Lecture 5: Counting

Nitin Saxena *

IIT Kanpur

Counting problems arise in almost every aspect of computer science. In this lecture we will learn some basic techniques and principles for counting.

1 Basic counting

There are two very simple rules used extensively to count.

1. *Sum rule:* We need to choose one element from two sets, first containing a elements and the other b. If the two sets are *disjoint* then it can be done in $a + b$ ways.
2. *Product rule:* We need to choose two elements, one *each* from two sets, first containing a elements and the other b. Then, the total number of ways is ab.

We have taken only two sets, but the rule can be generalized to multiple sets easily. Let us do some examples.

Example 1. How many bit strings are there of length 7?

There are 7 positions. In each of these positions, there are two possibilities. Applying product rule, there are $2^7 = 128$ possibilities.

Exercise 1. How many numbers are there between 1000 and 9999 (including both) which are divisible by 3?

Count in the range $\{1, \ldots, 1000\}$ and $\{1, \ldots, 9999\}$.

Example 2. How many palindrome words are there of length n?
If n is even then $26^{\frac{n}{2}}$. If n is odd then $26^{\frac{n+1}{2}}$.

Example 3. How many ways are there to put m balls into n bins.

1. Balls are distinct and bins are distinct:
   Every ball has n choices. Hence n^m.

   Exercise 2. Why is the answer not m^n by looking at the opposite argument?

2. Balls are identical but bins are distinct:
   This corresponds to ordered partitions of the number m (in $\leq n$ parts). So, consider a string of m 0's and $n-1$ 1's, and permute them. Every permutation corresponds to a way of placing m balls in n bins. The former count is
   $$\frac{(m+n-1)!}{m! \cdot (n-1)!} = \binom{m+n-1}{m}.$$
3. Balls are distinct but bins are not:
   This corresponds to (unordered) partitions of the set $\{1, \ldots, m\}$ (in $\leq n$ parts). This is a difficult counting problem. We will cover it under *Bell's number* later.
4. Balls (respectively bins) are identical:
   This corresponds to the (unordered) partitions of the number m (in $\leq n$ parts). Again, this is a difficult counting problem. We will cover it under *generating functions* later.

* Edited from Rajat Mittal's notes.

1.1 Counting in two ways

Let us consider the following problem. Show that in a conference, the number of members who shake hands an odd number of times is even.

We will count the number P of pairs (m_i, m_j), where member m_i shook hands with another member m_j. We know that P is twice the number of total handshakes. Hence, P is even.

Counting P another way, it is the sum of handshakes done by each member. If the number of members who shook hands an odd number of time is odd then P will be odd (why?). But, since P is even, it implies that the number of members who shook hands an odd number of times is even.

This trick is called *counting in two ways*. The idea is to count one particular quantity in two different ways. Since we know that any counting should give identical results, we can derive certain properties.

2 Inclusion exclusion

Suppose we have 100 students in an institute using various social-media websites. Say 44 use Facebook, 50 use Twitter and 56 use G+.

It is also known that how many of them use multiple websites. 27 use Facebook and Twitter, 31 are on Facebook and G+, also 34 use G+ and Twitter. There are 19 who use all three websites.

How many students are there who do not use any website?

This kind of question can be visualized easily using a *Venn-diagram*. This Venn-diagram can be easily drawn using the data above. (Hint: Start from the inner-most part of the Venn-diagram, and gradually deduce the numbers in the outer-parts.)

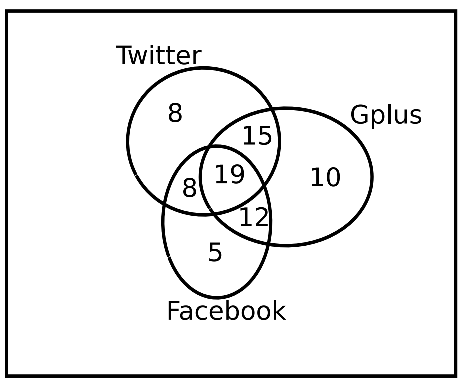

Fig. 1. Website Users

Looking at the diagram, it is quite clear that 23 $(= 100 - (56 + 8 + 8 + 5))$ students do not use any social media website. But, if there were twenty websites (or a lot of websites) then this calculation strategy is quite tedious!

We try to count the number of students using at least one website in a different way. Basically, we move from the outer-part towards the inner-parts (of an undrawn Venn-diagram). Our first guess would be to sum up students using any website $(44 + 50 + 56 = 150)$. But clearly this counts a student using two websites twice, so let us subtract the students who are in the intersection of two websites.

So the next guess would be $150 - (27 + 31 + 34) = 58$. Again, students using all three websites were counted 3 times in the beginning, then subtracted thrice, so were not counted at all. Hence the true count is $58 + 19 = 77$.

So there are $100 - 77 = 23$ students who do not use any of the three websites.

This argument can be generalized for more than three websites. Suppose there is a universe U with subsets $A_1, A_2, \cdots, A_n$. For our previous example, U will be the set of all students $(|U| = 100)$ and A_1 will be the set of students using Facebook and so on.

We are given their intersections $A_I := \cap_{i \in I} A_i$, for every subset $I \subseteq [n]$. Then the number of elements not in any of the sets A_i is given by,

$$\left| U - \bigcup_{i \in [n]} A_i \right| = \sum_{I \subseteq [n]} (-1)^{|I|} \cdot |A_I| \ .$$

Note 1. $A_\emptyset$ is the universe U itself. Why?

This is called the *principle of inclusion and exclusion*.

Exercise 3. Why is it called inclusion-exclusion?

Proof of the principle. We will show the equality by finding the count of an element of U in RHS. If an element $u \in U$ is not contained in any of the sets A_I then it will be counted exactly once (namely, by the term $|A_\emptyset|$).

So we only need to show that every other element is counted 0 times (overall). Suppose an element $u \in U$ is contained in A_j's for every $j \in J$, where J is the maximal nonempty subset of $[n]$. Then, the number of times u gets counted in RHS is,

$$c_J := \sum_{I \subseteq J} (-1)^{|I|} \cdot 1 \ .$$

Suppose $|J| = k$. We will be done if we can show that:

Exercise 4. $c_J = \sum_{i=0}^{k} (-1)^i \binom{k}{i} = 0$.

Consider the binomial expansion of $(1-1)^k$.

This proves that every element of U is counted exactly once if it is not in any of the A_i's, and not counted otherwise. This proves the inclusion-exclusion principle. □

References

1. P. J. Cameron. Combinatorics: Topics, Techniques and Algorithms. *Cambridge University Press*, 1994.
2. K. H. Rosen. Discrete Mathematics and Its Applications. *McGraw-Hill*, 1999.