INTRODUCTORY PROBABILITY AND STATISTICAL APPLICATIONS

PAUL L. MEYER

Introductory Probability and Statistical Applications

PAUL L. MEYER

Department of Mathematics, Washington State University

ADDISON-WESLEY, READING, MASSACHUSETTS

This book is in the **ADDISON-WESLEY SERIES IN STATISTICS**

Printed in the United States of America

Library of Congress Catalog Card No. 65-14525

ADDISON-WESLEY PUBLISHING COMPANY, INC.
READING, MASSACHUSETTS · Palo Alto · London
NEW YORK · DALLAS · ATLANTA · BARRINGTON, ILLINOIS

Lest men suspect your tale untrue
Keep probability in view.

JOHN GAY

Preface

This text is intended for a one-semester or two-quarter course in introductory probability theory and some of its applications. The prerequisite is one year of differential and integral calculus. No previous knowledge of probability or statistics is presumed. At Washington State University, the course for which this text was developed has been taught for a number of years, chiefly to students majoring in engineering or in the natural sciences. Most of these students can devote only one semester to the study of this subject area. However, since these students are familiar with the calculus, they may begin the study of this subject beyond the strictly elementary level.

Many mathematical topics may be introduced at varying stages of difficulty, and this is certainly true of probability. In this text, an attempt is made to take advantage of the reader's mathematical background without exceeding it. Precise mathematical language is used but care is taken not to become excessively immersed in unnecessary mathematical details. This is most certainly not a "cook book." Although a number of concepts are introduced and discussed in an informal manner, definitions and theorems are carefully stated. If a detailed proof of a theorem is not feasible or desirable at least an outline of the important ideas is provided. A distinctive feature of this text is the "Notes" following most of the theorems and definitions. In these Notes the particular result or concept being presented is discussed from an intuitive point of view.

Because of the self-imposed restriction of writing a relatively brief text on a very extensive subject area, a number of choices had to be made relating to the inclusion or exclusion of certain topics. There seems to be no obvious way of resolving this problem. I certainly do not claim that for some of the topics which are excluded a place might not have been found. Nor do I claim that none of the material could have been omitted. However, for the most part, the emphasis has been on fundamental notions, presented in considerable detail. Only Chapter 11 on reliability could be considered as a "luxury item." But even here I feel that the notions associated with reliability problems are of basic interest to many persons. In addition, reliability concepts represent an excellent vehicle for illustrating many of the ideas introduced earlier in the text.

Even though the coverage is limited by the available time, a fairly wide selection of topics has been achieved. In glancing through the Table of Contents it is evident that about three-quarters of the text deals with probabilistic topics while the last quarter is devoted to a discussion of statistical inference. Although there is nothing magic about this particular division of emphasis between probability and statistics, I do feel that a sound knowledge of the basic principles of probability is imperative for a proper understanding of statistical methods. Ideally, a course in probability should be followed by one in statistical theory and methodology. However, as I indicated earlier, most of the students who take this course do not have time for a two-semester exposure to this subject area and hence I felt compelled to discuss at least some of the more important aspects of the general area of statistical inference.

The potential success of a particular presentation of subject matter should not be judged only in terms of specific ideas learned and specific techniques acquired. The final judgment must also take into account how well the student is prepared to continue his study of the subject either through self-study or through additional formal course work. If this criterion is thought to be important, then it becomes clear that basic concepts and fundamental techniques should be emphasized while highly specialized methods and topics should be relegated to a secondary role. This, too, became an important factor in deciding which topics to include.

The importance of probability theory is difficult to overstate. The appropriate mathematical model for the study of a large number of observational phenomena is a probabilistic one rather than a deterministic one. In addition, the entire subject of statistical inference is based on probabilistic considerations. Statistical techniques are among the most important tools of scientists and engineers. In order to use these techniques intelligently a thorough understanding of probabilistic concepts is required.

It is hoped that in addition to the many specific methods and concepts with which the reader becomes familiar, he also develops a certain point of view: to think probabilistically, replacing questions such as "How long will this component function?" by "How probable is it that this component will function more than 100 hours?" In many situations the second question may be not only the more pertinent one but in fact the only meaningful one to ask.

Traditionally, many of the important concepts of probability are illustrated with the aid of various "games of chance": tossing coins or dice, drawing cards from a deck, spinning a roulette wheel, etc. Although I have not completely avoided referring to such games since they do serve well to illustrate basic notions, an attempt has been made to bring the student into contact with more pertinent illustrations of the applications of probability: the emission of α-particles from a radioactive source, lot sampling, the life length of electronic devices, and the associated problems of component and system reliability, etc.

I am reluctant to mention a most obvious feature of any text in mathematics: the problems. And yet, it might be worthwhile to point out that the working of the problems must be considered as an integral part of the course. Only by becoming personally involved in the setting up and solving of the exercises can the student really develop an understanding and appreciation of the ideas and a familiarity with the pertinent techniques. Hence over 330 problems have been included in the text, over half of which are provided with answers at the end of the book. In addition to the problems for the reader, there are many worked-out examples scattered throughout the text.

This book has been written in a fairly consecutive manner: the understanding of most chapters requires familiarity with the previous ones. However, it is possible to treat Chapters 10 and 11 somewhat lightly, particularly if one is interested in devoting more time to the statistical applications which are discussed in Chapters 13 through 15.

As must be true of anyone writing a text, the debts I owe are to many: To my colleagues for many stimulating and helpful conversations, to my own teachers for the knowledge of and interest in this subject, to the reviewers of early versions of the manuscript for many helpful suggestions and criticisms, to Addison-Wesley Publishing Company for its great help and cooperation from the early stages of this project to the very end, to Miss Carol Sloan for being a most efficient and alert typist, to D. Van Nostrand, Inc., The Free Press, Inc., and Macmillan Publishing Company for their permission to reprint Tables 3, 6, and 1, respectively; to McGraw-Hill Book Co., Inc., Oxford University Press, Inc., Pergamon Press, Ltd., and Prentice-Hall, Inc., for their permission to quote certain examples in the text, and finally to my wife not only for bearing up under the strain but also for "leaving me" and taking our two children with her to visit grandparents for two crucial summer months during which I was able to convert our home into a rather messy but quiet workshop from which emerged, miraculously, the final, final version of this book.

Pullman, Washington P. L. M.
April, 1965

Contents

CHAPTER 1 | Introduction to Probability

1.1 Mathematical Models

In this chapter we shall discuss the type of phenomenon with which we shall be concerned throughout this book. In addition, we shall formulate a mathematical model which will serve us to investigate, quite precisely, this phenomenon.

At the outset it is very important to distinguish between the observable phenomenon itself and the mathematical model for this phenomenon. We have, of course, no influence over what we observe. However, in choosing a model we can use our critical judgment. This has been particularly well expressed by Professor J. Neyman, who wrote:*

> "Whenever we use mathematics in order to study some observational phenomena we must essentially begin by building a mathematical model (deterministic or probabilistic) for these phenomena. Of necessity, the model must simplify matters and certain details must be ignored. The success of the model depends on whether or not the details ignored are really unimportant in the development of the phenomena studied. The solution of the mathematical problem may be correct and yet be in considerable disagreement with the observed data simply because the underlying assumptions made are not warranted. It is usually quite difficult to state with certainty, whether or not a given mathematical model is adequate *before* some observational data are obtained. In order to check the validity of a model, we must *deduce* a number of consequences of our model and then compare these *predicted* results with observations."

We shall keep the above ideas in mind while we consider some observational phenomena and models appropriate for their description. Let us first consider what might suitably be called a *deterministic model*. By this we shall mean a model which stipulates that the conditions under which an experiment is performed *determine* the outcome of the experiment. For example, if we insert a battery into a simple circuit, the mathematical model which would presumably describe the observable flow of current would be $I = E/R$, that is, Ohm's law. The model predicts the value of I as soon as E and R are given. Saying it dif-

---

* *University of California Publications in Statistics*, Vol. I, University of California Press, 1954.

ferently, if the above experiment were repeated a number of times, each time using the same circuit (that is, keeping E and R fixed), we would presumably expect to observe the same value for I. Any deviations that might occur would be so small that for most purposes the above description (that is, model) would suffice. The point is that the particular battery, wire, and ammeter used to generate and to observe the current, and our ability to use the measuring instrument, determine the outcome on each repetition. (There are certain factors which may well be different from repetition to repetition that will, however, not affect the outcome in a noticeable way. For instance, the temperature and humidity in the laboratory, or the height of the person reading the ammeter can reasonably be assumed to have no influence on the outcome.)

There are many examples of "experiments" in nature for which deterministic models are appropriate. For example, the gravitational laws describe quite precisely what happens to a falling body under certain conditions. Kepler's laws give us the behavior of the planets. In each situation, the model stipulates that the conditions under which certain phenomena take place determine the value of certain observable variables: the *magnitude* of the velocity, the *area* swept out during a certain time period, etc. These numbers appear in many of the formulas with which we are familiar. For example, we know that under certain conditions the distance traveled (vertically, above the ground) by an object is given by $s = -16t^2 + v_0t$, where v_0 is the initial velocity and t is the time traveled. The point on which we wish to focus our attention is not the particular form of the above equation (that is, quadratic) but rather on the fact that there is a definite relationship between t and s, determining uniquely the quantity on the left-hand side of the equation if those on the right-hand side are given.

For a large number of situations the deterministic mathematical model described above suffices. However, there are also many phenomena which require a different mathematical model for their investigation. These are what we shall call ***nondeterministic*** or ***probabilistic*** models. (Another quite commonly used term is ***stochastic*** model.) Later in this chapter we shall consider quite precisely how such probabilistic models may be described. For the moment let us consider a few examples.

Suppose that we have a piece of radioactive material which is emitting α-particles. With the aid of a counting device we may be able to record the number of such particles emitted during a specified time interval. It is clear that we cannot predict precisely the number of particles emitted, even if we knew the exact shape, dimension, chemical composition, and mass of the object under consideration. Thus there seems to be no reasonable deterministic model yielding the number of particles emitted, say n, as a function of various pertinent characteristics of the source material. We must consider, instead, a probabilistic model.

For another illustration consider the following meteorological situation. We wish to determine how much precipitation will fall as a result of a particular storm system passing through a specified locality. Instruments are available with which to record the precipitation that occurs. Meteorological observations may give us considerable information concerning the approaching storm system: barometric pressure at various points, changes in pressure, wind velocity, origin and direction of the storm, and various pertinent high-altitude readings But this information, valuable as it may be for predicting the general nature of the precipitation (light, medium, or heavy, say), simply does not make it possible to state very accurately *how much* precipitation will fall. Again we are dealing with a phenomenon which does not lend itself to a deterministic approach. A probabilistic model describes the situation more accurately.

In principle we might be able to state how much rain fell, if the theory had been worked out (which it has not). Hence we use a probabilistic model. In the example dealing with radioactive disintegration, we must use a probabilistic model *even in principle*.

At the risk of getting ahead of ourselves by discussing a concept which will be defined subsequently, let us simply state that in a deterministic model it is supposed that the actual outcome (whether numerical or otherwise) is determined from the conditions under which the experiment or procedure is carried out. In a nondeterministic model, however, the conditions of experimentation determine only the probabilistic behavior (more specifically, the probabilistic law) of the observable outcome.

Saying it differently, in a deterministic model we use "physical considerations" to predict the outcome, while in a probabilistic model we use the same kind of considerations to specify a probability distribution.

1.2 Introduction to Sets

In order to discuss the basic concepts of the probabilistic model which we wish to develop, it will be very convenient to have available some ideas and concepts of the mathematical theory of sets. This subject is a very extensive one, and much has been written about it. However, we shall need only a few basic notions.

A *set* is a collection of objects. Sets are usually designated by capital letters A, B, etc. In describing which objects are contained in the set A, three methods are available.

(a) We may list the members of A. For example, $A = \{1, 2, 3, 4\}$ describes the set consisting of the positive integers 1, 2, 3, and 4.

(b) We may describe the set A in words. For example, we might say that A consists of all real numbers between 0 and 1, inclusive.

(c) To describe the above set we can simply write $A = \{x \mid 0 \leq x \leq 1\}$; that is, A is the set of all x's, where x is a real number between 0 and 1, inclusive.

The individual objects making up the collection of the set A are called *members* or *elements* of A. When "a" is a member of A we write $a \in A$ and when "a" is not a member of A we write $a \notin A$.

There are two special sets which are often of interest. In most problems we are concerned with the study of a definite set of objects, and no others. For example, we may be concerned with all the real numbers, all items coming off a production line during a 24-hour period, etc. We define the *universal set* as the set of all objects under consideration. This set is usually designated by U.

Another set which must be singled out may rise as follows. Suppose that the set A is described as the set of all *real* numbers x satisfying the equation $x^2 + 1 = 0$. Of course, we know that there are no such numbers. That is, the set A contains no members at all! This situation occurs sufficiently often to warrant the introduction of a special name for such a set. Hence we define the *empty* or *null* set to be the set containing no members. We usually designate this set by $\emptyset$.

It may happen that when two sets A and B are considered, being a member of A implies being a member of B. In that case we say that A is a *subset* of B and we write $A \subset B$. A similar interpretation is given to $B \subset A$. And we say that two sets are the same, $A = B$, if and only if $A \subset B$ and $B \subset A$. That is, two sets are *equal* if and only if they contain the same members.

The following two properties of the empty set and the universal set are immediate.

(a) For every set A, we have $\emptyset \subset A$.

(b) Once the universal set has been agreed upon, then for every set A considered in the context of U, we have $A \subset U$.

EXAMPLE 1.1. Suppose that U=all real numbers, $A = \{x \mid x^2 + 2x - 3 = 0\}$, $B = \{x \mid (x - 2)(x^2 + 2x - 3) = 0\}$, and $C = \{x \mid x = -3, 1, 2\}$. Then $A \subset B$, and $B = C$.

Next we consider the important idea of *combining* given sets in order to form a new set. Two basic operations are considered. These operations parallel, in certain respects, the operation of addition and multiplication of numbers. Let A and B be two sets. We define C as the *union* of A and B (sometimes called the sum of A and B) as follows:

$$C = \{x \mid x \in A \text{ or } x \in B \text{ (or both)}\}.$$

We write this as $C = A \cup B$. Thus C consists of all elements which are in A, or in B, or in both.

We define D as the *intersection* of A and B (sometimes called the product of A and B) as follows:

$$D = \{x \mid x \in A \text{ and } x \in B\}.$$

We write this as $D = A \cap B$. Thus D consists of all elements which are in A *and* in B.

Finally we introduce the idea of the *complement* of a set A as follows: The set, denoted by $\overline{A}$, consisting of all elements *not* in A (but in the universal set U) is called the complement of A. That is, $\overline{A} = \{x \mid x \notin A\}$.

A graphic device known as a *Venn diagram* can be used to considerable advantage when we are combining sets as indicated above. In each diagram in Fig. 1.1, the *shaded* region represents the set under consideration.

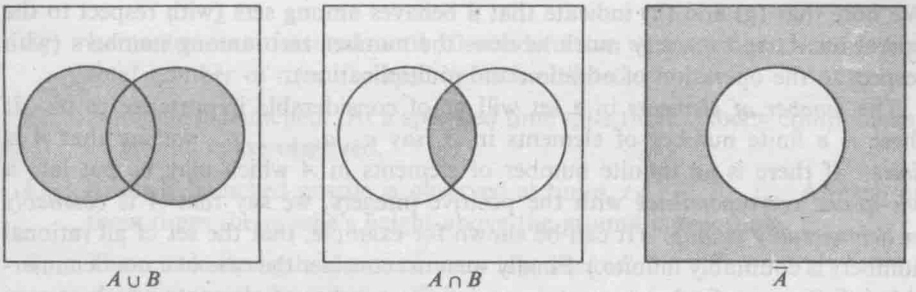

FIGURE 1.1

EXAMPLE 1.2. Suppose that $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$; $A = \{1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$. Then we find that $\overline{A} = \{5, 6, 7, 8, 9, 10\}$, $A \cup B = \{1, 2, 3, 4, 5, 6\}$, and $A \cap B = \{3, 4\}$. Note that in describing a set (such as $A \cup B$) we list an element exactly once.

The above operations of union and intersection defined for just two sets may be extended in an obvious way to any finite number of sets. Thus we define $A \cup B \cup C$ as $A \cup (B \cup C)$ or $(A \cup B) \cup C$, which are the same, as can easily be checked. Similarly, we define $A \cap B \cap C$ as $A \cap (B \cap C)$ or $(A \cap B) \cap C$, which again can be checked to be the same. And it is clear that we may continue these constructions of new sets for *any finite* number of given sets.

We asserted that certain sets were the same, for example $A \cap (B \cap C)$ and $(A \cap B) \cap C$. It turns out that there are a number of such *equivalent* sets, some of which are listed below. If we recall that two sets are the same whenever they contain the same members, it is easy to show that the assertions stated are true. The reader should convince himself of these with the aid of Venn diagrams.

$$\begin{aligned} &\text{(a) } A \cup B = B \cup A, && \text{(b) } A \cap B = B \cap A, \\ &\text{(c) } A \cup (B \cup C) = (A \cup B) \cup C, && \text{(d) } A \cap (B \cap C) = (A \cap B) \cap C. \end{aligned} \tag{1.1}$$

We refer to (a) and (b) as the *commutative* laws, and (c) and (d) as the *associative* laws.

There are a number of other such *set identities* involving union, intersection, and complementation. The most important of these are listed below. In each case, their validity may be checked with the aid of a Venn diagram.

$$\begin{aligned} &\text{(e) } A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \\ &\text{(f) } A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \\ &\text{(g) } A \cap \emptyset = \emptyset, \\ &\text{(h) } A \cup \emptyset = A, && \text{(i) } \overline{(A \cup B)} = \overline{A} \cap \overline{B}, \\ &\text{(j) } \overline{(A \cap B)} = \overline{A} \cup \overline{B}, && \text{(k) } \overline{\overline{A}} = A. \end{aligned} \tag{1.2}$$

We note that (g) and (h) indicate that $\emptyset$ behaves among sets (with respect to the operations $\cup$ and $\cap$) very much as does the number zero among numbers (with respect to the operation of addition and multiplication).

The *number of elements* in a set will be of considerable importance to us. If there is a finite number of elements in A, say $a_1, a_2, \ldots, a_n$, we say that A is *finite.* If there is an infinite number of elements in A which may be put into a *one-to-one correspondence* with the positive integers, we say that A is *countably* or *denumerably infinite.* (It can be shown for example, that the set of all rational numbers is countably infinite.) Finally we must consider the case of a nondenumerable infinite set. Such sets contain an infinite number of elements which cannot be enumerated. It can be shown, for instance, that for any two real numbers $b > a$, set $A = \{x \mid a \leq x \leq b\}$ has a nondenumerable number of elements. Since we may associate with each real number a point on the real number line, the above says that any (nondegenerate) interval contains more than a countable number of points.

The concepts introduced above, although representing only a brief glimpse into the theory of sets, are sufficient for our purpose: to describe with considerable rigor and precision, the basic ideas of probability theory.

1.3 Examples of Nondeterministic Experiments

We are now ready to discuss what we mean by a "random" or "nondeterministic" experiment. (More precisely, we shall give examples of phenomena for which nondeterministic models are appropriate. This is a distinction which the reader should keep in mind. Thus we shall repeatedly refer to nondeterministic or random experiments when in fact we are talking about a nondeterministic *model* for an experiment.) We shall not attempt to give a precise dictionary definition of this concept. Instead, we shall cite a large number of examples illustrating what we have in mind. In describing an experiment, we must specify not only what operation or procedure is to be carried out, but we must also specify what is to be observed. Note, for instance, the difference between E_2 and E_3 below.

E_1: Toss a die and observe the number that shows on top.

E_2: Toss a coin four times and observe the total number of heads obtained.

E_3: Toss a coin four times and observe the sequence of heads and tails obtained.

E_4: Manufacture items on a production line and count the number of defective items produced during a 24-hour period.

E_5: An airplane wing is assembled with a large number of rivets. The number of defective rivets is counted.

E_6: A light bulb is manufactured. It is then tested for its life length by inserting it into a socket and the time elapsed (in hours) until it burns out is recorded.

E_7: A lot of 10 items contains 3 defectives. One item is chosen after another (without replacing the chosen item) until the last defective item is obtained. The total number of items removed from the lot is counted.

E_8: Items are manufactured until 10 nondefective items are produced. The total number of manufactured items is counted.

E_9: A missile is launched. At a specified time t, its three velocity components, v_x, v_y, and v_z are observed.

E_{10}: A newly launched missile is observed at times, $t_1, t_2, \ldots, t_n$. At each of these times the missile's height above the ground is recorded.

E_{11}: The tensile strength of a steel beam is measured.

E_{12}: From an urn containing only black balls, a ball is chosen and its color noted.

E_{13}: A thermograph records temperature, continuously, over a 24-hour period. At a specified locality and on a specified date, such a thermograph is "read."

E_{14}: In the situation described in E_{13}, x and y, the minimum and maximum temperatures of the 24-hour period in question are recorded.

What do the above experiments have in common? The following features are pertinent for our characterization of a *random experiment*.

(a) Each experiment is capable of being repeated indefinitely under essentially unchanged conditions.

(b) Although we are in general not able to state what a *particular* outcome will be, we are able to describe the set of *all possible* outcomes of the experiment.

(c) As the experiment is performed repeatedly, the individual outcomes seem to occur in a haphazard manner. However, as the experiment is repeated a *large* number of times, a definite pattern or regularity appears. It is this regularity which makes it possible to construct a precise mathematical model with which to analyze the experiment. We will have much more to say about the nature and importance of this regularity later. For the moment, the reader need only think of the repeated tossings of a fair coin. Although heads and tails will appear, successively, in an almost arbitrary fashion, it is a well-known empirical fact that after a large number of tosses the proportion of heads and tails will be approximately equal.

It should be noted that all the experiments described above satisfy these general characteristics. (Of course, the last mentioned characteristic can only be verified by experimentation; we will leave it to the reader's intuition to believe that if the experiment were repeated a large number of times, the regularity referred to would be evident. For example, if a large number of light bulbs from the same manufacturer were tested, presumably the number of bulbs burning more than 100 hours, say, could be predicted with considerable accuracy.) Note that experiment E_{12} has the peculiar feature that only one outcome is possible. In general such experiments will not be of interest, for the very fact that we do not know which particular outcome will occur when an experiment is performed is what makes it of interest to us.