Coupon Problem

Ruizhe Li

March 2025

Preliminaries

0.1 Binomial inversion

Here, consider the binomial inversion formula:

$$f(n) = \sum_{i=m}^{n} \binom{n}{i} g(i) \Leftrightarrow g(n) = \sum_{i=m}^{n} (-1)^{n-i} \binom{n}{i} f(i)$$

Proof. Notice that f and g are symmertric, so we just need to prove the $\Leftarrow$ side. Substituing $g(i)$'s formulas into $f(n)$, we get:

$$f(n) = \sum_{i=m}^{n} \binom{n}{i} \sum_{j=m}^{i} (-1)^{i-j} \binom{i}{j} f(j) = \sum_{i=m}^{n} \sum_{j=m}^{i} (-1)^{i-j} \binom{n}{i} \binom{i}{j} f(j)$$

Hence,

$$f(n) = \sum_{j=m}^{n} f(j) \sum_{i=j}^{n} (-1)^{i-j} \binom{n}{i} \binom{i}{j}$$

Notice that:

$$\sum_{i=j}^{n} (-1)^{i-j} \binom{n}{i} \binom{i}{j} = \sum_{i=j}^{n} \binom{n}{j} \binom{n-j}{i-j} (-1)^{i-j} = \binom{n}{j} \sum_{i=j}^{n} \binom{n-j}{i-j} (-1)^{i-j}$$

What remains to prove is obviously. $\square$

Note 1. If you let $G(i) = (-1)^i g(i)$, then we have different formulas:

$$f(n) = \sum_{i=m}^{n} (-1)^i \binom{n}{i} G(i) \Leftrightarrow G(n) = \sum_{i=m}^{n} (-1)^i \binom{n}{i} f(i)$$

There is also a different inversion formula:

$$f(n) = \sum_{i=n}^{m} \binom{i}{n} g(i) \Leftrightarrow g(n) = \sum_{i=n}^{m} (-1)^{i-n} \binom{i}{n} f(i)$$

Proof. Similarly, we just need to prove the $\Leftarrow$ side. Substituing $g(i)$'s formulas into $f(n)$, we get:

$$f(n) = \sum_{i=n}^{m} \binom{i}{n} \sum_{j=i}^{m} (-1)^{j-i} \binom{j}{i} f(j) \tag{1}$$

$$= \sum_{j=n}^{m} f(j) \sum_{i=n}^{j} (-1)^{j-i} \binom{i}{n} \binom{j}{i} \tag{2}$$

$$= \sum_{j=n}^{m} \binom{j}{n} f(j) \sum_{i=n}^{j} \binom{j-n}{j-i} (-1)^{j-i} \tag{3}$$

$$= \sum_{j=n}^{m} \binom{j}{n} f(j) \sum_{t=0}^{j-n} \binom{j-n}{t} (-1)^{t} 1^{j-n-t} \tag{4}$$

$$= \sum_{j=n}^{m} \binom{j}{n} f(j) \delta_{jn} = f(n) \tag{5}$$

□

0.2 Min-max approach

Let k(S) denotes the k-th largest element of a finite set S,Then

$$k(S) = \sum_{T \subset S, |T| \geq k} (-1)^{|T|-k} \binom{|T|-1}{k-1} Min(T)$$

Proof. Assume there exists g satisfies

$$k(S) = \sum_{T \subset S, T \neq \phi} g(|T|) Min(T)$$

Arrange the elements of S from large to small, denoted by $x_1, x_2, \ldots, x_m$,Then for x_i ,its contribution to the left is just δ_{ik} , its contribution to right is$\sum_{j=0}^{i-1} \binom{i-1}{j} g(j+1)$

The inequality holds iff $\delta_{ik} = \sum_{j=0}^{i-1} \binom{i-1}{j} g(j+1)$

Set $F(i) = \delta_{i+1,k}$,i.e. $\delta_{ik} = F(i-1)$,$G(i) = g(i+1)$, we just need to prove :$F(i) = \sum_{j=0}^{i} \binom{i}{j} G(j)$

Using binomial inversion, we have: this is equivalent to: $G(i) = \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} F(j) = (-1)^{i-k+1} \binom{i}{k-1}$

So $g(i) = G(i-1) = (-1)^{i-k} \binom{i-1}{k-1}$ is enough.

(Here, consider the general definition of combinatorial numbers) Hence,

$$k(S) = \sum_{T \subset S, |T| \geq k} (-1)^{|T|-k} \binom{|T|-1}{k-1} Min(T)$$

□

1 The Coupon Problem

consider the following problem:

Core Problem 1 (Coupon Problem)**.** There are m different types of toys hidden in blind boxes. Each time a person buy a box and see a toy it is, independently of ones previously obtained, a type j toy with probability p_j, with $\sum_{j=1}^{m} p_j = 1$. Let N denote the number of boxes one needs to buy in order to have a complete collection of at least one of each type. Find $E[N]$

Proof. Let X_i denote the number of toys the man need to see to obtain the first toy of type i.

The total time spent to complete the collection is $X = \max\{X_1, \ldots, X_m\}$. It's obviously that:

$$min_{i \in \mathcal{A}} X_i \sim Geo(\sum_{i \in \mathcal{A}} p_i)$$

Hence, we can use the min-max approach to find the expected number of time needed:

$$\begin{aligned}
E[X] &= E[\max_{i=1,\ldots,m} X_i] && (6)\\
&= \sum_i \mathbb{E}[X_i] - \sum_{i<j} \mathbb{E}[\min(X_i, X_j)] + \sum_{i<j<k} \mathbb{E}[\min(X_i, X_j, X_k)] - \ldots && (7)\\
&= \sum_i \frac{1}{p_i} - \sum_{i<j} \frac{1}{p_i + p_j} + \sum_{i<j<k} \frac{1}{p_i + p_j + p_k} - \ldots && (8)\\
&\cdots + (-1)^{m-1} \frac{1}{p_1 + \cdots + p_m}. && (9)\\
& && (10)
\end{aligned}$$

So, after some calculating, we get:

$$E[X] = \int_0^{+\infty} \left(1 - \prod_{i=1}^{m} (1 - \mathrm{e}^{-p_i x})\right) dx.$$

□