M. Loève

Probability Theory II

4th Edition

M. Loève

Probability Theory II

4th Edition

Springer-Verlag
New York Heidelberg Berlin

M. Loève
Departments of Mathematics and Statistics
University of California at Berkeley
Berkeley, California 94720

AMS Subject Classifications
28–01, 60A05, 60Bxx, 60E05, 60Fxx

Library of Congress Cataloging in Publication Data

Loève, Michel, 1907–
Probability theory.
(Graduate texts in mathematics; 45-46)
Bibliography p. .
Includes index.
1. Probabilities. I. Title. II. Series.
QA273.L63 1977 519.2 76–28332

Originally published in the University Series in Higher Mathematics (D. Van Nostrand Company); edited by M. H. Stone, L. Nirenberg, and S. S. Chern.

Printed in the United States of America.

ISBN 0-387-90262-7 Springer-Verlag New York
ISBN 3-540-90262-7 Springer-Verlag Berlin Heidelberg

PREFACE TO THE FOURTH EDITION

This fourth edition contains several additions. The main ones concern three closely related topics: Brownian motion, functional limit distributions, and random walks. Besides the power and ingenuity of their methods and the depth and beauty of their results, their importance is fast growing in Analysis as well as in theoretical and applied Probability.

These additions increased the book to an unwieldy size and it had to be split into two volumes.

About half of the first volume is devoted to an elementary introduction, then to mathematical foundations and basic probability concepts and tools. The second half is devoted to a detailed study of Independence which played and continues to play a central role both by itself and as a catalyst.

The main additions consist of a section on convergence of probabilities on metric spaces and a chapter whose first section on domains of attraction completes the study of the Central limit problem, while the second one is devoted to random walks.

About a third of the second volume is devoted to conditioning and properties of sequences of various types of dependence. The other two thirds are devoted to random functions; the last Part on Elements of random analysis is more sophisticated.

The main addition consists of a chapter on Brownian motion and limit distributions.

It is strongly recommended that the reader begin with less involved portions. In particular, the starred ones ought to be left out until they are needed or unless the reader is especially interested in them.

I take this opportunity to thank Mrs. Rubalcava for her beautiful typing of all the editions since the inception of the book. I also wish to thank the editors of Springer-Verlag, New York, for their patience and care.

M. L.

January, 1977
Berkeley, California

PREFACE TO THE THIRD EDITION

This book is intended as a text for graduate students and as a reference for workers in Probability and Statistics. The prerequisite is honest calculus. The material covered in Parts Two to Five inclusive requires about three to four semesters of graduate study. The introductory part may serve as a text for an undergraduate course in elementary probability theory.

The Foundations are presented in:

the Introductory Part on the background of the concepts and problems, treated without advanced mathematical tools;
Part One on the Notions of Measure Theory that every probabilist and statistician requires;
Part Two on General Concepts and Tools of Probability Theory.

Random sequences whose general properties are given in the Foundations are studied in:

Part Three on Independence devoted essentially to sums of independent random variables and their limit properties;
Part Four on Dependence devoted to the operation of conditioning and limit properties of sums of dependent random variables. The last section introduces random functions of second order.

Random functions and processes are discussed in:

Part Five on Elements of random analysis devoted to the basic concepts of random analysis and to the martingale, decomposable, and Markov types of random functions.

Since the primary purpose of the book is didactic, methods are emphasized and the book is subdivided into:

unstarred portions, independent of the remainder; starred portions, which are more involved or more abstract;
complements and details, including illustrations and applications of the material in the text, which consist of propositions with fre-

quent hints; most of these propositions can be found in the articles and books referred to in the Bibliography.

Also, for teaching and reference purposes, it has proved useful to name most of the results.

Numerous historical remarks about results, methods, and the evolution of various fields are an intrinsic part of the text. The purpose is purely didactic: to attract attention to the basic contributions while introducing the ideas explored. Books and memoirs of authors whose contributions are referred to and discussed are cited in the Bibliography, which parallels the text in that it is organized by parts and, within parts, by chapters. Thus the interested student can pursue his study in the original literature.

This work owes much to the reactions of the students on whom it has been tried year after year. However, the book is definitely more concise than the lectures, and the reader will have to be armed permanently with patience, pen, and calculus. Besides, in mathematics, as in any form of poetry, the reader has to be a poet *in posse*.

This third edition differs from the second (1960) in a number of places. Modifications vary all the way from a prefix ("sub" martingale in lieu of "semi"-martingale) to an entire subsection (§36.2). To preserve pagination, some additions to the text proper (especially 9, p. 656) had to be put in the Complements and Details. It is hoped that moreover most of the errors have been eliminated and that readers will be kind enough to inform the author of those which remain.

I take this opportunity to thank those whose comments and criticisms led to corrections and improvements: for the first edition, E. Barankin, S. Bochner, E. Parzen, and H. Robbins; for the second edition, Y. S. Chow, R. Cogburn, J. L. Doob, J. Feldman, B. Jamison, J. Karush, P. A. Meyer, J. W. Pratt, B. A. Sevastianov, J. W. Woll; for the third edition, S. Dharmadhikari, J. Fabius, D. Freedman, A. Maitra, U. V. Prokhorov. My warm thanks go to Cogburn, whose constant help throughout the preparation of the second edition has been invaluable. This edition has been prepared with the partial support of the Office of Naval Research and of the National Science Foundation.

M. L.

April, 1962
Berkeley, California

CONTENTS OF VOLUME I

GRADUATE TEXTS IN MATHEMATICS VOL. 45

CONTENTS OF VOLUME II

GRADUATE TEXT IN MATHEMATICS VOL. 46

PART FOUR: DEPENDENCE

CHAPTER VIII: CONDITIONING

CHAPTER IX: FROM INDEPENDENCE TO DEPENDENCE

PART FIVE: ELEMENTS OF RANDOM ANALYSIS

CHAPTER XII: FOUNDATIONS; MARTINGALES AND DECOMPOSABILITY

CHAPTER XIII: BROWNIAN MOTION AND LIMIT DISTRIBUTIONS

CHAPTER XIV: MARKOV PROCESSES

Part Four

DEPENDENCE

For about two centuries probability theory has been concerned almost exclusively with independence. Yet, very particular forms of dependence appear already in the theory of games of chance. But a first general type of dependence—chains—was introduced only at the beginning of this century by Markov. Another type of dependence—stationarity—appears in ergodic theory, and a related type—second order stationarity—is then introduced in probability theory by Khintchine (1932). Centering at conditional expectations by P. Lévy (1935) gives rise to a new type of dependence—martingales.

At the very core of the study of dependence lies the concept of conditioning—with respect to a function—put in an abstract and rigorous form by Kolmogorov. In this part, the concept of conditioning is introduced in a more general form—with respect to a σ-field—and, as much as possible, the properties of various types of dependence are related to more general results, with emphasis given to the methods.

Chapter VIII

CONDITIONING

§27. CONCEPT OF CONDITIONING

The concept of "conditioning" can be expressed in terms of sub σ-fields of events. Conditional probabilities of events and conditional expectations of r.v.'s "given a σ-field $\mathcal{B}$," to be introduced and investigated in this chapter, are $\mathcal{B}$-measurable functions defined up to an equivalence. If $\mathcal{B}$ is determined by a countable partition of the sure event, then these functions are elementary. In this "elementary case," a constructive approach with a definite intuitive appeal is possible and there are no technical difficulties. In the general case, there is no suitable and rigorous constructive approach, and a descriptive one, requiring more powerful tools, especially the Radon-Nikodym theorem, has to be used.

The R.-N. theorem was obtained in its abstract form in 1930 and the concept of conditional probabilities and of conditional expectations of integrable r.v.'s "given" a measurable function, finite or not, numerical or not, was then put on a rigorous basis by Kolmogorov in 1933.

27.1. Elementary case. Investigation of the elementary case will give us an insight into the ideas involved in the intuitive notion of conditioning and will lead "naturally" to the notions and problems which appear in the general case.

The notion of conditional probability of an event A "given an event B" corresponds to that of frequencies of A in the repeated trials where B occurs; it is one of the oldest probability notions. For every event A, the relation

$$PB \cdot P_B A = PAB$$

defines the *conditional probability* (*c.pr.*) $P_B A$ *of* A *given* B as the ratio PAB/PB, provided B is a nonnull event; if B is null, so is AB, and the foregoing relation leaves $P_B A$ undetermined. In what follows, we assume that, unless otherwise stated, B is nonnull.

The function P_B on the σ-field $\mathfrak{A}$ of events, whose values are $P_B A$, $A \in \mathfrak{A}$, is called *conditional pr. given B*. The defining relation shows at once that since P on $\mathfrak{A}$ is normed, nonnegative, and σ-additive, so is P_B on $\mathfrak{A}$:

$$P_B \Omega = 1, \quad P_B \geqq 0, \quad P_B \sum A_j = \sum P_B A_j.$$

Thus, the conditioning expressed by "given B" means that the initial pr. space $(\Omega, \mathfrak{A}, P)$ is replaced by the pr. space $(\Omega, \mathfrak{A}, P_B)$. The expectation, if it exists, of a r.v. X on this new pr. space is called *conditional expectation* (*c.exp.*) *given B* and is denoted by $E_B X$; in symbols

$$E_B X = \int X \, dP_B.$$

Since $P_B = 0$ on $\{AB^c, A \in \mathfrak{A}\}$, the right-hand side reduces to $\int_B X \, dP_B$ and, since $P_B = \frac{1}{PB} P$ on $\{AB, A \in \mathfrak{A}\}$, it becomes $\frac{1}{PB} \int_B X \, dP$. Therefore, the c.exp. of X given B can be defined directly by

$$PBE_B X = \int_B X \, dP$$

and is determined if B is a nonnull event. In particular,

$$PBE_B I_A = \int_B I_A \, dP = PAB$$

so that the c.pr. $P_B A$ can be defined, thereafter, by

$$P_B A = E_B I_A.$$

Thus, if E_B is the *c.exp. given B*, with values $E_B X$ on the family $\mathcal{E}_B$ of all r.v.'s X whose integral on B exists, the c.pr. P_B becomes the restriction of E_B to the family $I_{\mathfrak{a}}$ of indicators of events. Furthermore, properties of P_B become particular cases of the immediate properties of E_B below.

If $X \geqq 0$ then $E_B X \geqq 0$, and if c is a constant then $E_B c = c$. If the X_j are nonnegative, or if the X_j are integrable and their consecutive sums are uniformly bounded by an integrable r.v., then $E_B \sum X_j = \sum E_B X_j$.

C.exp.'s (hence c.pr.'s) acquire their full meaning when reinterpreted as values of functions, as follows. The number $E_B X$ is no longer assigned to B but to every point of B, and similarly for $E_{B^c}X$, so that we have a two-valued function on Ω, with values E_BX for $\omega \in B$ and $E_{B^c}X$ for $\omega \in B^c$. More generally, let $\{B_j\}$ be a countable partition of Ω and let $\mathcal{B}$ be the minimal σ-field over this partition. Let $\mathcal{E}$ be the family of all r.v.'s X whose expectation EX exists, so that their indefinite integrals, hence c.exp.'s given any nonnull event, exist. Consider the elementary functions

$$E^{\mathcal{B}}X = \sum (E_{B_j}X)I_{B_j}, \quad X \in \mathcal{E}.$$

If some B_j are null, then the corresponding values $E_{B_j}X$ are undetermined, so that $E^{\mathcal{B}}X$ is undetermined on the null event which is the sum of null B_j. Such a possibility, together with the definition of $E_{B_j}X$, leads to the following

Constructive definition. *The elementary function $E^{\mathcal{B}}X$ defined up to an equivalence by*

$$(1) \qquad E^{\mathcal{B}}X = \sum \left(\frac{1}{PB_j}\int_{B_j} X\,dP\right) I_{B_j}, \quad X \in \mathcal{E},$$

is the c.exp. of X given $\mathcal{B}$.

Upon particularizing to indicators, the $\mathcal{B}$-measurable function $P^{\mathcal{B}}A$, defined up to an equivalence by setting

$$P^{\mathcal{B}}A = E^{\mathcal{B}}I_A, \quad A \in \mathcal{A},$$

will be the *c.pr. of A given* $\mathcal{B}$; the contraction of $E^{\mathcal{B}}$ on $I_{\mathcal{A}}$, to be denoted by $P^{\mathcal{B}}$, will be the *c.pr. given* $\mathcal{B}$, and its values are the $\mathcal{B}$-measurable functions $P^{\mathcal{B}}A$, $A \in \mathcal{A}$, defined up to an equivalence.

We say "given (the σ-field) $\mathcal{B}$" and not "given (the partition) $\{B_j\}$," because $E^{\mathcal{B}}X$ determines the c.exp. of X given an arbitrary nonnull event $B \in \mathcal{B}$. In fact, if $\sum'$ denotes the summation over some subclass of $\{B_j\}$, then every event $B \in \mathcal{B}$ is of the form $\sum' B_j$, and we have

$$PBE_BX = \int_{\Sigma' B_j} X\,dP = \sum{}' \int_{B_j} X\,dP = \sum{}' PB_jE_{B_j}X.$$

This relation can also be written as follows: If $P_{\mathcal{B}}$ is the restriction of P to $\mathcal{B}$, defined by

$$P_{\mathcal{B}}B = PB, \quad B \in \mathcal{B},$$

then the right-hand side becomes $\int_B (E^{\mathcal{B}}X)\,dP_{\mathcal{B}}$ while the left-hand side is $\int_B X\,dP$. This leads to the following