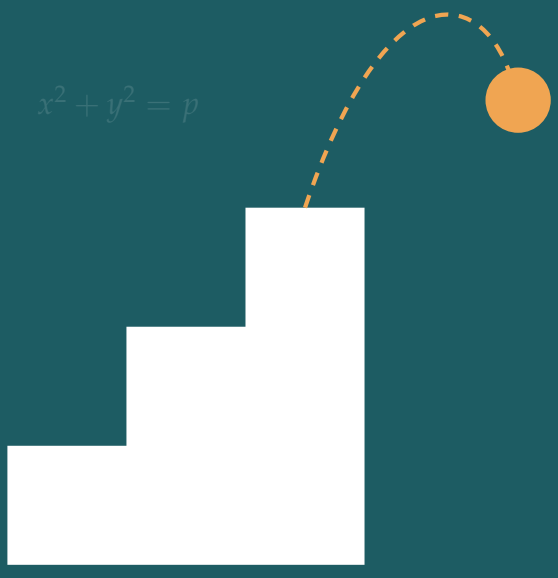
STEP 備試 · 大學數學

Beyond STEP

第 9 章

計數與機會



KEITH McNULTY

Beyond STEP

以歷屆 STEP 試題為門
走進大學數學

第 9 章

計數與機會

Keith McNulty

Beyond STEP：以歷屆 STEP 試題為門，走進大學數學

第 9 章：計數與機會

本冊是 *Beyond STEP* 全書 13 章中的第 9 章。全書、各章的單行本、每一節的可列印工作紙，以及每一道 STEP 試題的詳細解答，均可在 `beyondstep.keithmcnulty.org` 免費閱讀及下載。

2026 年 11 月首次於網上發表。

本 PDF 內凡屬作者本人創作的內容，包括正文、解答、插圖、練習及答案，均按照 Creative Commons CC0 1.0 Universal 授權條款貢獻予公有領域：你可以為任何目的複製、改編及分享，毋須徵求許可。詳見 `creativecommons.org/publicdomain/zero/1.0`。

本書討論的 STEP 試題，版權屬 Cambridge University Press & Assessment 所有。本書並無轉載這些試題：每一題均附連結，通往官方來源的題目，即 STEP Support Programme 的試題資料庫或 OCR 的歷屆試卷。

如有錯漏，歡迎到 `github.com/keithmcnulty/step_book_mcnulty` 提交 issue 指正。

章

9

計數與機會

計數是最古老的數學，但要數得巧妙，至今仍是最困難的事情之一。

本章內容

9.1	直線分割平面	STEP II 1990
9.2	小型版的拉姆齊定理	STEP I 1989
9.3	為籬笆樁着色	STEP II 1995
9.4	把數列看成函數	STEP I 2016
9.5	湊零錢	STEP I 1997
9.6	奇怪的骰子	STEP I 1993
9.7	相等的冪和	STEP I 1992
9.8	沒有人拿到自己的恤衫	STEP II 2016
9.9	生日問題	STEP II 2007
9.10	總能找贖	STEP I 2011
9.11	集齊一套	STEP II 1994
9.12	隨機子集	STEP III 2019
9.13	做了手腳的骰子	STEP I 1999
A	練習答案	第 89 頁

計數的藝術

組合數學是拓撲學的貧民窟。

— J. H. C. Whitehead，說於 1950 年代。今天已經沒有人抱持這種看法。

人人都會數數。但若要問：用硬幣湊出 £1 有多少種方法？安排十個人入座，使得沒有人坐在自己的位子上，有多少種方法？用四種顏色為一幅地圖着色，又有多少種方法？單純的點算很快就行不通了。情況多得無法逐一列出，而且各種情況還會互相重疊。組合數學（combinatorics）是關於計數與排列的數學，它是一門藝術：把這類問題整理得井井有條，使我們不必逐一列舉也能解決。

正如 Whitehead 那句話所反映的，在很長一段時間裏，組合數學只被視為一堆巧妙技巧的集合，而不是一門嚴肅的學科。這種情況在二十世紀下半葉改變了，推動力一方面來自計算機科學（computer science），因為點算一個算法（algorithm）所需的步驟數目，本身就是一個組合問題；另一方面則來自人們發現它與代數、幾何和概率（probability）之間有深刻的聯繫。今天，它是數學中最活躍的領域之一。2022 年，許埈珥（June Huh）獲頒數學界的最高榮譽菲爾茲獎（Fields Medal），部分原因是他運用代數幾何（algebraic geometry）的思想，證明了一個 1968 年提出的猜想，這個猜想所涉及的，正是 Section 9.3 中的色多項式（chromatic polynomial）。

計數為何重要

概率最簡單的形式就是計數：當所有結果的可能性均等時，一個事件的概率等於有利結果的數目除以結果的總數。因此，每一條關於紙牌、骰子和彩票的問題，都是計數問題；本章後半部分取材自概率試卷，正是這個緣故。但計數的重要性遠遠不止於遊戲。計算機科學家要點算一個算法所需的運算次數，以及它可能出錯的方式。化學家要點算分子可能具有的結構。生物學家要點算一條 DNA 鏈可以摺疊的方式。密碼學家要點算攻擊者必須嘗試的密鑰數目，以及（如 Section 9.9 所示）要有多少個，才會出現兩個相同的情況。

你在大學會遇到甚麼

大多數大學都會在一年級或二年級開設組合數學或離散數學（discrete mathematics）的課程，其主題正是本章的主題。你會學到基本的計數原理（乘法原理（product rule）、雙射（bijection）、容斥原理（inclusion–exclusion）、鴿巢原理（pigeonhole principle）），以及怎

樣建立和求解遞推關係（recurrence relation）。你會遇到*生成函數*（generating function），它把數列變成函數，把計數變成代數。你會研究圖（graph），從着色（colouring）到拉姆齊理論（Ramsey theory），還有偏序集（partially ordered set），例如子集所構成的格（lattice）。概率課程也會運用同樣的工具，再加上期望值的線性性質（linearity of expectation），來分析隨機結構。再往後，組合數學會與代數（對稱函數、表示論）、數論（分拆、加性組合）以及計算機科學（算法分析、複雜度理論）連繫起來。

貫穿本章的問題

三大問題

1. **怎樣不逐一列舉而完成計數？**有哪些結構（遞推關係、雙射、生成函數）能讓你準確地點算龐大的集合？
2. **甚麼時候某件事必然發生？**甚麼時候單憑數量之大，就足以迫使某種規律、巧合或配對出現？
3. **計數與概率怎樣互相幫助？**甚麼時候最好用計數來求概率，甚麼時候又最好用概率來求數目？

本章分為兩半。前七節取材自純數學試卷，先講遞推關係（直線分割平面、為籬笆樁着色）和鴿巢原理（拉姆齊定理），然後建立生成函數，並用它來點算硬幣、設計奇特的骰子，以及把數字分成冪和相等的隊伍。後六節是概率試卷中的經典計數問題：錯排（derangement）與 $1/e$ 的出現、生日問題、選票問題及其反射原理（reflection principle）、優惠券收集問題（coupon collector）、隨機子集，以及為何不可能把骰子灌鉛，使得每個總和出現的可能性均等。

本章的思想

§	課題	思想
9.1	平面的區域	多加一個時，有甚麼改變
9.2	拉姆齊理論	鴿巢原理
9.3	圖着色	耦合遞推關係；色多項式
9.4	生成函數	數列變成函數
9.5	分拆（partition）	透過係數來計數
9.6	Sicherman 骰子	多項式的因式分解
9.7	Thue–Morse	重根帶來平衡的和
9.8	錯排	容斥原理；$1/e$
9.9	巧合	補事件；平方根定律
9.10	選票問題	反射原理
9.11	優惠券收集問題	期望值的線性性質；$n \ln n$
9.12	隨機子集	每次只看一個元素
9.13	灌鉛骰子	以算術—幾何平均不等式（AM–GM）證明不可能

怎樣閱讀本章

永遠從小的情況入手。畫三條直線，列出三張牌的六種次序，數一數湊出 5p 的方法。小的情況能給你數據去猜想，最後也可以用來檢驗你的公式。但在找到理由之前，不要輕信任何規律：Section 9.1 中有一個著名的數列 $1, 2, 4, 8, 16, \ldots$，它的下一項是 31。

9.1 直線分割平面

題目

STEP II 1990，第 4 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/90-s2-q4`

引子

用 n 刀直線把一個薄餅切開，各塊不必大小相同，最多可以切成多少塊？一刀得 2 塊，兩刀得 4 塊，三刀得 7 塊（不是 8 塊：不妨試試），四刀得 11 塊。數列 $1, 2, 4, 7, 11, 16, \ldots$ 稱為*懶惰宴會承辦人數列*（lazy caterer's sequence），而本題要求的正是它的公式。

這道題是一種貫穿整個組合數學（combinatorics）的技巧的第一課：不要試圖一次過把所有東西數完，而應該問：多加一個對象時，有甚麼改變？而答案原來是歐拉公式（Euler's formula）$V - E + F = 2$ 的一個特殊情況，你在 Section 7.9 討論柏拉圖立體時已經見過這條公式。

動筆之前

> **停一停，試一試** 仔細畫出 $n = 1, 2, 3, 4$ 的情況，並保持各直線處於一般位置（general position）（沒有兩條平行，沒有三條共點）。數一數每種情況中線段和區域的數目。你能猜出當中的規律嗎？

你應該會得到：

n	0	1	2	3	4
線段	0	1	4	9	16
區域	1	2	4	7	11

線段的數目看來是 n^2。區域的數目則依次增加 $1, 2, 3, 4$：相鄰兩項之差正是自然數。

逐步解題

階段 1　數線段

> **想一想**　集中看其中一條直線。其餘的直線中有多少條與它相交？它們把這條直線切成多少段？

每條直線都與其餘 $n-1$ 條直線逐一相交（沒有一條與它平行），而且交於 $n-1$ 個不同的點（沒有三條直線交於同一點）。這 $n-1$ 個點把該直線切成 n 段：兩端各有一段無限長的，中間有 $n-2$ 段有限長的。共有 n 條直線，所以共有

$$n \times n = n^2 \text{ 條線段.} \qquad \blacksquare$$

這是直接點算，無需使用歸納法（induction）：我們是逐條直線去數線段的。

階段 2　加一條直線時會怎樣？

> **想一想**　假設已畫好 $n-1$ 條直線，現在加上第 n 條。它被其餘直線切成 n 段。每一段對它所穿過的區域有甚麼影響？

畫出新直線時，它的 n 段各自穿過一個原有的區域，並把該區域一分為二（圖 9.1.1）。不同的線段位於不同的區域內，因為新直線要由一個區域走到另一個區域，就必須越過一條原有的直線。因此，加上第 n 條直線後，區域的數目恰好增加 n：

$$R(n) = R(n-1) + n.$$

階段 3　解遞歸關係

由 $R(0) = 1$（沒有直線，只有一個區域）開始，

$$R(n) = 1 + (1 + 2 + \cdots + n) = 1 + \frac{n(n+1)}{2} = \frac{n^2 + n + 2}{2}.$$

若要寫成歸納法證明：公式對 $n = 1$ 成立，因為 $\tfrac{1}{2}(1+1+2) = 2$；而若 $R(n-1) = \tfrac{1}{2}\big((n-1)^2 + (n-1) + 2\big) = \tfrac{1}{2}(n^2 - n + 2)$，則 $R(n) = \tfrac{1}{2}(n^2 - n + 2) + n = \tfrac{1}{2}(n^2 + n + 2)$。 $\blacksquare$

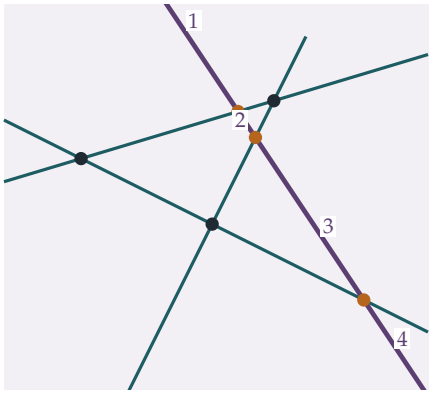


圖 9.1.1　三條直線（藍綠色）把平面分成 7 個區域。第四條直線（紫紅色）與其餘三條相交於琥珀色的點，這些點把它切成四段，編號為 1 至 4。每一段把一個區域一分為二，所以第四條直線增加了 4 個區域，合共 11 個。

回顧

- **每次只數一部分。**線段是逐條直線去數的；區域則是按每條新直線所增加的數目去數的。
- **把幾何上的一步化為遞歸關係。**「每條新線段把一個區域一分為二」這項觀察，化成了 $R(n) = R(n-1) + n$。
- **一般位置十分重要。**若有平行的直線，或有三條直線共點，區域的數目就會較少。試找出每項假設在哪裏用過。

背後的數學

喬裝的歐拉公式

以下是第二個證明，它不僅推導出答案，更解釋了答案從何而來。把這些直線組成的圖形看成一幅地圖：交點是頂點（vertices），線段是邊（edges），區域是面（faces）。圖中有 $V = \binom{n}{2}$ 個頂點（每一對直線對應一個）和 $E = n^2$ 條邊。無限長的線段只有一個端點，所以我們額外加上一個「在無窮遠處」的頂點，讓它們全部在那裏相遇；這樣，整幅圖就成為畫在球面上的地圖（想像把平面包在一個球上，無窮遠點位於球頂）。由 Section 7.9 的歐拉公式 $V - E + F = 2$ 可得

$$\left(\binom{n}{2} + 1\right) - n^2 + F = 2, \qquad \text{所以} \qquad F = n^2 - \frac{n(n-1)}{2} + 1 = \frac{n^2 + n + 2}{2}.$$

同一個數，卻來自完全不同的方向。這些區域還可以整齊地分成兩類：其中 $2n$ 個是無界的（每兩條相鄰的無限射線之間有一個），其餘 $\frac{1}{2}(n^2 + n + 2) - 2n = \binom{n-1}{2}$ 個是有界

的。

二項式係數與高維空間

這條公式可以寫成

$$R(n) = \binom{n}{0} + \binom{n}{1} + \binom{n}{2},$$

而這並非巧合。把圖形轉動，使得沒有一條直線是水平的。一個下方有界的區域必有一個最低點，而這個點必定是交點；每個交點又恰好是一個區域（即它正上方的那個）的最低點：這樣就得到 $\binom{n}{2}$ 個區域，每一對直線對應一個。其餘每個區域都向下無限延伸，因而會與一條畫在所有交點下方很遠處的水平線相遇；那 n 條直線把這條水平線切成 $n+1$ 個區間，每個這樣的區域內各有一個。合共 $\binom{n}{2} + n + 1$ 個。在三維空間中，n 個處於一般位置的平面把空間切成

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} = \frac{n^3 + 5n + 6}{6}$$

塊（即蛋糕數（cake numbers）$1, 2, 4, 8, 15, 26, \ldots$），而這個規律在任何維數都成立。證明用的是同一個歸納法：新加的平面被其餘平面切成若干區域，其排列方式就如平面上的直線所分成的區域，而每個這樣的區域都把一塊空間一分為二。

施泰納與超平面排列

直線和平面所分割出的區域數目的公式，是施泰納（Jakob Steiner）在 1826 年發現的。今天，它是*超平面排列*（hyperplane arrangements）理論的起點；這套理論點算任意一組直線、平面及其高維類比所形成的區域數目，包括並非處於一般位置的情況。Thomas Zaslavsky 在 1975 年證明了一條非凡的定理：利用一個記錄了哪些超平面相交的多項式，就可以數出任何排列的區域數目。同樣的想法也出現在機器學習之中：由「ReLU」單元構成的神經網絡，會用超平面把輸入空間切成多個區域，而區域的數目正好衡量它所能表示的函數有多複雜。

思維技巧

> **問一問：多加一個時，有甚麼改變**
>
> *要點算由 n 個對象組成的構形，可把它與 $n-1$ 個對象的構形比較。點算其中的改變，往往比點算整體容易得多。*
>
> 這是把歸納法用作發現的工具，而不僅是證明的工具。它是組合數學中最常見的手法之一，能把幾何問題化為可以求解的遞歸關係（recurrence relation）。

延伸練習

1 **圓。**證明：若 n 個圓中每兩個都交於兩點，且沒有三個圓共點，則它們把平面分成 $n^2 - n + 2$ 個區域。

2 **平行線。**若 n 條直線中有 k 條互相平行（其餘方面均處於一般位置），共有多少個區域？

3 **蛋糕數。**用歸納法證明：n 個處於一般位置的平面把空間分成 $\frac{1}{6}(n^3 + 5n + 6)$ 個區域。

4 **會失效的規律。**在圓周上取 n 個點，並把每一對點用弦連起來，其中沒有三條弦在圓內交於同一點。當 $n = 1, 2, 3, 4, 5$ 時，分別得到 $1, 2, 4, 8, 16$ 個區域。數一數 $n = 6$ 時的區域數目。（提示：使用歐拉公式；每四個點給出一個交點。）證明區域的數目是 $1 + \binom{n}{2} + \binom{n}{4}$。

延伸閱讀

- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley），第 1 章。開首討論的正是這個問題，還有「薄餅」和「之字形」的變化題。
- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific）。一本平易近人的入門教材，設有歸納法、圖和歐拉公式等章節。
- **Richard Guy, "The strong law of small numbers"**，*American Mathematical Monthly* 95 (1988)。收錄了許多看似明顯、其實不成立的規律，包括練習中的圓問題。

9.2 小型版的拉姆齊定理

題目

STEP I 1989，第 4 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/89-s1-q4`

引子

在任何一個六人派對上，要麼有三個人彼此都認識，要麼有三個人彼此都不認識。無論這六個人是誰，這一點都必定成立。換成五個人，就沒有這個保證了。這個小小的事實是拉姆齊定理（*Ramsey's theorem*）的第一個情形，它道出了一條貫穿現代組合數學（combinatorics）的原則：完全的無序是不可能的。任何足夠大的結構，無論怎樣着色或排列，都必定包含一個高度有序的部分。

證明只需四行。但順理成章的下一個問題（要多少人才能保證有四個互相認識的人或四個互不認識的人？五個呢？）卻通向數學中一些最困難的未解難題。

動筆之前

想一想 那六個點之所以重要，只在於哪些點對以哪種顏色相連：幾何只會令人分心。一共有多少條線段？每個點又連出多少條？

一共有 $\binom{6}{2} = 15$ 條線段，每個點連出 5 條。（一般位置（general position）這個條件只是確保任意三點都構成一個真正的三角形。這道題其實是關於六個頂點上的完全圖（*complete graph*）K_6，它的邊以兩種顏色着色。）

逐步解題

階段 1　五條線段，兩種顏色

想一想 從 A 連出五條線段，每條不是金色就是銀色。關於它們的顏色，你可以肯定甚麼？

其中至少有三條顏色相同。假如金色的至多兩條，銀色的也至多兩條，那麼線段至多只有四條。這就是鴿巢原理（*pigeonhole principle*）：把五件物件放進兩個盒子，必有一個盒子裏至少有三件。不妨設（有需要時把兩種顏色的名稱對調）AB、AC 和 AD 全是金色。

階段 2　看看另一端

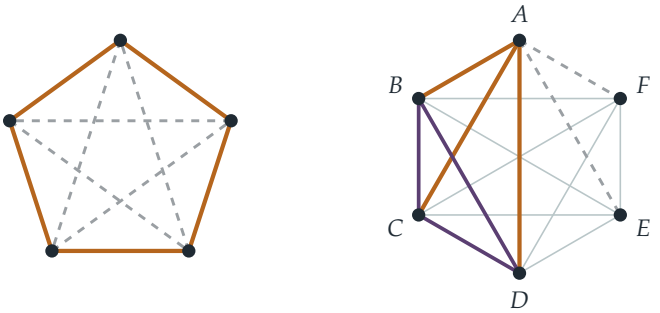
想一想 現在看看三角形 BCD。如果它有任何一條邊是金色，會怎樣？如果一條都不是呢？

如果三角形 BCD 有任何一條邊是金色，比如 BC，那麼 ABC 的三條邊 AB、AC 和 BC 全是金色。如果 BCD 沒有一條邊是金色，那麼 BC、CD 和 DB 全是銀色，BCD 就是一個銀色三角形。無論哪種情況，都有一個單色三角形（圖 9.2.1，右）。 ■

階段 3　五個點並不足夠

想一想 若只有五個點，每個點只連出四條線段。怎樣着色，才能使每個點連出的線段恰好每種顏色各兩條？

把五個點放成一個正五邊形（如果想要它們處於一般位置，可以把它們稍微提離平面）。把五邊形的五條邊塗成金色，五條對角線塗成銀色（圖 9.2.1，左）。金色的邊構成一個五圈，當中沒有三角形；銀色的邊構成一個五角星，它同樣是一個五圈，當中也沒有三角形。所以並沒有單色三角形。 ■



五個點：沒有單色三角形　　六個點：看看 BCD

圖 9.2.1　左：五個點，五邊形為金色（實線），五角星為銀色（虛線），沒有單色三角形。右：六個點時，從 A 連出的五條線段中有三條顏色相同，比如連到 B、C、D 的三條都是金色。這時三角形 BCD（紫紅色）決定了結果：只要有一條邊是金色，就與 A 構成一個金色三角形；沒有金色的邊，BCD 就是一個銀色三角形。

回顧

- **撇除無關的東西。**空間中的點變成了一個着色的圖；重要的只有邊的顏色。
- **鴿巢原理。**五條邊、兩種顏色，迫使其中三條同色。
- **對合適的對象分情況討論。**一旦找到三條金色的邊，它們另一端的三角形就決定了一切。
- **緊確性（sharpness）。**五邊形的例子說明「六」這個數不能再降低。一條好的定理，總會附帶一個例子，說明它已是最佳的結果。

背後的數學

拉姆齊數

拉姆齊數（*Ramsey number*）$R(s,t)$ 是滿足以下條件的最小的 N：把 K_N 的邊以金、銀兩色着色，無論怎樣着色，都必定包含一個金色的 K_s（即 s 個點，它們之間的所有邊都是金色）或一個銀色的 K_t。這道題已證明了 $R(3,3)=6$：六個點必定可行，而五邊形說明五個點並不可行。

弗蘭克·拉姆齊（Frank Ramsey）在 1928 年證明了這些數必定存在：對任意的 s 和 t，總有某個 N 符合要求。證明用的是歸納法（induction），把上面的論證加以推廣，並得出

$R(s,t) \le R(s-1,t) + R(s,t-1)$。但要找出確切的數值卻異常困難：

$$R(3,3) = 6, \qquad R(4,4) = 18, \qquad 43 \le R(5,5) \le 46.$$

$R(4,4)$ 的值是在 1955 年求得的。儘管已動用電腦作大規模搜尋，$R(5,5)$ 至今仍然未知；上界 46 到 2024 年才獲證明。電腦無法直接檢查 K_{43} 的每一種着色：一共有 2^{903} 種，遠遠多於宇宙中原子的數目。

艾狄胥與外星人

保羅·艾狄胥（Paul Erdős）對拉姆齊理論（Ramsey theory）發展的貢獻無人能及，他很喜歡講這個故事。假設外星人入侵地球，並威脅說除非我們告訴他們 $R(5,5)$ 的值，否則就毀滅地球。我們應當調動所有電腦和數學家，這樣或許能把它找出來。但假如他們問的是 $R(6,6)$，我們就應該設法消滅外星人。艾狄胥還在 1947 年證明了 $R(k,k) > 2^{k/2}$，所用的方法令人驚歎：他證明了，在少於 $2^{k/2}$ 個點上隨機着色的圖，有正的概率不包含任何單色的 K_k，所以好的着色必定存在，儘管沒有人知道怎樣把它構造出來。這種*概率方法*（*probabilistic method*）如今是組合數學中最強大的工具之一。

亂中有序

拉姆齊定理有很多近親，它們表達的都是同一個想法：足夠大的結構無法避開秩序。*范德瓦爾登定理*（*Van der Waerden's theorem*）說，如果把所有整數以有限種顏色着色，就存在任意長的單色等差數列（arithmetic progression）。*艾狄胥—塞凱賴什定理*（*Erdős–Szekeres theorem*）說，任何由 n^2+1 個相異的數組成的數列，都包含一個長度為 $n+1$ 的遞增或遞減子數列（subsequence）。而拉姆齊本人感興趣的是一個邏輯問題：他是在一篇關於判定邏輯命題真假的論文中，把這條定理作為引理（lemma）來證明的。

思維技巧

鴿巢原理

如果把多於 kn 件物件放進 n 個盒子，則必有一個盒子裏有多於 k 件物件。

這條原理顯而易見，但要用得好卻是一門藝術：功夫在於選定甚麼是物件、甚麼是盒子。在這道題中，物件是某一點上的五條邊，盒子是兩種顏色。在大學裏，鴿巢原理的論證會出現在數論（以分數逼近無理數）、分析（波爾查諾–魏爾斯特拉斯定理（Bolzano–Weierstrass theorem））之中，更遍及整個組合數學。

延伸練習

1 **至少兩個。**證明在 K_6 的任何雙色着色中，其實至少有兩個單色三角形。（提示：數一數「雙色角」的數目，即在同一個頂點上的一對不同顏色的邊。若一個點有 g 條金色的邊和 $5-g$ 條銀色的邊，它就是 $g(5-g) \le 6$ 個這樣的角的頂點。每個非單色的三角形恰好包含兩個這樣的角。）

2 **三種顏色。**如果把 K_{17} 的邊以三種顏色着色，證明其中有一個單色三角形。（提示：有一個點至少有六條同色的邊。）

3 **三或四。**證明 $R(3,4) \le 10$：在 K_{10} 的任何金、銀着色中，都有一個金色三角形或一個銀色的 K_4。（提示：一個點要麼至少有 4 條金色的邊，要麼至少有 6 條銀色的邊。）事實上 $R(3,4) = 9$。

4 **單調子數列。**就 $n = 2$ 的情形證明艾狄胥–塞凱賴什定理：任何由 5 個相異的數組成的數列，都包含一個長度為 3 的遞增或遞減子數列。找出一個由 4 個數組成、兩者皆無的數列。

延伸閱讀

- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），組合數學入門書，可讀其中關於鴿巢原理和拉姆齊理論的章節。
- **Ronald Graham & Joel Spencer,「Ramsey theory」**，*Scientific American*（1990 年 7 月）。由兩位頂尖研究者撰寫的科普介紹。
- **Noga Alon & Joel Spencer, *The Probabilistic Method***（Wiley）。講述艾狄胥這套方法的標準著作；第一章不難讀懂。
- **Paul Hoffman, *The Man Who Loved Only Numbers***（Fourth Estate）。艾狄胥的傳記。

9.3 為籬笆樁着色

題目

STEP II 1995，第 2 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/95-s2-q2`

引子

為排成一列的柱子髹色很容易：第一根有 3 種選擇，其後每根各有 2 種（只要不與前一根同色即可），共有 $3 \times 2^{n-1}$ 種髹法。把這一列彎成一個圓圈，問題卻變得出奇地棘手，因為最後一根柱子既要與相鄰的前一根不同色，又要與第一根不同色，而這究竟排除了一種顏色還是沒有排除任何顏色，取決於前面所有柱子的髹法。

題目用一個經典技巧繞過這個難處：按最後一根柱子的顏色把計數分成幾種情況，再找出各種情況之間如何互相轉化。答案 $2^n + 2(-1)^n$ 是*色多項式*（chromatic polynomial）的第一個例子，而色多項式正是四色定理（four colour theorem）背後的數學的一部分。

動筆之前

> **停一停，試一試** 數一數為圍成一圈的 3 根柱子髹色有多少種方法，4 根又有多少種。$3 \times 2^{n-1}$ 有沒有多算？多算了多少？

圍成一圈的三根柱子兩兩相鄰，所以需要三種不同的顏色：共 $3! = 6$ 種方法。圍成一圈的四根柱子：18 種方法（你可以在最後驗證這一點）。與排成一列時的 $3 \times 2^2 = 12$ 和 $3 \times 2^3 = 24$ 比較：圓圈少了的，正是最後一根柱子與第一根同色的那些着色。

逐步解題

階段 1　兩類着色

想一想 考慮 $n+1$ 根柱子的一種着色（colouring），其中第一根和最後一根都是紅色。把最後一根柱子拿走。餘下的首 n 根柱子的着色有甚麼特點？

第 n 根柱子與第 $n+1$ 根相鄰，而後者是紅色，所以第 n 根不是紅色。拿走第 $n+1$ 根後，剩下的是 n 根柱子的一種着色，其中第一根是紅色而最後一根不是紅色：這正是 s_n 種着色之一。反過來，s_n 種着色中的任何一種都恰好有一種方法延伸，就是把第 $n+1$ 根柱子髹成紅色。所以兩個集合完全對應，$r_{n+1} = s_n$。

想一想 $r_n + s_n$ 是甚麼？

每一種第一根柱子為紅色的着色，其最後一根柱子不是紅色就是非紅色，所以 $r_n + s_n$ 數的是 n 根柱子中第一根為紅色的全部着色。第一根之後，每根柱子有 2 種選擇，所以 $r_n + s_n = 2^{n-1}$。因此

$$r_{n+1} + r_n = s_n + r_n = 2^{n-1}. \qquad \blacksquare$$

階段 2 用數學歸納法（induction）證明公式

當 $n = 1$ 時只有一根柱子，髹成紅色，所以 $r_1 = 1$，而公式給出 $\frac{1}{3}(1+2) = 1$。假設公式對 r_n 成立。那麼

$$r_{n+1} = 2^{n-1} - r_n = 2^{n-1} - \frac{2^{n-1} + 2(-1)^{n-1}}{3} = \frac{2 \cdot 2^{n-1} - 2(-1)^{n-1}}{3} = \frac{2^n + 2(-1)^n}{3},$$

這正是 r_{n+1} 的公式。 $\blacksquare$

停一停，試一試 這條公式從何而來？遞歸關係（recurrence relation）$r_{n+1} + r_n = 2^{n-1}$ 就像一條微分方程（differential equation）：解是一個「特解」（particular solution）$c \cdot 2^n$ 加上一個滿足 $r_{n+1} + r_n = 0$ 的「餘函數」（complementary function）$A(-1)^n$。求 c 和 A。

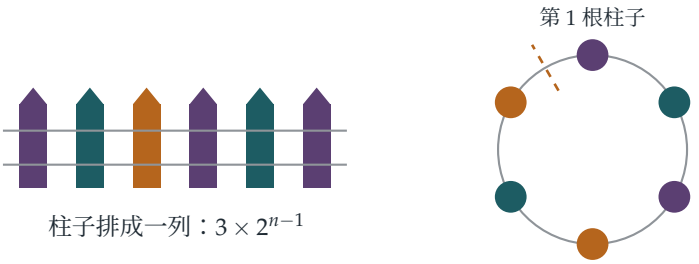
階段 3 圓圈

想一想 在 n 根柱子圍成的圓圈中，固定第 1 根柱子的顏色，例如紅色。現在在第 n 根與第 1 根柱子之間把圓圈剪開，拉直成一列。哪些直列着色對應於合乎要求的圓圈着色？

拉直之後，圓圈變成由紅色開始的一列 n 根柱子，其中最後一根還必須與第 1 根不同色：最後一根不是紅色。這些正好就是那 s_n 種着色。第 1 根柱子可以是 3 種顏色中的任何一種，所以圓圈着色的數目是

$$3s_n = 3(2^{n-1} - r_n) = 3 \cdot 2^{n-1} - 2^{n-1} - 2(-1)^{n-1} = 2^n + 2(-1)^n. \qquad \blacksquare$$

驗算：$n = 3$ 給出 $8 - 2 = 6$，而 $n = 4$ 給出 $16 + 2 = 18$。



柱子排成一列：$3 \times 2^{n-1}$

圍成一圈：$2^n + 2(-1)^n$

圖 9.3.1 左：排成一列時，第一根之後的每根柱子只須與前一根不同色。右：圍成一圈時，最後一根柱子還必須與第 1 根柱子不同色。沿虛線把圓圈剪開，圓圈着色就變成首尾兩根柱子不同色的直列着色：即題目中所數的 s_n 種着色之一。

回顧

- **分成互相轉化的情況。**記下最後一根柱子是否紅色，便得到兩個數列 r_n 和 s_n，它們由簡單的遞歸關係聯繫起來。
- **雙射（bijection）。**$r_{n+1} = s_n$ 是靠把着色一一配對來證明的，而不是靠計算。
- **剪開圓圈。**一個循環的問題變成了一個對兩端附加條件的直線問題。

背後的數學

色多項式

柱子和它們之間的相鄰關係構成一個圖（graph）：排成一列時是一條路徑（path），圍成一圈時是一個圈（cycle）。*正常着色*（proper colouring）給相鄰的頂點（vertex）不同的顏色。用 k 種顏色的正常着色數目是 k 的一個多項式，稱為該圖的*色多項式*，由伯克霍夫（George Birkhoff）於 1912 年引入。用 k 種顏色重複題目中的論證，可得

$$\text{路徑：} k(k-1)^{n-1}, \qquad \text{圈：} (k-1)^n + (-1)^n(k-1).$$

取 $k = 3$，圈的公式就是 $2^n + 2(-1)^n$。伯克霍夫希望透過證明每幅地圖所對應的圖的色多項式在 $k = 4$ 時為正數，從而證明四色定理（每幅地圖都可以用四種顏色着色，使相鄰的國家顏色不同）。這個方法始終未能成功；定理最終在 1976 年由阿佩爾（Kenneth Appel）和哈肯（Wolfgang Haken）借助大量電腦運算證明。色多項式至今仍是一個活躍的研究領域，並與統計物理（statistical physics）有出人意表的聯繫：在那裏它們以描

述磁性的*波茨模型*（Potts model）的形式出現。

用矩陣數出同一個答案

還有另一種方法可以看出答案，用的是第 7 章的線性代數（linear algebra）。設 $\mathbf{M}$ 為對角線上是 0、其餘位置是 1 的 3×3 矩陣：當顏色 j 可以跟在顏色 i 之後時，第 (i, j) 項為 1。那麼 $\mathbf{M}^{n-1}$ 的第 (i, j) 項數的是由顏色 i 開始、以顏色 j 結束的直列着色，所以 r_n 是 $\mathbf{M}^{n-1}$ 的一個對角線項。圓圈着色是經過 n 步後回到起點的一個序列，所以圓圈着色的數目就是 $\mathbf{M}^n$ 的*跡*（trace）。由於 $\mathbf{M} = \mathbf{J} - \mathbf{I}$，其中 $\mathbf{J}$ 的每一項都是 1，$\mathbf{M}$ 的特徵值（eigenvalue）是 $3 - 1 = 2$（特徵向量（eigenvector）為 $(1, 1, 1)$）和 -1（兩重），而 $\mathbf{M}^n$ 的跡是各特徵值的 n 次方之和：

$$\operatorname{tr} \mathbf{M}^n = 2^n + (-1)^n + (-1)^n = 2^n + 2(-1)^n.$$

這種*轉移矩陣*（transfer matrix）方法在組合數學（combinatorics）和統計物理中隨處可見，用來數算逐步構造出來的排列。

思維技巧

建立互相耦合的遞歸關係

當一個數列單靠自己無法滿足簡單的遞歸關係時，不妨多記錄一點資料（這裏是最後一根柱子的顏色），再找出聯繫各種情況的遞歸關係。

這就是*狀態*（state）的概念：只記下繼續下去所必須知道的資料，不多也不少。它是電腦科學中動態規劃（dynamic programming）的基礎，也是概率中馬可夫鏈（Markov chains）（Section 12.16）和轉移矩陣方法的基礎。

延伸練習

1 **k 種顏色。**用 k 種顏色重做這道題。證明 $r_{n+1} = s_n$ 仍然成立，並且 $r_{n+1} + r_n = (k-1)^{n-1}$，求 r_n，並由此推導出圈的色多項式 $(k-1)^n + (-1)^n(k-1)$。

2 **沒有兩根紅柱相鄰。**把排成一列的 n 根柱子髹成紅色或白色，而且沒有兩根相鄰的柱子同為紅色，共有多少種方法？（你應該會得到斐波那契數（Fibonacci numbers）。）

3 **輪子。**加一根位於中心的柱子，與圓圈上的每根柱子都相鄰（一個「輪子」）。用 4 種顏色為它着色有多少種方法？（提示：中心用去一種顏色，圓圈必須使用其餘三種。）

4 **項鏈。**如果圓圈的兩種着色在其中一種可由另一種旋轉而得時視為相同，那麼當 $n = 4$ 時有多少種不同的着色？$n = 5$ 時呢？（這是伯恩賽德引理（Burnside's lemma）的開端。）

延伸閱讀

- **Robin Wilson, *Introduction to Graph Theory***（Pearson）。標準的圖論入門教材；涵蓋着色和四色問題。
- **Robin Wilson, *Four Colours Suffice***（Penguin）。四色定理的故事，為一般讀者而寫。
- **Herbert Wilf, *generatingfunctionology***（可在網上免費閱讀）。講述遞歸關係及其解法，包括轉移矩陣方法。

9.4 把數列看成函數

題目

STEP I 2016，第 8 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/16-s1-q8`

引子

用計算機算一算 1/89：0.011235955…。開首的數字是 $0, 1, 1, 2, 3, 5$，正是斐波那契數（Fibonacci numbers）。這是巧合嗎？不是，而這道題目正好解釋了原因。

這個想法是組合數學（combinatorics）中最強大的想法之一。取一個無窮數列，把各項用作冪級數（power series）的係數，從而把整個數列包裝成單一個函數。這樣，關於數列的事實就變成關於函數的事實，你便可以用代數和微積分來處理。Herbert Wilf 有一句令人難忘的話：「生成函數（generating function）是一條晾衣繩，我們把一列數字掛在上面展示。」

動筆之前

想一想 哪個數列的生成函數是 $\dfrac{1}{1-x}$？$\dfrac{1}{1-2x}$ 又如何？如果把一個數列的生成函數乘以 x，數列會有甚麼變化？

等比級數（geometric series）$\frac{1}{1-x} = 1 + x + x^2 + \cdots$ 是 $1, 1, 1, \ldots$ 的生成函數，而 $\frac{1}{1-2x} = 1 + 2x + 4x^2 + \cdots$ 則是 2 的各次冪的生成函數。乘以 x 會把每一項向後移一位：$x \cdot \frac{1}{1-x} = x + x^2 + \cdots$ 對應的數列是 $0, 1, 1, 1, \ldots$。

逐步解題

階段 1　自然數與三角形數

想一想　用二項級數（binomial series）展開 $(1-x)^{-2}$。x^n 的係數是甚麼？

由二項級數可得

$$(1-x)^{-2} = 1 + 2x + 3x^2 + 4x^3 + \cdots = \sum_{n\ge 0} (n+1)x^n,$$

所以 $x(1-x)^{-2} = x + 2x^2 + 3x^3 + \cdots = \sum nx^n$：這就是 $u_n = n$ 的生成函數。

同樣地，$(1-x)^{-3}$ 中 x^n 的係數是 $\frac{(-3)(-4)\cdots(-3-n+1)}{n!}(-1)^n = \frac{(n+1)(n+2)}{2}$，所以

$$x(1-x)^{-3} = \sum_{n\ge 1} \frac{n(n+1)}{2}\, x^n = x + 3x^2 + 6x^3 + 10x^4 + \cdots,$$

這就是三角形數（triangle numbers）$u_n = \frac{1}{2}n(n+1)$ 的生成函數。

想一想　怎樣用 n 和 $\frac{1}{2}n(n+1)$ 來表示 n^2？

由於 $n^2 = 2\cdot\frac{1}{2}n(n+1) - n$，而生成函數可以像數列一樣相加，

$$\sum n^2 x^n = \frac{2x}{(1-x)^3} - \frac{x}{(1-x)^2} = \frac{2x - x(1-x)}{(1-x)^3} = \frac{x(1+x)}{(1-x)^3}. \qquad \blacksquare$$

階段 2　遞推關係變成方程

想一想　把 $u_n = ku_{n-1}$ 乘以 x^n，然後對所有 $n \ge 1$ 求和。兩邊用 f 表示，分別是甚麼？

左邊是 $\sum_{n\ge 1} u_n x^n = \mathrm{f}(x) - u_0 = \mathrm{f}(x) - a$。右邊是 $kx\sum_{n\ge 1} u_{n-1}x^{n-1} = kx\,\mathrm{f}(x)$。所以 $\mathrm{f}(x) = a + kx\,\mathrm{f}(x)$，於是

$$\mathrm{f}(x) = \frac{a}{1-kx}. \qquad \blacksquare$$

當然，$u_n = ak^n$；把 $\frac{a}{1-kx}$ 展開成等比級數，便可確認這一點。

階段 3　斐波那契數

對 $u_n = u_{n-1} + u_{n-2}$（其中 $n \ge 2$）作同樣的處理：

$$\sum_{n\ge 2} u_n x^n = x\sum_{n\ge 2} u_{n-1}x^{n-1} + x^2 \sum_{n\ge 2} u_{n-2}x^{n-2},$$

即

$$\mathrm{f}(x) - u_0 - u_1 x = x\big(\mathrm{f}(x) - u_0\big) + x^2\,\mathrm{f}(x).$$

代入 $u_0 = 0$ 和 $u_1 = 1$，上式即 $\mathrm{f}(x) - x = x\,\mathrm{f}(x) + x^2\,\mathrm{f}(x)$，所以

$$\mathrm{f}(x) = \frac{x}{1 - x - x^2}. \qquad \blacksquare$$

階段 4 **解釋** $1/89$

> **停一停，試一試** 暫且不理收斂（convergence）的問題（這裏沒有問題：級數在 $|x| < 0.618$ 時收斂），把 $x = \frac{1}{10}$ 代入 f。你得到甚麼？

$$\frac{1}{10}F_1 + \frac{1}{100}F_2 + \frac{1}{1000}F_3 + \cdots = \mathrm{f}\Big(\frac{1}{10}\Big) = \frac{1/10}{1 - 1/10 - 1/100} = \frac{10}{89}.$$

除以 10，得 $\frac{1}{89} = \sum F_n/10^{n+1}$。斐波那契數依次放在相繼的小數位上，當它們大於 9 時便會互相重疊（圖 9.4.1）。$1/89$ 開首的數字顯示出斐波那契數；其後的數字則因進位而變得模糊。

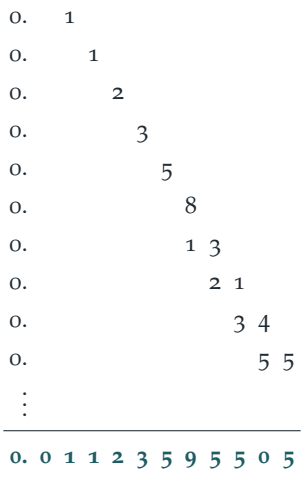


圖 9.4.1 把 $F_n/10^{n+1}$ 對 $n = 1, 2, 3, \ldots$ 相加。每個斐波那契數的末位都比前一個向右多移一個小數位。一旦它們有兩位數字，便會互相重疊，而進位會改變第七位起的數字：$\frac{1}{89} = 0.01123595505\ldots$。

回顧

- **把數列變成函數。**整個無窮數列變成了一個有理函數（rational function）。

- **遞推關係變成方程。**乘以 x^n 再求和，便把規則 $u_n = u_{n-1} + u_{n-2}$ 變成一條關於 f 的方程，可以用代數方法解出。
- **從舊的造出新的。**把生成函數相加、平移和求導，便能得出新的生成函數，無須從頭做起。

背後的數學

從生成函數到公式

生成函數不只是把數列編碼：它還可以把數列解出來。分母可以因式分解為 $1 - x - x^2 = (1 - \varphi x)(1 - \psi x)$，其中 $\varphi = \frac{1+\sqrt{5}}{2}$ 和 $\psi = \frac{1-\sqrt{5}}{2}$ 是 $t^2 = t + 1$ 的根。由部分分式（partial fractions）可得

$$\frac{x}{1 - x - x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1 - \varphi x} - \frac{1}{1 - \psi x}\right),$$

再把每個分式展開成等比級數，

$$F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}.$$

這就是*比內公式*（*Binet's formula*）。由於 $|\psi| < 1$，第二項非常小，所以 F_n 是最接近 $\varphi^n/\sqrt{5}$ 的整數，而相鄰兩個斐波那契數之比趨向黃金比（golden ratio）（這是 Section 2.2 的主題）。同一方法可以解任何常係數線性遞推關係（linear recurrence with constant coefficients）；練習中有一個來自 STEP I 2004 的例子。

生成函數為何如此有用

真正的威力在相乘時才顯現出來。乘積 $f(x)g(x)$ 中 x^n 的係數是 $\sum_k u_k v_{n-k}$，即對所有把 n 拆成兩部分的方法求和。要數算由兩個部分組成的東西，這正是我們所需要的。接下來的兩節會用它來數算找換零錢的方法（Section 9.5），以及找出點數古怪、但總和的分佈與普通骰子相同的骰子（Section 9.6）。在概率中，同一想法給出*概率生成函數*（*probability generating function*），它把獨立隨機變量之和變成乘積。

棣莫弗、歐拉與拉普拉斯

棣莫弗（Abraham de Moivre）在 1720 年代用生成函數來解線性遞推關係，所用的方法基本上就是上面的方法。歐拉（Leonhard Euler）在 1748 年用它們來研究整數的分拆（partitions），而拉普拉斯（Pierre-Simon Laplace）則以它們作為其著作 *Théorie analytique des probabilités*（1812）的基礎。到了二十世紀，生成函數成為組合數學的核心，而解析方法（把生成函數看成複變數的函數來研究）如今能為複雜得無法精確求解的計數問題給出精確的估計。

思維技巧

轉換表示方式

一個關於數列的問題，在把數列變成函數、矩陣或圖形之後，可能會變得容易。

生成函數把遞推關係翻譯成代數，把計數翻譯成乘法。在大學裏你會遇到許多這類翻譯：傅立葉級數（Fourier series）把函數變成數列，拉普拉斯變換（Laplace transforms）把微分方程變成代數，而矩陣則把幾何變成算術。

延伸練習

1 **每次兩位數字。**試做 STEP II 2022，第 3 題（載於 OCR 提供下載的 2022 年試卷，`www.ocr.org.uk/Images/696587-2022-past-papers.zip`）。它解釋了為何 $\frac{1}{89}$ 的小數展開式以斐波那契數開首，並要求找出一個分數，其小數每次兩位數字地顯示這個規律。

2 **解遞推關係。**試做 STEP I 2004，第 7 題（`step.maths.org/questions/04-s1-q7`）。它用生成函數和部分分式解一個線性遞推關係，然後把這個方法倒過來做。

3 **逐一數數。**證明 $\frac{1}{9801} = 0.00\,01\,02\,03\,04\ldots$，其中每兩位數字為一組，逐一向上數。這個規律最先在哪裏出錯？為甚麼？

4 **平方和。**解釋為何 $\frac{1}{1-x}f(x)$ 是部分和 $u_0 + u_1 + \cdots + u_n$ 的生成函數。利用這一點和第 (i) 部分的答案，求 $1^2 + 2^2 + \cdots + n^2$ 的公式。

延伸閱讀

- **Herbert Wilf, *generatingfunctionology***（A K Peters；網上免費閱讀）。經典的入門書。讀完本節後，最適合接着讀第 1 章。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***（Addison-Wesley），第 7 章。
- **Philippe Flajolet & Robert Sedgewick, *Analytic Combinatorics***（Cambridge University Press；網上免費閱讀）。遠遠超出大學一年級的程度，但開首數頁展示了這門學問的去向。

9.5 湊零錢

題目

STEP I 1997，第 1 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/97-s1-q1`

引子

這看來是全書最容易的一題：把所有可能逐一列出便是。可是列舉很快就會失控。用英國人口袋裏的硬幣，湊出 20p 有 41 種方法，湊出 50p 有 451 種，湊出 £1 更有 4563 種。這道題真正要求的是一套*系統*，也就是一種整理計數的方法，使得既無遺漏，亦無重複。

這套系統背後是一段優美的數學。每種硬幣對應一個等比級數（geometric series），湊出 n 便士的方法數就是這些級數之積的一個係數（coefficient）；把同一個想法一併用於所有正整數，便得到歐拉（Euler）的分拆（partition）理論，那是數論中最深刻的課題之一。

動筆之前

想一想 只用 2p 和 1p 硬幣，湊出 n 便士有多少種方法？為甚麼答案只取決於你用了多少枚 2p 硬幣？

一旦選定 2p 硬幣的數目，其餘的就必須全是 1p 硬幣，所以 2p 硬幣的每個可能數目 $0, 1, \ldots, \lfloor n/2 \rfloor$ 都恰好對應一種方法。這樣共有 $\lfloor n/2 \rfloor + 1$ 種方法，記作 $a(n)$。例如 $a(10) = 6$，$a(5) = 3$。

逐步解題

階段 1　十便士

想一想　按面值最大的硬幣來整理計數。你可以用多少枚 10p 硬幣？定下來之後，又可以用多少枚 5p 硬幣？

- 一枚 10p 硬幣：1 種方法。
- 不用 10p，用兩枚 5p：1 種方法。
- 不用 10p，用一枚 5p：餘下的 5p 由 2p 和 1p 湊成，有 $a(5) = 3$ 種方法。
- 不用 10p 也不用 5p：10p 全由 2p 和 1p 湊成，有 $a(10) = 6$ 種方法。

總數是 $1 + 1 + 3 + 6 = 11$。 ■

階段 2　逐層建立

想一想　設 $b(n)$ 為用 5p、2p 和 1p 硬幣湊出 n 便士的方法數。怎樣用函數 a 表示 $b(n)$？

選定 5p 硬幣的數目 k；餘下的 $n - 5k$ 便士由 2p 和 1p 硬幣湊成。所以

$$b(n) = a(n) + a(n-5) + a(n-10) + \cdots .$$

由此得

n	0	5	10	15	20
$a(n)$	1	3	6	8	11
$b(n) = a(n) + b(n-5)$	1	4	10	18	29

同樣地，設 $c(n)$ 為連 10p 硬幣也可使用時的方法數；則 $c(n) = b(n) + b(n-10) + \cdots$，所以 $c(10) = 10 + 1 = 11$（與第 1 階段的結果一致），而 $c(20) = b(20) + b(10) + b(0) = 29 + 10 + 1 = 40$。

階段 3　二十便士

最後輪到 20p 硬幣：我們要麼用它（一種方法），要麼不用（$c(20)$ 種方法）。所以共有

$$40 + 1 = 41 \text{ 種方法.} \qquad \blacksquare$$

停一停，試一試 留意當中的結構：每加入一種新硬幣就多一層，而每一層都是把上一層按新硬幣的面值為間隔取值再相加。你能看出電腦會怎樣為 £1 甚至 £100 進行這項計算嗎？

回顧

- **按最大的部分來整理。**先固定面值最大的硬幣的數目，計數便有條不紊。
- **建立表格。**每加入一種新硬幣，就多出一行，由上一行計算出來。這就是動態規劃（dynamic programming），是電腦計算這類數目的標準方法。
- **與已知的情況核對。**表格重現了 10p 的 11 種方法。

背後的數學

把硬幣看成等比級數

以下這個想法把問題化為代數。所用 5p 硬幣的數目可以是 $0, 1, 2, \ldots$，分別貢獻 $0, 5, 10, \ldots$ 便士。把這一點記錄成級數

$$1 + x^5 + x^{10} + x^{15} + \cdots = \frac{1}{1 - x^5}.$$

對每種硬幣都這樣做，然後相乘：

$$\frac{1}{(1-x)(1-x^2)(1-x^5)(1-x^{10})(1-x^{20})} = \sum_n w(n) x^n.$$

把各括號乘開時，每一項 x^n 都來自在每個括號中各選一項，也就是為每種硬幣各選一個數目，而總值為 n。所以 x^n 的係數 $w(n)$ 正好就是湊出 n 便士的方法數。第 2 階段的表格，其實就是把這個乘積逐個括號乘開。這正是 Section 9.4 中母函數（generating function）的想法，而乘法正是它大顯威力之處：母函數之積所計算的，就是各項選擇的組合數目。

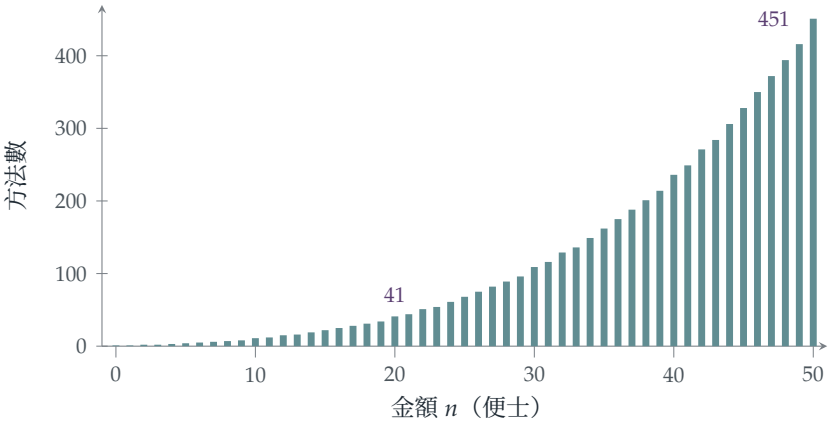


圖 9.5.1　用 1p、2p、5p、10p、20p 和 50p 硬幣湊出 n 便士的方法數。在 5 和 10 的倍數處出現的跳升，顯示新硬幣首次可用的位置。有六種硬幣時，方法數的增長大致與 n^5 相若：一般而言，增長有如一個多項式，其次數比硬幣種數少一。

分拆

假如每個正整數都可以作為硬幣的面值，又會怎樣？這時，不計次序而把 n 寫成正整數之和的方法數，就是**分拆數**（partition number）$p(n)$。例如 $p(4) = 5$，因為 $4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1$。歐拉在 1748 年看出

$$\sum_{n\geq 0} p(n)x^n = \frac{1}{(1-x)(1-x^2)(1-x^3)(1-x^4)\cdots},$$

並用它發現了一些非凡的恆等式（identity）。其中之一是：把 n 寫成*互不相同*的部分之和的方法數，等於把它寫成*奇數*部分之和的方法數。（用母函數來證明，只需一行代數：請在練習中一試。）

哈代、拉馬努金與 $p(200)$

分拆數增長得很快：$p(10) = 42$，$p(100) = 190\,569\,292$。1918 年，哈代（G. H. Hardy）與拉馬努金（Srinivasa Ramanujan）為 $p(n)$ 找到一條令人驚歎的公式，那是一個無窮級數（infinite series），其首項為

$$p(n) \approx \frac{1}{4n\sqrt{3}}\, e^{\pi\sqrt{2n/3}}.$$

為了檢驗這條公式，Percy MacMahon 利用歐拉的恆等式，徒手算出 $p(200) = 3\,972\,999\,029\,388$。哈代與拉馬努金的級數只取首幾項，所得結果與正確答案相差不足一。拉馬努金還發現 $p(5n+4)$ 必定可被 5 整除，$p(7n+5)$ 必定可被 7 整除，而 $p(11n+6)$ 必定可被 11 整除。至於為何再沒有其他這類規律，則要到二十一世紀才得

到解釋。

思維技巧

藉係數來計數

當一個對象由若干獨立的選擇構成時，每項選擇就是一個因式，而總大小爲 n 的對象的數目，就是乘積中 x^n 的係數。

這把計數化為代數。在大學裏，它是計數組合學（enumerative combinatorics）的基礎，而概率母函數（probability generating function）也正是這樣處理獨立隨機變量（independent random variables）之和的。

延伸練習

1 **兩種六。**試做 STEP II 2012，第 1 題（`step.maths.org/questions/12-s2-q1`）。該題利用等比級數之積來計數。做完之後，請描述一個以這些係數為答案的硬幣計數問題。

2 **互異與奇數。**證明 $(1+x)(1+x^2)(1+x^3)\cdots$ 是分拆成互不相同部分的母函數。利用 $1+x^k = \frac{1-x^{2k}}{1-x^k}$ 證明它等於 $\frac{1}{(1-x)(1-x^3)(1-x^5)\cdots}$，並由此推出歐拉定理。就 $n=6$ 驗證這個結果。

3 **一條公式。**證明只用 1p、2p 和 5p 硬幣湊出 n 便士的方法數，是最接近 $\frac{(n+4)^2}{20}$ 的整數。（先用較小的 n 驗證。）

4 **次序有別。**把 1p 和 2p 硬幣逐枚投入投幣口來湊出 n 便士，若次序有別，共有多少種方法？（這個數列（sequence）你以前見過。）

延伸閱讀

- **George Andrews & Kimmo Eriksson, *Integer Partitions***（Cambridge University Press）。一本為大學本科生而寫的入門書，篇幅短小，淺白易讀。
- **波利亞（George Pólya），「On picture-writing」**，*American Mathematical Monthly* 63 (1956)。波利亞以兌換一美元的問題（共有 292 種方法）來講解母函數的經典文章。
- **Robert Kanigel, *The Man Who Knew Infinity***（Abacus）。拉馬努金的生平，以及他與哈

代的合作。

9.6 奇怪的骰子

題目

STEP I 1993，第 1 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/93-s1-q1`

引子

兩顆六面骰子，各面標上正整數，除了平常的標法之外，還有沒有別的標法，可以令由 2 到 12 的每個點數和出現的次數，都與普通骰子完全相同？1978 年，Martin Gardner 在《科學人》（*Scientific American*）雜誌上宣布，這樣的標法恰好只有一種：一顆骰子標上 $1,2,2,3,3,4$，另一顆標上 $1,3,4,5,6,8$。這對骰子由 George Sicherman 上校發現，現在市面上有售，稱為 *Sicherman* 骰子。用它們來玩一局大富翁（Monopoly），與用普通骰子玩根本無從分辨。

本題是一道熱身題：從點數和的列表，推斷兩顆骰子各面標了甚麼數字。要有系統地解決這個問題，以至證明 Sicherman 骰子是唯一的一對奇異骰子，方法都是把每顆骰子化為一個多項式（polynomial）。

動筆之前

想一想 假設第一顆骰子有 p_i 個面顯示 i，第二顆有 q_j 個面顯示 j，那麼擲出點數和 t 共有多少種方法？

方法的數目是 $\sum_i p_i q_{t-i}$：先選定第一顆骰子的點數 i，第二顆就必須顯示 $t-i$。就題目中的第一對骰子而言，點數和 6 可以來自 $1+5$（有 1×1 種方法）、$2+4$（2×1 種）、$3+3$（2×2 種），此外再無其他可能，因為第一顆骰子沒有 4 或 5，而第二顆沒有 0。所以共有 $1+2+4=7$ 種方法。 ■

逐步解題

階段 1　兩端的情況

把兩顆骰子稱為 A 和 B，並以 a_i、b_i 分別表示兩顆骰子上顯示 i 的面數。

> **想一想**　點數和不可能是 2，也不可能是 12。這對 1 和 6 的分佈有甚麼啟示？擲出 3 卻確實有兩種方法。哪一顆骰子上有 1？

擲不出點數和 2，即 $a_1 b_1 = 0$：最多只有一顆骰子上有 1。點數和 3 只可能是 $1+2$ 或 $2+1$，而且有兩種方法，所以其中一顆骰子確實有 1；把它稱為 A，於是 $b_1 = 0$。那麼

$$\text{點數和3：}\quad a_1 b_2 = 2, \qquad \text{所以} (a_1, b_2) = (2,1) \text{ 或} (1,2).$$

同理，擲不出點數和 12，即 $a_6 b_6 = 0$。

階段 2　逐步向上推

> **想一想**　寫出點數和為 4 的方程，記住 $b_1 = 0$。在兩種情況下，這個方程分別說明了甚麼？

點數和 4 是 $1+3$ 或 $2+2$（由於 $b_1 = 0$，$3+1$ 不可能出現），所以 $a_1 b_3 + a_2 b_2 = 1$。

- **情況 (a)**，$a_1 = 2$，$b_2 = 1$：$2b_3 + a_2 = 1$，所以 $b_3 = 0$ 及 $a_2 = 1$。
- **情況 (b)**，$a_1 = 1$，$b_2 = 2$：$b_3 + 2a_2 = 1$，所以 $a_2 = 0$ 及 $b_3 = 1$。

點數和 5：$a_1 b_4 + a_2 b_3 + a_3 b_2 = 1$。

- 情況 (a)：$2b_4 + 0 + a_3 = 1$，所以 $b_4 = 0$，$a_3 = 1$。至此 A 有 $1,1,2,3$，B 有 2；B 其餘五個面都是 5 或 6。
- 情況 (b)：$b_4 + 0 + 2a_3 = 1$，所以 $a_3 = 0$，$b_4 = 1$。至此 A 有 1，B 有 $2,2,3,4$；A 其餘五個面取自 $4,5,6$，B 其餘兩個面取自 $5,6$。

階段 3　情況 (a) 不成立

> **想一想**　在情況 (a) 中，寫出點數和為 7 的方程，並利用 $b_5 + b_6 = 5$。

點數和 7：$a_1 b_6 + a_2 b_5 + a_3 b_4 + a_4 b_3 + a_5 b_2 = 2b_6 + b_5 + 0 + 0 + a_5 = 3$。代入 $b_6 = 5 - b_5$，得 $10 - b_5 + a_5 = 3$，所以 $b_5 = 7 + a_5 \geq 7$。但 B 只有六個面。情況 (a) 不可能成立。

階段 4 情況 (b) 成立

想一想 在情況 (b) 中，利用點數和 6 和 11，以及 $a_6b_6 = 0$ 這個事實。

點數和 6：$a_1b_5 + a_4b_2 = b_5 + 2a_4 = 4$（其餘各個乘積都含有 $a_2 = a_3 = 0$ 或 $b_1 = 0$）。由於 $b_5 \le 2$，而且 $b_5 = 4 - 2a_4$ 是偶數，所以要麼 $b_5 = 0, a_4 = 2$，要麼 $b_5 = 2, a_4 = 1$。

- 若 $b_5 = 0$，則 B 是 $2,2,3,4,6,6$，於是 $a_6 = 0$，而 A 是 $1,4,4,5,5,5$。但這樣點數和 7 就有 $a_1b_6 + a_4b_3 + a_5b_2 = 2 + 2 + 6 = 10$ 種方法，而不是 3 種。
- 若 $b_5 = 2$，則 B 是 $2,2,3,4,5,5$，其中 $b_6 = 0$，而 A 是 $1,4$ 再加上四個取自 5 和 6 的面。點數和 11：$a_6b_5 = 2a_6 = 6$，所以 $a_6 = 3$ 及 $a_5 = 1$。

唯一的可能是

$$A:\ 1,4,5,6,6,6, \qquad B:\ 2,2,3,4,5,5.$$

最後，我們必須驗證這對骰子確實給出整個列表，因為我們只用了其中一部分點數和。把乘積展開（或利用下文的多項式），即可確認每一項都吻合。 ■

回顧

- **從兩端入手。**最小和最大的點數和只牽涉一兩個乘積，所以提供的資訊最有力。
- **分情況討論，逐一排除。**每個點數和都把可能性收窄一些，直到只剩下一種情況。
- **檢查充分性（sufficiency）。**推論說明了骰子*必須*是甚麼樣子；最後的驗證則說明這對骰子確實可行。

背後的數學

把骰子看成多項式

把每顆骰子寫成一個多項式：每個顯示 i 的面對應一項 x^i。普通骰子是 $x + x^2 + x^3 + x^4 + x^5 + x^6$，而題目中的兩顆骰子是

$$A(x) = x + x^4 + x^5 + 3x^6, \qquad B(x) = 2x^2 + x^3 + x^4 + 2x^5.$$

相乘時，每一對面貢獻 $x^i x^j = x^{i+j}$，所以 $A(x)B(x)$ 中 x^t 的係數（coefficient），就是擲出點數和 t 的方法數目。這又是 Section 9.4 中母函數（generating function）的想法：點數和的列表其實就是乘積 $A(x)B(x)$，而我們逐個情況的推論，正是把它因式分解（factorise）的一種方法。

Sicherman 骰子

兩顆普通骰子的乘積是 $(x+x^2+\cdots+x^6)^2$。任何一對點數和分佈相同的骰子，其多項式 $P(x)$、$Q(x)$ 必須滿足 $P(x)Q(x)=(x+\cdots+x^6)^2$。現在作因式分解：

$$x+x^2+\cdots+x^6=x(1+x)(1+x+x^2)(1-x+x^2),$$

而這些因式不能再進一步分解（成為整數係數的多項式）。所以 P 和 Q 都必須由因式 x、$1+x$、$1+x+x^2$、$1-x+x^2$ 組成，而每個因式合共出現兩次。以下三個條件把選擇收窄：

- 每顆骰子有 6 個面，所以 $P(1)=Q(1)=6$；各因式在 $x=1$ 時的值分別是 1、2、3、1，所以每顆骰子各得一個 $1+x$ 和一個 $1+x+x^2$；
- 沒有任何面顯示 0，所以 $P(0)=Q(0)=0$：每顆骰子各得一個因式 x；
- 餘下唯一的自由，是兩個 $1-x+x^2$ 如何分配。每顆各得一個，就是普通骰子。兩個都歸同一顆骰子，則得

$$x(1+x)(1+x+x^2)=x+2x^2+2x^3+x^4,$$
$$x(1+x)(1+x+x^2)(1-x+x^2)^2=x+x^3+x^4+x^5+x^6+x^8.$$

這就是 Sicherman 骰子，1,2,2,3,3,4 和 1,3,4,5,6,8（圖 9.6.1），而以上論證亦證明了它們是唯一的另一種可能。關鍵在於：多項式與整數一樣，可以唯一地分解為不可約（irreducible）因式的乘積——數論與組合數學（combinatorics）在此相遇。

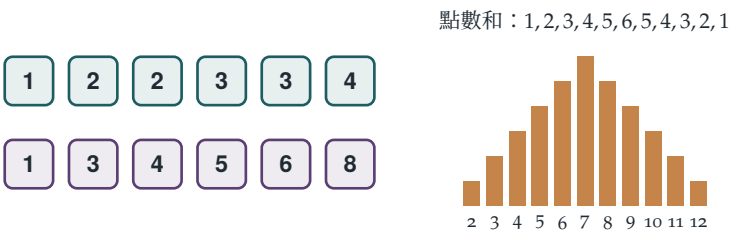


圖 9.6.1　Sicherman 骰子。兩顆一起擲，由 2 到 12 的每個點數和出現的頻數，都與兩顆普通骰子完全相同（右）：擲出 2 有一種方法，擲出 3 有兩種方法，如此類推，直至擲出 7 有六種方法。

概率母函數

在概率論中，同一個想法以*概率母函數*（probability generating function）的形式出現：取值 $0,1,2,\ldots$ 的隨機變量（random variable）X 有 $G(t)=E(t^X)=\sum P(X=k)t^k$，而對於獨立（independent）的 X 和 Y，$E(t^{X+Y})=E(t^X)E(t^Y)$。獨立隨機變量之和，對應於母函數之積。STEP III 2003 Q14（見練習）就是用這個方法找出 Sicherman 骰子，

而 Section 9.13 則用一個相關的論證，說明無論怎樣在兩顆骰子上做手腳，都不能令所有點數和出現的機會均等。

思維技巧

把計數問題編碼成多項式

當物件以數值相加的方式組合時，把每個物件表示成 x 的多項式。組合變成相乘，而有關組合的問題就變成有關因式分解的問題。

在這裏，一個關於骰子的問題化成了多項式的唯一分解（unique factorisation）。在大學裏，這類編碼隨處可見：在組合數學中，在概率論中，也在代數中——在代數裏，多項式本身就是研究對象，它們構成一個環（ring），其算術與整數的算術相似。

延伸練習

1 **四面體骰子。**試做 STEP III 2003，第 14 題（`step.maths.org/questions/03-s3-q14`），該題就四面骰子提出 Sicherman 的問題。（提示：$x + x^2 + x^3 + x^4 = x(1+x)(1+x^2)$。）

2 **驗證列表。**展開 $(x + x^4 + x^5 + 3x^6)(2x^2 + x^3 + x^4 + 2x^5)$，並與題目中的列表對照。

3 **三顆骰子。**是否存在三顆六面骰子，各面標上正整數，不全是標準骰子，而其點數和的分佈與三顆標準骰子相同？（提示：$(x + \cdots + x^6)^3$ 的因式現在包括三個 $1 - x + x^2$。）

4 **一枚硬幣與一個八面體。**一顆標上 $1, \ldots, 8$ 的八面骰子，和一顆標上 $1, 2$ 的兩面「骰子」（即一枚硬幣）：能否用另一種方式重新標號，而點數和的分佈保持不變？

延伸閱讀

- **Martin Gardner, *Penrose Tiles to Trapdoor Ciphers*** (Mathematical Association of America)。收錄了關於 Sicherman 骰子的那篇專欄文章。
- **Joseph Gallian & David Rusin, "Cyclotomic polynomials and nonstandard dice"**, *Discrete Mathematics* 27 (1979)。一般理論，適用於任意面數的骰子。

► **Herbert Wilf, *generatingfunctionology***（可於網上免費閱覽），第 1 章。

9.7 相等的冪和

題目

STEP I 1992，第 5 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/92-s1-q5`

引子

把 1 至 16 這些數分成兩隊，每隊八個：

$$\{1, 4, 6, 7, 10, 11, 13, 16\} \quad 和 \quad \{2, 3, 5, 8, 9, 12, 14, 15\}.$$

兩隊的和都是 68。兩隊的平方和都是 748。兩隊的立方和都是 9248。很難相信這不是純粹的巧合，但它其實是一個無窮無盡的規律的第三步：只要用正確的方法把 1 至 32 分開，連四次方之和也會相等。

這個分法由 Eugène Prouhet 於 1851 年發現，而它背後的規則，即 *Thue–Morse* 數列（Thue–Morse sequence），出現的地方多得驚人：數論、計算理論、國際象棋，以至最公平的輪流方式。本題透過一個多項式（polynomial）及其導數（derivative）把它發掘出來。

動筆之前

想一想 若多項式 $\mathrm{p}(x) = \sum c_k x^k$ 滿足 $\mathrm{p}(1) = 0$，這對各係數（coefficient）有甚麼含意？若 $\mathrm{p}'(1) = 0$ 又如何？

$\mathrm{p}(1) = \sum c_k$，所以 $\mathrm{p}(1) = 0$ 表示各係數之和為零。而 $\mathrm{p}'(x) = \sum k c_k x^{k-1}$，所以 $\mathrm{p}'(1) = \sum k c_k$。若所有係數都是 ± 1，這兩個條件就是說：$+1$ 與 -1 所在的位置數目相同，而且位置之和也相等。

逐步解題

階段 1　一個三重根

想一想　把 $1-x^2$ 和 $1-x^4$ 因式分解。當中共出現多少個因式 $1-x$？

由於 $1-x^2=(1-x)(1+x)$ 及 $1-x^4=(1-x)(1+x)(1+x^2)$，

$$p_0(x)=(1-x)^3(1+x)^2(1+x^2),$$

所以 $(1-x)^3$ 是一個因式（factor）。 ■

階段 2　在 1 處的導數

想一想　若 $(1-x)^3$ 整除 p_0，那麼 $(1-x)$ 的多少次方整除 p_0'？接着 $p_1=xp_0'$ 和 p_1' 又如何？

寫成 $p_0(x)=(1-x)^3q(x)$。根據積法則（product rule），$p_0'(x)=-3(1-x)^2q(x)+(1-x)^3q'(x)$，它有因式 $(1-x)^2$。所以 $p_0'(1)=0$。於是 $p_1(x)=xp_0'(x)$ 同樣有因式 $(1-x)^2$，而用相同的論證可知 p_1' 有因式 $(1-x)$，所以 $p_1'(1)=0$。 ■

階段 3　讀出各係數

展開可得 $(1-x)(1-x^2)=1-x-x^2+x^3$，然後

$$p_0(x)=1-x-x^2+x^3-x^4+x^5+x^6-x^7.$$

想一想　以係數 c_k 表示，$p_0'(1)$ 和 $p_1'(1)$ 分別是甚麼？

由於 $p_1(x)=xp_0'(x)=\sum kc_kx^k$，我們有 $p_1'(1)=\sum k^2c_k$。係數在 $k=0,3,5,6$ 時為 $+1$，在 $k=1,2,4,7$ 時為 -1。所以 $p_0'(1)=0$ 和 $p_1'(1)=0$ 即是說

$$0+3+5+6=1+2+4+7\quad(=14),$$
$$0^2+3^2+5^2+6^2=1^2+2^2+4^2+7^2\quad(=70),$$

略去 0 之後，這正是所需的恆等式。 ■

階段 4　再多一個因式

> **停一停，試一試** 要令立方和也平衡，我們需要第四個因式 $(1-x)$。應該用哪一個多項式取代 p_0？

取 $\mathrm{P}(x)=(1-x)(1-x^2)(1-x^4)(1-x^8)$，它有因式 $(1-x)^4$。它的係數來自 $\mathrm{p}_0(x)-x^8\mathrm{p}_0(x)$：先是 0 至 7 的正負號規律，然後在 8 至 15 重複同一規律，但正負號全部相反：

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
c_k	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+

把第 2 階段的論證再用一次，可知對於 $r=0,1,2,3$ 都有 $\sum c_k k^r=0$。取 $+$ 號的位置 $\{0,3,5,6,9,10,12,15\}$ 和 $-$ 號的位置 $\{1,2,4,7,8,11,13,14\}$。當 $r=1,2,3$ 時，來自 0 的一項為零，所以可以略去。因此以下的次序

$$a_1,\dots,a_8=1,2,4,7,8,11,13,14,\qquad a_9,\dots,a_{15}=3,5,6,9,10,12,15$$

符合要求：兩半的和都是 60，平方和都是 620，立方和都是 7200。 ■

回顧

- **導數能偵測重根（repeated root）。**因式 $(1-x)^m$ 會迫使首 $m-1$ 階導數在 $x=1$ 處全部為零。
- **先乘以 x 再求導**把 $\sum kc_k$ 變成 $\sum k^2c_k$，如此類推便得到更高次方。
- **倍增。**每多一個因式 $(1-x^{2^m})$，長度便增加一倍，同時多一個次方達到平衡。

背後的數學

Thue–Morse 數列

$(1-x)(1-x^2)(1-x^4)(1-x^8)\cdots$ 各係數的正負號構成 *Thue–Morse* 數列。以 0 代表 $+$、1 代表 $-$，寫出來就是：

$$0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\ \dots$$

它有三種等價的描述。

- **倍增：**由 0 開始，不斷把手上已有的部分的補（complement，即把 0 與 1 互換）接在後面：$0\to 01\to 0110\to 01101001\to\cdots$。
- **二進制數字：**把 k 寫成二進制（binary）時，若當中 1 的個數是偶數，第 k 項便是 0；若是奇數，便是 1。（這就是 x^k 的正負號為 $(-1)^{\text{1 的個數，取自}k}$ 的原因：要從乘積中得出 x^k，就要選取那些在二進制中組成 k 的因子 x^{2^m}，而每一個都帶來一個負號。）

► **多項式：**$\prod_m (1 - x^{2^m})$ 的各係數，一如題目所示。

按 Thue–Morse 數列把 $0, 1, \ldots, 2^N - 1$ 分開，所得的兩個集合對於每個 $r < N$ 都有相等的 r 次方之和（圖 9.7.1）。這就是 *Prouhet–Tarry–Escott 問題*（Prouhet–Tarry–Escott problem）：找出兩個不同的整數集合，使它們的 r 次方之和在 $r = 1, \ldots, k$ 時全部相等。Prouhet 的解用了 2^{k+1} 個數；究竟最少需要多少個數，是一個著名的未解決問題。

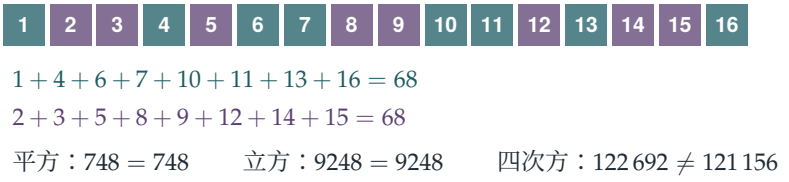


圖 9.7.1　按 Thue–Morse 數列分開的 1 至 16。兩個集合的和相等、平方和相等、立方和也相等，但四次方之和並不相等；要做到這一點，便需要 1 至 32。由左至右讀出顏色，得到的規律是藍綠–紫–紫–藍綠、紫–藍綠–藍綠–紫、...。

最公平的輪流方式

兩個小孩輪流從一堆糖果中挑選，最好的先被選走。若他們以 $ABABAB\ldots$ 的次序交替，那麼每一對糖果中較好的一顆總是歸 A，這是一種系統性的優勢。$ABBA$ 的次序較為公平，而 $ABBABAAB$ 更加公平：這正是 Thue–Morse 數列，而次方和平衡恰恰就是箇中原因。若糖果的價值由其位置的一個光滑函數給出，把該函數展開成多項式，便可看到 Thue–Morse 次序會把它的首幾項抵銷。同樣的想法曾被提議用於足球的互射十二碼（足球規則制定者自 2017 年起試行的「$ABBA$」次序），而網球決勝局的發球次序亦採用這個方式。

Thue、Morse 與 Euwe

這個數列曾多次被人發現。Prouhet 在 1851 年用過它。Axel Thue 在 1912 年研究沒有重複規律的數列時重新發現了它：它永遠不會連續三次出現同一段（即它是「無立方」(cube-free) 的）。Marston Morse 在 1921 年研究曲面上的測地線 (geodesics) 時再次發現它。而國際象棋世界冠軍尤偉（Max Euwe）在 1929 年用它證明了，當時生效的重複局面和棋規則並不能防止對局無限延續下去。時至今日，它是*自動數列*（automatic sequences）理論中的核心例子；自動數列是指由簡單機器讀取 n 的二進制數字而計算出來的數列。

思維技巧

把組合性質編碼成多項式

次方和是否平衡，關鍵在於 $+$ 號和 $-$ 號落在哪些位置。以這些正負號作爲係數的多項式，可把這個性質轉化爲一個關於根的命題。

本題從頭到尾都沒有要求你去搜尋：$x = 1$ 處的一個重根已經包辦了所有工作。利用代數來構造組合（combinatorial）對象，而不是靠反覆試驗去找，是大學組合數學中一再出現的想法。

延伸練習

1 **平移。**試做 STEP II 2025，第 8 題（載於 OCR 提供下載的 2025 年試卷，`www.ocr.org.uk/Images/750486-2025-past-papers.zip`）。該題說明怎樣把一個令和與平方和平衡的分法，透過平移其中一列再合併，變成一個更大而且連立方和也平衡的分法。

2 **平方的平方。**（同樣出自 STEP II, 2025。）集合 $A = \{1,3,4,5,9,11\}$ 和 $B = \{2,6,7,8,10\}$ 的和相等，平方和也相等。對於每個正整數 n，把 $\{n^2,(n+1)^2,\ldots,(n+11)^2\}$ 分成兩個和相等的集合。（提示：把 0 加入第二個集合。）

3 **平移不變性(translation invariance)。**證明：若兩個數集的 r 次方之和在 $r = 0,1,\ldots,k$ 時全部相等，那麼把每個元素都加上同一個數 c 所得的兩個集合亦然。為甚麼這表示題目中的集合 $\{1,\ldots,15\}$ 同樣可以換成任意 16 個連續整數？

4 **公平分糖果。**糖果的價值為 $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{8}$。比較 A 在 $ABABABAB$ 和 $ABBABAAB$ 兩種次序下所得的總值。

延伸閱讀

- **Jean-Paul Allouche & Jeffrey Shallit, “The ubiquitous Prouhet–Thue–Morse sequence”**，載於 *Sequences and Their Applications*（Springer, 1999）。綜述這個數列在各處的出現；可在網上免費閱讀。
- **Matt Parker, *Things to Make and Do in the Fourth Dimension***（Penguin）。內容包括等次方和及其他數字趣聞。
- **Joshua Cooper & Aaron Dutle, “Greedy Galois games”**，*American Mathematical Monthly*

120 (2013)。公平輪流與 Thue–Morse 數列。

9.8 沒有人拿到自己的恤衫

題目

STEP II 2016，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/16-s2-q12`

引子

一支十一人的足球隊匆忙更衣，各人隨手抓起一件球衣就穿。沒有任何人穿上自己球衣的機會有多大？再設想那是一支五十人的大軍，甚至是一百萬人的人羣。人多了，出錯的人也多了，機會是否隨之上升？還是因為有更多人可能碰巧拿對，機會反而下降？出人意料的是，這個機會幾乎完全不變：由五個人開始，它與 $1/e \approx 0.368$ 的差距不超過半個百分點。

這是組合概率（combinatorial probability）中最古老的問題，由 Pierre de Montmort 於 1708 年提出，稱為 *problème des rencontres*（相遇問題）。本題以容斥原理（inclusion–exclusion principle）解決這個問題；容斥原理是點算互相重疊的事物的工具，而所得的答案，是 e 在數學中最令人意想不到的現身之一。

動筆之前

停一停，試一試 先試 $n = 3$：列出卡 $1, 2, 3$ 的全部六種次序，數一數其中有多少種是沒有任何一張卡在自己位置上的。如果你有耐性，再試 $n = 4$（共有 24 種次序）。

當 $n = 3$ 時，只有 231 和 312 沒有任何卡在自己的位置上，所以概率是 $\frac{2}{6} = \frac{1}{3}$。當 $n = 4$ 時，24 種次序中有 9 種符合條件，概率為 $\frac{3}{8} = 0.375$。沒有任何一張卡留在原位的排列，稱為錯排（derangement）。

逐步解題

階段 1　三個事件的容斥原理

想一想　怎樣把 $A \cup B \cup C$ 寫成兩個事件的併集（union），使兩個事件的公式可以派上用場？

把 $A \cup B$ 看成單一事件，對它和 C 應用公式：

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A \cup B) + \mathrm{P}(C) - \mathrm{P}((A \cup B) \cap C).$$

現在 $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$，再應用一次公式，得 $\mathrm{P}((A \cap C) \cup (B \cap C)) = \mathrm{P}(A \cap C) + \mathrm{P}(B \cap C) - \mathrm{P}(A \cap B \cap C)$。把 $\mathrm{P}(A \cup B)$ 也展開，再合併各項，便得出結果。 ■

四個事件時，規律依舊：把四個單一事件的概率相加，減去六個兩兩交集（intersection）的概率，加上四個三重交集的概率，再減去 $\mathrm{P}(A \cap B \cap C \cap D)$。

階段 2　各個概率

想一想　如果卡 i 在位置 i，其餘的卡有多少種排法？如果卡 i 和卡 j 都固定了呢？

共有 $n!$ 種機會均等的次序。如果卡 i 固定在位置 i，其餘 $n-1$ 張卡可以任意排列：共 $(n-1)!$ 種排法。所以

$$\mathrm{P}(E_i) = \frac{(n-1)!}{n!} = \frac{1}{n}, \qquad \mathrm{P}(E_i \cap E_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},$$
$$\mathrm{P}(E_i \cap E_j \cap E_k) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}.$$

一般而言，指定的 k 張卡全部在自己位置上的概率是 $\frac{(n-k)!}{n!}$。

階段 3　加起來

想一想　在 $\mathrm{P}(E_1 \cup \cdots \cup E_n)$ 的容斥公式中，涉及 k 個事件交集的項有多少個？它們的總和是多少？

從這些事件中選出 k 個，有 $\binom{n}{k}$ 種方法，而每個交集的概率都是 $\frac{(n-k)!}{n!}$。所以各個 k 重項的總和是

$$\binom{n}{k}\frac{(n-k)!}{n!} = \frac{n!}{k!\,(n-k)!} \cdot \frac{(n-k)!}{n!} = \frac{1}{k!}.$$

除了 $k!$ 之外，全部約去。配上正負相間的符號，

$$\mathrm{P}(\text{最少有一張卡在自己的位置上}) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n+1}\frac{1}{n!}. \qquad \blacksquare$$

所以出現錯排的概率是

$$d_n = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

階段 4 恰好一張

想一想 要恰好有一張卡在原位，是哪一張卡？其餘的卡又必須怎樣？

先選定哪一張卡在自己的位置上（n 種選法）。卡 i 在原位的概率是 $\frac{1}{n}$；在這個條件下，其餘 $n-1$ 張卡必須在它們之間形成一個錯排，其概率為 d_{n-1}。所以

$$\mathrm{P}(\text{恰好一張}) = n \cdot \frac{1}{n} \cdot d_{n-1} = \sum_{k=0}^{n-1} \frac{(-1)^k}{k!}. \qquad \blacksquare$$

回顧

- **先多算，再修正。**把各個 $\mathrm{P}(E_i)$ 相加，會把有兩張卡固定的排列算了兩次；減去兩兩交集的項可以修正，但對三張卡固定的情況又減得太多；如此類推。
- **對稱性令各項相等。**每個 k 重交集的概率都相同，所以公式的每一層都是一個數乘以一個二項式係數（binomial coefficient）。
- **化為較小的情況。**「恰好一張」變成了 $n-1$ 張卡的錯排。

背後的數學

e 從何而來

級數 $\sum(-1)^k/k!$ 就是指數級數（exponential series）$e^x = \sum x^k/k!$ 在 $x = -1$ 時的值。所以當 $n \to \infty$ 時，

$$d_n \to e^{-1} \approx 0.3679,$$

而恰好有一張卡對上的概率也趨向同一個極限。收斂（convergence）的速度快得驚人（圖 9.8.1）：在 $1/n!$ 一項停下來所產生的誤差小於 $1/(n+1)!$，所以在 $n = 7$ 時，答案已經與 $1/e$ 吻合至小數點後四位。事實上，n 件物件的錯排數目 $n!\,d_n$ 恰好是最接近 $n!/e$ 的整數。

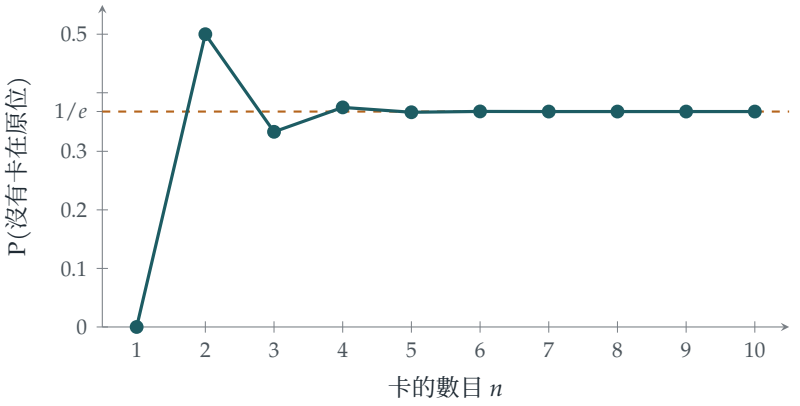


圖 9.8.1　把 n 張卡隨機洗勻後，沒有任何一張卡留在自己位置上的概率。它在 $1/e$ 的上下交替出現，而由 $n = 6$ 開始，在這個比例尺下已經與 $1/e$ 無法分辨。

對上的數目服從泊松分佈

結論還不止於此。設 N 為在自己位置上的卡的數目。那麼對每個 k，都有 $\mathrm{P}(N = k) \to \frac{e^{-1}}{k!}$：在極限之下，對上的數目服從平均值為 1 的泊松分佈（Poisson distribution）。平均值為甚麼是 1，不用任何計算也看得出來：N 是 $E_1, \ldots, E_n$ 的指示變量（indicator）之和，每個的期望值（expectation）都是 $\frac{1}{n}$，而即使這些事件並非獨立，期望值仍然可以相加。STEP III 2004 Q12（見練習）證明方差（variance）同樣是 1。每當有許多機會很微、又近乎獨立的事件時，泊松分佈就會出現；錯排是一個經典例子，也是第 12 章 Section 12.11 的主題。

Montmort 的 *treize* 遊戲

Pierre Rémond de Montmort 是一位法國貴族，他在著作 *Essay d'analyse sur les jeux de hazard*（1708）中研究各種博彩遊戲。在 *treize*（十三）這個遊戲中，玩家派出十三張牌，同時口中數着「一、二、三……」，只要有一次派出的牌與所數的數字相同，便算勝出。Montmort 求出了勝出的概率，並在與尼古拉·伯努利（Nicolaus Bernoulli）的通信中，於 1713 年或之前確立了一般公式。歐拉（Euler）在 1753 年重新發現了這條公式。容斥原理其後由棣莫弗（Abraham de Moivre）加以形式化，而它的一般形式則由西爾維斯特（James Joseph Sylvester）在十九世紀建立。

思維技巧

容斥原理

要點算互相重疊的集合的併集，先把各集合的大小相加，減去兩兩重疊的部分，加上三重重疊的部分，如此類推。

當對稱性令同一大小的重疊部分全部相等時，這個方法最見威力，本題正是如此。在大學裏，它會在組合數學（combinatorics）、概率論和數論（number theory）中出現；在數論中，它用來點算不能被一系列質數中任何一個整除的整數，並由此引出歐拉函數（Euler's totient function）和默比烏斯函數（Möbius function）。

延伸練習

1 **球衣。**試做 STEP III 2004，第 12 題（`step.maths.org/questions/04-s3-q12`）。這道題用指示變量點算穿上自己球衣的球員人數，求出平均值和方差，並把結果與泊松分佈比較。

2 **遞推關係。**設 D_n 為 n 件物件的錯排數目。考慮卡 1 去了哪個位置，以及原本在該位置的卡是否去了位置 1，證明 $D_n = (n-1)(D_{n-1} + D_{n-2})$。用 $D_3 = 2$、$D_4 = 9$ 驗算。

3 **恰好 k 張。**證明恰好有 k 張卡在自己位置上的概率是 $\frac{1}{k!}d_{n-k}$，並且它趨向 $\frac{e^{-1}}{k!}$。

4 **互質的整數。**用容斥原理點算由 1 至 30 的整數中，不能被 2、3 和 5 任何一個整除的有多少個。你的答案與 $30\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)$ 有甚麼關係？

延伸閱讀

- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。大學一年級教科書，涵蓋容斥原理和配對問題。
- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 IV 章。經典之作，載有配對問題的許多變化。
- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），關於容斥原理的部分。

9.9 生日問題

題目

STEP II 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-s2-q13`

引子

一個房間裏要有多少人，才會令其中兩人同一天生日的機會大於一半？大多數人會猜 180 左右，即 365 的一半。答案卻是 23。在一班三十人的課堂裏，有人同一天生日的概率（probability）是 70%；在五十人的派對上，則是 97%。

這個結果令人意外，是因為我們把兩個不同的問題混為一談：「有沒有人與我同一天生日？」（這確實需要大約 253 人）和「有沒有任何兩個人同一天生日？」。23 個人可以組成 $\binom{23}{2} = 253$ 對，而每一對都是一次出現相同生日的機會。這道題把這個直覺變得精確，而同樣的數學亦決定了保護你銀行戶口的密碼技術有多安全。

動筆之前

? **想一想** 為甚麼計算沒有兩位賓客同一天生日的概率會比較容易？請賓客排成一行：第二位避開第一位生日的機會是多少？第三位同時避開前兩位的機會又是多少？

「至少有兩人相同」包含許多互相重疊的情況（兩人相同、三人相同、兩對不同的人各自相同……）。它的補事件（complement）「全部不同」卻是單一而簡潔的事件。第二位賓客避開第一位生日的概率是 $\frac{364}{365}$；在此條件下，第三位同時避開前兩位的概率是 $\frac{363}{365}$；如此類推。

逐步解題

階段 1　一個有用的近似

想一想　$y = e^{-x}$ 在 $x = 0$ 處的切線（tangent line）是甚麼？

由於 $e^{-x} = 1 - x + \frac{x^2}{2} - \cdots$，當 x 很小時，我們有 $e^{-x} \approx 1 - x$。取 $x = r/n$，得

$$\frac{n-r}{n} = 1 - \frac{r}{n} \approx e^{-r/n}. \qquad \blacksquare$$

誤差約為 $\frac{r^2}{2n^2}$，當 r 遠小於 n 時，這個誤差很小。

階段 2　有人同一天生日的概率

k 個生日全部不同的概率是

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-(k-1)}{365} = \prod_{r=0}^{k-1} \frac{365-r}{365}.$$

對每個因子使用上述近似（approximation），再利用指數函數相乘時指數相加這個性質，

$$\prod_{r=0}^{k-1} e^{-r/365} = \exp\Big(-\frac{0+1+\cdots+(k-1)}{365}\Big) = \exp\Big(-\frac{k(k-1)}{730}\Big).$$

因此，至少有一對人同一天生日的概率約為 $1 - e^{-k(k-1)/730}$。 $\blacksquare$

想一想　看看指數：$\frac{k(k-1)}{730} = \binom{k}{2}\Big/365$。用文字表達，這說明了甚麼？

它是賓客的配對數目，乘以某一對人同一天生日的概率 $\frac{1}{365}$：也就是生日相同的配對數目的期望值（expected number）。這個近似告訴我們，*沒有任何配對相同的機會是* $e^{-(\text{相同配對的期望數目})}$，這又一次透露出泊松分佈（Poisson distribution）的影子，正如 Section 9.8 所見。

階段 3　二十三

我們需要 $1 - e^{-k(k-1)/730} \geq \frac{1}{2}$，即 $e^{-k(k-1)/730} \leq \frac{1}{2}$，亦即

$$\frac{k(k-1)}{730} \geq \ln 2 \approx \frac{253}{365}, \qquad \text{即} \qquad k(k-1) \geq 506.$$

由於 $22 \times 21 = 462$ 而 $23 \times 22 = 506$，最小的值是 $k = 23$。 $\blacksquare$

（23 位賓客時的精確概率是 0.507，22 位賓客時是 0.476，所以這個近似給出了正確答案。）

階段 4 與主人同一天生日

想一想 對每位賓客來說，不與主人同一天生日的概率是多少？在這方面，各位賓客之間是否互相獨立（independent）？

每位賓客各自獨立地以概率 $\frac{364}{365} \approx e^{-1/365}$ 避開主人的生日。因此 k 位賓客中無人與主人同一天生日的概率是 $\left(\frac{364}{365}\right)^k \approx e^{-k/365}$，而我們需要

$$e^{-k/365} \le \tfrac{1}{2}, \qquad \text{即} \qquad k \ge 365 \ln 2 \approx 253.$$

所以必須有 253 位賓客。 ■

這個對比正是重點所在：要與某一個指定的人生日相同，需要 253 位賓客；但要有某一對人生日相同，只需 23 人，因為 23 個人可以組成 253 對。

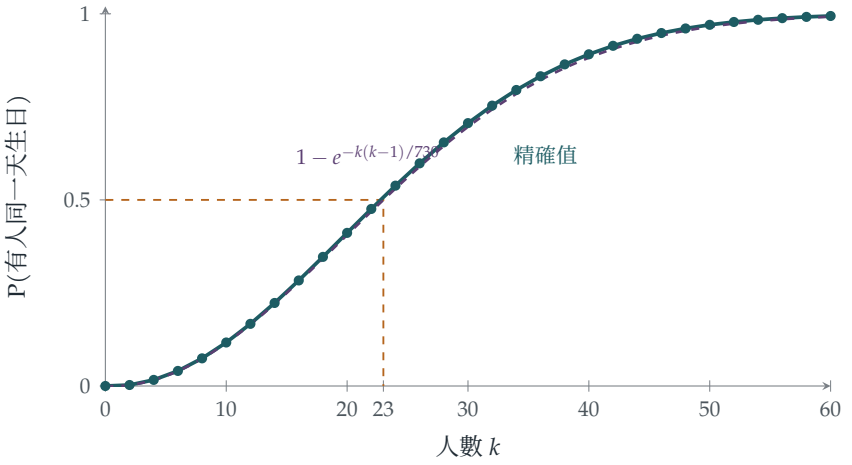


圖 9.9.1 k 個人中至少有兩人同一天生日的概率：精確值（藍綠色）與近似值 $1 - e^{-k(k-1)/730}$（紫色虛線）。兩者非常接近，並且都在 $k = 22$ 與 $k = 23$ 之間越過 $\frac{1}{2}$。

回顧

- **利用補事件。**「全部不同」只是一個乘積；「至少有一對相同」卻是一團亂麻。
- **先線性化，再取指數。**把每個因子 $1 - x$ 換成 e^{-x}，乘積就變成了和。
- **數配對。**指數正是配對的數目乘以一對人相同的機會。$k(k-1)$ 中的平方，就是這個悖論的全部解釋。

背後的數學

平方根定律

若把 365 天換成 n 個等可能的日子，同樣的論證給出

$$\mathrm{P}(\text{有人相同，人數為}k) \approx 1 - e^{-k^2/(2n)},$$

所以只要 k 達到 $\sqrt{n}$ 左右，而不是 n，就很可能出現相同的情況。更準確地說，概率在 $k \approx 1.18\sqrt{n}$ 時達到 $\frac{1}{2}$，而在首次出現重複之前你需要遇到的人數，期望值約為 $\sqrt{\pi n/2}$（當 $n = 365$ 時，約為 24.6）。這條*平方根定律*（square-root law）正是巧合如此常見的原因：在任何一大批事件之中，總很可能有某一對在某方面吻合，儘管任何一個指定的吻合都不大可能發生。

生日攻擊

電腦會計算檔案的*雜湊值*（hash），以檢查檔案有否被竄改：那是一個指紋，可能的取值有（比方說）$n = 2^{128}$ 個。偽造者若想得到兩份雜湊值相同的文件（一份無害的合約和一份欺詐的合約），並不需要試 2^{128} 份文件。根據生日問題的論證，為每份文件各自產生大約 $\sqrt{2^{128}} = 2^{64}$ 個變體，就很可能得到一對相同的雜湊值。這種*生日攻擊*（birthday attack）正是現代雜湊函數採用 256 位元或以上的原因：有一半的位元因平方根定律而白白損失。類似的計算亦決定了隨機識別碼必須有多長，才能令兩部電腦永遠不會碰巧選中同一個。

誰最先提出？

生日問題通常歸功於馮·米塞斯（Richard von Mises），他在 1939 年討論過這個問題，不過它在此之前已經流傳。據說 Harold Davenport 亦更早知道這個問題，只是覺得它太簡單而沒有發表。按 Warren Weaver 的說法，他在第二次世界大戰期間喜歡在軍官的晚宴上示範這個問題：他打賭在場二十二人之中有兩人同一天生日，結果輸了，直至一位女侍應透露她與其中一位軍官同一天生日。

思維技巧

計算補事件；小心地近似

當「至少一個」很複雜時，就去求「一個也沒有」的概率。當乘積難以處理時，把每個因子近似爲指數函數，然後把指數相加。

這兩個做法在大學裏都是標準手法。第二個，即 $1 - x \approx e^{-x}$，是泊松近似理論的起點，也是漸近方法（asymptotic methods）的起點；後者用來估計那些複雜得無法精確計算的量。

延伸練習

1 **首次重複。**試做 STEP II 2014，第 13 題（`step.maths.org/questions/14-s2-q13`）。這道題研究隨機數產生器（random number generator）首次出現重複的值，其實是換了個面貌的生日問題，並把有關概率化成兩個出人意表的恆等式（identities）。

2 **三人相同。**大約需要多少人，才會令某三人同一天生日的機會大於一半？（提示：同一天生日的三人組數目的期望值是 $\binom{k}{3}/365^2$。令它約等於 $\ln 2$。）真正的答案是 88。

3 **差之毫釐。**需要多少人，才會令兩人的生日在同一天或相鄰兩天的機會達到一半？

4 **不均勻的生日。**現實中的生日並非均勻（uniform）分佈。證明若各日的概率為 $p_1, \ldots, p_{365}$，則兩個人同一天生日的機會是 $\sum p_i^2$，並且當所有 p_i 相等時這個值最小。（所以不均勻只會令相同的情況更容易出現。）

延伸閱讀

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 II 章。生日問題及其各種近似。
- **Persi Diaconis & Frederick Mosteller,「Methods for studying coincidences」**，*Journal of the American Statistical Association* 84 (1989)。解釋巧合為何常見；非常易讀。
- **David Hand, *The Improbability Principle***（Bantam）。平方根定律，以及不大可能的事情為何時刻都在發生的其他原因。

9.10 總能找贖

題目

STEP I 2011，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/11-s1-q12`

引子

這個規律實在叫人難以抗拒：$\frac{m}{m+1}$、$\frac{m-1}{m+1}$、$\frac{m-2}{m+1}$。若有 n 個人拿着 £2 硬幣，答案想必就是 $\frac{m-n+1}{m+1}$。事實正是如此，而最巧妙的證明（proof）用上了組合數學（combinatorics）中數一數二優雅的想法：*反射原理*（reflection principle）。它把圖形的一部分沿一條直線反射，從而將一個困難的計數問題（從不缺零錢找贖的隊伍有多少種）變成一個容易的計數問題。

同一個問題有許多不同的面貌：點票過程中勝出者從未落後的選舉（1887 年的伯特蘭投票問題（Bertrand's ballot problem））、永不破產的賭徒、正確配對的括號。當 $m = n$ 時，數出來的就是*卡塔蘭數*（Catalan numbers），這大概是組合數學中除二項式係數（binomial coefficients）以外最常出現的數列（sequence）。

動筆之前

想一想 每位拿 £1 的顧客會給你一枚硬幣；每位拿 £2 的顧客會取走一枚（作為找贖）。你甚麼時候會無法繼續？

留意你手上 £1 硬幣的數目。每來一位 £1 顧客，數目加 1；每來一位 £2 顧客，數目減 1。你無法繼續的情況，恰好就是一位 £2 顧客來到而你手上沒有硬幣：數目會變成 -1。所以你能成功，當且僅當（if and only if）在隊伍中的每一刻，到目前為止的 £1 顧客人數都不少於到目前為止的 £2 顧客人數。

逐步解題

階段 1　一枚 £2 硬幣

當 $n = 1$ 時，只有 £2 顧客排第一才會失敗。£2 顧客排在 $m+1$ 個位置中任何一個的可能性均等，所以

$$\mathrm{P}(\text{成功}) = 1 - \frac{1}{m+1} = \frac{m}{m+1}. \qquad \blacksquare$$

階段 2　兩枚 £2 硬幣

> **想一想**　第一個人必須拿着 £1 硬幣。第二個人可以出現甚麼情況？在每種情況下，你餘下面對的是甚麼局面？

第一個人拿着 £1 硬幣的概率（probability）是 $\frac{m}{m+2}$；否則你立即失敗。這時你手上有一枚硬幣，而餘下的隊伍有 $m-1$ 個 £1 的人和兩個 £2 的人。

- 若第二個人拿着 £1 硬幣（概率為 $\frac{m-1}{m+1}$），你手上便有兩枚硬幣，無論兩枚 £2 硬幣何時出現都足夠找贖。成功是必然的。
- 若第二個人拿着 £2 硬幣（概率為 $\frac{2}{m+1}$），你手上又變回沒有硬幣，餘下 $m-1$ 個 £1 的人和一個 £2 的人。根據第 (i) 部分，以 $m-1$ 代替 m，你成功的概率是 $\frac{m-1}{m}$。

所以

$$\mathrm{P}(\text{成功}) = \frac{m}{m+2}\left(\frac{m-1}{m+1} + \frac{2}{m+1}\cdot\frac{m-1}{m}\right) = \frac{m}{m+2}\cdot\frac{m-1}{m+1}\cdot\frac{m+2}{m} = \frac{m-1}{m+1}. \qquad \blacksquare$$

> **停一停，試一試**　試用同樣的方法處理 $n = 3$。這是行得通的，但情況會愈分愈多。有沒有方法可以一次過數出所有成功的隊伍？

階段 3　把隊伍看成路徑

> **想一想**　以你手上的硬幣數目對已服務的人數作圖。一條隊伍看起來是甚麼樣子？一條成功的隊伍又是甚麼樣子？

一條隊伍就是一條路徑，由 m 步上升（£1）和 n 步下降（£2）組成，從高度 0 出發，終於高度 $m-n$。這樣的路徑共有 $\binom{m+n}{n}$ 條，每條出現的可能性均等。一條隊伍成功，恰好就是它的路徑從不低於 0。所以我們需要數出壞路徑，即那些觸及高度 -1 的路徑。

階段 4　反射

想一想 取一條壞路徑。找出它第一次觸及 -1 的時刻，把路徑在這一刻之前的部分沿直線 $y=-1$ 反射。新的路徑從哪裏開始，又在哪裏結束？

反射後的路徑從 -2 出發（即 0 對直線 $y=-1$ 的鏡像），終點與原來的路徑相同，即 $m-n$（圖 9.10.1）。反過來說，每條由 -2 到 $m-n$ 的路徑都必定在某處穿過 $y=-1$（因為 $m\geq n$，它起點在這條線之下，終點在這條線之上）；把它第一次到達這條線之前的部分反射，便得回一條由 0 出發的壞路徑。所以由 0 出發的壞路徑，與由 -2 到 $m-n$ 的*所有*路徑恰好一一對應。

一條以 $m+n$ 步由 -2 走到 $m-n$ 的路徑共上升了 $m-n+2$，所以它有 $m+1$ 步上升和 $n-1$ 步下降：這樣的路徑有 $\binom{m+n}{n-1}$ 條。因此

$$\mathrm{P}(\text{失敗})=\frac{\binom{m+n}{n-1}}{\binom{m+n}{n}}=\frac{n!\,m!}{(n-1)!\,(m+1)!}=\frac{n}{m+1},$$

而

$$\mathrm{P}(\text{成功})=1-\frac{n}{m+1}=\frac{m-n+1}{m+1}.$$

當 $n=3$ 時，這就是 $\frac{m-2}{m+1}$。 ■

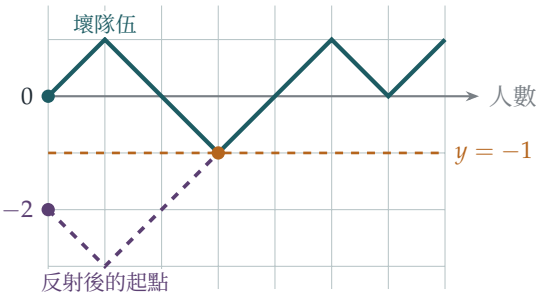


圖 9.10.1 一條有 $m=4$ 枚 £1 硬幣和 $n=3$ 枚 £2 硬幣的隊伍，畫成一條路徑（藍綠色）。它在第三個人時首次到達 -1（琥珀色），這時你無法找贖。把路徑在這一點之前的部分沿直線 $y=-1$ 反射，便得到一條由 -2 出發（紫色虛線）而終點相同的路徑。壞隊伍與由 -2 出發的路徑之間的這種配對，就是反射原理。

回顧

- **找出合適的狀態。**唯一要緊的是手上硬幣的數目，這把隊伍變成了一次遊走。
- **以最初幾步為條件。**第 (i) 和第 (ii) 部分的做法，是考慮最初幾個人的情況，從而化約為較小的情形。

- **用雙射（bijection）數出補集。**壞路徑難以直接數出，但反射把它們與一個容易數算的集合配對起來。

背後的數學

卡塔蘭數

當 $m = n$（£1 硬幣剛好足夠）時，成功隊伍的數目是

$$\binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n},$$

即第 n 個卡塔蘭數 C_n：$1, 1, 2, 5, 14, 42, 132, \ldots$。Richard Stanley 的著作 *Catalan Numbers* 列出了超過二百種由卡塔蘭數計算數目的事物。其中包括：把 n 對括號正確排列的方法數目（「(())()」是允許的，「())(()」則不允許，理由正是上述的硬幣論證，只要把「(」看成 £1、「)」看成 £2）、用互不相交的對角線把一個 $n+2$ 邊形分割成三角形的方法數目，以及為 $n+1$ 個因子的乘積加上括號的方法數目。

隨機遊走

隊伍就是隨機遊走（random walk），而反射原理是研究隨機遊走的基本工具之一。舉例說，它可以證明由 0 出發的公平隨機遊走必定終會返回 0，也可以證明它在最初 $2n$ 步一直保持嚴格為正的概率是 $\frac{1}{2}\binom{2n}{n}2^{-2n} \approx \frac{1}{2\sqrt{\pi n}}$。在連續的形式下，把同樣的反射論證應用於布朗運動（Brownian motion），便可得出其最大值的分佈（distribution）。這是某些金融期權定價的基礎，這類期權的回報取決於股價有否越過某個界限。

伯特蘭投票問題

1887 年，約瑟夫·伯特蘭（Joseph Bertrand）提出：在一場選舉中，候選人 A 得 a 票，B 得 $b < a$ 票，那麼 A 在整個點票過程中一直嚴格領先的概率是多少？答案是 $\frac{a-b}{a+b}$，同年由 Désiré André 證明，他的論證通常被視為反射原理（不過歷史學家指出，André 實際的證明用的是一個稍有不同的雙射）。我們的抽獎券問題是一個容許平手的變體：成功的意思是 £1 的數目從不低於 £2 的數目。其中一道練習會由我們的答案推導出伯特蘭的答案。

思維技巧

雙射計數

要數算一個集合，就找出它與一個你懂得數算的集合之間的一一對應（*one-to-one correspondence*）。

反射把壞隊伍與由另一個起點出發、不受限制的路徑配對起來。雙射是組合數學中最令人滿足的證明，因為它們解釋了兩個數爲何相等，而不只是證明它們相等。在大學裏，你會在組合數學中處處遇到雙射證明，也會在研究隨機遊走和布朗運動時再次遇上反射原理。

延伸練習

1 **分情況處理三枚。**仿照第 (ii) 部分，以隊伍中最初幾個人為條件來證明第 (iii) 部分，看看反射原理省卻了多少工夫。

2 **伯特蘭投票問題。**在一場選舉中，A 得 a 票，B 得 $b < a$ 票。證明 A 在整個點票過程中一直嚴格領先的概率是 $\frac{a-b}{a+b}$。（提示：第一票必須投給 A；此後，A 落後於 B 的票數絕不可多於…。）

3 **括號。**列出三對括號的 5 種正確排列，以及把五邊形分割成三角形的 5 種方法。你能找出兩者之間的一個雙射嗎？

4 **一張 £5 紙幣。**一條隊伍中有 m 個人拿着 £1 硬幣，n 個人拿着 £5 紙幣，而一張票的價錢是 £1。解釋為何反射原理不再給出簡單的答案。（這個推廣困難得多。）

延伸閱讀

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1**（Wiley），第 III 章。有關隨機遊走和反射原理的經典論述。
- **Richard Stanley, *Catalan Numbers***（Cambridge University Press）。超過二百種詮釋，並附練習。
- **Marc Renault, "Lost (and found) in translation: André's actual method and its application to the generalized ballot problem"**，*American Mathematical Monthly* 115 (2008)。反射原理的歷史。

9.11 集齊一套

題目

STEP II 1994，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/94-s2-q13`

引子

一本足球貼紙簿要集齊 600 張貼紙，貼紙每次隨機發售一張。要貼滿整本貼紙簿，你預計要買多少張貼紙？最初幾百張很快到手，因為幾乎每一張都是新的。但貼紙簿愈貼愈滿，新貼紙就愈來愈罕有，單是最後一張，平均就要買 600 次。總數大約是 4200 張，是貼紙簿大小的七倍。

這就是*集券問題*（coupon collector's problem）。它的答案結合了兩個想法：等候一次成功所需的時間，以及調和級數（harmonic series）$1 + \frac{1}{2} + \frac{1}{3} + \cdots$，後者的增長與 $\ln n$ 相若。你在 Section 2.3 見過調和級數發散（diverge）；在這裏，它那緩慢的增長正正就是集齊一套收藏品所要付出的代價。

動筆之前

想一想 如果每包各自獨立地以概率 p 開出一張新卡，那麼直至開出第一張新卡為止（連這一包計算在內），你開包數目的期望值（expectation）是多少？

開包數目 X 服從*幾何分佈*（geometric distribution）：$\mathrm{P}(X = k) = (1-p)^{k-1}p$。它的期望值是 $\frac{1}{p}$。（有一個快捷的看法：$\mathrm{E}(X) = 1 + (1-p)\mathrm{E}(X)$，因為你先開一包，然後有 $1-p$ 的概率回到起點。）如果有三分之一的機會開出新卡，你預計要等三包。

逐步解題

階段 1　等候一張新卡

手上已有 r 張不同的卡時，一包開出新卡的概率是 $\frac{n-r}{n}$，與之前各包獨立。因此，直至開出下一張新卡為止，所需包數的期望值是

$$\frac{1}{(n-r)/n} = \frac{n}{n-r}. \qquad \blacksquare$$

階段 2　比較和與積分

> **想一想**　在區間 $r \le x \le r+1$ 上，$\frac{1}{x}$ 介乎哪兩個值之間？這對曲線下的面積有甚麼啟示？

由於 $\frac{1}{x}$ 是遞減的，在 $[r, r+1]$ 上有 $\frac{1}{r+1} \le \frac{1}{x} \le \frac{1}{r}$。在長度為 1 的區間上積分，得

$$\frac{1}{r+1} \le \int_r^{r+1} \frac{\mathrm{d}x}{x} \le \frac{1}{r}. \qquad \blacksquare$$

用圖來看（圖 9.11.1）：曲線下的面積夾在高度為 $\frac{1}{r+1}$ 的長方形與高度為 $\frac{1}{r}$ 的長方形之間。把這些不等式由 $r=1$ 加到 $r=n-1$，各個積分便連接成 $\int_1^n \frac{\mathrm{d}x}{x} = \ln n$，於是

$$\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \le \ln n \le 1 + \frac{1}{2} + \cdots + \frac{1}{n-1}. \qquad \blacksquare$$

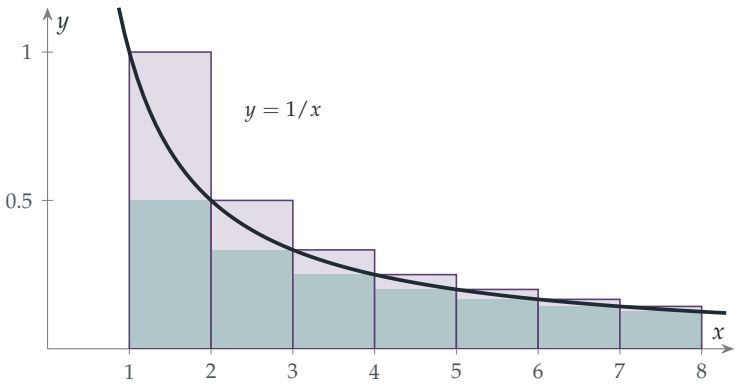


圖 9.11.1　在每個區間 $[r, r+1]$ 上，$y = 1/x$ 之下的面積介乎藍綠色長方形（高度 $\frac{1}{r+1}$）與較大的紫紅色長方形（高度 $\frac{1}{r}$）之間。全部加起來，$\ln n$ 便介乎 $\frac{1}{2} + \cdots + \frac{1}{n}$ 與 $1 + \frac{1}{2} + \cdots + \frac{1}{n-1}$ 之間。曲線與紫紅色長方形之間的空隙，加起來是一個有限的總和：歐拉常數（Euler's constant）$\gamma \approx 0.577$。

階段 3　集齊全套

想一想 總包數是等候第一張卡、第二張新卡……各段等候時間之和。它的期望值是多少？

設 X_r 為由 r 張不同的卡增至 $r+1$ 張所需的包數。總數是 $X_0 + X_1 + \cdots + X_{n-1}$，而期望值可以相加，所以

$$\mathrm{E}(\text{總數}) = \sum_{r=0}^{n-1} \frac{n}{n-r} = n\Big(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\Big) = nH_n,$$

其中 $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ 是第 n 個調和數（harmonic number）。由第 2 步，$\ln n + \frac{1}{n} \le H_n \le \ln n + 1$，所以當 n 很大時 H_n 約為 $\ln n$，而我預計大約要吃掉

$$n \ln n \text{ 包.} \qquad \blacksquare$$

回顧

- **把漫長的等候分成多個階段。**總時間拆成 n 段等候時間，每一段都屬於簡單的幾何分佈。
- **期望值可以相加。**我們從頭到尾都不需要總數的分佈，只需要各部分的期望值。
- **比較和與積分。**那幅圖把一個我們無法求值的和化成一個對數，誤差至多為 1。

背後的數學

這個估計有多準確？

這個比較還可以再精確一些。圖 9.11.1 中曲線與上方長方形之間的陰影空隙，總面積是有限的，所以 $H_n - \ln n$ 趨向一個極限，

$$\gamma = \lim_{n\to\infty} (H_n - \ln n) = 0.5772156649\ldots,$$

即歐拉常數。（沒有人知道 γ 是否有理數。）因此包數的期望值是 $n\ln n + \gamma n + \frac{1}{2} + \cdots$。以 600 張貼紙的貼紙簿為例，$600 \times (\ln 600 + 0.577) \approx 4185$，這就是引子中所引述的數字。

時間都花到哪裏去了？各個階段極不平均。集齊全套的前一半只需大約 $n\ln 2 \approx 0.69n$ 包，而單是最後一張卡，平均就要 n 包。大部分工夫都花在最後幾張卡上，這正是收藏者要互相交換重複卡的原因。

集中現象

期望值只是個開始：總數的變動有多大？總數的方差（variance）是 $\sum nr/(n-r)^2$，約為 $\frac{\pi^2}{6}n^2$（Section 2.4 中的和 $\sum 1/k^2 = \pi^2/6$ 在這裏出現了）。所以標準差（standard deviation）約為 $1.28n$，當 n 很大時遠小於平均值 $n\ln n$：總數集中在其平均值附近。艾狄胥（Paul Erdős）和雷尼（Alfréd Rényi）在 1961 年證明了一個更精確的結果：總數減去 $n\ln n$ 再除以 n，會有一個極限分佈（limiting distribution），即耿貝爾分佈（Gumbel distribution），這個分佈也描述了許多隨機量之中最大的一個（Section 12.25）。

無處不在的集券問題

這個問題可追溯至棣莫弗（Abraham de Moivre，1712 年）和拉普拉斯（Laplace）。如今它以許多不同的面貌出現。要檢查一個系統的每個組件，需要做多少次隨機測試？在一個網絡上進行隨機漫步（random walk），要多久才會走遍每個節點？電腦的雜湊表（hash table）要探查多少次？每一種情況的答案都涉及 $n\ln n$。在生物學中，同樣的計算可用來估計要有多少個定序「讀段」（reads）才能覆蓋整個基因組。

思維技巧

期望值的線性性質（linearity of expectation）

和的期望值等於期望值之和，不論各部分是否獨立。把一個複雜的隨機量拆成簡單的部分。

這是概率論中最有用的事實之一。在這裏，它把集券問題簡化為 n 段幾何等候；在 Section 9.8 中，它不用任何計算就給出了配對數目的期望值。到了大學，同一個想法推動着演算法分析，以及組合數學中的概率方法（probabilistic method）。

延伸練習

1 **半套。**證明集齊 $\frac{n}{2}$ 張不同的卡（n 為偶數）所需包數的期望值是 $n(H_n - H_{n/2})$，並且這個值約為 $n\ln 2$。

2 **不同的卡。**證明開了 k 包之後，你手上不同的卡的數目的期望值是 $n\big(1-(1-\frac{1}{n})^k\big)$。（提示：對每一張卡，求出你擁有它的概率，然後相加。）大約要開多少包，才能得到全套的 95%？

3 **兩位收藏者。**兩個小孩收集同一套卡，並共用所開的各包，因此對其中一人來說是重複的卡，對另一人來說可能是新的。證明集齊兩套所需包數的期望值約為 $n(\ln n + \ln\ln n)$。（這題較難；先說服自己這個數目遠小於 $2n\ln n$。）

4 **更精確的界。**改用梯形法則（trapezium rule）的圖像而不用長方形，證明 $H_n - \ln n$ 是遞減的，而 $H_n - \ln n - \frac{1}{2n}$ 是遞增的。由較小的 n 值推導出 $0.5 < \gamma < 0.7$。

延伸閱讀

- **Geoffrey Grimmett & David Stirzaker, *One Thousand Exercises in Probability***（Oxford University Press）。集券問題的許多變化。
- **Julian Havil, *Gamma: Exploring Euler's Constant***（Princeton University Press）。關於 γ 和調和級數的一切，寫得深入淺出。
- **Michael Mitzenmacher & Eli Upfal, *Probability and Computing***（Cambridge University Press）。集券問題及其在電腦科學中的應用。

9.12 隨機子集

題目

STEP III 2019，第 12 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 OCR 提供下載的 2019 年 STEP 試卷（zip 檔案）內的 STEP III 試卷：

`www.ocr.org.uk/Images/696590-2019-past-papers.zip`

引子

完全隨機地選出 $\{1,2,\ldots,100\}$ 的兩個子集（subset）。它們沒有任何共同元素的可能性有多大？直覺也許會告訴你：「相當大：大多數成對的集合都不相同」。事實上，這個概率是 $\left(\frac{3}{4}\right)^{100}$，約為 3×10^{-13}。隨機集合很大（大約佔全部元素的一半），所以它們幾乎總會重疊。

這道題很短，當中唯一的想法卻很有威力：不要把集合當作整體來想，而要每次只想一個元素。一個隨機子集不過是 n 次獨立（independent）的擲硬幣，每個元素擲一次；於是，凡是關於隨機子集的問題，都變成關於這些擲硬幣結果的問題。

動筆之前

? **想一想** 怎樣用 n 條是非題的一串答案，來描述 $\{1,\ldots,n\}$ 的一個子集？

對每個元素 i，回答「i 在集合裏嗎？」這條問題。子集決定了答案，答案也決定了子集。每個元素有 2 個選擇，所以一共有 2^n 個子集。均勻隨機（uniformly at random）地選一個子集，就等於獨立地擲一枚公平硬幣來回答每條問題。

逐步解題

階段 1 一個元素

按照擲硬幣的描述，$\mathrm{P}(1 \in A_1) = \frac{1}{2}$。事實上，$2^n$ 個子集中恰好有一半包含 1：把每個不包含 1 的子集，與加入 1 之後所得的子集配成一對即可。

階段 2　不相交的集合

想一想　要使 $A_1 \cap A_2 = \varnothing$，每個元素 i 必須滿足甚麼條件？就單一個元素而言，這件事的概率是多少？為甚麼可以把各個元素的概率相乘？

A_1 與 A_2 不相交（disjoint），當且僅當沒有元素同時屬於兩者。對單一個元素 i 來說，有四種等可能的情況：兩個集合都不在、只在 A_1、只在 A_2、兩個都在（圖 9.12.1 左）。只有最後一種是禁止的，所以 i 不同時屬於兩者的概率是 $\frac{3}{4}$。不同的元素由不同而且獨立的擲硬幣結果決定，所以

$$\mathrm{P}(A_1 \cap A_2 = \varnothing) = \left(\frac{3}{4}\right)^n. \qquad \blacksquare$$

對三個集合而言，元素 i 同時屬於三者的概率是 $\frac{1}{8}$，所以 $\mathrm{P}(A_1 \cap A_2 \cap A_3 = \varnothing) = \left(\frac{7}{8}\right)^n$。一般而言，

$$\mathrm{P}(A_1 \cap \cdots \cap A_m = \varnothing) = \left(1 - \frac{1}{2^m}\right)^n. \qquad \blacksquare$$

階段 3　子集

想一想　要使 $A_1 \subseteq A_2$，一個元素的四種情況之中，哪一種是禁止的？

$A_1 \subseteq A_2$ 不成立，當且僅當有某個元素在 A_1 裏而不在 A_2 裏。同樣是四種模式中有一種被禁止，所以

$$\mathrm{P}(A_1 \subseteq A_2) = \left(\frac{3}{4}\right)^n,$$

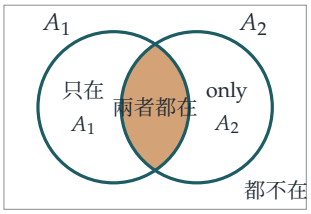
與不相交的情況相同。這並非巧合：$A_1 \subseteq A_2$ 當且僅當 A_1 與 A_2 的補集（complement）不相交，而隨機集合的補集也是一個隨機集合。

想一想　要使 $A_1 \subseteq A_2 \subseteq A_3$，一個元素的 8 種歸屬模式（在或不在每個集合裏）之中，哪些是容許的？

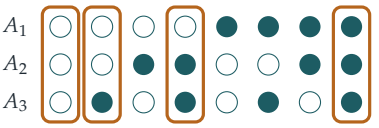
沿着 A_1, A_2, A_3 這條鏈（chain）走，一個元素一旦進入，就必須一直留在裏面。所以容許的模式是「不在、不在、不在」、「不在、不在、在」、「不在、在、在」和「在、在、在」：8 種之中的 4 種（圖 9.12.1 右）。因此 $\mathrm{P}(A_1 \subseteq A_2 \subseteq A_3) = \left(\frac{4}{8}\right)^n = \left(\frac{1}{2}\right)^n$。

對於由 m 個集合組成的鏈，一個容許的模式取決於該元素最先進入哪一個集合（$A_1, \ldots, A_m$ 其中之一，或者從不進入），所以 2^m 種模式之中有 $m+1$ 種是容許的：

$$\mathrm{P}(A_1 \subseteq A_2 \subseteq \cdots \subseteq A_m) = \left(\frac{m+1}{2^m}\right)^n. \qquad \blacksquare$$



每個元素：4 個等可能的位置



符合 $A_1 \subseteq A_2 \subseteq A_3$ 的模式：4 種（共 8 種）

圖 9.12.1　左：S 的一個元素相對於 A_1 和 A_2 可以處於四個位置之一，每個位置的概率都是 $\frac{1}{4}$。不相交這個條件禁止「兩者都在」（琥珀色）；$A_1 \subseteq A_2$ 則禁止「只在 A_1」。右：元素對 A_1、A_2、A_3 的八種歸屬模式（實心圓表示「在」）。鏈 $A_1 \subseteq A_2 \subseteq A_3$ 只容許加了方框的四種模式，即元素一旦進入便一直留在裏面。

回顧

- **按元素分解。**每個集合層面的條件（不相交、子集、鏈），都是分別加在每個元素上的條件。
- **元素之間互相獨立。**由於各個元素由獨立的擲硬幣結果決定，概率可以相乘。
- **數算模式。**每個答案都是（容許的模式數 ÷ 全部模式數）^{n}。

背後的數學

不用概率的計數

把每個概率乘以選擇的總數，就得到一個計數。滿足 $A_1 \subseteq A_2$ 的有序對 (A_1, A_2) 的數目是 $4^n \times \left(\frac{3}{4}\right)^n = 3^n$。你也可以直接數算這些有序對：先選定 A_2 的大小 k，然後以 $\binom{n}{k}$ 種方法選出 A_2 本身，再從它的 2^k 個子集中任選一個作為 A_1。所以

$$\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n,$$

這正是 $(1+2)^n$ 的二項式定理（binomial theorem）。用兩種方法數算同一樣東西，一次按元素、一次按大小，便證明了一條恆等式：這就是算兩次（double counting）的方法，你在 Section 7.9 討論柏拉圖立體時已經見過。同樣地，鏈 $A_1 \subseteq \cdots \subseteq A_m$ 的數目是 $(m+1)^n$。

子集的格

$\{1, \ldots, n\}$ 的所有子集按包含關係排序，形成一個稱為布爾格（Boolean lattice）的結構，可以把它想像成一個 n 維立方體：每個子集是一個頂點，以它那一串是非題的答案作標

記，而沿着一條棱移動，就是加入或移走一個元素。鏈就是向上走的路徑。有些深刻的結果描述了在給定條件下，子集族可以有多大。施佩納定理（Sperner's theorem，1928 年）指出，在沒有一個集合包含另一個集合的子集族之中，最大的一族有 $\binom{n}{\lfloor n/2 \rfloor}$ 個集合；艾狄胥—柯—拉多定理（Erdős–Ko–Rado theorem，1961 年）則說明，最多可以選出多少個 k 元集合，使得任意兩個都有共同元素。兩者都是組合數學（combinatorics）入門課程的標準內容，而把子集看成擲硬幣的觀點，給出了其中一些最漂亮的證明。

布爾的代數

布爾（George Boole）的 *The Laws of Thought*（《思維規律》，1854 年）把邏輯命題當作集合，把「與」、「或」、「非」當作交集、併集和補集。把一個集合編碼成一串是非題的答案，正是電腦表示集合的方法：$\{1, \ldots, 64\}$ 的一個子集就是單獨一個 64 位元的數，而交集就是「逐位元 AND」（bitwise AND）指令。本題的方法，即每次只看一個元素（一個位元），正是分析這類運算的方法。

思維技巧

每次只想一個元素

當對象是由每個元素各自的獨立選擇構成時，一個分別加在每個元素上的條件，其概率等於一個乘積。

這個方法把一個關於 2^n 個子集的問題，變成一個關於 4 種或 8 種模式的問題。到了大學，它會在組合數學（集合與多重集的生成函數（generating functions））、概率論（獨立的坐標）和計算機科學（逐位元的推理）中再次出現。

延伸練習

1 **覆蓋。**求 $\mathrm{P}(A_1 \cup A_2 = S)$ 和 $\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_m = S)$。

2 **大小相同。**證明 $\mathrm{P}(|A_1| = |A_2|) = \binom{2n}{n}/4^n$。（提示：$|A_1| = |A_2|$ 當且僅當 $|A_1| + |S \setminus A_2| = n$。）證明當 n 很大時，這個概率約為 $1/\sqrt{\pi n}$。

3 **分成三份。**證明 S 的不相交子集所組成的有序對 (A, B) 的數目是 3^n，並求由不相交子集所組成的無序對 $\{A, B\}$ 的數目。

4 **小 n 的施佩納定理。**對 $n = 3$ 和 $n = 4$，求 $\{1, \ldots, n\}$ 的子集族中，沒有一個集合包含另一個集合的最大一族。

延伸閱讀

- **Miklós Bóna, *A Walk Through Combinatorics***（World Scientific），其中關於子集和偏序集（partially ordered sets）的章節。
- **Béla Bollobás, *Combinatorics: Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability***（Cambridge University Press）。一本簡短而優雅的書，講述子集族，由施佩納定理說起。
- **Peter Cameron, *Combinatorics: Topics, Techniques, Algorithms***（Cambridge University Press）。一門入門課程，特別着重用兩種方法計數。

9.13 做了手腳的骰子

題目

STEP I 1999，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/99-s1-q12`

引子

擲兩顆公平骰子時，點數總和為 7 的可能性是總和為 2 的六倍。桌上遊戲的設計者有時會希望情況並非如此：能否在骰子上做手腳（加上配重，使某些面較常出現），令 2 至 12 的每個總和都有相同的機會，即 $\frac{1}{11}$？那十二個概率（probability）可以任你選擇，看來迴旋的餘地相當充裕。

答案是不能，而證明堪稱一絕：它只用到十一個總和中的三個，以及不等式（inequality）$x + \frac{1}{x} \geq 2$。這是證明某件事不可能的典範；與證明某件事辦得到相比，這往往更難，也更有趣。

動筆之前

想一想 對於 $x > 0$，你會怎樣證明 $x + \frac{1}{x} \geq 2$？等號何時成立？

兩邊同乘 $x > 0$：不等式變成 $x^2 + 1 \geq 2x$，即 $(x-1)^2 \geq 0$，這顯然成立。等號只在 $x = 1$ 時成立。（另一做法是 $x + \frac{1}{x} - 2 = (\sqrt{x} - \frac{1}{\sqrt{x}})^2$。）這是算術平均—幾何平均不等式（AM–GM inequality）最簡單的形式：x 與 $\frac{1}{x}$ 的平均值不小於它們的幾何平均（geometric mean）1。

逐步解題

階段 1　兩端的總和

想一想　擲出總和 2 有多少種方法？總和 12 呢？試以 q_i 和 r_j 表示它們的概率。

總和為 2 需要兩顆骰子都擲出 1，總和為 12 則需要兩顆都擲出 6，所以 $\mathrm{P}(2) = q_1 r_1$ 及 $\mathrm{P}(12) = q_6 r_6$。

為了引出矛盾，假設全部十一個總和的可能性相等，每個的概率都是 $\frac{1}{11}$。那麼

$$q_1 r_1 = \frac{1}{11}, \qquad q_6 r_6 = \frac{1}{11}.$$

特別地，q_1、r_1、q_6、r_6 四者全為正數。

階段 2　中間的總和

想一想　總和 7 可以由六種方法得出。其中哪兩種只涉及 1 和 6 這兩個面？試利用上面的等式寫出它們的概率。

總和 7 包括結果 $(1, 6)$ 和 $(6, 1)$，所以

$$\mathrm{P}(7) \geq q_1 r_6 + q_6 r_1.$$

現在利用第 1 步的等式引入一個比。由於 $q_1 = \frac{1}{11 r_1}$ 及 $q_6 = \frac{1}{11 r_6}$，

$$q_1 r_6 + q_6 r_1 = \frac{1}{11}\left(\frac{r_6}{r_1} + \frac{r_1}{r_6}\right) \geq \frac{1}{11} \times 2 = \frac{2}{11},$$

這裏用了 $x + \frac{1}{x} \geq 2$，其中 $x = r_6 / r_1 > 0$。因此 $\mathrm{P}(7) \geq \frac{2}{11}$，與 $\mathrm{P}(7) = \frac{1}{11}$ 矛盾。各個總和不可能全部等可能。 ■

停一停，試一試　看看這個證明為何行得通。要令 2 和 12 與 7 同樣可能出現，1 和 6 就必須有相當高的概率，於是單憑 $(1, 6)$ 和 $(6, 1)$ 這兩個「混合」結果，已足以令 7 的概率過高。圖 9.13.1 顯示了所涉及的方格。



圖 9.13.1　兩顆骰子的 36 個結果，各以其總和標示。證明只用到四個角落的方格。兩個琥珀色方格的概率之積，等於兩個深藍綠色方格的概率之積，即 $q_1r_1q_6r_6 = q_1r_6 \cdot q_6r_1$；而根據算術平均—幾何平均不等式，藍綠色那一對之和至少是琥珀色那一對的幾何平均的兩倍。

階段 3　同一顆骰子擲兩次

? **想一想**　如果把一顆骰子擲兩次，兩次點數相同的概率是多少？試以 q_j 表示。對公平骰子而言，這個概率是多少？

兩次擲出相同點數的概率是 $q_1^2 + q_2^2 + \cdots + q_6^2$。對公平骰子而言，這是 $6 \times \frac{1}{36} = \frac{1}{6}$，已經遠大於 $\frac{1}{36}$。

? **想一想**　在骰子上做手腳，能否令 $\sum q_j^2$ 小於 $\frac{1}{6}$？你知道 $\sum q_j = 1$。

寫成 $q_j = \frac{1}{6} + e_j$，其中各偏差 e_j 之和為零。那麼

$$\sum q_j^2 = \sum\left(\frac{1}{36} + \frac{e_j}{3} + e_j^2\right) = \frac{1}{6} + \frac{1}{3}\sum e_j + \sum e_j^2 = \frac{1}{6} + \sum e_j^2 \geq \frac{1}{6}.$$

所以擲出重複點數的概率必定不小於 $\frac{1}{6}$，而且只有公平骰子才令等號成立。它當然不可能小於 $\frac{1}{36}$。 ■

回顧

- **反證法（proof by contradiction）。**我們假設所有總和等可能，然後推出不可能成立的結論。
- **只用所需的條件。**十一條等式之中，三條已經足夠。選對等式正是關鍵的一步。
- **以不等式證明不可能。**第 (i) 部分的算術平均—幾何平均不等式，以及第 (ii) 部分的「平方和非負」，都一舉排除了一整類的可能性。

背後的數學

用多項式來證明

以下是一個完全不同的證明，用的是 Section 9.6 的生成函數（generating function）。設 $Q(x) = q_1x + \cdots + q_6x^6$ 及 $R(x) = r_1x + \cdots + r_6x^6$。假如所有總和的概率都是 $\frac{1}{11}$，則

$$Q(x)R(x) = \frac{x^2 + x^3 + \cdots + x^{12}}{11} = \frac{x^2}{11}\big(1 + x + \cdots + x^{10}\big).$$

現在 $Q(x)/x$ 是一個 5 次實係數多項式（polynomial），而每個奇數次的實多項式都有一個實根（它的圖像由 $-\infty$ 走到 $+\infty$，或者反過來）。所以左邊有一個非零的實根。但 $1 + x + \cdots + x^{10} = \frac{x^{11}-1}{x-1}$，它的根是 1 以外的 11 次單位根（roots of unity），當中沒有一個是實數。矛盾。同樣的論證表明，只要 $n - 1$ 是奇數，即 n 是偶數，無論怎樣在兩顆 n 面骰子上做手腳，都不能令各個總和等可能。（當 n 是奇數時，多項式的論證需要多花些工夫，但上面的算術平均—幾何平均證明對每個 $n \geq 2$ 都適用。）

碰撞概率

第 (ii) 部分中的量 $\sum q_j^2$，即兩次獨立抽取結果相同的機會，稱為一個分佈的*碰撞概率*（collision probability）。第 (ii) 部分表明，均勻分佈（uniform distribution）的碰撞概率最小。它量度一個分佈有多集中：在生態學中，$\sum q_j^2$ 是辛普森多樣性指數（Simpson's index of diversity），其中 q_j 是物種 j 在某個生境中所佔的比例；在經濟學中，它是量度市場集中度的赫芬達爾指數（Herfindahl index）；在密碼學中，它量度一個密碼分佈有多容易被猜中。它也解釋了 Section 9.9 結尾的那句話：各個生日的概率不相等，只會令生日相同的情況更有可能出現。

不可能性的證明

數學中一些最著名的結果，說的正是某件事辦不到：沒有分數的平方等於 2；沒有尺規作圖法能三等分任意角；沒有根式公式能解所有五次方程；沒有算法能判定每一個數學命題。每一個這樣的結果都改變了所屬的學科，因為要證明*沒有*任何方法行得通，就必須了解所有方法構成的整個空間。灌鉛骰子問題是一個縮影：要排除每一種做手腳的方式，我們需要一個每種方式都必須遵守的不變性質（算術平均—幾何平均不等式）。

思維技巧

以反證法和不等式證明不可能

要證明沒有任何參數的選取行得通，先假設有一組行得通，找出互相抵觸的那幾個條件，再用不等式證明它們不可能同時成立。

在大學裏，反證法配合不等式是證明某些對象不存在的標準方法，見於分析、組合數學和最優化；在這些領域中，算術平均—幾何平均不等式和柯西—施瓦茲不等式（Cauchy–Schwarz inequality）這類不等式為可達到的結果設下了界限。

延伸練習

1 **柯西—施瓦茲不等式。**利用柯西—施瓦茲不等式 $\left(\sum a_j b_j\right)^2 \le \sum a_j^2 \sum b_j^2$，取 $a_j = q_j$ 及 $b_j = 1$，為第 (ii) 部分給出一行的證明。

2 **兩顆不同的骰子。**如果把兩顆不同的灌鉛骰子各擲一次，它們點數相同的概率能否小於 $\frac{1}{6}$？能否等於 0？

3 **面上數字不同時的均勻總和。**證明如果兩顆骰子各面的數字另作安排，它們可以給出等可能的總和：試取一顆骰子的數字為 $0,1,2,3,4,5$，另一顆為 $0,6,12,18,24,30$。會出現哪些總和？為甚麼這與題目並不矛盾？

4 **三顆骰子。**三顆各面為 1 至 6 的六面灌鉛骰子，能否令 3 至 18 的所有總和等可能？（提示：利用總和 3、8、13 和 18，以及六個數的算術平均—幾何平均不等式。為甚麼多項式證明這次行不通？）

延伸閱讀

- **J. Michael Steele, *The Cauchy–Schwarz Master Class***（Cambridge University Press）。一本出色的不等式入門書，由算術平均—幾何平均不等式講起。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***（Oxford University Press）。講述生成函數，以及它們在上述這類證明中的應用。
- **Ian Stewart, *Galois Theory***（CRC Press），開首數章。最著名的不可能性證明：五次方程。

我們學到了甚麼

本章開首說過，計數是最古老的數學，而要數得巧妙，至今仍是最難的事情之一。我們提出了三個問題：怎樣不逐一列舉而數出答案？甚麼時候某件事必然發生？計數與概率（probability）怎樣互相幫助？以下是本章十三道題目所揭示的答案。

1. 怎樣不逐一列舉而數出答案？

本章逐步建立了一套工具。

- **遞歸關係（recurrences）。**在平面上多加一條直線（Section 9.1），或在圍欄上多加一根柱子（Section 9.3），數目都只會以簡單的方式改變，由此得出可以求解的遞歸關係。只要多記錄一點資料，即最後一根柱子的顏色，這些遞歸關係便能自成一體。
- **生成函數（generating functions）。**把數列（sequence）包裝成冪級數（power series），遞歸關係便化為代數運算（Section 9.4），既解釋了為何 $1/89 = 0.01123\ldots$，也把數算組合變成多項式相乘：Section 9.5 的硬幣、Section 9.6 的骰子，以及 Section 9.7 的等冪和，其中 $x = 1$ 處的重根（repeated root）包辦了一切。
- **雙射（bijections）。**反射原理（reflection principle）（Section 9.10）把不合格的隊列與不受限制的路徑一一配對，從而數出前者的數目；而取隨機集合的補集（complement），則把子集變成互不相交的集合對（Section 9.12）。
- **容斥原理（inclusion–exclusion）。**先多算、後修正，便求得錯排（derangements）的數目（Section 9.8）。
- **重複計數（double counting）。**以歐拉公式（Euler's formula）數算區域（Section 9.1），以及按大小數算子集對（Section 9.12），都是循不同途徑得出相同答案，並順道證明了一些恆等式。

2. 甚麼時候某件事必然發生？

有好幾節顯示，單憑規模已足以迫出結構。根據鴿巢原理（pigeonhole principle），六個人當中必有三人互相認識，或有三人互不相識（Section 9.2），但五個人則不足夠。二十三人當中有人同一天生日的機會已過半（Section 9.9），因為配對的數目隨人數的平方增長。一個大集合的隨機子集幾乎必定互相重疊（Section 9.12）。還有一些事情，無論怎樣嘗試都不可能發生：不論怎樣為兩顆骰子灌鉛，都無法令每個點數總和出現的機會均

等（Section 9.13）；而除了 Sicherman 的編號方式外，再沒有其他骰子編號能重現普通骰子的點數總和分佈（Section 9.6）。

3. 計數與概率怎樣互相幫助？

在本章的後半部分，概率是靠計數求得的：錯排用容斥原理，生日問題用連乘積，選票問題（ballot problem）用反射。但反過來，概率也幫了計數的忙。期望值的線性性質（linearity of expectation）給出了集換券收集問題（coupon collector）的 nH_n（Section 9.11），也給出了洗牌後配對數目的期望值，完全毋須計數。把隨機子集看成擲 n 次硬幣（Section 9.12），3^n 和 $(m+1)^n$ 這類數目便成了一行就能完成的計算。而艾狄胥（Erdős）的概率方法（probabilistic method）（Section 9.2）則透過證明隨機着色有正概率符合要求，證明了理想的着色確實存在。有兩個常數憑空出現：e 出現在錯排概率的極限中，而歐拉常數 γ 則出現在集齊一整套所需的代價中。

4. 計數的工具箱

方法	用於何處
多加一個時有何改變	直線（9.1）；圍欄柱子（9.3）
鴿巢原理	拉姆齊（Ramsey）（9.2）
耦合遞歸關係	圍欄柱子（9.3）
生成函數	9.4、9.5、9.6、9.7、9.13
從極端情況入手分類討論	奇異骰子（9.6）；灌鉛骰子（9.13）
容斥原理	錯排（9.8）
補集與指數近似	生日（9.9）
反射原理	選票問題（9.10）
期望值的線性性質	集換券（9.11）；配對（9.8）
每次處理一個元素	隨機子集（9.12）
以算術—幾何平均不等式（AM–GM）導出矛盾	灌鉛骰子（9.13）

下一步

- 錯排的泊松極限（Poisson limit）會在 Section 12.11 再次出現，而隨機變量（random variables）的生成函數則推動 Section 12.15 對家族樹的研究。
- 耦合遞歸關係在 Section 12.16 會變成馬可夫鏈（Markov chains）。
- 集換券收集問題中的岡貝爾分佈（Gumbel distribution），在 Section 12.25 用來描述眾多隨機量中的最大者。

如果你想在入讀大學前再進一步

- **Miklós Bóna, *A Walk Through Combinatorics***：一本平易近人的入門教材，幾乎涵蓋本章的每一個概念。
- **Herbert Wilf, *generatingfunctionology***：生成函數的經典著作，可於網上免費閱讀。
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics***：求和、遞歸關係與生成函數，筆調風趣，習題豐富。
- **Geoffrey Grimmett & Dominic Welsh, *Probability: An Introduction***：大學一年級概率論中與計數有關的部分。

懷特黑德（Whitehead）曾稱組合數學（combinatorics）為「拓撲學的貧民窟」。本章嘗試說明，為何今天再沒有人會這樣說。每一個巧妙的計數背後，都有一個影響遠超計數本身的想法：一個生成函數、一個雙射、一個遞歸關係、一個不等式。這些想法把計數與代數、分析和概率連繫起來；學會尋找這些想法，而不是再多找一個竅門，正是令計數成為數學的關鍵。

延伸練習答案

以下只是簡短的答案和提示，並非完整解答：這裏只給出結果和關鍵步驟，方便你核對自己的做法。請先認真做過每道練習，然後才翻看。

9.1 直線分割平面

1. 一個圓把平面分成 2 個區域。第 n 個圓與其餘各圓相交於 $2(n-1)$ 點，這些點把它分成 $2(n-1)$ 段弧，每段弧把一個區域一分為二。所以 $R_n = R_{n-1} + 2(n-1)$，由此得 $n^2 - n + 2$。

2. 每條直線所增加的區域數目，是 1 加上該直線上的交點數目，所以總數是 $1 + n +$ (交點數目)。那 k 條平行線少了 $\binom{k}{2}$ 個交點：$1 + n + \binom{n}{2} - \binom{k}{2}$。

3. 第 n 個平面被其餘平面截出 $n-1$ 條處於一般位置（general position）的直線，這些直線把它分成 $\frac{1}{2}(n^2 - n + 2)$ 塊，每一塊把空間中的一個區域一分為二。把這個數加到 $\frac{1}{6}\big((n-1)^3 + 5(n-1) + 6\big)$ 上，便得 $\frac{1}{6}(n^3 + 5n + 6)$。

4. 當 $n = 6$ 時有 31 個區域，而不是 32 個。在圓內，這個圖有 $V = n + \binom{n}{4}$ 個頂點（vertex）（每個交點由四個點決定）和 $E = n + \binom{n}{2} + 2\binom{n}{4}$ 條邊（edge）（圓上的各段弧，加上各條弦，而每個交點再多添兩段）。歐拉（Euler）公式 $V - E + F = 2$ 把圓外的區域也算在內，所以區域的數目是 $F - 1 = 1 + \binom{n}{2} + \binom{n}{4}$。

9.2 小型版的拉姆齊定理

1. 雙色角（bichromatic angle）最多有 $6 \times 6 = 36$ 個。每個非單色的三角形恰好含有兩個雙色角，所以這類三角形最多有 18 個。由於 $\binom{6}{3} = 20$，最少有 2 個三角形是單色的。

2. 每個點有 16 條邊，所以其中最少有 6 條同色，設為紅色，並設它們連到 $P_1, \ldots, P_6$ 各點。若 P_i 之間有任何一條邊是紅色，便有一個紅色三角形。若沒有，則各 P_i 構成一個以兩種顏色著色的 K_6，其中必有單色三角形。

3. 點 v 有 9 條邊，所以要麼最少有 4 條是金色，要麼最少有 6 條是銀色。若有 4 個金色鄰點：它們之間只要有一條金色邊，便與 v 構成金色三角形，否則它們構成一個銀色的 K_4。若有 6 個銀色鄰點：它們之中必有一個單色三角形；若是金色，證明完成，若是銀色，它便與 v 構成一個銀色的 K_4。

4. 給每一項標上 (i, d)：分別是以該項結尾的最長遞增子序列（subsequence）和最長遞減子序列的長度。對任意兩項，較後的一項不是 i 較大，就是 d 較大，所以各個標記互不相同。假如沒有長度為 3 的子序列，那麼 5 項就只有 4 個可能的標記。兩者皆無的例子：$2, 1, 4, 3$。

9.3 為籬笆樁着色

1. 與之前一樣，$r_{n+1} = s_n$（第 n 根柱不是紅色，所以第 $n+1$ 根柱可以是紅色），而 $r_n + s_n = (k-1)^{n-1}$，因為第一根之後的每根柱都有 $k-1$ 種選擇。所以 $r_{n+1} + r_n = (k-1)^{n-1}$，而 $r_n = \frac{(k-1)^{n-1}+(k-1)(-1)^{n-1}}{k}$。至於圍成一圈的情況：第 1 根柱的顏色有 k 種選擇，然後是一排 $n+1$ 根柱，最後一根與第一根同色，由此得 $kr_{n+1} = (k-1)^n + (-1)^n(k-1)$。

2. $2, 3, 5, 8, \ldots$：設共有 a_n 種方法，以白色結尾的一排可接在任何一排 $n-1$ 根柱之後，而以紅色結尾的一排則接在以白色結尾的一排 $n-2$ 根柱之後，所以 $a_n = a_{n-1} + a_{n-2}$。答案是 F_{n+2}。

3. 中心有 4 種選擇，然後是 n-圈的 3-著色：$4\big(2^n + 2(-1)^n\big)$。

4. 當 $n = 4$ 時：有 18 種著色。轉四分之一圈不會保持任何一種著色不變，轉半圈則保持其中 6 種（即相對的柱同色的那些）不變，所以有 $\frac{1}{4}(18 + 0 + 6 + 0) = 6$ 條項鏈。當 $n = 5$ 時：有 30 種著色，沒有任何旋轉能保持其中一種不變，所以是 6。

9.4 把數列看成函數

1. 餘下各項之和最多是 $F_7 \sum_{r\geq 7} 2^{r-7}10^{-r} = \frac{13\times 10^{-7}}{0.8} < 2 \times 10^{-6}$。由於 $\sum F_r 10^{-r} = \frac{10}{89}$，而首六項給出 0.112358，所以 $\frac{10}{89}$ 介乎 0.112358 與 0.112360 之間，於是 $\frac{1}{89} = 0.011235\ldots$。若改以兩位數為一組，則 $\sum F_r 100^{-r-1} = \frac{1}{9899}$。

2. 把遞推關係（recurrence）求和，得 $\mathrm{f}(x)(1 - 7x + 10x^2) = 2 - 7x$，而 $\frac{2-7x}{(1-2x)(1-5x)}$ 可如題目所述分拆。至於 b_n：由 $10p + 5q = 40$ 和 $40p + 10q = 100$ 得 $p = 1$，$q = 6$；$t^2 = t + 6$ 的根是 3 和 -2，而 $b_n = 4 \cdot 3^n + (-2)^n$。

3. $\sum nx^n = \frac{x}{(1-x)^2}$；在 $x = \frac{1}{100}$ 時它等於 $\frac{100}{9801}$，所以 $\frac{1}{9801} = \sum n \cdot 100^{-n-1}$。這個規律在 98 這一組出錯：100 那一組把 1 進位到 99 那一組，後者又再進位到 98 那一組，於是展開式是 $\ldots 96\,97\,99\,00\,01\ldots$。

4. 乘以 $\frac{1}{1-x} = 1 + x + x^2 + \cdots$ 之後，x^n 的係數變成 $u_0 + \cdots + u_n$。平方數的生成函數（generating function）是 $\frac{x(1+x)}{(1-x)^3}$，所以其部分和的生成函數是 $\frac{x+x^2}{(1-x)^4}$，而 x^n 的係數是 $\binom{n+2}{3} + \binom{n+1}{3} = \frac{n(n+1)(2n+1)}{6}$。

9.5 湊零錢

1. x^{24} 的係數就是 $6a+6b+3c=24$ 的解的數目：對由 0 到 4 的每個 $s=a+b$，各有 $s+1$ 種選擇，合共 15 種。一般而言，x^n 的係數是 0，除非 $n=3m$，這時它等於 $\frac{1}{2}(k+1)(k+2)$，其中 $k=\lfloor m/2\rfloor$。加上額外的因式後，係數變成部分和：x^{24}、x^{25} 和 x^{66} 的係數分別是 55、55 和 650。它們所數的，是用 1p 和 3p 硬幣以及兩種不同的 6p 硬幣付款的方法數目。

2. 每個部分 k 要麼用一次，要麼完全不用。這個乘積可以逐項相消：$\prod \frac{1-x^{2k}}{1-x^k}$ 把分母中指數為偶數的因式全部約去，剩下 $\prod \frac{1}{1-x^{\text{奇數}}}$，也就是分拆（partition）成奇數部分的生成函數。當 $n=6$ 時：6、$5+1$、$4+2$、$3+2+1$，以及 $5+1$、$3+3$、$3+1+1+1$、$1+1+1+1+1+1$：兩類各有四個。

3. 當 $n=0,1,2,\ldots$ 時，數值依次是 $1,1,2,2,3,4,5,6,7,8,10,\ldots$，而 $\frac{(n+4)^2}{20}$ 與它們的差距總不超過 $\frac{1}{2}$（已驗算至 $n=60$；證明則要用到 $\frac{1}{(1-x)(1-x^2)(1-x^5)}$ 的部分分式（partial fractions））。

4. 最後一枚硬幣是 1p 或 2p，所以 $c_n=c_{n-1}+c_{n-2}$：這就是斐波那契（Fibonacci）數，$c_n=F_{n+1}$。

9.6 奇怪的骰子

1. 每顆骰子都是 x 乘以 $1+x$、$1+x$、$1+x^2$、$1+x^2$ 這些因式中若干個的乘積，而且在 $x=1$ 時的值是 4，所以它要取其中兩個因式。除了標準的一對之外，唯一的選擇是 $x(1+x)^2$ 和 $x(1+x^2)^2$：各面是 $1,2,2,3$ 和 $1,3,3,5$。

2. 乘積是 $2x^3+x^4+x^5+4x^6+3x^7+8x^8+6x^9+5x^{10}+6x^{11}$，與表中數字吻合。

3. 可以：取 Sicherman 骰子對 $1,2,2,3,3,4$ 和 $1,3,4,5,6,8$，再加一顆標準骰子。

4. $(x+x^2)(x+\cdots+x^8)=x^2(1+x)^2(1+x^2)(1+x^4)$。硬幣需要一個值為 2 的因式：除了標準的選擇 $x(1+x)$ 之外，它可以是 $x(1+x^2)$（兩面是 $1,3$，配上八面體（octahedron）骰子 $1,2,2,3,5,6,6,7$），也可以是 $x(1+x^4)$（兩面是 $1,5$，配上 $1,2,2,3,3,4,4,5$）。

9.7 相等的冪和

1. $\sum(c+b_k)^3-\sum(c+a_k)^3 = 3c^2\left(\sum b_k-\sum a_k\right)+3c\left(\sum b_k^2-\sum a_k^2\right)+\sum b_k^3-\sum a_k^3$，而首兩個括號都等於零。由 $\{1,4,6,7\}$、$\{2,3,5,8\}$ 出發，取 $c=8$：得 $\{1,4,6,7,10,11,13,16\}$ 和 $\{2,3,5,8,9,12,14,15\}$。

2. A 和 $B\cup\{0\}$ 都有 6 個元素，和為 33，平方和為 253。於是 $\sum_{a\in A}(n+a)^2=\sum_{b\in B\cup\{0\}}(n+b)^2$，這樣便把 $\{n^2,\ldots,(n+11)^2\}$ 分成兩組。

3. 用二項式定理（binomial theorem）展開 $(x+c)^r$：式中只涉及不高於 r 次的冪，而這些冪的和都相等。題目中的兩個集合，在第二個加入 0 之後，合起來就是 $\{0,\dots,15\}$，再作平移便得任意 16 個連續整數。

4. 按 $ABABABAB$ 的次序，A 約得 1.676，B 約得 1.042。按 $ABBABAAB$ 的次序，A 約得 1.560，B 約得 1.158：仍不相等，但已接近得多。

9.8 沒有人拿到自己的恤衫

1. $\mathrm{E}(C_1)=\frac{1}{m}$，$\mathrm{Var}(C_1)=\frac{1}{m}-\frac{1}{m^2}$，而 $\mathrm{Cov}(C_1,C_2)=\frac{1}{m(m-1)}-\frac{1}{m^2}=\frac{1}{m^2(m-1)}$。於是 $\mathrm{E}(N)=1$，而 $\mathrm{Var}(N)=m\bigl(\frac{1}{m}-\frac{1}{m^2}\bigr)+m(m-1)\frac{1}{m^2(m-1)}=1$。當 $m=4$ 時：$\mathrm{P}(N=0)=\frac{9}{24}=0.375$，相比之下 $\mathrm{e}^{-1}\approx 0.368$。

2. 第 1 張牌落在某個位置 i（有 $n-1$ 種選擇）。若第 i 張牌落在位置 1，餘下的就是 $n-2$ 張牌的錯排（derangement）；若不是，則把位置 1 當作第 i 張牌的禁位，得到 $n-1$ 張牌的錯排。驗算：$D_3=2(1+0)=2$，$D_4=3(2+1)=9$。

3. 先選出 k 張留在原位的牌，再把其餘的牌錯排：$\binom{n}{k}\frac{D_{n-k}}{n!}=\frac{1}{k!}\cdot\frac{D_{n-k}}{(n-k)!}=\frac{d_{n-k}}{k!}$，它趨向 $\frac{\mathrm{e}^{-1}}{k!}$。

4. $30-(15+10+6)+(5+3+2)-1=8$，它等於 $30\cdot\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5}$：容斥原理（inclusion–exclusion）的和式可以因式分解（這就是歐拉的 $\varphi(30)$）。

9.9 生日問題

1. $\mathrm{P}(X=r)=\bigl(1-\frac{1}{n}\bigr)\cdots\bigl(1-\frac{r-2}{n}\bigr)\frac{r-1}{n}$，而這些概率之和為 1。至於期望值（expectation），$\mathrm{P}(Y=r)$ 在 $\mathrm{P}(Y\ge 1),\dots,\mathrm{P}(Y\ge r)$ 中各被計算一次。記 $Q_k=\mathrm{P}(X>k)=\prod_{j=1}^{k-1}\bigl(1-\frac{j}{n}\bigr)$。於是 $\mathrm{E}(X)=\sum(k+1)\frac{k}{n}Q_k$，同時 $\mathrm{E}(X)=1+\sum Q_k$；利用 $\sum\frac{k}{n}Q_k=1$ 並將兩式相減，得 $\sum Q_k\bigl(1-\frac{k^2}{n}\bigr)=0$。

2. 由 $\binom{k}{3}\approx 365^2\ln 2$ 得 $k\approx 82$；這個近似忽略了各個三人組之間的相依關係，真正的答案是 88。

3. 14 人。（每一對人「接近」的概率是 $\frac{3}{365}$，而由 $\binom{k}{2}\frac{3}{365}\approx\ln 2$ 得 $k\approx 14$；精確計算的結果是：14 人時概率約為 0.537，13 人時約為 0.483。）

4. 兩人同在第 i 天生日的概率是 p_i^2；把它們加起來。根據柯西–施瓦茲不等式（Cauchy–Schwarz），$1=\bigl(\sum p_i\bigr)^2\le 365\sum p_i^2$，等號僅在所有 p_i 都相等時成立。

9.10 總能找贖

1. 第一個人必須持有一枚 £1 硬幣。若按其後幾個人的情況取條件，便要分成許多情形（視乎每位 £2 顧客之後你手上有哪些硬幣），每種情形最後都歸結到第 (i) 和第

(ii) 部分的答案；把它們合起來，得 $\frac{m-2}{m+1}$。用反射原理（reflection principle）則一行便可完成。

2. 第一票必須投給 A。在以 A 開始的點票過程之中，壞的那些（即曾經出現平手的）可藉着反射第一次平手之前的部分，與以 B 開始的點票過程一一對應。所以好的點票過程的數目是 $\binom{a+b-1}{a-1} - \binom{a+b-1}{a}$，除以 $\binom{a+b}{a}$ 便得 $\frac{a-b}{a+b}$。

3. (((()))、((()))、((()))()、()((()))、()()()。一個雙射（bijection）：第一對括號把序列分成括號之內的部分和括號之後的部分；相應地，以某條固定的邊為底的三角形把多邊形分成兩個較小的多邊形。對兩部分分別遞歸處理。

4. 每位 £5 顧客需要找回四枚 £1 硬幣，所以路徑每步上升 1 或下降 4。路徑可以越過界線而不碰到它，於是反射再不能把壞的路徑與一個簡單的集合配對起來。

9.11 集齊一套

1. 各段等待時間的期望值之和是 $\frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{n/2+1} = n(H_n - H_{n/2}) \approx n(\ln n - \ln \frac{n}{2}) = n \ln 2$。

2. 每張卡缺少的概率是 $(1 - \frac{1}{n})^k$；對 n 張卡求和。要達到 95%，需 $(1 - \frac{1}{n})^k \approx \mathrm{e}^{-k/n} = 0.05$，所以 $k \approx n \ln 20 \approx 3n$。

3. 買了 t 包之後，某一張指定的卡出現少於兩次的概率約為 $\mathrm{e}^{-t/n}(1 + \frac{t}{n})$。當這類卡的期望數目 $n\mathrm{e}^{-t/n}(1 + \frac{t}{n})$ 降至約 1 時，兩個孩子便大致都集齊了；由此得 $\frac{t}{n} \approx \ln n + \ln \ln n$，遠小於 $2 \ln n$。

4. $\frac{1}{n+1} < \ln \frac{n+1}{n}$（$\frac{1}{x}$ 在 $[n, n+1]$ 上的曲線下面積），所以 $H_n - \ln n$ 遞減。由凸性（convexity）可知梯形（trapezium）高估了面積：$\ln \frac{n+1}{n} < \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right)$，所以 $H_n - \ln n - \frac{1}{2n}$ 遞增。取 $n = 1$，得 $\gamma > \frac{1}{2}$；取 $n = 4$，得 $\gamma < H_4 - \ln 4 \approx 0.697$。

9.12 隨機子集

1. 每個元素必須屬於最少一個集合：$\left(\frac{3}{4}\right)^n$，一般情況則是 $\left(1 - 2^{-m}\right)^n$。

2. $|S \setminus A_2|$ 與 $|A_2|$ 有相同的分佈（distribution），並且與 A_1 獨立（independent），所以所求概率就是 $|A_1| + |S \setminus A_2| = n$ 的概率。$(A_1, S \setminus A_2)$ 這一對集合對應於一個有 $2n$ 個元素的集合的隨機子集，其大小為 n 的概率是 $\binom{2n}{n}/4^n$。由斯特靈（Stirling）公式（或沃利斯（Wallis）乘積，見 Section 4.4）得 $\approx \frac{1}{\sqrt{\pi n}}$。

3. 每個元素放進 A、放進 B，或兩者都不放：3^n。滿足 $A = B$ 的有序對只有 $(\varnothing, \varnothing)$，所以共有 $\frac{3^n-1}{2} + 1 = \frac{3^n+1}{2}$ 個無序對。

4. 當 $n = 3$ 時，是 3 個集合（全部單元素集，或全部二元素集）。當 $n = 4$ 時，是 6 個集合（全部二元素集）。一般情況是 $\binom{n}{\lfloor n/2 \rfloor}$（Sperner 定理，Sperner's theorem）。

9.13 做了手腳的骰子

1. $1 = \left(\sum q_j\right)^2 \le 6\sum q_j^2$，所以 $\mathrm{P}(\text{相同}) = \sum q_j^2 \ge \frac{1}{6} > \frac{1}{36}$。

2. 兩問的答案都是「可以」：所求概率是 $\sum q_j r_j$，而比方說，若一顆骰子永遠擲出 1，另一顆永遠擲出 2，它便等於 0。

3. 點數總和是 $0, 1, \ldots, 35$，每個恰好出現一次：是均勻分佈。題目要求兩顆骰子都標上 1 至 6，而這裏的骰子並非如此。

4. 不可以。記 $\mathrm{P}(3) = q_1 r_1 s_1$ 和 $\mathrm{P}(18) = q_6 r_6 s_6$，兩者都是 $\frac{1}{16}$。$\mathrm{P}(8) + \mathrm{P}(13)$ 包含 $q_1 r_1 s_6, \ldots, q_1 r_6 s_6$ 這六項（兩個 1 和一個 6，或兩個 6 和一個 1），它們的乘積是 $(\mathrm{P}(3)\mathrm{P}(18))^3$。由算術–幾何平均不等式（AM–GM）得 $\mathrm{P}(8) + \mathrm{P}(13) \ge 6 \cdot \frac{1}{16} > \frac{2}{16}$。原來那個多項式證明照搬過來並不成立，因為 $1 + x + \cdots + x^{15}$ *確實*有一個實根 -1。（不過它可以補救：三個實係數五次多項式需要三個實根，而 -1 只是單根。）

Beyond STEP

STEP 是大多數中學生一生中應考過最艱深的數學考試，也是最出色的考試之一：題目簡短、陌生而深刻，當中最精彩的，其實是喬裝了的大學數學小品。

本書精選最有趣的 STEP 試題，由 1987 年首次開考之前的樣本試卷（specimen papers），到 2025 年的試卷，並以每一道題目作為一扇門。每道題目都會仔細解答，沿途提出優秀學生會問的問題，然後帶你穿過這扇門：這個想法從何而來，到了大學會演變成甚麼，以及接下來該讀些甚麼。

本書為正在準備 STEP、有天分又有志投身數學的學生而寫。他們想要的不止是模擬試卷：他們想開始像數學系本科生那樣思考，並一探前方有甚麼正等着自己。

- 共 13 章，涵蓋純數學（pure mathematics）、力學（mechanics）、概率（probability）與統計（statistics）
- 136 道 STEP 歷屆試題，每題均附引導式解答
- 每道題目背後的大學數學
- 附答案的延伸練習，以及延伸閱讀

全書共 13 章，本冊為第 9 章。 全書各章及所有詳細解答均可免費閱讀、下載及分享，網址為

beyondstep.keithmcnulty.org

文字內容已貢獻予公有領域（CC0 1.0）。STEP 試題 © Cambridge University Press & Assessment。