Combinatorics

Instructor: Jie Ma, Scribed by Tianchi Yang

1 Lecture 3. Inclusion and exclusion

This lecture is devoted to Inclusion-exclusion formula and its applications.

Let Ω be a ground set and let $A_1, A_2, ..., A_n$ be subsets of Ω. Write $A_i^c = \Omega \backslash A_i$. Throughout this lecture, we use the following notation.

Definition 1.1. *Let $A_\emptyset = \Omega$. For any nonempty subset $I \subseteq [n]$, let*

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

For any integer $k \geq 0$, let

$$S_k = \sum_{I \in \binom{[n]}{k}} |A_I|.$$

Now we introduce Inclusion-exclusion formula (in three equivalent forms) and give two proofs as following.

Theorem 1.2 (Inclusion-exclusion Formula)**.**

$$|A_1 \cup A_2 \cup ... \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} S_k$$

$$\Longleftrightarrow \left| \Omega \backslash \bigcup_{i=1}^{n} A_i \right| = |A_1^c \cap A_2^c \cap ... \cap A_n^c| = \sum_{k=0}^{n} (-1)^k S_k$$

$$\Longleftrightarrow \left| \Omega \backslash \bigcup_{i=1}^{n} A_i \right| = |A_1^c \cap A_2^c \cap ... \cap A_n^c| = \sum_{I \subseteq [n]} (-1)^{|I|} |A_I|.$$

Proof (first). For any subset $X \subseteq \Omega$, we define its characterization function $\mathbb{1}_X : \Omega \to \{0, 1\}$ by assigning

$$\mathbb{1}_X(x) = \begin{cases} 1, & x \in X \\ 0, & x \notin X. \end{cases}$$

Then $\sum_{x \in \Omega} \mathbb{1}_X(x) = |X|$. Let $A = A_1 \cup A_2 \cup ... \cup A_n$. Our key observation is that

$$(\mathbb{1}_A - \mathbb{1}_{A_1})(\mathbb{1}_A - \mathbb{1}_{A_2}) \cdots (\mathbb{1}_A - \mathbb{1}_{A_n})(x) \equiv 0$$

holds for any $x \in \Omega$. Next we expand this product into a summation of 2^n terms as following:

$$\sum_{I \subseteq [n]} (-1)^{|I|} (\prod_{i \in I} \mathbb{1}_{A_i}) \equiv 0 \iff \mathbb{1}_A(x) + \sum_{I \subseteq [n],\, I \neq \emptyset} (-1)^{|I|} \mathbb{1}_{A_I}(x) \equiv 0$$

holds for any $x \in \Omega$. Summing over all $x \in \Omega$, this gives that

$$|A| + \sum_{I \subseteq [n],\, I \neq \emptyset} (-1)^{|I|} |A_I| = 0,$$

which implies that

$$|A_1 \cup A_2 \cup ... \cup A_n| = |A| = \sum_{\substack{I \subseteq [n] \\ I \neq \emptyset}} (-1)^{|I|+1} |A_I| = \sum_{k=1}^{n} (-1)^{k+1} S_k,$$

finishing the proof. ∎

Proof (second). It suffices to prove that

$$\mathbb{1}_{A_1 \cup A_2 \cup ... \cup A_n}(x) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \binom{[n]}{k}} \mathbb{1}_{A_I}(x)$$

holds for all $x \in \Omega$. Denote by LHS (resp. RHS) the left (resp. right) side of the above equation.

Assume that x is contained in exactly ℓ subsets, say $A_1, A_2, \cdots, A_\ell$. If $\ell = 0$, then clearly $LHS = 0 = RHS$, so we are done. So we may assume that $\ell \geq 1$. In this case, we have $LHS = 1$ and

$$RHS = \ell - \binom{\ell}{2} + \binom{\ell}{3} + \cdots + (-1)^{\ell+1} \binom{\ell}{\ell} = 1.$$

Note that the above equation holds since $\sum_{i=0}^{\ell} (-1)^i \binom{\ell}{i} = (1-1)^\ell = 0$. This finishes the proof. ∎

Next, we will demonstrate the power of Inclusion-exclusion formula by using it to solve several problems.

Definition 1.3. *Let $\varphi(n)$ be the number of integers $m \in [n]$ which are relatively prime*[1] *to n.*

Theorem 1.4. *If we express $n = p_1^{a_1} p_2^{a_2} \cdots p_t^{a_t}$, where $p_1 \cdots p_t$ are distinct primes, then*

$$\varphi(n) = n \prod_{i=1}^{t} (1 - \frac{1}{p_i}).$$

Proof. Let $A_i = \{m \in [n] : p_i | m\}$ for $i \in \{1, 2, \cdots, t\}$. It implies

$$\varphi(n) = \big|\{m \in [n] : m \notin A_i \text{ for all } i \in [t]\}\big| = \big|[n] \backslash (A_1 \cup A_2 \cup \cdots \cup A_t)\big|.$$

By Inclusion-exclusion formula,

$$\varphi(n) = \sum_{I \subseteq [t]} (-1)^{|I|} |A_I|,$$

where $A_I = \cap_{i \in I} A_i = \{m \in [n] : (\prod_{i \in I} p_i) | m\}$ and thus $|A_I| = n / \prod_{i \in I} p_i$. We can derive that

$$\varphi(n) = \sum_{I \subseteq [t]} (-1)^{|I|} \frac{n}{\prod_{i \in I} p_i} = n(1 - \frac{1}{p_1})(1 - \frac{1}{p_2}) \cdots (1 - \frac{1}{p_t}),$$

as desired. ∎

[1] Here, "m is relatively prime to n" means that the greatest common divisor of m and n is 1.

Definition 1.5. *A permutation* $\sigma : [n] \to [n]$ *is called a* **derangment** *of* $[n]$ *if* $\sigma(i) \neq i$ *for all* $i \in [n]$.

Theorem 1.6. *Let* D_n *be the family of all derangment of* $[n]$. *Then*

$$|D_n| = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Proof. Let

$$A_i = \{\text{all permutations } \sigma : [n] \to [n] \text{ such that } \sigma(i) = i\}.$$

Then

$$D_n = A_1^c \cap A_2^c \cap \cdots \cap A_n^c \text{ and } |A_I| = (n - |I|)!.$$

By Inclusion-exclusion formula, we get

$$|D_n| = \sum_{I \subseteq [n]} (-1)^{|I|} |A_I| = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)! = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!} = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!},$$

as desired. ∎

Remark 1.7. *We have that*

$$|D_n| \to \frac{n!}{e} \text{ as } n \to \infty.$$

It is because $\sum_{k=0}^{+\infty} \frac{(-1)^k}{k!} = e^{-1}$ *(by the Taylor series of* $e^x = \sum_{k=0}^{+\infty} \frac{x^k}{k!}$*).*

Next we recall the definition of $S(n,k)$ and aim to give a precise formula for it. We know that

(1.) $S(n,k)$ is equal to the number of partitions of $[n]$ into k non-ordered non-empty set.

(2.) $S(n,k)k!$ is equal to the number of surjective functions $f : [n] \to [k]$.

Theorem 1.8. *We have*

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

Proof. For $i \in [k]$, let

$$A_i = \{\text{all functions } f : [n] \to [k] \backslash \{i\}\}.$$

Then

$$A_1^c \cap A_2^c \cap \cdots \cap A_k^c = \{\text{all surjective } f : [n] \to [k]\}.$$

So

$$S(n,k)k! = \#\text{surjective } f : [n] \to [k] = \sum_{i=0}^{k} (-1)^i S_i,$$

where

$$S_i = \sum_{I \in \binom{[k]}{i}} |A_I| = \binom{k}{i} (k-i)^n.$$

Finally, we get

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

∎

2 Generating functions

Definition 2.1. *The (ordinary)* **generating function** *for an infinity sequence* $\{a_0, a_1, \cdots\}$ *is a power series*

$$f(x) = \sum_{n\geq 0} a_n x^n.$$

We have two ways to view this power series.

(i). When the power series $\sum_{n\geq 0} a_n x^n$ converges (i.e. there exists a radius $R > 0$ of convergence), we view GF as a function of x and we can apply operations of calculus on it (including differentiation and integration). For example, we know that

$$a_n = \frac{f^{(n)}(0)}{n!}.$$

Recall the following sufficient condition on the radius of convergence that if $|a_n| \leq K^n$ for some $K > 0$, then $\sum_{n\geq 0} a_n x^n$ converges in the interval $(-\frac{1}{K}, \frac{1}{K})$.

(ii). When we are not sure of the convergence, we view the generating function as a formal object with additions and multiplications. Let $a(x) = \sum_{n\geq 0} a_n x^n$ and $b(x) = \sum_{n\geq 0} b_n x^n$.

Addition.

$$a(x) + b(x) = \sum_{n\geq 0} (a_n + b_n) x^n.$$

Multiplication. Let $c_n = \sum_{i=0}^{n} a_i b_{n-i}$. Then

$$a(x)b(x) = \sum_{n\geq 0} c_n x^n.$$

Example 2.2. *We see* $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ *holds for all* $-1 < x < 1$. *By the point view of (i), its first derivative gives*

$$\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1} = \sum_{n=0}^{\infty} (n+1) x^n.$$

Problem 2.3. *Let* $a_0 = 1$ *and* $a_n = 2a_{n-1}$ *for* $n \geq 1$. *Find* a_n.

Solution. Consider the generating function,

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = 1 + \sum_{n=1}^{\infty} a_n x^n = 1 + 2x \sum_{n=1}^{\infty} a_{n-1} x^{n-1} = 1 + 2x f(x).$$

So $f(x) = \frac{1}{1-2x}$, which implies that $f(x) = \sum_{n=0}^{+\infty} 2^n x^n$ and $a_n = 2^n$. ∎

From this problem, we see one of the basic ideas for using generating function: in order to find the general expression of a_n, we work on its generating function $f(x)$; once we find the formula of $f(x)$, then we can expand $f(x)$ into a power series and get a_n by choosing the coefficient of the right term.

Fact 2.4. *For $j \in [n]$, let $f_j(x) := \sum_{i \in I_j} x^i$, where $I_j \subset Z$. Let b_k be the number of solutions to $i_1 + i_2 + ... + i_n = k$ for $i_j \in I_j$. Then*

$$\prod_{j=1}^{n} f_j(x) = \sum_{k=0}^{\infty} b_k x^k.$$

Fact 2.5. *If $f(x) = \prod_{i=1}^{k} f_i(x)$ for polynomials $f_1, ..., f_k$, then*

$$[x^n]f = \sum_{i_1+i_2+\cdots+i_k=n} \prod_{j=1}^{k} \left([x^{i_j}]f_j\right),$$

where $[x^n]f$ is the coefficient of x^n in f.

Problem 2.6. *Let A_n be the set of strings of length n with entries from the set $\{a, b, c\}$ and with no "aa" occuring (in the consecutive positions). Find $|A_n|$ for $n \geq 1$.*

Solution. Let $a_n = |A_n|$. We first observe that $a_1 = 3, a_2 = 8$. For $n \geq 3$, we will find a_n by recursion as following. If the first string is 'a', the second string has two choices, 'b' or 'c'. Then the last $n-2$ strings have a_{n-2} choices. If the first string is 'b' or 'c', the last $n-1$ strings have a_{n-1} choices. They are all different. Totally, for $n \geq 3$, we have

$$a_n = 2a_{n-1} + 2a_{n-2}.$$

Set $a_0 = 1$, then $a_n = 2a_{n-1} + 2a_{n-2}$ holds for $n \geq 2$. The generating function of $\{a_n\}$ is

$$f(x) = \sum_{n\geq 0} a_n x^n = a_0 + a_1 x + \sum_{n\geq 2}(2a_{n-1} + 2a_{n-2})x^n = 1 + 3x + 2x(f(x) - 1) + 2x^2 f(x),$$

which implies that

$$f(x) = \frac{1+x}{1-2x-2x^2}.$$

By Partial Fraction Decomposition, we calculate that

$$f(x) = \frac{1-\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}+1+2x} + \frac{1+\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}-1-2x},$$

which implies that

$$a_n = \frac{1-\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}+1} \left(\frac{-2}{\sqrt{3}+1}\right)^n + \frac{1+\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}-1} \left(\frac{2}{\sqrt{3}-1}\right)^n.$$

Note that a_n must be an integer but its expression is of a combination of irrational terms! Observe that $\left|\frac{-2}{\sqrt{3}+1}\right| < 1$, so $\left(\frac{-2}{\sqrt{3}+1}\right)^n \to 0$ as $n \to \infty$. Thus, when n is sufficiently large, this integer a_n is about the value of the second term $\frac{1+\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}-1} \left(\frac{2}{\sqrt{3}-1}\right)^n$. Equivalently a_n will be the nearest integer to that. ∎

Definition 2.7. *For any real r and an integer $k \geq 0$, let*

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

Theorem 2.8 (Newton's Binomial Theorem)**.** *For any real number r and $x \in (-1, 1)$,*

$$(1+x)^r = \sum_{k=0}^{\infty} \binom{r}{k} x^k.$$

Proof. By Taylor series, it is obvious. ∎

Corollary 2.9. *Let $r = -n$ for some integer $n \geq 0$. Then*

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

Therefore

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

Problem 2.10. *Let a_n be the number of ways to pay n Yuan using 1-Yuan bills, 2-Yuan bills and 5-Yuan bills. What is the generating function of this sequence $\{a_n\}$?*

Solution. Observe that a_n is the number of integer solutions (i_1, i_2, i_3) to $i_1 + i_2 + i_3 = n$, where $i_1 \in I_1 := \{0, 1, 2, ...\}$, $i_2 \in I_2 := \{0, 2, 4, ...\}$ and $i_3 \in I_3 := \{0, 5, 10, ...\}$. Let $f_j(x) := \sum_{m \in I_j} x^m$ for $j = 1, 2, 3$. By Fact 2.4, we have

$$\sum_{n=0}^{+\infty} a_n x^n = f_1(x) f_2(x) f_3(x) = \frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^5}.$$

∎