PROBABILITY

An Introduction

SAMUEL GOLDBERG

PROBABILITY

An Introduction

SAMUEL GOLDBERG

DEPARTMENT OF MATHEMATICS
OBERLIN COLLEGE

PRENTICE-HALL, INC.

ENGLEWOOD CLIFFS, N. J.

Library of Congress Catalog Card Number: 60-10871

First printing .... *April, 1960*
Second printing .... *April, 1961*
Third printing .... *February, 1962*

PRINTED IN THE UNITED STATES OF AMERICA

71158—C

PREFACE

THIS BOOK is intended for all who require a mathematically sound, but elementary introduction to the theory of probability.

Probability concepts are now of great importance in a wide variety of fields. The theory of probability, as the foundation upon which the methods of statistics are based, should command the attention of those who want to understand as well as apply statistical techniques. Probabilistic theories, making explicit reference to the nature and effects of chance phenomena, are the rule rather than the exception in the physical and biological sciences. Less well known is the fact that probability concepts are finding increased use in the social sciences and business: psychologists develop stochastic models for learning; economists use the techniques of game theory to discuss competition and markets; expected values, variances, and other matters related to random variables turn out to be important in the problem of finding combinations of securities that best meet the needs of the investor; business managers, because their decisions must be made in the face of uncertainty, invoke the theory of probability as an aid in planning inventory, establishing quality control, designing market surveys, etc. We need not go on—it is clear that probability concepts and methods are now widely used and will see even more extensive use in the future.

One noteworthy indication of the importance of our subject is the recent decision of the Commission on Mathematics of the College Entrance Examination Board to recommend that a course in probability and statistical inference be offered in the twelfth grade of the secondary school. Thus, secondary school teachers of mathematics, at some point in their college or in-service training, or in summer institutes (such as those sponsored by the National Science Foundation), should achieve some mastery of the elements of probability theory. Parts of this book were used in courses offered in NSF Institutes held at Oberlin College in 1958 and 1959, and the final manuscript has benefited from the helpful comments of the many teachers who studied preliminary versions.

Although there are a number of excellent textbooks on probability, they are all written for readers who have the mathematical sophistication that comes with a working knowledge of the differential and integral calculus. It seemed to me worthwhile to bring the theory of probability to the attention of those who do not have the calculus prerequisite. It was with this aim in mind that I limited myself to those topics that are accessible to readers with only a good background in high school algebra and a little ability in the reading and manipulation of mathematical symbols. The consequent limitation to finite sample spaces, although severe, facilitates a careful logical treatment of the essentials needed by all who use probability concepts. Furthermore, I have found that an understanding of the basic definitions, theorems, and methods in the finite case makes it much easier for students with the necessary preparation to master the corresponding ideas in the infinite case. I am therefore hopeful that this volume, although written as a basic textbook for courses in probability and statistics for students without calculus, will also prove useful in courses for those who have previous training in calculus.

One further possible use of this book is worthy of mention here. There are many college students who, for one reason or another, can take at most one year of mathematics. These students are often offered a smorgasbord survey course in which they sample one topic after another and learn very little about lots of things. Many teachers, however, prefer to offer a course centering on a few main topics, going into each systematically and deeply enough to give the student a reasonable depth of knowledge in the chosen subjects. Although many topics vie for inclusion in such a program, I believe a strong case can be made for a course that concentrates on sets and probability in the finite case at first, proceeds to an introduction to the calculus, and then applies this calculus to the elements of probability in the infinite case. (In my own course, I also include applications of the differential calculus to simple problems in economics.) Such a course, if properly executed, can give the student a keen sense of the nature and achievements of mathematical thinking, while laying a firm foundation for further study in economics, statistics, operations research, or allied fields. Such a program would therefore be especially valuable for social science and business students, assuming they can devote only a year to mathematics at the college level. I have used this volume in preliminary form in roughly the first third of such a year course at Oberlin College, with students who present less than three years of high school mathematics for entrance. Teachers who share my point of view may also find this book useful in their own introductory mathematics courses.

Since the theory of probability is best formulated using the language and notation of sets, we devote the first chapter to the elementary mathematics of sets. Proofs of laws in the algebra of sets are simplified by the use of so-called membership tables, a device analagous to truth tables in logic. Here we also introduce Cartesian product sets, which are needed at many points throughout the book.

Chapter 2 develops the basic calculus of probability for experiments with only a finite number of possible outcomes (finite sample spaces). A probability measure is first introduced over the events of a sample space and then conditional probability, independent events, and independent trials are carefully defined. Illustrative and problem material is here limited to the simplest experimental situations, and more sophisticated combinatorial techniques are first treated in Chapter 3. The usual order of topics has been reversed because beginning students seem always to have difficulty with the use of permutation and combination formulas, and this difficulty often impairs the learning of the basic probability ideas when both are presented simultaneously. We present the basic ideas first and then, in Chapter 3, offer a set of exercises in which the previously mastered probability theory is applied to a wide variety of situations requiring the use of sophisticated counting techniques. It has been our experience that this procedure makes it considerably easier for the student to learn this basic material.

Chapter 4 is an introduction to the analytic theory of probability in the finite case. Random variables are defined as functions on sample spaces, and probability distributions, means, standard deviations, joint probability functions, covariance, and correlation are discussed. Independence of random variables is defined and, with these ideas extended to the multivariate case, applications to random sampling theory can be included. The sampling distribution of the sample mean is discussed and formulas for its mean and variance are derived for both sampling with and without replacement.

The most important probability function defined on a finite sample space, the binomial distribution, forms the subject matter of the final chapter. The basic properties of a Bernoulli process and a binomially distributed random variable are derived, and the use of tables of cumulative binomial probabilities is discussed. Applications to the testing of statistical hypotheses (significance tests), as well as to a more complex problem of decision-making under uncertainty serve to illustrate how probability methods are applied in statistical investigations.

For some classes, teachers may find it necessary to offer supplementary lessons on the method of mathematical induction and the use of summation signs, as these topics arise in the text. I have also found that it is wise to constantly remind the beginning student of the substitution principle, for example, that from $\mathrm{Var}(X) \geq 0$ for all X it follows that $\mathrm{Var}(2X - 3Y) \geq 0$. Much of the difficulty beginners have with mathematics stems from a lack of understanding of this principle, and it is well worth emphasis.

In all other respects, I have made every effort to have this book self-contained, clear, and readable. Throughout, stress is laid on the explanation of fundamental concepts and patterns of mathematical reasoning, as well as on techniques of problem-solving. Problems at the end of each section are designed to supplement the many worked-out illustrative examples in the text and to enable the reader to check his understanding of new definitions, theorems, and methods. From time to time, problems are included to challenge the better student—the sample variance, maximum likelihood estimation, the hypergeometric distribution, regression functions, and OC-curves for sampling inspection are introduced in problems that are written so as to guide the student toward an understanding of these important topics. Answers (often complete solutions) to half of the 360 problems are collected in a 21-page section at the end of the book. To facilitate computations, tables of ordinary logarithms, logarithms of factorials, and cumulative binomial probabilities are included in the text. A list of books suitable for supplementary reading appears at the end of each chapter. I trust that these features will serve to make the hard job of learning a little less hard.

Comments from readers are always welcome.

SAMUEL GOLDBERG

Cambridge, Mass.

ACKNOWLEDGMENTS

I TAKE this opportunity to gratefully acknowledge my debt to Professor William Feller who, as my teacher, first showed me the beauty and importance of the mathematical theory of probability.

Part of the material on sets and probability in Chapters 1 and 2 was prepared in preliminary form and tested in the classroom under a grant by the Carnegie Corporation of New York to Oberlin College for experimentation in freshman mathematics.

Assistance of various kinds was rendered at Oberlin College during the summers of 1958 and 1959, when I offered a probability course in National Science Foundation Institutes for secondary school teachers of mathematics, by Bruce T. Marcus, David Webster, and especially by Edward T. Wong.

The manuscript was completely rewritten while I held a visiting appointment at the Harvard University Graduate School of Business Administration to teach at the Institute of Basic Mathematics for Application to Business. The Institute, which was sponsored by the Ford Foundation, arranged for the final typing of the manuscript. W. Allen Spivey read part of the manuscript and offered helpful comments. Howard Raiffa made numerous valuable suggestions and the book is much the better for his counsel. Robert Schlaifer kindly gave permission to use the material in the final section of Chapter 5. William A. Ericson read the manuscript and prepared the solutions to problems.

I am grateful to all these friends for their help and to each goes my sincere thanks.

CONTENTS

Chapter 1

SETS

1. Examples of sets; basic notation

The concept of a set, whose fundamental role in mathematics was first pointed out in the work of the mathematician Georg Cantor (1845–1918), has significantly affected the structure and language of modern mathematics. In particular, the mathematical theory of probability is now most effectively formulated by using the terminology and notation of sets. For this reason, we devote Chapter 1 to the elementary mathematics of sets. Additional topics in set theory are included throughout the text, as the need for this material becomes apparent.

The notion of a set is sufficiently deep in the foundation of mathematics to defy being defined (at the level of this book) in terms of still more basic concepts. Hence, we can only aim here, by taking advantage of the reader's knowledge of the English language and his experience with the real and conceptual world, to make clear the denotation of the word "set."

A set is merely an aggregate or collection of objects of any sort: people, numbers, books, outcomes of experiments, geometrical figures, etc. Thus, we can speak of the set of all integers, or the set of all oceans, or the set of all possible sums when two dice are rolled and the number of dots on the uppermost faces are added, or the set consisting of the cities of Cambridge and Oberlin and all their residents, or the set of all straight lines (in a given plane) which pass through a given point.

The collection of objects must be *well-defined*, by which we mean that, for any object whatsoever, the question "Does this object belong to the collection?" has an unequivocal "yes" or "no" answer. It is not necessary that we personally have the knowledge required to decide which answer is correct. We must know only that, of the answers "yes" and "no," exactly one is correct.

Let us also agree that no object in a set is counted twice; i.e., the objects are *distinct*. It follows that, when listing the objects in a set, we do not repeat an object after it is once recorded. For example, according to this convention, the set of letters in the word "banana" is a set containing not six letters, but rather the three distinct letters b, a, and n.

The following definition summarizes our discussion to this point and introduces some additional terminology and notation.

Definition 1.1. A *set* is a well-defined collection of distinct objects. The individual objects that collectively make up a given set are called its *elements*, and each element *belongs to* or is a *member of* or is *contained in* the set. If a is an object and A a set, then we write $a \in A$ as an abbreviation for "a is an element of A" and $a \notin A$ for "a is not an element of A." If a set has a finite number of elements, then it is called a *finite* set; otherwise it is called an *infinite* set.

We are relying on the reader's knowledge of the positive integers $1, 2, 3, \cdots$, the so-called counting or natural numbers. This is an infinite set of numbers. To say that a set is finite means that one can enumerate the elements of the set in some order, then count these elements one by one until a *last element is reached.* Let us note that it is possible for a set, like the set of grains of sand on the Coney Island beach, to have a fantastically large number of elements and nevertheless be a finite set.

A set is ordinarily specified either by (i) listing all its elements and enclosing them in braces (the so-called *roster method* of defining the set), or by (ii) enclosing in braces a *defining property* and agreeing that those objects that have the property, and only those objects, are members of the set. We discuss these important ideas further and introduce additional notation in the following examples.

Example 1.1. The set whose elements are the integers 0, 5, and 12 is a finite set with three elements. If we denote this set by A, then it is conveniently written using the roster method: $A = \{0, 5, 12\}$. The statements "$5 \in A$" and "$6 \notin A$" are both true.

Example 1.2. If we write $V = \{a, e, i, o, u\}$, then we have defined the set V of vowels in the English alphabet by listing its five elements. To specify V by a defining property we write

$$V = \{x \mid x \text{ is a vowel in the English alphabet}\},$$

which is read "V is the set of those elements x such that x is a vowel in the English alphabet." Braces are always used when specifying a set; the vertical bar | is read "such that" or "for which." The symbol x is of course merely a place-holder; any other symbol will do just as well. For example, we can also write

$$V = \{* \mid * \text{ is a vowel in the English alphabet}\}.$$

A slight modification of this notation is often used. Let us first introduce the set A to stand for the set of all letters of the English alphabet. Then we write

$$V = \{* \in A \mid * \text{ is a vowel}\},$$

which is read "V is the set of those elements $*$ of A such that $*$ is a vowel."

Example 1.3. The set $B = \{-2, 2\}$ is the same set as $\{x \in R \mid x^2 = 4\}$, where R is the set of all real numbers. The set $\{x \in R \mid x^2 = -1\}$ has no elements, since the square of any real number is nonnegative. But if C is the set of all complex numbers, then $\{x \in C \mid x^2 = -1\}$ contains the elements $i = \sqrt{-1}$ and $-i$.

Example 1.4. A prime number is a positive integer greater than 1 but divisible only by 1 and itself. A proof of the fact that the set $\{p \mid p \text{ is a prime number}\}$ is an infinite set was given by Euclid (?330–275 B.C.) in the ninth book of his *Elements*. Strictly speaking, the roster method is unavailable for infinite sets, since it is not possible to list all the members and have explicitly before one a totality of elements making up an infinite set. The notation

$$\{2, 3, 5, 7, 11, 13, 17, 19, \cdots\},$$

in which some of the elements of the set are listed followed by three dots which take the place of *et cetera* and stand for obviously understood omissions of one or more elements, is an often used but logically unsatisfactory way out of this difficulty. (See Problem 1.3.) To specify an infinite set correctly, one must (as we did when we introduced the set of prime numbers) cite a defining property of the set.

Example 1.5. If a rectangular coordinate system (with x-axis and y-axis) is introduced in a plane, then each point of the plane has an x-coordinate and a y-coordinate, and can be represented, as in Figure 1(a), by an ordered pair of real numbers. In analytic geometry, one

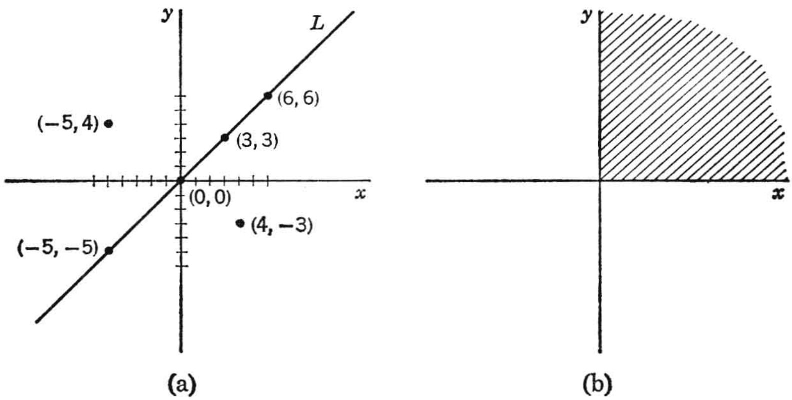

Figure 1

is interested in sets of points whose coordinates meet certain requirements. For example, the set $\{(x, y) \mid y = x\}$ is the set of all points (in a plane) with equal x- and y-coordinates. This infinite set of points makes up the straight line L, a portion of which is sketched in Figure 1(a), passing through the origin O and bisecting the first and third quadrants. We say that the line L is the *graph* of the set $\{(x, y) \mid y = x\}$. Similarly, the entire x-axis is the graph of the set $\{(x, y) \mid y = 0\}$, and the *positive* x-axis is the graph of the set $\{(x, y) \mid x > 0 \text{ and } y = 0\}$. The set $\{(x, y) \mid x > 0 \text{ and } y > 0\}$ is the set of points whose x- and y-coordinates are both positive. Thus, the graph of this set is the entire first quadrant (axes excluded), as indicated in Figure 1(b).

We see that a relation (in the form of equalities or inequalities between x and y) can be considered a *set-selector*, and the graph pictures the set of those points (from among all in the plane) selected by the requirement that their coordinates satisfy the given relation.

Although it may seem strange at first, it turns out to be convenient to talk about sets that have no members.

Definition 1.2. A set with no members is called an *empty* or *null* set.

The set $\{x \in R \mid x^2 = -1\}$ in Example 1.3 is an empty set. Another example is obtained by considering the set of all paths by which the line drawing of a house in Figure 2 can be traced without lifting one's pencil or retracing any line segment. Whether this set is empty or not is of some interest, since to assert that it is empty is to say that the figure cannot be traced under the prescribed conditions. (Let the reader convince himself that this set is indeed empty.) As our work develops, we shall see many other less frivolous reasons for introducing the notion of an empty set.

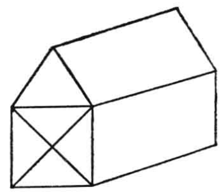

Figure 2

We conclude this ground-breaking section with one more definition.

Definition 1.3. Two sets A and B are said to be *equal* and we write $A = B$ if and only if they have exactly the same elements. If one of the sets has an element not in the other, they are unequal and we write $A \neq B$.

Thus $A = B$ means that every element of A is also an element of B and every element of B is an element of A. Equal sets are identical sets, and this identity is symbolized by the equality sign.

This definition has some interesting consequences. First, it is clear that the order in which we list the elements of a set is immaterial. For example, the set $\{a, b, c\}$ is equal to the set $\{c, a, b\}$, since they do indeed have exactly the same three elements.

Also, when sets are specified by defining properties, they can be equal even though the defining properties themselves are outwardly different. Thus, the set of all even prime numbers and the set of real numbers x such that $x + 3 = 5$ have different defining properties, yet they are equal sets, for each contains the number 2 as its only element.

Up to now, we have been careful to speak of a set having no members as *an* empty set. But it is clear from Definition 1.3 that any two empty sets are equal. For to be unequal it is necessary for one of the sets to contain an element not in the other, and this is impossible since neither set contains any elements. Therefore we are justified in referring to *the* empty set or *the* null set.* We denote the null set by the special symbol $\emptyset$.

PROBLEMS

1.1. We list eight sets. For each set, state whether it is finite or infinite. If finite, count the number of elements in the set. Where feasible, write the set using the roster method.

(a) The set of footnotes in Section 1.
(b) The set of letters in the word "probability."
(c) The set of odd positive integers.
(d) The set of prime numbers less than one million.
(e) The set of paths by which the following figure can be traced without lifting one's pencil or retracing any line segment:

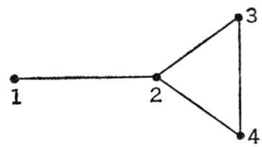

Figure 3

(f) The set of those points (in a given plane) that are exactly five units from the origin O.
(g) The set of real numbers satisfying the equation $x^2 - 3x + 2 = 0$.
(h) The set of possible outcomes of the experiment in which one card is selected from a standard deck of 52 cards.

1.2. The following paragraph was written by a student impressed with the technical vocabulary of set theory. Rewrite in more usual English prose.

Let C be the set of Mr. and Mrs. Smith's children. C was equal to $\emptyset$ until March 1, 1958. C contained exactly one element from that date until March 15, 1959 when it increased its membership by two!

* The following true story concerns the attempt of a well-known professor of mathematics to teach his five-year-old son the subtle distinction between "a" (or "an") and "the." One day the son answered the telephone, listened a moment and then said, "I'm sorry, but you have the wrong number." (Isn't this what most of us say when someone dials incorrectly?) The father, having overheard, immediately called the boy to him and gently instructed, "What you said would be correct if there were exactly one wrong number. But since there are many possible wrong numbers, it would be more accurate to say, 'I'm sorry, but you have *a* wrong number.' "