CS30 (Discrete Math in CS), Summer 2021 : Lecture 11

Topic: Combinatorics: The Four Fold Formula

Disclaimer: These notes have not gone through scrutiny and in all probability contain errors.
Please discuss in Piazza/email errors to deeparnab@dartmouth.edu

---

Suppose we are given (many copies of) n *distinct* items. We need to pick k items from these. How many ways can we do this? Well, this is an underspecified problem. To specify it, we need to fix two bits – (a) does the order in which we pick them matter?, and (b) can multiple copies of the same item be picked?

1. **Ordering matters, Repetition allowed.**

   Example: *How many* 4 *letter words are there (not necessarily in the English dictionary)?*

   The n distinct items here are the 26 alphabets $\{\mathtt{a}, \mathtt{b}, \mathtt{c}, \ldots, \mathtt{z}\}$. The parameter k is 4. Note that the order in which the letters are written matters: the word `stop` and `tops` are different. Furthermore, repetition is allowed: the word `good` is a valid four letter word (and so is `zxvx` as well; these don't need to be in the English dictionary).

   We know the answer to this by the *product principle*: the answer is $\mathbf{n^k}$.

2. **Ordering matters, Repetition not allowed.**

   Example: *How many ways can you rank the top-*10 *webpages of a Google search?*

   The n distinct items are the distinct webpages that Google returns as possible hits. But it can only display 10 on a page. How many choices are there? It can't show the same page more than once, and surely, the order matters (at least to the website owners).

   We know the answer to this also by the product principle. The first item has n choices, the second has $(n-1)$ choices, and so and so forth, till the kth item has $(n-k+1)$ choices. So the answer is

   $$n(n-1)(n-2)\cdots(n-k+2)(n-k+1) = \frac{\mathbf{n!}}{\mathbf{(n-k)!}}$$

   Often this is denoted as $P(n,k)$ or ${}^nP_k$.

   Since order matters, both the above examples are about *sequences* from a collection of distinct items. Next we move to sets.

3. **Ordering Doesn't matter, Repetition not allowed.**

   Example: *How many ways can you pick* 11 *players from a collection of* 30 *players?*

   You want to pick a squad for the Indian team. The talent is super, and you have 30 excellent players. You have to pick 11. It doesn't matter in which order you pick them, a team is a team. However, you can't pick the same player twice (doesn't make sense!). Here $n = 30$, and $k = 11$.

   Let us derive this answer in two ways (actually, they are the same way). First, we actually count the number of ways in which we can pick k items where repetition is not allowed, but ordering *matters*. We know the answer to this: the answer is $\frac{n!}{(n-k)!}$. Let A be the set consisting of all the sequences (since ordering matters). Now, we map this to the set S of objects we want: a collection of unordered sets of size k. Note that given any k elements (players), there are $k!$ orderings each of which are distinct elements in A which map to the same element in S. Thus, by the *division rule*, the answer we are interested in is $|A|/k! = \frac{n!}{(n-k)!k!}$.

   Another way to think about this problem is as follows.

Given a set of n distinct items, what is the number of subsets which have cardinality exactly k.

Recalling the bijection between subsets and n-length bit-strings from last time, we see that the number of subsets of cardinality k is exactly the number of n-length bit-strings with exactly k ones. This, as we calculated last time is $\frac{n!}{k!(n-k)!}$.

This number, that is, the number of ways of choosing a subset of k things from a set of n things is a very, very useful object. It is denoted as $\binom{n}{k}$ and is pronounced as "n-choose-k". They have some amazing properties, and we will see some of them.

$$\text{The answer in this case is} \quad \binom{\mathbf{n}}{\mathbf{k}} = \frac{\mathbf{n}!}{\mathbf{k}!(\mathbf{n}-\mathbf{k})!}$$

4. **Ordering doesn't matter, Repetition Allowed.**

   Example: *You are at a take-out restaurant where they have 6 entrees. You have chosen the 4 entree option (you are hungry after the CS 30 midterm). How many ways can you fill your dinner plate – note that you can take the Sesame Chicken more than once, but it will count towards your 4 entrees.*

   Let us cast this problem slightly differently. We claim that the number of choices is precisely equal to the cardinality of the following set

$$T := \{(x_1, x_2, \ldots, x_n) : \sum_{i=1}^{n} x_i = k, \quad x_i \geq 0, \quad x_i \in \mathbb{Z}\} \tag{1}$$

   In English, T is the collection of a sequence/vector of n non-negative integers which sum to exactly k. We claim that the number of ways to choose k items from n distinct choices where repetition is allowed but order doesn't matter is $|T|$.

   The proof goes by showing a bijective map. Given any unordered collection of k items from n distinct choices, we can get a vector in T where x_i denotes the number of copies of item i. Note that the number of copies is ≥ 0 and an integer, and they sum to k. Conversely, given any vector $x \in T$ we can get a collection of k items by picking x_i copies of the ith choice.

   Remark: *Before we move on to answer what (1) is, let's already note that this "translation" means that this question also captures questions of the form:* ***how many non-negative integer solutions are there to the equation*** $a + b + c = 100$***?*** *This is the same problem as above with $n = 3$ and $k = 100$; that is, picking 100 items from 3 different choices where repetition is allowed and order doesn't matter.*

   Back to figuring out what (1) is. It doesn't immediately leap to the eye how to solve this problem. The trick is to consider the following picture: given a vector $\vec{x}$ in T, consider x_1 circles followed by a vertical line, then x_2 circles followed by a vertical line, and so and so forth till you have x_{n-1} circles followed by a vertical line, and then x_n circles. We see that (or at least guess; proof coming up) the circle-vertical lines patterns are in one-to-one correspondence with $\vec{x} \in T$. Finally, think of each circle as a 0 and a vertical line as 1; the circle-line pattern is precisely a $(n + k - 1)$ dimensional bit-vector with exactly k zeros and $n - 1$ ones. (And by now we very well know how to calculate that!)

Theorem 1. There is a bijection from T to the set $Z := \{z \in \{0,1\}^{n+k-1} : z \text{ has } k \text{ zeros}\}$.

Proof. Given a $\vec{x} \in T$, map it to the vector in Z which has x_1 zeros followed a one, and then x_2 zeros followed by a one, and so on and so forth till you have x_{n-1} zeros followed by a one, and then x_n zeros. Note this is a vector in Z; the total number of ones is $n-1$ and the total number of zeros is $\sum_{i=1}^{n} x_i = k$.

This map is bijective: given any vector $z \in Z$, you can "reverse the map" to get the same vector in T; traverse the vector z from left-to-right; count the number of zeros till you see the first one, and set x_1 to this, then count the number of zeros till you see the second one, and set x_2 to that, and so and so forth, till you count the number of zeros till you see the $(n-1)$th one and set x_{n-1} to that, and then set x_n to be the number of trailing zeros.

To give an example, if $\vec{x} = (2,3,0,0,1)$, then the corresponding z is (0010001110). Here $k = 6$ and $n = 5$. ☐

In sum, the number of ways to choose k items from n distinct choices where repetitions are allowed and order doesn't matter is equal to the number of bit-strings of length $n+k-1$ which have exactly k zeros. And the answer to this, as we now know is

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

	Order Matters	Order Doesn't Matter
Repetition	$\mathbf{n^k}$	$\binom{\mathbf{n+k-1}}{\mathbf{k}}$
No Repetition	$\frac{\mathbf{n!}}{\mathbf{(n-k)!}}$	$\binom{\mathbf{n}}{\mathbf{k}} = \frac{\mathbf{n!}}{\mathbf{k!(n-k)!}}$

Table 1: Picking k things from n distinct items