詳細解答

STEP II 1999，第 14 題

概率與統計 • 最優停止

> **STEP II 1999，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/99-s2-q14

關鍵思路

骰子沒有記憶。每當你決定再擲一次，你就回到同一個遊戲的最開端，所以從那一刻起你所能期望的最好結果仍是同一個數 V，即整個遊戲的最佳期望得分。這樣一來，每次擲骰後的決定就變成一個簡單的比較：如果手上的點數勝過 V 就保留，否則再擲。把這一點寫下來，便得到一條關於 V 的方程，解出它就同時知道這個值和應採用的策略。

解答

設定

把擲出 1 的一擲稱為強制的：遊戲就此結束，得分為 1。擲出其他任何點數 $x \in \{2,3,4,5,6\}$ 之後，我們可選擇停止（得分為 x）或繼續。每個得分都介乎 1 與 6 之間，所以每個策略的期望得分（expected score）都在 $[1,6]$ 之內。設

$$V = \text{任何策略所能達到的最大期望得分}$$

（嚴格來說，是對所有策略取的最小上界（least upper bound）；我們稍後會看到它確實可以達到）。

一個得分為 4 的策略

考慮以下策略：一擲出 4、5 或 6 就停止（當然，擲出 1 時也要停止）。設 E 為其期望得分。以第一擲的結果為條件（condition）。擲出 $1,4,5,6$ 的概率（probability）各為 $\frac{1}{6}$，此時遊戲結束，得分就是該點數。擲出 2 或 3 的概率為 $\frac{2}{6}$；這時我們再擲，而由於骰子

沒有記憶，從那裏開始的期望得分仍然是 E。所以

$$E = \tfrac{1}{6}(1+4+5+6) + \tfrac{2}{6}E \quad\Longrightarrow\quad \tfrac{4}{6}E = \tfrac{16}{6} \quad\Longrightarrow\quad E = 4.$$

（遊戲以概率 1 結束，因為連擲 n 次而仍未停止的機會是 $(\tfrac{1}{3})^n \to 0$。）因此 $V \geq 4$。

沒有策略能勝過 4

任取一個策略；它甚至可以取決於之前所有擲骰結果的整段歷史。看看它的第一擲 x。

- 若 $x = 1$，得分為 1。
- 若 $x \geq 2$ 而該策略停止，得分為 x。
- 若 $x \geq 2$ 而該策略繼續，那麼它此後的做法只不過是對一個全新的同樣遊戲所用的某個策略（先前的擲骰結果不能影響之後的結果）。所以它從這裏開始的期望得分至多為 V。

在每一種情況下，已知 $x \geq 2$ 時的期望得分都至多為 $\max(x, V)$。因此每個策略的期望得分至多為

$$g(V) = \tfrac{1}{6}\Big(1 + \sum_{x=2}^{6} \max(x, V)\Big),$$

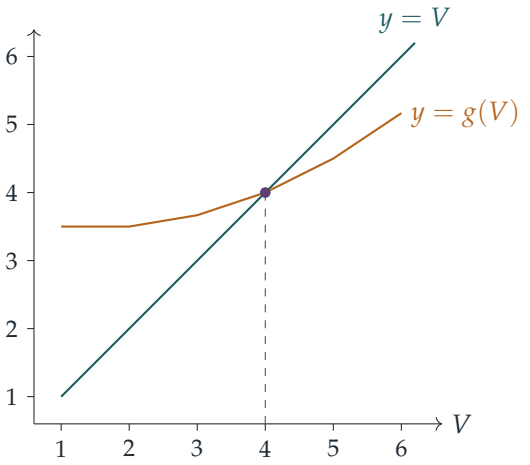
再對所有策略取最小上界，便得 $V \leq g(V)$。

? **想一想** 不等式 $V \leq g(V)$ 怎樣把 V 確定下來？

考慮 $h(V) = g(V) - V$。當 V 增加時，每一項 $\max(x, V)$ 的增長率是 0 或 1，而這樣的項只有五個，所以 g 的增長率至多為 $\frac{5}{6}$。因此 h 是嚴格遞減（strictly decreasing）的。在 $V = 4$ 時，

$$g(4) = \tfrac{1}{6}(1+4+4+4+4+5+6) = \tfrac{24}{6} = 4,$$

所以 $h(4) = 0$，而對每個 $V > 4$ 都有 $h(V) < 0$。由於 $h(V) \geq 0$，必定有 $V \leq 4$。



結論

綜合以上兩部分，$V = 4$：最佳的期望得分是 4，而做法是擲出 4、5 或 6 時停止，擲出 2 或 3 時再擲。這正是「若擲出的點數不小於 V 就停止」這條規則：2 或 3 比繼續下去所期望的 4 差，而 5 或 6 則較好。

擲出 4 是打和的情況。擲出 4 時停止得 4 分，繼續下去的期望得分也是 4。所以「只在 5 或 6 時停止」同樣好：用同樣的條件論證，

$$E' = \tfrac{1}{6}(1+5+6) + \tfrac{3}{6}E' \quad \Longrightarrow \quad E' = 4.$$

其他任何規則都嚴格地較差；例如擲出 3 或以上就停止，得 $\frac{1+3+4+5+6}{5} = \frac{19}{5}$，而只在 6 時停止，得 $\frac{1+6}{2} = \frac{7}{2}$。

驗算

對於任何形如「擲出 1 及集合 S 內的點數時停止」的規則，有一條巧妙的捷徑：每一擲令遊戲結束的概率都相同，而結束時的點數等可能地是 $\{1\} \cup S$ 中的任何一個，所以期望得分就是這些點數的平均值。對 $S = \{4,5,6\}$，它是 $\frac{16}{4} = 4$；對 $S = \{5,6\}$，它是 $\frac{12}{3} = 4$。我們精確地檢查了全部 32 條這樣的規則：只有這兩條得出 4，次佳的是 $\frac{19}{5}$。模擬 400 000 局遊戲，採用擲出 4 或以上就停止的規則，所得平均值為 3.998。

常見錯誤

- 把一擲的平均值 3.5 當作繼續下去的價值。繼續的價值高於 3.5，因為你保留了再擲的選擇權，所以「擲出 4 或以上就停止」不能靠與 3.5 比較來證明。

- 忘記了擲出 1 是強制的，於是把繼續的價值寫成彷彿可以毫無風險地一直擲下去。
- 只檢查閾值策略（threshold strategy），卻沒有解釋為何更巧妙、取決於歷史的規則不能做得更好。上面的無記憶性（memorylessness）論證涵蓋了所有策略。

STEP 之外

這是最優停止（optimal stopping）的一個小例子，而方程 $V = g(V)$ 是一條貝爾曼方程（Bellman equation），它是動態規劃（dynamic programming）和馬可夫決策過程（Markov decision process）的基本工具。同樣的推理可用來為金融中的美式期權（American option）定價，也可解決著名的祕書問題（secretary problem）。Grimmett 與 Welsh 的 *Probability: An Introduction* 涵蓋了這裏所用的條件論證，而 Blitzstein 與 Hwang 的 *Introduction to Probability* 則有許多同類型的題目。

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。