中国计算机学会
CHINA COMPUTER FEDERATION

概率论与概率方法

北京大学 管晏如

2024 年 5 月 17 日

引言

“概率”在 OI 中有哪些出现形式?

引言

“概率” 在 OI 中有哪些出现形式?

- 数据随机;

引言

“概率”在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;

引言

“概率”在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;

引言

“概率”在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;
- ……

引言

“概率”在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;
- ……
- 限于时间，今天主要会讨论后两个。

引言

"概率" 在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;
- ……
- 限于时间，今天主要会讨论后两个。

"概率方法" 又是什么?

引言

"概率"在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;
- ……
- 限于时间，今天主要会讨论后两个。

"概率方法"又是什么?

- 【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立？正确的概率足够高即可。

引言

"概率"在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;
- ……
- 限于时间，今天主要会讨论后两个。

"概率方法"又是什么?

- 【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立? 正确的概率足够高即可。
- 【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。(为什么存在? 找的方法有哪些?)

引言

“概率”在 OI 中有哪些出现形式?

- 数据随机;
- 求某个概率期望;
- 在算法中引入随机化，提升效率或准确率;
- ……
- 限于时间，今天主要会讨论后两个。

“概率方法”又是什么?

- 【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立? 正确的概率足够高即可。
- 【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。(为什么存在? 找的方法有哪些?)
- ……

目录

概率论基础

离散型随机变量

连续型随机变量

概率方法

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$；

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$;
- 抛 n 次硬币，$\Omega = \{H, T\}^n$;

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$；
- 抛 n 次硬币，$\Omega = \{H, T\}^n$；
- 掷骰子，$\Omega = \{1, 2, \ldots, 6\}$。

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$；
- 抛 n 次硬币，$\Omega = \{H, T\}^n$；
- 掷骰子，$\Omega = \{1, 2, \ldots, 6\}$。

事件：可认为是 Ω 的子集，通常用 $A, B, \ldots$ 表示。

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$；
- 抛 n 次硬币，$\Omega = \{H, T\}^n$；
- 掷骰子，$\Omega = \{1, 2, \ldots, 6\}$。

事件：可认为是 Ω 的子集，通常用 $A, B, \ldots$ 表示。

- 掷骰子得到的点数 ≤ 3，就是一个事件。

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$；
- 抛 n 次硬币，$\Omega = \{H, T\}^n$；
- 掷骰子，$\Omega = \{1, 2, \ldots, 6\}$。

事件：可认为是 Ω 的子集，通常用 $A, B, \ldots$ 表示。

- 掷骰子得到的点数 ≤ 3，就是一个事件。

概率空间：为 Ω 中的每个元素赋予一个发生的概率，用 $P : \Omega \to [0, 1]$ 表示。应有所有元素发生的概率之和 $= 1$。

概率论基础

样本空间：随机试验的所有结果构成的集合 Ω。

- 抛硬币，$\Omega = \{H, T\}$；
- 抛 n 次硬币，$\Omega = \{H, T\}^n$；
- 掷骰子，$\Omega = \{1, 2, \ldots, 6\}$。

事件：可认为是 Ω 的子集，通常用 $A, B, \ldots$ 表示。

- 掷骰子得到的点数 ≤ 3，就是一个事件。

概率空间：为 Ω 中的每个元素赋予一个发生的概率，用 $P : \Omega \to [0, 1]$ 表示。应有所有元素发生的概率之和 $= 1$。那么一个事件发生的概率 $P(A) = \sum_{a \in A} P(a)$。

概率论基础

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。
或者说，在 Ω 有限的情况之下，使用 $\sum$ 求和是合理的……

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。
或者说，在 Ω 有限的情况之下，使用 $\sum$ 求和是合理的……但如果无限呢？

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。
或者说，在 Ω 有限的情况之下，使用 $\sum$ 求和是合理的……但如果无限呢？
对于只有有限个（准确来说是可数个）a 满足 $P(x = a) > 0$ 的情形，我们称 X 是**离散型随机变量**。

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。
或者说，在 Ω 有限的情况之下，使用 $\sum$ 求和是合理的……但如果无限呢？
对于只有有限个（准确来说是可数个）a 满足 $P(x = a) > 0$ 的情形，我们称 X 是**离散型随机变量**。
否则，我们称其为**连续型随机变量**。

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X]=\sum P(X=a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。
或者说，在 Ω 有限的情况之下，使用 $\sum$ 求和是合理的……但如果无限呢？
对于只有有限个（准确来说是可数个）a 满足 $P(x=a)>0$ 的情形，我们称 X 是**离散型随机变量。**
否则，我们称其为**连续型随机变量。**

- X 等概率从 $1,2,\ldots,n$ 中随机选取——离散；
- X 在 $[0,1]$ 内等概率随机选取——连续。

概率论基础

随机变量：X 从一个概率空间中抽取，以 $P(a)$ 的概率取值为 a
期望：$E[X] = \sum P(X = a)a$（在有加法的前提之下）
注意！这里的写法并不是严谨的。
或者说，在 Ω 有限的情况之下，使用 $\sum$ 求和是合理的……但如果无限呢？
对于只有有限个（准确来说是可数个）a 满足 $P(x = a) > 0$ 的情形，我们称 X 是**离散型随机变量**。
否则，我们称其为**连续型随机变量**。

- X 等概率从 $1, 2, ..., n$ 中随机选取——离散；
- X 在 $[0, 1]$ 内等概率随机选取——连续。

在进一步探讨随机变量有关概念之前，我们先在离散的意义下，使用概率和期望的基本概念看几道例题。

AGC060C

题目大意

- 考虑一个排列 $P_1, P_2, \ldots, P_{2^n-1}$，其被称作“好的”，当且仅当 $P_i < P_{2i}, P_i < P_{2i+1}$。
- 给定正整数 A, B，令 $U = 2^A, V = 2^{B+1}-1$。
- 求在所有“好的”排列中随机选择，满足 $P_U < P_V$ 的概率。
- 答案对 998244353 取模。
- 【数据范围】$1 \le n \le 5000, 1 \le A, B < n$

AGC060C

- U, V 分别为第 A, B 层最左和最后的点，显然它们的 LCA 恰为 1。

AGC060C

- U, V 分别为第 A, B 层最左和最后的点，显然它们的 LCA 恰为 1。
- 只考虑 1 往下，最左和最右的两条链。不包括 1，各恰好 $n-1$ 个点。

AGC060C

- U, V 分别为第 A, B 层最左和最后的点，显然它们的 LCA 恰为 1。
- 只考虑 1 往下，最左和最右的两条链。不包括 1，各恰好 $n-1$ 个点。
- 模拟拓扑序的过程，即只关心这 $2(n-1)$ 个点移除的相对顺序，并要求 U 在 V 之前被移除。

AGC060C

- U, V 分别为第 A, B 层最左和最后的点，显然它们的 LCA 恰为 1。
- 只考虑 1 往下，最左和最右的两条链。不包括 1，各恰好 $n-1$ 个点。
- 模拟拓扑序的过程，即只关心这 $2(n-1)$ 个点移除的相对顺序，并要求 U 在 V 之前被移除。
- 设 $dp[i][j]$ 表示已移除左右分别 i, j 个点。

AGC060C

- U, V 分别为第 A, B 层最左和最后的点，显然它们的 LCA 恰为 1。
- 只考虑 1 往下，最左和最右的两条链。不包括 1，各恰好 $n-1$ 个点。
- 模拟拓扑序的过程，即只关心这 $2(n-1)$ 个点移除的相对顺序，并要求 U 在 V 之前被移除。
- 设 $dp[i][j]$ 表示已移除左右分别 i, j 个点。
- 注意此时我们多钦定了一些大小关系，可以把 $2(n-1)$ 个点之间的连边，改为钦定的一条链，那么这样得到的仍然是一棵树。

AGC060C

- U, V 分别为第 A, B 层最左和最后的点，显然它们的 LCA 恰为 1。
- 只考虑 1 往下，最左和最右的两条链。不包括 1，各恰好 $n-1$ 个点。
- 模拟拓扑序的过程，即只关心这 $2(n-1)$ 个点移除的相对顺序，并要求 U 在 V 之前被移除。
- 设 $dp[i][j]$ 表示已移除左右分别 i, j 个点。
- 注意此时我们多钦定了一些大小关系，可以把 $2(n-1)$ 个点之间的连边，改为钦定的一条链，那么这样得到的仍然是一棵树。
- 现在，需要连同中间挂着的子树，一起来计算拓扑序的概率。

AGC060C

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \to$ 左 $i+1 \to \ldots \to$ 右 j

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \to$ 左 $i+1 \to \ldots \to$ 右 j
- 原本左 i 的贡献是 $\frac{1}{2^{n-i}-1}$，现在的贡献是 $\frac{1}{2^{n-i}-1+2^{n-j}-1}$

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \rightarrow$ 左 $i+1 \rightarrow \ldots \rightarrow$ 右 j
- 原本左 i 的贡献是 $\frac{1}{2^{n-i}-1}$，现在的贡献是 $\frac{1}{2^{n-i}-1+2^{n-j}-1}$
- 一乘一除即可。

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \to$ 左 $i+1 \to \dots \to$ 右 j
- 原本左 i 的贡献是 $\frac{1}{2^{n-i}-1}$，现在的贡献是 $\frac{1}{2^{n-i}-1+2^{n-j}-1}$
- 一乘一除即可。
- 对于右边也是同理。

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \to$ 左 $i+1 \to \ldots \to$ 右 j
- 原本左 i 的贡献是 $\frac{1}{2^{n-i-1}}$，现在的贡献是 $\frac{1}{2^{n-i-1}+2^{n-j-1}}$
- 一乘一除即可。
- 对于右边也是同理。
- 转移：直接枚举下一个填左还是右，同时算上它的贡献。

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \to$ 左 $i+1 \to \dots \to$ 右 j
- 原本左 i 的贡献是 $\frac{1}{2^{n-i}-1}$，现在的贡献是 $\frac{1}{2^{n-i}-1+2^{n-j}-1}$
- 一乘一除即可。
- 对于右边也是同理。
- 转移：直接枚举下一个填左还是右，同时算上它的贡献。
- 注意要计算 $O(n^2)$ 个数的逆元，可以线性预处理。

AGC060C

- 具体地，假设左边第 i 个点的后继中，第一个右边的点为 j
  - 左 $i \to$ 左 $i+1 \to \ldots \to$ 右 j
- 原本左 i 的贡献是 $\frac{1}{2^{n-i-1}}$，现在的贡献是 $\frac{1}{2^{n-i-1}+2^{n-j-1}}$
- 一乘一除即可。
- 对于右边也是同理。
- 转移：直接枚举下一个填左还是右，同时算上它的贡献。
- 注意要计算 $O(n^2)$ 个数的逆元，可以线性预处理。
- 时间复杂度 $O(n^2)$。

条件概率

概率论基础

条件概率：$P(A|B)$ 定义为在 B 发生的条件下，A 发生的概率。

概率论基础

条件概率：$P(A|B)$ 定义为在 B 发生的条件下，A 发生的概率。

$$P(A|B) \triangleq \frac{P(A,B)}{P(B)}$$

概率论基础

条件概率：$P(A|B)$ 定义为在 B 发生的条件下，A 发生的概率。

$$P(A|B) \triangleq \frac{P(A,B)}{P(B)}$$

根据定义，可直接得到**贝叶斯公式**

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

CF1096E 加强版

题目大意

- 考察所有满足 $\sum_{i=1}^{p} a_i = s, a_1 \geq r$ 的非负整数序列 $\{a_i\}_{i=1}^{p}$。
- 定义 winner 是该序列的最大值，如有多个则从中等概率随机选择一个。
- 现在在所有可能的序列中等概率随机选择一个，问 a_1 是 winner 的概率是多少。
- 答案对 998244353 取模。
- 【数据范围】$1 \leq p \leq 10^7, 0 \leq r \leq s \leq 10^7$

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$
- 这样 p 个位置的地位都是一致的了。

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$
- 这样 p 个位置的地位都是一致的了。$|\Omega| = \binom{s+p-1}{p-1}$。

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$
- 这样 p 个位置的地位都是一致的了。$|\Omega| = \binom{s+p-1}{p-1}$。
- 令 winner 的下标是 x，则

$$ans = P(x = 1 | a_1 \geq r)$$

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$
- 这样 p 个位置的地位都是一致的了。$|\Omega| = \binom{s+p-1}{p-1}$。
- 令 winner 的下标是 x，则

$$
\begin{aligned}
ans &= P(x = 1 | a_1 \geq r) \\
&= P(x = 1, a_1 \geq r) / P(a_1 \geq r)
\end{aligned}
$$

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$
- 这样 p 个位置的地位都是一致的了。$|\Omega| = \binom{s+p-1}{p-1}$。
- 令 winner 的下标是 x，则

$$
\begin{aligned}
ans &= P(x = 1 | a_1 \geq r) \\
&= P(x = 1, a_1 \geq r) / P(a_1 \geq r) \\
&= \frac{1}{p} P(a_x \geq r) / P(a_1 \geq r)
\end{aligned}
$$

CF1096E 加强版

- 概率空间 $\Omega = \{(a_1, a_2, ..., a_p) : \sum_{i=1}^{p} a_i = s\}$
- 这样 p 个位置的地位都是一致的了。$|\Omega| = \binom{s+p-1}{p-1}$。
- 令 winner 的下标是 x，则

$$
\begin{aligned}
ans &= P(x = 1 | a_1 \geq r) \\
&= P(x = 1, a_1 \geq r) / P(a_1 \geq r) \\
&= \frac{1}{p} P(a_x \geq r) / P(a_1 \geq r)
\end{aligned}
$$

- 现在有多个最大值也不要紧了！我们只要分别计算最大值 $\geq r$ 的概率，以及 $a_1 \geq r$ 的概率。而这并不困难。

CF1096E 加强版

CF1096E 加强版

- 对于 $P(a_x \geq r)$，使用容斥原理，钦定有 i 个位置 $\geq r$：

$$
\begin{aligned}
P(\max_{i=1}^{p} a_i \geq r) &= 1 - P(\max_{i=1}^{p} a_i \leq r-1) \\
&= 1 - \frac{1}{|\Omega|}\sum_{i=0}^{p}(-1)^i\binom{p}{i}\binom{s-r\cdot i+p-1}{p-1}
\end{aligned}
$$

CF1096E 加强版

- 对于 $P(a_x \geq r)$，使用容斥原理，钦定有 i 个位置 $\geq r$：

$$P(\max_{i=1}^{p} a_i \geq r) = 1 - P(\max_{i=1}^{p} a_i \leq r-1)$$
$$= 1 - \frac{1}{|\Omega|}\sum_{i=0}^{p}(-1)^i\binom{p}{i}\binom{s-r\cdot i+p-1}{p-1}$$

- 对于 $P(a_1 \geq r)$，我们枚举 a_1 的真实值：

$$P(a_1 \geq r) = \frac{1}{|\Omega|}\sum_{a_1=r}^{s}\binom{s-a_1+p-2}{p-2}$$

CF1096E 加强版

- 对于 $P(a_x \geq r)$，使用容斥原理，钦定有 i 个位置 $\geq r$：

$$P(\max_{i=1}^{p} a_i \geq r) = 1 - P(\max_{i=1}^{p} a_i \leq r-1)$$
$$= 1 - \frac{1}{|\Omega|}\sum_{i=0}^{p}(-1)^i\binom{p}{i}\binom{s-r\cdot i+p-1}{p-1}$$

- 对于 $P(a_1 \geq r)$，我们枚举 a_1 的真实值：

$$P(a_1 \geq r) = \frac{1}{|\Omega|}\sum_{a_1=r}^{s}\binom{s-a_1+p-2}{p-2}$$

- 复杂度线性。

CF1096E 加强版

- 对于 $P(a_x \geq r)$，使用容斥原理，钦定有 i 个位置 $\geq r$：

$$
\begin{aligned}
P(\max_{i=1}^{p} a_i \geq r) &= 1 - P(\max_{i=1}^{p} a_i \leq r-1) \\
&= 1 - \frac{1}{|\Omega|}\sum_{i=0}^{p}(-1)^i \binom{p}{i}\binom{s-r\cdot i+p-1}{p-1}
\end{aligned}
$$

- 对于 $P(a_1 \geq r)$，我们枚举 a_1 的真实值：

$$
P(a_1 \geq r) = \frac{1}{|\Omega|}\sum_{a_1=r}^{s}\binom{s-a_1+p-2}{p-2}
$$

- 复杂度线性。注意额外定义 $\binom{x}{-1} = [x=0]$。

离散型随机变量 - 无限操作

- ARC150D
- ABC270Ex

ARC150D

题目大意

- 给定一棵 n 个点的树，初始时节点都是白色
- 定义一个点是坏的，当且仅当它到根路径上存在白点
- 每次等概率随机选择一个坏点，染成黑色
- 当所有点都是黑色（即不存在坏点）时，停止操作
- 求期望操作步数，对 998244353 取模
- 【数据范围】$2 \leq n \leq 2 \cdot 10^5$

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d>0$。

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d > 0$。
- 假设它在第 v 步被随机到，且此时它是坏点。

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d > 0$。
- 假设它在第 v 步被随机到，且此时它是坏点。
- 要求在前 $v-1$ 步中，这 d 个点不全被选过。

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d > 0$。
- 假设它在第 v 步被随机到，且此时它是坏点。
- 要求在前 $v - 1$ 步中，这 d 个点不全被选过。
  - 存在点未被选过 $=1-$ 不存在点未被选过

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d>0$。
- 假设它在第 v 步被随机到，且此时它是坏点。
- 要求在前 $v-1$ 步中，这 d 个点不全被选过。
  - 存在点未被选过 $=1-$ 不存在点未被选过
- 容斥，钦定 d 个点中的某 i 个点没被选过，同时要求第 v 步随机到的是我们要的点：

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d > 0$。
- 假设它在第 v 步被随机到，且此时它是坏点。
- 要求在前 $v - 1$ 步中，这 d 个点不全被选过。
  - 存在点未被选过 $=1-$ 不存在点未被选过
- 容斥，钦定 d 个点中的某 i 个点没被选过，同时要求第 v 步随机到的是我们要的点：

$$f(d) = -\frac{1}{n}\sum_{i=1}^{d}\binom{d}{i}(-1)^i\frac{1}{1-\frac{n-i}{n}} = -\sum_{i=1}^{d}\binom{d}{i}(-1)^i\frac{1}{i}$$

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d>0$。
- 假设它在第 v 步被随机到，且此时它是坏点。
- 要求在前 $v-1$ 步中，这 d 个点不全被选过。
  - 存在点未被选过 $=1-$ 不存在点未被选过
- 容斥，钦定 d 个点中的某 i 个点没被选过，同时要求第 v 步随机到的是我们要的点：

$$f(d)=-\frac{1}{n}\sum_{i=1}^{d}\binom{d}{i}(-1)^i\frac{1}{1-\frac{n-i}{n}}=-\sum_{i=1}^{d}\binom{d}{i}(-1)^i\frac{1}{i}$$

- NTT

ARC150D

- 可假设每次是从 n 个点中等概率随机的，凡是随机到好点就不计入答案。
- 于是只要计算每个点对答案的贡献。这只与其深度有关，设深度为 $d > 0$。
- 假设它在第 v 步被随机到，且此时它是坏点。
- 要求在前 $v - 1$ 步中，这 d 个点不全被选过。
  - 存在点未被选过 $= 1 -$ 不存在点未被选过
- 容斥，钦定 d 个点中的某 i 个点没被选过，同时要求第 v 步随机到的是我们要的点：

$$f(d) = -\frac{1}{n}\sum_{i=1}^{d}\binom{d}{i}(-1)^i\frac{1}{1-\frac{n-i}{n}} = -\sum_{i=1}^{d}\binom{d}{i}(-1)^i\frac{1}{i}$$

- NTT?

ARC150D

- 我们考虑 $f(d) - f(d-1)$，看看会发生什么

ARC150D

- 我们考虑 $f(d)-f(d-1)$，看看会发生什么

$$f(d)-f(d-1)=-\sum_{i=1}^{d-1}\left(\binom{d}{i}-\binom{d-1}{i}\right)(-1)^i\frac{1}{i}-(-1)^d\frac{1}{d}$$

ARC150D

- 我们考虑 $f(d) - f(d-1)$，看看会发生什么

$$f(d) - f(d-1) = -\sum_{i=1}^{d-1}\left(\binom{d}{i} - \binom{d-1}{i}\right)(-1)^i\frac{1}{i} - (-1)^d\frac{1}{d}$$

$$= -\sum_{i=1}^{d-1}\binom{d-1}{i-1}(-1)^i\frac{1}{i} - (-1)^d\frac{1}{d}$$

ARC150D

- 我们考虑 $f(d) - f(d-1)$，看看会发生什么

$$
\begin{aligned}
f(d) - f(d-1) &= -\sum_{i=1}^{d-1}\left(\binom{d}{i} - \binom{d-1}{i}\right)(-1)^i\frac{1}{i} - (-1)^d\frac{1}{d} \\
&= -\sum_{i=1}^{d-1}\binom{d-1}{i-1}(-1)^i\frac{1}{i} - (-1)^d\frac{1}{d} \\
&= -\sum_{i=1}^{d-1}\frac{1}{d}\binom{d}{i}(-1)^i - (-1)^d\frac{1}{d}
\end{aligned}
$$

ARC150D

- 我们考虑 $f(d) - f(d-1)$，看看会发生什么

$$
\begin{aligned}
f(d) - f(d-1) &= -\sum_{i=1}^{d-1}\left(\binom{d}{i} - \binom{d-1}{i}\right)(-1)^i\frac{1}{i} - (-1)^d\frac{1}{d} \\
&= -\sum_{i=1}^{d-1}\binom{d-1}{i-1}(-1)^i\frac{1}{i} - (-1)^d\frac{1}{d} \\
&= -\sum_{i=1}^{d-1}\frac{1}{d}\binom{d}{i}(-1)^i - (-1)^d\frac{1}{d} \\
&= \frac{1}{d}
\end{aligned}
$$

ARC150D

- 我们考虑 $f(d)-f(d-1)$，看看会发生什么

$$
\begin{aligned}
f(d)-f(d-1) &= -\sum_{i=1}^{d-1}\left(\binom{d}{i}-\binom{d-1}{i}\right)(-1)^i\frac{1}{i}-(-1)^d\frac{1}{d} \\
&= -\sum_{i=1}^{d-1}\binom{d-1}{i-1}(-1)^i\frac{1}{i}-(-1)^d\frac{1}{d} \\
&= -\sum_{i=1}^{d-1}\frac{1}{d}\binom{d}{i}(-1)^i-(-1)^d\frac{1}{d} \\
&= \frac{1}{d}
\end{aligned}
$$

- $f(d)=\sum_{i=1}^{d}\frac{1}{i}$，从而本题可以做到线性。

ABC270Ex

题目大意

- 高桥君有一个长度为 n 的序列 $a_1, a_2, ..., a_n$，满足 $0 = a_1 \le a_2 \le \ldots \le a_n$ 且 $a_n > 0$。
- 他还有 n 个计数器，用 $b_1, b_2, \ldots, b_n$ 表示，初值都是 0。
- 每次他会从 1 至 n 中等概率随机一个 i，然后将 b_i 设置为 0，其余 $b_j (j \neq i)$ 加一。
- 当所有 $b_i \ge a_i$ 时，停止。
- 求操作次数的期望，对 998244353 取模。
- 【数据范围】$1 \le n \le 2 \times 10^5, a_n \le 10^{18}$

ABC270Ex

- $ans = \sum_{v=0}^{+\infty} f(v)$，$f(v) =$ 进行了 v 步还未结束的概率。

ABC270Ex

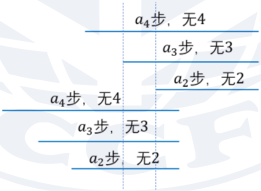
- $ans = \sum_{v=0}^{+\infty} f(v)$，$f(v) =$ 进行了 v 步还未结束的概率。
- 为保证在前 v 步不结束，考虑容斥：钦定 $S \subseteq \{1, 2, ..., v\}$，表示 $\forall i \in S$ 是终止态。容斥系数 $(-1)^{|S|}$。

ABC270Ex

- $ans=\sum_{v=0}^{+\infty} f(v)$，$f(v)=$ 进行了 v 步还未结束的概率。
- 为保证在前 v 步不结束，考虑容斥：钦定 $S\subseteq\{1,2,...,v\}$，表示 $\forall i\in S$ 是终止态。容斥系数 $(-1)^{|S|}$。
- 当钦定第 i 步是终止态时，$[i-a_2+1,i]$ 都不可以随机出 2，$[i-a_3+1,i]$ 都不可以随机出 3，以此类推，倒数的 a_n 步都有限制，再往前就没有限制了。

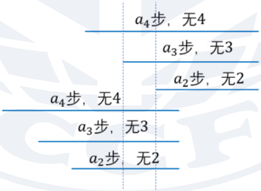
ABC270Ex

- $ans = \sum_{v=0}^{+\infty} f(v)$，$f(v) =$ 进行了 v 步还未结束的概率。
- 为保证在前 v 步不结束，考虑容斥：钦定 $S \subseteq \{1, 2, ..., v\}$，表示 $\forall i \in S$ 是终止态。容斥系数 $(-1)^{|S|}$。
- 当钦定第 i 步是终止态时，$[i - a_2 + 1, i]$ 都不可以随机出 2，$[i - a_3 + 1, i]$ 都不可以随机出 3，以此类推，倒数的 a_n 步都有限制，再往前就没有限制了。
- 而当 $j < i \in S, i - j < a_n$ 时，它们对应的限制区间就会有交。



ABC270Ex

- $ans=\sum_{v=0}^{+\infty} f(v)$，$f(v)=$ 进行了 v 步还未结束的概率。
- 为保证在前 v 步不结束，考虑容斥：钦定 $S\subseteq\{1,2,...,v\}$，表示 $\forall i\in S$ 是终止态。容斥系数 $(-1)^{|S|}$。
- 当钦定第 i 步是终止态时，$[i-a_2+1,i]$ 都不可以随机出 2，$[i-a_3+1,i]$ 都不可以随机出 3，以此类推，倒数的 a_n 步都有限制，再往前就没有限制了。
- 而当 $j<i\in S, i-j<a_n$ 时，它们对应的限制区间就会有交。



- 这不难处理，因为较靠前的限制更强。如上图虚线内，靠后限制为不能有 $3,4$，靠前限制为不能有 $2,3,4$。

ABC270Ex

ABC270Ex

- 因此，从上一次钦定 x 跳跃到下一次 y（即 x, y 是 S 中相邻元素）时：

ABC270Ex

- 因此，从上一次钦定 x 跳跃到下一次 y（即 x, y 是 S 中相邻元素）时：
  - 如果 $y - x < a_n$，则只有 $(x, y]$ 存在多出来的限制。其中 $i \in (x, y]$ 处的限制为：不能随机出 $a_j \geq y - i + 1$ 的 j。

ABC270Ex

- 因此，从上一次钦定 x 跳跃到下一次 y（即 x, y 是 S 中相邻元素）时：
  - 如果 $y - x < a_n$，则只有 $(x, y]$ 存在多出来的限制。其中 $i \in (x, y]$ 处的限制为：不能随机出 $a_j \geq y - i + 1$ 的 j。
  - 如果 $y - x \geq a_n$，则 $[y - a_n + 1, y]$ 有同上的限制。

ABC270Ex

- 因此，从上一次钦定 x 跳跃到下一次 y（即 x, y 是 S 中相邻元素）时：
  - 如果 $y - x < a_n$，则只有 $(x, y]$ 存在多出来的限制。其中 $i \in (x, y]$ 处的限制为：不能随机出 $a_j \geq y - i + 1$ 的 j。
  - 如果 $y - x \geq a_n$，则 $[y - a_n + 1, y]$ 有同上的限制。
- 从而这一段的概率只与 $y - x$ 有关：

$$G(y-x) = \prod_{i=x+1}^{y} \frac{\sum_{j=1}^{n}[a_j \leq y - i]}{n}$$

- 定义 $y - x \geq a_n$ 时的 $G(y - x) = G(a_n)$。

ABC270Ex

ABC270Ex

- 现在，我们可以写出当 $S=\{s_1,s_2,...,s_k\}$（其中 $a_n \le s_1 < \ldots s_k \le v$）时的贡献了

ABC270Ex

- 现在，我们可以写出当 $S=\{s_1,s_2,...,s_k\}$（其中 $a_n\le s_1<\dots s_k\le v$）时的贡献了

$$G(a_n)\cdot\prod_{i=1}^{k-1}G(s_{i+1}-s_i)\cdot(-1)^k$$

ABC270Ex

- 现在，我们可以写出当 $S = \{s_1, s_2, ..., s_k\}$（其中 $a_n \le s_1 < \ldots s_k \le v$）时的贡献了

$$G(a_n) \cdot \prod_{i=1}^{k-1} G(s_{i+1} - s_i) \cdot (-1)^k$$

  - 这里 $G(a_n)$ 表示第一段 $[s_1 - a_n + 1, s_1]$ 的限制，$G(s_{i+1} - s_i)$ 表示后面每多添加一个 s_i 时多出来的限制，$(-1)^k$ 表示容斥系数。

ABC270Ex

- 现在，我们可以写出当 $S = \{s_1, s_2, ..., s_k\}$（其中 $a_n \le s_1 < \dots s_k \le v$）时的贡献了

$$G(a_n) \cdot \prod_{i=1}^{k-1} G(s_{i+1} - s_i) \cdot (-1)^k$$

  - 这里 $G(a_n)$ 表示第一段 $[s_1 - a_n + 1, s_1]$ 的限制，$G(s_{i+1} - s_i)$ 表示后面每多添加一个 s_i 时多出来的限制，$(-1)^k$ 表示容斥系数。
- 因为第一段的存在，我们必须要求 $k \ge 1$。

ABC270Ex

- 现在，我们可以写出当 $S=\{s_1,s_2,...,s_k\}$（其中 $a_n \le s_1 < \dots s_k \le v$）时的贡献了

$$G(a_n)\cdot\prod_{i=1}^{k-1}G(s_{i+1}-s_i)\cdot(-1)^k$$

  - 这里 $G(a_n)$ 表示第一段 $[s_1-a_n+1,s_1]$ 的限制，$G(s_{i+1}-s_i)$ 表示后面每多添加一个 s_i 时多出来的限制，$(-1)^k$ 表示容斥系数。
- 因为第一段的存在，我们必须要求 $k\ge1$。对于 $k=0$ 即 $S=\emptyset$ 的情况，相当于对每个 v 存在 $+1$ 的贡献。这就会导致有无穷多个 1，怎么办？

ABC270Ex

- 现在，我们可以写出当 $S=\{s_1,s_2,\dots,s_k\}$（其中 $a_n \le s_1 < \dots s_k \le v$）时的贡献了

$$G(a_n)\cdot\prod_{i=1}^{k-1}G(s_{i+1}-s_i)\cdot(-1)^k$$

  - 这里 $G(a_n)$ 表示第一段 $[s_1-a_n+1,s_1]$ 的限制，$G(s_{i+1}-s_i)$ 表示后面每多添加一个 s_i 时多出来的限制，$(-1)^k$ 表示容斥系数。
- 因为第一段的存在，我们必须要求 $k\ge 1$。对于 $k=0$ 即 $S=\emptyset$ 的情况，相当于对每个 v 存在 $+1$ 的贡献。这就会导致有无穷多个 1，怎么办？
- 考虑设 $F(x)=\sum_{v\ge 0}f(v)x^v$，则 $ans=F(1)$。对每个 v 存在 $+1$，相当于 $\frac{1}{1-x}$。我们只要在后续运算中将其抵消即可。

- 最后我们来写出 $F(x)$。

ABC270Ex

- 最后我们来写出 $F(x)$。

$$F(x) = \frac{1}{1-x} + \sum_{v\geq 0} x^v \sum_{\emptyset\neq S\subseteq[V]} G(a_n) \cdot \prod_{i=1}^{k-1} G(s_{i+1} - s_i) \cdot (-1)^k$$

ABC270Ex

- 最后我们来写出 $F(x)$。

$$F(x) = \frac{1}{1-x} + \sum_{v \geq 0} x^v \sum_{\emptyset \neq S \subseteq [V]} G(a_n) \cdot \prod_{i=1}^{k-1} G(s_{i+1} - s_i) \cdot (-1)^k$$

$$= \frac{1}{1-x} - G(a_n) \sum_{S \neq \emptyset} x^{s_1} \prod_{i=1}^{k-1} (-G(s_{i+1} - s_i)) x^{s_{i+1} - s_i} \cdot \frac{1}{1-x}$$

- 最后我们来写出 $F(x)$。

$$
\begin{aligned}
F(x) &= \frac{1}{1-x} + \sum_{v\ge 0} x^v \sum_{\emptyset\ne S\subseteq[V]} G(a_n)\cdot\prod_{i=1}^{k-1} G(s_{i+1}-s_i)\cdot(-1)^k \\
&= \frac{1}{1-x} - G(a_n)\sum_{S\ne\emptyset} x^{s_1}\prod_{i=1}^{k-1}(-G(s_{i+1}-s_i))x^{s_{i+1}-s_i}\cdot\frac{1}{1-x} \\
&= \frac{1}{1-x} - \frac{G(a_n)x^{a_n}}{1-x}\cdot\frac{1}{1-x}\cdot\frac{1}{1+Q(x)}
\end{aligned}
$$

ABC270Ex

- 最后我们来写出 $F(x)$。

$$F(x) = \frac{1}{1-x} + \sum_{v\geq 0} x^v \sum_{\emptyset\neq S\subseteq[V]} G(a_n)\cdot\prod_{i=1}^{k-1} G(s_{i+1}-s_i)\cdot(-1)^k$$

$$= \frac{1}{1-x} - G(a_n)\sum_{S\neq\emptyset} x^{s_1}\prod_{i=1}^{k-1}(-G(s_{i+1}-s_i))x^{s_{i+1}-s_i}\cdot\frac{1}{1-x}$$

$$= \frac{1}{1-x} - \frac{G(a_n)x^{a_n}}{1-x}\cdot\frac{1}{1-x}\cdot\frac{1}{1+Q(x)}$$

- 其中 $Q(x)=\sum_{i=1}^{+\infty} G(i)x^i$。由于 $i\geq a_n$ 时 $G(i)=G(a_n)$，故设 $P(x)=\sum_{i=1}^{a_n} G(i)x^i$，有 $Q(x)=P(x)+G(a_n)\cdot\frac{x^{a_n+1}}{1-x}$。

- 最后我们来写出 $F(x)$。

$$F(x)=\frac{1}{1-x}+\sum_{v\geq 0}x^v\sum_{\emptyset\neq S\subseteq[V]}G(a_n)\cdot\prod_{i=1}^{k-1}G(s_{i+1}-s_i)\cdot(-1)^k$$
$$=\frac{1}{1-x}-G(a_n)\sum_{S\neq\emptyset}x^{s_1}\prod_{i=1}^{k-1}(-G(s_{i+1}-s_i))x^{s_{i+1}-s_i}\cdot\frac{1}{1-x}$$
$$=\frac{1}{1-x}-\frac{G(a_n)x^{a_n}}{1-x}\cdot\frac{1}{1-x}\cdot\frac{1}{1+Q(x)}$$

- 其中 $Q(x)=\sum_{i=1}^{+\infty}G(i)x^i$。由于 $i\geq a_n$ 时 $G(i)=G(a_n)$，故设 $P(x)=\sum_{i=1}^{a_n}G(i)x^i$，有 $Q(x)=P(x)+G(a_n)\cdot\frac{x^{a_n+1}}{1-x}$。
- 经过化简，可得 $F(1)=\frac{P(1)-G(a_n)+1}{G(a_n)}$。

- 最后我们来写出 $F(x)$。

$$
\begin{aligned}
F(x) &= \frac{1}{1-x} + \sum_{v\geq 0} x^v \sum_{\emptyset\neq S\subseteq[V]} G(a_n)\cdot\prod_{i=1}^{k-1} G(s_{i+1}-s_i)\cdot(-1)^k \\
&= \frac{1}{1-x} - G(a_n)\sum_{S\neq\emptyset} x^{s_1}\prod_{i=1}^{k-1}(-G(s_{i+1}-s_i))x^{s_{i+1}-s_i}\cdot\frac{1}{1-x} \\
&= \frac{1}{1-x} - \frac{G(a_n)x^{a_n}}{1-x}\cdot\frac{1}{1-x}\cdot\frac{1}{1+Q(x)}
\end{aligned}
$$

- 其中 $Q(x)=\sum_{i=1}^{+\infty} G(i)x^i$。由于 $i\geq a_n$ 时 $G(i)=G(a_n)$，故设 $P(x)=\sum_{i=1}^{a_n} G(i)x^i$，有 $Q(x)=P(x)+G(a_n)\cdot\frac{x^{a_n+1}}{1-x}$。
- 经过化简，可得 $F(1)=\frac{P(1)-G(a_n)+1}{G(a_n)}$。
- 剩下的问题就是如何计算 $P(1)=\sum_{k=1}^{a_n} G(k)$ 和 $G(a_n)$ 了。

ABC270Ex

- 我们回顾一下 $G(\cdot)$ 的式子：

ABC270Ex

- 我们回顾一下 $G(\cdot)$ 的式子：

$$G(k)=\prod_{i=1}^{k}\frac{\sum_{j=1}^{n}[a_j\le k-i]}{n}$$

ABC270Ex

- 我们回顾一下 $G(\cdot)$ 的式子：

$$G(k)=\prod_{i=1}^{k}\frac{\sum_{j=1}^{n}[a_j\leq k-i]}{n}$$

- 注意到 $G(?)$ 其实相当于一个后缀积。

ABC270Ex

- 我们回顾一下 $G(\cdot)$ 的式子：

$$G(k) = \prod_{i=1}^{k} \frac{\sum_{j=1}^{n} [a_j \le k - i]}{n}$$

- 注意到 $G(?)$ 其实相当于一个后缀积。
- 在 a_i 具有单调性的情况下，$\sum_{j=1}^{n} [a_j \le k - i]$ 是一个分段函数，每一段里对 $P(1) = \sum_{k=1}^{a_n} G(k)$ 的贡献就是一个等比数列求和问题。

ABC270Ex

- 我们回顾一下 $G(\cdot)$ 的式子：

$$G(k) = \prod_{i=1}^{k} \frac{\sum_{j=1}^{n} [a_j \le k - i]}{n}$$

- 注意到 $G(?)$ 其实相当于一个后缀积。
- 在 a_i 具有单调性的情况下，$\sum_{j=1}^{n} [a_j \le k - i]$ 是一个分段函数，每一段里对 $P(1) = \sum_{k=1}^{a_n} G(k)$ 的贡献就是一个等比数列求和问题。
- 时间复杂度 $O(n \log P)$。

连续型随机变量

分布函数：令 $F(x) = P(X \leq x)$。

连续型随机变量

分布函数：令 $F(x) = P(X \leq x)$。

- 非负性：$F(x) \in [0,1]$ 且 $F(-\infty) = 0, F(+\infty) = 1$
- 单调性：$F(x)$ 单调不减
- 右连续：$\lim_{x_0 \to x^+} F(x_0) = F(x)$
- 离散随机变量的分布函数一般是跳跃的

连续型随机变量

分布函数：令 $F(x) = P(X \leq x)$。

- 非负性：$F(x) \in [0, 1]$ 且 $F(-\infty) = 0, F(+\infty) = 1$
- 单调性：$F(x)$ 单调不减
- 右连续：$\lim_{x_0 \to x^+} F(x_0) = F(x)$
- 离散随机变量的分布函数一般是跳跃的

概率密度：对于连续型随机变量，概率密度函数 $f(x)$ 满足 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$

连续型随机变量

分布函数：令 $F(x) = P(X \leq x)$。

- 非负性：$F(x) \in [0,1]$ 且 $F(-\infty) = 0, F(+\infty) = 1$
- 单调性：$F(x)$ 单调不减
- 右连续：$\lim_{x_0 \to x^+} F(x_0) = F(x)$
- 离散随机变量的分布函数一般是跳跃的

概率密度：对于连续型随机变量，概率密度函数 $f(x)$ 满足 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$

- 直观理解：$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x)\mathrm{d}x$，因此 $P(X = a) = 0$。

连续型随机变量

分布函数：令 $F(x) = P(X \leq x)$。

- 非负性：$F(x) \in [0,1]$ 且 $F(-\infty) = 0, F(+\infty) = 1$
- 单调性：$F(x)$ 单调不减
- 右连续：$\lim_{x_0 \to x^+} F(x_0) = F(x)$
- 离散随机变量的分布函数一般是跳跃的

概率密度：对于连续型随机变量，概率密度函数 $f(x)$ 满足 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$

- 直观理解：$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x)\mathrm{d}x$，因此 $P(X = a) = 0$。
- 均匀分布 $U[a,b]$，其 $f(x) = \frac{1}{b-a}, \forall x \in [a,b]$；

连续型随机变量

分布函数：令 $F(x)=P(X\leq x)$。

- 非负性：$F(x)\in[0,1]$ 且 $F(-\infty)=0, F(+\infty)=1$
- 单调性：$F(x)$ 单调不减
- 右连续：$\lim_{x_0\to x^+}F(x_0)=F(x)$
- 离散随机变量的分布函数一般是跳跃的

概率密度：对于连续型随机变量，概率密度函数 $f(x)$ 满足 $F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t$

- 直观理解：$P(x_1<X\leq x_2)=\int_{x_1}^{x_2}f(x)\mathrm{d}x$，因此 $P(X=a)=0$。
- 均匀分布 $U[a,b]$，其 $f(x)=\frac{1}{b-a}, \forall x\in[a,b]$；
- 高斯分布 $N(0,1)$，其 $f(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \forall x\in\mathbb{R}$。

ARC113F

题目大意

给定 $0 = X_0 < X_1 < \ldots < X_n$，令随机变量 a_i 在 $[X_{i-1}, X_i]$ 内均匀随机取值。

求 $\min_{i=1}^{n-1}(a_{i+1} - a_i)$ 的期望。

答案对 998244353 取模。

【数据范围】$2 \leq n \leq 20, X_n \leq 10^6$

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i)\geq m$ 的概率 $f(m)$，则 $ans=\int_0^{X_n}f(m)\mathrm{d}m$。

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i) \ge m$ 的概率 $f(m)$，则 $ans = \int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1}-a_i \ge m \iff b_i = a_i - m \cdot i, b_i \le b_{i+1}, \forall 1 \le i < n$

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i)\ge m$ 的概率 $f(m)$，则 $ans=\int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1}-a_i\ge m \iff b_i=a_i-m\cdot i, b_i\le b_{i+1}, \forall 1\le i<n$
- $a_i\sim U[X_{i-1},X_i] \iff b_i\sim U[X_{i-1}-m\cdot i, X_i-m\cdot i]$

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i) \geq m$ 的概率 $f(m)$，则 $ans = \int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1} - a_i \geq m \iff b_i = a_i - m \cdot i, b_i \leq b_{i+1}, \forall 1 \leq i < n$
- $a_i \sim U[X_{i-1}, X_i] \iff b_i \sim U[X_{i-1} - m \cdot i, X_i - m \cdot i]$
- 因此，计算 $f(m)$ 等价于计算从上述 n 个区间内分别抽取 b_i 且满足 b_i 递增的概率。

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i) \ge m$ 的概率 $f(m)$，则 $ans = \int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1}-a_i \ge m \iff b_i = a_i - m\cdot i, b_i \le b_{i+1}, \forall 1 \le i < n$
- $a_i \sim U[X_{i-1}, X_i] \iff b_i \sim U[X_{i-1}-m\cdot i, X_i - m\cdot i]$
- 因此，计算 $f(m)$ 等价于计算从上述 n 个区间内分别抽取 b_i 且满足 b_i 递增的概率。
- 如果这些区间是确定的，那么只要对区间端点进行离散化，按照数轴方向从左往右进行 $O(n^3)$ DP 即可。

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i) \geq m$ 的概率 $f(m)$，则 $ans=\int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1}-a_i \geq m \iff b_i = a_i - m\cdot i, b_i \leq b_{i+1}, \forall 1 \leq i < n$
- $a_i \sim U[X_{i-1}, X_i] \iff b_i \sim U[X_{i-1}-m\cdot i, X_i - m\cdot i]$
- 因此，计算 $f(m)$ 等价于计算从上述 n 个区间内分别抽取 b_i 且满足 b_i 递增的概率。
- 如果这些区间是确定的，那么只要对区间端点进行离散化，按照数轴方向从左往右进行 $O(n^3)$ DP 即可。
- 进一步地，这一过程仅取决于这些端点的相对顺序。而随着 m 的变化，相对顺序最多改变 $O(n^2)$ 次。

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i)\geq m$ 的概率 $f(m)$，则 $ans=\int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1}-a_i\geq m \iff b_i=a_i-m\cdot i, b_i\leq b_{i+1}, \forall 1\leq i<n$
- $a_i\sim U[X_{i-1},X_i] \iff b_i\sim U[X_{i-1}-m\cdot i, X_i-m\cdot i]$
- 因此，计算 $f(m)$ 等价于计算从上述 n 个区间内分别抽取 b_i 且满足 b_i 递增的概率。
- 如果这些区间是确定的，那么只要对区间端点进行离散化，按照数轴方向从左往右进行 $O(n^3)$ DP 即可。
- 进一步地，这一过程仅取决于这些端点的相对顺序。而随着 m 的变化，相对顺序最多改变 $O(n^2)$ 次。
  - 求解分界点的方法：计算直线交点 + 排序离散化。

ARC113F

- 考虑固定 m，计算 $\min_{i=1}^{n-1}(a_{i+1}-a_i) \ge m$ 的概率 $f(m)$，则 $ans = \int_0^{X_n} f(m)\mathrm{d}m$。
- $a_{i+1}-a_i \ge m \iff b_i = a_i - m\cdot i, b_i \le b_{i+1}, \forall 1 \le i < n$
- $a_i \sim U[X_{i-1}, X_i] \iff b_i \sim U[X_{i-1}-m\cdot i, X_i - m\cdot i]$
- 因此，计算 $f(m)$ 等价于计算从上述 n 个区间内分别抽取 b_i 且满足 b_i 递增的概率。
- 如果这些区间是确定的，那么只要对区间端点进行离散化，按照数轴方向从左往右进行 $O(n^3)$ DP 即可。
- 进一步地，这一过程仅取决于这些端点的相对顺序。而随着 m 的变化，相对顺序最多改变 $O(n^2)$ 次。
  - 求解分界点的方法：计算直线交点 $+$ 排序离散化。
- 于是，我们可以单独考察每个相对顺序不变的区间 $[L, R]$，计算 $\int_L^R f(m)\mathrm{d}m$ 并最后加起来就是答案。

ARC113F

- 现在让我们回到 DP。

ARC113F

- 现在让我们回到 DP。
- 注意到，在每一步转移时，转移系数是一个关于 m 的一次多项式。

ARC113F

- 现在让我们回到 DP。
- 注意到，在每一步转移时，转移系数是一个关于 m 的一次多项式。
- 因此，最终的 $f(m)$ 一定是一个关于 m 的 n 次多项式。多项式的系数可以通过 DP 来维护，也可以通过 Lagrange 插值法得到。

ARC113F

- 现在让我们回到 DP。
- 注意到，在每一步转移时，转移系数是一个关于 m 的一次多项式。
- 因此，最终的 $f(m)$ 一定是一个关于 m 的 n 次多项式。多项式的系数可以通过 DP 来维护，也可以通过 Lagrange 插值法得到。
- 对于多项式的积分是容易的。

ARC113F

- 现在让我们回到 DP。
- 注意到，在每一步转移时，转移系数是一个关于 m 的一次多项式。
- 因此，最终的 $f(m)$ 一定是一个关于 m 的 n 次多项式。多项式的系数可以通过 DP 来维护，也可以通过 Lagrange 插值法得到。
- 对于多项式的积分是容易的。
- 总时间复杂度 $O(n^6)$

概率方法

概率方法

【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立？正确的概率足够高即可。

概率方法

【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立？正确的概率足够高即可。

- 矩阵乘法是很慢的，但是矩阵乘向量是足够快的！

概率方法

【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立？正确的概率足够高即可。

- 矩阵乘法是很慢的，但是矩阵乘向量是足够快的！
- 随机一个列向量 x，判定是否有 $ABx = x$。

概率方法

【问题 1】给定 $n \times n$ 的矩阵 A, B，如何在 $O(n^2)$ 内判断 $AB = I$ 是否成立？正确的概率足够高即可。

- 矩阵乘法是很慢的，但是矩阵乘向量是足够快的！
- 随机一个列向量 x，判定是否有 $ABx = x$。
- 相关题目：QOJ7612

概率方法

概率方法

【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。

概率方法

【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。

- 法 1：对每个顶点，以 $1/2$ 的概率决定其染黑还是白。选出所有连接的顶点不同色的边，则它们构成一个二分子图。

概率方法

【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。

- 法 1：对每个顶点，以 $1/2$ 的概率决定其染黑还是白。选出所有连接的顶点不同色的边，则它们构成一个二分子图。
- $\mathrm{E}[\text{ 选出的边数 }] = \frac{1}{2}m$，因此存在性得证。

概率方法

【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。

- 法 1：对每个顶点，以 $1/2$ 的概率决定其染黑还是白。选出所有连接的顶点不同色的边，则它们构成一个二分子图。
- $\mathrm{E}[$ 选出的边数 $] = \frac{1}{2}m$，因此存在性得证。
- 具体实现：多次重复（至少可以做到 $1/2$ 近似）。

概率方法

【问题 2】给定无向图 G，找到一个至少包含一半边的二分子图。

- 法 1：对每个顶点，以 $1/2$ 的概率决定其染黑还是白。选出所有连接的顶点不同色的边，则它们构成一个二分子图。
- $\mathrm{E}[\text{ 选出的边数 }] = \frac{1}{2}m$，因此存在性得证。
- 具体实现：多次重复（至少可以做到 $1/2$ 近似)。
- 法 2：增量法，依次加入每个点，令其颜色为当前邻居中较少的一种。则这样一定有至少一半的边满足顶点颜色不同。

概率方法

【问题 2.5】给定无向图 G，若其有 $2n$ 个点和 m 条边，则至少存在一个包含 $mn/(2n-1)$ 条边的二分子图；若其有 $2n+1$ 个点和 m 条边，则至少存在一个包含 $m(n+1)/(2n+1)$ 条边的二分子图。

概率方法

【问题 2.5】给定无向图 G，若其有 $2n$ 个点和 m 条边，则至少存在一个包含 $mn/(2n-1)$ 条边的二分子图；若其有 $2n+1$ 个点和 m 条边，则至少存在一个包含 $m(n+1)/(2n+1)$ 条边的二分子图。

- 对于 $2n$ 个点的情形，我们每次从 $\binom{2n}{n}$ 个大小为 n 的子集中等概率随机选择，并将它们染黑，余下点染白。

概率方法

【问题 2.5】给定无向图 G，若其有 $2n$ 个点和 m 条边，则至少存在一个包含 $mn/(2n-1)$ 条边的二分子图；若其有 $2n+1$ 个点和 m 条边，则至少存在一个包含 $m(n+1)/(2n+1)$ 条边的二分子图。

- 对于 $2n$ 个点的情形，我们每次从 $\binom{2n}{n}$ 个大小为 n 的子集中等概率随机选择，并将它们染黑，余下点染白。
- 则一条边的两个顶点颜色不同的概率是 $\frac{n}{2n-1}$。

概率方法

【问题 2.5】给定无向图 G，若其有 $2n$ 个点和 m 条边，则至少存在一个包含 $mn/(2n-1)$ 条边的二分子图；若其有 $2n+1$ 个点和 m 条边，则至少存在一个包含 $m(n+1)/(2n+1)$ 条边的二分子图。

- 对于 $2n$ 个点的情形，我们每次从 $\binom{2n}{n}$ 个大小为 n 的子集中等概率随机选择，并将它们染黑，余下点染白。
- 则一条边的两个顶点颜色不同的概率是 $\frac{n}{2n-1}$。
- 对于 $2n+1$ 个点的情形，我们选 n 个顶点染黑，$n+1$ 个顶点染白，则一条边的两个顶点颜色不同的概率是 $\frac{n+1}{2n+1}$。

概率方法

概率方法

使用类似的方法，我们可以得到许多有趣的结论，比如：
【问题 3】$\forall n \in Z_+$，存在一个竞赛图包含至少 $n!/2^{n-1}$ 条哈密尔顿路径。

概率方法

使用类似的方法，我们可以得到许多有趣的结论，比如：

【问题 3】$\forall n \in Z_+$，存在一个竞赛图包含至少 $n!/2^{n-1}$ 条哈密尔顿路径。

- 对于竞赛图中的每条边，以 $1/2, 1/2$ 的概率决定它的方向。

概率方法

使用类似的方法，我们可以得到许多有趣的结论，比如：
【问题 3】$\forall n \in Z_+$，存在一个竞赛图包含至少 $n!/2^{n-1}$ 条哈密尔顿路径。

- 对于竞赛图中的每条边，以 $1/2, 1/2$ 的概率决定它的方向。
- 则对于固定的一个排列，其构成哈密尔顿路径的概率是 $\frac{1}{2^{n-1}}$。

概率方法

使用类似的方法，我们可以得到许多有趣的结论，比如：
【问题 3】$\forall n \in Z_+$，存在一个竞赛图包含至少 $n!/2^{n-1}$ 条哈密尔顿路径。

- 对于竞赛图中的每条边，以 $1/2, 1/2$ 的概率决定它的方向。
- 则对于固定的一个排列，其构成哈密尔顿路径的概率是 $\frac{1}{2^{n-1}}$。
- $E[$ 哈密尔顿路径数量 $] = n!/2^{n-1}$，因此得证。

概率方法

概率方法

【问题 4】$\forall B=\{b_1,b_2,...,b_n\}, b_i\neq 0$，$\exists A\subseteq B$ 满足其是 sum-free 的（不存在 $a_1,a_2,a_3\in A, a_1+a_2=a_3$）且 $|A|>\frac{1}{3}n$。

【问题 4】$\forall B=\{b_1,b_2,...,b_n\}, b_i\neq 0$，$\exists A\subseteq B$ 满足其是 sum-free 的（不存在 $a_1,a_2,a_3\in A, a_1+a_2=a_3$）且 $|A|>\frac{1}{3}n$。

- 取质数 $p=3k+2$ 满足 $p>2\max|b_i|$。

概率方法

【问题 4】$\forall B = \{b_1, b_2, ..., b_n\}, b_i \neq 0$，$\exists A \subseteq B$ 满足其是 sum-free 的（不存在 $a_1, a_2, a_3 \in A, a_1 + a_2 = a_3$）且 $|A| > \frac{1}{3}n$。

- 取质数 $p = 3k + 2$ 满足 $p > 2\max|b_i|$。
- 考察集合 $C = \{k+1, k+2, \ldots, 2k+1\}$，其在 $\mod p$ 下是 sum-free 的，且 $\frac{|C|}{p-1} > \frac{1}{3}$。

概率方法

【问题 4】$\forall B = \{b_1, b_2, ..., b_n\}, b_i \neq 0$，$\exists A \subseteq B$ 满足其是 sum-free 的（不存在 $a_1, a_2, a_3 \in A, a_1 + a_2 = a_3$）且 $|A| > \frac{1}{3}n$。

- 取质数 $p = 3k + 2$ 满足 $p > 2\max|b_i|$。
- 考察集合 $C = \{k+1, k+2, \ldots, 2k+1\}$，其在 $\mod p$ 下是 sum-free 的，且 $\frac{|C|}{p-1} > \frac{1}{3}$。
- 现在从 $1, 2, \ldots, p-1$ 中等概率随机选取一个值作为 x。

概率方法

【问题 4】$\forall B=\{b_1,b_2,\ldots,b_n\}, b_i\neq 0$，$\exists A\subseteq B$ 满足其是 sum-free 的（不存在 $a_1,a_2,a_3\in A, a_1+a_2=a_3$）且 $|A|>\frac{1}{3}n$。

- 取质数 $p=3k+2$ 满足 $p>2\max|b_i|$。
- 考察集合 $C=\{k+1,k+2,\ldots,2k+1\}$，其在 $\mod p$ 下是 sum-free 的，且 $\frac{|C|}{p-1}>\frac{1}{3}$。
- 现在从 $1,2,\ldots,p-1$ 中等概率随机选取一个值作为 x。
- 令 $d_i\equiv xb_i \pmod p$，则固定 i，随着 x 的变化，d_i 会跑遍 $1,2,\ldots,p-1$，因此 $P(d_i\in C)=\frac{|C|}{p-1}>\frac{1}{3}$。

概率方法

【问题 4】$\forall B=\{b_1,b_2,...,b_n\}, b_i\neq 0$，$\exists A\subseteq B$ 满足其是 sum-free 的（不存在 $a_1,a_2,a_3\in A, a_1+a_2=a_3$）且 $|A|>\frac{1}{3}n$。

- 取质数 $p=3k+2$ 满足 $p>2\max|b_i|$。
- 考察集合 $C=\{k+1,k+2,\ldots,2k+1\}$，其在 $\mod p$ 下是 sum-free 的，且 $\frac{|C|}{p-1}>\frac{1}{3}$。
- 现在从 $1,2,\ldots,p-1$ 中等概率随机选取一个值作为 x。
- 令 $d_i\equiv xb_i \pmod p$，则固定 i，随着 x 的变化，d_i 会跑遍 $1,2,...,p-1$，因此 $P(d_i\in C)=\frac{|C|}{p-1}>\frac{1}{3}$。
- 这意味着 $E[\sum_{i=1}^{n}[d_i\in C]]>n/3$，因此 $\exists 1\leq x<p$ 和一个 $A\subseteq B$ 满足 $|A|>n/3$ 且 $\forall a\in A, xa \pmod p\in C$。

概率方法

【问题 4】$\forall B=\{b_1,b_2,...,b_n\}, b_i\neq 0$，$\exists A\subseteq B$ 满足其是 sum-free 的（不存在 $a_1,a_2,a_3\in A, a_1+a_2=a_3$）且 $|A|>\frac{1}{3}n$。

- 取质数 $p=3k+2$ 满足 $p>2\max|b_i|$。
- 考察集合 $C=\{k+1,k+2,\ldots,2k+1\}$，其在 $\mod p$ 下是 sum-free 的，且 $\frac{|C|}{p-1}>\frac{1}{3}$。
- 现在从 $1,2,\ldots,p-1$ 中等概率随机选取一个值作为 x。
- 令 $d_i\equiv xb_i \pmod p$，则固定 i，随着 x 的变化，d_i 会跑遍 $1,2,...,p-1$，因此 $P(d_i\in C)=\frac{|C|}{p-1}>\frac{1}{3}$。
- 这意味着 $E[\sum_{i=1}^{n}[d_i\in C]]>n/3$，因此 $\exists 1\leq x<p$ 和一个 $A\subseteq B$ 满足 $|A|>n/3$ 且 $\forall a\in A, xa \pmod p\in C$。
- A 是 sum-free 的，因为假如 $a_1+a_2=a_3\Rightarrow xa_1+xa_2\equiv xa_3 \pmod p$，与 C 在 $\mod p$ 下 sum-free 矛盾。

概率方法 - 例题

- CF1746F
- QOJ6413

CF1746F

题目大意

- 有一个长度为 n 的序列 $\{a_n\}$，你需要支持单点修改，并回答询问：
- 给定 l, r, k，回答是否 $\forall i, \sum_{j=l}^{r}[i = a_j]$ 均为 k 的倍数。
- 【数据范围】$1 \leq n, q \leq 3 \cdot 10^5$

CF1746F

- 先离散化

CF1746F

- 先离散化
- 然后以每个数 1/2 的概率，随机一个值域的子集

CF1746F

- 先离散化
- 然后以每个数 $1/2$ 的概率，随机一个值域的子集
- 统计这个集合内数出现的次数总和，是否是 k 的倍数

CF1746F

- 先离散化
- 然后以每个数 1/2 的概率，随机一个值域的子集
- 统计这个集合内数出现的次数总和，是否是 k 的倍数
- 如果答案是 YES，不会有问题

CF1746F

- 先离散化
- 然后以每个数 $1/2$ 的概率，随机一个值域的子集
- 统计这个集合内数出现的次数总和，是否是 k 的倍数
- 如果答案是 YES，不会有问题
- 如果答案是 NO，那么判错的概率 $\leq 1/2$

CF1746F

- 先离散化
- 然后以每个数 $1/2$ 的概率，随机一个值域的子集
- 统计这个集合内数出现的次数总和，是否是 k 的倍数
- 如果答案是 YES，不会有问题
- 如果答案是 NO，那么判错的概率 $\leq 1/2$（为什么?)

CF1746F

- 先离散化
- 然后以每个数 $1/2$ 的概率，随机一个值域的子集
- 统计这个集合内数出现的次数总和，是否是 k 的倍数
- 如果答案是 YES，不会有问题
- 如果答案是 NO，那么判错的概率 $\leq 1/2$（为什么?)
- 事先准备好 30 个固定的子集，询问时拿出来判断即可

QOJ6413

题目大意

- 给定一个连通无向图 G，称 v 是叶子当且仅当它度数 $=1$。保证每个点的邻居叶子个数不超过 2。
- 找一个点集 S 使得 S 是一个支配集，S 的补集也是支配集，$|S| = \lfloor n/2 \rfloor$。
- 可以证明存在！
- 【数据范围】$n, m \leq 2 \times 10^5$

QOJ6413

QOJ6413

- 首先进行一些删边操作。

QOJ6413

- 首先进行一些删边操作。
- 如果一条边删了以后，不产生孤立点，且每个点的邻居叶子个数还是 ≤ 2，我们就可以把它删掉。

QOJ6413

- 首先进行一些删边操作。
- 如果一条边删了以后，不产生孤立点，且每个点的邻居叶子个数还是 ≤ 2，我们就可以把它删掉。
  - 一方面是题目条件 $\Rightarrow$ 解的存在性，另一方面是支配集的限制只会更强。

QOJ6413

- 首先进行一些删边操作。
- 如果一条边删了以后，不产生孤立点，且每个点的邻居叶子个数还是 ≤ 2，我们就可以把它删掉。
  - 一方面是题目条件 $\Rightarrow$ 解的存在性，另一方面是支配集的限制只会更强。
- 容易在 $O(m \log m)$ 复杂度内完成所有删边操作。

QOJ6413

- 首先进行一些删边操作。
- 如果一条边删了以后，不产生孤立点，且每个点的邻居叶子个数还是 ≤ 2，我们就可以把它删掉。
  - 一方面是题目条件 $\Rightarrow$ 解的存在性，另一方面是支配集的限制只会更强。
- 容易在 $O(m \log m)$ 复杂度内完成所有删边操作。
- 下面来考察无法删边的结构是什么样的。

QOJ6413

- 事实上，余下的结构非常特殊。

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。
- (1) X 是独立集

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。
- (1) X 是独立集
  - 若 $u, v \in X, u \sim v$，则可以把 $u \sim v$ 这条边删除。

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。
- (1) X 是独立集
  - 若 $u, v \in X, u \sim v$，则可以把 $u \sim v$ 这条边删除。
- (2) Y 是独立集

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。
- (1) X 是独立集
  - 若 $u, v \in X, u \sim v$，则可以把 $u \sim v$ 这条边删除。
- (2) Y 是独立集
  - 否则，取 $G[Y]$ 的一个非孤立点连通块 C。由 G 的连通性，存在 $u \sim v, u \in X, v \in C$。

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。
- (1) X 是独立集
  - 若 $u, v \in X, u \sim v$，则可以把 $u \sim v$ 这条边删除。
- (2) Y 是独立集
  - 否则，取 $G[Y]$ 的一个非孤立点连通块 C。由 G 的连通性，存在 $u \sim v, u \in X, v \in C$。
  - 由 C 的连通性，存在 $w \in C, v \sim w$。

QOJ6413

- 事实上，余下的结构非常特殊。
- 由于只要对每个连通块求出一种合法方案，不难合并得到答案，因此不妨假设整张图连通。
- 不妨设点数 ≥ 4。
- 设 X 为度数至少为 3 的顶点集合， Y 是度数至多为 2 的顶点集合。
- (1) X 是独立集
  - 若 $u, v \in X, u \sim v$，则可以把 $u \sim v$ 这条边删除。
- (2) Y 是独立集
  - 否则，取 $G[Y]$ 的一个非孤立点连通块 C。由 G 的连通性，存在 $u \sim v, u \in X, v \in C$。
  - 由 C 的连通性，存在 $w \in C, v \sim w$。
  - 则经过检验，可以把 $u \sim v$ 这条边删除。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \ge 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。
  - 目前，仍有 A, B 大小差 ≤ 1。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。
  - 目前，仍有 A, B 大小差 ≤ 1。
  - 现在还剩下所有 $\deg = 2$ 的点，注意这些点分为两类：

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \ge 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。
  - 目前，仍有 A, B 大小差 ≤ 1。
  - 现在还剩下所有 $\deg = 2$ 的点，注意这些点分为两类：
  - $C =$ 两个邻居均在 A' 中（或均在 B'）中

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \ge 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。
  - 目前，仍有 A, B 大小差 ≤ 1。
  - 现在还剩下所有 $\deg = 2$ 的点，注意这些点分为两类：
  - $C =$ 两个邻居均在 A' 中（或均在 B'）中
  - $D =$ 两个邻居一个在 A'，一个在 B'

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。
  - 目前，仍有 A, B 大小差 ≤ 1。
  - 现在还剩下所有 $\deg = 2$ 的点，注意这些点分为两类：
  - $C =$ 两个邻居均在 A' 中（或均在 B'）中
  - $D =$ 两个邻居一个在 A'，一个在 B'
- 对于 C，其必须被放入与邻居相反的集合中；对于 D，随意。

QOJ6413

- (3) X 中每个点恰有 2 个叶子邻居
  - 设 $u \in X$ 的叶子邻居个数 ≤ 1。由于 $\deg(u) \geq 3$ 且 X 是独立集，$\exists v \sim u$ 使得 $\deg(v) = 2$。
  - 由于 Y 是独立集，v 的另一个邻居 $w \in X$。
  - 则经过检验，可以把 $v \sim w$ 这条边删除。
- 我们希望只对 X 进行染色，然后其他所有点的颜色可以自然确定。
- 设 X 被划分为 A' 和 B' 两个不交的集合且大小差 ≤ 1。令初始 $A = A', B = B'$。
  - 挂在 $u \in A'$ 上的叶子，放入 B；反之同理。
  - 目前，仍有 A, B 大小差 ≤ 1。
  - 现在还剩下所有 $\deg = 2$ 的点，注意这些点分为两类:
  - $C =$ 两个邻居均在 A' 中（或均在 B'）中
  - $D =$ 两个邻居一个在 A'，一个在 B'
- 对于 C，其必须被放入与邻居相反的集合中；对于 D，随意。
- 因此只要 $|D| \geq |C|$，就可得合法构造。

QOJ6413

QOJ6413

- 我们的问题转化成了：
  - 将 X 分成 A', B'，满足 $||A'| - |B'|| \le 1$，且 $|D| \ge |C|$，其中 C 表示 A', B' 各自内部的边，D 表示 A', B' 之间的边。

QOJ6413

- 我们的问题转化成了：
  - 将 X 分成 A', B'，满足 $||A'| - |B'|| \le 1$，且 $|D| \ge |C|$，其中 C 表示 A', B' 各自内部的边，D 表示 A', B' 之间的边。
- 这等价于找一个至少包含一半边的二分子图，只不过额外要求 $||A'| - |B'|| \le 1$。

QOJ6413

- 我们的问题转化成了：
  - 将 X 分成 A', B'，满足 $||A'| - |B'|| \le 1$，且 $|D| \ge |C|$，其中 C 表示 A', B' 各自内部的边，D 表示 A', B' 之间的边。
- 这等价于找一个至少包含一半边的二分子图，只不过额外要求 $||A'| - |B'|| \le 1$。
- 【法 1】概率方法。均匀随机考虑一个满足 $||A'| - |B'|| \le 1$ 的划分，与【问题 2.5】一致。

QOJ6413

- 我们的问题转化成了：
  - 将 X 分成 A', B'，满足 $||A'|-|B'|| \le 1$，且 $|D| \ge |C|$，其中 C 表示 A', B' 各自内部的边，D 表示 A', B' 之间的边。
- 这等价于找一个至少包含一半边的二分子图，只不过额外要求 $||A'|-|B'|| \le 1$。
- 【法 1】概率方法。均匀随机考虑一个满足 $||A'|-|B'|| \le 1$ 的划分，与【问题 2.5】一致。
- 【法 2】调整。从一个满足 $||A'|-|B'|| \le 1$ 的划分开始，每次交换两个点。

QOJ6413

- 我们的问题转化成了：
  - 将 X 分成 A', B'，满足 $||A'| - |B'|| \leq 1$，且 $|D| \geq |C|$，其中 C 表示 A', B' 各自内部的边，D 表示 A', B' 之间的边。
- 这等价于找一个至少包含一半边的二分子图，只不过额外要求 $||A'| - |B'|| \leq 1$。
- 【法 1】概率方法。均匀随机考虑一个满足 $||A'| - |B'|| \leq 1$ 的划分，与【问题 2.5】一致。
- 【法 2】调整。从一个满足 $||A'| - |B'|| \leq 1$ 的划分开始，每次交换两个点。
- 【法 3】增量法。每次加入两个新的点。

Q&A

Q&A

谢谢大家！