詳細解答

STEP II 2013，第 13 題

概率與統計 • 擲硬幣中的游程

> **STEP II 2013，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/13-s2-q13

關鍵思路

事件 $S = n$ 完全確定了首 $n+1$ 次擲幣的結果：先是 n 次相同的結果，然後一次打斷這段連串（run）的結果。$A = n$ 亦是如此：先是 n 次交替（alternating）出現的結果，然後一次重複。因此每個概率（probability）都是兩個乘積之和，各對應第一次擲幣的一種可能結果。比較 S 與 A 便歸結為將四項之差因式分解（factorisation）；若我們把 p 和 q 當作兩個獨立的字母，而不是以 $1-p$ 代替 q，因式分解會較容易看出來。

解答

全文中，各次擲幣互相獨立（independent），每次得正面的概率為 p，得反面的概率為 q，且 $p, q > 0$，$p \neq q$。

一般公式

$S = n$ 發生，當且僅當第 $1, \ldots, n$ 次擲幣全部與第一次相同，而第 $n+1$ 次不同。若第一次是正面，結果便是 $\mathrm{H}^n\mathrm{T}$，概率為 $p^n q$；若第一次是反面，結果便是 $\mathrm{T}^n\mathrm{H}$，概率為 $q^n p$。這兩種情況互斥（disjoint），所以

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

$A = n$ 發生，當且僅當第 $1, \ldots, n$ 次擲幣交替出現，而第 $n+1$ 次與第 n 次相同，從而結束交替。

▶ 當 $n = 2m$（偶數）：連串 $\mathrm{HTHT}\cdots\mathrm{HT}$ 有 m 個正面和 m 個反面，並以 T 結尾，所以

下一次必須是 T，得 $p^m q^m \cdot q$。以反面開始則得 $q^m p^m \cdot p$。因此

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

- 當 $n = 2m + 1$（奇數）：連串 HTH$\cdots$H 有 $m + 1$ 個正面和 m 個反面，並以 H 結尾，所以下一次必須是 H，得 $p^{m+2} q^m$。以反面開始則得 $q^{m+2} p^m$。因此

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

第 (i) 部分

$A = 1$ 表示第二次擲幣重複第一次的結果：HH 或 TT。所以 $\mathrm{P}(A = 1) = p^2 + q^2$（即上面 $m = 0$ 的情況）。$S = 1$ 表示第二次擲幣與第一次不同：HT 或 TH，所以

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

於是

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

由於 $p \neq q$，不等號是嚴格的。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。至於長度 3，

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

所以這兩者亦相等。

第 (iii) 部分

偶數長度。設 $n > 1$。將兩個概率都除以 $pq > 0$，

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

? **想一想** 怎樣才能判斷 $p^{2n-1} + q^{2n-1} - p^n q^{n-1} - p^{n-1} q^n$ 的正負號？

將每個純冪與含同一字母最高次冪的交叉項配對：

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

由於 $p, q > 0$ 且 $p \neq q$，對任何 $k \geq 1$，$p^k - q^k$ 均非零，且與 $p - q$ 同號（因為 $x \mapsto x^k$ 在 $x > 0$ 上嚴格遞增（strictly increasing））。這裏 $n \geq 2$，所以 n 和 $n-1$ 都至少是 1，兩個括號均非零且同號。它們的乘積為正，再乘回 $pq > 0$，得

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

（當 $n = 1$ 時，第二個括號是 $p^0 - q^0 = 0$，這正是第 (ii) 部分的等式。）

奇數長度。同樣地，仍設 $n > 1$，

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

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

而

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

根據同樣的論證，兩個括號均非零且與 $p - q$ 同號，所以

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

同樣，因子 $p^{n-1} - q^{n-1}$ 在 $n = 1$ 時為零，與 $\mathrm{P}(S = 3) = \mathrm{P}(A = 3)$ 吻合。

總括而言：長度為 1 的相同連串比交替連串出現的機會小，長度 2 和 3 的機會相等，而由 4 起的每個長度，相同連串出現的機會都較大。

驗算

- 取 $p = \frac{1}{3}$，$q = \frac{2}{3}$。則 $\mathrm{P}(S = 4) = pq(p^3 + q^3) = \frac{2}{9} \cdot \frac{1}{3} = \frac{2}{27} \approx 0.074$，而 $\mathrm{P}(A = 4) = p^2q^2 = \frac{4}{81} \approx 0.049$，所以 $S = 4$ 的機會確實較大。
- 將 $\mathrm{P}(S = n)$ 對所有 $n \geq 1$ 求和，得 $q \cdot \frac{p}{1-p} + p \cdot \frac{q}{1-q} = p + q = 1$，理應如此。
- 每條公式都已用電腦核實：對多個 p 值，以精確分數列舉所有 14 次擲幣的序列，驗證長度至 13 的情況。

常見錯誤

- 忘記那次結束連串的擲幣。$S = n$ 需要第 $n+1$ 次擲幣不同；缺了這一步，你計算的其實是 $\mathrm{P}(S \geq n)$。

- 交替連串的結尾字母弄錯：奇數長度的連串以開頭的同一字母結尾，偶數長度的則以另一字母結尾。
- 在第 (iii) 部分代入 $q = 1 - p$：這樣各式會變成 p 的多項式，因式分解難看得多。

STEP 之外

S 是一個幾何型隨機變量（geometric-type random variable）：給定第一次擲幣的結果，連串長度就是等到第一次改變的等候時間（waiting time）。這類連串是統計學中遊程檢定（runs test）的基礎，該檢定用來判斷一串結果看來是否隨機；而生成函數（generating functions）則提供一個簡潔的方法，求出最長連串的完整分佈。Grimmett 和 Welsh 的 *Probability: An Introduction*（《概率導論》）以合適的程度講解生成函數和等候時間。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。