Combinatorics

SECOND EDITION

RUSSELL MERRIS
California State University, Hayward

WILEY-
INTERSCIENCE
A JOHN WILEY & SONS, INC., PUBLICATION

Combinatorics

Second Edition

WILEY-INTERSCIENCE
SERIES IN DISCRETE MATHEMATICS AND OPTIMIZATION

A complete list of titles in this series appears at the end of this volume.

Combinatorics

SECOND EDITION

RUSSELL MERRIS
California State University, Hayward

WILEY-
INTERSCIENCE
A JOHN WILEY & SONS, INC., PUBLICATION

This book is printed on acid-free paper. ♾

Published simultaneously in Canada.

For ordering and customer service, call 1-800-CALL-WILEY.

Library of Congress Cataloging-in-Publication Data:

Merris, Russell, 1943–
Combinatorics / Russell Merris.–2nd ed.
p. cm. – (Wiley series in discrete mathematics and optimization)
Includes bibliographical references and index.
ISBN-0-471-26296-X (acid-free paper)
1. Combinatorial analysis I. Title. II. Series.

QA164.M47 2003
511′.6–dc21 2002192250

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

This book is dedicated to my wife, Karen Diehl Merris.

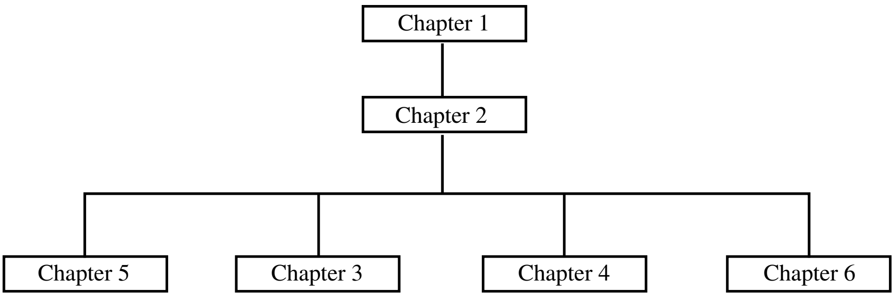
Contents

Note: Asterisks indicate optional sections that can be omitted without loss of continuity.

Preface

This book is intended to be used as the text for a course in combinatorics at the level of beginning upper division students. It has been shaped by two goals: to make some fairly deep mathematics accessible to students with a wide range of abilities, interests, and motivations and to create a pedagogical tool useful to the broad spectrum of instructors who bring a variety of perspectives and expectations to such a course.

The author's approach to the second goal has been to maximize flexibility. Following a basic foundation in Chapters 1 and 2, each instructor is free to pick and choose the most appropriate topics from the remaining four chapters. As summarized in the chart below, Chapters 3–6 are *completely independent of each other.* Flexibility is further enhanced by optional sections and appendices, by weaving some topics into the exercise sets of multiple sections, and by identifying various points of departure from each of the final four chapters. (The price of this flexibility is some redundancy, e.g., several definitions can be found in more than one place.)



Turning to the first goal, students using this book are expected to have been exposed to, even if they cannot recall them, such notions as equivalence relations, partial fractions, the Maclaurin series expansion for e^x, elementary row operations, determinants, and matrix inverses. A course designed around this book should have as specific prerequisites those portions of calculus and linear algebra commonly found among the lower division requirements for majors in the mathematical and computer sciences. Beyond these general prerequisites, the last two sections of Chapter 5 presume the reader to be familiar with the *definitions* of classical adjoint (adjugate) and characteristic roots (eigenvalues) of real matrices, and the first two sections of Chapter 6 make use of reduced row-echelon form, bases, dimension, rank, nullity, and orthogonality. (All of these topics are reviewed in Appendix A3.)

Strategies that promote student engagement are a lively writing style, timely and appropriate examples, interesting historical anecdotes, a variety of exercises (tempered and enlivened by suitable hints and answers), and judicious use of footnotes and appendices to touch on topics better suited to more advanced students. These are things about which there is general agreement, at least in principle.

There is less agreement about how to focus student energies on attainable objectives, in part because focusing on some things inevitably means neglecting others. If the course is approached as a *last chance* to expose students to this marvelous subject, it probably will be. If approached more invitingly, as a *first* course in combinatorics, it may be. To give some specific examples, highlighted in this book are binomial coefficients, Stirling numbers, Bell numbers, and partition numbers. These topics appear and reappear throughout the text. Beyond reinforcement in the service of retention, the tactic of overarching themes helps foster an image of combinatorics as a unified mathematical discipline. While other celebrated examples, e.g., Bernoulli numbers, Catalan numbers, and Fibonacci numbers, are generously represented, they appear almost entirely in the exercises. For the sake of argument, let us stipulate that these roles could just as well have been reversed. The issue is that beginning upper division students cannot be expected to absorb, much less appreciate, *all* of these special arrays and sequences in a single semester. On the other hand, the flexibility is there for willing admirers to rescue one or more of these justly famous combinatorial sequences from the relative obscurity of the exercises.

While the overall framework of the first edition has been retained, everything else has been revised, corrected, smoothed, or polished. The focus of many sections has been clarified, e.g., by eliminating peripheral topics or moving them to the exercises. Material new to the second edition includes an optional section on algorithms, several new examples, and many new exercises, some designed to guide students to discover and prove nontrivial results for themselves. Finally, the section of hints and answers has been expanded by an order of magnitude.

The material in Chapter 3, Pólya's theory of enumeration, is typically found closer to the end of comparable books, perhaps reflecting the notion that it is the *last* thing that should be taught in a junior-level course. The author has aspired, not only to make this theory accessible to students taking a first upper division mathematics course, but to make it possible for the subject to be addressed right after Chapter 2. Its placement in the middle of the book is intended to signal that it *can* be fitted in there, not that it must be. If it seems desirable to cover some but not all of Chapter 3, there are many natural places to exit in favor of something else, e.g., after the application of Bell numbers to transitivity in Section 3.3, after enumerating the overall number of color patterns in Section 3.5, after *stating* Pólya's theorem in Section 3.6, or after proving the theorem at the end of Section 3.6.

Optional Sections 1.3 and 1.10 can be omitted with the understanding that exercises in subsequent sections involving probability or algorithms should be assigned with discretion. With the same caveat, Section 1.4 can be omitted by those not intending to go on to Sections 6.1, 6.2, or 6.4. The material in Section 6.3, touching on mutually orthogonal Latin squares and their connection to finite projective planes, can be covered independently of Sections 1.4, 6.1, and 6.2.

The book contains much more material then can be covered in a single semester. Among the possible syllabi for a one semester course are the following:

- Chapters 1, 2, and 4 and Sections 3.1–3.3
- Chapters 1 (omitting Sections 1.3, 1.4, & 1.10), 2, and 3, and Sections 5.1 & 5.2
- Chapters 1 (omitting Sections 1.3 & 1.10), 2, and 6 and Sections 4.1 – 4.4
- Chapters 1 (omitting Sections 1.4 & 1.10) and 2 and Sections 3.1 – 3.3, 4.1 – 4.3, & 6.3
- Chapters 1 (omitting Sections 1.3 & 1.4) and 2 and Sections 4.1 – 4.3, 5.1, & 5.3–5.7
- Chapters 1 (omitting Sections 1.3, 1.4, & 1.10) and 2 and Sections 4.1 – 4.3, 5.1, 5.3–5.5, & 6.3

Many people have contributed observations, suggestions, corrections, and constructive criticisms at various stages of this project. Among those deserving special mention are former students David Abad, Darryl Allen, Steve Baldzikowski, Dale Baxley, Stanley Cheuk, Marla Dresch, Dane Franchi, Philip Horowitz, Rhian Merris, Todd Mullanix, Cedide Olcay, Glenn Orr, Hitesh Patel, Margaret Slack, Rob Smedfjeld, and Masahiro Yamaguchi; sometime collaborators Bob Grone, Tom Roby, and Bill Watkins; correspondents Mark Hunacek and Gerhard Ringel; reviewers Rob Beezer, John Emert, Myron Hood, Herbert Kasube, André Kézdy, Charles Landraitis, John Lawlor, and Wiley editors Heather Bergman, Christine Punzo, and Steve Quigley. I am especially grateful for the tireless assistance of Cynthia Johnson and Ken Rebman.

Despite everyone's best intentions, no book seems complete without some errors. An up-to-date errata, accessible from the Internet, will be maintained at URL

http://www.sci.csuhayward.edu/∼rmerris

Appropriate acknowledgment will be extended to the first person who communicates the specifics of a previously unlisted error to the author, preferably by e-mail addressed to

merris@csuhayward.edu

Hayward California RUSSELL MERRIS

1

The Mathematics of Choice

> It seems that mathematical ideas are arranged somehow in strata, the ideas in each stratum being linked by a complex of relations both among themselves and with those above and below. The lower the stratum, the deeper (and in general the more difficult) the idea. Thus, the idea of an irrational is deeper than the idea of an integer.
>
> — G. H. Hardy (*A Mathematician's Apology*)

Roughly speaking, the first chapter of this book is the top stratum, the surface layer of combinatorics. Even so, it is far from superficial. While the first main result, the so-called fundamental counting principle, is nearly self-evident, it has enormous implications throughout combinatorial enumeration. In the version presented here, one is faced with a sequence of decisions, each of which involves some number of choices. It is from situations like this that the chapter derives its name.

To the uninitiated, mathematics may appear to be "just so many numbers and formulas." In fact, the numbers and formulas should be regarded as shorthand notes, summarizing *ideas*. Some ideas from the first section are summarized by an algebraic formula for multinomial coefficients. Special cases of these numbers are addressed from a combinatorial perspective in Section 1.2.

Section 1.3 is an optional discussion of probability theory which can be omitted if probabilistic exercises in subsequent sections are approached with caution. Section 1.4 is an optional excursion into the theory of binary codes which can be omitted by those not planning to visit Chapter 6. Sections 1.3 and 1.4 are partly motivational, illustrating that even the most basic combinatorial ideas have real-life applications.

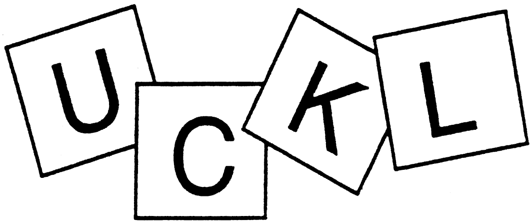
In Section 1.5, ideas behind the formulas for sums of powers of positive integers motivate the study of relations among binomial coefficients. Choice is again the topic in Section 1.6, this time with or without replacement, where order does or doesn't matter.

To better organize and understand the multinomial theorem from Section 1.7, one is led to symmetric polynomials and, in Section 1.8, to partitions of n. Elementary symmetric functions and their association with power sums lie at the

Combinatorics, *Second Edition*, by Russell Merris.
ISBN 0-471-26296-X

heart of Section 1.9. The final section of the chapter is an optional introduction to algorithms, the flavor of which can be sampled by venturing only as far as Algorithm 1.10.3. Those desiring not less but more attention to algorithms can find it in Appendix A2.

1.1. THE FUNDAMENTAL COUNTING PRINCIPLE



How many different four-letter words, including nonsense words, can be produced by rearranging the letters in LUCK? In the absence of a more inspired approach, there is always the brute-force strategy: Make a systematic list.

Once we become convinced that Fig. 1.1.1 accounts for every possible rearrangement and that no "word" is listed twice, the solution is obtained by counting the 24 words on the list.

While finding the brute-force strategy was effortless, implementing it required some work. Such an approach may be fine for an isolated problem, the *like* of which one does not expect to see again. But, just for the sake of argument, imagine yourself in the situation of having to solve a great many thinly disguised variations of this same problem. In that case, it would make sense to invest some effort in finding a strategy that requires less work to implement. Among the most powerful tools in this regard is the following commonsense principle.

1.1.1 Fundamental Counting Principle. Consider a (finite) sequence of decisions. Suppose the number of choices for each individual decision is independent of decisions made previously in the sequence. Then the number of ways to make the whole sequence of decisions is the product of these numbers of choices.

To state the principle symbolically, suppose c_i is the number of choices for decision i. If, for $1 \le i < n$, c_{i+1} does not depend on which choices are made in

LUCK	LUKC	LCUK	LCKU	LKUC	LKCU
ULCK	ULKC	UCLK	UCKL	UKLC	UKCL
CLUK	CLKU	CULK	CUKL	CKLU	CKUL
KLUC	KLCU	KULC	KUCL	KCLU	KCUL

Figure 1.1.1. The rearrangements of LUCK.

decisions $1, \ldots, i$, then the number of different ways to make the sequence of decisions is $c_1 \times c_2 \times \cdots \times c_n$.

Let's apply this principle to the word problem we just solved. Imagine yourself in the midst of making the brute-force list. Writing down one of the words involves a sequence of four decisions. Decision 1 is which of the four letters to write first, so $c_1 = 4$. (It is no accident that Fig. 1.1.1 consists of four rows!) For each way of making decision 1, there are $c_2 = 3$ choices for decision 2, namely which letter to write second. Notice that the specific letters comprising these three choices depend on how decision 1 was made, but their *number* does not. That is what is meant by the number of choices for decision 2 being independent of how the previous decision is made. Of course, $c_3 = 2$, but what about c_4? Facing no alternative, is it correct to say there is "no choice" for the last decision? If that were literally true, then c_4 would be zero. In fact, $c_4 = 1$. So, by the fundamental counting principle, the number of ways to make the sequence of decisions, i.e., the number of words on the final list, is

$$c_1 \times c_2 \times c_3 \times c_4 = 4 \times 3 \times 2 \times 1.$$

The product $n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$ is commonly written n! and read *n-factorial*.* The number of four-letter words that can be made up by rearranging the letters in the word LUCK is $4! = 24$.

What if the word had been LUCKY? The number of five-letter words that can be produced by rearranging the letters of the word LUCKY is $5! = 120$. A systematic list might consist of five rows each containing $4! = 24$ words.

Suppose the word had been LOOT? How many four-letter words, including nonsense words, can be constructed by rearranging the letters in LOOT? Why not apply the fundamental counting principle? Once again, imagine yourself in the midst of making a brute-force list. Writing down one of the words involves a sequence of four decisions. Decision 1 is which of the three letters L, O, or T to write first. This time, $c_1 = 3$. But, what about c_2? In this case, the number of choices for decision 2 depends on how decision 1 was made! If, e.g., *L* were chosen to be the first letter, then there would be two choices for the second letter, namely O or T. If, however, O were chosen first, then there would be three choices for the second decision, L, (the second) O, or T. Do we take $c_2 = 2$ or $c_2 = 3$? The answer is that *the fundamental counting principle does not apply to this problem* (at least not directly). The fundamental counting principle applies *only* when the *number* of choices for decision $i + 1$ is *independent* of how the previous i decisions are made.

To enumerate all possible rearrangements of the letters in LOOT, begin by distinguishing the two O's. maybe write the word as LOoT. Applying the fundamental counting principle, we find that there are $4! = 24$ different-*looking* four-letter words that can be made up from L, O, o, and T.

*The exclamation mark is used, not for emphasis, but because it is a convenient symbol common to most keyboards.

LOoT	LOTo	LoOT	LoTO	LTOo	LToO
OLoT	OLTo	OoLT	OoTL	OTLo	OToL
oLOT	oLTO	oOLT	oOTL	oTLO	oTOL
TLOo	TLoO	TOLo	TOoL	ToLO	ToOL

Figure 1.1.2. Rearrangements of LOoT.

Among the words in Fig. 1.1.2 are pairs like OLoT and oLOT, which look different only because the two O's have been distinguished. In fact, every word in the list occurs twice, once with "big O" coming before "little o", and once the other way around. Evidently, the number of different words (with indistinguishable O's) that can be produced from the letters in LOOT is not 4! but $4!/2 = 12$.

What about TOOT? First write it as TOot. Deduce that in any list of all possible rearrangements of the letters T, O, o, and t, there would be $4! = 24$ different-looking words. Dividing by 2 makes up for the fact that two of the letters are O's. Dividing by 2 again makes up for the two T's. The result, $24/(2 \times 2) = 6$, is the number of different words that can be made up by rearranging the letters in TOOT. Here they are

TTOO TOTO TOOT OTTO OTOT OOTT

All right, what if the word had been LULL? How many words can be produced by rearranging the letters in LULL? Is it too early to guess a pattern? Could the number we're looking for be $4!/3 = 8$? No. It is easy to see that the correct answer must be 4. Once the position of the letter U is known, the word is completely determined. Every other position is filled with an L. A complete list is ULLL, LULL, LLUL, LLLU.

To find out why 4!/3 is wrong, let's proceed as we did before. Begin by distinguishing the three L's, say L_1, L_2, and L_3. There are 4! different-looking words that can be made up by rearranging the four letters L_1, L_2, L_3, and U. If we were to make a list of these 24 words and then erase all the subscripts, how many times would, say, LLLU appear? The answer to this question can be obtained from the fundamental counting principle! There are three decisions: decision 1 has three choices, namely which of the three L's to write first. There are two choices for decision 2 (which of the two remaining L's to write second) and one choice for the third decision, which L to put last. Once the subscripts are erased, LLLU would appear 3! times on the list. We should divide $4! = 24$, not by 3, but by $3! = 6$. Indeed, $4!/3! = 4$ is the correct answer.

Whoops! if the answer corresponding to LULL is 4!/3!, why didn't we get 4!/2! for the answer to LOOT? In fact, we did: $2! = 2$.

Are you ready for MISSISSIPPI? It's the same problem! If the letters were all different, the answer would be 11!. Dividing 11! by 4! makes up for the fact that there are four I's. Dividing the quotient by another 4! compensates for the four S's.

Dividing that quotient by 2! makes up for the two P's. In fact, no harm is done if that quotient is divided by $1! = 1$ in honor of the single M. The result is

$$\frac{11!}{4!\,4!\,2!\,1!} = 34,650.$$

(Confirm the arithmetic.) The 11 letters in MISSISSIPPI can be (re)arranged in 34,650 different ways.*

There is a special notation that summarizes the solution to what we might call the "MISSISSIPPI problem."

1.1.2 Definition. The *multinomial coefficient*

$$\binom{n}{r_1, r_2, \ldots, r_k} = \frac{n!}{r_1!r_2!\cdots r_k!},$$

where $r_1 + r_2 + \cdots + r_k = n$.

So, "multinomial coefficient" is a *name* for the answer to the question, how many n-letter "words" can be assembled using r_1 copies of one letter, r_2 copies of a second (different) letter, r_3 copies of a third letter, ..., and r_k copies of a kth letter?

1.1.3 Example. After cancellation,

$$\binom{9}{4,3,1,1} = \frac{9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{4 \times 3 \times 2 \times 1 \times 3 \times 2 \times 1 \times 1 \times 1}$$
$$= 9 \times 8 \times 7 \times 5 = 2520.$$

Therefore, 2520 different words can be manufactured by rearranging the nine letters in the word SASSAFRAS. □

In real-life applications, the words need not be assembled from the English alphabet. Consider, e.g., POSTNET† barcodes commonly attached to U.S. mail by the Postal Service. In this scheme, various numerical delivery codes‡ are represented by "words" whose letters, or *bits*, come from the alphabet $\{\text{I}, \big|\}$. Corresponding, e.g., to a ZIP + 4 code is a 52-bit barcode that begins and ends with $\big|$. The 50-bit middle part is partitioned into ten 5-bit zones. The first nine of these zones are for the digits that comprise the ZIP + 4 code. The last zone accommodates a *parity*

* This number is roughly equal to the number of members of the Mathematical Association of America (MAA), the largest professional organization for mathematicians in the United States.

† Postal Numeric Encoding Technique.

‡ The original five-digit Zoning Improvement Plan (ZIP) code was introduced in 1964; ZIP+4 codes followed about 25 years later. The 11-digit Delivery Point Barcode (DPBC) is a more recent variation.

Figure 1.1.3. POSTNET barcodes.

check digit, chosen so that the sum of all ten digits is a multiple of 10. Finally, each digit is represented by one of the 5-bit barcodes in Fig. 1.1.3. Consider, e.g., the ZIP +4 code 20090-0973, for the Mathematical Association of America. Because the sum of these digits is 30, the parity check digit is 0. The corresponding 52-bit word can be found in Fig. 1.1.4.

20090-0973

Figure 1.1.4

We conclude this section with another application of the fundamental counting principle.

1.1.4 Example. Suppose you wanted to determine the number of positive integers that exactly divide $n = 12$. That isn't much of a problem; there are six of them, namely, 1, 2, 3, 4, 6, and 12. What about the analogous problem for $n = 360$ or for $n = 360,000$? Solving even the first of these by brute-force list making would be a lot of work. Having already found another strategy whose implementation requires a lot less work, let's take advantage of it.

Consider $360 = 2^3 \times 3^2 \times 5$, for example. If $360 = dq$ for positive integers d and q, then, by the uniqueness part of the *fundamental theorem of arithmetic*, the prime factors of d, together with the prime factors of q, are precisely the prime factors of 360, multiplicities included. It follows that the prime factorization of d must be of the form $d = 2^a \times 3^b \times 5^c$, where $0 \leq a \leq 3$, $0 \leq b \leq 2$, and $0 \leq c \leq 1$. Evidently, there are four choices for a (namely 0, 1, 2, or 3), three choices for b, and two choices for c. So, the number of possibile d's is $4 \times 3 \times 2 = 24$. □

1.1. EXERCISES

1 The Hawaiian alphabet consists of 12 letters, the vowels *a*, *e*, *i*, *o*, *u* and the consonants *h*, *k*, *l*, *m*, *n*, *p*, *w*.

(a) Show that 20,736 different 4-letter "words" could be constructed using the 12-letter Hawaiian alphabet.

(b) Show that 456,976 different 4-letter "words" could be produced using the 26-letter English alphabet.*

(c) How many four-letter "words" can be assembled using the Hawaiian alphabet if the second and last letters are vowels and the other 2 are consonants?

(d) How many four-letter "words" can be produced from the Hawaiian alphabet if the second and last letters are vowels but there are no restrictions on the other 2 letters?

2 Show that

(a) $3! \times 5! = 6!$.

(b) $6! \times 7! = 10!$.

(c) $(n+1) \times (n!) = (n+1)!$.

(d) $n^2 = n![1/(n-1)! + 1/(n-2)!]$.

(e) $n^3 = n![1/(n-1)! + 3/(n-2)! + 1/(n-3)!]$.

3 One brand of electric garage door opener permits the owner to select his or her own electronic "combination" by setting six different switches either in the "up" or the "down" position. How many different combinations are possible?

4 One generation back you have two ancestors, your (biological) parents. Two generations back you have four ancestors, your grandparents. Estimating 2^{10} as 10^3, approximately how many ancestors do you have

(a) 20 generations back?

(b) 40 generations back?

(c) In round numbers, what do you estimate is the total population of the planet?

(d) What's wrong?

5 Make a list of all the "words" that can be made up by rearranging the letters in

(a) TO. **(b)** TOO. **(c)** TWO.

6 Evaluate multinomial coefficient

(a) $\binom{6}{4,1,1}$. **(b)** $\binom{6}{3,3}$. **(c)** $\binom{6}{2,2,2}$.

*Based on these calculations, might it be reasonable to expect Hawaiian words, on average, to be longer than their English counterparts? Certainly such a conclusion would be warranted if both languages had the same vocabulary and both were equally "efficient" in avoiding long words when short ones are available. How efficient is English? Given that the total number of words defined in a typical "unabridged dictionary" is at most 350,000, one could, at least in principle, construct a new language with the same vocabulary as English but in which every word has four letters—and there would be 100,000 words to spare!

(d) $\binom{6}{3,2,1}$. **(e)** $\binom{6}{1,3,2}$. **(f)** $\binom{6}{1,1,1,1,1,1}$.

7 How many different "words" can be constructed by rearranging the letters in

(a) ALLELE? **(b)** BANANA? **(c)** PAPAYA?

(d) BUBBLE? **(e)** ALABAMA? **(f)** TENNESSEE?

(g) HALEAKALA? **(h)** KAMEHAMEHA? **(i)** MATHEMATICS?

8 Prove that

(a) $1 + 2 + 2^2 + 2^3 + \cdots + 2^n = 2^{n+1} - 1$.

(b) $1 \times 1! + 2 \times 2! + 3 \times 3! + \cdots + n \times n! = (n+1)! - 1$.

(c) $(2n)!/2^n$ is an integer.

9 Show that the barcodes in Fig. 1.1.3 comprise *all possible* five-letter words consisting of two |'s and three ı's.

10 Explain how the following barcodes fail the POSTNET standard:

(a)

(b)

(c)

11 "Read" the ZIP+4 Code

(a)

(b)

12 Given that the first nine zones correspond to the ZIP+4 delivery code 94542-2520, determine the parity check digit and the two "hidden digits" in the 62-bit DPBC

(*Hint:* Do you need to be told that the parity check digit is last?)

13 Write out the 52-bit POSTNET barcode for 20742-2461, the ZIP+4 code at the University of Maryland used by the Association for Women in Mathematics.

14 Write out all 24 divisors of 360. (See Example 1.1.4.)

15 Compute the number of positive integer divisors of

(a) 2^{10}. **(b)** 10^{10}. **(c)** 12^{10}. **(d)** 31^{10}.

(e) 360,000. **(f)** 10!.

16 Prove that the positive integer n has an odd number of positive-integer divisors if and only if it is a perfect square.

17 Let $D = \{d_1, d_2, d_3, d_4\}$ and $R = \{r_1, r_2, r_3, r_4, r_5, r_6\}$. Compute the number

(a) of different functions $f : D \to R$.

(b) of one-to-one functions $f : D \to R$.

18 The latest automobile license plates issued by the California Department of Motor Vehicles begin with a single numeric digit, followed by three letters, followed by three more digits. How many different license "numbers" are available using this scheme?

19 One brand of padlocks uses combinations consisting of three (not necessarily different) numbers chosen from $\{0, 1, 2, \ldots, 39\}$. If it takes five seconds to "dial in" a three-number combination, how long would it take to try all possible combinations?

20 The *International Standard Book Number* (ISBN) is a 10-digit numerical code for identifying books. The groupings of the digits (by means of hyphens) varies from one book to another. The first grouping indicates where the book was published. In ISBN 0-88175-083-2, the zero shows that the book was published in the English-speaking world. The code for the Netherlands is "90" as, e.g., in ISBN 90-5699-078-0. Like POSTNET, ISBN employs a check digit scheme. The first nine digits (ignoring hyphens) are multiplied, respectively, by 10, 9, 8, …, 2, and the resulting products summed to obtain S. In 0-88175-083-2, e.g.,

$$S = 10 \times 0 + 9 \times 8 + 8 \times 8 + 7 \times 1 + 6 \times 7 + 5 \times 5 + 4 \times 0 \\ + 3 \times 8 + 2 \times 3 = 240.$$

The last (check) digit, L, is chosen so that $S + L$ is a multiple of 11. (In our example, $L = 2$ and $S + L = 242 = 11 \times 22$.)

(a) Show that, when S is divided by 11, the quotient Q and remainder R satisfy $S = 11Q + R$.

(b) Show that $L = 11 - R$. (When $R = 1$, the check digit is X.)

(c) What is the value of the check digit, L, in ISBN 0-534-95154-L?

(d) Unlike POSTNET, the more sophisticated ISBN system can not only detect common errors, it can sometimes "correct" them. Suppose, e.g., that a single digit is wrong in ISBN 90-5599-078-0. Assuming the check digit is correct, can you identify the position of the erroneous digit?

(e) Now that you know the position of the (single) erroneous digit in part (d), can you recover the correct ISBN?

(f) What if it were expected that exactly two digits were wrong in part (d). Which two digits might they be?

21 A total of $9! = 362,880$ different nine-letter "words" can be produced by rearranging the letters in FULBRIGHT. Of these, how many contain the four-letter sequence GRIT?

22 In how many different ways can eight coins be arranged on an 8×8 checkerboard so that no two coins lie in the same row or column?

23 If A is a finite set, its *cardinality*, $o(A)$, is the number of elements in A. Compute

(a) $o(A)$ when A is the set consisting of all five-digit integers, each digit of which is 1, 2, or 3.

(b) $o(B)$, where $B = \{x \in A : \text{each of } 1, 2, \text{and } 3 \text{ is among the digits of } x\}$ and A is the set in part (a).

1.2. PASCAL'S TRIANGLE

Mathematics is the art of giving the same name to different things.

— Henri Poincaré (1854–1912)

In how many different ways can an r-element subset be chosen from an n-element set S? Denote the number by $C(n, r)$. Pronounced "n-choose-r", $C(n, r)$ is just a name for the answer. Let's find the number represented by this name.

Some facts about $C(n, r)$ are clear right away, e.g., the nature of the elements of S is immaterial. All that matters is that there are n of them. Because the only way to choose an n-element subset from S is to choose all of its elements, $C(n, n) = 1$. Having n single elements, S has n single-element subsets, i.e., $C(n, 1) = n$. For essentially the same reason, $C(n, n-1) = n$: A subset of S that contains all but one element is uniquely determined by the one element that is left out. Indeed, this idea has a nice generalization. A subset of S that contains all but r elements is uniquely determined by the r elements that are left out. This natural one-to-one correspondence between subsets and their complements yields the following *symmetry property*:

$$C(n, n-r) = C(n, r).$$

1.2.1 Example. By definition, there are $C(5, 2)$ ways to select two elements from $\{A, B, C, D, E\}$. One of these corresponds to the two-element subset $\{A, B\}$. The complement of $\{A, B\}$ is $\{C, D, E\}$. This pair is listed first in the following one-to-one correspondence between two-element subsets and their three-element complements:

$$\begin{array}{ll} \{A,B\} \leftrightarrow \{C,D,E\}, & \{B,D\} \leftrightarrow \{A,C,E\}; \\ \{A,C\} \leftrightarrow \{B,D,E\}, & \{B,E\} \leftrightarrow \{A,C,D\}; \\ \{A,D\} \leftrightarrow \{B,C,E\}, & \{C,D\} \leftrightarrow \{A,B,E\}; \\ \{A,E\} \leftrightarrow \{B,C,D\}, & \{C,E\} \leftrightarrow \{A,B,D\}; \\ \{B,C\} \leftrightarrow \{A,D,E\}, & \{D,E\} \leftrightarrow \{A,B,C\}. \end{array}$$

By counting these pairs, we find that $C(5,2) = C(5,3) = 10$. □

A special case of symmetry is $C(n,0) = C(n,n) = 1$. Given n objects, there is just one way to reject all of them and, hence, just one way to choose none of them. What if $n = 0$? How many ways are there to choose no elements from the empty set? To avoid a deep philosophical discussion, let us simply adopt as a convention that $C(0,0) = 1$.

A less obvious fact about choosing these numbers is the following.

1.2.2 Theorem (Pascal's Relation). *If* $1 \le r \le n$, *then*

$$C(n+1,r) = C(n,r-1) + C(n,r). \tag{1.1}$$

Together with Example 1.2.1, Pascal's relation implies, e.g., that $C(6,3) = C(5,2) + C(5,3) = 20$.

Proof. Consider the $(n+1)$-element set $\{x_1, x_2, \ldots, x_n, y\}$. Its r-element subsets can be partitioned into two families, those that contain y and those that do not. To count the subsets that contain y, simply observe that the remaining $r-1$ elements can be chosen from $\{x_1, x_2, \ldots, x_n\}$ in $C(n,r-1)$ ways. The r-element subsets that do not contain y are precisely the r-element subsets of $\{x_1, x_2, \ldots, x_n\}$, of which there are $C(n,r)$. ■

The proof of Theorem 1.2.2 used another self-evident fact that is worth mentioning explicitly. (A much deeper extension of this result will be discussed in Chapter 2.)

1.2.3 The Second Counting Principle. If a set can be expressed as the disjoint union of two (or more) subsets, then the number of elements in the set is the sum of the numbers of elements in the subsets.

So far, we have been viewing $C(n,r)$ as a single number. There are some advantages to looking at these choosing numbers collectively, as in Fig. 1.2.1. The triangular shape of this array is a consequence of not bothering to write $0 = C(n,r)$, $r > n$. Filling in the entries we know, i.e., $C(n,0) = C(n,n) = 1$, $C(n,1) = n = C(n,n-1)$, $C(5,2) = C(5,3) = 10$, and $C(6,3) = 20$, we obtain Fig. 1.2.2.

$n \backslash r$	0	1	2	3	4	5	6	7
0	$C(0,0)$							
1	$C(1,0)$	$C(1,1)$						
2	$C(2,0)$	$C(2,1)$	$C(2,2)$					
3	$C(3,0)$	$C(3,1)$	$C(3,2)$	$C(3,3)$				
4	$C(4,0)$	$C(4,1)$	$C(4,2)$	$C(4,3)$	$C(4,4)$			
5	$C(5,0)$	$C(5,1)$	$C(5,2)$	$C(5,3)$	$C(5,4)$	$C(5,5)$		
6	$C(6,0)$	$C(6,1)$	$C(6,2)$	$C(6,3)$	$C(6,4)$	$C(6,5)$	$C(6,6)$	
7	$C(7,0)$	$C(7,1)$	$C(7,2)$	$C(7,3)$	$C(7,4)$	$C(7,5)$	$C(7,6)$	$C(7,7)$
				...				

Figure 1.2.1. $C(n, r)$.

Given the fourth row of the array (corresponding to $n = 3$), we can use Pascal's relation to compute $C(4, 2) = C(3, 1) + C(3, 2) = 3 + 3 = 6$. Similarly, $C(6, 4) = C(6, 2) = C(5, 1) + C(5, 2) = 5 + 10 = 15$. Continuing in this way, one row at a time, we can complete as much of the array as we like.

$n \backslash r$	0	1	2	3	4	5	6	7
0	1							
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	$C(4,2)$	4	1			
5	1	5	10	10	5	1		
6	1	6	$C(6,2)$	20	$C(6,4)$	6	1	
7	1	7	$C(7,2)$	$C(7,3)$	$C(7,4)$	$C(7,5)$	7	1
				...				

Figure 1.2.2

Following Western tradition, we refer to the array in Fig. 1.2.3 as *Pascal's triangle.** (Take care not to forget, e.g., that $C(6, 3) = 20$ appears, not in the third column of the sixth row, but in the fourth column of the seventh!)

Pascal's triangle is the source of many interesting identities. One of these concerns the sum of the entries in each row:

$$\begin{aligned} 1 + 1 &= 2, \\ 1 + 2 + 1 &= 4, \\ 1 + 3 + 3 + 1 &= 8, \\ 1 + 4 + 6 + 4 + 1 &= 16, \end{aligned} \tag{1.2}$$

*After Blaise Pascal (1623–1662), who described it in the book *Traité du triangle arithmétique*. Rumored to have been included in a lost mathematical work by Omar Khayyam (ca. 1050–1130), author of the *Rubaiyat*, the triangle is also found in surviving works by the Arab astronomer al-Tusi (1265), the Chinese mathematician Chu Shih-Chieh (1303), and the Hindu writer Narayana Pandita (1365). The first European author to mention it was Petrus Apianus (1495–1552), who put it on the title page of his 1527 book, *Rechnung*.

$n \backslash r$	0	1	2	3	4	5	6	7
0	1							
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
5	1	5	10	10	5	1		
6	1	6	15	20	15	6	1	
7	1	7	21	35	35	21	7	1
				...				

Figure 1.2.3. Pascal's triangle.

and so on. Why should each row sum to a power of 2? In

$$C(n,0) + C(n,1) + \cdots + C(n,n) = \sum_{r=0}^{n} C(n,r),$$

$C(n,0)$ is the number of subsets of $S = \{x_1, x_2, \ldots, x_n\}$ that have no elements; $C(n,1)$ is the number of one-element subsets of S; $C(n,2)$ is the number of two-element subsets, and so on. Evidently, the sum of the numbers in row n of Pascal's triangle is the total number of subsets of S (even when $n = 0$ and $S = \emptyset$). The empirical evidence from Equations (1.2) suggests that an n-element set has a total of 2^n subsets. How might one go about proving this conjecture?

One way to do it is by mathematical induction. There is, however, another approach that is both easier and more revealing. Imagine youself in the process of listing the subsets of $S = \{x_1, x_2, \ldots, x_n\}$. Specifying a subset involves a sequence of decisions. Decision 1 is whether to include x_1. There are two choices, *Yes* or *No*. Decision 2, whether to put x_2 into the subset, also has two choices. Indeed, there are two choices for each of the n decisions. So, by the fundamental counting principle, S has a total of $2 \times 2 \times \cdots \times 2 = 2^n$ subsets.

There is more. Suppose, for example, that $n = 9$. Consider the sequence of decisions that produces the subset $\{x_2, x_3, x_6, x_8\}$, a sequence that might be recorded as NYYNNYNYN. The first letter of this word corresponds to *No*, as in "no to x_1"; the second letter corresponds to *Yes*, as in "yes to x_2"; because x_3 is in the subset, the third letter is Y; and so on for each of the nine letters. Similarly, $\{x_1, x_2, x_3\}$ corresponds to the nine-leter word YYYNNNNNN. In general, there is a one-to-one correspondence between subsets of $\{x_1, x_2, \ldots, x_n\}$, and n-letter words assembled from the alphabet $\{\mathrm{N}, \mathrm{Y}\}$. Moreover, in this correspondence, r-element subsets correspond to words with r Y's and $n - r$ N's.

We seem to have discovered a new way to think about $C(n,r)$. It is the number of n-letter words that can be produced by (re)arranging r Y's and $n - r$ N's. This interpretation can be verified directly. An n-letter word consists of n spaces, or locations, occupied by letters. Each of the words we are discussing is completely determined once the r locations of the Y's have been chosen (the remaining $n - r$ spaces being occupied by N's).

The significance of this new perspective is that we know how to count the number of n-letter words with r Y's and $n - r$ N's. That's the MISSISSIPPI problem! The answer is multinomial coefficient $\binom{n}{r, n-r}$. Evidently,

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

For things to work out properly when $r = 0$ and $r = n$, we need to adopt another convention. Define $0! = 1$. (So, $0!$ is *not* equal to the nonsensical $0 \times (0-1) \times (0-2) \times \cdots \times 1$.)

It is common in the mathematical literature to write $\binom{n}{r}$ instead of $\binom{n}{r, n-r}$, one justification being that the information conveyed by "$n - r$" is redundant. It can be computed from n and r. The same thing could, of course, be said about *any* multinomial coefficient. The last number in the second row is always redundant. So, that particular argument is not especially compelling. The honest reason for writing $\binom{n}{r}$ is tradition.

We now have two ways to look at $C(n, r) = \binom{n}{r}$. One is what we might call the *combinatorial definition*: n-choose-r is the number of ways to choose r things from a collection of n things. The alternative, what we might call the *algebraic definition*, is

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

Don't make the mistake of asuming, just because it is more familiar, that the algebraic definition will always be easiest. (Try giving an algebraic proof of the identity $\sum_{r=0}^{n} C(n, r) = 2^n$.) Some applications are easier to approach using algebraic methods, while the combinatorial definition is easier for others. Only by becoming familiar with both will you be in a position to choose the easiest approach in every situation!

1.2.4 Example. In the basic version of poker, each player is dealt five cards (as in Fig. 1.2.4) from a standard 52-card deck (no joker). How many different five-card poker hands are there? Because someone (in a fair game it might be *Lady Luck*) chooses five cards from the deck, the answer is $C(52, 5)$. The ways to find the number behind this name are: (1) Make an exhaustive list of all possible hands, (2) work out 52 rows of Pascal's triangle, or (3) use the algebraic definition

$$\begin{aligned} C(52, 5) &= \frac{52!}{5!\,47!} \\ &= \frac{52 \times 51 \times 50 \times 49 \times 48 \times 47!}{5 \times 4 \times 3 \times 2 \times 1 \times 47!} \\ &= \frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2 \times 1} \\ &= 52 \times 51 \times 10 \times 49 \times 2 \\ &= 2{,}598{,}960. \end{aligned}$$

□

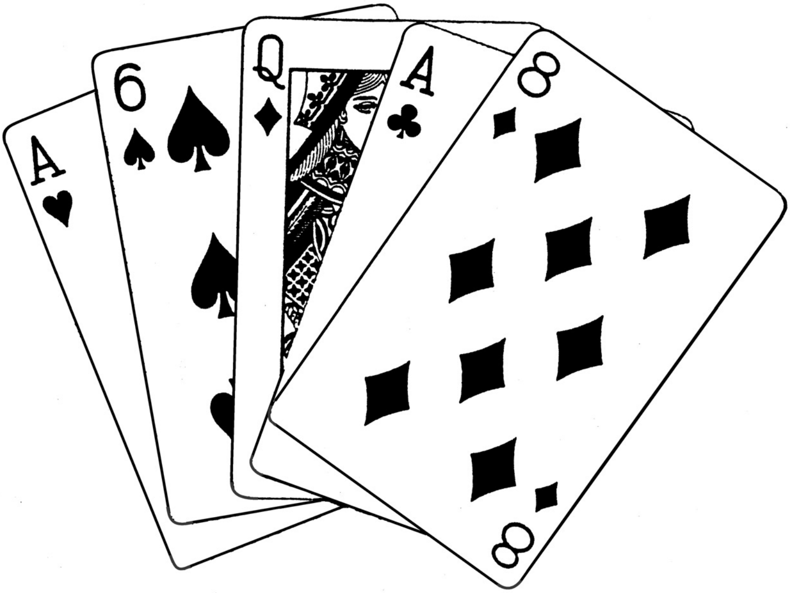


Figure 1.2.4. A five-card poker hand.

1.2.5 Example. The game of bridge uses the same 52 cards as poker.* The number of different 13-card bridge hands is

$$C(52,13) = \frac{52!}{13!\,39!} = \frac{52 \times 51 \times \cdots \times 40 \times 39!}{13! \times 39!} = \frac{52 \times 51 \times \cdots \times 40}{13!},$$

about 635,000,000,000. □

It may surprise you to learn that $C(52,13)$ is so much larger than $C(52,5)$. On the other hand, it does seem clear from Fig. 1.2.3 that the numbers in each row of Pascal's triangle increase, from left to right, up to the middle of the row and then decrease from the middle to the right-hand end. Rows for which this property holds are said to be *unimodal*.

1.2.6 Theorem. *The rows of Pascal's triangle are unimodal.*

*The actual, physical cards are typically slimmer to accommodate the larger, 13-card hands.

Proof. If $n > 2r + 1$, the ratio

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

implying that $C(n, r+1) > C(n, r)$. ■

1.2. EXERCISES

1 Compute

(a) $C(7,4)$. **(b)** $C(10,5)$. **(c)** $C(12,4)$.

(d) $C(101,2)$. **(e)** $C(101,99)$. **(f)** $C(12,6)$.

2 If n and r are integers satisfying $n > r \geq 0$, prove that

(a) $(r+1)C(n,r+1) = (n-r)C(n,r)$.

(b) $(r+1)C(n,r+1) = nC(n-1,r)$.

3 Write out rows 7 through 10 of Pascal's triangle and confirm that the sum of the numbers in the 10th row is $2^{10} = 1024$.

4 Consider the sequence of numbers 0, 0, 1, 3, 6, 10, 15, ... from the third $(r = 2)$ column of Pascal's triangle. Starting with $n = 0$, the nth term of the sequence is $a_n = C(n,2)$. Prove that, for all $n \geq 0$,

(a) $a_{n+1} - a_n = n$. **(b)** $a_{n+1} + a_n = n^2$.

5 Consider the sequence $b_0, b_1, b_2, b_3, \ldots$, where $b_n = C(n,3)$. Prove that, for all $n \geq 0$,

(a) $b_{n+1} - b_n = C(n,2)$.

(b) $b_{n+2} - b_n$ is a perfect square.

6 Poker is sometimes played with a joker. How many different five-card poker hands can be "chosen" from a deck of 53 cards?

7 Phrobana is a game played with a deck of 48 cards (no aces). How many different 12-card phrobana hands are there?

8 Give the inductive proof that an n-element set has 2^n subsets.

9 Let r_i be a positive integer, $1 \leq i \leq k$. If $n = r_1 + r_2 + \cdots + r_k$, prove that

$$\binom{n}{r_1, r_2, \ldots, r_k} = \binom{n-1}{r_1 - 1, r_2, \ldots, r_k} + \binom{n-1}{r_1, r_2 - 1, \ldots, r_k} + \cdots + \binom{n-1}{r_1, r_2, \ldots, r_k - 1}$$

(a) using algebraic arguments.
(b) using combinatorial arguments.

10 Suppose n, k, and r are integers that satisfy $n \geq k \geq r \geq 0$ and $k > 0$. Prove that

(a) $C(n,k)C(k,r) = C(n,r)C(n-r,k-r)$.

(b) $C(n,k)C(k,r) = C(n,k-r)C(n-k+r,r)$.

(c) $\sum_{j=0}^{n} C(n,j)C(j,r) = C(n,r)2^{n-r}$.

(d) $\sum_{j=k}^{n} (-1)^{j+k} C(n,j) = C(n-1,k-1)$.

11 Prove that $\left[\sum_{r=0}^{n} C(n,r)\right]^2 = \sum_{s=0}^{2n} C(2n,s)$.

12 Prove that $C(2n,n)$, $n > 0$, is always even.

13 Probably first studied by Leonhard Euler (1707–1783), the *Catalan sequence** 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, ... is defined by $c_n = C(2n,n)/(n+1)$, $n \geq 0$. Confirm that the Catalan numbers satisfy

(a) $c_2 = 2c_1$.
(b) $c_3 = 3c_2 - c_1$.
(c) $c_4 = 4c_3 - 3c_2$.
(d) $c_5 = 5c_4 - 6c_3 + c_2$.
(e) $c_6 = 6c_5 - 10c_4 + 4c_3$.
(f) $c_7 = 7c_6 - 15c_5 + 10c_4 - c_3$.
(g) Speculate about the general form of these equations.
(h) Prove or disprove your speculations from part (g).

14 Show that the Catalan numbers (Exercise 13) satisfy

(a) $c_n = C(2n-1,n-1) - C(2n-1,n+1)$.

(b) $c_n = C(2n,n) - C(2n,n-1)$.

(c) $c_{n+1} = \dfrac{4n+2}{n+2} c_n$.

15 One way to illustrate an r-element subset S of $\{1,2,\ldots,n\}$ is this: Let P_0 be the origin of the xy-plane. Setting $x_0 = y_0 = 0$, define

$$P_k = (x_k, y_k) = \begin{cases} (x_{k-1}+1, y_{k-1}) & \text{if} \quad k \in S, \\ (x_{k-1}, y_{k-1}+1) & \text{if} \quad k \notin S. \end{cases}$$

Finally, connect successive points by unit segments (either horizontal or vertical) to form a "path". Figure 1.2.5 illustrates the path corresponding to $S = \{3,4,6,8\}$ and $n = 8$.

*Euler was so prolific that more than one topic has come to be named for the first person to work on it *after* Euler, in this case, Eugene Catalan (1814–1894).

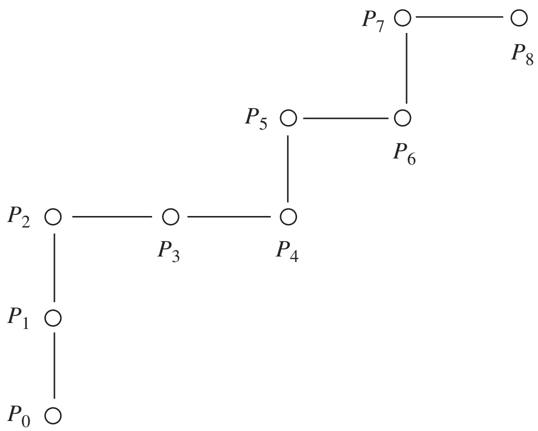


Figure 1.2.5

(a) Illustrate $E = \{2, 4, 6, 8\}$ when $n = 8$.

(b) Illustrate $E = \{2, 4, 6, 8\}$ when $n = 9$.

(c) Illustrate $D = \{1, 3, 5, 7\}$ when $n = 8$.

(d) Show that $P_n = (r, n - r)$ when S is an r-element set.

(e) A *lattice path* of *length* n in the xy-plane begins at the origin and consists of n unit "steps" each of which is either up or to the right. If r of the steps are to the right and $s = n - r$ of them are up, the lattice path terminates at the point (r, s). How many different lattice paths terminate at (r, s)?

16 Define $c_0 = 1$ and let c_n be the number of lattice paths of length $2n$ (Exercise 15) that terminate at (n, n) and never rise above the line $y = x$, i.e., such that $x_k \geq y_k$ for each point $P_k = (x_k, y_k)$. Show that

(a) $c_1 = 1$, $c_2 = 2$, and $c_3 = 5$.

(b) $c_{n+1} = \sum_{r=0}^{n} c_r c_{n-r}$. (*Hint:* Lattice paths "touch" the line $y = x$ for the last time at the point (n, n). Count those whose next-to-last touch is at the point (r, r)).

(c) c_n is the nth Catalan number of Exercises 13–14, $n \geq 1$.

17 Let X and Y be disjoint sets containing n and m elements, respectively. In how many different ways can an $(r + s)$-element subset Z be chosen from $X \cup Y$ if r of its elements must come from X and s of them from Y?

18 Packing for a vacation, a young man decides to take 3 long-sleeve shirts, 4 short-sleeve shirts, and 2 pairs of pants. If he owns 16 long-sleeve shirts, 20 short-sleeve shirts, and 13 pairs of pants, in how many different ways can he pack for the trip?

n \ r	0	1	2	3	4	5	6	7
0	C(0,0)							
1	C(1,0)	C(1,1)						
2	C(2,0)	C(2,1)	C(2,2)					
3	C(3,0)	C(3,1)	C(3,2)	**C(3,3)**				
4	C(4,0)	C(4,1)	**C(4,2)**	C(4,3)	C(4,4)			
5	C(5,0)	**C(5,1)**	C(5,2)	C(5,3)	C(5,4)	C(5,5)		
6	**C(6,0)**	C(6,1)	C(6,2)	C(6,3)	C(6,4)	C(6,5)	C(6,6)	
7	C(7,0)	C(7,1)	C(7,2)	C(7,3)	C(7,4)	C(7,5)	C(7,6)	C(7,7)
				...				

Figure 1.2.6

19 Suppose n is a positive integer and let $k = \lfloor n/2 \rfloor$, the greatest integer not larger than $n/2$. Define

$$F_n = C(n, 0) + C(n-1, 1) + C(n-2, 2) + \cdots + C(n-k, k).$$

Starting with $n = 0$, the sequence $\{F_n\}$ is

$$1, 1, 2, 3, 5, 8, 13, \ldots,$$

where, e.g., the $7th$ number in the sequence, $F_6 = 13$, is computed by summing the **boldface** numbers in Fig. 1.2.6.*

(a) Compute F_7 directly from the definition.

(b) Prove the recurrence $F_{n+2} = F_{n+1} + F_n$, $n \geq 0$.

(c) Compute F_7 using part (b) and the initial fragment of the sequence given above.

(d) Prove that $\sum_{i=0}^{n} F_i = F_{n+2} - 1$.

20 C. A. Tovey used the Fibonacci sequence (Exercise 19) to prove that infinitely many pairs (n, k) solve the equation $C(n, k) = C(n-1, k+1)$. The first pair is $C(2, 0) = C(1, 1)$. Find the second. (*Hint*: $n < 20$. Your solution need not make use of the Fibonacci sequence.)

21 The Buda side of the Danube is hilly and suburban while the Pest side is flat and urban. In short, Budapest is a divided city. Following the creation of a new commission on culture, suppose 6 candidates from Pest and 4 from Buda volunteer to serve. In how many ways can the mayor choose a 5-member commission.

*It was the French number theorist François Édouard Anatole Lucas (1842–1891) who named these numbers after Leonardo of Pisa (ca. 1180–1250), a man also known as Fibonacci.

(a) from the 10 candidates?

(b) if proportional representation dictates that 3 members come from Pest and 2 from Buda?

22 H. B. Mann and D. Shanks discovered a criterion for primality in terms of Pascal's triangle: Shift each of the $n + 1$ entries in row n to the right so that they begin in column $2n$. Circle the entries in row n that are multiples of n. Then r is prime if and only if all the entries in column r have been circled. Columns 0–11 are shown in Fig. 1.2.7. Continue the figure down to row 9 and out to column 20.

$n \backslash r$	0	1	2	3	4	5	6	7	8	9	10	11
1			①	①								
2					1	②	1					
3							1	③	③	1		
4									1	④	6	④
5											1	⑤

. . .

Figure 1.2.7

23 The superintendent of the Hardluck Elementary School District suggests that the Board of Education meet a \$5 million budget deficit by raising average class sizes, from 30 to 36 students, a 20% increase. A district teacher objects, pointing out that if the proposal is adopted, the potential for a *pair* of classmates to get into trouble will increase by 45%. What is the teacher talking about?

24 Strictly speaking, Theorem 1.2.6 establishes only half of the unimodality property. Prove the other half.

25 If n and r are nonnegative integers and x is an indeterminate, define $K(n, r) = (1 + x)^n x^r$.

(a) Show that $K(n + 1, r) = K(n, r) + K(n, r + 1)$.

(b) Compare and contrast the identity in part (a) with Pascal's relation.

(c) Since part (a) is a polynomial identity, it holds when numbers are substituted for x. Let $k(n, r)$ be the value of $K(n, r)$ when $x = 2$ and exhibit the numbers $k(n, r)$, $0 \leq n, r \leq 4$, in a 5×5 array, the rows of which are indexed by n and the columns by r. (*Hint:* Visually confirm that $k(n + 1, r) = k(n, r) + k(n, r + 1)$, $0 \leq n, r \leq 3$.)

26 Let S be an n-element set, where $n \geq 1$. If A is a subset of S, denote by $o(A)$ the *cardinality* of (number of elements in) A. Say that A is odd (even) if $o(A)$ is odd (even). Prove that the number of odd subsets of S is equal to the number of its even subsets.

27 Show that there are exactly seven different ways to factor $n = 63{,}000$ as a product of two relatively prime integers, each greater than one.

28 Suppose $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$, where $p_1, p_2, \ldots, p_r$ are distinct primes. Prove that there are exactly $2^{r-1} - 1$ different ways to factor n as a product of two relatively prime integers, each greater than one.

*1.3. ELEMENTARY PROBABILITY

> The theory of probabilities is basically only common sense reduced to calculation; it makes us appreciate with precision what reasonable minds feel by a kind of instinct, often being unable to account for it. ... It is remarkable that [this] science, which began with the consideration of games of chance, should have become the most important object of human knowledge.
>
> — Pierre Simon, Marquis de Laplace (1749–1827)

Elementary probability theory begins with the consideration of D equally likely "events" (or "outcomes"). If N of these are "noteworthy", then the probability of a noteworthy event is the fraction N/D. Maybe a brown paper bag contains a dozen jelly beans, say, 1 red, 2 orange, 2 blue, 3 green, and 4 purple. If a jelly bean is chosen at random from the bag, the probability that it will be blue is $\frac{2}{12} = \frac{1}{6}$; the probability that it will be green is $\frac{3}{12} = \frac{1}{4}$; the probability that it will be blue or green is $(2+3)/12 = \frac{5}{12}$; and the probability that it will be blue and green is $\frac{0}{12} = 0$.

Dice are commonly associated with games of chance. In a dice game, one is typically interested only in the numbers that rise to the top. If a single die is rolled, there are just six outcomes; if the die is "fair", each of them is equally likely. In computing the probability, say, of rolling a number greater than 4 with a single fair die, the denominator is $D = 6$. Since there are $N = 2$ noteworthy outcomes, namely 5 and 6, the probability we want is $P = \frac{2}{6} = \frac{1}{3}$.

The situation is more complicated when two dice are rolled. If all we care about is their sum, then there are 11 possible outcomes, anything from 2 to 12. But, the probability of rolling a sum, say, of 7 is not $\frac{1}{11}$ because these 11 outcomes are not equally likely. To help facilitate the discussion, assume that one of the dice is green and the other is red. Each time the dice are rolled, Lady Luck makes two decisions, choosing a number for the green die, and one for the red. Since there are 6 choices for each of them, the two decisions can be made in any one of $6^2 = 36$ ways. If both dice are fair, then *each of these* 36 *outcomes is equally likely*. Glancing at Fig. 1.3.1,

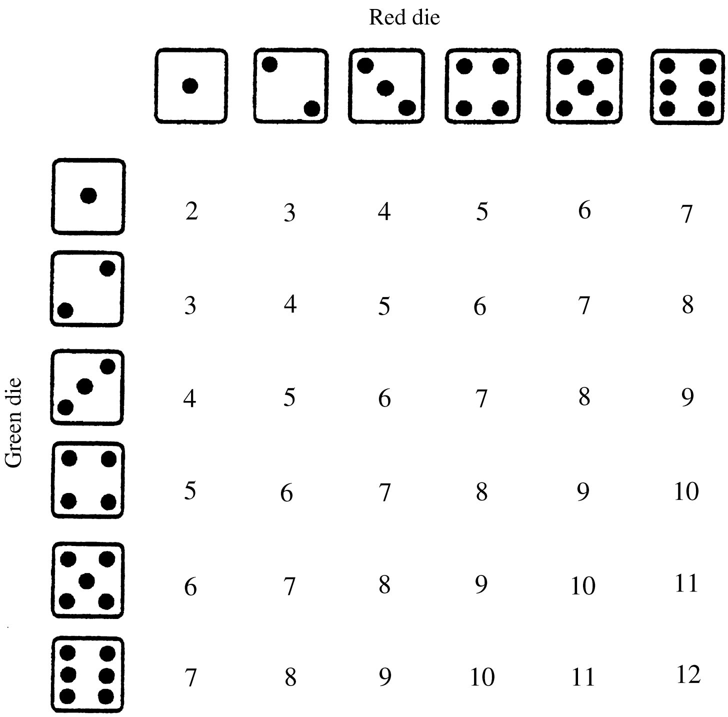


Figure 1.3.1. The 36 outcomes of rolling two dice.

one sees that there are six ways the dice can sum to 7, namely, a green 1 and a red 6, a green 2 and a red 5, a green 3 and a red 4, and so on. So, the probability of rolling a (sum of) 7 is not $\frac{1}{11}$ but $\frac{6}{36} = \frac{1}{6}$.

1.3.1 Example. Denote by $P(n)$ the probability of rolling (a sum of) n with two fair dice. Using Fig. 1.3.1, it is easy to see that $P(2) = \frac{1}{36} = P(12)$, $P(3) = \frac{2}{36} = \frac{1}{18} = P(11)$, $P(4) = \frac{3}{36} = \frac{1}{12} = P(10)$, and so on. What about $P(1)$? Since 1 is not among the outcomes, $P(1) = \frac{0}{36} = 0$. In fact, if P is some probability (any probability at all), then $0 \leq P \leq 1$. □

1.3.2 Example. A popular game at charity fundraisers is Chuck-a-Luck. The apparatus for the game consists of three dice housed in an hourglass-shaped cage. Once the patrons have placed their bets, the operator turns the cage and the dice roll to the bottom. If none of the dice comes up 1, the bets are lost. Otherwise, the operator matches, doubles, or triples each wager depending on the number of "aces" (1's) showing on the three dice.

Let's compute probabilities for various numbers of 1's. By the fundamental counting principle, there are $6^3 = 216$ possible outcomes (all of which are equally

Number of 1's	0	1	2	3
Probability	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$

Figure 1.3.2. Chuck-a-Luck probabilities.

likely if the dice are fair). Of these 216 outcomes, only one consists of three 1's. Thus, the probability that the bets will have to be tripled is $\frac{1}{216}$.

In how many ways can two 1's come up? Think of it as a sequence of two decisions. The first is which die should produce a number different from 1. The second is what number should appear on that die. There are three choices for the first decision and five for the second. So, there are $3 \times 5 = 15$ ways for the three dice to produce exactly two 1's. The probability that the bets will have to be doubled is $\frac{15}{216}$.

What about a single ace? This case can be approached as a sequence of three decisions. Decision 1 is which die should produce the 1 (three choices). The second decision is what number should appear on the second die (five choices, anything but 1). The third decision is the number on the third die (also five choices). Evidently, there are $3 \times 5 \times 5 = 75$ ways to get exactly one ace. So far, we have accounted for $1 + 15 + 75 = 91$ of the 216 possible outcomes. (In other words, the probability of getting *at least* one ace is $\frac{91}{216}$.) In the remaining $216 - 91 = 125$ outcomes, three are no 1's at all. These results are tabulated in Fig. 1.3.2. □

Some things, like determining which team kicks off to start a football game, are decided by tossing a coin. A fair coin is one in which each of the two possible outcomes, heads or tails, is equally likely. When a fair coin is tossed, the probability that it will come up heads is $\frac{1}{2}$.

Suppose four (fair) coins are tossed. What is the probability that half of them will be heads and half tails? Is it obvious that the answer is $\frac{3}{8}$? Once again, Lady Luck has a sequence of decisions to make, this time four of them. Since there are two choices for each decision, $D = 2^4$. With the noteworthies in **boldface**, these 16 outcomes are arrayed in Fig. 1.3.3. By inspection, $N = 6$, so the probability we seek is $\frac{6}{16} = \frac{3}{8}$.

HHHH	HTHH	THHH	**TTHH**
HHHT	**HTHT**	**THHT**	TTHT
HHTH	**HTTH**	**THTH**	TTTH
HHTT	HTTT	THTT	TTTT

Figure 1.3.3

1.3.3 Example. If 10 (fair) coins are tossed, what is the probability that half of them will be heads and half tails? Ten decisions, each with two choices, yields $D = 2^{10} = 1024$. To compute the numerator, imagine a systematic list analogous to Fig. 1.3.3. In the case of 10 coins, the noteworthy outcomes correspond to

10-letter "words" with five H's and five T's, so $N = \binom{10}{5,5} = C(10,5) = 252$, and the desired probability is $\frac{252}{1024} \doteq 0.246$. More generally, if n coins are tossed, the probability that exactly r of them will come up heads is $C(n,r)/2^n$.

What about the probability that *at most* r of them will come up heads? That's easy enough: $P = N/2^n$, where $N = N(n,r) = C(n,0) + C(n,1) + \cdots + C(n,r)$ is the number of n-letter words that can be assembled from the alphabet $\{H, T\}$ and that contain at most r H's. □

Here is a different kind of problem: Suppose two fair coins are tossed, say a dime and a quarter. If you are told (only) that one of them is heads, what is the probability that the other one is also heads? (Don't just guess, think about it.)

May we assume, without loss of generality, that the dime is heads? If so, because the quarter has a head of its own, so to speak, the answer should be $\frac{1}{2}$. To see why this is wrong, consider the equally likely outcomes when two fair coins are tossed, namely, HH, HT, TH, and TT. If all we know is that one (at least) of the coins is heads, then TT is eliminated. Since the remaining three possibilities are still equally likely, $D = 3$, and the answer is $\frac{1}{3}$.

There are two "morals" here. One is that the most reliable guide to navigating probability theory is *equal likelihood*. The other is that finding a correct answer often depends on having a precise understanding of the question, and that requires precise language.

1.3.4 Definition. A nonempty finite set E of equally likely outcomes is called a *sample space*. The number of elements in E is denoted $o(E)$. For any subset A of E, the probability of A is $P(A) = o(A)/o(E)$. If B is a subset of E, then $P(A \text{ or } B) = P(A \cup B)$, and $P(A \text{ and } B) = P(A \cap B)$.

In mathematical writing, an unqualified "or" is inclusive, as in "A or B or both".*

1.3.5 Theorem. *Let E be a fixed but arbitrary sample space. If A and B are subsets of E, then*

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$

Proof. The sum $o(A) + o(B)$ counts all the elements of A and all the elements of B. It even counts some elements twice, namely those in $A \cap B$. Subtracting $o(A \cap B)$ compensates for this double counting and yields

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

(Notice that this formula generalizes the second counting principle; it is a special case of the even more general principle of inclusion and exclusion, to be discussed in Chapter 2.) It remains to divide both sides by $o(E)$ and use Definition 1.3.4. ■

*The exclusive "or" can be expressed using phrases like "either A or B" or "A or B but not both".

1.3.6 Corollary. *Let E be a fixed but arbitrary sample space. If A and B are subsets of E, then $P(A \text{ or } B) \leq P(A) + P(B)$ with equality if and only if A and B are disjoint.*

Proof. $P(A \text{ and } B) = 0$ if and only if $o(A \cap B) = 0$ if and only if $A \cap B = \emptyset$. ■

A special case of this corollary involves the *complement*, $A^c = \{x \in E : x \notin A\}$. Since $A \cup A^c = E$ and $A \cap A^c = \emptyset$, $o(A) + o(A^c) = o(E)$. Dividing both sides of this equation by $o(E)$ yields the useful identity

$$P(A) + P(A^c) = 1.$$

1.3.7 Example. Suppose two fair dice are rolled, say a red one and a green one. What is the probability of rolling a 3 on the red die, call it a red 3, or a green 2? Let's abbreviate by setting $R3 = \text{red 3}$ and $G2 = \text{green 2}$ so that, e.g., $P(R3) = \frac{1}{6} = P(G2)$.

Solution 1: When both dice are rolled, only one of the $6^2 = 36$ equally likely outcomes corresponds to $R3$ and $G2$, so $P(R3 \text{ and } G2) = \frac{1}{36}$. Thus, by Theorem 1.3.5,

$$\begin{aligned} P(R3 \text{ or } G2) &= P(R3) + P(G2) - P(R3 \text{ and } G2) \\ &= \tfrac{1}{6} + \tfrac{1}{6} - \tfrac{1}{36} \\ &= \tfrac{11}{36}. \end{aligned}$$

Solution 2: Let P_c be the complementary probability that neither $R3$ nor $G2$ occurs. Then $P_c = N/D$, where $D = 36$. The evaluation of N can be viewed in terms of a sequence of two decisions. There are five choices for the "red" decision, anything but number 3, and five for the "green" one, anything but number 2. Hence, $N = 5 \times 5 = 25$, and $P_c = \frac{25}{36}$, so the probability we want is

$$P(R3 \text{ or } G2) = 1 - P_c = \tfrac{11}{36}.$$ □

1.3.8 Example. Suppose a single (fair) die is rolled twice. What is the probability that the first roll is a 3 or the second roll is a 2? Solution: $\frac{11}{36}$. This problem is equivalent to the one in Example 1.3.7. □

1.3.9 Example. Suppose a single (fair) die is rolled twice. What is the probability of getting a 3 or a 2?

Solution 1: Of the $6 \times 6 = 36$ equally likely outcomes, $4 \times 4 = 16$ involve neither a 3 nor a 2. The complementary probability is $P(2 \text{ or } 3) = 1 - \frac{16}{36} = \frac{5}{9}$.

Solution 2: There are two ways to roll a 3 *and* a 2; either the 3 comes first followed by the 2 or the other way around. So, $P(3 \text{ and } 2) = \frac{2}{36} = \frac{1}{18}$. Using Theorem 1.3.5, $P(3 \text{ or } 2) = \frac{1}{6} + \frac{1}{6} - \frac{1}{18} = \frac{5}{18}$.

Whoops! Since $\frac{5}{9} \neq \frac{5}{18}$, one (at least) of these "solutions" is incorrect. The probability computed in solution 1 is greater than $\frac{1}{2}$, which *seems* too large. On the other hand, it is not hard to spot an error in solution 2, namely, the incorrect application of Theorem 1.3.5. The calculation $P(3) = \frac{1}{6}$ would be valid had the die been rolled only *once*. For this problem, the correct interpretation of $P(3)$ is the probability that the first roll is 3 or the second roll is 3. That should be identical to the probability determined in Example 1.3.8. (Why?) Using the (correct) values $P(3) = P(2) = \frac{11}{36}$ in solution 2, we obtain $P(2 \text{ or } 3) = \frac{11}{36} + \frac{11}{36} - \frac{1}{18} = \frac{5}{9}$.

The next time you get a chance, roll a couple of dice and see if you can avoid both 2's and 3's more than 44 times out of 99. □

Another approach to $P(A \text{ and } B)$ emerges from the notion of "conditional probability".

1.3.10 Definition. Let E be a fixed but arbitrary sample space. If A and B are subsets of E, the *conditional probability*

$$P(B|A) = \begin{cases} P(B) & \text{if } A = \emptyset, \\ o(A \cap B)/o(A) & \text{otherwise.} \end{cases}$$

When A is not empty, $P(B|A)$ may be viewed as the probability of B given that A is certain (e.g., known already to have occurred). The problem of tossing two fair coins, a dime and a quarter, involved conditional probabilities. If h and t represent heads and tails, respectively, for the dime and H and T for the quarter, then the sample space $E = \{hH, hT, tH, tT\}$. If $A = \{hH, hT, tH\}$ and $B = \{hH\}$, then $P(B|A) = \frac{1}{3}$ is the probability that both coins are heads given that one of them is. If $C = \{hH, hT\}$, then $P(B|C) = \frac{1}{2}$ is the probability that both coins are heads given that the dime is.

1.3.11 Theorem. *Let E be a fixed but arbitrary sample space. If A and B are subsets of E, then*

$$P(A \text{ and } B) = P(A)P(B|A).$$

Proof. Let $D = o(E)$, $a = o(A)$, and $N = o(A \cap B)$. If $a = 0$, there is nothing to prove. Otherwise, $P(A) = a/D$, $P(B|A) = N/a$, and $P(A)P(B|A) = (a/D)(N/a) = N/D = P(A \text{ and } B)$. ■

1.3.12 Corollary (Bayes's* First Rule). *Let E be a fixed but arbitrary sample space. If A and B are subsets of E, then* $P(A)P(B|A) = P(B)P(A|B)$.

Proof. Because $P(A \text{ and } B) = P(B \text{ and } A)$, the result is immediate from Theorem 1.3.11. ■