Mathematical Foundations of Computer Science

This Lecture: Combinatorics - Principle of Inclusion and Exclusion

Department of Mathematical and Computational Sciences
National Institute of Technology Karnataka, Surathkal

MA714 (Odd Semester [2025-26])

Introduction

The Principle of Inclusion and Exclusion (PIE) is a fundamental combinatorial method used to count the number of elements in the union of several sets. It corrects the overcounting that occurs when simple addition of individual set sizes is used.

This principle is widely used in computer science, especially in areas such as database query optimization, counting solutions to constraint satisfaction problems, and analyzing algorithms.

Example 1

Let S represent a set of 100 students enrolled in the first year M.Tech. program at NITK. Then $|S| = 100$. *Let* c_1, c_2 *denote two properties satisfied by some of the students in S.*

c_1: *A student at NITK is among the 100 students in M.Tech. and is enrolled in Swimming Classes.*

c_2: *A student at NITK is among the 100 students in M.Tech. and is enrolled in Badminton Training.*

Suppose that 35 of the 100 students are enrolled in Swimming Classes and that 30 of them are enrolled in Badminton Training.

That is, $N(c_1) = 35$ *and* $N(c_2) = 30$.

If 9 of these are enroled in both Swimming Classes and Badminton Training, then we write $N(c_1c_2) = 9$.

Example 1 (contd.)

Using similar notations and $\overline{c}_i$ to denote "not c_i", we note:
$N(\overline{c}_1) = 100 - 35 = 65$ and similarly, $N(\overline{c}_2) = 70$.
Likewise, $N(c_1\overline{c}_2) = N(c_1) - N(c_1c_2) = 35 - 9 = 26$ and
$N(\overline{c}_1c_2) = N(c_2) - N(c_1c_2) = 30 - 9 = 21$.
Now, we want $N(\overline{c}_1\overline{c}_2)$. We note that $N(\overline{c}_1) = N(\overline{c}_1c_2) + N(\overline{c}_1\overline{c}_2)$.
Substituting the values, we get
$N(\overline{c}_1\overline{c}_2) = N(\overline{c}_1) - N(\overline{c}_1c_2) = 65 - 21 = 44$.
Actually, the computations have happened as follows:
$N(\overline{c}_1\overline{c}_2) = N(\overline{c}_1) - N(\overline{c}_1c_2) = [N - N(c_1)] - [N(c_2) - N(c_1c_2)] =$
$N - N(c_1) - N(c_2) + N(c_1c_2)$.

Note 1

$N(\overline{c}_1\overline{c}_2)$ is not the same as $N(\overline{c_1c_2})$. In the Example 1, $N(\overline{c}_1\overline{c}_2) = 44$, while $N(\overline{c_1c_2}) = N - N(c_1c_2) = 100 - 9 = 91$, which is $65 + 70 - 44 = N(\overline{c}_1) + N(\overline{c}_2) - N(\overline{c}_1\overline{c}_2)$.

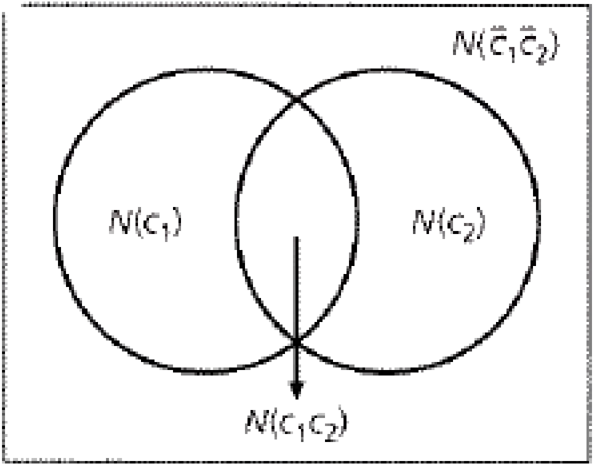

Figure: Pictorial depiction of the Example 1

Note 2

If we extend the example 1 with another property, say
c_3: *A student at NITK is among the 100 students in M.Tech. and is enrolled in Chess Training.*
then the number of students $N(\overline{c}_1, \overline{c}_2, \overline{c}_3)$ *will be obtained as*
$N(\overline{c}_1, \overline{c}_2, \overline{c}_3) = N - [N(c_1) + N(c_2) + N(c_3)] + [N(c_1 c_2) + N(c_2 c_3) + N(c_1 c_3)] - N(c_1 c_2 c_3)$.

This leads to define:

Definition 1

Consider a set S, with $|S| = N$, and conditions c_i, $(1 \leq i \leq t)$, each of which may be satisfied by some of the elements of S. Then the number of elements of S that satisfy ***none*** *of the conditions c_i, $(1 \leq i \leq t)$, is denoted by $\overline{N} = N(\overline{c}_1, \overline{c}_2, \cdots, \overline{c}_t)$, where*

$$\overline{N} = N - [N(c_1) + N(c_2) + \cdots + N(c_t)] + [N(c_1c_2) + N(c_1c_3) + \cdots + N(c_{t-1}c_t)] - [N(c_1c_2c_3) + N(c_1c_2c_4) + \cdots + N(c_{t-2}c_{t-1}c_t)] + \cdots + (-1)^t N(c_1c_2\cdots c_t)$$

$$= N - \sum_{1 \leq i \leq t} N(c_i) + \sum_{1 \leq i \leq j \leq t} N(c_ic_j) - \sum_{1 \leq i \leq j \leq k \leq t} N(c_ic_jc_k) + \cdots + (-1)^t N(c_1c_2\cdots c_t).$$

Note 3

As an immediate consequence of the definition 1, it follows that $N(c_1 \text{ or } c_2 \text{ or } \cdots c_t) = N - \overline{N}$.

Example 2

Determine the number of positive integers n where $1 \leq n \leq 100$ *and n is not divisible by 2, 3 or 5.*
Solution: Here $S = \{1, 2, \cdots, 100\}$ *and* $N = 100$. *For* $n \in S$, *n satisfies*

1. *condition* c_1 *if n is divisible by 2,*
2. *condition* c_2 *if n is divisible by 3, and*
3. *condition* c_3 *if n is divisible by 5.*

Then the answer to the problem is $N(\overline{c}_1\overline{c}_2\overline{c}_3)$.
Now, $N(c_1) = \lfloor \frac{100}{2} \rfloor = 50$, $N(c_2) = \lfloor \frac{100}{3} \rfloor = 33$, *and* $N(c_3) = \lfloor \frac{100}{5} \rfloor = 20$, $N(c_1c_2) = \lfloor \frac{100}{2\times3} \rfloor = 16$, $N(c_1c_3) = \lfloor \frac{100}{2\times5} \rfloor = 10$, $N(c_2c_3) = \lfloor \frac{100}{3\times5} \rfloor = 6$, *and* $N(c_1c_2c_3) = \frac{100}{2\times3\times5} = 3$.
Applying PIE, we get
$N(\overline{c}_1\overline{c}_2\overline{c}_3) = 100 - [50 + 33 + 20] + [16 + 10 + 6] - 3 = 26.$

Example 3

Find the number of non-negative integer solutions of $x_1 + x_2 + x_3 + x_4 = 18$ *where each* $x_i \leq 7$ *for* $1 \leq i \leq 4$.
Solution: Let S *be the set of non-negative integer solutions of* $x_1 + x_2 + x_3 + x_4 = 18$. *So,* $|S| = N = S_0 = \binom{4+18-1}{18} = \binom{21}{18}$.
Let c_i *be the condition that a solution* $x_i > 7$ *for* $1 \leq i \leq 4$. *Then we need to find* $N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4)$.
Here by symmetry, $N(c_1) = N(c_2) = N(c_3) = N(c_4)$. *To compute* $N(c_1)$, *we consider the integer solutions of* $x_1 + x_2 + x_3 + x_4 = 18$, *with each* $x_i \geq 0$ *for all* $1 \leq i \leq 4$. *Then we add 8 to the value of* x_1 *and get the solutions of* $x_1 + x_2 + x_3 + x_4 = 10$ *that satisfy the condition* c_1.
Hence $N(c_i) = \binom{4+10-1}{10} = \binom{13}{10}$ *for each* $1 \leq i \leq 4$, *and* $S_1 = \binom{4}{1}\binom{13}{10}$.
Likewise, $N(c_1c_2)$ *is the number of non-negative integer solutions of* $x_1 + x_2 + x_3 + x_4 = 2$. *So,* $N(c_1c_2) = \binom{4+2-1}{2} = \binom{5}{2}$, *and* $S_2 = \binom{4}{2}\binom{5}{2}$.
Clearly, $N(c_ic_jc_k) = 0$ *for every* i, j, k *and* $N(c_1c_2c_3c_4) = 0$, *we get*
$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) = S_0 - S_1 + S_2 - S_3 + S_4 = \binom{21}{18} - \binom{4}{1}\binom{13}{10} + \binom{4}{2}\binom{5}{2} - 0 + 0 = 246.$

Example 4

In how many ways can the 26 letters of English alphabet be permuted so that none of the patterns car, dog, pun or byte occurs?

Solution: Let S the set of all permutations of 26 letters, which means, $|S| = 26!$. *For each i with* $1 \leq i \leq 4$, *a permutation in S is said to satisfy condition* c_i *if the permutation contains the pattern car, dog, pun or byte respectively.*

Now, $N(c_1)$ *is the number of ways of permutations of the 24 sybmols* $car, b, d, e, f, \cdots, p, q, s, t, \cdots, x, y, z$. *So,* $N(c_1) = 24!$. *Similarly,* $N(c_2) = N(c_3) = 24!$ *and* $N(c_4) = 23!$.

For $N(c_1c_2)$, *we deal with the 22 symbols* $car, dog, b, e, f, h, i, \cdots, m, n, p, q, s, t, \cdots, x, y, z$. *So,* $N(c_1c_2) = 22!$. *Similarly, we obtain* $N(c_1c_3) = N(c_2c_3) = 22!$ *and* $N(c_ic_4) = 21!$ *for* $i \neq 4$.

Example 4 (contd...)

In a similar fashion, we get $N(c_1c_2c_3) = 20!$, $N(c_ic_jc_4) = 19!$ *for* $1 \leq i < j \leq 3$ *and* $N(c_1c_2c_3c_4) = 17!$.

So, the number of permutations of S that contain none of the given patterns is

$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4) =$

$26! - [3(24!) + 23!] + [3\ 22! + 3\ 21!)] - [20! + 3(19!)] + 17!$

Example 5

Six married couples are to be seated around a circular table. In how many ways can they be arranged to sit so that no wife sits next to her husband? (Note that two seating arrangements are considered the same if one is the rotaion of the other.)

Solution: For $1 \leq i \leq 6$, let c_i denote the condition where a seating arrangement has couple i seated next to each other.

To determine $N(c_1)$, consider arranging 11 distinct objects - namely, couple 1 (considered as one object), and the other 10 people. 11 distinct objects can be arranged around a circular table in $(11 - 1)! = 10!$ ways. However, here it matters whether wife in couple 1 is seated to the left or the right of husband. So, $N(c_1) = 2(10!)$. Similarly, $N(c_i) = 2(10!)$ for $2 \leq i \leq 6$, and $S_1 = \binom{6}{1}2(10!)$.

Example 5 (contd...)

To compute $N(c_i c_j)$ $(1 \leq i < j \leq 6)$*, we note that we are arranging 10 distinct objects - Couple i as one object, Couple j as another, while the rest 8 as 8 objects. This arrangement can be done in* $(10-1)! = 9!$ *ways. So, here* $N(c_i c_j) = 2^2(9!)$ *because there are two ways each for the couples i and j for the husband and wife to sit next to each other. Consequently,* $S_2 = \binom{6}{2}2^2(9!)$*.*

With similar reasoning, we get

$N(c_1c_2c_3) = 2^3(8!)$ *and* $S_3 = \binom{6}{3}2^3(8!)$, $N(c_1c_2c_3c_4) = 2^4(7!)$ *and* $S_4 = \binom{6}{4}2^4(7!)$

$N(c_1c_2c_3c_4c_5) = 2^5(6!)$ *and* $S_5 = \binom{6}{5}2^5(6!)$,

$N(c_1c_2c_3c_4c_5c_6) = 2^6(5!)$ *and* $S_65 = \binom{6}{6}2^6(5!)$.

Let S_0 *denote the total number of arrangements of 12 people. Hence* $S_0 = (12-1)! = 11!$.

Example 5 (contd...)

Now, the number of arrangements such that no wife sits next to her husband is

$N(\overline{c}_1\overline{c}_2\overline{c}_3\overline{c}_4\overline{c}_5\overline{c}_6) = \sum_{i=0}^{6}(-1)^i S_i = \sum_{i=0}^{6}(-1)^i \binom{6}{i} 2^i (11-i)! =$
$39916800 - 43545600 + 21772800 - 6451200 + 1209600 - 138240 + 7680 = 12771840.$

Consider a set S with $|S| = N$, and conditions $c_1, c_2, \cdots, c_t$ satisfied by some of the elements of S. Let $m \in \mathbb{Z}^+$ such that $1 \leq m \leq t$. Define E_m as the number of elements in S that satisfy *exactly* m of the t conditions.

We can write formulas such as

$E_1 = N(c_1\overline{c}_2\overline{c}_3\cdots\overline{c}_t) + N(\overline{c}_1c_2\overline{c}_3\cdots\overline{c}_t) + \cdots + N(\overline{c}_1\overline{c}_2\overline{c}_2\cdots\overline{c}_{t-1}\overline{c}_t)$

$E_2 = N(c_1c_2\overline{c}_3\cdots\overline{c}_t)+N(c_1\overline{c}_2c_3\cdots\overline{c}_t)+\cdots+N(\overline{c}_1\overline{c}_2\overline{c}_3\cdots\overline{c}_{t-2}c_{t-1}c_t)$

and so on $\cdots$

This is, however, a long and relatively cumbersome process! So, we have $\cdots$.

Theorem 1

*Under the same hypotheses made while stating the PIE, for each $1 \leq m \leq t$, the number of elements in S that satisfy **exactly** m of the conditions $c_1, c_2, \cdots, c_t$ is given by*

$$E_m = S_m - \binom{m+1}{1} S_{m+1} + \binom{m+2}{2} S_{m+2} - \cdots + (-1)^{t-m} \binom{t}{t-m} S_t \quad (1)$$

Corollary

*Let L_m denote the number of elements of S that satisfy **at least** m of the t conditions. Then*

$$L_m = S_m - \binom{m}{m-1} S_{m+1} + \binom{m+1}{m-1} S_{m+2} - \cdots + (-1)^{t-m} \binom{t-1}{m-1} S_t \quad (2)$$