詳細解答

STEP III 2002，第 12 題

概率與統計 • 以條件化求期望值

> **STEP III 2002，第 12 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/02-s3-q12

關鍵思路

按遊戲如何開始來取條件（conditioning）。每一局遊戲的開頭不外乎三種：H、TH 或 TT。在前兩種情況下，並沒有取得任何有用的進展（連續反面已被打斷），所以遊戲等於從頭再來，而餘下部分的期望長度仍然是 E。這種「更新」（renewal）給出一條關於 E 的線性方程，解之即可。同樣的想法——按第一個正面何時出現來分類——也能處理連續 r 次反面的情況。

解答

全文中，p 是擲出正面的概率，$q = 1 - p$ 是擲出反面的概率，各次投擲互相獨立（independent）。我們假設 $0 < p < 1$，這樣遊戲以概率 1 結束，而且下文所有期望值（expectation）都是有限的。

解釋關於 E 的方程

考慮最初的一兩次投擲。以下三種情況恰好有一種發生：

- **先擲出 H**（概率為 p）。已用去一次投擲，而且沒有正在累積的連續反面。由於各次投擲互相獨立，遊戲餘下的部分就是整個遊戲的一個全新複本，期望長度為 E。期望總長度：$1 + E$。
- **先 T 後 H**（概率為 qp）。已用去兩次投擲，而正面再次把連續反面清零，所以遊戲重新開始。期望總長度：$2 + E$。
- **先 T 後 T**（概率為 q^2）。遊戲恰好在 2 次投擲後結束。

這三種情況互斥（exclusive）而且窮盡（exhaustive）所有可能（它們的概率之和為 $p + qp + q^2 = p + q = 1$）。根據全期望定律（law of total expectation），

$$E = p(1 + E) + (1 - p)p(2 + E) + 2(1 - p)^2,$$

這正是題目給出的方程。

求 E

把含 E 的項合併：

$$E\,(1 - p - qp) = p + 2qp + 2q^2.$$

左邊，$1 - p - qp = q - qp = q(1 - p) = q^2$。右邊，$p + 2q(p + q) = p + 2q = 1 + q$。由於 $q \neq 0$，

$$E = \frac{1 + q}{q^2} = \frac{2 - p}{(1 - p)^2}.$$

期望獎金

設 W 為正面次數的期望值，亦即以英鎊計的期望獎金。沿用同樣的三種情況。擲出 H 之後，我們已贏得 £1，遊戲重新開始；擲出 TH 之後，我們同樣已贏得 £1，遊戲重新開始；擲出 TT 之後，我們一無所獲，遊戲結束。所以

$$W = p(1 + W) + qp(1 + W) + q^2 \cdot 0.$$

? **想一想** 在 TH 的情況中，為甚麼是 $1 + W$ 而不是 W？

在這種情況下，第二次投擲*確實*是正面，所以在遊戲重新開始之前已經賺得 £1。這一題最容易出錯的地方，就是漏掉了它。

合併各項，得 $W(1 - p - qp) = p(1 + q)$，所以 $Wq^2 = p(1 + q)$，而

$$W = \frac{p(1 + q)}{q^2} = \frac{p(2 - p)}{(1 - p)^2} \text{ 英鎊}.$$

留意 $W = pE$：平均而言，所有投擲中有 p 這個比例是正面，即使遊戲停止的時刻取決於投擲的結果。

連續 r 次反面的遊戲

現在設 E 為遊戲在連續出現 r 次反面時停止的情況下，投擲次數的期望值。按第一個正面來取條件。要麼最初 r 次投擲全是反面，要麼第一個正面出現在第 $j + 1$ 次投擲，其中 j 滿足 $0 \leq j \leq r - 1$：

- j **次反面後一次正面**（概率為 $q^j p$）。這用去了 $j+1$ 次投擲，連續反面已被打斷，遊戲重新開始。期望總長度：$j+1+E$。
- r **次反面**（概率為 q^r）。遊戲恰好在 r 次投擲後結束。

所以

$$E = \sum_{j=0}^{r-1} q^j p\,(j+1+E) + rq^r.$$

右邊 E 的係數是 $p\sum_{j=0}^{r-1} q^j = p\cdot\dfrac{1-q^r}{1-q} = 1-q^r$，因為 $1-q=p$。把它移到左邊，

$$q^r E = p\sum_{j=0}^{r-1}(j+1)q^j + rq^r. \qquad (*)$$

想一想 沒有公式表在手，怎樣求 $\sum(j+1)q^j$ 的和？

記 $S=\sum_{j=0}^{r-1}(j+1)q^j = 1+2q+3q^2+\cdots+rq^{r-1}$，然後減去 $qS = q+2q^2+\cdots+(r-1)q^{r-1}+rq^r$：

$$(1-q)S = 1+q+q^2+\cdots+q^{r-1}-rq^r = \frac{1-q^r}{p} - rq^r.$$

因此 $pS = \dfrac{1-q^r}{p} - rq^r$，代入 $(*)$ 後，rq^r 兩項互相抵消：

$$q^r E = \frac{1-q^r}{p} - rq^r + rq^r = \frac{1-q^r}{p}.$$

於是

$$E = \frac{1-q^r}{pq^r},$$

這就是所要證明的結果。當 $r=2$ 時，上式為 $\dfrac{(1-q)(1+q)}{pq^2} = \dfrac{1+q}{q^2}$，與第一個遊戲的結果一致。

驗算

- 若硬幣是公平的，$p=q=\frac{1}{2}$：第一個遊戲平均持續 $E=\frac{3/2}{1/4}=6$ 次投擲，贏得 $W=$£3，正好是 E 的一半，合乎預期。
- $r=1$：遊戲在第一次出現反面時停止，所以長度服從幾何分佈（geometric distribution），平均值為 $1/q$；公式給出 $\dfrac{1-q}{pq} = \dfrac{1}{q}$。
- 當 $q\to 0$（幾乎每次投擲都是正面）時，E 和 W 都趨向無窮大，這是理所當然的。

- 我們以分數作精確計算，對 $r = 1, \ldots, 7$ 及若干個 p 值解出「目前已連續出現 k 次反面」這一串狀態的方程，確認了這些公式。另外，以 $p = 0.4$ 模擬 2×10^5 局遊戲，在 $r = 2$ 時得出平均 4.438 次投擲（理論值 4.444）及獎金 £1.772（理論值 £1.778），在 $r = 3$ 時得出 9.076 次投擲（理論值 9.074）。

常見錯誤

- 在獎金的方程中，把 $qp(1 + W)$ 寫成 qpW，結果得出錯誤答案 p/q^2。
- 只按第一次投擲來取條件：擲出一次 T 之後，遊戲並不會重新開始，因為連續反面中的一次已經到手。這正是題目所給方程要往前看兩次投擲的原因。

STEP 之外

「一旦出現正面，未來的情形就與開局時完全一樣」這個論證，是馬可夫鏈（Markov chain）的*更新法*或*首步分析法*（first-step method），一般情況下的期望到達時間（expected hitting time）正是這樣計算的。$W = pE$ 這個簡潔的事實是沃爾德恆等式（Wald's identity）的一個特例：把隨機個數 N 個獨立同分佈（independent, identically distributed）的項相加，其和的期望值等於 $\mathbb{E}[N]$ 乘以單獨一項的平均值，前提是 N 為停時（stopping time）。這兩個課題都可以在 Grimmett 與 Welsh 的 *Probability: An Introduction*（概率論入門），或 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率論導論）中讀到。

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。