詳細解答

# STEP I 2000，第 12 題

概率與統計 • 幾何分佈

> **STEP I 2000，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/00-s1-q12

## 關鍵思路

每一種情況都先問：已知之前每一次嘗試都失敗，下一次成功的機會是多少？然後把首 $n-1$ 次嘗試失敗的概率（probability）與第 $n$ 次成功的概率相乘。三種規則得出的答案形態截然不同：一種是條件概率（conditional probability）不斷增大，一種是由一開始就保持不變，還有一種是第一次之後才保持不變。

全篇以 $N$ 表示所需的嘗試次數，因此我們要求的是 $\mathrm{P}(N=n)$，其中 $n=1,2,3,\ldots$。我們把能開門的那條鑰匙稱為正確鑰匙。

## 解答

### 第 (i) 部分

這裏我絕不重複使用任何鑰匙，所以我是按一個均勻隨機的次序逐一試鑰匙。根據對稱性（symmetry），正確鑰匙落在這個次序中 $k$ 個位置的任何一個，可能性都相同。因此

$$\mathrm{P}(N=n)=\frac{1}{k}\quad \text{其中} n=1,2,\ldots,k, \qquad \mathrm{P}(N=n)=0 \text{ 當} n>k.$$

如果覺得這個對稱性論證來得太取巧，可以改為把條件概率相乘。失敗 $j$ 次之後，還有 $k-j$ 條未試過的鑰匙，其中一條是正確的，所以對於 $1\le n\le k$

$$\mathrm{P}(N=n)=\frac{k-1}{k}\cdot\frac{k-2}{k-1}\cdots\frac{k-n+1}{k-n+2}\cdot\frac{1}{k-n+1}=\frac{1}{k},$$

因為這個乘積會逐項相消（telescope）。

### 第 (ii) 部分

現在每一次嘗試都是從全部 $k$ 條鑰匙中重新、獨立（independent）地選取，所以不論之前發生過甚麼，每一次嘗試成功的概率都是 $\frac{1}{k}$。要剛好嘗試 $n$ 次，首 $n-1$ 次必須失敗，而第 $n$ 次必須成功：

$$\mathrm{P}(N=n)=\left(\frac{k-1}{k}\right)^{n-1}\frac{1}{k},\qquad n=1,2,3,\ldots$$

這是參數為 $\frac{1}{k}$ 的幾何分佈（geometric distribution），而 $n$ 並沒有上限。

### 第 (iii) 部分

第一次嘗試成功的概率是 $\frac{1}{k}$，所以 $\mathrm{P}(N=1)=\frac{1}{k}$。

現在假設第 $m$ 次嘗試（對某個 $m\geq 1$）失敗了。那麼第 $m$ 次所試的是一條錯的鑰匙。在第 $m+1$ 次嘗試，我從其餘 $k-1$ 條鑰匙中均勻地選取。

> **想一想** 第 $m$ 次嘗試之前的經過，會否影響第 $m+1$ 次嘗試成功的機會？

不會。唯一有關係的是哪一條鑰匙被排除，而被排除的永遠是一條錯的鑰匙。所以可供選擇的 $k-1$ 條鑰匙永遠包括正確的那一條，而不論之前發生過甚麼，第 $m+1$ 次嘗試成功的概率都是 $\frac{1}{k-1}$，失敗的概率都是 $\frac{k-2}{k-1}$。

所以對於 $n\geq 2$：第一次嘗試失敗（概率為 $\frac{k-1}{k}$），第 $2,\ldots,n-1$ 次嘗試失敗（每次概率為 $\frac{k-2}{k-1}$），而第 $n$ 次嘗試成功（概率為 $\frac{1}{k-1}$）。因此

$$\mathrm{P}(N=n)=\frac{k-1}{k}\left(\frac{k-2}{k-1}\right)^{n-2}\frac{1}{k-1}=\frac{1}{k}\left(\frac{k-2}{k-1}\right)^{n-2},\qquad n\geq 2,$$

另外還有 $\mathrm{P}(N=1)=\frac{1}{k}$。

有趣的是，同樣有 $\mathrm{P}(N=2)=\frac{1}{k}$，與第 (i) 部分的結果一樣。（當 $k=2$ 時，這個規則迫使我輪流使用兩條鑰匙，所以 $N$ 是 1 或 2，概率各為 $\frac{1}{2}$；只要把 $0^0$ 理解為 1，公式便給出這個結果。）

## 驗算

- 每個分佈的總和都是 1。在 (iii) 中，$\frac{1}{k}+\frac{1}{k}\sum_{j\geq 0}\left(\frac{k-2}{k-1}\right)^j=\frac{1}{k}+\frac{1}{k}(k-1)=1$。
- 各個平均值（mean）都合乎情理：(i) 是 $\frac{k+1}{2}$，(ii) 是 $k$，(iii) 是 $1+\frac{(k-1)^2}{k}$。與 (ii) 相比，不重複使用上一條鑰匙略有幫助，但遠遠不及完全不重複使用任何鑰匙。
- 我們對 $k=2,\ldots,5$ 和 $n\leq 7$，以分數列舉每一個選擇序列，精確地核對了全部三條公式。就 $k=5$ 的 (iii) 進行 $2\times 10^5$ 次模擬（simulation）試驗，得出 $\mathrm{P}(N=3)\approx$

0.1496，而精確值是 $\frac{1}{5} \cdot \frac{3}{4} = 0.15$。

## 常見錯誤

- 在 (iii) 中，第一次嘗試也用了 $\frac{k-2}{k-1}$。第一次是從全部 $k$ 條鑰匙中選取，所以第一個因子是 $\frac{k-1}{k}$。
- 在 (i) 中，忘記說明當 $n > k$ 時 $\mathrm{P}(N = n) = 0$。
- 在 (iii) 中，以為被排除的鑰匙可能是正確的那一條。其實被排除的永遠只是剛剛試過而失敗的鑰匙。

## STEP 之外

第 (iii) 部分是一條小型的馬可夫鏈（*Markov chain*）。接下來發生甚麼只取決於當前的狀態（state）（這裏是指上一次試的是哪一條鑰匙），而與整段歷史無關。找出一個令過去變得無關重要的狀態，正是馬可夫鏈理論的核心技巧，它把繁複的條件概率化為簡單的乘積。Grimmett 與 Welsh 的 *Probability: An Introduction*（概率入門教材）正是以這個程度介紹馬可夫鏈。

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。