國立復旦大學商科研究所叢書

高等統計學

薛仲三 著

天津市立第二圖書館
藏書圖記

商務印書館發行

0001

112

臨時

五六年查訖

旧参
C 8
X Z 8

國立復旦大學商科研究所叢書

高等統計學

薛仲三著

商務印書館發行

18837

自　序

我國各大學商學院、農學院、教育或師範學院內所設之各科系，率將普通統計學列爲必修科，並將應用統計學——如經濟統計學、商務統計學、農業統計學、生物統計學、教育統計學、社會統計學、政治統計學、生命統計學、醫學統計學等——列爲必修科或選修科；商學院之統計學系及其附設之統計專修科，教育或師範學院之教育及心理等學系，復將高等統計學列爲必修科；統計學系或會計學系之統計組更將數理統計學列爲必修科，等等情形。統計學科之設置似已美備，無乃於內容之講授尙欠充實。是以，如何充實統計學科之內容，殊爲當務之急。著者以爲充實是科內容，應自選定適合課程標準之教科書始。因爲有了良善之教本，教師於講授時得有一貫之系統，可鏟除張皇補苴之苦；學生於學習時亦有途徑可尋，可減少無措手足之勞。

年來中外坊間出版有關統計學科之著作甚夥，其中可權作普通及應用統計學之課本者尙復不少；但求其適應高等及數理統計課程之標準堪作教本者，誠不數覯。爲濟目前之急，高等統計學及數理統計學之編輯，尤其是高等統計學，實刻不容緩。良以高等統計學介乎普通統計學與數理統計學之間，承上啓下，若不將高等統計學之內容及範疇釐訂於先，則關於數理統計學之一切實無法闡發於後。

天津圖書館藏書圖記
TIENTSIN LIBRARY

本書之主要目的有二：(1)作參考書之用，(2)作大學教科書之用.因此不免要發生下列各問題：如何能將其上之理論及方法便於應用各種特殊事實上,以與應用統計學收相得益彰之效,而不冒互相踰越之嫌？如何能使其與普通統計學密切銜接，庶使學者不至有扞格不入之歎？如何使其不侵犯數理統計學之境界,免貽人以疊牀架屋之譏？不但此也，就現行學制言，此學程每週授課三小時,須於一學年內講授完竣,如何能充分利用此既定之時數，使其內容儘量羅納一切統計學上實際問題之解答，普通公式之推演等事項？對此等等問題曾熟加考慮，並且得到尚覺滿意之解答，著者頗以此自慰，願以此告慰讀者.

本書約二十七萬字，共分八章：第一章略述統計學之基本概念，主要設備、實際運算上之規則與機巧等，俾學者得知統計學之津涯及統計實驗室是一個怎樣的"舞臺"，"演員"要演奏一些什麼技藝；第二及第三兩章以實例作出發點，推演統計常數之公式,以便計算之用,兼達理論互證之旨；第四章討論皮爾生各型頻數之配合方法及理論，係摹彷愛而德敦所著之頻數曲線與相關及焦恩思所著之初級統計學二書而寫，不過將彼等所用之公式皆一再加以推演，俾便應用；第五章研究曲線配合問題，係將各種配合方法及理論作詳盡之敍述；第六章爲尺度變換法之理論及其應用，係著者個人關於變換法初步研究所得之結果，用以激發學者對於此法作更進一步之追求；第七章爲相關問題，大部取材於閔茲數理統計學一書，蓋其上所講述之理論及所用之方法符號等與著者向日所主張及援用者多相脗合故也；第八章爲取樣問題，係用淺鮮實例鈎稽重要理論及原則. 總之，本書博採各家名著之精華，並參以個人從事

統計工作之經驗與夫教學研究之心得，而編定之者，因此略明各章之出處，用示不敢獨自掠美之意.

著者對於統計學之研究雖極感興趣，嘗有若干年之苦心；第以此科包括之理論極繁，涉及之事項彌廣，自愧學識與經驗有限，未能作周詳之貢獻！惟著者仍在孜孜上進，以本書內容之充實是求，倘蒙國內學者隨時賜教，指正書中之遺誤處，則不勝翹企歡迎之至！

本書凡三易稿：初稿由姜振黃、羅奎生、范光麟及黃士偉四位同學於暑假之暇分任繕校工作．復稿由田鳳調先生及范家榮女士擔任設計與抄寫，並且末稿由黃松岡、姜崇瑜、鄭忠賢、路基森、吳振英、唐尚治及徐元慶諸同學謄清，並承梁明策先生代爲繪圖．對於以上諸人著者實深感激．於付印之前，得蒙鄭郁華、褚鳳儀、李炳煥、李仲珩及單致任諸先生校正與補充，著者尤爲敬佩！然而書中一切缺點之責，仍由著者自負.

最後著者應當感謝商務印書館編審部駱尙熊先生，出版科姚心吾先生，及建華製版所周少山先生，因爲對於製圖、印刷、排版等事項，諸先生不惮煩勞的幫忙，實在超出他們應盡的本分之外.

民國三十六年十一月　　　　薛仲三序於江灣復旦

目　　次

第三章 動差、機率及常態曲線

第四章 皮爾生各型曲線

第五章 曲線配合

第六章 尺度變換法之理論及其應用

第七章 相關

第八章　抽樣問題

附錄

圖　　次

表次

高等統計學

第一章 緒論

第一節 名詞之沿革

統計一詞係由英文字 Statistics 或德文 Statistik 迻譯而來,並且該二字均係間接導源於拉丁字 Status. Status 含政治情形 (political state) 之意,蓋當時用 Statistics 或 Statistik 二字特指國家事業而言,殆無疑義. Statistics 一字遠在公元一七〇七年出版之萬象之原 (The Elements of Universal Erudition) 一書中即已見之,其中有一章曾用此字標題,並給以定義謂:"統計學乃示世界上各近代國家調劑政治之科學*". 彼時統計一詞之意義甚狹. 迨至公元一七八七年德人新莫滿 (E. A. W. Zimmerman) 於其所著之歐洲政況調查 (A Political Survey of The Present State of Europe) 一書的序言中,亦嘗引用此字;但其意義略比以前廣泛.

直至十八世紀之末,統計一詞之應用尚屬罕見;但此後不久,流傳就很廣了,引用者亦日衆. 如蘇格蘭統計紀錄 (Statistical Account) 之首任主筆辛可賴 (John Sinclair) 於公元一七九〇年發出之通告裏說:

---

* Willcox, W. F.: Ruarterly Publications of the American Statistical Association vol. 14, 1914 p. 278.

"德國曾舉行大規模之'統計訪問'"並且渠解釋"統計訪問"說:"統計訪問者,係關於人口、政況、出產情形等之訪問". 在統計紀錄緣起及發展裏說:'許多百姓猝然看見現在所用之新字 Statistics, 不免要驚愕的,以為一定由已有之英國字中能找一合適者,用以表明相同的意思. 但我於公元一七八六年漫遊北歐時,適值德國舉行政況訪問,僉謂此訪問為政況統計,……因此我想一個比較新穎的名詞,每易引人注意,便毅然採取之,並且希望今後多數人用牠,以至使牠成為大衆用語". 可見這個名詞在英國猶覺新奇時,在德國之文獻上已屢屢見之.

總之,在新英滿及華可賴時代(十八世紀),關於國家之特性幾乎完全偏重質的解釋,而忽視了量的敍述. 良以彼時可靠的數量論據 (quantitative data) 不可多得之故. 因此 Statistics 專指品質論據 (qualitative data) 而言. 降至十九世紀初葉,官方資料積聚日多,於是量的敍述一時膾炙人口. '統計'一詞之含義,於不知不覺中,有一致之轉移,卽是專指用"量"的方法說明國家之特性. 經過一再演變,以迄十九世紀末葉,統計不專指國家的資料而言,其他科學所屬之數量論據,亦得以統計名之矣,如數量論據屬於人類學者曰'人類統計',其屬於氣象學者曰'氣象統計'是.

第二節 定義及發展

1. 定義 統計資料可分之為兩類:第一類係受多種因子影響者;第二類係祇受一種因子影響者. 例如人之身長受種族、祖先、居處、幼年營養、年齡、性別等之影響,故身長之論據屬於第一種;再如氣體之體積,

若保持温度不變，所受之影響祇有一種因子——氣壓，因此氣體體積之論據屬於第二種。那麽，統計資料是指何種而言呢？乃係指着第一種說的。如此，物理上或化學上的論據似乎不在統計資料範圍之內了；其實則否。就像方纔所舉關於氣體與氣壓之例，我們無法控制温度不變，因此氣體之體積不僅受氣壓的影響，還要受温度的影響。不但此也，觀察者心思之綿密性，以及其所用之儀器的精確度，均能影響所得結果之良窳。是以統計研究中所指之論據，除包含社會學上者外，自然科學上之大部分亦包含在內。

基於上面之討論，我們可瞭解遊爾 (G. Udny Yule) 給予"統計" (statistics)，"統計方法" (statistical methods) 及"統計學" (theory of statistics) 各名詞所下之定義。茲分別譯述於下：

"統計"者，指數量論據而言，此論據係受多種因子影響至相當顯著之程度者。

"統計方法"者，乃是種種特殊方法，專用以處理受多種複雜因子影響之數量論據者。

"統計學"者，乃是一種專門科學用以闡發統計方法者。

由此可見，統計方法可應用於受多種複雜因子影響之數量論據上面，該數量論據不限於受多種雜複因子影響"至相當顯著之程度"；是卽統計方法應用之範圍，不限於統計上面，其他自然科學論據之處理及分析亦得應用之。

2. 發展　法人雷蒲拉斯 (Pierre Simon Laplace, 1749–1827) 及德人高斯 (Karl Friedrich Gauss, 1777–1855) 之觀察差誤論，實開統計學

研究之先河，獨惜其理論在當時並未惹人注意！洎乎十九世紀末葉，學者風起雲湧，方使其臻入輝煌之境。比人郭泰來 (Lambert Adolph Jacques Quetelet, 1796–1874) 乃其中之傑出者，曾將雷高二氏之方法應用於實際之問題上面，對於統計學之貢獻誠非淺鮮。雖然，對於二量相關問題，從未有人注意。迨英人葛爾登 (Francis Galton, 1822–1907) 研究遺傳問題時，始發覺二數相關之測度，乃問題之研究中不可或少者，遂有測量相關度等法之發明。當時所用之法固然不甚完備，但於實際之應用上尚覺滿意。葛氏研究興趣之熱烈，曾激起英國負有盛譽之生物學家皮爾生 (Karl Pearson, 1857–1936) 及魏而敦 (W. F. R. Weldon, 1860–1906) 得入其門牆，共同舜研，大張其軍，統計學之基礎遂得以日益鞏固。

就個人之貢獻言，統計學得有今日之長足發展，皮氏之功甚偉。茲舉其犖犖大者四端，如下：

(1) 用動差法 (method of moments) 以配合頻數曲線 (frequency curve).

(2) 用偏態頻數曲線 (skew frequency curve) 以描寫自然現象。

(3) 用相關理論 (theory of correlation) 以解決遺傳及進化問題。

(4) 用 X^2 測驗 (chi-square test) 以判斷配合之適度 (goodness of fit).

以國家為單位而論，關於統計學之發展，近百年來英國學者實處領導地位。除葛皮魏諸人外，如傅耳 (William Farr, 1807–1883)、薛伯 (W. F. Sheppard, 1863–1936)、愛奇渥斯 (F. Y. Edgeworth)、游爾、費緒、(R. A. Fisher)、愛爾德頓 (W. P. Elderton)、焦恩思 (D. C. Jones) 等，對於統計學上之貢獻亦夥。

此外另成一系，對於偏態曲線理論亦有相當相關發者，則爲斯堪地納維亞學派 (Scandinavian School). 其中代表人物爲葛蘭母 (J. P. Gram)、翟耳 (T. N. Thiele)、沙利爾 (C. V. L. Charlier)、魏克澤 (S. D. Wicksell)、費恩 (Arne Fisher) 等. 費氏曾挾此派理論，在美講學多年，一時美國學者趨之若鶩，後漸衰微，今則幾乎不復被人稱述矣.

第三節 統計方法之步驟

統計方法可分爲三個步驟：論據之搜集，論據之處理及論據之分析. 茲將各步驟之特點，撮要討論如下：

1. 論據之搜集 統計家第一必須思考的問題，爲如何搜集其論據. 許多科學，如經濟學及社會學，不能自己準備論據，必須仰賴官方所發表者. 但此等資料往往與其本人之目的相左. 調查員 A 欲研究蔗糖銷路暢滯情形，發現官方未將蔗糖及菜糖數字分列. 調查員 B 希望比較歷年物價漲落情形，發現當公元一九一四至一九一八年戰爭時的數字，竟付缺如. 調查員 C 想着研究人民貧苦狀況，不得不借助於間接的材料. 縱令論據之如何不完善，以及不適切研究者的需要，但不得不委曲利用. 自然在其他情形中，如研究氣象學、生物學等，調查者有時可以得到自己所欲得到的資料. 總之，無論自己搜集之初級資料 (primary data)，或取自已經發表之次級資料 (secondary data)，在未進行整理之先，應當審核他們的可靠性，並應設法彌補論據之缺憾. 否則，應用精密之統計方法於可疑的而且殘缺的論據上，未免虛費光陰.

2. 論據之處理 既經得到滿意的論據之後，統計家就要進行整理

牠們.處理論據之惟一法門,即是將牠們歸納就範.人的靈思不能領略大批錯綜雜陳的數字,因此一定要將可資利用之論據經過歸納作用,使其有條有理的呈現,俾讀者一見可得到清楚的印象.例如,我們討論一千個人的身長,若將一千個人的身長一一列舉,令人無法得其要領,必須將各數值歸納,在人的腦子裏總會生出一個清楚的圖樣.在本例中,就是要將這一羣人之身長分成若干組:譬如說,每相鄰兩組身長之差爲3吋,其最高與最低身長之差爲3呎,我們可將身長歸納爲十二組;換言之,即是用十二個數值代替一千個數值——一千人之身長.可見這樣做法,未免犧牲了一些事實,十二個數目告訴出來的事實,絕對沒有一千個數值告訴出來的多.然而,以其能給我們以明晰之觀念,這十二個數目或者正緊需要;所去掉的事實,或者無關宏旨.因爲我們祇要知道人們的平均身長爲若干吋,以及大多數人的身長爲何就夠了.故歸納爲處理論據時不可或少的手續.至於歸納到何種地步爲止?則視事實失之於犧牲,能否得之於明晰以爲斷.

3. 論據之分析 當統計家將論據歸納而成適宜形式之後,可以說已完成其初步工作;再進一步的工作就是分析.分析之目的在能以簡馭繁,即是要以統計常數或數學函數,描寫大批論據.

對於某種論據適用某種統計方法以分析之,爲本書以後各章所要研究者;至於如何搜集準確之論據,以及如何處理大批論據之手續,爲統計調查及方法中所有事,本書不復加以研討.

第四節 統計實驗室

我們大概都承認:科學如物理學、化學、生物學、心理學、地質學、工

程學等之研究，不應一味注重課堂上之授受，並應注重實驗室內之實習，俾學理與實際工作得以互相印證。我們更都知道：有些課程之實習時數往往超過演講時數，而以實習為重。統計學就是這樣的一門學科，最低限度實習應與演講並重。設有統計系、專科、以至開有統計課程之學校，均應有統計實驗室之籌設，並應實之以設備，予學生以實習之機。現在將統計實驗室之起碼設備，略述如下，用資參考。

1. 關於畫圖及製表者　統計工作有時要畫圖及製表。為使所畫之圖及所製之表美觀，並為節省學生時間計，統計實驗室內宜置備下列各種儀器，發給或借給學生應用：

a. 格紙　格紙，如普通格紙 (ordinary graph paper)、單對數格紙 (semi-logarithmic paper)、雙對數格紙 (double logarithmic paper)、極坐標格紙 (polar coordinate paper)、機率格紙 (probability paper)、製表紙 (tabulating paper) 等，有時在文具店不易購得，最好發給或以低價售給學生應用。

b. 文具　各種文具，如畫圖版 (drawing board)、畫圖筆 (ruling pen)、描圖紙 (tracing paper)、繪圖墨水 (india ink)、寫字模 (lettering guide)、丁字尺 (T square)、雲形板 (french curve)、三角板 (triangles)、三稜尺 (triangular ruler)、硬鉛筆 (hard pencil)、軟橡皮 (soft eraser)、硬橡皮 (hard eraser)、圖釘 (thumb tack) 等，均應設置，以備臨時之需。

c. 儀器　各種儀器，如圖畫器 (drawing instruments)、比例儀 (proportional divisor)、放大器 (pantograph)、求積儀 (polar planimeter) 等，亦宜置備，俾資應用。

2. 關於分類者　設有大批統計資料，即如有十萬份農村經濟狀況調查報告，並且其中包括之項目異常繁多，專恃人工分類，既費時間與精力，且易滋誤．是以統計實驗室內宜設備統計卡片 (manila cards)、打孔機 (punching machine)、驗孔機 (mechanical verifier)、分類機 (sorting machine)，以供學生實習之用，兼備研究之需：

a. 統計卡片　統計卡片為一種特製之卡片，其厚薄及大小均有一定．此卡片之主要用途，在能將原始記錄用其上所鑿之孔表示之，俾可利用分類機分類．

b. 打孔機　統計卡片之孔即用此機鑿打者．

c. 驗孔機　因為所鑿之孔代表原始記錄，孔一誤鑿，分類所得之結果即將隨之舛錯．故用打孔機打好孔洞之後，宜用驗孔機驗對所鑿之孔有無錯誤，誤則改之．最好兩人分任打孔與驗孔工作，以免發生雷同之謬誤．

d. 分類機　統計卡片上所鑿之孔經驗對無誤後，即宜用此分類機分類．

統計實驗室內至少宜設備分類機一架，打孔機及驗孔機各兩架，統計卡片幾十萬張．就實際上說，一架分類機應配以五架至十架打孔機；但在實驗室內係以表演實習或研究為目的，無需配以多架打孔機，只要有兩架即敷需要．兩架打孔機中最好有一架為電動複打機，以便照樣復打之用；一架為普通的，專為普通用手打孔之用．驗孔機既為驗對卡片上所打之孔有無錯誤而設，故其設備之架數應與打孔機相等．統計卡片有 45 行與 80 行者之別，昔用 50 行者，今則已相率改用 80 行者，蓋行

數多者代表之事實亦多故也. 原始報告上之記錄,無論怎樣複雜,用 80 行統計卡片可能將所有事項悉行納入. 用 80 行之統計卡片宜用 80 行之打孔機、驗孔機及分類機自不待言.

3. 關於計算者 對數表 (logarithmic table)、計算尺 (slide rule)、計算機 (calculating machine)、加減機 (adding machine)、巴樓表 (Barlow's Tables)、皮爾生表(Tables for Statisticians and Biometricians) 等,均為計算之工具,亦為統計工作中之不可少者. 茲分述之如下:

a. 對數表 對數表現在雖非計算之主要工具,然而成為對數與真數互相檢查之重要用表,價值並未少減.

b. 計算尺 計算尺之發明,計算上增加了一種工具,其功用雖不能與對數表相等,但用以計算三四位有效數字,尚覺方便.

c. 計算機 計算機為在統計計算上不可或少者,用牠計算大批數目,不但其準確之度遠非對數表及計算尺之所能企及,卽是運算速率之大亦非其他工具所敢比擬.

d. 加減機 此機專為做加法或減法之用.

e. 巴樓表 巴樓表一書中,列有數值之平方、立方、平方根、立方根、反商等,凡數值在一萬以內之平方、立方、平方根、立方根及反商等一檢卽得,便利從事統計工作或研究者甚多.

f. 皮爾生表 皮爾生所編輯之表內,列有種種統計工作上所需要之常數或函數之值,用之可減少自己計算之勞.

時至今日,可以說計算機為普通計算之惟一工具,統計實驗室內應當購置若干架,借給選習統計實習課程之學生應用. 最好每人得以借用

一架爲準;換言之,購置之架數宜與選習統計課程最大班級之學生人數相等.於計算機之外,宜輔之以加減機,最好購置二三架置於實驗室內,以備學生隨時應用.

計算機既能做加法及減法,爲什麼於計算機之外還要購置加減機呢?蓋通用之加減機,祇有十個打字鍵,容易打數;並且同時將欲加或減之數目及其結果,印於紙條上,容易與原來數值核對.如果紙條上所印之數目與原來欲加或減之數目相同,則示所得之結果無誤;如發現有歧異處,則立可更正之,得免重行加減之煩.

第五節 孟祿計算機

1. 與其他各種計算機之比較 計算機有多種.美國出品中最著名的就有三種:(1) 傅瑞登計算機 (Friden calculating machine)、(2) 麻詗德計算機 (Marchand calculating machine) 及 (3) 孟祿計算機 (Monroe calculating machine). 這三種著者均曾用過:著者在北平協和醫學院執教時所用者係第二種及第三種,而以用第三種之時間最久;留學美國在美國人口清查局 (Bureau of the Census) 實習,參加編製壽命表 (life table) 工作,計算壽命表上各函數時,所用者係第一種.就計算速率說,三者均甚高,幾乎不相上下;若強定優劣,則以第一種計算之速率爲最高,第二種次之,第三種又次之.就構造說,第三種比較簡單,不容易發生障礙,其他二種正與之相反.因此我國甚至美國最通用者不是首二種,卻是第三種.

2. 各部分之名稱、位置及功用 孟祿計算機有手搖與電動者之別.

不論手搖者或電動者又各有種種不同之樣式，但其構造之原理是大同小異的．祇要知道一種計算機之用法後，其他各種甚至其他不同牌號者，亦不難知之．於是舉一普通電動孟祿計算機爲例，說明其各部之名稱、位置及功用，如下：

孟祿計算機有一個可以左右移動的車，叫做收輪車 (carriage)；並有一個機體 (body)．收輪車上面有兩組可以旋轉的輪：一組叫做記次輪 (upper dial) (1)，一組叫做記值輪 (lower dial) (2)；在機體極右端上面有一突出之樁，叫做移車柄 (shifting bar) (3)；在其右側有一個搖把，叫做還原柄 (clear crank) (4)；在機體右上面有兩個開關：一個叫做減開關 (minus bar) (5)，一個叫做加開關 (plus bar) (6)；機體正面爲打字盤 (key board) (7)；右下側有一加減鍵 (non-repeat key) (8)，一乘除鍵 (repeat key) (9) 及一還原鍵 (clear key) (10)；並在左前面有一移車桿 (carriage shift lever) (11)．各部分之位置見圖 1.

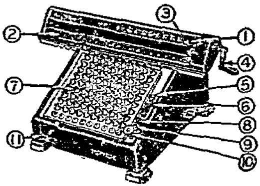


圖 1. 孟祿計算機。

普通十位電動孟祿計算機各部分之作用與十位手搖者沒有什麼不同，祇是前者以電動機 (moto.) 旋轉運算時所用的搖把（稱做運算柄，

operating crank) 而已。用手按加開關一下,便等於依時針方向旋轉運算柄一週;按減開關一下,等於反時針方向旋轉運算柄一週。

3. *用法* 計算機計算之效率遠非其他計算工具所能企及,已如上述。不但此也,計算機能連續運算,即能將很複雜之運算結果,如某數與其他二數乘積之差、兩兩數值乘積之代數和、直線式 $a+bx$ 之值等,無須寫出中間步驟,用一連續算法 (continuous operation) 求得之,計算上之效率更能增進一籌。現將連續算法之妙用分別舉例說明如下:

例 1. 試用連續算法求由 15 裏減去 3.141593 與 2.718282 相乘積之差。

運算手續為:

a. 清計算機 即是將記次輪、記值輪及打字盤上面的數字一一清除。這是每於進行運算以前必經之手續。

b. 定小數位 即是將記次輪、記值輪及打字盤上面的小數位一一定好。在本例裏(在以下各例裏亦然)所用的小數位:

在記次輪上面的為 6 位,

在記值輪上面的為 12 位,並且

在打字盤上面的為 6 位 (參閱第七節 4)。

c. 鎖乘除鍵 即是將乘除鍵用手指按下去。

d. 登被減數 即是將被減數登記在記值輪上面。在本例裏要登的數目為 15.000 000,000,000

e. 打被乘數 即是將被乘數打在打字盤上面。在本例裏要打的數目為 3.141593.

f. 移載輪車　即是將載輪車向右移動若干位。在本例裏應當將載輪車向右移動六位(何故?)通常載輪車的位置,係在機體極左端,即是記位輪或記次輪上面第一位數字與打字盤上第一行數字在一條直線上。

g. 做乘法　即是將減開關按下直等到記次輪上面現出紅色的(何故?) 2 爲止;將載輪車向左移動一位,並且再將減開關按下,直等到記次輪上面現出紅色的 7 爲止;將載輪車再向左移動一位,並且再將減開關按下,直等到記次輪上面現出紅色的 1 爲止;如此等等。於是在記次輪上面最後現出 2.718282.

h. 讀結果　記位輪上面所餘之數值即爲欲求之結果,讀出之爲 6.460264.

例 2　試用連續算法計算下列各數相乘積之代數和:

$$3.141593 \times 2.718282$$
$$-2.302585 \times 0.434294$$
$$+0.674490 \times 1.234567$$
$$-2.345678 \times 3.456789$$

運算手續爲:

a. 清計算機　(參閱例 1, *a.*)

b. 定小數位　(參閱例 1, *b.*)

c. 鎖乘除鍵　(參閱例 1, *c.*)

d. 做乘法　即是依次乘出來各雙的數值。因爲我們假設學生已經曉得了孟祿計算機的用法,所以關於做乘法的普通運算的手續,不復贅及。但有應注意者:本例是叫我們一下求出兩兩數值相乘積之和,應當

乘得第一變數値之後，祇要將記次輪及打字盤上面數字清除，千萬莫清記值輪上面的數字；第二變數値之積爲負數，當做乘法時應使記次輪上面現出紅色的數字，並且照樣將記次輪及打字盤上面數字清除，而不清除記值輪上面的數字；如此等直至將末變數値乘完爲止。

e. 讀結果　記値輪上之數値即爲欲求之結果，讀出之爲 0.263926.

例 3.　設 $x=5$、7、15、12……，求 $12+6x$ 之值。

此處 $12+6x$ 爲 $a+bx$ 之特殊形式，即 $a=12$, $b=6$. 故運算手續可分述之如下：

a. 將 12 登記在記值輪上面.

b. 將記次輪上面的數字 1 及打字盤上面的數字 12 清除.

c. 將 6 登記在打字盤上面.

d. 將乘除鍵鎖住.

e. 按加開關使記次輪上面現出黑色 5，如是記值輪上面現出數值 42，即爲欲求之第一個結果；再按加開關使記次輪上面的 5 變做 7，則得到第二個欲求之結果 54；餘類推.（若 a 之值爲負當如何運算?）

第六節　開方

用計算機求某數之平方根，也很迅速的．普通最常用之法爲"遞減法"．另有一種方法係算術學上通用之筆算開方法，但著者將其全盤搬到計算機上面來的，特名此法曰"試除法"．後法較前法既簡便又迅速，學者可比較之．倘學者經過比較之後，一定會欣賞後法的，並且會自動的放棄前法．現在將這兩種開平方法，分別舉例說明之如下：

例 1. 用遞減法，求數值 2114.833692757041 之平方根.

運算手續為：

a. 清計算機 (參閱例 1, *a*.)

b. 定小數位 (參閱例 1, *b*.)

c. 鎖乘除鍵 (參閱例 1, *c*.)

d. 登記數值 即是將欲開方之數登記在記值輪上面. 在本例裏應登之數為 2114.833692757041.

e. 清記次輪 即是將記次輪上面的數字 1 清除.

f. 移裁輪車 即是將裁輪車向右移動若干位. 在本例裏向右移動七位. (何故?)

g. 求首位數 即是在打字盤上某相當行依次打以 1、3、5 等奇數，並且每打一數就要按減開關一下，意思是要由記值輪上面之數值裏遞次減去所打之數值，直至不夠減(或聽到鈴響)時為止. 在本例裏要在打字盤第八行上依次打以 1、3、5 等數，並且依次在該行遞減以 1、3、5 等數，直減至數字 9 鈴響為止. 聽到鈴響之後馬上要按開關一下，使記值輪上面之數值還原，隨後將這個 9 變做 8 (=9-1, 即是打字盤上之數值減以 1)..

h. 求次位數 將裁輪車向左移一位；在打字盤次一行(即第七行)依次打以 1、3、5 等數字，並且同樣每打一數之後就要按減開關一下，意思是要由記值輪上面的餘數裏遞減以 81、83、85 等數值，直減至 91 鈴響為止. 同樣聽到鈴聲後馬上按加開關一下，使記值輪上面的數值還原，並且將最後所打之 1 變做 0 (=1-1, 即是將現在打字盤上之數值減以 1).

i. 讀結果 重複運算手續 *h*, 記次輪上面將現出紅色的數字 45.987321, 即所求之結果.

例 2. 用"試除法"求 2114.833692757041 的平方根.

運算手續爲:

a. 清計算機 (參閱例 1, *a.*)

b. 定小數位 (參閱例 1, *b.*)

c. 鎖乘除鍵 (參閱例 1, *c.*)

d. 登記數值 (參閱例 3, *d.*)

e. 清記次輪 (參閱例 3, *e.*)

f. 移載輪車 (參閱例 3, *f.*)

g. 求首位數 所給之數有四位整數,因知其由兩位整數自乘得之者.我們知道 21 的平方根爲 4, 因此在打字盤第八行上打以 4, 並用他去除記值輪上面的數值,直至記次輪上面現出來紅色的 4 爲止(打字盤上面所打之數字應與在記次輪上面所現出之紅色數字相同). 於是得到平方根之第一位數值 4.

h. 求次位數 求平方根次一位數值時用六個步驟: (1) 倍數,即是將打字盤第八位上面的數值 4 以 2 倍之,使之爲 8; (2) 移車,即是將載輪車向左移動一位; (3) 試除,即是將記值輪上面的餘數以 8 除之,得商數爲 6; (4) 還原,即是按加開關使記次輪上面現出之數字 6 變做 0; (5) 打數,即是在打字盤第七行上打以 6; (6) 再除,即是再做除法,如此所得之商數不夠 6 而爲 5. 倘遇此種例外情形,我們應當將記次輪上面的 5 還原,將打字盤上所打之 6 變做 5, 然後再做除法,同時要注意除

得之商數務必爲 5, 切莫超過. 於是得到平方根的第二位數值 5.

i. 求再次一位數 求再次一位數時亦用六個步驟: (1) 倍數, 卽是將打字盤第七位上的 5 以 2 倍之, 或者說將 85 變做 90; (2) 移車, 卽是將載輪車向左移動一位; (3) 試除, 卽是以 90 除記值輪上面的餘數, 得商數爲 9; (4) 還原, 卽是使記次輪上面 9 還原; (5) 打數, 卽是在打字盤第六行上面打以 9; (6) 再除, 卽是再做除法, 如此所得之商恰恰爲 9. 於是得平方根之第三位數值 9.

j. 讀答案 連續重複運算手續 *h* 或 *i*, 記次輪上面將現出紅色數字 45.987321, 卽爲欲求之答案.

"試除法" 之特點, 在能利用除法決定出來方根次一位數之值. 自然其第一及第二位數值不能用此法決定, 因決定出來者往往比實在應得之數值爲大; 但是從此以後, 用除法決定出來方根次一位數之值常是正確的, 就像決定第三位數值 9, 第四位數值 8 等等. 因爲用計算機做除法, 尤其用電動的, 是不假絲毫思索的, 並且很迅速的, 如此平方根次一位數之決定極爲捷便, 故 "試除法" 乃是一個比較合乎理想求平方根的法子. 總括一下說: 假設我們已經求得平方根首二位數之值, 除手續 *h* 所列之特殊情形外, 求平方根第三位及以後各位之運算手續如下:

(1) 倍數、(2) 移車、(3) 試除、* (4) 還原、* (5) 打數及 (6) 再除.

以上兩節所講的計算方面, 不過舉其犖犖大者, 學者如能隨機應變, 自能得到一切運算上之機巧. 至於一般習用之運算手續, 譬如利用反商法以求某頻數分配之百分分配, 利用乘除鍵以求由一個固定數值與各

* * 有經驗之計算者, 有時能用心算算得商數, 故此二步手續可省.

個變值之差，利用簡單的恆等式 $273\times(100-3)\equiv273\times97$ 以求 273 與 97 相乘之積等等，學者或已知之，不再詳述。

第七節 計算通則

全部算草都是有價值之永久記錄，均宜妥愼保存，以備日後覆核或參考之用。因此計算時必須遵守下列之規則：

1. 算草宜整潔，並宜排列有序。

2. 算草宜用墨筆書寫，每頁並宜標以頁數及日期。

3. 不要將原來寫在紙上的數字用橡皮擦掉，也不要將計算的中間步驟寫在另外一張廢紙上，旋卽拋棄。遇有錯誤之數字，應當用一條橫線畫掉；所有計算步驟均宜列出。

4. 在一般情形中應將有效數字算至六位，因爲這樣可使運算一致，並且可以免除每次判斷所得數值準確度之煩勞。

5. 在計算比率數、百分數等數值時，不得援用上面所列之規則，在此刻至多將有效數字算至比實在量度多一位就夠了。

6. 決定某數末位次一位數之捨入，以下列事實爲準：若末位次一位數

大於 5，應將末位數升以 1。

小於 5，應將末位不變。

等於 5，當末位數爲奇數時則升以 1；當其爲偶數時則仍舊不變。

7. 表示計算結果時，有效數字之位數不得超過實在量度一位或兩位。

8. 所有統計常數之單位應當標出,不可輕忽.

第八節 製表通則

怎樣製表原無一定標準,現在僅將製表時應當遵守之通則分述於下,提供參考:

1. 論據十分複雜者應當將其分列數表.

2. 重要之事實應當置於表之第一行或第一列.

3. 標題 (title) 應當簡賅.

4. 有空間及時間性之論據,標題時一定要將事實發現之地點及時間列入.

5. 資料來源應當註於標題之下或表底.

6. 如果一表佔數頁,須將題目完全逐頁標出.

7. 標目 (heading) 要確切,要寫出所用之量度單位,必要時加以定義或註解.

8. 冗長之表應當用字碼或字母註明行列次序.

9. 表之縱標目 (caption) 上端宜畫以雙橫線.

10. 表之橫標目 (stub) 與表體之間宜畫以單橫線.

11. 表體下端亦宜畫以單橫線.

12. 各縱行宜用單縱線隔分.

13. 表上所有之單線及雙線應當粗細一律.但為醒目起見,有時得變通之,畫以粗細不同之線.

14. 表之左右兩側一律不要畫線;但調查或登記表不在此限.

15. "總計"與表體之間宜用線分開.

16. 表中之數目字宜一律用阿拉伯字碼填寫,並宜對齊位次或小數位.

17. 較長之表應當每五列空一列.

18. "總計"或其他重要的數字可用重體字表示.

19. 負數,與本年相比較之上年數字、比率數、百分數、或平均數及與實際對照之估計數值,可用斜體字表示.

20. 通常將縱行總計列在表之底端,橫列總計列在表之右端,但有將其置於表體之上端或左端者.(當用分類機分類時,向是先得到總計,因此將總計置於表之上端或左端較為方便.美國人口清查局刊行之人口及生死統計年報,率將總計列在表之上端或左端,就是這個道理.)

第九節 繪圖通則

在自然與社會科學研究中,人多喜用圖表示其所得之結果;但其所用的方式彼此各異,殊乏一致之標準,不無遺憾!為補救這個遺憾起見,美國工程、統計、經濟、生物等學會代表,於公元一九一五年組織一個委員會,討論繪圖之統一方法,曾發表一篇初步報告.*茲將該委員提示各點,譯述於下:

1. 圖之普通排列宜自左而右.

2. 儘其可能用直線代表數量,因為用面積或體積代表容易使人誤解.

* Joint Committee on Standards for Graphic Representation Preliminary Report, Quart. Amer. Stat. Assoc., Vol. 14, pp. 790-7, 1915.

3. 畫曲線圖時，垂直尺度宜加意選擇，俾能將零線畫在圖上。

4. 如果垂直尺度的零線不能照常畫入圖內，宜利用水平斷線法將其畫入。

5. 尺度的零線應當畫粗些，叫牠和其他各坐標線截然有別。

6. 帶有百分數尺度之圖，常要設法將百分線顯出，其他用做比較之線也應該設法使之格外分明。

7. 圖之尺度關於時間者，並且所表示之時間段落是不完全的，最好不要將首線及末線加粗，因爲這兩條線並非表示時間之起止。

8. 當曲線畫在對數坐標上面時，圖之邊線的尺度宜爲十的乘冪。

9. 坐標線除爲便於瀏覽外不宜太多。

10. 所繪之曲線與坐標線應當有明晰的區別。

11. 代表一組觀察值之圖，要將代表各個觀察值所有各點清楚的畫出。

12. 圖之水平尺度須自左向右讀之，垂直尺度須自下向上讀之。

13. 尺度之數目字應當寫在縱軸的左邊及橫軸的下邊，或者各沿着坐標軸書寫。

14. 圖所代表之論據或公式常要列入。

15. 如果不便將論據列入圖中，可另表列出。

16. 圖中所有文字及數目字應當便於站在圖下或圖右之人閱覽。

17. 圖之標題宜力求清楚完備。爲使標題明白起見，不妨加入標目或說明。

問題 I

1. 閱讀附錄 IV 參考書 44 之緒論 (Introduction), 並做一簡單報告,字數以五百為限.

2. 瀏覽附錄 IV 參考書 26 及 42 一過,並做一簡短報告,字數不拘.

3. 簡述皮爾生、傅耳或其他某統計學家之一生事蹟.

4. 閱讀附錄 IV 參考書 12 以後,詳論斯堪地納維亞學派關於偏態曲線理論逐漸消沉不復見重於美國學者之原因.

5. 我國算盤發明甚早,其究竟始於何代?演進之情況如何?詳撰一文作有系統之敍述.

6. 試計劃一理想之統計實驗室.

7. 用各號寫字模寫 LETTERING GUIDE 0123456789 等字,以資練習.

8. 試用比例儀將圖 38_c 放大若干倍,以為演講之用.

9. 試用放大儀將圖 38_b 放大若干倍,以為某種目的之用.

10. 試根據一本比較精確的地圖,如丁文江等著中國分省地圖,用求積儀,求出我國各省之面積.

11. 假設有人說:"做加減法時用算盤比用加減機或計算機方便而且迅速". 你怎樣證明他所說的乃是無稽之談?

12. 求表 1 裏各數值之和, ΣX.

13. 求表 1 裏各數值平方之和, ΣX^2.

14. 求表 1 裏各數值之標準差：

$$\text{標準差}=\left[\frac{1}{147}\Sigma X^2-\left(\frac{1}{147}\Sigma X\right)^2\right]^{\frac{1}{2}}.$$

15. 驗算表 2 及 3 裏各數值有無錯誤.

16. 用遞減法求 2 之平方根，至嫻熟無誤爲止.

17. 用遞減法求 17 之平方根，至嫻熟無誤爲止.

18. 用試除法求 2 之平方根，至嫻熟無誤爲止.

19. 用試除法求 17 之平方根，至嫻熟無誤爲止.

20. 用遞減法及試除法分別求 2、3、5、7、10、11、13、17、19 及 23 之平方根，結果算至小數六位；並分別記載所用之時間，以便由計算速率決定該二法以何者爲優.

21. 試將下面戶口調查表上各項記錄編成號碼，以便用統計卡片表示原始事實：

戶口調查記錄表

稱謂	姓名	性別	年齡	出生年月日	婚姻狀況	籍貫	職業	服務地點

22. 用打孔機按教師指定已經編好號碼之事實打在統計卡片上，然後用驗孔機驗對所打之孔有無錯誤之處.

23. 將教師指定之統計卡片用分類機分類，並將結果填入下表：

年齡	人數		
	男女合計	男	女
0—			
5—			
10—			
15—			
20—			
25—			
30—			
35—			
40—			
45—			
50—			
55—			
60—			
65—			
70—			
75—			
80—			
85—			
90—			
95—			
100—			

24. 試按照製表通則所列各項繪製一實際表格.

25. 繪圖說明繪圖通則內所列各項之涵義.

第二章　頻數分配之分析

第一節　集中常數與離勢常數之意義及其種類

全部統計學內所研究者不外有關頻數分配之各分析問題，故潘耳(Raymond Pearl)給統計學所下之定義為："統計學為科學之一支派，研究事物種類發現之頻數，或事物品質發現之頻數者"*.研究頻數分配所用之方法因頻數分配所關變數之多寡而互有異同.一個變數之頻數分配得用集中常數,離勢常數等研究之,此即現在本章內所要研究者;得用其他常數——如動差、偏度、峯度等——研究之,此即將來在第三章內所要研究者;得用曲線配合法研究之,此即將來在第四章內所要研究者;並且得用測定樣本波動法研究之,此即將來在末章(第八章)一部分內所要研究者.兩個或多個變數之頻數分配得用曲線配合法研究之,此即將來在第五及第六兩章內所要研究者;得用常數——如相關係數、相關比等——研究之,此即將來在第七章內所要研究者;並且得用測定樣本波動法研究之,此即將來在第八章一部分內所要研究者.

大多數頻數分配表具有一種特徵,即頻數分配各組之頻數率是由少而多,迨達到最多時便逐漸減少,終詎於無.此種特徵即在統計學上所說的收斂趨勢 (convergent tendency), 並且得用常數或統計常數

---

* 見附錄 IV 所列之參考書 29.

(statistical constant) 描寫之. 描寫此種趨勢之統計常數叫做平均數 (mean, average)、或叫做地位常數 (constant of position) 或叫做集中常數 (constant of central tendency).

集中常數有五種,其爲

(1) 算術平均數 (arithmetic mean, arithmetic average),

(2) 中位數 (median),

(3) 衆數 (mode),

(4) 幾何平均數 (geometric mean, geometric average), 及

(5) 倒數平均數 (harmonic mean, harmonic average).

就中最常用者爲算術平均數,其他四者在特殊情形中亦有時用之.

兩個頻數分配之集中常數,有時彼此極相接近,甚或完全相同,但各變量散佈在集中常數左右之疏密情形未必一致,如下圖所示.該圖係兩條頻數曲線,其平均數以及頻數之和是完全相同的. 但一則高聳,顯示各變量密佈於集中常數 A 點之兩側;一則扁平,顯示各變量疏散於 A 點之兩側.如是於集中常數之外還需要其他統計常數,將各變量散佈在集中常數左右之疏密情形表示出來. 表示此種情形之統計常數叫做離勢常數 (constant of dispersion, variation, variability).

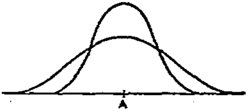


圖 2. 兩種離勢不同之頻數曲線。

通常最習知之離勢常數亦有五種,其爲

(1) 標準差 (standard devi tion),

(2) 平均差 (mean deviation),

(3) 四分位差 (quartile deviation),

(4) 全距 (range), 及

(5) 均互差 (mutual deviation).

其中亦以第一種最關重要,第二種次之,第三種更次之,第四種再次之,並且第五種幾乎無人問津.

此外更因離勢常數受原始度量時所用尺度之影響,有用離勢係數(coefficient of variation)表示頻數分配之離勢者.習見之離勢係數有三種,其爲

(1) 標準差係數 (coefficient of standard deviation),

(2) 平均差係數 (coefficient of mean deviation),

(3) 四分位差係數 (coefficient of quartile deviation).

其中常用者爲第一種,並且一般泛稱之離勢係數,即指第一種而言.

關於各集中常數,離勢常數及離勢係數分別檢討於下.

第二節　算學符號及其使用法

Σ 爲統計學上最常用之符號.現在爲使學者澈底明瞭該符號之意義及用法起見,於研討各種統計常數——集中常數、離勢常數等——之前,舉出幾個關於算學符號 Σ 之簡單定理如下:

設 X 表示變數, $X_1, X_2, \cdots, X_N$ 表示變量.則各變量之和, $X_1+X_2+\cdots+X_N$, 可用符號 $\sum_{i=1}^{N} X_i$ 表之;或在不致淆混情形中,用比較簡單之符號 ΣX 表之.即

$$\sum_{i=1}^{N} X_i = X_1 + X_2 + \cdots + X_N$$

或者
$$\Sigma X = X_1 + X_2 + \cdots + X_N,$$

彼處 $\sum_{i=1}^{N} X_i$ 讀做各 X_i 相加，i 由 1 至 N.

如果用 X_i 表示表 1 裏第 i 個觀察値，則 $X_1 = 105$ 公分，$X_2 = 122$ 公分，…，$X_N = 104$ 公分，$\sum_{i=1}^{N} X_i = 105 + 122 + \cdots + 104$；並且 $N = 147$.

如果 $X_1 = 1, X_2 = 2, \cdots, X_N = N$；則

$$\sum_{x=1}^{N} x = 1 + 2 + \cdots + N,$$

以及
$$\sum_{x=1}^{N} x^2 = 1^2 + 2^2 + \cdots + N^2.$$

定理 1. 兩項或多項代數和之總和等於各項總和之代數和．用符號表之爲

$$\sum_{i=1}^{N}(x_i + y_i - z_i) = \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} z_i.$$

定理 2. 在符號 Σ 下的因子若爲常數，則可將其提到符號前邊，當做 Σ 的因子．卽

$$\sum_{i=1}^{N} c x_i = c \sum_{i=1}^{N} x_i.$$

上面二定理可以當做練習題，留待學生自行證明．

定理 3. 在符號 Σ 下的算式若爲常數，則其結果爲 NC.

例 1. $\sum_{i=1}^{N} C = C + C + \cdots$ 至 N 項 $= NC$.

例 2. $\sum_{i=1}^{N}(X_i - C) = \sum_{i=1}^{N} X_i - \sum_{i=1}^{N} C$, 由定理 1;

$$= \sum_{i=1}^{N} X_i - NC,$$ 由本定理例 1.

定理 4. 由 1 起各整數 N 項之和爲 $\frac{1}{2}N(N+1)$. 用符號表之爲

$$\sum_{x=1}^{N} x=\frac{1}{2}N(N+1).$$

因爲 1, 2,…, N 爲一等差級數，其首項爲 1, 公差爲 1, 末項爲 N, 故其和可用等差級數求和之公式求得.

定理 5. 由 1 起各整數 N 項之平方和爲 $\frac{1}{6}N(N+1)(2N+1)$. 用符號表之爲

$$\sum_{x=1}^{N} x^2=\frac{1}{6}N(N+1)(2N+1).$$

證：讓我們利用恆等式 $x^3-(x-1)^3=3x^2-3x+1$. 使 $x=1, 2, 3,\cdots, N$, 並代入此恆等式之兩邊；則有

$$\left.\begin{array}{l} 1^3-0^3 \\ +2^3-1^3 \\ +3^3-2^3 \\ +\cdots\cdots\cdots \\ +N^3-(N-1)^3 \end{array}\right\}=N^3=3\sum_{x=1}^{N} x^2-3\sum_{x=1}^{N} x+N.$$

移項並化簡，我們有

$$\sum_{x=1}^{N} x^2=\frac{N^3-N}{3}+\sum_{x=1}^{N} x=\frac{1}{6}N(N+1)(2N+1). \qquad \text{證訖.}$$

第三節 算術平均數與標準差

1. 定義及基本公式 一組變量之算術平均數爲各變量之和除以變量個數所得之商；一組變量之標準差等於離均差 (deviation from mean, 變量與其平均數* 之差) 自乘平均數之平方根. 故標準差有時

* 在不至混淆情形中，算術平均數可簡稱之爲平均數；此處之平均數即指算術平均數而言，以下仿此。

被稱做由平均數算起之根均方差 (root-mean-square deviation from mean).

現在我們思考尚未分組之各變量,如表 1 所示者然. 設 $\bar{X}$ 表示 N 個變量 $X_1, X_2, \cdots, X_N$ 之平均數, σ_X 表示其標準差. 於是

$$(1)\quad \begin{cases} \bar{X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}X_i, \\ \sigma_X=\sqrt{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(X_i-\bar{X})^2}. \end{cases}$$

如果已給之觀察值中有些相等或相近者,如表 1 裏所列小學生之身長在 99 與 100 公分之間者有 3 人,在 100 與 101 公分之間者有 4 人,在 101 與 102 公分之間者有 13 人等等是. 就一般言之,設有 n 個不同變量或組值 $X_1, X_2, \cdots, X_n$; 並且發現之對應頻數各為 $F_{X_1}, F_{X_2}, \cdots, F_{X_N}$, 或各簡寫做 $F_1, F_2, \cdots, F_n$. 於是

$$(2)\quad \begin{cases} \bar{X}=\dfrac{1}{N}\Sigma XF, \\ \sigma_X=\sqrt{\dfrac{1}{N}\Sigma(X-\bar{X})^2F}; \end{cases}$$

彼處 $N=\Sigma F$.

平均數及標準差之單位與原始觀察值之單位同. 如果原始觀察時所用之單位為公分,則平均數及標準差之單位亦為公分. (學者宜驗證之.)

式 (1) 及 (2) 為計算 $\bar{X}$ 及 σ_X 之基本公式 (fundamental formula). 式 (1) 為式 (2) 之特殊情形,祇要使式 (2) 中之 $F_1=F_2=\cdots=1$, 式 (2) 就會變做式 (1) 的.

茲為說明用式 (1) 及 (2) 計算 $\bar{X}$ 及 σ_X 之手續，並為學者對於該公式得一具體觀念起見，各舉一實例如下：

例 1. 試應用式 (1) 計算表 1 裏所列 147 個兒童身長之平均數及標準差。

表 1. 小學校六歲男生之身長（單位為最末公分*）

105	105	105	109	106	111	105	105
122	111	110	109	113	111	105	111
115	107	101	109	115	109	107	108
99	104	112	108	115	111	117	103
100	106	108	107	109	103	105	113
114	100	97	106	108	105	106	112
122	104	104	105	103	103	116	104
106	101	104	105	99	102	118	
110	103	102	105	105	106	109	
106	104	101	105	106	105	112	
105	107	101	104	111	103	108	
105	116	110	103	109	101	117	
103	105	112	103	108	117	114	
106	106	115	103	102	109	125	
108	107	114	101	102	105	101	
106	103	113	101	101	115	104	
101	115	113	101	105	115	109	
122	104	112	101	110	107	110	
115	103	110	100	101	111	100	
107	104	110	99	109	105	109	

* 表示尺度之法最常用者有三種：第一種為用實在尺度 (exact scale) 表示，第二種為用最近尺度 (nearest scale) 表示，並且第三種為用最末尺度 (last scale) 表示。例如某物之長為 5.2 尺，若係用第一種方法表示者，則該物之長恰為 5.2 尺；若係用第二種方法表示者，則其長在 5.15 尺與 5.25 尺之間；若係用第三種方法表示者，則其長在 5.2 尺與 5.3 尺之間。

應用式 (1) 第一式計算平均數須求表 1 裏各數之和;應用式 (1) 第二式計算標準差,首須求表 1 裏各數值與平均數之差,次須求如此所得各差之平方,最後須求各差平方之和. 因此我們必須列出表 2 及 3, 以便分別計算各數值與其平均數之差 $(X-\bar{X})$ 及各差之平方 $(X-\bar{X})^2$.

因本例 $N=147$, 並由表 1 算得 $\Sigma X=15{,}783$; 於是 $\bar{X}=107.367347$ (用最末尺度表示者).

表 2. $(X-\bar{X})$ 之計算,取 $\bar{X}=107.37$

−2.37	−2.37	−2.37	1.63	−1.37	3.63	−2.37	−2.37
14.63	3.63	2.63	1.63	5.63	3.63	−2.37	3.63
7.63	−0.37	−6.37	1.63	7.63	1.63	−0.37	0.63
−8.37	−3.37	4.63	0.63	7.63	3.63	9.63	−4.37
−7.37	−1.37	0.63	−0.37	1.63	−1.37	−2.37	5.63
6.63	−7.37	−10.37	−1.37	0.63	−2.37	−1.37	4.63
14.63	−3.37	−3.37	−2.37	−4.37	−4.37	8.63	−3.37
−1.37	−6.37	−3.37	−2.37	−8.37	−5.37	10.63	
2.63	−4.37	−5.37	−2.37	−2.37	−1.37	1.63	
−1.37	−3.37	−6.37	−2.37	−1.37	−2.37	4.63	
−2.37	−0.37	−6.37	−3.37	3.63	−4.37	0.63	
−2.37	8.63	2.63	−4.37	1.63	−6.37	9.63	
−4.37	−2.37	4.63	−4.37	0.63	9.63	6.63	
−1.37	−1.37	7.63	−4.37	−5.37	1.63	17.63	
0.63	−0.37	6.63	−6.37	−5.37	−2.37	−6.37	
−1.37	−4.37	5.63	−6.37	−6.37	7.63	−3.37	
−6.37	7.63	5.63	−6.37	−2.37	7.63	1.63	
14.63	−3.37	4.63	−6.37	2.63	−0.37	2.63	
7.63	1.63	2.63	−7.37	−6.37	3.63	−7.37	
−0.37	−3.37	2.63	−8.37	1.63	−2.37	1.63	

表 3. $(X-\bar{X})^2$ 之計算，取 $\bar{X}=107.37$

5.6169	5.6169	5.6169	2.6569	1.8769	13.1769	5.6169	5.6169
214.0369	13.1769	6.9169	2.6569	31.6969	13.1769	5.6169	13.1769
58.2169	0.1369	40.5769	2.6569	58.2169	2.6569	0.1369	0.3969
70.0569	11.3569	21.4369	0.3969	58.2159	13.1769	92.7369	19.0969
54.3169	1.8769	0.3969	0.1369	2.6569	1.8769	5.6169	31.6969
43.9569	54.3169	107.5369	1.8769	0.3969	5.6169	1.[illegible]769	21.4369
214.0369	11.3569	11.3569	5.6169	19.0969	19.0969	74.4769	11.3569
1.8769	40.5769	11.3569	5.6169	70.0569	28.8369	112.9969	
6.9169	19.0969	28.8369	5.6169	5.6169	1.8769	2.6569	
1.8769	11.3569	40.5769	5.6169	1.8769	5.6169	21.4369	
5.6169	0.1369	40.5769	11.3569	13.1769	19.0969	0.3969	
5.6169	74.4769	6.9169	19.0969	2.6569	40.5769	92.7369	
19.0969	5.6169	21.4369	19.0969	0.3969	92.7369	43.9569	
1.8769	1.8769	58.2169	19.0969	28.8369	2.6569	310.8169	
0.3969	0.1369	43.9569	40.5769	28.8369	5.6169	40.5769	
1.8769	19.0969	31.6969	40.5769	40.5769	58.2169	11.3569	
40.5769	58.2169	31.6969	40.5769	5.6169	58.2169	2.6569	
214.0369	11.3569	21.4369	40.5769	6.9169	0.1369	6.9169	
58.2169	2.6569	6.9169	54.3169	40.5769	13.1769	54.3169	
0.1369	11.3569	6.9169	70.0569	2.6569	5.6169	2.6569	

將表 3 裏各數值加之，則得

$$\Sigma(X-\bar{X})^2=4118.1643.$$

於是

$$\frac{1}{N}\Sigma(X-\bar{X})^2=28.014723.$$

取上面已算得之平均數至小數點兩位，即取 $\bar{X}=107.37$ 公分(用最末尺度表示者)，並且計算標準差至小數點四位，我們有欲求之結果

$$\bar{X}=107.87 \text{ 公分}^*,$$

$$\sigma_X=5.2929 \text{ 公分}.$$

* 因為最末尺度表示長度之數值小於用實在尺度所表示者，在本例裏前者小於後者 0.5 公分，故 $\bar{X}=107.37+0.5=107.87$ 公分。
但所得標準差 σ_X 之結果當然仍為 5.2929 公分。(何故?)

例 2. 試應用式(2)計算表 4 首二行所列 147 個兒童身長之平均數及標準差。

爲應用式(2)計算 $\bar{X}$ 及 σ_X，我們宜填算表 4 末三行裏各數值。

表 4. XF 及 $(X-\bar{X})^2F$ 之計算

組段(身長)	頻數 F	中值 X	XF	$(X-\bar{X})$	$(X-\bar{X})^2F$
97 –	1	97.5	97.5	−10.37	107.5369
98 –	...	98.5	...	− 9.37	...
99 –	3	99.5	298.5	− 8.37	210.1707
100 –	4	100.5	402.0	− 7.37	217.2676
101 –	13	101.5	1319.5	− 6.37	527.4997
102 –	4	102.5	410.0	−5.37	115.3476
103 –	10	103.5	1035.0	−4.37	190.9690
104 –	10	104.5	1045.0	−3.37	113.5690
105 –	20	105.5	2110.0	−2.37	112.3380
106 –	12	106.5	1278.0	−1.37	22.5228
107 –	7	107.5	752.5	−0.37	.9583
108 –	7	108.5	759.5	0.63	2.7783
109 –	12	109.5	1314.0	1.63	31.8828
110 –	7	110.5	773.5	2.63	48.4183
111 –	7	111.5	780.5	3.63	92.2383
112 –	5	112.5	562.5	4.63	107.1845
113 –	4	113.5	454.0	5.63	126.7876
114 –	3	114.5	343.5	6.63	131.8707
115 –	8	115.5	924.0	7.63	465.7352
116 –	2	116.5	233.0	8.63	148.9538
117 –	3	117.5	352.5	9.63	278.2107
118 –	1	118.5	118.5	10.63	112.9969
119 –	...	119.5	...	11.63	...
120 –	...	120.5	...	12.63	...
121 –	...	121.5	...	13.63	...
122 –	3	122.5	367.5	14.63	642.1107
123 –	...	123.5	...	15.63	...
124 –	...	124.5	...	16.63	...
125 –	1	125.5	125.5	17.63	310.8169
總計	147	...	15,856.5	...	4118.1643

因爲　$\Sigma XF=15{,}856.5,\ \Sigma(X-\bar{X})^2F=4118.1643.$

故　$\bar{X}=107.87$ 公分，

並且　$\sigma_X=5.2929$ 公分.

此結果與例 1 所得者相同.

由以上二例可見，用基本公式求 $\bar{X}$ 及 σ_X 之值，所用之計算手續非常繁瑣.故在實際運算中宜將基本公式加以種種推演，以期計算手續簡化.

2. 由未分組論據計算 $\bar{X}$ 及 σ_X 之法　由上面例 1 得知，計算標準差之手續所以麻煩者，因平均數常是帶有不盡小數或有時其本身爲不盡小數，當求 $\Sigma(X-\bar{X})^2$ 之值時所用之計算手續極爲麻煩故也.

茲爲避免用不盡小數運算起見，將式 (1) 第二式加以推演如下：

因爲
$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2&=\frac{1}{N}\sum_{i=1}^{N}(X_i^2-2\bar{X}X_i+\bar{X}^2)\\&=\frac{1}{N}\left(\sum_{i=1}^{N}X_i^2-2\bar{X}\sum_{i=1}^{N}X_i+N\bar{X}^2\right)\\&=\frac{1}{N}\left(\sum_{i=1}^{N}X_i^2-2N\bar{X}^2+N\bar{X}^2\right)\\&=\frac{1}{N}\sum_{i=1}^{N}X_i^2-\bar{X}^2.\end{aligned}$$

故式 (1) 變做

$$(3)\qquad\begin{cases}\bar{X}=\nu_{1X},\\ \sigma_X=\sqrt{\nu_{2X}-\nu_{1X}^2};\end{cases}$$

彼處 $\nu_{1X}=\frac{1}{N}\sum_{i=1}^{N}X_i,\ \nu_{2X}=\frac{1}{N}\sum_{i=1}^{N}X_i^2$. 至於 ν 右下角所綴之第一個數碼 1 及 2 表示符號 Σ 下變量 X_i 之次數或 ν 之級數，所綴之第二個大寫

英文字母 X 表示 ν_1 及 ν_2 之値係以原始所用之量度爲單位．因此可以明瞭 σ 右下角所綴之 X 亦係表示 σ 之值以原始量度爲單位．

用式 (3) 由未分組論據求 X 及 σ_X 之法可舉例明之如下：

例 3. 試應用式(3)計算表1裏所列兒童身長之平均數及標準差．

本例觀察值之個數爲 147，即 $N=147$；並且表 1 裏各數值相加之和爲 15,783，即 $\Sigma X=15{,}783$．爲求 ΣX^2 之值須列表 5，該表裏之結果爲由表 1 裏各數值自乘得之者．

表 5. X^2 之計算

11025	11025	11025	11881	11236	12321	11025	11025
14884	12321	12100	11881	12769	12321	11025	12321
13225	11449	10201	11881	13225	11881	11449	11664
9801	10816	12544	11664	13225	12321	13689	10609
10000	11236	11664	11449	11881	11236	11025	12769
12996	10000	9409	11236	11664	11025	11236	12544
14884	10816	10816	11025	10609	10609	13456	10816
11236	10201	10816	11025	9801	10404	13924	
12100	10609	10404	11025	11025	11236	11881	
11236	10816	10201	11025	11236	11025	12544	
11025	11449	10201	10816	12321	10609	11664	
11025	13456	12100	10609	11881	10201	13689	
10609	11025	12544	10609	11664	13689	12996	
11236	11236	13225	10609	10404	11881	15625	
11664	11449	12996	10201	10404	11025	10201	
11236	10609	12769	10201	10201	13225	10816	
10201	13225	12769	10201	11025	13225	11881	
14884	10816	12544	10201	12100	11449	12100	
13225	11881	12100	10000	10201	12321	10000	
11449	10816	12100	9801	11881	11025	11881	

將表 5 裏各數值加之得 $\Sigma X^2=1{,}698{,}697$，

於是

$$\nu_{1X}=107.367347,$$

$$\nu_{2X}=11{,}555.761905,$$

$$\sqrt{\nu_{2X}-\nu_{1X}^2}=5.292892.$$

因之代入式 (2), 我們有

$$\bar{X}=107.87 \text{ 公分},$$

$$\sigma_X=5.2929 \text{ 分公}.$$

此結果與例 1 裏求得者完全相同.

用式(1)計算 $\bar{X}$ 及 σ_X 之法可稱做直接法 (direct method) 之冗長法 (long method); 用式(3)計算之法可稱做直接法之簡便法(short method).

3. 由頻數表計算 $\bar{X}$ 及 σ_X 之第一法 由上面例 2 可見, 計算 $\Sigma(X-\bar{X})^2F$ 之值時, 所用之計算手續, 因 $\bar{X}$ 帶有不盡小數, 仍嫌繁重. 於是將式 (2) 推演一下, 使成下列式 (4) 之形式推演時所用之手續與由式 (1) 化做式 (2) 時所用者同, 學者可試行推演之.

$$(4)\qquad \begin{cases} \bar{X}=\nu_{1X}, \\ \sigma_X=\sqrt{\nu_{2X}-\nu_{1X}^2}. \end{cases}$$

此與式 (3) 之形式完全相同; 不過此時式中之 $\nu_{1X}=\frac{1}{N}\sum_{i=1}^{n}X_iF_i$, $\nu_{2X}=\frac{1}{N}\sum_{i=1}^{n}X_i^2F_i$ 並且 $N=\sum_{i=1}^{n}F_i$ 罷了.

用式(4)由頻數分配表求 $\bar{X}$ 及 σ_X 之法, 可以名之曰間接法(indirect method) 之第一法, 或曰間接法之冗長法.

現在舉例說明第一法之應用於下:

例 4. 試應用第一法並根據表 6 裏所列之論據, 分別求小學校六歲男生身長之平均數及標準差.

為由表 6 求平均身長及身長之標準差, 我們須先計算表 6 第三及第四兩行裏各數值, 以便求得 ΣXF 及 ΣX^2F 之值. 由表 6 我們有 $\Sigma F=147$, $\Sigma XF=15{,}856.5$, $\Sigma X^2F=1{,}714{,}516.75$. 於是

$$\nu_{1X}= \quad 107.867347 \text{ 公分},$$

$$\nu_{2X}=11{,}663.379252 \text{ (公分)}^2,$$

$$\sqrt{\nu_{2X}-\nu_{1X}^2}= \quad 5.292892 \text{ 公分}.$$

因之代入式 (3) 得

$\bar{X}=107.87$ 公分,

$\sigma_X=5.2929$ 公分.

表 6. 應用第一法計算平均數及標準差

中值:X (1)	頻數:F (2)	XF (3)	X^2F (4)
97.5	1	97.5	9,506.25
98.5	...	......	....
99.5	3	298.5	29,700.75
100.5	4	402.0	40,401.00
101.5	13	1,319.5	133,929.25
102.5	4	410.0	42,025.00
103.5	10	1,035.0	107,122.50
104.5	10	1,045.0	109,202.50
105.5	20	2,110.0	222,605.00
106.5	12	1,278.0	136,107.00
107.5	7	752.5	80,893.75
108.5	7	759.5	82,405.75
109.5	12	1,314.0	143.883.00
110.5	7	773.5	85,471.75
111.5	7	780.5	87,025.75
112.5	5	562.5	63,281.25
113.5	4	454.0	51,529.00
114.5	3	343.5	39,330.75
115.5	8	924.0	106,722.00
116.5	2	233.0	27,144.50
117.5	3	352.5	41,418.75
118.5	1	118.5	14,042.25
119.5	...	......	....
120.5	...	......	....
121.5	...	......	....
122.5	3	367.5	45,018.75
123.5	...	......	....
124.5	...	......	....
125.5	1	125.5	15,750.25
總計	147	15,856.5	1,714,516.75

4. 由頻數表計算 $\bar{X}$ 及 σ_X 之第二法　如果 X_0 代表一任意數，h 代表組距，X 代表原始變數，x 代表助變數；則 X 與 x 之關係得寫出之如下：

$$X = X_0 + hx. \tag{5}$$

並且由此關係可以推得

$$\begin{aligned}\nu_{1X} &= \frac{1}{N}\sum_{i=1}^{n}(X_0 + hx_i)f_i \\ &= X_0 + \frac{h}{N}\sum_{i=1}^{n}x_if_i.\end{aligned}$$

$$\begin{aligned}\nu_{2X} &= \frac{1}{N}\Sigma(X_0 + hx_i)^2f_i \\ &= X_0^2 + \frac{2hX_0}{N}\sum_{i=1}^{n}x_if_i + \frac{h^2}{N}\sum_{i=1}^{n}x_i^2f_i,\end{aligned}$$

及

$$\nu_{2X} - \nu_{1X}^2 = \frac{h^2}{N}\sum_{i=1}^{n}x_i^2f_i - \left(\frac{h}{N}\sum_{i=1}^{n}x_if_i\right)^2.$$

因之式 (4) 變做

$$\begin{cases}\bar{X} = X_0 + h\nu_{1x}, \\ \sigma_X = h\sqrt{\nu_{2x} - \nu_{1x}^2};\end{cases} \tag{6}$$

彼處 $\nu_{1x} = \frac{1}{N}\sum_{i=1}^{n}x_if_i$，$\nu_{2x} = \frac{1}{N}\sum_{i=1}^{n}x_i^2f_i$，$N = \sum_{i=1}^{n}f_i$，$f_i$ 爲 f_{x_i} 之省略寫法，並且 $f_{x_i} = F_{X_0+hx_i}$. 至於 ν 之右下角所綴之第二個小寫英文字母 x 表示各 ν 之值以組距爲單位。

用式 (6) 由頻數分配求 $\bar{X}$ 及 σ_X 之法，可以名之曰間接法之第二法，或稱之曰簡便法.

第二法之應用可用下例說明之：

例 5. 試應用第二法並根據表 6 裏所列之論據，分別求小學校六歲男生身長之平均數及標準差。

表 7. 應用第二法計算平均數及標準差.

組值：X (1)	頻數：f (2)	助變數：x (3)	xf (4)	x^2f (5)
97.5	1	−10	−10	100
98.5	...	−9	...	...
99.5	3	−8	−24	192
100.5	4	−7	−28	196
101.5	13	−6	−78	468
102.5	4	−5	−20	100
103.5	10	−4	−40	160
104.5	10	−3	−30	90
105.5	20	−2	−40	80
106.5	12	−1	−12	12
107.5	7	0	0	0
108.5	7	1	7	7
109.5	12	2	24	48
110.5	7	3	21	63
111.5	7	4	28	112
112.5	5	5	25	125
113.5	4	6	24	144
114.5	3	7	21	147
115.5	8	8	64	512
116.5	2	9	18	162
117.5	3	10	30	300
118.5	1	11	11	121
119.5	...	12	...	...
120.5	...	13	...	...
121.5	...	14	...	...
122.5	3	15	45	675
123.5	...	16	...	...
124.5	...	17	...	...
125.5	1	18	18	324
總 計	147	...	54	4,138

應用公式 (6) 計算平均數及標準差，我們須先列表 7，填寫該表第三行裏各數值，並須計算第四，第五兩行裏數值，以求 Σxf 及 Σx^2f 之值．第三行各數值可逕行填寫之，* 即以任意組爲 0 點，由此向上各組依次填以 -1，-2，…等數，向下各組依次填以 1，2，…等數．有了第三行裏各數值，便可算得第四，第五兩行裏各數值．由表 7 我們有：$\Sigma xf=54$，$\Sigma x^2f=4{,}138$．因之將此等值代入式 (6)，並且不要忘記 $h=1$ 公分，$X_0=107.5$ 公分，$N=147$；則得

$$\nu_{1x}=0.367347,$$

$$\nu_{2x}=28.149660,$$

$$\sqrt{\nu_{2x}-\nu_{1x}^2}=5.292892.$$

故
$$\bar{X}=107.87\text{ 公分．}$$

$$\sigma_X=5.2929\text{ 公分．}$$

由例 4 及 5 可見，由頻數表計算平均數及標準差用第一法不如用第二法省事，故實際由已分組論據——頻數分配表——計算 $\bar{X}$ 及 σ_X 用式 (6) 而不用式 (4)．

5. 標準差之功用　假設變數 X 依常態型分配，其變量有 N 個，並且已算得其平均數 $\bar{X}$．於是 X 之分配情形可由已算得之標準差 σ_X 推知之．換言之，由 N，$\bar{X}$ 及 σ_X，可將 X 之頻數分配繁衍出來，X 之分配圖形可以描繪出來，此點將在第四章第十五節詳論之．現在我們僅將標準差之功用，根據上面例 1 裏算得之結果，解釋之如下：

---

* 實在說，該行裏各數值係由式 (5) 算得者．在本例中 $h=1$ 公分，以第十一組爲原點（卽 $X_0=107.5$ 公分），依次將 97.5，98.5，…，125.5 代入式 (5)，則得 -10，-9，…，18．

在上面例 1 裏曾經算得平均數及標準差各爲

$$\bar{X}=107.87 \text{ 公分},$$

$$\sigma_X=5.2929 \text{ 公分}.$$

所以

$\bar{X}-0.5\,\sigma_X=105.22$ 公分	$\bar{X}+0.5\,\sigma_X=110.52$ 公分
$\bar{X}-1.0\,\sigma_X=102.58$ 公分	$\bar{X}+1.0\,\sigma_X=113.16$ 公分
$\bar{X}-1.5\,\sigma_X=99.93$ 公分	$\bar{X}+1.5\,\sigma_X=115.81$ 公分
$\bar{X}-2.0\,\sigma_X=97.28$ 公分	$\bar{X}+2.0\,\sigma_X=118.46$ 公分
$\bar{X}-2.5\,\sigma_X=94.64$ 公分	$\bar{X}+2.5\,\sigma_X=121.10$ 公分
$\bar{X}-3.0\,\sigma_X=91.99$ 公分	$\bar{X}+3.0\,\sigma_X=123.75$ 公分
$\bar{X}-3.5\,\sigma_X=89.34$ 公分	$\bar{X}+3.5\,\sigma_X=126.40$ 公分
$\bar{X}-4.0\,\sigma_X=86.70$ 公分	$\bar{X}+4.0\,\sigma_X=129.04$ 公分

就理論說，任一變量在 $(\bar{X}-0.5\,\sigma_X)$ 與 $(\bar{X}+0.5\,\sigma_X)$ 之間的機會約爲 38%，在 $(\bar{X}-1.0\,\sigma_X)$ 與 $(\bar{X}+1.0\,\sigma_X)$ 之間的機會約爲 68%，如此等等(見附錄 III 表 B)。吾應要問：所得之觀察機會與算得之機會相隔多遠呢？爲回答此問題，我們檢查表 1 裏各變量：在 105.22 公分與 110.52 公分之間者有若干個，其佔總個數之百分數爲何；在 102.58 公分與 113.16 公分之間者有若干個，其佔總個數之百分數爲何；如此等等。茲將各變量在指定之組段內(各組段之組距均係 $0.5\,\sigma_X$)之頻數，及其佔總個數之百分數(即相對頻數)詳列於表 8，並將由表 B 查得之理論相對頻數亦行列入，俾資比較。

表 8 第二行所列之觀察頻數，如 56，係用簡單算學求得者。由表 I 裏查得之變量在 105 最末公分與 110 最末公分之間者有 65 項，其確爲

105 最末公分者有 20 項，確爲 110 最末公分者有 7 項。因是 20 項中小於 105.22 公分者應爲 6 項，7 項中大於 110.52 公分者應爲 3 項。故表 1 裏之各變量在 105.22 公分與 110.52 公分之間者應有 56 項，即由 65 減以 6 及 3 所得之差。

表 8. 在指定組段內之觀察頻數及觀察與理論之相對頻數。

指定之組段	觀察之頻數	相對頻數(%)	
		觀察者	理論者
在 $\bar{X}\pm0.5\,\sigma_X$ 之間	56	38.1	38.3
在 $\bar{X}\pm1.0\,\sigma_X$ 之間	100	68.0	68.2
在 $\bar{X}\pm1.5\,\sigma_X$ 之間	132	89.8	86.6
在 $\bar{X}\pm2.0\,\sigma_X$ 之間	142	96.6	95.4
在 $\bar{X}\pm2.5\,\sigma_X$ 之間	143	97.3	98.8
在 $\bar{X}\pm3.0\,\sigma_X$ 之間	146	99.3	99.7
在 $\bar{X}\pm3.5\,\sigma_X$ 之間	147	100.0	100.0
在 $\bar{X}\pm4.0\,\sigma_X$ 之間	147	100.0	100.0

由上表末兩行可見，各指定組段內觀察與理論之相對頻數，雖未能完全契合，但出入甚微。因是觀察之相對頻數若以理論者代替之，不至發生不合理之差誤。此種事實不獨在本例裏爲然，任何頻數分配，祇要其爲常態型或近於常態型，皆有同樣情形。

由於上面表演之事實得以判斷，由表 1 裏任取一變量，其落在 $(\bar{X}-0.5\,\sigma_X)$ 與 $(\bar{X}+0.5\,\sigma_X)$ 之間，落在 $(\bar{X}-1.0\,\sigma_X)$ 與 $(\bar{X}+1.0\,\sigma_X)$ 之間，…的機會。明確言之，一百個變量中落在 $(\bar{X}-0.5\,\sigma_X)$ 與 $(\bar{X}+0.5\,\sigma_X)$ 之間，$(\bar{X}-1.0\,\sigma_X)$ 與 $(\bar{X}+1.0\,\sigma_X)$ 之間，… 的機會各約爲 38.3%，68.3%，等等。

6. 平均數之性質 關於平均數之性質，可用下列之定理敍述之：

定理 6. 離均差之代數和等於零.

證：讓 x_i 代表第 i 個變量 X_i 與平均數 $\bar{X}$ 之差。於是

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N}(X_i - \bar{X})$$

$$= \sum_{i=1}^{N} X_i - N\bar{X}.$$

記得 $\frac{1}{N}\sum_{i=1}^{N} X_i = \bar{X}$，即 $\sum_{i=1}^{N} X_i = N\bar{X}$；所以

$$\sum_{i=1}^{N} x_i = N\bar{X} - N\bar{X} = 0.$$ 證訖.

定理 7. 各離均差之平方和必較各任意離差（變量與一任意數之差）之平方和爲小.

證：設 A 爲一任意數，d_i 爲第 i 個變量 X_i 與 A 之差，並且 $R = \sum_{i=1}^{n} d_i^2$. 這個定理說，R 之值以當 $A = \bar{X}$ 時爲最小.

由微積分學上之定理得知，R 值最小之必要及充分條件爲

$$\frac{dR}{dA} = 0 \text{ 及 } \frac{d^2R}{dA^2} > 0.$$

現在 $$R = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n}(X_i - A)^2.$$

故 R 之第一次及第二次微係數各爲

$$\frac{dR}{dA} = -2\sum_{i=1}^{n}(X_i - A),$$

$$\frac{d^2R}{dA^2} = 2N.$$

由此可見：當 $\frac{dR}{dA}=0$ 時，則 $A=\bar{X}$；並且因為 N 為正數，$\frac{d^2R}{dA^2}$ 之值永遠為正．於是得知 R 之值以當 $A=\bar{X}$ 時為最小． 證訖．

系． 標準差為根均方差中之最小者．

定理 8． 如果一組 N_1 個變量之平均數為 $\bar{X}_1$，另一組 N_2 個變量之平均數為 $\bar{X}_2$；則二組合計之平均數

$$\bar{X}=\frac{N_1\bar{X}_1+N_2\bar{X}_2}{N_1+N_2}. \tag{7}$$

證： 由定義
$$\bar{X}=\frac{\sum_{i=1}^{N_1}X_{1i}+\sum_{i=1}^{N_2}X_{2i}}{N_1+N_2},$$

$$\bar{X}_1=\frac{1}{N_1}\sum_{i=1}^{N_1}X_{1i},$$

並且
$$\bar{X}_2=\frac{1}{N_2}\sum_{i=1}^{N_2}X_{2i}.$$

所以
$$\bar{X}=\frac{N_1\bar{X}_1+N_2\bar{X}_2}{N_1+N_2}.$$
證訖．

若將此定理普遍化，則有下列之定理．

定理 9． 如果有 N 個變量，係由 k 個小組合併而成；則其平均數

$$\bar{X}=\frac{1}{N}\sum_{i=1}^{k}N_i\bar{X}_i, \tag{8}$$

彼處
$$\bar{X}_i=\frac{1}{N_i}\sum_{j=1}^{N_i}X_{ij},\ N=\sum_{j=1}^{k}N_j.$$

系． 設 $N_i=N_1$，$i=1,2,\cdots,k$；則 $N=kN_1$，因此本定理之結果變做

$$\bar{X}=\frac{1}{k}\sum_{i=1}^{k}\bar{X}_i. \tag{9}$$

此與式 (1) 第一式之形式相同．

7. 標準差之性質 關於標準差之性質，既有用處，且饒趣味，因敘述之並證明之如下：

定理 10. 對於重心 P 之二級動差 (moment of second order)* 等於對於任意點 Q 之二級動差減以 P, Q 二點距離平方之差。

此定理用符號敘述之或較清楚些。假設 $\bar{X}$ 爲 N 個變量之平均數，並且 X_0 爲一任意數值。則 P 點之坐標爲 $\bar{X}$，並且 Q 點之坐標爲 X_0。如此這個定理用算學符號表之爲

$$(10)\qquad \frac{1}{N}\Sigma(X-\bar{X})^2=\frac{1}{N}\Sigma(X-X_0)^2-(\bar{X}-X_0)^2.$$

證：因爲

$$(X-\bar{X})=(X-X_0)-(\bar{X}-X_0)$$

於是

$$\frac{1}{N}\Sigma(X-\bar{X})^2=\frac{1}{N}\Sigma\{(X-X_0)-(\bar{X}-X_0)\}^2$$

$$=\frac{1}{N}\Sigma(X-X_0)^2-\frac{2}{N}\Sigma(X-X_0)(\bar{X}-X_0)+(\bar{X}-X_0)^2$$

$$=\frac{1}{N}\Sigma(X-X_0)^2-2(\bar{X}-X_0)^2+(\bar{X}-X_0)^2$$

$$=\frac{1}{N}\Sigma(X-X_0)^2-(\bar{X}-X_0)^2.$$ 證訖。

定理 11. 假設 $X_{1i}(i=1, 2,\cdots, N_1)$ 代表 N_1 個變量，$X_{2i}(i=1, 2,\cdots, N_2)$ 代表另外 N_2 個變量，$\bar{X}$ 代表合計之平均數。於是合計之標準差的平方 $\sigma_X{}^2$，可由下式求之：

$$(11)\qquad N\sigma_X{}^2=\sum_{i=1}^{N_1}(X_{1i}-\bar{X})^2+\sum_{i=1}^{N_2}(X_{2i}-\bar{X})^2,$$

彼處 $N=N_1+N_2$。

* 動差之定義見下章第一節。

證：因為

$$\sum_{i=1}^{N_1+N_2}(X_i-\bar{X})^2=\sum_{i=1}^{N_1}(X_{1i}-\bar{X})^2+\sum_{i=1}^{N_2}(X_{2i}-\bar{X})^2,$$

即是
$$N\sigma_X{}^2=\sum_{i=1}^{N_1}(X_{1i}-\bar{X})^2+\sum_{i=1}^{N_2}(X_{2i}-\bar{X})^2.$$ 證訖。

定理 12. 讓 N_1, $\bar{X}_1$ 及 σ_1 各代一組變量之頻數總和，平均數及標準差，N_2, $\bar{X}_2$ 及 σ_2 各代表另一組變量之頻數總和，平均數及標準差。於是合計之變差 (variance, 即標準差之平方，其正式定義見下章第一節) σ^2 可由下式求之：

$$N\sigma^2=N_1\sigma_1{}^2+N_2\sigma_2{}^2+N_1d_1{}^2+N_2d_2{}^2.$$

彼處 $N=N_1+N_2$, $d_1=\bar{X}_1-\bar{X}$, $d_2=\bar{X}_2-\bar{X}$; 並且 $\bar{X}$ 為合計之平均數。

證：就第一組變量說，$\bar{X}$ 可視做任意點。如此，由定理 10，我們有

$$\frac{1}{N_1}\sum_{i=1}^{N_1}(X_{1i}-\bar{X}_1)^2=\frac{1}{N_1}\sum_{i=1}^{N_1}(X_{1i}-\bar{X})^2-(\bar{X}_1-\bar{X})^2,$$

即是
$$N_1\sigma_1{}^2=\sum_{i=1}^{N_1}(X_{1i}-\bar{X})^2-N_1d_1{}^2;$$

同樣，就第二組變量說，我們有

$$N_2\sigma_2{}^2=\sum_{i=1}^{N_2}(X_{2i}-\bar{X})^2-N_2d_2{}^2.$$

於是將末二式相加得

$$N_1\sigma_1{}^2+N_2\sigma_2{}^2=N\sigma^2-N_1d_1{}^2-N_2d_2{}^2,$$ 由定理 11

移項得

(12)
$$N\sigma^2=N_1\sigma_1{}^2+N_2\sigma_2{}^2+N_1d_1{}^2+N_2d_2{}^2.$$ 證訖。

若將 k 組變量合而為一，則上面定理可推廣之如下：

$$(13)\qquad N\sigma^2=\sum_{i=1}^{k}N_i\sigma_i^2+\sum_{i=1}^{k}N_id_i^2,$$

彼處 $N=\sum_{i=1}^{k}N_i$, 並且 $d_i=\bar{X}_i-\bar{X}$. 很顯然的, $\frac{1}{N}\sum_{i=1}^{k}N_id_i^2$ 爲各組平均數之變差,如此我們得到一重要關係:

$$(14)\qquad \sigma^2=\frac{1}{N}\sum_{i=1}^{k}N_i\sigma_i^2+\sigma^2_{\bar{X}_i}$$

此示總變差 (total variance) 等於兩個變差之和: 一爲各組變差之平均數,另一爲各組平均數之變差.

系 1. 方程式 (12) 可以寫成下列之形式:

$$(15)\qquad N\sigma^2=N_1(\sigma_1^2+\bar{X}_1^2)+N_2(\sigma_2^2+\bar{X}_2^2)-N\bar{X}^2.$$

證: 因爲

$$\begin{aligned}N_1d_1^2&=N_1(\bar{X}_1-\bar{X})^2\\&=N_1\bar{X}_1^2-(2N_1\bar{X}_1\bar{X}-N_1\bar{X}^2),\\N_2d_2^2&=N_2(\bar{X}_2-\bar{X})^2\\&=N_2\bar{X}_2^2-(2N_2\bar{X}_2\bar{X}-N_2\bar{X}^2),\end{aligned}$$

並且

$$\begin{aligned}(2N_1\bar{X}_1\bar{X}-N_1\bar{X}^2)&+(2N_2\bar{X}_2\bar{X}-N_2\bar{X}^2)\\&=2\bar{X}(N_1\bar{X}_1+N_2\bar{X}_2)-\bar{X}^2(N_1+N_2)\\&=2N\bar{X}^2-N\bar{X}^2=N\bar{X}^2,\end{aligned}$$

所以由式 (12) 得

$$\begin{aligned}N\sigma^2&=N_1\sigma_1^2+N_2\sigma_2^2+N_1\bar{X}_1^2+N_2\bar{X}_2^2-N\bar{X}^2\\&=N_1(\sigma_1^2+\bar{X}_1^2)+N_2(\sigma_2^2+\bar{X}_2^2)-N\bar{X}^2.\end{aligned}$$ 證訖.

就一般言之,若有 k 組變量,可將本定理推廣之如下:

(16)
$$N\sigma^2=\sum_1^k N_i(\sigma_i^2+\bar{X}_i^2)-N\bar{X}.$$

系 2. 方程式 (12) 仍然可以寫做下列之形式:

(17)
$$N\sigma^2=N_1\sigma_1^2+N_2\sigma_2^2+\frac{N_1N_2}{N}(\bar{X}_1-\bar{X}_2)^2.$$

此系之證明係證明

$$N_1d_1^2+N_2d_2^2=\frac{N_1N_2}{N}(\bar{X}_1-\bar{X}_2)^2.$$

留待學生試證之.

定理 13. 設有 k 組變量,每組之變差係以總平均數 $\bar{X}$ 爲原點,並設 $\mu_2^{(i)}$ 代表第 i 組變量之變差.於是合計之變差

(18)
$$\sigma_X^2=\frac{1}{N}\{N_1\mu_2^{(1)}+N_2\mu_2^{(2)}+\cdots+N_k\mu_2^{(k)}\}$$

彼處 N_i 代表第 i 組之頻數總和,並且 $\sum_{i=1}^k N_i=N.$

證: 我們可將式 (11) 寫成下形:

$$N\sigma_X^2=N_1\mu_2^{(1)}+N_2\mu_2^{(2)}.$$

如此將其一般化,可得式 (18).

下面定理爲求首 N 個整數標準差之法,即當 $X_1=1,\ X_2=2,\ \cdots,\ X_N=N$ 時求標準差之法. 當觀察值不以量度而以等級記載,求其標準差時,此定理是很有用處的.

定理 14. 首 N 個自然數之標準差

(19)
$$\sigma_X=\sqrt{\frac{N^2-1}{12}}.$$

證: 由式 (3) 第二式得

$$\sigma_X^2=\frac{1}{N}\sum_{i=1}^{N}X_i^2-\left(\frac{1}{N}\sum_{i=1}^{N}X_i\right)^2.$$

使 $X_i=i,\ i=1,2,\cdots,N$; 則

$$\sigma_X^2=\frac{1}{6}(N+1)(2N+1)-\frac{1}{4}(N+1)^2$$

$$=\frac{N^2-1}{12}$$

將末式兩邊開平方，即得欲證明之式 (19).

第四節　中位數與平均差

1. 定義及基本公式　中位數者乃是一個數值，在 N 個觀察值中比牠大的和比牠小的各佔有一半. 此定義稍嫌籠統，比較精確一點的可敍述之如下:

讓 $X_1, X_2, \cdots X_N$ 爲已按其大小次序排列過的 N 過觀察值，如此 $X_1\leq X_2\leq\cdots\leq X_N$; 並讓 $\tilde{X}$ 代表各觀察值之中位數. 於是

(20)　　$\tilde{X}=X_{(N+1)/2}$, 如果 N 爲奇數;

(21)*　　$\tilde{X}$ 爲在 $X_{N/2}$ 與 $X_{N/2}+1$ 中間之數，如果 N 爲偶數;

(22)　在頻數分配情形中，$\tilde{X}$ 爲在積分或累計頻數圖上與 $\frac{N}{2}$ 成對應之 X 值.

用離均差之平均數度量各觀察值對其集中常數之離勢，亦覺合理. 平均差即其一例. 平均差者，爲各觀察值與其平均數相減絕對值之平均數，通常以 MD 表之. 取離均差絕對值之理由甚爲簡單，因爲不如是各離均差之和將爲零了 (見上節定理 6), 如是 MD 亦將等於零了. 平均差之定義用算學符號表之爲

* 當 N 爲偶數時，普通取 $(X_{N/2}+X_{N/2+1})\div 2$ 當做中位之定義.

(23) $$MD=\frac{1}{N}\sum_{i=1}^{N}|X_i-\tilde{X}|,\ \text{或者}=\frac{1}{N}\sum_{i=1}^{n}|X_i-\tilde{X}|F_i.$$

茲分別舉例說明各公式之應用如下:

例 1. 試根據表 1 裏所列之論據,求各觀察值之中位數及其平均差.

爲求表 1 裏小學校六歲男生身長之中位數,我們須將這些學生之身長按其高矮次序排列一下,如表 9 所示.因爲 $N=147$ (其爲奇數),故由表得知 $X_{(N+1)/2}=X_{74}=106.5$ 公分,並且在表 10 裏算得 $\Sigma|X_i-\tilde{X}|=625.77$. 於是分別代入式 (20) 乃 (23), 我們有

$$\tilde{X}=106.5\text{ 公分};$$

$$MD=625.77\div147=4.26\text{ 公分}.$$

表 9. 中位數之計算.小學校六歲男生身長曾按其高矮次序排列過的

(論據見表 1)

97	101	104	105	107	109	112	117
99	102	104	105	107	109	112	117
99	102	104	105	107	109	113	118
99	102	104	105	107	110	113	122
100	102	104	105	108	110	113	122
100	103	105	106	108	110	113	122
100	103	105	106	108	110	114	125
100	103	105	106	108	110	114	
101	103	105	106	108	110	114	
101	103	105	106	108	110	115	
101	103	105	106	108	111	115	
101	103	105	106	109	111	115	
101	103	105	106	109	111	115	
101	103	105	106	109	111	115	
101	103	105	106	109	111	115	
101	104	105	106	109	111	115	
101	104	105	106	109	111	115	
101	104	105	107	109	112	116	
101	104	105	107	109	112	116	
101	104	105	107	109	112	117	

表 10. 平均差之計算. 表 9 裏各數減以平均數 (=107.37 公分) 之絕對值(論據見表 1)

10.37	6.37	3.37	2.37	0.37	1.63	4.63	9.63
8.37	5.37	3.37	2.37	0.37	1.63	4.63	9.63
8.37	5.37	3.37	2.37	0.37	1.63	5.63	10.63
8.37	5.37	3.37	2.37	0.37	2.63	5.63	14.63
7.37	5.37	3.37	2.37	0.63	2.63	5.63	14.63
7.37	4.37	2.37	1.37	0.63	2.63	5.63	14.63
7.37	4.37	2.37	1.37	0.63	2.63	6.63	17.63
7.37	4.37	2.37	1.37	0.63	2.63	6.63	
6.37	4.37	2.37	1.37	0.63	2.63	6.63	
6.37	4.37	2.37	1.37	0.63	2.63	7.63	
6.37	4.37	2.37	1.37	0.63	3.63	7.63	
6.37	4.37	2.37	1.37	1.63	3.63	7.63	
6.37	4.37	2.37	1.37	1.63	3.63	7.63	
6.37	4.37	2.37	1.37	1.63	3.63	7.63	
6.37	4.37	2.37	1.37	1.63	3.63	7,63	
6.37	3.37	2.37	1.37	1.63	3.63	7.63	
6.37	3.37	2.37	1.37	1.63	3.63	7.63	
6.37	3.37	2.37	0.37	1.63	4.63	8.63	
6.37	3.37	2.37	0.37	1.63	4.63	8.63	
6.37	3.37	2.37	0.37	1.63	4.63	9.63	

例 2. 求數值 10, 6, 5, 25, 15, 18, 37 及 20 之中位數及其平均差.

將已給數值之次序排列之,則有 5, 6, 10, 15, 18, 20, 25 及 37. 因為 $N=8$ (其為偶數),並且 $\bar{X}=17$; 故由式 (21) 及 (23) 得知

$$\check{X} \text{ 為一任意數在 } 15 \text{ 與 } 18 \text{ 之間,}$$

並且
$$MD=[(17-5)+(17-6)+\cdots+(37-17)]\div 8$$
$$=64\div 8=8.$$

2. 由頻數表求中位數及平均差 在頻數分配表情形中,中位數及平均差可分別用下式求得之:

(24) $$\tilde{X}=a+\frac{h}{c}\left(\frac{N}{2}-b\right)$$

(25) $$MD=\frac{h}{N}\left[\sum_{i=1}^{n}|x_if_i|+\frac{\bar{X}-X_r}{h}(2N_1-N)\right],$$

彼處 a 爲中位數所在組段之低限，b 爲小於 a 各變量發現之頻數，c 爲中位數所在組段之頻數，n 爲頻數分配之組數，N 爲各頻數之總和，N_1 爲組值小於平均數 $\bar{X}$ 各組發現頻數之和，X_r 爲恰小於 $\bar{X}$ 之組值，其餘各文字代表之意義同前。

式 (24) 及 (25) 可分別證明之如下：

證式 (24)：設下圖中之 X 代表一任意變數——自變數，Y 代表

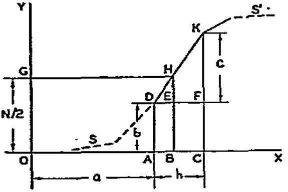


圖 3. 肩形圖。

在 X 以下之頻數——倚變數；$SDHKS'$ 爲一肩形圖；GH 及 DF 均平行於 OX；AD, BH 及 CK 均平行於 OY；OG 或 $BH=N/2$ (頻數總和之半)；$OA=a$, $AC=h$, $AD=b$, $FK=c$. 於是

$$DE:DF=EH:FK, \text{因 } \triangle DEH \text{ 與 } \triangle DFK \text{ 相似}$$

但 $$DE=AB=OB-OA=\tilde{X}-a,$$

$$DF=AC=h,$$

$$EH = BH - BE = BH - AD = N/2 - b,$$

並且 $FK = c.$

所以 $\bar{X} - a : h = N/2 - b : b,$

即 $\bar{X} = a + \frac{h}{c}\left(\frac{N}{2} - b\right).$ 證訖。

證式 (25): 因為

$$MD = \frac{1}{N}\left[\sum_{i=1}^{r}(\bar{X} - X_i)f_i + \sum_{i=r+1}^{n}(X_i - \bar{X})f_i\right]$$

$$= \frac{1}{N}\left[\sum_{i=1}^{r}(\bar{X} - X_r - hx_i)f_i + \sum_{i=r+1}^{n}(X_r + hx_i - \bar{X})f_i\right]$$

(因使 $X = X_r + hx$, $F_{X_r} + hx = f_x$)

$$= \frac{1}{N}\left[h\left(\sum_{i=r+1}^{n}x_i f_i - \sum_{i=1}^{r}x_i f_i\right) + (\bar{X} - X_r)(N_1 - \overline{N - N_1})\right]$$

(因使 $\sum_{i=1}^{r} f_i = N_1$, 自然 $\sum_{i=r+1}^{n} f_i = N - N_1$)

$$= \frac{h}{N}\left[\sum_{i=1}^{n}|x_i f_i| + \frac{\bar{X} - X_r}{h}(2N_1 - N)\right]$$

(因 $x_1, x_2, \cdots, r-1$ 均為負). 證訖。

茲舉例以明公式 (24) 及 (25) 之應用，如下：

例 3. 試根據表 4 裏所列之論據，計算中位數及平均差。

本例 $N = 147$, $a = 106.0$ 公分, $b = 65$ (見下表第四行附有°之累計頻數), $c = 12$ (見下表第三行附有*之頻數), $h = 1$ 公分, $\bar{X} = 107.87$ 公分(見前節例 1), $X_r = 107.5$ 公分 (見下表第二行), $N_1 = 84$ (見下表第四行附有+之累計頻數), $\Sigma|xf| = 618$ (見下表第六行總計). 將此等數分別代入式 (24) 及 (25), 我們有

$$\tilde{X}=106.0+\frac{73.5-65}{12}=106.71\text{ 公分},$$

$$MD=\frac{1}{147}\times[618+0.37\times(84\times2-147)]=4.26\text{ 公分}$$

表 11. 中位數及平均差之計算 (論據見表 4)

組段 (1)	組值:X (2)	頻數:f (3)	累計頻數:Σf (4)	助變數:x (5)	$\|xf\|$ (6)
97 –	97.5	1	1	–10	10
98 –	98.5	...	...	–9	...
99 –	99.5	3	4	–8	24
100 –	100.5	4	8	–7	28
101 –	101.5	13	21	–6	78
102 –	102.5	4	25	–5	20
103 –	103.5	10	35	–4	40
104 –	104.5	10	45	–3	30
105 –	105.5	20	65*	–2	40
106 –	106.5	12*	77	–1	12
107 –	X_r=107.5	7	84+	0	0
108 –	108.5	7	91	1	7
109 –	109.5	12	103	2	24
110 –	110.5	7	110	3	21
111 –	111.5	7	117	4	28
112 –	112.5	5	122	5	25
113 –	113.5	4	126	6	24
114 –	114.5	3	129	7	21
115 –	115.5	8	137	8	64
116 –	116.5	2	139	9	18
117 –	117.5	3	142	10	30
118 –	118.5	1	143	11	11
119 –	119.5	...	...	12	...
120 –	120.5	...	...	13	...
121 –	121.5	...	...	14	...
122 –	122.5	3	146	15	45
123 –	123.5	...	...	16	...
124 –	124.5	...	...	17	...
125 –	125.5	1	147	18	18
總計	...	147	...	...	618

3. 平均差之性質 關於平均差之性質可用下面定理敍述之：

定理 15. 各離中差(各變量與其中位數之差)絕對值之和必不大於各任意離差絕對值之和。

證：此定理可分兩種情形證明之。

a. N 為奇數 設有 N 個觀察值 $X_i=OL_i$ $(i=1, 2, \cdots, 2m+1)$, OL_{m+1} 依定義為中位數(參閱圖 $4a$). 於是對於 L_s 及 L_{s+1} 離差絕對值

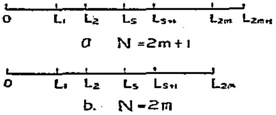


圖 4. N 個觀察值之位置。

之和 Δ_s 及 Δ_{s+1} 各為

$$\Delta_s=L_1L_s+L_2L_s+\cdots+L_sL_s+L_sL_{s+1}+\cdots+L_sL_{2m}+L_sL_{2m+1}.$$

$$\begin{aligned}\Delta_{s+1}&=L_1L_{s+1}+L_2L_{s+1}+\cdots+L_sL_{s+1}+L_{s+1}L_{s+1}+\cdots\\&\qquad+L_{s+1}L_{2m}+L_{s+1}L_{2m+1}\\&=L_1L_s+L_2L_s+\cdots+L_sL_s+L_sL_{s+1}+\cdots+L_sL_{2m}+L_sL_{2m+1}\\&\quad+L_sL_{s+1}+L_sL_{s+1}+\cdots+L_sL_{s+1}+L_{s+1}L_s+\cdots\\&\qquad+L_{s+1}L_s+L_{s+1}L_s.\\&=\Delta_s+(s-\overline{2m+1-s})L_sL_{s+1},\ \text{因}\ L_sL_{s+1}=-L_{s+1}L_s\\&=\Delta_s+(2s-\overline{2m+1})L_sL_{s+1}.\end{aligned}$$

由是知：若 s 之值大於 $(2m+1)/2$, 則 $\Delta_{s+1}>\Delta_s$; 若 s 之值小於 $(2m+1)/2$, 則 $\Delta_{s+1}<\Delta_s$. 就是說：分別使 $s=m+1, m+2, \cdots, 2m+1$, 可

得 $\Delta_{m+1}<\Delta_{m+2}<\cdots<\Delta_{2m+1}$ 之結果;並且分別使 $s=m, m-1,\cdots, 1$, 可得 $\Delta_{m+1}<\Delta_m<\cdots<\Delta_1$ 之結果.換句話說, Δ_{m+1} 爲各 Δ 中之最小者.不要忘記, Δ_{m+1} 爲對於中位數離差絕對值之和.此卽在 N 爲奇數情形中

$$\Delta_{m+1}<\Delta_i(i=1, 2,\cdots, m, m+2,\cdots, 2m+1),$$

亦卽離中差絕對值之和較任意離差絕對值之和爲小.

b. N 爲偶數　設有 N 個觀察值 $X_i=OL_i$ $(i=1, 2,\cdots, 2m)$. 於是和先前一樣,對於 L_s 及 L_{s+1} 各離差絕對值之和 Δ_s 及 Δ_{s+1} 各爲

$$\Delta_s=L_1L_s+L_2L_s+\cdots+L_sL_s+L_sL_{s+1}+\cdots+L_sL_{2m}.$$

$$\begin{aligned}\Delta_{s+1}&=L_1L_{s+1}+L_2L_{s+1}+\cdots+L_sL_{s+1}+L_{s+1}L_{s+1}+\cdots+L_{s+1}L_{2m}\\&=L_1L_s+L_2L_s+\cdots+L_sL_s+L_sL_{s+1}+\cdots+L_sL_{2m}+L_sL_{s+1}\\&\qquad+L_sL_{s+1}+\cdots+L_sL_{s+1}+L_{s+1}L_s+\cdots+L_{s+1}L_s\\&=\Delta_s+(s-\overline{2m-s})L_sL_{s+1},\quad 因\ L_sL_{s+1}=-L_{s+1}L_s\\&=\Delta_s+2(s-m)L_sL_{s+1}.\end{aligned}$$

由是知:若 s 之值大於 m, 則 $\Delta_{s+1}>\Delta_s$; 若 s 之值等於 m, 則 $\Delta_{s+1}=\Delta_s$; 若 s 之值小於 m, 則 $\Delta_{s+1}<\Delta_s$. 就是說:分別使 $s=m+1, m+2,\cdots, 2m$, 可得 $\Delta_{m+1}<\Delta_{m+2}<\cdots<\Delta_{2m}$ 之結果;使 $s=m$, 可得 $\Delta_{m+1}=\Delta_m$ 之結果;並且使 $s=m-1, m-2, \cdots, 1$, 可得$\Delta_m<\Delta_{m-1}<\cdots<\Delta_1$ 之結果.再設 Δ_θ 代表各觀察值對於 L_θ 點離差絕對值之和,彼處 $m<\theta<m+1$ (卽 OL_θ 在圖 4 *b* 情形中之中位數); 於是

$$\begin{aligned}\Delta_\theta&=L_1L_\theta+L_2L_\theta+\cdots+L_mL_\theta+L_\theta L_{m+1}+\cdots+L_\theta L_{2m}\\&=L_1L_m+L_2L_m+\cdots+L_mL_m+L_mL_{m+1}+\cdots+L_mL_{2m}\\&\qquad+L_mL_\theta+L_mL_\theta+\cdots+L_mL_\theta+L_\theta L_m+\cdots+L_\theta L_m\end{aligned}$$

$$= \Delta_m + (m - m)L_m L_g, \text{ 因 } L_m L_g = -L_g L_m$$

$$= \Delta_m$$

因之 $\Delta_g = \Delta_m = \Delta_{m+1}$, 因爲前此會證明 $\Delta_m = \Delta_{m+1}$.

此卽,在 N 爲偶數情形中, $\Delta_g \leq \Delta_i (i = 1, 2, \cdots, 2m)$.

最後我們可以總結說：各離中差絕對值之和不大於各任意離差絕對值之和. 證訖.

第五節

上下四分位數,四分位差,全距,均互差及離勢係數

1. 上下四分位數　如選擇一點代表中位數然,我們可另外加選兩點,如此其與中位數合在一起,將全部觀察值或頻數分配劃成四個相等部分. 此二點稱做四分位數 (quartile). 點在中位數之前者稱做下四分位數 (lower quartile), 或曰第一四分位數 (first quartile); 在其後者稱做上四分位數 (upper quartile), 或曰第三四分位數 (third quartile). 因之中位數有時被稱第二四分位數 (second quartile). 第一四分位數 Q_1 爲在累計曲線上與 $\Sigma f = N/4$ 成對應之 X 值; 即是全體觀察值中有四分之一小於 Q_1, 並且有四分之三大於 Q_1, 彼處 Σf 代表頻數之和(以下同此). 第二四分位數 Q_2 爲在累計曲線上與 $\Sigma f = N/2$ 成對應之 X 值; 即是全體觀察值中有二分之一小於 Q_2, 並且有二分之一大於 Q_2, 這個正與中位數之定義符合. 第三四分位數 Q_3 爲在累計曲線上與 $\Sigma f = 3N/4$ 成對應之 X 值; 即是在全體觀察值中有四分之三小於 Q_3, 並且有四分之一大於 Q_3. 由是知觀察值在 Q_1 與 Q_3 之間者有百分之五十

2. 四分位差　若將 Q_1 與 Q_3 之間的距離平均之，則得通常以 Q 代表之四分位差 (quartile deviation)，卽

(26) $$Q=(Q_3-Q_1)\div 2.$$

很顯然的，中位數不一定落在 Q_1 與 Q_3 之平分點上，卽是 $\tilde{X}$ 不一定等於 $(Q_1+Q_3)\div 2$；有時 $\tilde{X}$ 大於亦有時小於 $(Q_1+Q_3)\div 2$ (見圖 5).

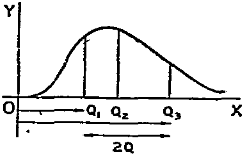


圖 5. 上下四分位數，中位數及四分位差。

Q 值爲描寫各變量散佈情形之良好常數，當取中位數描寫集中趨勢時，此値能示吾人以適宜之觀念。四分位數 Q_1 及 Q_3，和中位數 $\tilde{X}$ 一樣，可由累計頻數用比例法計算之。現在不復證明，僅將求 Q_1 及 Q_3 之公式分別列下：

(27) $$\begin{cases} Q_1=a_1+\dfrac{h}{c_1}(N/4-b_1), \\ Q_3=a_3+\dfrac{h}{c_3}(3N/4-b_3). \end{cases}$$

彼處 a_1 及 a_3 各代表 Q_1 及 Q_3 所在組段之低限；b_1 及 b_3 各爲各變量小於 a_1 及 a_3 發現之頻數；c_1 及 c_3 各代 Q_1 及 Q_3 所在組段之對應頻數.

例 1. 試根據表 11 裏所列之論據計算上下四分位數及四分位差.

由表 11 得知 $N=147$，因之 $N/4=36.75$ 及 $3N/4=110.25$；並且

得知 $a_1=104$ 公分，$a_3=111$ 公分，$b_1=35$，$b_3=110$，$c_1=10$，$c_3=7$. 將此等值分別代入式 (27)，我們有

$$Q_1=104+(36.75-35)\div 10=104.18\text{ 公分},$$

$$Q_3=111+(110.25-110)\div 7=111.04\text{ 公分}.$$

代入式 (26)，我們有

$$Q=(111.04-104.18)\div 2=3.43\text{ 公分}.$$

3. 全距　全距者爲觀察值中之最大者與最小者之差. 故知最大值與最小值之間包含觀察值之全體. 昔時人恆用全距以表示各變量之散佈情形. 但以該值本身不穩定，祇要最大值或最小值一有變化，其將隨之而變化；又以其所示離勢量度之概念過於粗率，現在用之者甚少.

4. 均互差　各觀察值間所有可能差量之和除以差量之個數，如此所得之平均數叫做均互差. 均互差可用 $\bar{D}$ 表之，其公式爲

(28) $\bar{D}=$

$$\frac{(N-1)(X_N-X_1)+(N-3)(X_{N-1}-X_2)+\cdots+(N-2r+1)(X_{N-r+1}-X_r)}{\dfrac{N(N-1)}{2}}$$

式 (28) 可證明之如下：

設 $X_1, X_2, \cdots, X_N$ 爲 N 個由小而大順序排列之觀察值. 於是各可能之差量可寫出之爲

$X_N-X_1,\quad X_N-X_2, \cdots,\quad X_N-X_{N-2}, X_N-X_{N-1}$

$X_{N-1}-X_1, X_{N-1}-X_2, \cdots, X_{N-1}-X_{N-2}$

$X_{N-2}-X_1, X_{N-2}-X_2, \cdots$

$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$

$X_3-X_1,\quad X_3-X_2.$

$X_2-X_1.$

將各差量加之，並且先將

第一行與第一列之差量加之，得

$$(X_N-X_1)+[(X_{N-1}-X_1)+(X_N-X_{N-1})]+\cdots$$
$$+(X_2-X_1)+(X_N-X_2)]$$
$$=(X_N-X_1)+(X_N-X_1)+\cdots+(X_N-X_1)$$
$$=(N-1)(X_N-X_1);$$

次將第二行與第二列之差量加之，但將第一行及第一列之差量除掉，得

$$(X_{N-1}-X_2)+[(X_{N-2}-X_2)+(X_{N-1}-X_{N-2})]+\cdots$$
$$+[(X_3-X_2)+(X_{N-1}-X_3)]$$
$$=(X_{N-1}-X_2)+(X_{N-1}-X_2)+\cdots+(X_{N-1}-X_2)$$
$$=(N-3)(X_{N-1}-X_2);$$

同理，將第 r 行與第 r 列之差量加之，但將第 r 行以前及第 r 列以上各差量除掉，我們有　$(N-2r+1)(X_{N-r+1}-X_r)$.

所以，將所有各差相加之結果為

$$\bar{D}=\frac{(N-1)(X_N-X_1)+(N-3)(X_{N-1}-X_2)+\cdots+(N-2r+1)(X_{N-r+1}-X_r)}{\frac{N(N-1)}{2}}$$

證訖.

例 2.　求數值 5,6,10,15,18,25 及 37 之均互差.

表 12.　均互差之計算

(1)	(2)	(2)−(1) (3)	(4)	(3)×(4) (5)
$X_1=5$	$X_N=37$	32	$(N-1)=7$	224
$X_2=6$	$X_{N-1}=25$	19	$(N-3)=5$	95
10	20	10	3	30
15	18	3	1	3
總　計	…	…	…	352

本例 $N(N-1)\div 2=8\times 7\div 2=28$. 故由式 (28) 得

$$\bar{D}=352\div 28=12.57.$$

5. 離勢係數　茲將三種常用之離勢係數分別予以定義如下:

標準差係數 $=100\sigma_X\div\bar{X}$.

平均差係數 $=MD\div\tilde{X}$.

四分位差係數 $=(Q_3-Q_1)\div(Q_3+Q_1)$.

彼處 $\bar{X}$, $\tilde{X}$, Q_1 及 Q_3 各代表平均數,中位數,下四分位數及上四分位數, σ_X 及 MD 各代表標準差及平均差(均與前同).

第六節　衆數

1. 定義及基本公式　設有頻數分配,各組之頻數中有一組發現頻數最多,該組之組值叫做衆數. 若以 X_m 代表發現頻數最多之觀察值或組值,並以 $\hat{X}$ 代表衆數則

(29) $$\hat{X}=X_m.$$

例 1. 根據表 4 裏所列之論據,計算衆數.

由表 4 一望即知最高頻數為 20; 並且其對應組之組值為 105.5 公分. 故由式 (29) 得知

$$\hat{X}=150.5 \text{ 公分}.$$

2. 密而思及鮑萊公式　觀察個數若是相當多,用公式(29)決定之衆數是很準確的. 然而密而思及鮑萊為求比較更準確之衆數起見,曾分別有下列公式之推薦:

(30) $$\hat{X}=a+\frac{hf_{+1}}{f_{-1}+f_{+1}},$$ (密而思公式, Mills' formula)

(31) $$\hat{X}=a+\frac{h(f_0-f_{-1})}{2f_0-f_{-1}-f_{+1}},$$ (鮑萊公式, Bowley's formula)

設達 a 為最多頻數所在組段之低限，f_{-1} 為最高頻數前一組之頻數，f_0 為最高頻數，f_{+1} 為最高頻數後一組之頻數。

例 2. 根據表 4 裏所列之論據，分別用公式 (30) 及 (31) 求衆數之值.

由表 4 得知：$a=105$ 公分，$f_{-1}=10$, $f_0=20$, $f_{+1}=12$, $h=1$ 公分。於是分別代入式 (30) 及 (31) 我們有

$$\check{X}=105+\frac{1\times 12}{10+12}=105.55 \text{ 公分},$$

$$\check{X}=105+\frac{1\times(20-10)}{40-10-12}=105.56 \text{ 公分}.$$

3. 皮爾生公式　用公式 (29) 求衆數，僅與衆數所在組段之頻數有關；用式 (30) 求之，僅與衆數所在上下組段之頻數有關；用式 (31) 求之，僅與衆數及其上下組段之頻數有關：這些公式均忽略了其他組段裏的頻數。經過悠久的經驗與仔細的觀察，在微偏態頻數分配的情形中，皮爾生曾發現平均數、中位數及衆數三者之關係如下：

(32) $$\check{X}=\bar{X}-3(\bar{X}-\tilde{X})\text{ 或 }=3\tilde{X}-2\bar{X}.$$

設處 $\bar{X}$ 代表平均數，$\tilde{X}$ 代表中位數及 $\check{X}$ 代表衆數(均和先前一樣).

例 3. 試根據表 4 所列之論據並應用式 (32)，計算衆數.

在第三節例 1 及第四節例 1 裏曾分別算得 $\bar{X}=107.87$ 公分，$\tilde{X}=106.50$ 公分. 故代入式 (32) 得

$$\check{X}=3\times 106.5-2\times 107.87=103.84 \text{ 公分}.$$

4. 理論與近似衆數之比較　用式 (29), (30), (31) 及 (32) 算得者通常叫做經驗或近似衆數 (empirical or approximate mode). 在精細分

析上,背係先求得代表頻數分配曲線之方程式 $Y=F(X)$, 然後由其縱坐標 (ordinate) Y 對於橫坐標 (abscissa) X 第一次微係數 (first derivative), 以決定衆數. 在此情形中,衆數爲由式 $\frac{dY}{dX}=0$ 求得之 X 值. 用此法決定之衆數通常稱做理論衆數. 此與曲線配合有關,容待下章詳述. 現在所欲說明者爲用皮爾生公式算得之衆數與用曲線方程式算得者極爲切近,此切近情形可由下列之表 13 見之. 表中所列之各近似衆數與理論衆數,除哥拉斯渦及丹地二段衆數近似值與理論值之差達 0.021 吋外,均相差甚微.

表 13. 近似衆數與理論衆數之比較(錄自參考書 44)

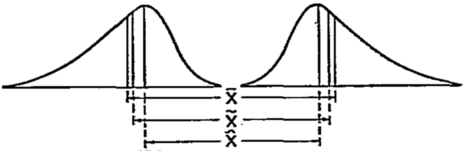
觀察臺名稱	氣壓高度(吋)			
	平均數	中位數	近似衆數	理論衆數
南安普敦 (Southampton)	29.961	30.000	30.038	30.039
倫敦地里 (Londonderry)	29.891	29.915	29.963	29.960
加馬耳姑 (Garmarthen)	29.953	29.974	30.018	30.013
哥拉斯渦 (Glasgow)	29.885	29.906	29.946	29.967
丹　地 (Duddee)	29.870	29.890	29.930	29.951

第七節　平均數,中位數及衆數之討論

平均數爲橫軸上之一點,此點就是在力學上我們所稱說的面積的重心. 設有一薄而勻之金屬片,切成之圖形與頻數分配之直方圖

(histogram) 相當。如果經過代表平均數之點，將此金屬片用一線繫之於水面上，則此金屬片不論怎樣轉動其底邊將永與此水平面平行。頻數分配之中位數亦為橫軸上之一點，經過此點作一直線，使與金屬片底邊成垂直，此垂線將分面積為二等分。頻數分配之衆數為與代表頻數分配圓滑曲線 (smoothed curve) 上最高點成對應之橫坐標。

在微偏態分配情形中，這三種集中常數——平均數，中位數及衆數——之位置，如圖 6 所示。由圖 6a 可見：衆數最大，中位數次之，平均數最小；凡是偏左的頻數分配都有這種現象。由圖 6b 可見：平均數最大，中位數次之，衆數最小，正與圖 6a 情形相反；凡是偏右的頻數分配都有這種現象。如果頻數分配為常態，此三種常數相等，或者說在圖上代表這三個常數之點重合。



a. 偏左　　b. 偏右

圖 6. 平均數 $\bar{X}$，中位數 $\tilde{X}$ 及衆數 $\hat{X}$ 之位置。

此時學者不免要這樣問：平均數，中位數及衆數三者以那一種較切實用呢？此問題之回答自然要依各集中常數之特點及資料之性質而定。在多數情形中，前者優於後二者，良以第一種之定義確鑿，計算之手續簡單，能用代數方法處理，並且便於理論上之研討與追溯。中位數之定義不若平均數之謹嚴，其於理論上之探討，亦不若平均數之易於尋繹；然而

牠是一富有代表性之數值，至少在下列三種之情形中，我們不能不承認中位數猶有勝於平均數之處：

a. 當頻數分配之兩端數值特別大或特別小，平均數將受少數異常數值影響，平均數失卻代表性時.

b. 當頻數分配之一端或兩端有開口時，例如某校註册部報告學生之成績爲：60 分以下者 21 人，60–69 分者 1060 人；在此情形中，無須知道 60 分以下者的事實，我們可以求此分配之中位數，但無法求其平均數。

c. 當觀察之事物不能以數量計，只能列其次序時.

至於衆數也是一最通常之數值，市場中商人極喜用之，因爲商人皆欲知某類貨物或某種品質之貨物易於傾銷.再有關於工資及生活用度，衆數能映出平均情形.在頻數分配之算學理論中，衆數之概念仍是很有用處的.然而，除非有大批觀察值衆數無實際上之意義.

有時頻數分配之衆數似乎不止一個，卽其有時不止一峯.此乃由於資料之粗雜所致，頻數分配未必眞呈雙峯或多峯態.眞正雙峯或多峯之頻數分配，不在本書討論之列.

第八節 · 幾何平均數

1. 定義及公式 N 個觀察值相乘積，開以 N 次方，所得之方根叫做幾何平均數；或者說，牠是 N 個觀察值對數平均數之逆對數.設 G 表示幾何平均數，$\log^{-1}$ 表示逆對數.於是

(33) $$G=(X_1X_2\cdots X_N)^{\frac{1}{N}}，或者$$

(34) $$G=\log^{-1}\left[\frac{1}{N}\sum_{i=1}^{N}\log X_i\right]$$

例 1. 求 3,6,12,24 及 48 之幾何平均數.

將已給之值代入式 (33), 則有

$$G=(3\times6\times12\times24\times48)^{\frac{1}{5}}=(3^5\times2^{10})^{\frac{1}{5}}=12.$$

例 2. 求 7.96,13.82,22.95 及 35.34 之幾何平均數.

log 7.96 =	0.90091
log 13.82 =	1.14051
log 22.95 =	1.36078
log 35.34 =	1.54827
總計 =	4.95047
總計 ÷ 4 =	1.23762

故由公式 (34) 得

$$G=\log^{-1}1.23762=17.28.$$

2. 定律及應用　當統計資料之一端數值有限,他端則否,並且一數與其次一數有成等比之趨勢,遵循下列之簡單指數定律 (simple exponential law) 時,幾何平均數乃是一個較為適宜之平均數.簡單指數定律之公式如下:

(35) $$y=ar^x.$$

彼處 a 及 r 均為常數, x 及 y 均為變數.

自然界時間數列之增加往往遵循此定律,因此有時被稱做成長定律 (law of growth); 但稱之為指數函數(exponential function) 較為貼切些.即如某城之人口數,培養器上細菌之數目等,依時之增加情形,係遵循指數定律,即式 (35) 中之 x 將代表時間, y 代表某城人口數.細菌

數目等;又如若干年後按複利計算之本利和亦遵循此定律,因複利之公式爲

(36) $$S=P(1+i)^N.$$

彼處 S 代表 N 年後以複利計算之本利和,P 代表本金,i 代表利率。此式與式(35)比較,得知此時之 $r=1+i$, $a=P$, 並且 S 及 N 均爲變數而與式(35)中之 x 及 y 成對應。

如果方程式(35)中之 y 代表人口,x 代表年數;則用幾何平均數估計各年人口數較用其他平均數估計者爲準確。例如,已知某城 1920 年人口數爲 2500, 1930 年人口數爲 5000. 我們希望知道 1925 年的人數及每年之平均增率.假設已給之事實祇此而已,我們用下式估計該年人口數較近事實:

$$G=(y_1\times y_2)^{\frac{1}{2}}=(2500\times 5000)^{\frac{1}{2}}=3535.$$

每年平均增加率可用式(36)求之如下:

$$5000=2500(1+i)^{10},$$

卽是 $$(1+i)^{10}=2.$$

所以 $$i={}^{10}\sqrt{2}-1=0.0718=7.18\%.$$

如此每年之平均增加率爲 7.18%.

幾何平均數又可用以求物價指數之平均數,因爲物價指數係某一時期物品價格與另一時期物品價格之比.普通言之,欲知率,比及百分比之變遷情形,以用幾何平均數表之,較爲適宜.

第九節 倒數平均數

1. 定義及公式 N 個觀察值反商平均數之反商,叫做倒數平均數。如果以 H 表示倒數平均數,則

$$H=\frac{1}{\frac{1}{N}\left(\frac{1}{X_1}+\frac{1}{X_2}+\frac{1}{X_3}+\cdots+\frac{1}{X_N}\right)}=\frac{N}{\sum_{i=1}^{N}\frac{1}{X_i}} \tag{37}$$

例 1. 有一工程，A 做之 4 分鐘完成，B 做之 5 分鐘完成，C 做之 6 分鐘完成，D 做之 10 分鐘完成，E 做之 12 分鐘完成，問他們工作之平均速率爲何?(因本例係用 t/d 形式表示速率，並因以 t 當做固定元素爲宜，故求平均速率時宜用式(37)，彼處 t 代表時間，d 代表工程，參閱本節 3 之規則 a)。

將已給之值代入式(37)，得

$$H=\frac{1}{\frac{1}{5}\left(\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{10}+\frac{1}{12}\right)}\ \frac{\text{分鐘}}{\text{工程}}$$

$$=\frac{25}{4}\ \frac{\text{分鐘}}{\text{工程}}=6\frac{1}{4}\ \frac{\text{分鐘}}{\text{工程}}$$

2. 算術平均數與倒數平均數之討論　在時間速率情形中，我們有兩個數量：一爲用 t 表示之時間，一爲用 d 表示之距離、工程、溫度等。於是我們可以做以下之觀察：

a. 速率可用 d/t 或用 t/d 形式表示或敍述之。卽如我們可以說車之速度爲每小時 30 里，也可以說車之速率爲每 2 分鐘 1 里。第二種表示速率之法固不習見，但其敍述速率之概念亦甚明晰。

b. 當求平均速率時，我們第一樣要決定以 d 抑或以 t 做討論上之基礎。論及以何做基礎或做固定元素 (fixed element) 較爲適宜，意見頗不一致。例如有一班學生，准許其在 15 分鐘之內盡量做題，並記載每個學生做得正確之題數，教育家一定要主張以時間爲固定元素，而以每單位時間內所能解答之題數爲變數爲宜；或有人說，學生解答一題之時

間 (t) 應爲適當之變數，而以題數 (d) 爲固定元素.

c. d/t 形式之倒數平均數爲 t/d 形式之算術平均數，此事實是很顯然的，可直接由定義證明之.

爲便於說明起見，我們舉實例如下:

假設有三種速率不同的車，其速率爲

I $\begin{cases}\text{甲車每小時行 15 里,}\\ \text{乙車每小時行 20 里,}\\ \text{丙車每小時行 30 里;}\end{cases}$

但是她們的速率仍然可以這樣說

II $\begin{cases}\text{甲車每 4 分鐘行一里,}\\ \text{乙車每 3 分鐘行一里,}\\ \text{丙車每 2 分鐘行一里.}\end{cases}$

在情形 I 裏之倒數平均速率爲每小時 20 里;在情形 II 裏之算術平均速率爲每 3 分鐘行一里，也是每小時行 20 里.同樣在情形 I 裏之算術平均速率爲每小時 $21\frac{2}{3}$ 里，適與在情形 II 裏之倒數平均數同.

因此便發生疑問:那個算法是正確呢? 每小時 20 里歟? 每小時 $21\frac{2}{3}$ 里歟? 此問題須俟我們同意以時間或距離二者那個爲固定元素後，方能做明瞭之答覆.下面的解釋能幫助我們了解此點.

在第一種情形中，讓 $X_i=d_i/t_i$ 表示第 i 個事物之速率，$i=1、2、3\cdots、N$, D 表示總距離(全程)，T 表示總時間(行全程的時間).於是平均速率爲

$$\frac{D}{T}=\frac{d_1+d_2+\cdots+d_N}{t_1+t_2+\cdots+t_N}$$

$$(38)\qquad =\frac{d_1+d_2+\cdots+d_N}{\frac{d_1}{X_1}+\frac{d_2}{X_2}+\cdots+\frac{d_N}{X_N}}$$

$$(39)\qquad =\frac{t_1X_1+t_2X_2+\cdots+t_NX_N}{t_1+t_2+\cdots+t_N}.$$

如果距離為固定元素，即是 d 為常數;則由式 (38) 得

$$\frac{D}{T}=\frac{Nd}{d\Sigma\frac{1}{X_i}}=\frac{N}{\Sigma\frac{1}{X_i}}.$$

此為倒數平均數.如果 t 為固定元素，則由式 (39) 得

$$\frac{D}{T}=\frac{t\Sigma X_i}{Nt}=\frac{1}{N}\Sigma X_i.$$

此為算術平均數.

在第二種情形中讓 $X_i=t_i/d_i$ 表示第 i 個事物之速率，於是平均速率為

$$\frac{T}{D}=\frac{t_1+t_2+\cdots+t_N}{d_1+d_2+\cdots+d_N}=\frac{\Sigma t_i}{\Sigma d_i}$$

$$(40)\qquad =\frac{d_1X_1+d_2X_2+\cdots+d_NX_N}{d_1+d_2+\cdots+d_N}$$

$$(41)\qquad =\frac{t_1+t_2+\cdots t_N}{\frac{t_1}{X_1}+\frac{t_2}{X_2}+\cdots\frac{t_N}{X_N}}$$

如果 d 為固定元素，則由式 (40) 得

$$\frac{T}{D}=\frac{d\Sigma X_i}{Nd}=\frac{1}{N}\Sigma X_i \qquad \text{(算術平均數);}$$

如果 t 為固定元素，則由式 (41) 得

$$\frac{T}{D}=\frac{Nt}{t\Sigma\frac{1}{X_i}}=\frac{N}{\Sigma\frac{1}{X_i}} \qquad \text{(倒數平均數).}$$

3. 求平均速度之規則 由於上面的討論,我們得到 a、b 兩條求平均速率的規則:

規則 a. 當元素 d 固定並且速率用 d/t 形式表示時,或者當元素 t 固定並且速率用 t/d 形式表示時,均宜用倒數平均數求其平均速率.

規則 b. 當固定元素爲 t 並且速率用 d/t 形式表示時,或者當固定元素爲 d 並且速率用 t/d 形式表示時,均宜用算術平均數求其平均速率.

在物價情形中,物價 B 是兩種數量之比,一量爲錢幣單位 (m),一量爲商品單位 (c), 故其可依同理討論; 不過將上述之時間易之以錢幣單位而已.物價可用兩種方法表示:一用每單位商品價值多少錢 (m/c) 表示,一用每元錢可購若干物品 (c/m) 表示.卽如 100 斛麥售價爲 75 元,可用每元購麥 100/75 斛 $=4/3$ 斛表示之,亦可用每斛售價 75/100 元 $=3/4$ 元表示之.

正確之平均數依物價之敍述法如何以及以物價或商品何者爲固定元素而定. 欲知其詳,學者可參閱 The Nature and Use of the Harmanic Mean—Ferger, Jour. Amer. Stat. Ass. vol. 26. 1931, p. 36-40

玆舉一實例以明平均速率規則之應用.

例 2. 有一駕車旅行者,在三處購買汽油,計在甲處每元可購汽油 5 加侖,在乙處每元可購 7 加侖,在丙處可購 6 加侖. 這裏之油價係用 c/m 形式表示;以 c 爲固定元素,並且以 m 爲變數似較適宜.若以 c/m 代 d/t, 卽以物價代速率,於是根據平均速率規則之第一條,我們宜求倒數平均數.因此

$$H=\frac{3}{\frac{1}{5}+\frac{1}{6}+\frac{1}{7}}=\frac{630}{107}\frac{\text{加侖}}{\text{元}}=5.89\frac{\text{加侖}}{\text{元}};$$

或者說 $$\frac{1}{H}=\frac{107}{630}\frac{\text{元}}{\text{加侖}}=.17\frac{\text{元}}{\text{加侖}}.$$

問題 II

1. 試將表 4 裏男生身長之頻數分配用直方圖及多邊圖表示之。

2. 試以 2 公分及 5 公分爲組距，直接由表 1 裏所列之論據製成頻數分配表。

3. 依照表 1 回答下列各問題：

 a. 學生身長在 105—107 公分者有若干人？

 b. 學生身長在 103 公分以下者有若干人？

 c. 學生之身長最矮者有若干人？

 d. 學生之身長在 99 至 109 公分之間者所佔之百分數爲何？

4. 依照表 6 回答上題所問之各問題。

5. 試求表 6 所列頻數分配之累計頻數，並製成累計頻數圖（肩形圖）。

6. 試將下式寫成展開形式：

$$(a)\ \sum_{i=1}^{k}x_if_i;\quad (b)\ \sum_{i=1}^{k}x_i^2f_i;\quad (c)\ \sum_{i=1}^{k}(x_i-\bar{x})f_i.$$

7. 試將下式寫成展開形式：

$$(a)\ \sum_{i=1}^{N_1}f_i;\quad (b)\ \sum_{i=n_1+1}^{n_1+n_2}f_i;\quad (c)\ \sum_{i=1}^{n_1}x_if_i+\sum_{i=n_1+1}^{n_1+n_2}x_if_i.$$

8. 如果 $n_1+n_2=k$, 試將上題 (c) 以一單獨和表之.

9. 證明.

(a) $$\sum_1^k (x_i+1)^2 f_i=\sum_1^k x_i^2 f_i+2\sum_1^k x_i f_i+N.$$

(b) $$\sum_{x=0}^n x(x-1)p=\sum_{x=2}^n x(x-1)p.$$

10. 用簡短符號表示:

(a) 各 x 之平方和除以其和之平方.

(b) 如果各 x 爲由 1 至 N 之整數, 用 N 表示 (a) 之值.

(c) 證明首 N 個整數之平均數爲 $\frac{1}{2}(N+1)$.

11. 某大學第一年級學生期考平均分數如下:

男生 501 人每人平均分數爲 75.50;

女生 356 人每人平均分數爲 76.39.

求該校男女生合計之平均分數.

12. 一個學生期終考試化學筆試成績爲 65%, 口試成績爲 85%, 實驗成績爲 80%. 如果將其成績各以 1, 2 及 3 權之, 問該生之平均成績爲何?

13. 第一學期之末某生受課之時數及實得之成績如下:

算學 4 小時, 實得分數爲 88,

英文 4 小時, 實得分數爲 80,

歷史 3 小時, 實得分數爲 85,

物理 4 小時, 實得分數爲 78.

問該生之平均分數爲何?

14. 某城人口於 5 年內由 225,000 增至 245,000. 問每年人口平均增加若干?每年之增加率爲何?

15. 某種細菌之數目在某日中午爲 4×10^6, 翌日中午爲 9×10^6. 如果細菌每小時之增加率爲一常數,問在翌日中夜其數應爲若干?

16. 有兩個正數 a 及 b, 其幾何平均數爲 $x=\sqrt{ab}$. 試在以 $(a+b)$ 爲直徑之圓上給出 x 之值來.

17. 已給兩組正變量

$$X_{11}、X_{12}、X_{13}、\cdots、X_{1N},$$
$$X_{21}、X_{22}、X_{23}、\cdots、X_{2N}.$$

證明兩組對應變值之比的幾何平均數等於其幾何平均數之比.

18. 證明正變數頻數分配幾何平均數之公式 (33) 可寫做

$$G=[X_1^{f_1}X_2^{f_2}\cdots X_k^{f_k}]^{\frac{1}{N}}.$$

彼處 X_i 爲觀察值, f_i 爲 X_i 值重複的次數,並且 $N=\sum_1^k f_i$.

19. 假想有兩個正變量 X_1 及 X_2, 證明牠們的幾何平均數等於牠們的算術平均數與倒數平均數之幾何平均數.

20. 假設有 20 隻船,每年在大西洋往返六次——即 $\frac{1}{6}$ 年(約 60 日)往返一次; 10 隻船每年往返四次——即 $\frac{1}{4}$ 年(約 90 日)往返一次,問此三十船往返一次平均所需之日數爲何?

21. 如果 $2a$ 爲兩個有理數 b 及 c 之倒數平均數,證明 a、b 及 c 之平方和爲一有理數之平方.

22. 如果 A、G 及 H 各代表 N 個正數之算術、幾何及倒數平均數,求證

$$H<G<A.$$

23. 核對表 6 各行數值有無錯誤.

24. 核對表 7 各行數值有無錯誤.

25. 理論頻數在 $\bar{X}-\sigma$ 與 $\bar{X}+\sigma$ 之間,在 $\bar{X}-2\sigma$ 與 $\bar{X}+2\sigma$ 之間,在 $\bar{X}-3\sigma$ 與 $\bar{X}+3\sigma$ 之間各爲 68.3%, 95.4% 及 99.7% (見附錄 III 表 B 及第三節 4). 試用本章第三節例 4 算得之 $\bar{X}=107.87$ 公分, $\sigma_X=5.2929$ 公分,驗證表 6 裏所列之各觀察值與此理論頻數之吻合情形.

第三章　動差·機率及常態曲線

第一節　動差之意義及其圖解

摘述資料之要點以及描寫資料之特徵，爲統計學上重要問題之一。關於此點費宣曾有言曰：

"如將大批數量資料堆在那裏，我們得不到什麽印象，宜將其處理之，俾能用少許數量代表原始之蕪雜數量，並宜儘其可能包含所有之切題事實."*

上文所說之少許數量往往得用動差表示之．動差的級數不一，如所謂一級、二級等等動差．動差的定義很簡單，所謂幾級動差乃是將各觀察值或中値方至第幾次之平均數．在統計計算中，除去常要計算一級動差(其即算術平均數)及二級動差(以平均數爲原點二級動差之平方根，即爲標準差)外，當研究頻數分配種種特徵時，我們仍然要計算三級及四級動差．各動差在統計學上之意義容後詳述．玆將動差名詞之淵源略加追溯，並將其在力學上之意義略加說明，如下

動差這個名詞本淵源於力學，在那裏稱做力之動差 (moment of force)．譬如我們有一根剛體槓桿 (rigid lever)，其一點 O 繫着外力 F_1

* Foundations of Theoretical Statistics, Phil. Trans. Roy. Soc., Vol. 222a (1932), P. 309.

及 F_2. 此 O 點人都知之爲支點 (fulcrum), 如圖 7 所示. 設 F_1 與 O 點之距離爲 X_1, F_2 與 O 點之距離爲 X_2.

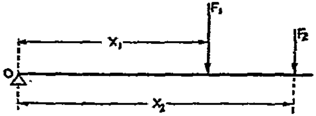


圖 7. 槓桿。

於是乘積 X_1F_1 及 X_2F_2 均被稱之爲力之動差. 同樣, 如果有 n 個力量 $F_1, F_2, \cdots\cdots F_n$ 共同作用於一根槓桿上, 力之方向相同, 並且其與支點 O 之距離各爲 $X_1, X_2\cdots\cdots X_n$; 則

$$X_1F_1+X_2F_2+\cdots\cdots+X_nF_n=\sum_{i=1}^{n}X_iF_i.$$

稱爲各力動差之總和 (total moment of forces).

如果將距離平方之, 則得二級動差之總和 $\sum_{i=1}^{n}X_i^2F_i$; 將距離立方之, 則得三級動差之總和 $\sum_{i=1}^{n}X_i^3F_i$; 如此等等.

在力學上之 X 代表支點與力點之距離, 在統計學上代表變量或組值; 在力學上 F 代表力量, 在統計學上代表頻數. 如將頻數分配視做以金屬薄片製成之直方圖 (參閱前章第七節), 則力學上 X 與 F 所代表之意義將與統計學上所代表者可以互通了. 不過在統計學上我們所用的動差都是以頻數之總和爲單位, 卽是將各級動差統以觀察個數 N 除之, 因此特稱之爲統計動差 (statistical moment) 或簡稱之曰動差, 以示與力之動差有別. 對於零點並以原始量度爲單位之一級統計動差爲

平均數，其與力學上重心之坐標成對應，已在前章第十節討論過了；以平均數爲原點並以原始量度爲單位之二級動差與力學上之旋轉惰徑 (radius of gyration) 成對應，在統計學上亦給一特別名稱，名之曰變差。

第二節 一般動差

以任意數爲原點，並以原始量度爲單位之動差可以叫做一般動差 (general moment). r 級一般動差 (general moment of rth order) 可用符號 λ_{rX} 代表之，即

$$(1)\qquad \lambda_{rX}=\frac{1}{N}\sum_{i=1}^{n}(X_i-X_0)^rF_i,$$

彼處 X_0 代表任意數，λ 代表以 X_0 爲原點之動差，F_i 爲與 X_i 成對應之頻數，$N=\sum_{i=1}^{n}F_i$；綴於 λ 右下角之 r 表示動差之級次，其下之大寫字母 X 表示動差係以原始量度爲單位．在特殊情形時，$r=1, 2, 3$ 及 4，我們有以原始量度爲單位之首四級一般動差

$$(2)\qquad \begin{cases}\lambda_{1X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-X_0)F_i,\\[2ex] \lambda_{2X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-X_0)^2F_i,\\[2ex] \lambda_{3X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-X_0)^3F_i,\\[2ex] \lambda_{4X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-X_0)^4F_i.\end{cases}$$

式 (1) 及 (2) 旣爲動差之一般形式，各特殊動差皆可據此推求之。

第三節　主要動差

前節所講者,係以任意數爲原點以原始量度爲單位之各級動差。然而在統計理論及應用上所需要者不是那些,乃是以平均數爲原點之各級動差。以平均數爲原點之各級動差叫做主要動差 (principal moment), 通常以 μ 記之,茲將各級主要動差用不同單位表示之如下:

1. 以原始量度爲單位　使式 (1) 中之 $X_a=\bar{X}$, 則得以原始量度爲單位之 r 級主要動差

$$\mu_{rX}=\frac{1}{N}\sum_{i=1}^{n}(X_i-\bar{X})^r F_i. \tag{3}$$

並且和先前一樣,使 $r=1, 2, 3$ 及 4, 則得以原始量度爲單位之首四級主要動差

$$\begin{cases}\mu_{1X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})F_i,\\ \mu_{2X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})^2F_i,\\ \mu_{3X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})^3F_i,\\ \mu_{4X}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(X_i-\bar{X})^4F_i;\end{cases} \tag{4}$$

彼處 μ 之右下角所綴之第一個數碼 r, 1, 2, 3 及 4 係代表動差之級次,第二個大寫字母 X 係代表該動差以原始量度爲單位(和前節一樣)。

2. 以組距爲單位　若將式 (1) 中之變數 X 以助變數 x 代之, X_a 以 x 之平均數 $\bar{x}$ 代之;則有以 $\bar{x}$ 爲原點並以組距爲單位之 r 級及首四級主要動差

$$(5)\qquad \mu_{rx}=\frac{1}{N}\sum_{i=1}^{n}(x_i-\bar{x})^r f_i,$$

及

$$(6)\qquad \begin{cases} \mu_{1x}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})f_i, \\ \mu_{2x}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^2 f_i, \\ \mu_{3x}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^3 f_i, \\ \mu_{4x}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(x_i-\bar{x})^4 f_i; \end{cases}$$

此處 μ 右下角所綴之第一個數碼 $r, 1, 2, 3$ 及 4 亦係代表動差之級次。第二個小寫字母 x 係代表該動差以原始量度爲單位，f_i 代表與 X_i 成對應之頻數，並且 $N=\sum_{i=1}^{n}f_i$.

3. 以標準差爲單位　若將式(1)及(2)中之 X 以 t 代之，$\bar{X}$ 以 t 之平均數 $\bar{t}$ 代之；於是我們有：

$$(7)\qquad \mu_{rt}=\frac{1}{N}\sum_{i=1}^{n}(t_i-\bar{t})^r \varphi_i,$$

$$(8)\qquad \begin{cases} \mu_{1t}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(t_i-\bar{t})\varphi_i, \\ \mu_{2t}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(t_i-\bar{t})^2\varphi_i, \\ \mu_{3t}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(t_i-\bar{t})^3\varphi_i, \\ \mu_{4t}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{n}(t_i-\bar{t})^4\varphi_i; \end{cases}$$

彼處如 μ 右下角之第一個數碼代表之意義與前同,第二個字母 t 係表示各級動差以標準差爲單位,φ_i 代表與 X_i 成對應之頻數,並且 $N=\sum_{i=1}^{n}\varphi_i$.

第四節　輔助動差

以 O 爲原點之動差叫做輔助動差 (crude moment), 通常以 ν 記之. 玆將各級輔助動差用不同單位表之如下:

1. 以原始量度爲單位　若使式 (1) 及 (2) 中之 $X_0=0$, 則得以原點並以原始量度爲單位之 r 級及首四級輔助動差

$$\nu_{rX}=\frac{1}{N}\sum_{i=1}^{n}X_i^rF_i, \tag{9}$$

$$\begin{cases}\nu_{1X}=\frac{1}{N}\sum_{i=1}^{n}X_iF_i,\\ \nu_{2X}=\frac{1}{N}\sum_{i=1}^{n}X_i^2F_i,\\ \nu_{3X}=\frac{1}{N}\sum_{i=1}^{n}X_i^3F_i,\\ \nu_{4X}=\frac{1}{N}\sum_{i=1}^{n}X_i^4F_i;\end{cases} \tag{10}$$

彼處 ν 右下角所綴之第一個數碼 r, 1, 2, 3 及 4 亦係代表動差之級次; 第二個大寫字母 X 亦係代表該動差以原始量度爲單位(和上面一樣).

2. 以組距爲單位　如將式 (1) 及 (2) 中之 X 以 x 代之, X_0 以 0 代之; 則得以組距爲單位之 r 級及首四級輔助動差

$$\nu_{rx}=\frac{1}{N}\sum_{i=1}^{n}x_i^rf_i \tag{11}$$

及

$$(12)\quad \begin{cases} \nu_{1x}=\frac{1}{N}\sum_{i=1}^{n}x_if_i, \\ \nu_{2x}=\frac{1}{N}\sum_{i=1}^{n}x_i^2f_i, \\ \nu_{3x}=\frac{1}{N}\sum_{i=1}^{n}x_i^3f_i, \\ \nu_{4x}=\frac{1}{N}\sum_{i=1}^{n}x_i^4f_i; \end{cases}$$

彼處 ν 右下角所綴之第一個數碼 r, 1, 2, 3 及 4 係代表動差之級次，所綴第二個小寫字母 x 係代表各動差以組距爲單位(和先前同).

第五節　μ_{rx} 與 ν_{rx} 之關係

前面曾經說過，在統計工作中，我們需要計算首四級主要動差，換言之，我們欲求 μ_{1X}, μ_{2X}, μ_{3X} 及 μ_{4X} 之值；或 μ_{1x}, μ_{2x}, μ_{3x} 及 μ_{4x} 之值；或 μ_{1t}, μ_{2t}, μ_{3t} 及 μ_{4t} 之值. 但 $\bar{X}$, $\bar{x}$ 及 $\bar{t}$ 之值率係帶有不盡小數，因之 $(X-\bar{X})$, $(x-\bar{x})$ 及 $(t-\bar{t})$ 帶有不盡小數. 若將此帶有不盡小數之數自乘至二次、三次以至四次以求二級、三級及四級之主要動差，則所用之運算手續，至爲紛繁. 首四級主要動差之值實際不要直接用式 (4), (6) 及 (8) 求之，乃是間接用式 (12) 求得 ν_{1x}, ν_{2x} ν_{3x} 及 ν_{4x}——輔助動差——之值後求之. 因此我們在此處先推求 μ_{rx} 與 ν_{rx} 之關係，再在下節推求 μ_{rX}, μ_{rx} 及 μ_{rt} 之相互關係.

將式 (5) 右方展開之，得

(13) $$\mu_{rx}=\frac{1}{N}\Sigma\left(x_i^r-r\bar{x}x_i^{r-1}+\frac{r(r-1)}{2!}\bar{x}^2x_i^{r-2}-\cdots\cdots\right)f_i$$

$$=\nu_{rx}-r\bar{x}\nu_{(r-1)x}+\frac{r(r-1)}{2!}\bar{x}^2\nu_{(r-2)x}-\cdots\cdots$$

使 $r=1, 2, 3$ 及 4, 並且因爲 $\nu_{1x}=\bar{x}$, $\nu_{0x}=1$; 於是我們有

(14) $$\begin{cases}\mu_{1x}=0,\\ \mu_{2x}=\nu_{2x}-\nu_{1x}^2,\\ \mu_{3x}=\nu_{3x}-3\nu_{1x}\nu_{2x}+2\nu_{1x}^3,\\ \mu_{4x}=\nu_{4x}-4\nu_{1x}\nu_{3x}+6\nu_{1x}^2\nu_{2x}-3\nu_{1x}^4.\end{cases}$$

這就是說，有了以組距爲單位各級輔助動差 ν_{rx} 之後可應用上列各式以求各級主要動差 μ_{rx}.

第六節 μ_{rX}, μ_{rx} 及 μ_{rt} 之互相關係

設 X 與 x 之關係爲 $X=X_0+hx$, 如前章公式 (5) 所表示者. 於是由式 (3) 得

(15) $$\mu_{rX}=\frac{h^r}{N}\Sigma(x_i-\bar{x})^rf_i,\quad f_i=F_{X_0+hx_i},$$

並且由式 (4) 得

(16) $$\begin{cases}\mu_{1X}=\frac{h}{N}\Sigma(x-\bar{x})f_i,\\ \mu_{2X}=\frac{h^2}{N}\Sigma(x-\bar{x})^2f_i,\\ \mu_{3X}=\frac{h^3}{N}\Sigma(x-\bar{x})^3f_i,\\ \mu_{4X}=\frac{h^4}{N}\Sigma(x-\bar{x})^4f_i\end{cases}$$

因之應用公式 (5) 及 (6), 我們得將公式 (15) 及 (16) 分別寫做

(17) $$\mu_{rX}=h^r\mu_{rx},$$

及

(18) $$\begin{cases}\mu_{1X}=h\mu_{1x},\\ \mu_{2X}=h^2\mu_{2x},\\ \mu_{3X}=h^3\mu_{3x},\\ \mu_{4X}=h^4\mu_{4x}.\end{cases}$$

如使 $X=X_0+\sigma_X t$, 即使前章式 (5) 中之 $h=\sigma_X$ 及 $x=t$, 則由式 (3) 得

(19) $$\mu_{rX}=\frac{\sigma_X{}^r}{N}(t_i-\bar{t})^r\varphi_i,\quad \varphi_i=F_{X_0+\sigma_X t_i},$$

並且由式 (4) 得

(20) $$\begin{cases}\mu_{1X}=\dfrac{\sigma_X}{N}\Sigma(t-\bar{t})\varphi_i,\\ \mu_{2X}=\dfrac{\sigma_X{}^2}{N}\Sigma(t-\bar{t})^2\varphi_i,\\ \mu_{3X}=\dfrac{\sigma_X{}^3}{N}\Sigma(t-\bar{t})^3\sigma_i,\\ \mu_{4X}=\dfrac{\sigma_X{}^4}{N}\Sigma(t-\bar{t})^4\varphi_i.\end{cases}$$

因之應用公式 (7) 及 (8), 我們得將公式 (19) 及 (20) 分別寫做

(21) $$\mu_{rX}=\sigma_X{}^r\mu_{rt},$$

及

$$
(22)\qquad \begin{cases} \mu_{1X} = \sigma_X \mu_{1t}, \\ \mu_{2X} = \sigma_X^2 \mu_{2t}, \\ \mu_{3X} = \sigma_X^3 \mu_{3t}, \\ \mu_{4X} = \sigma_X^4 \mu_{4t}. \end{cases}
$$

由式(17)及(21)我們有

$$
(23)\qquad h^r \mu_{rx} = \sigma_X^r \mu_{rt},
$$

並且由式(18)及(22)則得

$$
(24)\qquad \begin{cases} h\mu_{1x} = \sigma_X \mu_{1t}, \\ h^2\mu_{2x} = \sigma_X^2 \mu_{2t}, \\ h^3\mu_{3x} = \sigma_X^3 \mu_{3t}, \\ h^4\mu_{4x} = \sigma_X^4 \mu_{4x} \end{cases}
$$

由式(17), (21)及(23)可見：如果知道 μ_{rX}, μ_{rx} 及 μ_{rt} 三者之一，便可計算其餘；由式(18), (22)及(24)可見：如果知道以三種不同單位表示之首四級主要動差之一種，便可計算其餘.

第七節　β_1, β_2, κ_1 及 κ_2 之計算步驟

為測定頻數分配為常態，帶偏態或帶峯態以及配合皮爾生各型曲線，我們須計算以平均數為原點之首四級動差——首四級主要動差，並且由此以計算 β_1, β_2, κ_1 及 κ_2. 茲將計算 β_1, β_2, κ_1 及 κ_2 所用之公式列後：

$$
(25)\quad\begin{cases}
\beta_1 = \mu_{3x}^2 \div \mu_{2x}^3, \\
\beta_2 = \mu_{4x} \div \mu_{2x}^2, \\
\kappa_1 = 2\beta_2 - 3\beta_1 - 6, \\
\kappa_2 = (\beta_2+3)^2\beta_1 \div \{(8\kappa_1 + 12\beta_1 + 48)\kappa_1\}.
\end{cases}
$$

再由式 (18) 及 (22) 得知 β_1 及 β_2 之值與表示各級動差所用之單位無關,即是

$$\beta_1 = \mu_{3x}^2 \div \mu_{2x}^3 = \mu_{3t}^2 \div \mu_{2t}^3,$$

$$\beta_2 = \mu_{4x} \div \mu_{2x}^2 = \mu_{4t} \div \mu_{2t}^2.$$

因此 β_1、β_2、κ_1 及 κ_2 之值可由下列之式 (26) 或 (27) 求之:

$$
(26)\quad\begin{cases}
\beta_1 = \mu_{3x}^2 \div \mu_{2x}^3, \\
\beta_2 = \mu_{4x} \div \mu_{2x}^2, \\
\kappa_1 = 2\beta_2 - 3\beta_1 - 6, \\
\kappa_2 = (\beta_2+3)^2\beta_1 \div \{(8\kappa_1 + 12\beta_1 + 48)\kappa_1\}.
\end{cases}
$$

$$
(27)\quad\begin{cases}
\beta_1 = \mu_{3t}^2 \div \mu_{2t}^3, \\
\beta_2 = \mu_{4t} \div \mu_{2t}^2, \\
\kappa_1 = 2\beta_2 - 3\beta_1 - 6, \\
\kappa_2 = (\beta_2+3)^2\beta_1 \div \{(8\kappa_1 + 12\beta_1 + 48)\kappa_1\}.
\end{cases}
$$

兩式中以式 (26) 之應用較廣，蓋計算以組距為單位各級主要動差較為方便故也.

用公式 (14) 求 μ_{2x}, μ_{3x} 及 μ_{4x} 之值固已較便,然而未臻極致,蓋該

* 高次 β 之定義為 $\beta_{2r+1} = \frac{\mu_3\mu_{2r+3}}{\mu_2^{r+3}}$, $\beta_{2r} = \frac{\mu_{2r+2}}{\mu_2^{r+1}}$.

公式仍可使之簡化，運算手續更能使之輕鬆．因此將上列之公式(14)及(26)更加以進一步之推演．經此番推演後，所有各步結果均可用連續運算法完成之(參閱第一章第五節3)．茲將計算 $\beta_1, \beta_2, \kappa_1$ 及 κ_2 之步驟分列如下：

(28)
$$
\begin{cases}
N = \Sigma f, & \mu_{2x} = -\nu_{1x}^2 + \nu_{2x} - 0.083333,^{*} \\
\nu_{4x} = \dfrac{1}{N}\Sigma t^4 f, & F = \mu_{2x}^2, \\
\nu_{3x} = \dfrac{1}{N}\Sigma t^3 f, & G = \mu_{2x} F, \\
\nu_{2x} = \dfrac{1}{N}\Sigma t^2 f, & \mu_{3x} = \nu_{3x} + A\nu_{2x} + C\nu_{1x}, \\
\nu_{1x} = \dfrac{1}{N}\Sigma t f, & \beta_1 = \mu_{3x}^2 \div G, \\
A = -3\nu_{1x}, & \beta_2 = (\nu_{4x} + B\nu_{3x} + D\nu_{2x} + E\nu_{1x} \\
B = -4\nu_{1x}, & \qquad -0.5\mu_{2x} - 0.0125) \div F, \\
C = 2\nu_{1x}^2, & \kappa_1 = 2\beta_2 - 3\beta_1 - 6, \\
D = 6\nu_{1x}^2, & H = (8\kappa_1 + 12\beta_1 + 48)\kappa_1, \\
E = A\nu_{1x}^3, & \kappa_2 = (\beta_2 + 3)^2 \div H \times \beta_1.
\end{cases}
$$

上列各步運算手續可以自明，無須加以解釋；惟有數點應當注意．(i) 宜按公式次序進行運算；(ii) 宜用 N 之反商分別乘以 $\Sigma t^4 f$, $\Sigma t^3 f$, $\Sigma t^2 f$ 及 $\Sigma t f$ 以求 $\nu_{4x}, \nu_{3x}, \nu_{2x}$ 及 ν_{1x} 之值，並且各數值之小數應取至六位；(iii) 求得各 ν 值之後，宜將小數位定好，即是使記次輪，記值輪及打字盤上的小數各爲6位，12位及6位；(參閱第一章第五節例1運

* 薛伯校正數 (Sheppard's correction).

算手續 b); (iv) 宜將此時記值輪上面 ν_{1x} 值登記在打字盤上，分別乘以 3,4 及 ν_{1x} 以求 A, B 及 ν_{1x}^2 之值；(v) 宜將此時記值輪上面 ν_{1x}^2 之值登記在打字盤上，並分別乘之以 2, 6 及 A 以求 C, D 及 E 之值；(vi) 宜將打字盤上面之 ν_{1x}^2 反登在記值輪上面(何故?)，並加 ν_{2x} 之值於其上(如須用薛伯氏校正數時，更應減去 0.083333)，以求 ν_{2x} 之值；(vii) 宜將此時記值輪上面 μ_{2x} 之值登記在打字盤上，並依次乘以 μ_{2x} 及 F 以求 F 及 G 之值；(viii) 如此便可依次求得 β_1, β_2, κ_1 及 κ_2 之值.

第八節　β_1, β_2, κ_1 及 κ_2 之計算

為說明計算數值 β_1, β_2, κ_1 及 κ_2 各步之手續起見，我們舉一實在的例題如下：

例　試根據下表裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

下表第一至第七行之運算手續無待說明，惟第八行之列入須加以解釋. 第八行為便於說明如何驗對第四，第五，第六及第七行裏之各數值有無錯誤而設，實際運算無列入之必要. 因 $x \cdot x^3f = x^4f$, 故第七行與第八行裏各對應值應當完全一致；但求該兩行裏數值所用之運算手續截然不同. 一用第三行裏各數值乘以第六行裏各對應數值得之者，一用第二行裏各數值乘以第三行裏各對應數值的四次方得之者. 若第七及第八兩行各變對應值完全一致，則得斷定第四至第七各行裏算得之數值無誤；反之若遇有不合者，分明第四至第七各行裏之對應數值必有舛錯，因可據以校正之. 此法可用以代替沙利驗算法(Charlier's Check). 經過此番核對無誤後，我們便可將表中各行裏數值分別相加以求其總

和，並應用公式 (28) 以求 β_1、β_2、κ_1 及 κ_2 之值．

表 14. β_1、β_3、κ_1 及 κ_2 之計算(論據由參考書 25)

體重 X (1)	頻數 f (2)	助變數 x (3)	xf (2)×(3) (4)	x^2f (3)×(4) (5)	x^3f (3)×(5) (6)	x^4f (3)×(6) (7)	x^4f (2)×(3)^4 (8)
29.5	1	−5	−5	25	−125	625	625
33.5	14	−4	−56	224	−896	3,584	3,584
37.5	56	−3	−168	504	−1,512	4,536	4,536
41.5	172	−2	−344	688	−1,376	2,752	2,752
45.5	245	−1	−245	245	−245	245	245
49.5	263	0	0	0	0	0	0
53.5	156	1	156	156	156	156	156
57.5	67	2	134	268	536	1,072	1,072
61.5	23	3	69	207	621	1,863	1,863
65.5	3	4	12	48	192	768	768
總計	1,000	……	−447	2,365	−2,649	15,601	15,601

計算：

$\frac{1}{N}=$	0.001000	$\mu_{2x}=$	2.165191
$\nu_{4x}=$	15.601000	$F=$	4.688052
$\nu_{3x}=$	−2.649000	$G=$	10.150528
$\nu_{2x}=$	2.365000	$\mu_{3x}=$	0.343836
$\nu_{1x}=$	−0.447000	$\beta_1=$	0.011647
$A=$	1.341000	$\beta_2=$	2.896748
$B=$	1.788000	$\kappa_1=$	−0.241445
$C=$	0.399618	$H=$	−11.156740
$D=$	1.198854	$\kappa_2=$	−0.036300
$E=$	0.267944		

結果：　$\beta_1 = 0.011647$　　　$\kappa_1 = -0.241445$

　　　　$\beta_2 = 2.896748$　　　$\kappa_2 = -0.036300$

第九節　薛伯校正數

當由頻數分配計算動差時，我們所根據之假設爲：在一個組段裏之各個觀察值係以組段之組值代替．這就是用一個帶有幾分虛構性之組值去代替實在觀察值．很顯然的，如果所分之組很粗率的話，可以使所得之結果與事實大相徑庭，並且用算學能夠證明分組可以引入一貫之誤差，卽所謂之分組誤差 (grouping error)．這個誤差對於二級及四級之動差有影響，其與一級及三級之動差 (以平均數爲原點者) 無關．薛伯首先注意及此，並謂應由 $(-\nu_{1x}^2 + \nu_{2x})$ 內減去 0.083333 以得 μ_{2x} 之校正數，由 $(\nu_{4x} + B\nu_{3x} + D\nu_{2x} + E\nu_{1x})$ 內減去 $(0.5\mu_{2x} + 0.0125)$ 以得 μ_{4x} 之校正數．職是之故，數值 -0.083333 及 $-(0.5\mu_{2x} + 0.0125)$ 通常被稱之爲薛伯校正數 [Sheppard's correction，參閱公式 (28) 內附有 * 號之項]．

此種校正數來源之推演見第四章第九節．現在我們祇須瞭解一點：

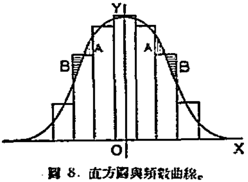


圖 8．直方圖與頻數曲線．

何故有些動差須加校正，並且有些無須呢？為瞭解此點，爰繪上圖．

上圖係代表實在頻數分配之圓滑曲線及其直方圖．因為動差由於以直方圖代表之頻數分配而計算者，故所得之結果不能與實在動差完全一致，就一個矩形說，我們是拋棄了曲線下面帶有陰影部份之面積 A，並且補添了曲線上面帶有陰影部分之面積 B．普通言之，B 值向是大於 A 值．如果我們將這些小塊面積想作與三角形面積相近，並依其在平均數之左右而定三角形底邊之正負，則矩形在 $\bar{X}$ 左者 B 值多於 A 值之數為負，在其右者 B 值多於 A 值之數為正，所有多於之數，不論為正為負，均與動差 $\frac{1}{N}\Sigma(X_i-\bar{X})^r f_i$ 有關．當 r 為奇數時，因正負號關係，在頻數分配之全距中，B 值多於或少於 A 值之數恰好正負相消；但在偶次動差情形中，所有多於 B 值之數均變成正數，如此集在一起的錯誤是很大的，並且影響各動差最後之結果也是很大的．

薛伯校正數僅僅應用於鐘形分配上，纔生效力，其對於 J 形或 U 形的分配不宜採用．

茲為討論薛伯校正數影響前例所得之結果起見，特舉下例．

例　試根據表 14 裏所列之論據，計算 β_1、β_2、κ_1 及 κ_2（此論據為鐘形分配，故可應用薛伯校正數）．

計算：

$\frac{1}{N}$，ν_{4x}、ν_{3x}、ν_{2x}、ν_{1x}、A，B，C，D，E 及 μ_{3x} 之數值與前例同；因此續前算得之數值，我們有

$$\mu_{2x}=2.081858 \qquad \beta_2=2.890238$$

$$F=4.334133 \qquad \kappa_1=-0.258830$$

$G=9.023049$ $H=-11.928591$

$\beta_1=0.013102$ $\kappa_2=-0.038108$

結果:

$\beta_1=0.013102$ $\kappa_1=-0.258830$

$\beta_2=2.890238$ $\kappa_2=-0.038108$

可見算得之結果與前例算得者,有些許之出入.在初步統計分析中,此種校正手續可省;惟論據本屬於皮爾生頻數曲線某型,往往因此些許之差以致改隸他型.故在用皮爾生頻數曲線方程式配合論據情形中,此些許之差不容忽視.

第十節 偏態與峯態

多數頻數分配是對稱的 (symmetrical), 但有些頻數呈偏斜狀態 (skew, asymmetrical) (見圖 9); 或呈峯錐狀態 (kurtic) (見圖 10). 很顯然的,平均數及標準差對於帶有偏態或峯態特徵之頻數分配,不能盡描寫之致,勢須採用其他量數描寫牠們.

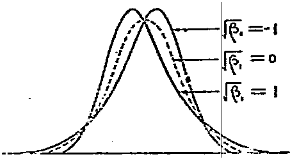


圖 9. 常態與偏態分配。

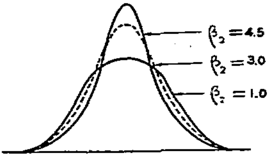


圖 10. 常態與峯態分配。

不對稱頻數分配之偏斜情形可用偏度係數(coefficient of skewness)測量之,此偏度係數為用標準差去除平均數及衆數之差得之者,即

(29) $$偏度係數=\frac{\bar{X}-\hat{X}}{\sigma_X},$$

恰巧由表示偏態分配之皮爾生型 III 曲線得以證明

(30) $$\frac{\bar{X}-\hat{X}}{\sigma_X}=\frac{\pm\sqrt{\beta_1}}{2}.$$

故頻數分配之偏斜情形得 $\sqrt{\beta_1}$ 測量之. 如果頻數分配為對稱者,則 $\sqrt{\beta_1}$ 之值為零;其長尾拖向平均數右方,宜取正號之 $\sqrt{\beta_1}$, 即 $+\sqrt{\beta_1}$; 其拖向平均數左方者,宜取負號之 $\sqrt{\beta_1}$, 即 $-\sqrt{\beta_1}$. 因此頻數曲線長尾在右方者曰正偏態曲線 (positively skew curve); 反之,在左方者曰負偏態曲線 (negatively skew curve, 參閱圖 9).

圖 10 顯示各扁平度數不同之曲線, 曲線之具有此特徵者, 叫做帶有峯態曲線. β_2 為測量此峯度 (kurtosis) 之常數. 用微積分方法可證明,在常態曲線情形中, $\beta_2=3$. 如果有一頻數分配: $\beta_2<3$ 者,稱做低闊峯分配 (platykurtic distribution); 如果 $\beta_2>3$ 者,稱之曰高狹峯分配 (leptokurtic distribution).

β_1 爲以標準差爲單位之三級動差，並且 β_2 爲以標準差爲單位之四級動差（學者驗證之），二者均爲抽象數，如此不同分配之偏度及峯度均可用 β_1 及 β_2 度量之。

第十一節　機率

1. 意義及種類　機率 (probability) 對於統計理論之關係極爲重要，但其包括之問題甚廣，在此不能作詳盡之論述，現在僅將有關機率問題之基本原則時常用之於統計學上者，加以研究。機率可分爲兩類：一爲通常所說的事前機率(priori probability)，一爲經驗機率(empirical probabitity)。

a. 事前機率　假設將所有事件成功及失敗之項目分爲彼此互斥 s 種，每種發現之機會均等；並假設成功項目有 X 種，則在一單獨試驗中成功之機率爲 X/s 此機率卽所稱之事前機率。事前機率者，乃是預先卽知有關影響事件發現等等實事之意；換言之，此問題可以用抽象方法解決之。下列各問題乃是事前機率問題：一盒內有白球 4 枚，紅球 5 枚，問取得同色兩球之機率爲何？一枚銅元連擲七次，正面向上至少三次之機率爲何？由箱中任取一收音機，箱中有收音機 100 具，其中 20 具有毛病，實在取得 2 具有毛病收音機之機率爲何？

b. 經驗機率　設有一組事件，其成功或失敗之項目事前莫由知之，因此其成功之機率亦無從推得。當遇此種情形時，須藉實驗及觀察以估計其機率，此機率卽所謂經驗機率。假設在 s 次試驗中觀察到某事件發現 X 次，比值 X/s 叫做事件成功之相對頻數 (relative frequency of success). 若所取之 s 值非常之大,此相對頻數之極限值叫做事件在一單獨試驗中之成功機率.用符號表之為

$$\lim_{s\to\infty}\frac{X}{s}=p$$

在統計應用上面,X/s 之極限值通常不能測定,但觀察之相對頻數 X/s (若 s 很大) 常常供給吾人以有價值之估計.例如,根據美國經驗死亡表 (American Experience Mortality Table), 年達 60 歲之 57,917 人中於一年內有 1,546 人死亡.如此,保險公司可取相對頻數 $1,546\div 57,917=0.026693$ 當做年 60 歲者不能生存至次一年之機率.

2. 公式及定理　重要之公式及定理與機率有關,並且在統計理論上常常應用者,撮要敍述於下.

定理 1. n 種不同事物,每次取出 r 種來,其不同之排列方法 $P(n, r)$ 可由下式求之:

$$P(n, r)=n(n-1)(n-2)\cdots\cdots(n-r+1).$$

系. 假設 n 種事物中有些相同者,卽如有 n_1 種皆屬於 T_1 型,有 n_2 種皆屬於 T_2 型,……,n_k 種皆屬於 T_k 型.於是不同之排列方法為

$$\frac{n!}{n_1!\ n_2!\cdots\cdots,\ n_k!},$$

彼處 $\sum_{1}^{r}n_1=n$, 符號 $n!=n(n-1)(n-2)\cdots\cdots3.2.1$ 讀做階乘 n (factorial n).

定理 2. n 種不同事物,每次取出 r 種來,其不同組合方法 $C(n, r)$ 可由下式求之:

$$C(n, r)=\frac{n(n-1)(n-2)\cdots\cdots(n-r+1)}{r!}$$

$$=\frac{n!}{r!(n-r)!}.$$

當 $r>n$ 時，我們知道 $C(n, r)$ 爲零；當 $r=n$ 時，其等於 1.

定理 3. n 種不同事物，每次取 1, 2……或 n 種排列之，其排列方法總共等於 2^n-1.

證：因 $C(n, r)$ 爲二項式 $(x+y)^n$ 展開式第 $(r+1)$ 項之係數，如此

$$(x+y)^n=x^n+C(n, 1)x^{n-1}y+C(n, 2)x^{n-2}y^2+\cdots\cdots$$

$$+C(n, r)x^{n-r}y^r+\cdots\cdots+y^n.$$

讓 $x=y=1$，我們有

$$2^n-1=C(n, 1)+C(n, 2)+C(n, 3)+\cdots\cdots+C(n, n).$$

在某一時間事件中之一種發現，其他任意一種將不能同時發現，則稱此等事件曰彼此互斥 (mutually exclusive). 某一種事件之發現影響他種事件之發現，則稱此二事件曰彼此相干 (dependent)；反之，一事件之發現不影響他種事件之發現，則稱此二事件曰彼此無干 (independent).

如果 p 爲在一單獨試驗中事件發現之成功機率，q 爲此事件在一單獨試驗中之失敗機率；於是 $p+q=1$.

定理 4. 各個事件彼此互斥其共同發現之機率爲各個機率相加之和

定理 5. 各個彼此無干事件在同一時間發現，其發現之機率爲各個機率相乘之積.

定理 6. 假設二事件是彼此相干的. p_1 代表第一種事件 E_1 發現之機率,並且 p_2 為繼 E_1 之後第二種事件 E_2 發現之機率.於是兩種事件發現之機率將為 p_1p_2. 此定理當可推廣至多種事件.

定理 7. 讓 p 為在一單獨試驗中事件發現之機率,並且 $q(=1-p)$ 為事件未發現之機率,當做試驗時 p 仍不變;於是在 s 次試驗中事件發現確為 X 次之機率為 $B(X)$, 此處 $B(X)$ 為 $(q+p)^s$ 展開式之第 $(X+1)$ 項. 即

$$(31)\qquad B(X)=C(s,X)p^Xq^{s-X}$$

$$=\frac{s!}{X!(s-X)!}p^Xq^{s-X}.$$

證: 由定理 5 得知事件依其特定次序發現 X 次未發現 $(s-X)$ 次之機率為 p^Xq^{s-X}. 因特定次序之方法有 $C(s,X)$ 或 $C(s,s-X)$ 種, 並且此方法係自由排列,而且又彼此互斥,所以由定理 4 得知所求之機率為 $C(s,X)p^Xq^{s-X}$.

系 1. 在 s 次試驗中事件至多發現 X 次之機率為 $\sum\limits_0^X B(X)$.

系 2. 在 s 次試驗中事件至少發現 X 次之機率為 $\sum\limits_X^s B(X)$.

此二系可直接利用定理 4 證明之.

定理 8. 在一組 s 次試驗中,每次之成功機率 p 不變, 如此當 s 無限制的加大時, 相對頻數 X/s 以 p 為其極限值之機率 Q_ε 與 1 相近.此定理將在第八章第四節證明之.

此定理可用另一種方法述之如下: 當 s 無限制的加大時,差量 $\frac{X}{s}$

$-p$ 之值大於既定正數 ϵ 之機率 $P_{\varepsilon}(=1-Q_{\varepsilon})$ 與零相近.

上列之定理通常稱爲伯敖里定理 (Theorem of Bernoulli).

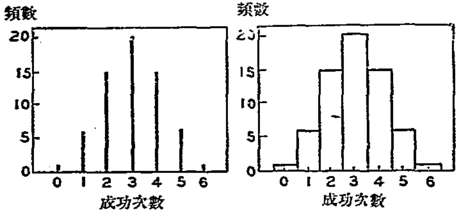
3. 伯敖里分配 間斷數值之頻數分配與表15第二行各項成比例者,通常稱之爲伯敖里分配,或稱之爲二項分配 (binomial distribution).

a. 頻數 假設在一單獨試驗中事件發現之機率爲 p, 其不發現之機率爲 q. 如是,若取一包含 s 次試驗之樣本,則事件發現 0, 1, 2, 3,…s 次之頻數與 $(q+p)^s$ 展開式之各項成比例. 即 $(q+p)^s$ 展開式之各項爲理論相對頻數. 如果我們取 N 組 s 次試驗,理論絕對頻數可用二項式 $N(q+p)^s$ 展開式之各項表之. 我們選取之 N 應使展開式之各項爲整數.

b. 直方圖 二項分配可用直方圖表之,此圖即以 $X=0, 1, 2,\cdots\cdots s$ 爲底邊中點並以與二項展開式各項成比例之數值爲高,各竪一矩形而成. 變量 X 爲事件發現之次數,或曰成功之次數;二項式各項數值爲事件發現 X 次之相對理論頻數,或曰事件成功之機率.

一羣間斷數值之相對頻數固然用縱線而不用面積表示較爲合理,如圖 11*a* 所示. 然而,因爲每個矩形之底爲 1, 矩形之高仍是代表面積,並且用面積代表頻數有助於分析之工作,故相對頻數遂用代表面積之矩形表示,如圖 11*b* 所示.

如果我們將矩形想做相對頻數或機率, 則其和等於 1; 如果將其想做絕對頻數,則直方圖總面積等於 N. 如此投擲 6 枚銅元,理論之絕對頻數爲 $64(\frac{1}{2}+\frac{1}{2})^6$ 展開式之各項,即爲 1, 6, 15, 20, 15, 6 及 1, 其和爲 64.



a. 用縱線表示頻數　　　　*b*. 用矩形表示頻數

圖 11. 二項分配 $N(q+p)^s$: $N=64, q=p=\frac{1}{2}, s=6$.

c. 平均數、標準差、偏度及峯度　我們講過：如果在N個樣本中，每個由 s 次試驗組成，並且每次試驗成功之機率為 p. 則 $N(q+p)^s$ 展開式之各項為供給成功之期望頻數，彼處 $q=1-p$. 現在我們進一步推求這個期望頻數分配之各次動差. 我們暫以首項為原點計算首四級動差如下：

由定義　$$\nu_1=\frac{\Sigma XB(X)}{\Sigma B(X)},\qquad \nu_2=\frac{\Sigma X^2B(X)}{\Sigma B(X)},$$

$$\nu_3=\frac{\Sigma X^3B(X)}{\Sigma B(X)},\qquad \nu_4=\frac{\Sigma X^4B(X)}{\Sigma B(X)}.$$

彼處 X 表示事件成功次數，$B(X)$ 表示與 X 成對應之機率，自然其為理論相對頻數. 下表（表 15）乃是一個二項分配頻數表. 很顯然的，第二行之和 $\Sigma B(X)$ 為 1. 於是

$$\nu_1=\sum_{X=0}^{s}XB(X)$$

$$=\sum_{r=0}^{s}\frac{rs!}{r!\,(s-r)!}\,p^rq^{s-r}$$

$$=\sum_{r=0}^{s}\frac{s!}{(r-1)!\,(s-r)!}\,p^rq^{s-r}$$

$$=sp\sum_{r=0}^{s}\frac{(s-1)!}{(r-1)!\,(s-r)!}\,p^{r-1}q^{s-r}$$

$$=sp(q+p)^{s-1}$$

$$=sp.$$

表 15.　二項分配 $(q+p)^s$

X	$B(X)$	$XB(X)$
0	q^s	0
1	spq^{s-1}	spq^{s-1}
2	$\frac{s!}{2!\,(s-2)!}p^2q^{s-2}$	$\frac{s!}{(s-2)!}p^2q^{s-2}$
3	$\frac{s!}{3!\,(s-3)!}p^3q^{s-3}$	$\frac{s!}{2!(s-3)!}p^3q^{s-3}$
⋮	⋮	⋮
r	$\frac{s!}{r!\,(s-r)!}p^rq^{s-r}$	$\frac{s!}{(r-1)!(s-r)!}p^rq^{s-r}$
⋮	⋮	⋮
s	p^s	sp^s
總計	$\Sigma B(X)$	$\Sigma XB(X)$

$$
\begin{aligned}
\nu_2 &= \sum_{X=0}^{s} X^2 B(X) \\
&= \sum_{r=0}^{s} \frac{r^2 s!}{r!\,(s-r)!}\, p^r q^{s-r} \\
&= \sum_{r=0}^{s} \frac{[r(r-1)+r]s!}{r!\,(s-r)!}\, p^r q^{s-r} \\
&= \sum_{r=0}^{s} \frac{s!}{(r-2)!\,(s-r)!}\, p^r q^{s-r} + \sum_{r=0}^{s} \frac{s!}{(r-1)!\,(s-r)!}\, p^r q^{s-r} \\
&= s(s-1)p^2 \sum_{r=0}^{s} \frac{(s-2)!}{(r-2)!\,(s-r)!}\, p^{r-2} q^{s-r} + \nu_1 \\
&= s(s-1)p^2 (q+p)^{s-2} + \nu_1 \\
&= s(s-1)p^2 + sp.
\end{aligned}
$$

$$
\begin{aligned}
\nu_3 &= \sum_{X=0}^{s} X^3 B(X) \\
&= \sum_{r=0}^{s} \frac{r^3 s!}{r!\,(s-r)!}\, p^r q^{s-r} \\
&= \sum_{r=0}^{s} \frac{[r(r-1)(r-2)+3r^2-2r]s!}{r!\,(s-r)!}\, p^r q^{s-r} \\
&= \sum_{r=0}^{s} \frac{s!}{(r-3)!\,(s-r)!}\, p^r q^{s-r} + 3\nu_2 - 2\nu_1 \\
&= s(s-1)(s-2)p^3 \sum_{r=0}^{s} \frac{(s-3)!}{(r-3)!\,(s-r)!}\, p^{r-3} q^{s-r} + 3\nu_2 - 2\nu_1 \\
&= s(s-1)(s-2)p^3 (q+p)^{s-3} + 3\nu_2 - 2\nu_1 \\
&= s(s-1)(s-2)p^3 + 3s(s-1)p^2 + sp.
\end{aligned}
$$

$$\nu_4 = \sum_{X=0}^{s} X^4 B(X)$$

$$= \sum_{r=0}^{s} \frac{r^4 s!}{r!(s-r)!} p^r q^{s-r}$$

$$= \sum_{r=0}^{s} \frac{[r(r-1)(r-2)(r-3) + 6r^3 - 11r^2 + 6r]s!}{r!(s-r)!} p^r q^{s-r}$$

$$= s(s-1)(s-2)(s-3)p^4 \sum_{r=0}^{s} \frac{(s-4)!}{(r-4)!(s-r)!} p^{r-4} q^{s-r} + 6\nu_3 - 11\nu_2 + 6\nu_1$$

$$= s(s-1)(s-2)(s-3)p^4 + 6\nu_3 - 11\nu_2 + 6\nu_1$$

$$= s(s-1)(s-2)(s-3)p^4 + 6s(s-1)(s-2)p^3 + 7s(s-1)p^2 + sp.$$

因之由式(14)我們有首四級主要動差

$$\mu_1 = 0, \qquad \mu_2 = spq,$$

$$\mu_3 = spq(q-p), \qquad \mu_4 = spq[1 + 3(s-2)pq].$$

記得 $\bar{X} = \nu_1$, $\sigma_X = \sqrt{\mu_2}$, $\beta_1 = \mu_3^2/\sigma_X^6$ 及 $\beta_2 = \mu_4/\sigma_X^4$；我們最後有

4)
$$\begin{cases} \bar{X} = sp, \\ \sigma_{sp} = \sqrt{spq} \text{ 或者 } \sigma_p = \sqrt{\dfrac{pq}{s}}, \\ \beta_1 = (q-p)^2/(spq), \\ \beta_2 = 1/spq - 6/s + 3. \end{cases}$$

d. 循環公式　循環公式 (recursion formula) 是一很精妙之公式，用以求伯努里分配分之各級主要動差是很省事的．其公式為

$$\mu_{k+1} = pq\left[sk\mu_{k-1} - \frac{d\mu_k}{dq}\right]$$

* 此循環公式係科瑞格 (A. T. Craig) 所發明，其證明見 Bulletin of the American Mathematical Society, Vol. 40, pp. 262—264.

我們知道 $\mu_0=1$, $\mu_1=0$, 如此這個公式當 $k\geq 1$ 時,可以應用.

$$k=1,\qquad \mu_2=pq(s\mu_0-0)$$

$$=spq.$$

$$k=2,\qquad \mu_3=pq\left(2s\mu_1-\frac{d\mu_2}{dq}\right)$$

$$=pq[0-(s-2sq)]$$

$$=spq(2q-1)$$

$$=spq(q-p).$$

$$k=3,\qquad \mu_4=pq[3s^2pq+s-6sq-6sq^2]$$

$$=spq[1+3spq-6pq]$$

$$=spq[1+3(s-2)pq].$$

4. 卜瓦松變態級數　設 p 值很小,則 $1-q$ 近於 1; 再設 s 之值相當大,如此得使 $s!$ 及 $(s-r)!$ 用史德齡近似值*替換之. 於是

$$B(X)=\frac{s!}{X!(s-X)!}p^Xq^{s-X}$$

$$=\frac{s^{s+\frac{1}{2}}e^{-s}p^Xq^{s-X}}{X!(s-X)^{s-X+\frac{1}{2}}e^{-s+X}}$$

$$=\frac{(sp)^Xe^{-X}(1-p)^{s-X}}{X!\left(1-\frac{X}{s}\right)^{s-X+\frac{1}{2}}}$$

我們知道,當 X 之值很小而 s 之值很大時.

---

* 史德齡之近似公式 (Stirling's Approximation formula) 為

$$n!=\sqrt{2\pi}\,n^{n+\frac{1}{2}}e^{-n}\left(1+\frac{1}{12n}+\cdots\cdots\right),$$ n 為相當大的數.

$$\left(1-\frac{X}{s}\right)^{s-X+\frac{1}{2}} \doteq \left(1-\frac{X}{s}\right)^{s} \doteq e^{-X},$$

並且
$$(1-p)^{s-X} \doteq (1-p)^{s} \doteq \left[(1-p)^{\frac{1}{p}}\right]^{sp} \doteq e^{-sp},$$

總之

(34)
$$B(X)=\frac{m^{X}e^{-m}}{X!},\quad m=sp.$$

此爲卜瓦松變冪級數之公項.如使 $x=0, 1, 2, \cdots\cdots, X$, 則得一級數

$$e^{-m}\left(1+m+\frac{m^{2}}{2!}+\frac{m^{3}}{3!}+\cdots\cdots+\frac{m^{X}}{X!}\right).$$

此即通常所稱之卜瓦松變冪級數 (Poisson exponential series).

其各項各爲在 s 次試驗中發現 $0, 1, 2, \cdots\cdots, s$ 次事件之機率,此級數內只含一個可變常數 m.

我們可以求當 $X=0, 1, 2, \cdots\cdots, s$ 時式 (34) 頻數分配之各級動差.由式 (34) 我們有

$$\nu_0=\sum_{X=0}^{s} B(X)=\sum_{X=0}^{s}\frac{m^{X}e^{-m}}{X!}$$

$$=e^{-m}\sum_{X=0}^{s}\frac{m^{X}}{X!}=e^{-m}e^{m}=1,$$

$$\nu_1=\sum_{X=0}^{s} XB(X)=e^{-m}\sum_{X=0}^{s} X\frac{m^{X}}{X!}$$

$$=me^{-m}\sum_{X=0}^{s}\frac{m^{X-1}}{(X-1)!}=me^{-m}e^{m}=m,$$

$$\nu_2=\sum_{X=0}^{s}X^2B(X)$$

$$=\sum_{X=0}^{s}[X(X-1)+X]B(X)$$

$$=m(m+1),$$

$$\nu_3=m(m^2+3m+1),$$

及 $$\nu_4=m(m^3+6m^2+7m+1).$$

於是 $$\mu_1=0,$$

$$\mu_2=m(m+1)-m^2=m.$$

$$\mu_3=m,$$

$$\mu_4=3m^2+m.$$

因之我們有

$$\bar{X}=m=sp,$$

$$\sigma_X=\sqrt{m},$$

$$\beta_1=\frac{1}{m},$$

$$\beta_2=3+\frac{1}{m}.$$

由上列各式可見，平均數，變差及 μ_3 均等於 m. 此爲卜瓦松變霖級數之特點.

第十二節　常態曲線

1. 來源　許多頻數分配可用所謂常態曲線描寫之，所以常態曲線爲頻數曲線中之最重要者. 常態曲線之方程式爲

$$(35)\qquad y=\frac{N}{\sigma_x\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_x^2}},\qquad x=\frac{X-\bar{X}}{h}.$$

此曲線本係法國著名算學家狄牧夫 (De Moivre, 1667—1754) 約在二百年以前所發明者,因此有時稱此曲線為狄牧夫曲線;其與機率理論有密切關係,因此又有時稱之為常態機率曲線.

常態曲線對於實際問題之應用,則自天文學家雷普拉斯 (Laplace, 1749—1827) 及高斯 (Gauss, 1777—1855) 始. 二人均係個別研究,對於狄氏之創作,事前殆未之知也.雷高二氏曾發現用以描寫物理上之觀察誤差甚為美滿,因此之故牠又被叫做雷普拉斯曲線,高斯曲線或差誤常態曲線 (normal curve of error). 所謂差誤者,係指觀察數值與實在數值之離差而言. 自此以後,學者更發現用牠描寫關於生物、教育及社會範圍內之資料,也是十分恰意.

2. 標準形式　式(35)之常數 N, $\bar{X}$ 及 σ_x 因頻數分配之不同而異,此等常數通常稱做參數 (parameter); 至若其中之 π 及 e ($\pi=3.14159\cdots$, $e=2.71828\cdots\cdots$見附錄 I) 乃常數 (constant) 也,其不因客觀條件之變而變.參數為測定曲線形狀,位置及離勢者,與曲線之基本性質無關.研究曲線特性最好將方程式內之參數消去.換言之,即使常態曲線方程式內不含有參數,亦即使曲線下總面積為 1, 取平均數為原點, 並且用標準差當做橫坐標量度之單位. 用算學話講, 即是使式 (35) 中之 $N=1$, $t=(X-\bar{X})/\sigma_X$, 將參數消去之. 若以 $\phi(t)$ 表示消去參數後之結果函數,則有

$$(36)\qquad \phi(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}.$$

此爲常態曲線之標準形式 (normal curve in standard form).

正像方程式 (35) 以 (x, y) 表示點之坐標一樣,用 t 表示曲線 (36) 之橫坐標, $\phi(t)$ 表示縱坐標. (x, y) 與 (t, ϕ) 兩個坐標系之關係爲

(37) $$X=\bar{X}+\sigma_X t,$$

及

(38) $$y=\frac{N}{\sigma_x}\phi(t)$$

$$=\frac{Nh}{\sigma_X}\phi(t), \quad h \text{ 爲組距.}$$

3. 縱坐標及面積 使常態曲線下之面積及其縱坐標不因論據之不同而異其值,爲將方程式 (35) 寫成標準形式 (36) 的理由之一,曲線 (36) 下之面積及其縱坐標列入附錄 III 表 B 中.由式 (36) 可見: $\phi(-t)=\phi(t)$, 即是 t 值爲負或爲正之縱坐標相等,亦即此曲線對於直線 $t=0$ 說是對稱的,所以列表時祇列 t 值爲正數之各縱坐標 $\phi(t)$ 足矣.方程式

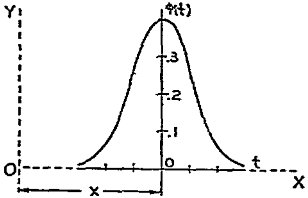


圖 12. 曲線 $\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$

36 之圖可畫出之，如圖 12 所示。

此曲線在其兩端與橫軸極相接近，但永不與之相交，卽其與橫軸在無窮遠點相切。雖然在實際上永不會遇到與無限大橫坐標成對應之觀察值，然而不能斷言在樣本所由來之宇宙內亦絕對不具有此特性之數值。故無限大情形在學理研究上甚有裨益；進而言之，在實際上用無限大全距表示觀察值之分配亦無困難，因在 $t=\pm 3$ 之外的面積僅爲萬分之 27，換言之，觀察值小於平均數減以或加以 3 倍標準差之發現機會甚小。

當 $t=a$ 及 $t=b$ 時曲線下之面積若以 $\int_a^b$ 表之，如此當 $t=0$ 及 $t=1$ 時之面積得用 $\int_0^1$ 表之，其等於 .3413 因爲曲線下總面積爲 1，故在直線 $t=0$ 兩側之面積均等於 0.5，故列表時只須由 0 至 t 各正數之面積 $\int_0^t$ 足矣。由 $t=-1$ 至 $t=0$ 之面積等於由 $t=0$ 至 $t=1$ 之面積，用符號敍述之爲 $\int_{-1}^0=\int_0^1$.

任何面積之值可由已列之表用加減法得之。例如，我們需要在 $t=-2$ 以下之面積，卽求 $\int_{-\infty}^{-2}$ 之面積。因爲由 $-\infty$ 至 -2 之面積等於由 0.5 減以由 -2 至 0 之面積，又因爲由 -2 至 0 之面積與由 0 至 2 之面積相同；於是

$$\int_{-\infty}^{-2}=.5-\int_{-2}^{0}=.5-\int_{0}^{2}=.5-.4772=0.0228.$$

4. 求縱坐標及面積之公式 將方程式 (36) 之右方展開我們得寫求縱坐標 $\phi(t)$ 之公式如下：

(39) $$\phi(t)=\frac{1}{\sqrt{2\pi}}\left[1-\frac{t^2}{2}+\frac{1}{2!}\left(\frac{t^2}{2}\right)^2-\frac{1}{3!}\left(\frac{t^2}{2}\right)^3+\cdots\cdots\right]$$

$$=.39894228-.19947114t^2-.04986779t^4$$

$$-.00831130t^6+.00103891t^8-.00010389t^{10}$$

$$+.00000866t^{12}-.00000062t^{14}+\cdots\cdots$$

將上式右方逐項積分之,則得求曲線(36)下面積之公式:

(40) $$\int_0^t\phi(t)dt=.39894228t-.06649038t^3+.00997356t^5$$

$$-.00118733t^7+.00011543t^9-.00000944t^{11}$$

$$+.00000067t^{13}-.00000004t^{15}-\cdots\cdots$$

無論 t 之值爲何,式(39)之右方是一收斂級數,加至某項之值與實在值之差小於所略去各項首項數值之絕對值. t 值若是不大,該級數收斂甚快,因之當 $t\leq\sqrt{2}$ 用以求面積 $\int_0^t\phi(t)dt$ 之值時,可以得到滿意結果.

若 t 之值大,用式(40)計算面積,需要計算的項數太多,頗不方便,所以最好能用 t 之降級數表示欲求之面積.爲達到此目的,我們寫

$$\frac{1}{\sqrt{2\pi}}\int_0^t e^{-t^2/2}dt=\int_0^\infty e^{-t^2/2}dt-\frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-t^2/2}dt$$

$$=0.5-\frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-t^2/2}dt$$

但 $$\frac{1}{\sqrt{2\pi}}\int_0^t e^{-t^2/2}dt=0.5-\frac{1}{\sqrt{2\pi}}\int_t^\infty\frac{1}{t}te^{-t^2/2}dt,$$

用部分積分法(integration by parts)積分之,則有

$$\int_t^\infty \frac{1}{t}\,t e^{-t^2/2}dt = -\frac{1}{t}e^{-t^2/2}\Big|_t^\infty - \int_t^\infty \frac{1}{t^2}e^{-t^2/2}dt$$

$$=\frac{1}{t}e^{-t^2/2} - \int_t^\infty \frac{1}{t^3}\,t e^{-t^2/2}dt;$$

將此種積分手續重複之，我們有

$$(41)\qquad \int_t^\infty e^{-t^2/2}dt = \frac{e^{-t^2/2}}{t}\left\{1-\frac{1}{t^2}+\frac{3}{t^4}-\frac{3\cdot5}{t^6}+\cdots\cdots\right\}.$$

於是我們得到欲求之結果

$$\frac{1}{\sqrt{2\pi}}\int_0^t e^{-t^2/2}dt = \frac{1}{2} - \frac{e^{-t^2/2}}{t\sqrt{2\pi}}\left\{1-\frac{1}{t^2}+\frac{3}{t^4}-\frac{3\cdot5}{t^6}+\cdots\cdots + (-1)^n T_{n+1}+\cdots\cdots\right\},$$

彼處 T_{n+1} 代表第 $(n+1)$ 項數，

$$T_{n+1} = \frac{1\cdot3\cdot5\cdots\cdots(2n-1)}{t^{2n}}.$$

式 (41) 表示之級數叫做半歛級數 (asymptotic, semi-convergent series)，其收歛至最小項以後便放散．公項 T_{n+1} 繼續減少直至 $n\leq t^2/2$，但用部分積分法求出許多項之後至 $n>t^2/2$ 時，T_{n+1} 值便起始增加．自然遇到 $n=t^2/2$ 不要再往下求．由式 (41) 所得之值與實在數值之差小於保留各項末項之值，當 $t>\sqrt{2}$ 時常態曲線下面積之值可用式 (41) 求之．

例：求 $\frac{1}{\sqrt{2\pi}}\int_0^3 e^{-t^2/2}dt$ 之值．

將 $t=3$ 代入式 (41) 內，我們有

$$=.5-\frac{e^{-\frac{9}{2}}}{3\sqrt{2\pi}}\left\{1-\frac{1}{9}+\frac{1}{81}-\frac{15}{729}+\frac{105}{6561}\right\}$$

$$=.5-.00136$$

$$=.49864.$$

此與表上所列者 .49865 之差僅爲 .00001.

常態曲線之縱坐標及其下之面積見附錄 III 表 B.

5. 性質　爲徹底了解常態曲線對於實際上之應用，必須先知其性質。下列標準常態曲線之性質頗關重要，而且極饒興趣，學者宜嫻習之.

a. 平均數、中位數及衆數在 $t=0$ 處重合，並且此時曲線之縱坐標爲最高，其值爲 $\frac{1}{\sqrt{2\pi}}=.3989$.

b. 橫坐標 X 之值可用式(37)由已給之 $\bar{X}$, σ_X 及 t 值求得；縱坐標 Y 之值可用式(38)由已給之 N, σ_x 及 $\phi(t)$ 值求得.

設 $\int_c^d y\,dX$ 表示在曲線(35)下 X 由 c 至 d 之面積，並且設 $\int_a^b \phi(t)dt$ 表示在曲線(36)下 t 由 a 至 b 之面積. 彼處 $t=\frac{X-\bar{X}}{\sigma_X}$, $a=(c-\bar{X})/\sigma_X$ 及 $b=(d-\bar{X})/\sigma_X$; 於是欲求之面積 $\int_c^d y\,dX$ 可由已給之之觀察總數 N, 組距 i 及面積 $\int_a^b \phi(t)\,dt$ 求得，即

$$\int_c^d y\,dX=N\int_a^b \phi(t)\,dt. \tag{42}$$

c. 此曲線在 $t=\pm 1$ 處轉變方向，即 $\left(\pm 1, \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}}\right)$ 爲曲線之反曲點.

d. 標準差 σ_X 比平均差 MD 約小 25%. 精確言之，

$$MD=0.798\,\sigma_X.$$

e. Q_1 及 Q_3 距原點的遠近相等．由定義 Q_3 為當 $\int_{-\infty}^{t}=.75$ 時之 t 值，亦即當 $\int_0^t=.25$ 時之 t 值，此時之 t 值等於 .6744898; Q_1 為當 $\int_t^0=.25$ 時之 t 值．所以 $Q_3=\bar{X}+.6745\,\sigma_X$ 並且 $Q_1=\bar{X}-.6745\,\sigma_X$.

f. 四分位差　　$Q=\frac{Q_3-Q_1}{2}=.6745\,\sigma_X$.

g. 如果變數 X 依常態曲線分配，由一組變量中任意選出若干個來，則將有一半變量在 $\bar{X}-Q$ 與 $\bar{X}+Q$ 之間，此 Q 值即通常所稱之機誤 (probable error).

h. 和 σ_X 有同樣情形，橫軸上之尺度可用 Q 度量之．

i. 常態曲線方程式除用標準形式表示外，物理及天文學家喜用下列之形式表示之：

$$y=\frac{pN}{\sqrt{\pi}}e^{-p^2x^2}, \tag{43}$$

可見此曲線之圖形為尖峭或扁平依參數 p 之值定之；將此式與式 (35) 比較之，得知 $p=\frac{1}{\sqrt{2}\,\sigma_X}$. 很顯然的，$p$ 與 σ_X 成反比例，因此之故 p 被叫做精確指數 (index of precision).

j. 此曲線之 $\beta_1=0$, $\beta_2=3$; 即其不呈偏態與峯態．

6. 配合　關於配合一般頻數曲線之理論及方法容在下章及第五章裏詳述．現在祇將其關於配合常態曲線者加以闡明如下：

a. 目的　我們知道，頻數分配不過為宇宙之一樣本而已．一人由一宇宙抽取一樣本得一頻數分配；另有一人由於同一宇宙另抽一樣本，得到另一頻數分配．此二頻數分配常顯有些微區別，此殆由於抽樣波動使然．配合常態曲線之第一個目的在使觀察之論據理想化，除掉抽樣

波動之不規則情形，得到一理論曲線，以便與觀察頻數或其他同類理論曲線相比較．其第二個目的在得到一有效方法以保存所搜集之論據．如果所得之曲線堪能代表原始論據，則可將原始論據拋棄之，並且由曲線方程式可隨時將其導出之．

b. 假設　方程式(35)配合已給之論據，我們假設

(i) 已給之頻數 N——即用直方圖所代表者——等於曲線下之面積．

(ii) 由已給之觀察分配算得之平均數等於由理論分配算得者．

(iii) 由觀察分配算得之標準差等於由理論分配算得者．

c. 步驟　配合常態曲線步驟可舉表 14 裏所列之論據為例，並說明各步運算之手續如下：

(i) 將已給及算得之統計常數

$$\left.\begin{aligned} N &= 1000 \\ \bar{X} &= 47.712000 \\ \sigma_X &= \ \ 5.771456 \\ \sigma_x &= \ \ 1.442864 \end{aligned}\right\}\text{(參閱本章第八及第九節)}$$

分別代入式(35),(37)(38)及(42)內，則得

$$y=\frac{1000}{1.442864\sqrt{2\pi}}e^{-\frac{x^2}{2(1.442864)^2}}, \tag{44}$$

$$t=\frac{X-47.71200}{5.771456}=-8.266891+.173267X, \tag{45}$$

$$y=\frac{1000}{1.442862}\phi(t)=693.07\,\phi(t). \tag{46}$$

(47) $$\int_{-\infty}^{X} y\,dx = 1000\int_{-\infty}^{t}\phi(t)dt.$$

(ii) 將式(45)中之 X 指定以適當之數值，如表 17 第一行所列者，以計算 t 之各對應值——第三行所列者．計算時宜用第一章第五節例3裏所述之方法，俾節省時間與勞力．於是我們便可應用附錄 III 表 B 查得縱坐標 $\phi(t)$ 之值——第四行所列者．由 $\phi(t)$ 之值，應用式(46)，便可算得其對應之 y 值，即直方圖上之曲線的縱坐標——第五行所列者，如此得將所求之曲線畫出，即畫出表 15 第一及第五兩行之對應值．直方圖及配合之曲線見圖 18.

再應用附錄 III 表 B 可查得面積 $A=\int_{-\infty}^{t}$ 之值，由此可算得 ΔA 之值，並且各 ΔA 之值以 N（本例 $N=1000$）乘之，由式(47)得知乘得之結果為欲求之理論頻數．此等數值分別列入表 16 末三行中．

表 16. 曲線(4)之縱坐標及其下面積之計算（論據由表 14）

體重 X (1)	人數 f (2)	t (3)	$\phi(t)$ (4)	y (5)	$A=\int_{-\infty}^{t}$ (6)	ΔA (7)	$N\cdot\Delta A$ (8)
$-\infty$	—	$-\infty$	.00000	0.00	.0000	.0002	0.2
27.5	1	−3.502	.00087	0.60	.0002	.0023	2.3
31.5	14	−2.809	.00772	5.35	.0025	.0147	14.7
35.5	56	−2.116	.04253	29.48	.0172	.0602	60.2
39.5	172	−1.423	.14495	100.46	.0774	.1553	155.3
43.5	245	−0.730	.30563	211.82	.2327	.2525	252.5
47.5	263	−0.037	.39866	276.30	.4852	.2589	258.9
51.5	153	−0.656	.32171	222.97	.7441	.1672	167.2
55.5	67	1.349	.16060	111.31	.9113	.0631	68.1
59.5	23	2.042	.04960	34.38	.9794	.0175	17.5
63.5	3	2.736	.00945	6.55	.9969	.0031	3.1
∞	—	∞	.00000	0.00	1.0000	.0000	0.0
		—	—	—	—	1.0000	1000.0

7. 修勻 在配合曲線下各組段內之面積叫做理論頻數. 用曲線測定理論頻數之手續叫做論據之修勻,即是用曲線配合法使論據變成圓滑之意. 由修勻頻數分配求得之平均數,標準差及總頻數,必各與由觀察頻數分配求得者相同,因修勻手續中皆假設滿足此條件故也. 如果所配合之曲線對於論據頗覺適切,則由理論頻數分配算得之偏度及峯態與由觀察頻數分配算得者亦不應有顯著之差別.

8. 機率格紙 常態曲線 $\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$ 下之累積面積 $A=\int_{-\infty}^{t}$, 當 t 由 $-\infty$ 變至 ∞ 時, A 由 0 變至 1; 然凡當 $t=\pm3$ (此爲在實際上常見之限度)時, A 由 .00135 變至 .99865. (驗證). 若將 A 看做 t 之函數,則各 (t, A) 之值可以畫出之,而成一圓滑曲線或直線視所用之格紙而定. 若所用之格紙爲算術格紙,則得一肩形曲線; 若用機率格紙,則得一直線. 如此將觀察分配之累計頻數畫在機率格紙上,查看其圖形是否與直線相近,便能決定觀察分配是否近於常態.

問 題 III

1. 證明

   (a) 零次動差恆爲 1.

   (b) 偶次動差恆爲正數,但奇次者可爲正數,爲負數或爲零.

   (c) 如果頻數分配爲對稱,以平均數爲原點各奇級動差爲零.

2. 自行推得式 (14), (18), (20), (22) 及 (24) 所列之結果.

3. 復推式 (26), (27), 及 (28) 裏所列之結果.

4. 證明式 (30) 裏所列之結果.

5. 驗算表 14 各行裏所列之數值有無錯誤.

6. 驗算根據表 14 算得 β_1, β_2, κ_1 及 κ_2 之值有無錯誤.

7. 設以標準單位 t 表示任意分配之變值，求此分配之平均數及其標準差.

8. 如果各變量一一加以常數，試證各次主要動差之值不變.

9. 假設各變量一一乘以常數，問 $\bar{X}$, σ_X, β_1 及 β_2 所受之影響如何?

10. 證明 $\mu_2 \ngtr \nu_2$.

11. 按照教師指定之頻數分配計算中位數、平均差及四分位差.

12. 證明在常態曲線情形中，平均差 MD 與標準差 σ 之關係為

$$MD=\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\sigma \doteq .798\,\sigma.$$

13. 如果 x 以 $y=f(x)$ 為其分配函數，總頻數為 1. 試用微積分方法證明對於 ν 平均差

$$M=\int_{-\infty}^{\infty}|x-\nu|f(x)dx$$

以當 ν 等於中位數時為最小(參閱前章定理 15).

啓示：證明時可利用公式

$$\frac{dH}{d\theta}=\int_a^b\frac{\partial f}{\partial\theta}dx-f(a,\theta)\frac{da}{d\theta}+f(b,\theta)\frac{db}{d\theta},$$

倘是

$$H(\theta)=\int_a^b f(x,\theta)dx.$$

14. 根據表 50 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

15. 根據表 51 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

16. 根據表 52 裏所列之論據，求 β_1, β_2, κ_1 及 κ^2 之值.

17. 根據表 53 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

18. 根據表 54 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

19. 根據表 55 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

20. 根據表 56 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

21. 根據表 57 裏所列之論據，求 β_1, β_2, κ_1 及 κ_2 之值.

22. 設已給 $N=5,000$, $\bar{X}=75$ 及 $\sigma_X=10$; 並設變數 X 依照常態曲線分配.求

(a) $\text{cum} f=800$ 時之 X 值.

(b) N 個變值中小於 80 者有若干.

第四章　皮爾生各型曲線

第一節　頻數函數及機率函數

1. 基本概念　前兩章曾用統計常數——集中常數、離勢常數、偏度、峯度等——描寫頻數分配之種種特徵，那是頻數分配之初步分析。現在作進一步之分析，即將用方程式描寫各種不同頻數分配之性質。茲先述頻數函數 (frequency function) 及機率函數 (probability function) 基本概念於下。

當連續變數 X 發現於 a 與 b 間之頻數得用

(1) $$\int_a^b f(X)dX$$

度量時，我們說 X 有一分配函數 $f(X)$；彼處 $f(X)$ 係一非負單值之函數 (non-negative and single valued function). 如果 X 的分配函數 $f(X)$ 各頻數之和為 N，於是

(2) $$\int_{-\infty}^{\infty} f(X)dX=N$$

並且 $Y=f(X)$ 叫做理論頻數曲線 (theoretical frequency curve) 或簡稱之曰頻數曲線 (frequency curve). 如果變數 X 之實在發現限於一定數距之內，則 $f(X)$ 在此數距之外者完全為零.

如取曲線下面積為 1，即取

(3) $$\int_{-\infty}^{\infty} f(X)dx=1$$

則 $Y=f(X)$ 叫做 X 之密度曲線 (density curve), 或叫做機率曲線. 因此, $f(X)dX$ 爲 X 在 X 與 $X+dX$ 間之機率; 至若在式 (2) 條件下之 $\int_a^b f(X)dX$ 表示 X 在 a 與 b 間之頻數. 所以分配曲線叫做頻數曲線或機率曲線, 依式 (2) 及式 (3) 代表之條件以爲斷.

2. 動差　如果 X 依頻數曲線 $Y=f(X)$ 分配, 頻數之和爲 N; 於是對於 Y 軸之 r 級補助動差爲

$$\nu_{rX}=\frac{1}{N}\int_{-\infty}^{\infty}X^r f(X)dX. \tag{4}$$

在特殊情形中, 當 $r=1$ 時, 我們得到平均數

$$\bar{X}=\frac{1}{N}\int_{-\infty}^{\infty}Xf(X)dX.$$

若取平均數做原點, 則一級主要動差

$$\mu_{1X}=\int_{-\infty}^{\infty}(X-\bar{X})f(X)dX=0$$

並且 r 級主要動差

$$\mu_{rX}=\frac{1}{N}\int_{-\infty}^{\infty}(X-\bar{X})^r f(X)dX. \tag{5}$$

在特殊情形中, 當 $r=2$ 時, 我們有

$$\mu_{2X}=\frac{1}{N}\int_{-\infty}^{\infty}(X-\bar{X})^2 f(X)dX=\sigma_X^2$$

此爲變差(見前章第一節).

所有各 μ 得用各 ν 表之(見前章第五節), 即

$$\mu_{rX}=\nu_{rX}-r\nu_{(r-1)X}\nu_{1X}+\frac{r(r-1)}{2!}\nu_{(r-2)X}\nu_{1X}^2-\cdots. \tag{6}$$

在特殊情形中

$$(7)\qquad \begin{cases}\mu_{2X}=\nu_{2X}-\nu_{1X}^2,\\ \mu_{3X}=\nu_{3X}-3\nu_{2X}\nu_{1X}+2\nu_{1X}^3,\\ \mu_{4X}=\nu_{4X}-4\nu_{3X}\nu_{1X}+6\nu_{2X}\nu_{1X}^2-3\nu_{1X}^4.\end{cases}$$

如用標準單位表示,則 r 級主要動差

$$(8)\qquad \mu_{rt}=\int_{-\infty}^{\infty}(t-\bar{t})^r\phi(t)dt=\frac{\mu_{rX}}{\sigma_X^r},$$

彼處　$t=\dfrac{X-\bar{X}}{\sigma_X},\quad \phi(t)=\dfrac{\sigma_X}{N}f(X)=\dfrac{\sigma_X}{N}f(\bar{X}+\sigma_X t).$

於是由式 (8) 及前章第七節我們有

$$(9)\qquad \begin{cases}\mu_{0t}=1,\\ \mu_{1t}=0,\\ \mu_{2t}=1,\\ \mu_{3t}^2=\beta_1,\\ \mu_{4t}=\beta_2.\end{cases}$$

第二節　Γ 函數與 β 函數

1. Γ 函數　皮爾生各型曲線方程式內所包含之參數往往得用 Γ 及 β 函數表之,因此在本節內先討論 Γ 函數之性質,並且在以後各節討論 β 函數之性質及 Γ 與 β 函數之相互關係.

下列之定積分 (definite integral)

$$(10)\qquad \Gamma(n)=\int_0^{\infty}x^{n-1}e^{-x}dx,\quad n>0$$

叫做正數 n 之 Γ 函數 (gamma function). 使 $x^n=u$ 及 $e^{-x}dx=dv$, 用部分積分法求得式 (10) 右方之積分爲

(11) $$\Gamma(n+1)=\left|-x^ne^{-x}\right|_0^\infty-\int_0^\infty nx^{n-1}(-e^{-x})dx$$

$$=n\int_0^\infty x^{n-1}e^{-x}dx,\ 因\left|-x^ne^{-x}\right|_0^\infty=0$$

$$=n\Gamma(n).$$

重複應用公式 (11), 我們得

(12) $$\Gamma(n+1)=n(n-1)\cdots(n-k)\Gamma(n-k),$$

彼處 k 為小於 n 之正整數. 如果 n 亦為正整數, 則 $k=n-1$ 當然也是正整數. 於是我們有

(13) $$\Gamma(n+1)=n!;\ 因由式\ (10)\ 得知\ \Gamma(1)=1.$$

式 (12) 可被想做 $n!$ 之一般形式, 因此 Γ 函數有時被叫做階乘函數. 式 (10) 函數之圖形, 如圖 13 所示, 係根據下列各值繪畫者, 其中有些數值可直接由式 (10) 或 (11) 算得, 餘者將在以後推求.

$$\Gamma(0)=\infty$$

$$\Gamma(1)=1$$

$$\Gamma(\tfrac{1}{2})=(\pi)^{\frac{1}{2}}$$

$$\Gamma(2)=1$$

$$\Gamma(3)=2$$

$$\Gamma(4)=6$$

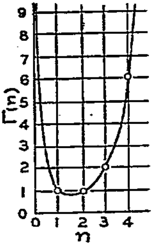


圖 13. Γ(n)之圖。

用替換變數法可得式 (10) 之另一形式. 例如, 使 $x=y^2$ 我們有

(14) $$\Gamma(n)=2\int_0^\infty y^{2n-1}e^{-y^2}dy,$$

並且由此式我們可以證明

$$(15) \qquad \int_0^\infty e^{-y^2}dy=\frac{\sqrt{\pi}}{2}.$$

因爲由式 (14)

$$(16) \qquad \Gamma(\tfrac{1}{2})=2\int_0^\infty e^{-y^2}dy,$$

又因式 (16) 代表之定積分與變數無關,故可將其寫做

$$\Gamma(\tfrac{1}{2})=2\int_0^\infty e^{-x^2}dx.$$

因之將最末二式相乘得

$$(17) \qquad [\Gamma(\tfrac{1}{2})]^2=4\int_0^\infty e^{-y^2}dy\int_0^\infty e^{-x^2}dx$$

$$=4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}dx\,dy,$$

彼處兩個積分所以能化成一個二重積分 (double integral) 者,蓋其一積分之被積分函數及其極限之變數與其他積分無關故也.

最好將式 (17) 化成極坐標然後再求其值,因此關於替換變數之一般情形儘先敍述一下.設 x 及 y 爲一點對於某一平面垂直軸之坐標, u 及 v 爲另一點對於另一平面任意軸之坐標.假設有一函數

$$z=f(x,y)$$

而以 x, y 爲其自變數,並且假設新變數 u 與 v 及 x 與 y 之關係爲

$$x=g(u,v) \text{ 及 } y=h(u,v).$$

再設 dA 代表函數 $f(x,y)$ 之面積元素 (area element) 於是根據高等微積分* 上證明之公式我們有

* 見 Mathematical Analysis, Goursat-Hedrick, Vol. 1.

$$dA = \left| J\left(\frac{x, y}{u, v}\right) \right| du\, dv$$

彼處 $\left| J\left(\frac{x, y}{u, v}\right) \right|$ 為下面行列式

$$\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

絕對值之簡寫符號。此行列式通常被叫做變換函數行列式 (functional determinant of transformation).

如果將式 (17) 變做極坐標 (polar coordinates), 即使

(18) $$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta; \end{cases}$$

則函數行列式變做

$$\begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r;$$

面積元素 $dx\, dy$ 變做 $r dr\, d\theta$; 積分之上下限在 r 為由 0 至 ∞, 在 θ 為由 0 至 $\pi/2$. 由式 (18), $x^2+y^2=r^2$. 如此式 (17) 變做

$$[\Gamma(\tfrac{1}{2})]^2 = 4\int_0^{\pi/2}\int_0^{\infty} e^{-r^2} r dr\, d\theta = 2\int_0^{\pi/2} d\theta = \pi.$$

所以

(19) $$\Gamma(\tfrac{1}{2}) = (\pi)^{\frac{1}{2}},$$

並且由式 (19) 及 (16) 可得式 (15) 所表示之結果.

比式 (15) 普通些之形式,可用下法得之:

使式 (14) 中之 $y = t/(2k)^{\frac{1}{2}}$, $k>0$, 我們有

(20) $$\int_0^{\infty} e^{-t^2/2k}dt=\tfrac{1}{2}(2\pi k)^{\frac{1}{2}}$$

並且

(21) $$\int_{-\infty}^{\infty} e^{-t^2/2k}dt=(2\pi k)^{\frac{1}{2}}.$$

式 (19) 之值可用另法求之如下. 式 (17) 右方代表在鐘形曲面

(22) $$z=e^{-(x^2+y^2)}$$

下之體積 v (見圖 14), 如此由式 (17) 我們有 $\Gamma(\frac{1}{2})=v^{\frac{1}{2}}$. 因為式 (22) 為一旋轉曲面 (surface of revolution), 我們可取半徑 r, 厚 dr 及高 z 之圓筒當做體積元素. 於是

$$dv=2\pi rzdr=2\pi re^{-r^2}dr,$$

$$v=2\pi\int_0^{\infty} e^{-r^2}rdr=\pi,$$

並且最後我們由末式及式(17)得到式 (19).

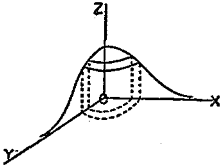


圖 14. $Z=e^{-(x^2+y^2)}$ 之軌跡。

2. β 函數　下面定積分

(23) $$\beta(m,n)=\int_0^1 x^{m-1}(1-x)^{n-1}dx$$

叫做正數 m 及 n 之 β 函數 (beta function). 使上式裏之 $x=\sin^2\theta$, 則得

(24) $$\beta(m,n)=2\int_0^{\pi/2}\sin^{2m-1}\theta\cos^{2n-1}\theta\, d\theta,$$

此為 β 函數之另一定義.

如果讓 $x=1-y$, 式 (23) 變做

$$\beta(m, n)=\int_0^1 (1-y)^{m-1} y^{n-1} dy$$
$$=\int_0^1 (1-x)^{m-1} x^{n-1} dx$$
$$=\beta(n, m).$$

由此可見 β 函數内之 m 及 n 可以彼此互相調換.

3. Γ 函數與 β 函數之關係 Γ 函數與 β 函數之關係可求得之如下.由式 (14) 我們可寫

$$\Gamma(n)\,\Gamma(m)=4\int_0^{\infty} x^{2n-1} e^{-x^2} dx \int_0^{\infty} y^{2m-1} e^{-y^2} dy$$
$$=4\int_0^{\infty}\int_0^{\infty} x^{2n-1} y^{2m-1} e^{-(x^2+y^2)} dx\, dy;$$

將其變為極坐標,則有

$$\Gamma(n)\,\Gamma(m)=4\int_0^{\pi/2}\int_0^{\infty} r^{2(m+n)-1} e^{-r^2} \sin^{2m-1}\theta \cos^{2n-1}\theta\, d\theta\, dr$$
$$=4\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2n-1}\theta\, d\theta \int_0^{\infty} r^{2(m+n)-1} e^{-r^2} dr$$
$$=2\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2n-1}\theta\, d\theta \times \int_0^{\infty} s^{m+n-1} e^{-s} ds,\quad s=r^2$$
$$=\beta(m, n)\,\Gamma(m+n),\ \text{由式 (24) 及 (10)}.$$

所以

(25) $$\beta(m, n)=\frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}$$

4. 不完全之 Γ 函數與 β 函數 下列之不定積分

(26) $$\Gamma_x(n+1)=\int_0^x x^n e^{-x} dx$$

叫做不完全 Γ 函數 (incomplete gamma function), 並且

(27) $$\beta_x(m, n)=\int_0^x x^{m-1}(1-x)^{n-1} dx$$

叫做不完全 β 函數 (incomplete beta function). 式 (26) 及 (27) 皆為有用之函數，皮爾生曾將此等函數之值，列表行世.

第三節　紅白球實驗

假設一袋內裝有 n 個球，其中有 np 個是紅色的，其餘 nq 個是白色的. 如此，在一單獨試驗中，取得一個紅球之機率為 p; 取不到紅球之機率為 $q=1-p$. 如果由袋內每次取出 s 個球，取出後再還回袋內. 現在我們要問確切取出 x 個紅球及 $(s-x)$ 個白球之機率為何? 由前章第十一節定理 7 得知所問之機率為 $B(x)=\frac{s!}{x!(s-x)!}p^xq^{s-x}$.

如果以縱坐標 y_x 代表 $B(x)$ 項，如此，可得 $(s+1)$ 個 (x, y) 點，並且經過這些點可以想像有一條能被算式代表之曲線. 於是

$$y_x=\frac{s!}{x!(s-x)!}p^xq^{s-x},$$

$$y_{x+1}=\frac{s!}{(x+1)!(s-x-1)!}p^{x+1}q^{s-x-1},$$

並且

$$\frac{y_{x+1}}{y_x}=\frac{p(s-x)}{q(x+1)}.$$

因之由末式得

$$(28)\qquad \frac{y_{x+1}-y_x}{y_{x+1}+y_x}=\frac{sp-q-x}{sp+q+(q-p)x}.$$

任意二縱坐標 y_x 及 y_{x+1} 之平均數可視做與縱坐標 $y_{x+\frac{1}{2}}$ 相等. 聯結曲線上兩點 (x, y_x) 及 $(x+1, y_{x+1})$ 成直線，其斜度與經過曲線上 $(x+\frac{1}{2}, y_{x+\frac{1}{2}})$ 點切線之斜度相近. 因此可將式 (28) 寫做

$$\frac{1}{y_{x'}}\frac{dy_{x'}}{dx'}=\frac{2(sp-q-x)}{sp+q+(q-p)x},$$

彼處 $x'=x+\frac{1}{2}$. 於是

$$\frac{1}{y_{x'}}\frac{dy_{x'}}{dx'}=\frac{2\{sp-q-(x'-\frac{1}{2})\}}{sp+q+(q-p)(x'-\frac{1}{2})};$$

去掉 x 右上角之撇 ($'$), 我們有

(29) $$\frac{1}{y}\frac{dy}{dx}=\frac{2\{sp-q-(x-\frac{1}{2})\}}{sp+q+(q-p)(x-\frac{1}{2})}.$$

如果 $p=q=\frac{1}{2}$, 則式 (29) 變做

(30) $$\frac{1}{y}\frac{dy}{dx}=\frac{2\left(\frac{s}{2}-x\right)}{\frac{s+1}{2}}=\frac{-\left(x-\frac{s}{2}\right)}{\frac{s+1}{4}},$$

其形式爲

(31) $$\frac{dy}{dx}=\frac{(x-a)y}{c_0},$$

積分後之形式爲

(32) $$y=y_0e^{\frac{(x-a)^2}{2c_0}};$$

其爲一常態曲線,彼處 a, c_0 及 y_0 爲待定之常數(參閱本章第十二節).

如果 $p\neq q$, 由式 (29), 我們有

$$-\frac{1}{y}\frac{dy}{dx}=\frac{\frac{q-p}{2}+(x-sp)}{spq+\frac{1}{4}+\frac{(x-sp)(q-p)}{2}}.$$

若使 $\sqrt{\beta_1}=\frac{q-p}{\sqrt{spq}}$ 及 $t=\frac{x-sp}{\sqrt{spq}}$, [參閱前章式 (32)]

則 $dx=(spq)^{\frac{1}{2}}dt$, 因之

(33) $$-\frac{1}{y}\frac{dy}{dt}=\frac{\sqrt{\beta_1}/2+t}{\left(1+\sqrt{\beta_1}t/2+\frac{1}{4spq}\right)\sqrt{spq}}.$$

如果 spq 很大，則以 $1/4spq$ 甚小可將其忽略之。如此式 (33) 變做

$$\frac{1}{y}\frac{dy}{dt}=-\frac{\sqrt{\beta_1}/2+t}{(1+\sqrt{\beta_1}t/2)\sqrt{spq}}, \tag{34}$$

其形式為

$$\frac{dy}{dx}=\frac{(x-a)y}{c_0+c_1x} \tag{35}$$

並且積分後之形式為皮爾生頻數曲線型 III. 關於此型曲線，將在第十四節詳加討論。很顯然的，如果 $\sqrt{\beta_1}\to 0$，則皮爾生頻數曲線型 III 變成了常態曲線。

當 $p=q$ 時，微分方程式之形式如式 (31) 所示；當 $p\neq q$ 時，其形式如式 (35) 所示，已如上述。現在再進一步研究，在何種情形之下我們可以得到下列之形式：

$$\frac{dy}{dx}=\frac{(x-a)y}{c_0+c_1x+c_2x^2}, \tag{36}$$

其右方分母為一二次式。為達到此目的，我們將紅白球實驗問題推廣之。將上面紅白球實驗中"取出再還回袋內"之限制除掉，於是各次試驗機率依其前次試驗之結果而定。如此，取出 s 個球來，攫得 0、1、2、…、s 個紅球之機率為一超比級數 (hypergeometric series).

$$\frac{1}{C(n,s)}\{C(np,0)C(nq,s)+C(np,1)C(nq,s-1)+\cdots \tag{37}$$
$$+C(np,x)C(nq,s-x)+\cdots+C(np,s)C(nq,0)\}.$$

若以 $H(x)$ 代表公項，則

$$H(x)=\frac{C(np,x)C(nq,s-x)}{C(n,s)}$$
$$=\frac{(np)!(nq)!s!(n-s)!}{(np-x)!(nq-s+x)!n!x!(s-x)!}.$$

用此級數各項代表頻數多邊形之縱坐標．我們可證明多邊形任一邊中點之斜度，以中點縱坐標除之，為一分數，該分數之分子為 x 的一次式乘以 y，分母為 x 的二次式，其形式如式 (36) 右方所示．

因為由超比級數生成式 (36)，並且伯努里級數生成式 (35)；於是我們很自然的希望伯努里級數為超比級數之特殊情形．為說明此點起見將 $H(x)$ 寫成下形：

$$H(x)=\frac{s!}{x!(s-x)!}\cdot\frac{p\{p-1/n\}\cdots\{p-(x-1)/n\}q\{q-1/n\}\cdots\{q-(s-x-1)/n\}}{\{1-1/n\}\cdots\{1-(x-1)/n\}\{1-x/n\}\cdots\{1-(s-1)/n\}}$$

很顯然的

$$\lim_{n\to\infty} H(x)=C(s,\ x)\,p^x q^{s-x}=B(x).$$

當 $n=\infty$ 時，即袋內之球增至無限多，如此雖不將取出之球還回，各次試驗之 p 值亦可視作不變；反之，袋內之球數有限，將其取出後還回，各次試驗之 p 值不變．故知若一袋內所裝紅白兩色之球數無限，第一次將球取出 x 個後不還回袋內就取第二次；另一袋內之球數有限，紅白兩色球數之比與前袋同，第一次將球亦取出 x 個，取出後還回再取第二次：則由此二袋內取得紅球(或白球)之機率相同．

第四節　求積公式

假設我們有縱坐標 $y_{-\frac{1}{2}}$、$y_{\frac{1}{2}}$、$y_{-1\frac{1}{2}}$、$y_{1\frac{1}{2}}$、…或 y_0、y_{-1}、y_1、y_{-2}、y_2、…或 y_0、y_1、$y_2\cdots y_{n-1}$．現在我們用求積公式計算與縱坐標 y_0、y_1、$y_2\cdots y_{n-1}$ 對應之實在面積．求積公式甚多，其形式因 y_x 之形式而異．吾人所欲求者祇以適應特殊目的為依歸．現在舉例明之如下：

例 1. 設 $y_x=a+bx+cx^2+dx^3+ex^4$. 試以 y_{-2}, y_{-1}, y_0, y_1 及 y_2 表示面積 $\int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx$.

因 $$y_x=a+bx+cx^2+dx^3+ex^4$$

於是 $$\int_{-\frac{1}{2}}^{\frac{1}{2}} dy\,dx=\int_{-\frac{1}{2}}^{\frac{1}{2}}(a+bx+cx^2+dx^3+ex^4)dx$$

$$=a+\frac{c}{12}+\frac{e}{80}$$

但 $$y_0=a$$

$$y_{-1}+y_1=2(a+c+e)$$

$$y_{-2}+y_2=2(a+4c+16e)$$

並且可假設

$$a+\frac{c}{12}+\frac{e}{80}=hy_0+k(y_{-1}+y_1)+l(y_{-2}+y_2)$$

故 $$a+\frac{c}{12}+\frac{e}{80}=ha+2k(a+c+e)+2l(a+4c+16e).$$

分別使此恆等式中之 $a=1$, $c=0$ 及 $e=0$; $a=0$, $c=1$ 及 $e=0$; $a=0$, $c=0$ 及 $e=1$: 我們有

$$h+2k+2l=1,$$

$$2k+8l=\frac{1}{12},$$

$$2k+32l=\frac{1}{80}.$$

以 h, k 及 l 當做未知數解之,得

$$h=\frac{5178}{5760},\quad k=\frac{308}{5760},\quad l=\frac{-17}{5760}.$$

所以

$$(38) \qquad \int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx \frac{1}{5760}\{5178y_0 + 308(y_{-1}+y_1) - 17(y_{-2}+y_2)\}.$$

例 2. 設 $y_x = a+bx+cx^2+dx^3$. 試以 y_{-1}、y_0 及 y_1 表示面積 $\int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx$.

因為 $y_x = a+bx+cx^2+dx^3$, 故

$$(39) \qquad \int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx = \frac{1}{24}\{y_{-1}+22y_0+y_1\}.$$

例 3. 設 $y_x = a+bx+cx^2+dx^3+ex^4$. 試以 $y_{-2\frac{1}{2}}$、$y_{-1\frac{1}{2}}$、$y_{1\frac{1}{2}}$ 及 $y_{2\frac{1}{2}}$ 表示面積 $\int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx$.

因為 $y_x = a+bx+cx^2+dx^3+ex^4$, 故

$$(40) \quad \int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx = \frac{1}{1440}\{802(y_{-\frac{1}{2}}+y_{\frac{1}{2}}) - 93(y_{-1\frac{1}{2}}+y_{1\frac{1}{2}}) + 11(y_{-2\frac{1}{2}}+y_{2\frac{1}{2}})\}.$$

例 4. 設 $y_x = a+bx+cx^2+dx^3$. 試以 y_0, y_1, y_2 及 y_3 表示面積 $\int_{-\frac{1}{2}}^{1\frac{1}{2}} y_x dx$.

因為 $y_x = a+bx+cx^2+dx^3$, 故

$$(41) \qquad \int_{-\frac{1}{2}}^{1\frac{1}{2}} y_x dx = \frac{1}{24}\{27y_0 + 17y_1 + 5y_2 - y_3\}.$$

第五節　釐正縱坐標之動差

假設各縱坐標 y_0、$y_1 \cdots y_{n-1}$ 是孤立的，想若計算牠們的動差，我們須用各孤立的縱坐標表示面積 $\int_{-\frac{1}{2}}^{n-\frac{1}{2}} y_x dx$. 因為

$$\int_{-\frac{1}{2}}^{n-\frac{1}{2}} y_x dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx + \int_{\frac{1}{2}}^{1} y_x dx + \cdots + \int_{n-1\frac{1}{2}}^{n-\frac{1}{2}} y_x dx,$$

於是應用式 (38) 於右邊各積分上，但居首及居末二積分除外，得

$$\int_{-\frac{1}{2}}^{n-\frac{1}{2}} y_x dx=\int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx+\int_{\frac{1}{2}}^{1\frac{1}{2}} y_x dx+\int_{n-2\frac{1}{2}}^{n-1\frac{1}{2}} y_x dx+\int_{n-1\frac{1}{2}}^{n-\frac{1}{2}} y_x dx$$

$$+\frac{1}{5760}\Big[\{5178y_2+308(y_1+y_3)-17(y_0+y_4)\}$$

$$+\{5178y_3+308(y_2+y_4)-17(y_1+y_5)\}$$

$$+\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+\{5178y_{n-5}+308(y_{n-6}+y_{n-4})-17(y_{n-7}+y_{n-3})\}\Big].$$

居首及居末二積分可應用式(41)求得之，即

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx+\int_{\frac{1}{2}}^{1\frac{1}{2}} y_x dx=\int_{-\frac{1}{2}}^{1\frac{1}{2}} y_x dx$$

$$=\frac{1}{24}\{27y_0+17y_1+5y_2-y_3\}$$

$$=\frac{1}{5760}\{6480y_0+4080y_1+1200y_2-240y_3\}$$

及
$$\int_{n-2\frac{1}{2}}^{n-1\frac{1}{2}} y_x dx+\int_{n-1\frac{1}{2}}^{n-\frac{1}{2}} y_x dx=\int_{n-2\frac{1}{2}}^{n-\frac{1}{2}} y_x dx$$

$$=\frac{1}{24}\{27y_{n-1}+17y_{n-2}+5y_{n-3}-y_{n-4}\}$$

$$=\frac{1}{5760}\{6480y_{n-1}+4080y_{n-2}+1200y_{n-3}-240y_{n-4}\}.$$

故欲求之面積

$$\int_{-\frac{1}{2}}^{n-\frac{1}{2}} y_x dx=\frac{1}{5760}\Big[\{6480y_0+4080y_1+1200y_2-240y_3\}$$

$$+\{5178y_2+308(y_1+y_3)-17(y_0+y_4)\}$$

$$+\{5178y_3+308(y_2+y_4)-17(y_1+y_5)\}$$

$$+\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+\{5178y_{n-3}+308(y_{n-4}+y_{n-2})-17(y_{n-5}+y_{n-1})\}$$

$$+\{6480y_{n-1}+4080y_{n-2}+1200y_{n-3}-240y_{n-4}\}\Big]$$

$$=\frac{1}{5760}\left[6463y_0+4371y_1+6669y_2+5537y_3\right.$$
$$+5760(y_4+y_5+\cdots+y_{n-6}+y_{n-5})$$
$$\left.+5537y_{n-4}+6669y_{n-3}+4371y_{n-2}+6463y_{n-1}\right].$$

由末式得知,若由縱坐標求動差我們宜將

第一及最末縱坐標以 $\frac{6463}{5760}(=1.1220486)$ 乘之,

第二及末第二縱坐標以 $\frac{4731}{5760}(=0.7588542)$ 乘之,

第三及末第三縱坐標以 $\frac{6669}{5760}(=1.1578125)$ 乘之,

第四及末第四縱坐標以 $\frac{5537}{5760}(=0.961284)$ 乘之;

其他各縱坐標仍舊不動.乘以相當常數後之縱坐標叫做矯正縱坐標.如果所給縱坐標之數目少於 8, 上式則不適用,求矯正縱坐標時須用其他公式.各級動差宜用矯正縱坐標求之,不宜用原始縱坐標.

下表第一及第二行列入原始及矯正縱坐標,其他各行曉示求首四級動差之步驟.於是以 207.41 當做頻數總和,並且以頻數總和爲單位之首四級動差係以 207.41 而非以 207.18 除 −391.15、1530.63、−4672.21 及 19168.59 得之,彼處所以用 207.41 而不用 207.18 除者,因前者爲頻數之和,而後者非也,僅係未矯正之一組等距離數值之和耳.

當已給縱坐標兩端之數值很小時,並且其與 x 軸有極相接近之趨勢,於是在首項之前及末項之後雖假設有數值存在,但極微渺,無足介意.如此與全組坐標成對應之積分,可以延出 $-\frac{1}{2}$ 及 $n-\frac{1}{2}$ 限度之外.並且延出部份所引入之面積得以忽略.如果面積向外如此伸張,則上面方

表 17. 矯正縱坐標首四級動差之計算(錄自參考書 II)

原始縱坐標 y_x	矯正縱坐標 y_x'	x	$y_x'x$	$y_x'x^2$	$y_x'x^3$	$y_x'x^4$
51.81	58.13	−4	−232.52	930.08	−3720.32	14881.28
43.74	33.19	−3	−99.57	298.71	−896.13	2688.39
35.42	41.01	−2	−82.02	164.04	−328.08	656.16
27.80	26.72	−1	−26.72	26.72	−26.72	26.72
20.42	20.42	0	0.00	0.00	0.00	0.00
13.79	13.26	1	13.26	13.26	13.26	13.26
8.22	9.52	2	19.04	38.08	76.16	152.32
4.29	3.26	3	9.78	29.34	88.02	264.06
1.69	1.90	4	7.60	30.40	121.60	486.40
207.18	207.41	······	−391.15	1530.63	−4672.21	19168.59

程式裏縱坐標由 y_0 至 y_{n-1} 之係數將為 1. 所有應當加權之縱坐標將為零矣. 如是頻數曲線一端若與 x 軸成高切, 計算各級動差時祗須矯正他端; 其兩端若均與 x 軸成高切, 無須矯正.

第六節 薛伯校正數之來源

前在第三章第九節曾討論過薛伯校正數存在之理由，今將該值之來源加以追溯如下:

假設各矩形面積聚在其底邊之中點，於是面積 $\int_{-\frac{1}{2}}^{\frac{1}{2}} y_x dx$、$\int_{\frac{1}{2}}^{1\frac{1}{2}} y_x dx$ … 至原點之距離應各與 y_0、y_1 … 至 y 軸之距離相同，故其 t 級觀察動差為

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} X^t y_x dx + \int_{\frac{1}{2}}^{1\frac{1}{2}} (X+1)^t y_x dx + \cdots + \int_{n-1\frac{1}{2}}^{n-\frac{1}{2}} (X+n-1)^t y_x dx,$$

彼處 X 為 y_0 與縱軸之距離.

將公式 (38) 應用於每個積分上面，並將同類項歸併之，則得公項之係數為

$$\frac{y_x}{5760}[5178h^t+308\{(h-1)^t+(h+1)^t\}-17\{(h-2)^t+(h+2)^t\}],$$

彼處 h 代表 $X+x$, $x=0$、1、2…；或將末式寫做

$$\frac{y_x}{5760}\{5760h^t+240t(t-1)h^{t-2}+3t(t-1)(t-2)(t-3)h^{t-4}+\cdots\}.$$

如果 $t=1$, 上式變做 hy_x.

如果 $t=2$, 上式變做 $\left(h^2+\frac{1}{12}\right)y_x$

如果 $t=3$, 上式變做 $\left(h^3+\frac{1}{4}h\right)y_x$

如果 $t=4$, 上式變做 $\left(h^4+\frac{1}{2}h^2+\frac{1}{80}\right)y_x$

我們曾經討論過，如果頻數為高切，$\int_{-\frac{1}{2}}^{n-\frac{1}{2}}(X+x)^t y_x dx$ 之值可由不矯正之縱坐標求之；就是說 $\sum_{x=0}^{n-1}(X+x)^2 y_x$ 為二級矯正動差，$\sum_{x=0}^{n-1}(X+x)^3 y_x$ 為三級矯正動差，如此等等；故若以 μ' 代表此矯正之主要動差，則 μ' 與 μ 之關係為

$$(42)\qquad \begin{cases}\mu_1=\mu_1'=0,\\ \mu_2=\mu_2'+\frac{1}{12},\\ \mu_3=\mu_3'+\frac{1}{4}\mu_1',\\ \mu_4=\mu_4'+\frac{1}{2}\mu_2'+\frac{1}{80};\end{cases}\qquad \text{即是}\begin{cases}\mu_1'=\mu_1=0,\\ \mu_2'=\mu_2-\frac{1}{12},\\ \mu_3'=\mu_3,\\ \mu_4'=\mu_4-\frac{1}{2}\mu_2'-\frac{1}{80}.\end{cases}$$

一級動差無須矯正，因爲 $\mu_1=\mu_1'$. 如果 $t=1$，則 $\mu_1'=\mu_1=0$. 故三級動差雖然應當以 $\frac{1}{4}\mu_1'$ 矯正之，但因 μ_1' 之值爲 0. 故矯正項之值爲 0. 二級動差應以 $-\frac{1}{12}$ 矯正之，四級動差應以 $-\frac{1}{2}\mu_2-\frac{1}{80}$ 矯正之.

第七節　常數與觀察動差之關係

皮爾生微分方程式內所含常數與觀察動差之關係可推得之如下：

將公式 (36) 兩邊各以 $(c_0+c_1x+c_2x^2)x^n dx$ 乘之，並將結果積分之；我們有

$$\int_u^v (c_0+c_1x+c_2x^2)x^n dy=\int_u^v (x-a)x^n y\,dx$$

彼處 u 爲曲線起點之橫坐標，v 爲其終點之橫坐標.

施行部分積分法於上式左方，則因 $\left|(c_0+c_1x+c_2x^2)x^n y\right|_u^v=0$,

$$\int_u^v (x-a)x^n y\,dx=-\int_u^v \{c_0 nx^{n-1}+c_1(n+1)x^n+c_2(n+2)x^{n+1}\}y\,dx$$

即是

(43)
$$c_0 n\int_u^v x^{n-1}y\,dx+\{c_1(n+1)-a\}\int_u^v x^n y\,dx+\{c_2(n+2)+1\}\int_u^v x^{n+1}y\,dx=0$$

假設 x 爲以組距爲單位之離均差，於是式 (43) 可寫做

(44)
$$c_0 n\,\mu_{n-1}+\{c_1(n+1)-a\}\mu_n+\{c_2(n+2)+1\}\mu_{n+1}=0,$$

彼處
$$\mu_s=\int_u^v x^s y\,dx\div\int_u^v y\,dx.$$

使 $n=0$、1、2 及 3，并且因爲 $\mu_0=1$ 及 $\mu_1=0$；我們有

$$c_1-a=0,$$
$$c_0+(3c_2+1)\mu_2=0,$$
$$(3c_1-a)\mu_2+(4c_2+1)\mu_3=0,$$
$$3c_0\mu_2+(4c_1-a)\mu_3+(5c_2+1)\mu_4=0.$$

以 a、c_0、c_1 及 c_2 當做未知數解之,並以 $\beta_1=\mu_3^2/\mu_2^3$, $\beta_2=\mu_4/\mu_2^2$, $\kappa_1=2\beta_2-3\beta_1-6$, $H=(8\kappa_1+12\beta_1+48)\kappa_1$, $K=10\beta_2-12\beta_1-18$, $L=(\beta_2+3)\mu_3/\mu_2$ 代入結果式中;則得欲求之關係

$$(45)\qquad \begin{cases} a=-\dfrac{L}{K} \\ c_0=-\dfrac{H\mu_2}{4\kappa_1 K} \\ c_1=-\dfrac{L}{K} \\ c_2=-\dfrac{\kappa_1}{K} \end{cases}$$

此關係爲本章後來推演計算時所用公式之張本,學者宜注意及之.至於 L、K、H、κ_1 及 μ_2 之值,宜用前章式 (28) 所給之實際運算公式求之.

第八節 皮爾生各型曲線概要

皮爾生頻數曲線共十三種:型 I 至 XII 及常態曲線.型 I、IV 及 VI 通常稱之爲主型;型 II、III、V、VII 至 XIII 及常態曲線通常稱之爲變型.就中以首六種及常態曲線應用較廣,故分別在以後各節討論之.其他六種不一一備述,學者欲知其詳可閱覽參考書 11 及 21.

茲爲便於學者披閱起見,特將皮爾生各型曲線之方程式、原點、特徵及準則列入表 18.表 18 裏所列之 $\beta_1=\mu_3^2/\mu_2^3$, $\beta_2=\mu_4/\mu_2^2$,

表 18. 皮爾生各型頻數曲線概要

型	方程式	原點	特徵	準則
主型：				
I	$y=y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2}$	在衆數處	不對稱，平常爲鈴形，全距有限	$\kappa_2<0$
IV	$y=y_0\left(1+\frac{x^2}{p_1^2}\right)^{-p_0}e^{-p_2\tan^{-1}\frac{x}{p_1}}$	在平均數前 $\frac{hc_1}{2c_2}$ 處	偏側鈴形，全距無限	$0<\kappa_2<1$
VI	$y=y_0(x-q_0)^{q_1}x^{-q_2}$	在平均數後 hr_2 處	偏側鈴形，亦可爲 J 形，一端無限	$\kappa_2>1$
變型：				
常態分配	$y=y_0e^{-x^2/2\sigma^2}$	在衆數處	對稱，鈴形，兩端無限	$\beta_1=0,\ \beta_2=3$
II	$y=y_0\left(1-\frac{x^2}{a^2}\right)^{m}$	在衆數處	對稱，通常爲鈴形，全距有限	$\beta_1=0,\ \beta_2<3$
III	$y=y_0\left(1+\frac{x}{s_0}\right)^{s_1}e^{-s_2x}$	在衆數處	通常爲鈴形，但可爲 J 形，一端無限	$\kappa_1=0$
V	$y=y_0x^{-t_0}e^{-t_1/x}$	在平均數前 $\frac{hc_1}{2c_2}$ 處	鈴形，一端無限	$\kappa_2=1$
VII	$y=y_0\left(1+\frac{x^2}{a^2}\right)^{-m}$	在衆數處	對稱，鈴形，全距無限	$\beta_1=0,\ \beta_2>3$
VIII	$y=y_0\left(1+\frac{x}{a}\right)^{-m}$	在曲線末端	x 由 $-a$ 至 0	$\kappa_2<0,\ \lambda=0,\ K<0$
IX	$y=y_0\left(1+\frac{x}{a}\right)^{m}$	在曲線末端	x 由 $-a$ 至 0	$\kappa_2<0,\ \lambda=0,\ K>0,\ \kappa_1<0$
X	$y=y_0e^{-x/\sigma}$	在曲線始端	x 由 0 至 ∞.	$\beta_1=4,\ \beta_2=9$
XI	$y=y_0x^{-m}$	在曲線始端前 b 處	J 形.	$\kappa_2>0,\ \lambda=0,\ \kappa_1>0$
XII	$y=y_0\left(\frac{\sigma\{\sqrt{3+\beta_1}+\sqrt{\beta_1}\}+x}{\sigma\{\sqrt{3+\beta_1}-\sqrt{\beta_1}\}-x}\right)^{\sqrt{\beta_1/(3+\beta_1)}}$	在平均數處	扭 J 形	$K=0$

$$\kappa_1=2\beta_2-3\beta_1-6,\quad \kappa_2=(\beta_2+3)^2\div H\times\beta_1,$$

$$\lambda=\frac{K^2-4\kappa_1\kappa_2(K-\kappa_1)}{4\kappa_1{}^2(1-\kappa_2)},$$

$$H=(8\kappa_1+12\beta_1+48)\kappa_1$$

並且

$$K=10\beta_2-12\beta_1-18.$$

當 $\frac{dy}{dx}=0$ 時,由式 (36) 得知 $x=a$. 但式中之 x 爲以組距爲單位之離均差,即 $x=\frac{X-\bar{X}}{h}$; 故 $X=\bar{X}+hx$. 由前第二章第六節 4 關於衆數之定義,得知皮爾生各型曲線之衆數 $\check{X}=\bar{X}+ha$.

第九節　皮爾生曲線型 I

1. 準則及方程式　屬於皮爾生曲線型 I 之頻數分配,以 $\kappa_2=\frac{c_1{}^2}{4c_0c_2}<0$ 爲判定之準則.如果 $\kappa_2<0$, 則 c_0 與 c_2 之符號相反, $c_0+c_1x+c_2x^2=0$ 兩根之大小不等並且符號相反,譬如各等於 $-r_1$ 及 r_2, 彼處 r_1 及 r_2 均爲正數.因此用部分分數法得

$$\frac{x-a}{c_0+c_1x+c_2x^2}=\frac{1}{c_2(r_1+r_2)}\left\{\frac{a+r_1}{x+r_1}-\frac{a-r_2}{x-r_2}\right\}.$$

將其代入式 (36) 內並將結果積分之,我們有

$$\log y=\frac{a+r_1}{c_2(r_1+r_2)}\log(x+r_1)-\frac{a-r_2}{c_2(r_1+r_2)}\log(x-r_2)+\text{積分常數}.$$

將式中之 x 以 $x+a$ 代之,即是將原點向右移動 a 單位,亦卽以

$$X_0=X+ah$$

爲現在曲線原點之横坐標,彼處 h 代表組距;我們有

$$\log y=m_1\log(a_1+x)+m_2\log(a_2-x)+\text{積分常數},$$

後處 $a_1=a+r_1$, $a_2=-(a-r_2)$, $m_1=\frac{a_1}{c_2(r_1+r_2)}$ 並且 $m_2=\frac{a_2}{c_2(r_1+r_2)}$.
於是得到所求之方程式

$$y=y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2},$$

式處 y_0 爲待定之常數.此爲皮爾生頻數曲線型 I 之方程式.

2. 原點橫坐標 X_0 原點橫坐標 X_0 可用下式求之:

47) $$X_0=\bar{X}+ha$$

式處 $\bar{X}$ 爲平均數, h 爲組距, $a=-\frac{L}{K}$

3. 常數 y_0 由上式可見:如果 $y=0$, $x=-a_1$ 或 a_2; 就是說 $(-a_1, 0)$ 爲曲線之始點, $(a_2, 0)$ 爲其終點.故計算 y_0 之公式可推求之如下:

我們知道 $$\int_{-a_1}^{a_2}y_0\left(1+\frac{x}{a_1}\right)^{m_1}\left(1-\frac{x}{a_2}\right)^{m_2}dx=N,$$

式處 N 表示頻數之總和;又因式中之 x 得用下式

$$z=(a_1+x)/(a_1+a_2)$$

替換之,於是 $$x=(a_1+a_2)z-a_1$$

$$dx=(a_1+a_2)dz$$

並且 $$\frac{(a_1+a_2)^{m_1+m_2+1}y_0}{a_1^{m_1}\,a_2^{m_2}}\int_0^1 z^{m_1}(1-z)^{m_2}dz=N.$$

因之 $$\frac{(a_1+a_2)^{m_1+m_2+1}y_0}{a_1^{m_1}\,a_2^{m_2}}\cdot\frac{\Gamma(m_1+1)\Gamma(m_2+1)}{\Gamma(m_1+m_2+2)}=N,$$

即 $$y_0=\frac{Na_1^{m_1}\,a_2^{m_2}}{(a_1+a_2)^{m_1+m_2+1}}\div\beta(m_1+1,\ m_2+1)$$

$$=\frac{Na_1^{m_1}\,a_2^{m_2}}{(a_1+a_2)^{m_1+m_2+1}}\div\frac{\Gamma(m_1+1)\Gamma(m_2+1)}{\Gamma(m_1+m_2+2)}$$

48) $$=\frac{Na_1^{m_1}\,a_2^{m_2}\Gamma(m_1+m_2+2)}{(a_1+a_2)^{m_1+m_2+1}\Gamma(m_1+1)\Gamma(m_2+1)}.$$

4. 公式 除 X_0 及 y_0 之值分別用式 (47) 及 (48) 計算外，式 (46) 中之 a_1、a_2、m_1 及 m_2 之值可依照下列各步驟計算之：

$$
(49)\quad\begin{cases}
K=10\beta_2-12\beta_1-18 & Q=(\frac{1}{2}\div\kappa_1)\times M\\
L=(\beta_2+3)\div\mu_2\times\mu_3 & a_1=-Q+(a-P)\\
a=-L\div K & a_2=-Q-(a-P)\\
M=\left|L^2-\mu_2 H\right|^{\frac{1}{2}} & m_1=K\div M\times a_1\\
P=(\frac{1}{2}\div\kappa_1)\times(-L) & m_2=K\div M\times a_2
\end{cases}
$$

5. 舉例 試以皮爾生頻數曲線之適宜型配合下列之論據.

a. 計算 β_1、β_2、κ_1 及 κ_2 [公式見前章式 (28)]

表 19. 燕麥總產量之分配 (論據由參考書 27)

總產量 (磅)	頻數 f	助變數 x	xf	x^2f	x^3f	x^4f
0.0—	3	−4	−12	48	−192	768
1.0—	50	−3	−150	450	−1,350	4,050
2.0—	106	−2	−212	424	−848	1,696
3.0—	109	−1	−109	109	−109	109
4.0—	80	0	0	0	0	0
5.0—	42	1	42	42	42	42
6.0—	7	2	14	28	56	112
7.0—	2	3	6	18	54	162
8.0—	1	4	4	16	64	256
總計	400	……	−417	1,135	−2,283	7,195

$N=400$		$\mu_2=$	1.750694
$\nu_4=$	17.987500	$F=$	3.064929
$\nu_3=$	-5.707500	$G=$	5.365753
$\nu_2=$	2.837500	$\mu_3=$	0.900791
$\nu_1=$	-1.042500	$\beta_1=$	0.151223
$A=$	3.127500	$\beta_2=$	2.984296
$B=$	4.170000	$\kappa_1=$	-0.485077
$C=$	2.173612	$H=$	-22.281556
$D=$	6.520836	$\kappa_2=$	-0.243052
$E=$	3.398986	(合乎型 I 之準則)	

b. 計算 a_1、a_2、m_1、m_2、X_0 及 y_0 [見式 (48)、(49) 及 (50)]

$K=$	10.028284	$Q=$	-7.177640
$L=$	3.079122	$a_1=$	3.696748
$a=$	-0.307044	$a_2=$	10.658532
$M=$	6.963417	$m_1=$	5.323827
$P=$	3.173848	$m_2=$	15.349757

$$\log N= \quad 2.6020600$$

$$m_1\log a_1= \quad 3.0229738$$

$$m_2\log a_2= \quad 15.7749065$$

$$\log\Gamma(m_1+m_2+2)= \quad 20.6104144$$

$$-(m_1+m_2+1)\log(a_1+a_2)= -25.0765903$$

$$-\log\Gamma(m_1+1)= -\ 2.3231733$$

$$\text{-} \log \Gamma(m_2+1) = -12.5345529$$

$$\log y_0 = \quad 2.0760371$$

$$y_0 = \quad 119.134384$$

$$X_0 = \quad 3.150456$$

c. 寫出方程式　根據算得之結果，得將方程式寫出之如下：

$$(50) \qquad y = 119.1344\left(1+\frac{x}{3.6967}\right)^{5.2238}\left(1-\frac{x}{10.6585}\right)^{15.5498}.$$

d. 計算縱坐標及面積　曲線(50)之縱坐標及其下面積可依據表20內所列之步驟計算之．該表內各行之運算手績可以自明，無須加以解說．

表 20.　曲線(50)縱坐標及其下面積之計算

縱坐標(公斤) X (1)	$\frac{X-X_0}{h}=x$ (2)	$\log\left(1+\frac{x}{a_1}\right)$ (3)	$\log\left(1-\frac{x}{a_2}\right)$ (4)	$m_1\times(3)$ (5)	$m_2\times(4)$ (6)	y (7)	面積* (8)
0.5	-2.650456	$\bar{1}$.451816[illegible]	.096448[illegible]	-2.[illegible]417	1.480458	4.4	6.0
1.5	-1.650456	$\bar{1}$.7431426	.062525[illegible]	-1.367457	.959751	46.6	47.3
2.5	-.650456	$\bar{1}$.9159503	.025726[illegible]	-.447464	.394897	105.[illegible]	103.6
3.5	.349544	.0392379	$\bar{1}$.985518[illegible]	.208895	-.222287	11[illegible].[illegible]	113.6
4.5	1.349544	.1351543	$\bar{1}$.941204[illegible]	.719534	-.902496	78.[illegible]	78.0
5.5	2.349544	.2136717	$\bar{1}$.891850[illegible]	1.137545	-1.660079	35.[illegible]	36.5
6.5	3.349544	.2801438	$\bar{1}$.836158[illegible]	1.491430	-2.514930	11.[illegible]	12.0
7.5	4.349544	.3377794	$\bar{1}$.772260[illegible]	1.798270	-3.495750	2.[illegible]	2.7
8.5	5.349544	.3886545	$\bar{1}$.6973122	2.06911[illegible]	-4.646197		.4
總計	......	......	......	...	......	......	40[illegible]

* 此行所列之面積係用式(39)計算者，所得結果已甚準確；倘需要更準確之結果時，可用式(38)計算之．

e. 比較配合情形　由於比較表 20 末行與表 19 第二行裏之數值，可見該二行數值極相切近；因此得知皮爾生曲線型 I 配合表 19 裏燕麥產量論據頗稱滿意，此滿意情形更可由圖 15 見之。圖 15 裏之矩形圖代表觀察論據，其中之光滑曲線代表配合之曲線，即根據表 20 第七行繪入者。

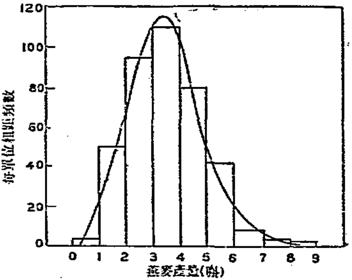


圖 15. 直方圖及式(50)之軌跡。

第十節　皮爾生曲線型 IV

1. 準則及方程式　關於皮爾生曲線型 IV 之頻數分配以 $0<k_2<1$ 為判定之準則。如果 $0<k_2<1$，於是 c_0 與 c_2 之符號相同，$c_1^2-4c_0c_2<0$，$c_0+c_1x+c_2x^2=0$ 之兩根均為虛數，並且 $\dfrac{x-a}{c_0+c_1x+c_2x^2}$ 可以寫成下形：

$$(51)\qquad \frac{\left(x+\dfrac{c_1}{2c_2}\right)-\left(a+\dfrac{c_1}{2c_2}\right)}{c_2\left[\left(x+\dfrac{c_1}{2c_2}\right)^2+\dfrac{4c_0c_2-c_1^2}{4c_2^2}\right]}.$$

將式(51)代入式(36),並以 $x+\frac{c_1}{2c_2}$ 代替式中之 x(即是將原點向左移動 $\frac{c_1}{2c_2}$ 單位,亦即以 $X_0=\bar{X}-\frac{hc_1}{2c_2}$ 為現在曲線之原點之橫坐標);我們有

$$\frac{dy}{y}=\frac{-(2p_0x+p_1p_2)}{x^2+p_1^2}dx.$$

並且積分之,

(52) $$y=y_0\left(1+\frac{x^2}{p_1^2}\right)^{-p_0}e^{-p_2\tan^{-1}x/p_1},$$

彼處 $p_0=-\frac{1}{2c_2}$, $p_1=\frac{\sqrt{4c_0c_2-c_1^2}}{2c_2}$, $p_2=\frac{a+\frac{c_1}{2c_2}}{c_2p_1}$ 並且 y_0 為待定之常數.

式(52)為皮爾生頻數曲線型 IV 之方程式.

2. 原點橫坐標 X_0 我們不難證明原點橫坐標

(53) $$X_0=\bar{X}-\frac{hp_1p_2}{r},\quad r=2p_0-2.$$

故 X_0 之值可用該式求之.

3. 常數 y_0 常數 y_0 之值可決定之如下:

因為 $$\int_{-\infty}^{\infty}y_0\left(1+\frac{x^2}{p_1^2}\right)^{-p_0}e^{-p_2\tan^{-1}x/p_1}dx=N$$

於是 $$y_0=N\div\left[\int_{-\infty}^{\infty}\left(1+\frac{x^2}{p_1^2}\right)^{-p_0}e^{-p_2\tan^{-1}x/p_1}dx\right]$$

$$=N\div\int_{-\pi/2}^{\pi/2}p_1\cos^{2p_0-2}\zeta\, e^{-p_2\zeta}d\zeta,\quad \zeta=\tan^{-1}\frac{x}{p_1}$$

$$=N\div\left[p_1e^{-p_2\pi/2}\int_0^{\pi}\sin^r\theta\, e^{p_2\theta}d\theta\right],\ r=2p_0-2\ \text{並且}\ \theta=\frac{\pi}{2}-\zeta$$

(54) $$=\frac{N}{p_1F(r,p_2)},$$

彼處 $$F(r,p_2)=e^{-p_2\pi/2}\int_0^{\pi}\sin^r\theta\, e^{p_2\theta}d\theta.$$

$$=\frac{(\cos\phi)^{r+1}}{\sqrt{r-1}\,e^{-p_2\phi}}H(r,\ p_2)$$

並且　$$\phi=\tan^{-1}\frac{p_2}{r}.$$

既然　$$F(r,p_2)=\frac{(\cos\phi)^{r+1}}{\sqrt{r-1}\,e^{-p_2\phi}}H(r,p_2),$$ 則

(55)　$$\log F(r,p_2)=\log\frac{\cos^{r+1}\phi}{\sqrt{r-1}\,e^{-p_2\phi}}+\log H(r,p_2).$$

故 $\log F(r,p_2)$ 之值直接或間接利用末式及附錄 III 表 F 或 G 求之，因此得以求得 y_0 之值.

雖然 y_0 之值可用式 (55) 求之，但當 r 相當大時，仍可利用下式以求其近似值：

(56)　$$y_0=\alpha\beta,$$

彼處　$$\alpha=\frac{N}{p_2}\sqrt{\frac{r}{2\pi}},\quad \beta=\frac{e^{\frac{\cos^2\phi}{3r}-\frac{1}{12r}-\phi r\tan\phi}}{(\cos\phi)^{r+1}}.$$

$\log H(r,p_2)$ 之值固可由表直接查得，但當 $\phi>60°$ 時，其值仍可用下式計算之：

(57)　$$\log H(r,p_2)=\frac{1}{2}q^2\log e+\log\sqrt{2\pi}+\frac{1}{2}\log\left(1-\frac{1}{r}\right).$$

彼處　$$q=\sqrt{\frac{1-4\cos^2\phi}{6r}}$$

求得 $\log H(r,p_2)$ 之值後再求 $\log F(r,p_2)$ 之值，由此以便計算 y_0 之值.

4. 公式　除 X_0 之值用式 (53)，y_0 之值用式 (54) 或 (56) 計算外，式 (52) 中 p_0、p_1、p_2 及 ϕ 之值可用下式計算之：

$$
(58)\quad \begin{cases}
K = 10\beta_2 - 12\beta_1 - 18 & p_1 = \left| L^2 - \mu_2 H \right|^{\frac{1}{2}} \times \frac{1}{2} \div \kappa_1 \\
p_0 = K \div \kappa_1 \times \frac{1}{2} & R = 2\kappa_1^2 p_1 \\
r = K \div \kappa_1 - 2 & p_2 = (2\kappa_1 - K) \div R \times L \\
\bar{L} = (\beta_2 \div 3) \div \mu_2 \times \mu_3 & \phi = \tan^{-1} \frac{p_2}{r}
\end{cases}
$$

5. 舉例　試以皮爾生頻數曲線之適宜型配合下表所列之論據。

a. 計算 β_1、β_2、κ_1 及 κ_2

表 21.　麥稈高度之分配（論據見參考書 27）

麥稈高度（公分）	頻數 f	步差數 x	xf	x^2f	x^3f	x^4f
25—	4	−6	−24	144	−864	5,184
30—	7	−5	−35	175	−875	4,375
35—	27	−4	−108	432	−1,728	6,912
40—	76	−3	−228	684	−2,052	6,156
45—	151	−2	−302	604	−1,208	2,416
50—	237	−1	−237	237	−237	237
55—	196	0	0	0	0	0
60—	83	1	83	83	83	83
65—	27	2	54	108	216	432
70—	15	3	45	135	405	1,215
75—	2	4	8	32	128	512
總　計	825	……	−744	2,634	−6,132	27,522

$N = 825$　　　　$\mu_2 = 2.379451$

$\nu_1 = 33.360000$　　　　$F = 5.661787$

$\nu_3 = -7.432727$　　　　$G = 13.471945$

$\nu_2 = 3.192727$　　　　$\mu_3 = -0.261805$

$\nu_1=-0.901818$　　$\beta_1=0.005088$

$A=2.705454$　　$\beta_2=3.557762$

$B=3.607272$　　$\kappa_1=1.100260$

$C=1.626552$　　$H=62.564234$

$D=4.879656$　　$\kappa_2=0.003497$

$E=2.200281$　　(合乎型 IV 之準則)

b. 計算 p_0、p_1、p_2、ϕ、X_0 及 y_0 [公式見式 (58) 及 (56)]

$K=17.516564$　　$\cos\phi=0.9982499$

$p_0=7.960193$　　$\frac{1}{3r}=0.023946$

$r=13.920386$　　$\alpha=221.857587$

$L=-0.721534$　　$\beta=0.995227$

$p_1=5.534972$　　$y_0=220.798661$

$R=13.400964$　　$\log y_0=2.3439965$

$p_2=0.824646$　　$X_0=54.630375$

$\phi=3°23'25''=0.059171$ 强

c. 寫出方程式　欲求之方程式可寫出之如下：

$$y=220.7987\left(1+\frac{x^2}{30.6360}\right)^{-7.9602}e^{-.8246\tan^{-1}\frac{x}{5.5350}} \tag{59}$$

d. 計算縱坐標及面積　曲線 (59) 之縱坐標及其下面積，可依據下面表 22 所列之步驟計算之.

e. 比較配合情形　由於比較表 21 第二行與表 22 最末行裏之數值，並且比較代表原始頻數之直方圖（根據表 21 第二行繪之者）與配

表 22.　曲線(59)縱坐標及其下面積之計算

麥桿高度 X (1)	$\frac{X-X_0}{h}=x$ (2)	$\frac{x}{p_1}$ (3)	$\tan^{-1}\frac{x}{p_1}$ (4)	$\log\left(1+\frac{x^2}{p_1^2}\right)$ (5)	$-p_1\log e\times(4)$ (6)	$-p_0\times(5)$ (7)	y (8)	面積 (9)
27.5	−5.426075	−.980325	−.775467	.2924658	.2777244	−2.3282434	2.0	2.1
32.5	−4.426075	−.799656	−.674533	.2146982	.2415767	−1.7090391	7.5	8.1
37.5	−3.426075	−.619937	−.554265	.1405677	.1985040	−1.1213341	26.4	27.7
42.5	2.426075	−.438317	−.413097	.0763207	.1479462	−.6075275	76.7	78.2
47.5	−1.426075	−.257648	−.252164	.0279131	.0903098	−.2221937	163.0	162.0
52.5	−.426075	−.076979	−.076828	.0025660	.0275151	−.0204259	224.4	220.3
57.5	.573925	.103691	.103324	.0046446	−.0370044	−.0369719	186.3	184.1
62.5	1.573925	.284360	.277051	.0337697	−.0992228	−.2688133	91.6	95.5
67.5	2.573925	.465029	.435288	.0850236	−.1558937	−.6768043	32.5	34.1
72.5	3.573925	.645699	.573351	.1513472	−.2053395	−1.2047529	8.6	9.3
77.5	4.573925	.826368	.690615	.2260642	−.2473363	−1.7995147	2.4	2.2
總計	……	……	……	……	……	……	……	523.9

合之頻數曲線(根據表 22 第八行給之者)得知,以皮爾生曲線型 IV 配合表 21 裏所列之論據尚覺適切(參閱圖 16).

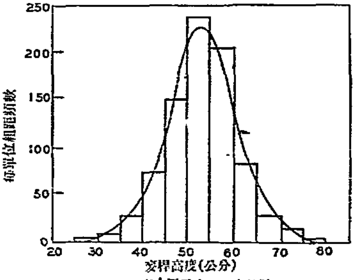


圖 16. 直方圖及式(59)之軌跡。

第十一節 皮爾生曲線型 VI

1. 準則及方程式 屬於皮爾生曲線型 VI 之頻數分配以 $k_2>1$ 爲判定之準則。如果 $k_2>1$，於是方程式 $c_0+c_1x+c_2x^2=0$ 之兩根不等，但其符號相同。如此可將 $c_0+c_1x+c_2x^2$ 分成兩個實因子 $(x-r_1)$ 及 $(x-r_2)$，彼處 r_1 及 r_2 之值同時爲正或同時爲負。用部分分數法得

$$\frac{x-a}{c_0+c_1x+c_2x^2}=\frac{1}{c_2(r_1-r_2)}\left[-\frac{a-r_1}{x-r_1}+\frac{a-r_2}{x-r_2}\right].$$

將其代入式 (36) 內積分之，我們有

$$\log y=-\frac{a-r_1}{c_2(r_1-r_2)}\log(x-r_1)+\frac{a-r_2}{c_2(r_1-r_2)}\log(x-r_2)+\text{積分常數}$$

將末式中之 x 以 $x+r_2$ 代之，即以 $X_0=\bar{X}+r_2h$ 爲現在曲線原點之縱坐標，亦即將原點向右移動 r_2 單位，我們有

(60) $$y=y_0(x-q_0)^{q_1}x^{-q_2},$$

彼處 y_0＝欲測定之常數；$q_0=r_1-r_2$，$r_1>r_2$；$q_1=-\frac{a-r_1}{c_2(r_1-r_2)}$ 並且 $q_2=-\frac{a-r_2}{c_2(r_1-r_2)}$。此爲皮爾生頻數曲線型 VI 之方程式。

2. 原點橫坐標 X_0 原點橫坐標 X_0 可用下式計算之：

(61) $$X_0=\bar{X}+hr_2$$

3. 常數 y_0 茲將常數 y_0 決定之如下：

因爲 $$\int_{q_0}^{\infty}y_0(x-q_0)^{q_1}x^{-q_2}dx=N$$

故 $$y_0=N\div\int_{q_0}^{\infty}(x-q_0)^{q_1}x^{-q_2}dx$$

$$=N\div\left[q_0^{q_1-q_2+1}\int_0^1(1-z)^{q_1}z^{q_2-q_1-2}dz\right],\quad z=\frac{q_0}{x}$$

$$=N\div[q_0{}^{q_1-q_2+1}\beta(q_1+1,\ q_2-q_1-1)]$$

$$(62)\qquad =\frac{Nq_0{}^{q_2-q_1-1}\,\Gamma(q_2)}{\Gamma(q_1+1)\,\Gamma(q_2-q_1-1)}.$$

4. 公式　除計算 X_0 及 y_0 分別公式 (61) 及 (62) 外，式 (60) 中之 q_0、q_1 及 q_2 之值可用下式計算之：

$$(63)\qquad \begin{cases} K=10\beta_2-12\beta_1-18 & (r_1,\ r_2)=P\pm Q,\ (r_1>r_2) \\ L=(\beta_2+3)\div\mu_2\times\mu_3 & r_1= \\ a=-L\div K & r_2= \\ M=\left\{L^2-\mu_2 H\right\}^{\frac{1}{2}} & q_0=r_1-r_2 \\ P=\frac{1}{2}\div\kappa_1\times(-L) & q_1=K\div M\times(a-r_1) \\ Q=\frac{1}{2}\div\kappa_1\times M & q_2=K\div M\times(a-r_2) \end{cases}$$

5. 舉例　試以皮爾生頻數曲線之適宜型配合下表裏所列之論據。

a. 計算 β_1、β_2、κ_1 及 κ_2

表 23. 麥稈數目之分配（論據見參考書 27）

麥稈數目 X	頻數 f	助變數 x	xf	x^2f	x^3f	x^4f
1	1	−4	−4	16	−64	256
2	57	−3	−171	513	−1,539	4,617
3	184	−2	−368	736	−1,472	2,944
4	111	−1	−111	111	−111	111
5	33	0	0	0	0	0
6	11	1	11	11	11	11
7	2	2	4	8	16	32
8	1	3	3	9	27	81
總計	400	……	−636	1,404	−3,132	8,052

$N=400$	$\mu_2=0.981900$
$\nu_4=20.130000$	$F=0.964128$
$\nu_3=-7.830000$	$G=0.946677$
$\nu_2=3.510000$	$\mu_3=0.873342$
$\nu_1=-1.590000$	$\beta_1=0.805688$
$A=4.770000$	$\beta_2=4.562794$
$B=6.360000$	$\kappa_1=0.708524$
$C=5.056200$	$H=44.875393$
$D=15.168600$	$\kappa_2=1.026888$
$E=12.059037$	(合乎型 VI 之準則)

b. 計算 q_0、q_1、q_2、y_0 及 X_0

$K=17.959684$	$q_2=84.818042$
$L=6.726658$	$\log N=2.6020600$
$a=-0.374542$	$(q_2-q_1-1)\log q_0=4.5400335$
$M=1.088476$	$\log\Gamma(q_2)=126.1698592$
$P=-4.746949$	$-\log\Gamma(q_1+1)=-80.9768850$
$Q=0.768129$	$-\log\Gamma(q_2-q_1-1)=-22.8907738$
$r_1=-3.978820$	$\log y_0=29.4442939$
$r_2=-5.515078$	$y_0=2.7816\times10^{29}$。
$q_0=1.536258$	$X_0=-2.105078$
$q=59.470025$	

寫出方程式　若將求得之常數代入式(60)內,則有下列方程式:

(64) $$y = 2.7816 \times 10^{-29}(x-1.5363)^{59.4703}x^{-51.8\,80}.$$

表 24. 曲線 (64) 縱坐標及其下面積之計算

麥桿數目 X	$\frac{X-X_0}{h}=x$	$x-q_0$	$q_1\log(x-q_0)$	$-q_2\log x$	y	面積
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1.0	3.105078	1.568820	11.6307431	−41.7366260	.2	2.8
2.0	4.105078	2.568820	24.3668734	−52.0207203	61.7	64.0
3.0	5.105078	3.568820	32.8586518	−60.0513773	178.5	170.8
4.0	6.105078	4.568820	39.2385604	−66.6407892	110.2	110.0
5.0	7.105078	5.568820	44.3505561	−72.2284067	36.9	38.5
6.0	8.105078	6.568820	48.6159902	−77.0790064	9.6	10.4
7.0	9.105078	7.568820	52.2758290	−81.3645652	2.3	2.5
8.0	10.105078	8.568820	55.4808352	−85.2030905	.5	.6
總計	…	…	…	…	…	399.9

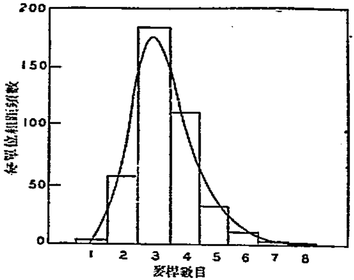


圖 17. 直方圖及式(64)之軌跡。

d. 計算縱坐標及面積 我們可計算上式之縱坐標及其下面積，其

手續可總括之如下:

c. 比較配合情形　用皮爾生頻數曲線型 VI 配合表 23 裏論據之適度情形,可由上面圖 17 見之. 圖 17 爲代表原始觀察頻數 (表 23 第二行)之直方圖及配合之曲線 [方程式 (64)].

第十二節　常態曲線

1. 準則及方程式　常態曲線皮爾生頻數曲線型之一. 其以 $\beta_1=0$ 及 $\beta_2=3$ 爲判定頻數分配屬於此型之準則. 如果 $\beta_1=0$ 並且 $\beta_2=3$, 於是由式 (45) 可證 $a=c_1=c_2=0$. 因之式 (36) 變做

$$\frac{dy}{y}=\frac{x}{c_0}dx.$$

積分之,我們有 $\log y=\frac{x^2}{2c_0}+$ 積分常數,即

(65)
$$y=y_0e^{x^2/2c_0}$$

彼處 y_0 爲欲測定之常數.

2. 原點橫坐標 X_0　原點在平均數處,即

(66)
$$X_0=\bar{X}.$$

3. 常數 y_0　常數 y_0 可決定之如下:

因爲
$$\int_{-\infty}^{\infty}ydx=N,$$

即是
$$y_0\int_{-\infty}^{\infty}e^{x^2/2c_0}dx=N.$$

所以
$$\begin{aligned}y_0&=N\div\int_{-\infty}^{\infty}e^{x^2/2c_0}dx\\&=N\div\sqrt{-2c_0}\int_0^{\infty}t^{-\frac{1}{2}}e^{-t}dt,\quad \sqrt{t}=\frac{x}{\sqrt{-2c_0}}\\&=N\div\sqrt{-2c_0\pi},\end{aligned}$$

彼處 c_0 尙待測定。

由式 (45) $c_0=Hu_2/4k_1K$; 但因 $H=k_1=0$, c_0 變成不定形式, c_0 之值不得不引用他法求之。

表 25. 曲線 (68) 縱坐標及其下面積之計算

體重 X (1)	頻數 f (2)	$x=\frac{X-X_0}{h}$ (3)	x^2 (4)	$\frac{x^2}{2c_0}\lg e$ (5)	縱坐標 y (6)	面 積 (7)
29.5	1	−4.553	20.729809	$\bar{3}$.837798	1.9	2.3
33.5	14	−3.553	12.623809	$\bar{2}$.683286	13.3	14.7
37.5	56	−2.553	6.517809	$\bar{1}$.320166	57.8	60.0
41.5	172	−1.553	2.411809	$\bar{1}$.748439	155.1	155.3
45.5	245	−0.553	0.305809	$\bar{1}$.968103	256.9	252.9
49.5	263	0.447	0.199809	$\bar{1}$.979159	263.5	259.2
53.5	156	1.447	2.093809	$\bar{1}$.781607	167.2	167.0
57.5	67	2.447	5.987809	$\bar{1}$.375448	65.6	67.8
61.5	23	3.447	11.881809	$\bar{2}$.760680	15.9	17.4
65.5	3	4.447	19.775809	$\bar{3}$.937304	2.4	2.9
總 計	1000	...	...	...	...	999.5

我們知道 $\sigma_x^2=\frac{1}{N}\int_{-\infty}^{\infty}x^2y\,dx$.

$$=\frac{y_0}{N}\int_{-\infty}^{\infty}x^2e^{x^2/2c_0}dx$$

$$=\frac{y_0}{N}\left[(-2c_0)^{\frac{3}{2}}\int_0^{\infty}t^{\frac{1}{2}}e^{-t}dt\right],\quad \sqrt{t}=\frac{x}{\sqrt{-2c_0}}$$

$$=-c_0,\quad 因\ y_0=N\div\sqrt{-2c_0\pi}.$$

於是,我們有 $\qquad c_0=-\sigma_x^2$.

因此
$$y_0=\frac{N}{\sigma_x\sqrt{2\pi}}.$$

4. 公式　c_0 之值可用下式求之:

$$c_0=-u_{2x}. \tag{67}$$

5. 舉例　試以皮爾生頻數曲線之適宜型配合表 14 裏所列之論據.

a. 計算 β_1 及 β_2　前已根據表 14 所列之論據算得 $\beta_1=0.013102$, $\beta_2=2.890238$ (見前章第九節), 因是知該論據(頻數數列)得用常態曲線配合之.

b. 計算 c_0, y_0 及 X_0 之值.

$c_0=-2.081858$　(見前章第九節),

$y_0=276.493389$,

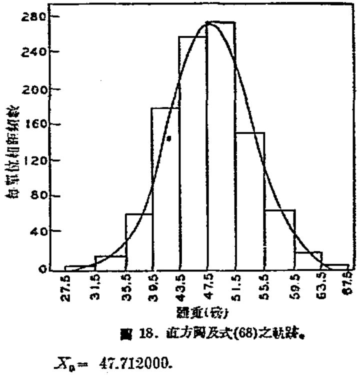


圖 18. 直方圖及式(68)之軌跡.

$X_0=$ 47.712000.

c. 寫出方程式　所求之常態曲線方程式得寫出之如下：

$$y = 276.493389\,e^{-0.240170x} \tag{68}$$

d. 計算縱坐標及面積　我們由式 (68) 得分別計算常態曲線之縱坐標及其下面積.計算之手續詳見表 25.

e. 比較配合情形　為比較配合情形,爰檢閱 18, 如上.

第十三節　皮爾生曲線型 II

1. 準則及方程式　屬於皮爾生曲線型 II 之頻數分配以 $\beta_1=0$, $\beta_2<3$ 為判定之準則. 如果 $\beta_1=0$, $\beta_2<3$; 於是 $c_1=a=0$, 方程式 $c_0+c_1x+c_2x^2=0$ 兩根之大小相等而符號相反.因之 $a_1=a_2$ (譬如 $=a'$) 並且 $m_1=m_2$ (譬如 $=m'$). 故式 (46) 變做

$$y = y_0\left(1-\frac{x^2}{a'^2}\right)^{m'} \tag{69}$$

此為皮爾生頻數曲線型 II 之方程式.

2. 原點橫坐標 X_0　原點橫坐標 X_0 可用下式求之:

$$X_0 = \bar{X} + ha'. \tag{70}$$

3. 常數 y_0　因 $a_1=a_2=a'$, $m_1=m_2=m'$; 故由式 (48) 得

$$y_0 = \frac{Na'^{2m'}\Gamma(2m'+2)}{(2a')^{2m'+1}[\Gamma(m'+1)]^2}. \tag{71}$$

4. 公式　除 X_0 及 y_0 之值分別用式 (70) 及 (71) 計算外,常數 a' 及 m' 之值可用下式求之:

$$\begin{cases}\beta_1=0, & a' = -4\mu_2\beta_2 \div \kappa_1', \\ \kappa_1' = 2\beta_2-6, & m' = (-5\beta_2+9)\div\kappa_1'.\end{cases} \tag{72}$$

5. 舉例　試以皮爾生頻數曲線之適宜型配合下列論據.

a. 計算 β_1、β_2、及 κ_1

表 26.　小麥平均穗數之分配(論據見參攷書 27)

小麥穗平均數 X	頻數 f	助變數 x	xf	x^2f	x^3f	x^4f
10—	8	−5	−40	200	−1,000	5,000
15—	28	−4	−112	448	−1,792	7,168
20—	57	−3	−171	513	−1,539	4,617
25—	100	−2	−200	400	−800	1,600
30—	115	−1	−115	115	−115	115
35—	102	0	0	0	0	0
40—	61	1	61	61	61	61
45—	26	2	52	104	208	416
50—	3	3	9	27	81	243
總計	500	...	−516	1,868	−4,896	19,220

$N=500$		$u_2'=$	2.587643
$\nu_4=$	38.440000	$F=$	6.695896
$\nu_3=$	−9.792600	$G=$	17.326588
$\nu_2=$	3.736000	$u_3'=$	−0.423554
$\nu_1=$	−1.032000	$\beta_1=$	0.010354
$A=$	3.096000	$\beta_2=$	2.560207
$B=$	4.128000	因爲	
$C=$	2.130048	$\beta_1=0$	
$D=$	6.390144	$\beta_2<3$	
$E=$	3.297314	故近乎型 II 準則.	

b. 計算 a', m', X_0 及 y_0

$\beta_1=0$		$\log(2m'+2)=$	6.3882366
$k_1'=$	$-.867586$	$-(2m'+1)\log 2a'=$	-10.2640820
$a'^2=$	30.615651	$-2\log\Gamma(m'+1)=$	-3.3203600
$m'=$	4.415741	$\log y_0=$	2.0643052
$\log N=$	2.698700	$y_0=$	115.959358
$2m'\log a'=$	6.5615412	$X_0=$	32.340000

c. 寫出方程式　根據算得之結果得將欲求之方程式寫出之如下：

$$y=115.9594\left(1-\frac{x^2}{30.6157}\right)^{4.4157} \tag{73}$$

d. 計算縱坐標及面積　曲線(73)之縱坐標及其下面積可按表 27 所列之手續計算之.

表 27.　曲線(73)縱坐標及其下面積之計算

小穗平均數 X (1)	$\frac{X-X_0}{h}=x$ (2)	$\frac{x}{a}$ (3)	$\log\left(1-\frac{x^2}{a^2}\right)$ (4)	$m'\times(4)$ (5)	y (6)	面積 (7)
12.5	-3.968	$-.717133$	$\bar{1}.6863862$	-1.3848373	4.8	5.5
17.5	-2.968	$-.536404$	$\bar{1}.8526452$	$-.6506806$	25.9	26.6
22.5	-1.968	$-.355675$	$\bar{1}.9412605$	$-.2583784$	64.0	64.0
27.5	$-.968$	$-.174946$	$\bar{1}.9865003$	$-.0596112$	101.1	100.2
32.5	.032	.005783	$\bar{1}.9999854$	$-.0000645$	115.9	114.6
37.5	1.032	.186512	$\bar{1}.9846233$	$-.0678995$	99.2	98.3
42.5	2.032	.367241	$\bar{1}.9370834$	$-.2778234$	61.2	61.2
47.5	3.032	.547970	$\bar{1}.8449298$	$-.6847498$	24.0	24.7
52.5	4.032	.728699	$\bar{1}.6711707$	-1.4520250	4.1	4.8
總計	...	...	...	...	...	499.9

e. 比較配合情形 配合之適度情形可由下圖見之:

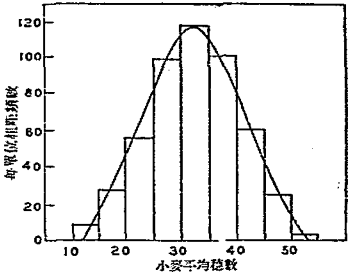


圖 19. 直方圖及式(73)之軌跡。

第十四節 皮爾生曲線型 III

1. 準則及方程式 關於皮爾生曲線型III之頻數分配以 $\kappa_1=0$ 為判定之準則.如果 $\kappa_1=0$, 則 $c_2=-\frac{\kappa_1}{\kappa}=0$. 所以式(36)變做

(74) $$\frac{1}{y}\frac{dy}{dx}=\frac{x-a}{c_0+c_1x}$$

將式(74)積分之得 $\log y=\frac{1}{c_1}x-\frac{a+\frac{c_0}{c_1}}{c_1}\log\left(x+\frac{c_0}{c_1}\right)+C.$

設此 C 為積分常數.將式中之 x 以 $(x+a)$ 代之,我們有

$$\log y=\frac{1}{c_1}(x+a)-\frac{a+\frac{c_0}{c_1}}{c_1}\log\left(x+a+\frac{c_0}{c_1}\right)+C.$$

$$\log y=-s_2(x+a)+s_1\log(x+s_0)+C.$$

彼處 $s_0=a+\frac{c_0}{c_1}$, $s_1=-\frac{s_0}{c_1}=s_0s_2$, 並且 $s_2=-\frac{1}{c_1}=\frac{s_1}{s_0}$. 於是

$$y=y_0\left(1+\frac{x}{s_0}\right)^{s_1}e^{-s_2x}, \tag{75}$$

彼處 y_0 爲欲測定之常數.此爲皮爾生頻數曲線型 III 之方程式.

2. 原點縱坐標 X_0 X_0 之值可用下式求之:

$$X=\bar{X}+ha. \tag{76}$$

3. 常數 y_0 y_0 計算之公式可推演之如下:

因爲 $$\int_{-s_0}^{\infty}y_0\left(1+\frac{x}{s_0}\right)^{s_1}e^{-s_2x}dx=N.$$

即是
$$\begin{aligned}
y_0&=N\div\left[\int_{-s_0}^{\infty}\left(1+\frac{x}{s_0}\right)^{s_1}e^{-s_2x}dx\right]\\
&=N\div\left[s_1^{-s_1}s_2^{-1}e^{s_1}\int_0^{\infty}z^{s_1}e^{-z}dz\right],\quad z=s_2(s_0+x)\\
&=N\div\left[s_1^{-s_1}s_2^{-1}e^{s_1}\Gamma(s_1+1)\right]\\
&=\frac{Ns_1^{s_1}s_2}{e^{s_1}\Gamma(s_1+1)}
\end{aligned}$$

$$=\frac{Ns_1^{s_1+1}}{s_0e^{s_1}\Gamma(s_1+1)}. \tag{77}$$

4. 公式 除 X_0 及 y_0 分別用式 (76) 及 (77) 計算外, s_0、s_1 及 s_2 之值可用下式計算之:

$$\begin{cases}
\beta_2=1.5\beta_1+3 & S=K\times L\\
K=40\beta_2-12\beta_1-18 & s_0=(\mu_2K^2-L^2)\div S\\
L=(\beta_2+3)\div\mu_2\times\mu_3 & s_1=s_0\div L\times K\\
a=-L\div K & s_2=s_1\div s_0.
\end{cases} \tag{78}$$

4. 舉例　試以皮爾生頻數曲線之適宜型配合下列論據。

a. 計算 β_1、β_2、κ_1 及 κ_2

表 28.　燕麥總產量之分配(論據見參攷書 27)

總產量(克)	頻數 f	助變數 x	xf	x^2f	x^3f	x^4f
0.0—	87	−3	−261	783	−2,349	7,047
2.0—	192	−2	−384	768	−1,536	3,072
4.0—	128	−1	−128	128	−128	128
6.0—	71	0	0	0	0	0
8.0—	12	1	12	12	12	12
10.0—	7	2	14	28	56	112
12.0—	3	3	9	27	81	243
總計	500	…	−738	1,746	−3,864	10,614

$N=500$　　$\mu_2=1.313424$

$\nu_4=21.228000$　　$F=1.725083$

$\nu_3=-7.728000$　　$G=2.265765$

$\nu_2=3.492000$　　$\mu_3'=1.303420$

$\nu_1=-1.476000$　　$\beta_1=.749815$

$A=4.428000$　　$\beta_2=4.062895$

$B=5.904000$　　$\kappa_1=-.123655$

$C=4.357152$　　因為

$D=13.071456$　　$\kappa_1 \doteq 0$

$E=9.646735$　　近乎型 III 準則

b. 計算 s_0、s_1、s_2、X_0 及 y_0

$\beta_2=$	4.124722	$\log N=$	2.6989700
$\kappa=$	14.249440	$(s_1+1)\log s_1=$	3.3979266
$L=$	7.070455	$-\log s_0=$	$-$ .3326036
$a=$	$-$ .496192	$-s_1\log e=$	$-$1.8825146
$s=$	100.750024	$-\log\Gamma(s_1+1)=$	$-$1.6043730
$s_0=$	2.150818	$\log y_0=$	2.2774054
$s_1=$	4.334650	$y_0=$	189.411087
$s_2=$	2.015349	$X_0=$	3.055616

c. 寫出方程式　所求之方程式爲

$$y=189.4111\left(1+\frac{x}{2.1508}\right)^{4.3346}\;{}^{-2.0153x} \tag{79}$$

d. 計算縱坐標及面積　由式 (79) 計算之縱坐標及面積分別列入表 29 末尾兩行.

表 29. 曲線 (79) 縱坐標及其下面積之計算

總產量 X (1)	$\frac{X-X_0}{h}=x$ (2)	$\log\left(1+\frac{x}{s_0}\right)$ (3)	$s_1\times(3)$ (4)	$-xs_2\log e$ (5)	y (6)	面積 (7)
1	$-$1.027808	$\bar{1}$.7177803	$-$1.2233236	.8995941	89.9	90.3
3	$-$.027808	$\bar{1}$.9943484	$-$.0233977	.0243391	189.3	182.9
5	.972192	.1619696	.7020815	.8509159	134.5	133.7
7	1.972192	.2826106	1.2250180	$-$1.7261709	59.7	61.2
9	2.972192	.3769213	1.6338219	$-$2.6014259	20.4	21.4
11	3.972192	.4543611	1.9694963	$-$3.4766809	5.9	6.3
13	4.972192	.5200595	2.2542759	$-$4.3519359	1.5	1.6
總計	...	...	...	...	...	497.4

e. 比較配合情形　下圖顯示用皮爾生曲線型 III 配合已給論據之美滿情形。

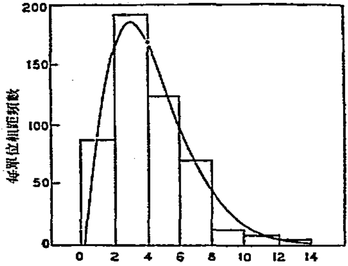


圖 20. 直方圖及式(79)之軌跡。

第十五節　皮爾生曲線型 V

1. 準則及方程式　關於皮爾生曲線型 V 之頻數分配以 $\kappa_2=1$ 為判定之準則。如果 $\kappa_2=1$，則 $c_0+c_1x+c_2x^2=0$ 之兩根相等，其均等於 $-\frac{c_1}{2c_0}$. 故用部分分數法得

$$\frac{x-a}{c_0+c_1x+c_2x^2}=\frac{1}{c_0}\left\{\frac{1}{x-r'}-\frac{a-r'}{(x-r')^2}\right\},\quad r'=-\frac{c_1}{2c_2}$$

如此，式 (36) 積分之結果為

$$\log y=\frac{1}{c_2}\log(x-r')+\frac{a-r'}{c_2(x-r')}+\text{積分常數}.$$

以 $x+r'$ 代替上式中之 x，即以 $X_0=\bar{X}+hr'$ 為現在曲線原點之橫坐標；我們有

$$\log y = \frac{1}{c_2}\log x + \frac{a-r'}{c_2 x} + \text{積分常數,即是}$$

(80) $$y = y_0 x^{-t_0} e^{-t_1/x}$$

彼處 $t_0 = -\frac{1}{c_2}$, $t_1 = -\frac{a-r'}{c_2}$, y_0 爲欲測定之常數. 此爲皮爾生頻數曲線型 V 之方程式.

2. 原點橫坐標 X_0　X_0 之值可用下式求之:

(81) $$X_0 = X + hP,$$

彼處 $$P = \tfrac{1}{2} + \kappa_1 \times (-L).$$

3. 常數 y_0　求 y_0 之值的公式可推求之如下:

因爲 $$\int_0^\infty y_0 x^{-t_0} e^{-t_1/x} dx = N,$$

於是 $$\begin{aligned} y_0 &= N \div \left[\int_0^\infty x^{-t_0} e^{-t_1/x} dx\right] \\ &= N \div \left[t_1^{-t_0+1}\int_0 z^{t_0-2} e^{-z} dz\right], \quad z = \frac{t_1}{x} \\ &= N \div [t_1^{-t_1+1}\Gamma(t_0-1)] \end{aligned}$$

(82) $$= \frac{N t_1^{t_0-1}}{\Gamma(t_0-1)}.$$

4. 公式　除 X_0 及 y_0 分别用式 (81) 及 (82) 計算外,常數 t_0 及 t_1 可用下式計算之:

(83) $$\begin{cases} T = -\beta_1 + 32 & a = -L \div \lambda \\ \beta_2 = \{\pm 6(\beta_1+4)^{\frac{3}{2}} + 39\beta_1 + 48\} \div T & \kappa_1 = 2\beta_2 - 3\beta_1 - 6 \\ (\text{此處 } \beta_2 \text{ 之值應與由論據算得者相近}) & t_0 = K \div \kappa_1 \\ \quad = 10\beta_2 - 12\beta_1 - 18 & P = \tfrac{1}{2} + \kappa_1 \times (-L) \\ L = (\beta_2 + 3) \div \mu_2 \times \mu_3 & t_1 = (a-P)t_0. \end{cases}$$

5. 舉例　試以皮爾生頻數曲線之適宜型配合下表所列之論據.

a. 計算 β_1、β_2、及 κ_1 及 κ_2

表 30. 死亡人口年齡之分配(論據見參考書 11)

年齡 X	死亡人口 f	助變數 x	xf	x^2f	x^3f	x^4f
30—	1	−9	−9	81	−729	6,561
35—	5	−8	−40	320	−2,560	20,480
40—	8	−7	−56	392	−2,744	19,208
45—	12	−6	−72	432	−2,592	15,552
50—	28	−5	−140	700	−3,500	17,500
55—	82	−4	−328	1,312	−5,248	20,992
60—	128	−3	−384	1,152	−3,456	10,368
65—	253	−2	−506	1,012	−2,024	4,048
70—	342	−1	−342	342	−342	342
75—	525	0	0	0	0	0
80—	438	1	438	438	438	438
85—	265	2	530	1,060	2,120	4,240
90—	53	3	159	477	1,431	4,293
95—	18	4	72	288	1,152	4,608
100—	4	5	20	100	500	2,500
總計	2,162	…	−658	8,106	−17,554	131,130

$N=2{,}162$　　　　$\mu_2'=3.573345$

$\nu_4=60.652174$　　　　$F=12.768794$

$\nu_3=-8.119334$　　　　$G=45.627306$

$\nu_2=3.749306$　　　　$\mu_3'=-4.752435$

$\nu_1 = -.304348$	$\beta_1 = .495002$
$A = .913044$	$\beta_2 = 3.996194$
$B = 1.217392$	$\kappa_1 = .507382$
$C = .185256$	$H = 29.427689$
$D = .555768$	$\kappa_2 = .823331$
$E = .084573$	近乎型 V 準則

b. 計算 t_0、t_1、X_0 及 y_0

$T = 31.504998$	$t_1 = -390.704299$
$\beta_2' = 3.951288$	$\log N = 3.3348557$
$\kappa = 15.572856$	$(t_0 - 1)\log t_1 = 94.0685412$
$L = -9.244993$	$-\log \Gamma(t_0 - 1) = -40.4705324$
$a = .593661$	$\log y_0 = 56.9328645$
$\kappa_1' = .417570$	$y_0 = 8.567704 \times 10^{56}$
$t_0 = 37.294001$	$X_0 = 131.328220$
$P = 11.069992$	

c. 寫出方程式　得將所求之方程式寫出之如下:

(84)
$$y = 8.567704 \times 10^{56} x^{-37.2940} e^{390.7018/x}.$$

d. 計算縱坐標及面積　由式 (84) 所計算之縱坐標面積及計算之手續見表 3x.

表 31. 曲線(84)縱坐標及其下面積之計算

年齡 X (1)	$\frac{X-X_0}{h}=x$ (2)	$-t_0\log\lvert x\rvert$ (3)	$\frac{-t_1\log e}{x}$ (4)	y (5)	面積 (6)
32.5	−19.765644	−48.3297061	−8.5846290	1.0	1.0
37.5	−18.765644	−47.4888196	−9.0420942	2.5	2.6
42.5	−17.765644	−46.6018751	−9.5510594	6.0	6.2
47.5	−16.765644	−45.6635319	−10.1207398	14.1	14.5
52.5	−15.765644	−44.6674613	−10.7626889	31.8	32.6
57.5	−14.765644	−43.6061602	−11.4915896	68.4	69.7
62.5	−13.765644	−42.4702741	−12.3263918	136.8	138.5
67.5	−12.765644	−41.2487576	−13.2919828	246.7	247.8
72.5	−11.765644	−39.9275430	−14.4217113	283.4	381.8
77.5	−10.765644	−38.4889045	−15.7613164	481.6	476.0
82.5	−9.765644	−36.9097423	−17.3752720	444.5	438.4
87.5	−8.765644	−35.1601800	−19.3574735	260.1	260.1
92.5	−7.765644	−33.1982806	−21.8501802	76.6	81.4
97.5	−6.765644	−30.9655484	−25.0797590	7.7	10.3
102.5	−5.765644	−28.3750586	−29.4296216	.1	.4
總計	···	···	···	···	2,161.3

e. 比較配合情形　為顯示用皮爾生曲線型 V 配合已給論據之美滿情形，爰繪下圖：

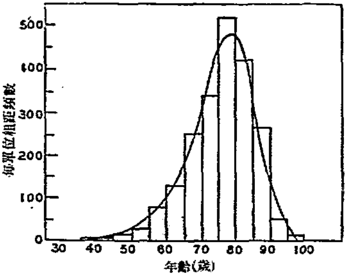


圖 21. 直方圖及式(84)之軌跡。

問 題 IV

1. 用代數方法及敍述方法證明:

   a. $P(n, r) = C(n, r)P(r, r)$. 及

   b. $C(n, r) = C(n, n-r)$.

2. 一箱內裝有五枚紅色橡球,其上各寫有 1, 2, 3, 4 及 5 號碼;裝有三枚白球,其上各寫有 1, 2, 及 3 號碼.由箱內任取兩球.問取得

   a. 一紅一白之機率爲何?

   b. 同色或同號之機率爲何?

3. 由裝有四枚白球 五枚紅球及六枚黑球之袋內任意取出三球試求

   a. 取不到黑球之機率.

   b. 取得二枚黑球之機率,及

c. 取得同樣顏色球之機率.

4. 狂風揚塵,灰塵顆粒大小相同,祇有褐黃之分,褐色者與黃色者之比爲 3 比 2. 假設五顆灰塵吹入眼中,試求兩褐三黃之機率.

5. 投擲六枚銅元一次,或者說一枚銅元投擲六次. 求擲得正面之機率:

a. 確爲三次,

b. 至多三次,

c. 至少三次,及

d. 至少一次.

6. 一個袋內裝有白球及黑球,白色者與黑色者之比爲 2 比 3. 連續取出三球,每次將球取出後還回袋中,並使與袋內其他各球混合均勻,如此不致影響下次取得白球或黑球之機會.求取得 0、1、2 及 3 個白球之機率.如將此實驗重複 125 次,求取得 0、1、2 及 3 個白球之理論頻數.

7. (*a*) 求 $C(18, x)$ 之值, x 由 0 至 18.

(*b*) 求 $2^x/3^{18}$ 之值, x 由 0 至 18.

(*c*) 證明 $\left(\frac{1}{3}+\frac{2}{3}\right)^{18}$ 可寫做

$$\sum_{x=0}^{18} f(x),\ \text{彼處}\ f(x)=C(18, x)2^x/3^{18}.$$

(*d*) 用 (*a*) 及 (*b*) 之結果,求 x 由 0 至 18 $f(x)$ 之值.

8. 試就下列之二項式,求 $\bar{x}$、σ、β_1 及 β_2 之值:

$$\left(\frac{1}{2}+\frac{1}{2}\right)^7,\ \left(\frac{1}{6}+\frac{5}{6}\right)^4,\ \left(\frac{1}{3}+\frac{2}{3}\right)^{18}.$$

9. 袋中有一圓鈔票三張，五圓鈔票四張。第一次任取三張鈔票，但每取出一張一圓鈔票要還入袋內三張兩圓鈔票，每取出一張五圓鈔票要還入一圓，兩圓及十圓鈔票各一張於袋內。第二次取出兩張鈔票。第一及第二次抽取鈔票各以 x 及 y 代表之。試列一表顯示 x 與 y 之聯合機率。

10. 當 $n=0$ 時，Γ 函數變做無限大，試證明之。

11. 證明 $\int_{-\infty}^{\infty}\phi(t)dt=1$，彼處 $\phi(t)=\frac{1}{(2\pi)^{1/2}}e^{-\frac{t^2}{2}}$.

12. 將 $\int_{-2}^{\infty}e^{-3x}(1+x/2)^7dx$ 化做 Γ 函數後，再求其值。

答　$(e^6 7!)/3(6)^7$.

13. 求 $\int_{6}^{\infty}e^{-2x}(x-6)^7dx$ 之值。　　答　$e^{-12}2^{-8}7!$.

14. 已知 $\int_{0}^{\infty}e^{-x}x^{-\frac{1}{2}}dx=(\pi)^{\frac{1}{2}}$，求 $\int_{0}^{\infty}e^{-x}x^{\frac{1}{2}}dx$ 之值。

15. 試求 10! 之實在值與用史德齡公式算得者之差。

16. 證明 $\int_{0}^{\pi/2}\cos^m\theta d\theta=\frac{1}{2}\beta[(m+1)/2,1/2]$.

17. 已給 $f(n)=n^{\frac{1}{2}}\beta(n/2,1/2)$，證明 $\lim\limits_{n\to\infty}f(n)=(2\pi)^{\frac{1}{2}}$.

18. 試以常態曲線配合二項分配 $(\frac{1}{3}+\frac{2}{3})^{18}$.

19. 試以常態曲線配合二項分配 $(\frac{1}{2}+\frac{1}{2})^{6}$.

20. 試以常態曲線配合第六章第七節例 1 裏所列之論據。

21. 舉例以明機率格紙之應用。

22. 討論皮爾生頻數曲線型 III.

23. 證明第三章式 (30).

24. 假設以 $\frac{x-\bar{x}}{\sigma}$ 代替式 (36) 中之 x, 求證

$$a=-\frac{(\beta_2+3)\sqrt{\beta_1}}{K}$$

$$c_0=-\frac{4\beta_2-3\beta_1}{K}$$

$$c_1=-\frac{(\beta_2+3)\sqrt{\beta_1}}{K}$$

$$c_2=-\frac{\kappa_1}{K}$$

25. 驗對式 (49).

26. 驗對式 (58).

27. 驗對式 (63).

28. 驗對式 (72).

29. 驗對式 (78).

30. 驗對式 (83).

31. 根據第九節所得之結果計算 X^2 [公式見第六章式 (30)], 並由查得之 P 值討論配合之適度情形.

32. 根據第十節所得之結果,計算 X^2 並由查得之 P 值討論配合之適度情形.

33. 根據第十一節所得之結果,計算 X^2 並由查得之 P 值討論配合之適度情形.

34. 根據第十二節所得之結果,計算 X^2 並由查得之 P 值討論配合之適度情形.

35. 根據第十三節所得之結果，計算 X^2 並由查得之 P 值討論配合之適度情形.

36. 根據第十四節所得之結果，計算 X^2 並由查得之 P 值討論配合之適度情形.

37. 根據第十五節所得之結果，計算 X^2 並由查得之 P 值討論配合之適度情形.

38. 試以皮爾生頻數曲線之一配合表 49 裏所列之論據.

39. 試以皮爾生頻數曲線之一配合表 50 裏所列之論據.

40. 試以皮爾生頻數曲線之一配合表 51 裏所列之論據.

41. 試以皮爾生頻數曲線之一配合表 52 裏所列之論據.

42. 試以皮爾生頻數曲線之一配合表 53 裏所列之論據.

43. 試以皮爾生頻數曲線之一配合表 54 裏所列之論據.

44. 試以皮爾生頻數曲線之一配合表 55 裏所列之論據.

45. 試以皮爾生頻數曲線之一配合表 56 裏所列之論據.

46. 試以皮爾生頻數曲線之一配合表 57 裏所列之論據.

47. 根據練習題 38 所得之結果計算 X^2 [公式見第六章式 (30)], 並由查得之 P 值討論配合之適度情形(與第六章第七節例 1 比較).

48. 根據練習題 39 所得之結果計算 X^2 之值，並由查得之 P 值推論配合之適度情形(與第六章第七節例 2 比較).

49. 根據練習題 40 所得之結果計算 X^2 之值，並由查得之 P 值推論配合之適度情形(與第六章第七節例 3 比較).

50. 根據練習題 41 所得之結果計算 X^2 之值，並由查得之 P 值推

論配合之適度情形(與第六章第七節例 4 比較).

51. 根據練習題 42 所得之結果計算 X^2 之值,並由查得之 P 值推論配合之適度情形(與第六章第七節例 5 比較).

52. 根據練習題 43 所得之結果計算 X^2 之值,並由查得之 P 值推論配合之適度情形(與第六章第七節例 6 比較).

53. 根據練習題 44 所得之結果計算 X^2 之值,並由查得之 P 值推論配合之適度情形(與第六章第七節例 7 比較).

54. 根據練習題 45 所得之結果計算 X^2 之值,並由查得之 P 值推論配合之適度情形(與第六章第七節例 8 比較).

55. 根據練習題 46 所得之結果計算 X^2 之值,並由查得之 P 值推論配合之適度情形(與第六章第七節例 9 比較).

第五章　曲線配合

第一節　實驗式

在前章裏曾經研究過各種頻數曲線配合之方法．另有一種統計資料，其不屬於頻數分配，但也可用各種適宜曲線配合之，配合此種統計資料之法爲本章所欲論述者．

當一變數爲他種變數之函數時，二者之間常是有一種關係存在，但此關係常是不知道的．即如第四節例1裏所列之論據(表32第一及第二兩行裏所列者)，我們並不知道各年度與紙煙產量之關係，我們所知道者祇是一組成雙的觀察數值而已．然而此等數值可視做點之坐標．如此，年度可視做橫坐標——即自變數；紙煙產量可視做縱坐標——即倚變數．我們的問題是，如何方能求出一算式 $Y=f(X)$ 表示論據所啓示之函數關係．配合觀察值所用之方程式，吾人所能找到者，叫做實驗式(empirical formula)，以示與用算學方法推得之理解式(rational formula)有別．

曲線配合(curve fitting)者乃係一種方法，舉凡尋求描寫已給論據最適宜之實驗式，找到適宜算式後用何步驟配合之，以及用何方法測定算式內所含之常數等等問題，皆在研究之列．自他方面言之，曲線配合係一種手續，用牠能將一羣參差不齊之對應值光滑之，因此曲線配合

方法或手續，又有時被稱做論據之修勻(smoothing the given data).

實驗式或爲直線、或爲二次、或爲高次、或爲指數、或爲對數、或爲超級函數(transcendental function)，如此等等，不一而足，將在以下各節擇其重要者討論之。

第二節　直線式

直線爲曲線中之最簡單者，其式曰直線式．如果有兩個變數，直線式之一般形式爲

$$LX+MY=N, \tag{1}$$

彼處 L, M 及 N 爲任意常數.

若 $M\neq 0$, 則上式可用 M 除之，並且得到下列最通用之形式

$$Y=a+bX, \tag{2}$$

彼處 $a=\frac{N}{M}$, $b=-\frac{L}{M}$

定理 1. 若直線方程式之形式如式(2)所表示者，則其中之 a 表示直線在 Y 軸上之截距(intercept on Y-axis)，並且 b 表示直線之斜率(slope).

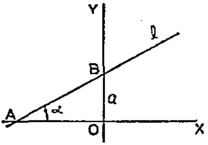


圖 22. 直線 $Y=a+bX$.

證: 如圖,設直線 l 與 X 軸及 Y 軸各相交於 A 及 B, 與 X 軸所成之傾角為 α, 並且其方程式為式 (2). 於是,若使式 (2) 中之 $X=0$, 則得 $Y=a=OB$, 即 l 在 Y 軸上之截距;若使式 (2) 中之 $Y=0$, 則得 $OA=-\frac{a}{b}$, 即 $AO=\frac{a}{b}$. 故直線 l 之斜率

$$\tan\alpha=\frac{OB}{AO}=\frac{a}{\frac{a}{b}}=b.$$ 證訖.

因為式 (2) 中之 a 為直線在 Y 軸上之截距,並且 b 為直線之斜率;故稱式 (2) 代表直線之方程式曰斜斷式 (slope-intercept form).

若 $L\neq 0$, 則式 (1) 兩邊可用 L 除之,並且式 (1) 變做

(3) $$X=a'+b'Y,$$

彼處 $a'=\frac{N}{L}$, $b'=-\frac{M}{L}$.

在特殊情形時,式 (1) 中之 L 或 M 或 N 為零. 如果 $L=0$, 於是 $Y=\frac{N}{M}$. 此為平行於 X 軸之直線; 如果 $M=0$, 則方程式之形式為 $X=\frac{N}{L}$. 此為平行於 Y 軸之直線. 如果 $N=0$, 則式 (1) 變做 $LX+MY=0$, 此為經過原點之直線.

直線之方程式乃是一變數等於他變數之函數, 其一變數對於他變數之變化率恆為一定. 其證明詳見下列之定理 (2) 中. 何謂變化率? 意即 Y 對於 X 之變化,亦即 X 變化一單位 Y 值之變化. 如此若 (X_1, Y_1) 及 (X_2, Y_2) 為兩變數值,則變化率為 $\frac{Y_2-Y_1}{X_2-X_1}$ 所表示之比值.

定理 2. 如果 $Y=f(X)$ 為一直線,則 Y 值對於 X 之變化率恆為一定.

證: 由假設，$f(X)$ 可以寫做 $a+bX$，其中 a 及 b 為常數. 如果 (X_1, Y_1) 及 (X_2, Y_2) 為在此直線上兩點，於是此二點之縱橫坐標一定能滿足所假設之方程式. 所以

$$Y_1 = a + bX_1,$$

$$Y_2 = a + bX_2.$$

減之得
$$Y_2 - Y_1 = b(X_2 - X_1),$$

由是
$$b = \frac{Y_2 - Y_1}{X_2 - X_1},$$

因 b 為常數，故知 Y 對於 X 之變化率 $\frac{Y_2 - Y_1}{X_2 - X_1}$ 恆為一常數. 證訖.

定理 . 如果 Y 對於 X 之變化率恆為一定，則 $Y=f(X)$ 為一直線式.

證: 讓 b 表示變化率，(X_1, Y_1) 為已給之一雙數值，並且 (X, Y) 為所取之任意一點. 於是我們有

$$b = \frac{Y - Y_1}{X - X_1},$$

由是

(4) $$Y - Y_1 = b(X - X_1),$$

即

(5) $$Y = (Y_1 - bX_1) + bX,$$

此式與式 (2) 比較，得知 $a = Y_1 - bX_1$，並且式 (5) 為一直線式. 證訖.

系. 因為 $b = \frac{Y_2 - Y_1}{X_2 - X_1}$.

故式 (4) 可寫做
$$Y - Y_1 = \frac{Y_2 - Y_1}{X_2 - X_1}(X - X_1).$$

其為用兩點表示之直線,故稱此直線之方程式曰兩點式 (two-point form).

定理 4. 若直線 l 法線與 X 軸所成之角為 α, 則式 (1) 可寫做

(6) $$X\cos\alpha+Y\sin\alpha=p.$$

其中 p 為由原點 O 至直線 l 之垂直距離. 用式 (6) 表示之直線方程式, 通常稱之曰法線式 (normal form).

證: 因由式 (1) 及圖 23 得知

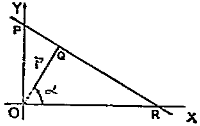


圖 23. 直線 $X\cos\alpha+Y\sin\alpha=p$.

$OP=\frac{N}{M}$, 即使式 (1) 中之 $X=0$ 得之者;

$OR=\frac{N}{L}$, 即使式 (1) 中之 $Y=0$ 得之者;

$$PQ=(\overline{OP}^2-p^2)^{\frac{1}{2}}$$

$$=\left(\frac{N^2}{M^2}-p^2\right)^{\frac{1}{2}};$$

$$QR=(\overline{PQ}^2-p^2)^{\frac{1}{2}}$$

$$=\left(\frac{N^2}{L^2}-p^2\right)^{\frac{1}{2}}.$$

但 $$PR=PQ+QR,$$

及 $$PR=(\overline{OP}^2+\overline{OR}^2)^{\frac{1}{2}},$$

於是 $$PQ+QR=(\overline{OI}^2+\overline{OR}^2)^{\frac{1}{2}}.$$

即 $$\left(\frac{N^2}{M^2}-p^2\right)^{\frac{1}{2}}+\left(\frac{N^2}{L^2}-p^2\right)^{\frac{1}{2}}=\left(\frac{N^2}{M^2}+\frac{N^2}{L^2}\right)^{\frac{1}{2}}.$$

平方之，並將同類項歸併之或消去之，得

$$\left(\frac{N^2}{L^2}-p^2\right)^{\frac{1}{2}}\left(\frac{N^2}{M^2}-p^2\right)^{\frac{1}{2}}=p^2,$$

再平方之並化簡之，我們有

$$(L^2+M^2)p^2=N^2.$$

於是 $$P=\pm\frac{N}{(L^2+M^2)^{\frac{1}{2}}}.$$

因為 $$\cos\alpha=\frac{p}{OR}=\frac{\pm\frac{N}{\sqrt{L^2+M^2}}}{\frac{N}{L}}=\frac{L}{\pm\sqrt{L^2+M^2}}.$$

並且 $$\sin\alpha=\pm\sqrt{1-\cos^2\alpha}=\pm\sqrt{1-\frac{L^2}{L^2+M^2}}=\frac{M}{\pm\sqrt{L^2+M^2}}.$$

可以式 (1) 之兩邊以 $\pm\frac{1}{\sqrt{L^2+M^2}}$ 除之，我們有式 (6) 彼處 $\cos\alpha=\frac{L}{\pm\sqrt{L^2+M^2}}$，$\sin\alpha=\frac{M}{\pm\sqrt{L^2+M^2}}$ 並且 $p=\frac{N}{\pm\sqrt{L^2+M^2}}$.

因為我們常是取 p 之值為正，故 p 為 $\frac{N}{\pm\sqrt{L^2+M^2}}$ 之絕對值. 即當 $N>0$ 時，$p=\frac{N}{+\sqrt{L^2+M^2}}$；當 $N<0$ 時，$p=\frac{N}{-\sqrt{L^2+M^2}}$.

系. 若直線方程式如式 (2) 之形式，並且其中之 a 值為正；則

$$\cos\alpha=\frac{1}{\sqrt{1+b^2}},$$

$$\sin\alpha=\frac{-b}{\sqrt{1+b^2}},$$

$$p=\frac{a}{\sqrt{1+b^2}}.$$

定理 5. 若二直線 l_1 及 l_2 之方程式各為

$$Y=a_1+b_1X,$$

及

$$Y=a_2+b_2X.$$

則 l_1 與 l_2 所夾之角 θ, 可由下式求之

(7) $$\tan\theta=\frac{b_2-b_1}{1+b_1b_2}.$$

證: 因已給二直線之方程式為

$$Y=a_1+b_1X,$$

$$Y=a_2+b_2X.$$

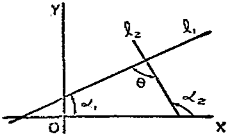


圖 24. 二直線所夾之角 θ.

於是由定理 1 得知

$$\tan\alpha_1=b_1,$$

$$\tan\alpha_2=b_2.$$

設α α_1 為第一條直線與 X 軸所成之角，α_2 為第二條直線與 X 軸所成之角 (見圖 24); 但由圖 24 可見 $\theta=\alpha_2-\alpha_1$.

所以
$$\tan\theta=\tan(\alpha_2-\alpha_1)$$
$$=\frac{\tan\alpha_2-\tan\alpha_1}{1+\tan\alpha_1\tan\alpha_2}$$
$$=\frac{b_2-b_1}{1+b_1b_2}.$$
證訖.

第三節　二次式

一變數 X 之二次式爲 X 之二次多項式,並且其形式爲 $Y=a+bX+cX^2$; 彼處 a, b 及 c 爲固定實數. 此函數之最低極限值在統計學上是有用處的,因此我們提出下列之定理.

定理 6. 設有一個二次式

(8)
$$f(X)=a+bX+cX^2.$$

並且式中之 c 值爲正. 求證當 $X=-\frac{b}{2c}$ 時, $f(X)$ 之最小值爲
$$(4ac-b^2)/4c.$$

證: 因爲
$$a+bX+cX^2=c\left(X^2+\frac{b}{c}X\right)+a$$
$$=c\left(X+\frac{b}{2c}\right)^2+\frac{4ac-b^2}{4c}.$$

並且因爲 $\left(X+\frac{b}{2c}\right)^2$ 之值永遠爲正; 故當 $a+bX+cX^2$ 有最小值時, c 之值須爲正,且須 $\left(X+\frac{b}{2c}\right)^2=0$, 即須 $X=-\frac{b}{2c}$. 於是得知,當 c 時, $\frac{4ac-b^2}{4c}$ 爲 $f(X)$ 之最小值. 證訖.

若 $c>0$, 則 $Y=a+bX+cX^2$ 代表之拋物線, 而以 (X_0, Y_0) 爲頂點(見圖 25); 彼處 $X_0=-\frac{b}{2c}$, $Y_0=(4ac-b^2)/4c$.

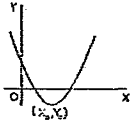


圖 25. 拋物線 $Y=a+bX+cX^2$

第四節 最小平方法

以曲線配合觀察論據之方法約有五種：一曰最小平方法 (method of least squares), 二曰動差法 (method of moment), 三曰分組法 (method of grouping), 四曰定點法 (fixed point method), 五曰幾近法 (method of approximation). 曲線配合之標準方法，人皆知之為最小平方法，茲僅先在本節敘述之；至於其他各種配合法則在以下各節依次討論之. 所謂最小平方法者，係用觀察值與理論值 (由實驗式算得者) 差量平方和之大小來判定配合曲線優劣之方法. 明確言之，即將此差量平方和

$$R=\Sigma d^2$$

縮至最小當做配合最美滿之條件；彼處 d 代表觀察值與理論值之差，此差量 d 即通常所稱之剩餘數. 剩餘數有種種不同之取法：1. 取其與 Y 軸方向相同者，而取 Y 之剩餘數；2. 取其與 X 軸方向相同者，即取 X 之剩餘數；3. 取其與垂直線 OQ 方向相同者 (見圖 24)，即取 p 之剩餘數；4. 取 $\log Y$ 之剩餘數等等. 茲分別一一論及之.

1. 以 Y 之剩餘數平方和最小為準則　若觀察之 Y 值易生誤差之論據而 X 則否，宜依據此準則配合之. 茲將配合方法論述如下.

設 $Y=f(X)$ 爲欲求之曲線方程式，並且

使 $$f(X)=a_1f_1+a_2f_2+\cdots\cdots+a_nf_n,$$

此處 $a_1, a_2, \cdots\cdots, a_n$ 爲常數，$f_1, f_2, \cdots\cdots, f_n$ 爲 X 之函數.

由式(9)可見，$f(X)$ 爲 $a_1, a_2, \cdots\cdots, a_n$ 之一次式. 凡算式能寫成式(9)形式者，得以 Σd_Y^2 之值縮至最小當做曲線最美滿配合之準則；此處 $d_Y=Y-(a_1f_1+a_2f_2+\cdots\cdots+a_nf_n)$. 茲證明之如下:

證：因爲 $d_Y=Y-(a_1f_1+a_2f_2+\cdots\cdots+a_nf_n)$,

於是 $$R_Y=\Sigma[Y-(a_1f_1+a_2f_2+\cdots\cdots+a_nf_n)]^2,$$

此處 $R_Y=\Sigma d_Y^2$.

求 R_Y 對於 $a_1, a_2, \cdots\cdots, a_n$ 之第一次偏微係數 (partial derivative)，並且分別使之各等於零 (此爲將 R_Y 之值縮至最小之必要條件)，我們有

$$(10)\quad \begin{cases} a_1\Sigma f_1f_1+a_2\Sigma f_1f_2+\cdots\cdots+a_n\Sigma f_1f_n=\Sigma Yf_1, \\ a_1\Sigma f_1f_2+a_2\Sigma f_2f_2+\cdots\cdots+a_n\Sigma f_2f_n=\Sigma Yf_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_1\Sigma f_1f_n+a_2\Sigma f_2f_n+\cdots\cdots+a_n\Sigma f_nf_n=\Sigma Yf_n. \end{cases}$$

此式通常被稱爲式(9)之略範方程式 (normal equations). 式(10)之數目共有 n 個，適足以求出 n 個常數 $a_1, a_2, \cdots\cdots, a_n$ 之值. 解之有

$$\begin{cases} a_1=\dfrac{1}{D}\begin{vmatrix} \Sigma Yf_1 & \Sigma f_2f_1\cdots\cdots f_nf_1 \\ \Sigma Yf_2 & \Sigma f_2f_2\cdots\cdots f_nf_2 \\ \cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ \Sigma Yf_n & \Sigma f_2f_n\cdots\cdots f_nf_n \end{vmatrix} \end{cases}$$

(11)
$$\left\{\begin{aligned} a_2 &= \frac{1}{D}\begin{vmatrix} \Sigma f_1f_1 & \Sigma fYf_1 & \cdots\cdots & \Sigma f_nf_1 \\ \Sigma f_1f_2 & \Sigma fYf_2 & \cdots\cdots & \Sigma f_nf_2 \\ \cdots & \cdots & \cdots & \cdots \\ \Sigma f_1f_n & \Sigma Yf_n & \cdots\cdots & \Sigma f_nf_n \end{vmatrix} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ a_n &= \frac{1}{D}\begin{vmatrix} \Sigma f_1f_1 & \Sigma f_2f_1 & \cdots\cdots & \Sigma Yf_1 \\ \Sigma f_1f_2 & \Sigma f_2f_2 & \cdots\cdots & \Sigma Yf_2 \\ \cdots & \cdots & \cdots & \cdots \\ \Sigma f_1f_n & \Sigma f_2f_n & \cdots\cdots & \Sigma Yf_n \end{vmatrix} \end{aligned}\right.$$

彼處
$$D=\begin{vmatrix} \Sigma f_1f_1 & \Sigma f_2f_1 & \cdots\cdots & \Sigma f_nf_1 \\ \Sigma f_1f_2 & \Sigma f_2f_2 & \cdots\cdots & \Sigma f_nf_2 \\ \cdots & \cdots & \cdots & \cdots \\ \Sigma f_1f_n & \Sigma f_2f_n & \cdots\cdots & \Sigma f_nf_n \end{vmatrix}.$$

a. 直線　因爲一般直線方程式可寫成式(2)之形式，故其爲式(9)之特殊情形．若定式(9)中之 $a_1=a$, $f_1=1$, $a_2=b$, $f_2=X$, 其他各 a 值均等於零；於是式(9)變成了式(2)，並且由式(10)得到式(2)之疇範方程式

(12)
$$\left\{\begin{aligned} Na+b\Sigma X &= \Sigma Y, \\ a\Sigma X+b\Sigma X^2 &= \Sigma XY. \end{aligned}\right.$$

此等疇範方程式可用前面定理6求得之．因爲

$$\begin{aligned} R_Y &= \Sigma[Y-(a+bX)]^2 \\ &= \Sigma(Y^2-2aY-2bXY+a^2+2abX+b^2X^2), \end{aligned}$$

可見 R_Y 爲 a 又爲 b 之二次式，因之由定理6得知 R_Y 之值當

$$Na+b\Sigma X=\Sigma Y$$

及 $$a\Sigma X+b\Sigma X^2=\Sigma XY$$

時為最小(學者試驗證之).這兩個方程式與式(12)完全相同.以式(12)中之 a 及 b 當做未知數解之,得

$$(13)\quad \begin{cases} a=\dfrac{1}{D}\begin{vmatrix}\Sigma Y & \Sigma X\\ \Sigma XY & \Sigma X^2\end{vmatrix} \\[2ex] b=\dfrac{1}{D}\begin{vmatrix}N & \Sigma Y\\ \Sigma X & \Sigma XY\end{vmatrix},\end{cases}$$

彼處 $$D=\begin{vmatrix}N & \Sigma X\\ \Sigma X & \Sigma X^2\end{vmatrix}.$$

若將算得之 a 及 b 值代入原方程式 $Y=a+bX$ 內,即得所欲求之方程式.

例 1. 試以 Y 之剩餘數平方和最小為準則,並以直線 $Y=a+bX$ 配合下表裏所列之論據:

表 32. 配合直線(2)之計算手續

(論據見參攷書 25)

年度	紙煙枝數以百億計 Y	助變數 x	x^2	XY
1923	66.7	0	0	0.0
1924	72.7	1	1	72.7
1925	82.3	2	4	164.6
1926	92.1	3	9	276.3
1927	93.0	4	16	372.0
1928	100.6	5	25	503.0
總計	507.4	15	55	1388.6

此處 $N=6$, $\Sigma x=15$, $\Sigma Y=507.4$, $\Sigma x^2=55$, $\Sigma xY=1388.6$. 於是

$$D=\begin{vmatrix} 6 & 15 \\ 15 & 55 \end{vmatrix}=105;$$

因之

$$a=\frac{1}{D}\begin{vmatrix} 507.4 & 15 \\ 1388.6 & 55 \end{vmatrix}=67.41,$$

$$b=\frac{1}{D}\begin{vmatrix} 6 & 507.4 \\ 15 & 1388.6 \end{vmatrix}=6.863;$$

並且所求之直線方程式爲

$$Y=67.41+6.863x. \tag{14}$$

如將 x 之值 0 及 5 分別代入式(14)內，則算得 Y 之值各爲 67.4 及 101.7，即得到兩點 (0,67.4) 及 (5,101.7). 有此兩點可以繪出式(14)代表之直線. 觀察值與配合之直線一併繪入圖 26 中. 由此圖可見，原始論據以直線配合之相當滿意.

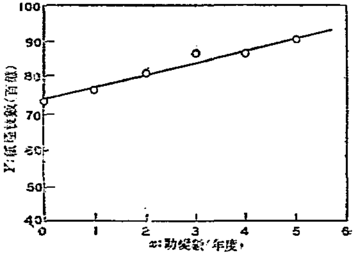


圖 26. 紙煙論據及式(14)之軌跡。

此例係用助變數 x 代替原始變數 X；所以如此者，爲運算上之便利計耳。以下各例有時也用助變數進行運算，其目的亦在此。用 x 代替 X，僅在變數 X 之組距相等時較爲方便。如欲以 X 表示 Y 亦可，其法系先將 X 與 x 之關係求出，將此關係代入求得之方程式內卽得。在本例中 X 與 x 之關係爲 $x=X-1923$，將其代入式(14)內得

$$Y=67.41+6.863(X-1923)$$

$$=-13130.14+6.863X. \tag{15}$$

自然由原始論據求得 ΣX, ΣY, ΣX^2, ΣXY 値後，代入下列之疇範方程式

$$Na+b\Sigma X=\Sigma Y,$$

$$a\Sigma X+b\Sigma X^2=\Sigma XY$$

內解之，亦可得與式(15)相同之結果。

b. 二次拋物線　因二次拋物線可寫成式(8)之形式，故其爲式(9)之特殊情形；卽是使式(9)中之 $a_1=a$, $a_2=b$, $a_3=c$, $f_1=1$, $f_2=X$, $f_3=X^2$，其他各 a 之値爲零，可得式(8)。將上面各數值代入式(10)內，可得疇範方程式

$$Na+b\Sigma X+c\Sigma X^2=\Sigma Y,$$

$$a\Sigma X+b\Sigma X^2+c\Sigma X^3=\Sigma XY,$$

$$a\Sigma X^2+b\Sigma X^3+c\Sigma X^4=\Sigma X^2Y.$$

以 a, b 及 c 當做未知數解之，或由式(11)得

$$
(16)\qquad
\begin{cases}
a=\dfrac{1}{D}\begin{vmatrix} \Sigma Y & \Sigma X & \Sigma X^2 \\ \Sigma XY & \Sigma X^2 & \Sigma X^3 \\ \Sigma X^2Y & \Sigma X^3 & \Sigma X^4 \end{vmatrix}, \\[2ex]
b=\dfrac{1}{D}\begin{vmatrix} N & \Sigma Y & \Sigma X^2 \\ \Sigma X & \Sigma XY & \Sigma Y^3 \\ \Sigma X^2 & \Sigma X^2Y & \Sigma X^4 \end{vmatrix}, \\[2ex]
c=\dfrac{1}{D}\begin{vmatrix} N & \Sigma X & \Sigma Y \\ \Sigma X & \Sigma X^2 & \Sigma XY \\ \Sigma X^2 & \Sigma X^3 & \Sigma X^2Y \end{vmatrix},
\end{cases}
$$

彼處

$$
D=\begin{vmatrix} N & \Sigma X & \Sigma X^2 \\ \Sigma X & \Sigma X^2 & \Sigma X^3 \\ \Sigma X^2 & \Sigma X^3 & \Sigma X^4 \end{vmatrix}.
$$

如果以居中二 X 值之平均數爲原點(當觀察值之個數爲偶數時),或以居中之X值爲原點(當觀察值之個數爲奇數時),則式 (13) 變做

$$
(17)\qquad
\begin{cases}
a=\dfrac{1}{D}\begin{vmatrix} \Sigma Y & 0 \\ \Sigma xY & \Sigma x^2 \end{vmatrix}, \\[2ex]
b=\dfrac{1}{D}\begin{vmatrix} N & \Sigma Y \\ 0 & \Sigma xY \end{vmatrix},
\end{cases}
$$

彼處

$$
D=\begin{vmatrix} N & 0 \\ 0 & \Sigma x^2 \end{vmatrix};
$$

並且式 (16) 變做

$$
(18)\quad
\begin{cases}
a=\dfrac{1}{D}\begin{vmatrix}\Sigma Y & 0 & \Sigma x^2\\ \Sigma xY & \Sigma x^2 & 0\\ \Sigma x^2Y & 0 & \Sigma x^4\end{vmatrix}\\[2ex]
b=\dfrac{1}{D}\begin{vmatrix}N & \Sigma Y & \Sigma x^2\\ 0 & \Sigma xY & 0\\ \Sigma x^2 & \Sigma x^2Y & \Sigma x^4\end{vmatrix}\\[2ex]
c=\dfrac{1}{D}\begin{vmatrix}N & 0 & \Sigma Y\\ 0 & \Sigma x^2 & \Sigma xY\\ \Sigma x^2 & 0 & \Sigma x^2Y\end{vmatrix}
\end{cases}
$$

設處

$$
D=\begin{vmatrix}N & 0 & \Sigma x^2\\ 0 & \Sigma x^2 & 0\\ \Sigma x^2 & 0 & \Sigma x^4\end{vmatrix}
$$

可見用式(17)計算 a 及 b 之手續比用式(13)省事些，用式(18)計算 a, b 及 c 之手續比用式(16)省事些。

例 2. 論據與例 1 同；但運算時以居中二 X 之平均數為原點而不以居首之 X 為原點。

表 33. 配合直線([illegible])之計算手續

年度 X	紙煙枝數以百億計 Y	助變數 x	x^2	xY
1923	66.7	−5	25	−333.5
1924	72.7	−3	9	−218.1
1925	82.3	−1	1	−82.3
1926	92.1	1	1	92.1
1927	93.0	3	9	279.0
1928	100.6	5	25	503.0
總計	507.4	0	70	240.2

將上表裏算得之值代入式 (17) 內得

$$D=\begin{vmatrix} 6 & 0 \\ 0 & 70 \end{vmatrix}=420.$$

因之
$$a=\frac{1}{D}\begin{vmatrix} 507.4 & 0 \\ 240.2 & 72 \end{vmatrix}=84.5667,$$

$$b=\frac{1}{D}\begin{vmatrix} 6 & 507.4 \\ 0 & 240.2 \end{vmatrix}=3.4314$$

並且所求之方程式為

$$Y=84.5667+3.4314\,x.$$

若將上式中之 x 以 $\dfrac{X-1925.5}{0.5}$ 代之,則得與式 (15) 相同之結果;彼處 1925.5 為 1925 與 1926 之平均數.

例 3. 試以拋物線配合下表裏所列之論據:

表 34. 配合拋物線之計算手續

(論據見參考書 25)

年度 X	每1000雙配偶離婚人數 Y	助變數 x	x^2	x^4	xY	x^2Y
1900	81	−3	9	81	−243	729
1905	84	−2	4	16	−168	336
1910	88	−1	1	1	− 88	88
1915	104	0	0	0	0	0
1920	134	1	1	1	134	134
1925	148	2	4	16	296	592
1930	170	3	9	81	510	1530
總數	809	0	28	196	441	3409

此處 $N=7$, $\Sigma x=0$, $\Sigma x^2=28$, $\Sigma x^3=0$, $\Sigma x^4=196$, $\Sigma Y=809$, $\Sigma xY=441$, 並且 $\Sigma x^2Y=3409$.

將此等值代入式 (18) 內得

$$a=107.33,\quad b=15.7500,\quad c=2.0595.$$

故所求之方程式為

$$Y=107.33+15.7500x+2.0595x^2, \tag{19}$$

或以 X 表之,因 $X=5x+1915$, 為

$$Y=296181.08-312.3654X+0.082380X^2.$$

設 $x=-3,\ -2,\ -1,\ 0,\ 1,\ 2$, 及 3, 並將其分別代入式 (19) 內;則可算得 Y 值各為 78.6, 84.7, 93.6, 107.3, 125.1, 147.1 及 173.1.

茲將配合之拋物線及原始論據繪入圖 27 中. 圖 27 顯示用拋物線配合原始論據尚稱滿意.

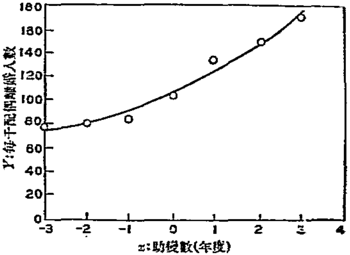


圖 27. 離婚論據及式(19)之軌跡.

c. 對數曲線　如果對數曲線之方程式為

(20) $$Y=a+bX+c\log X,$$

亦得用現在方法配合之，蓋使式(9)中之 $a_1=a$, $a_2=b$, $a_3=c$, $f_1=1$, $f_2=X$, $f_3=\log X$, 其他各 a 值均等於零，可得式(20). 因之由式(10)可直接寫出以式(20)配合論據之時範方程式

(21) $$\begin{cases} Na+b\Sigma X+c\Sigma\log X=\Sigma Y, \\ a\Sigma X+b\Sigma X^2+c\Sigma X\log X=\Sigma XY, \\ a\Sigma\log X+b\Sigma X\log X+c\Sigma(\log X)^2=\Sigma Y\log X. \end{cases}$$

以式(21)中之 a, b 及 c 當做未知數解之，得

(22) $$\begin{cases} a=\dfrac{1}{D}\begin{vmatrix} \Sigma Y & \Sigma X & \Sigma\log X \\ \Sigma XY & \Sigma X^2 & \Sigma X\log X \\ \Sigma Y\log X & \Sigma X\log X & \Sigma(\log X)^2 \end{vmatrix}, \\ b=\dfrac{1}{D}\begin{vmatrix} N & \Sigma Y & \Sigma\log X \\ \Sigma X & \Sigma XY & \Sigma X\log X \\ \Sigma\log X & \Sigma Y\log X & \Sigma(\log X)^2 \end{vmatrix}, \\ c=\dfrac{1}{D}\begin{vmatrix} N & \Sigma X & \Sigma Y \\ \Sigma X & \Sigma X^2 & \Sigma XY \\ \Sigma\log X & \Sigma X\log X & \Sigma Y\log X \end{vmatrix}; \end{cases}$$

彼處 $$D=\begin{vmatrix} N & \Sigma X & \Sigma\log X \\ \Sigma X & \Sigma X^2 & \Sigma X\log X \\ \Sigma\log X & \Sigma X\log X & \Sigma(\log X)^2 \end{vmatrix}.$$

例 4. 以對數曲線 $Y=a+bX+c\log X$ 配合表35裏所列之論據.

由表35得知 $N=20$, $\Sigma x=210$, $\Sigma Y=2673.7$, $\Sigma\log x=18.3861247$, 並且由表35直接得以算出 $\Sigma x^2=2870$, $\Sigma x\log x=230.0033043$, $\Sigma(\log x)^2=19.2694684$, $\Sigma xY=31640.5$, $\Sigma Y\log x=2677.6433$.

將此等值代入式(22)中,我們有

$$a=54.347,\quad b=1.555,\quad c=68.549.$$

是以欲求之方程式為

(23) $$Y=54.347+1.555\,x+68.549\log x.$$

表 35. 曲線(23)縱坐標之計算

(論據見參考書 29)

雞雛重量(公分) X	雞雛平均毛高(公釐) Y	助變數 x	log x	縱坐標 Y
0—19	58.8	1	.0000000	55.9
20—39	76.4	2	.3010300	78.1
40—59	91.1	3	.4771213	91.7
60—79	99.0	4	.6020600	101.8
80—99	108.1	5	.6989700	110.0
100—119	115.1	6	.7781513	117.0
120—139	122.7	7	.8450980	123.2
140—159	129.5	8	.9030900	128.7
160—179	135.0	9	.9542425	133.7
180—199	141.1	10	1.0000000	138.4
200—219	144.0	11	1.0413927	142.8
220—239	150.0	12	1.0791813	147.0
240—259	152.8	13	1.1139433	150.9
260—279	155.6	14	1.1461281	154.7
280—299	158.6	15	1.1760912	158.3
300—319	161.3	16	1.2041200	161.8
320—339	160.5	17	1.2304489	165.1
340—359	171.0	18	1.2552725	168.4
360—379	169.5	19	1.2787536	171.5
380—399	173.6	20	1.3010300	174.6
總　計	2,673.7	210	18.3861247	2673.6

茲將表 35 裏所列之原始觀察值及算得之對應值一併輸入圖 28

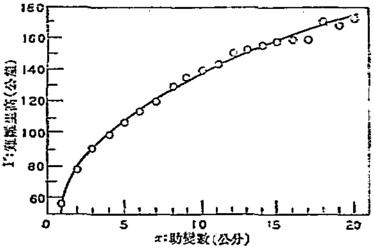


圖 28. 雞雛平均坐高論據及式(23)之軌跡。

中。由圖 28 可見用曲線式(20)配合表 35 中所列之論據尚稱滿意。

2. 以X之剩餘數平方和最小爲準則　若觀察之 X 值易生誤差之論據，而 Y 值則否，宜以此爲準則配合之。配合方法如下：

設 $X=F(Y)$ 爲欲求之曲線方程式，並且

$$(24)\qquad F(Y)=b_1F_1+b_2F_2+\cdots\cdots+b_nF_n,$$

其爲常數 $b_1, b_2, \cdots\cdots, b_n$ 之一次式；彼處 $F_1, F_2, \cdots\cdots, F_n$ 各爲 Y 之函數。於是得以 Σd_X^2 之值縮至最小當做配合最美滿之條件。

證：因爲 $d_X=X-(b_1F_1+b_2F_2+\cdots\cdots+b_nF_n)$,

故 $$R_X=\Sigma[X-(b_1F_1+b_2F_2+\cdots\cdots+b_nF_n)]^2,$$

彼處 $$R_X=\Sigma d_X^2.$$

和先前一樣，求 R_X 對於 $b_1, b_2, \cdots\cdots, b_n$ 之第一次偏微係數，並且使各微係數等於零（此爲將 R_X 值縮至最小之必要條件）；則得標範方程式

$$(25)\quad \begin{cases} b_1\Sigma F_1F_1+b_2\Sigma F_1F_2+\cdots\cdots+b_n\Sigma F_1F_n=\Sigma XF_1 \\ b_1\Sigma F_1F_2+b_2\Sigma F_2F_2+\cdots\cdots+b_n\Sigma F_2F_n=\Sigma XF_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b\,\Sigma F_1F_n+b_2\Sigma F_2F_n+\cdots\cdots+b_n\Sigma F_nF_n=\Sigma XF_n \end{cases}$$

該方程式之個數與待定常數之個數相等,適足以求出各常數之值.即

$$(26)\quad \begin{cases} b_1=\dfrac{1}{D}\begin{vmatrix} \Sigma XF_1 & \Sigma F_2F_1\cdots\cdots\Sigma F_nF_1 \\ \Sigma XF_2 & \Sigma F_2F_2\cdots\cdots\Sigma F_nF_2 \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \Sigma XF_n & \Sigma F_2F_n\cdots\cdots\Sigma F_nF_n \end{vmatrix} \\ b_2=\dfrac{1}{D}\begin{vmatrix} \Sigma F_1F_1 & \Sigma XF_1\cdots\cdots\Sigma F_nF_1 \\ \Sigma F_1F_2 & \Sigma XF_2\cdots\cdots\Sigma F_nF_2 \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \Sigma F_1F_n & \Sigma XF_n\cdots\cdots\Sigma F_nF_n \end{vmatrix} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_3=\dfrac{1}{D}\begin{vmatrix} \Sigma F_1F_1 & \Sigma F_2F_1\cdots\cdots\Sigma XF_1 \\ \Sigma F_1F_2 & \Sigma F_2F_2\cdots\cdots\Sigma XF_2 \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \Sigma F_1F_n & \Sigma F_2F_n\cdots\cdots\Sigma XF_n \end{vmatrix} \end{cases}$$

彼處
$$D=\begin{vmatrix} \Sigma F_1F_1 & \Sigma F_2F_1\cdots\cdots\Sigma F_nF_1 \\ \Sigma F_1F_2 & \Sigma F_2F_2\cdots\cdots\Sigma F_nF_2 \\ \cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\ \Sigma F_1F_n & \Sigma F_2F_n\cdots\cdots\Sigma F_nF_n \end{vmatrix}$$

因直線方程式可寫成如式(3)形式,故其爲式(24)之特殊情形,祇

要使 $b_1=a'$, $b_2=b'$, $F_1=1$, $F_2=Y$, 其他各 b 均爲零，式 (24) 即可變做式 (3). 於是由式 (25) 得到式 (3) 之疇範方程式

$$Na'+b'\Sigma Y=\Sigma X,$$

$$a'\Sigma Y+b'\Sigma Y^2=\Sigma XY.$$

解之或由式 (26), 得

$$(27)\qquad \begin{cases} a'=\dfrac{1}{D}\begin{vmatrix} \Sigma X & \Sigma Y \\ \Sigma XY & \Sigma Y^2 \end{vmatrix}, \\ b'=\dfrac{1}{D}\begin{vmatrix} N & \Sigma Y \\ \Sigma Y & \Sigma XY \end{vmatrix}; \end{cases}$$

彼處
$$D=\begin{vmatrix} N & \Sigma Y \\ \Sigma Y & \Sigma Y^2 \end{vmatrix}$$

將算得之 a' 及 b' 值代入式 (3), 即得所求之直線方程式.

例 5. 以直線 $X=a'+b'Y$ 配合例 1 裏所列之論據.

配合直線 $X=a'+b'Y$ 所需之數值 ΣY, ΣX 及 ΣXY 已在表 32 裏算得各爲 507.4, 15 及 1388.6; 至 ΣY^2 可由該表直接算得，其爲 43759.24. 將此等值代入式 (27) 中，則得 $D=5100.6800$, 因之

$$a'=-9.4359,\quad b'=0.1413.$$

於是所求之直線方程式爲

$$X=-9.4359+0.1413Y.$$

或者寫做

$$(28)\qquad Y=66.78+7.0771X.$$

3. 取 p 之剩餘數平方和最小爲準則　當觀察之 X 及 Y 值均易發

生誤差之論據，並且 (X, Y) 之趨勢爲一直線時，宜依據此準則配合之。以垂直剩餘數平方和最小原則配合直線之法，可敘述之如下：

假設直線方程式爲

$$Y=a+bX.$$

於是由定理 4 系得知某點至此直線上之垂直距離

$$d_p=\frac{a+bX-Y}{\pm\sqrt{1+b^2}}.$$

故

$$R_p=\frac{1}{b^2+1}\Sigma(a+bX-Y)^2,$$

彼處

$$R_p=\Sigma d_p^2.$$

如將 R_p 縮至最小，則其必要條件爲 $\frac{\partial R_p}{\partial a}=0$ 及 $\frac{\partial R_p}{\partial b}=0$；即是

$$(29)\qquad \begin{cases} a+b\bar{X}=\bar{Y}, \\ b^2r\sigma_X\sigma_Y+b(\sigma_X^2-\sigma_Y^2)-r\sigma_X\sigma_Y=0; \end{cases}$$

彼處

$$r=\frac{\nu_{XY}-\nu_{1X}\nu_{1Y}}{\sigma_X\sigma_Y}.$$

以式 (29) 第二式中之 b 當做未知數解之，則有

$$(30)\qquad b=\frac{-(\sigma_X^2-\sigma_Y^2)\pm[(\sigma_X^2-\sigma_Y^2)^2+4r^2\sigma_X^2\sigma_Y^2]^{\frac{1}{2}}}{2r\sigma_X\sigma_Y}.$$

由式 (29) 第一式得知直線 $Y=a+bX$ 經過 $(X, \bar{Y})$，於是

$$Y-\bar{Y}=b(X-X),$$

彼處 b 可用式 (30) 決定之。

設式 (30) 中之二值，一爲 b_1，一爲 b_2，並且

$$b_1=\frac{-(\sigma_X^2-\sigma_Y^2)+\sqrt{(\sigma_X^2-\sigma_Y^2)^2+4r^2\sigma_X^2\sigma_Y^2}}{2r\sigma_X\sigma_Y},$$

$$b_2=\frac{-(\sigma_X^2-\sigma_Y^2)-\sqrt{(\sigma_X^2-\sigma_Y^2)^2+4r^2\sigma_X^2\sigma_Y^2}}{2\ \sigma_X\sigma_Y}$$

因此，由本章定理 5，得知二直線

$$(31)\qquad \begin{cases} l_1:\ Y-\bar{Y}=b_1(X-\bar{X}) \\ l_2:\ Y-\bar{Y}=b_2(X-\bar{X}) \end{cases}$$

夾角 θ 之正切 (tangent)

$$\tan\theta=\frac{b_2-b_1}{1+b_1b_2}$$

$$=\frac{-\dfrac{\sqrt{(\sigma_X^2-\sigma_Y^2)^2+4r^2\sigma_X^2\sigma_Y^2}}{r\sigma_X\sigma_Y}}{1-\dfrac{r\sigma_X\sigma_Y}{r\sigma_X\sigma_Y}}$$

$$=\frac{-\sqrt{(\sigma_X^2-\sigma_Y^2)^2+4r^2\sigma_X^2\sigma_Y^2}}{r\sigma_X\sigma_Y-r\sigma_X\sigma_Y}$$

$$=\infty,\quad (\text{因分母爲零}).$$

這就是說二直線 l_1 與 l_2 互相垂直. 也就是說：若 l_1 配合論據爲最美滿者，則 l_2 爲最惡劣者；反之若 l_2 爲最美滿者，則 l_1 爲最惡劣者.

例 6. 試以垂直剩餘數平方和最小爲準則並以直線式 (31) 配合例 1 裏之論據.

由例 1 裏論據得知 $N=6$, $\Sigma x=15$, $\Sigma Y=507.4$, $\Sigma x^2=55$, $\Sigma Y^2=43759.24$, $\Sigma xY=1388.6$；於是

$$\nu_{1x}=2.500000,\qquad \nu_{1Y}=84.566667,$$

$$\nu_{2x}=9.766667,\qquad \nu_{2Y}=7293.2066667,$$

$$\nu_{xY}=231.433333.$$

$$\bar{x}=2.500000, \qquad \bar{Y}=84.566667,$$

$$\sigma_x=1.707825, \qquad \sigma_Y=11.903172,$$

$$\sigma_x\sigma_Y=\ 20.328533,$$

$$r=\ 0.984658,$$

$$2r\sigma_x\sigma_Y=\ 40.033305,$$

$$-(\sigma_x^2-\sigma_Y^2)=138.768837,$$

$$[(\sigma_x^2-\sigma_Y^2)^2+(2r\sigma_x\sigma_Y)^2]^{\frac{1}{2}}=144.428029.$$

因之
$$b_1=\frac{-(\sigma_x^2-\sigma_Y^2)+[(\sigma_x^2-\sigma_Y^2)^2+(2r\sigma_x\sigma_Y)^2]^{\frac{1}{2}}}{2\ \sigma_x\sigma_Y}=7.074031,$$

$$b_2=\frac{-(\sigma_x^2-\sigma_Y^2)+[(\sigma_x^2-\sigma_Y^2)^2+(2r\sigma_x\sigma_Y)^2]^{\frac{1}{2}}}{2r\sigma_x\sigma_Y}=-0.141362.$$

由此得到欲求之方程式

$$Y-84.566667=\ \ 7.074031(x-2.5),$$

$$Y-84.566667=-0.141362(x-2.5).$$

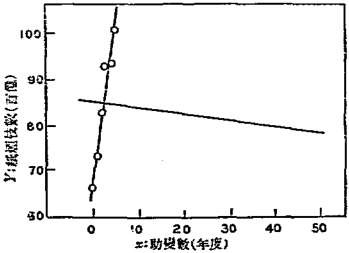


圖 29. 表32裏所列之數據及式(32)之軌跡。

即是

$$(32)\qquad \begin{cases} Y=66.881589+7.074031x, \\ Y=84.920072-0.141362x. \end{cases}$$

前者爲配合最完美之直線，後者爲最惡劣者，如圖 29 所示.

4. 取 $\log Y$ 之剩餘數平方和最小爲準則 設指數曲線方程式之形式如下:

$$(33)\qquad Y=ab^X,$$

如是取其兩邊對數我們有

$$\log Y=\log a+X\log b.$$

若以數值 $R_{\log Y}=\Sigma[\log Y-(\log a+X\log b)]^2$ 縮至最小當做最美滿配合之條件，彼處 $R_{\log Y}=\Sigma d^2_{\log Y}$. 則可得到配合指數形式之疇範方程式

$$Na+\log b\Sigma X=\Sigma\log Y,$$

$$\log a\Sigma X+\log b\Sigma X^2=\Sigma X\log Y.$$

解之，我們有

$$(34)\qquad \begin{cases} \log a=\dfrac{1}{D}\begin{vmatrix}\Sigma\log Y & \Sigma X\\ \Sigma X\log Y & \Sigma X^2\end{vmatrix}, \\ \log b=\dfrac{1}{D}\begin{vmatrix}N & \Sigma\log Y\\ \Sigma X & \Sigma X\log Y\end{vmatrix}. \end{cases}$$

即

$$a=\log^{-1}\left[\frac{1}{D}\begin{vmatrix}\Sigma\log Y & \Sigma X\\ \Sigma X\log Y & \Sigma X^2\end{vmatrix}\right],$$

$$b=\log^{-1}\left[\frac{1}{D}\begin{vmatrix}N & \Sigma\log Y\\ \Sigma X & \Sigma X\log Y\end{vmatrix}\right];$$

彼處
$$D=\begin{vmatrix} N & \Sigma X \\ \Sigma X & \Sigma X^2 \end{vmatrix}.$$

若將算得之 a 及 b 值代入式 (33) 內卽得欲求之方程式.

例 7. 試用式 (33) 配合下表裏所得之論據:

表 36. 配合指數曲線 $Y=ab^X$ 之計算手續

X	Y	X^2	$\log Y$	$X\log Y$
1	1.6	1	.2041	.2041
2	4.5	4	.6532	1.3064
3	13.8	9	1.1399	3.4197
4	40.2	16	1.6042	6.4168
5	125.0	25	2.0969	10.4845
15	...	55	5.6983	21.8315

此處 $N=5$, $\Sigma X=15$, $\Sigma X^2=55$, $\Sigma\log Y=5.6983$ 並且 $\Sigma X\log Y=21.8315$. 將此等值代入式 (34) 內,則得 $D=50$. 因之

$$\log a=\bar{1}.718680,$$

$$\log b=0.473660;$$

卽是
$$a=0.5233,$$

$$b=2.9762.$$

故所求之方程式為

$$Y=.5233(2.9762)^X. \tag{35}$$

將 $X=1,2,3,4$ 及 5 代入式 (35) 裏,則可算得各對應之 Y 值 1.56, 4.64, 13.81, 41.10 及 122.31. 原始論據及配合之曲線 (35) 一併繪入圖 30 中.

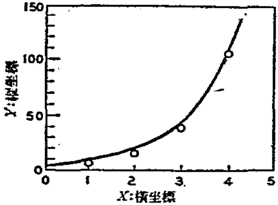
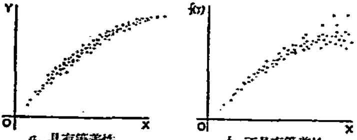


圖 30. 表36裏所列之論據及式(35)之軌跡。

此外還有取 $\frac{1}{Y}$ 剩餘數平方和最小當做配合最美滿之準則者，如配合曲線 $\frac{1}{Y}=a+bX$ 是；取 XY 剩餘數平方和最小當做準則者，如配合曲線 $(X-a)(Y-b)=c$ 是，如此等等。我們以 Y 之剩餘數平方和最小當做曲線配合之最美滿條件為常，尤其當論據 (X, Y) 之趨勢如圖 31 a 而 $[X, f(Y)]$ 之趨勢如圖 31 b 所示時，更不能不用曲線 (X, Y) 配合之：因前者在 X 的任一組段內各 Y 值之離勢常數相同，即各點之趨勢具有等差性 (homoscedasticity)；後者在 X 組段內各 $f(Y)$ 離勢常數之值隨 X 組段之變而變，即各點之趨勢不具有等差性 (heteroscedasticity)，故也。



a. 具有等差性　　b. 不具有等差性

圖 31. 各點之趨勢具有與不具有等差性者。

第五節 動差法

皮爾生配合其頻數曲線時所用之方法即係動差法，已見前章．此法仍可用以配合成雙變數顯示之趨勢曲線．如果有一函數 $X=f(X, a, b, c \cdots\cdots)$，則 $a, b, c, \cdots\cdots$ 可用此法測定之；其法係先求得函數之各級動差及縱坐標之各級動差，再使函數之動差與同級縱坐標之動差相等．如此得到許多方程式，然後由此方程式測定原函數中所含常數之值．因此通常稱此法為縱坐標動差法．

若用此配合直線，因直線方程式內含有兩個常數 a 及 b，須有兩個含有 a 及 b 之聯立方程式．如果使零級及一級觀察動差(即縱坐標動差)各等於零級及一級計算動差(即函數動差)，並且假設欲配合之直線方程式之形式為 $Y=a+bX$；於是我們有

$$\Sigma a+b\Sigma X=\Sigma Y,$$

$$a\Sigma X+b\Sigma X^2=\Sigma XY,$$

彼處 Σ 為由 1 至 N．此與用最小平方法所求得之略範方程式同．

如果配合之曲線為拋物線，

$$Y=a+bX+cX^2,$$

則其中常數之值可由下列方程式求得之：

$$Na+b\Sigma X+c\Sigma X^2=\Sigma Y,$$

$$a\Sigma X+b\Sigma X^2+c\Sigma X^3=\Sigma XY,$$

$$a\Sigma X^2+b\Sigma X^3+c\Sigma X^4=\Sigma X^2Y,$$

此亦與用最小平方法以拋物線 $Y=a+bX+cX^2$ 配合論據所用之略範方程式相同．

一般言之，凡以高次拋物線配合論據，測定常數所用之方程式，用動差法與用最小平方法得之者，可以完全相同（此問題留待學生研討與證明）。

第六節 分組法

已給之論據有時亦得用分組法配合之．其法係將觀察之 N 雙變量分成若干小組，組數之多寡與曲線方程式內所含常數之個數相等．設曲線方程式內含有三個常數，於是我們可將觀察變量分爲三小組，每組有 n 雙觀察變數．配合高培志曲線（Gompertz curve）往往應用此法．因此舉此曲線爲例，以便說明分組法之計算手續．

高培志曲線爲統計學上著明曲線之一．保險學家計算死亡率之高低，經濟學家預測商業之榮枯，人口學家估計人口之消長，以及教育家描寫教育上種種問題，均往往用之．此曲線係公元一八二五年高培志所發明，故名．其方程式爲

$$Y=KG^{A^X} \tag{36}$$

設 $X=X_0+hx$［見第二章式(4)］，並且 $x=0, 1, 2, \cdots\cdots, 3n-1$；於是以 X_0+hx 代替式中之 X，我們有

$$\begin{aligned} Y&=KG^{A^{X_0+hx}} \\ &=KG^{A^{X_0}}\times G^{A^{hx}} \end{aligned}$$

取兩邊對數得

$$y=k+ga^x, \tag{37}$$

彼處 $y=\log Y$，$k=\log KG^{A^{X_0}}$，$g=\log G$ 並且 $a=A^h$．式(37)爲高培志曲線方程式之另一寫法．

因爲 $x=0, 1, 2, \cdots\cdots, 3n-1$, 故得寫三組之 y 值如下:

$$\sum_{i=0}^{n-1} y_i = nk + g\sum_{i=0}^{n-1} a^i,$$

$$\sum_{i=n}^{2n-1} y_i = nk + g\sum_{i=n}^{2n-1} a^i,$$

$$\sum_{i=2n}^{3n-1} y_i = nk + g\sum_{i=2n}^{3n-1} a^i.$$

如果 s_1, s_2 及 s_3 各代表 $\sum_{i=0}^{n-1} y_i$, $\sum_{i=n}^{2n-1} y_i$ 及 $\sum_{i=2n}^{3n-1} y_i$ 則有

$$s_1 = nk + g(1 + a + a^2 + \cdots\cdots + a^{n-1}),$$

$$s_2 = nk + a^n g(1 + a + a^2 + \cdots\cdots + a^{n-1}),$$

$$s_3 = nk + a^{2n} g(1 + a + a^2 + \cdots\cdots + a^{n-1}).$$

兩兩相減,我們有

$$s_3 - s_2 = g a^n (a^n - 1)(1 + a + a^2 + \cdots\cdots + a^{n-1}),$$

$$s_2 - s_1 = g(a^n - 1)(1 + a + a^2 + \cdots\cdots + a^{n-1}).$$

因之
$$a^n = \frac{s_3 - s_2}{s_2 - s_1},$$

即
$$a = \sqrt[n]{\frac{s_3 - s_2}{s_2 - s_1}}.$$

但 $1 + a + a^2 + \cdots\cdots + a^{n-1} = \dfrac{a^n - 1}{a - 1}$, 故算式 $s_2 - s_1$ 可寫做

$$s_2 - s_1 = \frac{g(a^n - 1)^2}{a - 1}.$$

於是
$$g = \frac{(s_2 - s_1)(a - 1)}{(a^n - 1)^2}.$$

又算式 s_1 可寫做　　$s_1 = nk + \frac{g(a^n - 1)}{a - 1}$,

故　　$$k = \frac{1}{n}\left[s_1 - \frac{g(a^n - 1)}{a - 1}\right].$$

上面算式 s_1, s_2, s_3 表示函數 y 之和。如果將函數 y 易之以觀察 n,於是由算式 a^n 可求得 a 之值;有了 a 值,依次應用算式 g 及 k 可求得 g 及 k 之值;因之 K 及 G 之值可以求得之.

如果 $a<1$,於是,當 $X \to \infty$ 時,由式(37)得知 $y \to k$,並且由式(36)得知 $Y=K$ 爲高培志曲線之漸近綫.此爲唯一之漸近綫. K 值叫做高培志曲線之極限值.

例. 試以高培志曲線配合表 37 裏所列之論據.

由下表　　$s_1 = 26.374290$,

$s_2 = 30.031076$,

$s_3 = 32.126214$.

於是算得

$a = 0.911351$,

$k = 5.822848$,

$g = -1.777511$.

故所求曲線之方程式爲

(38)　　$\log Y = 5.822848 - 1.777\ 1\ (0.911351)^x$,

即是

(39)　　$Y = 10^{5.822848}10^{-1.777511(0.911351)^x}$.

用此法配合僅能求得各常數之近似值 (approximate value).

表37. 用分組法配合高培志曲線之計算手續（論據見參考28）

年度 X	人造絲產量（千磅）Y	x	$\log Y$
1920	10125	0	4.0053950
1921	14986	1	4.1756857
1922	24067	2	4.3814220
1923	34959	3	4.5435590
1924	36328	4	4.5602415
1925	51049	5	4.7079872
1926	62693	6	4.7972191
1927	75555	7	4.8782632
1928	97232	8	4.9878092
1929	121399	9	5.0842151
1930	127333	10	5.1049409
1931	150879	11	5.1786288
1932	134670	12	5.1292709
1933	213498	13	5.3293938
1934	208321	14	5.3187331
1935	257557	15	5.4108734
1936	277626	16	5.4434601
1937	312236	17	5.4944829

原始論據及配合之高培志曲線顯示於圖32中。

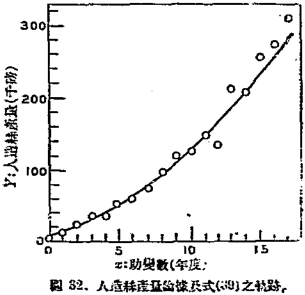


圖32. 人造絲產量論據及式(39)之軌跡。

將表 37 裏所列之 x 值分別代入式 (38) 內，則可算得 $\log Y$ 之值；因之可算得 Y 之值．茲將計算手續詳列於下：

表 38.　高培志曲線縱坐標之計算

年度 X	x	a^x	$k+ga^x$	縱坐標 Y
1920	0	1.000000	4.045[illegible]7	11100
1921	1	0.911351	4.202912	15956
1922	2	0.830561	4.346517	22208
1923	3	0.756933	4.477391	30019
1924	4	0.689832	4.596664	39506
1925	5	0.628679	4.705364	50742
1926	6	0.572948	4.804427	63742
1927	7	0.522157	4.894708	78471
1928	8	0.475868	4.976987	94839
1929	9	0.433683	5.051972	112712
1930	10	0.395237	5.120310	131920
1931	11	0.360200	5.182589	152261
1932	12	0.328269	5.239346	173519
1933	13	0.299168	5.291074	195467
1934	14	0.272647	5.338215	217879
1935	15	0.248477	5.381177	240534
1936	16	0.226450	5.420331	263227
1937	17	0.206375	5.456014	285768

第七節　定點法

定點法又可稱做圖解法．此法係由已給之觀察值 (X, Y) 中選取適宜者代入欲配合之方程式內，以測定方程式內所含各常數之值．因此稱之曰定點法．觀察值 (X, Y) 波動往往很大，選不出適宜點以為測定常數之依據．在此情形中，我們可先隨手任給一曲線，儘量切合各觀察值之趨勢，然後由所給之曲線上選取適宜之點並讀出其縱橫坐標，用以

代替觀察值，再依上法進行運算。曾有人稱此法曰圖解法者即以此故

定點法爲統計學上測定常數之值時最常用之法，因爲任何曲線方程式率能用此法測定其中所含各常數之值，非若上述之最小平方法所能配合之實驗式限於某種特殊形式也。不過用此法所配合之曲線常因個人主觀見解之不同而異其形式與位置，不免貽人以抨擊之隙。雖然吾人祇需得到近似結果時，此法尚覺稱意。

茲舉直線、指數曲線及蒲芮曲線 (Pearl-Reed curve) 爲例，以明配合時所需之運算手續。至於配合其他形式曲線之步驟，學者可以隅反。

1. 直線　用定點法配合直線時之運算手續至爲簡單。其法係先繪一直線（最好用一透明尺繪），使其儘量切合所繪之點；如此在所繪之直線上取得兩點，以估計其坐標 (X_1, Y_1) 及 (X_2, Y_2)；然後將此坐標代入直線方程式內，以其內所含之常數當做未知數解之，因可得到各常數之值；最終可得到欲求之方程式。

設式 (2) 爲所欲求之直線方程式。於是將 (X_1, Y_1) 及 (X_2, Y_2) 代入式內有

$$Y_1 = a + bX_1. \qquad Y_2 = a + bX_2.$$

解之得

$$(40) \qquad \begin{cases} a = \dfrac{1}{D}\begin{vmatrix} Y_1 & X_1 \\ Y_2 & X_2 \end{vmatrix}, \\[2ex] b = \dfrac{1}{D}\begin{vmatrix} 1 & Y_1 \\ 1 & Y_2 \end{vmatrix}, \end{cases}$$

彼處
$$D=\begin{vmatrix}1 & X_1\\1 & X_2\end{vmatrix}.$$

例 1. 試用定點法,以直線 $Y=a+bX$, 配合第四節例 1 裏所列之論據.

取 $(X_1, Y_1)=(0, 67)$; $(X_2, Y_2)=(5, 100)$. 於是由式(40)得
$$a=67,\ b=6.6,$$
及所所求之方程式
$$Y=67+6.6X.$$
此與第四節例 1 裏所得之結果無大差別.

2. 指數曲線 指數曲線爲不能用普通最小平方法(以 Y 而不以 $\log Y$ 之剩餘數平方和最小爲準則)配合之曲線;若用定點法配合之,則所用之手續甚爲簡易.

設式(33)爲欲求之指數曲線方程式,(X_1, Y_1), (X_2, Y_2) 爲所取之二點. 如是
$$Y_1=ab^{X_1},$$
$$Y_2=ab^{X_2}.$$
取其兩邊之對數,並分別乘以 X_2, X_1 後,得
$$X_2\log Y_1=X_2\log a+X_1X_2\log b,$$
$$X_1\log Y_2=X_1\log a+X_1X_2\log b.$$
以 a 及 b 當做未知數解之,得
$$(41)\qquad \begin{cases}a=\log^{-1}\left[\dfrac{1}{X_2-X_1}(X_2\log Y_1-X_1\log Y_2)\right]\\[2ex] b=\log^{-1}\left[\dfrac{1}{X_2-X_1}(\log Y_2-\log Y_1)\right]\end{cases}$$

例 2. 試用定點法,以曲線 $Y=ab^X$ 配合第四節例 7 要所列之論據.

取 $(X_1, Y_1)=(1, 16)$; $(X_2, Y_2)=(5, 125)$. 於是代入式 (41) 中得

$$a=0.538,$$

$$b=2.973.$$

故所求之方程式為　　$Y=0.538(2.973)^X$.

此與第四節例 7 所得之結果,不甚懸殊.

3. 蒲芮曲線.

a. 淵源　蒲芮曲線亦為著曲線之一，對於人口成長及經濟上之論據,極能盡其描寫之能事. 本係比利時算學家郛好斯德 (P. F. Verhulst) 於公元一八三八年所發明,當時稱做天算曲線 (logistic curve); 但此偉大之貢獻久已被人遺忘. 迨至公元一九二〇年美國生物及統計學家蒲爾 (Raymond Pearl) 及芮德 (L. J. Reed) 二氏，當時並未覩及郛氏之著作,獨自推得此曲線之方程式,用以描寫或估計世界各國人口,頗稱滿意. 故美國學者咸名此曲線曰蒲芮曲線.

b. 方程式　蒲芮曲線方程式之形式為

$$Y-d=\frac{K}{1+Ce^{rX}} \tag{42}$$

設處 Y 代表人口，X 代表時間，K, C, d 及 r 代表待定之常數. 但為後來討論便利計,假設 K, C 及 d 之值均為正,並 r 之值為負.

c. 性質　當 $X=-\infty$ 時，式 (42) 中之 $Y=d$; 當 $X=+\infty$ 時，$Y=K+d$. 於是知此曲線有兩條漸近線 $Y=d$ 及 $Y=K+d$. 前者為最低漸近線,後者為最高漸近線.

將式 (42) 中之 Y 對於 X 連續微分之,則得二次微係數

$$\frac{d^2Y}{dX^2} = -\frac{b}{K}(K+2d-2Y)\frac{dY}{dX}$$

使之等於零,我們有 $Y=\frac{K}{2}+d$

就是說蒲内曲線有一反曲點(見圖33),其坐標為

$$\left(-\frac{1}{r}\log_e C, \frac{K}{2}+d\right).$$

由式(42)得

(43) $$\frac{m-Y}{Y-d}=Ce^{rX},$$

彼處 m 代表 $K+d$. 取式(43)兩邊對數有

(44) $$\log\frac{m-Y}{Y-d}=\log C+rX\log e.$$

故知 $\log\frac{m-Y}{Y-d}$ 與 X 為一直線式之關係.

d. 圖形 根據上面之討論 得將式(42)代表之圖形畫出如下:

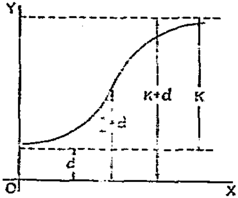


圖 33. 式(42)之軌跡;式中之 K, C 及 $d=+$, $r=-$

e. 公式 蒲内曲線方程式內含有四個常數,故須選取四組適宜觀察值或四個適宜等距離點以便計算牠們. 設 (X_1, Y_1), (X_1+h, Y_2),

(X_1+2h, Y_3) 及 (X_1+3h, Y_4) 爲四個等距離點. 於是將此四點之坐標代入式 (43) 中,我們有

$$\frac{m-Y_1}{Y_1-d}=Ce^{rX_1},$$

$$\frac{m-Y_2}{Y_2-d}=Ce^{rX_1+rh},$$

$$\frac{m-Y_3}{Y_3-d}=Ce^{rX_1+2rh},$$

$$\frac{m-Y_4}{Y_4-d}=Ce^{rX_1+3rh}.$$

若以前一式除其次一式,則得方程式

$$(45)\qquad \begin{cases} \dfrac{m-Y_2}{Y_2-d}\Big/\dfrac{m-Y_1}{Y_1-d}=e^{rh}, \\ \dfrac{m-Y_3}{Y_3-d}\Big/\dfrac{m-Y_2}{Y_2-d}=e^{rh}, \\ \dfrac{m-Y_4}{Y_4-d}\Big/\dfrac{m-Y_3}{Y_3-d}=e^{rh}. \end{cases}$$

此如

$$(46)\qquad \begin{cases} \dfrac{m-Y_2}{Y_2-d}\Big/\dfrac{m-Y_1}{Y_1-d}=\dfrac{m-Y_3}{Y_3-d}\Big/\dfrac{m-Y_2}{Y_2-d}, \\ \dfrac{m-Y_3}{Y_3-d}\Big/\dfrac{m-Y_2}{Y_2-d}=\dfrac{m-Y_4}{Y_4-d}\Big/\dfrac{m-Y_3}{Y_3-d}. \end{cases}$$

卽是

$$(47)\qquad \begin{cases} Qm+Qd-Pmd+R=0, \\ Q'm+Q'd-P'md+R'=0. \end{cases}$$

此處

$$P=2Y_2-(Y_1+Y_3);$$

$$Q=Y_2^2-Y_1Y_3,$$

$$R=2Y_1Y_2Y_3-Y_2^2(Y_1+Y_3);$$

$$P'=2Y_3-(Y_2+Y_4),$$

$$Q'=Y_3^2-Y_2Y_4,$$

$$R'=2Y_2Y_3Y_4-Y_3^2(Y_2+Y_4).$$

由式 (47) 得

$$\frac{Qm+R}{Pm-Q}=\frac{Q'm+R'}{P'm-Q'}(=d),$$

及

$$\frac{Qd+R}{Pd-Q}=\frac{Q'd+R'}{P'd-Q'}(=m).$$

因之

$$(48)\quad\begin{cases}(PQ'-P'Q)m^2+(PR'-P'R)m-(QR'-Q'R)=0,\\(PQ'-P'Q)d^2+(PR'-P'R)d-(QR'-Q'R)=0.\end{cases}$$

以 m 及 d 當做未知數解之，則得下列之四組答案：

第一組 $\begin{cases}m=\beta+\gamma,\\d=\beta+\gamma;\end{cases}$ 第二組 $\begin{cases}m=\beta+\gamma,\\d=\beta-\gamma;\end{cases}$

第三組 $\begin{cases}m=\beta-\gamma,\\d=\beta+\gamma;\end{cases}$ 第四組 $\begin{cases}m=\beta-\gamma,\\d=\beta-\gamma;\end{cases}$

彼處 $\gamma=\dfrac{\sqrt{(QR'-Q'R)\alpha+\alpha^2\beta^2}}{\alpha}$, $\beta=-\dfrac{(PR'-P'R)\times .5}{\alpha}$, 並且 $\alpha=PQ'-P'Q$. 但由式 (48) 之前一式減去後一式，得

$$(PQ'-P'Q)(m^2-d^2)+(PR'-P'R)(m-d)=0,$$

並且 $m=K+d$ 爲 Y 之最高極限值，d 爲其最低極限值，在一般情形中，$m-d\neq 0$; 於是末式得用 $(m-d)$ 除之，並且除得之結果爲

$$m+d=2\beta.$$

由此式可以斷定上列四組答案中，祇有居中兩組答案與此條件相合，並且首末兩組答案宜棄捨之．因為 $K(=m-d)$ 之值宜常為正，又因由第二組答案算得之 $K=2\gamma$，由第三組答案算得之 $K=-2\gamma$；於是，

$$
\text{(49)}\qquad \begin{cases} \text{當 } \gamma>0 \text{ 時,} \\ m=\beta+\gamma, \\ d=\beta-\gamma, \\ K=2\gamma. \end{cases}
$$

$$
\text{(50)}\qquad \begin{cases} \text{當 } \gamma<0 \text{ 時,} \\ m=\beta-\gamma, \\ d=\beta+\gamma, \\ K=-2\gamma. \end{cases}
$$

此為計算 m, d 及 K 之公式．

因為 $\log_e\dfrac{m-Y}{Y-d}$ 與 X 為一直線關係[見式(44)]，故 r 及 C 之值可由已給之兩點 (X_1, Y_1) 及 (X_1+h, Y_2) 測定之．將此兩點之坐標分別代入式(44)內，我們有

$$\log\frac{m-Y_1}{Y_1-d}=\log C+rX_1\log e.$$

$$\log\frac{m-Y_2}{Y_2-d}=\log C+r(X_1+h)\log e.$$

以 C 及 r 當做未知數解之，得

$$\log C=\frac{1}{D}\begin{vmatrix} \log\dfrac{m-Y_1}{Y_2-d} & X_1\log e \\ \log\dfrac{m-Y_2}{Y_2-d} & (X_1+h)\log e \end{vmatrix}$$

$$r=\frac{1}{D}\begin{vmatrix} 1 & \log\dfrac{m-Y_1}{Y_1-d} \\ 1 & \log\dfrac{m-Y_2}{Y_2-d} \end{vmatrix},$$

彼處

$$D=\begin{vmatrix} 1 & X_1\log e \\ 1 & (X_1+h)\log e \end{vmatrix};$$

或者說

$$(51)\qquad \begin{cases} C=\operatorname{antilog}\left\{\log\dfrac{m-Y_1}{Y_1-d}-\dfrac{\Delta_1 X_1}{h}\right\}, \\ r=\dfrac{\Delta_1}{h}\log_e 10, \end{cases}$$

彼處

$$\Delta_1=\log\frac{m-Y_2}{Y_2-d}-\log\frac{m-Y_1}{Y_1-d}.$$

f. 步驟　根據上面推得之公式,可將計算之步驟總括之如下:

計算 $P, Q, R; P', Q', R'$　茲將計算 $P, Q, R; P', Q', R'$ 之詳細手續列入下表:

表 39.　$P, Q, R; P', Q', R'$ 之計算手續

手　續	公	式
(1)	$Y_2\times Y_2=$	$Y_3\times Y_3=$
(2)	$Y_2\times 2=$	$Y_3\times 2=$
(3)	$Y_3\times Y_1=$	$Y_2\times Y_4=$
(4)	$Y_3+Y_1=$	$Y_2+Y_4=$
(2)−(4)	$P=$	$P'=$
(1)−(3)	$Q=$	$Q'=$
(2)×(3)−(1)×(4)	$R=$	$R'=$

* 計算式中 $Y_2\times Y_2$, $Y_2\times 2$, $Y_2\times Y_4$ 及 Y_2+Y_4, 可將 Y_2 登記在打字機上,然後分別乘以 Y_2, 2, Y_4 及加以 Y_4; 同樣計算 $Y_3\times Y_3$, $Y_3\times 2$, $Y_3\times Y_1$ 及 Y_3+Y_1, 可將 Y_3 登記在打字器上,然後分別乘以 Y_2, 2, Y_1 及加以 Y_1: 這樣便更能省事些。

(ii) 計算 α, β 及 γ

根據上表裏算得之各數值可進行計算

$$(52)\qquad \begin{cases} \alpha = PQ' - P'Q, \\ \beta = -(PR' - P'R) \times \frac{1}{2} \div \alpha, \\ \gamma = \dfrac{\sqrt{(QR' - Q'R)\alpha + \alpha^2\beta^2}}{\alpha} \end{cases}$$

(iii) 計算 m 及 d

當 $\gamma>0$ 時，我們用下式以求 m, d 及 K 之值：

$$(53)\qquad \begin{cases} m = \beta + \gamma, \\ d = \beta - \gamma, \\ K = 2\gamma. \end{cases}$$

反之 當 $\gamma<0$ 時，我們用下式以求 m, d 及 K 之值：

$$(54)\qquad \begin{cases} m = \beta - \gamma, \\ d = \beta + \gamma, \\ K = -2\gamma. \end{cases}$$

(iv) 計算 Δ_1 及 Δ_2 為計算 C 及 r, 須先計算 Δ_1 及 Δ_2. 茲將計算 Δ_1 及 Δ_2 之手續詳列於表 40 (表見下頁).

若專為計算 C 及 r, 我們祗須求出 Δ_1 之值就夠了. 上表所以將 Δ_2 之值求出者，目的在驗算已經求得之 m 及 d 之值有無舛錯耳. 因為我們知道，Δ_1 與 Δ_2 均為直線 $\log\frac{m-Y}{Y-d} = \log C + rX$ 之斜率，其必相等 (參閱定理 2). 故由 Δ_1 與 Δ_2 之相等，可確信算得之 m 及 d 無誤.

表 40. Δ_1 及 Δ_2 之計算手續

X	Y	$\frac{m-Y}{Y-d}$	$\log\frac{m-Y}{Y-d}$	Δ
X	Y_1	$\frac{m-Y_1}{Y_1-d}$	$\log\frac{m-Y_1}{Y_1-d}$	Δ_1
$X+h$	Y_2	$\frac{m-Y_2}{Y_2-d}$	$\log\frac{m-Y_2}{Y_2-d}$	Δ_2
$X+2h$	Y_3	$\frac{m-Y_3}{Y_3-d}$	$\log\frac{m-Y_3}{Y_3-d}$	

旣然求得 Δ_1 之値，並且 Δ_1 果與 Δ_2 相等，可將 Δ_1 代入式(51)以求 C 及 r 之値.

g. 舉例　茲為說明運算手續起見，舉例如下：

例 3.　用定點法以蒲芮曲線配合後面表 43 裏所列之論據.

取 $Y_1=1.89$, $Y_2=2.40$, $Y_3=3.80$, $Y_4=5.52$. 於是將此等數値代入表 39，我們有

表 41.　P, Q, R; P', Q', R' 之計算

手　續	結	果
(1)	5.7600	14.4400
(2)	4.8000	7.6000
(3)	7.1820	13.2480
(4)	5.6900	7.9200
(2) − (4)	$P=-0.8900$	$P'=-0.3200$
(1) − (3)	$Q=-1.4220$	$Q'=1.1920$
(2) × (3) − (1) × (4)	$R=1.6992$	$R'=-13.6800$

將上表算得之 P, Q, R, P', Q', R' 之值代入式 (52) 內得

$$\alpha = -1.515620,$$

$$\beta = 4.195124,$$

$$\alpha\beta = 6.359472,$$

$$\gamma = -2.470372 \text{ (其值為負).}$$

於是由式 (54) 得

$$m = 6.665496,$$

$$d = 1.724752,$$

$$K = 4.940744.$$

既然得到 m 及 d 之值，可依照表 40 裏所列之手續以求 Δ_1 及 Δ_2 之值，並應用式 (51) 以求 C 及 r 之值。

表 42. Δ_1 及 Δ_2 之計算

x	Y	$\dfrac{m-Y}{Y-d}$	$\log\dfrac{m-Y}{Y-d}$	Δ
-40	1.89	28.898964	1.4608823	
10	2.40	6.316932	.8005062	$-.660376 = \Delta_1$
60	3.80	1.380797	.1401299	$-.660376 = \Delta_2$

於是

$$C = \log^{-1}\left[1.460882 - \frac{-0.660376}{50}(-40)\right]$$

$$= \log^{-1}.932581 = 8.562114,$$

$$r = \frac{-0.660376}{50}\log_e 10 = -0.030413.$$

故所求之蒲內曲線方程式為

$$Y - 1.724752 = \frac{4.940744}{1+8.6521e^{-0.030413x}} \tag{55}$$

得到曲線方程式(55)之後，就該計算各年之人口數，以便與觀察人口數相比較.因爲由式(43)及(44)得知 Ce^{rx} 爲 $(\log C+rx\log_{10}e)$ 之逆對數. 所以若連續將 x 代入 $(\log c+rx\log_{10}e)$ 式內可求得與 x 成對應各 $\log Ce^{rx}$ 之值，並應用對數表可查得此等數值之逆對數，因之最後可得 Ce^{rX} 之值;連續以 $1+Ce^{rX}$ 去除 K, 並加 d 於其上，則得欲求之人口數. 算得之人口數列入表 43,第六行中.計算與觀察人口數之差列入該表末行. 差量最大者爲 0.119, 約佔該年中人口數百分之 2.5. 末行各差量平方平均數的平方根——即通常所稱之根均方差——與標準差頗有相同之處，其值爲 0.0618 百萬.

代表各原始觀察值之點及用定點法配合蒲芮曲線繪入圖 34 中.由表 43 末行或由 43 圖得知配合之情形尚稱滿意.

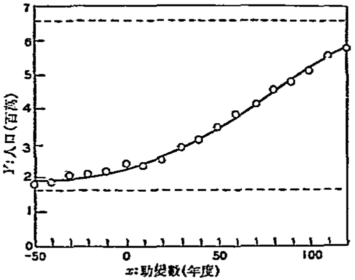


圖 34. 瑞典人口統計及式(55)之軌跡。

表43. 瑞典1750—1920年人口論據及由式(55)算得之人口數

(論據見參加書29)

年度 X (1)	人口(百萬) 觀察值 Y (2)	助變數 x (3)	Ce^{rx} 由式(55) (4)	$\frac{K}{1+Ce^{rx}}$ (5)	人口(百萬) 計算值 Y (6)	差量 (6)−(2) (7)
1750	1.763	−50	39.1800	0.122965	1.848	0.085
60	1.893	−40	28.9000	0.165242	1.890	−0.003
70	2.030	−30	21.3200	0.221359	1.946	−0.084
80	2.118	−20	15.7300	0.295322	2.020	−0.098
90	2.158	−10	11.6100	0.391811	2.117	−0.041
1800	2.347	0	8.5610	0.516760	2.242	−0.105
10	2.378	10	6.3170	0.675242	2.400	0.022
20	2.585	20	4.6600	0.872923	2.598	0.013
30	2.888	30	3.4380	1.113282	2.838	−0.050
40	3.139	40	2.5360	1.397269	3.122	−0.017
50	3.483	50	1.8710	1.720913	3.446	−0.037
60	3.860	60	1.3810	2.075070	3.800	−0.060
70	4.168	70	1.0102	2.457837	4.183	0.015
80	4.566	80	.7514	2.821025	4.546	−0.020
90	4.785	90	.5542	3.178962	4.904	0.119
1900	5.136	100	.4089	3.506809	5.232	0.096
10	5.522	110	.3017	3.795609	5.520	−0.002
20	5.766	120	.2226	4.041178	5.766	0.000

第八節　幾近法

用分組法或定點法所得之配合曲線固甚美滿，但常常不能盡合吾人之意. 吾人常欲得到者爲配合最美滿之曲線，即根據最小平方法原則得之者. 上面曾經講過，許多著名曲線如龔裴曲線. 高培志曲線及指數曲線，不能以常數之一次式表之，故其不能用最小平方法配合. 所幸得

到配合美滿之曲線後，該曲線可視作用泰來定理(Taylor's theorem)求得配合最美滿曲線之第一次幾近線．如用泰來定理，將式(42)展開，可得欲求益正項之一次式，此式之疇範方程式可依前法寫出之，因此得以測定益正項之值．通常用第二次幾近法求得之結果即能滿意；倘仍覺不慊於心，我們可重複上面手續，以求其第三、第四等次之幾近線，直至心滿意足爲止．此法叫做依最小平方法原則配合曲線之幾近法，或簡稱之爲幾近法．

現在我們舉岗蒲曲線爲例，說明用此法配合論據之計算手續．

例．設已求得 r 之幾近值 $\rho=-0.0235470$，C 之幾近值 $C'=7.086190$，d 之幾近值 $\sigma=1.56$ 及 K 之幾近值 $K'=6.1$．試以蒲岗曲線(42)用幾近法配合表 43 裏所列之瑞典入口論據．

設 h, i, j 及 k 各爲 ρ, C', σ 及 K 之益正值．於是

$$\begin{aligned} Y&=f(r, C, d, K)\\ &=f(\rho+h, C'+i, \sigma+j, K'+k), \end{aligned}$$

彼處 $f(r, C, d, K)$ 代表 $d+\dfrac{K}{1+Ce^{rX}}$.

用泰來定理將末式展開之，並棄捨一次以上之益正項；我們有

$$f(r, C, d, K)=f_0+\left(h\frac{\partial f_0}{\partial r}+i\frac{\partial f_0}{\partial C}+j\frac{\partial f_0}{\partial d}+k\frac{\partial f_0}{\partial K}\right),$$

彼處

$$f_0=f(r, C, d, K)_{r=\rho,\ C=C',\ d=\sigma,\ K=K'},$$

$$\frac{\partial f_0}{\partial r}=\left(\frac{\partial f}{\partial r}\right)_{r=\rho,\ C=C',\ d=\sigma,\ K=K'},$$

$$\frac{\partial f_0}{\partial c}=\left(\frac{\partial f}{\partial c}\right)_{r=\rho,\ C=C',\ d=\sigma,\ K=K'},$$

$$\frac{\partial f_0}{\partial d}=\left(\frac{\partial f}{\partial d}\right)_{r=\rho,\ C=C',\ d=c,\ K=K'},$$

$$\frac{\partial f_0}{\partial K}=\left(\frac{\partial f}{\partial K}\right)_{r=\rho,\ C=C',\ d=c,\ K=K'},$$

設 $\frac{\partial f_0}{\partial r}=\xi$, $\frac{\partial f_0}{\partial C}=\eta$, $\frac{\partial f_0}{\partial K}=\zeta$, 因為 $\frac{\partial f_0}{\partial d}=1$, 所以

(56)
$$Y=f_0+h\xi+i\eta+j+k\zeta.$$

因 Y 之觀察值與其計算值之差

$$d_Y=Y-(f_0+h\xi+i\eta+j+k\zeta),$$

故式 (56) 之常範方程式為

(57)
$$\begin{cases} h\Sigma\xi^2+i\Sigma\xi\eta+j\Sigma\xi+k\Sigma\xi\zeta+\Sigma\xi\lambda=0,\\ h\Sigma\xi\eta+i\Sigma\eta^2+j\Sigma\eta+k\Sigma\eta\zeta+\Sigma\eta\lambda=0,\\ h\Sigma\xi+i\Sigma\eta+Nj+k\Sigma\zeta+\Sigma\lambda=0,\\ h\Sigma\xi\zeta+i\Sigma\eta\zeta+j\Sigma\zeta+k\Sigma\zeta^2+\Sigma\zeta\lambda=0; \end{cases}$$

設處
$$\lambda=f_0-Y.$$

但
$$\xi=\frac{-K'e^{\rho X}C'X}{(1+C'e^{\rho X})^2},$$

$$\eta=\frac{-K'e^{\rho X}}{(1+C'e^{\rho X})^2},$$

$$\zeta=\frac{1}{1+C'e^{\rho X}};$$

故 $\eta=-K'\zeta^2e^{\rho X}$, $\xi=\eta C'X$ 並且 $\lambda=c+K'\zeta-Y$. 因此,當實際運算時,我們宜先求 ζ 之值,次求 η 之值,再次求 ξ 之值,最後求 λ 之值.

為求 $\Sigma\xi^2$, $\Sigma\xi\eta$,……之值,爰列下表:

表 44.　以幾近法配合蕭內曲線之計算手續

取 $\rho=-0.0235470,\ c'=7.086190,\ \sigma=1.56$ 並且 $K'=6.1$

年度 X	Y	x	ξ	η	ζ	λ	s
1750	1.763	−50	12.17[illegible]822	−.03134340	.01165668	.05116675	12.2272[illegible]
60	1.893	−40	12.061526	−.04255293	.05215234	−.01487073	12.05[illegible]
70	2.030	−30	11.137352	−.05228991	.06 097 3	−.07290 07	11.071[illegible]
80	2.118	−20	9.079659	−.06406587	.08098142	−.06101334	9.03276[illegible]
90	2 1 8	−10	5.50[illegible]860	−.07769845	.10032510	.01398311	5.542470
1800	2.347	0	0.000000	−.09329148	.1236676[illegible]	−.03262740	− 0.002210
10	2.378	10	− 7.812537	−.11067354	.15152616	.10631141	− 7.695674
20	2.58	20	−18.343844	−.12943376	.18134096	.09947986	−18.189[illegible]
30	2.888	30	−31.647[illegible]99	−.14886984	.22239903	.02863408	−31.545436
40	3.139	40	−47.611473	−.16797275	.2617[illegible]439	.04210178	−47.471[illegible]
50	3.483	50	− 65.71492	−.1874732	.31414720	−.00670208	−65.[illegible]
60	3.860	60	− 85.02076	−.1999682	.36694809	−.06161667	−84.91540
70	4.168	70	−104.22807	−.2101231	.42314817	−.02671616	−104.04182
80	4.566	80	−121.83133	−.2149098	.48141004	−.06939876	−121.63423
90	4.787	90	−136.363[illegible]6	−.21[illegible]8[illegible]71	.54018201	.07011026	−135.96703
1900	5.136	100	−1 6 65862	−.2069610	.[illegible]9785724	.07092916	−146.19679
10	5.522	110	−152.0[illegible]367	−.1950701	6[illegible]294601	02097066	−151.5[illegible]780
20	5.901	120	−152.47080	−.1792051	.70422178	−.01824711	−151.99[illegible]
總計	…	…	−1019 82392	−2.5269531	5.36877[illegible]09	.10670974	−1016.87761

上表($s=\xi+\eta+\zeta+\lambda$)一行為驗對各行數值，各行數值之平方及兩行數值之相乘積有無錯誤而設．因 $s=\xi+\eta+\zeta+\lambda$，故 $\Sigma s=\Sigma\xi+\Sigma\eta+\Sigma\zeta+\Sigma\lambda$，$\Sigma\xi s=\Sigma\xi^2+\Sigma\xi\eta+\Sigma\xi\zeta+\Sigma\xi\lambda$，$\Sigma\eta s=\Sigma\xi\eta+\Sigma\eta^2+\Sigma\eta\zeta+\Sigma\eta\lambda$，$\Sigma\zeta s=\Sigma\xi\zeta+\Sigma\eta\zeta+\Sigma\zeta^2+\Sigma\zeta\lambda$．於是

$\Sigma\xi=-1019.823920$	$\Sigma\xi^2=127921.79$
$\Sigma\eta=-\quad 2.52695410$	$\Sigma\xi\eta=\quad 207.20940$
$\Sigma\zeta=\quad 5.36877209$	$\Sigma\xi\zeta=-\quad 543.70734$
$\Sigma\lambda=\quad 0.10650974$	$\Sigma\xi\lambda=-\quad 5.310004$
-1016.87559	$127,579.98$
$\Sigma s=-1016.87561$	$\Sigma\xi s=127,579.99$
$\Sigma\xi\eta=207.20940$	$\Sigma\xi\zeta=-543.70734$
$\Sigma\eta^2=\quad .42744085$	$\Sigma\eta\zeta=-\quad .96685$
$\Sigma\eta\zeta=-\quad .96685321$	$\Sigma\zeta^2=\quad 2.4332851$
$\Sigma\eta\lambda=-\quad .018225622$	$\Sigma s\lambda=\quad .031328884$
206.65176	-542.20957
$\Sigma\eta s=206.65174$	$\Sigma\zeta s=-542.20955$

將此等值代入式 (57) 內，我們有

$$127921.79h+207.20940i-1019.8239j$$
$$-543.70734k-5.310004=0,$$

$$207.20940h+42744085i-2.5269531j$$
$$-.96685321k-.018225622=0,$$

$$-1019.8239h-2.5269531i+18j$$
$$+5.3687721k+.10650974=0,$$

$$-543.70734h-.96685321i+5.3687721j$$
$$+2.4332851k+.031328884=0.$$

解上面聯立方程式可按下表所列之步驟進行，各步驟可以由表上自明，無須再加申說。

表 45. 解聯立方程式之步驟

步驟	首項反商	h	i	j	k	
(1)	.0000078172765	127921.79	207.20940	−1019.8239	−543.70734	−5.3 0001
(2)	.00 8760359	207.20940	.42744085	−2.5260 31	−.96685321	−.018225032
(3)	−.000980 6145	−1019.8239	−2.5260 31	18.	5.3687721	.10650974
(4)	.0018392248	−543.70734	−.96685321	5.3687721	2.43328 1	.0313 8884
(5): 反商 (1)×(1)		1	.0016198132	−.0079722454	−.0042503106	−.0000415098
(6): 反商 (2)×(2)		1	.0020628440	−.0121951661	−.0016660681	−.0000879 75
(7): 反商 (3)×(3)		1	.0024778328	−.0176 01061	−.00526441 0	.000 044393
(8): 反商 (4)×(4)		1	.0017789604	−.00987 3788	−.0044753583	−.0000576209
(9): (8)−(5)	6311.250687		.0001581472	−.0019041133	−.0002250477	−.0000161111
(10): (6)−(5)	2257.174825		.0004430317	−.0042229210	−.0004157577	.0000464477
(11): (7)−(5)	1165.474541		.0008580196	−.0096778607	−.0010141004	−.0000629205
(12): 反商 (9)×(9)			1	−12.0048707475	−1.4203324512	−.1016811900
(13): 反商 (10)×(10)			1	−4.5318709691	.9384378437	.1048405791
(14): 反商 (11)×(11)			1	−11.2793002571	1.1819081982	.0733427301
(15): (13)−(12)	.4013721101			2.4749697581	.8189437 0	−.0031593882
(16): (14)−(12)	1.3782828671			.7255401704	.2381242530	−.0283384608
(17): 反商 (15)×(15)				1	.19486 6454	−.0012775685
(18): 反商 (16)×(16)				1	.3286160630	.0390584150
(19): (18)−(17)	7.476556271				.1337514176	.0403359835
(20): 反商 (19)×(19)					1	.30157425

由手續 (20)　　　$k=-.301574.$

由手續 (18)　　　$j=\ .0600437.$

由手續 (14)　　　$i=\ .394161.$

由手續 (8)　　　$h=-.00140006.$

故　　$r=-.0235470-.00140006=-.02494706,$

$C=7.086190+.394161=7.480351,$

$d=1.56+.0600437=1.6200437,$

$K=6.1-.301574=5.798426;$

並且所求之方程式爲

(58)　　$$Y-1.62004=\frac{5.798426}{1+7.480351e^{-.02494706x}}$$

由上式算得之人口數列入表 46 第三行中．計算手續詳見表 43，茲不詳及．爲比較配合情形之優劣起見，曾將原來觀察之人口數及觀察人口數與算得人口數之差量亦行分別列入表 43 第五及第六行中．由於如此算得之差量，可見配合之情形尙佳．由新曲線［曲線(58)，用幾近法配合者］算得之根均方差爲 0.0562，其較由舊曲線［曲線(55)，用定點法算得者，根均方差爲 0.0619］算得者稍小，換言之，新曲線配合論據較舊者爲優．

觀察値與配合之曲線顯示於圖 35 中，由此圖亦可見配合情形美滿．

表46. 觀察人口與由式(58)算得人口之比較

年度 X (1)	x (2)	人口數 計算者 (3)	人口數 觀察者 (4)	人口數 差量 (5)
1750	-50	1.8344	1.763	+.0714
60	-40	1.8923	1.893	-.0007
70	-30	1.9649	2.030	-.0651
80	-20	2.0553	2.118	-.0627
90	-10	2.1670	2.158	+.0090
1800	0	2.3037	2.347	-.0433
10	10	2.4691	2.378	+.0911
20	20	2.6663	2.585	+.0813
30	30	2.8974	2.888	+.0094
40	40	3.1631	3.139	+.0241
50	50	3.4614	3.483	-.0216
60	60	3.7881	3.860	-.0719
70	70	4.1359	4.168	-.0321
80	80	4.4953	4.566	-.0707
90	90	4.8554	4.785	+.0704
1900	100	5.2053	5.136	+.0693
10	110	5.5352	5.522	+.0132
20	120	5.8377	5.904	-.0663

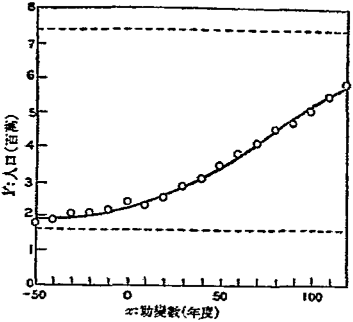


圖 35: 瑞典人口論據及式(58)之軌跡。

問題 V

1. 應用定理 6 證明第二次動差以對於平均數為最小.

2. 保壽險 $1000, 各年齡 (X) 應繳之保險費若干元 (Y) 如下表所示:

X	20	25	30	35	40	45	50	55	60
Y	18.78	21.02	23.86	27.54	32.36	38.83	47.68	59.88	76.94

試繪一圖表示 Y 為 X 的函數. 由圖上估計 32 歲及 34 時應繳之保費, 並估計應繳保費 $ 52 之年齡.

3. 試以 Y 之剩餘數平方和最小為準則配合下列論據:

X	6	7	7	8	8	8	9	9	10
Y	5	5	4	5	4	3	4	3	3

4. 證明由上題方程式算得之 $\Sigma d_Y=0$.

5. 試以 X 之剩餘數平方和最小爲準則配合練習題 2 裏所列之論據.

6. 證明由上題方程式算得之 $\Sigma d_X=0$.

7. 試以 r 之剩餘數平方和最小爲準則配合練習題 2 裏所列之論據

8. 證明由上題方程式算得之 $\Sigma d_p=0$.

9. 詳述單對數與複對數格紙之主要用途.

10. 假設 Y 對於 X 之變化率與 Y 之最高極限值成正比例.問 Y 爲 X 的何種函數?

11. 假設 Y 對於 X 之變化率與 Y 之最高極限值減以 Y 之差成正比例.問 Y 爲 X 之何種函數?

12. 假設 $\log Y$ 對於 X 之變化率與 X 之比恆爲一定.問 Y 爲 X 的何種函數?

13. 假設 $\log Y$ 對於 X 之變化率與 Y 之最高極限值減以 Y 之差成正比例.問 Y 爲 X 的何種函數?

14. 求配合下列各點最美滿拋物線之方程式:

(−4, 2) (0, 8) (4, 9) (8, 11) (12, 8) (16, 5).

15. 假設各 t 值成一等差級數,並且 $\Sigma t=0$. 證明 $\Sigma t^3=0$.

16. 在金屬棒延長實驗中得到下列之論據,彼處 X 代表重量, Y 代表尺度:

X	1	2	3	4	5
Y	14	27	40	55	68

試用最小平方理論測定一直線方程式.

17. 若將上題之 X 各加以 6, 上題之 Y 各加以 300, 試求配合最美滿指數曲線之方程式,並將已給之論據及所得之曲線繪在單對數格紙上.

18. 區別函數 $Y=Ae^{-BX}$ 及 $Y=Ke^{-kX^2}$ 所表示之曲線形式,彼處 A, B, K 及 k 爲正實數. 如將其繪在單對數格紙上其形式如何?

19. 試由式(35)計算 $\Sigma d_{\log Y}$ 及 $\Sigma d^2_{\log Y}$, 並計算 Σd_Y 及 Σd_Y^2 之值.此練習題表演之重要事實爲何?

20. 如果以 Σd_Y^2 最小爲準則,試寫出以曲線 $Y=ab^X$, 用幾近法,配合已給論據所需之疇範方程式.

21. 試用上題所得之疇範方程式及式(35)內 a 及 b 之值(卽 $a=.5233$, $b=2.9762$, 配合表 36 裏所列之論據.

22. 由上題所得之結果計算 $\Sigma d_{\log Y}$ 及 $\Sigma d^2_{\log Y}$ 之值;並計算 Σd_Y 及 Σd_Y^2 之值.這些值與練習題 19 所得之各對應值相比較,你所能得到之結論爲何?

23. 在何種情形之下,配合論據時可以 $\Sigma d^2_{\log Y}$ 最小爲準則?並在何種情形之下不可? 試在文獻中覓一實例說明之.

24. 討論曲線配合分組法之優劣各點.

25. 試以最小平方法作準則，用高培志曲線配合表 36 裏所列之論據.

26. 試用定理 2 證明 $\Delta_1=\Delta_2$ (參閱表 40).

27. 試直接由式 (46) 推求計算 m 之公式.

28. 假設蒲芮曲線方程式內之常數 C, d, K 及 r 不受任何限制,試討論圖形之可能樣式.

第六章　尺度變換法之理論及其應用

第一節　頻數曲線方程式與面積

尺度變換法(scale transformation)者係用新尺度表示原始尺度之方法也.用此法能將在原始尺度上變數之分配化成一種在新尺度上者,並且由於知道其在新尺度上分配之形式能將其在原始尺度上分配形式測得.茲將尺度變換法所根據之理論詳述於下.設

(1)
$$Y=f(X)$$

爲在原始尺度(舊尺度)上機率曲線之方程式,該方程式爲我們正欲尋求者;設

(2)
$$y=\phi(t)$$

爲在變換尺度(新尺度)上機率曲線之方程式,其形式爲已知之者;並且設

(3)
$$t=\Omega(X)$$

爲欲測定之變換方程式(equation of transformation),彼處 t 爲 X 之單值函數.因此我們有下列之定理.

定理 1.　如果在原始尺度上機率曲線(1)及在變換尺度上機率曲線(2)下之對應面積相等,則

(4)
$$f(X)=\phi(t)t',$$

彼處 t' 爲 t 對於 X 之第一次微係數.

證: 因爲等面積之條件用算式表之爲

$$\int_{X_i}^{X_j} f(X)dX = \int_{t_i}^{t_j} \phi(t)dt. \tag{5}$$

彼處 $t_i=\Omega(X_i)$, $t_j=\Omega(X_j)$; 並且用幾何解釋之如圖 36 所示. 此圖分上下兩幅: 上幅爲在 X 尺度上之機率曲線——原始曲線 (original curve); 下幅爲在 t 尺度上之機率曲線——變換曲線 (transformed curve). 該二曲線帶有陰影部份的面積是相等的, 並且 X 軸上之尺度與 t 軸上之尺度一一成對應.

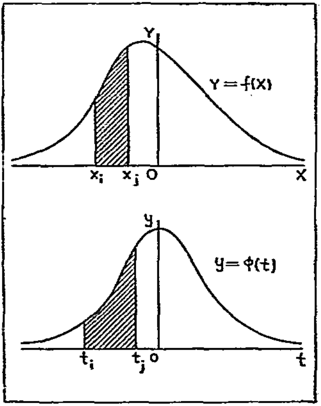


圖 36. 尺度變換後對於機率曲線形式之影響。

因爲方程式(5)不論 X_j 及與之成對應的 t_j 爲何值常是成立的.吾人可以將彼處之 X_j 及 t_j 各以 X 及 t 代之,即

$$\int_{X_i}^{X} f(X)dX=\int_{t_i}^{t}\phi(t)dt.$$

將末式兩邊對於 X 微分之,我們有

$$f(X)=\phi(t)t'.$$　證訖.

系　$Y=\phi(t)t'$ 爲在原始尺度上機率曲線之方程式.

定理 2.　如果方程式(4)正確,式(1)代表原始機率曲線方程式,並且式(2)代表變換機率曲線方程式;則在曲線(1)及(2)下之對應面積相等.

證:　由積分上定理得

$$\int_{t_i}^{t_j}\phi(t)dt=\int_{X_i}^{X_j}\phi(t)t'dX.$$

但由式(4)　$$\phi(t)t'=f(X).$$

故　$$\int_{t_i}^{t_j}\phi(t)dt=\int_{X_i}^{X_j}f(X)dX.$$

即新舊二曲線下之對應面積相等.　證訖.

第二節　$\phi(t)$ 之選擇

因爲 $\phi(t)$ 爲 t 之任意函數,故吾人須選擇一種適宜者,以便吾人由 t 之形式[即 $\Omega(X)$ 之形式]能以確定原始機率函數 $f(X)$ 之形式. $\phi(t)$ 之最方便而且最有意義之形式當爲常態函數,因此取

(7)　$$\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

爲已知之變換機率函數，或者說取

$$(8) \qquad y=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

爲已知之變換機率曲線方程式.

曲線(8)下面之面積及其縱坐標，可由已做成之表直接查得之，上面所謂取常態函數爲變換函數較爲便利者，即指此而言. 於是由式(4)得知函數 $f(X)$ 可寫做

$$(9) \qquad f(X)=\frac{t'}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

或者說欲測之原始機率曲線方程式爲

$$Y=\frac{t'}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$$

即欲測之原始頻數曲線方程式爲

$$(10) \qquad Y=\frac{Nt'}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}.$$

此式之軌跡因 t' 值之不同而異(見圖 41 至 49).

第三節　t 或 $\Omega(X)$ 之測定

如果欲將原來曲線方程式明確寫出，自然我們必須知道函數 $\Omega(X)$ 的形式，即須明確寫出方程式(3). 如何纔能寫出式(3)呢？對於此問題之實際解答纔做由觀察論證測定一個算式，明確表示變換後變數之尺度與原來變數尺度之關係，如此在新尺度上之機率曲線爲常態.

變換方程式之測定當然以等面積變換之基本條件，如式(5)所陳述

者爲根據．每有一 X 值我們可以說有百分之幾觀察值小於 X 值，即每有一 X 值可以算得與之成對應之相對累計頻數，因此假定各觀察值之累計百分數合乎常態分配，我們便可檢得各 t 之值，即檢得當 $\int_{-\infty}^{t}\phi(t)dt=\int_{-\infty}^{X}f(X)dX$ 時之 t 值．於是可得一羣 X 與 t 之對應點．例如觀察值有 10% 在 X_i 以下，於是因爲在常態曲線下面積在 $t_i=-1.28$ 點以下者有 10%（見圖 37），因之 $X=X_i$, $t=t_i=-1.28$，其爲一雙在兩個尺度上之對應點．此等一羣對應點 (X_i, t_i) 表顯之趨勢能使吾人用曲線配合法測定之，就是說用算式將 t 與 X 之關係明白表示出來．

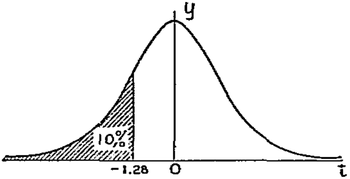


圖 37．當面積 $\int_{-\infty}^{t}=10\%$ 時，$t=-1.28$ 之位置。

第四節　舉例

在未思索正文之先，關於如何由 X 值得到對應 t 值之手續，用一實例解說之．正文內欲講述者爲測定變換方程式之簡易代數式．表 47 係公元一九一八年美國 234 郡黑人嬰兒死亡率之統計，圖 38 a 爲此頻數分配之直方圖．

表 47. 美國登記區黑人嬰兒死亡率及其對應之標準尺度

組段 (嬰兒死亡率) (1)	頻數 f (2)	組段之高限 X (3)	在 X 以下頻數 絕對數 (4)	在 X 以下頻數 百分數 (5)	標準尺度 t (6)
−20−	0	0	0	.0000	$-\infty$
0−	1	20	1	.4274	−2.6295
20−	1	40	2	.8547	−2.3847
40−	5	60	7	2.9915	−1.8821
60−	22	80	29	12.3932	−1.1556
80−	19	100	48	20.5128	−.8235
100−	42	120	90	38.4615	−.2934
120−	37	140	127	54.2735	.1073
140−	30	160	157	67.0940	.4425
160−	19	180	176	75.2137	.6812
180−	17	200	193	82.4786	.9338
200−	10	220	203	86.7521	1.1148
220−	12	240	215	91.8803	1.3971
240−	8	260	223	95.2991	1.6746
260−	4	280	227	97.0085	1.8821
280−	2	300	229	97.8632	2.0263
300−	1	320	230	98.2906	2.1179
320−	…	340	…	…	…
340−	…	360	…	…	…
360−	1	380	231	98.7179	2.2315
380−	…	400	…	…	…
400−	1	420	232	99.1453	2.3847
420−	…	440	…	…	…
440−	1	460	233	99.5727	2.6307
460−	1	480	234	100.0000	∞
總計	234	…	…	…	…

由圖 38a 可見這個頻數分數，呈現偏態．用變換法求得方程式以描寫頻數分配之步驟，曉示於表 47 第三、第四、第五及第六行中．第三行爲嬰兒死亡率之組段；第四及第五兩行爲死亡率在 X 以下之頻數及

其百分數，即嬰兒死亡率在某數值以下之郡數及其百分數；並且第六行之值由常態曲線下面積表查得之標準尺度 t. 圖 38b 顯示頻數分配在 t 尺度上時爲常態之情形.

X 與 t 成功一羣具有相關模式的數列，如圖 38c 所示. 吾人的主要目的是要用曲線配合法測定一方程式以描寫此種模式 (pattern). 如何求出描寫此模式曲線之步驟容後講述，現在僅聲明下列之方程式

$$(11)\qquad \log(d-t)=r_0+r_1X+r_2X^2+r_3X^3$$

$$d=2.631$$

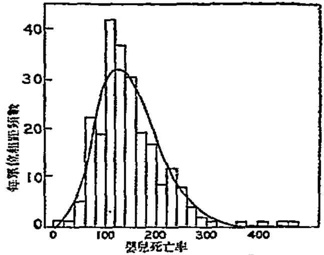
$$r_0=.7802$$

$$r_1=-.2729\times10^{-2}$$

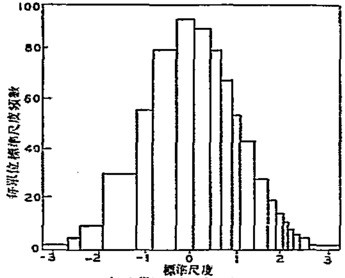
$$r_2=.3226\times10^{-5}$$

$$r_3=.1695\times10^{-7}$$

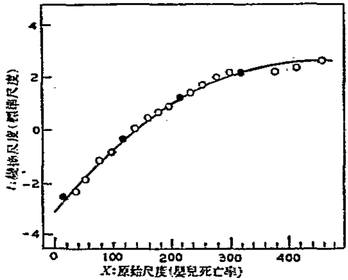
供給此等對應點之描寫甚爲滿意，其滿意情形可由繪在圖 38a 及 38c 裏面之圓滑曲線見之.



a 在原始尺度上之直方圖及配合之曲線



b. 在變換尺度上之直方圖



c. 原始尺度與標準尺度之關係

圖 38. 美國登記區 1918 年 234 都市人嬰兒死亡率。

圖 38 c 裏所給者爲方程式 (11) 代表之曲線及表 47 內之觀察點 (X, t).

因爲方程式 (10)　　$Y=\frac{Nt'}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$

中之 t 值可由方程所得之式 (11) 算得,故原始頻數曲線可以畫出.圖 28a 裏所繪者為觀察頻數分配之直方圖及上面式(10)所代表之原始頻數曲線.

第五節 變換方程式

1. 一般形式 如前所述,關於應用變換法之實在問題為選取一適宜變換尺度方程式,即測定適合於實際問題 $\Omega(X)$ 之形式.為得到關於本問題之線索起見,吾人曾普遍考查各種不同頻數分配之直方圖,及與之成對應之標準尺度 t 與原始尺度 X 關係之趨勢圖,於是發現 X 與 t 有一明確之聯繫如下:

(i) 如果頻數分配曲線之一端或兩端與其橫軸或高度相切 (high contact), 則 X 與 t 之趨勢曲線有一條或兩條水平漸近線 (horizontal asymptote); 反之則否.

(ii) 如果頻數分配曲線之一端或兩端突然而起 (rises abruptly), 則 X 與 t 之趨勢曲線有一條或兩條垂直漸近線 (vertical asymptote); 反之則否.

因此可按 X 與 t 之趨勢曲線有無水平漸近線,將頻數分配分為兩組:第一組與第二組.所有各頻數分配其趨勢曲線無水平漸近線者,則屬於第一組 (參閱圖 39); 其有水平漸近線者,則屬於第二組(參閱圖 40).更可按 X 與 t 之趨勢曲線有無垂直漸近線將屬於第一組之頻數分配分為四種,其為:由頻數分配所得之趨勢曲線之兩端無垂直漸近線者,其始端有一垂直漸近線者,其末端有一垂直漸近線者,並且其始末兩端均有垂直漸近線者 (參閱圖 39 及例 1 至 4); 將屬於第二組者分為五種,其為: 由頻數分配所得之趨勢曲線之始末兩端無垂直漸近線但在其始

端有一水平漸近線者，其末端有一垂直漸近線者，其始末兩端無垂直漸近線但在其末端有一水平漸近線者，其始端有一垂直漸近線者，並且其始末兩端無垂直漸近線者(參閱圖 40 及例 5 至 9).

仔細研究上面所述之事實，可推想到變差尺度方程式之算學形式．吾人知道，0 之對數近於 $-\infty$，故可假定頻數分配之趨勢曲線方程式中含有對數項 $\log(X-a)$、$\log(b-X)$、$\log(t-c)$ 及 $\log(d-t)$，彼處 a、b、c 及 d 均為待定之常數．什麼緣故呢？因為若曲線方程式內包含對數項 $\log(X-a)$ 及 $\log(b-X)$ 此曲線必定有垂直(與 X 軸成垂直)漸近線 $X=c$ 及 $X=d$；若其內包含對數項 $\log(t-c)$ 及 $(d-t)$ 必定有水平(與 t 軸成垂直)漸近線 $t=c$ 及 $t=d$．更進而言之，若欲使此曲線有種種不同曲率，則方程式內仍須含有 X 及 t 之代數項，譬如含有 X 及 t 之有理整式 (rational integral expression).

由是吾人得將上面種種情形用下面所列之算式概括之：

$$\log(t-c)^{\gamma}(d-t)^{\delta}+s_1t+s_2t^2+\cdots$$
$$=\log(X-a)^{\alpha}(b-X)^{\beta}+r_0+r_1X+r_2X^2+\cdots, \tag{12}$$

彼處 a、b、c 及 d，α、β、γ 及 δ，r_0、r_1、r_2、…，s_1、s_2、…均為常數：其中 a 可以取做 0；b、c、d、r_0、r_2、…等為待定之常數；α、β、γ、δ、s_1、s_2、…等為須受限制之常數．茲分別討論如下：

(i) 若使 $a=0$，則曲線 (12) 亦不失其普遍性，因 $a=0$ 僅係將在 X 軸上之原點移一位置．

(ii) 上面已經說過，式 (12) 中包含對數項，意思為該式代表之曲線有漸近線，但漸近線同時不能有兩條，故上式中之常數 α, β, γ, δ 之值不能不受限制，即是四個常數之中至多有兩個不為零．更因在實際上，用常數 γ 及 δ 增加曲線之曲率乃屬無需，只要取 γ 之值為 0 或為 $+1$，並且 δ 之值為 0 或為 ±1 就夠了．

(iii) 祇要我們增加常數 r_0、r_1、r_2、… 之個數，式(12)能代表各種形式之曲線。然而在代表普通頻數分配曲線方程式內，僅包含 r_0、r_1、r_2 及 r_3 即足敷實際上問題之需（見例 1 至 6，例 8 至 9）；僅僅於不習見各分配情形中，如雙峯頻數分配需要多個任意常數（見例 7）。

(iv) 除此之外，我們可取 s_1 爲 0 或爲 1，其他各 s 均爲 0。於是，第一組頻數分配之變換尺度方程式中之 γ 及 δ 均爲 0，因該曲線沒有水平漸近線；第二組頻數分配之變換尺度方程式中之 γ 及 δ 至少有一不爲 0，因該曲線在始端末端或始末兩端有水平漸近線。故關於第一組頻數分配趨勢曲線之方程式爲

$$(13)\qquad t=\log X^{\alpha}(b-X)^{\beta}+r_0+r_1X+r_2X^2+r_3X^3,$$

並且關於第二組者爲

$$(14)\qquad \log(t-c)^{\gamma}(d-t)^{\delta}=\log X^{\alpha}(b-X)^{\beta}+r_0+r_1X+r_2X^2+r_3X^3,$$

PS 此處 α、β、γ 及 δ 四數至多有兩數不爲 0。

如此我們得將變換尺度方程式分成下列各種情形。

3. 各種情形　關於第一組四種不同情形之趨勢曲線可由式(13)得之，其爲

$$(15)\qquad t=r_0+r_1X+r_2X^2+r_3X^3,$$

係使方程式(13)中之 α 及 β 均等於 0 得之者；

$$(16)\qquad t=\log X^{\alpha}+r_0+r_1X+r_2X^2+r_3X^3,$$

係使式(13)中之 $\beta=0$ 得之者；

$$(17)\qquad t=\log(b-X)^{\beta}+r_0+r_1X+r_2X^2+r_3X^3,$$

係使式(13)中之 $\alpha=0$ 得之者；及

$$(18)\qquad t=\log X^{\alpha}(b-X)^{\beta}+r_0+r_1X+r_2X^2+r_3X^3,$$

其即式(13)。

關於第二組之五種不同情形之趨勢曲線可由式(14)得之，其爲

(19) $$\log(t-c)=r_0+r_1X+r_2X^2+r_3X^3,$$

係使式(14)中之 $\alpha=0$、$\beta=0$、$\gamma=1$ 及 $\delta=0$ 得之者；

(20) $$\log(t-c)=\log(b-X)^{\beta}+r_0+r_1X+r_2X^2+r_3X^3,$$

係使式(14)中之 $\alpha=0$、$\gamma=1$ 及 $\delta=0$ 得之者；

(21) $$\log(d-t)=r_0+r_1X+r_2X^2+r_3X^3,$$

係使式(14)中之 $\alpha=0$、$\beta=0$、$\gamma=0$ 及 $\delta=1$ 得之者；

(22) $$\log(d-t)=\log X^{\alpha}+r_0+r_1X+r_2X^2+r_3X^3,$$

即使式(14)中之 $\beta=0$、$\gamma=0$ 及 $\delta=1$ 得之者；及

(23) $$\log\frac{d-t}{t-c}=r_0+r_1X+r_2X^2+r_3X^3,$$

係使式(14)中之 $\alpha=0$、$\beta=0$、$\gamma=-1$ 及 $\delta=1$ 得之者.

上述各種情形可用下列之表總括之.

表 48. 變換式與漸近線

組 別	變 換 式	水平漸近線		垂直漸近線	
		在始端	在末端	經過原點	遠離原點
第一組					
i	(15)	–	–	–	–
ii	(16)	–	–	$X=0$	–
iii	(17)	–	–	–	$X=b$
iv	(18)	–	–	$X=0$	$X=b$
第二組					
i	(19)	$t=c$	–	–	–
ii	(20)	$t=c$	–	–	$X=b$
iii	(21)	–	$t=d$	–	–
iv	(22)	–	$t=d$	$X=0$	–
v	(23)	$t=c$	$t=d$	–	–

3. 略圖　為增加具體之觀念起見，繪第一組及第二組頻數曲線趨勢曲線之略圖如下：

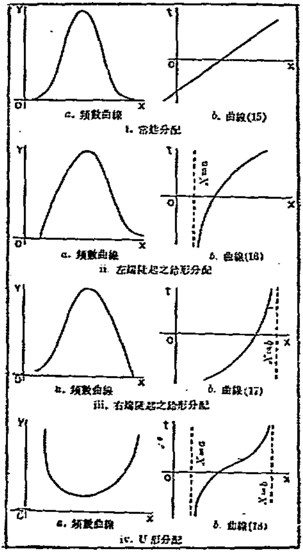


圖 89. 第一組頻數分配及趨勢曲線之草圖。

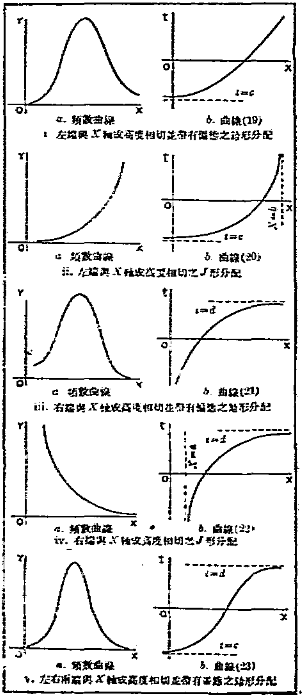


圖 40. 第二組頻數分配及變換曲線之示圖。

第六節　常數之測定

測定包含於變換尺度方程式內常數之方法，依所需準確之程度而異．茲分述各種方法於下：

1. 定點法　用定點法以求常數之值，需要選取之點數與常數之個數相等（見第五章第七節）．因爲公式中之 a, b, c, 及 d 之值可由圖上估計之，故需要測定之常數至多有六個，因之吾人祇須選出適宜之六點來，用以計算該常數之值．設 (X_0, t_1), (X_0+h, t_2), (X_0+2h, t_3), (X_0+3h, t_4), (X_0+4h, t_5), (X_0+5h, t_6) 爲六個等距離點．將此六點之坐標依次代入式 (12) 內，以 $\psi(t)$ 代表式 (12) 左方各項之和，以 $r(X)$ 代表 $r_0+r_1X+r_2X^2+r_3X^3$ 並使式中之 $a=0$；我們有

$$
\begin{aligned}
\psi(t_1)&=\alpha\log X_0 &&+\beta\log(b-X_0) &&+r(X_0).\\
\psi(t_2)&=\alpha\log(X_0+h) &&+\beta\log(b-X_0-h) &&+r(X_0+h).\\
\psi(t_3)&=\alpha\log(X_0+2h) &&+\beta\log(b-X_0-2h) &&+r(X_0+2h).\\
\psi(t_4)&=\alpha\log(X_0+3h) &&+\beta\log(b-X_0-3h) &&+r(X_0+3h).\\
\psi(t_5)&=\alpha\log(X_0+4h) &&+\beta\log(b-X_0-4h) &&+r(X_0+4h)\\
\psi(t_6)&=\alpha\log(X_0+5h) &&+\beta\log(b-X_0-5h) &&+r(X_0+5h).
\end{aligned}
$$

用 Δ、Δ^2、Δ^3、…各代表第一次、第二次、…差量，最後我們可得

$$
(24)\quad
\begin{cases}
\beta = \dfrac{1}{D}\begin{vmatrix} \Delta^4 \log X_0 & \Delta^4 \psi(t_1) \\ \Delta^4 \log (X_0+h) & \Delta^4 \psi(t_2) \end{vmatrix}, \\[2ex]
\alpha = \dfrac{1}{D}\begin{vmatrix} \Delta^4 \log t_1 & \Delta^4 \log(\beta - X_0) \\ \Delta^4 \log t_2 & \Delta^4 \log(\beta - X_0 - h) \end{vmatrix}, \\[2ex]
r_3 = \dfrac{1}{6h^3}[\Delta^3 t_1 - \alpha\Delta^3 \log X_0 - \beta\Delta^3 \log(\beta - X_0)], \\[2ex]
r_2 = \dfrac{1}{2h^2}[\Delta^2 t_1 - \alpha\Delta^2 \log X_0 - \beta\Delta^2 \log(\beta - X_0)] - 3r_3(X_0+h), \\[2ex]
r_1 = \dfrac{1}{h}[\Delta t_1 - \alpha\Delta \log X_0 - \beta \log(\beta - X_0)] \\
\qquad\qquad - r_3(3X_0^2 + 3X_0 h + h^2) - r_2(2X_0 + h), \\[1ex]
r_0 = \psi(t_1) - \alpha \log X_0 - \beta \log(\beta - X_0) - r_1 X_0 - r_2 X_0^2 - r_3 X_0^3;
\end{cases}
$$

彼處
$$
D = \begin{vmatrix} \Delta^4 \log X_0 & \Delta^4 \log(b - X_0) \\ \Delta^4 \log(X_0 + h) & \Delta^4 \log(b - X_0 - h) \end{vmatrix}.
$$

上面為測定常數之一般公式，各種特殊情形之公式可由該公式推得之．各常數之值仍然可用直接代入法求得自不待言，但解如此所得之聯立方程式時，包含不必需運算上之麻煩．將一般公式應用於特殊情形上將在下節例1中解釋之．

2. 最小平方法　如前所述，式(12)內包含六個常數，如此用最小平方法原則測定常數時需要六個規範方程式以代替用定點法測定常數時所需之六點．因為式(12)就所含之常數說是成直線關係，如果以$\psi(x)$剩餘數平方和最小為準則，如是可將規範方程式直接寫出之如下:

$$
(25)\quad
\begin{cases}
\alpha\Sigma A^2f+\beta\Sigma ABf+r_0\Sigma Af+r_1\Sigma AXf+r_2\Sigma AX^2f+r_3\Sigma AX^3f=\Sigma A\psi(t)f\\
\alpha\Sigma ABf+\beta\Sigma B^2f+r_0\Sigma Bf+r_1\Sigma BXf+r_2\Sigma BX^2f+r_3\Sigma BX^3f=\Sigma B\psi(t)f\\
\alpha\Sigma Af+\beta\Sigma Bf+r_0\Sigma f+r_1\Sigma Xf+r_2\Sigma X^2f+r_3\Sigma X^3f=\Sigma\psi(t)f\\
\alpha\Sigma AXf+\beta\Sigma BXf+r_0\Sigma Xf+r_1\Sigma X^2f+r_2\Sigma X^3f+r_3\Sigma X^4f=\Sigma X\psi(t)f\\
\alpha\Sigma AX^2f+\beta\Sigma BX^2f+r_0\Sigma X^2f+r_1\Sigma X^3f+r_2\Sigma X^4f+r_3\Sigma X^5f=\Sigma X^2\psi(t)f\\
\alpha\Sigma AX^3f+\beta\Sigma BX^3f+r_0\Sigma X^3f+r_1\Sigma X^4f+r_2\Sigma X^5f+r_3\Sigma X^6f=\Sigma X^3\psi(t)f
\end{cases}
$$

彼處 f 爲在已給 X 值前一組內之頻數，A 代表 $\log X$ 並且 B 代表 $\log(b-X)$. 上式稱做加權略範方程式，因爲皆將頻數(或權數)計入.

不加權之略範方程式爲:

$$
(26)\quad
\begin{cases}
\alpha\Sigma A^2+\beta\Sigma AB+r_0\Sigma A+r_1\Sigma AX+r_2\Sigma AX^2+r_3\Sigma AX^3=\Sigma A\psi(t)\\
\alpha\Sigma AB+\beta\Sigma B^2+r_0\Sigma B+r_1\Sigma BX+r_2\Sigma BX^2+r_3\Sigma BX^3=\Sigma B\psi(t)\\
\alpha\Sigma A+\beta\Sigma B+r_0\Sigma 1+r_1\Sigma X+r_2\Sigma X^2+r_3\Sigma X^3=\Sigma\psi(t)\\
\alpha\Sigma AX+\beta\Sigma BX+r_0\Sigma X+r_1\Sigma X^2+r_2\Sigma X^3+r_3\Sigma X^4=\Sigma X\psi(t)\\
\alpha\Sigma AX^2+\beta\Sigma BX^2+r_0\Sigma X^2+r_1\Sigma X^3+r_2\Sigma X^4+r_3\Sigma X^5=\Sigma X^2\psi(t)\\
\alpha\Sigma AX^3+\beta\Sigma BX^3+r_0\Sigma X^3+r_1\Sigma X^4+r_2\Sigma X^5+r_3\Sigma X^6=\Sigma X^3\psi(t)
\end{cases}
$$

係使方程式 (25) 中各不同頻數 f 等於 1 得之者.

關於用加權或不加權最小平方法與點法所得結果之比較，見下面例 1 中.

第七節　變換法之應用

爲曉示變換法之應用起見，特舉九種頻數分配實例如下. 但因各例所應用之手續大致相同，故只在例 1 中詳列運算手續(表 49a、b 及 c)，其餘八例祇各用一總表分別列出原始論據，簡要手續及所得之結果. 總表共有八個(表 50 至 57)，每表列有七行. 第一及第三行所列者爲原始論據；第二行爲助變數；第四行爲修勻之頻數；第五及第六兩行爲觀察與

計算之標準尺度；第七行為欲求之頻數曲線之縱坐標.在每表之下列有算得之 X^2, 自由度 n 及機率 P 各值；變換尺度方程式及算得的各常數之值.

九例之論據及結果分繪於九圖(圖 41 至 49)中.每圖分上下兩幅,上幅包含觀察頻數之直方圖及配合之頻數曲線,下幅包含原始度量所用之尺度 X 與標準尺度 t 之點圖及配合之趨勢曲線.

例 1. (常態頻數分配之例) 參加某種考試 514 人成績之分配(論據見參考書 21).

a. 計算標準尺度 由原始論據計算標準尺度之手續詳見表 49 a.

表 49 a. 由已給之論據計算標準尺度

組 段 (分數) (1)	頻 數 f_0 (2)	組段高限 X (3)	助變數 x (4)	在 X 以下頻數之百分數 (5)	標準尺度 t (6)
−4.5 −	−	.5	0	.0000	−∞
.5 −	5	5.5	1	.9728	−2.3367
5.5 −	9	10.5	2	2.7237	−1.9231
10.5 −	28	15.5	3	8.1712	−1.3937
15.5 −	49	20.5	4	17.7043	− .9267
20.5 −	58	25.5	5	28.9883	− .5537
25.5 −	82	30.5	6	44.9416	− .1237
30.5 −	87	35.5	7	61.5677	.3020
35.5 −	79	40.5	8	77.2374	.7467
40.5 −	50	45.5	9	86.9650	1.1248
45.5 −	37	50.5	10	94.1634	1.5687
50.5 −	21	55.5	11	98.2490	2.1081
55.5 −	6	60.5	12	99.4163	2.5219
60.5 −	3	65.5	13	100.0000	∞
總 計	514	…	…	…	…

上表第一行及第二行爲原始論據，第三行爲第一行組段之低限；第四行爲與 X 成對應之助變數 x，即以三行第一數爲原點以組距爲單位照第二章式(5)求得之 x 值；第五行爲由第二行算得之累計頻數之百分數；第六行爲與第五行成對應之標準尺度 t. 計算 t 值之手續已見前第三及第四兩節中，不復述及.

b. 選擇變換尺度方程式並測定常數　由圖 41a 可見本例頻數分配爲一鈴形並且對稱. 由圖 41b 可見標準單位 t 與助變數(代表原始量度單位) x 之關係爲一直線而無漸近線，故知其與表 48 裏第一種情形符合，並且應當以

$$(27)\qquad t = r_0 + r_1 x$$

當做變換尺度方程式.

當用定點法求 r_0 及 r_1 之值時我們可取

$$(x_0, t_1)=(5, -.5537) \text{ 及 } (x_0+h, t_2)=(8, .7467)$$

此即在圖 41b 中以黑點表示之者. 將此等值代入式(24)中得

$$r_1 = \frac{1}{h} \times \Delta_t = \frac{1}{3} \times 1.3004 = .433467,$$

$$r_0 = t - r_1 x_0 = -.5537 - 5 \times (.435467) = -2.721035.$$

當用加權最小平方法求常數之值時，我們宜先由式(25)寫出常範方程式，

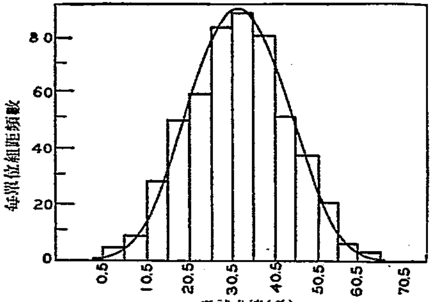
$$r_0\Sigma f + r_1\Sigma xf = \Sigma tf$$

及

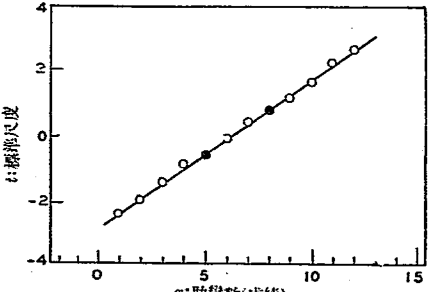
$$r_0\Sigma xf + r_1\Sigma x^2 f = \Sigma xtf,$$

然後由此方程式求之.

f 爲在已給 X 值前一組段內之頻數，故最末組段內之數值不能計入，因與此組段成對應之 t 值爲 ∞ 故也. 於是



a. 直方圖及式(29)第一式之軌跡



b. 論據及式(28)第一式之軌跡

圖 41. 某種考試 514 人成績之分配(論據見表 49 *a*)。

$\Sigma f=511, \Sigma xf=3{,}449, \Sigma x^2f=35{,}953, \Sigma tf=102.9866$ 及 $\Sigma xtf=1842.8498.$

將此等值代入瞬範方程式內解之，我們有

$$r_1= \quad .429232$$

$$r_0=-2.695570$$

當用不加權最小平方法求常數之值時，我們宜先由式(26)寫出瞬範方程式

$$nr_0 \;+ r_1\Sigma x \;=\Sigma t$$

$$r_0\Sigma x+r_1\Sigma x^2=\Sigma xt$$

然後由現在所寫之瞬範方程式求之。

與先前一樣，最末組段裏之頻數亦不宜計入，我們有

$$n=12,\ \Sigma x=78,\ \Sigma x^2=650,\ \Sigma t=1.1112 \text{ 及 } \Sigma xt=69.7478.$$

將此等值代入瞬範方程式內解之，則得

$$r_1= \quad .437238$$

$$r_0=-2.749445$$

故用上面三種方法得到之變換尺度方程式各爲

$$(28)\qquad \begin{cases} t=-2.7210+.4335x, \\ t=-2.6956+.4292x, \\ t=-2.7494+.4372x. \end{cases}$$

由此可見，用三種不同配合法算得常數之值相差甚微，其差別有無重要性容後詳述。現在我們僅發明用定點法所得之結果雖稍遜於用加權最小平方法所得者，但較優於用不加權最小平方法所得者。

c. 頻數曲線方程式　既得變換尺度方程式(28)，由式(10)可寫用此三種配合法所得機率曲線之方程式，該方程式為

$$Y=\frac{.4335}{\sqrt{2\pi}}e^{-\frac{t^2}{2}},$$

$$Y=\frac{.4292}{\sqrt{2\pi}}e^{-\frac{t^2}{2}},$$

$$Y=\frac{.4372}{\sqrt{2\pi}}e^{-\frac{t^2}{2}};$$

即所得之頻數曲線方程式各為

$$(29)\qquad \begin{cases} Y=\dfrac{514\times .4335}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}, \\ Y=\dfrac{514\times .4292}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}, \\ Y=\dfrac{514\times .4372}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}. \end{cases}$$

d. 頻數曲線縱坐標及面積　頻數曲線之縱坐標及面積，可依照表49*b*裏所列之步驟由方程式(29)算得．表49*b*第一行所列者為助變數x；第二行為由(28)首式算得之標準單位t；第三及第五兩行為由常態曲線表(附錄III表B)直接查得之縱坐標及面積；第四行為欲求之縱坐標Y，即為第三行裏各數值乘以Ni'($=222.8190$)得之者；第六行為在x以下之頻數，即第五行裏各數值乘以N($=514$)得之者；並且第六行為第五行裏各數值依次遞減，即由次一組段之頻數減去其前一組段之頻數而得．

表 49b. 曲線式 (29) 第一式縱坐標及其下面積之計算

組距數 x	得算之 t [由式(28)首式]	縱坐標 $\phi(t)$ [由表B查得者]	所求之 Y [由式(29)首式]	面積 $\int_{-\infty}^{t}\phi(t)dt$. [由表B查得者]	計算之頻數 在 x 以下	計算之頻數 在 x 與 $x+1$ 之間
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0	−2.7210	.009844	2.2	.003054	–	5.7
1	−2.2876	.029146	6.5	.011081	5.7	10.7
2	−1.8541	.071522	15.9	.031864	16.4	23.6
3	−1.4206	.145440	32.4	.077717	39.9	43.2
4	−.9872	.245068	54.6	.161775	83.2	65.9
5	−.5537	.342242	76.3	.289895	149.1	83.4
6	−.1202	396070	88.2	.452163	232.4	87.8
7	.3132	.379843	84.6	.622935	320.2	76.8
8	.7467	.301880	67.3	.772375	397.0	55.9
9	1.1802	.198816	44.3	.881040	452.9	33.7
10	1.6136	.108525	24.2	.946691	486.6	17.0
11	2.0471	.049084	10.9	.979675	503.6	7.0
12	2.4806	.018396	4.1	.993442	510.6	2.5
13	2.9140	.005716	1.3	.998216	513.1	.9
總計	...	...	...	...	...	514.1

e. X^2 測驗　方程式配合之完美程度可取 X^2 值當做量度之準則。如果 X^2 之值大，則謂配合之完美程度低；反之則稱做高。X^2 之公式爲

(30)
$$X^2=\sum\frac{(f_0-f_c)^2}{f_c}.$$

式中 f_0 代表 x 與 $x+1$ 間之觀察頻數、f_c 爲與 f_0 成對應之理論頻數，k 爲組數.

表 49c. 測驗配合適度情形 X^2 之計算

x	觀察之頻數 f_0	計算之頻數 f_c	f_0-f_c	$(f_0-f_c)^2$	$(f_0-f_c)^2/f_c$
0	5	5.7	$-.7$	.49	.09
1	9	10.7	-1.7	2.89	.27
2	28	23.6	4.4	19.36	.82
3	49	43.2	5.8	33.64	.78
4	58	65.9	-7.9	64.61	.98
5	82	83.4	-1.4	1.96	.02
6	87	87.8	$-.8$	.64	.01
7	79	76.8	2.2	4.84	.06
8	50	55.9	-5.9	34.81	.62
9	37	33.7	3.3	10.89	.32
10	21	17.0	4.0	16.00	.94
11	6	7.0			
12	3	3.4	-1.4	2.25	.18
總計	514	514.1	$-.1$	…	$5.09=X^2$

表 49c 曉示計算 X^2 之手續. 因為推求 X^2 試驗公式時係假設各組段內之頻數須相當多,故吾人必須將頻數分配兩端頻數之少者合併之,須使每組段裏之理論頻數大於 5. 現在所以將表 49c 末兩組段裏之頻數合併者,即以此故。

自由度 (degrees of freedom) n, 係由組數 k 減去在配合手續中失掉

之自由度得之者.在本例中 $n=9$. 本表自由度之數目本為 13 (因為已給之頻數分配有 13 組). 但各頻數受其和之限制,失去自由度 1; 末兩組頻數要被合併又失去自由度 1; 配合之 (x, t) 曲線內有常數 2, 測定常數之值時需要兩變數值,因此其又失去自由度 2: 合計其失掉 4 個自由度.故只剩 9 個自由度.數值 X^2 及 n 能使吾人求得 P 值,其為 .823. P 值為 .823, 意思就是說, 100 同樣大小的樣本,取自用此曲線代表之宇宙內,算得之 X^2 值必均與此宇宙 X^2 之值有差別, 100 個 X^2 值之中約有 82 個等於或大於現在所得者;或者說如果以常態曲線配合由同一宇宙內抽取之 100 個樣本,即 100 個頻數分配,論及配合適度之情形有 82 個遜於現在之樣本,故現在之樣本以常態曲線配合之相當滿意.

用定點法所求之結果與用加權及不加權縱坐標最小平方法所求者究竟相差情形如何,是值得注意的.用加權縱坐標最小平方法配合算得之 P 值為 .861, 用不加權縱坐標最小平方法算得者 .816, 可見用定點法算得之值(其為 .823) 介於二者之間,蓋用定點法得到常數之值,係選與較大頻數成對應之 (x, t) 點計算者,如此該點並非完全未賦之以權. 定點法需要之計算簡單,得到之結果相當準確,其準確度足敷實際上之需要,故在以後各例中求變換尺度方程式裏各常數之值時悉用定點法.

例 2. 458 男童肺活量量度 (measure of lung capacity) 之分配(頻數分配始端突起之例,圖見下面,論據見參考書 20).

表 50. 458 男童肺活量量度之分配、論據、步驟及結果。

肺活量 (立方英寸) (1)	助變數 x (2)	頻數 觀察的 (3)	頻數 計算的 (4)	標準尺度 觀察的 (5)	標準尺度 計算的 (6)	縱坐標 Y (7)
44.5 –	0	–	–	$-\infty$	$-\infty$	–
59.5 –	1	2	2.0	–2.622	–2.620	7.4
74.5 –	2	17	15.6	–1.734	–1.768	23.2
89.5 –	3	24	28.4	–1.317	–1.275	32.9
104.5 –	4	30	36.0	–.997	–.918	38.1
119.5 –	5	56	40.7	–.578	–.624	41.1
134.5 –	6	35	42.0	–.364	–.363	42.6
149.5 –	7	38	42.5	–.148	–.122	43.1
164.5 –	8	57	42.6	.165	.114	42.5
179.5 –	9	34	42.0	.358	.348	40.5
194.5 –	10	29	39.2	.533	.587	37.0
209.5 –	11	43	34.0	.831	.831	32.2
224.5 –	12	35	28.9	1.142	1.084	26.4
239.5 –	13	21	23.6	1.400	1.350	19.9
254.5 –	14	15	17.0	1.664	1.626	13.7
269.5 –	15	9	11.0	1.905	1.916	8.6
284.5 –	16	6	6.5	2.163	2.220	4.8
299.5 –	17	5	3.5	–	2.539	2.4
314.5 –	18	–	1.6	2.622	2.874	1.0
329.5 –	19	2	.7	–	3.223	.4
總計	–	458	457.8	–	–	–

X^2 測驗: $X^2=25.3$, $n=11$, $P=.008$.

變換式: 變換式之形式爲

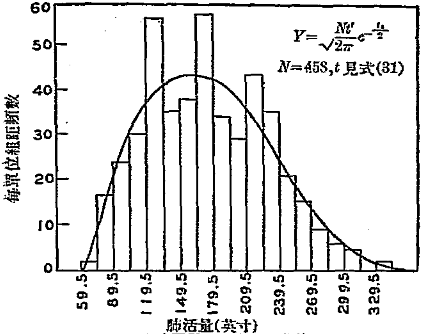
(31) $$t=\log x^{a}+r_0+r_1x+r_2x^2.$$

所取之值: $x=1$, $t_1=-2.620$, $t_3=.831$,

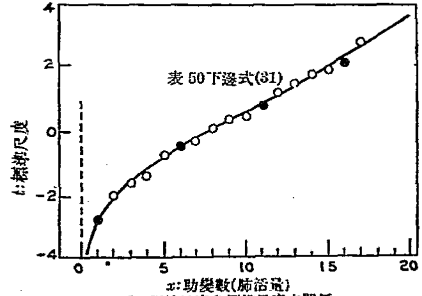
$h=5$, $t_2=-.364$, $t_4=2.220$,

所求常數之值: $a=3.028504$, $r_1=-.089903$,

$r_2=.009968$, $r_0=-2.540065$.



a. 直方圖及配合之頻數曲線



b. 原始尺度與標準尺度之關係

圖 42. 458 男童肺活量量度之分配（論據見表50，

例 3. 458 男童骨化比量度 (measure of ossification ratio) 之分配(末端立止頻數分配之例，圖見下面，論據見參考書 20).

表 51. 458 男童骨化比量度之分配.論據、步驟及結果

骨化比	助變數 x	頻數 f		標準尺度		縱坐標 Y
		觀察的	計算的	觀察的	計算的	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
.245 –	0	–	–	–	–4.151	–
.295 –	1	1	.2	–2.850	–3.396	.4
.345 –	2	1	1.1	–2.622	–2.771	2.2
.395 –	3	11	4.2	–1.905	–2.261	6.5
.445 –	4	5	9.3	–1.759	–1.850	12.1
.495 –	5	7	14.6	–1.602	–1.522	16.7
.545 –	6	15	18.1	–1.357	–1.262	19.1
.595 –	7	27	19.5	–1.053	–1.053	19.7
.645 –	8	20	19.8	– .878	– .880	19.9
.695 –	9	21	20.2	– .720	– .727	20.8
.745 –	10	21	22.0	–.578	–.578	23.6
.795 –	11	25	26.1	–.423	–.416	29.0
.845 –	12	27	33.0	–.266	–.226	37.6
.895 –	13	50	42.9	.011	.011	43.9
.945 –	14	50	54.0	.289	.312	59.0
.995 –	15	69	61.5	.720	.696	62.1
1.045 –	16	54	57.5	1.186	1.186	50.0
1.095 –	17	33	38.0	1.686	1.814	24.9
1.145 –	18	19	14.1	2.622	2.649	5.4
1.195 –	19	2	1.8	∞	∞	–
總計	–	458	457.9	–	–	–

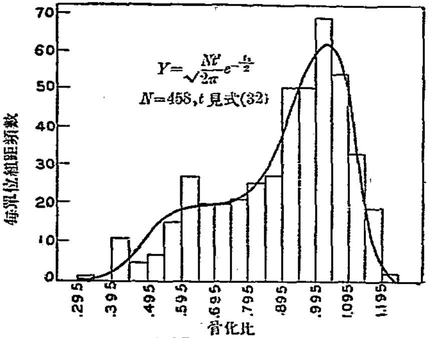
X^2 測驗: $X^2=10.13,\quad n=7,\quad P=.182.$

變換式: 變換式之形式為

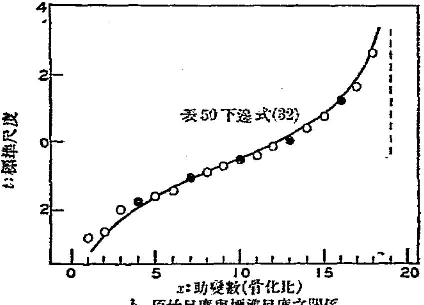
(32) $t=\log(b-x)^{\beta}+r_0+r_1x+r_2x^2+r_3x^3.$

所取之值: $x_0=4,\quad t_1=-1.850,\quad t_4=.011,$
$h=3,\quad t_2=-1.053,\quad t_5=1.186.$
$b=19,\quad t_3=-.578,$

所求常數之值: $\beta=-.710564,\quad r_1=.809073,$
$r_2=-.073384,\quad r_3=.002590,$
$r_0=-3.242216.$



a. 直方圖及配合之頻數曲線



b. 原始尺度與標準尺度之關係

圖 43．458 男童骨化比量度之分配。

例 4. 格林威池 (Greedwich) 1890 至 1904 年 (1901 年除外) 七月陰雲疊度 (degree of cloudness) 之分配 (U 形頻數分配之例，圖見下面，論據見參考書 45)。

表 52. 格林威池陰雲疊度之分配、論據、步驟及結果

陰雲疊度	助變數 x	頻數 f		標準尺度 t		縱坐標 Y
		觀察的	計算的	觀察的	計算的	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
−.5 −	0	−	−	$-\infty$	$-\infty$	∞
.5 −	1	320	320.0	−.8905	−.8905	175.5
1.5 −	2	129	130.3	−.6378	−.6355	99.3
2.5 −	3	74	81.1	−.5102	−.4960	68.2
3.5 −	4	68	59.5	−.3999	−.3999	53.4
4.5 −	5	45	49.2	−.3296	−.3230	46.8
5.5 −	6	45	46.3	−.2609	−.2525	46.8
6.5 −	7	55	49.6	−.1785	−.1785	53.6
7.5 −	8	65	60.7	−.0827	−.0889	70.4
8.5 −	9	90	87.0	.0490	.0383	109.4
9.5 −	10	148	155.2	.2685	.2685	227.4
10.5 −	11	676	676.1	∞	∞	∞
總 計	−	1715	1715.0	−	−	−

X^2 測驗： $X^2 = 3.56$, $n = 5$, $P = .616$.

變換式： 變換式之形式為

(33) $$t = \log x^{\alpha}(b-x)^{\beta} + r_0 + r_1 x.$$

所取之值： $x_0 = 1$, $t_1 = -.8905$,

$l = 3$, $t_2 = -.3999$,

$b = 11$, $t_3 = -.1785$,

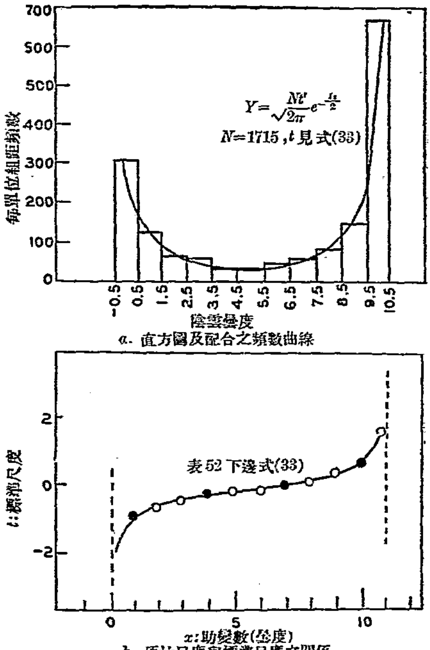
$t_4 = .2685$.

所求常數之值： $\beta = -.864545$,

$\alpha = .962050$,

$r_1 = -.074177$,

$r_0 = .048222$.



a. 直方圖及配合之頻數曲線

b. 原始尺度與標準尺度之關係

圖 44. 格林威池1890—1904年七月陰雲量度之分配(論據見表52)

例 5. 南安普敦氣壓高度 (barometric height) 之分配始端與 X 軸成高度相切(頻數分配之例,圖見下面,論據由參考書 35.)

表 53. 南安普敦氣壓高度之分配.論據、步驟及結果

氣壓高度 (英寸) (1)	助變數 x (2)	頻數 f 觀察的 (3)	頻數 f 計算的 (4)	標準尺度 t 觀察的 (5)	標準尺度 t 計算的 (6)	縱坐標 Y (7)
28.35 –	0	0	…	–	–3.3354	.5
28.45 –	1	2	1.9	–3.2013	–3.2232	.8
28.55 –	2	1	1.0	–3.0823	–3.0978	1.3
28.65 –	3	–	1.7	–	–2.9587	2.1
28.75 –	4	3	2.8	–2.8699	–2.8055	3.7
28.85 –	5	7	4.9	–2.6160	–2.6376	6.3
28.95 –	6	7.5	8.4	–2.4565	–2.4550	10.9
29.05 –	7	15.5	14.4	–2.2470	–2.2575	18.7
29.15 –	8	23.5	24.7	–2.0163	–2.0150	31.7
29.25 –	9	50	41.3	–1.7808	–1.8178	52.4
29.35 –	10	65	67.1	–1.5572	–1.5760	83.8
29.45 –	11	98.5	104.9	–1.3199	–1.3200	121.4
29.55 –	12	150.5	156.0	–1.0581	–1.0502	170.7
29.65 –	13	201	218.1	–.7936	–.7673	251.3
29.75 –	14	288	253.8	–.4891	–.4717	314.3
29.85 –	15	339.5	339.8	–.1803	–.16[illegible]0	360.5
29.95 –	16	382.5	370.3	.1491	.1550	373.8
30.05 –	17	395.5	363.8	.5093	.4848	347.2
30.15 –	18	315	318.9	.8507	.8247	286.1
30.25 –	19	237	247.2	1.1934	1.1742	207.2
30.35 –	20	140.5	168.0	1.4888	1.5327	130.7
30.45 –	21	107.5	99.2	1.8594	1.9000	71.2
30.55 –	22	52.5	50.5	2.2110	2.2757	33.2
30.65 –	23	30.5	22.0	2.7391	2.6598	13.2
30.75 –	24	4	8.1	2.9271	3.0522	4.4
30.85 –	25	4	2.5	3.3960	3.4531	1.2
30.95 –	26	1	.6	–	3.8631	.3
總 計	–	2922.0	2922.1	–	–	–

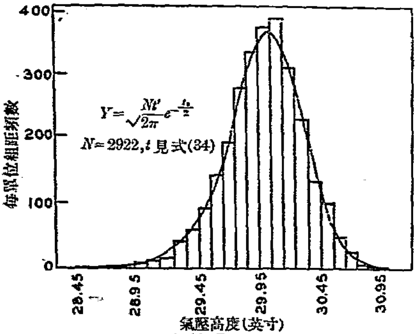
X^2 測驗: $X^2=17.89$, $n=15$, $P=.269$.

變換式: 變換式之形式為

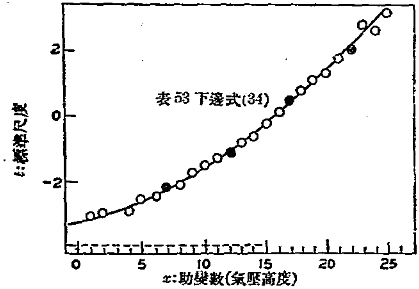
(34) $\log(t-c)=r_0+r_1x+r_2x^2+r_3x^3$.

所取之值: $x_0=6$, $h=5$, $c=-4$, $t_1=-2.455$, $t_2=-1.320$, $t_3=.155$, $t_4=1.900$.

所求常數之值: $r_3=.000014$, $r_1=.069208$, $r_2=-.011443$, $r_0=-.177445$.



a. 直方圖及配合之頻數曲線



b. 原始尺度與標準尺度之關係

圖 45. 南安普敦氣壓高度之分配(論據見表53)

例 6. 某種頻數分配(始端與X軸成高度相切J形頻數分配之例。圖見下面，論據由參考書 11.)

表 54. 某種頻數分配.論據、步驟及結果

單位	助變數 x	頻數 f		標準尺度 t		縱坐標 Y
		觀察的	計算的	觀察的	計算的	
(1)	(2)	(3)	(4)	(5)	(9)	(7)
-.5 -	0	-	-	-	-3.6989	-
.5 -	1	1	-	-3.6685	-3.6110	.5
1.5 -	2	1	.6	-3.4873	-3.5044	.8
2.5 -	3	1	1.1	-3.3773	-3.3771	1.5
3.5 -	4	1	2.1	-3.2974	-3.2278	2.9
4.5 -	5	2	4.1	-3.1814	-3.0552	3.7
5.5 -	6	9	8.2	-2.9059	-2.8589	11.4
6.5 -	7	17	16.7	-2.6601	-2.6389	23.3
7.5 -	8	36	33.9	-2.3954	-2.3954	47.3
8.5 -	9	71	68.3	-2.1208	-2.1286	94.5
9.5 -	10	132	134.7	-1.8373	-1.8373	184.0
10.5 -	11	266	258.4	-1.5098	-1.5172	349.4
11.5 -	12	480	487.6	-1.1545	-1.1546	660.1
12.5 -	13	982	943.0	- .6934	- .7087	1317.8
13.5 -	14	2028	2066.7	- .0211	- .0211	2471.2
14.5 -	15	4165	4165.3	+∞	+∞	∞
總計	-	8192	8190.7	-	-	-

X^2 測驗: $X^2=2.42$, $n=4$, $P=.661$.

變換式: 變換式之形式為

(35) $\log(t-c)=\log(b-x)^{\beta}+r_0+r_1x+r_2x^2$.

所取之值: $x_0=8$, $t_1=-2.3954$,

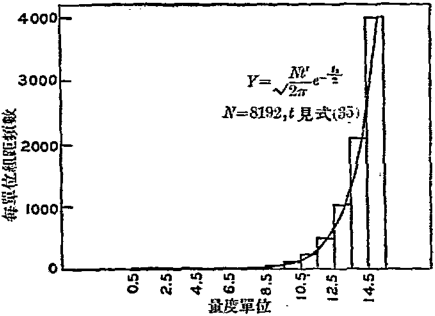
$h=2$, $t_2=-1.8373$,

$b=15$, $t_3=-1.1545$,

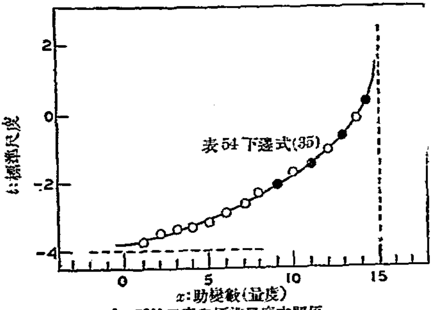
$c=-4$, $t_4=-.0211$.

所求常數之值: $\beta=-.205563$, $r_1=.108367$

$r_2=-.003254$, $r_0=-.279592$.



a. 直方圖及配合之頻數曲線



b. 原始尺度與標準尺度之關係

圖 46. 某種頻數分配(論據見表 54).

例 7. 458 男童胸圍量度 (measure of chest girth) 之分配 (末端與 X 軸成高度相切頻數分配之例，圖見下面，論據見參考書 20.)

表 55. 458 男童胸圍量度之分配.論據、步驟及結果

胸圍 (英寸) (1)	助變數 x (2)	頻數 f 觀察的 (3)	頻數 f 計算的 (4)	標準尺度 t 觀察的 (5)	標準尺度 t 計算的 (6)	縱坐標 Y (7)
19.95 –	0	–	…	–	–6.0	–
20.95 –	1	1	.1	–2.850	–3.599	.5
21.95 –	2	4	4.9	–2.393	–2.293	12.8
22.95 –	3	23	23.6	–1.545	–1.334	32.9
23.95 –	4	39	37.0	–1.052	–1.066	39.0
24.95 –	5	37	28.5	–.748	–.748	37.2
25.95 –	6	34	36.1	–.521	–.507	35.7
26.95 –	7	34	26.4	–.317	–.291	37.4
27.95 –	8	44	39.5	–.071	–.071	41.7
28.95 –	9	45	43.7	.176	.169	46.1
29.95 –	10	43	47.5	.423	.442	47.4
30.95 –	11	49	45.6	.741	.740	43.2
31.95 –	12	42	38.9	1.091	1.059	33.8
32.95 –	13	30	28.2	1.461	1.384	22.6
33.95 –	14	14	17.7	1.734	1.700	13.2
34.95 –	15	8	9.8	1.977	1.991	6.9
35.95 –	16	4	4.9	2.163	2.243	3.4
36.95 –	17	1	2.4	2.223	2.450	1.6
37.95 –	18	4	1.2	2.622	2.606	.8
38.95 –	19	–	.6	–	2.715	.4
39.95 –	20	–	.3	–	2.783	.2
40.95 –	21	–	.1	–	2.821	.1
41.95 –	22	1	.1	2.622	2.839	–
42.95 –	23	1	–	2.850	2.847	–
總計	–	458	457.1		–	–

X^2 測驗: $X^2=3.36$, $n=7$, $P=.847$

變換式: 變換式之形式為

(36) $\log(d-t)=r_0+r_1x+r_2x^2+r_3x^3+r_4x^4+r_5x^5$

所取值: $x=2$, $t_1=-2.293$, $t_4=.741$,

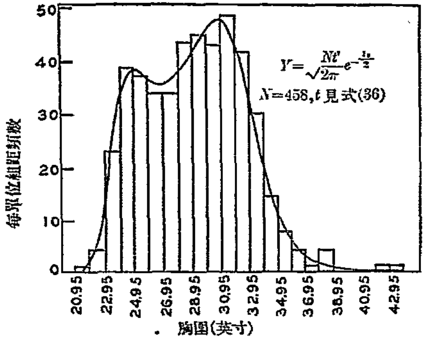
$h=3$, $t_2=-.748$, $t_5=1.700$,

$d=2.850$, $t_3=-.071$, $t_6=2.450$.

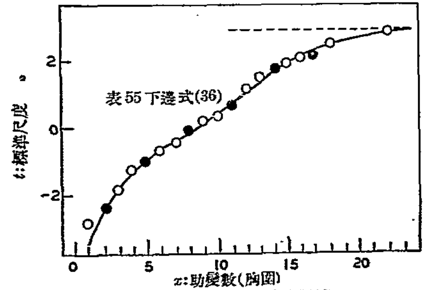
所求常數之值: $r_5=-.000002$, $r_2=.026080$,

$r_4=.000087$, $r_1=-.161959$,

$r_3=-.002266$, $r_0=.947597$.



a. 直方圖及配合之頻數曲線



b. 原始尺度與標準尺度之關係

圖 47. 458 男童胸圍量度之分配(論據見表 55).

例 8. 英國 1915 年某種財產年値之分配（末端與 X 軸成高度相切 J 形頻數分配之例，圖見下面，論據見參考書 45.）

表 56. 某種不動產之年値之分配、論據、步驟及結果

年値（百鎊）(1)	助變數 x (2)	頻數 f 觀察的 (3)	頻數 f 計算的 (4)	標準尺度 t 觀察的 (5)	標準尺度 t 計算的 (6)	縱坐標 Y (7)
−2−	0	…	…	$-\infty$	$-\infty$	$+\infty$
0−	1	2006.5	2006.5	.8793	.8793	336.9
2−	2	227.5	213.0	1.2946	1.2614	137.1
4−	3	89	101.1	1.5399	1.5321	74.0
6−	4	55	57.4	1.7556	1.7556	43.9
8−	5	39.5	34.7	1.9840	1.9 01	27.0
10−	6	21	21.6	2.1663	2.1240	17.0
12−	7	7.5	13.7	2.2534	2.2800	10.9
14−	8	11	8.8	2.4241	2.4205	7.0
16−	9	6	5.7	2.5589	2.5469	4.6
18−	10	−	3.8	−	2.6600	3.1
20−	11	5	2.5	2.7234	2.7608	2.1
22−	12	2	1.7	2.8171	2.8·99	1.4
24−	13	−	1.2	−	2.9283	1.0
26−	14	2	.8	2.9148	2.9967	.7
28−	15	−	.6	−	3.0560	.5
30−	16	1	.4	3.0338	3.1070	.4
32−	17	−	.3	−	3.1505	.3
34−	18	−	.2	−	3.1873	.2
36−	19	−	.2	−	3.2182	.2
38−	20	1	.1	3.1529	3.2440	.1
40−	21	−	.1	−	3.2654	.1
42−	22	−	.1	−	3.2829	.1
44−	23	1	.1	3.3·00	3.2971	.1
46−	24	−	−	−	3.3087	−
48−	25	1	−	−	3.3179	−
總計	−	2476.0	2474.6	−	−	−

X^2 測驗：　　$X^2=6.61$,　$n=4$,　$P=.161$.

變換式：變換式之形式爲

(37)　　$\log(d-t)=a\log x+r_0+r_1x+r_2x^2$.

所取之値：　$d=3.35$,　　$t_1=.8793$,　　$t_4=2.6600$.

　　$x_0=1$,　　$t_2=1.7556$,

　　$h=3$,　　$t_3=2.2800$,

所求常數之値：　$a=-.126740$,　　$r_1=-.030058$,

　　$r_2=-.001583$,　　$r_0=-.424461$.

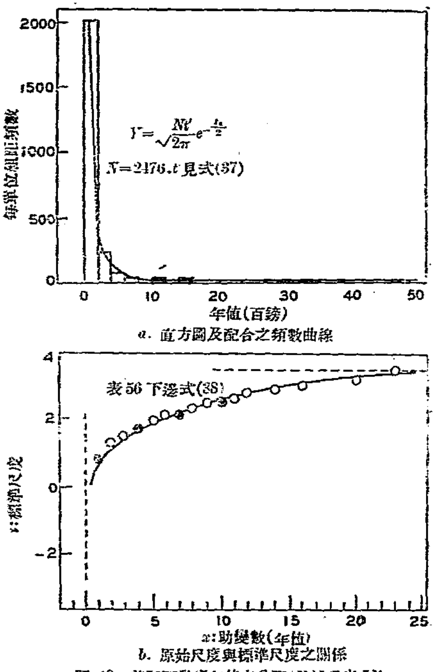


圖 48. 某種不動產年值之分配(論據見表56)

例 9. 在某種經驗中病人年齡之分配(兩端與 X 軸成高度相切頻數分配之例，圖見下面，論據由參考書 21.)

表 57. 病人年齡之分配、論據、步驟及結果

年 齡	助變數 x	在 x 與 $x+1$ 間之頻數 f		標準尺度 t		縱坐標 Y
		觀察的	計算的	觀察的	計的算	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
−2.5−	0	−	...	$-\infty$	−3.2188	3.3
2.5−	1	10	11.4	−3.0642	−3.0237	8.7
7.5−	2	13	15.3	−2.8051	−2.7461	24.8
12.5−	3	41	44.3	−2.4578	−2.4198	71.7
17.5−	4	115	124.0	−2.0631	−2.0775	193.8
22.5−	5	326	310.0	−1.5967	−1.5967	453.6
27.5−	6	675	652.2	−1.1316	−1.1435	874.1
32.5−	7	1113	1118.5	− .6729	− .6789	1357.0
37.5−	8	1528	1545.4	− .2085	− .2085	1690.1
42.5−	9	1692	1720.6	.2592	.2673	1687.1
47.5−	10	1530	1535.9	.7368	.7494	1338.5
52.5−	11	1122	1087.4	1.2370	1.2370	831.0
57.5−	12	610	601.7	1.7347	1.7244	399.3
62.5−	13	255	259.6	2.2102	2.1985	150.3
67.5−	14	86	89.8	2.6395	2.6395	46.7
72.5−	15	26	26.6	3.0091	3.0248	13.2
77.5−	16	8	7.5	3.3281	3.3361	3.8
82.5−	17	2	2.2	3.5175	3.5657	1.2
87.5−	18	1	.7	3.6975	3.7190	.4
92.5−	19	1	.3	−	3.8110	.2
總 計	−	9154	9153.4	−	−	−

X^2 測驗: $X^2=5.47,\quad n=9,\quad P=.797.$

變換式: 變換式之形式為

$$\log\frac{d-t}{t-c}=r_0+r_1x+r_2x^2+r_3x^3. \tag{38}$$

所取之值: $x_0=5,\quad t_1=-1.5967,$

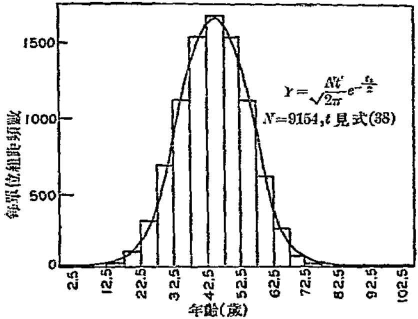
$h=3,\quad t_2=-0.2058,$

$c=-3.5,\quad t_3=1.2370,$

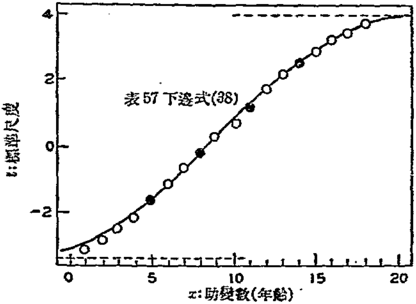
$d=3.9,\quad t_4=2.6395.$

所求常數之值: $r_3=-.000672,\quad r_1=-.257390,$

$r_2=.017130,\quad r_0=1.403344.$



a. 直方圖及配合之頻數曲線



b. 原始尺度與標準尺度之關係

圖 49. 在某種經驗中病人年齡之分配(論據見表 57)

問題 IV

1. 驗算表 49*a* 各行裏數值有無錯誤.

2. 驗算表 49*b* 各行裏數值有無錯誤.

3. 證明用加權最小平方法求得式 (27) 裏之 r_1 值等於標準差之反商 $\frac{1}{\sigma}$, 即是證明式 (29) 之第二式可寫做 $Y=\frac{514}{\sigma\sqrt{2\pi}}e^{-t^2/2}$.

4. 試將例 2 裏之論據繪一在標準尺度上之直方圖.

5. 試將例 3 裏之論據繪一在標準尺度上之直方圖.

6. 試將例 4 裏之論據繪一在標準尺度上之直方圖.

7. 試將例 5 裏之論據繪一在標準尺度上之直方圖.

8. 試將例 6 裏之論據繪一在標準尺度上之直方圖.

9. 試將例 7 裏之論據繪一在標準尺度上之直方圖.

10. 試將例 8 裏之論據繪一在標準尺度上之直方圖.

11. 試將例 9 裏之論據繪一在標準尺度上之直方圖.

12. 計算例 2 變換方程式裏各常數之值.

13. 計算例 3 變換方程式裏各常數之值.

14. 計算例 4 變換方程式裏各常數之值.

15. 計算例 5 變換方程式裏各常數之值.

16. 計算例 6 變換方程式裏各常數之值.

17. 計算例 7 變換方程式裏各常數之值.

18. 計算例 8 變換方程式裏各常數之值.

19. 計算例 9 變換方程式裏各常數之值.

20. 驗算表 50 各行裏之數值有無錯誤.

21. 驗算表 51 各行裏之數值有無錯誤.

22. 驗算表 52 各行裏之數值有無錯誤.

23. 驗算表 53 各行裏之數值有無錯誤.

24. 驗算表 54 各行裏之數值有無錯誤.

25. 驗算表 55 各行裏之數值有無錯誤.

26. 驗算表 56 各行裏之數值有無錯誤.

27. 驗算表 57 各行裏之數值有無錯誤.

28. 驗算例 1 至 9 裏 X^2, n 及 P 之值有無錯誤.

29. 用變換尺度法配合第四章第九節裏所列之論據.

30. 用變換變數法配合第四章第十節裏所列之論據.

31. 用變換變數法配合第四章第十一節裏所列之論據.

32. 用變換變數法配合第四章第十二節裏所列之論據.

33. 用變換變數法配合第四章第十三節裏所列之論據.

34. 用變換變數法配合第四章第十四節裏所列之論據.

35. 用變換變數法配合第四章第十五節裏所列之論據.

第七章　相關

第一節　相關之意義及種類

關於一個變數之頻數分配等問題已在第二、第三、第四及第六章討論過了．現在我們進一步研究兩個及多個變數共同變化之各有關問題．統計學上最重要而且最有用之技術，為兩個及多個有聯繫變數之分析．許多問題非用此法分析莫辯．表顯二數及多數共同變化事實，以及度量其相關程度種種方法，均係英國生物學家葛爾同及皮爾生所首創（參閱第一章第二節）．

相關之種類可略別之為：直線相關 (linear correlation)、非直線相關 (non-linear correlation)、複相關 (multiple correlation)、淨相關 (net correlation, partial correlotion)、等級相關 (rank correlation)、品質相關 (association of attributes) 等．我們將在以下兩節裏敍述直線及非直線相關之意義，討論與此兩種相關之有關各問題；關於二數相關之基本理論則在第四節闡明之；等級、品質等相關統以其他相關名之，其特殊性質及有關各問題將在第五節論述之．至於複相關及淨相關之意義等，將分別在第六及第七兩節討論之．

在某種研究中，常常遇到兩組有聯繫之觀察值，結果生成 N 雙對應值 (X_i, Y_i)，$i=1, 2, 3, \cdots, N$．即如若以 X 以代表七月的雨量，則 Y 可代表某區穀類之平均產量；若以 X 代表物價指數，則 Y 可代表同時期之工人指數；或者研究學校兒童之發育情形，量度每人之身長與體重，如是 X

代表每個人之身長，Y代表其體重；或者我們用十二枚骰子連續擲之，擲完第一次後留五枚骰子於桌上（譬如說）不動，再擲第二次，研究每相連兩次所得么點之頻數分配．則第一次投擲十二枚骰子所得么點之數可以 X 代表之，第二次投擲七枚骰子和其他留滯未動的五枚共得么點之數可用 Y 代表之．下列之表 58 及 59 即二數相關表之例．表 58 裏所列者爲體重與身長之相關，並且表 59 爲出生率與死亡率之相關．

當已給之成雙數値用點代表時，則在坐標系上得到一羣（X, Y）點，此點代表之圖形（圖 50）即所謂之散佈圖（scatter diagram）．若散佈圖中各點之趨勢成一直線形式者，則稱此二數之相關曰直線相關；其趨勢成一曲線形式者曰非直線相關．直線或非直線二變數相關有時總稱之爲單相關（simple correlation）．

表 58. 體重與身長之相關（論據見參考書 28）．

體重(磅) X	身長(吋) Y	體重(磅) X	長身(吋) Y
140	67	127	61
120	62	206	75
191	73	123	62
140	65	137	64
172	71	187	69
190	69	145	65
175	68	197	74
210	75	185	72
132	64	157	66
190	72	194	73
125	61	142	63
192	70	190	71
200	74	142	67
136	63	168	68
162	66	188	70

表 59. 出生率與死率之相關(論據見參考書 45).

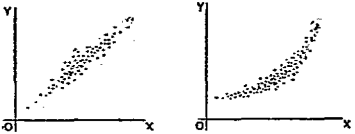
年度	出生率	死亡率
1876－1880	35.3	20.8
1881－1885	33.5	19.4
1886－1890	31.4	18.9
1891－1895	30.5	18.7
1896－1900	29.3	17.7
1901－1905	28.2	16.0
1906－1910	26.3	14.7
1911－1915	23.6	14.3
1916－1920	20.1	14.4
1921－1925	19.9	12.2
1926－1930	16.7	12.1

第二節　直線相關

1. 相關問題　現在我們要思考之第一個問題爲如何測定在直線相關情形中 X 與 Y 間之聯繫,卽測定 X 與 Y 之相關程度.相關程度有兩個極端;一爲完全相關 (perfect correlation),一爲零相關(zero correlation).通常之相關程度介乎此二極端相關之間.第二個問題爲如何求出聯繫兩個變數之趨勢直線,並且指定變數之一如何由此直線以估計其他變數之平均值.我們在本節裏以此二問題爲研究之中心,其他問題則附帶研究之.

2. 相關係數　由圖 50 可見,X 值在指定之一小段 ΔX 內,其對應 Y 值之變動甚大.所謂正相關 (positive correlation) 者,如果指定之

X 值比 $\bar{X}$ 大，其對應之平均 Y 值比 $\bar{Y}$ 大；並且如果指定之 X 值比 $\bar{X}$ 小，其對應之平均 Y 值比 $\bar{Y}$ 小。從他方面言之，當 X 值增加時，Y 值或有減少之趨勢。在此情形中，指定 X 大於 $\bar{X}$，其對應之平均 Y 值小於 $\bar{Y}$；並且指定 X 小於 $\bar{X}$，其對應 Y 值之平均數大於 $\bar{Y}$，X 與 Y 之關係如此者稱之曰負相關 (negative correlation)。如果指定 X 大於 $\bar{X}$，並且對應之 Y 值大於 $\bar{Y}$ 者與小於 $\bar{Y}$ 者相等，X 與 Y 之關係如此者即上面所稱之曰零相關。



a. 直線相關　　*b* 非直線相關

圖 50　散佈圖。

在圖 51 中有兩組坐標系：一組坐標系以 OX 及 OY 爲軸，以 O 爲原點；一組坐標系以 $O'x$ 及 $O'y$ 爲軸，以 O' 爲原點並且 O' 點之坐標爲 $(\bar{X}, \bar{Y})$。於是散佈圖之點分配於 xy 平面的四個象限 (quadrant) 中，在此四個象限中各點坐標之符號如下：

在第一象限中，x 及 y 皆爲正；

在第二象限中，x 爲負 y 爲正；

在第三象限中，x 及 y 皆爲負；

在第四象限中，x 爲正 y 爲負。

如此，在第一及第三象限中，各點坐標乘積 xy 之值爲正；在第二及第四

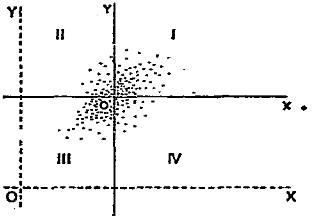


圖 51. 散佈圖與象限。

象限中，各點坐標乘積 xy 之値為負．此等乘積之代數和能將象限內各點之分配情形描寫出來．若此和為正，點之趨勢經過第一及第三象限；若為負，點之趨勢經過第二及第四象限；並且若為零，各點平均分配於四個象限內，卽乘積 xy 之正者與負者相抵．很自然的，將各乘積 xy 加之，並以 N 除之，除得之商可做為二量相關之量度．再進一步說，如果我們起始各以其標準單位表示 x 及 y，可得與原來尺度無關之相關量度．如此用符號表之，我們有

$$r=\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\sigma_X}\right)\left(\frac{y_i}{\sigma_Y}\right).$$

此卽

(1)
$$r=\frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i-\bar{X}}{\sigma_X}\right)\left(\frac{Y_i-\bar{Y}}{\sigma_Y}\right).$$

彼處 r 有時稱之為總相關係數 (coefficient of total correlation)，更有時稱之為積差相關係數 (product moment coefficient of correlation)，通常稱之為相關係數 (correlation coefficent).

於是我們可將相關係數予以如下之定義:

兩組變數各以其標準為單位表之者,其相關係數為各對應離均差乘積之算術平均數.

3. r之其他形式 用式(1)解釋相關係數之由來是很方便的.但因求 $\left(\frac{X_i-\bar X}{\sigma_X}\right)\left(\frac{Y-\bar Y}{\sigma_Y}\right)$ 之值時須做許多除法,式(1)在實際運算上則不適用;因此將其推演一下,以應需要.我們知道 σ_X 及 σ_Y 均為常數,故式(1)可寫做下列之形式:

(2) $$r=\frac{\frac{1}{N}\Sigma(X-\bar X)(Y-\bar Y)}{\sigma_X\sigma_Y}.$$

用此式求 r 之值只須做最後一次除法,故此式實較式(1)為優.但 $\bar X$ 及 $\bar Y$ 之值往往帶有不盡小數,故當計算數值 $\Sigma(X-\bar X)(Y-\bar Y)$ 時仍感麻煩;於是算式 $\Sigma(X-\bar X)(Y-\bar Y)$ 有再行簡化之必要.因為

$$\Sigma(X-\bar X)(Y-\bar Y)=\Sigma XY-\bar X\Sigma Y-\bar Y\Sigma X+\bar X\bar Y$$
$$=\Sigma XY-\bar X\bar Y.$$

所以我們最後得寫

(3) $$r=\frac{\nu_{XY}-\nu_{1X}\nu_{1Y}}{\sqrt{\nu_{2X}-\nu_{1X}^2}\sqrt{\nu_{2Y}-\nu_{1Y}^2}},$$

式中 $\nu_{1X}=\frac{1}{N}\Sigma X$, $\nu_{2X}=\frac{1}{N}\Sigma X^2$, $\nu_{1Y}=\frac{1}{N}\Sigma Y$, $\nu_{2Y}=\frac{1}{N}\Sigma Y^2$, $\nu_{XY}=\frac{1}{N}\Sigma XY$.

定理 1. 相關係數與原點所在之位置及所用之單位無關.

證: 設

(4) $$x=\frac{X-X_0}{h},\quad y=\frac{Y-Y_0}{k}.$$

於是 $\nu_{1X}=X_0+h\nu_{1x}$, $\nu_{2X}=\frac{1}{N}\Sigma(X_0+hx)^2$, $\nu_{1Y}=Y_0+k\nu_{1y}$, $\nu_{2Y}=\frac{1}{N}\Sigma(Y_0+ky)^2$, 並且 $\nu_{XY}=\frac{1}{N}\Sigma(X_0+hx)(Y_0+ky)$.

將各式代入式(3)內並化簡之,我們有

$$(5) \qquad r=\frac{\nu_{xy}-\nu_{1x}\nu_{1y}}{\sqrt{\nu_{2x}-\nu_{1x}^2}\sqrt{\nu_{2y}-\nu_{1y}^2}},$$

彼處 $\nu_{1x}=\frac{1}{N}\Sigma x$, $\nu_{2x}=\frac{1}{N}\Sigma x^2$, $\nu_{1y}=\frac{1}{N}\Sigma y$, $\nu_{2y}=\frac{1}{N}\Sigma y^2$ 並且 $\nu_{xy}=\frac{1}{N}\Sigma xy$.

由末式可見 r 與常數 X_0, Y_0, h 及 k 無關. 證訖.

此式與式(4)形式完全一樣,不過前式中之 X 及 Y(大寫字母)易之以 x 及 y(小寫字母)而已.當已給 X 或 Y 值之位數相當多時,用式(5)計算 r 較用式(4)為便.

4. 舉例 為說明用公式(5)計算 r 之手續起見,舉例如下:

例 1. 試根據下表裏所列之論據計算相關係數 r.

表 60. 未分組論據相關係數之計算(論據見參考書25).

月份	貸款(萬萬) X	物價指數 Y	x $(=X-5.0)$	y $(=Y-250)$	x^2	y^2	xy
1	5.33	248	.33	−2	.1089	4	−0.66
2	5.67	248	.67	−2	.4489	4	−1.34
3	5.65	243	.65	−7	.4225	49	−4.55
4	5.56	249	.56	−1	.3136	1	−0.56
5	5.53	235	.53	−15	.2809	225	−7.95
6	5.28	265	.28	15	.0784	225	4.20
7	5.77	282	.77	32	.5929	1024	24.64
8	6.02	303	1.02	53	1.0404	2809	54.06
9	6.35	290	1.35	40	1.8225	1600	54.00
10	6.80	230	1.80	−20	3.2430	400	−36.00
11	4.88	201	−.12	−49	.0144	2401	5.88
12	3.45	206	−1.55	−44	2.4025	1936	68.20
總計	…	…	6.29	0	10.7659	10678	159.92

計算:　$\nu_{1x}=.524167,\qquad \nu_{1y}=0.000000,$

$\nu_{2x}=.897158,\qquad \nu_{2y}=889.833333,$

$\nu_{xy}=13.326667;$

$\sigma_x=.788924,\qquad \sigma_y=29.830074,$

$\sigma_x\sigma_y=23.533661.$

結果:　$r=0.57.$

5. 迴歸線 r 之性質可用配合散佈圖之直線研究之. 當配合直線時以使各點至此直線縱距離剩餘數平方和最小為準則,並且 $Y=a+bX$ 為該配合之直線;則式中之 a 及 b 可分別用下式計算之:

$$a=\frac{\begin{vmatrix}\Sigma Y & \Sigma X\\ \Sigma XY & \Sigma X^2\end{vmatrix}}{\begin{vmatrix}N & \Sigma X\\ \Sigma X & \Sigma X^2\end{vmatrix}},\qquad b=\frac{\begin{vmatrix}N & \Sigma Y\\ \Sigma X & \Sigma XY\end{vmatrix}}{\begin{vmatrix}N & \Sigma X\\ \Sigma X & \Sigma X^2\end{vmatrix}}.$$

因為

$$\frac{\begin{vmatrix}\Sigma Y & \Sigma X\\ \Sigma XY & \Sigma X^2\end{vmatrix}}{\begin{vmatrix}N & \Sigma X\\ \Sigma X & \Sigma X^2\end{vmatrix}}=\frac{\Sigma X^2\Sigma Y-\Sigma X\Sigma XY}{N\Sigma X^2-(\Sigma X)^2},$$

故

$$\frac{\begin{vmatrix}\Sigma Y & \Sigma X\\ \Sigma XY & \Sigma X^2\end{vmatrix}}{\begin{vmatrix}N & \Sigma X\\ \Sigma X & \Sigma X^2\end{vmatrix}}=\frac{\nu_{2X}\nu_{1Y}-\nu_{1X}\nu_{XY}}{\nu_{2X}-\nu_{1X}^2}$$

$$=\frac{(\nu_{2X}\nu_{1Y}-\nu_{1X}^2\nu_{1Y})-(\nu_{1X}\nu_{XY}-\nu_{1X}^2\nu_{1Y})}{\sigma_X^2}$$

$$=\frac{\nu_{1Y}\sigma_X^2-\nu_{1X}(\nu_{XY}-\nu_{1X}\nu_{1Y})}{\sigma_X^2}$$

$$=\nu_{1Y}-\frac{r\sigma_Y\nu_{1X}}{\sigma_X}$$

$$=\bar{Y}-\frac{r\sigma_Y\bar{X}}{\sigma_X}.$$

又因爲 $$\frac{\begin{vmatrix} N & \Sigma Y \\ \Sigma X & \Sigma XY \end{vmatrix}}{\begin{vmatrix} N & \Sigma X \\ \Sigma X & \Sigma X^2 \end{vmatrix}}=\frac{N\Sigma XY-\Sigma X\Sigma Y}{N\Sigma X^2-(\Sigma X)^2}$$

故 $$\frac{\begin{vmatrix} N & \Sigma Y \\ \Sigma X & \Sigma XY \end{vmatrix}}{\begin{vmatrix} N & \Sigma X \\ \Sigma X & \Sigma X^2 \end{vmatrix}}=\frac{\nu_{XY}-\nu_{1X}\nu_{1Y}}{\sigma_X^2}$$

$$=\frac{r\sigma_X\sigma_Y}{\sigma_X^2}$$

$$=\frac{r\sigma_Y}{\sigma_X}$$

於是 $$a=\bar{Y}-\frac{r\sigma_Y\bar{X}}{\sigma_X},\quad b=\frac{r\sigma_Y}{\sigma_X}.$$

因之

(6) $$Y-\bar{Y}=r\frac{\sigma_Y}{\sigma_X}(X-\bar{X}).$$

此卽通常所稱之 Y 倚 X 之迴歸線 (line of regression of Y on X). 可見此線經過平均點 $(\bar{X},\bar{Y})$.

同理如果配合時以使各點至直線橫距離剩餘數平方和最小爲準則,並且 $X=a'+b'Y$ 爲欲配合之直線;則可得第二條迴歸線. 我們可依同法測定式中之 $a'=\bar{X}-\frac{r\sigma_X\bar{Y}}{\sigma_Y}$, 並且 $b'=\frac{r\sigma_X}{\sigma_Y}$(參閱第五章第四節 2). 於是得到所求之方程式

(7) $$X-\bar{X}=\frac{r\sigma_X}{\sigma_Y}(Y-\bar{Y}).$$

此卽通常所稱之 X 倚 Y 之迴歸線 (line of regression of X on Y). 可見其亦經過平均點 $(\bar{X},\bar{Y})$. 迴歸線方程式 (6) 及 (7) 各寫之如下:

(8) $$y=\frac{r\sigma_Y}{\sigma_X}x,$$

(9) $$x=\frac{r\sigma_X}{\sigma_Y}y;$$

彼處 $$x=X-\bar{X},\quad y=Y-\bar{Y}.$$

由式(8)及(9)得知：如果 $r=-1$，則二迴歸線(8)及(9)均變做

$$y=\frac{-\sigma_Y}{\sigma_X}x.$$

此即二迴歸線在第二及第四象限內重合，其與 x 軸及 y 軸所成之角度各爲 α 及 $\beta=90°-\alpha$（見圖 52a）。如果 $-1<r<0$，則迴歸線(8)與 $O'x$ 成一角度 α'，迴歸線(9)與 $O'y$ 成一角度 β'；此二迴歸線亦在第二及第四象限內，但不重合（見圖 52b）。如果 $r=0$，則迴歸線(8)變做 $y=0$，迴歸線(9)變做 $x=0$；即各變做 x 軸及 y 軸，彼此互相成垂直（見圖 52c）。如果 $0<r<1$，則迴歸線(8)與 $O'x$ 成一角度 α''，迴歸線(9)與 $O'y$ 成一角度 β''，並且二迴歸線在第一第三象限內（見圖 52d）。最後如果 $r=1$，則二迴歸線均變做 $y=\frac{r\sigma_Y}{\sigma_X}x$；即二迴歸線在第一及第三象限内重合，其與 X 軸及 Y 軸所成之角度各爲 α''' 及 $\beta'''=90°-\alpha'''$（見圖 52e）。

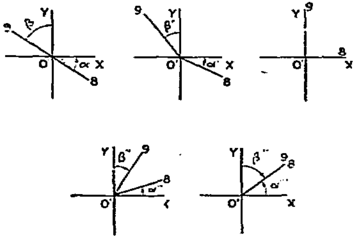


圖 52 相關係數 r 與迴歸線(8)及(9)之關係。

用方程式 (6) 可由指定之 X 值以估計 Y 值. 同理, 用式 (7) 可由指定之 Y 值以估計 X 值. 如果 X 值之變化易於控制, 則以用式 (6) 估計 Y 值爲宜; 如果 Y 值之變化易於控制, 則以用式 (7) 估計 X 值爲宜.

數量 $b = r\frac{\sigma_Y}{\sigma_X}$ 叫做 Y 倚 X 之迴歸係數 (regression coefficient of Y on X), 其爲 X 變化一單位 Y 值之變化也. 同樣, $b' = r\frac{\sigma_X}{\sigma_Y}$ 叫做 X 倚 Y 之迴歸係數 (regression coefficient of X on Y), 其爲 Y 變化一單位 X 值之變化也. 如此 r 之數值等於 $\pm\sqrt{bb'}$, 卽 r 爲 b 及 b' 之比例中項. 至於根號前之正負號須由原始論據決定之: 卽一變數增加或減少他變數隨之, 則宜取正號; 反之, 一變數增加他變數反減少, 或一變數減少他變數反增加, 則宜取負號.

就幾何學上意義說, b 爲直線 (6) 之斜率, $\frac{1}{b'}$ 爲直線 (7) 之斜率, 並且二直線相交於 $(\bar{X}, \bar{Y})$ 點.

6. 估計標準誤　各點散佈於 Y 倚 X 迴歸線兩側之情形, 可用 S_Y^2 度量之; 彼處 $S_Y^2 = \frac{1}{N}\Sigma d_Y^2$ 並且 d_Y 爲 Y 之觀察值與其由 Y 倚 X 迴歸線算得之對應值之差. S_Y 通常被稱做 Y 值之估計標準誤 (standard error of estimate).

定理 2.　設 S_Y, σ_Y 及 r 代表之意義同上. 則 $S_Y^2 = \sigma_Y^2(1 - r^2)$.

證:　因爲　$NS_Y^2 = \Sigma d_Y^2$

$$= \Sigma\left[Y - \left\{\bar{Y} + r\frac{\sigma_Y}{\sigma_X}(X - \bar{X})\right\}\right]^2$$

$$= \Sigma\left[y - r\frac{\sigma_Y}{\sigma_X}x\right]^2, \quad x = X - \bar{X} \text{ 並且 } y = Y - \bar{Y}$$

$$= \Sigma y^2 - 2r\frac{\sigma_Y}{\sigma_X}\Sigma xy + r^2\frac{\sigma_Y^2}{\sigma_X^2}\Sigma x^2$$

$$=N\sigma_Y^2-2Nr\frac{\sigma_Y}{\sigma_X}r\sigma_X\sigma_Y+Nr^2\frac{\sigma_Y^2}{\sigma_X^2}\sigma_X^2$$
$$=N\sigma_Y^2(1-r^2).$$

上式兩方以 N 除之，我們有

(10) $$S_Y^2=\sigma_Y^2(1-r^2).$$ 證訖.

系　估計標準誤 S_Y 可用下式計算之：

(11) $$S_Y=\sigma_Y\sqrt{1-r^2}.$$

應用同樣手續，我們取 X 之觀察值與其由 X 倚 Y 算得之對應值的差量，則得描寫散佈於 X 倚 Y 迴歸線兩側情形之常數

(12) $$S_X=\sigma_X\sqrt{1-r^2}.$$

彼處 S_X 通常被稱做 X 值之估計標準誤.

7. 相關係數與估計標準誤之性質　r 及 S_Y 或 S_X 之意義可研討之如下：

當用標準單位表示時，迴歸線 (6) 及 (7) 各變為

(13) $$t_Y=rt_X.$$

(14) $$t_X=rt_Y \text{ 或 } t_Y=\frac{1}{r}t_X.$$

彼處 $t_X=\dfrac{X-\bar{X}}{\sigma_X}$ 及 $t_Y=\dfrac{Y-\bar{Y}}{\sigma_Y}$.

在此形式中，$r=\dfrac{dt_Y}{dt_X}$ 為直線 (13) 之斜率，$\dfrac{1}{r}=\dfrac{dt_Y}{dt_X}$ 為直線(14)之斜率. 可見：當 r 為正數時，一變數 t_X 增加或減少，他變數 t_Y 亦隨之增加或減少；當 r 為負數時，一變數之增加（或減少）他變數反減少（或增加）. 此等增加(或減少)之程度依 r 絕對值之小大定之. 因此當兩變數以標準單位表示並且迴歸線為直線時，r 為度量二量相關之常數[參閱式(1)].

當 $r=-1$ 時,二迴歸線(13)及(14)之斜率相同,即二迴歸線之一與 t_X 軸成 $-45°$ 角,另一與 t_Y 軸成 $+45°$ 角,即其在第二及第四象限內重合(見圖 53a). 當 $-1<r<0$ 時,二迴歸線亦在第二及第四象限內但不重合,其一與 t_X 軸成 $-\alpha$ 度角,並且其他與 t_Y 軸成 α 度角(見圖 53b). 當 $r=0$ 時,二迴歸線各與 t_x 及 t_y 軸重合:即直線(13)與 t_x 軸所成之角度為 0,直線(14)與 t_y 軸所成之角度亦為 0;二式各變做 $t_Y=0$ 及 $t_X=0$,即在原來單位上各變做 $Y=\bar{Y}$ 及 $X=\bar{X}$,彼此互成垂直(見圖 53c). 當 $0<r<1$ 時,二迴歸線之斜率不同在第一及第三象限內,直線與 t_x 軸成 α' 度角,並且直線(14)與 t_y 軸成 $-\alpha'$ 度角(見圖 53d). 最後,當 $r=+1$ 時,二迴歸線在第一及第三象限內重合,直式(13)與 t_X 軸成 45° 角,直線(14)與 t_Y 軸成 $-45°$ 角,重合之迴歸線平分象限,但不以標準單位表示之方程式,除 $\sigma_Y=\sigma_X$ 外,則否(見圖 53e).

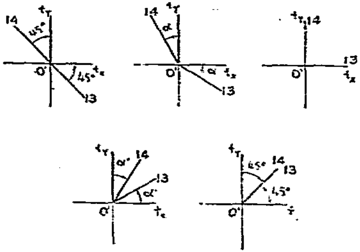


圖 53. 相關係數 r 與迴歸線(13)及(14)之關係。

很顯然的式(10)左方 $S_Y{}^2$ 及其右方 $\sigma_Y{}^2$ 均為正數或為 0. 故

$-1 \leq r \leq +1$. 如果各觀察點均在迴歸線上,則式(10)左方必為 0. 但因 σ_Y^2 不能為 0, 則 $1-r^2$ 必為 0, 即 $r=\pm1$. 此即在本節開始將所稱之完全相關. 當 $r=+1$ 時 稱二數為成正完全相關(positively perfect correlation);反之 若 $r=-1$,則稱之為成負完全相關(negatively perfect correlation). 如果各點漫然散佈在迴歸線兩側,則 $S_Y^2=\sigma_Y^2$, 因之 $r=0$. 此即以前所稱之成零相關.

應當注意者:在直線相關情形中,當 X 與 Y 不相關聯,則 $r=0$, 即變數 X 與 Y 彼此獨立的;但在非直線相關情形中,當 $r=0$ 時, X 與 Y 或者成高度相關. 關於此點容在下節研究之.

S_Y 之用途在其能判斷由 Y 倚 X 迴歸線估計 Y 值之確度. 為領略這個意義起見,我們研究當 $|r|$ 由 0 變至 1 時式(6)及式(10)之變化情形. 當 $r=0$ 時,式(6)變做 $Y=\bar{Y}$. 就是說,由 X 估計 Y 值其估計最美滿之 Y 值為 Y 分配之平均數;換言之, Y 之估計值即其本身之平均數,與 X 值漫無關係. 在此情形中 $S_Y=\sigma_Y$. 此為必然之結果,因為環繞直線 $Y=\bar{Y}$ 之標準差與環繞平均數 $\bar{Y}$ 各 Y 之標準差定然相同. 但 r 之值增加, S_Y 之值減少,當 $|r|$ 由 0 增至 1 時, S_Y 由 σ_Y 減至 0. $|r|$ 之增加為 S_Y 較 σ_Y 進步之表示. 其進步之意義可由圖 54 中見之,彼處在迴歸線 RR' 兩側各繪有平行於 RR' 之直線,其去 RR' 之縱距離為 S_Y. 若已給 $|r|\neq0$, 此二平行線表示各觀察值對於迴歸線之平均離勢. 當 $r=0$ 時,二平行線表示各觀察值對於 $Y=\bar{Y}$ 之平均離勢,並且其去直線 $Y=\bar{Y}$ 之距離為 σ_Y. 當 $|r|$ 由 0 增加,迴歸線由水平位置轉到最末當 $|r|=1$ 時之位置,同時 S_Y 由 σ_Y 減少而為 0. S_Y 之減少情形

可由式 (10) 知之，卽當 $|r|$ 如此增加時，S_Y 之值按比例 $(1-r^2)^{\frac{1}{2}}$ 而減少.

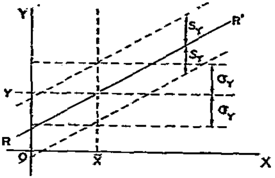


圖 51. 當 σ_Y 固定時 S_Y 因 r 之增加按數值 $\sqrt{1-r^2}$ 而減少之圖解。

關於 X 倚 Y 之迴歸線 (7) 可做同樣之分析，當 r 由 0 變至 1 時，此迴歸線由垂直位置 $X=\bar{X}$ 轉向直線 (6) 並且最後與之重合. 當直線 (7) 轉動時，S_X 按比例 $(1-r^2)^{\frac{1}{2}}$ 由 σ_X 逐漸減少以至於 0.

由直線 (13) 可以證明

(15) $$\frac{1}{N}\delta_Y^2=1-r^2,$$

彼處 δ_Y 爲 t_Y 之觀察值與其由式 (13) 算得之 t_Y 值之差.

由上式可見，$\frac{1}{N}\Sigma\delta^2$ 之最大值爲 1. 所以，若取

(16) $$1-\frac{1}{N}\Sigma\delta_Y^2$$

當做配合適度之量數，由式 (15) 及 (16) 可見 r^2 爲式 (13) 配合適度之量數. 類乎此亦可以討論式 (14).

8. 進一步之討論　假設已給一組 N 雙有關聯之變量 (X_1, Y_1), (X_2, Y_2), …, (X_N, Y_N), 並且已經求得迴歸線方程式 (6) 內常數之值. 於是將各已知之 X 值代入此式內，可得一組估計之 Y 值，譬如其值以 $_EY$ 代

之.因此我們有下列之定理.

定理 3. 假設 ${}_EY$, $\bar{Y}$, r 及 σ_Y 代表之意同上.求證(i) ${}_EY$ 之平均數 ${}_E\bar{Y}$ 與 $\bar{Y}$ 相等. (ii) ${}_EY$ 之變差等於 $r^2\sigma_Y^2$. 現在證明之如下:

證(i): 因 $${}_EY=\bar{Y}+r\frac{\sigma_Y}{\sigma_X}(X-\bar{X}).$$

於是 $${}_E\bar{Y}=\frac{1}{N}\sum_1^N\bar{Y}+\frac{r\sigma_Y}{\sigma_X}\frac{1}{N}\sum_{i=1}^N(X_i-\bar{X}).$$

但由第二章定理6得知 $\sum_{i=1}^N(X_i-\bar{X})=0$, 如此 ${}_E\bar{Y}=\bar{Y}$.

證(ii): 由定義 $$\sigma_{EY}^2=\frac{1}{N}\sum_1^N({}_EY-{}_E\bar{Y})^2.$$

我們知道由式(6)算得之 Y 與 ${}_EY$ 同,並且由本定理前段得知 $\bar{Y}={}_E\bar{Y}$;

所以 $$\sigma_{EY}^2=\frac{1}{N}\sum_{i=1}^N\left[r\frac{\sigma_Y}{\sigma_X}(X_i-\bar{X})\right]^2$$

$$=r^2\frac{\sigma_Y^2}{\sigma_X^2}\frac{1}{N}\sum_{i=1}^N(X_i-\bar{X})^2$$

(17) $$=r^2\sigma_Y^2.$$ 證之.

由於這個定理及式(10)可得

(18) $$S_Y^2=\sigma_Y^2-\sigma_{EY}^2.$$

此關係能幫助吾人了解 Y 及 S_Y 之意義.習慣上稱 σ_{EY}^2 為由 X 估計 Y 之變差;即是稱做 Y 倚 X 迴歸線 Y 之變差. 所以式(18)所顯示之 Y 的變化為 Y 依 X 的變化打以相當折扣者. S_Y^2 有時被叫做剩餘變差. 式(18)之關係可用一直角三角形之邊表之,即是如果以 σ_Y 為直徑畫一半圓弧,在其內作一內接直角三角形,如圖 55 所示;則此三角形其他

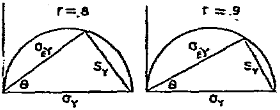


圖 55. S_Y, σ_Y 及 σ_{EY} 關係之圖示。

二邊分別代表 σ_{EY} 及 S_Y. 由圖可見，$\cos\theta = \sigma_{EY}/\sigma_Y$. 故由式 (17) 我們有 $\cos\theta = r$. $\cos\theta$ 之特殊值可由表查得之. 當 $r = .8$ 時，$\theta = 36°52'$; 當 $r = .9$ 時，$\theta = 25°50'$. 於是知：算得 σ_Y 及 σ_{EY} 之值後，應用三角函數表及式 (17) 可以直接查 r 之值.

定理 4. Y 之估計值與實在值之相關係數等於 X 與 Y 之相關係數.

證：我們祇要能證明

$$\frac{\frac{1}{N}\Sigma Y_E Y - \bar{Y}_E \bar{Y}}{\sigma_{EY}\sigma_Y} = \frac{\nu_{XY} - \nu_{1X}\nu_{1Y}}{\sigma_X \sigma_Y}.$$

就夠了. 將 $_EY$, $_E\bar{Y}$, σ_{EY} 代入上式並將結果化簡卽得. 其餘步驟留待學生補充.

9. 餘相關係數 相關係數愈大，則由式 (10) 或 (11) 得知估計標準誤愈小，故由方程式估計之 Y 值因估計標準誤之減少而進步；反之，相關係數愈小，則估計標準誤愈大，故由方程式估計之 Y 值因估計標準之過大不能有所進步. 此種進步及不能進步之情形，可用下式度量之：

$$(19) \qquad \kappa = (1 - r^2)^{\frac{1}{2}},$$

彼處之 κ 有時被叫做餘相關係數 (coefficient of alienation). 恰巧 κ 與 r 之基本關係可用以 1 爲半徑之圓表之，卽 $\kappa^2 + r^2 = 1$.

若以 κ' 代表 $1-\kappa$, 則由式(19)得

$$(20) \qquad \kappa' = 1-(1-r^2)^{\frac{1}{2}},$$

彼處之 κ' 通常被叫做進步因子(improvement factor), 因 $\kappa' = 1 - S_1/\sigma_Y$ 即示其因 $|r|$ 之增加而減少故也.

很顯然的 $$\kappa^2 = \frac{1}{N}\Sigma\delta_Y^2 = \frac{S_Y^2}{\sigma_Y^2},$$ [參閱式(15)及(10)]

並且 $$\kappa' = 1-\kappa.$$

表 61. r 值及其對應之 κ 與 κ' 值

r	κ [由式(19)算得]	κ' [由式(20)算得]
.1	.995	.005
.2	.980	.020
.3	.954	.046
.4	.917	.083
.5	.866	.134
.6	.800	.200
.7	.714	.286
.8	.600	.400
.9	.436	.564
.92	.392	.608
.94	.341	.659
.96	.280	.720
.98	.198	.811
1.00	.000	1.000

表 61 所列者為 r 值及其對應之 κ 和 κ' 值,在事實無關聯的情形

中，Y 之最好估計值為 $\bar{Y}$；在事實有關聯的情形中，若指定 r 值之後則用方程式(6)估計之 Y 值比只用 $\bar{Y}$ 估計者為優．上表裏所列之 κ 行數值能告訴我們以方程估計 Y 值之優異程度．例如，當表 61 中之 $r=.5$ 時，由 κ 行得知標準誤 S_Y 約為 σ_Y 之 86.6%，由 κ' 行得知 S_Y 比由 $\bar{Y}$ 計算者僅減少 13.4%；是以當 $|r|$ 由 1 降至 .8，S_Y/σ_Y 由 0 增至 60.0%；自他方面觀之，當 $|r|$ 由 0 增至 .8，估計誤之進步僅為 40.0%．相關係數 $r=.9$，預測之 Y 值比根據平均數猜度僅僅好了 56.4%．

由於普通之 r 值，指定 X 值後以推測 Y 之個別值未必可靠，此為顯然事實；然而，當 N 值很大時，我們指定 X 值後用以估計 Y 之平均值是相當可靠的．關於此點容後詳加說明．

10. 相關表　當所研究之樣本相當大時，成雙之觀察值可用散佈圖(見圖 50 及 51)代表，前已言之；但最好用相關表以代替散佈圖．我們可將散佈圖劃分大小適宜之矩形，各點落在矩形之內者，可以想像集中於矩形中心，於是在矩形內寫一數目以表示點數．故相關表者實係兩組變量 X 及 Y 之頻數分配表，其由若干行及若干列組成，每行或列之本身各為一頻數分配．

表 62 裏的 X 代表 100 名學生平時成績，Y 代表其大考成績．當將論據歸組之後，以組值代替了變量．是以在此相關表裏學生平時成績在 85 與 90 之間其大考成績在 80 與 85 之間者有 9 人；置大考成績於不論，學生平時成績在 85 與 90 之間者有 24 人；置平時成績於不論，學生大考成績在 80 與 85 之間者有 28 人．下表二變數 X 及 Y 之組距相同，但在大多數情形中則未必全是如此．

a. 實例 為使學者對於相關表得有一具體觀念,特舉一相關表之實例於下:

表 62. 相關表之一例

大考分數 Y	平時分數:X							各列總計	y
	65 –	70 –	75 –	80 –	85 –	90 –	95 –		
60 –	1	..	..	..	..	..	..	1	–3
65 –	3	2	..	..	..	..	..	5	–2
70 –	2	3	5	6	1	1	..	18	–1
75 –	3	3	7	6	4	..	..	23	0
80 –	4	4	6	4	9	1	..	28	1
85 –	..	..	1	3	8	1	5	18	2
90 –	..	..	..	1	2	3	1	7	3
各行總計	13	12	19	20	24	6	6	100	...
x	–3	–2	–1	0	1	2	3	...	

b. 一般形式 一般言之,相關表之形式如下:

表 63. 相關表之一般形式

變數 Y	變數:X					Ψ_Y
	X_1	X_2	X_3	...	X_m	
Y_1	F_{11}	F_{21}	F_{31}	...	F_{m1}	Ψ_1
Y_2	F_{12}	F_{22}	F_{32}	...	F_{m2}	Ψ_2
Y_3	F_{13}	F_{23}	F_{33}	...	F_{m3}	Ψ_3
...	$\vdots$	$\vdots$			$\vdots$	
Y_n	F_{1n}	F_{2n}	F_{3n}	...	F_{mn}	Ψ_n
Φ_X	Φ_1	Φ_2	Φ_3	...	Φ_m	N

上表係由 m 行及 n 列頻數組成,即由 $m\times n$ 個方格(矩形)組成,此

等方格內所包含之頻數可用 F_{XY} 代表之;故 $F_{X_iY_j}$ 或簡寫做 F_{ij} 為方格內第 i 行與第 j 列交會處之頻數.相關表每行為 Y 之頻數分配,與指定之變數或組值 X 成對應;其各列為 X 之頻數分配,與指定之變數或組值 Y 成對應. X 行各頻數之和可用 Φ_X 代表之,故 Φ_{X_i} 或簡寫之為 Φ_i 為第 i 行內各頻數之和;同樣, Y 列各頻數之和可用 Ψ_Y 代表之,故 Ψ_{Y_j} 或簡做 Ψ_j 為第 j 列內各頻數之和. 更進一步講,如果我們用 $\sum\limits_{Y}F_{XY}$ 代表 $F_{X1}+F_{X2}+\cdots+F_{Xn}$, $\sum\limits_{X}F_{XY}$ 代表 $F_{1Y}+F_{2Y}+\cdots+F_{mY}$, $\sum\limits_{Y}YF_{XY}$ 代表 $Y_1F_{X1}+Y_2F_{X2}+\cdots\cdots+Y_nF_{Xn}$, $\sum\limits_{X}XF_{XY}$ 代表 $X_1F_{1Y}+X_2F_{2Y}+\cdots+X_mF_{mY}$, 如此等等;則不難證明下列各式是成立的:

$$
(21)\quad
\begin{cases}
N=\sum\limits_{X}\Phi_X=\sum\limits_{Y}\Psi_Y=\sum\limits_{Y}\sum\limits_{X}F_{XY}=\sum\limits_{X}\sum\limits_{Y}F_{XY},\\
\nu_{1X}=\dfrac{1}{N}\sum\limits_{X}X\Phi_X=\dfrac{1}{N}\sum\limits_{X}X\sum\limits_{Y}F_{XY}=\dfrac{1}{N}\sum\limits_{Y}\sum\limits_{X}XF_{XY}=\dfrac{1}{N}\sum\limits_{X}\sum\limits_{Y}XF_{XY},\\
\nu_{2X}=\dfrac{1}{N}\sum X^2\Phi_X=\dfrac{1}{N}\sum\limits_{X}X^2\sum\limits_{Y}F_{XY}=\dfrac{1}{N}\sum\limits_{Y}\sum\limits_{X}X^2F_{XY}=\dfrac{1}{N}\sum\limits_{X}\sum\limits_{Y}X^2F_{XY},\\
\nu_{1Y}=\dfrac{1}{N}\sum Y\Psi_Y=\dfrac{1}{N}\sum\limits_{Y}Y\sum\limits_{X}F_{XY}=\dfrac{1}{N}\sum\limits_{Y}\sum\limits_{X}YF_{XY}=\dfrac{1}{N}\sum\limits_{X}\sum\limits_{Y}YF_{XY},\\
\nu_{2Y}=\dfrac{1}{N}\sum Y^2\Psi_Y=\dfrac{1}{N}\sum\limits_{Y}Y^2\sum\limits_{X}F_{XY}=\dfrac{1}{N}\sum\limits_{Y}\sum\limits_{X}Y^2F_{XY}=\dfrac{1}{N}\sum\limits_{X}\sum\limits_{Y}Y^2F_{XY},\\
\nu_{XY}=\dfrac{1}{N}\sum\limits_{Y}Y\sum\limits_{X}XF_{XY}=\dfrac{1}{N}\sum\limits_{X}X\sum\limits_{Y}YF_{XY}=\dfrac{1}{N}\sum\limits_{Y}\sum\limits_{X}XYF_{XY}\\
\qquad\qquad=\sum\limits_{X}\sum\limits_{Y}XYF_{XY}.
\end{cases}
$$

現在我們證明 $\sum\limits_{Y}\sum\limits_{X}XF_{XY}=\sum\limits_{X}\sum\limits_{Y}XF_{XY}$ 以示一例,其他各式可依同法證明.

證: 因為 $\sum\limits_{Y}\sum\limits_{X}XF_{XY}=\sum\limits_{Y}(X_1F_{1Y}+X_2F_{2Y}+\cdots+X_mF_{mY})$

$$=X_1\sum\limits_{Y}F_{1Y}+X_2\sum\limits_{Y}F_{2Y}+\cdots+X_m\sum\limits_{Y}F_{mY},$$

所以

(22) $$\sum_{Y}\sum_{X} XF_{XY}=X_1F_{11}+X_1F_{12}+\cdots+X_1F_{1n}$$
$$+X_2F_{21}+X_2F_{22}+\cdots+X_2F_{2n}$$
$$+\cdots\cdots\cdots\cdots\cdots\cdots$$
$$+X_mF_{m1}+X_mF_{m2}+\cdots+X_mF_{mn};$$

又因 $$\sum_{X}\sum_{Y} XF_{XY}=\sum_{X}(XF_{X1}+XF_{X2}+\cdots+XF_{Xn})$$
$$=\sum_{X}XF_{X1}+\sum_{X}XF_{X2}+\cdots+\sum_{X}XF_{Xn}$$

所以

(23) $$\sum_{X}\sum_{Y} XF_{XY}=X_1F_{11}+X_2F_{21}+\cdots+X_mF_{m1}$$
$$+X_1F_{12}+X_2F_{22}+\cdots+X_mF_{m2}$$
$$+\cdots\cdots\cdots\cdots\cdots\cdots$$
$$+X_1F_{1n}+X_2F_{2n}+\cdots+X_mF_{mn}.$$

但式 (22) 之右方與式 (23) 之右方完全相同,故

$$\sum_{Y}\sum_{X} XF_{XY}=\sum_{X}\sum_{Y} XF_{XY}.$$ 證訖.

於是由式 (21) 我們有

(24) $$\begin{cases} \bar{X}=\nu_{1X} \\ \sigma_X=\sqrt{\nu_{2X}-\nu_{1X}^2} \\ \bar{Y}=\nu_{1Y} \\ \sigma_Y=\sqrt{\nu_{2Y}-\nu_{1Y}^2} \\ r=\dfrac{\nu_{XY}-\nu_{1X}\nu_{1Y}}{\sigma_X\sigma_Y} \end{cases}$$

此式與式 (3) 之形式完全一樣,不過式中之 ν_{1X}, ν_{2X}, ν_{1Y}, ν_{2Y} 及 ν_{XY}

各代表 $\frac{1}{N}\sum_X\sum_Y XF_{XY}$, $\frac{1}{N}\sum_X\sum_Y X^2F_{XY}$, $\frac{1}{N}\sum_X\sum_Y F_{XY}$, $\frac{1}{N}\sum_X\sum_Y Y^2F_{XY}$ 及 $\frac{1}{N}\sum_X\sum_Y XYF_{XY}$ 罷了.

相關表方格內之頻數 F_{XY} 不變,其右邊及底邊總和亦不變;反之則謬.例如表 62 右下角四個方格內之頻數 $\begin{bmatrix}1 & 5\\ 3 & 1\end{bmatrix}$ 以 $\begin{bmatrix}2 & 4\\ 2 & 2\end{bmatrix}$ 易之,其邊和不變.故邊和頻數分配 Φ_X 及 Ψ_Y 固定之後,不足以推斷方格內頻數分配之形式.

若 X 行各 Y 之平均數以 $\bar{Y}_X$ 記之,並且 Y 列各 X 之平均數以 $\bar{X}_Y$ 記之;則可敍述下列之定理.

定理 5. 平均數 $\bar{Y}$ 等於各行平均數 $\bar{Y}_X$ 之加權平均數.

證: 此定理祇要能證明 $\frac{1}{N}\sum_X \bar{Y}_X\Phi_X=\bar{Y}$ 就夠了,

彼處 $\bar{Y}_X=\frac{1}{\Phi_X}\sum_Y YF_{XY}$. 留待學生自行證明.

11. 計算 r 之簡捷法 用式 (24) 末式計算 r 不甚方便,因是將其簡化之,以應需要.將式 (4) 中之 X 及 Y 代入式 (21) 中,並且讓 $f_{xy}=F_{X_0+hx,\ Y_0+ky}$, $\phi_x=\Phi_{X_0+hx}$, $\psi_y=\Psi_{Y_0+ky}$;我們有

$$
(25)\qquad
\begin{cases}
\nu_{1X}=X_0+h\nu_{1x},\\
\nu_{2X}=X_0^2+2hX_0\nu_{1x}+h^2\nu_{2x},\\
\nu_{1Y}=Y_0+k\nu_{1y},\\
\nu_{2Y}=Y_0^2+2kY_0\nu_{1y}+k^2\nu_{2y},\\
\nu_{XY}=X_0Y_0+hY_0\nu_{1x}+kX_0\nu_{1y}+hk\nu_{xy}.
\end{cases}
$$

所以式 (24) 變做

$$(26)\quad \begin{cases} \bar{X}=X_0+h\nu_{1x}, \\ \sigma_X=h\sigma_x, \\ \bar{Y}=Y_0+k\nu_{1y}, \\ \sigma_Y=k\sigma_y, \\ r=\dfrac{\nu_{xy}-\nu_{1x}\nu_{1y}}{\sigma_x\sigma_y}; \end{cases}$$

彼處 $\nu_{1x}=\frac{1}{N}\sum_x x\phi_x$, $\nu_{2x}=\frac{1}{N}\sum_x x^2\phi_x$, $\nu_{1y}=\frac{1}{N}\sum_y y\psi_y$, $\nu_{2y}=\frac{1}{N}\sum_y y^2\psi_y$, $\nu_{xy}=\frac{1}{N}\sum_y y\sum_x xfxy$, $\sigma_x=\sqrt{\nu_{2x}-\nu_{1x}^2}$, $\sigma_y=\sqrt{\nu_{2y}-\nu_{1y}^2}$ 並且 $N=\sum_x\sum_y f_{xy}$. 式(26) 與式 (3) 之形式完全一樣,不過式 (3) 中之 ν_{1X}, ν_{2X}, ν_{1Y}, ν_{2Y} 及 ν_{XY} 分別以 ν_{1x}, ν_{2x}, ν_{1y}, ν_{2y} 及 ν_{xy} 易之罷了.應用式 (26) 之末式由相關表計算 r 之手續較應用式 (24) 末式簡便的多.

例 2. 計算下表裏所列論據之相關係數.

計算相關係數之手續如下(見下面表 64).

a. 計算各行及各列裏數值　下表各行及各列裏數值除 $\sum_x xf_{xy}$ 外無須加以說明. $\sum_x xf_{xy}$ 行之第一個數值係 4 與 (−6), 2 與 (−5), 1 與 (−4) 及 9 與 (−2) 相乘積之和,即 $4\times(-6)+2\times(-5)+1\times(-4)+9\times(-2)=-56$; 第二個數值係 1 與 (−6), 9 與 (−5), 19 與 (−4), 19 與 (−3) 及 29 與 (−2) 相乘積之和,即 $1\times(-6)+9\times(-5)+19\times(-4)+19\times(-3)+29\times(-2)=-242$; 其餘可依同法計算.實際計算時可將 x 列之值 −6, −5, −4, −3, −2, −1, 0, 1, 2, 3, 4 及 5 (卽用粗黑線包圍之列)另外用一紙條抄寫下來,然後將此紙條置於表體第一列 4, 2, 10 之下,使表上之 4 與紙條上之 −6 相對,因之 2 與 −5, 1 與 −4

表 64. 聯合銀行與商業銀行貼現率之相關(論據見參考書 28)

商業銀行貼現率(%) Y	聯合銀行貼現率(%):X 1.25-	1.75-	2.25-	2.75-	3.25-	3.75-	4.25-	4.75-	5.25-	5.75-	6.25-	6.75-	ϕ_y	y	$y\phi_y$	$y^2\phi_y$	$\sum_x xf_{xy}$	$y\sum_x xf_{xy}$
3.25-	4	2	1	…	9	…	…	…	…	…	…	…	16	-4	-64	256	-56	224
3.75-	1	9	19	19	29	…	…	…	…	…	…	…	77	-3	-231	693	-242	726
4.25-	…	…	16	27	122	96	3	…	…	…	…	…	264	-2	-528	1056	-485	970
4.75-	…	…	4	25	111	198	65	1	…	…	…	…	402*	-1	-402	402	-508	508
5.25-	…	…	…	1	90	164	175	45	…	…	…	…	475	0	0	0	-302	0
5.75-	…	…	…	…	9	21	146	150	8	32	…	…	366	1	366	366	223	223
6.25-	…	…	…	…	…	…	4	22	10	22	1	3	62	2	124	284	127	254
6.75-	…	…	…	…	…	…	…	5	4	63	9	36	117	3	351	1053	418	1254
7.25-	…	…	…	…	…	…	…	…	…	7	9	1	17	4	68	272	62	248
7.75-	…	…	…	…	…	…	…	…	…	1	1	2	4	5	20	100	17	85
ϕ_x	5	11	40	72	370	477	393	223	22	125	20	42	1800	—	-296	4416	-740	4492
x	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	—					
$x\phi_x$	30	-55	-160	-216	-740	-477	0	223	44	375	80	210	-740					
$x^2\phi_x$	180	275	640	648	1480	477	0	223	88	1125	320	1050	6506					

及 9 與 -2 一一相對，將此相對之數字一一乘之，並將乘得之數加之，於是得到欲求之第一個數值 -56. 將紙條向下平移一列，使表上之 1 與紙條上之 -6 相對，因之 9 與 -5, 19 與 -4, 19 與 -3 及 29 與 -2 一一相對，同樣將此等相對之數字一一乘之，並將乘得之積加之，於是得到欲求之第二個數值 -242. 餘可類推. 因是我們有

$N=1800$, $\sum_x x\phi_x=-746$, $\sum_x x^2\phi_x=6506$, $\sum_y y\psi_y=-296$, $\sum_y y^2\psi_y=4446$, $\sum_y y\sum_x xf_{xy}=4492$.

b. 計算　根據上面算得之數值，我們有

$$\nu_{1x}=-.414444 \qquad \nu_{1y}=-.164444$$

$$\nu_{2x}=3.614444 \qquad \nu_{2y}=2.470000$$

$$\nu_{xy}=2.495556$$

$$\sigma_x=1.855446 \qquad \sigma_y=1.562996$$

$$\sigma_x\sigma_y=2.900055.$$

於是應用式 (26) 則得

$$r=0.837020.$$

c. 結果　於是得到欲求之相關係數

$$r=0.837.$$

12. 相關表上之迴歸線　相關表上各方格內之頻數可以想做位於方格中心之點，已如上述. 設 F_{ij} 為在坐標 (X_i, Y_j) 格內之點數，並且 Φ_i 為當橫坐標為 X_i 時所有之點數. 現在我們更假設每行各點均集中於該行平均點處，並且此觀察平均點之縱坐標以 ${}_0\bar{Y}_X$ 表之. 於是

(27) $$ {}_0\bar{Y}_X=\frac{1}{\Phi_X}\sum_Y YF_{XY}.$$

因之，$\sum_X {}_0\bar{Y}_X\Phi_X$ 代表 X 行各變量之總和.

每行各有其平均點 $(X, \bar{Y}_X)$. 假設此等點可用直線 $\bar{Y}_X=a+bX$ 表之，式中常數 a 及 b 之値依據最小平方原則測定之；${}_0\bar{Y}_X$ 及 ${}_c\bar{Y}_X$ 各爲每行觀察平均數及計算平均數；並且 ${}_0\bar{Y}_X-{}_c\bar{Y}_X$ 之平方以該行頻數 Φ_X 權之. 於是 a 及 b 之値將與配合 Y 倚 X 迴歸線時所得者相同. 此事實可證明之如下：

因爲以最小平方法原則測定直線 ${}_c\bar{Y}_X=a+bX$ 內常數之嚌範方程式爲

$$\begin{cases}\sum_X({}_0\bar{Y}_X-a-bX)\Phi_X=0,\\ \sum_X({}_0\bar{Y}_X-a-bX)X\Phi_X=0;\end{cases}$$

展開之，我們有

$$\begin{cases}\sum_X {}_0\bar{Y}_X\Phi_X-a\sum_X\Phi_X-b\sum_X X\Phi_X=0,\\ \sum_X X_0\bar{Y}_X\Phi_X-a\sum_X X\Phi_X-b\sum_X X^2\Phi_X=0.\end{cases} \tag{28}$$

又由式 (27) 及 (21)，我們有

$$\sum_X {}_0\bar{Y}_X\Phi_X=\sum_X\left(\sum_Y Yf_{XY}\right)=\sum_X\sum_Y YF_{XY}=N\nu_{1Y},$$

並且 $$\sum_X X\bar{Y}_X\Phi_X=\sum_X X\left(\sum_Y YF_{XY}\right)=\sum_X\sum_Y XYF_{XY}=N\nu_{XY}.$$

因之方程式 (28) 變做

$$\nu_{1Y}-a-b\nu_{1X}=0,$$

$$\nu_{XY}-a\nu_{1X}-b\nu_{2X}=0.$$

以 a 及 b 當做未知數解之，得

$$a=\nu_{1Y}-\frac{r\sigma_Y\nu_{1X}}{\sigma_X},$$

$$b=\frac{r\sigma_Y}{\sigma_X}.$$

故所求之方程式為

$$_{c}\bar{Y}_{X}=r\frac{\sigma_{Y}}{\sigma_{X}}X+\bar{Y}-\frac{\sigma_{Y}}{\sigma_{X}}r\bar{X},$$

或者寫做

(29) $$_{c}\bar{Y}_{X}-\bar{Y}=r\frac{\sigma_{Y}}{\sigma_{X}}(X-\bar{X}).$$

此為由已給之 X 值求各行平均數 $_{c}\bar{Y}_{X}$ 之公式(見圖 56). 可見用此式求得之 $_{c}\bar{Y}_{X}$ 值與用式 (6) 求得之 Y 值相同. 因式 (6) 與 (29) 之形式完全相同, 祇要將式 (6) 中之 Y 以 $_{c}\bar{Y}_{X}$ 代之即可得到式 (29). 但以相關表做觀點, 迴歸線方程式乃係用以由 X 值估計 Y 值之平均數者.

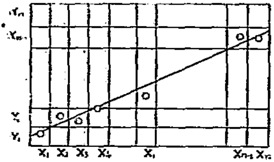


圖 56. 解釋迴歸線為配合各行平均數最美滿之直線。

13. 應用 相關表之論據和其他論據同, 常被人看做組成宇宙論證之樣本. 迴歸方程式雖由有限論據——樣本——算得者, 但對於富有代表性之樣本, 用以由已給之 X 值估計 Y 值之平均數常是有價值的. 茲舉例以明相關係數及迴歸線之應用如下:

例 3. 假設某工廠募用處實行一種制度, 凡應募者皆須受智力測驗, 如是積聚了好多關於應募者智力與錄用後工廠產量相關之論據, 如表 65 所示. (a) 試計算各行及各列之平均數, 以示工人智力與工廠產

量之關係．(b) 試計算兩個變數之相關係數，以示工人智力與工廠產量之相關程度．(c) 試由迴歸線方程式由於工人智力估計工廠產量．

表 65.　工人智力與工廠產量之關係（論據見參考書 25）

估標準產量之百分數 Y	工人智力：X								Φ_Y	$\bar{X}_Y$
	22.5	27.5	32.5	37.5	42.5	47.5	52.5	57.5		
55	2	3	3	4	1	1	...	...	14	33.2
65	2	5	5	8	8	6	1	...	35	38.0
75	2	1	5	6	16	8	5	...	43	41.6
85	1	3	12	11	7	12	7	1	54	40.7
95	...	2	1	10	12	9	8	2	44	44.0
105	...	...	5	7	8	11	8	7	46	45.9
115	...	...	1	3	1	4	4	4	17	48.1
125	...	...	...	...	2	3	2	...	7	47.5
Φ_X	7	14	32	49	55	54	35	14	260	-
$\bar{Y}_X$	67.9	72.1	1.0	84.8	85.7	90.9	95.6	105.0	-	

由表 65 算得各行以及各列之平均數，分別列入該表末列及末行中．由表 65 又可算得 X 之平均數 $\bar{X}$，其標準差 σ_X；Y 之平均數 $\bar{Y}$ 及其標準差 σ_Y，如下：

$$\bar{X}=42.17, \qquad \bar{Y}=87.31,$$

$$\sigma_X=8.40, \qquad \sigma_Y=17.41.$$

於是我們算得欲求之相關係數

$$r=.417,$$

並且 Y 倚 X 迴歸線之方程式

$$\bar{Y}_X-87.31=.864(X-42.17)$$

或將此方程式寫成下列之形式

$$\bar{Y}_X = 50.88 + .864X. \tag{30}$$

此爲配合各行加權平均數最美滿之直線，其美滿情形可由圖 57 見之。用式 (30) 可由指定之 X 值，以求各行之希望值 $\bar{Y}_X$. 該式用處可依據工人智力測驗等第，以測定計算工廠之平均生產能力；換言之，由於已雇用工人記錄得到之迴歸線可用以決定工廠將來如何選拔應募者，庶幾對於產量能達預期之標準。

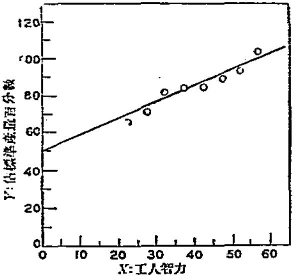


圖 57. 各行平均數代表之點及式(30)之軌跡(論據見表 65)。

14. S_Y 在相關表上之意義　S_Y 代表估計標準誤，其定義已見本節 6 欲了解 S_Y 與估計值 ${}_e\bar{Y}_X$ 之關係，第一樣要知道 S_Y 在相關表上之意義。讓我們以 $S_{Y \cdot X}$ 表示 X 行各觀察值對於迴歸線縱坐標之標準差。如此我們有

$$S_{Y \cdot X}^2 = \frac{1}{f_{X}}\sum({}_{o}Y - {}_{e}\bar{Y}_X)^2 f_{XY}, \tag{31}$$

彼處 ${}_oY$ 表示 Y 之觀察值，${}_e\bar{Y}_X$ 表示由迴歸線方程式算得之縱坐標；即

如將 $X=32.5$ 代入式(30)內算得 ${}_c\bar{Y}_X=78.96$. 於是，我們求 32 個 Y 值與 48.96 離差之平方；將此各個平方加之，以 32 除之，即得當 $X=32.5$ 時 $S_{Y\cdot X}^2$ 之值；將所得之值開方之，我們有 $S_{Y\cdot X}$ 之值 15.96.

S_Y 與 ${}_c\bar{Y}_X$

之關係可用下列之定理敍述之：

定理 6. X 行各 Y 值對於 ${}_c\bar{Y}_X$ 變差 $S_{Y\cdot X}^2$ 之加權算術平均數等於 $S_Y^2=\sigma_Y^2(1-r^2)$.

證：由式(31)我們有

$$\frac{1}{N}\sum_X S_{Y\cdot X}^2\Phi_X=\frac{1}{N}\sum_X\sum_Y({}_0Y-{}_c\bar{Y}_X)^2F_{XY}.$$

將式(29)代入上式，並將 Y 左下角所綴之小字母 0 去掉，我們有

$$\begin{aligned}
\frac{1}{N}\sum_X S_{Y\cdot X}^2\Phi_X&=\frac{1}{N}\sum_X\sum_Y\left[Y-\left\{\bar{Y}+r\frac{\sigma_Y}{\sigma_X}(X-\bar{X})\right\}\right]^2F_{XY}\\
&=\frac{1}{N}\sum_X\sum_Y\left[(Y-\bar{Y})-r\frac{\sigma_Y}{\sigma_X}(X-\bar{X})\right]^2F_{XY}\\
&=\frac{1}{N}\sum_X\sum_Y\left[(Y-\bar{Y})^2-2r\frac{\sigma_Y}{\sigma_X}(X-\bar{X})(Y-\bar{Y})\right.\\
&\qquad\left.+r^2\frac{\sigma_Y^2}{\sigma_X^2}(X-\bar{X})^2\right]F_{XY}\\
&=\sigma_Y^2-2r\frac{\sigma_Y}{\sigma_X}r\sigma_X\sigma_Y+r^2\frac{\sigma_Y^2}{\sigma_X^2}\sigma_X^2\\
&=\sigma_Y^2-r^2\sigma_Y^2\\
&=\sigma_Y^2(1-r^2)\\
&=S_Y^2,\quad \text{由式(10).}
\end{aligned}$$

證訖

根據表 64 下面求得之結果 ($\sigma_y=1.562996$ 及 $r=.837$)，我們用式(10)算得 $S_Y=.4276$; 並且求得 Y 倚 X 之迴歸線 $Y=2.391+.7051X$,

該迴歸線即圖 58 RR' 所代表之直線.在 RR' 兩側各繪有兩條與之平行之直線：一距 RR' 之縱距離爲 $\pm S_Y$，即圖中斷線所代表者；一距 RR' 縱之距離爲 $\pm 3S_Y$，即圖中之點線所代表者.如是，就理論上說，兩線斷線內所包圍之點數，即數值 (X, Y) 之個數，宜爲百分之 68.3；在點線所包圍區域內之點數宜爲百分之 99.7. 故估計標準誤與迴歸線關係若有標準差與平均數然(參閱第二章問題 25).

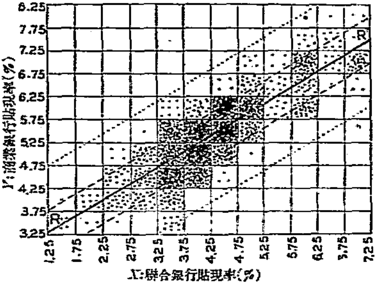


圖 58. 解釋估計標準誤之意義(論據見表 64)。

第三節　非直線相關

1. 相關比　我們曾經討論過簡單相關中之直線相關. 但相關不以直線爲限，其兩條迴歸線全非直線者有之，其中之一爲直線其他則否者亦或有之；例如下面所列之表，Y 倚 X 之迴歸線爲直線，而 X 倚 Y 之迴歸線爲曲線.

表 66. 非直線相關之一例

變量 Y	變量 X 0	1	2	3	各列總計
0	1	1	...	...	2
1	...	1	1	2	4
2	...	...	...	2	2
各行總計	1	2	4	4	8

若各行或各列之平均數之趨勢爲曲線,二量相關之程度不能用相關係數 r 度量之,因爲此時之 $r=0$ 僅表示缺乏直線相關,不足以推斷一般相關之不存在.因此用相關比 (correlation ratio) 代替相關係數.

爲說明相關比如何導出起見,我們取式 (10) 當做相關係數之定義;

即取 $$r^2=1-\frac{S_Y^2}{\sigma_Y^2},\quad [\text{參閱式 (10)}]$$

亦即取

(32) $$r=\pm\sqrt{1-\frac{S_Y^2}{\sigma_Y^2}}$$

當做相關係數之定義,彼處 S_Y 爲 X 行各 Y 值對於 Y_X 變差之加權平均數參閱定理 6.

以式 (32) 右方爲相關係數 r 定義之優點在能推得 r 之豐富意義.當 $S_Y=0$ 時,則 $r=\pm1$,即散佈圖上所有各點將落在迴歸線上,正與上面所講的相符.

最應牢記者,式 (32) 中之變差 S_Y^2 係以迴歸線爲軸算之者;現在讓我們用 $S_Y'^2$ 代表與 S_Y^2 成對應之變差,但以該行平均數爲計算之依據.

於是，當迴歸線確爲直線時，$S_Y^2=S_Y'^2$；但當其非爲直線時，$S_Y^2\neq S_Y'^2$. 這個事實暗示我們說，在非直線相關情形中，用以量度相關程度之公式與 $[1-S_Y^2/\sigma_Y^2]^{\frac{1}{2}}$ 有密切關係，即是用 S_Y' 以代替式 (10) 中之 S_Y. 我們於是取 $1-S_Y'^2/\sigma_Y^2$ 當做相關程度之量度，即取

$$(33)\qquad \eta_{YX}^2=1-\frac{S_Y'^2}{\sigma_Y^2},$$

彼處 η_{YX} 爲 Y 倚 X 之相關比，$S_Y'^2$ 爲各行變量各對其平均數變差之加權平均數，各行平均數可能距迴歸線很遠，亦或很近.

概括言之，Y 倚 X 之相關比爲度量各變量對於各行平均數之聚散情形.

類似上面討論，可寫 X 倚 Y 相關比平方之公式如下：

$$(34)\qquad \eta_{XY}^2=1-\frac{S_X'^2}{\sigma_X^2},$$

彼處 $S_X'^2$ 爲各列變量各對其平均數變差之加權平均數.

由式 (33) 得知 $\eta_{YX}^2\leq 1$，並且僅當每行所有各點均在該行平均數上時，$\eta_{YX}=1$. 由第二章定理 7 得知式 (33) 裏之 $S_Y'^2$ 一定小於或等於式 (32) 裏之 S_Y^2，蓋式 (32) 裏之離差係以迴歸線爲軸算之者. 於是由式 (33) 及 (32) 得以推斷 $\eta_{YX}^2\geq r^2$. 所以

$$1\geq\eta_{YX}^2\geq r^2.$$

更進而言之，當 Y 倚 X 之迴歸線爲直線，由於樣本算得之 $\eta_{YX}^2=r^2$ 縱然不等於零，但亦不能大於隨機選樣之波動. 所以 $\eta_{YX}^2-r^2$ 可以當做 Y 倚 X 直線性測驗之準則. 此點容在下面討論.

相關比可用各行平均數之標準差表之. 因此之故，讓 $\bar{Y}_X$ 表示 X 行

各 Y 值之平均數；$\sigma_{\overline{Y}_X}$ 為各行平均數之標準差．但求此標準差時每行離均差之平方和 $\Sigma(\overline{Y}_X-\overline{Y})^2$ 務須各以該行頻數 Φ_X 權之．於是得證

$$(35)\qquad \eta_{YX}^2=\frac{\sigma_{\overline{Y}_X}^{\ 2}}{\sigma_Y^2}.$$

就是說，Y 倚 X 之相關比為各行平均數標準差與變數 Y 標準差之比．

為證明式(35)起見，我們須證明 $\sigma_Y^2-S_Y'^2=\sigma_{\overline{Y}_X}^{\ 2}$．我們知道一行內各點對於平均數之離勢，可用該行標準差度量之．讓 $\sigma_{Y\cdot X}$ 表示在 X 行內各 Y 對其平均數 $\overline{Y}_X$ 之標準差，於是

$$(36)\qquad \sigma_{Y\cdot X}^2=\frac{1}{\Phi_X}\sum_Y(Y-\overline{Y}_X)^2F_{XY}.$$

兩邊乘以 Φ_X 乘之，我們有

$$\sigma_{Y\cdot X}^2\Phi_X=\sum_Y(Y-\overline{Y}_X)^2F_{XY}.$$

因之，施用運算符號 $\sum\limits_X$ 於上式之兩邊，並以 N 除所得之結果得

$$\frac{1}{N}\sum_X\sigma_{Y\cdot X}^2\Phi_X=\frac{1}{N}\sum_X\sum_Y(Y-\overline{Y}_X)^2F_{XY}.$$

即是全表各點對於各行平均數之離勢，可用各行變差 $\sigma_{Y\cdot X}^2$ 之加權平均數度量之．此加權平均數我們可用 $S_Y'^2$ 表之［參閱式(34)］，即是

$$(37)\qquad S_Y'^2=\frac{1}{N}\sum_X\sigma_{Y\cdot X}^2\Phi_X.$$

應用第二章式(13)並用本章符號，我們有

$$N\sigma_Y^2=\sum_X\sigma_{Y\cdot X}^2\Phi_X+\sum_X(\overline{Y}_X-\overline{Y})^2\Phi_X;$$

又因為原設 $\qquad \sum\limits_X\sigma_{Y\cdot X}^2\Phi_X=S_Y'^2,$

並且由定義 $\qquad \sum\limits_X(\overline{Y}_X-\overline{Y})^2\Phi_X=N\sigma_{\overline{Y}_X}^{\ 2}.$

於是將此二式代入末式內，並將結果以 N 除之；則得

$$\sigma_Y^2 = S_Y'^2 + \sigma_{\overline{Y}_X}^2.$$

該式可寫做 $\sigma_Y^2 - S_Y'^2 = \sigma_{\overline{Y}_X}^2$，因之得式 (35). 證訖.

2. η^2 之計算 茲爲說明公式 (33) 及 (35) 之應用起見，特舉實例明之如下:

例 1. 分別應用公式 (33) 及 (35) 裏的關係，並根據表 65 裏所列之論據，計算 η_{YX}^2.

因爲 $S_Y'^2 = 246.45$,

$\sigma_Y^2 = 303.11$,

及 $\sigma_{\overline{Y}_X} = 56.66$;

故由式 (33) 得 $\eta_{YX}^2 = .1869$,

並且由式 (35) 得 $\eta_{YX}^2 = .1869$.

此二結果完全相同.

爲了解 η_{YX}^2 之意義，式 (33) 及 (35) 是有裨益的，然而實際運算時常不用牠們，蓋以其所用之運算手續至爲繁留故也. 相關比之值由相關表直接計算之尚覺方便. 茲將實際運算公式推演之如下:

各行平均數之標準差 $\sigma_{\overline{Y}_X}$ 可用 (x, y) 單位表之，即

$$\sigma_{\overline{Y}_X}^2 = k^2 \sigma_{\overline{y}_x}^2,$$

表之，彼處 $x = \dfrac{X - X_0}{h}$, $y = \dfrac{Y - Y_0}{k}$, $\sigma_{\overline{y}_x}^2 = \dfrac{1}{N}\sum_x (\overline{y}_x - \overline{y})^2 \phi_x = \dfrac{1}{N}\sum \overline{y}_x^2 \phi_x - \overline{y}^2$, 並且 $\phi_x = \Phi_{X_0 + hx}$.

因爲我們知道 $\overline{y}_x = \dfrac{r}{\phi_x}$,

彼處 $r = \sum_y y f_{xy}$, $f_{xy} = F_{X_0 + hx,\ Y_0 + ky}$.

於是　　$$\sigma_{\bar{y}_x}{}^2=\frac{1}{N}\sum_x\frac{r^2}{\phi_x}-\bar{y}^2.$$

因之　　$$\eta_{YX}{}^2=\frac{\sigma_{\overline{Y}_X}{}^2}{\sigma_Y{}^2}$$

$$=\frac{k^2\left\{\frac{1}{N}\sum_x\frac{r^2}{\phi_x}-\bar{y}^2\right\}}{k^2\sigma_y{}^2}$$

(38)　　$$=\frac{1}{\sigma_y{}^2}\left\{\frac{1}{N}\sum_x\frac{r^2}{\phi_x}-\bar{y}^2\right\}.$$

同理可得 X 倚 Y 之相關比的平方

(39)　　$$\eta_{XY}{}^2=\frac{1}{\sigma_x{}^2}\left\{\frac{1}{N}\sum_y\frac{u^2}{\psi_y}-\bar{x}^2\right\},$$

彼處 $u=\sum_x xf_{xy}$, $\psi_y=\Psi_{Y_0+ky}$.

例 2. 試根據表 62 裏所列之論據並應用公式 (38), 計算 $\eta_{YX}{}^2$.

因為當 $x=-3, -2, -1, 0, 1, 2$ 及 3 時之 r 值已算得各為 $-7, -3, 3, 7, 30, 11$ 及 13, 並且 ϕ_x 之值各為 13, 12, 19, 20, 24, 6 及 6; 於是 r^2/ϕ_x 之值各為 3.77, .75, .47, 2.45, 37.50, 20.17 及 28.17, 其和為 93.28. 又由表 62 得知 $N=100$, 並且算得 $\bar{y}=0.54$, $\sigma_y{}^2=1.8084$. 將此等值代入式 (38) 內,我們有

$$\eta_{YX}{}^2=\frac{1}{1.8084}[93.28-(.54)^2]=.3546.$$

η 之值與論據之分組不無關係,當組段很窄時,即分組分得過於精細時,每行只有一點,於是 $S_Y'=0$, 因之由式 (33) 得知 $\eta=1$. 從他方面言之,分組分得太粗率,能使 η 與 r 相近. 相關比僅對於分組適宜之論據纔有意義,此點不可不注意及之。

3. 直線性之測驗 在式 (34) 下面我們曾經提說過，$\eta_{YX}^2 - r^2$ 可當做 Y 倚 X 直線性測驗之準則. 現在我們討論爲什麼可以拿牠當做準則?

爲討論方便計，我們將上面推得的重要公式彙錄如下:

$$(40)\quad \begin{cases} S_{Y\cdot X}^2 = \dfrac{1}{\Phi_X}\sum\limits_Y (Y - {}_c\bar{Y}_X)^2 F_{XY}, & [\text{式 (31)}] \\ S_Y^2 = \dfrac{1}{N}\sum\limits_X S_{Y\cdot X}^2 \Phi_X, & [\text{定理 6}] \\ r^2 = 1 - \dfrac{S_Y^2}{\sigma_Y^2}, & [\text{式 (10)}] \end{cases}$$

$$(41)\quad \begin{cases} \sigma_{Y\cdot X}^2 = \dfrac{1}{\Phi_X}\sum\limits_Y (Y - {}_0\bar{Y}_X)^2 F_{XY}, & [\text{式 (36)}] \\ S_Y'^2 = \dfrac{1}{N}\Sigma \sigma_{Y\cdot X}^2 \Phi_X, & [\text{式 (37)}] \\ \eta_{YX}^2 = 1 - \dfrac{S_Y'^2}{\sigma_Y^2} = \dfrac{\sigma_{\bar{Y}_X}^2}{\sigma_Y^2}, & [\text{式 (35)}] \end{cases}$$

記得 $S_{Y\cdot X}^2$ 爲 X 行之變差，如此其爲對於迴歸線——配合各行平均數最美滿者——標準差之平方，S_Y^2 爲 $S_{Y\cdot X}^2$ 之加權平均數，並且 r^2 得用 S_Y^2 表之；同樣，$\sigma_{Y\cdot X}^2$ 爲在某一行中對於迴歸線——經過平均數者——標準差之平方，$S_Y'^2$ 爲 $\sigma_{Y\cdot X}^2$ 之加權平均數，並且 η^2 得用 $S_Y'^2$ 表之. 如果迴歸線爲直線，各行平均數將落在配合最美滿之直線上. 於是 $S_{Y\cdot X}^2 = \sigma_{Y\cdot X}^2$，並且 $S_Y^2 = S_Y'^2$，因之 $r^2 = \eta_{YX}^2$.

很可注意的是，$\sigma_{Y\cdot X}^2$ 既爲在 X 行中各 Y 對其平均數之變差，故 $\sigma_{Y\cdot X}^2$ 得用主要動差符號 μ_{2Y_X} 記之. 因此 $S_{Y\cdot X}^2$ 得用符號 λ_{2Y_X} 或 ν_{2Y_X} 記之.其爲在 X 行中各 Y 對於任意數之二級動差. 因爲 $\mu^2 \leq \nu^2$ (參閱第二章定理 7); 故 $\sigma_{Y\cdot X}^2 \leq S_{Y\cdot X}^2$，$S_Y'^2 \leq S_Y^2$，因之 $r^2 \leq \eta_{YX}^2$. 如

果屬於某行各個 Y 值與該行 Y 之平均數相等，很顯然的 $\sigma_{\bar{Y}_X}{}^2=0$. 在此情形中，$S_Y'^2=0$ 並且 $\eta_{Y\cdot X}{}^2=1$. 自他方言之，某行之離勢量數 $\sigma_{Y\cdot X}$ 不能大於全表之離勢量數 σ_Y，因之 $\sigma_{Y\cdot X}{}^2\leq\sigma_Y{}^2$，如此 $S_Y'^2\leq\sigma_Y{}^2$. 所以 $\eta_{YX}{}^2\leq 1$.

將式 (35) 寫做

$$S_Y'=\sigma_Y\sqrt{1-\eta_{YX}{}^2}.$$

可見 S_Y' 為對於迴歸線（其為平均數之軌跡）離勢之量度，以與 $S_Y=\sigma_Y\sqrt{1-r^2}$ 成對應，彼處 S_Y 為對於配合最美滿直線離勢之量度。如果 $r=1$，於是 Y 與 X 之關係為一直線函數 (linear functon). 如果 $\eta_{YX}{}^2=1$，則 Y 為 X 的單值函數 (single valued function). 反之，$r^2=0$，X 與 Y 不一定沒有關係：因為若 $\eta_{YX}{}^2=0$，則 $r^2=0$；但 $r^2=0$，$\eta_{YX}{}^2$ 未必等於零。

在一個最理想的相關表裏，祇當 $\eta_{YX}{}^2-r^2=0$ 時，Y 倚 X 迴歸線為一直線。但在一個觀察之相關表裏樣本之波動在所不免，所以由樣本算得 $\eta_{YX}{}^2-r^2$ 之值不一定就是零；然而 Y 與 X 果為一直線關係，其與零之差別不能大於因機會而起之波動。自然要發生問題，η^2 與 r^2 之差別應當如何我們纔能說迴歸線不是直線呢？為解決此問題柏可門 (J. Blakeman) 曾建議好幾個測驗公式，* 茲取其比較適用且易於計算者如下：

(42) $$\sigma_\delta=\frac{2}{\sqrt{N}}\left[\delta^{\frac{1}{2}}\{(1-\eta^2)^2-(1-r^2)^2+1\}^{\frac{1}{2}}\right].$$

彼處 $\delta=\eta^2-r^2$，σ_δ 為 δ 之標準差。

* Biometrika, vol. 4, pp. 332－350, 190 .

例 3. 測驗由表 62 裏論據算得之 Y 倚 X 迴歸線是否爲一直線。

由表 62 並用式 (42) 算得

$$\sigma_\delta = .2(.3546-.3364)^{\frac{1}{2}}\{(1-.3546)^2-(1-.3364)^2+1\}^{\frac{1}{2}}$$
$$=.2\times.1349\times.9880$$
$$=.0267.$$

但 $\delta=\eta_{YX}^2-r^2=.0182$, 於是 $\dfrac{\delta}{\sigma_\delta}=\dfrac{.0182}{.0267}=0.69$, 即 δ 僅爲其標準差之 0.69 倍,因知由表 62 求得 Y 倚 X 之迴歸線爲直線 (參閱下章第四節, 7).

因爲相關比有兩個,一爲 η_{YX}, 一爲 η_{XY}, 故有兩個測驗. 一條迴歸線爲直線,其他很可能爲曲線,須視測驗之結果如何而定.

第四節 二數聯合分配之一般情形

1. 基本概念 一個變數之機率分配函數已在第四章及第六章裏討論過了. 現在我們進一步討論兩個變數之機率分配. 如果在 XY 平面上一個區域內 $F(X, Y)$ 之二重積分 (double integral) 爲度量變數 (X, Y) 發現之相對頻數,則稱 $F(X, Y)$ 爲連續變數 (X, Y) 之聯合機率分配函數. 我們應當注意, $F(X, Y)$ 爲一連續單値並且非負的函數. 如果 (X, Y) 之値限於既定之區域內,我們說其在此區域外之 $F(X, Y)$ 完全爲零. 於是我們有

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(X, Y)dYdX=1.$$

就幾何上說,上列之積分代表曲面 $Z=F(X, Y)$ 下之體積,其値等於 1.

因此 $F(X, Y)dYdX$ 為 X 發現在組段 $(X, X+dX)$ 之內同時 Y 發現在組段 $(Y, Y+dY)$ 之內的機率,並且

$$\int_a^b \int_c^d F(X, Y)dYdX$$

代表 X 在 a 與 b 之間, Y 在 c 與 d 之間的機率.

我們分兩種情形研究:(a) 當機率上所說的二變數彼此各不相干時;(b) 當牠們彼此有關聯時.假設不論 Y 值為何, $\Phi(X)dX$ 為 X 在 dX 行內發現之機率.於是

$$\Phi(X)dX = dX\int_{-\infty}^{\infty} F(X, Y)dY. \tag{43}$$

很顯然的,由式 (43) 可求得 $\Phi(X)$, 因為 X 在任意組段 (a, b) 內發現之相對頻數為 X 在 XY 平面上組段 $a<X<b$ 內變數 (X, Y) 發現之相對頻數,如此

$$\int_a^b \int_{-\infty}^{\infty} F(X, Y)dYdX = \int_a^b \Phi(X)dX.$$

同理,如果不論 X 之值為何, $\Psi(Y)dY$ 為 Y 發現於 dY 內之機率,我們有

$$\Psi(Y)dY = dY\int_{-\infty}^{\infty} F(X, Y)dX. \tag{44}$$

依照習慣, $\Phi(X)$ 及 $\Psi(Y)$ 常被稱之為邊緣分配之機率函數 (probability function of marginal distribution).

2. 定義 X 與 Y 之不相干特性可用下列定義敍述之:當 $F(X, Y) = \Phi(X)\Psi(Y)$ 時,稱做二變數 X 及 Y 彼此各不相干;如果 $F(X, Y) \neq \Phi(X)\Psi(Y)$, 我們說 X 與 Y 有關聯.

3. 動差 假設

(45) $$\nu_{mn}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(X,Y)X^mY^n dYdX$$

爲對於 X 及 Y 公共原點之一般積差 (product moment).

如果 $m=0$ 及 $n=1$, 我們有

$$\nu_{01}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(X,Y)YdYdX.$$

設 $F(X,Y)$ 爲一可以互換積分次序之函數.於是 ν_{01} 變做

$$\nu_{01}=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}F(X,Y)dX\right]YdY=\int_{-\infty}^{\infty}\Psi(Y)YdY,$$

其爲各 Y 值之平均數 $\bar{Y}$. 同理,各 X 值之平均數 $\bar{X}$ 爲

$$\nu_{10}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(X,Y)XdYdX=\int_{-\infty}^{\infty}\Phi(X)XdX.$$

我們現在可將對於平均數 $(\bar{X},\bar{Y})$ 之積差予以定義如下:

(46) $$\mu_{mn}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(X-\bar{X})^m(Y-\bar{Y})^nf(X,Y)dYdX.$$

若使 $m=n=1$, 則由上式得聯合分配之共差 (co-variance)

$$\mu_{11}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(X-\bar{X})(Y-\bar{Y})F(X,Y)dYdX;$$

若使 $m=2$, $n=0$; 則得 X 之變差

$$\mu_{20}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(X-\bar{X})^2F(X,Y)dYdX$$

$$=\int_{-\infty}^{\infty}(X-\bar{X})^2\Phi(X)dX=\sigma_X^2;$$

並且若使 $m=0$, $n=2$; 則得 Y 之變差

$$\mu_{02}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(Y-\bar{Y})^2F(X,Y)dYdX$$

$$=\int_{-\infty}^{\infty}(Y-\bar{Y})^2\Psi(Y)dY=\sigma_Y^2.$$

由是不難證明下列之關係:

$$(47)\qquad \begin{cases}\mu_{11}=\nu_{11}-\nu_{10}\nu_{01}\\ \mu_{20}=\nu_{20}-\nu_{10}^2\\ \mu_{02}=\nu_{02}-\nu_{01}^2\end{cases}$$

若以 ρ_{XY} 表示 X 與 Y 之相關係數,則

$$(48)\qquad \rho_{XY}=\frac{\mu_{11}}{\sigma_X\sigma_Y}.$$

4. 迴歸線 如將聯合機率函數 $F(X,Y)$ 中之 Y 值固定,則由式(44)得

$$\frac{1}{\Psi(Y)}\int_{-\infty}^{\infty}F(X,Y)dX=1.$$

即是

$$\int_{-\infty}^{\infty}\frac{F(X,Y)}{\Psi(Y)}dX=1.$$

因此,X 在無限小組段內之機率爲 $\frac{F(X,Y)}{\Psi(Y)}dX$; 並且 $\frac{F(X,Y)}{\Psi(Y)}$ 通常被稱之爲 Y 値固定後 X 之機率函數 (probability function), 或者稱爲在 Y 列內各 X 之機率密度 (probability density).

同樣, $F(X,Y)/\Phi(X)$ 被稱之爲 X 値固定後 Y 之機率函數,或稱爲在 X 行內各 Y 之機率密度,因爲,當 X 被固定時,積分

$$\int_{-\infty}^{\infty}\frac{F(X,Y)}{\Phi(X)}dY=1.$$

在 X 行內各 Y 値之平均數爲

$$(49)\qquad \bar{Y}_X=\int\frac{YF(X,Y)}{\Phi(X)}dY,$$

彼處積分要經過在 X 行內所有各 Y 值. 同理, Y 列各 X 值之平均數為

$$(50)\qquad \bar{X}_Y=\int\frac{XF(X,Y)}{\Psi(Y)}dX,$$

彼處積分經過在 Y 列內所有各 X 值.

在 X 行各 Y 值之變差

$$\sigma_{YX}{}^2=\int(Y_X-\bar{Y}_X)^2\frac{F(X,Y)}{\Phi(X)}dY.$$

彼處積分經過 X 行內所有各 Y 值.

同理, 在 Y 列各 X 值之變差

$$\sigma_{XY}{}^2=\int(X_Y-X_Y)^2\frac{F(X,Y)}{\Psi(Y)}dX.$$

彼處積分經過 Y 列內所有各 X 值.

每取一 X 行則得一平均數 $\bar{Y}_X$, 連續變化 X 可得各平均數之軌跡, 此即 Y 倚 X 之迴歸線. 此迴歸線之方程式為式 (49), 自然彼處之 X 為變數. 同理式 (50) 為 X 倚 Y 之迴歸線. 迴歸線之意義最大並且應用最廣者厥為直線.

茲特就迴歸線為直線時加以討論. 如果 Y 倚 X 迴歸線之形式為

$$Y_X=AX+B,$$

則稱此 Y 倚 X 之迴歸線為直線. 同理, 如果 X 倚 Y 迴歸線方程式之形式為

$$X_Y=CY+D,$$

則稱此 X 倚 Y 之迴歸線為直線. 一條迴歸線為直線, 他條未必為直線 (見前節).

現在我們研究當迴歸線爲直線時聯合機率函數 $F(X,Y)$ 及邊和機率函數 $\Phi(X)$ 及 $\Psi(Y)$ 之涵義.假設 $\bar{Y}_X$ 與 X 之關係爲一直線,即

$$\int_{-\infty}^{\infty}\frac{YF(X,Y)}{\Phi(X)}dY=A+BX;$$

或者

(51) $$\int_{-\infty}^{\infty}YF(X,Y)dY=A\Phi(X)+BX\Phi(X).$$

將上式兩邊乘以 dX,並且對於 X 積分之;我們有

$$\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}YF(X,Y)dY\right]dX=A\int_{-\infty}^{\infty}\Phi(X)dX+B\int_{-\infty}^{\infty}X\Phi(X)dX,$$

即是

(52) $$\nu_{01}=A+B\nu_{10}.$$

將式 (51) 兩邊乘以 XdX 並且對於 X 積分之,我們有

$$\int_{-\infty}^{\infty}X\left[\int_{-\infty}^{\infty}YF(X,Y)dY\right]dX=A\int_{-\infty}^{\infty}X\Phi(X)dX+B\int_{-\infty}^{\infty}X^2\Phi(X)dX.$$

因爲上式左方爲 ν_{11},於是我們有

(53) $$\nu_{11}=A\nu_{10}+B\nu_{20}.$$

將式 (52) 及 (53) 當做 A 及 B 之聯立方程式解之,得

$$A=\nu_{01}-\nu_{10}\frac{\sigma_Y}{\sigma_X}\rho=\bar{Y}-\bar{X}\frac{\sigma_Y}{\sigma_X}\rho,$$

$$B=\frac{\nu_{11}-\nu_{10}\nu_{01}}{\nu_{20}-\nu_{10}^2}=\frac{\mu_{11}}{\mu_{20}}=\rho\frac{\sigma_Y}{\sigma_X}.$$

所以 Y 倚 X 迴歸線之方程式 (50) 變做

(54) $$\bar{Y}_X-\bar{Y}=\rho\frac{\sigma_Y}{\sigma_X}(X-\bar{X}).$$

如果 X 倚 Y 迴歸線爲直線,同樣可得 X 倚 Y 迴歸線之方程式爲

(55) $$X_Y-\bar{X}=\rho\frac{\sigma_X}{\sigma_Y}(Y-\bar{Y}).$$

數值 $A=\rho\frac{\sigma_Y}{\sigma_X}$ 及 $B=\rho\frac{\sigma_X}{\sigma_Y}$ 各叫做迴歸係數. 很顯然的，其積等於

例　假設以已給　$F(X,Y)=\frac{2}{a^2}$
當做二數聯合機率函數，彼處 $0\leq X\leq Y$ 並且 $0\leq Y\leq a$. 求 (i) $\Phi(X)$ 及 $\Psi(Y)$; (ii) $\bar{X}$, $\bar{Y}$, σ_X^2 及 σ_Y^2; (iii) Y 倚 X 及 X 倚 Y 迴歸線之方程式; (iv) 相關係數 ρ.

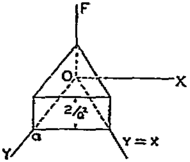


圖 59.　$F(X,Y)=\frac{2}{a^2}$之圖。

因為已給曲面 $F(X,Y)=\frac{2}{a^2}$ 下體積為

$$\int_0^a\int_0^Y\frac{2}{a^2}dXdY=\frac{2}{a^2}\int_0^a YdY$$

$$=\frac{2}{a^2}\times\frac{a^2}{2}=1.$$

故曲面之圖形如圖 59 所示.

(i) 邊和分配為

$$\Phi(X)=\int_X^a\frac{2}{a^2}dY=\frac{2}{a^2}(a-X),$$

$$\Psi(Y)=\int_0^Y\frac{2}{a^2}dX=\frac{2Y}{a^2}.$$

(ii) 平均數

$$\bar{X}=\int_0^a X\frac{2}{a^2}(a-X)dX=\frac{a}{3},$$

$$\bar{Y}=\int_0^a Y\frac{2Y}{a^2}dY=\frac{2a}{3}.$$

因爲

$$\nu_{20}=\int_0^a X^2\frac{2}{a^2}(a-X)dX=\frac{a^2}{6},$$

$$\nu_{02}=\int_0^a Y^2\frac{2}{a^2}YdY=\frac{a^2}{2}.$$

並且已求得

$$\bar{X}=\frac{a}{3},$$

$$\bar{Y}=\frac{2a}{3}.$$

故變差

$$\sigma_X^2=\frac{a^2}{6}-\frac{a^2}{9}=\frac{a^2}{18},$$

$$\sigma_Y^2=\frac{a^2}{2}-\frac{4a^2}{9}=\frac{a^2}{18}.$$

(iii) 迴歸線爲

$$\bar{Y}_X=\int_X^a Y\frac{\frac{2}{a^2}}{\frac{2(a-X)}{a^2}}dY=\frac{a+X}{2},$$

$$\bar{X}_Y=\int_0^Y X\frac{\frac{2}{a^2}}{\frac{2Y}{a^2}}dX=\frac{Y}{2}.$$

(iv) 由迴歸線得 $\rho^2=\frac{1}{4}$. 因 $\rho\frac{\sigma_Y}{\sigma_X}$ 爲正數,故 $\rho=\frac{1}{2}$.

5. 估計標準誤 我們知道,在 X 行裏各 Y 值之機率密度爲 $\frac{F(X,Y)}{\Phi(X)}$, 於是此行之變差

$$S_{Y\cdot X}^2=\int_{-\infty}^{\infty}(Y-{}_e\bar{Y}_X)^2\frac{F(X,Y)}{\Phi(X)}dY$$

每有一行即有一數值 $S_{Y\cdot X}^2$，若將此值各以該行頻數權之，則得 S_Y^2，S_Y 即所稱之估計標準誤.我們可證 $S_Y^2=\sigma_Y^2(1-\rho^2)$.

證：因爲 $$S_Y^2=\int_{-\infty}^{\infty}\Phi(X)S_{Y\cdot X}^2dX$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(Y-\bar{Y}_X)^2F(X,Y)dYdX.$$

將式(54)所給之 $\bar{Y}_X$ 值代入上式，我們有

$$S_Y^2=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{(Y-\bar{Y})-\rho\frac{\sigma_Y}{\sigma_X}(X-\bar{X})\right\}^2F(X,Y)dYdX$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{(Y-\bar{Y})^2-2\rho\frac{\sigma_Y}{\sigma_X}(X-\bar{X})(Y-\bar{Y})+\rho^2\frac{\sigma_Y^2}{\sigma_X^2}(X-\bar{X})^2\right\}$$

$$F(X,Y)dYdX$$

$$=\sigma_Y^2(1-\rho^2).$$ 證訖.

6 常態相關曲面之方程式 設二數相關曲面之方程式爲

$$Z=f(x,x),$$

彼處 x 及 y 各爲變數 X 及 Y 各以組距 h 及 k 爲單位之離均差，卽 $x=\frac{X-\bar{X}}{h}$，$y=\frac{Y-\bar{Y}}{k}$ 此曲面方程式可分兩種情形推求之：第一種情形係假設 x 與 y 各不相干，第二種情形係假設 x 與 y 彼此有關聯.

第一種情形 設二變數 X 及 Y 之分配各合乎常態定律.於是變數 X 之離均差發現於 x 與 $x+\delta x$ 間的機率爲

$$\frac{\delta x}{\sigma_x\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_x^2}}$$

並且變數 Y 之離均差發現於 y 與 δy 間的機率爲

$$\frac{\delta y}{\sigma_y\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma_y^2}}$$

因之 X 與 Y 各自的離均差聯合發現於 x 與 $x+\delta x$ 及 y 與 $y+\delta y$ 間的機率爲

$$\left(\frac{\delta x}{\sigma_x\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_x^2}}\right)\left(\frac{\delta y}{\sigma_y\sqrt{2\pi}}e^{-\frac{y^2}{2\sigma_y^2}}\right)=\frac{\delta x\delta y}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}$$

(參閱本節 1 之定義).

設 N 爲觀察值之個數,則聯合發現於 x 與 $x+\delta x$ 及 y 與 $y+\delta y$ 間之頻數爲

$$\frac{N\delta x\delta y}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}.$$

以 $Z\delta x\delta y$ 代表此頻數,我們得到各不相干二變數頻數曲面之方程式

$$Z=\frac{N}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}. \tag{56}$$

第二種情形　現在我們求兩個相關變數頻數曲面之方程式. 假設 $y=r\frac{\sigma_y}{\sigma_x}x$ 爲 y 倚 x 之迴歸線,並設 ζ 爲 y 之計算值與其觀察值之差. 於是每有一點 (x, y), 卽有一點 (x, ζ) 與之成對應. 但由定義得知 X 與 ζ 之相關係數

$$r_{x\zeta}=\frac{\frac{1}{N}\Sigma x\zeta}{\sigma_x\sigma_\zeta}.$$

因

$$\Sigma x\zeta=\Sigma x\left(y-r\frac{\sigma_y}{\sigma_x}x\right).$$

$$=\Sigma xy-r\frac{\sigma_y}{\sigma_x}\Sigma x^2$$

$$=Nr\sigma_x\sigma_y-Nr\sigma_x\sigma_y.$$

故 $$r_{x\zeta}=0.$$

此示二變數 x 與 ζ 各不相干，因之二數 x 與 ζ 共同發現之機率很容易由式 (56) 寫出，因其等於 X 的離均差發現於 x 與 $x+\delta x$ 間之機率乘以 ζ 離均差發現於 ζ 與 $\zeta+\delta\zeta$ 間之機率的相乘積. 即

$$\left(\frac{\delta x}{\sigma_x\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma_x^2}}\right)\left(\frac{\delta\zeta}{\sigma_\zeta\sqrt{2\pi}}e^{-\frac{\zeta^2}{2\sigma_\zeta^2}}\right)$$

$$=\frac{\delta x\delta\zeta}{2\pi\sigma_x\sigma_\zeta}e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{\zeta^2}{\sigma_\zeta^2}\right)}$$

$$=\frac{\delta x\delta\zeta}{2\pi\sigma_x\sigma_\zeta}e^{-\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2}+\frac{\left(y-r\frac{\sigma_y}{\sigma_x}x\right)^2}{\sigma_\zeta^2}\right\}}$$

$$=\frac{\delta x\delta\zeta}{2\pi\sigma_x\sigma_\zeta}e^{-\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_\zeta^2}-\frac{2r\sigma_y xy}{\sigma_x\sigma_\zeta^2}+\frac{r^2\sigma_y^2x^2}{\sigma_x^2\sigma_\zeta^2}\right\}}$$

$$=\frac{\delta x\delta\zeta}{2\pi\sigma_x\sigma_\zeta}e^{-\frac{1}{2}\left\{x^2\left(\frac{1}{\sigma_x^2}+\frac{r^2\sigma_y^2}{\sigma_x^2\sigma_\zeta^2}\right)-2rxy\frac{\sigma_y}{\sigma_x\sigma_\zeta^2}+\frac{y^2}{\sigma_\zeta^2}\right\}}$$

$$=\frac{\delta x\delta\zeta}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}e^{-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_x^2}-\frac{2rxy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2}\right)},$$

[因為我們可證 $\sigma_\zeta^2=\sigma_y^2(1-r^2)$]

類似上面，用 $Z\delta x\delta\zeta$ 代表 X 與 ζ 各自離均差 x 至 $x+\delta x$, ζ 至 $\zeta+\delta\zeta$ 聯合發現之頻數，我們最後有

$$(57)\qquad Z=\frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}e^{-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_x^2}-\frac{2rxy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2}\right)}.$$

此即所欲求之常態曲面方程式.

若 X 與 Y 各不相干，則 $r=0$. 因之式 (57) 變成了式 (56).

7. 常態曲面之性質　常態曲面之重要性質可用下列定理敘述之.

定理 7. 如果二變數 x 及 y 常態相關，則 x 之邊緣分配 $\phi(x)$ 及 y 之邊緣分配 $\psi(y)$ 均爲常態.

證：因

$$\phi(x)=\int_{-\infty}^{\infty}f(x,y)dy$$

$$=\frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}e^{-\frac{x^2}{2\sigma_x^2}}\int_{-\infty}^{\infty}e^{-\frac{1}{2(1-r^2)}\left\{\frac{y}{\sigma_y}-\frac{rx}{\sigma_x}\right\}^2}dy$$

$$=\frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}e^{-\frac{x^2}{2\sigma_x^2}}\int_{-\infty}^{\infty}e^{-\frac{\zeta^2}{2\sigma_y^2(1-r^2)}}d\zeta,$$

$$\left(\text{因 } \zeta=y-\frac{r\sigma_y}{\sigma_x}x\right)$$

$$=\frac{N}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}e^{-\frac{x^2}{2\sigma_x^2}}\left\{2\pi(1-r^2)\right\}^{\frac{1}{2}}\sigma y$$

$$=\frac{N}{\sqrt{2\pi}\sigma_x}e^{-\frac{x^2}{2\sigma_x^2}};$$

同理可證

$$\psi(y)=\int_{-\infty}^{\infty}f(x,y)dx$$

$$=\frac{N}{\sqrt{2\pi}\sigma_y}e^{-\frac{y^2}{2\sigma_y^2}}$$

這就是說，如果二變數爲常態相關，每個變數之邊緣分配爲常態.

證訖.

定理 8. 常態相關曲面之迴歸線爲直線.

啓示　由 $\bar{y}_x=\int_{-\infty}^{\infty}\frac{yf(x,y)}{\phi(x)}dx$ 及 $\bar{x}_y=\int_{-\infty}^{\infty}\frac{xf(x,y)}{\Psi(y)}dx$ 證起. 後步留待學生自行證明.

定理 9. 如果 X 與 Y 爲常態相關，X 行之各頻數爲常態分配，其變差 $S_Y'^2$ 不因 X 值之變而變，並且 $S_Y'^2=\sigma_Y^2(1-r^2)$；Y 列之各頻數爲常態分配,其變差 $S_X'^2$ 不因 Y 值之變而變,並且 $S_X'^2=\sigma_X^2(1-r^2)$. 此定理之證明留待學生練習.

第五節. 其他相關

I. 等級相關.

a. 差量之變差 差量之變差與求等級相關係數公式有關，因此於推演計算相關係數之前，我們求出一公式，表示 $(X-Y)$ 之變差. 讓 X 及 Y 表示兩個對應變數，各有 n 個變量；並且讓 Z 表示如此做成之第三個變數，彼處 $Z_i=X_i-Y_i$. 於是 Z 之平均數 $\bar{Z}=\bar{X}-\bar{Y}$，並且由定義 Z 之變差

$$\sigma_Z^2=\frac{1}{N}\Sigma Z^2-\bar{Z}^2$$

以 $X-Y$ 代 Z, 上式變做

$$\sigma_Z^2=\frac{1}{N}\Sigma(X-Y)^2-(\bar{X}-\bar{Y})^2$$

$$=\frac{1}{N}\Sigma(X^2-2XY+Y^2)-\bar{X}^2-\bar{Y}^2+2\bar{X}\bar{Y}$$

$$=\left\{\frac{1}{N}\Sigma X^2-\bar{X}^2\right\}-2\left\{\frac{1}{N}\Sigma XY-\bar{X}\bar{Y}\right\}+\left\{\frac{1}{N}\Sigma Y^2-\bar{Y}^2\right\}$$

$$=\sigma_X^2-2r\sigma_X\sigma_Y+\sigma_Y^2. \tag{58}$$

此即所欲求者.

如果 X 與 Y 無關，我們得式 (58) 之特殊情形:

(59) $$\sigma_Z^2=\sigma_X^2+\sigma_Y^2.$$

以式 (58) 中之 r 當做未知數解之，於是得

(60) $$r=\frac{\sigma_X^2+\sigma_Y^2-\sigma_Z^2}{2\sigma_X\sigma_Y}.$$

此為相關係數之另一表示法，式內祇含標準差。在某種情形中，X 及 Y 以等級表示時，式 (60) 頗可利用。

b. 公式　所謂等級者乃係按數值大小或品質高低所排列之次第也；換言之，等級者為變量排列後之地位也。如果 X 及 Y 表示同一事物具有兩種特徵，則等級係續聯且不重複之數值，即是 X 及 Y 皆為由 1 至 N 之連續整數。故 $\bar{X}=\bar{Y}$，$\sigma_X^2=\sigma_Y^2=\frac{1}{12}(N^2-1)$，見第二章定理 14。更進而言之

$$\begin{aligned}\sigma_Z^2&=\frac{1}{N}\Sigma Z^2-\bar{Z}^2\\&=\frac{1}{N}\Sigma(X-Y)^2-(\bar{X}-\bar{Y})^2\\&=\frac{1}{N}\Sigma(X-Y)^2.\ (\text{因 }\bar{X}-\bar{Y}=0).\end{aligned}$$

當 X 及 Y 表示變數等級而非變數本身時，其相關程度通常用 ρ 表示之。因是式 (60) 變做

$$\rho=\frac{2\left(\frac{N^2-1}{12}\right)-\frac{1}{N}\Sigma(X-Y)^2}{2\left(\frac{N^2-1}{12}\right)}$$

(61) $$=1-\frac{6\Sigma(X-Y)^2}{N(N^2-1)}.$$

如果有兩個或多個事物具有特徵之量度相同，可予以個別等級，或

予以相同等級而附以分數，例如有兩個相同變量予以相同等級但在表等級之數值後面各附以 $\frac{1}{2}$；若有三個相同變量，予以相同等級但在表示等級數值後面各附以 $\frac{1}{3}$；等等。

例 1. 假設我們有兩種測驗結果，已按其成績等級排列，計算此論據之等級相關。

表 67. 計算等級相關係數

學生	第一種測驗		第二種測驗		$X-Y$	$(X-Y)^2$
	分數	等級 X	分數	等級 Y		
A	92	1	85	2	-1	1
B	86	2	76	4	-2	4
C	84	3	93	1	2	4
D	78	4	68	6	-2	4
E	71	5	67	7	-2	4
F	69	6	83	3	3	9
G	66	7	54	9	-2	4
H	58	8	70	5	3	9
I	53	9	43	10	-1	1
J	45	10	59	8	2	4
總計	...	...	...	...	0	44

應用式 (61)，我們有

$$\rho=1-\frac{6\times 44}{10\times 99}=0.733.$$

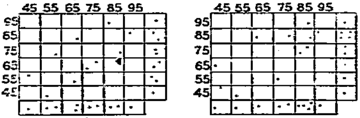
c. ρ 與 r 之比較 根據表 67 並應用公式 (3) 算得之相關係數 $r=.729$，此值與方程算得之 ρ 值相比，則知二者頗相近；但用此兩種方

法算得之結果未必常是如此.例如若將表 67 裏第一種測驗分數 92、86、84、78、71、69、66、58、53 及 45 各改做 92、89、87、86、83、77、71、62、53 及 40,第二種測驗分數 85、76、93、68、67、83、54、70、43 及 59 各改做 88、85、93、79、70、87、52、84、41 及 64,則其等級不變,因之 ρ 之值仍為 .733,但 r 之值變為 .636 矣.

當用公式求等級相關時,變數分配之性質不可輕忽.表示 ρ 與 r 關係之種種公式,均限於變數分配在某種情形時方能適用.皮爾生公式

$$r = 2\sin\left(\frac{\pi}{6}\rho\right) \tag{62}$$

限於在常態分配時用之.現在我們將上面所說的兩種分配分別繪成散佈圖,並比較其邊緣總計(marginal total).由圖 60 a 可見表 67 裏之邊緣總計呈對稱之象,但將測驗分數改易後則否(見圖 60 b).因此,論據如圖 60 b 情形時不得援用皮爾生公式由已算得之 ρ 值以求 r 之值.



a. 原來論證　　b. 改變後論據

圖 60. 兩種測驗成績之頻數分配。

學者應當牢記,由算得之 r 或 ρ 祇能作宇宙真正值之估計耳,根據估計之結果不過僅知宇宙真正值之輪廓而已.我們可以這樣總括說,ρ 值僅僅表示二數相關之存在,而不能表示其確切程度.

2. 品質相關.

a. 列聯係數 列聯係數爲皮爾生所創。茲將計算列聯係數敍述於下:

設 a_{ij} 爲第 i 行第 j 列之觀察頻數; A_{ij} 爲與 a_{ij} 成對應之理論頻數: $X^2=\sum_{i=1}^{s}\sum_{j=1}^{t}\frac{(a_{ij}-A_{ij})^2}{A_{ij}}$, 自然式中之 $A_{ij}=\frac{a_i a_j}{N}$; $\phi^2=\frac{X^2}{N}$. 於是

(63) $$C=\sqrt{\frac{\phi^2}{1+\phi^2}}.$$

彼處 C 通常稱之爲列聯係數 (coefficient of contingency).

因爲 $\phi^2=\frac{X^2}{N}$, 故式 (63) 可以寫做

(64) $$C=\sqrt{\frac{X^2}{N+X^2}}.$$

例 2. 根據下表裏所列之論據並應用式 64, 計算列聯係數.

表 68. 6800 男人之髮色與睛色

睛色	髮色				總計
	黃	褐	黑	紅	
藍	1768 (1169.5)	807 (1088.0)	189 (505.5)	47 (48.0)	2811 (2811.0)
碧	946 (1303.0)	1387 (1212.3)	746 (563.3)	53 (53.4)	3132 (3132.0)
褐	115 (356.5)	438 (331.7)	288 (154.2)	16 (14.6)	857 (857.0)
總計	2829 (2829.0)	2632 (2632.0)	1223 (1223.0)	116 (116.0)	6800 (6800.0)

上表括號裏之數值爲算得之理論頻數.

表 69. 計算 $\frac{(a_{ij}-A_{ij})^2}{A_{ij}}$

瞳 色	髮色				總 計
	黃	褐	黑	紅	
藍	303.6	72.6	198.2	0.0	577.1
碧	97.8	25.2	59.3	0.0	182.3
褐	163.7	34.1	116.1	0.1	313.9
總 計	567.7	131.9	373.6	0.1	$1073.3 = X^2$

於是算得 $$C=\sqrt{\frac{1073.3}{7873.3}}=0.37.$$

此示髮色與瞳色有聯帶關係.

b. 四格表 表之形式如下面所列者叫做四格表.

表 70. 四格表之一般形式

品質 B	品質 A		總 計
	+	−	
+	a	b	$a+b$
−	c	d	$c+d$
總 計	$a+c$	$b+d$	$a+b+c+d=N$

度量品質 A 與品質 B 聯繫強度之係數曾有多種,不克一一敍述.*

* 學者欲知其詳可參閱下列之刊物:
Proc. Roy. Soc. of Medicine, vol. 8, 1915, p. 113.
Phil. Trans. Roy. Soc. Series A, vol. 194, 1900, p. 257.
Biometrika, vol. 2, 1903, p. 121.
Biometrika, vol. 9, 1913, pp 159-332.
Jour. Roy. Stat. Soc., vol. 75, 1912, pp. 579-642.

茲舉其簡單者如下：

$$
(5)\qquad \begin{cases} Q=\dfrac{ad-bc}{ad+bc}, \\[2ex] \omega=\dfrac{\sqrt{ad}-\sqrt{bc}}{\sqrt{ad}+\sqrt{bc}}, \\[2ex] r=\cos\dfrac{\sqrt{bc}}{\sqrt{ad}+\sqrt{bc}}\pi; \end{cases}
$$

彼處 Q 通常稱做相聯係數 (coefficient of association), ω 通常稱做束縛係數(coefficient of colligation), 並且 r 稱做四格表相關係數(fourfold correlation coefficient).

上列三式有一致之特點：

(i) 如果 a 或 d 為 0, 則 $Q=\omega=r=-1$;

(ii) 如果 $ad=bc$, 則 $Q=\omega=r=0$;

(iii) 如果 b 或 c 為 0, 則 $Q=\omega=r=1$;

(iv) 如果在普通情況之下, Q、ω 及 r 之值在 -1 與 $+1$ 之間. 實在說這些係數就是根據此四特點創製者.

第六節　複相關

1. 符號　前五節所研究者係有關二數相關各問題. 在二數相關情形中卽或倘含有其他因子, 因二變數之重要性質特別顯著, 其他得以忽略之. 但是在農業、生物、經濟、教育及心理學上之論據, 常會有兩種以上之變數互相影響, 其影響程度無畸輕畸重之分. 因之研究三個或以上變數之相關勢爲必須. 例如農品產量之與土壤肥沃、雨量多寡、溫度高低等

之關係;學生成績之與天才、健康、學習鐘點等之關係;胸圍之與身長、體重等之關係是.

複相關內所研究之問題爲三個或以上變數之相關理論及應用. 茲爲論述便利計,我們研究之相關問題暫以三個變數爲限,然而所用之研究方法得以推廣之.如果三個變數相關問題之公式一經成立,則立刻可使其普遍化,推得 k 個變數之公式.

在兩個變數 X 及 Y 情形中,我們可用平行於坐標軸之直線將 XY 平面分成若干小長方格.同樣若有三個變數 X、Y 及 Z, 我們可用平行於坐標軸之平面將長方體切成若干小長方塊.

我們用小長方塊中心表示坐標 (X, Y, Z), 並用 F_{XYZ} 表示在此中心點之頻數.一雙 (X, Y) 值固定一 Z 柱體 (column) (圖 61). 此柱體內各頻數之和

(66) $$\sum_Z F_{XYZ} = F_{XY}$$

通常叫做柱和 (column total); 彼處符號 $\sum$ 之下所錄之變數 Z, 表示沿

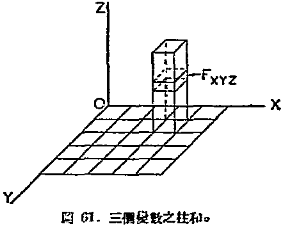


圖 61. 三個變數之柱和。

若該變數方向各頻數之和以下同此．現在我們思攷所有各 Y 值相同之柱體；此等柱體內各頻數之和

$$\sum_{X} F_{XY} = F_{Y} \tag{67}$$

可以稱之爲壁和 (slab total)（圖 62）．

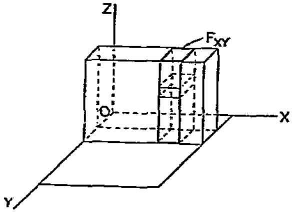


圖 62．三個變數之壁和。

最後，如果我們將各壁和加之，則得總頻數 N．如此

$$\sum_{Y} F_{Y} = N. \tag{68}$$

借助式（68）我們可用二重和（double sum）

$$\sum_{X}\sum_{Z} F_{XYZ} = F_{Y} \tag{69}$$

表示式（67）．再借助式（69）可用三重和（triple sum）

$$\sum_{X}\sum_{Y}\sum_{Z} F_{XYZ} = N$$

表示式（68）．

如果我們臆想各頻數之值（柱和）寫在 XY 平面方格中心點處，並且將各頻數（各柱和）集合之而成一二量相關表（圖 63），此表之相關係數 r_{XY}

叫做總相關係數(參閱第二節 2),所以別於後來欲研究之複相關及淨相關係數;並且此迴歸線叫做 Y 倚 X 及 X 倚 Y 之總迴歸線(total regression curve).

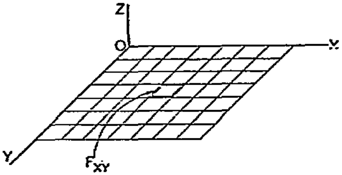


圖 63. 在方格(X,Y)內之總頻數。

依同樣情形,一雙 (Y, Z) 值固定一平行於 OX 之柱體. 在此柱體——X 柱體——內各頻數之和爲

$$\sum_X F_{XYZ}=F_{YZ};$$

如使 Z 不變並將各柱體內之頻數加之,則得一垂直 OZ 之壁體,含於其內各頻數之和爲

$$\sum_Y F_{YZ}=F_Z;$$

最後將此垂直壁體內之頻數加之得

$$\sum_Z F_Z=N.$$

如果我們將 F_{YZ} 各值之和寫在 YZ 平面方格中心點處,如此便成一個二數相關表,如圖 64 所示. r_{YZ} 爲此表 Y 與 Z 之總相關係數,並且其迴歸線爲 Y 倚 Z 及 Z 倚 Y 之總迴歸線.

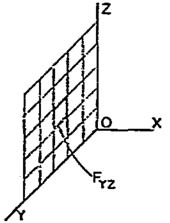


圖 64. 在方格(Y,Z)內之總頻數。

再依同樣情形，一雙 (X, Z) 值固定一 Y 柱體平行於 OY. 在此柱體內頻數之和爲

$$\sum_{Y} F_{XYZ} = F_{XZ};$$

如使 X 不變並將所有柱和加之，則得一垂直於 OX 之壁體並且含於其內各頻數之和爲

$$\sum_{Z} F_{XZ} = F_{X};$$

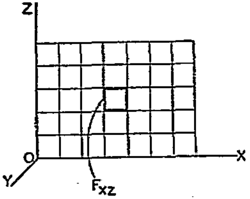


圖 65. 在方格(X,Z)內之總一頻數。

如將此等壁和加之，則得

$$\sum_X F_Z = N.$$

在 XZ 平面上各柱和 F_{XZ} 組成一個二量相關表，其相關係數 r_{XZ} 爲 X 與 Z 之總相關係數；並且此表之迴歸線爲 X 倚 Z 及 Z 倚 X 之總迴歸線（圖 65）．

2. 迴歸線　立在 (X, Y) 點上柱體各 Z 之平均數 $\bar{Z}_{XY}$ 之定義爲

$$\bar{Z}_{XY} = \frac{1}{F_{XY}} \sum Z F_{XYZ}.$$

同樣立在 (Y, Z) 點上柱體各 X 之平均數 $\bar{X}_{YZ}$ 之定義爲

$$\bar{X}_{YZ} = \frac{1}{F_{YZ}} \sum_X X F_{XYZ}$$

並且立在 (X, Z) 點上柱體各 Y 之平均數 $\bar{Y}_{XZ}$ 之定義爲

$$\bar{Y}_{XZ} = \frac{1}{F_{XZ}} \sum_Y Y F_{XYZ}.$$

Z 倚 X 與 Y 之迴歸面係一平面，其爲依最小平方原則配合各 Z 柱體平均數最美滿之平面．Y 倚 X 與 Z 及 X 倚 Y 與 Z 之迴歸面均係依相同原則所得之平面．

三個變數最好以習慣上通用之字母表示之，以便將所得之方程式推廣．因此我們以 X_1 代 Z、以 X_2 代 X、以 X_3 代 Y，並以 r_{23} 代 r_{XY}、以 r_{31} 成 r_{YZ}，以 r_{12} 代 r_{ZX}．

我們現在推演 X_1 倚 X_2 及 X_3 迴歸面之方程式，並根據最小平方法原則測定方程式內所含常數之值．爲簡便計，我們取各變數之平均數做

原點,即使 $x_1=X_1-\bar{X}_1$, $x_2=X_2-\bar{X}_2$, $x_3=X_3-\bar{X}_3$ 加上這個限制與方程式之普遍性無損.讓所求之方程式爲

(70) $$x_1=A+Bx_2+Cx_3.$$

於是剩餘數之平方和

$$U=\sum_{1,2,3}(x_1-A-Bx_2-Cx_3)^2f,$$

彼處 f 爲 f_{x_1,x_2,x_3} 之略,$f_{x_1,x_2,x_3}=F_{X_1,X_2,X_3}$; 並且 $\sum\limits_{1,2,3}$ 爲 $\sum\limits_{x_1}\sum\limits_{x_2}\sum\limits_{x_3}$ 之略.如果將 U 縮至最小,則 U 對於 A、B 及 C 之第一次偏微係數須各等於零.因之我們有

$$\sum_{1,2,3}(x_1-A-Bx_2-Cx_3)f=0,$$

$$\sum_{1,2,3}x_2(x_1-A-Bx_2-Cx_3)f=0,$$

$$\sum_{1,2,3}x_3(x_1-A-Bx_2-Cx_3)f=0.$$

如果以平均數爲原點,則 $\sum\limits_{1,2,3}x_1f=\sum\limits_{1,2,3}x_2f=\sum\limits_{1,2,3}x_3f=0$. 於是末三式之首式變作 $A=0$, 並且次二式可寫成下列之形式:

(71) $$\begin{cases}B\sum\limits_{1,2,3}x_2^2f+C\sum\limits_{1,2,3}x_2x_3f=\sum\limits_{1,2,3}x_1x_2f,\\ B\sum\limits_{1,2,3}x_2x_3f+C\sum\limits_{1,2,3}x_3^2f=\sum\limits_{1,2,3}x_1x_3f.\end{cases}$$

讓 σ_i^2 爲 x_i 之變差,並且 r_{ij} 爲 x_i 與 x_j 之相關係數.故由定義得

$$\sum_{1,2,3}x_i^2f=N\sigma_i^2,$$

$$\sum_{1,2,3}x_ix_jf=N\sigma_i\sigma_jr_{ij}.$$

如此式(71)各變做

$$B\sigma_2^2 + C\sigma_2\sigma_3 r_{23} = \sigma_1\sigma_2 r_{12},$$

$$B\sigma_2\sigma_3 r_{23} + C\sigma_3^2 = \sigma_1\sigma_3 r_{13}.$$

以 B 及 C 當做未知數解之,我們有

$$B = \frac{\sigma_1}{\sigma_2}\,\frac{\begin{vmatrix} r_{12} & r_{23} \\ r_{13} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{23} \\ r_{23} & 1 \end{vmatrix}}, \quad C = \frac{\sigma_1}{\sigma_3}\,\frac{\begin{vmatrix} 1 & r_{12} \\ r_{23} & r_{13} \end{vmatrix}}{\begin{vmatrix} 1 & r_{23} \\ r_{23} & 1 \end{vmatrix}}.$$

爲簡便並爲推廣至 k 個變數起見,我們設

$$R = \begin{vmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{vmatrix},$$

並且使 R_{ij} 代表 r_{ij} 之伴因子 (cofactor), 卽附以符號 $(-1)^{i+j}$ 之伴列式 (minor). 因此

$$R_{11} = \begin{vmatrix} r_{22} & r_{23} \\ r_{32} & r_{33} \end{vmatrix}, \quad R_{12} = -\begin{vmatrix} r_{21} & r_{23} \\ r_{31} & r_{33} \end{vmatrix}, \quad R_{13} = \begin{vmatrix} r_{21} & r_{22} \\ r_{31} & r_{32} \end{vmatrix}.$$

很顯明的,$r_{11} = r_{22} = r_{33} = 1$, 並且 $r_{12} = r_{21}$ 等等.故

$$B = -\frac{\sigma_1 R_{12}}{\sigma_2 R_{11}}, \quad C = -\frac{\sigma_1 R_{13}}{\sigma_3 R_{11}}.$$

因之式(70)變做

$$(72) \qquad \frac{x_1}{\sigma_1}R_{11} + \frac{x_2}{\sigma_2}R_{12} + \frac{x_3}{\sigma_3}R_{13} = 0.$$

假設實在迴歸面近於平面並且每個 x_1 柱體近於對稱的話,應用此方程式可由指定之 x_2 及 x_3 值來得 x_1 之最可能值.此爲最重要之方程

式，用牠能顯示出來 x_1 受 x_2 及 x_3 變化之影響.

其他兩個類似之方程式：x_2 倚 x_1 與 x_3 迴歸平面及 x_3 倚 x_1 與 x_2 迴歸平面，可參照式 (72) 交換 x 及 R 之下標 (subscript) 得之．其為

(73) $$\frac{x_2}{\sigma_2}R_{22}+\frac{x_3}{\sigma_3}R_{23}+\frac{x_1}{\sigma_1}R_{21}=0,$$

彼處 x_2 為倚變數；並且

(74) $$\frac{x_3}{\sigma_3}R_{33}+\frac{x_1}{\sigma_1}R_{31}+\frac{x_2}{\sigma_2}R_{32}=0,$$

彼處 x_3 為倚變數．若以 0 為原點式 (72) 變做

$$\frac{X_1-\bar{X}_1}{\sigma_1}R_{11}+\frac{X_2-\bar{X}_2}{\sigma_2}R_{12}+\frac{X_3-\bar{X}_3}{\sigma_3}R_{13}=0,$$

彼處 $X_i-\bar{X}_i=x_i$, $i=1$、2 及 3. 同樣可將式 (73) 及 (74) 以 0 為原點類似上面之公式寫出．

式 (72) 可以推廣至 k 個變數．於是我們有 x_1 倚 x_2、x_3、…及 x_k 之超級平面 (hyperplane)

(75) $$\frac{x_1}{\sigma_1}R_{11}+\frac{x_2}{\sigma_2}R_{12}+\cdots+\frac{x_k}{\sigma_k}R_{1k}=0,$$

彼處 R_{ij} 為下面行列式中 r_{ij} 之伴因子：

(76) $$R=\begin{vmatrix} r_{11} & r_{12}\cdots r_{1k} \\ r_{21} & r_{22}\cdots r_{2k} \\ \cdots & \cdots\cdots \\ r_{k1} & r_{k2}\cdots r_{kk}\end{vmatrix}.$$

當迴歸面方程式用標準單位表示時，式 (75) 變做

$$t_1=-\frac{1}{R_{11}}\sum_{i=2}^{k}R_{1i}t_i,\ \text{彼處}\ t_i=\frac{x_i}{\sigma_i}$$

3. 估計標準誤　由前面知道，估計標準誤 $S_Y{}^2=\sigma_Y(1-r^2)$ 為度量相關表各 X 行 Y 之平均數對於 Y 倚 X 迴歸線之離勢。現在我們尋求在三個變數情形中之類似算式。為達此目的，讓

$$S_{1\cdot 23}{}^2=\frac{1}{N}\sum_{1,2,3}\delta^2 f,$$

彼處 δ 為迴歸平面與點 (x_1, x_2, x_3) 間平行於 x_1 軸之距離。即是 δ 等於 x_1 之估計值與其觀察值之差。如果 x_1 之估計值係用式(72)算得者，於是我們可寫

$$NS_{1\cdot 23}{}^2=\frac{\sigma_1{}^2}{R_{11}{}^2}\sum_{1,2,3}\left(R_{11}\frac{x_1}{\sigma_1}+R_{12}\frac{x_2}{\sigma_2}+R_{13}\frac{x_3}{\sigma_3}\right)^2 f$$

$$=\frac{\sigma_1{}^2N}{R_{11}{}^2}(R_{11}{}^2+R_{12}{}^2+R_{13}{}^2+2R_{11}R_{13}r_{13}+2R_{12}R_{13}r_{23}+2R_{11}R_{12}r_{12})$$

$$=\frac{\sigma_1{}^2N}{R_{11}{}^2}\{R_{11}(R_{11}+r_{12}R_{12}+r_{13}R_{13})+R_{12}(R_{12}+Rr_{12}R_{11}+r_{23}R_{13})$$

$$+R_{13}(R_{13}+r_{13}R_{11}+r_{23}R_{12})\}$$

因為

$$(77)\qquad\begin{cases}R_{11}+r_{12}R_{12}+r_{13}R_{13}=R,\\ R_{12}+r_{12}R_{11}+r_{23}R_{13}=0,\\ R_{13}+r_{13}R_{11}+r_{23}R_{12}=0;\end{cases}$$

所以

$$S_{1\cdot 23}{}^2=\frac{\sigma_1{}^2R}{R_{11}}.$$

其平方根為

$$(78)\qquad S_{1\cdot 23}=\sigma_1\left(\frac{R}{R_{11}}\right)^{\frac{1}{2}}.$$

此為由指定之 x_2 及 x_3 估計 x_1 之標準誤。

4. 估計值之標準差 我們欲求與 σ_{EY} 相類似之公式(參閱第二節 8). 因爲由式(72)得知估計值之平均數爲零.故 x_1 估計值之變差 σ_{E1}^2

爲 $$\sigma_{E1}^2=\frac{1}{N}\sum\left(\frac{\sigma_1 R_{12}}{\sigma_2 R_{11}}x_2+\frac{\sigma_1 R_{13}}{\sigma_3 R_{11}}x_3\right)^2 f$$

$$=\frac{\sigma_1^2}{R_{11}^2}(R_{12}^2+R_{13}^2+2R_{12}R_{13}r_{23})$$

$$=\frac{\sigma_1^2}{R_{11}^2}\{R_{12}(R_{12}+R_{13}r_{23})+R_{13}(R_{13}+R_{12}r_{23})\}$$

$$=\frac{\sigma_1^2}{R_{11}^2}\{-R_{12}R_{11}r_{12}-R_{13}R_{11}r_{13}\}.\ \text{應用式(77)末二式}$$

$$=\frac{\sigma_1^2}{R_{11}}(R_{11}-R),\ \text{因}\ r_{11}R_{11}+r_{12}R_{12}+r_{13}R_{13}=R$$

$$=\sigma_1^2\left(1-\frac{R}{R_{11}}\right).$$

因之我們有

$$\sigma_{E1}=\sigma_1\left(1-\frac{R}{R_{11}}\right)^{\frac{1}{2}}. \tag{79}$$

此結果與 $\sigma_{EY}=r\sigma_Y$ 成對應(見式(17)). 我們一定可以想到因子 $\left(1-\frac{R}{R_{11}}\right)^{\frac{1}{2}}$ 可能與 r 成一種對應. 實際說 $\sqrt{1-\frac{R}{R_{11}}}$ 爲 x_1 倚 x_2 與 x_3 複相關係數之公式(見下面式(80)).

5. 複相關係數 x_1 之觀察值與由式(72)算得與 x_1 成對應估計值之相關係數叫做 x_1 倚 x_2 與 x_3 之複關係數 (multiple correlation coefficient), 通常以 $r_{1\cdot 23}$ 記之:如此我們有

$$r_{1\cdot 23}=\frac{\sum\limits_{1,2,3} x_{1E}x_1 f}{N\sigma_1\sigma_{E1}}.$$

彼處 ${}_0x_1$ 及 ${}_Ex_1$ 各表示 x_1 的觀察與估計值

應用式 (72), 可將上式寫之爲

$$N\sigma_1\sigma_{E1}r_{1\cdot23}=\frac{\sigma_1^2}{R_{11}}\sum\frac{x_1}{\sigma_1}\left(-\frac{R_{12}x_2}{\sigma_2}-\frac{R_{13}x_3}{\sigma_3}\right)f$$

$$=\frac{N\sigma_1^2}{R_{11}}(-R_{12}r_{12}-R_{13}r_{13})$$

$$=\frac{N\sigma_1^2}{R_{11}}(R_{11}-R)$$

$$=N\sigma_1^2\left(1-\frac{R}{R_{11}}\right)$$

再借助式 (79), 則得欲求之公式

(80) $$r_{1\cdot23}=\left(1-\frac{R}{R_{11}}\right)^{\frac{1}{2}}.$$

同樣可得 x_2 倚 x_1 與 x_3 及 x_3 倚 x_1 與 x_2 複相關係數之公式各爲

$$r_{2\cdot31}=\left(1-\frac{R}{R_{22}}\right)^{\frac{1}{2}}$$

及 $$r_{3\cdot12}=\left(1-\frac{R}{R_{33}}\right)^{\frac{1}{2}}.$$

將式 (78) 寫做下列之形式:

$$S_{1\cdot23}^2=\sigma_1^2\left\{1-\left(1-\frac{R}{R_{11}}\right)\right\}$$

我們有

(81) $$S_{1\cdot23}^2=\sigma_1^2(1-r_{1\cdot23}^2).$$

此與在簡單相關中 S_Y^2 的算式相類似. 由式 (81) 可見

$$-1\leq r_{1\cdot23}^2\leq 1$$

式 (80) 以及其類似之公式均可推廣至 k 個變數. 因此 x_1 倚 $(k-1)$ 個變數複相關係數之公式為

$$r_{1\cdot 23\cdots k}=\left(1-\frac{R}{R_{11}}\right)^{\frac{1}{2}}, \tag{82}$$

彼處 R_{ij} 為行列式 (76) 中 r_{ij} 之伴因子.

為適應計算目的我們須將複相關係數以簡單相關係數表之.

例 1. 設已給三個變數之兩兩簡單相關係數為

$$r_{12}=.8,\qquad r_{13}=-.7,\qquad r_{23}=-.9.$$

求 x_1 倚 x_2 與 x_3 複相關係數 $r_{1\cdot 23}$.

因為 $$R=\begin{vmatrix}1 & .8 & -.7\\ .8 & 1 & -.9\\ -.7 & -.9 & 1\end{vmatrix}=.068,\quad R_{11}=.19,$$

故 $$r_{1\cdot 23}=.8013.$$

例 2. 假設曾求得 $r_{12}=.6$, $r_{13}=-.4$, $r_{23}=.7$. 評論此結果.

因 $R=-.346$ 並且 $R_{11}=.51$, 故 $r_{1\cdot 23}=1.29$. 此結果不合理. 細察已給 r_{12} 及 r_{23} 之值均為正, 即 x_1 之值隨 x_2 值與 x_3 值之增加而增加; 但因 r_{13} 為負, 意即 x_1 之值不隨 x_3 值之增加而增加, 此為不可能之事, 故已給之結果一定有矛盾之處.

6. 定理 下列定理本身是很有興趣的, 並且對於理論之解釋及應用也有莫大神益, 因敘述之並證明之如下:

定理 10. 試證三平面 (72), (73) 及 (74) 重合之必要及充分條件為

$$r_{12}^2+r_{13}^2+r_{23}^2-2r_{12}r_{13}r_{23}=1.$$

證: 由解析幾何學上得知,兩個一次方程式代表同一平面之必要及充分條件爲其對應係數成比例,即

$$\frac{R_{11}}{R_{12}}=\frac{R_{12}}{R_{22}}=\frac{R_{13}}{R_{23}},$$

及
$$\frac{R_{12}}{R_{13}}=\frac{R_{22}}{R_{23}}=\frac{R_{23}}{R_{33}}.$$

當將上面關係用 r_{ij} 表示時,可見式(83)之關係是成立的. 證訖.

系 式(83)所表示之關係爲三個變數完全相關之必要及充分條件.

例 3. 已給 $r_{12}=.6$, $r_{13}=.4$; 並且假設 $r_{1\cdot 23}=1$. 試求 r_{23} 之值. 將已給之數值代入式(83)內,我們有

$$.36+.16-r_{23}^2-48r_{23}=1,$$

即是
$$r_{23}^2+48r_{23}+.48=0.$$

解之並棄掉 r_{23} 的負值,得 $r_{23}=.97$.

此例顯示之事實說,雖然 r_{12} 及 r_{13} 之各別值很小,三個變數 x_1、x_2 及 x_3 之相關未嘗不可以很高. 實在說,兩個變數各自與第三個變數有聯繫,很顯然的其不足以適應預測之目的,當將三個變數合攏在一起用複相關預測時,或者是很有價值的. 從他方面言之,用複相關迴歸方程式,由 $r_{1\cdot 23}$, 所能得到美滿所得之預測值用單相關迴歸方程式,由 r_{12} 或 r_{13} 預測,亦可能得到. 此種情形可由下列之二定理中瞭解之[9].

定理 11. 如果 $r_{23}=1$, 則 (i) $r_{1\cdot 23}^2=r_{12}^2=r_{13}^2$,

(ii) $S_{1\cdot 23}^2=\sigma_1^2(1-r_{12}^2)$.

證: 當 $r_{23}=1$ 時,於是 $R_{11}=0$, 及由式(80)得知 $r_{1\cdot 23}$ 變做無限

大.此爲不可能之事.當 $r_{23}=1$ 時,仍然可以證明 $r_{12}=r_{13}$ 學者試由式(80)驗證之.如此我們養看,當 $r_{13}=r_{12}$ 時 式(80)之變化如何.如使式(77)第一式中之 $r_{13}=r_{12}$, 則 $R_{13}=R_{12}$, 因之 $R-R_{11}=2r_{12}R_{12}$. 將此結果代入式(80)中,我們立刻可得

$$r_{1\cdot23}^2=\frac{-2r_{12}R_{12}}{R_{11}}=\frac{2r_{12}^2}{1+r_{23}}$$

牢記 $r_{13}=r_{12}$. 如果讓末式內之 $r_{23}=1$, 我們便得定理之第一個結論.第二個結論可由第一個及式(81)證得之　　證訖.

上面定理告訴我們說,複相關不比簡單迴歸線——x_1 倚 x_2 及 x_1 倚 x_3——有什麼好處,因爲 $S_{1\cdot23}^2=S_1^2$.

定理 12. 當 $r_{23}=0$ 時,則 $r_{1\cdot23}^2=r_{12}^2+r_{13}^2$.

證: 當 $r_{23}=0$ 時,頗易證明 $R_{11}=1$ 並且 $R=1-r_{12}^2-r_{13}^2$. 如此由式(80)我們有

$$r_{1\cdot23}^2=r_{12}^2+r_{13}^2.$$

證訖.

於是由上面定理得知,當 $r_{23}=0$ 時估計標準誤之公式(81)變做

$$S_{1\cdot23}^2=\sigma_1^2(1-r_{12}^2-r_{13}^2).$$

因之,當 x_2 及 x_3 完全不相干時,複迴歸線供給之預測值,比 x_1 倚 x_2 或 x_1 倚 x_3 簡單迴歸線供給者爲優;如果 $r_{12}+r_{13}$ 等於常數並且 r_{12} 幾乎與 r_{13} 相等更優.當 r_{12} 眞正與 r_{13} 相等時,則其最大值爲 $\left(\frac{1}{2}\right)^{\frac{1}{2}}=.707$. 此定理顯示之事實爲,當二變數各與第三變數成高度相關但其彼此各不相干時,有一良好之迴歸方程式,由已給之二變數以估計第三變數之

第七節 淨相關

1. 意義 保持其他變數不變而度量兩個變數之相關，在統計分析上是很重要的。在三個變數情形中，假想有一與 x_1x_2 平面平行之壁體（圖 62）。此壁體上面之二量相關表表示 x_3 值固定後 x_1 與 x_2 之相關。在此分配中 x_1 與 x_2 相關係數叫做偏相關係數 (partial correlation coefficient)，或曰淨相關係數 (net correlation coefficient)，通常以 $r_{12\cdot3}$ 記之；在壁體上面之迴歸曲線叫做偏相關迴歸曲線 (partial regression curve)，或叫淨相關曲線 (net regression curve)。例如研究父、母及子之身長的關係，當指定母之身長為一特別值——譬如為 62 英寸——時，問父子身長之相關情形若何？此為求淨相關係數之問題。

為用總相關係數 r_{ij} 表示淨相關係數 $r_{12\cdot3}$ 起見，像先前表示複相關係數 $r_{1\cdot23}$ 一樣，我們假定一理想情況。假設所研究之分配是這樣：各總迴歸曲線為直線，並且迴歸面為平面。於是在 x_3 處 x_1 倚 x_2 之淨相關迴歸線為 x_1 倚 x_2 及 x_3 迴歸面之截線，因為此線經過立於 (x_2, x_3) 點上各 x_1 柱體之平均數。

2. 公式 在兩個變數情形中，我們知道 S_Y^2 為 X 行各 Y 變差平均數。進一步說，當分配為常態時，得證明各行變差為常數，該常數即為 S_Y^2。在三個變數情形中，若其分配合乎上述之理想，和兩個變數頗有類似之處。追想在普通直線迴歸線情形中，$S_{1\cdot23}^2$ 可視做在 (x_2, x_3) 點處各 x_1 柱體變差之平均數，因為，當迴歸線為直線時，各柱體之平均數在迴歸平面上。現在假設分配在 x_1 方向之離勢相同，如此各柱體之變差相同。

在此假設之下，$S_{1\cdot 23}^2$ 爲每個 x_1 柱體之離勢．讓 $\sigma_{1\cdot 3}^2$ 代表在 x_3 處各 x_1 之離勢．記得 $r_{12\cdot 3}$ 爲 x_1 與 x_2 之相關表中之相關係數，並且此表中之迴歸線爲直線及各行之離勢相同：所以此表各柱體之離勢 $S_{1\cdot 23}^2$ 可寫做

(84) $$S_{1\cdot 23}^2=\sigma_{1\cdot 3}^2(1-r_{12\cdot 3}^2).$$

現在我們研究二量 x_1 與 x_3 之相關表．在此表中，r_{13} 爲 x_1 與 x_3 之總相關係數，並且 $\sigma_{1\cdot 3}^2$ 爲在 x_3 行各 x_1 之變差．因爲已假設在此各行之 $\sigma_{1\cdot 3}^2$ 爲常數，我們可寫

(85) $$\sigma_{1\cdot 3}^2=\sigma_1^2(1-r_{13}^2)$$

由式 (81)、(84) 及 (85)，我們得

$$(1-r_{1\cdot 23}^2)=(1-r_{13}^2)(1-r_{12\cdot 3}^2),$$

即是
$$\frac{R}{R_{11}}=R_{22}(1-r_{12\cdot 3}^2).$$

解之，有
$$r_{12\cdot 3}^2=\frac{R_{11}R_{22}-R}{R_{11}R_{22}}.$$

將各 R 展開，可以驗證

$$R_{11}R_{22}-R=(-R_{12})^2.$$

爲一恆等式．所以我們得最後結果

$$r_{12\cdot 3}=\frac{-R_{12}}{(R_{11}R_{12})^{\frac{1}{2}}}.$$

如欲以各 r 表示上式右方，則有

(86) $$r_{12\cdot 3}=\frac{r_{12}-r_{13}r_{23}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{23}^2}}.$$

因式 (86) 不包含 x_3，指定 x_3 以一值與指定 x_3 以另一值之 $r_{12\cdot 3}$ 值

是相同的．所以，不但各變差在 x_1 方向分配須相等，並且垂直於 x_3 軸各壁體之 $r_{12\cdot 3}$ 亦均須相同．很顯然的，此等條件在實際問題中未必具備．如此，在應用時，$r_{12\cdot 3}$ 當做各淨相關係數之一種平均值，該值由指定 x_3 以所有各值即能得之．淨相關主要用途，在除掉第三個變數之影響後使吾人得以測驗其餘二變數相關之程度．

3. 推求淨相關係數公式之另一法　淨相關係數公式可用另一法推求之．和先前一樣，假設 x_1、x_2 及 x_3 爲以算術平均數做原點之變數，並假設我們所欲求者爲除掉第三個變數 x_3 之影響後 x_1 與 x_2 之相關係數．讓我們由 x_1 減去 x_1 受 x_3 影響之部分，即減去 x_1 倚 x_3 迴歸線所能限定之部分，並讓 $x_{1\cdot 3}$ 表示如此所得之剩餘數；同樣，讓由 x_2 減去 x_2 受 x_3 影響之部分，即減去 x_2 倚 x_3 迴歸線所能限定之部分，並讓 $x_{2\cdot 3}$ 表示如此所得之剩餘數．於是，我們有

$$(87)\qquad \begin{cases} x_{1\cdot 3}=x_1-r_{13}\dfrac{\sigma_1}{\sigma_3}x_3, \\ x_{2\cdot 3}=x_2-r_{23}\dfrac{\sigma_2}{\sigma_3}x_3. \end{cases}$$

我們將在下面證明 $x_{1\cdot 3}$ 與 $x_{2\cdot 3}$ 之簡單相關係數爲 $r_{12\cdot 3}$．由定義，$x_{1\cdot 3}$ 與 $x_{2\cdot 3}$ 之相關係數爲

$$(88)\qquad \frac{\sum\limits_{1,2,3} x_{1\cdot 3}x_{2\cdot 3}f}{N\sigma_{1\cdot 3}\sigma_{2\cdot 3}}.$$

借助式 (87)，則分數式 (88) 之分子爲

$$\sum_{1,2,3} x_1x_2f-r_{13}\frac{\sigma_1}{\sigma_3}\sum_{1,2,3} x_2x_3f-r_{23}\frac{\sigma_2}{\sigma_3}\sum_{1,2,3} x_1x_3f+r_{13}r_{23}\frac{\sigma_1\sigma_2}{\sigma_3^2}\Sigma x_3^2f$$

$$= N(\sigma_1\sigma_2 r_{12} - \sigma_1\sigma_2 r_{13} r_{23} - \sigma_1\sigma_2 r_{13} r_{23} + \sigma_1\sigma_2 r_{13} r_{23})$$

$$= N\sigma_1\sigma_2(r_{12} - r_{13} r_{23})$$

現在借助式(85),我們有

$$\sigma_{1\cdot 3} = \sigma_1(1 - r_{13}^2)^{\frac{1}{2}}$$

並依同理 $$\sigma_{2\cdot 3} = \sigma_2(1 - r_{23}^2)^{\frac{1}{2}}.$$

將此等結果代入式(88)內我們得到欲求之式(86).

因爲推求 $r_{12\cdot 3}$ 所用之方法,係將居首二變數受第三個變數之影響除掉後求居首二變數之相關程度,故偏相關又有時被稱做淨相關,偏相關係數有時被稱做淨相關係數.

4. 一般情形 在 N 個變數情形中,如果使 $(N-2)$ 個變數 X_3、X_4、…、X_N 固定,則二變數 X_1 及 X_2 之相關係數 $r_{12\cdot 34\cdots n}$ 可由下式求之:

$$r_{12\cdot 34\cdots n} = \frac{r_{12\cdot 34\cdots(n-1)} - r_{1n\cdot 34\cdots(n-1)}\, r_{2n\cdot 34\cdots(n-1)}}{\sqrt{1 - r_{1n\cdot 34\cdots(n-1)}^2}\,\sqrt{1 - r_{2n\cdot 34\cdots(n-1)}^2}}.$$

式中 r 之下標係被句點(·)分爲兩部,前部如 12 可稱做主下標(primary subscript),後部如 $34\cdots n$ 可稱做次下標(secondary subscript).係數附以 S 個次下標者可以名之爲 S 次係數,如 $r_{12\cdot 3}$ 可以名之爲一次相關係數,$r_{12\cdot 34}$ 可以名之爲二次相關係數;因此總相關係數 r_{12} 可以名之爲零次相關係數.

問題 VII

1. 試用式(3)計算表 58 裏體重與身長之相關係數.
2. 試用式(5)計算表 58 裏體重與身長之相關係數.

3. 下表爲檢查20個學生體格時所得的論據. X 代表身長(英寸), Y 代表胸圍(英寸),並且 Z 代表體重(磅). 求相關係數 (a) X 與 Y 的, (b) X 與 Z 的及 (c) Y 與 Z 的.

X	Y	Z	X	Y	Z
68.5	33.6	148	65.3	33.0	136
67.2	35.0	144	65.1	34.0	144
67.7	30.2	145	64.8	37.3	170
63.8	30.0	108	69.6	33.4	154
69.9	33.0	130	68.2	31.5	122
64.7	31.0	112	68.2	32.0	141
68.4	33.0	134	72.3	35.0	159
66.4	30.2	112	67.8	33.7	134
69.1	33.3	143	71.3	31.5	136
71.0	32.3	136	63.5	33.6	126

4. 舉出幾個負相關之例.

5. 試用式(6)及(7)求表 58 裏體重倚身長及身長倚體重之迴歸線.

6. 試求表 59 裏出生率與死亡率之相關係數及由死亡率估計出生率或由出生率估計死亡率之迴歸線.

7. 以直線並依第五章第四節 1 及 2 所述之條件爲準則配合上題裏之論據.現在所得之二直線與上題所得之二迴歸線一一比較,你所得到的結論爲何?

8. 由於 1000 個學生身長與體重量度算得之結果如下:

$$\bar{X}=150.00\text{ 磅},\quad \bar{Y}=68.00\text{ 英寸},$$

$$\sigma_X=\ 20.00\text{ 磅},\quad \sigma_Y=\ 2.50\text{ 英寸}.$$

$$r=.60.$$

如果知道甲生之身長爲5英尺，乙生之體重爲200磅，試由甲生之身長估計其體重，並由乙生之體重估計其身長。

9. 設 $N=1,000$; $\Sigma X=150,000$; $\Sigma X^2=22,725,000$; $\Sigma Y=70,000$; $\Sigma Y^2=4,936,000$; $\Sigma XY=10,522,500$. 求 $\bar{X}$、$\bar{Y}$、σ_X、σ_Y、r 及迴歸線。

10. 證明各估計 Y 值之變差(對於 Y 值之平均數者)與各 Y 值之變差 σ_Y^2 之比等於 r^2.

11. 已給兩個相關變數分配之標準差 σ_X 及 σ_Y: 問

(a) 如果 $r=0$ 由 X 估計 Y 之標準誤爲何?

(b) 如果 r 增至 .25, (a) 裏之 S_Y 減少若干?

(c) 若使 S_Y 爲 (a) 裏者之半，r 之值應增至何數?

(d) 若使 S_Y 爲 (a) 裏者三分之一，r 之值應增至何數?

(e) S_Y 減至零，r 之值爲何?

(f) 不論 r 之值爲何，由 X 估計 Y 之標準誤與 Y 分配之標準差之比爲何?

12. 如果所有各點恰在 Y 倚 X 之迴歸線上，證明 $S_Y^2=0$ 並且因此 $r=\pm1$.

13. 證明 S_Y^2 可用下面關係計算之:

$$NS_Y^2=\Sigma y'^2-\frac{(\Sigma x'y')^2}{\Sigma x'^2}.$$

彼處 x' 及 y' 各表示 $X-\bar{X}$ 及 $Y-\bar{Y}$.

14. 證明二迴歸線夾角之正切由式 (6) 及 (7) 算得者爲

$$\tan\theta=\frac{\sigma_X\sigma_Y}{\sigma_X^2+\sigma_Y^2}\left\{\frac{1-r^2}{r}\right\}$$

並且由式 (13) 及 (14) 算得者爲

$$\tan\theta=\frac{1-r^2}{2r}.$$

問當 $r=1$ 及 0 時其值爲何?

15. 證明式 (15).

16. 完成定理 4 之證明.

17. 任證式 (21) 末五式中之一.

18. 假設 $U=\sum\limits_X XF_{XY}$, $V=\sum\limits_Y YF_{XY}$.

求證 (a) $\sum\limits_X V=\sum\limits_Y X\Phi_X$, (b) $\sum\limits_Y V=\sum\limits_Y Y\Psi_Y$, (c) $\sum\limits_Y YV=\sum\limits_X XV$.

19. 分別應用式 (33) 及 (35) 計算表 64 裏所列論據之相關比 η_{YX}.

20. 應用式 (38) 計算表 65 裏論據相關比 η_{yx} 並與第三節例 1 裏所得之結果比較之.

21. 討論 S_Y 對於估計 Y 值上之用途.

22. 證明 ρ 之定義可寫做下列之形式:

$$\rho=\frac{1}{\sigma_x\sigma_y}\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\,dxdy-\bar{x}\bar{y}\right).$$

23. 已給 $f(x,y)=2/a^2$, $0\leq x\leq y$, $0\leq y\leq a$. 證明二迴歸線均爲直線.求 ρ 之值.

24. (a) 如果 $\rho=.6$ 證明橢圓長軸與短軸之比爲 2. (b) 常態相關曲面 y 倚 x 迴歸線之斜度以 S_y 及 σ_x 單位表之爲 $\rho/(1-\rho^2)^{\frac{1}{2}}$.

25. 推求式 (55).

26. 假設有售貨員十二人，商店經理按他們的服務年數及工作效率分成等第，如下表所示。問售貨員服務年數與工作效率之關係如何?

售貨員	服務年數	服務年數等第	工作效率等第
1	5	7.5	6
2	2	11.5	12
3	10	2	1
4	8	4	9
5	6	6	8
6	4	9	5
7	12	1	2
8	2	11.5	10
9	7	5	3
10	5	7.5	7
11	9	3	4
12	3	10	11

27. 證明下列橢圓之面積為 $\pi\lambda^2\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}$:

$$\frac{x^2}{\sigma_x^2}-\frac{2\rho xy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2}=\lambda^2,$$

彼處 $\lambda^2=2(1-\rho^2)\log_e\frac{\kappa}{c}$, $\frac{1}{\kappa}=2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}$, $0<c<\kappa$, $-\infty\leq x\leq\infty$, 並且 $-\infty\leq y\leq\infty$.

28. 如果

$$f(x,y)=\kappa e^{-P},$$

彼處 $\frac{1}{\kappa}=2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}$, $p=\frac{1}{2(1-\rho^2)}\left\{\frac{x^2}{\sigma_x^2}-\frac{2\rho xy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma_y^2}\right\}$, $-\infty\leq x\leq\infty$, $-\infty\leq y\leq\infty$, 證明

(a) 經過所有 x 值之平均數 $\bar{y}_x$ 爲零.

(b) $\bar{y}_x$ 之變差等於 $\rho^2\sigma_y^2$.

(c) $\bar{y}_x$ 與 y 間之相關係數等於 ρ.

29. 假設 X_1 代表榮譽分數, X_2 代表普通智力, X_3 代表學習鐘點; $\bar{X}_1=18.5$, $\bar{X}_2=100.6$, $\bar{X}_3=24$; $\sigma_1=11.2$, $\sigma_2=15.8$, $\sigma_3=6$; $r_{12}=.60$, $r_{13}=.32$, $r_{23}=-.35$ (論據見參考書 25). 求複相關係數及迴歸方程式.

30. 在上題論據裏如果學習鐘點相同,榮譽分數與普通智力之相關達到某種程度?

31. 假設 $X_1=$ 草子收穫量 (cwt), $X_2=$ 雨量(吋); $X_3=$ 春季在 $42°F$ 以上之溫度; $\bar{X}_1=28.02$, $\bar{X}_2=4.91$, $\bar{X}_3=594$; $\sigma_1=4.42$, $\sigma_2=1.10$, $\sigma_3=85$; $r_{12}=.80$, $r_{13}=-.40$, $r_{23}=-.56$ (論據見參考書 44). 求淨相關係數及 X_1 倚 X_2 及 X_3 迴歸方程式.

32. 推求並解釋 $\sigma_1^2=\sigma_{r1}^2+S_{1\cdot23}^2$. 在單相關情形中其對應之關係爲何?

33. 下面所列者爲美國依阿華州二十五郡地價與收穫之論據(見參考書 25).

$X_1=1920$ 年每畝平均地價 (\$),

$X_2=1910-1919$ 年每畝平均產穀量(斛),

$X_3=$ 產雜糧土地之百分數,

$X_4=$ 產玉蜀黍土地之百分數.

郡	X_1	X_2	X_3	X_4
1	87	40	11	14
2	133	36	13	30
3	174	34	19	30
4	385	41	33	39
5	363	39	25	33
6	274	42	23	34
7	235	40	22	37
8	104	31	9	20
9	141	36	13	27
10	208	34	17	40
11	115	30	18	19
12	271	40	23	31
13	163	37	14	25
14	193	41	13	28
15	203	38	24	31
16	279	38	31	35
17	179	24	16	26
18	244	45	19	34
19	165	34	20	30
20	257	40	20	38
21	252	41	22	35
22	280	42	21	41
23	167	35	16	23
24	168	33	18	24
25	135	36	18	21

(a) 求 X_1 倚 X_2, X_3 及 X_4 之直線迴歸方程式.

(b) 用 (a) 裏得到之方程式估計 X_1 之首五值.

(c) 計算 $S_{1.234}$ 及 $r_{1.234}$ 之值.

第八章　抽樣問題

第一節　抽樣問題之要點及解決方法

1. 要點　本章用淺鮮實例闡述有關抽樣問題之重要理論. 我們可將研究之問題分爲第一及第二兩類:第一類係用量度——如集中常數、離勢常數、相關係數等——以總括一羣度量之切要事實;第二類係用由宇宙內隨機抽取之樣本以探求或推測全宇宙*之事實,以及各有關問題. 然而這兩類問題之處理方法有時是不可分的. 所謂樣本(或者精確些稱之曰隨機樣本)者意思爲一小組變量,其中任一變量由宇宙內抽取之機會是相同的且是自由的. 由於這樣抽取之樣本,試行推測關於全宇宙之事實. 我們先思考無限(或有限)宇宙,並且依照機率定理研究樣本之姿態. 總之,抽樣問題之中心興趣在推得準則,以輔助用常識判斷平均數、比率數、係數等波動情形之不逮.

爲避淆混起見,可將量度分爲兩種:由於樣本頻數分配算得者,稱之爲統計常數;由於宇宙頻數分配算得者,稱之爲真正常數. 後者爲吾人所欲探求者,但受時間及空間之限制無從知道,祇能依機率理論做有根據之估計耳. 統計常數之變化隨樣本而異,此爲習知之事實. 追究推論所能容受之離差,爲在本章研究抽樣問題之主眼.

---

* 宇宙本係論理學上習用之名詞,表示事物之全體,今借用之表示同樣意義.

2. 解決方法 抽樣全部問題係以頻數分配及機率做基礎，解決抽樣問題自當從這方面着想.茲將解決抽樣問題所用之方法，提要敍述如下：

在抽樣問題中所欲研究之特定事項爲：如已給 N 個變量樣本之平均數及標準差，問由此以估計宇宙平均數及標準差之準確程度如何？又如已給兩個樣本之平均數或其他統計常數，問彼此有無顯著之差別？此差由於抽樣機會而起，抑或兩個統計常數來自不同宇宙？爲回答此種問題我們須研究宇宙內各變量之分配，須研究由包含 N 個變量樣本算得各統計常數之分配，並須研究表示此二分配所用之算式.頻數及機率分配曲線之研究在統計學上所以佔有重要地位者，卽在於此.

算得之統計常數因樣本而異.例如，由 100 個觀察值算得之平均數與由另外 100 個同樣性質觀察值求得者，未必完全相等. 實在說，如果取許多個樣本，每樣本包含 100 個觀察值，我們於是得到許多平均數而成一頻數分配. 此分配本身有其平均數(平均數之平均數)、標準差及高級動差，以描寫其種種特性. 用以描寫由 N 個事物組成各樣本平均數分配定律之函數叫做分配函數 (distribution function)，其圖叫做平均數曲線 (curve of means). 對於平均數之討論旣如上述，其他各統計常數可同樣討論之.

總括言之，用統計方法研究樣本之終極目標，不外關於宇宙之說明及各樣本統計常數分配函數之測定，此樣本係宇宙內抽取者並且其內包含事物之個數不變. 測定各種統計常數分配函數爲近代統計家備極關心問題之一. 在大多數情形中，必須假定宇宙爲常態形式，如此可以

用解析法得到統計常數之抽樣分配．許多顯著性測驗均以此假定爲其基礎．關於自由抽樣分配之事實，可由抽樣分配各級動差或其期望值知之，因此於下節論述之．

第二節 期望值

1. 意義 設 X 爲受分配機率函數 $\Phi(X)$ 限制之連續變數，$G(X)$ 爲 X 之任意函數．於是 $G(X)$ 之期望值 (expected value)，如果下式右邊積分存在的話，得用運算符號 E 表示之，卽

(1) $$E\{G(X)\}=\int_{-\infty}^{\infty}G(X)\Phi(X)dX.$$

在特殊情形中，如果

$$G(X)=X^k,\ k=1,\ 2,\cdots\cdots,$$

我們有 $$E(X^k)=\int_{-\infty}^{\infty}X^k\Phi(X)dX.$$

若 $k=1$，$E(X^k)$ 卽爲 $\Phi(X)$ 所代表之宇宙內各 X 的平均數．此後我們用 $\tilde{X}$ 代表各 X 之宇宙平均數(以前曾用 $\tilde{X}$ 代表過中位數，切勿與之相混)．因此我們寫

(2) $$E(X)=\tilde{X},$$

如果 $G(X)=(X-\tilde{X})^2$，並且以 $\sigma_X{}^2$ 代表 X 之變差；則

(3) $$\sigma_X{}^2=E(X-\tilde{X})^2,$$

$\sigma_X{}^2$ 之平方根(正值)叫做 X 分配之標準差或簡稱之爲標準差(參閱第一之章第三節)．自然可給 Y 以類似之定義．

設 $F(X, Y)$ 爲 X 與 Y 聯合分配之機率函數．於是

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}XYF(X, Y)dY\,dX.$$

如果 X 與 Y 爲有聯繫之變數，於是由前章第四節得知

$$F(X, Y)\doteq\Phi(X)\Psi(Y),$$

彼處 $\Phi(X)$ 及 $\Psi(Y)$ 各爲 X 及 Y 邊和分配之機率函數．如果 ρ 爲在 $F(X, Y)$ 所代表之二量宇宙內 X 與 Y 之相關係數，則

$$(4)\qquad \rho=\frac{E(XY)-\tilde{X}\tilde{Y}}{\sigma_X\sigma_Y},$$

彼處 $\tilde{X}$, $\tilde{Y}$, σ_X, σ_Y, ρ, 爲由宇宙算得之真正常數 (true constants).

2. 定理　下面所列之各定理不難由定義推得，因此僅敘述之，不加證明．

定理 1.　一常數與一變數相乘積之期望值等於常數與變數期望值之相乘積．即是

$$E(CX)=CE(X).$$

定理 2.　一變數與其期望值離差之期望值等於零．即是

$$E(X-\tilde{X})=0.$$

定理 3.　兩個或兩個以上變數和之期望值等於各變數期望值之和．用符號表之爲

$$E(X+Y+Z)=E(X)+E(Y)+E(Z).$$

定理 4.　如果 X 與 Y 爲各不相干之變數，於是牠們乘積之期望值等於各期望值之乘積．即是

$$E(XY)=E(X)E(Y).$$

定理 5. 兩個各不相干之變數各與其期望值離差相乘積之期望值等於零. 即是 $E\{(X-\bar{X})(Y-\bar{Y})\}=0.$

定理 6. 兩個彼此有關聯之變數各與其期望值離差相乘積之期望值爲 $E\{(X-\bar{X})(Y-\bar{Y})\}=\rho\sigma_X\sigma_Y.$

第三節 直線函數之標準差

假設某變數爲兩個或兩個以上不相干變數之直線函數,每個不相干變數係由宇宙內任意抽取者. 試用各不相干變數分配之某種動差表示此函數之標準差. 爲達到這個目的,讓

$$\omega=C_1X_1+C_2X_2+\cdots\cdots+C_NX_N, \tag{5}$$

彼處變數 X_k, $k=1, 2, \cdots\cdots, N$, 之分配是任意的,並且各常數 C 之值也是任意的. 讓 σ_k 代表 X_k 所屬宇宙之標準差,並且讓 ρ_{ij} 代表 X_i 與 X_j 之相關係數(如果相關存在的話). 於是我們試用 σ_k 及 $\rho_{ij}(i, j=1, 2\cdots, N)$ 表示 ω 之標準差 σ_ω.

第一種情形. 我們先假設在各宇宙內之變數彼此有關聯,即是不論 i 及 j 之值爲何, $\rho_{ij}\neq 0$. 由式 (5) 及定理 3 我們有

$$E(\omega)=C_1E(X_1)+C_2E(X_2)+\cdots\cdots+C_NE(X_N), \tag{6}$$

即是 $$\bar{\omega}=C_1\bar{X}_1+C_2\bar{X}_2+\cdots\cdots+C_N\bar{Y}_N,$$

因之,由式 (5) 及末式,我們有

$$\omega-\bar{\omega}=C_1(X_1-\bar{X}_1)+C_2(X_2-\bar{X}_2)+\cdots\cdots+C_N(X_N-\bar{X}_N),$$

$$E(\omega-\bar{\omega})^2=\Sigma C_i^2E(X_i-\bar{X}_i)^2+\Sigma C_iC_jE(X_i-\bar{X}_i)(Y_j-\bar{Y}_j),$$

並且,由式 (3) 及定理 6, 上式變爲

(7) $$\sigma_\omega^2=\Sigma C_i^2\sigma_i^2+\Sigma C_iC_j\rho_{ij}\sigma_i\sigma_j.$$

如果 $C_1=1$, $C_2=-1$ 並且 $N=2$; 我們有

(8) $$\sigma_\omega^2=\sigma_1^2-2\rho_{12}\sigma_1\sigma_2+\sigma_2^2,\text{ 彼處 } \omega=X_1-X_2,$$

此為式 (7) 之特殊情形.

第二種情形. 假設式(5)中之各 X 彼此不相干,如此 $\rho_{ij}=0$. 於是式(7)變做

(9) $$\sigma_\omega^2=C_1^2\sigma_1^2+C_2^2\sigma_2^2+\cdots\cdots C_N^2\sigma_N^2.$$

並且式 (8) 變做

(10) $$\sigma_\omega^2=\sigma_1^2+\sigma_2^2.$$

第四節 平均數

1. 符號 在本章裏所用之符號與前不盡相同，即如以前用 $\tilde{X}$ 表示中位數，今則用以表示宇宙平均數，因此將現在採取之符號列表區別如下:

表 71. 符號及其代表之意義

常數	屬於		
	宇宙分配	樣本分配	各平均數分配
平均數	$\tilde{X}$	$\bar{X}$	$E(\bar{X})=\tilde{X}$
標準差	σ_X	s_X	$\sigma_{\bar{X}}$
變差	σ_X^2	s_X^2	$\sigma_{\bar{X}}^2$
偏度	$\sqrt{\tilde{\beta}_{1X}}$	$\sqrt{\beta_{1X}}$	$\sqrt{\tilde{\beta}_{1\bar{X}}}$
峯度	$\tilde{\beta}_{2X}$	β_{2X}	$\tilde{\beta}_{2\bar{X}}$

2. 平均數之平均數 關於平均數之平均數的定理如下：

定理 7. 如果由任意宇宙內抽取包含 N 個事物之樣本，並且 $\bar{X}$ 代表由樣本算得之平均數；於是所有各可能平均數之平均數等於宇宙之平均數. 即是

$$(11) \qquad E(\bar{X})=(\tilde{X}).$$

證：使式 (5) 內之 $C_1=C_2=\cdots=C_N=\frac{1}{N}$，並且讓 $X_1, X_2\cdots\cdots, X_N$ 組成之樣本係由以 $\tilde{X}$ 爲平均數以 $\sigma_X{}^2$ 爲變差之宇宙內抽取者. 於是由式 (5) 得知 $\omega=\bar{X}$，由式 (2) 得知 $E(X_i)=\tilde{X}(i=1, 2,\cdots\cdots, N)$ 並且由式 (6) 得知 $E(\omega)=\tilde{X}$ 因之 $E(\bar{X})=\tilde{X}$. 證訖.

爲瞭解此定理之具體意義起見，特舉一實例說明如下：

假設我們的樣本是由很多人口中抽取者，或者說由宇宙內抽取者，並且我們知道此宇宙內每個事物之量度爲由 0 至 9，例如記載每個人在某固定時間內患普通感冒之最多次數爲 9，全人口平均患感冒次數 $\tilde{X}$ 及標準差 σ_X 係已知者；讓 $\tilde{X}=4.5$ 並且 $\sigma_X=2.87$. 由於這樣的宇宙內，任意抽取 5 人之樣本. 我們可以計算每個樣本之平均數. 我們要問由於這個小樣本求得之平均數與實在平均數之差究竟達到何種程度呢？

在表 72 裏列有 100 個 5 人樣本，其由宇宙內隨機抽取者. 例如在表 72 第一個樣本中患感冒 2 次者 3 人，幸免者 1 人，不幸患 4 次者 1 人. 由於每個樣本可計算其平均數，因是共有 100 個平均數；這些平均數之平均數 4.43，其與宇宙平均數 4.50 頗相近. 可見定理與事實符合 [參閱式(11)].

表72.　100個5人樣本,每人感冒之平均次數(取自平均數為4.5標準差為2.87之宇宙內)

樣本號數	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
	2	6	3	9	7	5	3	0	5	4	5	2	4	1	2	6	6	9	2	3	5	3	9	6	6
	4	9	1	1	7	5	1	0	2	3	9	2	3	5	8	6	5	8	2	3	1	6	1	6	2
	2	7	3	2	3	1	6	3	7	1	2	7	1	7	0	8	8	9	7	8	6	5	5	3	7
	0	5	1	6	6	8	4	3	8	6	9	2	9	3	2	8	3	5	6	5	2	2	4	7	2
	2	9	1	9	7	6	8	6	6	6	8	1	1	5	3	7	1	1	7	1	5	4	2	5	1
平均數	2.0	7.2	1.8	5.4	6.0	5.0	4.4	2.4	5.6	4.0	6.6	2.8	3.6	4.2	3.0	7.0	4.6	6.4	4.8	4.0	3.8	4.0	4.2	5.4	3.6

樣本號數	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
	7	7	8	7	2	4	1	3	3	9	4	2	5	9	5	9	7	3	0	2	1	4	9	4	4
	5	0	0	2	5	2	6	5	9	7	1	5	1	7	8	3	6	6	9	3	1	1	8	4	1
	3	8	6	6	4	6	4	7	6	1	7	1	9	5	5	9	1	5	0	6	0	4	1	4	4
	1	5	0	5	9	8	4	4	9	5	3	5	2	3	7	9	9	2	7	4	2	6	7	2	7
	6	5	4	9	2	7	3	5	9	2	4	4	5	8	7	7	3	2	7	0	7	1	8	1	6
平均數	4.4	5.0	3.6	5.8	5.2	5.4	3.6	4.8	7.2	4.8	3.8	2.4	4.4	6.4	6.4	7.4	5.2	3.6	4.6	3.0	2.2	3.6	6.6	3.0	4.4

樣本號數	1	2	3	4		6		8	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
	9	9	7	0	9	1	0	6	1	9	1	7	3	1	0	3	4	0	3	6	2	7	8	0	1
	0	9	8	9	4	0	6	7	5	2	7	8	7	5	7	6	1	6	3	4	0	2	9	6	9
	1	3	5	3	9	5	2	8	8	4	3	2	0	9	8	5	8	3	0	7	3	8	7	9	2
	0	4	4	1	1	1	5	3	4	7	7	4	5	6	6	8	3	3	9	4	7	0	7	2	2
	6	3	3	2	8	4	6	9	3	3	8	3	8	0	6	0	2	9	5	2	5	1	2	9	0
平均數	5.0	5.6	5.4	3.0	6.2	2.2	3.8	6.6	4.2	5.0	5.2	4.8	4.6	4.2	5.4	4.4	3.6	4.2	5.6	4.6	3.4	3.6	6.6	5.2	2.8

樣本號數	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
	2	3	2	5	9	1	2	2	9	0	3	9	6	5	6	4	4	2	8	4	1	7	2	8	0
	1	9	8	1	4	3	4	7	6	9	3	4	9	3	5	0	6	3	5	4	1	9	9	6	1
	0	1	2	3	7	1	2	1	2	3	8	1	0	1	4	4	3	1	2	1	0	7	2	3	5
	4	2	4	3	8	1	7	5	7	9	0	7	2	5	2	4	9	7	3	2	0	3	5	8	2
	1	1	0	2	1	6	9	9	2	7	2	5	7	0	1	2	1	7	0	4	4	3	0	9	6
平均數	1.6	3.2	3.2	2.8	5.8	2.4	4.8	4.8	5.2	5.6	3.2	5.2	4.8	2.8	3.6	2.8	4.6	4.0	3.6	3.0	1.2	5.8	3.6	6.8	3.4

3. 平均數之標準差　關於平均數之標準差之定理如下：

定理 8. 由任意宇宙內抽取許多樣本，其平均數分配之變差等於宇宙變差除以樣本包含事物之個數. 用符號表之為

$$\sigma_{\bar{X}}^2=\frac{\sigma_X^2}{N}, \tag{12}$$

或者

$$\sigma_{\bar{X}}=\frac{\sigma_X}{\sqrt{N}}. \tag{13}$$

證： 和證明定理 7 一樣，讓 $\omega=X$, $C_1=C_2\cdots\cdots=C_N=\frac{1}{N}$; 於是式 (9) 變做

$$\sigma_{\bar{X}}^2=\frac{1}{N^2}\sum_1^N\sigma_i^2, \tag{14}$$

因為 $\sigma_i^2=\sigma_X^2$, $(i=1,2,\cdots\cdots,N)$, 如此式 (14) 化做式 (12). 證訖.

現在我們以表 72 為例，討論此定理之應用.

表 72 裏所列之 100 個平均數能做成一個頻數分配，第一個樣本的均數為 2.0, 第二個的為 7.2, 等等. 各平均數之歸組分配列入表 73 第二行中. 其中有一樣本之平均數低至 1.2, 另一樣本之平均數高達 7.4. 該表第三、第四及第五行各列有 10 人，20 人及 50 人樣本平均數之分配. 研究此等分配，得知 5 人樣本平均數之全距很寬，並且樣本加大平均數之全距漸狹. 5 人樣本平均數全距之寬，乃在意料之中. 各樣本平均數之值由組段 .75—1.25 (有一值為 1.2) 至 7.25—7.75 (有一值為 7.4), 所以由 5 人樣本算得之平均數，因機會關係，距實在平均數 4.50 甚遠. 從他方面言之，平均數之極端值 (表 73 第二行) 則不多見，並且大部份

平均數靠近宇宙平均數(4.50); 100 平均數有 39 在 3.75 與 5.25 之間——卽是,樣本平均數有 39% 在眞正平均數 ±3/4 組段之間.

若以 100 個 10 人樣本代替 100 個 5 人樣本,其平均數分配之離勢較小,如表 73 第三行所示,所得之極端值在 1.75—6.75 之間,並且 100 平均數有 58 個在實在平均數± 3/4 組段之間;若以 100 個 20 人樣本代替 100 個 5 人樣本(表 73 第四行),其平均數之遠離程度更形縮小,所得之極端值在 2.75 與 6.25 之間,並且平均數有百分之 77 在實

表 73. 100 個 5 人、10 人、20 人及 50 人樣本平均數

(平均每人患感冒次數)之分配

樣本平均數 (1)	頻		數	
	5 人樣本 (2)	10 人樣本 (3)	20 人樣本 (4)	50 人樣本 (5)
0.75—	1	…	…	…
1.25—	1	…	…	…
1.75—	4	1	…	…
2.25—	2	2	…	…
2.75—	12	5	2	1
3.25—	15	8	9	5
3.75—	12	16	24	22
4.25—	10 }39	26 }58	31 }77	45 }91
4.75—	17	16	22	24
5.25—	8	15	10	3
5.75—	6	8	2	…
6.25—	7	3	…	…
6.75—	4	…	…	…
7.25—	1	…	…	…
總計	100	100	100	100
平均數	4.43	4.61	4.50	4.48

在平均數 ±3/4 組段之間；最後若以 100 個 50 人樣本代替 100 個 5 人樣本(表 73 第 5 行)，其平均數有百分之 91 在眞正平均數 ±3/4 組段之間，並且有百分之 45 在 4.25 與 4.75 之間，4.25 及 4.75 與眞正平均數相去無幾，數值在眞正平均數 ±3/4 組段限度之外者雖然也有，但罕見耳.

由表 73 及式 (13) 可見平均數之準確度至少要與樣本之大 N 有關. 樣本愈大所得之結果愈準確，即所得之數值愈與實在數值 (表示宇宙特性之數值)相切近. 然而樣本之大並不是影響數值準確性之惟一因子. 稍加思索或由式 (13) 便知其與宇宙數值之離勢 σ_x 亦有關係. 如果宇宙內每個事物僅有一值——即如上例宇宙內每個人患感冒確爲 3 次，於是不論樣本之大如何，所得之平均數一定必與實在平均數完全相同. 從他方面言之，如果患感冒次數由 0 至 900 次，而非由 0 至 9 次，各樣本平均數之離勢前者必大於後者. 可見由於樣本算得數值之準確性受兩個因子影響：一爲樣本之大；一爲在宇宙內事物特性之離勢，此宇宙爲樣本之所從出.

樣本平均數遠離眞正平均數之程度，得用眞正平均數加以或減以 3/4 組段權衡之. 已在上面說過了. 精密言之，此離勢宜用標準差表示之. 現在根據表 73 裏頻數分配直接算得之結果列入表 74，以便與由式 (13) 算得者比較：

由下表第三行可見，對於總平均數各平均數之標準差，由表 73 頻數分配算得者，因樣本內人數之增多而減少. 在實際上，我們不知道平均數之標準差，因爲通常我們祇有一個樣本. 就像上面表 72 患感冒之

表 74. 由表 73 及由式(13)分別算得之標準差

每樣本內事物之個數 N	100 平均數之平均數 $\bar{X}$	標準差: $\sigma_{\bar{x}}$	
		由表 73	由式(13)
5	4.43	1.36	1.28
10	4.61	0.91	0.91
20	4.50	0.61	0.64
50	4.48	0.44	0.41

例,我們祇計算一個樣本之單獨平均數——患感冒之平均次數.我們之問題爲:這個平均數之準確度如何?——即是如果我們取另一羣病人樣本,所得患感冒之平均次數與現在所得者之差別如何?如將同類樣本重複抽取之,各樣本之標準差應如何?

式(13)告訴我們說,平均數之標準差約等於樣本標準差除以樣本個數平方根之商.茲將用此式算得之結果列入表 73 末行中.其與同表第三行裏的結果相比較,可知用兩種方法算得者無大出入.此些微之出入,僅係由於取樣之波動而生,無足介意.

樣本大小不同,致各樣本平均數各與眞正平均數之離勢亦異,即如上面所指出者,5 人樣本 100 個平均數中有 39 個在眞正平均數加以或減以 3/4 組段之間,100 個樣本之總平均數 4.43 與宇宙之眞正平均數 4.50 不十分相近,因爲總共祇 500 個觀察值,其本身乃係一樣本.觀察值個數愈多,其平均數與眞正平均數愈相近,即如 100 個 50 人樣本之總平均數爲 4.48,其與 4.50 近得多了.

實在說，既知 100 個平均數在總平均數（或眞正平均數）加以或減以 3/4 組段之間有多少個，我們當然可以查驗 100 個中有若干個在總平均數加以或減以一倍標準差之間．由附錄 III 表 B 得知 100 平均數在總平均數加以或減以一倍標準差之間者均有 68 個，總平均數加以或減以 2 倍標準差之間者約有 95 個等等．如果樣本相當大，常能得到這樣的結果．於是我們可得下列之結論：

a. 如果由於宇宙內取一羣樣本，則各樣本之平均數未必與宇宙眞正平均數相等；但散佈在其左右．

b. 我們可用樣本平均數之標準差量度平均數之離勢；平均數在眞正平均數——樣本平均數之平均數——減以或加以 2 倍標準差之外者，必不多見．

假設僅有一個單獨平均數．現在我們要解說在實際上如何利用這個僅有的事實．在第二章第三節例 1 裏算得 147 個小學生身長之平均數爲 107.87 公分，我們想要知道牠的準確度——即是測得這個平均數與實在平均數（全體小學校學生之平均身長）相差多遠．

假設眞正平均數爲 $\bar{X}$，於是由上面的解說及式 (13) 得知觀察平均數大約在 $\bar{X}-2(\sigma_X/\sqrt{N})$ 與 $\bar{X}+2(\sigma_X/\sqrt{N})$ 之間．然而我們不知道 σ_X 之值，因之不得不由樣本估計之．我們算得 147 個量度之標準差爲 5.2929 公分．所以由此得到平均數標準差之估計値爲 .4365 公分．

我們得總結說（假設樣本爲隨機抽取者），觀察平均數或與眞正平均數有別，其差別大至 $\pm 2\sigma_{\bar{X}}=\pm 2\times(.4365)$ 公分，則不常見．換言之，眞正平均數大約在 $107.87\pm.8730$ 公分之間，或在 107.00 與 108.74 公分

之間，其值若在此界限之外，不易得到平均數爲 107.87 公分的樣本.

符號±後面之數值，如 ±.4365，叫做標準誤，其爲量度前面數值準確度之用. 如果樣本很小，估計值是沒用的，因爲以少數觀察值——如表 72 所列之任一樣本——之標準差代替宇宙標準差，易生嚴重之錯誤.

刊物上常有用機誤以代替標準誤者. 前者並非優於後者，不過機誤等於 .6744898 乘以標準誤耳. 所以標準誤之二倍約與機誤之三倍相等. 因數 .6744898 是這樣來的：即如上面各樣本平均數之例，100 個平均數在總平均數減以機誤與平均數加以機誤之間者恰恰有 50 個，在此界限之外者亦有 50 個. 如有一樣本，其平均數與眞正平均數之差大至 .6744898 乘以標準誤之機會爲 50%.

4. 偏態及峯態　關於頻數分配之形式的定理如下：

定理 9. 描寫樣本平均數分配偏態及峯態之動差與宇宙成對應之動差的關係爲

$$
(15)\qquad \begin{cases} \bar{\beta}_{1\bar{X}} = \dfrac{\bar{\beta}_{1X}}{N} \\ \bar{\beta}_{2\bar{X}} = 3 + \dfrac{1}{N}(\bar{\beta}_{2X} - 3) \end{cases}
$$

證：由式 (5) 及 (7) 我們有

$$E(\omega - \bar{\omega})^3 = E[C_1(X_1 - \bar{X}_1) + C_2(X_2 - \bar{X}_2) + \cdots\cdots + C_N(X_N - \bar{X}_N)]^3$$

及 $$E(\omega - \bar{\omega})^4 = E[C_1(X_1 - \bar{X}_1) + C_2(X_2 - \bar{X}_2) + \cdots\cdots + C_N(X_N - \bar{X}_N)]^4$$

應用期望值定理並將此二式展開之可得式 (15) 裏之結果.

定理 9 告訴我們關於平均數分配之形式. 如果宇宙爲常態，則 $\bar{\beta}_{1X} = 0$ 並且 $\bar{\beta}_{2X} = 3$，於是由式 (15) 得知 $\bar{\beta}_{1\bar{X}} = 0$ 並且 $\bar{\beta}_{2\bar{X}} = 3$. 就是

說,由於常態宇宙抽取許多樣本,$\bar{X}$之分配亦呈常態,不帶偏態與峯態.

5. 二量聯合分配 關於兩個相關變數 X 與 Y 聯合分配之三個重要定理如下:

定理 10. 讓 $(X_1, Y_1), (X_2, Y_2), \cdots\cdots, (X_N, Y_N)$ 爲一 N 雙變量樣本,由於受 (X, Y) 限制之分配內自由抽取者;並且讓 $(\bar{X}, \bar{Y})$ 表示由樣本算得之平均數.則所有各可能樣本平均數間之相關係數 R 等於宇宙相關係數 ρ.

證: 由定義

$$R=\frac{E(\bar{X}\bar{Y})-\tilde{X}\tilde{Y}}{\sigma_{\bar{X}}\sigma_{\bar{Y}}} \tag{16}$$

並且 $$E(\bar{X}\bar{Y})=\frac{1}{N^2}E\{(X_1+X_2+\cdots\cdots+X_N)(Y_1+Y_2+\cdots\cdots+Y_N)\}$$
$$=E(s)/N^2,$$

彼處
$$\begin{aligned} s= \ & X_1Y_1+X_1Y_2+\cdots\cdots+X_1Y_N \\ & +X_2Y_1+X_2Y_2+\cdots\cdots+X_2Y_N \\ & +\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & +X_NY_1+X_NY_2+\cdots\cdots+X_NY_N. \end{aligned}$$

我們將 s 分成兩部份:一部份叫做 u,代表 X_1Y_1, X_2Y_2 等項;另一部份叫做 v,代表 X_1Y_2, X_1Y_3 等項,卽是使

$$u=\Sigma X_iY_j \text{(共有 } N \text{ 項, } i=j)$$

並且 $$v=\Sigma X_iY_j \text{(共有 } N^2-N \text{ 項, } i\neq j).$$

在 u 式中, $E(u)=E\{\sum_1^N(X_iY_j)\}=\sum_1^N\{E(X_iY_j)\}=NE(XY)$;

在 v 式中,因 $i\neq j$, X_i 必與 Y_j 沒有關係,所以

$$E(X_iY_j)=E(X_i)E(Y_j)=\bar{X}\bar{Y},$$

並且
$$E(r)=(N^2-N)\bar{X}\bar{Y}.$$

如此我們有
$$E(s)=NE(XY)+(N^2-N)\bar{X}\bar{Y},$$

所以

$$E(\bar{X}\bar{Y})=\frac{1}{N}\{E(XY)+(N-1)\bar{X}\bar{Y}\}. \tag{17}$$

將上式代入式(16)內，並記得 $\sigma_{\bar{X}}=\frac{\sigma_X}{\sqrt{N}}$，$\sigma_{\bar{Y}}=\frac{\sigma_Y}{\sqrt{N}}$，則有

$$R=\frac{\frac{1}{N}\{E(XY)+(N-1)\bar{X}\bar{Y}\}-\bar{X}\bar{Y}}{\frac{1}{N}\sigma_X\sigma_Y}$$

$$=\frac{E(XY)-\bar{X}\bar{Y}}{\sigma_X\sigma_Y}=\rho.$$
證訖.

定理11. 讓 $\bar{X}$ 爲受 $g(X)$ 限制 N 個事物樣本之平均數，並且 $\bar{Y}$ 爲受 $h(Y)$ 限制 N 個事物樣本之平均數，彼處 $g(X)$ 及 $h(Y)$ 爲受 $f(X,Y)$ 限制之宇宙的邊和分配.讓 $\omega=\bar{X}-\bar{Y}$.求證 ω 樣本分配之變差

$$\sigma_\omega^2=\frac{1}{N}(\sigma_X^2-2\rho\sigma_X\sigma_Y+\sigma_Y^2). \tag{18}$$

此定理可由式(7)及定理10證明之.

定理12. 讓 $\bar{X},\bar{Y}$ 各爲 X 及 Y 之平均數，s_X 及 s_Y 各爲 X 及 Y 之標準差，並且 r 爲樣本之相關係數.假設 N 是如此大足以取 s_X^2，s_Y^2 及 r 各當做 σ_X^2，σ_Y^2 及 ρ 之良好估計.如此我們可寫

$$\sigma_{\bar{X}}^2=\frac{s_X^2}{N},\quad \sigma_{\bar{Y}}^2=\frac{s_Y^2}{N},\quad \rho=r.$$

於是 $\omega=\bar{X}-\bar{Y}$ 樣本分配之變差可用下式計算之:

(19) $$\sigma_d^2=\frac{1}{N}\left\{\Sigma(X_i-Y_i)^2-\frac{(\Sigma X_i-\Sigma Y_i)^2}{N}\right\}.$$

此式可由式 (18) 證明之.

6. 非常態及有限宇宙　由解析方面着想，關於由非常態宇宙內抽取許多樣本，其統計常數之實在分配莫由知之.最近芮志 (H. L. Rietz) 發表一篇論文 (見參考書 38, p. 114)，曾將前人關於此點之貢獻及發展做一概括之敍述. 學者宜閱覽之.

論及平均數,定理 3 告訴我們說,當 $N\to\infty$ 時, $\bar{\beta}_{1\bar{x}}\to 0$ 並且 $\bar{\beta}_{2\bar{x}}\to 3$. 如此，卽令樣本不十分大，各樣本平均數 $\bar{X}$ 之分配近於常態形式. $\bar{\beta}_{1\bar{x}}=0$ 並且 $\bar{\beta}_{2\bar{x}}=3$ 為常態分配之必需條件，而非充分條件. 卽令 N 之値較小，由於實驗上之證據，亦足以假定 $\bar{X}$ 之分配與常態相接近.

我們所思考的宇宙係假定其為無限的，卽是將宇宙內事物分成若干組，各組與樣本比較都是足夠大的. 有限宇宙亦可滿足此條件，例如在紅白球實驗中，袋內球數固然不是多至無限，但是若將球取出後還回原袋再採取之，則與以後取得紅球或白球之機率自然不生影響是.

如果 M 為有限宇宙之總頻數, $\bar{X}$ 分配之首四級動差

(20)* $$\begin{cases} E(\bar{X})=\bar{X} \\ \sigma_{\bar{x}}^2=\dfrac{M-N}{N(M-1)}\sigma_x^2 \\ \bar{\beta}_{1\bar{x}}=\dfrac{(M-1)(M-2N)}{N(M-N)(M-2)}\beta_{1x} \\ \bar{\beta}_{2\bar{x}}=\dfrac{(M-1)\{(M^2-6MN+M+6N^2)\bar{\beta}_{2x}+3M(M-N-1)(N-1)\}}{N(M-2)(M-3)(M-N)} \end{cases}$$

* Church, Biometrika, Vol. 18, p. 357, 1926.

式(20)比式(12)及(15)普遍些.若使 $M\to\infty$, 式(20)可變做式(12)及(15). 因此得到之結論爲，由於任意有限宇宙內抽取之樣本，其平均數之分配實際與常態相近.

7. 哲乞夫定理　在式(1)中以 ω 代 X, 讓 $G(\omega)=(\omega-\bar{\omega})^2$, 並且在算式 $E\{(\omega-\bar{\omega})^2\}$ 中所有 ω 大於 $\bar{\omega}+\varepsilon\sigma$ 者均以 $\bar{\omega}+\varepsilon\sigma$ 代之,所有 ω 小於 $\bar{\omega}-\varepsilon\sigma$ 者均以 $\bar{\omega}-\varepsilon\sigma$ 代之,彼處 ε 爲正數. 於是

$$E\{(\omega-\bar{\omega})^2\}=\int_{-\infty}^{\infty}(\omega-\bar{\omega})^2\Phi(\omega)d\omega$$

$$=\int_{-\infty}^{\bar{\omega}-\varepsilon\sigma}+\int_{\bar{\omega}-\varepsilon\sigma}^{\bar{\omega}+\varepsilon\sigma}+\int_{\bar{\omega}+\varepsilon\sigma}^{\infty}$$

$$\geq\kappa^2+(\varepsilon\sigma)^2P_\varepsilon$$

如此

(21)　$$E\{(\omega-\bar{\omega})^2\}\geq(\varepsilon\sigma)^2P_\varepsilon,$$

彼處 $\kappa^2=\int_{\bar{\omega}+\varepsilon\sigma}^{\bar{\omega}-\varepsilon\sigma}(\omega-\bar{\omega})^2\Phi(\omega)d\omega\geq0$, 並且 P_ε 爲 ω 在組段 $(\bar{\omega}-\varepsilon\sigma,\ \bar{\omega}+\varepsilon\sigma)$ 之外的機率.因爲 $E\{(\omega-\bar{\omega})^2\}=\sigma^2$, 於是由式(21)我們有

(22)　$$P_\varepsilon\leq\frac{1}{\varepsilon^2}.$$

所以有下面的定理.

定理 13.　設由宇宙 $\Phi(\omega)$ 任取一值 ω, 則 ω 與其期望值之差大於標準差乘以 ε 之機率不能大於 $1/\varepsilon^2$.

此定理稱爲哲乞夫定理(Tchebycheff's theorem), 或稱爲哲乞夫準則(Tchebycheff's criterion), 或稱爲哲乞夫不等式(Tchebycheff's inequality).

8. 哲乞夫定理之應用 哲乞夫定理最顯著之性質爲其與 ω 之分配種類無關,此定理之優點卽在此. 正以其如此寬泛,用以判定某已知分配之抽樣波動,無實在價值. 然而自有其用途在. 茲舉兩個重要用途——一關於大數樣本,一關於樣本波動之機率尺度——於下,以見一斑.

a. 伯努里定理之證明 伯努里定理現在可以證明了. 讓 $\omega=X/s$, X 爲在 s 次試驗中之成功次數; 於是 $\bar{\omega}=p$. 讓 P_ε 爲 X/s 在組段 $(p-\Delta, p+\Delta)$ 之外的機率,彼處 Δ 爲正數. 我們可以取 $\Delta=\varepsilon(pq/s)^{\frac{1}{2}}$. 因此

$$\frac{1}{\varepsilon}=\frac{(pq/s)^{\frac{1}{2}}}{\Delta}$$

並且由定理 13, 我們得到不等式

$$P_\varepsilon\leq\frac{p(1-p)}{s\Delta^2}.$$

指定 Δ 值後,祇要加大 s 值能使 P_ε 很小. 如此 s 愈加大,用 X/s 估計之值亦愈可靠.

b. 大樣本穩定性之證明 哲乞夫不等式仍能用以證明大樣本平均數之穩定性. 假設宇宙 $\Phi(X)$ 之變差爲 σ^2, 由這樣的宇宙內抽取 N 個単物樣本. 讓

$$\omega=\frac{1}{N}(X_1+X_2+\cdots\cdots+X_N).$$

假設 c^2 爲一常數,並且 $\sigma^2\leq c^2$. 因 $\omega=\bar{X}$, 由式 (12) 我們有

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{N}$$

$$\leq\frac{c^2}{N}.$$

設 P_ε 爲 $(X-\bar{X})^2>h^2$ 之機率. 卽是, P_ε 爲

$$(X-\bar{X})^2>\frac{Nh^2}{c^2}\cdot\frac{c^2}{N}$$

$$>\frac{Nh^2}{c^2}\sigma_{\bar{x}}^2$$

之機率. 所以由定理 13, 得

$$P_\varepsilon\leq\frac{c^2}{Nh^2}.$$

因上式中之 c 及 h 之值不變, 如此祇要加大 N 值便可使 P_ε 之值很小. 於是有下面的定理.

定理 14. 樣本平均數與宇宙平均數之差大於已知正數 h 之機率, 祇要加多觀察值之個數, 能使之很小.

在此定理之條件下, 我們依統計術語說: X 收斂至 $\bar{X}$. 此與數學上所說的收斂不同, 切勿與之淆混.

c. 機率尺度　假設 σ_X 爲已知者. 我們知道各 X 對於 $\bar{X}$ 之分配爲(或近於)常態, 而以 $\sigma_{\bar{x}}=\sigma_X/\sqrt{N}$ 爲其標準差. 如將 X 之分配用變換式

$$t=\frac{X-\bar{X}}{\sigma_X/\sqrt{N}} \tag{23}$$

化做標準單位, 則知 t 對於零點之分配近於常態 而以 1 爲其標準差.

因之一個隨機樣本平均數 X 與 $\bar{X}$ 之差大至 $|\varepsilon|$ 的機率得用常態機率尺度表示之,彼處 ε 係用 $\sigma_{\bar{x}}$ 單位表之者. 如此我們有下面的定理.

定理 15. 假設由於無限宇宙抽取一隨機樣本,其平均數 X 在宇宙平均數 $\bar{X}$ 加以 ε 或減以 ε 之內的機率 Q_ε 與 $2\int_0^\varepsilon \phi(t)dt$ 相近,彼處 ε 爲 t 之觀察値由式 (23) 算得者, 並且 $\phi(t)$ 爲常態曲線函數. 於是 $P_\varepsilon=1-Q_\varepsilon$ 爲 X 不在 $\bar{X}-\varepsilon$ 與 $\bar{X}+\varepsilon$ 之內的相近機率(見圖 66). 如果宇宙爲常態, P_ε 爲實在機率.

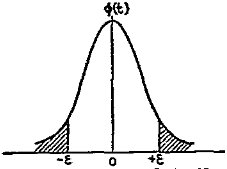


圖 66. 離差 E 與機率 P_ε 之關係。
P_ε 爲圖上帶有陰影部分,Q_ε 爲曲線下空白部分。

9. 顯著性測驗 研究者常顧慮到,機率 P_ε 小至某種程度方可承認其所得之結果有顯著性,非係由於樣本波動呢?通常所採用之規則爲:

如果 $P_\varepsilon \geq .05$, 稱 ε 爲不顯著;

如果 $P_\varepsilon \leq .01$, 稱 ε 爲顯著;

如果 $.05 > P_\varepsilon > .01$, 則

ε 有無顯著性是很難確定的,尚有待於事實之補充. 更有人採取較爲嚴格之水準.

例 1. 假設 100 個人身長之平均數 $X=70.56$ 英寸. 問此數與宇宙

平均數 $\bar{X}=69.943$ 英寸有顯著之差別否？已給宇宙標準差 $\sigma_X=3.115$ 英寸.

因為
$$\varepsilon=70.56-69.943=0.617,$$
$$t=\frac{70.56-69.943}{3.115/\sqrt{100}}=1.98.$$

故參照常態機率尺度得知，觀察與理論平均數差量大至 ε 之機率 $P_\varepsilon=.0477$. 於是我們得到的結論為：已給之統計常數 $\bar{X}=70.65$, 雖然其有來自另一宇宙之可能. 就是說在此情形中，$\bar{X}$ 與 $\bar{X}$ 可能為兩個不同民族之平均身長；但因 $\bar{X}=70.56$ 並非例外，謂其來自同一宇宙未始不可.

例 2. 投擲 12 枚骰子 26,306 次，且以擲得 5 點或 6 點算做成功. 觀察分配之平均數為 4.0524 點. 投擲真的骰子而得 5 點或 6 點之機率為 $\frac{1}{3}$, 如此其頻數分配與 $(\frac{2}{3}+\frac{1}{3})^{12}$ 展開式各項成比例. 所以，若骰子是真的，期望平均數為 $sp=12\times\frac{1}{3}=4$. 試用觀察及理論平均數之差當做判斷準則，測驗骰子之真偽.

因
$$N=26,306,$$
$$\bar{X}=4.0524,$$
$$\sigma_X=\sqrt{12\times\tfrac{1}{3}\times\tfrac{2}{3}}=1.633,$$
$$\bar{X}=4.$$

故
$$\varepsilon=4.0524-4=.0524,$$
$$\frac{\sigma_X}{\sqrt{N}}=.010.$$

因之
$$t=\frac{.0524}{.010}=5.2.$$

離差在 $t=\pm 5$ 外之機率是非常小的，如此我們得說，骰子不是完全立方體都是帶有偏誤的.

第五節 比例數

在實際統計工作上頗關重要之另一數值為百分數.例如，由於肺炎病人樣本我們要計算其死亡率.讓我們假設，由於觀察大批資料之經驗，肺炎死亡率假如為百分之 20. 我們任選患肺炎者 100 人以血清治療之.如果血清治療無效，約有 20 人死亡.因為機會作祟，我們所得之觀察死亡率或者確為 20%，或者大於此數，或者小於此數.假設觀察到 10 人死亡；所取之樣本係 100 病人，此事件之發現是否由於機會使然？如果此事件易由機會而起.於是我們更可推論說，血清雖有價值，但其效不顯，最低限度需其有顯著功效之證據不足.在未得到結論之前，我們須將樣本加大.自他方面言之，如果此事件非係偶然——即是非由於機會使然——我們可以推論說，血清對於治療肺炎有效.先假設血清治療無效，於是觀察大樣本之死亡率應當為 20%（或者與之相近）.樣本之大不同，所得之百分率當亦各異，究竟由於不同樣本算得之百分率當與 20% 相差多遠呢？

1. 樣本之大與所得結果之關係 現在我們取大小不同樣本，分別加以研究如下：

a. 一人樣本 如果我們的樣本僅由一人組成，則算得之死亡率不是 0，即是 100%. 若病人死去，其死亡率大於以往之經驗；倘其痊愈自然不能歸功於治療.因為依照以往經驗，5 個患者即或不予治療，亦僅

有 4 人痊愈(死亡率爲百分之 20, 即每 5 人有 1 人死亡).

b. 二人樣本 如將樣本增至二人,事件之發現有三種可能:二人痊愈,一愈一死及二人全死.

以過去之經驗作根據,我們可以計算每種事件發現之機率,因爲一人痊愈之機會 $\frac{4}{5}$,另一人痊愈之機會亦爲 $\frac{4}{5}$,故二人皆愈之機會爲這兩個不相干機會之相乘積——即是爲 $\frac{4}{5}\times\frac{4}{5}=\frac{16}{25}$.因爲一人痊愈之機會仍爲 $\frac{4}{5}$,另一人死亡之機會爲 $\frac{1}{5}$,故二人一愈一死之機會爲 $\frac{4}{5}\times\frac{1}{5}=\frac{4}{25}$;但是此比值應以 2 乘之,因爲事件發生之方法有二——A 愈 B 死,或 A 死 B 愈.故一愈一死之總機率爲 $2\times\frac{4}{5}\times\frac{1}{5}=\frac{8}{25}$.最後因爲每人死亡之機率 $\frac{1}{5}$,故二人皆死亡之機會爲 $\frac{1}{5}\times\frac{1}{5}=\frac{1}{25}$.此等數值可用下表表示之.

表 75. 二人樣本,每人死亡率爲 1/5(=20%),

二人發現事件之機率及其死亡百分率

事　件	發現機率	死亡百分率
二人皆愈	16/25 = 0.64	0
一愈一死	8/25 = 0.32	50
二人皆死	1/25 = 0.04	100
總　計	25/25 = 1.00	......

很顯然的,因爲二人皆愈之死亡率爲 0,與已往經驗之死亡率(百分之 20)相比,得知僅僅二人皆愈指示治療有效.至於一人死亡之死亡率百

分之 50, 二人皆死之死亡率百分之 100, 二率皆遜於已往經驗. 但二人皆愈爲比較容易發現之事件,故不能說二人皆愈不是機會使然;即或治療無效,此事件在 100 次試驗中有發現 64 次之可能.故用二人作樣本,通常之死亡率爲百分之 20, 二人皆愈之機會是相當大的, 如果觀察到此種結果,我們不能推論血清之治療有效.

c. 三人樣本 若將樣本增大至三人, 四種事件可以發生: 三人皆愈,二愈一死,一愈二死及三人皆死.

每事件之機率可計算之如前. 三人皆愈之機會爲 $\left(\frac{4}{5}\right)^3=\frac{64}{125}$; 二愈一死之機會爲 $3\left(\frac{4}{5}\right)^2\left(\frac{1}{5}\right)=\frac{48}{125}$; 一愈二死之機會爲 $3\times\frac{4}{5}\times\left(\frac{1}{5}\right)^2=\frac{12}{125}$; 最後,三人皆死之機會爲 $\left(\frac{1}{5}\right)^3=\frac{1}{125}$. 列表則有:

表 76. 三人樣本,每人死亡率爲 1/5(=20%),
三人發現事件之機率及其死亡百分率

事 件	發 現 機 率	死亡百分率
三人皆愈	64/125=0.512	0.0
二愈一死	48/125=0.384	33.3
一愈二死	12/125=0.096	66.7
三人皆死	1/125=0.008	100.0
總 計	125/12. =1.000	......

僅僅三人皆愈指示治療有效. 其他事件所得之死亡率均較已經之經驗爲高.但三人皆愈易由機會而起;在 100 次試驗中,即令治療無效,可能

發現之次數爲51. 故用三人作樣本，三人皆愈之機會也是相當大的，我們也不能推論特殊治療有何價值.

d. 四人樣本 如將樣本增至四人，則有五種可能事件發生：四人皆愈，三愈一死，二愈二死，一愈三死及四人皆死. 今已往經驗做根據，各種事件發現之機率爲何？四人皆愈之機率爲 $\left(\frac{4}{5}\right)^4=\frac{256}{625}$；三愈一死之機會爲 $4\left(\frac{4}{5}\right)^3\left(\frac{1}{5}\right)=\frac{256}{625}$；二愈二死之機會爲 $6\left(\frac{4}{5}\right)^2\left(\frac{1}{5}\right)^2=\frac{96}{625}$；僅一人痊愈並且三人死亡之機會爲 $4\left(\frac{4}{5}\right)\left(\frac{1}{5}\right)^3=\frac{16}{625}$；最後，四人皆死之機會爲 $\left(\frac{1}{5}\right)^4=\frac{1}{625}$. 列表:

表77. 四人樣本，每人死亡率爲1/5(=20%)，
四人發現事件之機率及其死亡百分率

事件	發現機率	死亡百分率
四人皆愈	256/625＝0.4096	0
三愈一死	256/625＝0.4096	25
二愈二死	96/625＝0.1536	50
一愈三死	16/625＝0.0256	75
四人皆死	1/625＝0.0016	100
總計	625/625＝1.0000	......

又是所有病人皆愈之死亡率爲推一之死亡率較已往經驗爲低，但此亦是極易由於機會而生；即或治療無效，發現之機會幾在100次試驗中有41次.

e. 十人樣本　樣本之大不拘，我們可用此法終能算得各種事件之死亡率，並斷定其是否由於機會而生；然而當樣本增加至過大，計算各率之手續便極紛繁．所幸我們無須算出所有各率，例如治療十個病人，那麼所要的結果只限於較已往經驗爲佳者：十個病人皆愈，或僅一人死亡者，卽是死亡率爲百分之 0 或 10 者．如果十人死二人死亡率爲 20%，此與已往經驗相同．如果三人或者三人以上死亡，則是比已往經驗爲高．10 人皆愈之機會爲$(4/5)^{10}=.1047$，9 愈 1 死之機會爲 $10(4/5)^9(1/5)=0.2687$，故 10 人皆愈及 9 愈 1 死之發生機率爲 $0.1074+0.2684=0.3758$．換言之，卽或治療無效，在 100 次實驗中約有 38 次較已往經驗爲佳，故發生 10 人皆愈及 9 愈 1 死之事件未必非由機會使然．因此，我們不能推論血清確能減低死亡率．

f. 百人樣本　現在我們話歸前題——卽是 100 個病人樣本，其中僅有 10 人死亡，我們所欲求者爲：假令血清治療完全無效，我們能否希望死亡率爲 10% 或較此結果爲低？可能用上法同樣計算之．茲將計算結果列入表 78 中．

表 78 內各機率之和爲等於 0.005696，此爲欲求之機率；並且我們可以推論說：在 10000 次實驗中，每次用 100 人作實驗，約有 57 次死亡率爲 10% 或較此爲低．卽是，結果爲 10% 或低於 10% 之機率極小，其僅爲 .005696．結果如此者，顯示用血清治療之病人其死亡率的確降低．但是用此法計算非常繁瑣，用一簡單方法速算乃切實際之需要．

表78.　百人樣本，每人死亡率為 1/5(=20%)，

百人發現事件之機率及其死亡百分率

事　件	發　現　機　率		死亡百分率
100 人皆愈	$(4/5)^{100}$	$=0.000000$	0
99 愈 1 死	$100(4/5)^{99}(1/5)$	$=0.000000$	1
98 愈 2 死	$4,950(4/5)^{98}(1/5)^2$	$=0.000000$	2
97 愈 3 死	$161,700(4/5)^{97}(1/5)^3$	$=0.000001$	3
96 愈 4 死	$3,921,225(4/5)^{96}(1/5)^4$	$=0.000003$	4
95 愈 5 死	$75,287,520(4/5)^{95}(1/5)^5$	$=0.000015$	5
94 愈 6 死	$1,192,052,400(4/5)^{94}(1/5)^6$	$=0.000059$	6
93 愈 7 死	$16,007,560,800(4/5)^{93}(1/5)^7$	$=0.000199$	7
92 愈 8 死	$186,087,894,300(4/5)^{92}(1/5)^8$	$=0.000578$	8
91 愈 9 死	$1,902,231,808,400(4/5)^{91}(1/5)^9$	$=0.001478$	9
90 愈 10 死	$17,310,309,456,440(4/5)^{90}(1/5)^{10}$	$=0.003363$	10

9. 一般情形　試我們重談二人樣本．如果有 100 個醫院，各有病人 2，其死亡百分率為 0 者有 64 個，為 50 者有 32 個，為 100 者有 4 個．由於這些數字，我們得以計算 100 個醫院之平均死亡百分率，以及此頻數分配之標準差．

$$\text{平均數}=(0\times64+50\times32+100\times4)\div100=20,$$

$$\text{標準差}=\sqrt{\frac{0^2\times64+(50)^2\times32+(100)^2\times4}{100}-(20)^2}=28.3.$$

同樣，在三人樣本中，我們可以算得平均死亡率及標準差各為 20% 及 23.1%；在四人樣本中，平均死亡率及標準差各為 20% 及 20.0%；在十人樣本中，死亡率及標準差各為 20% 及 12.6%；在百人樣本中，平均

死亡率及標準差各爲 20.0% 及 4.0%. 各樣本之平均死亡率都是 20%，但每個樣本之觀察平均數未必恰是 20%，因爲機會從中作祟，觀察平均數與眞正平均數常是有差別，此差別卽以離勢常數度量之者. 我們無須知道各個事件之機率，就能算得標準差. 於是由第三章式(32)算得

$$2\text{ 人樣本之標準差}=\sqrt{\frac{20\times 80}{2}}=28.3,$$

$$3\text{ 人樣本之標準差}=\sqrt{\frac{20\times 80}{3}}=23.1,$$

$$4\text{ 人樣本之標準差}=\sqrt{\frac{20\times 80}{4}}=20.0,$$

$$10\text{ 人樣本之標準差}=\sqrt{\frac{20\times 80}{10}}=12.6,$$

$$100\text{ 人樣本之標準差}=\sqrt{\frac{20\times 80}{100}}=4.0.$$

此等數值與上面由頻數分配直接算得者相同.

2. 由已知條件計算樣本之大　如果限定 X 與 $\bar{X}$ 之差不大於 $|\varepsilon|$，並且 X 在 $\bar{X}-|\varepsilon|$ 與 $\bar{X}+|\varepsilon|$ 之間的機率爲一固定值；則樣本之大可由下式推得:

$$N=\frac{t^2pq}{\varepsilon^2}, \tag{24}$$

彼處 N 爲樣本之大，t 爲標準單位，p 爲宇宙內事物具有某種品質之比例數，$q=1-p$，ε 爲 X 與 $\bar{X}$ 之差.

例. 假設某地瘧疾染患率，據有悠久之經驗者談，向爲 20%. 欲證實所言不誤，爰舉行一瘧疾調查. 若限定所得之百分率在 19.8% 與

20.2% 之間，卽是容忍 1% 的誤差；並且限定 100 個樣本中有 99 個樣本的百分率在 19.8% 與 20.2% 之間．問所取之樣本之大應當如何？——卽是應當調查若干人？

因爲 $p=20\%$, $\varepsilon=.2\%$, 並且 100 個樣本的百分率有 99 個在指定限度之內，我們應取 $t=2.58$; 故

$$N=\frac{(2.58)^2\times.2\times.8}{(.002)^2}=266,256.$$

3. 兩個比例數之差 假設有兩個隨機樣本：在第一個樣本內有 N_1 個事物，其具有某種品質者有 X_1 個；在第二個樣本內有 N_2 個事物，其具有該種品質者有 X_2 個．我的問題爲，觀察所得之比率數 $\frac{X_1}{N_1}$ 與 $\frac{X_2}{N_2}$ 之差是偶然的，抑或二樣本所屬之宇宙亦有同樣差異存在？

a. 定理 下列之定理可以回答上面所問之問題．

定理 16. 如果 $\frac{X_1}{N_1}$ 及 $\frac{X_2}{N_2}$ 爲隨機並且不相干之樣本，由於無限宇宙內抽取者，宇宙內具有某種品質事物所佔之比率爲 p, 則所得觀察比率之差大至 $\omega=\left|\frac{X_1}{N_1}-\frac{X_2}{N_2}\right|$ 之機率爲與 P_ε 相近，彼處 P_ε 之定義見定理 15, 並且

$$\varepsilon=\frac{\omega}{\sigma_\omega},\quad \sigma_\omega=\sqrt{pq\left(\frac{1}{N_1}+\frac{1}{N_2}\right)}.$$

證： 由伯努里定理，$\frac{X_1}{N_1}$ 對於期望值 p 之變化而以 $\frac{pq}{N_1}$ 爲其變差，彼處 $q=1-p$; 同樣，$\frac{X_2}{N_2}$ 對於 p 之變化而以 $\frac{pq}{N_2}$ 爲其變差．於是

$$\omega=\left|\left(\frac{X_1}{N_1}-p\right)-\left(\frac{X_2}{N_2}-p\right)\right|$$

$$E(\omega)=E\left\{\left(\frac{X_1}{N_1}-p\right)-E\left(\frac{X_2}{N_2}-p\right)\right\}=0,$$

且並由式(10)

$$\sigma_\omega^2=\frac{pq}{N_1}+\frac{pq}{N_2}=pq\left(\frac{1}{N_1}+\frac{1}{N_2}\right) \tag{25}$$

所以，σ_ω^2 爲 ω 對於零點離勢之變差，且並比值

$$t=\frac{\omega}{\sigma_\omega} \tag{26}$$

表示以標準爲單位 ω 對於零點之離勢。　　證訖。

關於 t 分配形式之事實可由高級動差得之．不難證明

$$\begin{cases}\bar{\beta}_1=\dfrac{1-4pq}{pq}\times\dfrac{(N_1-N_2)^2}{N_1N_2(N_1+N_2)},\\ \bar{\beta}_2=3+\dfrac{1-6pq}{pq}\times\dfrac{N_1^2-N_1N_2+N_2^2}{N_1N_2(N_1+N_2)}\end{cases} \tag{27}$$

若取樣本之大無限，p 及 q 之値固定，可見 $\bar{\beta}_1\to 0$ 及 $\bar{\beta}_2\to 3$. 卽或樣本不十分大，t 之分配去常態形式亦不太遠。

b. 規則　下列之規則對於實際工作上是有裨益的，特敍述之如下：

規則 I.　假設 $N_1<N_2$. 如果 $N_1p>5$, 得用常態幾率尺度．如果 $N_1p\leq 5$, 須考驗 $\bar{\beta}_1$. 若 $\bar{\beta}_1<.04$, 用常態幾率尺度仍是很準確的．但是若 $\bar{\beta}_1\geq .04$, 對於所得之結果不可輕易置信．

規則 II.　當應用定理 16 時，須估計 p 之値．因此常取

$$\bar{p}=\frac{X_1+X_2}{N_1+N_2} \tag{28}$$

當做 p 之最佳估計值。

c. 對照組　在前面標準差之計算是以已往經驗比率數作根據——例如由已往經驗中知道死亡率爲百分之 20. 在實際統計工作中，可資根據之已往經驗實不多覯.通常爲代替已往經驗起見，作實驗的人往往取一對照組以與實驗組相比較. 例如，爲適應此目的起見，作實驗的人常按下列之方法收搜資料:

50 個患肺炎者用普通法治療，其死亡率爲……　　30%,

50 個患肺炎者於用普通法之外更用某種特殊方法治療，

其死亡率爲……　　10%.

由此資料以研究此兩種死亡率之差是否僅由於機會而起.很顯然的，兩個百分率因爲機會作祟，樣本與樣本不同；如果某種特殊方法一點效驗沒有，實驗許許多多次，其結果必定有時實驗組死亡率低、有時對照組低、更有時無甚軒輊.就是說，兩組死亡率絕對不應該有顯然之差別.我們當前的問題是，其差別達到何種程度，幾不是出於偶然的呢?換言之，我們需要去求兩個百分率之差的標準差. 假使樣本係由每人患感冒次數宇宙內抽取者，並且計算各樣本內患感冒者之百分率.譬如宇宙內未患感冒之百分率爲 10.

d. 5 人樣本　如由宇宙內任取二樣本，每樣本包含 5 人，我們未必觀察到二樣本未患感冒之百分率爲 10; 在一樣本中，我們得到之百分率或爲 20, 在他樣本中或爲 60, 其差爲百分之 40. 此差是否易由機會而起?爲解決此問題 由宇宙內取 100 雙樣本，各包含 5 人. 取得樣本後，計算 5 人中未患感冒者之百分數，並記載其差量——例如在第一樣本內 5 人中有 1 人未患感冒，至第二樣本內 5 人中有 2 人未患感冒;

故其百分率各為 20 及 40，其差為 20. 此法可得到各差量之分配，如表 79 所示.

由表 79 可見，幾乎樣本之半數 (45 例)，無差別，但有者距 0 甚遠. 譬如有 6 例其差為 40，並且有一例其差大至 60. 100 雙百分率之差的平均數很近於 0，其為 -0.6，但是最常見之差量為百分之 20. 差量之散佈情形和通常一樣，可用其標準差量度之；其為 18.0. 差量的分配是相當對稱的，我們知道像這樣的分配，數值在平均數減以或加以 2 倍標準差之外者，則不多見. 於是我們說：若有 5 人樣本，其期望之百分率為 10，則兩個樣本百分率之差固宜為 0，實際很容易觀察到二量之差非 0，而為 $\pm 2\times 18.0\% = \pm 36.0\%$. 但差量常在此限界之內，在其外者則少見.

表 79. 100 個差量之頻數分配 (5 人樣本)

差 量	頻 數
-60	1
-40	2
-20	26
0	45
20	22
40	4
60	…
總 計	100

c. 20 人樣本　如將樣本增大至 20，則得另一差量分配. 其最大之

差量僅爲 25 (見表 80), 差量之平均數也是與 0 相近——卽是 −1.3, 但標準差變做 9.53; 樣本加大 4 倍, 差量之標準差減少 2 倍. 於是我們可以推論說, 差量在 $\pm 2\times 9.5\% = \pm 19.0\%$ 之間者, 易由機會而起, 大於此者比較罕見.

表 80. 100 個差量之頻數分配 (20 人樣本)

差　量	頻　數
−25	...
−20	6
−15	6
−10	15
−5	14
0	27
5	14
10	12
15	3
20	2
25	1
總　計	100

f. 50 人樣本　最後 50 人樣本差量分配之頻數如下:

可見現在之一百變差量中其無超過 14% 者; 其平均數爲 +0.9, 其標準差爲 5.7. 在此例中, 每個樣本預料之百分率爲 10%, 每雙樣本百分率之差將必爲 0, 其差量大都落在 $\pm 2\times 5.7 = \pm 11.4$ 之間, 並且較大之差量則不多見.

表 81.　100 個差量之頻數分配 (50 人樣本)

差　量	頻　數
−14	1
−12	…
−10	4
− 8	6
− 6	7
− 4	2
− 2	19
0	6
2	17
4	17
6	10
8	2
10	3
12	5
14	1
總　計	100

由表 79, 80 及 81 算得之結果,可知量差之標準差或標準誤因樣本之加大而減少.

g. 差量標準誤　每個百分數之標準誤為 $\sqrt{\frac{pq}{N}}$ [見第三章式 (32)], 彼處 p 為宇宙內事物具有某種品質之百分率,例如曾染患感冒者之百分率; q 為事物不具有此種品質之百分率,例如未染患感冒一次以上之百分率; 並且 N 為樣本中包含事物之個數. 兩個百分率差量之標準誤

已在本節證明爲 $\sqrt{pq/N_1+pq/N_2}$, 彼處 N_1 及 N_2 各爲樣本內包含事物之個數. 例如由於 $p=10\%$ ——如此 $q=90\%$ 之宇宙內抽取 5 人樣本, 各雙樣本百分率差量之標準誤爲 $\sqrt{\frac{.10\times .90}{5}+\frac{.10\times .90}{5}}=19.0\%$. 換言之, 由同一宇宙內抽取包含 5 人之樣本, 每雙樣本之差量應爲 0, 但實際最易得到之差量爲 $(\pm 2\times 19.0\%=38.0\%)$. 此爲理論值. 其與前面由 100 個實際 5 人樣本算得者 $(\pm 2\times 18.0\%=36.0\%)$, 相差極微. 同理由 100 個 20 人實際樣本算得之標準誤爲 9.5. 根據理論算得者爲 $\sqrt{\frac{.10\times .90}{20}+\frac{.10\times .90}{20}}=9.5\%$. 最後, 100 個實際 5 人樣本差量之標準差爲 5.7%. 理論者爲 $\sqrt{\frac{.10\times .90}{50}+\frac{.10\times .90}{50}}=6.0\%$.

很顯然的, 如果知道宇宙未患感冒之百分率, 我們得以計算差量之標準差, 如此斷定該差量是否由於機會使然. 例如, 由宇宙內抽取 2 個 50人樣本, 其用甲種維生素治療者, 經過一定時間感染到感冒百分率爲 4%, 其未經治療者, 在相同時間內, 感冒百分率爲 14%. 標準差可示差量是否由於機會而起. 此處二樣本之差量爲 $14\%-4\%=10\%$. 並且其標準差爲 6.0%, 換言之, 由同一宇宙內取兩個 50 人樣本, 所得之差量每易爲 $\pm 2\times 6.0\%=\pm 12.0\%$; 故差量爲 10% 者, 謂爲由於機會而起未始不可. 若謂染患感冒百分率之降低爲甲種維生素之功, 實覺不妥, 因爲即令甲種維生素治療無效, 亦易得到同樣之差量.

在實際上我們無法知道宇宙內之 p 值; 計算標準誤時, 不得不用計算值代替之. 以計算值代替時, 有兩條不同之途徑可循:

(i) 我們有兩個樣本, 包含 50 人, 並且其一未患感冒者之百分率

爲 4%，其他爲 14%. 設兩個樣本由同一宇宙內抽取者，宇宙內未患感冒者之百分率爲 "x"，我們問 x 與 4 及 x 與 14 之差別是否僅由於機會而起？讓我們暫時假定所有樣本係由同一宇宙內抽取者，我們估計 "x" 之值，最好之估計法由於全體觀察值——卽是 100 個觀察值中其未患感冒者之百分率——估計之.其爲 9%$\left(=\frac{4+14}{100\times100}\%\right)$. 我們的問題現在變爲："如果兩個 50 人樣本，由於同一宇宙內抽取者，宇宙內未患感冒者之百分率爲 9%，我們得到之差量爲 10% 而未得到 0，此差量 10% 是否由於機會使然？"若答以是，於是我們必須承認觀察之差量或者是眞實的，但是他很能僅由機會使然，重復實驗之或者就要消失. 自他方面言之，若答以否，我們可推論說，我們若說兩個樣本由同一宇宙內抽取者非也——卽是兩個樣本百分率之差量，起之於機會並非易易，我們須尋求生成這個差量之原由.

在上例裏，其觀察之差量爲 $14\%-4\%=10\%$. 根據上面之假定,差量之準標差爲 $\sqrt{\frac{.09\times.91}{50}+\frac{.09\times.91}{50}}=5.7\%$. 觀察差量小於其標準差之二倍，故我們得以推論說：此量差，僅由於機會而生，非係罕見之事——卽是得到如此大之差量，乃常事耳. 如果其差爲 15%，我們庶幾可以推論說：此差未必由於機會使然，因其大於標準差之二倍，如此必有生成此差量之原由，或謂其由於甲種維生素之效果使然.

(ii) 保留各樣本係由於同一宇宙內抽取者之假定，並且和前先一樣我們不知道宇宙內未患感冒者之百分率. 但樣本供給我們兩個估計值，如此供給我們兩個標準誤. 第一個估計值爲 4%，標準誤爲

$\sqrt{\frac{.04\times .96}{50}}$；第二個估計值為 14%，標準誤為 $\sqrt{\frac{.14\times .86}{50}}$. 兩個百分率差量之標準誤於是算得為 $\sqrt{\frac{.04\times .96}{50}+\frac{.14\times .86}{50}}=5.6\%$. 因為觀察之差量為 10%，而小於標準誤之二倍，與兩個估計值來自同一宇宙之原始假定相符合. 換言之，根據所得之事實不足以證明這兩個觀察百分率之差非由於機會而起.

可見由兩個方法所得之結果極為相近，在實際上常常如此；但在統計工作中我們率循途徑 (i).

第六節　χ^2 分配及其應用

1. χ^2 之分配　χ^2 分配得用皮爾生頻數曲線型 III 描寫之. 現在我們證明之如下：

如將第四章式 (36) 中之 X 以 $t\left(=\frac{X-\bar{X}}{\sigma}\right)$ 代之，並且使其中之 $c_2=0$；則同章式 (46) 中之 $a=\frac{1}{2}\alpha_3$, $c_0=-1$, $c_1=a$. 於是該章式 (74) 變做

$$y=y_0\left(1+\frac{t}{s_0}\right)^{s_1}e^{-s_2t}, \tag{29}$$

此時式中之 $y_0=\frac{Ns_1^{s_1+1}}{s_0e^{s_1}\Gamma(s_1+1)}$, $s_0=a-\frac{1}{a}$, $s_1=\frac{1}{a^2}-1$, $s_2=-\frac{1}{a}=\frac{2}{\alpha_3}$.

使
$$s_2=m,\ 則\ s_1=m^2-1,\ s_0=m-\frac{1}{m}$$

並且式 (29) 變做

$$y=y_0s_0^{-s_1}\left(m-\frac{1}{m}+t\right)^{m^2-1}e^{-mt};$$

再使 $m^2=\frac{n}{2}$ 及 $m\left(m-\frac{1}{m}+t\right)=\frac{X^2}{2}$，我們於是有

$$y=y_0s_0^{-s_1}\left(\frac{X^2}{2m}\right)^{\frac{n}{2}-1}e^{-\frac{X^2}{2}+m^2-1} \tag{30}$$

$$=y_0'\left(\frac{X^2}{2}\right)^{\frac{n}{2}-1}e^{-\frac{X^2}{2}}.$$

彼處 $y_0'=y_0s_0^{-s_1}m^{1-n/2}e^{m^2-1}$，其可測定之為 $\frac{N}{\Gamma\left(\frac{n}{2}\right)}$。式(30)即為 X^2 之機率分配.

有了式(30)我們可用下式以求 P 之值(參第六章第七節 c):

$$P=\frac{\int_{X^2}^{\infty}K\left(\frac{X^2}{2}\right)^{\frac{n}{2}-1}e^{-\frac{X^2}{2}}dX^2}{\int_0^{\infty}K\left(\frac{X^2}{2}\right)^{\frac{n}{2}-1}e^{-\frac{X^2}{2}}dX^2} \tag{31}$$

$$=\frac{\int_X^{\infty}x^{n-1}e^{-x^2/2}dx}{\int_0^{\infty}x^{n-1}e^{-x^2/2}dx}，n \text{ 為自由度}.$$

2. X^2 之應用　用 X^2 測驗品度配合頻數分配之適度情形已詳見第六章第七節例 1 裏，不再贅述。現在僅將其對於其他方面之應用分別敍述之並舉例明說之，如下:

四格表　X^2 測驗最普通之應用為應用於四格表上面，表內事物具有兩種品質可能組合之頻數係已知者.在此情形中，由第六章式(30)很容易證明

$$X^2=\frac{(ad-bc)^2N}{(a+b)(c+d)(b+d)(a+c)}. \tag{32}$$

彼處 a 代表事物兼具甲乙兩種品質之頻數，b 代表事物具有甲種品質但不具有乙種品質之頻數，c 代表事物不具有甲種品質但具有乙種品質之頻數，d 代表事物不具有甲種或乙種品質之頻數，並且 $N=a+b+c+d$.

例 1. 設有論據如表 82 所示. 問人之脾臟用透視法看出腫大與用顯微鏡由其血片查到瘧原蟲有無聯繫?

表 82. 檢查 5501 人脾臟及血片之結果

瘧原蟲	脾臟 曾腫大	脾臟 未腫大	總計
有	740	743	1483
無	1287	2731	4018
總計	2027	3474	5501

$$X^2=\frac{(740\times 2731-743\times 1287)^2\times 5501}{1483\times 4018\times 3474\times 2027}$$

$$=\frac{(1064699)^2\times 5501}{5958694\times 7041798}=148.6.$$

因爲若給四格表之各邊綠頻數 (marginal frequency)，則與表格內四個頻數祇有一個是自由的，即是四格表之自由度 $n=1$. 於是由附錄 III 表 C 查得 $X^2=148.6$ 之 P 值時，我們得到 $X^2=15.1321$, $P=0.0001$. 如此 $X^2=148.6$ 之 P 值必較 0.0001 爲小，即小於萬分之一. 就是說，實在觀察之頻數與理論頻數迥乎不同，理論頻數係假設脾腫大與血片中瘧原蟲之有無沒有聯繫得之者. 所以我們可以得推論說，兩種品質——脾腫大與血液中存有瘧原蟲——確有顯著之聯繫.

b. 兩個觀察樣本之比較 X^2 測驗之另外一種應用爲用以比較兩個觀察樣本之頻數分配. 假設我們有下列之問題: 已給兩個現象之頻數分配, 其特徵僅由樣本知之. 一方面我們要問, 這兩個樣本由於同一宇宙隨機抽取之機率爲何? 另一方面要問, 一個分配實在與另一分配不同之度大於僅由機會而起之機率爲何?

皮爾生曾經證明: 如果兩個樣本所從出之宇宙的頻數分配爲

$$m_1, m_2, m_3, \cdots\cdots, m_p, \cdots\cdots, m_s,$$

其和等於 M, 並且兩個樣本之頻數分配

第一個樣本爲 $f_1, f_2, f_3, \cdots\cdots, f_p, \cdots\cdots, f_s$,

第二個樣本爲 $f_1', f_2', f_3', \cdots\cdots f_p', \cdots\cdots, f_s'$,

其和各等於 N 及 N', 彼處 N 及 N' 相差或者很大, 亦或很小; 則

$$X^2 = NN' \sum \frac{\left(\frac{f_p}{N} - \frac{f_p'}{N'}\right)^2}{f_p + f_p'} \tag{33}$$

例 2. 馬克端納 (David Mac Donald) 曾得表 83 裏所列的結果.* 問猩紅熱與痲疹之感染是否因兒童髮色而異? 換言之, 猩紅熱與痲疹對於髮色分配之彼此不同是否易由機會而起?

運算之詳細詳見表 84.

由表 84 我們算得

$$X^2 = NN' \times .0000211 = 1864 \times 661 \times .0000211 = 26.00,$$

並且自由度 $r = 4$, 於是由附錄 III 表 A 查得 P 值小於 .01; 換言之 如

* Biometrika, vol. 8, pp. 13—39, 1911.

表 83. 感染猩紅熱與麻疹對於髮色之關係

髮色	病人數目	
	猩紅熱	麻疹
黑色	12	0
微黑	289	85
中常	1109	367
微紅	360	184
紅色	94	25
總計	1864	661

果表 84 中之兩個分配為隨機樣本由於同一宇宙內抽取者,在 100 次試驗中僅有一次之值等於或大於 26.00. 我們可以推論說,這兩個樣本的確不同. 或者說猩紅熱與麻疹之感染,不是不因髮色而異,即是猩紅熱與麻疹各選擇髮色不同的人而進攻之.

表 84. 測驗顯著性之計算

手續	髮色					總計
	黑色	微黑	中常	微紅	紅色	
(1): 猩紅熱	12	289	1109	360	94	1864
(2): 麻疹	0	85	367	184	25	661
(3): (1)+(2)	12	374	1476	544	119	2525
(4): (1)/1864	.0064	.1551	.5950	.1931	.0504	1.0000
(5): (2)/661	.0000	.1286	.5552	.2784	.0378	1.0000
(6): (4)−(5)	.0064	.0265	.0398	−.0853	.0126	.0000
(7): $(6)^2$	.000041	.000702	.001584	.007276	.000159	......
(8): (7)÷(3)	.0000034	.0000019	.0000011	.0000134	.0000013	.0000211

第七節 標準誤

關於平均數$\bar{X}$,百分數或比例數p,二組或兩個平均數之差(m_1-m_2)及兩個百分數或比例數之差(p_1-p_2)標準誤之公式各爲

(34) $$\sigma_{\bar{x}}=\frac{\sigma_X}{\sqrt{N}},\quad [\text{式 (13)}]$$

(35) $$\sigma_p=\sqrt{\frac{pq}{N}},\quad [\text{由第三章式 (32) 得}]$$

(36) $$\sigma_{m_1-m_2}=\sqrt{\sigma_{m_1}^2+\sigma_{m_2}^2},\quad [\text{由式 (10) 得}]$$

(37) $$\sigma_{p_1-p_2}=pq\sqrt{\frac{1}{N_1}+\frac{1}{N_2}},\quad [\text{由式 (25) 得}]$$

前已詳加研討.現在導演其他重要統計常數標準誤之公式於下,以作本章之結束.

1. 某組頻數之標準誤 假設我們計算動差時所根據之論據係已經分過組的,譬如說分成了n組;並且各組之期望頻數爲$f_1, f_2, \cdots\cdots, f_n$,彼處$f_1+f_2+\cdots\cdots+f_n=N$.讓我們先計算第$i$組頻數$f_i$之樣本波動標準誤.如果取樣不存偏見,則樣本第$i$組頻數必與宇宙第$i$組頻數成比例.由宇宙抽取一事物落在第$i$組之機率爲$\frac{f_i}{N}$,其不落在第$i$組之機率爲$1-\frac{f_i}{N}$.由宇宙取出$N$個事物,落在各組之頻數可由二項式

$$N\left\{\left(1-\frac{f_i}{N}\right)+\frac{f_i}{N}\right\}^N$$

求之;以f_i爲該分配之期望値(平均數),並以

(38) $$\sigma_{f_i}=\sqrt{N\frac{f_i}{N}\left(1-\frac{f_i}{N}\right)}$$

$$=\sqrt{f_i\left(1-\frac{f_i}{N}\right)}$$

為其標準差. 此為式 (35) 之另一種寫法.

2. 對於定點 q 級動差之標準誤　由定義, 對於定點 X_0 之 q 級動差 λ_q, 彼處

$$N\nu_q = \Sigma(x_i^q f_i)$$

並且 $x = X - X_0$.

如此用 $\delta\nu_q$ 代表 ν_q 伴 δf_i 而生離差, 我們有

$$N\delta\nu_q = \Sigma x_i^q \delta f_i.$$

將上式兩邊平方之,

$$N^2(\delta\nu_q)^2 = (x_1^q\delta f_1 + x_2^q\delta f_2 + \cdots + x_n^q\delta f_n)^2$$

$$= \Sigma\{x_i^{2q}(\delta f_i)^2\} + 2\Sigma'(x_i^q x_j^q \delta f_i \delta f_j),$$

彼處 Σ' 表示 i 及 j 經過一切值各項之和, 但將 $i=j$ 之項除外.

因為末式對於任意樣本是有效的, 故將所有各樣本之離差合計之

得
$$N\sigma_{\nu_q}^2 = \Sigma\{x_i^{2q}\sigma_{f_i}^2\} + 2N\Sigma'(x_i^q x_j^q r_{\delta f_i\delta f_j}\sigma_{f_i}\sigma_{f_j}).$$

但
$$r_{\delta f_i\delta f_j} = -\frac{f_i f_j}{N\sigma_{f_i}\sigma_{f_j}},\ \text{(證明見腳註*)}$$

---

* 假設二組之頻數為 f_i 及 f_j, δf_i 為第 i 組觀察頻數與期望或理論頻之離差, 並且 δf_j 為第 j 組觀察頻數與理論頻數之離差. 因為總頻數為 N, $N-f_i$ 為頻數分配於 i 組以外其他各組者. 如果在第 i 組得到一離差 δf_i, 於是 $(-\delta f_i)$ 之分配於其他各組. 所以分配於第 j 組之離差應為 $\frac{f_j}{N-f_i}(-\delta f_i)$. 平均言之,

$$\delta f_j = \frac{f_j}{N-f_i}(-\delta f_i).$$

所以
$$\delta f_i \delta f_j = -\frac{f_j}{N-f_i}(\delta f_i)^2.$$

將兩邊所有各離差加之, 我們有

$$\sigma_{f_i}\sigma_{f_j} r_{\delta f_i\delta f_j} = -\frac{f_j}{N-f_i}\sigma_{f_i}^2.$$

如此由式 (38),

(39)
$$\sigma_{f_i}\sigma_{f_j} r_{\delta f_i\delta f_j} = -\frac{f_i f_j}{N}.$$

$$\sigma_{f_i}=\sqrt{f_i\left(1-\frac{f_i}{N}\right)},\quad \text{[見式 (38)]}$$

因之,我們有

$$\begin{aligned}N^2\sigma_{\nu_q}{}^2&=\Sigma\left\{x_i^{2q}f_i\left(1-\frac{f_i}{N}\right)\right\}-2\Sigma'\left(x_i^q x_j^q\frac{f_if_j}{N}\right)\\&=\Sigma(x_i^{2q}f_i)-\frac{1}{N}\Sigma(x_i^qf_i)\Sigma(x_j^qf_j)\\&=N\nu_{2q}-N\nu_q^2.\end{aligned}$$

所以

$$\sigma_{\nu_q}=\sqrt{\frac{\nu_{2q}-\nu_q^2}{N}}. \tag{40}$$

3. 對於平均數 q 級動差之標準誤 前面已經討論過了對於定點(或定數)動差之標準誤.在實際中,我們最常用者爲對於樣本平均數之動差.因爲此平均數本身受樣本波動之影響的,對於平均數動差之標準誤往往與對於定數者不同.

如果 m 爲平均數,由定義,我們有

$$\begin{aligned}N\mu_q&=\Sigma\{(x_i-m)^qf_i\}\\&=\Sigma(x_i^qf_i)-qm\Sigma(x_i^{q-1}f_i)+T,\end{aligned} \tag{41}$$

彼處 T 代表含有 m^2 及二次以上各項之式.

現在讓 m 變至 $m+\delta m$, f_i 變至 $f_i+\delta f_i$, 並 μ_q 變至 $\mu_r+\delta\mu_q$. 我們有

$$N(\mu_q+\delta\mu_q)=\Sigma\{x_i^q(f_i+\delta f_i)\}-q(m+\delta m)\Sigma\{x_i^{q-1}(f_i+\delta f_i)\}+T. \tag{42}$$

由式 (42) 減以式 (41),

$$\begin{aligned}N\delta\mu_q&=\Sigma(x_i^q\delta f_i)-q\delta m\Sigma(x_i^{q-1}f_i)-q\Sigma(x_i^{q-1}\delta m\delta f_i)+U\\&=N\delta\nu_q-Nq\nu_{q-1}\delta m-Nq\delta m\delta\nu_{q-1}+U.\end{aligned}$$

彼處 U 代表含有 m 及其高次項之式．因爲帶有 $\delta m \delta \nu_{q-1}$ 之項較其他各項爲小，我們可以棄捨之，於是將末式平方之並將所有各樣本加之，得

$$\sigma^2_{\mu_q}=\sigma_{\nu_q}{}^2+q^2\nu_{q-1}{}^2\sigma_m{}^2-2q\nu_{q-1}\sigma_m\sigma_{\nu_q}r_{m\nu_q}+U.$$

但
$$\sigma_{\nu_q}\sigma_{\nu_r}r_{\nu_q\nu_r}=\frac{\nu_{q+r}-\nu_q\nu_r}{N},\quad \text{(證明見脚註*)}$$

$$\sigma_{\nu_q}=\sqrt{\frac{\nu_{2q}-\nu_q^2}{N}},\quad \text{[見式 (40)]}$$

於是
$$\sigma^2_{\mu_q}=\frac{\nu_{2q}-\nu_q{}^2+q^2\nu_{q-1}{}^2\nu_2-2q\nu_{q-1}\nu_{q+1}}{N}+U.$$

現在使 $m=0$. 上式中之 U 項不見，輔助動差各變做主要動差，並且上式變做

$$(43)\qquad \sigma^2_{\mu_q}=\sqrt{\frac{\mu_{2q}-\mu_q{}^2+q^2\mu_2\mu_{q-1}{}^2-2q\mu_{q-1}\mu_{q+1}}{N}}.$$

4. 變差之標準誤　依據本節已得之結果我們能夠討論許多統計常數之標準誤，該結果係用皮爾生方法推求者（見參考書 36）．

由式 (43)，使 $q=2$，因 $\mu_1=0$，我們有

$$(44)\qquad \sigma_{\mu_2}=\sqrt{\frac{\mu_4-\mu_2{}^2}{N}}$$

---

* 因爲

$$N\delta\nu_q=\Sigma x_i{}^q\delta f_i.$$
$$N\delta\nu_r=\Sigma x_i{}^r\delta f_i.$$

相乘得
$$N^2\delta\nu_q\delta\nu_r=\Sigma(x_i{}^{q+r}\delta f_i^2)+\Sigma'\{(x_i{}^qx_j{}^r+x_i{}^rx_j{}^q)(\delta f_i\delta f_j)\}.$$

將所有各樣本加之，

$$N^2\sigma_{\nu_q}\sigma_{\nu_r}r_{\nu_q\nu_r}=\Sigma(x_i{}^{q+r}\sigma^2_{f_i})+\Sigma'\{(x_i{}^qx_j{}^r+x_i{}^rx_j{}^q)(\sigma_{f_i}\sigma_{f_j}r_{f_if_j})\}.$$

由式 (38) 及 (39) 我們有

$$(45)\qquad \sigma_{\nu_q}\sigma_{\nu_r}r_{\nu_q\nu_r}=\frac{\nu_{q+r}-\nu_q\nu_r}{N}.$$

此卽變差 μ_2 之標準誤.

如果宇宙爲常態,

$$\mu_2=\sigma^2,\ \mu_4=3\sigma^4\ (\text{學者試證之});\text{則}$$

(46)
$$\sigma_{\mu_2}=\sqrt{\frac{3\sigma^4-\sigma^4}{N}}=\sigma^2\sqrt{\frac{2}{N}}.$$

5. 標準差之標準誤　如果 μ_2 爲變差,我們有

$$\mu_2=\sigma^2.$$

所以
$$\mu_2+\delta\mu_2=(\sigma+\delta\sigma)^2=\sigma^2+2\sigma\delta\sigma+(\delta\sigma)^2.$$

$(\delta\sigma)^2$ 與 $\delta\sigma$ 相比較可以捨棄之,於是

$$\delta\mu_2=2\sigma\delta\sigma.$$

將末式平方之並將各樣本相加之,得

$$\sigma_{\mu_2}^2=4\sigma^2\sigma_\sigma^2.$$

因之
$$\sigma_\sigma=\frac{\sigma_{\mu_2}}{2\sigma}$$

(47)
$$=\sqrt{\frac{\mu_4-\mu_2^2}{4N\sigma^2}}.$$

如果宇宙爲常態,式 (47) 變做

(48)
$$\sigma_\sigma=\frac{\sigma}{\sqrt{2N}}.$$

6. 對於平均數三級及四級動差之標準誤.

由式 (43), 使 $q=3$,

(49)
$$\sigma_{\mu_3}=\sqrt{\frac{\mu_6-\mu_3^2-6\mu_4\mu_2+9\mu_2^3}{N}}.$$

如果分配爲常態,則

$$\mu_6=15\sigma^6,\ \mu_4=3\sigma^4,\ \mu_3=0,\ \mu_2=\sigma^2.$$

所以

$$\sigma_{\mu_3} = \frac{\sigma^3}{\sqrt{N}}\sqrt{15-18+9} = \sigma^3\sqrt{\frac{6}{N}}. \tag{50}$$

同樣，由方程式(43)，使 $q=4$,

$$\sigma_{\mu_4} = \sqrt{\frac{\mu_8 - \mu_4^2 - 8\mu_5\mu_3 + 16\mu_2\mu_3^2}{N}}. \tag{51}$$

如果分配為常態，$\mu_8 = 105\sigma^8$, $\mu_5 = 0$.

所以

$$\sigma_{\mu_4} = \frac{\sigma^4}{\sqrt{N}}\sqrt{105-9} = \sigma^4\sqrt{\frac{96}{N}}. \tag{52}$$

7. 離勢係數之標準誤　設離勢係數(即標準差係數) V 之定義(見第二章第五節 5)，用現在符號表之，為

$$V = \frac{100\sigma}{m} = \frac{100\sqrt{\mu_2}}{m},$$

式中 m 及 σ 各代表變數 X 的平均數及標準差. 於是

$$\begin{aligned} V + \delta V &= \frac{100\sqrt{\mu_2 + \delta\mu_2}}{m + \delta m} \\ &= \frac{100\sqrt{\mu_2}}{m}\left(1 + \frac{\delta\mu_2}{\mu_2}\right)^{\frac{1}{2}}\left(1 + \frac{\delta m}{m}\right)^{-1} \\ &= V\left(1 + \frac{\delta\mu_2}{2\mu_2} + \cdots\right)\left(1 - \frac{\delta m}{m} + \cdots\right). \end{aligned}$$

如將較小於 $\delta\mu_2$ 及 δm 之量捨棄之，則有

$$V + \delta V = V\left(1 + \frac{\delta\mu_2}{2\mu_2} - \frac{\delta m}{m}\right).$$

因之
$$\frac{\delta V}{V} = \frac{\delta\mu_2}{2\mu_2} - \frac{\delta m}{m}$$

將此式平方之,並將所有各樣本相加之,我們有

$$\frac{\sigma_V^2}{V^2}=\frac{\sigma_{\mu_2}^2}{4\mu_2^2}+\frac{\sigma_m^2}{m^2}-\frac{1}{\mu_2 m}\sigma_{\mu_2}\sigma_m r_{\mu_2 m}.$$

如果分配爲常態,則 $\sigma_{\mu_2}^2=\frac{2\sigma^4}{N}$, $\sigma_m^2=\frac{\sigma^2}{N}$, 並且 $r_{\mu_2 m}=0$

所以
$$\frac{\sigma_V^2}{V^2}=\frac{1}{2N}+\frac{\sigma^2}{m^2N}.$$

卽是

(53)
$$\sigma_V=\frac{V}{\sqrt{2N}}\sqrt{1+2\left(\frac{V}{100}\right)^2}.$$

8. β_1 及 β_2 之標準誤　β_1 及 β_2 之標準誤可依上面相同之法求之.因爲

$$\beta_1=\frac{\mu_3^2}{\mu_2^3},$$

故
$$\beta_1+\delta\beta_1=\frac{(\mu_3+\delta\mu_3)^2}{(\mu_2+\delta\mu_2)^3}$$

變化之得

$$\delta\beta_1=\frac{2\mu_3\delta\mu_3}{\mu_2^3}-\frac{3\mu_3^2}{\mu_2^4}\delta\mu_2.$$

將此式平方之,並將所有各樣本相加之,則有

$$\sigma_{\beta_1}^2=\frac{4\mu_3^2}{\mu_2^6}\sigma_{\mu_3}^2+\frac{9\mu_3^4}{\mu_2^8}\sigma_{\mu_2}^2-\frac{12\mu_3^3}{\mu_2^7}\sigma_{\mu_3}\sigma_{\mu_2}r_{\mu_3\mu_2}.$$

$$N\sigma_{\beta_1}^2=\frac{4\mu_3^2}{\mu_2^6}(\mu_6-\mu_3^2-6\mu_4\mu_2+9\mu_2^3)$$
$$+\frac{9\mu_3^4}{\mu_2^8}(\mu_4-\mu_2^2)-\frac{12\mu_3^3}{\mu_2^7}(\mu_5-4\mu_2\mu_3).$$

但　$\beta_{2r+1}=\frac{\mu_3\mu_{2r+3}}{\mu_2^{r+3}}$,　$\beta_{2r}=\frac{\mu_{2r+2}}{\mu_2^{r+1}}$,　[見第三章式 (25) 之脚註]

故

$$(54)\qquad \sigma_{\beta_1}{}^2=\frac{\beta_1}{N}(4\beta_4-24\beta_2{}^2+36+9\beta_1\beta_2-12\beta_3+35\beta_1).$$

同樣

$$(55)\qquad \sigma_{\beta_2}{}^2=\frac{1}{N}(\beta_6-4\beta_2\beta_4+4\beta_2{}^3-\beta_2{}^2+16\beta_2\beta_1-8\beta_3+16\beta_1).$$

9. 相關係數及迴歸係數之標準誤　茲將求相關係數 r 及迴歸係數 b_{12} 之公式列下:

$$(56)\qquad \sigma_r=\frac{1-r^2}{\sqrt{N}}.$$

此公式之證明詳見 Phil. Trans. Roy. Soc., Series A, vol. 191, 1898, p. 229., 茲從略.

$$(57)\qquad \sigma_{b_{12}}=\frac{\sigma_1\sqrt{1-r_{12}}}{\sigma_2\sqrt{N}}=\frac{\sigma_{1\cdot2}}{\sigma_2\sqrt{N}}.$$

此式仍適用於任意次數之迴歸係數,即是

$$(58)\qquad \sigma_{b_{12\cdot k}}=\frac{\sigma_{1\cdot2k}}{\sigma_{2\cdot k}\sqrt{N}}.$$

彼處 k 表示次下標 (secondary subscript) 之集團,但 1 或 2 除外.

問題 VIII

1. 假設變數 ω 分配爲常態,我們隨意選取一個數值,則此值在 $E(\omega)\pm3\sigma_\omega$ 之內與在其外之比爲 369 : 1.

2. (a) 假設有一有限宇宙,其中包含五個變數 x_1、x_2、x_3、x_4 及 x_5. 取得三個不同變數樣本之數目爲 $C(5, 3)=10$. 將牠們寫出來

(b) 設 $\bar{x}_i$ 代表第 i 個樣本平均數.寫出這 10 個不同樣本平均

數，例如

$$\bar{x}_1=\frac{x_1+x_2+x_3}{3}.$$

(c) 證明 10 值 $\bar{x}_i(i=1、2、3、\cdots、10)$ 之平均數為 5 值 $x_i(i=1、2、3、\cdots、5)$ 之平均數.

即證明
$$\frac{1}{10}\Sigma\bar{x}_i=\frac{1}{5}\Sigma x_i=\bar{x}.$$

3. 證明 ω^2 之期望值大於 ω 期望值之平方.

4. 假設 1000 同年齡人之平均體重為 140 磅，標準差為 20.0 磅. 問此樣本平均數之標準差為何? 問此平均數與同年齡宇宙平均數之差不大於五磅之機會為何?

5. 美國人口之平均死亡年齡為 59.13 歲，芝加哥城為 58.98 歲，假設芝加哥人口數為 24,000，標準差為 10 歲，並且宇宙之分配為常態. 問該二年齡之差是否由於機會而起?

6. 一個共濟團體欲確定其團體同人之平均死亡年齡為不比期望年齡 59.13 歲大至 1 歲. 所謂 '確定' 者，意思為 Q_ϵ 必須等於或小於 .001. 此團體應有若干人? (假設標準差為 10 歲).

7. 已給 $\omega=\sum_1^k(f_i+x_i).$

如果各 x 之值是不相干的，並且 $\sum_1^k f_i$ 為常數. 證明

$$\sigma_\omega^2=\sum_1^k\sigma_i^2.$$

彼處 σ_i^2 代表 x_i 之變差.

8. 求 $x^2+y^2=r^2$ 在第一象限內縱坐標之平均數.

(a) 沿 x 軸等分,

(b) 沿圓周等分.

9. 求曲線 $y=a+bx^2$ 由 0 至 x 沿 x 軸等分之縱坐標的平均數.

10. 投擲銅元 128 次，並記載正面之頻數. 應用 X^2 測驗探討觀察與理論分配之脗合情形.

11 假設有 300 隻白鼠，各個患有癌病，病況之嚴重性相同，任意分為兩組，一組有白鼠 100 隻，一組有 200 隻. 第一組用血清治療，第二組不用血清治療，除此之外，二組其他治療情形相同. 用血清治療之組有 8 隻白鼠死亡，他組有 25 隻死亡. 測驗二組白鼠死亡率之差的顯著性.

12. 某教授擔任同樣課程兩班，一班有學生 20 人，一班有學生 30 人. 小班學生成績在乙等以上者有 4 人，大班有 8 人. 這個差別除去樣本波動之外還有其他解說否?

13. 假設 X 依照頻數曲線 $Y=Ce^{-X/a}$ 分配，$0\leq X\leq\infty$，a 為正數，C 為待定之常數(調定時使常態曲線下之面積等於 N). 計算 ν_{rX}($r=1$、2、3 及 4). 然後求 μ_{rX}，最後得 $\bar{X}=a$, $\sigma_X=a$, $\beta_1=4$, $\beta_2=9$.

14. 假設下表(摘錄自參考書 45)英國人之身長為常態分配，並且算得其平均數及標準差各為 67.46 吋及 2.57 吋. 如果任選一人，選取之法與選 8585 人時同，問此人在 5 呎與 6 呎之間的機率為何?

身長(吋)	成年男人數目
57—	2
58—	4
59—	14
60—	41
61—	83
62—	169
63—	394
64—	669
65—	990
66—	1223
67—	1329
68—	1230
69—	1063
70—	646
71—	392
72—	202
73—	79
74—	32
75—	16
76—	5
77—	2
總計	8,585

15. 由下頁之表(爲據由參考書 45)求中位數(154.7 磅),兩個四分位數(142.5 磅及 168.4 磅)之標準誤.

16. 求上題裏爲據四分位差之標準誤.

17. 練習題 15 分配之標準誤爲 21.3 磅,求平均數之標準誤並與在該練習題內算得之中位數標準誤比較之.

18. 兩個宇宙之平均數相同,但其一之標準誤二倍於其他.由各宇宙內任取 500 人樣本,其平均數之差的機率不能超過 0.3σ、彼選 σ 爲

兩個標準差之小者;假設平均數之差爲常態分配.求差量大於 0.15σ 之機率.

體重(磅,最近尺度)	人數
90—	2
100—	34
110—	152
120—	390
130—	867
140—	1623
150—	1559
160—	1326
170—	787
180—	476
190—	263
200—	107
210—	85
220—	41
230—	16
240—	11
250—	8
260—	1
270—	...
280—	1
總計	7,749

19. 1000 塊農田樣本,某年之平均產麥量爲每畝 2000 磅,而以 192 磅爲其標準差. 1000 塊農田樣本下年之平均產麥量爲每畝 2100 磅而以 224 磅爲其標準差.證明由這論據可以總結說,農田二年之產麥量相同.

20. 下表由參考書 44) 爲投擲骰子實驗之結果:

投擲 12 枚骰子 4096 次，擲得 6 點視做成功

成功數目	頻 數
0…………	447
1…………	1145
2…………	1181
3…………	796
4…………	380
5…………	115
6…………	24
7及以上…	8
總 計	4096

假設骰子不偏斜，求 X^2 之值，並且由此證明論據與此假設一致。

21. 證明式(33)。

22. 設有兩個調查員在某鎮內分別進行調查，其目的在欲知該鎮居民之貧富情形，因此將貧富分爲貧窮、中常及小康三組(貧富之界按其收入之多寡定之，兩個調查員同)。所得之結果爲：

調查員	收入			計 總
	貧 窮	中 常	小 康	
A	140	100	15	255
B	140	50	20	210
總 計	280	150	35	465

證明至少有一調查員玩忽調查工作。

附　　錄

I. 算學常數

在統計工作中，爲求算得結果準確計，我們往往需要許多小數位之算學常數，因此將其常用者列舉如下，俾資應用：

$e=$	2.71828 18284 59045	$\frac{1}{e}=$	0.36787 94411 71442
$e^2=$	7.38905 60989 30650	$\frac{1}{e^2}=$	0.13533 52832 36613
$\sqrt{e}=$	1.64872 12707 00128	$\frac{1}{\sqrt{e}}=$	0.60653 06597 12633
$\log_{10}e=$	0.43429 44819 03252	$\log_e 10=$	2.30258 50929 94046
$\pi=$	3.14159 26535 89793	$\frac{1}{\pi}=$	0.31830 98861 83791
$\pi^2=$	9.86960 44010 89359	$\frac{1}{\pi^2}=$	0.10132 11836 42338
$\sqrt{\pi}=$	1.77245 38509 05516	$\frac{1}{\sqrt{\pi}}=$	0.56418 95835 47756
$\sqrt{2\pi}=$	2.50662 82746 31001	$\frac{1}{\sqrt{2\pi}}=$	0.39894 22804 01433
$\sqrt{\frac{\pi}{2}}=$	1.25331 41373 15500	$\sqrt{\frac{2}{\pi}}=$	0.79788 45608 02865
$\sqrt[3]{\pi}=$	1.46459 18875 61523	$\sqrt[3]{\frac{1}{\pi}}=$	0.68278 40632 55296
1弳$=$	$57^\circ.29577$ 95130 82321	$1^\circ=$	0.01745 32925 19943 弳

$1\text{弳} = 3437'.746\ 77078\ 49393 \qquad 1' = 0.00029\ 08882\ 08666\ \text{弳}$

$1\text{弳} = 206264''.80624\ 70964 \qquad 1'' = 0.00000\ 48481\ 36811\ \text{弳}$

$\sqrt{2} = 1.41421\ 35623\ 73095 \qquad \sqrt[3]{2} = 1.25992\ 10498\ 94873$

$\sqrt{3} = 1.73205\ 08075\ 68877 \qquad \sqrt[3]{10} = 2.15443\ 46900\ 31884$

$\sqrt{10} = 3.16227\ 76601\ 68379 \qquad \sqrt[3]{100} = 4.64158\ 88336\ 12779$

II. 希臘字母及讀音

在統計學上,往往用希臘字母代表算得之常數或運算之符號,就像,用 σ 代表標準差, Σ 代表許多數值相加之和,等等.因此將希臘字母及讀音列表如下:

希臘字母及讀音

字母		讀音	字母		讀音
Α	α	alpha	Ν	ν	nu
Β	β	beta	Ξ	ξ	xi
Γ	γ	gamma	Ο	ο	omicron
Δ	δ	delta	Π	π	pi
Ε	ε	epsilon	Ρ	ρ	rho
Ζ	ζ	zeta	Σ	σ, ς	sigma
Η	η	eta	Τ	τ	tau
Θ	θ	theta	Υ	υ	upsilon
Ι	ι	iota	Φ	φ	phi
Κ	κ	kappa	Χ	χ	chi (ki)
Λ	λ	lambda	Ψ	ψ	psi
Μ	μ	mu	Ω	ω	omega

III. 統計用表

表 A. X^2 之值：自由度由 1 至 30，機率由 .99 至 .01

自由度	機率: P					
n	.99	.98	.95	.90	.80	.70
1	.000157	.000628	.00393	.0158	.0642	.148
2	.0201	.0404	.103	.211	.446	.713
3	.115	.185	.352	.584	1.005	1.424
4	.297	.429	.711	1.064	1.649	2.195
5	.554	.752	1.145	1.610	2.343	3.000
6	.872	1.134	1.635	2.204	3.070	3.828
7	1.239	1.564	2.167	2.833	3.822	4.671
8	1.646	2.032	2.733	3.490	4.594	5.527
9	2.088	2.532	3.325	4.168	5.380	6.393
10	2.558	3.059	3.940	4.865	6.179	7.267
11	3.053	3.609	4.575	5.578	6.989	8.148
12	3.571	4.178	5.226	6.304	7.807	9.034
13	4.107	4.765	5.892	7.042	8.634	9.926
14	4.660	5.368	6.571	7.790	9.467	10.821
15	5.229	5.985	7.261	8.547	10.307	11.721
16	5.812	6.614	7.962	9.312	11.152	12.624
17	6.408	7.255	8.672	10.085	12.002	13.531
18	7.015	7.906	9.390	10.865	12.857	14.440
19	7.633	8.567	10.117	11.651	13.716	15.352
20	8.260	9.237	10.851	12.443	14.578	16.266
21	8.897	9.915	11.591	13.240	15.445	17.182
22	9.542	10.600	12.338	14.041	16.314	18.101
23	10.196	11.293	13.091	14.848	17.187	19.021
24	10.856	11.992	13.848	15.659	18.062	19.943
25	11.524	12.697	14.611	16.473	18.940	20.867
26	12.198	13.409	15.379	17.292	19.820	21.792
27	12.879	14.125	16.151	18.114	20.703	22.719
28	13.565	14.847	16.928	18.939	21.588	23.647
29	14.256	15.574	17.708	19.768	22.475	24.577
30	14.953	16.306	18.493	20.599	23.364	25.508

表 A. χ^2 之值: 自由度由 1 至 30, 機率由 .99 至 .01 (續完)

自由度 n	機率: P .50	.30	.20	.10	.05	.02	.01
1	.455	1.074	1.642	2.706	3.841	5.412	6.635
2	1.386	2.408	3.219	4.605	5.991	7.824	9.210
3	2.366	3.665	4.642	6.251	7.815	9.837	11.341
4	3.357	4.878	5.989	7.779	9.488	11.668	13.277
5	4.351	6.064	7.289	9.236	11.070	13.388	15.086
6	5.348	7.231	8.558	10.645	12.592	15.033	16.812
7	6.346	8.383	9.803	12.017	14.067	16.622	18.475
8	7.344	9.524	11.030	13.362	15.507	18.168	20.090
9	8.343	10.656	12.242	14.684	16.919	19.679	21.666
10	9.342	11.781	13.442	15.987	18.307	21.161	23.209
11	10.341	12.899	14.631	17.275	19.675	22.618	24.725
12	11.340	14.011	15.812	18.549	21.026	24.054	26.217
13	12.340	15.119	16.985	19.812	22.362	25.472	27.688
14	13.339	16.222	18.151	21.064	23.685	26.873	29.141
15	14.339	17.322	19.311	22.307	24.996	28.259	30.578
16	15.338	18.418	20.465	23.542	26.296	29.633	32.000
17	16.338	19.511	21.615	24.769	27.587	30.995	33.409
18	17.338	20.601	22.760	25.989	28.869	32.346	34.805
19	18.338	21.689	23.900	27.204	30.144	33.687	36.191
20	19.337	22.775	25.038	28.412	31.410	35.020	37.566
21	20.337	23.858	26.171	29.615	32.671	36.343	38.932
22	21.337	24.939	27.301	30.813	33.924	37.659	40.289
23	22.337	26.018	28.429	32.007	35.172	38.968	41.638
24	23.337	27.096	29.553	33.196	36.415	40.270	42.980
25	24.337	28.172	30.675	34.382	37.652	41.566	44.314
26	25.336	29.246	31.795	35.563	38.885	42.856	45.642
27	26.336	30.319	32.912	36.741	40.113	44.140	46.963
28	27.336	31.391	34.027	37.916	41.337	45.419	48.278
29	28.336	32.461	35.139	39.087	42.557	46.693	49.588
30	29.336	33.530	36.250	40.256	43.773	47.962	50.892

表 B. 常態曲線下面積及縱坐標：橫坐標由 .00 至 4.50

t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$	t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$	t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$
.00	.5000000	.3989423						
.01	.5039894	.3989223	.51	.6949743	.3502919	1.01	.8437524	.2395511
.02	.5079783	.3988625	.52	.6984682	.3484925	1.02	.8461358	.2371320
.03	.5119665	.3987628	.53	.7019440	.3466677	1.03	.8484950	.2347138
.04	.5159534	.3986233	.54	.7054015	.3448180	1.04	.8508300	.2322970
.05	.5199388	.3984439	.55	.7088403	.3429439	1.05	.8531409	.2298821
.06	.5239222	.3982248	.56	.7122603	.3410458	1.06	.8554277	.2274696
.07	.5279032	.3979661	.57	.7156612	.3391243	1.07	.8576903	.2250599
.08	.5318814	.3976677	.58	.7190427	.3371799	1.08	.8599289	.2226535
.09	.5358564	.3973298	.59	.7224047	.3352132	1.09	.8621434	.2202508
.10	.5398278	.3969525	.60	.7257469	.3332246	1.10	.8643339	.2178522
.11	.5437953	.3965360	.61	.7290691	.3312147	1.11	.8665005	.2154582
.12	.5477584	.3960802	.62	.7323711	.3291840	1.12	.8686431	.2130691
.13	.5517168	.3955854	.63	.7356527	.3271330	1.13	.8707619	.2106856
.14	.5556700	.3950517	.64	.7389137	.3250623	1.14	.8728568	.2083078
.15	.5596177	.3944793	.65	.7421539	.3229724	1.15	.8749281	.2059363
.16	.5635595	.3938684	.66	.7453731	.3208638	1.16	.8769756	.2035714
.17	.5674949	.3932190	.67	.7485711	.3187371	1.17	.8789995	.2012135
.18	.5714237	.3925315	.68	.7517478	.3165929	1.18	.8809999	.1988631
.19	.5753454	.3918060	.69	.7549029	.3144317	1.19	.8829768	.1965205
.20	.5792597	.3910427	.70	.7580363	.3122539	1.20	.8849303	.1941861
.21	.5831662	.3902419	.71	.7611479	.3100603	1.21	.8868606	.1918602
.22	.5870644	.3894038	.72	.7642375	.3078513	1.22	.8887676	.1895432
.23	.5909541	.3885286	.73	.7673049	.3056274	1.23	.8906514	.1872354
.24	.5948349	.3876166	.74	.7703500	.3033893	1.24	.8925123	.1849373
.25	.5987063	.3866681	.75	.7733726	.3011374	1.25	.8943502	.1826491
.26	.6025681	.3856834	.76	.7763727	.2988724	1.26	.8961653	.1803712
.27	.6064199	.3846627	.77	.7793501	.2965948	1.27	.8979577	.1781038
.28	.6102612	.3836063	.78	.7823046	.2943050	1.28	.8997274	.1758474
.29	.6140919	.3825146	.79	.7852361	.2920038	1.29	.9014747	.1736022
.30	.6179114	.3813878	.80	.7881446	.2896916	1.30	.9031995	.1713686
.31	.6217195	.3802264	.81	.7910299	.2873689	1.31	.9049021	.1691468
.32	.6255158	.3790305	.82	.7938919	.2850364	1.32	.9065825	.1669370
.33	.6293000	.3778007	.83	.7967306	.2826945	1.33	.9082409	.1647397
.34	.6330717	.3765372	.84	.7995458	.2803438	1.34	.9098773	.1625551
.35	.6368307	.3752403	.85	.8023375	.2779849	1.35	.9114920	.1603833
.36	.6405764	.3739106	.86	.8051055	.2756182	1.36	.9130850	.1582248
.37	.6443088	.3725483	.87	.8078498	.2732444	1.37	.9146565	.1560797
.38	.6480273	.3711539	.88	.8105703	.2708640	1.38	.9162067	.1539483
.39	.6517317	.3697277	.89	.8132671	.2684774	1.39	.9177356	.1518308
.40	.6554217	.3682701	.90	.8159399	.2660852	1.40	.9192433	.1497275
.41	.6590970	.3667817	.91	.818[illegible]887	.2636880	1.41	.9207302	.1476385
.42	.6627573	.3652627	.92	.8216136	.2612863	1.42	.9221962	.145[illegible]41
.43	.6664022	.3637136	.93	.8238145	.2588805	1.43	.9236415	.143[illegible]046
.44	.6700314	.3621349	.94	.8263912	.2564713	1.44	.9250663	.141460[illegible]
.45	.6736448	.3605270	.95	.8289439	.2540591	1.45	.9264707	.1394306
.46	.6772419	.3588903	.96	.8314724	.2516443	1.46	.9278550	.1374165
.47	.6808225	.3572253	.97	.8339768	.2492277	1.47	.9292191	.1354181
.48	.6843863	.3555325	.98	.8364569	.2468095	1.48	.9305634	.1334353
.49	.6879331	.3538124	.99	.8389129	.2443904	1.49	.9318879	.131[illegible]684
.50	.6914625	.3520653	1.00	.8413447	.2419707	1.50	.9331928	.129[illegible]176

表 *B*. 常態曲線下面積及縱坐標：橫坐標由 .00 至 4.50 (續前)

t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$	t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$	t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$
1.51	.9344783	.1275830	2.01	.9777844	.0529192	2.51	.9939634	.0170947
1.52	.9357445	.1256646	2.02	.9783083	.0518636	2.52	.9941323	.0166701
1.53	.9369916	.1237628	2.03	.9788217	.0508239	2.53	.9942969	.0162545
1.54	.9382198	.1218775	2.04	.9793248	.0498001	2.54	.9944574	.0158476
1.55	.9394292	.1200090	2.05	.9798178	.0487920	2.55	.9946139	.0154493
1.56	.9406201	.1181573	2.06	.9803007	.0477996	2.56	.9947664	.0150596
1.57	.9417924	.1163225	2.07	.9807738	.0468226	2.57	.9949151	.0146782
1.58	.9429466	.1145048	2.08	.9812372	.0458611	2.58	.9950600	.0143051
1.59	.9440826	.1127042	2.09	.9816911	.0449148	2.59	.9952012	.0139401
1.60	.9452007	.1109208	2.10	.9821356	.0439836	2.60	.9953388	.0135830
1.61	.9463011	.1091548	2.11	.9825708	.0430674	2.61	.9954729	.0132337
1.62	.9473839	.1074061	2.12	.9829970	.0421661	2.62	.9956035	.0128921
1.63	.9484493	.1056748	2.13	.9834142	.0412795	2.63	.9957308	.0125581
1.64	.9494974	.1039611	2.14	.9838226	.0404076	2.64	.9958547	.0122315
1.65	.9505285	.1022649	2.15	.9842224	.0395500	2.65	.9959754	.0119122
1.66	.9515428	.1005864	2.16	.9846137	.0387069	2.66	.9960930	.0116001
1.67	.9525403	.0989255	2.17	.9849966	.0378779	2.67	.9962074	.0112951
1.68	.9535213	.0972823	2.18	.9853713	.0370629	2.68	.9963189	.0109969
1.69	.9544860	.0956568	2.19	.9857379	.0362619	2.69	.9964274	.0107056
1.70	.9554345	.0940491	2.20	.9860966	.0354746	2.70	.9965330	.0104209
1.71	.9563671	.0924591	2.21	.9864474	.0317009	2.71	.9966358	.0101428
1.72	.9572838	.0908870	2.22	.9867906	.0339408	2.72	.9967359	.0098712
1.73	.9581849	.0893326	2.23	.9871263	.0331939	2.73	.9968333	.0096058
1.74	.9590705	.0877961	2.24	.9874545	.0324603	2.74	.9969280	.0093466
1.75	.9599408	.0862773	2.25	.9877755	.0317397	2.75	.9970202	.0090936
1.76	.9607961	.0847764	2.26	.9880894	.0310319	2.76	.9971099	.0088465
1.77	.9616364	.0832932	2.27	.9883962	.0303370	2.77	.9971972	.0086052
1.78	.9624620	.0818278	2.28	.9886962	.0296546	2.78	.9972821	.0083697
1.79	.9632730	.0803801	2.29	.9889893	.0289847	2.79	.9973646	.0081398
1.80	.9640697	.0789502	2.30	.9892759	.0283270	2.80	.9974449	.0079155
1.81	.9648521	.0775379	2.31	.9895559	.0276816	2.81	.9975229	.0076965
1.82	.9656205	.0761433	2.32	.9898296	.0270481	2.82	.9975988	.0074829
1.83	.9663750	.0747663	2.33	.9500969	.0264265	2.83	.9976726	.0072744
1.84	.9671159	.0734068	2.34	.9903581	.0258165	2.84	.9977443	.0070711
1.85	.9678432	.0720649	2.35	.9906133	.0252182	2.85	.9978140	.0068728
1.86	.9685572	.0707404	2.36	.9908625	.0246313	2.86	.9978818	.0066793
1.87	.9692581	.0694333	2.37	.9911060	.0240556	2.87	.9979476	.0064907
1.88	.9699460	.0681436	2.38	.9913437	.0234910	2.88	.9980116	.0063067
1.89	.9706210	.0668711	2.39	.9915758	.0229374	2.89	.9980738	.0061274
1.90	.9712834	.0656158	2.40	.9918025	.0223945	2.90	.9981342	.0059525
1.91	.9719334	.0643777	2.41	.9920237	.0218624	2.91	.9981929	.0057821
1.92	.9725711	.0631566	2.42	.9922397	.0213407	2.92	.9982498	.0056160
1.93	.9731966	.0619524	2.43	.9924506	.0208294	2.93	.9983052	.0054541
1.94	.9738102	.0607652	2.44	.9926564	.0203284	2.94	.9983589	.0052963
1.95	.9744119	.0595947	2.45	.9928572	.0198374	2.95	.9984111	.0051426
1.96	.9750021	.0584409	2.46	.9930531	.0193563	2.96	.9984618	.0049929
1.97	.9755808	.0573038	2.47	.9932443	.0188850	2.97	.9985110	.0048470
1.98	.9761482	.0761831	2.48	.9934309	.0184233	2.98	.9985586	.0047050
1.99	.9767045	.0550789	2.49	.9936128	.0179711	2.99	.9986071	.0045666
2.00	.9772499	.0539910	2.50	.9937903	.0175283	3.00	.9986501	.0044318

表 B. 常態曲線下之面積及縱坐標：横坐標由 .00 至 4.50 (續完)

t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$	t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$	t	$\int_{-\infty}^{t}\phi(t)dt$	$\phi(t)$
3.01	.9986938	.0043007	3.51	.9997759	.0008426	4.01	.9999696	.0001286
3.02	.9987361	.0041729	3.52	.9997842	.0008135	4.02	.9999709	.0001235
3.03	.9987772	.0040486	3.53	.9997922	.0007853	4.03	.9999721	.0001186
3.04	.9988171	.0039276	3.54	.9997999	.0007581	4.04	.9999733	.0001140
3.05	.9988558	.0038098	3.55	.9998074	.0007317	4.05	.9999744	.0001094
3.06	.9988933	.0036951	3.56	.9998146	.0007061	4.06	.9999755	.0001051
3.07	.9989297	.0035836	3.57	.9998215	.0006814	4.07	.9999765	.0001009
3.08	.9989650	.0034751	3.58	.9998282	.0006575	4.08	.9999775	.0000969
3.09	.9989992	.0033695	3.59	.9998347	.0006343	4.09	.9999784	.0000930
3.10	.9990324	.0032668	3.60	.9998409	.0006119	4.10	.9999793	.0000893
3.11	.9990646	.0031669	3.61	.9998469	.0005902	4.11	.9999802	.0000857
3.12	.9990957	.0030698	3.62	.9998527	.0005693	4.12	.9999811	.0000822
3.13	.9991260	.0029754	3.63	.9998583	.0005490	4.13	.9999819	.0000789
3.14	.9991553	.0028835	3.64	.9998637	.0005294	4.14	.9999826	.0000757
3.15	.9991836	.0027943	3.65	.9998689	.0005105	4.15	.9999834	.0000725
3.16	.9992112	.0027075	3.66	.9998739	.0004921	4.16	.9999841	.0000697
3.17	.9992378	.0026231	3.67	.9998787	.0004744	4.17	.9999848	.0000668
3.18	.9992636	.0025412	3.68	.9998834	.0004573	4.18	.9999854	.0000641
3.19	.9992886	.0024615	3.69	.9998879	.0004408	4.19	.9999861	.0000615
3.20	.9993129	.0023841	3.70	.9998922	.0004248	4.20	.9999867	.0000589
3.21	.9993363	.0023089	3.71	.9998964	.0004093	4.21	.9999872	.0000565
3.22	.9993590	.0022358	3.72	.9999004	.0003944	4.22	.9999878	.0000542
3.23	.9993810	.0021649	3.73	.9999043	.0003800	4.23	.9999883	.0000519
3.24	.9994024	.0020960	3.74	.9999080	.0003661	4.24	.9999888	.0000498
3.25	.9994230	.0020290	3.75	.9999116	.0003526	4.25	.9999893	.0000477
3.26	.9994429	.0019641	3.76	.9999150	.0003396	4.26	.9999898	.0000457
3.27	.9994623	.0019010	3.77	.9999184	.0003271	4.27	.9999902	.0000438
3.28	.9994810	.0018397	3.78	.9999216	.0003149	4.28	.9999907	.0000420
3.29	.9994991	.0017803	3.79	.9999247	.0003032	4.29	.9999911	.0000402
3.30	.9995166	.0017226	3.80	.9999277	.0002919	4.30	.9999915	.0000385
3.31	.9995335	.0016666	3.81	.9999305	.0002810	4.31	.9999918	.0000369
3.32	.9995499	.0016122	3.82	.9999333	.0002705	4.32	.9999922	.0000354
3.33	.9995658	.0015595	3.83	.9999359	.0002604	4.33	.9999925	.0000339
3.34	.9995811	.0015084	3.84	.9999385	.0002506	4.34	.9999929	.0000324
3.35	.9995959	.0014587	3.85	.9999409	.0002411	4.35	.9999932	.0000310
3.36	.9996103	.0014106	3.86	.9999433	.0002320	4.36	.9999935	.0000297
3.37	.9996242	.0013639	3.87	.9999456	.0002232	4.37	.9999938	.0000284
3.38	.9996376	.0013187	3.88	.9999478	.0002147	4.38	.9999941	.0000272
3.39	.9996505	.0012748	3.89	.9999499	.0002065	4.39	.9999943	.0000261
3.40	.9996631	.0012322	3.90	.9999519	.0001987	4.40	.9999946	.0000249
3.41	.9996752	.0011910	3.91	.9999539	.0001910	4.41	.9999948	.0000239
3.42	.9996869	.0011510	3.92	.9999557	.0001837	4.42	.9999951	.0000228
3.43	.9996982	.0011122	3.93	.9999575	.0001766	4.43	.9999973	.0000218
3.44	.9997091	.0010747	3.94	.9999593	.0001698	4.44	.9999955	.0000209
3.45	.9997197	.0010383	3.95	.9999609	.0001633	4.45	.9999957	.0000200
3.46	.9997299	.0010030	3.96	.9999625	.0001569	4.46	.9999959	.0000191
3.47	.9997398	.0009689	3.97	.9999641	.0001508	4.47	.9999961	.0000183
3.48	.9997493	.0009358	3.98	.9999655	.0001449	4.48	.9999963	.0000175
3.49	.9997585	.0009037	3.99	.9999670	.0001393	4.49	.9999964	.0000167
3.50	.9997674	.0008727	4.00	.9999683	.0001338	4.50	.9999966	.0000160

表 C. X^2 及 P 之值：自由度 = 1, t 由 .00 至 4.49

t	X^2	P	t	X^2	P	t	X^2	P
0.00	0.0000	1.0000	0.50	0.2500	0.6171	1.00	1.0000	0.3173
0.01	0.0001	0.9920	0.51	0.2601	0.6101	1.01	1.0201	0.3125
0.02	0.0004	0.9840	0.52	0.2704	0.6031	1.02	1.0404	0.3077
0.03	0.0009	0.9760	0.53	0.2809	0.5961	1.03	1.0609	0.3030
0.04	0.0016	0.9681	0.54	0.2916	0.5892	1.04	1.0816	0.2983
0.05	0.0025	0.9601	0.55	0.3025	0.5823	1.05	1.1025	0.2937
0.06	0.0036	0.9522	0.56	0.3136	0.5755	1.06	1.1236	0.2891
0.07	0.0049	0.9442	0.57	0.3249	0.5687	1.07	1.1449	0.2846
0.08	0.0064	0.9362	0.58	0.3364	0.5619	1.08	1.1664	0.2801
0.09	0.0081	0.9283	0.59	0.3841	0.5552	1.09	1.1881	0.2757
0.10	0.0100	0.9203	0.60	0.3600	0.5485	1.10	1.2100	0.2713
0.11	0.0121	0.9124	0.61	0.3721	0.5419	1.11	1.2321	0.2670
0.12	0.0144	0.9045	0.62	0.3844	0.5353	1.12	1.2544	0.2627
0.13	0.0169	0.8966	0.63	0.3969	0.5287	1.13	1.2769	0.2585
0.14	0.0196	0.8887	0.64	0.4096	0.5222	1.14	1.2996	0.2543
0.15	0.0225	0.8808	0.65	0.4225	0.5157	1.15	1.3225	0.2501
0.16	0.0256	0.8729	0.66	0.4356	0.5093	1.16	1.3456	0.2460
0.17	0.0289	0.8650	0.67	0.4489	0.5029	1.17	1.3689	0.2420
0.18	0.0324	0.8572	0.68	0.4624	0.4965	1.18	1.3924	0.2380
0.19	0.0361	0.8493	0.69	0.4761	0.4902	1.19	1.4161	0.2340
0.20	0.0400	0.8415	0.70	0.4900	0.4839	1.20	1.4400	0.2301
0.21	0.0441	0.8337	0.71	0.5041	0.4777	1.21	1.4641	0.2263
0.22	0.0484	0.8259	0.72	0.5184	0.4715	1.22	1.4884	0.2225
0.23	0.0529	0.8181	0.73	0.5329	0.4654	1.23	1.5129	0.2187
0.24	0.0576	0.8103	0.74	0.5476	0.4593	1.24	1.5376	0.2150
0.25	0.0625	0.8026	0.75	0.5625	0.4533	1.25	1.5625	0.2113
0.26	0.0676	0.7949	0.76	0.5776	0.4473	1.26	1.5876	0.2077
0.27	0.0729	0.7872	0.77	0.5929	0.4413	1.27	1.6129	0.2041
0.28	0.0784	0.7795	0.78	0.6084	0.4354	1.28	1.6384	0.2005
0.29	0.0841	0.7718	0.79	0.6241	0.4295	1.29	1.6641	0.1971
0.30	0.0900	0.7642	0.80	0.6400	0.4237	1.30	1.6900	0.1936
0.31	0.0961	0.7566	0.81	0.6561	0.4179	1.31	1.7161	0.1902
0.32	0.1024	0.7490	0.82	0.6724	0.4122	1.32	1.7424	0.1868
0.33	0.1089	0.7414	0.83	0.6889	0.4065	1.33	1.7689	0.1835
0.34	0.1156	0.7339	0.84	0.7056	0.4009	1.34	1.7956	0.1802
0.35	0.1225	0.7263	0.85	0.7225	0.3953	1.35	1.8225	0.1770
0.36	0.1296	0.7188	0.86	0.7396	0.3898	1.36	1.8496	0.1738
0.37	0.1369	0.7114	0.87	0.7569	0.3843	1.37	1.8769	0.1707
0.38	0.1444	0.7039	0.88	0.7744	0.3789	1.38	1.9044	0.1676
0.39	0.1521	0.6965	0.89	0.7921	0.3735	1.39	1.9321	0.1645
0.40	0.1600	0.6892	0.90	0.8100	0.3681	1.40	1.9600	0.1615
0.41	0.1681	0.6818	0.91	0.8281	0.3628	1.41	1.9881	0.1585
0.42	0.1764	0.6745	0.92	0.8464	0.3576	1.42	2.0164	0.1556
0.43	0.1849	0.6672	0.93	0.8649	0.3524	1.43	2.0449	0.1527
0.44	0.1936	0.6599	0.94	0.8836	0.3472	1.44	2.0736	0.1499
0.45	0.2025	0.6527	0.95	0.9025	0.3421	1.45	2.1025	0.1471
0.46	0.2116	0.6455	0.96	0.9216	0.3371	1.46	2.1316	0.1443
0.47	0.2209	0.6384	0.97	0.9409	0.3320	1.47	2.1609	0.1416
0.48	0.2304	0.6312	0.98	0.9604	0.3271	1.48	2.1904	0.1389
0.49	0.2401	0.6241	0.99	0.9801	0.3222	1.49	2.2201	0.1362

表 C. X^2 及 P 之值: 自由度=1, t 由 .00 至 4.49 (續前)

t	X^2	P	t	X^2	P	t	X^2	P
1.50	2.2500	0.1336	2.00	4.0000	0.0455	2.50	6.2500	0.0124
1.51	2.2801	0.1310	2.01	4.0401	0.0444	2.51	6.3001	0.0121
1.52	2.3104	0.1285	2.02	4.0804	0.0434	2.52	6.3504	0.0117
1.53	2.3409	0.1260	2.03	4.1209	0.0424	2.53	6.4009	0.0114
1.54	2.3716	0.1236	2.04	4.1616	0.0414	2.54	6.4516	0.0111
1.55	2.4025	0.1211	2.05	4.2025	0.0404	2.55	6.5025	0.0108
1.56	2.4336	0.1188	2.06	4.2436	0.0394	2.56	6.5536	0.0105
1.57	2.4649	0.1164	2.07	4.2849	0.0385	2.57	6.6049	0.0102
1.58	2.4964	0.1141	2.08	4.3264	0.0375	2.58	6.6564	0.0099
1.59	2.5281	0.1118	2.09	4.3681	0.0366	2.59	6.7081	0.0096
1.60	2.5600	0.1096	2.10	4.4100	0.0357	2.60	6.7600	0.0093
1.61	2.5921	0.1074	2.11	4.4521	0.0349	2.61	6.8121	0.0091
1.62	2.6244	0.1052	2.12	4.4944	0.0340	2.62	6.8644	0.0088
1.63	2.6569	0.1031	2.13	4.5369	0.0332	2.63	6.9169	0.0085
1.64	2.6896	0.1010	2.14	4.5796	0.0324	2.64	6.9696	0.0083
1.65	2.7225	0.0989	2.15	4.6225	0.0316	2.65	7.0225	0.0080
1.66	2.7556	0.0969	2.16	4.6656	0.0308	2.66	7.0756	0.0078
1.67	2.7889	0.0949	2.17	4.7089	0.0300	2.67	7.1289	0.0076
1.68	2.8224	0.0930	2.18	4.7524	0.0293	2.68	7.1824	0.0074
1.69	2.8561	0.0910	2.19	4.7961	0.0285	2.69	7.2361	0.0071
1.70	2.8900	0.0891	2.20	4.8400	0.0278	2.70	7.2900	0.0069
1.71	2.9241	0.0873	2.21	4.8841	0.0271	2.71	7.3441	0.0067
1.72	2.9584	0.0854	2.22	4.9284	0.0264	2.72	7.3984	0.0065
1.73	2.9929	0.0836	2.23	4.9729	0.0257	2.73	7.4529	0.0063
1.74	3.0276	0.0819	2.24	5.0176	0.0251	2.74	7.5076	0.0061
1.75	3.0625	0.0801	2.25	5.0625	0.0244	2.75	7.5625	0.0060
1.76	3.0976	0.0784	2.26	5.1076	0.0238	2.76	7.6176	0.0058
1.77	3.1329	0.0767	2.27	5.1529	0.0232	2.77	7.6729	0.0056
1.78	3.1684	0.0751	2.28	5.1984	0.0226	2.78	7.7284	0.0054
1.79	3.2041	0.0735	2.29	5.2441	0.0220	2.79	7.7841	0.0053
1.80	3.2400	0.0719	2.30	5.2900	0.0214	2.80	7.8400	0.0051
1.81	3.2761	0.0703	2.31	5.3361	0.0209	2.81	7.8961	0.0050
1.82	3.3124	0.0688	2.32	5.3824	0.0203	2.82	7.9524	0.0048
1.83	3.3489	0.0672	2.33	5.4289	0.0198	2.83	8.0089	0.0047
1.84	3.3856	0.0658	2.34	5.4756	0.0193	2.84	8.0656	0.0045
1.85	3.4225	0.0643	2.35	5.5225	0.0188	2.85	8.1225	0.0044
1.86	3.4596	0.0629	2.36	5.5696	0.0183	2.86	8.1796	0.0042
1.87	3.4969	0.0615	2.37	5.6169	0.0178	2.87	8.2369	0.0041
1.88	3.5344	0.0601	2.38	5.6644	0.0173	2.88	8.2944	0.0040
1.89	3.5721	0.0588	2.39	5.7121	0.0168	2.89	8.3521	0.0039
1.90	3.6100	0.0574	2.40	5.7600	0.0164	2.90	8.4100	0.0037
1.91	3.6481	0.0561	2.41	5.8081	0.0160	2.91	8.4681	0.0036
1.92	3.6864	0.0549	2.42	5.8564	0.0155	2.92	8.5264	0.0035
1.93	3.7249	0.0536	2.43	5.9049	0.0151	2.93	8.5849	0.0034
1.94	3.7636	0.0524	2.44	5.9536	0.0147	2.94	8.6436	0.0033
1.95	3.8025	0.0512	2.45	6.0025	0.0143	2.95	8.7025	0.0032
1.96	3.8416	0.0500	2.46	6.0516	0.0139	2.96	8.7616	0.0031
1.97	3.8809	0.0488	2.47	6.1009	0.0135	2.97	8.8209	0.0030
1.98	3.9204	0.0477	2.48	6.1504	0.0131	2.98	8.8804	0.0029
1.99	3.9601	0.0466	2.49	6.2001	0.0128	2.99	8.9401	0.0028

表 C. X^2 及 P 之値: 自由度=1, t 由 .00 至 4.49 (續完)

t	X^2	P	t	X^2	P	t	X^2	P
3.00	9.0000	0.0027	3.50	12.2500	0.0005	4.00	16.0000	0.0000
3.01	9.0601	0.0026	3.51	12.3201	0.0004	4.01	16.0801	0.0000
3.02	9.1204	0.0025	3.52	12.3904	0.0004	4.02	16.1604	0.0000
3.03	9.1809	0.0024	3.53	12.4609	0.0004	4.03	16.2409	0.0000
3.04	9.2416	0.0024	3.54	12.5316	0.0004	4.04	16.3216	0.0000
3.05	9.3025	0.0023	3.55	12.6025	0.0004	4.05	16.4025	0.0000
3.06	9.3636	0.0022	3.56	12.6736	0.0004	4.06	16.4836	0.0000
3.07	9.4249	0.0021	3.57	12.7449	0.0004	4.07	16.5649	0.0000
3.08	9.4864	0.0021	3.58	12.8164	0.0003	4.08	16.6464	0.0000
3.09	9.5481	0.0020	3.59	12.8881	0.0003	4.09	16.7281	0.0000
3.10	9.6100	0.0019	3.60	12.9600	0.0003	4.10	16.8100	0.0000
3.11	9.6721	0.0019	3.61	13.0321	0.0003	4.11	16.8921	0.0000
3.12	9.7344	0.0018	3.62	13.1044	0.0003	4.12	16.9744	0.0000
3.13	9.7969	0.0017	3.63	13.1769	0.0003	4.13	17.0569	0.0000
3.14	9.8596	0.0017	3.64	13.2496	0.0003	4.14	17.1396	0.0000
3.15	9.9225	0.0016	3.65	13.3225	0.0003	4.15	17.2225	0.0000
3.16	9.9856	0.0016	3.66	13.3956	0.0003	4.16	17.3056	0.0000
3.17	10.0489	0.0015	3.67	13.4689	0.0002	4.17	17.3889	0.0000
3.18	10.1124	0.0015	3.68	13.5424	0.0002	4.18	17.4724	0.0000
3.19	10.1761	0.0014	3.69	13.6161	0.0002	4.19	17.5561	0.0000
3.20	10.2400	0.0014	3.70	13.6900	0.0002	4.20	17.6400	0.0000
3.21	10.3041	0.0013	3.71	13.7641	0.0002	4.21	17.7241	0.0000
3.22	10.3684	0.0013	3.72	13.8384	0.0002	4.22	17.8084	0.0000
3.23	10.4329	0.0012	3.73	13.9129	0.0002	4.23	17.8929	0.0000
3.24	10.4976	0.0012	3.74	13.9876	0.0002	4.24	17.9776	0.0000
3.25	10.5625	0.0012	3.75	14.0625	0.0002	4.25	18.0625	0.0000
3.26	10.6276	0.0011	3.76	14.1376	0.0002	4.26	18.1476	0.0000
3.27	10.6929	0.0011	3.77	14.2129	0.0002	4.27	18.2329	0.0000
3.28	10.7584	0.0010	3.78	14.2884	0.0002	4.28	18.3184	0.0000
3.29	10.8241	0.0010	3.79	14.3641	0.0002	4.29	18.4041	0.0000
3.30	10.8900	0.0010	3.80	14.4400	0.0001	4.30	18.4900	0.0000
3.31	10.9561	0.0009	3.81	14.5161	0.0001	4.31	18.5761	0.0000
3.32	11.0224	0.0009	3.82	14.5924	0.0001	4.32	18.6624	0.0000
3.33	11.0889	0.0009	3.83	14.6689	0.0001	4.33	18.7489	0.0000
3.34	11.1556	0.0008	3.84	14.7456	0.0001	4.34	18.8356	0.0000
3.35	11.2225	0.0008	3.85	14.8225	0.0001	4.35	18.9225	0.0000
3.36	11.2896	0.0008	3.86	14.8996	0.0001	4.36	19.0096	0.0000
3.37	11.3569	0.0008	3.87	14.9769	0.0001	4.37	19.0969	0.0000
3.38	11.4244	0.0007	3.88	15.0544	0.0001	4.38	19.1844	0.0000
3.39	11.4921	0.0007	3.89	15.1321	0.0001	4.39	19.2721	0.0000
3.40	11.5600	0.0007	3.90	15.2100	0.0000	4.40	19.3600	0.0000
3.41	11.6281	0.0006	3.91	15.2881	0.0000	4.41	19.4481	0.0000
3.42	11.6964	0.0006	3.92	15.3664	0.0000	4.42	19.5364	0.0000
3.43	11.7649	0.0006	3.93	15.4449	0.0000	4.43	19.6249	0.0000
3.44	11.8336	0.0006	3.94	15.5236	0.0000	4.44	19.7136	0.0000
3.45	11.9025	0.0006	3.95	15.6025	0.0000	4.45	19.8025	0.0000
3.46	11.9716	0.0005	3.96	15.6816	0.0000	4.46	19.8916	0.0000
3.47	12.0409	0.0005	3.97	15.7609	0.0000	4.47	19.9809	0.0000
3.48	12.1104	0.0005	3.98	15.8404	0.0000	4.48	20.0704	0.0000
3.49	12.1801	0.0005	3.99	15.9201	0.0000	4.49	20.1601	0.0000

表 D. 常態曲線之橫坐標：面積由 .000 至 .499

面積	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
.00	∞	3.0902	2.8782	2.7478	2.6521	2.5758	2.5121	2.4573	2.4089	2.3656
.01	2.3263	2.2904	2.2571	2.2262	2.1973	2.1701	2.1444	2.1201	2.0969	2.0749
.02	2.0537	2.0335	2.0141	1.9954	1.9774	1.9600	1.9431	1.9268	1.9110	1.8957
.03	1.8808	1.8663	1.8522	1.8384	1.8250	1.8119	1.7991	1.7866	1.7744	1.7624
.04	1.7507	1.7392	1.7279	1.7169	1.7060	1.6954	1.6849	1.6747	1.6646	1.6546
.05	1.6449	1.6352	1.6258	1.6164	1.6072	1.5982	1.5893	1.5805	1.5718	1.5632
.06	1.5548	1.5464	1.5382	1.5301	1.5220	1.5141	1.5063	1.4985	1.4909	1.4833
.07	1.4758	1.4684	1.4611	1.4538	1.4466	1.4395	1.4325	1.4255	1.4187	1.4118
.08	1.4051	1.3984	1.3917	1.3852	1.3787	1.3722	1.3658	1.3595	1.3532	1.3469
.09	1.3408	1.3346	1.3285	1.3225	1.3165	1.3106	1.3047	1.2988	1.2930	1.2873
.10	1.2816	1.2759	1.2702	1.2646	1.2591	1.2536	1.2481	1.2426	1.2372	1.2319
.11	1.2265	1.2212	1.2160	1.2107	1.2055	1.2004	1.1952	1.1901	1.1850	1.1800
.12	1.1750	1.1700	1.1650	1.1601	1.1552	1.1503	1.1455	1.1407	1.1359	1.1311
.13	1.1264	1.1217	1.1170	1.1123	1.1077	1.1031	1.0985	1.0939	1.0893	1.0848
.14	1.0803	1.0758	1.0714	1.0669	1.0625	1.0581	1.0537	1.0494	1.0450	1.0407
.15	1.0364	1.0322	1.0279	1.0237	1.0194	1.0152	1.0110	1.0069	1.0027	0.9986
.16	0.9945	0.9904	0.9863	0.9822	0.9782	0.9741	0.9701	0.9661	0.9621	0.9581
.17	0.9542	0.9502	0.9463	0.9424	0.9385	0.9346	0.9307	0.9269	0.9230	0.9192
.18	0.9154	0.9116	0.9078	0.9040	0.9002	0.8965	0.8927	0.8890	0.8853	0.8816
.19	0.8779	0.8742	0.8705	0.8669	0.8633	0.8596	0.8560	0.8524	0.8488	0.8452
.20	0.8416	0.8381	0.8345	0.8310	0.8274	0.8239	0.8204	0.8169	0.8134	0.8099
.21	0.8064	0.8030	0.7995	0.7961	0.7926	0.7892	0.7858	0.7824	0.7790	0.7756
.22	0.7722	0.7688	0.7655	0.7621	0.7588	0.7554	0.7521	0.7488	0.7454	0.7421
.23	0.7388	0.7356	0.7323	0.7290	0.7257	0.7225	0.7192	0.7160	0.7128	0.7095
.24	0.7063	0.7031	0.6999	0.6967	0.6935	0.6903	0.6871	0.6840	0.6808	0.6776
.25	0.6745	0.6713	0.6682	0.6651	0.6620	0.6588	0.6557	0.6526	0.6495	0.6464
.26	0.6433	0.6403	0.6372	0.6341	0.6311	0.6280	0.6250	0.6219	0.6189	0.6158
.27	0.6128	0.6098	0.6068	0.6038	0.6008	0.5978	0.5948	0.5918	0.5888	0.5858
.28	0.5828	0.5799	0.5769	0.5740	0.5710	0.5681	0.5651	0.5622	0.5592	0.5563
.29	0.5534	0.5505	0.5476	0.5446	0.5417	0.5388	0.5359	0.5330	0.5302	0.5273
.30	0.5244	0.5215	0.5187	0.5158	0.5129	0.5101	0.5072	0.5044	0.5015	0.4987
.31	0.4959	0.4930	0.4902	0.4874	0.4845	0.4817	0.4789	0.4761	0.4733	0.4705
.32	0.4677	0.4649	0.4621	0.4593	0.4565	0.4538	0.4510	0.4482	0.4454	0.4427
.33	0.4399	0.4372	0.4344	0.4316	0.4289	0.4261	0.4234	0.4207	0.4179	0.4152
.34	0.4125	0.4097	0.4070	0.4043	0.4016	0.3989	0.3961	0.3934	0.3907	0.3880
.35	0.3853	0.3826	0.3799	0.3772	0.3745	0.3719	0.3692	0.3665	0.3638	0.3611
.36	0.3585	0.3558	0.3531	0.3505	0.3478	0.3451	0.3425	0.3398	0.3372	0.3345
.37	0.3319	0.3292	0.3266	0.3239	0.3213	0.3186	0.3160	0.3134	0.3107	0.3081
.38	0.3055	0.3029	0.3002	0.2976	0.2950	0.2924	0.2898	0.2871	0.2845	0.2819
.39	0.2793	0.2767	0.2741	0.2715	0.2689	0.2663	0.2637	0.2611	0.2585	0.2559
.40	0.2533	0.2508	0.2482	0.2456	0.2430	0.2404	0.2378	0.2353	0.2327	0.2301
.41	0.2275	0.2250	0.2224	0.2198	0.2173	0.2147	0.2121	0.2096	0.2070	0.2045
.42	0.2019	0.1993	0.1968	0.1942	0.1917	0.1891	0.1866	0.1840	0.1815	0.1789
.43	0.1764	0.1738	0.1713	0.1687	0.1662	0.1637	0.1611	0.1586	0.1560	0.1535
.44	0.1510	0.1484	0.1459	0.1434	0.1408	0.1383	0.1358	0.1332	0.1307	0.1282
.45	0.1257	0.1231	0.1206	0.1181	0.1156	0.1130	0.1105	0.1080	0.1055	0.1030
.46	0.1004	0.0979	0.0954	0.0929	0.0904	0.0878	0.0853	0.0828	0.0803	0.0778
.47	0.0753	0.0728	0.0702	0.0677	0.0652	0.0627	0.0602	0.0577	0.0552	0.0527
.48	0.0502	0.0476	0.0451	0.0426	0.0401	0.0376	0.0351	0.0326	0.0301	0.0276
.49	0.0251	0.0226	0.0201	0.0175	0.0150	0.0125	0.0100	0.0075	0.0050	0.0025

表 E. $\log \Gamma(n)$ 之值：n 由 1.00 至 1.99

n	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
1.00	---	9750	9 00	9251	9003	8755	8509	8263	8017	7773
1.01	$\bar{1}$.997529	7285	7043	6801	6560	6320	6030	5841	5602	5365
1.02	5128	4892	4656	4421	4187	3953	3721	3489	3257	3026
1.03	2796	2567	2338	2110	1883	1656	1430	1205	0981	0757
1.04	0533	0311	0089	9868	$\bar{9}$647	$\bar{9}$427	$\bar{9}$208	$\bar{8}$989	$\bar{8}$772	$\bar{8}$554
1.05	$\bar{1}$.988335	8122	7907	7692	7478	7265	7053	6841	6629	6419
1.06	6205	6000	5791	5583	5376	5169	4963	4758	4553	4349
1.07	4145	3943	3741	3539	3338	3138	2939	2740	2541	2344
1.08	2147	1951	1755	1560	1365	1172	0978	0786	0594	0403
1.09	0212	0022	$\bar{9}$833	$\bar{9}$644	$\bar{9}$456	$\bar{9}$269	$\bar{9}$082	$\bar{8}$896	$\bar{8}$710	$\bar{8}$525
1.10	$\bar{1}$.978341	8157	7974	7791	7610	7428	7248	7068	6888	6709
1.11	6531	6354	6177	6000	5825	5650	5475	5301	5128	4955
1.12	4783	4612	4441	4271	4101	3932	3764	3596	3429	3262
1.13	3096	2931	2766	2602	2438	2275	2113	1951	1790	1629
1.14	1469	1309	1150	0992	0835	0677	0521	0365	0210	0055
1.15	$\bar{1}$.969901	9747	9594	9442	9290	9139	8988	8838	8688	8539
1.16	8390	8243	8096	7949	7803	7658	7513	7369	7225	7082
1.17	6939	6797	6655	6514	6374	6234	6095	5957	5818	5681
1.18	5544	5408	5272	5137	5002	4868	4734	4601	4469	4337
1.19	4205	4075	3944	3815	3686	3557	3429	3302	3175	3048
1.20	2922	2797	2672	2548	2425	2302	2179	2057	1936	1815
1.21	1695	1575	1456	1337	1219	1101	0984	0867	0751	0636
1.22	0521	0407	0293	0180	0067	$\bar{9}$955	$\bar{9}$843	$\bar{9}$732	$\bar{9}$621	$\bar{9}$511
1.23	$\bar{1}$.959401	9292	9184	9076	8968	8861	8755	8649	8544	8439
1.24	8335	8231	8128	8025	7923	7821	7720	7620	7520	7420
1.25	7321	7223	7125	7027	6930	6834	6738	6642	6547	6453
1.26	6359	6267	6173	6081	5989	5898	5807	5716	5627	5537
1.27	5449	5360	5273	5185	5099	5013	4927	4842	4757	4673
1.28	4589	4506	4423	4341	4259	4178	4097	4017	3938	3858
1.29	3780	3702	3624	3547	3470	3394	3318	3243	3168	3094
1.30	3020	2947	2874	2802	2730	2659	2588	2518	2448	2379
1.31	2310	2242	2174	2106	2040	1973	1907	1842	1777	1712
1.32	1648	1585	1522	1459	1397	1336	1275	1214	1154	1094
1.33	1035	0977	0918	0861	0803	0747	0690	0634	0579	0524
1.34	0470	0416	0362	0309	0257	0205	0153	0102	0051	0001
1.35	$\bar{1}$.949951	9902	9853	9805	9757	9710	9663	9617	9571	9525
1.36	9480	9435	9391	9348	9304	9262	9219	9178	9136	9095
1.37	9054	9015	8975	8936	8898	8859	8822	8785	8748	8711
1.38	8676	8640	8605	8571	8537	8503	8470	8437	8405	8373
1.39	8342	8311	8280	8250	8221	8192	8163	8135	8107	8080
1.40	8053	8026	8000	7975	7950	7925	7901	7877	7854	7831
1.41	7808	7786	7765	7744	7723	7703	7683	7664	7645	7626
1.42	7608	7590	7573	7556	7540	7524	7509	7494	7479	7465
1.43	7451	7438	7425	7413	7401	7389	7378	7368	7358	7348
1.44	7338	7329	7321	7312	7305	7298	7291	7284	7278	7273
1.45	7268	7263	7259	7255	7251	7248	7246	7244	7242	7241
1.46	7240	7239	7239	7240	7241	7242	7243	7245	7248	7251
1.47	7254	7258	7262	7266	7271	7277	7282	7289	7295	7302
1.48	7310	7317	7326	7334	7343	7353	7363	7373	7384	7395
1.49	7407	7419	7431	7444	7457	7471	7485	7499	7514	7529

表 E. $\log\Gamma(n)$ 之值: n 由 1.00 至 1.99 (續完)

n	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
1.50	$\bar{1}$.947545	7561	7577	7594	7612	7629	7647	7666	7685	7704
1.51	7724	7744	7764	7785	7806	7828	7850	7873	7896	7919
1.52	7943	7967	7991	8016	8041	8067	8093	8120	8146	8174
1.53	8201	8229	8258	8287	8316	8346	8376	8406	8437	8468
1.54	8500	8532	8561	8597	8630	8664	8698	8732	8767	8802
1.55	8837	8873	8910	8946	8983	9021	9059	9097	9135	9174
1.56	9214	9254	9294	9334	9375	9417	9458	9500	9543	9586
1.57	9629	9672	9716	9761	9806	9851	9896	9942	9989	0035
1.58	$\bar{1}$.950082	0130	0177	0225	0274	0323	0372	0422	0472	0523
1.59	0573	0624	0676	0728	0780	0833	0886	0939	0993	1047
1.60	1102	1157	1212	1268	1324	1380	1437	1494	1552	1610
1.61	1668	1727	1786	1845	1905	1965	2025	2086	2147	2209
1.62	2271	2333	2396	2459	2522	2586	2650	2715	2780	2845
1.63	2911	2977	3043	3110	3177	3244	3312	3380	3449	3517
1.64	3587	3656	3726	3797	3867	3938	4010	4081	4154	4226
1.65	4299	4372	4446	4519	4594	4668	4743	4819	4894	4970
1.66	5047	5124	5201	5278	5356	5434	5513	5592	5671	5740
1.67	5830	5911	5991	6072	6154	6235	6317	6400	6182	6566
1.68	6649	6733	6817	6901	6986	7072	7157	7243	7322	7416
1.69	7503	7590	7678	7766	7854	7943	8032	8122	8211	8301
1.70	8391	8482	8573	8664	8756	8848	8941	9034	9127	9220
1.71	9314	9409	9502	9598	9693	9788	9884	9980	$\bar{0}$077	$\bar{0}$174
1.72	$\bar{1}$.960271	0369	0467	0565	0664	0763	0862	0961	1061	1162
1.73	1262	1363	1464	1565	1668	1770	1873	1976	2079	2183
1.74	2287	2391	2496	2601	2706	2812	2918	3024	3131	3238
1.75	3345	3453	3561	3669	3778	3887	3996	4015	4215	4326
1.76	4436	4547	4659	4770	4882	4994	5107	5220	5333	5447
1.77	5561	5675	5789	5904	6019	6135	6251	6367	6484	6600
1.78	6718	6835	6953	7071	7189	7308	7427	7547	7666	7787
1.79	7907	8028	8149	8270	8392	8514	8636	8759	8882	9005
1.80	9129	9253	9377	9501	9626	9751	9877	$\bar{0}$003	$\bar{0}$129	$\bar{0}$255
1.81	$\bar{1}$.970383	0509	0637	0765	0893	1021	1150	1279	1408	1538
1.82	1668	1798	1929	2060	2191	2322	2454	2586	2719	2852
1.83	2985	3118	3252	3386	3520	3655	3790	3925	4061	4197
1.84	4333	4470	4606	4744	4881	5019	5157	5295	5434	5573
1.85	5712	5852	5992	6132	6273	6414	6555	6697	6838	6980
1.86	7123	7266	7408	7552	7696	7840	7984	8128	8273	8419
1.87	8564	8710	8856	9002	9149	9296	9443	9591	9739	9887
1.88	$\bar{1}$.980036	0184	0333	0483	0633	0783	0933	1084	1234	1386
1.89	1537	1689	1841	1994	2147	2299	2453	2607	2761	2915
1.90	3069	3224	3379	3535	3690	3846	4003	4159	4316	4474
1.91	4631	4789	4947	5105	5264	5423	5582	5742	5902	6062
1.92	6223	6383	6544	6706	6867	7029	7192	7354	7517	7680
1.93	7844	8007	8171	8336	8500	8665	8830	8996	9161	9327
1.94	9494	9660	9827	9995	$\bar{0}$162	$\bar{0}$330	$\bar{0}$498	$\bar{0}$666	$\bar{0}$835	$\bar{1}$004
1.95	$\bar{1}$.991173	1343	1512	1683	1853	2024	2195	2366	2537	2709
1.96	2881	3054	3227	3399	3573	3746	3920	4094	4260	4443
1.97	4618	4794	4969	5145	5321	5498	5674	5851	6029	6206
1.98	6384	6562	6740	6919	7098	7277	7457	7637	7817	7997
1.99	8178	8359	8540	8722	8903	9085	9268	9450	9633	9816

表 F. $\log F(r, r_2)$ 及 $\log H(r, r_2)$ 之值：ϕ 由 0° 至 45°，$r=1, 2$

$\phi°$	$r=1$			$r=2$			
	$\log F(r, p_2)$	Δ	Δ^2	$\log F(r, p_2)$	$\log H(r, p_2)$	Δ	Δ^2
0	.3010300			.1961199	.1961199		
1	0609	309		2052	1391	192	
2	1538	923	619	4614	1966	575	383
3	3087	1550	621	8890	2924	958	383
4	5262	2175	625	.1974890	4264	1340	382
5	8067	2805	630	.1982627	5985	1722	381
6	.3021508	3441	636	.1992118	8087	2102	380
7	5594	4086	645	.2003385	.1970567	2480	379
8	.3030335	4740	654	.2016452	3424	2857	377
9	5741	5406	666	.2031349	6655	3231	375
10	.3041825	6085	679	.2048110	.1980260	3604	373
11	8603	6778	693	.2066774	4234	3974	370
12	.3056091	7488	710	.2087382	8574	4341	367
13	.3064307	8216	728	.2109935	.1993280	4705	364
14	.3073271	8964	749	.2134631	8344	5064	360
15	.3083006	9735	771	.2161383	.2003764	5420	356
16	.3093538	10532	796	.2190303	9537	5773	352
17	.3104892	11353	822	.2221462	.2015657	6120	347
18	.3117098	12206	853	.2254936	.2022120	6463	343
19	.3130189	13091	885	.2290807	8921	6801	338
20	.3144200	14011	919	.2329167	.2036054	7133	332
21	.3159169	14969	958	.2370114	.2043514	7460	327
22	.3175137	15968	999	.2413755	.2051294	7780	320
23	.3192150	17013	1045	.2460203	9387	8093	313
24	.3210256	18106	1093	.2509584	.2067787	8400	307
25	.3229507	19252	1146	.2562034	.2076487	8700	300
26	.3249963	20456	1204	.2617697	.2085478	8991	291
27	.3271685	21722	1266	.2676733	.2094753	9275	284
28	.3294740	23055	1334	.2739311	.2104302	9549	275
29	.3319202	24462	1407	.2805618	.2114118	9816	266
30	.3345150	25948	1486	.2875852	.2124190	10072	256
31	.3372672	27521	1573	.2950232	.2134509	10319	247
32	.3401860	29188	1667	.3028992	.2145064	10556	236
33	.3132818	30958	1769	.3112388	.2155846	10782	226
34	.3465656	32838	1881	.3200695	.2166842	10996	213
35	.3500496	34840	2002	.3294214	.2178041	11199	203
36	.3537469	36974	2134	.3393271	.2189431	11390	191
37	.3576722	39252	2279	.3498221	.2201000	11569	178
38	.3618410	41639	2436	.3609451	.2212734	11734	165
39	3662708	44298	2609	.3727382	.2224621	11886	152
40	3709805	47096	2799	3852475	.2236644	12023	138
41	.3759908	50103	3007	.3985232	.2248791	12148	125
42	.3813246	53338	3235	.4126205	.2261048	12256	109
43	.3870070	56824	3486	.4275995	.2273397	12350	94
44	.3930658	60588	3764	.4435266	.2285824	12427	77
45	.3995316	64658	4069	.4604745	.2298313	12489	62

表 G. $\log H(r, r_2)$ 之值: ϕ 由 0° 至 45°, r 由 3 至 50

$\phi°$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$
0	.2754537	.3097418	.3290589	.3414849	.3501576	.3565570	.3614744
1	4674	7523	0673	4921	1639	5624	4793
2	5084	7840	0930	5137	1824	5788	4938
3	5676	8365	1357	5496	2134	6059	5180
4	6721	9103	1956	5999	2567	6440	5520
5	7943	.3100049	2723	6644	3123	6923	5955
6	9444	1204	3661	7432	3801	7524	6485
7	.2761209	2565	4765	8362	4601	8226	7111
8	3242	4132	6037	9428	5522	9034	7831
9	5540	5904	7474	.3420635	6562	9947	8644
10	8101	7877	9075	1980	7720	.3570964	9550
11	.2770923	.3110051	.3300333	3461	8995	2093	.3620547
12	4003	2423	2761	5075	.3510386	3304	1635
13	7338	4990	4842	6823	1891	4625	2812
14	.2780926	7751	7079	8701	3508	6044	4075
15	4762	.3120701	9469	.3430707	5235	7560	5426
16	8844	3838	.3312010	2839	7071	9170	6860
17	.2793167	7159	4760	5095	9012	.3580874	8378
18	7726	.3130660	7532	7472	.3521058	2663	9976
19	.2802519	4337	.3320508	9968	3206	4[illegible]53	.3631654
20	7539	8186	3621	.3442579	5452	6524	3409
21	.2812782	.3142204	6870	5302	7795	8579	5239
22	8243	6385	.33302[illegible]0	8135	.3530232	.3590716	7142
23	.2823915	.3150726	3757	.3451074	2760	2933	9115
24	9794	5222	7387	4115	5375	5246	.3641157
25	.2835873	9866	.3341137	7256	8075	7594	3264
26	.2842145	.3164655	5001	.3460492	.3540857	.3600033	5435
27	8604	9583	8976	3819	3717	2540	7666
28	.2855244	.3174644	.3353057	7235	6652	5112	9955
29	.2862057	9833	7239	.3470734	9659	7747	.3652300
30	9035	.3185143	.3361516	4311	.3552732	.3610441	4697
31	.2876170	.3190569	5884	7965	5871	3191	7144
32	.2883456	6104	.3370339	.3481689	9070	5993	9636
33	.2890883	.3201741	4874	5480	.3562324	8844	.3662173
34	8443	7474	948[illegible]	9332	5632	.3621741	4750
35	.2906127	.3213297	.3384164	.3493242	8988	4631	7365
36	.2913927	9021	8908	7204	.3572388	7679	.3670013
37	.2921832	.3225181	.3393709	.3501213	5829	.3630671	2633
38	9834	.3231228	8563	5265	9306	3715	5400
39	.2937923	7335	.3403463	9354	.3582814	6787	8131
40	.2946089	.3243495	8404	.3513477	6350	9883	.3680884
41	.2954321	9700	.3413378	7627	9910	.3642993	3654
42	.2962610	.3255943	8381	.3521799	.3593488	6130	6438
43	.2970945	.3262216	.3423405	5989	7080	9274	9233
44	9316	8510	8446	.3530191	.3600682	.3652427	.3692036
45	.2987710	.3274317	.3433493	4399	4290	5584	4242

表 G. $\log H(r, r_2)$ 之値: ϕ 由 0° 至 45°, r 由 3 至 50 (續前)

$\phi°$	$r=10$	$r=11$	$r=12$	$r=13$	$r=14$	$r=15$	$r=16$
0	.3653717	.3685367	.3711582	.3733653	.3752489	.3768754	.3782941
1	3761	5407	1619	3686	2520	8783	2963
2	3892	5527	1729	3787	2614	8871	3051
3	4111	5725	1912	3956	2771	9018	3188
4	4416	6004	2166	4192	2990	9222	3380
5	4808	6361	2494	4494	3271	9435	3626
6	5257	6796	2893	4864	3614	9805	3927
7	5851	7310	3364	5289	4019	.3770183	4281
8	6500	7900	3907	5800	4484	0617	4688
9	7233	8563	4519	6365	5010	1108	5149
10	8030	9311	5201	6395	5595	1655	5661
11	8949	.3690129	5952	7639	6239	2256	6226
12	9929	1022	6771	8445	6942	2912	6841
13	.3660990	1987	7656	9263	7702	3622	7507
14	2123	3024	8608	.3740142	8518	4334	8222
15	3346	4132	9624	1081	9390	5199	8985
16	4639	5303	.3720703	2078	.3760317	6064	9796
17	6007	6553	1845	3132	1297	6979	.3790654
18	7447	7864	3048	4243	2329	7942	1558
19	8960	9240	4310	5400	3412	8953	2506
20	.3670541	.3700679	5630	6528	4544	.3780011	3498
21	2190	2179	7006	7839	5725	1113	4532
22	3904	3739	8437	9281	6952	2259	5606
23	5682	5357	9921	.3750591	8225	3443	6721
24	7522	7030	.3731456	2008	9542	4577	7874
25	9420	8757	3040	3472	.3770901	5946	9063
26	.3681376	.3710536	4672	4978	2301	7252	.3800289
27	3386	2365	6349	6527	3739	8595	1548
28	5448	4241	8069	8115	5215	9973	2839
29	7559	6163	9831	9742	6726	.3791383	4162
30	9718	8125	.3741631	.3761405	8270	2825	5514
31	.3691922	.3720129	3469	3102	9846	4296	6893
32	4167	2171	5341	4831	.3781452	5795	8299
33	6451	4249	7246	6589	3085	7320	9729
34	8772	6359	9182	8376	4745	8869	.3811181
35	.3701126	8500	.3751145	.3770188	6428	.3800140	2654
36	3510	.3730669	3133	2024	8133	2031	4146
37	5923	2862	5144	3881	9857	3641	5655
38	8360	5078	7176	5757	.3791593	5267	7180
39	.3710819	7314	9226	7649	3356	6907	8718
40	3297	9567	.3761291	9556	5127	8560	.3820267
41	5790	.3741834	3370	.3781475	6909	.3810223	1826
42	8296	4113	5458	3403	8699	1894	3393
43	.3720813	6490	7555	5338	.3800497	3572	4966
44	3335	8693	9657	7279	2299	5254	6543
45	5861	.3750990	.3771[illegible]63	9222	4103	6938	8122

表 G. $\log H(r, p_2)$ 之值: ϕ 由 0° 至 45°, r 由 3 至 50 (續前)

$\phi°$	$r=17$	$r=18$	$r=19$	$r=20$	$r=21$	$r=22$	$r=23$
0	.3795425	.3806194	.3816376	.3825253	.3833271	.3840548	.3847182
1	5450	6518	6399	5275	3292	0568	7202
2	5528	6591	6469	5341	3354	0628	7259
3	5657	6713	6584	5451	3459	0728	7355
4	5838	6884	6746	5605	3606	0867	7488
5	6070	7103	6954	5802	3793	1047	7660
6	6452	7370	7207	6042	4022	1266	7869
7	6686	7685	7505	6326	4293	1523	8116
8	7070	8048	7849	6652	4634	1820	8400
9	7503	8457	8237	7021	4955	2155	8721
10	7986	8913	8669	7432	5316	2529	9078
11	8517	9415	9144	7883	5776	2940	9471
12	9096	9962	9663	8376	6245	3358	9899
13	9723	.3810554	.3820224	8909	6753	3872	.3850363
14	.3800386	1190	0826	9482	7299	4336	0861
15	1115	1869	1470	.3830093	7881	4949	1393
16	1878	2591	2154	0743	8500	5540	1958
17	2686	3354	2877	1430	9154	6164	2556
18	3537	4157	3638	2153	9843	6822	3185
19	4430	5001	4437	2912	.3840566	7513	3845
20	5363	5882	5273	3706	1322	8235	4536
21	6336	6802	6144	4534	2111	8987	5256
22	7348	7757	7049	5394	2930	9770	6004
23	8397	8749	7988	6286	3780	.3850391	6780
24	9482	9774	8960	7209	4659	1420	7583
25	.3810602	.3820832	9962	8162	5566	2286	8411
26	1756	1921	.3830994	9142	6500	3178	9264
27	2941	3041	2055	.3840150	7460	4094	.3860141
28	4157	4189	3143	1184	8445	5034	1040
29	5702	5365	4257	2243	9453	5996	1961
30	6674	6567	5396	3325	.3850484	6980	2902
31	7973	7794	6558	4429	1535	7984	3862
32	9296	9043	7742	5554	2607	9007	4840
33	.3820642	.3830314	8946	6698	3696	.3860047	5836
34	2009	1605	.3840170	7860	4803	1104	6846
35	3395	2915	1411	9039	5926	2175	7871
36	4800	4241	2667	.3850233	7063	3261	8910
37	6220	5583	3938	1440	8213	4359	9960
38	7655	6938	5222	2660	9375	5468	.3871021
39	9102	8305	6518	3891	.3860547	6586	2091
40	.3830561	9683	7823	5130	1728	7713	3163
41	2028	.3841069	9136	6378	2916	8848	4254
42	3503	2462	.3850455	7632	4110	9987	5344
43	4984	3860	1780	8890	5308	.3871131	6438
44	6458	5262	3108	.3860151	6510	2278	7535
45	7953	6665	4437	1414	7712	3426	8633

表 G. $\log H(r, p_2)$ 之值：ϕ 由 0° 至 45°, r 由 3 至 50（續前）

$\phi°$	$r=24$	$r=25$	$r=26$	$r=27$	$r=28$	$r=29$	$r=30$
0	.3853256	.3858838	.3863984	.3868744	.3873160	.3877268	.3881099
1	3275	8855	4001	8760	3176	7283	1113
2	3330	8908	4052	8809	3223	7328	1157
3	3421	8996	4136	8891	3301	7404	1231
4	3549	9119	4254	9005	3411	7510	1333
5	3714	9277	4406	9151	3552	7647	1465
6	3915	9470	4591	9329	3724	7813	1625
7	4153	9697	4810	9539	3927	8008	1815
8	4423	9958	5061	9781	4160	8234	2032
9	4730	.3860253	5345	.3870055	4424	8488	2278
10	5073	0582	5661	0359	4717	8772	2552
11	5450	0943	6009	0694	5041	9084	2854
12	5860	1338	6388	1059	5393	9424	3183
13	6305	1764	6798	1454	5774	9792	3538
14	6782	2223	7239	1879	6183	.3880187	3920
15	7292	2712	7710	2332	6620	0609	4328
16	7834	3232	8210	2814	7084	1057	4762
17	8406	3782	8738	3323	7576	1531	5220
18	9009	4361	9295	3859	8093	2031	5703
19	9642	4969	9880	4422	8635	2555	6209
20	.3860304	5694	.3870491	5010	9203	3103	6739
21	0994	6267	1128	5624	9794	3674	7291
22	1711	6955	1790	6261	.3880409	4263	7865
23	2455	7661	2477	6723	1047	4883	8460
24	3225	8408	3187	7607	1706	5520	9076
25	4018	9170	3920	8312	2387	6177	9711
26	4836	9955	4674	9039	3088	6854	.3890366
27	5676	.3870762	5470	9786	3808	7550	1038
28	6538	1589	6245	.3880552	4547	8263	1727
29	7420	2436	7060	1337	5303	8993	2433
30	8322	3202	7893	2138	6077	9740	3155
31	9242	4185	8742	2956	6865	.3890501	3891
32	.3870180	5085	9608	3790	7669	1277	4642
33	1134	6001	.3880458	4638	8487	2067	5405
34	2102	6931	1382	5499	9317	2863	6180
35	3085	7874	2289	6372	.3890159	3682	6965
36	4080	8829	3208	7257	1012	4505	7762
37	5086	9795	4137	8151	1875	5338	8567
38	6103	.3880771	5075	9055	2746	6179	9380
39	7128	1756	6022	9967	3625	7028	.3900201
40	8162	2748	6975	.3890885	4511	7883	1028
41	9201	3746	7935	1809	5402	8744	1859
42	.3880246	4749	8900	2738	6298	9608	2695
43	1295	5756	9868	3670	7197	.3900476	3534
44	2346	6765	.3890838	4604	8098	1346	4375
45	3398	7775	1810	5540	9000	2217	5217

表 G. $\log H(r, r_2)$ 之值：ϕ 由 0° 至 45°，r 由 3 至 50 (續前)

$\varphi°$	$r=31$	$r=32$	$r=33$	$r=34$	$r=35$	$r=36$	$r=37$
0	.3884679	.3888034	.3891183	.3894146	.3896937	.3899572	.3902063
1	4694	8048	1197	4159	6949	9584	2074
2	4736	8089	1237	4198	6987	9620	2110
3	4808	8158	1304	4262	7050	9682	2170
4	4906	8254	1397	4353	7138	9767	2273
5	5034	8378	1516	4469	7251	9877	2360
6	5189	8528	1662	4611	7389	.3900011	2490
7	5372	8706	1835	4778	7551	0168	2643
8	5583	8910	2033	4970	7738	0350	2820
9	5822	9140	2256	5187	7948	0555	3020
10	6086	9397	2505	5429	8183	0783	3242
11	6378	9680	2780	5695	8442	1035	3486
12	6696	9988	3078	5985	8724	1309	3753
13	7041	.3890322	3402	6299	9029	1605	4041
14	7410	0680	3749	6636	9356	1924	4351
15	7805	1062	4120	6996	9706	2264	4682
16	8225	1469	4514	7379	.3900078	2625	5034
17	8668	1899	4931	7783	0471	3007	5405
18	9136	2351	5370	8209	0884	3409	5797
19	9626	2826	5830	8656	1319	3832	6208
20	.3890138	3323	6312	9123	1773	4273	6637
21	0673	3840	6814	9611	2246	4733	7085
22	1228	4379	7335	.3900117	2738	5212	7551
23	1804	4937	7877	0643	3248	5708	8033
24	2400	5514	8437	1186	3776	6221	8532
25	3015	6109	9014	1746	4321	6750	9048
26	3648	6723	9609	2324	4881	7296	9578
27	4299	7353	.3900220	2917	4558	7856	.3910123
28	4966	8000	0817	3525	6049	8431	0682
29	5649	8661	1459	4148	6654	9019	1255
30	6348	9338	2145	4785	7273	9621	1840
31	7060	.3900028	2815	5435	7904	.3910234	2437
32	7786	0732	3497	6097	8547	0859	3046
33	8525	1447	3190	6770	9201	1495	3664
34	9275	2174	4895	7454	9865	2141	4293
35	.3900035	2911	5610	8148	.3910530	2796	4930
36	0806	3657	6333	8850	1222	3460	5576
37	1585	4412	7065	9561	1912	4131	6229
38	2372	5174	7805	.3910278	2609	4809	6888
39	3166	5944	8551	1002	3312	5492	7553
40	3966	6719	9302	1732	4021	6181	8224
41	4771	7498	.3910058	2466	4734	6874	8898
42	5580	8282	0818	3203	5450	7571	9576
43	6392	9061	1581	3943	6169	8270	.3920256
44	7206	9857	2345	4686	6890	8971	0938
45	8021	.3910646	3111	5429	7612	9673	1621

表 G.　$\log H(r, \nu_2)$ 之值：ϕ 由 0° 至 45°，r 由 3 至 50 (續前

$\phi°$	$r=38$	$r=39$	$r=40$	$r=41$	$r=42$	$r=43$	$r=44$
0	.3904421	.3906678	.3908782	.3910801	.3912724	.3914556	.3916305
1	4433	6669	8793	0812	2734	4566	6315
2	4468	6703	8826	0844	2766	4597	6345
3	4526	6760	8881	0898	2818	4618	6395
4	4607	6838	8958	0973	2891	4720	6465
5	4711	6940	9057	1069	2985	4812	6554
6	4837	7063	9177	1187	3100	4924	6664
7	4987	7209	9319	1325	3235	5056	6793
8	5159	7377	9493	1485	3391	5208	6942
9	5353	7566	9667	1665	3567	5380	7109
10	5569	7777	9873	1865	3763	5571	7296
11	5808	8009	.3910099	2086	3978	5781	7502
12	6067	8262	0346	2327	4213	6011	7726
13	6343	8535	0612	2587	4467	6259	7769
14	6650	8829	0899	2867	4740	6526	8229
15	6972	9143	1205	3165	5032	6810	8508
16	7314	9477	1530	3483	5341	7113	8803
17	7676	9829	1874	3818	5669	7433	9116
18	8057	.3910201	2236	4172	6014	7770	9445
19	8457	0591	2616	4542	6376	8123	9791
20	8876	0998	3014	4930	6754	8493	.3920152
21	9312	1423	3428	5334	7149	8878	0528
22	9765	1865	3859	5754	7559	9279	0920
23	.3910235	2323	4305	6190	7984	9694	1326
24	0721	2796	4767	6640	8424	.3920124	1746
25	1223	3285	5243	7105	8878	0567	2179
26	1739	3788	5734	7584	9345	1024	2625
27	2270	4306	6238	8076	9826	1493	3084
28	2815	4836	6756	8581	.3920318	1974	3554
29	3372	5379	7285	9097	0823	2467	4035
30	3942	5935	7827	9626	1338	2970	4528
31	4523	6501	8379	.3920164	1864	3484	5030
32	5115	7078	8942	0713	2400	4007	5541
33	5718	7665	9514	1272	2945	4540	6062
34	6330	8261	.3920095	1839	3499	5081	6590
35	6950	8865	0685	2414	4067	5629	7126
36	7579	9479	1282	2997	4629	6185	7669
37	8215	.3920098	1886	3586	5204	6747	8218
38	8857	0724	2496	4181	5785	7314	8773
39	9505	1355	3112	4782	6371	7886	9332
40	.3920157	1991	3732	5386	6962	8463	9896
41	0814	2631	4355	5995	7556	9044	.3930463
42	1474	3274	4982	6607	8153	9627	1033
43	2136	3919	5612	7221	8752	.3930212	1605
44	2800	4566	6242	7836	9353	0799	2178
45	3465	5213	6874	8452	9954	1386	2752

表 G. $\log H(r, r_2)$ 之值: ϕ 由 0° 至 45°, r 由 3 至 50 (續完)

$\phi°$	$r=45$	$r=46$	$r=47$	$r=48$	$r=49$	$r=50$
0	.3917975	.3919572	.3921100	.3922565	.3923969	.3925316
1	7984	9581	1110	2574	3978	5325
2	8014	9610	1138	2601	4005	5352
3	8063	9658	1185	2647	4050	5396
4	8131	9725	1250	2711	4113	5457
5	8219	9811	1334	2794	4193	5536
6	8326	9915	1437	2894	4291	5633
7	8452	.3920039	1557	3012	4407	5746
8	8598	0181	1697	3149	4541	5877
9	8762	0342	1854	3302	4692	6025
10	8944	0520	2029	3474	4859	6189
11	9145	0717	2221	3662	5044	6370
12	9365	0932	2431	3868	5245	6569
13	9602	1164	2658	4090	5463	6781
14	9857	1413	2902	4329	5697	7010
15	.3920129	1679	3163	4584	5947	7255
16	0418	1962	3440	4855	6213	7516
17	0724	2261	3732	5142	6493	7791
18	1046	2576	4041	5444	6789	8080
19	1383	2906	4364	5760	7099	8384
20	1737	3252	4702	6092	7424	8702
21	2105	3612	5055	6437	7762	9034
22	2488	3987	5421	6796	8114	9378
23	2885	4375	5801	7168	8478	9736
24	3295	4776	6194	7553	8855	.3930105
25	3719	5191	6600	7950	9244	0486
26	4155	5628	7018	8359	9645	0879
27	4603	6066	7447	8779	.3930057	1282
28	5063	6506	7887	9210	.0479	1696
29	5534	6967	8338	9652	0911	2120
30	6015	7437	8799	.3930103	1353	2553
31	6506	7918	9269	0563	1804	2995
32	7006	8407	9748	1032	2263	3445
33	7515	8905	.3930235	1509	2731	3903
34	8032	9410	0729	1993	3205	4368
35	8556	9923	1231	2485	3685	4840
36	9087	.3930442	1740	2982	4174	5318
37	9624	0967	2254	3486	4667	5801
38	.3930166	1498	2773	3994	5165	6289
39	0713	2033	3297	4507	5667	6781
40	1264	2572	3824	5024	6174	7277
41	1818	3115	4355	5543	6683	7776
42	2376	3660	4889	6066	7195	8278
43	2935	4207	5424	6590	7703	8781
44	3496	4755	5961	7116	8223	9286
45	4057	5304	6498	7642	8739	9791

IV. 參考書

本書引用之參攷書如下:

1. Baten: Mathematical Statistics, (John Wiley and Sons Inc. .

2. Burgess, R. W.: An Introduction to the Mathematics of Statistics, *Boston (Houghton Mifflin Co.), 1927.*

3. Bowley. A. L: F Y. Edgeworth's Contributions to Mathematical Statistics. London (P. S. King & Son), 1928.

4. Brunt, D.: Combination of Observations, London (Cambridge University Press).

5. Camp: *The Mathematical Part of Elementary Statistics*, (D. C. Heath Co.).

6. Charlier, C. V. L.: Researches into the Theory of Probability. Acta Universitatis Lundensis, 1905.

7. Comrie, L. J. (Editor): Barlow's Tables of Squares, Cubes, Square Roots, Cube Roots and Reciprocals, London (E. & F. Spon, Ltd.), 1919.

8. *Coolidge: A Treatise on Probability.*

9. Davis and Nelson: Elements of Statistics, (Principia).

10. Edgeworth, F. Y.: On the Representation of Statistics by Mathematical Formulae. Jour. Roy. Stat. Soc., 63:72, 1900.

11. Elderton W. P.: Frequency Curves and Correlation, London (Cambridge University *Press*), 1938.

12. Fisher, Arne The Mathematical Theory of Probability.

13. Fisher, R. A.: Statistical Methods for Research Workers, London (Oliver & Boyd), 1938.

14. Fry: Probability and Its Engineering Uses.

15. Glover, J. W.: Tables of Applied Mathematics in Finance, Insurance, Statistics, Ann Arbor, Mich. (George Wahr), 1923.

16. Gram, J. P.: On the Development of Series by Means of the Method of Least squares. Doctor's Thesis in Danish. University of Copenhagen, 1879.

17. Greenwood, M.: Epidemic & Crowd-Diseases, New York (Mac. Millan Co.) 1935.

18. Haskell, A. C.: How to Make and Use Graphic Charts, New York (Codex Book Co.), 1919.

19. Hill, A. B.: Principles of Medical Statistics, London (The Lancet Ltd.), 1937.

20. Holzinger, K. J. and B. C. Mitchell.: Exercise Manual in Statistics, New York (Ginn and Co.), 1929.

21. Jones, D. C.: A First Course in Statistics, London (G. Bell & Sons), 1921.

22. Kapteyn, J. C.: Skew Frequency Curves in Biology and Statistics, New York. (E. Steiger and Co.), 1903. p. 45.

23. Kapteyn, J. C. and M. J. Van Uven: Skew Frequency Curves in

Biology and Statistics. Groningen (Hoitsema), 1916, p. 69.

24. Kelley, T. L. Statistical Methods, New York (Mac Millan Co.).

25. Kenney, J. F.: Mathematics of Statistics, New York (D. Van Nostrand Co.), 1941.

26. Koren, John: The History of Statistics, Their Development and Progress in Many Countries.

27. Love: Statistical Methods for Agricultural Research, Shanghai (Commercial Press Ltd.), 1937.

28. Mills, F. C.: Statistical Methods Applied to Economics and Business, New York (Henry Holt), 1938.

29. Pearl R.: An Introduction to Medical Biometry and Statistics, Philadelphia (W. B. Saunders Co.), 1940.

30. Pearl R.: Biometric Data on Infant Mortality in the United States Birth Registration Area, 1915–1918. Am., Jour. Hyg., 1:419, 1921.

31. Pearson, Karl: Tables for Statisticians and Biometricians, Part I, London (Cambridge University Press), 1914.

32. Pearson, Karl: Skew Variation in Homogeneous Material, Philos. Trans., A, Part I, 186:343, 1895.

33. Pearson, Karl: Supplement to a Memoir on Skew Variation, Philos. Trans., A 197:443, 1901.

34. Pearson, Karl: Second Supplement to a Menoir on Skew Variation, Philos. Trans., A. 216:429, 1916.

435. Pearson, Karl: The Fifteen Constants Bivariate Frequency Surface, *Biometrika*, 17:268, 1925.

436. Pearson, Karl: on the Probable Errors of Frequency Constants, Part 1, Biometrika, vol. 2, 1903, p. 273, Part 2, ibid., vol. 9, 1913, p. 1, and Part 3, ibid., vol. 13, 1920, p. 113.

437. Richard Son: An Introduction to Statistical Analysis, Harcourt (Brace & Co.).

438. Rietz, H. L.: Mathematical Statistics, Chicago (Open Court Publishing Co.), 1927.

439. Snedecor: Statistical Method, Collegiate Press Inc., Ames, Iowa.

440. Thiele, T. N.: Theory of Observations, London (Charls and Edwin Layton), 1889, p. 142.

441. Tippett, L. H. C.: Methods of Statistics, London (Williams & Norgate).

442. Walker. Mathematics Essential for Elementary Statistics (Henry Holt & Co.).

443. Walker, H. M.: Studies in the History of Statistical Methods with Special Reference to Certain Educational Problems, Baltimore (Williams & Wilkins Co.), 1929.

444. Whittaker, E. T. & G. Robinson: Calculus of Observations, London (Blackie & Co.).

445. Yule, G. U. & M. G. Kendall: An Introduction to the Theory of Statistics, London (Griffin & Co.), 1937.

V. 中西譯名對照表

本對照表所列之中文統計譯名,大體以教育部公布之統計學名詞(民國三十五年五月,國立編譯館編訂,上海正中書局印行)一書為準;但著者認為該書上名詞有修正必要者,則另擬譯之.即如 logistic curve 一詞該書譯為"羅吉斯曲線,"本表譯為蒲芮曲線,蓋該曲線常被稱之為 Pearl-Reed curve 故也;又如 beta function 及 gamma function 二詞該書各譯為 beta 函數及珈瑪函數,本表各譯為 β 函數及 Γ 函數是.又本對照表內曾附列外國地名及人名及其中文譯名.

A

abscissa, 橫坐標.
adding machine, 加算機.
approximate mode, 近似衆數.
area, 面積.
area element, 面積元素.
arithmetic average, 算術平均數.
arithmetic mean, 算術平均數.
average, 平均數.
asymptote, 漸近線.

B

barometric height, 氣壓高度.
beta function, β 函數.
binomial distribution, 二項分配.
Barlow's Tables, 巴樓表.
Bernoulli distribution, 伯努里分配.
Bernoulli series, 伯努里級數.
Bernoulli theorem, 伯努里定理.
Bowley, A. L., 鮑萊.
Bureau of the Census, 人口清查局.

C

calculating machine, 計算機.
caption, 縱標目.
carriage, 載輪車.
carriage shift lever, 移車桿.
centroid, 重心.
chest girth, 胸圍.
chi-square, χ^2 測驗.
clear crank, 還原柄.
clear key, 還原鍵.
coefficient of alienation, 疏相關係數.
coefficient of association, 相聯係數.
coefficient of average deviation, 平均差係數.
coefficient of colligation, 束聯係數.
coefficient of contingency, 列聯係數.
coefficient of correlation, 相關係數.

coefficient of dispersion, 離勢係數.

coefficient of mean deviation, 平均差係數.

coefficient of quartile deviation, 四分位差係數.

coefficient of skewness, 偏度係數.

coefficient of standard deviation, 標準差係數.

coefficient of total correlation, 總相關係數.

coefficient of variability, 離勢係數.

coefficient of variation, 離勢係數.

cofactor, 餘因子.

column, 柱列.

column total, 柱和.

constant, 常數.

constant of central tendency, 集中常數.

constant of dispersion, 離勢常數.

constant of position, 地位常數.

continuous operation, 連續運算.

correlation, 相關.

correlation coefficient, 相關係數.

correlation ratio, 相關比.

convergent tendency, 收斂趨勢.

co-variance, 共差.

crude moment, 輔助動差.

cummulative frequency diagram, 累計頻數圖.

curve fitting, 曲線配合.

Carmarthen, 加馬耳站.

Charlier check, 沙利耳驗算法.

Charlier, C. V. L., 沙利耳.

Craig, A. T., 科雷格.

D

degree of cloudiness, 陰雲各度.

degree of freedom, 自由度.

density curve, 密度曲線.

deviation from average, 離均差.

deviation from mean, 離均差.

distribution function, 分配函數.

double integral, 二重積分.

double logarithmic paper, 雙對數格紙.

double sum, 二重和.

drawing board, 畫圖板.

drawing instrument, 圖畫器.

De Moivre, 狄摩夫.

Duddee, 丹笛.

E

empirical formula, 實驗式.

empirical mode, 經驗眾數.

empirical probability, 經驗機率.

equation of transformation, 變換方程式

exact scale, 實在尺度.

expected value, 期望值.

exponential curve, 指數曲線.

Edgeworth, F. Y., 愛奇渥斯.

Elderton, W. P., 愛爾德頓.

F

first quartile, 第一四分位數.

fixed element, 固定元素.

fixed point method, 定點法.

fourfold table, 四格表.

fourfold table correlation coefficient, 四格表相關係數.

french curve, 雲形板.

frequency curve, 頻數曲線.

frequency distribution, 頻數分配.

frequency function, 頻數函數.

functional determinant, 函數行列式.
Farr, William, 傅耳.
Fisher, Arne, 費恩.
Fisher, R.A., 費楷.
Friden calculating machine, 傅瑞登計算機.

G

gamma function, Γ 函數.
general moment, 一般動差.
geometric average, 幾何平均數.
geometric mean, 幾何平均數.
goodness of fit, 配合適度.
graphic method, 圖解法.
grouping error, 分組誤差.
Galton, Francis, 高爾登.
Gauss, Karl Friedrich, 高斯.
Glasgow, 哥拉斯誇.
Gompertz curve, 高培志曲線.
Gram, J. P., 葛爾珩.
Greek alphabet, 希臘字母.
Greenwich, 哥林威治.

H

hard eraser, 硬橡皮.
hard pencil, 硬鉛筆.
harmonic average, 倒數平均數.
harmonic mean, 倒數平均數.
heading, 標目.
high contact, 高度相切.
histogram, 直方圖.
horizontal asymptote, 水平漸近線.
horizontal axis, 橫軸.
hypergeometric series, 超比度數.

I

improvement factor, 進步因子.
incomplete beta function, 不完全 β 函數.
incomplete gamma function, 不完全 Γ 函數.
index of precision, 精確指數.
india ink, 繪圖墨水.
integration by parts, 部分積分法.
intelligence, 智力.
intercept, 截距.

J

J-shaped frequency distribution, J 形頻數分配.
Jones, D. C., 瓊斯.

K

key board, 打字盤.
kurtosis, 峯度.

L

last scale, 最末尺度.
law of growth, 成長定律.
leptokurtic distribution, 高狹峯分配.
lettering guide, 寫字板.
life table, 壽命表.
linear correlation, 直線相關.
linear function, 直線函數.
line of regression, 迴歸線.
line polygon, 多邊圖.
logarithmic curve, 對數曲線.
logarithmic table, 對數表.
logistic curve, 蒲芮曲線.
lower dial, 記數輪.
lower quartile, 下四分位數.

lung capacity, 肺活量.
Laplace, Pierre Simon, 雷普拉斯.
Laplacian curve, 雷普拉斯曲線.
Londonderry, 倫敦迪亞.

M

manila cards, 統計卡片.
marginal frequency, 邊際頻數.
mean, 平均數.
mean deviation, 平均差.
machanical verifier, 驗孔機.
median, 中位數.
method of approximation, 逐近法.
method of grouping, 分組法.
method of least squares, 最小平方法.
method of moments, 動差法.
minor, 伴列式.
minus bar, 減數器.
mode, 衆數.
moment, 動差.
mortality table, 死亡表.
motor, 電動機.
multiple correlation, 複相關.
multiple correlation coefficient, 複相關係數.
multiple regression line, 複迴歸線.
mutual deviation, 均互差.
Mac Donald, David, 馬克唐納.
Marchand calculating machine, 馬強德計算機.
Mills, C. F. 密而思.
Monroe calculating machine, 孟祿計算機.

N

nearest scale, 最近尺度.
negative correlation, 負相關.
net correlation, 淨相關.
net correlation coefficient, 淨相關係數.
non-linear correlation, 非直線相關.
non-repeat key, 加減鍵.
normal correlation surface, 常態相關曲面.
normal curve, 常態曲線.
normal curve of error, 常態差誤曲線.
normal equations, 標準方程式.
normal form, 法線式.
normal law, 常態定律.
normal probability curve, 常態機率曲線.

O

ogive, 肩形圖.
ordinary graph paper, 普通格紙.
ordinate, 縱坐標.
original curve, 原始曲線.
operating crank, 運算柄.

P

pantograph, 放大器.
parabola of second order, 二次拋物線.
parameter, 參數.
pattern, 模式.
partial correlation, 淨相關.
partial correlation coefficient, 偏相關係數.
partial derivative, 偏微係數.
perfect correlation, 完全相關.
platykurtic distribution, 扁峯分配.

plus bar, 加閘閥.
polar coordinate, 極坐標.
polar coordinate paper, 極坐標格紙.
polar planimeter, 求積儀.
positive correlation, 正相關.
principal moment, 主要動差.
priori probability, 事前機率.
product moment, 積差.
product moment coefficient of correlation, 積差相關係數.
probability, 機率.
probability density, 機率密度.
probability function, 機率函數.
probability paper, 機率格紙.
probability scale, 機率尺度.
probable error, 機誤.
proportion, 比例數.
proportional divisor, 比例儀.
punching machine, 打孔機.
Pearl, Raymond, 蒲耳.
Pearl-Reed curve, 蒲芮曲線.
Pearson, Karl, 皮爾生.
Pearson's differential equation, 皮爾生微分方程式.
Pearson's formula, 皮爾生公式.
Pearson's frequency curve, 皮爾生頻數曲線.
Poisson exponential series, 卜瓦松指數級數.

Q

qualitative data, 品質論據.
quadrant, 象限.
quadrature formula, 求積公式.
quantitative data, 數量論據.
quartile, 四分位數.
quartile deviation, 四分位差.
Quetelet, Lambert Adolph Jacques, 郭泰來.

R

radian, 弳.
rank correlation, 等級相關.
range, 全距.
rational formula, 理解式.
rational integral expression, 有理整式.
reciprocal, 反商.
recursion formula, 循環公式.
regression coefficient, 迴歸係數.
relative frequency, 相對頻數.
repeat key, 重除鍵.
residual, 剩餘數.
rigid lever, 剛體槓桿.
root-mean-square-deviation, 根均方差.
ruling pen, 鴨嘴筆.
Reed, L. J. 芮德.

S

scale transformation, 尺度變換法.
scatter diagram, 散佈圖.
secondary subscript, 次下標.
second quartile, 第二四分位數.
semi-logarithmic paper, 半對數格紙.
semi-convergent series, 半斂級數.
shifting bar, 移車桿.
significant test, 顯著性測驗.
simple correlation, 單相關.
simple exponential law, 簡單指數定律.
size of sample, 樣本之大.
skew, 偏態.

skew frequency curve, 偏態頻數曲線.
skew frequency distribution, 偏態頻數分配.
skewness, 偏度.
slab total, 壁和.
slide rule, 計算尺.
slope, 斜率.
slope intercept form, 斜截式.
smoothed curve, 平滑曲線.
soft eraser, 軟橡皮.
sorting machine, 分類機.
square root, 平方根.
standard deviation, 標準差.
standard error, 標準誤.
standard scale, 標準尺度.
standard error of estimate, 估計標準誤.
statistical constant, 統計常數.
statistical laboratory, 統計實驗室.
statistical method, 統計方法.
statistical moment, 統計動差.
statistics, 統計.
straight line, 直線.
stub, 橫標目.
successive subtraction method, 遞減法.
surface of revolution, 旋轉曲面.
Scandinavian School, 斯堪地納維亞學派.
Sterling's approximation formula, 史特靈近似值公式.
Sheppard, W. F., 薛伯.
Sheppard's correction, 薛伯校正數.
Sinclair, John, 辛可萊.
Southampton, 南安普敦.

T

tabulating paper, 製表紙.
theorem of Bernoulli, 伯努里定理.
theory of statistics, 統計學.
third quartile, 第三四分位數.
thumb tacks, 圖釘.
title, 標題.
total regression curve, 總迴歸線.
total variance, 總變差.
tracing paper, 描圖紙.
transcendental function, 超越函數.
transformed curve, 變換曲線.
triangles, 三角板.
triangular ruler, 三稜尺.
triple sum, 三重和.
true constant, 真正常數.
trying division method. 試除法.
two point form, 兩點式.
Tables for Statisticians & Biometricians, 皮爾生表.
Taylor's theorem, 泰來定理.
Tchebycheff's criterion, 哲迭夫準則.
Tchebycheff's inequality, 哲迭夫不等式.
Tchebycheff's theorem, 哲迭夫定理.
Thiele, T. N., 蒂耳.
T-square, 丁字尺.

U

upper dial, 頂尖台.
upper quartile, 上四分位數.
U-shaped frequency distribution, U形頻數分配.

V

variance, 變差.
vertical asymptote, 垂直漸近線.
Verhulst. P. F., 費赫斯特.

VI. 索 引

本索引係筆畫多寡次序而編列者. 各名詞後邊附註之阿拉伯字碼代表數頁.

六　　畫

七　　畫

八　　畫

九　畫

十　畫

十 三 畫

十 四 畫

十 五 畫

十 六 畫

十 七 畫

十 八 畫

十九畫

二十二畫

二十三畫

雜 組

中華民國三十七年一月初版

(60821·2)

國立復旦大學商科研究所叢書 高等統計學一冊

定價國幣拾肆元

印刷地點外另加運費

版權所有 翻印必究

著作者 薛仲三

發行人 朱經農 上海河南中路

印刷所 商務印書館印刷廠

發行所 商務印書館 各地

山 参
G a
X Z S
几也寺臺記