Yu. A. Rozanov

Markov Random Fields

Translated by Constance M. Elson

Yu. A. Rozanov

Markov Random Fields

Translated by Constance M. Elson

Springer-Verlag
New York Heidelberg Berlin

Yu. A. Rozanov
Steklov Mathematics Institute
UL. Vavilov 42
Moscow 117333
U.S.S.R.

Constance M. Elson (*Translator*)
403 Turner Place
Ithaca, NY 14850
U.S.A.

AMS Subject Classifications (1980): 60G60

Library of Congress Cataloging in Publication Data

Rozanov, I͡U. A. (I͡Urii Anatol'evich), 1934–
Markov random fields.

(Applications of mathematics)
Bibliography: p.
Includes index.
1. Random fields. 2. Vector fields.
I. Title. II. Series.
QA274.45.R68 519.2 82-3303
AACR2

With 1 Illustration

The original Russian title is Random Vector Fields, published by Nauka 1980.

Typeset by Composition House Ltd., Salisbury, England.
Printed and bound by R. R. Donnelley & Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90708-4 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90708-4 Springer-Verlag Berlin Heidelberg New York

Contents

CHAPTER 1

General Facts About Probability Distributions

§1. Probability Spaces

1. Measurable Spaces

Let X be an arbitrary set. When we consider elements $x \in X$ and sets $A \subseteq X$, we call X a *space*.

We use standard notation for set operations: $\cup$ for union, $\cap$ for intersection (also called the product and sometimes indicated by a dot), A^c for the complement of A, $A_1 \setminus A_2 = A_1 \cdot A_2^c$ for the difference of A_1 and A_2, $A_1 \circ A_2 = (A_1 \setminus A_2) \cup (A_2 \setminus A_1)$ for the symmetric difference, $\varnothing$ for the empty set.

Collections of Sets. When looking at collections of sets, we will use the following terminology.

A collection $\mathfrak{G}$ of subsets of the space X is called a *semi-ring* when for any sets A, A_1 in $\mathfrak{G}$ their intersection is also in $\mathfrak{G}$ and when $A_1 \subseteq A$, then A can be represented as a finite union of disjoint sets $A_1, \ldots, A_n$ in $\mathfrak{G}$, $A = \bigcup_1^n A_i$. We also require that $\varnothing \in \mathfrak{G}$ and the space X itself be represented as a countable union of disjoint sets $A_1, \ldots \in \mathfrak{G}$: $X = \bigcup_1^\infty A_i$.

A semi-ring $\mathfrak{G}$ is a *ring* if for any two sets A_1, A_2, it also contains their union.

Let $\mathfrak{G}$ be an arbitrary semi-ring. Then the collection of all sets $A \subseteq X$ which can be represented as a finite union of intersections of sets in $\mathfrak{G}$ is a ring. If the ring $\mathfrak{G}$ also includes the set X, then it is called an *algebra*.

An algebra is invariant with respect to the operations union, intersection and complement, taken a finite number of times. The collection of sets is called a *σ-algebra* if this invariance holds when the operations are taken a countable number of times.

The intersection of an arbitrary number of σ-algebras is again a σ-algebra. For any collection of sets $\mathfrak{G}$, there is a σ-algebra $\mathscr{A}$ containing $\mathfrak{G}$. The minimal such σ-algebra is called the *σ-algebra generated by the collection* $\mathfrak{G}$.

EXAMPLE (Union of σ-algebras). Let $\mathscr{A} = \mathscr{A}_1 \vee \mathscr{A}_2$ be the minimal σ-algebra containing both $\mathscr{A}_1$ and $\mathscr{A}_2$. It is generated by the semi-ring $\mathfrak{G} = \mathscr{A}_1 \cdot \mathscr{A}_2$ of sets of the form $A = A_1 \cdot A_2$, $A_i \in \mathscr{A}_i$.

We call a σ-algebra $\mathscr{A}$ *separable* if it is generated by some countable collection of sets $\mathfrak{G}$. Notice that in the case when $\mathfrak{G}$ is a countable collection, the algebra it generates is countable, consisting of all sets which can be derived from $\mathfrak{G}$ by finite intersections, unions, and complements.

When we speak of X as a *measurable space* we will mean that it is equipped with a particular σ-algebra $\mathscr{A}$ of sets $A \subseteq X$. We indicate a measurable space by the pair $(X, \mathscr{A})$. In the case where X is a topological space, then frequently the σ-algebra $\mathscr{A}$ is generated by a complete neighborhood system (basis) of X. Usually we will deal with the *Borel σ-algebra*, generated by all open (closed) sets, or the *Baire σ-algebra*, which is the σ-algebra generated by inverse images of open (closed) sets in $\mathbb{R}$ under continuous mappings $\varphi: X \to \mathbb{R}$.

If X is a metric space with metric ρ, and if $F \subseteq X$ is any closed set, then the function $\varphi(x) = \inf_{x' \in F} (x, x')$, $x \in X$, is continuous and F is the pre-image of $\{0\}$ under φ, $F = \{x: \varphi(x) = 0\}$; hence each Borel set is Baire. This is also true for compact X with countable basis: such a space is metrizable.

EXAMPLE. The system of half-open intervals (x', x'') on the real line $X = \mathbb{R}$ forms a semi-ring and the σ-algebra it generates is the collection of all Borel sets. The same is true of the countable semi-ring of half-open intervals with rational endpoints.

EXAMPLE (The semi-ring generated by closed sets). The collection $\mathfrak{G}$ of all sets of the form $A = G_1 \setminus G_2$, where G_1 and G_2 are closed sets, is a semi-ring: for any $A', A'' \in \mathfrak{G}$, $A' \cap A'' = G_1' \cdot G_1'' \setminus (G_2' \cup G_2'') \in \mathfrak{G}$; furthermore if $A'' \subseteq A'$, we can assume $G_2' \subseteq G_2'' \subseteq G_1'' \subseteq G_1'$ and we have $A' \setminus A'' = A_1 \cup A_2$ where $A_1 = G_1' \setminus G_1''$ and $A_2 = G_2'' \setminus G_2'$ are disjoint.

EXAMPLE (The semi-ring of Baire sets). Let F be a closed Baire set in X which is the inverse image of some closed set B on the real line Y, $F = \{\varphi \in B\}$. If one takes any continuous function ψ on Y, mapping the closed set B to 0 and strictly positive outside B (for instance, $\psi(y)$ could be the distance from the point $y \in Y$ to the set $B \subseteq Y$), then the composition $\psi \circ \varphi$ is continuous on X and the closed Baire set F is precisely the null set $\{\psi \circ \varphi = 0\}$. The system $\mathfrak{G}$ of all closed Baire sets F which are null sets of continuous functions φ on the real line contains the intersection $F_1 \cap F_2$ and union $F_1 \cup F_2$ for any $F_1, F_2 \in \mathfrak{G}$. For example, if $F_i = \{\varphi_i = 0\}$ then $F_1 \cup F_2 =$

$\{\varphi_1\varphi_2 = 0\}$ and $F_1 \cap F_2 = \{|\varphi_1| + |\varphi_2| = 0\}$. The collection of all sets A which can be represented as a difference $F_1 \backslash F_2$ of two sets $F_2 \subseteq F_1$ in $\mathfrak{G}$ is a semi-ring which generates the entire σ-algebra of Baire sets in the space X.

Standard Borel σ-algebras. Let $(X, \mathscr{A})$ be a measurable space; we call $\mathscr{A}$ a standard Borel σ-algebra if it is isomorphic to a Borel σ-algebra $\mathscr{B}$ on some Borel subset Y of a complete separable metric space. (Two σ-algebras $\mathscr{A}$ and $\mathscr{B}$ are Borel isomorphic if there is a one-to-one mapping $\varphi: X \to Y$ and $\mathscr{A}$ consists of all $A \subseteq X$ of the form $A = \{x: \varphi(x) \in B\}$, $B \in \mathscr{B}$.) The following holds: *a standard Borel σ-algebra is isomorphic to a Borel σ-algebra on some compact metric space.*

Products of Spaces. The product of measurable spaces $(X_1, \mathscr{A}_1)$ *and* $(X_2, \mathscr{A}_2)$ is the space $X = X_1 \times X_2$ of all pairs (x_1, x_2), $x_i \in X_i$, with σ-algebra $\mathscr{A}$ generated by the semi-ring $\mathfrak{G} = \mathscr{A}_1 \times \mathscr{A}_2$ of sets $A \subseteq X$ of the form $A = A_1 \times A_2$, $A_i \in \mathscr{A}_i$; more precisely, A is the set of all pairs (x_1, x_2), $x_i \in A_i$.

We define a finite product $X = \prod_{t \in T} X_t$ of measurable spaces $(X_t, \mathscr{A}_t)$ in the same way. Here T is a finite index set and X is the set of elements $x = \{x_t, t \in T\}$, each a tuple of "coordinates" $x_t \in X_t$, with σ-algebra $\mathscr{A}$ generated by the semi-ring $\mathfrak{G} = \prod_{t \in T} \mathscr{A}_t$ of sets $A = \prod_{t \in T} A_t$, $A_t \in \mathscr{A}_t$. Each such A is a set of elements x with corresponding coordinates $x_t \in A_t$.

Let T be an arbitrary index set and $(X_t, \mathscr{A}_t)$, $t \in T$, be an arbitrary family of measurable spaces. We define the product $X = \prod_{t \in T} X_t$ to be the space of elements $x = \{x_t, t \in T\}$, given by means of "coordinates" $x_t \in X_t$, with σ-algebra $\mathscr{A}$ generated by the semi-ring $\mathfrak{G} = \prod_{t \in T} \mathscr{A}_t$ of cylinder sets. A *cylinder set* $A \subseteq X$ is of the form

$$A = \{x: x_S \in A_S\}, \tag{1.1}$$

where S is a finite subset of T. Here the symbol x_S indicates the point in the space $X_S = \prod_{t \in S} X_t$ whose S-coordinates are the same as those of x and $A_S \subseteq X_S$ is a set in the semi-ring $\prod_{t \in S} \mathscr{A}_t$. We call $(X, \mathscr{A})$ a coordinate space.

If the X_t are topological spaces, then the cylinder sets (1.1) with A_S of the form $A_S = \prod_S A_t$, A_t open in X_t, form a basis for the topological space $X = \prod_T X_t$; this is the *Tychonov product*. A commonly used example is the coordinate space $X = E^T$; here each X_t is some fixed space E and $\mathscr{A}_t$ is some fixed σ-algebra $\mathscr{B}$. The elements $x = \{x(t), t \in T\}$, of this space are all possible functions on the set T with values in the "phase space" E.

2. Distributions and Measures

A non-negative function $P = P(A)$ defined on the semi-ring $\mathfrak{G}$ of sets A in the space X is a *distribution* if $P(\varphi) = 0$ and

$$P(A) = \sum_k P(A_k), \quad \text{whenever } A = \bigcup_k A_k, \text{ a countable union of disjoint sets } A_1, \ldots, \text{ in } \mathfrak{G}. \tag{1.2}$$

In case (1.2) is true only for a finite number of sets, the function P is usually called a *weak distribution.* Every weak distribution P can be uniquely extended from the semi-ring $\mathfrak{S}$ to the ring of all sets $A \subseteq X$ which are a finite union of disjoint sets $A_1, \ldots, A_n \in \mathfrak{S}$; the extension is done using (1.2), which gives the finite additivity. A (weak) distribution P on a semi-ring $\mathfrak{S}$ is called bounded if the function $P(A)$ is bounded. We consider only bounded distribution. A weak distribution P on a ring $\mathfrak{S}$ is a distribution iff it is continuous in the following sense: for every monotone sequence of sets $A_1 \supseteq A_2 \supseteq \cdots$ whose intersection $\bigcap_n A_n = \varnothing$, $\lim_{n\to\infty} P(A_n) = 0$.

Each distribution extends uniquely to a measure, i.e., a countably additive function P on the σ-algebra $\mathscr{A}$ generated by the original semi-ring $\mathfrak{S}$. The extension is defined by

$$P(A) = \inf \sum_k P(A_k), \tag{1.3}$$

where the inf is taken over all sets $A_1, A_2, \ldots \in \mathfrak{S}$ whose union contains the set A.

A measure P on a topological space X is *Borel* (*Baire*) if it is defined on the Borel (Baire) sets.

For any set $A \subseteq X$, define $P(A)$ by means of (1.3); for $A_1, A_2 \subseteq X$, the "distance" $\rho(A_1, A_2) = P(A_1 \circ A_2)$ indicates to what extent the sets A_1, A_2 differ from one another. Let P be a measure on the σ-algebra $\mathscr{A}$. A set $A \subseteq X$ is called *measurable* if $\exists$ some $A' \in \mathscr{A}$ such that $P(A \circ A') = 0$. If $\mathfrak{S}$ is a ring generating $\mathscr{A}$, then a set A is measurable iff it can be approximated by sets $A_\varepsilon \in \mathfrak{S}$ in the sense that

$$P(A \circ A_\varepsilon) \leq \varepsilon, \qquad \text{for any } \varepsilon > 0. \tag{1.4}$$

The collection of all measurable sets is a σ-algebra and (1.3) defines the measure P on it. This extension of the original measure P is *complete* in the sense that any subset A' of a set A of measure zero is measurable and $P(A') = 0$.

All of the above observations apply to unbounded distributions and measures with minor restrictions; in discussing unbounded measures it is important to stress that X must be σ-finite, i.e., representable as a countable union of sets of finite measure.

Let $\mathscr{A}_1, \mathscr{A}_2$ be two collections of sets having the property that for each $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$, one can find $A'_1 \in \mathscr{A}_2$ and $A'_2 \in \mathscr{A}_1$ differing from A_1, A_2 by sets of measure 0, $P(A_1 \circ A'_1) = P(A_2 \circ A'_2) = 0$. We indicate this situation by the equality

$$\mathscr{A}_1 = \mathscr{A}_2 \ (\mathrm{mod}\ 0).$$

Tight Measures. A Borel measure P on a topological space X is *regular* if for every measurable set A,

$$P(A) = \sup_{F \subseteq A} P(F), \tag{1.5}$$

where the sup is taken over all closed sets $F \subseteq A$. (1.5) is equivalent to

$$P(A) = \inf_{G \supseteq A} P(G), \tag{1.6}$$

where the inf is taken over all open G containing A.

Let the measure P have the property that $P(X) = \sup_{F \subseteq X} P(F)$, where the sup is taken over compact F. Such P is said to be *tight*. Every measure on a complete separable metric space is tight. For such measures (1.5) can be restated with "compact" replacing "closed": i.e., a regular tight measure is *Radon*.

EXAMPLE Let X be the real line. Then the Borel and Baire sets coincide. The measure $P(A)$ of each measurable set A is defined by (1.3), where the inf is taken over all disjoint half-open intervals $A_k = (x'_k, x''_k]$ whose union contains A. At the same time each interval $(x', x'']$ is the intersection of a countable number of open intervals and (1.6) clearly holds with the inf taken over all open sets G containing A.

Equation (1.6) is also true in any topological space X on which the Borel and Baire sets coincide. Every set $F \in \mathscr{A}$ which is the null set of a continuous function φ on the real line Y is the intersection of a countable number of open sets of the form $G_\delta = \{|\varphi| < \delta\}$, $F = \bigcap G_\delta$. By the continuity of the measure P, $P(F) = \inf P(G_\delta)$. For the difference of such sets, $A = F_1 \setminus F_2$ with $F_2 \subseteq F_1$, we have $P(A) = P(F_1) - P(F_2) = \inf P(G)$, with the inf taken over all open sets G of the form $G = G_1 \setminus F_2$, with $F_1 \subseteq G_1$. Since sets of the form $F_1 \setminus F_2$ are a semi-ring generating $\mathscr{A}$, we have for any P-measurable set A, $P(A) = \inf \sum_k P(A_k)$, where the A_k are a countable disjoint covering of A and each $A_k = F_{1k} \setminus F_{2k}$. Clearly $P(A)$ coincides with $\inf_{G \supseteq A} P(G)$, the inf taken over all unions G of appropriate sets.

A weak distribution P on a semi-ring $\mathfrak{G}$ is *tight* if each set $A \in \mathfrak{G}$ can be arbitrarily closely approximated in the sense (1.4) by compact sets $F_\varepsilon \subseteq A$: $P(A \setminus F_\varepsilon) \leq \varepsilon$ for any $\varepsilon > 0$. Such a weak distribution is a distribution and extends to a tight measure on the σ-algebra generated by $\mathfrak{G}$.

We will show why this is true. We assume $\mathfrak{G}$ is the ring formed by finite unions of sets in the original semi-ring. It is sufficient to establish that P is continuous. If $\lim_n P(A_n) \neq 0$ for some sequence $A_1 \supseteq A_2 \supseteq \cdots$, then one can find a sequence of approximating compacts $F_n \subseteq A_n$ with $F_1 \supseteq F_2 \supseteq \cdots$, and with $P(F_n) = P(A_n) - P(A_n \setminus F_n) > 0$ and whose intersection is non-empty, $\varnothing \neq \bigcap_n F_n \subseteq \bigcap_n A_n$. Hence for any sequence $A_1 \supseteq A_2 \supseteq \cdots$ whose intersection is empty, we have $\lim P(A_n) = 0$.

Products of Measures. Let P_1, P_2 be measures on measurable spaces $(X_i, \mathscr{A}_i)$, $i = 1, 2$. The equation

$$P(A) = P_1(A_1)P_2(A_2)$$

defines a distribution on the semi-ring $\mathfrak{G} = \mathscr{A}_1 \times \mathscr{A}_2$ (sets of the form $A = A_1 \times A_2$, $A_i \in \mathscr{A}_i$) in the product space $X = X_1 \times X_2$. The corresponding measure $P = P_1 \times P_2$ on the σ-algebra $\mathscr{A}$ generated by $\mathscr{A}_1 \times \mathscr{A}_2$ is the *product* of the *measures* P_1 and P_2.

For an arbitrary family of measure spaces $(X_t, \mathscr{A}_t)$, $t \in T$, with $P_t(X_t) = 1$ for all but a finite number of t, we define the *product measure* in a similar way: $P = \prod_{t \in T} P_t$ on the coordinate space $X = \prod_{t \in T} X_t$ with σ-algebra $\mathscr{A}$ generated by the semi-ring $\mathfrak{G} = \prod_{t \in T} \mathscr{A}_t$ of cylinder sets of the form (1.1).

Let $X = \prod_{t \in T} X_t$ be a coordinate space and P a distribution on the semi-ring $\mathfrak{G} = \prod_T \mathscr{A}_t$ of cylinder sets (1.1). Then

$$P_S(A_S) = P(A), \qquad A \in \mathfrak{G} \tag{1.7}$$

defines the projection of the distribution P on the space $X_S = \prod_{t \in S} X_T$ and the corresponding semi-ring $\prod_{t \in S} \mathscr{A}_t$. It satisfies the following consistency condition: for $S_1 \subseteq S_2$, the distribution P_{S_1} is the projection of the distribution P_{S_2}.

Let P_S, $S \subseteq T$, be a family of distributions parametrized by finite subsets $S \subseteq T$ and satisfying the consistency conditions described above. Then equation (1.7) defines a weak distribution $P = P_T$ on the space $X = \prod_{t \in T} X_t$ and the semi-ring $\prod_{t \in T} \mathscr{A}_t$. For an arbitrary $S \subseteq T$, let P_S denote the projection of P on the space X_S and semi-ring $\prod_{t \in S} \mathscr{A}_t$. Clearly P_T is a distribution $\Leftrightarrow$ for any countable $S \subseteq T$, P_S is a distribution since then equation (1.2) will hold for countably many cylinder sets in $\prod_{t \in T} \mathscr{A}_t$.

In the case of a topological space, we saw that a weak distribution P_S is a distribution if it is tight. Suppose that the distributions P_t corresponding to singleton sets $S = \{t\}$, are tight. Then $P = P_S$ will be tight for countable $S \subseteq T$ since each set $A = \prod_{t \in S} A_t$, $A_t \in \mathscr{A}_t$, can be approximated arbitrarily closely by compacts of the form $F = \prod_{t \in S} F_t = \bigcap_{t \in S} \{x_t \in F_t\}$ for a suitable choice of $F_t \subseteq X_t$:

$$A \backslash F = \bigcup_{t \in S} \{x_t \in A_t \backslash F_t\}, \qquad P(A \backslash F) \leq \sum_{t \in S} P_t(A_t \backslash F_t).$$

In particular, if $X = E^T$, where E is a complete separable metric space, then for a consistent family of distributions P_S corresponding to finite $S \subseteq T$, equation (1.7) gives a distribution P on cylinder sets and it can be extended to a measure on the σ-algebra generated by the semi-ring $\mathfrak{G} = \prod_{t \in T} \mathscr{A}_t$.

Let $X = \prod_{t \in T} X_t$ be an arbitrary coordinate space with measure P on the σ-algebra $\mathscr{A}(T)$ generated by the semi-ring $\prod_{t \in T} \mathscr{A}_t$. Then for each measurable $A \subseteq X$, $\exists$ some countable $S \subseteq T$ and set A' in the σ-algebra $\mathscr{A}(S)$ generated by the semi-ring $\prod_{t \in S} \mathscr{A}_t$ such that $A = A'$ (mod 0); that is, A and A' differ by a set of measure 0.

Mappings and Measures. Let $(X, \mathscr{A})$ be a measurable space with measure P on the σ-algebra $\mathscr{A}$ and let $\varphi(x)$, $x \in X$, be a function taking values in a space Y. The equation

$$P^{\varphi}(B) = P\{\varphi \in B\} \tag{1.8}$$

defines a measure P^φ on the σ-algebra $\mathscr{B}$ consisting of all sets $B \subseteq \mathscr{B}$ whose pre-images $A = \{\varphi \in B\}$ belong to the σ-algebra $\mathscr{A}$.

Let X be an arbitrary space and $(Y, \mathscr{B})$ a measurable space with measure Q on the σ-algebra $\mathscr{B}$. We write $\varphi(A)$ for the image of a set $A \subseteq X$ under the map $\varphi: X \to Y$. When the set $\varphi(X)$ is measurable and $A = \{\varphi \in B\}$, then

$$P(A) = Q(B \cap \varphi(X)) \tag{1.9}$$

defines a measure P on the σ-algebra $\mathscr{A}^\varphi$, consisting of all pre-images $A = \{\varphi \in B\}$, $B \in \mathscr{B}$; we will say that the *σ-algebra $\mathscr{A}^\varphi$ is generated by the function φ*.

A map φ from a measurable space $(X, \mathscr{A})$ with complete measure P on the σ-algebra $\mathscr{A}$ to the measurable space $(Y, \mathscr{B})$ is called *measurable* if for each $B \in \mathscr{B}$, the pre-image $A = \{\varphi \in B\}$ is a measurable set. When speaking of a real measurable function φ, we will mean a map to the real line $Y = \mathbb{R}$ with the Borel σ-algebra $\mathscr{B}$.

Let $(X, \mathscr{A})$ be a topological space with tight Borel measure P. Then for any real measurable function φ, the image $B = \varphi(X)$ is a measurable set of the real line (with respect to the corresponding Borel measure P^φ).

We will show this. A measurable function φ is the uniform limit of piecewise constant functions φ_n defined by $\varphi_n(x) = y_{kn}$ if $x \in A_{kn}$, where $A_{1n}, \ldots$ are disjoint measurable sets in X; one can take approximating compact sets $F_{kn} \subseteq A_{kn}$ whose finite unions $F_n = \bigcup_{k \le m_k} F_{kn}$ are such that the intersection $X_\varepsilon = \bigcap_n F_n$ approximates the space X to within any previously specified $\varepsilon > 0$:

$$P(X \setminus X_\varepsilon) < \varepsilon.$$

Each function φ_n on the compact set X_ε takes only a finite number of different values y_{kn}; moreover, the pre-images $\{\varphi_n = y_{kn}\} = F_{kn} \cap X_\varepsilon$ are compact. It is clear that all functions $\varphi_n(x)$, $x \in X_\varepsilon$, as well as their uniform limit $\varphi(x)$, are continuous on X_ε. Let $B = \varphi(X)$. The image $B_\varepsilon = \varphi(X_\varepsilon)$ is compact, since φ is continuous, and we have

$$P^\varphi(B \setminus B_\varepsilon) = P(\{\varphi \in B\} - \{\varphi \in B_\varepsilon\}) \le P(X \setminus X_\varepsilon) < \varepsilon,$$

where ε can be chosen to be arbitrarily small. By (1.4) the set B is measurable.

A measure P on a measurable space (X, A) will be called *perfect* if for each measurable real function φ on X the image $B = \varphi(X)$ is measurable with respect to the Borel measure P^φ.

The Weak Topology on the Space of Measures. Let X be a topological space and $\mathscr{M}(X)$ the collection of all Borel measures P on the space X, normalized so $P(X) = 1$. The *weak topology* on $\mathscr{M}(X)$ is the topology generated by neighborhoods of $P \in \mathscr{M}(X)$ of the type

$$\left\{\tilde{P}: \left| \int_X f\, d\tilde{P} - \int_X f\, dP \right| < \varepsilon\right\},$$

where $\varepsilon > 0$ and f is any bounded continuous functions on the space X. We will say that a sequence of measures P_n *converges weakly* to the measure P if convergence takes place in the weak topology; in other words, P_n converges weakly to P if for any bounded continuous function f,

$$\int_X f(x)P_n(dx) \underset{n}{\to} \int_X f(x)P(dx).$$

In the case where X is compact metric, the space $\mathscr{M}(X)$ is also compact with respect to the weak topology.

3. Probability Spaces

An arbitrary set Ω, together with a σ-algebra $\mathscr{A}$ of subsets of Ω and a positive measure P defined on $\mathscr{A}$ and normalized so $P(\Omega) = 1$, is a *probability space*. In speaking of a probability space $(\Omega, \mathscr{A}, P)$, the elements $\omega \in \Omega$ are usually called elementary events, the sets $A \in \mathscr{A}$ are events, and the measure $P(A)$ is the *probability* of event A occurring.

The concept of independence is of fundamental importance. Events $A_1, \ldots, A_n$ are called *independent* if

$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdots P(A_n); \tag{1.10}$$

σ-algebras $\mathscr{A}_1, \ldots, \mathscr{A}_n \subseteq \mathscr{A}$ are independent if (1.10) holds for any events $A_1 \in \mathscr{A}_1, \ldots, A_n \in \mathscr{A}_n$.

A measurable function $\xi = \xi(\omega)$, $\omega \in \Omega$, on a probability space $(\Omega, \mathscr{A}, P)$ taking values in a measure space $(X, \mathscr{B})$ is called a *random variable*. The probability measure P^ξ defined on the space X by $P^\xi(B) = P\{\xi \in B\}$, $B \in \mathscr{B}$, is the probability distribution of the random variable ξ.

We will say that the σ-algebra $\mathscr{A}$ is generated by a family of variables ξ if it is generated by all possible events of the form $\{\xi \in B\}$, $B \in \mathscr{B}$. Random variables ξ_1, ξ_2 with values in X will be called equivalent if $P\{\xi_1 \neq \xi_2\} = 0$, in other words if ξ_1 and ξ_2 are equal with probability 1. All random variables equivalent to the random variable ξ have the same probability distribution P^ξ.

Let ξ_k, $k = 1, \ldots, n$, be random variables with values in spaces $(X_k, \mathscr{B}_k)$. Their *joint probability distribution* $P^{\xi_1 \cdots \xi_n}$ is defined as the distribution on the semi-ring of sets of the form $B_1 \times \cdots \times B_n$, $B_i \in \mathscr{B}_i$, in the product space $X_1 \times \cdots \times X_n$ given by

$$P^{\xi_1 \cdots \xi_n}(B_1 \times \cdots \times B_n) = P\{\xi_1 \in B_1, \ldots, \xi_n \in B_n\},$$

for all $B_1 \in \mathscr{B}_1, \ldots, B_n \in \mathscr{B}_n$.

We call random variables $\xi_1, \ldots, \xi_n$ *independent* if

$$P^{\xi_1, \ldots, \xi_n}(B_1 \times \cdots \times B_n) = P^{\xi_1}(B_1) \cdots P^{\xi_n}(B_n), \qquad B_1 \in \mathscr{B}_1, \ldots, B_n \in \mathscr{B}_n.$$

Random Functions. Let $(E, \mathscr{B})$ be a measurable space and T an arbitrary set. The family of random variables $\xi(t)$, $t \in T$, with values in $(E, \mathscr{B})$ is a *random function on the set T with phase space $(E, \mathscr{B})$.* The distributions

$$P_S(B_1 \times \cdots \times B_n) = P\{\xi(t_1) \in B_1, \ldots, \xi(t_n) \in B_n\}, \qquad S = (t_1, \ldots, t_n),$$

on the products E^S are called the finite dimensional distributions of the random function $\xi = \xi(t)$, $t \in T$. Recall that each random variable $\xi(t) = \xi(\omega, t)$, $\omega \in \Omega$, is defined on the probability space (Ω, A, P); for each fixed $\omega \in \Omega$, the function $\xi(\omega, \cdot) = \xi(\omega, t)$, $t \in T$, is called a *trajectory.*

Let E be a compact separable metric space. For a given family of consistent probability distributions P_S, $S \subseteq T$, on E one can define a probability space and a family of random variables $\xi(t)$, $t \in T$, with finite dimensional distributions P_S, $S \subseteq T$: for Ω, take the coordinate space $X = E^T$, and for each $t \in T$ define $\xi(t) = \xi(\omega, t)$ as a function of $\omega = x \in X$ by the equation

$$\xi(\omega, t) = x(t), \quad \text{where} \quad x = \{x(t), t \in T\};$$

the corresponding probability measure P is defined on the σ-algebra $\mathscr{A} = \mathscr{A}(T)$ generated by all cylinder sets by means of the given distributions P_S, $S \subseteq T$, by equation (1.7).

A random function $\xi(t)$, $t \in T$, into the phase space $(E, \mathscr{B})$ gives a measurable mapping

$$\omega \to x = \xi(\omega, t), \qquad t \in T \tag{1.11}$$

from the probability space $(\Omega, \mathscr{A}, P)$ to the function space $X = E^T$ with probability distribution P^ξ on the σ-algebra $\mathscr{A}(T)$.

Random functions into the space E are called *equivalent* if for all $t \in T$, $\xi_1(t)$ and $\xi_2(t)$ are equivalent. The finite dimensional distributions of equivalent random functions coincide. In the class of all equivalent random functions one usually distinguishes a suitable representative having particular properties for the trajectory (that is, with trajectories in a particular function space X).

Let T be a topological space; a random function $\xi(t)$, $t \in T$, into a metric phase space E with distance function ρ is called *stochastically continuous* if for any $\varepsilon > 0$,

$$\lim_{s \to t} P\{\rho(\xi(s), \xi(t)) \geq \varepsilon\} = 0.$$

When speaking of random variables or random functions we will, as a rule, mean real (or complex) valued variables ξ. In this case, we let $E\xi$ denote the *mathematical expectation* of the random variable ξ,

$$E\xi = \int_\Omega \xi(\omega) P(d\omega).$$

We frequently consider the spaces $L^p(\mathscr{A}) = L^p(\Omega, \mathscr{A}, P)$, $p = 1, 2$ of all random variables ξ such that $E|\xi|^p < \infty$, with corresponding norm $\|\xi\| =$

$(E|\xi^p|)^{1/p}$; when $p = 2$ this gives the scalar product $(\xi_1, \xi_2) = E\xi_1 \cdot \overline{\xi_2}$. Convergence in the spaces $L^p(\mathscr{A})$ will be called *convergence in mean* ($p = 1$) and *in mean square* ($p = 2$).

In speaking of random variables $\xi \in L^p(\mathscr{A})$ we will not distinguish between equivalent random variables. In accordance with this we will not distinguish between σ-algebras which differ only by events of probability zero.

Let T be a domain in Euclidean space $\mathbb{R}^d$ and $\xi(t)$, $t \in T$, be a random function with finite second moments $E|\xi(t)|^2 < \infty$. In speaking of continuity, differentiability, or integrability we will mean the existence of these properties for $\xi(t)$, $t \in T$, regarded as a function on T with values in the Hilbert space $L^2(\Omega, \mathscr{A}, P)$.

Random Measures. Let T be a measurable space with a ring of measurable sets $\mathfrak{G}$; to each set $\Delta \in \mathfrak{G}$ associate a real or complex random variable $\eta(\Delta)$ with mean zero, $E\eta(\Delta) = 0$, and finite second moment, $E|\eta(\Delta)|^2 < \infty$. This defines a function on $\mathfrak{G}$ with values in $L^2(\Omega, \mathscr{A}, P)$ which we require to be additive: for disjoint $\Delta_1, \Delta_2 \in \mathfrak{G}$, $\eta(\Delta_1 \cup \Delta_2) = \eta(\Delta_1) + \eta(\Delta_2)$.

Suppose, in addition, that $E\eta(\Delta_1)\overline{\eta(\Delta_2)} = 0$ when Δ_1 and Δ_2 are disjoint and that the real-valued additive function $\mu(\Delta) = E|\eta(\Delta)|^2$ is a continuous distribution on the ring $\mathfrak{G}$. A random function $\eta(\Delta)$, $\Delta \in \mathfrak{G}$, having these properties is usually called a *random* (or *stochastic*) *orthogonal measure*. To characterize these, we will use the symbolic notation

$$E|\eta(dt)|^2 = \mu(dt).$$

For a measurable function $\varphi(t)$, $t \in T$, square-integrable with respect to the measure $\mu(dt)$, a standard construction defines the *stochastic integral*

$$\int_T \varphi(t)\eta(dt) \in L^2(\Omega, \mathscr{A}, P)$$

having the property that

$$E\int_T \varphi(t)\eta(dt) = 0,$$

$$E\left|\int_T \varphi(t)\eta(dt)\right|^2 = \int_T |\varphi(t)|^2\mu(dt),$$

and

$$E\left[\int_T \varphi_1(t)\eta(dt)\right]\left[\overline{\int_T \varphi_2(t)\eta(dt)}\right] = \int_T \varphi_1(t)\overline{\varphi_2(t)}\mu(dt).$$

Generalized Random Functions. Let T be an open domain in d-dimensional Euclidean space $\mathbb{R}^d$ and $C_0^\infty(T)$ the space of infinitely differentiable functions $u = u(t)$, $t \in T$, with compact support $\operatorname{Supp} u \subseteq T$. We can regard $C_0^\infty(T)$ as the union of topological spaces $C_0^\infty(T_{\text{loc}})$, T_{loc} a compact subset of T, each having a neighborhood basis at the origin of the form $\{u\colon \|u\|_l < \varepsilon\}$; here

$$\|u\|_l^2 = \sum_{|k| \le l} \|D^k u\|^2, \qquad l = 0, 1, \ldots,$$

and

$$\|D^k u\|^2 = \int_T |D^k u|^2 \, dt,$$

where

$$D^k u(t) = \frac{\partial^{|k|} u(t)}{\partial t_1^{k_1} \cdots \partial t_d^{k_d}}, \qquad t = (t_1, \ldots, t_d) \in \mathbb{R}^d,$$

$$k = (k_1, \ldots, k_d) \quad \text{and} \quad |k| = k_1 + \cdots + k_d.$$

Convergence of a sequence $u_n \to u$ in the space $C_0^\infty(T)$ means that the functions u_n all have support $\operatorname{Supp} u_n \subseteq T_{\text{loc}}$ for some compact $T_{\text{loc}} \subseteq T$ and that $u_n \to u$ in the topological space $C_0^\infty(T_{\text{loc}})$.

Consider a continuous linear map from the space $C_0^\infty(T)$ into $L^2(\Omega, \mathscr{A}, P)$, under which the functions $u \in C_0^\infty(T)$ correspond to random variables denoted by $(u, \xi) \in L^2(\Omega, \mathscr{A}, P)$. We will call this continuous linear operator $\xi = (u, \xi)$, $u \in C_0^\infty(T)$, a *generalized random function*. For $\xi = (u, \xi)$, we define the operations differentiation, multiplication by a C^∞ function, etc., as they are usually defined for ordinary generalized functions; that is,

$$D^k \xi = (-1)^{|k|}(D^k u, \xi),$$

$$a \cdot \xi = (\bar{a} \cdot u, \xi) \quad \text{for } a = a(t),\ t \in T, \quad \text{an infinitely differentiable function.}$$

An example of a generalized random function is given by the operator $(u, \xi) = \int_T u(t)\overline{\xi(t)}\, dt$, $u \in C_0^\infty(T)$, where the function $\xi(t) \in L^2(\Omega, \mathscr{A}, P)$, $t \in T$, is required to be integrable on every bounded domain $S \subseteq T$ and in particular,

$$\|(u, \xi)\| \le \int_T |u(t)| \, \|\xi(t)\| \, dt.$$

It is in the above sense that we will speak of a generalized random function henceforth.

Another example is offered by so-called "white noise" $\dot{\eta}(t)$, $t \in T$; this is a generalized random function of the form

$$(u, \dot{\eta}) = \int_T u(t)\dot{\eta}(t)\, dt = \int_T u(t)\eta(dt), \qquad u \in C_0^\infty(T).$$

The first expression only makes sense when interpreted according to the second (stochastic) integral, in which $\eta(dt)$ is the orthogonal random measure for which $E|\eta(dt)|^2$ is the Lebesgue measure dt.