詳細解答

STEP I 2007，第 12 題

概率與統計 • 條件概率

> **STEP I 2007，第 12 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-q12

關鍵思路

抽第二粒糖果時，我們並不知道袋中的情況，因為那取決於之前發生了甚麼事。所以我們要考慮條件概率（conditional probability）：按第二次抽取之前發生的所有情況分類，求出每種情況下抽到紅色的機會，再用全概率定理（law of total probability）加起來。第 (ii) 部分有一條好用的捷徑：一個袋剛收到一粒隨機糖果後，袋中紅色糖果的期望數目（expected number）很容易寫出來，所以我們不必就轉移的那粒糖果的顏色重複分類。

解答

第 (i) 部分

設 R_1 為第一粒糖果是紅色的事件，R_2 為第二粒糖果是紅色的事件。每粒糖果被首先抽中的可能性均等，所以

$$\mathrm{P}(R_1) = \frac{a}{N}.$$

至於 R_2，按第一粒糖果的顏色分類。若第一粒是紅色，則餘下 $N-1$ 粒糖果，其中 $a-1$ 粒是紅色；若不是紅色，則餘下 $N-1$ 粒，其中 a 粒是紅色。（這裏 $N \geq 2$ 保證了第二粒糖果存在。）由全概率定理，

$$\begin{aligned}\mathrm{P}(R_2) &= \mathrm{P}(R_1)\,\mathrm{P}(R_2 \mid R_1) + \mathrm{P}(R_1')\,\mathrm{P}(R_2 \mid R_1') \\ &= \frac{a}{N}\cdot\frac{a-1}{N-1} + \frac{N-a}{N}\cdot\frac{a}{N-1} \\ &= \frac{a[(a-1)+(N-a)]}{N(N-1)} = \frac{a(N-1)}{N(N-1)} = \frac{a}{N}.\end{aligned}$$

所以 $\mathrm{P}(R_2) = \mathrm{P}(R_1)$，正是所求。（用對稱性（symmetry）也可得出同一結論：由兩粒不同糖果組成的每個有序對出現的可能性均等，所以第二粒糖果是某一粒指定糖果的機會，與第一粒是該糖果的機會相同。）

第 (ii) 部分

這裏的「第一粒糖果」是第一次擲硬幣後轉移的那一粒，「第二粒糖果」是第二次擲硬幣後選出的那一粒。設 R_1、R_2 分別為這兩粒糖果是紅色的事件。兩次擲硬幣互相獨立（independent），而每次擲硬幣的結果亦與選中哪粒糖果無關。

第一粒糖果。 按第一次擲硬幣的結果分類，

$$\mathrm{P}(R_1) = p \cdot \frac{a}{N} + q \cdot \frac{b}{N} = \frac{pa + qb}{N}.$$

第二粒糖果。 兩次擲硬幣共有四種情況。

正面、正面。先從第 1 袋取出一粒糖果，然後再從第 1 袋取出另一粒。這正是把第 (i) 部分的情況套用於第 1 袋（有 N 粒糖果，其中 a 粒紅色），所以第二粒糖果是紅色的概率為 a/N。

正面、反面。一粒糖果由第 1 袋移到第 2 袋，第 2 袋現在有 $N+1$ 粒糖果；然後我們從第 2 袋抽取。第 2 袋有 $b+1$ 粒紅色糖果的概率為 a/N，有 b 粒紅色糖果的概率為 $1 - a/N$，所以

$$\mathrm{P}(R_2 \mid \mathrm{HT}) = \frac{a}{N} \cdot \frac{b+1}{N+1} + \left(1 - \frac{a}{N}\right) \frac{b}{N+1} = \frac{b + \frac{a}{N}}{N+1} = \frac{Nb + a}{N(N+1)}.$$

? **想一想** 為甚麼第二行可以化簡得如此乾淨俐落？

分子就是轉移之後第 2 袋中紅色糖果的期望數目：原有的 b 粒，加上新來的糖果是紅色的機會 a/N。由於在這種情況下第 2 袋的大小固定為 $N+1$，抽到紅色的機會就是這個期望數目除以 $N+1$。

反面、正面。把兩個袋對調，用同樣的論證可知，第 1 袋現在有 $N+1$ 粒糖果，其中紅色糖果的期望數目為 $a + b/N$，所以

$$\mathrm{P}(R_2 \mid \mathrm{TH}) = \frac{Na + b}{N(N+1)}.$$

反面、反面。把第 (i) 部分套用於第 2 袋，得概率為 b/N。

現在把四種情況合併，它們的概率分別為 p^2、pq、qp、q^2：

$$\begin{aligned}\mathrm{P}(R_2) &= p^2\frac{a}{N} + pq\,\frac{(Nb+a)+(Na+b)}{N(N+1)} + q^2\frac{b}{N}\\ &= p^2\frac{a}{N} + pq\,\frac{(N+1)(a+b)}{N(N+1)} + q^2\frac{b}{N}\\ &= \frac{p^2a + pqa + pqb + q^2b}{N}\\ &= \frac{pa(p+q)+qb(p+q)}{N} = \frac{pa+qb}{N},\end{aligned}$$

最後一步用了 $p+q=1$。這等於 $\mathrm{P}(R_1)$，正是我們要證明的結果。

驗算

- 若 $p=1$，一切都在第 1 袋中發生，兩粒糖果的答案 a/N 正好就是第 (i) 部分。若 $p=0$，則得到第 2 袋的第 (i) 部分。
- 若 $a=b=0$，則沒有紅色糖果，兩個概率都是 0；若 $a=b=N$，則每粒糖果都是紅色，兩個概率都是 1。
- 取 $N=5$、$a=2$、$b=4$、$p=0.3$，公式給出 $(0.6+2.8)/5=0.68$。把整個過程模擬 200 000 次，第一粒糖果的結果為 0.679，第二粒為 0.681；而對所有較小的 N、a、b 逐一窮舉計算，結果亦吻合。

常見錯誤

- 忘記了轉移之後，收到糖果的袋有 $N+1$ 粒糖果而不是 N 粒，而送出糖果的袋則有 $N-1$ 粒。
- 在 HH 和 TT 兩種情況中，把第二次抽取當作與第一次獨立。其實兩者並不獨立，但第 (i) 部分告訴我們答案始終不變，所以應直接引用，而不必重新計算。

STEP 之外

這兩部分都是同一個深刻想法的例子：如果你從甕中抽取一件東西而不看抽到的是甚麼，你對甕的認識並沒有改變，所以下一次抽取的分佈（distribution）依然相同。嚴格來說，這一連串抽取是**可交換的**（exchangeable），而在第 (ii) 部分中，$p\times$(第 1 袋中紅色所佔比例) $+\,q\times$(第 2 袋中紅色所佔比例) 這個量，即下一粒抽出的糖果是紅色的機會，是一個**鞅**（martingale）：它未來的期望值等於它現在的值。鞅是現代概率論的核心，由賭博到金融都有它的身影。Grimmett 與 Welsh 合著的 *Probability:*

An Introduction（概率論入門教材）是認識條件期望（conditional expectation）和這些想法的好起點。

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。