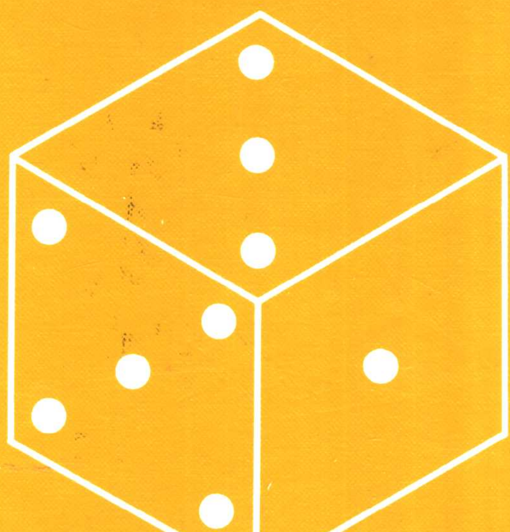
SPRINGER TEXTS IN STATISTICS

Gunnar Blom

Probability and Statistics

Theory and Applications



Springer-Verlag

Gunnar Blom

Probability and Statistics

Theory and Applications

With 107 Illustrations

Springer-Verlag
New York Berlin Heidelberg
London Paris Tokyo

Gunnar Blom
Department of Mathematical Statistics
University of Lund and
Lund Institute of Technology
S-221 00 Lund
Sweden

Mathematics Subject Classifications (1980): 60-XX, 60-01, 62-01

Library of Congress Cataloging-in-Publication Data
Blom, Gunnar.
[Sannolikhetsteori och statistikteori med tillämpningar. English]
Probability and statistics : theory and applications / Gunnar Blom.
p. cm. — (Springer texts in statistics)
Translation of: Sannolikhetsteori och statistikteori med tillämpningar. Lund, 1980.
Bibliography: p.
Includes index.
ISBN 0-387-96852-0 (alk. paper)
1. Probabilities. 2. Mathematical statistics. I. Title. II. Series.
QA273.B592413 1989 88-29766
519.2—dc19 CIP

Original Swedish edition: *Sannolikhetsteori och statistikteori med tillämpningar*. Studentlitteratur, Lund, 1980.

Printed on acid-free paper

Typeset by Asco Trade Typesetting Ltd., Hong Kong.
Printed and bound by R.R. Donnelley & Sons, Harrisonburg, Virginia.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-96852-0 Springer-Verlag New York Berlin Heidelberg
ISBN 3-540-96852-0 Springer-Verlag Berlin Heidelberg New York

Preface

This is a somewhat extended and modified translation of the third edition of the text, first published in 1969. The Swedish edition has been used for many years at the Royal Institute of Technology in Stockholm, and at the School of Engineering at Linköping University. It is also used in elementary courses for students of mathematics and science.

The book is not intended for students interested only in theory, nor is it suited for those seeking only statistical recipes. Indeed, it is designed to be intermediate between these extremes. I have given much thought to the question of dividing the space, in an appropriate way, between mathematical arguments and practical applications. Mathematical niceties have been left aside entirely, and many results are obtained by analogy.

The students I have in mind should have three ingredients in their course: elementary probability theory with applications, statistical theory with applications, and something about the planning of practical investigations. When pouring these three ingredients into the soup, I have tried to draw upon my experience as a university teacher and on my earlier years as an industrial statistician.

The programme may sound bold, and the reader should not expect too much from this book. Today, probability, statistics and the planning of investigations cover vast areas and, in 356 pages, only the most basic problems can be discussed. If the reader gains a good understanding of probabilistic and statistical reasoning, the main purpose of the book has been fulfilled.

Certain sections of the book may be omitted, as noted in the text. However, the reader is advised not to skip them all, for they are meant as spices in the soup.

A working knowledge of elementary calculus, in particular derivatives and Riemann integrals, is an essential prerequisite.

Many theoretical and applied problems are solved in the text. Exercises with answers are given at the end of most chapters. Most of the exercises are easy, many very easy, and are thought to be especially suitable for those using the book for self-study. For the benefit of those who like more difficult exercises there are also some rather hard problems indicated by asterisks. Furthermore, there is a collection of 80 problems of varying difficulty at the end of the book. Readers who solve the majority of these will have acquired a good knowledge of the subject.

Numerous references to the literature are given at the end of the book. They are meant to help and guide those students who wish to extend the knowledge they have acquired.

The tables at the end of the book are taken from a separate collection of tables published by AB Studentlitteratur in Lund, Sweden.

My sincere thanks are due to Herbert A. David and Theodore A. Vessey for their invaluable help with the translation. I also want to express my gratitude to Krzysztof Nowicki and Dennis Sandell for checking the answers of many exercises and to Jan-Eric Englund, Jan Lanke and Dennis Sandell for their assistance with the proof-reading.

Lund, Sweden G.B.
November, 1988

Contents

[1] Special section, which may be omitted.

[1] Special section, which may be omitted.

[1] Special section, which may be omitted.

CHAPTER 1

Introduction to Probability Theory

1.1. On the Usefulness of Probability Theory

We shall give three examples intended to show how probability theory may be used in practice.

Example 1. Batch with Defective Units

A car manufacturer buys certain parts (units) from a subcontractor. From each batch of 1,000 units a sample of 75 is inspected. If two or fewer defective units are found, the whole batch is accepted; otherwise, it is returned to the supplier.

Let us assume that there are 20 defective units in the whole batch. What is the probability that the batch will pass inspection? Such a question is of interest, for the answer gives the car manufacturer an idea of the efficiency of the inspection. Probability theory provides the answer. □

Example 2. Measuring a Physical Constant

At a laboratory, five measurements were made of a physical constant. The following values were obtained:

2.13 2.10 2.05 2.11 2.14.

Because of errors in measurement, there is a certain uncontrollable variation among the values, which can be termed random.

In order to find out how much the measurements deviate from the true value of the constant, the following question is natural: What is the probability that the arithmetic mean of the five measurements differs from the true value by more than a certain given quantity? To answer this question, one needs probability theory. □

Example 3. Animal Experiment

A research worker at a pharmaceutical industry gives 12 rabbits an injection of a certain dose of insulin. After a certain time has elapsed, the decrease in the percentage of blood sugar is determined. The following result is obtained (computed as a percentage decrease of the initial value):

11.2	21.2	−4.0	18.7	2.8	27.2
25.1	25.8	2.2	28.3	23.7	−2.2.

There is considerable variation among the data, as is generally the case when biological phenomena are studied. Look at the extreme values: one animal has 28% *less* blood-sugar than before the injection, and another has 4% *higher* blood-sugar than before. The variability is large, in spite of the fact that the animals have been reared in the same way and have also been treated alike in all other respects. For various reasons, the rabbits react very differently to the same dose. Nevertheless, important information can be extracted from such animal experiments. For this purpose, probability theory is again required. □

A common feature of the three examples, and of all other situations investigated by probability theory, is that *variability* is present. In Example 1 the number of defective units varies from sample to sample, if the inspection is repeated several times by taking new samples from the same batch. In Example 2, there are variations from one measurement to the next, and in Example 3, from one animal to the next.

Variability is a very general phenomenon. The variations can be small, as among cog-wheels manufactured with great precision or between identical twins "as like as two peas in a pod"; or large, as among the stones in Scandinavian moraines or among the profits in dollars of all the firms in Western Europe during a certain year. Differences among people, physical and psychological, and variations with regard to geography, mineral resources, weather and economy, affect the world in a fundamental way. Knowledge of such things is often important for the actions of individuals and society.

It can be disastrous to neglect variability. Only one, almost trivial example, will be mentioned. If you plan to wade across a river, it is not enough to know that the average depth is 3 feet; it is necessary to know something about the variations in depth!

Because of the omnipresence of variability, it is natural that probability theory—the basic science of variability—has a multitude of applications in technological research, physics, chemistry, biology, medicine and the geosciences. In recent years the interest in applied probability theory has increased considerably, in view of the investigation of models for automatic control, queueing models, models for production and epidemiological models (models for the spread of diseases). Probability theory is also of basic importance for statistical theory.

It should be added that probability theory is not only a useful, but also an enjoyable and elegant science. The enjoyment will be experienced by anyone interested in games of chance, stakes and lotteries, upon discovering that probability theory permits the solution of many associated problems. Many research problems in probability can also give one, with the right mathematical ability and interest, an opportunity for both great effort and great satisfaction.

1.2. Models, Especially Random Models

The concept of a *model* is of basic importance in science. A model is intended to describe essential properties of a phenomenon, without copying all of the details. A fundamental, but common, error is to confuse the model with the reality which it is intended to depict. It is very important to keep in mind that every model is approximate.

A good motto for anyone dealing with models is therefore: *Distinguish between the model and reality.*

Models are of different kinds. There are *physical* models (Example: a house built to the scale 1:50), *analogue* models (Examples: map, drawing, atomic model with balls on rods), and *abstract* models.

Abstract models are often used in mathematical and experimental sciences. Such models belong to the world of mathematics and can be formulated in mathematical terms. An abstract model can be *deterministic* or *stochastic.* Instead of a stochastic model we mostly say a *random* model; the term *chance* model is also used. The word stochastic derives from the Greek "stochastikos", able to aim at, to hit (the mark); "stochos" means target, conjecture.

In a deterministic model a phenomenon is approximated by mathematical functions. For example, the area of a round table is determined by conceptually replacing the table-top by a circle; we then obtain πr^2, where r is the radius of the circle. Clearly, this is a model; exactly circular tables exist only in an abstract world.

Let us consider a model with more far-reaching implications. A widely debated deterministic model of great consequence, for individuals and their views of life, presupposes that information about the positions and movements of all atoms at a given moment makes it possible to predict the whole subsequent development of the world.

Simpler, but important, examples of deterministic models are Euclidean geometry which was used for measuring areas of land in ancient Egypt and classical mechanics which made it possible for Newton to explain the movements of the planets around the sun.

In the following we consider only random models. Such models are used in probability theory for describing *random trials* (often called random experiments). By that we mean, somewhat vaguely, every experiment that can be repeated under similar conditions, but whose outcome cannot be predicted exactly in advance, even when the same experiment has been performed many times. It is just the unpredictable variation of the outcomes which is described by a random model.

Let us take a classical example of a random experiment. Throw a die repeatedly and each time note the number of points obtained. Suppose the result is 1, 5, 1, 6, 3, 2, 4, 6, 1. Clearly, this may be termed a random experiment, only if we perform the throws appropriately so that the outcome cannot be predicted. For example, if the die is held just above the table each time, the outcome is determined in advance. It follows from this simple remark that certain conditions have to be fulfilled by a random trial, and also that these conditions cannot be strictly formulated.

Another example of a random trial is afforded by radioactive decay. The result of a famous series of such experiments was published at the beginning of this century by Rutherford and Geiger. For a certain radioactive source they recorded the number of atoms that decayed in successive minutes, for example, 24, 31, 29, 16, 19, 25, ...

Before probability theory can be applied to such trials as we have described here, it is necessary to construct a random model which shows how the result may vary. In our examples, the models are rather simple, but sometimes complicated models are required; for example, when industrial processes are regulated using principles of automatic control.

In this book, we discuss some basic properties of certain important random models, and in this way lay a foundation for probability theory.

1.3. Some Historical Notes

Probability theory has its roots in sixteenth-century Italy and seventeenth-century France and was originally a theory for games of chance. At the beginning of the sixteenth century, the Italian Cardano wrote a treatise about such games, in which he solved several problems of probability. Beginning in 1654, the famous correspondence between Pascal and Fermat occurred concerning certain games of chance. James Bernoulli and de Moivre are also pioneers in the field of probability, in a somewhat later time-period. Laplace, in his *Théorie Analytique des Probabilités*, published in 1812, described probability theory as a mathematical theory of wide scope and extended its application to astronomy and mechanics. Several physicists, among them Maxwell and Boltzmann, used probability theory in the creation of statistical mechanics.

Only in our own time has probability theory become a science, in the modern sense of the word. A large number of mathematicians and probabilists have contributed to its development, among them many Russians such as Chebyshev, Markov, Liapounov and Kolmogorov. Important contributions to the theory have also been made by the Frenchman Lévy, the Austrian von Mises, the Americans Doob and Feller, and the Swede Cramér.

CHAPTER 2

Elements of Probability Theory

2.1. Introduction

This chapter is important. In order to come to grips with it rapidly, the reader is advised, on the first reading, to skip all remarks and most of the examples in the text. It is a good idea to read the chapter several times, in order to lay a solid foundation for what follows.

The most important sections are §2.2 and §2.3, which deal, respectively, with events and with the definition of probability; in the latter section, Kolmogorov's axioms are of particular importance. In §2.4 the classical definition of probability is introduced, and we give several examples of probabilities in discrete sample spaces. In §2.5 and §2.6 we discuss conditional probability and the important concept of independent events. §2.7 contains some results from combinatorics. The last section, §2.8, deals with some classical problems of probability.

2.2. Events

We begin with three definitions.

Definition. The result of a random trial is called an *outcome*. □

Definition. The set of all possible outcomes is called the *sample space*. □

Definition. A collection of outcomes is called an *event*. □

Thus an event is a subset of the sample space (or possibly the whole sample space).

Outcomes are denoted by numbers or letters, in this book often by u_1, u_2, ... The sample space is denoted by Ω (omega). Events are denoted by capital letters A, B, C, ...

Example 1. Throwing Dice

(a) Throwing a single die

Let us throw a die. We introduce six outcomes, which we denote by 1, 2, ..., 6. The sample space Ω consists of these outcomes and can be written $\Omega = \{1, 2, \ldots, 6\}$. Examples of events are A = "number of points is odd" and B = "number of points is at most 2". Using the symbols for the outcomes, we can write these events

$$A = \{1, 3, 5\}; \qquad B = \{1, 2\}.$$

Note that we have only assumed that the die has six sides; these sides need not necessarily be alike.

(b) Throwing two dice

Let us throw two dice at the same time. It is then convenient to introduce 36 outcomes $u_1, \ldots, u_{36}$. These outcomes can also be written (1, 1), (1, 2), ..., (6, 6), where the first digit represents one die and the second digit the other. (We suppose that the dice have different colours so that they are distinguishable.) The sample space consists of these 36 outcomes. As an example of an event we can take A = "the sum of the two dice is at most 3"; this event can be written

$$A = \{(1, 1), (1, 2), (2, 1)\}.$$

Hence the event A contains three outcomes.

A figure or diagram is often useful in illustrating how Ω and A look (see Fig. 2.1, where A is represented by the enclosed points). □

Example 2. Radioactive Decay

A sample of a radioactive substance is available. A Geiger–Müller counter registers the number of particles decaying during a certain time interval. This number can be 0, 1, 2, ..., which are all possible outcomes. Thus the sample space contains a denumerably infinite number of outcomes. □

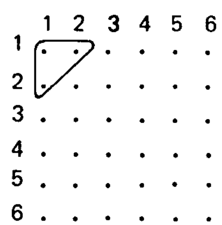


Fig. 2.1.

Example 3. Strength of Material

From a batch of reinforcing bars a certain number are taken, and their tensile strengths are determined. The result (in kp/mm^2) is, say, 62.4, 69.6, 65.0, 63.1, 76.8, ...

The possible values (= outcomes) are, at least theoretically, all numbers between 0 and ∞. We therefore let the sample space consist of all nonnegative numbers. It is impossible to write all these outcomes as one sequence. The event A might be, for example, "tensile strength of a chosen bar is between 60.0 and 65.0".

In practice, the tensile strength is stated with a certain number of decimal places, perhaps with one decimal place as we did above. Then Ω consists only of the numbers 0.0, 0.1, 0.2, ..., and the event A consists of just a finite number of outcomes 60.0, 60.1, ..., 65.0. From a mathematical point of view, it is generally easier to disregard the discreteness and to retain the continuous character of the sample space. □

We shall give a further definition.

Definition. If the number of outcomes is finite or denumerably infinite, Ω is said to be a *discrete* sample space. More particularly, if the number is finite, Ω is said to be a *finite* sample space.

If the number of outcomes is neither finite nor denumerably infinite, Ω is said to be a *continuous* sample space. □

In Examples 1 and 2 the sample space is discrete, in Example 1 it is actually finite; in Example 3 it is continuous.

Example 4. Coin Tosses

A coin is tossed repeatedly. Let a tail be denoted by 0 and a head by 1.

(a) A given number of tosses

The result of n tosses can be represented by a binary n-digit number. (For example, if $n = 3$, the possible outcomes are 000, 001, 010, 011, 100, 101, 110, 111.) We let each such number be an outcome. The number of possible outcomes is 2^n. As an illustration we use a "tree diagram" (a fallen tree); see Fig. 2.2 for the case $n = 3$.

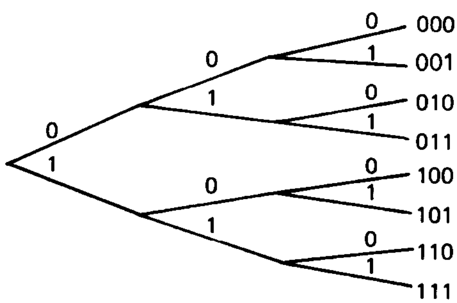


Fig. 2.2. Three tosses with a coin.