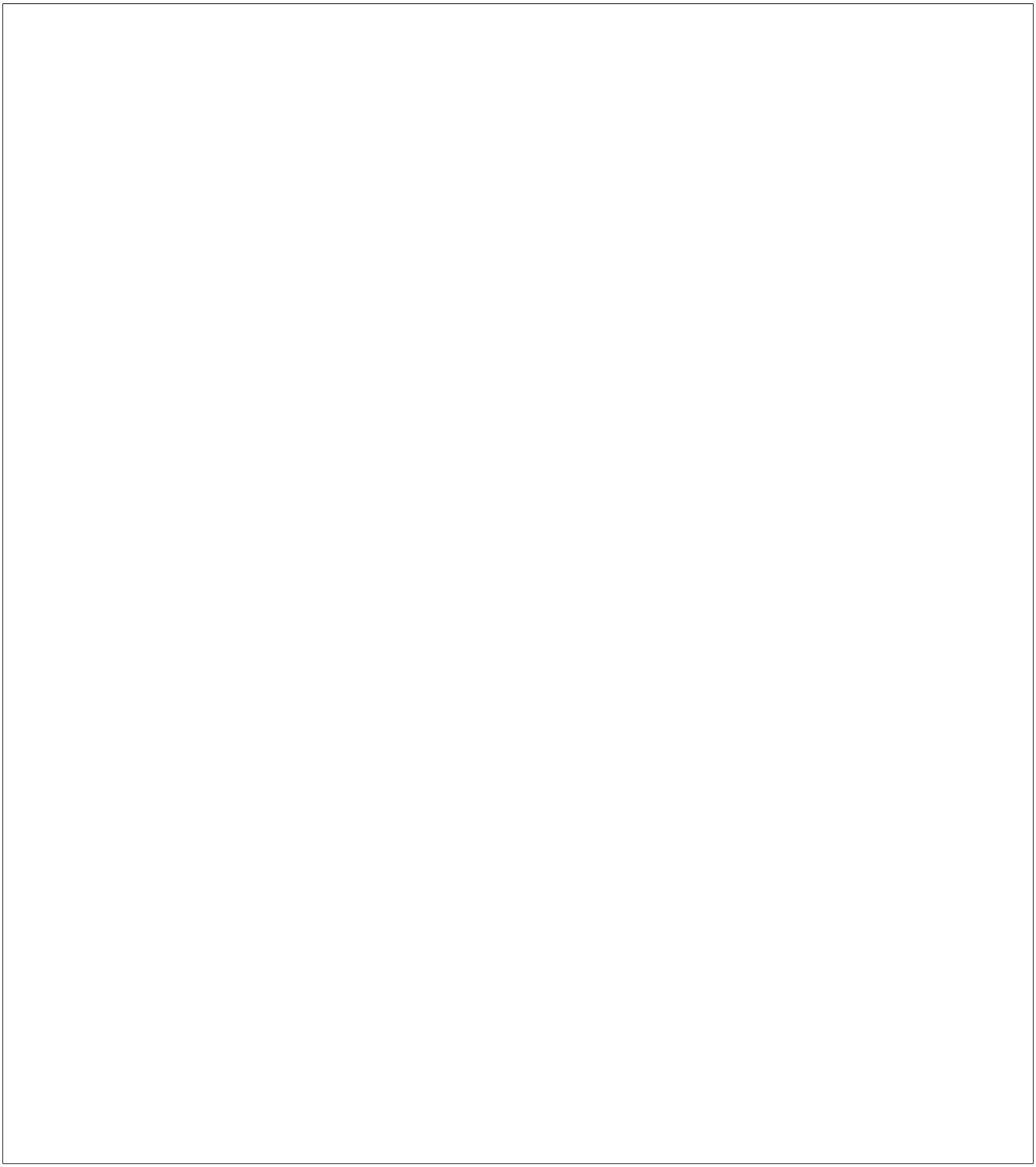
Mathematics for Computer Science

revised Wednesday 8th September, 2010, 00:40

Eric Lehman
Google Inc.

F Thomson Leighton
Department of Mathematics and CSAIL, MIT
Akamai Technologies

Albert R Meyer
Massachusets Institute of Technology



Contents

II Structures

III Counting

2. hop, hop
3. hop, step, step
4. step, hop step
5. step, step, hop

Working through this problem will demonstrate the major features of our first cookbook method for solving recurrences. We'll fill in the details of the general solution afterward.

Finding a Recurrence

As special cases, there is 1 way to climb 0 stairs (do nothing) and 1 way to climb 1 stair (step up). In general, an ascent of n stairs consists of either a step followed by an ascent of the remaining $n-1$ stairs or a hop followed by an ascent of $n-2$ stairs. So the total number of ways to climb n stairs is equal to the number of ways to climb $n-1$ plus the number of ways to climb $n-2$. These observations define a recurrence:

$$\begin{aligned} f(0) &= 1 \\ f(1) &= 1 \\ f(n) &= f(n-1) + f(n-2) \qquad \text{for } n \geq 2. \end{aligned}$$

Here, $f(n)$ denotes the number of ways to climb n stairs. Also, we've switched from subscript notation to functional notation, from T_n to f_n. Here the change is cosmetic, but the expressiveness of functions will be useful later.

This is the Fibonacci recurrence, the most famous of all recurrence equations. Fibonacci numbers arise in all sorts of applications and in nature. Fibonacci introduced the numbers in 1202 to study rabbit reproduction. Fibonacci numbers also appear, oddly enough, in the spiral patterns on the faces of sunflowers. And the input numbers that make Euclid's GCD algorithm require the greatest number of steps are consecutive Fibonacci numbers.

Solving the Recurrence

The Fibonacci recurrence belongs to the class of linear recurrences, which are essentially all solvable with a technique that you can learn in an hour. This is somewhat amazing, since the Fibonacci recurrence remained unsolved for almost six centuries!

In general, a *homogeneous linear recurrence* has the form

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \ldots + a_d f(n-d)$$

where $a_1, a_2, \ldots, a_d$ and d are constants. The *order* of the recurrence is d. Commonly, the value of the function f is also specified at a few points; these are called *boundary conditions*. For example, the Fibonacci recurrence has order $d = 2$ with coefficients $a_1 = a_2 = 1$ and $g(n) = 0$. The boundary conditions are $f(0) = 1$ and $f(1) = 1$. The word "homogeneous" sounds scary, but effectively means "the simpler kind". We'll consider linear recurrences with a more complicated form later.

Let's try to solve the Fibonacci recurrence with the benefit centuries of hindsight. In general, linear recurrences tend to have exponential solutions. So let's guess that

$$f(n) = x^n$$

where x is a parameter introduced to improve our odds of making a correct guess. We'll figure out the best value for x later. To further improve our odds, let's neglect the boundary conditions, $f(0) = 0$ and $f(1) = 1$, for now. Plugging this guess into the recurrence $f(n) = f(n-1) + f(n-2)$ gives

$$x^n = x^{n-1} + x^{n-2}.$$

Dividing both sides by x^{n-2} leaves a quadratic equation:

$$x^2 = x + 1.$$

Solving this equation gives *two* plausible values for the parameter x:

$$x = \frac{1 \pm \sqrt{5}}{2}.$$

This suggests that there are at least two different solutions to the recurrence, neglecting the boundary conditions.

$$f(n) = \left(\frac{1+\sqrt{5}}{2}\right)^n \quad \text{or} \quad f(n) = \left(\frac{1-\sqrt{5}}{2}\right)^n$$

A charming features of homogeneous linear recurrences is that any linear combination of solutions is another solution.

Theorem 10.3.1. *If $f(n)$ and $g(n)$ are both solutions to a homogeneous linear recurrence, then $h(n) = sf(n) + tg(n)$ is also a solution for all $s, t \in \mathbb{R}$.*

Proof.

$$\begin{aligned}
h(n) &= sf(n) + tg(n) \\
&= s\left(a_1 f(n-1) + \ldots + a_d f(n-d)\right) + t\left(a_1 g(n-1) + \ldots + a_d g(n-d)\right) \\
&= a_1(sf(n-1) + tg(n-1)) + \ldots + a_d(sf(n-d) + tg(n-d)) \\
&= a_1 h(n-1) + \ldots + a_d h(n-d)
\end{aligned}$$

The first step uses the definition of the function h, and the second uses the fact that f and g are solutions to the recurrence. In the last two steps, we rearrange terms and use the definition of h again. Since the first expression is equal to the last, h is also a solution to the recurrence. ■

The phenomenon described in this theorem—a linear combination of solutions is another solution—also holds for many differential equations and physical systems. In fact, linear recurrences are so similar to linear differential equations that you can safely snooze through that topic in some future math class.

Returning to the Fibonacci recurrence, this theorem implies that

$$f(n) = s\left(\frac{1+\sqrt{5}}{2}\right)^n + t\left(\frac{1-\sqrt{5}}{2}\right)^n$$

is a solution for all real numbers s and t. The theorem expanded two solutions to a whole spectrum of possibilities! Now, given all these options to choose from, we can find one solution that satisfies the boundary conditions, $f(0) = 1$ and $f(1) = 1$. Each boundary condition puts some constraints on the parameters s and t. In particular, the first boundary condition implies that

$$f(0) = s\left(\frac{1+\sqrt{5}}{2}\right)^0 + t\left(\frac{1-\sqrt{5}}{2}\right)^0 = s + t = 1.$$

Similarly, the second boundary condition implies that

$$f(1) = s\left(\frac{1+\sqrt{5}}{2}\right)^1 + t\left(\frac{1-\sqrt{5}}{2}\right)^1 = 1.$$

Now we have two linear equations in two unknowns. The system is not degenerate, so there is a unique solution:

$$s = \frac{1}{\sqrt{5}} \cdot \frac{1+\sqrt{5}}{2} \qquad t = -\frac{1}{\sqrt{5}} \cdot \frac{1-\sqrt{5}}{2}.$$

These values of s and t identify a solution to the Fibonacci recurrence that also satisfies the boundary conditions:

$$\begin{aligned} f(n) &= \frac{1}{\sqrt{5}} \cdot \frac{1+\sqrt{5}}{2}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \cdot \frac{1-\sqrt{5}}{2}\left(\frac{1-\sqrt{5}}{2}\right)^n \\ &= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}. \end{aligned}$$

It is easy to see why no one stumbled across this solution for almost six centuries! All Fibonacci numbers are integers, but this expression is full of square roots of five! Amazingly, the square roots always cancel out. This expression really does give the Fibonacci numbers if we plug in $n = 0, 1, 2$, etc.

This closed-form for Fibonacci numbers has some interesting corollaries. The first term tends to infinity because the base of the exponential, $(1 + \sqrt{5})/2 = 1.618\ldots$ is greater than one. This value is often denoted ϕ and called the "golden ratio". The second term tends to zero, because $(1 - \sqrt{5})/2 = -0.618033988\ldots$ has absolute value less than 1. This implies that the nth Fibonacci number is:

$$f(n) = \frac{\phi^{n+1}}{\sqrt{5}} + o(1).$$

Remarkably, this expression involving irrational numbers is actually very close to an integer for all large n—namely, a Fibonacci number! For example:

$$\frac{\phi^{20}}{\sqrt{5}} = 6765.000029\ldots \approx f(19).$$

This also implies that the ratio of consecutive Fibonacci numbers rapidly approaches the golden ratio. For example:

$$\frac{f(20)}{f(19)} = \frac{10946}{6765} = 1.618033998\ldots.$$

10.3.2 Solving Homogeneous Linear Recurrences

The method we used to solve the Fibonacci recurrence can be extended to solve any homogeneous linear recurrence; that is, a recurrence of the form

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \ldots + a_d f(n-d)$$

where $a_1, a_2, \ldots, a_d$ and d are constants. Substituting the guess $f(n) = x^n$, as with the Fibonacci recurrence, gives

$$x^n = a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_d x^{n-d}.$$

Dividing by x^{n-d} gives

$$x^d = a_1 x^{d-1} + a_2 x^{d-2} + \ldots + a_{d-1} x + a_d.$$

This is called the *characteristic equation* of the recurrence. The characteristic equation can be read off quickly since the coefficients of the equation are the same as the coefficients of the recurrence.

The solutions to a linear recurrence are defined by the roots of the characteristic equation. Neglecting boundary conditions for the moment:

- If r is a nonrepeated root of the characteristic equation, then r^n is a solution to the recurrence.
- If r is a repeated root with multiplicity k then r^n, nr^n, n^2r^n, ..., $n^{k-1}r^n$ are all solutions to the recurrence.

Theorem 10.3.1 implies that every linear combination of these solutions is also a solution.

For example, suppose that the characteristic equation of a recurrence has roots s, t, and u twice. These four roots imply four distinct solutions:

$$f(n) = s^n \qquad f(n) = t^n \qquad f(n) = u^n \qquad f(n) = nu^n.$$

Furthermore, every linear combination

$$f(n) = a \cdot s^n + b \cdot t^n + c \cdot u^n + d \cdot nu^n \tag{10.1}$$

is also a solution.

All that remains is to select a solution consistent with the boundary conditions by choosing the constants appropriately. Each boundary condition implies a linear equation involving these constants. So we can determine the constants by solving a system of linear equations. For example, suppose our boundary conditions were $f(0) = 0$, $f(1) = 1$, $f(2) = 4$, and $f(3) = 9$. Then we would obtain four equations in four unknowns:

$$\begin{aligned}
f(0) = 0 \quad &\Rightarrow \quad a \cdot s^0 + b \cdot t^0 + c \cdot u^0 + d \cdot 0u^0 = 0 \\
f(1) = 1 \quad &\Rightarrow \quad a \cdot s^1 + b \cdot t^1 + c \cdot u^1 + d \cdot 1u^1 = 1 \\
f(2) = 4 \quad &\Rightarrow \quad a \cdot s^2 + b \cdot t^2 + c \cdot u^2 + d \cdot 2u^2 = 4 \\
f(3) = 9 \quad &\Rightarrow \quad a \cdot s^3 + b \cdot t^3 + c \cdot u^3 + d \cdot 3u^3 = 9
\end{aligned}$$

This looks nasty, but remember that s, t, and u are just constants. Solving this system gives values for a, b, c, and d that define a solution to the recurrence consistent with the boundary conditions.

10.3.3 Solving General Linear Recurrences

We can now solve all linear homogeneous recurrences, which have the form

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \ldots + a_d f(n-d).$$

Many recurrences that arise in practice do not quite fit this mold. For example, the Towers of Hanoi problem led to this recurrence:

$$\begin{aligned}
&f(1) = 1 \\
&f(n) = 2f(n-1) + 1 \qquad (\text{for } n \geq 2).
\end{aligned}$$

The problem is the extra $+1$; that is not allowed in a homogeneous linear recurrence. In general, adding an extra function $g(n)$ to the right side of a linear recurrence gives an *inhomogeneous linear recurrence*:

$$f(n) = a_1 f(n-1) + a_2 f(n-2) + \ldots + a_d f(n-d) + g(n).$$

Solving inhomogeneous linear recurrences is neither very different nor very difficult. We can divide the whole job into five steps:

1. Replace $g(n)$ by 0, leaving a homogeneous recurrence. As before, find roots of the characteristic equation.
2. Write down the solution to the homogeneous recurrence, but do not yet use the boundary conditions to determine coefficients. This is called the *homogeneous solution*.
3. Now restore $g(n)$ and find a single solution to the recurrence, ignoring boundary conditions. This is called a *particular solution*. We'll explain how to find a particular solution shortly.
4. Add the homogeneous and particular solutions together to obtain the *general solution*.
5. Now use the boundary conditions to determine constants by the usual method of generating and solving a system of linear equations.

As an example, let's consider a variation of the Towers of Hanoi problem. Suppose that moving a disk takes time proportional to its size. Specifically, moving the smallest disk takes 1 second, the next-smallest takes 2 seconds, and moving the nth disk then requires n seconds instead of 1. So, in this variation, the time to complete the job is given by a recurrence with a $+n$ term instead of a $+1$:

$$\begin{aligned} f(1) &= 1 \\ f(n) &= 2f(n-1) + n \qquad \text{for } n \geq 2. \end{aligned}$$

Clearly, this will take longer, but how much longer? Let's solve the recurrence with the method described above.

In Steps 1 and 2, dropping the $+n$ leaves the homogeneous recurrence $f(n) = 2f(n-1)$. The characteristic equation is $x = 2$. So the homogeneous solution is $f(n) = c2^n$.

In Step 3, we must find a solution to the full recurrence $f(n) = 2f(n-1) + n$, without regard to the boundary condition. Let's guess that there is a solution of the

form $f(n) = an + b$ for some constants a and b. Substituting this guess into the recurrence gives

$$\begin{aligned} an + b &= 2(a(n-1) + b) + n \\ 0 &= (a+1)n + (b - 2a). \end{aligned}$$

The second equation is a simplification of the first. The second equation holds for all n if both $a + 1 = 0$ (which implies $a = -1$) and $b - 2a = 0$ (which implies that $b = -2$). So $f(n) = an + b = -n - 2$ is a particular solution.

In the Step 4, we add the homogeneous and particular solutions to obtain the general solution

$$f(n) = c2^n - n - 2.$$

Finally, in step 5, we use the boundary condition, $f(1) = 1$, determine the value of the constant c:

$$\begin{aligned} f(1) = 1 \quad &\Rightarrow \quad c2^1 - 1 - 2 = 1 \\ &\Rightarrow \quad c = 2. \end{aligned}$$

Therefore, the function $f(n) = 2 \cdot 2^n - n - 2$ solves this variant of the Towers of Hanoi recurrence. For comparison, the solution to the original Towers of Hanoi problem was $2^n - 1$. So if moving disks takes time proportional to their size, then the monks will need about twice as much time to solve the whole puzzle.

10.3.4 How to Guess a Particular Solution

Finding a particular solution can be the hardest part of solving inhomogeneous recurrences. This involves guessing, and you might guess wrong.[1] However, some rules of thumb make this job fairly easy most of the time.

- Generally, look for a particular solution with the same form as the inhomogeneous term $g(n)$.
- If $g(n)$ is a constant, then guess a particular solution $f(n) = c$. If this doesn't work, try polynomials of progressively higher degree: $f(n) = bn + c$, then $f(n) = an^2 + bn + c$, etc.
- More generally, if $g(n)$ is a polynomial, try a polynomial of the same degree, then a polynomial of degree one higher, then two higher, etc. For example, if $g(n) = 6n + 5$, then try $f(n) = bn + c$ and then $f(n) = an^2 + bn + c$.

[1] In Chapter 12, we will show how to solve linear recurrences with generating functions—it's a little more complicated, but it does not require guessing.

- If $g(n)$ is an exponential, such as 3^n, then first guess that $f(n) = c3^n$. Failing that, try $f(n) = bn3^n + c3^n$ and then $an^2 3^n + bn3^n + c3^n$, etc.

The entire process is summarized on the following page.

10.4 Divide-and-Conquer Recurrences

We now have a recipe for solving general linear recurrences. But the Merge Sort recurrence, which we encountered earlier, is not linear:

$$\begin{aligned} T(1) &= 0 \\ T(n) &= 2T(n/2) + n - 1 \qquad (\text{for } n \geq 2). \end{aligned}$$

In particular, $T(n)$ is not a linear combination of a fixed number of immediately preceding terms; rather, $T(n)$ is a function of $T(n/2)$, a term halfway back in the sequence.

Merge Sort is an example of a divide-and-conquer algorithm: it divides the input, "conquers" the pieces, and combines the results. Analysis of such algorithms commonly leads to *divide-and-conquer* recurrences, which have this form:

$$T(n) = \sum_{i=1}^{k} a_i T(b_i n) + g(n)$$

Here $a_1, \ldots a_k$ are positive constants, $b_1, \ldots, b_k$ are constants between 0 and 1, and $g(n)$ is a nonnegative function. For example, setting $a_1 = 2$, $b_1 = 1/2$, and $g(n) = n - 1$ gives the Merge Sort recurrence.

10.4.1 The Akra-Bazzi Formula

The solution to virtually all divide and conquer solutions is given by the amazing *Akra-Bazzi formula*. Quite simply, the asymptotic solution to the general divide-and-conquer recurrence

$$T(n) = \sum_{i=1}^{k} a_i T(b_i n) + g(n)$$

is

$$T(n) = \Theta\left(n^p \left(1 + \int_1^n \frac{g(u)}{u^{p+1}}\, du\right)\right) \tag{10.2}$$

Short Guide to Solving Linear Recurrences

A linear recurrence is an equation

$$f(n) = \underbrace{a_1 f(n-1) + a_2 f(n-2) + \ldots + a_d f(n-d)}_{\text{homogeneous part}} \quad \underbrace{+\, g(n)}_{\text{inhomogeneous part}}$$

together with boundary conditions such as $f(0) = b_0$, $f(1) = b_1$, etc. Linear recurrences are solved as follows:

1. Find the roots of the characteristic equation

$$x^n = a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{k-1} x + a_k.$$

2. Write down the homogeneous solution. Each root generates one term and the homogeneous solution is their sum. A nonrepeated root r generates the term $c r^n$, where c is a constant to be determined later. A root r with multiplicity k generates the terms

$$d_1 r^n \qquad d_2 n r^n \qquad d_3 n^2 r^n \qquad \ldots \qquad d_k n^{k-1} r^n$$

where $d_1, \ldots d_k$ are constants to be determined later.

3. Find a particular solution. This is a solution to the full recurrence that need not be consistent with the boundary conditions. Use guess-and-verify. If $g(n)$ is a constant or a polynomial, try a polynomial of the same degree, then of one higher degree, then two higher. For example, if $g(n) = n$, then try $f(n) = bn + c$ and then $an^2 + bn + c$. If $g(n)$ is an exponential, such as 3^n, then first guess $f(n) = c3^n$. Failing that, try $f(n) = (bn + c)3^n$ and then $(an^2 + bn + c)3^n$, etc.

4. Form the general solution, which is the sum of the homogeneous solution and the particular solution. Here is a typical general solution:

$$f(n) = \underbrace{c2^n + d(-1)^n}_{\text{homogeneous solution}} + \underbrace{3n + 1}_{\text{inhomogeneous solution}}.$$

5. Substitute the boundary conditions into the general solution. Each boundary condition gives a linear equation in the unknown constants. For example, substituting $f(1) = 2$ into the general solution above gives

$$\begin{aligned} 2 &= c \cdot 2^1 + d \cdot (-1)^1 + 3 \cdot 1 + 1 \\ \Rightarrow \quad -2 &= 2c - d. \end{aligned}$$

Determine the values of these constants by solving the resulting system of linear equations.

where p satisfies

$$\sum_{i=1}^{k} a_i b_i^p = 1. \tag{10.3}$$

A rarely-troublesome requirement is that the function $g(n)$ must not grow or oscillate too quickly. Specifically, $|g'(n)|$ must be bounded by some polynomial. So, for example, the Akra-Bazzi formula is valid when $g(n) = x^2 \log n$, but not when $g(n) = 2^n$.

Let's solve the Merge Sort recurrence again, using the Akra-Bazzi formula instead of plug-and-chug. First, we find the value p that satisfies

$$2 \cdot (1/2)^p = 1.$$

Looks like $p = 1$ does the job. Then we compute the integral:

$$\begin{aligned} T(n) &= \Theta\left(n\left(1 + \int_1^n \frac{u-1}{u^2} du\right)\right) \\ &= \Theta\left(n\left(1 + \left(\log u + \frac{1}{u}\right)_1^n\right)\right) \\ &= \Theta\left(n\left(\log n + \frac{1}{n}\right)\right) \\ &= \Theta\left(n \log n\right). \end{aligned}$$

The first step is integration and the second is simplification. We can drop the $1/n$ term in the last step, because the $\log n$ term dominates. We're done!

Let's try a scary-looking recurrence:

$$T(n) = 2T(n/2) + 8/9T(3n/4) + n^2.$$

Here, $a_1 = 2$, $b_1 = 1/2$, $a_2 = 8/9$, and $b_2 = 3/4$. So we find the value p that satisfies

$$2 \cdot (1/2)^p + (8/9)(3/4)^p = 1.$$

Equations of this form don't always have closed-form solutions, so you may need to approximate p numerically sometimes. But in this case the solution is simple: $p = 2$. Then we integrate:

$$\begin{aligned} T(n) &= \Theta\left(n^2\left(1 + \int_1^n \frac{u^2}{u^3} du\right)\right) \\ &= \Theta\left(n^2(1 + \log n)\right) \\ &= \Theta\left(n^2 \log n\right). \end{aligned}$$

That was easy!

10.4.2 Two Technical Issues

Until now, we've swept a couple issues related to divide-and-conquer recurrences under the rug. Let's address those issues now.

First, the Akra-Bazzi formula makes no use of boundary conditions. To see why, let's go back to Merge Sort. During the plug-and-chug analysis, we found that

$$T_n = nT_1 + n\log n - n + 1.$$

This expresses the nth term as a function of the first term, whose value is specified in a boundary condition. But notice that $T_n = \Theta(n \log n)$ for *every* value of T_1. The boundary condition doesn't matter!

This is the typical situation: *the asymptotic solution to a divide-and-conquer recurrence is independent of the boundary conditions*. Intuitively, if the bottom-level operation in a recursive algorithm takes, say, twice as long, then the overall running time will at most double. This matters in practice, but the factor of 2 is concealed by asymptotic notation. There are corner-case exceptions. For example, the solution to $T(n) = 2T(n/2)$ is either $\Theta(n)$ or zero, depending on whether $T(1)$ is zero. These cases are of little practical interest, so we won't consider them further.

There is a second nagging issue with divide-and-conquer recurrences that does not arise with linear recurrences. Specifically, dividing a problem of size n may create subproblems of non-integer size. For example, the Merge Sort recurrence contains the term $T(n/2)$. So what if n is 15? How long does it take to sort seven-and-a-half items? Previously, we dodged this issue by analyzing Merge Sort only when the size of the input was a power of 2. But then we don't know what happens for an input of size, say, 100.

Of course, a practical implementation of Merge Sort would split the input *approximately* in half, sort the halves recursively, and merge the results. For example, a list of 15 numbers would be split into lists of 7 and 8. More generally, a list of n numbers would be split into approximate halves of size $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$. So the maximum number of comparisons is actually given by this recurrence:

$$\begin{aligned} T(1) &= 0 \\ T(n) &= T\left(\lceil n/2 \rceil\right) + T\left(\lfloor n/2 \rfloor\right) + n - 1 \qquad (\text{for } n \geq 2). \end{aligned}$$

This may be rigorously correct, but the ceiling and floor operations make the recurrence hard to solve exactly.

Fortunately, *the asymptotic solution to a divide and conquer recurrence is unaffected by floors and ceilings*. More precisely, the solution is not changed by replacing a term $T(b_i n)$ with either $T(\lceil b_i n \rceil)$ or $T(\lfloor b_i n \rfloor)$. So leaving floors and

ceilings out of divide-and-conquer recurrences makes sense in many contexts; those are complications that make no difference.

10.4.3 The Akra-Bazzi Theorem

The Akra-Bazzi formula together with our assertions about boundary conditions and integrality all follow from the *Akra-Bazzi Theorem*, which is stated below.

Theorem 10.4.1 (Akra-Bazzi)**.** *Suppose that the function $T : \mathbb{R} \to \mathbb{R}$ satisfies the recurrence*

$$T(x) = \begin{cases} \text{is nonnegative and bounded} & \text{for } 0 \le x \le x_0 \\ \sum_{i=1}^{k} a_i T(b_i x + h_i(x)) + g(x) & \text{for } x > x_0 \end{cases}$$

where:

1. *$a_1, \ldots, a_k$ are positive constants.*
2. *$b_1, \ldots, b_k$ are constants between 0 and 1.*
3. *x_0 is large enough so that T is well-defined.*
4. *$g(x)$ is a nonnegative function such that $|g'(x)|$ is bounded by a polynomial.*
5. *$|h_i(x)| = O(x/\log^2 x)$.*

Then

$$T(x) = \Theta\left(x^p \left(1 + \int_1^x \frac{g(u)}{u^{p+1}}\, du\right)\right)$$

where p satisfies

$$\sum_{i=1}^{k} a_i b_i^p = 1.$$

The Akra-Bazzi theorem can be proved using a complicated induction argument, though we won't do that here. But let's at least go over the statement of the theorem.

All the recurrences we've considered were defined over the integers, and that is the common case. But the Akra-Bazzi theorem applies more generally to functions defined over the real numbers.

The Akra-Bazzi formula is lifted directed from the theorem statement, except that the recurrence in the theorem includes extra functions, h_i. These functions

extend the theorem to address floors, ceilings, and other small adjustments to the sizes of subproblems. The trick is illustrated by this combination of parameters

$$\begin{array}{lll} a_1 = 1 & b_1 = 1/2 & h_1(x) = \left\lceil \frac{x}{2} \right\rceil - \frac{x}{2} \\ a_2 = 1 & b_2 = 1/2 & h_2(x) = \left\lfloor \frac{x}{2} \right\rfloor - \frac{x}{2} \\ & g(x) = x - 1 & \end{array}$$

which corresponds the recurrence

$$\begin{aligned} T(x) &= 1 \cdot T\left(\frac{x}{2} + \left(\left\lceil \frac{x}{2} \right\rceil - \frac{x}{2}\right)\right) + \cdot T\left(\frac{x}{2} + \left(\left\lfloor \frac{x}{2} \right\rfloor - \frac{x}{2}\right)\right) + x - 1 \\ &= T\left(\left\lceil \frac{x}{2} \right\rceil\right) + T\left(\left\lfloor \frac{x}{2} \right\rfloor\right) + x - 1. \end{aligned}$$

This is the rigorously correct Merge Sort recurrence valid for all input sizes, complete with floor and ceiling operators. In this case, the functions $h_1(x)$ and $h_2(x)$ are both at most 1, which is easily $O(x/\log^2 x)$ as required by the theorem statement. These functions h_i do not affect—or even appear in—the asymptotic solution to the recurrence. This justifies our earlier claim that applying floor and ceiling operators to the size of a subproblem does not alter the asymptotic solution to a divide-and-conquer recurrence.

10.4.4 The Master Theorem

There is a special case of the Akra-Bazzi formula known as the Master Theorem that handles some of the recurrences that commonly arise in computer science. It is called the *Master* Theorem because it was proved long before Akra and Bazzi arrived on the scene and, for many years, it was the final word on solving divide-and-conquer recurrences. We include the Master Theorem here because it is still widely referenced in algorithms courses and you can use it without having to know anything about integration.

Theorem 10.4.2 (Master Theorem)**.** *Let T be a recurrence of the form*

$$T(n) = aT\left(\frac{n}{b}\right) + g(n).$$

Case 1: *If $g(n) = O\left(n^{\log_b(a)-\epsilon}\right)$ for some constant $\epsilon > 0$, then*

$$T(n) = \Theta\left(n^{\log_b(a)}\right).$$

Case 2: *If* $g(n) = \Theta\left(n^{\log_b(a)} \log^k(n)\right)$ *for some constant* $k \geq 0$*, then*

$$T(n) = \Theta\left(n^{\log_b(a)} \log^{k+1}(n)\right).$$

Case 3: *If* $g(n) = \Omega\left(n^{\log_b(a)+\epsilon}\right)$ *for some constant* $\epsilon > 0$ *and* $ag(n/b) < cg(n)$ *for some constant* $c < 1$ *and sufficiently large* n*, then*

$$T(n) = \Theta(g(n)).$$

The Master Theorem can be proved by induction on n or, more easily, as a corollary of Theorem 10.4.1. We will not include the details here.

10.4.5 Pitfalls with Asymptotic Notation and Induction

We've seen that asymptotic notation is quite useful, particularly in connection with recurrences. And induction is our favorite proof technique. But mixing the two is risky business; there is great potential for subtle errors and false conclusions!

False Claim. *If*

$$T(1) = 1 \quad \textit{and}$$
$$T(n) = 2T(n/2) + n,$$

then $T(n) = O(n)$.

The Akra-Bazzi theorem implies that the correct solution is $T(n) = \Theta(n \log n)$ and so this claim is false. But where does the following "proof" go astray?

Bogus proof. The proof is by strong induction. Let $P(n)$ be the proposition that $T(n) = O(n)$.

Base case: $P(1)$ is true because $T(1) = 1 = O(1)$.

Inductive step: For $n \geq 2$, assume $P(1)$, $P(2)$, ..., $P(n-1)$ to prove $P(n)$. We have

$$\begin{aligned} T(n) &= 2 \cdot T(n/2) + n \\ &= 2 \cdot O(n/2) + n \\ &= O(n). \end{aligned}$$

The first equation is the recurrence, the second uses the assumption $P(n/2)$, and the third is a simplification. ■

Where's the bug? The proof is already far off the mark in the second sentence, which defines the induction hypothesis. The statement "$T(n) = O(n)$" is either true or false; it's validity does not depend on a particular value of n. Thus the very idea of trying to prove that the statement holds for $n = 1, 2, \ldots$, is wrong-headed.

The safe way to reason inductively about asymptotic phenomena is to *work directly with the definition of the asymptotic notation*. Let's try to prove the claim above in this way. Remember that $f(n) = O(n)$ means that there exist constants n_0 and $c > 0$ such that $|f(n)| \leq cn$ for all $n \geq n_0$. (Let's not worry about the absolute value for now.) If all goes well, the proof attempt should fail in some blatantly obvious way, instead of in a subtle, hard-to-detect way like the earlier argument. Since our perverse goal is to demonstrate that the proof won't work for *any* constants n_0 and c, we'll leave these as variables and assume only that they're chosen so that the base case holds; that is, $T(n_0) \leq cn$.

Proof Attempt. We use strong induction. Let $P(n)$ be the proposition that $T(n) \leq cn$.

Base case: $P(n_0)$ is true, because $T(n_0) \leq cn$.

Inductive step: For $n > n_0$, assume that $P(n_0), \ldots, P(n-1)$ are true in order to prove $P(n)$. We reason as follows:

$$\begin{aligned} T(n) &= 2T(n/2) + n \\ &\leq 2c(n/2) + n \\ &= cn + n \\ &= (c+1)n \\ &\not\leq cn. \end{aligned}$$

■

The first equation is the recurrence. Then we use induction and simplify until the argument collapses!

In general, it is a good idea to stay away from asymptotic notation altogether while you are doing the induction. Once the induction is over and done with, then you can safely use big-Oh to simplify your result.

10.5 A Feel for Recurrences

We've guessed and verified, plugged and chugged, found roots, computed integrals, and solved linear systems and exponential equations. Now let's step back and look for some rules of thumb. What kinds of recurrences have what sorts of solutions?

Here are some recurrences we solved earlier:

	Recurrence	Solution
Towers of Hanoi	$T_n = 2T_{n-1} + 1$	$T_n \sim 2^n$
Merge Sort	$T_n = 2T_{n/2} + n - 1$	$T_n \sim n \log n$
Hanoi variation	$T_n = 2T_{n-1} + n$	$T_n \sim 2 \cdot 2^n$
Fibonacci	$T_n = T_{n-1} + T_{n-2}$	$T_n \sim (1.618\ldots)^{n+1}/\sqrt{5}$

Notice that the recurrence equations for Towers of Hanoi and Merge Sort are somewhat similar, but the solutions are radically different. Merge Sorting $n = 64$ items takes a few hundred comparisons, while moving $n = 64$ disks takes more than 10^{19} steps!

Each recurrence has one strength and one weakness. In the Towers of Hanoi, we broke a problem of size n into two subproblem of size $n - 1$ (which is large), but needed only 1 additional step (which is small). In Merge Sort, we divided the problem of size n into two subproblems of size $n/2$ (which is small), but needed $(n - 1)$ additional steps (which is large). Yet, Merge Sort is faster by a mile!

This suggests that *generating smaller subproblems is far more important to algorithmic speed than reducing the additional steps per recursive call*. For example, shifting to the variation of Towers of Hanoi increased the last term from $+1$ to $+n$, but the solution only doubled. And one of the two subproblems in the Fibonacci recurrence is just *slightly* smaller than in Towers of Hanoi (size $n - 2$ instead of $n-1$). Yet the solution is exponentially smaller! More generally, linear recurrences (which have big subproblems) typically have exponential solutions, while divide-and-conquer recurrences (which have small subproblems) usually have solutions bounded above by a polynomial.

All the examples listed above break a problem of size n into two smaller problems. How does the number of subproblems affect the solution? For example, suppose we increased the number of subproblems in Towers of Hanoi from 2 to 3, giving this recurrence:

$$T_n = 3T_{n-1} + 1$$

This increases the root of the characteristic equation from 2 to 3, which raises the solution exponentially, from $\Theta(2^n)$ to $\Theta(3^n)$.

Divide-and-conquer recurrences are also sensitive to the number of subproblems. For example, for this generalization of the Merge Sort recurrence:

$$\begin{aligned} T_1 &= 0 \\ T_n &= aT_{n/2} + n - 1. \end{aligned}$$

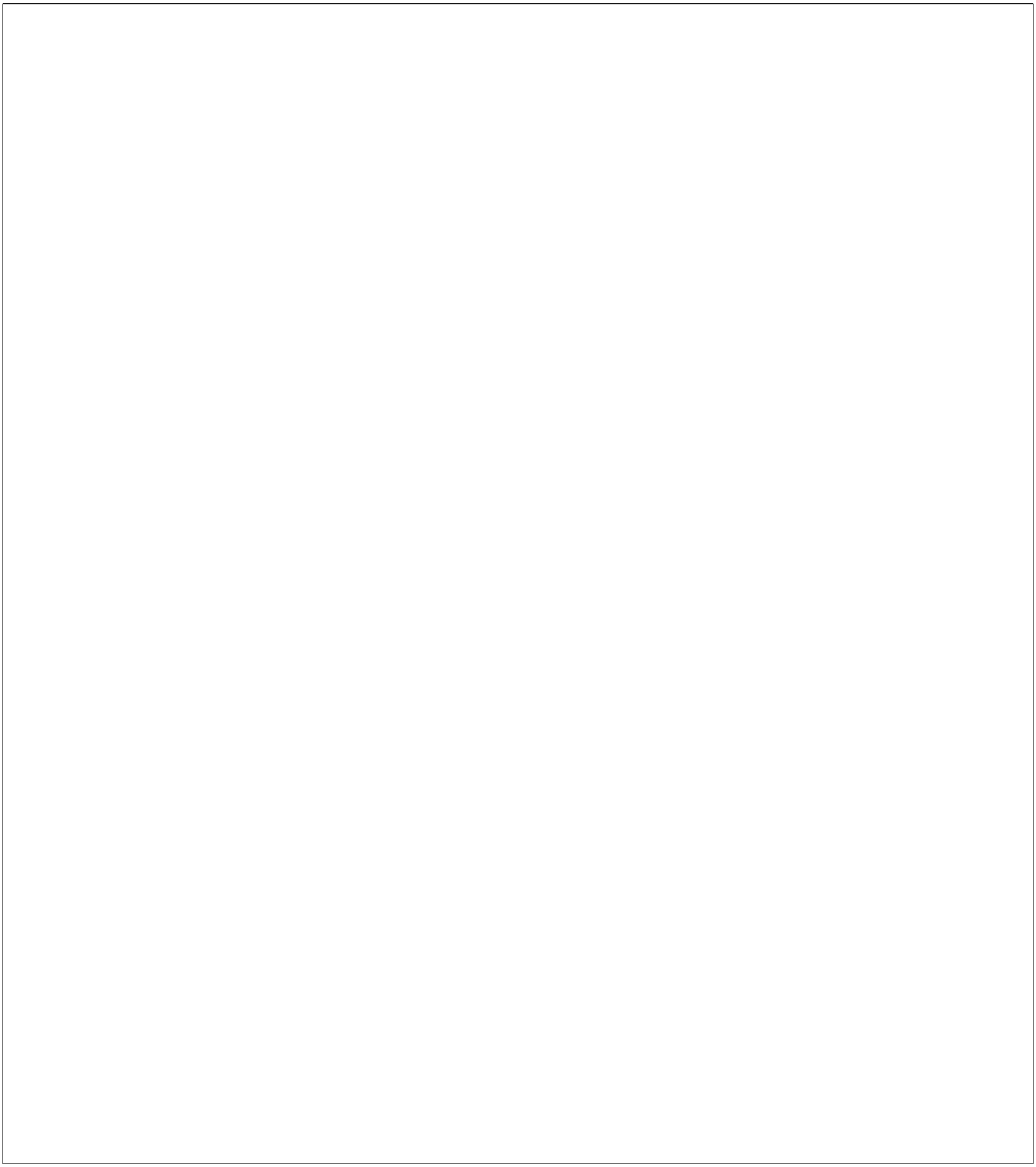
the Akra-Bazzi formula gives:

$$T_n = \begin{cases} \Theta(n) & \text{for } a < 2 \\ \Theta(n \log n) & \text{for } a = 2 \\ \Theta(n^{\log a}) & \text{for } a > 2. \end{cases}$$

So the solution takes on three completely different forms as a goes from 1.99 to 2.01!

How do boundary conditions affect the solution to a recurrence? We've seen that they are almost irrelevant for divide-and-conquer recurrences. For linear recurrences, the solution is usually dominated by an exponential whose base is determined by the number and size of subproblems. Boundary conditions matter greatly only when they give the dominant term a zero coefficient, which changes the asymptotic solution.

So now we have a rule of thumb! The performance of a recursive procedure is usually dictated by the size and number of subproblems, rather than the amount of work per recursive call or time spent at the base of the recursion. In particular, if subproblems are smaller than the original by an additive factor, the solution is most often exponential. But if the subproblems are only a fraction the size of the original, then the solution is typically bounded by a polynomial.



11 Cardinality Rules

11.1 Counting One Thing by Counting Another

How do you count the number of people in a crowded room? You could count heads, since for each person there is exactly one head. Alternatively, you could count ears and divide by two. Of course, you might have to adjust the calculation if someone lost an ear in a pirate raid or someone was born with three ears. The point here is that you can often *count one thing by counting another*, though some fudge factors may be required. This is a central theme of counting, from the easiest problems to the hardest.

In more formal terms, every counting problem comes down to determining the size of some set. The *size* or *cardinality* of a finite set S is the number of elements in S and it is denoted by $|S|$. In these terms, we're claiming that we can often find the size of one set by finding the size of a related set. We've already seen a general statement of this idea in the Mapping Rule of Theorem 7.2.1. Of particular interest here is part 3 of Theorem 7.2.1, where we state that if there is a bijection between two sets, then the sets have the same size. This important fact is commonly known as the *Bijection Rule*.

11.1.1 The Bijection Rule

Rule 11.1.1 (Bijection Rule)**.** *If there is a bijection $f : A \to B$ between A and B, then $|A| = |B|$.*

The Bijection Rule acts as a magnifier of counting ability; if you figure out the size of one set, then you can immediately determine the sizes of many other sets via bijections. For example, consider the two sets mentioned at the beginning of Part III:

$A =$ all ways to select a dozen doughnuts when five varieties are available

$B =$ all 16-bit sequences with exactly 4 ones

Let's consider a particular element of set A:

$$\underbrace{0\,0}_{\text{chocolate}} \quad \underbrace{}_{\text{lemon-filled}} \quad \underbrace{0\,0\,0\,0\,0\,0}_{\text{sugar}} \quad \underbrace{0\,0}_{\text{glazed}} \quad \underbrace{0\,0}_{\text{plain}}$$

We've depicted each doughnut with a 0 and left a gap between the different varieties. Thus, the selection above contains two chocolate doughnuts, no lemon-filled,

six sugar, two glazed, and two plain. Now let's put a 1 into each of the four gaps:

$$\underbrace{0\,0}_{\text{chocolate}}\quad 1\quad \underbrace{}_{\text{lemon-filled}}\quad 1\quad \underbrace{0\,0\,0\,0\,0\,0}_{\text{sugar}}\quad 1\quad \underbrace{0\,0}_{\text{glazed}}\quad 1\quad \underbrace{0\,0}_{\text{plain}}$$

We've just formed a 16-bit number with exactly 4 ones—an element of B!

This example suggests a bijection from set A to set B: map a dozen doughnuts consisting of:

$$c \text{ chocolate, } l \text{ lemon-filled, } s \text{ sugar, } g \text{ glazed, and } p \text{ plain}$$

to the sequence:

$$\underbrace{0\ldots0}_{c}\quad 1\quad \underbrace{0\ldots0}_{l}\quad 1\quad \underbrace{0\ldots0}_{s}\quad 1\quad \underbrace{0\ldots0}_{g}\quad 1\quad \underbrace{0\ldots0}_{p}$$

The resulting sequence always has 16 bits and exactly 4 ones, and thus is an element of B. Moreover, the mapping is a bijection; every such bit sequence is mapped to by exactly one order of a dozen doughnuts. Therefore, $|A| = |B|$ by the Bijection Rule!

This example demonstrates the magnifying power of the bijection rule. We managed to prove that two very different sets are actually the same size—even though we don't know exactly how big either one is. But as soon as we figure out the size of one set, we'll immediately know the size of the other.

This particular bijection might seem frighteningly ingenious if you've not seen it before. But you'll use essentially this same argument over and over, and soon you'll consider it routine.

11.2 Counting Sequences

The Bijection Rule lets us count one thing by counting another. This suggests a general strategy: get really good at counting just a *few* things and then use bijections to count *everything else*. This is the strategy we'll follow. In particular, we'll get really good at counting *sequences*. When we want to determine the size of some other set T, we'll find a bijection from T to a set of sequences S. Then we'll use our super-ninja sequence-counting skills to determine $|S|$, which immediately gives us $|T|$. We'll need to hone this idea somewhat as we go along, but that's pretty much the plan!

11.2.1 The Product Rule

The *Product Rule* gives the size of a product of sets. Recall that if $P_1, P_2, \ldots, P_n$ are sets, then

$$P_1 \times P_2 \times \ldots \times P_n$$

is the set of all sequences whose first term is drawn from P_1, second term is drawn from P_2 and so forth.

Rule 11.2.1 (Product Rule)**.** *If $P_1, P_2, \ldots P_n$ are sets, then:*

$$|P_1 \times P_2 \times \ldots \times P_n| = |P_1| \cdot |P_2| \cdots |P_n|$$

For example, suppose a *daily diet* consists of a breakfast selected from set B, a lunch from set L, and a dinner from set D where:

$$\begin{aligned} B &= \{\text{pancakes, bacon and eggs, bagel, Doritos}\} \\ L &= \{\text{burger and fries, garden salad, Doritos}\} \\ D &= \{\text{macaroni, pizza, frozen burrito, pasta, Doritos}\} \end{aligned}$$

Then $B \times L \times D$ is the set of all possible daily diets. Here are some sample elements:

$$\begin{gathered} (\text{pancakes, burger and fries, pizza}) \\ (\text{bacon and eggs, garden salad, pasta}) \\ (\text{Doritos, Doritos, frozen burrito}) \end{gathered}$$

The Product Rule tells us how many different daily diets are possible:

$$\begin{aligned} |B \times L \times D| &= |B| \cdot |L| \cdot |D| \\ &= 4 \cdot 3 \cdot 5 \\ &= 60. \end{aligned}$$

11.2.2 Subsets of an n-element Set

How many different subsets of an n-element set X are there? For example, the set $X = \{x_1, x_2, x_3\}$ has eight different subsets:

$$\begin{array}{cccc} \emptyset & \{x_1\} & \{x_2\} & \{x_1, x_2\} \\ \{x_3\} & \{x_1, x_3\} & \{x_2, x_3\} & \{x_1, x_2, x_3\}. \end{array}$$

There is a natural bijection from subsets of X to n-bit sequences. Let $x_1, x_2, \ldots, x_n$ be the elements of X. Then a particular subset of X maps to the sequence $(b_1, \ldots, b_n)$

where $b_i = 1$ if and only if x_i is in that subset. For example, if $n = 10$, then the subset $\{x_2, x_3, x_5, x_7, x_{10}\}$ maps to a 10-bit sequence as follows:

$$\begin{array}{rccccccccccc} \text{subset:} & \{ & & x_2, & x_3, & & x_5, & & x_7, & & & x_{10} & \} \\ \text{sequence:} & (& 0, & 1, & 1, & 0, & 1, & 0, & 1, & 0, & 0, & 1 &) \end{array}$$

We just used a bijection to transform the original problem into a question about sequences—*exactly according to plan!* Now if we answer the sequence question, then we've solved our original problem as well.

But how many different n-bit sequences are there? For example, there are 8 different 3-bit sequences:

$$\begin{array}{llll} (0,0,0) & (0,0,1) & (0,1,0) & (0,1,1) \\ (1,0,0) & (1,0,1) & (1,1,0) & (1,1,1) \end{array}$$

Well, we can write the set of all n-bit sequences as a product of sets:

$$\underbrace{\{0,1\} \times \{0,1\} \times \ldots \times \{0,1\}}_{n \text{ terms}} = \{0,1\}^n$$

Then Product Rule gives the answer:

$$\begin{aligned} |\{0,1\}^n| &= |\{0,1\}|^n \\ &= 2^n \end{aligned}$$

This means that the number of subsets of an n-element set X is also 2^n. We'll put this answer to use shortly.

11.2.3 The Sum Rule

Linus allocates his big sister Lucy a quota of 20 crabby days, 40 irritable days, and 60 generally surly days. On how many days can Lucy be out-of-sorts one way or another? Let set C be her crabby days, I be her irritable days, and S be the generally surly. In these terms, the answer to the question is $|C \cup I \cup S|$. Now assuming that she is permitted at most one bad quality each day, the size of this union of sets is given by the *Sum Rule*:

Rule 11.2.2 (Sum Rule)**.** *If $A_1, A_2, \ldots, A_n$ are* disjoint *sets, then:*

$$|A_1 \cup A_2 \cup \ldots \cup A_n| = |A_1| + |A_2| + \ldots + |A_n|$$

Thus, according to Linus' budget, Lucy can be out-of-sorts for:

$$\begin{aligned} |C \cup I \cup S| &= |C| + |I| + |S| \\ &= 20 + 40 + 60 \\ &= 120 \text{ days} \end{aligned}$$

Notice that the Sum Rule holds only for a union of *disjoint* sets. Finding the size of a union of intersecting sets is a more complicated problem that we'll take up later.

11.2.4 Counting Passwords

Few counting problems can be solved with a single rule. More often, a solution is a flurry of sums, products, bijections, and other methods. For example, the sum and product rules together are useful for solving problems involving passwords, telephone numbers, and license plates. For example, on a certain computer system, a valid password is a sequence of between six and eight symbols. The first symbol must be a letter (which can be lowercase or uppercase), and the remaining symbols must be either letters or digits. How many different passwords are possible?

Let's define two sets, corresponding to valid symbols in the first and subsequent positions in the password.

$$F = \{a, b, \ldots, z, A, B, \ldots, Z\}$$
$$S = \{a, b, \ldots, z, A, B, \ldots, Z, 0, 1, \ldots, 9\}$$

In these terms, the set of all possible passwords is:[1]

$$(F \times S^5) \cup (F \times S^6) \cup (F \times S^7)$$

Thus, the length-six passwords are in the set $F \times S^5$, the length-seven passwords are in $F \times S^6$, and the length-eight passwords are in $F \times S^7$. Since these sets are disjoint, we can apply the Sum Rule and count the total number of possible passwords as follows:

$$\begin{aligned}
&|(F \times S^5) \cup (F \times S^6) \cup (F \times S^7)| \\
&\quad = |F \times S^5| + |F \times S^6| + |F \times S^7| && \text{Sum Rule} \\
&\quad = |F| \cdot |S|^5 + |F| \cdot |S|^6 + |F| \cdot |S|^7 && \text{Product Rule} \\
&\quad = 52 \cdot 62^5 + 52 \cdot 62^6 + 52 \cdot 62^7 \\
&\quad \approx 1.8 \cdot 10^{14} \text{ different passwords.}
\end{aligned}$$

11.3 The Generalized Product Rule

We realize everyone has been working pretty hard this term, and we're considering awarding some prizes for *truly exceptional* coursework. Here are some possible

[1] The notation S^5 means $S \times S \times S \times S \times S$.

categories:

Best Administrative Critique We asserted that the quiz was closed-book. On the cover page, one strong candidate for this award wrote, "There is no book."

Awkward Question Award "Okay, the left sock, right sock, and pants are in an antichain, but how—even with assistance—could I put on all three at once?"

Best Collaboration Statement Inspired by a student who wrote "I worked alone" on Quiz 1.

In how many ways can, say, three different prizes be awarded to n people? This is easy to answer using our strategy of translating the problem about awards into a problem about sequences. Let P be the set of n people taking the course. Then there is a bijection from ways of awarding the three prizes to the set $P^3 ::= P \times P \times P$. In particular, the assignment:

"person x wins prize #1, y wins prize #2, and z wins prize #3"

maps to the sequence (x, y, z). By the Product Rule, we have $|P^3| = |P|^3 = n^3$, so there are n^3 ways to award the prizes to a class of n people.

But what if the three prizes must be awarded to *different* students? As before, we could map the assignment

"person x wins prize #1, y wins prize #2, and z wins prize #3"

to the triple $(x, y, z) \in P^3$. But this function is *no longer a bijection*. For example, no valid assignment maps to the triple (Dave, Dave, Becky) because Dave is not allowed to receive two awards. However, there *is* a bijection from prize assignments to the set:

$$S = \{(x, y, z) \in P^3 \mid x, y, \text{and } z \text{ are different people}\}$$

This reduces the original problem to a problem of counting sequences. Unfortunately, the Product Rule is of no help in counting sequences of this type because the entries depend on one another; in particular, they must all be different. However, a slightly sharper tool does the trick.

Rule 11.3.1 (Generalized Product Rule)**.** *Let S be a set of length-k sequences. If there are:*

- *n_1 possible first entries,*
- *n_2 possible second entries for each first entry,*

- *n_3 possible third entries for each combination of first and second entries, etc.*

then:

$$|S| = n_1 \cdot n_2 \cdot n_3 \cdots n_k$$

In the awards example, S consists of sequences (x, y, z). There are n ways to choose x, the recipient of prize #1. For each of these, there are $n-1$ ways to choose y, the recipient of prize #2, since everyone except for person x is eligible. For each combination of x and y, there are $n-2$ ways to choose z, the recipient of prize #3, because everyone except for x and y is eligible. Thus, according to the Generalized Product Rule, there are

$$|S| = n \cdot (n-1) \cdot (n-2)$$

ways to award the 3 prizes to different people.

11.3.1 Defective Dollar Bills

A dollar bill is *defective* if some digit appears more than once in the 8-digit serial number. If you check your wallet, you'll be sad to discover that defective bills are all-too-common. In fact, how common are *nondefective* bills? Assuming that the digit portions of serial numbers all occur equally often, we could answer this question by computing

$$\text{fraction of nondefective bills} = \frac{|\{\text{serial \#'s with all digits different}\}|}{|\{\text{serial numbers}\}|}. \tag{11.1}$$

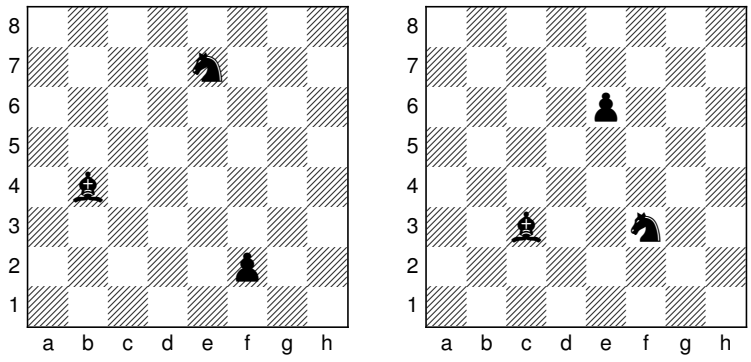
Let's first consider the denominator. Here there are no restrictions; there are are 10 possible first digits, 10 possible second digits, 10 third digits, and so on. Thus, the total number of 8-digit serial numbers is 10^8 by the Product Rule.

Next, let's turn to the numerator. Now we're not permitted to use any digit twice. So there are still 10 possible first digits, but only 9 possible second digits, 8 possible third digits, and so forth. Thus, by the Generalized Product Rule, there are

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 = \frac{10!}{2} = 1{,}814{,}400$$

serial numbers with all digits different. Plugging these results into Equation 11.1, we find:

$$\text{fraction of nondefective bills} = \frac{1{,}814{,}400}{100{,}000{,}000} = 1.8144\%$$



(a) valid (b) invalid

Figure 11.1 Two ways of placing a pawn (♙), a knight (♘), and a bishop (♗) on a chessboard. The configuration shown in (b) is invalid because the bishop and the knight are in the same row.

11.3.2 A Chess Problem

In how many different ways can we place a pawn (P), a knight (N), and a bishop (B) on a chessboard so that no two pieces share a row or a column? A valid configuration is shown in Figure 11.1(a), and an invalid configuration is shown in Figure 11.1(b).

First, we map this problem about chess pieces to a question about sequences. There is a bijection from configurations to sequences

$$(r_P, c_P, r_N, c_N, r_B, c_B)$$

where r_P, r_N, and r_B are distinct rows and c_P, c_N, and c_B are distinct columns. In particular, r_P is the pawn's row, c_P is the pawn's column, r_N is the knight's row, etc. Now we can count the number of such sequences using the Generalized Product Rule:

- r_P is one of 8 rows
- c_P is one of 8 columns
- r_N is one of 7 rows (any one but r_P)
- c_N is one of 7 columns (any one but c_P)
- r_B is one of 6 rows (any one but r_P or r_N)
- c_B is one of 6 columns (any one but c_P or c_N)

Thus, the total number of configurations is $(8 \cdot 7 \cdot 6)^2$.

11.3.3 Permutations

A *permutation* of a set S is a sequence that contains every element of S exactly once. For example, here are all the permutations of the set $\{a, b, c\}$:

$$\begin{array}{lll} (a,b,c) & (a,c,b) & (b,a,c) \\ (b,c,a) & (c,a,b) & (c,b,a) \end{array}$$

How many permutations of an n-element set are there? Well, there are n choices for the first element. For each of these, there are $n-1$ remaining choices for the second element. For every combination of the first two elements, there are $n-2$ ways to choose the third element, and so forth. Thus, there are a total of

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

permutations of an n-element set. In particular, this formula says that there are $3! = 6$ permutations of the 3-element set $\{a, b, c\}$, which is the number we found above.

Permutations will come up again in this course approximately 1.6 bazillion times. In fact, permutations are the reason why factorial comes up so often and why we taught you Stirling's approximation:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

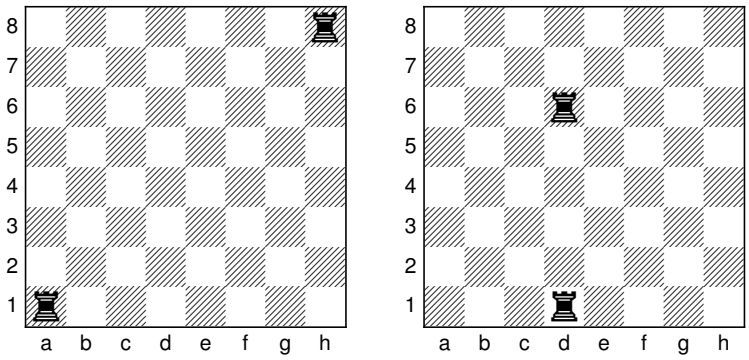
11.4 The Division Rule

Counting ears and dividing by two is a silly way to count the number of people in a room, but this approach is representative of a powerful counting principle.

A *k-to-1 function* maps exactly k elements of the domain to every element of the codomain. For example, the function mapping each ear to its owner is 2-to-1. Similarly, the function mapping each finger to its owner is 10-to-1, and the function mapping each finger and toe to its owner is 20-to-1. The general rule is:

Rule 11.4.1 (Division Rule)**.** *If $f : A \to B$ is k-to-1, then $|A| = k \cdot |B|$.*

For example, suppose A is the set of ears in the room and B is the set of people. There is a 2-to-1 mapping from ears to people, so by the Division Rule, $|A| = 2 \cdot |B|$. Equivalently, $|B| = |A|/2$, expressing what we knew all along: the number of people is half the number of ears. Unlikely as it may seem, many counting problems are made much easier by initially counting every item multiple times and then correcting the answer using the Division Rule. Let's look at some examples.



(a) valid (b) invalid

Figure 11.2 Two ways to place 2 rooks (♖) on a chessboard. The configuration in (b) is invalid because the rooks are in the same column.

11.4.1 Another Chess Problem

In how many different ways can you place two identical rooks on a chessboard so that they do not share a row or column? A valid configuration is shown in Figure 11.2(a), and an invalid configuration is shown in Figure 11.2(b).

Let A be the set of all sequences

$$(r_1, c_1, r_2, c_2)$$

where r_1 and r_2 are distinct rows and c_1 and c_2 are distinct columns. Let B be the set of all valid rook configurations. There is a natural function f from set A to set B; in particular, f maps the sequence (r_1, c_1, r_2, c_2) to a configuration with one rook in row r_1, column c_1 and the other rook in row r_2, column c_2.

But now there's a snag. Consider the sequences:

$$(1, 1, 8, 8) \qquad \text{and} \qquad (8, 8, 1, 1)$$

The first sequence maps to a configuration with a rook in the lower-left corner and a rook in the upper-right corner. The second sequence maps to a configuration with a rook in the upper-right corner and a rook in the lower-left corner. The problem is that those are two different ways of describing the *same* configuration! In fact, this arrangement is shown in Figure 11.2(a).

More generally, the function f maps exactly two sequences to *every* board configuration; that is f is a 2-to-1 function. Thus, by the quotient rule, $|A| = 2 \cdot |B|$.

Rearranging terms gives:

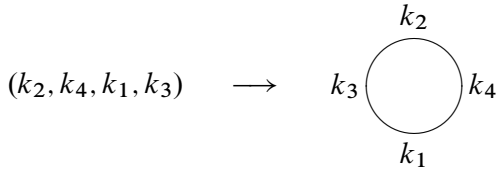
$$|B| = \frac{|A|}{2} = \frac{(8 \cdot 7)^2}{2}.$$

On the second line, we've computed the size of A using the General Product Rule just as in the earlier chess problem.

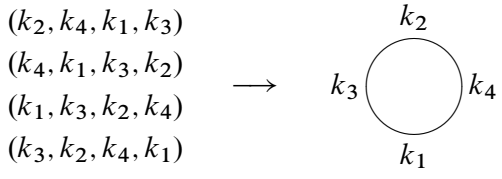
11.4.2 Knights of the Round Table

In how many ways can King Arthur seat n different knights at his round table? Two seatings are considered equivalent if one can be obtained from the other by rotation. For example, the following two arrangements are equivalent:

Let A be all the permutations of the knights, and let B be the set of all possible seating arrangements at the round table. We can map each permutation in set A to a circular seating arrangement in set B by seating the first knight in the permutation anywhere, putting the second knight to his left, the third knight to the left of the second, and so forth all the way around the table. For example:



This mapping is actually an n-to-1 function from A to B, since all n cyclic shifts of the original sequence map to the same seating arrangement. In the example, $n = 4$ different sequences map to the same seating arrangement:



Therefore, by the division rule, the number of circular seating arrangements is:

$$|B| = \frac{|A|}{n} = \frac{n!}{n} = (n-1)!$$

Note that $|A| = n!$ since there are $n!$ permutations of n knights.

11.5 Counting Subsets

How many k-element subsets of an n-element set are there? This question arises all the time in various guises:

- In how many ways can I select 5 books from my collection of 100 to bring on vacation?
- How many different 13-card Bridge hands can be dealt from a 52-card deck?
- In how many ways can I select 5 toppings for my pizza if there are 14 available toppings?

This number comes up so often that there is a special notation for it:

$$\binom{n}{k} ::= \text{ the number of } k\text{-element subsets of an } n\text{-element set.}$$

The expression $\binom{n}{k}$ is read "n choose k." Now we can immediately express the answers to all three questions above:

- I can select 5 books from 100 in $\binom{100}{5}$ ways.
- There are $\binom{52}{13}$ different Bridge hands.
- There are $\binom{14}{5}$ different 5-topping pizzas, if 14 toppings are available.

11.5.1 The Subset Rule

We can derive a simple formula for the n-choose-k number using the Division Rule. We do this by mapping any permutation of an n-element set $\{a_1, \ldots, a_n\}$ into a k-element subset simply by taking the first k elements of the permutation. That is, the permutation $a_1 a_2 \ldots a_n$ will map to the set $\{a_1, a_2, \ldots, a_k\}$.

Notice that any other permutation with the same first k elements $a_1, \ldots, a_k$ *in any order* and the same remaining elements $n - k$ elements *in any order* will also map to this set. What's more, a permutation can only map to $\{a_1, a_2, \ldots, a_k\}$ if its first k elements are the elements $a_1, \ldots, a_k$ in some order. Since there are $k!$ possible permutations of the first k elements and $(n-k)!$ permutations of the remaining elements, we conclude from the Product Rule that exactly $k!(n-k)!$ permutations of the n-element set map to the the particular subset, S. In other words, the mapping from permutations to k-element subsets is $k!(n-k)!$-to-1.

But we know there are $n!$ permutations of an n-element set, so by the Division Rule, we conclude that

$$n! = k!(n-k)!\binom{n}{k}$$

which proves:

Rule 11.5.1 (Subset Rule)**.** *The number of k-element subsets of an n-element set is*

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

Notice that this works even for 0-element subsets: $n!/0!n! = 1$. Here we use the fact that 0! is a *product* of 0 terms, which by convention[2] equals 1.

11.5.2 Bit Sequences

How many n-bit sequences contain exactly k ones? We've already seen the straightforward bijection between subsets of an n-element set and n-bit sequences. For example, here is a 3-element subset of $\{x_1, x_2, \ldots, x_8\}$ and the associated 8-bit sequence:

$$\begin{array}{llllllllll} \{ & x_1, & & & x_4, & x_5 & & & & \} \\ (& 1, & 0, & 0, & 1, & 1, & 0, & 0, & 0 &) \end{array}$$

Notice that this sequence has exactly 3 ones, each corresponding to an element of the 3-element subset. More generally, the n-bit sequences corresponding to a k-element subset will have exactly k ones. So by the Bijection Rule,

[2] We don't use it here, but a *sum* of zero terms equals 0.

The number of n-bit sequences with exactly k ones is $\binom{n}{k}$.

11.6 Sequences with Repetitions

11.6.1 Sequences of Subsets

Choosing a k-element subset of an n-element set is the same as splitting the set into a pair of subsets: the first subset of size k and the second subset consisting of the remaining $n - k$ elements. So the Subset Rule can be understood as a rule for counting the number of such splits into pairs of subsets.

We can generalize this to splits into more than two subsets. Namely, let A be an n-element set and $k_1, k_2, \ldots, k_m$ be nonnegative integers whose sum is n. A $(k_1, k_2, \ldots, k_m)$-*split of* A is a sequence

$$(A_1, A_2, \ldots, A_m)$$

where the A_i are disjoint subsets of A and $|A_i| = k_i$ for $i = 1, \ldots, m$.

Rule 11.6.1 (Subset Split Rule)**.** *The number of* $(k_1, k_2, \ldots, k_m)$*-splits of an* n*-element set is*

$$\binom{n}{k_1, \ldots, k_m} ::= \frac{n!}{k_1!\, k_2!\, \cdots k_m!}$$

The proof of this Rule is essentially the same as for the Subset Rule. Namely, we map any permutation $a_1 a_2 \ldots a_n$ of an n-element set A into a $(k_1, k_2, \ldots, k_m)$-split by letting the 1st subset in the split be the first k_1 elements of the permutation, the 2nd subset of the split be the next k_2 elements, ..., and the mth subset of the split be the final k_m elements of the permutation. This map is a $k_1!\, k_2!\, \cdots\, k_m!$-to-1 function from the $n!$ permutations to the $(k_1, k_2, \ldots, k_m)$-splits of A, and the Subset Split Rule now follows from the Division Rule.

11.6.2 The Bookkeeper Rule

We can also generalize our count of n-bit sequences with k ones to counting sequences of n letters over an alphabet with more than two letters. For example, how many sequences can be formed by permuting the letters in the 10-letter word BOOKKEEPER?

Notice that there are 1 B, 2 O's, 2 K's, 3 E's, 1 P, and 1 R in BOOKKEEPER. This leads to a straightforward bijection between permutations of BOOKKEEPER and

(1,2,2,3,1,1)-splits of $\{1, 2, \ldots, 10\}$. Namely, map a permutation to the sequence of sets of positions where each of the different letters occur.

For example, in the permutation BOOKKEEPER itself, the B is in the 1st position, the O's occur in the 2nd and 3rd positions, K's in 4th and 5th, the E's in the 6th, 7th and 9th, P in the 8th, and R is in the 10th position. So BOOKKEEPER maps to

$$(\{1\}, \{2, 3\}, \{4, 5\}, \{6, 7, 9\}, \{8\}, \{10\}).$$

From this bijection and the Subset Split Rule, we conclude that the number of ways to rearrange the letters in the word BOOKKEEPER is:

$$\frac{\overbrace{10!}^{\text{total letters}}}{\underbrace{1!}_{\text{B's}}\underbrace{2!}_{\text{O's}}\underbrace{2!}_{\text{K's}}\underbrace{3!}_{\text{E's}}\underbrace{1!}_{\text{P's}}\underbrace{1!}_{\text{R's}}}$$

This example generalizes directly to an exceptionally useful counting principle which we will call the

Rule 11.6.2 (Bookkeeper Rule)**.** *Let $l_1, \ldots, l_m$ be distinct elements. The number of sequences with k_1 occurrences of l_1, and k_2 occurrences of l_2, ..., and k_m occurrences of l_m is*

$$\frac{(k_1 + k_2 + \ldots + k_m)!}{k_1!\, k_2!\, \ldots\, k_m!}$$

For example, suppose you are planning a 20-mile walk, which should include 5 northward miles, 5 eastward miles, 5 southward miles, and 5 westward miles. How many different walks are possible?

There is a bijection between such walks and sequences with 5 N's, 5 E's, 5 S's, and 5 W's. By the Bookkeeper Rule, the number of such sequences is:

$$\frac{20!}{5!^4}.$$

11.6.3 The Binomial Theorem

Counting gives insight into one of the basic theorems of algebra. A *binomial* is a sum of two terms, such as $a + b$. Now consider its 4th power, $(a + b)^4$.

If we multiply out this 4th power expression completely, we get

$$\begin{aligned}(a+b)^4 = \quad & aaaa + aaab + aaba + aabb \\ & + abaa + abab + abba + abbb \\ & + baaa + baab + baba + babb \\ & + bbaa + bbab + bbba + bbbb\end{aligned}$$

Notice that there is one term for every sequence of a's and b's. So there are 2^4 terms, and the number of terms with k copies of b and $n-k$ copies of a is:

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

by the Bookkeeper Rule. Hence, the coefficient of $a^{n-k}b^k$ is $\binom{n}{k}$. So for $n = 4$, this means:

$$(a+b)^4 = \binom{4}{0}\cdot a^4b^0 + \binom{4}{1}\cdot a^3b^1 + \binom{4}{2}\cdot a^2b^2 + \binom{4}{3}\cdot a^1b^3 + \binom{4}{4}\cdot a^0b^4$$

In general, this reasoning gives the Binomial Theorem:

Theorem 11.6.3 (Binomial Theorem)**.** *For all $n \in \mathbb{N}$ and $a, b \in \mathbb{R}$:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}b^k$$

The expression $\binom{n}{k}$ is often called a "binomial coefficient" in honor of its appearance here.

This reasoning about binomials extends nicely to *multinomials*, which are sums of two or more terms. For example, suppose we wanted the coefficient of

$$bo^2k^2e^3pr$$

in the expansion of $(b+o+k+e+p+r)^{10}$. Each term in this expansion is a product of 10 variables where each variable is one of b, o, k, e, p, or r. Now, the coefficient of $bo^2k^2e^3pr$ is the number of those terms with exactly 1 b, 2 o's, 2 k's, 3 e's, 1 p, and 1 r. And the number of such terms is precisely the number of rearrangements of the word BOOKKEEPER:

$$\binom{10}{1,2,2,3,1,1} = \frac{10!}{1!\,2!\,2!\,3!\,1!\,1!}.$$

The expression on the left is called a "multinomial coefficient." This reasoning extends to a general theorem.

Definition 11.6.4. For $n, k_1, \ldots, k_m \in \mathbb{N}$, such that $k_1 + k_2 + \cdots + k_m = n$, define the *multinomial coefficient*

$$\binom{n}{k_1, k_2, \ldots, k_m} ::= \frac{n!}{k_1!\,k_2!\,\ldots k_m!}.$$

Theorem 11.6.5 (Multinomial Theorem)**.** *For all* $n \in \mathbb{N}$,

$$(z_1 + z_2 + \cdots + z_m)^n = \sum_{\substack{k_1,\ldots,k_m \in \mathbb{N} \\ k_1+\cdots+k_m=n}} \binom{n}{k_1, k_2, \ldots, k_m} z_1^{k_1} z_2^{k_2} \cdots z_m^{k_m}.$$

You'll be better off remembering the reasoning behind the Multinomial Theorem rather than this ugly formal statement.

11.6.4 A Word about Words

Someday you might refer to the Subset Split Rule or the Bookkeeper Rule in front of a roomful of colleagues and discover that they're all staring back at you blankly. This is not because they're dumb, but rather because we made up the name "Bookkeeper Rule". However, the rule is excellent and the name is apt, so we suggest that you play through: "You know? The Bookkeeper Rule? Don't you guys know *anything???*"

The Bookkeeper Rule is sometimes called the "formula for permutations with indistinguishable objects." The size k subsets of an n-element set are sometimes called *k-combinations*. Other similar-sounding descriptions are "combinations with repetition, permutations with repetition, r-permutations, permutations with indistinguishable objects," and so on. However, the counting rules we've taught you are sufficient to solve all these sorts of problems without knowing this jargon, so we won't burden you with it.

11.7 Counting Practice: Poker Hands

Five-Card Draw is a card game in which each player is initially dealt a *hand* consisting of 5 cards from a deck of 52 cards.[3] (Then the game gets complicated, but let's not worry about that.) The number of different hands in Five-Card Draw is the

[3] There are 52 cards in a standard deck. Each card has a *suit* and a *rank*. There are four suits:

♠ (spades) ♡ (hearts) ♣ (clubs) ◇ (diamonds)

And there are 13 ranks, listed here from lowest to highest:

$$\overset{\text{Ace}}{A}, 2, 3, 4, 5, 6, 7, 8, 9, \overset{\text{Jack}}{J}, \overset{\text{Queen}}{Q}, \overset{\text{King}}{K}.$$

Thus, for example, 8♡ is the 8 of hearts and A♠ is the ace of spades.

number of 5-element subsets of a 52-element set, which is

$$\binom{52}{5} = 2,598,960.$$

Let's get some counting practice by working out the number of hands with various special properties.

11.7.1 Hands with a Four-of-a-Kind

A *Four-of-a-Kind* is a set of four cards with the same rank. How many different hands contain a Four-of-a-Kind? Here are a couple examples:

$$\{8\spadesuit,\ 8\diamondsuit,\ Q\heartsuit,\ 8\heartsuit,\ 8\clubsuit\}$$
$$\{A\clubsuit,\ 2\clubsuit,\ 2\heartsuit,\ 2\diamondsuit,\ 2\spadesuit\}$$

As usual, the first step is to map this question to a sequence-counting problem. A hand with a Four-of-a-Kind is completely described by a sequence specifying:

1. The rank of the four cards.
2. The rank of the extra card.
3. The suit of the extra card.

Thus, there is a bijection between hands with a Four-of-a-Kind and sequences consisting of two distinct ranks followed by a suit. For example, the three hands above are associated with the following sequences:

$$(8, Q, \heartsuit) \leftrightarrow \{8\spadesuit,\ 8\diamondsuit,\ 8\heartsuit,\ 8\clubsuit,\ Q\heartsuit\}$$
$$(2, A, \clubsuit) \leftrightarrow \{2\clubsuit,\ 2\heartsuit,\ 2\diamondsuit,\ 2\spadesuit,\ A\clubsuit\}$$

Now we need only count the sequences. There are 13 ways to choose the first rank, 12 ways to choose the second rank, and 4 ways to choose the suit. Thus, by the Generalized Product Rule, there are $13 \cdot 12 \cdot 4 = 624$ hands with a Four-of-a-Kind. This means that only 1 hand in about 4165 has a Four-of-a-Kind. Not surprisingly, Four-of-a-Kind is considered to be a very good poker hand!

11.7.2 Hands with a Full House

A *Full House* is a hand with three cards of one rank and two cards of another rank. Here are some examples:

$$\{2\spadesuit,\ 2\clubsuit,\ 2\diamondsuit,\ J\clubsuit,\ J\diamondsuit\}$$
$$\{5\diamondsuit,\ 5\clubsuit,\ 5\heartsuit,\ 7\heartsuit,\ 7\clubsuit\}$$

Again, we shift to a problem about sequences. There is a bijection between Full Houses and sequences specifying:

1. The rank of the triple, which can be chosen in 13 ways.
2. The suits of the triple, which can be selected in $\binom{4}{3}$ ways.
3. The rank of the pair, which can be chosen in 12 ways.
4. The suits of the pair, which can be selected in $\binom{4}{2}$ ways.

The example hands correspond to sequences as shown below:

$$(2, \{\spadesuit, \clubsuit, \diamondsuit\}, J, \{\clubsuit, \diamondsuit\}) \leftrightarrow \{2\spadesuit,\ 2\clubsuit,\ 2\diamondsuit,\ J\clubsuit,\ J\diamondsuit\}$$
$$(5, \{\diamondsuit, \clubsuit, \heartsuit\}, 7, \{\heartsuit, \clubsuit\}) \leftrightarrow \{5\diamondsuit,\ 5\clubsuit,\ 5\heartsuit,\ 7\heartsuit,\ 7\clubsuit\}$$

By the Generalized Product Rule, the number of Full Houses is:

$$13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}.$$

We're on a roll—but we're about to hit a speed bump.

11.7.3 Hands with Two Pairs

How many hands have *Two Pairs*; that is, two cards of one rank, two cards of another rank, and one card of a third rank? Here are examples:

$$\{3\diamondsuit,\ 3\spadesuit,\ Q\diamondsuit,\ Q\heartsuit,\ A\clubsuit\}$$
$$\{9\heartsuit,\ 9\diamondsuit,\ 5\heartsuit,\ 5\clubsuit,\ K\spadesuit\}$$

Each hand with Two Pairs is described by a sequence consisting of:

1. The rank of the first pair, which can be chosen in 13 ways.
2. The suits of the first pair, which can be selected $\binom{4}{2}$ ways.

3. The rank of the second pair, which can be chosen in 12 ways.
4. The suits of the second pair, which can be selected in $\binom{4}{2}$ ways.
5. The rank of the extra card, which can be chosen in 11 ways.
6. The suit of the extra card, which can be selected in $\binom{4}{1} = 4$ ways.

Thus, it might appear that the number of hands with Two Pairs is:

$$13 \cdot \binom{4}{2} \cdot 12 \cdot \binom{4}{2} \cdot 11 \cdot 4.$$

Wrong answer! The problem is that there is *not* a bijection from such sequences to hands with Two Pairs. This is actually a 2-to-1 mapping. For example, here are the pairs of sequences that map to the hands given above:

$$\begin{array}{lll} (3, \{\diamondsuit, \spadesuit\}, Q, \{\diamondsuit, \heartsuit\}, A, \clubsuit) & \searrow & \\ & & \{3\diamondsuit,\ 3\spadesuit,\ Q\diamondsuit,\ Q\heartsuit,\ A\clubsuit\} \\ (Q, \{\diamondsuit, \heartsuit\}, 3, \{\diamondsuit, \spadesuit\}, A, \clubsuit) & \nearrow & \\ & & \\ (9, \{\heartsuit, \diamondsuit\}, 5, \{\heartsuit, \clubsuit\}, K, \spadesuit) & \searrow & \\ & & \{9\heartsuit,\ 9\diamondsuit,\ 5\heartsuit,\ 5\clubsuit,\ K\spadesuit\} \\ (5, \{\heartsuit, \clubsuit\}, 9, \{\heartsuit, \diamondsuit\}, K, \spadesuit) & \nearrow & \end{array}$$

The problem is that nothing distinguishes the first pair from the second. A pair of 5's and a pair of 9's is the same as a pair of 9's and a pair of 5's. We avoided this difficulty in counting Full Houses because, for example, a pair of 6's and a triple of kings is different from a pair of kings and a triple of 6's.

We ran into precisely this difficulty last time, when we went from counting arrangements of *different* pieces on a chessboard to counting arrangements of two *identical* rooks. The solution then was to apply the Division Rule, and we can do the same here. In this case, the Division rule says there are twice as many sequences as hands, so the number of hands with Two Pairs is actually:

$$\frac{13 \cdot \binom{4}{2} \cdot 12 \cdot \binom{4}{2} \cdot 11 \cdot 4}{2}.$$

Another Approach

The preceding example was disturbing! One could easily overlook the fact that the mapping was 2-to-1 on an exam, fail the course, and turn to a life of crime. You can make the world a safer place in two ways:

1. Whenever you use a mapping $f : A \to B$ to translate one counting problem to another, check that the same number elements in A are mapped to each element in B. If k elements of A map to each of element of B, then apply the Division Rule using the constant k.

2. As an extra check, try solving the same problem in a different way. Multiple approaches are often available—and all had better give the same answer! (Sometimes different approaches give answers that *look* different, but turn out to be the same after some algebra.)

We already used the first method; let's try the second. There is a bijection between hands with two pairs and sequences that specify:

1. The ranks of the two pairs, which can be chosen in $\binom{13}{2}$ ways.

2. The suits of the lower-rank pair, which can be selected in $\binom{4}{2}$ ways.

3. The suits of the higher-rank pair, which can be selected in $\binom{4}{2}$ ways.

4. The rank of the extra card, which can be chosen in 11 ways.

5. The suit of the extra card, which can be selected in $\binom{4}{1} = 4$ ways.

For example, the following sequences and hands correspond:

$$(\{3, Q\}, \{\diamondsuit, \spadesuit\}, \{\diamondsuit, \heartsuit\}, A, \clubsuit) \leftrightarrow \{3\diamondsuit,\ 3\spadesuit,\ Q\diamondsuit,\ Q\heartsuit,\ A\clubsuit\}$$
$$(\{9, 5\}, \{\heartsuit, \clubsuit\}, \{\heartsuit, \diamondsuit\}, K, \spadesuit) \leftrightarrow \{9\heartsuit,\ 9\diamondsuit,\ 5\heartsuit,\ 5\clubsuit,\ K\spadesuit\}$$

Thus, the number of hands with two pairs is:

$$\binom{13}{2} \cdot \binom{4}{2} \cdot \binom{4}{2} \cdot 11 \cdot 4.$$

This is the same answer we got before, though in a slightly different form.

11.7.4 Hands with Every Suit

How many hands contain at least one card from every suit? Here is an example of such a hand:

$$\{7\diamondsuit,\ K\clubsuit,\ 3\diamondsuit,\ A\heartsuit,\ 2\spadesuit\}$$

Each such hand is described by a sequence that specifies:

1. The ranks of the diamond, the club, the heart, and the spade, which can be selected in $13 \cdot 13 \cdot 13 \cdot 13 = 13^4$ ways.

2. The suit of the extra card, which can be selected in 4 ways.

3. The rank of the extra card, which can be selected in 12 ways.

For example, the hand above is described by the sequence:

$$(7, K, A, 2, \diamondsuit, 3) \leftrightarrow \{7\diamondsuit,\ K\clubsuit,\ A\heartsuit,\ 2\spadesuit,\ 3\diamondsuit\}.$$

Are there other sequences that correspond to the same hand? There is one more! We could equally well regard either the $3\diamondsuit$ or the $7\diamondsuit$ as the extra card, so this is actually a 2-to-1 mapping. Here are the two sequences corresponding to the example hand:

$$\begin{array}{ll} (7, K, A, 2, \diamondsuit, 3) & \searrow \\ & \{7\diamondsuit,\ K\clubsuit,\ A\heartsuit,\ 2\spadesuit,\ 3\diamondsuit\} \\ (3, K, A, 2, \diamondsuit, 7) & \nearrow \end{array}$$

Therefore, the number of hands with every suit is:

$$\frac{13^4 \cdot 4 \cdot 12}{2}.$$

11.8 Inclusion-Exclusion

How big is a union of sets? For example, suppose there are 60 math majors, 200 EECS majors, and 40 physics majors. How many students are there in these three departments? Let M be the set of math majors, E be the set of EECS majors, and P be the set of physics majors. In these terms, we're asking for $|M \cup E \cup P|$.

The Sum Rule says that if M, E, and P are disjoint, then the sum of their sizes is

$$|M \cup E \cup P| = |M| + |E| + |P|.$$

However, the sets M, E, and P might *not* be disjoint. For example, there might be a student majoring in both math and physics. Such a student would be counted twice on the right side of this equation, once as an element of M and once as an element of P. Worse, there might be a triple-major[4] counted *three* times on the right side!

Our most-complicated counting rule determines the size of a union of sets that are not necessarily disjoint. Before we state the rule, let's build some intuition by considering some easier special cases: unions of just two or three sets.

[4] . . . though not at MIT anymore.

11.8.1 Union of Two Sets

For two sets, S_1 and S_2, the *Inclusion-Exclusion Rule* is that the size of their union is:

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2| \tag{11.2}$$

Intuitively, each element of S_1 is accounted for in the first term, and each element of S_2 is accounted for in the second term. Elements in *both* S_1 and S_2 are counted *twice*—once in the first term and once in the second. This double-counting is corrected by the final term.

11.8.2 Union of Three Sets

So how many students are there in the math, EECS, and physics departments? In other words, what is $|M \cup E \cup P|$ if:

$$\begin{aligned} |M| &= 60 \\ |E| &= 200 \\ |P| &= 40. \end{aligned}$$

The size of a union of three sets is given by a more complicated Inclusion-Exclusion formula:

$$\begin{aligned} |S_1 \cup S_2 \cup S_3| = {} & |S_1| + |S_2| + |S_3| \\ & - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| \\ & + |S_1 \cap S_2 \cap S_3|. \end{aligned}$$

Remarkably, the expression on the right accounts for each element in the union of S_1, S_2, and S_3 exactly once. For example, suppose that x is an element of all three sets. Then x is counted three times (by the $|S_1|$, $|S_2|$, and $|S_3|$ terms), subtracted off three times (by the $|S_1 \cap S_2|$, $|S_1 \cap S_3|$, and $|S_2 \cap S_3|$ terms), and then counted once more (by the $|S_1 \cap S_2 \cap S_3|$ term). The net effect is that x is counted just once.

If x is in two sets (say, S_1 and S_2), then x is counted twice (by the $|S_1|$ and $|S_2|$ terms) and subtracted once (by the $|S_1 \cap S_2|$ term). In this case, x does not factor into any of the other terms, since $x \notin S_3$.

So we can't answer the original question without knowing the sizes of the various intersections. Let's suppose that there are:

4	math - EECS double majors
3	math - physics double majors
11	EECS - physics double majors
2	triple majors

Then $|M \cap E| = 4+2$, $|M \cap P| = 3+2$, $|E \cap P| = 11+2$, and $|M \cap E \cap P| = 2$. Plugging all this into the formula gives:

$$\begin{aligned}|M \cup E \cup P| &= |M| + |E| + |P| - |M \cap E| - |M \cap P| - |E \cap P| + |M \cap E \cap P| \\ &= 60 + 200 + 40 - 6 - 5 - 13 + 2 \\ &= 278\end{aligned}$$

11.8.3 Sequences with 42, 04, or 60

In how many permutations of the set $\{0, 1, 2, \ldots, 9\}$ do either 4 and 2, 0 and 4, or 6 and 0 appear consecutively? For example, none of these pairs appears in:

$$(7, 2, 9, 5, 4, 1, 3, 8, 0, 6).$$

The 06 at the end doesn't count; we need 60. On the other hand, both 04 and 60 appear consecutively in this permutation:

$$(7, 2, 5, \underline{6}, \underline{0}, \underline{4}, 3, 8, 1, 9).$$

Let P_{42} be the set of all permutations in which 42 appears. Define P_{60} and P_{04} similarly. Thus, for example, the permutation above is contained in both P_{60} and P_{04}, but not P_{42}. In these terms, we're looking for the size of the set $P_{42} \cup P_{04} \cup P_{60}$.

First, we must determine the sizes of the individual sets, such as P_{60}. We can use a trick: group the 6 and 0 together as a single symbol. Then there is a natural bijection between permutations of $\{0, 1, 2, \ldots 9\}$ containing 6 and 0 consecutively and permutations of:

$$\{60, 1, 2, 3, 4, 5, 7, 8, 9\}.$$

For example, the following two sequences correspond:

$$(7, 2, 5, \underline{6}, \underline{0}, 4, 3, 8, 1, 9) \longleftrightarrow (7, 2, 5, \underline{60}, 4, 3, 8, 1, 9).$$

There are 9! permutations of the set containing 60, so $|P_{60}| = 9!$ by the Bijection Rule. Similarly, $|P_{04}| = |P_{42}| = 9!$ as well.

Next, we must determine the sizes of the two-way intersections, such as $P_{42} \cap P_{60}$. Using the grouping trick again, there is a bijection with permutations of the set:

$$\{42, 60, 1, 3, 5, 7, 8, 9\}.$$

Thus, $|P_{42} \cap P_{60}| = 8!$. Similarly, $|P_{60} \cap P_{04}| = 8!$ by a bijection with the set:

$$\{604, 1, 2, 3, 5, 7, 8, 9\}.$$

And $|P_{42} \cap P_{04}| = 8!$ as well by a similar argument. Finally, note that $|P_{60} \cap P_{04} \cap P_{42}| = 7!$ by a bijection with the set:

$$\{6042, 1, 3, 5, 7, 8, 9\}.$$

Plugging all this into the formula gives:

$$|P_{42} \cup P_{04} \cup P_{60}| = 9! + 9! + 9! - 8! - 8! - 8! + 7!.$$

11.8.4 Union of *n* Sets

The size of a union of n sets is given by the following rule.

Rule 11.8.1 (Inclusion-Exclusion)**.**

$$|S_1 \cup S_2 \cup \cdots \cup S_n| =$$

	the sum of the sizes of the individual sets
minus	*the sizes of all two-way intersections*
plus	*the sizes of all three-way intersections*
minus	*the sizes of all four-way intersections*
plus	*the sizes of all five-way intersections, etc.*

The formulas for unions of two and three sets are special cases of this general rule.

This way of expressing Inclusion-Exclusion is easy to understand and nearly as precise as expressing it in mathematical symbols, but we'll need the symbolic version below, so let's work on deciphering it now.

We already have a standard notation for the sum of sizes of the individual sets, namely,

$$\sum_{i=1}^{n} |S_i|.$$

A "two-way intersection" is a set of the form $S_i \cap S_j$ for $i \neq j$. We regard $S_j \cap S_i$ as the same two-way intersection as $S_i \cap S_j$, so we can assume that $i < j$. Now we can express the sum of the sizes of the two-way intersections as

$$\sum_{1 \leq i < j \leq n} |S_i \cap S_j|.$$

Similarly, the sum of the sizes of the three-way intersections is

$$\sum_{1 \leq i < j < k \leq n} |S_i \cap S_j \cap S_k|.$$

These sums have alternating signs in the Inclusion-Exclusion formula, with the sum of the k-way intersections getting the sign $(-1)^{k-1}$. This finally leads to a symbolic version of the rule:

Rule (Inclusion-Exclusion)**.**

$$\begin{aligned}|\bigcup_{i=1}^{n} S_i| = &\sum_{i=1}^{n} |S_i| \\ &- \sum_{1\le i<j\le n} |S_i \cap S_j| \\ &+ \sum_{1\le i<j<k\le n} |S_i \cap S_j \cap S_k| + \cdots \\ &+ (-1)^{n-1} |\bigcap_{i=1}^{n} S_i|.\end{aligned}$$

11.8.5 Computing Euler's Function

As an example, let's use Inclusion-Exclusion to calculate Euler's function, $\phi(n)$. By definition, $\phi(n)$ is the number of nonnegative integers less than a positive integer n that are relatively prime to n. But the set S of nonnegative integers less than n that are *not* relatively prime to n will be easier to count.

Suppose the prime factorization of n is $p_1^{e_1} \cdots p_m^{e_m}$ for distinct primes p_i. This means that the integers in S are precisely the nonnegative integers less than n that are divisible by at least one of the p_i's. Letting C_i be the set of nonnegative integers less than n that are divisible by p_i, we have

$$S = \bigcup_{i=1}^{m} C_i.$$

We'll be able to find the size of this union using Inclusion-Exclusion because the intersections of the C_i's are easy to count. For example, $C_1 \cap C_2 \cap C_3$ is the set of nonnegative integers less than n that are divisible by each of p_1, p_2 and p_3. But since the p_i's are distinct primes, being divisible by each of these primes is the same as being divisible by their product. Now observe that if r is a positive divisor of n, then exactly n/r nonnegative integers less than n are divisible by r, namely, $0, r, 2r, \ldots, ((n/r) - 1)r$. So exactly $n/p_1 p_2 p_3$ nonnegative integers less than n are divisible by all three primes p_1, p_2, p_3. In other words,

$$|C_1 \cap C_2 \cap C_3| = \frac{n}{p_1 p_2 p_3}.$$

Reasoning this way about all the intersections among the C_i's and applying Inclusion-Exclusion, we get

$$\begin{aligned}
|S| &= |\bigcup_{i=1}^{m} C_i| \\
&= \sum_{i=1}^{m} |C_i| - \sum_{1\le i<j\le m} |C_i \cap C_j| + \sum_{1\le i<j<k\le m} |C_i \cap C_j \cap C_k| - \cdots + (-1)^{m-1} |\bigcap_{i=1}^{m} C_i| \\
&= \sum_{i=1}^{m} \frac{n}{p_i} - \sum_{1\le i<j\le m} \frac{n}{p_i p_j} + \sum_{1\le i<j<k\le m} \frac{n}{p_i p_j p_k} - \cdots + (-1)^{m-1} \frac{n}{p_1 p_2 \cdots p_n} \\
&= n\left(\sum_{i=1}^{m} \frac{1}{p_i} - \sum_{1\le i<j\le m} \frac{1}{p_i p_j} + \sum_{1\le i<j<k\le m} \frac{1}{p_i p_j p_k} - \cdots + (-1)^{m-1} \frac{1}{p_1 p_2 \cdots p_n}\right)
\end{aligned}$$

But $\phi(n) = n - |S|$ by definition, so

$$\begin{aligned}
\phi(n) &= n\left(1 - \sum_{i=1}^{m} \frac{1}{p_i} + \sum_{1\le i<j\le m} \frac{1}{p_i p_j} - \sum_{1\le i<j<k\le m} \frac{1}{p_i p_j p_k} + \cdots + (-1)^{m} \frac{1}{p_1 p_2 \cdots p_n}\right) \\
&= n \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right). \qquad (11.3)
\end{aligned}$$

Yikes! That was pretty hairy. Are you getting tired of all that nasty algebra? If so, then good news is on the way. In the next section, we will show you how to prove some heavy-duty formulas without using any algebra at all. Just a few words and you are done. No kidding.

11.9 Combinatorial Proofs

Suppose you have n different T-shirts, but only want to keep k. You could equally well select the k shirts you want to keep or select the complementary set of $n-k$ shirts you want to throw out. Thus, the number of ways to select k shirts from among n must be equal to the number of ways to select $n-k$ shirts from among n. Therefore:

$$\binom{n}{k} = \binom{n}{n-k}.$$

This is easy to prove algebraically, since both sides are equal to:

$$\frac{n!}{k!\,(n-k)!}.$$

But we didn't really have to resort to algebra; we just used counting principles.

Hmmm...

11.9.1 Pascal's Identity

Jay, famed Math for Computer Science Teaching Assistant, has decided to try out for the US Olympic boxing team. After all, he's watched all of the *Rocky* movies and spent hours in front of a mirror sneering, "Yo, you wanna piece a' *me*?!" Jay figures that n people (including himself) are competing for spots on the team and only k will be selected. As part of maneuvering for a spot on the team, he needs to work out how many different teams are possible. There are two cases to consider:

- Jay *is* selected for the team, and his $k-1$ teammates are selected from among the other $n-1$ competitors. The number of different teams that can be formed in this way is:
$$\binom{n-1}{k-1}.$$
- Jay is *not* selected for the team, and all k team members are selected from among the other $n-1$ competitors. The number of teams that can be formed this way is:
$$\binom{n-1}{k}.$$

All teams of the first type contain Jay, and no team of the second type does; therefore, the two sets of teams are disjoint. Thus, by the Sum Rule, the total number of possible Olympic boxing teams is:

$$\binom{n-1}{k-1}+\binom{n-1}{k}.$$

Jeremy, equally-famed Teaching Assistant, thinks Jay isn't so tough and so he might as well also try out. He reasons that n people (including himself) are trying out for k spots. Thus, the number of ways to select the team is simply:

$$\binom{n}{k}.$$

Jeremy and Jay each correctly counted the number of possible boxing teams. Thus, their answers must be equal. So we know:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

This is called *Pascal's Identity*. And we proved it *without any algebra!* Instead, we relied purely on counting techniques.

11.9.2 Finding a Combinatorial Proof

A *combinatorial proof* is an argument that establishes an algebraic fact by relying on counting principles. Many such proofs follow the same basic outline:

1. Define a set S.
2. Show that $|S| = n$ by counting one way.
3. Show that $|S| = m$ by counting another way.
4. Conclude that $n = m$.

In the preceding example, S was the set of all possible Olympic boxing teams. Jay computed

$$|S| = \binom{n-1}{k-1} + \binom{n-1}{k}$$

by counting one way, and Jeremy computed

$$|S| = \binom{n}{k}$$

by counting another way. Equating these two expressions gave Pascal's Identity.

More typically, the set S is defined in terms of simple sequences or sets rather than an elaborate story. Here is a less colorful example of a combinatorial argument.

Theorem 11.9.1.

$$\sum_{r=0}^{n} \binom{n}{r}\binom{2n}{n-r} = \binom{3n}{n}$$

Proof. We give a combinatorial proof. Let S be all n-card hands that can be dealt from a deck containing n red cards (numbered $1, \ldots, n$) and $2n$ black cards (numbered $1, \ldots, 2n$). First, note that every $3n$-element set has

$$|S| = \binom{3n}{n}$$

n-element subsets.

From another perspective, the number of hands with exactly r red cards is

$$\binom{n}{r}\binom{2n}{n-r}$$

since there are $\binom{n}{r}$ ways to choose the r red cards and $\binom{2n}{n-r}$ ways to choose the $n-r$ black cards. Since the number of red cards can be anywhere from 0 to n, the total number of n-card hands is:

$$|S| = \sum_{r=0}^{n} \binom{n}{r}\binom{2n}{n-r}.$$

Equating these two expressions for $|S|$ proves the theorem. ■

Combinatorial proofs are almost magical. Theorem 11.9.1 looks pretty scary, but we proved it without any algebraic manipulations at all. The key to constructing a combinatorial proof is choosing the set S properly, which can be tricky. Generally, the simpler side of the equation should provide some guidance. For example, the right side of Theorem 11.9.1 is $\binom{3n}{n}$, which suggests that it will be helpful to choose S to be all n-element subsets of some $3n$-element set.

11.10 The Pigeonhole Principle

Here is an old puzzle:

> A drawer in a dark room contains red socks, green socks, and blue socks. How many socks must you withdraw to be sure that you have a matching pair?

For example, picking out three socks is not enough; you might end up with one red, one green, and one blue. The solution relies on the *Pigeonhole Principle*, which is a friendly name for the contrapositive of the injective case of the Mapping Rule.

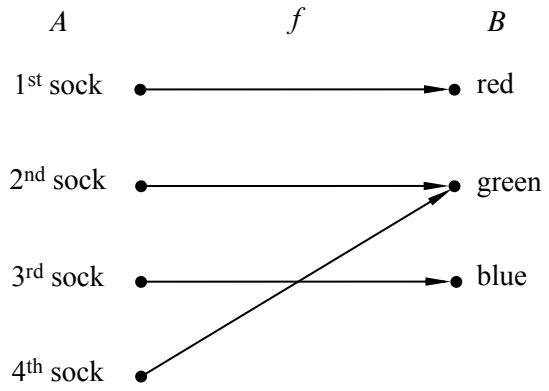


Figure 11.3 One possible mapping of four socks to three colors.

Rule 11.10.1 (Pigeonhole Principle)**.** *If $|X| > |Y|$, then for every total function*[5] *$f : X \to Y$, there exist two different elements of X that are mapped to the same element of Y.*

What this abstract mathematical statement has to do with selecting footwear under poor lighting conditions is maybe not obvious. However, let A be the set of socks you pick out, let B be the set of colors available, and let f map each sock to its color. The Pigeonhole Principle says that if $|A| > |B| = 3$, then at least two elements of A (that is, at least two socks) must be mapped to the same element of B (that is, the same color). Therefore, four socks are enough to ensure a matched pair. For example, one possible mapping of four socks to three colors is shown in Figure 11.3.

Not surprisingly, the pigeonhole principle is often described in terms of pigeons:

> *If there are more pigeons than holes they occupy, then at least two pigeons must be in the same hole.*

In this case, the pigeons form set A, the pigeonholes are set B, and f describes which hole each pigeon flies into.

Mathematicians have come up with many ingenious applications for the pigeonhole principle. If there were a cookbook procedure for generating such arguments, we'd give it to you. Unfortunately, there isn't one. One helpful tip, though: when you try to solve a problem with the pigeonhole principle, the key is to clearly identify three things:

[5] This Mapping Rule applies even if f is a total injective *relation*. Recall that a function is total if $\forall x \in X\ \exists y \in Y.\ f(x) = y$.

1. The set A (the pigeons).
2. The set B (the pigeonholes).
3. The function f (the rule for assigning pigeons to pigeonholes).

11.10.1 Hairs on Heads

There are a number of generalizations of the pigeonhole principle. For example:

Rule 11.10.2 (Generalized Pigeonhole Principle)**.** *If $|X| > k \cdot |Y|$, then every total function $f : X \to Y$ maps at least $k + 1$ different elements of X to the same element of Y.*

For example, if you pick two people at random, surely they are extremely unlikely to have *exactly* the same number of hairs on their heads. However, in the remarkable city of Boston, Massachusetts there are actually *three* people who have exactly the same number of hairs! Of course, there are many bald people in Boston, and they all have zero hairs. But we're talking about non-bald people; say a person is non-bald if they have at least ten thousand hairs on their head.

Boston has about 500,000 non-bald people, and the number of hairs on a person's head is at most 200,000. Let A be the set of non-bald people in Boston, let $B = \{10,000, 10,001, \ldots, 200,000\}$, and let f map a person to the number of hairs on his or her head. Since $|A| > 2|B|$, the Generalized Pigeonhole Principle implies that at least three people have exactly the same number of hairs. We don't know who they are, but we know they exist!

11.10.2 Subsets with the Same Sum

For your reading pleasure, we have displayed ninety 25-digit numbers in Figure 11.4. Are there two different subsets of these 25-digit numbers that have the same sum? For example, maybe the sum of the last ten numbers in the first column is equal to the sum of the first eleven numbers in the second column?

Finding two subsets with the same sum may seem like a silly puzzle, but solving these sorts of problems turns out to be useful in diverse applications such as finding good ways to fit packages into shipping containers and decoding secret messages.

It turns out that it is hard to find different subsets with the same sum, which is why this problem arises in cryptography. But it is easy to prove that two such subsets *exist*. That's where the Pigeonhole Principle comes in.

Let A be the collection of all subsets of the 90 numbers in the list. Now the sum of any subset of numbers is at most $90 \cdot 10^{25}$, since there are only 90 numbers and every 25-digit number is less than 10^{25}. So let B be the set of integers $\{0, 1, \ldots, 90 \cdot 10^{25}\}$, and let f map each subset of numbers (in A) to its sum (in B).

00204801353855029644480 38	31710048321735013941130 17
57632573310834796474093 98	82473310000429953116460 21
04894459918669156762409 92	32082344215973686470192 65
58009491235489891226286 63	84962439971234759227663 10
10826620324303796513709 81	34372546563551578648691 13
60429008011992802180260 01	85183991406760026607474 77
11784808947697061789949 93	35748833930586539237113 65
61161717891377378967014 05	85436912834701914523337 63
12531273516832396938513 27	36449099460404801899691 49
61448689730015823697235 12	86753092583741370924613 52
13015051292340778110690 11	37900441327370840944172 46
62473145938511692347461 52	86943211123639968672966 65
13115671111438664338821 94	38703321274379713555228 15
68144289442668749634882 74	87723212036084772458511 54
14700294527212035876862 14	40805058045778014513631 00
68708529455438868491478 81	87914221617225825463410 91
15782710472862574994338 86	41672834610257023481249 20
69149555081209500937323 97	90626280245921262839732 85
16382439218521762431923 54	42359968311237777882112 49
69496324513659871524235 41	91378455669255263498977 94
17635802191319859631023 65	46709394457494390421112 20
71282111436136198284156 50	91537629668031892919344 19
18262277956018422310296 94	48153793518653842796134 27
71739200836518623079253 94	92708801940776364069842 49
18439718626751020372014 20	48370529482129226044421 90
72156548742117556762205 87	93243014807221034903792 04
23969511937221345261772 37	51063894238550185506715 30
72569328471643910402330 50	94360908321466951471405 81
27813945682685998010963 54	51423681920047692180699 10
73328226570752354316203 17	94753081597345382490132 38
27966051967136104054080 19	51812340961301440840418 56
74264418295415734449641 39	94923766239174869749232 02
29310163947619752631903 47	51982673981256179943913 48
76321981265318093271863 21	95119725587798802882529 79
29334580582944051551972 96	53175929403162312197583 72
77121544322119128823105 11	96024134246191871125522 64
30755144104909759203153 48	53843581267717941283569 47
78589186642402623566100 10	96312171149061292194611 11
81494367168713711619320 35	31576931053251112843219 93
31114749852527934528600 17	54392117122489019954234 41
78981567867632129631786 79	99081898531027533359813 19
31456215879361201184387 01	56103798260928381927604 58
81475910170375733378486 16	99132374763417642998139 87
31489012556288811031985 49	56323175554652286776760 44
56921683746370196174237 12	81760638316825365713067 91

Figure 11.4 Ninety 25-digit numbers. Can you find two different subsets of these numbers that have the same sum?

We proved that an n-element set has 2^n different subsets in Section 11.2. Therefore:

$$|A| = 2^{90} \geq 1.237 \times 10^{27}$$

On the other hand:

$$|B| = 90 \cdot 10^{25} + 1 \leq 0.901 \times 10^{27}.$$

Both quantities are enormous, but $|A|$ is a bit greater than $|B|$. This means that f maps at least two elements of A to the same element of B. In other words, by the Pigeonhole Principle, two different subsets must have the same sum!

Notice that this proof gives no indication *which* two sets of numbers have the same sum. This frustrating variety of argument is called a *nonconstructive proof*. To see if was possible to actually *find* two different subsets of the ninety 25-digit numbers with the same sum, we offered a $100 prize to the first student who did it. We didn't expect to have to pay off this bet, but we underestimated the ingenuity and initiative of the students. One computer science major wrote a program that cleverly searched only among a reasonably small set of "plausible" sets, sorted them by their sums, and actually found a couple with the same sum. He won the prize. A few days later, a math major figured out how to reformulate the sum problem as a "lattice basis reduction" problem; then he found a software package implementing an efficient basis reduction procedure, and using it, he very quickly found lots of pairs of subsets with the same sum. He didn't win the prize, but he got a standing ovation from the class—staff included.

11.11 A Magic Trick

There is a Magician and an Assistant. The Assistant goes into the audience with a deck of 52 cards while the Magician looks away.

Five audience members each select one card from the deck. The Assistant then gathers up the five cards and holds up four of them so the Magician can see them. The Magician concentrates for a short time and then correctly names the secret, fifth card!

Since we don't really believe the Magician can read minds, we know the Assistant has somehow communicated the secret card to the Magician. Since real Magicians and Assistants are not to be trusted, we can expect that the Assistant would illegitimately signal the Magician with coded phrases or body language, but they don't have to cheat in this way. In fact, the Magician and Assistant could be

Sets with Distinct Subset Sums

How can we construct a set of n positive integers such that all its subsets have *distinct* sums? One way is to use powers of two:

$$\{1, 2, 4, 8, 16\}$$

This approach is so natural that one suspects all other such sets must involve larger numbers. (For example, we could safely replace 16 by 17, but not by 15.) Remarkably, there are examples involving *smaller* numbers. Here is one:

$$\{6, 9, 11, 12, 13\}$$

One of the top mathematicians of the Twentieth Century, Paul Erdős, conjectured in 1931 that there are no such sets involving *significantly* smaller numbers. More precisely, he conjectured that the largest number in such a set must be greater than $c2^n$ for some constant $c > 0$. He offered $500 to anyone who could prove or disprove his conjecture, but the problem remains unsolved.

kept out of sight of each other while some audience member holds up the 4 cards designated by the Assistant for the Magician to see.

Of course, without cheating, there is still an obvious way the Assistant can communicate to the Magician: he can choose any of the $4! = 24$ permutations of the 4 cards as the order in which to hold up the cards. However, this alone won't quite work: there are 48 cards remaining in the deck, so the Assistant doesn't have enough choices of orders to indicate exactly what the secret card is (though he could narrow it down to two cards).

11.11.1 The Secret

The method the Assistant can use to communicate the fifth card exactly is a nice application of what we know about counting and matching.

The Assistant has a second legitimate way to communicate: he can choose *which of the five cards to keep hidden*. Of course, it's not clear how the Magician could determine which of these five possibilities the Assistant selected by looking at the four visible cards, but there is a way, as we'll now explain.

The problem facing the Magician and Assistant is actually a bipartite matching problem. Put all the *sets* of 5 cards in a collection X on the left. And put all the *sequences* of 4 distinct cards in a collection Y on the right. These are the two sets of vertices in the bipartite graph. There is an edge between a set of 5 cards and a sequence of 4 if every card in the sequence is also in the set. In other words, if the audience selects a set of 5 cards, then the Assistant must reveal a sequence of 4 cards that is adjacent in the bipartite graph. Some edges are shown in the diagram in Figure 11.5.

For example,

$$\{8\heartsuit, K\spadesuit, Q\spadesuit, 2\diamondsuit, 6\diamondsuit\} \tag{11.4}$$

is an element of X on the left. If the audience selects this set of 5 cards, then there are many different 4-card sequences on the right in set Y that the Assistant could choose to reveal, including $(8\heartsuit, K\spadesuit, Q\spadesuit, 2\diamondsuit)$, $(K\spadesuit, 8\heartsuit, Q\spadesuit, 2\diamondsuit)$, and $(K\spadesuit, 8\heartsuit, 6\diamondsuit, Q\spadesuit)$.

What the Magician and his Assistant need to perform the trick is a *matching* for the X vertices. If they agree in advance on some matching, then when the audience selects a set of 5 cards, the Assistant reveals the matching sequence of 4 cards. The Magician uses the matching to find the audience's chosen set of 5 cards, and so he can name the one not already revealed.

For example, suppose the Assistant and Magician agree on a matching containing the two bold edges in Figure 11.5. If the audience selects the set

$$\{8\heartsuit, K\spadesuit, Q\spadesuit, 9\clubsuit, 6\diamondsuit\}, \tag{11.5}$$

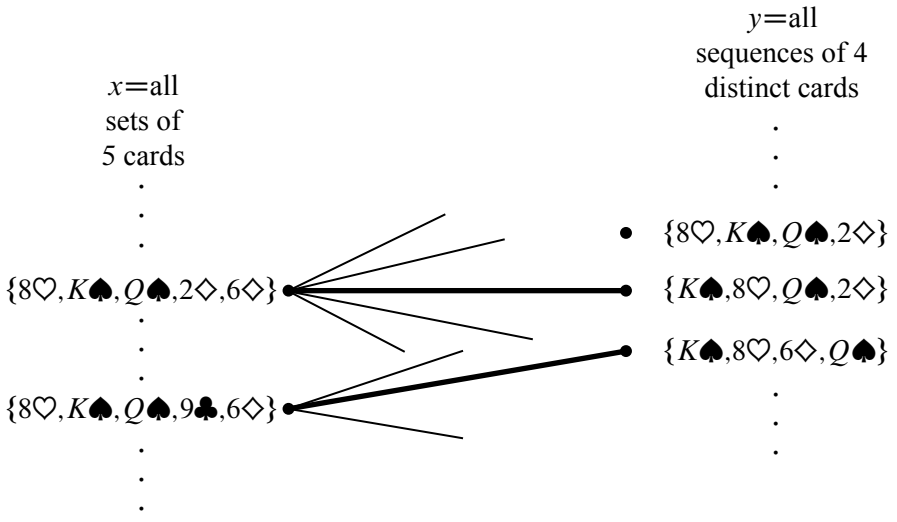


Figure 11.5 The bipartite graph where the nodes on the left correspond to *sets* of 5 cards and the nodes on the right correspond to *sequences* of 4 cards. There is an edge between a set and a sequence whenever all the cards in the sequence are contained in the set.

then the Assistant reveals the corresponding sequence

$$(K\spadesuit, 8\heartsuit, 6\diamondsuit, Q\spadesuit). \tag{11.6}$$

Using the matching, the Magician sees that the hand (11.5) is matched to the sequence (11.6), so he can name the one card in the corresponding set not already revealed, namely, the 9♣. Notice that the fact that the sets are *matched*, that is, that different sets are paired with *distinct* sequences, is essential. For example, if the audience picked the previous hand (11.4), it would be possible for the Assistant to reveal the same sequence (11.6), but he better not do that; if he did, then the Magician would have no way to tell if the remaining card was the 9♣ or the 2◇.

So how can we be sure the needed matching can be found? The answer is that each vertex on the left has degree $5 \cdot 4! = 120$, since there are five ways to select the card kept secret and there are 4! permutations of the remaining 4 cards. In addition, each vertex on the right has degree 48, since there are 48 possibilities for the fifth card. So this graph is *degree-constrained* according to Definition 5.2.6, and therefore satisfies Hall's matching condition.

In fact, this reasoning shows that the Magician could still pull off the trick if 120 cards were left instead of 48, that is, the trick would work with a deck as large as 124 different cards—without any magic!

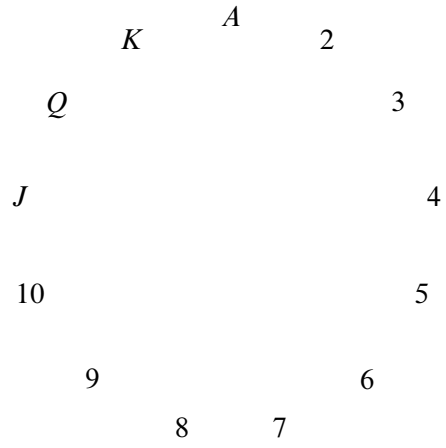


Figure 11.6 The 13 card ranks arranged in cyclic order.

11.11.2 The Real Secret

But wait a minute! It's all very well in principle to have the Magician and his Assistant agree on a matching, but how are they supposed to remember a matching with $\binom{52}{5} = 2,598,960$ edges? For the trick to work in practice, there has to be a way to match hands and card sequences mentally and on the fly.

We'll describe one approach. As a running example, suppose that the audience selects:

$$10\heartsuit \quad 9\diamond \quad 3\heartsuit \quad Q\spadesuit \quad J\diamond.$$

- The Assistant picks out two cards of the same suit. In the example, the assistant might choose the $3\heartsuit$ and $10\heartsuit$. This is always possible because of the Pigeonhole Principle—there are five cards and 4 suits so two cards must be in the same suit.

- The Assistant locates the ranks of these two cards on the cycle shown in Figure 11.6. For any two distinct ranks on this cycle, one is always between 1 and 6 hops clockwise from the other. For example, the $3\heartsuit$ is 6 hops clockwise from the $10\heartsuit$.

- The more counterclockwise of these two cards is revealed first, and the other becomes the secret card. Thus, in our example, the $10\heartsuit$ would be revealed, and the $3\heartsuit$ would be the secret card. Therefore:

  - The suit of the secret card is the same as the suit of the first card revealed.

- The rank of the secret card is between 1 and 6 hops clockwise from the rank of the first card revealed.

- All that remains is to communicate a number between 1 and 6. The Magician and Assistant agree beforehand on an ordering of all the cards in the deck from smallest to largest such as:

$$A\clubsuit\ A\diamondsuit\ A\heartsuit\ A\spadesuit\ 2\clubsuit\ 2\diamondsuit\ 2\heartsuit\ 2\spadesuit\ \ldots\ K\heartsuit\ K\spadesuit$$

The order in which the last three cards are revealed communicates the number according to the following scheme:

(small, medium, large)	= 1
(small, large, medium)	= 2
(medium, small, large)	= 3
(medium, large, small)	= 4
(large, small, medium)	= 5
(large, medium, small)	= 6

In the example, the Assistant wants to send 6 and so reveals the remaining three cards in large, medium, small order. Here is the complete sequence that the Magician sees:

$$10\heartsuit \quad Q\spadesuit \quad J\diamondsuit \quad 9\diamondsuit$$

- The Magician starts with the first card, $10\heartsuit$, and hops 6 ranks clockwise to reach $3\heartsuit$, which is the secret card!

So that's how the trick can work with a standard deck of 52 cards. On the other hand, Hall's Theorem implies that the Magician and Assistant can *in principle* perform the trick with a deck of up to 124 cards. It turns out that there is a method which they could actually learn to use with a reasonable amount of practice for a 124-card deck, but we won't explain it here.[6]

11.11.3 The Same Trick with Four Cards?

Suppose that the audience selects only *four* cards and the Assistant reveals a sequence of *three* to the Magician. Can the Magician determine the fourth card?

Let X be all the sets of four cards that the audience might select, and let Y be all the sequences of three cards that the Assistant might reveal. Now, on one hand, we have

$$|X| = \binom{52}{4} = 270,725$$

[6] See *The Best Card Trick* by Michael Kleber for more information.

by the Subset Rule. On the other hand, we have

$$|Y| = 52 \cdot 51 \cdot 50 = 132,600$$

by the Generalized Product Rule. Thus, by the Pigeonhole Principle, the Assistant must reveal the *same* sequence of three cards for at least

$$\left\lceil \frac{270,725}{132,600} \right\rceil = 3$$

different four-card hands. This is bad news for the Magician: if he sees that sequence of three, then there are at least three possibilities for the fourth card which he cannot distinguish. So there is no legitimate way for the Assistant to communicate exactly what the fourth card is!

11.11.4 Never Say Never

No sooner than we finished proving that the Magician can't pull off the trick with four cards instead of five, a student showed us a way that it might be doable after all. The idea is to place the three cards on a table one at a time instead of revealing them all at once. This provides the Magician with two completely independent sequences of three cards: one for the *temporal* order in which the cards are placed on the table, and one for the *spatial* order in which they appear once placed.

For example, suppose the audience selects

$$10\heartsuit \quad 9\diamondsuit \quad 3\heartsuit \quad Q\spadesuit$$

and the assistant decides to reveal

$$10\heartsuit \quad 9\diamondsuit \quad Q\spadesuit.$$

The assistant might decide to reveal the $Q\spadesuit$ first, the $10\heartsuit$ second, and the $9\diamondsuit$ third, thereby production the *temporal* sequence

$$(Q\spadesuit, 10\heartsuit, 9\diamondsuit).$$

If the $Q\spadesuit$ is placed in the middle position on the table, the $10\heartsuit$ is placed in the rightmost position on the table, and the $9\diamondsuit$ is placed in the leftmost position on the table, the *spatial* sequence would be

$$(9\diamondsuit, Q\spadesuit, 10\heartsuit).$$

In this version of the card trick, X consists of all sets of 4 cards and Y consists of all *pairs* of sequences of the same 3 cards. As before, we can create a bipartite

graph where an edge connects a set S of 4 cards in X with a pair of sequences in Y if the 3 cards in the sequences are in S.

The degree of every node in X is then

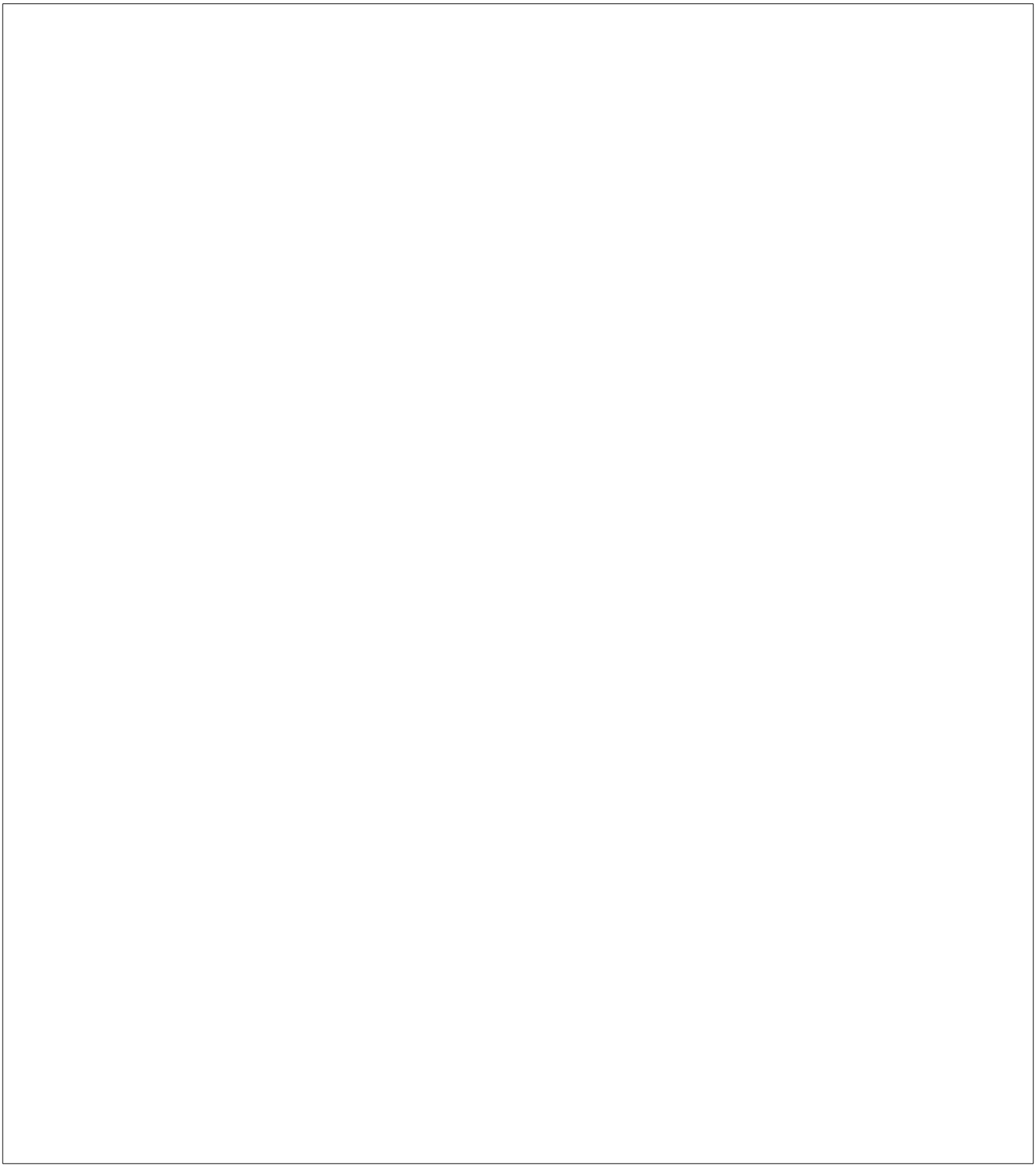
$$4 \cdot 3! \cdot 3! = 144$$

since there are 4 choices for which card is not revealed and 3! orders for each sequence in the pair.

The degree of every node in Y is 49 since there are $52 - 3 = 49$ possible choices for the 4th card. Since $144 \geq 49$, we can use Hall's Theorem to establish the existing of a matching for X.

Hence, the magic trick *is* doable with 4 cards—the assistant just has to convey more information. Can you figure out a convenient way to pull off the trick on the fly?

So what about the 3-card version? Surely that is not doable...



12 Generating Functions

Generating Functions are one of the most surprising and useful inventions in Discrete Math. Roughly speaking, generating functions transform problems about *sequences* into problems about *functions*. This is great because we've got piles of mathematical machinery for manipulating functions. Thanks to generating functions, we can then apply all that machinery to problems about sequences. In this way, we can use generating functions to solve all sorts of counting problems. They can also be used to find closed-form expressions for sums and to solve recurrences. In fact, many of the problems we addressed in Chapters 9–11 can be formulated and solved using generating functions.

12.1 Definitions and Examples

The *ordinary generating function* for the sequence[1] $\langle g_0, g_1, g_2, g_3 \dots \rangle$ is the power series:

$$G(x) = g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots .$$

There are a few other kinds of generating functions in common use, but ordinary generating functions are enough to illustrate the power of the idea, so we'll stick to them and from now on, *generating function* will mean the ordinary kind.

A generating function is a "formal" power series in the sense that we usually regard x as a placeholder rather than a number. Only in rare cases will we actually evaluate a generating function by letting x take a real number value, so we generally ignore the issue of convergence.

Throughout this chapter, we'll indicate the correspondence between a sequence and its generating function with a double-sided arrow as follows:

$$\langle g_0, g_1, g_2, g_3, \dots \rangle \longleftrightarrow g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots .$$

For example, here are some sequences and their generating functions:

$$\langle 0, 0, 0, 0, \dots \rangle \longleftrightarrow 0 + 0x + 0x^2 + 0x^3 + \cdots = 0$$
$$\langle 1, 0, 0, 0, \dots \rangle \longleftrightarrow 1 + 0x + 0x^2 + 0x^3 + \cdots = 1$$
$$\langle 3, 2, 1, 0, \dots \rangle \longleftrightarrow 3 + 2x + 1x^2 + 0x^3 + \cdots = 3 + 2x + x^2$$

[1] In this chapter, we'll put sequences in angle brackets to more clearly distinguish them from the many other mathematical expressions floating around.

The pattern here is simple: the ith term in the sequence (indexing from 0) is the coefficient of x^i in the generating function.

Recall that the sum of an infinite geometric series is:

$$1 + z + z^2 + z^3 + \cdots = \frac{1}{1-z}.$$

This equation does not hold when $|z| \geq 1$, but as remarked, we won't worry about convergence issues for now. This formula gives closed form generating functions for a whole range of sequences. For example:

$$\langle 1, 1, 1, 1, \ldots \rangle \longleftrightarrow 1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x}$$

$$\langle 1, -1, 1, -1, \ldots \rangle \longleftrightarrow 1 - x + x^2 - x^3 + x^4 - \cdots = \frac{1}{1+x}$$

$$\langle 1, a, a^2, a^3, \ldots \rangle \longleftrightarrow 1 + ax + a^2x^2 + a^3x^3 + \cdots = \frac{1}{1-ax}$$

$$\langle 1, 0, 1, 0, 1, 0, \ldots \rangle \longleftrightarrow 1 + x^2 + x^4 + x^6 + x^8 + \cdots = \frac{1}{1-x^2}$$

12.2 Operations on Generating Functions

The magic of generating functions is that we can carry out all sorts of manipulations on sequences by performing mathematical operations on their associated generating functions. Let's experiment with various operations and characterize their effects in terms of sequences.

12.2.1 Scaling

Multiplying a generating function by a constant scales every term in the associated sequence by the same constant. For example, we noted above that:

$$\langle 1, 0, 1, 0, 1, 0, \ldots \rangle \longleftrightarrow 1 + x^2 + x^4 + x^6 + \cdots = \frac{1}{1-x^2}.$$

Multiplying the generating function by 2 gives

$$\frac{2}{1-x^2} = 2 + 2x^2 + 2x^4 + 2x^6 + \cdots$$

which generates the sequence:

$$\langle 2, 0, 2, 0, 2, 0, \ldots \rangle .$$

Rule 12.2.1 (Scaling Rule)**.** *If*

$$\langle f_0, f_1, f_2, \ldots \rangle \longleftrightarrow F(x),$$

then

$$\langle cf_0,\ cf_1,\ cf_2,\ \ldots \rangle \longleftrightarrow c \cdot F(x).$$

The idea behind this rule is that:

$$\begin{aligned} \langle cf_0, cf_1, cf_2, \ldots \rangle &\longleftrightarrow cf_0 + cf_1 x + cf_2 x^2 + \cdots \\ &= c \cdot (f_0 + f_1 x + f_2 x^2 + \cdots) \\ &= cF(x). \end{aligned}$$

12.2.2 Addition

Adding generating functions corresponds to adding the two sequences term by term. For example, adding two of our earlier examples gives:

$$\begin{array}{rl} \langle 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ \ldots \rangle & \longleftrightarrow \dfrac{1}{1-x} \\ +\ \langle 1, -1, 1, -1, 1, -1, \ldots \rangle & \longleftrightarrow \dfrac{1}{1+x} \\ \hline \langle 2,\ 0,\ 2,\ 0,\ 2,\ 0,\ \ldots \rangle & \longleftrightarrow \dfrac{1}{1-x} + \dfrac{1}{1+x} \end{array}$$

We've now derived two different expressions that both generate the sequence $\langle 2, 0, 2, 0, \ldots \rangle$. They are, of course, equal:

$$\frac{1}{1-x} + \frac{1}{1+x} = \frac{(1+x)+(1-x)}{(1-x)(1+x)} = \frac{2}{1-x^2}.$$

Rule 12.2.2 (Addition Rule)**.** *If*

$$\begin{aligned} \langle f_0, f_1, f_2, \ldots \rangle &\longleftrightarrow F(x) \qquad \textit{and} \\ \langle g_0, g_1, g_2, \ldots \rangle &\longleftrightarrow G(x), \end{aligned}$$

then

$$\langle f_0 + g_0,\ f_1 + g_1,\ f_2 + g_2,\ \ldots \rangle \longleftrightarrow F(x) + G(x).$$

The idea behind this rule is that:

$$\begin{aligned}\langle f_0 + g_0,\ f_1 + g_1,\ f_2 + g_2,\ \ldots\rangle &\longleftrightarrow \sum_{n=0}^{\infty}(f_n + g_n)x^n \\ &= \left(\sum_{n=0}^{\infty} f_n x^n\right) + \left(\sum_{n=0}^{\infty} g_n x^n\right) \\ &= F(x) + G(x).\end{aligned}$$

12.2.3 Right Shifting

Let's start over again with a simple sequence and its generating function:

$$\langle 1, 1, 1, 1, \ldots\rangle \longleftrightarrow \frac{1}{1-x}.$$

Now let's *right-shift* the sequence by adding k leading zeros:

$$\begin{aligned}\langle\overbrace{0, 0, \ldots, 0}^{k \text{ zeroes}}, 1, 1, 1, \ldots\rangle &\longleftrightarrow x^k + x^{k+1} + x^{k+2} + x^{k+3} + \cdots \\ &= x^k \cdot (1 + x + x^2 + x^3 + \cdots) \\ &= \frac{x^k}{1-x}.\end{aligned}$$

Evidently, adding k leading zeros to the sequence corresponds to multiplying the generating function by x^k. This holds true in general.

Rule 12.2.3 (Right-Shift Rule)**.** *If* $\langle f_0, f_1, f_2, \ldots\rangle \longleftrightarrow F(x)$*, then:*

$$\langle\overbrace{0, 0, \ldots, 0}^{k \text{ zeroes}}, f_0, f_1, f_2, \ldots\rangle \longleftrightarrow x^k \cdot F(x).$$

The idea behind this rule is that:

$$\begin{aligned}\langle\overbrace{0, 0, \ldots, 0}^{k \text{ zeroes}}, f_0, f_1, f_2, \ldots\rangle &\longleftrightarrow f_0 x^k + f_1 x^{k+1} + f_2 x^{k+2} + \cdots \\ &= x^k \cdot (f_0 + f_1 x + f_2 x^2 + f_3 x^3 + \cdots) \\ &= x^k \cdot F(x).\end{aligned}$$

12.2.4 Differentiation

What happens if we take the *derivative* of a generating function? As an example, let's differentiate the now-familiar generating function for an infinite sequence of 1's:

$$1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x}$$

$$\text{IMPLIES} \quad \frac{d}{dx}\left(1 + x + x^2 + x^3 + x^4 + \cdots\right) = \frac{d}{dx}\left(\frac{1}{1-x}\right)$$

$$\text{IMPLIES} \quad 1 + 2x + 3x^2 + 4x^3 + \cdots = \frac{1}{(1-x)^2}$$

$$\text{IMPLIES} \quad \langle 1, 2, 3, 4, \ldots \rangle \longleftrightarrow \frac{1}{(1-x)^2}. \tag{12.1}$$

We found a generating function for the sequence $\langle 1, 2, 3, 4, \ldots \rangle$ of positive integers!

In general, differentiating a generating function has two effects on the corresponding sequence: each term is multiplied by its index and the entire sequence is shifted left one place.

Rule 12.2.4 (Derivative Rule)**.** *If*

$$\langle f_0, f_1, f_2, f_3, \ldots \rangle \longleftrightarrow F(x),$$

then

$$\langle f_1, 2f_2, 3f_3, \ldots \rangle \longleftrightarrow F'(x).$$

The idea behind this rule is that:

$$\begin{aligned}\langle f_1, 2f_2, 3f_3, \ldots \rangle &\longleftrightarrow f_1 + 2f_2 x + 3f_3 x^2 + \cdots \\ &= \frac{d}{dx}\left(f_0 + f_1 x + f_2 x^2 + f_3 x^3 + \cdots\right) \\ &= \frac{d}{dx} F(x).\end{aligned}$$

The Derivative Rule is very useful. In fact, there is frequent, independent need for each of differentiation's two effects, multiplying terms by their index and left-shifting one place. Typically, we want just one effect and must somehow cancel out the other. For example, let's try to find the generating function for the sequence of squares, $\langle 0, 1, 4, 9, 16, \ldots \rangle$. If we could start with the sequence $\langle 1, 1, 1, 1, \ldots \rangle$ and multiply each term by its index two times, then we'd have the desired result:

$$\langle 0 \cdot 0,\ 1 \cdot 1,\ 2 \cdot 2,\ 3 \cdot 3,\ \ldots \rangle = \langle 0, 1, 4, 9, \ldots \rangle.$$

A challenge is that differentiation not only multiplies each term by its index, but also shifts the whole sequence left one place. However, the Right-Shift Rule 12.2.3 tells how to cancel out this unwanted left-shift: multiply the generating function by x.

Our procedure, therefore, is to begin with the generating function for $\langle 1, 1, 1, 1, \ldots \rangle$, differentiate, multiply by x, and then differentiate and multiply by x once more. Then

$$\begin{aligned}
&\langle 1,1,1,1,\ldots\rangle \longleftrightarrow \frac{1}{1-x}\\
\text{Derivative Rule:}\quad &\langle 1,2,3,4,\ldots\rangle \longleftrightarrow \frac{d}{dx}\frac{1}{1-x} = \frac{1}{(1-x)^2}\\
\text{Right-shift Rule:}\quad &\langle 0,1,2,3,\ldots\rangle \longleftrightarrow x\cdot\frac{1}{(1-x)^2} = \frac{x}{(1-x)^2}\\
\text{Derivative Rule:}\quad &\langle 1,4,9,16,\ldots\rangle \longleftrightarrow \frac{d}{dx}\frac{x}{(1-x)^2} = \frac{1+x}{(1-x)^3}\\
\text{Right-shift Rule:}\quad &\langle 0,1,4,9,\ldots\rangle \longleftrightarrow x\cdot\frac{1+x}{(1-x)^3} = \frac{x(1+x)}{(1-x)^3}
\end{aligned}$$

Thus, the generating function for squares is:

$$\frac{x(1+x)}{(1-x)^3}. \tag{12.2}$$

12.2.5 Products

Rule 12.2.5 (Product Rule)**.** *If*

$$\langle a_0, a_1, a_2, \ldots\rangle \longleftrightarrow A(x), \quad \textit{and} \quad \langle b_0, b_1, b_2, \ldots\rangle \longleftrightarrow B(x),$$

then

$$\langle c_0,\ c_1,\ c_2,\ \ldots\rangle \longleftrightarrow A(x)\cdot B(x),$$

where

$$c_n ::= a_0b_n + a_1b_{n-1} + a_2b_{n-2} + \cdots + a_nb_0.$$

To understand this rule, let

$$C(x) ::= A(x)\cdot B(x) = \sum_{n=0}^{\infty} c_n x^n.$$

We can evaluate the product $A(x) \cdot B(x)$ by using a table to identify all the cross-terms from the product of the sums:

	b_0x^0	b_1x^1	b_2x^2	b_3x^3	...
a_0x^0	$a_0b_0x^0$	$a_0b_1x^1$	$a_0b_2x^2$	$a_0b_3x^3$	...
a_1x^1	$a_1b_0x^1$	$a_1b_1x^2$	$a_1b_2x^3$	...	
a_2x^2	$a_2b_0x^2$	$a_2b_1x^3$	...		
a_3x^3	$a_3b_0x^3$	...			
$\vdots$	...				

Notice that all terms involving the same power of x lie on a diagonal. Collecting these terms together, we find that the coefficient of x^n in the product is the sum of all the terms on the $(n+1)$st diagonal, namely,

$$a_0b_n + a_1b_{n-1} + a_2b_{n-2} + \cdots + a_nb_0. \tag{12.3}$$

This expression (12.3) may be familiar from a signal processing course; the sequence $\langle c_0, c_1, c_2, \ldots \rangle$ is called the *convolution* of sequences $\langle a_0, a_1, a_2, \ldots \rangle$ and $\langle b_0, b_1, b_2, \ldots \rangle$.

12.3 Evaluating Sums

The product rule looks complicated. But it is surprisingly useful. For example, suppose that we set

$$B(x) = \frac{1}{1-x}.$$

Then $b_i = 1$ for $i \geq 0$ and the nth coefficient of $A(x)B(x)$ is

$$a_0 \cdot 1 + a_1 \cdot 1 + a_2 \cdot 1 + \cdots + a_n \cdot 1 = \sum_{i=0}^{n} a_i.$$

In other words, given any sequence $\langle a_0, a_1, a_2, \ldots \rangle$, we can compute

$$s_n = \sum_{i=0}^{n} a_i$$

for all n by simply multiplying the sequence's generating function by $1/(1-x)$. This is the Summation Rule.

Rule 12.3.1 (Summation Rule)**.** *If*

$$\langle a_0, a_1, a_2, \ldots \rangle \longleftrightarrow A(x),$$

then

$$\langle s_0, s_1, s_2, \ldots \rangle \longleftrightarrow \frac{A(x)}{1-x}$$

where

$$s_n = \sum_{i=0}^{n} a_i \quad \text{for } n \geq 0.$$

The Summation Rule sounds powerful, and it is! We know from Chapter 9 that computing sums is often not easy. But multiplying by $1/(1-x)$ is about as easy as it gets.

For example, suppose that we want to compute the sum of the first n squares

$$s_n = \sum_{i=0}^{n} i^2$$

and we forgot the method in Chapter 9. All we need to do is compute the generating function for $\langle 0, 1, 4, 9, \ldots \rangle$ and multiply by $1/(1-x)$. We already computed the generating function for $\langle 0, 1, 4, 9, \ldots \rangle$ in Equation 12.2—it is

$$\frac{x(1+x)}{(1-x)^3}.$$

Hence, the generating function for $\langle s_0, s_1, s_2, \ldots \rangle$ is

$$\frac{x(1+x)}{(1-x)^4}.$$

This means that $\sum_{i=0}^{n} i^2$ is the coefficient of x^n in $x(1+x)/(1-x)^4$.

That was pretty easy, but there is one problem—we have no idea how to determine the coefficient of x^n in $x(1+x)/(1-x)^4$! And without that, this whole endeavor (while magical) would be useless. Fortunately, there is a straightforward way to produce the sequence of coefficients from a generating function.

12.4 Extracting Coefficients

12.4.1 Taylor Series

Given a sequence of coefficients $\langle f_0, f_1, f_2, \ldots \rangle$, computing the generating function $F(x)$ is easy since

$$F(x) = f_0 + f_1 x + f_2 x^2 + \cdots .$$

To compute the sequence of coefficients from the generating function, we need to compute the *Taylor Series* for the generating function.

Rule 12.4.1 (Taylor Series)**.** *Let $F(x)$ be the generating function for the sequence*

$$\langle f_0, f_1, f_2, \ldots \rangle.$$

Then

$$f_0 = F(0)$$

and

$$f_n = \frac{F^{(n)}(0)}{n!}$$

for $n \geq 1$, where $F^{(n)}(0)$ is the nth derivative of $F(x)$ evaluated at $x = 0$.

This is because if

$$F(x) = f_0 + f_1 x + f_2 x^2 + \cdots ,$$

then

$$\begin{aligned} F(0) &= f_0 + f_1 \cdot 0 + f_2 \cdot 0^2 + \cdots \\ &= f_0. \end{aligned}$$

Also,

$$\begin{aligned} F'(x) &= \frac{d}{dx}(F(x)) \\ &= f_1 + 2f_2 x + 3f_3 x^2 + 4f_4 x^3 + \cdots \end{aligned}$$

and so

$$F'(0) = f_1,$$

as desired. Taking second derivatives, we find that

$$\begin{aligned} F''(x) &= \frac{d}{dx}(F'(x)) \\ &= 2f_2 + 3 \cdot 2f_3 x + 4 \cdot 3f_4 x^2 + \cdots \end{aligned}$$

and so

$$F''(0) = 2f_2,$$

which means that

$$f_2 = \frac{F''(0)}{2}.$$

In general,

$$\begin{aligned} F^{(n)} = n!f_n + (n+1)!f_{n+1}x + \frac{(n+2)!}{2}f_{n+2}x^2 + \cdots \\ + \frac{(n+k)!}{k!}f_{n+k}x^k + \cdots \end{aligned}$$

and so

$$F^{(n)}(0) = n!f_n$$

and

$$f_n = \frac{F^{(n)}(0)}{n!},$$

as claimed.

This means that

$$\left\langle F(0), F'(0), \frac{F''(0)}{2!}, \frac{F'''(0)}{3!}, \ldots, \frac{F^{(n)}(0)}{n!}, \ldots \right\rangle \longleftrightarrow F(x). \tag{12.4}$$

The sequence on the left-hand side of Equation 12.4 gives the well-known Taylor Series expansion for a function

$$F(x) = F(0) + F'(0)x + \frac{F''(0)}{2!}x^2 + \frac{F'''(0)}{3!}x^3 + \cdots + \frac{F^{(n)}(0)}{n!}x^n + \cdots.$$

12.4.2 Examples

Let's try this out on a familiar example:

$$F(x) = \frac{1}{1-x}.$$

Computing derivatives, we find that

$$\begin{aligned} F'(x) &= \frac{1}{(1-x)^2}, \\ F''(x) &= \frac{2}{(1-x)^3}, \\ F'''(x) &= \frac{2\cdot 3}{(1-x)^4}, \\ &\vdots \\ F^{(n)} &= \frac{n!}{(1-x)^{n+1}}. \end{aligned}$$

This means that the coefficient of x^n in $1/(1-x)$ is

$$\frac{F^{(n)}(0)}{n!} = \frac{n!}{n!\,(1-0)^{n+1}} = 1.$$

In other words, we have reconfirmed what we already knew; namely, that

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots.$$

Using a similar approach, we can establish some other well-known series:

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots, \\ e^{ax} &= 1 + ax + \frac{a^2}{2!}x^2 + \frac{a^3}{3!}x^3 + \cdots + \frac{a^n}{n!}x^n + \cdots, \\ \ln(1-x) &= -ax - \frac{a^2}{2}x^2 - \frac{a^3}{3}x^3 - \cdots - \frac{a^n}{n}x^n - \cdots. \end{aligned}$$

But what about the series for

$$F(x) = \frac{x(1+x)}{(1-x)^4}? \tag{12.5}$$

In particular, we need to know the coefficient of x^n in $F(x)$ to determine

$$s_n = \sum_{i=0}^{n} i^2.$$

While it is theoretically possible to compute the nth derivative of $F(x)$, the result is a bloody mess. Maybe these generating functions weren't such a great idea after all....

12.4.3 Massage Helps

In times of stress, a little massage can often help relieve the tension. The same is true for polynomials with painful derivatives. For example, let's take a closer look at Equation 12.5. If we massage it a little bit, we find that

$$F(x) = \frac{x + x^2}{(1-x)^4} = \frac{x}{(1-x)^4} + \frac{x^2}{(1-x)^4}. \tag{12.6}$$

The goal is to find the coefficient of x^n in $F(x)$. If you stare at Equation 12.6 long enough (or if you combine the Right-Shift Rule with the Addition Rule), you will notice that the coefficient of x^n in $F(x)$ is just the sum of

$$\text{the coefficient of } x^{n-1} \text{ in } \frac{1}{(1-x)^4} \text{ and}$$

$$\text{the coefficient of } x^{n-2} \text{ in } \frac{1}{(1-x)^4}.$$

Maybe there is some hope after all. Let's see if we can produce the coefficients for $1/(1-x)^4$. We'll start by looking at the derivatives:

$$\begin{aligned} F'(x) &= \frac{4}{(1-x)^5}, \\ F''(x) &= \frac{4 \cdot 5}{(1-x)^6}, \\ F'''(x) &= \frac{4 \cdot 5 \cdot 6}{(1-x)^7}, \\ &\vdots \\ F^{(n)}(x) &= \frac{(n+3)!}{6(1-x)^{n+4}}. \end{aligned}$$

This means that the nth coefficient of $1/(1-x)^4$ is

$$\frac{F^{(n)}(0)}{n!} = \frac{(n+3)!}{6n!} = \frac{(n+3)(n+2)(n+1)}{6}. \tag{12.7}$$

We are now almost done. Equation 12.7 means that the coefficient of x^{n-1} in $1/(1-x)^4$ is

$$\frac{(n+2)(n+1)n}{6} \tag{12.8}$$

and the coefficient[2] of x^{n-2} is

$$\frac{(n+1)n(n-1)}{6}. \tag{12.9}$$

Adding these values produces the desired sum

$$\begin{aligned}\sum_{i=0}^{n} i^2 &= \frac{(n+2)(n+1)n}{6} + \frac{(n+1)n(n-1)}{6} \\ &= \frac{(2n+1)(n+1)n}{6}.\end{aligned}$$

This matches Equation 9.14 from Chapter 9. Using generating functions to get the result may have seemed to be more complicated, but at least there was no need for guessing or solving a linear system of equations over 4 variables.

You might argue that the massage step was a little tricky. After all, how were you supposed to know that by converting $F(x)$ into the form shown in Equation 12.6, it would be sufficient to compute derivatives of $1/(1-x)^4$, which is easy, instead of derivatives of $x(1+x)/(1-x)^4$, which could be harder than solving a 64-disk Tower of Hanoi problem step-by-step?

The good news is that this sort of massage works for any generating function that is a ratio of polynomials. Even better, you probably already know how to do it from calculus—it's the method of *partial fractions*!

12.4.4 Partial Fractions

The idea behind partial fractions is to express a ratio of polynomials as a sum of a polynomial and terms of the form

$$\frac{cx^a}{(1-\alpha x)^b} \tag{12.10}$$

where a and b are integers and $b > a \geq 0$. That's because it is easy to compute derivatives of $1/(1-\alpha x)^b$ and thus it is easy to compute the coefficients of Equation 12.10. Let's see why.

Lemma 12.4.2. *If $b \in \mathbb{N}^+$, then the nth derivative of $1/(1-\alpha x)^b$ is*

$$\frac{(n+b-1)!\,\alpha^n}{(b-1)!\,(1-\alpha x)^{b+n}}.$$

[2]To be precise, Equation 12.8 holds for $n \geq 1$ and Equation 12.9 holds for $n \geq 2$. But since Equation 12.8 is 0 for $n = 1$ and Equation 12.9 is 0 for $n = 1, 2$, both equations hold for all $n \geq 0$.

Proof. The proof is by induction on n. The induction hypothesis $P(n)$ is the statement of the lemma.

Base case ($n = 1$): The first derivative is

$$\frac{b\alpha}{(1-\alpha x)^{b+1}}.$$

This matches

$$\frac{(1+b-1)!\,\alpha^1}{(b-1)!\,(1-\alpha x)^{b+1}} = \frac{b\alpha}{(1-\alpha x)^{b+1}},$$

and so $P(1)$ is true.

Induction step: We next assume $P(n)$ to prove $P(n+1)$ for $n \geq 1$. $P(n)$ implies that the nth derivative of $1/(1-\alpha x)^b$ is

$$\frac{(n+b-1)!\,\alpha^n}{(b-1)!\,(1-\alpha x)^{b+n}}.$$

Taking one more derivative reveals that the $(n+1)$st derivative is

$$\frac{(n+b-1)!\,(b+n)\alpha^{n+1}}{(b-1)!\,(1-\alpha x)^{b+n+1}} = \frac{(n+b)!\,\alpha^{n+1}}{(b-1)!\,(1-\alpha x)^{b+n+1}},$$

which means that $P(n+1)$ is true. Hence, the induction is complete. ■

Corollary 12.4.3. *If $a, b \in \mathbb{N}$ and $b > a \geq 0$, then for any $n \geq a$, the coefficient of x^n in*

$$\frac{cx^a}{(1-\alpha x)^b}$$

is

$$\frac{c(n-a+b-1)!\,\alpha^{n-a}}{(n-a)!\,(b-1)!}.$$

Proof. By the Taylor Series Rule, the nth coefficient of

$$\frac{1}{(1-\alpha x)^b}$$

is the nth derivative of this expression evaluated at $x = 0$ and then divided by $n!$. By Lemma 12.4.2, this is

$$\frac{(n+b-1)!\,\alpha^n}{n!\,(b-1)!\,(1-0)^{b+n}} = \frac{(n+b-1)!\,\alpha^n}{n!\,(b-1)!}.$$

By the Scaling Rule and the Right-Shift Rule, the coefficient of x^n in

$$\frac{cx^\alpha}{(1-\alpha x)^b}$$

is thus

$$\frac{c(n-a+b-1)!\,\alpha^{n-a}}{(n-a)!\,(b-1)!}.$$

as claimed. ■

Massaging a ratio of polynomials into a sum of a polynomial and terms of the form in Equation 12.10 takes a bit of work but is generally straightforward. We will show you the process by means of an example.

Suppose our generating function is the ratio

$$F(x) = \frac{4x^3+2x^2+3x+6}{2x^3-3x^2+1}. \tag{12.11}$$

The first step in massaging $F(x)$ is to get the degree of the numerator to be less than the degree of the denominator. This can be accomplished by dividing the numerator by the denominator and taking the remainder, just as in the Fundamental Theorem of Arithmetic—only now we have polynomials instead of numbers. In this case we have

$$\frac{4x^3+2x^2+3x+6}{2x^3-3x^2+1} = 2 + \frac{8x^2+3x+4}{2x^3-3x^2+1}.$$

The next step is to factor the denominator. This will produce the values of α for Equation 12.10. In this case,

$$\begin{aligned} 2x^3-3x^2+1 &= (2x+1)(x^2-2x+1) \\ &= (2x+1)(x-1)^2 \\ &= (1-x)^2(1+2x). \end{aligned}$$

We next find values c_1, c_2, c_3 so that

$$\frac{8x^2+3x+4}{2x^3-3x^2+1} = \frac{c_1}{1+2x} + \frac{c_2}{(1-x)^2} + \frac{c_3x}{(1-x)^2}. \tag{12.12}$$

This is done by cranking through the algebra:

$$\begin{aligned} \frac{c_1}{1+2x} + \frac{c_2}{(1-x)^2} + \frac{c_3x}{(1-x)^2} &= \frac{c_1(1-x)^2 + c_2(1+2x) + c_3x(1+2x)}{(1+2x)(1-x)^2} \\ &= \frac{c_1 - 2c_1x + c_1x^2 + c_2 + 2c_2x + c_3x + 2c_3x^2}{2x^3-3x^2+1} \\ &= \frac{c_1+c_2+(-2c_1+2c_2+c_3)x + (c_1+2c_3)x^2}{2x^3-3x^2+1}. \end{aligned}$$

For Equation 12.12 to hold, we need

$$\begin{aligned} 8 &= c_1 + 2c_3, \\ 3 &= -2c_1 + 2c_2 + c_3, \\ 4 &= c_1 + c_2. \end{aligned}$$

Solving these equations, we find that $c_1 = 2$, $c_2 = 2$, and $c_3 = 3$. Hence,

$$\begin{aligned} F(x) &= \frac{4x^3 + 2x^2 + 3x + 6}{2x^3 - 3x^2 + 1} \\ &= 2 + \frac{2}{1+2x} + \frac{2}{(1-x)^2} + \frac{3x}{(1-x)^2}. \end{aligned}$$

Our massage is done! We can now compute the coefficients of $F(x)$ using Corollary 12.4.3 and the Sum Rule. The result is

$$f_0 = 2 + 2 + 2 = 6$$

and

$$\begin{aligned} f_n &= \frac{2(n-0+1-1)!\,(-2)^{n-0}}{(n-0)!\,(1-1)!} \\ &\quad + \frac{2(n-0+2-1)!\,(1)^{n-0}}{(n-0)!\,(2-1)!} \\ &\quad + \frac{3(n-1+2-1)!\,(1)^{n-1}}{(n-1)!\,(2-1)!} \\ &= (-1)^n 2^{n+1} + 2(n+1) + 3n \\ &= (-1)^n 2^{n+1} + 5n + 2 \end{aligned}$$

for $n \geq 1$.

Aren't you glad that you know that? Actually, this method turns out to be useful in solving linear recurrences, as we'll see in the next section.

12.5 Solving Linear Recurrences

Generating functions can be used to find a solution to any linear recurrence. We'll show you how this is done by means of a familiar example, the Fibonacci recurrence, so that you can more easily understand the similarities and differences of this approach and the method we showed you in Chapter 10.

12.5.1 Finding the Generating Function

Let's begin by recalling the definition of the Fibonacci numbers:

$$\begin{aligned} f_0 &= 0 \\ f_1 &= 1 \\ f_n &= f_{n-1} + f_{n-2} \qquad \text{for } n \geq 2. \end{aligned}$$

We can expand the final clause into an infinite sequence of equations. Thus, the Fibonacci numbers are defined by:

$$\begin{aligned} f_0 &= 0 \\ f_1 &= 1 \\ f_2 &= f_1 + f_0 \\ f_3 &= f_2 + f_1 \\ f_4 &= f_3 + f_2 \\ &\vdots \end{aligned}$$

The overall plan is to *define* a function $F(x)$ that generates the sequence on the left side of the equality symbols, which are the Fibonacci numbers. Then we *derive* a function that generates the sequence on the right side. Finally, we equate the two and solve for $F(x)$. Let's try this. First, we define:

$$F(x) = f_0 + f_1 x + f_2 x^2 + f_3 x^3 + f_4 x^4 + \cdots .$$

Now we need to derive a generating function for the sequence:

$$\langle 0,\ 1,\ f_1 + f_0,\ f_2 + f_1,\ f_3 + f_2,\ \ldots \rangle .$$

One approach is to break this into a sum of three sequences for which we know generating functions and then apply the Addition Rule:

$$\begin{array}{rccccccl} \langle 0, & 1, & 0, & 0, & 0, & \ldots \rangle & \longleftrightarrow & x \\ \langle 0, & f_0, & f_1, & f_2, & f_3, & \ldots \rangle & \longleftrightarrow & xF(x) \\ +\langle 0, & 0, & f_0, & f_1, & f_2, & \ldots \rangle & \longleftrightarrow & x^2F(x) \\ \hline \langle 0, & 1 + f_0, & f_1 + f_0, & f_2 + f_1, & f_3 + f_2, & \ldots \rangle & \longleftrightarrow & x + xF(x) + x^2F(x) \end{array}$$

This sequence is almost identical to the right sides of the Fibonacci equations. The one blemish is that the second term is $1 + f_0$ instead of simply 1. However, this amounts to nothing, since $f_0 = 0$ anyway.

If we equate $F(x)$ with the new function $x + xF(x) + x^2F(x)$, then we're implicitly writing down *all* of the equations that define the Fibonacci numbers in one fell swoop:

$$\begin{array}{cccccccccc} F(x) & = & f_0 + & f_1 & x + & f_2 & x^2 + & f_3 & x^3 + \cdots \\ \| & & \| & \| & & \| & & \| & \\ x + xF(x) + x^2F(x) & = & 0 + & (1 + f_0)x & + & (f_1 + f_0)x^2 & + & (f_2 + f_1)x^3 & + \cdots \end{array}$$

Solving for $F(x)$ gives the generating function for the Fibonacci sequence:

$$F(x) = x + xF(x) + x^2F(x)$$

so

$$F(x) = \frac{x}{1 - x - x^2}. \tag{12.13}$$

This is pretty cool. After all, who would have thought that the Fibonacci numbers are precisely the coefficients of such a simple function? Even better, this function is a ratio of polynomials and so we can use the method of partial fractions from Section 12.4.4 to find a closed-form expression for the nth Fibonacci number.

12.5.2 Extracting the Coefficients

Repeated differentiation of Equation 12.13 would be very painful. But it is easy to use the method of partial fractions to compute the coefficients. Since the degree of the numerator in Equation 12.13 is less than the degree of the denominator, the first step is to factor the denominator:

$$1 - x - x^2 = (1 - \alpha_1 x)(1 - \alpha_2 x)$$

where $\alpha_1 = (1 + \sqrt{5})/2$ and $\alpha_2 = (1 - \sqrt{5})/2$. These are the same as the roots of the characteristic equation for the Fibonacci recurrence that we found in Chapter 10. That is not a coincidence.

The next step is to find c_1 and c_2 that satisfy

$$\begin{aligned} \frac{x}{1 - x - x^2} &= \frac{c_1}{1 - \alpha_1 x} + \frac{c_2}{1 - \alpha_2 x} \\ &= \frac{c_1(1 - \alpha_2 x) + c_2(1 - \alpha_1 x)}{(1 - \alpha_1 x)(1 - \alpha_2 x)} \\ &= \frac{c_1 + c_2 - (c_1\alpha_2 + c_2\alpha_1)x}{1 - x - x^2}. \end{aligned}$$

Hence,

$$c_1 + c_2 = 0 \qquad \text{and} \qquad -(c_1\alpha_2 + c_2\alpha_1) = 1.$$

Solving these equations, we find that

$$c_1 = \frac{1}{\alpha_1 - \alpha_2} = \frac{1}{\sqrt{5}}$$
$$c_2 = \frac{-1}{\alpha_1 - \alpha_2} = \frac{-1}{\sqrt{5}}.$$

We can now use Corollary 12.4.3 and the Sum Rule to conclude that

$$\begin{aligned} f_n &= \frac{\alpha_1^n}{\sqrt{5}} - \frac{\alpha_2^n}{\sqrt{5}} \\ &= \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right). \end{aligned}$$

This is exactly the same formula we derived for the nth Fibonacci number in Chapter 10.

12.5.3 General Linear Recurrences

The method that we just used to solve the Fibonacci recurrence can also be used to solve general linear recurrences of the form

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_d f_{n-d} + g_n$$

for $n \geq d$. The generating function for $\langle f_0, f_1, f_2, \ldots \rangle$ is

$$F(x) = \frac{h(x) + G(x)}{1 - a_1 x - a_2 x^2 - \cdots - a_d x^d}$$

where $G(x)$ is the generating function for the sequence

$$\langle \overbrace{0, 0, \ldots, 0}^{d}, g_d, g_{d+1}, g_{d+2}, \ldots \rangle$$

and $h(x)$ is a polynomial of degree at most $d-1$ that is based on the values of f_0, $f_1, \ldots, f_{d-1}$. In particular,

$$h(x) = \sum_{i=0}^{d-1} h_i x^i$$

where

$$h_i = f_0 - a_1 f_{i-1} - a_2 f_{i-2} - \cdots - a_i f_0$$

for $0 \leq i < d$.

To solve the recurrence, we use the method of partial fractions described in Section 12.4.4 to find a closed-form expression for $F(x)$. This can be easy or hard to do depending on $G(x)$.

12.6 Counting with Generating Functions

Generating functions are particularly useful for solving counting problems. In particular, problems involving choosing items from a set often lead to nice generating functions by letting the coefficient of x^n be the number of ways to choose n items.

12.6.1 Choosing Distinct Items from a Set

The generating function for binomial coefficients follows directly from the Binomial Theorem:

$$\left\langle \binom{k}{0}, \binom{k}{1}, \binom{k}{2}, \ldots, \binom{k}{k}, 0, 0, 0, \ldots \right\rangle \longleftrightarrow \binom{k}{0} + \binom{k}{1}x + \binom{k}{2}x^2 + \cdots + \binom{k}{k}x^k = (1+x)^k$$

Thus, the coefficient of x^n in $(1+x)^k$ is $\binom{k}{n}$, the number of ways to choose n distinct items[3] from a set of size k. For example, the coefficient of x^2 is $\binom{k}{2}$, the number of ways to choose 2 items from a set with k elements. Similarly, the coefficient of x^{k+1} is the number of ways to choose $k+1$ items from a size k set, which is zero.

12.6.2 Building Generating Functions that Count

Often we can translate the description of a counting problem directly into a generating function for the solution. For example, we could figure out that $(1+x)^k$ generates the number of ways to select n distinct items from a k-element set without resorting to the Binomial Theorem or even fussing with binomial coefficients! Let's see how.

First, consider a single-element set $\{a_1\}$. The generating function for the number of ways to select n elements from this set is simply $1+x$: we have 1 way to select zero elements, 1 way to select one element, and 0 ways to select more than one element. Similarly, the number of ways to select n elements from the set $\{a_2\}$ is also given by the generating function $1+x$. The fact that the elements differ in the two cases is irrelevant.

Now here is the main trick: *the generating function for choosing elements from a union of disjoint sets is the product of the generating functions for choosing from each set.* We'll justify this in a moment, but let's first look at an example. According to this principle, the generating function for the number of ways to select

[3] Watch out for the reversal of the roles that k and n played in earlier examples; we're led to this reversal because we've been using n to refer to the power of x in a power series.

n elements from the $\{a_1, a_2\}$ is:

$$\underbrace{(1+x)}_{\text{select from } \{a_1\}} \cdot \underbrace{(1+x)}_{\text{select from } \{a_2\}} = \underbrace{(1+x)^2}_{\text{select from } \{a_1, a_2\}} = 1 + 2x + x^2.$$

Sure enough, for the set $\{a_1, a_2\}$, we have 1 way to select zero elements, 2 ways to select one element, 1 way to select two elements, and 0 ways to select more than two elements.

Repeated application of this rule gives the generating function for selecting n items from a k-element set $\{a_1, a_2, \ldots, a_k\}$:

$$\underbrace{(1+x)}_{\text{select from } \{a_1\}} \cdot \underbrace{(1+x)}_{\text{select from } \{a_2\}} \cdots \underbrace{(1+x)}_{\text{select from } \{a_k\}} = \underbrace{(1+x)^k}_{\substack{\text{select from} \\ \{a_1, a_2, \ldots, a_k\}}}$$

This is the same generating function that we obtained by using the Binomial Theorem. But this time around, we translated directly from the counting problem to the generating function.

We can extend these ideas to a general principle:

Rule 12.6.1 (Convolution Rule)**.** *Let $A(x)$ be the generating function for selecting items from set $\mathcal{A}$, and let $B(x)$ be the generating function for selecting items from set $\mathcal{B}$. If $\mathcal{A}$ and $\mathcal{B}$ are disjoint, then the generating function for selecting items from the union $\mathcal{A} \cup \mathcal{B}$ is the product $A(x) \cdot B(x)$.*

This rule is rather ambiguous: what exactly are the rules governing the selection of items from a set? Remarkably, the Convolution Rule remains valid under *many* interpretations of selection. For example, we could insist that distinct items be selected or we might allow the same item to be picked a limited number of times or any number of times. Informally, the only restrictions are that (1) the order in which items are selected is disregarded and (2) restrictions on the selection of items from sets $\mathcal{A}$ and $\mathcal{B}$ also apply in selecting items from $\mathcal{A} \cup \mathcal{B}$. (Formally, there must be a bijection between n-element selections from $\mathcal{A} \cup \mathcal{B}$ and ordered pairs of selections from $\mathcal{A}$ and $\mathcal{B}$ containing a total of n elements.)

To count the number of ways to select n items from $\mathcal{A} \cup \mathcal{B}$, we observe that we can select n items by choosing j items from $\mathcal{A}$ and $n - j$ items from $\mathcal{B}$, where j is any number from 0 to n. This can be done in $a_j b_{n-j}$ ways. Summing over all the possible values of j gives a total of

$$a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0$$

ways to select n items from $\mathcal{A} \cup \mathcal{B}$. By the Product Rule, this is precisely the coefficient of x^n in the series for $A(x)B(x)$.

12.6.3 Choosing Items with Repetition

The first counting problem we considered was the number of ways to select a dozen doughnuts when five flavors were available. We can generalize this question as follows: in how many ways can we select n items from a k-element set if we're allowed to pick the same item multiple times? In these terms, the doughnut problem asks how many ways we can select $n = 12$ doughnuts from the set of $k = 5$ flavors

$$\{\text{chocolate, lemon-filled, sugar, glazed, plain}\}$$

where, of course, we're allowed to pick several doughnuts of the same flavor. Let's approach this question from a generating functions perspective.

Suppose we make n choices (with repetition allowed) of items from a set containing a single item. Then there is one way to choose zero items, one way to choose one item, one way to choose two items, etc. Thus, the generating function for choosing n elements with repetition from a 1-element set is:

$$\langle 1, 1, 1, 1, \ldots \rangle \longleftrightarrow 1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}.$$

The Convolution Rule says that the generating function for selecting items from a union of disjoint sets is the product of the generating functions for selecting items from each set:

$$\underbrace{\frac{1}{1-x}}_{\text{choose } a_1\text{'s}} \cdot \underbrace{\frac{1}{1-x}}_{\text{choose } a_2\text{'s}} \cdots \underbrace{\frac{1}{1-x}}_{\text{choose } a_k\text{'s}} = \underbrace{\frac{1}{(1-x)^k}}_{\substack{\text{repeatedly choose from} \\ \{a_1, a_2, \ldots, a_k\}}}$$

Therefore, the generating function for choosing items from a k-element set with repetition allowed is $1/(1-x)^k$. Computing derivatives and applying the Taylor Series Rule, we can find that the coefficient of x^n in $1/(1-x)^k$ is

$$\binom{n+k-1}{n}.$$

This is the Bookkeeper Rule from Chapter 11—namely there are $\binom{n+k-1}{n}$ ways to select n items with replication from a set of k items.

12.6.4 Fruit Salad

In this chapter, we have covered a lot of methods and rules for using generating functions. We'll now do an example that demonstrates how the rules and methods can be combined to solve a more challenging problem—making fruit salad.

In how many ways can we make a salad with n fruits subject to the following constraints?

- The number of apples must be even.
- The number of bananas must be a multiple of 5.
- There can be at most four oranges.
- There can be at most one pear.

For example, there are 7 ways to make a salad with 6 fruits:

Apples	6	4	4	2	2	0	0
Bananas	0	0	0	0	0	5	5
Oranges	0	2	1	4	3	1	0
Pears	0	0	1	0	1	0	1

These constraints are so complicated that the problem seems hopeless! But generating functions can solve the problem in a straightforward way.

Let's first construct a generating function for choosing apples. We can choose a set of 0 apples in one way, a set of 1 apple in zero ways (since the number of apples must be even), a set of 2 apples in one way, a set of 3 apples in zero ways, and so forth. So we have:

$$A(x) = 1 + x^2 + x^4 + x^6 + \cdots = \frac{1}{1-x^2}.$$

Similarly, the generating function for choosing bananas is:

$$B(x) = 1 + x^5 + x^{10} + x^{15} + \cdots = \frac{1}{1-x^5}.$$

We can choose a set of 0 oranges in one way, a set of 1 orange in one way, and so on. However, we can not choose more than four oranges, so we have the generating function:

$$O(x) = 1 + x + x^2 + x^3 + x^4 = \frac{1-x^5}{1-x}.$$

Here we're using the geometric sum formula. Finally, we can choose only zero or one pear, so we have:

$$P(x) = 1 + x.$$

The Convolution Rule says that the generating function for choosing from among all four kinds of fruit is:

$$\begin{aligned} A(x)B(x)O(x)P(x) &= \frac{1}{1-x^2}\frac{1}{1-x^5}\frac{1-x^5}{1-x}(1+x) \\ &= \frac{1}{(1-x)^2} \\ &= 1+2x+3x^2+4x^3+\cdots. \end{aligned}$$

Almost everything cancels! We're left with $1/(1-x)^2$, which we found a power series for earlier: the coefficient of x^n is simply $n+1$. Thus, the number of ways to make a salad with n fruits is just $n+1$. This is consistent with the example we worked out at the start, since there were 7 different salads containing 6 fruits. *Amazing!*

13 Infinite Sets

So you might be wondering how much is there to say about an infinite set other than, well, it has an infinite number of elements. Of course, an infinite set does have an infinite number of elements, but it turns out that not all infinite sets have the same size—some are bigger than others! And, understanding infinity is not as easy as you might think. Some of the toughest questions in mathematics involve infinite sets.

Why should you care? Indeed, isn't computer science only about finite sets? Not exactly. For example, we deal with the set of natural numbers $\mathbb{N}$ all the time and it is an infinite set. In fact, that is why we have induction: to reason about predicates over $\mathbb{N}$. Infinite sets are also important in Part IV of the text when we talk about random variables over potentially infinite sample spaces.

So sit back and open your mind for a few moments while we take a very brief look at *infinity*.

13.1 Injections, Surjections, and Bijections

We know from Theorem 7.2.1 that if there is an injection or surjection between two finite sets, then we can say something about the relative sizes of the two sets. The same is true for infinite sets. In fact, relations are the primary tool for determining the relative size of infinite sets.

Definition 13.1.1. Given any two sets A and B, we say that

A surj B	iff there is a surjection from A to B,
A inj B	iff there is an injection from A to B,
A bij B	iff there is a bijection between A and B, and
A strict B	iff there is a surjection from A to B but there is *no* bijection from B to A.

Restating Theorem 7.2.1 with this new terminology, we have:

Theorem 13.1.2. *For any pair of* finite *sets A and B,*

$$\begin{aligned} |A| \geq |B| \quad &\textit{iff} \quad A \text{ surj } B, \\ |A| \leq |B| \quad &\textit{iff} \quad A \text{ inj } B, \\ |A| = |B| \quad &\textit{iff} \quad A \text{ bij } B, \\ |A| > |B| \quad &\textit{iff} \quad A \text{ strict } B. \end{aligned}$$

Theorem 13.1.2 suggests a way to generalize size comparisons to infinite sets; namely, we can think of the relation surj as an "at least as big" relation between sets, even if they are infinite. Similarly, the relation bij can be regarded as a "same size" relation between (possibly infinite) sets, and strict can be thought of as a "strictly bigger" relation between sets.

Note that we haven't, and won't, define what the size of an infinite set is. The definition of infinite "sizes" is cumbersome and technical, and we can get by just fine without it. All we need are the "as big as" and "same size" relations, surj and bij, between sets.

But there's something else to watch out for. We've referred to surj as an "as big as" relation and bij as a "same size" relation on sets. Most of the "as big as" and "same size" properties of surj and bij on finite sets do carry over to infinite sets, but *some important ones don't*—as we're about to show. So you have to be careful: don't assume that surj has any particular "as big as" property on *infinite* sets until it's been proved.

Let's begin with some familiar properties of the "as big as" and "same size" relations on finite sets that do carry over exactly to infinite sets:

Theorem 13.1.3. *For any sets, A, B, and C,*

1. A surj B *and* B surj C IMPLIES A surj C.
2. A bij B *and* B bij C IMPLIES A bij C.
3. A bij B IMPLIES B bij A.

Parts 1 and 2 of Theorem 13.1.3 follow immediately from the fact that compositions of surjections are surjections, and likewise for bijections. Part 3 follows from the fact that the inverse of a bijection is a bijection. We'll leave a proof of these facts to the problems.

Another familiar property of finite sets carries over to infinite sets, but this time it's not so obvious:

Theorem 13.1.4 (Schröder-Bernstein)**.** *For any pair of sets A and B, if A* surj B *and B* surj A*, then A* bij B.

The Schröder-Bernstein Theorem says that if A is at least as big as B and, conversely, B is at least as big as A, then A is the same size as B. Phrased this way, you might be tempted to take this theorem for granted, but that would be a mistake. For infinite sets A and B, the Schröder-Bernstein Theorem is actually pretty technical. Just because there is a surjective function $f : A \to B$—which need not be a bijection—and a surjective function $g : B \to A$—which also need not

be a bijection—it's not at all clear that there must be a bijection $h : A \to B$. The challenge is to construct h from parts of both f and g. We'll leave the actual construction to the problems.

13.1.1 Infinity Is Different

A basic property of finite sets that does *not* carry over to infinite sets is that adding something new makes a set bigger. That is, if A is a finite set and $b \notin A$, then $|A \cup \{b\}| = |A| + 1$, and so A and $A \cup \{b\}$ are not the same size. But if A is infinite, then these two sets *are* the same size!

Theorem 13.1.5. *Let A be a set and $b \notin A$. Then A is infinite iff A* bij $A \cup \{b\}$.

Proof. Since A is *not* the same size as $A \cup \{b\}$ when A is finite, we only have to show that $A \cup \{b\}$ *is* the same size as A when A is infinite.

That is, we have to find a bijection between $A \cup \{b\}$ and A when A is infinite. Since A is infinite, it certainly has at least one element; call it a_0. Since A is infinite, it has at least two elements, and one of them must not be equal to a_0; call this new element a_1. Since A is infinite, it has at least three elements, one of which must not equal a_0 or a_1; call this new element a_2. Continuing in this way, we conclude that there is an infinite sequence $a_0, a_1, a_2, \ldots, a_n, \ldots$, of different elements of A. Now it's easy to define a bijection $f : A \cup \{b\} \to A$:

$$\begin{aligned} f(b) &::= a_0, \\ f(a_n) &::= a_{n+1} && \text{for } n \in \mathbb{N}, \\ f(a) &::= a && \text{for } a \in A - \{b, a_0, a_1, \ldots\}. \end{aligned}$$

■

13.2 Countable Sets

13.2.1 Definitions

A set C is *countable* iff its elements can be listed in order, that is, the distinct elements in C are precisely

$$c_0, c_1, \ldots, c_n, \ldots.$$

This means that if we defined a function f on the nonnegative integers by the rule that $f(i) ::= c_i$, then f would be a bijection from $\mathbb{N}$ to C. More formally,

Definition 13.2.1. A set C is *countably infinite* iff $\mathbb{N}$ bij C. A set is *countable* iff it is finite or countably infinite.

Discrete mathematics is often defined as the mathematics of countable sets and so it is probably worth spending a little time understanding what it means to be countable and why countable sets are so special. For example, a small modification of the proof of Theorem 13.1.5 shows that countably infinite sets are the "smallest" infinite sets; namely, if A is any infinite set, then A surj $\mathbb{N}$.

13.2.2 Unions

Since adding one new element to an infinite set doesn't change its size, it's obvious that neither will adding any *finite* number of elements. It's a common mistake to think that this proves that you can throw in countably infinitely many new elements—just because it's ok to do something any finite number of times doesn't make it ok to do it an infinite number of times.

For example, suppose that you have two countably infinite sets $A = \{a_0, a_1, a_2, \ldots\}$ and $B = \{b_0, b_1, b_2, \ldots\}$. You might try to show that $A \cup B$ is countable by making the following "list" for $A \cup B$:

$$a_0, a_1, a_2, \ldots, b_0, b_1, b_2, \ldots \tag{13.1}$$

But this is not a valid argument because Equation 13.1 is not a list. The key property required for listing the elements in a countable set is that for any element in the set, you can determine its finite index in the list. For example, a_i shows up in position i in Equation 13.1, but there is no index in the supposed "list" for any of the b_i. Hence, Equation 13.1 is not a valid list for the purposes of showing that $A \cup B$ is countable when A is infinite. Equation 13.1 is only useful when A is finite.

It turns out you really can add a countably infinite number of new elements to a countable set and still wind up with just a countably infinite set, but another argument is needed to prove this.

Theorem 13.2.2. *If A and B are countable sets, then so is $A \cup B$.*

Proof. Suppose the list of distinct elements of A is $a_0, a_1, \ldots$, and the list of B is $b_0, b_1, \ldots$. Then a valid way to list all the elements of $A \cup B$ is

$$a_0, b_0, a_1, b_1, \ldots, a_n, b_n, \ldots. \tag{13.2}$$

Of course this list will contain duplicates if A and B have elements in common, but then deleting all but the first occurrence of each element in Equation 13.2 leaves a list of all the distinct elements of A and B. ■

Note that the list in Equation 13.2 does not have the same defect as the purported "list" in Equation 13.1, since every item in $A \cup B$ has a finite index in the list created in Theorem 13.2.2.

	b_0	b_1	b_2	b_3	$\ldots$
a_0	c_0	c_1	c_4	c_9	
a_1	c_3	c_2	c_5	c_{10}	
a_2	c_8	c_7	c_6	c_{11}	
a_3	c_{15}	c_{14}	c_{13}	c_{12}	
$\vdots$					$\ddots$

Figure 13.1 A listing of the elements of $C = A \times B$ where $A = \{a_0, a_1, a_2, \ldots\}$ and $B = \{b_0, b_1, b_2, \ldots\}$ are countably infinite sets. For example, $c_5 = (a_1, b_2)$.

13.2.3 Cross Products

Somewhat surprisingly, cross products of countable sets are also countable. At first, you might be tempted to think that "infinity times infinity" (whatever that means) somehow results in a larger infinity, but this is not the case.

Theorem 13.2.3. *The cross product of two countable sets is countable.*

Proof. Let A and B be any pair of countable sets. To show that $C = A \times B$ is also countable, we need to find a listing of the elements

$$\{(a, b) \mid a \in A, b \in B\}.$$

There are many such listings. One is shown in Figure 13.1 for the case when A and B are both infinite sets. In this listing, (a_i, b_j) is the kth element in the list for C where

$$\begin{aligned} &a_i \text{ is the } i\text{th element in } A,\\ &b_j \text{ is the } j\text{th element in } B, \text{ and}\\ &k = \max(i, j)^2 + i + \max(i - j, 0). \end{aligned}$$

The task of finding a listing when one or both of A and B are finite is left to the problems at the end of the chapter.

■

13.2.4 $\mathbb{Q}$ Is Countable

Theorem 13.2.3 also has a surprising Corollary; namely that the set of rational numbers is countable.

Corollary 13.2.4. *The set of rational numbers $\mathbb{Q}$ is countable.*

Proof. Since $\mathbb{Z}\times\mathbb{Z}$ is countable by Theorem 13.2.3, it suffices to find a surjection f from $\mathbb{Z}\times\mathbb{Z}$ to $\mathbb{Q}$. This is easy to to since

$$f(a,b) = \begin{cases} a/b & \text{if } b \neq 0 \\ 0 & \text{if } b = 0 \end{cases}$$

is one such surjection. ■

At this point, you may be thinking that every set is countable. That is *not* the case. In fact, as we will shortly see, there are many infinite sets that are uncountable, including the set of real numbers $\mathbb{R}$.

13.3 Power Sets Are Strictly Bigger

It turns out that the ideas behind Russell's Paradox, which caused so much trouble for the early efforts to formulate Set Theory, also lead to a correct and astonishing fact discovered by Georg Cantor in the late nineteenth century: infinite sets are *not all the same size*.

Theorem 13.3.1. *For any set A, the power set $\mathcal{P}(A)$ is strictly bigger than A.*

Proof. First of all, $\mathcal{P}(A)$ is as big as A: for example, the partial function $f : \mathcal{P}(A) \to A$ where $f(\{a\}) ::= a$ for $a \in A$ is a surjection.

To show that $\mathcal{P}(A)$ is strictly bigger than A, we have to show that if g is a function from A to $\mathcal{P}(A)$, then g is not a surjection. So, mimicking Russell's Paradox, define

$$A_g ::= \{\, a \in A \mid a \notin g(a) \,\}.$$

A_g is a well-defined subset of A, which means it is a member of $\mathcal{P}(A)$. But A_g can't be in the range of g, because if it were, we would have

$$A_g = g(a_0)$$

for some $a_0 \in A$. So by definition of A_g,

$$a \in g(a_0) \quad \text{iff} \quad a \in A_g \quad \text{iff} \quad a \notin g(a)$$

for all $a \in A$. Now letting $a = a_0$ yields the contradiction

$$a_0 \in g(a_0) \quad \text{iff} \quad a_0 \notin g(a_0).$$

So g is not a surjection, because there is an element in the power set of A, namely the set A_g, that is not in the range of g. ■

13.3.1 $\mathbb{R}$ Is Uncountable

To prove that the set of real numbers is uncountable, we will show that there is a surjection from $\mathbb{R}$ to $\mathcal{P}(\mathbb{N})$ and then apply Theorem 13.3.1 to $\mathcal{P}(\mathbb{N})$.

Lemma 13.3.2. $\mathbb{R}$ surj $\mathcal{P}(\mathbb{N})$.

Proof. Let $A \subset \mathbb{N}$ be any subset of the natural numbers. Since $\mathbb{N}$ is countable, this means that A is countable and thus that $A = \{a_0, a_1, a_2, \dots\}$. For each $i \geq 0$, define $\text{bin}(a_i)$ to be the binary representation of a_i. Let x_A be the real number using only digits 0, 1, 2 as follows:

$$x_A ::= 0.2\,\text{bin}(a_0)2\,\text{bin}(a_1)2\,\text{bin}(a_2)2\dots \tag{13.3}$$

We can then define a surjection $f : \mathbb{R} \to \mathcal{P}(\mathbb{N})$ as follows:

$$f(x) = \begin{cases} A & \text{if } x = x_A \text{ for some } A \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\mathbb{R}$ surj $\mathcal{P}(\mathbb{N})$. ■

Corollary 13.3.3. $\mathbb{R}$ *is uncountable.*

Proof. By contradiction. Assume $\mathbb{R}$ is countable. Then $\mathbb{N}$ surj $\mathbb{R}$. By Lemma 13.3.2, $\mathbb{R}$ surj $\mathcal{P}(\mathbb{N})$. Hence $\mathbb{N}$ surj $\mathcal{P}(\mathbb{N})$. This contradicts Theorem 13.3.1 for the case when $A = \mathbb{N}$. ■

So the set of rational numbers and the set of natural numbers have the same size, but the set of real numbers is strictly larger. In fact, $\mathbb{R}$ bij $\mathcal{P}(N)$, but we won't prove that here.

Is there anything bigger?

13.3.2 Even Larger Infinities

There are lots of different sizes of infinite sets. For example, starting with the infinite set $\mathbb{N}$ of nonnegative integers, we can build the infinite sequence of sets

$$\mathbb{N},\ \mathcal{P}(\mathbb{N}),\ \mathcal{P}(\mathcal{P}(\mathbb{N})),\ \mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N}))), \dots$$

By Theorem 13.3.1, each of these sets is strictly bigger than all the preceding ones. But that's not all, the union of all the sets in the sequence is strictly bigger than each set in the sequence. In this way, you can keep going, building still bigger infinities.

13.3.3 The Continuum Hypothesis

Georg Cantor was the mathematician who first developed the theory of infinite sizes (because he thought he needed it in his study of Fourier series). Cantor raised the question whether there is a set whose size is strictly between the "smallest" infinite set, $\mathbb{N}$, and $\mathcal{P}(\mathbb{N})$. He guessed not:

Cantor's Continuum Hypothesis. *There is no set A such that $\mathcal{P}(\mathbb{N})$ is strictly bigger than A and A is strictly bigger than $\mathbb{N}$.*

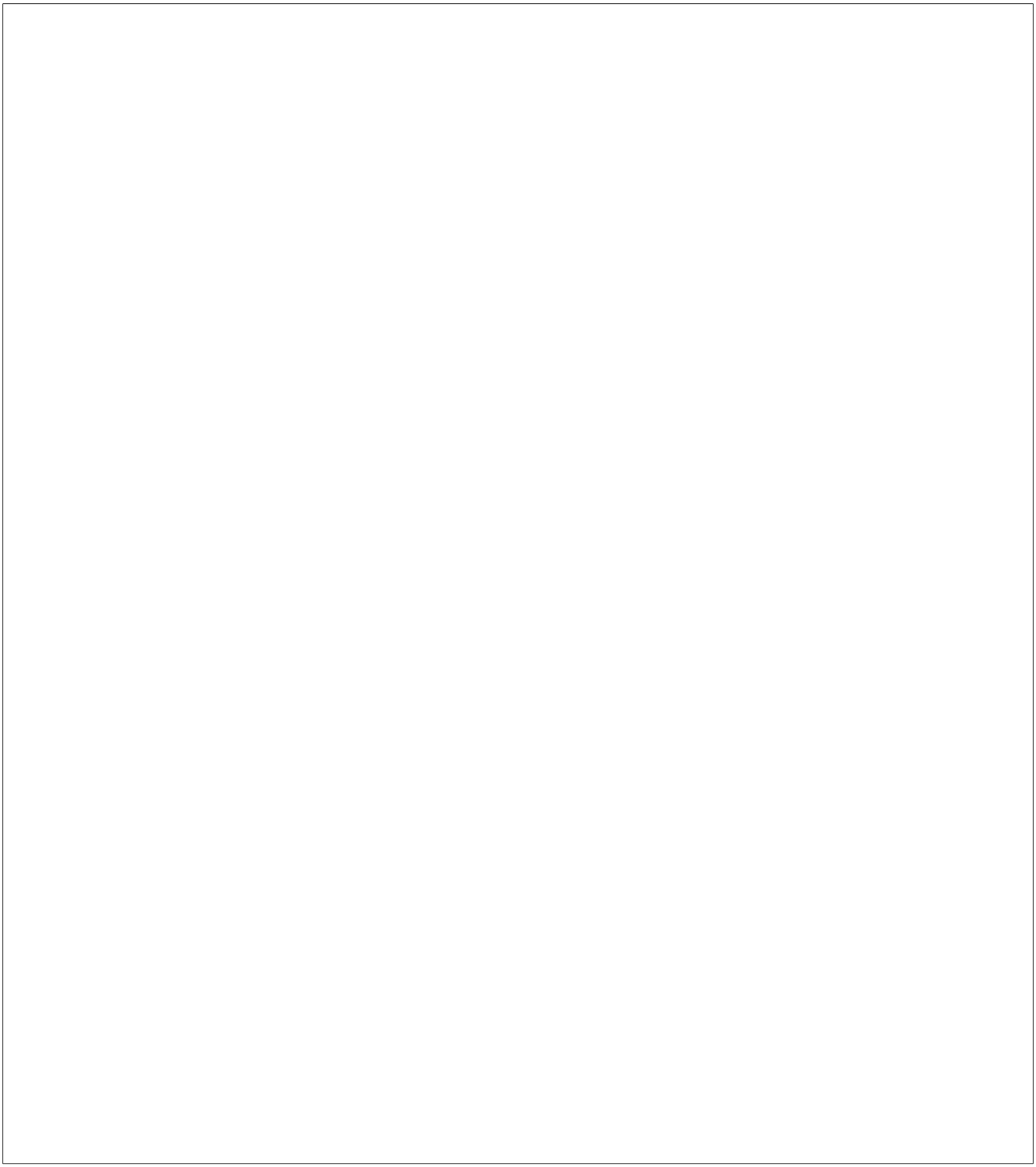
The Continuum Hypothesis remains an open problem a century later. Its difficulty arises from one of the deepest results in modern Set Theory—discovered in part by Gödel in the 1930s and Paul Cohen in the 1960s—namely, the ZFC axioms are not sufficient to settle the Continuum Hypothesis: there are two collections of sets, each obeying the laws of ZFC, and in one collection, the Continuum Hypothesis is true, and in the other, it is false. So settling the Continuum Hypothesis requires a new understanding of what sets should be to arrive at persuasive new axioms that extend ZFC and are strong enough to determine the truth of the Continuum Hypothesis one way or the other.

13.4 Infinities in Computer Science

If the romance of different size infinities and continuum hypotheses doesn't appeal to you, not knowing about them is not going to lower your professional abilities as a computer scientist. These abstract issues about infinite sets rarely come up in mainstream mathematics, and they don't come up at all in computer science, where the focus is generally on countable, and often just finite, sets. In practice, only logicians and set theorists have to worry about collections that are too big to be sets. In fact, at the end of the 19th century, even the general mathematical community doubted the relevance of what they called "Cantor's paradise" of unfamiliar sets of arbitrary infinite size.

That said, it is worth noting that the proof of Theorem 13.3.1 gives the simplest form of what is known as a "diagonal argument." Diagonal arguments are used to prove many fundamental results about the limitations of computation, such as the undecidability of the Halting Problem for programs and the inherent, unavoidable inefficiency (exponential time or worse) of procedures for other computational problems. So computer scientists do need to study diagonal arguments in order to understand the logical limits of computation. Ad a well-educated computer scientist will be comfortable dealing with countable sets, finite as well as infinite.

IV Probability



Introduction

Probability is one of the most important disciplines in all of the sciences. It is also one of the least well understood.

Probability is especially important in computer science—it arises in virtually every branch of the field. In algorithm design and game theory, for example, *randomized* algorithms and strategies (those that use a random number generator as a key input for decision making) frequently outperform deterministic algorithms and strategies. In information theory and signal processing, an understanding of randomness is critical for filtering out noise and compressing data. In cryptography and digital rights management, probability is crucial for achieving security. The list of examples is long.

Given the impact that probability has on computer science, it seems strange that probability should be so misunderstood by so many. Perhaps the trouble is that basic human intuition is wrong as often as it is right when it comes to problems involving random events. As a consequence, many students develop a fear of probability. Indeed, we have witnessed many graduate oral exams where a student will solve the most horrendous calculation, only to then be tripped up by the simplest probability question. Indeed, even some faculty will start squirming if you ask them a question that starts "What is the probability that. . . ?"

Our goal in the remaining chapters is to equip you with the tools that will enable you to easily and confidently solve problems involving probability.

We begin in Chapter 14 with the basic definitions and an elementary 4-step process that can be used to determine the probability that a specified event occurs. We illustrate the method on two famous problems where your intuition will probably fail you.

In Chapter 15, we describe conditional probability and the notion of independence. Both notions are important, and sometimes misused, in practice. We will

consider the probability of having a disease given that you tested positive, and the probability that a suspect is guilty given that his blood type matches the blood found at the scene of the crime.

We study random variables and distributions in Chapter 17. Random variables provide a more quantitative way to measure random events. For example, instead of determining the probability that it will rain, we may want to determine *how much* or *how long* it is likely to rain. This is closely related to the notion of the expected value of a random variables, which we will consider in Chapter 18.

In Chapter 19, we examine the probability that a random variable deviates significantly from its expected value. This is especially important in practice, where things are generally fine if they are going according to expectation, and you would like to be assured that the probability of deviating from the expectation is very low.

We conclude in Chapter 20 by combining the tools we have acquired to solve problems involving more complex random processes. We will see why you will probably never get very far ahead at the casino, and how two Stanford graduate students became gazillionaires by combining graph theory and probability theory to design a better search engine for the web.

14 Events and Probability Spaces

14.1 Let's Make a Deal

In the September 9, 1990 issue of *Parade* magazine, columnist Marilyn vos Savant responded to this letter:

> *Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?*
>
> Craig. F. Whitaker
> Columbia, MD

The letter describes a situation like one faced by contestants in the 1970's game show *Let's Make a Deal*, hosted by Monty Hall and Carol Merrill. Marilyn replied that the contestant should indeed switch. She explained that if the car was behind either of the two unpicked doors—which is twice as likely as the the car being behind the picked door—the contestant wins by switching. But she soon received a torrent of letters, many from mathematicians, telling her that she was wrong. The problem became known as the *Monty Hall Problem* and it generated thousands of hours of heated debate.

This incident highlights a fact about probability: the subject uncovers lots of examples where ordinary intuition leads to completely wrong conclusions. So until you've studied probabilities enough to have refined your intuition, a way to avoid errors is to fall back on a rigorous, systematic approach such as the Four Step Method that we will describe shortly. First, let's make sure we really understand the setup for this problem. This is always a good thing to do when you are dealing with probability.

14.1.1 Clarifying the Problem

Craig's original letter to Marilyn vos Savant is a bit vague, so we must make some assumptions in order to have any hope of modeling the game formally. For example, we will assume that:

1. The car is equally likely to be hidden behind each of the three doors.
2. The player is equally likely to pick each of the three doors, regardless of the car's location.
3. After the player picks a door, the host *must* open a different door with a goat behind it and offer the player the choice of staying with the original door or switching.
4. If the host has a choice of which door to open, then he is equally likely to select each of them.

In making these assumptions, we're reading a lot into Craig Whitaker's letter. Other interpretations are at least as defensible, and some actually lead to different answers. But let's accept these assumptions for now and address the question, "What is the probability that a player who switches wins the car?"

14.2 The Four Step Method

Every probability problem involves some sort of randomized experiment, process, or game. And each such problem involves two distinct challenges:

1. How do we model the situation mathematically?
2. How do we solve the resulting mathematical problem?

In this section, we introduce a four step approach to questions of the form, "What is the probability that...?" In this approach, we build a probabilistic model step-by-step, formalizing the original question in terms of that model. Remarkably, the structured thinking that this approach imposes provides simple solutions to many famously-confusing problems. For example, as you'll see, the four step method cuts through the confusion surrounding the Monty Hall problem like a Ginsu knife.

14.2.1 Step 1: Find the Sample Space

Our first objective is to identify all the possible outcomes of the experiment. A typical experiment involves several randomly-determined quantities. For example, the Monty Hall game involves three such quantities:

1. The door concealing the car.
2. The door initially chosen by the player.

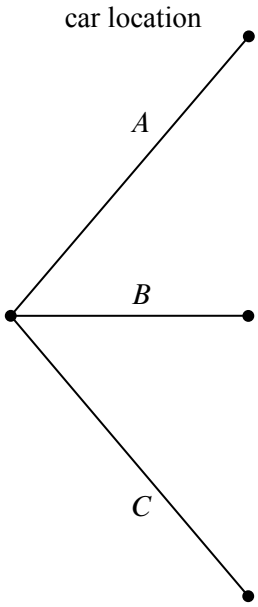


Figure 14.1 The first level in a tree diagram for the Monty Hall Problem. The branches correspond to the door behind which the car is located.

3. The door that the host opens to reveal a goat.

Every possible combination of these randomly-determined quantities is called an *outcome*. The set of all possible outcomes is called the *sample space* for the experiment.

A *tree diagram* is a graphical tool that can help us work through the four step approach when the number of outcomes is not too large or the problem is nicely structured. In particular, we can use a tree diagram to help understand the sample space of an experiment. The first randomly-determined quantity in our experiment is the door concealing the prize. We represent this as a tree with three branches, as shown in Figure 14.1. In this diagram, the doors are called A, B, and C instead of 1, 2, and 3, because we'll be adding a lot of other numbers to the picture later.

For each possible location of the prize, the player could initially choose any of the three doors. We represent this in a second layer added to the tree. Then a third layer represents the possibilities of the final step when the host opens a door to reveal a goat, as shown in Figure 14.2.

Notice that the third layer reflects the fact that the host has either one choice or two, depending on the position of the car and the door initially selected by the player. For example, if the prize is behind door A and the player picks door B, then

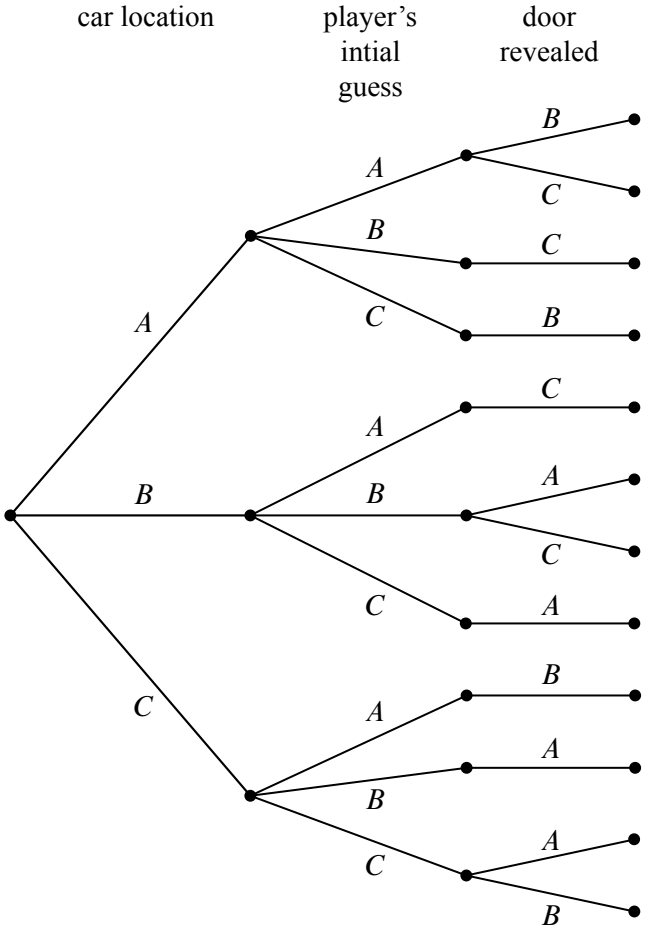


Figure 14.2 The full tree diagram for the Monty Hall Problem. The second level indicates the door initially chosen by the player. The third level indicates the door revealed by Monty Hall.

the host must open door C. However, if the prize is behind door A and the player picks door A, then the host could open either door B or door C.

Now let's relate this picture to the terms we introduced earlier: the leaves of the tree represent *outcomes* of the experiment, and the set of all leaves represents the *sample space*. Thus, for this experiment, the sample space consists of 12 outcomes. For reference, we've labeled each outcome in Figure 14.3 with a triple of doors indicating:

(door concealing prize, door initially chosen, door opened to reveal a goat).

In these terms, the sample space is the set

$$\mathcal{S} = \left\{ \begin{array}{l} (A,A,B),\ (A,A,C),\ (A,B,C),\ (A,C,B),\ (B,A,C),\ (B,B,A), \\ (B,B,C),\ (B,C,A),\ (C,A,B),\ (C,B,A),\ (C,C,A),\ (C,C,B) \end{array} \right\}$$

The tree diagram has a broader interpretation as well: we can regard the whole experiment as following a path from the root to a leaf, where the branch taken at each stage is "randomly" determined. Keep this interpretation in mind; we'll use it again later.

14.2.2 Step 2: Define Events of Interest

Our objective is to answer questions of the form "What is the probability that ... ?", where, for example, the missing phrase might be "the player wins by switching", "the player initially picked the door concealing the prize", or "the prize is behind door C". Each of these phrases characterizes a set of outcomes. For example, the outcomes specified by "the prize is behind door C" is:

$$\{(C,A,B),(C,B,A),(C,C,A),(C,C,B)\}.$$

A set of outcomes is called an *event* and it is a subset of the sample space. So the event that the player initially picked the door concealing the prize is the set:

$$\{(A,A,B),(A,A,C),(B,B,A),(B,B,C),(C,C,A),(C,C,B)\}.$$

And what we're really after, the event that the player wins by switching, is the set of outcomes:

$$\{(A,B,C),(A,C,B),(B,A,C),(B,C,A),(C,A,B),(C,B,A)\}.$$

These outcomes are denoted with a check mark in Figure 14.4.

Notice that exactly half of the outcomes are checked, meaning that the player wins by switching in half of all outcomes. You might be tempted to conclude that a player who switches wins with probability $1/2$. *This is wrong.* The reason is that these outcomes are not all equally likely, as we'll see shortly.

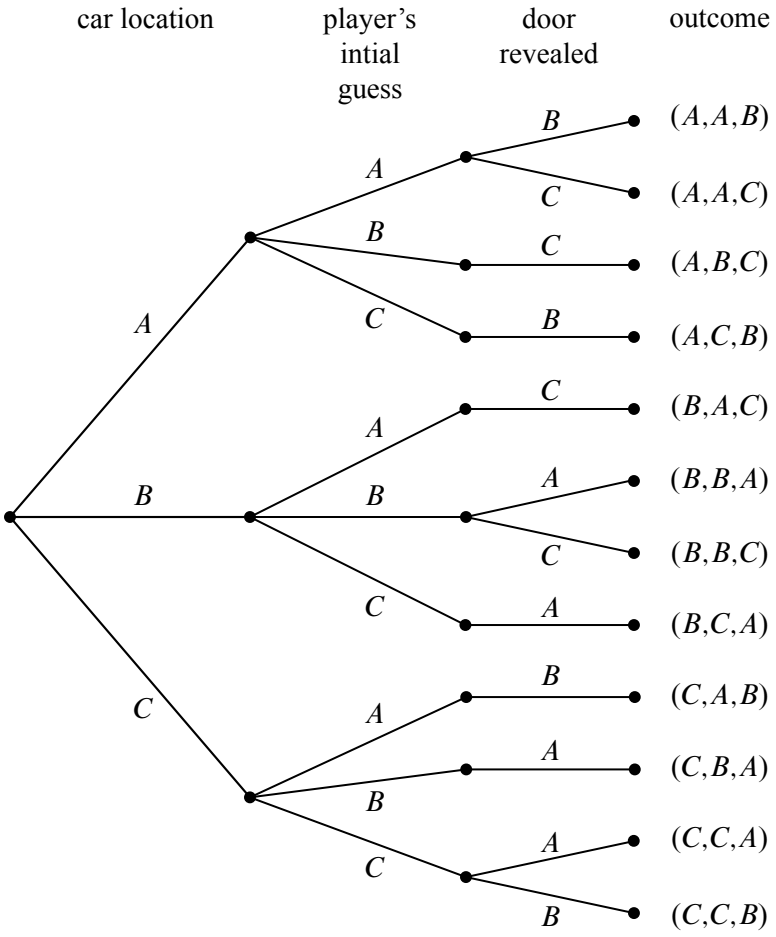


Figure 14.3 The tree diagram for the Monty Hal Problem with the outcomes labeled for each path from root to leaf. For example, outcome (A, A, B) corresponds to the car being behind door A, the player initially choosing door A, and Monty Hall revealing the goat behind door B.

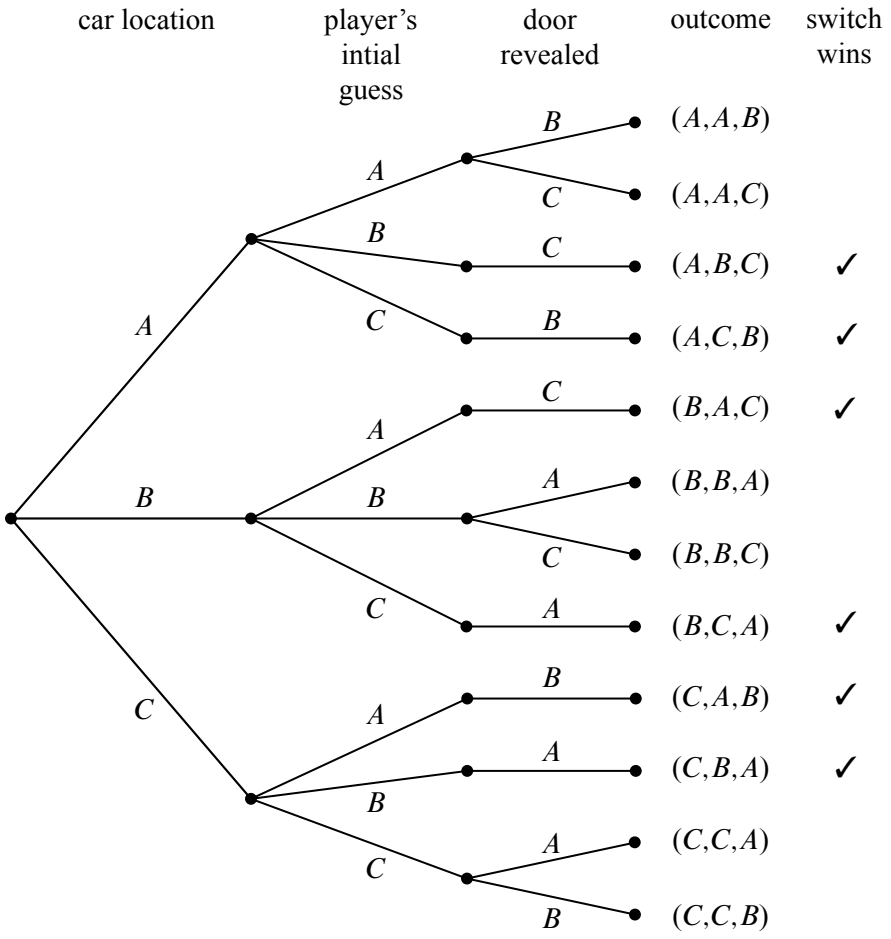


Figure 14.4 The tree diagram for the Monty Hall Problem where the outcomes in the event where the player wins by switching are denoted with a check mark.

14.2.3 Step 3: Determine Outcome Probabilities

So far we've enumerated all the possible outcomes of the experiment. Now we must start assessing the likelihood of those outcomes. In particular, the goal of this step is to assign each outcome a probability, indicating the fraction of the time this outcome is expected to occur. The sum of all outcome probabilities must be one, reflecting the fact that there always is an outcome.

Ultimately, outcome probabilities are determined by the phenomenon we're modeling and thus are not quantities that we can derive mathematically. However, mathematics can help us compute the probability of every outcome *based on fewer and more elementary modeling decisions.* In particular, we'll break the task of determining outcome probabilities into two stages.

Step 3a: Assign Edge Probabilities

First, we record a probability on each *edge* of the tree diagram. These edge-probabilities are determined by the assumptions we made at the outset: that the prize is equally likely to be behind each door, that the player is equally likely to pick each door, and that the host is equally likely to reveal each goat, if he has a choice. Notice that when the host has no choice regarding which door to open, the single branch is assigned probability 1. For example, see Figure 14.5.

Step 3b: Compute Outcome Probabilities

Our next job is to convert edge probabilities into outcome probabilities. This is a purely mechanical process: *the probability of an outcome is equal to the product of the edge-probabilities on the path from the root to that outcome*. For example, the probability of the topmost outcome in Figure 14.5, (A, A, B), is

$$\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{18}.$$

There's an easy, intuitive justification for this rule. As the steps in an experiment progress randomly along a path from the root of the tree to a leaf, the probabilities on the edges indicate how likely the path is to proceed along each branch. For example, a path starting at the root in our example is equally likely to go down each of the three top-level branches.

How likely is such a path to arrive at the topmost outcome, (A, A, B)? Well, there is a 1-in-3 chance that a path would follow the A-branch at the top level, a 1-in-3 chance it would continue along the A-branch at the second level, and 1-in-2 chance it would follow the B-branch at the third level. Thus, it seems that 1 path in 18 should arrive at the (A, A, B) leaf, which is precisely the probability we assign it.

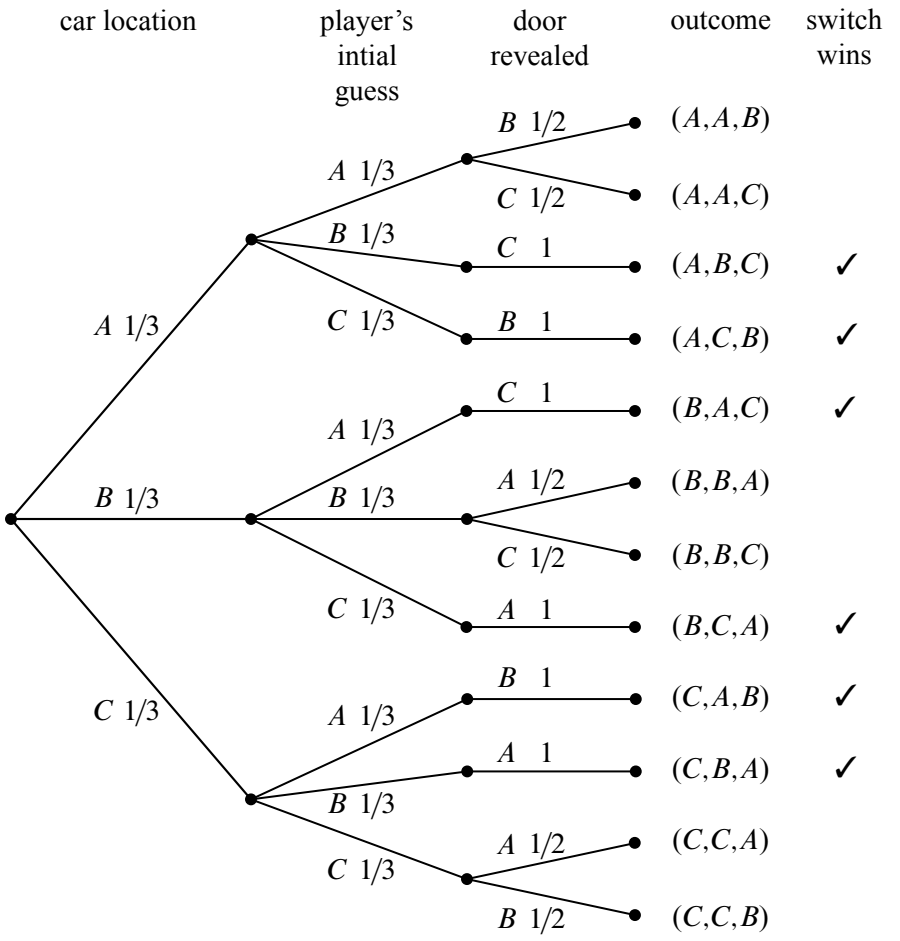


Figure 14.5 The tree diagram for the Monty Hall Problem where edge weights denote the probability of that branch being taken given that we are at the parent of that branch. For example, if the car is behind door A, then there is a 1/3 chance that the player's initial selection is door B.

We have illustrated all of the outcome probabilities in Figure 14.6.

Specifying the probability of each outcome amounts to defining a function that maps each outcome to a probability. This function is usually called **Pr**. In these terms, we've just determined that:

$$\begin{aligned}\Pr[(A,A,B)] &= \frac{1}{18},\\ \Pr[(A,A,C)] &= \frac{1}{18},\\ \Pr[(A,B,C)] &= \frac{1}{9},\\ &\text{etc.}\end{aligned}$$

14.2.4 Step 4: Compute Event Probabilities

We now have a probability for each *outcome*, but we want to determine the probability of an *event*. The probability of an event E is denoted by $\Pr[E]$ and it is the sum of the probabilities of the outcomes in E. For example, the probability of the event that the player wins by switching is:[1]

$$\begin{aligned}\Pr[\text{switching wins}] &= \Pr[(A,B,C)] + \Pr[(A,C,B)] + \Pr[(B,A,C)] +\\ &\qquad \Pr[(B,C,A)] + \Pr[(C,A,B)] + \Pr[(C,B,A)]\\ &= \frac{1}{9}+\frac{1}{9}+\frac{1}{9}+\frac{1}{9}+\frac{1}{9}+\frac{1}{9}\\ &= \frac{2}{3}.\end{aligned}$$

It seems Marilyn's answer is correct! A player who switches doors wins the car with probability $2/3$. In contrast, a player who stays with his or her original door wins with probability $1/3$, since staying wins if and only if switching loses.

We're done with the problem! We didn't need any appeals to intuition or ingenious analogies. In fact, no mathematics more difficult than adding and multiplying fractions was required. The only hard part was resisting the temptation to leap to an "intuitively obvious" answer.

14.2.5 An Alternative Interpretation of the Monty Hall Problem

Was Marilyn really right? Our analysis indicates that she was. But a more accurate conclusion is that her answer is correct *provided we accept her interpretation of the*

[1] "Switching wins" is shorthand for the set of outcomes where switching wins; namely, $\{(A,B,C),\ (A,C,B),\ (B,A,C),\ (B,C,A),\ (C,A,B),\ (C,B,A)\}$. We will frequently use such shorthand to denote events.

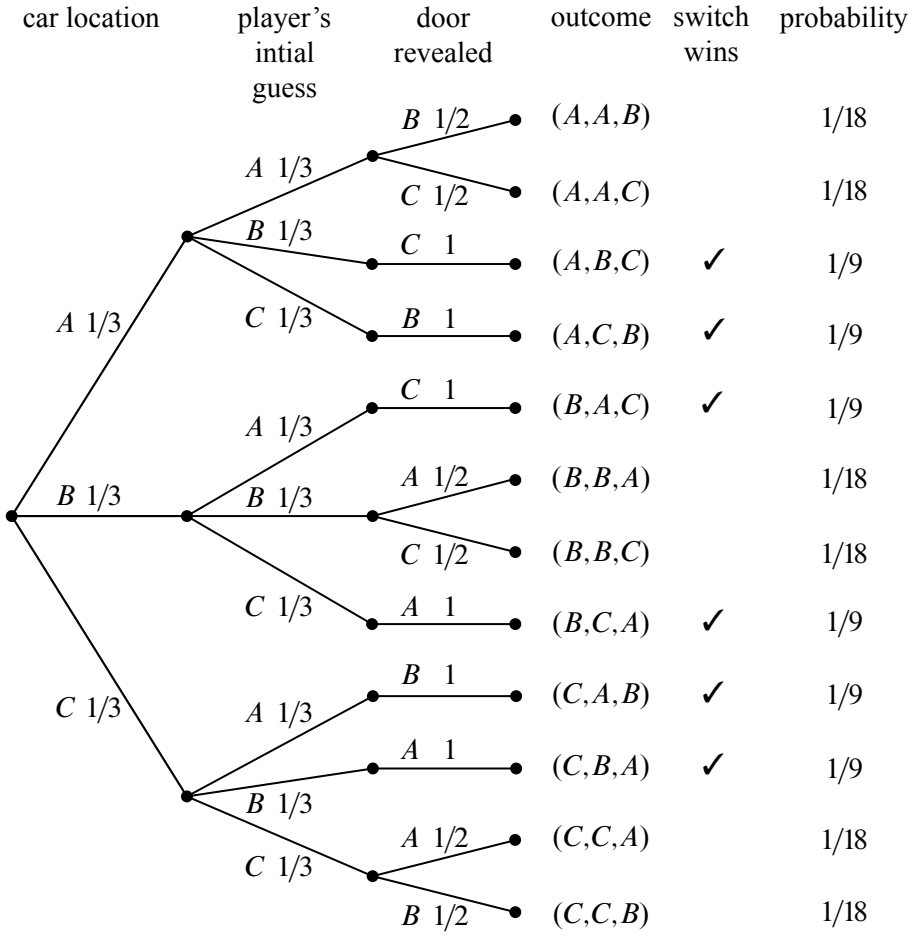
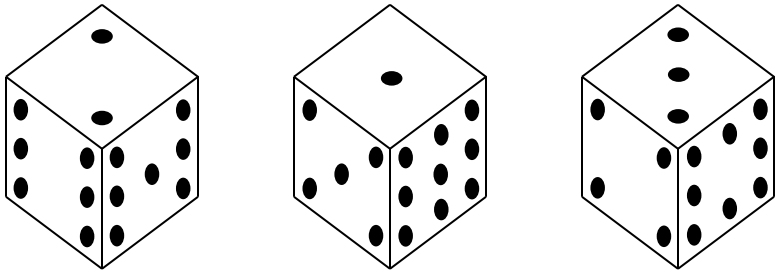


Figure 14.6 The rightmost column shows the outcome probabilities for the Monty Hall Problem. Each outcome probability is simply the product of the probabilities on the branches on the path from the root to the leaf for that outcome.



a *b* *c*

Figure 14.7 The strange dice. The number of pips on each concealed face is the same as the number on the opposite face. For example, when you roll die A, the probabilities of getting a 2, 6, or 7 are each $1/3$.

question. There is an equally plausible interpretation in which Marilyn's answer is wrong. Notice that Craig Whitaker's original letter does not say that the host is *required* to reveal a goat and offer the player the option to switch, merely that he *did* these things. In fact, on the *Let's Make a Deal* show, Monty Hall sometimes simply opened the door that the contestant picked initially. Therefore, if he wanted to, Monty could give the option of switching only to contestants who picked the correct door initially. In this case, switching never works!

14.3 Strange Dice

The four-step method is surprisingly powerful. Let's get some more practice with it. Imagine, if you will, the following scenario.

It's a typical Saturday night. You're at your favorite pub, contemplating the true meaning of infinite cardinalities, when a burly-looking biker plops down on the stool next to you. Just as you are about to get your mind around $\mathcal{P}(\mathcal{P}(\mathbb{R}))$, biker dude slaps three strange-looking dice on the bar and challenges you to a \$100 wager.

The rules are simple. Each player selects one die and rolls it once. The player with the lower value pays the other player \$100.

Naturally, you are skeptical. A quick inspection reveals that these are not ordinary dice. They each have six sides, but the numbers on the dice are different, as shown in Figure 14.7.

Biker dude notices your hesitation and so he offers to let you pick a die first, and

then he will choose his die from the two that are left. That seals the deal since you figure that you now have an advantage.

But which of the dice should you choose? Die B is appealing because it has a 9, which is a sure winner if it comes up. Then again, die A has two fairly large numbers and die B has an 8 and no really small values.

In the end, you choose die B because it has a 9, and then biker dude selects die A. Let's see what the probability is that you will win.[2] Not surprisingly, we will use the four-step method to compute this probability.

14.3.1 Die A versus Die B

Step 1: Find the sample space.
The sample space for this experiment is worked out in the tree diagram shown in Figure 14.8.[3]

For this experiment, the sample space is a set of nine outcomes:

$$\mathcal{S} = \{(2,1),\ (2,5),\ (2,9),\ (6,1),\ (6,5),\ (6,9),\ (7,1),\ (7,5),\ (7,9)\}.$$

Step 2: Define events of interest.
We are interested in the event that the number on die A is greater than the number on die B. This event is a set of five outcomes:

$$\{(2,1),\ (6,1),\ (6,5),\ (7,1),\ (7,5)\}.$$

These outcomes are marked A in the tree diagram in Figure 14.8.

Step 3: Determine outcome probabilities.
To find outcome probabilities, we first assign probabilities to edges in the tree diagram. Each number on each die comes up with probability $1/3$, regardless of the value of the other die. Therefore, we assign all edges probability $1/3$. The probability of an outcome is the product of the probabilities on the corresponding root-to-leaf path, which means that every outcome has probability $1/9$. These probabilities are recorded on the right side of the tree diagram in Figure 14.8.

Step 4: Compute event probabilities.
The probability of an event is the sum of the probabilities of the outcomes in that event. In this case, all the outcome probabilities are the same. In general, when the probability of every outcome is the same, we say that the sample space is *uniform*. Computing event probabilities for uniform sample spaces is particularly easy since

[2] Of course, you probably should have done this before picking die B in the first place.

[3] Actually, the whole probability space is worked out in this one picture. But pretend that each component sort of fades in—nyyrrroom!—as you read about the corresponding step below.

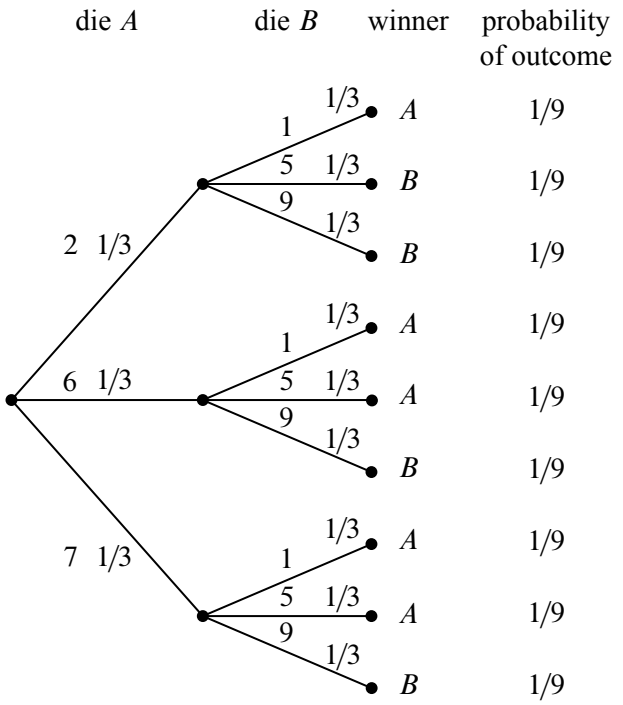


Figure 14.8 The tree diagram for one roll of die A versus die B. Die A wins with probability $5/9$.

you just have to compute the number of outcomes in the event. In particular, for any event E in a uniform sample space $\mathcal{S}$,

$$\Pr[E] = \frac{|E|}{|\mathcal{S}|}. \tag{14.1}$$

In this case, E is the event that die A beats die B, so $|E| = 5$, $|\mathcal{S}| = 9$, and

$$\Pr[E] = 5/9.$$

This is bad news for you. Die A beats die B more than half the time and, not surprisingly, you just lost \$100.

Biker dude consoles you on your "bad luck" and, given that he's a sensitive guy beneath all that leather, he offers to go double or nothing.[4] Given that your wallet only has \$25 in it, this sounds like a good plan. Plus, you figure that choosing die A will give *you* the advantage.

So you choose A, and then biker dude chooses C. Can you guess who is more likely to win? (Hint: it is generally not a good idea to gamble with someone you don't know in a bar, especially when you are gambling with strange dice.)

14.3.2 Die A versus Die C

We can construct the three diagram and outcome probabilities as before. The result is shown in Figure 14.9 and there is bad news again. Die C will beat die A with probability $5/9$, and you lose once again.

You now owe the biker dude \$200 and he asks for his money. You reply that you need to go to the bathroom.

Being a sensitive guy, biker dude nods understandingly and offers yet another wager. This time, he'll let you have die C. He'll even let you raise the wager to \$200 so you can win your money back.

This is too good a deal to pass up. You know that die C is likely to beat die A and that die A is likely to beat die B, and so die C is *surely* the best. Whether biker dude picks A or B, the odds are *surely* in your favor this time. Biker dude must really be a nice guy.

So you pick C, and then biker dude picks B. Let's use the tree method to figure out the probability that you win.

[4] *Double or nothing* is slang for doing another wager after you have lost the first. If you lose again, you will owe biker dude *double* what you owed him before. If you win, you will now be even and you will owe him *nothing*.

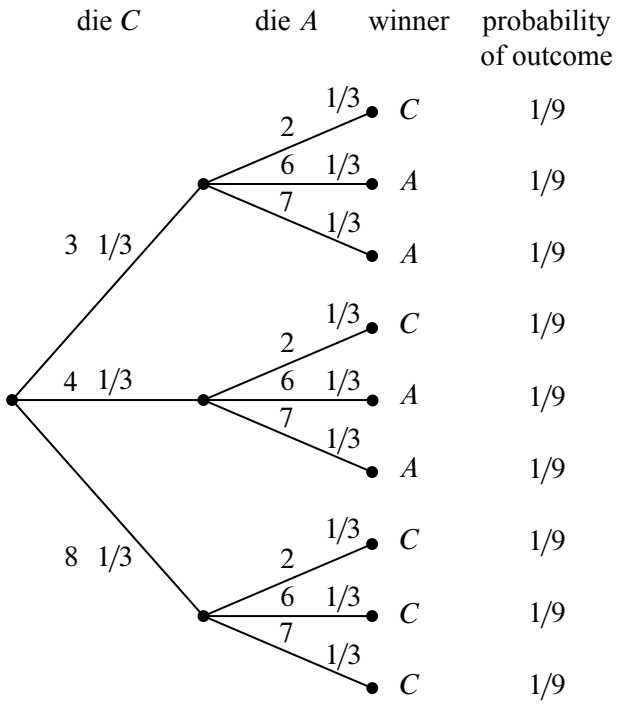


Figure 14.9 The tree diagram for one roll of die C versus die A. Die C wins with probability $5/9$.

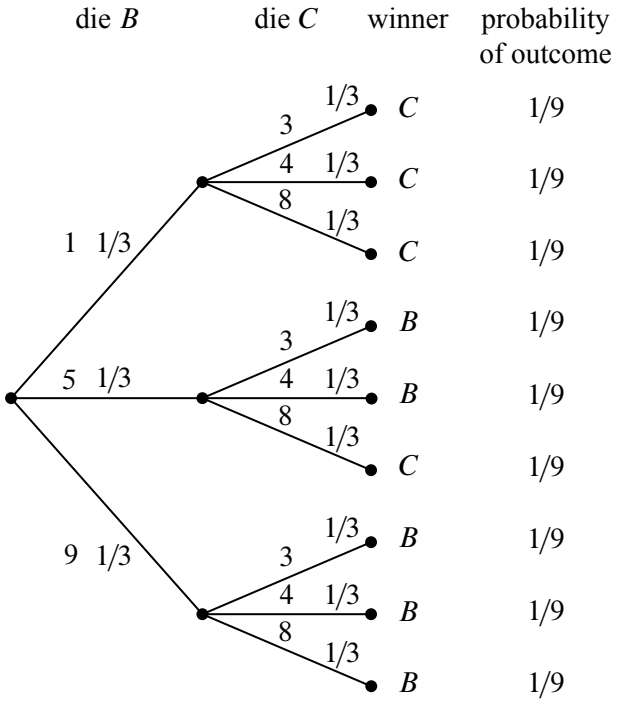


Figure 14.10 The tree diagram for one roll of die B versus die C. Die B wins with probability $5/9$.

14.3.3 Die B versus Die C

The tree diagram and outcome probabilities for B versus C are shown in Figure 14.10. But surely there is a mistake! The data in Figure 14.10 shows that *die* B wins with probability $5/9$. How is it possible that

C beats A with probability $5/9$,

A beats B with probability $5/9$, and

B beats C with probability $5/9$?

The problem is not with the math, but with your intuition. It *seems* that the "likely-to-beat" relation should be transitive. But it is not, and whatever die you pick, biker dude can pick one of the others and be likely to win. So picking first is a big disadvantage and you now owe biker dude \$400.

Just when you think matters can't get worse, biker dude offers you one final wager for \$1,000. This time, you demand to choose second. Biker dude agrees,

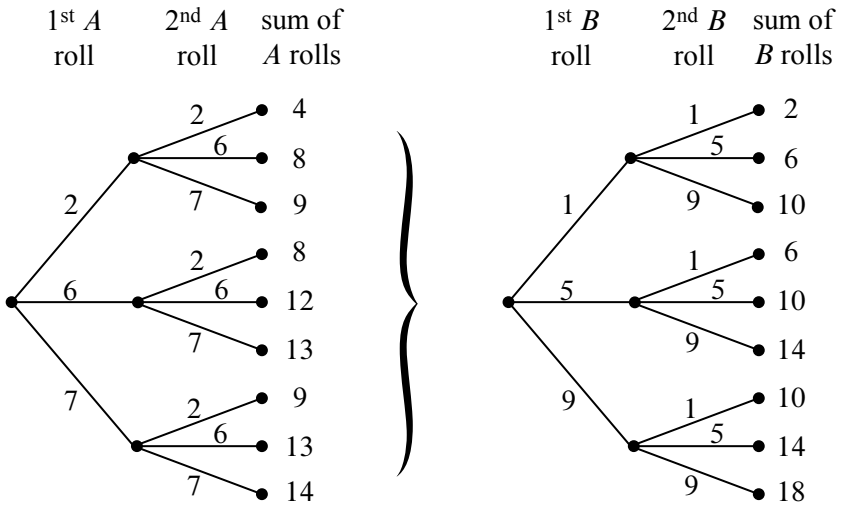


Figure 14.11 Parts of the tree diagram for die B versus die A where each die is rolled twice. The first two levels are shown in (a). The last two levels consist of nine copies of the tree in (b).

but with the condition that instead of rolling each die once, you each roll your die twice and your score is the sum of your rolls.

Believing that you finally have a winning wager, you agree.[5] Biker dude chooses die B and, of course, you grab die A. That's because you know that die A will beat die B with probability $5/9$ on one roll and so *surely* two rolls of die A are likely to beat two rolls of die B, right?

Wrong!

14.3.4 Rolling Twice

If each player rolls twice, the tree diagram will have four levels and $3^4 = 81$ outcomes. This means that it will take a while to write down the entire tree diagram. We can, however, easily write down the first two levels (as we have done in Figure 14.11(a)) and then notice that the remaining two levels consist of nine identical copies of the tree in Figure 14.11(b).

The probability of each outcome is $(1/3)^4 = 1/81$ and so, once again, we have a uniform probability space. By Equation 14.1, this means that the probability that A wins is the number of outcomes where A beats B divided by 81.

To compute the number of outcomes where A beats B, we observe that the sum

[5] Did we mention that playing strange gambling games with strangers in a bar is a bad idea?

of the two rolls of die A is equally likely to be any element of the following multiset:

$$\mathcal{S}_A = \{4, 8, 8, 9, 9, 12, 13, 13, 14\}.$$

The sum of two rolls of die B is equally likely to be any element of the following multiset:

$$\mathcal{S}_B = \{2, 6, 6, 10, 10, 10, 14, 14, 18\}.$$

We can treat each outcome as a pair $(x, y) \in \mathcal{S}_A \times \mathcal{S}_B$, where A wins iff $x > y$. If $x = 4$, there is only one y (namely $y = 2$) for which $x > y$. If $x = 8$, there are three values of y for which $x > y$. Continuing the count in this way, the number of pairs for which $x > y$ is

$$1 + 3 + 3 + 3 + 3 + 6 + 6 + 6 + 6 = 37.$$

A similar count shows that there are 42 pairs for which $x > y$, and there are two pairs ((14, 14), (14, 14)) which result in ties. This means that A *loses* to B with probability $42/81 > 1/2$ and ties with probability $2/81$. Die A wins with probability only $37/81$.

How can it be that A is more likely than B to win with 1 roll, but B is more likely to win with 2 rolls?!? Well, why not? The only reason we'd think otherwise is our (faulty) intuition. In fact, the die strength reverses no matter which two die we picked. So for 1 roll,

$$A \succ B \succ C \succ A,$$

but for two rolls,

$$A \prec B \prec C \prec A,$$

where we have used the symbols $\succ$ and $\prec$ to denote which die is more likely to result in the larger value. This is surprising even to us, but at least we don't owe biker dude $1400.

14.3.5 Even Stranger Dice

Now that we know that strange things can happen with strange dice, it is natural, at least for mathematicians, to ask how strange things can get. It turns out that things can get very strange. In fact, mathematicians[6] recently made the following discovery:

Theorem 14.3.1. *For any $n \geq 2$, there is a set of n dice $D_1, D_2, \ldots, D_n$ such that for any n-node tournament graph[7] G, there is a number of rolls k such that if each*

[6] Reference Ron Graham paper.

[7] Recall that a tournament graph is a directed graph for which there is precisely one directed edge between any two distinct nodes. In other words, for every pair of distinct nodes u and v, either u beats v or v beats u, but not both.

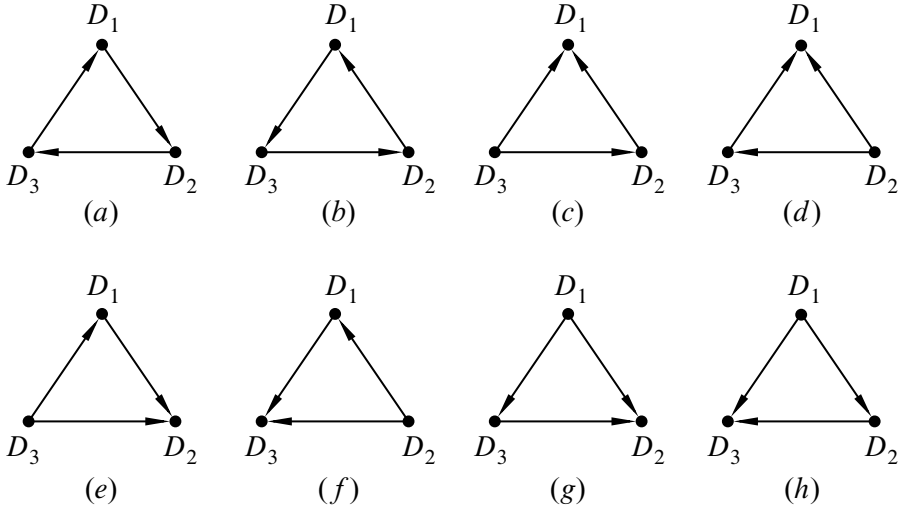


Figure 14.12 All possible relative strengths for three dice D_1, D_2, and D_3. The edge $D_i \to D_j$ denotes that the sum of rolls for D_i is likely to be greater than the sum of rolls for D_j.

die is rolled k times, then for all $i \neq j$, the sum of the k rolls for D_i will exceed the sum for D_j with probability greater than $1/2$ iff $D_i \to D_j$ is in G.

It will probably take a few attempts at reading Theorem 14.3.1 to understand what it is saying. The idea is that for some sets of dice, by rolling them different numbers of times, the dice have varying strengths relative to each other. (This is what we observed for the dice in Figure 14.7.) Theorem 14.3.1 says that there is a set of (very) strange dice where *every* possible collection of relative strengths can be observed by varying the number of rolls. For example, the eight possible relative strengths for $n = 3$ dice are shown in Figure 14.12.

Our analysis for the dice in Figure 14.7 showed that for 1 roll, we have the relative strengths shown in Figure 14.12(a), and for two rolls, we have the (reverse) relative strengths shown in Figure 14.12(b). Can you figure out what other relative strengths are possible for the dice in Figure 14.7 by using more rolls? This might be worth doing if you are prone to gambling with strangers in bars.

14.4 Set Theory and Probability

The study of probability is very closely tied to set theory. That is because any set can be a sample space and any subset can be an event. This means that most of the rules and identities that we have developed for sets extend very naturally to probability. We'll cover several examples in this section, but first let's review some definitions that should already be familiar.

14.4.1 Probability Spaces

Definition 14.4.1. A countable[8] *sample space* $\mathcal{S}$ is a nonempty countable set. An element $w \in \mathcal{S}$ is called an *outcome*. A subset of $\mathcal{S}$ is called an *event*.

Definition 14.4.2. A *probability function* on a sample space $\mathcal{S}$ is a total function $\Pr : \mathcal{S} \to \mathbb{R}$ such that

- $\Pr[w] \geq 0$ for all $w \in \mathcal{S}$, and
- $\sum_{w \in \mathcal{S}} \Pr[w] = 1$.

A sample space together with a probability function is called a *probability space*. For any event $E \subseteq \mathcal{S}$, the *probability of* E is defined to be the sum of the probabilities of the outcomes in E:

$$\Pr[E] ::= \sum_{w \in E} \Pr[w].$$

14.4.2 Probability Rules from Set Theory

An immediate consequence of the definition of event probability is that for *disjoint* events E and F,

$$\Pr[E \cup F] = \Pr[E] + \Pr[F].$$

This generalizes to a countable number of events, as follows.

Rule 14.4.3 (Sum Rule)**.** *If* $\{E_0, E_1, \ldots\}$ *is collection of disjoint events, then*

$$\Pr\left[\bigcup_{n \in \mathbb{N}} E_n\right] = \sum_{n \in \mathbb{N}} \Pr[E_n].$$

[8] Yes, sample spaces can be infinite. We'll see some examples shortly. If you did not read Chapter 13, don't worry—*countable* means that you can list the elements of the sample space as w_1, w_2, $w_3, \ldots$.

The Sum Rule lets us analyze a complicated event by breaking it down into simpler cases. For example, if the probability that a randomly chosen MIT student is native to the United States is 60%, to Canada is 5%, and to Mexico is 5%, then the probability that a random MIT student is native to North America is 70%.

Another consequence of the Sum Rule is that $\Pr[A] + \Pr[\overline{A}] = 1$, which follows because $\Pr[\mathcal{S}] = 1$ and $\mathcal{S}$ is the union of the disjoint sets A and $\overline{A}$. This equation often comes up in the form:

Rule 14.4.4 (Complement Rule)**.**

$$\Pr[\overline{A}] = 1 - \Pr[A].$$

Sometimes the easiest way to compute the probability of an event is to compute the probability of its complement and then apply this formula.

Some further basic facts about probability parallel facts about cardinalities of finite sets. In particular:

$$\begin{aligned}
&\Pr[B - A] = \Pr[B] - \Pr[A \cap B], && \text{(Difference Rule)}\\
&\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B], && \text{(Inclusion-Exclusion)}\\
&\Pr[A \cup B] \leq \Pr[A] + \Pr[B], && \text{(Boole's Inequality}\\
&\text{If } A \subseteq B, \text{ then } \Pr[A] \leq \Pr[B]. && \text{(Monotonicity)}
\end{aligned}$$

The Difference Rule follows from the Sum Rule because B is the union of the disjoint sets $B - A$ and $A \cap B$. Inclusion-Exclusion then follows from the Sum and Difference Rules, because $A \cup B$ is the union of the disjoint sets A and $B - A$. Boole's inequality is an immediate consequence of Inclusion-Exclusion since probabilities are nonnegative. Monotonicity follows from the definition of event probability and the fact that outcome probabilities are nonnegative.

The two-event Inclusion-Exclusion equation above generalizes to n events in the same way as the corresponding Inclusion-Exclusion rule for n sets. Boole's inequality also generalizes to

$$\Pr[E_1 \cup \cdots \cup E_n] \leq \Pr[E_1] + \cdots + \Pr[E_n]. \qquad \text{(Union Bound)}$$

This simple Union Bound is useful in many calculations. For example, suppose that E_i is the event that the i-th critical component in a spacecraft fails. Then $E_1 \cup \cdots \cup E_n$ is the event that *some* critical component fails. If $\sum_{i=1}^{n} \Pr[E_i]$ is small, then the Union Bound can give an adequate upper bound on this vital probability.

14.4.3 Uniform Probability Spaces

Definition 14.4.5. A finite probability space $\mathcal{S}$, Pr is said to be *uniform* if $\Pr[w]$ is the same for every outcome $w \in \mathcal{S}$.

As we saw in the strange dice problem, uniform sample spaces are particularly easy to work with. That's because for any event $E \subseteq \mathcal{S}$,

$$\Pr[E] = \frac{|E|}{|\mathcal{S}|}. \tag{14.2}$$

This means that once we know the cardinality of E and $\mathcal{S}$, we can immediately obtain $\Pr[E]$. That's great news because we developed lots of tools for computing the cardinality of a set in Part III.

For example, suppose that you select five cards at random from a standard deck of 52 cards. What is the probability of having a full house? Normally, this question would take some effort to answer. But from the analysis in Section 11.7.2, we know that

$$|\mathcal{S}| = \binom{13}{5}$$

and

$$|E| = 13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}$$

where E is the event that we have a full house. Since every five-card hand is equally likely, we can apply Equation 14.2 to find that

$$\begin{aligned}
\Pr[E] &= \frac{13 \cdot 12 \cdot \binom{4}{3} \cdot \binom{4}{2}}{\binom{13}{5}} \\
&= \frac{13 \cdot 12 \cdot 4 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48} \\
&= \frac{18}{12495} \\
&\approx \frac{1}{694}.
\end{aligned}$$

14.5 Infinite Probability Spaces

General probability theory deals with uncountable sets like $\mathbb{R}$, but in computer science, it is usually sufficient to restrict our attention to countable probability spaces.

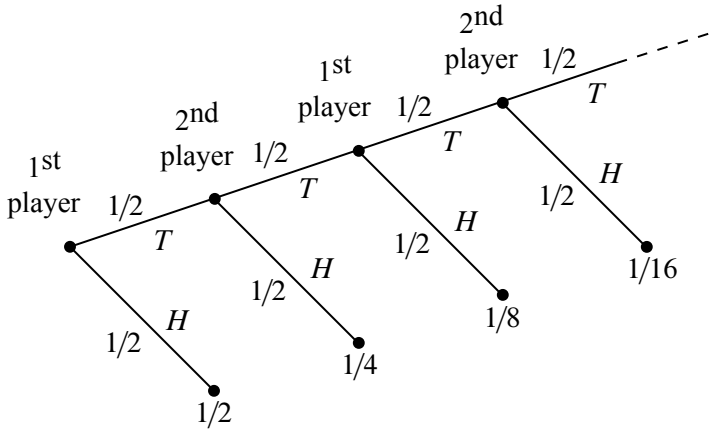


Figure 14.13 The tree diagram for the game where players take turns flipping a fair coin. The first player to flip heads wins.

It's also a lot easier—infinite sample spaces are hard enough to work with without having to deal with uncountable spaces.

Infinite probability spaces are fairly common. For example, two players take turns flipping a fair coin. Whoever flips heads first is declared the winner. What is the probability that the first player wins? A tree diagram for this problem is shown in Figure 14.13.

The event that the first player wins contains an infinite number of outcomes, but we can still sum their probabilities:

$$\begin{aligned}\Pr[\text{first player wins}] &= \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \frac{1}{128} + \cdots \\ &= \frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{1}{4}\right)^n \\ &= \frac{1}{2}\left(\frac{1}{1-1/4}\right) = \frac{2}{3}.\end{aligned}$$

Similarly, we can compute the probability that the second player wins:

$$\Pr[\text{second player wins}] = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \frac{1}{256} + \cdots = \frac{1}{3}.$$

In this case, the sample space is the infinite set

$$\mathcal{S} ::= \{\, \mathrm{T}^n\mathrm{H} \mid n \in \mathbb{N} \,\},$$

where T^n stands for a length n string of T's. The probability function is

$$\Pr[T^n H] ::= \frac{1}{2^{n+1}}.$$

To verify that this is a probability space, we just have to check that all the probabilities are nonnegative and that they sum to 1. Nonnegativity is obvious, and applying the formula for the sum of a geometric series, we find that

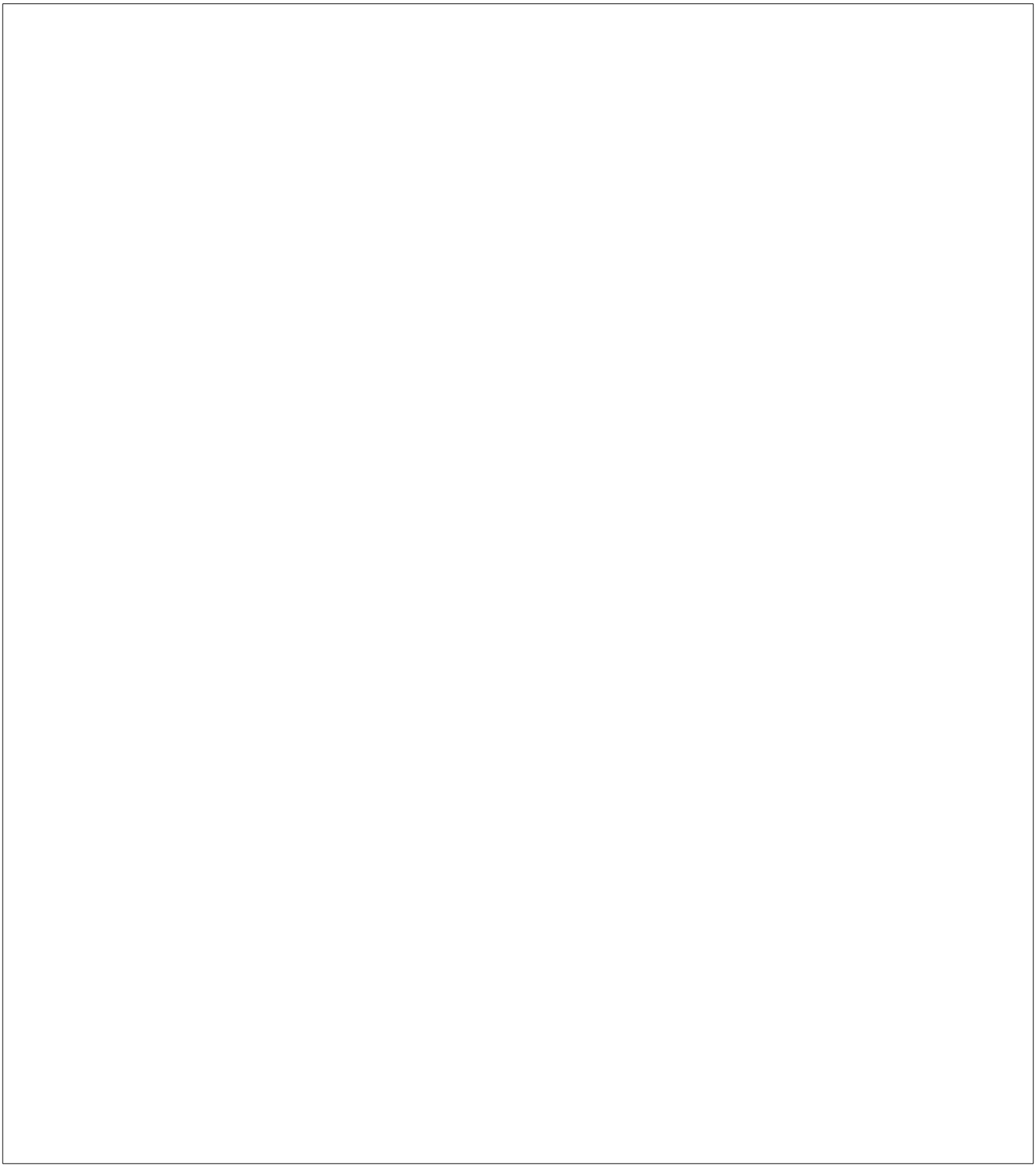
$$\sum_{n\in\mathbb{N}} \Pr[T^n H] = \sum_{n\in\mathbb{N}} \frac{1}{2^{n+1}} = 1.$$

Notice that this model does not have an outcome corresponding to the possibility that both players keep flipping tails forever.[9] That's because the probability of flipping forever would be

$$\lim_{n\to\infty} \frac{1}{2^{n+1}} = 0,$$

and outcomes with probability zero will have no impact on our calculations.

[9] In the diagram, flipping forever corresponds to following the infinite path in the tree without ever reaching a leaf or outcome. Some texts deal with this case by adding a special "infinite" sample point w_{forever} to the sample space, but we will follow the more traditional approach of excluding such sample points, as long as they collectively have probability 0.



15 Conditional Probability

15.1 Definition

Suppose that we pick a random person in the world. Everyone has an equal chance of being selected. Let A be the event that the person is an MIT student, and let B be the event that the person lives in Cambridge. What are the probabilities of these events? Intuitively, we're picking a random point in the big ellipse shown in Figure 15.1 and asking how likely that point is to fall into region A or B.

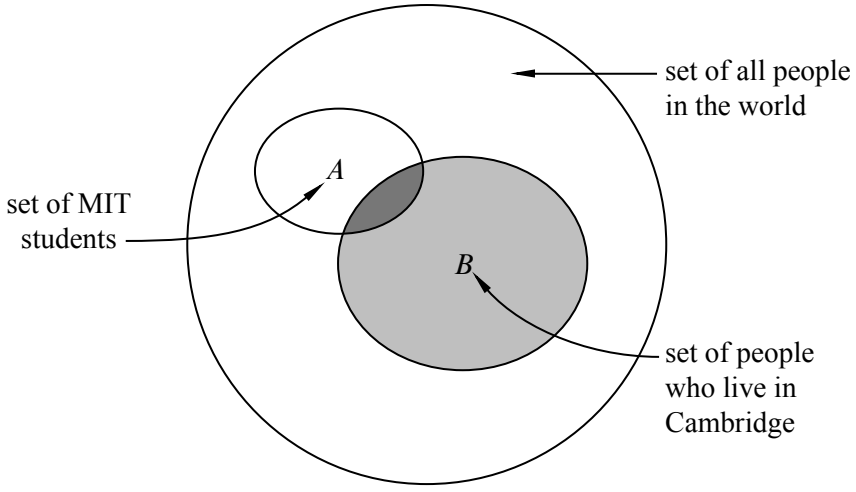


Figure 15.1 Selecting a random person. A is the event that the person is an MIT student. B is the even that the person lives in Cambridge.

The vast majority of people in the world neither live in Cambridge nor are MIT students, so events A and B both have low probability. But what about the probability that a person is an MIT student, *given* that the person lives in Cambridge? This should be much greater—but what is it exactly?

What we're asking for is called a *conditional probability*; that is, the probability that one event happens, given that some other event definitely happens. Questions about conditional probabilities come up all the time:

- What is the probability that it will rain this afternoon, given that it is cloudy this morning?

- What is the probability that two rolled dice sum to 10, given that both are odd?
- What is the probability that I'll get four-of-a-kind in Texas No Limit Hold 'Em Poker, given that I'm initially dealt two queens?

There is a special notation for conditional probabilities. In general, $\Pr[A \mid B]$ denotes the probability of event A, given that event B happens. So, in our example, $\Pr[A \mid B]$ is the probability that a random person is an MIT student, given that he or she is a Cambridge resident.

How do we compute $\Pr[A \mid B]$? Since we are *given* that the person lives in Cambridge, we can forget about everyone in the world who does not. Thus, all outcomes outside event B are irrelevant. So, intuitively, $\Pr[A \mid B]$ should be the fraction of Cambridge residents that are also MIT students; that is, the answer should be the probability that the person is in set $A \cap B$ (the darkly shaded region in Figure 15.1) divided by the probability that the person is in set B (the lightly shaded region). This motivates the definition of conditional probability:

Definition 15.1.1.

$$\Pr[A \mid B] ::= \frac{\Pr[A \cap B]}{\Pr[B]}$$

If $\Pr[B] = 0$, then the conditional probability $\Pr[A \mid B]$ is undefined.

Pure probability is often counterintuitive, but conditional probability is even worse! Conditioning can subtly alter probabilities and produce unexpected results in randomized algorithms and computer systems as well as in betting games. Yet, the mathematical definition of conditional probability given above is very simple and should give you no trouble—provided that you rely on formal reasoning and not intuition. The four-step method will also be very helpful as we will see in the next examples.

15.2 Using the Four-Step Method to Determine Conditional Probability

15.2.1 The "Halting Problem"

The *Halting Problem* was the first example of a property that could not be tested by any program. It was introduced by Alan Turing in his seminal 1936 paper. The problem is to determine whether a Turing machine halts on a given . . . yadda yadda

yadda . . . more importantly, it was the name of the MIT EECS department's famed C-league hockey team.

In a best-of-three tournament, the Halting Problem wins the first game with probability $1/2$. In subsequent games, their probability of winning is determined by the outcome of the previous game. If the Halting Problem won the previous game, then they are invigorated by victory and win the current game with probability $2/3$. If they lost the previous game, then they are demoralized by defeat and win the current game with probability only $1/3$. What is the probability that the Halting Problem wins the tournament, given that they win the first game?

This is a question about a conditional probability. Let A be the event that the Halting Problem wins the tournament, and let B be the event that they win the first game. Our goal is then to determine the conditional probability $\Pr[A \mid B]$.

We can tackle conditional probability questions just like ordinary probability problems: using a tree diagram and the four step method. A complete tree diagram is shown in Figure 15.2.

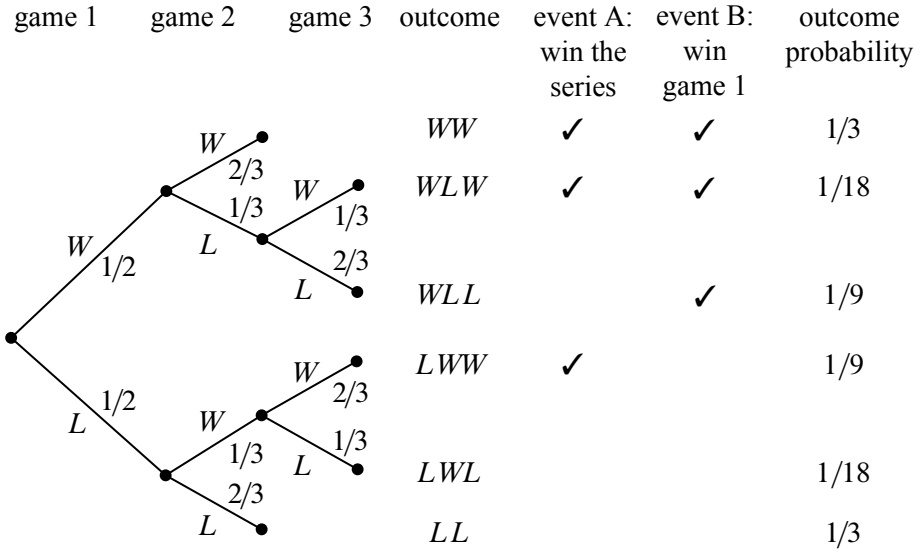


Figure 15.2 The tree diagram for computing the probability that the "Halting Problem" wins two out of three games given that they won the first game.

Step 1: Find the Sample Space

Each internal vertex in the tree diagram has two children, one corresponding to a win for the Halting Problem (labeled W) and one corresponding to a loss (la-

beled L). The complete sample space is:

$$\mathcal{S} = \{WW,\ WLW,\ WLL,\ LWW,\ LWL,\ LL\}.$$

Step 2: Define Events of Interest

The event that the Halting Problem wins the whole tournament is:

$$T = \{WW,\ WLW,\ LWW\}.$$

And the event that the Halting Problem wins the first game is:

$$F = \{WW,\ WLW,\ WLL\}.$$

The outcomes in these events are indicated with check marks in the tree diagram in Figure 15.2.

Step 3: Determine Outcome Probabilities

Next, we must assign a probability to each outcome. We begin by labeling edges as specified in the problem statement. Specifically, The Halting Problem has a $1/2$ chance of winning the first game, so the two edges leaving the root are each assigned probability $1/2$. Other edges are labeled $1/3$ or $2/3$ based on the outcome of the preceding game. We then find the probability of each outcome by multiplying all probabilities along the corresponding root-to-leaf path. For example, the probability of outcome WLL is:

$$\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{2}{3} = \frac{1}{9}.$$

Step 4: Compute Event Probabilities

We can now compute the probability that The Halting Problem wins the tournament, given that they win the first game:

$$\begin{aligned}
\Pr\left[A \mid B\right] &= \frac{\Pr[A \cap B]}{\Pr[B]} \\
&= \frac{\Pr[\{WW, WLW\}]}{\Pr[\{WW, WLW, WLL\}]} \\
&= \frac{1/3 + 1/18}{1/3 + 1/18 + 1/9} \\
&= \frac{7}{9}.
\end{aligned}$$

We're done! If the Halting Problem wins the first game, then they win the whole tournament with probability $7/9$.

15.2.2 Why Tree Diagrams Work

We've now settled into a routine of solving probability problems using tree diagrams. But we've left a big question unaddressed: what is the mathematical justification behind those funny little pictures? Why do they work?

The answer involves conditional probabilities. In fact, the probabilities that we've been recording on the edges of tree diagrams *are* conditional probabilities. For example, consider the uppermost path in the tree diagram for the Halting Problem, which corresponds to the outcome WW. The first edge is labeled $1/2$, which is the probability that the Halting Problem wins the first game. The second edge is labeled $2/3$, which is the probability that the Halting Problem wins the second game, *given* that they won the first—that's a conditional probability! More generally, on each edge of a tree diagram, we record the probability that the experiment proceeds along that path, given that it reaches the parent vertex.

So we've been using conditional probabilities all along. But why can we multiply edge probabilities to get outcome probabilities? For example, we concluded that:

$$\Pr[WW] = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

Why is this correct?

The answer goes back to Definition 15.1.1 of conditional probability which could be written in a form called the *Product Rule* for probabilities:

Rule (Product Rule for 2 Events)**.** *If* $\Pr[E_1] \neq 0$*, then:*

$$\Pr[E_1 \cap E_2] = \Pr[E_1] \cdot \Pr\left[E_2 \mid E_1\right].$$

Multiplying edge probabilities in a tree diagram amounts to evaluating the right side of this equation. For example:

$$\begin{aligned}
&\Pr[\text{win first game} \cap \text{win second game}] \\
&\quad = \Pr[\text{win first game}] \cdot \Pr\left[\text{win second game} \mid \text{win first game}\right] \\
&\quad = \frac{1}{2} \cdot \frac{2}{3}.
\end{aligned}$$

So the Product Rule is the formal justification for multiplying edge probabilities to get outcome probabilities! Of course to justify multiplying edge probabilities along longer paths, we need a Product Rule for n events.

Rule (Product Rule for n Events)**.**

$$\begin{aligned}
\Pr[E_1 \cap E_2 \cap \ldots \cap E_n] = \Pr[E_1] \cdot \Pr\left[E_2 \mid E_1\right] \cdot \Pr\left[E_3 \mid E_1 \cap E_2\right] \cdots \\
\cdot \Pr\left[E_n \mid E_1 \cap E_2 \cap \ldots \cap E_{n-1}\right]
\end{aligned}$$

provided that

$$\Pr[E_1 \cap E_2 \cap \cdots \cap E_{n-1}] \neq 0.$$

This rule follows from the definition of conditional probability and induction on n.

15.2.3 Medical Testing

There is an unpleasant condition called *BO* suffered by 10% of the population. There are no prior symptoms; victims just suddenly start to stink. Fortunately, there is a test for latent *BO* before things start to smell. The test is not perfect, however:

- If you have the condition, there is a 10% chance that the test will say you do not. These are called "false negatives".
- If you do not have the condition, there is a 30% chance that the test will say you do. These are "false positives".

Suppose a random person is tested for latent *BO*. If the test is positive, then what is the probability that the person has the condition?

Step 1: Find the Sample Space

The sample space is found with the tree diagram in Figure 15.3.

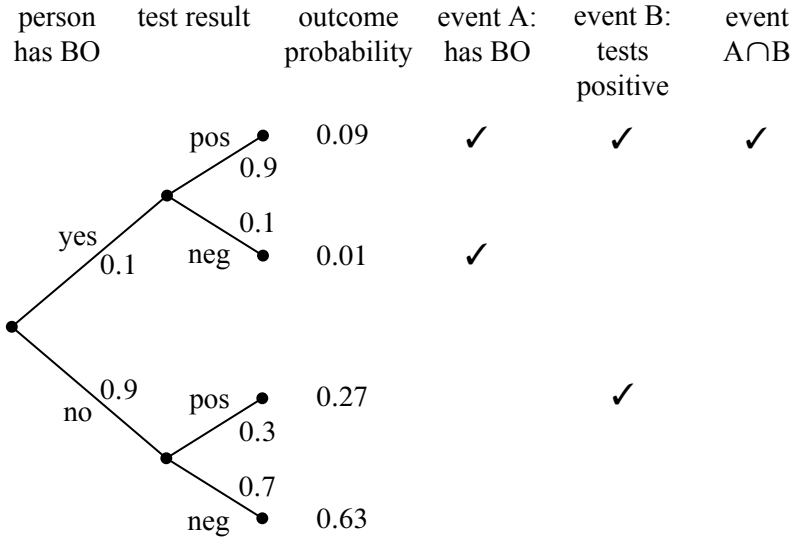


Figure 15.3 The tree diagram for the BO problem.

Step 2: Define Events of Interest

Let A be the event that the person has BO. Let B be the event that the test was positive. The outcomes in each event are marked in the tree diagram. We want to find $\Pr[A \mid B]$, the probability that a person has BO, given that the test was positive.

Step 3: Find Outcome Probabilities

First, we assign probabilities to edges. These probabilities are drawn directly from the problem statement. By the Product Rule, the probability of an outcome is the product of the probabilities on the corresponding root-to-leaf path. All probabilities are shown in Figure 15.3.

Step 4: Compute Event Probabilities

From Definition 15.1.1, we have

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{0.09}{0.09 + 0.27} = \frac{1}{4}.$$

So, if you test positive, then there is only a 25% chance that you have the condition!

This answer is initially surprising, but makes sense on reflection. There are two ways you could test positive. First, it could be that you have the condition and the test is correct. Second, it could be that you are healthy and the test is incorrect. The problem is that almost everyone is healthy; therefore, most of the positive results arise from incorrect tests of healthy people!

We can also compute the probability that the test is correct for a random person. This event consists of two outcomes. The person could have the condition and test positive (probability 0.09), or the person could be healthy and test negative (probability 0.63). Therefore, the test is correct with probability $0.09 + 0.63 = 0.72$. This is a relief; the test is correct almost three-quarters of the time.

But wait! There is a simple way to make the test correct 90% of the time: always return a negative result! This "test" gives the right answer for all healthy people and the wrong answer only for the 10% that actually have the condition. So a better strategy by this measure is to completely ignore the test result!

There is a similar paradox in weather forecasting. During winter, almost all days in Boston are wet and overcast. Predicting miserable weather every day may be more accurate than really trying to get it right!

15.3 *A Posteriori* Probabilities

If you think about it too much, the medical testing problem we just considered could start to trouble you. The concern would be that by the time you take the test, you either have the BO condition or you don't—you just don't know which it is. So you may wonder if a statement like "If you tested positive, then you have the condition with probability 25%" makes sense.

In fact, such a statement does make sense. It means that 25% of the people who test positive actually have the condition. It is true that any particular person has it or they don't, but a *randomly selected* person among those who test positive will have the condition with probability 25%.

Anyway, if the medical testing example bothers you, you will definitely be worried by the following examples, which go even further down this path.

15.3.1 The "Halting Problem," in Reverse

Suppose that we turn the hockey question around: what is the probability that the Halting Problem won their first game, given that they won the series?

This seems like an absurd question! After all, if the Halting Problem won the series, then the winner of the first game has already been determined. Therefore, who won the first game is a question of fact, not a question of probability. However, our mathematical theory of probability contains no notion of one event preceding another—there is no notion of time at all. Therefore, from a mathematical perspective, this is a perfectly valid question. And this is also a meaningful question from a practical perspective. Suppose that you're told that the Halting Problem won the series, but not told the results of individual games. Then, from your perspective, it makes perfect sense to wonder how likely it is that The Halting Problem won the first game.

A conditional probability $\Pr[B \mid A]$ is called *a posteriori* if event B precedes event A in time. Here are some other examples of a posteriori probabilities:

- The probability it was cloudy this morning, given that it rained in the afternoon.
- The probability that I was initially dealt two queens in Texas No Limit Hold 'Em poker, given that I eventually got four-of-a-kind.

Mathematically, a posteriori probabilities are *no different* from ordinary probabilities; the distinction is only at a higher, philosophical level. Our only reason for drawing attention to them is to say, "Don't let them rattle you."

Let's return to the original problem. The probability that the Halting Problem won their first game, given that they won the series is $\Pr[B \mid A]$. We can compute this using the definition of conditional probability and the tree diagram in Figure 15.2:

$$\Pr[B \mid A] = \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{1/3 + 1/18}{1/3 + 1/18 + 1/9} = \frac{7}{9}.$$

This answer is suspicious! In the preceding section, we showed that $\Pr[A \mid B]$ was also $7/9$. Could it be true that $\Pr[A \mid B] = \Pr[B \mid A]$ in general? Some reflection suggests this is unlikely. For example, the probability that I feel uneasy, given that I was abducted by aliens, is pretty large. But the probability that I was abducted by aliens, given that I feel uneasy, is rather small.

Let's work out the general conditions under which $\Pr[A \mid B] = \Pr[B \mid A]$. By the definition of conditional probability, this equation holds if an only if:

$$\frac{\Pr[A \cap B]}{\Pr[B]} = \frac{\Pr[A \cap B]}{\Pr[A]}$$

This equation, in turn, holds only if the denominators are equal or the numerator is 0; namely if

$$\Pr[B] = \Pr[A] \qquad \text{or} \qquad \Pr[A \cap B] = 0.$$

The former condition holds in the hockey example; the probability that the Halting Problem wins the series (event A) is equal to the probability that it wins the first game (event B) since both probabilities are $1/2$.

In general, such pairs of probabilities are related by Bayes' Rule:

Theorem 15.3.1 (Bayes' Rule)**.** *If* $\Pr[A]$ *and* $\Pr[B]$ *are nonzero, then:*

$$\Pr[B \mid A] = \frac{\Pr[A \mid B] \cdot \Pr[B]}{\Pr[A]} \tag{15.1}$$

Proof. When $\Pr[A]$ and $\Pr[B]$ are nonzero, we have

$$\Pr[A \mid B] \cdot \Pr[B] = \Pr[A \cap B] = \Pr[B \mid A] \cdot \Pr[A]$$

by definition of conditional probability. Dividing by $\Pr[A]$ gives (15.1). ∎

Next, let's look at a problem that even bothers us.

15.3.2 A Coin Problem

Suppose that someone hands you either a fair coin or a trick coin with heads on both sides. You flip the coin 100 times and see heads every time. What can you say about the probability that you flipped the fair coin? Remarkably, nothing!

In order to make sense out of this outrageous claim, let's formalize the problem. The sample space is worked out in the tree diagram shown in Figure 15.4. We do not know the probability p that you were handed the fair coin initially—you were just given one coin or the other. Let A be the event that you were handed the

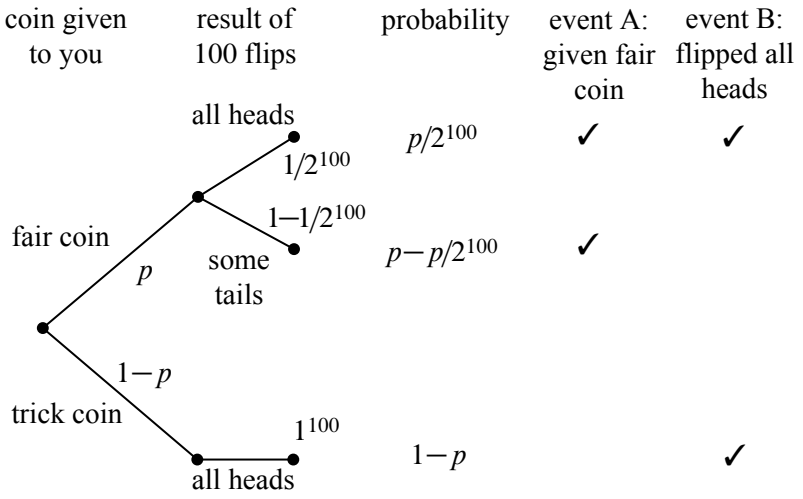


Figure 15.4 The tree diagram for the coin-flipping problem.

fair coin, and let B be the event that you flipped 100 straight heads. We're looking for $\Pr[A \mid B]$, the probability that you were handed the fair coin, given that you flipped 100 heads. The outcome probabilities are worked out in Figure 15.4. Plugging the results into the definition of conditional probability gives:

$$\begin{aligned}\Pr[A \mid B] &= \frac{\Pr[A \cap B]}{\Pr[B]} \\ &= \frac{p/2^{100}}{1-p+p/2^{100}} \\ &= \frac{p}{2^{100}(1-p)+p}.\end{aligned}$$

This expression is very small for moderate values of p because of the 2^{100} term in the denominator. For example, if $p = 1/2$, then the probability that you were given the fair coin is essentially zero.

But we *do not know* the probability p that you were given the fair coin. And perhaps the value of p is *not* moderate; in fact, maybe $p = 1 - 2^{-100}$. Then there is nearly an even chance that you have the fair coin, given that you flipped 100 heads. In fact, maybe you were handed the fair coin with probability $p = 1$. Then the probability that you were given the fair coin is, well, 1!

Of course, it is extremely unlikely that you would flip 100 straight heads, but in this case, that is a given from the assumption of the conditional probability. And so if you really did see 100 straight heads, it would be very tempting to also assume that p is not close to 1 and hence that you are very likely to have flipped the trick coin.

We will encounter a very similar issue when we look at methods for estimation by sampling in Section 17.5.5.

15.4 Conditional Identities

15.4.1 The Law of Total Probability

Breaking a probability calculation into cases simplifies many problems. The idea is to calculate the probability of an event A by splitting into two cases based on whether or not another event E occurs. That is, calculate the probability of $A \cap E$ and $A \cap \overline{E}$. By the Sum Rule, the sum of these probabilities equals $\Pr[A]$. Expressing the intersection probabilities as conditional probabilities yields:

Rule 15.4.1 (Law of Total Probability, single event)**.** *If* $\Pr[E]$ *and* $\Pr[\overline{E}]$ *are nonzero, then*

$$\Pr[A] = \Pr\left[A \mid E\right] \cdot \Pr[E] + \Pr\left[A \mid \overline{E}\right] \cdot \Pr[\overline{E}].$$

For example, suppose we conduct the following experiment. First, we flip a fair coin. If heads comes up, then we roll one die and take the result. If tails comes up, then we roll two dice and take the sum of the two results. What is the probability that this process yields a 2? Let E be the event that the coin comes up heads, and let A be the event that we get a 2 overall. Assuming that the coin is fair, $\Pr[E] = \Pr[\overline{E}] = 1/2$. There are now two cases. If we flip heads, then we roll a 2 on a single die with probability $\Pr\left[A \mid E\right] = 1/6$. On the other hand, if we flip tails, then we get a sum of 2 on two dice with probability $\Pr\left[A \mid \overline{E}\right] = 1/36$. Therefore, the probability that the whole process yields a 2 is

$$\Pr[A] = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{36} = \frac{7}{72}.$$

There is also a form of the rule to handle more than two cases.

Rule 15.4.2 (Law of Total Probability)**.** *If $E_1, \ldots, E_n$ are disjoint events whose union is the whole sample space, then:*

$$\Pr[A] = \sum_{i=1}^{n} \Pr\left[A \mid E_i\right] \cdot \Pr[E_i].$$

15.4.2 Conditioning on a Single Event

The probability rules that we derived in Chapter 14 extend to probabilities conditioned on the same event. For example, the Inclusion-Exclusion formula for two sets holds when all probabilities are conditioned on an event C:

$$\Pr\left[A \cup B \mid C\right] = \Pr\left[A \mid C\right] + \Pr\left[B \mid C\right] - \Pr\left[A \cap B \mid C\right].$$

This follows from the fact that if $\Pr[C] \neq 0$, then

$$\begin{aligned}
\Pr\left[A \cup B \mid C\right] &= \frac{\Pr[(A \cup B) \cap C]}{\Pr[C]} \\
&= \frac{\Pr[(A \cap C) \cup (B \cap C)]}{\Pr[C]} \\
&= \frac{\Pr[A \cap C] + \Pr[B \cap C] - \Pr[A \cap B \cap C]}{\Pr[C]} \\
&= \Pr\left[A \mid C\right] + \Pr\left[B \mid C\right] - \Pr\left[A \cap B \mid C\right].
\end{aligned}$$

It is important not to mix up events before and after the conditioning bar. For example, the following is *not* a valid identity:

False Claim.

$$\Pr\left[A \mid B \cup C\right] = \Pr\left[A \mid B\right] + \Pr\left[A \mid C\right] - \Pr\left[A \mid B \cap C\right]. \tag{15.2}$$

A counterexample is shown in Figure 15.5. In this case, $\Pr\left[A \mid B\right] = 1/2$, $\Pr\left[A \mid C\right] = 1/2$, $\Pr\left[A \mid B \cap C\right] = 1$, and $\Pr\left[A \mid B \cup C\right] = 1/3$. However, since $1/3 \neq 1/2 + 1/2 - 1$, Equation 15.2 does not hold.

So you're convinced that this equation is false in general, right? Let's see if you *really* believe that.

15.4.3 Discrimination Lawsuit

Several years ago there was a sex discrimination lawsuit against a famous university. A female math professor was denied tenure, allegedly because she was

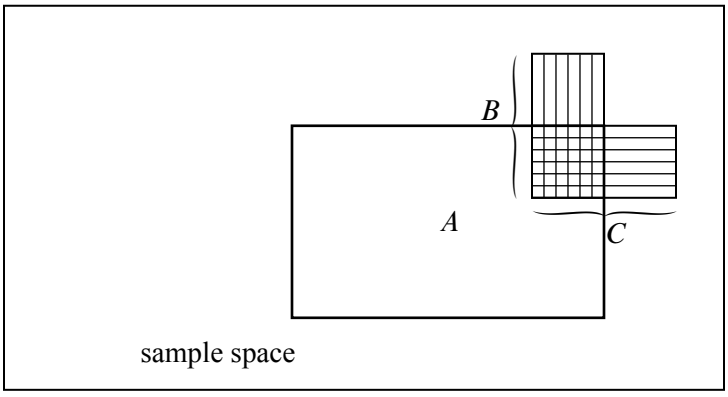


Figure 15.5 A counterexample to Equation 15.2. Event A is the gray rectangle, event B is the rectangle with vertical stripes, and event C is the rectangle with horizontal stripes. $B \cap C$ lies entirely within A while $B - C$ and $C - B$ are entirely outside of A.

a woman. She argued that in every one of the university's 22 departments, the percentage of male applicants accepted was greater than the percentage of female applicants accepted. This sounds very suspicious!

However, the university's lawyers argued that across the university as a whole, the percentage of male applicants accepted was actually *lower* than the percentage of female applicants accepted. This suggests that if there was any sex discrimination, then it was against men! Surely, at least one party in the dispute must be lying.

Let's simplify the problem and express both arguments in terms of conditional probabilities. To simplify matters, suppose that there are only two departments, EE and CS, and consider the experiment where we pick a random applicant. Define the following events:

- Let A be the event that the applicant is accepted.
- Let F_{EE} the event that the applicant is a female applying to EE.
- Let F_{CS} the event that the applicant is a female applying to CS.
- Let M_{EE} the event that the applicant is a male applying to EE.
- Let M_{CS} the event that the applicant is a male applying to CS.

Assume that all applicants are either male or female, and that no applicant applied to both departments. That is, the events F_{EE}, F_{CS}, M_{EE}, and M_{CS} are all disjoint.

CS	0 females accepted, 1 applied	0%
	50 males accepted, 100 applied	50%
EE	70 females accepted, 100 applied	70%
	1 male accepted, 1 applied	100%
Overall	70 females accepted, 101 applied	$\approx 70\%$
	51 males accepted, 101 applied	$\approx 51\%$

Table 15.1 A scenario where females are less likely to be admitted than males in each department, but more likely to be admitted overall.

In these terms, the plaintiff is making the following argument:

$$\Pr[A \mid F_{EE}] < \Pr[A \mid M_{EE}] \quad \text{and}$$
$$\Pr[A \mid F_{CS}] < \Pr[A \mid M_{CS}].$$

That is, in both departments, the probability that a woman is accepted for tenure is less than the probability that a man is accepted. The university retorts that overall, a woman applicant is *more* likely to be accepted than a man; namely that

$$\Pr[A \mid F_{EE} \cup F_{CS}] > \Pr[A \mid M_{EE} \cup M_{CS}].$$

It is easy to believe that these two positions are contradictory. In fact, we might even try to prove this by adding the plaintiff's two inequalities and then arguing as follows:

$$\begin{aligned} & \Pr[A \mid F_{EE}] + \Pr[A \mid F_{CS}] < \Pr[A \mid M_{EE}] + \Pr[A \mid M_{CS}] \\ \Rightarrow \quad & \Pr[A \mid F_{EE} \cup F_{CS}] < \Pr[A \mid M_{EE} \cup M_{CS}]. \end{aligned}$$

The second line exactly contradicts the university's position! But there is a big problem with this argument; the second inequality follows from the first only if we accept the false identity (15.2). This argument is bogus! Maybe the two parties do not hold contradictory positions after all!

In fact, Table 15.1 shows a set of application statistics for which the assertions of both the plaintiff and the university hold. In this case, a higher percentage of males were accepted in both departments, but overall a higher percentage of females were accepted! Bizarre!

16 Independence

16.1 Definitions

Suppose that we flip two fair coins simultaneously on opposite sides of a room. Intuitively, the way one coin lands does not affect the way the other coin lands. The mathematical concept that captures this intuition is called *independence*:

Definition 16.1.1. Events A and B are independent if $\Pr[B] = 0$ or if

$$\Pr\left[A \mid B\right] = \Pr[A]. \tag{16.1}$$

In other words, A and B are independent if knowing that B happens does not alter the probability that A happens, as is the case with flipping two coins on opposite sides of a room.

16.1.1 Potential Pitfall

Students sometimes get the idea that disjoint events are independent. The *opposite* is true: if $A \cap B = \emptyset$, then knowing that A happens means you know that B does not happen. So disjoint events are *never* independent—unless one of them has probability zero.

16.1.2 Alternative Formulation

Sometimes it is useful to express independence in an alternate form:

Theorem 16.1.2. *A and B are independent if and only if*

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B]. \tag{16.2}$$

Proof. There are two cases to consider depending on whether or not $\Pr[B] = 0$.

Case 1 ($\Pr[B] = 0$)**:** If $\Pr[B] = 0$, A and B are independent by Definition 16.1.1. In addition, Equation 16.2 holds since both sides are 0. Hence, the theorem is true in this case.

Case 2 ($\Pr[B] > 0$)**:** By Definition 15.1.1,

$$\Pr[A \cap B] = \Pr\left[A \mid B\right] \Pr[B].$$

So Equation 16.2 holds if

$$\Pr\left[A \mid B\right] = \Pr[A],$$

which, by Definition 16.1.1, is true iff A and B are independent. Hence, the theorem is true in this case as well. ■

16.2 Independence Is an Assumption

Generally, independence is something that you *assume* in modeling a phenomenon. For example, consider the experiment of flipping two fair coins. Let A be the event that the first coin comes up heads, and let B be the event that the second coin is heads. If we assume that A and B are independent, then the probability that both coins come up heads is:

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

In this example, the assumption of independence is reasonable. The result of one coin toss should have negligible impact on the outcome of the other coin toss. And if we were to repeat the experiment many times, we would be likely to have $A \cap B$ about 1/4 of the time.

There are, of course, many examples of events where assuming independence is *not* justified, For example, let C be the event that tomorrow is cloudy and R be the event that tomorrow is rainy. Perhaps $\Pr[C] = 1/5$ and $\Pr[R] = 1/10$ in Boston. If these events were independent, then we could conclude that the probability of a rainy, cloudy day was quite small:

$$\Pr[R \cap C] = \Pr[R] \cdot \Pr[C] = \frac{1}{5} \cdot \frac{1}{10} = \frac{1}{50}.$$

Unfortunately, these events are definitely not independent; in particular, every rainy day is cloudy. Thus, the probability of a rainy, cloudy day is actually $1/10$.

Deciding when to *assume* that events are independent is a tricky business. In practice, there are strong motivations to assume independence since many useful formulas (such as Equation 16.2) only hold if the events are independent. But you need to be careful lest you end up deriving false conclusions. We'll see several famous examples where (false) assumptions of independence led to trouble over the next several chapters. This problem gets even trickier when there are more than two events in play.

16.3 Mutual Independence

16.3.1 Definition

We have defined what it means for two events to be independent. What if there are more than two events? For example, how can we say that the flips of n coins are all independent of one another?

Events $E_1, \ldots, E_n$ are said to be *mutually independent* if and only if the probability of any event E_i is unaffected by knowledge of the other events. More formally:

Definition 16.3.1. A set of events $E_1, E_2, \ldots, E_n$, is *mutually independent* if $\forall i \in [1,n]$ and $\forall S \subseteq [1,n] - \{i\}$, either

$$\Pr\left[\bigcap_{j\in S} E_j\right] = 0 \quad \text{or} \quad \Pr[E_i] = \Pr\left[E_i \,\middle|\, \bigcap_{j\in S} E_j\right].$$

In other words, no matter which other events are known to occur, the probability that E_i occurs is unchanged for any i.

For example, if we toss 100 fair coins at different times, we might reasonably assume that the tosses are mutually independent since the probability that the ith coin is heads should be $1/2$, no matter which other coin tosses came out heads.

16.3.2 Alternative Formulation

Just as Theorem 16.1.2 provided an alternative definition of independence for two events, there is an alternative definition for mutual independence.

Theorem 16.3.2. *A set of events $E_1, E_2, \ldots, E_n$ is mutually independent iff $\forall S \subseteq [1,n]$,*

$$\Pr\left[\bigcap_{j\in S} E_j\right] = \prod_{j\in S} \Pr[E_j].$$

The proof of Theorem 16.3.2 uses induction and reasoning similar to the proof of Theorem 16.1.2. We will not include the details here.

Theorem 16.3.2 says that $E_1, E_2, \ldots, E_n$ are mutually independent if and only

if all of the following equations hold for all distinct i, j, k, and l:

$$\begin{aligned}
\Pr[E_i \cap E_j] &= \Pr[E_i] \cdot \Pr[E_j] \\
\Pr[E_i \cap E_j \cap E_k] &= \Pr[E_i] \cdot \Pr[E_j] \cdot \Pr[E_k] \\
\Pr[E_i \cap E_j \cap E_k \cap E_l] &= \Pr[E_i] \cdot \Pr[E_j] \cdot \Pr[E_k] \cdot \Pr[E_l] \\
&\vdots \\
\Pr[E_1 \cap \cdots \cap E_n] &= \Pr[E_1] \cdots \Pr[E_n].
\end{aligned}$$

For example, if we toss n fair coins, the tosses are mutually independent iff for all $m \in [1, n]$ and every subset of m coins, the probability that every coin in the subset comes up heads is 2^{-m}.

16.3.3 DNA Testing

Assumptions about independence are routinely made in practice. Frequently, such assumptions are quite reasonable. Sometimes, however, the reasonableness of an independence assumption is not so clear, and the consequences of a faulty assumption can be severe.

For example, consider the following testimony from the O. J. Simpson murder trial on May 15, 1995:

Mr. Clarke: When you make these estimations of frequency—and I believe you touched a little bit on a concept called independence?

Dr. Cotton: Yes, I did.

Mr. Clarke: And what is that again?

Dr. Cotton: It means whether or not you inherit one allele that you have is not—does not affect the second allele that you might get. That is, if you inherit a band at 5,000 base pairs, that doesn't mean you'll automatically or with some probability inherit one at 6,000. What you inherit from one parent is what you inherit from the other.

Mr. Clarke: Why is that important?

Dr. Cotton: Mathematically that's important because if that were not the case, it would be improper to multiply the frequencies between the different genetic locations.

Mr. Clarke: How do you—well, first of all, are these markers independent that you've described in your testing in this case?

Presumably, this dialogue was as confusing to you as it was for the jury. Essentially, the jury was told that genetic markers in blood found at the crime scene matched Simpson's. Furthermore, they were told that the probability that the markers would be found in a randomly-selected person was at most 1 in 170 million. This astronomical figure was derived from statistics such as:

- 1 person in 100 has marker A.
- 1 person in 50 marker B.
- 1 person in 40 has marker C.
- 1 person in 5 has marker D.
- 1 person in 170 has marker E.

Then these numbers were multiplied to give the probability that a randomly-selected person would have all five markers:

$$\begin{aligned}\Pr[A \cap B \cap C \cap D \cap E] &= \Pr[A] \cdot \Pr[B] \cdot \Pr[C] \cdot \Pr[D] \cdot \Pr[E] \\ &= \frac{1}{100} \cdot \frac{1}{50} \cdot \frac{1}{40} \cdot \frac{1}{5} \cdot \frac{1}{170} \\ &= \frac{1}{170{,}000{,}000}.\end{aligned}$$

The defense pointed out that this assumes that the markers appear mutually independently. Furthermore, all the statistics were based on just a few hundred blood samples.

After the trial, the jury was widely mocked for failing to "understand" the DNA evidence. If you were a juror, would *you* accept the 1 in 170 million calculation?

16.4 Pairwise Independence

The definition of mutual independence seems awfully complicated—there are so many subsets of events to consider! Here's an example that illustrates the subtlety of independence when more than two events are involved. Suppose that we flip three fair, mutually-independent coins. Define the following events:

- A_1 is the event that coin 1 matches coin 2.
- A_2 is the event that coin 2 matches coin 3.

- A_3 is the event that coin 3 matches coin 1.

Are A_1, A_2, A_3 mutually independent?

The sample space for this experiment is:

$$\{HHH,\ HHT,\ HTH,\ HTT,\ THH,\ THT,\ TTH,\ TTT\}.$$

Every outcome has probability $(1/2)^3 = 1/8$ by our assumption that the coins are mutually independent.

To see if events A_1, A_2, and A_3 are mutually independent, we must check a sequence of equalities. It will be helpful first to compute the probability of each event A_i:

$$\begin{aligned}\Pr[A_1] &= \Pr[HHH] + \Pr[HHT] + \Pr[TTH] + \Pr[TTT] \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} \\ &= \frac{1}{2}.\end{aligned}$$

By symmetry, $\Pr[A_2] = \Pr[A_3] = 1/2$ as well. Now we can begin checking all the equalities required for mutual independence in Theorem 16.3.2:

$$\begin{aligned}\Pr[A_1 \cap A_2] &= \Pr[HHH] + \Pr[TTT] \\ &= \frac{1}{8} + \frac{1}{8} \\ &= \frac{1}{4} \\ &= \frac{1}{2} \cdot \frac{1}{2} \\ &= \Pr[A_1]\Pr[A_2].\end{aligned}$$

By symmetry, $\Pr[A_1 \cap A_3] = \Pr[A_1] \cdot \Pr[A_3]$ and $\Pr[A_2 \cap A_3] = \Pr[A_2] \cdot \Pr[A_3]$ must hold also. Finally, we must check one last condition:

$$\begin{aligned}\Pr[A_1 \cap A_2 \cap A_3] &= \Pr[HHH] + \Pr[TTT] \\ &= \frac{1}{8} + \frac{1}{8} \\ &= \frac{1}{4} \\ &\neq \Pr[A_1]\Pr[A_2]\Pr[A_3] = \frac{1}{8}.\end{aligned}$$

The three events A_1, A_2, and A_3 are not mutually independent even though any two of them are independent! This not-quite mutual independence seems weird at first, but it happens. It even generalizes:

Definition 16.4.1. A set $A_1, A_2, \ldots,$ of events is *k-way independent* iff every set of k of these events is mutually independent. The set is *pairwise independent* iff it is 2-way independent.

So the sets A_1, A_2, A_3 above are pairwise independent, but not mutually independent. Pairwise independence is a much weaker property than mutual independence.

For example, suppose that the prosecutors in the O. J. Simpson trial were wrong and markers A, B, C, D, and E appear only *pairwise* independently. Then the probability that a randomly-selected person has all five markers is no more than:

$$\begin{aligned}\Pr[A \cap B \cap C \cap D \cap E] &\leq \Pr[A \cap E] \\ &= \Pr[A] \cdot \Pr[E] \\ &= \frac{1}{100} \cdot \frac{1}{170} \\ &= \frac{1}{17{,}000}.\end{aligned}$$

The first line uses the fact that $A \cap B \cap C \cap D \cap E$ is a subset of $A \cap E$. (We picked out the A and E markers because they're the rarest.) We use pairwise independence on the second line. Now the probability of a random match is 1 in 17,000—a far cry from 1 in 170 million! And this is the strongest conclusion we can reach assuming only pairwise independence.

On the other hand, the 1 in 17,000 bound that we get by assuming pairwise independence is a lot better than the bound that we would have if there were no independence at all. For example, if the markers are dependent, then it is possible that

everyone with marker E has marker A,

everyone with marker A has marker B,

everyone with marker B has marker C, and

everyone with marker C has marker D.

In such a scenario, the probability of a match is

$$\Pr[E] = 1/170.$$

So a stronger independence assumption leads to a smaller bound on the probability of a match. The trick is to figure out what independence assumption is reasonable. Assuming that the markers are *mutually* independent may well *not* be reasonable unless you have examined hundreds of millions of blood samples. Otherwise, how would you know that marker D does not show up more frequently whenever the other four markers are simultaneously present?

We will conclude our discussion of independence with a useful, and somewhat famous, example known as the Birthday Paradox.

16.5 The Birthday Paradox

Suppose that there are 100 students in a class. What is the probability that some birthday is shared by two people? Comparing 100 students to the 365 possible birthdays, you might guess the probability lies somewhere around $1/3$—but you'd be wrong: the probability that there will be two people in the class with matching birthdays is actually $0.999999692\ldots$. In other words, the probability that all 100 birthdays are different is less than 1 in 3,000,000.

Why is this probability so small? The answer involves a phenomenon known as the *Birthday Paradox* (or the *Birthday Principle*), which is surprisingly important in computer science, as we'll see later.

Before delving into the analysis, we'll need to make some modeling assumptions:

- For each student, all possible birthdays are equally likely. The idea underlying this assumption is that each student's birthday is determined by a random process involving parents, fate, and, um, some issues that we discussed earlier in the context of graph theory. The assumption is not completely accurate, however; a disproportionate number of babies are born in August and September, for example.

- Birthdays are mutually independent. This isn't perfectly accurate either. For example, if there are twins in the class, then their birthdays are surely not independent.

We'll stick with these assumptions, despite their limitations. Part of the reason is to simplify the analysis. But the bigger reason is that our conclusions will apply to many situations in computer science where twins, leap days, and romantic holidays are not considerations. After all, whether or not two items collide in a hash table really has nothing to do with human reproductive preferences. Also, in pursuit of

generality, let's switch from specific numbers to variables. Let m be the number of people in the room, and let N be the number of days in a year.

We can solve this problem using the standard four-step method. However, a tree diagram will be of little value because the sample space is so enormous. This time we'll have to proceed without the visual aid!

Step 1: Find the Sample Space

Let's number the people in the room from 1 to m. An outcome of the experiment is a sequence $(b_1, \ldots, b_m)$ where b_i is the birthday of the ith person. The sample space is the set of all such sequences:

$$\mathcal{S} = \{\, (b_1, \ldots, b_m) \mid b_i \in \{1, \ldots N\}\,\}.$$

Step 2: Define Events of Interest

Our goal is to determine the probability of the event A in which some pair of people have the same birthday. This event is a little awkward to study directly, however. So we'll use a common trick, which is to analyze the *complementary* event $\overline{A}$, in which all m people have different birthdays:

$$\overline{A} = \{\, (b_1, \ldots, b_m) \in \mathcal{S} \mid \text{all } b_i \text{ are distinct}\,\}.$$

If we can compute $\Pr[\overline{A}]$, then we can compute what really want, $\Pr[A]$, using the identity

$$\Pr[A] + \Pr[\overline{A}] = 1.$$

Step 3: Assign Outcome Probabilities

We need to compute the probability that m people have a particular combination of birthdays $(b_1, \ldots, b_m)$. There are N possible birthdays and all of them are equally likely for each student. Therefore, the probability that the ith person was born on day b_i is $1/N$. Since we're assuming that birthdays are mutually independent, we can multiply probabilities. Therefore, the probability that the first person was born on day b_1, the second on b_2, and so forth is $(1/N)^m$. This is the probability of every outcome in the sample space, which means that the sample space is uniform. That's good news, because, as we have seen, it means that the analysis will be simpler.

Step 4: Compute Event Probabilities

We're interested in the probability of the event $\overline{A}$ in which everyone has a different birthday:

$$\overline{A} = \{\, (b_1, \ldots, b_n) \mid \text{all } b_i \text{ are distinct}\,\}.$$

This is a gigantic set. In fact, there are N choices for b_i, $N-1$ choices for b_2, and so forth. Therefore, by the Generalized Product Rule,

$$|\overline{A}| = \frac{N!}{(N-m)!} = N(N-1)(N-2)\cdots(N-m+1).$$

Since the sample space is uniform, we can conclude that

$$\Pr[\overline{A}] = \frac{|\overline{A}|}{N^m} = \frac{N!}{N^m(N-m)!}. \tag{16.3}$$

We're done!

Or are we? While correct, it would certainly be nicer to have a closed-form expression for Equation 16.3. That means finding an approximation for $N!$ and $(N-m)!$. But this is what we learned how to do in Section 9.6. In fact, since N and $N-m$ are each at least 100, we know from Corollary 9.6.2 that

$$\sqrt{2\pi N}\left(\frac{N}{e}\right)^N \quad \text{and} \quad \sqrt{2\pi(N-m)}\left(\frac{N-m}{e}\right)^{N-m}$$

are excellent approximations (accurate to within .09%) of $N!$ and $(N-m)!$, respectively. Plugging these values into Equation 16.3 means that (to within .2%)[1]

$$\begin{aligned}
\Pr[\overline{A}] &= \frac{\sqrt{2\pi N}\left(\frac{N}{e}\right)^N}{N^m\sqrt{2\pi(N-m)}\left(\frac{N-m}{e}\right)^{N-m}} \\
&= \sqrt{\frac{N}{N-m}}\,\frac{e^{N\ln(N)-N}}{e^{m\ln(N)}e^{(N-m)\ln(N-m)-(N-m)}} \\
&= \sqrt{\frac{N}{N-m}}\,e^{(N-m)\ln(N)-(N-m)\ln(N-m)-m} \\
&= \sqrt{\frac{N}{N-m}}\,e^{(N-m)\ln\left(\frac{N}{N-m}\right)-m} \\
&= e^{\left(N-m+\frac{1}{2}\right)\ln\left(\frac{N}{N-m}\right)-m}.
\end{aligned} \tag{16.4}$$

[1] If there are two terms that can be off by .09%, then the ratio can be off by at most a factor of $(1.0009)^2 < 1.002$.

We can now evaluate Equation 16.4 for $m = 100$ and $N = 365$ to find that the probability that all 100 birthdays are different is[2]

$$3.07\ldots \cdot 10^{-7}.$$

We can also plug in other values of m to find the number of people so that the probability of a matching birthday will be about $1/2$. In particular, for $m = 23$ and $N = 365$, Equation 16.4 reveals that the probability that all the birthdays differ is $0.49\ldots$. So if you are in a room with 23 other people, the probability that some pair of people share a birthday will be a little better than $1/2$. It is because 23 seems like such a small number of people for a match that the phenomenon is called the *Birthday Paradox*.

16.5.1 Applications to Hashing

Hashing is frequently used in computer science to map large strings of data into short strings of data. In a typical scenario, you have a set of m items and you would like to assign each item to a number from 1 to N where no pair of items is assigned to the same number and N is as small as possible. For example, the items might be messages, addresses, or variables. The numbers might represent storage locations, devices, indices, or digital signatures.

If two items are assigned to the same number, then a *collision* is said to occur. Collisions are generally bad. For example, collisions can correspond to two variables being stored in the same place or two messages being assigned the same digital signature. Just imagine if you were doing electronic banking and your digital signature for a \$10 check were the same as your signature for a \$10 million dollar check. In fact, finding collisions is a common technique in breaking cryptographic codes.[3]

In practice, the assignment of a number to an item is done using a hash function

$$h : S \to [1, N],$$

where S is the set of items and $m = |S|$. Typically, the values of $h(S)$ are assigned randomly and are assumed to be equally likely in $[1, N]$ and mutually independent.

For efficiency purposes, it is generally desirable to make N as small as necessary to accommodate the hashing of m items without collisions. Ideally, N would be only a little larger than m. Unfortunately, this is not possible for random hash functions. To see why, let's take a closer look at Equation 16.4.

[2] The possible .2% error is so small that it is lost in the ... after 3.07.

[3] Such techniques are often referred to as *birthday attacks* because of the association of such attacks with the Birthday Paradox.

By Theorem 9.6.1 and the derivation of Equation 16.4, we know that the probability that there are no collisions for a random hash function is

$$\sim e^{(N-m+\frac{1}{2})\ln(\frac{N}{N-m})-m}. \tag{16.5}$$

For any m, we now need to find a value of N for which this expression is at least 1/2. That will tell us how big the hash table needs to be in order to have at least a 50% chance of avoiding collisions. This means that we need to find a value of N for which

$$\left(N-m+\frac{1}{2}\right)\ln\left(\frac{N}{N-m}\right)-m \sim \ln\left(\frac{1}{2}\right). \tag{16.6}$$

To simplify Equation 16.6, we need to get rid of the $\ln\left(\frac{N}{N-m}\right)$ term. We can do this by using the Taylor Series expansion for

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots$$

to find that[4]

$$\begin{aligned}
\ln\left(\frac{N}{N-m}\right) &= -\ln\left(\frac{N-m}{N}\right) \\
&= -\ln\left(1-\frac{m}{N}\right) \\
&= -\left(-\frac{m}{N} - \frac{m^2}{2N^2} - \frac{m^3}{3N^3} - \cdots\right) \\
&= \frac{m}{N} + \frac{m^2}{2N^2} + \frac{m^3}{3N^3} + \cdots.
\end{aligned}$$

[4]This may not look like a simplification, but stick with us here.

Hence,

$$\begin{aligned}\left(N-m+\frac{1}{2}\right)\ln\left(\frac{N}{N-m}\right)-m &= \left(N-m+\frac{1}{2}\right)\left(\frac{m}{N}+\frac{m^2}{2N^2}+\frac{m^3}{3N^3}+\cdots\right)-m\\ &= \left(m+\frac{m^2}{2N}+\frac{m^3}{3N^2}+\cdots\right)\\ &\quad -\left(\frac{m^2}{N}+\frac{m^3}{2N^2}+\frac{m^4}{3N^3}+\cdots\right)\\ &\quad +\frac{1}{2}\left(\frac{m}{N}+\frac{m^2}{2N^2}+\frac{m^3}{3N^3}+\cdots\right)-m\\ &= -\left(\frac{m^2}{2N}+\frac{m^3}{6N^2}+\frac{m^4}{12N^3}+\cdots\right)\\ &\quad +\frac{1}{2}\left(\frac{m}{N}+\frac{m^2}{2N^2}+\frac{m^3}{3N^3}+\cdots\right).\end{aligned} \tag{16.7}$$

If N grows faster than m^2, then the value in Equation 16.7 tends to 0 and Equation 16.6 cannot be satisfied. If N grows more slowly than m^2, then the value in Equation 16.7 diverges to negative infinity, and, once again, Equation 16.6 cannot be satisfied. This suggests that we should focus on the case where $N = \Theta(m^2)$, when Equation 16.7 simplifies to

$$\sim \frac{-m^2}{2N}$$

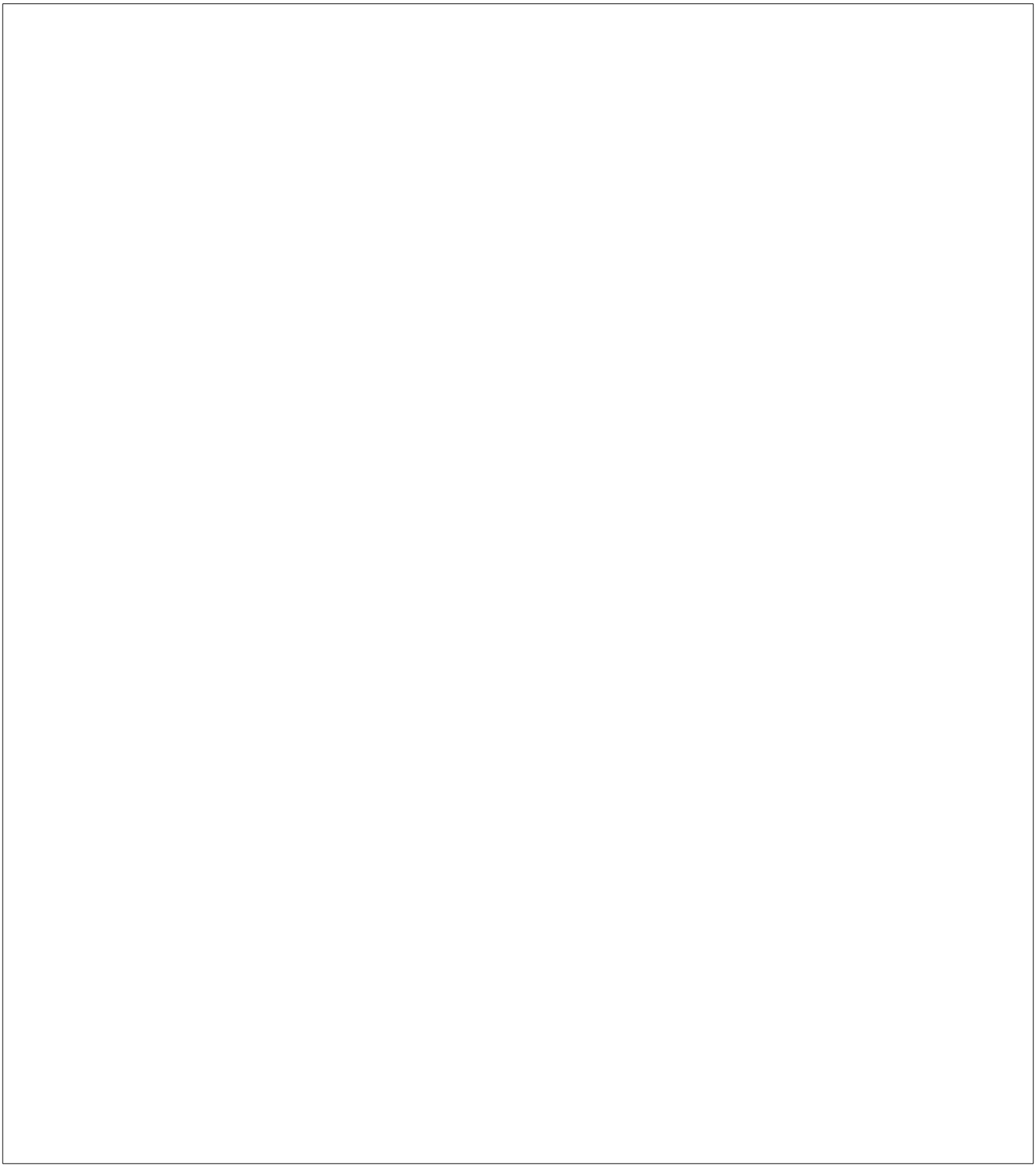
and Equation 16.6 becomes

$$\frac{-m^2}{2N} \sim \ln\left(\frac{1}{2}\right). \tag{16.8}$$

Equation 16.8 is satisfied when

$$N \sim \frac{m^2}{2\ln(2)}. \tag{16.9}$$

In other words, N needs to grow quadratically with m in order to avoid collisions. This unfortunate fact is known as the *Birthday Principle* and it limits the efficiency of hashing in practice—either N is quadratic in the number of items being hashed or you need to be able to deal with collisions.



17 Random Variables and Distributions

Thus far, we have focused on probabilities of events. For example, we computed the probability that you win the Monty Hall game, or that you have a rare medical condition given that you tested positive. But, in many cases we would like to more more. For example, *how many* contestants must play the Monty Hall game until one of them finally wins? *How long* will this condition last? *How much* will I lose gambling with strange dice all night? To answer such questions, we need to work with random variables.

17.1 Definitions and Examples

Definition 17.1.1. A *random variable* R on a probability space is a total function whose domain is the sample space.

The codomain of R can be anything, but will usually be a subset of the real numbers. Notice that the name "random variable" is a misnomer; random variables are actually functions!

For example, suppose we toss three independent[1], unbiased coins. Let C be the number of heads that appear. Let $M = 1$ if the three coins come up all heads or all tails, and let $M = 0$ otherwise. Every outcome of the three coin flips uniquely determines the values of C and M. For example, if we flip heads, tails, heads, then $C = 2$ and $M = 0$. If we flip tails, tails, tails, then $C = 0$ and $M = 1$. In effect, C counts the number of heads, and M indicates whether all the coins match.

Since each outcome uniquely determines C and M, we can regard them as functions mapping outcomes to numbers. For this experiment, the sample space is

$$\mathcal{S} = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

and C is a function that maps each outcome in the sample space to a number as

[1] Going forward, when we talk about flipping independent coins, we will assume that they are *mutually* independent.

follows:

$$\begin{aligned} C(HHH) &= 3 & C(THH) &= 2 \\ C(HHT) &= 2 & C(THT) &= 1 \\ C(HTH) &= 2 & C(TTH) &= 1 \\ C(HTT) &= 1 & C(TTT) &= 0. \end{aligned}$$

Similarly, M is a function mapping each outcome another way:

$$\begin{aligned} M(HHH) &= 1 & M(THH) &= 0 \\ M(HHT) &= 0 & M(THT) &= 0 \\ M(HTH) &= 0 & M(TTH) &= 0 \\ M(HTT) &= 0 & M(TTT) &= 1. \end{aligned}$$

So C and M are random variables.

17.1.1 Indicator Random Variables

An *indicator random variable* is a random variable that maps every outcome to either 0 or 1. Indicator random variables are also called *Bernoulli variables*. The random variable M is an example. If all three coins match, then $M = 1$; otherwise, $M = 0$.

Indicator random variables are closely related to events. In particular, an indicator random variable partitions the sample space into those outcomes mapped to 1 and those outcomes mapped to 0. For example, the indicator M partitions the sample space into two blocks as follows:

$$\underbrace{HHH \quad TTT}_{M=1} \quad \underbrace{HHT \quad HTH \quad HTT \quad THH \quad THT \quad TTH}_{M=0}.$$

In the same way, an event E partitions the sample space into those outcomes in E and those not in E. So E is naturally associated with an indicator random variable, I_E, where $I_E(w) = 1$ for outcomes $w \in E$ and $I_E(w) = 0$ for outcomes $w \notin E$. Thus, $M = I_E$ where E is the event that all three coins match.

17.1.2 Random Variables and Events

There is a strong relationship between events and more general random variables as well. A random variable that takes on several values partitions the sample space into several blocks. For example, C partitions the sample space as follows:

$$\underbrace{TTT}_{C=0} \quad \underbrace{TTH \quad THT \quad HTT}_{C=1} \quad \underbrace{THH \quad HTH \quad HHT}_{C=2} \quad \underbrace{HHH}_{C=3}.$$

Each block is a subset of the sample space and is therefore an event. Thus, we can regard an equation or inequality involving a random variable as an event. For example, the event that $C = 2$ consists of the outcomes THH, HTH, and HHT. The event $C \leq 1$ consists of the outcomes TTT, TTH, THT, and HTT.

Naturally enough, we can talk about the probability of events defined by properties of random variables. For example,

$$\begin{aligned}\Pr[C = 2] &= \Pr[THH] + \Pr[HTH] + \Pr[HHT] \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{8} \\ &= \frac{3}{8}.\end{aligned}$$

As another example:

$$\begin{aligned}\Pr[M = 1] &= \Pr[TTT] + \Pr[HHH] \\ &= \frac{1}{8} + \frac{1}{8} \\ &= \frac{1}{4}.\end{aligned}$$

17.1.3 Functions of Random Variables

Random variables can be combined to form other random variables. For example, suppose that you roll two unbiased, independent 6-sided dice. Let D_i be the random variable denoting the outcome of the ith die for $i = 1, 2$. For example,

$$\Pr[D_1 = 3] = 1/6.$$

Then let $T = D_1 + D_2$. T is also a random variable and it denotes the sum of the two dice. For example,

$$\Pr[T = 2] = 1/36$$

and

$$\Pr[T = 7] = 1/6.$$

Random variables can be combined in complicated ways, as we will see in Chapter 19. For example,

$$Y = e^T$$

is also a random variable. In this case,

$$\Pr[Y = e^2] = 1/36$$

and

$$\Pr[Y = e^7] = 1/6.$$

17.1.4 Conditional Probability

Mixing conditional probabilities and events involving random variables creates no new difficulties. For example, $\Pr\left[C \geq 2 \mid M = 0\right]$ is the probability that at least two coins are heads ($C \geq 2$) given that not all three coins are the same ($M = 0$). We can compute this probability using the definition of conditional probability:

$$\begin{aligned}
\Pr\left[C \geq 2 \mid M = 0\right] &= \frac{\Pr[C \geq 2 \cap M = 0]}{\Pr[M = 0]} \\
&= \frac{\Pr[\{THH, HTH, HHT\}]}{\Pr[\{THH, HTH, HHT, HTT, THT, TTH\}]} \\
&= \frac{3/8}{6/8} \\
&= \frac{1}{2}.
\end{aligned}$$

The expression $C \geq 2 \cap M = 0$ on the first line may look odd; what is the set operation $\cap$ doing between an inequality and an equality? But recall that, in this context, $C \geq 2$ and $M = 0$ are *events*, and so they are *sets* of outcomes.

17.1.5 Independence

The notion of independence carries over from events to random variables as well. Random variables R_1 and R_2 are *independent* iff for all x_1 in the codomain of R_1, and x_2 in the codomain of R_2 for which $\Pr[R_2 = X_2] > 0$, we have:

$$\Pr\left[R_1 = x_1 \mid R_2 = x_2\right] = \Pr[R_1 = x_1].$$

As with events, we can formulate independence for random variables in an equivalent and perhaps more useful way: random variables R_1 and R_2 are independent if for all x_1 and x_2

$$\Pr[R_1 = x_1 \cap R_2 = x_2] = \Pr[R_1 = x_1] \cdot \Pr[R_2 = x_2].$$

For example, are C and M independent? Intuitively, the answer should be "no". The number of heads, C, completely determines whether all three coins match; that is, whether $M = 1$. But, to verify this intuition, we must find some $x_1, x_2 \in \mathbb{R}$ such that:

$$\Pr[C = x_1 \cap M = x_2] \neq \Pr[C = x_1] \cdot \Pr[M = x_2].$$

One appropriate choice of values is $x_1 = 2$ and $x_2 = 1$. In this case, we have:

$$\Pr[C = 2 \cap M = 1] = 0$$

and

$$\Pr[M = 1] \cdot \Pr[C = 2] = \frac{1}{4} \cdot \frac{3}{8} \neq 0.$$

The first probability is zero because we never have exactly two heads ($C = 2$) when all three coins match ($M = 1$). The other two probabilities were computed earlier.

On the other hand, let F be the indicator variable for the event that the first flip is a Head, so

$$\text{“}F = 1\text{”} = \{HHH, HTH, HHT, HTT\}.$$

Then F is independent of M, since

$$\Pr[M = 1] = 1/4 = \Pr\left[M = 1 \mid F = 1\right] = \Pr\left[M = 1 \mid F = 0\right]$$

and

$$\Pr[M = 0] = 3/4 = \Pr\left[M = 0 \mid F = 1\right] = \Pr\left[M = 0 \mid F = 0\right].$$

This example is an instance of a simple lemma:

Lemma 17.1.2. *Two events are independent iff their indicator variables are independent.*

As with events, the notion of independence generalizes to more than two random variables.

Definition 17.1.3. Random variables $R_1, R_2, \ldots, R_n$ are *mutually independent* iff

$$\begin{aligned}&\Pr[R_1 = x_1 \cap R_2 = x_2 \cap \cdots \cap R_n = x_n] \\ &= \Pr[R_1 = x_1] \cdot \Pr[R_2 = x_2] \cdots \Pr[R_n = x_n].\end{aligned}$$

for all $x_1, x_2, \ldots, x_n$.

A consequence of Definition 17.1.3 is that the probability that any *subset* of the variables takes a particular set of values is equal to the product of the probabilities that the individual variables take their values. Thus, for example, if $R_1, R_2, \ldots, R_{100}$ are mutually independent random variables, then it follows that:

$$\begin{aligned}&\Pr[R_1 = 7 \cap R_7 = 9.1 \cap R_{23} = \pi] \\ &\quad = \Pr[R_1 = 7] \cdot \Pr[R_7 = 9.1] \cdot \Pr[R_{23} = \pi].\end{aligned}$$

The proof is based on summing over all possible values for all of the other random variables.

17.2 Distribution Functions

A random variable maps outcomes to values. Often, random variables that show up for different spaces of outcomes wind up behaving in much the same way because they have the same probability of having any given value. Hence, random variables on different probability spaces may wind up having the same *probability density function*.

Definition 17.2.1. Let R be a random variable with codomain V. The *probability density function (pdf)* of R is a function $\mathrm{PDF}_R : V \to [0,1]$ defined by:

$$\mathrm{PDF}_R(x) ::= \begin{cases} \Pr[R = x] & \text{if } x \in \mathrm{range}(R) \\ 0 & \text{if } x \notin \mathrm{range}(R). \end{cases}$$

A consequence of this definition is that

$$\sum_{x \in \mathrm{range}(R)} \mathrm{PDF}_R(x) = 1.$$

This is because R has a value for each outcome, so summing the probabilities over all outcomes is the same as summing over the probabilities of each value in the range of R.

As an example, suppose that you roll two unbiased, independent, 6-sided dice. Let T be the random variable that equals the sum of the two rolls. This random variable takes on values in the set $V = \{2, 3, \ldots, 12\}$. A plot of the probability density function for T is shown in Figure 17.1: The lump in the middle indicates that sums close to 7 are the most likely. The total area of all the rectangles is 1 since the dice must take on exactly one of the sums in $V = \{2, 3, \ldots, 12\}$.

A closely-related concept to a PDF is the *cumulative distribution function (cdf)* for a random variable whose codomain is the real numbers. This is a function $\mathrm{CDF}_R : \mathbb{R} \to [0,1]$ defined by:

$$\mathrm{CDF}_R(x) = \Pr[R \leq x].$$

As an example, the cumulative distribution function for the random variable T is shown in Figure 17.2: The height of the ith bar in the cumulative distribution function is equal to the *sum* of the heights of the leftmost i bars in the probability

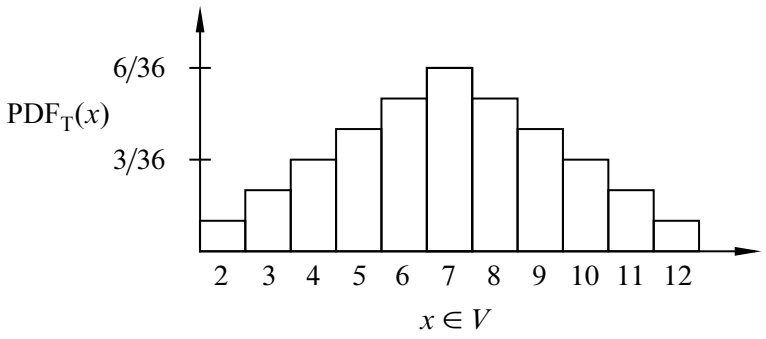


Figure 17.1 The probability density function for the sum of two 6-sided dice.

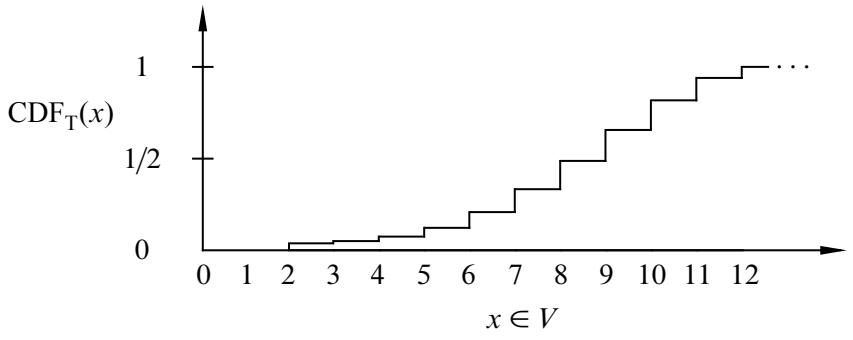


Figure 17.2 The cumulative distribution function for the sum of two 6-sided dice.

density function. This follows from the definitions of pdf and cdf:

$$\begin{aligned}\mathrm{CDF}_R(x) &= \Pr[R \le x] \\ &= \sum_{y \le x} \Pr[R = y] \\ &= \sum_{y \le x} \mathrm{PDF}_R(y).\end{aligned}$$

In summary, $\mathrm{PDF}_R(x)$ measures the probability that $R = x$ and $\mathrm{CDF}_R(x)$ measures the probability that $R \le x$. Both PDF_R and CDF_R capture the same information about the random variable R—you can derive one from the other—but sometimes one is more convenient.

One of the really interesting things about density functions and distribution functions is that many random variables turn out to have the *same* pdf and cdf. In other words, even though R and S are different random variables on different probability spaces, it is often the case that

$$\mathrm{PDF}_R = \mathrm{PDF}_s.$$

In fact, some pdfs are so common that they are given special names. For example, the three most important distributions in computer science are the *Bernoulli distribution*, the *uniform distribution*, and the *binomial distribution*. We look more closely at these common distributions in the next several sections.

17.3 Bernoulli Distributions

The Bernoulli distribution is the simplest and most common distribution function. That's because it is the distribution function for an indicator random variable. Specifically, the *Bernoulli distribution* has a probability density function of the form $f_p : \{0, 1\} \to [0, 1]$ where

$$\begin{aligned}f_p(0) &= p, \quad \text{and} \\ f_p(1) &= 1 - p,\end{aligned}$$

for some $p \in [0, 1]$. The corresponding cumulative distribution function is $F_p : \mathbb{R} \to [0, 1]$ where:

$$F_p(x) = \begin{cases} 0 & \text{if } x < 0 \\ p & \text{if } 0 \le x < 1 \\ 1 & \text{if } 1 \le x. \end{cases}$$

17.4 Uniform Distributions

17.4.1 Definition

A random variable that takes on each possible value with the same probability is said to be *uniform*. If the sample space is $\{1, 2, \ldots, n\}$, then the *uniform distribution* has a pdf of the form

$$f_n : \{1, 2, \ldots, n\} \to [0, 1]$$

where

$$f_n(k) = \frac{1}{n}$$

for some $n \in \mathbb{N}^+$. The cumulative distribution function is then $F_n : \mathbb{R} \to [0, 1]$ where

$$F_n(x) = \begin{cases} 0 & \text{if } x < 1 \\ k/n & \text{if } k \le x < k + 1 \text{ for } 1 \le k < n \\ 1 & \text{if } n \le x. \end{cases}$$

Uniform distributions arise frequently in practice. For example, the number rolled on a fair die is uniform on the set $\{1, 2, \ldots, 6\}$. If $p = 1/2$, then an indicator random variable is uniform on the set $\{0, 1\}$.

17.4.2 The Numbers Game

Enough definitions—let's play a game! I have two envelopes. Each contains an integer in the range $0, 1, \ldots, 100$, and the numbers are distinct. To win the game, you must determine which envelope contains the larger number. To give you a fighting chance, we'll let you peek at the number in one envelope selected at random. Can you devise a strategy that gives you a better than 50% chance of winning?

For example, you could just pick an envelope at random and guess that it contains the larger number. But this strategy wins only 50% of the time. Your challenge is to do better.

So you might try to be more clever. Suppose you peek in one envelope and see the number 12. Since 12 is a small number, you might guess that that the number in the other envelope is larger. But perhaps we've been tricky and put small numbers in *both* envelopes. Then your guess might not be so good!

An important point here is that the numbers in the envelopes may *not* be random. We're picking the numbers and we're choosing them in a way that we think will defeat your guessing strategy. We'll only use randomization to choose the numbers if that serves our purpose, which is to make you lose!

Intuition Behind the Winning Strategy

Amazingly, there is a strategy that wins more than 50% of the time, regardless of what numbers we put in the envelopes!

Suppose that you somehow knew a number x that was in between the numbers in the envelopes. Now you peek in one envelope and see a number. If it is bigger than x, then you know you're peeking at the higher number. If it is smaller than x, then you're peeking at the lower number. In other words, if you know a number x between the numbers in the envelopes, then you are certain to win the game.

The only flaw with this brilliant strategy is that you do *not* know such an x. Oh well.

But what if you try to *guess* x? There is some probability that you guess correctly. In this case, you win 100% of the time. On the other hand, if you guess incorrectly, then you're no worse off than before; your chance of winning is still 50%. Combining these two cases, your overall chance of winning is better than 50%!

Informal arguments about probability, like this one, often sound plausible, but do not hold up under close scrutiny. In contrast, this argument sounds completely implausible—but is actually correct!

Analysis of the Winning Strategy

For generality, suppose that we can choose numbers from the set $\{0, 1, \ldots, n\}$. Call the lower number L and the higher number H.

Your goal is to guess a number x between L and H. To avoid confusing equality cases, you select x at random from among the half-integers:

$$\left\{\frac{1}{2},\ 1\frac{1}{2},\ 2\frac{1}{2},\ \ldots,\ n-\frac{1}{2}\right\}$$

But what probability distribution should you use?

The uniform distribution turns out to be your best bet. An informal justification is that if we figured out that you were unlikely to pick some number—say $50\frac{1}{2}$—then we'd always put 50 and 51 in the envelopes. Then you'd be unlikely to pick an x between L and H and would have less chance of winning.

After you've selected the number x, you peek into an envelope and see some number T. If $T > x$, then you guess that you're looking at the larger number. If $T < x$, then you guess that the other number is larger.

All that remains is to determine the probability that this strategy succeeds. We can do this with the usual four step method and a tree diagram.

Step 1: Find the sample space.
You either choose x too low ($< L$), too high ($> H$), or just right ($L < x < H$). Then you either peek at the lower number ($T = L$) or the higher number ($T = H$). This gives a total of six possible outcomes, as show in Figure 17.3.

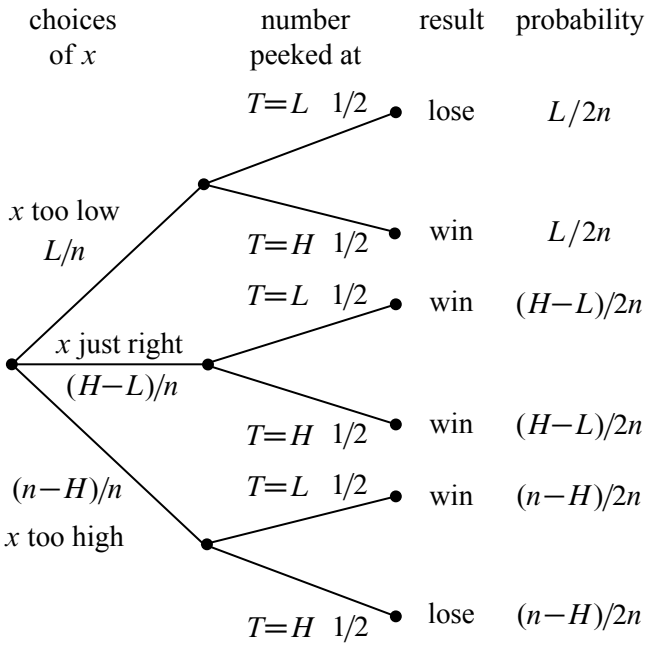


Figure 17.3 The tree diagram for the numbers game.

Step 2: Define events of interest.
The four outcomes in the event that you win are marked in the tree diagram.

Step 3: Assign outcome probabilities.
First, we assign edge probabilities. Your guess x is too low with probability L/n, too high with probability $(n - H)/n$, and just right with probability $(H - L)/n$. Next, you peek at either the lower or higher number with equal probability. Multiplying along root-to-leaf paths gives the outcome probabilities.

Step 4: Compute event probabilities.
The probability of the event that you win is the sum of the probabilities of the four outcomes in that event:

$$\begin{aligned}\Pr[\text{win}] &= \frac{L}{2n} + \frac{H-L}{2n} + \frac{H-L}{2n} + \frac{n-H}{2n} \\ &= \frac{1}{2} + \frac{H-L}{2n} \\ &\geq \frac{1}{2} + \frac{1}{2n}\end{aligned}$$

The final inequality relies on the fact that the higher number H is at least 1 greater than the lower number L since they are required to be distinct.

Sure enough, you win with this strategy more than half the time, regardless of the numbers in the envelopes! For example, if I choose numbers in the range $0, 1, \ldots, 100$, then you win with probability at least $\frac{1}{2} + \frac{1}{200} = 50.5\%$. Even better, if I'm allowed only numbers in the range $0, \ldots, 10$, then your probability of winning rises to 55%! By Las Vegas standards, those are great odds!

17.4.3 Randomized Algorithms

The best strategy to win the numbers game is an example of a *randomized algorithm*—it uses random numbers to influence decisions. Protocols and algorithms that make use of random numbers are very important in computer science. There are many problems for which the best known solutions are based on a random number generator.

For example, the most commonly-used protocol for deciding when to send a broadcast on a shared bus or Ethernet is a randomized algorithm known as *exponential backoff*. One of the most commonly-used sorting algorithms used in practice, called *quicksort*, uses random numbers. You'll see many more examples if you take an algorithms course. In each case, randomness is used to improve the probability that the algorithm runs quickly or otherwise performs well.

17.5 Binomial Distributions

17.5.1 Definitions

The third commonly-used distribution in computer science is the *binomial distribution*. The standard example of a random variable with a binomial distribution is the number of heads that come up in n independent flips of a coin. If the coin is

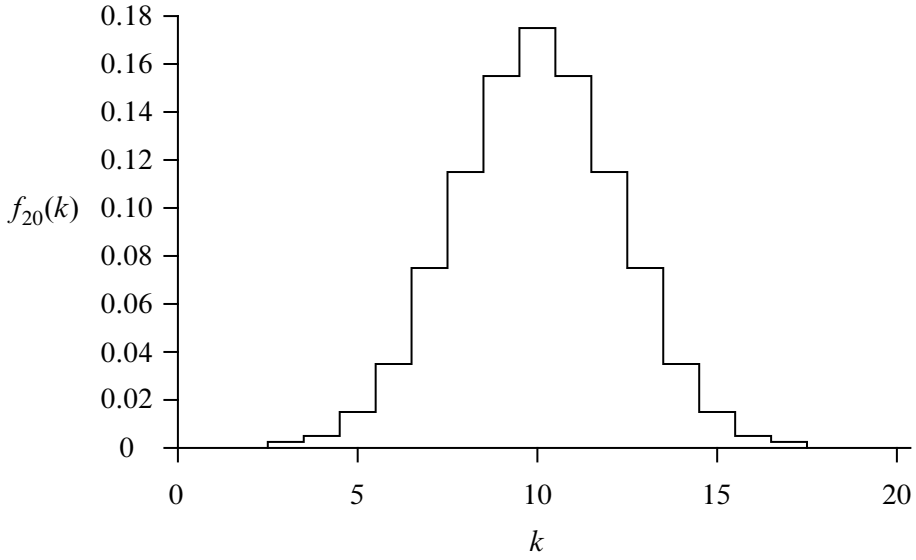


Figure 17.4 The pdf for the unbiased binomial distribution for $n = 20$, $f_{20}(k)$.

fair, then the number of heads has an *unbiased binomial distribution*, specified by the pdf

$$f_n : \{1, 2, \ldots, n\} \to [0, 1]$$

where

$$f_n(k) = \binom{n}{k} 2^{-n}$$

for some $n \in \mathbb{N}^+$. This is because there are $\binom{n}{k}$ sequences of n coin tosses with exactly k heads, and each such sequence has probability 2^{-n}.

A plot of $f_{20}(k)$ is shown in Figure 17.4. The most likely outcome is $k = 10$ heads, and the probability falls off rapidly for larger and smaller values of k. The falloff regions to the left and right of the main hump are called the *tails of the distribution*. We'll talk a lot more about these tails shortly.

The cumulative distribution function for the unbiased binomial distribution is $F_n : \mathbb{R} \to [0, 1]$ where

$$F_n(x) = \begin{cases} 0 & \text{if } x < 1 \\ \sum_{i=0}^{k} \binom{n}{i} 2^{-n} & \text{if } k \le x < k+1 \text{ for } 1 \le k < n \\ 1 & \text{if } n \le x. \end{cases}$$

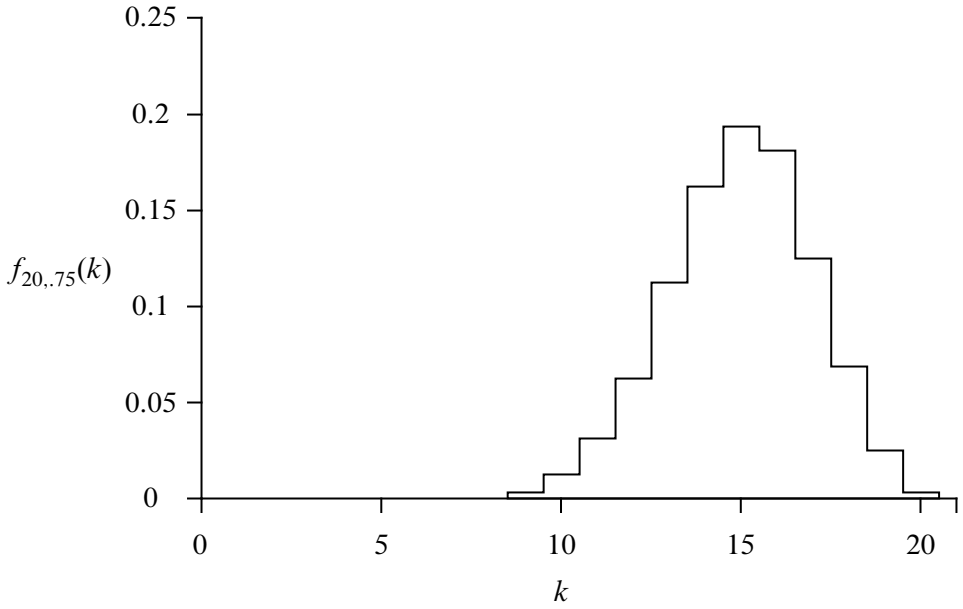


Figure 17.5 The pdf for the general binomial distribution $f_{n,p}(k)$ for $n = 20$ and $p = .75$.

The General Binomial Distribution

If the coins are biased so that each coin is heads with probability p, then the number of heads has a *general binomial density function* specified by the pdf

$$f_{n,p} : \{1, 2, \ldots, n\} \to [0, 1]$$

where

$$f_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

for some $n \in \mathbb{N}^+$ and $p \in [0, 1]$. This is because there are $\binom{n}{k}$ sequences with k heads and $n - k$ tails, but now the probability of each such sequence is $p^k(1-p)^{n-k}$.

For example, the plot in Figure 17.5 shows the probability density function $f_{n,p}(k)$ corresponding to flipping $n = 20$ independent coins that are heads with probability $p = 0.75$. The graph shows that we are most likely to get $k = 15$ heads, as you might expect. Once again, the probability falls off quickly for larger and smaller values of k.

The cumulative distribution function for the general binomial distribution is $F_{n,p} : \mathbb{R} \to [0,1]$ where

$$F_{n,p}(x) = \begin{cases} 0 & \text{if } x < 1 \\ \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} & \text{if } k \le x < k+1 \text{ for } 1 \le k < n \\ 1 & \text{if } n \le x. \end{cases} \tag{17.1}$$

17.5.2 Approximating the Probability Density Function

Computing the general binomial density function is daunting when k and n are large. Fortunately, there is an approximate closed-form formula for this function based on an approximation for the binomial coefficient. In the formula below, k is replaced by αn where α is a number between 0 and 1.

Lemma 17.5.1.

$$\binom{n}{\alpha n} \sim \frac{2^{nH(\alpha)}}{\sqrt{2\pi\alpha(1-\alpha)n}} \tag{17.2}$$

and

$$\binom{n}{\alpha n} < \frac{2^{nH(\alpha)}}{\sqrt{2\pi\alpha(1-\alpha)n}} \tag{17.3}$$

where $H(\alpha)$ is the entropy function[2]

$$H(\alpha) ::= \alpha \log\left(\frac{1}{\alpha}\right) + (1-\alpha)\log\left(\frac{1}{1-\alpha}\right).$$

Moreover, if $\alpha n > 10$ and $(1-\alpha)n > 10$, then the left and right sides of Equation 17.2 differ by at most 2%. If $\alpha n > 100$ and $(1-\alpha)n > 100$, then the difference is at most 0.2%.

The graph of H is shown in Figure 17.6.

Lemma (17.5.1) provides an excellent approximation for binomial coefficients. We'll skip its derivation, which consists of plugging in Theorem 9.6.1 for the factorials in the binomial coefficient and then simplifying.

Now let's plug Equation 17.2 into the general binomial density function. The probability of flipping αn heads in n tosses of a coin that comes up heads with

[2] $\log(x)$ means $\log_2(x)$.

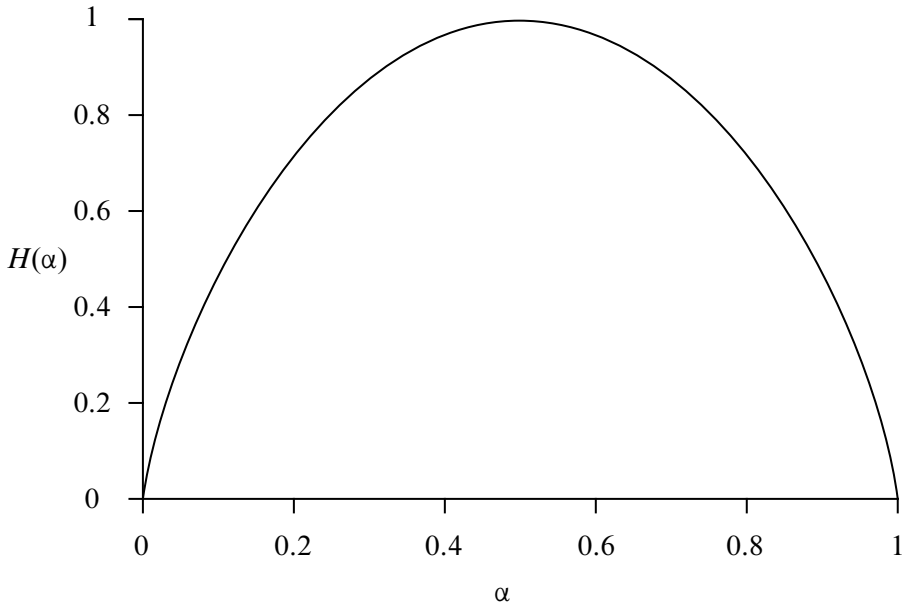


Figure 17.6 The Entropy Function

probability p is:

$$f_{n,p}(\alpha n) \sim \frac{2^{nH(\alpha)} p^{\alpha n} (1-p)^{(1-\alpha)n}}{\sqrt{2\pi\alpha(1-\alpha)n}}$$

$$= \frac{2^{n\left(\alpha \log\left(\frac{p}{\alpha}\right) + (1-\alpha)\log\left(\frac{1-p}{1-\alpha}\right)\right)}}{\sqrt{2\pi\alpha(1-\alpha)n}}, \tag{17.4}$$

where the margin of error in the approximation is the same as in Lemma 17.5.1. From Equation 17.3, we also find that

$$f_{n,p}(\alpha n) < \frac{2^{n\left(\alpha \log\left(\frac{p}{\alpha}\right) + (1-\alpha)\log\left(\frac{1-p}{1-\alpha}\right)\right)}}{\sqrt{2\pi\alpha(1-\alpha)n}}. \tag{17.5}$$

The formula in Equations 17.4 and 17.5 is as ugly as a bowling shoe, but it's useful because it's easy to evaluate. For example, suppose we flip a fair coin n times. What is the probability of getting *exactly* pn heads? Plugging $\alpha = p$ into Equation 17.4 gives:

$$f_{n,p}(pn) \sim \frac{1}{\sqrt{2\pi p(1-p)n}}.$$

Thus, for example, if we flip a fair coin (where $p = 1/2$) $n = 100$ times, the probability of getting exactly 50 heads is within 2% of 0.079, which is about 8%.

17.5.3 Approximating the Cumulative Distribution Function

In many fields, including computer science, probability analyses come down to getting small bounds on the tails of the binomial distribution. In a typical application, you want to bound the tails in order to show that there is very small probability that too many *bad* things happen. For example, we might like to know that it is very unlikely that too many bits are corrupted in a message, or that too many servers or communication links become overloaded, or that a randomized algorithm runs for too long.

So it is usually good news that the binomial distribution has small tails. To get a feel for their size, consider the probability of flipping at most 25 heads in 100 independent tosses of a fair coin.

The probability of getting at most αn heads is given by the binomial cumulative distribution function

$$F_{n,p}(\alpha n) = \sum_{i=0}^{\alpha n} \binom{n}{i} p^i (1-p)^{n-i}. \tag{17.6}$$

We can bound this sum by bounding the ratio of successive terms.

In particular, for $i \leq \alpha n$,

$$\begin{aligned}
\frac{\binom{n}{i-1} p^{i-1}(1-p)^{n-(i-1)}}{\binom{n}{i} p^i (1-p)^{n-i}} &= \frac{\frac{n!p^{i-1}(1-p)^{n-i+1}}{(i-1)!(n-i+1)!}}{\frac{n!p^i(1-p)^{n-i}}{i!(n-i)!}} \\
&= \frac{i(1-p)}{(n-i+1)p} \\
&\leq \frac{\alpha n(1-p)}{(n-\alpha n+1)p} \\
&\leq \frac{\alpha(1-p)}{(1-\alpha)p}.
\end{aligned}$$

This means that for $\alpha < p$,

$$\begin{aligned} F_{n,p}(\alpha n) &< f_{n,p}(\alpha n) \sum_{i=0}^{\infty} \left[\frac{\alpha(1-p)}{(1-\alpha)p}\right]^i \\ &= \frac{f_{n,p}(\alpha n)}{1 - \dfrac{\alpha(1-p)}{(1-\alpha)p}} \\ &= \left(\frac{1-\alpha}{1-\alpha/p}\right) f_{n,p}(\alpha n). \end{aligned} \tag{17.7}$$

In other words, the probability of at most αn heads is at most

$$\frac{1-\alpha}{1-\alpha/p}$$

times the probability of exactly αn heads. For our scenario, where $p = 1/2$ and $\alpha = 1/4$,

$$\frac{1-\alpha}{1-\alpha/p} = \frac{3/4}{1/2} = \frac{3}{2}.$$

Plugging $n = 100$, $\alpha = 1/4$, and $p = 1/2$ into Equation 17.5, we find that the probability of at most 25 heads in 100 coin flips is

$$F_{100,1/2}(25) < \frac{3}{2} \cdot \frac{2^{100\left(\frac{1}{4}\log(2) + \frac{3}{4}\log\left(\frac{2}{3}\right)\right)}}{\sqrt{75\pi/2}} \leq 3 \cdot 10^{-7}.$$

This says that flipping 25 or fewer heads is extremely unlikely, which is consistent with our earlier claim that the tails of the binomial distribution are very small. In fact, notice that the probability of flipping *25 or fewer* heads is only 50% more than the probability of flipping *exactly 25* heads. Thus, flipping exactly 25 heads is twice as likely as flipping any number between 0 and 24!

Caveat. The upper bound on $F_{n,p}(\alpha n)$ in Equation 17.7 holds only if $\alpha < p$. If this is not the case in your problem, then try thinking in complementary terms; that is, look at the number of tails flipped instead of the number of heads. In fact, this is precisely what we will do in the next example.

17.5.4 Noisy Channels

Suppose you are sending packets of data across a communication channel and that each packet is lost with probability $p = .01$. Also suppose that packet losses are independent. You need to figure out how much redundancy (or error correction) to

build into your communication protocol. Since redundancy is expensive overheard, you would like to use as little as possible. On the other hand, you never want to be caught short. Would it be safe for you to assume that in any batch of 10,000 packets, only 200 (or 2%) are lost? Let's find out.

The noisy channel is analogous to flipping $n = 10{,}000$ independent coins, each with probability $p = .01$ of coming up heads, and asking for the probability that there are at least αn heads where $\alpha = .02$. Since $\alpha > p$, we cannot use Equation 17.7. So we need to recast the problem by looking at the numbers of tails. In this case, the probability of tails is $p = .99$ and we are asking for the probability of at most αn tails where $\alpha = .98$.

Now we can use Equations 17.5 and 17.7 to find that the probability of losing 2% or more of the 10,000 packets is at most

$$\left(\frac{1-.98}{1-.98/.99}\right)\frac{2^{10000\left(.98\log\left(\frac{.99}{.98}\right)+.02\log\left(\frac{.01}{.02}\right)\right)}}{\sqrt{2\pi(.98)(1-.98)10000}} < 2^{-60}.$$

This is good news. It says that planning on at most 2% packet loss in a batch of 10,000 packets should be very safe, at least for the next few millennia.

17.5.5 Estimation by Sampling

Sampling is a very common technique for estimating the fraction of elements in a set that have a certain property. For example, suppose that you would like to know how many Americans plan to vote for the Republican candidate in the next presidential election. It is infeasible to ask every American how they intend to vote, so pollsters will typically contact n Americans selected at random and then compute the fraction of *those* Americans that will vote Republican. This value is then used as the *estimate* of the number of all Americans that will vote Republican. For example, if 45% of the n contacted voters report that they will vote Republican, the pollster reports that 45% of all Americans will vote Republican. In addition, the pollster will usually also provide some sort of qualifying statement such as

> "There is a 95% probability that the poll is accurate to within ± 4 percentage points."

The qualifying statement is often the source of confusion and misinterpretation. For example, many people interpret the qualifying statement to mean that there is a 95% chance that between 41% and 49% of Americans intend to vote Republican. But this is wrong! The fraction of Americans that intend to vote Republican is a fixed (and unknown) value p that is *not* a random variable. Since p is not a random variable, we cannot say anything about the probability that $.41 \leq p \leq .49$.

To obtain a correct interpretation of the qualifying statement and the results of the poll, it is helpful to introduce some notation.

Define R_i to be the indicator random variable for the ith contacted American in the sample. In particular, set $R_i = 1$ if the ith contacted American intends to vote Republican and $R_i = 0$ otherwise. For the purposes of the analysis, we will assume that the ith contacted American is selected uniformly at random (with replacement) from the set of all Americans.[3] We will also assume that every contacted person responds honestly about whether or not they intend to vote Republican and that there are only two options—each American intends to vote Republican or they don't. Thus,

$$\Pr[R_i = 1] = p \tag{17.8}$$

where p is the (unknown) fraction of Americans that intend to vote Republican.

We next define

$$T = R_1 + R_2 + \cdots + R_n$$

to be the number of contacted Americans who intend to vote Republican. Then T/n is a random variable that is the estimate of the fraction of Americans that intend to vote Republican.

We are now ready to provide the correct interpretation of the qualifying statement. The poll results mean that

$$\Pr\left[\,|T/n - p| \leq .04\right] \geq .95. \tag{17.9}$$

In other words, there is a 95% chance that the sample group will produce an estimate that is within ± 4 percentage points of the correct value for the overall population. So either we were "unlucky" in selecting the people to poll or the results of the poll will be correct to within ± 4 points.

How Many People Do We Need to Contact?

There remains an important question: how many people n do we need to contact to make sure that Equation 17.9 is true? In general, we would like n to be as small as possible in order to minimize the cost of the poll.

Surprisingly, the answer depends only on the desired *accuracy* and *confidence* of the poll and not on the number of items in the set being sampled. In this case, the desired accuracy is .04, the desired confidence is .95, and the set being sampled is the set of Americans. It's a good thing that n won't depend on the size of the set being sampled—there are over 300 million Americans!

[3] This means that someone could be contacted multiple times.

The task of finding an n that satisfies Equation 17.9 is made tractable by observing that T has a general binomial distribution with parameters n and p and then applying Equations 17.5 and 17.7. Let's see how this works.

Since we will be using bounds on the tails of the binomial distribution, we first do the standard conversion

$$\Pr\left[\,|T/n - p| \le .04\right] \;=\; 1 - \Pr\left[\,|T/n - p| > .04\right].$$

We then proceed to upper bound

$$\begin{aligned}\Pr\left[\,|T/n - p| > .04\right] &= \Pr[T < (p - .04)n] + \Pr[T > (p + .04)n] \\ &= F_{n,p}((p - 0.4)n) + F_{n,1-p}((1 - p - .04)n). \quad (17.10)\end{aligned}$$

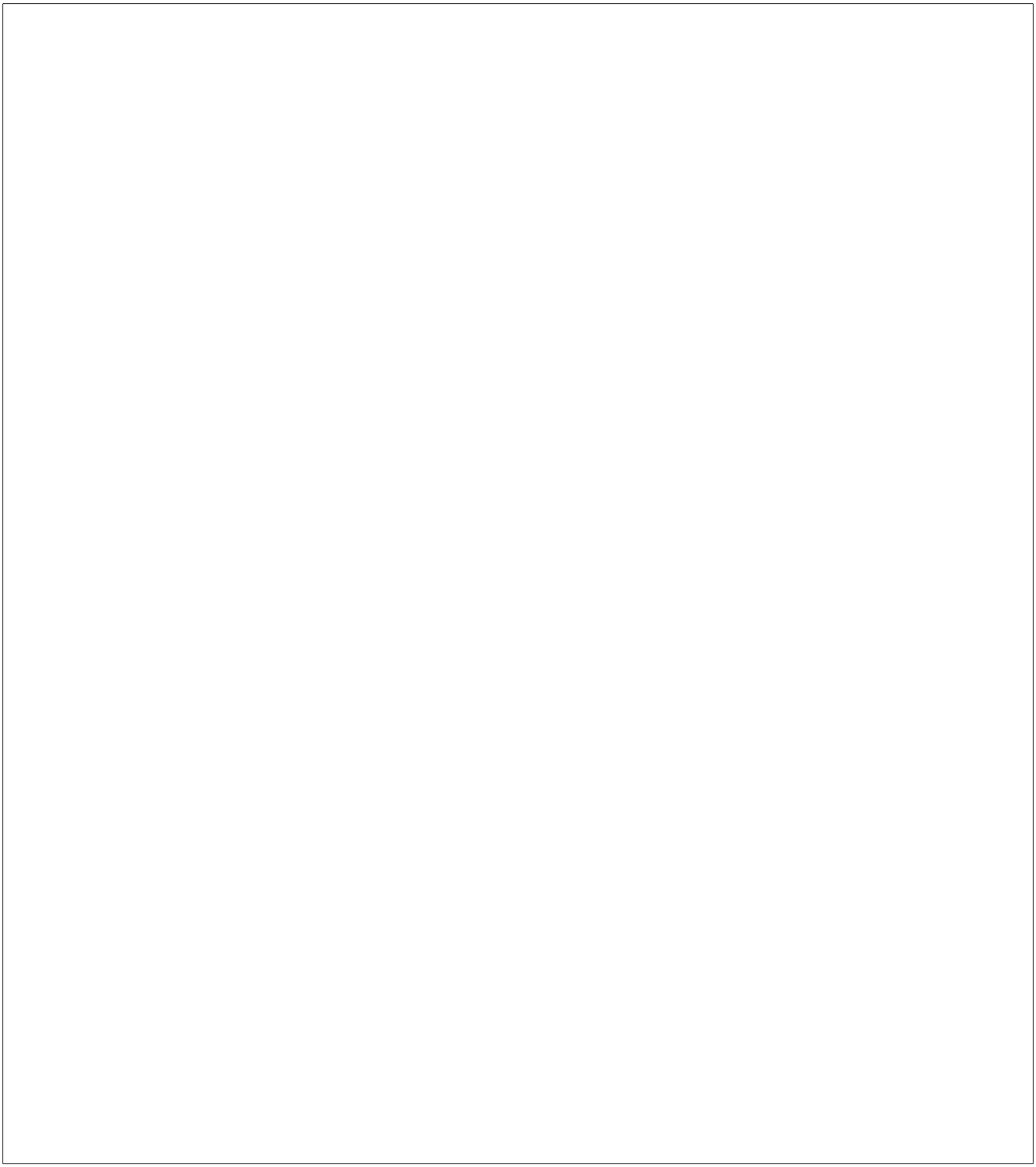
We don't know the true value of p, but it turns out that the expression on the righthand side of Equation 17.10 is maximized when $p = 1/2$ and so

$$\begin{aligned}\Pr\left[\,|T/n - p| > .04\right] &\le 2F_{n,1/2}(.46n) \\ &< 2\left(\frac{1 - .46}{1 - (.46/.5)}\right) f_{n,1/2}(.46n) \\ &< 13.5 \cdot \frac{2^{n\left(.46\log\left(\frac{.5}{.46}\right) + .54\log\left(\frac{.5}{.54}\right)\right)}}{\sqrt{2\pi \cdot 0.46 \cdot 0.54 \cdot n}} \\ &< \frac{10.81 \cdot 2^{-.00462n}}{\sqrt{n}}. \quad (17.11)\end{aligned}$$

The second line comes from Equation 17.7 using $\alpha = .46$. The third line comes from Equation 17.5.

Equation 17.11 provides bounds on the confidence of the poll for different values of n. For example, if $n = 665$, the bound in Equation 17.11 evaluates to .04978.... Hence, if the pollster contacts 665 Americans, the poll will be accurate to within ± 4 percentage points with at least 95% probability.

Since the bound in Equation 17.11 is exponential in n, the confidence increases greatly as n increases. For example, if $n = 6{,}650$ Americans are contacted, the poll will be accurate to within ± 4 points with probability at least $1 - 10^{-10}$. Of course, most pollsters are not willing to pay the added cost of polling 10 times as many people when they already have a confidence level of 95% from polling 665 people.



18 Expectation

18.1 Definitions and Examples

The *expectation* or *expected value* of a random variable is a single number that tells you a lot about the behavior of the variable. Roughly, the expectation is the average value of the random variable where each value is weighted according to its probability. Formally, the expected value (also known as the *average* or *mean*) of a random variable is defined as follows.

Definition 18.1.1. If R is a random variable defined on a sample space $\mathcal{S}$, then the expectation of R is

$$\mathrm{Ex}[R] ::= \sum_{w \in \mathcal{S}} R(w) \Pr[w]. \tag{18.1}$$

For example, suppose $\mathcal{S}$ is the set of students in a class, and we select a student uniformly at random. Let R be the selected student's exam score. Then $\mathrm{Ex}[R]$ is just the class average—the first thing everyone wants to know after getting their test back! For similar reasons, the first thing you usually want to know about a random variable is its expected value.

Let's work through some examples.

18.1.1 The Expected Value of a Uniform Random Variable

Let R be the value that comes up with you roll a fair 6-sided die. The the expected value of R is

$$\mathrm{Ex}[R] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

This calculation shows that the name "expected value" is a little misleading; the random variable might *never* actually take on that value. You don't ever expect to roll a $3\frac{1}{2}$ on an ordinary die!

Also note that the mean of a random variable is not the same as the *median*. The median is the midpoint of a distribution.

Definition 18.1.2. The *median*[1] of a random variable R is the value $x \in \mathrm{range}(R)$

[1] Some texts define the median to be the value of $x \in \mathrm{range}(R)$ for which $\Pr[R \leq x] < 1/2$ and $\Pr[R > x] \leq 1/2$. The difference in definitions is not important.

such that

$$\Pr[R \leq x] \leq \frac{1}{2} \qquad \text{and}$$
$$\Pr[R > x] < \frac{1}{2}.$$

In this text, we will not devote much attention to the median. Rather, we will focus on the expected value, which is much more interesting and useful.

Rolling a 6-sided die provides an example of a uniform random variable. In general, if R_n is a random variable with a uniform distribution on $\{1, 2, \ldots, n\}$, then

$$\text{Ex}[R_n] = \sum_{i=1}^{n} i \cdot \frac{1}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}.$$

18.1.2 The Expected Value of an Indicator Random Variable

The expected value of an indicator random variable for an event is just the probability of that event.

Lemma 18.1.3. *If I_A is the indicator random variable for event A, then*

$$\text{Ex}[I_A] = \Pr[A].$$

Proof.

$$\begin{aligned}\text{Ex}[I_A] &= 1 \cdot \Pr[I_A = 1] + 0 \cdot \Pr[I_A = 0] \\ &= \Pr[I_A = 1] \\ &= \Pr[A]. \qquad \text{(def of } I_A\text{)}\end{aligned}$$

■

For example, if A is the event that a coin with bias p comes up heads, then $\text{Ex}[I_A] = \Pr[I_A = 1] = p$.

18.1.3 Alternate Definitions

There are several equivalent ways to define expectation.

Theorem 18.1.4. *If R is a random variable defined on a sample space S then*

$$\text{Ex}[R] = \sum_{x \in \text{range}(R)} x \cdot \Pr[R = x]. \tag{18.2}$$

The proof of Theorem 18.1.4, like many of the elementary proofs about expectation in this chapter, follows by judicious regrouping of terms in the Equation 18.1:

Proof.

$$\begin{aligned}
\mathrm{Ex}[R] &= \sum_{\omega\in\mathcal{S}} R(\omega)\Pr[\omega] && \text{(Def 18.1.1 of expectation)}\\
&= \sum_{x\in\mathrm{range}(R)}\ \sum_{\omega\in[R=x]} R(\omega)\Pr[\omega]\\
&= \sum_{x\in\mathrm{range}(R)}\ \sum_{\omega\in[R=x]} x\Pr[\omega] && \text{(def of the event } [R=x])\\
&= \sum_{x\in\mathrm{range}(R)} x\left(\sum_{\omega\in[R=x]}\Pr[\omega]\right) && \text{(distributing } x \text{ over the inner sum)}\\
&= \sum_{x\in\mathrm{range}(R)} x\cdot\Pr[R=x]. && \text{(def of } \Pr[R=x])
\end{aligned}$$

The first equality follows because the events $[R = x]$ for $x \in \mathrm{range}(R)$ partition the sample space $\mathcal{S}$, so summing over the outcomes in $[R = x]$ for $x \in \mathrm{range}(R)$ is the same as summing over $\mathcal{S}$. ■

In general, Equation 18.2 is more useful than Equation 18.1 for calculating expected values and has the advantage that it does not depend on the sample space, but only on the density function of the random variable. It is especially useful when the range of the random variable is $\mathbb{N}$, as we will see from the following corollary.

Corollary 18.1.5. *If the range of a random variable R is $\mathbb{N}$, then*

$$\mathrm{Ex}[R] = \sum_{i=1}^{\infty} i\Pr[R=i] = \sum_{i=0}^{\infty}\Pr[R>i].$$

Proof. The first equality follows directly from Theorem 18.1.4 and the fact that $\mathrm{range}(R) = \mathbb{N}$. The second equality is derived by adding the following equations:

$$\begin{array}{lllll}
\Pr[R>0] = & \Pr[R=1] & + \Pr[R=2] & + \Pr[R=3] & +\cdots\\
\Pr[R>1] = & & \Pr[R=2] & + \Pr[R=3] & +\cdots\\
\Pr[R>2] = & & & \Pr[R=3] & +\cdots\\
& & \vdots & &
\end{array}$$

$$\begin{aligned}
\sum_{i=0}^{\infty}\Pr[R>i] &= 1\cdot\Pr[R=1] + 2\cdot\Pr[R=2] + 3\cdot\Pr[R=3] + \cdots\\
&= \sum_{i=1}^{\infty} i\Pr[R=i].
\end{aligned}$$

■

18.1.4 Mean Time to Failure

The mean time to failure is a critical parameter in the design of most any system. For example, suppose that a computer program crashes at the end of each hour of use with probability p, if it has not crashed already. What is the expected time until the program crashes?

If we let C be the number of hours until the crash, then the answer to our problem is $\text{Ex}[C]$. C is a random variable with values in $\mathbb{N}$ and so we can use Corollary 18.1.5 to determine that

$$\text{Ex}[C] = \sum_{i=0}^{\infty} \Pr[C > i]. \tag{18.3}$$

$\Pr[C > i]$ is easy to evaluate: a crash happens later than the ith hour iff the system did not crash during the first i hours, which happens with probability $(1-p)^i$. Plugging this into Equation 18.3 gives:

$$\begin{aligned} \text{Ex}[C] &= \sum_{i=0}^{\infty} (1-p)^i \\ &= \frac{1}{1-(1-p)} && \text{(sum of geometric series)} \\ &= \frac{1}{p}. \end{aligned} \tag{18.4}$$

For example, if there is a 1% chance that the program crashes at the end of each hour, then the expected time until the program crashes is $1/0.01 = 100$ hours.

The general principle here is well-worth remembering:

> *If a system fails at each time step with probability p, then the expected number of steps up to (and including) the first failure is $1/p$.*

Making Babies

As a related example, suppose a couple really wants to have a baby girl. For simplicity, assume that there is a 50% chance that each child they have is a girl, and that the genders of their children are mutually independent. If the couple insists on having children until they get a girl, then how many baby boys should they expect first?

The question, "How many hours until the program crashes?" is mathematically the same as the question, "How many children must the couple have until they get a girl?" In this case, a crash corresponds to having a girl, so we should set

$p = 1/2$. By the preceding analysis, the couple should expect a baby girl after having $1/p = 2$ children. Since the last of these will be the girl, they should expect just one boy.

18.1.5 Dealing with Infinity

The analysis of the mean time to failure was easy enough. But if you think about it further, you might start to wonder about the case when the computer program *never* fails. For example, what if the program runs forever? How do we handle outcomes with an infinite value?

These are good questions and we wonder about them too. Indeed, mathematicians have gone to a lot of work to reason about sample spaces with an infinite number of outcomes or outcomes with infinite value.

To keep matters simple in this text, we will follow the common convention of ignoring the contribution of outcomes that have probability zero when computing expected values. This means that we can safely ignore the "never-fail" outcome, because it has probability

$$\lim_{n\to\infty}(1-p)^n = 0.$$

In general, when we are computing expectations for infinite sample spaces, we will generally focus our attention on a subset of outcomes that occur with collective probability one. For the most part, this will allow us to ignore the "infinite" outcomes because they will typically happen with probability zero.[2]

This assumption does *not* mean that the expected value of a random variable is always finite, however. Indeed, there are many examples where the expected value is infinite. And where infinity raises its ugly head, trouble is sure to follow. Let's see an example.

18.1.6 Pitfall: Computing Expectations by Sampling

Suppose that you are trying to estimate a parameter such as the average delay across a communication channel. So you set up an experiment to measure how long it takes to send a test packet from one end to the other and you run the experiment 100 times.

You record the latency, rounded to the nearest millisecond, for each of the hundred experiments, and then compute the average of the 100 measurements. Suppose that this average is 8.3 ms.

Because you are careful, you repeat the entire process twice more and get averages of 7.8 ms and 7.9 ms. You conclude that the average latency across the channel

[2] If this still bothers you, you might consider taking a course on measure theory.

is

$$\frac{7.8 + 7.9 + 8.3}{3} = 8\,\text{ms}.$$

You might be right but you might also be horribly wrong. In fact, the expected latency might well be *infinite*. Here's how.

Let D be a random variable that denotes the time it takes for the packet to cross the channel. Suppose that

$$\Pr[D = i] = \begin{cases} 0 & \text{for } i = 0 \\ \frac{1}{i} - \frac{1}{i+1} & \text{for } i \in \mathbb{N}^+. \end{cases} \tag{18.5}$$

It is easy to check that

$$\sum_{i=0}^{\infty} \Pr[D = i] = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots = 1$$

and so D is, in fact, a random variable.

From Equation 18.5, we might expect that D is likely to be small. Indeed, $D = 1$ with probability $1/2$, $D = 2$ with probability $1/6$, and so forth. So if we took 100 samples of D, about 50 would be 1 ms, about 16 would be 2 ms, and very few would be large. In summary, it might well be the case that the average of the 100 measurements would be under 10 ms, just as in our example.

This sort of reasoning and the calculation of expected values by averaging experimental values is very common in practice. It can easily lead to incorrect conclusions, however. For example, using Corollary 18.1.5, we can quickly (and accurately) determine that

$$\begin{aligned} \text{Ex}[D] &= \sum_{i=1}^{\infty} i \Pr[D = i] \\ &= \sum_{i=1}^{\infty} i \left(\frac{1}{i} - \frac{1}{i+1}\right) \\ &= \sum_{i=1}^{\infty} i \left(\frac{1}{i(i+1)}\right) \\ &= \sum_{i=1}^{\infty} \left(\frac{1}{i+1}\right) \\ &= \infty. \end{aligned}$$

Uh-oh! The expected time to cross the communication channel is *infinite*! This result is a far cry from the 10 ms that we calculated. What went wrong?

It is true that most of the time, the value of D will be small. But sometimes D will be very large and this happens with sufficient probability that the expected value of D is unbounded. In fact, if you keep repeating the experiment, you are likely to see some outcomes and averages that are much larger than 10 ms. In practice, such "outliers" are sometimes discarded, which masks the true behavior of D.

In general, the best way to compute an expected value in practice is to first use the experimental data to figure out the distribution as best you can, and then to use Theorem 18.1.4 or Corollary 18.1.5 to compute its expectation. This method will help you identify cases where the expectation is infinite, and will generally be more accurate than a simple averaging of the data.

18.1.7 Conditional Expectation

Just like event probabilities, expectations can be conditioned on some event. Given a random variable R, the expected value of R conditioned on an event A is the (probability-weighted) average value of R over outcomes in A. More formally:

Definition 18.1.6. The *conditional expectation* $\mathrm{Ex}[R \mid A]$ of a random variable R given event A is:

$$\mathrm{Ex}[R \mid A] ::= \sum_{r \in \mathrm{range}(R)} r \cdot \Pr\left[R = r \mid A\right]. \tag{18.6}$$

For example, we can compute the expected value of a roll of a fair die, *given*, for example, that the number rolled is at least 4. We do this by letting R be the outcome of a roll of the die. Then by equation (18.6),

$$\mathrm{Ex}[R \mid R \geq 4] = \sum_{i=1}^{6} i \cdot \Pr\left[R = i \mid R \geq 4\right] = 1{\cdot}0 + 2{\cdot}0 + 3{\cdot}0 + 4{\cdot}\tfrac{1}{3} + 5{\cdot}\tfrac{1}{3} + 6{\cdot}\tfrac{1}{3} = 5.$$

As another example, consider the channel latency problem from Section 18.1.6. The expected latency for this problem was infinite. But what if we look at the

expected latency conditioned on the latency not exceeding n. Then

$$\begin{aligned}
\mathrm{Ex}[D] &= \sum_{i=1}^{\infty} i \Pr\left[D = i \mid D \le n\right] \\
&= \sum_{i=1}^{\infty} i \frac{\Pr[D = i \wedge D \le n]}{\Pr[D \le n]} \\
&= \sum_{i=1}^{n} \frac{i \Pr[D = i]}{\Pr[D \le n]} \\
&= \frac{1}{\Pr[D \le n]} \sum_{i=1}^{n} i \left(\frac{1}{i(i+1)}\right) \\
&= \frac{1}{\Pr[D \le n]} \sum_{i=1}^{n} \frac{1}{i+1} \\
&= \frac{1}{\Pr[D \le n]} (H_{n+1} - 1),
\end{aligned}$$

where H_{n+1} is the $(n+1)$st Harmonic number

$$H_{n+1} = \ln(n+1) + \gamma + \frac{1}{2n} + \frac{1}{12n^2} + \frac{\epsilon(n)}{120n^4}$$

and $0 \le \epsilon(n) \le 1$. The second equality follows from the definition of conditional expectation, the third equality follows from the fact that $\Pr[D = i \wedge D \le n] = 0$ for $i > n$, and the fourth equality follows from the definition of D in Equation 18.5.

To compute $\Pr[D \le n]$, we observe that

$$\begin{aligned}
\Pr[D \le n] &= 1 - \Pr[D > n] \\
&= 1 - \sum_{i=n+1}^{\infty} \left(\frac{1}{i} - \frac{1}{i+1}\right) \\
&= 1 - \left[\left(\frac{1}{n+1} - \frac{1}{n+2}\right) + \left(\frac{1}{n+2} - \frac{1}{n+3}\right) \right. \\
&\qquad \left. + \left(\frac{1}{n+3} - \frac{1}{n+4}\right) + \cdots\right] \\
&= 1 - \frac{1}{n+1} \\
&= \frac{n}{n+1}.
\end{aligned}$$

Hence,

$$\mathrm{Ex}[D] = \frac{n+1}{n}(H_{n+1} - 1). \tag{18.7}$$

For $n = 1000$, this is about 6.5. This explains why the expected value of D appears to be finite when you try to evaluate it experimentally. If you compute 100 samples of D, it is likely that all of them will be at most 1000 ms. If you condition on not having any outcomes greater than 1000 ms, then the conditional expected value will be about 6.5 ms, which would be a commonly observed result in practice. Yet we know that $\mathrm{Ex}[D]$ is infinite. For this reason, expectations computed in practice are often really just conditional expectations where the condition is that rare "outlier" sample points are eliminated from the analysis.

18.1.8 The Law of Total Expectation

Another useful feature of conditional expectation is that it lets us divide complicated expectation calculations into simpler cases. We can then find the desired expectation by calculating the conditional expectation in each simple case and averaging them, weighing each case by its probability.

For example, suppose that 49.8% of the people in the world are male and the rest female—which is more or less true. Also suppose the expected height of a randomly chosen male is $5' 11''$, while the expected height of a randomly chosen female is $5' 5''$. What is the expected height of a randomly chosen individual? We can calculate this by averaging the heights of men and women. Namely, let H be the height (in feet) of a randomly chosen person, and let M be the event that the person is male and F the event that the person is female. Then

$$\begin{aligned}\mathrm{Ex}[H] &= \mathrm{Ex}[H \mid M]\Pr[M] + \mathrm{Ex}[H \mid F]\Pr[F] \\ &= (5 + 11/12) \cdot 0.498 + (5 + 5/12) \cdot 0.502 \\ &= 5.665\end{aligned}$$

which is a little less than 5' 8".

This method is justified by the Law of *Total Expectation*.

Theorem 18.1.7 (Law of Total Expectation)**.** *Let R be a random variable on a sample space $\mathcal{S}$ and suppose that $A_1, A_2, \ldots,$ is a partition of $\mathcal{S}$. Then*

$$\mathrm{Ex}[R] = \sum_i \mathrm{Ex}[R \mid A_i]\Pr[A_i].$$

Proof.

$$\begin{aligned}
\mathrm{Ex}[R] &= \sum_{r \in \mathrm{range}(R)} r \cdot \Pr[R = r] && \text{(Equation 18.2)} \\
&= \sum_r r \cdot \sum_i \Pr\left[R = r \mid A_i\right] \Pr[A_i] && \text{(Law of Total Probability)} \\
&= \sum_r \sum_i r \cdot \Pr\left[R = r \mid A_i\right] \Pr[A_i] && \text{(distribute constant } r\text{)} \\
&= \sum_i \sum_r r \cdot \Pr\left[R = r \mid A_i\right] \Pr[A_i] && \text{(exchange order of summation)} \\
&= \sum_i \Pr[A_i] \sum_r r \cdot \Pr\left[R = r \mid A_i\right] && \text{(factor constant } \Pr[A_i]\text{)} \\
&= \sum_i \Pr[A_i] \, \mathrm{Ex}[R \mid A_i]. && \text{(Def 18.1.6 of cond. expectation)}
\end{aligned}$$

■

As a more interesting application of the Law of Total Expectation, let's take another look at the mean time to failure of a system that fails with probability p at each step. We'll define A to be the event that the system fails on the first step and $\overline{A}$ to be the complementary event (namely, that the system does not fail on the first step). Then the mean time to failure $\mathrm{Ex}[C]$ is

$$\mathrm{Ex}[C] = \mathrm{Ex}[C \mid A] \Pr[A] + \mathrm{Ex}[C \mid \overline{A}] \Pr[\overline{A}]. \tag{18.8}$$

Since A is the condition that the system crashes on the first step, we know that

$$\mathrm{Ex}[C \mid A] = 1. \tag{18.9}$$

Since $\overline{A}$ is the condition that the system does *not* crash on the first step, conditioning on $\overline{A}$ is equivalent to taking a first step without failure and then starting over without conditioning. Hence,

$$\mathrm{Ex}[C \mid \overline{A}] = 1 + \mathrm{Ex}[C]. \tag{18.10}$$

Plugging Equations 18.9 and 18.10 into Equation 18.8, we find that

$$\begin{aligned}
\mathrm{Ex}[C] &= 1 \cdot p + (1 + \mathrm{Ex}[C])(1 - p) \\
&= p + 1 - p + (1 - p)\,\mathrm{Ex}[C] \\
&= 1 + (1 - p)\,\mathrm{Ex}[C].
\end{aligned}$$

Rearranging terms, we find that

$$1 = \mathrm{Ex}[C] - (1-p)\,\mathrm{Ex}[C] = p\,\mathrm{Ex}[C],$$

and thus that

$$\mathrm{Ex}[C] = \frac{1}{p},$$

as expected.

We will use this sort of analysis extensively in Chapter 20 when we examine the expected behavior of random walks.

18.1.9 Expectations of Functions

Expectations can also be defined for functions of random variables.

Definition 18.1.8. Let $R : \mathcal{S} \to V$ be a random variable and $f : V \to \mathbb{R}$ be a total function on the range of R. Then

$$\mathrm{Ex}[f(R)] = \sum_{w \in \mathcal{S}} f(R(w))\,\Pr[w]. \tag{18.11}$$

Equivalently,

$$\mathrm{Ex}[f(R)] = \sum_{r \in \mathrm{range}(R)} f(r)\,\Pr[R = r]. \tag{18.12}$$

For example, suppose that R is the value obtained by rolling a fair 6-sided die. Then

$$\mathrm{Ex}\left[\frac{1}{R}\right] = \frac{1}{1}\cdot\frac{1}{6} + \frac{1}{2}\cdot\frac{1}{6} + \frac{1}{3}\cdot\frac{1}{6} + \frac{1}{4}\cdot\frac{1}{6} + \frac{1}{5}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{49}{120}.$$

18.2 Expected Returns in Gambling Games

Some of the most interesting examples of expectation can be explained in terms of gambling games. For straightforward games where you win $\$A$ with probability p and you lose $\$B$ with probability $1-p$, it is easy to compute your *expected return* or *winnings*. It is simply

$$pA - (1-p)B.$$

For example, if you are flipping a fair coin and you win \$1 for heads and you lose \$1 for tails, then your expected winnings are

$$\frac{1}{2}\cdot 1 - \left(1 - \frac{1}{2}\right)\cdot 1 = 0.$$

In such cases, the game is said to be *fair* since your expected return is zero.

Some gambling games are more complicated and thus more interesting. For example, consider the following game where the winners split a pot. This sort of game is representative of many poker games, betting pools, and lotteries.

18.2.1 Splitting the Pot

After your last encounter with biker dude, one thing lead to another and you have dropped out of school and become a Hell's Angel. It's late on a Friday night and, feeling nostalgic for the old days, you drop by your old hangout, where you encounter two of your former TAs, Eric and Nick. Eric and Nick propose that you join them in a simple wager. Each player will put $2 on the bar and secretly write "heads" or "tails" on their napkin. Then one player will flip a fair coin. The $6 on the bar will then be divided equally among the players who correctly predicted the outcome of the coin toss.

After your life-altering encounter with strange dice, you are more than a little skeptical. So Eric and Nick agree to let you be the one to flip the coin. This certainly seems fair. How can you lose?

But you have learned your lesson and so before agreeing, you go through the four-step method and write out the tree diagram to compute your expected return. The tree diagram is shown in Figure 18.1.

The "payoff" values in Figure 18.1 are computed by dividing the $6 pot[3] among those players who guessed correctly and then subtracting the $2 that you put into the pot at the beginning. For example, if all three players guessed correctly, then you payoff is $0, since you just get back your $2 wager. If you and Nick guess correctly and Eric guessed wrong, then your payoff is

$$\frac{6}{2} - 2 = 1.$$

In the case that everyone is wrong, you all agree to split the pot and so, again, your payoff is zero.

To compute your expected return, you use Equation 18.1 in the definition of expected value. This yields

$$\begin{aligned} \mathrm{Ex}[\text{payoff}] &= 0 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 4 \cdot \frac{1}{8} \\ &\quad + (-2) \cdot \frac{1}{8} + (-2) \cdot \frac{1}{8} + (-2) \cdot \frac{1}{8} + 0 \cdot \frac{1}{8} \\ &= 0. \end{aligned}$$

[3] The money invested in a wager is commonly referred to as the *pot*.

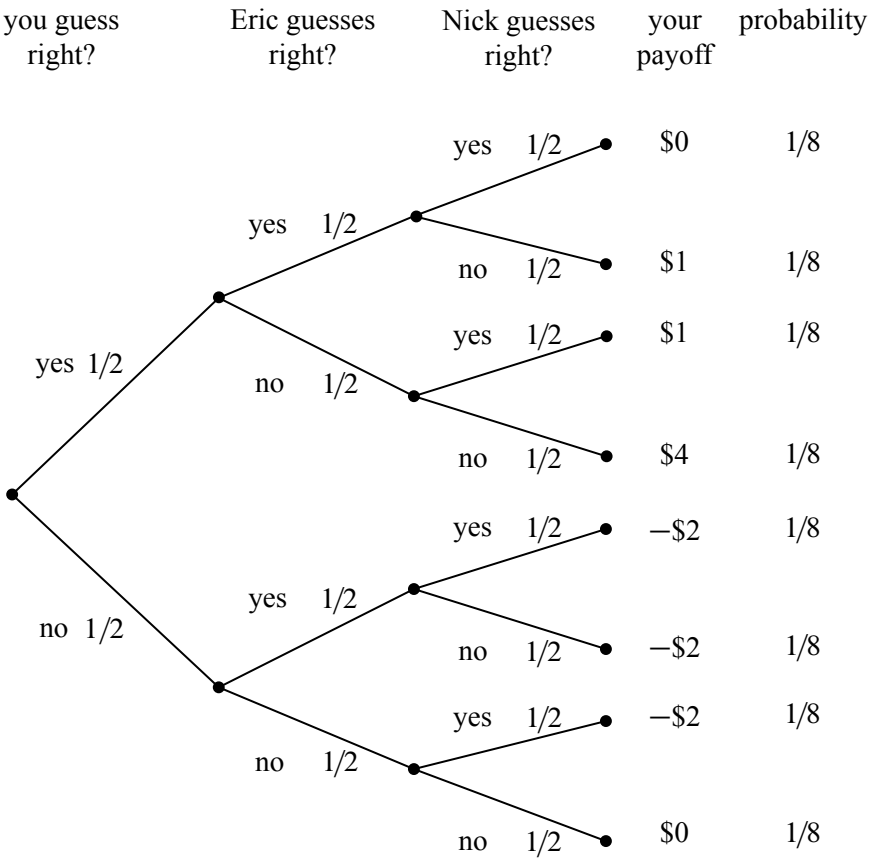


Figure 18.1 The tree diagram for the game where three players each wager \$2 and then guess the outcome of a fair coin toss. The winners split the pot.

This confirms that the game is fair. So, for old time's sake, you break your solemn vow to never ever engage in strange gambling games.

18.2.2 The Impact of Collusion

Needless to say, things are not turning out well for you. The more times you play the game, the more money you seem to be losing. After 1000 wagers, you have lost over $500. As Nick and Eric are consoling you on your "bad luck," you do a back-of-the-napkin calculation using the bounds on the tails of the binomial distribution from Section 17.5 that suggests that the probability of losing $500 in 1000 wagers is less than the probability of a Vietnamese Monk waltzing in and handing you one of those golden disks. How can this be?

It is possible that you are truly very very unlucky. But it is more likely that something is wrong with the tree diagram in Figure 18.1 and that "something" just might have something to do with the possibility that Nick and Eric are colluding against you.

To be sure, Nick and Eric can only guess the outcome of the coin toss with probability $1/2$, but what if Nick and Eric always guess differently? In other words, what if Nick always guesses "tails" when Eric guesses "heads," and vice-versa? This would result in a slightly different tree diagram, as shown in Figure 18.2.

The payoffs for each outcome are the same in Figures 18.1 and 18.2, but the probabilities of the outcomes are different. For example, it is no longer possible for all three players to guess correctly, since Nick and Eric are always guessing differently. More importantly, the outcome where your payoff is $4 is also no longer possible. Since Nick and Eric are always guessing differently, one of them will always get a share of the pot. As you might imagine, this is not good for you!

When we use Equation 18.1 to compute your expected return in the collusion scenario, we find that

$$\begin{aligned}\text{Ex}[\text{payoff}] &= 0\cdot 0 + 1\cdot\frac{1}{4} + 1\cdot\frac{1}{4} + 4\cdot 0 \\ &\qquad + (-2)\cdot 0 + (-2)\cdot\frac{1}{4} + (-2)\cdot\frac{1}{4} + 0\cdot 0 \\ &= -\frac{1}{2}.\end{aligned}$$

This is very bad indeed. By colluding, Nick and Eric have made it so that you expect to lose $.50 every time you play. No wonder you lost $500 over the course of 1000 wagers.

Maybe it would be a good idea to go back to school—your Hell's Angels buds may not be too happy that you just lost their $500.

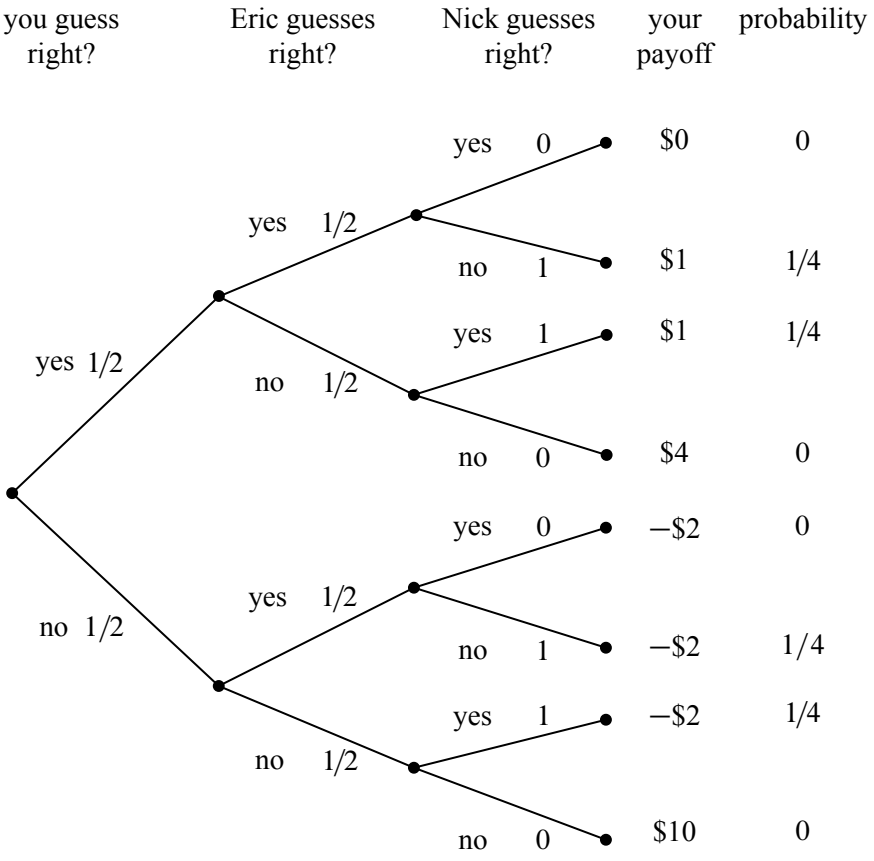


Figure 18.2 The revised tree diagram reflecting the scenario where Nick always guesses the opposite of Eric.

18.2.3 How to Win the Lottery

Similar opportunities to "collude" arise in many betting games. For example, consider the typical weekly football betting pool, where each participant wagers $10 and the participants that pick the most games correctly split a large pot. The pool seems fair if you think of it as in Figure 18.1. But, in fact, if two or more players collude by guessing differently, they can get an "unfair" advantage at your expense!

In some cases, the collusion is inadvertent and you can profit from it. For example, many years ago, a former MIT Professor of Mathematics named Herman Chernoff figured out a way to make money by playing the state lottery. This was surprising since state lotteries typically have very poor expected returns. That's because the state usually takes a large share of the wagers before distributing the rest of the pot among the winners. Hence, anyone who buys a lottery ticket is expected to *lose* money. So how did Chernoff find a way to make money? It turned out to be easy!

In a typical state lottery,

- all players pay $1 to play and select 4 numbers from 1 to 36,
- the state draws 4 numbers from 1 to 36 uniformly at random,
- the states divides 1/2 of the money collected among the people who guessed correctly and spends the other half redecorating the governor's residence.

This is a lot like the game you played with Nick and Eric, except that there are more players and more choices. Chernoff discovered that a small set of numbers was selected by a large fraction of the population. Apparently many people think the same way; they pick the same numbers not on purpose as in the previous game with Nick and Eric, but based on Manny's batting average or today's date.

It was as if the players were colluding to lose! If any one of them guessed correctly, then they'd have to split the pot with many other players. By selecting numbers uniformly at random, Chernoff was unlikely to get one of these favored sequences. So if he won, he'd likely get the whole pot! By analyzing actual state lottery data, he determined that he could win an average of 7 cents on the dollar. In other words, his expected return was not $-\$.50$ as you might think, but $+\$.07$.[4]

Inadvertent collusion often arises in betting pools and is a phenomenon that you can take advantage of. For example, suppose you enter a Super Bowl betting pool where the goal is to get closest to the total number of points scored in the game. Also suppose that the average Super Bowl has a total of 30 point scored and that

[4]Most lotteries now offer randomized tickets to help smooth out the distribution of selected sequences.

everyone knows this. Then most people will guess around 30 points. Where should you guess? Well, you should guess just outside of this range because you get to cover a lot more ground and you don't share the pot if you win. Of course, if you are in a pool with math students and they all know this strategy, then maybe you should guess 30 points after all.

18.3 Expectations of Sums

18.3.1 Linearity of Expectation

Expected values obey a simple, very helpful rule called *Linearity of Expectation*. Its simplest form says that the expected value of a sum of random variables is the sum of the expected values of the variables.

Theorem 18.3.1. *For any random variables R_1 and R_2,*

$$\mathrm{Ex}[R_1 + R_2] = \mathrm{Ex}[R_1] + \mathrm{Ex}[R_2].$$

Proof. Let $T ::= R_1 + R_2$. The proof follows straightforwardly by rearranging terms in Equation (18.1):

$$\begin{aligned}
\mathrm{Ex}[T] &= \sum_{\omega\in S} T(\omega)\cdot \Pr[\omega] && \text{(Definition 18.1.1)}\\
&= \sum_{\omega\in S} (R_1(\omega) + R_2(\omega))\cdot \Pr[\omega] && \text{(definition of } T\text{)}\\
&= \sum_{\omega\in S} R_1(\omega)\Pr[\omega] + \sum_{\omega\in S} R_2(\omega)\Pr[\omega] && \text{(rearranging terms)}\\
&= \mathrm{Ex}[R_1] + \mathrm{Ex}[R_2]. && \text{(Definition 18.1.1)}
\end{aligned}$$

■

A small extension of this proof, which we leave to the reader, implies

Theorem 18.3.2. *For random variables R_1, R_2 and constants $a_1, a_2 \in \mathbb{R}$,*

$$\mathrm{Ex}[a_1 R_1 + a_2 R_2] = a_1\,\mathrm{Ex}[R_1] + a_2\,\mathrm{Ex}[R_2].$$

In other words, expectation is a linear function. A routine induction extends the result to more than two variables:

Corollary 18.3.3 (Linearity of Expectation)**.** *For any random variables $R_1, \ldots, R_k$ and constants $a_1, \ldots, a_k \in \mathbb{R}$,*

$$\mathrm{Ex}[\sum_{i=1}^{k} a_i R_i] = \sum_{i=1}^{k} a_i\,\mathrm{Ex}[R_i].$$

The great thing about linearity of expectation is that *no independence is required*. This is really useful, because dealing with independence is a pain, and we often need to work with random variables that are not known to be independent.

As an example, let's compute the expected value of the sum of two fair dice. Let the random variable R_1 be the number on the first die, and let R_2 be the number on the second die. We observed earlier that the expected value of one die is 3.5. We can find the expected value of the sum using linearity of expectation:

$$\mathrm{Ex}[R_1 + R_2] = \mathrm{Ex}[R_1] + \mathrm{Ex}[R_2] = 3.5 + 3.5 = 7.$$

Notice that we did *not* have to assume that the two dice were independent. The expected sum of two dice is 7, even if they are glued together (provided each individual die remains fair after the gluing). Proving that this expected sum is 7 with a tree diagram would be a bother: there are 36 cases. And if we did not assume that the dice were independent, the job would be really tough!

18.3.2 Sums of Indicator Random Variables

Linearity of expectation is especially useful when you have a sum of indicator random variables. As an example, suppose there is a dinner party where n men check their hats. The hats are mixed up during dinner, so that afterward each man receives a random hat. In particular, each man gets his own hat with probability $1/n$. What is the expected number of men who get their own hat?

Letting G be the number of men that get their own hat, we want to find the expectation of G. But all we know about G is that the probability that a man gets his own hat back is $1/n$. There are many different probability distributions of hat permutations with this property, so we don't know enough about the distribution of G to calculate its expectation directly. But linearity of expectation makes the problem really easy.

The trick[5] is to express G as a sum of indicator variables. In particular, let G_i be an indicator for the event that the ith man gets his own hat. That is, $G_i = 1$ if the ith man gets his own hat, and $G_i = 0$ otherwise. The number of men that get their own hat is then the sum of these indicator random variables:

$$G = G_1 + G_2 + \cdots + G_n. \tag{18.13}$$

These indicator variables are *not* mutually independent. For example, if $n - 1$ men all get their own hats, then the last man is certain to receive his own hat. But, since we plan to use linearity of expectation, we don't have worry about independence!

[5] We are going to use this trick a lot so it is important to understand it.

Since G_i is an indicator random variable, we know from Lemma 18.1.3 that

$$\mathrm{Ex}[G_i] = \Pr[G_i = 1] = 1/n. \tag{18.14}$$

By Linearity of Expectation and Equation 18.13, this means that

$$\begin{aligned}
\mathrm{Ex}[G] &= \mathrm{Ex}[G_1 + G_2 + \cdots + G_n] \\
&= \mathrm{Ex}[G_1] + \mathrm{Ex}[G_2] + \cdots + \mathrm{Ex}[G_n] \\
&= \overbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}^{n} \\
&= 1.
\end{aligned}$$

So even though we don't know much about how hats are scrambled, we've figured out that on average, just one man gets his own hat back!

More generally, Linearity of Expectation provides a very good method for computing the expected number of events that will happen.

Theorem 18.3.4. *Given any collection of n events $A_1, A_2, \ldots, A_n \subseteq \mathcal{S}$, the expected number of events that will occur is*

$$\sum_{i=1}^{n} \Pr[A_i].$$

For example, A_i could be the event that the ith man gets the right hat back. But in general, it could be any subset of the sample space, and we are asking for the expected number of events that will contain a random sample point.

Proof. Define R_i to be the indicator random variable for A_i, where $R_i(w) = 1$ if $w \in A_i$ and $R_i(w) = 0$ if $w \notin A_i$. Let $R = R_1 + R_2 + \cdots + R_n$. Then

$$\begin{aligned}
\mathrm{Ex}[R] &= \sum_{i=1}^{n} \mathrm{Ex}[R_i] && \text{(by Linearity of Expectation)} \\
&= \sum_{i=1}^{n} \Pr[R_i = 1] && \text{(by Lemma 18.1.3)} \\
&= \sum_{i=1}^{n} \sum_{w \in A_i} \Pr[w] && \text{(definition of indicator variable)} \\
&= \sum_{i=1}^{n} \Pr[A_i]. && \blacksquare
\end{aligned}$$

So whenever you are asked for the expected number of events that occur, all you have to do is sum the probabilities that each event occurs. Independence is not needed.

18.3.3 Expectation of a Binomial Distribution

Suppose that we independently flip n biased coins, each with probability p of coming up heads. What is the expected number of heads?

Let J be the random variable denoting the number of heads. Then J has a binomial distribution with parameters n, p, and

$$\Pr[J = k] = \binom{n}{k} k^p (n-k)^{1-p}.$$

Applying Equation 18.2, this means that

$$\begin{aligned} \operatorname{Ex}[J] &= \sum_{k=0}^{n} k \Pr[J = k] \\ &= \sum_{k=0}^{n} k \binom{n}{k} k^p (n-k)^{1-p}. \end{aligned} \tag{18.15}$$

Ouch! This is one nasty looking sum. Let's try another approach.

Since we have just learned about linearity of expectation for sums of indicator random variables, maybe Theorem 18.3.4 will be helpful. But how do we express J as a sum of indicator random variables? It turns out to be easy. Let J_i be the indicator random variable for the ith coin. In particular, define

$$J_i = \begin{cases} 1 & \text{if the } i\text{th coin is heads} \\ 0 & \text{if the } i\text{th coin is tails.} \end{cases}$$

Then the number of heads is simply

$$J = J_1 + J_2 + \cdots + J_n.$$

By Theorem 18.3.4,

$$\begin{aligned} \operatorname{Ex}[J] &= \sum_{i=1}^{n} \Pr[J_i] \\ &= np. \end{aligned} \tag{18.16}$$

That really was easy. If we flip n mutually independent coins, we expect to get pn heads. Hence the expected value of a binomial distribution with parameters n and p is simply pn.

But what if the coins are not mutually independent? It doesn't matter—the answer is still pn because Linearity of Expectation and Theorem 18.3.4 do not assume any independence.

If you are not yet convinced that Linearity of Expectation and Theorem 18.3.4 are powerful tools, consider this: without even trying, we have used them to prove a very complicated identity, namely[6]

$$\sum_{k=0}^{n} k \binom{n}{k} k^p (n-k)^{1-p} = pn.$$

If you are still not convinced, then take a look at the next problem.

18.3.4 The Coupon Collector Problem

Every time we purchase a kid's meal at Taco Bell, we are graciously presented with a miniature "Racin' Rocket" car together with a launching device which enables us to project our new vehicle across any tabletop or smooth floor at high velocity. Truly, our delight knows no bounds.

There are n different types of Racin' Rocket cars (blue, green, red, gray, etc.). The type of car awarded to us each day by the kind woman at the Taco Bell register appears to be selected uniformly and independently at random. What is the expected number of kid's meals that we must purchase in order to acquire at least one of each type of Racin' Rocket car?

The same mathematical question shows up in many guises: for example, what is the expected number of people you must poll in order to find at least one person with each possible birthday? Here, instead of collecting Racin' Rocket cars, you're collecting birthdays. The general question is commonly called the *coupon collector problem* after yet another interpretation.

A clever application of linearity of expectation leads to a simple solution to the coupon collector problem. Suppose there are five different types of Racin' Rocket cars, and we receive this sequence:

blue green green red blue orange blue orange gray.

Let's partition the sequence into 5 segments:

$$\underbrace{\text{blue}}_{X_0} \quad \underbrace{\text{green}}_{X_1} \quad \underbrace{\text{green red}}_{X_2} \quad \underbrace{\text{blue orange}}_{X_3} \quad \underbrace{\text{blue orange gray}}_{X_4}.$$

[6] This follows by combining Equations 18.15 and 18.16.

The rule is that a segment ends whenever we get a new kind of car. For example, the middle segment ends when we get a red car for the first time. In this way, we can break the problem of collecting every type of car into stages. Then we can analyze each stage individually and assemble the results using linearity of expectation.

Let's return to the general case where we're collecting n Racin' Rockets. Let X_k be the length of the kth segment. The total number of kid's meals we must purchase to get all n Racin' Rockets is the sum of the lengths of all these segments:

$$T = X_0 + X_1 + \cdots + X_{n-1}$$

Now let's focus our attention on X_k, the length of the kth segment. At the beginning of segment k, we have k different types of car, and the segment ends when we acquire a new type. When we own k types, each kid's meal contains a type that we already have with probability k/n. Therefore, each meal contains a new type of car with probability $1-k/n = (n-k)/n$. Thus, the expected number of meals until we get a new kind of car is $n/(n-k)$ by the "mean time to failure" formula in Equation 18.4. This means that

$$\mathrm{Ex}[X_k] = \frac{n}{n-k}.$$

Linearity of expectation, together with this observation, solves the coupon collector problem:

$$\begin{aligned}
\mathrm{Ex}[T] &= \mathrm{Ex}[X_0 + X_1 + \cdots + X_{n-1}] \\
&= \mathrm{Ex}[X_0] + \mathrm{Ex}[X_1] + \cdots + \mathrm{Ex}[X_{n-1}] \\
&= \frac{n}{n-0} + \frac{n}{n-1} + \cdots + \frac{n}{3} + \frac{n}{2} + \frac{n}{1} \\
&= n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{3} + \frac{1}{2} + \frac{1}{1}\right) \\
&= n\left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-1} + \frac{1}{n}\right) \\
&= nH_n && (18.17) \\
&\sim n\ln n. && (18.18)
\end{aligned}$$

Wow! It's those Harmonic Numbers again!

We can use Equation 18.18 to answer some concrete questions. For example, the expected number of die rolls required to see every number from 1 to 6 is:

$$6H_6 = 14.7\ldots.$$

And the expected number of people you must poll to find at least one person with each possible birthday is:

$$365H_{365} = 2364.6\ldots.$$

18.3.5 Infinite Sums

Linearity of expectation also works for an infinite number of random variables provided that the variables satisfy some stringent absolute convergence criteria.

Theorem 18.3.5 (Linearity of Expectation)**.** *Let $R_0, R_1, \ldots,$ be random variables such that*

$$\sum_{i=0}^{\infty} \mathrm{Ex}[|R_i|]$$

converges. Then

$$\mathrm{Ex}\left[\sum_{i=0}^{\infty} R_i\right] = \sum_{i=0}^{\infty} \mathrm{Ex}[R_i].$$

Proof. Let $T ::= \sum_{i=0}^{\infty} R_i$.

We leave it to the reader to verify that, under the given convergence hypothesis, all the sums in the following derivation are absolutely convergent, which justifies rearranging them as follows:

$$\begin{aligned}
\sum_{i=0}^{\infty} \mathrm{Ex}[R_i] &= \sum_{i=0}^{\infty} \sum_{s\in S} R_i(s)\cdot \Pr[s] && \text{(Def. 18.1.1)}\\
&= \sum_{s\in S} \sum_{i=0}^{\infty} R_i(s)\cdot \Pr[s] && \text{(exchanging order of summation)}\\
&= \sum_{s\in S} \left[\sum_{i=0}^{\infty} R_i(s)\right]\cdot \Pr[s] && \text{(factoring out } \Pr[s])\\
&= \sum_{s\in S} T(s)\cdot \Pr[s] && \text{(Def. of } T)\\
&= \mathrm{Ex}[T] && \text{(Def. 18.1.1)}\\
&= \mathrm{Ex}[\sum_{i=0}^{\infty} R_i]. && \text{(Def. of } T). \ \blacksquare
\end{aligned}$$

18.4 Expectations of Products

While the expectation of a sum is the sum of the expectations, the same is usually not true for products. For example, suppose that we roll a fair 6-sided die and denote the outcome with the random variable R. Does $\text{Ex}[R \cdot R] = \text{Ex}[R] \cdot \text{Ex}[R]$?

We know that $\text{Ex}[R] = 3\frac{1}{2}$ and thus $\text{Ex}[R]^2 = 12\frac{1}{4}$. Let's compute $\text{Ex}[R^2]$ to see if we get the same result.

$$\begin{aligned}
\text{Ex}[R^2] &= \sum_{w \in S} R^2(w) \Pr[w] \\
&= \sum_{i=1}^{6} i^2 \cdot \Pr[R_i = i] \\
&= \frac{1^2}{6} + \frac{2^2}{6} + \frac{3^2}{6} + \frac{4^2}{6} + \frac{5^2}{6} + \frac{6^2}{6} \\
&= 15\ 1/6 \\
&\neq 12\ 1/4.
\end{aligned}$$

Hence,

$$\text{Ex}[R \cdot R] \neq \text{Ex}[R] \cdot \text{Ex}[R]$$

and so the expectation of a product is not always equal to the product of the expectations.

There is a special case when such a relationship *does* hold however; namely, when the random variables in the product are *independent*.

Theorem 18.4.1. *For any two* independent *random variables R_1, R_2,*

$$\text{Ex}[R_1 \cdot R_2] = \text{Ex}[R_1] \cdot \text{Ex}[R_2].$$

Proof. The event $[R_1 \cdot R_2 = r]$ can be split up into events of the form $[R_1 =$

r_1 and $R_2 = r_2]$ where $r_1 \cdot r_2 = r$. So

$$\begin{aligned}
&\mathrm{Ex}[R_1 \cdot R_2] \\
&= \sum_{r \in \mathrm{range}(R_1 \cdot R_2)} r \cdot \Pr[R_1 \cdot R_2 = r] && \text{(Theorem 18.1.4)} \\
&= \sum_{r_1 \in \mathrm{range}(R_1)} \sum_{r_2 \in \mathrm{range}(R_2)} r_1 r_2 \cdot \Pr[R_1 = r_1 \text{ and } R_2 = r_2] \\
&= \sum_{r_1 \in \mathrm{range}(R_1)} \sum_{r_2 \in \mathrm{range}(R_2)} r_1 r_2 \cdot \Pr[R_1 = r_1] \cdot \Pr[R_2 = r_2] && \text{(independence of } R_1, R_2) \\
&= \sum_{r_1 \in \mathrm{range}(R_1)} r_1 \Pr[R_1 = r_1] \left(\sum_{r_2 \in \mathrm{range}(R_2)} r_2 \Pr[R_2 = r_2] \right) && \text{(factor out } r_1 \Pr[R_1 = r_1]) \\
&= \sum_{r_1 \in \mathrm{range}(R_1)} r_1 \Pr[R_1 = r_1] \cdot \mathrm{Ex}[R_2] && \text{(Theorem 18.1.4)} \\
&= \mathrm{Ex}[R_2] \left(\sum_{r_1 \in \mathrm{range}(R_1)} r_1 \Pr[R_1 = r_1] \right) && \text{(factor out } \mathrm{Ex}[R_2]) \\
&= \mathrm{Ex}[R_2] \cdot \mathrm{Ex}[R_1]. && \text{(Theorem 18.1.4)}
\end{aligned}$$

■

For example, let R_1 and R_2 be random variables denoting the result of rolling two independent and fair 6-sided dice. Then

$$\mathrm{Ex}[R_1 \cdot R_2] = \mathrm{Ex}[R_1]\,\mathrm{Ex}[R_2] = 3\frac{1}{2} \cdot 3\frac{1}{2} = 12\frac{1}{4}.$$

Theorem 18.4.1 extends by induction to a collection of mutually independent random variables.

Corollary 18.4.2. *If random variables $R_1, R_2, \ldots, R_k$ are mutually independent, then*

$$\mathrm{Ex}\left[\prod_{i=1}^{k} R_i\right] = \prod_{i=1}^{k} \mathrm{Ex}[R_i].$$

18.5 Expectations of Quotients

If S and T are random variables, we know from Linearity of Expectation that

$$\mathrm{Ex}[S+T] = \mathrm{Ex}[S] + \mathrm{Ex}[T].$$

If S and T are independent, we know from Theorem 18.4.1 that

$$\mathrm{Ex}[ST] = \mathrm{Ex}[S]\,\mathrm{Ex}[T].$$

Is it also true that

$$\mathrm{Ex}[S/T] = \mathrm{Ex}[S]/\mathrm{Ex}[T]? \tag{18.19}$$

Of course, we have to worry about the situation when $\mathrm{Ex}[T] = 0$, but what if we assume that T is always positive? As we will soon see, Equation 18.19 is usually not true, but let's see if we can prove it anyway.

False Claim 18.5.1. *If S and T are independent random variables with $T > 0$, then*

$$\mathrm{Ex}[S/T] = \mathrm{Ex}[S]/\mathrm{Ex}[T]. \tag{18.20}$$

Bogus proof.

$$\begin{aligned}
\mathrm{Ex}\left[\frac{S}{T}\right] &= \mathrm{Ex}\left[S \cdot \frac{1}{T}\right] \\
&= \mathrm{Ex}[S] \cdot \mathrm{Ex}\left[\frac{1}{T}\right] \qquad \text{(independence of } S \text{ and } T\text{)} \qquad (18.21)\\
&= \mathrm{Ex}[S] \cdot \frac{1}{\mathrm{Ex}[T]}. \qquad (18.22)\\
&= \frac{\mathrm{Ex}[S]}{\mathrm{Ex}[T]}. \qquad \blacksquare
\end{aligned}$$

Note that line 18.21 uses the fact that if S and T are independent, then so are S and $1/T$. This holds because functions of independent random variables are independent. It is a fact that needs proof, which we will leave to the reader, but it is not the bug. The bug is in line (18.22), which assumes

False Claim 18.5.2.

$$\mathrm{Ex}\left[\frac{1}{T}\right] = \frac{1}{\mathrm{Ex}[T]}.$$

Benchmark	RISC	CISC	CISC/RISC
E-string search	150	120	0.8
F-bit test	120	180	1.5
Ackerman	150	300	2.0
Rec 2-sort	2800	1400	0.5
Average			1.2

Table 18.1 Sample program lengths for benchmark problems using RISC and CISC compilers.

Here is a counterexample. Define T so that

$$\Pr[T = 1] = \frac{1}{2} \quad \text{and} \quad \Pr[T = 2] = \frac{1}{2}.$$

Then

$$\mathrm{Ex}[T] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}$$

and

$$\frac{1}{\mathrm{Ex}[T]} = \frac{2}{3}$$

and

$$\mathrm{Ex}\left[\frac{1}{T}\right] = \frac{1}{1} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4} \neq \frac{1}{\mathrm{Ex}[1/T]}.$$

This means that Claim 18.5.1 is also false since we could define $S = 1$ with probability 1. In fact, both Claims 18.5.1 and 18.5.2 are untrue for most all choices of S and T. Unfortunately, the fact that they are false does not keep them from being widely used in practice! Let's see an example.

18.5.1 A RISC Paradox

The data in Table 18.1 is representative of data in a paper by some famous professors. They wanted to show that programs on a RISC processor are generally shorter than programs on a CISC processor. For this purpose, they applied a RISC compiler and then a CISC compiler to some benchmark source programs and made a table of compiled program lengths.

Each row in Table 18.1 contains the data for one benchmark. The numbers in the second and third columns are program lengths for each type of compiler. The fourth column contains the ratio of the CISC program length to the RISC program length. Averaging this ratio over all benchmarks gives the value 1.2 in the lower right. The conclusion is that CISC programs are 20% longer on average.

Benchmark	RISC	CISC	RISC/CISC
E-string search	150	120	1.25
F-bit test	120	180	0.67
Ackerman	150	300	0.5
Rec 2-sort	2800	1400	2.0
Average			1.1

Table 18.2 The same data as in Table 18.1, but with the opposite ratio in the last column.

However, some critics of their paper took the same data and argued this way: redo the final column, taking the other ratio, RISC/CISC instead of CISC/RISC, as shown in Table 18.2.

From Table 18.2, we would conclude that RISC programs are 10% longer than CISC programs on average! We are using the same reasoning as in the paper, so this conclusion is equally justifiable—yet the result is opposite. What is going on?

A Probabilistic Interpretation

To resolve these contradictory conclusions, we can model the RISC vs. CISC debate with the machinery of probability theory.

Let the sample space be the set of benchmark programs. Let the random variable R be the length of the compiled RISC program, and let the random variable C be the length of the compiled CISC program. We would like to compare the average length $\mathrm{Ex}[R]$ of a RISC program to the average length $\mathrm{Ex}[C]$ of a CISC program.

To compare average program lengths, we must assign a probability to each sample point; in effect, this assigns a "weight" to each benchmark. One might like to weigh benchmarks based on how frequently similar programs arise in practice. Lacking such data, however, we will assign all benchmarks equal weight; that is, our sample space is uniform.

In terms of our probability model, the paper computes C/R for each sample point, and then averages to obtain $\mathrm{Ex}[C/R] = 1.2$. This much is correct. The authors then conclude that CISC programs are 20% longer on average; that is, they conclude that $\mathrm{Ex}[C] = 1.2\ \mathrm{Ex}[R]$. Therein lies the problem. The authors have implicitly used False Claim 18.5.1 to assume that $\mathrm{Ex}[C/R] = \mathrm{Ex}[C]/\mathrm{Ex}[R]$. By using the same false logic, the critics can arrive at the opposite conclusion; namely, that RISC programs are 10% longer on average.

The Proper Quotient

We can compute $\mathrm{Ex}[R]$ and $\mathrm{Ex}[C]$ as follows:

$$\begin{aligned}\mathrm{Ex}[R] &= \sum_{i \in \mathrm{Range(R)}} i \cdot \Pr[R = i] \\ &= \frac{150}{4} + \frac{120}{4} + \frac{150}{4} + \frac{2800}{4} \\ &= 805,\end{aligned}$$

$$\begin{aligned}\mathrm{Ex}[C] &= \sum_{i \in \mathrm{Range(C)}} i \cdot \Pr[C = i] \\ &= \frac{120}{4} + \frac{180}{4} + \frac{300}{4} + \frac{1400}{4} \\ &= 500\end{aligned}$$

Now since $\mathrm{Ex}[R]/\mathrm{Ex}[C] = 1.61$, we conclude that the *average RISC program* is 61% longer than the *average CISC program*. This is a third answer, completely different from the other two! Furthermore, this answer makes RISC look really bad in terms of code length. This one is the correct conclusion, under our assumption that the benchmarks deserve equal weight. Neither of the earlier results were correct—not surprising since both were based on the same False Claim.

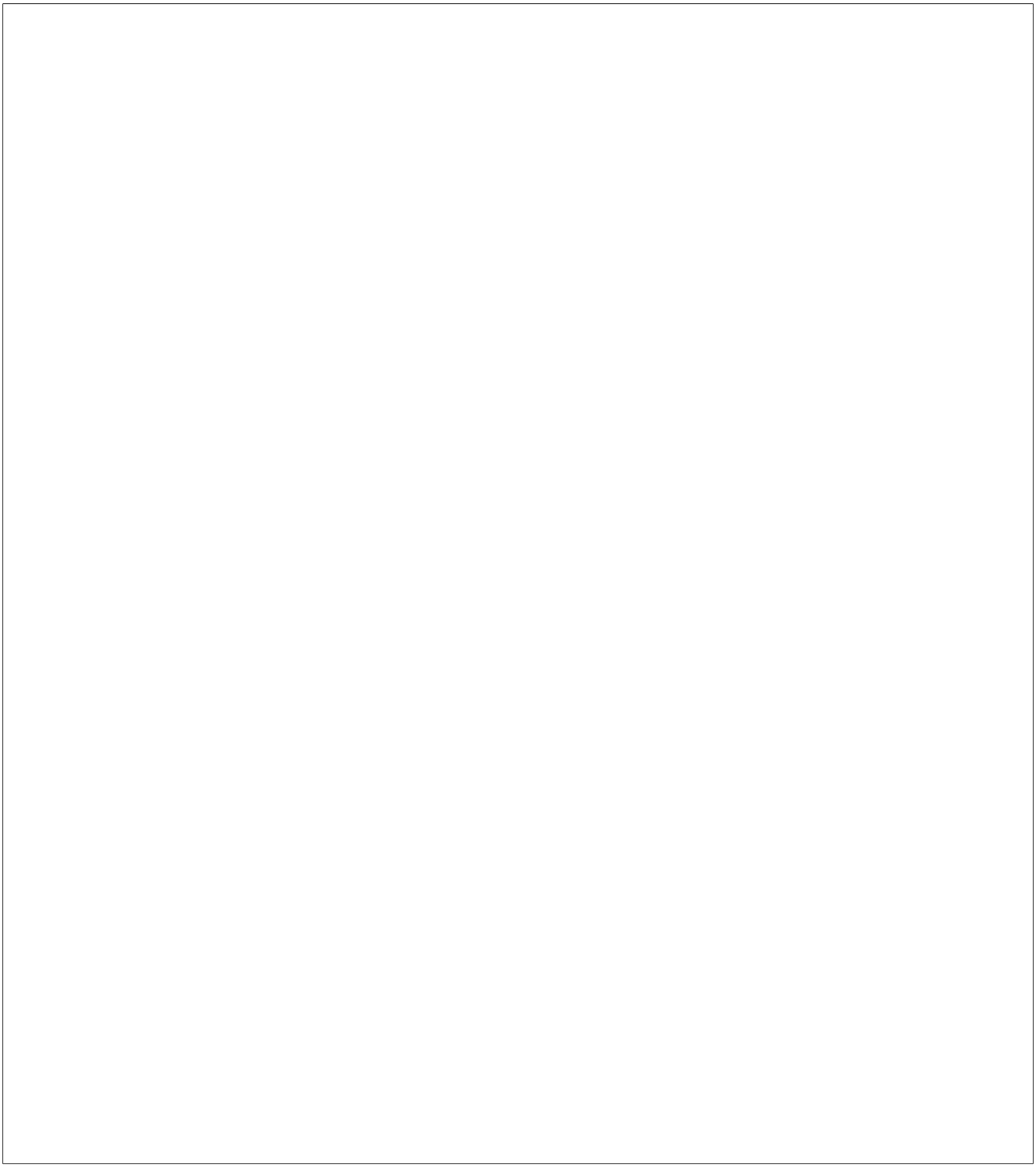
A Simpler Example

The source of the problem is clearer in the following, simpler example. Suppose the data were as follows.

Benchmark	Processor A	Processor B	B/A	A/B
Problem 1	2	1	1/2	2
Problem 2	1	2	2	1/2
Average			1.25	1.25

Now the data for the processors A and B is exactly symmetric; the two processors are equivalent. Yet, from the third column we would conclude that Processor B programs are 25% longer on average, and from the fourth column we would conclude that Processor A programs are 25% longer on average. Both conclusions are obviously wrong.

The moral is that one must be very careful in summarizing data, we must not take an average of ratios blindly!



19 Deviations

In some cases, a random variable is likely to be very close to its expected value. For example, if we flip 100 fair, mutually-independent coins, it is very likely that we will get about 50 heads. In fact, we proved in Section 17.5 that the probability of getting fewer than 25 or more than 75 heads are each less than $3 \cdot 10^{-7}$. In such cases, the mean provides a lot of information about the random variable.

In other cases, a random variable is likely to be *far* from its expected value. For example, suppose we flipped 100 fair coins that are glued together so that they all come out "heads" or they call all come out "tails." In this case, the expected value of the number of heads is still 50, but the actual number of heads is guaranteed to be far from this value—it will be 0 or 100, each with probability $1/2$.

Mathematicians have developed a variety of measures and methods to help us understand how a random variable performs in comparison to its mean. The simplest and most widely used measure is called the *variance* of the random variable. The variance is a single value associated with the random variable that is large for random variables that are likely to deviate significantly from the mean and that is small otherwise.

19.1 Variance

19.1.1 Definition and Examples

Consider the following two gambling games:

Game A: You win \$2 with probability $2/3$ and lose \$1 with probability $1/3$.

Game B: You win \$1002 with probability $2/3$ and lose \$2001 with probability $1/3$.

Which game would you rather play? Which game is better financially? We have the same probability, 2/3, of winning each game, but that does not tell the whole story. What about the expected return for each game? Let random variables A and B be the payoffs for the two games. For example, A is 2 with probability 2/3 and -1 with

probability 1/3. We can compute the expected payoff for each game as follows:

$$\begin{aligned}\mathrm{Ex}[A] &= 2 \cdot \frac{2}{3} + (-1) \cdot \frac{1}{3} = 1,\\ \mathrm{Ex}[B] &= 1002 \cdot \frac{2}{3} + (-2001) \cdot \frac{1}{3} = 1.\end{aligned}$$

The expected payoff is the same for both games, but they are obviously very different! The stakes are a lot higher for Game B and so it is likely to deviate much farther from its mean than is Game A. This fact is captured by the notion of *variance*.

Definition 19.1.1. The *variance* $\mathrm{Var}[R]$ of a random variable R is

$$\mathrm{Var}[R] ::= \mathrm{Ex}[(R - \mathrm{Ex}[R])^2].$$

In words, the variance of a random variable R is the expectation of the square of the amount by which R differs from its expectation.

Yikes! That's a mouthful. Try saying that 10 times in a row!

Let's look at this definition more carefully. We'll start with $R - \mathrm{Ex}[R]$. That's the amount by which R differs from its expectation and it is obviously an important measure. Next, we square this value. More on why we do that in a moment. Finally, we take the the expected value of the square. If the square is likely to be large, then the variance will be large. If it is likely to be small, then the variance will be small. That's just the kind of statistic we are looking for. Let's see how it works out for our two gambling games.

We'll start with Game A:

$$\begin{aligned}A - \mathrm{Ex}[A] &= \begin{cases}1 & \text{with probability } \frac{2}{3}\\ -2 & \text{with probability } \frac{1}{3}\end{cases}\\ (A - \mathrm{Ex}[A])^2 &= \begin{cases}1 & \text{with probability } \frac{2}{3}\\ 4 & \text{with probability } \frac{1}{3}\end{cases}\\ \mathrm{Ex}[(A - \mathrm{Ex}[A])^2] &= 1 \cdot \frac{2}{3} + 4 \cdot \frac{1}{3}\\ \mathrm{Var}[A] &= 2. \qquad (19.1)\end{aligned}$$

For Game B, we have

$$B - \mathrm{Ex}[B] = \begin{cases} 1001 & \text{with probability } \frac{2}{3} \\ -2002 & \text{with probability } \frac{1}{3} \end{cases}$$

$$(B - \mathrm{Ex}[B])^2 = \begin{cases} 1{,}002{,}001 & \text{with probability } \frac{2}{3} \\ 4{,}008{,}004 & \text{with probability } \frac{1}{3} \end{cases}$$

$$\mathrm{Ex}[(B - \mathrm{Ex}[B])^2] = 1{,}002{,}001 \cdot \frac{2}{3} + 4{,}008{,}004 \cdot \frac{1}{3}$$

$$\mathrm{Var}[B] = 2{,}004{,}002.$$

The variance of Game A is 2 and the variance of Game B is more than two million! Intuitively, this means that the payoff in Game A is usually close to the expected value of \$1, but the payoff in Game B can deviate very far from this expected value.

High variance is often associated with high risk. For example, in ten rounds of Game A, we expect to make \$10, but could conceivably lose \$10 instead. On the other hand, in ten rounds of Game B, we also expect to make \$10, but could actually lose more than \$20,000!

Why Bother Squaring?

The variance is the average *of the square* of the deviation from the mean. For this reason, variance is sometimes called the "mean squared deviation." But why bother squaring? Why not simply compute the average deviation from the mean? That is, why not define variance to be $\mathrm{Ex}[R - \mathrm{Ex}[R]]$?

The problem with this definition is that the positive and negative deviations from the mean exactly cancel. By linearity of expectation, we have:

$$\mathrm{Ex}\left[R - \mathrm{Ex}[R]\right] = \mathrm{Ex}[R] - \mathrm{Ex}\left[\mathrm{Ex}[R]\right].$$

Since $\mathrm{Ex}[R]$ is a constant, its expected value is itself. Therefore

$$\mathrm{Ex}\left[R - \mathrm{Ex}[R]\right] = \mathrm{Ex}[R] - \mathrm{Ex}[R] = 0.$$

By this definition, every random variable would have zero variance, which would not be very useful! Because of the square in the conventional definition, both positive and negative deviations from the mean increase the variance, and they do not cancel.

Of course, we could also prevent positive and negative deviations from canceling by taking an absolute value. In other words, we could compute $\mathrm{Ex}[\,|R - \mathrm{Ex}[R]|\,]$. But this measure doesn't have the many useful properties that variance has, and so mathematicians went with squaring.

19.1.2 Standard Deviation

Because of its definition in terms of the square of a random variable, the variance of a random variable may be very far from a typical deviation from the mean. For example, in Game B above, the deviation from the mean is 1001 in one outcome and -2002 in the other. But the variance is a whopping 2,004,002.

From a dimensional analysis viewpoint, the "units" of variance are wrong: if the random variable is in dollars, then the expectation is also in dollars, but the variance is in square dollars.

For these reasons, people often describe the deviation of a random variable using *standard deviation* instead of variance.

Definition 19.1.2. The *standard deviation* σ_R of a random variable R is the square root of the variance:

$$\sigma_R ::= \sqrt{\text{Var}[R]} = \sqrt{\text{Ex}[(R - \text{Ex}[R])^2]}.$$

So the standard deviation is the square root of the mean of the square of the deviation, or the *root mean square* for short. It has the same units—dollars in our example—as the original random variable and as the mean. Intuitively, it measures the average deviation from the mean, since we can think of the square root on the outside as roughly canceling the square on the inside.

For example, the standard deviations for A and B are

$$\sigma_A = \sqrt{\text{Var}[A]} = \sqrt{2} \approx 1.41,$$
$$\sigma_B = \sqrt{\text{Var}[B]} = \sqrt{2,004,002} \approx 1416.$$

The random variable B actually deviates from the mean by either positive 1001 or negative 2002; therefore, the standard deviation of 1416 describes this situation reasonably well.

19.1.3 An Alternative Formulation

Applying linearity of expectation to the formula for variance yields a convenient alternative formula.

Lemma 19.1.3. *For any random variable R,*

$$\text{Var}[R] = \text{Ex}[R^2] - \text{Ex}^2[R].$$

Here we use the notation $\text{Ex}^2[R]$ as shorthand for $(\text{Ex}[R])^2$. Remember that $\text{Ex}[R^2]$ is generally not equal to $\text{Ex}^2[R]$. We know the expected value of a product is the product of the expected values for independent variables, but not in general. And R is not independent of itself unless it is constant.

Proof of Lemma 19.1.3. Let $\mu = \text{Ex}[R]$. Then

$$\begin{aligned}
\text{Var}[R] &= \text{Ex}[(R - \text{Ex}[R])^2] && \text{(Definition 19.1.1 of variance)} \\
&= \text{Ex}[(R - \mu)^2] && \text{(definition of } \mu\text{)} \\
&= \text{Ex}[R^2 - 2\mu R + \mu^2] \\
&= \text{Ex}[R^2] - 2\mu\,\text{Ex}[R] + \mu^2 && \text{(linearity of expectation)} \\
&= \text{Ex}[R^2] - 2\mu^2 + \mu^2 && \text{(definition of } \mu\text{)} \\
&= \text{Ex}[R^2] - \mu^2 \\
&= \text{Ex}[R^2] - \text{Ex}^2[R]. && \text{(definition of } \mu\text{)}
\end{aligned}$$

■

For example, let's take another look at Game A from Section 19.1 where you win \$2 with probability $2/3$ and lose \$1 with probability $1/3$. Then

$$\text{Ex}[A] = 2 \cdot \frac{2}{3} + (-1) \cdot \frac{1}{3} = 1$$

and

$$\text{Ex}[A^2] = 4 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = 3.$$

By Lemma 19.1.3, this means that

$$\text{Var}[A] = \text{Ex}[A^2] - \text{Ex}^2[A] = 3 - 1^2 = 2,$$

confirming the result in Equation 19.1.

The alternate formulation of variance given in Lemma 19.1.3 has a cute implication:

Corollary 19.1.4. *If R is a random variable, then* $\text{Ex}[R^2] \geq \text{Ex}^2[R]$.

Proof. We defined $\text{Var}[R]$ as an average of a squared expression, so $\text{Var}[R]$ is nonnegative. Then we proved that $\text{Var}[R] = \text{Ex}[R^2] - \text{Ex}^2[R]$. This implies that $\text{Ex}[R^2] - \text{Ex}^2[R]$ is nonnegative. Therefore, $\text{Ex}[R^2] \geq \text{Ex}^2[R]$. ■

In words, the expectation of a square is at least the square of the expectation. The two are equal exactly when the variance is zero:

$$\text{Ex}[R^2] = \text{Ex}^2[R] \quad \text{iff} \quad \text{Ex}[R^2] - \text{Ex}^2[R] = 0 \quad \text{iff} \quad \text{Var}[R] = 0.$$

This happens precisely when

$$\Pr\left[R = \text{Ex}[R]\right] = 1;$$

namely, when R is a constant.[1]

[1] Technically, R could deviate from its mean on some sample points with probability 0, but we are ignoring events of probability 0 when computing expectations and variances.

19.1.4 Indicator Random Variables

Computing the variance of an indicator random variable is straightforward given Lemma 19.1.3.

Lemma 19.1.5. *Let B be an indicator random variable for which* $\Pr[B = 1] = p$. *Then*

$$\text{Var}[B] = p - p^2 = p(1-p). \tag{19.2}$$

Proof. By Lemma 18.1.3, $\text{Ex}[B] = p$. But since B only takes values 0 and 1, $B^2 = B$. So

$$\text{Var}[B] = \text{Ex}[B^2] - \text{Ex}^2[B] = p - p^2,$$

as claimed. ■

For example, let R be the number of heads when you flip a single fair coin. Then

$$\text{Var}[R] = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \tag{19.3}$$

and

$$\sigma_R = \sqrt{\frac{1}{4}} = \frac{1}{2}.$$

19.1.5 Mean Time to Failure

As another example, consider the mean time to failure problem, described in Section 18.1.4. If the system crashes at each step with probability p, then we already know that the mean time to failure is $1/p$. In other words, if C is the number of steps up to and including the step when the first crash occurs, then

$$\text{Ex}[C] = \frac{1}{p}.$$

What about the variance of C? To use Lemma 19.1.3, we need to compute $\text{Ex}[C^2]$. As in Section 18.1.4, we can do this by summing over all the sample points or we can use the Law of Total Expectation. The latter approach is simpler, so we'll do that. The analysis breaks into two cases: the system crashes in the first step or it doesn't. Hence,

$$\begin{aligned}
\text{Ex}[C^2] &= 1^2 \cdot p + \text{Ex}[(C+1)^2](1-p) \\
&= p + \text{Ex}[C^2](1-p) + 2\,\text{Ex}[C](1-p) + (1-p) \\
&= 1 + \text{Ex}[C^2](1-p) + 2\left(\frac{1-p}{p}\right).
\end{aligned}$$

Simplifying, we find that

$$p \operatorname{Ex}[C^2] = \frac{2-p}{p}$$

and that

$$\operatorname{Ex}[C^2] = \frac{2-p}{p^2}.$$

Using Lemma 19.1.3, we conclude that

$$\begin{aligned}\operatorname{Var}[C] &= \operatorname{Ex}[C^2] - \operatorname{Ex}^2[C] \\ &= \frac{2-p}{p^2} - \frac{1}{p^2} \\ &= \frac{1-p}{p^2}.\end{aligned}$$

19.1.6 Uniform Random Variables

Computing the variance of a uniform random variable is also straightforward given Lemma 19.1.3. For example, we can compute the variance of the outcome of a fair die R as follows:

$$\begin{aligned}\operatorname{Ex}[R^2] &= \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6}, \\ \operatorname{Ex}^2[R] &= \left(3\frac{1}{2}\right)^2 = \frac{49}{4}, \\ \operatorname{Var}[R] &= \operatorname{Ex}[R^2] - \operatorname{Ex}^2[R] = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.\end{aligned}$$

For a general uniform random variable R on $\{1, 2, 3, \ldots n\}$, the variance can be

computed as follows:

$$\begin{aligned}
\operatorname{Ex}[R] &= \frac{1}{n}(1+2+\cdots+n) \\
&= \frac{1}{n}\cdot\frac{n(n+1)}{2} \\
&= \frac{n+1}{2}.
\end{aligned}$$

$$\begin{aligned}
\operatorname{Ex}[R^2] &= \frac{1}{n}(1^2+2^2+\cdots+n^2) \\
&= \frac{1}{n}\cdot\frac{(2n+1)n(n+1)}{6} \\
&= \frac{(2n+1)(n+1)}{6}.
\end{aligned}$$

$$\begin{aligned}
\operatorname{Var}[R] &= \operatorname{Ex}[R^2]-\operatorname{Ex}^2[R] \\
&= \frac{(2n+1)(n+1)}{6}-\left(\frac{n+1}{2}\right)^2 \\
&= \frac{n^2-1}{12}.
\end{aligned}$$

19.1.7 Dealing with Constants

It helps to know how to calculate the variance of $aR+b$:

Theorem 19.1.6. *Let R be a random variable, and let a and b be constants. Then*

$$\operatorname{Var}[aR+b] = a^2\operatorname{Var}[R]. \tag{19.4}$$

Proof. Beginning with Lemma 19.1.3 and repeatedly applying linearity of expectation, we have:

$$\begin{aligned}
\operatorname{Var}[aR] &= \operatorname{Ex}[(aR+b)^2]-\operatorname{Ex}^2[aR+b] \\
&= \operatorname{Ex}[a^2R^2+2abR+b^2]-(a\operatorname{Ex}[R]+b)^2 \\
&= a^2\operatorname{Ex}[R^2]+2ab\operatorname{Ex}[R]+b^2-a^2\operatorname{Ex}^2[R]-2ab\operatorname{Ex}[R]-b^2 \\
&= a^2\operatorname{Ex}[R^2]-a^2\operatorname{Ex}^2[R] \\
&= a^2\left(\operatorname{Ex}[R^2]-\operatorname{Ex}^2[R]\right) \\
&= a^2\operatorname{Var}[R] \qquad \text{(by Lemma 19.1.3).}
\end{aligned}$$

■

Corollary 19.1.7.

$$\sigma_{aR+b} = |a|\,\sigma_R.$$

19.1.8 Variance of a Sum

In general, the variance of a sum is not equal to the sum of the variances, but variances do add for *independent* random variables. In fact, *mutual* independence is not necessary: *pairwise* independence will do.

Theorem 19.1.8. *If R_1 and R_2 are independent random variables, then*

$$\mathrm{Var}[R_1 + R_2] = \mathrm{Var}[R_1] + \mathrm{Var}[R_2]. \tag{19.5}$$

Proof. As with the proof of Theorem 19.1.6, this proof uses repeated applications of Lemma 19.1.3 and Linearity of Expectation.

$$\begin{aligned}
\mathrm{Var}[R_1 + R_2] &= \mathrm{Ex}[(R_1 + R_2)^2] - \mathrm{Ex}^2[R_1 + R_2] \\
&= \mathrm{Ex}[R_1^2 + 2R_1R_2 + R^2] - (\mathrm{Ex}[R_1] + \mathrm{Ex}[R_2])^2 \\
&= \mathrm{Ex}[R_1^2] + 2\,\mathrm{Ex}[R_1R_2] + \mathrm{Ex}[R_2^2] \\
&\qquad - \mathrm{Ex}^2[R_1] - 2\,\mathrm{Ex}[R_1]\,\mathrm{Ex}[R_2] - \mathrm{Ex}^2[R_2] \\
&= \mathrm{Var}[R_1] + \mathrm{Var}[R_2] + 2(\mathrm{Ex}[R_1R_2] - \mathrm{Ex}[R_1]\,\mathrm{Ex}[R_2]) \\
&= \mathrm{Var}[R_1] + \mathrm{Var}[R_2].
\end{aligned}$$

The last step follows because

$$\mathrm{Ex}[R_1R_2] = \mathrm{Ex}[R_1]\,\mathrm{Ex}[R_2]$$

when R_1 and R_2 are independent. ■

Note that Theorem 19.1.8 does not necessarily hold if R_1 and R_2 are dependent since then it would generally not be true that

$$\mathrm{Ex}[R_1R_2] = \mathrm{Ex}[R_1]\,\mathrm{Ex}[R_2] \tag{19.6}$$

in the last step of the proof. For example, suppose that $R_1 = R_2 = R$. Then Equation 19.6 holds only if R is essentially constant.

The proof of Theorem 19.1.8 carries over straightforwardly to the sum of any finite number of variables.

Theorem 19.1.9 (Pairwise Independent Additivity of Variance)**.** *If $R_1, R_2, \ldots, R_n$ are* pairwise *independent random variables, then*

$$\mathrm{Var}[R_1 + R_2 + \cdots + R_n] = \mathrm{Var}[R_1] + \mathrm{Var}[R_2] + \cdots + \mathrm{Var}[R_n]. \tag{19.7}$$

Unfortunately, there is no product rule for computing variances, even if the random variables are mutually independent. However, we can use Theorem 19.1.9 to quickly compute the variance of a random variable with a general binomial distribution.

19.1.9 Binomial Distributions

Lemma 19.1.10 (Variance of the Binomial Distribution)**.** *If J has a binomial distribution with parameters n and p, then*

$$\mathrm{Var}[J] = np(1-p). \tag{19.8}$$

Proof. From the definition of the binomial distribution, we can think of J as being the number of "heads" when you flip n mutually independent coins, each of which is "heads" with probability p. Thus J can be expressed as the sum of n mutually independent indicator variables J_i where

$$\Pr[J_i = 1] = p$$

for $1 \le i \le n$. From Lemma 19.1.5, we know that

$$\mathrm{Var}[J_i] = p(1-p).$$

By Theorem 19.1.9, this means that

$$\mathrm{Var}[J] = \sum_{i=1}^{n} \mathrm{Var}[J_i] = np(1-p). \qquad \blacksquare$$

For example, suppose we flip n mutually independent[2] fair coins. Let R be the number of heads. Then Theorem 19.1.9 tells us that

$$\mathrm{Var}[R] = n\left(\frac{1}{2}\right)\left(1-\frac{1}{2}\right) = \frac{n}{4}.$$

Hence,

$$\sigma_R = \frac{\sqrt{n}}{2}.$$

This value is small compared with

$$\mathrm{Ex}[R] = \frac{n}{2},$$

which should not be surprising since we already knew from Section 17.5 that R is unlikely to stray very far from its mean.

[2] Actually, we only need to assume pairwise independence for this to be true using Theorem 19.1.9.

19.2 Markov's Theorem

The variance of a random variable gives us a rough idea of the amount by which a random variable is likely to deviate from its mean. But it does not directly give us specific bounds on the probability that the deviation exceeds a specified threshold. To obtain such specific bounds, we'll need to work a little harder.

In this section, we derive a famous result known as Markov's Theorem that gives an upper bound on the probability that a random variable exceeds a specified threshold. In the next section, we give a similar but stronger result known as Chebyshev's Theorem. The difference between these results is that Markov's Theorem depends only on the mean of the random variable, whereas Chebyshev's Theorem makes use of the mean *and* the variance. Basically, the more you know about a random variable, the better bounds you can derive on the probability that it deviates from its mean.

19.2.1 A Motivating Example

The idea behind Markov's Theorem can be explained with a simple example involving *intelligence quotients*, or IQs. This quantity was devised so that the average IQ measurement would be 100. From this fact alone we can conclude that at most 1/3 the population can have an IQ of 300 or more, because if more than a third had an IQ of at least 300, then the average IQ would have to be *more* than $(1/3)300 = 100$, contradicting the fact that the average is 100. So the probability that a randomly chosen person has an IQ of 300 or more is at most 1/3. Of course this is not a very strong conclusion since no IQ over 200 has ever been recorded.

By the same logic, we can also conclude that at most 2/3 of the population can have an IQ of 150 or more. IQ's over 150 have certainly been recorded, although a much smaller fraction than 2/3 of the population actually has an IQ that high.

Although these conclusions about IQ are weak, they are actually the strongest general conclusions that can be reached about a random variable using *only* the fact that it is nonnegative and its mean is 100. For example, if we choose a random variable equal to 300 with probability 1/3, and 0 with probability 2/3, then its mean is 100, and the probability of a value of 300 or more really is 1/3. So we can't hope to get a better upper bound based solely on this limited amount of information.

Markov's Theorem characterizes the bounds that can be achieved with this kind of analysis

19.2.2 The Theorem

Theorem 19.2.1 (Markov's Theorem)**.** *If R is a nonnegative random variable, then for all $x > 0$,*

$$\Pr[R \geq x] \leq \frac{\mathrm{Ex}[R]}{x}.$$

Proof. For any $x > 0$

$$\begin{aligned}
\mathrm{Ex}[R] &= \sum_{y \in \mathrm{range}(R)} y \Pr[R = y] \\
&\geq \sum_{\substack{y \geq x, \\ y \in \mathrm{range}(R)}} y \Pr[R = y] \qquad \text{(because } R \geq 0\text{)} \\
&\geq \sum_{\substack{y \geq x, \\ y \in \mathrm{range}(R)}} x \Pr[R = y] \\
&= x \sum_{\substack{y \geq x, \\ y \in \mathrm{range}(R)}} \Pr[R = y] \\
&= x \Pr[R \geq x]. \qquad (19.9)
\end{aligned}$$

Hence,

$$\Pr[R \geq x] \leq \frac{\mathrm{Ex}[R]}{x}.$$ ■

Corollary 19.2.2. *If R is a nonnegative random variable, then for all $c \geq 1$,*

$$\Pr\left[R \geq c \cdot \mathrm{Ex}[R]\right] \leq \frac{1}{c}. \qquad (19.10)$$

Proof. Set $x = c\,\mathrm{Ex}[R]$ in Theorem 19.2.1. ■

As an example, suppose we flip 100 fair coins and use Markov's Theorem to compute the probability of getting all heads:

$$\Pr[\text{heads} \geq 100] \leq \frac{\mathrm{Ex}[\text{heads}]}{100} = \frac{50}{100} = \frac{1}{2}.$$

If the coins are mutually independent, then the actual probability of getting all heads is a minuscule 1 in 2^{100}. In this case, Markov's Theorem looks very weak. However, in applying Markov's Theorem, we made no independence assumptions. In fact, if all the coins are glued together, then probability of throwing all heads is exactly $1/2$. In this nasty case, Markov's Theorem is actually tight!

The Chinese Appetizer Problem

Suppose that n people are seated at a circular table and that each person has an appetizer in front of them on a rotating Chinese banquet tray. Just as everyone is about to dig in, some joker spins the tray so that each person receives a random appetizer. We are interested in the number of people R that get their same appetizer as before, assuming that the n appetizers are all different.

Each person gets their original appetizer with probability $1/n$. Hence, by Linearity of Expectation,

$$\mathrm{Ex}[R] = n \cdot \frac{1}{n} = 1.$$

What is the probability that all n people get their original appetizer back? Markov's Theorem tells us that

$$\Pr[R = n] = \Pr[R \geq n] \leq \frac{\mathrm{Ex}[R]}{n} = \frac{1}{n}.$$

In fact, this bound is tight sine everyone gets their original appetizers back if and only if the rotating tray returns to its original configuration, which happens with probability $1/n$.

The Chinese Appetizer problem is similar to the Hat Check problem that we studied in Section 18.3.2, except that no distribution was specified in the Hat Check problem—we were told only that each person gets their correct hat back with probability $1/n$. If the hats are scrambled according to uniformly random permutations, then the probability that everyone gets the right hat back is $1/n!$, which is much less than the $1/n$ upper bound given by Markov's Theorem. So, in this case, the bound given by Markov's Theorem is not close to the actual probability.

What is the probability that at least two people get their right hats back? Markov's Theorem tells us that

$$\Pr[R \geq 2] \leq \frac{\mathrm{Ex}[R]}{2} = \frac{1}{2}.$$

In this case, Markov's Theorem is not too far off from the right answer if the hats are distributed according to a random permutation[3] but it is not very close to the correct answer $1/n$ for the case when the hats are distributed as in the Chinese Appetizer problem.

Why R Must be Nonnegative

Remember that Markov's Theorem applies only to nonnegative random variables! Indeed, the theorem is false if this restriction is removed. For example, let R be -10

[3] Proving this requires some effort.

with probability $1/2$ and 10 with probability $1/2$. Then

$$\mathrm{Ex}[R] = -10 \cdot \frac{1}{2} + 10 \cdot \frac{1}{2} = 0.$$

Suppose that we now tried to compute $\Pr[R \geq 5]$ using Markov's Theorem:

$$\Pr[R \geq 5] \leq \frac{\mathrm{Ex}[R]}{5} = \frac{0}{5} = 0.$$

This is the wrong answer! Obviously, R is at least 5 with probability $1/2$.

On the other hand, we can still apply Markov's Theorem indirectly to derive a bound on the probability that an arbitrary variable like R is 5 or more. For example, given any random variable, R with expectation 0 and values ≥ -10, we can conclude that $\Pr[R \geq 5] \leq 2/3$. To prove this fact, we define $T ::= R + 10$. Then T is a nonnegative random variable with expectation $\mathrm{Ex}[R + 10] = \mathrm{Ex}[R] + 10 = 10$, so Markov's Theorem applies and tells us that $\Pr[T \geq 15] \leq 10/15 = 2/3$. But $T \geq 15$ iff $R \geq 5$, so $\Pr[R \geq 5] \leq 2/3$, as claimed.

19.2.3 Markov's Theorem for Bounded Variables

Suppose we learn that the average IQ among MIT students is 150 (which is not true, by the way). What can we say about the probability that an MIT student has an IQ of more than 200? Markov's Theorem immediately tells us that no more than $150/200$ or $3/4$ of the students can have such a high IQ. That's because if R is the IQ of a random MIT student, then

$$\Pr[R > 200] \leq \frac{\mathrm{Ex}[R]}{200} = \frac{150}{200} = \frac{3}{4}.$$

But let's also suppose that no MIT student has an IQ less than 100 (which may be true). This means that if we let $T ::= R - 100$, then T is nonnegative and $\mathrm{Ex}[T] = 50$, so we can apply Markov's Theorem to T and conclude:

$$\Pr[R > 200] = \Pr[T > 100] \leq \frac{\mathrm{Ex}[T]}{100} = \frac{50}{100} = \frac{1}{2}.$$

So only half, not 3/4, of the students can be as amazing as they think they are. A bit of a relief!

More generally, we can get better bounds applying Markov's Theorem to $R - l$ instead of R for any lower bound l on R, even when l is negative.

Theorem 19.2.3. *Let R be a random variable for which $R \geq l$ for some $l \in \mathbb{R}$. Then for all $x \geq l$,*

$$\Pr[R \geq x] \leq \frac{\mathrm{Ex}[R] - l}{x - l}.$$

Proof. Define

$$T ::= R - l.$$

Then T is a nonnegative random variable with mean

$$\mathrm{Ex}[T] = \mathrm{Ex}[R - l] = \mathrm{Ex}[R] - l.$$

Hence, Markov's Theorem implies that

$$\begin{aligned}\Pr[T \geq x - l] &\leq \frac{\mathrm{Ex}[T]}{x - l} \\ &= \frac{\mathrm{Ex}[R] - l}{x - l}.\end{aligned}$$

The result then follows from the fact that

$$\begin{aligned}\Pr[R \geq x] &= \Pr[R - l \geq x - l] \\ &= \Pr[T \geq x - l].\end{aligned}$$

■

19.2.4 Deviations Below the Mean

Markov's Theorem says that a random variable is unlikely to greatly exceed the mean. Correspondingly, there is a variation of Markov's Theorem that says a random variable is unlikely to be much smaller than its mean.

Theorem 19.2.4. *Let $u \in \mathbb{R}$ and let R be a random variable such that $R \leq u$. Then for all $x < u$,*

$$\Pr[R \leq x] \leq \frac{u - \mathrm{Ex}[R]}{u - x}.$$

Proof. The proof is similar to that of Theorem 19.2.3. Define

$$S ::= u - R.$$

Then S is a nonnegative random variable with mean

$$\mathrm{Ex}[S] = \mathrm{Ex}[u - R] = u - \mathrm{Ex}[R].$$

Hence, Markov's Theorem implies that

$$\Pr[S \geq u - x] \leq \frac{\mathrm{Ex}[S]}{u - x} = \frac{u - \mathrm{Ex}[R]}{u - x}.$$

The result then follows from the fact that

$$\Pr[R \leq x] = \Pr[u - S \leq x] = \Pr[S \geq u - x].$$

■

For example, suppose that the class average on a midterm was 75/100. What fraction of the class scored below 50?

There is not enough information here to answer the question exactly, but Theorem 19.2.4 gives an upper bound. Let R be the score of a random student. Since 100 is the highest possible score, we can set $u = 100$ to meet the condition in the theorem that $R \leq u$. Applying Theorem 19.2.4, we find:

$$\Pr[R \leq 50] \leq \frac{100 - 75}{100 - 50} = \frac{1}{2}.$$

That is, at most half of the class scored 50 or worse. This makes sense; if more than half of the class scored 50 or worse, then the class average could not be 75, even if everyone else scored 100. As with Markov's Theorem, Theorem 19.2.4 often gives weak results. In fact, based on the data given, the *entire* class could have scored *above* 50.

19.2.5 Using Markov's Theorem to Analyze Non-Random Events

In the previous example, we used a theorem about a random variable to conclude facts about non-random data. For example, we concluded that if the average score on a test is 75, then at most $1/2$ the class scored 50 or worse. There is no randomness in this problem, so how can we apply Theorem 19.2.4 to reach this conclusion?

The explanation is not difficult. For any set of scores $S = \{s_1, s_2, \ldots, s_n\}$, we introduce a random variable R such that

$$\Pr[R = s_i] = \frac{(\#\text{ of students with score } s_i)}{n}.$$

We then use Theorem 19.2.4 to conclude that $\Pr[R \leq 50] \leq 1/2$. To see why this means (with certainty) that at most $1/2$ of the students scored 50 or less, we observe that

$$\begin{aligned}
\Pr[R \leq 50] &= \sum_{s_i \leq 50} \Pr[R = s_i] \\
&= \sum_{s_i \leq 50} \frac{(\#\text{ of students with score } s_i)}{n} \\
&= \frac{1}{n}(\#\text{ of students with score 50 or less}).
\end{aligned}$$

So, if $\Pr[R \leq 50] \leq 1/2$, then the number of students with score 50 or less is at most $n/2$.

19.3 Chebyshev's Theorem

As we have just seen, Markov's Theorem can be extended by applying it to functions of a random variable R such as $R - l$ and $u - R$. Even stronger results can be obtained by applying Markov's Theorem to powers of R.

Lemma 19.3.1. *For any random variable R, $\alpha \in \mathbb{R}^+$, and $x > 0$,*

$$\Pr[|R| \geq x] \leq \frac{\mathrm{Ex}[|R|^\alpha]}{x^\alpha}.$$

Proof. The event $|R| \geq x$ is the same as the event $|R|^\alpha \geq x^\alpha$. Since $|R|^\alpha$ is nonnegative, the result follows immediately from Markov's Theorem. ■

Similarly,

$$\Pr[|R - \mathrm{Ex}[R]| \geq x] \leq \frac{\mathrm{Ex}[(R - \mathrm{Ex}[R])^\alpha]}{x^\alpha}. \tag{19.11}$$

The restatement of Equation 19.11 for $\alpha = 2$ is known as *Chebyshev's Theorem*.

Theorem 19.3.2 (Chebyshev)**.** *Let R be a random variable and $x \in \mathbb{R}^+$. Then*

$$\Pr[|R - \mathrm{Ex}[R]| \geq x] \leq \frac{\mathrm{Var}[R]}{x^2}.$$

Proof. Define

$$T ::= R - \mathrm{Ex}[R].$$

Then

$$\begin{aligned}
\Pr\left[\,|R - \mathrm{Ex}[R]| \geq x\right] &= \Pr[|T| \geq x] \\
&= \Pr[T^2 \geq x^2] \\
&\leq \frac{\mathrm{Ex}[T^2]}{x^2} && \text{(by Markov's Theorem)} \\
&= \frac{\mathrm{Ex}[(R - \mathrm{Ex}[R])^2]}{x^2} \\
&= \frac{\mathrm{Var}[R]}{x^2}. && \text{(by Definition 19.1.1)}
\end{aligned}$$

■

Corollary 19.3.3. *Let R be a random variable, and let c be a positive real number.*

$$\Pr[|R - \mathrm{Ex}[R]| \geq c\sigma_R] \leq \frac{1}{c^2}.$$

Proof. Substituting $x = c\sigma_R$ in Chebyshev's Theorem gives:

$$\Pr[|R - \mathrm{Ex}[R]| \geq c\sigma_R] \leq \frac{\mathrm{Var}[R]}{(c\sigma_R)^2} = \frac{\sigma_R^2}{(c\sigma_R)^2} = \frac{1}{c^2}. \qquad \blacksquare$$

As an example, suppose that, in addition to the national average IQ being 100, we also know the standard deviation of IQ's is 10. How rare is an IQ of 300 or more?

Let the random variable R be the IQ of a random person. So we are supposing that $\mathrm{Ex}[R] = 100$, $\sigma_R = 10$, and R is nonnegative. We want to compute $\Pr[R \geq 300]$.

We have already seen that Markov's Theorem 19.2.1 gives a coarse bound, namely,

$$\Pr[R \geq 300] \leq \frac{1}{3}.$$

Now we apply Corollary 19.3.3 to the same problem:

$$\Pr[R \geq 300] \leq \Pr\left[\,|R - 100| \geq 20\sigma_R\right] \leq \frac{1}{400}. \tag{19.12}$$

So Chebyshev's Theorem implies that at most one person in four hundred has an IQ of 300 or more. We have gotten a much tighter bound using the additional information, namely the standard deviation of R, than we could get knowing only the expectation.

More generally, Corollary 19.3.3 tells us that a random variable is never likely to stray by more than a few standard deviations from its mean. For example, plugging $c = 3$ into Corollary 19.3.3, we find that the probability that a random variable strays from the mean by more than 3σ is at most $1/9$.

This fact has a nice pictorial characterization for pdf's with a "bell-curve" shape; namely, the width of the bell is $O(\sigma)$, as shown in Figure 19.1.

19.3.1 Bounds on One-Sided Errors

Corollary 19.3.3 gives bounds on the probability of deviating from the mean in *either* direction. If you only care about deviations in one direction, as was the case in the IQ example, then slightly better bounds can be obtained.

Theorem 19.3.4. *For any random variable R and any $c > 0$,*

$$\Pr[R - \mathrm{Ex}[R] \geq c\sigma_R] \leq \frac{1}{c^2 + 1}$$

and

$$\Pr[R - \mathrm{Ex}[R] \leq -c\sigma_R] \leq \frac{1}{c^2 + 1}.$$

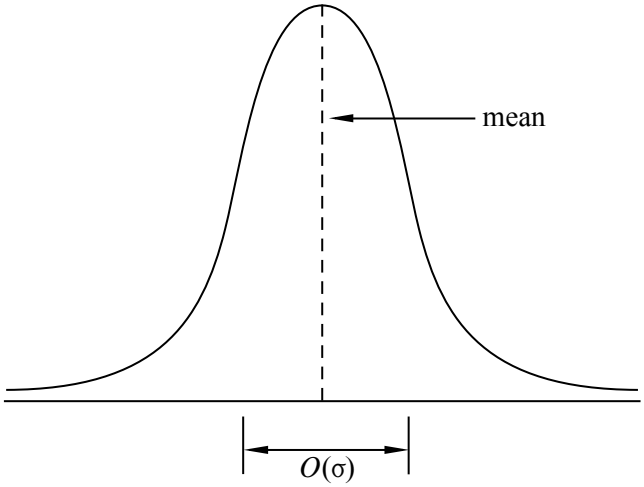


Figure 19.1 If the pdf of a random variable is "bell-shaped," then the width of the bell is $O(\sigma)$.

The proof of Theorem 19.3.4 is trickier than the proof of Chebyshev's Theorem and we will not give the details here. Nor will we prove the fact that the bounds in Theorem 19.3.4 are the best bounds that you can obtain if you know only the mean and standard deviation of the random variable R.

Returning to the IQ example, Theorem 19.3.4 tells us that

$$\Pr[R \geq 300] \leq \Pr[R - 100 \geq 20\sigma_R] \leq \frac{1}{401},$$

which is a *very slight* improvement over Equation 19.12.

As another example, suppose we give an exam. What fraction of the class can score more than 2 standard deviations from the average? If R is the score of a random student, then

$$\Pr[|R - \mathrm{Ex}[R]| \geq 2\sigma_R] \leq \frac{1}{4}.$$

For one-sided error, the fraction that could be 2 standard deviations or more above the average is at most

$$\frac{1}{2^2 + 1} = \frac{1}{5}.$$

This results holds no matter what the test scores are, and is again a deterministic fact derived using probabilistic tools.

19.4 Bounds for Sums of Random Variables

If all you know about a random variable is its mean and variance, then Chebyshev's Theorem is the best you can do when it comes to bounding the probability that the random variable deviates from its mean. In some cases, however, we know more—for example, that the random variable has a binomial distribution—and then it is possible to prove much stronger bounds. Instead of polynomially small bounds such as $1/c^2$, we can sometimes even obtain exponentially small bounds such as $1/e^c$. As we will soon discover, this is the case whenever the random variable T is the sum of n mutually independent random variables $T_1, T_2, \ldots, T_n$ where $0 \leq T_i \leq 1$. A random variable with a binomial distribution is just one of many examples of such a T. Here is another.

19.4.1 A Motivating Example

Fussbook is a new social networking site oriented toward unpleasant people.

Like all major web services, Fussbook has a load balancing problem. Specifically, Fussbook receives 24,000 forum posts every 10 minutes. Each post is assigned to one of m computers for processing, and each computer works sequentially through its assigned tasks. Processing an average post takes a computer $1/4$ second. Some posts, such as pointless grammar critiques and snide witticisms, are easier. But the most protracted harangues require 1 full second.

Balancing the work load across the m computers is vital; if any computer is assigned more than 10 minutes of work in a 10-minute interval, then that computer is overloaded and system performance suffers. That would be bad, because Fussbook users are *not* a tolerant bunch.

An early idea was to assign each computer an alphabetic range of forum topics. ("That oughta work!", one programmer said.) But after the computer handling the "*pr*ivacy" and "*pr*eferred text editor" threads melted, the drawback of an ad hoc approach was clear: there are no guarantees.

If the length of every task were known in advance, then finding a balanced distribution would be a kind of "bin packing" problem. Such problems are hard to solve exactly, though approximation algorithms can come close. But in this case, task lengths are not known in advance, which is typical for workload problems in the real world.

So the load balancing problem seems sort of hopeless, because there is no data available to guide decisions. Heck, we might as well assign tasks to computers at random!

As it turns out, random assignment not only balances load reasonably well, but

also permits provable performance guarantees in place of "That oughta work!" assertions. In general, a randomized approach to a problem is worth considering when a deterministic solution is hard to compute or requires unavailable information.

Some arithmetic shows that Fussbook's traffic is sufficient to keep $m = 10$ computers running at 100% capacity with perfect load balancing. Surely, more than 10 servers are needed to cope with random fluctuations in task length and imperfect load balance. But how many is enough? 11? 15? 20? 100? We'll answer that question with a new mathematical tool.

19.4.2 The Chernoff Bound

The Chernoff[4] bound is a hammer that you can use to nail a great many problems. Roughly, the Chernoff bound says that certain random variables are very unlikely to significantly exceed their expectation. For example, if the expected load on a computer is just a bit below its capacity, then that computer is unlikely to be overloaded, provided the conditions of the Chernoff bound are satisfied.

More precisely, the Chernoff Bound says that *the sum of lots of little, independent random variables is unlikely to significantly exceed the mean of the sum*. The Markov and Chebyshev bounds lead to the same kind of conclusion but typically provide much weaker bounds. In particular, the Markov and Chebyshev bounds are polynomial, while the Chernoff bound is exponential.

Here is the theorem. The proof will come later in Section 19.4.3.

Theorem 19.4.1 (Chernoff Bound)**.** *Let $T_1, \ldots T_n$ be mutually independent random variables such that $0 \leq T_i \leq 1$ for all i. Let $T = T_1 + \cdots + T_n$. Then for all $c \geq 1$,*

$$\Pr[T \geq c \operatorname{Ex}[T]] \leq e^{-k \operatorname{Ex}[T]} \tag{19.13}$$

where $k = c \ln(c) - c + 1$.

The Chernoff bound applies only to distributions of sums of independent random variables that take on values in the interval $[0, 1]$. The binomial distribution is of course such a distribution, but there are lots of other distributions because the Chernoff bound allows the variables in the sum to have differing, arbitrary, and even unknown distributions over the range $[0, 1]$. Furthermore, there is no direct dependence on the number of random variables in the sum or their expectations. In short, the Chernoff bound gives strong results for lots of problems based on little information—no wonder it is widely used!

[4] Yes, this is the same Chernoff who figured out how to beat the state lottery. So you might want to pay attention—this guy knows a thing or two.

More Examples

The Chernoff bound is pretty easy to apply, though the details can be daunting at first. Let's walk through a simple example to get the hang of it.

What is the probability that the number of heads that come up in 1000 independent tosses of a fair coin exceeds the expectation by 20% or more? Let T_i be an indicator variable for the event that the i-th coin is heads. Then the total number of heads is

$$T = T_1 + \cdots + T_{1000}.$$

The Chernoff bound requires that the random variables T_i be mutually independent and take on values in the range $[0, 1]$. Both conditions hold here. In fact, this example is similar to many applications of the Chernoff bound in that every T_i is *either* 0 or 1, since they're indicators.

The goal is to bound the probability that the number of heads exceeds its expectation by 20% or more; that is, to bound $\Pr[T \geq c \operatorname{Ex}[T]]$ where c = 1.2. To that end, we compute k as defined in the theorem:

$$k = c \ln(c) - c + 1 = 0.0187\ldots.$$

Plugging this value into the Chernoff bound gives:

$$\begin{aligned}\Pr\left[T \geq 1.2 \operatorname{Ex}[T]\right] &\leq e^{-k \operatorname{Ex}[T]} \\ &= e^{-(0.0187\ldots)\cdot 500} \\ &< 0.0000834.\end{aligned}$$

So the probability of getting 20% or more extra heads on 1000 coins is less than 1 in 10,000.[5]

The bound becomes much stronger as the number of coins increases, because the expected number of heads appears in the exponent of the upper bound. For example, the probability of getting at least 20% extra heads on a million coins is at most

$$e^{-(0.0187\ldots)\cdot 500000} < e^{-9392}$$

which is pretty darn small.

Alternatively, the bound also becomes stronger for larger deviations. For example, suppose we're interested in the odds of getting 30% or more extra heads in 1000 tosses, rather than 20%. In that case, $c = 1.3$ instead of 1.2. Consequently, the parameter k rises from 0.0187 to about 0.0410, which may seem insignificant.

[5] Since we are analyzing a binomial distribution here, we can get somewhat better bounds using the methods from Section 17.5, but it is much easier to use the Chernoff bounds, and they provide results that are nearly as good.

But because k appears in the exponent of the upper bound, the final probability decreases from around 1 in 10,000 to about 1 in a billion!

Pick-4

Pick-4 is a lottery game where you pick a 4-digit number between 0000 and 9999. If your number comes up in a random drawing, then you win $5,000. Your chance of winning is 1 in 10,000. And if 10 million people play, then the expected number of winners is 1000. The lottery operator's nightmare is that the number of winners is much greater; say, 2000 or more. What is the probability that will happen?

Let T_i be an indicator for the event that the i-th player wins. Then $T = T_1 + \cdots + T_n$ is the total number of winners. If we assume[6] that the players' picks and the winning number are random, independent and uniform, then the indicators T_i are independent, as required by the Chernoff bound.

Since 2000 winners would be twice the expected number, we choose $c = 2$, compute $k = c\ln(c) - c + 1 = 0.386\ldots$, and plug these values into the Chernoff bound:

$$\begin{aligned}\Pr[T \geq 2000] &= \Pr\left[T \geq 2\,\mathrm{Ex}[T]\right] \\ &\leq e^{-k\,\mathrm{Ex}[T]} \\ &= e^{-(0.386\ldots)\cdot 1000} \\ &< e^{-386}.\end{aligned}$$

So there is almost no chance that the lottery operator pays out double. In fact, the number of winners won't even be 10% higher than expected very often. To prove that, let $c = 1.1$, compute $k = c\ln(c) - c + 1 = 0.00484\ldots$, and plug in again:

$$\begin{aligned}\Pr\left[T \geq 1.1\,\mathrm{Ex}[T]\right] &\leq e^{-k\,\mathrm{Ex}[T]} \\ &= e^{-(0.00484)\cdot 1000} \\ &< 0.01.\end{aligned}$$

So the Pick-4 lottery may be exciting for the players, but the lottery operator has little doubt about the outcome!

Randomized Load Balancing

Now let's return to Fussbook and its load balancing problem. Specifically, we need to determine how many machines suffice to ensure that no server is overloaded;

[6] As we noted in Chapter 18, human choices are often not uniform and they can be highly dependent. For example, lots of people will pick an important date. So the lottery folks should not get too much comfort from the analysis that follows, unless they assign random 4-digit numbers to each player.

that is, assigned to do more than 10 minutes of work in a 10-minute interval.

To begin, let's find the probability that the first server is overloaded. Let T_i be the number of seconds that the first server spends on the i-th task. So T_i is zero if the task is assigned to another machine, and otherwise T_i is the length of the task. Then $T = \sum_{i=1}^{n} T_i$ is the total length of tasks assigned to the server, where $n = 24{,}000$. We need an upper bound on $\Pr[T \geq 600]$; that is, the probability that the first server is assigned more than 600 seconds (or, equivalently, 10 minutes) of work.

The Chernoff bound is applicable only if the T_i are mutually independent and take on values in the range $[0, 1]$. The first condition is satisfied if we assume that task lengths and assignments are independent. And the second condition is satisfied because processing even the most interminable harangue takes at most 1 second.

In all, there are 24,000 tasks, each with an expected length of 1/4 second. Since tasks are assigned to computers at random, the expected load on the first server is:

$$\begin{aligned} \operatorname{Ex}[T] &= \frac{24{,}000 \text{ tasks} \cdot 1/4 \text{ second per task}}{m \text{ machines}} \\ &= 6000/m \text{ seconds.} \end{aligned} \tag{19.14}$$

For example, if there are $m = 10$ machines, then the expected load on the first server is 600 seconds, which is 100% of its capacity.

Now we can use the Chernoff bound to upper bound the probability that the first server is overloaded:

$$\begin{aligned} \Pr\left[T \geq 600\right] &= \Pr\left[T \geq \frac{m}{10} \operatorname{Ex}[T]\right] \\ &= \Pr\left[T \geq c \operatorname{Ex}[T]\right] \\ &\leq e^{-(c \ln(c) - c + 1) \cdot 6000/m}, \end{aligned}$$

where $c = m/10$. The first equality follows from Equation 19.14.

The probability that *some* server is overloaded is at most m times the probability that the first server is overloaded by the Sum Rule in Section 14.4.2. So

$$\begin{aligned} \Pr[\text{some server is overloaded}] &\leq \sum_{i=1}^{m} \Pr[\text{server } i \text{ is overloaded}] \\ &= m \Pr[\text{the first server is overloaded}] \\ &\leq m e^{-(c \ln(c) - c + 1) \cdot 6000/m}, \end{aligned}$$

where $c = m/10$. Some values of this upper bound are tabulated below:

$$\begin{aligned} m &= 11: \quad 0.784\ldots \\ m &= 12: \quad 0.000999\ldots \\ m &= 13: \quad 0.0000000760\ldots \end{aligned}$$

These values suggest that a system with $m = 11$ machines might suffer immediate overload, $m = 12$ machines could fail in a few days, but $m = 13$ should be fine for a century or two!

19.4.3 Proof of the Chernoff Bound

The proof of the Chernoff bound is somewhat involved. Heck, even *Chernoff* didn't come up with it! His friend, Herman Rubin, showed him the argument. Thinking the bound not very significant, Chernoff did not credit Rubin in print. He felt pretty bad when it became famous![7]

Here is the theorem again, for reference:

Theorem 19.4.2 (Chernoff Bound)**.** *Let $T_1, \ldots, T_n$ be mutually independent random variables such that $0 \le T_i \le 1$ for all i. Let $T = T_1 + \cdots + T_n$. Then for all $c \ge 1$,*

$$\Pr[T \ge c \operatorname{Ex}[T]] \le e^{-k \operatorname{Ex}[T]} \tag{19.13}$$

where $k = c \ln(c) - c + 1$.

Proof. For clarity, we'll go through the proof "top down"; that is, we'll use facts that are proved immediately afterward.

The key step is to exponentiate both sides of the inequality $T \ge c \operatorname{Ex}[T]$ and then apply the Markov bound:

$$\begin{aligned}
\Pr[T \ge c \operatorname{Ex}[T]] &= \Pr[c^T \ge c^{c \operatorname{Ex}[T]}] \\
&\le \frac{\operatorname{Ex}[c^T]}{c^{c \operatorname{Ex}[T]}} && \text{(by Markov)} \\
&\le \frac{e^{(c-1)\operatorname{Ex}[T]}}{c^{c \operatorname{Ex}[T]}} \\
&= e^{-(c \ln(c) - c + 1)\operatorname{Ex}[T]}.
\end{aligned}$$

In the third step, the numerator is rewritten using the inequality

$$\operatorname{Ex}[c^T] \le e^{(c-1)\operatorname{Ex}[T]}$$

which is proved below in Lemma 19.4.3. The final step is simplification, using the fact that c^c is equal to $e^{c \ln(c)}$. ■

[7] See "A Conversation with Herman Chernoff," *Statistical Science* 1996, Vol. 11, No. 4, pp 335–350.

Algebra aside, there is a brilliant idea in this proof: in this context, exponentiating somehow supercharges the Markov bound. This is not true in general! One unfortunate side-effect is that we have to bound some nasty expectations involving exponentials in order to complete the proof. This is done in the two lemmas below, where variables take on values as in Theorem 19.4.1.

Lemma 19.4.3.

$$\operatorname{Ex}[c^T] \le e^{(c-1)\operatorname{Ex}[T]}.$$

Proof.

$$\begin{aligned}
\operatorname{Ex}[c^T] &= \operatorname{Ex}[c^{T_1+\cdots+T_n}] \\
&= \operatorname{Ex}[c^{T_1}\cdots c^{T_n}] \\
&= \operatorname{Ex}[c^{T_1}]\cdots\operatorname{Ex}[c^{T_n}] \\
&\le e^{(c-1)\operatorname{Ex}[T_1]}\cdots e^{(c-1)\operatorname{Ex}[T_n]} \\
&= e^{(c-1)(\operatorname{Ex}[T_1]+\cdots+\operatorname{Ex}[T_n])} \\
&= e^{(c-1)\operatorname{Ex}[T_1+\cdots+T_n]} \\
&= e^{(c-1)\operatorname{Ex}[T]}.
\end{aligned}$$

The first step uses the definition of T, and the second is just algebra. The third step uses the fact that the expectation of a product of independent random variables is the product of the expectations. This is where the requirement that the T_i be independent is used. Then we bound each term using the inequality

$$\operatorname{Ex}[c^{T_i}] \le e^{(c-1)\operatorname{Ex}[T_i]},$$

which is proved in Lemma 19.4.4. The last steps are simplifications using algebra and linearity of expectation. ■

Lemma 19.4.4.

$$\operatorname{Ex}[c^{T_i}] \le e^{(c-1)\operatorname{Ex}[T_i]}$$

Proof. All summations below range over values v taken by the random variable T_i,

which are all required to be in the interval [0, 1].

$$\begin{aligned}
\operatorname{Ex}[c^{T_i}] &= \sum_v c^v \Pr[T_i = v] \\
&\le \sum_v (1 + (c-1)v) \Pr[T_i = v] \\
&= \sum_v \Pr[T_i = v] + (c-1)v \Pr[T_i = v] \\
&= \sum_v \Pr[T_i = v] + \sum (c-1)v \Pr[T_i = v] \\
&= 1 + (c-1) \sum_v v \Pr[T_i = v] \\
&= 1 + (c-1) \operatorname{Ex}[T_i] \\
&\le e^{(c-1)\operatorname{Ex}[T_i]}.
\end{aligned}$$

The first step uses the definition of expectation. The second step relies on the inequality $c^v \le 1 + (c-1)v$, which holds for all v in [0, 1] and $c \ge 1$. This follows from the general principle that a convex function, namely c^v, is less than the linear function, $1 + (c-1)v$, between their points of intersection, namely $v = 0$ and 1. This inequality is why the variables T_i are restricted to the interval [0, 1]. We then multiply out inside the summation and split into two sums. The first sum adds the probabilities of all possible outcomes, so it is equal to 1. After pulling the constant $c - 1$ out of the second sum, we're left with the definition of $\operatorname{Ex}[T_i]$. The final step uses the standard inequality $1 + z \le e^z$, which holds for all $z > 0$. ■

19.5 Mutually Independent Events

Suppose that we have a collection of mutually independent events $A_1, A_2, \ldots, A_n$, and we want to know how many of the events are likely to occur.

Let T_i be the indicator random variable for A_i and define

$$p_i = \Pr[T_i = 1] = \Pr\left[A_i\right]$$

for $1 \le i \le n$. Define

$$T = T_1 + T_2 + \cdots + T_n$$

to be the number of events that occur.

We know from Linearity of Expectation that

$$\begin{aligned}\operatorname{Ex}[T] &= \operatorname{Ex}[T_1] + \operatorname{Ex}[T_2] + \cdots + \operatorname{Ex}[T_n] \\ &= \sum_{i=1}^{n} p_i .\end{aligned}$$

This is true even if the events are *not* independent.

By Theorem 19.1.9, we also know that

$$\begin{aligned}\operatorname{Var}[T] &= \operatorname{Var}[T_1] + \operatorname{Var}[T_2] + \cdots + \operatorname{Var}[T_n] \\ &= \sum_{i=1}^{n} p_i(1-p_i),\end{aligned}$$

and thus that

$$\sigma_T = \sqrt{\sum_{i=1}^{n} p_i(1-p_i)}.$$

This is true even if the events are only pairwise independent.

Markov's Theorem tells us that for any $c > 1$,

$$\Pr[T \geq c \operatorname{Ex}[T]] \leq \frac{1}{c}.$$

Chebyshev's Theorem gives us the stronger result that

$$\Pr[|T - \operatorname{Ex}[T]| \geq c\sigma_T] \leq \frac{1}{c^2}.$$

The Chernoff Bound gives us an even stronger result; namely, that for any $c > 0$,

$$\Pr[T - \operatorname{Ex}[T] \geq c \operatorname{Ex}[T]] \leq e^{-(c\ln(c)-c+1)\operatorname{Ex}[T]}.$$

In this case, the probability of exceeding the mean by $c \operatorname{Ex}[T]$ decreases as an exponentially small function of the deviation.

By considering the random variable $n - T$, we can also use the Chernoff Bound to prove that the probability that T is much lower than $\operatorname{Ex}[T]$ is also exponentially small.

19.5.1 Murphy's Law

Suppose we want to know the probability that at least 1 event occurs. If $\text{Ex}[T] < 1$, then Markov's Theorem tells us that

$$\Pr[T \geq 1] \leq \text{Ex}[T].$$

On the other hand, if $\text{Ex}[T] \geq 1$, then we can obtain a lower bound on $\Pr[T \geq 1]$ using a result that we call Murphy's Law[8].

Theorem 19.5.1 (Murphy's Law)**.** *Let $A_1, A_2, \ldots, A_n$ be mutually independent events. Let T_i be the indicator random variable for A_i and define*

$$T ::= T_1 + T + 2 + \cdots + T_n$$

to be the number of events that occur. Then

$$\Pr[T = 0] \leq e^{-\text{Ex}[T]}.$$

Proof.

$$\begin{aligned}
\Pr[T = 0] &= \Pr[\overline{A}_1 \wedge \overline{A}_2 \wedge \cdots \wedge \overline{A}_n] \\
&= \prod_{i=1}^{n} \Pr[\overline{A}_i] && \text{(by independence of } A_i\text{)} \\
&= \prod_{i=1}^{n} (1 - \Pr[A_i]) \\
&\leq \prod_{i=1}^{n} e^{-\Pr[A_i]} && \text{(since } \forall x. 1 - x \leq e^{-x}\text{)} \\
&= e^{-\sum_{i=1}^{n} \Pr[A_i]} \\
&= e^{-\sum_{i=1}^{n} \text{Ex}[T_i]} && \text{(since } T_i \text{ is an indicator for } A_i\text{)} \\
&= e^{-\text{Ex}[T]} && \text{(Linearity of Expectation)}
\end{aligned}$$

■

For example, given any set of mutually independent events, if you expect 10 of them to happen, then at least one of them will happen with probability at least $1 - e^{-10}$. The probability that none of them happen is at most $e^{-10} < 1/22000$.

So if there are a lot of independent things that can go wrong and their probabilities sum to a number much greater than 1, then Theorem 19.5.1 proves that some of them surely will go wrong.

[8] This is in reference and deference to the famous saying that "If something can go wrong, it will go wrong."

This result can help to explain "coincidences," "miracles," and crazy events that seem to have been very unlikely to happen. Such events do happen, in part, because there are so many possible unlikely events that the sum of their probabilities is greater than one. For example, someone *does* win the lottery.

In fact, if there are 100,000 random tickets in Pick-4, Theorem 19.5.1 says that the probability that there is no winner is less than $e^{-10} < 1/22000$. More generally, there are literally millions of one-in-a-million possible events and so some of them will surely occur.

19.5.2 Another Magic Trick

Theorem 19.5.1 is surprisingly powerful. In fact, it is so powerful that it can enable us to read your mind. Here's how.

You choose a secret number n from 1 to 9. Then we randomly shuffle an ordinary deck of 52 cards and display the cards one at a time. You watch as we reveal the cards and when we reveal the nth card, that card becomes your *secret card*. If the card is an Ace, a 10, or a face card, then you assign that card a *value* of 1. Otherwise, you assign that card a value that is its number. For example, the $J\heartsuit$ gets assigned a value $v_1 = 1$ and the $4\diamondsuit$ gets assigned a value $v_1 = 4$. You do all of this in your mind so that we can't tell when the nth card shows up.

We keep revealing the cards, and when the $(n + v_1)$th card shows up, that card becomes your *new* secret card. You compute its value v_2 using the same scheme as for v_1. For example, if your new secret card is the $10\clubsuit$, then $v_2 = 1$. The $(n + v_1 + v_2)$th card will then become your next secret card, and so forth.

We proceed in this fashion until all 52 cards have been revealed, whereupon we read your mind by predicting your last secret card! How is this possible?

For the purposes of illustration, suppose that your secret number was $n = 3$ and the deck consisted of the 11 cards:

$$3\diamondsuit \quad 5\spadesuit \quad 2\diamondsuit \quad 3\clubsuit \quad 10\clubsuit \quad Q\diamondsuit \quad 3\heartsuit \quad 7\spadesuit \quad 6\clubsuit \quad 4\diamondsuit \quad 2\heartsuit.$$

Then your secret cards would be

$$2\diamondsuit,\ 10\clubsuit,\ Q\diamondsuit,\ 3\heartsuit,\ 4\diamondsuit$$

since $v_1 = 2,\ v_2 = 1,\ v_3 = 1,\ v_4 = 3$, and $v_5 = 4$. In this example, your last secret card is the $4\diamondsuit$.

To make the trick work, we follow the same rules as you, except that we start with $n = 1$. With the 11-card deck shown above, our secret cards would be

$$3\diamondsuit,\ 3\clubsuit,\ 3\heartsuit,\ 4\diamondsuit.$$

We have the same last secret card as you do! That is *not* a coincidence. In fact, this is how we predict your last card—we just guess that it is the same as our last card. And, we will be right with probability greater than 90%.

To see why the trick is likely to work, you need to notice that if we ever share a secret card, then we will surely have the same *last* secret card. That's because we will perform exactly the same steps as the cards are revealed.

Each time we get a new secret card, there is always a chance that it was one of your secret cards. For any given step, the chance of a match is small but we get a lot of chances. In fact, the number of chances will typically outweigh the inverse of the probability of a match on any given step and so, at least informally, Murphy's Law suggests that we are likely to eventually get a match, whereupon we can read your mind.

The details of the proof are complicated and we will not present them here. One of the main complications is that when you are revealing cards from a deck without replacement, the probability of getting a match on a given step is conditional based on the cards that have already been revealed.

19.5.3 The Subprime Mortgage Disaster

Throughout the last few chapters, we have seen many examples where powerful conclusions can be drawn about a collection of events if the events are independent. Of course, such conclusions are totally invalid if the events have dependencies. Unforeseen dependencies can result in disaster in practice. For example, misguided assumptions about the independence of loans (combined with a large amount of greed) triggered the global financial meltdown in 2008–2009.

In what follows, we'll explain some of what went wrong. You may notice that we have changed the names of the key participants. That is not to protect the innocent, since innocents are few and far between in this sordid tale. Rather, we changed the names to protect ourselves.[9] In fact, just to be on the safe side, we'll forget about what really happened here on Earth and instead tell you a fairy tale that took place in a land far, far away.

The central players in our story are the major Wall Street firms, of which Golden Scoundrels (commonly referred to as "Golden") is the biggest and most aggressive. Firms such as Golden ostensibly exist to make markets; they purport to create an open and orderly market in which buyers and sellers can be brought together and through which capitalism can flourish. It all sounds good, but the fees that can be had from facilitating transactions in a truly open and orderly market are often just not enough to satisfy the ever-increasing need to make more. So the employees at

[9]For a much more detailed accounting of these events (and one that does name names), you may enjoy reading *The Big Short* by Michael Lewis.

such firms are always trying to figure out a way to create new opportunities to make even more money.

One day, they came up with a whopper. Suppose they bought a collection of 1000 (say) subprime mortgage loans from all around the country and packaged them up into a single entity called a *bond*. A *mortgage loan* is a loan to a homeowner using the house as collateral; if the homeowner stops paying on the loan (in which case the loan is said to be in *default*), then the owner of the loan takes ownership of the house. A mortgage loan is classified as *subprime* if the homeowner does not have a very good credit history. Subprime loans are considered to be more risky than *prime* loans since they are more likely to default. Defaults are bad for everyone; the homeowner loses the home and the loan owner gets stuck trying to sell the house, which can take years and often results in very high losses.

Of course, a bond consisting of 1000 subprime loans doesn't sound very appealing to investors, so to dress it up, Golden sells the bond in *tranches*. The idea behind the tranches is to provide a way to assign losses from defaults. In a typical scenario, there would be 10 tranches and they are prioritized from 1 to 10. The defaults are assessed against the lowest tranches first. For example, suppose that there were 150 defaults in the collection of 1000 loans (an impossibly high number of defaults according to Golden). Then the lowest tranche would absorb the first 100 defaults (effectively wiping them out since all 100 of "their" loans would be in default) and the second-lowest tranche would be assigned the next 50 defaults, (wiping out half of their investment). The remaining 8 tranches would be doing great—none of "their" loans would be in default.

Because they are taking on more risk, the lower tranches would get more of the interest payments. The top tranche would get the lowest rate of return and would also be the safest. The lowest tranche would get the most interest, but also be the most exposed.

But how much should you pay for a tranche? Suppose the probability that any given loan defaults in a year is 1%. In other words, suppose you expect 10 of the 1000 loans to default in each year. If the defaults are independent, then we can use the Chernoff bound to conclude that the chances of more than 100 defaults (10%) in the 1000-loan collection is exceedingly tiny. This means that every tranche but the lowest is essentially risk-free. That is excellent news for Golden since they can buy 1000 cheap[10] subprime loans and then sell the top 9 tranches at premium rates, thereby making a large and instant profit on 900 of the 1000 loans. It is like turning a bunch of junk into a bunch of gold with a little junk left over.

There remains the problem of the lowest tranche, which is expected to have 10 defaults in a pool of 100 loans for a default rate of 10%. This isn't so good

[10] They are *subprime* loans after all.

so the first thing to do is to give the tranche a better sounding name than "lowest tranche." "Mezzanine" tranche sounds much less ominous and so that is what they used.

By the Chernoff bound, the default rate in the Mezzanine tranche is very unlikely to be much greater than 10%, and so the risk of owning this tranche can be addressed in part by increasing the interest payments for the tranche by 10%. But Golden had an even better idea (whopper number two)—rather than pay the extra 10%, why not collect together a bunch of mezzanine tranches from a bunch of bonds and then package them together into a "super bond" and then create tranches in the super-bond? The technical name for such a super bond is a *collateralized debt obligation* or CDO. This way, 90% of the mezzanine tranches instantly became essentially "risk-free," or so Golden claimed as they were marketing them.

The only problem now is getting the pension funds and other big investors to buy the CDOs at the same price as if they were AAA-rated "risk-free" bonds. This was a little tricky because 1) it was virtually impossible for the buyer to figure out exactly what loans they were effectively buying since they were buying a tranche of a collection of tranches, and 2) if you could ever figure out what it was, you would discover that it was the junk of the junk when it comes to loans.

The solution was to enlist the help of the big bond-rating agencies: Substandard and Prevaricators (S&P) and Mopey's. If Golden could get AAA ratings[11] on their tranches, then the pension funds and other big investors would buy them at premium rates.

It turned out to be easier than you might think (or hope) to convince S&P and Mopey's to give high ratings to the CDO tranches. After all, the ratings agencies are trying to make money too and they make money by rating bonds. And Golden was only going to pay them if their bonds and CDOs got good ratings. And, since defaults were assumed to be essentially independent, there was a good argument as to why all but the mezzanine tranche of a bond or CDO would be essentially risk-free.[12]

So the stage is set for Golden to make a bundle of money. Cheap junk loans come in the back door and exit as expensive AAA-rated bonds and CDOs out the front door. The remaining challenge is to ramp up the new money-making machine. That

[11] AAA ratings are the best you can get and are supposed to imply that there is virtually no chance of default.

[12] The logic gets a little fuzzy when you keep slicing and dicing the tranches—after a few iterations, you should be able to conclude that the mezzanine tranche of a CDO is sure to have 100% defaults, but it required effort to see what was going on under the covers and effort costs money, and so the ratings agencies considered the risk of the mezzanine tranche of one CDO to be the same as the mezzanine tranche of any other, even though they could have wildly different probabilities of sustaining large numbers of defaults.

means creating more (preferably, many, many more) junk loans to fuel the machine.

This is where Joe enters the scene. Joe is a migrant laborer earning $15,000 per year. Joe's credit history is not great (since he has never had a loan or credit card) but it is also not bad (since he has never missed a payment on a credit card and never defaulted on a loan). In short, Joe is a perfect candidate for a subprime mortgage loan on a $750,000 home.

When Loans-Я-Us approaches Joe for a home loan,[13] Joe dutifully explains that while he would love to own a $750,000 home, he doesn't have enough money to pay for food, let alone the interest payments on the mortgage. "No problem!" replies Loans-Я-Us. It is Joe's lucky day. The interest rates are super-low for the first 2 years and Joe can take out a second loan to cover them during that period. "What happens after 2 years?" Joe wants to know. "No problem!" replies Loans-Я-Us. Joe can refinance—his home will surely be worth more in 2 years. Indeed, Joe can even make money while he enjoys the comforts of his new home. If all goes well, he can even ease off on the laborer work, and maybe even by a second home. Joe is sold. In fact, millions of Joes are sold and, before long, the subprime loan business is booming.

It turns out that there were a few folks out there who really did their math homework when they were in college. They were running hedge funds and, as the money-making machine was cranking away, they realized that a disaster was looming. They knew that loan defaults are not independent—in fact, they are very dependent. Once home values stop rising, or a recession hits, or it comes time for Joe to refinance, defaults will occur at much higher rates than projected and the CDOs and many tranches of the underlying bonds will become worthless. And there is so much money invested in these bonds and CDOs that the economy could be ruined.

Unfortunately, the folks who figured out what was going to happen didn't alert anyone. They didn't go to the newspapers. They didn't call the See no Evil Commission. They didn't even call 911. Instead, they worked with Golden to find a new way to make even more money—betting against the CDO market.

If you think a stock is going to decline, you can profit from the decline by borrowing the stock and selling it. After the stock declines in value, you buy it back and return it to the person that lent it to you. Your profit is the decline in price. This process is called *shorting* the stock.

So the hedge funds wanted to short the CDOs. Unfortunately, there was no established way to borrow a tranche of a CDO. Always looking for a new way to make money, the investment houses came up with an even bigger whopper than the

[13] Yes, we know it is supposed to go the other way around—Joe is supposed to approach the loan company—but these are extraordinary times.

CDO—they invented the *credit default swap*.

The idea behind the credit default swap is to provide a kind of insurance against the event that a bond or CDO suffers a certain number of defaults. Since the hedge funds believe that the CDOs were going to have lots of defaults, they want to buy the insurance. The trick is to find someone dumb enough to sell the insurance. That's where the world's largest insurance company, Awful Insurance Group (AIG), enters the fray. AIG sells insurance on just about anything and they, too, are looking for new ways to make money, so why not sell insurance on CDO defaults?

Golden has a new business! They buy the CDO insurance from AIG for an astonishingly low price (about $2 annually for every $1000 of CDO value) and sell it to the hedge funds for a much higher price (about $20 annually for every $1000 of CDO value). If a CDO sustains defaults, then AIG needs to pay the value of the CDO ($1000 in this hypothetical example) to the hedge funds who own the insurance. Until that time, the hedge funds are paying the annual fee for the insurance, 90% of which is pocketed by Golden. This is a great business; Golden pockets 90% of the money and AIG takes all the risk. The only risk that Golden has is if AIG goes down, but AIG is "too big to fail. . . ."

Golden's new credit default swap business is even better than the CDO business. The only trouble now is that there are only so many Joes out there who can take out subprime loans. This means that there is a hard limit on how many billions Golden can make. This challenge led to whopper number four.

If the hedge funds want to buy insurance and AIG wants to sell it, who really cares if there is only one insurance policy per loan or CDO? Indeed, why not just sell lots of credit default swaps on the same set of junk CDOs? This way, the profits could be unlimited! And so it went. "Synthetic" CDOs were created and soon the "insurance" quickly turned into a very high-stakes (and very stupid, at least for AIG) bet. The odds were weighted heavily in favor of the folks who did their math homework (the hedge funds); the hedge funds had figured out that the failure of the CDOs was a virtual certainty, whereas AIG believed that failure was virtually impossible.

Of course, we all know how the story ends. The holders of the CDOs and subprime debt and the sellers of insurance got wiped out, losing hundreds of billions of dollars. Since many of these folks were deemed by the Government as "too big to fail," they were bailed out using nearly a trillion dollars of taxpayer money. The executives who presided over the disaster were given huge bonuses because, well, that's how it works for executives in the land far, far away. The story also ends well for the hedge funds that bought the insurance—they made many, many billions of dollars.

So everyone involved in the disaster ends up very rich. Everyone except Joe, of

course. Joe got kicked out of his home and lost his job in the recession.

Too bad for Joe that it isn't just a fairy tale.

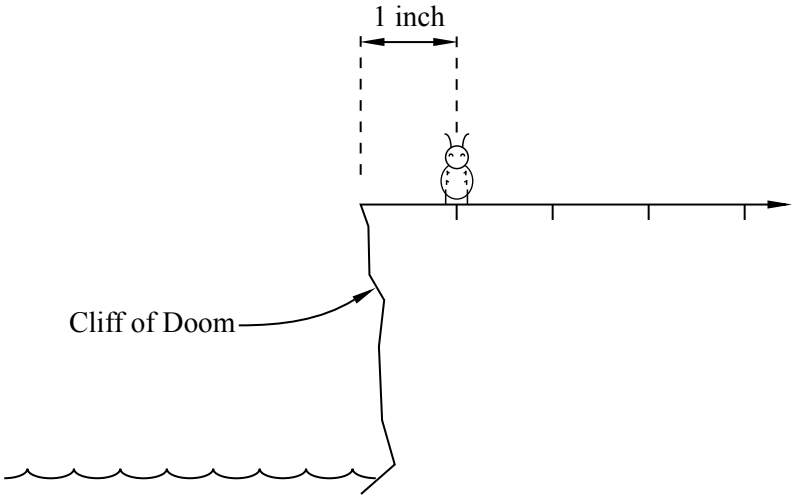
20 Random Walks

Random Walks are used to model situations in which an object moves in a sequence of steps in randomly chosen directions. Many phenomena can be modeled as a random walk and we will see several examples in this chapter. Among other things, we'll see why it is rare that you leave the casino with more money than you entered with and we'll see how the Google search engine uses random walks through the graph of the world-wide web links to determine the relative importance of websites.

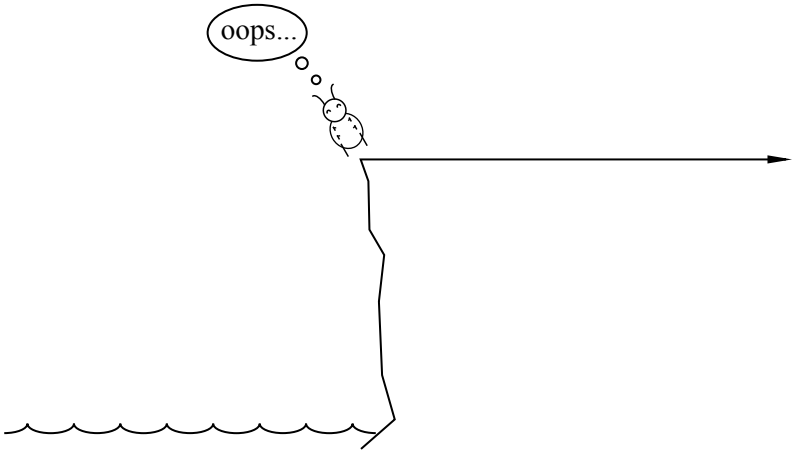
20.1 Unbiased Random Walks

20.1.1 A Bug's Life

There is a small flea named Stencil. To his right, there is an endless flat plateau. One inch to his left is the Cliff of Doom, which drops to a raging sea filled with flea-eating monsters.



Each second, Stencil hops 1 inch to the right or 1 inch to the left with equal probability, independent of the direction of all previous hops. If he ever lands on the very edge of the cliff, then he teeters over and falls into the sea. So, for example, if Stencil's first hop is to the left, he's fishbait. On the other hand, if his first few hops are to the right, then he may bounce around happily on the plateau for quite



some time.

Our job is to analyze the life of Stencil. Does he have any chance of avoiding a fatal plunge? If not, how long will he hop around before he takes the plunge?

Stencil's movement is an example of a *random walk*. A typical *one-dimensional* random walk involves some value that randomly wavers up and down over time. The walk is said to be *unbiased* if the value is equally likely to move up or down. If the walk ends when a certain value is reached, then that value is called a *boundary condition* or *absorbing barrier*. For example, the Cliff of Doom is a boundary condition in the example above.

Many natural phenomena are nicely modeled by random walks. However, for some reason, they are traditionally discussed in the context of some social vice. For example, the value is often regarded as the position of a drunkard who randomly staggers left, staggers right, or just wobbles in place during each time step. Or the value is the wealth of a gambler who is continually winning and losing bets. So discussing random walks in terms of fleas is actually sort of elevating the discourse.

20.1.2 A Simpler Problem

Let's begin with a simpler problem. Suppose that Stencil is on a small island; now, not only is the Cliff of Doom 1 inch to his left, but also there is another boundary condition, the Pit of Disaster, 2 inches to his right! For example, see Figure 20.1

In the figure, we've worked out a tree diagram for Stencil's possible fates. In

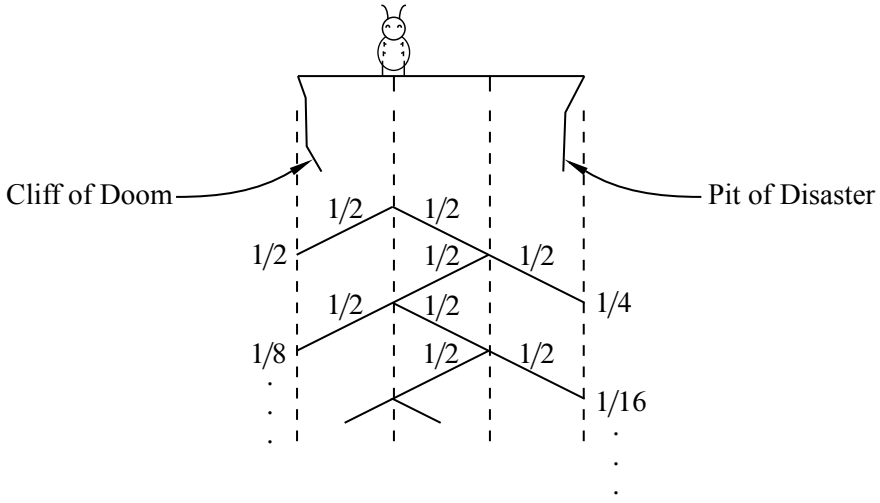


Figure 20.1 An unbiased, one-dimensional random walk with absorbing barriers at positions 0 and 3. The walk begins at position 1. The tree diagram shows the probabilities of hitting each barrier.

particular, he falls off the Cliff of Doom on the left side with probability:

$$\begin{aligned}\frac{1}{2}+\frac{1}{8}+\frac{1}{32}+\ldots &= \frac{1}{2}\left(1+\frac{1}{4}+\frac{1}{16}+\ldots\right)\\ &= \frac{1}{2}\cdot\frac{1}{1-1/4}\\ &= \frac{2}{3}.\end{aligned}$$

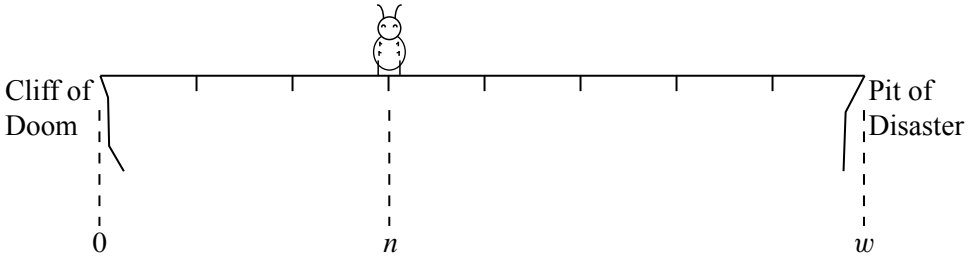
Similarly, he falls into the Pit of Disaster on the right side with probability:

$$\frac{1}{4}+\frac{1}{16}+\frac{1}{64}+\ldots=\frac{1}{3}.$$

There is a remaining possibility: Stencil *could* hop back and forth in the middle of the island forever. However, we've already identified two disjoint events with probabilities $2/3$ and $1/3$, so this happy alternative must have probability 0.

20.1.3 A Big Island

Putting Stencil on such a tiny island was sort of cruel. Sure, he's probably carrying bubonic plague, but there's no reason to pick on the little fella. So suppose that we instead place him n inches from the left side of an island w inches across: In



other words, Stencil starts at position n and his random walk ends if he ever reaches positions 0 or w.

Now he has three possible fates: he could fall off the Cliff of Doom, fall into the Pit of Disaster, or hop around on the island forever. We could compute the probabilities of these three events with a horrific summation, but fortunately there's a far easier method: we can use a linear recurrence.

Let R_n be the probability that Stencil falls to the right into the Pit of Disaster, given that he starts at position n. In a couple special cases, the value of R_n is easy to determine. If he starts at position w, he falls into the Pit of Disaster immediately, so $R_w = 1$. On the other hand, if he starts at position 0, then he falls from the Cliff of Doom immediately, so $R_0 = 0$.

Now suppose that our frolicking friend starts somewhere in the middle of the island; that is, $0 < n < w$. Then we can break the analysis of his fate into two cases based on the direction of his first hop:

- If his first hop is to the left, then he lands at position $n - 1$ and eventually falls into the Pit of Disaster with probability R_{n-1}.
- On the other hand, if his first hop is to the right, then he lands at position $n+1$ and eventually falls into the Pit of Disaster with probability R_{n+1}.

Therefore, by the Total Probability Theorem, we have:

$$R_n = \frac{1}{2}R_{n-1} + \frac{1}{2}R_{n+1}.$$

Solving the Recurrence

Let's assemble all our observations about R_n, the probability that Stencil falls into the Pit of Disaster if he starts at position n:

$$\begin{aligned} R_0 &= 1 \\ R_w &= 0 \\ R_n &= \frac{1}{2}R_{n-1} + \frac{1}{2}R_{n+1} \qquad (0 < n < w). \end{aligned}$$

This is just a linear recurrence—and we know how to solve those! Uh, right? Remember Chapter 10 or Chapter 12?

There is one unusual complication: in a normal recurrence, R_n is written a function of preceding terms. In this recurrence equation, however, R_n is a function of both a preceding term (R_{n-1}) and a *following* term (R_{n+1}). This is no big deal, however, since we can just rearrange the terms in the recurrence equation:

$$R_{n+1} = 2R_n - R_{n-1}.$$

Now we're back on familiar territory.

Let's solve the recurrence. The characteristic equation is:

$$x^2 - 2x + 1 = 0.$$

This equation has a double root at $x = 1$. There is no inhomogeneous part, so the general solution has the form:

$$R_n = a \cdot 1^n + b \cdot n1^n = a + bn.$$

Substituting in the boundary conditions $R_0 = 0$ and $R_w = 1$ gives two linear equations:

$$\begin{aligned} 0 &= a, \\ 1 &= a + bw. \end{aligned}$$

The solution to this system is $a = 0$, $b = 1/w$. Therefore, the solution to the recurrence is:

$$R_n = n/w.$$

20.1.4 Death Is Certain

Our analysis shows that if we place Stencil n inches from the left side of an island w inches across, then he falls off the right side with probability n/w. For example, if Stencil is $n = 4$ inches from the left side of an island $w = 12$ inches across, then he falls off the right side with probability $n/w = 4/12 = 1/3$.

We can compute the probability that he falls off the *left* side by exploiting the symmetry of the problem: the probability that he falls off the *left* side starting at position n is the same as the probability that he falls of the *right* side starting at position $w - n$, which is $(w - n)/n$.

This is bad news. The probability that Stencil eventually falls off one side or the other is:

$$\frac{n}{w} + \frac{w-n}{w} = 1.$$

There's no hope! The probability that Stencil hops around on the island forever is zero.

And there's even worse news. Let's go back to the original problem where Stencil is 1 inch from the left edge of an *infinite* plateau. In this case, the probability that he eventually falls into the sea is:

$$\lim_{w\to\infty} \frac{w-1}{w} = 1.$$

So even if there were no Pit of Disaster, Stencil still falls off the Cliff of Doom with probability 1. And since

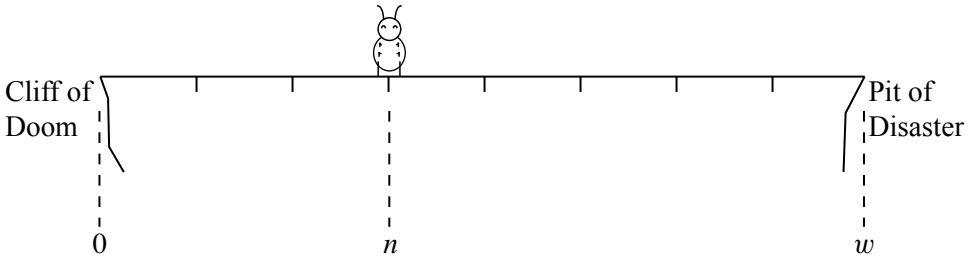
$$\lim_{w\to\infty} \frac{w-n}{w} = 1$$

for any finite n, this is true no matter where Stencil starts. Our little friend is doomed!

Hey, you know how in the movies they often make it look like the hero dies, but then he comes back in the end and everything turns out okay? Well, we're not sayin' anything, just pointing that out.

20.1.5 Life Expectancy

On the bright side, Stencil may get to hop around for a while before he goes over an edge. Let's use the same setup as before, where he starts out n inches from the left side of an island w inches across: What is the expected number of hops he takes



before falling off an edge?

Let X_n be Stencil's expected lifespan, measured in hops. If he starts at either edge of the island, then he dies immediately:

$$X_0 = 0,$$
$$X_w = 0.$$

If he starts somewhere in the middle of the island ($0 < n < w$), then we can again break down the analysis into two cases based on his first hop:

- If his first hop is to the left, then he lands at position $n-1$ and can expect to live for another X_{n-1} steps.
- If his first hop is to the right, then he lands at position $n+1$ and can expect to live for another X_{n+1} steps.

Thus, by the Law of Total Expectation and Linearity of Expectation, Stencil's expected lifespan is:

$$X_n = 1 + \frac{1}{2}X_{n-1} + \frac{1}{2}X_{n+1}.$$

The leading 1 accounts for his first hop.

Solving the Recurrence

Once again, Stencil's fate hinges on a recurrence equation:

$$\begin{array}{lll} X_0 & = 0 & \\ X_w & = 0 & \\ X_n & = 1 + \frac{1}{2}X_{n-1} + \frac{1}{2}X_{n+1} & (0 < n < w). \end{array}$$

We can rewrite the last line as:

$$X_{n+1} = 2X_n - X_{n-1} - 2. \tag{20.1}$$

As before, the characteristic equation is:

$$x^2 - 2x + 1 = 0.$$

There is a double-root at 1, so the homogeneous solution has the form:

$$X_n = a + bn.$$

But this time, there's an inhomogeneous term, so we also need to find a particular solution. Since this term is a constant, we should try a particular solution of the form $X_n = c$ and then try $X_n = c + dn$ and then $X_n = c + dn + en^2$ and so forth. As it turns out, the first two possibilities don't work, but the third does. Substituting $X_n = c + dn + en^2$ into Equation 20.1 gives

$$c + d(n+1) + e(n+1)^2 = 2(c + dn + en^2) - \big(c + d(n-1) + e(n-1)^2\big) - 2,$$

which simplifies to $e = -1$. Since all the c and d terms cancel, $X_n = c + dn - n^2$ is a particular solution for all c and d. For simplicity, let's take $c = d = 0$. Thus, our particular solution is $X_n = -n^2$.

Adding the homogeneous and particular solutions gives the general form of the solution:

$$X_n = a + bn - n^2.$$

Substituting in the boundary conditions $X_0 = 0$ and $X_w = 0$ gives two linear equations:

$$0 = a,$$
$$0 = a + bw - w^2.$$

The solution to this system is $a = 0$ and $b = w$. Therefore, the solution to the recurrence equation is:

$$X_n = wn - n^2 = n(w - n).$$

Interpreting the Solution

Stencil's expected lifespan is $X_n = n(w-n)$, which is the *product* of the distances to the two edges. Thus, for example, if he's 4 inches from the left edge and 8 inches from the right cliff, then his expected lifespan is $4 \cdot 8 = 32$.

Let's return to the original problem where Stencil has the Cliff of Doom 1 inch to his left and an infinite plateau to this right. (Also, cue the "hero returns" theme music.) In this case, his expected lifespan is:

$$\lim_{w\to\infty} 1(w-1) = \infty$$

Yes, Stencil is certain to eventually fall off the Cliff of Doom—but his expected lifespan is infinite! This sounds almost like a contradiction, but both answers are correct!

Here's an informal explanation. It turns out that the probability p_k that Stencil falls from the Cliff of Doom on the kth step is $\Theta(1/k^{3/2})$. You can verify by the integration bound that $\sum_{i=1}^{\infty} 1/k^{3/2}$ converges.

On the other hand, the expected time until Stencil falls over the edge is

$$\begin{aligned}\sum_{i=1}^{\infty} k p_k &\geq c \sum_{k=1}^{\infty} \frac{k}{k^{3/2}} \\ &= c \sum_{k=1}^{\infty} \frac{1}{\sqrt{k}} \\ &= \infty,\end{aligned}$$

where c is a constant that comes from the Θ notation. So our answers are compatible.

20.1.6 Application to Fair Gambling Games

We took the high road for a while, but let's now discuss random walks in a more conventional setting—gambling.

A gambler goes to Las Vegas with $\$n$ in her pocket. Her plan is to make only \$1 bets and somehow she has found a casino that will offer her truly even odds[1]; namely, she will win or lose \$1 on each bet with probability $1/2$. She'll play until she is broke or she has won $\$m$. In the latter case, she will go home with

$$w = n + m$$

dollars. What's the probability that she goes home a winner?

This is identical to the flea problem that we just analyzed. Going broke is analogous to falling off the Cliff of Doom. Going home a winner is analogous to falling into the Pit of Disaster, just a lot more fun.

Our analysis of Stencil's life tells us everything we want to know about the gambler's prospects:

- The gambler goes home broke with probability
$$\frac{n}{w} = \frac{m}{n+m},$$
- the gambler goes home a winner with probability
$$\frac{w-n}{w} = \frac{n}{n+m},$$
- the gambler goes home with probability
$$\frac{n}{n+m} + \frac{m}{n+m} = 1,$$
- and the number of bets before the gambler goes home is expected to be
$$n(w-n) = nm.$$

If the gambler gets greedy and keeps playing until she goes broke, then

- the gambler eventually goes broke with probability 1, and
- the number of bets before the gambler goes broke is expected to be infinite.

The bottom line here is clear: when gambling, quit while you are ahead—if you play until you go broke, you will certainly go broke.

And that's the good news! Matters get much worse for the more typical scenario where the odds are against you. Let's see why.

[1] Don't worry, we'll get to the more realistic scenario when she is more likely to lose than win in a moment, but let's just fantasize about the fair scenario for a bit.

20.2 Gambler's Ruin

So far, we have considered *unbiased* random walks, where the probability of moving up or down (or left or right) is $1/2$. Unfortunately, things are never quite this simple (or fair) in real casinos.

For example, suppose the gambler goes to Las Vegas and makes \$1 bets on red or black in roulette. In this case, she will win \$1 with probability

$$\frac{18}{38} \approx 0.473$$

and she will lose \$1 with probability

$$\frac{20}{38} \approx 0.527.$$

That's because the casinos add those bothersome green 0 and 00 to give the house a slight advantage.

At first glance (or after a few drinks), $18/38$ seems awfully close to $1/2$ and so our intuition tells us that the game is "almost fair." So we might expect the analysis we just did for the fair game to be "almost right" for the real game. For example, if the gambler starts with \$100 and quits when she gets ahead by \$100 in the fair game, then she goes home a winner with probability

$$\frac{100}{200} = .5.$$

And, if she wants to improve her chances of going home a winner, she could bring more money. If she brings \$1000 and quits when she gets ahead by \$100 in the fair game, then she goes home a winner with probability

$$\frac{1000}{1100} \approx .91.$$

So, given that the real game is "almost fair," we might expect the probabilities of going home a winner in these two scenarios to be "almost 50% and 91%," respectively.

Unfortunately for the gambler, all this "almost" reasoning will almost surely lead to disaster. Here are the grim facts for the real game where the gambler wins \$1 with probability $18/38$.

n = starting wealth	probability she reaches n + \$100 before \$0
\$100	1 in 37649.619496...
\$1000	1 in 37648.619496...
\$1,000,000,000	1 in 37648.619496...

Except on the very low end, the amount of money she brings makes almost no difference![2] She is almost certain to go broke before winning $100. Let's see why.

20.2.1 Finding a Recurrence

We can approach the gambling problem the same way we studied the life of Stencil. Suppose that the gambler starts with n dollars. She wins each bet with probability p and plays until she either goes bankrupt or has $w = n + m$ dollars in her pocket. (To be clear, w is the total amount of money she wants to end up with, not the amount by which she wants to increase her wealth, which is m.) Our objective is to compute R_n, the probability that she goes home a winner.

As usual, we begin by identifying some boundary conditions. If she starts with no money, then she's bankrupt immediately so $R_0 = 0$. On the other hand, if she starts with w dollars, then she's an instant winner, so $R_w = 1$.

Now we divide the analysis of the general situation into two cases based on the outcome of her first bet:

- She wins her first bet with probability p. She then has $n + 1$ dollars and probability R_{n+1} of reaching her goal of w dollars.
- She loses her first bet with probability $1 - p$. This leaves her with $n - 1$ dollars and probability R_{n-1} of reaching her goal.

Plugging these facts into the Total Probability Theorem gives the equation:

$$R_n = pR_{n+1} + (1-p)R_{n-1}. \tag{20.2}$$

20.2.2 Solving the Recurrence

Rearranging the terms in Equation 20.2 gives us a recurrence for R_n, the probability that the gambler reaches her goal of w dollars if she starts with n:

$$\begin{aligned} R_0 &= 0 \\ R_w &= 1 \\ pR_{n+1} - R_n + (1-p)R_{n-1} &= 0 \qquad (0 < n < w). \end{aligned}$$

The characteristic equation is:

$$px^2 - x + (1-p) = 0.$$

[2] The fact that only one digit changes from the first case to the second is a peripheral bit of bizarreness that we'll leave in your hands.

The quadratic formula gives the roots:

$$\begin{aligned} x &= \frac{1 \pm \sqrt{1-4p(1-p)}}{2p} \\ &= \frac{1 \pm \sqrt{(1-2p)^2}}{2p} \\ &= \frac{1 \pm (1-2p)}{2p} \\ &= \frac{1-p}{p} \text{ or } 1. \end{aligned}$$

There's an important point lurking here. If the gambler is equally likely to win or lose each bet, then $p = 1/2$, and the characteristic equation has a double root at $x = 1$. This is the situation we considered in the flea problem. The double root led to a general solution of the form:

$$R_n = a + bn$$

Now suppose that the gambler is *not* equally likely to win or lose each bet; that is, $p \neq 1/2$. Then the two roots of the characteristic equation are different, which means that the solution has a completely different form:

$$R_n = a \cdot \left(\frac{1-p}{p}\right)^n + b \cdot 1^n$$

In mathematical terms, this is where the fair game and the "almost fair" game take off in completely different directions: in one case we get a linear solution and in the other we get an exponential solution! This is going to be bad news for anyone playing the "almost fair" game.

Anyway, substituting the boundary conditions into the general form of the solution gives a system of linear equations:

$$\begin{aligned} 0 &= a + b \\ 1 &= a \cdot \left(\frac{1-p}{p}\right)^w + b. \end{aligned}$$

Solving this system, gives:

$$a = \frac{1}{\left(\frac{1-p}{p}\right)^w - 1}, \qquad b = -\frac{1}{\left(\frac{1-p}{p}\right)^w - 1}.$$

Substituting these values back into the general solution gives:

$$R_n = \left(\frac{1}{\left(\frac{1-p}{p}\right)^w - 1}\right) \cdot \left(\frac{1-p}{p}\right)^n - \frac{1}{\left(\frac{1-p}{p}\right)^w - 1}$$
$$= \frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}.$$

(Suddenly, Stencil's life doesn't seem so bad, huh?)

20.2.3 Bad News!

We have an answer! But it's not good news. If the gambler starts with n dollars and wins each bet with probability p, then the probability she reaches w dollars before going broke is:

$$\frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}.$$

Let's try to make sense of this expression. If the game is biased against her, as with roulette, then $1-p$ (the probability she loses) is greater than p (the probability she wins). If n, her starting wealth, is also reasonably large, then both exponentiated fractions are big numbers and the -1's don't make much difference. Thus, her probability of reaching w dollars is very close to:

$$\left(\frac{1-p}{p}\right)^{n-w} = \left(\frac{1-p}{p}\right)^m.$$

In particular, if she is hoping to come out $m = \$100$ ahead in roulette, then $p = 18/38$ and her probability of success is:

$$\left(\frac{10}{9}\right)^{-100} = 1 \text{ in } 37648.619496.$$

This explains the strange number we arrived at earlier! In fact, this number does not change no matter how large n gets, so even if the gambler starts with a trillion dollars, she is sill not likely to ever get ahead by even \$100.

20.2.4 But Why?

Why does the gambler's starting wealth have so little impact on her probability of coming out ahead? Intuitively, there are two forces at work. First, the gambler's

wealth has random upward and downward *swings* due to runs of good and bad luck. Second, her wealth has a steady, downward *drift* because she has a small expected loss on every bet. The situation is illustrated in Figure 20.2.

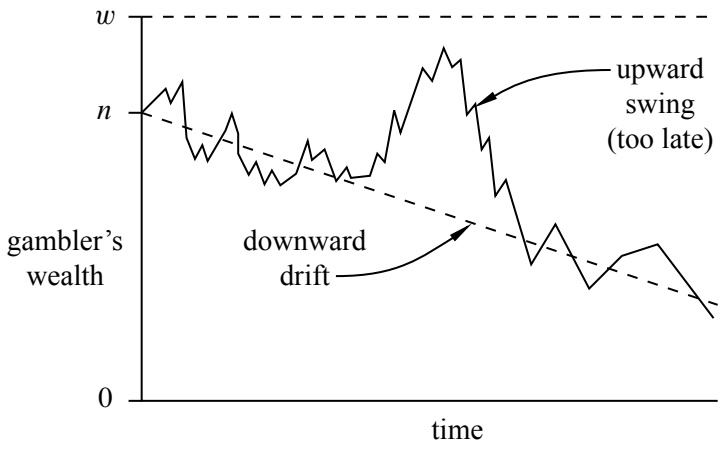


Figure 20.2 In a biased random walk, the downward drift usually dominates swings of good luck.

For example, in roulette, the gambler wins a dollar with probability $18/38$ and loses a dollar with probability $20/38$. Therefore, her expected return on each bet is

$$1 \cdot \frac{18}{38} + (-1) \cdot \frac{20}{38} = \frac{-2}{38} = -\frac{1}{19}.$$

Thus, her expected wealth drifts downward by a little over 5 cents per bet.

One might think that if the gambler starts with a billion dollars, then she will play for a long time, so at some point she should have a lucky, upward swing that puts her \$100 ahead. The problem is that her capital is steadily drifting downward. And after her capital drifts down a few hundred dollars, she needs a huge upward swing to save herself. And such a huge swing is extremely improbable. So if she does not have a lucky, upward swing early on, she's doomed forever. As a rule of thumb, *drift* dominates *swings* over the long term.

20.2.5 Expected Playing Time

Even though casino gamblers are destined to lose, some of them enjoy the process. So let's figure out how long their enjoyment is expected to last.

Let X_n be the expected number of bets before going home (broke or a winner).

Reasoning as in Section 20.1.5, we can set up a recurrence for X_n:

$$\begin{aligned} X_0 &= 0, \\ X_w &= 0, \\ X_n &= 1 + (1-p)X_{n-1} + pX_{n+1}. \end{aligned} \tag{20.3}$$

This is the same as the recurrence for R_n in Equation 20.2 except for the inhomogeneous part.

To find the particular solution, we try $X_n = c$ (which doesn't work) and then $X_n = c + dn$ (which does work as long as $p \neq 1/2$). Plugging $X_n = c + dn$ into Equation 20.3 yields:

$$\begin{aligned} c + dn &= 1 + (1-p)(c + d(n-1)) + p(c + d(n+1)) \\ &= 1 + c + dn - (1-p)d + pd \end{aligned}$$

and thus that

$$d = \frac{1}{1-2p}.$$

Since c is arbitrary, we will set $c = 0$ and our particular solution is

$$X_n = \frac{n}{1-2p}.$$

The characteristic equation for Equation 20.3 is

$$px^2 - x + (1-p) = 0.$$

We have already determined that the roots for this equation are

$$\frac{1-p}{p} \quad \text{and} \quad 1.$$

Hence, the general solution to the recurrence is

$$X_n = a\left(\frac{1-p}{p}\right)^n + b + \frac{n}{1-2p}.$$

Plugging in the boundary conditions, we find that

$$\begin{aligned} 0 &= a + b, \\ 0 &= a\left(\frac{1-p}{p}\right)^w + b + \frac{w}{1-2p}. \end{aligned}$$

Hence

$$a = \frac{-\left(\frac{w}{1-2p}\right)}{\left(\frac{1-p}{p}\right)^w - 1} \qquad \text{and} \qquad b = \frac{\left(\frac{w}{1-2p}\right)}{\left(\frac{1-p}{p}\right)^w - 1}.$$

The final solution to the recurrence is then

$$X_n = \frac{-\left(\frac{w}{1-2p}\right)\left(\frac{1-p}{p}\right)^n}{\left(\frac{1-p}{p}\right)^w - 1} + \frac{\left(\frac{w}{1-2p}\right)}{\left(\frac{1-p}{p}\right)^w - 1} + \frac{n}{1-2p}$$

$$= \frac{n}{1-2p} - \left(\frac{w}{1-2p}\right)\left[\frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}\right].$$

Yikes! The gambler won't have any fun at all if she is thinking about this equation. Let's see if we can make it simpler in the case when $m = w - n$ is large.

Since $p < 1/2$, $(1-p)/p > 1$ and for large m,

$$\lim_{m\to\infty}\left(\frac{w}{1-2p}\right)\left[\frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^w - 1}\right] = \lim_{m\to\infty}\left(\frac{w}{1-2p}\right)\left(\frac{1-p}{p}\right)^{-m} = 0.$$

This means that as m gets large,

$$X_n \sim \frac{n}{1-2p},$$

which is much simpler. It says that if the gambler starts with $\$n$, she will expect to make about $n/(1-2p)$ bets before she goes home broke. This seems to make sense since she expects to lose

$$1\cdot(1-p) + (-1)p = 1-2p$$

dollars on every bet and she started with n dollars.[3]

[3] Be careful, it is tempting to use such a direct and simple argument instead of all those nasty recurrences, but such an argument is not correct. There are examples where the expected duration of a process is not close to the starting point divided by the expected decrease at each step.

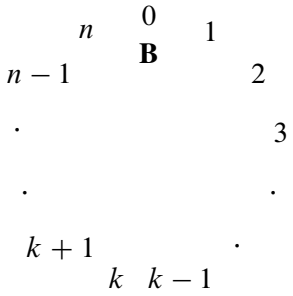


Figure 20.3 $n+1$ people sitting in a circle. The B indicates the person with the broccoli—in this case, person 0.

20.3 Walking in Circles

So far, we have considered random walks on a line. Now we'll look at a problem where the random walk is on a circle. Going from a line to a circle may not seem like such a big change, but as we have seen so often with probability, small changes can have large consequences that are often beyond the grasp of our intuition.

20.3.1 Pass the Broccoli

Suppose there are $n+1$ people, numbered $0, 1, \ldots, n$, sitting in a circle as shown in Figure 20.3. The B in Figure 20.3 indicates that person 0 has a big stalk of nutritious broccoli, which provides 250% of the US recommended daily allowance of vitamin C and is also a good source of vitamin A and iron. (Typical for a random walk problem, this game originally involved a pitcher of beer instead of a broccoli. We're taking the high road again.)

Person 0 passes the broccoli either to the person on his left or the person on his right with equal probability. Then, that person also passes the broccoli left or right at random and so on. After a while, everyone in an arc of the circle has touched the broccoli and everyone outside that arc has not. Eventually, the arc grows until all but one person has touched the broccoli. That final person is declared the winner because they have avoided the brocolli for the longest time.

Suppose that you are allowed to position yourself anywhere in the circle. Where should you stand in order to maximize the probability that you win? You shouldn't be person 0; you can't win in that position. The answer is "intuitively obvious": you should sit as far as possible from person 0, which would be position $n/2$ or $(n+1)/2$ depending on whether n is even or odd.

20.3.2 There Is No Escape

Let's try to verify this intuition. Suppose that you sit at position $k \neq 0$. At some point, the broccoli is going to end up in the hands of one of your neighbors. This has to happen eventually; the game can't end until at least one of them touches it. Let's say that person $k-1$ gets the broccoli first. Now let's cut the circle between yourself and your other neighbor, person $k+1$:

$$\underset{\mathbf{B}}{k} \quad (k-1) \quad \ldots \quad 3 \quad 2 \quad 1 \quad 0 \quad n \quad (n-1) \quad \ldots \quad (k+1).$$

There are two possibilities. If the broccoli reaches you before it reaches person $k+1$, then you lose. But if the broccoli reaches person $k+1$ before it reaches you, then every other person has touched the broccoli and you win. So we need to compute the probability that the broccoli hops $n-1$ people to the right before it takes 1 hop to the left. This will be the probability that you win.

But this is just the flea problem all over again. From the analysis in Section 20.1.3, we know that the probability of moving $n-1$ steps rightward before moving one step leftward is simply $1/n$. This means that wherever you sit (aside from position 0, of course), your probability of getting the broccoli last is $1/n$.

So our intuition was completely wrong (again)! It doesn't matter where you sit. Being close to the broccoli or far away at the start makes no difference; there is no escape—you still get the broccoli last with probability $1/n$.

Enough with the bad news: Stencil's doomed, you go home broke from the casino, and you can't escape the broccoli. Let's see how to use probability to *make* some money.