Counting

Sam Elder

Sunday, November 5, 2006

1 Introduction

Counting is a task that you will have to do again and again in your math careers. And that's not a joke. It appears at ALL levels of competition, from CSU Math Day to the USAMO as well as college. While one of the most fundamental ideas, counting can get quite difficult. For example, consider this problem: Given n different markers, each separated from their caps, in how many ways can you put the caps on the markers so that none of them match? For example, if $n = 3$, there are two ways, one for each cap we can put on the first marker. Small cases seem easy, but how do we solve it in the general case. We'll look at this problem (called the number of *derangements*) a couple of ways today.

While most of these lectures will be devoted (almost) entirely to olympiad math, this topic is so diverse, I have relaxed that condition today. The topics covered here will span the entire spectrum.

The rest of the handout is divided into several sections. First, we'll establish the basics of the area of math known as combinatorics. After that, we'll look at several tricks that help you count things, such as overcounting. Then we'll look at the more advanced concepts of the Principle of Inclusion-Exclusion, generating functions, and the ubiquitous Catalan numbers. I've included a dozen problems that we will try to do along the way. They'll be homework if we don't get to them.

2 Combinatorics Basics

You should all be familiar with the following:

- Factorials: $n! = n(n-1)(n-2)\cdots(2)(1)$.
- Permutations: ${}_n\mathrm{P}_r = n!/(n-r)! = n(n-1)\cdots(n-r+1)$.
- Combinations: $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$. This last expression even works for n any complex number, not just nonnegative integers $\geq k$. We'll use this fact in one of the problems.
- Binomial Theorem: $(x+1)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$.
- Combinatorial identities: $\binom{n}{k} = \binom{n}{n-k}$, $\sum_{i=0}^{n} \binom{n}{i} = 2^n$.
- Number of ways to arrange n letters with k_i repeats of letter i: $\frac{n!}{k_1!k_2!\cdots k_j!}$.

If you need a brief refresher on anything here, please tell me. Otherwise, we'll move on.

3 Counting Tricks

There is no prescriptive way to count things. When you were young, you probably figured out that instead of counting the jelly beans one by one, it was much faster to count them in groups of 3, 4 or 5. That's a simple trick. In olympiad math, you often find some way to count something entirely different from the original problem.

3.1 Overcounting

When we *overcount*, we simply count more than we should and then subtract or divide (as appropriate) to get what we want. For example, suppose we need to give out medals for the top 5 finishers in a 10-person race. In how many ways can you assure that at least one of three friends will get a medal? The easiest way to do this is to see how many ways they can not get a medal (${}_7\mathrm{P}_5$) and subtract that from the total number of ways (${}_{10}\mathrm{P}_5$), getting $10\cdot9\cdot8\cdot7\cdot6-7\cdot6\cdot5\cdot4\cdot3 = 30240-2520 = 27720$. That's the subtract method. Division is also easy: Just use the number of ways to arrange n letters from the previous section as an example.

3.2 Bijections

A *bijection*, also called a 1-to-1 correspondence, is a way of "translating" one problem into another. Whenever you want to get a bijection, you have to prove both sides of it. Given an object in A, give an algorithm to get a unique object in B, then do the same from B to A.

For example, consider this problem: How many solutions are there to $x_1 + x_2 + \cdots + x_k = n$ for nonnegative integers $x_1, x_2, \ldots, x_k$? We translate this to a list of $n + k - 1$ boxes, n of which are filled and $k - 1$ of which are empty. Given the list of boxes, count out how many boxes are filled at the front end of the list and let x_1 be this value. Then remove the first empty box and repeat $k - 1$ times. There are a total of n filled boxes, so $x_1 + x_2 + \cdots + x_k = n$, as desired. Moreover, no two lists of boxes can give the same set of nonnegative integers, since if they first differ at a point encountered in the i-th step, they will have different values for x_i. The other direction of the bijection is simple: Given a set of k nonnegative integers summing to n, fill the first x_1 boxes and leave the $(x_1 + 1)$-th box empty. Ignoring the first $x_1 + 1$ boxes, repeat this process for each of the other $k - 1$ numbers, of course not making another empty box for the last step. The result is that $k - 1$ boxes will be empty and n will be filled, yielding $n + k - 1$ total boxes, as desired. And if two sequences first differ in x_i, that will result in different sequences of boxes starting with the i-th empty box. So there is a bijection from the number of lists of n filled and $k - 1$ empty boxes and the number of sets of k nonnegative integers summing to n.

Now, we count the number of lists of boxes. All we have to do is pick n to be filled, so in general, there are $\displaystyle\binom{n+k-1}{n}$ ways to fill the boxes, and thus ways to choose the nonnegative integers.

3.3 Counting in two ways

Many problems ask you to prove that two quantities, usually in terms of n (and possibly k or i) are equal. One method, similar to the bijection method is to invent a counting problem, the answer to which can be calculated in two different ways, yielding one side of the equation each time. It's best to see these in action, so a large portion of the problems at the end are of this type.

Here's a simple example: $\binom{n}{k} = \binom{n}{n-k}$. For the LHS, we count the number of ways to *pick* k items from a set of n, and for the RHS, we count the number of ways to *not pick* $n - k$ items from a set of n. These are the same.

4 Principle of Inclusion-Exclusion

Most introductions to PIE begin with an analogy of so many students taking these different classes and we want to know how many aren't taking any or are taking all of them or whatever. Instead of starting that way, I will start in a more general context: Let $A_1, A_2, \ldots, A_n$ be subsets of the universal set X. We want to calculate the number of elements in X that are not in any of the A_i. We start with $|X|$, the cardinality (number of elements in) set X. We then exclude the number of elements in each of the A_i, or $\sum |A_i|$. But if an element is in two, we excluded it one too many times, so we add back in the $|A_i \cap A_j|$ for $1 \le i < j \le n$. Now if an element is in three sets, A_i, A_j, A_k, then we have included it once, excluded it three times and included it three times, which means we need to exclude it one more time. So we subtract $|A_i \cap A_j \cap A_k|$ for all $1 \le i < j < k \le n$. Continuing this process, we simply alternate excluding and including cardinalities of more and more intersecting sets, and we cap it all off with the cardinality of the intersection of all n sets multiplied by $(-1)^n$ to account for the alternating inclusion/exclusion. With that, we get our total number of elements in X but not in the A_i.

Formally, we can define PIE as follows[1]: Again let $A_1, A_2, \ldots, A_n$ be subsets of X. Let $I \subseteq \{1, 2, \cdots, n\}$ be a subset of the index set. Then let $A_I = \bigcap_{i \in I} A_i$, basically defining the intersection of $|I|$ sets from I. PIE says that the number of elements in X but not in the A_i is

$$\sum_{I \subseteq \{1,2,\cdots,n\}} (-1)^{|I|} |A_I|.$$

We prove this simply by looking at the contribution of each element to the sum. If $x \in X$ but $x \notin A_i$ for any i, then it contributes exactly once (when I is empty). If x is in k of the A_i, it contributes $\sum_{i=0}^{k} (-1)^i \binom{k}{i} = 0$ times (by Binomial Theorem with $x = -1$). We are done.

PIE shows up in all sorts of contexts. For a powerful example, let's take a stab at the derangements problem again. Suppose there are $f(k)$ ways to put the letters in the envelopes so that *at least* k of them match. PIE says that the total number of derangements is $f(0) - f(1) + f(2) - \cdots$. So we try to find a formula for $f(k)$. First, we choose in $\binom{n}{k}$ ways those that we are going to match up. Then we put the remaining $n-k$ anywhere ($(n-k)!$ ways), for a total of $\binom{n}{k}(n-k)! = \frac{n!}{k!}$ ways total. In all, we get the number of derangements to be

$$\sum_{k=0}^{n} (-1)^k \frac{n!}{k!} = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} = n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\right).$$

Recalling from calculus that $e^x = 1 + \dfrac{x}{1!} + \dfrac{x^2}{2!} + \cdots$, with $x = -1$ we get $\dfrac{1}{e} = 1 - \dfrac{1}{1!} + \dfrac{1}{2!} - \dfrac{1}{3!} + \cdots$. From there we just need to prove that our partial sum is approximately equal to the infinite series, or at least within $1/n!$ of it, which is not hard to do. Basically, the difference between them is an alternating series, so we can throw everything we know from calculus about alternating series in its direction. In conclusion, we get the number of derangements $d(n)$ to be $\lfloor n!/e \rfloor$.

As you can see, many times PIE will give us an answer in sum notation, that we may or may not be able to simplify. In this case, we got lucky and were able to reduce it to a simple expression. This is almost always true: PIE will give us an answer, but it isn't always pretty.

---

[1] This explanation is from Peter Cameron's *Combinatorics: Topics, Techniques, Algorithms.*

5 Generating Functions

Generating functions are one of weirdest topics in mathematics. The entire field is based on the crazy idea of taking a sequence of numbers and sticking it into an infinite power series polynomial, otherwise known as a generating function. Formally, that means that given the sequence $\{a_n\}$, its *generating function* $A(x)$ is defined to be

$$A(x) = \sum_{n\geq 0} a_n x^n.$$

To avoid the pages and pages of formal baggage, we just say that these generating functions behave exactly like polynomials regarding addition, subtraction, multiplication and division. We'll look at multiplication real quick. Formally,

$$f(x)g(x) = \sum_{n\geq 0}(a_n b_0 + a_{n-1}b_1 + \cdots + a_0 b_n)x^n.$$

A useful trick is to turn $1 + x + x^2 + \cdots$ into $\frac{1}{1-x}$. Note that we never said what x was. To put it simply, *we don't care about convergence.* In fact, there isn't much we actually do care about. A lot of it is selective; we only ever plug in x if we want to. In this world of generating functions, $1 + 2 + 4 + 8 + \cdots$ does equal -1. It doesn't really matter that it doesn't really make sense. Are you confused yet?

So what can we do with generating functions? We can prove things about them that make no sense elsewhere. In a way, they're kind of like mods. If a generating function has a nonzero constant term, then it has a multiplicative inverse (another generating function). How do we prove that? Suppose the constant term of $A(x)$ is a_0. Then we let the constant term of $B(x)$ be $b_0 = 1/a_0$. Defining b_n for $i > 0$ such that $a_n b_0 + a_{n-1}b_1 + \cdots + a_0 b_n = 0$, we see from our definition of multiplication, $A(x)B(x) = 1$, as desired.

Name a sequence, and it likely has a simple generating function. In fact, for the Fibonacci numbers F_n, $F(x) = \dfrac{1}{1 - x - x^2}$. Another interesting example is the Catalan numbers, whose generating function actually makes things easy relating to them. We'll look at those in the problems if we get to them.

6 Problems

Here is where you get to apply these concepts. It's great to know about generating functions in general, but knowing where and how to use them is a whole different issue.

1. Prove that
$$\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2.$$

2. (Hockey Stick Theorem) Prove that
$$\binom{n}{k} + \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k}{k} = \binom{n+1}{k+1}.$$

3. (AoPS) What fraction of n-letter "words" have at least one repeated letter?

4. Show that
$$\sum_{k=0}^{n} \binom{n}{k}\binom{n-k}{m-k} = 2^m \binom{n}{m}.$$

5. There are n pieces of candy in a pile. One is allowed to separate a pile into two piles, and add the product of the sizes of the resulting piles to a running total. The process terminates when each piece of candy is in its own pile. Show that the final sum is independent of the sequence of operations performed.

6. Show that for nonnegative a, b, k,
$$\binom{a+b}{k} = \sum_{i=0}^{k} \binom{a}{i}\binom{b}{k-i}.$$

7. We'll define the Catalan numbers C_n as the number of expressions of n open parentheses and n closed parentheses such that at any point the number of open parentheses to the left is at least that of the number of closed parentheses. We also define $C_0 = 1$ as a convention. For example, $C_3 = 5$ since ()()(), ()(()), (())(), ((())), and (()()) are all of the ways to do this. Prove that $C_n = C_{n-1}C_0 + C_{n-2}C_1 + \cdots + C_0C_{n-1}$.

8. Show that the generating function for the Catalan numbers is
$$C(x) = \frac{1-\sqrt{1-4x}}{2x}.$$

9. Using the previous problem, prove that $C_n = \dfrac{1}{n+1}\dbinom{2n}{n}$.

10. Prove that the generating function for the Fibonacci numbers is $\dfrac{1}{1-x-x^2}$.

11. In how many ways can you get a total of 21 on 6 different dice?

12. What is the generating function for the number of partitions of $n = n_1 + n_2 + \ldots + n_k$ with $n_i \geq n_j$ whenever $i < j$?