解答例

STEP I 2013、第 12 問

確率・統計 ・ 非復元抽出

STEP I 2013、問 12

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/13-s1-q12

鍵となる考え方

k 錠を「無作為に」選ぶとは、容器に残っている錠剤から作られる k 錠の組がどれも同様に確からしいということである。したがって、各日の確率は個数の比、すなわち条件を満たす選び方の数を $\binom{\text{残りの錠剤数}}{k}$ で割ったものになる。最終日についての問いでは、対称性（symmetry）が決め手となる。最終日のために残される錠剤は、1 日目に選ばれる錠剤とまったく同じように無作為なのである。(iii) では、各日がうまくつながる。うまくいった日の翌日には、容器の中はちょうど日数の短い休暇の場合と同じ状態になっているからである。

解答

以下では一貫して、錠剤は見た目が同じであっても区別できる（たとえば番号が付いている）ものと考える。こう考えても確率は変わらず、しかも N 錠から k 錠を選ぶ $\binom{N}{k}$ 通りの選び方がすべて同様に確からしくなる。

小問 (i)

$k = 3, n = 3$ のとき錠剤は 9 錠あり、各種類 3 錠ずつである。そこから 3 錠を選ぶ方法は $\binom{9}{3} = 84$ 通りである。各種類を 1 錠ずつ得るには、種類 1 の 3 錠から 1 錠、種類 2 の 3 錠から 1 錠、種類 3 の 3 錠から 1 錠を選べばよいので、$3^3 = 27$ 通りある。よって

$$\mathrm{P}(1\text{日目に各種類1錠ずつ}) = \frac{27}{84} = \frac{9}{28},$$

となり、求める結果が得られる。

考えてみよう 最終日の確率を求めるのに、なぜ 1 日目と 2 日目に何が起こるかを追いかける必要がないのだろうか。

休暇の間に、9 錠の錠剤は無作為に 3 錠ずつの 3 つのグループに分けられる。1 日目のグループ、2 日目のグループ、そして 3 日目のグループ（残ったもの）である。グループを 1, 2, 3 の順に選んでも 3, 2, 1 の順に選んでも、得られる分布は同じである。錠剤をラベル付きの 3 つのグループに分ける方法はどれも同様に確からしいからである。特に、3 日目のために残される組は、1 日目に選ばれる組とまったく同様に、3 錠の組として一様（uniformly）に無作為である。したがって、最終日に各種類 1 錠ずつが残っている確率もまた

$$\frac{9}{28}.$$

である。

小問 (ii)

今度は錠剤が $3n$ 錠あり、各種類 n 錠ずつである。同様に確からしい選び方は $\binom{3n}{3}$ 通りあり、そのうち各種類を 1 錠ずつ含むものは n^3 通りである。よって

$$\begin{aligned}\mathrm{P}(\text{1 日目に各種類 1 錠ずつ}) &= \frac{n^3}{\binom{3n}{3}} = \frac{6n^3}{3n(3n-1)(3n-2)} \\ &= \frac{2n^2}{(3n-1)(3n-2)}.\end{aligned}$$

$n = 3$ のとき、これは $\frac{18}{8\cdot 7} = \frac{9}{28}$ となり、(i) と一致する。

小問 (iii)

今度は $k = 2$ で、錠剤は $2n$ 錠あり、2 つの種類それぞれ n 錠ずつである。$m = n, n-1, \ldots, 1$ に対して、その日の始めに容器に各種類 m 錠ずつ入っている日に各種類を 1 錠ずつ選ぶという事象を G_m とする（したがって 1 日目は $m = n$ である）。

容器に各種類 m 錠ずつ入っているとき、2 錠の選び方は $\binom{2m}{2} = m(2m-1)$ 通りあり、そのうち各種類 1 錠ずつとなるものは $m \cdot m$ 通りである。よって

$$\mathrm{P}(\text{各種類 1 錠ずつ} \mid m \text{ 錠ずつ残っている}) = \frac{m^2}{m(2m-1)} = \frac{m}{2m-1}.$$

それまでの日がすべてうまくいっていれば、各日にちょうど各種類 1 錠ずつが取り除かれているので、次の日は各種類 $m-1$ 錠ずつで始まる。条件付き確率（conditional probability）を 1 日ずつ掛け合わせると

$$\mathrm{P}(\text{毎日各種類 1 錠ずつ}) = \prod_{m=1}^{n} \frac{m}{2m-1} = \frac{n!}{1 \times 3 \times 5 \times \cdots \times (2n-1)}.$$

奇数の積を整理するために、分子と分母に偶数の積 $2\times4\times\cdots\times(2n)=2^n\,n!$ を掛けると

$$1\times3\times\cdots\times(2n-1)=\frac{(2n)!}{2^n\,n!}.$$

よって

$$\boxed{\mathrm{P}(\text{毎日各種類 1 錠ずつ})=\frac{2^n\,(n!)^2}{(2n)!}}$$

（検算として、$n=1$ のときは 1 となり、これは正しい。各種類 1 錠ずつしかなければ、両方を取るしかないからである。）

次に、$n!$ と $(2n)!$ にスターリングの近似（Stirling's approximation）を用いる。

$$(n!)^2\approx2n\pi\left(\frac{n}{\mathrm{e}}\right)^{2n},$$

$$(2n)!\approx\sqrt{4n\pi}\left(\frac{2n}{\mathrm{e}}\right)^{2n}=2\sqrt{n\pi}\,2^{2n}\left(\frac{n}{\mathrm{e}}\right)^{2n}.$$

因数 $(n/\mathrm{e})^{2n}$ が約分されて

$$\frac{2^n\,(n!)^2}{(2n)!}\approx\frac{2^n\cdot2n\pi}{2\sqrt{n\pi}\,2^{2n}}=2^{-n}\,\frac{n\pi}{\sqrt{n\pi}}=2^{-n}\sqrt{n\pi},$$

が残り、これが求める近似である。

検算

- 錠剤の取り出し方をすべてコンピュータで数え上げると、(i) で 1 日目と 3 日目のいずれについても $\frac{9}{28}$ となること、$n\le7$ に対して (ii) の公式、$n\le4$ に対して (iii) の公式が正しいことが確かめられる。
- $n=10$ のとき、(iii) の正確な確率は $\frac{2^{10}(10!)^2}{20!}\approx0.005542$ であり、一方 $2^{-10}\sqrt{10\pi}\approx0.005474$ である。近似値は約 1.3% 小さい。n が大きくなるとこの比は 1 に近づく（$n=50$ では約 1.0025 である）。これはスターリングの近似から予想されるとおりである。
- (iii) の別の数え方：$2n$ 錠を n 組のペアの順序付きの列に分ける方法は $(2n)!/2^n$ 通りあり、どれも同様に確からしい。条件を満たす分け方は、種類 1 の錠剤と種類 2 の錠剤を日ごとにペアにするもので、$n!\times n!$ 通りある。その比はやはり $2^n(n!)^2/(2n)!$ である。

よくある間違い

- 錠剤を 1 錠ずつ取り出すと考えて $\frac{3}{9}\times\frac{3}{8}\times\frac{3}{7}$ と書いてしまう。これは種類 1, 2, 3 がこの順に出る確率である。種類の出る順序は 3! 通りあるので、$6\times\frac{27}{504}=\frac{9}{28}$ とな

る。（あるいは、最初の錠剤は何でもよいので $1 \times \frac{6}{8} \times \frac{3}{7} = \frac{9}{28}$ となる。）

- 最終日の確率を求めるのに、1 日目に起こりうることすべてについて条件付けしようとする。それでも解けるが長くなり、対称性を使えば答えはすぐに得られる。
- $(2n)!$ にスターリングの近似を用いるとき、平方根の中も含めてすべての n を $2n$ に置き換える必要がある。$\sqrt{2(2n)\pi} = 2\sqrt{n\pi}$ であり、$\sqrt{2n\pi}$ ではない。

STEP のその先へ

(iii) の答えは $2^n / \binom{2n}{n}$ であり、評価式 $\binom{2n}{n} \approx 4^n / \sqrt{n\pi}$ は確率論のいたるところに現れる。これは単純ランダムウォーク（simple random walk）が $2n$ ステップ後に出発点に戻っている確率を与え、1 次元や 2 次元のそのようなランダムウォークが確率 1 で出発点に戻ることの証明の原動力となる。(i) の対称性の議論は**交換可能性**（exchangeability）という考え方である。すなわち、無作為な取り出しを行う順序を変えても、それらの同時分布（joint distribution）は変わらない。どちらの考え方も、Grimmett and Welsh, *Probability: An Introduction*（確率論入門）や、Blitzstein and Hwang, *Introduction to Probability*（確率論入門）で展開されている。

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。