

A study of sequences of equivalent events as special stable sequences

By A. RÉNYI and P. RÉVÉSZ (Budapest)

Introduction

Let $\{\Omega, S, \mathbf{P}\}$ be a probability space and $A_1, A_2, \dots$ be a finite or infinite sequence of events (i. e. $A_1, A_2, \dots$ are subsets of Ω belonging to the σ -algebra S) in this space. The events A_n are called equivalent (or symmetrically dependent, see [1]) if the probability of the event¹⁾ $A_{i_1}A_{i_2}\dots A_{i_k}$ ($i_j \neq i_l$ if $j \neq l$) depends only on k and it does not depend on the indices $i_1, i_2, \dots, i_k$. The simplest examples of sequences of equivalent events are the sequences of mutually independent events having the same probability. A more general example is the following: Let $A_n(t)$ ($0 \leq t \leq 1$; $n=1, 2, \dots$) for every fixed t be a sequence of independent events, such that $\mathbf{P}(A_n(t))=t$ ($n=1, 2, \dots$) and let $\lambda(\omega)$ be a random variable with values in the interval $[0, 1]$. Then $A_n(\lambda)$ is a sequence of equivalent events, provided that $A_n(\lambda) = \bigcup_{0 \leq t \leq 1} (A_n(t) \cap \{\lambda(\omega)=t\})$ are events ($n=1, 2, \dots$) i. e. they are elements of S .

One can ask which sequences of equivalent events can be represented in this form. In connection with this question an important result is due to DE FINETTI [2]. (See also [3], [4] and for generalisations [5].) It is the following:

Theorem 1. *Let $A_1, A_2, \dots$ be an infinite sequence of equivalent events and put*

$$\mathbf{P}(A_1 A_2 \dots A_k) = \omega_k.$$

Then there exists a distribution function $F(x)$ such that

$$F(0)=0, \quad F(1+0) = 1$$

and

$$\omega_k = \int_0^{1+0} x^k dF(x) \quad (k=1, 2, \dots).$$

In view of this theorem a very natural conjecture is the following: every infinite sequence of equivalent events can be represented in the mentioned form; or with other words: if $A_1, A_2, \dots$ is an infinite sequence of equivalent events, then there exists a random variable $\lambda(\omega)$ ($0 \leq \lambda(\omega) \leq 1$) such that the events $A_1, A_2, \dots$ are

¹⁾ Here and in what follows the product of events denotes the joint occurrence of these events (that is the intersection of the corresponding sets).

under the condition $\lambda(\omega)=x$ independent, and have the probability x , i. e.

$$\mathbf{P}(A_{i_1}A_{i_2}\dots A_{i_k}|\lambda(\omega)=x) = x^k \text{ (with probability 1).}$$

In § 2. of this paper we prove the above mentioned conjecture (Theorem 2.). Before doing this in § 1 we give a new proof of DE FINETTI's theorem. § 3. contains some remarks on equivalent random variables.

§ 1. A new proof of de Finetti's theorem

We recall some definitions and results of paper [6]. A sequence of events $A_1, A_2, \dots$ is called a *stable* sequence if

$$(1.1) \quad \lim_{n \rightarrow \infty} \mathbf{P}(A_n B) = \mathbf{Q}(B)$$

exists for every event $B \in S$. It is easy to see that $\mathbf{Q}(B)$ is a measure which is absolutely continuous with respect to $\mathbf{P}$. Let the Radon—Nikodym derivative of $\mathbf{Q}$ with respect to $\mathbf{P}$ be $\lambda(\omega)$, i. e.

$$\mathbf{Q}(B) = \int_B \lambda(\omega) d\mathbf{P}, \quad \text{for } B \in S.$$

The random variable $\lambda(\omega)$ is called the *local density* of the stable sequence $A_1, A_2, \dots$. Clearly $0 \leq \lambda(\omega) \leq 1$.

In [6] it is proved that if $\{A_n\}$ is a stable sequence of events with the indicator functions $\alpha_n(\omega)$, i. e.

$$\alpha_n(\omega) = \begin{cases} 1 & \text{if } \omega \in A_n \\ 0 & \text{if } \omega \in \overline{A_n} \end{cases}$$

then α_n converges to λ weakly i. e. for every element g of the Hilbert-space $L^2_{\mathbf{P}}(\Omega)$

$$(1.2) \quad \lim_{n \rightarrow \infty} (g, \alpha_n) = \lim_{n \rightarrow \infty} \int_{\Omega} g \alpha_n d\mathbf{P} = \int_{\Omega} g \lambda d\mathbf{P} = (g, \lambda).$$

If λ is constant (with probability 1) the stable sequence $A_1, A_2, \dots$ is called *mixing* (see [10]). It is shown further in [6] that in order that a sequence $\{A_n\}$ should be stable it is sufficient that the limit

$$\lim_{n \rightarrow \infty} \mathbf{P}(A_k A_n) = \mathbf{Q}(A_k)$$

should exist for each $k = 1, 2, \dots$. It follows evidently that every sequence of equivalent events is stable. Using these facts a very simple proof of DE FINETTI's theorem can be given.

THE PROOF OF DE FINETTI'S THEOREM. Let the local density of the sequence $A_1, A_2, \dots$ of equivalent events be $\lambda(\omega)$, and we denote the indicator function of A_n by α_n . Then we have

$$\omega_k = \mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k}) = \int_{\Omega} \alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_k} d\mathbf{P} = (\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{k-1}}, \alpha_{i_k}).$$

Thus by (1.2)

$$\omega_k = \lim_{i_k \rightarrow \infty} (\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{k-1}}, \alpha_{i_k}) = (\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{k-1}}, \lambda) = (\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{k-2}}, \lambda, \alpha_{i_{k-1}}).$$

Applying the same argument again, we obtain

$$\omega_k = \lim_{i_{k-1} \rightarrow \infty} (\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{k-2}}, \lambda, \alpha_{i_{k-1}}) = (\alpha_{i_1} \alpha_{i_2} \dots \alpha_{i_{k-2}}, \lambda, \lambda).$$

Applying the same argument again $k-2$ times, we obtain that

$$(1.3) \quad \omega_k = \mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k}) = \int_{\Omega} \lambda^k(\omega) d\mathbf{P} = \int_0^1 x^k dF_{\lambda}(x)$$

where $F_{\lambda}(x)$ is the distribution function of $\lambda(\omega)$. Thus the theorem of DE FINETTI is proved.

As a matter of fact we have proved somewhat more, namely that for any $k \geq 1$, for any set of k different positive integers $i_1, i_2, \dots, i_k$ and for any r with $0 \leq r \leq k$ we have

$$(1.4) \quad \mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k}) = \int_{\Omega A_{i_1} A_{i_2} \dots A_{i_r}} \lambda^{k-r}(\omega) d\mathbf{P}.$$

Clearly (1.4) reduces for $r=0$ to (1.3), while in the other extreme case $r=k$ (1.4) is trivial. We shall need (1.4) in the course of the proof of Theorem 2 in § 2.

Now we show how to any distribution function $F(x)$ such that $F(0)=0$ and $F(1+0)=1$ a sequence $\{A_n\}$ of equivalent events can be constructed such that

$$\mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k}) = \omega_k = \int_0^1 x^k dF(x)$$

for any $k=1, 2, \dots$ and any set of different integers $i_1, i_2, \dots, i_k$. Such a sequence $\{A_n\}$ can easily be constructed in any nonatomic probability space, by using the evident fact that the sequence $\{\omega_k\}$ is absolutely monotonic, i. e. (putting $\omega_0=1$)

$$\Delta^r \omega_n = \sum_{j=0}^r \binom{r}{j} (-1)^j \omega_{n+j} = \int_0^1 x^n (1-x)^r dF(x) \geq 0, \quad \text{for } r \geq 0, n \geq 0.$$

However, our example gives much more, as it exhibits explicitly the local density $\lambda(\omega)$ and shows that the events A_n are in fact conditionally independent under the condition $\lambda(\omega)=x$ and have the probability x , for any x with $0 < x < 1$.

Let the probability space Ω be the unit square of the plane, i. e.

$$\Omega = I_1 \times I_2$$

where I_1 and I_2 are unit intervals. Let the probability measure $\mathbf{P}$ on Ω be the product measure

$$\mathbf{P} = \mu_1 \times \mu_2$$

where μ_1 is the Lebesgue—Stieltjes measure on I_1 defined by the distribution function $F(x)$ and μ_2 is the ordinary Lebesgue-measure on I_2 .

To define the events A_n we need first to define a set of polynomials

$$p_k^{(n)}(x) \quad (k=0, 1, 2, \dots, 2^n; n=1, 2, \dots)$$

as follows: $p_0^{(n)}(x) \equiv 0$, and

$$p_{k+1}^{(n)}(x) = \sum_{j=0}^k x^{\alpha_n(j)} (1-x)^{\beta_n(j)} \quad \text{for } k=0, 1, \dots, 2^n-1$$

where $\alpha_n(j)$ resp. $\beta_n(j)$ denote the number of zeros resp. ones in the dyadic expansion of $j/2^n$; more exactly if

$$\frac{j}{2^n} = \sum_{i=1}^n \frac{\varepsilon_i}{2^i} \quad \text{where } \varepsilon_i \text{ is 0 or 1}$$

then

$$\beta_n(j) = \sum_{i=1}^n \varepsilon_i \quad \text{and} \quad \alpha_n(j) = n - \beta_n(j).$$

Thus — for instance —

$$\begin{aligned} p_0^{(3)}(x) &\equiv 0 & p_4^{(3)}(x) &= x^3 + 2x^2(1-x) + x(1-x)^2 \\ p_1^{(3)}(x) &= x^3 & p_5^{(3)}(x) &= x^3 + 3x^2(1-x) + x(1-x)^2 \\ p_2^{(3)}(x) &= x^3 + x^2(1-x) & p_6^{(3)}(x) &= x^3 + 3x^2(1-x) + 2x(1-x)^2 \\ p_3^{(3)}(x) &= x^3 + 2x^2(1-x) & p_7^{(3)}(x) &= x^3 + 3x^2(1-x) + 3x(1-x)^2 \\ p_8^{(3)}(x) &= x^3 + 3x^2(1-x) + 3x(1-x)^2 + (1-x)^3 \equiv 1. \end{aligned}$$

In general we have $p_{2^n}^{(n)}(x) \equiv 1$ for $n=1, 2, \dots$.

Now let $B_k^{(n)}$ be the set of all points (x, y) for which $p_{2k}^{(n)}(x) \leq y < p_{2k+1}^{(n)}(x)$ and let A_n be the union of the sets $B_k^{(n)}$ ($k=0, 1, 2, \dots, 2^{n-1}-1$).

It is easy to verify that the events A_n are equivalent and

$$\mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k}) = \omega_k = \int_0^1 x^k dF(x) \quad \text{for } k=1, 2, \dots \text{ and } i_1 < i_2 < \dots < i_k.$$

Clearly $\lambda(x, y) = x$ and the events A_n are independent under the condition $\lambda = x$ and all have the probability x ($0 < x < 1$).

Let us mention that the well known theorem of HAUSDORFF [9] — according to which a necessary and sufficient condition for a sequence $\{\omega_k\}$ to be absolutely monotonic is that ω_k should be the k -th moment of a distribution function in the interval $(0, 1)$ — follows easily from DE FINETTI's theorem. Thus the theorem of HAUSDORFF in question can be proved in a purely probabilistic way.

§ 2. The general form of a sequence of equivalent events

The aim of this § is to prove the following

Theorem 2. *Let $A_1, A_2, \dots$ be a sequence of equivalent events. Let $\lambda(\omega)$ be the local density of the sequence $\{A_n\}$ considered as a stable sequence. Then we have²⁾*

$$(2.1) \quad \mathbf{P}(A_{i_1} A_{i_2} \dots A_{i_k} | \lambda(\omega)) = \lambda^k(\omega) \quad \text{with probability 1}$$

for $k=1, 2, \dots$ and $i_1 < i_2 < \dots < i_k$.

PROOF. First of all we prove (2.1) for $k=1$. Let us assume that

$$(2.2) \quad \mathbf{P}(A_n | \lambda) = \lambda(\omega) + \varepsilon_n(\omega).$$

Here $\varepsilon_n(\omega)$ is a Baire-function of $\lambda(\omega)$ by the definition of conditional probability (see [7]). Put $\varepsilon_n(\omega) = g_n(\lambda(\omega))$ and let $\alpha_n(\omega)$ denote the indicator function of A_n . Let us denote by $\mathbf{M}(\xi)$ the expectation of the random variable ξ and by $\mathbf{M}(\xi|\eta)$ the conditional expectation of ξ with respect to the random variable η . In what follows we shall use repeatedly the following well known properties of conditional expectations (see [7]): for any ξ for which $\mathbf{M}(\xi)$ exists and any η

$$(2.3) \quad \mathbf{M}(\xi) = \mathbf{M}(\mathbf{M}(\xi|\eta))$$

further

$$(2.4) \quad \mathbf{M}(g(\eta)\xi|\eta) = g(\eta)\mathbf{M}(\xi|\eta)$$

where $g(x)$ is a Baire function. Now we have by (1.3), (2.2) and (2.3)

$$\mathbf{P}(A_n) = \int_{\Omega} \lambda d\mathbf{P} = \mathbf{M}(\mathbf{M}(\alpha_n|\lambda)) = \mathbf{M}(\lambda + \varepsilon_n)$$

therefore $\mathbf{M}(\varepsilon_n) = 0$ ($n=1, 2, \dots$).

Similarly, using (1.4), (2.2), (2.3) and (2.4) we have

$$\begin{aligned} \mathbf{P}(A_k A_l) &= \int_{\Omega} \lambda^2 d\mathbf{P} = \int_{A_k} \lambda d\mathbf{P} = \mathbf{M}(\lambda \alpha_k) = \mathbf{M}(\mathbf{M}(\lambda \alpha_k | \lambda)) = \\ &= \mathbf{M}(\lambda \mathbf{M}(\alpha_k | \lambda)) = \mathbf{M}(\lambda(\lambda + \varepsilon_k)) = \mathbf{M}(\lambda^2) + \mathbf{M}(\lambda \varepsilon_k). \end{aligned}$$

Therefore $\mathbf{M}(\lambda \varepsilon_k) = 0$ ($k=1, 2, \dots$). Similarly we obtain

$$\mathbf{M}(\lambda^n \varepsilon_k) = \int_0^1 x^n g_k(x) dF_{\lambda}(x) = 0 \quad (n=0, 1, 2, \dots; k=1, 2, \dots)$$

where $F_{\lambda}(x)$ is the distribution function of $\lambda(\omega)$. The fact that the sequence $\{x^n\}$ is a complete sequence in the space $L^2_{F_{\lambda}}(0, 1)$ (the space of functions in the interval $[0, 1]$ which are square integrable with respect to the measure defined by the distribution function $F_{\lambda}(x)$) ([8]), implies that $g_n(x)$ is equal to 0 almost everywhere with respect to the measure defined by $F_{\lambda}(x)$; thus we have

$$\mathbf{P}(\varepsilon_r = 0) = 1 \quad (r=1, 2, \dots)$$

²⁾ $\mathbf{P}(B|\lambda)$ denotes the conditional probability of the event B under the condition that the random variable λ takes on a fixed value.

The proof for $k=2$ is completely similar to the above proof. Let us put

$$\mathbf{P}(A_i A_j | \lambda) = \lambda^2 + \varepsilon_{ij}$$

where ε_{ij} is a Baire-function of λ . With these notations we have

$$\mathbf{P}(A_i A_j) = \mathbf{M}(\alpha_i \alpha_j) = \int_{\Omega} \lambda^2 d\mathbf{P} = \mathbf{M}[\mathbf{M}(\alpha_i \alpha_j | \lambda)] = \mathbf{M}(\lambda^2 + \varepsilon_{ij})$$

and therefore

$$\mathbf{M}(\varepsilon_{ij}) = 0.$$

Similarly

$$\begin{aligned} \mathbf{P}(A_i A_j A_k) &= \int_{\Omega} \lambda^3 d\mathbf{P} = \int_{A_i A_j} \lambda d\mathbf{P} = \mathbf{M}(\alpha_i \alpha_j \lambda) = \mathbf{M}[\mathbf{M}(\alpha_i \alpha_j \lambda | \lambda)] = \\ &= \mathbf{M}(\lambda(\mathbf{M}(\alpha_i \alpha_j | \lambda))) = \mathbf{M}(\lambda(\lambda^2 + \varepsilon_{ij})) = \mathbf{M}(\lambda^3) + \mathbf{M}(\lambda \varepsilon_{ij}) \end{aligned}$$

so

$$\mathbf{M}(\lambda \varepsilon_{ij}) = 0$$

and in general we obtain

$$\mathbf{M}(\varepsilon_{ij} \lambda^n) = 0 \text{ for } n=0, 1, \dots \text{ i. e. } \mathbf{P}(\varepsilon_{ij}=0) = 1.$$

The proof of (2.1) for any value of k is essentially the same.

§ 3. Equivalent random variables

Let $\xi_1, \xi_2, \dots$ be a finite or infinite sequence of random variables. The random variables $\{\xi_n\}$ are called equivalent if the distribution function

$$F_n(x_1, x_2, \dots, x_n) = \mathbf{P}\{\xi_{i_1} < x_1, \xi_{i_2} < x_2, \dots, \xi_{i_n} < x_n\}$$

depends only on n and $x_1, x_2, \dots, x_n$, and it does not depend on the sequence of different integers $i_1, i_2, \dots, i_n$. One can ask what are the generalizations of Theorems 1. and 2. for equivalent random variables.

A sequence $\{\xi_n\}$ of random variables is called a stable sequence (see [6]) if the sequence of events $A_n(x) = \{\xi_n < x\}$ ($n=1, 2, \dots$) is stable for every x belonging to a set X which is everywhere dense on the real line. Let $\lambda_x(\omega)$ denote the local density of the stable sequence $\{A_n(x)\}$. If $\lambda_x(\omega)$ is a constant for every $x \in X$, the sequence is mixing (see [11]). Clearly any sequence of equivalent random variables is stable in the above sense.

Let us denote the event $\{\xi_n < x\}$ by $A_n(x)$. It is evident that $A_n(x)$ ($n=1, 2, \dots$) is a sequence of equivalent events for every x , if $\{\xi_n\}$ is a sequence of equivalent random variables. Let the local density of the sequence $A_n(x)$ be $\lambda_x(\omega)$.

The following result is valid (see [2] and [5])

Theorem 3. *If $\xi_1, \xi_2, \dots$ is an infinite sequence of equivalent random variables then*

$$P(\xi_{i_1} < x_1, \xi_{i_2} < x_2, \dots, \xi_{i_n} < x_n) = F_n(x_1, \dots, x_n) = \int_{\Omega} \lambda_{x_1}(\omega) \dots \lambda_{x_n}(\omega) d\mathbf{P}.$$

Clearly $\lambda(\omega)$ as a function of x is a distribution function for almost all ω .

The proof of this theorem is exactly the same as the proof of Theorem 1.

The evident generalization of Theorem 2 is the following statement: the equivalent random variables $\xi_1, \xi_2, \dots$ are independent and identically distributed with respect to F where F is the least σ -algebra containing all the σ -algebras A_x ($-\infty < x < +\infty$) where A_x is the smallest σ -algebra with respect to which $\lambda_x(\omega)$ is measurable.

Bibliography

- [1] B. DE FINETTI, *Mem. della R. Acc. Naz. dei Lincei* (6), 4, **86** (1930).
- [2] B. DE FINETTI, La prévision: ses lois logiques, ses sources subjectives, *Ann. Inst. H. Poincaré*, **7** (1937), 1-68.
- [3] А. Я. ХИНЧИН, о классах эквивалентных событий, *Математический Сборник*, **39** (1932), 40-43.
- [4] А. Я. ХИНЧИН, О классах эквивалентных событий, *Доклады Акад. наук СССР* **85** (1952) 713-714.
- [5] E. HEWITT — L. J. SAVAGE, Symmetric measures on Cartesian products, *Trans. Amer. Math. Soc.*, **80** (1955), 470-501.
- [6] A. RÉNYI, On stable sequences of events, *SANKHYA* (in print).
- [7] A. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin, 1933.
- [8] G. ALEXITS, Convergence problems of orthogonal series, *Budapest*, 1961.
- [9] F. HAUSDORFF, Momentprobleme für ein endliches Intervall, *Math. Z.* **16** (1923), 220-248.
- [10] A. RÉNYI, On mixing sequences of events, *Acta Math. Acad. Sci. Hungar.* **9** (1958), 215-228.
- [11] A. RÉNYI—P. RÉVÉSZ, On mixing sequences of random variables, *Acta Math. Acad. Sci. Hungar.* **9** (1958), 389-394.

(Received April 17, 1963.)