REMARKS ON THE CONVERGENCE OF MEASURABLE SETS AND MEASURABLE FUNCTIONS

BY

E. MARCZEWSKI (WROCLAW)

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.)

Many logical connections 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\(^1\) that, if \(\sigma\)-measure defined in a \(\sigma\)-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 similar analysis of the notion of atom of a measure.

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

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

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

\(^1\) Presented to the Polish Mathematical Society, Wroclaw Section, on January 30, 1953.

\(^2\) I write i.p. [=in probability] instead of i. m. [=in measure] to avoid confusion with the convergence in the mean.

\(^3\) In a conference On the convergence of random events, held in Wroclaw, on December 18, 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 \(\mu\) 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:

\[X=X_1+X_2+X_3+\ldots,\]

where \(X_3\) 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 \(X_3\) is void, then \(\mu\) is called purely atomic. E.g. any finite \(\sigma\)-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: 2

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

Theorem (iii) implies the following one:

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

To prove this, let

\[E^c=X-E, \quad E=E\]

for every set \(E\).

By (ii) there is a set \(E_1\) with \(\mu(E_1)=p_1\). If \(E_i\) are already defined for \(i\leq n\), and if they are stochastically independent sets with \(\mu(E_i)=p_i\), then there is, in view of (iii), a set \(E_{n+1}\) such that

\[\mu(E_{n+1}|E_1,\ldots,E_n)=p_{n+1}\]

for every system \(s_1, s_2, \ldots, s_n\) 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 \(\mathcal{P}\) the characteristic function of the set \(E\). We say that a sequence of sets \(\{E_n\}\) converges in measure (converges almost everywhere) if

\[\lim_{n\to\infty} \mu(E_n \Delta E) = 0\]

or

\[\lim_{n\to\infty} \mathcal{P}(E_n \Delta E) = 0\]

(see e.g. H. Hahn and R. Rosenthal, Set Functions, Albuquerque 1948, p. 51 and P. Halmos, Measure Theory, New York 1950, p. 174 (3)).
serves i.p. to \( E \) if the sequence \([E_n]\) converges i.p. to \( c_E \), or, in other terms, if \( \mu(E - E_n) \to 0 \).\(^4\)

We write, as usual,
\[
\lim_{n \to \infty} E_n = \bigcup_{n=1}^{\infty} E_n, \quad \lim_{n \to \infty} E_n = \bigcap_{n=1}^{\infty} E_n
\]
and we say that a sequence \([E_n]\) of sets converges to \( E \) if and only if
\[
E = \lim_{n \to \infty} E_n - \lim_{n \to \infty} E_n.
\]

A sequence \([E_n]\) converges to \( E \) if and only if the sequence of functions \([c_{E_n}]\) converges a.e. 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
\[
\mu(E - \lim_{n \to \infty} E_n) = \mu(E - \lim_{n \to \infty} E_n) = 0.
\]

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 \( \mu \).

It follows from (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
\[
e_{n} = f_{n}(x), \quad e_{n} = f(x) \quad \text{for } x \in A^*.
\]

Since \( f_n \) converge i.p. to \( f \), there is, for every \( \varepsilon > 0 \), a number \( n_\varepsilon \) such that
\[
|f_n(x) - f(x)| < \varepsilon \quad \text{for } n > n_\varepsilon
\]
outside a set \( Z_\varepsilon \) with \( \mu(Z_\varepsilon) < \mu(A^*) \). Consequently
\[
|e_n - e| < \varepsilon \quad \text{for } n > n_\varepsilon
\]
which implies convergence of \( e_n \) to \( f \) a.e. on \( A \), q.e.d.

(ii) If \( \mu \) is atomless, then there is a sequence \([E_n]\) of measurable sets convergent i.p. to the void set and such that
\[
\lim_{n \to \infty} E_n = 0, \quad \lim_{n \to \infty} E_n = X.
\]

Without any loss of generality we may suppose \( \mu(X) = 1 \). By (i) there exists a sequence \([E_n]\) of Bernoulli trials (i.e. a sequence of a.e.

\(^4\) \( A - B \) denotes the symmetric difference of \( A \) and \( B \). For this kind of \( \mu \) a.e. see e.g. Hahn and Rosenthal, i.e., p. 32.

... 

The following theorem is obvious:

**Theorem II.** The following statements are equivalent:
- (a) \( \mu \) is strictly positive,
- (b) \( \mu \) is strictly positive,
- (c) convergence a.e. of measurable functions implies their convergence everywhere.

\(^4\) This is a modification of a well-known construction. Cf. e.g. Halmos, I.e., p. 84 (6).

\(^5\) This theorem fails for Boolean \( \sigma \)-algebras. See section 4.
(c) convergence a. e. of measurable sets implies their convergence.

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

**Theorem III.** The following statements are equivalent:

(a) \( \mu \) is strictly positive;

(b) convergence t. p. of measurable functions implies their convergence everywhere;

(c) convergence t. p. of measurable sets implies their convergence.  

4. **Convergences in Boolean algebras.** If we replace the \( \sigma \)-field \( M \) of sets by a Boolean \( \sigma \)-algebra \( B \) with a finite \( \sigma \)-measure \( \mu \), we can also consider different kinds of convergences of sequences consisting of elements of \( B \).

Theorems 3(i) and III cannot be extended to Boolean algebras. In fact, the ordinary Lebesgue measure in the algebra \( L(N) \) of Lebesgue measureable sets modulo sets of measure zero is an atomless and strictly positive \( \sigma \)-measure. Obviously, in view of 2(ii), in this algebra convergence t. 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 \( \sigma \)-algebras:

**Theorem I.** \( \mu \) is purely atomic if and only if convergence t. p. implies convergence a.e.

**Theorem II.** \( \mu \) is strictly positive if and only if convergence a.e. implies convergence.

**Theorem III.** \( \mu \) is purely atomic and strictly positive if and only if convergence t. 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 \( \mu \), 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 \( 0 \leq p \leq 1 \):

\[ p^+ = \inf (p, 1-p), \quad p^0 = p, \quad p^* = \max (p, 1-p). \]

--\)

Obviously \( p^* + p^0 = 1 \), and the conditions

\[ \sum_{n=1}^{\infty} p_n^* = \infty, \quad \prod_{n=1}^{\infty} p_n^* = 0 \]

are equivalent. The following conditions are equivalent too:

\[ \sum_{n=1}^{\infty} p_n^* < \infty, \quad \prod_{n=1}^{\infty} p_n^* > 0. \]

I shall prove

(i) If \( \mu (X) = 1 \), if the set \( E_n \in M \) are stochastically independent and if the sequence \( p_n = \mu (E_n) \) satisfies (3), then \( \mu \) is atomless.

For every sequence \( \{ \xi_n \} \) consisting of numbers 0 and 1 we have:

\[ \mu (\xi_0^1 \xi_1^0 \xi_2^1 \ldots) = \mu (\xi_0^1) \mu (\xi_1^0) \mu (\xi_2^1) \ldots = p_1^0 p_2^0 p_3^1 \ldots, \]

whence, by (3), \( \mu (E_1^0 E_2^0 E_3^1 \ldots) = 0 \).

Let \( A \) be an atom of \( \mu \) and let us put \( \xi_n = 0 \) or \( \xi_n = 1 \) according as

\[ \mu (\bar{A} E_n) = 0 \quad \text{or} \quad \mu (A E_n) = 0. \]

Then

\[ \mu (A - E_1^0 E_2^0 E_3^1 \ldots) = 0. \]

whence \( \mu (A) = 0 \).

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

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

\[ \mu (\lim \bar{E}_n) = 1. \]

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

\[ \sum_{n=1}^{\infty} p_{g_n}^* < \infty. \]

Then, it suffices to consider the sequence \( \{ E_n^h \} \) and to make the following simple assertion:

\[ \text{if} \sum_{n=1}^{\infty} \mu (M_n) < \infty, \quad \text{then} \mu (\lim \bar{M}_n) = 0. \]

--\)

\[ \text{Cf. e. g. Halmos, I. e., p. 40 (8).} \]
Then putting \( b_n = 1 - g_n \), we obtain (5). Consequently, denoting by \( J \) the set of all sequences \( [x_n] \) consisting of 0 and 1 and such that \( j_n = b_n \) for sufficiently large \( n \), we have

\[
\mu(\bigoplus_{O \in J} E^O_1 \cdots) = \mu(\limsup_{n \to \infty} E^n_n) = 1.
\]

Now, let us denote by \( J' \) the set of all \([x_n] \in J\) such that \( \mu(E^O_1 E^O_2 \cdots) > 0 \). Since the set \( J \) is denumerable, the preceding equality implies the following one:

\[
\mu(\bigoplus_{O \in J'} E^O_1 E^O_2 \cdots) = 1.
\]

It follows from the hypotheses, that every set belonging to \( M \) is the sum of a family of disjoint sets of the form \( E^O_1 E^O_2 \cdots \), whence every set \( E^O_1 E^O_2 \cdots \) with \( [x_n] \in J' \) is an atom of \( \mu \). Therefore, in view of the preceding formula, \( \mu \) is purely atomic.

Combining (i) and (ii) we obtain

**Theorem IV.** There is a purely atomic measure \( \mu \) and a sequence of stochastically independent sets \( [E_n] \) with \( \mu(E_n) = p_n \), if and only if

\[
\sum_{n=1}^{\infty} p_n < \infty.
\]

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 \( M \) as the \( \sigma \)-field spanned by \([E_n]\). It follows from (ii) that the Lebesgue measure, considered on \( M \) only, is purely atomic.

---

2) This construction of a sequence of sets independent with respect to a purely atomic measure was found also by S. Zborowsky.

---

**ON THE AXIOMATIC TREATMENT OF PROBABILITY**

BY

J. LOŠ (TORUŚ)

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 definition 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 probability. 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 mathematicalization of the calculus of probability.