解答例

STEP I 2013、第 13 問

確率・統計 ・ 数え上げと期待値

> **STEP I 2013、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/13-s1-q13

鍵となる考え方

数 $1, \ldots, 52$ は 26 個の互いに交わらない（disjoint）「相方の組」$\{k, 53-k\}$ に分かれる。そして捨てられる組とは、両方の要素を選んでしまった相方の組にほかならない。したがって、完全な組をちょうど k 組含む選び方を数えるには、まずどの組を完全な組にするかを選び、残りの枠を**異なる**組から来る数で埋めればよい。その際、各組でどちらの相方を取るかは自由に選べる。期待値（expectation）については、26 組にわたる線形性（linearity）を使えば手早く確かめられる。

解答

選び方を数える

整数を 26 組 $\{1,52\}, \{2,51\}, \ldots, \{26,27\}$ に分けて書く。1 から 52 までの二つの整数の和が 53 になるのは、それらがこれらの組の一つをなすときであり、そのときに限る。また、各整数はちょうど一つの組に属する。よって X は、選ばれた七つの数の中に両方の要素が含まれる組の個数である。$2X \leq 7$ であるから、X のとりうる値は $0, 1, 2, 3$ である。

$\binom{52}{7}$ 通りの選び方はすべて同様に確からしいので、$X = k$ となる選び方を数える。

- 完全な組を k 組選ぶ：$\binom{26}{k}$ 通り。これらで七つの数のうち $2k$ 個を使う。
- 残りの $7-2k$ 個の数は、残る $26-k$ 組のうちの**異なる** $7-2k$ 組から来なければならない（同じ組から二つ取ると、完全な組がもう一つできてしまう）：$\binom{26-k}{7-2k}$ 通り。
- それらの各組から、二つの要素のうち一方を取る：2^{7-2k} 通り。

$X = k$ となる各選び方は、この方法でちょうど一回ずつ現れるから、

$$\mathrm{P}(X = k) = \frac{\binom{26}{k}\binom{26-k}{7-2k}\,2^{7-2k}}{\binom{52}{7}}, \qquad k = 0, 1, 2, 3.$$

具体的な数値

$$\begin{aligned}
X = 0: &\quad \binom{26}{7}2^7 = 657\,800 \times 128 = 84\,198\,400,\\
X = 1: &\quad 26\binom{25}{5}2^5 = 26 \times 53\,130 \times 32 = 44\,204\,160,\\
X = 2: &\quad \binom{26}{2}\binom{24}{3}2^3 = 325 \times 2024 \times 8 = 5\,262\,400,\\
X = 3: &\quad \binom{26}{3}\binom{23}{1}2 = 2600 \times 23 \times 2 = 119\,600,
\end{aligned}$$

であり、$\binom{52}{7} = 133\,784\,560$ である。これら四つの個数の和は確かに $\binom{52}{7}$ に等しく、数え漏れや重複がないことのよい確認になる。

考えてみよう 九桁の数を含む分数は、どうやって約分すればよいだろうか？

問題文の注がヒントである。分母は $5593 = 7 \times 17 \times 47$ になるはずだ。実際 $\binom{52}{7} = 23\,920 \times 5593$ であり、四つの個数はどれも 23 920 の倍数である：

$$\begin{aligned}
84\,198\,400 &= 3520 \times 23\,920, & 44\,204\,160 &= 1848 \times 23\,920,\\
5\,262\,400 &= 220 \times 23\,920, & 119\,600 &= 5 \times 23\,920.
\end{aligned}$$

よって X の分布（distribution）は

k	0	1	2	3
$\mathrm{P}(X = k)$	$\frac{3520}{5593}$	$\frac{1848}{5593}$	$\frac{220}{5593}$	$\frac{5}{5593}$

である（二番目は既約分数で表すと $\frac{264}{799}$ である）。そして当然ながら $3520 + 1848 + 220 + 5 = 5593$ となっている。

期待値

$$\mathrm{E}(X) = \frac{0 \times 3520 + 1 \times 1848 + 2 \times 220 + 3 \times 5}{5593} = \frac{2303}{5593}.$$

ここで $2303 = 7^2 \times 47$、$5593 = 7 \times 17 \times 47$ であるから、

$$\mathrm{E}(X) = \frac{7^2 \times 47}{7 \times 17 \times 47} = \frac{7}{17},$$

となり、示すべきことが示された。

期待値へのより速い道

$j = 1, \ldots, 26$ に対し、第 j 組の両方の要素が選ばれていれば $I_j = 1$、そうでなければ 0 と定める。すると $X = I_1 + \cdots + I_{26}$ である。ある組の両方の要素を選ぶと、残りの 50 個からさらに 5 個を選ぶことになるので、

$$\mathrm{P}(I_j = 1) = \frac{\binom{50}{5}}{\binom{52}{7}} = \frac{7 \times 6}{52 \times 51} = \frac{7}{442}.$$

期待値の線形性（これには独立性（independence）は必要ない。実際、I_j たちは互いに従属している）により、

$$\mathrm{E}(X) = 26 \times \frac{7}{442} = \frac{7}{17}.$$

検算

- 四つの確率の和は 1 であり、二つの方法は同じ $\mathrm{E}(X) = \frac{7}{17} \approx 0.412$ を与える。
- 無作為な選び方 300 000 回のコンピュータシミュレーションでは、$X = 0, 1, 2, 3$ の相対度数（relative frequency）はおよそ 0.630、0.329、0.039、0.001 となり、厳密な値 0.6294、0.3304、0.0393、0.0009 とよく合った。
- 妥当性の確認：$\mathrm{P}(X = 3)$ はごく小さい。七つの数だけで完全な組が三組できるのだから、そうなるはずである。

よくある間違い

- $X = 0$ となる選び方の数を、$\binom{52}{7}$ から何かを大ざっぱに引いたものとしたり、$\binom{26}{7}$ だけとしたりして、各組のどちらの要素を取るかに対応する因数 2^7 を忘れる。
- $X = 1$ のとき、残りの五つの数の選び方を $\binom{50}{5}$ とする。これでは二つ目の完全な組ができる場合も含まれてしまい、数えすぎになる。

STEP のその先へ

個数を指示変数（indicator variable）の和として表し、期待値の線形性を使うのは、大学一年次の確率論で最も役に立つ技法の一つである。分布を求めるのが絶望的な場合でも、期待値は求められるのだ。今回は分布全体を求めることができたが、たとえば無作為な置換（permutation）の不動点（fixed point）の個数の期待値などでは、指示変数を使うのが自然な入り口となる。Blitzstein and Hwang, *Introduction to Probability*（確率論入門）は、この考え方を何度も繰り返し用いている。

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。