INTRODUCTION TO PROBABILITY AND STATISTICS

INTRODUCTION TO PROBABILITY AND STATISTICS

FROM A BAYESIAN VIEWPOINT

PART 1
PROBABILITY

BY

D. V. LINDLEY

Head of the Department of Statistics
University College London

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE

LONDON NEW YORK NEW ROCHELLE

MELBOURNE SYDNEY

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521055628

First published 1965
Reprinted 1969, 1976
First paperback edition 1980
Re-issued in this digitally printed version 2008

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-05562-8 hardback
ISBN 978-0-521-29867-4 paperback

To

M. P. MESHENBERG

in gratitude

CONTENTS

PREFACE

The content of the two parts of this book is the minimum that, in my view, any mathematician ought to know about random phenomena—probability and statistics. The first part deals with probability, the deductive aspect of randomness. The second part is devoted to statistics, the inferential side of our subject.

The book is intended for students of mathematics at a university. The mathematical prerequisite is a sound knowledge of calculus, plus familiarity with the algebra of vectors and matrices. The temptation to assume a knowledge of measure theory and general integration has been resisted and, for example, the concept of a Borel field is not used. The treatment would have been better had these ideas been used, but against this, the number of students able to study random phenomena by means of the book would have been substantially reduced. In any case the intent is only to provide an introduction to the subject, and at that level the measure theory concepts do not appreciably assist the understanding. A statistical specialist should, of course, continue his study further; but only, in my view, at a postgraduate level with the prerequisite of an honours degree in pure mathematics, when he will necessarily know the appropriate measure theory.

A similar approach has been adopted in the level of the proofs offered. Where a rigorous proof is available at this level, I have tried to give it. Otherwise the proof has been omitted (for example, the convergence theorem for characteristic functions) or a proof that omits certain points of refinement has been given, with a clear indication of the presence of gaps (for example, the limiting properties of maximum likelihood). Probability and statistics are branches of applied mathematics—in the proper sense of that term, and not in the narrow meaning that is common, where it means only applications to physics. This being so, some slight indulgence in the nature of the rigour is perhaps permissible. The applied nature of the subject means

that the student using this book needs to supplement it with some experience of practical data handling. No attempt has been made to provide such experience in the present book, because it would have made the book too large, and in any case other books that do provide it are readily available. The student should be trained in the use of various computers and be given exercises in the handling of data. In this way he will obtain the necessary understanding of the practical stimuli that have led to the mathematics, and the use of the mathematical results in understanding the numerical data. These two aspects of the subject, the mathematical and the practical, are complementary, and both are necessary for a full understanding of our subject. The fact that only one aspect is fully discussed here ought not to lead to neglect of the other.

The book is divided into eight chapters, and each chapter into six sections. Equations and theorems are numbered in the decimal notation: thus equation 3.5.1 refers to equation 1 of section 5 of chapter 3. Within §3.5 it would be referred to simply as equation (1). Each section begins with a formal list of definitions, with statements and proofs of theorems. This is followed by discussion of these, examples and other illustrative material. In the discussion an attempt has been made to go beyond the usual limits of a formal treatise and to place the ideas in their proper contexts; and to emphasize ideas that are of wide use, as distinct from those of only immediate value. At the end of each chapter there is a large set of exercises, some of which are easy, but many of which are difficult. Most of these have been taken from examinations papers, and I am grateful for permission from the Universities of London, Cambridge, Aberdeen, Wales, Manchester and Leicester to use the questions in this way. (In order to fit into the Bayesian framework some minor alterations of language have had to be made in these questions. But otherwise they have been left as originally set.)

The first part of the book, the first four chapters, is devoted to probability. The axioms of probability are stated in chapter 1 and elementary deductions made from them. In chapters 2 and 3 these results are applied to the study of random variables in one and higher dimensions respectively. Chapter 4 provides

an introduction to the study of stochastic processes, including queueing theory, renewal theory and Markov chains.

The axiomatic structure used here is not the usual one associated with the name of Kolmogorov. Instead one based on the ideas of Renyi has been used. The essential difference between the two approaches is that Renyi's is stated in terms of conditional probabilities, whereas Kolmogorov's is in terms of absolute probabilities, and conditional probabilities are defined in terms of them. Our treatment always refers to the probability of A, given B, and not simply to the probability of A. In my experience students benefit from having to think of probability as a function of two arguments, A and B, right from the beginning. The conditioning event, B, is then not easily forgotten and misunderstandings are avoided. These ideas are particularly important in Bayesian inference where one's views are influenced by the changes in the conditioning event.

Another novelty in chapter 1 is the extensive discussion of probability as a degree of belief in section 6. The reason for this is that the treatment of statistics in the second part of the book is based on this form of probability. An attempt, at a modest mathematical level, has been made to justify the axiomatic structure for beliefs: in the hope that this will help to convince students of the reasonableness of the idea that beliefs can be measured numerically. The content of the first three chapters is a prerequisite for the second part on statistics. The study of stochastic processes in chapter 4 is not used in the latter part.

I am extremely grateful to J. W. Pratt, H. V. Roberts, M. Stone, D. J. Bartholomew; and particularly to D. R. Cox and A. M. Walker who made valuable comments on an early version of the manuscript; and to D. A. East who gave substantially of his time at various stages and generously helped with the proof-reading. Mrs M. V. Bloor and Miss C. A. Davies made life easier by their efficient and accurate typing. I am most grateful to the University Press for the excellence of their printing.

D. V. L.

Aberystwyth
April 1964