计数与期望问题选讲

陈立杰

清华大学

May 11, 2017

K Perm Counting

- 给你两个数 N 和 K。
- 问有多少个 $1 \ldots N$ 的排列 $\{a_i\}$ 满足 $|a_i - i| \neq K$。
- $N \leq 2000$。

K Perm Counting ：讨论

- 考虑使用容斥原理来计算有 k 个 a_i 不合法的方案的数量。
- 简单的 dp 就可以了。

Dictionary

- 给你 n 个由小写字母和? 号组成的字符串 $s_1, s_2, \ldots, s_n$。问有多少种将? 号变成小写字母的方法，使得 $s_1 < s_2 < \ldots < s_n$。
- $n \leq 20$，字符串长度都 ≤ 50。

Dictionary ：讨论

- 考虑一位一位地确定所有串，那么就被分成了很多子问题，就可以做了。

Leftmost Ball

- 一个人有 $N \cdot K$ 个球，N 种颜色每种有 K 个，颜色分别记为 1 到 N。
- 他把所有球从左到右派排起来，对每个颜色把它出现的最左边的球给染成 0 号颜色，这样就得到一个由 0 到 N 的颜色的球组成的序列。
- 问有多少序列是可能的，求答案对 $10^9 + 7$ 取模。
- $1 \leq N, K \leq 2000$。

Leftmost Ball ：讨论

- 注意到，一个序列合法当且仅当每个 0 都能往后和一个某颜色的最靠左的球匹配。
- 那么我们不妨考虑一个由白球和某个颜色的最靠左的球组成的序列，并考虑对根据这个序列进行计算。
- 不妨考虑从右往左来 dp，假设目前有 x 个可以匹配的某颜色的最靠左的球。
- 假设当前想要插入一个白球，那么就将 x-1，插入一个白球。
- 否则我们插入一个新的颜色的球，并且在它后面选择一些位置插入 $K-2$ 个与它同色的球。

Tournament

- 问所有 n 个点的竞赛图的强联通分量的数量的和。
- 竞赛图就是对于任意两对不同的点 a 和 b，或者 a 到 b 有有向边，或者 b 到 a 有有向边。
- $n \leq 10^5$。
- 输出答案对 $10^9 + 7$ 取模。

Tournament : 讨论

- 考虑一个竞赛图，将它的强连通分量都缩成点，缩成点之后形成的 DAG 就是一条全序的链。
- 那么就能注意到，强连通分量的个数就是竞赛图中“割”的数量加上一。
- 一个割就是一个点集 S，使得 S 中所有的点都往 $[n] \setminus S$ 中的点连边。
- 那我们不妨求所有竞赛图割的数量的和，我们可以枚举割的大小，然后计算有多少竞赛图的割是这个。

Dancing IOI Team

- 有两个序列 $a_1 \geq a_2 \geq a_3 \ldots \geq a_n$，和 $b_1, b_2, \ldots, b_n$。和一个整数 k。
- 考虑所有 $n!$ 种将 $\{a_i\}$ 和 $\{b_i\}$ 两两配对的方式，不妨假设 a_i 与 b_{p_i} 配对。
- 问有多少种配对的方式，满足对于所有的 $1 \leq i \leq k$ 和 $k+1 \leq j \leq n$，有

$$a_i + b_{p_i} \geq a_j + b_{p_j}.$$

- 输出答案对 10^9+7 取模。
- $n \leq 100$。$0 \leq a_i, b_i \leq 1000$。$a_i$ 两两不同，b_i 两两不同。

- 不妨令 $S_i = a_i + b_{p_i}$，原问题就是求有多少排列 p_i 使得 S_1 到 S_k 都大于等于 S_{k+1} 到 S_n。
- 不妨先枚举 S_1 到 S_k 的最小值 T，那么现在问题就变成了，问有多少个排列 $\{p_i\}$ 满足
  1. 对于 $i \in [1 \ldots k]$，有 $a_i + b_{p_i} \geq T$。
  2. 对于 $i \in [k+1 \ldots n]$，有 $a_i + b_{p_i} \leq T$。
  3. 存在一个 $i \in [1 \ldots k]$，有 $a_i + b_{p_i} = T$。
- 不妨令以上值为 $f(T)$。

- 第三个条件可以较为简单地解决，不妨令 $g(A, B)$ 为满足下列要求的排列的数量
  1. 对于 $i \in [1 \dots k]$，有 $a_i + b_{p_i} \geq A$。
  2. 对于 $i \in [k+1 \dots n]$，有 $a_i + b_{p_i} \leq B$。
- 那么容易看出 $f(T) = g(T, T) - g(T+1, T)$。

Dancing IOI Team ： 讨论

- 假设目前考虑到第 i 个 b_i，我们同时记录一下之前有多少个 b_i 被配到了组 A，假设有 j 个。
- 假设 b_i 被分到了组 A，假设组 A 中有 x 个可以与 b_i 匹配的 a_i，那么 b_i 就可以选择 $x-j$ 个可能的 a_i。
- 假设 b_i 被分到了组 B，这个时候我们把它标记成待分配的。
- 注意到一个组 B 的 a_i 只能和 $b_j \le B - a_i$ 的 b_j 匹配，我们不妨将 B 组的 a_i 以 $(B - a_i)$ 为关键字和所有 b_j 一起排序。
- 碰到一个 B 组的 a_i 的时候，我们选择一个待分配的 b_i 分配给它再乘上方案数就可以了。

Dancing IOI Team ： 讨论

- 假设目前考虑到第 i 个 b_i，我们同时记录一下之前有多少个 b_i 被配到了组 A，假设有 j 个。
- 假设 b_i 被分到了组 A，假设组 A 中有 x 个可以与 b_i 匹配的 a_i，那么 b_i 就可以选择 $x - j$ 个可能的 a_i。
- 假设 b_i 被分到了组 B，这个时候我们把它标记成待分配的。
- 注意到一个组 B 的 a_i 只能和 $b_j \leq B - a_i$ 的 b_j 匹配，我们不妨将 B 组的 a_i 以 $(B - a_i)$ 为关键字和所有 b_j 一起排序。
- 碰到一个 B 组的 a_i 的时候，我们选择一个待分配的 b_i 分配给它再乘上方案数就可以了。

Game with Marbles

- 袋子里有 r 个红球，g 个绿球和 b 个蓝球。每回合你等概率随机取出袋子中的一个球。
- 如果是红球，就扔掉。
- 如果是绿球或者蓝球，就把球放回袋子里。
- 问到当进行到刚好拿出过 k 次蓝球的时候，期望经过了多少回合？
- $1 \leq \mathsf{r}, \mathsf{g}, \mathsf{b}, \mathsf{k} \leq 10^9$。

Game with Marbles

- 不妨分别计算结束之前，期望拿出多少个各种颜色的球。
- 显然期望拿出正好 k 个蓝球。
- 考虑期望拿出多少个绿色的球，注意到这个时候可以完全无数红球的存在，容易看出期望拿出 $\frac{k \cdot g}{b}$ 个绿球。
- 考虑期望拿出多少个红色的球，不妨单独考虑每个红色的球，计算它被拿出来的概率，注意到这个时候我们可以无视其它所有的红色球和绿色球的存在，容易看出单个红色球被拿出来的概率恰好是 $1 - (\frac{b}{b+1})^K$。
- 简单求和就可以计算答案了。

Shortest Path System

- 考虑一个 n 个点的无向图，边带正整数权值，并且其每对点间的最短路径都是唯一的。
- 我们定义其最短路径系统为所有 $\binom{n}{2}$ 对最短路径的集合。
- 如果可以通过给点重标号来将一个最短路径系统变为另一个，我们就称这两个最短路径系统是本质相同的。
- n 个点的图有多少种本质不同的最短路径系统？
- $2 \le n \le 6$。

Shortest Path System ： 讨论

- 注意到范围很小，所以可以打表。
- 注意到最短路径系统必然满足，假如 p 是一条 a 到 b 的最短路，那么 p 的任何子路径也是对应的点之间的最短路。
- 不妨通过爆搜得出满足上面条件的最短路径系统的集合。
- 但上述只是必要条件，接下来可以用线性规划来验证是否存在边权的值满足这个最短路径系统。

Izhevsk Training Camp (Simplified)

- 给你三个排列 1 到 n 的排列 $\{a_i\}, \{b_i\}, \{c_i\}$。
- 问有多少对 (x, y) 满足 $a_x < a_y$，$b_x < b_y$，$c_x < c_y$。
- $n \leq 5 \cdot 10^6$。

Izhevsk Training Camp (Simplified)

- n 很大，爆搜是不可取的。
- 考虑任意一对数 (x, y)，可以看出要么有一个 a, b, c 都比对方大，要么有一个 a, b, c 中有恰好两个比对方大。
- 不妨令前一种对的集合 A，后一种的集合为 B，我们有 $A + B = \binom{n}{2}$。
- 考虑 $\{a_i\}$ 和 b_i，我们计算出有多少对数 (x, y) 满足其中有一个 a 和 b 都比对方大。$\{b_i\}$ 和 $\{c_i\}$，$\{c_i\}$ 和 $\{a_i\}$ 也同理。然后再将三者加起来，得到值 X。
- 注意到，对于 A 中的数对，上述过程恰好将它计算了 3 次，对于 B 中的数对则恰好计算了 1 次，所有我们又有 $3|A| + |B| = X$。
- 于是就能反推出 $|A|$，是我们想要计算的值。

Petya and Arrays

- 有多少个长度为 N 的整数数组满足以下条件?
  - 每个数都在 1 到 P 之间。
  - A 没有出现在数组里。
  - 没有一个连续的子数组的和被 P 整数。
- 输出答案对 $10^9 + 7$ 取模。
- $\mathsf{N} \le 85$, $\mathsf{A}, \mathsf{P} \le 10^9$。

Petya and Arrays ：讨论

- 注意到前缀和对 P 取模之后唯一确定了这个序列，不妨令 S_i 表示这个整数数组前 i 个的值，同时有 $S_0 = 0$。
- 那我们需要满足的条件是
  1. 所有 S_i 都两两不同。
  2. 不存在相邻的两个 S_i 和 S_{i+1} 差是 A。

Petya and Arrays ：讨论

- 首先注意到如果 x 往 $x + A \pmod{P}$ 连边的话，就形成了一些环，0 到 $P - 1$ 的数就形成了一些环。
- 我们不妨考虑一个环一个环的往当前序列插入数字，我们需要知道之前的序列中有多少个相邻的位置差 A，这些位置需要以后被新插进的数分开来。
- 假设我们一个环一个环的考虑，假设我们已经考虑了前 i 个环，对于一个新的环，我们考虑在它里面选择几段数，插入哪些位置等等来进行动态规划。

City United

- 给你一个大小为 n 的图，问有多少个联通的点集?
- 一个点集 S 被称为联通的，当且仅当对于其中任意两个点 a,b，存在只经过 S 中的点的原图中的路径连接它们。
- 两个点 a 和 b 只有在 $1 \le |a - b| \le 13$ 的时候才会有边。
- 输出答案对 2 取模。
- $n \le 50$。

City United

- 考虑这样一个问题，给一个图，将所有点黑白染色，满足每个联通块的颜色要一样。
- 对于图 S，如果其联通块数量为 k，那么显然染色方案数就为 2^k。
- 不妨考虑下面这个问题，求一个图的所有子集对应的图的染色方案数的和，对 4 取模。
- 这样的话，注意到如果不连通的子集，由于联通块数量至少是 2 就被忽略了，否则就加了 2，于是就能解决原问题了。
- 回到原问题，注意到染色方案数等价于给所有子集中的点黑白染色，并且满足两个有边的点颜色相同，我们只需要记录最后 13 个点中每个点是选择了还是没有选择，选择了的话是黑色还是白色就可以了。

Connected Spanning Subgraph

- 给你一个大小为 n，m 条边的连通图，问有多少个连通的边集?
- 一个边集和 S 被称为连通的，当且仅当对于原图中任意的两个点，存在只经过 S 中的边的路径连接它们。
- 输出答案对 2 取模。
- $n \le 10^5, m \le 2 \cdot 10^5$。

Connected Spanning Subgraph

- 如果原图不连通，那么显然答案为 0。
- 考虑一种黑色和白色的染色方案，注意到首先黑色和白色可以互换，因此每个方案都被算了两次。其次如果存在一条边在同色点之间，那么这条边可选可不选，于是就又乘了 2。
- 所以在计算过程中，由于我们需要方案数量对 4 取模，如果黑色和白色的染法没有将原图分成一个二分图，那么就被自动忽略了。
- 所以可以看出，答案是 1 当且仅当原图是一个二分图。

BBQ Hard

- 有 n 个数对 (A_i, B_i)，要求

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\binom{A_i + B_i + A_j + B_j}{A_i + A_j}$$

- 输出答案对 $10^9 + 7$ 取模。
- $n \le 2 \cdot 10^5$，$A_i, B_i \le 2000$。

BBQ Hard

- 把每个点 (A_i, B_j) 看成平面上的两个点 $X_i = (-A_i, -B_i)$ 和 $Y_i = (A_i, B_i)$。
- 假设有一个人，只能往上或者往右走。
- 那么可以看出 $\binom{A_i+B_i+A_j+B_j}{A_i+A_j}$ 是从点 X_i 走到 Y_j 的方案数。
- 不妨考虑一个 dp，令 $dp_{x,y}$ 表示所有 X 中且在 (x, y) 左下方的 $\{X\}$ 中的点，从它们走到 (x, y) 的方案数的和。
- 容易看出 dp 可以在 $O(N^2)$ 的时间求解，答案也就可以计算出来了。

Number of Solutions

- 有 n 个布尔变量 $x_1, x_2, \ldots, x_n$。
- 再给你一个 2SAT 形式的约束。
- 问你有多少组解。
- $n \leq 50$。

Number of Solutions

- 如果两个变量之间有约束我们就在它们之间连一条边。
- 如果当前图最大度数是 2，那么可以注意到它就是一堆环和链，这个时候我们直接 dp 就能求解了。
- 否则假设最大度数是 3(3 以上类似)，不妨令 x_i 为一个度数为 3 的变量，假设 $x_i = 0$ 与 a 个变量的某个赋值冲突，$x_i = 1$ 与 b 个变量的某个赋值冲突，那么有 $a + b = 3$。
- 那么注意到在两种情况下剩下的图需要考虑的点的个数分别是 $n - 1 - a$ 和 $n - 1 - b$。进行简单的归纳可以发现可以通过 $n = 50$ 的情况。

Mixed Drinks

- 给你三个 1 到 n 的排列 $\{a_i\}$，$\{b_i\}$，$\{c_i\}$。
- 称三元组 (x, y, z) 为合法的，当且仅当存在一个下标的集合 $S \subseteq [n]$ 满足

$$(x, y, z) = (\max_{i \in S} a_i, \max_{i \in S} b_i, \max_{i \in S} c_i).$$

- 问合法三元组的数量。
- $n \leq 10^5$。

Mixed Drinks ：讨论

- 考虑一个合法三元组对应的下标的集合 S，不妨只保留那些在位置上达到 S 中最大值的下标。那么可以看出 $|S| \leq 3$，称其为最简下标集合。
- 注意到合法三元组和最简下标集合之间存在一一对应的关系，所以只需要计算最简下标集合的数量。
- 大小为 1 的所有下标集合都是最简下标集合。
- 大小为 2 的情况的话，考虑两个下标 $\{x, y\}$，只要不是其中一个 a, b, c 都比另一个大就合法。

Mixed Drinks ：讨论

- 最后考虑大小为 3 的情况。注意到这意味着下标 $\{x, y, z\}$，每个都在 a, b, c 中的一项上是三个中最大的。
- 直接计算比较困难，不妨考虑计算不合法的大小为 3 的下标集合。
- 第一种情况是存在一个在 a, b, c 中都是最大的。这种情况可以简单的在 $O(n\log^2 n)$ 的时间内计算，不妨令这种情况的下标集合的集合为 A。
- 第二种情况是一个在 a, b, c 中的两项上最大，另一个在 a, b, c 中的另一项最大，不妨令这种情况的下标集合的集合为 B。
- 考虑枚举 a, b, c 中的两个，然后计算有多少个三元组满足其中存在一个在 a, b 中都是最大的，把算出的三个值加起来得到 X。
- 那么可以注意到，计算 X 的过程中, B 中的三元组恰好被算了一次，而 A 中的三元组则被计算了 3 次，所以我们有 $|A| \cdot 3 + |B| = X$，然后就能解出 $|B|$ 了。

Mixed Drinks：讨论

- 考虑一个合法三元组对应的下标的集合 S，不妨只保留那些在位置上达到 S 中最大值的下标。那么可以看出 $|S| \leq 3$，称其为最简下标集合。
- 注意到合法三元组和最简下标集合之间存在一一对应的关系，所以只需要计算最简下标集合的数量。
- 大小为 1 的所有下标集合都是最简下标集合。
- 大小为 2 的情况的话，考虑两个下标 $\{x, y\}$，只要不是其中一个 a, b, c 都比另一个大就合法。