Counting and the Pigeonhole Principle

Combinatorics and Counting

Combinatorics is the mathematical discipline that studies arrangements of objects. Determining numbers of arrangements is called counting.

The two foundational principles of counting are the product rule (aka the multiplication principle) and the sum rule. (These are NOT the same product rule and sum rule you may know from calculus.)

The Product Rule

Product Rule: If a procedure consists of carrying out two tasks, and if task 1 can be performed in n ways, and if for every way of performing task 1, there are m ways of performing task 2, then the procedure can be performed in nm ways.

Example: you order a pizza and have a choice of 1 out of 3 toppings and 1 out of 4 kinds of cheeses. Each cheese can be combined with any of the toppings. Then you have a choice between 12 different pizzas.

The product rule also works for chains of more than two tasks. In that case, the number of ways of performing the entire procedure is equal to the product of all the numbers of ways to carry out the individual tasks.

Application of the Product Rule to Phone Numbers

Problem: If a phone number consists of 6 digits, and the first digit can’t be a zero, how many different phone numbers are there?

Solution: there are 9 choices for the first digit. For any choice of the first digit, there are 10 choices for the second digit. For any choice of the second digit, there are 10 choices for the third digit, and so on. Therefore, there are $9 \cdot 10^5 = 900{,}000$ different phone numbers.

The Sum Rule

If a task can be performed in n ways of one type, or m ways of a different type, and if these two types are mutually exclusive, then the task can be performed in $n + m$ ways.

Example: you have a choice of one extra topping for a pizza. The choices are 5 kinds of extra cheese and 8 kinds of vegetables. Again, you may not choose one of each, just one. Then there are $8 + 5 = 13$ ways of choosing the topping.

Application of the Product and Sum Rules to Passwords

Sometimes, we need to use both product and sum rules to solve a counting problem.

Problem: If a password must consists of 20 characters (upper or lowercase letters of the English alphabet, or digits (0-9), or any of 7 permitted special characters), how many different passwords are there?

Solution: by the sum rule, there are $26 + 26 + 10 + 7 = 69$ choices for each character. By the product rule, there are

$$69^{20} = 5983865382382622672294674517005559601$$

$\approx 6 \cdot 10^{36}$ many different passwords.

(If an attacker could attempt one trillion or 10^{12} passwords per second, it would take her 10^{36} divided by 10^{12}, or 10^{24} seconds. This is over two million times the known age of the universe of roughly 14 billion years.)

Counting Injective Functions

We can use the multiplication principle to count the number of injective functions f from a set of n elements $S = \{x_1, x_2, \dots, x_n\}$ to a set of n elements $T = \{y_1, y_2, \dots, y_n\}$.

Designing such a function f means picking the outputs $f(x_1)$, $f(x_2)$, …, $f(x_n)$. For $f(x_1)$, we are free to choose any of the elements in T, so there are n choices. Once $f(x_1)$ has been chosen, it can't be used as another output since the function is supposed to be 1-1. Therefore, there are only $n - 1$ choices for $f(x_2)$, then only $n - 2$ choices for $f(x_3)$, and so on.

Therefore, by the product rule, the total number of injective functions from S to T is $n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 2 \cdot 1$. You may already know that this quantity is known as "n factorial" and written $n!$.

The Pigeonhole Principle

The Pigeonhole principle: if there are n boxes, and you need to reach into these boxes $n + 1$ times, then you need to reach into at least one of the boxes twice.

The Pigeonhole principle can be proved by contradiction: what if you could reach into the boxes $n + 1$ times without repeating a box? Since there are only n boxes, and you can reach into each box at most once, you reached at most n times into a box. That is a contradiction.

Simple Applications of the Pigeonhole Principle

If you have 367 people in a room, two must have the same birthday.

If you have 51 people in a room who were born in one of the 50 states of the United States, then two must have been born in the same state.

If the choices in an election were the conservative party and the liberal party, then whenever you have 3 voters in a room, 2 of them must have voted for the same party.

Combining the Pigeonhole Principle with the Product Rule

In a medical survey, respondents were asked for

- Their age group (18-30, 31-49, 50+)
- Their income (below or at least 50k per year)
- Whether they exercise regularly (yes/no)

What number of people surveyed would guarantee that two people surveyed have all matching answers to these questions?

Solution: there are 3 choices for the age group, 2 for income and 2 for exercising. By the product rule, there are 12 ways of answering this survey. Therefore, if 13 or more people were surveyed, it is guaranteed that two will have all matching answers.

An advanced application of the Pigeonhole Principle

Sometimes, it is not so obvious what the “boxes” should be in an application of the Pigeonhole principle. The following is an example of that.

What size subset of $\{1,2,3,4,5,6\}$ will guarantee that two of its elements add up to 7?

We can only get 7 as 1+6, 2+5, 3+4. These three groups of two numbers are the boxes. Therefore, if we select 4 unique numbers, we will have to reach into one of the three boxes twice, and therefore obtain a pair that adds up to 7.

The Generalized Pigeonhole Principle

Let us illustrate a generalization of the Pigeonhole principle with an example first. If there are 10 drawers, and you reach into these drawers 21 times, then you must have reached into at least one of them at least 3 times.

The Generalized Pigeonhole principle: if there are n boxes and you reach into these boxes $nm + 1$ times, then you need to reach into at least one of the boxes $m + 1$ times.

Once again, we prove this by contradiction: what if you could reach into the boxes $nm + 1$ times and only reach into each box at most m times? Since there are only n boxes, then you reached in at most nm times, which is less than $nm + 1$. That's a contradiction.

Application of the Generalized Pigeonhole Principle

If you have 15 people in a room, 3 must have been born on the same day of the week.

The same conclusion would hold if we had 16, 17, 18, 19, 20 or 21 people in a room. Only at 22 people do we reach the threshold for a stronger conclusion:

If you have 22 people in a room, 4 must have been born on the same day of the week.

This observation gives rise to the following slightly more powerful version of the Generalized Pigeonhole principle:

If there are n boxes and you reach into these boxes m times, then you need to reach into at least one of the boxes $\left\lceil \frac{m}{n} \right\rceil$ times.

In the day of the week example above, the boxes are days of the week, so $n = 7$, and $m = 15, 16, 17, \ldots, 22$.