解答例

STEP II 2004、第 13 問

確率・統計 ・ 最適停止

> **STEP II 2004、問 13**
>
> 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/04-s2-q13

鍵となる考え方

1 回ずつの取り出しに注目して考える。白球を n 個無事に集めた時点で、もう 1 回取り出すと £1 が加わるか、手元の £n がすべて失われるかのどちらかである。したがって、その 1 回がもたらす効果の期待値（expected value）は簡単に書き下せる。目標を N 個としたときの報酬の総額の期待値が増えるのは、追加の 1 回による利得の期待値が正であるときであり、かつそのときに限る。よって最良の目標は、その利得がまだ非負である最後の個数である。

解答

以下では $1 \le r < b$ を仮定する。すなわち赤球も白球も少なくとも 1 個ずつある。($r = 0$ ならば、すべての球を取り出せばよい。)

次の 1 回で得られる利得の期待値

最初の n 回の取り出しがすべて白であったとする。袋には $b-n$ 個の球、すなわち赤球 r 個と白球 $b-n-r$ 個が残っており、次にどの球が出るのも同様に確からしい（白球が残っているためには $n \le b-r-1$ が必要である。いずれにせよ $b-n \ge r \ge 1$ である)。

- 確率 $\dfrac{b-n-r}{b-n}$ で球は白であり、報酬は £n から £$(n+1)$ になる。利得は $+1$ である。
- 確率 $\dfrac{r}{b-n}$ で球は赤であり、本来得られたはずの £n を失う。利得は $-n$ である。

したがって、次の 1 回による利得の期待値は

$$G_n = \frac{(b-n-r)\cdot 1 + r\cdot(-n)}{b-n} = \frac{b-r-n(r+1)}{b-n}.$$

である。分母は正であるから、$G_n = 0$ となるのは $b-r = n(r+1)$ のとき、すなわち

$$n = \frac{b-r}{r+1},$$

のときであり、かつそのときに限る。これが示すべきことであった。さらに役立つこととして、$n < \dfrac{b-r}{r+1}$ のとき $G_n > 0$、$n > \dfrac{b-r}{r+1}$ のとき $G_n < 0$ である。

最良の球の個数

最初の N 回の取り出しがすべて白である確率を P_N とし、N 個の球を取り出すことを目標としたときの報酬の総額の期待値を E_N とする。その N 個がすべて白ならば £N を得て、そうでなければ何も得られないので、

$$E_N = NP_N.$$

である。

> **考えてみよう** 1 回の取り出しによる利得の期待値は、報酬の総額の期待値 E_N とどう関係しているか。

目標を N とする場合と $N+1$ とする場合を比べる。二つの方針に違いが生じるのは最初の N 個がすべて白である場合だけであり（そうでなければ、どちらの方針でもすでに負けている）、これは確率 P_N で起こる。そしてその場合、追加の 1 回は期待値 G_N の分だけ報酬を変化させる。よって

$$E_{N+1} - E_N = P_N\, G_N = P_N\, \frac{b-r-N(r+1)}{b-N}.$$

である。（$P_{N+1} = P_N \cdot \frac{b-r-N}{b-N}$ から直接確かめることもできる。）

$N \le b-r$ のとき $P_N > 0$ であるから、$E_{N+1} - E_N$ の符号は $b-r-N(r+1)$ の符号と一致する。$k = N+1$ とおくと、これは

$$(b+1) - k(r+1).$$

の符号である。$m = \dfrac{b+1}{r+1}$ とおく。すると、$k < m$ のとき $E_k > E_{k-1}$、$k = m$ のとき $E_k = E_{k-1}$、$k > m$ のとき $E_k < E_{k-1}$ である。したがって数列 $E_0, E_1, E_2, \ldots$ は、少なくとも $k-1 \le b-r$ の範囲では、$E_{\lfloor m \rfloor}$ まで狭義に増加し（m が整数ならば $E_{m-1} = E_m$ と等しくなる）、その後は狭義に減少する。それより先では $P_N = 0$ かつ $E_N = 0$ であり、これより良くなることはない。

$\lfloor m \rfloor$ が意味のある目標であること、すなわち $\lfloor m \rfloor \le b - r$ を確かめておく。$b = r + 1$ ならば $m = \frac{r+2}{r+1}$ であるから $\lfloor m \rfloor = 1 = b - r$ である。$b \ge r + 2$ ならば

$$(b - r)(r + 1) - (b + 1) = r(b - r - 1) - 1 \ge r - 1 \ge 0,$$

であるから $m \le b - r$ となる。いずれの場合も、E_N の最大値は

$$n = \left\lfloor \frac{b+1}{r+1} \right\rfloor,$$

すなわち $\frac{b+1}{r+1}$ の整数部分（integer part）において達成される。これが示すべきことであった。

この目標は、どのような停止規則（stopping rule）と比べても最良である。続けるかどうかを決める前に彼女が知りうるのは、これまでの球がすべて白であったということだけである（そうでなければゲームは終わっている）。したがって、どの規則も前もって個数 N を一つ選ぶことに帰着する。

$r = 1$ の場合

赤球が 1 個のとき、最初の N 回で赤球を避けられるのは、赤球が最後の $b - N$ 個の位置のいずれかにあるときであり、かつそのときに限る。どの位置も同様に確からしいので、

$$P_N = \frac{b - N}{b}, \qquad E_N = \frac{N(b - N)}{b}.$$

である。最良の目標は $n = \left\lfloor \frac{b+1}{2} \right\rfloor$ である。

b が偶数のとき。 このとき $\frac{b+1}{2} = \frac{b}{2} + \frac{1}{2}$ であるから $n = \frac{b}{2}$ であり、

$$E_n = \frac{\frac{b}{2} \cdot \frac{b}{2}}{b} = \frac{b}{4}.$$

となる。報酬の総額の期待値は £$\frac{1}{4}b$ であり、これが示すべきことであった。

b が奇数のとき。 このとき $n = \frac{b+1}{2}$ であるから $b - n = \frac{b-1}{2}$ であり、

$$E_n = \frac{\frac{b+1}{2} \cdot \frac{b-1}{2}}{b} = \frac{b^2 - 1}{4b}.$$

となる。報酬の総額の期待値は £$\dfrac{b^2 - 1}{4b}$ である。（この場合 $m = \frac{b+1}{2}$ は整数なので、対称な式 $N(b - N)/b$ からわかるように、$\frac{b-1}{2}$ 個を目標としても同じ値が得られる。）

検算

- $r = 1$ のとき、$E_N = N(b - N)/b$ は N について上に凸の放物線であり、頂点は $N = b/2$ にある。これは二つの答えと整合する。b が偶数ならば頂点でちょうど $b/4$ となり、b が奇数ならばそれよりわずかに小さい $\frac{b}{4} - \frac{1}{4b}$ となる。

- $b = 7$、$r = 2$ とすると、公式から $n = \lfloor 8/3 \rfloor = 2$ が得られ、直接計算すると $E_0, \ldots, E_5 = 0,\ \frac{5}{7},\ \frac{20}{21},\ \frac{6}{7},\ \frac{4}{7},\ \frac{5}{21}$ となって $N = 2$ で最大になる。これは球の並べ方 7! 通りをすべて数え上げて確認した。また、$b \le 24$ のすべての場合について、n の公式を最適停止の完全な計算と照合した。
- $b = 9$、$r = 1$ のとき答えは $\frac{80}{36} = \frac{20}{9} \approx 2.22$ であり、$\frac{9}{4}$ をわずかに下回る。

よくある間違い

- 赤球が出るとすでに集めた £n を失うことを忘れ、利得の期待値を単に $\frac{b-n-r}{b-n}$ と書いてしまう。
- 「すでに n 個取り出した」ことと「目標が n 個」であることを混同する。利得 G_n は $n = \frac{b-r}{r+1}$ で 0 になる。だからこそ最良の**目標**はそれより 1 大きい $\frac{b-r}{r+1} + 1 = \frac{b+1}{r+1}$ となるのである。
- $\frac{b+1}{r+1}$ が整数のときに値が等しくなることを見落とす。このときは n と $n-1$ の両方が最適である。

STEP のその先へ

これは小さな**最適停止**（optimal stopping）問題であり、ここでの議論は「1 段先読み」の規則、すなわち「もう 1 段進んだときの利得の期待値が負になったら直ちに止める」という規則である。この規則は、問題が**単調**（monotone）であるとき、つまり一度続けることが不利に見えたらその後もずっと不利であるときには常に最適である。ここでは G_n が n について減少しているので、まさにその状況になっている。同じ考え方は、大学 1・2 年次の確率論で学ぶ秘書問題（secretary problem）やマルチンゲール（martingale）の手法へとつながる。その背後にある条件付き期待値（conditional expectation）の道具立てを身につけるには、Grimmett と Welsh の *Probability: An Introduction*（確率論入門）がよい出発点である。

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。