Lisa Yan
CS 109

Lecture Notes #2
September 25, 2019

Combinatorics

Based on a handout by Mehran Sahami and Chris Piech

As we mentioned last class, the principles of counting are core to probability. Counting is like the foundation of a house (where the house is all the great things we will do later in CS109, such as machine learning). Houses are awesome. Foundations, on the other hand, are pretty much just concrete in a hole. But don't make a house without a foundation. It won't turn out well.

1 Permutations

Permutation Rule: A permutation is an ordered arrangement of n distinct objects. Those n objects can be permuted in $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1 = n!$ ways.

This changes slightly if you are permuting a subset of distinct objects, or if some of your objects are indistinct. We will handle those cases shortly!

1.1 Example 1

Part A: iPhones used to have 4-digit passcodes. Suppose there are 4 smudges over 4 digits on the screen. How many distinct passcodes are possible?

Solution: Since the order of digits in the code is important, we should use permutations. And since there are exactly four smudges we know that each number is distinct. Thus, we can plug in the permutation formula: $4! = 24$.

Part B: What if there are 3 smudges over 3 digits on the screen?

Solution: One of 3 digits is repeated, but we don't know which one. We can solve this by making three cases, one for each digit that could be repeated (each with the same number of permutations). Let A, B, C represent the 3 digits, with C repeated twice. We can initially pretend the two C's are distinct. Then each case will have $4!$ permutations:

$$A \ B \ C_1 \ C_2$$

However, then we need to eliminate the double-counting of the permutations of the identical digits (one A, one B, and two C's):

$$\frac{4!}{2! \cdot 1! \cdot 1!}$$

Adding up the three cases for the different repeated digits gives

$$3 \cdot \frac{4!}{2! \cdot 1! \cdot 1!} = 3 \cdot 12 = 36$$

Part C: What if there are 2 smudges over 2 digits on the screen?

Solution: There are two possibilities: 2 digits used twice each, or 1 digit used 3 times, and other digit used once.

$$\frac{4!}{2! \cdot 2!} + 2 \cdot \frac{4!}{3! \cdot 1!} = 6 + (2 \cdot 4) = 6 + 8 = 14$$

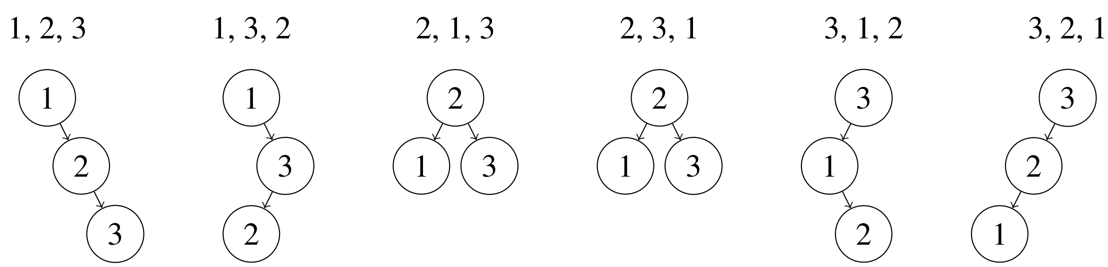
1.2 Example 2

Recall the definition of a **binary search tree (BST)**, which is a binary tree that satisfies the following three properties for *every* node n in the tree:

1. n's value is greater than all the values in its left subtree.
2. n's value is less than all the values in its right subtree.
3. both n's left and right subtrees are binary search trees.

Problem: How many possible binary search trees are there which contain the three values 1, 2, and 3, and have a degenerate structure (i.e., each node in the BST has at most one child)?

Solution: We start by considering the fact that the three values in the BST (1, 2, and 3) may have been inserted in any one of 3! (=6) orderings (permutations). For each of the 3! ways the values could have been ordered when being inserted into the BST, we can determine what the resulting structure would be and determine which of them are degenerate. Below we consider each possible ordering of the three values and the resulting BST structure.



We see that there are 4 degenerate BSTs here (the first two and last two).

2 Permutations of Indistinct Objects

Permutation of Indistinct Objects: Generally when there are n objects and

n_1 are the same (indistinguishable),

n_2 are the same,

...

and n_r are the same,

then there are

$$\frac{n!}{n_1!n_2!\ldots n_r!}$$

distinct permutations of the objects.

2.1 Example 3

Problem: How many distinct bit strings can be formed from three 0's and two 1's?

Solution: 5 total digits would give 5! permutations. But that is assuming the 0's and 1's are distinguishable (to make that explicit, let's give each one a subscript). Here is a subset of the permutations.

$$\begin{matrix} 0_1 & 1_1 & 1_2 & 0_2 & 0_3 \\ 0_1 & 1_1 & 1_2 & 0_3 & 0_2 \\ 0_2 & 1_1 & 1_2 & 0_1 & 0_3 \\ 0_2 & 1_1 & 1_2 & 0_3 & 0_1 \\ 0_3 & 1_1 & 1_2 & 0_1 & 0_2 \\ 0_3 & 1_1 & 1_2 & 0_2 & 0_1 \end{matrix}$$

If identical digits are indistinguishable, then all the listed permutations are the same. For any given permutation, there are 3! ways of rearranging the 0's and 2! ways of rearranging the 1's (resulting in indistinguishable strings). We have over-counted. Using the formula for permutations of indistinct objects, we can correct for the over-counting:

Total $= \frac{5!}{3! \cdot 2!} = \frac{160}{6 \cdot 2} = \frac{120}{12} = 10.$

3 Combinations

Combinations: A combination is an unordered selection of r objects from a set of n objects. If all objects are distinct, then the number of ways of making the selection is:

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

,
where $\binom{n}{r}$ is defined as a binomial coefficient and is often read as "n choose r".

Consider this general way to produce combinations: To select r unordered objects from a set of n objects, e.g., "7 choose 3",

1. First consider permutations of all n objects. There are $n!$ ways to do that.
2. Then select the first r in the permutation. There is one way to do that.
3. Note that the order of r selected objects is irrelevant. There are $r!$ ways to permute them. The selection remains unchanged.
4. Note that the order of $(n-r)$ unselected objects is irrelevant. There are $(n-r)!$ ways to permute them. The selection remains unchanged.

Total $= \frac{n!}{r! \cdot (n-r)!} = \binom{n}{r} = \binom{n}{n-r}$ e.g., $\frac{7!}{3!4!} = 35$

which is the combinations formula.

A useful recursive identity for combinations is as follows:

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}, 0 \leq r \leq n.$$

This identity can be proved via a combinatorial argument. When we select a group of size r from n distinct objects, then any particular object (say, object 1) will either be part of that group or not part of that group. We can then define sets A and B, where A is the number of ways of selecting a group that contains object 1, and B is the number of ways of selecting a group that does not contain object 1. For Set A, if we decide to include object 1, then we must select $r-1$ of the remaining $n-1$ objects (since the membership of object 1 in our selection is already decided), or $1 \times \binom{n-1}{r-1}$. For Set B, if we decide to exclude object 1, then we only have $n-1$ possible objects to select from to create a group of size r, or $n - 1r$. These sets are mutually exclusive, and therefore by the Sum Rule of Counting the total number of possibilities are as above.

3.1 Example 4

Problem: In the Hunger Games, how many ways are there of choosing 2 villagers from district 12, which has a population of 8,000?

Solution: This is a straightforward combinations problem. $\binom{8000}{2} = 31{,}996{,}000$.

3.2 Example 5

Part A: How many ways are there to select 3 books from a set of 6?

Solution: If each of the books are distinct, then this is another straightforward combination problem. There are $\binom{6}{3} = \frac{6!}{3!3!} = 20$ ways.

Part B: How many ways are there to select 3 books if there are two books that should not both be chosen together? For example, if you are choosing 3 out of 6 probability books, don't choose both the 8th and 9th edition of the Ross textbook).

Solution: This problem is easier to solve if we split it up into cases. Consider the following three different cases:

Case 1: Select the 8th Ed. and 2 other non-9th Ed.: There are $\binom{4}{2}$ ways of doing so.

Case 2: Select the 9th Ed. and 2 other non-8th Ed.: There are $\binom{4}{2}$ ways of doing so.

Case 3: Select 3 from the books that are neither the 8th nor the 8th edition: There are $\binom{4}{3}$ ways of doing so.

Using our old friend the Sum Rule of Counting, we can add the cases:

$$\text{Total} = 2 \cdot \binom{4}{2} + \binom{4}{3} = 16.$$

Alternatively, we could have calculated all the ways of selecting 3 books from 6, and then subtract the "forbidden" ones (i.e., the selections that break the constraint). Chris Piech calls this the Forbidden City method.

> Forbidden Case: Select 8th edition and 9th edition and 1 other book. There are $\binom{4}{1}$ ways of doing so (which equals 4).
>
> Total = All possibilities − forbidden = 20 – 4 = 16.

Two different ways to get the same right answer!

4 Selecting multiple groups of objects

Selecting multiple groups of objects: If n objects are distinct, then the number of ways of selecting r groups of objects, such that group i has size n_i, and $\sum_{i=1}^{r} n_i = n$, is:

$$\frac{n!}{n_1!n_2!\cdots n_r!} = \binom{n}{n_1, n_2, \ldots, n_r},$$

where $\binom{n}{n_1,n_2,\ldots,n_r}$ is defined as a multinomial coefficient.

This situation is a generalization of the combination, where $\binom{n}{r}$ is defined as a binomial coefficient. One way to see this is that the task of selecting r unordered objects from a set of n distinct objects is analogous to the task of separating n objects into two groups 1 and 2, with respective element counts $n_1 = r$ and $n_2 = n - r$. Therefore it is true that the binomial coefficient $\binom{n}{r} = \binom{n}{r, n-r}$, where the latter is the multinomial coefficient.

4.1 Example 6

Problem: Company Camazon has 13 new servers that they would like to assign to 3 datacenters, where Datacenter A, B, and C have 6, 4, and 3 empty server racks, respectively. How many different divisions of the servers are possible?

Solution: This is a straightforward application of our multinomial coefficient representation. Setting $n_1 = 6, n_2 = 4, n_3 = 3$, $\binom{13}{6,4,3} = 60,060$.

Another way to do this problem would be from first principles of combinations as a multipart experiment. We first select the 6 servers to be assigned to Datacenter A, in $\binom{13}{6}$ ways. Now out of the 7 servers remaining, we select the 4 servers to be assigned to Datacenter B, in $\binom{7}{4}$ ways. Finally, we select the 3 servers out of the remaining 3 servers, in $\binom{3}{3}$ ways. By the Product Rule of Counting, the total number of ways to assign all servers would be $\binom{13}{6}\binom{7}{4}\binom{3}{3} = \frac{13!}{6!4!3!} = 60,060$.

5 Bucketing / Group Assignment

You have probably heard about the dreaded "balls and urns" probability examples. What are those all about? They are for counting the many different ways that we can think of stuffing elements into containers. (It turns out that Jacob Bernoulli was into voting and ancient Rome. And in ancient Rome they used urns for ballot boxes.) This "bucketing" or "group assignment" process is a useful metaphor for many counting problems.

Note that this bucketing problem is different from the previous combinations problem. In combinations, we have *n distinct* (distinguishable) objects to put in r distinct groups, and we are fixing the number of distinct objects in group i to be n_i (where $\sum_{i=1}^{r} n_i = n$) for *every* outcome that we count. By contrast, in the bucketing problem we still have n objects to put in r distinct groups, but (1) our objects can be distinct or indistinct, and (2) for each outcome we can vary the number of objects in each distinct group i.

5.1 Example 7

Problem: Say you want to put n distinguishable balls into r urns. (No! Wait! Don't say that!) Okay, fine. No urns. Say we are going to put n strings into r buckets of a hash table where all outcomes are equally likely. How many possible ways are there of doing this?

Solution: You can think of this as n independent experiments each with r outcomes. Using our friend the General Principle of Counting, this comes out to r^n.

While the previous example allowed us to put n distinguishable objects into r distinct groups, the more interesting problem is to work with n indistinguishable objects. This task has a direct analogy to the number of ways to solve the following positive integer equation:

$$x_1 + x_2 + \cdots + x_r = n, \text{ where } 0 \leq x_i \leq n \text{ for all } i = 1, \ldots, r$$

.

Divider Method: Suppose you want to place n indistinguishable items into r containers. The divider method works by imagining that you are going to solve this problem by sorting two types of objects, your n original elements and $(r - 1)$ dividers. Thus, you are permuting $n + r - 1$ objects, n of which are same (your elements) and $r - 1$ of which are same (the dividers). Thus the total number of outcomes is:

$$\frac{(n + r - 1)!}{n!(r - 1)!} = \binom{n + r - 1}{n} = \binom{n + r - 1}{r - 1}.$$

5.2 Example 8

Part A: Say you are a startup incubator and you have \$10 million to invest in 4 companies (in \$1 million increments). How many ways can you allocate this money?

Solution: This is just like putting 10 balls into 4 urns. Using the Divider Method we get:

$$\text{Total ways} = \binom{10+4-1}{10} = \binom{13}{10} = 286.$$

This problem is analogous to solving the integer equation $x_1 + x_2 + x_3 + x_4 = 10$, where x_i represents the investment in company i such that $0 \leq x_i \leq 10$ for all $i = 1, 2, 3, 4$.

Part B: What if you don't have to invest all \$10 M? (The economy is tight, say, and you might want to save your money.)

Solution: Imagine that you have an extra company: yourself. Now you are investing \$10 million in 5 companies. Thus, the answer is the same as putting 10 balls into 5 urns.

$$\text{Total ways} = \binom{10+5-1}{10} = \binom{14}{10} = 1001.$$

This problem is analogous to solving the integer equation $x_1 + x_2 + x_3 + x_4 + x_5 = 10$, such that $0 \leq x_i \leq 10$ for all $i = 1, 2, 3, 4, 5$.

Part C: What if you know you want to invest at least \$3 million in Company 1?

Solution: There is one way to give \$3 million to Company 1. The number of ways of investing the remaining money is the same as putting 7 balls into 4 urns.

$$\text{Total ways} = \binom{7+4-1}{7} = \binom{10}{7} = 120.$$

This problem is analogous to solving the integer equation $x_1 + x_2 + x_3 + x_4 = 10$, where $3 \leq x_1 \leq 10$ and $0 \leq x_2, x_3, x_4 \leq 10$. To translate this problem into the integer solution equation that we can solve via the divider method, we need to adjust the bounds on x_1 such that the problem becomes $x_1 + x_2 + x_3 + x_4 = 7$, where x_i is defined as in Part A.