详细解答

STEP I 1990，第 14 题

概率与统计 · 期望值与最优停止

STEP I 1990，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/90-s1-q14

关键思路

由于每次抽出的球都会放回，各次抽球互相独立（independent），所以每次抽到白球、红球或黑球的概率都固定不变。对 Juggins 而言，我们只需知道连续三次抽球都“存活”的机会，以及一次存活抽球的平均得分。对 Muggins 而言，问题在于取舍：多抽一次，平均可多得少许分数，却要冒失去已累积的全部分数的风险，而这个风险随 N 增大而增大。比较“再抽一次的期望得分”与“就此停手的得分”，便得出一个临界值（threshold）。

解答

在任何一次抽球中，设 W、R、B 分别为抽到白球、红球、黑球的事件。由于球会放回，每次抽球都与其他各次独立，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 最终得零分，当且仅当他首三次抽球中出现黑球。否则三次抽球全都不是黑球，他便停手。因此

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

至于期望得分（expected score），设 S 为他的最终得分。若首三次抽球全都不是黑球（概率为 $\frac{64}{125}$），则 $S=X_1+X_2+X_3$，其中 X_i 是已知第 i 次抽球不是黑球时该次的得分。已知某次抽球不是黑球，则它是白球的概率为 $\frac{1/2}{4/5}=\frac{5}{8}$，是红球的概率为 $\frac{3/10}{4/5}=\frac{3}{8}$。因此

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

在其余所有情况下 $S = 0$。所以

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

想一想 为甚么这里用 $\frac{5}{8}$ 和 $\frac{3}{8}$，而不是 $\frac{1}{2}$ 和 $\frac{3}{10}$？

我们已经以“存活”为条件（conditioning），即已知三次抽球每次都不是黑球，所以各概率必须重新按比例调整，使其总和为 1。换一个不用条件概率的说法：$\mathrm{E}(S)$ 是对 $2^3 = 8$ 个白／红序列求（概率）×（得分）的总和。每次抽球的期望贡献为 $1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{10} = \frac{11}{10}$，而只有当另外两次抽球也存活时（概率为 $\frac{16}{25}$），这份贡献才得以保留。由此同样得出 $3 \cdot \frac{11}{10} \cdot \frac{16}{25} = \frac{264}{125}$。

Muggins 的抉择

假设 Muggins 现有 N 分，并再抽一次（然后停手）。他的得分会变成 $N+1$、$N+2$ 或 0，所以

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

平均而言，再抽一次胜过停手，当且仅当

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

抽一次球的平均得益是 $\frac{11}{10}$，平均损失则是 $\frac{1}{5}N$，即抽到黑球的机会乘以所押上的分数。由于得分是整数，这提示了以下策略：

> 得分为 5 或以下时继续抽球；一旦达到 6 或以上便立即停手。

想一想 只比较“多抽一次”与“停手”，看来有点短视。在 $N \geq 6$ 时先抽一次，并打算稍后再抽，会不会更有利？

不会。得分一旦达到 6 或以上，其后任何一次抽球都只会令得分仍然是 6 或以上（或以 0 分结束游戏），所以在往后每一个决策点，“只看一次抽球”的比较仍然会说“停手”。从这样的得分出发，每多抽一次都会降低最终得分的期望值，所以由此开始的最佳方案就是立即停手。在 6 分以下，多抽一次已经胜过停手，而可以继续抽下去这个选择只会锦上添花。因此这个临界值法则不仅合理，更是最佳策略。

采用这个策略，设 $V(N)$ 为由 N 分开始的最终得分期望值。则当 $N \geq 6$ 时 $V(N) = N$，而当 $N \leq 5$ 时 $V(N) = \frac{1}{2}V(N+1) + \frac{3}{10}V(N+2)$。由上而下逐一计算：

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 平均每局约得 2.29 分，而 Juggins 则是 2.112 分。

验算

- 以电脑精确地列举了所有最多三次抽球的序列，确认 $\mathrm{P}(0) = \frac{61}{125} = 0.488$ 及 $\mathrm{E}(S) = \frac{264}{125}$。以蒙地卡罗（Monte Carlo）方法模拟 2×10^5 局，得出 0.488 和 2.111。
- 在 $N = 5$ 时，抽一次得 $\frac{4}{5} \cdot 5 + 1.1 = 5.1 > 5$。在 $N = 6$ 时则得 $5.9 < 6$。所以转捩点在 5 与 6 之间。
- 对所有策略进行价值迭代（value iteration），得出相同的法则。若改为在 ≥ 5 或 ≥ 7 时停手，平均得分较低，分别为 2.261 和 2.265，而在 6 时停手则为 2.292。

常见错误

- 把 $\mathrm{E}(S)$ 算成 $3 \times \frac{11}{10}$，忘记了其后抽到黑球会把先前的分数一笔勾销。
- 把条件平均值 $\frac{11}{8}$ 再乘以无条件概率，以致把存活因子重复计算了一次。
- 把临界值写成 $N \leq 5.5$，却没有指出 N 是整数，所以法则应是“5 分时抽，6 分时停”。

STEP 之外

Muggins 的问题是一个最优停止（optimal stopping）问题。“当现在停手胜过多走一步时便停手”这个法则称为一步前瞻（one-step look-ahead）法则；只要像本题一样，法则指示停手的那些状态所组成的集合一经进入便无法离开，这个法则就是最优的。同样的想法可引申至马可夫决策过程（Markov decision processes）和动态规划（dynamic programming）。Grimmett 与 Welsh 的 *Probability: An Introduction*（概率入门）涵盖期望值和马可夫链（Markov chain）的背景知识，而 Blitzstein 与 Hwang 的 *Introduction to Probability*（概率导论）则有许多类似的首步分析（first-step analysis）。

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。