REMARKS ON THE CONVERGENCE OF MEASURABLE SETS AND MEASURABLE FUNCTIONS *)

BY

E. MARCZEWSKI (WROCŁAW)

In the Measure Theory and the Probability Theory one deals with different kinds of convergences of a sequence of functions, in particular with the convergence almost everywhere (convergence a. e.) and the convergence in measure (convergence i. p.)[1].

Many logical connexions between them are well known, *e. g.*, convergence a. e. implies convergence i. p. and, more precisely, a sequence $\{f_n\}$ of measurable functions converges i. p. to f if and only if every subsequence of $\{f_n\}$ possesses a subsequence convergent a. e. to f.

The purpose of this paper is to investigate *when convergence i. p. implies convergence a. e.* or, respectively, convergence everywhere. (That is a development of an idea of Professor Josef Novák who has proved[2]) that, if a σ-measure defined in a σ-field of sets is strictly positive, then the convergence i. p. of a sequence of sets implies the convergence of this sequence in the sense of the General Theory of Sets, *i. e.* in the sense of the equality (2)). The answer to this question is easily obtained here by a simple analysis of the notion of *atom* of a measure.

Incidentally I treat some questions concerning the *stochastic independence* with respect to purely atomic measures.

1. Atomless and purely atomic measures. In Sections 1-3 and 5 I denote by μ a finite *σ-measure* (*i. e.* a countably additive and non negative set function) defined in a *σ-field* (*i. e.* countably additive and complementative class) $\boldsymbol{M}$ of subsets of a fixed abstract set X. Sets belonging to $\boldsymbol{M}$ are called *measurable*. Any real function f defined on X is called *measurable* if $f^{-1}(G)\in\boldsymbol{M}$ for every open set G.

A set $A\in\boldsymbol{M}$ is called an *atom* of μ, if $\mu(A)>0$ and if, moreover, the relations $A\supset B\in\boldsymbol{M}$ imply either $\mu(B)=0$ or $\mu(A)=\mu(B)$.

---

*) Presented to the Polish Mathematical Society, Wrocław Section, on January 30, 1953.

[1]) I write i. p. (=in probability) instead of i. m. (=in measure) to avoid confusion with the convergence in the mean.

[2]) In a conference *On the convergence of random events*, held in Wrocław, on December 15, 1952.

It is easy to see that

(i) *For every atom A, every measurable function is constant a. e. on A.*

If a set B contains no atom, then B and the measure μ on B are called *atomless*. *E. g.* the Lebesgue measure is atomless. It is easy to see that

(ii) *There is a decomposition of X into disjoint sets:*

$$X=X_0+X_1+X_2+\ldots,$$

where X_0 is either void or an atomless set of positive measure, and each of the sets $X_1, X_2, \ldots$ is either a void set or an atom.

This decomposition will be used several times in this paper. If X_0 is void, then μ is called *purely atomic*. *E. g.* any finite σ-measure defined in the field of all subsets of an almost denumerable set is purely atomic.

The following "intermediate value theorem" is valid for atomless measures[3]):

(iii) *If μ is atomless, then for every $B\in\boldsymbol{M}$ and every number c with $0<c<\mu(B)$ there is a set $C\in\boldsymbol{M}$ such that $B\supset C$ and $\mu(C)=c$.*

Theorem (iii) implies the following one:

(iv) *If μ is atomless and $\mu(X)=1$, then, for every sequence p_n with $0\leqslant p_n\leqslant 1$, there exists a sequence $\{E_n\}$ of stochastically independent sets with $\mu(E_n)=p_n$.*

To prove this, let

$$E^0=X-E, \qquad E^1=E \tag{1}$$

for every set $E\subset X$.

By (iii) there is a set E_1 with $\mu(E_1)=p_1$. If E_j are already defined for $j\leqslant n$, and if they are stochastically independent sets with $\mu(E_j)=p_j$, then there is, in view of (iii), a set E_{n+1} such that

$$\mu(E_1^{i_1}\cdot\ldots\cdot E_n^{i_n}E_{n+1})=p_{n+1}\mu(E_1^{i_1}\cdot\ldots\cdot E_n^{i_n})$$

for every system $i_1, i_2, \ldots, i_n$ of numbers 0 and 1. It is easy to see that $\{E_n\}$ is the required sequence.

Theorem (iv) is important for subsequent proofs.

2. Convergence in measure and convergence almost everywhere. We adopt the current definitions of these notions for sequences of measurable functions. Next, let us denote by c_E the characteristic function of the set E. We say that a sequence of sets $\{E_n\}$ *converges in measure (con-*

---

[3]) Cf. *e. g.* H. Hahn and A. Rosenthal, *Set Functions*, Albuquerque 1948, p. 51 and P. R. Halmos, *Measure Theory*, New York 1950, p. 174 (2).

verges i. p.) to E if the sequence $\{c_{E_n}\}$ converges i. p. to c_E, or, in other terms, if $\mu(E \dot{-} E_n) \to 0$ [4]).

We write, as usual,

$$\underline{\lim_n} E_n = \sum_{k=1}^{\infty} \prod_{n=k}^{\infty} E_n, \qquad \overline{\lim_n} E_n = \prod_{k=1}^{\infty} \sum_{n=k}^{\infty} E_n$$

and we say that a sequence $\{E_n\}$ of sets *converges to E* if and only if

$$E = \underline{\lim_n} E_n = \overline{\lim_n} E_n. \tag{2}$$

A sequence $\{E_n\}$ converges to E if and only if the sequence of functions $\{c_{E_n}\}$ converges to c_E everywhere.

We say that a sequence $\{E_n\}$ of sets *converges a. e. to E*, if the sequence of functions $\{c_{E_n}\}$ converges a. e. to c_E, or, in other terms, if

$$\mu(E \dot{-} \underline{\lim_n} E_n) = \mu(E \dot{-} \overline{\lim_n} E_n) = 0.$$

We shall prove that

(i) *The convergence i. p. of a sequence of measurable functions $\{f_n\}$ to a measurable function f implies the convergence of $\{f_n\}$ to f a. e. on every atom A of μ.*

It follows from 1(i), that there is an atom $A^* \subset A$ such that $\mu(A - A^*) = 0$ and that $f, f_1, f_2, \ldots$ are constant on A^*. Let

$$c = f(x), \qquad c_n = f_n(x) \quad \text{for} \quad x \in A^*.$$

Since f_n converge i. p. to f, there is, for every $\varepsilon > 0$, a number n_0 such that

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for} \quad n > n_0$$

outside a set Z_n with $\mu(Z_n) < \mu(A^*)$. Consequently

$$|c_n - c| < \varepsilon \quad \text{for} \quad n > n_0,$$

which implies convergence of f_n to f a. e. on A, *q. e. d.*

(ii) *If μ is atomless, then there is a sequence $\{E_n\}$ of measurable sets convergent i. p. to the void set and such that*

$$\underline{\lim_n} E_n = 0, \qquad \overline{\lim_n} E_n = X.$$

Without any loss of generality we may suppose $\mu(X) = 1$. By 1 (iv) there exists a sequence $\{B_n\}$ of Bernoulli trials (*i. e.* a sequence of stochastically independent sets with the same measure) with $\mu(B_n) = 1/2$. Denoting by B' the complement of B, the sequence

$$B_1, B_1', B_1B_2, B_1'B_2, B_1B_2', B_1'B_2', B_1B_2B_3, B_1'B_2B_3, \ldots$$

obviously satisfies the conditions of the thesis [5]).

It follows easily from (ii) that

(iii) *If convergence i. p. of measurable sets implies their convergence a. e., then μ is purely atomic.*

Combining (i), (ii) and (iii), we obtain

THEOREM I. *The following statements are equivalent:*

(a_1) *μ is purely atomic;*

(b_1) *convergence i. p. of measurable functions implies their convergence a. e.;*

(c_1) *convergence i. p. of measurable sets implies their convergence a. e.*

3. Strictly positive measures. If $\mu(N) = 0$ implies $N = 0$, then μ is called *strictly positive*.

(i) *If μ is strictly positive, then it is purely atomic* [6]).

Let us suppose that μ is not purely atomic, or, in other terms, $\mu(X_0) > 0$, and let us consider a σ-measure ν defined in X_0 as follows:

$$\nu(E) = \frac{1}{\mu(X_0)} \mu(E) \quad \text{for} \quad X_0 \supset E \in \boldsymbol{M}.$$

The measure ν is an atomless σ-measure in X_0 with $\nu(X_0) = 1$. Thus, by 1(iv), there exists a sequence $\{B_n\}$ of Bernoulli trials with $\nu(B_n) = 1/2$. By using the notation (1) we have

$$X_0 = \sum B_1^{i_1} B_2^{i_2} \ldots$$

where $\{i_n\}$ runs over all infinite sequences consisting of the numbers 0 and 1. Consequently there is a sequence $\{i_n\}$ such that $B_1^{i_1} B_2^{i_2} \ldots \neq 0$. It follows from the definition of Bernoulli trials that the measure ν vanishes for this intersection, whence the measures ν and μ are not strictly positive.

The following theorem is obvious:

THEOREM II. *The following statements are equivalent:*

(a_2) *μ is strictly positive;*

(b_2) *convergence a. e. of measurable functions implies their convergence everywhere;*

[4]) $A \dot{-} B$ denotes the symmetric difference of A and B. For this kind of convergence see e. g. Hahn and Rosenthal, l. c., p. 32.

[5]) This is a modification of a well-known construction. Cf. e. g. Halmos, l. c., p. 94 (6).

[6]) This theorem fails for Boolean σ-algebras. See section 4.

(c_2) *convergence a. e. of measurable sets implies their convergence.*

Combining this theorem with 3(i) and 2(i), we obtain

THEOREM III. *The following statements are equivalent:*

(a_2) *μ is strictly positive;*

(b_3) *convergence i. p. of measurable functions implies their convergence everywhere;*

(c_3) *convergence i. p. of measurable sets implies their convergence*[7]).

4. Convergences in Boolean algebras. If we replace the σ-field $\boldsymbol{M}$ of sets by a Boolean σ-algebra $\boldsymbol{B}$ with a finite σ-measure μ, we can also consider different kinds of convergences of sequences consisting of elements of $\boldsymbol{B}$.

Theorems 3(i) and III cannot be extended to Boolean algebras. In fact, the ordinary Lebesgue measure in the algebra $\boldsymbol{L}|\boldsymbol{N}$ of Lebesgue measurable sets modulo sets of measure zero is an atomless and strictly positive σ-measure. Obviously, in view of 2(ii), in this algebra convergence i. p. does not imply convergence (in the sense of the equality (2)).

Nevertheless, by using the same arguments as in the preceding sections, we obtain the following theorems for Boolean σ-algebras:

THEOREM I*. *μ is purely atomic if and only if convergence i. p. implies convergence a. e.*

THEOREM II*. *μ is strictly positive if and only if convergence a. e. implies convergence.*

THEOREM III*. *μ is purely atomic and strictly positive if and only if convergence i. p. implies convergence.*

5. Atoms and stochastic independence. In connection with Theorem I it seems interesting to investigate the properties of purely atomic measures. Theorem 1(iv) states that, for every atomless μ, there is a sequence of stochastically independent sets with prescribed measures, in particular a sequence of Bernoulli trials. In this section I shall prove that the existence of a sequence of sets having prescribed measures $\{p_n\}$ and stochastically independent with respect to a purely atomic measure depends on arithmetic properties of $\{p_n\}$. In particular there exists no sequence of Bernoulli trials with respect to a purely atomic measure (which follows directly from (i)).

Let us put for every number p with $0 \leqslant p \leqslant 1$:

$$p^{(0)} = 1-p, \qquad p^{(1)} = p,$$
$$p_* = \min(p, 1-p), \qquad p^* = \max(p, 1-p).$$

[7]) The implication (a_2)→(c_3) was proved by Professor Novak; cf. the introduction and the footnote [2]).

Obviously $p_* + p^* = 1$, and the conditions

$$\sum_{n=1}^{\infty} (p_n)_* = \infty, \qquad \prod_{n=1}^{\infty} p_n^* = 0 \tag{3}$$

are equivalent. The following conditions are equivalent too:

$$\sum_{n=1}^{\infty} (p_n)_* < \infty, \qquad \prod_{n=1}^{\infty} p_n^* > 0. \tag{4}$$

I shall prove

(i) *If $\mu(X)=1$, if the sets $E_n \epsilon \boldsymbol{M}$ are stochastically independent and if the sequence $p_n = \mu(E_n)$ satisfies* (3), *then μ is atomless.*

For every sequence $\{i_n\}$ consisting of numbers 0 and 1 we have:

$$\mu(E_1^{i_1} E_2^{i_2} \ldots) = \mu(E_1^{i_1})\mu(E_2^{i_2}) \ldots = p_1^{(i_1)} p_2^{(i_2)} \ldots,$$

whence, by (3), $\mu(E_1^{i_1} E_2^{i_2} \ldots) = 0$.

Let A be an atom of μ, and let us put $i_n = 0$ or $i_n = 1$ according as $\mu(AE_n) = 0$ or $\mu(A - E_n) = 0$. Then

$$\mu(A - E_1^{i_1} E_2^{i_2} \ldots) = 0$$

whence $\mu(A) = 0$.

(ii) *If $\mu(X)=1$, if $\{E_n\}$ is a sequence of sets such that $\boldsymbol{M}$ is the σ-field spanned by $\{E_n\}$ and if the sequence $p_n = \mu(E_n)$ satisfies* (4), *then μ is purely atomic.*

At first we shall prove that there exists a sequence $\{h_n\}$ consisting of numbers 0 and 1 such that

$$\mu(\underline{\lim}_n E_n^{h_n}) = 1. \tag{5}$$

Since (4), there is a sequence $\{g_n\}$ consisting of 0 and 1 and such that

$$\sum_{n=1}^{\infty} p_n^{(g_n)} < \infty.$$

Then, it suffices to consider the sequence $\{E_n^{g_n}\}$ and to make the following simple statement:

$$\text{if} \quad \sum_{n=1}^{\infty} \mu(M_n) < \infty, \qquad \text{then} \quad \mu(\overline{\lim}_n M_n) = 0 \tag{8}$$ [8]).

[8]) Cf. e. g. Halmos, l. c., p. 40 (6).

Then putting $h_n = 1 - g_n$, we obtain (5). Consequently, denoting by J the set of all sequences $\{j_n\}$ consisting of 0 and 1 and such that $j_n = h_n$ for sufficiently large n, we have

$$\mu\Big(\sum_{\{j_n\} \in J} E_1^{j_1} E_2^{j_2} \ldots\Big) = \mu(\underline{\lim}_n E_n^{h_n}) = 1.$$

Now, let us denote by J^* the set of all $\{j_n\} \in J$ such that $\mu(E_1^{j_1} E_2^{j_2} \ldots) > 0$. Since the set J is denumerable, the preceding equality implies the following one:

$$\mu\Big(\sum_{\{j_n\} \in J^*} E_1^{j_1} E_2^{j_2} \ldots\Big) = 1.$$

It follows from the hypotheses, that every set belonging to $\boldsymbol{M}$ is the sum of a family of disjoint sets of the form $E_1^{i_1} E_2^{i_2} \ldots$, whence every set $E_1^{j_1} E_2^{j_2} \ldots$ with $\{j_n\} \in J^*$ is an atom of μ. Therefore, in view of the preceding formula, μ is purely atomic.

Combining (i) and (ii) we obtain

THEOREM IV. *There is a purely atomic measure μ and a sequence of stochastically independent sets $\{E_n\}$ with $\mu(E_n) = p_n$, if and only if*

$$\sum_{n=1}^{\infty} (p_n)_* < \infty \quad (9)$$ [9]).

The necessity of this condition follows directly from (i). To prove the sufficiency, it suffices to consider the Lebesgue measure in the unit interval, to apply 1(iv) and to define $\boldsymbol{M}$ as the σ-field spanned by $\{E_n\}$. It follows from (ii) that the Lebesgue measure, considered on $\boldsymbol{M}$ only, is purely atomic.

[9]) This construction of a sequence of sets independent with respect to a purely atomic measure was found also by S. Zubrzycki.

ON THE AXIOMATIC TREATMENT OF PROBABILITY

BY

J. ŁOŚ (TORUŃ)

The calculus of probability is a branch of mathematics whose foundations have so far not been fully investigated. There are of course, many such branches, but the calculus of probability is unique among them as regards the specific course of the development of its fundamental principles. This is bound with what prof. Steinhaus calls the "tavern" origin of the calculus of probability. A theory of gambling games at first, it gradually extended its range of applicability, becoming finally a mathematical theory of great practical and theoretical importance.

It was at a very early stage of the development of the calculus of probability that mathematicians felt the need of formulating its foundations more precisely. The first attempt in this direction was probably the definiton of "classical probability" given by Laplace. However it was the introduction of axiomatic methods, which made it possible to investigate the principles of probability along new lines.

The first axiomatic of probability was given by Bohlmann [2] about the year 1904. Since that time there have appeared (and still appear) numerous axiomatics, suggesting new methods of treatment or — more frequently — distorting treatments already known by means of the terminology which they adopt.

In principle it is the aim of every axiomatic of the calculus of probability to answer the following two questions:

1° What are events, *i. e.* what are those objects supposed to be probable?

2° What kind of function of events should probability be?

Rather a paradoxical point of view could be ventured, namely, that the answers to the above-mentioned questions should not be given by probabilists. The first should be answered by algebraists and the other by real function specialists.

And even if it were not true, experience shows that certain parts of algebra (lattice theory, and especially the theory of Boolean algebras) and certain parts of the theory of functions (measure theory) control the foundations of probability to such an extent that they almost absorb them. This is a useful process of complete mathematization of the calculus of probability.