機率與統計：從入門到精通

第 1 章　什麼是機率？

What Is Probability?

難度：入門

機率是「不確定性的數學」——一套讓我們能夠精確推理「無法確定之事」的語言。擲一枚硬幣、明天的天氣、一封電子郵件是否為垃圾郵件、句子中的下一個詞是什麼：這些都充滿不確定性，而機率讓我們得以**量化**這種不確定性，並且對它進行**計算**。本章奠定整個學科的地基：結果與事件的詞彙、支配所有機率運算的三條公理、計數的技藝，以及關於「機率究竟是什麼意思」的兩大學派。

學習目標

讀完本章後，你將能夠：

- 正確使用**樣本空間**、**結果**、**事件**等詞彙描述隨機現象，並以集合運算組合事件。
- 敘述 Kolmogorov 三大公理，並由公理推導出補集規則、單調性與排容原理（inclusion–exclusion）。
- 在等可能的情境下，用**排列**與**組合**計算機率，處理撲克牌、生日問題等經典例子。
- 分辨**頻率學派**與**貝氏學派**對機率的詮釋，並說明兩者如何同時貫穿現代機器學習。
- 用 Python / NumPy 以模擬驗證理論計算的結果。

1.1　隨機試驗、樣本空間與事件

機率理論的第一步，是替「不確定的事情」建立一套精確的詞彙。日常語言說「擲骰子可能出現很多情況」，數學語言則把「所有可能情況」整理成一個集合，並把「我們關心的情況」定義為它的子集合。

定義 1.1　隨機試驗、樣本空間與事件

隨機試驗（random experiment）是任何結果不確定的過程。**樣本空間**（sample space）Ω 是所有可能結果所成的集合。**事件**（event）A 是 Ω 的一個子集合——也就是我們關心的一組結果。

例 1.1　擲一顆骰子

擲一顆公正的六面骰。樣本空間為

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

事件「出現偶數」是 $A = \{2, 4, 6\}$；事件「大於 4」是 $B = \{5, 6\}$。事件之間可以用集合運算組合：

- $A \cup B = \{2, 4, 5, 6\}$：「$A$ 或 B 發生」（至少一個發生）；
- $A \cap B = \{6\}$：「A 且 B 同時發生」；
- $A^c = \{1, 3, 5\}$：「A 不發生」（補集）。

例 1.2　擲兩枚硬幣

擲兩枚硬幣，以 H 表示正面、T 表示反面，則

$$\Omega = \{\mathrm{HH},\ \mathrm{HT},\ \mathrm{TH},\ \mathrm{TT}\}, \qquad |\Omega| = 4.$$

注意 HT 與 TH 是**不同的結果**：第一枚正、第二枚反，與第一枚反、第二枚正，是兩回事。事件「恰好一個正面」為 $\{\mathrm{HT}, \mathrm{TH}\}$，機率是 $2/4 = 1/2$；若誤把樣本空間寫成「0 個正面、1 個正面、2 個正面」三種並假設等可能，就會得到錯誤答案 $1/3$。**寫對樣本空間，是所有機率計算的第一步。**

例 1.3　連續的樣本空間

公車每 10 分鐘一班，你隨機抵達站牌，等待時間 T 可以是 $[0, 10]$ 之間的任何實數，故

$$\Omega = [0, 10] \subset \mathbb{R}.$$

這是一個**連續**樣本空間，含有不可數無限多個結果。事件「等待不超過 3 分鐘」是子區間 $[0, 3]$，直觀上其機率為長度比 $3/10$。連續情形的嚴謹處理要等到第 3 章的機率密度函數，但「事件 = 子集合」的觀念完全相同。

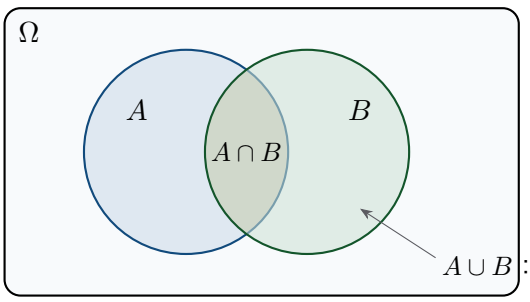


圖 1.1: 把 Ω 想成「所有可能發生之事」構成的空間，事件是從中劃出的區域。

直觀理解

把 Ω 想像成「一切可能發生之事」構成的空間，而事件是從中劃出的一塊區域。問「A 的機率是多少？」就是問「全部的可能性中，有多大比例落在 A 這塊區域裡？」機率的邏輯就是集合的邏輯——聯集、交集、補集——再加上一個度量「大小」的數字。熟悉集合運算的讀者會發現，De Morgan 定律

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c$$

在機率計算裡隨處可見：「兩者皆非」等於「並非至少其一」。

1.2 機率公理

現代機率論建立在三條規則之上，由 Kolmogorov 於 1933 年提出。其餘一切性質都由它們推導而得。

定義 1.2　Kolmogorov 公理

機率 $\mathbb{P}$ 為每個事件指定一個數字，並滿足：

1. **非負性**：對任何事件 A，$\mathbb{P}(A) \geq 0$。
2. **規範性**：$\mathbb{P}(\Omega) = 1$——「總有某件事會發生」。
3. **可加性**：若 $A_1, A_2, \ldots$ **互斥**（disjoint，兩兩交集為空），則

$$\mathbb{P}\Big(\bigcup_{i} A_i\Big) = \sum_{i} \mathbb{P}(A_i).$$

命題 1.3　由公理導出的基本性質

對任何事件 A, B：

1. **補集規則**：$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$；特別地 $\mathbb{P}(\varnothing) = 0$。
2. **單調性**：若 $A \subseteq B$，則 $\mathbb{P}(A) \leq \mathbb{P}(B)$；因此 $0 \leq \mathbb{P}(A) \leq 1$。
3. **排容原理**（inclusion–exclusion）：

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

證明　(1) A 與 A^c 互斥且 $A \cup A^c = \Omega$，由可加性與規範性得 $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1$。取 $A = \Omega$ 得 $\mathbb{P}(\varnothing) = 0$。

(2) 若 $A \subseteq B$，把 B 拆成互斥兩塊：$B = A \cup (B \cap A^c)$。由可加性 $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \geq \mathbb{P}(A)$，最後一步用了非負性。

(3) 把 $A \cup B$ 拆成三塊互斥區域：只在 A、只在 B、同時在兩者，即 $A \cup B = (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B)$。分別記三塊機率為 x, y, z，則 $\mathbb{P}(A \cup B) = x + y + z$，而 $\mathbb{P}(A) = x + z$、

$\mathbb{P}(B) = y + z$，故 $\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = (x+z) + (y+z) - z = x + y + z$。 ■

常見陷阱

排容原理正是初學者最常跌倒的地方：**只有事件互斥時，機率才能直接相加**。若事件有重疊，必須扣掉被重複計算的交集。忘記這件事是最常見的機率錯誤，而且它會以更隱蔽的形式一再出現——例如在建模時假設特徵彼此獨立（實際上並不獨立），本質上犯的是同一類「忽略重疊 / 關聯」的錯。

例 1.4　排容原理：紅心或 A

從一副標準 52 張撲克牌中隨機抽一張，求抽到「紅心或 A」的機率。

解　設 $A =$「紅心」、$B =$「A（Ace）」。則 $\mathbb{P}(A) = \frac{13}{52}$，$\mathbb{P}(B) = \frac{4}{52}$，而兩者的交集是「紅心 A」這一張牌，$\mathbb{P}(A \cap B) = \frac{1}{52}$。由排容原理

$$\mathbb{P}(A \cup B) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13} \approx 0.308.$$

若直接相加 $\frac{13}{52} + \frac{4}{52} = \frac{17}{52}$，就把紅心 A 算了兩次。

例 1.5　三事件的排容原理

三個事件的版本為

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

某班學生中，60% 修數學、45% 修物理、40% 修化學；25% 同時修數學與物理、20% 同時修數學與化學、15% 同時修物理與化學；10% 三科全修。隨機抽一位學生，至少修一科的機率為

$$0.60 + 0.45 + 0.40 - 0.25 - 0.20 - 0.15 + 0.10 = 0.95,$$

因此三科都沒修的機率是 $1 - 0.95 = 0.05$。「加總—扣重疊—補回扣過頭的部分」，正是排容原理的節奏。

1.3　等可能結果與計數

當有限樣本空間中的 N 個結果**等可能**時，機率化約為計數：

$$\mathbb{P}(A) = \frac{A \text{ 中的結果數}}{\text{總結果數}} = \frac{|A|}{|\Omega|}.$$

所以我們必須學會「數得準」。一切計數的起點是**乘法原理**：若一件事分兩步完成，第一步有 m 種做法、第二步有 n 種做法，則總共有 $m \times n$ 種做法。由此得到兩大主力工具：

定義 1.4　排列與組合

從 n 個相異物件中取 k 個：

- **排列**（permutation，順序重要）：

$$P(n,k) = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1).$$

- **組合**（combination，順序不重要）：

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

兩者相差 $k!$ 倍，因為每一組 k 個物件都有 $k!$ 種排序方式。

例 1.6　排列 vs. 組合

8 位選手參加賽跑。

解　**(a) 前三名（金、銀、銅牌）有幾種可能？**順序重要，用排列：

$$P(8,3) = 8 \times 7 \times 6 = 336 \text{ 種}.$$

(b) 選 3 人進入決賽有幾種選法？順序不重要，用組合：

$$\binom{8}{3} = \frac{8!}{3!\,5!} = \frac{8\times7\times6}{3\times2\times1} = 56 \text{ 種}.$$

$336 = 56 \times 3!$：每一組決賽名單對應 $3! = 6$ 種頒獎排序，正是兩者的差別。

例 1.7　五張撲克牌與同花

一副標準撲克牌共 52 張，發 5 張牌的可能手牌數為

$$\binom{52}{5} = 2{,}598{,}960,$$

每一種特定手牌的機率就是它的倒數。進一步問：拿到**同花**（flush，五張同一花色，含同花順）的機率是多少？

解　分兩步計數：先選花色（4 種），再從該花色的 13 張中選 5 張：

$$|A| = \binom{4}{1}\binom{13}{5} = 4 \times 1287 = 5148, \qquad \mathbb{P}(A) = \frac{5148}{2{,}598{,}960} \approx 0.00198,$$

大約 0.2%，即平均每 505 手才出現一次。若要排除同花順（straight flush，同花且連號），再扣掉每花色 10 種、共 40 種連號組合：$(5148 - 40)/2{,}598{,}960 \approx 0.00197$。

例 1.8　生日問題

一個房間裡有 n 個人（假設每人生日獨立且均勻分布在 365 天中）。至少有兩人同一天生日的機率是多少？$n = 23$ 時答案會讓多數人吃驚。

解　直接算「至少兩人同生日」很難（情況太多），改算補集「**所有人生日互不相同**」。把 n 個人依序排開：第一人可以是任何一天（365/365），第二人要避開第一人（364/365），第三人要避開前兩人（363/365），依此類推：

$$\mathbb{P}(\text{全不相同}) = \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{365-n+1}{365} = \prod_{k=0}^{n-1} \frac{365-k}{365}.$$

因此

$$\mathbb{P}(\text{至少兩人同生日}) = 1 - \prod_{k=0}^{n-1} \frac{365-k}{365}.$$

代入 $n = 23$ 得 $1 - 0.4927 \approx \mathbf{0.507}$：只要 23 人，同生日的機率就超過一半！$n = 50$ 時已高達 0.970。直覺會失準，是因為我們該數的不是「人數」，而是「**人與人的配對數**」：23 人有 $\binom{23}{2} = 253$ 對，每一對都是一次「撞生日」的機會。

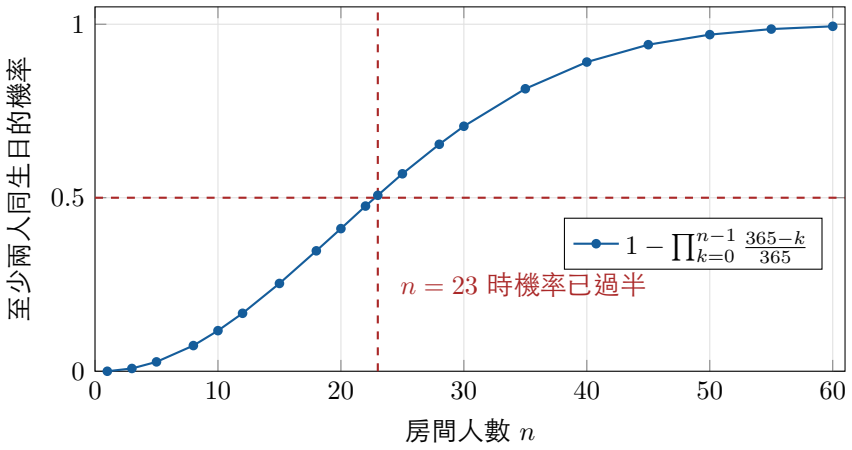


圖 1.2: 生日問題：機率隨人數增加的速度遠超直覺，因為配對數 $\binom{n}{2}$ 以平方成長。

直觀理解

組合數 $\binom{n}{k}$ 回答「從 n 個東西中挑出 k 個、不管順序，有幾種挑法」。它的舞台遠不止牌桌：**二項分布**（第 4 章）用它來數「n 次試驗中恰好 k 次成功」有幾種發生方式；任何「對子集合求和」的計算——包括模型計數與集成方法（ensemble）中的組合項——都會出現同樣的係數。把 $\binom{n}{k}$ 練熟，第 4 章會輕鬆許多。

1.4　機率是什麼意思？兩大詮釋

公理告訴我們機率**如何運算**，卻沒有告訴我們機率**是什麼意思**。兩個學派給出不同回答——而這條分界線直直穿過機器學習的心臟。

核心觀念

頻率學派（frequentist）：機率是**長期頻率**。「$\mathbb{P}(\text{正面}) = 0.5$」的意思是：重複拋擲很多次，正面的比例會趨近 0.5。機率是**世界本身的性質**，靠重複試驗顯露出來。

貝氏學派（Bayesian）：機率是**信念的程度**（degree of belief）。「$\mathbb{P}(\text{明天下雨}) = 0.3$」是在現有資訊下對明天的一致性信心陳述——即使「明天」只會發生一次。機率是**觀察者知識的性質**，隨著證據到來而更新。

直觀理解

兩種觀點沒有誰「正確」，它們是不同的工具。頻率學派的推理支撐假設檢定與信賴區間（第 8 章）；貝氏學派的推理支撐後驗更新與多數現代機率式機器學習（第 9 章）。深度學習其實悄悄地**兩者並用**：它用（頻率學派的）最大概似估計來訓練參數，卻把 softmax 的輸出解讀為（帶貝氏色彩的）類別信念。你將學會流利使用這兩種方言。

例 1.9　兩種詮釋的實戰對照

判斷下列敘述各自屬於哪種詮釋較自然：

(a)「這顆骰子擲出 6 的機率是 1/6。」

(b)「這位病人罹患此疾病的機率是 70%。」

(c)「這封郵件是垃圾郵件的機率是 0.92。」

解　(a) 頻率詮釋最自然：骰子可以重複擲，1/6 是長期頻率的預測，也可由對稱性（等可能）推得。(b) 貝氏詮釋較自然：「這位病人」只有一位、生病與否已成事實，機率描述的是**醫師在現有證據下的信念**，會隨檢驗結果更新。(c) 機器學習分類器的輸出：這封郵件也只有一封，0.92 是模型的信念程度（貝氏味）；但模型的參數是靠大量歷史郵件以頻率學派的方法（最大概似）訓練出來的。一句 0.92 裡，兩個學派各占一半。

這如何驅動機器學習／ AI

模型做出的每一個預測都是一個機率。分類器並不輸出「貓」，而是輸出一個分布：$\mathbb{P}(\text{貓}) = 0.83$、$\mathbb{P}(\text{狗}) = 0.12$、……，並且**恰好**服從本章的公理（每項非負、總和為一）。這裡的樣本空間就是類別集合；softmax 層本質上是一台「把任意實數向量變成合法機率分布」的機器。生成式模型更進一步，直接在「所有可能的圖片」或「所有可能的句子」上放一個機率分布。把地基打對——什麼是事件、機率何時可以相加、那個數字是什麼意思——正是「能夠健全地推理不確定性的模型」與「自欺欺人的模型」之間的分水嶺。

1.5 Python 實作

理論之後，動手驗證。下面的程式碼示範：頻率觀點下長期頻率趨近機率、組合計數、排容原理，以及生日問題的蒙地卡羅模擬。

```python
import numpy as np
from math import comb, prod

rng = np.random.default_rng(42)

# --- 1. 頻率觀點：長期頻率趨近機率 ---
flips = rng.integers(0, 2, size=1_000_000)   # 0/1 = 反面/正面
print(flips.mean())                          # -> 約 0.5

# --- 2. 計數：五張手牌總數與同花機率 ---
total = comb(52, 5)
flush = 4 * comb(13, 5)
print(total, flush / total)                  # 2598960, 約 0.00198

# --- 3. 排容原理：P(A or B) 必須扣掉交集 ---
PA, PB, PAB = 0.5, 0.4, 0.2
print(PA + PB - PAB)                         # 0.7，而不是 0.9

# --- 4. 生日問題：理論 vs. 模擬 ---
n = 23
p_theory = 1 - prod((365 - k) / 365 for k in range(n))
bdays = rng.integers(0, 365, size=(100_000, n))
p_sim = np.mean([len(np.unique(row)) < n for row in bdays])
print(round(p_theory, 4), round(p_sim, 4))   # 約 0.5073, 0.507x
```

本章總結

- 樣本空間 Ω 收納所有可能結果；**事件**是 Ω 的子集合；機率度量事件的「大小」。寫對樣本空間是一切計算的第一步。
- **Kolmogorov 三公理**（非負性、規範性、互斥事件可加性）生成所有規則：補集規則、單調性、排容原理。
- 機率**只有在事件互斥時**才能直接相加；有重疊時必須使用 $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$。
- 等可能結果下，機率就是計數：**排列** $P(n,k)$ 管順序、**組合** $\binom{n}{k}$ 不管順序，兩者相差 $k!$ 倍。
- 「難算正面就算補集」（生日問題）與「先分步再相乘」（乘法原理）是計數的兩大心法。

- **頻率學派**（長期頻率）與**貝氏學派**（信念程度）兩種詮釋都貫穿機器學習：訓練用頻率式的最大概似，預測輸出讀作貝氏式的信念。

1.6 練習題

建議先獨立完成再對照詳解。題目依難度標示為**基礎**、**進階**、**挑戰**三級。

練習 1.1 (基礎)　擲兩顆公正骰子，記錄點數對 (i, j)。(a) 樣本空間共有幾個結果？(b) 求「點數和等於 7」的機率。(c) 求「點數和至少 10」的機率。

練習 1.2 (基礎)　從一副 52 張撲克牌隨機抽一張。設 $A =$「紅心」、$B =$「人頭牌（J、Q、K）」。求 $\mathbb{P}(A)$、$\mathbb{P}(B)$、$\mathbb{P}(A \cap B)$、$\mathbb{P}(A \cup B)$，以及「既不是紅心也不是人頭牌」的機率，並指出最後一問與 De Morgan 定律的關係。

練習 1.3 (基礎)　僅使用 Kolmogorov 三公理，證明：(a) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$；(b) 若 $A \subseteq B$ 則 $\mathbb{P}(A) \le \mathbb{P}(B)$；(c) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$。

練習 1.4 (基礎)　某班 40 位學生中，25 人修統計、18 人修程式設計、10 人兩科都修。隨機抽出一位學生，求：(a) 至少修一科的機率；(b) 恰好修一科的機率；(c) 兩科都沒修的機率。

練習 1.5 (基礎)　社團有 8 位成員。(a) 選出 3 人組成委員會，有幾種選法？(b) 選出主席、副主席、書記各一人（一人不得兼任），有幾種選法？(c) 說明兩個答案為何差 $3! = 6$ 倍。

練習 1.6 (進階)　求五張撲克手牌是**葫蘆**（full house：三張同點數＋兩張另一同點數）的機率。

練習 1.7 (進階)　寫出生日問題的一般公式，驗證 $n = 23$ 時機率首次超過 $1/2$（計算 $n = 22$ 與 $n = 23$ 的數值），並用「配對數 $\binom{n}{2}$」的觀點解釋為何人數不用很多。

練習 1.8 (進階)　擲一顆公正骰子兩次。設 $A =$「第一次為偶數」、$B =$「第二次為偶數」、$C =$「兩次點數和為偶數」。(a) 分別求 $\mathbb{P}(A), \mathbb{P}(B), \mathbb{P}(C)$。(b) 求所有兩兩交集與三重交集的機率。(c) 用三事件排容原理求 $\mathbb{P}(A \cup B \cup C)$，並解釋所得結果的直觀意義。

練習 1.9 (進階)　6 個人隨機圍一張圓桌而坐（旋轉後相同的座位安排視為同一種）。求甲、乙兩人相鄰而坐的機率。

練習 1.10 (進階)　概念題：(a)「這枚公正硬幣擲出正面的機率是 1/2」與「這場官司原告勝訴的機率是 70%」分別適合哪種詮釋？為什麼？(b) 一個影像分類器對某張照片輸出 $\mathbb{P}(\text{貓}) = 0.83$。從詮釋的角度說明：這個數字在什麼意義下是頻率式的？在什麼意義下是貝氏式的？

練習 1.11 (挑戰)　**錯排問題（matching problem）**：n 封信隨機放入 n 個寫好地址的信封（每個信封恰一封信）。(a) 用排容原理證明「沒有任何一封信裝對信封」的機率為 $\sum_{k=0}^{n} \frac{(-1)^k}{k!}$。(b) 說明 $n \to \infty$ 時此機率趨近 $e^{-1} \approx 0.3679$。

練習 1.12 (挑戰)　求五張撲克手牌是**順子**（straight：五張點數連續，如 A-2-3-4-5 到 10-J-Q-K-A，共 10 種起點；花色不限但**不得**全部同花色，全同花色者稱為同花順，不算在內）的機率。

練習 1.13 (挑戰·程式)　用 NumPy 寫一個模擬：擲兩顆骰子 10^6 次，估計「點數和為 7」的機率並與理論值 1/6 比較；再模擬練習 1.8 中的事件 $A \cup B \cup C$，驗證其機率為 1。

1.7　練習題詳解

練習 1.1 詳解　(a) 每顆骰子 6 種結果，由乘法原理 $|\Omega| = 6 \times 6 = 36$，且每個點數對等可能。

(b) 和為 7 的結果：$(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)$，共 6 個，故

$$\mathbb{P}(\text{和} = 7) = \frac{6}{36} = \frac{1}{6} \approx 0.167.$$

（有趣的事實：7 是兩骰之和中機率最大的值。）

(c) 和至少 10 的結果：和為 10 有 $(4,6), (5,5), (6,4)$ 共 3 個；和為 11 有 $(5,6), (6,5)$ 共 2 個；和為 12 有 $(6,6)$ 共 1 個。合計 $3 + 2 + 1 = 6$ 個，

$$\mathbb{P}(\text{和} \geq 10) = \frac{6}{36} = \frac{1}{6}.$$

練習 1.2 詳解　紅心 13 張，人頭牌每花色 3 張共 12 張，紅心人頭牌（紅心 J、Q、K）3 張：

$$\mathbb{P}(A) = \frac{13}{52} = \frac{1}{4}, \qquad \mathbb{P}(B) = \frac{12}{52} = \frac{3}{13}, \qquad \mathbb{P}(A \cap B) = \frac{3}{52}.$$

由排容原理

$$\mathbb{P}(A \cup B) = \frac{13 + 12 - 3}{52} = \frac{22}{52} = \frac{11}{26} \approx 0.423.$$

「既非紅心也非人頭牌」即 $A^c \cap B^c$。由 De Morgan 定律 $A^c \cap B^c = (A \cup B)^c$，故

$$\mathbb{P}(A^c \cap B^c) = 1 - \mathbb{P}(A \cup B) = 1 - \frac{11}{26} = \frac{15}{26} \approx 0.577.$$

De Morgan 把「兩者皆非」轉換成「並非至少其一」，再套用補集規則，是這類問題的標準解法。

練習 1.3 詳解　(a) A 與 A^c 互斥且聯集為 Ω。由公理 3（可加性）：$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega)$；由公理 2（規範性）右邊為 1，移項即得 $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$。
(b) 因 $A \subseteq B$，可寫 $B = A \cup (B \cap A^c)$，兩塊互斥。由可加性

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c),$$

而公理 1 保證 $\mathbb{P}(B \cap A^c) \geq 0$，故 $\mathbb{P}(B) \geq \mathbb{P}(A)$。
(c) 將 $A \cup B$ 分割為互斥三塊：$A \cap B^c$、$A^c \cap B$、$A \cap B$，記其機率為 x, y, z。可加性給出

$$\mathbb{P}(A \cup B) = x + y + z, \qquad \mathbb{P}(A) = x + z, \qquad \mathbb{P}(B) = y + z.$$

於是

$$\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = (x + z) + (y + z) - z = x + y + z = \mathbb{P}(A \cup B). \qquad \blacksquare$$

練習 1.4 詳解　設 $S =$「修統計」、$C =$「修程式設計」，$\mathbb{P}(S) = \frac{25}{40}$、$\mathbb{P}(C) = \frac{18}{40}$、$\mathbb{P}(S \cap C) = \frac{10}{40}$。
(a) 排容原理：

$$\mathbb{P}(S \cup C) = \frac{25 + 18 - 10}{40} = \frac{33}{40} = 0.825.$$

(b)「恰好一科」= 只修統計 + 只修程式設計：

$$\frac{25 - 10}{40} + \frac{18 - 10}{40} = \frac{15 + 8}{40} = \frac{23}{40} = 0.575.$$

（也可用 $\mathbb{P}(S \cup C) - \mathbb{P}(S \cap C) = \frac{33}{40} - \frac{10}{40}$ 驗證。）
(c) 補集規則：

$$\mathbb{P}(\text{都沒修}) = 1 - \mathbb{P}(S \cup C) = 1 - \frac{33}{40} = \frac{7}{40} = 0.175.$$

練習 1.5 詳解　(a) 委員會不分職位，順序不重要，用組合：

$$\binom{8}{3} = \frac{8 \times 7 \times 6}{3!} = 56 \text{ 種}.$$

(b) 三個職位互不相同，順序重要，用排列：

$$P(8, 3) = 8 \times 7 \times 6 = 336 \text{ 種}.$$

(c) 每一組 3 人的委員會，可以用 $3! = 6$ 種方式分配主席 / 副主席 / 書記三個職位，因此「有序」數量恰為「無序」數量的 6 倍：$336 = 56 \times 6$。這正是公式 $\binom{n}{k} = P(n,k)/k!$ 的內容。

練習 1.6 詳解　分四步建構一手葫蘆（三條 + 一對）：

1. 選「三條」的點數：$\binom{13}{1} = 13$ 種；
2. 從該點數的 4 張牌中選 3 張：$\binom{4}{3} = 4$ 種；
3. 選「一對」的點數（不得與三條相同）：$\binom{12}{1} = 12$ 種；
4. 從該點數的 4 張牌中選 2 張：$\binom{4}{2} = 6$ 種。

由乘法原理，葫蘆總數為

$$13 \times 4 \times 12 \times 6 = 3744,$$

機率為

$$\mathbb{P}(\text{葫蘆}) = \frac{3744}{2{,}598{,}960} \approx 0.00144,$$

約 0.14%，平均每 694 手出現一次。注意步驟 1、3 **不能**寫成 $\binom{13}{2}$：三條的點數與一對的點數角色不同（3 張 vs. 2 張），順序有別，必須依序選取。

練習 1.7 詳解　一般公式（$n \le 365$）：

$$\mathbb{P}(\text{至少兩人同生日}) = 1 - \prod_{k=0}^{n-1} \frac{365-k}{365}.$$

數值驗證：

$$n = 22: \quad 1 - 0.5243 = 0.4757 < \tfrac{1}{2}, \qquad n = 23: \quad 1 - 0.4927 = 0.5073 > \tfrac{1}{2},$$

故 $n = 23$ 是機率首次超過一半的人數。
直觀解釋：撞生日的「機會」不是 n 次，而是每一**對**人都有一次機會，共 $\binom{n}{2}$ 對。$n = 23$ 時有 $\binom{23}{2} = 253$ 對；每對同生日的機率約 $1/365$，粗略估計

$$\mathbb{P}(\text{至少一對相同}) \approx 1 - \left(1 - \frac{1}{365}\right)^{253} \approx 1 - e^{-253/365} \approx 0.500,$$

與精確值 0.5073 非常接近。配對數以 n^2 級成長，這就是為什麼人數不需要接近 365。

練習 1.8 詳解　(a) 骰子偶數有 $\{2,4,6\}$，故 $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$。和為偶數 $\iff$ 兩次同奇偶（偶偶或奇奇），

$$\mathbb{P}(C) = \tfrac{1}{2}\cdot\tfrac{1}{2} + \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{2}.$$

(b) $A \cap B$ =「兩次皆偶」：$\mathbb{P} = \frac{1}{4}$。$A \cap C$：第一次偶且和偶 $\Rightarrow$ 第二次也必須是偶，故 $A \cap C = A \cap B$，$\mathbb{P} = \frac{1}{4}$。同理 $B \cap C = A \cap B$，$\mathbb{P} = \frac{1}{4}$。$A \cap B \cap C$ =「兩次皆偶」（此時和自

動為偶），$\mathbb{P} = \frac{1}{4}$。
(c) 排容原理：

$$\mathbb{P}(A \cup B \cup C) = \underbrace{\tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2}}_{3/2} - \underbrace{\tfrac{1}{4} - \tfrac{1}{4} - \tfrac{1}{4}}_{\text{扣}3/4} + \tfrac{1}{4} = \tfrac{3}{2} - \tfrac{3}{4} + \tfrac{1}{4} = 1.$$

機率為 1 表示 $A \cup B \cup C = \Omega$：三個事件**必然至少發生一個**。直觀驗證——若 A 與 B 都不發生，表示兩次都是奇數，那麼和 = 奇 + 奇 = 偶，C 必然發生。排容原理算出的「1」正好印證了這個邏輯事實。

練習 1.9 詳解　圓桌排列把旋轉視為相同：固定其中一人當「基準」消除旋轉自由度，其餘 5 人可任意排列，總數為 $(6-1)! = 120$ 種。
甲乙相鄰的排法：把甲乙「綁」成一個單位，圓桌上共 5 個單位，圓排列數 $(5-1)! = 24$；甲乙在單位內可交換左右，乘上 2，得 $24 \times 2 = 48$ 種。

$$\mathbb{P}(\text{甲乙相鄰}) = \frac{48}{120} = \frac{2}{5}.$$

快速驗算：固定甲的座位，乙落在剩下 5 個座位中的任一個，等可能；其中與甲相鄰的座位恰有 2 個，故 $\mathbb{P} = 2/5$。兩法一致。

練習 1.10 詳解　(a)「公正硬幣出正面機率 1/2」適合**頻率詮釋**：拋硬幣可以無限重複，1/2 描述長期頻率（亦可由物理對稱性支持）。「原告勝訴機率 70%」適合**貝氏詮釋**：這場官司獨一無二、不可重複，70% 表達的是在現有證據下對結果的信念強度，新證據出現時應當更新。
(b) 頻率式的一面：模型的參數是用大量歷史資料、以最大概似（頻率學派方法，第 7 章）訓練而得；而且若模型**校準良好**（第 13 章），那麼在所有被它標為 0.83 的照片中，長期而言約有 83% 真的是貓——這是一個可用頻率檢驗的性質。貝氏式的一面：對**這一張**特定的照片而言，「是貓」已是既定事實，0.83 表達的是模型基於這張圖的特徵所持的信念程度，功能上等同於貝氏後驗機率。實務上我們同時使用兩種讀法。

練習 1.11 詳解　(a) 設 $A_i =$「第 i 封信裝對信封」。所求為 $\mathbb{P}\big(\bigcap_i A_i^c\big) = 1 - \mathbb{P}\big(\bigcup_i A_i\big)$。
排容原理的一般形式：

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{i_1 < \cdots < i_k} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}).$$

指定 k 封特定的信全部裝對時，其餘 $n-k$ 封信可任意排列，故

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \frac{(n-k)!}{n!},$$

而選出這 k 封信的方式有 $\binom{n}{k}$ 種，因此第 k 層的貢獻為

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

代入排容原理：

$$\mathbb{P}\Big(\bigcup_{i} A_i\Big) = \sum_{k=1}^{n}\frac{(-1)^{k+1}}{k!}, \qquad \mathbb{P}(\text{全部裝錯}) = 1 - \sum_{k=1}^{n}\frac{(-1)^{k+1}}{k!} = \sum_{k=0}^{n}\frac{(-1)^{k}}{k!}.$$

(b) 由指數函數的泰勒展開 $e^{x} = \sum_{k=0}^{\infty} x^{k}/k!$，取 $x=-1$：

$$\lim_{n\to\infty}\sum_{k=0}^{n}\frac{(-1)^{k}}{k!} = e^{-1} \approx 0.3679.$$

收斂極快：$n=6$ 時已準確到小數第三位。無論 10 封信還是 1000 萬封信，「全部裝錯」的機率都約為 36.8%——一個反直覺的美麗結果。

練習 1.12 詳解　分步計數順子（不含同花順）：

1. 選起點：順子可為 A-2-3-4-5、2-3-4-5-6、……、10-J-Q-K-A，共 10 種點數組合；
2. 五張牌各自獨立選花色：$4^5 = 1024$ 種；
3. 扣除五張全同花色的情形（即同花順）：每種點數組合有 4 種全同花，故扣 4。

順子總數：

$$10 \times (4^5 - 4) = 10 \times 1020 = 10{,}200,$$

機率：

$$\mathbb{P}(\text{順子}) = \frac{10{,}200}{2{,}598{,}960} \approx 0.00392,$$

約 0.39%，平均每 255 手一次。（對照：同花 $\approx 0.197\%$、葫蘆 $\approx 0.144\%$——這正是撲克牌型大小排序「順子 < 同花 < 葫蘆」的由來：越稀有，牌型越大。）

練習 1.13 詳解　參考程式與輸出：

```python
import numpy as np
rng = np.random.default_rng(0)
N = 1_000_000

# (a) 兩骰之和為 7
d1 = rng.integers(1, 7, size=N)
d2 = rng.integers(1, 7, size=N)
p_seven = np.mean(d1 + d2 == 7)
print(round(p_seven, 4), round(1/6, 4))      # 約 0.1667 vs 0.1667

```

```python
# (b) 練習 1.8：A = 第一骰偶數, B = 第二骰偶數, C = 和為偶數
A = (d1 % 2 == 0)
B = (d2 % 2 == 0)
C = ((d1 + d2) % 2 == 0)
print(np.mean(A | B | C))                    # 1.0，與排容原理的結論一致
```

模擬結果 $\hat{p} \approx 0.1667$ 與理論值 $1/6$ 相符（誤差約 10^{-3} 量級，來自有限樣本的隨機波動，第 6 章的中央極限定理會告訴我們誤差正比於 $1/\sqrt{N}$）。第二個輸出恆為 1.0：$A \cup B \cup C$ 涵蓋整個樣本空間，每一次模擬都至少發生其中一個事件。

下一章：條件機率與貝氏定理——機率如何吸收證據。