详细解答

STEP II 样本试卷，第 16 题

概率与统计 • 淘汰赛的概率

STEP II 样本试卷，第 16 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/spec-s2-q16

本题在 *Beyond STEP* 第 12.7 节淘汰赛中谁会遇上谁中有完整的引导式讲解：https://beyondstep.keithmcnulty.org/zh-hans/ch12/knockout.html

关键思路

逐轮追踪两位选定的选手 A 和 B。在任何一轮中，只要两人仍未出局又未曾交手，就只有两件事要紧：A 是否抽中与 B 对赛？若否，两人是否都胜出？这两者都是简单的概率（probability），而由此得出的乘积会漂亮地逐项相消。至于最后一部分，还有一个巧妙的对称性（symmetry）论证：数一数整个淘汰赛中的比赛场数。

解答

记 $N = 2^n$，并称两位选定的选手为 A 和 B。经过 $r-1$ 轮后，余下 2^{n-r+1} 位选手，所以第 r 轮有 $M_r = 2^{n-r+1}$ 位选手，而第 n 轮（此时 $M_n = 2$）就是决赛。

我们会反复用到以下事实：若 A 和 B 同在 M 位随机配对的选手之中，则 A 的对手等可能地是其余 $M-1$ 位选手中的任何一位，所以

$$\mathrm{P}(A \text{ 抽中与 } B \text{ 对赛}) = \frac{1}{M-1}.$$

（决赛不设抽签，但代入 $M = 2$ 时公式给出 1，结果正确。）

第 (i) 部分

在第一轮中 $M_1 = N$，所以

$$\mathrm{P}(A \text{ 和 } B \text{ 在第一轮相遇}) = \frac{1}{2^n - 1}.$$

第 (ii) 部分

设 S_r 为“A 和 B 在第 r 轮开始时都仍在赛事之中”这个事件（event）（这必然表示两人尚未交手，因为每场比赛的负方即告出局）。在 S_r 发生的条件下，两人一同晋身第 $r+1$ 轮，当且仅当他们在第 r 轮没有抽中对赛，然后各自在（不同的）比赛中胜出。两场比赛的结果互相独立（independent），胜出的概率各为 $\frac{1}{2}$，所以

$$\mathrm{P}(S_{r+1} \mid S_r) = \frac{M_r - 2}{M_r - 1} \cdot \frac{1}{4}.$$

? **想一想** 怎样把它化成一个我们真正算得出来的乘积?

记 $M = 2^k$，并留意 $M - 2 = 2(2^{k-1} - 1)$，所以

$$\frac{M-2}{M-1} = \frac{2\,(2^{k-1} - 1)}{2^k - 1}.$$

某一轮的分子（撇除因子 2 不计）正是下一轮的分母。由于 S_1 是必然事件，把 $k = n, n-1, \ldots, n-r+2$ 各轮的结果相乘，便得到一个裂项相消的乘积（telescoping product）：

$$\mathrm{P}(S_r) = \frac{1}{4^{r-1}} \prod_{k=n-r+2}^{n} \frac{2\,(2^{k-1} - 1)}{2^k - 1} = \frac{2^{r-1}}{4^{r-1}} \cdot \frac{2^{n-r+1} - 1}{2^n - 1} = \frac{2^{n-r+1} - 1}{2^{r-1}\,(2^n - 1)}.$$

在 S_r 发生的条件下，两人在第 r 轮相遇的概率为 $1/(M_r - 1) = 1/(2^{n-r+1} - 1)$，所以

$$\mathrm{P}(A \text{ 和 } B \text{ 在第 } r \text{ 轮相遇}) = \frac{1}{2^{r-1}\,(2^n - 1)}.$$

这条公式简洁得令人欣喜：每过一轮，在该轮相遇的机会就减半。取 $r = 1$ 便重新得到第 (i) 部分的结果。至于决赛，取 $r = n$：

$$\mathrm{P}(A \text{ 和 } B \text{ 在决赛相遇}) = \frac{1}{2^{n-1}\,(2^n - 1)}.$$

再用对称性核对一下是否合理：赛制对每位选手一视同仁，所以决赛的两位选手等可能地是 $\binom{N}{2} = 2^{n-1}(2^n - 1)$ 对选手中的任何一对，由此得出相同的答案。

第 (iii) 部分

对 $r = 1, \ldots, n$，“在第 r 轮相遇”这些事件互斥（mutually exclusive）（两位选手最多只能交手一次，因为负方会离场），所以我们可利用等比级数（geometric series）把概率相

加：

$$\mathrm{P}(A \text{ 和 } B \text{ 相遇}) = \frac{1}{2^n-1}\sum_{r=1}^{n}\frac{1}{2^{r-1}} = \frac{1}{2^n-1}\cdot 2\left(1-\frac{1}{2^n}\right)$$
$$= \frac{2}{2^n}\cdot\frac{2^n-1}{2^n-1} = \frac{1}{2^{n-1}}.$$

想一想 有没有方法可以完全不用求和就看出这个答案？

有。每场比赛恰好淘汰一位选手，而除冠军外所有人都会被淘汰，所以比赛总是恰好有 2^n-1 场。每场比赛由一对选手进行，而根据对称性，$\binom{N}{2}$ 对选手中每一对交手的概率都相同，设为 p。比赛场数的期望值（expected value）等于所有选手对的这些概率之和，所以

$$\binom{2^n}{2}p = 2^n-1 \quad\Longrightarrow\quad p = \frac{2(2^n-1)}{2^n(2^n-1)} = \frac{2}{2^n} = \frac{1}{2^{n-1}},$$

与求和所得的结果一致。

验算

- $n=1$：只有两位选手，他们必定在第一轮相遇，而这一轮就是决赛。三条公式都给出 1。
- $n=2$：四位选手。第 (i) 部分给出 $\frac{1}{3}$，决赛给出 $\frac{1}{6}$，而总和是 $\frac{1}{3}+\frac{1}{6}=\frac{1}{2}$，与第 (iii) 部分所说的相符。
- $n=3$：把每一种抽签结果和每一种赛果逐一列举，准确算出 $\frac{1}{7}$、$\frac{1}{28}$ 和 $\frac{1}{4}$。以 16 位选手模拟 2×10^5 次淘汰赛，得出 0.0664、0.0085 和 0.1247，而准确值则为 $\frac{1}{15}\approx0.0667$、$\frac{1}{120}\approx0.0083$ 和 $\frac{1}{8}$。

常见错误

- 忘记了两位选手若要在较后的一轮相遇，必须先在之前的每一轮都避开对方：因子 $\frac{M-2}{M-1}$ 很容易漏掉。
- 把第 (ii) 部分答作 $\left(\frac{1}{2^{n-1}}\right)^2$，即假如两人互相独立时双双打入决赛的机会。但他们并不独立：若两人早早抽中对赛，最多只有一人能够晋级。

STEP 之外

第 (iii) 部分的论证，其实是配合指示变量（indicator variable）使用的期望值的线性性质（linearity of expectation）：不论各事件是否独立，比赛场数的期望值都等于所有选手

对各自交手的概率之和。这是概率论和组合数学（combinatorics）中最强大的工具之一，也是用来证明具有指定性质的对象确实存在的“概率方法”（probabilistic method）的基础。Blitzstein 与 Hwang 合著的 *Introduction to Probability*（概率入门教科书）有一整章精彩的例子，示范以指示变量立论的方法。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。