

單元 4：實驗的機率模型：離散情況 (課本 §2.4)

定義 1. 一實驗就是一觀察結果 (值) 被製造出的過程.

註 1. 習慣用語上, 一實驗包含由完全無法掌控狀態下 (completely uncontrollable situations) 所獲得的觀察值 (如, 某特定股票每日行情的觀察值) 以及可控制實驗狀況下 (controlled laboratory conditions) 的觀察值.

註 2. 實驗的例子, 如: 投擲銅板及骰子, 度量個人的智商分數 (IQ), 決定食物中每單位立方公分中細菌的數量.

註 3. 執行實驗後, 會產生一個或多個結果, 統稱為事件 (events) (亦即, 事件為實驗執行後的結果的描述), 常以大寫字母表示之. 如, 實驗為計數食物中的細菌數: 一些有興趣的事件可能為

A : 剛好有 110 個細菌

B : 超過 200 個細菌

C : 細菌數介於 100 與 300 之間

或實驗為投擲一公正的骰子：一些與此實驗關聯的事件為

A ：觀察到奇數點

B ：觀察到小於 5 的點數

C ：出現 2 或 3

E_1 ：出現 1, E_2 ：出現 2, ..., E_6 ：出現 6

事件間的分類 (差異):

(1) 觀察到事件 $A \Leftrightarrow$ 在同時，會觀察到事件 E_1 或 E_3 或 E_5 . 此種可被分解 (decomposed) 的事件稱作複合事件 (compound event) (可由多種途徑發生).

(2) 不能再被分解的事件，如事件 $E_1, \dots, E_6$, 稱作簡單事件 (simple events) (僅能由一種途徑發生).

註 4. 習慣上，可將 “不同的單一點 (distinct point), 稱作樣本點 (sample point)” 與 “一實驗的簡單事件” 關聯之.

定義 2. 簡單事件 (simple event) 是一不可分解的事件. 每一簡單事件對應一個且唯一的一個樣本點

(sample point). 含下標的字母 E 被用於表示一簡單事件或其對應的樣本點.

定義 3. 一實驗 (關聯) 的樣本空間 (sample space) 是所有可能樣本點的集合, 並以 S 表示之.

註. 投擲骰子實驗的樣本空間

$$S = \{E_1, E_2, \dots, E_6\}$$

有限的, 其 Venn diagram 如下圖所示.

食物樣本中, 計數細菌的實驗的

$$S = \{E_0, E_1, E_2, E_3, \dots\}$$

其中 $E_i =$ 觀察到 i 個細菌的事件, $i = 0, 1, 2, \dots$, 乃可數無限的 (countably infinite).

定義 4. 一含有限個或可數無限個不同樣本點的樣本空間稱作離散樣本空間.

註 1 針對離散樣本空間的實驗, 複合事件 (compound event) 可被視為樣本點的集合 (collection of

sample points) 或單一樣本點集合 (set of single sample point) 對應的簡單事件 (simple event) 的聯集, 如投擲一骰子,

$$\begin{aligned} A \text{ (出現奇數)} &= \{E_1, E_3, E_5\} \text{ (樣本點的集合)} \\ &\stackrel{\text{或}}{=} E_1 \cup E_3 \cup E_5 \text{ (簡單事件的聯集)} \end{aligned}$$

以及

$$\begin{aligned} B \text{ (小於 5 點)} &= \{E_1, E_2, E_3, E_4\} \\ &\quad \text{(collection of sample points)} \\ &\stackrel{\text{或}}{=} E_1 \cup E_2 \cup E_3 \cup E_4 \\ &\quad \text{(union of simple events)} \end{aligned}$$

註 2. 簡單事件 E_i 被包含於事件 $A \Leftrightarrow$ 當 E_i 發生時, A 發生.

定義 5. 在離散樣本空間 S 中, 一事件就是一個樣本點的集合, 亦即, S 的任一子集合就是一事件.

註 1. 事件間的關係可清楚地由 Venn diagram 看出: 如投骰子實驗 (die-tossing experiment), 其事件間的關係如下圖所示.

註 2. 計數細菌的實驗,

$$B \text{ (超過 200 個細菌)} = \{E_{201}, E_{202}, E_{203}, \dots\}$$

其中 $E_i =$ 出現 i 個細菌的簡單事件.

註 3. 針對離散樣本空間的實驗, 其機率模型的建構可經由指定樣本空間 S 中每一簡單事件一個數值機率 (numerical probability) 而達成. 此數字 (亦即, 數值機率) 表示重複實驗一次時, 相信此事件發生的度量且須與機率的相對頻率概念相一致.

雖然相對頻率的概念不能提供機率的嚴格定義, 但任何可應用到真實世界的定義都應與事件的相對頻率的直覺觀念相吻合. 檢視機率的相對頻率概念的三要件:

- (1) 任一事件的相對頻率必須 ≥ 0 ; 負的相對頻率沒有任何意義.
- (2) 整個樣本空間 S 的相對頻率必須為 1, 因為實驗的每一種可能結果都是 S 中的一點, 故每次執行實驗時, S 必要發生.

(3) 若 2 事件互斥, 其聯集的相對頻率是此 2 事件對應的相對頻率的和. 如, 投一公正的骰子, 投擲總次數中的 $1/6 + 1/6 = 1/3$ 會是 1 點或 2 點.

此三要件構成如下機率定義的基礎.

定義 6. 設 S 為一實驗 (關聯) 的樣本空間. 針對 S 中的每一事件 A (亦即, A 為 S 的一子集合), 指定一數 $P(A)$, 稱作 A 的機率, 以致於下列公設成立:

公設 1: $P(A) \geq 0$

公設 2: $P(S) = 1$

公設 3: 若 $A_1, A_2, A_3, \dots$ 形成一個兩兩互斥事件的序列 (sequence of pairwise mutually exclusive events) (亦即, $A_i \cap A_j = \emptyset$, 若 $i \neq j$), 則

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

(稱此性質為可數可加性, countable additivity)

註 1. 針對有限個兩兩互斥事件 $A_1, A_2, \dots, A_n$, 由公設 3 可導出

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^n P(A_i)$$

<證> 取 $A_{n+1} = A_{n+2} = \dots = \emptyset$ 並根據 $P(\emptyset) = 0$ (因為 $\emptyset$ 為一不會發生, 不存在的事件), 可得

$$\begin{aligned} & P(A_1 \cup A_2 \cup \dots \cup A_n) \\ &= P(A_1 \cup \dots \cup A_n \cup A_{n+1} \cup A_{n+2} \cup \dots) \\ &= \sum_{i=1}^{\infty} P(A_i) \text{ (根據公設 3)} \\ &= \sum_{i=1}^n P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset) \\ &= \sum_{i=1}^n P(A_i) \end{aligned}$$

註 2. 定義僅敘述指定機率時, 必須滿足的條件, 卻無說明如何指定 “特定的機率” 到 “事件” 上. 若機率模型是有實用的目的, 則 “特定的機率” 的指定必須與實際一致才行, 如投擲銅板 1000 次後, 出現 800 次正面. 再擲

銅板 1 次, 可能的結果為正面或反面, 所以根據機率的定義, 可指定

$$P(\text{正面}) = p_1 \geq 0 \text{ 以及 } P(\text{反面}) = p_2 \geq 0$$

且

$$p_1 + p_2 = 1$$

即可 (如, $p_1 = p_2 = 1/2$, 但不實際). 因此, 一個比較合理的機率指定為

$$P(\text{正面}) = 0.8, P(\text{反面}) = 0.2$$

註 3. 針對離散樣本空間, 只要指定機率給每一個簡單事件就夠了. (因為其它事件只是這些簡單事件的聯集)

例 1. 設有 5 台完全相同的電腦終端機, 其中有 2 台是有瑕疵的. 今任選 2 台並運送出.

(a) 試列出此實驗的樣本空間.

<解> 令 D_1, D_2 表示 2 台有問題的, 且 G_1, G_2, G_3 表示 3 台好的. 則 10 個樣本點 (又稱簡單事件)

爲

$$\begin{aligned} E_1 &= \{D_1, D_2\}, E_2 = \{D_1, G_1\}, E_3 = \{D_1, G_2\}, \\ E_4 &= \{D_1, G_3\}, E_5 = \{D_2, G_1\}, E_6 = \{D_2, G_2\}, \\ E_7 &= \{D_2, G_3\}, E_8 = \{G_1, G_2\}, E_9 = \{G_1, G_3\}, \\ E_{10} &= \{G_2, G_3\} \end{aligned}$$

故,

$$S = \{E_1, E_2, \dots, E_{10}\}$$

(b) 令 $A =$ 送出的兩台均是好的事件. 試列出 A 中的樣本點.

<解> $A = \{E_8, E_9, E_{10}\}$

(c) 試建構一描述事件 A 的 Venn-diagram.

<解> 如下圖.

(d) 試對簡單事件指定出合用的機率.

<解> 因爲是任選終端機, 任一對被選中的可能性都是一樣的. 所以, 合理的機率指定爲

$$P(E_i) = \frac{1}{10}, \quad i = 1, \dots, 10$$

(e) 試求事件 A 的機率.

<解> 因爲

$$A = E_8 \cup E_9 \cup E_{10}$$

故,

$$\begin{aligned} P(A) &= P(E_8) + P(E_9) + P(E_{10}) \text{ (互斥)} \\ &= \frac{3}{10} \end{aligned}$$

註. 有些實驗其樣本空間是不可數的 (uncountable), 故爲非離散的, 如實驗爲度量糖尿病患 (diabetic patient) 的血糖指數 (blood glucose level), 則其樣本空間將包含實數的一區間, 故含有不可數個數值 (uncountable number of values). (此種情況將於 ch4 中探討) (本章剩餘章節將專注於離散樣本空間上的事件機率的計算)