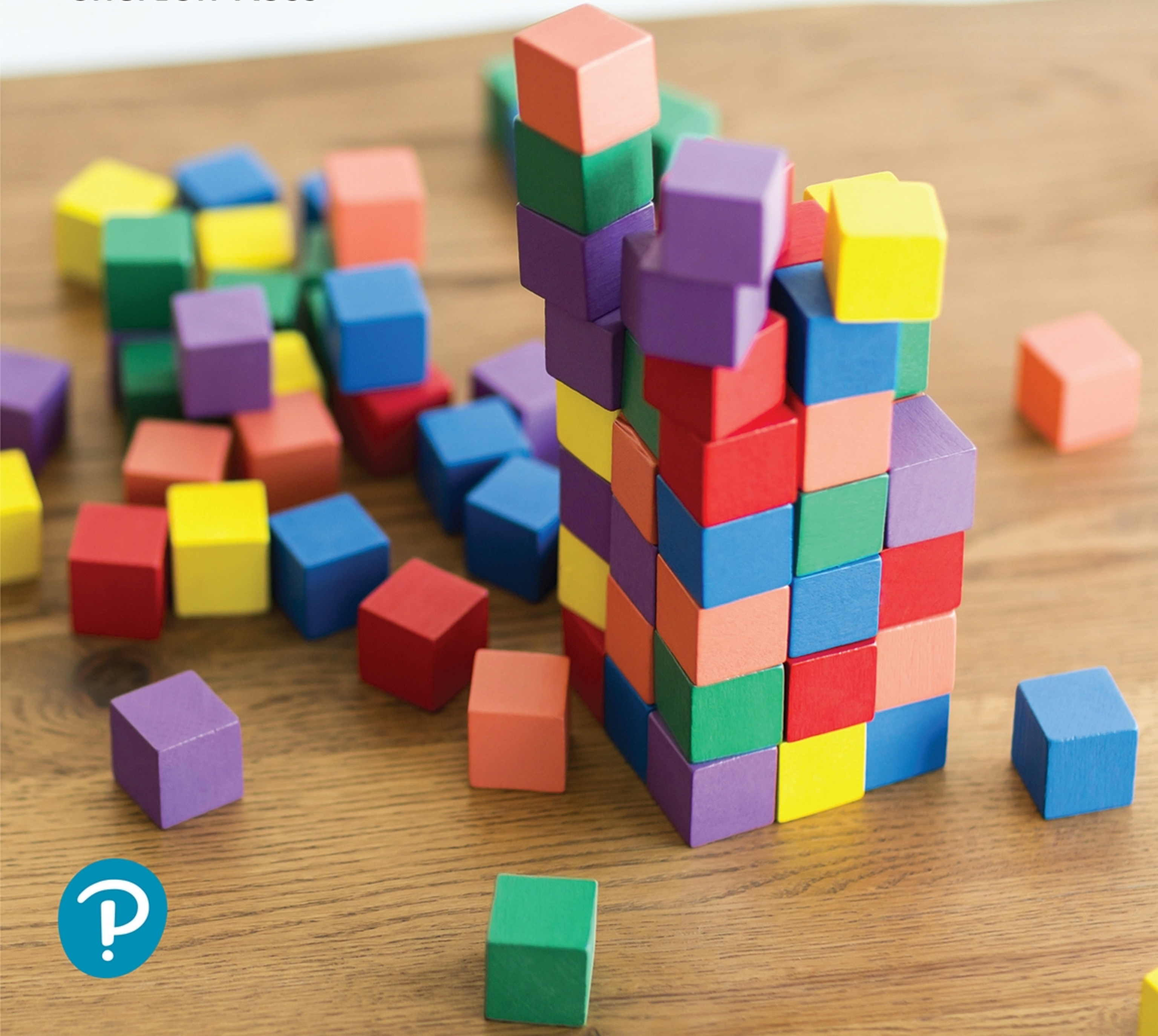
GLOBAL EDITION

A First Course in Probability

TENTH EDITION

Sheldon Ross



A First Course in Probability

Tenth Edition

Global Edition

Sheldon Ross
University of Southern California

Pearson

Director, Portfolio Management: Deirdre Lynch
Courseware Portfolio Manager: Suzanna Bainbridge
Courseware Portfolio Management Assistant: Morgan Danna
Assistant Editors, Global Edition: Tanima Ghosh and Shaoni Mukherjee
Content Producer: Tara Corpuz
Managing Producer: Scott Disanno
Producer: Jon Wooding
Product Marketing Manager: Yvonne Vannatta
Product Marketing Assistant: Jon Bryant
Field Marketing Manager: Evan St. Cyr
Senior Author Support/Technology Specialist: Joe Vetere
Manager, Rights and Permissions: Gina Cheselka
Cover Design: Lumina Datamatics
Manufacturing Buyer: Carol Melville, LSC Communications
Manufacturing Buyer, Global Edition: Kay Holman
Cover Image: sukiyaki/Shutterstock

Pearson Education Limited
KAO Two
KAO Park
Harlow
CM17 9SR
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

ISBN 10: 1-292-26920-0
ISBN 13: 978-1-292-26920-7
ebook ISBN 13: 978-1-292-26923-8

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

For Rebecca

CONTENTS

Preface

"We see that the theory of probability is at bottom only common sense reduced to calculation; it makes us appreciate with exactitude what reasonable minds feel by a sort of instinct, often without being able to account for it.... It is remarkable that this science, which originated in the consideration of games of chance, should have become the most important object of human knowledge.... The most important questions of life are, for the most part, really only problems of probability." So said the famous French mathematician and astronomer (the "Newton of France") Pierre-Simon, Marquis de Laplace. Although many people believe that the famous marquis, who was also one of the great contributors to the development of probability, might have exaggerated somewhat, it is nevertheless true that probability theory has become a tool of fundamental importance to nearly all scientists, engineers, medical practitioners, jurists, and industrialists. In fact, the enlightened individual had learned to ask not "Is it so?" but rather "What is the probability that it is so?"

General Approach and Mathematical Level

This book is intended as an elementary introduction to the theory of probability for students in mathematics, statistics, engineering, and the sciences (including computer science, biology, the social sciences, and management science) who possess the prerequisite knowledge of elementary calculus. It attempts to present not only the mathematics of probability theory, but also, through numerous examples, the many diverse possible applications of this subject.

Content and Course Planning

Chapter 1 presents the basic principles of combinatorial analysis, which are most useful in computing probabilities.

Chapter 2 handles the axioms of probability theory and shows how they can be applied to compute various probabilities of interest.

Chapter 3 deals with the extremely important subjects of conditional probability and independence of events. By a series of examples, we illustrate how conditional probabilities come into play not only when some partial information is available, but also as a tool to enable us to compute probabilities more easily, even when no partial information is present. This extremely important technique of obtaining probabilities by "conditioning" reappears in Chapter 7, where we use it to obtain expectations.

The concept of random variables is introduced in Chapters 4, 5, and 6. Discrete random variables are dealt with in Chapter 4, continuous random variables in Chapter 5, and jointly distributed random variables in Chapter 6. The important concepts of the expected value and the variance of a random variable are introduced in Chapters 4 and 5, and these quantities are then determined for many of the common types of random variables.

Additional properties of the expected value are considered in Chapter 7. Many examples illustrating the usefulness of the result that the expected value of a sum of random variables is equal to the sum of their expected values are presented. Sections on conditional expectation, including its use in prediction, and on moment-generating functions are contained in this chapter. In addition, the final section introduces the multivariate normal distribution and presents a simple proof concerning the joint distribution of the sample mean and sample variance of a sample from a normal distribution.

Chapter 8 presents the major theoretical results of probability theory. In particular, we prove the strong law of large numbers and the central limit theorem. Our proof of the strong law is a relatively simple one that assumes that the random variables have a finite fourth moment, and our proof of the central limit theorem assumes Levy's continuity theorem. This chapter also presents such probability inequalities as Markov's inequality, Chebyshev's inequality, and Chernoff bounds. The final section of Chapter 8 gives a bound on the error involved when a probability concerning a sum of independent Bernoulli random variables is approximated by the corresponding probability of a Poisson random variable having the same expected value.

Chapter 9 presents some additional topics, such as Markov chains, the Poisson process, and an introduction to information and coding theory, and Chapter 10 considers simulation.

As in the previous edition, three sets of exercises are given at the end of each chapter. They are designated as **Problems, Theoretical Exercises**, and **Self-Test Problems and Exercises**. This last set of exercises, for which complete solutions appear in Solutions to Self-Test Problems and Exercises, is designed to help students test their comprehension and study for exams.

Changes for the Tenth Edition

The tenth edition continues the evolution and fine tuning of the text. Aside from a multitude of small changes made to increase the clarity of the text, the new edition includes many new and updated problems, exercises, and text material chosen both for inherent interest and for their use in building student intuition about probability. Illustrative of these goals are Examples 4n of Chapter 3, which deals with computing NCAA basketball tournament win probabilities, and Example 5b of Chapter 4, which introduces the friendship paradox. There is also new material on the Pareto distribution (introduced in Section 5.6.5), on Poisson limit results (in Section 8.5), and on the Lorenz curve (in Section 8.7).

Acknowledgments

I would like to thank the following people who have graciously taken the time to contact me with comments for improving the text: Amir Ardestani, Polytechnic University of Teheran; Joe Blitzstein, Harvard University; Peter Nuesch, University of Lausaunne; Joseph Mitchell, SUNY, Stony Brook; Alan Chambless, actuary; Robert Kriner; Israel David, Ben-Gurion University; T. Lim, George Mason University; Wei Chen, Rutgers; D. Monrad, University of Illinois; W. Rosenberger, George Mason University; E. Ionides, University of Michigan; J. Corvino, Lafayette College;

T. Seppalainen, University of Wisconsin; Jack Goldberg; University of Michigan; Sunil Dhar, New Jersey Institute of Technology; Vladislav Kargin, Stanford University; Marlene Miller; Ahmad Parsian; and Fritz Scholz, University of Washington.

I would also like to especially thank the reviewers of the ninth and tenth editions: Richard Laugesen, University of Illinois; Stacey Hancock, Clark University; Stefan Heinz, University of Wyoming; and Brian Thelen, University of Michigan; Mark Ward, Purdue University. I would like to thank the accuracy checker, Stacey Hancock (Montana State University), for her careful review.

Finally, I would like to thank the following reviewers for their many helpful comments. Reviewers of the tenth edition are marked with an asterisk.

K. B. Athreya, *Iowa State University*
Richard Bass, *University of Connecticut*
Robert Bauer, *University of Illinois at Urbana-Champaign*
Phillip Beckwith, *Michigan Tech*
Arthur Benjamin, *Harvey Mudd College*
Geoffrey Berresford, *Long Island University*
Baidurya Bhattacharya, *University of Delaware*
Howard Bird, *St. Cloud State University*
Shahar Boneh, *Metropolitan State College of Denver*
Jean Cadet, *State University of New York at Stony Brook*
Steven Chiappari, *Santa Clara University*
Nicolas Christou, *University of California, Los Angeles*
James Clay, *University of Arizona at Tucson*
Francis Conlan, *University of Santa Clara*
Justin Corvino, *Lafayette College*
Jay DeVore, *California Polytechnic University, San Luis Obispo*
Scott Emerson, *University of Washington*
Thomas R. Fischer, *Texas A & M University*
Anant Godbole, *Michigan Technical University*
Zakkula Govindarajulu, *University of Kentucky*
Richard Groeneveld, *Iowa State University*
*Stacey Hancock, *Clark University*
Mike Hardy, *Massachusetts Institute of Technology*
Bernard Harris, *University of Wisconsin*
Larry Harris, *University of Kentucky*
David Heath, *Cornell University*
*Stefan Heinz, *University of Wyoming*
Stephen Herschkorn, *Rutgers University*
Julia L. Higle, *University of Arizona*
Mark Huber, *Duke University*
Edward Ionides, *University of Michigan*
Anastasia Ivanova, *University of North Carolina*
Hamid Jafarkhani, *University of California, Irvine*
Chuanshu Ji, *University of North Carolina, Chapel Hill*
Robert Keener, *University of Michigan*
*Richard Laugesen, *University of Illinois*
Fred Leysieffer, *Florida State University*
Thomas Liggett, *University of California, Los Angeles*
Helmut Mayer, *University of Georgia*
Bill McCormick, *University of Georgia*
Ian McKeague, *Florida State University*
R. Miller, *Stanford University*
Ditlev Monrad, *University of Illinois*
Robb J. Muirhead, *University of Michigan*
Joe Naus, *Rutgers University*
Nhu Nguyen, *New Mexico State University*
Ellen O'Brien, *George Mason University*
N. U. Prabhu, *Cornell University*
Kathryn Prewitt, *Arizona State University*
Jim Propp, *University of Wisconsin*
William F. Rosenberger, *George Mason University*
Myra Samuels, *Purdue University*
I. R. Savage, *Yale University*
Art Schwartz, *University of Michigan at Ann Arbor*
Therese Shelton, *Southwestern University*
Malcolm Sherman, *State University of New York at Albany*
Murad Taqqu, *Boston University*
*Brian Thelen, *University of Michigan*
Eli Upfal, *Brown University*
Ed Wheeler, *University of Tennessee*
Allen Webster, *Bradley University*

S. R.
smross@usc.edu

Acknowledgments for the Global Edition

Pearson would like to thank and acknowledge the following people for their contributions to the Global Edition.

Contributors

Boudjemaa Anchouche, *Kuwait University*
Polina Dolmatova, *American University of Central Asia*
Lino Sant, *University of Malta*
Monique Sciortino, *University of Malta*
David Paul Suda, *University of Malta*

Reviewers

Vikas Arora
Olivia Tze Kar Choi, *The University of Hong Kong*
Ajay Jasra, *King Abdullah University of Science and Technology*
Ranjita Pandey, *University of Delhi*
Dharini Pathmanathan, *Universiti Malaya*
Nur Anisah Binti Mohamed A. Rahman, *Universiti Malaya*
Nelson Gomes Rodrigues Antunes, *Universidade do Algarve*
Yuan Wang, *Sheffield Hallam University*
Femin Yalçin, *İzmir Kâtip Çelebi University*

This page intentionally left blank

Chapter 1

Combinatorial Analysis

Contents

1.1 Introduction

Here is a typical problem of interest involving probability: A communication system is to consist of n seemingly identical antennas that are to be lined up in a linear order. The resulting system will then be able to receive all incoming signals—and will be called *functional*—as long as no two consecutive antennas are defective. If it turns out that exactly m of the n antennas are defective, what is the probability that the resulting system will be functional? For instance, in the special case where $n = 4$ and $m = 2$, there are 6 possible system configurations, namely,

$$\begin{array}{cccc} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{array}$$

where 1 means that the antenna is working and 0 that it is defective. Because the resulting system will be functional in the first 3 arrangements and not functional in the remaining 3, it seems reasonable to take $\frac{3}{6} = \frac{1}{2}$ as the desired probability. In the case of general n and m, we could compute the probability that the system is functional in a similar fashion. That is, we could count the number of configurations that result in the system's being functional and then divide by the total number of all possible configurations.

From the preceding discussion, we see that it would be useful to have an effective method for counting the number of ways that things can occur. In fact, many problems in probability theory can be solved simply by counting the number of different ways that a certain event can occur. The mathematical theory of counting is formally known as *combinatorial analysis*.

1.2 The Basic Principle of Counting

The basic principle of counting will be fundamental to all our work. Loosely put, it states that if one experiment can result in any of m possible outcomes and if another experiment can result in any of n possible outcomes, then there are mn possible outcomes of the two experiments.

The basic principle of counting

Suppose that two experiments are to be performed. Then if experiment 1 can result in any one of m possible outcomes and if, for each outcome of experiment 1, there are n possible outcomes of experiment 2, then together there are mn possible outcomes of the two experiments.

Proof of the Basic Principle: The basic principle may be proven by enumerating all the possible outcomes of the two experiments; that is,

$$
\begin{array}{l}
(1,1),\ (1,2),\ \ldots,\ (1,n) \\
(2,1),\ (2,2),\ \ldots,\ (2,n) \\
\quad\vdots \\
(m,1),\ (m,2),\ \ldots,\ (m,n)
\end{array}
$$

where we say that the outcome is (i, j) if experiment 1 results in its ith possible outcome and experiment 2 then results in its jth possible outcome. Hence, the set of possible outcomes consists of m rows, each containing n elements. This proves the result.

Example 2a

A small community consists of 10 females, each of whom has 3 children. If one female and one of her children are to be chosen as parent and child of the year, how many different choices are possible?

Solution By regarding the choice of the female as the outcome of the first experiment and the subsequent choice of one of her children as the outcome of the second experiment, we see from the basic principle that there are $10 \times 3 = 30$ possible choices. ■

When there are more than two experiments to be performed, the basic principle can be generalized.

The generalized basic principle of counting

If r experiments that are to be performed are such that the first one may result in any of n_1 possible outcomes; and if, for each of these n_1 possible outcomes, there are n_2 possible outcomes of the second experiment; and if, for each of the possible outcomes of the first two experiments, there are n_3 possible outcomes of the third experiment; and if ..., then there is a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes of the r experiments.

Example 2b A college planning committee consists of 3 freshmen, 4 sophomores, 5 juniors, and 2 seniors. A subcommittee of 4, consisting of 1 person from each class, is to be chosen. How many different subcommittees are possible?

Solution We may regard the choice of a subcommittee as the combined outcome of the four separate experiments of choosing a single representative from each of the classes. It then follows from the generalized version of the basic principle that there are $3 \times 4 \times 5 \times 2 = 120$ possible subcommittees. ■

Example 2c How many different 7-place license plates are possible if the first 3 places are to be occupied by letters and the final 4 by numbers?

Solution By the generalized version of the basic principle, the answer is $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 175{,}760{,}000$. ■

Example 2d How many functions defined on n points are possible if each functional value is either 0 or 1?

Solution Let the points be $1, 2, \ldots, n$. Since $f(i)$ must be either 0 or 1 for each $i = 1, 2, \ldots, n$, it follows that there are 2^n possible functions. ■

Example 2e In Example 2c, how many license plates would be possible if repetition among letters or numbers were prohibited?

Solution In this case, there would be $26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7 = 78{,}624{,}000$ possible license plates. ■

1.3 Permutations

How many different ordered arrangements of the letters a, b, and c are possible? By direct enumeration we see that there are 6, namely, *abc, acb, bac, bca, cab*, and *cba*. Each arrangement is known as a *permutation*. Thus, there are 6 possible permutations of a set of 3 objects. This result could also have been obtained from the basic principle, since the first object in the permutation can be any of the 3, the second object in the permutation can then be chosen from any of the remaining 2, and the third object in the permutation is then the remaining 1. Thus, there are $3 \cdot 2 \cdot 1 = 6$ possible permutations.

Suppose now that we have n objects. Reasoning similar to that we have just used for the 3 letters then shows that there are

$$n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1 = n!$$

different permutations of the n objects.

Whereas $n!$ (read as "n factorial") is defined to equal $1 \cdot 2 \cdots n$ when n is a positive integer, it is convenient to define 0! to equal 1.

Example 3a How many different batting orders are possible for a baseball team consisting of 9 players?

Solution There are $9! = 362{,}880$ possible batting orders. ■

Example 3b

A class in probability theory consists of 6 sophomores and 4 juniors. An examination is given, and the students are ranked according to their performance. Assume that no two students obtain the same score.

(a) How many different rankings are possible?

(b) If the sophomores are ranked just among themselves and the juniors just among themselves, how many different rankings are possible?

Solution (a) Because each ranking corresponds to a particular ordered arrangement of the 10 people, the answer to this part is $10! = 3{,}628{,}800$.

(b) Since there are $6!$ possible rankings of the sophomores among themselves and $4!$ possible rankings of the juniors among themselves, it follows from the basic principle that there are $(6!)(4!) = (720)(24) = 17{,}280$ possible rankings in this case. ■

Example 3c

Ms. Jones has 10 books that she is going to put on her bookshelf. Of these, 4 are mathematics books, 3 are chemistry books, 2 are history books, and 1 is a language book. Ms. Jones wants to arrange her books so that all the books dealing with the same subject are together on the shelf. How many different arrangements are possible?

Solution There are $4!\ 3!\ 2!\ 1!$ arrangements such that the mathematics books are first in line, then the chemistry books, then the history books, and then the language book. Similarly, for each possible ordering of the subjects, there are $4!\ 3!\ 2!\ 1!$ possible arrangements. Hence, as there are $4!$ possible orderings of the subjects, the desired answer is $4!\ 4!\ 3!\ 2!\ 1! = 6912$. ■

We shall now determine the number of permutations of a set of n objects when certain of the objects are indistinguishable from one another. To set this situation straight in our minds, consider the following example.

Example 3d

How many different letter arrangements can be formed from the letters *PEPPER*?

Solution We first note that there are $6!$ permutations of the letters $P_1E_1P_2P_3E_2R$ when the $3P$'s and the $2E$'s are distinguished from one another. However, consider any one of these permutations—for instance, $P_1P_2E_1P_3E_2R$. If we now permute the P's among themselves and the E's among themselves, then the resultant arrangement would still be of the form $PPEPER$. That is, all $3!\ 2!$ permutations

$$
\begin{array}{ll}
P_1P_2E_1P_3E_2R & P_1P_2E_2P_3E_1R \\
P_1P_3E_1P_2E_2R & P_1P_3E_2P_2E_1R \\
P_2P_1E_1P_3E_2R & P_2P_1E_2P_3E_1R \\
P_2P_3E_1P_1E_2R & P_2P_3E_2P_1E_1R \\
P_3P_1E_1P_2E_2R & P_3P_1E_2P_2E_1R \\
P_3P_2E_1P_1E_2R & P_3P_2E_2P_1E_1R
\end{array}
$$

are of the form $PPEPER$. Hence, there are $6!/(3!\ 2!) = 60$ possible letter arrangements of the letters *PEPPER*. ■

In general, the same reasoning as that used in Example 3d shows that there are

$$\frac{n!}{n_1!\, n_2!\, \cdots\, n_r!}$$

different permutations of n objects, of which n_1 are alike, n_2 are alike, $\ldots, n_r$ are alike.

Example 3e A chess tournament has 10 competitors, of which 4 are Russian, 3 are from the United States, 2 are from Great Britain, and 1 is from Brazil. If the tournament result lists just the nationalities of the players in the order in which they placed, how many outcomes are possible?

Solution There are

$$\frac{10!}{4!\,3!\,2!\,1!} = 12{,}600$$

possible outcomes. ■

Example 3f How many different signals, each consisting of 9 flags hung in a line, can be made from a set of 4 white flags, 3 red flags, and 2 blue flags if all flags of the same color are identical?

Solution There are

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

different signals. ■

1.4 Combinations

We are often interested in determining the number of different groups of r objects that could be formed from a total of n objects. For instance, how many different groups of 3 could be selected from the 5 items A, B, C, D, and E? To answer this question, reason as follows: Since there are 5 ways to select the initial item, 4 ways to then select the next item, and 3 ways to select the final item, there are thus $5 \cdot 4 \cdot 3$ ways of selecting the group of 3 when the order in which the items are selected is relevant. However, since every group of 3—say, the group consisting of items A, B, and C—will be counted 6 times (that is, all of the permutations ABC, ACB, BAC, BCA, CAB, and CBA will be counted when the order of selection is relevant), it follows that the total number of groups that can be formed is

$$\frac{5 \cdot 4 \cdot 3}{3 \cdot 2 \cdot 1} = 10$$

In general, as $n(n-1)\cdots(n-r+1)$ represents the number of different ways that a group of r items could be selected from n items when the order of selection is relevant, and as each group of r items will be counted $r!$ times in this count, it follows that the number of different groups of r items that could be formed from a set of n items is

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

Notation and terminology

We define $\binom{n}{r}$, for $r \leq n$, by

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

and say that $\binom{n}{r}$ (read as "n choose r") represents the number of possible combinations of n objects taken r at a time.

Thus, $\binom{n}{r}$ represents the number of different groups of size r that could be selected from a set of n objects when the order of selection is not considered relevant.

Equivalently, $\binom{n}{r}$ is the number of subsets of size r that can be chosen from a set of size n. Using that $0! = 1$, note that $\binom{n}{n} = \binom{n}{0} = \frac{n!}{0!n!} = 1$, which is consistent with the preceding interpretation because in a set of size n there is exactly 1 subset of size n (namely, the entire set), and exactly one subset of size 0 (namely the empty set). A useful convention is to define $\binom{n}{r}$ equal to 0 when either $r > n$ or $r < 0$.

Example 4a

A committee of 3 is to be formed from a group of 20 people. How many different committees are possible?

Solution There are $\binom{20}{3} = \frac{20 \cdot 19 \cdot 18}{3 \cdot 2 \cdot 1} = 1140$ possible committees. ■

Example 4b

From a group of 5 women and 7 men, how many different committees consisting of 2 women and 3 men can be formed? What if 2 of the men are feuding and refuse to serve on the committee together?

Solution As there are $\binom{5}{2}$ possible groups of 2 women, and $\binom{7}{3}$ possible groups of 3 men, it follows from the basic principle that there are $\binom{5}{2}\binom{7}{3} = \frac{5 \cdot 4}{2 \cdot 1} \cdot \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 350$ possible committees consisting of 2 women and 3 men.

Now suppose that 2 of the men refuse to serve together. Because a total of $\binom{2}{2}\binom{5}{1} = 5$ out of the $\binom{7}{3} = 35$ possible groups of 3 men contain both of the feuding men, it follows that there are $35 - 5 = 30$ groups that do not contain both of the feuding men. Because there are still $\binom{5}{2} = 10$ ways to choose the 2 women, there are $30 \cdot 10 = 300$ possible committees in this case. ■

Example 4c

Consider a set of n antennas of which m are defective and $n - m$ are functional and assume that all of the defectives and all of the functionals are considered indistinguishable. How many linear orderings are there in which no two defectives are consecutive?

Solution Imagine that the $n - m$ functional antennas are lined up among themselves. Now, if no two defectives are to be consecutive, then the spaces between the

$\wedge 1 \wedge 1 \wedge 1 \ldots \wedge 1 \wedge 1 \wedge$

1 = functional

$\wedge$ = place for at most one defective

Figure 1.1 No consecutive defectives.

functional antennas must each contain at most one defective antenna. That is, in the $n - m + 1$ possible positions—represented in Figure 1.1 by carets—between the $n - m$ functional antennas, we must select m of these in which to put the defective antennas. Hence, there are $\binom{n-m+1}{m}$ possible orderings in which there is at least one functional antenna between any two defective ones. ■

A useful combinatorial identity, known as *Pascal's identity*, is

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \qquad 1 \le r \le n \tag{4.1}$$

Equation (4.1) may be proved analytically or by the following combinatorial argument: Consider a group of n objects, and fix attention on some particular one of these objects—call it object 1. Now, there are $\binom{n-1}{r-1}$ groups of size r that contain object 1 (since each such group is formed by selecting $r - 1$ from the remaining $n - 1$ objects). Also, there are $\binom{n-1}{r}$ groups of size r that do not contain object 1. As there is a total of $\binom{n}{r}$ groups of size r, Equation (4.1) follows.

The values $\binom{n}{r}$ are often referred to as *binomial coefficients* because of their prominence in the binomial theorem.

The binomial theorem

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \tag{4.2}$$

We shall present two proofs of the binomial theorem. The first is a proof by mathematical induction, and the second is a proof based on combinatorial considerations.

Proof of the Binomial Theorem by Induction: When $n = 1$, Equation (4.2) reduces to

$$x + y = \binom{1}{0} x^0 y^1 + \binom{1}{1} x^1 y^0 = y + x$$

Assume Equation (4.2) for $n - 1$. Now,

$$\begin{aligned}
(x + y)^n &= (x + y)(x + y)^{n-1} \\
&= (x + y)\sum_{k=0}^{n-1}\binom{n-1}{k}x^k y^{n-1-k} \\
&= \sum_{k=0}^{n-1}\binom{n-1}{k}x^{k+1}y^{n-1-k} + \sum_{k=0}^{n-1}\binom{n-1}{k}x^k y^{n-k}
\end{aligned}$$

Letting $i = k + 1$ in the first sum and $i = k$ in the second sum, we find that

$$\begin{aligned}
(x + y)^n &= \sum_{i=1}^{n}\binom{n-1}{i-1}x^i y^{n-i} + \sum_{i=0}^{n-1}\binom{n-1}{i}x^i y^{n-i} \\
&= \sum_{i=1}^{n-1}\binom{n-1}{i-1}x^i y^{n-i} + x^n + y^n + \sum_{i=1}^{n-1}\binom{n-1}{i}x^i y^{n-i} \\
&= x^n + \sum_{i=1}^{n-1}\left[\binom{n-1}{i-1} + \binom{n-1}{i}\right]x^i y^{n-i} + y^n \\
&= x^n + \sum_{i=1}^{n-1}\binom{n}{i}x^i y^{n-i} + y^n \\
&= \sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i}
\end{aligned}$$

where the next-to-last equality follows by Equation (4.1). By induction, the theorem is now proved.

Combinatorial Proof of the Binomial Theorem: Consider the product

$$(x_1 + y_1)(x_2 + y_2)\cdots(x_n + y_n)$$

Its expansion consists of the sum of 2^n terms, each term being the product of n factors. Furthermore, each of the 2^n terms in the sum will contain as a factor either x_i or y_i for each $i = 1, 2, \ldots, n$. For example,

$$(x_1 + y_1)(x_2 + y_2) = x_1x_2 + x_1y_2 + y_1x_2 + y_1y_2$$

Now, how many of the 2^n terms in the sum will have k of the x_i's and $(n - k)$ of the y_i's as factors? As each term consisting of k of the x_i's and $(n - k)$ of the y_i's corresponds to a choice of a group of k from the n values $x_1, x_2, \ldots, x_n$, there are $\binom{n}{k}$ such terms. Thus, letting $x_i = x, y_i = y, i = 1, \ldots, n$, we see that

$$(x + y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$$

Example 4d Expand $(x + y)^3$.

Solution

$$(x + y)^3 = \binom{3}{0}x^0y^3 + \binom{3}{1}x^1y^2 + \binom{3}{2}x^2y^1 + \binom{3}{3}x^3y^0$$
$$= y^3 + 3xy^2 + 3x^2y + x^3$$ ■

Example 4e How many subsets are there of a set consisting of n elements?

Solution Since there are $\binom{n}{k}$ subsets of size k, the desired answer is

$$\sum_{k=0}^{n}\binom{n}{k} = (1 + 1)^n = 2^n$$

This result could also have been obtained by assigning either the number 0 or the number 1 to each element in the set. To each assignment of numbers, there corresponds, in a one-to-one fashion, a subset, namely, that subset consisting of all elements that were assigned the value 1. As there are 2^n possible assignments, the result follows.

Note that we have included the set consisting of 0 elements (that is, the null set) as a subset of the original set. Hence, the number of subsets that contain at least 1 element is $2^n - 1$. ■

1.5 Multinomial Coefficients

In this section, we consider the following problem: A set of n distinct items is to be divided into r distinct groups of respective sizes $n_1, n_2, \ldots, n_r$, where $\sum_{i=1}^{r} n_i = n$. How many different divisions are possible? To answer this question, we note that there are $\binom{n}{n_1}$ possible choices for the first group; for each choice of the first group, there are $\binom{n - n_1}{n_2}$ possible choices for the second group; for each choice of the first two groups, there are $\binom{n - n_1 - n_2}{n_3}$ possible choices for the third group; and so on. It then follows from the generalized version of the basic counting principle that there are

$$\binom{n}{n_1}\binom{n - n_1}{n_2}\cdots\binom{n - n_1 - n_2 - \cdots - n_{r-1}}{n_r}$$
$$= \frac{n!}{(n - n_1)!\,n_1!}\,\frac{(n - n_1)!}{(n - n_1 - n_2)!\,n_2!}\cdots\frac{(n - n_1 - n_2 - \cdots - n_{r-1})!}{0!\,n_r!}$$
$$= \frac{n!}{n_1!\,n_2!\cdots n_r!}$$

possible divisions.

Another way to see this result is to consider the n values $1, 1, \ldots, 1, 2, \ldots, 2, \ldots, r, \ldots, r$, where i appears n_i times, for $i = 1, \ldots, r$. Every permutation of these values corresponds to a division of the n items into the r groups in the following manner: Let the permutation $i_1, i_2, \ldots, i_n$ correspond to assigning item 1 to group i_1, item 2 to group i_2, and so on. For instance, if $n = 8$ and if $n_1 = 4$, $n_2 = 3$, and $n_3 = 1$, then the permutation $1, 1, 2, 3, 2, 1, 2, 1$ corresponds to assigning items $1, 2, 6, 8$ to the first group, items $3, 5, 7$ to the second group, and item 4 to the third group. Because every permutation yields a division of the items and every possible division results from some permutation, it follows that the number of divisions of n items into r distinct groups of sizes $n_1, n_2, \ldots, n_r$ is the same as the number of permutations of n items of which n_1 are alike, and n_2 are alike, ..., and n_r are alike, which was shown in Section 1.3 to equal $\dfrac{n!}{n_1!n_2!\cdots n_r!}$.

Notation

If $n_1 + n_2 + \cdots + n_r = n$, we define $\dbinom{n}{n_1, n_2, \ldots, n_r}$ by

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

Thus, $\dbinom{n}{n_1, n_2, \ldots, n_r}$ represents the number of possible divisions of n distinct objects into r distinct groups of respective sizes $n_1, n_2, \ldots, n_r$.

Example 5a

A police department in a small city consists of 10 officers. If the department policy is to have 5 of the officers patrolling the streets, 2 of the officers working full time at the station, and 3 of the officers on reserve at the station, how many different divisions of the 10 officers into the 3 groups are possible?

Solution There are $\dfrac{10!}{5!\,2!\,3!} = 2520$ possible divisions. ■

Example 5b

Ten children are to be divided into an A team and a B team of 5 each. The A team will play in one league and the B team in another. How many different divisions are possible?

Solution There are $\dfrac{10!}{5!\,5!} = 252$ possible divisions. ■

Example 5c

In order to play a game of basketball, 10 children at a playground divide themselves into two teams of 5 each. How many different divisions are possible?

Solution Note that this example is different from Example 5b because now the order of the two teams is irrelevant. That is, there is no A or B team, but just a division consisting of 2 groups of 5 each. Hence, the desired answer is

$$\frac{10!/(5!\,5!)}{2!} = 126$$ ■

The proof of the following theorem, which generalizes the binomial theorem, is left as an exercise.

The multinomial theorem

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{\substack{(n_1,\ldots,n_r):\\ n_1 + \cdots + n_r = n}} \binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$$

That is, the sum is over all nonnegative integer-valued vectors $(n_1, n_2, \ldots, n_r)$ such that $n_1 + n_2 + \cdots + n_r = n$.

The numbers $\binom{n}{n_1, n_2, \ldots, n_r}$ are known as *multinomial coefficients*.

Example 5d In the first round of a knockout tournament involving $n = 2^m$ players, the n players are divided into $n/2$ pairs, with each of these pairs then playing a game. The losers of the games are eliminated while the winners go on to the next round, where the process is repeated until only a single player remains. Suppose we have a knockout tournament of 8 players.

(a) How many possible outcomes are there for the initial round? (For instance, one outcome is that 1 beats 2, 3 beats 4, 5 beats 6, and 7 beats 8.)

(b) How many outcomes of the tournament are possible, where an outcome gives complete information for all rounds?

Solution One way to determine the number of possible outcomes for the initial round is to first determine the number of possible pairings for that round. To do so, note that the number of ways to divide the 8 players into a *first* pair, a *second* pair, a *third* pair, and a *fourth* pair is $\binom{8}{2,2,2,2} = \frac{8!}{2^4}$. Thus, the number of possible pairings when there is no ordering of the 4 pairs is $\frac{8!}{2^4\,4!}$. For each such pairing, there are 2 possible choices from each pair as to the winner of that game, showing that there are $\frac{8!2^4}{2^4\,4!} = \frac{8!}{4!}$ possible results of round 1. [Another way to see this is to note that there are $\binom{8}{4}$ possible choices of the 4 winners and, for each such choice, there are 4! ways to pair the 4 winners with the 4 losers, showing that there are $4!\binom{8}{4} = \frac{8!}{4!}$ possible results for the first round.]

Similarly, for each result of round 1, there are $\frac{4!}{2!}$ possible outcomes of round 2, and for each of the outcomes of the first two rounds, there are $\frac{2!}{1!}$ possible outcomes of round 3. Consequently, by the generalized basic principle of counting, there are $\frac{8!}{4!}\frac{4!}{2!}\frac{2!}{1!} = 8!$ possible outcomes of the tournament. Indeed, the same argument can be used to show that a knockout tournament of $n = 2^m$ players has $n!$ possible outcomes.

Knowing the preceding result, it is not difficult to come up with a more direct argument by showing that there is a one-to-one correspondence between the set of possible tournament results and the set of permutations of $1, \ldots, n$. To obtain such a correspondence, rank the players as follows for any tournament result: Give the tournament winner rank 1, and give the final-round loser rank 2. For the two players who lost in the next-to-last round, give rank 3 to the one who lost to the player ranked 1 and give rank 4 to the one who lost to the player ranked 2. For the four players who lost in the second-to-last round, give rank 5 to the one who lost to player ranked 1, rank 6 to the one who lost to the player ranked 2, rank 7 to the one who lost to the player ranked 3, and rank 8 to the one who lost to the player ranked 4. Continuing on in this manner gives a rank to each player. (A more succinct description is to give the winner of the tournament rank 1 and let the rank of a player who lost in a round having 2^k matches be 2^k plus the rank of the player who beat him, for $k = 0, \ldots, m - 1$.) In this manner, the result of the tournament can be represented by a permutation $i_1, i_2, \ldots, i_n$, where i_j is the player who was given rank j. Because different tournament results give rise to different permutations, and because there is a tournament result for each permutation, it follows that there are the same number of possible tournament results as there are permutations of $1, \ldots, n$. ■

Example 5e

$$\begin{aligned}
(x_1 + x_2 + x_3)^2 &= \binom{2}{2,0,0} x_1^2 x_2^0 x_3^0 + \binom{2}{0,2,0} x_1^0 x_2^2 x_3^0 \\
&\quad + \binom{2}{0,0,2} x_1^0 x_2^0 x_3^2 + \binom{2}{1,1,0} x_1^1 x_2^1 x_3^0 \\
&\quad + \binom{2}{1,0,1} x_1^1 x_2^0 x_3^1 + \binom{2}{0,1,1} x_1^0 x_2^1 x_3^1 \\
&= x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3
\end{aligned}$$
■

*1.6 The Number of Integer Solutions of Equations

An individual has gone fishing at Lake Ticonderoga, which contains four types of fish: lake trout, catfish, bass, and bluefish. If we take the result of the fishing trip to be the numbers of each type of fish caught, let us determine the number of possible outcomes when a total of 10 fish are caught. To do so, note that we can denote the outcome of the fishing trip by the vector (x_1, x_2, x_3, x_4) where x_1 is the number of trout that are caught, x_2 is the number of catfish, x_3 is the number of bass, and x_4 is the number of bluefish. Thus, the number of possible outcomes when a total of 10 fish are caught is the number of nonnegative integer vectors (x_1, x_2, x_3, x_4) that sum to 10.

More generally, if we supposed there were r types of fish and that a total of n were caught, then the number of possible outcomes would be the number of nonnegative integer-valued vectors $x_1, \ldots, x_r$ such that

$$x_1 + x_2 + \ldots + x_r = n \tag{6.1}$$

To compute this number, let us start by considering the number of positive integer-valued vectors $x_1, \ldots, x_r$ that satisfy the preceding. To determine this number, suppose that we have n consecutive zeroes lined up in a row:

$$0\,0\,0 \ldots 0\,0$$

* Asterisks denote material that is optional.

$$0 \wedge 0 \wedge 0 \wedge \ldots \wedge 0 \wedge 0$$

n objects 0

Choose $r - 1$ of the spaces $\wedge$.

Figure 1.2 Number of positive solutions.

Note that any selection of $r - 1$ of the $n - 1$ spaces between adjacent zeroes (see Figure 1.2) corresponds to a positive solution of (6.1) by letting x_1 be the number of zeroes before the first chosen space, x_2 be the number of zeroes between the first and second chosen space, ..., and x_n being the number of zeroes following the last chosen space.
For instance, if we have $n = 8$ and $r = 3$, then (with the choices represented by dots) the choice

$$0 . 0\,0\,0\,0 . 0\,0\,0$$

corresponds to the solution $x_1 = 1, x_2 = 4, x_3 = 3$. As positive solutions of (6.1) correspond, in a one-to-one fashion, to choices of $r - 1$ of the adjacent spaces, it follows that the number of differerent positive solutions is equal to the number of different selections of $r - 1$ of the $n - 1$ adjacent spaces. Consequently, we have the following proposition.

Proposition 6.1

There are $\binom{n-1}{r-1}$ distinct positive integer-valued vectors $(x_1, x_2, \ldots, x_r)$ satisfying the equation

$$x_1 + x_2 + \cdots + x_r = n, \quad x_i > 0, \quad i = 1, \ldots, r$$

To obtain the number of nonnegative (as opposed to positive) solutions, note that the number of nonnegative solutions of $x_1 + x_2 + \cdots + x_r = n$ is the same as the number of positive solutions of $y_1 + \cdots + y_r = n + r$ (seen by letting $y_i = x_i + 1$, $i = 1, \ldots, r$). Hence, from Proposition 6.1, we obtain the following proposition.

Proposition 6.2

There are $\binom{n+r-1}{r-1}$ distinct nonnegative integer-valued vectors $(x_1, x_2, \ldots, x_r)$ satisfying the equation

$$x_1 + x_2 + \cdots + x_r = n$$

Thus, using Proposition 6.2, we see that there are $\binom{13}{3} = 286$ possible outcomes when a total of 10 Lake Ticonderoga fish are caught.

Example 6a

How many distinct nonnegative integer-valued solutions of $x_1 + x_2 = 3$ are possible?

Solution There are $\binom{3 + 2 - 1}{2 - 1} = 4$ such solutions: (0, 3), (1, 2), (2, 1), (3, 0). ■

Example 6b

An investor has \$20,000 to invest among 4 possible investments. Each investment must be in units of \$1000. If the total \$20,000 is to be invested, how many different investment strategies are possible? What if not all the money needs to be invested?

Solution If we let x_i, $i = 1, 2, 3, 4$, denote the number of thousands invested in investment i, then, when all is to be invested, x_1, x_2, x_3, x_4 are integers satisfying the equation

$$x_1 + x_2 + x_3 + x_4 = 20 \quad x_i \geq 0$$

Hence, by Proposition 6.2, there are $\binom{23}{3} = 1771$ possible investment strategies. If not all of the money needs to be invested, then if we let x_5 denote the amount kept in reserve, a strategy is a nonnegative integer-valued vector $(x_1, x_2, x_3, x_4, x_5)$ satisfying the equation

$$x_1 + x_2 + x_3 + x_4 + x_5 = 20$$

Hence, by Proposition 6.2, there are now $\binom{24}{4} = 10{,}626$ possible strategies. ■

Example 6c

How many terms are there in the multinomial expansion of $(x_1 + x_2 + \cdots + x_r)^n$?

Solution

$$(x_1 + x_2 + \cdots + x_r)^n = \sum \binom{n}{n_1, \ldots, n_r} x_1^{n_1} \cdots x_r^{n_r}$$

where the sum is over all nonnegative integer-valued $(n_1, \ldots, n_r)$ such that $n_1 + \cdots + n_r = n$. Hence, by Proposition 6.2, there are $\binom{n + r - 1}{r - 1}$ such terms. ■

Example 6d

Let us consider again Example 4c, in which we have a set of n items, of which m are (indistinguishable and) defective and the remaining $n - m$ are (also indistinguishable and) functional. Our objective is to determine the number of linear orderings in which no two defectives are next to each other. To determine this number, let us imagine that the defective items are lined up among themselves and the functional ones are now to be put in position. Let us denote x_1 as the number of functional items to the left of the first defective, x_2 as the number of functional items between the first two defectives, and so on. That is, schematically, we have

$$x_1 \, 0 \, x_2 \, 0 \cdots x_m \, 0 \, x_{m+1}$$

Now, there will be at least one functional item between any pair of defectives as long as $x_i > 0$, $i = 2, \ldots, m$. Hence, the number of outcomes satisfying the condition is the number of vectors $x_1, \ldots, x_{m+1}$ that satisfy the equation

$$x_1 + \cdots + x_{m+1} = n - m, \quad x_1 \geq 0, \; x_{m+1} \geq 0, \; x_i > 0, \; i = 2, \ldots, m$$

But, on letting $y_1 = x_1 + 1, y_i = x_i,\ i = 2, \ldots, m,\ y_{m+1} = x_{m+1} + 1$, we see that this number is equal to the number of positive vectors $(y_1, \ldots, y_{m+1})$ that satisfy the equation

$$y_1 + y_2 + \cdots + y_{m+1} = n - m + 2$$

Hence, by Proposition 6.1, there are $\binom{n - m + 1}{m}$ such outcomes, in agreement with the results of Example 4c.

Suppose now that we are interested in the number of outcomes in which each pair of defective items is separated by at least 2 functional items. By the same reasoning as that applied previously, this would equal the number of vectors satisfying the equation

$$x_1 + \cdots + x_{m+1} = n - m, \quad x_1 \geq 0,\ x_{m+1} \geq 0,\ x_i \geq 2,\ i = 2, \ldots, m$$

Upon letting $y_1 = x_1 + 1, y_i = x_i - 1, i = 2, \ldots, m, y_{m+1} = x_{m+1} + 1$, we see that this is the same as the number of positive solutions of the equation

$$y_1 + \cdots + y_{m+1} = n - 2m + 3$$

Hence, from Proposition 6.1, there are $\binom{n - 2m + 2}{m}$ such outcomes. ■

Summary

The basic principle of counting states that if an experiment consisting of two phases is such that there are n possible outcomes of phase 1 and, for each of these n outcomes, there are m possible outcomes of phase 2, then there are nm possible outcomes of the experiment.

There are $n! = n(n - 1)\cdots 3 \cdot 2 \cdot 1$ possible linear orderings of n items. The quantity 0! is defined to equal 1.

Let

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

when $0 \leq i \leq n$, and let it equal 0 otherwise. This quantity represents the number of different subgroups of size i that can be chosen from a set of size n. It is often called a *binomial coefficient* because of its prominence in the binomial theorem, which states that

$$(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

For nonnegative integers $n_1, \ldots, n_r$ summing to n,

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

is the number of divisions of n items into r distinct nonoverlapping subgroups of sizes $n_1, n_2 \ldots, n_r$. These quantities are called *multinomial coefficients*.

Problems

1.1. (a) How many different 7-place license plates are possible if the first 2 places are for letters and the other 5 for numbers?

(b) Repeat part (a) under the assumption that no letter or number can be repeated in a single license plate.

1.2. How many outcome sequences are possible when a die is rolled four times, where we say, for instance, that the outcome is 3, 4, 3, 1 if the first roll landed on 3, the second on 4, the third on 3, and the fourth on 1?

1.3. Ten employees of a company are to be assigned to 10 different managerial posts, one to each post. In how many ways can these posts be filled?

1.4. John, Jim, Jay, and Jack have formed a band consisting of 4 instruments. If each of the boys can play all 4 instruments, how many different arrangements are possible? What if John and Jim can play all 4 instruments, but Jay and Jack can each play only piano and drums?

1.5. A safe can be opened by inserting a code consisting of three digits between 0 and 9. How many codes are possible? How many codes are possible with no digit repeated? How many codes starting with a 1 are possible?

1.6. A well-known nursery rhyme starts as follows:

"As I was going to St. Ives
I met a man with 7 wives.
Each wife had 7 sacks.
Each sack had 7 cats.
Each cat had 7 kittens..."
How many kittens did the traveler meet?

1.7. (a) In how many ways can 3 boys and 3 girls sit in a row?

(b) In how many ways can 3 boys and 3 girls sit in a row if the boys and the girls are each to sit together?

(c) In how many ways if only the boys must sit together?

(d) In how many ways if no two people of the same sex are allowed to sit together?

1.8. When all letters are used, how many different letter arrangements can be made from the letters

(a) Partying?

(b) Dancing?

(c) Acting?

(d) Singing?

1.9. A box contains 13 balls, of which 4 are yellow, 4 are green, 3 are red, and 2 are blue. Find the number of ways in which these balls can be arranged in a line.

1.10. In how many ways can 8 people be seated in a row if

(a) there are no restrictions on the seating arrangement?

(b) persons A and B must sit next to each other?

(c) there are 4 men and 4 women and no 2 men or 2 women can sit next to each other?

(d) there are 5 men and they must sit next to one another?

(e) there are 4 opposite-sexed couples and each couple must sit together?

1.11. In how many ways can 3 novels, 2 mathematics books, and 1 chemistry book be arranged on a bookshelf if

(a) the books can be arranged in any order?

(b) the mathematics books must be together and the novels must be together?

(c) the novels must be together, but the other books can be arranged in any order?

1.12. How many 3 digit numbers xyz, with x, y, z all ranging from 0 to 9 have at least 2 of their digits equal. How many have exactly 2 equal digits.

1.13. How many different letter configurations of length 4 or 5 can be formed using the letters of the word A C H I E V E?

1.14. Five separate awards (best scholarship, best leadership qualities, and so on) are to be presented to selected students from a class of 30. How many different outcomes are possible if

(a) a student can receive any number of awards?

(b) each student can receive at most 1 award?

1.15. Consider a group of 20 people. If everyone shakes hands with everyone else, how many handshakes take place?

1.16. How many distinct triangles can be drawn by joining any 8 dots on a piece of paper? Note that the dots are in such a way that no 3 of them form a straight line.

1.17. A dance class consists of 22 students, of which 10 are women and 12 are men. If 5 men and 5 women are to be chosen and then paired off, how many results are possible?

1.18. A team consisting of 5 players is to be chosen from a class of 12 boys and 9 girls. How many choices are possible if

(a) all players are of the same gender?

(b) the team includes both genders?

1.19. Seven different gifts are to be distributed among 10 children. How many distinct results are possible if no child is to receive more than one gift?

1.20. A team of 9, consisting of 2 mathematicians, 3 statisticians, and 4 physicists, is to be selected from a faculty of 10 mathematicians, 8 statisticians, and 7 physicists. How many teams are possible?

1.21. From a group of 8 women and 6 men, a committee consisting of 3 men and 3 women is to be formed. How many different committees are possible if

(a) 2 of the men refuse to serve together?

(b) 2 of the women refuse to serve together?

(c) 1 man and 1 woman refuse to serve together?

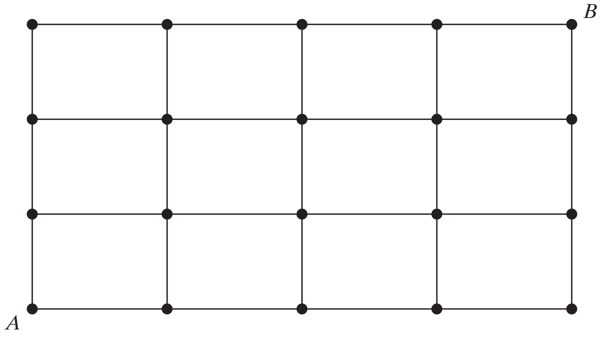
1.22. A person has 8 friends, of whom 5 will be invited to a party.

(a) How many choices are there if 2 of the friends are feuding and will not attend together?

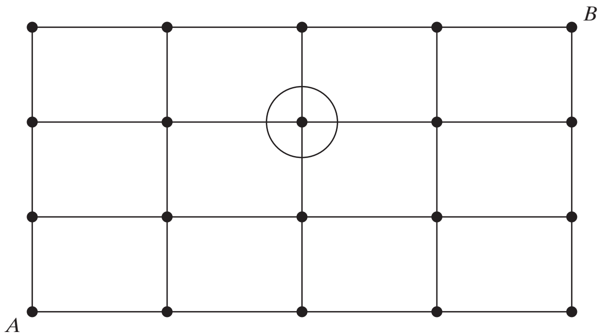
(b) How many choices if 2 of the friends will only attend together?

1.23. Consider the grid of points shown at the top of the next column. Suppose that, starting at the point labeled A, you can go one step up or one step to the right at each move. This procedure is continued until the point labeled B is reached. How many different paths from A to B are possible?

Hint: Note that to reach B from A, you must take 4 steps to the right and 3 steps upward.



24. In Problem 23, how many different paths are there from A to B that go through the point circled in the following lattice?



25. A psychology laboratory conducting dream research contains 3 rooms, with 2 beds in each room. If 3 sets of identical twins are to be assigned to these 6 beds so that each set of twins sleeps in different beds in the same room, how many assignments are possible?

26. (a) Show $\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n$

(b) Simplify $\sum_{k=0}^{n} \binom{n}{k} x^k$

27. Expand $(4x - 3y)^4$.

28. The game of bridge is played by 4 players, each of whom is dealt 13 cards. How many bridge deals are possible?

29. Expand $(x_1 + 2x_2 + 3x_3)^4$.

30. If 12 people are to be divided into 3 committees of respective sizes 3, 4, and 5, how many divisions are possible?

31. If 10 gifts are to be distributed among 3 friends, how many distributions are possible? What if each friend should receive at least 3 gifts?

32. Ten weight lifters are competing in a team weightlifting contest. Of the lifters, 3 are from the United States, 4 are from Russia, 2 are from China, and 1 is from Canada. If the scoring takes account of the countries that the lifters represent, but not their individual identities, how many different outcomes are possible from the point of view of scores? How many different outcomes correspond to results in which the United States has 1 competitor in the top three and 2 in the bottom three?

33. Delegates from 10 countries, including Russia, France, England, and the United States, are to be seated in a row. How many different seating arrangements are possible if the French and English delegates are to be seated next to each other and the Russian and U.S. delegates are not to be next to each other?

***34.** If 8 identical blackboards are to be divided among 4 schools, how many divisions are possible? How many if each school must receive at least 1 blackboard?

***35.** An elevator starts at the basement with 8 people (not including the elevator operator) and discharges them all by the time it reaches the top floor, number 6. In how many ways could the operator have perceived the people leaving the elevator if all people look alike to him? What if the 8 people consisted of 5 men and 3 women and the operator could tell a man from a woman?

***36.** We have $20,000 that must be invested among 4 possible opportunities. Each investment must be integral in units of $1000, and there are minimal investments that need to be made if one is to invest in these opportunities. The minimal investments are $2000, $2000, $3000, and $4000. How many different investment strategies are available if

(a) an investment must be made in each opportunity?

(b) investments must be made in at least 3 of the 4 opportunities?

***37.** Suppose that 10 fish are caught at a lake that contains 5 distinct types of fish.

(a) How many different outcomes are possible, where an outcome specifies the numbers of caught fish of each of the 5 types?

(b) How many outcomes are possible when 3 of the 10 fish caught are trout?

(c) How many when at least 2 of the 10 are trout?

Theoretical Exercises

1.1. Prove the generalized version of the basic counting principle.

1.2. Two experiments are to be performed. The first can result in any one of m possible outcomes. If the first experiment results in outcome i, then the second experiment can result in any of n_i possible outcomes, $i = 1, 2, \ldots, m$. What is the number of possible outcomes of the two experiments?

1.3. In how many ways can r objects be selected from a set of n objects if the order of selection is considered relevant?

1.4. There are $\binom{n}{r}$ different linear arrangements of n balls of which r are black and $n - r$ are white. Give a combinatorial explanation of this fact.

1.5. Determine the number of vectors $(x_1, \ldots, x_n)$, such that each x_i is either 0 or 1 and

$$\sum_{i=1}^{n} x_i \geq k$$

1.6. How many vectors $x_1, \ldots, x_k$ are there for which each x_i is a positive integer such that $1 \leq x_i \leq n$ and $x_1 < x_2 < \cdots < x_k$?

1.7. Give an analytic proof of Equation (4.1).

1.8. Prove that

$$\binom{n+m}{r} = \binom{n}{0}\binom{m}{r} + \binom{n}{1}\binom{m}{r-1} + \cdots + \binom{n}{r}\binom{m}{0}$$

Hint: Consider a group of n men and m women. How many groups of size r are possible?

1.9. Use Theoretical Exercise 8 to prove that

$$\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$$

1.10. From a group of n people, suppose that we want to choose a committee of $k, k \leq n$, one of whom is to be designated as chairperson.

(a) By focusing first on the choice of the committee and then on the choice of the chair, argue that there are $\binom{n}{k} k$ possible choices.

(b) By focusing first on the choice of the nonchair committee members and then on the choice of the chair, argue that there are $\binom{n}{k-1}(n - k + 1)$ possible choices.

(c) By focusing first on the choice of the chair and then on the choice of the other committee members, argue that there are $n\binom{n-1}{k-1}$ possible choices.

(d) Conclude from parts (a), (b), and (c) that

$$k\binom{n}{k} = (n - k + 1)\binom{n}{k-1} = n\binom{n-1}{k-1}$$

(e) Use the factorial definition of $\binom{m}{r}$ to verify the identity in part (d).

1.11. The following identity is known as Fermat's combinatorial identity:

$$\binom{n}{k} = \sum_{i=k}^{n} \binom{i-1}{k-1} \qquad n \geq k$$

Give a combinatorial argument (no computations are needed) to establish this identity.
Hint: Consider the set of numbers 1 through n. How many subsets of size k have i as their highest numbered member?

1.12. Consider the following combinatorial identity:

$$\sum_{k=1}^{n} k\binom{n}{k} = n \cdot 2^{n-1}$$

(a) Present a combinatorial argument for this identity by considering a set of n people and determining, in two ways, the number of possible selections of a committee of any size and a chairperson for the committee.
Hint:

(i) How many possible selections are there of a committee of size k and its chairperson?

(ii) How many possible selections are there of a chairperson and the other committee members?

(b) Verify the following identity for $n = 1, 2, 3, 4, 5$:

$$\sum_{k=1}^{n} \binom{n}{k} k^2 = 2^{n-2} n(n + 1)$$

For a combinatorial proof of the preceding, consider a set of n people and argue that both sides of the identity represent the number of different selections of a committee, its chairperson, and its secretary (possibly the same as the chairperson).

Hint:

(i) How many different selections result in the committee containing exactly k people?
(ii) How many different selections are there in which the chairperson and the secretary are the same? (ANSWER: $n2^{n-1}$.)
(iii) How many different selections result in the chairperson and the secretary being different?

(c) Now argue that

$$\sum_{k=1}^{n} \binom{n}{k} k^3 = 2^{n-3} n^2 (n + 3)$$

1.13. Show that, for $n > 0$,

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0$$

Hint: Use the binomial theorem.

1.14. From a set of n people, a committee of size j is to be chosen, and from this committee, a subcommittee of size $i, i \leq j$, is also to be chosen.

(a) Derive a combinatorial identity by computing, in two ways, the number of possible choices of the committee and subcommittee—first by supposing that the committee is chosen first and then the subcommittee is chosen, and second by supposing that the subcommittee is chosen first and then the remaining members of the committee are chosen.

(b) Use part (a) to prove the following combinatorial identity:

$$\sum_{j=i}^{n} \binom{n}{j} \binom{j}{i} = \binom{n}{i} 2^{n-i} \quad i \leq n$$

(c) Use part (a) and Theoretical Exercise 13 to show that

$$\sum_{j=i}^{n} \binom{n}{j} \binom{j}{i} (-1)^{n-j} = 0 \quad i < n$$

1.15. Let $H_k(n)$ be the number of vectors $x_1, \ldots, x_k$ for which each x_i is a positive integer satisfying $1 \leq x_i \leq n$ and $x_1 \leq x_2 \leq \cdots \leq x_k$.

(a) Without any computations, argue that

$$H_1(n) = n$$
$$H_k(n) = \sum_{j=1}^{n} H_{k-1}(j) \quad k > 1$$

Hint: How many vectors are there in which $x_k = j$?

(b) Use the preceding recursion to compute $H_3(5)$.
Hint: First compute $H_2(n)$ for $n = 1, 2, 3, 4, 5$.

1.16. Consider a tournament of n contestants in which the outcome is an ordering of these contestants, with ties allowed. That is, the outcome partitions the players into groups, with the first group consisting of the players who tied for first place, the next group being those who tied for the next-best position, and so on. Let $N(n)$ denote the number of different possible outcomes. For instance, $N(2) = 3$, since, in a tournament with 2 contestants, player 1 could be uniquely first, player 2 could be uniquely first, or they could tie for first.

(a) List all the possible outcomes when $n = 3$.

(b) With $N(0)$ defined to equal 1, argue, without any computations, that

$$N(n) = \sum_{i=1}^{n} \binom{n}{i} N(n - i)$$

Hint: How many outcomes are there in which i players tie for last place?

(c) Show that the formula of part (b) is equivalent to the following:

$$N(n) = \sum_{i=0}^{n-1} \binom{n}{i} N(i)$$

(d) Use the recursion to find $N(3)$ and $N(4)$.

1.17. Present a combinatorial explanation of why $\binom{n}{r} = \binom{n}{r, n - r}$.

1.18. Argue that

$$\begin{aligned}\binom{n}{n_1, n_2, \ldots, n_r} &= \binom{n - 1}{n_1 - 1, n_2, \ldots, n_r} \\ &\quad + \binom{n - 1}{n_1, n_2 - 1, \ldots, n_r} + \cdots \\ &\quad + \binom{n - 1}{n_1, n_2, \ldots, n_r - 1}\end{aligned}$$

Hint: Use an argument similar to the one used to establish Equation (4.1).

1.19. Prove the multinomial theorem.

***1.20.** In how many ways can n identical balls be distributed into r urns so that the ith urn contains at least m_i balls, for each $i = 1, \ldots, r$? Assume that $n \geq \sum_{i=1}^{r} m_i$.

***1.21.** Argue that there are exactly $\binom{r}{k} \binom{n - 1}{n - r + k}$ solutions of

$$x_1 + x_2 + \cdots + x_r = n$$

for which exactly k of the x_i are equal to 0.

***1.22.** Consider a function $f(x_1, \ldots, x_n)$ of n variables. How many different partial derivatives of order r does f possess?

***1.23.** Determine the number of vectors $(x_1, \ldots, x_n)$ such that each x_i is a nonnegative integer and

$$\sum_{i=1}^{n} x_i \le k$$

Self-Test Problems and Exercises

1.1. How many different linear arrangements are there of the letters A, B, C, D, E, F for which

(a) A and B are next to each other?
(b) A is before B?
(c) A is before B and B is before C?
(d) A is before B and C is before D?
(e) A and B are next to each other and C and D are also next to each other?
(f) E is not last in line?

1.2. If 4 Americans, 3 French people, and 3 British people are to be seated in a row, how many seating arrangements are possible when people of the same nationality must sit next to each other?

1.3. A president, treasurer, and secretary, all different, are to be chosen from a club consisting of 10 people. How many different choices of officers are possible if

(a) there are no restrictions?
(b) A and B will not serve together?
(c) C and D will serve together or not at all?
(d) E must be an officer?
(e) F will serve only if he is president?

1.4. A student is to answer 7 out of 10 questions in an examination. How many choices has she? How many if she must answer at least 3 of the first 5 questions?

1.5. In how many ways can a man divide 7 gifts among his 3 children if the eldest is to receive 3 gifts and the others 2 each?

1.6. How many different 7-place license plates are possible when 3 of the entries are letters and 4 are digits? Assume that repetition of letters and numbers is allowed and that there is no restriction on where the letters or numbers can be placed.

1.7. Give a combinatorial explanation of the identity

$$\binom{n}{r} = \binom{n}{n-r}$$

1.8. Consider n-digit numbers where each digit is one of the 10 integers $0, 1, \ldots, 9$. How many such numbers are there for which

(a) no two consecutive digits are equal?
(b) 0 appears as a digit a total of i times, $i = 0, \ldots, n$?

1.9. Consider three classes, each consisting of n students. From this group of $3n$ students, a group of 3 students is to be chosen.

(a) How many choices are possible?
(b) How many choices are there in which all 3 students are in the same class?
(c) How many choices are there in which 2 of the 3 students are in the same class and the other student is in a different class?
(d) How many choices are there in which all 3 students are in different classes?
(e) Using the results of parts (a) through (d), write a combinatorial identity.

1.10. How many 5-digit numbers can be formed from the integers $1, 2, \ldots, 9$ if no digit can appear more than twice? (For instance, 41434 is not allowed.)

1.11. From 10 opposite-sexed couples, we want to select a group of 6 people that is not allowed to contain a married couple.
(a) How many choices are there?
(b) How many choices are there if the group must also consist of 3 men and 3 women?

1.12. A committee of 6 people is to be chosen from a group consisting of 7 men and 8 women. If the committee must consist of at least 3 women and at least 2 men, how many different committees are possible?

***1.13.** An art collection on auction consisted of 4 Dalis, 5 van Goghs, and 6 Picassos. At the auction were 5 art collectors. If a reporter noted only the number of Dalis, van Goghs, and Picassos acquired by each collector, how many different results could have been recorded if all of the works were sold?

***1.14.** Determine the number of vectors $(x_1, \ldots, x_n)$ such that each x_i is a positive integer and

$$\sum_{i=1}^{n} x_i \leq k$$

where $k \geq n$.

1.15. A total of n students are enrolled in a review course for the actuarial examination in probability. The posted results of the examination will list the names of those who passed, in decreasing order of their scores. For instance, the posted result will be "Brown, Cho" if Brown and Cho are the only ones to pass, with Brown receiving the higher score. Assuming that all scores are distinct (no ties), how many posted results are possible?

1.16. How many subsets of size 4 of the set $S = \{1, 2, \ldots, 20\}$ contain at least one of the elements $1, 2, 3, 4, 5$?

1.17. Give an analytic verification of

$$\binom{n}{2} = \binom{k}{2} + k(n - k) + \binom{n-k}{2}, \quad 1 \leq k \leq n$$

Now, give a combinatorial argument for this identity.

1.18. In a certain community, there are 3 families consisting of a single parent and 1 child, 3 families consisting of a single parent and 2 children, 5 families consisting of 2 parents and a single child, 7 families consisting of 2 parents and 2 children, and 6 families consisting of 2 parents and 3 children. If a parent and child from the same family are to be chosen, how many possible choices are there?

1.19. If there are no restrictions on where the digits and letters are placed, how many 8-place license plates consisting of 5 letters and 3 digits are possible if no repetitions of letters or digits are allowed? What if the 3 digits must be consecutive?

1.20. Verify the identity

$$\sum_{x_1+\ldots+x_r=n,\ x_i \geq 0} \frac{n!}{x_1!x_2!\cdots x_r!} = r^n$$

(a) by a combinatorial argument that first notes that r^n is the number of different n letter sequences that can be formed from an alphabet consisting of r letters, and then determines how many of these letter sequences have letter 1 a total of x_1 times and letter 2 a total of x_2 times and ... and letter r a total of x_r times;

(b) by using the multinomial theorem.

1.21. Simplify $n - \binom{n}{2} + \binom{n}{3} - \ldots + (-1)^{n+1}\binom{n}{n}$

Chapter 2

AXIOMS OF PROBABILITY

Contents

2.1 Introduction

In this chapter, we introduce the concept of the probability of an event and then show how probabilities can be computed in certain situations. As a preliminary, however, we need to discuss the concept of the sample space and the events of an experiment.

2.2 Sample Space and Events

Consider an experiment whose outcome is not predictable with certainty. However, although the outcome of the experiment will not be known in advance, let us suppose that the set of all possible outcomes is known. This set of all possible outcomes of an experiment is known as the *sample space* of the experiment and is denoted by S. Following are some examples:

1. If the outcome of an experiment consists of the determination of the sex of a newborn child, then

$$S = \{g, b\}$$

where the outcome g means that the child is a girl and b that it is a boy.

2. If the outcome of an experiment is the order in which seven people $\{P_1, P_2, ..., P_7\}$ are waiting in a queue to withdraw cash from an ATM machine, then

$$S = \{\text{all } 7! \text{ permutations of } (P_1, P_2, P_3, P_4, P_5, P_6, P_7)\}$$

The outcome $(P_5, P_4, P_6, P_1, P_7, P_2P_3)$ means, for instance, that P_5 is first in the queue, P_4 is second, P_6 is third, and so on.

3. If the experiment consists of flipping two coins, then the sample space consists of the following four points:

$$S = \{(h, h), (h, t), (t, h), (t, t)\}$$

The outcome will be (h, h) if both coins are heads, (h, t) if the first coin is heads and the second tails, (t, h) if the first is tails and the second heads, and (t, t) if both coins are tails.

4. If the experiment consists of tossing two dice, then the sample space consists of the 36 points

$$S = \{(i, j)\colon i, j = 1, 2, 3, 4, 5, 6\}$$

where the outcome (i, j) is said to occur if i appears on the leftmost die and j on the other die.

5. If the experiment consists of measuring (in hours) the lifetime of a transistor, then the sample space consists of all nonnegative real numbers; that is,

$$S = \{x\colon 0 \leq x < \infty\}$$

Any subset E of the sample space is known as an *event*. In other words, an event is a set consisting of possible outcomes of the experiment. If the outcome of the experiment is contained in E, then we say that E has occurred. Following are some examples of events.

In the preceding Example 1, if $E = \{g\}$, then E is the event that the child is a girl. Similarly, if $F = \{b\}$, then F is the event that the child is a boy.

In Example 2, if

$$E = \{\text{all outcomes in } S \text{ starting with } P_3\}$$

then E is the event that the person P_3 is first in the queue.

In Example 3, if $E = \{(h, h), (h, t)\}$, then E is the event that a head appears on the first coin.

In Example 4, if $E = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$, then E is the event that the sum of the dice equals 7.

In Example 5, if $E = \{x\colon 0 \leq x \leq 5\}$, then E is the event that the transistor does not last longer than 5 hours.

For any two events E and F of a sample space S, we define the new event $E \cup F$ to consist of all outcomes that are either in E or in F or in both E and F. That is, the event $E \cup F$ will occur if *either* E or F occurs. For instance, in Example 1, if $E = \{g\}$ is the event that the child is a girl and $F = \{b\}$ is the event that the child is a boy, then

$$E \cup F = \{g, b\}$$

is the whole sample space S. In Example 3, if $E = \{(h, h), (h, t)\}$ is the event that the first coin lands heads, and $F = \{(t, h), (h, h)\}$ is the event that the second coin lands heads, then

$$E \cup F = \{(h, h), (h, t), (t, h)\}$$

is the event that at least one of the coins lands heads and thus will occur provided that both coins do not land tails.

The event $E \cup F$ is called the *union* of the event E and the event F.

Similarly, for any two events E and F, we may also define the new event EF, called the *intersection* of E and F, to consist of all outcomes that are both in E and in F. That is, the event EF (sometimes written $E \cap F$) will occur only if both E and F occur. For instance, in Example 3, if $E = \{(h,h),(h,t),(t,h)\}$ is the event that at least 1 head occurs and $F = \{(h,t),(t,h),(t,t)\}$ is the event that at least 1 tail occurs, then

$$EF = \{(h,t),(t,h)\}$$

is the event that exactly 1 head and 1 tail occur. In Example 4, if $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ is the event that the sum of the dice is 7 and $F = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$ is the event that the sum is 6, then the event EF does not contain any outcomes and hence could not occur. To give such an event a name, we shall refer to it as the null event and denote it by Ø. (That is, Ø refers to the event consisting of no outcomes.) If $EF = \emptyset$, then E and F are said to be *mutually exclusive*.

We define unions and intersections of more than two events in a similar manner. If $E_1, E_2, \ldots$ are events, then the union of these events, denoted by $\bigcup_{n=1}^{\infty} E_n$, is defined to be that event that consists of all outcomes that are in E_n for at least one value of $n = 1,2,\ldots$. Similarly, the intersection of the events E_n, denoted by $\bigcap_{n=1}^{\infty} E_n$, is defined to be the event consisting of those outcomes that are in all of the events $E_n, n = 1,2,\ldots$.

Finally, for any event E, we define the new event E^c, referred to as the *complement* of E, to consist of all outcomes in the sample space S that are not in E. That is, E^c will occur if and only if E does not occur. In Example 4, if event $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, then E^c will occur when the sum of the dice does not equal 7. Note that because the experiment must result in some outcome, it follows that $S^c = \emptyset$.

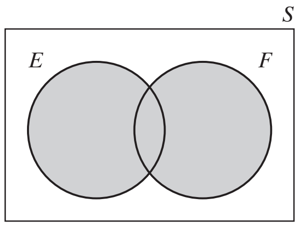
For any two events E and F, if all of the outcomes in E are also in F, then we say that E is *contained* in F, or E is a *subset* of F, and write $E \subset F$ (or equivalently, $F \supset E$, which we sometimes say as F is a *superset* of E). Thus, if $E \subset F$, then the occurrence of E implies the occurrence of F. If $E \subset F$ and $F \subset E$, we say that E and F are equal and write $E = F$.

A graphical representation that is useful for illustrating logical relations among events is the Venn diagram. The sample space S is represented as consisting of all the outcomes in a large rectangle, and the events $E, F, G, \ldots$ are represented as consisting of all the outcomes in given circles within the rectangle. Events of interest can then be indicated by shading appropriate regions of the diagram. For instance, in the three Venn diagrams shown in Figure 2.1, the shaded areas represent, respectively, the events $E \cup F$, EF, and E^c. The Venn diagram in Figure 2.2 indicates that $E \subset F$.

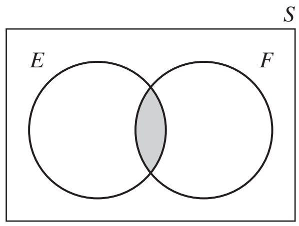
The operations of forming unions, intersections, and complements of events obey certain rules similar to the rules of algebra. We list a few of these rules:

Commutative laws	$E \cup F = F \cup E$	$EF = FE$
Associative laws	$(E \cup F) \cup G = E \cup (F \cup G)$	$(EF)G = E(FG)$
Distributive laws	$(E \cup F)G = EG \cup FG$	$EF \cup G = (E \cup G)(F \cup G)$

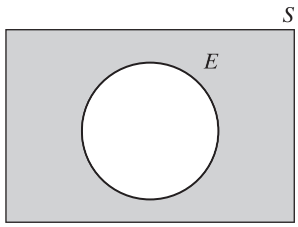
These relations are verified by showing that any outcome that is contained in the event on the left side of the equality sign is also contained in the event on the right side, and vice versa. One way of showing this is by means of Venn diagrams. For instance, the distributive law may be verified by the sequence of diagrams in Figure 2.3.



(a) Shaded region: $E \cup F$.



(b) Shaded region: EF.



(c) Shaded region: E^c.

Figure 2.1 Venn diagrams.

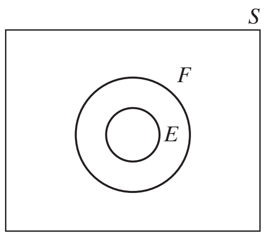
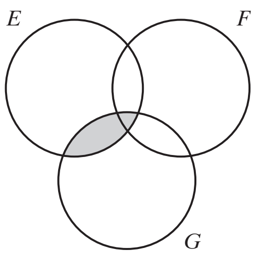
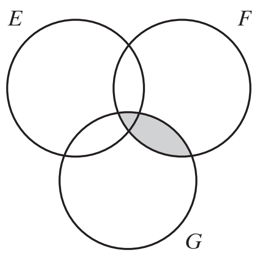


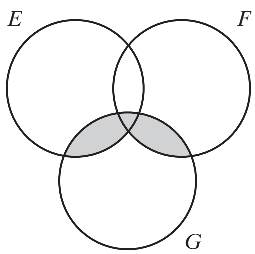
Figure 2.2 $E \subset F$.



(a) Shaded region: EG.



(b) Shaded region: FG.



(c) Shaded region: $(E \cup F)G$.

Figure 2.3 $(E \cup F)G = EG \cup FG$.

The following useful relationships among the three basic operations of forming unions, intersections, and complements are known as *DeMorgan's laws*:

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i^c$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i^c$$

For instance, for two events E and F, DeMorgan's laws state that

$$(E \cup F)^c = E^c F^c \quad \text{and} \quad (EF)^c = E^c \cup F^c$$

which can be easily proven by using Venn diagrams (see Theoretical Exercise 7).

To prove DeMorgan's laws for general n, suppose first that x is an outcome of $\left(\bigcup_{i=1}^{n} E_i\right)^c$. Then x is not contained in $\bigcup_{i=1}^{n} E_i$, which means that x is not contained in any of the events $E_i, i = 1, 2, \ldots, n$, implying that x is contained in E_i^c for all $i = 1, 2, \ldots, n$ and thus is contained in $\bigcap_{i=1}^{n} E_i^c$. To go the other way, suppose that x is an outcome of $\bigcap_{i=1}^{n} E_i^c$. Then x is contained in E_i^c for all $i = 1, 2, \ldots, n$, which means that x is not contained in E_i for any $i = 1, 2, \ldots, n$, implying that x is not contained in $\bigcup_{i}^{n} E_i$, in turn implying that x is contained in $\left(\bigcup_{1}^{n} E_i\right)^c$. This proves the first of DeMorgan's laws.

To prove the second of DeMorgan's laws, we use the first law to obtain

$$\left(\bigcup_{i=1}^{n} E_i^c\right)^c = \bigcap_{i=1}^{n} (E_i^c)^c$$

which, since $(E^c)^c = E$, is equivalent to

$$\left(\bigcup_{1}^{n} E_i^c\right)^c = \bigcap_{1}^{n} E_i$$

Taking complements of both sides of the preceding equation yields the result we seek, namely,

$$\bigcup_{1}^{n} E_i^c = \left(\bigcap_{1}^{n} E_i\right)^c$$

2.3 Axioms of Probability

One way of defining the probability of an event is in terms of its long run *relative frequency*. Such a definition usually goes as follows: We suppose that an experiment, whose sample space is S, is repeatedly performed under exactly the same conditions. For each event E of the sample space S, we define $n(E)$ to be the number of times in the first n repetitions of the experiment that the event E occurs. Then $P(E)$, the *probability* of the event E, is defined as

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

That is, $P(E)$ is defined as the (limiting) proportion of time that E occurs. It is thus the limiting relative frequency of E.

Although the preceding definition is certainly intuitively pleasing and should always be kept in mind by the reader, it possesses a serious drawback: How do we know that $n(E)/n$ will converge to some constant limiting value that will be the same for each possible sequence of repetitions of the experiment? For example, suppose that the experiment to be repeatedly performed consists of flipping a coin. How do we know that the proportion of heads obtained in the first n flips will converge to some value as n gets large? Also, even if it does converge to some value, how do we know that, if the experiment is repeatedly performed a second time, we shall obtain the same limiting proportion of heads?

Proponents of the relative frequency definition of probability usually answer this objection by stating that the convergence of $n(E)/n$ to a constant limiting value is an assumption, or an *axiom*, of the system. However, to assume that $n(E)/n$ will necessarily converge to some constant value seems to be an extraordinarily complicated assumption. For, although we might indeed hope that such a constant limiting frequency exists, it does not at all seem to be a priori evident that this need be the case. In fact, would it not be more reasonable to assume a set of simpler and more self-evident axioms about probability and then attempt to prove that such a constant limiting frequency does in some sense exist? The latter approach is the modern axiomatic approach to probability theory that we shall adopt in this text. In particular, we shall assume that, for each event E in the sample space S, there exists a value $P(E)$, referred to as the probability of E. We shall then assume that all these probabilities satisfy a certain set of axioms, which, we hope the reader will agree, is in accordance with our intuitive notion of probability.

Consider an experiment whose sample space is S. For each event E of the sample space S, we assume that a number $P(E)$ is defined and satisfies the following three axioms:

The three axioms of probability

Axiom 1

$$0 \leq P(E) \leq 1$$

Axiom 2

$$P(S) = 1$$

Axiom 3

For any sequence of mutually exclusive events $E_1, E_2, \ldots$ (that is, events for which $E_i E_j = \emptyset$ when $i \neq j$),

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

We refer to $P(E)$ as the *probability* of the event E.

Thus, Axiom 1 states that the probability that the outcome of the experiment is an outcome in E is some number between 0 and 1. Axiom 2 states that, with probability 1, the outcome will be a point in the sample space S. Axiom 3 states that, for any sequence of mutually exclusive events, the probability of at least one of these events occurring is just the sum of their respective probabilities.

If we consider a sequence of events $E_1, E_2, \ldots$, where $E_1 = S$ and $E_i = \emptyset$ for $i > 1$, then, because the events are mutually exclusive and because $S = \bigcup_{i=1}^{\infty} E_i$, we have, from Axiom 3,

$$P(S) = \sum_{i=1}^{\infty} P(E_i) = P(S) + \sum_{i=2}^{\infty} P(\emptyset)$$

implying that

$$P(\emptyset) = 0$$

That is, the null event has probability 0 of occurring.

Note that it follows that, for any finite sequence of mutually exclusive events E_1, $E_2, \ldots, E_n$,

$$P\left(\bigcup_{1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) \tag{3.1}$$

This equation follows from Axiom 3 by defining E_i as the null event for all values of i greater than n. Axiom 3 is equivalent to Equation (3.1) when the sample space is finite. (Why?) However, the added generality of Axiom 3 is necessary when the sample space consists of an infinite number of points.

Example 3a

If our experiment consists of tossing a coin and if we assume that a head is as likely to appear as a tail, then we would have

$$P(\{H\}) = P(\{T\}) = \frac{1}{2}$$

On the other hand, if the coin were biased and we believed that a head were twice as likely to appear as a tail, then we would have

$$P(\{H\}) = \frac{2}{3} \quad P(\{T\}) = \frac{1}{3}$$ ■

Example 3b

If a die is rolled and we suppose that all six sides are equally likely to appear, then we would have $P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = P(\{5\}) = P(\{6\}) = \frac{1}{6}$. From Axiom 3, it would thus follow that the probability of rolling an even number would equal

$$P(\{2, 4, 6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = \frac{1}{2}$$ ■

The assumption of the existence of a set function P, defined on the events of a sample space S and satisfying Axioms 1, 2, and 3, constitutes the modern mathematical approach to probability theory. It is hoped that the reader will agree that the axioms are natural and in accordance with our intuitive concept of probability as related to chance and randomness. Furthermore, using these axioms, we shall be able to prove that if an experiment is repeated over and over again, then, with probability

1, the proportion of time during which any specific event E occurs will equal $P(E)$. This result, known as the strong law of large numbers, is presented in Chapter 8. In addition, we present another possible interpretation of probability—as being a measure of belief—in Section 2.7.

Technical Remark. We have supposed that $P(E)$ is defined for all the events E of the sample space. Actually, when the sample space is an uncountably infinite set, $P(E)$ is defined only for a class of events called *measurable*. However, this restriction need not concern us, as all events of any practical interest are measurable.

2.4 Some Simple Propositions

In this section, we prove some simple propositions regarding probabilities. We first note that since E and E^c are always mutually exclusive and since $E \cup E^c = S$, we have, by Axioms 2 and 3,

$$1 = P(S) = P(E \cup E^c) = P(E) + P(E^c)$$

Or, equivalently, we have Proposition 4.1.

Proposition 4.1

$$P(E^c) = 1 - P(E)$$

In words, Proposition 4.1 states that the probability that an event does not occur is 1 minus the probability that it does occur. For instance, if the probability of obtaining a head on the toss of a coin is $\frac{3}{8}$, then the probability of obtaining a tail must be $\frac{5}{8}$.

Our second proposition states that if the event E is contained in the event F, then the probability of E is no greater than the probability of F.

Proposition 4.2 If $E \subset F$, then $P(E) \leq P(F)$.

Proof Since $E \subset F$, it follows that we can express F as

$$F = E \cup E^c F$$

Hence, because E and $E^c F$ are mutually exclusive, we obtain, from Axiom 3,

$$P(F) = P(E) + P(E^c F)$$

which proves the result, since $P(E^c F) \geq 0$. □

Proposition 4.2 tells us, for instance, that the probability of rolling a 1 with a die is less than or equal to the probability of rolling an odd value with the die.

The next proposition gives the relationship between the probability of the union of two events, expressed in terms of the individual probabilities, and the probability of the intersection of the events.

Proposition 4.3

$$P(E \cup F) = P(E) + P(F) - P(EF)$$

Proof To derive a formula for $P(E \cup F)$, we first note that $E \cup F$ can be written as the union of the two disjoint events E and $E^c F$. Thus, from Axiom 3, we obtain

$$\begin{aligned} P(E \cup F) &= P(E \cup E^c F) \\ &= P(E) + P(E^c F) \end{aligned}$$

Furthermore, since $F = EF \cup E^cF$, we again obtain from Axiom 3

$$P(F) = P(EF) + P(E^cF)$$

or, equivalently,

$$P(E^cF) = P(F) - P(EF)$$

thereby completing the proof. □

Proposition 4.3 could also have been proved by making use of the Venn diagram in Figure 2.4.

Let us divide $E \cup F$ into three mutually exclusive sections, as shown in Figure 2.5. In words, section I represents all the points in E that are not in F (that is, EF^c), section II represents all points both in E and in F (that is, EF), and section III represents all points in F that are not in E (that is, E^cF).

From Figure 2.5, we see that

$$\begin{aligned} E \cup F &= \mathrm{I} \cup \mathrm{II} \cup \mathrm{III} \\ E &= \mathrm{I} \cup \mathrm{II} \\ F &= \mathrm{II} \cup \mathrm{III} \end{aligned}$$

As I, II, and III are mutually exclusive, it follows from Axiom 3 that

$$\begin{aligned} P(E \cup F) &= P(\mathrm{I}) + P(\mathrm{II}) + P(\mathrm{III}) \\ P(E) &= P(\mathrm{I}) + P(\mathrm{II}) \\ P(F) &= P(\mathrm{II}) + P(\mathrm{III}) \end{aligned}$$

which shows that

$$P(E \cup F) = P(E) + P(F) - P(\mathrm{II})$$

and Proposition 4.3 is proved, since $\mathrm{II} = EF$.

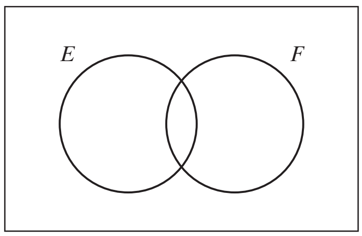


Figure 2.4 Venn diagram.

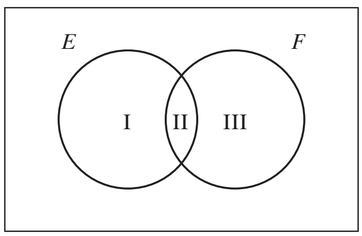


Figure 2.5 Venn diagram in sections.

Example 4a J is taking two books along on her holiday vacation. With probability .5, she will like the first book; with probability .4, she will like the second book; and with probability .3, she will like both books. What is the probability that she likes neither book?

Solution Let B_i denote the event that J likes book $i, i = 1, 2$. Then the probability that she likes at least one of the books is

$$P(B_1 \cup B_2) = P(B_1) + P(B_2) - P(B_1B_2) = .5 + .4 - .3 = .6$$

Because the event that J likes neither book is the complement of the event that she likes at least one of them, we obtain the result

$$P(B_1^cB_2^c) = P\left((B_1 \cup B_2)^c\right) = 1 - P(B_1 \cup B_2) = .4 \qquad \blacksquare$$

We may also calculate the probability that any one of the three events E, F, and G occurs, namely,

$$P(E \cup F \cup G) = P[(E \cup F) \cup G]$$

which, by Proposition 4.3, equals

$$P(E \cup F) + P(G) - P[(E \cup F)G]$$

Now, it follows from the distributive law that the events $(E \cup F)G$ and $EG \cup FG$ are equivalent; hence, from the preceding equations, we obtain

$$\begin{aligned}
&P(E \cup F \cup G)\\
&\quad = P(E) + P(F) - P(EF) + P(G) - P(EG \cup FG)\\
&\quad = P(E) + P(F) - P(EF) + P(G) - P(EG) - P(FG) + P(EGFG)\\
&\quad = P(E) + P(F) + P(G) - P(EF) - P(EG) - P(FG) + P(EFG)
\end{aligned}$$

In fact, the following proposition, known as the *inclusion–exclusion identity*, can be proved by mathematical induction:

Proposition 4.4

$$\begin{aligned}
P(E_1 \cup E_2 \cup \cdots \cup E_n) = &\sum_{i=1}^{n} P(E_i) - \sum_{i_1<i_2} P(E_{i_1}E_{i_2}) + \cdots\\
&+ (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} P(E_{i_1}E_{i_2}\cdots E_{i_r})\\
&+ \cdots + (-1)^{n+1}P(E_1E_2\cdots E_n)
\end{aligned}$$

The summation $\sum\limits_{i_1<i_2<\cdots<i_r} P(E_{i_1}E_{i_2}\cdots E_{i_r})$ is taken over all of the $\binom{n}{r}$ possible subsets of size r of the set $\{1, 2, \ldots, n\}$.

In words, Proposition 4.4 states that the probability of the union of n events equals the sum of the probabilities of these events taken one at a time, minus the sum of the probabilities of these events taken two at a time, plus the sum of the probabilities of these events taken three at a time, and so on.

Remarks 1. For a noninductive argument for Proposition 4.4, note first that if an outcome of the sample space is not a member of any of the sets E_i, then its probability does not contribute anything to either side of the equality. Now, suppose that an outcome is in exactly m of the events E_i, where $m > 0$. Then, since it is in $\bigcup\limits_i E_i$, its probability is counted once in $P\left(\bigcup_i E_i\right)$; also, as this outcome is contained in $\binom{m}{k}$ subsets of the type $E_{i_1}E_{i_2}\cdots E_{i_k}$, its probability is counted

$$\binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \cdots \pm \binom{m}{m}$$

times on the right of the equality sign in Proposition 4.4. Thus, for $m > 0$, we must show that

$$1 = \binom{m}{1} - \binom{m}{2} + \binom{m}{3} - \cdots \pm \binom{m}{m}$$

However, since $1 = \binom{m}{0}$, the preceding equation is equivalent to

$$\sum_{i=0}^{m} \binom{m}{i} (-1)^i = 0$$

and the latter equation follows from the binomial theorem, since

$$0 = (-1 + 1)^m = \sum_{i=0}^{m} \binom{m}{i} (-1)^i (1)^{m-i}$$

2. The following is a succinct way of writing the inclusion–exclusion identity:

$$P(\cup_{i=1}^n E_i) = \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1 < \cdots < i_r} P(E_{i_1} \cdots E_{i_r})$$

3. In the inclusion–exclusion identity, going out one term results in an upper bound on the probability of the union, going out two terms results in a lower bound on the probability, going out three terms results in an upper bound on the probability, going out four terms results in a lower bound, and so on. That is, for events $E_1, \ldots, E_n$, we have

$$P(\cup_{i=1}^n E_i) \le \sum_{i=1}^{n} P(E_i) \tag{4.1}$$

$$P(\cup_{i=1}^n E_i) \ge \sum_{i=1}^{n} P(E_i) - \sum_{j<i} P(E_i E_j) \tag{4.2}$$

$$P(\cup_{i=1}^n E_i) \le \sum_{i=1}^{n} P(E_i) - \sum_{j<i} P(E_i E_j) + \sum_{k<j<i} P(E_i E_j E_k) \tag{4.3}$$

and so on. To prove the validity of these bounds, note the identity

$$\cup_{i=1}^n E_i = E_1 \cup E_1^c E_2 \cup E_1^c E_2^c E_3 \cup \cdots \cup E_1^c \cdots E_{n-1}^c E_n$$

That is, at least one of the events E_i occurs if E_1 occurs, or if E_1 does not occur but E_2 does, or if E_1 and E_2 do not occur but E_3 does, and so on. Because the right-hand side is the union of disjoint events, we obtain

$$P(\cup_{i=1}^n E_i) = P(E_1) + P(E_1^c E_2) + P(E_1^c E_2^c E_3) + \ldots + P(E_1^c \cdots E_{n-1}^c E_n)$$
$$= P(E_1) + \sum_{i=2}^{n} P(E_1^c \cdots E_{i-1}^c E_i) \tag{4.4}$$

Now, let $B_i = E_1^c \cdots E_{i-1}^c = (\cup_{j<i} E_j)^c$ be the event that none of the first $i - 1$ events occurs. Applying the identity

$$P(E_i) = P(B_i E_i) + P(B_i^c E_i)$$

shows that

$$P(E_i) = P(E_1^c \cdots E_{i-1}^c E_i) + P(E_i \cup_{j<i} E_j)$$

or, equivalently,

$$P(E_1^c \cdots E_{i-1}^c E_i) = P(E_i) - P(\cup_{j<i} E_i E_j)$$

Substituting this equation into (4.4) yields

$$P(\cup_{i=1}^n E_i) = \sum_i P(E_i) - \sum_i P(\cup_{j<i} E_i E_j) \tag{4.5}$$

Because probabilities are always nonnegative, Inequality (4.1) follows directly from Equation (4.5). Now, fixing i and applying Inequality (4.1) to $P(\cup_{j<i} E_i E_j)$ yields

$$P(\cup_{j<i} E_i E_j) \le \sum_{j<i} P(E_i E_j)$$

which, by Equation (4.5), gives Inequality (4.2). Similarly, fixing i and applying Inequality (4.2) to $P(\cup_{j<i} E_i E_j)$ yields

$$P(\cup_{j<i} E_i E_j) \ge \sum_{j<i} P(E_i E_j) - \sum_{k<j<i} P(E_i E_j E_i E_k)$$
$$= \sum_{j<i} P(E_i E_j) - \sum_{k<j<i} P(E_i E_j E_k)$$

which, by Equation (4.5), gives Inequality (4.3). The next inclusion–exclusion inequality is now obtained by fixing i and applying Inequality (4.3) to $P(\cup_{j<i} E_i E_j)$, and so on.

The first inclusion-exclusion inequality, namely that

$$P(\cup_{i=1}^n E_i) \le \sum_{i=1}^{n} P(E_i)$$

is known as *Boole's inequality*.

2.5 Sample Spaces Having Equally Likely Outcomes

In many experiments, it is natural to assume that all outcomes in the sample space are equally likely to occur. That is, consider an experiment whose sample space S is a finite set, say, $S = \{1, 2, \ldots, N\}$. Then, it is often natural to assume that

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{N\})$$

which implies, from Axioms 2 and 3 (why?), that

$$P(\{i\}) = \frac{1}{N} \quad i = 1, 2, \ldots, N$$

From this equation, it follows from Axiom 3 that, for any event E,

$$P(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } S}$$

In words, if we assume that all outcomes of an experiment are equally likely to occur, then the probability of any event E equals the proportion of outcomes in the sample space that are contained in E.

Example 5a

If two dice are rolled, what is the probability that the sum of the upturned faces will equal 7?

Solution We shall solve this problem under the assumption that all of the 36 possible outcomes are equally likely. Since there are 6 possible outcomes—namely, (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), and (6, 1)—that result in the sum of the dice being equal to 7, the desired probability is $\frac{6}{36} = \frac{1}{6}$. ■

Example 5b

If 3 balls are "randomly drawn" from a bowl containing 6 white and 5 black balls, what is the probability that one of the balls is white and the other two black?

Solution If we regard the balls as being distinguishable and the order in which they are selected as being relevant, then the sample space consists of $11 \cdot 10 \cdot 9 = 990$ outcomes. Furthermore, there are $6 \cdot 5 \cdot 4 = 120$ outcomes in which the first ball selected is white and the other two are black; $5 \cdot 6 \cdot 4 = 120$ outcomes in which the first is black, the second is white, and the third is black; and $5 \cdot 4 \cdot 6 = 120$ in which the first two are black and the third is white. Hence, assuming that "randomly drawn" means that each outcome in the sample space is equally likely to occur, we see that the desired probability is

$$\frac{120 + 120 + 120}{990} = \frac{4}{11}$$

This problem could also have been solved by regarding the outcome of the experiment as the unordered set of drawn balls. From this point of view, there are $\binom{11}{3} = 165$ outcomes in the sample space. Now, each set of 3 balls corresponds to 3! outcomes when the order of selection is noted. As a result, if all outcomes are assumed equally likely when the order of selection is noted, then it follows that they remain equally likely when the outcome is taken to be the unordered set of selected balls. Hence, using the latter representation of the experiment, we see that the desired probability is

$$\frac{\binom{6}{1}\binom{5}{2}}{\binom{11}{3}} = \frac{4}{11}$$

which, of course, agrees with the answer obtained previously. ■

When the experiment consists of a random selection of k items from a set of n items, we have the flexibility of either letting the outcome of the experiment be the ordered selection of the k items or letting it be the unordered set of items selected. In the former case, we would assume that each new selection is equally likely to be any of the so far unselected items of the set, and in the latter case, we would assume that all $\binom{n}{k}$ possible subsets of k items are equally likely to be the set selected. For instance, suppose 5 people are to be randomly selected from a group of 20 individuals consisting of 10 married couples, and we want to determine $P(N)$, the probability that the 5 chosen are all unrelated. (That is, no two are married to each other.) If we regard the sample space as the set of 5 people chosen, then there are $\binom{20}{5}$ equally likely outcomes. An outcome that does not contain a married couple can be thought of as being the result of a six-stage experiment: In the first stage, 5 of the 10 couples to have a member in the group are chosen; in the next 5 stages, 1 of the 2 members of each of these couples is selected. Thus, there are $\binom{10}{5}2^5$ possible outcomes in which the 5 members selected are unrelated, yielding the desired probability of

$$P(N) = \frac{\binom{10}{5}2^5}{\binom{20}{5}}$$

In contrast, we could let the outcome of the experiment be the *ordered* selection of the 5 individuals. In this setting, there are $20 \cdot 19 \cdot 18 \cdot 17 \cdot 16$ equally likely outcomes, of which $20 \cdot 18 \cdot 16 \cdot 14 \cdot 12$ outcomes result in a group of 5 unrelated individuals, yielding the result

$$P(N) = \frac{20 \cdot 18 \cdot 16 \cdot 14 \cdot 12}{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16}$$

We leave it for the reader to verify that the two answers are identical.

Example 5c A committee of 5 is to be selected from a group of 6 men and 9 women. If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women?

Solution Because each of the $\binom{15}{5}$ possible committees is equally likely to be selected, the desired probability is

$$\frac{\binom{6}{3}\binom{9}{2}}{\binom{15}{5}} = \frac{240}{1001} \qquad \blacksquare$$

Example 5d An urn contains n balls, one of which is special. If k of these balls are withdrawn one at a time, with each selection being equally likely to be any of the balls that remain at the time, what is the probability that the special ball is chosen?

Solution Since all of the balls are treated in an identical manner, it follows that the set of k balls selected is equally likely to be any of the $\binom{n}{k}$ sets of k balls. Therefore,

$$P\{\text{special ball is selected}\} = \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{k}{n}$$

We could also have obtained this result by letting A_i denote the event that the special ball is the ith ball to be chosen, $i = 1, \ldots, k$. Then, since each one of the n balls is equally likely to be the ith ball chosen, it follows that $P(A_i) = 1/n$. Hence, because these events are clearly mutually exclusive, we have

$$P\{\text{special ball is selected}\} = P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i) = \frac{k}{n}$$

We could also have argued that $P(A_i) = 1/n$, by noting that there are $n(n-1)\cdots(n-k+1) = n!/(n-k)!$ equally likely outcomes of the experiment, of which $(n-1)(n-2)\cdots(n-i+1)(1)(n-i)\cdots(n-k+1) = (n-1)!/(n-k)!$ result in the special ball being the ith one chosen. From this reasoning, it follows that

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$$ ■

Example 5e Suppose that $n + m$ balls, of which n are red and m are blue, are arranged in a linear order in such a way that all $(n + m)!$ possible orderings are equally likely. If we record the result of this experiment by listing only the colors of the successive balls, show that all the possible results remain equally likely.

Solution Consider any one of the $(n + m)!$ possible orderings, and note that any permutation of the red balls among themselves and of the blue balls among themselves does not change the sequence of colors. As a result, every ordering of colorings corresponds to $n!\, m!$ different orderings of the $n + m$ balls, so every ordering of the colors has probability $\frac{n!m!}{(n+m)!}$ of occurring.

For example, suppose that there are 2 red balls, numbered r_1, r_2, and 2 blue balls, numbered b_1, b_2. Then, of the 4! possible orderings, there will be 2! 2! orderings that result in any specified color combination. For instance, the following orderings result in the successive balls alternating in color, with a red ball first:

$$r_1, b_1, r_2, b_2 \quad r_1, b_2, r_2, b_1 \quad r_2, b_1, r_1, b_2 \quad r_2, b_2, r_1, b_1$$

Therefore, each of the possible orderings of the colors has probability $\frac{4}{24} = \frac{1}{6}$ of occurring. ■

Example 5f A woman gets a set of coupons to buy items from 5 different shops located on 4 different streets. There are 13 shops on each of these streets, numbered 1 to 13.

If the set of coupons allows you to buy from 5 shops that are consecutively numbered and are not all on the same street, we say that the set is straight. For instance, if the coupons are meant to be bought from shops 5,6,7, and 8 on the second street and shop 9 on the fourth street, then the set is straight. What is the probability that the set of coupons is straight?

Solution We start by assuming that all $\binom{52}{5}$ possible sets of coupons are equally likely. To determine the number of outcomes that are straights, let us first determine the number of possible outcomes for which the set of outcomes consists of a shop numbered 1, 2, 3, 4, and 5 (the streets being irrelevant). Since there are 4 shops numbered 1 (each on a different street), and similarly for the shop 2, 3, 4, and 5, it follows that there are 4^5 outcomes leading to exactly one shop 1, 2, 3, 4, and 5. Since, in 4 of these outcomes all the shops will be on the same street, it follows that there are $4^5 - 4$ sets that make up a straight of the form shop 1, 2, 3, 4, and 5. Similarly, there are $4^5 - 4$ sets that make up a straight of the form shop 9, 10, 11, 12, and 13. Thus, there are $9(4^5 - 4)$ sets that are straights, and it follows that the desired probability is

$$\frac{9(4^5 - 4)}{\binom{52}{5}} = \frac{9180}{2,598,960} \approx 0.003532. \qquad \blacksquare$$

Example 5g A company has 52 employees, four aged 20, four aged 21,..., and four aged 32. A committee is said to be perfect if it consists of 5 employees of two different age groups (say three aged 24 and 2 aged 31). A 5-member committee is selected randomly, what is the probability it is perfect?

Solution We assume that all $\binom{52}{5}$ possible committees are equally likely. To determine the number of possible perfect committees, we first note that there are $\binom{4}{2}\binom{4}{3}$ different combinations of, say, two 23 years old and three 27 years old members. Because there are 13 different choices for the kind of pair and, after a pair has been chosen, there are 12 other choices of the remaining 3 members, it follows that the probability of a committee consisting of two different age groups is

$$\frac{13 \times 12 \times \binom{4}{2} \times \binom{4}{3}}{\binom{52}{5}} \approx .0014. \qquad \blacksquare$$

Example 5h In the game of bridge, the entire deck of 52 cards is dealt out to 4 players. What is the probability that

(a) one of the players receives all 13 spades;
(b) each player receives 1 ace?

Solution (a) Letting E_i be the event that hand i has all 13 spades, then

$$P(E_i) = \frac{1}{\binom{52}{13}}, \quad i = 1, 2, 3, 4$$

Because the events E_i, $i = 1, 2, 3, 4$, are mutually exclusive, the probability that one of the hands is dealt all 13 spades is

$$P(\cup_{i=1}^{4} E_i) = \sum_{i=1}^{4} P(E_i) = 4\Big/\binom{52}{13} \approx 6.3 \times 10^{-12}$$

(b) Let the outcome of the experiment be the sets of 13 cards of each of the players 1, 2, 3, 4. To determine the number of outcomes in which each of the distinct players receives exactly 1 ace, put aside the aces and note that there are $\binom{48}{12, 12, 12, 12}$ possible divisions of the other 48 cards when each player is to receive 12. Because there are 4! ways of dividing the 4 aces so that each player receives 1, we see that the number of possible outcomes in which each player receives exactly 1 ace is $4!\binom{48}{12, 12, 12, 12}$.

As there are $\binom{52}{13,13,13,13}$ possible hands, the desired probability is thus

$$\frac{4!\binom{48}{12,12,12,12}}{\binom{52}{13,13,13,13}} \approx .1055$$ ■

Some results in probability are quite surprising when initially encountered. Our next two examples illustrate this phenomenon.

Example 5i

If n people are present in a room, what is the probability that no two of them celebrate their birthday on the same day of the year? How large need n be so that this probability is less than $\frac{1}{2}$?

Solution As each person can celebrate his or her birthday on any one of 365 days, there are a total of $(365)^n$ possible outcomes. (We are ignoring the possibility of someone having been born on February 29.) Assuming that each outcome is equally likely, we see that the desired probability is $(365)(364)(363)\ldots(365 - n + 1)/(365)^n$. It is a rather surprising fact that when $n \geq 23$, this probability is less than $\frac{1}{2}$. That is, if there are 23 or more people in a room, then the probability that at least two of them have the same birthday exceeds $\frac{1}{2}$. Many people are initially surprised by this result, since 23 seems so small in relation to 365, the number of days of the year. However, every pair of individuals has probability $\dfrac{365}{(365)^2} = \dfrac{1}{365}$ of having the same birthday, and in a group of 23 people, there are $\binom{23}{2} = 253$ different pairs of individuals. Looked at this way, the result no longer seems so surprising.

When there are 50 people in the room, the probability that at least two share the same birthday is approximately .970, and with 100 persons in the room, the odds are better than 3,000,000:1. (That is, the probability is greater than $\dfrac{3 \times 10^6}{3 \times 10^6 + 1}$ that at least two people have the same birthday.) ■

Example 5j

A deck of 52 playing cards is shuffled, and the cards are turned up one at a time until the first ace appears. Is the next card—that is, the card following the first ace—more likely to be the ace of spades or the two of clubs?

Solution To determine the probability that the card following the first ace is the ace of spades, we need to calculate how many of the (52)! possible orderings of the cards have the ace of spades immediately following the first ace. To begin, note that each ordering of the 52 cards can be obtained by first ordering the 51 cards different from the ace of spades and then inserting the ace of spades into that ordering. Furthermore, for each of the (51)! orderings of the other cards, there is only one place where the ace of spades can be placed so that it follows the first ace. For instance, if the ordering of the other 51 cards is

$$4c,\ 6h,\ Jd,\ 5s,\ Ac,\ 7d, \ldots, Kh$$

then the only insertion of the ace of spades into this ordering that results in its following the first ace is

$$4c,\ 6h,\ Jd,\ 5s,\ Ac,\ As,\ 7d, \ldots, Kh$$

Therefore, there are (51)! orderings that result in the ace of spades following the first ace, so

$$P\{\text{the ace of spades follows the first ace}\} = \frac{(51)!}{(52)!} = \frac{1}{52}$$

In fact, by exactly the same argument, it follows that the probability that the two of clubs (or any other specified card) follows the first ace is also $\frac{1}{52}$. In other words, each of the 52 cards of the deck is equally likely to be the one that follows the first ace!

Many people find this result rather surprising. Indeed, a common reaction is to suppose initially that it is more likely that the two of clubs (rather than the ace of spades) follows the first ace, since that first ace might itself be the ace of spades. This reaction is often followed by the realization that the two of clubs might itself appear before the first ace, thus negating its chance of immediately following the first ace. However, as there is one chance in four that the ace of spades will be the first ace (because all 4 aces are equally likely to be first) and only one chance in five that the two of clubs will appear before the first ace (because each of the set of 5 cards consisting of the two of clubs and the 4 aces is equally likely to be the first of this set to appear), it again appears that the two of clubs is more likely. However, this is not the case, and our more complete analysis shows that they are equally likely. ■

Example 5k

A football team consists of 20 offensive and 20 defensive players. The players are to be paired in groups of 2 for the purpose of determining roommates. If the pairing is done at random, what is the probability that there are no offensive–defensive roommate pairs? What is the probability that there are $2i$ offensive–defensive roommate pairs, $i = 1, 2, \ldots, 10$?

Solution There are

$$\binom{40}{2,2,\ldots,2} = \frac{(40)!}{(2!)^{20}}$$

ways of dividing the 40 players into 20 *ordered* pairs of two each. (That is, there are $(40)!/2^{20}$ ways of dividing the players into a *first* pair, a *second* pair, and so on.) Hence, there are $(40)!/2^{20}(20)!$ ways of dividing the players into (unordered) pairs of 2 each. Furthermore, since a division will result in no offensive–defensive pairs if the offensive (and defensive) players are paired among themselves, it follows that there are $[(20)!/2^{10}(10)!]^2$ such divisions. Hence, the probability of no offensive–defensive roommate pairs, call it P_0, is given by

$$P_0 = \frac{\left(\dfrac{(20)!}{2^{10}(10)!}\right)^2}{\dfrac{(40)!}{2^{20}(20)!}} = \frac{[(20)!]^3}{[(10)!]^2(40)!}$$

To determine P_{2i}, the probability that there are $2i$ offensive–defensive pairs, we first note that there are $\binom{20}{2i}^2$ ways of selecting the $2i$ offensive players and the $2i$ defensive players who are to be in the offensive–defensive pairs. These $4i$ players can then be paired up into $(2i)!$ possible offensive–defensive pairs. (This is so because the first offensive player can be paired with any of the $2i$ defensive players, the second offensive player with any of the remaining $2i - 1$ defensive players, and so on.) As the remaining $20 - 2i$ offensive (and defensive) players must be paired among themselves, it follows that there are

$$\binom{20}{2i}^2 (2i)!\left[\frac{(20 - 2i)!}{2^{10-i}(10 - i)!}\right]^2$$

divisions that lead to $2i$ offensive–defensive pairs. Hence,

$$P_{2i} = \frac{\binom{20}{2i}^2 (2i)!\left[\dfrac{(20 - 2i)!}{2^{10-i}(10 - i)!}\right]^2}{\dfrac{(40)!}{2^{20}(20)!}} \qquad i = 0, 1, \ldots, 10$$

The $P_{2i}, i = 0, 1, \ldots, 10$, can now be computed, or they can be approximated by making use of a result of Stirling, which shows that $n!$ can be approximated by $n^{n+1/2}e^{-n}\sqrt{2\pi}$. For instance, we obtain

$$P_0 \approx 1.3403 \times 10^{-6}$$
$$P_{10} \approx .345861$$
$$P_{20} \approx 7.6068 \times 10^{-6}$$

■

Our next three examples illustrate the usefulness of the inclusion–exclusion identity (Proposition 4.4). In Example 5l, the introduction of probability enables us to obtain a quick solution to a counting problem.

Example 5l

A total of 36 members of a club play tennis, 28 play squash, and 18 play badminton. Furthermore, 22 of the members play both tennis and squash, 12 play both tennis and badminton, 9 play both squash and badminton, and 4 play all three sports. How many members of this club play at least one of three sports?

Solution Let N denote the number of members of the club, and introduce probability by assuming that a member of the club is randomly selected. If, for any subset C of members of the club, we let $P(C)$ denote the probability that the selected member is contained in C, then

$$P(C) = \frac{\text{number of members in } C}{N}$$