

統計學

科技資料分析方法

著者
方純

興業圖書股份有限公司

4

學 計 統

著 者
編 方
純

(國立成功大學統計學教授)

興業圖書股份有限公司

大專用書 **統 計 學**

版權所有・翻印必究

中華民國六十三年九月一版
中華民國六十四年九月修訂二版

中華民國六十六年 三版

中華民國六十七年 四版

中華民國六十八年 五版

中華民國六十九年 再版

編著者：方 純
發行人：王 志 康

出版者：
印刷者：興業圖書股份有限公司

地 址：台南市勝利路一一八號

電 話：（〇六二）三七三二五三號

郵 撥：南字三一五七三號

本公司經行政院新聞局核准登記，發給出版事業登記證局版台業字第零肆壹零號

定價貳佰元

近代統計學術發展之里程碑

(代序)

- 1713 年 白努里 (J. Bernoulli) 詳論二項分配 (The binomial distribution)。
- 1733 年 棣美弗 (De Moivre) 發現常態機率分配 (The normal probability distribution)。
- 1763 年 貝氏 (R. Bayes) 提供應用於歸納推計之一機率定理，目前命名貝氏定理。
- 1754 年 奧國於 1754 年，美國於 1790 年，英國與法國於 1801 年分別舉辦第一次人口普查。
- 1809 年 高斯 (Gauss) 闡述觀測值之誤差理論。
- 1812 年 拉普拉斯 (Laplace) 出版“機率基本論”(Theorie analytique des probalites)。
- 1837 年 卜瓦松 (Poisson) 發表目前以其名命名之卜氏分配。
- 1863 年 介達士 (W. S. Jevons)、拉斯培勒 (E. Laspyres)
- 1874 年)、裴許 (H. Pausche) 及艾幾華斯 (Edge worth) 等發表測量經濟變化之物價指數理論與公式。
- 1876 年 赫爾瑪 (Helmert) 導出亥方 (χ^2) 分配 (Chi-square distribution)。
- 1886 年 高爾登 (F. Galton) 首創迴歸 (regression) 觀念。
- 1888 年 何瑞禮 (H. Hollerith) 發明電動打孔卡片分類列表體系。
- 1896 年 皮爾遜 (K. Pearson) 首創線性相關係數 r 之用法。

- 1900 年 皮爾遜 (K. Pearson) 根據亥方分配，明確系統地闡明配合適度檢定法 (test of goodness of fit)。
- 1908 年 高賽德 (W. S. Gosset) 用筆名“學生 (Student)”發表 t 分配。提供一種根據樣本資料估計均數 (mean) 的有效檢定法。
- 1911 年 費雪 (I. Fisher) 發表指數之理想公式，並創時間數列分析方法。
- 1915 年 費休 (R. A. Fisher) 導出線性相關係數 r 之抽樣分配。
- 1921 年 費休 (R. A. Fisher) 首創極大可能原則 (The Principle of maximum likelihood)，並導出 r 之 z 分配。
- 1924 年 費休 (R. A. Fisher) 推廣常態， t 及亥方分配至 z 分配，並藉使變異數分析 (analysis of variance) 技術系統化。
- 1924 年 修哈特 (W. A. Shewhart) 應用統計學於大量生產之品質管制，發明第一張管制圖。
- 1925 年 費休 (R. A. Fisher) 擴展 t 分配範圍至包涵迴歸係數。
- 1927 年 鐵貝特 (L. H. C. Tippett) 接受皮爾遜建議，編製發表第一本隨機號碼表 (table of random numbers)。
- 1929 年 道奇 (H. F. Dodge) 與羅明 (H. G. Roming) 首創抽樣檢驗之基本概念與理論。
- 1928 年 牛曼 (J. Neyman) 與小皮爾遜 (E. S. Pearson 係 1933 K. Pearson 之子) 發展統計假說之檢定理論 (theory of testing statistical hypotheses)，導入工業生產、廣被採用。

- 1933 年 柯莫果夫 (A.Kolmogorov) 發表“機率論基礎”(Foundations of the theory of probability) 奠定統計理論之邏輯基礎。
- 1934 年 牛曼 (J.Neyman) 傾力研究分層隨機抽樣之最優配置 (optimum allocation) 條件。
- 1935 年 費休 (R.A.Fisher) 名著“實驗設計 (The Design of Experiments) 出版。
- 1937 年 史迺德 (G.W. Snedecor) 改進其師費休之 z 分配，並命名 F 分配以資紀念。
- 1940 年 發展從廣大地區 (全體) 抽樣之方法。
- 前 後
- 1943 年 華德 (A.Wald) 創造逐次抽查法 (sequential sampling) 美國列為國防機密，至戰後 1945 年始允公開。
- 1944 年 艾肯 (H.Aiken) 設計 Mark 1 號巨型機械計算機問世。此機為電子計算機前身，由 IBM 公司製造贈送哈佛大學。
- 1945 年 艾克梯 (J. Eckert) 與麥克萊 (J. Mauchly) 發明公認第一部電子計算機“電動數字累積計算器 (ENIAC)”，但無內儲記憶。
- 1945 年 華德 (A.Wald) 完成統計推計學之一般形式。提供最佳判定之統計定義，並決定最強統計判定函數 (Most powerful statistical decision function)。
- 1950 年
- 1946 年 美國品質管制學會在紐約成立，並出版“工業品質管”(Industrial Quality Control) 月刊。自 1968 年起更名“品質發展”(Quality Progress) 繼續發行。

- 1952 年 英國完成第一部內儲程式電子計算機。
- 1953 年 I B M裝設電晶體與積體印刷電路之第二代、第三代計算機相繼問世，廣泛應用於軍事與商業。
- 1960 年 美國Hewlett Packard公司袖珍型H. P. 65 內儲程式計算器問世，使近代電子計算機開始走進私人家庭。
- 1972 年 H. P. 65 為第一部不需特殊計算機語言，即可作計算程式（一百個步驟內），採磁卡插入按鈕式掌上（僅重11兩）計算機。其功能除解決一般數學、工程與醫藥等技術問題外，可做變異分析、迴歸、相關分析及各種機率分配等 37 種有關統計分析程式問題。

方 純

民國六十三年九月成功大學

統計學

目次

近代統計學術發展之里程碑（代序）

第一章	集合與概率	1—32
1.1	集合與子集	1
1.2	集合的運算	4
1.3	樣本空間	8
1.4	樣點的點計	10
1.5	機 率	15
1.6	幾個機率定理	17
1.7	條件機率	20
1.8	貝氏定理	24
	習 題	27—32
第二章	隨機變數	33—81
2.1	隨機變數的概念	33
2.2	不連續機率分配	34
2.3	連續機率分配	39
2.4	經驗分配	42
2.5	連合機率分配	46
2.6	數學期望值	56
2.7	期望值的定理	62
2.8	特殊的數學期望值	66

2.9	變異數的特性	71
2.10	乞比希夫定理	73
	習 題	76—81
第 三 章	重要不連續機率分配	82—111
3.1	概 論	82
3.2	均質分配	82
3.3	二項分配與多項分配	85
3.4	超幾何分配	92
3.5	卜瓦松分配	100
3.6	負二項分與幾何分配	105
	習 題	108—111
第 四 章	重要連續機率分配	112—142
4.1	常態分配	112
4.2	常態曲線下的面積	115
4.3	二項分配的常態近似值	124
4.4	伽瑪指數與亥方分配	130
4.5	韋布分配	135
	習 題	139—142
第 五 章	隨機變數的函數	143—181
5.1	變數的轉換	143
5.2	動差母函數	153
5.3	隨機抽樣	160
5.4	抽樣原理	162
5.5	均數的抽樣分配	169

5.6	$(n-1)S^2/\sigma^2$ 的抽樣分配	173
5.7	t 分配	176
5.8	F 分配	181
	習題	185—190
第 六 章	估計原理	191—228
6.1	引言	191
6.2	傳統的估計方法	192
6.3	均數的估計	195
6.4	二均數差的估計	200
6.5	比率的估計	210
6.6	兩比率差的估計	215
6.7	變異數的估計	217
6.8	二變異數比的估計	219
6.9	貝氏估計方法	222
6.10	判定原理	228
	習題	231—237
第 七 章	假說的檢定	238—296
7.1	統計假說	238
7.2	型 I 與型 II 誤差	239
7.3	單尾與雙尾檢定	248
7.4	均數與變異數的檢定	250
7.5	檢定均數所用樣本大小的選取	257
7.6	魏克生雙樣本檢定法	262
7.7	魏克生成對觀測值檢定	269
7.8	比率的檢定	274

7.9	兩比率差的檢定	217
7.10	配合適度之檢定	280
7.11	獨定性之檢定	284
	習題	288—296
第 八 章	線型迴歸與相關	297—326
8.1	線型迴歸	297
8.2	簡線型迴歸	299
8.3	信賴界限與顯著性檢定	302
8.4	迴歸模式之選擇	313
8.5	變異數分析法	314
8.6	反應之重複觀測	316
8.7	相 關	322
	習題	327—332
第 九 題	複線型迴歸	333—361
9.1	緒 言	333
9.2	最小平方估計量	335
9.3	誤差變異數之估計	339
9.4	複線型迴歸之推計	343
9.5	模式之配合度	347
9.6	正交性之特例	352
	習題	358—361
第 十 題	多項式迴歸	362—379
10.1	緒 言	362
10.2	單個自變數	363

10.3	假說檢定	364
10.4	正交多項式之應用	365
10.5	等級距觀測值實驗	369
10.6	多個自變數	373
	習題	378 — 379
第十一章	變異數分析	380 — 431
11.1	緒言	380
11.2	一因子變異數分析	382
11.3	多個變異數相等之檢定	390
11.4	單自由度之對比	393
11.5	多變矩檢定法	397
11.6	區集內一組處理之比較	399
11.7	隨機完全區集實驗	400
11.8	關於隨機完全區集實驗的附加說明	409
11.9	隨機影響模式	411
11.10	變異數分析之迴歸近似	417
11.11	變異數分析測驗的檢定力	421
	習題	428 — 431
第十二章	因子實驗	432 — 465
12.1	二因子實驗	432
12.2	二因子實驗的交互作用	432
12.3	二因子變異數分析	434
12.4	三因子實驗	445
12.5	特殊多因子模式之討論	454
12.6	型Ⅱ模式的因子實驗	456

12.7	樣本大小之選擇	460
	習題	462—465
第十三章	2 ^k 因子實驗	466—493
13.1	引言	466
13.2	變異數分析	467
13.3	葉特之計算對比法	470
13.4	不完全區集之因子實驗	475
13.5	部分混合	481
13.6	分段因子實驗	483
13.7	分段因子實驗的分析	486
	習題	489—493
	參考書目	494—495
	統計表	496—528
	習題答案	529—538

第一章 集合與概率

1-1 集合與子集

研究概率與統計需要數學，而**集合**(Sets)概念是各種數學的基礎。所謂集合可視為有定義事物的總括。例如，台灣省的河流，周期表中的化學元素，立法院的委員，鞋店每月收入等，均成集合。一線段也可視為無限個點組成的集合。集合中每一個體稱為該集合的**元素**(Element)。但香花、美女、大數、重石等均不成集合。

通常以大寫字母如A、B、X或Y表集合，小寫字母如a、b、x或y表集合中的元素。為使集合有明確意義，對任一已知個體，必需有判斷該個體是否為某集合中元素的方法。

描述集合的方法有二：

第一，若某集合僅含有限個元素，則將各元素列舉在大括號內，且各元素間用逗點分開。因此如集合A含有某工廠過去發生意外事件次數，則可寫成

$$A = \{ 3, 1, 4, 2 \},$$

又如從生產線上任抽兩件產品，檢查結果則可寫成

$$B = \{ D, N \},$$

其中D、N分別表劣與良。

第二，集合可用敘句或規則說明。例如，C表世界上人口超過百萬的城市。若x為C的任一元素，則可寫成

$$C = \{x \mid x \text{ 是人口超過百萬的城市}\},$$

讀作“ C 是由所有 x 組成元集合，而 x 表世界上人口超過百萬的城市”。同樣，若 P 表半徑是2的圓上點 (x, y) 的集合，則可寫成

$$P = \{(x, y) \mid x^2 + y^2 = 4\}。$$

一集合宜用何種方式描敘，應隨討論問題的不同而異。要將藍眼珠的人列舉一集合內實在困難；但另一方面，想用一敘句或規則說明集合

$$Y = \{\text{書、人、哲學家、鉸釘}\},$$

也不容易。

集合符號中，符號 $\in$ 的意思是“是……的一元素”或“屬於”； $\notin$ 的意思是“不是……的一元素”或“不屬於”。若 x 是集合 A 中的一元素，而 y 不是則寫成

$$x \in A, y \notin A。$$

例 1.1 設 $A = \{2, 4, 6, 8\}$ 與 $B = \{x \mid x \text{ 可被 } 3 \text{ 整除}\}$ 。則
 $4 \in A, 9 \in B, 8 \notin B$ 與 $3 \notin A$ 。

定義 1.1 若二集合中的元素完全相同，則稱此二集合相等。

若集合 A 等於或恆等於集合 B ，則屬於 A 的每一元素也屬於 B ；而屬於 B 的每一元素，也必屬於 A ，並以 $A = B$ 表示。但若 A 或 B 中至少有一元素非二集合所共有，則此二集合不等，並寫成 $A \neq B$ 。

例 1.2 設 $A = \{1, 3, 5\}$ ， $B = \{3, 1, 5\}$ 與 $C = \{1, 3, 5, 7\}$ ，則 $A = B$ ， $A \neq C$ 與 $B \neq C$ 。注意：集合並不因所含元素順序不同而異。

定義 1.2 不含元素的集合稱為零 (null) 集合或空 (empty) 集合，以符號 ϕ 表示。

若 A 表生物實驗中，用肉眼觀察發現的微生物集合，則 A 必為空集合。又若 $B = \{x \mid x \text{ 為 } 7 \text{ 的非質因數}\}$ ，則 B 也是空集合，因 7 的可能因數僅有質數 1 與 7。

設集合 $A = \{x \mid 0 < x < 5\}$ ，集合 $B = \{y \mid 2 \leq y \leq 4\}$ ，其中 x 與 y 均表整數。顯見集合 B 中每一元素都是集合 A 的元素。集合 B 稱為 A 的**子集** (Subset)。用符號寫成 $B \subset A$ ；其中符號 $\subset$ 意義是“是……的子集”或“包含於”。

定義 1.3 若集合 A 的每一元素都是集合 B 的元素，則稱 A 是 B 的子集。

據此定義，每一集合是其本身的子集。任一集合的子集，不是該集合的本身，則稱其為該集合的**真子集** (Proper Subset)。故若 $B \subset A$ 且 $B \neq A$ ，則稱 B 是 A 的真子集。

例 1.3 集合 $B = \{2, 4\}$ 是集合 $A = \{1, 2, 3, 4, 5\}$ 的真子集。但集合 $C = \{3, 2, 5, 1, 4\}$ 是 A 的子集而不是真子集，因 $A = C$ 。

許多問題中，所有集合都是某一種特殊集合的子集，此特殊集合稱為**全集** (Universal Set)，並常用 U 表示。在研討某項數學問題時，可認定全部實數是全集。若所有大專學生的智商 ($I.Q.$) 為全集，某所大學學生的智商則為其子集。

例 1.4 全集 $U = \{1, 2, 3\}$ 的所有子集是 $\{1, 2, 3\}$, $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1\}$, $\{2\}$, $\{3\}$, ϕ 。注意含三個元素的全集計有 $2^3 = 8$ 個子集。一般而言，含 n 個元素的全集，計有 2^n 個子集。

各子集與對應全集間的關係，可用維恩圖解(Venn Diagram)表示。在維恩圖解中，用一矩形表全集，其中各圖形分別表一子集。如在圖 1.1 中， A ， B 與 C 都是全集 U 的子集。顯見 $B \subset A$ ， B 與 C 無共同元素， A 與 C 至少有一共同元素。

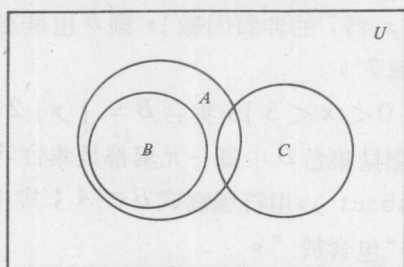


圖 1.1 U 的子集

有時為清楚起見，將圖解中各面積分別用不同的陰影或線條區分。如圖 1.2 中，全部大學生為一全集，讀數學與讀物理的學生是其子集，並分別塗抹不同方向斜線，交叉線部份面積表選讀數學與物理兩課的學生，其餘空白部份表示未選讀此兩科之學生。

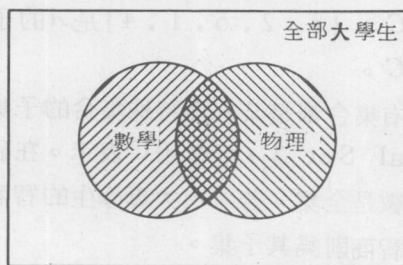


圖 1.2 用不同斜線區分子集

1-2 集合的運算

現在準備討論幾種集合的運算，運算結果可產生新集合，而諸新