解答例

STEP II 2024、第 12 問

確率・統計 ・ 離散順序統計量

STEP II 2024、問 12

先に進む前に、まず問題を読んでおこう。問題は、OCR が配布している 2024 年の STEP 過去問一式（zip ファイル）のうち、STEP II の問題用紙にある：

www.ocr.org.uk/Images/725753-2024-past-papers.zip

鍵となる考え方

第 1 のゲームにおける勝者とは、**ただ 1 人だけが持つ**最小の数を持つプレイヤーのことであり、そのようなプレイヤーは高々 1 人しか存在しない。したがって「勝者がいる」という事象は、誰がどの数で勝つかによって互いに排反（disjoint）な場合に分けられ、各場合は独立（independent）な確率（probability）の単純な積になる。$\sum i^2$ と $\sum i^3$ の公式が与えられているのは、現れる和が多項式の和になるというヒントである。第 2 のゲームも同様に扱えるが、今度はプレイヤーの順序対（一方が最小、他方が最大）と、和をとるべき 2 つの数を考えることになる。

解答

以下を通して、k 個の数は独立で、それぞれ $\{1,2,\ldots,n\}$ 上の一様分布（uniform distribution）に従う。したがって、特定のプレイヤーが特定の数を受け取る確率は $\frac{1}{n}$ であり、s 個の値からなる与えられた集合に属する数を受け取る確率は $\frac{s}{n}$ である。

小問 (i)

$1 \le a \le n-1$ を満たす a を固定する。Ada は a を受け取らなければならず（確率 $\frac{1}{n}$）、他の $k-1$ 人のプレイヤーはそれぞれ $n-a$ 個の数 $a+1,\ldots,n$ のいずれかを受け取らなければならない（それぞれ確率 $\frac{n-a}{n}$）。独立性より、

$$\mathrm{P}(\text{Ada の数が}a\text{, 他の全員はそれより大きい}) = \frac{1}{n}\left(\frac{n-a}{n}\right)^{k-1} = \frac{(n-a)^{k-1}}{n^k}.$$

（$a = n$ のときはこれが 0 になるので、a は $n-1$ までとしている。）

勝者がいるという事象を W とする。勝者は他の全員の数より真に小さい数を持つので、勝者が 2 人いることはありえない。もし 2 人いれば、互いに相手より小さくなければならないからである。よって W は、k 人のプレイヤーと $a = 1, \ldots, n-1$ のすべてにわたる事象「このプレイヤーが a を受け取り、他の全員がそれより大きい数を受け取る」の互いに排反な和集合である。対称性（symmetry）より各プレイヤーの寄与は Ada の寄与と等しいので、

$$\mathrm{P}(W) = k \sum_{a=1}^{n-1} \frac{(n-a)^{k-1}}{n^k} = \frac{k}{n^k} \sum_{j=1}^{n-1} j^{k-1},$$

となる。ここで $j = n - a$ と置き換えた。j も 1 から $n-1$ までを動く。

$k = 4$ のとき、$\sum i^3$ の公式を n の代わりに $n-1$ として用いると、

$$\mathrm{P}(W) = \frac{4}{n^4} \cdot \frac{1}{4}(n-1)^2 n^2 = \frac{(n-1)^2}{n^2},$$

となり、示すべき式が得られる。

小問 (ii)

$1 \le d \le n-2$ かつ $1 \le a \le n-d-1$ を満たす d と a を固定する。これらの条件により、$a + d + 1 \le n$ であり、かつ a と $a+d+1$ の間に（両端を除いて）少なくとも 1 つの数が存在することが保証される。Ada は a を、Bob は $a+d+1$ を受け取らなければならず（それぞれ確率 $\frac{1}{n}$）、他の $k-2$ 人のプレイヤーはそれぞれ d 個の数 $a+1, \ldots, a+d$ のいずれかを受け取らなければならない。独立性より、

$$\mathrm{P}(\text{Ada の数が}a,\ \text{Bob の数が}a+d+1,\ \text{他は真にその間}) = \frac{1}{n^2}\left(\frac{d}{n}\right)^{k-2} = \frac{d^{k-2}}{n^k}.$$

考えてみよう 勝者が 2 人になるあらゆる場合が、プレイヤーと a、d にわたる和の中にちょうど 1 回ずつ現れるのはなぜか。

勝者が 2 人いるとは、あるプレイヤーがただ 1 人だけの最小の数を持ち、それとは別のプレイヤーがただ 1 人だけの最大の数を持つということである。小問 (i) と同様に、小さい側の勝者は一意的であり、大きい側の勝者も一意的である。よって結果から、相異なるプレイヤーの**順序対**（小さい側の勝者、大きい側の勝者）と、小さい側の数 a、および大きい側の数までの間隔 $d \ge 1$ が定まる。$k-2 \ge 2$ 人の他のプレイヤーの数が真にその間にあるので、間隔は少なくとも 1 でなければならない。これらの選び方が異なれば、互いに排反な事象になる。プレイヤーの順序対は $k(k-1)$ 通りあり、それぞれ (Ada, Bob) と同じだけ寄与するので、

$$\mathrm{P}(\text{勝者が 2 人}) = k(k-1) \sum_{d=1}^{n-2} \sum_{a=1}^{n-d-1} \frac{d^{k-2}}{n^k} = \frac{k(k-1)}{n^k} \sum_{d=1}^{n-2} (n-1-d)\, d^{k-2},$$

となる。内側の和は $n-d-1$ 個の等しい項からなるからである。

ここで $k=4$ とし、$m=n-1$ とおくと、d は 1 から $m-1$ までを動く。与えられた 2 つの公式を n の代わりに $m-1$ として用いると、

$$\begin{aligned}\sum_{d=1}^{m-1}(m-d)\,d^2 &= m\cdot\tfrac{1}{6}(m-1)m(2m-1)-\tfrac{1}{4}(m-1)^2m^2\\ &= \tfrac{1}{12}m^2(m-1)[2(2m-1)-3(m-1)]\\ &= \tfrac{1}{12}m^2(m-1)(m+1).\end{aligned}$$

したがって、$m-1=n-2$ かつ $m+1=n$ であるから、

$$\mathrm{P}(\text{勝者が 2 人})=\frac{12}{n^4}\cdot\frac{1}{12}(n-1)^2(n-2)n=\frac{(n-2)(n-1)^2}{n^3},$$

となり、これが求める結果である。

$k=4$ のときの勝者 1 人と 2 人の比較

あるプレイヤーがただ 1 人だけの最小の数を持つという事象を L、あるプレイヤーがただ 1 人だけの最大の数を持つという事象を H とする。小問 (i) より $\mathrm{P}(L)=\frac{(n-1)^2}{n^2}$ である。すべての数 x を $n+1-x$ に置き換えると、一様分布に従う数は一様分布に従う数に移り、「最小」と「最大」が入れ替わるので、$\mathrm{P}(H)=\mathrm{P}(L)$ も成り立つ。

勝者がちょうど 2 人であるという事象は $L\cap H$ であるので、$T=\mathrm{P}(L\cap H)=\frac{(n-2)(n-1)^2}{n^3}$ と書く。勝者がちょうど 1 人であるとは、L または H が起こるが両方は起こらないということであるから、

$$\mathrm{P}(\text{ちょうど 1 人})=\mathrm{P}(L)+\mathrm{P}(H)-2\,\mathrm{P}(L\cap H)=\frac{2(n-1)^2}{n^2}-2T.$$

勝者が 2 人である確率が 1 人である確率より大きくなるのは、次の場合であり、かつその場合に限る。

$$\begin{aligned}T>\frac{2(n-1)^2}{n^2}-2T &\iff 3\cdot\frac{(n-2)(n-1)^2}{n^3}>\frac{2(n-1)^2}{n^2}\\ &\iff 3(n-2)>2n \iff n>6,\end{aligned}$$

ここで正の量 $\frac{(n-1)^2}{n^3}$ で割った（$n\geq 3$ なので正である）。$n=6$ では 2 つの確率は等しいので、求める最小値は

$$n=7.$$

である。

検算

- $n = 3$、$k = 4$ のとき、結果は全部で 81 通りある。1 で勝つには残り 3 つの数が $\{2,3\}$ に属する必要があり、$4 \cdot 2^3 = 32$ 通りである。2 で勝つには残り 3 つがすべて 3 である必要があり、4 通りである。よって $\mathrm{P}(W) = \frac{36}{81} = \frac{4}{9} = \frac{(3-1)^2}{3^2}$ となる。
- $n = 6$ では「ちょうど 1 人」と「ちょうど 2 人」の確率はともに $\frac{25}{54} \approx 0.463$ である。$n = 7$ ではそれぞれ $\frac{144}{343} \approx 0.420$ と $\frac{180}{343} \approx 0.525$ である。これらの値と、$n \leq 8$、$k = 2,3,4$ に対する両小問の公式は、すべての結果をコンピュータで数え上げることによって確認した。
- $n \to \infty$ のとき同じ数が重なることはまれになるので、プレイヤーが 4 人ならば、ほぼ確実にただ 1 人の最小とただ 1 人の最大がともに存在するはずである。実際、$\frac{(n-2)(n-1)^2}{n^3} \to 1$ である。

よくある間違い

- 小問 (i) の因子 k、あるいは小問 (ii) の因子 $k(k-1)$ を忘れること。小問 (ii) では、一方が小さい側、他方が大きい側であるから、勝者の組は順序対である。
- 「勝者がちょうど 1 人」を $\mathrm{P}(L) + \mathrm{P}(H) - \mathrm{P}(L \cap H)$ と計算してしまうこと。これは「少なくとも 1 人」の確率である。重なりの部分は 2 回引かなければならない。
- $n = 6$ と答えること。このとき 2 つの確率は等しく、大小が厳密についていない。

STEP のその先へ

事象を「誰が」と「どこで」によってラベル付けされた互いに排反な部分に分けるという手法は、標本の最小値や最大値といった順序統計量（order statistics）の分布を求めることの離散版である。連続的な標本に対しては、同じ議論によって最小値と最大値の同時密度（joint density）が得られ、この問題を面白くしている同点は起こらなくなる。これらの考え方に触れるには、Grimmett and Welsh, *Probability: An Introduction*（確率論の入門書）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。