詳細解答

STEP I 2013，第 13 題

概率與統計 • 計數與期望值

> **STEP I 2013，第 13 題**
>
> 繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：
>
> step.maths.org/questions/13-s1-q13

關鍵思路

數字 $1,\ldots,52$ 可分成 26 個互不相交的「伙伴對」$\{k, 53-k\}$，而被棄掉的一對，其實就是我們把兩個成員都選中了的伙伴對。因此，要數出恰好有 k 個完整對的選法，我們先選定哪些對是完整的，再用來自不同對的數字填滿其餘位置，每次可自由選取該對中的哪一個成員。至於期望值（expectation），對這 26 個對運用線性性質（linearity）便可迅速驗證結果。

解答

數算選法

把這些整數寫成 26 個對 $\{1,52\},\{2,51\},\ldots,\{26,27\}$。由 1 至 52 中的兩個整數之和為 53，當且僅當它們組成其中一個對，而每個整數恰好屬於一個對。所以 X 就是兩個成員都在所選七個數之中的對的數目。由於 $2X \le 7$，X 的取值為 $0,1,2,3$。

所有 $\binom{52}{7}$ 種選法的可能性相等，因此我們數算 $X=k$ 的選法數目。

- 選出 k 個完整對：共 $\binom{26}{k}$ 種方法。這些對用去七個數中的 $2k$ 個。
- 其餘 $7-2k$ 個數必須來自餘下 $26-k$ 個對中 $7-2k$ 個不同的對（若有兩個來自同一對，便會多出一個完整對）：共 $\binom{26-k}{7-2k}$ 種方法。
- 從每個這樣的對中取其兩個成員之一：共 2^{7-2k} 種方法。

每個 $X=k$ 的選法以這種方式恰好出現一次，所以

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

具體數值

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

而 $\binom{52}{7} = 133\,784\,560$。這四個數目加起來確實等於 $\binom{52}{7}$，這是一個很好的驗證，說明沒有遺漏或重複計算。

想一想 分數中有九位數時，我們該如何化簡？

題目中的附註就是提示：$5593 = 7\times 17\times 47$ 應該是分母。事實上 $\binom{52}{7} = 23\,920\times 5593$，而這四個數目每一個都是 23 920 的倍數：

$$84\,198\,400 = 3520\times 23\,920, \qquad 44\,204\,160 = 1848\times 23\,920,$$
$$5\,262\,400 = 220\times 23\,920, \qquad 119\,600 = 5\times 23\,920.$$

所以 X 的分佈（distribution）為

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

（第二個的最簡形式為 $\frac{264}{799}$），而 $3520+1848+220+5 = 5593$，理應如此。

期望值

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

由於 $2303 = 7^2\times 47$ 而 $5593 = 7\times 17\times 47$，所以

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

正如所求。

求期望值的捷徑

對於 $j = 1,\ldots,26$，若第 j 個對的兩個成員都被選中，則設 $I_j = 1$，否則為 0，於是 $X = I_1 + \cdots + I_{26}$。選中某一指定對的兩個成員後，須從其餘 50 個數中再選 5 個，所以

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

由期望值的線性性質（這不需要獨立性（independence）；事實上這些 I_j 並不獨立），

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

驗算

- 四個概率之和為 1，而兩種方法都給出相同的 $\mathrm{E}(X) = \frac{7}{17} \approx 0.412$。
- 以電腦模擬 300 000 次隨機選取，得到 $X = 0, 1, 2, 3$ 的相對頻率約為 0.630、0.329、0.039、0.001，而準確值為 0.6294、0.3304、0.0393、0.0009。
- 合理性檢查：$\mathrm{P}(X = 3)$ 非常小，這合乎預期，因為只有七個數卻要有三個完整對。

常見錯誤

- 把 $X = 0$ 的選法數目算成 $\binom{52}{7}$ 減去某個粗略的數，或只算 $\binom{26}{7}$，忘記了每個對取哪一個成員所帶來的因子 2^7。
- 對於 $X = 1$，把其餘五個數的選法算成 $\binom{50}{5}$：這容許出現第二個完整對，因而多算了。

STEP 之外

把一個計數寫成指示變量（indicator variable）之和，再運用期望值的線性性質，是大學一年級概率論中最有用的技巧之一：即使分佈根本無法求出，它仍能給出期望值。本題的完整分佈尚可求得，但例如要求隨機排列（permutation）的不動點（fixed point）的期望數目，指示變量便是最自然的切入點。Blitzstein 與 Hwang 的 *Introduction to Probability*（《概率導論》）一書反覆運用這個想法。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。