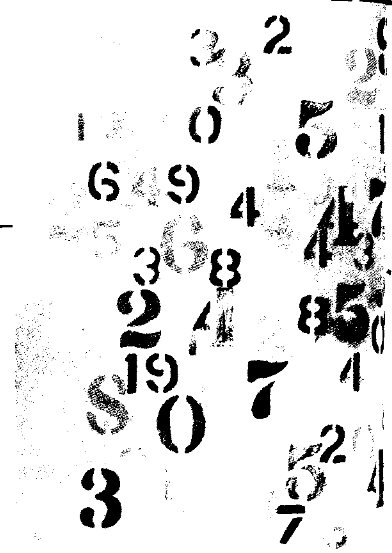
Chapter 4

Collections



In this chapter we're going to introduce a new, fundamental idea in counting. This will also be the last new formula: using this and the ideas we've already introduced in combination, we'll be able to count all the objects we want, at least until the final (and optional) chapter of this part of the book.

There's nothing mysterious about it. Basically, in the last couple of chapters we've considered a range of problems in which we count the number of ways to make a sequence of choices. In each instance, either the choices were made from different collections of objects (shirts and pants; meat and vegetable toppings on our pizza; action thrillers and light romantic comedies) or, if they were selections made from the same collection of objects, the order mattered: when we're counting four-letter words, "POOL" is not the same as "POLO."

What we want to look at now are situations where we choose a collection of objects from the same pool, and *the order doesn't matter*. We'll start by revisiting some of the problems we've dealt with, and show how slight variations will put us in this kind of situation.

4.1 Back to the House of Pizza

It's a new day, and once more you head over to the House of Pizza for lunch. Today, though, you're feeling both hungry and carnivorous: a pizza with three meat toppings sounds about right. Assuming that the HoP is still offering seven meat toppings, how many different pizzas will fit the bill?

On to the video store. Your roommates have once more selected you to go forth and rent the evening's entertainment; this night everyone's in the mood for a festival of movies based on comic strips or video games, and your mission is to return with four of the 23 such movies that the House of Videos carries. How many different choices do you have at the video store?

Finally, back to our class of 15 students. This time, we're not going to select officers: we're just going to choose a committee of four students. There are no distinctions among the four members of the committee; we just have to select four students from among the 15 in the class. How many different committees can be formed?

You get the idea? In each case we're choosing a collection of a specified number of objects from a common pool (toppings, videos, students), and the order doesn't matter: ordering a pizza with sausage, pepperoni and hamburger gets you pretty much the same pizza as ordering one with hamburger, pepperoni and sausage. This sort of situation comes up constantly: when you're dealt a hand of five cards in draw poker, or of 13 cards in bridge, it doesn't matter in what order you receive the cards; possible hands consist of collections of five or 13 cards out of the deck of 52. By way of language, in this sort of situation—where we make a series of selections from a common pool, but all that matters is the totality of objects selected, not the order in which they're selected—we'll refer to choosing a *collection* of objects. In settings where the selections are made from different pools, or where the order does matter, we'll refer to choosing a *sequence*.

Now, as you'll recognize, all of the problems above are really the same problem with different numbers substituted. In fact, there are only two numbers involved: in each of these cases, the number of possible choices depends really only on the number of objects in the pool we're selecting from; and the number of objects to be selected to form our collection.

What we need to do, then, is to find a formula for the number of such collections. We'll do that in the following section, and then we'll see how to combine that formula with the others we've derived to solve a large range of counting problems.

4.2 Binomial Coefficients

The good news: the formula for the number of collections is very simple to write down and to remember. The bad: it's not quite as straightforward to derive as the ones we've done up to now; in fact, figuring it out requires a somewhat indirect argument. What we'll do is show how to find the answer in a particular case, and once we've done that it'll be pretty clear how to replace the particular numbers in that example with arbitrary ones.

Let's take the case of choosing a committee of four students from among a class of 15—that is, the problem of counting the number of possible committees that can be formed. Again, this is a situation where the order of selection doesn't matter: egos aside, choosing Dave and then Rebecca has the same outcome as choosing Rebecca and then Dave; all that matters in the end is who is on the committee and who is not. And since possible committees don't correspond to sequences of choices, the multiplication principle doesn't seem to apply.

But it does apply!—in a sort of weird, backhanded way. To see how, let's focus for a moment on a different problem, the class officer problem: that is, counting the number of ways we can choose a president, vice president, secretary and treasurer for the class, assuming no student can occupy more than one office. As we saw, the multiplication principle works just fine here: we choose a president (15 choices), then the vice president (14), then the secretary (13) and finally the treasurer (12), for total of

$$15 \cdot 14 \cdot 13 \cdot 12 \qquad \text{or} \qquad \frac{15!}{11!}$$

possible slates.

But now suppose we want to solve the same problem in a different, somewhat warped way (though again using the multiplication principle). Suppose that instead of choosing the slate one officer at a time, we break the process up into two steps: first we choose a committee of four students who will be the class officers, and *then* choose which of those four will be president, vice president, secretary and treasurer.

That may seem like an unnecessarily complicated way to proceed. After all, we already know the answer to the class officer problem, while we don't know the number of committees. Bear with us! Let's look anyway at what it tells us.

The one thing we do know is, having selected the four members of the committee, how many ways there are of assigning to the four of them the jobs of president, vice president, secretary and treasurer: by what we already know, this is just $4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$. So if we do break up the process of selecting class officers into two stages, choosing a committee and then assigning them the four jobs, what the multiplication principle tells us is that

$$\left\{\begin{array}{c}\text{the number of ways of}\\ \text{choosing a committee}\end{array}\right\} \cdot 4! = \left\{\begin{array}{c}\text{the number of ways of}\\ \text{selecting class officers}\end{array}\right\}.$$

Now that, if you think about it a moment, tells us something. Since we know that the number of ways of selecting class officers is $15!/11!$, we can solve this equation for the number of committees:

$$\begin{aligned}\left\{\begin{array}{c}\text{the number of ways of}\\ \text{choosing a committee}\end{array}\right\} &= \frac{1}{4!} \cdot \left\{\begin{array}{c}\text{the number of ways of}\\ \text{selecting class officers}\end{array}\right\}\\ &= \frac{15!}{4!11!}.\end{aligned}$$

In English: since every choice of committee corresponds to $4! = 24$ different possible choices of class officers, the number of possible committees is simply $(1/24)^{\text{th}}$ the number of slates.

You can probably see from this that it's going to be the same when we count the number of ways of choosing a collection of any number k of objects from a pool of any number n. We know that the number of ways of choosing a *sequence* of k objects without repetition—a first, then a second different from the first, then a third different from the first two and so on—is just

$$n \cdot (n-1) \cdot \cdots \cdot (n-k+1) = \frac{n!}{(n-k)!}.$$

At the same time, for each possible collection of k objects from the pool, there are

$$k \cdot (k-1) \cdot \cdots \cdot 2 \cdot 1 = k!$$

ways of putting them in order—choosing a first, a second, and so on. The conclusion, then, is that

$$\begin{aligned}&\left\{\begin{array}{l}\text{the number of ways of choosing}\\ \text{a collection of } k \text{ objects, without}\\ \text{repetition, from a pool of } n \text{ objects}\end{array}\right\}\\ &\qquad = \frac{1}{k!} \cdot \left\{\begin{array}{l}\text{the number of ways of choosing a}\\ \text{sequence of } k \text{ objects, without}\\ \text{repetition, from a pool of } n \text{ objects}\end{array}\right\}\\ &\qquad = \frac{n!}{k!(n-k)!}.\end{aligned}$$

or, in other words,

> The number of ways of choosing a collection of k objects, without repetition, from among n objects is
>
> $$\frac{n!}{k!(n-k)!}.$$

So, for example, if the House of Pizza offers seven meat toppings, the number of possible pizzas you can order with three meat toppings is

$$\frac{7!}{3!4!} = \frac{5{,}040}{6 \cdot 24} = 35;$$

and if you're sent to the video store with instructions to return with an assortment of exactly four of their 23 movies based on comic strips or video games, your choice is among

$$\frac{23!}{4!19!} = 8{,}855$$

such assortments.

The numbers that appear in this setting are so ubiquitous in math that they have a name and a notation of their own. They're called *binomial coefficients* (for reasons we'll explain in Chapter 6), and written in this way:

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

There are a number of things we can say right off the bat about binomial coefficients. To begin with, there is the basic observation that

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

This is obvious from the above formula: we see that

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

just by rearranging the factors $k!$ and $(n-k)!$ in the denominator. It's also clear from the interpretation of these numbers: after all, specifying which four kids in the class of 15 are to be put on the committee is tantamount to specifying which 11 to leave off it; and in general choosing which k objects to take from a pool of n is the same as choosing which $n-k$ not to take.

Second, as we've pointed out, the standard formula for the binomial coefficients

$$\begin{aligned}\binom{n}{k} &= \frac{n!}{k!(n-k)!}\\ &= \frac{n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1}{k \cdot (k-1) \cdots 2 \cdot 1 \cdot (n-k) \cdot (n-k-1) \cdots 2 \cdot 1}\end{aligned}$$

is in some ways not the most efficient way to represent the number—it's certainly not how you would calculate it in practice—since there are factors that appear in both the numerator and the denominator, and can be cancelled. Doing this gives us two alternative ways of writing the binomial coefficient:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k \cdot (k-1) \cdots 2 \cdot 1}$$

$$= \frac{n \cdot (n-1) \cdots (k+1)}{(n-k) \cdot (n-k-1) \cdots 2 \cdot 1}.$$

This is not just an aesthetic issue, it's a practical one as well. Suppose, for example, you wanted to count the number of possible five-card hands from a standard deck of 52—that is, you wanted to evaluate the binomial coefficient $\binom{52}{5}$—and you wanted to carry out the calculation on your calculator. If you write the binomial coefficient as

$$\binom{52}{5} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

your calculator'll have no trouble multiplying and dividing out the factors. But if you write

$$\binom{52}{5} = \frac{52!}{5!47!}$$

you're in trouble: when you punch in 52! in your calculator you'll probably get an error message; most calculators can't handle numbers that large. Or, even worse, you won't get an error message; your calculator will simply switch to scientific notation. In effect, the calculator will round off the number and not tell you; and these roundoff errors can and often do become significant.

Before we go on, let's look at some special cases of binomial coefficients. To begin with, note that for any n,

$$\binom{n}{1} = n,$$

corresponding to the statement that "there are n ways of choosing one object from among n." (Well, that's not exactly news.) Also, by our convention that $0! = 1$, we see that

$$\binom{n}{0} = \frac{n!}{0!n!} = 1$$

and similarly $\binom{n}{n} = 1$. Again, think of this simply as a convention; it makes the various formulas we're going to discover in Chapter 6 work.

The first interesting case is the number of ways of choosing a pair of objects from among n:

$$\binom{n}{2} = \frac{n(n-1)}{2},$$

so that the number of ways of choosing two objects from among three is $3 \cdot 2/2 = 3$ (remember, this is the same as the number of ways of picking one object); the number of ways of choosing two objects from among four is $4 \cdot 3/2 = 6$; and in general we can make a table

n	# of ways of choosing two objects from among n
3	3
4	6
5	10
6	15
7	21
8	28

and so on. We can make a similar table for the binomial coefficients $\binom{n}{3}$:

n	# of ways of choosing three objects from among n
4	4
5	10
6	20
7	35
8	56
9	84

Mathematicians have found many fascinating patterns in these numbers, as well as other interpretations of them. We'll take a look at a few of these in Chapter 6.

There's one final remark we want to make about the binomial coefficients, even though it may be of interest only to math nerds. From the formula

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

it's clear that $\binom{n}{k}$ is a fraction, but it's far from clear that it's actually a whole number. Of course, we know it's a whole number from the interpretation as the number of ways of choosing k objects from n, but that just raises the question: can we see *why* the formula above always yields a whole number? In some cases we can do this. For example, when we look at the formula

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

and ask, "why is this a whole number?" we have an answer: no matter what n is, either n or $n-1$ must be even. Thus the product $n(n-1)$—the numerator of our fraction—must be even, and so the quotient is a whole number. Likewise, consider the formula

$$\binom{n}{3} = \frac{n(n-1)(n-2)}{6}.$$

Of the three factors n, $n-1$ and $n-2$ that appear in the numerator, at least one must be divisible by 3, and at least one must be even. The numerator must thus be divisible by 6 (again, think about it; if it's not clear now it will be in Chapter 8), and so the quotient is whole.

But it gets less and less obvious as k increases. For example, when we say that

$$\binom{n}{4} = \frac{n(n-1)(n-2)(n-3)}{24}.$$

is a whole number, we're saying in effect that *the product of any four whole numbers in a row is divisible by* 24. Think about this for a moment: could you convince yourself of this fact without using the interpretation of $\binom{n}{4}$? Could you convince someone else?

In fact we'll see in Part III of this book why this is true; for now we'll just leave it as something for you to mull over (or not).

Now it's time for you to do some exercises.

Exercise 4.2.1 Suppose now that the menu at the House of Pizza lists eight meat toppings. How many different pizzas can you order there with two (different) meat toppings? How many with three?

Exercise 4.2.2 Suppose you are given an exam with 10 problems, and are asked to do exactly seven of them. In how many ways can you choose which seven to do?

4.3 Examples

We can combine the formula we have now for the number of collections with other formulas and techniques. Here are some examples, in the form of solved problems.

EXAMPLE 4.3.1 Suppose once more we're asked to choose a collection of four students from our class of 15 to form a committee, but this time we have a restriction: we don't want the committee to consist of all boys or all girls. (Recall that there are eight boys and seven girls in the class.) How many different committees can we form?

SOLUTION This clearly calls for the subtraction principle. We know that the total number of possible committees is

$$\binom{15}{4} = \frac{15 \cdot 14 \cdot 13 \cdot 12}{4 \cdot 3 \cdot 2 \cdot 1} = 1{,}365.$$

The number of committees consisting of all boys is similarly

$$\binom{8}{4} = \frac{8 \cdot 7 \cdot 6 \cdot 5}{4 \cdot 3 \cdot 2 \cdot 1} = 70,$$

and the number of committees consisting of all girls is

$$\binom{7}{4} = \frac{7 \cdot 6 \cdot 5 \cdot 4}{4 \cdot 3 \cdot 2 \cdot 1} = 35.$$

If we exclude those, we see that the number of allowable committees is

$$\binom{15}{4} - \binom{8}{4} - \binom{7}{4} = 1{,}365 - 70 - 35 = 1{,}260.$$

■

Note that even though this is a multiple subtraction, we don't need to add any terms back in, since there are no committees that belong to both excluded classes—that is, a committee can't simultaneously consist of all boys and all girls.

EXAMPLE 4.3.2 One more committee: now suppose we require that the committee includes exactly two boys and two girls. How many possibilities are there now?

SOLUTION This, by contrast, is a case for the multiplication principle: to choose an allowable committee subject to this restriction, we simply have to choose two among the eight boys, and (independently) two among the seven girls. The answer is thus

$$\binom{8}{2} \cdot \binom{7}{2} = 28 \cdot 21 = 588$$

■

Exercise 4.3.3 Back to the House of Pizza: given that they feature seven meat toppings and four vegetable toppings, how many pizzas can be ordered with two meat and two vegetable toppings?

Exercise 4.3.4 Sam's ice cream shop offers only vanilla ice cream, but has 17 different possible toppings to choose from.

1. How many different sundaes can be formed with *exactly* three toppings?
2. How many different sundaes are there with *at least* two toppings?
3. How many different sundaes can be formed with no restriction on the number of toppings?

EXAMPLE 4.3.5 There are 10 players on a basketball team, and the coach is going to divide them up into two teams of five—the Red team and the Blue team, say—for a practice scrimmage. She's going to do it randomly, meaning that all of the $\binom{10}{5}$ ways of assigning the players to the two teams are equally likely.

Two of the players, Sarah and Rebecca, are friends. "I hope we wind up on the same team," Rebecca says. "Well, we have a 50-50 chance," Sarah replies.

Is Sarah right?

SOLUTION What this problem is asking us to do is to count, among the $\binom{10}{5}$ ways of assigning players to the two teams, how many result in Sarah and Rebecca being on the same team, and how many result in their being on opposing teams. Sarah is, in effect, saying that these two numbers will be equal; we'll calculate both and see if she's right.

Let's start by counting the number of ways of choosing the teams that result in Sarah and Rebecca winding up on the same team. We can specify such an assignment in two stages: first, we decide which team, Red or Blue, gets the Rebecca/Sarah duo. Obviously, there are two possibilities. Having done that, we have to take the remaining eight players and divide them into two groups: three will go to the team that already has Sarah and Rebecca; five will go the other team. The number of ways of doing that is $\binom{8}{3}$, so the total number of team assignments with the two as teammates is

$$2 \cdot \binom{8}{3} = 2 \cdot 56 = 112.$$

Now let's count the number of assignments that result in Sarah and Rebecca opposing each other. Again, we can specify such a choice in two steps: first, we specify which team, Red or Blue, Rebecca is on; Sarah will necessarily go the other. We then have to take the remaining eight players and assign four of them to each of the two teams; there are $\binom{8}{4}$ ways of doing this, and so the number of team assignments with Rebecca and Sarah opposite is

$$2 \cdot \binom{8}{4} = 2 \cdot 70 = 140.$$

The conclusion, then, is that Sarah is wrong: it is more likely that she and Rebecca will wind up on opposing teams.

We're not done! One thing we should always look for in doing these problems is a way to check the accuracy both of our analysis and of our calculations. Here we have a perfect way to do that. We've said that there are a total of $\binom{10}{5}$ ways of assigning the 10 players to the two teams, of which 112 result in the two friends being teammates and 140 result in their being on opposite teams. Before we're satisfied that we've got the correct answer, we should check that in fact $\binom{10}{5}$ is equal to the sum $112 + 140 = 252$. Let's do it:

$$\binom{10}{5} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{30{,}240}{120} = 252.$$

Having done this, we can be much more confident both that our analysis was correct, and that we didn't make any numerical mistakes. We can also say that the probability that Rebecca and Sarah wind up on the same team is

$$\frac{112}{252} = \frac{4}{9} \sim .444$$

or about 44%. We'll talk a lot more about probability in Chapter 5. ■

Exercise 4.3.6 Suppose now that the 10 players are to be divided into teams of six and four. Is it more likely that Sarah and Rebecca will be teammates or opponents?

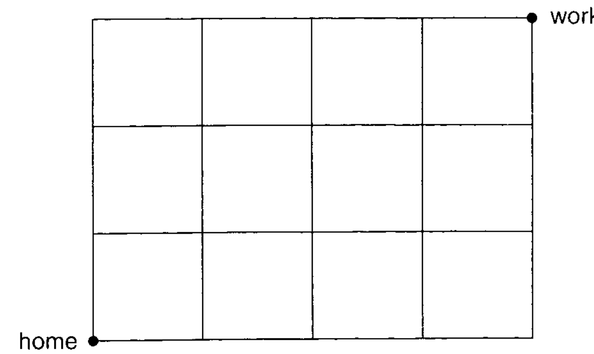
Note: this problem is just slightly trickier than the one we did above, since in one of the two calculations involved we can't use the multiplication principle. Be sure to check your answer!

EXAMPLE 4.3.7 Let's suppose you're playing Scrabble, and you have in your rack the letters "E, E, E, E, N, N, N" (this sort of thing seems to happen to us a lot). How many ways are there of arranging these tiles in your rack? In other words, how many seven-letter words (once more in the sense of arbitrary strings of letters) are there that contain exactly three Ns and four Es?

SOLUTION This is actually pretty simpleminded, but it'll lead to something more interesting in the next problem. The point is, if we think of our seven-letter word as having seven places to fill with Es and Ns, to specify such a word means just to specify which three of the places are to be filled with Ns; or, equivalently, which four are to be filled with Es. The answer is thus

$$\binom{7}{3} = \binom{7}{4} = 35.$$ ■

EXAMPLE 4.3.8 This problem may not seem like it has much to do with collections, but we'll see in a moment it does. Suppose we live in a city laid out in a rectangular grid, and that our job is located three blocks north and four blocks east of our apartment, as shown in the picture below.



Clearly, we have to walk seven blocks to get to work each morning. But there are many different paths we could take. How many different paths can we take, if we stay on the grid? Think about it before you look at the answer.

SOLUTION To specify a path from our home to work, we have to give a series of directions like, "Go one block North, then three blocks East, then another block North, then another block East, then another block North," or, for short

N, E, E, E, N, E, N.

In other words, paths correspond exactly to words consisting of exactly 3 Ns and 4 Es. So this is exactly the same problem as the last one and the answer is 35 paths. ■

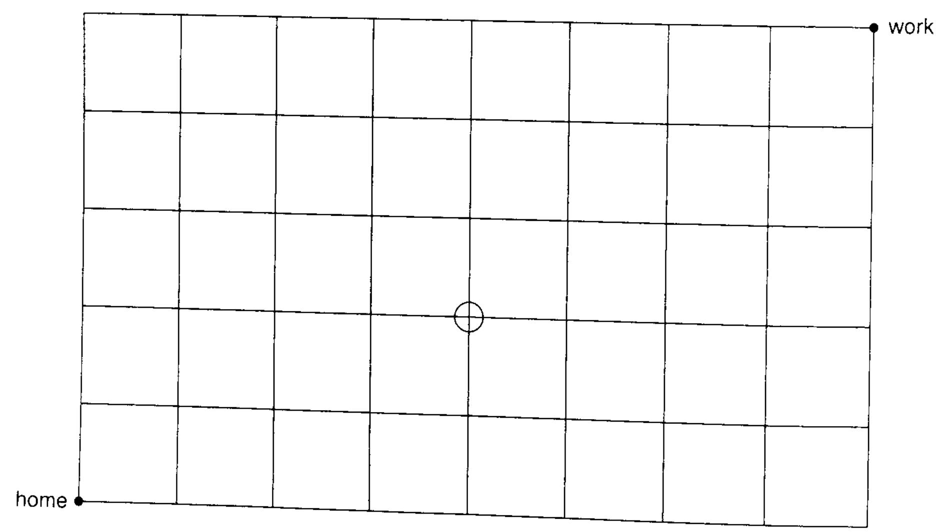
In general, if we have a k by l rectangular grid, by the same logic the paths from one corner to the opposite one (with no doubling back) correspond to words formed with k Ns (or Ss) and l Es (or Ws). The number of such paths is thus given by the binomial coefficient

$$\binom{k+l}{k},$$

or, equivalently, the binomial coefficient $\binom{k+l}{l}$. Note the symmetry here: the number of paths in a $k \times l$ grid is the same as the number of paths in an $l \times k$ grid, as the formula verifies.

Exercise 4.3.9

1. Consider the grid below. How many paths of shortest possible length (that is, 13 blocks) are there from the point labeled "home" to the point labeled "work"?



2. Suppose that the small circle on the diagram represents Mike's House of Donuts and Coffee, a crucial stop on the way if you're to arrive at work awake. How many paths from home to work (again of minimal length) pass through Mike's?
3. Suppose by contrast that you're on a strict diet, and that you must at all costs *avoid* passing through the intersection where Mike's is located. Now how many possible paths do you have?

One last item before we conclude this section: we promised in Chapter 2 to show you how to do Example 2.5.2, and now it's time. We'll start by doing a simpler version of the problem that nonetheless introduces the essential ideas.

EXAMPLE 4.3.10 Let's say a class consists of seven boys and seven girls, and we want to line them up so that no two boys are next to each other, and no two girls are next to each other. How many ways are there of doing this?

SOLUTION To start with, the key feature is to separate the problem into two phases:

- First, we have to choose which places in line are to be occupied by boys and which by girls. That is, we have to choose a sequence of genders or, if you like, a 14-letter word consisting of seven Bs and seven Gs; *but with no two Bs adjacent and no two Gs adjacent*. Once we've done that,
- We have to assign an actual student of the appropriate gender to each place.

In Example 2.5.1, the first step didn't exist: since there were eight boys and seven girls, the only possible arrangement of genders was to alternate BGBGBGBGBGBGBGB. Once we realized that, we simply had to assign the eight boys to the 8 Bs in that sequence, and the seven girls to the seven Gs, for a total of $8! \cdot 7!$ choices.

In the present circumstances, by contrast, there are different possible gender sequences. But not many: since we have to alternate boys and girls, the sequence of genders is determined once we specify the first; that is, it's got to be either GBGBGBGBGBGBGB or BGBGBGBGBGBGBG. So there are just 2 choices for the gender sequence.

The second step is essentially the same in both Example 2.5.1 and the present problem. In the present circumstances, once we've decided on a particular arrangement of genders there'll be exactly 7! ways of assigning the seven boys in class to the seven places in line designated for boys, and 7! ways of assigning the girls' places; so there'll be $7! \cdot 7!$ ways of assigning the 14 students to appropriate places in line. The answer is thus the number of gender arrangements times $7! \cdot 7!$; that is,

$$2 \cdot 7! \cdot 7! = 50{,}803{,}200. \qquad \blacksquare$$

Now we're ready to tackle Example 2.5.2. First, recall the problem: we have a class of 6 boys and 9 girls, and we want to know how many ways to line them up assuming we don't want any two boys next to each other in line. If you haven't thought about the problem, take some time now to do so, especially in light of the example we've just worked out.

Ready? Here goes. The first step, just as in the last problem, is to separate the problem into two phases: specifying which places in line are to be occupied by boys and which by girls; and then assigning an actual student of the appropriate gender to each place. Moreover, the second step is essentially the same in both examples: once we've decided on a particular arrangement of genders there'll be exactly 6! ways of assigning the six boys in class to the six places in line designated for boys, and 9! ways of assigning the girls' places; so there'll be $6! \cdot 9!$ ways of assigning the 15 students to appropriate places in line. The answer is thus $6! \cdot 9!$ times the number of gender arrangements—that is, $6! \cdot 9!$ times the number of 15-letter words consisting of 6 Bs and 9 Gs, with no two Bs in a row.

OK, then, how do we figure out that number? Here is where it gets slightly tricky. The first step is to consider two possibilities: either the sequence ends in a B or it ends in a G. Suppose first that the sequence ends in a G. In that case, we observe, every B in the sequence is necessarily followed by a G. In other words, instead of arranging 6 Bs and 9 Gs—subject to the requirement that no B follow another—we can pair off one G with each B to form six BGs, with three Gs left over, and *count arbitrary arrangements of six BGs with three Gs*. We know how many of those there are; it's just

$$\binom{9}{6} = 84.$$

Next, we consider arrangements ending in a B, and we count those similarly: we take the remaining five Bs and pair each with a G to form five BGs, with four Gs left over. Again, we can take arbitrary arrangements of these five BGs and four Gs, with a B stuck on at the end; and there are

$$\binom{9}{5} = 126$$

of these. Altogether, then, there are $84 + 126 = 210$ possible gender sequences; and since each gives rise to $6! \cdot 9!$ possible lineups, the total number is

$$210 \cdot 6! \cdot 9! = 54,867,456,000.$$

4.4 Multinomials

Say it's now our job to assign college students to dorm rooms. We have a group of nine students to assign, and three rooms to assign them to: one a quad, one a triple and one a double. The standard question: how many different ways can we do it?

"That's nothing new," you might say, "we already know how to do it." And you'd be right: to assign the nine students we could start by choosing four of the nine and assign then to the quad. That leaves five students to be assigned to the triple and double; choose three of the remaining five to go to the triple and you're done, since the remaining two necessarily go in the double. Since we had $\binom{9}{4} = 126$ ways of making the first choice and $\binom{5}{3} = 10$ ways of making the second, the answer would be

$$\binom{9}{4} \cdot \binom{5}{3}.$$

But wait: what would happen if we did the assignments in a different order? Suppose, for example, that we started by assigning two of the students to the double, and then chose four of the remaining seven to go the quad? We'd have $\binom{9}{2}$ choices at the first stage, and $\binom{7}{4}$ at the second, so the correct answer would be

$$\binom{9}{2} \cdot \binom{7}{4}.$$

What's up with that?

The answer is, there's no mistake. When we look closer, we see that

$$\binom{9}{4} \cdot \binom{5}{3} = \frac{9!}{4!5!} \cdot \frac{5!}{3!2!} = \frac{9!}{4!3!2!}$$

while on the other hand

$$\binom{9}{2} \cdot \binom{7}{4} = \frac{9!}{2!7!} \cdot \frac{7!}{4!3!} = \frac{9!}{4!3!2!}.$$

Hence both approaches lead to the same answer, which is equal to 1,260.

So there's really nothing new here. But the sort of numbers that arise here—the number of ways of distributing some number n of objects into three (or more) collections of specified size—are so common that they, like the binomial coefficients, deserve a name and a notation of their own. Think of it this way: maybe the correct (or at any rate the symmetric) way to think of the binomial coefficient $\binom{n}{k}$ is as the number of ways of distributing a group of n objects into two collections, of size k and $n-k$. Well, in the same vein, whenever we have a number n and three numbers a, b and c that add up to n, we can ask how many ways there are of distributing n objects into three collections, of sizes a, b and c. We can answer this completely analogously to the way we just did the last problem: we first choose which a of our n objects are to go into the first group; then which b of the remaining $n-a$ are to go into the second. That'll leave c objects, which have to go into the third; so by the multiplication principle the number of ways is

$$\binom{n}{a} \cdot \binom{n-a}{b} = \frac{n!}{a!(n-a)!} \cdot \frac{(n-a)!}{b!(n-a-b)!} = \frac{n!}{a!(n-a)!} \cdot \frac{(n-a)!}{b!c!} = \frac{n!}{a!b!c!}.$$

This number is called a *multinomial coefficient*, and is typically denoted by the symbol

$$\binom{n}{a,\ b,\ c} = \frac{n!}{a!b!c!}.$$

Similarly, if a, b, c and d are four numbers adding up to n, the number of ways of distributing n objects into groups of size a, b, c and d is

$$\binom{n}{a,\ b,\ c,\ d} = \frac{n!}{a!b!c!d!},$$

and so on. The most general form of this problem would be: suppose we have n different objects, which we want to distribute into k collections. The number of objects in each collection is specified: the first collection is to have a_1 of the objects, the second a_2 and so on; the k^{th} and last collection is to have a_k of the objects. We ask: how many ways are there of assigning the n objects to the k collections? The answer, as we've suggested, is that

> The number of ways of distributing n objects into groups of size $a_1, a_2, \ldots, a_k$ is
> $$\frac{n!}{a_1! \cdot a_2! \cdot \cdots \cdot a_k!}.$$

Again, the number $\frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}$ that appears here is called a multinomial coefficient and denoted $\binom{n}{a_1, a_2, \ldots, a_k}$. Note that in this setting our old friend the binomial coefficient $\binom{n}{k}$ could also be written as $\binom{n}{k,\ n-k}$; but it's easier (and unambiguous) to just drop the $n-k$.

Multinomial coefficients are thus a straightforward generalization of binomial coefficients, and are almost as ubiquitous (though, as we just saw, you don't *really* need to know them: if you just know about binomial coefficients and the multiplication principle, you can solve any problem involving multinomial coefficients).

Exercise 4.4.1 The job has fallen to you to assign 18 incoming freshmen to rooms in one particular dormitory. There are six rooms: two quads, two triples, and two doubles.

1. In how many ways can the 18 freshmen be assigned to the rooms?
2. After you submitted your list of assignments from part (1) to the Dean, she complained that some of them put men and women in the same room. If we designate one of the quads, one of the triples, and one of the doubles for women, in how many ways can the rooms be assigned to nine women and nine men?

One classic example of multinomial coefficients is in counting anagrams. By an *anagram* of a word, we mean a rearrangement of its letters: for example, "SAPS" is an anagram of "PASS." (Note that each letter must appear the same number of times in the anagram as in the original word.) In keeping with our conventions, by an anagram we'll mean an arbitrary rearrangement of the letters, not necessarily a word in the English language.

So: how many anagrams does a word have? In some cases this is easy: if a four-letter word, say, has all different letters (that is, none repeated) then an anagram of the word is simply an ordering of its letters, and so there are 4! of them. For example, the word "STOP" has 24 anagrams

STOP	STPO	SOTP	SOPT	SPTO	SPOT
TSOP	TSPO	TPSO	TPOS	TOSP	TOPS
OSTP	OSPT	OTSP	OTPS	OPST	OPTS
PSTO	PSOT	PTSO	PTOS	POST	POTS

By the same token, an n-letter word with n different letters will have $n!$ anagrams.[1] At the other extreme, the answer is also relatively easy: a word consisting of only one letter repeated n times has no anagrams other than itself; and, as we saw in the example of the Scrabble tiles, if a word consists of k repetitions of one letter and l repetitions of another, it has $\binom{k+l}{k}$ anagrams.

In general, the right way to think about anagrams (from a mathematical point of view) is the way we described in Example 4.3.7. Suppose, for example, we want to count the anagrams of the word "CHEESES." Any such anagram is again a seven-letter word. If we think of it as having seven slots to fill with the letters C, H, E and S then to specify an anagram we have to specify

1. Which one of those seven slots is to be assigned the letter C;
2. Which one of those seven slots is to be assigned the letter H;
3. Which three of those seven slots are to be assigned the letter E; and of course
4. Which two of those seven slots are to be assigned the letter S.

When we think of it in this way, the answer is clear: it's just the multinomial coefficient

$$\binom{7}{1,\ 1,\ 3,\ 2} = \frac{7!}{1!1!3!2!} = 420.$$

[1] For the verbally oriented among you, here's a question. Note that, of the 24 rearrangements of the letters STOP, six of them (STOP, SPOT, OPTS, POST, POTS and TOPS) are actual words in English. Are there four-letter words with more? What five-letter word has the most English word anagrams?

Exercise 4.4.2 How many anagrams does the word "MISSISSIPPI" have? In how many of them are the two Ps next to each other?

Exercise 4.4.3 Consider the following six-letter words: "TOTTER," "TURRET," "RETORT," "PEPPER" and "TSETSE." Which one has the most anagrams, and which the fewest? (You should try and figure out the answer before you actually calculate out the numbers in each case.)

4.5 Something's Missing

At the Bright Horizons School, prizes are given out to the students to reward excellence. All the prizes are identical, though the school may choose to give more than one prize to a given student.

In Ms. Wickersham's class at B.H.S., there are 14 students: Alicia, Barton, Carolina, and so on up to Mark and Nancy. Ms. W. has eight prizes to award, and has to decide how to give them out—that is, how many prizes each child should get. In how many ways can she do this?

For a slightly different formulation, suppose for the moment that you're the chief distributor for the National Widget Importing Co. The NWI has 14 warehouses, called (the NWI is not a very fanciful outfit) Warehouse A, Warehouse B and so on up to Warehouse N.

One day, eight containers of widgets show up at the docks, and it's your job to say how many of the eight should go to each of the 14 warehouses. How many ways are there of doing this?

Or: you're in the dining hall one day, and there's a massive fruit bowl, featuring unlimited quantities of each of 14 different fruits: apples, bananas, cherries and so on up to nectarines. Feeling a mite peckish, you decide to help yourself to eight servings of fruit, possibly taking more than one serving of a given kind. How many different assortments can you select?

Well, what is the point here? Actually, there are a couple: one, we don't know how to solve this problem; and two, we should. Think about it: we've derived, so far in this book, three formulas for counting the number of ways of making a series of k selections from a pool of n objects:

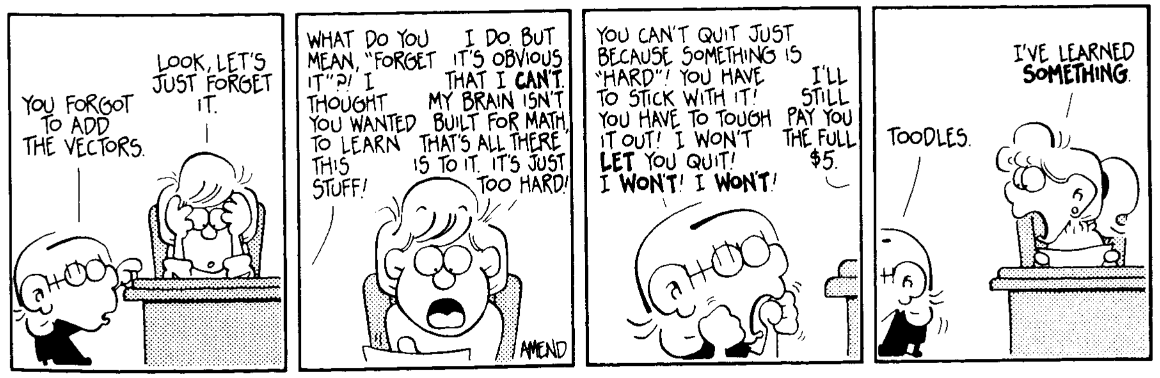
- We know the number of ways of choosing a sequence (that is, the order does matter) from the pool, with repetitions allowed: it's n^k.
- We know the number of ways of choosing a sequence from the pool, with no repetitions: it's $n!/(n-k)!$.
- We know the number of ways of choosing a collection (that is, the order doesn't matter) from the pool, with no repetitions: it's $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

If we arrange these formulas in a table, as here:

	repetitions allowed	without repetitions
sequences	n^k	$\dfrac{n!}{(n-k)!}$
collections	??	$\dfrac{n!}{k!(n-k)!}$

it's clear that something's missing: we don't have a formula for the number of collections of k objects, chosen from a pool of n, with repetitions allowed. That's what all those problems we just listed (or that single problem we repeated three times) involve.

So: are we going to tell you the answer, already? Well, yes and no. We are going to work out the formula in Chapter 7, at the end of this part of the book. But we thought it'd be nice to leave you something to think about and work on in the meantime. So we'll leave it as a challenge: can you solve the problem(s) above before we get to Chapter 7?



FOXTROT © Bill Amend. Reprinted with permission of UNIVERSAL PRESS SYNDICATE. All rights reserved.