CHAPTER 4

機　率

章前導讀

本書第 1 章即已指出統計學的範圍包括敘述統計與推論統計，而第 2 章、3 章專門介紹敘述統計的觀念與方法。從本章開始，將逐漸深入推論統計的領域。在探討推論統計之前，需先瞭解一些必備的課題，此即機率理論 (probability theory)。第 4、5、6 章將專注於機率理論相關課題的介紹，包括機率的基本概念、隨機變數與機率分配，以及一些常用的機率分配。

版權所有

華泰文化
HWA TAI PUBLISHING

4.1 導論

第 1 章曾介紹母體與樣本的概念，其中母體是指我們感興趣而欲加以探討的全部對象，而樣本只是母體的一部分，它只能提供部分的訊息。所謂推論統計即是根據此部分的訊息，以推測母體的某些特徵。換句話說，推論統計即是根據已知的樣本特徵來推測未知的母體特徵，這在科學邏輯的術語上稱為**歸納推理** (inductive reasoning)。但是，從另一方面，亦可能已事先知道母體的某些特徵〔例如，某校全體學生（母體）近視的比例〕，若自此一母體隨機抽出一個樣本（可能包含 50 名學生），則可根據對母體某些特徵的瞭解來探討樣本的某一特徵（例如，此樣本學生近視的比例），這種過程與前述推論統計的過程相反，稱為**演繹推理** (deductive reasoning)，此即機率理論探討的主題。換句話說，機率理論即根據已知的母體特徵來探討未知的樣本特徵；而統計推論則是根據已知的樣本特徵來推測未知的母體特徵，參見圖 4.1。由於機率理論較為簡單且發展得較早，因此我們先討論機率理論；而事實上，機率理論乃是推論統計所應具備的基礎。

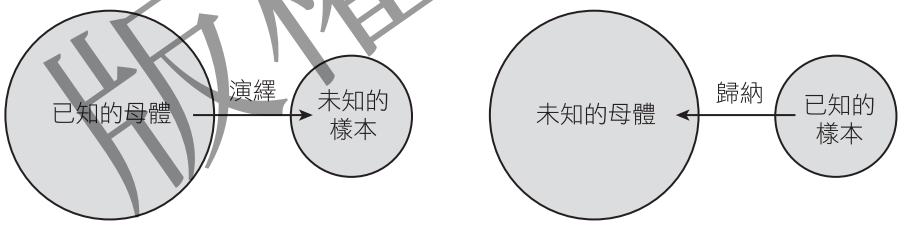


圖 4.1　演繹法（機率理論）與歸納法（統計推論）的比較

機率理論主要在探討「如何由一已知的母體特徵去推斷一未知的樣本特徵」。例如，假定某校學生（母體）的近視比例為 60%，則我們似乎不可能以隨機抽樣的方式，抽出近視比例剛好為 60%的樣本。但是，抽出比例為「可能地」(likely) 或「接近於」(close to) 60%近視的樣本卻是有可能的。

版權所有
華泰文化
HWA TAI PUBLISHING

機率理論即在於更精確地定義「可能地」與「接近於」的隱含意義，並以具體的數字來衡量此抽象的意義。

機率是衡量某一事件可能發生的程度（機會大小），並針對此一不確定事件發生之可能性賦予一量化的數值。

日常生活中，我們常聽到下列的用語：

「本週末可能會下雨，因此郊遊活動很可能要取消」；
「明天某公司的股票可能會漲」；
「今天的棒球賽，甲隊不太可能贏乙隊」。

諸如此類：「很可能」、「可能」、「不太可能」等等的用語，基本上都表示了不確定性程度的大小，而機率便是利用一量化的數值來說明事件將發生之可能性，其方式是：當一事件極不可能發生時，其機率值接近於0；相反地，當一事件極可能發生時，其機率值接近於1。由此可知，機率是一個數值，用以測度某一事件之不確定的程度，其值的範圍介於0與1之間。

綜合上述，機率乃是統計推論的基礎，可藉機率值的大小及其觀念，來表示推論結果的準確性。此外，機率亦是統計決策理論的重要工具，因為決策往往是在不確定性的情況下做出判斷，而機率則專門在探討不確定性的問題。因此藉由機率理論，可提供決策抉擇的參考。例如，投資決策、保險公司估計投保人的保費、經理人員預測產品的需求量等等，這些決策都需要根據事件發生的機率來決定。

承接上述機率理論之觀念，以下各節便依據圖4.2的架構逐一說明隨機實驗的整體概念與實例說明。

版權所有
華泰文化
HWA TAI PUBLISHING

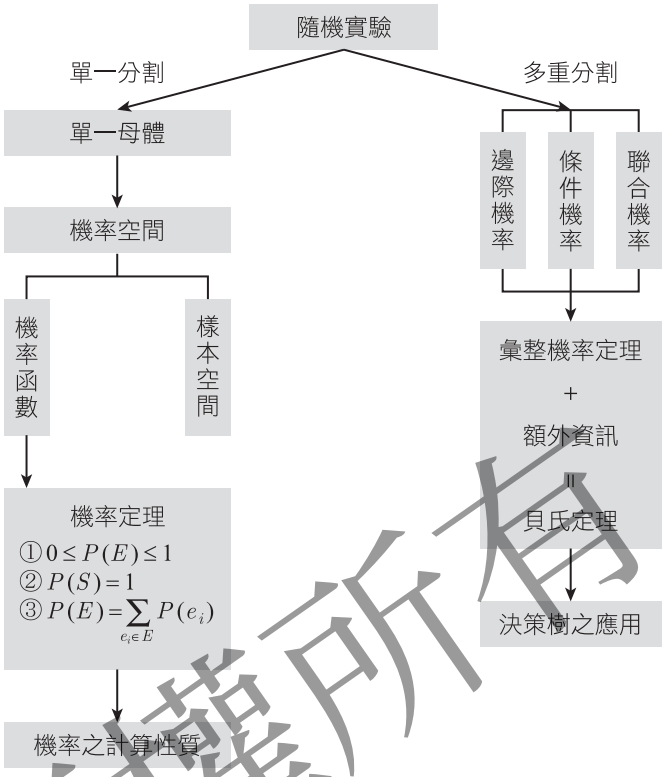


圖 4.2 隨機實驗之整體概念架構

版權所有

隨堂練習

1. 試寫出下列各隨機試驗的樣本空間：
   (a) 自一製造過程中隨機抽出四個產品檢驗，每次檢驗結果不是良品 (G) 就是不良品 (D)。
   (b) 20 名學生參加統計學期中考，調查其將通過此次考試的學生人數。
   (c) 檢驗一些產品，直到發現第二個不良品出現，以 D 表示不良品，G 表示良品。
   (d) 記錄 30 名學生 100 公尺速度，已知最快 12 秒，最慢 20 秒。

4.2 隨機試驗、樣本空間與事件

機率是針對不確定事件發生之可能性賦予一量化的數值；在此所謂的不確定性，是指我們知道某現象會有一些結果出現，但究竟是其中哪一種結果出現，則不得而知。因此，我們欲探討機率必先瞭解有哪些可能結

華泰文化
HWA TAI PUBLISHING

果，以及這些可能結果之組合的觀念。以下介紹幾個名詞的意義，並由此引申出更清晰的機率概念。

首先我們說明**隨機試驗** (random experiment) 的意義：

> 觀察一可產生各種**可能結果** (outcome) 的過程，稱為試驗；而若各種可能結果的出現（或發生）具有不確定性，則此一過程便稱為隨機試驗。

試驗一詞的意義極為廣泛，舉凡觀察、實驗、調查、檢驗、抽樣等，皆可謂試驗。例如，擲一骰子、自一副撲克牌抽出一張、對一群顧客進行意見調查及檢驗一批產品等等，皆是試驗的例子。上述這些試驗的例子皆具有一共同的特性，即每次試驗時都可知道有哪些可能結果，但究竟是其中哪一種結果會出現，則具有不確定性。例如，擲一個骰子，則其可能結果為 1、2、3、4、5 或 6，但出現 1 點、2 點或其他，每一次都不確定。此外，拋二個硬幣、檢驗三個產品、討論明天股票行情等等，由於各種可能結果的出現皆具有不確定性，因此皆屬於隨機試驗的例子。

嚴格來說，隨機試驗必須能滿足下列三個條件：

1. 試驗可以在相同的條件下重複進行；
2. 試驗的所有可能結果是明確可知的，且不止一個；
3. 每次試驗必定僅出現這些可能結果中的一個，但試驗之前無法肯定該次試驗會出現哪一個可能結果。

> 一隨機試驗之各種可能結果的集合，稱為**樣本空間** (sample space)；樣本空間內的每一元素，稱為**樣本點** (sample point)。

樣本空間是依據隨機試驗可能發生的各種可能結果來界定，可用集合的觀念來定義，一般以符號 S 表示。例如，擲一骰子的樣本空間為 $S_1 = \{1, 2, 3, 4, 5, 6\}$；拋二個硬幣的樣本空間為 $S_2 = \{(H、H), (H、T), (T、H), (T、T)\}$，其中 H 代表正面，T 代表反面。

版權所有

華泰文化
HWA TAI PUBLISHING

例題 4.1

自一批產品隨機抽出三件加以檢驗，由於每次檢驗皆僅有二種可能結果，即良品 G 與不良品 D，且每次的可能結果皆不確定，因此它是隨機試驗。此隨機試驗的樣本空間可用樹狀圖的方式列舉出來，如圖 4.3 所示。

圖 4.3 中，樣本點 GGG 代表三件產品皆為良品。樣本空間則為：

$$S = \{GGG, GGD, GDG, GDD, DGG, DGD, DDG, DDD\}$$

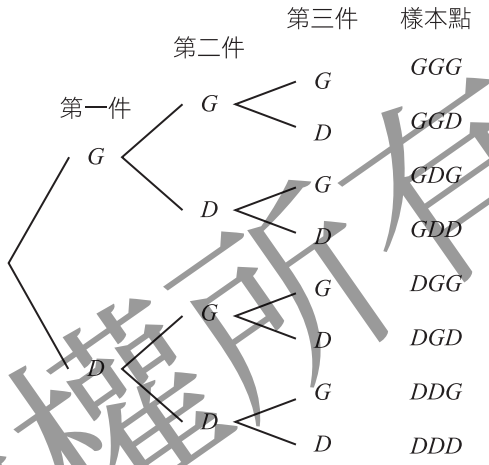


圖 4.3 檢驗三件產品之樣本空間的樹狀圖

樣本空間為一集合，而依據集合所含元素的個數，可定義出**有限樣本空間** (finite sample space) 與**無限樣本空間** (infinite sample space)；亦即：

僅含有限個樣本點的樣本空間，稱為有限樣本空間；含有無限多個樣本點的樣本空間，稱為無限樣本空間。

有限樣本空間與無限樣本空間的例子，參見表 4.1。

版權所有
華泰文化 HWA TAI PUBLISHING

表 4.1　有限樣本空間與無限樣本空間之舉例

隨機試驗	樣本空間	有限或無限
1. 擲一骰子	$S_1 = \{1, 2, 3, 4, 5, 6\}$	有限
2. 調查某班級近視人數（設有 45 名學生）	$S_2 = \{0, 1, 2, 3, \cdots, 45\}$	有限
3. 擲一硬幣兩次	$S_3 = \{HH, HT, TH, TT\}$	有限
4. 檢驗一批產品，直到發現有不良品為止，記錄檢驗的產品數	$S_4 = \{1, 2, \cdots, n, \cdots\}$	無限
5. 觀察某一燈管之使用壽命	$S_5 = \{t \mid t > 0\}$ t 表壽命時間	無限

在計算隨機試驗可能出現的結果（樣本空間）時，除了可利用例題 4.1 所示的樹枝圖外，亦可利用乘數定理、排列 (permutation) 與組合 (combinations) 來求出樣本空間的個數與組合形式。

1. 乘數定理

設一個隨機試驗包含 k 個試驗 $E_1, E_2, \cdots, E_k$，若每一試驗 E_j 有 n_j 種結果，$j = 1, 2, \cdots, k$，則該隨機試驗有 $n_1 \times n_2 \times \cdots \times n_k$ 種可能結果。

2. 排列

自一個含有 m 個元素的集合中，一次抽取 n 個元素（或每抽取一個，抽出不放回，連續抽取 n 個），則共有 P_n^m 個不同排列的樣本組，其排列數為：

$$P_n^m = \frac{m!}{(m-n)!}$$

（其中 $m!$ 視為是 m 階乘，代表：$m! = m(m-1)(m-2)\cdots(1)$）

注意：(1) 若自 m 個不同的項目中，取出 n 個項目加以排列，且允許每個項目均可重複排列，則其排列數為 n^r。

(2) 假設 m 物中，包含了 k 種不同的東西，其中 r_1 屬某一類，r_2 歸為另一類，……最終 r_k 個又是另外一類，在此情況之下，若 $r_1 + r_2 + \cdots + r_k = m$，則所有 m 物之排列數為：

$$P_{r_1, r_2, \cdots, r_k}^m = \frac{m!}{r_1! r_2! \cdots r_k!}$$

版權所有
華泰文化
HWA TAI PUBLISHING

例題 4.2

2013年初，某大學社會服務隊募集了3套棉被、5個電鍋及12個電暖爐，欲送給學校附近社區的20位獨居老人。請問共有幾種可能的分送方式？

解

$$P_{3,5,12}^{20} = \frac{20!}{3!5!12!} = 7{,}054{,}320$$

3. 組合

自一個含有 m 個元素的集合中，一次抽取 n 個元素，若不考慮 n 個元素被抽中元素的順序，則共有 C_n^m 個不同組合，其組合數為：

$$C_n^m = \frac{m!}{n!(m-n)!}$$

例題 4.3

目前台灣地區盛行的大樂透，乃從1號到49號任選6個號碼，若所下注的6個號碼與威力彩開出的號碼完全相同即中頭彩。試問簽注威力彩共有幾種可能的號碼組合？

解

$$C_6^{49} = \frac{49!}{6!43!} = 13{,}983{,}816$$

最後我們來介紹事件的觀念，其定義如下：

事件 (event) 乃樣本空間的部分集合（子集）；每一樣本點皆為樣本空間的子集，故亦皆為事件，稱為**簡單事件** (simple event)；而含有兩個以上樣本點的事件，稱為**複合事件** (compound event)。

根據上面的定義，空集合 (ϕ) 與樣本空間本身 (S) 乃是二個特殊的子集，故亦為事件，但此二事件有其特別的意涵。空集合所代表的事件，因它不含任何樣本點，故一般稱為**不可能事件**；而事件 S 包含樣本空間內的所有樣本點，必然會發生，故一般稱為**必然事件**。

版權所有
華泰文化 HWA TAI PUBLISHING

例題 4.4

擲一硬幣兩次，其樣本空間為 $\{HH, HT, TH, TT\}$，若分別以 e_1, e_2, e_3, e_4 表示 HH, HT, TH, TT，則 S 亦可寫成

$$S = \{e_1, e_2, e_3, e_4\}$$

上式 S 中的各元素之排列順序並不重要，我們所關心的只是這些元素（樣本點）所組成的集合。由於 $\{e_1\}, \{e_2\}, \{e_3\}, \{e_4\}$ 皆為 S 的子集，且只包含一個樣本點，故皆為簡單事件。

假如出現一個正面之事件為 E，則事件 E 包含二個樣本點 e_2 與 e_3，亦即 $E = \{e_2, e_3\}$，此時 E 即為複合事件。當某次試驗出現 HH 或 TT 時，則稱事件 E 沒有發生（因為 HH 或 TT 皆不屬於 E）。

例題 4.5

假設一箱子中有 10 個球，其中 7 個紅球 (R)，3 個為白球 (W)，現在自此箱中隨機抽出一球。此時可能出現兩種不同的結果，即所抽出的球不是紅球便是白球，因此其樣本空間為 $S_1 = \{R, W\}$，只含有二個樣本點。

但是從另一個角度來看，我們可將每一次抽出一球視為一簡單事件，而該球的顏色視為其特性。換句話說，我們可以將 7 個紅球附上標籤 $e_1, e_2, \cdots, e_7$，而 3 個白球附上標籤 e_8, e_9, e_{10}。由於分析每一次抽球乃是隨機的，故其樣本空間共有 10 個樣本點，即：

$$S_2 = \{\underbrace{e_1, e_2, \cdots, e_7}_{\text{紅球}}, \underbrace{e_8, e_9, e_{10}}_{\text{白球}}\}$$

如果我們令 A 代表所抽出之球為紅球的事件，則 A 所包含的元素有

$$A = \{e_1, e_2, \cdots, e_7\}$$

由上述可知，第二種定義樣本空間的方式可能比第一種定義方式來得簡易，而且第二種方式之優點更是多於第一種。

例題 4.5 說明了一個重要的觀念：「就某一特定的問題中，樣本空間之形成並非是唯一的。」因此，當我們考慮如何選擇樣本空間時，應以便於機率問題之處理為原則，而非以簡易地指出樣本空間諸元素為評判標準。

版權所有
華泰文化
HWA TAI PUBLISHING

4.3 機率測度的方法

在一項隨機試驗中，如果其出現的結果是事件 E 所包含的樣本點之一，則稱事件 E 發生；否則，事件 E 沒有發生。至於如何賦予一事件一個機率值呢？一般而言，賦予機率值的方式有三種，以下分別說明之。

4.3.1 古典方法

使用古典方法測度事件的機率，乃基於下列限制條件與假設：

1. 限制條件：必須是有限的樣本空間。
2. 基本假設：假設樣本空間內的每一樣本點其出現的機會皆相同。

若符合上述的假設條件，則可依古典方法測度事件的機率：

古典方法的機率測度

在一有限的樣本空間 S 中，某一事件 E 的機率 $P(E)$ 定義為：

$$P(E)=\frac{n(E)}{n(S)} \tag{4-1}$$

式中 $n(S)$ 與 $n(E)$ 分別代表樣本空間與事件所包含的樣本點個數。

例題 4.6

擲一公正硬幣兩次，試求出現一個正面的機率？

解

正如例題 4.4 所示，其樣本空間包含 4 個樣本點，即 $S=\{HH, HT, TH, TT\}$。又所謂公正硬幣乃表示此 4 個樣本點中，每一樣本點出現之機會相等，因此符合古典方法的假設條件。令事件 E 代表出現一個正面，即 $E=\{HT, TH\}$，於是

$$P(E)=\frac{n(E)}{n(S)}=\frac{2}{4}=\frac{1}{2}$$

例題 4.7

在例題 4.5 中，如果隨機自箱中抽出一個球，則抽出紅球的機率為何？

版權所有
華泰文化 HWA TAI PUBLISHING