Chapter Two
The Inclusion-Exclusion Principle

The sum principle for counting is the simplest of the basic counting principles. It states that if $A_1, A_2, \ldots, A_n$ are finite sets that are pairwise disjoint, then :

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n|$$

The Inclusion-Exclusion Principle, in its simplest form, provides a formula to calculate $|A_1 \cup A_2 \cup \cdots \cup A_n|$ when we allow the sets $A_1, A_2, \ldots, A_n$ to overlap.

In what follows, we assume that U is a given finite universal set and that $A_1, A_2, \ldots, A_n$ are subsets of U. For each $i = 1, 2, \ldots, n$, we define :

$$\alpha_i = \sum_{\{j_1, j_2, \ldots, j_i\} \subseteq \{1,2,\ldots,n\}} |A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_i}|$$

where the sum is taken over all possible subsets of indices $\{j_1, j_2, \ldots, j_i\} \subseteq \{1, 2, \ldots, n\}$.

Theorem (2.1) (The Inclusion-Exclusion Principle)

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \alpha_1 - \alpha_2 + \cdots + (-1)^{n-1}\alpha_n \tag{1}$$

Proof

Let $x \in A_1 \cup A_2 \cup \cdots \cup A_n$. When calculating $|A_1 \cup A_2 \cup \cdots \cup A_n|$, x is counted exactly once. The situation is different when calculating each $\alpha_1, \alpha_2, \ldots, \alpha_n$. We will prove that the contribution of x to the calculation of the number $\alpha_1 - \alpha_2 + \cdots + (-1)^{n-1}\alpha_n$ is equal to 1.

Assume that x belongs only to m sets among $A_1, \ldots, A_n$. Therefore, the contribution of x to the calculation of $\alpha_1 = |A_1| + |A_2| + \cdots + |A_n|$ is m. Also, the contribution of x to the calculation of α_2 is $\binom{m}{2}$, because the contribution of x to $|A_i \cap A_j|$ is 0 if $\{i, j\} \not\subseteq \{j_1, \ldots, j_m\}$ and 1 if $\{i, j\} \subseteq \{j_1, \ldots, j_m\}$.

Similarly, the contribution of x to the calculation of α_i is $\binom{m}{i}$ for each $1 \leq i \leq n$. Thus, the contribution of x to $\alpha_1 - \alpha_2 + \cdots + (-1)^{n-1}\alpha_n$ is :

$$\binom{m}{1} - \binom{m}{2} + \cdots + (-1)^{m-1}\binom{m}{m}$$

since $\binom{m}{i} = 0$ for each $m < i \leq n$. However, from the Binomial Theorem, we know that :

$$(1-1)^m = \binom{m}{0} - \binom{m}{1} + \cdots + (-1)^m \binom{m}{m} = 0$$

Therefore :

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

This completes the proof. ■

In many problems, we calculate the number of elements that do not belong to any of the sets $A_1, A_2, \ldots, A_n$ using the following result of the Inclusion-Exclusion Principle.

Corollary (2.1)

If U is a finite universal set and $A_1, A_2, \ldots, A_n$ are subsets of U, then :

$$|U - (A_1 \cup A_2 \cup \cdots \cup A_n)| = |U| - \alpha_1 + \alpha_2 - \cdots + (-1)^n \alpha_n \tag{2}$$

Now, we rely on the Inclusion-Exclusion Principle and its corollary to present a collection of various theorems and examples.

Theorem (2.2)

The number of surjective (onto) functions from the set $A = \{a_1, a_2, \ldots, a_m\}$ to the set $B = \{b_1, b_2, \ldots, b_n\}$, where $m \geq n$, is equal to :

$$n^m - \binom{n}{1}(n-1)^m + \binom{n}{2}(n-2)^m - \cdots + (-1)^{n-1}\binom{n}{n-1}1^m = \sum_{k=0}^{n-1}(-1)^k \binom{n}{k}(n-k)^m$$

Proof Let U be the set of functions from A to B. For each $1 \leq k \leq n$, let $A_k = \{f \in U : b_k \notin R(f)\}$, where $R(f)$ denotes the range of f. Thus, we need to calculate the number $|U - (A_1 \cup A_2 \cup \cdots \cup A_n)|$.

It is clear that $|U| = n^m$. Now, we calculate α_1 from the relation $\alpha_1 = |A_1| + |A_2| + \cdots + |A_n|$. From the definition of A_k, it follows that $|A_k|$ is the number of functions from A to $B - \{b_k\}$. Thus, $|A_k| = (n-1)^m$ for each $1 \leq k \leq n$. Therefore, $\alpha_1 = \binom{n}{1}(n-1)^m$.
To calculate α_2 :
We notice that $|A_i \cap A_j|$, where $1 \leq i < j \leq n$, is equal to the number of functions from A to $B - \{b_i, b_j\}$. Thus, $|A_i \cap A_j| = (n-2)^m$, and consequently :

$$\alpha_2 = \binom{n}{2}(n-2)^m$$

Similarly, we find that $\alpha_k = \binom{n}{k}(n-k)^m$ for each $1 \leq k \leq n$. Therefore :

$$\begin{aligned}
|U - (A_1 \cup A_2 \cup \cdots \cup A_n)| &= |U| - \alpha_1 + \alpha_2 - \cdots + (-1)^n \alpha_n \\
&= n^m - \binom{n}{1}(n-1)^m + \binom{n}{2}(n-2)^m - \cdots + (-1)^{n-1}\binom{n}{n-1}1^m \\
&= \sum_{k=0}^{n-1}(-1)^k \binom{n}{k}(n-k)^m
\end{aligned}$$

Example (2.3)

Find the number of integers x such that $1 \leq x \leq 500$, where x is not divisible by 5, not divisible by 6, and not divisible by 8.

Solution Let $U = \{1, 2, \ldots, 500\}$, and let $A_1 = \{x \in U : 5|x\}$, $A_2 = \{x \in U : 6|x\}$, and $A_3 = \{x \in U : 8|x\}$. We want to calculate the number $|U - (A_1 \cup A_2 \cup A_3)|$. We note that :

$$|A_1| = \left\lfloor \frac{500}{5} \right\rfloor = 100, \quad |A_2| = \left\lfloor \frac{500}{6} \right\rfloor = 83, \quad |A_3| = \left\lfloor \frac{500}{8} \right\rfloor = 62$$

As is known, $a|n$ and $b|n$ if and only if $\text{lcm}(a, b)|n$. Therefore, we find :

$$\begin{aligned}
|A_1 \cap A_2| &= \left\lfloor \frac{500}{\text{lcm}(5,6)} \right\rfloor = \left\lfloor \frac{500}{30} \right\rfloor = 16 \\
|A_1 \cap A_3| &= \left\lfloor \frac{500}{\text{lcm}(5,8)} \right\rfloor = \left\lfloor \frac{500}{40} \right\rfloor = 12 \\
|A_2 \cap A_3| &= \left\lfloor \frac{500}{\text{lcm}(6,8)} \right\rfloor = \left\lfloor \frac{500}{24} \right\rfloor = 20 \\
|A_1 \cap A_2 \cap A_3| &= \left\lfloor \frac{500}{\text{lcm}(5,6,8)} \right\rfloor = \left\lfloor \frac{500}{120} \right\rfloor = 4
\end{aligned}$$

Therefore :

$$\begin{aligned}
|U - (A_1 \cup A_2 \cup A_3)| &= |U| - (\alpha_1 - \alpha_2 + \alpha_3) \\
&= 500 - (100 + 83 + 62) + (16 + 12 + 20) - 4 = 299
\end{aligned}$$

Euler's Phi Function

For a positive integer n, Euler's phi function $\phi(n)$ is the number of positive integers less than or equal to n that are relatively prime to n. Thus,

$$\phi(n) = \big|\{x \in \mathbb{Z} : 1 \leq x \leq n \text{ and } \gcd(x, n) = 1\}\big|.$$

In particular, $\phi(1) = 1$.

Example (2.4)

Compute $\phi(40)$.

Direct solution. Since $40 = 2^3 \cdot 5$, an integer is relatively prime to 40 if and only if it is divisible by neither 2 nor 5. The positive integers less than or equal to 40 satisfying this condition are

$$1, 3, 7, 9, 11, 13, 17, 19, 21, 23, 27, 29, 31, 33, 37, 39.$$

Therefore,

$$\boxed{\phi(40) = 16.}$$

Solution using the Inclusion–Exclusion Principle. Let $U = \{1, 2, \ldots, 40\}$. We wish to count the integers in U that are relatively prime to 40. Since $40 = 2^3 \cdot 5$, an integer is relatively prime to 40 if and only if it is divisible by neither 2 nor 5.

Define

$$A_1 = \{x \in U : 2 \mid x\} \qquad \text{and} \qquad A_2 = \{x \in U : 5 \mid x\}.$$

Then $A_1 \cap A_2$ consists of the multiples of 10 in U. Hence,

$$|A_1| = \frac{40}{2} = 20, \qquad |A_2| = \frac{40}{5} = 8, \qquad |A_1 \cap A_2| = \frac{40}{10} = 4.$$

By the inclusion–exclusion principle,

$$\begin{aligned} \phi(40) &= |U \setminus (A_1 \cup A_2)| \\ &= |U| - |A_1| - |A_2| + |A_1 \cap A_2| \\ &= 40 - 20 - 8 + 4 \\ &= \boxed{16}. \end{aligned}$$

A General Formula for Euler's Phi Function

Let $n \geq 2$ be an integer with prime factorization

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k},$$

where $p_1, \ldots, p_k$ are distinct primes and $e_1, \ldots, e_k$ are positive integers. Then

$$\boxed{\phi(n) = n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right).}$$

Proof using the Inclusion–Exclusion Principle. Let $U = \{1, 2, \ldots, n\}$. An integer $x \in U$ is relatively prime to n if and only if none of $p_1, \ldots, p_k$ divides x.

For each $i \in \{1, \ldots, k\}$, define

$$A_i = \{x \in U : p_i \mid x\}.$$

Thus,

$$\phi(n) = \left| U \setminus \bigcup_{i=1}^{k} A_i \right|.$$

For any indices $1 \leq i_1 < \cdots < i_r \leq k$, the intersection $A_{i_1} \cap \cdots \cap A_{i_r}$ consists of the multiples of $p_{i_1} \cdots p_{i_r}$ in U. Since this product divides n,

$$|A_{i_1} \cap \cdots \cap A_{i_r}| = \frac{n}{p_{i_1} \cdots p_{i_r}}.$$

In particular, we notice that $|A_i| = \frac{n}{p_i}$, $|A_i \cap A_j| = \frac{n}{p_i p_j}$, and so on.

By the inclusion–exclusion principle, we first subtract the number of multiples of each prime p_i, then add back the number of multiples of each product $p_i p_j$, and continue with

alternating signs. Thus,

$$\phi(n) = n - \sum_{i=1}^{k} |A_i| + \sum_{1\le i<j\le k} |A_i \cap A_j| - \cdots + (-1)^k |A_1 \cap \cdots \cap A_k|$$

$$= n - \sum_{i=1}^{k} \frac{n}{p_i} + \sum_{1\le i<j\le k} \frac{n}{p_i p_j} - \cdots + (-1)^k \frac{n}{p_1 p_2 \cdots p_k}.$$

We can write $\phi(n)$ in another way, as follows

$$\phi(n) = n + \sum_{r=1}^{k} (-1)^r \sum_{1\le i_1<\cdots<i_r\le k} \frac{n}{p_{i_1} \cdots p_{i_r}}.$$

Factoring out n, we obtain

$$\phi(n) = n\left(1 - \sum_{i=1}^{k} \frac{1}{p_i} + \sum_{1\le i<j\le k} \frac{1}{p_i p_j} - \cdots + (-1)^k \frac{1}{p_1 p_2 \cdots p_k}\right)$$

$$= n\left(1 + \sum_{r=1}^{k} (-1)^r \sum_{1\le i_1<\cdots<i_r\le k} \frac{1}{p_{i_1} \cdots p_{i_r}}\right).$$

To recognize the expression in parentheses, consider first the product of two factors :

$$\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) = 1 - \frac{1}{p_1} - \frac{1}{p_2} + \frac{1}{p_1 p_2}.$$

With three factors, we similarly obtain

$$\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\left(1 - \frac{1}{p_3}\right)$$
$$= 1 - \left(\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3}\right) + \left(\frac{1}{p_1 p_2} + \frac{1}{p_1 p_3} + \frac{1}{p_2 p_3}\right) - \frac{1}{p_1 p_2 p_3}.$$

In general, when expanding the product of k factors, we choose either 1 or $-1/p_i$ from each factor. Choosing $-1/p_i$ from exactly r factors gives a term with sign $(-1)^r$. Therefore,

$$\prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right) = 1 - \sum_{i=1}^{k} \frac{1}{p_i} + \sum_{1\le i<j\le k} \frac{1}{p_i p_j} - \cdots + (-1)^k \frac{1}{p_1 p_2 \cdots p_k}.$$

Consequently,

$$\boxed{\phi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_k}\right) = n\prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right).}$$

Application to Example (2.4). Since $40 = 2^3 \cdot 5$, the formula gives

$$\phi(40) = 40\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{5}\right) = 40 \cdot \frac{1}{2} \cdot \frac{4}{5} = 16.$$

Example (2.5)

Find the number of integer solutions to the equation $X_1 + X_2 + X_3 = 13$ subject to the conditions $0 \leq X_1 \leq 6$, $0 \leq X_2 \leq 9$, and $0 \leq X_3 \leq 3$.

Solution Let U be the set of all non-negative integer solutions to the equation $X_1 + X_2 + X_3 = 13$.
Let A_1 be the set of integer solutions with $X_1 \geq 7$, A_2 with $X_2 \geq 10$, and A_3 with $X_3 \geq 4$. Thus, we need to calculate the number $|U - (A_1 \cup A_2 \cup A_3)|$.

It is clear that $|U| = \binom{3-1+13}{13} = \binom{15}{13} = 105$. Similarly, we find that :

$$|A_1| = \binom{3-1+13-7}{13-7} = \binom{8}{6} = 28$$
$$|A_2| = \binom{3-1+13-10}{13-10} = \binom{5}{3} = 10$$
$$|A_3| = \binom{3-1+13-4}{13-4} = \binom{11}{9} = 55$$

We also note that $|A_1 \cap A_2| = 0$ because $7 + 10 = 17 > 13$.
Furthermore :

$$|A_1 \cap A_3| = \binom{3-1+13-7-4}{13-7-4} = \binom{4}{2} = 6$$
$$|A_2 \cap A_3| = 0 \quad (\text{since } 10 + 4 = 14 > 13)$$
$$|A_1 \cap A_2 \cap A_3| = 0$$

Therefore, the number of required solutions is :

$$105 - (28 + 10 + 55) + (0 + 6 + 0) - 0 = 18$$

From the definition of the union of sets, it follows that the Inclusion-Exclusion Principle provides the number of elements that belong to at least one of the sets $A_1, A_2, \ldots, A_n$.

Exercises

1. A survey of 100 students showed that 32 students are enrolled in Course A, 44 students are enrolled in Course B, 47 students are enrolled in Course C, 11 students are enrolled in both Courses B and C, 12 students are enrolled in both Courses A and C, 12 students are enrolled in both Courses A and B, and 3 students are enrolled in all three courses. Find the number of students who are not enrolled in any of the three courses.
2. Tests were conducted on 200 groundwater samples to detect the presence of salts A, B, and C. It was found that 14 samples contain salt A, 10 samples contain salt B, 8 samples contain salt C, 6 samples contain salts A and B, 6 samples contain salts B and C, 4 samples contain salts A and C, and 2 samples contain salts A and B but do not contain salt C. Find the number of samples containing at least one of the three salts.

3. Find the number of integer solutions to the equation $X_1 + X_2 + X_3 + X_4 + X_5 = 20$ subject to :
   (a) $0 \leq X_i \leq 10$ for all $i = 1, 2, \ldots, 5$.
   (b) $0 \leq X_i \leq 8$ for all $i = 1, 2, \ldots, 5$.
4. Find the number of integer solutions to the equation $X_1 + X_2 + X_3 + X_4 = 30$ subject to :
   (a) $0 \leq X_i \leq 8$ for all $i = 1, 2, 3, 4$.
   (b) $10 \leq X_i \leq 20$ for all $i = 1, 2, 3, 4$.
   (c) $0 \leq X_1 \leq 6,\ 0 \leq X_2 \leq 9,\ 0 \leq X_3 \leq 15,\ 0 \leq X_4 \leq 18$.
5. Find the number of integer solutions to the equation $X_1 + X_2 + X_3 = 30$ subject to :
$$5 \leq X_1 < 11, \quad 6 < X_2 \leq 14, \quad 10 \leq X_3 \leq 24$$
6. If $A = \{1, 2, \ldots, 999999\}$, what is the number of integers belonging to A such that the sum of the digits of each number equals 15 ?
7. Find the number of multisets of size 15 chosen from the set $A = \{a_1, a_2, a_3\}$ such that the repetition of a_1 is less than 5, the repetition of a_2 is less than 7, and the repetition of a_3 is less than 6.
8. Find the number of integers n, where $1 \leq n \leq 2000$, such that :
   (a) $2 \nmid n,\ 3 \nmid n,\ 5 \nmid n$.
   (b) $2 \nmid n,\ 3 \nmid n,\ 5 \nmid n,\ 7 \nmid n$.
   (c) $2 \nmid n,\ 3 \nmid n,\ 5 \nmid n,\ 7 \nmid n$.
9. If $n = p_1^{a_1} p_2^{a_2}$ where both p_1 and p_2 are prime numbers, prove that :
$$\phi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)$$
   Then calculate $\phi(135)$.