

統計學

作者：胡孝繩

再版：1976

統計學

作者：胡孝繩

再版：1976

自序

本書係原理與應用並重。舉凡現代統計學的重要理論與技術，從易至難，均有介紹。不但敘述大樣本統計，而且重視小樣本方法。前者所討論的，包括二項分配，Poisson 分配，常態分配及 χ^2 分配；後者則包括 t 分配， F 分配及 Z 分配等。對各種分配，除探究原理及函數導源之外，並皆說明各種實用方法。其次，對變異數分析方法，且予明晰申論。而於統計工作上時常遭遇的問題，譬如選樣方法及編造指數等問題，亦有詳盡敘述。在相關與迴歸方面，諸如直線相關，非直線相關，等級相關、品質相關、複相關與淨相關等，莫不周密闡釋。至有關時間數列者，除詳細剖析長期趨勢、季節波動、循環波動及不規則波動之外，對於漸近成長曲線，包括指數曲線、Compertz 曲線及 Logistic 曲線等，尤加詳明討論。並於時間數列的相關與景氣預測，並列有各種測定方法。此外，並學實驗的統計設計，扼要說明其理論與應用。

全書共分三十章，一百六十九節，內容包括「普通統計」、「高等統計」、「統計分析」及「經濟統計」等課程的教材。可供大就統計學系兩學年「統計學」課程，循序講授之用。至一學年的大學「普通統計學」課程可選讀第一至十章、第十二章第一節、第十三章、第十九章及第二十五章第一至六節，其餘部份可酌量取捨。一學年的「高等統計學」課程，可選讀第十至十八章，第二十至三十章。一學期的「統計分析」課程，可選讀第十至十八章，餘可酌量取捨。此外，一學期的「經濟統計」課程，則可選讀第二十至二十四章及第二十九章，餘可酌量取捨。凡讀普通統計部份，僅須高等代數基礎；至於其他部份，則僅須普通微積分基礎；蓋各種虛玄無用的數學方法，已予盡量避免，藉求適應廣泛讀者的研究。

本書完稿之後，曾充南洋大學經政學系「統計學」及「高等統計學」兩課程的講義。現為星馬學界的普遍需要，乃予出版，尚祈高明指正。唯以當前各國統計學課本觀之，作者可以告慰讀者的則為：本書內容，已具現代最新標準。

胡孝繩 一九六〇年一月十五日
於南洋大學

再 版 序 言

本書初版迄今已十多年，大部份係供東南亞若干大學統計學課本及參攷書之用，其間曾充星加坡南洋大學及香港中文大學統計學中文課本達十餘年之久。只因初版之後，課本需求殷切，除部份在星加坡市面出售之外，甚少在香港坊間出售。屢承本港大專書院教師、同學及工商人士欲函購本書，只因冊數已盡，無法供應，作者深感萬分抱歉。

今將本書增訂，再版問世。本書為修讀統計學的基本書籍，雖然實用與理論並重，但注重學習統計應用的讀者，誠可將本書各種統計公式及分配理論的數學導源部份，予以遺去，然後循序閱讀，仍不失本書的完整性及聯貫性。

本書適合用作大專統計學中文課本，教師可就學生數學興趣，對書中數學引導，自由取捨。其次本書對統計方法，敘述頗詳，工商界人士如欲研習日常所需統計實務知識，可專讀本書應用部份，將數學部份遺去無妨，因數學引導只是探究原理而已，係備若干讀者滿足窮理的願望。

本書再版，係在香港印行，將可遍本港坊間，普遍供應。

胡 寧 經

1976年8月15日於

香港中文大學

目 錄

史 略	1
第一章 緒 論	9
一、統計學的性質	9
二、統計資料	13
三、變數與函數	25
四、數的概念	28
第二章 統計圖示法	39
一、圖示的意義與方法	39
二、直交坐標的圖示	40
三、對數值的圖示	41
四、對數與半對數的圖紙	43
五、各種圖形	45
第三章 次 數 分 配	53
一、原資料	53
二、次數表	55
三、組限的選定	58
四、組距的大小	61
五、不等組距的應用	63
六、組中點的位置	65
七、累積次數表	67
八、次數分配圖示：直方圖與次數折線	69
九、次數分配圖示：次數曲線	72
十、累積次數分配圖示：穹形	76
十一、Lorenz 曲線	78
第四章 集 中 趨 勢	83
一、平均數	83

二、算術均數	84
三、幾何均數	88
四、調和均數	90
五、平方均數	92
六、中位數	94
七、衆數	98
八、百分位數、十分位數及百分位數	101
九、各種平均數的關係與優劣	102
第五章 離勢數量	113
一、分派的離勢	113
二、全距	114
三、標準差與變異數	114
四、平均差	121
五、四分位差	123
六、對離勢的測定	124
第六章 分配的動勢	127
一、動勢的概念	127
二、動勢與分配性質關係的分析	129
三、動勢的校正	131
第七章 偏態與峯度	133
一、偏態概述	133
二、偏態的測定方法	134
三、峯度的種類	139
四、峯度的測定	140
第八章 統計推論與或然率	145
一、演繹與歸納	145
二、統計的推論	146
三、或然率的意義	147
四、加法定理與乘法定理	148

四、或然率的公理.....	148
五、加法定理與乘法定理.....	149
六、排列與組合.....	152
七、從或然率分析機遇變數與函數.....	155
第九章 選樣方法	159
一、樣本與選樣.....	159
二、任意抽樣.....	161
三、非任意抽樣.....	170
四、母體數值的估計.....	173
五、真確性與樣本大小.....	178
第十章 二項分配	183
一、二項羣體.....	183
二、二項分配公式.....	183
三、二項分配的性質.....	186
四、二項分配的平均數與標準差.....	188
五、二項分配的中數.....	192
六、理論或然率與觀察或然率.....	193
七、統計或然率.....	195
第十一章 Poisson 分配	197
一、Poisson 分配的導源.....	197
二、Poisson 分配的應用.....	198
第十二章 常態分配	201
一、何謂常態分配.....	201
二、常態分配公式的導源.....	204
三、常態分配曲線的性質.....	208
四、常態分配的配合曲線.....	213
五、中央極限定理.....	217
第十三章 大樣本統計分析	223
一、樣本平均數的可靠性 (一).....	223
二、樣本平均數的可靠性 (二).....	226

三、兩樣本平均數間差異的顯著性	231
四、樣本的二項分配百分數可靠性	233
五、二個樣本百分數差異的顯著性	235
六、樣本標準差的可靠性	237
第十四章 χ^2 的分配	253
一、 χ^2 分配函數與或然率表	253
二、 χ^2 的測驗	259
第十五章 小樣本的分配——t 的分配	271
一、t 分配的概念	271
二、t 分配函數的導源	272
第十六章 小樣本的分配——F 及 Z 的分配	279
一、F 分配函數	279
二、F 分配的應用	282
三、Z 的分配函數	284
四、Z 分配的應用——測驗二 σ 間差異顯著性，當 N 甚小 及 $N_1 \neq N_2$	287
第十七章 小樣本統計分析	289
一、小樣本平均數的可靠性	289
二、兩個小樣本平均數間差異的顯著性	292
三、小樣本標準差的可靠性	294
四、兩個大小不同小樣本標準差差異的顯著性	296
第十八章 變易數的分析與可能性標準	299
一、變易數的分析	299
二、變易數分析的應用	304
三、可能性標準的應用	306
第十九章 指數的理論與構造	311
一、指數意義與功用	311
二、相對量的性質	313

三、指數的資料問題.....	320
四、基期問題.....	322
五、簡單指數的構造.....	327
六、加權指數的基本公式.....	338
七、加權指數的偏誤、測驗與交叉.....	342
八、加權指數的改造公式.....	352
九、指數的整編與平縮.....	357
十、各種重要的指數.....	366
第二十章 時間數列總說	375
一、時間數列的性質.....	375
二、四種波動內容.....	376
三、時間數列資料的整理.....	381
第二十一章 長期趨勢（一）	385
一、計算長期趨勢的目的與方法.....	385
二、隨手法.....	386
三、部份平均法.....	387
四、選點法.....	389
五、移動平均法.....	389
六、最小平方法——直線趨勢.....	393
第二十二章 長期趨勢（二）	407
一、計算曲線趨勢的方法.....	407
二、加權移動平均.....	407
三、配合多項式.....	411
四、應用對數.....	413
五、配合漸近成長曲線.....	423
六、趨勢曲線的選擇.....	435
第二十三章 季節波動	441
一、測定季節波動的目的與方法.....	441
二、原資料平均法.....	443

三、全年比率平均法.....	445
四、平均數趨勢整理法.....	446
五、趨勢比率法.....	450
六、十二個月移動平均法.....	454
七、圖解法.....	461
八、環比法.....	464
九、各種測定方法的選用.....	468
十、漸變季節性的測定.....	469
十一、季節性的消除.....	471
第二十四章 循環波動	475
一、測定循環波動的意義與目的.....	475
二、殘餘法.....	476
三、直接法.....	487
四、循環平均法.....	489
第二十五章 單 相 關	499
一、相關與迴歸的意義.....	499
二、直線相關與分佈圖.....	499
三、相關係數的原理分析.....	501
四、 r 的解釋.....	506
五、迴歸直線與最小平方方法.....	508
六、 r 的簡單計算法與計算例.....	512
七、估計標準誤.....	513
八、相關係數與離勢程度的關係.....	518
九、分組資料與相關表.....	520
十、母體相關係數估計數.....	526
十一、相關係數的可靠性.....	527
第二十六章 等級相關與品質相關	537
一、等級相關.....	537
二、品質相關.....	538

第二十七章 非直線相關	543
一、非直線相關與分佈	543
二、相關的測驗	544
三、非直線形式的改換——對數直線	554
四、非直線形式的改換——倒數直線	561
五、一般曲線迴歸的計算	566
六、分組資料曲線迴歸的計算	574
七、曲線相關指數的可靠性	583
八、相關率	584
九、相關係數、相關指數與相關率的比較	593
第二十八章 複相關與淨相關	599
一、概 說	599
二、基本計算法	602
三、複相關測定數的計算	612
四、淨相關測定數的計算	621
五、複相關係數、淨相關係數及估計標準誤的其他計算法	624
六、Beta 係數	628
七、複相關與淨相關顯著性的測驗	630
八、複相關係數與淨相關係數的標準誤	632
九、複相關係數與淨相關係數的母體估計數	633
十、曲線複相關	634
第二十九章 時間數列的相關與景氣預測	643
一、時間數列相關的性質	643
二、時間數列相關的測定	644
三、時間落後的計算	658
四、景氣預測的方法	664
第三十章 實驗的統計設計	669
一、實驗設計的重要	669
二、如何使實驗成為敏感	669

三、實驗變數的排列	670
四、實驗設計的數學模型	672
五、應用說明	675
附 錄 (一)	679
附表 A 常態或然率曲線的縱坐標	679
附表 B 常態或然率曲線下的面積	680
附表 C 一定自由度及一定 P 數值的 χ^2 數值	681
附表 D n 自由度及 P 顯著性水準的 t 數值	683
附表 E 各 n_1 及 n_2 特定數值的 F 數值	685
附表 F 各 n_1 及 n_2 特定數值的 Z 數值	690
附表 G 各 N 與 K 特定數值的 L 數值	694
附表 H 自然順序數字各次自乘的連和	696
附表 I 單數順序數字各次自乘的連和	697
附表 J 對數表	698
附表 K 平方、開方及倒數	700
附 錄 (二)	703
a. Chi-square 分配函數的 P 數值計算式導源	703
b. 證明 $\chi^2 = \sum x_i^2 / \sigma_i^2 = \sum (f_i - f_o)^2 / f_o$	707

統計學史略

統計之事，淵源遠矣。伏羲畫八卦，（註一）爲圖線記事的濫觴。紀元前二千二百五十年，大禹治水，依山川土質，人口產物，及貢賦多寡，分全國爲九州，纂編禹貢九州篇，頗具統計雛形。而三代行井田，計口授地，已備土地與戶口的統計。王制云：「冢宰以三十年之通制國用，量入以爲出。」又云：「三年耕，必有一年之食。九年耕，必有三年之食。以三十年之通，雖有凶旱水溢，民無菜色。」足見我國古代對生產與消費的統計，業露端倪。春秋時代常以兵車乘數，論諸侯實力，顯已注意軍力調查與比較，漢代行口錢，民年十五至五十六，出錢百二十，以食天子，全國戶口與年齡的統計數字，成爲有司征賦的資料。而明初編製黃冊與魚鱗冊兩種。黃冊乃全國戶口名冊，記載各戶所有的丁糧，爲核定賦役的根據。魚鱗冊係全國土地圖籍，繪其地形，記明地味、位置及所有人，深具現代統計圖表的意味，而規模之大，殊屬空前。故我國歷代，對統計事務，頗爲重視，奈彼時缺乏學者，從事專門研究，未能成爲有系統的學問耳。

而西方各國統計之舉，首以紀元前三〇五〇年，埃及建造金字塔，爲徵集建築費，乃對全國人口與財產，舉行普查，是爲統計的嚆矢。嗣後羅馬斯第二（Romeses II）復對埃及土地，辦理調查。然皆屬實務的推行，未涉學理的探求。至統計學理研究的演進，則可溯自亞里斯多德（Aristotle）時代。迄今約歷二千三百年。其間可劃爲四個時期：（一）初期乃含濃厚哲學色彩，並以希臘「城邦政情」（Matters of state），爲其研究對象，除有助斯學名詞的鑄造外，對今日統計學的內容，關係甚微。（二）第二階段爲部份應用數字，而帶有強烈的社會學氣質，乃對衛生、保險、國內外貿易及行政設施等的大量數字測定數（Numerical measurements），從事數學與哲學的研究。（三）第三階段始自十六世紀。由於應用數學方法，解決賭徒們在機遇戲法（Game of chance）上的各項疑問，或然率（Probability）的數學理論，隨而發展，成爲統計學的一個基本理論。（四）晚近時期從十九世紀末葉以還，乃對第二及第三兩階段的研究內容，予以合併、改進及擴展，使統計學蔚爲數學性質的學科，足以處理廣泛的問題，並得從相當小量的數字測定數，導出真確與有用的推論。

當亞里斯多德時代，統計意旨，厥為對「城邦」(States) 的比較敘述。亞里斯多德曾經撰寫城邦紀要，達一百五十八種以上。(註二) 其中包括各邦的歷史、公共行政、藝術、科學與宗教等等。當時通稱此種著作為「城邦政情」論文 (Treatises on "matters of state")。因其名稱包含「城邦 state」一字，遂成為「統計學」(Statistics) 一詞的起源。雖然統計學一詞的正式鑄造，遠在亞里斯多德逝世之後，而究其原委，實萌於氏之論文名稱。此種城邦紀要的論著，在亞里斯多德死後相當長久時間中，仍有人繼續編纂，一方面乃因有閒階級以此為娛，一方面則緣各邦統治者渴望明瞭其領域的潛在力量或資源，包括臣民人數，財富數量等等；因而粗畧的國勢調查 (Census)，或經採

而城邦紀要形態的統計工作，流傳後世，斷斷續續，當十七世紀及十八世紀時期，充分興趣，乃貫注於用此方法，以研究國家政治、社會與經濟的狀況。1660 年德人 Herman Corning 任 Helmstadt 大學的自然律 (Nature law) 教授，曾以國家土地、生產及其他資源的數字資料，為其「國情紀要」課程的教材。1746 年 Gottfried Achenwall 執教於 Marbury 大學，所編統計講義，則詳述英、法、德、俄、荷蘭、丹麥、瑞典、西班牙、葡萄牙等國的物產、人口、土地、兵力等狀況，並分別加以比較；且將拉丁文的 Status 一字，蛻化為指統計學問的名詞。於是，在德國乃有 Staatenkunde 一門學問的產生；而 Staatenkunde 的名稱，遂成為 Statistics 一詞鑄成的直接模型。蓋 Statisti 原為意大利文，意指政治家 (Statesman)。而統計名稱則係將研究政情的學問與政治家二詞，變造而成。但此時德人的努力，仍不脫「城邦紀要」的窠臼，Staatenkunde 之學，既不具若何數學性質，亦無多大用處，不過表現純粹哲學形態的統計分析，發展至此，已達最後階段或尾聲；並指出統計學所具社會學的與政治學的世系而已。一般人常稱此等專以文字記述政情為統計內容的學者，為記述統計學派。然為便利各種事實用數字表現，及互相比較，殊有採用表格及幾何圖形的必要。1741 年丹麥人 I. P. Acherson 編纂歐洲各國比較統計表，對各國的面積、人口、教育、財政、軍隊、政體、度量衡、及貨幣等，悉用數字，列表比較。1782 年，德人 A. F. W. Crome 並繼之應用幾何圖形，以顯示統計結果。遂有圖表統計學派之稱。文字記述派

與圖表記述派，曾互相批評，引起激烈辯論。然歷史發展，乃有利於圖表派，而不利於文字派。

其次，現代型式統計學理，各演進階段中，具有較大貢獻的一個階段，當推 1660 年代以英人 John Graunt 爲首，所開創者。Graunt 氏對生命統計、保險統計及經濟統計，從事部份的數學研究。此一學問，曾被稱爲「政治算術」(Political arithmetic)。Graunt 氏由生命統計，發現人口出生與死亡率，時常不變，遂倡立「大量恒靜」定律 (Law of inertia of large numbers)，成爲統計學基本原理。蓋是時歐洲各國，發生週期性的瘟疫，每次減去人口，竟達十分之一之多；社會上遂普遍注意人口出生與死亡的數字，各地人口的估計、農業生產及國際貿易的數量，與乎對付由於人口死亡及殘廢，所引起經濟狀況的管理方法等等。似此，農業生產、對外貿易、及公共管理諸問題，既形複雜，亟欲若干形式的數學測定數，作爲處理根據的資料，並需客觀方法，以分析此等測定數。統計學的數學性質，漸見加重。迨十八世紀，各國受戰爭威脅及工業革命的刺激，對人口與資源的測定，極感興趣。由是有關經濟、社會、及政治等問題之統計資料的搜集與解釋，成爲政治數學家努力的對象。而且此等問題，日形重要，並益複雜。政治數學家乃設計方法，以估計各政治區域住民人數，並設計方法，以概括表現羣體資料。彼等所應用的數學分析，遂奠定現代統計方法的基礎。世人通稱其爲數理統計學派。

唯現代統計學演進的第三階段，則基於或然率數理論的發展。倘無或然率理論，統計學殊難有今日成就，將不足獲得普遍的信賴，與發揮廣泛的功用。但或然率原理的研究，乃緣許多數學家爲賭徒們解決機遇戲法的疑問，所觸發者。或謂現代統計學，係由充滿灰色與烏煙酒氣的豪賭，所附帶育成，似亦非過言。職業賭徒，志在贏錢。爲求增加賭勝的機會，彼等熱望洞悉，賭牌、擲骰、旋轉鋼錢或其他戲法，各種可能出現的結果。彼等曾忖度各種賭具的出現事件（譬如擲二骰出現八點的事件），係非完全碰巧，而乃循某種未知的定律。假使彼等能發現此中奧妙，加以運用，並保密不令對方知曉，則彼等可靠此機遇戲法的智巧，獲得優裕生活。爲探求賭徒們此等問題的解答，今日所謂或然率原理，從而誕生。

職業賭徒們，因無法導出支配機遇戲法的一般原理或定律，乃轉請當時數

學家，代為解釋。或因當時賭徒，在社會上，頗擁地位，各種有關機遇數法問題的論述，竟在知識社會的雜誌上，時常披露，且歷續數年之久。最早者為 1477 年，有關投擲三個骰子，可能出現的三種結果，曾在但丁神曲 (Dante's Divine Comedy) 中，加以考究。1494 年數學家 Pacioli 首先公表當二個賭徒的賭博行為，被中途干涉時，所剩籌碼，究應如何分配的解法。此似乎已將警士捉賭，亦列為賭博方法之一矣。而賭徒兼數學家的 Candan, 1501 年出生，並著有賭徒手冊 (Gambler's manual) 一種，似為專備自用的賭博寶鑑。有一個意大利貴族請天文學家 Galileo (1564—1642) 解釋下列問題：投擲三個骰子，出現 9 點與 10 點，各有六種不同組合法，但在經驗上，發現 10 點出現的次數多於 9 點，是何緣故？由於計出三個骰產生總和等 9 或 10 的共有途徑，Galileo 能够明白答覆此問題，而使對方信服。其解答成為或然率原理 (Theory of probability)，應用的首次發表。(註三)而另一才華洋溢的賭徒 Chevalier de Méré 提出一個賭博難題，請數學家 Pascal 解答。1654 年 Pascal 乃就此問題與另一數學家 Fermate 從事通訊討論。幾經研商，兩位數學家終對此問題，獲得一般化的解法，成為或然率原理的基礎。至 1657 年荷蘭著名物理學家兼數學家 Huygens 發表短文，論機遇數法。氏用數學術語討論牌戲與骰戲的贏賭機會。數學教授 Jacques Bernoulli 曾撰一鉅著，專論或然率原理。此書在 1713 年，公表於世。氏置或然率於數學基礎之上，並提出或然率原理，可應用於社會、倫理、及經濟事務之上的意見。然此意見，是時未受珍視，致或然率原理的普遍實用，遠在距此一世紀之後。同時期，另有兩個數學家，Montmort 與 De Moivre，亦發表機遇數法的文章。De Moivre 半生時間都為貴族解答賭博的數學難題，而得到不少金錢。氏喜愛純理數學 (Pure mathematics)，約當 1733 年導出一個極重要的公式，即後來所謂「常態曲線」(Normal curve)，並計出該曲線各不同間隔的或然率。此種常態或然率並經列成一表，首在 1799 年由 Kramp 所發表。今純理論成為現代統計學的軀幹。然 De Moivre 初不知其工作如是重要。氏且自由懶惰將其理論建立在賭博基礎之上，於是在其隨後出版物中，乃試圖將或然率原理與神學 (Theology) 聯結一起，認為神意秩序 (Divine order)，係表現於人間事物的規則性上。其實或然率原理，頗難逸出機遇數法的範疇。

試觀吾人今日言或然率原理，仍然不離以橋牌、骰子、或轉錢等等賭博為說明基礎，可見一斑。

由於數學家對賭博上特殊問題的解法，加以一般化，並對各種理論，予以系統化，結果乃形成，可以適用於許多場合的一種統計方法。其應用範圍，廣被自然科學與社會科學。在物理學與天文學，分別著重於實驗的觀察誤差 (Errors of observation)，與天體運動或位置的觀察誤差。而在社會科學研究的統計方面，及政府施政的統計方面，則大量資料被收集、組織、列表及解釋。從數理觀點上言，觀察誤差理論 (Theory of errors of observation) 乃十九世紀初由 C. F. Gauss (1777—1855) 及 P. S. Laplace (1740—1827) 所發展。而 A. M. Legendre (1752—1833) 所創最小平方法則 (Principle of least squares) 亦有功於誤差理論的建立。諸氏工作均指出今日吾人所稱誤差定律 (Law of error)，常態定律 (Normal law) 或常態分配 (Normal distribution)，確然成立。

迨十九世紀中葉，許多統計學理的新發展，幾皆直接或間接由二個偉大的天才所推動，一為比利時統計學家 Adolphe Quételet (1796—1874)，一為英國人類學家 Sir Francis Galton (1822—1911)。Quételet 主要功績在使統計方法，獲得普遍應用。蓋 Quételet 對各種科學均有造詣，諸如天文學、數學、物理學、人類學、行政、生命及社會統計學，氣象學等等，無所不通。氏將統計方法應用於所有此等研究範圍之上；並努力使人理解統計學係一種可用於任何科學的一般研究方法。氏且強調常態分配的用途，主張此種分配狀態，可適用於許多科學範疇。氏深信行政官與立法官必須受統計知識的指導，亦即必須受事實的指導，否則將趨失敗，其熱情澎湃，尤甚於當時著名博愛主義者 Florence Nightingale (1820—1910) 之所為。而 Galton 對遺傳定律，尤感興趣。耐心搜集甚多資料，從豌豆至人類，應有盡有，以無窮精力，鑽研所搜集資料所隱藏的模型與關係。氏最重要的貢獻，厥為領導導線與相關 (Regression and correlation) 概念的發展，以解決二個或二個以上測定數，同時從分別羣體計出，如何探索相互關係或共變 (Co-variation) 的問題。此外，氏並發明中數 (Median)、四分數 (Quartiles)、百分數 (Percentiles)、及四分差 (Quartile deviation) 等。氏的遺傳研究，實為導致其發現相關概