详细解答

STEP II 1999，第 14 题

概率与统计 · 最优停止

> **STEP II 1999，第 14 题**
>
> 继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：
>
> step.maths.org/questions/99-s2-q14

关键思路

骰子没有记忆。每当你决定再掷一次，你就回到同一个游戏的最开端，所以从那一刻起你所能期望的最好结果仍是同一个数 V，即整个游戏的最佳期望得分。这样一来，每次掷骰后的决定就变成一个简单的比较：如果手上的点数胜过 V 就保留，否则再掷。把这一点写下来，便得到一条关于 V 的方程，解出它就同时知道这个值和应采用的策略。

解答

设置

把掷出 1 的一掷称为强制的：游戏就此结束，得分为 1。掷出其他任何点数 $x \in \{2,3,4,5,6\}$ 之后，我们可选择停止（得分为 x）或继续。每个得分都介乎 1 与 6 之间，所以每个策略的期望得分（expected score）都在 $[1,6]$ 之内。设

$$V = \text{任何策略所能达到的最大期望得分}$$

（严格来说，是对所有策略取的最小上界（least upper bound）；我们稍后会看到它确实可以达到）。

一个得分为 4 的策略

考虑以下策略：一掷出 4、5 或 6 就停止（当然，掷出 1 时也要停止）。设 E 为其期望得分。以第一掷的结果为条件（condition）。掷出 $1,4,5,6$ 的概率（probability）各为 $\frac{1}{6}$，此时游戏结束，得分就是该点数。掷出 2 或 3 的概率为 $\frac{2}{6}$；这时我们再掷，而由于骰子没有

记忆，从那里开始的期望得分仍然是 E。所以

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

（游戏以概率 1 结束，因为连掷 n 次而仍未停止的机会是 $(\tfrac{1}{3})^n \to 0$。）因此 $V \geq 4$。

没有策略能胜过 4

任取一个策略；它甚至可以取决于之前所有掷骰结果的整段历史。看看它的第一掷 x。

- 若 $x = 1$，得分为 1。
- 若 $x \geq 2$ 而该策略停止，得分为 x。
- 若 $x \geq 2$ 而该策略继续，那么它此后的做法只不过是对一个全新的同样游戏所用的某个策略（先前的掷骰结果不能影响之后的结果）。所以它从这里开始的期望得分至多为 V。

在每一种情况下，已知 $x \geq 2$ 时的期望得分都至多为 $\max(x, V)$。因此每个策略的期望得分至多为

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

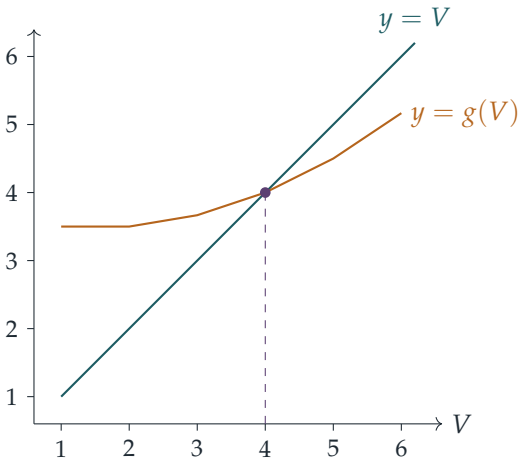
再对所有策略取最小上界，便得 $V \leq g(V)$。

? **想一想** 不等式 $V \leq g(V)$ 怎样把 V 确定下来?

考虑 $h(V) = g(V) - V$。当 V 增加时，每一项 $\max(x, V)$ 的增长率是 0 或 1，而这样的项只有五个，所以 g 的增长率至多为 $\frac{5}{6}$。因此 h 是严格递减（strictly decreasing）的。在 $V = 4$ 时，

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

所以 $h(4) = 0$，而对每个 $V > 4$ 都有 $h(V) < 0$。由于 $h(V) \geq 0$，必定有 $V \leq 4$。



结论

综合以上两部分，$V=4$：最佳的期望得分是 4，而做法是掷出 4、5 或 6 时停止，掷出 2 或 3 时再掷。这正是“若掷出的点数不小于 V 就停止”这条规则：2 或 3 比继续下去所期望的 4 差，而 5 或 6 则较好。

掷出 4 是打和的情况。掷出 4 时停止得 4 分，继续下去的期望得分也是 4。所以“只在 5 或 6 时停止”同样好：用同样的条件论证，

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

其他任何规则都严格地较差；例如掷出 3 或以上就停止，得 $\frac{1+3+4+5+6}{5} = \frac{19}{5}$，而只在 6 时停止，得 $\frac{1+6}{2} = \frac{7}{2}$。

验算

对于任何形如“掷出 1 及集合 S 内的点数时停止”的规则，有一条巧妙的捷径：每一掷令游戏结束的概率都相同，而结束时的点数等可能地是 $\{1\} \cup S$ 中的任何一个，所以期望得分就是这些点数的平均值。对 $S=\{4,5,6\}$，它是 $\frac{16}{4}=4$；对 $S=\{5,6\}$，它是 $\frac{12}{3}=4$。我们精确地检查了全部 32 条这样的规则：只有这两条得出 4，次佳的是 $\frac{19}{5}$。模拟 400 000 局游戏，采用掷出 4 或以上就停止的规则，所得平均值为 3.998。

常见错误

- 把一掷的平均值 3.5 当作继续下去的价值。继续的价值高于 3.5，因为你保留了再掷的选择权，所以“掷出 4 或以上就停止”不能靠与 3.5 比较来证明。

- 忘记了掷出 1 是强制的，于是把继续的价值写成仿佛可以毫无风险地一直掷下去。
- 只检查阈值策略（threshold strategy），却没有解释为何更巧妙、取决于历史的规则不能做得更好。上面的无记忆性（memorylessness）论证涵盖了所有策略。

STEP 之外

这是最优停止（optimal stopping）的一个小例子，而方程 $V = g(V)$ 是一条贝尔曼方程（Bellman equation），它是动态规划（dynamic programming）和马可夫决策过程（Markov decision process）的基本工具。同样的推理可用来为金融中的美式期权（American option）定价，也可解决著名的秘书问题（secretary problem）。Grimmett 与 Welsh 的 *Probability: An Introduction* 涵盖了这里所用的条件论证，而 Blitzstein 与 Hwang 的 *Introduction to Probability* 则有许多同类型的题目。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。