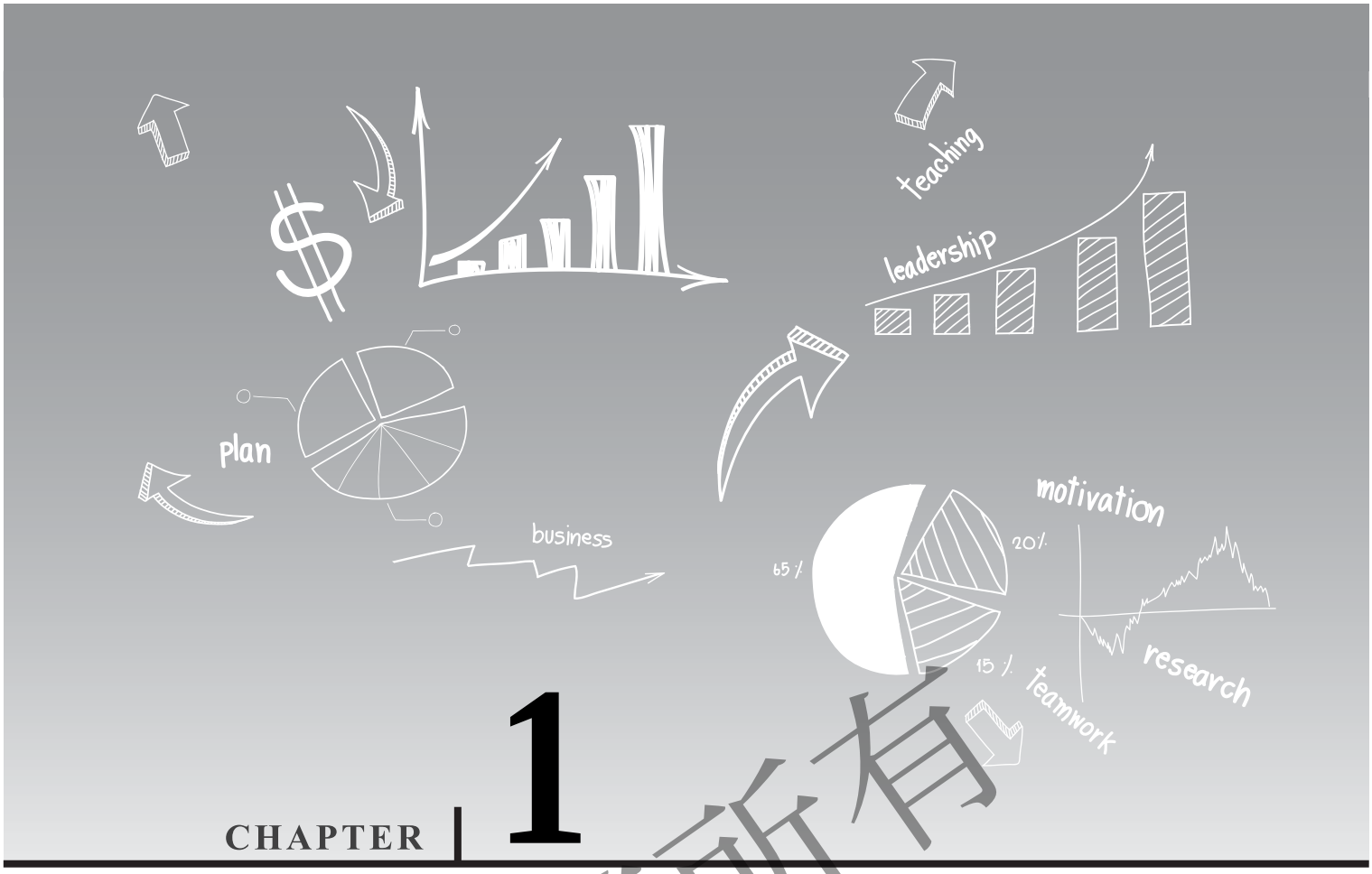


CHAPTER 1

緒　論

版權所有

1.1　統計學的定義及重要性
1.2　統計資料
1.3　統計的應用
1.4　統計學的內容與本書的架構

華泰文化
HWA TAI PUBLISHING

1.1 統計學的定義及重要性

「**統計**」(statistics)，此名詞源自於拉丁字「status」，意指狀態 (state)。其原本的意思在於顯示一個國家的經濟、人口、政治情勢之國情資料及圖形。也就是這樣，讓很多人對統計產生誤解，以為它只不過是一些數字的蒐集彙總與一些圖表表現而已。實際上，這些都只能當作統計的一小部分。現今的統計，隨著人類文明的進步與電腦的日益普及，它已發展成一門獨立的學門——稱為「統計學」。

如何運用統計學的觀念與方法，來幫助我們解決問題及提供做決策的參考依據是非常重要的。譬如，李登輝總統在「1998 年台北國際統計學術研討會」的書面賀詞中曾指出，我國統計數字的指揮控管與檢驗政策功能優異，在此次金融風暴中，適時提供決策者正確資訊，使我國經濟受創程度遠低於許多亞洲國家；《朝鮮日報》日前報導，韓國電子通信研究院 (Electronics and Telecommunications Research Institute, ETRI) 在 2008 年至 2010 年間，以智慧型手機的「大國」——韓國的全國 2,000 多名小學生為對象進行研究調查，並公布一份「關於電磁波對人體影響的綜合研究」的調查結果報告指出，青少年使用手機的時間愈長，發生「注意力不足過動症」(Attention Deficit Hyperactivity Disorder, ADHD) 的可能性愈高；並且也指出，青少年對手機電磁波的吸收量要比成人多 40%，更容易受到電磁波的傷害，而在國內，台南市環境保護聯盟也在 2008 年 8 月到 10 月間，首次針對 14 縣市學童進行手機使用調查，發現青少年學童攜帶手機比例頗高，並推估台灣 6 至 18 歲的青少年中，約 220 萬學童擁有手機，約 140 萬青少年學童帶手機上學，但有七成青少年學童不知非緊急情況不使用手機的安全預警觀念，所以此份報導的統計數據資訊提醒青少年在享受科技所衍生的樂趣與便利之際，也應重視培養正確、安全的使用概念與習慣（資料來源： 2013 年 6 月 24 日，「恐怖手機電磁波青少年健康殺手？」， Yahoo 奇摩新聞 (NOWnews)）。另外，統計大師 R. A. Fisher 曾說：統計科學是塑造二十世紀人類進步的表徵，舉凡自然或社會現象，都可利用統計的科學方法，來尋求規律並分析問題的癥結所在，以謀求解決的應對政策，因此，統計可說是人類知識累積的重要憑據。面對時代腳步不斷的快速成長及講求速度和追求效率的「知識引導」時代來臨，我們知道「腦力勞動」取代「體力勞動」的新時代已經到來，統計的重要性將會更為凸顯、重要。所以，在進入二十一世紀及其未來，我們更知道統計學將會被當作是資料整理分析與各種科學研究中相當重要的一門必修學科。因此，在此我們對統計學下一個定義如下：

版權所有

華泰文化
HWA TAI PUBLISHING

「統計學」是經由蒐集、整理、呈現、分析及解釋資料，並藉著科學推論方法，在不確定性（即隨機性）的情況下由樣本資料所獲得的結果，來推論**母體的特徵**(population characteristic)，從而做出聰明決策的一門學科。

在上述的定義中，提到兩個專有名詞「樣本」與「母體」，我們會在解釋名詞做介紹，以便有助於讀者更了解統計學的定義。另外，也提供一些相關的專有名詞解釋說明。

例 1.1

如果想了解淡江大學學生之體重情形，就目前淡江大學有 28,000 名學生，則 28,000 名淡江大學學生之體重所構成的集合，即為此處所謂的「母體」。

例 1.2

承例 1.1，現從 28,000 名淡江大學學生中抽取出 2,000 名學生調查其體重，則這 2,000 名學生的體重所構成的集合，即為此處所謂的「樣本」。

例 1.3

承例 1.1、1.2，從 28,000 名學生中取得 2,000 名學生的程序或過程，即為此處所謂的「抽樣」。

例 1.4

承例 1.1，28,000 名淡江大學學生的平均體重，即為此處所謂的「參數」。

版權所有

華泰文化
HWA TAI PUBLISHING

例 1.5

承例 1.2，現從 28,000 名淡江大學學生中抽取出 2,000 名學生的平均體重，即為此處所謂的一個「統計量」。

1.2 統計資料

「統計資料」乃是由一個或多個變數之計量值或計數值所組成的，所謂「**變數**」(variable) 是指一種具有不同值或結果的特性之衡量，凡一切可計量的特性皆稱為變數，而變數的計量值或計數值稱為「變量」。例如：身高為一變數，其變量可能為 50 至 200（公分）；另外，血型、年齡、性別、物價、國別、所得、重量、膚色、分數、教育程度、等級、溫度、銷售量、……，皆為變數。

統計資料的型態是根據其變量的特性，而區分為「**計量資料**」(quantitative data) 與「**計質資料**」(qualitative data)。而計量資料（也稱之為「屬量資料」），是根據數字尺度所測量得出的資料，例如：年齡、物價、所得、身高、重量、溫度、分數、銷售量、……。至於，計質資料（也稱之為「**屬性資料**」(attribute data) 或「**類別資料**」(categorical data)），是根據資料的屬性或類別之尺度來區分的資料，例如：血型、性別、國別、等級、教育程度、膚色、……。另外，計量資料還可以再區分為「**離散資料**」（discrete data，或「間斷資料」）與「**連續資料**」(continuous data)。而離散資料的特性是**可數的** (countable)，且具有一最小的計數單位，而相鄰兩個最小計數單位的離散資料之間必有間隙 (gap)。例如：淡江大學統計學系大一 *A* 班的人數即為離散資料，其最小計數單位為 1 個人，此班可能有 60 個人，但絕不會有 60.8 人。至於，其他離散資料例子有意外事件次數、產品良品個數、排球比賽得分數、每日國產汽車銷售的台數、……。關於連續資料是由量測 (measure) 得出的，這是不可數的，且任二個連續資料之間有無限個值，亦即不存在有間隙。例如：人的體重即為連續資料，體重 50 公斤與 100 公斤之間存在有許多其他值，如 50.1、 60.25、 80.78、……。至於，其他連續資料例子有身高、長度、時間、面積、濃度、速度、……。另外，統計資料也可以根據資料的「**尺度測度**」(measure of scale) 或「**水準測度**」(measure of level) 不同，而區分為四類：「**名目尺度**」(nominal scale)、「**順序尺度**」(ordinal scale)、「**區間尺**

版權所有
華泰文化
HWA TAI PUBLISHING

度」(interval scale)及「**比例尺度**」(ratio scale)。至於，詳細的敘述可以參考第 2 章 2.2 節。

例 1.6

試判別下列資料的型態：

(1) *A* 班學生中女學生的人數。

(2) 某次國際電話的通話時間。

(3) *A* 公司之員工的教育程度。

(4) 大一新生的宗教信仰。

(5) 袋中彩球的顏色。

(6) *A* 員工一年內生病請假的天數。

(7) 台灣地區 1 月份的降雨量。

解 計質資料：(3)、(4)、(5)。

計量資料：(1)、(2)、(6)、(7)，其中離散資料為 (1)、(6)，連續資料為 (2)、(7)。

例 1.7

試區分下列哪些資料是屬於計量資料，哪些是計質資料；若為計量資料，請指出其為離散資料或連續資料：

(1) 某私立大學學生的體重。

(2) 季節。

(3) 某國小教師的年齡。

(4) 某證券公司一個月室內電話的通話數。

(5) 某人開某廠牌汽車的行車速度。

(6) 某一化學公司員工的職位等級。

(7) 每通國際長途電話的通話時間。

(8) 性別。

(9) 某國立大學碩士班學生的人數。

(10) 血型。

版權所有

華泰文化
HWA TAI PUBLISHING