
ELEMENTARY APPLICATIONS OF PROBABILITY THEORY

H.C. Tuckwell

Chapman and Hall

51.71
T898

Elementary Applications of Probability Theory

Henry C. Tuckwell

Monash University
Victoria
Australia

London New York
CHAPMAN AND HALL

First published in 1988 by Chapman and Hall Ltd

11 New Fetter Lane, London EC4P 4EE
Published in the USA by Chapman and Hall
29 West 35th Street, New York NY 10001

© 1988 H.C. Tuckwell

Printed in Great Britain
by J.W. Arrowsmith, Bristol

ISBN 0 412 30480 5 (cased)
0 412 30490 2 (paperback)

This title is available in both hardbound and paperback editions. The paperback edition is sold subject to the condition that it shall not, by way of trade or otherwise, be lent, resold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

All rights reserved. No part of this book may be reprinted, or reproduced or utilized in any form or by any electronic, mechanical or other means, now known or hereafter invented, including photocopying and recording, or in any information storage and retrieval system; without permission in writing from the publisher.

British Library Cataloguing in Publication Data

Tuckwell, Henry C.
Elementary applications of probability
theory.
1. Probabilities 2. Mathematical statistics
I. Title
519.2 QA273

ISBN 0 412 30480 5
ISBN 0 412 30490 2 Pbk

Library of Congress Cataloging in Publication Data

Tuckwell, Henry C. (Henry Clavering), 1943–
Elementary applications of probability theory.

Bibliography: p.
Includes index.
1. Probabilities. I. Title.
QA273.T84 1988 519.2 87–22407

ISBN 0 412 30480 5
ISBN 0 412 30490 2 (pbk.)

OTHER STATISTICS TEXTS FROM CHAPMAN AND HALL

The Analysis of Time Series

C. Chatfield

Statistics for Technology

C. Chatfield

Introduction to Multivariate Analysis

C. Chatfield and A.J. Collins

Applied Statistics

D.R. Cox and E.J. Snell

An Introduction to Statistical Modelling

A.J. Dobson

Introduction to Optimization Methods and their Application in Statistics

B.S. Everitt

Multivariate Statistics—A Practical Approach

B. Flury and H. Riedwyl

Multivariate Analysis of Variance and Repeated Measures

D.J. Hand and C.C. Taylor

Multivariate Statistical Methods – a primer

Bryan F. Manley

Statistical Methods in Agriculture and Experimental Biology

R. Mead and R.N. Curnow

Elements of Simulation

B.J.T. Morgan

Essential Statistics

D.G. Rees

Decision Analysis: A Bayesian Approach

J.Q. Smith

Applied Statistics: A Handbook of BMDP Analyses

E.J. Snell

Intermediate Statistical Methods

G.B. Wetherill

Further information of the complete range of Chapman and Hall statistics books is available from the publishers.

Preface

This book concerns applications of probability theory. It has been written in the hope that the techniques presented will be useful for problems in diverse areas. A majority of the examples come from the biological sciences but the concepts and techniques employed are not limited to that field. To illustrate, birth and death processes (Chapter 9) have applications to chemical reactions, and branching processes (Chapter 10) have applications in physics but neither of these specific applications is developed in the text.

The book is based on an undergraduate course taught to students who have had one introductory course in probability and statistics. Hence it does not contain a lengthy introduction to probability and random variables, for which there are many excellent books. Prerequisites also include an elementary knowledge of calculus, including first-order differential equations, and linear algebra.

The basic plan of the book is as follows.

Chapter 1: a review of basic probability theory;

Chapters 2–5: random variables and their applications;

Chapter 6: sequences of random variables and concepts of convergence;

Chapters 7–10: theory and properties of basic random processes.

The outline is now given in more detail.

Chapter 1 contains a brief review of some of the basic material which will be needed in later chapters; for example, the basic probability laws, conditional probability, change of variables, etc. It is intended that Chapter 1 be used as a reference rather than a basis for instruction. Students might be advised to study this chapter as the material is called upon.

Chapter 2 illustrates the interplay between geometry and probability. It begins with an historically interesting problem and then addresses the problem of finding the density of the distance between two randomly chosen points. The second such case, when the points occur within a circle, is not easy but the result is useful.

Chapter 3 begins with the properties of the hypergeometric distribution. An important application is developed, namely the estimation of animal

populations by the capture–recapture method. The Poisson distribution is then reviewed and one-dimensional Poisson point processes introduced together with some of their basic properties. There follows a generalization to two dimensions, which enables one to study spatial distributions of plants and to develop methods to estimate their population numbers. The chapter concludes with the compound Poisson distribution which is illustrated by application to a neurophysiological model.

Chapter 4 introduces several of the basic concepts of reliability theory. The relevant properties of the standard failure time distributions are given. The interesting spare parts problem is next and the concluding sections discuss methods for determining the reliability of complex systems.

Chapter 5 commences by explaining the usefulness of computer simulation. There follows an outline of the theory of random number generation using the linear congruential method and the probability integral transformation. The polar method for normal random variables is given. Finally, tests for the distribution and independence properties of random numbers are described.

Chapter 6 deals with sequences of random variables. Some methods for studying convergence in distribution and convergence in probability are developed. In particular, characteristic functions and Chebyshev's inequality are the main tools invoked. The principal applications are to proving a central limit theorem and a weak law of large numbers. Several uses for the latter are detailed.

Chapter 7 starts with the definition of random (stochastic) processes and introduces the important Markov property. The rest of the chapter is mainly concerned with the elementary properties of simple random walks. Included are the unrestricted process and that in the presence of absorbing barriers. For the latter the probability of absorption and the expected time of absorption are determined using the difference equation approach. The concluding section briefly introduces the Wiener process, so fundamental in advanced probability. The concept of martingale and its usefulness are discussed in the exercises.

Chapter 8 is on Markov chains. However, the theory is motivated by examples in population genetics, so the Hardy–Weinberg principle is discussed first. Elementary general Markov chain theory is developed for absorbing Markov chains and those with stationary distributions.

Chapter 9 concerns birth and death processes, which are motivated by demographic considerations. The Poisson process is discussed as a birth process because of its fundamental role. There follow the properties of the Yule process, a simple death process and the simple birth and death process. The treatment of the latter only states rather than derives the equation satisfied by the probability generating function but this enables one to derive the satisfying result concerning the probability of extinction.

Chapter 10 contains a brief introduction to the theory of branching

processes, focusing on the standard Galton–Watson process. It is motivated by the phenomenon of cell division. The mean and variance are derived and the probability of extinction determined.

It should be mentioned that references are sometimes not to the latest editions of books; for example, those of Hoel, Pielou, Strickberger and Watson.

In the author's view there is ample material for a one-quarter or one-semester course. In fact some material might have to be omitted in such a course. Alternatively, the material could be presented in two courses, with a division at Chapter 6, supplemented by further reading in specialist areas (e.g. ecology, genetics, reliability, psychology) and project work (e.g. simulation).

I thank the many Monash students who have taken the course in applied probability on which this book is based. In particular, Derryn Griffiths made many useful suggestions. It is also a pleasure to acknowledge the helpful criticisms of Dr James A. Koziol of Scripps Clinic and Research Foundation, La Jolla; and Drs Fima Klebaner and Geoffrey A. Watterson at Monash University. I am also grateful to Barbara Young for her excellent typing and to Jean Sheldon for her splendid artwork.

Henry C. Tuckwell
Los Angeles, April 1987

Contents

Preface	xi
1 A review of basic probability theory	1
1.1 Probability and random variables	1
1.2 Mean and variance	5
1.3 Conditional probability and independence	5
1.4 Law of total probability	7
1.5 Change of variables	8
1.6 Two-dimensional random variables	10
1.7 Hypothesis testing – the χ^2 goodness of fit test	11
1.8 Notation	13
References	14
2 Geometric probability	16
2.1 Buffon's needle problem	16
2.2 The distance between two random points on a line segment	19
2.3 The distance between two points dropped randomly in a circle	21
2.4 Sum of two random variables	25
References	27
Exercises	27
3 Some applications of the hypergeometric and Poisson distributions	30
3.1 The hypergeometric distribution	30
3.2 Estimating a population from capture–recapture data	33
3.3 The Poisson distribution	37
3.4 Homogeneous Poisson point process in one dimension	38
3.5 Occurrence of Poisson processes in Nature	41
3.6 Poisson point processes in two dimensions	44
3.7 Compound Poisson random variables	48
3.8 The delta function	50

3.9	An application in neurobiology	52
	References	56
	Exercises	57
4	Reliability theory	61
4.1	Failure time distributions	61
4.2	Reliability function and failure rate function	63
4.3	The spare parts problem	69
4.4	Complex systems	70
4.5	Series and parallel systems	72
4.6	Combinations and other structures	75
	Further reading	77
	References	77
	Exercises	77
5	Simulation and random numbers	81
5.1	The need for simulation	81
5.2	The usefulness of a random sample from a uniform distribution	83
5.3	Generation of uniform $(0, 1)$ random numbers	86
5.4	Generation of random numbers from a normal distribution	88
5.5	Statistical tests for random numbers	90
5.6	Testing for independence	92
	References	96
	Exercises	96
6	Convergence of sequences of random variables: the central limit theorem and the laws of large numbers	98
6.1	Characteristic functions	98
6.2	Examples	101
6.3	Convergence in distribution	104
6.4	The central limit theorem	107
6.5	The Poisson approximation to the binomial distribution	110
6.6	Convergence in probability	111
6.7	Chebyshev's inequality	113
6.8	The weak law of large numbers	115
	References	119
	Exercises	119
7	Simple random walks	123
7.1	Random processes – definitions and classifications	123
7.2	Unrestricted simple random walk	126
7.3	Random walk with absorbing states	131

7.4	The probabilities of absorption at 0	132
7.5	Absorption at $c > 0$	137
7.6	The case $c = \infty$	138
7.7	How long will absorption take?	139
7.8	Smoothing the random walk – the Wiener process and Brownian motion	142
	References	145
	Exercises	145
8	Population genetics and Markov chains	148
8.1	Genes and their frequencies in populations	148
8.2	The Hardy–Weinberg principle	150
8.3	Random mating in finite populations: a Markov chain model	153
8.4	General description of Markov chains	154
8.5	Temporally homogeneous Markov chains	155
8.6	Random genetic drift	158
8.7	Markov chains with absorbing states	160
8.8	Absorption probabilities	162
8.9	The mean time to absorption	167
8.10	Mutation	171
8.11	Stationary distributions	173
8.12	Approach to a stationary distribution as $n \rightarrow \infty$	174
	References	178
	Exercises	179
9	Population growth I: birth and death processes	183
9.1	Introduction	183
9.2	Simple Poisson processes	185
9.3	Markov chains in continuous time	187
9.4	The Yule process	188
9.5	Mean and variance for the Yule process	192
9.6	A simple death process	194
9.7	Simple birth and death process	196
9.8	Mean and variance for the birth and death process	199
	References	201
	Exercises	201
10	Population growth II: branching processes	204
10.1	Cell division	204
10.2	The Galton–Watson branching process	205
10.3	Mean and variance for the Galton–Watson process	207
10.4	Probability generating functions of sums of random variables	209

x	Contents	
10.5	The probability of extinction	212
	References	216
	Exercises	217
	<i>Appendix</i> Table of critical values of the χ^2 -distribution	219
	Index	220

1

A review of basic probability theory

This is a book about the applications of probability. It is hoped to convey that this subject is both a fascinating and important one. The examples are drawn mainly from the biological sciences but some originate in the engineering, physical, social and statistical sciences. Furthermore, the techniques are not limited to any one area.

The reader is assumed to be familiar with the elements of probability or to be studying it concomitantly. In this chapter we will briefly review some of this basic material. This will establish notation and provide a convenient reference place for some formulas and theorems which are needed later at various points.

1.1 PROBABILITY AND RANDOM VARIABLES

When an experiment is performed whose outcome is uncertain, the collection of possible **elementary outcomes** is called a **sample space**, often denoted by Ω . Points in Ω , denoted in the discrete case by ω_i , $i = 1, 2, \dots$ have an associated probability $P\{\omega_i\}$. This enables the probability of any subset A of Ω , called an **event**, to be ascertained by finding the total probability associated with all the points in the given subset:

$$P\{A\} = \sum_{\omega_i \in A} P\{\omega_i\}$$

We always have

$$0 \leq P\{A\} \leq 1,$$

and in particular $P\{\Omega\} = 1$ and $P\{\emptyset\} = 0$, where \emptyset is the empty set relative to Ω .

A **random variable** is a real-valued function defined on the elements of a sample space. Roughly speaking it is an observable which takes on numerical values with certain probabilities.

Discrete random variables take on finitely many or countably infinitely many values. Their probability laws are often called **probability mass functions**. The following discrete random variables are frequently encountered.

2 Basic probability theory

Binomial

A **binomial** random variable X with parameters n and p has the probability law

$$p_k = \Pr \{X = k\} = \binom{n}{k} p^k q^{n-k} \quad (1.1)$$

$$\doteq b(k; n, p), \quad k = 0, 1, 2, \dots, n,$$

where $0 \leq p \leq 1$, $q = 1 - p$ and n is a positive integer (\doteq means we are defining a new symbol). The **binomial coefficients** are

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

being the number of ways of choosing k items, without regard for order, from n distinguishable items.

When $n = 1$, so we have

$$\Pr \{X = 1\} = p = 1 - \Pr \{X = 0\},$$

the random variable is called **Bernoulli**.

Note the following.

Convention

Random variables are always designated by capital letters (e.g. X, Y) whereas symbols for the values they take on, as in $\Pr \{X = k\}$, are always designated by lowercase letters.

The converse, however, is not true. Sometimes we use capital letters for non-random quantities.

Poisson

A **Poisson** random variable with parameter $\lambda > 0$ takes on non-negative integer values and has the probability law

$$p_k = \Pr \{X = k\} = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots \quad (1.2)$$

For any random variable the total probability mass is unity. Hence if p_k is given by either (1.1) or (1.2),

$$\sum_k p_k = 1$$

where summation is over the possible values k as indicated.

For any random variable X , the **distribution function** is

$$F(x) = \Pr \{X \leq x\}, \quad -\infty < x < \infty.$$

Continuous random variables take on a continuum of values. Usually the probability law of a continuous random variable can be expressed through its **probability density function**, $f(x)$, which is the derivative of the distribution function. Thus

$$\begin{aligned} f(x) &= \frac{d}{dx} F(x) \\ &= \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\Pr \{X \leq x + \Delta x\} - \Pr \{X \leq x\}}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\Pr \{x < X \leq x + \Delta x\}}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{\Pr \{X \in (x, x + \Delta x]\}}{\Delta x} \end{aligned} \quad (1.3)$$

The last two expressions in (1.3) often provide a convenient prescription for calculating probability density functions. Often the latter is abbreviated to p.d.f. but we will usually just say ‘density’.

If the interval (x_1, x_2) is in the range of X then the probability that X takes values in this interval is obtained by integrating the probability density over (x_1, x_2) .

$$\Pr \{x_1 < X < x_2\} = \int_{x_1}^{x_2} f(x) dx.$$

The following continuous random variables are frequently encountered.

Normal (or Gaussian)

A random variable with density

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}, \quad -\infty < x < \infty, \quad (1.4)$$

where $-\infty < \mu < \infty$ and $0 < \sigma^2 < \infty$,

is called **normal**. The quantities μ and σ^2 are the mean and variance (elaborated upon below) and such a random variable is often designated

4 Basic probability theory

$N(\mu, \sigma)$. If $\mu = 0$ and $\sigma = 1$ the random variable is called a **standard normal** random variable, for which the usual symbol is Z .

Uniform

A random variable with constant density

$$f(x) = \frac{1}{b-a}, \quad -\infty < a \leq x \leq b < \infty,$$

is said to be **uniformly distributed** on (a, b) and is denoted $U(a, b)$. If $a = 0, b = 1$ the density is unity on the unit interval,

$$f(x) = 1, \quad 0 \leq x \leq 1$$

and the random variable is designated $U(0, 1)$.

Gamma

A random variable is said to have a **gamma density** (or gamma distribution) with parameters λ and ρ if

$$f(x) = \frac{\lambda(\lambda x)^{\rho-1} e^{-\lambda x}}{\Gamma(\rho)}, \quad x \geq 0; \quad \lambda, \rho > 0.$$

The quantity $\Gamma(\rho)$ is the **gamma function** defined as

$$\Gamma(\rho) = \int_0^{\infty} x^{\rho-1} e^{-x} dx, \quad \rho > 0.$$

When $\rho = 1$ the gamma density is that of an **exponentially distributed random variable**

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

For continuous random variables the density must integrate to unity:

$$\int f(x) dx = 1$$

where the interval of integration is the whole range of values of X .

1.2 MEAN AND VARIANCE

Let X be a discrete random variable with

$$\Pr \{X = x_k\} = p_k, \quad k = 1, 2, \dots$$

The **mean, average** or **expectation** of X is

$$E(X) = \sum_k p_k x_k.$$

For a binomial random variable $E(X) = np$ whereas a Poisson random variable has mean $E(X) = \lambda$.

For a continuous random variable with density $f(x)$,

$$E(X) = \int x f(x) dx.$$

If X is normal with density given by (1.4) then $E(X) = \mu$; a uniform (a, b) random variable has mean $E(X) = \frac{1}{2}(a + b)$; and a gamma variate has mean $E(X) = \rho/\lambda$.

The **n th moment** of X is the expected value of X^n :

$$E(X^n) = \begin{cases} \sum_k p_k x_k^n & \text{if } X \text{ is discrete,} \\ \int x^n f(x) dx & \text{if } X \text{ is continuous.} \end{cases}$$

If $n = 2$ we obtain the **second moment** $E(X^2)$. The **variance**, which measures the degree of dispersion of the probability mass of a random variable about its mean, is

$$\begin{aligned} \text{Var}(X) &= E[(X - E(X))^2] \\ &= E(X^2) - E^2(X). \end{aligned}$$

The variances of the above-mentioned random variables are:

binomial, npq ; Poisson, λ ; normal, σ^2 ; uniform, $\frac{1}{12}(b - a)^2$; gamma, ρ/λ^2 .

The square root of the variance is called the **standard deviation**.

1.3 CONDITIONAL PROBABILITY AND INDEPENDENCE

Let A and B be two random events. The **conditional probability** of A given B is, provided $\Pr \{B\} \neq 0$,

$$\Pr\{A|B\} = \frac{\Pr\{AB\}}{\Pr\{B\}}$$

where AB is the **intersection** of A and B , being the event that both A and B occur (sometimes written $A \cap B$). Thus only the occurrences of A which are simultaneous with those of B are taken into account. Similarly, if X, Y are random variables defined on the same sample space, taking on values $x_i, i = 1, 2, \dots, y_j, j = 1, 2, \dots$, then the conditional probability that $X = x_i$ given $Y = y_j$ is, if $\Pr\{Y = y_j\} \neq 0$,

$$\Pr\{X = x_i | Y = y_j\} = \frac{\Pr\{X = x_i, Y = y_j\}}{\Pr\{Y = y_j\}},$$

the comma between $X = x_i$ and $Y = y_j$ meaning ‘and’.

The **conditional expectation** of X given $Y = y_j$ is

$$E(X | Y = y_j) = \sum_i x_i \Pr\{X = x_i | Y = y_j\}.$$

The **expected value** of XY is

$$E(XY) = \sum_{i,j} x_i y_j \Pr\{X = x_i, Y = y_j\},$$

and the **covariance** of X, Y is

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E(XY) - E(X)E(Y). \end{aligned}$$

The covariance is a measure of the linear dependence of X on Y .

If X, Y are independent then the value of Y should have no effect on the probability that X takes on any of its values. Thus we define X, Y as **independent** if

$$\Pr\{X = x_i | Y = y_j\} = \Pr\{X = x_i\}, \quad \text{all } i, j.$$

Equivalently X, Y are independent if

$$\Pr\{X = x_i, Y = y_j\} = \Pr\{X = x_i\} \Pr\{Y = y_j\},$$

with a similar formula for arbitrary independent events.

Hence for independent random variables

$$E(XY) = E(X)E(Y),$$

so their covariance is zero. Note, however, that $\text{Cov}(X, Y) = 0$ does not always imply X, Y are independent. The covariance is often normalized by defining the **correlation coefficient**

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$