Combinatorics

In combinatorics, we consider problems of the following kind: how may ways we can from an n-element set $\{x_1, x_2, \ldots, x_n\}$ *choose* k elements?

The answer depends on two things:

- Is repetition of chosen elements allowed?
- Does the order of elements has significance?

From these, we can build 4 different cases.

1. Selection without repetition when order has significance
2. Selection without repetition when order does not matter
3. Selection with repetition when order has significance
4. Selection with repetition when order does not matter

The *factorial* function is defined by the product:

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-2) \cdot (n-1) \cdot n$$

The factorial of 0 is 1, or in symbols, $0! = 1$.

Proposition 1. *There are $n!$ different ways of arranging n distinct objects into a sequence.*

Example 2. Find the number of 5 "words" that can be formed with using all the letters of the word "CHAIR".

Because "CHAIR" contains 5 distinct letters, the number of words that can be formed with these 5 letters is

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120.$$

The *Binomial coefficient* is defined as

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

A Selections without repetition

Let us begin with selecting k elements from an n-element set $\{x_1, x_2, \ldots, x_n\}$ without repetition when **order has significance**. If $n = 3$ and $k = 2$, then the task is to choose from the set $\{x_1, x_2, x_3\}$ 2 elements, and the possible cases are:

$$x_1x_2,\ x_1x_3,\ x_2x_1,\ x_2x_3,\ x_3x_1,\ x_3x_2.$$

In general case, we can deduce the number of combinations as follows:

- The first element is chosen from the set of n elements
- The second element is chosen from the set of $n-1$ elements
- The third element is chosen from the set of $n-2$ elements
- The kth element is chosen from the set of $n-(k-1) = (n-k+1)$ elements

A permutation is an arrangement of objects, without repetition, and order being important.

Proposition 3. *The number of possibilities to select k elements from an n-element set is*

$$n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

when order has significance.

Consider next how to choose k elements from an n-element set $\{x_1, x_2, \ldots, x_n\}$ without repetition when **order does not matter**. If $n = 3$ and $k = 2$, there are only the possible combinations:

$$x_1x_2,\ x_1x_3,\ x_2x_3.$$

Because the order does not matter, the selections x_1x_2 and x_2x_1 are the same, similarly as x_2x_3 and x_3x_2.

Because we are selecting k elements, these elements can have $k!$ different combinations. Because all these $k!$ combinations are considered the same, we need no divide the number given in Proposition 3 by $k!$ to get the number of different combinations.

Proposition 4. *The number of possibilities to choose k elements from an n-element set is*

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

when order does not matter.

B Selections with repetition

Consider selection of k elements from the set $\{x_1, x_2, \ldots, x_n\}$, when the order of elements has significance and repetition is allowed. If $n = 3$ and $k = 2$, there are nine different possibilities:

$$x_1x_1,\ x_1x_2,\ x_1x_3,\ x_2x_1,\ x_2x_2,\ x_2x_3,\ x_3x_1,\ x_3x_2,\ x_3x_3.$$

In the general case, the number of selections is n^k, because the first element can be chosen in n different ways. Because we are considering sections with repetition, also the second element can be chosen in n different ways. Because we are selecting k elements, the number of different selections is

$$\underbrace{n \cdot n \cdots n}_{k} = n^k.$$

Proposition 5. *We can choose k items from n items set with repetition in*

$$n^k$$

different ways when order of the elements has significance.

Finally, we consider selection of k elements with repetition from the $\{x_1, x_2, \ldots, x_n\}$ when the **order does not matter**. If $n = 3$ and $k = 4$, then there are 15 different choices:

$x_1x_1x_1x_1$	$x_1x_1x_1x_2$	$x_1x_1x_1x_3$	$x_1x_1x_2x_2$	$x_1x_1x_2x_3$
$x_1x_1x_3x_3$	$x_1x_2x_2x_2$	$x_1x_2x_2x_3$	$x_1x_2x_3x_3$	$x_1x_3x_3x_3$
$x_2x_2x_2x_2$	$x_2x_2x_2x_3$	$x_2x_2x_3x_3$	$x_2x_3x_3x_3$	$x_3x_3x_3x_3$

Each such selection can be represented as a sequence consisting of the digits 0 and 1. For instance, the selection $x_1x_2x_3x_3$ corresponds to 101011. The idea is that "zeros" act like as separators between different elements and "ones" tell how many times each element appears in the selection:

$$\begin{array}{cccccc} 1 & 0 & 1 & 0 & 1 & 1 \\ x_1 & & x_2 & & x_3 & x_3 \end{array}$$

Because in this case, the separator "0" appears exactly twice in the sequence of the length $4 + 2$, the number of selections is

$$\binom{6}{2} = \frac{6!}{2!\,4!} = \frac{5 \cdot 6}{2} = 15.$$

Because this is a selection in which order does not matter, we may consider the case in which each selection elements x_1 come first, then elements x_2, and so on. For such a selection, we may attach a sequence of zeros and ones, whose length is $k + n - 1$, and which is formed as explained above. Each such sequence contains $n - 1$ zeros and k ones.

It is clear that there is a bijection between selections and sequences. The $n-1$ zeros can be in any of the $n + k - 1$ positions. Therefore, the number of choices is

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

Proposition 6. *We can choose k items from n items set with repetition in*

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

different ways, when the order does not matter.

Our results are summarized in the following table:

	Order has significance	Order has no significance
Without repetition	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$
With repetition	n^k	$\binom{n+k-1}{k}$

C Examples

Example 7. We have already considered representing subsets as bit vectors. Let U be a finite set in which the elements have a fixed order $U = \{x_1, x_2, \ldots, x_n\}$. The *bit vector* of the subset $A \subseteq U$ is

$$\text{bit}(A) = (v_1, \ldots, v_n)$$

defined as

$$v_i = \begin{cases} 1 & \text{if } x_i \in A \\ 0 & \text{if } x_i \notin A \end{cases}$$

Now it is clear that as a mapping "bit" forms a bijection between the subsets of A and sequences of $\{0, 1\}$-vectors of length n.

Let $U = \{1, 2, 3, 4, 5\}$. If $A = \{1, 3, 4\}$ and $B = \{2, 5\}$, then

$$\text{bit}(A) = (1, 0, 1, 1, 0) \quad \text{and} \quad \text{bit}(B) = (0, 1, 0, 0, 1).$$

In particular, $\text{bit}(\emptyset) = (0, 0, 0, 0, 0)$ and $\text{bit}(U) = (1, 1, 1, 1, 1)$.

It is easy to see that the number of $\{0, 1\}$-vectors of length n is 2^n. This is because we selection with repetition n elements from 2 possible ones when the order has significance. Therefore, the size of $\wp(U)$ is 2^U.

Example 8. In a Finnish lottery called "Lotto" 7 numbers out of 40 are drawn. This can be seen as selection without repetition, when the order has no significance: the row 14, 7, 40, 25, 4, 5, 37 is the same Lotto row as 4, 5, 7, 14, 25, 37, 40. The number of different Lotto rows is therefore

$$\binom{40}{7} = \frac{40!}{7!\,33!} = \frac{34 \cdot 35 \cdot 36 \cdot 37 \cdot 38 \cdot 39 \cdot 40}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7} = 18\,643\,560.$$

Let us consider the numbers

$$\binom{n}{k}$$

in a detailed manner. These numbers are found in **Pascal's triangle** (Blaise Pascal 1623–1662). The rows of Pascal's triangle are conventionally enumerated starting with

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1

Figure 1: Pascal's triangle

row $n = 0$ at the top (the 0th row). The entries in each row are numbered from the left beginning with $k = 0$. The triangle may be constructed in the following manner: In row 0 (the topmost row), there is a unique nonzero entry 1. Each entry of each subsequent row is constructed by adding the number above and to the left with the number above and to the right, treating blank entries as 0 The entry in the nth row and kth column of Pascal's triangle is

$$\binom{n}{k},$$

where $k = 0, \dots, n$. For instance,

$$\binom{6}{2} = 15.$$

The construction of the previous paragraph may be written as follows:

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

for any non-negative integer n and any integer $0 \le k \le n$. Proving this formula is left as an exercise.

Proposition 9. *For all $n \in \mathbb{N}$,*

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n.$$

Proof. Let k be an integer between 0 and n, Now $\binom{n}{k}$ is the number of k-sized subsets when the size of the "universe" is n. If instance, if $U = \{a, b, c, d\}$, then its 2-sized subsets are

$$\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}.$$

There are $\binom{4}{2} = 6$ of them. This is selection without repetition, when the order of the elements is not significant. (that is, $\{a, b\}$ is the same set as $\{b, a\}$). In the left-hand side of the formula there is the sum of the sizes of all subsets of an n-sized set This is the number of all subset of an n-sized set, that is, 2^n. □

The next formulas are well known:

$$(a + b)^2 = a^2 + 2ab + b^2 \quad \text{and} \quad (a - b)^2 = a^2 - 2ab + b^2.$$

Let us consider the general form $(a + b)^n$.

Proposition 10. *For all $n \in \mathbb{N}$,*

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n.$$

Proof. Let us consider the product

$$\underbrace{(a + b)(a + b) \cdots (a + b)}_{n}.$$

When we compute the product, we select from each $(a+b)$ either a or b. This means that the number of the terms $a^{n-k}b^k$ equals the number of ways we can select k number of b's (and $n-k$ number of a's). According to Proposition 4 the number of such selections is $\binom{n}{k}$. □

Example 11. According to Proposition 10, we have

$$\begin{aligned} (a+b)^4 &= \binom{4}{0}a^4 + \binom{4}{1}a^{4-1}b + \binom{4}{2}a^{4-2}b^2 + \binom{4}{3}ab^{4-1} + \binom{4}{4}b^4 \\ &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \end{aligned}$$

for all $a, b \in \mathbb{R}$.

Induction

The **Induction Principle** is a powerful tool for proving statements related to integers.

Let $V(n)$ be an assertion (=predicate) concerning the integer n. When we want to prove by induction that the statement $V(n)$ is valid for all $n \in \mathbb{N}$, the procedure is the following:

1. We show that $V(0)$ is valid.
2. We assume that $V(k)$ is valid for some value $k \geq 0$. This is so-called **induction hypothesis**. Then we prove (usually by applying the induction hypothesis) that also $V(k+1)$ is valid.

The validity of the induction principle can be justified informally by the following infinite sequence:

- Because $V(0)$ valid and $V(0) \Rightarrow V(1)$, $V(1)$ is valid
- Because $V(1)$ valid and $V(1) \Rightarrow V(2)$, $V(2)$ is valid

$\vdots$

- Because $V(k-1)$ valid and $V(k-1) \Rightarrow V(k)$, $V(k)$ is valid

$\vdots$

Therefore, $V(n)$ is valid for all integers n.

The **well-ordering principle** states that:

Every nonempty set of natural numbers has a smallest element.

The validity of the principle of mathematical induction follows from the well-ordering principle, as we can see in the following proposition.

Proposition 12. *Let S be a subset of $\mathbb{N}$ such that*

(i) $0 \in S$;

(ii) $(\forall k \in \mathbb{N})\, k \in S \Rightarrow k+1 \in S$.

Then $S = \mathbb{N}$.

Proof. Suppose that the subset S satisfies conditions (i) and (ii). Assume for contradiction that S is not the whole $\mathbb{N}$. Thus, the set $\mathbb{N} \setminus S$ is non-empty. Therefore, by the well-ordering principle, S has a smallest element m.

It follows from condition (i) that m is not 0 (for $0 \in S$ and $m \notin S$). On the other hand, $m - 1 \in S$, since m is the smallest element of the set $\mathbb{N} \setminus S$. But from condition (ii) it follows that that $(m-1)+1 = m$ belongs to the set S. This contradicts with that $m \in \mathbb{N} \setminus S$. From the contradiction we can conclude that that the counter-assumption was incorrect. Thus $S = \mathbb{N}$. □

How the induction principle is obtained as a result of the previous proposition? If $V(n)$ is a statement which should be proved, we may denote

$$S = \{n \in \mathbb{N} \mid V(n) \text{ is true } \},$$

Now the previous theorem says that $V(n)$ is true for all values of $n = 0, 1, 2, \ldots$ (i.e. $S = \mathbb{N}$) provided that

1. $V(0)$ is valid.
2. The implication $V(k) \Rightarrow V(k+1)$ is valid for all $n \geq 0$.

Example 13. We show that

$$1 + 2 + 4 + \cdots + 2^n = 2^{n+1} - 1$$

for all $n \geq 0$.

The Base Case For $n = 0$, $2^0 = 1$ and $2^{0+1} - 1 = 1$.

Inductive Step Assume that $V(k)$ is valid, that is,

$$1 + 2 + 4 + \cdots + 2^k = 2^{k+1} - 1$$

We prove that also $V(k+1)$ is valid:

$$\begin{aligned} 1 + 2 + 4 \cdots + 2^k + 2^{k+1} &= (2^{k+1} - 1) + 2^{k+1} \\ &= 2 \cdot 2^{k+1} - 1 \\ &= 2^{k+2} - 1 \\ &= 2^{(k+1)+1} - 1 \end{aligned}$$

We have now proved that $V(n)$ is true for all $n \geq 0$.

There are several **variations** of the induction principle. The induction principle can be used, for example, to prove that a claim is valid for all integers $n \geq n_0$, where n_0 is any integer (note that n_0 can also be negative). The structure of such a proof is the following:

1. $V(n_0)$ is valid
2. The implication $V(k) \Rightarrow V(k+1)$ is valid for all $k \geq n_0$.

The correctness of the above can be established as follows. Proving that $V(n)$ is valid for all $n \geq n_0$ may be proved by showing that the statement $U(n) = V(n+n_0)$ is valid for all $n \geq 0$. This is because:

1. $U(0)$ is valid $\iff$ $V(n_0)$ is valid.
2. The implication $U(m) \Rightarrow U(m+1)$ is valid for all $m \geq 0$ if and only if the implication $V(k) \Rightarrow V(k+1)$ is valid for all $k \geq n_0$ – think that $k = m + n_0$.

Example 14. We prove that for all $n \geq 1$,

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Base case The claim holds for $n = 1$

Inductive step Assume that the claim holds for $n = k$:

$$1 + \cdots + k \;=\; \frac{k(k+1)}{2}.$$

We prove that the claim also holds for $n = k + 1$:

$$(1 + 2 + \cdots + k) + (k + 1) = \frac{k(k+1)}{2} + (k + 1).$$

The right hand side simplifies as:

$$\begin{aligned}\frac{k(k+1)}{2} + (k+1) &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{(k+1)((k+1) + 1)}{2}.\end{aligned}$$

We have that

$$1 + 2 + \cdots + k + (k+1) \;=\; \frac{(k+1)((k+1) + 1)}{2}.$$

This completes the proof.

Example 15. Note that the *base case* is important. Without verifying that we can "prove anything", like that for all $n \in \mathbb{N}$, $n = n + 5$.

Example 16. For all $n \geq 1$, $3^n - 1$ is a multiple of 2.

Example 17. For all $n \geq 1$, $1 + 3 + 5 + \cdots + (2n - 1) = n^2$.

Another practical modification of the induction principle is that the induction assumption assumes that "$V(i)$ is valid for **all** values $i \leq k$" instead of "$V(k)$ is valid" instead of. In this case, we prove that

1. $V(0)$ is valid.
2. We show that if for any k, the claims $V(0), V(1), \ldots, V(k)$ are valid, then $V(k+1)$ is also valid.

The validity of this proof structure can be seen considering the statement

$$U(k) = \text{“}V(i) \text{ is valid for all } i \leq k\text{”}.$$

Of course, these two generalizations can be combined. This is called **strong induction**.

Theorem 18. *Any natural number $n > 1$ can be written as a product of primes.*

Proof. Let $V(n) = n$ can be written as a product of primes.

Base case $V(2)$ is the proposition that 2 can be written as a product of primes. This is true, since 2 can be written as the product of one prime, itself. Remember that 1 is not prime.

Inductive Step The inductive hypothesis states that all integers from 2 to k can be written as a product of primes.

We have two cases (i) $k + 1$ is prime and (ii) $k + 1$ is not prime.

(i) $k + 1$ can be written as the product of one prime, itself.

(ii) By the definition of prime numbers, there exist integers a and b such that $2 \leq a, b < k + 1$ and $k + 1 = ab$. By the inductive hypothesis, both a and b can be written as a product of primes.Hence $k + 1$ can be written as a product of primes. □