ON MIXING SEQUENCES OF SETS

By

A. RÉNYI (Budapest), member of the Academy

Introduction

Let $[\Omega, \mathcal{A}, \mu]$ be a measure space. By other words, let Ω be an arbitrary abstract set, $\mathcal{A}$ a σ-algebra of subsets of Ω and $\mu(A)$ $(A \in \mathcal{A})$ a measure defined in Ω and on $\mathcal{A}$. We shall denote the elements of $\mathcal{A}$ by capital letters $A, B, C, \ldots$. The elements of Ω will be denoted by ω. We denote by $A+B$ the union and by AB the intersection of the sets A and B.

We shall call a sequence A_n $(n=0, 1, \ldots)$ of measurable sets *strongly mixing with density* α if for any $B \in \mathcal{A}$, such that $\mu(B) < +\infty$, we have

$$\lim_{n\to\infty} \mu(A_n B) = \alpha \mu(B) \tag{1}$$

where $0 < \alpha < 1$ and the value of α does not depend on B.

Evidently, in the case when $\mu(\Omega) < +\infty$, we have, choosing in (1) $B = \Omega$,

$$\lim_{n\to\infty} \mu(A_n) = \alpha \mu(\Omega). \tag{2}$$

Thus if $\mu(\Omega) < +\infty$, (1) can also be written in the form

$$\lim_{n\to\infty} \mu(A_n B) = \frac{\mu(B)}{\mu(\Omega)} \lim_{n\to\infty} \mu(A_n). \tag{3}$$

The term "strongly mixing" has been chosen in accordance with the well-known definition of a strongly mixing measure preserving transformation of a measure space in ergodic theory (see [1], [2]). As a matter of fact, if T is a measurable transformation of the measure space $[\Omega, \mathcal{A}, \mu]$ preserving the measure μ and $\mu(\Omega) < +\infty$, then T is called strongly mixing if for any $A \in \mathcal{A}$ and $B \in \mathcal{A}$ we have

$$\lim_{n\to\infty} \mu(T^{-n} A \cdot B) = \frac{\mu(A)\mu(B)}{\mu(\Omega)}. \tag{4}$$

Taking into account that in this case $\mu(T^{-n} A) = \mu(A)$ $(n=0, 1, \ldots)$ and using the terminology introduced above, we may say that the measure preserving transformation T is strongly mixing if and only if for any $A \in \mathcal{A}$ the sequence $A_n = T^{-n} A$ $(n=0, 1, \ldots)$ is strongly mixing with density $\dfrac{\mu(A)}{\mu(\Omega)}$.

This is, however, a very special way of obtaining strongly mixing sequences of sets, as will be seen from the examples given below.

The notion of strongly mixing sequences of sets is especially important in probability theory. In the present paper we shall mostly deal with these applications, and therefore we suppose in general that the measure space considered is a probability space, i. e. $\mu(\Omega)=1$. To avoid misunderstandings we shall denote probability measures by $\mathbf{P}$ (or $\mathbf{Q}$). If $[\Omega, \mathcal{A}, \mathbf{P}]$ is a probability space, then, as usual, the elements of $\mathcal{A}$ will be called *events*. Thus a sequence A_n $(n=0,1,\dots)$ of events will be called strongly mixing with density α if for any event $B \in \mathcal{A}$ we have

$$\lim_{n\to\infty} \mathbf{P}(A_n B) = \alpha \mathbf{P}(B) \tag{5}$$

where $0<\alpha<1$. As (5) is trivially satisfied (with every value of α) for any sequence A_n of events if the event B has probability 0, it suffices to suppose that (5) holds if $\mathbf{P}(B)>0$. By using the usual notation $\mathbf{P}(A|B)$ for the conditional probability of the event A with respect to the event B, defined in the case $\mathbf{P}(B)>0$ by

$$\mathbf{P}(A|B)=\frac{\mathbf{P}(AB)}{\mathbf{P}(B)}, \tag{6}$$

we may write (5) in the following equivalent form:

$$\lim_{n\to\infty} \mathbf{P}(A_n|B)=\alpha \tag{5*}$$

for every event B for which $\mathbf{P}(B)>0$. Thus *a sequence A_n $(n=0,1,\dots)$ of events is strongly mixing with density α $(0<\alpha<1)$ if* (5*) *is satisfied for every B which has a positive probability.*

It is easy to show that the following theorem[1] holds:

THEOREM 1. *If $[\Omega, \mathcal{A}, \mathbf{P}]$ is a probability space and the sequence A_n $(n=0,1,\dots)$ of events is strongly mixing with density α, then*

$$\lim_{n\to\infty} \mathbf{Q}(A_n)=\alpha \tag{7}$$

holds for any probability measure $\mathbf{Q}$ in Ω and on $\mathcal{A}$ which is absolutely continuous with respect to the measure $\mathbf{P}$.

PROOF. By the Radon—Nikodym theorem there exists a non-negative and measurable function $\chi(\omega)$ on Ω which is integrable with respect to the

[1] This theorem is, of course, known. (It has been used e. g. implicitly in [3].) We state it here for the sake of reference, as we did not find it explicitly formulated in the literature. For the same reason we sketch its simple proof.

measure $\mathbf{P}$ and is such that for any $A \in \mathfrak{A}$ we have

$$\mathbf{Q}(A) = \int_A \chi(\omega)\, d\mathbf{P}. \tag{8}$$

Clearly, (7) holds if $\chi(\omega)$ is a step function (i. e. if $\chi(\omega)$ takes on only a finite number of different values). As to any integrable χ and any $\varepsilon > 0$ there can be found a step function χ_1 such that $\int |\chi(\omega) - \chi_1(\omega)|\, d\mathbf{P} < \varepsilon$, it follows easily that (7) holds in the general case too.

Thus Theorem 1 is proved.

In § 1 of the present paper we shall give the following necessary and sufficient condition for a sequence of events being strongly mixing:

THEOREM 2. *The sequence A_n of events, such that $A_0 = \Omega$ and $\mathbf{P}(A_n) > 0$ $(n = 1, 2, \ldots)$,[2] is strongly mixing with density α if (and only if)*

$$\lim_{n\to\infty} \mathbf{P}(A_n | A_k) = \alpha \tag{9}$$

for $k = 0, 1, \ldots$ where $0 < \alpha < 1$ and α does not depend on k.

Thus the strongly mixing property of the sequence A_n depends on the relative positions of the sets A_n only.

Theorem 2 will be proved in § 1 by means of Lemma 1, relating to sequences of elements of an arbitrary Hilbert space.

Theorem 2 is fairly general and when applied to different types of sequences of events, leads to some interesting special cases. One of these is the following:

THEOREM 4. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of independent random variables on the probability space $[\Omega, \mathfrak{A}, \mathbf{P}]$ and let us suppose that there can be found a sequence C_n of real numbers and another sequence D_n of positive numbers such that $\lim_{n\to\infty} D_n = +\infty$, further a distribution function $F(x)$ such that putting $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$ $(n = 1, 2, \ldots)$ we have*

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\zeta_n - C_n}{D_n} < x\right) = F(x) \tag{10}$$

for every real x which is a point of continuity of the distribution function $F(x)$. Let $\mathbf{Q}$ be an arbitrary probability measure in Ω and on $\mathfrak{A}$ which is

[2] The supposition $\mathbf{P}(A_n) > 0$ for every $n \geqq 1$ is made only to make a simple formulation of our result possible; it is not an essential restriction. As a matter of fact, according to the definition, in a strongly mixing sequence of events there can occur only a finite number of events having the probability 0, and these may be omitted as the strongly mixing character of a sequence of events is not influenced by the change of a finite number of elements of the sequence. The condition $A_0 = \Omega$ is not a restriction either; it has been supposed only to include the condition $\lim_{n\to\infty} \mathbf{P}(A_n) = \alpha$ into (9).

absolutely continuous with respect to **P**. *Then we have*

$$\lim_{n\to\infty} \mathbf{Q}\left(\frac{\zeta_n - C_n}{D_n} < x\right) = F(x) \tag{11}$$

in every point of continuity x of $F(x)$.

Thus the fact that the distribution of $\frac{\zeta_n - C_n}{D_n}$ tends to a limiting distribution as well as this limiting distribution itself, are invariant against the change of the underlying probability measure, provided that this measure **P** is replaced by a probability measure **Q** which is absolutely continuous with respect to **P**.

Note that with respect to the measure **Q** the random variables ξ_n are, in general, not independent. Thus Theorem 3 may be considered as a result extending the validity of the limit theorems of probability theory, valid for independent random variables, to certain sequences of "almost independent" random variables.

The first result of the type of Theorem 4 has been given by the author of the present paper in [3] where there were two restrictions: it has been supposed that the random variables ξ_n have discrete distributions and that the probability space $[\Omega, \mathcal{A}, \mathbf{P}]$ is isomorphic to the probability space for which Ω is the interval (0, 1) and **P** the ordinary Lebesgue measure. In a subsequent paper [4] A. N. KOLMOGOROV has proved a more general result. He dropped the supposition that the variables ξ_n are discrete, and concerning the probability space he supposed that the measure **P** is perfect (for the definition of perfect measures see [5]). Theorem 4 does not contain any restriction concerning the probability space, thus it is more general than the result of KOLMOGOROV mentioned above. It has been pointed out by E. MARCZEWSKI (oral communication) that Theorem 4 can be deduced also from certain results of E. SPARRE-ANDERSEN and B. JESSEN [6]. P. RÉVÉSZ (oral communication) has shown that Theorem 3 can be proved also by using certain limit theorems of J. L. DOOB on martingales [7]. However, the proof given in § 2 of the present paper, which shows that Theorem 4 is a special case of Theorem 2, is in some sense the most natural approach. As a matter of fact, Theorem 2 is a source of a large number of similar results which can not all be obtained by the other methods mentioned. We can obtain e. g. by means of Theorem 2 results similar to Theorem 4 for general Markov chains instead of the special Markov chains formed by partial sums of independent random variables.[3]

[3] This question will be discussed in a forthcoming joint paper of P. RÉVÉSZ and the author.

Theorem 2 when applied to ergodic theory leads to a criterion (Theorem 3) for a measure preserving transformation defined on the measure space $[\Omega, \mathcal{A}, \mu]$ being strongly mixing.

In § 3 we consider weakly mixing sequences of sets and events, respectively, and obtain similar results as in the case of strong mixing.

My thanks are due to Mr. P. RÉVÉSZ for his valuable remarks which I utilized in preparing the present paper.

§ 1. A criterion for the strongly mixing property of a sequence of events

Let $\mathcal{H}$ be an arbitrary Hilbert space. We denote the elements of $\mathcal{H}$ by small letters (e. g. f, g). The inner product of the elements f and g will be denoted by (f, g) and the norm $(f, f)^{1/2}$ of f by $\|f\|$. We first prove the following

LEMMA 1. *Let f_n $(n=0, 1, \ldots)$ be a sequence of elements of a Hilbert space $\mathcal{H}$. Let us suppose that*

$$(1.1) \qquad \|f_n\| \leqq K \qquad (n=0, 1, \ldots)$$

where K is a positive constant not depending on n. Let us suppose further that for any $k=0, 1, \ldots$ we have

$$(1.2) \qquad \lim_{n\to\infty} (f_k, f_n) = 0.$$

Then for any $g \in \mathcal{H}$ we have

$$(1.3) \qquad \lim_{n\to\infty} (g, f_n) = 0.$$

PROOF. Let us denote by $\mathcal{H}_1$ the least subspace of $\mathcal{H}$ which contains the elements $f_0, f_1, \ldots, f_n, \ldots$. Clearly, (1.3) holds if g is a finite linear combination of the elements $f_0, f_1, \ldots, f_n, \ldots$, i. e. if $g = \sum_{k=1}^{n} c_k f_k$. It follows that (1.3) holds also if g is an arbitrary element of $\mathcal{H}_1$, because in this case for any $\varepsilon > 0$ there exists a finite linear combination $g_1 = \sum_{k=1}^{n} c_k f_k$ such that $\|g - g_1\| < \varepsilon$ which implies that

$$(1.4) \qquad |(g, f_n) - (g_1, f_n)| \leqq K\varepsilon$$

and thus

$$(1.5) \qquad \limsup_{n\to\infty} |(g, f_n)| \leqq K\varepsilon.$$

As $\varepsilon > 0$ is arbitrary, (1.5) implies (1.3). Now, let $\mathcal{H}_2$ denote the set of those elements of $\mathcal{H}$ which are orthogonal to every element of the sequence

f_n. Clearly, (1.3) holds if $g \in \mathcal{H}_2$. As by a well-known theorem (see [8], p. 8) every $g \in \mathcal{H}$ can be represented in the form $g = g_1 + g_2$ where $g_1 \in \mathcal{H}_1$ and $g_2 \in \mathcal{H}_2$, it follows[4] that (1.3) holds for any $g \in \mathcal{H}$. Thus our lemma is proved.[5]

It should be mentioned that our lemma contains as a special case the well-known fact that if $\{f_n\}$ is an orthonormal system, the Fourier coefficients (g, f_n) of an arbitrary $g \in \mathcal{H}$ are tending to 0 for $n \to \infty$. This fact is usually proved by means of BESSEL's inequality which gives, of course, much more. The corresponding stronger result under the supposition (1.2) will be given in a forthcoming paper.

We now deduce from our lemma

THEOREM 2. *Let* $[\Omega, \mathcal{A}, \mathbf{P}]$ *be a probability space. Let* A_n $(n = 0, 1, \ldots)$ *be a sequence of events such that* $A_0 = \Omega$ *and* $\mathbf{P}(A_n) > 0$ $(n = 1, 2, \ldots)$. *The sequence* A_n *of events is strongly mixing with density* α $(0 < \alpha < 1)$ *if (and only if)*

$$\lim_{n \to \infty} \mathbf{P}(A_n | A_k) = \alpha \tag{1.6}$$

for $k = 0, 1, \ldots$.

PROOF OF THEOREM 2. Let $\mathcal{H}$ denote the Hilbert space of all real random variables $\xi = \xi(\omega)$ $(\omega \in \Omega)$ such that $\int_\Omega \xi^2 d\mathbf{P}$ exists. Let us define the inner product by $(\xi, \eta) = \int_\Omega \xi \eta \, d\mathbf{P}$ and, correspondingly, the norm by $\|\xi\| = \left(\int_\Omega \xi^2 d\mathbf{P}\right)^{1/2}$.

Let the random variables $\alpha_n = \alpha_n(\omega)$ be defined as follows:

$$\alpha_n(\omega) = \begin{cases} 1 - \alpha & \text{if} \quad \omega \in A_n, \\ -\alpha & \text{if} \quad \omega \in \bar{A}_n \end{cases} \qquad (n = 0, 1, \ldots).$$

Then we have

$$(\alpha_k, \alpha_n) = \mathbf{P}(A_k A_n) - \alpha \mathbf{P}(A_k) - \alpha \mathbf{P}(A_n) + \alpha^2. \tag{1.7}$$

As $A_0 = \Omega$, it follows from (1.6) that

$$\lim_{n \to \infty} \mathbf{P}(A_n) = \alpha. \tag{1.8}$$

Further it follows from (1.6) that

$$\lim_{n \to \infty} \mathbf{P}(A_n A_k) = \alpha \mathbf{P}(A_k) \qquad (k = 1, 2, \ldots). \tag{1.9}$$

[4] B. SZ.-NAGY kindly called my attention to the fact that the idea of the above proof of our lemma is the same as that of the standard proof of the theorem (see e. g. [8], p. 10) that if f_n is an arbitrary sequence of elements of $\mathcal{H}$ such that $\|f_n\|$ is bounded, then there exists a subsequence of the sequence f_n which converges weakly to an element f of $\mathcal{H}$.

[5] Another proof of Lemma 1 has been found independently by P. RÉVÉSZ.

From (1. 7), (1. 8) and (1. 9) we obtain

(1. 10) $$\lim_{n\to\infty} (\alpha_k, \alpha_n) = 0 \qquad (k=0, 1, \ldots).$$

Taking into account that

(1. 11) $$\|\alpha_n\|^2 = (1-\alpha)^2 \mathbf{P}(A_n) + \alpha^2 (1-\mathbf{P}(A_n)) \leqq 1,$$

we see that the sequence α_n satisfies the conditions of Lemma 1. Thus we have for any $g \in \mathcal{H}$

(1. 12) $$\lim_{n\to\infty} (g, \alpha_n) = 0.$$

Choosing for $g = g(\omega)$ the random variable defined by

(1. 13) $$g(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in B, \\ 0 & \text{if} \quad \omega \bar{\in} B, \end{cases}$$

we have

(1. 14) $$(g, \alpha_n) = \mathbf{P}(A_n B) - \alpha \, \mathbf{P}(B)$$

and thus by virtue of (1. 12) we obtain, provided that $\mathbf{P}(B) > 0$,

(1. 15) $$\lim_{n\to\infty} \mathbf{P}(A_n | B) = \alpha.$$

Thus Theorem 2 is proved.

By combining Theorem 2 with Theorem 1 it follows[6] that if $\mathbf{Q}$ is any probability measure in Ω and on $\mathcal{A}$ which is absolutely continuous with respect to $\mathbf{P}$, and A_n satisfies the conditions of Theorem 2, then we have

(1. 16) $$\lim_{n\to\infty} \mathbf{Q}(A_n) = \alpha.$$

Let us now consider the application of Theorem 2 to ergodic theory. Let $[\Omega, \mathcal{A}, \mu]$ be a measure space and suppose that $\mu(\Omega) < +\infty$. A measure preserving (not necessarily one-to-one) transformation T of this space is called strongly mixing if, denoting by $T^{-1}A$ the inverse image of the set A (i. e. the set of those $\omega \in \Omega$ for which $T\omega \in A$) and defining $T^{-n}A$ by the recursion $T^{-n}A = T^{-(n-1)}(T^{-1}A)$ $(n=2, 3, \ldots)$, we have for any $A \in \mathcal{A}$ and $B \in \mathcal{A}$

(1. 17) $$\lim_{n\to\infty} \mu(T^{-n}A \cdot B) = \frac{\mu(A)\mu(B)}{\mu(\Omega)}.$$

[6] It should be mentioned that (1.16) could be deduced directly from the above proof of Theorem 2 without making use of Theorem 1. As a matter of fact, let $\chi = \chi(\omega)$ be a function, the existence of which is ensured by the Radon—Nikodym theorem, such that

$$\mathbf{Q}(A) = \int_A \chi(\omega)\, d\mathbf{P} \quad \text{for} \quad A \in \mathcal{A}.$$

If χ belongs to $\mathcal{H}$, i. e. if $\int_\Omega \chi^2 d\mathbf{P}$ exists, then applying (1. 12) to this function χ we directly obtain (1. 16). The general case follows by remarking that to any integrable random variable χ and any $\varepsilon > 0$ there can be found a $\chi_1 \in \mathcal{H}$ such that $\int |\chi - \chi_1| \, d\mathbf{P} < \varepsilon$.

Applying Theorem 2 to the sequence $A_n = T^{-n}A$ of sets of the probability space $[\Omega, \mathcal{A}, \mathbf{P}]$ where $\mathbf{P}(A) = \frac{\mu(A)}{\mu(\Omega)}$ and taking into account that by virtue of the supposed measure preserving property of the transformation T

$$\mu(T^{-n}A \cdot T^{-k}A) = \mu(T^{-k}(T^{-(n-k)}A \cdot A)) = \mu(T^{-(n-k)}A \cdot A) \tag{1.18}$$

for $n \geqq k$, we obtain the following

THEOREM 3. *Let $[\Omega, \mathcal{A}, \mu]$ be a measure space, $\mu(\Omega) < +\infty$ and T a measure preserving transformation of Ω onto itself. A necessary and sufficient condition for T being strongly mixing (i. e. for the validity of* (1. 17)*) is that for any $A \in \mathcal{A}$ with $\mu(A) > 0$ we should have*

$$\lim_{n\to\infty} \mu(T^{-n}A \cdot A) = \frac{\mu^2(A)}{\mu(\Omega)}. \tag{1.19}$$

By other words, if (1. 17) is valid for $B = A$, it is always valid.

§ 2. The invariance of the limiting distribution of sums of independent random variables

In this § we prove

THEOREM 4. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of independent random variables defined on the probability space $[\Omega, \mathcal{A}, \mathbf{P}]$. Let us suppose that there can be found a sequence C_n of real numbers and another sequence D_n of positive numbers for which $\lim_{n\to\infty} D_n = +\infty$, further a distribution function $F(x)$ such that putting $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$ we have*

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{\zeta_n - C_n}{D_n} < x\right) = F(x) \tag{2.1}$$

in every point of continuity x of the distribution function $F(x)$. Let $\mathbf{Q}$ be an arbitrary measure in Ω and on $\mathcal{A}$ which is absolutely continuous with respect to $\mathbf{P}$. Then we have

$$\lim_{n\to\infty} \mathbf{Q}\left(\frac{\zeta_n - C_n}{D_n} < x\right) = F(x) \tag{2.2}$$

if x is any point of continuity of $F(x)$.

PROOF OF THEOREM 4. Let x be a point of continuity of $F(x)$ and $F(x) > 0$. Clearly, it suffices to consider such values of x. In this case evidently $\mathbf{P}\left(\frac{\zeta_n - C_n}{D_n} < x\right) > 0$ for $n > n_0$. Let us put $A_0 = \Omega$ and denote by A_n the event that the inequality $\frac{\zeta_{n+n_0} - C_{n+n_0}}{D_{n+n_0}} < x$ takes place $(n = 1, 2, \ldots)$.

According to Theorems 1 and 2, if we show that condition (1.6) is fulfilled for these events, (2.2) follows. Let us put $\zeta_n^* = \frac{\zeta_n - C_n}{D_n}$. Thus it suffices to prove that

$$\lim_{n\to\infty} \mathbf{P}(\zeta_n^* < x \mid \zeta_k^* < x) = F(x) \tag{2.3}$$

for any $k > n_0$.

Now we need the following simple lemma (see e. g. [9], p. 254):

LEMMA 2. *If ϑ_n and ε_n are random variables such that $\lim_{n\to\infty} \mathbf{P}(\vartheta_n < x) = F(x)$ in every point of continuity x of the distribution function $F(x)$, and $\lim_{n\to\infty} \mathbf{P}(|\varepsilon_n| \geqq \delta) = 0$ for any $\delta > 0$, then we have $\lim_{n\to\infty} \mathbf{P}(\vartheta_n + \varepsilon_n < x) = F(x)$ in any point of continuity of $F(x)$.*

Applying Lemma 2 to $\vartheta_n = \zeta_n^*$ and $\varepsilon_n = -\frac{\zeta_k}{D_n}$, it follows from (2.1) that

$$\lim_{n\to\infty} \mathbf{P}\left(\zeta_n^* - \frac{\zeta_k}{D_n} < x\right) = F(x). \tag{2.4}$$

As $\zeta_n^* - \frac{\zeta_k}{D_n} = \frac{\zeta_n - \zeta_k - C_n}{D_n}$ is independent of ζ_k^*, we have

$$\mathbf{P}\left(\zeta_n^* - \frac{\zeta_k}{D_n} < x \,\middle|\, \zeta_k^* < x\right) = \mathbf{P}\left(\zeta_n^* - \frac{\zeta_k}{D_n} < x\right) \tag{2.5}$$

and thus from (2.4)

$$\lim_{n\to\infty} \mathbf{P}\left(\zeta_n^* - \frac{\zeta_k}{D_n} < x \,\middle|\, \zeta_k^* < x\right) = F(x). \tag{2.6}$$

Applying again Lemma 2 to the random variables $\vartheta_n = \zeta_n^* - \frac{\zeta_k}{D_n}$ and $\varepsilon_n = \frac{\zeta_k}{D_n}$ on the probability space $[\Omega, \mathcal{A}, \mathbf{P}']$ where $\mathbf{P}'(A) = \mathbf{P}(A \mid \zeta_k^* < x)$, we obtain (2.3). Thus Theorem 3 is proved.

Let us call a sequence η_n $(n = 1, 2, \ldots)$ of random variables *a mixing sequence with the limiting distribution function* $F(x)$ if for every $B \in \mathcal{A}$ with $\mathbf{P}(B) > 0$ and for every real x which is a point of continuity of $F(x)$ we have

$$\lim_{n\to\infty} \mathbf{P}(\eta_n < x \mid B) = F(x). \tag{2.7}$$

The assertion of Theorem 3 can be expressed by saying that if the random variables ξ_n are independent and putting $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$ the random variables $\zeta_n^* = \frac{\zeta_n - C_n}{D_n}$, where $D_n \to \infty$, have the limiting distribution function $F(x)$, then ζ_n^* *is a mixing sequence of random variables with the limiting distribution function* $F(x)$.

Mixing sequences of random variables have remarkable properties. For instance, if η_n is a mixing sequence of random variables, then η_n is in the limit independent of any random variable ϑ. As a matter of fact, if $\mathbf{P}(\vartheta < y) > 0$, we have

$$(2.8)\quad \lim_{n\to\infty} \mathbf{P}(\eta_n < x, \vartheta < y) = \mathbf{P}(\vartheta < y) \lim_{n\to\infty} \mathbf{P}(\eta_n < x \mid \vartheta < y) = \mathbf{P}(\vartheta < y) F(x).$$

Thus we obtain the following consequence of Theorem 4:

COROLLARY 1. *If the random variables ξ_n are independent, $\zeta_n = \xi_1 + \xi_2 + \cdots + \xi_n$, and there can be found sequences of real numbers C_n and $D_n > 0$ such that $D_n \to +\infty$ and $\zeta_n^* = \dfrac{\zeta_n - C_n}{D_n}$ has the limiting distribution $F(x)$, then ζ_n^* is in the limit independent of any random variable.*

Another interesting property of mixing sequences of random variables with a non-degenerate limiting distribution is that they can not be stochastically convergent to some random variable. As a matter of fact, let us suppose in contrary to our statement that η_n is a mixing sequence of random variables with the non-degenerate limiting distribution $F(x)$, and that η_n tends stochastically to the random variable η_∞, i. e. for any $\delta > 0$ we have

$$\lim_{n\to\infty} \mathbf{P}(|\eta_n - \eta_\infty| \geqq \delta) = 0.$$

Then evidently by Lemma 2

$$(2.9)\qquad \mathbf{P}(\eta_\infty < x) = \lim_{n\to\infty} \mathbf{P}(\eta_n < x) = F(x),$$

further by Theorem 4 and Lemma 2

$$(2.10)\qquad \mathbf{P}(\eta_\infty < x, \eta_k < y) = \lim_{n\to\infty} \mathbf{P}(\eta_n < x, \eta_k < y) = F(x)\,\mathbf{P}(\eta_k < y),$$

and therefore, applying again Theorem 3 and Lemma 2,

$$(2.11)\qquad \mathbf{P}(\eta_\infty < x, \eta_\infty < y) = \lim_{k\to\infty} \mathbf{P}(\eta_\infty < x, \eta_k < y) = F(x) F(y).$$

Thus η_∞ would be independent of itself which is clearly impossible, as by (2.9) and the supposition that $F(x)$ is a non-degenerate distribution, η_∞ is not a constant.

Thus we obtain the following

COROLLARY 2. *If $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ are independent random variables, further there can be found real sequences C_n and $D_n > 0$ with $D_n \to +\infty$ such that putting $\zeta_n = \xi_1 + \cdots + \xi_n$ and $\zeta_n^* = \dfrac{\zeta_n - C_n}{D_n}$ the limiting distribution of ζ_n^* exists and is non-degenerate, then the random variables ζ_n^* can not converge stochastically to a random variable.*

A special case of Corollary 2 has been mentioned in the textbook on probability theory of the author ([10], p. 534, Exercise 21).

Let us consider an example. Let

$$t=\sum_{n=1}^{\infty}\frac{\varepsilon_n(t)}{2^n} \tag{2.12}$$

be the dyadic expansion of the real number t $(0<t<1)$ where each $\varepsilon_n(t)$ is equal to 0 or 1. The functions $\varepsilon_n(x)$ may be considered as random variables on the probability space $[\Omega, \mathcal{A}, \mathbf{P}]$, where Ω is the interval $(0, 1)$, $\mathcal{A}$ the set of all Lebesgue measurable subsets of Ω and $\mathbf{P}$ the ordinary Lebesgue measure. The random variables $\varepsilon_n(t)$ are clearly independent and each takes on the values 0 and 1 with probability $\frac{1}{2}$. It follows by the Moivre—Laplace theorem that putting

$$S_n(t)=\varepsilon_1(t)+\cdots+\varepsilon_n(t) \tag{2.13}$$

we have

$$\lim_{n\to\infty}\mathbf{P}\left(\frac{S_n(t)-\frac{n}{2}}{\frac{1}{2}\sqrt{n}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du. \tag{2.14}$$

Our Theorem 4 gives in this case the following result: If $\mathbf{Q}$ is any probability measure defined on the Lebesgue measurable subsets of the interval $(0, 1)$, which is absolutely continuous with respect to the Lebesgue measure, i. e. if

$$\mathbf{Q}(A)=\int_A q(t)\,dt \tag{2.15}$$

where $q(t)\geqq 0$ and $\int_0^1 q(t)\,dt=1$, then we have

$$\lim_{n\to\infty}\mathbf{Q}\left(\frac{S_n(t)-\frac{n}{2}}{\frac{1}{2}\sqrt{n}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du. \tag{2.16}$$

Let us define the set $E_n(x)$ as the set of those t's $(0<t<1)$ for which $\dfrac{S_n(t)-\frac{n}{2}}{\frac{1}{2}\sqrt{n}}<x$ $(n=1, 2, \ldots)$. Then, clearly, $E_n(x)$ is a strongly mixing sequence of sets. This example gives some idea about the structure of strongly mixing sequences of sets, as the sets $E_n(x)$ can easily be constructed.

§ 3. Weak mixing

A measure preserving transformation T of the measure space $[\Omega, \mathfrak{A}, \mu]$ with $\mu(\Omega) < +\infty$ is called weakly mixing (see [1], [2]) if for any $A \in \mathfrak{A}$ and $B \in \mathfrak{A}$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \left| \mu(T^{-n}A\cdot B) - \frac{\mu(A)\mu(B)}{\mu(\Omega)} \right| = 0, \tag{3.1}$$

i. e. if $\mu(T^{-n}A\cdot B)$ is strongly $(C, 1)$-summable to the limit $\frac{\mu(A)\mu(B)}{\mu(\Omega)}$. Generalizing this notion, we shall say that *the sequence A_n $(n=0, 1, \ldots)$ of sets is weakly mixing with density α $(0<\alpha<1)$ if for every $B \in \mathfrak{A}$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu(A_n B) - \alpha\mu(B)| = 0. \tag{3.2}$$

A sequence of sets in a probability space which is weakly mixing with density α, will correspondingly be called a weakly mixing sequence of events (with density α). By the same method as used in the preceding §§ we can obtain analogous results for weak mixing.

The analogue of Theorem 1 runs as follows:

THEOREM 5. *If the sequence A_n $(n=0, 1, \ldots)$ of events of the probability space $[\Omega, \mathfrak{A}, \mathbf{P}]$ is weakly mixing with density α $(0<\alpha<1)$, we have for any probability measure $\mathbf{Q}$ in Ω and on $\mathfrak{A}$, which is absolutely continuous with respect to $\mathbf{P}$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mathbf{Q}(A_n) - \alpha| = 0. \tag{3.3}$$

The proof of Theorem 5 runs along the same lines as that of Theorem 1. Instead of Lemma 1 we need the following analogous

LEMMA 3. *Let f_n $(n=0, 1, \ldots)$ be a sequence of elements of the Hilbert space $\mathcal{H}$. Let us suppose that*

$$\frac{1}{N} \sum_{n=0}^{N-1} \|f_n\| \leq K \qquad (N=1, 2, \ldots), \tag{3.4}$$

further that for any $k=0, 1, \ldots$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |(f_k, f_n)| = 0. \tag{3.5}$$

Then we have for any $g \in \mathcal{H}$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |(g, f_n)| = 0. \tag{3.6}$$

PROOF OF LEMMA 3. Clearly, (3.6) holds if g is a finite linear combination of the elements of the sequence f_n, because

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|\left(\sum_{k=0}^{r} c_k f_k, f_n\right)\right| \leqq \sum_{k=0}^{r}|c_k|\frac{1}{N}\sum_{n=0}^{N-1}|(f_k, f_n)|.$$

If $\mathcal{H}_1$ is the least subspace of $\mathcal{H}$ containing the sequence f_n and $g \in \mathcal{H}_1$, we may find for any $\varepsilon > 0$ coefficients $c_0, c_1, \ldots, c_r$ such that putting $g_1 = \sum_{k=0}^{r} c_k f_k$ we have $\|g - g_1\| < \varepsilon$. It follows that

$$\big||(g, f_n)| - |(g_1, f_n)|\big| \leqq |(g, f_n) - (g_1, f_n)| \leqq \varepsilon \|f_n\|$$

and thus

$$\left|\frac{1}{N}\sum_{n=0}^{N-1}|(g, f_n)| - \frac{1}{N}\sum_{n=0}^{N-1}|(g_1, f_n)|\right| \leqq K\varepsilon$$

what proves that (3.6) holds for any $g \in \mathcal{H}_1$. But if we denote again by $\mathcal{H}_2$ the set of those elements $\mathcal{H}$ which are orthogonal to every f_n, (3.6) evidently holds for $g \in \mathcal{H}_2$ too, and as every $g \in \mathcal{H}$ can be represented in the form $g = g_1 + g_2$ with $g_1 \in \mathcal{H}_1$ and $g_2 \in \mathcal{H}_2$, it follows that (3.6) holds for every $g \in \mathcal{H}$. Thus Lemma 3 is proved. From Lemma 3 we may deduce the following result which is analogous to Theorem 2:

THEOREM 6. *The sequence A_n $(n = 0, 1, \ldots)$ of events belonging to the probability space $[\Omega, \mathcal{A}, \mathbf{P}]$, for which $A_0 = \Omega$ and $\mathbf{P}(A_n) > 0$ $(n = 1, 2, \ldots)$, is weakly mixing with density α $(0 < \alpha < 1)$ if (and only if) we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}|\mathbf{P}(A_n|A_k) - \alpha| = 0 \quad \text{for} \quad k = 0, 1, \ldots. \tag{3.7}$$

The analogue of Theorem 3 for weakly mixing transformations may be stated as follows:

THEOREM 7. *The measure preserving transformation T of the measure space $[\Omega, \mathcal{A}, \mu]$, for which $\mu(\Omega) < +\infty$, is weakly mixing if for any $A \in \mathcal{A}$ for which $\mu(A) > 0$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\left|\mu(T^{-n}A\cdot A) - \frac{\mu^2(A)}{\mu(\Omega)}\right| = 0. \tag{3.8}$$

The analogue of Theorem 4 for weakly mixing sequences of events is evidently also valid.

(Received 6 February 1958)

References

[1] E. Hopf, *Ergodentheorie*, Ergebnisse der Math., V. 2 (Berlin, 1937), p. 36.
[2] P. R. Halmos, *Lectures on ergodic theory*, The Mathematical Society of Japan, (1956), p. 36.
[3] A. Rényi, Contributions to the theory of independent random variables (in Russian, with Summary in English), *Acta Math. Acad. Sci. Hung.*, **1** (1950), pp. 99—108.
[4] A. H. Колмогоров, Теорема о сходимости условных математических ожиданий и некоторые ее применения, Comptes Rendus du Premier Congrès des Mathématiciens Hongrois (Budapest, 1952), pp. 367—386.
[5] B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables* (Cambridge, 1954).
[6] E. Sparre-Andersen and B. Jessen, Some limit theorems on integrals in an abstract set, *Kgl. Danske Videnskabernes Selskab*, **22** (1946), pp. 3—29.
[7] J. L. Doob, *Stochastic processes* (New York, 1953).
[8] B. Sz.-Nagy, *Spektraldarstellung linearer Transformationen des Hilbertschen Raumes*, Ergebnisse der Math., V. 5 (Berlin, 1942).
[9] H. Cramér, *Mathematical methods of statistics* (Princeton, 1946).
[10] A. Rényi, *Valószínűségszámítás* (Budapest, 1954).