组合数学第 7 讲

授课时间: 2024 年 10 月 12 日　授课教师: 孙晓明　记录人: 李驭（助教）　陈子珩（助教）

1　一维随机游走问题扩展

例 1　一个人从数轴上 1 的位置出发，每次以 p 的概率向负方向移动一格，以 $1-p$ 的概率向正方向移动一格，走至原点时停下. 试求这个人走回原点的概率以及他走回原点的期望步数。

解　记 p_k 为这个人在第 k 步走到原点的概率. 显然，只有走奇数步才有可能回到原点，因此 $p_{2k}=0$. 现在考虑求 p_{2k+1}，由于从 1 走回原点需要向左走 $k+1$ 步，向右走 k 步，且不能在第 $2k+1$ 步之前到达原点. 根据对 Catalan 数的讨论，有

$$p_{2k+1}=p^{k+1}(1-p)^kC_k,$$

其中 C_k 为第 k 个卡特兰数.

则在有限步内回到 0 的概率为

$$\begin{aligned}\sum_{k\ge0}p_{2k+1}&=\sum_{k\ge0}p^{k+1}(1-p)^kC_k\\&=p\sum_{k\ge0}(p(1-p))^k\,C_k\\&=pC\left(p(1-p)\right)\\&=p\frac{1-(1-4p(1-p))^{\frac12}}{2p(1-p)}\\&=\frac{1-|2p-1|}{2(1-p)}\\&=\begin{cases}1&\text{if } p\ge\frac12;\\\frac{p}{1-p}&\text{if } p<\frac12.\end{cases}\end{aligned}$$

由上式不难看出，当 $p\ge\frac12$ 时，此人将会以概率 1 回到原点；当 $p<\frac12$ 是，此人将会以大于 0 的概率永远无法回到原点.

现在计算游走次数 T 的期望. $p<\frac12$ 时，因为此人有大于 0 的概率无法回到原点，所以他有大于 0 的概率走 $+\infty$ 步，又因为游走次数必为正数，所以 $\mathbb{E}(T)=+\infty$.

$p\ge\frac12$ 时，

$$\begin{aligned}\mathbb{E}(T)&=\sum_{k\ge0}(2k+1)p_{2k+1}\\&=1+\sum_{k\ge0}2kp_{2k+1}\\&=1+\sum_{k\ge0}2p(p(1-p))^kkC_k.\end{aligned}$$

变换 Catalan 数的生成函数，

$$
\begin{aligned}
C(x) &= \frac{1-(1-4x)^{\frac{1}{2}}}{2x} = \sum_{k\geq 0} C_k x^k \\
xC'(x) &= \sum_{k\geq 0} C_k k x^k.
\end{aligned}
$$

从而

$$
\begin{aligned}
\mathbb{E}(T) &= 1 + 2p\left(\frac{1}{(1-4x)^{\frac{1}{2}}} - \frac{1-(1-4x)^{\frac{1}{2}}}{2x}\right)\Bigg|_{x=p(1-p)} \\
&= 1 + 2p\left(\frac{1}{(1-4p(1-p))^{\frac{1}{2}}} - \frac{1-(1-4p(1-p))^{\frac{1}{2}}}{2p(1-p)}\right) \\
&= \frac{1}{2p-1}.
\end{aligned}
$$

这个结果说明，$p < \frac{1}{2}$ 时，我们总能期望该随机游走会在有限步后停止；$p > \frac{1}{2}$ 时，回到原点所花费的游走次数的期望是 $+\infty$；而 $p = \frac{1}{2}$ 时，该情况非常特殊：虽然此人会以概率 1 回到原点，但是他回到原点所花费的游走次数的期望却是 $+\infty$.

2 容斥原理

考虑多个集合的并集的元素数量：

- 对于两个集合 A, B 的情况，易知

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

- 对于三个集合 A, B, C 的情况，通过分析韦恩图可得

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

一般地，我们给出 n 个集合并集元素数量的计算公式：

定理 1 (容斥原理并集形式). 给定 n 个集合 $A_1, A_2, \ldots, A_n$，则有

$$\left|\bigcup_{i=1}^n A_i\right| = \sum_{1\leq i\leq n} |A_i| - \sum_{1\leq i<j\leq n} |A_i \cap A_j| + \sum_{1\leq i<j<k\leq n} |A_i \cap A_j \cap A_k| + \cdots + (-1)^{n-1}\left|\bigcap_{i=1}^n A_i\right|. \tag{1}$$

定理 2 (容斥原理交集形式). 给定全集 I 和 I 的 n 个子集 $A_1, A_2, \ldots, A_n$，则有

$$\left|\bigcap_{i=1}^n \bar{A}_i\right| = |I| - \sum_{1\leq i\leq n} |A_i| + \sum_{1\leq i<j\leq n} |A_i \cap A_j| - \sum_{1\leq i<j<k\leq n} |A_i \cap A_j \cap A_k| + \cdots + (-1)^n\left|\bigcap_{i=1}^n A_i\right|.$$

证明 因为 $\bigcup_{i=1}^n A_i$ 和 $\bigcap\limits_{i=1}^n \bar{A}_i$ 互为补集，所以若定理 1成立则定理 2自动成立. 于是我们仅需证明定理 1.

$$\left|\bigcup_{i=1}^n A_i\right| = \left|\left\{x \middle| x \in \bigcup_{i=1}^n A_i\right\}\right| = \sum_{x\in\cup_{i=1}^n A_i} 1.$$

则对任意 $x\in\cup_{i=1}^n A_i$，其在式(1)的左侧被计数 1 次，即“贡献”为 1. 同样的，我们考虑 x 在式(1)的右侧的“贡献”：对任意 $x\in\cup_{i=1}^n A_i$，设其恰好属于 k 个集合，即存在 $i_1,i_2,\ldots,i_k\in\{1,2,\ldots,n\}$，使得 $x\in A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}$，且对任意 $m\notin\{i_1,i_2,\cdots,i_k\}$ 有 $x\notin A_m$. 则

- 对右端的第一个求和号，x 在其中的贡献为 $\binom{k}{1}$；
- 对右端的第二个求和号，x 在其中的贡献为 $-\binom{k}{2}$（由于求和号前面是负号，因此贡献为负）；
- $\cdots$
- 对右端的第 n 个求和号，x 在其中的贡献为 $(-1)^{n-1}\binom{k}{n}$.

故 x 在右端的总贡献即为

$$\begin{aligned}&\binom{k}{1}-\binom{k}{2}+\cdots+(-1)^{n-1}\binom{k}{n}\\ =&-\left(\binom{k}{0}-\binom{k}{1}+\binom{k}{2}+\cdots+(-1)^k\binom{k}{k}\right)+\binom{k}{0}\\ =&(1-1)^k+1=1.\end{aligned}$$

因此，任意元素 x 在式(1)左右两端的贡献相等，则式(1)成立.

□

例 2　计算 1 至 300 中与 300 互素的数的个数.

解　对 300 进行质因数分解，我们得到 $300=5^2\times3\times2^2$，则与 300 互素等价于不含 $2,3,5$ 为素因子. 记

$$\begin{aligned}I&=\{1,2,3,\ldots,300\},\\ A&=\{x|x\in I,2\mid x\},\\ B&=\{x|x\in I,3\mid x\},\\ C&=\{x|x\in I,5\mid x\}.\end{aligned}$$

则目标转化为求 $\left|\bar{A}\cap\bar{B}\cap\bar{C}\right|$.

由容斥定理，上式即为

$$\begin{aligned}|\bar{A}\cap\bar{B}\cap\bar{C}|&=|I|-|A|-|B|-|C|+|A\cap B|+|A\cap C|+|B\cap C|-|A\cap B\cap C|\\ &=300-\frac{300}{2}-\frac{300}{3}-\frac{300}{5}+\frac{300}{2\times3}+\frac{300}{2\times5}+\frac{300}{3\times5}-\frac{300}{2\times3\times5}\\ &=80.\end{aligned}$$

注　将上题中的 300 改为一般的 $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$，其中 $p_1,p_2,\ldots,p_k$ 为 n 的素因子. 由容斥定理可知，与 n 互素的数的个数

$$\begin{aligned}\varphi(n)&:=n-\frac{n}{p_1}-\frac{n}{p_2}-\cdots-\frac{n}{p_k}+\frac{n}{p_1p_2}+\cdots+\frac{n}{p_{k-1}p_k}-\cdots\\ &=n\left(1-\frac{1}{p_1}-\frac{1}{p_2}-\cdots-\frac{1}{p_k}+\frac{1}{p_1p_2}+\cdots+\frac{1}{p_{k-1}p_k}-\cdots\right)\\ &=n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_k}\right).\end{aligned}$$

$\varphi(n)$ 被称为**欧拉函数**.

例 3 计算 1 至 300 中素数的个数.

解 注意到，一个数 n 是合数当且仅当其能被写作两个数之积，因此 n 一定有不大于 $\sqrt{n}$ 的素因子. 不大于 $\sqrt{300}$ 的素数有 $2,3,5,7,11,13,17$. 记

$$\begin{aligned}
I &= \{1,2,3,\dots,300\}, \\
A_1 &= \{x|x\in I, 2\mid x\}, \\
A_2 &= \{x|x\in I, 3\mid x\}, \\
A_3 &= \{x|x\in I, 5\mid x\}, \\
A_4 &= \{x|x\in I, 7\mid x\}, \\
A_5 &= \{x|x\in I, 11\mid x\}, \\
A_6 &= \{x|x\in I, 13\mid x\}, \\
A_7 &= \{x|x\in I, 17\mid x\}.
\end{aligned}$$

则 I 中素数个数为

$$\left|\bigcap_{i=1}^{7}\bar{A}_i\right| - 1 + 7 = |I| - \sum_{1\le i\le 7}|A_i| + \sum_{1\le i<j\le 7}|A_i\cap A_j| - \cdots + (-1)^7\left|\bigcap_{i=1}^{7}A_i\right| - 1 + 7 = 62,$$

其中上式中的 -1 是由于 1 被计算在内且其不是素数，$+7$ 是由于 2，3，5，7，11，13，17 这七个素数未被计入.

例 4 试求不定方程 $x_1+x_2+x_3+x_4=20$ 的解的个数，其中 $x_1\le 7, x_2\le 6, x_3\le 8$, 且 x_1,x_2,x_3,x_4 均为正整数.

解 记

$$\begin{aligned}
I &= \left\{(x_1,x_2,x_3,x_4)\middle|x_1,x_2,x_3,x_4\in\mathbb{N}^+, x_1+x_2+x_3+x_4=20\right\}, \\
A &= \left\{(x_1,x_2,x_3,x_4)\middle|x_1,x_2,x_3,x_4\in\mathbb{N}^+, x_1+x_2+x_3+x_4=20, x_1\ge 8\right\}, \\
B &= \left\{(x_1,x_2,x_3,x_4)\middle|x_1,x_2,x_3,x_4\in\mathbb{N}^+, x_1+x_2+x_3+x_4=20, x_2\ge 7\right\}, \\
C &= \left\{(x_1,x_2,x_3,x_4)\middle|x_1,x_2,x_3,x_4\in\mathbb{N}^+, x_1+x_2+x_3+x_4=20, x_3\ge 9\right\}.
\end{aligned}$$

则所求为

$$|\bar{A}\cap\bar{B}\cap\bar{C}| = |I| - |A| - |B| - |C| + |A\cap B| + |A\cap C| + |B\cap C| - |A\cap B\cap C|.$$

注意到 $|A| = |\{(x_1,x_2,x_3,x_4)|x_1,x_2,x_3,x_4\in\mathbb{N}^+, x_1+x_2+x_3+x_4=13\}|$（在方程中用 x_1-7 代替 x）. 同理 $|A\cap B| = |\{(x_1,x_2,x_3,x_4)|x_1,x_2,x_3,x_4\in\mathbb{N}^+, x_1+x_2+x_3+x_4=7\}|$ 相等. 则

$$\begin{aligned}
|\bar{A}\cap\bar{B}\cap\bar{C}| &= |I| - |A| - |B| - |C| + |A\cap B| + |A\cap C| + |B\cap C| - |A\cap B\cap C| \\
&= \binom{19}{3} - \binom{12}{3} - \binom{13}{3} - \binom{11}{3} + \binom{6}{3} + \binom{4}{3} + \binom{5}{3} - 0 \\
&= 332.
\end{aligned}$$

例 5 现有 n 个信封以及与之一一对应的 n 张信纸，现将每个信封装入一张信纸，试求所有信封装的均不是原本对应信纸的情况总数.

解 记 I 为放置信纸的所有情况，由乘法原理 $|I| = n!$. 记 A_i 为第 i 个信封装入与之对应的信纸时，放置信纸的所有情况，故只需求 $\left|\cap_{i=1}^{n} \bar{A}_i\right|$.

注意到，在有 k 个信封放置了与之对应的信纸后，剩余的所有情况数即为 $(n-k)!$，因此

$$
\begin{aligned}
\left|\bigcap_{i=1}^{n} \bar{A}_i\right| &= |I| - \sum_{i=1}^{n} |A_i| + \sum_{1\le i<j\le n} |A_i \cap A_j| - \cdots + (-1)^n \left|\bigcap_{i=1}^{n} A_i\right| \\
&= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots + (-1)^n \\
&= \sum_{k=0}^{n} (-1)^k \binom{n}{k}(n-k)! \\
&= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \\
&\sim n! e^{-1}.
\end{aligned}
$$

注 此问题被称作装错信封问题，也被称作全错位排列问题. 该问题最早由瑞士数学家约翰 · 伯努利提出，瑞士数学家欧拉对此问题进行了研究并给出了解答.

3 选做题

1. 考虑整数轴 $\mathbb{Z}$ 上动点 Q 的离散时间随机游走，Q 每个时刻以概率 $1/2$ 向右移动一格，以 $1/2$ 概率向左移动一格。直至 Q 走到 0 时，随机游走结束。设 Q 的起始位置是 1000000，动点 Q 一定能回到原点吗?

2. 考虑 $\mathbb{Z}^2$ 上动点 Q 的离散时间随机游走，Q 从原点出发，每个时刻分别以 $1/4$ 的概率向 4 个方向移动一格。动点 Q 一定能回到原点吗?

3. 进一步扩展到 $\mathbb{Z}^d$ 上动点 Q 的离散时间随机游走，Q 从原点出发，每个时刻分别以 $1/2d$ 的概率向 $2d$ 个方向移动一格。动点 Q 一定能回到原点吗?

4. 如果去掉以上问题中的各方向等概率条件，试估计动点 Q 回到原点的概率。