

大學用書

醫學與生物統計方法

郭祖超 著

正中書局印行

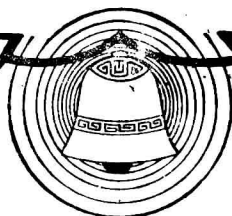
大學用書

醫學與生物統計方法

郭祖超著



正中書局印行



版權所有
翻印必究

中華民國三十七年九月初版

醫學與生物統計方法

全一冊 定價金圓券二元二角五分

(精裝本定價另加金圓券七角五分)

(外埠酌加運費雜費)

著	者	郭	祖	超
發	行	人	蔣	志
印	刷	所	正	中
發	行	所	正	中

(2327)

校整
向自

滬·本

2/2 - 0.15

序 言

三十二年春著者承乏中央大學醫學院統計學講席時，即感醫學統計方法之參考書籍，在我國甚為缺乏。從事研究工作者每借重於西籍，戰時西書難得，無待贅述。即戰後交通暢達，購置便利，而各書內容是否適合時代，符我國情，亦未敢必。余有鑒及此，爰不揣謏陋，在我國之醫學與生物科學雜誌中，儘量搜集統計資料，歷時兩載，粗具規模。經整理選擇後，乃以現代統計方法處理之，最後走筆為文。書中所列參考文獻，計有專著三十冊，論文九十二篇，其作者共一百四十一人。集此百餘位專家數十年之心血，始有本書之產生。著者得承諸先進之餘蔭，享文化之累積，其感與幸為何如耶。

統計資料首貴真實，故本書所舉例題悉係國內各專家之實驗結果，俾讀者有親切真實之感，而無不合國情之弊。即習題資料亦必詳註出處，使初學者認識統計數字不得杜撰虛構，並冀於日常訓練中養成其求真務實之習慣，此尤著者所深致意焉。

最近二十餘年來，歐美統計學之進步極速。此項新發展，就著者所知者，儘量列入本書，如自由度(degrees of freedom)、 t 測驗(t -test)、變異數分析(analysis of variance)、共變數(covariance)以及有關之實驗設計(design of experiments)等，均為國內流行之統計學書籍大多未備者，惟掛一漏萬，在所難免。尚祈海內賢達，不吝指教，曷勝欣幸。

書中所用醫學名詞，係根據中華醫學會出版之高氏醫學辭彙第九版；度量衡譯名則根據部頒名詞；統計學譯名亦大多採用部頒統計學名詞，其爲該書所未備或尙待商榷者則另譯之，以供參考。又章節前之編號在短畫以左者表示章次，右者表示節次，如1-1爲第一章第一節，7-4爲第七章第四節，餘類推。

本書所列統計資料，除大多錄自已發表之研究論文者外，其中關於肺活量及血液循環與呼吸速率者，蒙蔡翹、吳襄兩教授所供給，關於營養實驗者則蒙鄭集教授暨戴重光、王德寶二先生所供給，其他如成都進益高級助產職業學校附屬產院，及四川省立成都高級醫事職業學校附屬產院等，均供給不少寶貴資料。又在寫作過程中，蒙鄭集教授之多方鼓勵，俞煥文教授之惠借書籍，以及華大醫學院圖書室予以參考借閱上之便利。師友之賜，所惠良多。書成之日，均當深致謝意者也。

三十四年九月二日

欣逢抗戰勝利日，於錦城。

目 次

第一章 緒論... .. 1

醫學上何以需要統計學——常識的判斷和統計的處理——實驗設計的重要——學習統計學時應注意之點——參考文獻

第二章 平均數... .. 12

小樣本中均數之求法——次數表之製法——大樣本中均數之求法——中位數——衆數——直方圖與多邊圖——練習題——參考文獻

第三章 離勢量數... .. 20

全距——四分位數間距——均差——標準差——自由度——小樣本中標準差之計算法——大樣本中標準差之計算法——差異係數——練習題——參考文獻

第四章 均數之顯著性... .. 52

一個臨床的實驗——樣本均數的分配—— t 分配——均數顯著性之測驗——可信限——個別的比較與團體的比較——兩均數相差之顯著性——機差與機誤——練習題——參考文獻

第五章 直線迴歸... .. 77

迴歸方程式之計算法——標準估計誤差——修正值——用縮簡法求迴歸方程式——大樣本中迴歸方程式之計算法——迴歸係數之顯著性——迴歸之歷史的意義——練習題——參考文獻

第六章 相關... .. 105

相關係數——小樣本中相關係數之計算法——大樣本中相關係數之計算法——相關係數之抽樣變異——相關係數之顯著性——兩相關係數相差之顯著性與相關之合併——相關與迴歸——相關與離均差之平方和——練習題——參考文獻

第七章 常態曲線... .. 131

二項分配——二項分配之應用——常態曲線——常態曲線之面積——二項分配之均數及標準差——常態曲線之配合法——常態性之測驗——練習題——參考文獻

第八章 計數資料... ..169

計數資料之處理法——四格表中計算 χ^2 之簡法—— χ^2 之總和—— $R \times C$ 表中 χ^2 之計算法——獨立性與聯繫性之測驗——連續性之校正——四格表中 χ^2 機率之直接計算法—— χ^2 在遺傳學上之應用——曲數配合之適度——練習題——參考文獻

第九章 變異數分析... ..200

大小相等的三組間之比較——大小不等的三組間之比較——依一個標準分類者——依兩個標準分類的三組間之比較——交互影響之顯著性——依三個標準分類之多組間之比較——練習題——參考文獻

第十章 共變數分析... ..240

修正均數之顯著性——共變數分析之計算法——各組人數不等時之共變數分析法——修正均數與共變數分析之其他用途——各組組內之共變數——兩組資料之共變數——隨機區組之共變數——練習題——參考文獻

第十一章 多元迴歸... ..271

含有三個變數之資料——三元迴歸方程式之計算法——多元相關——估計之誤差——顯著性測驗——部分相關——含有四個以上變數之資料——四元迴歸之計算法——四元迴歸之顯著性測量——多元迴歸之一般解法——多元共變數——練習題——參考文獻

第十二章 曲線迴歸... ..308

引言——對數曲線之配合法——簡單拋物線之配合法——與直線迴歸相離之測驗——正交多項式——Logistic 曲線——練習題——參考文獻

第十三章 單一自由度... ..341

計算離均差平方和之新法——兩組以上之個別比較——均衡的比較——不均衡的比較——迴歸中之個別比較——正交多項式與迴歸中之個別比較——迴轉實驗——析因實驗——練習題——參考文獻

公式彙錄... ..366

附錄 中英文統計學名詞索引... ..379

第一章 緒 論

1-1. 醫學上何以需要統計學 生物界的現象千差萬別，而在千差萬別之中也有共同的原則可尋，若干事物之間更有相互的關係存在着。要推求共同的原則，搜尋事物的關係，研究差別的現象，進而測知所得結果的是否可靠：凡此種種，都得借助於統計學。最近二三十年來，統計學突飛猛進，更替實驗設計打下了鞏固的基礎。因此研究工作者，從研究計畫的開始，到最後結果的整理，隨時都需要統計學的知識做參考。

醫學的對象是人，人體的構造非常複雜，人與人之間又差別很大；而且醫療的結果，常受着無數獨立變化的小因子影響着。如疾病的嚴重程度，患病部分的血管分布，組織反應的力量，病人過去的歷史，體質的強弱，環境的情形，醫學的處理及其他種種因子，都會影響疾病的痊癒與否。即使我們認為一批病情相同的患者，他們彼此之間還有許多很小的差別存在着，也許這些差別目前還是無法查出的。在有些病人的情境中，剛好使他痊癒的因子占着優勢；有些病人卻反是。因此同樣的處理方法，對於各個病人未必產生同樣的效果。這些無數獨立變化的小因子，總稱為‘機遇’（chance）。臨床的診斷、病情的預測，都需要機遇、機率（probability）等概念做參考。可是醫學上對於這些概念，無論中外都不大注意。P. G. Edge 氏⁽¹⁾，在中華醫

學會第一次年會時，曾宣讀一篇論文，題為‘統計學與醫學研究’，其中有一段話大意如下：‘凡細菌學家、原生動物學家、病理學家、免疫學家以及醫學領域內其他專家的觀點和實驗結果，都受到相當的重視。但當有人建議採用統計學做一種工具的時候，卻常常被偏見和蔑視所阻止。統計學何以這樣受人懷疑，其原因何在？這是值得研究的一個問題。’氏認為‘這原因可以分做三方面：一部分由於醫學界保守分子的成見；一部分由於錯誤的見解，認為這是一門新穎而沒有試用過的科學；還有一部分卻應由統計學家自己負責，當他們呈現統計結果的時候，其所用方法，每使人感到統計學是些神祕而很艱難的數學計算，超乎常人知識以外的’。這三方面總括起來可以說是人的因子。D. Mainland 氏⁽²⁾更從日常生活和科學性質兩方面闡明了醫學上忽視機遇的原因：‘一為一切思想和活動所共有的現象，便是很多東西的計數，如錢幣、書籍、候診室中的病人等，只要數得不錯，就可得到確定的報告；又日常很多的測量，如距離、高度、重量等，在通常環境之下，可以正確到所需要的程度。還有一個原因卻是醫學與牙醫方面所特有的。因為醫學上所用量的方法，主要是從物理和化學方面來的，而在這兩門科學中，實驗誤差(experimental error)常比所測量的相差為小。’(舉例來說，把同一件東西用極精細的天秤稱上若干次，每次的結果，可能有細微的出入，這種出入叫做實驗誤差。若以輕重不同的甲乙兩物分別稱上幾次，甲物有它的實驗誤差，乙物也有它的實驗誤差，但這些‘實驗誤差’和甲乙兩物重量的‘相差’比較起來是無足輕重的。這種情形在物理化學方面是如此，而在生物科學卻不然，因為生物科學中的實驗誤差常比較的大。)本書作

者認為這些困難是有方法可以克服的。只要使人認識生物現象變異 (variation) 之大，機遇和實驗誤差占着很重要的勢力，所以不容忽視；同時用具體的例子表明統計學對於醫學研究確有幫助，經過相當時候，統計學在醫學上的應用或許可以逐漸展開吧。

1-2. 常識的判斷和統計的處理 有人說：我們過去許多的實驗結果，只要求得各組的平均數，必要時再算幾個百分數，根據平均數或百分數的大小，憑我們的常識，就可以看出各組的差別。例如我國二十歲男子的平均體重為 54 仟克（即公斤），同年齡女子的平均體重為 48 仟克，男比女重是沒有問題的，何必再要經過一番統計的計算，既添麻煩，又耗時間。這種僅憑數字表面值的推論，就是所謂常識的判斷。當兩組間的相差很大，而各組內的個別差異 (individual difference) 比較小的時候，常識的判斷和統計處理所得的結論，可能是一樣的。但是現在有十八個男嬰的平均體重為 2868 克（即公分），十四個女嬰的平均體重為 2878 克，男女相差 10 克。你能說初生的女嬰確是重於男嬰麼？又如某種疾病，用甲種方法處理了五個病人完全痊癒，另用乙種方法處理了五十個病人，結果只醫好了四十五人。倘以治愈的百分數計，前者為 100%，後者為 90%，你能說甲種處理法確比乙種好麼？常識判斷底主要缺點，是沒有把機遇考慮在內，其所下的結論，難免有錯誤的，尤其在數字表面值剛好迎合研究者的希望的時候，更容易得到似是而非的論斷。本書作者曾檢視過與醫學有關的各種實驗研究報告，其中經驗豐富的專家，以敏銳的觀察下審慎的結論的固然很多，可是也有僅憑表面的看法，匆遽地下了斷語，但經統計處理以後，所得結論恰恰相反的。在本書中將略提一

二例以說明統計處理在研究工作上的重要。統計處理除顧到機遇和實驗誤差以外，並可供給我們更多的知識，而這些知識僅憑常識的判斷是無從知道的。讀者看了後面各章，自可明瞭。

1-3. 實驗設計的重要 一般人有個錯誤的觀念，認為統計只是最後的整理和計算，至於事先的計畫、調查、實驗和紀錄等是與統計無關的。殊不知在計畫一個實驗或調查的時候，就應該洞觀全局，知道將來的結果怎樣用統計方法去處理？那些資料(data)需要詳細記載，那些項目可以略去？記載的表式如何纔能便於整理與統計？尤其重要的問題，如參加實驗的被試者(subjects)應有多少？這些被試者有沒有代表性？實驗的步驟應該怎樣安排纔能免除持久的偏性(bias)？——偏性是統計學上的名詞，現在舉兩個例子來說明它的意義。在打靶時，鎗彈打在紅星四周的機遇是相等的；但若因為鎗的內彈道欠正確，結果鎗子都打在紅星左邊。這是物理方面的偏性現象。又如調查某地中上等家庭的兒童，發現該地兒童對於白喉免疫性比他處為低。這裏面也有偏性存在。因為貧苦兒童終日邂逅街頭，與帶菌者時常接觸，獲得自動的後天免疫性(active acquired immunity)的機會較多，富家子弟保護周到，反而不容易獲得。祇根據富家子弟所得的結果，不能代表這社會中的全體兒童的。——現代統計學家，如 R. A. Fisher 氏等，對於上面這些問題，已有特殊的貢獻。氏所著 *The Design of Experiments* (3) 一書，詳述各種實驗的設計，其原理和方法可應用到各方面，為從事研究工作者所不可不讀的。

我國過去所發表的各種實驗，設計得好的固然不少，而計畫未臻完善的也很多。其中比較普遍的缺點有如下述：

(1) 實驗的方法不一致 有一個營養實驗,其控制組(control group)和實驗組(experimental group)的實驗時間各分五期,但前者每期七天,後者則每期六天。照理控制組與實驗組間,除所實驗的因子可以有計畫的變化而外,其他的因子應盡量使之相等。該實驗的目的,是在比較幾種食品的營養價值,那麼飼養‘時間’顯然是被控制的因子,而不是可變化的因子。兩組時間不等,無疑是設計上的一個缺點。

(2) 各組被試者的數目不一致 某實驗比較平臥與坐着時循環方面的變化。其女性被試者平臥時有41人,而坐着時僅7人,相差在五倍以上。但在日常生活中,每人都有平臥和坐着的時候,所以坐臥二組的人數應該一樣多。這實驗最好令同一批被試者在不同時間受平臥與坐着時的兩種測量,這樣的設計當更為合理。

(3) 各組被試者性別年齡等的分配不一致 例如關於動物生長的實驗,假使第一組裏雄的動物多,而第二組裏雌的多。這兩組生長的快慢,可能受到性別的影響,卻未必全是處理方法(如飼料)不同的結果。年齡方面也是如此,年幼時生長很快,成年後增加很少;若各組年齡的分配不同,結果就不容易比較。

(4) 各組被試者錯雜互用 第一組裏用過的一部分動物,移到第二、第三組裏來用。這裏就有兩種情形發生:(甲)一部分是一個動物先後受到幾種不同的處理;(乙)另一部分是不同的動物分別受到一種不同的處理。統計學上對於這兩種情形的計算方法是不同的,混在一起,徒然增加計算時的麻煩。

要補救上述幾種缺點,作者有下面的建議:(1)實驗的步驟必須

一致；(2)相互比較的各組所用被試者的數目應相等，如果要維持不等的比例，必須有充分的理由；(3)各組被試者的性別、年齡等分配，以及動物的窩別、原始體重等，應儘量使之相等；(4)若不同的被試者分別受到一種處理，那麼各組實驗應同時舉行，實驗時間必須一律，若同一被試者先後受到不同的處理，則應使全體被試者參加所有實驗，不能缺少或遺漏；(5)事先應有一通盤的計畫，不可陸續拼湊，因為拼湊成功的實驗，在季節、材料、技術等方面，都不容易完全一致。總之，實驗的要旨在控制別的因子，而使一個實驗的因子單獨變化。若因控制不周密，而有別的因子夾雜在內，致影響實驗的結果，這便是上文所說的‘偏性’。

參加實驗的被試者，只是從所欲研究的大團體中抽出來的小團體。假想的大團體叫做全體(population 或譯母體)，手頭的小團體叫做樣本(sample)。實驗或調查時所能得到的都是樣本，而並不是全體。樣本有大有小，但其為樣本則一。我們要使樣本為一‘具體而微’的全體，那麼這樣本纔有代表性。例如某社會的人口，中年人(十五至四十九歲)占50%，十五歲以下的占33%，五十歲以上的占17%。現在要研究某種疾病與年齡的關係，那麼被調查的一部分人口其年齡分配最好也近於這樣的比例，而後這樣本可以說是全體的縮影。假使我們對於全體的詳細情形，事先並不知道，那麼，用什麼方法纔能使樣本具有代表性呢？統計學上有一個‘隨機化’(randomization)的原則，便是解決這個問題的。例如我們要抽取某地戶口的百分之五作為研究的樣本(這過程叫做抽樣 sampling)，可先將該地所有的門牌號數順次排起來——若有兩街相鄰，甲街的門牌號數只有155號，

則乙街的 1 號可作為 156 號，餘類推。——然後每二十個抽取一個，結果就得百分之五。在開始抽取以前，先用二十個小玻璃球（取其圓滑，易於混和；如無玻璃球，改用豌豆亦可），每球上貼一號碼，從 1 起，至 20 為止。這些球放在袋裏徹底混和，然後取出一個，看它的號碼，假定是 13。這便是要調查的第一個門牌號數。以後每隔二十家就去調查一家。所以被調查的門牌號數是 13, 33, 53, 73, ……。當二十個小玻璃球在袋內徹底混和後，被抽取的機遇是相等的，換句話說，任何一家被調查的機遇是相等的，這樣就沒有好惡存於其間。所以隨機化的原則，可以避免偏性的產生，沒有偏性的樣本，纔能代表它的全體。

實驗設計還有一個‘重複’(replication)的原則，便是說，參加實驗的被試者，應當有好幾個，而不能祇有一個。根據一個被試者所下的結論是相當危險的，在適當範圍內重複愈多，則愈正確。重複除增加正確性以外，並可使我們知道機遇的變異(chance variation)有多少。假定我們選擇患某種病的十個病人，他們的性別、年齡、病情等都相同，今施以同一種處理方法，而各病人的反應並不完全一樣——有的復原快，有的復原慢。各種可以控制的因子都相等，而所生的效果並不完全相等，這叫做機遇的變異。如果另有十個同樣的病人受到第二種處理方法，同理，他們也有機遇的變異。把兩種處理結果的相差和他們自身的機遇變異相比較，便可知道相差的是否可靠。

1-4. 學習統計學時應注意之點 任何科學都有它的特性，明白了它的特性，並知道本人應有的準備，那麼在學習歷程中，可以減少很多不必要的困難。

(1) 資料必須正確 調查的結果、實驗的紀錄，總稱為資料(data)。資料必須正確可靠，而後統計的處理，纔有價值。如果資料不可靠，隨你用怎樣精細的方法去統計，這些工夫都是白費的。記得某學生做畢業論文，研究食物成分與牙病的關係，所得結果與文獻所載很有出入。他問我統計學上有無方法使結果更為合理。我從他所敘述的實驗情形發現了癥結所在。原來他估計食物分量的方法是根據兒童的口頭報告：‘今天飯吃了幾碗，肉吃了幾塊，青菜吃了幾匙，……’。這種報告是非常靠不住的。因為我們日常用膳，飯的碗數比較容易記得，而菜肴的分量卻很少人能夠記得清的。現在勉強要兒童報告，他難免隨便說個數目了事。縱使由於事前一再吩咐，他在吃飯時真能用心記下，但是碗有大有小，盛的飯有淺有滿，湯匙的情形亦復如此。你有什麼方法來作正確的估計呢？要研究這個問題，非將各個兒童每天所進的食物作仔細的稱量不可。統計學對於不正確的資料是愛莫能助的。

(2) 應有整齊清潔並愛好數字的習慣 誰都會說，整齊清潔是習慣，但未必人人都能做到。就本書作者幾年來的教學經驗，知道統計班的學生書寫潦草，數目脫漏，墨污滿紙，凌亂無序等情形很多；如數目的位置應該對齊，在小學裏早有這樣的訓練，可是大學生所寫的統計表上，同一直行的數目愈寫愈斜，末了一數的單位數竟不和第一數的千位數或萬位數在一直線上；作圖畫線，理應用尺，可是很多人往往順手一揮，管它是曲是直；如果要打一張格子，每格寬為一厘米，打好後一看，格子的寬窄不一，相差二三毫米是常事。這種馬虎草率的習慣，實不配從事科學的研究。又統計資料既然都

是數字，那麼學習統計者對於數目應養成愛好的態度，而不宜有厭惡的心理，這正與學醫者不該討厭病人是一樣的理由。如果他一見數目就感頭痛，這時他的心理上早已築了一道很厚的圍牆，那裏還能吸收統計學的知識呢！

(3) 應養成隨時核對與推理的習慣 數目的抄寫和計算，一不留心就會發生錯誤，這情形在有經驗者尚免不了，何況初學？不過有經驗者一遇錯誤，自己就會發覺立即改正。發覺的方法，主要的是靠核對(check)和推理(reasoning)兩方面。核對可以把原來的計算過程重複一遍，也可以用不同的方法再算一遍。但重複計算，同一的錯誤有重複發生的可能，故統計上常用不同的方法來核對。例如一行數目相加，第一次可由上而下，核對時則可自下而上；又如原來用的是減法，核對時可用加法還原，看結果是否等於固有的被減數；又乘法與除法，開方與平方都可彼此核對；諸如此類，不勝枚舉。推理也是發現錯誤的好方法。例如求若干個數值的平均，各數都在 50 以上，而求得的平均數只有 48，這一望而知是錯了；又如 25.5×3.2 等於 81.6，而不是 8.16，因為 25 的 3 倍是 75；又如我國成年男子的身高約在 160—190 厘米之間，若某人的身高紀錄在 200 厘米以上，其正確性也是值得考慮的。所以豐富的知識、過去的經驗，都能增加我們的推理力，而有利於錯誤的發現。

(4) 要明瞭公式的用途和計算的意義 初學者看見了統計公式，每喜歡追問來源。固然公式的來源為講授統計學者所不可不知，但統計學上有很多公式是用高深的數學原理演化出來的，初學者如果沒有充分的數學根基而要追究公式來源是非常困難的。本書作者

對於求知慾特強的同學會講解過不少次的公式演化，終因他們的數學根基不夠，而不能充分了解。本來公式的導源是屬於數理統計學 (mathematical statistics) 的範圍，在統計方法 (statistical methods) 中可略而不提。初學者必須明瞭的，是各種公式的正當用途和計算的意義；至於公式的來源，必須有了相當根基纔能研究。

一批資料應該用什麼方法去處理，要看研究的問題而定。假如研究身高體重生長的情形，那麼就要分別製成生長曲線；假如想從身高來推算體重，就要求出它的方程式；假如研究性別差異，就要看男女的平均數有無真正的相差。所以我們先有待解答的問題，然後去選擇統計的方法，方法用得適當，所需的答案就能充分表現出來；方法用得不適當，那麼他的問題也許只能解答一部分，甚至會弄到牛頭不對馬嘴。又統計學上同一類的公式往往有好幾個形式，應該用那一個形式，亦依資料的情形而定。公式用得適當，計算時就非常簡便；否則就吃力不討好。總之，每個統計公式都有它的正當用途，一切計算都有它的意義，在計算以前，必須弄清楚為什麼要這樣算；切不可依樣畫葫蘆的盲算，盲算會使人墮入五里霧中！

(5) 學習要循序要多做練習 統計學是一門應用數學，數學最注重先後的次序，故學習時必須循序前進，前面的都懂了，然後再看後面；千萬不要在中間翻出一段來看，因為這樣看法，對於統計學沒有相當根基的人是有害無益的。還有數學要做練習題，統計學也要做練習題，只有實地練習纔能真正了解。本書各章之末，附有練習題若干個，希望初學者能逐一演算。全書的練習題做完以後，還要自己去找一批資料，仔細地分析與統計，因為書本上的練習題是經過