詳細解答

STEP I 2007，第 13 題

概率與統計 • 條件概率

STEP I 2007，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/07-s1-q13

本題在 *Beyond STEP* 第 12.1 節精確地取條件中有完整的引導式講解：https://beyondstep.keithmcnulty.org/zh-hant/ch12/conditioning.html

關鍵思路

雖然圓片是逐個取出的，但每一組四個圓片出現的可能性都相同，所以大部分工夫都花在數組合（combination）上。真正值得留意的是各個條件。知道第一個取出的是哪一個圓片，或者只知道某個指定的圓片確實被取出，餘下三個圓片都仍然是從剩下十個之中隨機選出的。只知道有某個有號碼的圓片被取出，則是性質完全不同的資訊，而最後兩部分正是為了顯示這個分別而設。

解答

設 N 為取出的四個圓片之中有號碼的圓片數目。由於袋子經過搖勻，每一個由四個不同圓片組成的有序序列出現的可能性都相同，因此每一個由四個圓片組成的無序集合出現的可能性也相同。這樣的集合共有

$$\binom{11}{4} = 330$$

個。要組成一個恰好有 k 個有號碼圓片的集合，須從 5 個有號碼的圓片中選 k 個，再從 6 個空白圓片中選 $4-k$ 個，所以

$$\mathrm{P}(N=k) = \frac{\binom{5}{k}\binom{6}{4-k}}{330}.$$

第 (i) 部分

四個全部有號碼，即 $k = 4$：

$$\mathrm{P}(N = 4) = \frac{\binom{5}{4}\binom{6}{0}}{330} = \frac{5}{330} = \frac{1}{66}.$$

換一個算法，按次序逐個抽取，$\frac{5}{11} \cdot \frac{4}{10} \cdot \frac{3}{9} \cdot \frac{2}{8} = \frac{1}{66}$。

第 (ii) 部分

設 T_1 為「3」號圓片最先被取出這個事件（event）。已知 T_1 發生，搖勻的袋子裏現在有 10 個圓片，其中 4 個有號碼、6 個空白，而接着取出的三個圓片是從這些圓片中隨機選出的。四個圓片全部有號碼，當且僅當這三個全部有號碼，所以

$$\mathrm{P}(N = 4 \mid T_1) = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{4}{120} = \frac{1}{30}.$$

第 (iii) 部分

仍然已知 T_1 發生，我們需要從餘下十個圓片抽出的三個之中，恰好再有一個有號碼的圓片：

$$\mathrm{P}(N = 2 \mid T_1) = \frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} = \frac{4 \times 15}{120} = \frac{1}{2}.$$

第 (iv) 部分

現在設 T 為「3」號圓片在某一次被取出這個事件。

? **想一想** 「3」號圓片是最先取出，還是在其他某一次取出，有沒有分別？

沒有分別。已知 T 發生，取出的四個圓片就是「3」號圓片連同另外三個，而根據對稱性（symmetry），從其餘十個圓片中選出三個的每一個集合，出現的可能性都相同。這正是第 (iii) 部分的情況，當中抽取的次序從來都無關重要。所以答案同樣是 $\frac{1}{2}$。

直接驗證如下：$\mathrm{P}(T) = \binom{10}{3}/\binom{11}{4} = \frac{120}{330} = \frac{4}{11}$，而包含「3」號圓片、另外一個有號碼圓片和兩個空白圓片的集合共有 $\binom{4}{1}\binom{6}{2} = 60$ 個，所以

$$\mathrm{P}(N = 2 \mid T) = \frac{60/330}{120/330} = \frac{1}{2}.$$

第 (v) 部分

設 F 為第一個圓片有號碼這個事件。按第一個圓片上的號碼把 F 分拆：F 是「j 號圓片最先被取出」這五個事件（$j = 1, \ldots, 5$）的不相交併集（disjoint union），每個事件的概率

都是 $\frac{1}{11}$。第 (iii) 部分的論證並沒有用到號碼是 3 這一點，所以對每個 j 都有

$$\mathrm{P}(N = 2 \mid j \text{ 號圓片最先}) = \tfrac{1}{2}.$$

因此，根據全概率公式（law of total probability），

$$\mathrm{P}(N = 2 \mid F) = \sum_{j=1}^{5} \frac{1/11}{5/11} \cdot \frac{1}{2} = \frac{1}{2}.$$

（直接的做法：已知 F 發生，其餘三個圓片是從 4 個有號碼和 6 個空白的圓片中隨機抽出的三個，這又回到第 (iii) 部分。）

第 (vi) 部分

設 A 為最少有一個有號碼的圓片被取出這個事件，即 A 就是事件 $N \geq 1$。這與之前的資訊並不屬於同一類：它沒有指明某一個圓片，再讓其餘的保持隨機。我們改為使用條件概率（conditional probability）的定義。由於 $N = 2$ 蘊含 A，

$$\mathrm{P}(N = 2 \mid A) = \frac{\mathrm{P}(N = 2)}{\mathrm{P}(N \geq 1)}.$$

想一想 為甚麼不可以直接說「已知有一個有號碼的圓片，所以其餘三個是從剩下十個之中隨機抽出的」？

因為我們不知道已知的是哪一個圓片。一個有兩個有號碼圓片的集合，與一個有四個的集合，同樣滿足 A；並沒有哪一個圓片可以讓我們單獨抽出來放在一旁。所以這個條件是按 N 本身的值為各個集合重新加權，我們必須老老實實地數。

我們有

$$\mathrm{P}(N = 2) = \frac{\binom{5}{2}\binom{6}{2}}{330} = \frac{10 \times 15}{330} = \frac{150}{330}, \qquad \mathrm{P}(N = 0) = \frac{\binom{6}{4}}{330} = \frac{15}{330},$$

所以 $\mathrm{P}(N \geq 1) = \frac{315}{330}$，而

$$\mathrm{P}(N = 2 \mid A) = \frac{150}{315} = \frac{10}{21}.$$

驗算

N 的分佈（distribution）加起來應該正確：當 $k = 0, \ldots, 4$ 時，$\binom{5}{k}\binom{6}{4-k}$ 的數目分別是 $15, 100, 150, 60, 5$，總和為 330。在第 (ii) 和第 (iii) 部分，其餘有號碼圓片數目的條件分佈，在 120 個之中的數目分別是 $20, 60, 36, 4$，總和同樣正確。我們亦以精確分數列出全部 $11 \times 10 \times 9 \times 8 = 7920$ 個有序抽取結果，核對了全部六個答案（$\frac{1}{66}$、$\frac{1}{30}$、$\frac{1}{2}$、$\frac{1}{2}$、$\frac{1}{2}$、$\frac{10}{21}$）。留意 $\frac{10}{21} \approx 0.476$ 略小於 $\frac{1}{2}$，這是合理的：以 $N \geq 1$ 為條件只剔除了 $N = 0$ 的情況，所以與知道某個指定圓片被取出相比，它在 $N = 1$ 上留下了較多的權重。

常見錯誤

- 把第 (iv) 和第 (vi) 部分當作同一類條件。「3 號圓片被取出」指明了一個圓片；「有一個有號碼的圓片被取出」則沒有，兩者的答案並不相同。
- 在第 (vi) 部分，忘記了條件是 $N \geq 1$，反而除以 $\mathrm{P}(N = 1)$ 或 $\frac{5}{11}$。
- 分子數的是有序抽取，分母數的卻是無序集合。兩者選定其一，然後貫徹始終。

STEP 之外

第 (v) 和第 (vi) 部分之間的對比，與著名的「男孩還是女孩」悖論（"boy or girl" paradox）是同一個陷阱：得知某個指定的孩子是男孩，與得知最少有一個是男孩，會得出不同的答案。在大學裏，這一點會有嚴格的表述：以一個事件為條件，就是把概率測度（probability measure）限制在該事件上再重新歸一化（renormalise），而這裏的數目正是超幾何分佈（hypergeometric distribution）。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門），以及 Blitzstein 與 Hwang 的 *Introduction to Probability*（概率導論），都仔細討論了這些悖論。

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。