詳細解答

STEP III 2007，第 13 題

概率與統計 • 隨機漫步的期望值

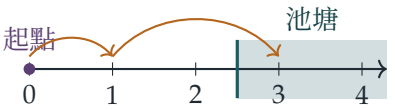
> **STEP III 2007，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/07-s3-q13

關鍵思路

那半米的作用，是消除「剛好落在池邊」的任何疑問：青蛙由距離池邊 $(n-\frac{1}{2})$ m 處出發，牠身處池中，當且僅當牠已跳的總距離至少為 n。因此整個問題的核心，就是由 1 和 2 組成的累計總和首次達到 n 的時刻。當 n 較小時，我們可以直接列出所有可能情況；至於一般公式，u_n 的三個值可給出三條方程，用以求出三個常數。

解答

設 S_j 為首 j 次跳躍所跳的總距離（以米為單位），並設 T 為到達池塘所需的跳躍次數。青蛙由距離池邊 $(n-\frac{1}{2})$ m 處出發，跳了 j 次後身處池中，當且僅當 $S_j > n-\frac{1}{2}$，亦即 $S_j \geq n$，因為 S_j 是整數。所以 $p_n(j) = \mathrm{P}(T=j)$，其中 T 是首個滿足 $S_j \geq n$ 的 j。



圖中所示為 $n=3$ 的情況（池邊距離起點 2.5 m），以及「短跳、長跳」這個跳躍序列，青蛙跳兩次便到達池塘。

第 (i) 部分

當 $n=2$ 時，池邊距離起點 1.5 m。若第一跳是長跳（概率（probability）為 q），青蛙便直接跳進池中，所以 $p_2(1)=q$。否則第一跳是短跳（概率為 p），青蛙距離池邊 0.5 m，

而第二跳無論長短都會把牠帶進池中。各次跳躍互相獨立（independent），所以

$$p_2(2) = p \times 1 = p,$$

這正是所要證明的結果。（此外，當 $j \geq 3$ 時 $p_2(j) = 0$，而 $p + q = 1$ 則確認了這個分佈（distribution）是完整的。）

第 (ii) 部分

由距離池邊 $\frac{1}{2}$ m 處出發，任何一跳都會落入池中，所以 $u_1 = 1$。

利用第 (i) 部分，

$$u_2 = 1 \cdot p_2(1) + 2 \cdot p_2(2) = q + 2p = 2 - q.$$

當 $n = 3$ 時，我們需要 $S_j \geq 3$。

- 只跳一次永遠不夠，因為 $S_1 \leq 2$：$p_3(1) = 0$。
- 跳兩次仍不夠，只會在兩次都是短跳時發生（$S_2 = 2$）；其餘每一種組合都有 $S_2 \geq 3$。所以 $p_3(2) = 1 - p^2$。
- 兩次短跳之後，青蛙距離池邊 0.5 m，第三跳必定落入池中：$p_3(3) = p^2$。

因此

$$\begin{aligned} u_3 &= 2(1 - p^2) + 3p^2 = 2 + p^2 \\ &= 2 + (1 - q)^2 = 3 - 2q + q^2, \end{aligned}$$

這正是所要證明的結果。

想一想 有沒有更快的方法看出這些值？這個方法能否推廣至所有 n？

有：以第一跳的結果作條件（conditioning）來考慮。第一跳之後，青蛙的處境與原來相同，只是 n 減少了 1 或 2，所以當 $n \geq 2$ 時

$$u_n = 1 + p\,u_{n-1} + q\,u_{n-2}, \qquad u_0 = 0,$$

其中 $u_0 = 0$ 對應由 1.5 m 處出發而第一跳是長跳的情況。由此再次得出 $u_2 = 1 + p$ 和 $u_3 = 1 + p(1 + p) + q = 2 + p^2$。這也解釋了第 (iii) 部分的形式為何正確：輔助方程（auxiliary equation）$\lambda^2 - p\lambda - q = 0$ 可因式分解為 $(\lambda - 1)(\lambda + q) = 0$，由此得出 $(-q)^{n-1}$ 這一項。而重根 1（源自那個「+1」）則解釋了 $B + Cn$ 的出現。

第 (iii) 部分

把 $n = 1, 2, 3$ 代入 $u_n = A(-q)^{n-1} + B + Cn$，並利用第 (ii) 部分：

$$A + B + C = 1, \tag{1}$$

$$-qA + B + 2C = 2 - q, \tag{2}$$

$$q^2A + B + 3C = 3 - 2q + q^2. \tag{3}$$

兩式相減以消去 B：

$$(2) - (1): \quad -(1+q)A + C = 1 - q, \tag{4}$$

$$(3) - (2): \quad q(1+q)A + C = 1 - q + q^2. \tag{5}$$

現在 $q \times (4) + (5)$ 可消去 A：

$$(1+q)C = q - q^2 + 1 - q + q^2 = 1, \qquad \text{所以} \qquad C = \frac{1}{1+q},$$

這正是所要證明的結果。然後由 (4) 得

$$(1+q)A = \frac{1}{1+q} - (1-q) = \frac{1-(1-q^2)}{1+q} = \frac{q^2}{1+q}, \qquad \text{所以} \qquad A = \frac{q^2}{(1+q)^2},$$

再由 (1) 得

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

所以

$$u_n = \frac{q^2(-q)^{n-1} + q + (1+q)n}{(1+q)^2}.$$

n 很大時。 由於 $0 \le q \le 1$，$A(-q)^{n-1}$ 和 B 這兩項都是有界（bounded）的（各自的絕對值最多為 $\frac{1}{4}$），而 Cn 則無止境地增大。因此當 n 很大時

$$u_n \approx Cn = \frac{n}{1+q} = \frac{n}{p+2q},$$

這裏用到 $1 = p + q$，故 $1 + q = p + 2q$。

為何這在意料之中。 每一跳的平均長度是 $1 \cdot p + 2 \cdot q = p + 2q$ 米。青蛙須跳過大約 n 米，而根據大數定律（law of large numbers），當 N 很大時，N 次跳躍的總長度接近 $N(p+2q)$。所以所需的跳躍次數接近 $n/(p+2q)$。這個近似相當準確，因為起點那半米的偏移，以及終點處多跳出的距離（青蛙落點最多超出池邊 1.5 m）都是有界的。它們只會令次數改變一個有界的量，與 n 相比微不足道。這個有界的修正，正正就是 $A(-q)^{n-1} + B$ 這部分。

驗算

- 在公式中取 $n = 0$，得 $\left(q^2 \cdot (-1/q) + q\right)/(1+q)^2 = 0 = u_0$，與應有的結果相符，而且公式滿足上面的遞推關係（recurrence）。
- $q = 0$（所有跳躍都是短跳）：$u_n = n$，這是正確的，因為青蛙恰好需要 n 次單位長度的跳躍。
- 以分數精確枚舉所有跳躍序列，結果在 $n = 1, \ldots, 12$ 及若干個 p 值下均與公式吻合。以 $q = 0.7$、$n = 200$ 進行模擬，得出平均跳躍次數為 117.90，而公式給出 117.89，$n/(p+2q) \approx 117.65$。

常見錯誤

- 把「落入池中」當作恰好 $S_j = n$。由於有那半米，青蛙可以跳過頭，凡是滿足 $S_j \geq n$ 的序列都要計算在內。
- 寫成 $p_3(2) = 2pq + q^2$ 時，忘記了長跳在先的情況。事實上 $pq + qp + q^2 = 1 - p^2$ 已涵蓋兩種次序。

STEP 之外

u_n 的遞推關係是一條更新方程（renewal equation）。結果 $u_n \sim n/\mu$（其中 μ 為每步的平均長度）就是基本更新定理（elementary renewal theorem）。由沃爾德恆等式（Wald's identity）$\mathrm{E}(S_T) = \mu\,\mathrm{E}(T)$ 可得出一個精確的版本：這裏 S_T 不是 n 就是 $n+1$，所以 $(1+q)u_n = n + \mathrm{P}(\text{跳過頭})$。Grimmett 與 Welsh 合著的 *Probability: An Introduction*（概率論入門）以同樣的思路講述隨機漫步（random walk）和母函數（generating function）。

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。