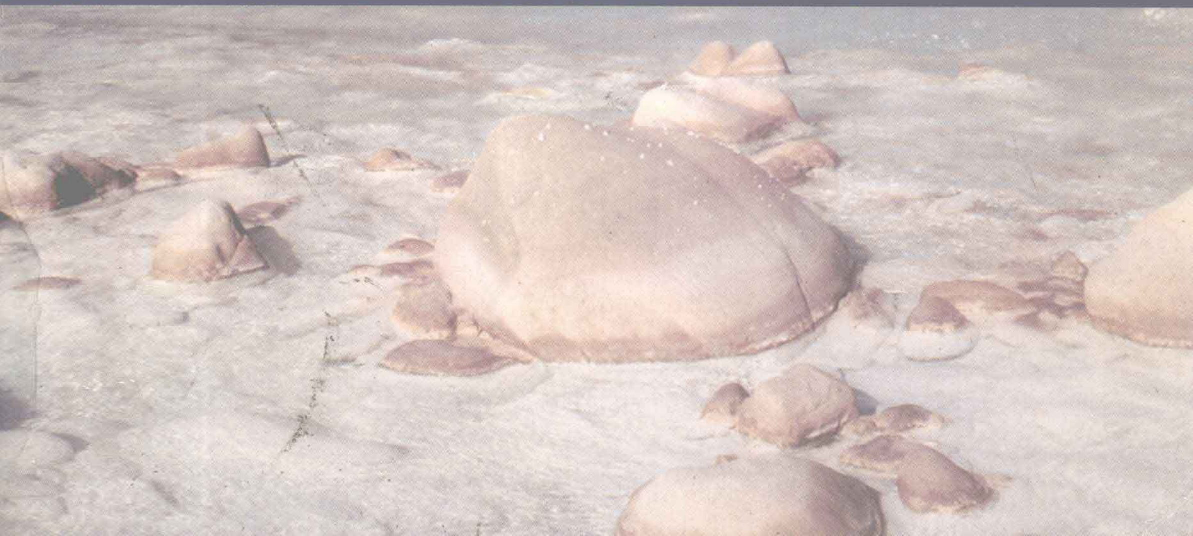
OXFORD

Probability and Random Processes

GEOFFREY GRIMMETT and DAVID STIRZAKER

Third Edition



Probability and Random Processes

GEOFFREY R. GRIMMETT
Statistical Laboratory, University of Cambridge

and

DAVID R. STIRZAKER
Mathematical Institute, University of Oxford

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Athens Auckland Bangkok Bogotá Buenos Aires Cape Town
Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul Karachi
Kolkata Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
Paris São Paulo Shanghai Singapore Taipei Tokyo Toronto Warsaw

with associated companies in Berlin Ibadan

Published in the United States
by Oxford University Press Inc., New York

First edition 1982
Second edition 1992
Third edition 2001

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

Library of Congress Cataloging in Publication Data
Data available

ISBN 0 19 857223 9 [hardback]
ISBN 0 19 857222 0 [paperback]

10 9 8 7 6 5 4 3 2 1

Typeset by the authors
Printed in Great Britain
on acid-free paper by Biddles Ltd, Guildford & King's Lynn

Probability and Random Processes

TEXTS FROM OXFORD UNIVERSITY PRESS

Discrete Mathematics
Norman L. Biggs

Introduction to Algebra
Peter J. Cameron

Mathematical Logic: A Course with Exercises: Parts I and II
René Cori and Daniel Lascar

Probability and Random Processes: Third Edition
Geoffrey R. Grimmett and David R. Stirzaker

One Thousand Exercises in Probability
Geoffrey R. Grimmett and David R. Stirzaker

Probability: An Introduction
Geoffrey Grimmett and Dominic Welsh

A First Course in Coding Theory
Raymond Hill

A Course in Group Theory
John F. Humphries

Invitation to Discrete Mathematics
Jiri Matousek and Jaroslav Nesetril

Visual Complex Analysis
H. A. Priestley

Elementary Geometry
John Roe

Geometry Ancient and Modern
John Silvester

Codes and Cryptography
Dominic Welsh

Lastly, numbers are applicable even to such things as seem to be governed by no rule, I mean such as depend on chance: the quantity of probability and proportion of it in any two proposed cases being subject to calculation as much as anything else. Upon this depend the principles of game. We find sharpers know enough of this to cheat some men that would take it very ill to be thought bubbles; and one gamester exceeds another, as he has a greater sagacity and readiness in calculating his probability to win or lose in any particular case. To understand the theory of chance thoroughly, requires a great knowledge of numbers, and a pretty competent one of Algebra.

John Arbuthnot
An essay on the usefulness of mathematical learning
25 November 1700

To this may be added, that some of the problems about chance having a great appearance of simplicity, the mind is easily drawn into a belief, that their solution may be attained by the mere strength of natural good sense; which generally proving otherwise, and the mistakes occasioned thereby being not infrequent, it is presumed that a book of this kind, which teaches to distinguish truth from what seems so nearly to resemble it, will be looked on as a help to good reasoning.

Abraham de Moivre
The Doctrine of Chances
1717

Preface to the Third Edition

This book provides an extensive introduction to probability and random processes. It is intended for those working in the many and varied applications of the subject as well as for those studying more theoretical aspects. We hope it will be found suitable for mathematics undergraduates at all levels, as well as for graduate students and others with interests in these fields.

In particular, we aim:

- to give a rigorous introduction to probability theory while limiting the amount of measure theory in the early chapters;
- to discuss the most important random processes in some depth, with many examples;
- to include various topics which are suitable for undergraduate courses, but are not routinely taught;
- to impart to the beginner the flavour of more advanced work, thereby whetting the appetite for more.

The ordering and numbering of material in this third edition has for the most part been preserved from the second. However, a good many minor alterations and additions have been made in the pursuit of clearer exposition. Furthermore, we have included new sections on sampling and Markov chain Monte Carlo, coupling and its applications, geometrical probability, spatial Poisson processes, stochastic calculus and the Itô integral, Itô's formula and applications, including the Black–Scholes formula, networks of queues, and renewal–reward theorems and applications. In a mild manifestation of millennial mania, the number of exercises and problems has been increased to exceed 1000. These are not merely drill exercises, but complement and illustrate the text, or are entertaining, or (usually, we hope) both. In a companion volume *One Thousand Exercises in Probability* (Oxford University Press, 2001), we give worked solutions to almost all exercises and problems.

The basic layout of the book remains unchanged. Chapters 1–5 begin with the foundations of probability theory, move through the elementary properties of random variables, and finish with the weak law of large numbers and the central limit theorem; on route, the reader meets random walks, branching processes, and characteristic functions. This material is suitable for about two lecture courses at a moderately elementary level. The rest of the book is largely concerned with random processes. Chapter 6 deals with Markov chains, treating discrete-time chains in some detail (and including an easy proof of the ergodic theorem for chains with countably infinite state spaces) and treating continuous-time chains largely by example. Chapter 7 contains a general discussion of convergence, together with simple but rigorous accounts of the strong law of large numbers, and martingale convergence. Each of these two chapters could be used as a basis for a lecture course. Chapters 8–13 are more fragmented and provide suitable material for about five shorter lecture courses on: stationary processes and ergodic theory; renewal processes; queues; martingales; diffusions and stochastic integration with applications to finance.

We thank those who have read and commented upon sections of this and earlier editions, and we make special mention of Dominic Welsh, Brian Davies, Tim Brown, Sean Collins, Stephen Suen, Geoff Eagleson, Harry Reuter, David Green, and Bernard Silverman for their contributions to the first edition.

Of great value in the preparation of the second and third editions were the detailed criticisms of Michel Dekking, Frank den Hollander, Torgny Lindvall, and the suggestions of Alan Bain, Erwin Bolthausen, Peter Clifford, Frank Kelly, Doug Kennedy, Colin McDiarmid, and Volker Priebe. Richard Buxton has helped us with classical matters, and Andy Burbanks with the design of the front cover, which depicts a favourite confluence of the authors.

This edition having been reset in its entirety, we would welcome help in thinning the errors should any remain after the excellent TEX-ing of Sarah Shea-Simonds and Julia Blackwell.

Cambridge and Oxford G. R. G.
April 2001 D. R. S.

Contents

4 Continuous random variables

5 Generating functions and their applications

6 Markov chains

1

Events and their probabilities

Summary. Any experiment involving randomness can be modelled as a probability space. Such a space comprises a set Ω of possible outcomes of the experiment, a set $\mathcal{F}$ of events, and a probability measure $\mathbb{P}$. The definition and basic properties of a probability space are explored, and the concepts of conditional probability and independence are introduced. Many examples involving modelling and calculation are included.

1.1 Introduction

Much of our life is based on the belief that the future is largely unpredictable. For example, games of chance such as dice or roulette would have few adherents if their outcomes were known in advance. We express this belief in chance behaviour by the use of words such as 'random' or 'probability', and we seek, by way of gaming and other experience, to assign quantitative as well as qualitative meanings to such usages. Our main acquaintance with statements about probability relies on a wealth of concepts, some more reasonable than others. A mathematical theory of probability will incorporate those concepts of chance which are expressed and implicit in common rational understanding. Such a theory will formalize these concepts as a collection of axioms, which should lead directly to conclusions in agreement with practical experimentation. This chapter contains the essential ingredients of this construction.

1.2 Events as sets

Many everyday statements take the form 'the chance (or probability) of A is p', where A is some event (such as 'the sun shining tomorrow', 'Cambridge winning the Boat Race', ...) and p is a number or adjective describing quantity (such as 'one-eighth', 'low', ...). The occurrence or non-occurrence of A depends upon the chain of circumstances involved. This chain is called an *experiment* or *trial*; the result of an experiment is called its *outcome*. In general, we cannot predict with certainty the outcome of an experiment in advance of its completion; we can only list the collection of possible outcomes.

(1) Definition. The set of all possible outcomes of an experiment is called the **sample space** and is denoted by Ω.

(2) Example. A coin is tossed. There are two possible outcomes, heads (denoted by H) and tails (denoted by T), so that $\Omega = \{\text{H}, \text{T}\}$. We may be interested in the possible occurrences of the following events:

(a) the outcome is a head;

(b) the outcome is either a head or a tail;

(c) the outcome is both a head and a tail (this seems very unlikely to occur);

(d) the outcome is not a head. ●

(3) Example. A die is thrown once. There are six possible outcomes depending on which of the numbers 1, 2, 3, 4, 5, or 6 is uppermost. Thus $\Omega = \{1, 2, 3, 4, 5, 6\}$. We may be interested in the following events:

(a) the outcome is the number 1;

(b) the outcome is an even number;

(c) the outcome is even but does not exceed 3;

(d) the outcome is not even. ●

We see immediately that each of the events of these examples can be specified as a subset A of the appropriate sample space Ω. In the first example they can be rewritten as

(a) $A = \{\text{H}\}$, (b) $A = \{\text{H}\} \cup \{\text{T}\}$,

(c) $A = \{\text{H}\} \cap \{\text{T}\}$, (d) $A = \{\text{H}\}^{\text{c}}$,

whilst those of the second example become

(a) $A = \{1\}$, (b) $A = \{2, 4, 6\}$,

(c) $A = \{2, 4, 6\} \cap \{1, 2, 3\}$, (d) $A = \{2, 4, 6\}^{\text{c}}$.

The *complement* of a subset A of Ω is denoted here and subsequently by A^{c}; from now on, subsets of Ω containing a single member, such as $\{\text{H}\}$, will usually be written without the containing braces.

Henceforth we think of *events* as subsets of the sample space Ω. Whenever A and B are events in which we are interested, then we can reasonably concern ourselves also with the events $A \cup B$, $A \cap B$, and A^{c}, representing 'A or B', 'A and B', and 'not A' respectively. Events A and B are called *disjoint* if their intersection is the empty set $\varnothing$; $\varnothing$ is called the *impossible event*. The set Ω is called the *certain event*, since some member of Ω will certainly occur.

Thus events are subsets of Ω, but need all the subsets of Ω be events? The answer is *no*, but some of the reasons for this are too difficult to be discussed here. It suffices for us to think of the collection of events as a subcollection $\mathcal{F}$ of the set of all subsets of Ω. This subcollection should have certain properties in accordance with the earlier discussion:

(a) if $A, B \in \mathcal{F}$ then $A \cup B \in \mathcal{F}$ and $A \cap B \in \mathcal{F}$;

(b) if $A \in \mathcal{F}$ then $A^{\text{c}} \in \mathcal{F}$;

(c) the empty set $\varnothing$ belongs to $\mathcal{F}$.

Any collection $\mathcal{F}$ of subsets of Ω which satisfies these three conditions is called a *field*. It follows from the properties of a field $\mathcal{F}$ that

$$\text{if} \quad A_1, A_2, \ldots, A_n \in \mathcal{F} \quad \text{then} \quad \bigcup_{i=1}^{n} A_i \in \mathcal{F};$$

Typical notation	Set jargon	Probability jargon
Ω	Collection of objects	Sample space
ω	Member of Ω	Elementary event, outcome
A	Subset of Ω	Event that some outcome in A occurs
A^{c}	Complement of A	Event that no outcome in A occurs
$A \cap B$	Intersection	Both A and B
$A \cup B$	Union	Either A or B or both
$A \setminus B$	Difference	A, but not B
$A \triangle B$	Symmetric difference	Either A or B, but not both
$A \subseteq B$	Inclusion	If A, then B
$\varnothing$	Empty set	Impossible event
Ω	Whole space	Certain event

Table 1.1. The jargon of set theory and probability theory.

that is to say, $\mathcal{F}$ is closed under finite unions and hence under finite intersections also (see Problem (1.8.3)). This is fine when Ω is a finite set, but we require slightly more to deal with the common situation when Ω is infinite, as the following example indicates.

(4) Example. A coin is tossed repeatedly until the first head turns up; we are concerned with the number of tosses before this happens. The set of all possible outcomes is the set $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$, where ω_i denotes the outcome when the first $i - 1$ tosses are tails and the ith toss is a head. We may seek to assign a probability to the event A, that the first head occurs after an even number of tosses, that is, $A = \{\omega_2, \omega_4, \omega_6, \dots\}$. This is an infinite countable union of members of Ω and we require that such a set belong to $\mathcal{F}$ in order that we can discuss its probability. ●

Thus we also require that the collection of events be closed under the operation of taking countable unions. Any collection of subsets of Ω with these properties is called a *σ-field*.

(5) Definition. A collection $\mathcal{F}$ of subsets of Ω is called a **σ-field** if it satisfies the following conditions:

(a) $\varnothing \in \mathcal{F}$;
(b) if $A_1, A_2, \dots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$;
(c) if $A \in \mathcal{F}$ then $A^{c} \in \mathcal{F}$.

It follows from Problem (1.8.3) that σ-fields are closed under the operation of taking countable intersections. Here are some examples of σ-fields.

(6) Example. The smallest σ-field associated with Ω is the collection $\mathcal{F} = \{\varnothing, \Omega\}$. ●

(7) Example. If A is any subset of Ω then $\mathcal{F} = \{\varnothing, A, A^{c}, \Omega\}$ is a σ-field. ●

(8) Example. The *power set* of Ω, which is written $\{0, 1\}^{\Omega}$ and contains all subsets of Ω, is obviously a σ-field. For reasons beyond the scope of this book, when Ω is infinite, its power set is too large a collection for probabilities to be assigned reasonably to all its members. ●

To recapitulate, with any experiment we may associate a pair $(\Omega, \mathcal{F})$, where Ω is the set of all possible outcomes or *elementary events* and $\mathcal{F}$ is a σ-field of subsets of Ω which contains all the events in whose occurrences we may be interested; henceforth, to call a set A an *event* is equivalent to asserting that A belongs to the σ-field in question. We usually translate statements about combinations of events into set-theoretic jargon; for example, the event that both A and B occur is written as $A \cap B$. Table 1.1 is a translation chart.

Exercises for Section 1.2

1. Let $\{A_i : i \in I\}$ be a collection of sets. Prove 'De Morgan's Laws'†:

$$\Big(\bigcup_i A_i\Big)^{c} = \bigcap_i A_i^{c}, \qquad \Big(\bigcap_i A_i\Big)^{c} = \bigcup_i A_i^{c}.$$

2. Let A and B belong to some σ-field $\mathcal{F}$. Show that $\mathcal{F}$ contains the sets $A \cap B$, $A \setminus B$, and $A \triangle B$.

3. A conventional knock-out tournament (such as that at Wimbledon) begins with 2^n competitors and has n rounds. There are no play-offs for the positions $2, 3, \ldots, 2^n - 1$, and the initial table of draws is specified. Give a concise description of the sample space of all possible outcomes.

4. Let $\mathcal{F}$ be a σ-field of subsets of Ω and suppose that $B \in \mathcal{F}$. Show that $\mathcal{G} = \{A \cap B : A \in \mathcal{F}\}$ is a σ-field of subsets of B.

5. Which of the following are identically true? For those that are not, say when they are true.
(a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
(b) $A \cap (B \cap C) = (A \cap B) \cap C$;
(c) $(A \cup B) \cap C = A \cup (B \cap C)$;
(d) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

1.3 Probability

We wish to be able to discuss the likelihoods of the occurrences of events. Suppose that we repeat an experiment a large number N of times, keeping the initial conditions as equal as possible, and suppose that A is some event which may or may not occur on each repetition. Our experience of most scientific experimentation is that the proportion of times that A occurs settles down to some value as N becomes larger and larger; that is to say, writing $N(A)$ for the number of occurrences of A in the N trials, the ratio $N(A)/N$ appears to converge to a constant limit as N increases. We can think of the ultimate value of this ratio as being the probability $\mathbb{P}(A)$ that A occurs on any particular trial‡; it may happen that the empirical ratio does not behave in a coherent manner and our intuition fails us at this level, but we shall not discuss this here. In practice, N may be taken to be large but finite, and the ratio $N(A)/N$ may be taken as an approximation to $\mathbb{P}(A)$. Clearly, the ratio is a number between zero and one; if $A = \varnothing$ then $N(\varnothing) = 0$ and the ratio is 0, whilst if $A = \Omega$ then $N(\Omega) = N$ and the

†Augustus De Morgan is well known for having given the first clear statement of the principle of mathematical induction. He applauded probability theory with the words: "The tendency of our study is to substitute the satisfaction of mental exercise for the pernicious enjoyment of an immoral stimulus".

‡This superficial discussion of probabilities is inadequate in many ways; questioning readers may care to discuss the philosophical and empirical aspects of the subject amongst themselves (see Appendix III).

ratio is 1. Furthermore, suppose that A and B are two disjoint events, each of which may or may not occur at each trial. Then

$$N(A \cup B) = N(A) + N(B)$$

and so the ratio $N(A \cup B)/N$ is the sum of the two ratios $N(A)/N$ and $N(B)/N$. We now think of these ratios as representing the probabilities of the appropriate events. The above relations become

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B), \qquad \mathbb{P}(\varnothing) = 0, \qquad \mathbb{P}(\Omega) = 1.$$

This discussion suggests that the probability function $\mathbb{P}$ should be *finitely additive*, which is to say that

$$\text{if } A_1, A_2, \dots, A_n \text{ are disjoint events, then } \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i);$$

a glance at Example (1.2.4) suggests the more extensive property that $\mathbb{P}$ be *countably additive*, in that the corresponding property should hold for countable collections $A_1, A_2, \dots$ of disjoint events.

These relations are sufficient to specify the desirable properties of a probability function $\mathbb{P}$ applied to the set of events. Any such assignment of likelihoods to the members of $\mathcal{F}$ is called a *probability measure*. Some individuals refer informally to $\mathbb{P}$ as a 'probability distribution', especially when the sample space is finite or countably infinite; this practice is best avoided since the term 'probability distribution' is reserved for another purpose to be encountered in Chapter 2.

(1) Definition. A **probability measure** $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ satisfying

(a) $\mathbb{P}(\varnothing) = 0, \quad \mathbb{P}(\Omega) = 1$;

(b) if $A_1, A_2, \dots$ is a collection of disjoint members of $\mathcal{F}$, in that $A_i \cap A_j = \varnothing$ for all pairs i, j satisfying $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$, comprising a set Ω, a σ-field $\mathcal{F}$ of subsets of Ω, and a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, is called a **probability space**.

A probability measure is a special example of what is called a *measure* on the pair $(\Omega, \mathcal{F})$. A measure is a function $\mu : \mathcal{F} \to [0, \infty)$ satisfying $\mu(\varnothing) = 0$ together with (b) above. A measure μ is a probability measure if $\mu(\Omega) = 1$.

We can associate a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with any experiment, and all questions associated with the experiment can be reformulated in terms of this space. It may seem natural to ask for the numerical value of the probability $\mathbb{P}(A)$ of some event A. The answer to such a question must be contained in the description of the experiment in question. For example, the assertion that a *fair* coin is tossed once is equivalent to saying that heads and tails have an equal probability of occurring; actually, this is the definition of fairness.