

18卷
C 80
CVC

C B A 學 計 統

著 吳 錫 恭

第九九號

版 出 局 書 界 世

八 二 九 號 特 雜

834421



A B C 叢書發刊旨趣

徐蔚南

西文 A B C 一語的解釋，就是各種學術的階梯和綱領。西洋一種學術都有一種 A B C；例如相對論便有英國當代大哲學家羅素出來編輯一本相對論 A B C；進化論便有進化論 A B C；心理學便有心理學 A B C。我們現在發刊這部 A B C 叢書有兩種目的：

第一 正如西洋 A B C 書籍一樣，就是我們要把各種學術通俗起來，普遍起來，使人人都有獲得各種學術的機會，使人人都能找到各種學術的門徑。我們要把各種學術從智識階級的掌握中解放出來，散還給全體民衆。A B C 叢書是通俗的大學教育，是新智識的泉源。

第二 我們要使中學生大學生得到一都有系統的優良的教科書

或參考書。我們知道近年來青年們對於一切學術都想去下一番工夫。可是沒有適宜的書籍來啓發他們的興趣，以致他們求智的勇氣都消失了。這部ABC叢書，每冊都寫得非常淺顯而且有味，青年們看時，絕不會感到一點疲倦，所以不特可以啓發他們的智識慾，并且可以使他們於極經濟的時間內收到很大的效果。ABC叢書是講堂裏實用的教本，是學生必辦的參考書。

我們爲要達到上述的兩重目的，特約海內當代聞名的科學家、文學家、藝術家以及力學的專門研究者來編這部叢書。

現在這部ABC叢書一本一本的出版了，我們就把發刊這部叢書的宗旨寫出來，海內明達之士幸速而教之！

序言

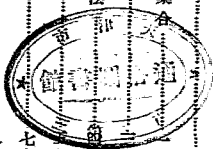
統計學這一個名辭，輸入中國，還沒有好多年。中文統計學書籍的稀少確實是件不容否認的事實。但是我們的學者，卻又偏偏只肯供給鴻篇巨著，使站在分利時代的青年，簡直無力購買。可是統計學並不是一種裝飾品，它是種很切實的學問，是現代人應備的學問。本書就認清了這個需要，對於必要的公式等等，儘量的加以介紹。本書是在很匆促的時間中，所擬成的，萬一有錯誤的話，務請讀者諸君，不吝指教。

中華民國十七年七月二十日 編者序於上海工整會統計室



目次

| | | |
|-----|------------|-----|
| 第一章 | 導言 | 一 |
| 第二章 | 統計學發達史略 | 八 |
| 第三章 | 統計資料的來源和集合 | 一九 |
| 第四章 | 統計表編製法 | 二五 |
| 第五章 | 統計圖的種類和作法 | 三七 |
| 第六章 | 平均數 | 三八 |
| 第七章 | 變量 | 三七 |
| 第八章 | 相互關係 | 八九 |
| 第九章 | 統計書報述要 | 一〇二 |



統計學ABC

社會學學士 蔡毓聰

統計學 A B C

第一章 導言

最近百數十年來，歐美的統計學者出得很多；但是大體各抱各的意見，各有各的注意點，在別一方面，便竭力排斥他人的意見並攻擊他人的注意點。這種現象，實在可以說，是一般科學在未成熟而將成熟之間所常發生的，倘若因了這種現象而竟然否認統計是一種科學，那就大錯而特錯了。

統計的科學性，久已為西洋學者所公認；但是事實上，統計並不僅僅是一種科學，牠也是一種方法，一種技術。英國的著名統計學家茄羅米教授 (Prof. Harry Jerome) 便承認統計是一種技術，他說：「統計是從大量的數字事實中間把顯著

的真實表現出來的技術。「他的所謂『數字事實』，就是用數字記載的事實，舉個例來說，人的年齡便是用數字記載的，例如：二十三歲，三十五歲，以及其他等等。其他用數字記載的事實，真是多得，這裏也不預備再舉了。總之，「數字事實」是合於統計的。現在我們且把茄羅米教授的意思，向上面所講的年齡這一項來引述一下：倘若我們要知道某一羣人中間的年齡，應用了統計的方法，我們便可以真實地把牠表示出來。

在統計這一個標題下面，我們還可以把牠分做許多步驟，其中最最重要的計有三個，便是：集合資料，分析事實，和表現結果。

最初當然向各方面竭力徵求所需要的資料，多多益善，

然後再仔細地去考察這些資料是否可靠，並且把不可靠的全數去除。

第二步是分析事實。要真實地把我們所聚合的資料顯示出來，我們必須先將所認為合格的資料分類一下。然後用文字或圖表把這個分類記述明晰。只不過根據醫生所報告的嬰孩死亡率，或者政府及殯葬人所報告的，都是並不能滿足我們的要求。我們所要知道的，並不僅僅是每千個嬰孩中的死亡率，而尤其要知道的，還是他們致死的原因。獲得了他們致死的原因，我們便可以進一步，在可能範圍以內，儘量地設法把死亡率降低。

良好的聚集和科學的分析之後，統計人員就可以得到重要的結果。統計員便應該把這個結果表現出來，有的用文字

，有的用表格，有的用各式各樣的圖解。統計表的編製，可以說是統計實踐上的最重要的一步；大多數，絕對大多數的統計事實，在手續完成以前，都是經過了製表這一個步驟的。

統計對於一般社會科學研究者的重要，在現代尤其顯著，大抵有三個原因，現在且分別敘述如下：

(一) 現在是一個大數的時代 一國的人民，有的多到幾萬萬；戰爭起來了，荷鎗上戰線的有時也竟然有幾千萬；此外每年所收穫的糧食，所徵收的稅銀，結婚的數目，離婚的數目，生育的數目，死亡的數目，以及其他種種，那一個不是數量很大的。倘若只選了其中一二特殊的事實，我們便可以說是顯然地錯誤了的。至於這一二個事實究竟是否是特

殊的，當然非大量觀察不可，因為這是一個大數的時代。

(二) 現代科學的根據是事實。科學之所以和哲學不同，因為牠有事實上的根據。現代的社會學說，大都完全是根據事實的，簡直不容參加個人的偏見。自然科學，自然更不要說了，不容參加絲毫的個人偏見。自然科學是可以實驗的，所以成績是很不錯。可惜社會事實並不是這樣，使得幾個人各方面的情形完全相同，可以說是絕對不可能的，那麼我們可以肯定地說，社會事實是不能實驗的。社會事實既不能應用實驗法，但是同時社會科學又須符合現代科學的條件——根據事實，那麼最妥適的辦法，當然是採用統計法了。

(三) 商業的程序漸漸地成為科學的步驟，使統計法的應用有了很大的收穫。

總而言之，統計法的本身是並沒有多大意義的，牠是一種工具，是一種求事實的真實的工具。

統計法的應用，近年來愈形廣大，現在擇要列舉如下：

(一) 在社會現象的分析上面 統計法應用得很廣，藉之得能明瞭牠們的關鍵，以謀改進的辦法。恩其爾 (Engel) 的費用分配律，便是這種分析的結果。

(二) 經濟衝突上的應用統計 生活費、工資、完稅、等等的統計，在經濟衝突上，尤其是勞資衝突上，是極有用處的。各大都會的生活費指數，在中國政治稍為澄清之後，大概都要編製起來的。

(三) 行政上的應用 人口統計，疾病統計，結婚統計，歲種穀類統計，都是屬於這一類的。

(四) 商業上的應用 壽險、火險、投資等等，均需要商業統計。其餘如物價、生產量，亦均與商業統計有莫大的關係。

(五) 慈善事業上的應用 每人最低的平均生活費額，各種事業的實效，最經濟最有效果的慈善事業應該如何辦理，該地方最需要的慈善事業是那一種，都是待統計來解決的問題。

(六) 教育上的應用 兒童年齡統計，家長職業統計，列年學生多寡統計，學生籍貫統計，都是極有意義的教育上資料。

(七) 心理學上的應用 例如：閱讀能力的統計，反應速度的統計，記憶能力的統計。都是統計應用在心理學上的

明顯表現。

(八) 生物學上的應用 關於生物長度，生物生產地，以及生物質量上的差異，都在在可以用統計法來研究的。

(九) 其他 如物理學、化學、地質學、氣象學、天文學、人類學等等，截至最近，亦均相當地採用統計法了。

第二章 統計學發達史畧

統計的歷史，實在是很久遠了。禹貢、周禮是中國古代的兩本統計冊。土壤的肥瘠，物產的多寡，田賦的高下，運輸的路徑，以及兵力等等，都已有相當的敘述。戰國以後，更有相當的進步。當時政客，往往能舉各國的兵數物產，以徵實其政策之合用。近人有在敦煌發現戶籍調查的文件，則中國的戶籍統計，當由來已久。惜中國民族好靜性太甚，以

致統計學在中國不僅不見昌明，不見發達，而反見退步。

統計學在西洋的起源，也是由來頗久。埃及時代，亦曾舉行全國人口與財富統計，以爲徵收金字塔建築費的根據。古希臘與羅馬，亦已有相當的統計，但是定期舉行的卻還很少。

中古以來，西洋的著名帝王，大都均能利用統計，以爲行政上的參考。自此以後，政府中人以及一般學者，遂羣起研究。統計法遂亦因之日漸發達。十七世紀以後，統計學漸趨成熟，各學派均先後成立。現在我們把其中較重要的幾派，分別敘述在下面，或可因之而知統計學發達的程序，對於統計學便可以有一個更深刻的了解。

(一) 官廳統計學派 統計學這一個名辭，在英文上爲

Statistias，在法文上為 Statistique。兩者均由意大利文政治家 (Statista) 一字推行而成。統計學與官廳關係之密切，由此可見一斑。

統計學的始祖，是馬堡 (Marburg) 大學教授阿痕華爾氏 (Gott. Achenwall)。但是阿痕華爾氏還不是官廳統計學派的主將，而是官廳統計學派的先導。阿氏以為統計是國家顯明事件的結晶體；就是說，我們知道了某一國的統計，便知道了該一國的真相。

由此我們可以知道，阿氏當時所定的統計範圍，異常廣泛，簡直無所不包。其資料的整理和蒐集的方法，又均無一定的程式，當然遠不及現代的統計方法。

其後，格丁根 (Göttingen) 大學教授休爾麥爾始將統計範

圍稍為縮小。氏以為並非任何事物均可納入統計範圍以內；歷史地理即不屬於統計，同時氏將政治地理屬於統計範圍，因其與國家治安有關。休氏曾言：「統計是停頓的歷史，歷史是流動的統計。」這二句話確是極著名的格言。

官廳統計學派 (Cameralists) 的起始，則在十八世紀的皮奧氏。氏著有國家學述要一書，內含計算資料不少，較阿痕華爾氏進步多多。採用比較的方法，實為該學派的特長。

(二) 表記統計學派 十八世紀中葉，丹麥人安歌遜 (Andersen) 編製統計表，舉凡各國的重要現象，如面積、人口、宗教、財政、軍隊、度量衡、貨幣等等，均一一以數字表出之。其研究的方法，僅將該項國家政治上的重要現象，藉數字羅列於一表中而已。未能對於一種事實的資料，盡量的

蒐集，實係該學派的缺點，因之根本的說明，事實上是很難辦到，所以表記統計學派並沒有多大的貢獻。

(三) 政算學派 政算學派 (Political Arithmetician) 是上列三學派中對於統計學最有貢獻的一個。該學派已知大量觀察的重要，並承認社會現象中間，有一定的規則，也有因果的關係。

該學派的始創者是英人葛蘭溫脫 (Capt. J. Graunt 1620-1674)。氏對於人口生死的統計，有詳密的研究。氏以為人民的生育率與死亡率均甚固定；男孩的生育率高於女孩的生育率，為十四與十三之比，頗正確可靠。

葛氏以後，政算學派的柱石，當推配第 (W. Petty 1623-1687)。配氏對於數學造就至深，實係最初注重平均數的一人。

其所著政治算術一書，即係該學派的代表作品。其所包含的範圍，亦極廣泛，土地的闢度及價格、人民住宅、農產、製造、漁業、工業、商業、水手、兵士、利息、租稅、收入、海員、港灣、船舶等等，均有相當的敘述。

其後有赫雷 (Edmund Halley 1656-1742) 氏者，亦係政算學派的健者。赫氏研究人口統計，極有心得；其所製人壽表，不僅有功於統計事業，人壽保險的基礎，亦由此植立。保險的期限以及保險費的多寡，都取決於該項人壽統計表。但當時取材範圍太小，且忽略人口增加這一點，不得不認為遺憾。

(四) 大數觀察的唱導 普爾有教士叫做蘇斯密爾 (John P. Ssemihloh 1707-1787) 的，於一七四一年刊行一書，從人口生育死亡及增加等各方面，證明男女變化上的自然秩序。在

蘇氏以前，學者於人類生死的現象，只有少數的個別的研究；綜合的研究，實始自蘇氏。但蘇氏身為教士，故不免有宗教的偏見，因欲趨合一夫一妻制的教義起見，特表明在結婚年齡之男女的數目，幾於相等。其後英人馬爾薩斯 (C. B. Mal. thus 1766-1834) 之人口統計，因欲顧全其自身的人口學說起見，特引取與彼學說相符合的各國，實亦犯同樣的錯誤。彼等大都先有成見在胸，然後以觀察所得資料為補充；其所倡的學說，因之亦常偏於演繹的。這實在和統計的本旨相反，是一個很大的缺點。西教士和帝國主義者所造的關於中國的種種統計，大都亦先有成見在胸，陷於同樣的錯誤。

蘇氏之後，法人蓋利，在一八二九年刊行法國犯罪統計一書，饒有價值，開近世社會統計的開河。同時有比人刻特

雷 (Lambert Adolphe Jacques Quételet 1796-1874) 者，為著名的天文家與數學家，對於統計學極具興味，在天文及氣候學範圍以內，曾作極廣泛的統計研究。因研究氣候的關係，刻氏進而以統計的方法研究植物界、動物界以及人類的一切現象。刻氏實係統計學界罕有的傑出人才，不僅使統計的應用擴大，並且使統計的理論亦有一番改進。

中庸 (Mean or Norm) 存在這一原則的發現，係刻氏在統計學史上不可消滅的貢獻。刻氏統計大自然中各種事物的結果，認為世界上一切形形色色的事物中間，都有中庸的存在。中庸就是中等的一部份，大多數的事物，都和牠相近；離中庸愈遠的，其數目也愈少。天才是很少的，但是同時白癡也不不多。大多數的人，都是相去不遠的；智慧上講，是這樣的

；體格上講，也是這樣的。

利氏所著的社會物理學 (Physique Sociale) 與人體學 (Somatologie) 兩書，都是很具精彩的著作。後者於測定人數和才能的差別上，尤有獨到的地方。

利氏曾主持比利時的行政統計，能以學術的興味參加在內，所以極有成效。一八四八年比國所試行的人口統計，就是利氏所主持的，到現在還公認為一個良好的人口統計。

(五) 統計學上的派別，僅上述的三四種而已。但是統計學得能發展到現在這一個地步，初不僅僅乎依靠了這三四個學派。行政統計的公開和統計學者的合作實是在其中最重要的兩個原因。採集資料的權力，向來是只有官廳享有的；但是官廳所獲得的資料，勢必又須經過學者的手，方可成為

實際的統計。事實上却很可憐，官廳方面，往往以為其所獲得的統計資料是國家的秘密，不願意使民間知道，不允許學者去研究。日積月累，官廳方面的統計資料逐漸豐富，但官廳所刊印的統計，因執政者的偏見，仍很少在民間流行。當時該項資料，大都有一共同的缺點，就是沒有系統的大數觀察。

最初公開的官廳統計，便是人口統計，但齊整的刊印也並不多見。其他統計書表，直至阿痕華爾時代，還只有官廳中人，得能閱及，及至皮與氏時代，統計資料方纔可為統計家私人所獲得。

皮與氏的著作，頗多引用統計資料，其一七九〇年在漢堡 (Hamburg) 出版普王弗蘭特第二治世紀中的附錄，詳載人口

、貿易、財政、軍隊等等的狀況，極足珍貴。

歐洲各國，其後設置統計局者亦頗不少，然統計公布後，時時發生反動，故雖特設統計專局，亦仍不能無間斷的刊印各種統計。普國統計局亦遭同樣的命運，各種統計結果均不在官報發表。直至一八五〇年，柏林統計局始有普魯士官廳統計及報告的刊行。

在這個時候以先，英國於一八三二年以來，有貿易統計表的刊行。法國的貿易統計表，刊印較早，起於一八一六年。法國又於一八二六年起，有犯罪統計表的刊行。

(六) 學術統計的勃興 至十九世紀中葉，官廳統計的結果，大都公布，統計學界便頓時起了一個很大的變更。統計技術，頓時有了一個猛烈的進步。恩其爾 (Ernst Engel) 撰

美洽 (Rumelin) / 格尼司 (Kries) / 華格納 (Adolf Wagner) / 披爾遜 (Karl Pearson) / 布勞克 (Mr. Block) 等等，都是對於統計法有極大貢獻的。厄治衛斯 (Francis Edgeworth) / 桑戴克 (E. L. Thorndike) / 猶爾 (G. U. Yule) / 高爾頓 (Francis Galton) / 丹溫泊脫 (C. B. Davenport) / 薄荷雷 (Bowley) / 霍勾 (R. H. Hooker) / 亞當斯 (Thomas S. Adams) / 潘遜斯 (Warren Persons) 等等，都是統計學的功臣。

第三章 統計資料的來源和集合

統計資料，大概可以分做二大類，就是：直接的和間接的。直接的資料就是直接集合得來的；間接的資料是根據別人的資料並非直接集合得來的。直接集合比較上是費錢很多的，所以統計上所採用的資料，往往是間接的。

直接的集合，通常有兩個方法，就是：

(一) 派人和對方會面 Personal Visit。

(二) 填寫表格法 Questionnaire method。

第一種方法——派人與對方會面——有二個優點：就是詳盡和真確。不過第二點是全以調查員的能力為比例的。限期不誤，也可以算是這種方法的第三個優點。倘若我們預備在十五天以內，將全上海的絲廠工會調查清楚，全上海絲廠工會只有九十七個，那麼我們規定每天調查七家，十四天便可以完全完畢。

這裏所要注意的，所謂「對方」也者，初不僅僅乎指點被調查者的本身而已，有時對於與被調查者有關係的各方面，也須加以相當的注意。據歐美個案調查家的意見，有關係

的各方面大概可以分成二十二個，雖然他們在程度上是很有差別的。現在把他們介紹在下面：

- 一 親戚（最重要）
- 二 朋友（最重要）
- 三 從前的鄰居
- 四 現在的鄰居（次要）
- 五 從前的雇主（次要）
- 六 現在的雇主
- 七 醫生（重要）
- 八 醫院及療病院
- 九 衛生局
- 十 看護

- 十一 醫藥補助機關（次要）
- 十二 學校及教員（如調查小孩，極重要）
- 十三 教會、教堂、及教士。（重要）
- 十四 從前的房東（次要）
- 十五 現在的房東（重要）
- 十六 公安局
- 十七 生育登記處
- 十八 結婚登記處
- 十九 法庭
- 二十 律師
- 二十一 慈善機關（次要）
- 二十二 工會及其他職業上的組合（重要）

下面七條原則，對於資料的採用都各有牠的相當價值，極值得我們去注意的，現在錄在下面：

一 和被調查者認識的年限愈長的人，所說的話也愈較可靠，因為他所知道的内容亦愈加透澈。

二 先找熟悉被調查者歷史的人，就是被調查者的親戚、朋友，一切從前的伴侶，然後方纔再找其他的人。

三 倘若所調查的是一件事，能夠找到當時目視的人，那是最好的了。

四 對於各種補充憑證之有特殊價值的，概予承認。

五 對於所訪問者的地位，是否有偏見的，應加以注意。

六 日常與被調查者相接近的人所說的話，較為可信。

七 資料如發現相衝突時，應互相比較，以便相當承認

在通常的調查（即集合資料）中間，事實上是並不需要這樣的詳細和周密。這種派人和對方會面的方法，同時也有兩個弱點。一，太費時間，因為每一個家庭或工會的調查，往往有化費到二三個鐘點以上的。二，結果的能否完備，不能預測結果的優劣。

第二種方法——填寫表格法，就是先將所需要的資料，編成問題式的表格，寄給所需要調查的對方，靜待牠們的寄回。這種方法也有幾個優點，其中最大的一個，便是省時間。通常發了一次表格以後，再發一次前往催促的信函，便可以得到相當的結果。所發的表格和前往催促的信函，當然

都是印刷的。那麼，一紙填就的表格的時間上代價，大概只不過書寫二個信封的時間罷了，最多也不過七八分鐘。不過被調查者有否填寫表格的能力，這是事前應該鄭重考慮一下，否則竟然會使書寫信封的時間完全是白費了的。第二個優點，可以無須優良的調查員。所以就時間上和經費上着想，第二種方法都比較上來得經濟。第三個優點，便是便於研究一種制度。藉着郵寄的方法，很遠地方的制度，往往在很短時期以內，便可以得到答案。至於弱點呢，正是第一種方法優點的反映。第一，表格不能十分詳細；否則足以引起對方的嫌惡，以至於不高興填寫。第二個弱點，被調查者有充分的考慮時間，為顧全體面或其他種種起見，便可以儘量作偽。第三個弱點，是不能限期。上海工整會這次發給各工會臨

時工會調查表等三張表格，原定十日結束，後來展期了二十天，還不能結束。割愛呢，往往亦為事實所不許。再展期呢，展到那一天，全數可以收回，又誰都不敢擔任。

所以在實際上的應用起見，採用折衷辦法。先將表格發給對方，到了一個相當期間以後；未填來的，派人去調查；填寫不完全的，也派人去調查。這種辦法，比較上是最妥當，時間上和經費上，所費比較上也還不多，確實是最可採用的辦法。

直接的集合資料，上面已經有了足夠的敘述，現在我們且來談談間接的集合資料。間接所獲得的資料，大都已經經過他人的一番整理，例如公家或官廳的報告和統計。

間接所集合的資料，大都是官廳，公共機關，以及各種

法團的紀錄。這些記錄有的是已經印成了冊子的，所以收集方面並不至於會感覺到困難。但是牠的價值却是很高，特別是政府的人口調查報告。西洋學者的言論，引用該國人口統計的，是很多很多。人口統計，通常能夠告訴我們：該區域中男子的總數，女子的總數；農民的總數，商民的總數，工人的總數；幼童的總數，成年的總數；識字者的總數，不識字者的總數，以及其他種種有價值的統計。

官廳方面的報告，以公安局、法院、衛生局、教育局、社會局、以及中央的勞工司等等的記錄，為最有價值。

公共機關方面，如：醫院、公園、圖書館、博物院、體育場、公共大會堂等地方的統計，以及慈善機關，學校等等的記錄，都是很有價值的資料。

工會、商會、商民協會、農民協會、學生會、律師公會、會計師公會以及國民黨黨部等等法團的記錄，有時也有極有價值的參考資料。

不過該項間接獲得的資料，必須經過嚴格的檢定，方始可以應用。該項檢定，大檢可以分做二個：一，對於文件的外表批評，包括着對於作者，資料的性質和分類，以及文件的形式等等的批評；二對於文件的內容批評，包含着作者（即敘述者或編製者）所敘述的確實意義，以及作者的忠實及真確之類的問題。

文件的外形，通常是最先受人注意的。在外表批評中間，我們最少應該得到下面幾個問題的滿意的答覆：這個文件在什麼地方和在什麼時候產生着的？當時有什麼目標和作用

沒有？敘述者或編製者是誰？他所採用着的是些什麼資料？這些資料是從那裏得來的？可靠不可靠？

當正式批評的時候，對於各種資料的分類是萬萬不能少的，因為由此我們可以知道，敘述者或編製者對於資料的選擇，並不是一件任意的或者偶然的事實。附有評語的書報目錄，有時也必須編製。至於分類的方法，却有許多種類。中間最良好的，要推用活葉本這一個方法了。這個方法的實施辦法，就是在檢閱這些文件的時候，用一册活葉的本子，把所有的要點一個一個地都記了下來。然後用大小一律的紙片，每張上面都記下一個要點。再次把要點內容相同的或相類似都放在一處就可以供檢查的用處了。著作者的姓名，書報的名稱，出版的日期，出版的地址，付印的次數、冊數、編

數、章數、節數、以及頁數等等，是都應該仔細地記在上面，並且選擇幾個較重要的排列法，依着這些程序去排列。

史家對於檢驗文件中事實的相對價值，也發見了幾個有價值的原則：一，當原稿尚保存着的時候，以原稿為主；二，當唯一的抄本或印本尚保存着的時候，只得該抄本或印本為主；三，當許多抄本或印本同時保存着的時候，那麼，我們必須將相異的地方互相比較一下。在通常應用中間，我們務須記憶着，原稿是最可靠的，因為原稿的誤點在抄本或印本中間通常是依舊存在的，同時抄本或印本中間的誤點，却不純粹是從原稿裏得來的。

文件的外表批評，是有範圍的並且有危險的。所謂有範圍也者，就是講，外表批評不過是對於內容批評的一種預備

工作。所以外表批評，並不是最後的目標，只不過是求取最後目標的途徑。至於外表批評的危險，是很容易養成只愛涉獵而不求精深以及吹毛求疵的人。

內容批評是更重要的工作。所以統計名家薄荷雷說，當我們念一個統計表以前，最好先坐下來，並且靜靜地思考一番，為所需要的目標而聚集的統計應該聚些什麼資料，那一種資料是應該採用的。假使所研究的對象是工資，那麼，除了家長所得的工資以外，工人家庭中間其餘份子的補家費，失業時期的年金，以及其他種種關於該家庭收入的問題，也都應該鄭重地考慮一下的。

文件檢驗的心理方面批評，大概集中在二點：一，文件內容的分析，使作者的原意不至於失去或誤解；二，對於文

件產生時候之情形的分析，藉以明瞭當時有無阻止作者言論的妨礙，可以決定這種文件的能否採用。倘若沒有這二種分析，一切文件就很難以當作參考或採用的資料了。

採用活葉本或硬紙片，對於這類分析是極有助力的。每張紙上面必須注明：從文件中間那一部份摘錄下來的，全文的大意是怎樣，作者的目的是和見解是怎樣。

在文字方面，還有兩點是應該格外注意着的。第一，當採用外國人所著的文件的時候，對於該文件的譯者的忠實與否也須相當地留意着。第二，當採用術語的時候，對於各個名辭的範圍，也不可加以注意。譬如：外國人民這一個名辭，是只指在外國產生着的人們，還是只指外國籍父母所產生的人們，還是二種都應用着的，都包括着的。此外，本國

之入外國藉的，是否就算外國人民。這些都是不得不辨判清楚的。

有時作者在同一文件中間，前後有了矛盾的敘述。那麼，當採用的時候，倘若必須採用的話，非格外留意不可。

當作者的真意明白了以後，我們應該更進一步，去估量作者忠實的程度和所敘述的事實是否真確。

有些人以為資料衆多的記載，其中所敘述的事實必定都是真確的。這個觀念，我們可以說是完全錯誤的。事實上通常所謂資料豐富的洋洋巨著，往往只把些雜亂無章的資料儘量地收入着罷了。

有時作者因為自身利害的關係，存心欺詐着的也是常常有的。公文裏面大部份是這樣的，國會或議會裏面的報告裏

面，尤其是這樣。報紙大都是有黨派色彩的，所以也常常是不可靠的。社論中間的黨派色彩，尤其顯著，是大家所知道的。其實登在消息欄中的資料，也時時受着黨派的支配。長期廣告費，往往影響到報紙上的記載和言論。在於私家的刊物，例如：公司的招股書，工廠的報告書等等，更含有不少的欺騙。

作者是否懷着旁觀者的態度，實在也很值得研究。作者的有沒固執的意見，或是偏於一方面的同情，委實是個大可注意的問題。

有時作者是個神經衰弱，受了重大刺激的，或者懷着偏見極深的人，那麼，他的觀察是很難真確的了。

總括講一句話，在採用現成的統計資料的時候，我們應

該格外的注意，格外的仔細，才不至於跑到錯誤的路上。

第四章 統計表編製法

統計表的編製，是統計實踐上所必不可少的步驟。當資料集合以後，第二個步驟便是編製統計表。等統計表編製完竣之後，才可以着手畫製統計圖。

朱君毅教授在所著的教育統計學一書中間，把統計表的功用舉了五個，現在我們來把牠說明一下：

(一) 提綱挈要 資料未分類並製成統計表以前，簡直毫無倫次；讀者倘若要翻閱一遍，費了很多的時間，往往還得不到綱要。製成了統計表以後，把所有的事實都變成了數字的，並且循着自然的程序；只須費了很短的一個時期，最多也不過十分鐘，資料中的重要事實便能夠使讀者有一個深

切的印象。這種費十分鐘才能閱畢的統計表，原稿大概非二小時不能閱畢，所以閱讀統計表，時間上比較上可以說是很經濟的了。

(二) 輔助記憶 事實的排列，既然有合理的，系統的次序，自然容易記憶。和通常蕪雜未整的資料相比較，記憶的便利上自然有很大的區別了。

(三) 便利比較 資料未經編成統計表以前，全體間的彼此關係，大都是很茫然的。事實上有時費了很多的力量，但是成效卻沒有什麼。一經製成了統計表，全體的關係，便可以一目了然。比較或對照，都是成爲很方便的事情。

(四) 便利總結 倘若要知道許多事物的總數，統計表實在是必不可少的東西，因爲我們把這些事物都列在幾條縱

線或橫線上面，要計算實在是很便利。反轉來講，倘若我們把這些事物散記在許多張紙的上面，不是費了極多的時間和精力，恐怕是一無所獲。

(五) 免除重複 事實的報告，倘若不是製成統計表，同一的事實，往往有二次、三次，甚而至於五六次的重複。但是當我們採用統計表的時候，各項事實大都有類可歸。但是一經歸類，便不至於再會反覆發現。

統計表可以分成一重的，二重的，三重的等三種。其採用的標準，可以完全根據牠們事實上的需要。現在列舉如下：

(一) 一重的統計表（即單純的統計表）這種表是只表示一種事實的，如第一表內人數的分配，是以學級為標準的

第一表

復旦大學

十六年秋季

分級統計表

| 數 人 | 級 學 |
|-----|-----|
| 8 | 生究研 |
| 108 | 級年四 |
| 168 | 級年三 |
| 208 | 級年二 |
| 272 | 級年一 |
| 68 | 科 預 |
| 99 | 生別特 |
| 931 | 數 總 |

如：第二表中的級別和科別。
 (二) 二重的統計表 這種表所表示的事實有二種，例

第二表

復旦大學

十六年秋季

| 總 數 | 級 科 別 別 |
|-----|------------|
| 931 | 計 統 |
| 187 | 科 文 |
| 114 | 科 工 理 |
| 326 | 科 商 |
| 15 | 科 學 物 生 |
| 141 | 科 學 科 會 社 |
| 66 | 科 學 文 國 中 |
| 82 | 科 預 |

C B A 學 計 統

子。計表上的，這張表就叫做三重的統計表。第三表便是一個例子。

(三) 三重的統計表 有時有三種事實必須列在一張統計表上的，這張表就叫做三重的統計表。第三表便是一個例子。

學生分科分級統計表

| 特別生 | 預科 | 一年級 | 二年級 | 三年級 | 四年級 | 研究生 |
|-----|----|-----|-----|-----|-----|-----|
| 99 | 68 | 272 | 208 | 168 | 108 | 8 |
| 16 | | 61 | 38 | 42 | 27 | 3 |
| 16 | | 34 | 32 | 19 | 13 | |
| 15 | | 100 | 97 | 64 | 48 | 2 |
| 4 | | 4 | 2 | 1 | 4 | |
| 14 | | 43 | 31 | 34 | 16 | 3 |
| 20 | | 30 | 8 | 8 | | |
| 14 | 68 | | | | | |

第三表
復旦大學
十六年
秋季學生
統計表

| 別 | 性 | 生正是否 | | 類別 | |
|----|-----|------|-----|-----|-------|
| | | 特別生 | 正式生 | 學 | 科 |
| 50 | 881 | 99 | 832 | 931 | 數總 |
| 3 | 184 | 16 | 171 | 187 | 科文 |
| 1 | 113 | 16 | 98 | 114 | 科工理 |
| 2 | 324 | 15 | 311 | 326 | 科商 |
| 3 | 12 | 4 | 11 | 15 | 科學物生 |
| 6 | 135 | 14 | 127 | 141 | 科學科會社 |
| 15 | 51 | 20 | 46 | 66 | 科學文國中 |
| 20 | 62 | 14 | 68 | 82 | 科預 |

這一種三重的統計表，是比較上簡單的了，是可以伸長的，是可以應用到四重，五重，以至於五重以上的。就事實來講，四重的統計表，也是已經很少的了，五重的以及五重

第四表
復旦大學
十六年秋季

| 式正 | | 類 | |
|-----|-----|-----|-----------|
| 男 | 總 | 別 | 科 |
| | | 數 | 別 |
| 802 | 832 | 931 | 共 總 |
| 169 | 171 | 187 | 科 文 |
| 97 | 98 | 114 | 科 工 理 |
| 309 | 311 | 326 | 科 商 |
| 9 | 11 | 15 | 科 學 物 生 |
| 124 | 127 | 141 | 科 學 科 會 社 |
| 40 | 46 | 66 | 科 學 文 國 中 |
| 54 | 68 | 82 | 科 預 |

以上的當然是更少了。
通常應用的三重的統計表，比較上來得複雜。第四表也是一張三重的統計表，但是所包含的似乎比第三表來得複雜，雖然同樣是三重的，包含着科別，是否正式生，性別等三項。

學生科別，
是否正式生
，性別統計
表

| 生 別 特 | | | 生 |
|-------|----|----|----|
| 女 | 男 | 總 | 女 |
| 20 | 79 | 99 | 30 |
| 1 | 15 | 16 | 2 |
| | 16 | 16 | 1 |
| | 15 | 15 | 2 |
| 1 | 3 | 4 | 2 |
| 3 | 11 | 14 | 3 |
| 9 | 11 | 20 | 6 |
| 6 | 8 | 14 | 14 |

統計表的通常形式，大概有上面所製的四種。他們所講到的，都是些次數分配。(Frequency distribution) 此外還有幾種高深的統計表，因其有時也很有用處，附例如下：

(一)「以上」式統計表("More than" form) 這種表可以考查在某某一個指定組以上的總數，第五表便是一個例子。

第五表

復旦大學

| 類 別 | 級 別 |
|-------|-----|
| 生 究 研 | |
| 級 年 四 | |
| 級 年 三 | |
| 級 年 二 | |
| 級 年 一 | |
| 科 預 | |

統一正十
計以式六
表上生年
「級秋
式次季

| | |
|--------------|------|
| 該級及以 上之數人 | 該級數人 |
| 8 | 8 |
| 116 | 103 |
| 284 | 168 |
| 492 | 203 |
| 764 | 272 |
| 832 | 63 |

(二) 「不及」式統計表 ("Less than" form) 這種表的用意

和編製，都和「以上」式的剛剛相反。這種表所查考的，是在某某一個指定級以下的總數，第六表便是一個例子。

第六表

復旦大學
十六年秋季
正式生級次
「不及」式
統計表

| | | |
|--------------|------|----------|
| 不及該級 之學生數 | 該級人數 | 類別 級別 |
| 0 | 68 | 科預 |
| 68 | 272 | 級年一 |
| 340 | 203 | 級年二 |
| 548 | 168 | 級年三 |
| 716 | 103 | 級年四 |
| 824 | 8 | 生究研 |

根據第五表和第六表，一個復旦大學的二年級生便可以知道，在同一校裏面有四百九十二人，至少和他自己在內的

，大概要比他還高一些；同時班次比他低的，也還有三百四十人。

此外，有為研究目標而集合的統計表，編製上較為複雜多多，現在也不預備舉例了。

特別是初學統計的人們以及未習統計的，往往以為統計表是很容易編製的。其實，這是錯誤，製表也不是一件很容易的事情。製表也有幾條公共遵守的規則，現在我們且分別敘述在下面：

- (一) 表的號數和標題，必須寫在表的上面。(用圖表時候，情形不同，應該寫在空格較多的地方或者圖的下面)
- (二) 標題必須很詳盡，能由之無須另製說明更好。否則，對於標題中的重要文字，無論如何，不得遺漏。

(三) 說明項 (如級別、科別、性別等等) 必須寫在表的上面。

(四) 數碼字以阿拉伯字爲宜；萬一必須用中國字的時候，千百十等字務必省去。

(五) 爲便利和阿拉伯字碼同時閱讀起見，說明文字務必自左至右。說明項，非有不得已的緣故，以寫在表的左邊爲宜。

(六) 讀表時宜自上而下，自左而右。

(七) 小項目 (Sub-heading) 應該放在大項目 (Main heading) 底下，並且以向內稍退爲宜。

(八) 項目和事實中間，較重要的時候，應該用實線或虛線 (就是點線) 畫出來。

(九) 重要的線應該畫雙行的，或者加料粗大，務須所
要顯著地表示的東西引起相當的注意。

(十) 倘若表是很長的話，就是有二十組以上的話，則
每隔五行，應留空行一行或半行。

(十一) 表中的字，切勿寫得太小，以致使閱者感覺眼
倦。

(十二) 倘若事實中斷，或者資料缺少了一部份的，應
該在原有地位加畫虛線。

(十三) 總數應該劃一條線，或者留一空地，以引起注
目。

(十四) 同位的數字，必須排列在一條線上，不要互相
參差。

(十五) 編製的時候，應該根據數目的次序規定排列的
前後。

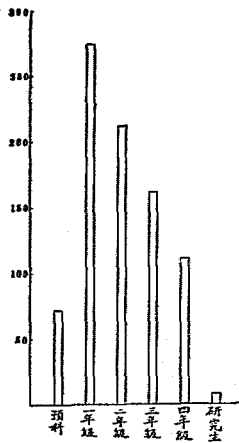
(十六) 表和說明附註等，以相接近為宜，不宜離遠。
以上所說的十六條，大體係根據麥若而教授(W. A. McCull) 所著的如何在教育上測驗一書，都極切實可行。

第五章 統計圖的種類和作法

統計圖大概可以分成二大類：第一類是通常的，第二類是技術的。我們先來談談通常的。通常的統計圖大都是切於實用，約略地可以分做四種；就是：直線統計圖(Bardagram)，曲線統計圖(Curve diagram)，扇形統計圖(Circular diagram)，和長方形統計圖。

第一圖便是直線統計圖，其所能表示的事實，大都是極

第一圖
復旦大
學十六年
秋季各級
正式生人
數表



簡單的。該項直線，可以空心的（如第一圖），也可以實心的（假設我們把第一圖中的空心填實，便成一張實心的了）

曲線統計圖，便是將各該直線的上端連接起來，同時並

且不把直線畫出來就成了。當我們比較許多時候的同一現象的時候，曲線統計表可以說是很適合的。有些曲線統計圖，所以也叫做歷史圖。

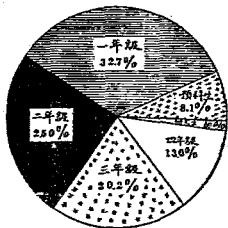
直線統計圖中的直線，可以酌量事實上的需要，畫成豎立的或橫列的，不必拘定任何一式。

直線統計圖中的直線，在可能範圍以內，以二十條為限。過多，不僅使閱者不易記憶，甚且使閱者不易認清應注意的地方。

統計上所要知道的是大概趨勢 (General tendency)。倘若畫通常曲線，有時還不大顯著，則須設計使其平滑 (Smooth)。

第二圖即係扇形統計圖。其整個的外形是一個圓圈。當繪製扇形統計圖以前，必須先將所須表示的事物合成

第二圖
復旦大學
十六年
秋季各級
正式生人
數比較表

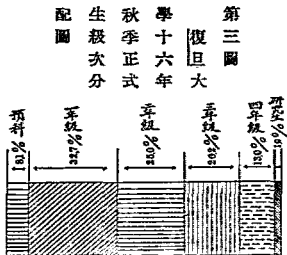


全面積之三成又十分之六，即第二圖中的預科生，佔學生全體之百分之八。一，該在三百六十度中佔二十九度又十分之二。餘類推。

該項辦法，最便通常閱者；但是有許多學者，卻還認為

百分比。但是圖內所表示事物的部份是扇形的面積，該部份所佔據的面積必須等於全面積的百分之幾，該項百分比自當和上項所述的相符合。查一圓周是三百六十度，以百除之，則每百分之一，該佔

不滿意，他們以為在扇形的統計圖中間，二個相差不十分大

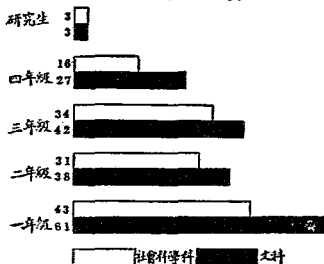


的事物，誰大誰小，一時不易剖明。不過無論如何，扇形統計圖在各處展覽會中間，確實是已經極流行的了。在實際應用在展覽會的時候，各部份通常是用各種顏色指示着的，一來可以醒目，二來又可以省力，真是很方便的辦法。

現在我們所要講的，便是長方形統計圖 (Rectangular diagram)。

第三圖就是長方形統計圖。其長處約有三點，就是：

第四圖
十六年秋
季復旦大
學社會科
學科與文
科學生比
較表



在第四圖中間，該校各該科各級學生的人數和與別科的差異，都表示得很清楚。

一 便於比較。
二 標字方向相同，閱讀時實較扇形統計圖便利多多。
三 可以有較細密的分類。
第四圖是第三圖附圖，也就是長方形統計圖第一項長處的另一個說明。

較通常的四種統計圖，上面都相當地敘述過了；現在我們可以進一步，來談談較高深的，或者較有技術的統計圖。

在這一類統計圖中間，以利用地圖的為最普遍。即就利用地圖而論，還可以分成四種，現在一一地分別敘述在下面

：

一 點圈地圖 以點或小圈來表示地圖上的事實，例如：人口的分配以及各地物產的數量等等。人數或產額較多的地方，該地應畫的點子或小圈也應較多，成為正比例的。

二 直綫地圖 就是以直綫的密度來表示事實的地圖。繪製該項地圖以前，必須先將各區域對於該事物的數量計算清楚，然後根據各該計算的結果，來決定直綫的密度。例如：甲省的人口以面積除之，等於乙省的兩倍，則地圖上同一

面積的直綫，甲省的應該等於乙省的兩倍。也有以顏色的深淺，來表示事物的多寡的，不過事實上，適當的顏色是極不容易配合的；所以現在還是保持直綫的佔多數。

三 平面地圖 該種地圖有二個特質；

(一) 表示某一固定地點所獨有的事實，

(二) 表示牽連於一定基線的事實。

現在我們可以舉一個例，譬如說滬寧鐵路。鐵路的搭客和所運輸的貨物都是獨有的事實，當我們把這些事實畫在平面地圖的時候，我們先把滬寧路截成幾段，並且把搭客和運貨合成車費和運費，然後在中間劃一條鐵路的直綫，左面表示某一段的車費，右面表示某一段的運費。各段間的營業狀態，便可以一覽無遺了。

四 形象地圖 就是根據統計資料，將一定的事實，在地圖上塑成凹凸的狀態。在展覽會場上，這種地圖倒隨時可見，通常是用得很少很少。

其餘用立體或其他方式表現事實的統計圖還多得很，這裏也不預備多述了。倘若要參考，請參閱拙著社會調查之原理及方法第十七第十八兩章。

最後，作圖也有幾條原則。現在我們把美國機械工程學會在一九一九年所規定的十七條，照譯如下：

- 一 圖的排列，應自左至右。
- 二 最好用線來代表數量，因面積和體積均易引起誤解。
- 三 曲線統計圖上的零度橫坐標，最好將其另行畫出。

四 倘零度橫座標不能另行畫出，則該橫坐標和其他的橫線中間，應作波紋的線使其隔離。

五 零度橫坐標應較通常線稍為粗大，使人較易識別。

六 倘若曲線圖以百分當做標準的，則所有的百分線應較通常線稍為寬大，藉以表示區別。其他以比較為目的的各線，也應該比通常的稍為寬大。

七 倘若圖是表示年月日期的，那麼兩旁的界線不宜於粗大，表示時間的始終，是不能加以限制的。

八 倘若曲線圖是畫在對數的 (logarithmic) 格紙上面，那麼縱橫兩條界線應該各畫在對數級上十數位的場所。

九 縱橫線除必要的以外，不應該太多。

十 圖上的曲線應該和其他的線不同，所以表示區別。

十一 倘若採用曲線代表各種事實的觀察，在可能範圍以內，應該於曲線上表明這一類觀察的各點。

十二 圖上量數的讀法，應該自左而右，自下面上。

十三 量表上的數字，應該放在縱坐標的左面，橫坐標的下面，或者縱橫軸的上面。

十四 圖上有時應該載明所代表的數目或者方程式。

十五 倘若數目不能表出，則可另列一表表出。

十六 凡標字和數目字，應該放在圖的下面或者左面。

十七 圖的題目，應該詳備明晰。必要的時候，不妨多加說明。

上文的縱坐標是英文 *Ordinate* 一字的翻譯，橫坐標是英文 *Abscissa* 一字的翻譯；合併附註，以便讀者來日閱讀英文

的統計學書籍。

第六章 平均數

平均數的本質大概可以分成下列六端：

一 平均數應有切實的數量，觀察者的主觀推測是不適用的。

二 平均數應該根據事實的全部觀察，否則難免受到片面事實的影響，而成為不足以代表全體的一部份。

三 平均數應該是一個簡明易曉的數目，並不具有數學上的抽象性質的。

四 平均數應該是計算很方便的。倘若有二種或二種以上的平均數，都是合乎事理的；那麼，我們應該選擇其中計算最方便的一個。但是我們並不是只根據計算的方便與否，

最重要的還是是否合乎事理這一點。

五 平均數應該是固定的數目，帶有普遍性的。倘若把同一組的事實分做若干部份，每部份都摘取幾個平均數；那麼，因為取樣的不調，各部份的同一種平均數，是決不能完全相同的了。但是那一種平均數的差別最少，那一種便是最固定，最合用的了。

六 平均數應該可以用代數的方法統算。

現在我們已經將平均數的本質敘述過了，可以進一步來敘述平均數的種類。統計學上的平均數共有五種：

- 一 算術平均數（就是通常所應用的平均數）
- 二 中數 (Median)
- 三 衆數 (Mode)

四 倒數平均數

五 幾何平均數 (或稱對數平均數)

算術平均數，就是先將各個的數目相加起來，得一總數，然後以次數除之。計算算術平均數的方法有五，現在分述如下：

(1) 量數未經分類的。

假定 M 是算術平均數

N 是次數的總數

X 是量數

Σ 是總和的記號

如此，我們可以有一個公式，就是：

$$M = \frac{1}{N} \Sigma (X) \dots \dots \dots \text{公式 1}$$

第七表 某大學三十個三年級生的已讀學分

| 學號 生數 | 已學 績分 | 學數 生數 | 已學 績分 |
|----------|----------|----------|----------|
| 1 | 108 | 16 | 94 |
| 2 | 106 | 17 | 93 |
| 3 | 102 | 18 | 93 |
| 4 | 99 | 19 | 92 |
| 5 | 98 | 20 | 91 |
| 6 | 98 | 21 | 91 |
| 7 | 97 | 22 | 91 |
| 8 | 96 | 23 | 91 |
| 9 | 96 | 24 | 90 |
| 10 | 96 | 25 | 89 |
| 11 | 96 | 26 | 88 |
| 12 | 95 | 27 | 88 |
| 13 | 95 | 28 | 88 |
| 14 | 94 | 29 | 87 |
| 15 | 94 | 30 | 87 |
| | | | 2823 |

$$N = 30, \quad \Sigma(X) = 2823$$

$$\therefore M = 2823 \div 30 = 94.1$$

(二) 量數已經分類，但組距是一個單位的。
我們也有一個公式，就是：

$$M = \frac{\sum_{i=1}^n f_i(X_i) \dots \dots \dots}{\sum_{i=1}^n f_i} \quad \text{公式二}$$

這公式中的 f 是表示每組的次數，其餘都和以前相同。

第八表 名稱與第七表同

| 已學 讀分 | 次 數 | 已分數 讀乘 舉次 |
|----------|--------|-----------------|
| 108 | 1 | 108 |
| 106 | 1 | 106 |
| 102 | 1 | 102 |
| 99 | 1 | 99 |
| 98 | 2 | 196 |
| 97 | 1 | 97 |
| 96 | 4 | 384 |
| 95 | 2 | 190 |
| 94 | 3 | 282 |
| 93 | 2 | 186 |
| 92 | 1 | 92 |
| 91 | 4 | 364 |
| 90 | 1 | 90 |
| 89 | 1 | 89 |
| 88 | 3 | 264 |
| 87 | 2 | 174 |
| | 30 | 2823 |

$$\Sigma (fX) = 2823$$

$$\therefore M = \frac{2823}{30} = 94.1$$

(三) 量數已經分類，但是組距是不僅一個單位的。這一個方法先有一個假定，就是每組內的量數，分配得很均勻，其平均數必定和該組的中點數相同。第九表便是一個說明。

(四) 簡捷法(A) —— 量數未經分類的。
倘若量數是很多的，則可用簡捷法。其公式如下：

| 組距 | 組距中點 | 次數 | 次數×組距中點 |
|-------|------|-----|---------|
| 不及5元 | 2.5 | 10 | 25.0 |
| 5—10 | 7.5 | 22 | 165.0 |
| 10—15 | 12.5 | 25 | 312.5 |
| 15—20 | 17.5 | 42 | 735.0 |
| 20—25 | 22.5 | 85 | 1912.5 |
| 25—30 | 27.5 | 35 | 962.5 |
| 30—35 | 32.5 | 21 | 682.5 |
| 35—40 | 37.5 | 7 | 262.5 |
| 40—45 | 42.5 | 5 | 212.5 |
| 45—50 | 47.5 | 4 | 190.0 |
| 50—55 | 52.5 | 2 | 105.0 |
| 55—60 | 57.5 | 2 | 115.0 |
| | | 260 | 5680.0 |

第九表

某大公司中下級職員的每月薪金表

$$\therefore M = \frac{5680}{260} = 21.846 +$$

$$M = A + \frac{1}{N} \sum (fa) \dots \dots \dots \text{公式三}$$

在這個公式中間，各記號所代表的如下：

M 是真正算術平均數 A 是假定的算術平均數

N 是次數的總數 \sum 是總和的記號

f 是次數 e 是置數和假定平均數的

相差。

$\frac{1}{N} \sum (fa)$ 是校正的數目，有時也寫做 C。

這種簡捷計算法（見第十表）和前面所講的方法有二點

不同：

- A. 先假定了一個平均數；
- B. 所計算的是真正平均數和假定平均數的相差，就是

C B A 學 計 統

| 差 數 + | 已 得 學 分 | 假 平 數 定 均 | 差 - |
|----------|---------|--------------|--------|
| 14 | 94 | 94 | 0 |
| 12 | 93 | | -1 |
| 8 | 93 | | -1 |
| 5 | 92 | | -2 |
| 4 | 91 | | -3 |
| 4 | 91 | | -3 |
| 3 | 91 | | -3 |
| 2 | 91 | | -3 |
| 2 | 90 | | -4 |
| 2 | 89 | | -5 |
| 2 | 88 | | -6 |
| 1 | 88 | | -6 |
| 1 | 88 | | -6 |
| 0 | 87 | | -7 |
| 0 | 87 | -7 | |
| 60 | | | -57 |

正數，而不是真正平均數。
第十表 資料同第七表

$$C = \frac{60 - 57}{30} = \frac{3}{30} = 0.1$$

$$\therefore M = A + C = 94 + 0.1 = 94.1$$

| 已得率分 | 假平數 定均 |
|------|--------------------------------|
| 108 | (Assumed Average) 94 |
| 106 | |
| 102 | |
| 99 | |
| 98 | |
| 98 | |
| 97 | |
| 96 | |
| 96 | |
| 96 | |
| 96 | |
| 95 | |
| 95 | |
| 94 | |
| 94 | |

在通常應用的時候，非不得已，項目切勿排成二行或二行以上。差數之下，應列兩行，一為正號。（即十），一為負號（即一）；第十表因次序已經排列整齊，所以是例外，各列其一已夠。

(五) 簡捷法 (B) —— 量數已分類且相距在二個單位以上的。

第十一表 資料同第九表

G B A 學 計 統

一 算術平均數的優劣點 算術平均數的優點有四：
 其計算的方法，請將第十一表細加觀察，即得。
 算術平均數所根據的事實，是全部的，可不致陷於

| 組 距 | 次 數 | 差 數 | 次×差 |
|-------|-----|-----|------|
| 不及5元 | 10 | -4 | -40 |
| 5-10 | 22 | -3 | -66 |
| 10-15 | 25 | -2 | -50 |
| 15-20 | 42 | -1 | -42 |
| 20-25 | 85 | 0 | -198 |
| 25-30 | 35 | 1 | 35 |
| 30-35 | 21 | 2 | 42 |
| 35-40 | 7 | 3 | 21 |
| 40-45 | 5 | 4 | 20 |
| 45-50 | 4 | 5 | 20 |
| 50-55 | 2 | 6 | 12 |
| 55-60 | 2 | 7 | 14 |
| | 260 | | 164 |

$$164 - 198 = -34,$$

$$-34 \div 260 = -0.13,$$

$$-0.13 \times 5 = -0.65,$$

$$\text{假定平均數} = 22.5$$

$$\therefore M = 22.5 - 0.65 = 21.85$$

部份的錯誤。

二 算術平均數是一個客觀的數量。

三 算術平均數的計算法是很容易，稍有算術智識者即能爲之。

四 算術平均數是很簡明的。

同時算術平均數也有四個劣點，或者說是四個缺點；就是：

一 算術平均數，非經過計算手續，一時不能決定。

二 算術平均數常常受到極端數量的影響。如計算全廠工資的時候，假將廠長的薪金也計算在內，便使全體數量大受獎勵。

三 如將極端數量截去，則算術平均數即陷於不正確。

四 算術平均數往往並非事實上所表示的。例如，第十表上的平均工資是二十一元八角五分；但是事實上，或僅只有二十一元五角和二十二元的，竟沒有一個職員的工資是二十一元八角五分的。

中數便是中間的一個數量。所以必須經過二步手續：一，把量數依大小排列起來；二，從一端數起，數到行次的一半；才可求得中數。

中數的位置有一個公式，就是：

$$\text{Median 的位置} = \frac{N+1}{2} \dots \dots \dots \text{公式四}$$

在這裏，Median 是代表中數，N 是次數的總數。設有下列二十五個數量，欲求其中數，其程序如下：

| | |
|----|-----|
| 1 | 108 |
| 2 | 106 |
| 3 | 102 |
| 4 | 99 |
| 5 | 98 |
| 6 | 98 |
| 7 | 97 |
| 8 | 96 |
| 9 | 96 |
| 10 | 96 |
| 11 | 96 |
| 12 | 95 |
| 13 | 95 |
| 14 | 94 |
| 15 | 94 |
| 16 | 94 |
| 17 | 93 |
| 18 | 93 |
| 19 | 92 |
| 20 | 91 |
| 21 | 91 |
| 22 | 91 |
| 23 | 91 |
| 24 | 90 |
| 25 | 89 |

此項排列，或由小而大，或由大而小，並不拘定。
就上例而講：

$$\text{Median 的位置} = \frac{N+1}{2} = \frac{25+1}{2} = 13$$

則第十三列，即係該數量的中數。其中數是九十五。

其數量是偶數的，則就中間的二個平均之。例如：第七表共有三十個數量，其中數則應為第十五和第十六兩數量的平均數；即

$$(94+94) \div 2 = 94$$

倘若有許多次數的，如第八表，則應將其次數加併起來

。在三十個數量中間，其中間的兩個應為第十五個和第十六個，我們便要去查考第十五和第十六的位置，（請參照第八表）

1 1 1 1 2 1 4 2
+ + + + + + +
————— 13
則第十五和第十六二數，均在 94
一列內，中數當為 94。

計算中數的方法，上面業已相當敘述，現在再來敘述一些中數的優劣點。中數的優點有：

- 一 是一個客觀的數量。
 - 二 所根據的事實是全部的。
 - 三 計算是很方便，很容易的。
 - 四 是固定的，可不受取樣的變動。
- 中數的缺點是：

一 不能用代數法計算的。

二 計算法不及算術平均數這般普通。

三 知道了中數，不能因此得到總數。

衆數，或作密集數，是許多數量中次數最多的一種。在第九表中，次數最多的是八十五，那麼衆數必在二十元與二十五元之間。二十元至二十五元的中點數是二十二元五角，那麼，牠們的衆數是二十二元五角。

在理論上，我們有一個求衆數的公式，就是：

$$Mo = M - 3(M - Mdn) \dots \dots \dots \text{公式五}$$

在這個公式中間：Mo是衆數，M是算術平均數，

Mdn是中數。

就第八表的事實來講：算術平均數是52.1，中數是

94. 那麼，衆數應該是 $94.1 - 3(94.1 - 94) = 94.1 - 0.3 = 93.8$ 。

衆數有三個優點，就是：

- 一 極端數量的影響可以免除。
- 二 衆數是最普通的狀態，所根據的是多數的事實。
- 三 通常的衆數，可以一目了然，極爲方便。

衆數也有幾個缺點，其中最顯明的是：

- 一 衆數的理論公式，使計算複雜。
- 二 有時有二個或二個以上的特出處，衆數不易認清。

倒數平均數 Harmonic Mean (簡寫作 H.M.) 是一種救濟算術

平均數所不及的平均數。例如：有甲乙丙學生三人，甲每小時能解通常四則題十個，乙每小時能解八個，丙每小時能解六個，問三人平均每小時能解答算題幾個？一位粗心的計算

員，一定以為：

$$\begin{array}{r} 10 \\ + 8 \\ + 6 \\ \hline 3/24 \end{array} \left(8 \right)$$

公式是 $\frac{10+8+6}{3} = \frac{24}{3} = 8$

可是錯了，這是算術平均數，並不是這裏所需要的。這裏所需要的是倒數平均數。倒數平均數是怎樣呢？我們先來看看牠的公式，然後再來解答上面這個問題。

$$H.M. = \frac{1}{\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + \dots + \frac{1}{X_N}} \quad \text{公式六}$$

其實，我們倘若把這個公式稍為倒動一些，亦何嘗不可。我們可以把牠改作：

$$H.M. = \frac{N}{\frac{N}{X_1} + \frac{N}{X_2} + \frac{N}{X_3} + \dots + \frac{N}{X_N}} \quad \text{公式七}$$

現在我們便可以根據了公式，來解答上面的算題。

$$\frac{\sum \frac{1}{X} = \frac{1}{10} + \frac{1}{8} + \frac{1}{6} = \frac{12+15+20}{120} = \frac{47}{120}}$$

$$N \div \sum \frac{1}{X} = 3 \div \frac{47}{120} = 3 \times \frac{120}{47} = \frac{360}{47} \dots 7.66057 +$$

那麼，他們平均每小時能解答 7.66 算題。

幾何平均數 (Geometric Mean) 又稱對數平均數 (Logarithmic Mean)，能減少較大數目的影響，並且增加較小數目的力量，所以計算物價指數的時候，常採用之。所以幾何平均數通常比算術平均數為小，但是猶大於倒數平均數。

幾何平均數，有一個極通行的公式，就是：

$$G.M. = \sqrt[n]{a_1 \times a_2 \times a_3 \times a_4 \dots \times a_n \dots} \quad \text{公式八}$$

在這個公式中間， n 是表示變量的次數。

$a_1, a_2, a_3, a_4, \dots, a_n$ 都是表示變量的。

倘若用對數表，則計算較易多多，公式亦略有變更，附

錄於下：

$$\text{Log G.M.} = \frac{\text{Log } a_1 + \text{Log } a_2 + \text{Log } a_3 + \dots + \text{Log } a_n}{n} \quad \text{公式(1)}$$

例如：求下列三數的幾何平均數：1, 10, 100。

$$\text{G.M.} = \sqrt[3]{1 \times 10 \times 100} = \sqrt[3]{1000}$$

$$\text{Log G.M.} = \frac{\text{Log} 1 + \text{Log} 10 + \text{Log} 100}{3} = \frac{0 + 1 + 2}{3} = 1$$

$$\therefore \text{G.M.} = 10.$$

各種平均數，上面都已有的敘述；此外尚有統計係數者，亦不可不知。以係數乘總數，當可得其相連帶的數目

。例如：有夫之少婦，每年每千人能生育嬰孩二百八十六人，則該項數目即係係數 (Coefficients)。那麼，該地方若有該項有夫之少婦三萬八千五百人，則每年即能生育嬰孩一萬一千〇十一人。在戶口統計中間，該項係數應用頗廣；惟計算時，對於分子（例如：所生育嬰孩的數目）分母（例如：有夫之少婦的總數）都應該加以充分注意，使不致於會陷於錯誤。

第七章 變量

變量 (Dispersion or Variability) 所欲表示的，就是差異。在一羣事物中間，體量在事實上必定是參差不齊的。同一枝幹上的樹葉，其長度和闊度是各各不相等的，有的祇有三四寸，有的却有七八寸。比較上相差很少的事物，我們就可以說

他是整齊的。整齊的數目是集中於平均數附近的，和平均數是相去不遠的；但是大部份的自然現象，決不會全體集中在一個數量上的。舉個例來說，我們以為娼妓的年齡是集中於二十一歲至二十五歲的。其實，決不是這樣的，有的娼妓，只不過十四五歲，有的却過了四十歲。即使我們不管這些，二十一歲至二十五歲，相去也有五年啊。這種差異的數目，就是叫做離中差 (Deviation)。計算變量的法則，通常有四種，現在分述如下：

一 全距離 (Range) 全距離就是從較大一端的數量中間減去較小一端的數量。例如：娼妓是自十四歲起直到四十一歲為止的，那麼，全距離是二十七年。事實上，全距離有二個困難：A. 常受少數極端數量的影響；例如：在娼妓中間，

年齡在三十三歲以上的只不過極端的少數。倘若我們把三十三歲以上減除了，全距離便縮短為十八年，可是又陷於不真確。B. 內部分配的實情，我們不能由之知道。

二 四分位差 (Quartile deviation or Semi-interquartile Range 簡稱作 Q) 中數是全數量的中間一項，所以中數是先將全數量項數平分的。倘若將這個平分開的兩部份再各求其中數，則全體分為四部，所以這兩部份的中數是叫做四分位數，統計學上的記號是 Q_1 和 Q_3 。這四部是相等的，從量數最小的一端數起；倘若分配是完全對稱 (Symmetric) 的，那麼：

$$\text{Median} - Q_1 = Q_3 - \text{Median} \dots \dots \dots \text{公式十}$$

倘若分配是偏向於一端的，或者完全不對稱的時候，那麼：

$$Q = \frac{Q_2 - Q_1}{2} \dots \dots \dots \text{公式十一}$$

至於 Q_1 和 Q_2 的求法，和求中數法亦略有出入，其公式如下。

$$Q_1 = L + \frac{1}{f} \times 0 \dots \dots \dots \text{公式十二} \quad Q_2 = H + \frac{1}{f} \times 0 \dots \dots \dots \text{公式十三}$$

公式中的 L 是指數量的下關， H 是指數量的上關。

就第九表而言，則

$$Q_1 = 15 + \frac{8}{45} \times 5 = 15.95$$

$$Q_2 = 30 - \frac{11}{35} \times 5 = 28.43$$

$$\text{求中數} \quad Q = \frac{Q_2 - Q_1}{2} = \frac{28.43 - 15.95}{2} = \frac{12.48}{2} = 0.94$$

四分位差的數量，剛剛和統計學上的概差 (Probable error) 相等，所以應用得頗廣。

三 標準差 (Standard Deviation, 簡稱作 S. D.) 標準差通常係以希臘字母 σ (讀作雪格麥) 為符號的。其求法頗為繁複，至少有下列五個步驟，就是：

- A 先求各數量和算術平均數的差數。
- B 使所求得的差數自乘一次。
- C 求各自乘積數的總和。
- D 再以次數的總和除積數的總和。
- E 再求其平方根。

如此，則全數的百分之六十八以上，均可包含在內，實乃真正根據各數量的離中差而測定變量的程度的。

標準差有一個公式，就是：

$$\sigma = \sqrt{\frac{\sum X^2}{N} \dots \dots \dots \text{公式十四}}$$

在這個公式中間， $\sum X^2$ 是各數量和算術平均數的差數，其他，都已經在上面說明過了。

其計算法如下：

第十二表 標準差的求法（資料根據第八表）

| 次 數 | Xf^2 |
|------|--------|
| 1 | 193.21 |
| 1 | 141.61 |
| 1 | 62.41 |
| 1 | 24.01 |
| 2 | 30.42 |
| 1 | 8.41 |
| 4 | 14.44 |
| 2 | 1.62 |
| 3 | 0.03 |
| 2 | 2.42 |
| 1 | 4.41 |
| 4 | 38.44 |
| 1 | 16.81 |
| 1 | 26.01 |
| 3 | 111.63 |
| 2 | 100.82 |
| N=30 | 776.70 |

但是，這並不是唯一的計算法，我們還有較為簡捷的。
 第十三表 標準差的捷求法（資料同上表）

$$\begin{aligned} \therefore s &= \sqrt{\frac{\sum X^2}{30} - \frac{(\sum X)^2}{30^2}} \\ &= \sqrt{\frac{776.70}{30} - \frac{26.89^2}{30}} \end{aligned}$$

| 已讀學分 | 差數 X | 差數方 X ² |
|----------|-------|--------------------|
| 108 | +13.9 | 193.21 |
| 106 | +11.9 | 141.61 |
| 102 | + 7.9 | 62.41 |
| 99 | + 4.9 | 24.01 |
| 98 | + 3.9 | 15.21 |
| 97 | + 2.9 | 8.41 |
| 96 | + 1.9 | 3.61 |
| 95 | + 0.9 | 0.81 |
| 94 | - 0.1 | 0.01 |
| 93 | - 1.1 | 1.21 |
| 92 | - 2.1 | 4.41 |
| 91 | - 3.1 | 9.61 |
| 90 | - 4.1 | 16.81 |
| 89 | - 5.1 | 26.01 |
| 88 | - 6.1 | 37.21 |
| 87 | - 7.1 | 50.41 |
| M = 94.1 | | |

| 已讀學分 | 次 數 | 與設均之 假平數差 | 次×數 數差 | 次×數 數差方 |
|------|-----|--------------|-----------|------------|
| 108 | 1 | +14 | 14 | 193 |
| 106 | 1 | +12 | 12 | 144 |
| 102 | 1 | + 8 | 8 | 64 |
| 99 | 1 | + 5 | 5 | 25 |
| 98 | 2 | + 4 | 8 | 32 |
| 97 | 1 | + 3 | 3 | 9 |
| 96 | 4 | + 2 | 8 | 16 |
| 95 | 2 | + 1 | 2 | 2 |
| 94 | 3 | 0 | +60 | 0 |
| 23 | 2 | - 1 | - 2 | 2 |
| 92 | 1 | - 2 | - 2 | 4 |
| 91 | 4 | - 3 | -12 | 36 |
| 90 | 1 | - 4 | - 4 | 16 |
| 89 | 1 | - 5 | - 5 | 25 |
| 88 | 3 | - 6 | -18 | 108 |
| 87 | 2 | - 7 | -14 | 98 |
| | 30 | | -57 | 777 |

$$0.3) + 5(0.1 \ 3.)777(25.9$$

$$\Rightarrow s = 0.1$$

$$\therefore s^2 = 0.1^2 = 0.01$$

$$\frac{\sum fe^2}{N} = 25.9$$

$$s^2 = 25.9 - 0.01 = 25.89$$

$$\therefore s = \sqrt{25.89} = 5.088$$

標準差的特點，較顯著的有六個，現在分述如下：

A 根據着全部事實；

B 可以用切實數量表示；

C 可以用代數法計算；

D 計算還容易；

E 取樣變動，可以不受影響；

F 正負各一標準差等於全數的百分之六八·二六。

其餘的幾個，則比較上較為次要的了。

四 平均差 (Average Deviation, 簡稱做 A. D.) 平均差也是

變量的一種。其計算的方法，先求各數益和平均數的差別，各差數相加，不計正負符號，得一總和，然後再以次數的總和除之，即得。求平均差的時候，中數和算術平均數，都可

以用做標準。但是就理論上講，中數似較好，因為中數通常

總比算術平均數來得簡單。

$$A. D. = \frac{\sum fX}{N} \dots \dots \dots \text{公式十五}$$

第十四表 平均差的求法

| 已學分 | 次 數 | 差 數 | 次×數 |
|-----|-----|-----|-----|
| 108 | 1 | 14 | 14 |
| 106 | 1 | 12 | 12 |
| 102 | 1 | 8 | 8 |
| 99 | 1 | 5 | 5 |
| 98 | 2 | 4 | 8 |
| 97 | 1 | 3 | 3 |
| 96 | 4 | 2 | 8 |
| 95 | 2 | 1 | 2 |
| 94 | 3 | 0 | 0 |
| 93 | 2 | 1 | 2 |
| 92 | 1 | 2 | 2 |
| 91 | 4 | 3 | 12 |
| 90 | 1 | 4 | 4 |
| 89 | 1 | 5 | 5 |
| 88 | 3 | 6 | 18 |
| 87 | 2 | 7 | 14 |
| | 30 | | 117 |

中數=94

$$\therefore A. D. = \frac{117}{30} = 3.9$$

平均差的特點有四點，就是：

A 取樣變動，亦須受影響；

B 算法比較上來得繁瑣；

C 11 A.D. 全數的 57.5% ；

D 在常態分配圖或稍不對稱的圖上 A.D. = 0.79798 。

數量分配不對稱，叫做偏態。(Skewness) 偏態的情形不一；有的略略偏態一些的，有的是極端偏態的。我們要比較其偏態的大小和性質，有二個公式是可以應用的。

(一) 普爾孫氏 Pearsonian formula

$$\text{偏態性 (Skewness)} = \frac{M_3 - 3M_1M_0}{\sigma^3} \dots\dots\dots \text{公式十六}$$

是但 $M_0 = M - 3(M - Mdn) \dots\dots\dots$ 見公式五

$$\therefore \text{偏態性} = \frac{M - (M - 3)(M - Md_n)}{9}$$

$$= \frac{3(M - Md_n)}{9} \dots \dots \dots \text{公式十七}$$

根據公式十六，可知偏態愈大。那麼算術平均數離衆數便愈遠，而二個數目的差別也愈大。公式十六乃是用算術平均數和衆數相差的大小，而推算偏態性的多少。公式十七乃是用算術平均數和中數相差的大小，而推算偏態性的多少。

(II) 數而式 Yule's formula 這個公式是利用四分位差，推算偏態性的多少。其公式如下：

$$\text{偏態性} = \frac{(Q_3 - Md_n) - (Md_n - Q_1)}{9} \dots \dots \dots \text{公式十八}$$

$$= \frac{Q_3 + Q_1 - 2Md_n}{9} \dots \dots \dots \text{公式十九}$$

舉例如下：在第九表內，

$$Mdn = 21.82,$$

$$Q_1 = 16.96,$$

$$Q_3 = 28.43,$$

$$Q_4 = 6.24.$$

$$\text{價格性} = \frac{(28.43 - 21.82) - (21.82 - 16.96)}{6.24}$$

$$= \frac{6.61 - 4.87}{6.24} = \frac{0.74}{6.24} = 0.118$$

第八章 相互關係

凡表示二種統計事項間的關係，叫做相互關係。二種統計的循環相合的，叫做正相關；相反的，叫做負相關。建築物特許狀的發給數目和生鐵生產量，是有密接的相互關係的。股票價格和物價也很有相關的。夫婦的年齡，顯然也是有關的。

相互關係的求法，大概有三種，就是：

一 均方相關的求法 均方相關法的公式，是聾而孫所得的，並且為求直線性相關法中的最完善的。現在分述如下：

A 通行算法 就是應用聾而孫的求均方相關的公式，其公式如下：

$$r = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}} \quad \text{公式一十}$$

在這個公式中間：

Y 是相互關係的係數。

X 是 X 行中的任何數量和諸 X 數量的平均數的差數。

Y 是 Y 行中的任何數量和諸 Y 數量的平均數的差數。

| Y | X ² | Y ² | +XY | -XY |
|-----|----------------|----------------|-----|------|
| + 9 | 49 | 81 | | -63 |
| - 3 | 1 | 9 | | - 3 |
| 0 | 4 | | | |
| + 6 | 4 | 36 | +12 | |
| - 3 | 1 | 9 | | - 3 |
| - 6 | | 36 | | |
| + 5 | 1 | 25 | + 5 | |
| + 2 | 4 | 4 | | - 4 |
| +10 | 9 | 100 | +30 | |
| + 1 | 49 | 1 | + 7 | |
| - 3 | 9 | 9 | + 9 | |
| - 7 | 16 | 49 | | -28 |
| - 1 | 36 | 1 | + 6 | |
| - 1 | | 1 | | |
| - 9 | 9 | 81 | +27 | |
| | 192 | 442 | 96 | -101 |

十五表便是一個很清楚的例子。

應用上面這個公式，（即公式二十）來求均方相關，第

第十五表 均方相關的通常求法

B 簡捷計算法 在原則上，簡捷計算法和通行計算法並沒有什麼不同；但是公式的排列，當然和前面的稍有不同，其公式如下：

$$y = \frac{98-101}{\sqrt{152} \sqrt{442}} = -\frac{5}{201.81} = -0.0176 \pm P. E.$$

| 人數 | X | Y | X |
|----|----|----|-----|
| 1 | 75 | 87 | - 7 |
| 2 | 83 | 75 | + 1 |
| 3 | 84 | 78 | + 2 |
| 4 | 84 | 84 | + 2 |
| 5 | 83 | 75 | + 1 |
| 6 | 82 | 72 | 0 |
| 7 | 83 | 83 | + 1 |
| 8 | 80 | 80 | - 2 |
| 9 | 85 | 88 | + 3 |
| 10 | 89 | 79 | + 7 |
| 11 | 79 | 75 | - 3 |
| 12 | 86 | 71 | + 4 |
| 13 | 76 | 77 | - 6 |
| 14 | 82 | 77 | 0 |
| 15 | 79 | 96 | - 3 |

M=82 M=78

$$y = \frac{\sum (e_{ij}) - N\delta_x\delta_y}{N \sqrt{\frac{\sum e_{ij}^2}{N} - \delta_x^2} \sqrt{\frac{\sum e_{ij}^2}{N} - \delta_y^2}} \dots \dots \dots \text{公式二十一}$$

有時為便利起見，將上式以 N 除之，亦可應用，公式如下：

$$r = \frac{\frac{\sum (e_{ij})}{N} - \delta_x\delta_y}{\sqrt{\frac{\sum e_{ij}^2}{N} - \delta_x^2} \sqrt{\frac{\sum e_{ij}^2}{N} - \delta_y^2}} \dots \dots \dots \text{公式二十二}$$

公式中的各符號的讀法及意義，茲為說明如下：

e 讀作 Xi，係指 X 行之任何數量和假設平均數的差數。

N 讀作意帶，係指 Y 行之任何數量和假設平均數的差數。

δ 讀作分面帶 δx 係指 X 行假設平均數的校正數。

第十五表) 資料同第 計算法 (關的簡捷 均方相 第十六數

| X | Y | ϵ | n | ϵ^2 | n^2 | $+\epsilon n$ | $-\epsilon n$ |
|----|----|------------|-----|--------------|-------|---------------|---------------|
| 75 | 87 | -5 | +7 | 25 | 49 | | -35 |
| 83 | 75 | +3 | -5 | 9 | 25 | | -15 |
| 84 | 78 | +4 | -2 | 16 | 4 | | -8 |
| 84 | 84 | +4 | +4 | 16 | 16 | 16 | |
| 83 | 75 | +3 | -5 | 9 | 25 | | -15 |
| 82 | 72 | +2 | -8 | 4 | 64 | | -16 |
| 83 | 83 | +3 | +3 | 9 | 9 | 9 | |
| 80 | 80 | 0 | 0 | 0 | 0 | | |
| 85 | 88 | +5 | +8 | 25 | 64 | 40 | |
| 89 | 79 | +9 | -1 | 81 | 1 | | -9 |
| 79 | 75 | -1 | -5 | 1 | 25 | 5 | |
| 86 | 71 | +6 | -9 | 36 | 81 | | -54 |
| 76 | 77 | -4 | -3 | 16 | 9 | 12 | |
| 82 | 77 | +2 | -3 | 4 | 9 | | -6 |
| 79 | 69 | -1 | -11 | 1 | 121 | 11 | |

δy 係指 Y 行假設平均數的校正數。

| | | | | | | | |
|--------------|--------------|-----|-----|-----|-----|----|------|
| $\delta 0 =$ | $\delta 0 =$ | 41 | 22 | 252 | 502 | 93 | -158 |
| 假設 | 假設 | -11 | -52 | | | | |
| 平均 | 平均 | 30 | -30 | | | | |
| 數 | 數 | | | | | | |

$$b_x = \frac{30}{16} = 2$$

$$\therefore b_x^2 = 2^2 = 4,$$

$$b_y = \frac{-30}{16} = -2$$

$$\therefore b_y^2 = (-2)^2 = 4,$$

$$\frac{\sum x^2}{N} = \frac{262}{16} = 16.8, \quad \frac{\sum x^2}{N} = \frac{502}{16} = 33.467, \quad \sum en = 93 \sim 168 = -95.$$

$$\therefore r = \frac{-65 - 16(2)(2)}{16\sqrt{16.8 - 4}\sqrt{33.467 - 4}} = \frac{-125}{16\sqrt{12.8 \times 29.467}} = \frac{-125}{201.316}$$

$$= -0.429 \pm R.E.$$

CBA 考計算

二 等級相關的計算法 學生的分數，往往參加着教師的偏見；所以有些人以為學生成績的次序，比較上更為可靠。且計算亦較方便，所以採用的很多。英人斯畢門(Spearman)對於該項計算法，有一個公式，就是現在通常所應用的。

A 斯畢門等級相關法

斯畢門的公式如下：

$$P = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} \quad \text{公式二十三}$$

在這個公式中間：

P (讀做羅) 係斯畢門等級相關的係數符號。價值較普而孫氏的 γ 略小。

D 係指 X 行的數量和 Y 行的數量在次第上的差數。

N 即係次數的總和。

茲有某級學生十五人，求其對於 X、Y 兩學程的等級相關。

我們便可以根據上面的公式，來計算一下：（見第十七表）

第十七表

斯畢門

等級相關

的求法

| 學生 | 等X級 | 等Y級 | D | D ² |
|----|-----|-----|---|----------------|
| A | 1 | 3 | 2 | 4 |
| B | 2 | 1 | 1 | 1 |
| C | 3 | 4 | 1 | 1 |
| D | 4 | 2 | 2 | 4 |
| E | 5 | 7 | 2 | 4 |
| F | 6 | 6 | 0 | 0 |
| G | 7 | 8 | 1 | 1 |
| H | 8 | 5 | 3 | 9 |
| I | 9 | 9 | 0 | 0 |
| J | 10 | 10 | 0 | 0 |
| K | 11 | 13 | 2 | 4 |
| L | 12 | 11 | 1 | 1 |
| M | 13 | 12 | 1 | 1 |
| N | 14 | 15 | 1 | 1 |
| O | 15 | 14 | 1 | 1 |

N=15 ΣD²=32

$$P = 1 - \frac{0.5 \sum D^2}{N(N^2 - 1)} = 1 - \frac{0.5 \times 32}{15(15^2 - 1)} = 1 - \frac{16}{3360} = 0.943 \quad Y = 0.948$$

Y 的來處，似覺滑稽，然較完備的統計學書籍，大都均附有許多備檢查的表格，從 P 查 Y，或從 Y 查 P，都屬有之。

B 等級相關的簡捷計算法

該項簡捷計算法的公式如下：

$$R = I + \frac{G}{N-1} \quad \text{公式二十四}$$

在這個公式中間：

R 係斯畢門捷算法的相關係數的特殊符號，其價值較 Y 稍大。

G 為英語 Gains 的縮寫，係指 Y 數量的次序超過 X 數量的次序的等級。反之，X 數量的超過 Y 數量的，亦可。

第十八表
等級相
關的簡捷

| 等級 Y | G (Y 大於 X) |
|------|---------------|
| 3 | 2 |
| 1 | 1 |
| 4 | 1 |
| 2 | 2 |
| 7 | 2 |
| 6 | |
| 8 | 1 |
| 5 | |
| 9 | |
| 16 | |
| 13 | 2 |
| 11 | |
| 12 | |
| 15 | 1 |
| 14 | |

$\Sigma G = 9$

計算法

| 學生 | 等級 X |
|----|------|
| A | 1 |
| B | 2 |
| C | 3 |
| D | 4 |
| E | 5 |
| F | 6 |
| G | 7 |
| H | 8 |
| I | 9 |
| J | 10 |
| K | 11 |
| L | 12 |
| M | 13 |
| N | 14 |
| O | 15 |

N=16

$$R = 1 - \frac{6 \sum G}{N^2 - 1} = 1 - \frac{6 \times 9}{256 - 1} = 0.759$$

若 $R = 0.759$, 則 $\gamma = 0.936$

至於 R 求 γ 的方法，另有表可查；本書限於篇幅，未及錄入。

三 異號相關的計算法 異號相關的計算法，是最敏捷的一個方法。所可惜者，所求得的數目，不甚精確耳。該項

計算法的公式，為英人薛伯遜 (Sheppard) 的求得，其公式如下：

$$D = \frac{u + \left(\frac{\frac{u}{2} + \frac{l}{2} + 1}{p} \right)}{N} \dots\dots\dots \text{公式二十五}$$

在這個公式中間：

U 係薛伯遜異號相關的係數符號，其價值和 Y 的價值剛相反。

I 係每雙相對數量的差數符號相同的，如十和一一。

n 係符號相異的，如十，一和一，十。

d 係 ○，○；十，○；○，十；一，○；和 ○，一；五

種情形的符號，即有 ○ 者。

第十九表 薛伯遜異號相關的計算法

| 數人 | X | Y | X | y |
|----|----|----|---|---|
| A | 75 | 87 | - | + |
| B | 83 | 75 | + | - |
| C | 84 | 78 | + | 0 |
| D | 84 | 84 | + | + |
| E | 83 | 75 | + | - |
| F | 82 | 72 | 0 | - |
| G | 83 | 83 | + | + |
| H | 80 | 80 | - | + |
| I | 85 | 88 | + | + |
| J | 89 | 79 | + | + |
| K | 79 | 75 | - | - |
| L | 86 | 71 | + | - |
| M | 76 | 77 | - | - |
| N | 82 | 77 | 0 | - |
| O | 79 | 69 | - | - |

$$\begin{aligned}
 & \frac{u}{N} + \left(\frac{\frac{u}{n+1} + 1}{2} \right) d \\
 & \frac{b}{N} + \left(\frac{\frac{b}{5+7+1} + 1}{2} \right) s \\
 & \frac{b}{N} + \left(\frac{b+6}{24} \right) s
 \end{aligned}$$

$$\begin{aligned}
 M &= 82 & M &= 78 \\
 N &= 15 \\
 I &= 7 \\
 n &= \dots \\
 d &= 3
 \end{aligned}$$

$$\frac{5+1.375}{15} = \frac{6.375}{15} = 0.425$$

$$\text{答 } U = 0.425,$$

$$JII \quad r = 0.3833$$

相關尚有自身相關等數種，因其比較上不甚重要，概從略。

第九章 統計書報述要

一 統計學書籍

A Secrist, II., *An Introduction to Statistical Methods* 該書對於統計學上的重要各點，都有相當的敘述；內容尚稱齊備，頗合初學者及高級中學教科之用。該書初版於一九一七年，改訂本出版於一九二七年，內容修改頗多，篇幅亦較前增加；充作高中教本，倘教授支配得法，亦不致感覺資料過多。

B King, W. L., *Elements of Statistical Method* 該書以簡單明瞭見長，亦係一統計學的通俗教本。惟該書出版於一九一二年，迄今未見修改，不得不認為遺憾。

C Elderton, W. P. and E. M., *Primer of Statistics* 該書對於初步的統計理論，解說至為明瞭。頗合初學時參考之用。

D Bowley, A. L., *Elements of Statistics* 此書是統計學界的名著。計分二冊，下冊專從數學方面討論。亦有合併成一冊者，為數似較少；初學者祇讀上冊一冊已足。該書的第四版，於一九二〇年出版。

E Jerome, H., *Statistical Method* 本書著者，是惠斯康辛大學的助教授，對於統計學造詣極深。本書共分五部；第一部，導言；第二部，圖表編製法；第三部，分析；第四部，統

計的來源和集合；第五部，附錄。其中，第三部敘述最詳，幾佔全書二分之一。附錄多至七十頁，真可謂應有盡有。探作專門以上學校之教本，當能適合。

F Kelley, F. L., *Statistical Method* 本書共分十三章。先從製表着手，旋論製圖，及集中趨向和變量的測定。其論相關至五章之多，可謂得未曾有。綜觀全書，實亦一專門以上學校的教本。

G Davenport, C. B., *Statistical Methods* 本書皮面金脊，頗為美觀。著者係一優生學者，所以本書對於生物變量，加以特殊的注意。附有計算多種，頗切實用。

H Rugg, H. O., *Statistical Methods Applied to Education* 本書為教育統計學中的通行本。凡研究教育及辦理教育行政的人，

倘若沒有學過統計學的話，很可以看一下子。

I Davies, G. R., Introduction to Economic Statistics 本書為大學商科及經濟科學生而寫，所以內容偏重於經濟現象。有相當根底的人，讀之當亦不嫌艱難。

其他統計書籍，所在尚多，這裏為篇幅所限，恕不備舉了。

二 統計學定期刊物

A 白拍孫氏統計冊 (Babson's Statistical Tables and Charts) 所編指數，極饒價值；西洋大商號，往往購買一份，以備經理等的參考。簡稱白拍孫。(Babson's) 是一種月刊。

B 經濟統計評論 (Review of Economic Statistics) 為哈佛大學經濟研究委員會 (Harvard University Committee on Economic Research)

所編製，是季刊性質。此外，該會尚編製週刊一種。二者總稱，則為哈佛經濟服務。(Harvard Economic Service) 該兩項刊物，均極受經濟學者之贊許，所以聲譽極好。

C 卜老克瑪亞經濟服務社 (Brookline Economic Service) 刊行統計定期刊至五種之多。內有月刊，兩月刊各二，週刊一種。這五種刊物，銷路都很不錯，頗為商人所信仰。

D 財政狀況評論週刊 (Weekly Review of Financial Conditions) 為Moody's Investor's Service 所發行。銷路亦頗廣。

此外，尚有統計定期刊物極多，惟大都祇行銷於一地或一業，概從略。年鑑及政府刊物，大都含有資料不多。學統計的，對於這兩方面，也應該加以注意。

ABC叢書目錄

文藝部

國學組

文字學ABC一册 平裝五角
精裝六角

持志大學國學主任胡撰安著

修辭學ABC一册 平裝五角
精裝六角

復旦大學中國文學系主任陳望道著

詩詞學ABC一册 平裝五角
精裝六角

持志大學文辭教授胡撰安著

詩經學ABC一册 平裝五角
精裝六角

南開大學教授金公亮著

元劇研究ABC上下册 平裝五角
精裝六角

元曲專家吳撰安著

文法解剖ABC一册 平裝五角
精裝六角

郭步陶著

文學組

文藝論ABC一册 平裝五角
精裝六角

前暨南大學文科主任夏徵尊著

文藝批評ABC一冊

平裝五角
精裝六角

復旦大學教授傅東藩著

文化評價ABC一冊

平裝五角
精裝六角

巴黎大學碩士葉法無著

詩歌原理ABC一冊

平裝五角
精裝六角

復旦大學教授傅東藩著

小說研究ABC一冊

平裝五角
精裝六角

文學批評家玄珠著

農氏文學ABC一冊

平裝五角
精裝六角

復旦大學教授謝六逸著

兒童文學ABC一冊

平裝五角
精裝六角

小臥月羅編譯徐潤孚著

報

學ABC一冊

平裝五角
精裝六角

時事新報編輯蔣公毅著

演說

學ABC一冊

平裝五角
精裝六角

復旦文科主任余繩秋著

西洋文學組

希羅文學ABC一冊

平裝五角
精裝六角

附歐文學專家曾孟侯著

英國文學ABC

上冊
平裝五角
精裝六角

真美華雜誌編輯曾虛白著

美國文學ABC一冊

平裝五角
精裝六角

真美華雜誌編輯曾虛白著

德國文學 A B C 一册 平裝五角
精裝六角

大學院秘書李金髮黃貞奇著

法國文學 A B C 一册 平裝五角
精裝六角

南歐文學專家曾孟侯著

俄國文學 A B C 一册 平裝五角
精裝六角

文學士吳謙中學教授汪儷然著

近代文學 A B C 一册 平裝五角
精裝六角

文學家吳雲著

騎士文學 A B C 一册 平裝五角
精裝六角

文學批評家玄珠著

神話組

神話學 A B C 一册 平裝五角
精裝六角

復旦大學教授謝六逸著

童話學 A B C 一册 平裝五角
精裝六角

趙景深著

中國神話研究 A B C 下册 平裝五角
精裝六角

文學批評家玄珠著

希臘神話 A B C 一册 平裝五角
精裝六角

文學士吳謙中學教授汪儷然著

藝術組

藝術論 A B C 一册 平裝五角
精裝六角

復旦大學中國文學系主任陳望道著

戲劇

培 A B C 一册
胡美戲劇家洪深著

平裝五角
精裝六角

獨劇

幕 A B C 一册
文學士麥基譯著

平裝五角
精裝六角

歌

劇 A B C 一册
音樂家張若谷著

平裝五角
精裝六角

音

樂 A B C 一册
音樂家張若谷著

平裝五角
精裝六角

國

畫 A B C 一册
名畫家朱應鵬著

平裝五角
精裝六角

洋

畫 A B C 一册
名洋畫家做過一著

平裝五角
精裝六角

圖案

畫 A B C 一册
中華藝術大學教授陳之佛著

平裝五角
精裝六角

構圖

法 A B C 一册
上海美術專門學校教授羅子位著

平裝五角
精裝六角

政治經濟部

政治組

黨

義 A B C 一册
世界書局編輯朱明新著

平裝五角
精裝六角

政治

學 A B C 一册
前政治大學校長張君勱著

平裝五角
精裝六角

國際政治 A B C 一册 平裝五角
精裝六角

巴黎大學吳頌皋著

外 交 A B C 一册 平裝五角
精裝六角

浙江反省院教授常春林著

都市論 A B C 一册 平裝五角
精裝六角

工學士楊哲明著

市政管理 A B C 一册 平裝五角
精裝六角

工學士楊哲明著

市政計畫 A B C 一册 平裝五角
精裝六角

工學士楊哲明著

市政工程 A B C 一册 平裝五角
精裝六角

工學士楊哲明著

道路學 A B C 一册 平裝五角
精裝六角

工學士楊哲明著

經濟組

經濟學 A B C 一册 平裝五角
精裝六角

經濟學博士李權時著

財政學 A B C 一册 平裝五角
精裝六角

經濟學博士李權時著

貨幣學 A B C 一册 平裝五角
精裝六角

德國柏林大學沈漢理著

統計學 A B C 一册 平裝五角
精裝六角

社會學博士蔡鏡題著

分 配

論 A B C 一册 平裝五角
精裝六角

法政大學教授嚴善光著

信用合作 A B C 一册

平裝五角
精裝六角

侯厚培著

農業合作 A B C 一册

平裝五角
精裝六角

法政大學教授王世穎著

生活進化史 A B C 一册

平裝五角
精裝六角

勞働大學教授劉叔罕著

商 業 組

銀 行

學 A B C 一册 平裝五角
精裝六角

復旦大學商學士顧世勳著

廣 告

學 A B C 一册 平裝五角
精裝六角

復旦大學商學士顧世勳著

商業經營 A B C 一册

平裝五角
精裝六角

明治大學商學士沈長明著

工商管理 A B C 一册

平裝五角
精裝六角

復旦大學商學士張家泰著

售 貨 術 A B C 一册

平裝五角
精裝六角

復旦大學商學士張家泰著

社 會 組

社 會

學 A B C 一册 平裝五角
精裝六角

社會學博士孫本文著

社會思想史 A B C 一册

平裝五角
續裝六角

法政大學教授檢選稿著

人口

論 A B C 一册

平裝五角
續裝六角

社會學博士塚本文著

人類

學 A B C 一册

平裝五角
續裝六角

巴黎大學學士尼密蘭著

優生

學 A B C 一册

平裝五角
續裝六角

日本帝國大學學士華汝成著

犯罪

學 A B C 一册

平裝五角
續裝六角

留美社會學碩士鹿成一著

婦女運動 A B C 一册

平裝五角
續裝六角

湯彬華女士著

哲學部

哲學組

哲

學 A B C 一册

平裝五角
續裝六角

前時本新編主筆張東蓀著

西洋

哲學 A B C 一册

平裝五角
續裝六角

廣學會編譯謝頌燕著

宗教

學 A B C 一册

平裝五角
續裝六角

廣學會編譯謝頌燕著

人生

觀 A B C 一册

平裝五角
續裝六角

前時本新編主筆張東蓀著

論理

學A B C 一册

平裝五角
精裝六角

前紹興女師校長朱兆琛著

倫理問題

A B C 一册

平裝五角
精裝六角

巴黎大學碩士葉法無著

中倫理想

A B C 一册

平裝五角
精裝六角

嶺南大學哲學教授謝扶雅著

教育史地部

教育組

教育學

A B C 一册

平裝五角
精裝六角

米西根大學碩士黃廷就編著

教育史

A B C 一册

平裝五角
精裝六角

李滄音

藝術教育

A B C 一册

平裝五角
精裝六角

美術專門學校教授野子恒著

職業教育

A B C 一册

平裝五角
精裝六角

中華職業教育社澤文安著

小學行政

A B C 一册

平裝五角
精裝六角

世界書局編輯魏汝心著

各科教學

A B C 一册

平裝五角
精裝六角

世界書局編輯所長溫雲六著

教育測驗

A B C 一册

平裝五角
精裝六角

世界書局編輯朱明新著

教育心理學 A B C 一册

平裝五角
精裝六角

前紹興女師校長朱兆翠著

圖書館學 A B C 一册

平裝五角
精裝六角

復旦大學圖書館主任沈學植著

田徑賽 A B C 一册

平裝五角
精裝六角

時事新報編輯部著

辯論術 A B C 一册

平裝五角
精裝六角

世界書局編輯部著

史地組

東洋史 A B C 一册

平裝五角
精裝六角

文學家傅彥長著

西洋史 A B C 一册

平裝五角
精裝六角

文學家傅彥長著

日本史 A B C 一册

平裝五角
精裝六角

北平女高師教授李宗武著

人文地理 A B C 一册

平裝五角
精裝六角

北平女高師教授李宗武著

自然地理 A B C 一册

平裝五角
精裝六角

雅安中學校長王益厚著

海洋學 A B C 一册

平裝五角
精裝六角

雅安中學校長王益厚著

科學部

自然科學組

- 進化論 A B C 一册 平裝五角
精裝六角
- 張默宗著
- 心理學 A B C 一册 平裝五角
精裝六角
- 復旦心理學院郭任遠著
- 變態心理學 A B C 一册 平裝五角
精裝六角
- 心理學學士黃捷慶著
- 性學 A B C 一册 平裝五角
精裝六角
- 南洋大學學院樂福沅著
- 衛生學 A B C 一册 平裝五角
精裝六角
- 沈繼香著劉清凡博士校訂

應用科學組

- 科學論 A B C 一册 平裝五角
精裝六角
- 北大理學士王爾森著
- 電學 A B C 一册 平裝五角
精裝六角
- 北大理學士王爾森著
- 攝影學 A B C 一册 平裝五角
精裝六角
- 化學教師吳靜山著
- 測量術 A B C 一册 平裝五角
精裝六角
- 東南大學理學士赫國均著
- 附告 本叢書的書目是沒有限制的
隨時加編隨時發表

中華民國十七年八月再版

不 准 翻 印

發行所

統計學ABC (全一冊)

(平裝五角 精裝六角)

(外埠酌加郵費運費)

著者 蔡 毓 聰

作 者 蔡 毓 聰

出版者 ABC叢書社

印刷者 世界書局

發行所 世界書局

上海四馬路 世界書局



