(The Joy of) Counting

Counting problems arise throughout Mathematics and Computer Science.

Most counting problems can be rephrased as finding the cardinality of a set. In these notes, I will try to give both a formal definition and an intuitive definition of the different properties we will study.

Types of counting problems:

- Number of licence plates in California.
- Number of steps / operations in an algorithm
- Number of words of a given length
- Number of bit strings of a given length.

Reminder

Cartesian product: Let A and B be sets. The Cartesian product of A and B, denoted $A \times B$, is the set of all ordered pairs (a, b), where $a \in A$ and $b \in B$.

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

Basic counting principles:

1) The product rule

Let us suppose we want to explore the number of possibilities for a list whose elements are pairs of 2 values. If there are n_1 ways to choose the first value and n_2 ways to choose the second value after each option for the first value, then there are $n_1 \times n_2$ ways to build the list.

In formal set theory, $|A \times B| = |A| \times |B|$

and more generally,

$$|A_1 \times A_2 \times \dots A_n| = |A_1| \times |A_2| \times \dots |A_n|$$

Examples:

a) Number of 4 "digit" pins, in which each "digit" can be a number from 0 to 9, or a letter from a to z.

Let A be the set $\{0, \dots, 9, a, \dots, z\}$.

A pin is an element of $A \times A \times A \times A$.

Therefore, the number of possible pins is $N = |A|^4 = 36^4 = 1679616$

b) Number N_n of bit strings of size N.

Each bit can take 2 values: 0 or 1

According to the product rule, $N_n = 2^n$

c) What is the value of k after the following piece of code: ③

$k \leftarrow 0$

for ($i_1 \leftarrow 1$; $i_1 \leq N_1$; Step = 1)

  for ($i_2 \leftarrow 1$; $i_2 \leq N_2$; Step = 1)

    for ($i_3 \leftarrow 1$; $i_3 \leq N_3$; Step = 1)

      for ($i_4 \leftarrow 1$; $i_4 \leq N_4$; Step = 1)

        $k \leftarrow k + 1$

k is a counter. k counts the number of quadruplets (i_1, i_2, i_3, i_4) that can be formed. Using the product rule: $k = N_1 \times N_2 \times N_3 \times N_4$

d) Number of subsets of a set S with $|S| = N$.

Let $\mathcal{P}(S)$ be the set of subsets of S and let B be the set of all bit strings of size N.

Let
$$f: \mathcal{P}(S) \longrightarrow B$$
$$A \longrightarrow f(A)$$

where $f(A)$ is the bit string representation of A, defined according to membership.

f is clearly a bijection.

Therefore $|\mathcal{P}(S)| = |B| = N_n = 2^N$ (problem b above)

2) The sum rule

If the list to be counted can be splitted into two disjoint lists of size m_1 and m_2, then the original list contains $m_1 + m_2$ elements.

In formal set theory, $|A \cup B| = |A| + |B|$ if $A \cap B = \emptyset$

Examples:

a) You need to choose one dessert among 12 types of fruits, 6 types of cakes, and 4 types of candies. How many choices do you have?

The list of dessert can be splitted into 3 disjoint lists of size 12, 6, and 4, respectively. Therefore, the total number of options is 22.

b) What is the value of k after:

$k \leftarrow 0$

For ($i_1 \leftarrow 1$; $i_1 \leq N_1$; Step $= 1$)

  $k \leftarrow k+1$

For ($i_2 \leftarrow 2$; $i_2 \leq N_2$), Step $= 1$

  $k \leftarrow k+1$

$k = N_1 + N_2$

c) The extended sum rule

When the list to count cannot be (a is not) divided into two disjoint sublists, we cannot use the sum rules directly, as it may lead to overcounting. If it is divided into two lists with n_1 and n_2 elements, respectively, and those two lists have n common elements, then the initial list has $n_1 + n_2 - n$ elements.

In formal set theory, $|A \cup B| = |A| + |B| - |A \cap B|$ (inclusion - exclusion principle).

d) Rule of complement

If a list L has n objects, and m of those have a given property P, then the number of objects in L that do not have the property P is $n - m$.

In formal set theory, $|\overline{A}| = |\mathcal{U}| - |A|$

Examples

a) How many bit strings of length 8 either start with 1, or end with 00?

Let us call L the list of such bit strings.
L can be divided into 2 sublists L_1 and L_2 corresponding to bit strings starting with 1, and ending with 00, respectively.
We use the product rule to compute $|L_1|$ and $|L_2|$:

$$|L_1| = 2^7 \quad \text{and} \quad |L_2| = 2^6$$

L_1 and L_2 are not disjoint. $L_1 \cap L_2$ contains the bit strings that start with 1 and end with 00.

$$|L_1 \cap L_2| = 2^5$$

Therefore: $|L| = 2^7 + 2^6 - 2^5 = 2^5(2^2+2-1) = 5 \times 2^5 = 160$.

b) How many positive integers less than 1,000 consist of different digits from {1, 3, 7, 9}.

Let us call L the list of such integers. L can be divided into 3 lists, L_1, the integers with 1 digit, L_2, the integers with 2 digits, and L_3 the integers with 3 digits.

L_1: one digit from {1, 3, 7, 9} : $|L_1| = 4$

L_2: two distinct digits from {1, 3, 7, 9} : $|L_2| = 4 \times 3$

L_3: three distinct digits from {1, 3, 7, 9} : $|L_3| = 4 \times 3 \times 2$

Therefore $|L| = |L_1| + |L_2| + |L_3| = 4 + 12 + 24 = 40$

c) Find the number of passwords of length 6, 7, or 8 characters where each character is an uppercase letter or a digit. Each password must contain a digit.

Let L be the set of password. L can be splitted into 3 sublists, L_6, L_7, and L_8, corresponding to passwords with 6, 7, or 8 characters.

Let us find $|L_6|$, $|L_7|$, and $|L_8|$.

Let U_6 be the set of passwords of length 6 that may, or may not contain a digit. $\overline{L_6}$ is the subset of U_6 whose elements do not contain a digit.

$$|U_6| = 36^6 \quad \text{and} \quad |\overline{L_6}| = 26^6$$

Therefore, according to the rule of complement,

$$|L_6| = 36^6 - 26^6$$

Similarly,

$$|L_7| = 36^7 - 26^7$$

$$|L_8| = 36^8 - 26^8$$

Therefore, $|L| = |L_6| + |L_7| + |L_8|$

$$|L| = 36^8 + 36^7 + 36^6 - 26^8 - 26^7 - 26^6$$

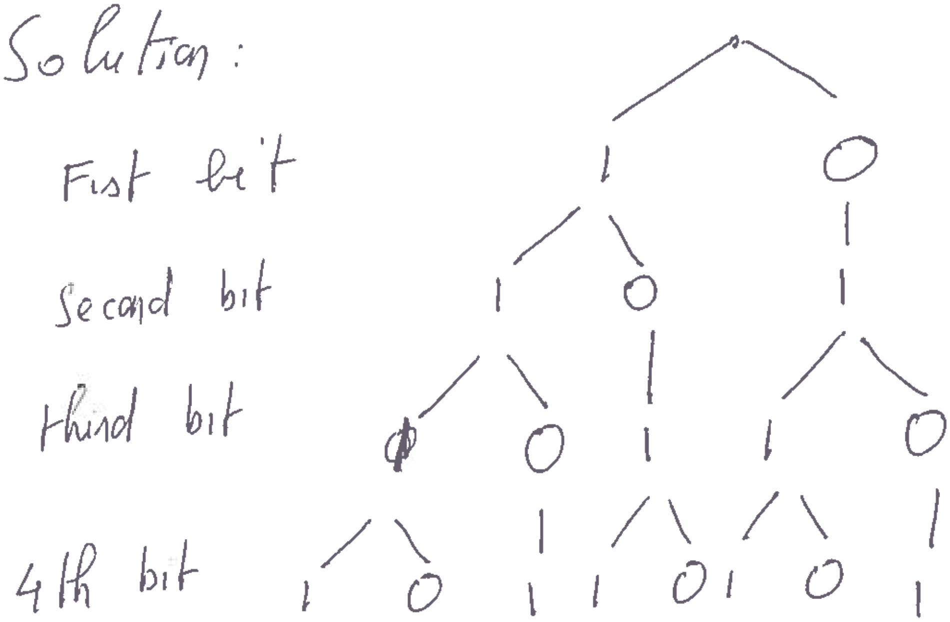
$$|L| = 2,684,483,063,360.$$

Tree diagram

Simple counting problems can be solved using tree diagrams. A tree consists of a root, a number of branches leaving the root, with each branch representing a possible choice. The end point of the branches are the leaves. The number of leaves is the solution to the counting problem.

Example: How many bit strings of length 4 do not have 2 consecutive 0s?

Solution:



(1111) (1110) (1101) (1011) (1010) (0111) (0110) (0101)

→ 8 solutions.

The pigeonhole principle

(also known as the Dirichlet drawer principle).

Theorem: If $k+1$ objects or more are placed into k boxes, then there is at least one box that contains two or more elements.

Proof (by contradiction): Suppose each box contains at most one object. Then according to the sum rule, the total number of object would be at most k → contradiction.

A formal expression of the pigeonhole principle:

Let f be a function from a set A to a set B with $|A| > |B|$. Then f cannot be injective (one to one).

Examples

a) Among 13 people, at least two have their birthday the same month.

b) Prove that for any 11 positive integers, some pair of them will have a difference divisible by 10!

Proof (using the concept of "boxes")

Let us place each integer n in a box B_i, where i is the remainder of the division of n by 10

There are 10 such boxes. According to the pigeon hole principle, 2 of the 11 integers fall in the same box: those two numbers i and j have the same remainders when divisible by 10 : i mod 10 = j mod 10, therefore $i \equiv j$ [10], i.e their difference is divisible by 10!

The generalized Pigeon hole principle:

If N objects are placed into k boxes, then there is a least one box containing $\lceil N/k \rceil$ objects.

Example Show that if seven distinct integers are chosen from the first 10 positive integers, then there must be at least two pairs of these integers with the sum 11.

Let S be the set of pairs of integers among the 10 first, whose sum is 11:

$S = \{ (1, 10), (2, 9), (3,8), (4,7), (5,6) \}$

$|S| = 5$.

Each element of S is a "box". An integer ≤ 10 is placed in one box based on its value.

According to the generalized pigeonhole principle, among 7 integers, at least 2 are in the same box.

Their sum is eleven.

The 5 remaining integers are placed into 4 remaining boxes, and again at least 2 of them are in the same box. Their sum is also 11.

Permutations

A permutation of a set of distinct objects is an ordered arrangement of these objects.
An ordered arrangement of r elements of a set is called a r-permutation.

Example: $S = \{a, b, c\}$. $\{c, a, b\}$ is a permutation of S. $\{b, a\}$ is a 2-permutation of S.

Theorem: The number of r-permutations of a set with n distinct elements is

$$P(n, r) = n(n-1)\ldots(n-r+1)$$

Proof: product rule!

$P(n, n)$ is written $n!$ (factorial n)

i.e $n! = n \times (n-1) \times \ldots 1$

Note that by definition, $0! = 1$.

Note also that

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

Examples

a) How many ways are there to arrange 5 people in a line?

Call the persons ①, ②, ③, ④ and ⑤. An arrangement of the persons in a line is a permutation of the set $S = \{①, ②, ③, ④, ⑤\}$ with $|S| = 5$.

$$P(5,5) = 5! = 120.$$

b) How many arrangements are there of the letters in the words MATCH?

$$P(5, 5) = 120$$

c) How many arrangements have the letters M, A side by side?

$S = \{ [MA], T, C, H\}$ $\quad P(4,4) = 24.$

Combinations

Let us now consider the question: how many subsets of size r are there in a set S of size n?

This differs from searching r-permutations in S: r-permutations are ordered arrangements while sets and subsets are not ordered.

For example, let $S = \{a, b, c\}$.

There are 6 2-permulations of S: $(a,b), (b,a)$ (a,c) (c,a) (b,c) (c,b)

but only 3 subsets of 2 elements: $\{a,b\}, \{a,c\}, \{b,c\}$

Definition: Subsets of size r from a set S are called r-combinations of S. We write $C(n, r)$, or C_n^r, or $\binom{n}{r}$ (read "n choose r") for the number of r-combinations in a set S of size n.

Theorem: $$C(n, r) = \frac{P(n, r)}{r!} = \frac{n!}{r!(n-r)!}$$

Proof Let s be a r-permutation of a set S with $|S| = n$, and let $A(s)$ be the corresponding subset of S containing all elements of s. All r-permutations of $A(s)$ are r-permutations of S. There are $r!$ r-permutations that can be associated to $A(s)$.

Therefore, $C(n, r) = \frac{P(n, r)}{r!}$

Examples:

a) How many subsets of 2 elements are there in a set of N elements?

$$C(N, 2) = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2}$$

b) How many subsets of 3 elements are there in a set of N elements

$$C(N, 3) = \frac{N!}{3!(N-3)!} = \frac{N(N-1)(N-2)}{6}$$

c) How many subsets of $(N-2)$ elements are there in a set of N elements?

$$C(N, N-2) = \frac{N!}{2!(N-2)!} = \frac{N(N-1)}{2} = C(N,2).$$

Theorem: Let N and r be non-negative integers with $r \leq N$

$$C(N, r) = C(N, N-r)$$

Proof:

$$C(N,r) = \frac{N!}{r!(N-r)!}$$

$$C(N, N-r) = \frac{N!}{(N-r)!(N-N+r)!} = C(N,r)$$

Examples:

a) How many bit strings of length n contains exactly r 1s?

The position of the r 1s in the bit string of length n form a r-combination of $\{1, \dots, n\}$. Hence there are $C(n,r)$ bit strings of length n with exactly r 1s.

b) A club has 25 members:

- How many ways are there to choose 3 members of the club to serve on a committee $\rightarrow C(25,3)$
- How many ways are there to choose a president, a vice president, and a treasurer for the club? $\rightarrow P(25,3)$

c) A club of 10 women and 8 men is forming a 5-person committee. Clearly, there are $C(18,5)$ possible committees. How many:

- contains exactly 3 women?

$$C(10,3) \times C(8,2) = 3360$$

- contains at least 3 women?

$$C(10,3) \times C(8,2) + C(10,4) \times C(8,1) + C(10,5) = 5292$$

The binomial theorem

Let a and b be two real numbers, and let n be a non negative integer. Then:

$$(a+b)^n = \sum_{i=0}^{n} C(n,i)\, a^{n-i}\, b^i$$

Example:

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$$

Interesting properties

$$\sum_{i=0}^{n} C(n, i) = 2^n$$

$$\sum_{i=0}^{n} C(n, i)(-1)^i = 0$$

Pascal's identity

Let n and k be positive integers with $n \geq k$

$$C(n+1, k) = C(n, k-1) + C(n, k)$$

Proof:

$$C(n, k-1) + C(n, k) = \frac{n!}{(n-k+1)!\,(k-1)!} + \frac{n!}{(n-k)!\,k!}$$

$$= \frac{n!\,k + n!\,(n-k+1)}{(n-k+1)!\,k!} = \frac{(n+1)!}{(n+1-k)!\,k!} = C(n+1, k)$$

Applications:

0 $C(0,0)=1$

1 $C(1,0)=1$ — $C(1,1)=1$

2 $C(2,0)=1$ — $C(2,2)=2$ — $C(2,2)=1$

3 $C(3,0)=1$ — $C(3,1)=3$ — $C(3,2)=3$ — $C(3,3)=1$

4 $C(4,0)=1$ $C(4,1)=4$ $C(4,2)=6$ $C(4,3)=4$ $C(4,4)=$