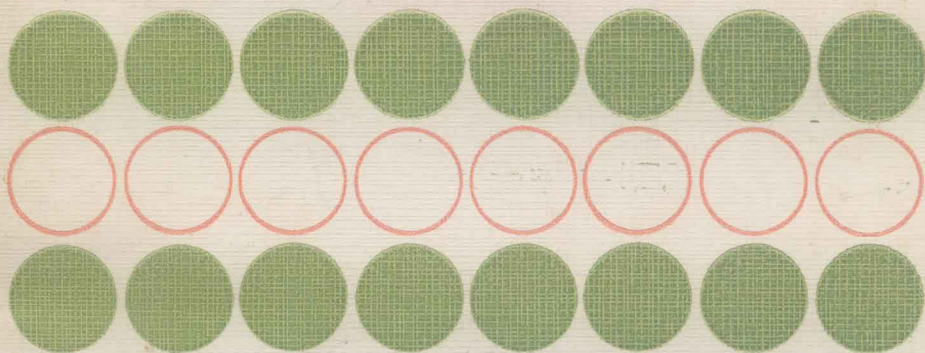


義精學計統計

編著者

宋汝濬



東華書局印行

統計學精義

宋汝澹 著

東華書局印行



版權所有・翻印必究

中華民國 六十 年 九 月 初 版
中華民國 七十 年 九 月 四 版

大專
用書 **統 計 學 精 義**

定價 新台幣壹百陸拾元整

(外埠酌加運費滙費)

編 著 者 宋 汝 潛
發 行 人 卓 鑫 森
出 版 者 臺灣東華書局股份有限公司
 (臺北市博愛路一〇五號)
印 刷 者 廣 益 印 書 局
 (臺中市中正路九十九號)

內政部登記證 內版臺業字第一〇三一號
(60024)

謹以此書獻給我的母親

統計學精義

自序

本書內容是從統計的基本概念起到常態分配為止，以統計方法為經，統計理論為緯，著重比較研究，討論儘量求深入，說理盡量求透徹，在系統和解釋方面有些地方稍有超出傳統，並兼顧統計方法之應用於實際問題。

魯峰 宋汝濬

中華民國六十年三月於臺灣省立臺中商業專科學校

統計學精義

目 錄

自序

第一章 緒論

第一節	統計和統計學的意義	1
第二節	統計的應用	2
第三節	統計方法的特質和應用上的限制	4
第四節	統計數字和近似數值	7
第五節	比率	15

第二章 統計資料的蒐集

第一節	統計資料的意義和種類	21
第二節	初級靜態資料的調查	22
第三節	抽樣的方法	25
第四節	調查的事前準備	31
第五節	初級動態資料的登記	34
第六節	次級資料的蒐集和利用	36
第七節	初級資料的審核	36

第三章 統計資料的整理

第一節	整理的意義和目的	39
第二節	分類的原則和標準	39
第三節	歸類的方法	42
第四節	統計數列	46
第五節	次數分配	51
第六節	相關表	62

第四章 統計表

第一節	統計表的意義和功用	67
-----	-----------	----

第二節	統計表的種類	67
第三節	統計表的構成要素	71
第四節	製表通則	72

第五章 統計圖

第一節	統計圖的意義和功用	75
第二節	統計圖的要件	76
第三節	統計圖的種類	76
第四節	統計圖的構成要素	77
第五節	畫圖的步驟和通則	78
第六節	點圖	84
第七節	曲線圖	86
第八節	條圖	123
第九節	面積圖	131
第十節	體積圖	134
第十一節	像形圖	136
第十二節	統計地圖	138

第六章 平均數

第一節	平均數概說	141
第二節	算術平均數	143
第三節	調和平均數	155
第四節	幾何平均數	157
第五節	中位數	161
第六節	四分位數及其他分割數	170
第七節	衆數	174
第八節	各種平均數的比較	181

第七章 離勢

第一節	離勢概說	185
第二節	全距	187
第三節	四分位差	187

第四節	平均差	188
第五節	標準差	200
第六節	均互差	209
第七節	相對離勢	218
第八節	各種離勢的比較	221
第八章 偏斜度峯度和動差		
第一節	偏斜度	223
第二節	峯度	230
第三節	動差	234
第九章 簡相關		
第一節	相關概說	237
第二節	散布圖和相關表	240
第三節	相關係數	242
第四節	直線迴歸方程式	255
第五節	估計標準誤	268
第六節	曲線相關	276
第十章 指數		
第一節	指數的意義性質和應用	287
第二節	指數的種類	292
第三節	價比	296
第四節	簡單物價指數計算方法	300
第五節	加權物價指數計算方法	308
第六節	指數公式的測驗	316
第七節	指數公式的選擇	323
第八節	編算物價指數的實際問題	324
第九節	我國現有重要的物價指數	341
第十節	其他各種指數的編算問題	348
第十一章 時間數列的分析		

第一節	時間數列的變動因素	365
第二節	長期趨勢	367
第三節	季節變動	407
第四節	循環變動	427
第五節	時間數列的相關	442

第十二章 常態分配

第一節	常態分配的意義	447
第二節	常態分配曲線方程式	456
第三節	配合常態分配	461

附錄

一、	常態曲線下面積及縱坐標表	469
二、	對數表	475
三、	隨機號碼表	499
四、	索引	507
五、	習題	517

統計學精義

第一章 緒論

第一節 統計和統計學的意義

統計一詞，英文爲 Statistics，是由拉丁文 Status 演變而來。Status 在拉丁文是國家的意思。德國赫姆斯特（Helmstadt）大學教授康倫（Herman Conring, 1606—1681）於1660年首先採用 Statistics 一詞，以數字表示土地及其生產和國家及其資源，敘述國勢，所以當時稱 Statistics 爲國家之學，並不含有近代所謂之統計方法。後來統計應用的對象逐漸由國家而推廣到社會和自然各種現象，而 Statistics 一詞仍一直沿用到現在。

在我國文字中，將 Statistics 依其特質譯爲「統計」含有「有系統大量的計數」的意思，但並不是一般所謂「總計」或「合計」之義。

關於統計 Statistics 一詞的涵義，有不同的說法：一說是尤爾（G. U. Yule）和坎德爾（M. G. Kendall）曾對統計（Statistics，指 Quantitative data）、統計方法（Statistical Method）和統計理論（Statistical theory）下過定義，並認爲 Statistics 一詞可以作爲一門科學和資料雙重的解釋（註一）。另一說是克勞斯頓（F. E. Croxton）和考丹（D. J. Cowden）認爲 Statistics 一詞含有兩種意義，一爲統計資料，一爲統計原理和方法，並且說牠不是一門科學，是科學的方法（註二）。雖然統計學是方法還是科學尚有爭論，統計一詞宜於指統計資料或統計學，而統計學又包括統計方法和統計理論。

（註一）見 G. U. Yule & M. G. Kendall: An Introduction to the Theory of Statistics, Introduction, P. xvi, 14th ed. 1950.

（註二）見 F. E. Croxton & D. J. Cowden: Applied General Statistics, Ch. 1, P. 1, 2nd ed. 1962.

一、統計資料：例如說根據「統計」，民國五十八年臺灣地區平均每人所得爲新臺幣10,305元，相當於四十七年3,434元的三倍。這統計二字就是指統計資料而言。統計的對象爲群體現象 (Mass phenomena)，包括社會現象和自然現象，又各有靜態和動態的分別。統計須以數字記錄，沒有數字就不能成爲統計。所以，群體現象的數字記錄，叫做統計資料。處理統計資料的工作，也叫做統計。

二、統計學：例如說某君是研究「統計」的，現在某學校教「統計」，這統計二字是指統計學而言。統計學是以數字記錄群體現象從而分析其數字間關係的方法之研究。從橫的方面看，統計學包括統計方法和統計理論，前者是指統計資料的蒐集、整理、表示、分析、解釋以及推測等方法，後者是闡明統計方法的原理、證明公式的來源。從縱的方面看，統計學包括敘述統計學 (Descriptive statistics) 和推測統計學 (Inductive statistics)。本書以敘述統計學爲主，略及推測方面，以統計方法爲經，以統計理論爲緯，著重比較研究。

除純理統計學外，尚有應用統計學，討論統計方法在各種實際問題上怎樣應用，如經濟統計、人口統計、教育統計等。

第二節 統計的應用

統計方法的功用爲研究問題和解答問題，應用範圍很廣，可概括的分爲三方面：

一、應用在行政方面：統計方法應用最早的就是在行政方面。古代的國家皆有關於人口和財富等事實的調查和記錄，以便明瞭國家的現勢，作爲治理國家的參考。例如大約西元前 3,050 年，埃及 (Egypt) 因修建金字塔 (Pyramid)，對全國的人口和財產曾經舉辦全面的調查。在我國古代，夏禹治水，先將山川河流地勢調查明白，治水纔能成功。春秋時代齊桓公的宰相管仲曾說「不明計數而欲舉大事，猶如無舟楫而欲經水險」。古時政治家對統計數字這樣的重視，可見歷代雖沒有統計之名，也是有實際的應用。後來政府的職掌日益增加，行政設施在在都需要依據事實，作完善的計劃。於是世界各國都先後設置獨立的統計機關，專門

辦理全國的統計工作。例如十八世紀中葉瑞典 (Sweden) 最先有人口登記局的設立，其後法國 (France) 於 1796 年、普魯士 (Prussia) 於 1805 年、奧國 (Austria) 於 1829 年、美國於 1900 年也相繼設置統計局。到現在，幾乎沒有一個走向現代化的國家不設有專門辦理統計的機關了。我國有專門統計機關的設置，是開始於光緒三十三年九月各部院設立統計處，各省設調查局。建立現代化的統計制度，是民國二十年四月一日國民政府設立主計處，下分歲計、會計和統計三局，是為超然主計。統計局是全國最高的統計機關，各級政府各機關普遍設置統計機構，形成一個周密的統計網。凡關於土地、人民、資源、政治、經濟、社會和文化等情形，皆由統計機關予以調查或登記，編成完整的統計資料，作為行政上設計、執行和考核的依據。因為設計、執行和考核為行政三聯制的三個步驟，執行須根據設計，設計須根據統計，考核須以執行的結果和設計相比較，執行的結果又須靠統計來表明，所以統計是增進行政效率必要的工具。民國三十七年行憲以後，主計處改為隸屬於行政院的主計部，不久又改名為主計處，以表示其超然主計的地位。

二、應用在學術研究方面：自從比利時 (Belgium) 凱特萊 (A. Quetelet 1796—1874) 將統計由國勢學說導入人類社會法則的研究，產生了群體現象的理論，統計方法的應用範圍逐漸的擴大，凡是不能實驗的社會現象以及不能單靠實驗方法研究的自然現象，都可用統計方法來觀察分析。如英國皮爾生 (Karl Pearson 1857—1936) 創立各種測定相關的方法，研究遺傳和進化的問題。奇馮士 (W. S. Jevons 1835—1882) 創立物價指數的算式理論，有助於經濟循環的研究。所以統計方法遂成為研究各種科學的最科學的歸納方法。

三、應用在企業經營方面：統計方法應用在企業的經營，是開始於二十世紀。企業家日常所處理的問題，可分為三類：第一類是製造品是否合乎預定規格的問題，需要以 Jacques Bernoulli (1654—1705) 創立的機率理論 (Theory of probability) 為基礎，應用統計的品質管制 (Quality control) 和抽樣檢查 (Sampling inspection)，纔能獲得經濟有效的處理。第二類是促進企業內部管理效率的問題，需要以統計方

法處理會計方法所不能作到的各種問題。如財務狀況的記錄和成本的分攤，原是屬於會計的範圍，而成本的各種因素，需要運用統計方法加以周密的分析，作為減低成本之參考，成本以外的各項支出，也需要作統計的分析，以便明瞭盈虧發生的原因而利於改進。第三類是企業家對於其外在市場情況的瞭解以及怎樣適應的問題。企業家能否獲得利益，常取決於其商品是否為顧客所歡迎，所以需要舉辦消費者意見調查 (Survey of consumers' point of view)，作為改進之參考。企業採購原料和其他生產必需的各項物品，銷售製品，必須明瞭市場上各種貨物的供需狀況、價格的變動情形和趨勢，纔能夠買的低賣的高，獲得最高的利潤。但物價現象至為複雜，需用統計方法分析其過去的變動情形，預測其未來的變動趨勢，以便於採取和變動相適應的對策。所以，較大規模的企業常自設機構，專門辦理與其營業有關的統計工作。

第三節 統計方法的特質和應用上的限制

一、統計方法以群體現象 (Mass phenomena) 為對象，不以典型現象 (Typical phenomena) 為對象：人、事或物能自成一個單位的，稱為一個個體，許多個體集合而成為一個群體。例如一個人是一個個體，許多人就是人的群體。由許多人組成一個社團，這個社團又是一個個體，許多社團又是社團的群體。各個體間在某種性質或成分上若完全相同，這種相同的性質或成分稱為典型現象。例如人的生理構造是典型的，研究生理學宜用實驗 (人體解剖) 的方法。各個體間在時間上、空間上、屬性上或數量上常有差異性 (Variation)，具有差異性的現象稱為群體現象。群體現象有的是自然現象，如人的性別、身高、體重等；有的是社會現象，如人的行業、職業、教育程度等；有的是靜態現象，如某年月日各地區的人口數等；有的是動態現象，如某地區歷年的人口數等。無論是自然現象或社會現象，靜態現象或動態現象，凡是屬於群體現象，因為其中各個體間既然有差異，其分配情形或變動情形不能用實驗的方法來研究，需要使用統計方法。所以統計的對象是群體現象，不是典型現象。

二、統計方法著重大量觀察，單獨個體的觀察不能成為統計：統計的對象既是群體現象，而群體是大量的，所以需要大量觀察。靜態方面觀察的範圍愈廣，動態方面觀察的時期愈長，統計的結果愈正確可靠。單獨個體的觀察或偶然的觀察，不能成為統計。例如某地方某天在來白米每公斤新臺幣五元八角，，這不能成為統計，因為單獨米的價格不足以代表一般物價水準，某一天的米價又不足以代表米價的變動情形。

三、統計方法須以數字記錄，以文字描寫不能成為統計：統計方法是對群體現象作大量的觀察，觀察的方法是點計（Counting）或測量（Measuring），點計或測量的結果，必須以數字記錄。數字比其他任何形容詞較為具體和確實，所以統計所著重的是經驗數量的多寡，離開了數字就沒有統計的存在。例如說房屋有的是磚造的，有的是石造的，有的是木造的，有的是鋼筋水泥造的，有的是其他材料造的，這只是說房屋在構造上有差異性的存在，不能成為統計；若說磚造的、石造的、木造的、鋼筋水泥造的和其他材料造的房屋各有多少間或多少坪，纔構成統計資料。又如某地區歷年產米若干公斤，以數字表示，是為統計；若以豐收、歉收或平常等詞敘述，除非各詞皆代表相當的數量，不能成為統計。尤其重要的，數字記錄必須確實，記錄不確實的統計資料沒有價值。

四、統計方法著重比較，但須以同質（Homogeneity）為基礎：統計方法的作用在於比較，以便明瞭群體和群體間或一群體內各部分間的相互關係。比較須以同質為基礎，不同質的項目不能比較。同質的項目相比較就是站在同一出發點上相比較，包括以下五方面：

（一）時期方面的同質：靜態現象的比較，須以同一時間為基礎。例如比較各地區的人口數，應採取各該地區同時間的人口數字作為比較。但在比較各國的人口數時，因為各國舉辦戶口普查的時間不一定相同，無法取得同時間的資料時，應儘可能採用時間上最接近的數字，並加附註說明。

（二）空間方面的同質：動態現象的比較，須以同一地方為基礎。例如比較各月份米價的變動情形，應蒐集同一地方各月份的米價作為比較。若將一月份臺北市的米價、二月份高雄市的米價和三月份花蓮市的

米價連在一起，不能表示米價變動的真相。

(三) 屬性方面的同質：例如一年級學生國文平均成績不能和二年級學生國文平均成績相比較，因為二年級學生業經讀完一年級的課程，且二年級國文教材比一年級國文教材較為高深，那麼二年級學生的國文程度應高於一年級學生的國文程度，是不待比較而自明的了。然而，同年級各班學生的國文平均成績是可以相比較的。

(四) 單位方面的同質：例如調查物價，同一物品的計價單位必須相同，纔能比較其價格的高低。若單位不同，應換算為相同的單位。

(五) 機會方面的同質：例如各中學畢業學生考取大專學校的人數不同，不足以正確地比較各中學學生程度的高低。因為各中學畢業學生的人數不一定相同，參加考試的人數也不一定相同，應該以參加考試人數除考取人數化為百分比，以百分比的大小來比較各中學學生程度的高低，或許較為合理。但各中學畢業學生的升學志願情形又不一定相同：有些中學的畢業學生全體參加考試，有些中學的畢業學生只有一部分參加考試，有些畢業學生不願升學，程度雖高而不參加考試。至於考試的結果，有些中學的畢業學生大多考取較前志願的學校，有些中學的畢業學生大多考取較後志願的學校，所以關於百分比的應用，應有慎重的態度。例如某中學畢業學生有二人參加大專學校招生考試，結果有一人考取，若說該中學畢業學生的升學考取率為50%，是很不妥當的。

五、統計方法能事後消除影響的結果：一種現象的變動，常受多種因素的綜合影響。科學研究的目的，常在根據所得的資料，從事分析比較，尋求每一因素對於研究的現象所產生影響的大小。有些自然現象的變動因素雖然複雜，可以事先加以控制，單獨觀察某一種因素的單純影響。例如物理學中測定溫度和壓力對氣體體積的影響時，可以先控制一定的溫度，單獨測定壓力對氣體體積的影響，然後再控制一定的壓力，單獨測定溫度對氣體體積的影響。凡是能事先控制影響因素的現象的研究，僅用實驗方法即可，無需採用統計方法。所有社會現象和一部分自然現象的變動因素，不是人力能够事先控制的，只能對於各種有關聯的因素，加以周密地觀察和確實地記錄，再將記錄所得的數量資料，運用

統計方法加以分析。例如測定經濟循環，可以從時間數列中消除長期趨勢、季節變動和不規則變動。凡是不能事先控制影響因素的現象的研究，使用統計方法最為適宜。

第四節 統計數字和近似數值

一、精確數值 (Exact numbers) 和近似數值 (Approximate numbers): 統計上的數值是由點計或測量而來，很難達到絕對的精確。精確數值是絕對精確毫無誤差的數值，只有在觀察的範圍很少，小到可以控制時，由細心地點計可以獲得，例如某學校的學生人數或某種商品在倉庫裏的儲存數量，經過細心地點計，可以獲得完全精確的數值。若觀察的範圍很大，大到無法控制，例如一省或一縣的人口數，縱使細心地點計，也不一定能夠十分精確。由此可見，當我們判斷一個統計數值的精確程度時，要看統計資料的性質以及取得統計資料的方法而定。不精確或相對精確而含有誤差的數值，叫做近似數值。凡由測量所得的數量，直接或間接估計所得的數值，割捨尾數的數值，簡化尾數的約略數 (Round numbers) (註)，以及在大到無法控制的範圍內點計所得的數值，皆為近似數值。所以，統計資料的數字，大半為近似數值，對近似數值和誤差的性質，應有相當的瞭解，處理這類資料時，纔不致於錯誤。

二、尾數的割捨和簡化：在有些情形下，十分精確的數值，並不為一般人所需要或欣賞，精確的程度若超過了需要，往往是一種浪費，所以有些機關所發表的統計數字，多已簡化或割捨其尾數。割捨或簡化尾數的方法，通常是採用四捨五入法，就是被割捨的尾數若等於或大於五，其前一位數加一，若小於五，其前一位數不變。

採取四捨五入的方法割捨尾數，入的機會 $5/9$ 較捨的機會 $4/9$ 為大，若有一群數值，經分別四捨五入割捨尾數後相加，其總和可能偏高。因此可採取另一更精密的四捨五入法，就是被割捨的尾數若大於五，其前

(註) Round number 的原義為 10、100、1,000、10,000 等數，擴大其涵義，又指 20、160、140,000 等數，遠東最新實用英漢辭典譯為約略數，也有譯為圓整數的。

一位數加一，若小於五，其前一位數不變，若恰為五，其前一位數為奇數時加一，為偶數時不變。例如第 1.1 表第 (1) 行六數的總和為 368.127，第 (2) 行依照四捨五入法割捨其三位尾數相加的總和為 369，第 (3) 行依照更精密四捨五入法割捨其三位尾數相加的總和為 368，和原來六數的總和更相接近。

第 1.1 表 割捨尾數方法的比較

原 數 (1)	四捨五入法 (2)	更精密 四捨五入法 (3)
51.510	52	52
63.276	63	63
98.500	99	98
41.849	42	42
23.500	24	24
89.492	89	89
368.127	369	368

整數部分尾數的簡化方法和小數部分尾數的割捨方法相同。例如民國五十九年六月底臺灣區常住人口為 14,500,872 人，設在某種情況下簡化為精確至最近萬人即認為合用，可寫為 14,500,000 人或 1,450 萬人。

三、誤差：近似數值誤差的大小，可用絕對誤差 (Absolute error) 和相對誤差 (Relative error) 表示。絕對誤差為近似數值與其真值的差額，相對誤差為絕對誤差對其真值的比。如果我們所知道的數值是真值，就根本沒有求誤差的必要。既然真值為未知數，絕對誤差和相對誤差皆無法計算，所以採用一種所謂可能誤差 (Possible error) 來代替絕對誤差，以可能誤差對近似值的比，作為相對誤差的近似值。

可能誤差為統計數值最末一位數的半個單位。例如測量某人的身高，規定精確至最近公厘，若記錄為 168.6 公分，其真實高度可能在 $168.6 - 0.05$ 公分和 $168.6 + 0.05$ 公分之間，或寫為 168.6 ± 0.05 公分，這 0.05 公分是 168.6 公分的可能誤差。凡曾割捨尾數的數值，其可能誤差也等於其最末一位數的半個單位。如前述某人的身高，依照四捨五入法割捨其一位小數變為 169 公分，其可能誤差為 0.5 公分。凡曾簡化尾數的數值，其可能誤差為其最後一位精確數字的半個單位。例如前述民國五十九年六月底臺灣區常住人口數若已簡化為 1,450 萬人，其可能誤差為五千人。

用可能誤差代替絕對誤差，用近似數值代替真值，則可能誤差對近