Chapter 1. Combinatorial Theory

1.1: So You Think You Can Count?

Slides (Google Drive) Alex Tsun Video (YouTube)

Before we jump into probability, we must first learn a little bit of combinatorics, or more informally, counting. You might wonder how this is relevant to probability, and we'll see how very soon. You might also think that counting is for kindergarteners, but it is actually a lot harder than you think!

To motivate us, let's consider how easy or difficult it is for a robber to randomly guess your PIN code. Every debit card has a PIN code that their owners use to withdraw cash from ATMs or to complete transactions. How secure are these PINs, and how safe can we feel?

1.1.1 Sum Rule

First though, we will count baby outfits. Let's say that a baby outfit consists of *either* a top *or* a bottom (but not both), and we have 3 tops and 4 bottoms. How many baby outfits are possible? We simply add $3 + 4 = 7$, and we have found the answer using the sum rule.

Definition 1.1.1: Sum Rule

If an experiment can either end up being one of N outcomes, or one of M outcomes (where there is no overlap), then the number of possible outcomes of the experiment is:

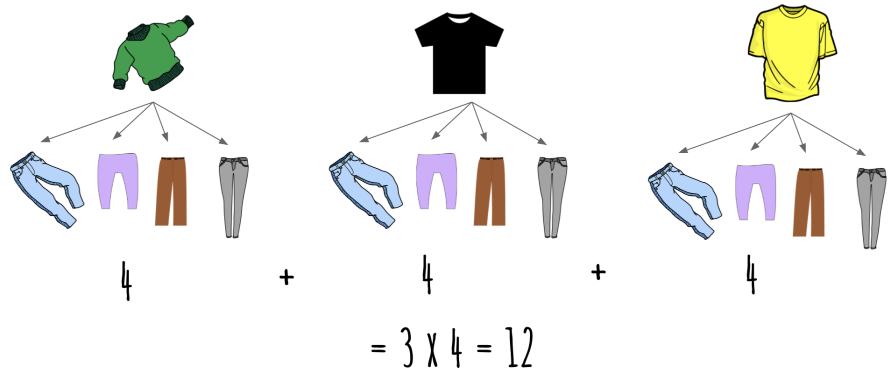
$$N + M$$

We'll see some examples of the Sum Rule combined with the Product Rule (next), so that they can be a bit more complex!

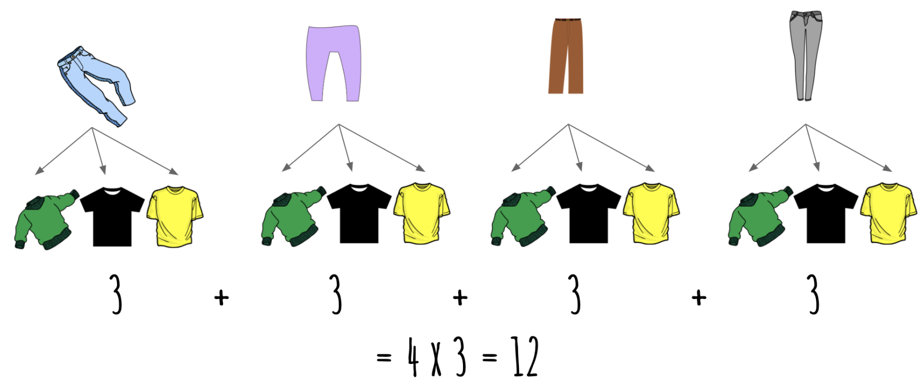
1.1.2 Product Rule

Now we will count real outfits. Let's say that a real outfit consists of *both* a top *and* a bottom, and again, we still have 3 tops and 4 bottoms. then how many outfits are possible?

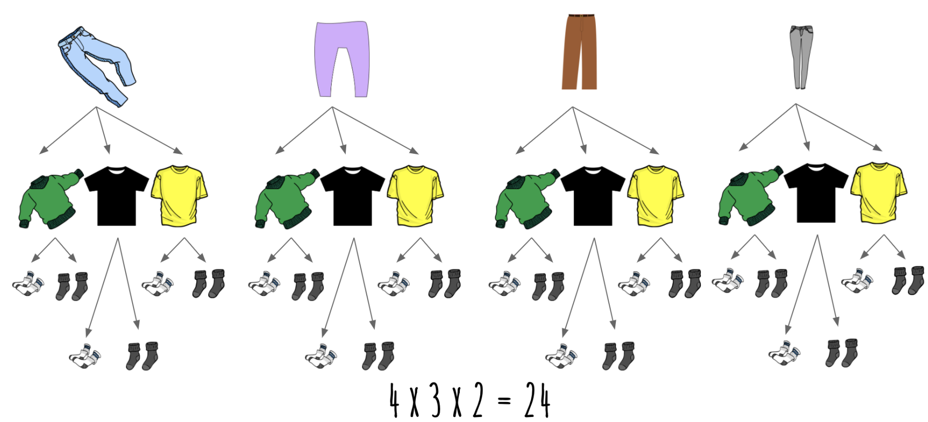
Well, we can consider this from first picking out a top. Once we have our top, we have 4 choices for our bottom. This means we have 4 choices of bottom for each top, which we have 3 of. So, we have a total of $4 + 4 + 4 = 3 \times 4 = 12$ outfit choices.



We could also do this in reverse and first pick out a bottom. Once we have our bottom, we have 3 choices for our top. This means we have 3 choices of top for each bottom, which we have 4 of. So, we still have a total of $3+3+3+3 = 4 \times 3 = 12$ outfit choices. (This makes sense - the number of outfits should be the same no matter how I count!)



What if we also wanted to add socks to the outfit, and we had 2 different pairs of socks? Then, for each of the 12 choices outlined above, we now have 2 choices of sock. This brings us to a total of 24 possible outfits.



This could be calculated more directly rather than drawing out each of these unique outfits, by multiplying our choices: 3 tops $\times$ 4 bottoms $\times$ 2 socks $=$ 24 outfits.

Definition 1.1.2: Product Rule

If an experiment has N_1 outcomes for the first stage, N_2 outcomes for the second stage, ..., and N_m outcomes for the m^{th} stage, then the total number of outcomes of the experiment is $N_1 \times N_2 \times \cdots \times N_m$.

If this still sounds "simple" to you or you just want to practice, see the examples below! There are some pretty interesting scenarios we can count, and they are more difficult than you might expect.

Example(s)

Flamingos Fanny and Freddy have three offspring: Happy, Glee, and Joy. These five flamingos are to be distributed to seven different zoos so that no zoo gets both a parent and a child :(. It is not required that every zoo gets a flamingo. In how many different ways can this be done?

Solution There are two disjoint (mutually exclusive) cases we can consider that cover every possibility. We can use the sum rule to add them up since they don't overlap!

1. **Case 1: The parents end up in the same zoo.** There are 7 choices of zoo they could end up at. Then, the three offspring can go to any of the 6 other zoos, for a total of $7 \times 6 \times 6 \times 6 = 7 \times 6^3$ possibilities (by the product rule).

2. **Case 2: The parents end up in different zoos.** There are 7 choices for Fanny and 6 for Freddy. Then, the three offspring can go to any of the 5 other zoos, for a total of $7 \times 6 \times 5^3$ possibilities.

The result, by the sum rule, is $7 \times 6^3 + 7 \times 6 \times 5^3$. (Note: This may not be the only way to solve this problem. Often, counting problems have two or more approaches, and it is instructive to try different methods to get the same answer. If they differ, at least one of them is wrong, so try to find out which one and why!) □

1.1.3 Permutations

Back to the example of the debit card. There are 10 possible digits for the each of the 4 digits of a PIN. So how many possible 4-digit PINs are there? This can be solved as $10 \times 10 \times 10 \times 10 = 10^4 = 10,000$. So, there is a one in ten thousand chance that a robber can guess your pin code (randomly).

Let's consider a stronger case where you must use each digit exactly once, so the PIN is exactly 10 digits long. How many such PINs exist?

Well, we have 10 choices for the first digit, 9 choices for the second digit, and so forth, until we only have 2 choices for the ninth digit, and 1 choice for the tenth digit. This means there are 362,880 possible PINs in this scenario as follows:

$$10 \times 9 \times \cdots \times 2 \times 1 = \prod_{i=1}^{10} i = 362,880$$

This formula/pattern seems like it would appear often! Wouldn't it be great if there were a shorthand for this?

Definition 1.1.3: Permutation

The number of orderings of N **distinct** objects, is called a permutation, and mathematically defined as:

$$N! = N \times (N-1) \times (N-2) \times \dots 3 \times 2 \times 1 = \prod_{j=1}^{N} j$$

$N!$ is read as "N factorial". It is important to note that $0! = 1$ since there is one way to arrange 0 objects.

Example(s)

A standard 52-card deck consists of one of each combination of: 13 different ranks (Ace, 2, 3,...,10, Jack, Queen, King) and 4 different suits (clubs, diamonds, hearts, spades), since $13 \times 4 = 52$. In how many ways a 52-card deck be dealt to thirteen players, four to each, so that every player has one card of each suit?

Solution This is a great example where we can try two equivalent approaches. Each person usually has different preferences, and sometimes one way is significantly easier to understand than another. Read them both, understand why they both make sense and are equal, and figure out which approach is more intuitive for you!

Let's assign each player one at a time. The first player has 13 choices for the club, 13 for the heart, 13 for the diamond, and 13 for the spade, for a total of 13^4 ways. The second player has 12^4 choices (since there are only 12 of each suit remaining). And so on, so the answer is $13^4 \times 12^4 \times 11^4 \times ... \times 2^4 \times 1^4$.

Alternatively, we can assign each suit one at a time. For the clubs suit, there are 13! ways to distribute them to the 13 different players. Then, the diamonds suit can be assigned in 13! ways as well, and same for the other two suits. By the product rule, the total number of ways is $(13!)^4$. Check that this different order of assigning cards gave the same answer as earlier! (Expand the factorials.) □

Example(s)

A group of n families, each with m members, are to be lined up for a photograph. In how many ways can the nm people be arranged if members of a family must stay together?

Solution We first choose the ordering of the families, of which there are $n!$. Then, in the first family, we have $m!$ ways to arrange them. The second family also has $m!$ ways to be arranged. And so on. By the product rule, the number of orderings is $n! \times (m!)^n$. □

1.1.4 Complementary Counting

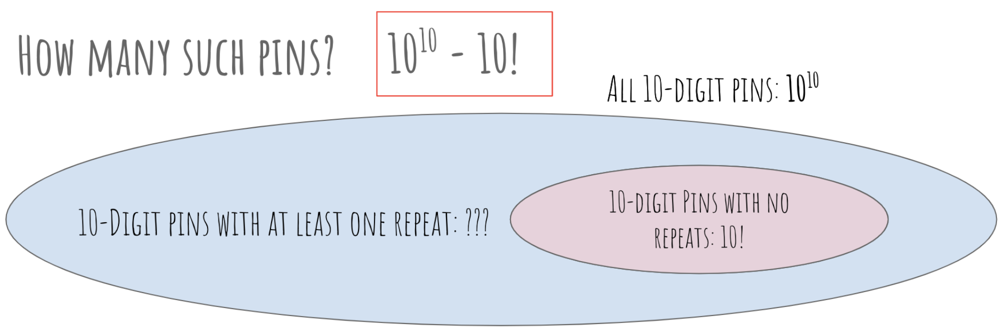
Now, let's consider an even trickier PIN requirement. Suppose we are still making a 10-digit PIN, but now at least one digit has to be repeated at least once. How many such PINs exist?

Some examples of this PIN would be 1111111111, 01234556788, or 9876598765, but the list goes on!

Let's try our "normal" approach. If we try this, we'll end up getting stuck. Consider placing the first digit - we have 10 choices. How many choices do we have for the second digit? Is this a repeated digit or not? We can try to find a product of counts of choices for each digit in different scenarios but this can become complicated as we move around which digits are repeated...

Another approach might be to count how many PINs **don't** satisfy this property, and subtract it from the total number of PINs. This strategy is called complementary counting, as we are counting the size of the complement of the set of interest. The number of possible 10-digit PINs, with no stipulations, is 10^{10} (from the product rule, multiplying 10 choices with itself for each of 10 positions). Then, we found above that the 10-digit PINs with no repeats has 10! possibilities (each digit used exactly once). Well, consider that the 10-digit PINs with at least one repeat will be all other possibilities (they could have one, two, or more repeats but certainly won't have none). This means that we can count this by taking the difference of all the possible 10-digit PINs and those with no repeats. That is:

$$10^{10} - 10!$$



Definition 1.1.4: Complementary Counting

Let $\mathcal{U}$ be a (finite) universal set, and S a subset of interest. Let $S^C = \mathcal{U} \setminus S$ denote the set difference (complement of S). Then,

$$| S | = | \mathcal{U} | - | S^C |$$

Informally, to find the number of ways to do something, we could count the number of ways to NOT to do that thing, and subtract it from the total. That is, the complement of the subset of interest is also of interest!

1.1.5 Exercises

1. Suppose we have 6 people who want to line up in a row for a picture, but two of them, A and B, refuse to sit next to each other. How many ways can they sit in a row?

   Solution: There are two equivalent approaches. The first approach is to solve it directly. However, depending on where A sits, B has a different number of options (whether A sits at the end or the middle). So we have two disjoint (non-overlapping) cases:

(a) Case 1: A sits at one of the two end seats. Then, A has 2 choices for where to sit, and B has 4. (See this diagram where A sits at the right end: $-----A$.) Then, there are 4! ways for the remaining people to sit, for a total of $2 \times 4 \times 4!$ ways.

(b) Case 2: A sits in one of the middle 4 seats. Then, A has 4 choices of seat, but B only has three choices for where to sit. (See this diagram where A sits in a middle seat: $-A----$.) Again, there are 4! ways to seat the rest, for a total of $4 \times 3 \times 4!$ ways.

Hence our total by the sum rule is $2 \times 4 \times 4! + 4 \times 3 \times 4! = 480$.

The alternative approach is complementary counting. We can count the total orderings, of which there are 6!, and subtract the cases where A and B *do* sit next to each other. There's a trick we can do to guarantee this: let's treat A and B as a *single entity*. Then, along with the remaining 4 people, there are only 5 entities. We order the entities in 5! ways, but also multiply by 2! since we could have the seating AB or BA. Hence, the number of ways they do sit together is $2 \times 5! = 240$, and the ways they do not sit together is $6! - 240 = 720 = 240 = 480$.

Decide which approach you liked better - oftentimes, one method will be easier than another!