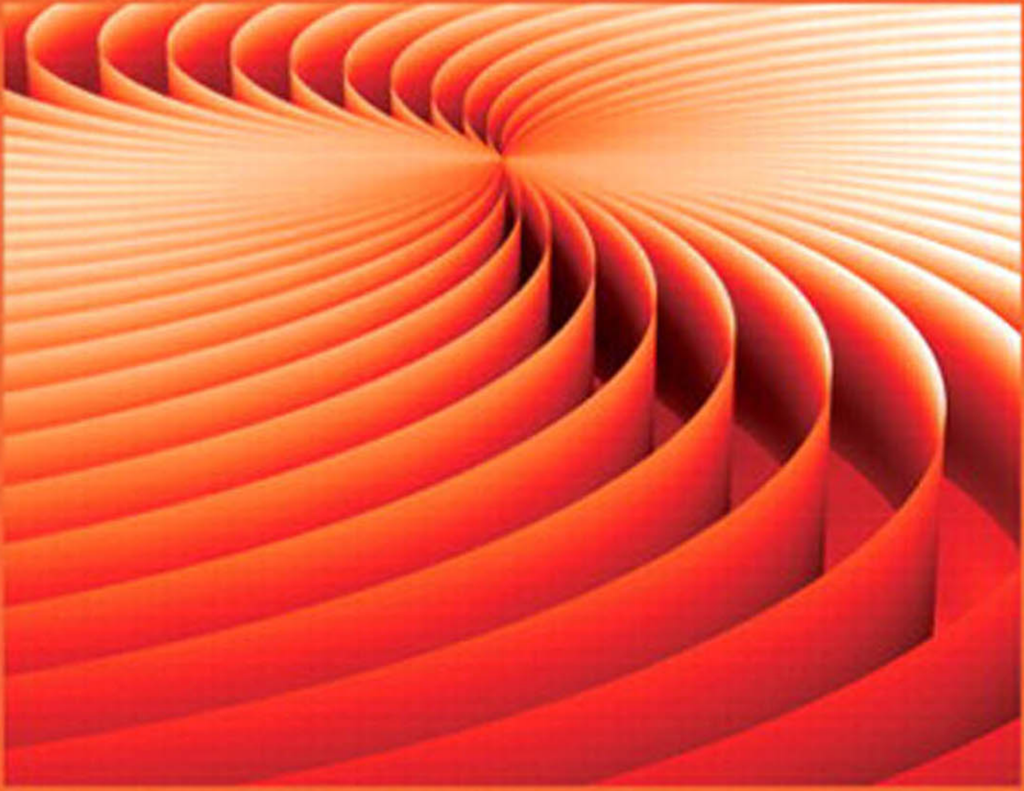
Probability and Statistical Inference

NINTH EDITION



Hogg | Tanis | Zimmerman

Probability and Statistical Inference

Ninth Edition

Robert V. Hogg

Elliot A. Tanis

Dale L. Zimmerman

PEARSON

Boston Columbus Indianapolis New York San Francisco
Upper Saddle River Amsterdam Cape Town Dubai
London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul
Singapore Taipei Tokyo

Editor in Chief: Deirdre Lynch
Acquisitions Editor: Christopher Cummings
Sponsoring Editor: Christina Lepre
Assistant Editor: Sonia Ashraf
Marketing Manager: Erin Lane
Marketing Assistant: Kathleen DeChavez
Senior Managing Editor: Karen Wernholm
Senior Production Editor: Beth Houston
Procurement Manager: Vincent Scelta
Procurement Specialist: Carol Melville
Associate Director of Design, USHE EMSS/HSC/EDU: Andrea Nix
Art Director: Heather Scott
Interior Designer: Tamara Newnam
Cover Designer: Heather Scott
Cover Image: Agsandrew/Shutterstock
Full-Service Project Management: Integra Software Services
Composition: Integra Software Services

Library of Congress Cataloging-in-Publication Data
Hogg, Robert V.
Probability and Statistical Inference/
Robert V. Hogg, Elliot A. Tanis, Dale Zimmerman. – 9th ed.
p. cm.
ISBN 978-0-321-92327-1
1. Mathematical statistics. I. Hogg, Robert V., II. Tanis, Elliot A. III. Title.
QA276.H59 2013
519.5–dc23
2011034906

10 9 8 7 6 5 4 3 2 1 EBM 17 16 15 14 13

PEARSON

www.pearsonhighered.com

ISBN-10: 0-321-92327-8
ISBN-13: 978-0-321-92327-1

Contents

Preface

In this Ninth Edition of *Probability and Statistical Inference*, Bob Hogg and Elliot Tanis are excited to add a third person to their writing team to contribute to the continued success of this text. Dale Zimmerman is the Robert V. Hogg Professor in the Department of Statistics and Actuarial Science at the University of Iowa. Dale has rewritten several parts of the text, making the terminology more consistent and contributing much to a substantial revision. The text is designed for a two-semester course, but it can be adapted for a one-semester course. A good calculus background is needed, but no previous study of probability or statistics is required.

CONTENT AND COURSE PLANNING

In this revision, the first five chapters on probability are much the same as in the eighth edition. They include the following topics: probability, conditional probability, independence, Bayes' theorem, discrete and continuous distributions, certain mathematical expectations, bivariate distributions along with marginal and conditional distributions, correlation, functions of random variables and their distributions, including the moment-generating function technique, and the central limit theorem. While this strong probability coverage of the course is important for all students, it has been particularly helpful to actuarial students who are studying for Exam P in the Society of Actuaries' series (or Exam 1 of the Casualty Actuarial Society).

The greatest change to this edition is in the statistical inference coverage, now Chapters 6–9. The first two of these chapters provide an excellent presentation of estimation. Chapter 6 covers point estimation, including descriptive and order statistics, maximum likelihood estimators and their distributions, sufficient statistics, and Bayesian estimation. Interval estimation is covered in Chapter 7, including the topics of confidence intervals for means and proportions, distribution-free confidence intervals for percentiles, confidence intervals for regression coefficients, and resampling methods (in particular, bootstrapping).

The last two chapters are about tests of statistical hypotheses. Chapter 8 considers terminology and standard tests on means and proportions, the Wilcoxon tests, the power of a test, best critical regions (Neyman/Pearson) and likelihood ratio tests. The topics in Chapter 9 are standard chi-square tests, analysis of variance including general factorial designs, and some procedures associated with regression, correlation, and statistical quality control.

The first semester of the course should contain most of the topics in Chapters 1–5. The second semester includes some topics omitted there and many of those in Chapters 6–9. A more basic course might omit some of the (optional) starred sections, but we believe that the order of topics will give the instructor the flexibility needed in his or her course. The usual nonparametric and Bayesian techniques are placed at appropriate places in the text rather than in separate chapters. We find that many persons like the applications associated with statistical quality control in the last section. Overall, one of the authors, Hogg, believes that the presentation (at a somewhat reduced mathematical level) is much like that given in the earlier editions of Hogg and Craig (see References).

The Prologue suggests many fields in which statistical methods can be used. In the Epilogue, the importance of understanding variation is stressed, particularly for its need in continuous quality improvement as described in the usual Six-Sigma programs. At the end of each chapter we give some interesting historical comments, which have proved to be very worthwhile in the past editions.

The answers given in this text for questions that involve the standard distributions were calculated using our probability tables which, of course, are rounded off for printing. If you use a statistical package, your answers may differ slightly from those given.

ANCILLARIES

Data sets from this textbook are available on Pearson Education's Math & Statistics Student Resources website: http://www.pearsonhighered.com/mathstatsresources.

An **Instructor's Solutions Manual** containing worked-out solutions to the even-numbered exercises in the text is available for download from Pearson Education's Instructor Resource Center at www.pearsonhighered.com/irc. Some of the numerical exercises were solved with *Maple*. For additional exercises that involve simulations, a separate manual, *Probability & Statistics: Explorations with MAPLE*, second edition, by Zaven Karian and Elliot Tanis, is also available for download from Pearson Education's Instructor Resource Center. Several exercises in that manual also make use of the power of *Maple* as a computer algebra system.

If you find any errors in this text, please send them to tanis@hope.edu so that they can be corrected in a future printing. These **errata** will also be posted on http://www.math.hope.edu/tanis/.

ACKNOWLEDGMENTS

We wish to thank our colleagues, students, and friends for many suggestions and for their generosity in supplying data for exercises and examples. In particular, we would like to thank the reviewers of the eighth edition who made suggestions for this edition. They are Steven T. Garren from James Madison University, Daniel C. Weiner from Boston University, and Kyle Siegrist from the University of Alabama in Huntsville. Mark Mills from Central College in Iowa also made some helpful comments. We also acknowledge the excellent suggestions from our copy editor, Kristen Cassereau Ng, and the fine work of our accuracy checkers, Kyle Siegrist and Steven Garren. We also thank the University of Iowa and Hope College for providing office space and encouragement. Finally, our families, through nine editions, have been most understanding during the preparation of all of this material. We would especially like to thank our wives, Ann, Elaine, and Bridget. We truly appreciate their patience and needed their love.

Robert V. Hogg

Elliot A. Tanis
tanis@hope.edu

Dale L. Zimmerman
dale-zimmerman@uiowa.edu

Prologue

The discipline of statistics deals with the collection and analysis of data. Advances in computing technology, particularly in relation to changes in science and business, have increased the need for more statistical scientists to examine the huge amount of data being collected. We know that data are not equivalent to information. Once data (hopefully of high quality) are collected, there is a strong need for statisticians to make sense of them. That is, data must be analyzed in order to provide information upon which decisions can be made. In light of this great demand, opportunities for the discipline of statistics have never been greater, and there is a special need for more bright young persons to go into statistical science.

If we think of fields in which data play a major part, the list is almost endless: accounting, actuarial science, atmospheric science, biological science, economics, educational measurement, environmental science, epidemiology, finance, genetics, manufacturing, marketing, medicine, pharmaceutical industries, psychology, sociology, sports, and on and on. Because statistics is useful in all of these areas, it really should be taught as an applied science. Nevertheless, to go very far in such an applied science, it is necessary to understand the importance of creating models for each situation under study. Now, no model is ever exactly right, but some are extremely useful as an approximation to the real situation. Most appropriate models in statistics require a certain mathematical background in probability. Accordingly, while alluding to applications in the examples and the exercises, this textbook is really about the mathematics needed for the appreciation of probabilistic models necessary for statistical inferences.

In a sense, statistical techniques are really the heart of the scientific method. Observations are made that suggest conjectures. These conjectures are tested, and data are collected and analyzed, providing information about the truth of the conjectures. Sometimes the conjectures are supported by the data, but often the conjectures need to be modified and more data must be collected to test the modifications, and so on. Clearly, in this iterative process, statistics plays a major role with its emphasis on the proper design and analysis of experiments and the resulting inferences upon which decisions can be made. Through statistics, information is provided that is relevant to taking certain actions, including improving manufactured products, providing better services, marketing new products or services, forecasting energy needs, classifying diseases better, and so on.

Statisticians recognize that there are often small errors in their inferences, and they attempt to quantify the probabilities of those mistakes and make them as small as possible. That these uncertainties even exist is due to the fact that there is variation in the data. Even though experiments are repeated under seemingly the same conditions, the results vary from trial to trial. We try to improve the quality of the data by making them as reliable as possible, but the data simply do not fall on given patterns. In light of this uncertainty, the statistician tries to determine the pattern in the best possible way, always explaining the error structures of the statistical estimates.

This is an important lesson to be learned: Variation is almost everywhere. It is the statistician's job to understand variation. Often, as in manufacturing, the desire is to reduce variation because the products will be more consistent. In other words, car doors will fit better in the manufacturing of automobiles if the variation is decreased by making each door closer to its target values.

Many statisticians in industry have stressed the need for "statistical thinking" in everyday operations. This need is based upon three points (two of which have been mentioned in the preceding paragraph): (1) Variation exists in all processes; (2) understanding and reducing undesirable variation is a key to success; and (3) all work occurs in a system of interconnected processes. W. Edwards Deming, an esteemed statistician and quality improvement "guru," stressed these three points, particularly the third one. He would carefully note that you could not maximize the total operation by maximizing the individual components unless they are independent of each other. However, in most instances, they are highly dependent, and persons in different departments must work together in creating the best products and services. If not, what one unit does to better itself could very well hurt others. He often cited an orchestra as an illustration of the need for the members to work together to create an outcome that is consistent and desirable.

Any student of statistics should understand the nature of variability and the necessity for creating probabilistic models of that variability. We cannot avoid making inferences and decisions in the face of this uncertainty; however, these inferences and decisions are greatly influenced by the probabilistic models selected. Some persons are better model builders than others and accordingly will make better inferences and decisions. The assumptions needed for each statistical model are carefully examined; it is hoped that thereby the reader will become a better model builder.

Finally, we must mention how modern statistical analyses have become dependent upon the computer. Statisticians and computer scientists really should work together in areas of exploratory data analysis and "data mining." Statistical software development is critical today, for the best of it is needed in complicated data analyses. In light of this growing relationship between these two fields, it is good advice for bright students to take substantial offerings in statistics and in computer science.

Students majoring in statistics, computer science, or a joint program are in great demand in the workplace and in graduate programs. Clearly, they can earn advanced degrees in statistics or computer science or both. But, more important, they are highly desirable candidates for graduate work in other areas: actuarial science, industrial engineering, finance, marketing, accounting, management science, psychology, economics, law, sociology, medicine, health sciences, etc. So many fields have been "mathematized" that their programs are begging for majors in statistics or computer science. Often, such students become "stars" in these other areas. We truly hope that we can interest students enough that they want to study more statistics. If they do, they will find that the opportunities for very successful careers are numerous.

Chapter 1

PROBABILITY

1.1 PROPERTIES OF PROBABILITY

It is usually difficult to explain to the general public what statisticians do. Many think of us as "math nerds" who seem to enjoy dealing with numbers. And there is some truth to that concept. But if we consider the bigger picture, many recognize that statisticians can be extremely helpful in many investigations.

Consider the following:

1. There is some problem or situation that needs to be considered; so statisticians are often asked to work with investigators or research scientists.
2. Suppose that some measure (or measures) is needed to help us understand the situation better. The measurement problem is often extremely difficult, and creating good measures is a valuable skill. As an illustration, in higher education, how do we measure good teaching? This is a question to which we have not found a satisfactory answer, although several measures, such as student evaluations, have been used in the past.
3. After the measuring instrument has been developed, we must collect data through observation, possibly the results of a survey or an experiment.
4. Using these data, statisticians summarize the results, often with descriptive statistics and graphical methods.
5. These summaries are then used to analyze the situation. Here it is possible that statisticians make what are called statistical inferences.
6. Finally, a report is presented, along with some recommendations that are based upon the data and the analysis of them. Frequently such a recommendation might be to perform the survey or experiment again, possibly changing some of the questions or factors involved. This is how statistics is used in what is referred to as the scientific method, because often the analysis of the data suggests other experiments. Accordingly, the scientist must consider different possibilities in his or her search for an answer and thus performs similar experiments over and over again.

The discipline of statistics deals with the *collection* and *analysis of data*. When measurements are taken, even seemingly under the same conditions, the results usually vary. Despite this variability, a statistician tries to find a pattern; yet due to the "noise," not all of the data fit into the pattern. In the face of the variability, the statistician must still determine the best way to describe the pattern. Accordingly, statisticians know that mistakes will be made in data analysis, and they try to minimize those errors as much as possible and then give bounds on the possible errors. By considering these bounds, decision makers can decide how much confidence they want to place in the data and in their analysis of them. If the bounds are wide, perhaps more data should be collected. If, however, the bounds are narrow, the person involved in the study might want to make a decision and proceed accordingly.

Variability is a fact of life, and proper statistical methods can help us understand data collected under inherent variability. Because of this variability, many decisions have to be made that involve uncertainties. In medical research, interest may center on the effectiveness of a new vaccine for mumps; an agronomist must decide whether an increase in yield can be attributed to a new strain of wheat; a meteorologist is interested in predicting the probability of rain; the state legislature must decide whether decreasing speed limits will result in fewer accidents; the admissions officer of a college must predict the college performance of an incoming freshman; a biologist is interested in estimating the clutch size for a particular type of bird; an economist desires to estimate the unemployment rate; an environmentalist tests whether new controls have resulted in a reduction in pollution.

In reviewing the preceding (relatively short) list of possible areas of applications of statistics, the reader should recognize that good statistics is closely associated with careful thinking in many investigations. As an illustration, students should appreciate how statistics is used in the endless cycle of the scientific method. We observe nature and ask questions, we run experiments and collect data that shed light on these questions, we analyze the data and compare the results of the analysis with what we previously thought, we raise new questions, and on and on. Or if you like, statistics is clearly part of the important "plan–do–study–act" cycle: Questions are raised and investigations planned and carried out. The resulting data are studied and analyzed and then acted upon, often raising new questions.

There are many aspects of statistics. Some people get interested in the subject by collecting data and trying to make sense out of their observations. In some cases the answers are obvious and little training in statistical methods is necessary. But if a person goes very far in many investigations, he or she soon realizes that there is a need for some theory to help describe the error structure associated with the various estimates of the patterns. That is, at some point appropriate probability and mathematical models are required to make sense out of complicated data sets. Statistics and the probabilistic foundation on which statistical methods are based can provide the models to help people do this. So in this book, we are more concerned with the mathematical, rather than the applied, aspects of statistics. Still, we give enough real examples so that the reader can get a good sense of a number of important applications of statistical methods.

In the study of statistics, we consider experiments for which the outcome cannot be predicted with certainty. Such experiments are called **random experiments**. Although the specific outcome of a random experiment cannot be predicted with certainty before the experiment is performed, the collection of all possible outcomes *is* known and can be described and perhaps listed. The collection of all possible outcomes is denoted by S and is called the **outcome space**. Given an outcome space S, let A be a part of the collection of outcomes in S; that is, $A \subset S$. Then A is called an **event**. When the random experiment is performed and the outcome of the experiment is in A, we say that **event** A **has occurred**.

Since, in studying probability, the words *set* and *event* are interchangeable, the reader might want to review **algebra of sets**. Here we remind the reader of some terminology:

- $\emptyset$ denotes the **null** or **empty** set;
- $A \subset B$ means A is a **subset** of B;
- $A \cup B$ is the **union** of A and B;
- $A \cap B$ is the **intersection** of A and B;
- A' is the **complement** of A (i.e., all elements in S that are not in A).

Some of these sets are depicted by the shaded regions in Figure 1.1-1, in which S is the interior of the rectangles. Such figures are called **Venn diagrams**.

Special terminology associated with events that is often used by statisticians includes the following:

1. $A_1, A_2, \ldots, A_k$ are **mutually exclusive events** means that $A_i \cap A_j = \emptyset$, $i \neq j$; that is, $A_1, A_2, \ldots, A_k$ are disjoint sets;
2. $A_1, A_2, \ldots, A_k$ are **exhaustive events** means that $A_1 \cup A_2 \cup \cdots \cup A_k = S$.

So if $A_1, A_2, \ldots, A_k$ are **mutually exclusive and exhaustive** events, we know that $A_i \cap A_j = \emptyset$, $i \neq j$, and $A_1 \cup A_2 \cup \cdots \cup A_k = S$.

Set operations satisfy several properties. For example, if A, B, and C are subsets of S, we have the following:

Commutative Laws

$$A \cup B = B \cup A$$

$$A \cap B = B \cap A$$

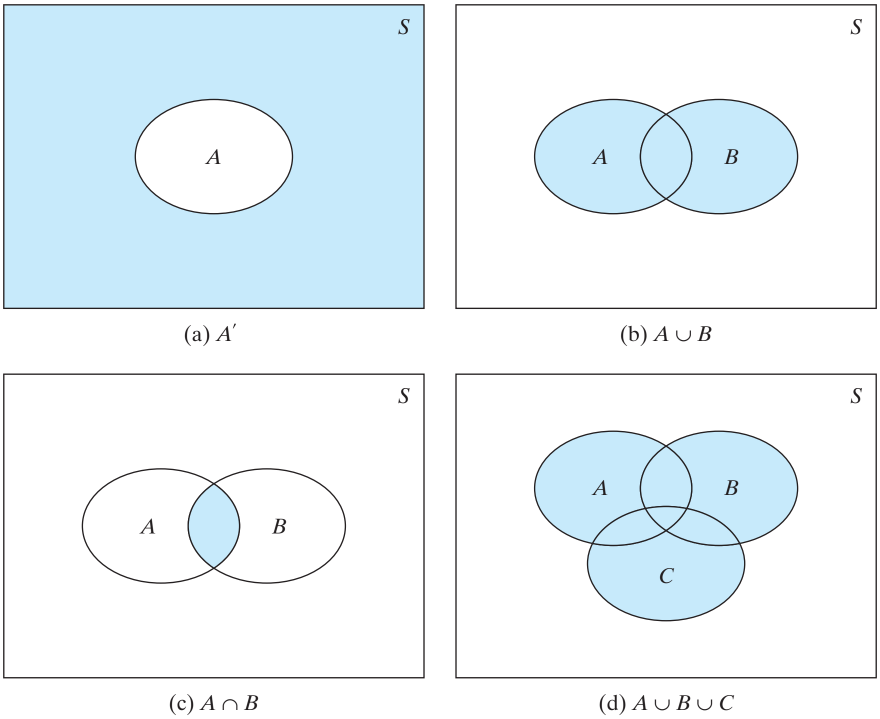


Figure 1.1-1 Algebra of sets

Associative Laws

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

Distributive Laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

De Morgan's Laws

$$(A \cup B)' = A' \cap B'$$
$$(A \cap B)' = A' \cup B'$$

A Venn diagram will be used to justify the first of De Morgan's laws. In Figure 1.1-2(a), $A \cup B$ is represented by horizontal lines, and thus $(A \cup B)'$ is the region represented by vertical lines. In Figure 1.1-2(b), A' is indicated with horizontal lines, and B' is indicated with vertical lines. An element belongs to $A' \cap B'$ if it belongs to both A' and B'. Thus the crosshatched region represents $A' \cap B'$. Clearly, this crosshatched region is the same as that shaded with vertical lines in Figure 1.1-2(a).

We are interested in defining what is meant by the probability of event A, denoted by $P(A)$ and often called the chance of A occurring. To help us understand what is meant by the probability of A, consider repeating the experiment a number of times—say, n times. Count the number of times that event A actually occurred throughout these n performances; this number is called the frequency of event A and is denoted by $\mathcal{N}(A)$. The ratio $\mathcal{N}(A)/n$ is called the **relative frequency** of event A in these n repetitions of the experiment. A relative frequency is usually very unstable for small values of n, but it tends to stabilize as n increases. This suggests that we associate with event A a number—say, p—that is equal to the number about which the relative frequency tends to stabilize. This number p can then be taken as the number that the relative frequency of event A will be near in future performances of the experiment. Thus, although we cannot predict the outcome of a random experiment with certainty, if we know p, for a large value of n, we can predict fairly accurately the relative frequency associated with event A. The number p assigned to event A is

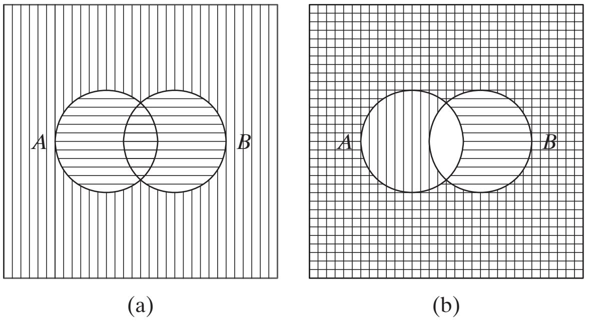


Figure 1.1-2 Venn diagrams illustrating De Morgan's laws

called the **probability** of event A and is denoted by $P(A)$. That is, $P(A)$ represents the proportion of outcomes of a random experiment that terminate in the event A as the number of trials of that experiment increases without bound.

The next example will help to illustrate some of the ideas just presented.

Example 1.1-1 A fair six-sided die is rolled six times. If the face numbered k is the outcome on roll k for $k = 1, 2, \ldots, 6$, we say that a match has occurred. The experiment is called a success if at least one match occurs during the six trials. Otherwise, the experiment is called a failure. The sample space is $S = \{\text{success, failure}\}$. Let $A = \{\text{success}\}$. We would like to assign a value to $P(A)$. Accordingly, this experiment was simulated 500 times on a computer. Figure 1.1-3 depicts the results of this simulation, and the following table summarizes a few of the results:

n	$\mathcal{N}(A)$	$\mathcal{N}(A)/n$
50	37	0.740
100	69	0.690
250	172	0.688
500	330	0.660

The probability of event A is not intuitively obvious, but it will be shown in Example 1.4-6 that $P(A) = 1 - (1 - 1/6)^6 = 0.665$. This assignment is certainly supported by the simulation (although not proved by it).

Example 1.1-1 shows that at times intuition cannot be used to assign probabilities, although simulation can perhaps help to assign a probability empirically. The next example illustrates where intuition can help in assigning a probability to an event.

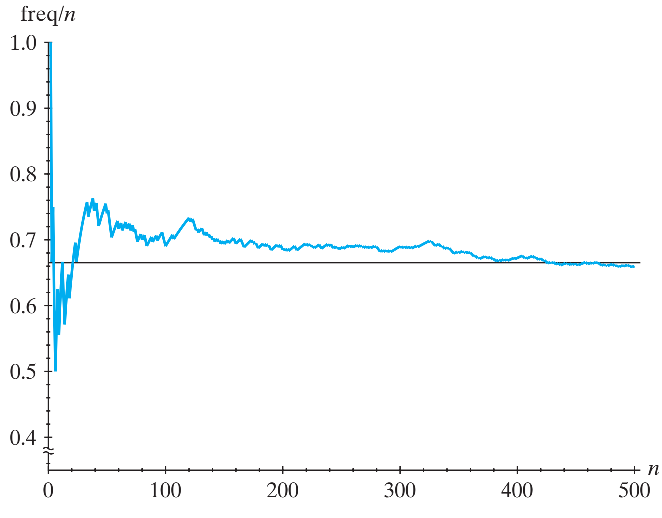


Figure 1.1-3 Fraction of experiments having at least one match

Example 1.1-2

A disk 2 inches in diameter is thrown at random on a tiled floor, where each tile is a square with sides 4 inches in length. Let C be the event that the disk will land entirely on one tile. In order to assign a value to $P(C)$, consider the center of the disk. In what region must the center lie to ensure that the disk lies entirely on one tile? If you draw a picture, it should be clear that the center must lie within a square having sides of length 2 and with its center coincident with the center of a tile. Since the area of this square is 4 and the area of a tile is 16, it makes sense to let $P(C) = 4/16$. ■

Sometimes the nature of an experiment is such that the probability of A can be assigned easily. For example, when a state lottery randomly selects a three-digit integer, we would expect each of the 1000 possible three-digit numbers to have the same chance of being selected, namely, 1/1000. If we let $A = \{233, 323, 332\}$, then it makes sense to let $P(A) = 3/1000$. Or if we let $B = \{234, 243, 324, 342, 423, 432\}$, then we would let $P(B) = 6/1000$, the probability of the event B. Probabilities of events associated with many random experiments are perhaps not quite as obvious and straightforward as was seen in Example 1.1-1.

So we wish to associate with A a number $P(A)$ about which the relative frequency $\mathcal{N}(A)/n$ of the event A tends to stabilize with large n. A function such as $P(A)$ that is evaluated for a set A is called a **set function**. In this section, we consider the probability set function $P(A)$ and discuss some of its properties. In succeeding sections, we shall describe how the probability set function is defined for particular experiments.

To help decide what properties the probability set function should satisfy, consider properties possessed by the relative frequency $\mathcal{N}(A)/n$. For example, $\mathcal{N}(A)/n$ is always nonnegative. If $A = S$, the sample space, then the outcome of the experiment will always belong to S, and thus $\mathcal{N}(S)/n = 1$. Also, if A and B are two mutually exclusive events, then $\mathcal{N}(A \cup B)/n = \mathcal{N}(A)/n + \mathcal{N}(B)/n$. Hopefully, these remarks will help to motivate the following definition.

Definition 1.1-1

Probability is a real-valued set function P that assigns, to each event A in the sample space S, a number $P(A)$, called the probability of the event A, such that the following properties are satisfied:

(a) $P(A) \geq 0$;

(b) $P(S) = 1$;

(c) if $A_1, A_2, A_3, \ldots$ are events and $A_i \cap A_j = \emptyset$, $i \neq j$, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_k) = P(A_1) + P(A_2) + \cdots + P(A_k)$$

for each positive integer k, and

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

for an infinite, but countable, number of events.

The theorems that follow give some other important properties of the probability set function. When one considers these theorems, it is important to understand the theoretical concepts and proofs. However, if the reader keeps the relative frequency concept in mind, the theorems should also have some intuitive appeal.

Theorem 1.1-1 For each event A,

$$P(A) = 1 - P(A').$$

Proof [See Figure 1.1-1(a).] We have

$$S = A \cup A' \qquad \text{and} \qquad A \cap A' = \emptyset.$$

Thus, from properties (b) and (c), it follows that

$$1 = P(A) + P(A').$$

Hence

$$P(A) = 1 - P(A'). \qquad \square$$

Example 1.1-3 A fair coin is flipped successively until the same face is observed on successive flips. Let $A = \{x : x = 3, 4, 5, \ldots\}$; that is, A is the event that it will take three or more flips of the coin to observe the same face on two consecutive flips. To find $P(A)$, we first find the probability of $A' = \{x : x = 2\}$, the complement of A. In two flips of a coin, the possible outcomes are $\{HH, HT, TH, TT\}$, and we assume that each of these four points has the same chance of being observed. Thus,

$$P(A') = P(\{HH, TT\}) = \frac{2}{4}.$$

It follows from Theorem 1.1-1 that

$$P(A) = 1 - P(A') = 1 - \frac{2}{4} = \frac{2}{4}. \qquad \blacksquare$$

Theorem 1.1-2 $P(\emptyset) = 0.$

Proof In Theorem 1.1-1, take $A = \emptyset$ so that $A' = S$. Then

$$P(\emptyset) = 1 - P(S) = 1 - 1 = 0. \qquad \square$$

Theorem 1.1-3 If events A and B are such that $A \subset B$, then $P(A) \le P(B)$.

Proof We have

$$B = A \cup (B \cap A') \qquad \text{and} \qquad A \cap (B \cap A') = \emptyset.$$

Hence, from property (c),

$$P(B) = P(A) + P(B \cap A') \ge P(A)$$

because, from property (a),

$$P(B \cap A') \ge 0. \qquad \square$$

Theorem 1.1-4 For each event A, $P(A) \le 1$.

Proof Since $A \subset S$, we have, by Theorem 1.1-3 and property (b),

$$P(A) \le P(S) = 1,$$

which gives the desired result. □

Property (a), along with Theorem 1.1-4, shows that, for each event A,

$$0 \le P(A) \le 1.$$

Theorem 1.1-5 If A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Proof [See Figure 1.1-1(b).] The event $A \cup B$ can be represented as a union of mutually exclusive events, namely,

$$A \cup B = A \cup (A' \cap B).$$

Hence, by property (c),

$$P(A \cup B) = P(A) + P(A' \cap B). \tag{1.1-1}$$

However,

$$B = (A \cap B) \cup (A' \cap B),$$

which is a union of mutually exclusive events. Thus,

$$P(B) = P(A \cap B) + P(A' \cap B)$$

and

$$P(A' \cap B) = P(B) - P(A \cap B).$$

If the right-hand side of this equation is substituted into Equation 1.1-1, we obtain

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

which is the desired result. □

Example 1.1-4 A faculty leader was meeting two students in Paris, one arriving by train from Amsterdam and the other arriving by train from Brussels at approximately the same time. Let A and B be the events that the respective trains are on time. Suppose we know from past experience that $P(A) = 0.93$, $P(B) = 0.89$, and $P(A \cap B) = 0.87$. Then

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0.93 + 0.89 - 0.87 = 0.95 \end{aligned}$$

is the probability that at least one train is on time. ■

Theorem 1.1-6 If A, B, and C are any three events, then

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Proof [See Figure 1.1-1(d).] Write

$$A \cup B \cup C = A \cup (B \cup C)$$

and apply Theorem 1.1-5. The details are left as an exercise. □

Example 1.1-5 A survey was taken of a group's viewing habits of sporting events on TV during the last year. Let $A = \{\text{watched football}\}$, $B = \{\text{watched basketball}\}$, $C = \{\text{watched baseball}\}$. The results indicate that if a person is selected at random from the surveyed group, then $P(A) = 0.43$, $P(B) = 0.40$, $P(C) = 0.32$, $P(A \cap B) = 0.29$, $P(A \cap C) = 0.22$, $P(B \cap C) = 0.20$, and $P(A \cap B \cap C) = 0.15$. It then follows that

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C) \\ &= 0.43 + 0.40 + 0.32 - 0.29 - 0.22 - 0.20 + 0.15 \\ &= 0.59 \end{aligned}$$

is the probability that this person watched at least one of these sports. ■

Let a probability set function be defined on a sample space S. Let $S = \{e_1, e_2, \ldots, e_m\}$, where each e_i is a possible outcome of the experiment. The integer m is called the total number of ways in which the random experiment can terminate. If each of these outcomes has the same probability of occurring, we say that the m outcomes are **equally likely**. That is,

$$P(\{e_i\}) = \frac{1}{m}, \qquad i = 1, 2, \ldots, m.$$

If the number of outcomes in an event A is h, then the integer h is called the number of ways that are favorable to the event A. In this case, $P(A)$ is equal to the number of ways favorable to the event A divided by the total number of ways in which the experiment can terminate. That is, under this assumption of equally likely outcomes, we have

$$P(A) = \frac{h}{m} = \frac{N(A)}{N(S)},$$

where $h = N(A)$ is the number of ways A can occur and $m = N(S)$ is the number of ways S can occur. Exercise 1.1-15 considers this assignment of probability in a more theoretical manner.

It should be emphasized that in order to assign the probability h/m to the event A, we must assume that each of the outcomes $e_1, e_2, \ldots, e_m$ has the same probability $1/m$. This assumption is then an important part of our probability model; if it is not realistic in an application, then the probability of the event A cannot be computed in this way. Actually, we have used this result in the simple case given in Example 1.1-3 because it seemed realistic to assume that each of the possible outcomes in $S = \{HH, HT, TH, TT\}$ had the same chance of being observed.

Example 1.1-6 Let a card be drawn at random from an ordinary deck of 52 playing cards. Then the sample space S is the set of $m = 52$ different cards, and it is reasonable to assume that each of these cards has the same probability of selection, 1/52. Accordingly, if A is the set of outcomes that are kings, then $P(A) = 4/52 = 1/13$ because there are $h = 4$ kings in the deck. That is, 1/13 is the probability of drawing a card that is a king, provided that each of the 52 cards has the same probability of being drawn. ■

In Example 1.1-6, the computations are very easy because there is no difficulty in the determination of the appropriate values of h and m. However, instead of drawing only one card, suppose that 13 are taken at random and without replacement. Then we can think of each possible 13-card hand as being an outcome in a sample space, and it is reasonable to assume that each of these outcomes has the same probability. For example, using the preceding method to assign the probability of a hand consisting of seven spades and six hearts, we must be able to count the number h of all such hands as well as the number m of possible 13-card hands. In these more complicated situations, we need better methods of determining h and m. We discuss some of these counting techniques in Section 1.2.

Exercises

1.1-1. Of a group of patients having injuries, 28% visit both a physical therapist and a chiropractor and 8% visit neither. Say that the probability of visiting a physical therapist exceeds the probability of visiting a chiropractor by 16%. What is the probability of a randomly selected person from this group visiting a physical therapist?

1.1-2. An insurance company looks at its auto insurance customers and finds that (a) all insure at least one car, (b) 85% insure more than one car, (c) 23% insure a sports car, and (d) 17% insure more than one car, including a sports car. Find the probability that a customer selected at random insures exactly one car and it is not a sports car.

1.1-3. Draw one card at random from a standard deck of cards. The sample space S is the collection of the 52 cards. Assume that the probability set function assigns 1/52 to each of the 52 outcomes. Let

$$A = \{x : x \text{ is a jack, queen, or king}\},$$

$$B = \{x : x \text{ is a 9, 10, or jack and } x \text{ is red}\},$$

$$C = \{x : x \text{ is a club}\},$$

$$D = \{x : x \text{ is a diamond, a heart, or a spade}\}.$$

Find **(a)** $P(A)$, **(b)** $P(A \cap B)$, **(c)** $P(A \cup B)$, **(d)** $P(C \cup D)$, and **(e)** $P(C \cap D)$.

1.1-4. A fair coin is tossed four times, and the sequence of heads and tails is observed.

(a) List each of the 16 sequences in the sample space S.

(b) Let events A, B, C, and D be given by $A = \{\text{at least 3 heads}\}$, $B = \{\text{at most 2 heads}\}$, $C = \{\text{heads on the third toss}\}$, and $D = \{\text{1 head and 3 tails}\}$. If the probability set function assigns 1/16 to each outcome in the sample space, find **(i)** $P(A)$, **(ii)** $P(A \cap B)$, **(iii)** $P(B)$, **(iv)** $P(A \cap C)$, **(v)** $P(D)$, **(vi)** $P(A \cup C)$, and **(vii)** $P(B \cap D)$.

1.1-5. Consider the trial on which a 3 is first observed in successive rolls of a six-sided die. Let A be the event that 3 is observed on the first trial. Let B be the event that at least two trials are required to observe a 3. Assuming that each side has probability 1/6, find **(a)** $P(A)$, **(b)** $P(B)$, and **(c)** $P(A \cup B)$.

1.1-6. If $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.3$, find **(a)** $P(A \cup B)$, **(b)** $P(A \cap B')$, and **(c)** $P(A' \cup B')$.

1.1-7. Given that $P(A \cup B) = 0.76$ and $P(A \cup B') = 0.87$, find $P(A)$.

1.1-8. During a visit to a primary care physician's office, the probability of having neither lab work nor referral to a specialist is 0.21. Of those coming to that office, the probability of having lab work is 0.41 and the probability of having a referral is 0.53. What is the probability of having both lab work and a referral?

1.1-9. Roll a fair six-sided die three times. Let $A_1 = \{\text{1 or 2 on the first roll}\}$, $A_2 = \{\text{3 or 4 on the second roll}\}$, and $A_3 = \{\text{5 or 6 on the third roll}\}$. It is given that $P(A_i) = 1/3$, $i = 1, 2, 3$; $P(A_i \cap A_j) = (1/3)^2$, $i \neq j$; and $P(A_1 \cap A_2 \cap A_3) = (1/3)^3$.

(a) Use Theorem 1.1-6 to find $P(A_1 \cup A_2 \cup A_3)$.

(b) Show that $P(A_1 \cup A_2 \cup A_3) = 1 - (1 - 1/3)^3$.

1.1-10. Prove Theorem 1.1-6.

1.1-11. A typical roulette wheel used in a casino has 38 slots that are numbered $1, 2, 3, \ldots, 36, 0, 00$, respectively. The 0 and 00 slots are colored green. Half of the remaining slots are red and half are black. Also, half of the integers between 1 and 36 inclusive are odd, half are even, and 0 and 00 are defined to be neither odd nor even. A ball is rolled around the wheel and ends up in one of the slots; we assume that each slot has equal probability of 1/38, and we are interested in the number of the slot into which the ball falls.

(a) Define the sample space S.

(b) Let $A = \{0, 00\}$. Give the value of $P(A)$.

(c) Let $B = \{14, 15, 17, 18\}$. Give the value of $P(B)$.

(d) Let $D = \{x : x \text{ is odd}\}$. Give the value of $P(D)$.

1.1-12. Let x equal a number that is selected randomly from the closed interval from zero to one, $[0, 1]$. Use your intuition to assign values to

(a) $P(\{x: 0 \le x \le 1/3\})$.

(b) $P(\{x: 1/3 \le x \le 1\})$.

(c) $P(\{x: x = 1/3\})$.

(d) $P(\{x: 1/2 < x < 5\})$.

1.1-13. Divide a line segment into two parts by selecting a point at random. Use your intuition to assign a probability to the event that the longer segment is at least two times longer than the shorter segment.

1.1-14. Let the interval $[-r, r]$ be the base of a semicircle. If a point is selected at random from this interval, assign a probability to the event that the length of the perpendicular segment from the point to the semicircle is less than $r/2$.

1.1-15. Let $S = A_1 \cup A_2 \cup \cdots \cup A_m$, where events $A_1, A_2, \ldots, A_m$ are mutually exclusive and exhaustive.

(a) If $P(A_1) = P(A_2) = \cdots = P(A_m)$, show that $P(A_i) = 1/m$, $i = 1, 2, \ldots, m$.

(b) If $A = A_1 \cup A_2 \cup \cdots \cup A_h$, where $h < m$, and (a) holds, prove that $P(A) = h/m$.

1.1-16. Let p_n, $n = 0, 1, 2, \ldots$, be the probability that an automobile policyholder will file for n claims in a five-year period. The actuary involved makes the assumption that $p_{n+1} = (1/4)p_n$. What is the probability that the holder will file two or more claims during this period?

1.2 METHODS OF ENUMERATION

In this section, we develop counting techniques that are useful in determining the number of outcomes associated with the events of certain random experiments. We begin with a consideration of the multiplication principle.

Multiplication Principle: Suppose that an experiment (or procedure) E_1 has n_1 outcomes and, for each of these possible outcomes, an experiment (procedure) E_2 has n_2 possible outcomes. Then the composite experiment (procedure) E_1E_2 that consists of performing first E_1 and then E_2 has n_1n_2 possible outcomes.

Example 1.2-1 Let E_1 denote the selection of a rat from a cage containing one female (F) rat and one male (M) rat. Let E_2 denote the administering of either drug A (A), drug B (B), or a placebo (P) to the selected rat. Then the outcome for the composite experiment can be denoted by an ordered pair, such as (F, P). In fact, the set of all possible outcomes, namely, $(2)(3) = 6$ of them, can be denoted by the following rectangular array:

(F, A) (F, B) (F, P)
(M, A) (M, B) (M, P)

■

Another way of illustrating the multiplication principle is with a tree diagram like that in Figure 1.2-1. The diagram shows that there are $n_1 = 2$ possibilities (branches) for the sex of the rat and that, for each of these outcomes, there are $n_2 = 3$ possibilities (branches) for the drug.

Clearly, the multiplication principle can be extended to a sequence of more than two experiments or procedures. Suppose that the experiment E_i has

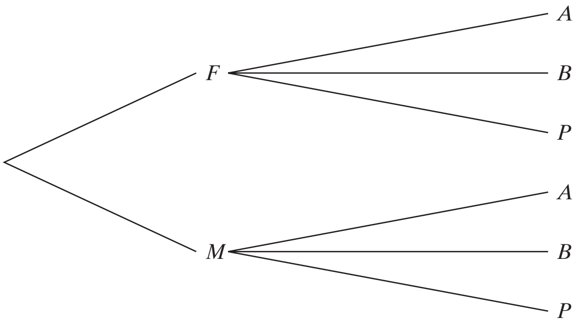


Figure 1.2-1 Tree diagram

n_i ($i = 1, 2, \ldots, m$) possible outcomes after previous experiments have been performed. Then the composite experiment $E_1 E_2 \cdots E_m$ that consists of performing E_1, then $E_2, \ldots,$ and finally E_m has $n_1 n_2 \cdots n_m$ possible outcomes.

Example 1.2-2

A certain food service gives the following choices for dinner: E_1, soup or tomato juice; E_2, steak or shrimp; E_3, French fried potatoes, mashed potatoes, or a baked potato; E_4, corn or peas; E_5, jello, tossed salad, cottage cheese, or coleslaw; E_6, cake, cookies, pudding, brownie, vanilla ice cream, chocolate ice cream, or orange sherbet; E_7, coffee, tea, milk, or punch. How many different dinner selections are possible if one of the listed choices is made for each of $E_1, E_2, \ldots,$ and E_7? By the multiplication principle, there are

$$(2)(2)(3)(2)(4)(7)(4) = 2688$$

different combinations. ■

Although the multiplication principle is fairly simple and easy to understand, it will be extremely useful as we now develop various counting techniques.

Suppose that n positions are to be filled with n different objects. There are n choices for filling the first position, $n-1$ for the second, $\ldots$, and 1 choice for the last position. So, by the multiplication principle, there are

$$n(n-1)\cdots(2)(1) = n!$$

possible arrangements. The symbol $n!$ is read "n factorial." We define $0! = 1$; that is, we say that zero positions can be filled with zero objects in one way.

Definition 1.2-1

Each of the $n!$ arrangements (in a row) of n different objects is called a **permutation** of the n objects.

Example 1.2-3

The number of permutations of the four letters a, b, c, and d is clearly $4! = 24$. However, the number of possible four-letter code words using the four letters a, b, c, and d if letters may be repeated is $4^4 = 256$, because in this case each selection can be performed in four ways. ■

If only r positions are to be filled with objects selected from n different objects, $r \le n$, then the number of possible ordered arrangements is

$$_nP_r = n(n-1)(n-2)\cdots(n-r+1).$$

That is, there are n ways to fill the first position, $(n-1)$ ways to fill the second, and so on, until there are $[n-(r-1)] = (n-r+1)$ ways to fill the rth position.

In terms of factorials, we have

$$_nP_r = \frac{n(n-1)\cdots(n-r+1)(n-r)\cdots(3)(2)(1)}{(n-r)\cdots(3)(2)(1)} = \frac{n!}{(n-r)!}.$$

Definition 1.2-2
Each of the $_nP_r$ arrangements is called a **permutation of n objects taken r at a time**.

Example 1.2-4 The number of possible four-letter code words, selecting from the 26 letters in the alphabet, in which all four letters are different is

$$_{26}P_4 = (26)(25)(24)(23) = \frac{26!}{22!} = 358{,}800.$$

■

Example 1.2-5 The number of ways of selecting a president, a vice president, a secretary, and a treasurer in a club consisting of 10 persons is

$$_{10}P_4 = 10 \cdot 9 \cdot 8 \cdot 7 = \frac{10!}{6!} = 5040.$$

■

Suppose that a set contains n objects. Consider the problem of drawing r objects from this set. The order in which the objects are drawn may or may not be important. In addition, it is possible that a drawn object is replaced before the next object is drawn. Accordingly, we give some definitions and show how the multiplication principle can be used to count the number of possibilities.

Definition 1.2-3
If r objects are selected from a set of n objects, and if the order of selection is noted, then the selected set of r objects is called an **ordered sample of size** r.

Definition 1.2-4
Sampling with replacement occurs when an object is selected and then replaced before the next object is selected.

By the multiplication principle, the number of possible ordered samples of size r taken from a set of n objects is n^r when sampling with replacement.

Example 1.2-6 A die is rolled seven times. The number of possible ordered samples is $6^7 = 279{,}936$. Note that rolling a die is equivalent to sampling with replacement from the set $\{1,2,3,4,5,6\}$.

■

Example 1.2-7

An urn contains 10 balls numbered 0, 1, 2, …, 9. If 4 balls are selected one at a time and with replacement, then the number of possible ordered samples is $10^4 = 10{,}000$. Note that this is the number of four-digit integers between 0000 and 9999, inclusive. ■

Definition 1.2-5

Sampling without replacement occurs when an object is not replaced after it has been selected.

By the multiplication principle, the number of possible ordered samples of size r taken from a set of n objects without replacement is

$$n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!},$$

which is equivalent to ${}_nP_r$, the number of permutations of n objects taken r at a time.

Example 1.2-8

The number of ordered samples of 5 cards that can be drawn without replacement from a standard deck of 52 playing cards is

$$(52)(51)(50)(49)(48) = \frac{52!}{47!} = 311{,}875{,}200.$$ ■

REMARK Note that it must be true that $r \le n$ when sampling without replacement, but r can exceed n when sampling with replacement. ■

Often the order of selection is not important and interest centers only on the selected set of r objects. That is, we are interested in the number of subsets of size r that can be selected from a set of n different objects. In order to find the number of (unordered) subsets of size r, we count, in two different ways, the number of ordered subsets of size r that can be taken from the n distinguishable objects. Then, equating the two answers, we are able to count the number of (unordered) subsets of size r.

Let C denote the number of (unordered) subsets of size r that can be selected from n different objects. We can obtain each of the ${}_nP_r$ ordered subsets by first selecting one of the C unordered subsets of r objects and then ordering these r objects. Since the latter ordering can be carried out in $r!$ ways, the multiplication principle yields $(C)(r!)$ ordered subsets; so $(C)(r!)$ must equal ${}_nP_r$. Thus, we have

$$(C)(r!) = \frac{n!}{(n-r)!},$$

or

$$C = \frac{n!}{r!\,(n-r)!}.$$

We denote this answer by either ${}_nC_r$ or $\binom{n}{r}$; that is,

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

Accordingly, a set of n different objects possesses

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

unordered subsets of size $r \le n$.

We could also say that the number of ways in which r objects can be selected without replacement from n objects when the order of selection is disregarded is $\binom{n}{r} = {}_nC_r$, and the latter expression can be read as "n choose r." This result motivates the next definition.

Definition 1.2-6
Each of the ${}_nC_r$ unordered subsets is called a **combination of n objects taken r at a time**, where

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

Example 1.2-9 The number of possible 5-card hands (in 5-card poker) drawn from a deck of 52 playing cards is

$${}_{52}C_5 = \binom{52}{5} = \frac{52!}{5!\,47!} = 2{,}598{,}960.$$

Example 1.2-10 The number of possible 13-card hands (in bridge) that can be selected from a deck of 52 playing cards is

$${}_{52}C_{13} = \binom{52}{13} = \frac{52!}{13!\,39!} = 635{,}013{,}559{,}600.$$

The numbers $\binom{n}{r}$ are frequently called **binomial coefficients**, since they arise in the expansion of a binomial. We illustrate this property by giving a justification of the binomial expansion

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} b^r a^{n-r}. \tag{1.2-1}$$

For each summand in the expansion of

$$(a+b)^n = (a+b)(a+b)\cdots(a+b),$$

either an a or a b is selected from each of the n factors. One possible product is then $b^r a^{n-r}$; this occurs when b is selected from each of r factors and a from each of the remaining $n-r$ factors. But the latter operation can be completed in $\binom{n}{r}$ ways, which then must be the coefficient of $b^r a^{n-r}$, as shown in Equation 1.2-1.

The binomial coefficients are given in Table I in Appendix B for selected values of n and r. Note that for some combinations of n and r, the table uses the fact that

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

That is, the number of ways in which r objects can be selected out of n objects is equal to the number of ways in which $n - r$ objects can be selected out of n objects.

Example 1.2-11 Assume that each of the $\binom{52}{5} = 2{,}598{,}960$ five-card hands drawn from a deck of 52 playing cards has the same probability of being selected. Then the number of possible 5-card hands that are all spades (event A) is

$$N(A) = \binom{13}{5}\binom{39}{0},$$

because the 5 spades can be selected from the 13 spades in $\binom{13}{5}$ ways, after which zero nonspades can be selected in $\binom{39}{0} = 1$ way. We have

$$\binom{13}{5} = \frac{13!}{5!8!} = 1287$$

from Table I in Appendix B. Thus, the probability of an all-spade five-card hand is

$$P(A) = \frac{N(A)}{N(S)} = \frac{1287}{2{,}598{,}960} = 0.000495.$$

Suppose now that the event B is the set of outcomes in which exactly three cards are kings and exactly two cards are queens. We can select the three kings in any one of $\binom{4}{3}$ ways and the two queens in any one of $\binom{4}{2}$ ways. By the multiplication principle, the number of outcomes in B is

$$N(B) = \binom{4}{3}\binom{4}{2}\binom{44}{0},$$

where $\binom{44}{0}$ gives the number of ways in which 0 cards are selected out of the nonkings and nonqueens and of course is equal to 1. Thus,

$$P(B) = \frac{N(B)}{N(S)} = \frac{\binom{4}{3}\binom{4}{2}\binom{44}{0}}{\binom{52}{5}} = \frac{24}{2{,}598{,}960} = 0.0000092.$$

Finally, let C be the set of outcomes in which there are exactly two kings, two queens, and one jack. Then

$$P(C) = \frac{N(C)}{N(S)} = \frac{\binom{4}{2}\binom{4}{2}\binom{4}{1}\binom{40}{0}}{\binom{52}{5}} = \frac{144}{2{,}598{,}960} = 0.000055$$

because the numerator of this fraction is the number of outcomes in C. ■

Now suppose that a set contains n objects of two types: r of one type and $n-r$ of the other type. The number of permutations of n different objects is $n!$. However, in this case, the objects are not all distinguishable. To count the number of distinguishable arrangements, first select r out of the n positions for the objects of the first type.

This can be done in $\binom{n}{r}$ ways. Then fill in the remaining positions with the objects of the second type. Thus, the number of distinguishable arrangements is

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}. $$

Definition 1.2-7
Each of the ${}_nC_r$ permutations of n objects, r of one type and $n-r$ of another type, is called a **distinguishable permutation**.

Example 1.2-12 A coin is flipped 10 times and the sequence of heads and tails is observed. The number of possible 10-tuplets that result in four heads and six tails is

$$ \binom{10}{4} = \frac{10!}{4!\,6!} = \frac{10!}{6!\,4!} = \binom{10}{6} = 210. $$

■

Example 1.2-13 In an orchid show, seven orchids are to be placed along one side of the greenhouse. There are four lavender orchids and three white orchids. Considering only the color of the orchids, we see that the number of lineups of the orchids is

$$ \binom{7}{4} = \frac{7!}{4!\,3!} = 35. $$

If the colors of the seven orchids are white, lavender, yellow, mauve, crimson, orange, and pink, the number of different displays is $7! = 5040$. ■

The foregoing results can be extended. Suppose that in a set of n objects, n_1 are similar, n_2 are similar, $\ldots$, n_s are similar, where $n_1 + n_2 + \cdots + n_s = n$. Then the number of distinguishable permutations of the n objects is (see Exercise 1.2-15)

$$ \binom{n}{n_1, n_2, \ldots, n_s} = \frac{n!}{n_1!\,n_2!\,\cdots\,n_s!}. \qquad (1.2\text{-}2) $$

Example 1.2-14 Among nine orchids for a line of orchids along one wall, three are white, four lavender, and two yellow. The number of different color displays is then

$$ \binom{9}{3, 4, 2} = \frac{9!}{3!\,4!\,2!} = 1260. $$

■

The argument used in determining the binomial coefficients in the expansion of $(a+b)^n$ can be extended to find the expansion of $(a_1+a_2+\cdots+a_s)^n$. The coefficient of ${a_1}^{n_1}{a_2}^{n_2}\cdots{a_s}^{n_s}$, where $n_1 + n_2 + \cdots + n_s = n$, is

$$ \binom{n}{n_1, n_2, \ldots, n_s} = \frac{n!}{n_1!\,n_2!\,\cdots\,n_s!}. $$

This is sometimes called a **multinomial coefficient**.

When r objects are selected out of n objects, we are often interested in the number of possible outcomes. We have seen that for ordered samples, there are

n^r possible outcomes when sampling with replacement and ${}_nP_r$ outcomes when sampling without replacement. For unordered samples, there are ${}_nC_r$ outcomes when sampling without replacement. Each of the preceding outcomes is equally likely, provided that the experiment is performed in a fair manner.

REMARK Although not needed as often in the study of probability, it is interesting to count the number of possible samples of size r that can be selected out of n objects when the order is irrelevant and when sampling with replacement. For example, if a six-sided die is rolled 10 times (or 10 six-sided dice are rolled once), how many possible unordered outcomes are there? To count the number of possible outcomes, think of listing r 0's for the r objects that are to be selected. Then insert $(n - 1)$ |'s to partition the r objects into n sets, the first set giving objects of the first kind, and so on. So if $n = 6$ and $r = 10$ in the die illustration, a possible outcome is

$$0\,0\,|\,|\,0\,0\,0\,|\,0\,|\,0\,0\,0\,|\,0,$$

which says there are two 1's, zero 2's, three 3's, one 4, three 5's, and one 6. In general, each outcome is a permutation of r 0's and $(n - 1)$ |'s. Each distinguishable permutation is equivalent to an unordered sample. The number of distinguishable permutations, and hence the number of unordered samples of size r that can be selected out of n objects when sampling with replacement, is

$$_{n-1+r}C_r = \frac{(n-1+r)!}{r!\,(n-1)!}.$$

Exercises

1.2-1. A boy found a bicycle lock for which the combination was unknown. The correct combination is a four-digit number, $d_1d_2d_3d_4$, where d_i, $i = 1, 2, 3, 4$, is selected from 1, 2, 3, 4, 5, 6, 7, and 8. How many different lock combinations are possible with such a lock?

1.2-2. In designing an experiment, the researcher can often choose many different levels of the various factors in order to try to find the best combination at which to operate. As an illustration, suppose the researcher is studying a certain chemical reaction and can choose four levels of temperature, five different pressures, and two different catalysts.

(a) To consider all possible combinations, how many experiments would need to be conducted?

(b) Often in preliminary experimentation, each factor is restricted to two levels. With the three factors noted, how many experiments would need to be run to cover all possible combinations with each of the three factors at two levels? (NOTE: This is often called a 2^3 design.)

1.2-3. How many different license plates are possible if a state uses

(a) Two letters followed by a four-digit integer (leading zeros are permissible and the letters and digits can be repeated)?

(b) Three letters followed by a three-digit integer? (In practice, it is possible that certain "spellings" are ruled out.)

1.2-4. The "eating club" is hosting a make-your-own sundae at which the following are provided:

Ice Cream Flavors	Toppings
Chocolate	Caramel
Cookies 'n' cream	Hot fudge
Strawberry	Marshmallow
Vanilla	M&M's
	Nuts
	Strawberries

(a) How many sundaes are possible using one flavor of ice cream and three different toppings?

(b) How many sundaes are possible using one flavor of ice cream and from zero to six toppings?

(c) How many different combinations of flavors of three scoops of ice cream are possible if it is permissible to make all three scoops the same flavor?

1.2-5. How many four-letter code words are possible using the letters in IOWA if

(a) The letters may not be repeated?

(b) The letters may be repeated?

1.2-6. Suppose that Novak Djokovic and Roger Federer are playing a tennis match in which the first player to win three sets wins the match. Using **D** and **F** for the winning player of a set, in how many ways could this tennis match end?

1.2-7. In a state lottery, four digits are drawn at random one at a time with replacement from 0 to 9. Suppose that you win if any permutation of your selected integers is drawn. Give the probability of winning if you select

(a) 6, 7, 8, 9.

(b) 6, 7, 8, 8.

(c) 7, 7, 8, 8.

(d) 7, 8, 8, 8.

1.2-8. How many different varieties of pizza can be made if you have the following choice: small, medium, or large size; thin 'n' crispy, hand-tossed, or pan crust; and 12 toppings (cheese is automatic), from which you may select from 0 to 12?

1.2-9. The World Series in baseball continues until either the American League team or the National League team wins four games. How many different orders are possible (e.g., *ANNAAA* means the American League team wins in six games) if the series goes

(a) Four games?

(b) Five games?

(c) Six games?

(d) Seven games?

1.2-10. Pascal's triangle gives a method for calculating the binomial coefficients; it begins as follows:

$$\begin{array}{ccccccccccc} & & & & & 1 & & & & & \\ & & & & 1 & & 1 & & & & \\ & & & 1 & & 2 & & 1 & & & \\ & & 1 & & 3 & & 3 & & 1 & & \\ & 1 & & 4 & & 6 & & 4 & & 1 & \\ 1 & & 5 & & 10 & & 10 & & 5 & & 1 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

The nth row of this triangle gives the coefficients for $(a+b)^{n-1}$. To find an entry in the table other than a 1 on the boundary, add the two nearest numbers in the row directly above. The equation

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1},$$

called **Pascal's equation**, explains why Pascal's triangle works. Prove that this equation is correct.

1.2-11. Three students (S) and six faculty members (F) are on a panel discussing a new college policy.

(a) In how many different ways can the nine participants be lined up at a table in the front of the auditorium?

(b) How many lineups are possible, considering only the labels S and F?

(c) For each of the nine participants, you are to decide whether the participant did a good job or a poor job stating his or her opinion of the new policy; that is, give each of the nine participants a grade of G or P. How many different "scorecards" are possible?

1.2-12. Prove

$$\sum_{r=0}^{n}(-1)^r\binom{n}{r} = 0 \qquad \text{and} \qquad \sum_{r=0}^{n}\binom{n}{r} = 2^n.$$

HINT: Consider $(1-1)^n$ and $(1+1)^n$, or use Pascal's equation and proof by induction.

1.2-13. A bridge hand is found by taking 13 cards at random and without replacement from a deck of 52 playing cards. Find the probability of drawing each of the following hands.

(a) One in which there are 5 spades, 4 hearts, 3 diamonds, and 1 club.

(b) One in which there are 5 spades, 4 hearts, 2 diamonds, and 2 clubs.

(c) One in which there are 5 spades, 4 hearts, 1 diamond, and 3 clubs.

(d) Suppose you are dealt 5 cards of one suit, 4 cards of another. Would the probability of having the other suits split 3 and 1 be greater than the probability of having them split 2 and 2?

1.2-14. A bag of 36 dum-dum pops (suckers) contains up to 10 flavors. That is, there are from 0 to 36 suckers of each of 10 flavors in the bag. How many different flavor combinations are possible?

1.2-15. Prove Equation 1.2-2. HINT: First select n_1 positions in $\binom{n}{n_1}$ ways. Then select n_2 from the remaining $n-n_1$ positions in $\binom{n-n_1}{n_2}$ ways, and so on. Finally, use the multiplication rule.

1.2-16. A box of candy hearts contains 52 hearts, of which 19 are white, 10 are tan, 7 are pink, 3 are purple, 5 are yellow, 2 are orange, and 6 are green. If you select nine pieces of candy randomly from the box, without replacement, give the probability that

(a) Three of the hearts are white.

(b) Three are white, two are tan, one is pink, one is yellow, and two are green.

1.2-17. A poker hand is defined as drawing 5 cards at random without replacement from a deck of 52 playing cards. Find the probability of each of the following poker hands:

(a) Four of a kind (four cards of equal face value and one card of a different value).

(b) Full house (one pair and one triple of cards with equal face value).

(c) Three of a kind (three equal face values plus two cards of different values).

(d) Two pairs (two pairs of equal face value plus one card of a different value).

(e) One pair (one pair of equal face value plus three cards of different values).

1.3 CONDITIONAL PROBABILITY

We introduce the idea of conditional probability by means of an example.

Example 1.3-1

Suppose that we are given 20 tulip bulbs that are similar in appearance and told that 8 will bloom early, 12 will bloom late, 13 will be red, and 7 will be yellow, in accordance with the various combinations listed in Table 1.3-1. If one bulb is selected at random, the probability that it will produce a red tulip (R) is given by $P(R) = 13/20$, under the assumption that each bulb is "equally likely." Suppose, however, that close examination of the bulb will reveal whether it will bloom early (E) or late (L). If we consider an outcome only if it results in a tulip bulb that will bloom early, only eight outcomes in the sample space are now of interest. Thus, under this limitation, it is natural to assign the probability 5/8 to R; that is, $P(R\,|\,E) = 5/8$, where $P(R\,|\,E)$ is read as the probability of R given that E has occurred. Note that

$$P(R\,|\,E) = \frac{5}{8} = \frac{N(R \cap E)}{N(E)} = \frac{N(R \cap E)/20}{N(E)/20} = \frac{P(R \cap E)}{P(E)},$$

where $N(R \cap E)$ and $N(E)$ are the numbers of outcomes in events $R \cap E$ and E, respectively. ■

This example illustrates a number of common situations. That is, in some random experiments, we are interested only in those outcomes which are elements of a subset B of the sample space S. This means, for our purposes, that the sample space is effectively the subset B. We are now confronted with the problem of defining a probability set function with B as the "new" sample space. That is, for a given event

Table 1.3-1 Tulip combinations

	Early (E)	Late (L)	Totals
Red (R)	5	8	13
Yellow (Y)	3	4	7
Totals	8	12	20