详细解答

STEP II 2025，第 11 题

概率与统计 · 期望值与二项式求和

STEP II 2025，第 11 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 2025 年 STEP 试卷（zip 文件）内的 STEP II 试卷：

www.ocr.org.uk/Images/750486-2025-past-papers.zip

关键思路

两个游戏都在第一次掷出 $2k$ 时结束，所以两者都可以写成对掷骰次数 r 的求和。Ali 的和近似几何级数（geometric series），而第 (i) 部分正是计算掷骰次数期望值（expected value）所需的工具。至于 Zen，一局胜出的游戏由 $2k$ 之前出现了哪些数字决定，因为一旦知道这个集合，递增的次序就随之确定。数算这些集合会引入二项式系数（binomial coefficients），所以这个和就是第 (iii) 部分的那个和。

解答

第 (i) 部分

当 $|x| < 1$ 时，几何级数给出

$$\sum_{r=0}^{\infty} x^r = \frac{1}{1-x}.$$

幂级数（power series）在其收敛区间（interval of convergence）内可以逐项微分（differentiate term by term），因此两边对 x 求导，得

$$\sum_{r=1}^{\infty} r x^{r-1} = \frac{1}{(1-x)^2}, \qquad |x| < 1,$$

即为所求。

第 (ii) 部分

每一次掷骰的结果可分为三类：

$$\mathrm{P}(\text{掷出}2k) = \frac{1}{2k}, \qquad \mathrm{P}(\text{偶数，但不是}2k) = \frac{k-1}{2k}, \qquad \mathrm{P}(\text{奇数}) = \frac{1}{2}.$$

设 R 为掷骰次数。Ali 以 $R = r$ 胜出，当且仅当第 $1, \ldots, r-1$ 次掷出的都是 $2k$ 以外的偶数，而第 r 次掷出 $2k$。各次掷骰互相独立（independent），所以设 $x = \dfrac{k-1}{2k}$，便有

$$\mathrm{P}(\text{Ali 胜出且} R = r) = \frac{1}{2k}\, x^{r-1}.$$

(a) 对几何级数求和（注意 $0 \le x < 1$）：

$$\mathrm{P}(\text{Ali 胜出}) = \frac{1}{2k} \cdot \frac{1}{1-x} = \frac{1}{2k} \cdot \frac{2k}{k+1} = \frac{1}{k+1}.$$

这个结果有一个快捷的看法。$2k$ 以外的偶数不影响结果，可以不理会。游戏的胜负取决于奇数和 $2k$ 哪一个先出现。共有 k 个奇数面和一个 $2k$ 面，出现机会均等，所以 $2k$ 先出现的概率是 $\frac{1}{k+1}$。

(b) 若 Ali 以 $R = r$ 胜出，他赚得 £r，否则一无所得，所以由第 (i) 部分，

$$\mathrm{E}(\text{Ali 的收入}) = \sum_{r=1}^{\infty} r \cdot \frac{1}{2k}\, x^{r-1} = \frac{1}{2k} \cdot \frac{1}{(1-x)^2} = \frac{1}{2k}\left(\frac{2k}{k+1}\right)^2 = \frac{2k}{(k+1)^2} \text{ 英镑.}$$

第 (iii) 部分

把该表达式记为 S。它的通项是 $(r+1)\binom{n}{r}x^r$，其中 $r = 0, 1, \ldots, n$。因子 $r+1$ 提示我们对 x^{r+1} 求导：

$$S = \sum_{r=0}^{n} (r+1)\binom{n}{r} x^r = \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{r=0}^{n} \binom{n}{r} x^{r+1}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(x(1+x)^n\Big),$$

这里用了二项式定理（binomial theorem）。由乘积法则（product rule），

$$S = (1+x)^n + nx(1+x)^{n-1} = (1+x)^{n-1}(1+(n+1)x).$$

第 (iv) 部分

假设 Zen 恰好掷了 r 次而胜出。最后一次掷出 $2k$，而首 $r-1$ 次掷出的是取自 $\{1, 2, \ldots, 2k-1\}$ 的严格递增（strictly increasing）的数。（无论如何它们都必须小于 $2k$，因为游戏在第一次掷出 $2k$ 时便停止。）

> **想一想** 长度为 r 的胜出序列有多少个?

一个严格递增序列，只要知道其中出现了哪些数，便完全确定，因为把一组互不相同的数按递增次序排列只有一种方法。所以长度为 r 的胜出序列与 $\{1, \ldots, 2k-1\}$ 的 $(r-1)$ 元子集（subset）一一对应。这样的子集共有 $\binom{2k-1}{r-1}$ 个，且 $1 \le r \le 2k$。

每一个特定的 r 次掷骰序列的概率都是 $\left(\frac{1}{2k}\right)^r$。因此

$$\mathrm{E}(\text{Zen 的收入}) = \sum_{r=1}^{2k} r\binom{2k-1}{r-1}\left(\frac{1}{2k}\right)^r = \frac{1}{2k}\sum_{j=0}^{2k-1}(j+1)\binom{2k-1}{j}\left(\frac{1}{2k}\right)^j,$$

其中 $j = r - 1$。这正是第 (iii) 部分中 $n = 2k-1$、$x = \frac{1}{2k}$ 时的和，此时 $1 + (n+1)x = 1 + 2k \cdot \frac{1}{2k} = 2$。所以

$$\text{E(Zen 的收入)} = \frac{1}{2k}\left(1 + \frac{1}{2k}\right)^{2k-2} \cdot 2 = \frac{1}{k}\left(1 + \frac{1}{2k}\right)^{2k-2} \text{ 英镑.}$$

第 (v) 部分

把 Zen 的预期收入除以 Ali 的预期收入：

$$\rho = \frac{1}{k}\left(1 + \frac{1}{2k}\right)^{2k-2} \cdot \frac{(k+1)^2}{2k} = \frac{(k+1)^2}{2k^2} \cdot \frac{\left(1 + \frac{1}{2k}\right)^{2k}}{\left(1 + \frac{1}{2k}\right)^{2}}.$$

现在设 k 很大。那么

- $\dfrac{(k+1)^2}{k^2} = \left(1 + \dfrac{1}{k}\right)^2 \approx 1$；
- $\left(1 + \dfrac{1}{2k}\right)^2 \approx 1$；
- 由题目给出的近似式，取 $n = 2k$，得 $\left(1 + \dfrac{1}{2k}\right)^{2k} \approx \mathrm{e}$。

所以

$$\rho \approx \frac{\mathrm{e}}{2} \approx \frac{2.71828}{2} = 1.359\ldots,$$

即当 k 很大时，Zen 的预期收入约为 Ali 的 1.359 倍，也就是多出稍高于 35%，即为所求。

验算

- 当 $k = 1$ 时，骰子就是一枚两面为 1, 2 的硬币。Ali 只有在第一次掷出 2 时才胜出，概率为 $\frac{1}{2} = \frac{1}{k+1}$，此时赚得 £1，所以预期收入为 $\frac{1}{2} = \frac{2k}{(k+1)^2}$。Zen 以序列 2（概率 $\frac{1}{2}$，赚得 £1）和 1, 2（概率 $\frac{1}{4}$，赚得 £2）胜出，所以她的预期收入是 $\frac{1}{2} + \frac{1}{2} = 1$。公式与此吻合：$k = 1$ 时指数 $2k - 2$ 为 0，所以公式给出 $\frac{1}{1} \cdot 1 = 1$。
- 当 $k = 3$ 时，Ali 的预期收入是 $\frac{3}{8}$，Zen 的是 $\frac{1}{3}\left(\frac{7}{6}\right)^4 = \frac{2401}{3888} \approx 0.618$。两者都与直接列举所有胜出序列的结果相符，也与游戏的电脑模拟结果相符。
- 比率 ρ 在 $k = 10$ 时约为 1.456，在 $k = 100$ 时约为 1.369，在 $k = 10^4$ 时约为 1.3592。它递减趋向 $\mathrm{e}/2 = 1.35914\ldots$。

常见错误

- 在第 (iv) 部分用阶乘去数有序序列，而不是数子集。每个集合只有一种递增排列。
- 忘记 Zen 游戏中较早的掷骰不能包含 $2k$。这使范围变成 $\{1,\ldots,2k-1\}$，所以 $n = 2k-1$，而不是 $2k$。
- 在第 (ii) 部分把继续游戏的概率写成 $\frac{k}{2k}$，而不是 $\frac{k-1}{2k}$。掷出 $2k$ 虽然是偶数，但它会结束游戏。

STEP 之外

第 (iii) 部分是*生成函数*（generating functions）的一个小例子：把一个数列编码为一个幂级数的系数，然后对这个级数运用微积分。概率生成函数（probability generating functions）把这个想法更进一步。若 $G(s) = \mathrm{E}(s^X)$，则 $\mathrm{E}(X) = G'(1)$，这正是本题两次用到的求导技巧。极限 $\left(1+\frac{1}{n}\right)^n \to \mathrm{e}$ 在概率论中随处可见，例如以泊松分布近似二项分布（Poisson approximation to the binomial）。Grimmett 与 Welsh 的 *Probability: An Introduction*（《概率导论》）有一章很好地讲解生成函数。

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。