PROBABILITY AND MATHEMATICAL STATISTICS

A Series of Monographs and Textbooks

FOUNDATIONS OF STOCHASTIC ANALYSIS

M. M. RAO

Foundations of Stochastic Analysis

This is a volume in
PROBABILITY AND MATHEMATICAL STATISTICS
A Series of Monographs and Textbooks

Editors: Z. W. Birnbaum and E. Lukacs

A complete list of titles in this series appears at the end of this volume.

Foundations
of
Stochastic Analysis

M. M. RAO

Department of Mathematics
University of California
Riverside, California

1981

ACADEMIC PRESS

A Subsidiary of Harcourt Brace Jovanovich, Publishers

New York London Toronto Sydney San Francisco

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Rao, Malempati Madhusudana, Date.
Foundations of stochastic analysis.

(Probability and mathematical statistics)
Bibliography: p.
Includes index.
1. Stochastic analysis. I. Title. II. Series.
QA274.2.R36 519.2 81-10831
ISBN 0-12-580850-X AACR2

PRINTED IN THE UNITED STATES OF AMERICA

81 82 83 84 9 8 7 6 5 4 3 2 1

To my sister, Jamantam,
for her help and encouragement
throughout my education

This page intentionally left blank

Contents

Chapter IV Martingales and Likelihood Ratios

Chapter V Abstract Martingales and Applications

Preface

Stochastic analysis consists of a study of different types of stochastic processes and of their transformations, arising from diverse applications. A basic problem in such studies is the existence of probability spaces supporting these processes when only their finite-dimensional distributions can be specified by the experimenter. The first solution to this problem is provided by the fundamental existence theorem of Kolmogorov (1933), according to which such a process, or equivalently a probability space, exists if and only if the set of all finite-dimensional distributions forms a compatible family. This result has been analyzed and abstracted by Bochner (1955), who showed it to be a problem on projective systems of probability spaces and who then presented sufficient conditions for such a system to admit a limit. The latter becomes the desired probability space, and this abstraction has greatly enlarged the scope of Kolmogorov's idea. One of the purposes of this book is to present the foundations of this theory of Kolmogorov and Bochner and to indicate its impact on the growth of the subject.

An elementary but important observation is that a projective system uniquely associates with itself a set martingale. In many cases the latter can be represented by a (point) martingale. On the other hand, a (point) martingale trivially defines a projective system of (signed) measure spaces. Thus the Kolmogorov–Bochner theory naturally leads to the study of martingales in terms of the basic (and independent) work due to Doob and Andersen–Jessen. However, to analyze and study the latter subject in detail, it is necessary to turn to the theory of conditional expectations and probabilities, which also appears in the desired generality in Kolmogorov's *Foundations* (1933) for the first time. This concept seems simple on the surface, but it is actually a functional operation and is nontrivial. To facilitate dealing with conditional expectations, which are immensely important in stochastic analyses, a detailed structural study of these operators is desirable. But such a general and comprehensive treatment has not yet appeared in book form. Consequently, after presenting the basic Kolmogorov–Bochner theorem in Chapter I, I devote Chapter II to this subject. The rest of the book treats aspects of martingales, certain extensions of projective limits, and applications to ergodic theory, to harmonic analysis, as well as to (Gaussian) likelihood ratios. The topics considered here are well suited for showing the natural interplay between real and abstract methods in stochastic analysis. I have tried to make this explicit. In so doing, I attempted to motivate the ideas at each turn so that one can see the appropriateness of a given method.

As the above description implies, a prerequisite for this book is a standard measure theory course such as that given in the Hewitt–Stromberg or Royden textbooks. No prior knowledge of probability (other than that it is a normed measure) is assumed. Therefore most of the results are proved in detail (at the risk of some repetitions), and certain elementary facts from probability are included. Actually, the present account may be regarded as an updating of Kolmogorov's *Foundations* (English translation, Chelsea, 1950, 74 pp.) referred to above, and thus a perusal of its first 56 pages will be useful. The treatment and the point of view of the present book are better explained by the brief outline that follows. A more detailed summary appears at the beginning of each chapter.

After introducing the subject, the main result proved in Chapter I is the basic Kolmogorov–Bochner existence theorem referred to above. To facilitate later work and to fix some notation and terminology, a résumé of real and abstract analysis is included here. Occasionally, some needed results that are not readily found in textbooks are presented in full detail. Most of these (particularly Section 4) can be omitted, and the reader may refer to them only when they are invoked. Chapter II is devoted entirely to conditional expectations and probabilities containing several characterizations of these operators and measures. The general viewpoint emphasizes that the Kolmogorov foundations are adequate for all the known applications. This is contrasted with (and is shown to include) the new foundations proposed by Rényi (1955). Then the integral representation of Reynolds operators is given as an application of these ideas, to be used later for a unified study of ergodic-martingale theories. Chapter III contains extensions of the Kolmogorov–Bochner theorem. The existence theorem of Prokhorov and certain other results of Choksi are also proved here. A treatment of direct limits of measures is necessary. This topic and infinite product conditional probabilities (Tulcea's theorem) are discussed. The work in this chapter is somewhat technical, and the reader might postpone the study of it until later. Chapters IV and V contain several aspects of (discrete) martingale theory. These include both scalar- and vector-valued martingales, their basic convergence, and many applications. The latter deal with ergodic theory, likelihood ratios, the Gaussian dichotomy theorem, and some results on the convergence of "partial sums" in harmonic analysis on a locally compact group. At the end of each chapter there is a problem section containing several facts, including important results in information theory, and many additions to the text. Most of these are provided with copious hints.

References to the literature are interspersed in the text with (I hope) due credits to various authors, backed up by an extensive bibliography. However, I have not always given the earliest reference of a given result. For instance, all the early work by Doob is referenced to his well-known treatise, and similarly, certain others with references to the monumental work of Dunford–Schwartz, from which an interested reader can trace the original source.

The arrangement of the material is such that this book can be used as a textbook for study following a standard real variable course. For this purpose, the following selections, based on my experience, are suggested: A solid semester's course can be given using Sections 1–3 of Chapter I, Chapter II (minus Section 6), Sections 1 and 2 of Chapter III, and most of Chapter IV. Then one can use any of the omitted sections with a view to covering Chapter V for the second semester. (This may be appropriately divided for a quarter system.) There is a sufficient amount of material for a year's treatment, and several possible extensions and open problems are pointed out, both in the text and in the Complements sections of the book. For ease of reference, theorems, lemmas, definitions, and the like are all consecutively numbered. Thus II.4.2 refers to the second item in Section 4 of Chapter II. In a given chapter (or section) the corresponding chapter (and section) number is omitted.

Several colleagues and students made helpful suggestions while the book was in progress. For reading parts of an earlier draft and giving me their comments and corrections, I am grateful to George Chi, Nicolae Dinculeanu, Jerome Goldstein, William Hudson, Tom S. Pitcher, J. Jerry Uhl, Jr., and Grant V. Welland. This work is part of a project that was started in 1968 with a sabbatical leave from Carnegie-Mellon University, continued at the Institute for Advanced Study during 1970-1972, and completed at the University of California at Riverside. This research was in part supported by the Grants AFOSR-69-1647, ARO-D-31-124-70-G100, and by the National Science Foundation. I wish to express my gratitude to these institutions and agencies as well as to the UCR research fund toward the preparation of the final version. I should like to thank Mrs. Joyce Kepler for typing the final and earlier drafts of the manuscript with diligence and speed. Also D. M. Rao assisted me in checking the proofs and preparing the Index. Finally, I appreciate the cooperation of the staff of Academic Press in the publication of this volume.

This page intentionally left blank

CHAPTER

I

Introduction and Generalities

This chapter is devoted to a motivational introduction and to preliminaries on real and abstract analysis to be used in the rest of the book. The main probabilistic result is the Kolmogorov–Bochner theorem on the existence of general, not necessarily scalar valued stochastic processes. Also included is a result on the existence of suprema for sets of measurable functions. Several useful complements are included as problems.

1.1 INTRODUCING A STOCHASTIC PROCESS

Stochastic analysis, in a general sense, is a study of the structural and inferential properties of stochastic processes. The latter object may be described as an indexed family of random variables $\{X_t, t \in T\}$ on a probability space. This brief statement implies much more and contains certain hidden conditions on the family. To explain this point clearly and precisely, we use the axiomatic theory of probability, due to Kolmogorov, and show how the basic probability space may be constructed, with the available initial information, in order that a stochastic process may be defined on it. Other axiomatic approaches, notably Rényi's, are also available, but the methods developed for the Kolmogorov model are adequate for all our purposes. This will become more evident in Chapter II, which elaborates on conditional probabilities, where Rényi's model is discussed and compared.

Thus, if (Ω, Σ, P) is a probability space, a mapping $X_t: \Omega \to \mathbb{R}$ (real line) is a (real) *random variable* if X_t is a measurable function. To fix the notation and for precision, we shall present a *résumé* of the main results from real analysis in Section 2, which will then be freely used in the book. Let T be an index set and $\{X_t, t \in T\}$ be a family of random variables on (Ω, Σ, P). If $t_1, \ldots, t_n$ are n points from T and $x_1, \ldots, x_n$ are in $\mathbb{R}$ or are $\pm\infty$, define the function $F_{t_1, \ldots, t_n}$, called the *n-dimensional (joint) distribution function* of $(X_{t_1}, \ldots, X_{t_n})$, by the equation

$$F_{t_1, \ldots, t_n}(x_1, \ldots, x_n) = P\left[\bigcap_{i=1}^{n} \{\omega : X_{t_i}(\omega) < x_i\}\right]. \tag{1}$$

As n and the t points vary, we get a family of multidimensional distribution functions $\{F_{t_1, \ldots, t_n}, t_i \in T, n \geq 1\}$. Since $\{\omega : X_t(\omega) < \infty\} = \Omega$, from (1) we get at once the following pair of relations:

$$F_{t_1, \ldots, t_n}(x_1, \ldots, x_{n-1}, \infty) = F_{t_1, \ldots, t_{n-1}}(x_1, \ldots, x_{n-1}), \tag{2}$$

$$F_{t_{i_1}, \ldots, t_{i_n}}(x_{i_1}, \ldots, x_{i_n}) = F_{t_1, \ldots, t_n}(x_1, \ldots, x_n), \tag{3}$$

where $(i_1, \ldots, i_n)$ is any permutation of $(1, \ldots, n)$. The functions $\{F_t, t \in T\}$ are monotone, nondecreasing, nonnegative, and left continuous. Moreover, $F_{t_1, \ldots, t_n}(+\infty, \ldots, +\infty) = 1$ and $F_{t_1, \ldots, t_n}(x_1, x_2, \ldots, x_{n-1}, -\infty) = 0$. The relations (2) and (3) are called the *Kolmogorov compatibility conditions* of the family $\{F_{t_1, \ldots, t_n} : t_i \in T, n \geq 1\}$. Thus any indexed family of (real) random variables on a probability space (or equivalently a stochastic process) determines a compatible collection of finite-dimensional distribution functions whose cardinality is that of D, the directed set (by inclusion) of all finite subsets of T.

The preceding description shows that even if the question of existence of a probability space (Ω, Σ, P) is not settled, it is simple to exhibit compatible families of distribution functions. It will then be natural to inquire into their relation to some (or any) probability space. To see that such families exist, let $f_1, \ldots, f_n$ be positive, measurable functions on the line each of which has integral equal to 1. Define $F_{1,2,\ldots,n}$ $(= F_n$, say):

$$F_n(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_1(t_1) \cdots f_n(t_n)\, dt_n \cdots dt_1. \tag{4}$$

It is clear that $\{F_n, n \geq 1\}$ is a family of distribution functions satisfying (2) and (3) with $T = \mathbb{N}$ there. A less simple collection is the Gaussian family of distribution functions given by

$$G_n(x_1, \ldots, x_n) = C_n \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} \exp[-\tfrac{1}{2}(t - \alpha)K^{-1}(t - \alpha)']\, dt_n \cdots dt_1, \tag{5}$$

where $K = (k_{ij})$ is a real symmetric positive definite matrix, $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a point of $\mathbb{R}^n$, $C_n = [(2\pi)^n \det(K)]^{-1/2}$, $\det(K) = \text{determinant}(K)$, and a prime denotes the transpose. An easy computation, which we omit, shows that the family $\{G_n, n \geq 1\}$ of (5) satisfies (2) and (3). Thus one can find many compatible families of distribution functions on $\{\mathbb{R}^n, n \geq 1\}$. A fundamental theorem of Kolmogorov states that every such compatible family of distribution functions yields a probability space and a stochastic process on it such that the (joint) finite-dimensional distributions on the process are precisely the given distributions. We shall prove this (in a slightly more general form) in Section 3. Thus the existence of a probability space is equivalent to the selection of a compatible family of distributions. Depending on the type of this family (i.e., *Gaussian*, *Poisson*, etc.), the probability space (Ω, Σ, P), or the stochastic process, is referred to by the same name. Let us first recall some measure theoretical results for convenient reference.

1.2 RÉSUMÉ OF REAL ANALYSIS

In this section we present an account of certain results from measure theory, mostly without proofs. Our purpose is to fix some notation and to make certain concepts precise since the reader is expected to have this background. (The omitted proofs may be found in Halmos [1], Hewitt–Stromberg [1], Royden [1], Sion [1], or Zaanen [1].)

Most of the references to measure will be to the abstract theory set forth by Càrathéodory as follows. Let $\mathscr{A}$ be a collection of subsets of a point set Ω for which $\varnothing \in \mathscr{A}$ and let $\tau: \mathscr{A} \to \bar{\mathbb{R}}^+$ be a function such that $\tau(\varnothing) = 0$. We define the set function μ on Ω by

$$\mu(A) = \inf\left\{\sum_{i=1}^{\infty} \tau(B_i) : B_i \in \mathscr{A}, A \subset \bigcup_{i=1}^{\infty} B_i\right\}, \qquad A \subset \Omega, \tag{1}$$

where $\inf(\varnothing) = +\infty$. We say that μ is *generated* by the pair $(\tau, \mathscr{A})$. Then μ is an outer measure. Let $\mathscr{M}_\mu = \{A \subset \Omega : \mu(T) = \mu(A \cap T) + \mu(A^c \cap T)$ for all $T \subset \Omega\}$. The following results holds.

1. Theorem (a) *The restriction of μ to $\mathscr{M}_\mu$, denoted by $\mu | \mathscr{M}_\mu$, is σ-additive, and $\mathscr{M}_\mu$ is a σ-algebra, containing the class of its μ-null sets (i.e., $\mathscr{M}_\mu$ is complete)*;

(b) *if $\mathscr{A}$ is a semi-ring and τ is additive, then $\mathscr{A} \subset \mathscr{M}_\mu$ and μ is an $\mathscr{A}_\sigma$-outer measure, i.e., for any $A \subset \Omega$, $\mu(A) = \inf\{\mu(B) : B \in \mathscr{A}_\sigma, B \supset A\}$, where $\mathscr{A}_\sigma$ is the closure of $\mathscr{A}$ under countable unions*;

(c) *under the hypothesis of (b), $\mu | \mathscr{A} = \tau$ iff τ is σ-additive*; and

(d) *if $\mu(\Omega) < \infty$, for each $A \subset \Omega$ there exists a $B \in \mathscr{A}_{\sigma\delta}$ (the closure of $\mathscr{A}_\sigma$ under countable intersections), $B \supset A$, such that $\mu(B) = \mu(A)$.*

[*Hence each A has a measurable cover B if μ is finite and the hypothesis of (b) holds.*]