

# 統計學

(第三次修訂初版)

**RONALD E. WALPOLE** 著

祁和福校訂

劉瑞雪譯述

# 統計學

(第三次修訂初版)

**RONALD E. WALPOLE** 著

祁和福校訂

劉瑞雪譯述

## 第三次修訂版譯者序

本書爲原著於1982年第三次修訂版之譯本，修訂後之內容，在著者序（請看著者序）中，均已說明，於此不另詳述。此次改版較第二版多兩章（共十三章），內容亦增加很多，篇幅約增二百頁，對基本觀念之說明，更加清楚。

本書初版於一九六八年問世，本人於一九七〇年譯成中文，第二版於一九七五年修訂，譯文於一九七六年八月完成。多年來採用此中文本爲教材者不少，承採用教師之建議，最好增加時間數列與指數，以免講授時再予補充，茲值改版，乃將此項材料列爲本書之第十四章（第二版中列爲第十二章），以供需用者參考。

譯者

民國七十二年二月

## 初版譯者序

隨著科學的進步與工商業之發展，社會對知識的需要，愈來愈新穎，愈廣泛，愈深入。大學基本應用學科教材之內容，除應滿足此種需要外，更應有所增進。如此，學術與實際應用，方可收互相促進之效。

近數年來，歐美各國新出版之統計書籍，編寫方式與以前大不相同，內容亦有很多差別。編寫重點，在詳細說明基本概念與統計推論方法之應用，而敘述統計僅講述其必須應用之方法。

本書係一九六八年美國麥克米蘭公司出版之大學普通統計學教本，其對基本概念之講解，簡明扼要，且用有系統之符號表示樣本與全體之關係；此項符號系統，對基本概念之了解，可有不少之幫助。至統計推論方法之應用，列舉甚豐，對目前之一般應用或無不足之感。

按此方式編寫之中文統計書籍，尚不甚多，故將此書譯成中文，以供參考。本書文字係按意譯，倘有與原意不合之處，敬祈讀者隨時指正。

此書譯成後，承國立政治大學統計研究所所長祁和福教授詳為校閱，並予訂正，衷心感激，謹此致謝。

譯者

## 著 者 序

本書第三版與第一版及第二版之寫法相同，均按教科書之形式編寫，可作大學各學系統計學之教本，例題與習題均選自各應用學科；如社會學，經濟學，心理學，企業管理學與其他理工學科。讀者如有高中代數之知識，即可瞭解本書之敘述統計與統計推論的內容；倘讀完大學一年級之數學並有一學期之初等函數論或有限數學的基礎，可獲益更多。

第三版中之習題與例題的數字測度單位，皆改用萬國制單位（公尺，公斤，公升）。前兩版之習題，均置於每章之末，本版習題改置於適當章節之後；且有很多例題與習題皆為新增。在習題中所說之基本觀念或方法，教師可斟酌取捨，與本書同時出版者，有學生學習指導，其中有各種類型之題解，另行發售。

關於敘述統計之觀念，以前分散於各章說明，本版則集中於開始三章。為補充前兩版之不足，新增百分位數，十分位數與四分位數，且分別按分組與未分組資料說明計算方法，又說明皮爾生偏斜係數，及增加用隨機數字表取隨機樣本之方法。現代統計之重點為統計推論，其理論基礎為機率原理；故於第四章即講述機率之粗淺概念，並用集合符號闡釋。第五章介紹隨機變數及其期望數與離散、連續機率分配並敘明此等分配之各種特性，關於期望數之討論，本版已完全重寫修訂。

離散機率分配導論，說明於第六章，常態分配介紹於第七章，關於抽樣理論，估計原理，假設之檢定，分別討論於第八、九、十各章，第八章業經修訂並增添模擬試驗及系統、分層與集體抽樣方法。卡方與F分配之應用，則在第九章中研究，關於二項母數與常態平均數之

貝氏估計法，亦於第九章新增一節單獨討論。第十一章先討論配合曲線之一般方法，次討論迴歸原理；再討論增加篇幅之相關分析與顯著性檢定。另增一節討論偏相關與複相關。第十二章討論變異數分析與試驗設計。試驗設計包括隨機區集與拉丁方塊設計，此為前兩版所無。有些作者以迴歸方法討論變異數分析，本書未用此法研究。此二章分別單獨進行講授，無相互應用之處。為加重無母數檢定之重要性，已增加材料，單獨講述於第十三章。

本書所含之材料甚豐，可隨時間之長短酌量取捨。如每週授課三小時，一學期課程應包括第一章至第五章的大部份，第六章之6.1，6.2與6.3；第七章至第十章，其中可略去第九章之9.9與9.10節；第十章之10.8，10.9與10.10節；第十一章中應講授11.1，11.2，11.3，11.4與11.8諸節。

對於協助本人寫成此書之各位，均甚感激。尤對龍大海核與金·馬克思二位將手稿打字與校對並作解答手冊；對威廉D. 歐格寫大部份學生學習指導；更為感激；對麥克米蘭出版公司協助出版，對用此書前兩版為教材之諸教師的建議與批評，均在此表由衷之感激。

承蒙爾立勿公司允許自R. A. Fisher之Statistical Methods for Research Workers書中抽印表A.5，皮爾生(E.S. Pearson)教授與Biometrik 期刊社之允許，自Biometrik Tables for Statisticans中抽印表A.6與表A.7，奴突蘭公司之同意自Molin之Poisson Exponential Binomial Limit中抽印表A.3，又分別得數理統計期刊社，印地亞那大學教育研究期刊社，美國賽納密公司出版物與Biometrics 期刊社之同意，抽印表A.8，表A.9，表A.10，表A.11與表A.14，美國統計學會期刊社之同意，抽印表A.3，本人對此項概允抽印，均表誠摯之謝意。

# 目 錄

## 第一章 緒 論 1

- |               |   |           |    |
|---------------|---|-----------|----|
| 1.1 敘述統計與推理統計 | 1 | 1.2 全體與樣本 | 5  |
| 1.3 統計學者之問題   | 8 | 1.4 總和符號  | 10 |

## 第二章 統計量數 18

- |            |    |            |    |
|------------|----|------------|----|
| 2.1 母數與統計數 | 18 | 2.2 中心位置量數 | 19 |
| 2.3 契畢懈夫定理 | 34 | 2.3 離散量數   | 27 |
| 2.4 $z$ 分數 | 35 |            |    |

## 第三章 資料之統計敘述 41

- |                |    |        |    |
|----------------|----|--------|----|
| 3.1 次數分配       | 41 | 3.2 圖示 | 46 |
| 3.3 對稱與偏斜      | 49 |        |    |
| 3.4 百分、十分與四分位數 | 53 |        |    |

## 第四章 機 率 61

- |           |    |            |    |
|-----------|----|------------|----|
| 4.1 樣本空間  | 61 | 4.2 事 件    | 63 |
| 4.3 事件之運算 | 66 | 4.4 樣本點之點數 | 73 |
| 4.5 機 率   | 82 | 4.6 相加法則   | 85 |
| 4.7 條件機率  | 91 | 4.8 相乘法則   | 94 |

## 4.9 貝氏法則 100

## 第五章 隨機變數之機率分配 107

- |                    |                  |
|--------------------|------------------|
| 5.1 隨機變數之概念 107    | 5.2 離散機率分配 110   |
| 5.3 連續機率分配 112     | 5.4 聯合機率分配 117   |
| 5.5 隨機變數的平均數 125   | 5.6 隨機變數的變異數 132 |
| 5.7 平均數與變異數的性質 136 |                  |

## 第六章 常用的離散機率分配 145

- |               |                  |
|---------------|------------------|
| 6.1 均勻分配 145  | 6.2 二項與多項分配 147  |
| 6.3 超幾何分配 157 | 6.4 負二項與幾何分配 165 |
| 6.5 卜瓦松分配 167 |                  |

## 第七章 常態分配 172

- |                    |                  |
|--------------------|------------------|
| 7.1 常態曲線 172       | 7.2 常態曲線下之面積 174 |
| 7.3 常態分配之應用 179    |                  |
| 7.4 二項分配之常態近似值 187 |                  |

## 第八章 抽樣原理 195

- |                |                  |
|----------------|------------------|
| 8.1 抽樣分配 195   | 8.2 平均數的抽樣分配 197 |
| 8.3 $t$ 分配 204 | 8.4 二平均數差的分配 208 |
| 8.5 模擬試驗 218   | 8.6 抽樣方法 220     |

## 第九章 母數之估計 226

- |           |     |              |     |
|-----------|-----|--------------|-----|
| 9.1 統計推論  | 226 | 9.2 傳統的估計方法  | 227 |
| 9.3 估計平均數 | 230 | 9.4 估計兩平均數之差 | 235 |
| 9.5 估計比率  | 249 | 9.6 估計兩比率之差  | 253 |
| 9.7 估計變異數 | 257 | 9.8 估計兩變異數之比 | 260 |
| 9.9 貝氏估計法 | 265 | 9.10 抉擇理論    | 269 |

## 第十章 假設之檢定 275

- |               |     |               |     |
|---------------|-----|---------------|-----|
| 10.1 統計假設     | 275 | 10.2 檢定統計假設   | 276 |
| 10.3 一尾與兩尾之檢定 | 285 | 10.4 對平均數之檢定  | 289 |
| 10.5 對變異數之檢定  | 301 | 10.6 對比率之檢定   | 305 |
| 10.7 檢定兩比率之差  | 308 | 10.8 配合度之檢定   | 313 |
| 10.9 獨立性之檢定   | 316 | 10.10 多個比率之檢定 | 319 |

## 第十一章 迴歸與相關 326

- |               |     |           |     |
|---------------|-----|-----------|-----|
| 11.1 直線迴歸     | 326 | 11.2 迴歸分析 | 333 |
| 11.3 對迴歸係數之推論 | 337 | 11.4 預測   | 340 |
| 11.5 迴歸直線性之檢定 | 344 | 11.6 指數迴歸 | 346 |
| 11.7 複迴歸      | 347 | 11.8 直線相關 | 354 |
| 11.9 複相關與偏相關  | 357 |           |     |

## 第十二章 變異數分析 364

|                 |     |            |     |
|-----------------|-----|------------|-----|
| 12.1 變異數分析方法    | 364 | 12.2 一因子分類 | 365 |
| 12.3 數個變異數相等之檢定 | 374 | 12.5 二因子分類 | 381 |
| 12.4 若干平均數距之檢定  | 376 | 12.7 試驗設計  | 400 |
| 12.6 有交感的二因子分類  | 388 | 12.8 拉丁方塊  | 402 |

### 第十三章 無母數統計 411

|                |     |               |     |
|----------------|-----|---------------|-----|
| 13.1 無母數檢定     | 411 | 13.2 符號檢定     | 412 |
| 13.3 維克生符號等級檢定 | 416 | 13.4 維克生等級和檢定 | 423 |
| 13.5 柯瓦二氏檢定    | 426 | 13.6 連檢定      | 431 |
| 13.7 等級相關係數    | 434 |               |     |

### 第十四章 時間數列與指數 441

|                |     |               |     |
|----------------|-----|---------------|-----|
| 14.1 前言        | 441 |               |     |
| 14.2 直線長期趨勢之估計 | 444 |               |     |
| 14.3 季節變化之估計   | 452 | 14.4 循環變動之決定  | 457 |
| 14.5 指數之意義與功用  | 459 | 14.6 物價指數編製法  | 460 |
| 14.7 物價指數與物值指數 | 466 | 14.8 物價與指數之緊縮 | 467 |

### 附 錄

|        |     |
|--------|-----|
| 統計計算用表 | 471 |
|--------|-----|

|      |     |
|------|-----|
| 習題答案 | 502 |
|------|-----|

# 第一章 緒 論

## 1.1 敘述統計與推理統計

統計學之研究，從基本方面說，係為陳示與解釋以科學方法研究之機會結果（Chance Outcomes）。例如，為欲設置交叉路口之交通號誌，每月記載其意外事故之損失；又如，為研究民意對某一案件是否贊同，將調查答案按“是”與“否”分類；再如，觸媒劑之用量在化學反應後，殘留物多寡之記載等。因此，統計人員研究者為點數或測定之數量資料或按某種標準的分類資料。

記載此類資料，不論其為數字的或類別的，皆稱為觀測資料（Observation）。如某交叉路口去年一月至四月發生意外事故之件數為3，1，0與2構成一組觀測值；同樣，化學試驗結果，測得殘留物有2.5，3.1與1.8克亦成一組觀測值。

**例1** 為海登山谷高地學生檢查血液，以決定各人之血型，血型按AB，A，B或O分為四種，再就各種有無Rh抗原體各分為兩種，共八種。此種結果為敘明或分類表此資料非數量或點數資料。△

**定義** 統計方法含有蒐集與敘明一組統計資料而能得有意義之訊息者，稱為敘述統計（Decriptive Statistics）。

上述定義提供之訊息僅為蒐集資料，並未對較多的資料作推論或

結論。如報紙、期刊上之統計表、統計圖與簡易及有關之計算，皆為敘述統計。

例2 表 1.1 為魯奴克時報與世界新聞星期日版刊載的資料。

表 1.1 國家聯盟棒球季節資料

| 隊 名  | <i>AB</i> | <i>R</i> | <i>H</i> | <i>HR</i> | <i>RBI</i> | <i>Pct</i> |
|------|-----------|----------|----------|-----------|------------|------------|
| 聖路易  | 2720      | 362      | 775      | 54        | 343        | .285       |
| 芝加哥  | 2542      | 337      | 693      | 74        | 313        | .273       |
| 匹次堡  | 2646      | 346      | 715      | 73        | 313        | .270       |
| 費 城  | 2761      | 356      | 742      | 62        | 335        | .269       |
| 西西那提 | 2841      | 361      | 756      | 67        | 339        | .266       |
| 蒙屈爾  | 2518      | 325      | 660      | 62        | 298        | .262       |
| 洛杉磯  | 2821      | 364      | 725      | 96        | 356        | .257       |
| 阿突蘭他 | 2758      | 364      | 707      | 72        | 334        | .256       |
| 舊金山  | 2711      | 389      | 680      | 72        | 357        | .251       |
| 休士頓  | 2773      | 311      | 691      | 32        | 294        | .249       |
| 紐 約  | 2640      | 294      | 652      | 41        | 281        | .247       |
| 聖地牙哥 | 2825      | 315      | 679      | 53        | 293        | .240       |

此表係蒐集棒球季節各場比賽的資料整理而成，由此表一看，可知各隊之攻守情形，打擊 (*AB*) 次數最多者為西西那提隊，打中 (*H*) 最多者為聖路易隊，舊金山隊為跑出 (*R*) 最多者，亦為上壘 (*RBI*) 最多之隊，跑回本壘 (*HR*) 最多之隊為洛杉磯隊。 △

例3 圖 1.1 為條圖 (Bar Chart)，示各學科大學男女畢業生 1978 年起薪比較圖，由圖知，各學科女生之起薪稍低於男生，但工程方面之學生除外，因女性工程師在全美國工程師中之比率少於 2%。 △

例4 圖 1.2 為圓餅 (Pie Chart) 形圖，示投資者每月分配其資金於股票市場，圖中各塊面積之大小，需計算每月投資於每類股票部份在總投資中之比率，然後乘以  $360^\circ$ 。 △

圖 1.1

大學畢業生起薪  
圖，1977-78。

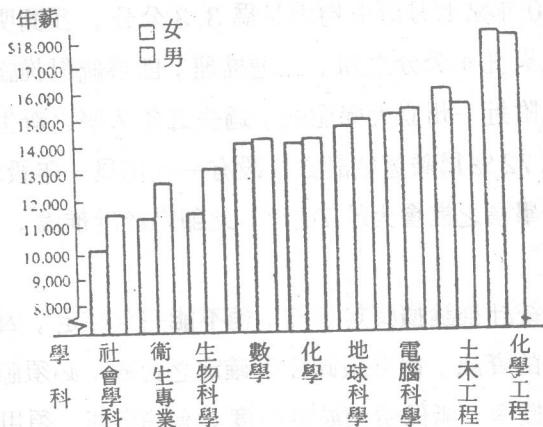
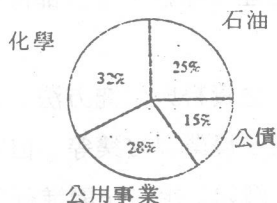


圖 1.2

圓餅形圖。



由例 2 至例 4，均將統計資料列表或以圖形說明實際情形，不論列表或以圖形說明，均有多種用途，但研究統計之最終目的，為作抉擇並根據部份少量資訊對全部大量資料作一結論，此即為統計推論部份。

**定義** 分析部份少量資料致能對全部資料作預測或作結論之統計方法，稱為統計推論 (Statistical Inference)。

敘述統計與統計推論之區別，再以下列二例說明之。

**例 5** 在勃拉錫湖之遊覽區域，有 30 年的七月份雨量記載，任何數值說明資料本身者，皆為敘述統計之範圍；如 30 年來七月份之平均

雨量或雨量最少之數值，此 30 年以外之數值，不作任何討論。若以過去 30 年來七月份平均雨量為 3.3 公分，預測明年七月份雨量可望在 3.2 與 3.4 公分之間，此種推理，即為統計推論之範圍。△

**例 6** 附近一州立大學記載，過去五年入學之新生，最後畢業者有 72%，此 72% 為敘述統計數。設有一人現為一年級新生，如認為該生在該校畢業之機會大於 70%，此即為統計推論，因其是否畢業不能確定。△

按統計推論擴展其意義，恒不能完全確定，因其所根據者，為部份資料的資訊，欲克服此種不確定之含義，必須瞭解機率理論。本書中所論機率，祇保留至最低限度，有些定理，須用微積分證明者，亦均略去，僅說明定理以備應用。大部份統計方法，以高中代數之概念均可瞭解。

本書中所提出之資料與研究方法，可適用於任一學科：如各種科學、心理學、商業、農業與醫藥等。但不論學科為何，蒐集與分析資料之方法皆相同。例如，用三元素進行化學試驗，並測定所製成品之總量；以同樣方法用三種肥料作試驗，測定稻穀之收穫量；亦可以相同方法，分析三部機器所產之不良品數。很多統計方法，原為發展農業而產生，現今將其應用於其他各學科，已獲普遍之證明與優良的效果。

現今每種工業均延請統計人員指導品質管制工作並協助其建立銷售制度，商業方面，則以研習統計者為決策之負責人，與分析時間數列及指數編製人員。就事實方面說，如應用適當，則統計為有力之工具；如應用不當，常能產生錯誤之結果。因此，應用統計者必須慎用正確統計方法，就可有之資訊發揮最大之功效。

進行一項統計研究工作，有賴蒐集資訊之處頗多。有此原因，自研究工作開始，則須延聘統計學者負責設計、執行與分析並解釋最後

結果。

## 1.2 全體與樣本

數字資料為統計之依據，非數字資料亦須將其數字化。如例 1 所述血型之分類共有八種，可將每種血型以一數字代表，八種血型以 1 至 8 分別代表一種血型；如此，每一數字代表每一學生之血型，在任一研究問題中，觀測值可能不多，亦可能很多以致無限。在血型分類中，觀測值之個數與海登山谷高地學生人數相等。但如擲二骰無限多次，並記載每次擲得兩骰點數和，每兩骰點數和即為一觀測值，故觀測值有無限多個。

全部觀測值，不論其有限或無限，而成研究對象者稱為全體 ( Population )。從前用 Population 一字表統計研究中，人口之觀測值。現今之統計學者，用此字表與研究有關之全部觀測值，可為人之觀測值，可為動物之觀測值，亦可為其他任事物之觀測值。

**定義** 全部觀測值而為研究對象者稱為全體。

全體中觀測值之總數稱為全體之大小 ( Size )。設全校有 600 名學生，按學生之血型分類，則稱全體之大小為 600。又如一堆卡片上之數字，某城市居民之高度，某湖中魚之長度等皆係全體之大小為有限之例。每例之觀測總數皆為有限。但如擲骰試驗，可產生之全體為無限大。同樣，從過去至未來測定每天大氣壓力之觀測值，或測定湖中每一點之深度等，皆為觀測值之總數為無限之例。有些有限全體很大，從理論方面說，可假定其為無限。此項假定並不錯誤。如分佈於全國之某種蓄電池壽命的全體，雖屬有限，確實很多，可視為無限。

當不易或不能獲得全體之數值予以研究時，僅根據部份數值或資料，對全體作一結論，則為統計推論之範圍。例如，欲決定某牌燈泡之平均壽命，當不能檢驗全部燈泡，因檢驗後，燈泡均全部毀損，不能再賣。研究全部觀測值，費時費錢常為人力與財力所不許，故必須自全體中取部份觀測值予以研究而對全體作一結論。因此，應研究抽樣理論。

**定義** 樣本為全體之部份集合。

若欲由樣本對全體之推論有用，則抽出之樣本應能代表全體。常有抽樣者自全體中抽最易取出之元素為樣本，此種抽樣方法，易導致錯誤之推論。由任何樣本而得之推論，高估或低估全體之特徵數值者，均稱樣本有偏誤。欲消除抽樣過程有任何偏誤，則所取之樣本應為簡單隨機樣本或簡稱為隨機樣本 ( Simple Random Sample )。

**定義** 自全體中取  $n$  個觀測值為樣本，取出每  $n$  個觀測值之機率皆相等，如此取出之樣本，稱為簡單隨機樣本。

當有限全體不大時，抽取簡單隨機樣本，並不困難；但若全體很大或無限，抽取樣本頗為不易。如有限全體很小，則將構成全體每一觀測單位寫在卡片上，並置於大容器內。倘取五個觀測單位為樣本，則從容器內取出五張卡片，即得含五個觀測值單位之樣本。如將卡片澈底混合，全體中每五張卡片之部份集合，每一部份集合被取出之機會皆相等。

如有限全體很大，用上述方法取一簡單隨機樣本，工作頗為繁重；且代表各單位之籤條或卡片，於抽樣前能否充分混合，亦成問題。抽

取簡單隨機樣本較好且更有效之方法，為用隨機數字（Random Numbers）表，本書附錄中表 A.12 即為此表。此表之構成，係將 0, 1, 2, ..., 9 十個數字按機率原理出現一長串數字，出現之先後次序隨機會而定。如一輪盤分十個相等之扇形上有一數，共有十個數字，轉動輪盤，任其自然停止，指針處之數即為隨機數。如此重複轉動輪盤，即得隨機數字表，現今均以電腦程式產製隨機數字表。

隨機數字表中數字可以任何方法組合，表 A.12 中之數字已將五個數字併為一組，橫行亦予編號以便應用。形式雖然如此，而每一位置之數字皆隨機產生且互相獨立。用此表時，先隨機選取一起點，然後按須用之位數就預定的方向依次看下去。茲舉例說明用隨機數字表取一簡單隨機樣本。

**例 7** 用表 A.12，自 80 隻老鼠中取 7 隻為樣本，以作癌症研究試驗中腫瘤生長率之研究。

**解** 先將老鼠標以 01, 02, ..., 80 個號碼，次任選第 28 橫行與 16 及 17 兩縱行數字向下讀，若兩位數字大於標號 80 者則略去，如此選擇七隻老鼠標號為

19   48   73   79   26   60   40    $\Delta$

茲設全體很大，不易或不可能標號，上述抽樣方法當不可能進行，如自加馬西亞進口藍山牌咖啡，欲對愛用此種牌號咖啡的美國人之比率作結論，用上述方法抽樣即不可能得結論。然可用其他可行之方法抽樣，如自美國各處——城市、鄉村、富庶與貧窮——分別取大量居民調查，如此抽樣調查，當可得一結論。但此種抽樣方法，不能符合簡單隨機抽樣之定義。因此，估計值可能有偏誤，估計全體比率之正確方法，如全體已分為若干區域，抽樣方法將說明於 8.6 節。本書所論之樣本，不論抽樣方法為何，在處理時，皆以簡單隨機樣本處理之。