解答例

STEP II 1999、第 14 問

確率・統計 ・ 最適停止

> **STEP II 1999、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/99-s2-q14

鍵となる考え方

サイコロには記憶がない。もう一度投げると決めたときはいつでも、同じゲームのまさに最初の状態に戻るのであるから、その時点から先に望みうる最善の結果は同じ数 V、すなわちゲーム全体の最大の期待得点（expected score）である。これにより、各回の投げの後の判断は単純な比較になる。今の目が V を上回るならそれを確定し、そうでなければもう一度投げればよい。これを式に書くと V についての方程式が一つ得られ、それを解けば値と戦略（strategy）の両方がわかる。

解答

設定

1 の目が出た投げを**強制終了**と呼ぶことにする。このときゲームはそこで終わり、得点は 1 である。それ以外の目 $x \in \{2,3,4,5,6\}$ が出た後は、やめる（得点 x）か続けるかを選ぶ。得点はすべて 1 以上 6 以下であるから、どの戦略の期待得点も $[1,6]$ に属する。

$$V = \text{どの戦略でも達成しうる期待得点の最大値}$$

とおく（厳密には、すべての戦略にわたる最小上界（least upper bound）である。これが実際に達成されることは後でわかる）。

得点 4 を与える戦略

「**4、5、6 のいずれかが出たらすぐにやめる**」（もちろん 1 が出たときも終わる）という戦略を考える。その期待得点を E とする。最初の投げで条件付け（conditioning）を

行う。それぞれ確率 $\frac{1}{6}$ で 1,4,5,6 が出て、ゲームはその目を得点として終わる。確率 $\frac{2}{6}$ で 2 または 3 が出る。このときはもう一度投げ、サイコロには記憶がないので、そこからの期待得点は再び E である。したがって

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

(やめることなく n 回投げる確率は $(\frac{1}{3})^n \to 0$ であるから、ゲームは確率 1 で終わる。)
よって $V \geq 4$ である。

4 を上回る戦略は存在しない

任意の戦略をとる。それはこれまでの投げの履歴全体に依存していてもよい。その最初の投げの目 x に注目する。

- $x = 1$ ならば得点は 1 である。
- $x \geq 2$ で戦略がやめるならば、得点は x である。
- $x \geq 2$ で戦略が続けるならば、それ以降の振る舞いは、新しく始めた同じゲームに対する何らかの戦略にすぎない(それまでの投げはそれ以降の投げに影響しない)。したがって、そこからの期待得点は高々 V である。

いずれの場合も、$x \geq 2$ のもとでの期待得点は高々 $\max(x, V)$ である。よって、どの戦略の期待得点も高々

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

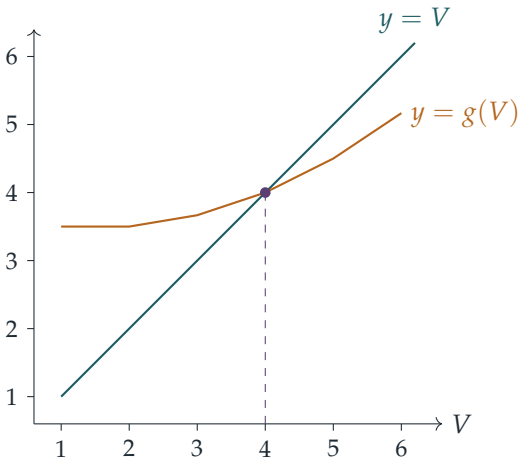
であり、戦略全体にわたる最小上界をとれば $V \leq g(V)$ を得る。

考えてみよう 不等式 $V \leq g(V)$ からどのようにして V が決まるか。

$h(V) = g(V) - V$ を考える。各項 $\max(x, V)$ は V が増加するとき変化率 0 または 1 で増加し、そのような項は五つしかないので、g の増加率は高々 $\frac{5}{6}$ である。よって h は狭義単調減少である。$V = 4$ では

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

であるから $h(4) = 0$ となり、$V > 4$ を満たすすべての V に対して $h(V) < 0$ である。$h(V) \geq 0$ であるから、$V \leq 4$ でなければならない。



結論

二つの部分を合わせると $V=4$ である。すなわち、最大の期待得点は 4 であり、それは 4、5、6 が出たらやめ、2 または 3 が出たらもう一度投げることで達成される。これはまさに「出た目が V 以上ならやめる」という規則である。2 や 3 は続けた場合に期待される 4 より悪く、5 や 6 はそれより良い。

4 は引き分けである。4 でやめれば 4 が得られ、続けても期待得点は 4 である。したがって「5 または 6 のときだけやめる」も同じく最善である。実際、同様の条件付けにより

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

これ以外の規則はすべて真に劣る。たとえば 3 以上でやめると $\frac{1+3+4+5+6}{5} = \frac{19}{5}$ であり、6 のときだけやめると $\frac{1+6}{2} = \frac{7}{2}$ である。

検算

「1 および集合 S に属する目でやめる」という任意の規則に対する便利な近道がある。各回の投げはどれも同じ確率でゲームを終わらせ、終わったときの目は $\{1\} \cup S$ のどの目も同様に確からしいので、期待得点はそれらの目の平均である。$S=\{4,5,6\}$ ならば $\frac{16}{4}=4$、$S=\{5,6\}$ ならば $\frac{12}{3}=4$ である。このような規則 32 通りすべてを正確に調べたところ、4 を与えるのはこの二つだけであり、次に良いのは $\frac{19}{5}$ であった。4 以上でやめるゲームを 400 000 回シミュレーションしたところ、平均は 3.998 であった。

よくある間違い

- 1 回の投げの平均である 3.5 を、続けることの価値として用いてしまう。続けた場合にはさらにもう一度投げるという選択肢が残るので、その価値は 3.5 より大きい。したがって、「4 以上でやめる」ことは 3.5 との比較では正当化されない。
- 1 が出たら強制終了であることを忘れ、あたかも危険なしに何度でも投げられるかのように続けた場合の価値を書いてしまう。
- しきい値型の戦略だけを調べ、履歴に依存するもっと巧妙な規則がそれより良くならない理由を説明しない。上の無記憶性（memorylessness）による議論は、すべての戦略を扱っている。

STEP のその先へ

これは**最適停止**（optimal stopping）の小さな例であり、方程式 $V = g(V)$ は**ベルマン方程式**（Bellman equation）、すなわち動的計画法（dynamic programming）やマルコフ決定過程（Markov decision processes）の基本的な道具である。同じ推論によって、金融におけるアメリカン・オプション（American options）の価格付けが行われ、有名な秘書問題（secretary problem）も解かれる。Grimmett と Welsh の *Probability: An Introduction*（確率論入門）はここで用いた条件付けの議論を扱っており、Blitzstein と Hwang の *Introduction to Probability*（確率入門）には同じ趣向の問題が数多く収められている。

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。