詳細解答

STEP II 1997，第 12 題

概率與統計 • 二項分佈的奇偶概率

> **STEP II 1997，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/97-s2-q12

關鍵思路

只需追蹤訊息此刻是對還是錯。每聽錯一次，訊息便對錯互換，所以傳回 Albert 的訊息正確，當且僅當耳語中出錯的次數是偶數。要抽出二項展開式（binomial expansion）中的偶數項，正好可以用 $(a+b)^n + (a-b)^n$，因為奇數項會互相抵消。至於修改後的遊戲，要留意一旦失敗一次，Albert 傳出去的便是錯的訊息，所以從那時起，每一輪都需要有奇數次聽錯。

解答

三名玩家

若有 n 名玩家，便有 n 次耳語：Albert 傳給 Bertha，Bertha 傳給下一位，如此類推繞一圈傳到 Zebedee，然後 Zebedee 傳回 Albert。每次耳語被正確聽到的概率（probability）是 p，而且各次互相獨立（independent）。Albert 聽到的是否真相，只取決於有多少次耳語被聽錯，與硬幣擲出的結果無關。

三名玩家時有三次耳語，每次正確的概率是 $\frac{2}{3}$。傳到 Albert 的訊息正確，條件是沒有耳語出錯，或者恰好有兩次出錯（第二次錯誤把第一次抵消了）。所以

$$\Pr(\text{獲勝}) = \left(\tfrac{2}{3}\right)^3 + 3\left(\tfrac{1}{3}\right)^2\left(\tfrac{2}{3}\right) = \frac{8}{27} + \frac{6}{27} = \frac{14}{27}.$$

n 名玩家

設 $q = 1 - p$。聽錯的耳語次數服從 $\mathrm{B}(n, q)$ 分佈（distribution），而當這個次數是偶數時遊戲便獲勝：

$$P = \sum_{k\text{ 為偶數}} \binom{n}{k} q^k p^{n-k}.$$

現在

$$(p+q)^n + (p-q)^n = \sum_{k=0}^{n} \binom{n}{k} p^{n-k} q^k (1 + (-1)^k) = 2 \sum_{k\text{ 為偶數}} \binom{n}{k} p^{n-k} q^k,$$

因為 k 為奇數的項互相抵消，而 k 為偶數的項則出現兩次。由於 $p + q = 1$ 及 $p - q = 2p - 1$，

$$\boxed{P = \tfrac{1}{2}(1 + (2p-1)^n).}$$

當 $n = 3$、$p = \frac{2}{3}$ 時，上式給出 $\frac{1}{2}(1 + \frac{1}{27}) = \frac{14}{27}$，與第一部分的結果一致。

修改後的遊戲

設 N 為 Albert 向 Bertha 耳語的次數。把訊息繞圈傳足一周稱為一輪，所以每一輪都由 Albert 耳語開始。

> **想一想** 一輪失敗之後，下一輪是否只是把第一輪原封不動地重來一次？

不是。在第一輪，Albert 傳出的是正確的結果。如果他從 Zebedee 那裏聽錯了，他傳下去的是他聽到的內容，而那是錯的。由錯的訊息出發，這一輪結束時 Albert 聽到真相，當且僅當有奇數次耳語出錯，其概率為 $1 - P$。如果這一輪又失敗，Albert 便再一次聽到（並傳出）錯的訊息，所以之後的每一輪都與這一輪相同。由於各次聆聽互相獨立，各輪也互相獨立。

因此第一輪成功的概率是 P，而之後每一輪成功的概率是 $1 - P$，失敗的概率是 P。故此 $\Pr(N = 1) = P$，而對於 $k \geq 2$，

$$\Pr(N = k) = \underbrace{(1-P)}_{\text{第 1 輪失敗}} \underbrace{P^{k-2}}_{\text{第 2,…,}k-1\text{ 輪失敗}} \underbrace{(1-P)}_{\text{第 }k\text{ 輪成功}} = (1-P)^2 P^{k-2}.$$

假設 $0 < p < 1$，於是 $0 < P < 1$。那麼，在題目給出的和式中取 $x = P$，可得

$$\sum_{k=2}^{\infty} kP^{k-2} = \frac{1}{P}\left(\sum_{k=1}^{\infty} kP^{k-1} - 1\right) = \frac{1}{P}\left(\frac{1}{(1-P)^2} - 1\right),$$

因此

$$\begin{aligned}\mathrm{E}(N) &= P + (1-P)^2 \cdot \frac{1}{P}\left(\frac{1}{(1-P)^2} - 1\right)\\ &= P + \frac{1-(1-P)^2}{P}\\ &= P + (2-P) = 2.\end{aligned}$$

Albert 向 Bertha 耳語次數的期望值（expectation）是 2，不論玩家有多少人，也不論 p 是多少。

還有一個更快捷的方法看出這個結果。已知第一輪失敗，其後所需的輪數服從幾何分佈（geometric distribution），成功概率為 $1-P$，所以其平均值是 $1/(1-P)$。故此 $\mathrm{E}(N) = 1 + (1-P)\cdot\frac{1}{1-P} = 2$。

（邊界情況：若 $p = 1$，或 $p = 0$ 而 n 為偶數，則 $P = 1$，Albert 只耳語一次。若 $p = 0$ 而 n 為奇數，則 $P = 0$，第二輪必定成功，所以 $N = 2$。）

驗算

- 以準確分數列出聽對與聽錯的全部 2^n 種組合，對 $n \le 8$ 及若干個 p 值都得出 $\frac{1}{2}(1 + (2p-1)^n)$，當中包括三名玩家時的 $\Pr(\text{獲勝}) = \frac{14}{27}$。
- $p = \frac{1}{2}$ 給出 $P = \frac{1}{2}$：訊息純粹是雜訊，理應如此。
- 以 $(n, p) = (3, \frac{2}{3})$、$(5, 0.9)$ 及 $(4, 0.3)$ 模擬 2×10^5 局修改後的遊戲，所得的平均耳語次數分別為 2.001、1.996 及 1.994。

常見錯誤

- 把耳語次數數成 $n-1$ 而不是 n：最後一次耳語，即由 Zebedee 傳回 Albert 的那一次，也要計算在內。
- 把修改後的遊戲中每一輪的獲勝概率都當作 P。這樣會得出 $1/P$，是錯的，因為失敗之後 Albert 傳出的是錯的訊息。

STEP 之外

訊息的對／錯狀態是一條兩狀態的馬可夫鏈（Markov chain），而 $2p-1$ 是其轉移矩陣（transition matrix）的第二個特徵值（eigenvalue）。這就是為甚麼 $(2p-1)^n$ 決定了訊息「忘記」硬幣結果的速度。同一個模型稱為二元對稱通道（binary symmetric channel），是資訊論（information theory）和糾錯碼（error-correcting codes）的起點。Grimmett 與

Welsh 的 *Probability: An Introduction*（概率論入門）以這個程度講解生成函數（generating functions）和馬可夫鏈。

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。