Statistics

統計學

張光昭　莊瑞珠　黃必祥
廖本煌　齊學平　著

三民書局

Statistics

統計學

張光昭　莊瑞珠　黃必祥
廖本煌　齊學平　著

三民書局

國家圖書館出版品預行編目資料

統計學／張光昭,莊瑞珠,黃必祥,廖本煌,齊學平著.
——初版一刷.——臺北市：三民，2009
面；　公分

ISBN 978-957-14-5140-4　(平裝)

1.統計學

510　　　　97024055

© 統計學

著作人	張光昭　莊瑞珠　黃必祥 廖本煌　齊學平
責任編輯	楊朝孔
美術設計	謝岱均
發行人	劉振強
著作財產權人	三民書局股份有限公司
發行所	三民書局股份有限公司 地址　臺北市復興北路386號 電話　(02)25006600 郵撥帳號　0009998-5
門市部	(復北店) 臺北市復興北路386號 (重南店) 臺北市重慶南路一段61號
出版日期	初版一刷　2009年6月
編號	S 510430

行政院新聞局登記證局版臺業字第〇二〇〇號

ISBN　978-957-14-5140-4　(平裝)

http://www.sanmin.com.tw　三民網路書店

序 言

統計學是一門有趣又有用的課程，它的應用範圍極其廣泛，國內大專院校的許多科系均將其列為必修科目。通常開授於大學一年級的統計學課程，其內容較為淺顯，僅涵蓋「不涉及微積分」之統計基本概念與常用統計方法，因此俗稱為「初等統計學」(以下簡稱「初統」)。初統是一般大專院校許多科系（包括統計系、經濟系、會計系、公衛系、心理系、農藝系等等）的必修基礎科目，因此可說是一門市場廣大的重要課程。雖然初統的內容並不涉及微積分等難度稍高的數學，但對於統計學的初學者而言該課程還是有其相當的難度，尤其是一些基本且重要的統計學名詞與觀念如有限母體、統計量、抽樣分佈、均方誤差等等，經常是大多數學生「有學沒有會」的部分，而許多統計學教師卻未曾察覺這種學習情況不佳的現象。有鑑於此，本書作者針對上述諸多統計學的基本觀念以接近口述式的筆法並搭配簡易而特殊的計算範例，用細部分解的方式來詳細說明，例如在第二章對於有限母體的詳細介紹以及在第六章 6.1.2 節針對抽樣分佈與均方誤差的精闢解說，絕對是坊間大多數統計學教科書中見不到的內容。此外，本書還包含一些具有趣味性的精彩例題，例如第四章 4.3 節之例題 4.3 的「彩券收集問題」，則是本書第一作者的兒時親身經歷，可說是一個難得一見之生活化的機率統計學實例。

由於初統之內容不涉及微積分，因此修習該課程之學子往往只是「知其然而不知其所以然」，例如只知道最小平方估計量之數學公式，卻不瞭解該公式之牽涉到偏微分的詳細數學推導過程。或許對於「非統計系」的大專學生而言，「知其然而不知其所以然」的初統就已足夠（甚至已經使得學子們頭痛不已）；但對於統計系的學生而言，初統僅是最基本的一門統計課程而已，而真正「夠分量」且最重要的統計課程，則非「數理統計學」莫屬了！但是數理統計學之內容，就一般大專學生而言，往往又過於艱澀，不僅一般大專學生不易看懂，就連研究生甚至教師們都感到很吃力呢！那麼，在初統之內容十分淺顯而數理統計學又過於艱深之情況下，一些具有智慧與眼光的早期國內統計學者就精心設計出「高等統計學」(以下簡稱「高統」)

這門特殊課程，開授於大學統計系二年級，其難度介於初等統計學與數理統計學之間，而其內容雖然牽涉到大量的微積分，但僅限於一般之計算或較簡易之證明與推導，而並不涵蓋數理統計學中較為抽象且難懂之「充分性」(sufficiency) 與「完全性」(completeness) 等概念，因此高統是銜接大學統計系一年級的初統與三年級的數理統計學的一門「橋樑」課程。這門精心設計的橋樑課程，的確為一些早期國內統計系奠定了良好的基礎，就以本書第一作者任教多年的輔仁大學統計資訊系（原名為統計系）而言，在十餘年前徐步堯主任的六年任期內，由於高統及高等微積分等優良必修課程之開授，使得社會組背景的輔大統計系學生數學程度大幅提升，因此修習大三的數理統計學亦能得心應手，進而培育出多位「博士級」的傑出系友，例如任教於美國 University of California at Santa Barbara 統計系的徐世健教授、國內國立中央大學統計研究所的陳玉英教授、逢甲大學統計與精算研究所的林麗芬教授、國立清華大學經濟系的黃明堆教授、國立政治大學統計系的鄭宗記教授、輔仁大學統計資訊系的梁德馨教授與侯家鼎教授等等諸位優秀統計學者，都是徐步堯主任時期培育出的傑出系友。

針對初統之內容會讓學生產生「知其然而不知其所以然」的問題，本書作者於撰文過程中摻入非常少量的微積分（其中大部分列為「選讀」)，以期讀者能夠稍微地「知其所以然」。如果初統的難度計量標準是 1.0，高統的難度計量標準是 2.0，那麼本書的難度就大約是 1.1 吧！因此，適合閱讀本書的讀者群極為廣泛，包括大專院校任何科系的一或二年級學生、技職學校的中高年級學生、任何具有高中數學程度的初學者等等。

本書作者希望莘莘學子能夠藉由閱讀本書而順利地進入優良學習狀況，並一掃對於統計學的困惑與畏懼，進而產生學習興趣與信心。

張光昭、莊瑞珠、黃必祥、廖本煌、齊學平

（以上作者群依撰寫章節次序排列）

2009 年 8 月

統計學

目次

第一章 預備知識

學習重點

1. 空集合是任何一個集合的部份集合。
2. 一個集合 S 之所有部份集合所形成的集合稱為 S 的冪集合。
3. 函數是兩個集合之間的一種特殊對應關係，它容許「多對一」，但不容許「一對多」。
4. 兩個事件如果不可能同時發生，則稱此二事件為互斥。
5. 獨立事件與互斥事件是兩個完全不相同的概念。

前言 統計學是一門既有趣又有用的課程，對於大專院校低年級的初學者而言，所需要的主要預備知識有：國中和高中數學的集合與函數之概念、排列組合理論，還有一些簡易的機率觀念等。若初學者具有微積分課程的基礎，或是於修習統計學課程的同時也修習微積分，則將有相當大的助益。在本書的第一章，我們將逐一復習、整理與介紹初學者所需的預備知識，以期莘莘學子能夠順利地進入優良學習狀況並成功地讀完本書。

專用字詞與術語

集合 (set)、元素 (element)、空集合 (empty set)、部分集合 (subset)、真部分集合 (proper subset)、有限集合 (finite set)、冪集合 (power set)、有理數 (rational numbers)、無理數 (irrational numbers)、實數 (real numbers)、交集 (intersection)、聯集 (union)、差集 (difference)、補集 (complement)、范氏圖形 (Venn diagram)、交換性 (commutative property)、結合性 (associative property)、分配性 (distributive property)、迪摩根定律 (DeMorgan's laws)、函數 (function)、定義域 (domain)、對應域 (counterdomain/codomain)、值域 (range)、映至 (into)、映成 (onto)、一對一 (one-to-one)、二元關係 (binary relation)、乘積集合 (product set/Cartesian product)、序對 (ordered pair)、自變數 (independent variable)、應變數 (dependent variable)、斜率 (slope)、截距 (intercept)、相乘法則 (multiplication rule)、排列 (permutations)、組合 (combinations)、n 階乘 (n factorial)、環狀排列 (circular permutations)、二項式定理 (binomial theorem)、二項係數 (binomial coefficients)、機率 (probability)、隨機實驗／統計實驗 (random experiment/statistical experiment)、樣本空間／出象空間 (sample space/outcome space)、樣本點／出象 (sample point/outcome)、事件 (event)、空事件 (null event)、必然事件 (sure event)、基本事件 (elementary event)、古典機率 (classical probability)、互斥事件 (mutually exclusive events)、集合函數 (set function)、機率函數 (probability function)、條件機率 (conditional probability)、獨立事件 (independent events)、分割 (partition)、總機率定理 (theorem of total probability)、貝氏定理 (Bayes' theorem)。

1.1 集合與函數

一、集　合

在本書的第一章，我們首先復習一下國中數學課程中的集合與函數之概念。所謂「集合」，就是由一群「元素」所構成的群體；這些元素可能是一群人、一群動物、一群數字、一群物品等等。在數學的書籍中，通常使用 A、B、C、D 等大寫英文字母以及大括號 $\{\cdots\}$ 來表示集合。若 a 為集合 A 的一個元素，則謂「a 屬於 A」，以符號「$a \in A$」表示之。

例題 1.1

我們舉幾個很簡單的集合如後：$A=\{1, 2, 3, 4, 5, 6\}$、$B=\{2, 4, 6, 8, 10\}$、$C=\{$孔子 , 孟子 , 莊子 , 荀子$\}$、$D=\{$孔子 , 孟子 , 莊子 , 荀子 , 蘇格拉底 , 柏拉圖$\}$、$E=\{$臺北市 , 臺中市 , 臺南市 , 高雄市 , 花蓮市$\}$。集合 A 有六個元素，可用符號表示為 $n(A)=6$；同理，$n(B)=5$、$n(C)=4$ 等等。

（一）各類集合介紹

1. 空集合

若一個集合中沒有任何元素，則稱之為空集合，以符號 $\varnothing$ 表示。

2. 子集合（部分集合）

若某集合中的所有元素都在另一個集合中，則這個某集合就是另一個集合的部分集合（或稱子集合），若 A 為 B 的部分集合，則記為 $A \subseteq B$。空集合是任何一個集合的部分集合；任何一個集合是它本身的部分集合。

3. 真部分集合

若某集合 C 中的所有元素都在另一個集合 D 中，並且 $C \neq D$，則集合 C 就是集合 D 的一個真部分集合，譬如例題 1.1 之 C 集合中的所有元素（共四個：孔子、孟

子、莊子、荀子）都在 D 集合中，因此 C 是 D 的一個真部分集合，記為 $C \subset D$ 或 $D \supset C$。

4. 有限集合

若一個集合 S 中的元素總個數為有限個，亦即 S 為一個**有限集合**❶，則 S 的部分集合共有 $2^{n(S)}$ 個（其中「真部分集合」共有 $2^{n(S)}-1$ 個)。

5. 冪集合

一個集合 S 之所有部分集合所形成的集合稱為 S 的**冪集合**，以符號 2^S 表示，所以 $2^S=\{A \mid A \subseteq S\}$，且 $n(2^S)=2^{n(S)}$。

在數學的課程中，有一些常用的數字類型集合如下：

$\mathbb{N}=\{1,2,3,\cdots\}$ = 所有正整數（或自然數）所形成的集合

$\mathbb{I}=\{\cdots,-3,-2,-1,0,1,2,3,\cdots\}$ = 所有整數（正或負）所形成的集合

$\mathbb{E}=\{\cdots,-6,-4,-2,0,2,4,6,\cdots\}$ = 所有偶數（正或負）所形成的集合

若一個集合中的所有元素具有某種共同型態或屬性，則可用 $\{\cdots \mid \cdots\}$ 之符號來表示，例如上述之所有偶數所形成的集合可寫為：

$$\mathbb{E}=\{\cdots,-6,-4,-2,0,2,4,6,\cdots\}=\{x \mid x=2i,\ i \in \mathbb{I}\}$$

又如：

$$\mathbb{Q}=\{x \mid x=\frac{m}{n};\ m,n \in \mathbb{I},\ n \neq 0\} = \text{所有「有理數」所形成的集合}$$

還有一些常用的數字類型集合很不容易使用大括號之符號來表示，例如：$\mathbb{Q}$ = 所有「**實數**」所形成的集合與 $\mathbb{Q}'$ = 所有「**無理數**」所形成的集合，就是兩個不易使用 $\{\cdots\}$ 或 $\{\cdots \mid \cdots\}$ 之符號來表示的集合。

（二）常用的集合運算

1. 交　集

集合與集合之間有一些常用的運算，如**交集**、**聯集**、**差集**（或稱**補集**）等等。

❶ 請參閱 Bartle (1976) 一書 22–23 頁。

集合 A 與 B 的交集就是由它們共同擁有的元素所構成的集合，可用符號與數學式表示為：

$$A \cap B = \{x \mid x \in A \text{ 且 } x \in B\}$$

2. 聯　集

集合 A 與 B 的聯集就是由它們合併之後的所有元素所構成的集合；若 A 與 B 擁有共同的元素，則這些共同的元素在聯集中不可重複出現。A 與 B 的聯集可用符號與數學式表示為：

$$A \cup B = \{x \mid x \in A \text{ 或 } x \in B\}$$

3. 推展至 n 個集合

如果有很多個（n 個）集合 $A_1, A_2, \cdots, A_n$，則它們的交集與聯集可表示如下：

$$A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i = \{x \mid x \in A_i, \forall i, i = 1, 2, \cdots, n\}$$

$$A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^{n} A_i = \{x \mid \exists i \text{ 使得 } x \in A_i, i = 1, 2, \cdots, n\}$$

在以上數學式中的 ∀ 與 ∃ 都是常用的數學符號，分別代表「所有的（或任何一個）」與「存在某至少一個」之意。

4. 差集（補集）

集合 A 與 B 的差集就是由所有屬於 A 但卻不屬於 B 的元素所構成的集合，可用符號與數學式表示為：

$$A - B = \{x \mid x \in A \text{ 且 } x \notin B\}$$

在某些數學書籍中❷，將差集 $A - B$ 稱為「集合 B 對於集合 A 的補集」。在數學與統計學書籍中時常使用**范氏圖**形來表示集合之間的運算結果，如下所示：

❷ 例如 Bartle (1976)。

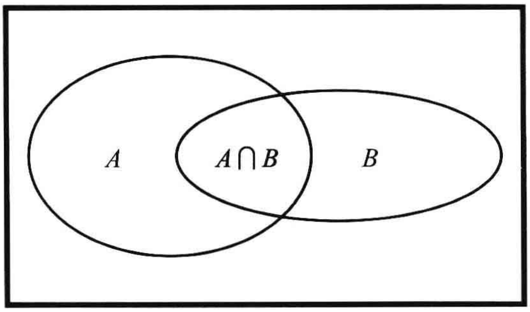

圖 1.1 $A \cap B$ 之范氏圖形

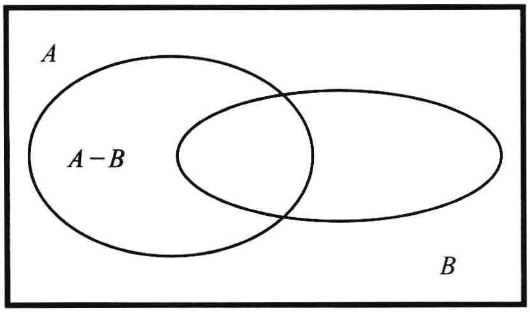

圖 1.2 $A - B$ 之范氏圖形

（三）集合運算的特性

1. 交換性與結合性

關於交集、聯集、差集這三種運算，交集與聯集都具有交換性與結合性，可用數學式表示為：

$$A \cap B = B \cap A、(A \cap B) \cap C = A \cap (B \cap C)$$

$$A \cup B = B \cup A、(A \cup B) \cup C = A \cup (B \cup C)$$

但差集則不具有以上兩種性質，因此在一般情況下：

$$A - B \neq B - A、(A - B) - C \neq A - (B - C)$$

2. 分配性

此外，交集與聯集之間還具有分配性如下：

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

在例題 1.1 中，可以容易地看出：$A \cap B = \{2, 4, 6\}$、$A \cap C = \varnothing$、$A \cup B = \{1, 2, 3, 4, 5, 6, 8, 10\}$、$A - B = \{1, 3, 5\}$ 等等。

3. 重要公式

關於交集、聯集、差集這三種運算，還有幾個基本且重要的公式如下：

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) \tag{1.1}$$

$$A=(A\cap B)\cup(A-B) \tag{1.2}$$

$$A-(B\cup C)=(A-B)\cap(A-C) \tag{1.3}$$

$$A-(B\cap C)=(A-B)\cup(A-C) \tag{1.4}$$

公式 (1.3) 與公式 (1.4) 就是在集合論課程中赫赫有名的迪摩根定律。此外，公式 (1.1) 可以延伸到三個集合之情況如下：

$$n(A\cup B\cup C)=n(A)+n(B)+n(C)-n(A\cap B)-n(A\cap C)-n(B\cap C)+n(A\cap B\cap C) \tag{1.1a}$$

公式 (1.1) 還可繼續延伸到很多個集合之情況。

二、函　數

關於國中數學之集合論的基本概念，我們復習到此處暫時告一段落。接下來，我們復習一下函數的基本概念。函數有不止一種的定義或說明方式，分別敘述如下。

（一）第一種函數說明方式

函數是兩個集合之間的一種特殊「對應關係」，這種對應關係從兩個集合中的第一個集合出發（此第一個集合稱為定義域），對應到第二個集合（此第二個集合稱為對應域），且必須滿足以下兩個條件：

(1)定義域中的每一個元素都必須對應到對應域中的某一個元素。

(2)定義域中的每一個元素對應到對應域中的元素不得超過一個。

此二條件可合併為「定義域中的每一個元素都必須對應到對應域中的某恰好一個元素」。由以上條件可知：定義域中的若干個不同的元素有可能對應到對應域中的同一個元素；換言之，以上條件**容許「多對一」，但不容許「一對多」**。此外，以上條件不表示函數的對應一定會涵蓋對應域中的所有元素，而對應域中被涵蓋（即被對應到）的元素稱為「函數值」，所有被對應到的元素（即所有的函數值）所形成的集合稱為函數的值域，值域是對應域的部分集合。凡是滿足以上條件(1)與條件(2)的函數通稱為映至函數。若一個函數的值域與其對應域相同，則此種類型的函數稱為映成函數。若一個函數之定義域中的任意兩個不相同元素一定對應到對應域中的兩個

不相同元素，則此種類型的函數稱為一對一函數。函數的符號通常為英文斜體小寫字母 f、g、h 等等，有些情況下也會用到大寫字母 F、G、H 或 X、Y、Z 等。

例題 1.2

我們利用例題 1.1 中的集合來製造幾個很簡單的函數。首先我們來製造一個很普通的映至函數 f，其定義域與對應域分別為 A 與 B，可用符號表示為 $f: A \to B$，其對應情況如下：定義域 A 中的前四個元素 $1, 2, 3, 4$ 均對應到對應域 B 中的同一個元素 8，定義域 A 中的最後兩個元素 $5, 6$ 均對應到對應域 B 中的同一個元素 10，這些對應可用「箭號」的方式表示為：

$$f: \begin{cases} 1, 2, 3, 4 \to 8 \\ \quad\ 5, 6 \to 10 \end{cases}$$

或用數學式表示為「$f(1) = f(2) = f(3) = f(4) = 8,\ f(5) = f(6) = 10$」。從以上函數 f 的對應情況可看出，f 既不是映成也不是一對一函數，而僅只是一個「陽春型」的映至函數。接著，我們來製造一個映成函數 g，其定義域與對應域分別為 A 與 C，其對應情況為：

$$g:\quad A \quad \to C$$

$$g: \begin{cases} 1, 2 \to \text{孔子} \\ 3, 4 \to \text{孟子} \\ \quad 5 \to \text{莊子} \\ \quad 6 \to \text{荀子} \end{cases}$$

或用數學式表示為「$g(1) = g(2) =$ 孔子，$f(3) = f(4) =$ 孟子，$f(5) =$ 莊子，$f(6) =$ 荀子」。最後，我們來製造一個一對一函數 h，其定義域與對應域分別為 C 與 E，其對應情況為：

$$h:\quad C \quad \to E$$

$$h: \begin{cases} \text{孔子} \to \text{臺北市} \\ \text{孟子} \to \text{臺中市} \\ \text{莊子} \to \text{臺南市} \\ \text{荀子} \to \text{花蓮市} \end{cases}$$