解答例

STEP I 1990、第 14 問

確率・統計 ・ 期待値と最適停止

> **STEP I 1990、問 14**
>
> 先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：
>
> step.maths.org/questions/90-s1-q14

鍵となる考え方

玉は毎回元に戻すので、各回の取り出しは独立（independent）であり、どの回も白・赤・黒が同じ一定の確率で出る。Juggins については、3 回の取り出しを生き残る確率と、生き残った 1 回あたりの平均得点さえ分かればよい。Muggins についてはトレードオフの問題になる。もう 1 回引けば平均的には少し得点が増えるが、それまでに積み上げた得点をすべて失う危険があり、その危険の大きさは N とともに大きくなる。「引いた場合の得点の期待値（expected value）」と「やめた場合の得点」を比べると、閾値（threshold）が得られる。

解答

各回の取り出しで、玉が白・赤・黒である事象をそれぞれ W, R, B とする。玉は元に戻すので、各回の取り出しは互いに独立であり、

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

である。

Juggins の戦略

Juggins の最終得点が 0 になるのは、最初の 3 回のうちに黒玉が出るとき、かつそのときに限る。そうでなければ 3 回とも黒以外であり、そこでやめる。よって

$$\mathrm{P}(\text{最終得点が0}) = 1 - \mathrm{P}(3\text{回とも黒以外}) = 1 - \left(\frac{4}{5}\right)^3 = 1 - \frac{64}{125} = \frac{61}{125}.$$

である。

得点の期待値を求めるため、最終得点を S とする。最初の 3 回がすべて黒以外であれば（確率 $\frac{64}{125}$）、$S = X_1 + X_2 + X_3$ である。ここで X_i は、i 回目が黒でないという条件のもとでの i 回目の得点である。1 回の取り出しが黒でないという条件のもとでは、白である確率は $\frac{1/2}{4/5} = \frac{5}{8}$、赤である確率は $\frac{3/10}{4/5} = \frac{3}{8}$ である。したがって

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

それ以外の場合はすべて $S = 0$ である。ゆえに

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

考えてみよう ここで $\frac{1}{2}$ と $\frac{3}{10}$ ではなく、$\frac{5}{8}$ と $\frac{3}{8}$ を用いるのはなぜか。

すでに生き残ることを条件としている（条件付き確率（conditional probability））ので、3 回の取り出しはいずれも黒でないと分かっており、確率の和が 1 になるように比率を調整し直さなければならない。同じことを条件付けなしで考えることもできる。$\mathrm{E}(S)$ は、白・赤からなる $2^3 = 8$ 通りの列すべてについての（確率）×（得点）の和である。各回の取り出しは期待値として $1 \cdot \frac{1}{2} + 2 \cdot \frac{3}{10} = \frac{11}{10}$ だけ寄与し、その得点が残るのは他の 2 回も生き残る場合（確率 $\frac{16}{25}$）に限られる。これより再び $3 \cdot \frac{11}{10} \cdot \frac{16}{25} = \frac{264}{125}$ が得られる。

Muggins の判断

Muggins が N 点を持っていて、もう 1 回引いて（そこでやめる）とする。得点は $N+1$, $N+2$, 0 のいずれかになるので、

$$\mathrm{E}(\text{もう 1 回引いた後の得点}) = \tfrac{1}{2}(N+1) + \tfrac{3}{10}(N+2) + \tfrac{1}{5} \cdot 0 = \tfrac{4}{5}N + \tfrac{11}{10}.$$

平均的に、引くほうがやめるより有利になるのは

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

のとき、かつそのときに限る。1 回引くことによる平均の利得は $\frac{11}{10}$ であり、平均の損失は $\frac{1}{5}N$、すなわち黒玉が出る確率と失うかもしれない得点との積である。得点は整数なので、次の戦略が示唆される。

得点が 5 以下である間は引き続け、6 以上になったらただちにやめる。

考えてみよう もう 1 回だけ引く場合とやめる場合を比べるのは、近視眼的に見える。$N \geq 6$ で引き、後でさらに引くつもりでいることが得になることはないだろうか。

ない。得点がいったん 6 以上になれば、それ以降どれだけ引いても得点は 6 以上のままである（あるいは 0 でゲームが終わる）。したがって、その後のどの判断の時点でも、1 回分の比較は再び「やめる」を示す。そのような得点から余分に 1 回引くたびに最終得点の期待値は下がるので、そこからの最善の方針はただちにやめることである。6 未満では、もう 1 回引くことがすでにやめることより有利であり、さらに続けるという選択肢はそれをいっそう有利にすることしかない。よって、この閾値による規則は単に妥当な戦略であるだけでなく、最善の戦略である。

この戦略を用いるとき、得点 N から始めたときの最終得点の期待値を $V(N)$ とする。すると $N \geq 6$ では $V(N) = N$、$N \leq 5$ では $V(N) = \frac{1}{2}V(N+1) + \frac{3}{10}V(N+2)$ である。大きい N から順に計算すると、

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

となる。よって Muggins は 1 ゲームあたり平均約 2.29 点を得ることになり、Juggins の 2.112 点を上回る。

検算

- 3 回までの取り出しのすべての列をコンピュータで厳密に数え上げた。その結果、$\mathrm{P}(0) = \frac{61}{125} = 0.488$ および $\mathrm{E}(S) = \frac{264}{125}$ が確認された。2×10^5 ゲームのモンテカルロ法（Monte Carlo method）によるシミュレーションでは、0.488 と 2.111 が得られた。
- $N = 5$ で引くと $\frac{4}{5} \cdot 5 + 1.1 = 5.1 > 5$ となる。$N = 6$ で引くと $5.9 < 6$ となる。よって切り替わりは 5 と 6 の間にある。
- あらゆる戦略にわたる価値反復（value iteration）でも同じ規則が得られた。代わりに ≥ 5 でやめる場合や ≥ 7 でやめる場合は、平均がそれぞれ 2.261, 2.265 となり、6 でやめる場合の 2.292 より低い。

よくある間違い

- $\mathrm{E}(S)$ を $3 \times \frac{11}{10}$ と計算してしまう。これは、後で黒玉が出るとそれまでの得点が消えてしまうことを忘れている。
- 条件付き平均 $\frac{11}{8}$ に、さらに条件なしの確率まで掛けてしまう。これは生き残りの因子を二重に数えている。
- 閾値を $N \leq 5.5$ と書いたまま、N が整数であることに触れない。N は整数なので、規則は「5 では引き、6 ではやめる」となる。

STEP のその先へ

Muggins の問題は**最適停止**（optimal stopping）問題である。「今やめるほうがもう 1 段階進むより有利なときにやめる」という規則は**一段先読み**（one-step look-ahead）規則とよばれ、この問題のように、その規則が「やめる」と示す状態の集合から決して抜け出せないときには、つねに最適となる。同じ考え方はマルコフ決定過程（Markov decision process）や動的計画法（dynamic programming）へとつながる。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）は期待値とマルコフ連鎖（Markov chain）の基礎を扱っており、Blitzstein と Hwang の *Introduction to Probability*（確率論入門）には、同様の第一段階解析（first-step analysis）の例が数多く載っている。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。