解答例

STEP I 2001、第 14 問

確率・統計 ・ ランダムな順列

> **STEP I 2001、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/01-s1-q14

この問題は、*Beyond STEP* の第 12.12 節「**いつ探すのをやめるか**」で詳しく解説されている：https://beyondstep.keithmcnulty.org/ja/ch12/secretary.html

鍵となる考え方

面接の順序全体を考えるのではなく、**特定の候補者がどこに位置するか**だけを考える。最初の n 人の中で最良の者が全体の最良であるのは、最上位の候補者が最初に面接される n 人の中にいるときであり、そのときに限る。そして対称性（symmetry）により、最上位の候補者は N 個の位置のどこにも等しい確率でいる。小問 (ii) では場合をもう一つ加えるか、うまく余事象（complement）を使えばよい。

解答

候補者に順位でラベルを付け、候補者 1 が全体で最良、候補者 2 が二番目に良い者、以下同様とする。$N!$ 通りの順序はすべて同様に確からしいので、どの候補者もどの位置に来る確率も等しく、最初の n 回の面接で会う候補者の集合は、大きさ n の $\binom{N}{n}$ 個の部分集合のどれである確率も等しい。以下を通して $1 \le n \le N$ とする。

小問 (i)

候補者 1 が最初に面接される n 人の中にいるという事象を A とする。A が起これば、候補者 1 は確かに最初の n 人の中で最良である。起こらなければ、最初の n 人の中で最良の者は候補者 1 より劣る誰かである。したがって、求める事象はちょうど A である。

候補者 1 は N 個の位置のそれぞれに等しい確率でおり、そのうち n 個が最初の n 個に含まれるので、

$$\mathrm{P}(\text{最初の}n\text{ 人の中の最良が全体の最良}) = \frac{n}{N}.$$

小問 (ii)

ここでは二番目に良い候補者が存在するように $N \geq 2$ とする。最初の n 人の中で最良の者が全体で最良または二番目であるための必要十分条件は、次のいずれかが成り立つことである。

- 候補者 1 が最初の n 人の中にいる（事象 A）。
- 候補者 1 は最初の n 人の中に**いない**が、候補者 2 はいる。

考えてみよう 二番目の場合、最初の n 人の中で最良の者が候補者 2 であり、さらに良い誰かではないとなぜわかるのか？

候補者 2 より良い候補者は候補者 1 だけであり、この場合その候補者は早い段階で面接されていない。逆に、1 も 2 も最初の n 人の中にいなければ、最初の n 人の中で最良の者の順位は 3 以下である。したがって、上の二つの場合を合わせたものがちょうど求める事象であり、しかもこれらは互いに排反（mutually exclusive）である。

一つ目の場合の確率は小問 (i) より $\frac{n}{N}$ である。二つ目については、$\mathrm{P}(A') = \frac{N-n}{N}$ であり、候補者 1 が最後の $N-n$ 個の位置のどれかにいるという条件の下で、候補者 2 は残りの $N-1$ 個の位置のどこにも等しい確率でいて、そのうち n 個が最初の n 個に含まれる。よって

$$\begin{aligned}\mathrm{P}(\text{最初の}n\text{ 人の中の最良が全体で一番目または二番目}) &= \frac{n}{N} + \frac{N-n}{N}\cdot\frac{n}{N-1} \\ &= \frac{n\big[(N-1)+(N-n)\big]}{N(N-1)} \\ &= \frac{n(2N-n-1)}{N(N-1)}.\end{aligned}$$

別解として、余事象を使うこともできる。この事象が起こらないのは、早い段階の n 人がすべて順位 3 から N までの $N-2$ 人の候補者から選ばれるときであり、その確率は

$$\frac{\binom{N-2}{n}}{\binom{N}{n}} = \frac{(N-n)(N-n-1)}{N(N-1)},$$

である。そして $N(N-1)-(N-n)(N-n-1) = 2Nn-n^2-n = n(2N-n-1)$ となり、同じ答えが得られる。

$N = 4$、$n = 2$ の場合

公式から、(i) は $\frac{2}{4} = \frac{1}{2}$、(ii) は $\frac{2 \cdot 5}{4 \cdot 3} = \frac{5}{6}$ となる。

起こりうる場合を書き出すにあたり、最初に面接される二人が**誰であるか**だけが問題であることに注意する。$\binom{4}{2} = 6$ 組のペアはそれぞれ、24 通りの順序のうちちょうど $2! \times 2! = 4$ 通りから生じる（最初の二人を並べ、次に残りの二人を並べる）ので、六つのペアは同様に確からしい。

最初に面接される二人	$\{1,2\}$	$\{1,3\}$	$\{1,4\}$	$\{2,3\}$	$\{2,4\}$	$\{3,4\}$
その中の最良	1	1	1	2	2	3

最初の二人の中で最良の者が全体の最良であるのは 6 通り中 3 通りで確率は $\frac{1}{2}$、全体で最良または二番目であるのは 6 通り中 5 通りで確率は $\frac{5}{6}$ である。どちらも公式と一致する。

検算

- $n = N$ のとき：全員に会うので、どちらの確率も 1 になるはずである。実際、$\frac{N}{N} = 1$ かつ $\frac{N(N-1)}{N(N-1)} = 1$ である。
- $n = 1$ のとき：最初の候補者は一様（uniformly）にランダムな一人なので、(ii) は $\frac{2}{N}$ になるはずである。公式は $\frac{2N-2}{N(N-1)} = \frac{2}{N}$ を与える。
- $2N - n - 1 \geq N - 1$ であるから、(ii) の答えは (i) の答え以上である。
- $N \leq 7$ とすべての n について $N!$ 通りの順序をすべて数え上げ（enumeration）た結果は、二つの公式と完全に一致する（例えば $N = 5$、$n = 2$ では $\frac{2}{5}$ と $\frac{7}{10}$ となる）。

よくある間違い

- (ii) で、「全体で二番目」が数に入るのは全体の最良が最初の n 人の中に**いない**ときだけであることを忘れる。単に $\frac{n}{N} + \frac{n}{N}$ と足すと重複して数えてしまう。
- 候補者 1 の位置で条件付け（conditioning）をした後に、候補者 2 について $\frac{n}{N}$ を使ってしまう。候補者 1 の位置が決まれば、残る位置は $N - 1$ 個しかない。

STEP のその先へ

これは有名な**秘書問題**（secretary problem）の設定である。各候補者をその場で採用するか不採用にするかを決めなければならないとき、最善の戦略は最初の n 人を見送

り、その後それら全員より良い最初の候補者を採用することである。小問 (i) の議論を拡張すると、最適な n はおよそ N/e であり、成功確率は約 $1/e$ であることが示される。ここで用いた対称性の議論（「どの候補者もどこにでも等しい確率でいる」）は、ランダム置換（random permutations）の研究の出発点である。Blitzstein と Hwang の *Introduction to Probability*（確率論入門）には、このような対称性の議論についての優れた解説がある。

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。