组合数学入门

wwlw

2024.6.8

从组合数说起

符号 $\binom{n}{m}$ 读作 n 选 m，例如 $\binom{4}{2}=6$，有

$$\binom{n}{m}=\frac{n(n-1)\ldots(n-m+1)}{m(m-1)\ldots(1)}=\frac{n!}{m!(n-m)!}$$

可以定义广义二项式系数如下

$$\binom{r}{k}=\begin{cases}\frac{r(r-1)\ldots(r-k+1)}{k(k-1)\ldots(1)}=\frac{r^{\underline{k}}}{k!}, & \text{整数}k\geq 0\\ 0, & \text{整数}k<0\end{cases}$$

从组合数说起

求组合数的几种方法

- 预处理阶乘及其逆元，预处理 $O(n)$ ，查询 $O(1)$ 。
- 当 n 很小，或者题目没有给定模数，用加法公式预处理（可能用到高精度），预处理 $O(n^2)$ ，查询 $O(1)$。
- 当 m 很小，可通过定义 $\frac{n^{\underline{m}}}{m!}$ 暴力计算，预处理 $O(m)$，查询 $O(m)$。
- 当模数 Mod 很小，套用 Lucas 定理，预处理 $O(Mod^2)$ ，查询 $O(\log_{Mod} n)$

$$\binom{n}{m} \equiv \binom{n \bmod p}{m \bmod p}\binom{\lfloor n/p \rfloor}{\lfloor m/p \rfloor} \pmod p$$

$\binom{n}{m} \bmod 2$ 等价于什么？

恒等变换

1. 对称性

$$\binom{n}{k} = \binom{n}{n-k}, \text{整数} n \geq 0, k \text{是整数}$$

恒等变换

2.二项式定理

$$\begin{aligned}(x+y)^n &= \overbrace{(x+y)(x+y)\ldots(x+y)}^{n\text{组}} \\ &= \sum_k \binom{n}{k} x^k y^{n-k}, \text{整数} n \geq 0 \text{或者} |\frac{x}{y}| \leq 1\end{aligned}$$

恒等变换

3. 吸收恒等式

$$k\binom{r}{k} = r\binom{r-1}{k-1}, \text{整数} k$$

推论

$$(r-k)\binom{r}{k} = r\binom{r-1}{k}, k\text{是整数}$$

恒等变换

小练习

冒险家协会有 n 个成员，现在需要任选一些人成立一个特别任务小队，并选出一位队长。共有多少种可能？

$1 \le n \le 10^9$

恒等变换

小练习

冒险家协会有 n 个成员，现在需要任选一些人成立一个特别任务小队，并选出一位队长。共有多少种可能？
$1 \le n \le 10^9$

$$\sum_{k \le n} k \binom{n}{k} = n \sum_{k \le n} \binom{n-1}{k-1} = n2^{n-1}$$

恒等变换

4.加法公式

$$\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}, k\text{是整数}$$

$$\left(f(u,v) = f(u-1,v) + f(u-1,v-1)\right)$$

恒等变换

小练习

给定序列 a ，现在需支持两种操作:

- 区间加
- 区间查询，选不超过 k 个值不大于 y 的数有多少种方案。

$1 \le n, Q \le 10^5$

考虑问题的简化版，假设已经知道区间内小于等于 y 的值的个数 m，答案是什么?

恒等变换

小练习

给定序列 a，现在需支持两种操作:

- 区间加
- 区间查询，选不超过 k 个值不大于 y 的数有多少种方案。

$1 \le n, Q \le 10^5$

考虑问题的简化版，假设已经知道区间内小于等于 y 的值的个数 m，答案是什么？
需要快速查询 $\sum_{i \le k} \binom{m}{i}$，即组合数一行的前缀和。

恒等变换

小练习

给定序列 a ，现在需支持两种操作:

- 区间加
- 区间查询，选不超过 k 个值不大于 y 的数有多少种方案。

$1 \le n, Q \le 10^5$

考虑问题的简化版，假设已经知道区间内小于等于 y 的值的个数 m，答案是什么？

需要快速查询 $\sum_{i \le k} \binom{m}{i}$，即组合数一行的前缀和。

考虑分块打表，令

$S(n, L, R) = \sum_{L \le i \le R} \binom{n}{i} = S(n-1, L-1, R-1) + S(n-1, L, R)$

单次查询 $O(\sqrt{n})$，预处理 $O(n\sqrt{n})$

恒等变换

5.平行求和法

$$\sum_{k\le n}\binom{r+k}{k}=\binom{r}{0}+\binom{r+1}{1}+\cdots+\binom{r+n}{n}$$
$$=\binom{r+n+1}{n},n\text{是整数}$$

杨辉三角斜线和

恒等变换

6. 上指标求和法

$$\sum_{0\le k\le n}\binom{k}{m}=\binom{0}{m}+\binom{1}{m}+\cdots+\binom{n}{m}$$
$$=\binom{n+1}{m+1}, \text{整数} m,n\ge 0$$

杨辉三角列和

恒等变换

小练习 CF938E

Q 次询问，固定 n，每次给出 x，求

$$\sum_{i \le x} i!(n-i-1)!\binom{x}{i}$$

$1 \le n, Q \le 10^5$

恒等变换

$$
\begin{aligned}
\sum_{i\le x} i!(n-i-1)!\binom{x}{i} &= \sum_{i\le x}\frac{x!(n-i-1)!}{(x-i)!} \\
&= x!(n-x-1)!\sum_{i\le x}\frac{(n-i-1)!}{(n-x-1)!(x-i)!} \\
&= x!(n-x-1)!\sum_{i\le x}\binom{n-1-i}{x-i} \\
&= x!(n-x-1)!\binom{n}{x} = \frac{n!}{n-x}
\end{aligned}
$$

恒等变换

7. 上指标反转

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, k\text{是整数}$$

恒等变换

7. 上指标反转

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, k\text{是整数}$$

小练习，化简

$$\sum_{k \le m} (-1)^k \binom{r}{k}$$

恒等变换

8. 三项式版恒等式

$$\binom{r}{m}\binom{m}{k}=\binom{r}{k}\binom{r-k}{m-k}, m, k\text{是整数}$$

恒等变换

8. 三项式版恒等式

$$\binom{r}{m}\binom{m}{k}=\binom{r}{k}\binom{r-k}{m-k}, m, k\text{是整数}$$

三项式系数

$$\binom{a+b+c}{a,b,c}=\frac{(a+b+c)!}{a!b!c!}=\binom{a+b+c}{b+c}\binom{b+c}{c}$$

恒等变换

8. 三项式版恒等式

$$\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}, m, k\text{是整数}$$

三项式系数

$$\binom{a+b+c}{a,b,c} = \frac{(a+b+c)!}{a!b!c!} = \binom{a+b+c}{b+c}\binom{b+c}{c}$$

多项式系数

$$\binom{a_1+a_2+\cdots+a_m}{a_1,a_2,\ldots,a_m} = \frac{(a_1+a_2+\cdots+a_m)!}{a_1!a_2!\ldots a_m!}$$

恒等变换

9. 范德蒙德卷积

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, n\text{是整数}$$

恒等变换

推论，通过运用上指标反转和对称性，可以得到

$$\sum_k \binom{r}{m+k}\binom{s}{n-k} = \binom{r+s}{m+n}, m, n\text{是整数}$$

$$\sum_k \binom{l}{m+k}\binom{s}{n+k} = \binom{l+s}{l-m+n}, \text{整数}l \geq 0, m, n\text{是整数}$$

$$\sum_k \binom{l}{m+k}\binom{s+k}{n}(-1)^k = (-1)^{l+m}\binom{s-m}{n-l}, \text{整数}l \geq 0, m, n\text{是整数}$$

$$\sum_{k\leq l} \binom{l-k}{m}\binom{s}{k-n}(-1)^k = (-1)^{l+m}\binom{s-m-1}{l-m-n}, \text{整数}l, m, n \geq 0$$

$$\sum_{-q\leq k\leq l} \binom{l-k}{m}\binom{q+k}{n} = \binom{l+q+1}{m+n+1}, \text{整数}m, n \geq 0, \text{整数}l+q \geq 0$$

恒等变换

小练习

你的衣服有两个口袋，左边口袋装了 x 个普通硬币和 y 个量子硬币，右边口袋装了 a 个普通硬币和 b 个量子硬币。普通硬币面值为 1 元，而量子硬币的面值等概率的为 0 元或者 1 元，只有当你拿出来观测的时候，才会坍缩为某一具体面值。现在你决定把两个口袋里的硬币全部拿出来观测。你想知道左口袋面值和大于等于右口袋面值和的概率是多少。Q 次询问。

$1 \le Q, x, y, a, b \le 10^5$

恒等变换

令 k,i 分别表示左/右口袋量子硬币 1 的个数，

$$
\begin{aligned}
P =& 2^{-(y+b)}\sum_{i}\binom{b}{i}\sum_{k>i}\binom{y}{k+a-x}\\
=& 2^{-(y+b)}\sum_{i}\binom{b}{i}\sum_{k>0}\binom{y}{k+i+a-x}\\
=& 2^{-(y+b)}\sum_{k>0}\sum_{i}\binom{b}{b-i}\binom{y}{k+i+a-x}\\
=& 2^{-(y+b)}\sum_{k>b+a-x}\binom{b+y}{k}
\end{aligned}
$$

分块打表，复杂度 $O(Q\sqrt{b+y})$

错位排列

问题引入

有多少个长为 n 的排列 P 满足 $P_i \neq i$ ？

这样的数 D_n 一般称为错位排列数。容易通过组合意义证明

$$D_n = (n-1)(D_{n-1} + D_{n-2})$$

初值 $D_1 = 1, D_2 = 1$，可以线性预处理。一般为某个计数问题的子问题。

卡特兰数

问题引入

- 一个栈（无穷大）的进栈序列为 $1, 2, 3, \cdots, n$ 有多少个不同的出栈序列？
- n 对括号最多能生成多少个合法的括号串。

这样的数 C_n 一般称作卡特兰数。一般为某个计数问题的子问题。可以通过折线法证明

$$C_n = \binom{2n}{n} - \binom{2n}{n-1}$$

折线法的应用较为广泛，可以通过 "JLOI2015 骗我呢" 进行学习。

斯特林数

第二类斯特林数

$\left\{ n \atop k \right\}$，也可记做 $S(n,k)$，表示将 n 个不同的元素，划分为 k 个互不区分的非空子集的方案数。

加法公式

$$\left\{ n \atop k \right\} = \left\{ n-1 \atop k-1 \right\} + k \left\{ n-1 \atop k \right\}$$

边界 $\left\{ n \atop 0 \right\} = [n=0]$，于是可以 $O(n^2)$ 预处理。

斯特林数

幂函数展开

$$i^K = \sum_{j=0}^{K} \left\{ {K \atop j} \right\} i^{\underline{j}}$$

当题目所求和式带 i^K 项时，可以考虑将其展开为斯特林数。

斯特林数

小练习

有 m 张牌，其中有一张是王牌。将这些牌均匀随机打乱 n 次，设有 X 次第一张为王牌，求 X^k 的期望值。
答案对 998244353 取模。
$1 \le n, m < 998244353, 1 \le K \le 5 \times 10^3$

斯特林数

小练习

有 m 张牌，其中有一张是王牌。将这些牌均匀随机打乱 n 次，设有 X 次第一张为王牌，求 X^k 的期望值。
答案对 998244353 取模。
$1 \leq n, m < 998244353, 1 \leq K \leq 5 \times 10^3$

对于每次随机打乱，第一张为王牌的概率均为 $\frac{1}{m}$，记为 p，那么

$$E(X^K) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} i^K$$

$$\sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \sum_{j=0}^{K} \left\{ {K \atop j} \right\} i^{\underline{j}} = (*)$$

斯特林数

$$
\begin{aligned}
(*) &= \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \sum_{j=0}^{K} \left\{ {K \atop j} \right\} i^{\underline{j}} \\
&= \sum_{j=0}^{K} \left\{ {K \atop j} \right\} \sum_{i=j}^{n} p^i (1-p)^{n-i} \binom{n}{i} \binom{i}{j} j! \\
&= \cdots \\
&= \sum_{j=0}^{K} \left\{ {K \atop j} \right\} n^{\underline{j}} p^j
\end{aligned}
$$

复杂度 $O(K^2)$

容斥原理

问题引入

假设班里有 10 个学生喜欢数学，15 个学生喜欢语文，21 个学生喜欢编程，班里至少喜欢一门学科的有多少个学生呢？

容斥原理

问题引入

假设班里有 10 个学生喜欢数学，15 个学生喜欢语文，21 个学生喜欢编程，班里至少喜欢一门学科的有多少个学生呢？

把喜欢语文、数学、编程的学生集合分别用 A, B, C 表示，那么

$$\begin{aligned}|A \cup B \cup C| = &+ |A| + |B| + |C| \\ &- |A \cap B| - |B \cap C| - |C \cap A| \\ &+ |A \cap B \cap C|\end{aligned}$$

把上述问题推广到一般情况，就是容斥原理。
（通过观察得出式子，由归纳法证明）

容斥原理

容斥原理常用于集合的计数问题，而对于两个集合的函数 $f(S), g(S)$，若

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

那么就有

$$g(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} f(T)$$

容斥原理

小练习

给出 n 个数，选出其中 k 个，使得它们的或恰好是 r，求方案数。

$1 \leq n \leq 10^6, 1 \leq r < 2^{20}$

容斥原理

令 cnt_S 表示有多少个数是 S 的子集。
令 $g(S)$ 表示选 k 个数恰好或起来是 S 的方案数。
令 $f(S)$ 表示选 k 个数或起来得到的数 T 是 S 子集的方案数，那么

$$f(S) = \sum_{T \subseteq S} g(T) = \binom{cnt_S}{k}$$

应用容斥原理得到

$$g(r) = \sum_{T \subseteq r} (-1)^{|r|-|T|} \binom{cnt_T}{k}$$

预处理阶乘及逆元，通过高维前缀和预处理 cnt_S，枚举 r 子集求出 $g(r)$，复杂度 $O(n + r\log r + r)$

反演

对于某些题，我们可能会得到形如下式的式子，对任何 n，满足

$$F(n) = \sum_{i=0}^{n} a_{n,i} \times G(i)$$

其中 $F(n)$ 可以通过另外一个表达式轻松算出，但是题目要算的是某个 $G(n)$。
对于用 F 来表示 G 的过程，数学上称为反演，有

$$G(n) = \sum_{i=0}^{n} b_{n,i} \times F(i)$$

其中 a 和 b 是相关的，满足

$$\sum_{i=j}^{n} a_{n,i} \times b_{i,j} = [j = n]$$

二项式反演

如果有

$$g_n = \sum_{i=0}^{n} \binom{n}{i} f_i$$

那么

$$f_n = \sum_{i=0}^{n} \binom{n}{i} (-1)^{n-i} g_i$$

上面这对式子被称为二项式反演。
变式：

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

二项式反演

小练习

有 m 种物品，每种有 a_i 个，现在把这些物品分给 n 个人，每个人至少拿到一个，求有多少种分法。但对 10^9+7 取模。
$1 \le n, m \le 10^3, 1 \le \sum a_i \le 10^6$

二项式反演

小练习

有 m 种物品，每种有 a_i 个，现在把这些物品分给 n 个人，每个人至少拿到一个，求有多少种分法。但对 10^9+7 取模。

$1 \le n, m \le 10^3, 1 \le \sum a_i \le 10^6$

如果没有每个人都必须分到物品的限制的话，考虑对每个物品用插板法，表示一种物品的分配方法，每个物品的方案数乘起来就是总的方案，即

$$\prod_{i=1}^{m} \binom{a_i+n-1}{n-1}$$

二项式反演

令 g_k 表示恰好有 k 个人没有得到物品的方案数，那么所求即 g_0。
令 f_k 表示强制不给其中 k 个人分物品，并将物品分给剩下的人的有重复的总方案数。那么先选定 k 个人，再分配，方案是

$$f_k = \binom{n}{k} \prod_{j=1}^{m} \binom{a_j + n - k - 1}{n - k - 1}$$

这样算出的 f_k 是有重复方案的，并且并不保证方案中恰好只有 k 个人没分到物品。具体而言，对于一个最终恰好有 i 个人没有分到物品的方案，在 f_k 中被重复计算了 $\binom{i}{k}$ 次，于是

$$f_k = \sum_{i=k}^{n} \binom{i}{k} g_i$$

二项式反演

反演得到

$$g_k = \sum_{i=k}^{n} (-1)^{i-k} f_i$$

$$= \sum_{i=k}^{n} (-1)^{i-k} \binom{n}{i} \prod_{j=1}^{m} \binom{a_j + n - i - 1}{n - i - 1}$$

答案是

$$g_0 = \sum_{i=0}^{n} (-1)^{i} \binom{n}{i} \prod_{j=1}^{m} \binom{a_j + n - i - 1}{n - i - 1}$$

预处理阶乘及其逆元，复杂度 $O(nm + \sum a_i)$

Min-Max 反演

对于满足全序关系并且其中元素满足可加减性的序列 $\{x_i\}$，设其长度为 n，并设 $S=\{1,2,3,\cdots,n\}$，则有：

$$\max_{i\in S} x_i = \sum_{T\subseteq S} (-1)^{|T|-1} \min_{j\in T} x_j$$

$$\min_{i\in S} x_i = \sum_{T\subseteq S} (-1)^{|T|-1} \max_{j\in T} x_j$$

Min-Max 反演

小练习

初始一个 $x=0$，每一秒钟选择一个值 $y\in[0,2^n)$，选到 i 的概率为 p_i，然后进行运算 $x|=y$，其中 | 表示按位或运算。求 x 变成 2^n-1 的期望时间。

$0\le p_i\le 1,\sum p_i=1,1\le n\le 20$

题目可以翻译成，每次会选择一个集合，求使得所有元素都被选到过的期望时间。

Min-Max 反演

小练习

初始一个 $x=0$，每一秒钟选择一个值 $y\in[0,2^n)$，选到 i 的概率为 p_i，然后进行运算 $x|=y$，其中 $|$ 表示按位或运算。求 x 变成 2^n-1 的期望时间。

$0\le p_i\le 1,\sum p_i=1,1\le n\le 20$

题目可以翻译成，每次会选择一个集合，求使得所有元素都被选到过的期望时间。

定义 $\min(S)$ 表示第一次选中了集合 S 中的元素的时间，$\max(S)$ 表示集合 S 中所有元素都被选到过的时间。那么答案是 $E(\max(U))$，U 是全集。由于期望的线性性，可得

$$E(\max(S))=\sum_{T\subseteq S}(-1)^{|T|-1}E(\min(T))$$

Min-Max 反演

考虑如何计算 $\min(S)$，令 $p(S)$ 表示选到 S 中的元素的概率，那么

$$p(S) = \sum_{S\cap T\neq\emptyset} p_T = \sum_{T\subseteq(U-S)} p_T$$

令

$$\begin{aligned}P(x) =& p(S)x + (1-p(S))p(S)x^2 + (1-p(S))^2 p(S)x^3 + \cdots \\ =& \frac{p(S)x}{1-(1-p(S))x}\end{aligned}$$

那么

$$E(\min(S)) = \frac{\mathrm{d}}{\mathrm{d}x}P(x)\Big|_{x=1} = \frac{1}{p(S)}$$

高位前缀和预处理所有 $p(S)$，枚举子集求答案，复杂度 $O(n2^n)$