详细解答

STEP II 2013，第 13 题

概率与统计 • 掷硬币中的游程

> **STEP II 2013，第 13 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/13-s2-q13

关键思路

事件 $S = n$ 完全确定了首 $n+1$ 次掷币的结果：先是 n 次相同的结果，然后一次打断这段连串（run）的结果。$A = n$ 亦是如此：先是 n 次交替（alternating）出现的结果，然后一次重复。因此每个概率（probability）都是两个乘积之和，各对应第一次掷币的一种可能结果。比较 S 与 A 便归结为将四项之差因式分解（factorisation）；若我们把 p 和 q 当作两个独立的字母，而不是以 $1-p$ 代替 q，因式分解会较容易看出来。

解答

全文中，各次掷币互相独立（independent），每次得正面的概率为 p，得反面的概率为 q，且 $p, q > 0$，$p \neq q$。

一般公式

$S = n$ 发生，当且仅当第 $1, \ldots, n$ 次掷币全部与第一次相同，而第 $n+1$ 次不同。若第一次是正面，结果便是 $\mathrm{H}^n\mathrm{T}$，概率为 $p^n q$；若第一次是反面，结果便是 $\mathrm{T}^n\mathrm{H}$，概率为 $q^n p$。这两种情况互斥（disjoint），所以

$$\mathrm{P}(S = n) = p^n q + q^n p = pq\left(p^{n-1} + q^{n-1}\right).$$

$A = n$ 发生，当且仅当第 $1, \ldots, n$ 次掷币交替出现，而第 $n+1$ 次与第 n 次相同，从而结束交替。

▶ 当 $n = 2m$（偶数）：连串 $\mathrm{HTHT}\cdots\mathrm{HT}$ 有 m 个正面和 m 个反面，并以 T 结尾，所以

下一次必须是 T，得 $p^m q^m \cdot q$。以反面开始则得 $q^m p^m \cdot p$。因此

$$\mathrm{P}(A = 2m) = p^m q^m (p + q).$$

- 当 $n = 2m + 1$（奇数）：连串 HTH$\cdots$H 有 $m + 1$ 个正面和 m 个反面，并以 H 结尾，所以下一次必须是 H，得 $p^{m+2} q^m$。以反面开始则得 $q^{m+2} p^m$。因此

$$\mathrm{P}(A = 2m + 1) = p^m q^m \left(p^2 + q^2\right).$$

第 (i) 部分

$A = 1$ 表示第二次掷币重复第一次的结果：HH 或 TT。所以 $\mathrm{P}(A = 1) = p^2 + q^2$（即上面 $m = 0$ 的情况）。$S = 1$ 表示第二次掷币与第一次不同：HT 或 TH，所以

$$\mathrm{P}(S = 1) = pq + qp = 2pq.$$

于是

$$\mathrm{P}(A = 1) - \mathrm{P}(S = 1) = p^2 - 2pq + q^2 = (p - q)^2 > 0,$$

由于 $p \neq q$，不等号是严格的。所以 $\mathrm{P}(S = 1) < \mathrm{P}(A = 1)$。

第 (ii) 部分

由公式，

$$\mathrm{P}(S = 2) = p^2 q + q^2 p = pq(p + q) = \mathrm{P}(A = 2),$$

此即所求。至于长度 3，

$$\mathrm{P}(S = 3) = p^3 q + q^3 p = pq\left(p^2 + q^2\right) = \mathrm{P}(A = 3),$$

所以这两者亦相等。

第 (iii) 部分

偶数长度。设 $n > 1$。将两个概率都除以 $pq > 0$，

$$\frac{\mathrm{P}(S = 2n)}{pq} = p^{2n-1} + q^{2n-1}, \qquad \frac{\mathrm{P}(A = 2n)}{pq} = p^{n-1} q^{n-1}(p + q) = p^n q^{n-1} + p^{n-1} q^n.$$

? **想一想** 怎样才能判断 $p^{2n-1} + q^{2n-1} - p^n q^{n-1} - p^{n-1} q^n$ 的正负号?

将每个纯幂与含同一字母最高次幂的交叉项配对：

$$\begin{aligned} p^{2n-1} + q^{2n-1} - p^n q^{n-1} - p^{n-1} q^n &= p^n\left(p^{n-1} - q^{n-1}\right) - q^n\left(p^{n-1} - q^{n-1}\right) \\ &= \left(p^n - q^n\right)\left(p^{n-1} - q^{n-1}\right). \end{aligned}$$

由于 $p, q > 0$ 且 $p \neq q$，对任何 $k \geq 1$，$p^k - q^k$ 均非零，且与 $p - q$ 同号（因为 $x \mapsto x^k$ 在 $x > 0$ 上严格递增（strictly increasing））。这里 $n \geq 2$，所以 n 和 $n-1$ 都至少是 1，两个括号均非零且同号。它们的乘积为正，再乘回 $pq > 0$，得

$$\mathrm{P}(S = 2n) > \mathrm{P}(A = 2n) \qquad (n > 1).$$

（当 $n = 1$ 时，第二个括号是 $p^0 - q^0 = 0$，这正是第 (ii) 部分的等式。）

奇数长度。同样地，仍设 $n > 1$，

$$\frac{\mathrm{P}(S = 2n+1)}{pq} = p^{2n} + q^{2n},$$

$$\frac{\mathrm{P}(A = 2n+1)}{pq} = p^{n-1}q^{n-1}\left(p^2 + q^2\right) = p^{n+1}q^{n-1} + p^{n-1}q^{n+1},$$

而

$$\begin{aligned} p^{2n} + q^{2n} - p^{n+1}q^{n-1} - p^{n-1}q^{n+1} &= p^{n+1}\left(p^{n-1} - q^{n-1}\right) - q^{n+1}\left(p^{n-1} - q^{n-1}\right) \\ &= \left(p^{n+1} - q^{n+1}\right)\left(p^{n-1} - q^{n-1}\right). \end{aligned}$$

根据同样的论证，两个括号均非零且与 $p - q$ 同号，所以

$$\mathrm{P}(S = 2n+1) > \mathrm{P}(A = 2n+1) \qquad (n > 1).$$

同样，因子 $p^{n-1} - q^{n-1}$ 在 $n = 1$ 时为零，与 $\mathrm{P}(S = 3) = \mathrm{P}(A = 3)$ 吻合。

总括而言：长度为 1 的相同连串比交替连串出现的机会小，长度 2 和 3 的机会相等，而由 4 起的每个长度，相同连串出现的机会都较大。

验算

- 取 $p = \frac{1}{3}$，$q = \frac{2}{3}$。则 $\mathrm{P}(S = 4) = pq(p^3 + q^3) = \frac{2}{9} \cdot \frac{1}{3} = \frac{2}{27} \approx 0.074$，而 $\mathrm{P}(A = 4) = p^2q^2 = \frac{4}{81} \approx 0.049$，所以 $S = 4$ 的机会确实较大。
- 将 $\mathrm{P}(S = n)$ 对所有 $n \geq 1$ 求和，得 $q \cdot \frac{p}{1-p} + p \cdot \frac{q}{1-q} = p + q = 1$，理应如此。
- 每条公式都已用电脑核实：对多个 p 值，以精确分数列举所有 14 次掷币的序列，验证长度至 13 的情况。

常见错误

- 忘记那次结束连串的掷币。$S = n$ 需要第 $n+1$ 次掷币不同；缺了这一步，你计算的其实是 $\mathrm{P}(S \geq n)$。

- 交替连串的结尾字母弄错：奇数长度的连串以开头的同一字母结尾，偶数长度的则以另一字母结尾。
- 在第 (iii) 部分代入 $q = 1 - p$：这样各式会变成 p 的多项式，因式分解难看得多。

STEP 之外

S 是一个几何型随机变量（geometric-type random variable）：给定第一次掷币的结果，连串长度就是等到第一次改变的等候时间（waiting time）。这类连串是统计学中游程检定（runs test）的基础，该检定用来判断一串结果看来是否随机；而生成函数（generating functions）则提供一个简洁的方法，求出最长连串的完整分布。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率导论》）以合适的程度讲解生成函数和等候时间。

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。