

(The joy of) Counting

Counting problems arise throughout Mathematics, Computer Science, and Data 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. First, however, I need to provide enough background on set theory

1 Basic set theory

1.1 What are sets?

Definition

A set is an unordered collection of objects

Examples:

- The set S of all numbers between 0 and 10:

$$S = \{1, 3, 5, 7, 9\}$$

- The days of the week:

$$W = \{ \text{"Monday"}, \text{"Tuesday"}, \text{"Wednesday"}, \text{"Thursday"}, \text{"Friday"}, \text{"Saturday"}, \text{"Sunday"} \}$$

- Sets of numbers:

$$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$$

1.1.1 How to describe a set?

A set is well defined if there is a clear rule that defines membership. Let us consider for example the set S of all even integers between 2 and 20. There are two methods for describing S symbolically:

- a) **The roster notation** is a complete, or implied listing of all elements of S :

i) $S = \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20\}$ (*complete*)

ii) $S = \{2, 4, 6, 8, \dots, 20\}$ (*implied, where the dots ... defines a implicit rule for the corresponding numbers*)

- b) **The set builder notation** is a symbolic representation used when the roster notation is cumbersome:

$$S = \{x | x \in \mathbb{N} \text{ and } 2 \leq x \leq 20 \text{ and } x \text{ is even}\}$$

The set builder notation clearly identifies that a set is defined within a universe of discourse, or domain (in the example above, this domain is \mathbb{N}).

There is an additional method that is useful to get intuitions about sets: **The Venn diagram:**

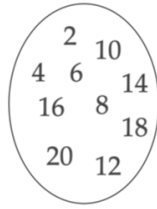


Figure 1: A Venn diagram representation of the set of all even integers between 2 and 20

1.1.2 Paradoxes

The objects, or elements of a set do not have to be numbers. We can even build sets of sets. This leads to some paradoxes.

The barber's paradox. Suppose there is a town with one barber, and all the men in this town keep clean shaven. Let S be the set of men that are shaved by the barber. If we try to define S with “the barber shaves all, and only those men who do not shave themselves”, we would reach a paradox, as we could not decide if the barber belongs to S !!

Russell's paradox. Let M be the set of all sets that do not contain themselves as member: $M = \{A \mid A \notin A\}$. M seems well defined...however, just like for the barber's paradox, it leads to a paradox when we try to check if M belongs to the M . Indeed, if $M \in M$, then by definition $M \notin M$! Similarly, if $M \notin M$, then M should belong to M !! We leave these paradoxes to logisticians!

1.2 Definitions

All those definitions relate to sets A and B in a domain D .

Definition 1: Subset

A set A is a subset of a set B if and only if every element of A is also an element of B . We write $A \subset B$.

Notes:

- a) Using quantifiers, $\forall x \in A, x \in B$
- b) A set is a subset of itself: $A \subset A$
- c) There is one special subset, the empty set, noted \emptyset that belongs to all sets: *forall* $A \in D, \emptyset \subset A$.

Definition 2: Union

The union of two sets A and B , denoted $A \cup B$, is the set that contains those elements of the domain D that are either in A , or in B , or in both.

Notes:

- Using the set builder notation, $A \cup B = \{x \in D | x \in A \vee x \in B\}$

Definition 3: Intersection

The intersection of two sets A and B , denoted $A \cap B$, is the set that contains those elements of the domain D that are both in A and in B .

Notes:

- Using the set builder notation, $A \cap B = \{x \in D | x \in A \wedge x \in B\}$

Definition 4: Complement

The complement of a set A , denoted \bar{A} , is the set that contains those elements of the domain D that are not in A .

Notes:

- Using the set builder notation, $\bar{A} = \{x \in D | x \notin A\}$

Definition 5: Difference

The difference of two sets A and B , denoted $A \setminus B$ and sometimes $A - B$, is the set that contains those elements of the domain D that are in A but in B .

Notes:

- Using the set builder notation, $A \setminus B = \{x \in D | x \in A \wedge x \notin B\}$

Definition 6: Symmetric difference

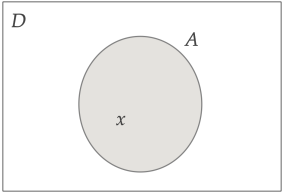
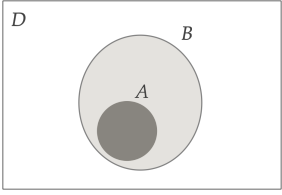
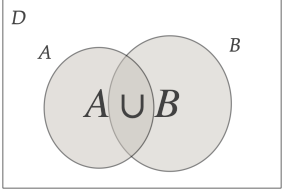
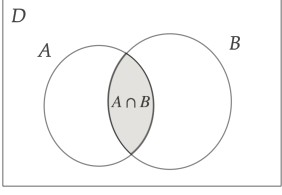
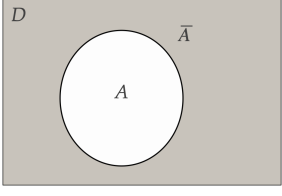
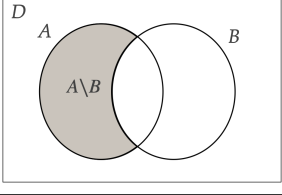
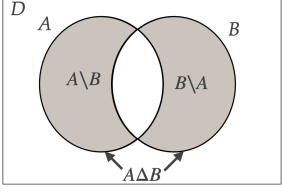
The symmetric difference of two sets A and B , denoted $A \Delta B$ is the set that contains those elements of the domain D that are either in A or in B but not in both.

Notes:

- Using the set builder notation, $A \Delta B = \{x \in D | x \in A \oplus x \in B\}$
- We can easily show that $A \Delta B = (A \setminus B) \cup (B \setminus A)$.

1.3 Summary

Let D be a domain. Let A and B be sets in D , and let x be an element of A .

Symbol	Meaning	Example	Reads as	Venn diagram
\in	element of	$x \in A$	x is an element of A	
\subset	subset	$A \subset B$	A is a subset of B	
\cup	union	$A \cup B$	A union B	
\cap	intersection	$A \cap B$	A intersects B	
	Complement	\bar{A}	Complement of A	
\setminus	Difference	$A \setminus B$	Difference of A and B	
Δ	Symmetric Difference	$A \Delta B$	Symm. Diff. of A and B	

1.4 Cardinality

1.4.1 Definition

If a set S is finite, the cardinality of S , denoted $|S|$, is the number of elements of S .

Notes:

- $|\emptyset| = 0$

1.4.2 Addition principle

If A and B are two sets in a domain D that are disjoint, i.e. $A \cap B = \emptyset$, then

$$|A \cup B| = |A| + |B|$$

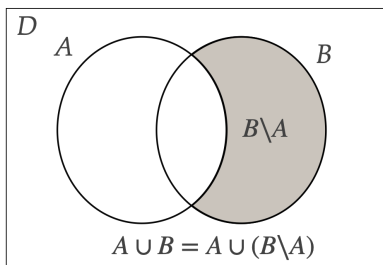
Proof: The proof is intuitive. We will assume it here.

1.4.3 Inclusion-exclusion principle

Let A and B be two sets in a domain D , then

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Proof: We provide the main elements:

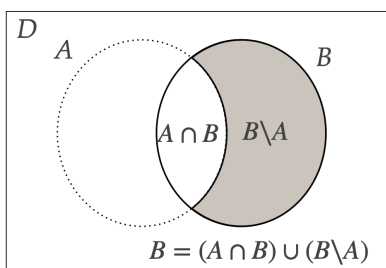


Note that

$$\begin{aligned} A \cap (B \setminus A) &= \{x \in D \mid x \in A \wedge (x \in (B \setminus A))\} \\ &= \{x \in D \mid x \in A \wedge (x \in (B \wedge x \notin A))\} \\ &= \emptyset \end{aligned}$$

Based on the addition principle, we have

$$|A \cup B| = |A| + |B \setminus A| \tag{1}$$



Note that

$$\begin{aligned}
 B &= (A \cap B) \cup (B \setminus A) \\
 &= \{x \in D \mid (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \in B)\} \\
 &= \emptyset
 \end{aligned}$$

Based on the addition principle, we have

$$|B| = |B \setminus A| + |A \cap B| \quad (2)$$

Combining equations 1 and 2, we get:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

which concludes the proof.

1.4.4 Some interesting properties of cardinality

a) **Inclusion-exclusion with 3 sets:**

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Let us define $D = B \cup C$, and let us apply the inclusion exclusion principle on $A \cup D$:

$$\begin{aligned}
 |A \cup D| &= |A| + |D| - |A \cap D| \\
 &= |A| + |B \cup C| - |A \cap (B \cup C)|
 \end{aligned} \quad (3)$$

Note that:

- Inclusion exclusion principle on $B \cup C$:
 $|B \cup C| = |B| + |C| - |B \cap C|$
- Distributivity:
 $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- Inclusion exclusion principle on $(A \cap B) \cup (A \cap C)$:
 $|A \cap (B \cup C)| = |A \cap B| + |A \cap C| - |A \cap B \cap C|$

Replacing in equation 3, we get:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

which concludes the proof.

b) **Cardinality and complement:** $|\bar{A}| = |D| - |A|$

We note that:

- Property of complement: $D = A \cup \bar{A}$

$$- A \cap \bar{A} = \{x \in D | x \in A \wedge x \notin A\} = \emptyset$$

Applying the addition principle:

$$|D| = |A| + |\bar{A}|$$

which concludes the proof.

1.5 Cartesian product

1.5.1 Definition

Let A and B be sets in possibly different domains D_A and D_B . The cartesian product of A and B , denoted $A \times B$ is the set of all ordered pairs (a, b) where a belongs to A and b belongs to B .

Notes:

- Set notation:

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

- Cartesian product and cardinality:

$$|A \times B| = |A| |B|$$

2 Basic counting principles

2.1 The product rule

2.1.1 Definition

Let us suppose that we want to explore the number of possibilities for a list L whose elements are pairs of 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 $n_1 n_2$ ways to build the list.

2.1.2 Set theory representation

The list L is build from ordered pairs of values, where the first value belongs to a set A and the second value belongs to a set B . Then $L = A \times B$, and

$$|L| = |A| \times |B|$$

Note that this generalizes to lists of multiple values. If $L = A_1 \times A_2 \times \dots \times A_N$, where N is a natural number representing the number of values, then

$$|L| = |A_1| \times |A_2| \times \dots \times |A_N|$$

2.1.3 Examples

- a) **Number of passwords:** Consider a password with 4 “digits”, where each digit can be a number from 0 to 9, or a letter from a to z (lower case). How many such passwords are there?

Let A be the set $\{0, 1, \dots, 9, a, b, \dots, z\}$. Note that $|A| = 36$. A password is a list of 4 values, with each value taken from A . Therefore, a password is an element of $A \times A \times A \times A = A^4$. The total number of possible passwords is therefore $N = |A|^4 = 36^4 = 1,679,616$.

- a) **Number of bits of a given size:** What is the number N_N of bit strings of size N ?

A bit string is a list of bits, where each bit belongs to the set $S = \{0, 1\}$. Note that $|S| = 2$. A bit string of length N then belongs to S^N , and the number of such bit string is $N_N = |S|^N = 2^N$.

2.2 The sum rule

2.2.1 Definition

If the list L can be split into two **disjoint** lists of size n_1 and n_2 , respectively, then there are $n_1 + n_2$ ways to build the list.

2.2.2 Set theory representation

The list L is build from the union of a set A and of a set B , with $A \cap B = \emptyset$. Then $L = A \cup B$, and

$$|L| = |A| + |B|$$

Note that this generalizes to multiple sublists. If $L = A_1 \cup A_2 \cup \dots \cup A_N$, where N is a natural number representing the number of sublists, and those subsets are disjoint, i.e. $A_i \cap A_j = \emptyset, \forall (i, j) \in [1, N] \times [1, N]$, then

$$|L| = |A_1| + |A_2| + \dots + |A_N|$$

2.2.3 Example

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

The list of desserts can be split into 3 disjoint lists of size 12 (fruits), 6 (cakes), and 4 (candies), respectively. Therefore, the total number of options is 22.

2.3 The extended sum rule

2.3.1 Definition

If the list L can be split into two lists of size n_1 and n_2 , respectively, and those two lists have n common elements, then there are $n_1 + n_2 - n$ ways to build the list.

2.3.2 Set theory representation

The list L is built from the union of a set A and of a set B . Then $L = A \cup B$, and

$$|L| = |A| + |B| - |A \cap B|$$

Note that this generalizes to multiple sublists, but the formula becomes then more complicated. We had seen for example that for three sublists A_1 , A_2 , and A_3 , $L = A_1 \cup A_2 \cup A_3$, and

$$|L| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|$$

2.4 The complement rule

2.4.1 Definition

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 $n - m$.

2.4.2 Set theory representation

The list L represents the domain. The object of that domain that possesses a property P represents a set A in L , and those that do not have the property P belong to the complement \bar{A} of A in L . In addition,

$$|\bar{A}| = |L| - |A|$$

2.5 Examples using one, or a combination of the rules defined above

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, or ending with 00, respectively. We note that L_1 and L_2 are not disjoint, and we call L_{12} there intersection. We use the product rule to compute the number of elements in L_1 , L_2 , and L_{12} :

- i) L_1 contains bit strings with 8 bits, but with the first bit set to 1. Therefore $|L_1| = 2^7$.
- i) L_2 contains bit strings with 8 bits, but with the last two bits set to 0. Therefore $|L_2| = 2^6$.
- i) L_3 contains bit strings with 8 bits, but with the first bit set to 1 and the last two bits set to 0. Therefore $|L_{12}| = 2^5$.

Based on the extended sum rule,

$$|L| = |L_1| + |L_2| - |L_{12}| = 2^7 + 2^6 - 2^5 = 160$$

b) *How many positive integer strictly less than 1,000 consist of digits from {1, 3, 7, 9}?*

Let us call L the list of such integers. L can be divided into 3 sublists, L_1 , L_2 , and L_3 , corresponding to integers with 1, 2, and 3 digits, respectively. We use the product rule to find the cardinality of those sublists:

- i) L_1 contains integers with 1 digit, with this digit being 1, 3, 7, or 9. Therefore $|L_1| = 4$.
- i) L_2 contains integers with 2 digits, with each digit being 1, 3, 7, or 9. Therefore $|L_2| = 4^2$.
- i) L_3 contains integers with 3 digits, with each digit being 1, 3, 7, or 9. Therefore $|L_3| = 4^3$.

Based on the sum rule,

$$|L| = |L_1| + |L_2| + |L_3| = 4 + 16 + 64 = 84$$

- c) ***How many positive integer strictly less than 1,000 consist of distinct digits from {1, 3, 7, 9}?***

Let us call L the list of such integers. L can be divided into 3 sublists, L_1 , L_2 , and L_3 , corresponding to integers with 1, 2, and 3 digits, respectively. We use the product rule to find the cardinality of those sublists:

- i) L_1 contains integers with 1 digit, with this digit being 1, 3, 7, or 9. Therefore $|L_1| = 4$.
- i) L_2 contains integers with 2 digits, with each digit being 1, 3, 7, or 9, but distinct from each other. Therefore $|L_2| = 4 \times 3 = 12$.
- i) L_3 contains integers with 3 digits, with each digit being 1, 3, 7, or 9, but distinct from each other. Therefore $|L_3| = 4 \times 3 \times 2 = 24$.

Based on the sum rule,

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

- 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 at least one digit***

Let us call L the list of such password. L can be divided into 3 sublists, L_6 , L_7 , and L_8 , corresponding to passwords of length 6, 7, and 8, respectively. Let us find the cardinality of those sublists.

Let D_6 be the set of passwords of length 6 that may, or may not contain a digit. By the product rule, $|D_6| = 36^6$. $\overline{L_6}$ is the subset of D_6 that contains passwords who do not contain a digit. $|\overline{L_6}| = 26^6$. Based on the complement rule,

$$|L_6| = |D_6| - |\overline{L_6}| = 36^6 - 26^6$$

Using similar reasonings,

$$|L_7| = |D_7| - |\overline{L_7}| = 36^7 - 26^7$$

$$|L_8| = |D_8| - |\overline{L_8}| = 36^8 - 26^8$$

Finally, using the sum rule,

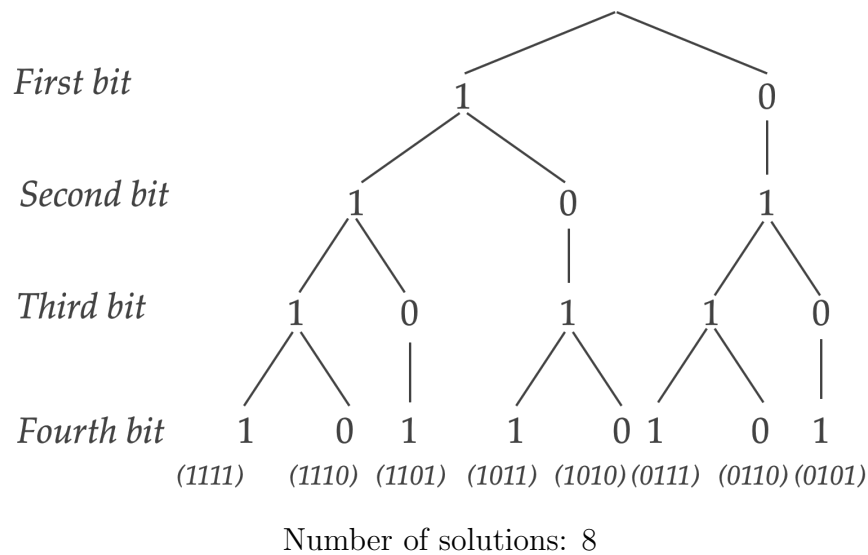
$$|L| = |L_6| + |L_7| + |L_8| = 36^8 + 36^7 + 36^6 - 26^8 - 26^7 - 26^6 = 2,684,483,063,360.$$

2.6 Tree diagrams

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 points 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, based on a tree:



2.7 The Pigeonhole principle

(also known as the Dirichlet's drawer principle)

Theorem

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

Proof:

We use a proof by contradiction. Suppose that each box contains at most one object. Then according to the sum rule, the total number of objects would be at most k , in contradiction with the fact that there are $k + 1$ objects at least.

Examples:

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

This is a direct application of the pigeonhole principle, with 13 “objects (the people), and 12 boxes (the 12 months of the year).

- b) For any 11 positive integers, some pair of them will have a difference divisible by 10.

Let us place each integer n into a box B_i , where i is the last digit of n (i.e. if $n = 121$, it is placed in the box B_1 , if $n = 237$, it is placed in the box B_7 , ...). There are 10 such boxes. According to the pigeonhole principle, at least 2 of the integers fall in the same box. Let n_1 and n_2 be two of those integers. Those two numbers end with the same last digit, and their difference has then its last digit equal to 0, i.e. it is divisible by 10.

2.8 The generalized Pigeonhole principle

Theorem

If N objects are placed into k boxes, then there is at least one box that contains $\lceil \frac{N}{k} \rceil$ objects, where $\lceil \cdot \rceil$ is the ceiling function.

Examples:

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

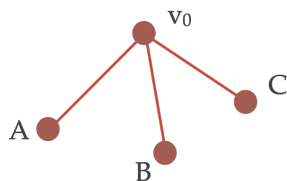
Let S be the set of pairs of integers among the first 10 integers whose sum is 11:

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

Note that $|S| = 5$. Each element of S can be thought of as a box. An integer smaller or equal to 10 is placed in one box based on its value. According to the generalized pigeonhole principle, among 7 integers, at least $\lceil \frac{7}{5} \rceil = 2$ fall in the same box. Their sum is necessarily 11. The 5 remaining integers are then placed into the four remaining boxes, and again 2 of them fall in the same box. Their sum is also 11.

- b) Consider 6 vertices, all connected by edges, such that those edges are either red or blue. Prove that at least 3 of those vertices form a triangle with all edges of the same color.

Consider one of the size vertices, say v_0 . It is connected to all 5 remaining vertices by edges. Those edges can be either red or blue. According to the generalized pigeonhole principle, at least three of those edges will have the same color, \mathcal{C} (where \mathcal{C} can be either red or blue). Let us name those vertices A , B , and C .



The vertex v_0 is connected to the vertices A , B , C , with 3 edges of the same color.

Let us consider now the edges AB , BC , and AC . If AB is of color \mathcal{C} , then the triangle v_0AB has its three edges colored in \mathcal{C} . If AC is of color \mathcal{C} , then the triangle v_0AC has its three edges

of color \mathcal{C} . If BC is of color \mathcal{C} , then the triangle v_0BC has its three edges of color \mathcal{C} . If none of those edges is of color \mathcal{C} , then they all three have the same color, and the triangle ABC satisfies the condition!

3 Permutations - Combinations

3.1 Permutations

Definition

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

Examples: Let $S = \{1, 2, 3\}$. the set $\{2, 3, 1\}$ is a permutation of S , The set $\{3, 1\}$ is a 2-permutation of S .

Theorem

The number of R permutation of a set with n distinct elements is:

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

This theorem is a direct application of the product rule.

Notes:

- $P(n, n)$ is written as $n!$ (factorial n , or n factorial), i.e.

$$n! = n(n - 1) \dots 1$$

- By definition $0! = 1$
- We have:

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

Examples of applications:

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

Call the persons $A, B, C, D,$ and E . An arrangement of these five persons in a line is a permutation of the set $S = \{A, B, C, D, E\}$, with $|S| = 5$. There are $P(5, 5) = 5! = 120$ such arrangement.

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

Same as in a): $P(5, 5) = 5! = 120$.

c) How many of the arrangements of the letters in the word *MATCH* have the letter *M*, *A*, side by side?

We define the new set $S = \{MA, T, C, H\}$. The number of arrangements of the letters in the word *MATCH* that have the letter *M*, *A*, side by side is the number of permutations of this set, namely $P(4, 4) = 4! = 24$.

3.2 Combinations

Let us 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 , as r -permutations are ordered arrangements, while sets and subsets are not ordered. For example, consider the set $S = \{a, b, c\}$. There are 6 2-permutations of S , $\{a, b\}$, $\{b, a\}$, $\{a, c\}$, $\{c, a\}$, $\{b, c\}$, $\{c, b\}$, but only 3 subsets of 2 elements in S , $\{a, b\}$, $\{a, c\}$, $\{b, c\}$.

Definition

Subsets of size r from a set S are called r -combinations of S . We write $C(n, r)$, or $\binom{n}{r}$, and 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 n elements, 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 of applications:

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?

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

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

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

d) How many bit strings of length n contain exactly r 1s?

The position of the R 1s in the bit string of length n forms a r -combination of $\{1, 2, \dots, n\}$. Hence there are $C(n, r)$ such bit strings of length n with exactly r 1s.

e) A club has 25 members:

– How many ways are there to choose 3 members of the club to serve on a committee?

$$C(25, 3)$$

– How many ways are there to choose a president, a vice president, and a treasurer for the club?

$$P(25, 3)$$

e) A club a 10 women and 8 men is forming a committee of 5 persons. Clearly, there are $C(18, 5)$ such committees. How many of those

– contain exactly 3 women:

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

– contain 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 integers. Then

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

Properties of combinations

Let n and r be non-negative integers with $r \leq n$. Then

$$\begin{aligned}C(n, r) &= C(n, n - r) \\ \sum_{i=0}^n C(n, i) &= 2^n \\ \sum_{i=0}^n C(n, i)(-1)^i &= 0\end{aligned}$$