PROBABILITY MEASURES IN INFINITE CARTESIAN PRODUCTS

BY
ANTONÍN ŠPAČEK

Introduction

The main purpose of this paper is to give a general scheme for the construction of probability measures in infinite Cartesian products of measurable spaces by using an appropriate imbedding of this Cartesian product into a standard random process of the "function space type" according to the terminology of Doob [3]. The application of this imbedding principle is based upon a method which is similar to that used in [8]. The final general result is expressed by Theorem 2 of §1. The significance of this theorem lies in the fact that it contains the abstract form of a relevant condition under which the construction of probability measures in infinite Cartesian products becomes always possible.

In §2 the imbedding principle is applied to the particular case of Cartesian products of separable metric spaces in order to obtain a generalized version of the well known Kolmogorov extension theorem [5]. An important step in the realization of the imbedding is the Banach-Mazur representation theorem [1]. The relevant condition under which the main theorem of §2 holds is a special form of absolute measurability. We shall see that the well known counterexamples of E. Sparre Andersen and B. Jessen [9] and that of P. R. Halmos [4] must violate this condition. One of the last contributions to the solution of problems of the above-mentioned type seems to be the paper of Blackwell [2]. D. Blackwell has proved the extension theorem under the assumption that the components of the Cartesian product are analytic sets, and we shall indicate in §2 the relation to our results.

In order to show that the application of the imbedding principle is not limited to the traditional cases of separable metric spaces, we shall establish in §3 a theorem concerning the construction of probability measures in infinite Cartesian products of sets of Schwartz distributions. The proof is based on the application of Theorem 2 of §1 and on a number of results of K. Winkelbauer [11] concerning random Schwartz distributions.

Throughout this paper the standard terminology and notation as well as a number of auxiliary results from Doob [3] and Halmos [4] are used without further particular reference.

1. The abstract scheme

Let us denote by F the set of all mappings of a fixed set $X \neq 0$ into a fixed set $R \neq 0$, i.e., $F = R^X$. For each $x \in X$, let τ_x be a mapping of F into R such that $\tau_x(f) = f(x)$ for every $f \in F$. Clearly,

$$\mathfrak{U}(x) = \{\tau_x^{-1}(D) : D \subset R\}$$

Received January 24, 1959.

is a complete algebra of subsets of F, and it is isomorphic to the complete algebra of all subsets of R under τ_x^{-1} for every $x \in X$. For each $A \subset X$, we shall denote by $\mathfrak{U}(A)$ the smallest complete algebra of subsets of F generated by the union

$$\bigcup_{x \in A} \mathfrak{U}(x).$$

For each $A \subset X$, the coincidence on A of mappings from F is an equivalence relation which induces a partition of F into equivalence sets. This partition coincides with the class of all atoms of $\mathfrak{U}(A)$. These atoms will be called *A-atoms,* and the complete algebra $\mathfrak{U}(A)$ coincides with the class of all unions of A-atoms. In particular, the X-atoms are one-element sets, and $\mathfrak{U}(X)$ is the class of all subsets of F. For each $x \in X$, we have $\mathfrak{U}(\{x\}) = \mathfrak{U}(x)$. We see at once that

$$\{0, F\} = \mathfrak{U}(0) \subset \mathfrak{U}(A_0) \subset \mathfrak{U}(A_1) \quad \text{for } A_0 \subset A_1 \subset X,$$

and that

$$\mathfrak{U}(A_0) \cap \mathfrak{U}(A_1) = \{0, F\} \qquad \text{for } A_0, A_1 \subset X, \quad A_0 \cap A_1 = 0.$$

Let us now consider a fixed sigma-algebra $\mathfrak{R}$ of subsets of R. Clearly,

$$\mathfrak{F}(x) = \{\tau_x^{-1}(D) : D \in \mathfrak{R}\}$$

is a sigma-algebra of subsets of F, and it is isomorphic to the sigma-algebra $\mathfrak{R}$ under τ_x^{-1} for every $x \in X$. For each $A \subset X$, we shall denote by $\mathfrak{F}(A)$ the smallest sigma-algebra of subsets of F generated by the union

$$\bigcup_{x \in A} \mathfrak{F}(x).$$

Clearly, $\mathfrak{F}(A) = \mathfrak{U}(A) \cap \mathfrak{F}(X)$; hence, $\mathfrak{F}(A) \subset \mathfrak{U}(A)$ for every $A \subset X$. In particular, $\mathfrak{F}(\{x\}) = \mathfrak{F}(x)$ for every $x \in X$. We see at once that

$$\{0, F\} = \mathfrak{F}(0) \subset \mathfrak{F}(A_0) \subset \mathfrak{F}(A_1) \quad \text{for } A_0 \subset A_1 \subset X,$$

and that

$$\mathfrak{F}(A_0) \cap \mathfrak{F}(A_1) = \{0, F\} \qquad \text{for } A_0, A_1 \subset X, \quad A_0 \cap A_1 = 0.$$

For each $A \subset X$, the sets from $\mathfrak{U}(A)$ will be called *A-cylinders,* and the sets from $\mathfrak{F}(A)$ *measurable A-cylinders.* The A-atoms of $\mathfrak{U}(A)$ are A-cylinders, but they are not in general measurable. We shall see that if R has at least two elements and A is nondenumerable, then the A-atoms are never measurable.

A class $\mathfrak{X}$ of subsets of the set X is said to be *sigma-directed* if to each denumerable subclass of $\mathfrak{X}$ there exists a set from $\mathfrak{X}$ including all sets from this denumerable subclass. Clearly, $\mathfrak{X}$ is sigma-directed if and only if each denumerable union of sets from $\mathfrak{X}$ is included in at least one set from $\mathfrak{X}$. The class $\mathfrak{X}$ of subsets of X is said to be a *covering class* if the union of all sets from $\mathfrak{X}$ coincides with X. For instance, the class of all denumerable subsets of X is a sigma-directed covering class whatever the power of X.

The following statement is obviously true.

(A) *If $\mathfrak{X}$ is a sigma-directed covering class of subsets of X, then*

$$\mathfrak{F}(X) = \bigcup_{A \in \mathfrak{X}} \mathfrak{F}(A).$$

For instance, our assertion that the atoms are under a wide variety of circumstances nonmeasurable is a simple consequence of (A).

We shall now introduce the concept of "abstract property" which, as will be seen later, is fundamental for this paper. Its significance will become completely clear from the application of our general results to the main problem of the construction of probability measures in Cartesian products of various concrete measurable spaces.

If $\mathfrak{X}$ is a sigma-directed covering class of subsets of X, then the mapping P of $\mathfrak{X} \cup \{X\}$ into $\mathfrak{U}(X)$ is said to be an *abstract property* with respect to $\mathfrak{X}$ if $P(A)$ is an A-cylinder, i.e., $P(A) \in \mathfrak{U}(A)$ for every $A \in \mathfrak{X} \cup \{X\}$.

The property P with respect to $\mathfrak{X}$ is said to be *extensible* if

$$P(X) \cap \bigcap_{x \in A} \{f : f \in F, f(x) = f_0(x)\} \neq 0$$

for every $A \in \mathfrak{X}$ and every $f_0 \in P(A)$, or in other words, if to each $A \in \mathfrak{X}$ and to each $f_0 \in P(A)$, there exists an $f \in P(X)$ such that $f(x) = f_0(x)$ for every $x \in A$.

(B) *If $\mathfrak{X}$ is a sigma-directed covering class of subsets of X and P an extensible property with respect to $\mathfrak{X}$, then to each $E \in \mathfrak{F}(X)$ for which $P(X) \subset E$, there corresponds a set $A_E \in \mathfrak{X}$ such that $P(A_E) \subset E$.*

The proof is very simple. If $E \in \mathfrak{F}(X)$, then by (A) there exists an $A_E \in \mathfrak{X}$ such that $E \in \mathfrak{F}(A_E)$. If $f_0 \in P(A_E)$, then, according to the extensibility of P, there exists an $f \in P(X)$ such that $f(x) = f_0(x)$ for every $x \in A_E$; hence f and f_0 belong to the same A_E-atom. By hypothesis $P(X) \subset E$; hence $f \in E$. Since $\mathfrak{U}(A_E)$ is the class of all unions of A_E-atoms and $\mathfrak{F}(A_E) \subset \mathfrak{U}(A_E)$, it follows that $f_0 \in E$, Q.E.D.

The property P with respect to $\mathfrak{X}$ is said to be *hereditary* if $P(A_1) \subset P(A_0)$ for $A_0, A_1 \in \mathfrak{X}$, $A_0 \subset A_1$.

We shall say that the property P with respect to $\mathfrak{X}$ is *measurable* if $P(A) \in \mathfrak{F}(X)$, i.e., $P(A)$ is a measurable cylinder for every $A \in \mathfrak{X}$.

Clearly, the conditions of extensibility, heredity, and measurability are consistent and mutually independent.

We shall now mention a simple auxiliary result which is a generalization of a theorem in [8].

THEOREM 1. *Let μ be a probability measure in $\mathfrak{F}(X)$, $\bar{\mu}$ the outer measure in $\mathfrak{U}(X)$ induced by μ, $\mathfrak{X}$ a sigma-directed covering class of subsets of X, and P a hereditary and extensible property with respect to $\mathfrak{X}$. Then a necessary and sufficient condition for $\bar{\mu}(P(X)) = 1$ is that $\bar{\mu}(P(A)) = 1$ for every $A \in \mathfrak{X}$.*

The necessity of the condition follows at once from the heredity of P and from the fact that $\bar{\mu}$ is monotone. The sufficiency is a simple consequence of (B). Indeed, since P is hereditary, the assertion in (B) can be completed by $P(X) \subset P(A_E)$, and the desired result follows at once from the definition of outer measure.

Clearly, if in addition P is measurable, i.e., $P(A)$ is a measurable cylinder, then the condition $\bar{\mu}(P(A)) = 1$ can be replaced by $\mu(P(A)) = 1$ for every $A \in \mathfrak{X}$.

We shall now repeat the construction of the Cartesian power replacing the set X by the Cartesian product $X \times Y$ where $Y \neq 0$. Let us denote by G the set of all mappings of $X \times Y$ into R, i.e., $G = R^{X \times Y}$. For each $x \in X$, $y \in Y$, we shall define the mapping $\sigma_{x,y}$ of G into R such that $\sigma_{x,y}(g) = g(x, y)$ for every $g \in G$. Clearly,

$$\mathfrak{B}(x, y) = \{\sigma_{x,y}^{-1}(D) : D \subset R\}$$

is a complete algebra of subsets of G, and

$$\mathfrak{G}(x, y) = \{\sigma_{x,y}^{-1}(D) : D \in \mathfrak{R}\}$$

is a sigma-algebra of subsets of G for every $x \in X$, $y \in Y$. For each $C \subset X \times Y$ we shall denote by $\mathfrak{B}(C)$ the smallest complete algebra of subsets of G generated by the union

$$\bigcup_{(x,y) \in C} \mathfrak{B}(x, y),$$

and by $\mathfrak{G}(C)$ the smallest sigma-algebra of subsets of G generated by the union

$$\bigcup_{(x,y) \in C} \mathfrak{G}(x, y).$$

Clearly, $\mathfrak{B}(\{x\} \times \{y\}) = \mathfrak{B}(x, y)$, and $\mathfrak{G}(\{x\} \times \{y\}) = \mathfrak{G}(x, y)$ for $x \in X$, $y \in Y$. The properties of $\mathfrak{B}(C)$ and $\mathfrak{G}(C)$ for $C \subset X \times Y$ are analogous to the properties of $\mathfrak{U}(A)$ and $\mathfrak{F}(A)$ for $A \subset X$.

For each $y \in Y$, let us define a mapping t_y of G into F such that $t_y(g) = g(\cdot, y) \in F$ for $g \in G$. Clearly, $\tau_x(t_y(g)) = \sigma_{x,y}(g)$ for every $x \in X$, $y \in Y$, and $g \in G$. Since $t_y^{-1}(\tau_x^{-1}(D)) = \sigma_{x,y}^{-1}(D)$ for every $x \in X$, $y \in Y$, and $D \subset R$, it follows from the definitions of $\mathfrak{U}(x)$, $\mathfrak{B}(x, y)$, $\mathfrak{F}(x)$, and $\mathfrak{G}(x, y)$ that the complete algebras $\mathfrak{U}(x)$ and $\mathfrak{B}(x, y)$ and the sigma-algebras $\mathfrak{F}(x)$ and $\mathfrak{G}(x, y)$ are isomorphic under t_y^{-1} for every $x \in X$, $y \in Y$, and, using the definitions of $\mathfrak{U}(A)$, $\mathfrak{B}(A \times \{y\})$, $\mathfrak{F}(A)$, and $\mathfrak{G}(A \times \{y\})$, we obtain at once the following lemma.

(C) *The complete algebras* $\mathfrak{U}(A)$ *and* $\mathfrak{B}(A \times \{y\})$ *and the sigma-algebras* $\mathfrak{F}(A)$ *and* $\mathfrak{G}(A \times \{y\})$ *are isomorphic under* t_y^{-1} *for every* $A \subset X$ *and every* $y \in Y$.

Now we shall introduce the fundamental concept of absolute measurability. For this purpose we shall assume that there is given an extensible, hereditary, and measurable property P with respect to a sigma-directed covering class $\mathfrak{X}$ of subsets of X. A subset S of $P(X)$ is said to be *absolutely measurable* if to each probability measure φ in $\mathfrak{F}(X)$, there exists a set $E_\varphi \in \mathfrak{F}(X)$ such that $P(X) \cap E_\varphi \subset S$ and $\bar{\varphi}(S) = \varphi(E_\varphi)$, $\bar{\varphi}$ denoting the outer measure in $\mathfrak{U}(X)$ induced by φ. Using this definition of absolute measurability, we can now formulate the following

THEOREM 2. *Let $\mathfrak{X}$ be a sigma-directed covering class of subsets of X, let P_y be an extensible, hereditary, and measurable property with respect to $\mathfrak{X}$, and suppose that the subset S_y of $P_y(X)$ is absolutely measurable for every $y \in Y$. Let μ be a probability measure in $\mathfrak{S}(X \times Y)$ and $\bar{\mu}$ the outer measure in $\mathfrak{B}(X \times Y)$ induced by μ. Then*

$$\bar{\mu}(\bigcap_{y \in Y} t_y^{-1}(S_y)) = 1$$

if and only if

$$\bar{\mu}(t_y^{-1}(S_y)) = 1 \tag{1}$$

for every $y \in Y$.

The necessity of the condition (1) is obvious, and its sufficiency will essentially result from Theorem 1, from (C), and from the following two obvious facts:

(D) *If $\mathfrak{Y}$ is a sigma-directed covering class of subsets of Y, then*

$$\{X \times B : B \in \mathfrak{Y}\} = X \times \mathfrak{Y}$$

is a sigma-directed covering class of subsets of $X \times Y$, and if in addition $\mathfrak{X}$ is a sigma-directed covering class of subsets of X, then

$$\{A \times B : A \in \mathfrak{X}, B \in \mathfrak{Y}\} = \mathfrak{X} \times \mathfrak{Y}$$

is a sigma-directed covering class of subsets of $X \times Y$.

(E) *If $\mathfrak{Y}$ is a sigma-directed covering class of denumerable subsets of Y, if $\mathfrak{X}$ is a sigma-directed covering class of subsets of X, and if P_y is an extensible, hereditary, and measurable property with respect to $\mathfrak{X}$ for every $y \in Y$, then the mapping Q of $(\mathfrak{X} \times \mathfrak{Y}) \cup \{X \times Y\}$ into $\mathfrak{B}(X \times Y)$ defined by*

$$Q(A \times B) = \bigcap_{y \in B} t_y^{-1}(P_y(A))$$

for $A \in \mathfrak{X} \cup \{X\}$, $B \in \mathfrak{Y} \cup \{Y\}$ is an abstract property with respect to $\mathfrak{X} \times \mathfrak{Y}$, and it is extensible, hereditary, and measurable.

In order to prove the sufficiency of the condition (1), we can assume that $\mathfrak{Y}$ is the class of all denumerable subsets of Y which evidently is a sigma-directed covering class. Let us suppose that (1) holds. By hypothesis, $S_y \subset P_y(X)$; hence by (C)

$$t_y^{-1}(S_y) \subset t_y^{-1}(P_y(X)) \in \mathfrak{B}(X \times \{y\}),$$

and using (1) we have

$$\bar{\mu}(t_y^{-1}(P_y(X))) = 1 \tag{2}$$

for all $y \in Y$. The properties P_y for $y \in Y$ are measurable; hence by (C)

$$t_y^{-1}(P_y(A)) \in \mathfrak{S}(A \times \{y\}),$$

and using in addition the heredity of P_y for all $y \in Y$, we obtain by (2)

$$\mu(t_y^{-1}(P_y(A))) = 1$$

for every $y \in Y$ and every $A \in \mathfrak{X}$. Since the sets from $\mathfrak{Y}$ are denumerable, therefore, by the definition of Q, we can state that

$$Q(A \times B) \in \mathfrak{S}(A \times B),$$
$$\mu(Q(A \times B)) = 1 \tag{3}$$

for every $A \in \mathfrak{X}$ and every $B \in \mathfrak{Y}$. By (D), $\mathfrak{X} \times \mathfrak{Y}$ is a sigma-directed covering class of subsets of $X \times Y$, and by (E), Q is an extensible, hereditary, and measurable property with respect to $\mathfrak{X} \times \mathfrak{Y}$; hence, using (3) and Theorem 1, we see that

$$\bar{\mu}(Q(X \times Y)) = \bar{\mu}(\textstyle\bigcap_{y \in Y} t_y^{-1}(P_y(X))) = 1. \tag{4}$$

By hypothesis, the subsets S_y of $P_y(X)$ are absolutely measurable for all $y \in Y$, i.e., to each $y \in Y$, there corresponds a set $E_y \in \mathfrak{F}(X)$ such that

$$P_y(X) \cap E_y \subset S_y \tag{5}$$

and such that by (C) and by (1)

$$t_y^{-1}(E_y) \in \mathfrak{S}(X \times \{y\}), \tag{6}$$

$$\mu(t_y^{-1}(E_y)) = 1. \tag{7}$$

Since the sets from $\mathfrak{Y}$ are denumerable, it follows from (6) and (7) that

$$\textstyle\bigcap_{y \in B} t_y^{-1}(E_y) \in \mathfrak{S}(X \times Y), \tag{8}$$

$$\mu(\textstyle\bigcap_{y \in B} t_y^{-1}(E_y)) = 1 \tag{9}$$

for every $B \in \mathfrak{Y}$. Now let us define

$$W(X \times B) = Q(X \times B) \cap \textstyle\bigcap_{y \in B} t_y^{-1}(E_y) = \bigcap_{y \in B} t_y^{-1}(P_y(X) \cap E_y)$$

for every $B \in \mathfrak{Y} \cup \{Y\}$. Clearly, W is hereditary, and using (4) and (7) we see that $t_y^{-1}(P_y(X) \cap E_y) \neq 0$; hence it is also extensible with respect to $X \times \mathfrak{Y}$. By (4), (8), and (9), $\bar{\mu}(W(X \times B)) = 1$ for every $B \in \mathfrak{Y}$; hence a second application of Theorem 1 furnishes

$$\bar{\mu}(W(X \times Y)) = \bar{\mu}(\textstyle\bigcap_{y \in Y} t_y^{-1}(P_y(X) \cap E_y)) = 1. \tag{10}$$

Finally, by (5)

$$\textstyle\bigcap_{y \in Y} t_y^{-1}(P_y(X) \cap E_y) \subset \bigcap_{y \in Y} S_y$$

holds; hence (10) together with the last inclusion implies

$$\bar{\mu}(\textstyle\bigcap_{y \in Y} t_y^{-1}(S_y)) = 1,$$

Q.E.D.

The theorem just proved expresses the main result on an intermediate level of generality. Its application to various concrete particular cases is very simple, as will be shown by two typical examples considered in the next two sections.

2. Separable metric spaces

The construction of probability measures in infinite Cartesian products of separable metric spaces is now possible by using the results of §1 and the following theorem due to S. Banach and S. Mazur [1].

Each separable metric space is isometric to a subset of the space of all real-valued continuous functions in the closed unit interval supplied with the usual metric.

According to this theorem, we can restrict ourselves to subsets of the space of all real-valued continuous functions in the closed unit interval, which evidently is an admissible formal simplification.

For this purpose we shall first specialize the assumptions concerning the sets R and X, the class $\mathfrak{X}$, and the sigma-algebra $\mathfrak{R}$. We shall suppose that R is the space of all real numbers, $X = \{x : x \in R, 0 \leqq x \leqq 1\}$, $\mathfrak{X}$ is the class of all denumerable dense subsets of X, and $\mathfrak{R}$ the sigma-algebra of all Borel subsets of R. Clearly, F becomes the set of all real-valued functions defined in X. For each $A \in \mathfrak{X} \cup \{X\}$ let

$$P(A) = \{f : f \in F, F \text{ is uniformly continuous in } A\}$$
$$= \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \{f : f \in F, \sup_{x_1, x_2 \in A, |x_1 - x_2| < 1/n} |f(x_1) - f(x_2)| < 1/m\}.$$

Clearly, the uniform continuity P is an extensible, hereditary, and measurable property with respect to $\mathfrak{X}$. The uniformity is relevant for the extensibility. Since, however, X is a compact space, we have

$$P(X) = \{f : f \in F, f \text{ is continuous in } X\}.$$

Let us define the metric ρ in $P(X)$ as usual by

$$\rho(f_1, f_2) = \max_{x \in X} |f_1(x) - f_2(x)|$$

for all pairs $f_1, f_2 \in P(X)$. It is well known that the metric space $P(X)$ is complete and separable with respect to ρ. Using the separability of $P(X)$, we obtain at once

$$P(X) \cap \mathfrak{F}(X) = \mathfrak{B}, \tag{11}$$

where $\mathfrak{B}$ denotes the sigma-algebra of all Borel subsets of $P(X)$.

Let S be a separable metric space, $\bar{S}$ its complete extension, and $\mathfrak{S}$ the sigma-algebra of all Borel subsets of $\bar{S}$. According to the Banach-Mazur theorem, we can always assume that $S \subset P(X)$, and that $\bar{S}$ is the closure of S in $P(X)$. Since $\bar{S} \subset P(X)$ and $\mathfrak{S} = \bar{S} \cap \mathfrak{B}$, therefore, by (11)

$$\mathfrak{S} = \bar{S} \cap \mathfrak{F}(X). \tag{12}$$

We shall say that the separable metric space S is *metrically absolutely measurable* if it is a measurable subset of $\bar{S}$ with respect to every probability measure in $\mathfrak{S}$, i.e., if for each probability measure in $\mathfrak{S}$ the corresponding inner and outer measures of S coincide.

It is well known that S is measurable with respect to a probability measure η in $\mathfrak{S}$ if and only if there exists a subset H_η of S such that $H_\eta \in \mathfrak{S}$

and $\eta(H_\eta) = \bar{\eta}(S)$, where $\bar{\eta}$ denotes the outer measure induced by η; hence by (11)

(F) *If S is metrically absolutely measurable, then it is absolutely measurable in the sense of* §1.

Let us denote by $\mathfrak{N}$ the class of all finite subsets of Y. With each $y \in Y$ we shall associate a separable metric space S_y. By the Banach-Mazur theorem it is legitimate to assume that $S_y \subset P(X)$ for every $y \in Y$. It is easy to verify that the union of the sigma-algebras

$$(13) \qquad (\bigcap_{y \in Y} t_y^{-1}(S_y)) \cap \mathfrak{G}(X \times B)$$

for $B \in \mathfrak{N}$, i.e.,

$$(14) \qquad \bigcup_{B \in \mathfrak{N}} ((\bigcap_{y \in Y} t_y^{-1}(S_y)) \cap \mathfrak{G}(X \times B))$$

is an algebra of subsets of

$$(15) \qquad \bigcap_{y \in Y} t_y^{-1}(S_y),$$

and this algebra (14) is a base of the sigma-algebra

$$(16) \qquad (\bigcap_{y \in Y} t_y^{-1}(S_y)) \cap \mathfrak{G}(X \times Y)$$

of subsets of (15).

The set (15) together with the sigma-algebra (16) is a measurable space which corresponds exactly to the Cartesian product of the measurable spaces $(S_y, S_y \cap \mathfrak{B})$ in accordance with the usual definition.

The main result of this section is the following generalization of the Kolmogorov theorem [5]:

THEOREM 3. *Let ψ be a real-valued set function in the algebra* (14) *which is a probability measure in the sigma-algebra* (13) *for every $B \in \mathfrak{N}$. If S_y is metrically absolutely measurable for every $y \in Y$, then there exists exactly one probability measure λ in the sigma-algebra* (16) *which coincides with ψ on the algebra* (14).

The assertion of this theorem is not true without the assumption of metrical absolute measurability of the components, as has been shown by Sparre Andersen and Jessen [9] and by Halmos [4]. On the other hand, in the same way as in the original version of the Kolmogorov theorem, the power of Y is completely irrelevant.

The proof of Theorem 3 is very simple. We shall first define a real-valued set function ν in the algebra

$$(17) \qquad \bigcup_{B \in \mathfrak{N}} \mathfrak{G}(X \times B),$$

using the identical mapping I of (15) into G. Clearly, $I^{-1}(E)$ belongs to (14) whenever E is from (17); hence, putting $\nu(E) = \psi(I^{-1}(E))$, we see at once that ν is a real-valued set function in the algebra (17), and it is a probability measure in the sigma-algebra $\mathfrak{G}(X \times B)$ for every $B \in \mathfrak{N}$. In order to satisfy formally the assumptions of the original version of the Kolmogorov theorem, note that to each finite set $C \subset X \times Y$ there exists a finite set $B_C \in \mathfrak{R}$ such that $C \subset X \times B_C$, and remember that R is the space of all real numbers and $\mathfrak{R}$ the sigma-algebra of all Borel subsets of R. Since $\mathfrak{S}(C) \subset \mathfrak{S}(X \times B_C)$, therefore, ν is a probability measure in $\mathfrak{S}(C)$ for every finite set $C \subset X \times Y$. We see that all assumptions of the Kolmogorov theorem are satisfied; hence there exists exactly one probability measure μ in the sigma-algebra $\mathfrak{S}(X \times Y)$ which coincides with ν on the algebra (17). Now we shall construct a probability measure λ in the sigma-algebra (16) which coincides with ψ on the algebra (14). By a well known lemma of Doob this is possible if and only if

$$(18) \qquad \bar{\mu}(\textstyle\bigcap_{y \in Y} t_y^{-1}(S_y)) = 1.$$

Under this condition the function λ defined by the equation $\mu = \lambda I^{-1}$ possesses all the desired properties, so that we can restrict ourselves to the verification of (18). By hypothesis, ψ is a probability measure in (13) for every $B \in \mathfrak{R}$; hence in particular, $\bar{\mu}(t_y^{-1}(S_y)) = 1$ for every $y \in Y$. Since S_y is assumed to be metrically absolutely measurable, therefore, by (F) it is absolutely measurable in the sense of §1 for every $y \in Y$, and (18) is an immediate consequence of Theorem 2. The uniqueness of the extension is evident.

The theorem just proved is nothing else but a simple consequence of the general result contained in Theorem 2. The assumption of absolute measurability is relevant, and each counterexample of the type considered by Sparre Andersen and Jessen [9] and by Halmos [4] must evidently violate the condition of absolute measurability. Since by [6] in a complete separable metric space every analytic set is absolutely measurable, Blackwell's generalization [2] of the Kolmogorov extension theorem is a corollary of Theorem 3.

3. Schwartz distributions

The main purpose of this section is to establish a theorem concerning the construction of probability measures in infinite Cartesian products of sets of Schwartz distributions. This theorem will be easily obtained by using the general results contained in §1. We shall restrict ourselves to the case of Schwartz distributions which are generalizations of real-valued functions of one real variable. This restriction is, however, completely irrelevant and is used only for the sake of simplicity.

Let R be the space of all real numbers and $\mathfrak{R}$ the sigma-algebra of all Borel subsets of R. If x is a real-valued function defined in R, then the closure of the set

$$\{r : r \in R,\ x(r) \neq 0\}$$

is said to be the support of x. We shall denote by X the set of all real-valued functions in R with compact supports and derivable any number of times. For each $x \in X$ and each $k = 0, 1, 2, 3, \cdots$, we shall denote by x^k the k^{th} derivative of x, and in particular $x^0 = x$, i.e., the 0^{th} derivative of x is the function x itself.

Clearly, with respect to the addition of functions and multiplication of functions by real numbers, the set X becomes a linear space.

The topology in X is determined by the following definition of convergence: We shall say that the sequence $x_1, x_2, x_3, \cdots$ of functions from X converges to the function $x \in X$ as $n \to \infty$, and we shall write $x_n \to x$, if there exists a positive integer m such that the supports of all $x, x_1, x_2, x_3, \cdots$ are contained in the interval

$$J_m = \{r : r \in R, \ -m \leqq r \leqq m\},$$

and the sequence $x_1^k, x_2^k, x_3^k, \cdots$ converges to x^k uniformly in J_m for all $k = 0, 1, 2, 3, \cdots$.

A real functional defined in X is said to be a *Schwartz distribution* if it is linear, i.e., if it is additive and continuous. This is the usual definition of Schwartz distributions [7].

For $m = 1, 2, 3, \cdots$, let X_m be the set of all functions from X the supports of which are included in the interval J_m; we shall denote by $\mathfrak{X}$ the class of all denumerable subgroups A of X which satisfy the following condition: to each positive integer m and to each $x \in X_m$, there exists a sequence $x_1, x_2, x_3, \cdots$ of functions from $A \cap X_m$ such that $x_n \to x$ as $n \to \infty$.

It was shown by K. Winkelbauer in [11] that $\mathfrak{X}$ is a sigma-directed covering class of subsets of X.

We shall preserve the notation F, $\mathfrak{U}(A)$, and $\mathfrak{F}(A)$ for $A \subset X$ as introduced in §1. In particular, F becomes the set of all real functionals defined in X.

Now let us define the property P with respect to $\mathfrak{X}$ as follows:

$$P(A) = (\bigcap_{x_1, x_2 \in A} \{f : f \in F, f(x_1 + x_2) = f(x_1) + f(x_2)\})$$
$$\cap \ (\bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{x \in A \cap X_m} \{f : f \in F, \ | f(x) | \leqq n \max_{r \in R} | x^k(r) | \})$$

for every $A \in \mathfrak{X} \cup \{X\}$.

It has been shown in [11] that the property P with respect to $\mathfrak{X}$ is extensible, hereditary, and measurable, and that

$$P(X) = \{f : f \in F, f \text{ is linear in } X\},$$

i.e., $P(X)$ coincides with the space of all Schwartz distributions.

A subset S of the space $P(X)$ of all Schwartz distributions is said to be *absolutely measurable* if it is absolutely measurable in the sense of the definition in §1.

We shall denote by $\mathfrak{N}$ the class of all finite subsets of Y as in §2, and with each $y \in Y$ we shall associate a set S_y of Schwartz distributions, i.e., of linear functionals in the function space X. Without any danger of confusion we can now use the notation (13), (14), (15), and (16) of §2 in order to formulate the analogue of Theorem 3 for infinite Cartesian products of sets of Schwartz distributions. Using our new interpretations of the notation of §2, we see that the set (15) together with the sigma-algebra (16) becomes the Cartesian product of the measurable spaces $(S_y, S_y \cap \mathfrak{F}(X))$ of Schwartz distributions in accordance with the usual definition, and the main result can now be simply established as follows:

THEOREM 4. *Let ψ be a real-valued set function in the algebra* (14) *which is a probability measure in the sigma-algebra* (13) *for every $B \in \mathfrak{N}$. If for each $y \in Y$ the set S_y of Schwartz distributions is absolutely measurable, then there exists exactly one probability measure λ in the sigma-algebra* (16) *which coincides with ψ on the algebra* (14).

The proof can be omitted because it is essentially the same as that of Theorem 3.

Conclusion

The application of the general results of §1 is not limited to the two particular cases of Cartesian products considered in the last two sections, which serve to illustrate the simplicity of the imbedding method. The last two special theorems show that the original version of Kolmogorov's theorem furnishes much more than may appear at first sight. On the other hand, the construction of probability measures in infinite Cartesian products of measurable spaces is not only of interest in itself, but it enables us to answer questions concerning the existence of very general random processes with prescribed probability distributions, as, for instance, random Schwartz distributions depending on an arbitrary parameter considered in [10]. An appropriate weakening of the conditions used in the last three theorems will probably furnish new conditions which will be not only sufficient, but also necessary.

REFERENCES

1. S. BANACH, *Théorie des opérations linéaires*, Warszawa, 1932.
2. D. BLACKWELL, *On a class of probability spaces*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1956, vol. II, pp. 1–6.
3. J. L. DOOB, *Stochastic processes*, New York, 1953.
4. P. R. HALMOS, *Measure theory*, New York, 1950.
5. A. N. KOLMOGOROFF, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin, 1933.
6. S. SAKS, *Theory of the integral*, 2nd rev. ed., Warszawa, 1937.
7. L. SCHWARTZ, *Théorie des distributions, tome I*, Paris, 1950.
8. A. ŠPAČEK, *Regularity properties of random transforms*, Czechoslovak Math. J., vol. 5 (1955), pp. 143–151.
9. E. SPARRE ANDERSEN AND B. JESSEN, *On the introduction of measures in infinite product sets*, Danske Videnskabernes Selskab Matematisk-Fysiske Meddelelser, vol. 25, no. 4, 1948.
10. M. ULLRICH, *Some theorems on random Schwartz distributions*, Transactions of the First Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, 1957, pp. 273–291.
11. K. WINKELBAUER, *On the theory of random Schwartz distributions*, Czechoslovak Math. J., vol. 6 (1956), pp. 517–521 (Russian with French summary).

CZECHOSLOVAK ACADEMY OF SCIENCES
PRAGUE, CZECHOSLOVAKIA