-random element in $\mathfrak{X}$ such that $E(Y)$ exists and such that $E(\|Y\|^2)$ exists and is finite. Then we can suppose without loss of generality $E(Y) = \theta$. We shall call

$$f(x^*) = e^{-(1/2)E((x^*,\, Y))^2} \tag{5.3}$$

a *normal, positive definite function.*

5.2. *Central limit theorems.* Let $\mathfrak{X}$ be any Banach space and $X_1, \cdots, X_j, \cdots$ be an infinite sequence of independent L-random elements in $\mathfrak{X}$ with the *same* probability law such that $E(X_j) = \theta$ and $E(\|X_j\|^2) < +\infty$. Let us put

$$Z_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j . \tag{5.4}$$

THEOREM 5.2. *The characteristic function of Z_n tends, as $n \to +\infty$, to a limit which is a normal, positive definite function.*

THEOREM 5.3. *If $\mathfrak{X}$ is a reflexive, separable $\mathfrak{S}$-space, the characteristic function of Z_n tends, as $n \to +\infty$, to the characteristic function of a Laplacian L-random element Z such that $E(\|Z\|^2) < +\infty$.*

THEOREM 5.4. *If $\mathfrak{X}$ is a reflexive, separable $\mathfrak{S}$-space which has a basis, the distribution function of $\|Z_n\|$ tends, as $n \to +\infty$, to the distribution function of $\|Z\|$. More generally, if $f(x)$ is a numerical-valued function of $x \in \mathfrak{X}$, uniformly continuous in x (with the strong topology $\mathfrak{X}$) on any finite sphere of $\mathfrak{X}$, the distribution function of $f(Z_n)$ tends, as $n \to +\infty$, to the distribution function of $f(Z)$.*

Under some conditions these theorems may be extended to the more complicated case where the X_j's do not have the same probability law [3].

6. Applications

6.1. *Preliminary remarks.* The preceding results may be applied to several important problems. We have considered [2] the convergence of empirical probability distributions to theoretical distributions and used the central limit theorems in a Banach space to study the addition of independent random functions and functionals of random functions derived from a Poisson process [3]. We shall now give a new application of the theory of L-random elements in a Banach space, the study of the "statistical functions" of von Mises [10]. This application will be deduced from the following statements. Let $\mathfrak{X}$ be a separable Banach space and $X_1, \cdots, X_j, \cdots$ be an infinite sequence of independent L-random elements in $\mathfrak{X}$ with the same probability law. Let $\mathfrak{A}$ be any topological Hausdorf space and $a(x)$ be a function on $\mathfrak{X}$ to $\mathfrak{A}$.

THEOREM 6.1. *If $E(\|X_j\|) < +\infty$ with $E(X_j) = \theta$, and if $a(x)$ is continuous at $x = \theta$ (with the strong topology in $\mathfrak{X}$), then*

$$\lim_{n\to\infty} a\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) = a(\theta), \quad a.s. \tag{6.1}$$

This follows immediately from the strong law of large numbers in Banach spaces.

THEOREM 6.2. *Let $E(\|X_j\|^\alpha) < +\infty$ where $\alpha \geqq 1$, let $E(X_j) = \theta$, and let $\mathfrak{A}$ be a metric space in which we shall denote by $[a', a'']$ the distance between two elements a' and a''. Finally let there exist two positive numbers λ and M such that $[a(x), a(\theta)] \leqq M \cdot \|x\|^\lambda$. Then*

$$\lim_{n \to +\infty} E\left\{\left[a\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right),\ a(\theta)\right]^{\alpha/\lambda}\right\} = 0. \tag{6.2}$$

This theorem follows immediately from the law of large numbers in mean of order α in Banach spaces.

If $\mathfrak{X}$ is a separable $\mathfrak{G}$-space and if $\alpha = 2$, then we have more precisely

$$E\left\{\left[a\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right),\ a(\theta)\right]^{2/\lambda}\right\} < M^{2/\lambda} K E(\|X_j\|^2) \cdot \frac{1}{n}. \tag{6.3}$$

Let us suppose that we have such a case and that, moreover, $\mathfrak{A}$ is a reflexive, separable $\mathfrak{G}$-space which has a basis. Finally let us assume that the function $a(x)$ is differentiable at $x = \theta$. By this we mean that there exists a continuous linear function T on $\mathfrak{X}$ to $\mathfrak{A}$ and a positive numerical-valued function $\Omega(\alpha)$ on the positive number α, with $\lim_{\alpha \to 0+} \Omega(\alpha) = 0$, such that, if we put

$$a(x) - a(\theta) = T(x) + \|x\| \omega(x), \tag{6.4}$$

we have

$$\|\omega(x)\| \leqq \Omega(\|x\|). \tag{6.5}$$

Under these assumptions, if

$$\Delta_n = a\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) - a(\theta), \tag{6.6}$$

we can write

$$\sqrt{n}\Delta_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} T(X_j) + \left\|\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_j\right\| \omega\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right). \tag{6.7}$$

Almost surely $\left\|\frac{1}{n}\sum_{j=1}^{n} X_j\right\| \to 0$ as $n \to +\infty$; therefore,

$$\lim_{n \to +\infty} \left\|\omega\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)\right\| = 0, \text{ a.s.} \tag{6.8}$$

On the other hand, $E\left(\left\|(1/\sqrt{n})\sum_{j=1}^{n} X_j\right\|^2\right)$ is bounded [it is less than $KE(\|X_j\|^2)$]. If we put

$$R_n = \left\|\frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_j\right\| \omega\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right), \tag{6.9}$$

$$Z_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} T(X_j), \tag{6.10}$$

$\|R_n\|$ tends in probability to zero as $n \to +\infty$. Let $f(a)$ be a numerical-valued function defined on $\mathfrak{A}$, uniformly continuous (with the strong topology in $\mathfrak{A}$) on any finite sphere of $\mathfrak{A}$. From the central limit theorems we know that if $E(\|T(X_j)\|^2) < +\infty$ there exists a Laplacian L-random element Z in $\mathfrak{A}$ such that the distribution function of $f(Z_n)$ tends, as $n \to +\infty$, to the distribution function of $f(Z)$. Then, obviously, the distribution function of $f(\sqrt{n}\Delta_n) = f(Z_n + R_n)$ tends also, as $n \to +\infty$, to the distribution function of $f(Z)$.

6.2. *Application to von Mises' statistical functions.* Let us consider an infinite sequence of independent, numerical-valued random variables U_j, $j = 1, 2, \cdots$, with the same probability distribution $F(u)$. Without significant loss of generality we shall suppose that the U_j's are uniformly distributed on (0, 1), that is, that

$$F(u) = \begin{cases} 0 & \text{if } u \leqq 0 , \\ u & \text{if } 0 < u \leqq 1 , \\ 1 & \text{if } u > 1 . \end{cases} \tag{6.11}$$

Let us put

$$X_j(u) = \begin{cases} 0 & \text{if } u \leqq U_j , \\ 1 & \text{if } u > U_j , \end{cases} \tag{6.12}$$

and

$$F_n(u) = \frac{1}{n}\sum_{j=1}^{n} X_j(u) . \tag{6.13}$$

The $X_j(u)$'s and $F_n(u)$ may be considered as random elements in the space $L_\alpha(0, 1)$, which is a separable $\mathfrak{S}$-space if $\alpha \geqq 2$. On the other hand, let $\mathcal{M}$ be the space of distribution functions on (0, 1) obtained by putting the weight $1/k$ on each of any k points on (0, 1), where k is any positive integer. Here $F_n(u)$ is a random element in $\mathcal{M}$. Let $a(m)$ be a function on $\mathcal{M}$ to the space $\mathfrak{A}$ of real numbers. Von Mises [10] has considered, and called *statistical functions*, random variables which possess the property that, as $n \to +\infty$, a.s. $a[F_n(u)]$ tends to $a[F(u)]$ and $\sqrt{n}\{a[F_n(u)] - a[F(u)]\}$ is asymptotically a Laplacian random variable. Such results may be obtained immediately from the above theorems. For instance, let us consider the quantity

$$A_n = \frac{1}{n}\sum_{j=1}^{n}\Big(U_j - \frac{1}{n}\sum_{k=1}^{n} U_k\Big)^2 , \tag{6.14}$$

which may be used to estimate the variance of the U_j's. It appears as a statistical function when we write (6.14) in the following way,

$$\begin{aligned} A_n &= \int_0^1 \Big[u - \int_0^1 v\,dF_n(v)\Big]^2 dF_n(u) \\ &= \Big[\int_0^1 F_n(u)\,du\Big]^2 - 2\int_0^1 \Big[u - 1 + \int_0^1 F_n(v)\,dv\Big] F_n(u)\,du . \end{aligned} \tag{6.15}$$

The corresponding function $a(m)$ may be extended to a function $a(\varphi)$ of the space $L_2(0, 1)$ to the space $\mathfrak{A}$ by the formula

$$a(\varphi) = \Big[\int_0^1 \phi(u)\,du\Big]^2 - 2\int_0^1 \Big[u - 1 + \int_0^1 \varphi(v)\,dv\Big]\varphi(u)\,du , \tag{6.16}$$

where φ is any function in $L_2(0, 1)$. Let us look at the continuity and differentiability of $a(\varphi)$ at $\varphi = \varphi_0$, where $\varphi_0(u) = F(u) = u$ on $(0, 1)$. For this, let us compute $a(\varphi_0 + \varphi)$. We find

$$a(\varphi_0+\varphi) = \tfrac{1}{12} - 2\int^1 (u - \tfrac{1}{2})\, \varphi(u)\, du - 2\left[\int_0^1 \varphi(v)\, dv\right]^2 . \tag{6.17}$$

In (6.17) the first two terms on the right are linear and continuous with the strong topology in $L_2(0, 1)$, and the last term is less than or equal to $\|\varphi\|^2$ where $\|\varphi\|$ denotes the norm of φ considered as an element in $L_2(0, 1)$. Then, obviously, the function $a(\varphi)$ is differentiable at $\varphi = \varphi_0$ (in the sense indicated above), and we have

$$A_n - a(F) = a(F_n) - a(F) = a[\varphi_0 + (F_n - F)] - a(\varphi_0) . \tag{6.18}$$

It follows that, as $n \to +\infty$,

(a) a.s. $A_n \to a(F) = 1/12$,

(b) $\sqrt{n}[A_n - a(F)]$ is asymptotically Laplacian.

Of course, this example is given not for its own interest but to point out how the method works.

7. L^*-random elements in Banach spaces

In problems involving random elements in a Banach space it may happen that the Banach space is defined directly. This is the case, for instance, in the application to random functions mentioned above. But it may also happen that the Banach space is more conveniently defined as the dual space of another Banach space. This is the case, for instance, in the application mentioned above of the convergence of empirical probability distributions to theoretical distributions. The results given in the preceding sections concern the case where the Banach space is defined directly. When the Banach space is defined as the dual space of another, we can obviously apply these results, but it is more convenient to introduce somewhat different definitions.

Let $\mathfrak{X}^*$ be the dual space of a *separable* Banach space $\mathfrak{X}$; this does not imply that $\mathfrak{X}^*$ is separable. A random element in $\mathfrak{X}^*$ is a function $x^*(u) = X^*$ on the fundamental probability space $\mathcal{U}$ to $\mathfrak{X}^*$. We shall say that X^* is an L^*-random element if $\langle x^*(u), x\rangle$ is measurable for all given $x \in \mathfrak{X}$. It is natural to compare this definition with the definition of an L-random element in $\mathfrak{X}^*$. Let $\mathfrak{X}^{**}$ be the dual space of $\mathfrak{X}^*$, that is to say, the bidual of $\mathfrak{X}$; it is known that $\mathfrak{X} \subset \mathfrak{X}^{**}$, but generally $\mathfrak{X}$ is smaller than $\mathfrak{X}^{**}$. For x^{**} any given element in $\mathfrak{X}^{**}$, the above definition of an L^*-random element X^* means that $\langle x^{**}, X^*\rangle$ must be measurable for any fixed x^{**} belonging to $\mathfrak{X}$. In order that X^* be an L-random element in $\mathfrak{X}^*$ it would be necessary that $\langle x^{**}, X^*\rangle$ be measurable for all $x^{**} \in \mathfrak{X}^{**}$, even for x^{**} not belonging to $\mathfrak{X}$. Hence, under the present conditions (that $\mathfrak{X}^*$ is the dual of a separable Banach space), the definition of L^*-random elements is less restrictive than that of L-random elements. Obviously, the two definitions become equivalent if $\mathfrak{X} = \mathfrak{X}^{**}$, that is to say, when $\mathfrak{X}$, and therefore $\mathfrak{X}^*$, are reflexive.

The mathematical expectation of an L^*-random element in $\mathfrak{X}^*$ is the element $E(X^*) \in \mathfrak{X}^*$, if one exists (and then it is unique), such that

$$\langle E(X^*), x\rangle = E(\langle X^*, x\rangle) , \tag{7.1}$$

for all $x \in \mathfrak{X}$. We have proved that if an L-random element takes its values in a *separable* Banach space, then its norm is measurable and a sufficient condition for the existence of its mathematical expectation is that the mathematical expectation of its norm be finite. We shall prove now, *without any restriction* on $\mathfrak{X}^*$ (the dual space of a separable Banach space $\mathfrak{X}$), that the norm of an L^*-random element X^* is measurable. Indeed, let $x_1, \cdots, x_j, \cdots$ be a denumerable sequence dense on the sphere $\|x\| = 1$ in $\mathfrak{X}$, and let R be a positive number. The condition $\|X^*\| = \|x^*(u)\| < R$ is equivalent to

$$|\langle x^*(u), x_j\rangle| < R, \qquad j = 1, 2, \cdots. \tag{7.2}$$

Moreover, the condition $E(\|X^*\|) < +\infty$ implies that $E(\langle X^*, x\rangle)$ exists, since $|\langle X^*, x\rangle| \leqq \|X^*\| \cdot \|x\|$. Let us put

$$l(x) = E(\langle X^*, x\rangle); \tag{7.3}$$

then

$$|l(x)| \leqq E(\|X^*\|) \cdot \|x\|. \tag{7.4}$$

Hence $l(x)$ is a bounded, linear, numerical-valued function of x and there exists a linear functional $E(X^*) \in \mathfrak{X}^*$ such that

$$\langle E(X^*), x\rangle = l(x) = E(\langle X^*, x\rangle). \tag{7.5}$$

This proves that the condition $E(\|X^*\|) < +\infty$ implies that $E(X^*)$ exists and that

$$\|E(X^*)\| \leqq E(\|X^*\|). \tag{7.6}$$

This property enables us to avoid the assumption of separability on $\mathfrak{X}^*$. Let $X_1^*, \cdots, X_j^*, \cdots$ be a strictly stationary sequence of L^*-random elements in $\mathfrak{X}^*$ such that $E(\|X_j^*\|) < +\infty$. Let us consider

$$Z_n^* = \frac{1}{n}\sum_{j=1}^{n} X_j^*. \tag{7.7}$$

There exists an L^*-random element Z^* in $\mathfrak{X}^*$ such that $E(\|Z^*\|) < +\infty$ and such that almost surely Z_n^* tends weakly to Z^* as $n \to +\infty$. By this we mean that

$$\lim_{n\to\infty} (\langle Z_n^*, x\rangle) = \langle Z^*, x\rangle, \text{ a.s.} \tag{7.8}$$

for all $x \in \mathfrak{X}$. It is natural to ask whether it is possible to find strong laws of large numbers or laws in mean of order α, as in section 3, without assuming that $\mathfrak{X}^*$ is separable. Some counterexamples show that this is impossible.

By definition the characteristic function of an L^*-random element X^* is the function $\varphi(x)$ defined for all $x \in \mathfrak{X}$ by

$$\varphi(x) = E(e^{i\langle X^*, x\rangle}). \tag{7.9}$$

Here $\varphi(x)$ is a positive definite function, with $\varphi(\theta) = 1$, and defines a measure on a Borel field B^* in $\mathfrak{X}^*$. Conversely, a given positive definite function is the characteristic function of an L^*-random element in $\mathfrak{X}^*$ if a certain condition C*, analogous to the condition C of section 4, is fulfilled. But this condition C* does *not* assume either that $\mathfrak{X}^*$ is separable or that it is reflexive but merely that $\mathfrak{X}^*$ is the dual space of a separable Banach space.

These are the first results concerning this second point of view, which needs further development.

REFERENCES

[1] R. Fortet and E. Mourier, "Résultats complémentaires sur les éléments aléatoires prenant leurs valeurs dans un espace de Banach," *Bull. des Sci. Math.*, 2d Ser., Vol. 78 (1954), pp. 14–30.

[2] ———, "Convergence de la répartition empirique vers la répartition théorique," *Annales Sci. de l'École Normale Sup. de Paris*, Vol. 70 (1953), pp. 267–285.

[3] R. FORTET and E. MOURIER, "Les fonctions aléatoires comme éléments aléatoires dans des espaces de Banach," *Studia Mathematica*, Vol. 15 (1955), pp. 62–79.
[4] M. FRÉCHET, "Généralisations de la loi de probabilité de Laplace," *Ann. de l'Inst. Henri Poincaré*, Vol. 12 (1951), pp. 1–29.
[5] I. M. GELFAND, "Generalized stochastic processes," *Doklady Akad. Nauk SSSR*, Vol. 100 (1955), pp. 853–856.
[6] E. HILLE, *Functional Analysis and Semi-groups*, Amer. Math. Soc. Colloquium Publications, Vol. 31, New York, 1948.
[7] A. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse Mathematik, Berlin, Springer, 1933.
[8] ———, "La transformation de Laplace dans les espaces linéaires," *C. R. Acad. Sci., Paris*, Vol. 200 (1935), pp. 1717–1718.
[9] E. MOURIER, "Éléments aléatoires dans un espace de Banach," *Ann. de l'Inst. Henri Poincaré*, Vol. 13 (1953), pp. 161–244.
[10] R. VON MISES, "Les lois de probabilité pour les fonctions statistiques," *Ann. de l'Inst. Henri Poincaré*, Vol. 6 (1936), pp. 185–212.
[11] B. J. PETTIS, "On integration in vector space," *Trans. Amer. Math. Soc.*, Vol. 44 (1938), pp. 277–304.
[12] K. YOSIDA and S. KAKUTANI, "Operator-theoretical treatment of Markoff's process and mean ergodic theorem," *Annals of Math.*, Vol. 42 (1941), pp. 188–228.