Counting I

Margaret M. Fleck

6 April 2009

This lecture starts the topic of counting, covering sections 5.1 and 5.3.

1 Introduction

Many applications require counting, or estimating, the size of a finite set. For small examples, you can just list all the elements. For examples with a simple structure, you could probably improvise the right answer. But when it's not so obvious, some techniques are often helpful.

2 Product rule

> **The product rule**: if you have p choices for pinning down one feature of the object and then q choices for a second feature, and your options for the second feature don't depend on what you chose for the first one, then you have pq options total.

So if T-shirts can come in 4 colors and 3 sizes, there are $4 \cdot 3 = 12$ types of T-shirts.

In general, if the objects are determined by a sequence of independent decisions, the number of different objects is the product of the number of options for each decision. So if the T-shirts come in 4 colors, 5 sizes, and 2 types of necklines, there are $4 \cdot 5 \cdot 2 = 40$ types of shirts.

Yup. That's pretty much what you thought. Nothing hard here.

3 The sum rule, inclusion/exclusion

> **The sum rule**: suppose your task can be done in one of two ways, which are mutually exclusive. If the first way has p choices and the second way has q choices, then you have $p+q$ choices for how to do the task.

Example: it's late evening and you want to watch TV. You have 37 programs on cable, 57 DVD's on the shelf, and 12 movies stored in I-tunes. So you have $37 + 57 + 12 = 106$ options for what to watch.

Weeeellll, maybe. This analysis assumes that there is no overlap between the movies available on the three media. If any movies are in more than one collection, this will double-count them. So we would need to subtract off the number that have been double-counted. For example, if the only overlap is that 2 movies are on I-tunes and also on DVD, you would have only $(37 + 57 + 12) - 2 = 104$ options.

The formal name for this correction is the "Inclusion-Exclusion Principle".

Suppose you have two sets A and B. Then

$$|A \cup B| = |A| + |B| - |A \cap B|$$

We can use this basic formula to derive the formula for three sets A, B, and C:

$$\begin{aligned}
|A \cup B \cup C| &= |A| + |B \cup C| - |A \cap (B \cup C)| \\
&= |A| + |B| + |C| - |B \cap C| - |A \cap (B \cup C)| \\
&= |A| + |B| + |C| - |B \cap C| - |(A \cap B) \cup (A \cap C)| \\
&= |A| + |B| + |C| - |B \cap C| - (|A \cap B| + |A \cap C| - |(A \cap B) \cap (A \cap C)|) \\
&= |A| + |B| + |C| - |B \cap C| - |A \cap B| - |A \cap C| + |A \cap B \cap C|
\end{aligned}$$

4 Combining these two rules

Let's see what happens on a more complex example. Suppose S contains all 5-digit decimal numbers that start with 2 one's or end in 2 zeros, where we don't allow leading zeros. How many numbers does S contain?

Let T be the set of 5-digit numbers starting in 2 one's. We know the first two digits and we have three independent choices (10 options each) for the last three. So there are 1000 numbers in T.

Let R be the set of 5-digit numbers ending in 2 zeros. We have 9 options for the first digit, since it can't be zero. We have 10 options each for the second and third digits, and the last two are fixed. So we have 900 numbers in R.

What's the size of $T \cap R$? Numbers in this set start with 2 one's and end with 2 zeros, so the only choice is the middle digit. So it contains 10 numbers. So

$$|S| = |T| + |R| - |T \cap R| = 1000 + 900 - 10 = 1890$$

5 Non-independent decisions

Now, suppose that describing our set of possibilities involves a sequence of decisions that aren't independent, i.e. where later decisions depend on earlier ones. Dependencies can create a lot of very hard-to-analyze situations. However, there is a range of standard situations for which nice formulas exist.

For example, suppose that we have 7 Scrabble tiles, all different from one another, and we want to form a 4-letter word. We have 7 choices for the first letter in the word. But then we have only 6 choices for the second letter, because we've used up one time. So the number of 4-letter words we can form is $7 \cdot 6 \cdot 5 \cdot 4 = 840$. In math jargon, each possible word is called a *4-permutation* of the set of 7 tiles.

To get a sense of the general formula, notice what happens if we want to form a word using all 7 letters. We then have $7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ ways to do it. This is just 7!.

In general, suppose we have a set S of n objects. Then a *permutation* of S is a way of putting the elements of S into an order. There are $n!$ ways to do this, because we have n choices for which element to put first, then $n-1$ choices for which to put second, and so on.

Now, suppose that we only want to pick k objects from S, but the order still matters to us. This is called a k-permutation of S. Then there are $n(n-1)\ldots(n-1+1) = \frac{n!}{(n-k)!}$ different k-permutations . This number is called $P(n,k)$.

6 Combinations

The permutations formula applies when we care about the order in which we are selecting the objects, e.g. we are putting them into an arrangement or choosing them for a series of different roles. If we simply want to select a subset of the objects, we need a different formula. An unordered set of k elements is called a k-combination.

Example: How many ways can I select a 7-card hand from a 60-card deck of magic cards (assuming no two cards are identical)?

One way to analyze this problem is to figure out how many ways we can select an ordered list of 7 cards, which is $P(60,7)$. This over-counts the number of possibilities, so we have to divide by the number of different orders in which the same 7-cards might appear. That's just 7!. So our total number of hands is $\frac{P(60,7)}{7!}$ This is $\frac{60\cdot 59\cdot 58\cdot 57\cdot 56\cdot 55\cdot 54}{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2}$. Probably not worth simplifying or multiplying this out unless you really have to. (Get a computer to do it.)

In general, suppose that we have a set S with n elements and we want to choose an unordered subset of k elements. We have $\frac{n!}{(n-k)!}$ ways to choose k elements in some particular order. Since there are $k!$ ways to put each subset into an order, we need to divide by $k!$ so that we will only count each subset once. So the general formula for the number of possible subsets is $\frac{n!}{k!(n-k)!}$.

The expression $\frac{n!}{k!(n-k)!}$ is often written $C(n,k)$ or $\binom{n}{r}$. This is pronounced "n choose r." It is also sometimes called a "binomial coefficient," for reasons that will become obvious in a couple lectures. So the shorthand answer to our question about magic cards would be $\binom{60}{7}$.

Notice that $\binom{n}{r}$ is only defined when $n \geq r \geq 0$. What is $\binom{0}{0}$? This is $\frac{0!}{0!0!} = \frac{1}{1\cdot 1} = 1$.

7 Applying the combinations formula

> Problem: Baskin-Robbins advertises having 31 flavors of ice cream. Suppose you want to order a sugar cone with 4 scoops, all differently flavored. How many choices do you have for the four flavors?

The first question you have to ask yourself is: do I care about the order? Is it "the same" choice if they put the blueberry on top of the chocolate or under the chocolate? If the problem wasn't clear on this point, you may have to ask for clarification.

If we think the order matters, we apply the permutations formula. Our answer is then $P(31, 4)$, which is $31 \cdot 30 \cdot 29 \cdot 28$. If we think the order doesn't matter, we apply the combinations formula. Our answer would then be $\binom{31}{4}$ which is $\frac{31!}{4!(31-4)!}$.

Also notice that these formulas only work if all items in the set are distinguishable and we don't get to pick duplicates of the same item. When those assumptions fail, we have to restructure the problem or use a more complex formula. We'll probably some examples involving duplicate objects next Monday, depending on how fast lectures go.

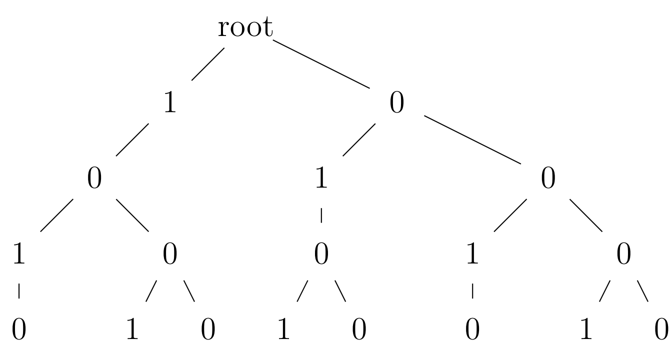
8 Tackling less obvious problems

The formulas for counting are not especially difficult to use. The biggest problem is knowing how to describe your problem in terms of a known formula. This sometimes takes some fiddling around.

For example, suppose we have a set of 7 adults and 3 kids. Let's call the kids A, B, and C. They need to stand in line to board an airplane and no two kids can stand next to each other because they will fight with one another and cause trouble. We have two choices for who is first. But then the later choices depend in a complex way on the earlier ones. This isn't going to work real well.

The trick for this problem is to place the 7 adults in line, with gaps between them. Each gap might be left empty or filled with one kid. There are 8 gaps, into which we have to put the 3 kids. So, we have 7! ways to assign adults to positions. Then we have 8 gaps in which we can put kid A, 7 for kid B, and 6 for kid C. That is $7! \cdot 8 \cdot 7 \cdot 6$ ways to line them all up.

When the problem is small and the dependencies among the decisions are complex, you can resort to drawing a tree diagram enumerating all the possibilities. For example, suppose we want to find all bit strings (strings of 0's and 1's) of length four which do not have two consecutive ones. We can draw the following tree, in which each branch represents one possible string. We can see that there are 8 strings with these properites.



9 The bit string viewpoint

Applying the combinations formula can also require reworking how you think about the problem. In particular, it often helps to see placing objects in certain positions of an arrangement as choosing a subset of the positions.

Example: How many 16-digit bit strings contain exactly 5 zeros?

Solution: The string contains 16 positions. We need to pick 5 of these to be the ones with the zeros. So we have $\binom{16}{5}$ ways to do this.

More complex example: How many 10-digit strings from the 26-letter ASCII alphabet contain exactly 3 A's?

Solution: We need to pick a subset of the 10 positions in which to put the three A's. There are $\binom{10}{3}$ ways to do this. After we've done that, we have seven positions to fill with our choice of any character except A. We have 25^7 ways to do that. So our total number of strings is $\binom{10}{3}25^7$

Solutions like this are somewhat custom, and not easy to generalize to apparently similar problems. You have to think about each situation carefully when your problem involves multiple identical objects.

Modified example: How many 10-digit strings from the 26-letter ASCII alphabet contain no more than 3 A's?

Solution: We do the above analysis to count the number of strings with exactly 3 A's, exactly 2 A's, exactly 1 A, and no A's. Then add up these four results. So the total number of strings is $\binom{10}{3}25^7 + \binom{10}{2}25^8 + \binom{10}{1}25^9 + 25^{10}$