9.12 随机子集

工作纸

题目

STEP III 2019，第 12 题

继续阅读之前，请先看题目（原题为英文）。题目载于 OCR 提供下载的 2019 年 STEP 试卷（zip 文件）内的 STEP III 试卷：

www.ocr.org.uk/Images/696590-2019-past-papers.zip

提示

请先自行尝试解题。如果遇到困难，请逐一阅读提示，每读完一个便回到题目再试。标有*停一停，试一试*的提示是关键步骤。每个提示之后均留有空位供你演算。

> ? **想一想**　怎样用 n 条是非题的一串答案，来描述 $\{1, \ldots, n\}$ 的一个子集？

阶段 1　一个元素

涵盖：第 (i) 部分

> ? **想一想**　把每个不含 1 的子集（subset）与加入 1 后所得的子集配对。含有 1 的子集占全部子集的几分之几？

阶段 2 互斥集合

涵盖：第 (ii) 部分

停一停，试一试 A_1 与 A_2 互斥（disjoint），当且仅当没有任何元素同时属于两者。就单一个元素而言，列出四种等可能（equally likely）的情况。哪一种是不容许的？为甚么可以把 n 个元素的结果相乘？

想一想 若有 m 个集合，某个指定元素同时属于所有集合的概率是多少？

阶段 3　子集

涵盖：第 (iii) 部分

想一想　若要 $A_1 \subseteq A_2$，一个元素的四种情况之中，哪一种是不容许的?

停一停，试一试　对于一条链（chain）$A_1 \subseteq A_2 \subseteq A_3$，列出一个元素“属于或不属于”各集合的 8 种模式。一旦某元素属于链中的一个集合，它还必定属于哪些集合？数出容许的模式有多少种，然后对 m 个集合作同样的处理。

背后的数学

不用概率的计数

把每个概率乘以选择的总数，就得到一个计数。满足 $A_1 \subseteq A_2$ 的有序对 (A_1, A_2) 的数目是 $4^n \times (\tfrac{3}{4})^n = 3^n$。你也可以直接数算这些有序对：先选定 A_2 的大小 k，然后以 $\binom{n}{k}$ 种方法选出 A_2 本身，再从它的 2^k 个子集中任选一个作为 A_1。所以

$$\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n,$$

这正是 $(1+2)^n$ 的二项式定理（binomial theorem）。用两种方法数算同一样东西，一次按元素、一次按大小，便证明了一条恒等式：这就是算两次（double counting）的方

法，你在 Section 7.9 讨论柏拉图立体时已经见过。同样地，链 $A_1 \subseteq \cdots \subseteq A_m$ 的数目是 $(m+1)^n$。

子集的格

$\{1, \ldots, n\}$ 的所有子集按包含关系排序，形成一个称为布尔格（Boolean lattice）的结构，可以把它想像成一个 n 维立方体：每个子集是一个顶点，以它那一串是非题的答案作标记，而沿着一条棱移动，就是加入或移走一个元素。链就是向上走的路径。有些深刻的结果描述了在给定条件下，子集族可以有多大。施佩纳定理（Sperner's theorem，1928 年）指出，在没有一个集合包含另一个集合的子集族之中，最大的一族有 $\binom{n}{\lfloor n/2 \rfloor}$ 个集合；艾狄胥—柯—拉多定理（Erdős–Ko–Rado theorem，1961 年）则说明，最多可以选出多少个 k 元集合，使得任意两个都有共同元素。两者都是组合数学（combinatorics）入门课程的标准内容，而把子集看成掷硬币的观点，给出了其中一些最漂亮的证明。

布尔的代数

布尔（George Boole）的 *The Laws of Thought*（《思维规律》，1854 年）把逻辑命题当作集合，把“与”、“或”、“非”当作交集、并集和补集。把一个集合编码成一串是非题的答案，正是电脑表示集合的方法：$\{1, \ldots, 64\}$ 的一个子集就是单独一个 64 位元的数，而交集就是“逐位元 AND”（bitwise AND）指令。本题的方法，即每次只看一个元素（一个位元），正是分析这类运算的方法。

延伸练习

1 **覆盖。** 求 $\mathrm{P}(A_1 \cup A_2 = S)$ 和 $\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_m = S)$。

2 **大小相同。** 证明 $\mathrm{P}(|A_1| = |A_2|) = \binom{2n}{n}/4^n$。（提示：$|A_1| = |A_2|$ 当且仅当 $|A_1| + |S \setminus A_2| = n$。）证明当 n 很大时，这个概率约为 $1/\sqrt{\pi n}$。

3 **分成三份。** 证明 S 的不相交子集所组成的有序对 (A, B) 的数目是 3^n，并求由不相交子集所组成的无序对 $\{A, B\}$ 的数目。

4 **小 n 的施佩纳定理。** 对 $n = 3$ 和 $n = 4$，求 $\{1, \ldots, n\}$ 的子集族中，没有一个集合包含另一个集合的最大一族。

延伸阅读

- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），其中关于子集和偏序集（partially ordered sets）的章节。

- **Béla Bollobás,** ***Combinatorics: Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability***（Cambridge University Press）。一本简短而优雅的书，讲述子集族，由施佩纳定理说起。
- **Peter Cameron,** ***Combinatorics: Topics, Techniques, Algorithms*** (Cambridge University Press)。一门入门课程，特别着重用两种方法计数。

完整的引导式解答见第 9.12 节：

https://beyondstep.keithmcnulty.org/zh-hans/ch09/random-subsets.html

每星期一道 STEP 题目，完整讲解：The STEP Up，以电子邮件免费发送，订阅网址：keithmcnulty.substack.com