詳細解答

STEP I 1990，第 14 題

概率與統計 • 期望值與最優停止

STEP I 1990，第 14 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/90-s1-q14

關鍵思路

由於每次抽出的球都會放回，各次抽球互相獨立（independent），所以每次抽到白球、紅球或黑球的概率都固定不變。對 Juggins 而言，我們只需知道連續三次抽球都「存活」的機會，以及一次存活抽球的平均得分。對 Muggins 而言，問題在於取捨：多抽一次，平均可多得少許分數，卻要冒失去已累積的全部分數的風險，而這個風險隨 N 增大而增大。比較「再抽一次的期望得分」與「就此停手的得分」，便得出一個臨界值（threshold）。

解答

在任何一次抽球中，設 W、R、B 分別為抽到白球、紅球、黑球的事件。由於球會放回，每次抽球都與其他各次獨立，且

$$\mathrm{P}(W)=\frac{5}{10}=\frac{1}{2},\qquad \mathrm{P}(R)=\frac{3}{10},\qquad \mathrm{P}(B)=\frac{2}{10}=\frac{1}{5}.$$

Juggins 的策略

Juggins 最終得零分，當且僅當他首三次抽球中出現黑球。否則三次抽球全都不是黑球，他便停手。因此

$$\mathrm{P}(\text{最終得分為0})=1-\mathrm{P}(\text{三次抽球均非黑球})=1-\left(\frac{4}{5}\right)^3=1-\frac{64}{125}=\frac{61}{125}.$$

至於期望得分（expected score），設 S 為他的最終得分。若首三次抽球全都不是黑球（概率為 $\frac{64}{125}$），則 $S=X_1+X_2+X_3$，其中 X_i 是已知第 i 次抽球不是黑球時該次的得分。已知某次抽球不是黑球，則它是白球的概率為 $\frac{1/2}{4/5}=\frac{5}{8}$，是紅球的概率為 $\frac{3/10}{4/5}=\frac{3}{8}$。因此

$$\mathrm{E}(X_i)=1\cdot\frac{5}{8}+2\cdot\frac{3}{8}=\frac{11}{8}.$$

在其餘所有情況下 $S = 0$。所以

$$\mathrm{E}(S) = \tfrac{64}{125} \cdot 3 \cdot \tfrac{11}{8} = \tfrac{264}{125} = 2.112.$$

想一想 為甚麼這裏用 $\frac{5}{8}$ 和 $\frac{3}{8}$，而不是 $\frac{1}{2}$ 和 $\frac{3}{10}$？

我們已經以「存活」為條件（conditioning），即已知三次抽球每次都不是黑球，所以各概率必須重新按比例調整，使其總和為 1。換一個不用條件概率的說法：$\mathrm{E}(S)$ 是對 $2^3 = 8$ 個白／紅序列求（概率）×（得分）的總和。每次抽球的期望貢獻為 $1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{10} = \frac{11}{10}$，而只有當另外兩次抽球也存活時（概率為 $\frac{16}{25}$），這份貢獻才得以保留。由此同樣得出 $3 \cdot \frac{11}{10} \cdot \frac{16}{25} = \frac{264}{125}$。

Muggins 的抉擇

假設 Muggins 現有 N 分，並再抽一次（然後停手）。他的得分會變成 $N+1$、$N+2$ 或 0，所以

$$\mathrm{E}(\text{再抽一次後的得分}) = \tfrac{1}{2}(N+1) + \tfrac{3}{10}(N+2) + \tfrac{1}{5} \cdot 0 = \tfrac{4}{5}N + \tfrac{11}{10}.$$

平均而言，再抽一次勝過停手，當且僅當

$$\tfrac{4}{5}N + \tfrac{11}{10} > N \iff \tfrac{11}{10} > \tfrac{1}{5}N \iff N < 5.5.$$

抽一次球的平均得益是 $\frac{11}{10}$，平均損失則是 $\frac{1}{5}N$，即抽到黑球的機會乘以所押上的分數。由於得分是整數，這提示了以下策略：

得分爲 5 或以下時繼續抽球；一旦達到 6 或以上便立即停手。

想一想 只比較「多抽一次」與「停手」，看來有點短視。在 $N \geq 6$ 時先抽一次，並打算稍後再抽，會不會更有利？

不會。得分一旦達到 6 或以上，其後任何一次抽球都只會令得分仍然是 6 或以上（或以 0 分結束遊戲），所以在往後每一個決策點，「只看一次抽球」的比較仍然會說「停手」。從這樣的得分出發，每多抽一次都會降低最終得分的期望值，所以由此開始的最佳方案就是立即停手。在 6 分以下，多抽一次已經勝過停手，而可以繼續抽下去這個選擇只會錦上添花。因此這個臨界值法則不僅合理，更是最佳策略。

採用這個策略，設 $V(N)$ 為由 N 分開始的最終得分期望值。則當 $N \geq 6$ 時 $V(N) = N$，而當 $N \leq 5$ 時 $V(N) = \frac{1}{2}V(N+1) + \frac{3}{10}V(N+2)$。由上而下逐一計算：

N	6	5	4	3	2	1	0
$V(N)$	6	5.1	4.35	3.705	3.1575	2.690	2.292

所以 Muggins 平均每局約得 2.29 分，而 Juggins 則是 2.112 分。

驗算

- 以電腦精確地列舉了所有最多三次抽球的序列，確認 $\mathrm{P}(0) = \frac{61}{125} = 0.488$ 及 $\mathrm{E}(S) = \frac{264}{125}$。以蒙地卡羅（Monte Carlo）方法模擬 2×10^5 局，得出 0.488 和 2.111。
- 在 $N = 5$ 時，抽一次得 $\frac{4}{5} \cdot 5 + 1.1 = 5.1 > 5$。在 $N = 6$ 時則得 $5.9 < 6$。所以轉捩點在 5 與 6 之間。
- 對所有策略進行價值迭代（value iteration），得出相同的法則。若改為在 ≥ 5 或 ≥ 7 時停手，平均得分較低，分別為 2.261 和 2.265，而在 6 時停手則為 2.292。

常見錯誤

- 把 $\mathrm{E}(S)$ 算成 $3 \times \frac{11}{10}$，忘記了其後抽到黑球會把先前的分數一筆勾銷。
- 把條件平均值 $\frac{11}{8}$ 再乘以無條件概率，以致把存活因子重複計算了一次。
- 把臨界值寫成 $N \leq 5.5$，卻沒有指出 N 是整數，所以法則應是「5 分時抽，6 分時停」。

STEP 之外

Muggins 的問題是一個最優停止（optimal stopping）問題。「當現在停手勝過多走一步時便停手」這個法則稱為一步前瞻（one-step look-ahead）法則；只要像本題一樣，法則指示停手的那些狀態所組成的集合一經進入便無法離開，這個法則就是最優的。同樣的想法可引申至馬可夫決策過程（Markov decision processes）和動態規劃（dynamic programming）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門）涵蓋期望值和馬可夫鏈（Markov chain）的背景知識，而 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論）則有許多類似的首步分析（first-step analysis）。

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。