Combinatorics for Dummies

Dody Eid

Introduction

How many ways are there to arrange a list of things? This is the fundamental question of **combinatorics**. Here is a nice example. Imagine you have a group of 10 people, each with a unique letter (A, B, C, ..., J), and you want to choose 4 people to be the rulers of your group. How many ways could you do this? By the end of this blog, you will be able to solve this problem and its more complicated variants.

Theorem 0: The Multiplication Rule

It cannot be overstated how important it is to understand the "multiplication rule" of combinatorics, so let's do that here with a specific example. Imagine I have three options for breakfast: eggs, ham, or waffles. Then, at lunch, I have three more options: salad, chicken, or steak. Question is: how many possible breakfast-lunch combinations exist? To start exploring this question, it is a good idea to write out a couple possible combinations to see what we are dealing with. Here are three:

- eggs-salad
- waffles-steak
- eggs-chicken

We need to see how many possible pairs like the ones above we can make. Well, there are 3 options for breakfast, and then for each of those 3 options, there are another 3 options for lunch! Therefore, we have $3 * 3 = 9$ total breakfast-lunch combo possibilities.

We can use this logic more generally. Say I have a options for breakfast and then b options for lunch (a and b are just whole numbers here). Then, I have $a * b$ options for breakfast-lunch combos. We can get even more general by realizing breakfast is just some arbitrary example. This idea holds for any kind of event that has a certain number of options. If you have a options for one event and b options for another event, you have $a * b$ total possibilities when you do both events. This is known as the multiplication rule.

What if I added a dinner trial that also has 3 outcomes: salmon, beef, and pasta. Now, how many breakfast-lunch-dinner combos exist? We know that we have $3*3$ options for breakfast lunch combos, and for each breakfast-lunch combo there are 3 dinner options, so we can multiply again: breakfast-lunch-dinner combos $= (3 * 3) * 3 = 3 * 3 * 3 = 27$.

Let's now do the most general case. Say I have k events with n possible outcomes for each event. Then, there are n^k total possibilities for the outcome of all k events!

Question 1: Sampling with replacement, order matters

We now have the tools to solve this question: From a group of 10 people (A, B, C, $\ldots$, J), How many ways can I choose a board consisting of a President, Vice President, Treasurer, and Secretary such that people are allowed to hold more than one position? Well, let's follow the logic of the multiplication rule. This process can be construed as the combination of 4 processes: choosing a President, choosing a Vice President, choosing a Treasurer, and choosing a Secretary. When making a choice for the President, I have 10 options. Then, when making a choice for the Vice President, I also have 10 options (remember that people can hold multiple positions, so my choice for President is eligible to be Vice President). Then, when making a choice for the Treasurer, I have 10 options. Finally, I have 10 options for making a choice for the Secretary. So, using the multiplication rule, I have $10 * 10 * 10 * 10 = 10^4 = 10,000$ possibilities for the board!

We can now generalize. This situation is called **sampling with replacement where order matters**. We are **sampling** because we are picking a subgroup out of a bigger group (in our specific case, picking a board out of a group of 10 people). We are doing it **with replacement** because after we pick something from the bigger group to fill one slot, we toss it back into the bigger pile to be an eligible candidate for the next pick (i.e. if we pick someone as the President, they are still eligible to be Vice President, Treasurer, and Secretary). And **order matters**: if we pick A to be President and B to be Vice President, this is a different situation than if B was chosen to be President and A was chosen to be Vice President.

Let's write a general formula for sampling with replacement where order matters. Recall that there were 10 options for each position on the board, and there were 4 positions, so we did $10 * 10 * 10 * 10$ to get $10^4 = 10,000$ possibilities (by the multiplication rule). If, instead, we had n options and k positions to fill, we would have n^k possibilities. **So, when we want to pick k things from a group of n things with replacement such that order matters, we have n^k possibilities!**

Question 2: Sampling without replacement, order matters

We now move on to a slightly harder combinatorics problem. What if I wanted to pick a President, Vice President, Treasurer, and Secretary from a group of 10 people, but no person can hold more than one title? This situation is more generally known as **sampling without replacement**, because once I pick someone from the group of 10 to fill the Presidency, he or she is no longer eligible for another slot. Since every position on the board has a title, the order in which I choose people still matters.

We can again use the multiplication rule to solve this problem. For the first slot (Presidency), there are 10 options available. Then, for the second slot (Vice Presidency), we will have 9 options, because someone was already chosen for the Presidency and that person is ineligible for another position. We don't know who *exactly* was chosen for the Presidency, but we know someone will be, and that will leave us with 9 options for the Vice Presidency. Similarly, for Treasurer, we have only 8 options available to us, because 2 people already occupy the first two slots and they are ineligible for another position. Finally, we will have 7 options for the Secretary. Using the multiplication rule, this means we have $10*9*8*7 = 5,040$ total possibilities for our board!

We can now generalize this finding. Say we are picking k people from a group of n to fill a board where each position on the board has a title (in the specific example we just did, $k = 4$ and $n = 10$). For the first position, we will have n options. For the second position, we will have $n - 1$ options. For the third position, we will have $n - 2$ options. This continues until we fill all k positions. Notice that at the kth position, we will have $n - (k - 1)$ options. This is clear if you look at the relationship between position number and what we subtract from n. Recall: in the *first* position, we have n minus *zero* options. In the *second* position, we have n minus *one* options. In the *third* position, we have n minus *two* options. So, in the *kth* position, we have n minus *k-1* options! Now, using the multiplication rule, the total possibilities are:

$$n * (n - 1) * (n - 2) * \ldots * (n - (k - 1))$$

This is the same as:

$$n(n - 1)(n - 2) \ldots (n - k + 1)$$

We can simplify this expression in a fancy way. Notice that the expression almost looks like $n!$, except we don't multiply all the way down to 1 (unless $k = n$, but this is not necessarily always the case). So, let's aim to rewrite the expression using $n!$. Imagine we kept multiplying down to 1, but wanted to cancel out all these additional terms by dividing:

$$\frac{n(n - 1)(n - 2) \ldots (n - k + 1)(n - k)(n - k - 1)(n - k - 2)...(1)}{(n - k)(n - k - 1)(n - k - 2)...(1)}$$

Rewrite using factorials:

$$\frac{n!}{(n - k)!}$$

This is our final answer! It has a special name: **permutation**. Permutations have their own notation, too:

$${}_nP_k = \frac{n!}{(n - k)!}$$

To bring things back to earth, **the above equation shows how many ways there are to choose k things from a group of n without replacement such that order matters.**

Question 3: Sampling without replacement, order does not matter

The next situation we will be tackling is extremely common in probability theory: how many ways can we choose k things from a group of n things without replacement such that order doesn't matter? We can again use a specific example to discover the general rule. Imagine I have a club of 10 people (A, B, C, ..., J) and I want to choose 4 to be on the board of the club, except this time no one has a label.

Our approach to solving this specific problem may follow the same reasoning as previous problems. It could go something like this: "Well, I have 10 options for the first slot, and then 9 options for the second slot, and then 8 options for the third slot, and then 7 options for the last slot, so I have $10*9*8*7 = 5,040$ total possibilities!" This seems reasonable, but it actually **overcounts** how many possibilities we have. The key is to remember what a possibility actually looks like. Note that sthe outcome CDGB is the same outcome as BGCD, since **order does not matter**. In fact, all outcomes that have B,C,D, and G are the same, no matter where they are located. So, when we use the multiplication rule and do $10 * 9 * 8 * 7 = 5,040$, we are not getting rid of all the duplicate possibilities.

How can we get rid of these duplicates? Well, how many different ways are there to arrange any 4 letters? Let's imagine we know what this number is, call it f. This will tell us *exactly* how many duplicates we have for each 4 letter group. So, after we do $10 * 9 * 8 * 7 = 5,040$, we know that one possibility is showing up f many times. Therefore, we should divide $5,040$ by f to get the true number of possibilities.

What is f? f is the number of ways we can arrange 4 letters. We already know how to do this. Imagine 4 slots for the letters. We have 4 options for the first slot, then 3 options for the second slot, then 2 options for the third slot, and finally 1 option for the last slot. So, by the multiplication rule, we have $4 * 3 * 2 * 1$ options. So, $f = 4 * 3 * 2 * 1 = 4!$.

Now that we know what f is, we can divide $10 * 9 * 8 * 7$ by f to get our final answer:

$$\frac{(10 * 9 * 8 * 7)}{f} = \frac{(10 * 9 * 8 * 7)}{4!} = \frac{5,040}{24} = 210$$

We can now generalize: how many ways can we choose k things out of a group of n things without replacement such that order does not matter?

I have n options for the first slot, $n-1$ options for the second slot, $n-2$ options for the third slot, and so on... By the time I get to the kth slot, I will have $n-(k-1) = n-k+1$ options (see prior section for the explanation). After all this, I will have overcounted, because I thus far have assumed that order matters. As we saw with the specific example, each unique possibility appears $k!$ times, so we should divide everything by $k!$.

Therefore, the number of ways to choose k things out of a group of n people without replacement such that order matters is:

$$\frac{n*(n-1)*(n-2)*\ldots*(n-k+1))}{k!}$$

By the reasoning of the prior section, we can rewrite the expression in the numerator to be $\frac{n!}{(n-k)!}$. So, a simplified version of our answer is:

$$\frac{n!}{k!(n-k)!}$$

The above equation is how many ways to choose k things out of a group of n without replacement such that order does not matter. We commonly say this as "n choose k" and it has its own notation:

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

The last part of the above equation is called a **combination**.

Question 4: Sampling with replacement, order does not matter

This is the trickiest one. Let's begin with a specific example. We want to choose 4 people from a group of 10 (A, B, C, ..., J) to be on the board, and we can choose a person more than once, and the order of choice doesn't matter (there are no titles).

Let's again think of slots. We have 4 slots. In the first slot, there are 10 possibilities. Then, in the second slot, there will be 10 possibilities since we are sampling with replacement. Similar logic shows that there are 10 possibilities for the third and fourth slots as well. So, we might conclude that by the multiplication rule we have $10^4 = 10,000$ possibilities. Call this our original count.

Have we overcounted? Yes, but it's not so simple to correct for this. Take the board possibility ADDE. This is the same as AEDD, DDAE, DADE, and … a lot more. So, we have overcounted! How can we correct for this? Like before, we might think "let's find how many duplicates exist in our original count for each possibility." Unfortunately, this will not work, because the number of duplicates is not consistent across possibilities.

To see this, look at the possibility AAAA. This shows up in our original count just once. However, the case AAAB shows up 4 times in the original count: BAAA, ABAA, AABA, and AAAB. So, the number of duplicates is not consistent across the possibilities in our original count.

What to do? Well, we need to fundamentally change our counting method. Instead of there being 4 slots, let's imagine 10 buckets, each bucket labeled with a letter. We have 4 chips and we can put these chips in the buckets. If I put 4 chips in the A bucket, then I have the outcome AAAA. If I put two chips in the B bucket and two in the D bucket, I have BBDD. We can visualize the chips and buckets for the BBDD possibility like so:

$$|cc||cc||||||$$

The bars separate the buckets and a "c" is a chip. Above, there are two c's in the B bucket and two c's in the D bucket. Here is what AAAA would look like:

$$cccc|||||||||$$

Notice that we can go from a bucket-chip drawing to a unique possibility:

$$c||||cc|c||||$$

represents AEEF. And,

$$c||||cc|||||c$$

represents AEEJ.

So, we can reframe the original question to be "how many ways are there of arranging 9 lines and 4 c's in one row?" Well, I have 15 slots, and I need to choose 4 of them to put the c's in (the rest of the slots will be automatically filled by lines). Order doesn't matter, and there is no replacement because once I choose a slot for one c I can't put another c in that slot. So, I have

15 choose 4 ways to arrange my c's and lines, which means I have 15 choose 4 total possibilities. With notation, 15 choose 4 is:

$$\binom{15}{4} = \frac{15!}{4!(15-4)!} = 1,365$$

We can now generalize. Imagine I have n things and I want to pick k of them with replacement such that order does not matter. This means I have n buckets and k chips. So, I have $n-1$ lines and k c's. Therefore, I have $n-1+k$ slots and I need to choose k of them for my c's. Thus, my total possibilities are:

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

Notice that I could equivalently pick where my $n-1$ lines are placed and the c's will automatically fill in the remaining slots. So, we can also say that our total possibilities are:

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

There is notation for this situation, it's called n **multichoose** k, and it looks like this:

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

The above equation displays how many ways to choose k things from a group of n with replacement such that order does not matter.

Congratulations

You now understand the 4 main methods of solving combinatorics problems. Here is a table of the final formulas: Choose k things from a group of n:

	With Replacement	Without Replacement
Order Matters	n^k	$\frac{n!}{(n-k)!}$
Order Doesn't Matter	$\binom{n+k-1}{k}$	$\binom{n}{k}$