

(大 專 用 書)

統計學概要

(增訂版)

鄭堯梓編著

(大 專 用 書)

統計學概要

(增訂版)

鄭堯梓編著

U8-47
Z450.107
Z455

序

統計學是一門新興科學，在最近數十年其內容進步迅速並日趨完備，適用對象亦不斷推廣，已成為各學科研習上不可缺少之分析工具。

本人曾在民國二十五年編著第一部統計學（上下二冊，由上海商務印書館出版）問世，不幸紙型毀于戰火，以致絕版多年，來台後再行執筆時，以上列著書內容，已嫌陳舊，特從新改寫，加入新的統計方法（統計推論方法），並已於民國四十七年出版統計學（一）（敘述篇），民國五十年出版統計學（二）（推測篇），民國五十六年出版數理統計學，前二者內容較為廣泛，包羅所有統計方法，惟仍只敘述應用，不探求由來，而後者則為高深理論，專述各種公式之來源。上列諸書均以專攻統計學者為對象，對於只須略涉統計學智識而期能即刻應用者，深感不甚切合，因此特濃縮統計學（一）（二）編成本書。其內容分配為敘述統計學約三分之一，推測統計學約三分之二，使二者能兼顧並蓄。其分量期在作教本時，依每週三小時，能在一學年內全部教畢，並能學習到統計學之重要部門為目標，以定取捨。再特明示各項統計方法之計算步驟，並用表解體裁，將其步驟用數字實例互相對照實行演算，使讀者一見即知進程序與計算方法。書內所附與實例亦相當衆多，共有一百八十餘題，務使讀者明理即能應用，求用者觀例即能做解。本尚擬加入迴歸分析多元相關及抽樣調查等項，後以所占篇幅太多，使教授時間不克分配，只有割愛，有志研究者可接讀拙著統計學（一）（二）即得。

本書旨在能瞭解統計學大意，其取材範圍是否適宜尚可探討，所述內容或有錯誤之處，均希讀者能多予指教，容待補充與訂正。

民國六十年六月

鄭堯梓 序于國立政治大學
統計研究所

增訂版序

根據多年講授經驗，深覺統計學概要內容，仍須包含比率，時間數列分析與抽樣調查等項。因此在此次重印時，特將指數一章改為比率與指數，時間數列一章內添列季節變動與循環變動之求法，並將抽樣調查方法概念另列一章，成為第十六章，附在最後，使其內容趨于完備，以滿足讀者需要。

再原著內 4、3 節之四分差，在分組時之計算方法有所更改，著者則覺其較原有者為條理分明並易于計算。

民國 66 年 8 月

著者序于台北

目 錄

第一章	緒 言	1~3
1.1	統計之意義	1
1.2	母體與樣本	2
1.3	統計學之內容	2
	練習題	3
第二章	調查、整理與製圖	4~35
2.1	統計工作步驟	4
2.2	擬訂調查方案	4
2.3	資料蒐集上各要項	5
2.4	統計表之編製	11
2.5	資料之整理	16
2.6	統計之發表	19
2.7	統計數字誤差之發生原因	20
2.8	統計圖之繪製	21
	練習題	35
第三章	次數分配	36~46
3.1	變量	36
3.2	次數分配表	36
3.3	次數分配圖	41
3.4	累加次數分配	43
3.5	次數曲線之模型	44
	練習題	45
第四章	地位量數	47~74
4.1	意義	47
4.2	算術均數	47
4.3	中位數	56
4.4	衆數	60

4.5	幾何均數	67
4.6	調和均數	69
4.7	各地位量數間之關係	71
	練習題	72
第五章	差 量	75~98
5.1	意義	75
5.2	全距	76
5.3	四分差	76
5.4	標準差	81
5.5	平均差之定義及計算公式	92
5.6	均互差之定義及計算公式	93
5.7	相對的差量	94
	練習題	95
第六章	動差、偏度及峰度	99~113
6.1	動差	99
6.2	偏度	104
6.3	峰度	108
6.4	總述	110
	練習題	111
第七章	初等機率論	114~144
7.1	事件與機率變數	114
7.2	用集合符號以表示事件之發生	116
7.3	機率定義	117
7.4	公理體系	119
7.5	Bayes 定理	127
7.6	分配函數	129
7.7	希望數 $E(X)$ 及變異數 $V(X)$	132
7.8	樣本統計量與母數間之對照表	137
	練習題	137
第八章	三個重要機率分配	145~167
8.1	二項分配	145
8.2	波氏分配	152

8.3	常態分配	156
8.4	總表	165
	練習題	166
第九章	常態母體下三個統計量之機率分配	168~179
9.1	χ^2 -分配	168
9.2	F -分配	172
9.3	t -分配	175
9.4	總表	177
	練習題	178
第十章	統計的推定	180~198
10.1	定義	180
10.2	點推定	180
10.3	區間推定	184
	練習題	197
第十一章	統計假設之檢定	199~233
11.1	前言	199
11.2	統計假設檢定之意義與名稱	199
11.3	母均數 μ 之檢定	202
11.4	母均數差之檢定	207
11.5	母變異數之檢定	217
11.6	百分率之檢定	221
11.7	適合度之檢定	227
11.8	獨立性之檢定	228
	練習題	232
第十二章	變異數分析	236~249
12.1	意義	234
12.2	一因子變異數分析	234
12.3	二因子變異數分析	242
	練習題	248
第十三章	迴歸與相關	250~282
13.1	意義	250
13.2	相關表之作成	250

13.3	迴歸直線方程式之求法	253
13.4	相關係數 r	258
13.5	二迴線直線相交角與相關係數間之關係	267
13.6	估計標準誤 S_{e_y} 與 S_{e_x}	268
13.7	互變數 $C(X, Y)$	271
13.8	r 分配	272
13.9	母相關係數 ρ 之信賴區間	274
13.10	母相關係數之檢定	276
	練習題	279
第十四章	比率與指數	283-297
14.1	比率	283
14.2	指數	289
14.3	綜合指數之計算方法	291
14.4	指數算式之測驗	294
14.5	生產指數(物量指數)	295
	練習題	296
第十五章	時間數列	298-318
15.1	意義	298
15.2	長期趨勢	299
15.3	季節變動	308
15.4	循環變動	316
	練習題	317
第十六章	抽樣調查	319-334
16.1	前言	319
16.2	選定抽出方法	319
16.3	隨機抽樣方法	319
16.4	層別抽出法	329
16.5	部落抽出法	331
16.6	系統抽出法	332
	練習題	333

數 值 表	335-356
表一 平方表	336
表二 二項機率右尾累加數值表	340
表三 Poisson 機率右尾累加數值表	342
表四 標準型常態機率面積數值表	344
表五 卡方數值表	347
表六 F 數值表	348
表七 t 數值表	352
表八 $\rho = 0$ 時 r_a 數值表	353
表九 Z 變換表	354
表十 亂數表	355
索 引	357-361
習題答案	362-401

統計學概要

第一章 緒 言

1.1 統計之意義

統計一詞有二種用法：

(1) 作動詞用，表示相加之意義，例如將1加2為3之結果取名為統計得3。

(2) 作名詞用，對此又可分為

(a) 在單數時，取名為統計量 (Statistic)，用以表示研究對象之特徵數，例如總數、比率數、平均數等是。

(b) 在複數時：(i) 取名為統計數字集團 (Statistics)，吾人所述之統計，則指此而言。不過此數字集團並非是抽象的數字（例如只書學生三十名），而是表示一定「時」「地」及「某種品質」之數字集團，例如民國五十六年十二月卅一日零時正全台灣省人口數及其品質內容等數字始可稱為一種統計字學。(ii) 統計學亦用此名詞。

對此統計數字可依狀態之不同又可分為

(1) 靜態統計：一定瞬間內之現象狀態稱為靜態，由此所數計得之統計數字，稱為靜態統計。例如戶口普查必須定一標準時刻（第二次台灣地區戶口普查定在民國五十五年十二月十六日午前零時正），在此時刻內所查計得之統計數字稱為靜態統計。

(2) 動態統計：由一定期間內所登記得之資料以求得之統計數字。例如生命統計是將平時所登記之出生死亡狀態所求得之統計，故稱為動態統計。

再依目的之不同分為

(1) 直接統計：專為得統計資料特于現象上舉行觀測後所得之統計，稱為直接統計或稱原始資料。例如依戶口普查所得之人口統計，是專為得人口統計起見，特于人口現象上施行戶口普查之觀測方法，故稱為直接統計。

(2) 間接統計：於已為別項目的所存在之資料中，收集資料以行作成

之統計，稱為間接統計或稱為次級資料。例如貿易統計是於已為別項目的所舉辦之每月每日貿易報告中搜集其所需要之資料，故稱為間接統計。

1.2 母體與樣本

母體 (Parent population) 係指調查對象之總集團，樣本 (Samples) 是為由母體所抽出之部分集團。

【例1】 在檢查一袋米之品質時，由此一袋米任抽出 400 粒米，此 400 粒米稱為樣本，其一袋米是母體。

【例2】 某病院對於一定健康狀態之病人 20 名施以某種新式治療法，以觀察其結果。當此時之 20 名病人是為樣本，而同樣健康狀態之所有人民是為其母體。

此母體在事實上很難掌握，吾人所得資料多只其一部份樣本而已，因此發生由樣本以推論母體之問題。

由母體所求得該現象之特徵數稱為母數 (Parameter)，由樣本所算得之特徵數稱為統計量 (Statistic)。

1.3 統計學之內容

統計學係為研究統計數字之製造方法與分析方法的一種學問，其目的在能求出可表示母體狀態之模型。至分析方法又稱為統計方法 (Statistical methods)，內容可包含

(1) 敘述 (Description)：應用簡單之統計量 (例如均數、比率及相關係數等) 以表示複雜現象之性質與相互間之關係。

(2) 推定 (Estimation)：根據少數樣本所查得資料之統計量以推算其母體之母數及其所在範圍。

(3) 檢定假說 (Test of Hypothesis)：根據樣本所得結果以檢定所給予統計假說之是否成立，為實驗科學提供重要之研究方法。

(4) 判定 (Decision)：根據樣本資料以判定母體所屬模型或用以確定此後行動之決策。

(5) 預測 (Prediction)：根據已有資料以測定此後發展之趨勢。

上列(1)只敘述現象內容而已，可專稱其為敘述統計學 (Descriptive Statistics)，而(2)，(3)，(4)及(5)則為最近二十年所發展之推測統計學 (Inductive Statistics)，完全應用機率理論 (Theory of probability)

以進行其研究。本書之大部分亦注重在此方面。

第一章 練習題

- (1) 統計與會計均係數字之處理，兩者主要不同之點何在？
- (2) 應用統計方法於各部門科學如經濟學、生物學或教育等之研究，其統計理論是否有所不同？
- (3) 試述統計之功效。
- (4) 試述下列各項何者為真，何者則否。
 - (a) 好幾個世紀以前，政府已開始收集統計資料。
 - (b) 數量資料上之數字，應該絕對精確，尾數不能有微小差異。
 - (c) 機率的數學理論，萌芽於賭博遊戲。
 - (d) 統計 (Statistic) 一詞的單數用法，其意為數字資料。
- (5) 試就政府發表之統計資料中列舉其直接統計 (原始資料) 及間接統計 (次級資料) 之名稱。
- (6) 由戶口普查所得之戶口統計，與由人事登記所得之出生死亡統計中何者為靜態統計？何者為動態統計？
- (7) 統計學之內容可依時代而不同，試就最近趨勢加以申述。
- (8) 設由一批製成品內抽出 200 件製品舉行檢驗，問何者為母體？何者為樣本？

第二章 調查、整理與製圖

2.1 統計工作步驟

舉辦統計工作之進行步驟如次：

擬訂方案→蒐集資料→整理資料→^{編製統計}_{圖 表}→分析研究→應 用

茲將其最初四項步驟在本章內加以說明。

2.2 擬訂調查方案 (Planning of survey)

無論何種工作均有其最終目的，統計工作亦然，吾人須有確定之調查目的，俾能正確把握問題中心，而後始能據此以擬訂方案。

在擬訂統計方案（即計劃書）之前，必須先認清對象，瞭解環境，求計劃之能切合實際，並須注意本身之人力、物力、財力、時間及一切法令規章與已有資料。

方案內容之最重要者，為決定何人主持（Who），對象何物（What），標準時刻（When），地域如何（Where），如何進行（How）等五項。統計專家所稱之五W主義即屬於此。例如民國 55 年台閩地區戶口普查工作上之

何人主持：行政院戶口普查處。

對象何物：常住人口兼查現在人口。

標準時刻：民國 55 年 12 月 16 日午前零時正。

地域如何：台閩地區。

如何調查：對於常住人口採預查復核制，現在人口採當場查填制。而後再決定要多少錢要多少時間，遇受經費與時間之限制，則須變更其內容與範圍。

方案初步擬訂妥當後應舉行試查（Pilot survey）數次，研究所擬方案是否適合實際，若有不妥處須即加修改。而後再注意

- (1) 充分訓練督導與調查人員；
- (2) 籌撥充裕經費；

(3) 注重宣傳工作，使被調查者均能瞭解其重要性與內容。

茲以人口普查為例以說明擬訂方案時必須計劃之各種事項，（摘自一九五八年聯合國出版之統計叢書 M27 號）以資參考。

- (1) 普查之法律根據。
- (2) 經費和工作人員之估計。
- (3) 普查對象及工作預定進度表。
- (4) 普查機構及其行政事務，（組織及辦事細則）。
- (5) 調查工作之準備，（地域分劃之決定，準備普查地圖，編製名冊等）。
- (6) 準備普查問項。
- (7) 確定統計表內容。
- (8) 決定調查之基本程序和管制方法。
- (9) 計劃抽樣工作之應用範圍。
- (10) 資料整理計劃。
- (11) 出版計劃。
- (12) 舉辦試查。
- (13) 宣傳。
- (14) 訓練調查人員。
- (15) 調查人員查勘地區。
- (16) 調查表件分送至各工作處所。
- (17) 調查工作之實施（監督與資料之收回）。
- (18) 問項表之審核與點收。
- (19) 製表和編纂。
- (20) 公佈（分初步速報與最後結果）。
- (21) 研究與檢討。

2.3 資料蒐集上各要項

(一) 蒐集機構之組織 (Organ of survey)

現代科學的職務組織皆如金字塔，凡決定事業全體之根本方針與分發重要命令之首要者，實據金字塔之頂端，其下依次分段，人數亦依次增加。

統計蒐集機關之組織亦宜如金字塔，有一指導者據其頂上以定一切方針，此指導者之責任重大，故最少須具備下列數項條件：

- (1) 對於數字的研究須有特殊之興趣；
- (2) 對於現代統計學之理論與實務上須有充分之研究；
- (3) 非只模仿而已要有創造力；
- (4) 須有努力統計事業發達之決心；
- (5) 尊重統計事業並確信其為終生事業，全力工作。對此一點則以優待統計人員為先決條件。

至調查工作人員可分為

督導員 (Supervisor)

指導員 (Crew leader)

審查員 (Editor)

調查員 (Enumerator or Interviewer)

最下層工作人員則為調查員。對此等人員，事前宜正確規定其業務之分掌，明白其責任之歸屬，使全體工作人員成為有機體的一團，而後工作始能圓滑推動。各國戶口普查於此意義上，特設立普查區 (Enumeration districts 簡書為 E. D.)，以定普查員之工作範圍。當劃分此等普查區時，對於境界線宜與行政上地理上傳統上之境界線相一致，對於其面積則宜考察普查項目之性質與普查員之人數以定其廣狹。

使用調查員時，對於調查員之選擇宜十分注意，須選擇其對於工作，有興趣、熱心、勤快、身體強健、溫厚、篤實、信望昭著、能嚴守秘密、熟悉地方情形與語言及精通調查項目之人士，如警員、稅務人員、新聞記者或通訊員等易為被調查者所誤解，非不得已時萬不能使用此等人員為調查員。

(二) 蒐集方法

最基本的直接蒐集方法有二，(1)為調查法，(2)為登記法，兩者性質不同，效用亦異。調查法用於對事實真相之採訪，是為採取靜態統計資料之方法，並依施行方法之不同又可分為全部調查 (Complete enumeration) (普查 Census) 與部分調查 (Partial enumeration) (抽查 Sampling survey) 二種。至調查方式又可分為①親自調查，②派員調查，③通信調查 (郵送或電話)。登記法為對事實真象之不斷記載，是為動態統計常川記錄方法。

至於間接調查方法 (Indirect survey) 可分為

- (1) 推算：根據原始資料加以分析與估計。

(2) 轉用 (Diversion)：利用已依別項目的所存在之資料，抄錄其所需要部分。

(3) 詢問調查法 (Enquête)：向所欲研究之問題有特殊認識之少數專家，加以詢問。

(4) 表式調查 (Tabular investigation)：用統計表式樣表令調查對象填報，在行政統計上多用此項調查方法。

(三) 調查項目 (Enumeration items)

當決定單位與品質之調查項目時須注意下列諸點：

(1) 須選擇適合於研究目的之調查事項。

當蒐集統計資料時，對於調查之目的須預行確定，並依此以定調查手段。

(2) 前後宜相同。

對於同一種類之調查，雖在不同時間與空間上舉行時，對於調查項目仍須前後相同，以便於比較。

當比較時為明瞭真相計，有充分考察該現象之前發條件的必要。例如將學校內之死亡率與全國死亡率相比較時，學校內之死亡率當然比全國為低，對此並非是一入學校即能減少死亡，實因全國人口內有老弱及幼孩等死亡率高大者混在其中，以致所算得之死亡率較學校為高。由此知在調查各現象時，對於前發現象（例如死亡率之前發現象為健康條件、年齡、職業等）有充分研究之必要。若前發現象有重要變化時應將其狀態詳細加以研究。

(3) 項目不宜過於太多。

同時調查多數項目，其結果多不優良，故對於項目須十分精選，抱少問多得主義，求其適用已足。

(4) 調查上之三不可。

調查中有三不可事，第一關於智識道德等不能調查事項。蓋智識道德屬於精神範圍，是無形之物，不易從調查求得正確之測度，故當蒐集資料之前，須先考察其欲調查之項目，能否依確實數量以行表示一事。第二關於個人之內行私事等不許調查事項。蓋調查個人秘密實為統計上所大忌，若強欲調查，則反有害其他正當之調查項目，故必須嚴禁。但已得被調查者之同意時，則不在此限內。第三關於事實上雖可調查，而實際上勞多功少之不利調查項目。例如調查全國鞋數，事雖可能，但問有何用處。此三

種不能、不許、不利項目，當決定調查項目時宜留意切避。

(5) 對於調查項目宜附正確明瞭之說明。

大量觀察是於大量現象上精查各單位，故對於所定調查項目須有正確明瞭之須知，對此須

(a) 注意調查員及被調查者之智識程度，使無論何人不致發生錯誤給予詳細之說明。

(b) 對於調查單位之規定，須預先與以周詳之研究，並指定明確範圍，使不致發生疑問及誤解等事。至發生疑問與誤解之對策，亦須預先研究妥當。

(c) 對於說明之用語，須求其最通俗化。

(d) 對於調查項目之定義須有理論與實際之根據，使任何人看了均能達到同一結論。

(e) 對於調查項目，宜避用預設分類之方法。

(f) 宜用具體之確實項目。例如問教育程度時不應詢問「教育程度優良，可及不可」等之問句，而應問其「何種學校畢業」「何種考試合格」「在某種學校受過教育幾年」較為易於得到正確答案。

四 調查時間

此時間得分靜動二態以行說明。

關於數計靜態之時間上所起問題，第一是在何年？第二是在何日？第三是為一日內之何時刻？吾人欲解決此等問題。則須先明瞭下列各點：

(1) 鑑察對象之性質：在人口普查上則選擇人類比較靜止狀態之深夜，所謂夜間調查。在職業調查上則採用失業者最少經濟活動最興旺之時期。

(2) 圖比較上之便利：例如 1872 年在俄都開國際統計會議時議決在西曆年號遇零的一年年末各國一齊舉行戶口普查，其用意即在此。

(3) 求調查上之方便：對此可分為被調查者，調查人員及調查機關三方面加以說明，依社會上習慣，在年末節期，被調查者相當繁忙，對此等時期為圖被調查者之方便，應加避免。雨季，颱風期及大寒大熱等時期是妨礙調查人員之行動，故為圖調查人員之方便，又須避免採用此等時期。至對於機關方面，在官署應切避在舉辦選舉等之時期，在私的機關則宜避開決算期及營業最盛之多忙時期。

對於動態之時期可分為事實發生時即行調查登記及經過一定時間後始行調查載錄其發生事象二種。直接統計是常將事象發生與調查在同一時期