

# Inclusion–exclusion principle

In combinatorics, a branch of mathematics, the **inclusion–exclusion principle** is a counting technique which generalizes the familiar method of obtaining the number of elements in the union of two finite sets; symbolically expressed as

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

where *A* and *B* are two finite sets and *|S|* indicates the cardinality of a set *S* (which may be considered as the number of elements of the set, if the set is finite). The formula expresses the fact that the sum of the sizes of the two sets may be too large since some elements may be counted twice. The double-counted elements are those in the intersection of the two sets and the count is corrected by subtracting the size of the intersection.

The inclusion-exclusion principle, being a generalization of the two-set case, is perhaps more clearly seen in the case of three sets, which for the sets *A*, *B* and *C* is given by

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

This formula can be verified by counting how many times each region in the Venn diagram figure is included in the right-hand side of the formula. In this case, when removing the contributions of over-counted elements, the number of elements in the mutual intersection of the three sets has been subtracted too often, so must be added back in to get the correct total.

Generalizing the results of these examples gives the principle of inclusion–exclusion. To find the cardinality of the union of *n* sets:

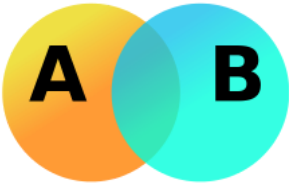
1. Include the cardinalities of the sets.
2. Exclude the cardinalities of the pairwise intersections.
3. Include the cardinalities of the triple-wise intersections.
4. Exclude the cardinalities of the quadruple-wise intersections.
5. Include the cardinalities of the quintuple-wise intersections.
6. Continue, until the cardinality of the *n*-tuple-wise intersection is included (if *n* is odd) or excluded (*n* even).

The name comes from the idea that the principle is based on over-generous *inclusion*, followed by compensating *exclusion*. This concept is attributed to Abraham de Moivre (1718),<sup>[1]</sup> although it first appears in a paper of Daniel da Silva (1854)<sup>[2]</sup> and later in a paper by J. J. Sylvester (1883).<sup>[3]</sup> Sometimes the principle is referred to as the formula of Da Silva or Sylvester, due to these publications. The principle can be viewed as an example of the sieve method extensively used in number theory and is sometimes referred to as the *sieve formula*.<sup>[4]</sup>

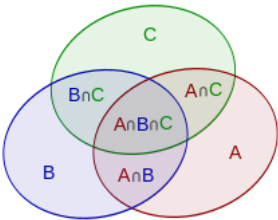
As finite probabilities are computed as counts relative to the cardinality of the probability space, the formulas for the principle of inclusion–exclusion remain valid when the cardinalities of the sets are replaced by finite probabilities. More generally, both versions of the principle can be put under the common umbrella of measure theory.

In a very abstract setting, the principle of inclusion–exclusion can be expressed as the calculation of the inverse of a certain matrix.<sup>[5]</sup> This inverse has a special structure, making the principle an extremely valuable technique in combinatorics and related areas of mathematics. As Gian-Carlo Rota put it:<sup>[6]</sup>

"One of the most useful principles of enumeration in discrete probability and combinatorial theory is the celebrated principle of inclusion–exclusion. When skillfully applied, this principle has yielded the solution to many a combinatorial problem."



Venn diagram showing the union of sets *A* and *B* as everything not in white



Inclusion–exclusion illustrated by a Venn diagram for three sets

## Contents

### Statement

### Examples

- Counting integers
- Counting derangements

### A special case

### A generalization

### In probability

- Special case

### Other forms

### Applications

- Counting derangements
- Counting intersections
- Graph coloring
- Bipartite graph perfect matchings
- Number of onto functions
- Permutations with forbidden positions
- Stirling numbers of the second kind
- Rook polynomials
- Euler's phi function

## Diluted inclusion–exclusion principle

### Proof of main statement

#### Algebraic proof

### See also

### Notes

### References

## Statement

In its general form, the principle of inclusion–exclusion states that for finite sets  $A_1, \dots, A_n$ , one has the identity

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| - \dots + (-1)^{n+1} |A_1 \cap \dots \cap A_n|.$$

This can be compactly written as

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{k=1}^n (-1)^{k+1} \left( \sum_{1 \leq i_1 < \dots < i_k \leq n} |A_{i_1} \cap \dots \cap A_{i_k}| \right)$$

or

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{\emptyset \neq J \subseteq \{1, \dots, n\}} (-1)^{|J|+1} \left| \bigcap_{j \in J} A_j \right|.$$

In words, to count the number of elements in a finite union of finite sets, first sum the cardinalities of the individual sets, then subtract the number of elements that appear in at least two sets, then add back the number of elements that appear in at least three sets, then subtract the number of elements that appear in at least four sets, and so on. This process always ends since there can be no elements that appear in more than the number of sets in the union. (For example, if  $n = 4$ , there can be no elements that appear in more than 4 sets; equivalently, there can be no elements that appear in at least 5 sets.)

In applications it is common to see the principle expressed in its complementary form. That is, letting  $S$  be a finite universal set containing all of the  $A_i$  and letting  $\bar{A}_i$  denote the complement of  $A_i$  in  $S$ , by De Morgan's laws we have

$$\left| \bigcap_{i=1}^n \bar{A}_i \right| = \left| S - \bigcup_{i=1}^n A_i \right| = |S| - \sum_{i=1}^n |A_i| + \sum_{1 \leq i < j \leq n} |A_i \cap A_j| - \dots + (-1)^n |A_1 \cap \dots \cap A_n|.$$

As another variant of the statement, let  $P_1, \dots, P_n$  be a list of properties that elements of a set  $S$  may or may not have, then the principle of inclusion–exclusion provides a way to calculate the number of elements of  $S$  which have none of the properties. Just let  $A_i$  be the subset of elements of  $S$  which have the property  $P_i$  and use the principle in its complementary form. This variant is due to J. J. Sylvester.<sup>[1]</sup>

Notice that if you take into account only the first  $m < n$  sums on the right (in the general form of the principle), then you will get an overestimate if  $m$  is odd and an underestimate if  $m$  is even.

## Examples

### Counting integers

As a simple example of the use of the principle of inclusion–exclusion, consider the question:<sup>[7]</sup>

How many integers in  $\{1, \dots, 100\}$  are not divisible by 2, 3 or 5?

Let  $S = \{1, \dots, 100\}$  and  $P_1$  the property that an integer is divisible by 2,  $P_2$  the property that an integer is divisible by 3 and  $P_3$  the property that an integer is divisible by 5. Letting  $A_i$  be the subset of  $S$  whose elements have property  $P_i$  we have by elementary counting:  $|A_1| = 50$ ,  $|A_2| = 33$ , and  $|A_3| = 20$ . There are 16 of these integers divisible by 6, 10 divisible by 10, and 6 divisible by 15. Finally, there are just 3 integers divisible by 30, so the number of integers not divisible by any of 2, 3 or 5 is given by:

$$100 - (50 + 33 + 20) + (16 + 10 + 6) - 3 = 26.$$

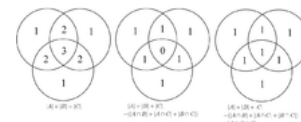
### Counting derangements

A more complex example is the following.

Suppose there is a deck of  $n$  cards numbered from 1 to  $n$ . Suppose a card numbered  $m$  is in the correct position if it is the  $m$ th card in the deck. How many ways,  $W$ , can the cards be shuffled with at least 1 card being in the correct position?

Begin by defining set  $A_m$ , which is all of the orderings of cards with the  $m$ th card correct. Then the number of orders,  $W$ , with *at least* one card being in the correct position,  $m$ , is

$$W = \left| \bigcup_{m=1}^n A_m \right|.$$



Each term of the inclusion–exclusion formula gradually corrects the count until finally each portion of the Venn diagram is counted exactly once.

Apply the principle of inclusion–exclusion,

$$W = \sum_{m_1=1}^n |A_{m_1}| - \sum_{1 \leq m_1 < m_2 \leq n} |A_{m_1} \cap A_{m_2}| + \cdots + (-1)^{p-1} \sum_{1 \leq m_1 < \cdots < m_p \leq n} |A_{m_1} \cap \cdots \cap A_{m_p}| + \cdots$$

Each value  $A_{m_1} \cap \cdots \cap A_{m_p}$  represents the set of shuffles having at least  $p$  values  $m_1, \dots, m_p$  in the correct position. Note that the number of shuffles with at least  $p$  values correct only depends on  $p$ , not on the particular values of  $\mathbf{m}$ . For example, the number of shuffles having the 1st, 3rd, and 17th cards in the correct position is the same as the number of shuffles having the 2nd, 5th, and 13th cards in the correct positions. It only matters that of the  $n$  cards, 3 were chosen to be in the correct position. Thus there are  $\binom{n}{p}$  equal terms in the  $p$ th summation (see [combination](#)).

$$W = \binom{n}{1}|A_1| - \binom{n}{2}|A_1 \cap A_2| + \cdots + (-1)^{p-1} \binom{n}{p}|A_1 \cap \cdots \cap A_p| + \cdots$$

$|A_1 \cap \cdots \cap A_p|$  is the number of orderings having  $p$  elements in the correct position, which is equal to the number of ways of ordering the remaining  $n - p$  elements, or  $(n - p)!$ . Thus we finally get:

$$\begin{aligned} W &= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \cdots + (-1)^{p-1} \binom{n}{p}(n-p)! + \cdots \\ &= \sum_{p=1}^n (-1)^{p-1} \binom{n}{p} (n-p)! \\ &= \sum_{p=1}^n (-1)^{p-1} \frac{n!}{p!(n-p)!} (n-p)! \\ &= \sum_{p=1}^n (-1)^{p-1} \frac{n!}{p!} \end{aligned}$$

A permutation where *no* card is in the correct position is called a [derangement](#). Taking  $n!$  to be the total number of permutations, the probability  $Q$  that a random shuffle produces a derangement is given by

$$Q = 1 - \frac{W}{n!} = \sum_{p=0}^n \frac{(-1)^p}{p!},$$

a truncation to  $n + 1$  terms of the [Taylor expansion](#) of  $e^{-1}$ . Thus the probability of guessing an order for a shuffled deck of cards and being incorrect about every card is approximately  $e^{-1}$  or 37%.

## A special case

The situation that appears in the derangement example above occurs often enough to merit special attention.<sup>[8]</sup> Namely, when the size of the intersection sets appearing in the formulas for the principle of inclusion–exclusion depend only on the number of sets in the intersections and not on which sets appear. More formally, if the intersection

$$A_J := \bigcap_{j \in J} A_j$$

has the same cardinality, say  $\alpha_k = |A_J|$ , for every  $k$ -element subset  $J$  of  $\{1, \dots, n\}$ , then

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} \alpha_k.$$

Or, in the complementary form, where the universal set  $S$  has cardinality  $\alpha_0$ ,

$$\left| S \setminus \bigcup_{i=1}^n A_i \right| = \sum_{k=0}^n (-1)^k \binom{n}{k} \alpha_k.$$

## A generalization

Given a [family](#) (repeats allowed) of subsets  $A_1, A_2, \dots, A_n$  of a universal set  $S$ , the principle of inclusion–exclusion calculates the number of elements of  $S$  in none of these subsets. A generalization of this concept would calculate the number of elements of  $S$  which appear in exactly some fixed  $m$  of these sets.

Let  $N = [n] = \{1, 2, \dots, n\}$ . If we define  $A_\emptyset = S$ , then the principle of inclusion–exclusion can be written as, using the notation of the previous section; the number of elements of  $S$  contained in none of the  $A_i$  is:

$$\sum_{J \subseteq [n]} (-1)^{|J|} |A_J|.$$

If  $I$  is a fixed subset of the index set  $N$ , then the number of elements which belong to  $A_i$  for all  $i$  in  $I$  and for no other values is:<sup>[9]</sup>

$$\sum_{I \subseteq J} (-1)^{|J|-|I|} |A_J|.$$

Define the sets

$$B_k = A_{I \cup \{k\}} \text{ for } k \in N \setminus I.$$

We seek the number of elements in none of the  $B_k$  which, by the principle of inclusion–exclusion (with  $B_\emptyset = A_I$ ), is

$$\sum_{K \subseteq N \setminus I} (-1)^{|K|} |B_K|.$$

The correspondence  $K \leftrightarrow J = I \cup K$  between subsets of  $N \setminus I$  and subsets of  $N$  containing  $I$  is a bijection and if  $J$  and  $K$  correspond under this map then  $B_K = A_J$ , showing that the result is valid.

## In probability

In probability, for events  $A_1, \dots, A_n$  in a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , the inclusion–exclusion principle becomes for  $n = 2$

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2),$$

for  $n = 3$

$$\mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3)$$

and in general

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i < j} \mathbb{P}(A_i \cap A_j) + \sum_{i < j < k} \mathbb{P}(A_i \cap A_j \cap A_k) + \dots + (-1)^{n-1} \sum_{i < \dots < n} \mathbb{P}\left(\bigcap_{i=1}^n A_i\right),$$

which can be written in closed form as

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n \left( (-1)^{k-1} \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \mathbb{P}(A_I) \right),$$

where the last sum runs over all subsets  $I$  of the indices  $1, \dots, n$  which contain exactly  $k$  elements, and

$$A_I := \bigcap_{i \in I} A_i$$

denotes the intersection of all those  $A_i$  with index in  $I$ .

According to the Bonferroni inequalities, the sum of the first terms in the formula is alternately an upper bound and a lower bound for the LHS. This can be used in cases where the full formula is too cumbersome.

For a general measure space  $(S, \Sigma, \mu)$  and measurable subsets  $A_1, \dots, A_n$  of finite measure, the above identities also hold when the probability measure  $\mathbb{P}$  is replaced by the measure  $\mu$ .

## Special case

If, in the probabilistic version of the inclusion–exclusion principle, the probability of the intersection  $A_I$  only depends on the cardinality of  $I$ , meaning that for every  $k$  in  $\{1, \dots, n\}$  there is an  $a_k$  such that

$$a_k = \mathbb{P}(A_I) \text{ for every } I \subset \{1, \dots, n\} \text{ with } |I| = k,$$

then the above formula simplifies to

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} a_k$$

due to the combinatorial interpretation of the binomial coefficient  $\binom{n}{k}$ . For example, if the events  $A_i$  are independent and identically distributed, then  $\mathbb{P}(A_i) = p$  for all  $i$ , and we have  $a_k = p^k$ , in which case the expression above simplifies to

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = 1 - (1 - p)^n.$$

(This result can also be derived more simply by considering the intersection of the complements of the events  $A_i$ .)

An analogous simplification is possible in the case of a general measure space  $(S, \Sigma, \mu)$  and measurable subsets  $A_1, \dots, A_n$  of finite measure.

## Other forms

The principle is sometimes stated in the form<sup>[10]</sup> that says that if

$$g(A) = \sum_{S \subseteq A} f(S)$$



The first occurrence of the problem of counting the number of derangements is in an early book on games of chance: *Essai d'analyse sur les jeux de hazard* by P. R. de Montmort (1678 – 1719) and was known as either "Montmort's problem" or by the name he gave it, "*problème des rencontres*."<sup>[11]</sup> The problem is also known as the *hatcheck problem*.

The number of derangements is also known as the subfactorial of  $n$ , written  $!n$ . It follows that if all bijections are assigned the same probability then the probability that a random bijection is a derangement quickly approaches  $1/e$  as  $n$  grows.

## Counting intersections

The principle of inclusion–exclusion, combined with De Morgan's law, can be used to count the cardinality of the intersection of sets as well. Let  $\overline{A_k}$  represent the complement of  $A_k$  with respect to some universal set  $A$  such that  $A_k \subseteq A$  for each  $k$ . Then we have

$$\bigcap_{i=1}^n A_i = \overline{\bigcup_{i=1}^n \overline{A_i}}$$

thereby turning the problem of finding an intersection into the problem of finding a union.

## Graph coloring

The inclusion exclusion principle forms the basis of algorithms for a number of NP-hard graph partitioning problems, such as graph coloring.<sup>[12]</sup>

A well known application of the principle is the construction of the chromatic polynomial of a graph.<sup>[13]</sup>

## Bipartite graph perfect matchings

The number of perfect matchings of a bipartite graph can be calculated using the principle.<sup>[14]</sup>

## Number of onto functions

Given finite sets  $A$  and  $B$ , how many surjective functions (onto functions) are there from  $A$  to  $B$ ? Without any loss of generality we may take  $A = \{1, \dots, k\}$  and  $B = \{1, \dots, n\}$ , since only the cardinalities of the sets matter. By using  $S$  as the set of all functions from  $A$  to  $B$ , and defining, for each  $i$  in  $B$ , the property  $P_i$  as "the function misses the element  $i$  in  $B$ " ( $i$  is not in the image of the function), the principle of inclusion–exclusion gives the number of onto functions between  $A$  and  $B$  as:<sup>[15]</sup>

$$\sum_{j=0}^n \binom{n}{j} (-1)^j (n-j)^k.$$

## Permutations with forbidden positions

A permutation of the set  $S = \{1, \dots, n\}$  where each element of  $S$  is restricted to not being in certain positions (here the permutation is considered as an ordering of the elements of  $S$ ) is called a *permutation with forbidden positions*. For example, with  $S = \{1, 2, 3, 4\}$ , the permutations with the restriction that the element 1 can not be in positions 1 or 3, and the element 2 can not be in position 4 are: 2134, 2143, 3124, 4123, 2341, 2431, 3241, 3421, 4231 and 4321. By letting  $A_i$  be the set of positions that the element  $i$  is not allowed to be in, and the property  $P_i$  to be the property that a permutation puts element  $i$  into a position in  $A_i$ , the principle of inclusion–exclusion can be used to count the number of permutations which satisfy all the restrictions.<sup>[16]</sup>

In the given example, there are  $12 = 2(3!)$  permutations with property  $P_1$ ,  $6 = 3!$  permutations with property  $P_2$  and no permutations have properties  $P_3$  or  $P_4$  as there are no restrictions for these two elements. The number of permutations satisfying the restrictions is thus:

$$4! - (12 + 6 + 0 + 0) + (4) = 24 - 18 + 4 = 10.$$

The final 4 in this computation is the number of permutations having both properties  $P_1$  and  $P_2$ . There are no other non-zero contributions to the formula.

## Stirling numbers of the second kind

The Stirling numbers of the second kind,  $S(n, k)$  count the number of partitions of a set of  $n$  elements into  $k$  non-empty subsets (indistinguishable boxes). An explicit formula for them can be obtained by applying the principle of inclusion–exclusion to a very closely related problem, namely, counting the number of partitions of an  $n$ -set into  $k$  non-empty but distinguishable boxes (ordered non-empty subsets). Using the universal set consisting of all partitions of the  $n$ -set into  $k$  (possibly empty) distinguishable boxes,  $A_1, A_2, \dots, A_k$ , and the properties  $P_i$  meaning that the partition has box  $A_i$  empty, the principle of inclusion–exclusion gives an answer for the related result. Dividing by  $k!$  to remove the artificial ordering gives the Stirling number of the second kind:<sup>[17]</sup>

$$S(n, k) = \frac{1}{k!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n.$$

## Rook polynomials

A rook polynomial is the generating function of the number of ways to place non-attacking rooks on a *board*  $B$  that looks like a subset of the squares of a checkerboard; that is, no two rooks may be in the same row or column. The board  $B$  is any subset of the squares of a rectangular board with  $n$  rows and  $m$  columns; we think of it as the squares in which one is allowed to put a rook. The coefficient,  $r_k(B)$  of  $x^k$  in the rook polynomial  $R_B(x)$  is the number of ways  $k$

rooks, none of which attacks another, can be arranged in the squares of  $B$ . For any board  $B$ , there is a complementary board  $B'$  consisting of the squares of the rectangular board that are not in  $B$ . This complementary board also has a rook polynomial  $R_{B'}(x)$  with coefficients  $r_k(B')$ .

It is sometimes convenient to be able to calculate the highest coefficient of a rook polynomial in terms of the coefficients of the rook polynomial of the complementary board. Without loss of generality we can assume that  $n \leq m$ , so this coefficient is  $r_n(B)$ . The number of ways to place  $n$  non-attacking rooks on the complete  $n \times m$  "checkerboard" (without regard as to whether the rooks are placed in the squares of the board  $B$ ) is given by the falling factorial:

$$(m)_n = m(m-1)(m-2) \cdots (m-n+1).$$

Letting  $P_i$  be the property that an assignment of  $n$  non-attacking rooks on the complete board has a rook in column  $i$  which is not in a square of the board  $B$ , then by the principle of inclusion–exclusion we have:<sup>[18]</sup>

$$r_n(B) = \sum_{t=0}^n (-1)^t (m-t)_{n-t} r_t(B').$$

## Euler's phi function

Euler's totient or phi function,  $\varphi(n)$  is an arithmetic function that counts the number of positive integers less than or equal to  $n$  that are relatively prime to  $n$ . That is, if  $n$  is a positive integer, then  $\varphi(n)$  is the number of integers  $k$  in the range  $1 \leq k \leq n$  which have no common factor with  $n$  other than 1. The principle of inclusion–exclusion is used to obtain a formula for  $\varphi(n)$ . Let  $S$  be the set  $\{1, \dots, n\}$  and define the property  $P_i$  to be that a number in  $S$  is divisible by the prime number  $p_i$ , for  $1 \leq i \leq r$ , where the prime factorization of

$$n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}.$$

Then,<sup>[19]</sup>

$$\varphi(n) = n - \sum_{i=1}^r \frac{n}{p_i} + \sum_{1 \leq i < j \leq r} \frac{n}{p_i p_j} - \cdots = n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right).$$

## Diluted inclusion–exclusion principle

In many cases where the principle could give an exact formula (in particular, counting prime numbers using the sieve of Eratosthenes), the formula arising does not offer useful content because the number of terms in it is excessive. If each term individually can be estimated accurately, the accumulation of errors may imply that the inclusion–exclusion formula is not directly applicable. In number theory, this difficulty was addressed by Viggo Brun. After a slow start, his ideas were taken up by others, and a large variety of sieve methods developed. These for example may try to find upper bounds for the "sieved" sets, rather than an exact formula.

Let  $A_1, \dots, A_n$  be arbitrary sets and  $p_1, \dots, p_n$  real numbers in the closed unit interval  $[0, 1]$ . Then, for every even number  $k$  in  $\{0, \dots, n\}$ , the indicator functions satisfy the inequality:<sup>[20]</sup>

$$1_{A_1 \cup \dots \cup A_n} \geq \sum_{j=1}^k (-1)^{j-1} \sum_{1 \leq i_1 < \dots < i_j \leq n} p_{i_1} \cdots p_{i_j} 1_{A_{i_1} \cap \dots \cap A_{i_j}}.$$

## Proof of main statement

Choose an element contained in the union of all sets and let  $A_1, A_2, \dots, A_t$  be the individual sets containing it. (Note that  $t > 0$ .) Since the element is counted precisely once by the left-hand side of equation (1), we need to show that it is counted precisely once by the right-hand side. On the right-hand side, the only non-zero contributions occur when all the subsets in a particular term contain the chosen element, that is, all the subsets are selected from  $A_1, A_2, \dots, A_t$ . The contribution is one for each of these sets (plus or minus depending on the term) and therefore is just the (signed) number of these subsets used in the term. We then have:

$$|\{A_i \mid 1 \leq i \leq t\}| - |\{A_i \cap A_j \mid 1 \leq i < j \leq t\}| + \cdots + (-1)^{t+1} |\{A_1 \cap A_2 \cap \dots \cap A_t\}| = \binom{t}{1} - \binom{t}{2} + \cdots + (-1)^{t+1} \binom{t}{t}.$$

By the binomial theorem,

$$0 = (1-1)^t = \binom{t}{0} - \binom{t}{1} + \binom{t}{2} - \cdots + (-1)^t \binom{t}{t}.$$

Using the fact that  $\binom{t}{0} = 1$  and rearranging terms, we have

$$1 = \binom{t}{1} - \binom{t}{2} + \cdots + (-1)^{t+1} \binom{t}{t},$$

and so, the chosen element is counted only once by the right-hand side of equation (1).

## Algebraic proof

An algebraic proof can be obtained using indicator functions (also known as characteristic functions). The indicator function of a subset  $S$  of a set  $X$  is the function

$$\mathbf{1}_S : X \rightarrow \{0, 1\}$$

$$\mathbf{1}_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

If  $A$  and  $B$  are two subsets of  $X$ , then

$$\mathbf{1}_A \cdot \mathbf{1}_B = \mathbf{1}_{A \cap B}.$$

Let  $A$  denote the union  $\bigcup_{i=1}^n A_i$  of the sets  $A_1, \dots, A_n$ . To prove the inclusion–exclusion principle in general, we first verify the identity

$$\mathbf{1}_A = \sum_{k=1}^n (-1)^{k-1} \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \mathbf{1}_{A_I} \quad (*)$$

for indicator functions, where:

$$A_I = \bigcap_{i \in I} A_i.$$

The following function

$$(\mathbf{1}_A - \mathbf{1}_{A_1})(\mathbf{1}_A - \mathbf{1}_{A_2}) \cdots (\mathbf{1}_A - \mathbf{1}_{A_n}),$$

is identically zero because: if  $x$  is not in  $A$ , then all factors are  $0 - 0 = 0$ ; and otherwise, if  $x$  does belong to some  $A_m$ , then the corresponding  $m^{\text{th}}$  factor is  $1 - 1 = 0$ . By expanding the product on the left-hand side, equation  $(*)$  follows.

To prove the inclusion–exclusion principle for the cardinality of sets, sum the equation  $(*)$  over all  $x$  in the union of  $A_1, \dots, A_n$ . To derive the version used in probability, take the expectation in  $(*)$ . In general, integrate the equation  $(*)$  with respect to  $\mu$ . Always use linearity in these derivations.

## See also

- Boole's inequality
- Combinatorial principles
- Maximum-minimums identity – Relates the maximum element of a set of numbers and the minima of its non-empty subsets
- Necklace problem
- Pigeonhole principle – If there are more items than boxes holding them, one box must contain at least two items
- Schuette–Nesbitt formula

## Notes

- Roberts & Tesman 2009, pg. 405
- Mazur 2010, pg. 94
- van Lint & Wilson 1992, pg. 77
- van Lint & Wilson 1992, pg. 77
- Stanley 1986, pg. 64
- Rota, Gian-Carlo (1964), "On the foundations of combinatoial theory I. Theory of Möbius functions", *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **2** (4): 340–368, doi:10.1007/BF00531932 (https://doi.org/10.1007%2FBF00531932), S2CID 121334025 (http s://api.semanticscholar.org/CorpusID:121334025)
- Mazur 2010, pp. 83–4, 88
- Brualdi 2010, pp. 167–8
- Cameron 1994, pg. 78
- Graham, Grötschel & Lovász 1995, pg. 1049
- van Lint & Wilson 1992, pp. 77-8
- Björklund, Husfeldt & Koivisto 2009
- Gross 2008, pp. 211–13
- Gross 2008, pp. 208–10
- Mazur 2008, pp.84-5, 90
- Brualdi 2010, pp. 177–81
- Brualdi 2010, pp. 282–7
- Roberts & Tesman 2009, pp.419–20
- van Lint & Wilson 1992, pg. 73
- (Fernández, Fröhlich & Alan D. 1992, Proposition 12.6)

## References

- Allenby, R.B.J.T.; Slomson, Alan (2010), *How to Count: An Introduction to Combinatorics* (http://www.crcpress.com/product/isbn/9781420082609), Discrete Mathematics and Its Applications (2 ed.), CRC Press, pp. 51–60, ISBN 9781420082609
- Björklund, A.; Husfeldt, T.; Koivisto, M. (2009), "Set partitioning via inclusion–exclusion", *SIAM Journal on Computing*, **39** (2): 546–563, doi:10.1137/070683933 (https://doi.org/10.1137%2F070683933)
- Brualdi, Richard A. (2010), *Introductory Combinatorics* (5th ed.), Prentice–Hall, ISBN 9780136020400
- Cameron, Peter J. (1994), *Combinatorics: Topics, Techniques, Algorithms*, Cambridge University Press, ISBN 0-521-45761-0
- Fernández, Roberto; Fröhlich, Jürg; Alan D., Sokal (1992), *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory*, Texts an Monographs in Physics, Berlin: Springer-Verlag, pp. xviii+444, ISBN 3-540-54358-9, MR 1219313 (https://www.ams.org/mathscinet-getitem?mr=1219313), Zbl 0761.60061 (https://zbmath.org/?format=complete&q=an:0761.60061)
- Graham, R.L.; Grötschel, M.; Lovász, L. (1995), *Hand Book of Combinatorics (volume-2)*, MIT Press – North Holland, ISBN 9780262071710
- Gross, Jonathan L. (2008), *Combinatorial Methods with Computer Applications*, Chapman&Hall/CRC, ISBN 9781584887430
- "Inclusion-and-exclusion principle" (https://www.encyclopediaofmath.org/index.php?title=Inclusion-and-exclusion\_principle), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- Mazur, David R. (2010), *Combinatorics A Guided Tour*, The Mathematical Association of America, ISBN 9780883857625
- Roberts, Fred S.; Tesman, Barry (2009), *Applied Combinatorics* (2nd ed.), CRC Press, ISBN 9781420099829
- Stanley, Richard P. (1986), *Enumerative Combinatorics Volume I*, Wadsworth & Brooks/Cole, ISBN 0534065465
- van Lint, J.H.; Wilson, R.M. (1992), *A Course in Combinatorics*, Cambridge University Press, ISBN 0521422604



*This article incorporates material from principle of inclusion–exclusion on [PlanetMath](#), which is licensed under the [Creative Commons Attribution/Share-Alike License](#).*

---

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Inclusion–exclusion\\_principle&oldid=1102804782](https://en.wikipedia.org/w/index.php?title=Inclusion–exclusion_principle&oldid=1102804782)"

---

This page was last edited on 7 August 2022, at 00:26 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License 3.0; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.