解答例

# STEP II 2017、第 13 問

確率・統計 ・ 期待値と無限級数

> **STEP II 2017、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/17-s2-q13

## 鍵となる考え方

どの設定でも、扉が $k$ 回目の試行で開くのは、最初の $k-1$ 回の試行がすべて失敗し、$k$ 回目が成功するときであり、またそのときに限る。したがって、その確率（probability）は「失敗」の因子の積に「成功」の因子を一つ掛けたものになる。設定ごとに変わるのは、各回に鍵を取り出すとき袋の中に鍵が何本あるかである。確率を書き下してしまえば、期待値（expectation）は和として求まる。(i) では等比級数（geometric series）を微分したもの、(ii) では等差級数（arithmetic series）、そして (iii) では打ち消し合う和（telescoping sum）であり、最後のものは調和級数（harmonic series）の一部を残す。

## 解答

以下を通して、扉が開く試行の回数を $X$ とし、$q = 1 - \frac{1}{n}$ とおく。

### 小問 (i)

鍵は毎回元に戻されるので、各回の取り出しは $n$ 本の鍵がすべて入った袋から、それ以前の取り出しとは独立に行われる。よって各試行は確率 $\frac{1}{n}$ で成功し、

$$\mathrm{P}(X = k) = q^{k-1} \cdot \frac{1}{n}, \qquad k = 1, 2, 3, \ldots$$

である。$0 < q < 1$ であるから二項展開（binomial expansion）が使えて、

$$(1-q)^{-2} = 1 + 2q + 3q^2 + \cdots = \sum_{k=1}^{\infty} kq^{k-1}.$$

したがって

$$\mathrm{E}(X) = \sum_{k=1}^{\infty} k\, q^{k-1} \frac{1}{n} = \frac{1}{n}(1-q)^{-2} = \frac{1}{n} \cdot n^2 = n.$$

試行回数の期待値は $n$ である。

**小問 (ii)**

今度は失敗するたびに袋から間違った鍵が 1 本取り除かれるので、$j$ 回目の試行のときには鍵は $n-j+1$ 本ある。扉が $k$ 回目の試行で開くには、$k-1$ 回失敗した後に成功する必要がある：

$$\mathrm{P}(X=k) = \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdots \frac{n-k+1}{n-k+2} \cdot \frac{1}{n-k+1} = \frac{1}{n},$$

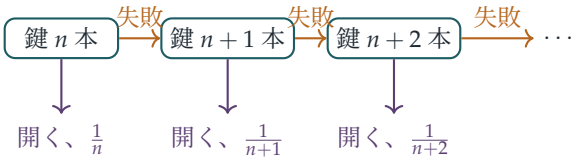
積が打ち消し合うからである。（これは直接考えても明らかである。鍵を元に戻さずに 1 本ずつ取り出すと、正しい鍵は $n$ 個の位置のうち一様にランダムな位置に来る。）よって $X$ は $\{1,2,\ldots,n\}$ 上の一様分布（uniform distribution）に従い、

$$\mathrm{E}(X) = \frac{1}{n} \sum_{k=1}^{n} k = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

**小問 (iii)**

**確率。** 袋には最初 $n$ 本の鍵があり、失敗するたびに間違った鍵が 1 本加えられ、正しい鍵は常にちょうど 1 本含まれている。よって $j$ 回目の試行のときには鍵は $n+j-1$ 本あり、条件付き確率（conditional probability）は

$$\mathrm{P}(\text{試行}j\text{回目が失敗} \mid \text{それ以前の試行がすべて失敗}) = \frac{n+j-2}{n+j-1}.$$




$k-1$ 回の失敗とその後の成功について条件付き確率を掛け合わせると、

$$\begin{aligned}\mathrm{P}(X=k) &= \frac{n-1}{n} \cdot \frac{n}{n+1} \cdots \frac{n+k-3}{n+k-2} \cdot \frac{1}{n+k-1} \\ &= \frac{n-1}{n+k-2} \cdot \frac{1}{n+k-1} = \frac{n-1}{(n+k-1)(n+k-2)},\end{aligned}$$

となる。各分子が一つ前の分母と約分され、最初の分子 $n-1$ と最後の失敗の因子の分母 $n+k-2$ だけが残るからである。$k=1$ のときは失敗の因子はなく、公式は $\frac{n-1}{n(n-1)} = \frac{1}{n}$ を与えるが、これは正しい値である。これが求める結果である。

**扉は必ず開く。** 平均を計算する前に、$X$ が確率 1 で有限であることを確かめておく価値がある。$\dfrac{n-1}{(n+k-1)(n+k-2)} = (n-1)\left(\dfrac{1}{n+k-2} - \dfrac{1}{n+k-1}\right)$ であるから、和は打ち消し合う：

$$\sum_{k=1}^{N} \mathrm{P}(X=k) = (n-1)\left(\frac{1}{n-1} - \frac{1}{n+N-1}\right) \to 1 \quad \text{極限} N \to \infty.$$

**期待値。** $\sum k\,\mathrm{P}(X=k)$ が必要なので、$\dfrac{k}{(n+k-1)(n+k-2)}$ を部分分数（partial fractions）に分解する。

$$k = A(n+k-2) + B(n+k-1)$$

とおき、$k = 1-n$ を代入すると $1-n = -A$ より $A = n-1$ を得る。$k = 2-n$ を代入すると $B = 2-n$ を得る。したがって

$$k\,\mathrm{P}(X=k) = (n-1)\left(\frac{n-1}{n+k-1} - \frac{n-2}{n+k-2}\right).$$

**考えてみよう** 係数 $n-1$ と $n-2$ が異なるので、これはきれいには打ち消し合わない。では、何が残るのだろうか？

最初の $N$ 項の和を $S_N$ とする。第 1 の分数では分母は $m = n, \ldots, n+N-1$ を動き、第 2 の分数では $m = n-1, \ldots, n+N-2$ を動く。共通の範囲 $m = n, \ldots, n+N-2$ では各 $\frac{1}{m}$ が係数 $(n-1)-(n-2) = 1$ で現れ、両端が残る：

$$S_N = (n-1)\left(\sum_{m=n}^{n+N-2} \frac{1}{m} + \frac{n-1}{n+N-1} - \frac{n-2}{n-1}\right).$$

最後の 2 項は有界である（中央の項は正で、最後の項は定数である）。一方、

$$\sum_{m=n}^{n+N-2} \frac{1}{m} = \sum_{m=1}^{n+N-2} \frac{1}{m} - \sum_{m=1}^{n-1} \frac{1}{m} \to \infty$$

が $N \to \infty$ のとき成り立つ。調和級数は発散（diverge）し、引いているのは一定の有限な量にすぎないからである。$n-1 > 0$ であるから $S_N \to \infty$ となる。項 $k\,\mathrm{P}(X=k)$ はすべて正なので、$\mathrm{E}(X)$ を定める級数は無限大に発散する。すなわち、扉はいずれ必ず開くにもかかわらず、試行回数の期待値は無限大である。

## 検算

- $n = 2$ のとき、三つの答えは 2、$\frac{3}{2}$、$\infty$ である。設定 2 は同じ鍵を二度試すことがないので、当然ながら最も効率がよい。設定 1 はその中間である。

- (iii) で $n = 3$ のとき、公式は $k = 1, 2, 3$ に対して $\frac{1}{3}, \frac{1}{6}, \frac{1}{10}$ を与える。直接計算すると $\frac{1}{3}$、$\frac{2}{3} \cdot \frac{1}{4} = \frac{1}{6}$、$\frac{2}{3} \cdot \frac{3}{4} \cdot \frac{1}{5} = \frac{1}{10}$ である。$2 \times 10^5$ 回のゲームのコンピュータシミュレーションでは、相対度数は 0.334、0.168、0.099 であった。
- 同じく $n = 3$ のとき、$N = 10^4$ での部分和 $S_N$ は約 15.6 であり、$2 \ln N$ のように増え続ける。極限は存在しない。

## よくある間違い

- (iii) で、$k$ 回目の試行のときの袋の中の鍵の本数を $n + k - 1$ ではなく $n + k$ としてしまう。$k$ 回目の取り出しの前に加えられた鍵は $k - 1$ 本だけである。
- 平均が無限大なので扉は決して開かないかもしれない、と結論してしまう。確率の和は 1 である。$1/k^2$ のように減衰する裾が重すぎて、$\sum k \, \mathrm{P}(X = k)$ が収束しないというだけのことである。
- 二つの係数が等しいかのように部分分数を打ち消し合わせ、有限の答えを得てしまう。

## STEP のその先へ

設定 3 では、$\mathrm{P}(X = k)$ が $1/k^2$ のように振る舞うため、確率 1 で有限であるにもかかわらず平均が無限大となる確率変数（random variable）が得られる。このような裾の重い分布（heavy-tailed distribution）や、有名なサンクトペテルブルクのパラドックス（St Petersburg paradox）は、大学 1 年の確率論の講義で $\mathrm{E}(X)$ は $\sum k \, \mathrm{P}(X = k)$ が収束するときに限り存在すると強調される理由を示している。失敗するたびに間違った鍵が加えられる袋は、「富める者はますます富む」過程の標準的なモデルであるポリア（Pólya）の壺の仲間である。Grimmett and Welsh, *Probability: An Introduction*（確率論入門）は、離散確率変数の期待値とこうした微妙な点を扱っている。

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。