統計方法

統計方法

STATISTICAL METHODS

美國F. C. MILLS著

李黃孝貞
陸宗蔚 譯

中華書局印行

譯本序

美國密爾斯教授（Prof. F. C. Mills）所著統計方法（Statistical Methods Applied to Economics and Business）爲研究統計學者最適用之初級教本，出版已十有餘年而聲價勿替，吾國各大學亦多採用之。孝貞宗蔚年來曾數度講授是書，爰爲合譯，以貢其斆學相長之一得。

是書着重於經濟及商業統計方法之研討。其優點有二：一、理論精闢，無蕪雜膚淺之弊，並注意於應用所設實例，足爲舉一反三之助；二、對於統計量數（statistical measures）之意義及其應用之方法，敍論不厭其詳，而對於各種量數在應用時所受之限制，尤再三致意，良以統計量數多係根據數學公式，應用時必須合於若干條件，並非任何資料皆可適用，且對於結論之解釋，亦常受資料性質上之限制，此在經濟及商業統計爲尤然。英人恆言所謂統計可以證明任何事物者，蓋指統計資料之解釋而言，解釋苟不得當，則數字斑斑，具文而已，何足以語於眞理之左證哉。

譯本係根據 1935 年九月版之原本，分任譯述，初稿各自起草，脫稿以後，互爲校勘，以求一貫。行文雖取直譯法，然艱澀語氣亦所力避。其中術語大體以書末所載中國統計學社編訂之統計譯名爲藍本。

李黃孝貞
陸宗蔚

三十年九月一日

原著序

近十年來經濟界與研究社會科學者，對於數量法（quantitative methods）之應用，漸見深切之注意。在昔恃個人之智慧以解答經濟上之問題，或憑未證實之假說以論斷社會問題者，已不復見於今日。社會學家及嚴正之經濟學者亦步自然科學家之後，基於事實，以施其觀察與分析，靡然成風，遠非古昔所能企及。此種觀察既屬於數量性質，則欲整理之，解釋之，自當應用適當方法。本書卽爲論述前項觀察之各種整理與分析之方法而作，而尤偏重於實際經濟資料之研究。

本書之論述，必欲限於經濟方面者，其故有二：一、普通統計方法之應用雖廣，然任何方面之研究，各有其特殊問題，而尤以在經濟方面所遇之困難與特殊問題爲多，於是切應此種特殊需要之方法，應運而生，欲以此種方法筆之於書，則此書所及之範圍，固不得不加以限制。二、方法之研究，如以專門論題爲例證，解釋當愈顯明，蓋抽象之論述，每爲讀者所厭倦。本書之專爲適合經濟界從事數量研究者之需要而作者，職此之由。

本書於方法之闡明，不厭其詳，而於純屬數學方面之演算，則務求簡略。又本書係備學者研究之用，非供斯學已有造詣之敎師閱讀，故凡所解釋，概以適切前者之需要爲準。此書爲統計學之入門，非窮究高深學理之作，對於意義深奧之定理，其解釋每略而不詳。

今日一般所用之數量分析法，乃係積聚多年來各界人士之貢獻，逐漸推演之結晶。對於統計學之發展，曾有相當貢獻者，在本書未遑遍舉，書中若干處，對於前人雖間有徵引，然近世統計學家所享受前人之功業，實非著者數言之所能盡。

本書各種資料之蒐集，備承各方之贊助。第十一、十三、十四、三章內之一部分統計資料，爲安得孫(H. E. Anderson)、凱洛(H. B. Killough)

兩氏所供給。哈佛大學經濟研究委員會潘遜斯教授(Prof. Warren M. Persons)慨允將彼對於物價指數研究所得之結果，刊入本書。第七章內之建築費指數（the index of building costs）與第九章內之商情指數(the index of business conditions）係美國電話電報公司統計部所編製。此項資料之採入，事前嘗得該部主任安德魯氏(Seymour L.Andrew)之特許。本書油印初稿中頗有意義晦澀之處，賴加利福尼亞大學毛物雷教授(Prof. A. H. Mowbray)之指示，得以修正。他如哥侖比亞大學摩爾(Prof. Henry L. Moore)、布朗（Prof.Theodore H. Brown）兩教授暨經濟學院舒爾茲氏(Henry Schultz）將原稿一部分爲之校訂，均所心感。塔特氏(Herbert F. Tutt)擔任統計圖表之繪製，襄助良多。在編印本書期間內之各種工作，如資料之蒐集，統計圖之繪製，以及印刷事務之管理等，隨時得達文波特氏(Donald H. Davenport)之襄助。對於余妻所給與不斷的助力，更爲銘感。

密爾斯(Frederick Cecil Mills)

一九二四年十一月

統計方法

目　次

統計圖索引

統計方法

第一章　統計方法與經濟及企業問題

今日普通所用'企業'(business)之術語，包含人類活動之意義，不一而足。吾人對於處理經濟資料與解決企業問題之方法既欲有所闡明，則對於各種企業之活動，凡可應用統計方法者先當作一簡略之敍述。

企業活動之類別

企業家日常所遇之事件，可分爲三類。多數企業家所從事活動之範圍，儘有僅限於其中一類者，然企業界應須處理之各問題，每兼三類而有之。

第一類爲發生於生產程序方面者，如關於物理、化學、工程學、牧畜學、航海術等之問題。解答此種問題所需之基本專門智識，爲吾人經濟生活之基礎。此類活動乃屬於處理原料、控制天然力之區域者也。

第二類爲關於企業個體內部組織與管理方面之活動。凡管理關於滿足人類慾望一切有機與無機物質之職務，各有其執行管理之單位，如農場、礦山、工廠、鐵道、百貨商店等是。企業家在組織此種單位，調劑其各部分間之工作，或監督每個組織內各員工之日常工作時，即遇種種新問題。此種問題縱不及前述第一類生產技術問題之緊要，但與一般企業家關係尤爲密切。科學方法對於解決此種問題，應用較難，且關於此種問題，既無成系統之智識，以爲引導，又乏有訓練之專家，以託付解答之

職責。

前述兩類之經濟活動，性質較爲集中，且可受人力之控制。如鋼鐵製造家有其熔鍊鋼鐵之技術問題，與其管理上之各種職務；農夫、礦主所遇問題亦同。凡執行此等業務，雖各有其特殊問題，惟此各種問題之原素，多少皆受人力之控制；故縱有困難發生，亦爲普通企業上本身之困難，既非由於問題所含原素之驟變，亦非由於新因素之驟然滲入，蓋可斷言。

至於第三類之經濟活動，其特殊之點，在於其所含之原素非任何個人所能控制。此類活動包括商品之買賣及與物價有關之一切經濟活動。在現時經濟組織之下，此種活動爲企業家所視爲最重要之業務，蓋企業家經濟活動之最終目的，在於銷售其產品以獲取贏利，產品銷售前之一切預備手續，如生產之技術、與夫企業本身之組織與管理，不過爲達到此最終目的之階梯。且企業家在賣買商品之時，恆遇種種問題，其要素非人力所得控制，如購置原料、採集生產必需之其他要素、與銷售製品時，企業家必須與各種市場——商品市場、勞力市場、金融市場等——互相交易，而在此等市場之中，其物價制度 (the system of prices) 實非企業家所能左右。故對於第一二兩類之企業活動，企業家雖常能控制，然對於處理其最後與其最重要部分，即以善價出售其產品之問題時，其控制力漸見薄弱。夫企業活動之原動力，在於獲得金錢之贏利，而欲獲得金錢之贏利，在於商品買賣之得宜，而欲求商品買賣之得宜，須在人力不能控制之物價制度中，攫取優先之地位，此物價制度所以爲近代企業最主要之因素也。

近代企業家實處於物價環境 (an environment of prices) 之中。此處'環境'之一名詞，非不切之喻；此企業家所服務之物價世界，係各個獨立部分所組成之互相團結、互相關聯之制度，而又爲包括企業上一切活動之制度。此種制度既非人力所能控制，則惟有適應之、瞭解之、以爲活

動之南針，苟對此茫無所知，則企業上之主要問題，將無法解決矣。

經濟與企業問題之數量性質

各種企業問題之屬於上述第一類者，其性質大率爲數量的，久爲吾人所公認。此種問題須賴自然科學界所應用之精密方法，以求解決，卽其他兩類之純屬於經濟及企業之問題，亦有應用數量法之必要，蓋在解答此種問題時，雖常包含品質上與不涉量度之觀察，而此種觀察大抵可以數量爲基礎。各種事實經數量上之比較，卽可作爲處理企業上問題與解釋經濟理論之根據。統計方法實卽以數量爲主體，而比較事實之方法也。

前節所述之三類問題，其中第二、第三、兩類，屬於本文討論範圍之內。統計方法雖亦可用以解答生產技術問題之一部分，然本書宗旨，不在討論此種問題，惟其他兩類問題——屬於企業個體內部組織與管理之問題及企業家與物價制度發生關係之商品買賣問題——之解答，則統計分析方法，至爲適用。

統計方法與企業內部管理問題

一般企業家在其組織之管理方面，必須處理各種數量資料，資料之屬於數量性質者，因其爲用各種數量單位以表示者也。彼須計算煤之噸數，煤氣之立方英尺數，工力之仟瓦特小時數 (kilowatt hour)，生鐵產量之噸數，皮鞋出產之雙數，機械小時數 (machine hour) 及人工小時數 (man hour)，以及用貨幣單位所表示之工資、生產費用及商品售價之數。近代企業組織之規模日見擴大，因此處理之資料亦日就繁複，而其眞實意義，遂漸難判明。僅憑個人之智力或固定的管理方法，實不能將大量之資料 (masses of data) 作有效的分析，而管理規模較大之企業組織。報償遞減法則 (law of decreasing returns) 之能在企業中

發生作用者，多由於企業管理方面之困難也。

吾人處理大量資料時，首須整理、簡縮、而分析之。吾人須將資料化繁就簡，而後可着手處理；又須將資料加以分析，而後可瞭解問題之要素。統計方法乃爲整理、簡縮與分析大量資料之利便而推演者也。

試舉例以明之。設吾人所欲解決之問題爲成本之分析 (allocation of costs)，卽會計學中所謂成本會計是也。如欲將該問題之各因素，作適當之分析，則非應用統計方法不爲功。會計方法祇限於處理金錢方面之事實，對於各項支出之總分析，殊非此法所能盡事，蓋除成本以外，各項支出之本身，亦應作統計分析，以明其變動，比較其今昔，而察其有無浪費之處。又銷貨紀錄(sales record) 之分析，亦須將資料化繁就簡，以簡單明瞭之形式表出之，而後可斷定其意義。至於市場之分析，進貨紀錄 (purchasing record) 之考查，以及商品之研究，皆須利用數量法，以數量法之應用，不限於任何度量之單位也。在企業內部管理之各方面，統計方法可補會計方法之不逮，增廣管理人員之智識，並可促進企業管理上之效率。

統計方法與企業本身以外之問題

企業家在市場中從事買賣，卽與物價制度發生密切關係，故用數量法以分析物價之變動，實爲切要之圖。但物價資料至爲繁多，欲處理之，必須化繁就簡，且商業具有週期變動之特性，如企業家欲採取與週期變動適應之政策，更宜應用適當之工具，以分析此種現象。此外如關於經濟分配等各種問題，雖與企業家關係較少，而爲經濟學家所推重。此種問題，在在與價値、價格等發生關聯，且其解決之途徑亦繫於數量之研究。

然則此種數量法果爲何種方法乎？且應用此法以研究之問題，又何以異於他種研究之問題乎？科學之研究，無論採取何種方法，必須經愼

重之觀察、合理之推考、與準確之證實。數量法之異於他法者，卽在觀察、推考、證實三方面，皆可以測度(measurement)爲依據，故較諸與數量無涉之分析方法，更爲確切。一種科學在不能應用數量以前，不論其研究者之悟解力如何靈敏，工作如何勤奮，其觀察與其所得之結論，必缺乏準確性。吾人藉測度法(methods of measurement)始可用精確之單位，分析各問題之原素，實予科學研究以種種利便。數學及其支系之統計與會計，皆爲近代經濟學家採用之利器，且因研究機關與研究方法之發展，企業界之實際應用亦漸見推廣矣。

統計家所用之工具僅爲切合於特殊研究所用之數學方法。此種特殊研究，在統計方法發達之初期，不屬於經濟方面，而屬於社會、政治、與人體測量方面，而另一方面之發展(關於機率原理 the theory of probabilities 者)，則由論理學之領域，而及於賭博。統計方法在昔爲用殊狹，今則應用之廣不可同日而語。斯學經相當之演進，應用漸博，獲得良好之結果，在經濟上之應用，僅其一例耳。經濟學家自利用統計方法後，頗能得心應手，而其工作之準確性大見增加；企業方面亦與較爲深奧之經濟科學，同受其利。

至於統計方法應用上之限制，本章中可暫置勿論。統計方法之應用固有時而窮，且應用統計方法者每忽視其應用上之限制。關於此點，容俟後文討論，惟此處所當致意者，統計方法僅爲一種工具，吾人應善爲利用之，由統計分析所獲之結果，應善爲解釋之，蓋吾人尚未發明一種完善之自動機，可將混雜之事實投入機械之一端，各種問題之答案卽可由另一端得之；換言之，卽吾人之理解力與批判力，並不因所用統計研究方法之進步遽可廢棄也。

第二章　圖示法

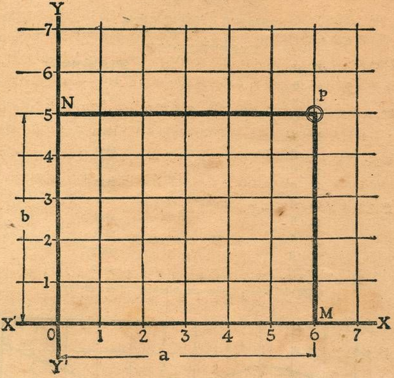
吾人欲將整理、分析及解釋企業與經濟事實之方法作一研究，必先討論若干基本之原則，而此項基本原則涉及統計學之範圍者少，涉及數學之範圍者多。本章爰將若干數學概念予以解釋，後此各章均將時加引證。此種數理雖或為讀者所共喻，在此實有簡略敍述之必要也。

統計分析大抵以基於實際測度而以貨幣或數量單位表示之資料為對象。哲學家笛卡兒氏（Descartes）所發明之坐標幾何法（method of coördinate geometry），對於前項資料之處理甚切實用。茲將解析幾何學之基本原理，簡述如次。

垂直之縱橫坐標

試於平面上繪一交叉成直角之兩直綫，則在該平面之任何一點，即可依兩綫之交叉點以說明其地位。吾人稱此兩直綫中之一綫（縱綫）為 $Y'Y$，另一綫（橫綫）為 $X'X$，其交叉點（原點 origin）為 O（參看圖一）。設 P 為平面上之任何一點，試作與 $Y'Y$ 綫平行之 PM 綫，與 $X'X$ 綫交於 M 點；又作與 $X'X$ 綫平行之 PN 綫，與 $Y'Y$ 綫交於 N 點。今假定 OM 之長度為 a 單位，ON 之長度為 b 單位，a 與 b 即為 P 之縱橫坐標（coördinates），說明以原點為標準時 P 點之地位。圖中 a 為 6，b 為 5。沿 x 軸（x-axis）之距離 a，稱為 P 點之橫坐標（abscissa）；沿 y 軸（y-axis）之距離 b，稱為 P 點之縱坐標（ordinate）（記法應先記橫坐標，後記縱坐標）。在同一平面上其他任何一點之縱橫坐標，均可依此決定；反之，任何兩實數，如以一數作為橫坐標，另一數作為縱坐標，亦可在平面上決定一點之地位。

平面上一點之地位，或在原點O之左，或在其右，或在其上，或在其下。通常以在原點右方之橫坐標為正，在原點左方之橫坐標為負；在原



圖一. 垂直縱橫坐標圖中點之位置

點上方之縱坐標爲正，在原點下方之縱坐標爲負。經濟統計所處理之數值(value)，大率均在圖中右上方之象限（quadrant）內，以該象限內之縱橫坐標均爲正也。

此種縱橫坐標之概念爲數學之基本原理，亦爲統計學上之基本要點。試以表一數字爲例以說明此法對於表述商業資料之用途。

此種資料可用垂直縱橫坐標圖表示之。沿 x 軸作爲月份，沿 y 軸作爲價格，如圖二。定橫坐標時，以一九二三年一月爲原點，故一九二三年一月之橫坐標之數值（x-value）爲 0，二月之橫坐標爲 1，餘類推。一

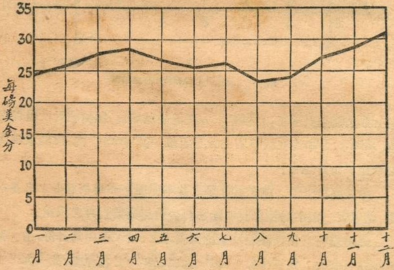
表　一

1923年各月美國棉花價格

（生產者所得之平均價格）

月　份	每磅價格(美金分)
一　月	24.5
二　月	25.9
三　月	27.7
四　月	28.4
五　月	26.9
六　月	25.6
七　月	26.2
八　月	23.5
九　月	24.1
十　月	27.2
十一月	28.8
十二月	31.0

（根據"Weather, Crops and Markets", Dec. 29, 1923）



圖二.　1923年各月美國棉花價格

九二三年一月份棉花價格之橫坐標及縱坐標爲 0，24.5；二月份爲 1，25.9；十二月份爲 11，31.0。如用直綫將各點相連，則對於棉花價格之變動趨勢，觀圖可易於領悟矣。

自變量與倚變量

依縱橫坐標制，任何一點之地位必涉及兩數值，故每點聯合兩個因素而表示其相互之關係。在前例中此兩因素爲月份與棉花價格。棉花價格隨時日之推移而有變動，其斷續綫(broken line)係表明此種變動之方向與大小(magnitude)。時間與價格皆爲變量(variable)，蓋其數量均不固定，而具有變化之特徵。如圖一中橫坐標之值爲 6，縱坐標之值爲 5，皆係固定者；而在圖二中縱橫坐標均不固定，橫坐標變動於 0 與 11 之間，縱坐標變動於 23.5 與 31.0 之間。通常以 x 與 y 兩符號表示此兩種變量，以 x 代表橫軸(horizontal axis)上之變量，y 代表縱軸(vertical axis)上之變量(註)。

在表示棉花價格與時間關係之圖二中，時間之變動係任意擇定一個月爲單位，並假定此時間之因素變動在先，而後測度逐個時間單位間價格變動之情形。此依定量爲增減之變量—時間—謂之自變量(independent variable)，恆繪於 x 軸上；其另一變量，謂之倚變量(dependent variable)，恆繪於 y 軸上。相倚之意義，有眞實的，有形式的。如倚變量之數值可依自變量之數值確定者，則兩數值之關係爲眞實的相倚；如兩者之間並無此種關係，則其相倚不過形式的而已。時間之變量，在製圖時常視爲自變量。

函數關係

(註)英文二十六字母之末數個，通常用爲代表變量之符號，字母之前數個則用爲代表常數之符號。常數者謂在某種情形下其值不變之數量也。

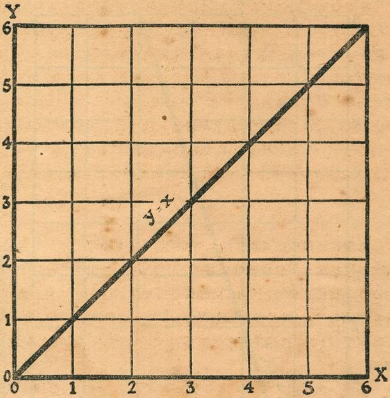
如兩變量之關係完全相倚，即已知 x 值可明確決定 y 值時，y 稱為 x 之函數（function）。此種關係大率以 $y=f(x)$ 之方程式表示之。例如物體由空中下墜時，其在某時間內之速度為下墜時間之函數；一定容積之氣體之壓力為其熱度之函數；依固定利率計息，其本利和之增加為時間之函數。如將自變量之各數值依直綫圖（rectilinear chart）之 x 軸繪之，將其函數（即倚變量）之各數值依 y 軸繪之，即得函數之圖解，成一曲綫（curve）之形狀(註)。此種屬於函數關係（functional relationship）之概念在統計學中甚為重要。茲將較為簡單之各種函數略述之。

直 綫

兩變量之關係如其數值始終相等，則其關係顯為 $y=x$ 之形式。樹齡與樹幹橫斷面所表現輪數之關係，即為其一例。樹木生長二十年時，則樹幹之橫斷面即有二十輪。此種關係可用數對假定之 x 與 y 值，以縱橫坐標圖表示之。將此數點繪於圖中，並繪一綫經過各該點，此綫即為經過原點（假定縱橫軸所用之尺度相同）而等分直角 XOY 之直綫（參看圖三）。

任何一次方程式（即不含 xy 或一次冪以上之 x 或 y 等項者）亦可同樣以直綫表示。普通直綫之方程式均可化為 $y=a+bx$ 之形式。式中 a 為常數（constant），表示原點及直綫與 y 軸交點間之距離；b 亦為常數，表示直綫之斜度（slope，即直綫與橫軸構成角度之正切 tangent of the angle）。常數 a 稱為 y 截綫（y-intercept）。自直綫方程式觀之，x 值為 0 時，y 即等於此常數。前例（圖三）中 a 值為 0，b 值為 1。綫之地位係由 a，b 兩數值之大小及其正負號決定之，故欲決定一直綫，應就所用資料中求 a，b 之數值。此項問題在統計方法中，性質不一，當分別討論之。

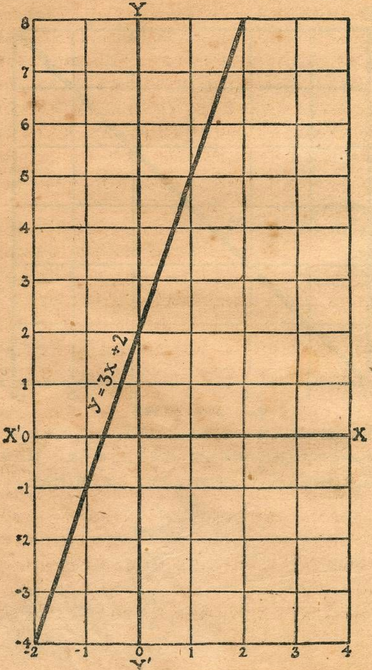
(註)通常所謂"curve"，係指縱橫坐標圖中之任何種綫，或為直綫，或為曲綫。



圖三　方程式 $y=x$ 之圖解

關於上述各點，可繪一次簡單方程式之圖解以說明之。例如吾人欲作函數 $y=2+3x$ 之圖解，則可假定 x 值爲各個不同之數字，以決定各個 y 值。今將假定之各 x 值與由方程式決定之各個 y 值，列表於下：

x	y $(2+3x)$
-4	-10
-2	-4
0	2
2	8
4	14



圖四· 方程式 $y=3x+2$ 之圖解

將表中各值繪於圖上，更將各點接連，卽得圖四所示之綫。此函數既爲直綫關係(卽圖爲一直綫形式)，故若已知任何兩點，卽可確定綫之位置。圖中 y 截綫爲 2，該綫與橫軸構成角度之正切(該綫之斜度)爲 3，卽 x 之係數 (coefficient)。此綫上任何一點之縱橫坐標，均適合於前項方程式，而適合於此式之 x 及 y 值，又必爲綫上之一點，足證此綫爲表示該方程式之綫。又若一變量依定量遞增時，其另一變量亦依定量遞增，則兩者之關係，可用直綫表示之。例如上表內 x 值依定量 2 遞增時，則 y 值依定量 6 遞增。此種依定量遞增之級數，謂之算術級數 (arithmetic series)。

在自然科學方面，兩變量之具有直綫關係者，其例不少概見。在經濟界中，按單利率計息時(卽利息並不併入本金生息)，其本利和之增加，卽爲其一例。設 r 表示單利率，x 表示年數，y 爲 x 年後本金一元所有之本利和，則表示此種關係之方程式爲

$$y = 1 + rx。$$

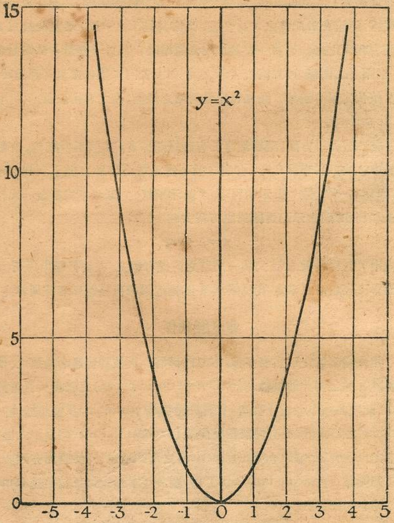
前式中 r 既爲常數，故該式爲一簡單之直綫方程式。在統計學中，表示此種簡單直綫關係之事實，實屬絕無僅有，惟近似於直綫關係者亦所常見。

非直綫關係

非直綫函數之形式有多種，本章內僅將少數較爲普通者論述之。學者對於各種主要非週期性曲綫 (non-periodic curve) 之普通性質應熟審之。此項曲綫中最爲重要者，有屬於拋物綫(parabola)及雙曲綫(hyperbola)之形式者，亦有屬於指數曲綫 (exponential curve) 之形式者；其中尤以遞升冪級數 (potential series) 之形式最爲普通，應用亦較廣。在週期函數 (periodic functions) 中，關於基本形式之正弦曲綫 (sine curve) 亦將簡述之。

拋物綫或雙曲綫形式之函數關係，在自然科學中甚爲普通，此種曲

綫在經濟資料中亦常可採用之。其普通方程式爲 $y=ax^b$。該式中指數 b 之符號爲正時，此綫爲拋物綫；符號爲負時，此綫爲雙曲綫。茲舉例題兩則以說明其形式如下：



圖五. 拋物綫：方程式 $y=x^2$ 之圖解

例題：試作一函數 $y=x^2$ 之圖解。

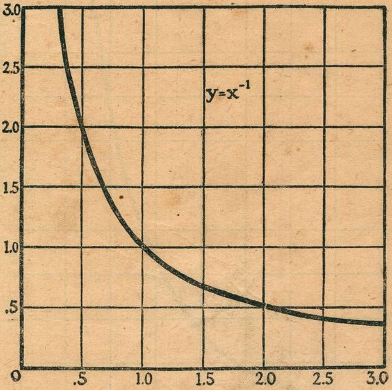
x	-5	-4	-3	-2	-1	0	1	2	3	4	5
$y(=x^2)$	25	16	9	4	1	0	1	4	9	16	25

依題製成之圖解如圖五。

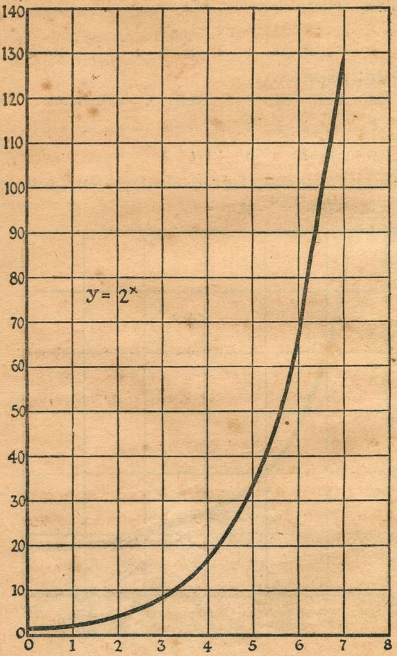
例題：試作一函數 $y=x^{-1}$ 之圖解，式中 x 值假定為正數。

x	$\frac{1}{3}$	$\frac{1}{2}$	1	2	3	4	5
$y(=x^{-1})$	3	2	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$

表示此函數之圖解為等邊雙曲綫（equilateral hyperbola），如圖六所示。此式亦可寫作 $y=\frac{1}{x}$，或 $xy=1$。



圖六． 等邊雙曲綫：方程式 $y=x^{-1}$ 之圖解（x 值爲正數）



圖七．指數曲線：方程式 $y=2^x$ 之圖解（x 值爲正數）

在此種方程式中，x 之值如按幾何級數 (geometric progression) 增加，則 y 之值亦按幾何級數增加。如在前述拋物綫之一例中($y=x^2$)，假定各 x 值成一幾何級數 (geometric series)(註)，則由此式求得相對於各 x 值之 y 值，亦具有幾何級數之形式。

x	1	2	4	8	16	32
y	1	4	16	64	256	1024

此外另有一種函數，其形式可以方程式 $y=ab^x$ 代表之。在此式中，其一變量爲指數 (exponent)，故表示此式之圖解，稱爲指數曲綫(exponential curve)。茲舉例說明於下：

例題：試作一函數 $y=2^x$ 之圖解，式中 x 值爲正數。

x	0	1	2	3	4	5	6
$y(=2^x)$	1	2	4	8	16	32	64

依題製成之圖解，如圖七所示。

兩變量之相互關係，如各按定量遞增 (組成算術級數 arithmetic series)時，可以直綫表示之；如各按幾何級數遞增時，可以拋物綫或雙曲綫表示之，已如前述。指數曲綫則爲前兩種之混合形式，因其所表示兩變量之相互關係，其一變量按算術級數遞增時，另一變量則按幾何級數遞增也。上列之數字卽係表示此種關係。

表示下列關係之曲綫，在統計研究方面應用頗廣。

$$y=a+bx+cx^2+dx^3+\cdots\cdots$$

所謂遞升冪級數(potential series)者，卽指此種形式之方程式而言。此種曲綫雖非圓錐截面形之拋物綫，但如該曲綫方程式右邊之行列係至 x 之二次冪爲止者，則此曲綫通常亦稱爲二次拋物綫 (a second degree parabola)；如至 x 之三次冪爲止者，稱爲三次拋物綫 (a third degree parabola)，餘類推。按此種級數並無劃一或單純之形式，後文當詳論之。

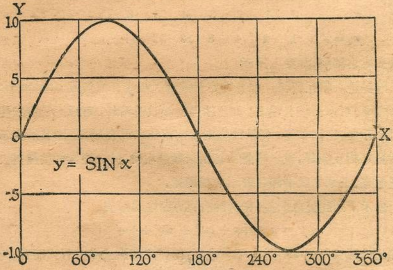
(註)某一級數如其中每項數字，係可由前一項數字乘以常數得之者，謂之幾何級數。

週期函數（periodic functions）爲另一種曲綫，常用以表示電學與氣象學中之各種關係者，惟其範圍並不以此爲限。此種關係之特徵，厥爲在自變量變動之固定時期內，倚變量之數值重覆不變。玆將代表週期函數基本形式之正弦曲綫（sine curve）舉例說明於下：

例題：試作一函數 $y=\sin x$ 之圖解。

x（角度）	0°	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°	360°	390°……
$y(=\sin x)$	.000	.500	.866	1.000	.866	.500	.000	−.500	−.866	−1.000	−.866	−.500	.000	.500……

依題製成之圖解，如圖八所示。



圖八. 正弦曲綫：方程式 $y=\sin x$ 之圖解

求數學方程式以表示兩變量之相互關係，對於統計工作上之重要性，在未詳論本題之前，尙不能有所表現，惟其基本目的在於根據觀察所得之現象，以決定自然或經濟的法則。其另一種更較實用之目的，則在求一公式，俾藉此公式，由已知一變量之值，以推求另一變量之值。這

到此種目的之方法，在本書各節中將詳論之(註)。

對數

對數 (logarithms) 之爲用，在普通數學演算方面，固占重要地位，而於統計資料之處理，其重要程度正復相等。茲將關於對數之性質及其便利計算手續之方法簡述之，惟本文所論述者，祇限於以 10 爲基數 (base) 之普通對數。

任何正數均可以 10 之羃數 (power) 表示之。如

$$1,000=10\times10\times10=10^3$$

$$10,000=10\times10\times10\times10=10^4$$

前式中 10 之羃數（即 10 之右上角較小之數字），表示因數 10 之自乘次數。眞數化爲 10 之自乘次數如成爲整數時，其羃數亦爲整數，其他數字則其羃數必含分數矣。如 100 等於 10 之二次羃，即 10^2；110 等於10之 2.04139 次羃，即 $10^{2.04139}$。

此 10 之羃數，即係將 10 提升至與某數相等時乘羃之指數，稱爲某數之對數。如 100 之對數爲 2；110 之對數爲 2.04139；998 之對數爲 2.99913 等是。此項對數均以 10 爲基數，惟對數法中任何數字均可作爲基數。通常如

$$a=b^c$$

則 $$\log_b a=c$$

此式可讀如"以 b 爲基數之 a 的對數爲 c"。在用普通對數時，眞數、基數、與對數三者之關係，可還下式記憶之。

$$100=10^2$$

$$\log_{10}100=2$$

一數之對數，常具整數與小數兩部分，其整數稱爲首數(character-

(註)關於各種曲線之形式，當於下文時間數列之分析一節中詳論之。

istic)，其小數稱爲尾數 (mantissa)。在求對數時，首數可由觀察斷定之，尾數則須查閱對數表得之。首數須視眞數之小數點位次而有變更，尾數則否，凡數字排列相同之眞數，不論其小數點位次之先後，其尾數常屬相同。此點可以下列數字舉證之。

$$\log 8450 = 3.92686$$
$$\log 845 = 2.92686$$
$$\log 84.5 = 1.92686$$
$$\log 8.45 = .92686$$
$$\log .845 = 9.92686-10$$
$$\log .0845 = 8.92686-10$$

由對數求眞數時(此眞數 natural mumber 謂之反對數 anti-logarithm)，可由尾數决定各位數字之排列，由首數决定小數點之位置。例如欲求 2.17609 之反對數，可將其小數點以後之尾數 .17609，查閱對數表得相對於該尾數之眞數爲 1500。因其首數爲 2，故眞數必在 100 至 1,000 之間，而可斷定此數爲 150。

下列各項數字，爲 10 之乘羃及相對於各乘羃之眞數，讀者略加研究，即可記憶 10 之倍數與其對數之關係，則求對數時其首數自易决定矣。

.0001	.001	.01	.1	1	10	100	1,000	10,000
10^{-4}	10^{-3}	10^{-2}	10^{-1}	10^{0}	10^{1}	10^{2}	10^{3}	10^{4}

上列兩行數字中，其下行內 10 之羃數，即係上行內各數之對數。

自 0 起至 1 止之一切數字，其對數皆爲負數；例如 .845 之對數爲 $-1+.92686$，可寫作$9.92686-10$。自0起至無限大數間之一切正數，其對數由負而正，包括一切正負數值。由是知負數旣無正對數，亦無負對數。

以 10 之羃數表示各種數字，其優點在於乘，除，求乘羃，或開方根時，可使計算手續便捷不少。

求乘積用對數相加　各因數的對數之和，等於其乘積之對數。

其公式爲：

$$a^{b}\times a^{c}=a^{(b+c)}$$

例如：

$$10^{2}\times10^{3}=(10\times10)\times(10\times10\times10)=10^{5}=100{,}000$$
$$100\times1{,}000=100{,}000$$

求商數用對數相減　如以甲數除乙數，可用乙數之對數，減去甲數之對數，即得所求商數之對數。

其公式爲：

$$a^{b}\div a^{c}=a^{(b-c)}$$

例如：

$$10^{5}\div10^{2}=\frac{10\times10\times10\times10\times10}{10\times10}=10^{3}=1{,}000$$
$$100{,}000\div100=1{,}000$$

求乘冪用對數相乘　欲求某數之乘冪，可將該數之對數乘其冪數，其乘積即爲所求乘冪之對數。

其公式爲：

$$(a^{b})^{c}=a^{bc}$$

例如：

$$(10^{3})^{2}=(10\times10\times10)\times(10\times10\times10)=10^{6}=1{,}000{,}000$$
$$1000^{2}=1{,}000{,}000$$

開方根用對數相除　欲開某數之方根，可將該數之對數除以方根之指數，其商數即爲所求方根之對數。

其公式爲：

$$\sqrt[b]{a^{c}}=a^{\left(\frac{c}{b}\right)}$$

例如：

$$\sqrt[3]{10^{6}}=10^{\frac{6}{3}}=10^{2}=100$$
$$\sqrt[3]{1{,}000{,}000}=100$$

茲將各式綜述於下:

$$\log(a\times b)=\log a+\log b$$

$$\log(a\div b)=\log a-\log b$$

$$\log a^{b}\quad=b\times\log a$$

$$\log\sqrt[b]{a}\quad=\log a\div b$$

計算尺(slide rule)爲利用對數之優點所製成之工具,所以便利計算手續者也。研究統計學者,應明瞭其使用法。

對數方程式

採用垂直縱橫坐標制作圖之方法及其優點,已如前述。作圖時有須用對數以代替眞數者,蓋以對數作圖常可使各種重要關係明白表現,而又可將處理資料之手續化繁爲簡;且利用對數時,可將複雜之曲綫化爲直綫,在處理與解釋方面較爲簡易,其所獲之利便爲尤著焉。

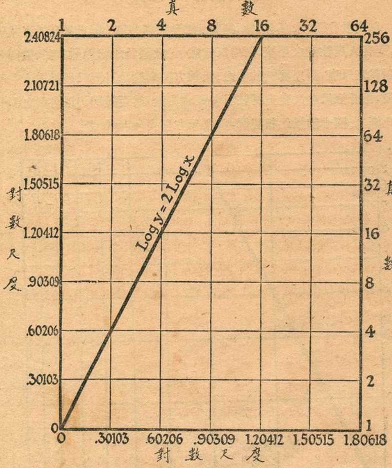
直綫方程式之普通形式爲 $y=a+bx$, a 與 b 皆爲常數,爲表示某綫與縱軸相交之截綫與其斜度。用對數將方程式化簡之法,係將 x 或 y,或同時將 x 與 y 之眞數各代以 $\log x$ 或 $\log y$,如是則可將高次方程式化爲較簡之形式。

此法可以方程式 $y=x^2$ 說明之。如以此式繪入垂直縱橫坐標圖中,卽得拋物綫式之曲綫(參看圖五);若將此式化爲對數之形式,卽成爲 $\log y=2\log x$。此用 $\log y$ 與 $\log x$ 爲變量之方程式,卽屬於直綫之形式。假定 $\log x$ 之值爲正,則所得圖解如圖九。

爲便於表示此中關係起見,可同時將相對於各對數之眞數,分別註明於圖之右方及上方之尺度上。眞數尺度上之各數成爲幾何級數,其對數則爲算術級數。惟須注意者,圖上相等之縱距離在對數尺度上係表示相等之絕對增量,而在眞數尺度上則係表示相等之增加百分率。

方程式 $y=5x^3$ 可依此化爲直綫形式,如 $\log y=\log 5+3\log x$。其

他屬於 $y=ax^b$ 形式之一切方程式，即一切單純拋物綫與雙曲綫，亦皆可同樣化爲直綫形式，如 $\log y=\log a+b\log x$。在圖解方面，此即以 y 之對數與 x 之對數作圖也。



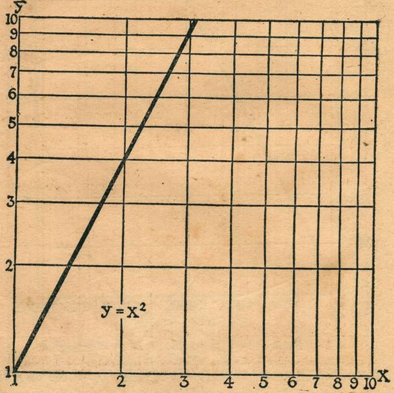
圖九．　方程式 $\log y=2\log x$ 之圖解（方程式 $y=x^2$ 之對數形式）

至於指數曲綫 $y=ab^x$ 形式之方程式，其變化方式與前式略異。此式以對數形式表示之爲 $\log y=\log a+x\log b$，亦屬直綫形式，其常數爲

$\log a$ 與 $\log b$，其變量爲 x 與 $\log y$。依此方程式將 x 之眞數與 y 之對數作圖，即得一直綫。茲將此種形式之曲綫說明於下。

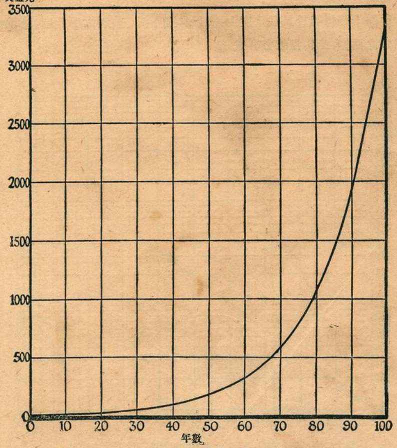
雙對數與單對數圖

以對數作圖，未必有利而無弊。如點數過多時，則查閱對數必甚煩瑣，而吾人所注意之原值，圖中反無載明。欲避免此種困難，可於製圖時用對數定尺度，而用眞數註明之。此種方法與製造計算尺所採取者完全相同，蓋計算尺在尺度上所註明者雖爲眞數，其距離則仍以各數之對數爲比例也。圖十即爲此種圖解，係表示方程式 $y=x^2$ 者。



圖十. 方程式 $y=x^2$ 之圖解（繪於雙對數尺度之圖紙上者）

如將此種圖解之形式略加變化，以眞數爲橫軸之尺度，以對數爲縱軸之尺度，即成統計工作上重要之圖解。在此種混合尺度上所作之圖，自與按照 x 之眞數與 y 之對數所繪製者相同。上述指數形式之曲綫化爲直綫時，即須用混合尺度之圖制繪之。此種單對數式(或稱比率式)之圖紙，可按照計算尺或對數表繪製，即製就之圖紙亦可購得。在經濟統計製圖方面，此種圖制尤有特殊之價值，蓋在經濟統計中，時間每爲變



圖十一． 複利法則：本金十元、複利六厘計息、百年間之本利和(繪於算術尺度上者)

量之一，而此變量自以採用真數尺度繪製，較爲適當也。

表示複利法則之方程式，卽屬此類曲綫。設以 r 表示利率，p 表示本金，y 表示 x 年後之本利和（依複利計息，每年結息一次），則所得方程式應爲：

$$y=p(1+r)^x$$

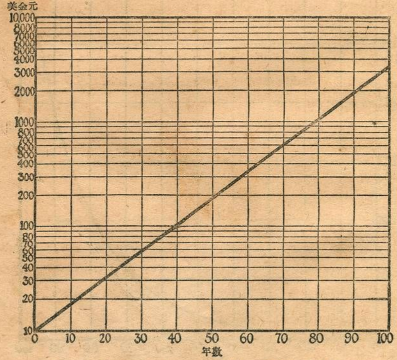
以對數形式表示之，其式爲：

$$\log y=\log p+x\log(1+r)$$

卽變爲直綫之方程式矣。

圖十一係繪於眞數尺度上，表示本金十元、複利六厘計息、本利和增加之曲綫。此爲指數方程式

$$y=10(1+.06)^x$$



圖十二 複利法則：本金十元、複利六厘計息、百年間之本利和

（繪於單對數尺度，卽比率尺度上者）

[illegible]

之圖解，y 表示第 x 年年末之本利和。圖十二係用同一資料繪於單對數尺度之圖紙上者，其指數曲綫已化爲直綫。

單對數圖紙 (semi-logarithmic paper) 之用途，並不限於由指數曲綫化爲直綫，蓋有多種資料如用單對數圖紙作圖，其意義最能明白顯示也。其優點將於下文詳述之。

統計圖之功用

根據觀察或統計調查所得之數字結果，加以分析及解釋時，應先作圖，以明其大概，此爲其初步工作之一。此種手續不特具有科學上之價值，可爲更進一步研究之張本，且各項數字可賴圖示而益明瞭；蓋僅依邏輯之順序以研究事物，失之抽象，而藉具體之圖形以刺戟視官，則爲達到理解與意想之捷徑。例如吾人欲解釋一行列之數字，或爲困難之事，而將同一資料示以圖形，則其意義可由圖上窺見一斑。此圖示法之所以不特在實驗室與製圖室中占極重要之地位，即在日常商業活動上，其重要程度亦復相若也。

至於現時工程師與統計家所用之圖，種類繁多，因不在本書範圍以內，故不備述；惟關於圖示法較重要之原理，應略加詮釋，日常通用之各種主要圖形，亦將予以例示。其他例證，當於本書後數章內舉示之。

選擇圖形之條件

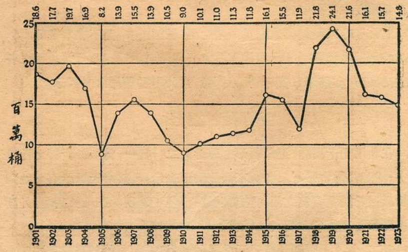
在應用圖示法時，選擇圖形之條件有二：其一，爲關於資料之性質者。一種資料通常雖可用各種不同之圖形表示，然其中大抵祇有一種圖形爲最適於該項資料之用，亦有某種圖形絕不適用於某種資料者，是則選擇圖形時應先參酌資料之性質。

其二，爲關於圖解所欲表示之目的者，則較第一條件更爲重要。普通所用之各種圖形各有其特殊之功用。圖解之目的，或爲表明某種特

徵，或爲揭示某種關係，未有一種圖形能合於各種目的者，故在作圖目的未曾確定以前，選擇圖形殆無標準可言。茲將少數標準圖形縷述於下，以供選擇圖形之參考。

表示時間數列之圖解

時間數列之圖解，其注重之點在於表示各種數值在時間上之變動與其長期趨勢，以及長期趨勢綫兩邊數值之變動情形。如圖示之目的注重於絕對數值之變動(absolute variation)，即注重於各值實際單位之變化時，則用一簡單圖形如圖十三者，即可達此目的。此圖爲一九〇一年至一九二三年間美國麵粉每年對外輸出數量之圖解。圖之縱橫軸均係用算術尺度。圖中所示各點表示各年之數量，惟爲便於解釋起見，各點之間均用直綫連之。此圖可說明逐年變動之簡單情形，如一九〇七年至一九一〇年之向下趨勢，一九一一年至一九一九年之向上趨勢，及一九一九年後之下降趨勢，均可於圖中見之。關於製圖方法，應注意下列各點。



圖十三.　1901—1923年美國麵粉輸出數量

1. 圖之標題應載明圖中所示之資料及其包含之時期。

2. 縱尺度應自零綫 (zero line) 起，使讀者對於變量之大小，可得一明確之印象。

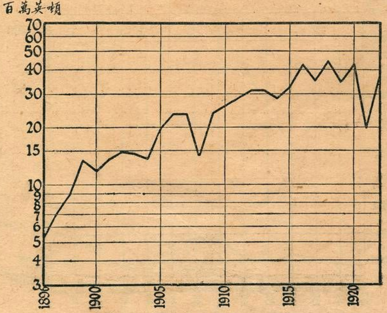
3. 零綫及連接各點之綫，應較縱橫尺度之綫稍粗，以求醒目。

4. 表示尺度之數字，應置於圖之左邊及下邊，惟縱尺度可重複置於圖之右邊，以求讀者之便利。數字之排列，應自下而上，以基綫作底綫；或自左而右，以圖之右邊綫作底綫。

5. 各點之 y 值可分別註明於圖頂。此種辦法雖有裨助，但非必要，因圖外常附有載明數字之表也。

比率圖之優點

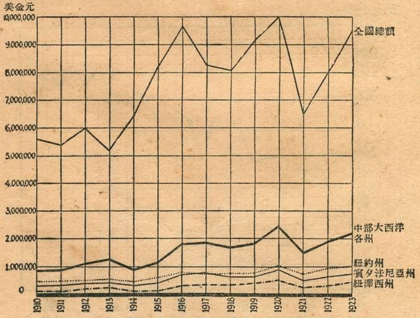
作圖之目的，如不欲注重絕對變動 (absolute variation)，而注重



圖十四.　1896—1922年美國鋼鐵產量(繪於單對數尺度上者)

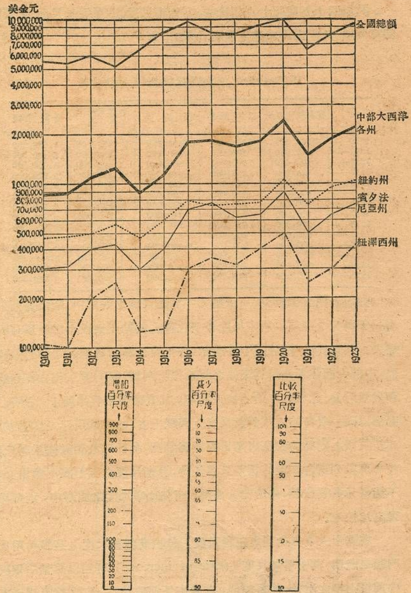
於相對變動 (relative variation)時，則所用之圖以取 y 軸用對數尺度，x 軸用算術尺度之單對數圖形最爲相宜。在普通算術尺度之圖形中，相等之縱距離表示相等之絕對增減量；而在單對數之圖形中，則相等之縱距離表示相等之變動百分率。吾人所以用單對數圖 (semi-logarithmic chart，或稱比率圖 ratio chart)表示時間數列者，因數量之變動必須與其所由變動之基數相比，始有意義；例如由基數 100 增加 100，與由 10,000 增加 10,000，其重要程度彼此相等，蓋兩數各增一倍也。然依絕對增量而言，則後者適百倍於前者，用算術尺度所表示兩者之變動，亦適成此比例。若用單對數尺度表示，則此兩種變動之重要程度彼此相等。

此種單對數圖形如圖十四所示。該圖爲自一八九六年至一九二二年美國鋼鐵產量之圖解。繪曲綫時雖用絕對數量，然因其縱軸爲比率尺度，故該圖所表示逐年產量之變動，與其變動百分率成正比例。



圖十五. 1910—1923年美國亞根公司之全國銷貨總額及某數州之銷貨額

(繪於算術尺度上者)



圖十六. 1910—1923年美國亞根公司之全國銷貨總額及某數州之銷貨額附增加、減少及比較之尺度(繪於單對數圖紙上者)

試以圖十五與圖十六相比較，則比率尺度(卽單對數尺度)之優點，卽可判明。此兩圖所表述之資料，列如下表：

表 二

1910—1923 年美國亞根公司之銷貨額

	賓夕法尼亞州	紐澤西州	紐約州	中部大西洋各州	全國總額
1910……	$305,000	$105,000	$465,000	$ 875,000	$5,600,000
1911……	310,000	100,000	480,000	890,000	5,400,000
1912……	400,000	200,000	500,000	1,100,000	6,000,000
1913……	425,000	250,000	575,000	1,250,000	5,200,000
1914……	300,000	125,000	465,000	890,000	6,400,000
1915……	400,000	130,000	600,000	1,130,000	8,200,000
1916……	700,000	300,000	800,000	1,800,000	9,700,000
1917……	760,000	350,000	740,000	1,850,000	8,300,000
1918……	630,000	320,000	750,000	1,700,000	8,100,000
1919……	650,000	400,000	775,000	1,825,000	9,200,000
1920……	900,000	500,000	1,050,000	2,450,000	10,000,000
1921……	500,000	250,000	750,000	1,500,000	6,500,000
1922……	650,000	300,000	950,000	1,900,000	8,000,000
1923……	750,000	425,000	1,025,000	2,200,000	9,500,000

如將表內五個數列，一併繪入一算術尺度之圖形中，則因數列中旣有一千萬元之一項，所定尺度勢須收縮，其結果則較小之數量在圖中遂失去其重要性。觀於圖十五所示，各年間全國銷貨總額之變動似甚劇烈，中部大西洋各州銷貨總額之變動似較和緩，紐約等三州之各州銷貨額幾無變動。此種圖形最易使人誤解，其實際情形當如圖十六所示。該圖係以同一資料繪入單對數尺度之圖紙上者。依該圖所示，則紐約等三州銷貨額之變動均較全國銷貨總額爲劇烈，故欲以數個數列互爲比較，而各個數列所含之數值相差又甚鉅時，算術尺度之圖形頗不適用，以其不能表示眞實意義，徒足使人曲解，而在比率尺度之圖形中，則可作合理之比較也。

在圖十六下方所刊之三種尺度，係欲將對數尺度之某種功用爲更明顯之說明。其中增加尺度 (scale of increase) 可用以測量一個數列內前後兩期之增量。前嘗說明，在圖中任何地位，一定之縱距離係表示一定之增加百分率，例如沿縱尺度自十萬元至二十萬元之距離，等於

自二百萬元至四百萬元之距離。吾人可量任何兩點之縱距離，即將此距離合於增加尺度上，由下至上讀之，卽得增加百分率。例如欲定紐澤西州一九一三年銷貨額較一九一二年之增加百分率，可先量此兩年間之縱距離，將此縱距離合於增加尺度上可得25%之增加率。

減少尺度之用法與增加尺度相似。吾人可量任何兩點之縱距離，將此距離合於減少尺度上，由上至下讀之，卽得減少百分率。各尺度上所繪之箭形，所以表示讀尺度時之方向也。

比較尺度係用以決定在任何時期兩個數列之百分比者。例如吾人欲知一九一二年中部大西洋各州之銷貨總額占同年全國銷貨總額之百分比時，則可量此兩點間之縱距離，將此縱距離合於比較尺度上，由上至下讀之，可得此百分比約爲18%。

繪上述三種尺度時可將尺度上應用之各數，依比率尺度繪之。吾人如欲採用標準形式之單對數圖紙，作若干圖解，可用預製之尺，以供繪製此種尺度之用，則尤爲便利矣。

用單對數尺度作圖，其主要優點略如下述：

1. 指數形式之曲綫，如繪於單對數圖上，卽成一直綫。例如依複利計息所有本利和之增加，如繪入此圖中，卽爲一直綫。
2. 任何數列中，如其增加率或減少率固定不變，則在單對數尺度之圖中爲一直綫。
3. 在此種圖形中，表示相等的變動百分率之數綫，其斜度必相等，故兩種數列如其增加率或減少率彼此相等，則代表此種數列之兩綫彼此平行。
4. 比較兩個或數個數列之變動率時，可由比較各綫之斜度得之。
5. 單對數尺度之圖形，既可用絕對數値作圖，同時又可作相對變動之比較。
6. 在比較若干數列，而各數列所含之數量相差又極鉅時，則繪於單

對數圖上較爲相宜。

7.變動百分率及各項數值間之百分比，皆可直接由圖中得之

比較各組次數之圖解

圖示之目的，如爲比較各組次數(frequency)，則當用另一種圖解以表示之。次數者，卽各類事物發現之次數也。試以下列數字爲例說明之。

表　三

1920年新英格蘭各州農場數

州　　別	農場數
緬因(Maine)	48,227
紐罕普什爾(New Hampshire)	20,523
弗蒙特(Vermont)	29,075
馬薩諸塞(Massachusetts)	32,001
羅得島(Rhode Island)	4,083
康涅狄格(Connecticut)	22,655

如欲比較上述六州一九二〇年之農場數，可用圖十七所示之條形

圖十七. 1920年新英格蘭各州農場數

圖(bar diagram)表示之。此種圖形雖簡，然頗合此項比較之用。

此種表示次數分配 (frequency distribution) 之圖形，容於下章內再行例示，並將說明處理某種資料時，由單純條形圖化爲次數多邊圖 (frequency polygon) 或次數曲綫 (frequency curve) 之方法。此種次數曲綫乃圖形中之極重要者，另於後文內詳述之。

表示成分之圖解

在製圖時，有須將總量依其成分 (component parts) 各別圖示者，如是則吾人不僅可明瞭總量之變動，且可明瞭各個成分之變動。下表數字爲屬於此類資料之一簡例。

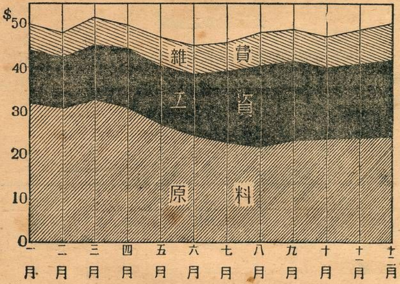
表　四

1923年某公司之生產費

(此項數字係表示製造每一單位之物品所需之各項生產費及其總數)

月份	原料費	工資	雜費	生產費總計
一月	$32.00	$12.00	$6.00	$50.00
二月	31.00	11.00	6.00	48.00
三月	33.00	12.00	6.50	51.50
四月	31.00	13.00	6.00	50.00
五月	27.00	13.00	7.00	47.00
六月	24.50	13.50	7.00	45.00
七月	23.00	16.00	7.00	46.00
八月	22.00	18.50	8.00	48.50
九月	23.50	18.00	7.50	49.00
十月	24.00	16.00	7.50	47.50
十一月	24.00	17.00	8.00	49.00
十二月	24.50	17.50	8.50	50.50

前項數字繪入成分圖中，當如圖十八所示。由圖可知各月生產費總數之變動較爲穩定，但其各個成分頗有變動甚烈者。是以吾人但知生產費之總數，猶未能窺其全豹，而須同時明瞭各項費用之變動。此種圖形不特表示總數之變動，且表示其各個成分之變動。



圖十八． 1923年各月某公司生產費用之分析

5.

累積圖

吾人於展開一個數列時，往往不注重於數列中各個別項目之數值，而注意於其累積之數值。吾人在擬定每年生產數量之預算時，卽有如此情形。在此種情形之下，吾人欲知實際生產量之總數與預計生產量總數之比 此則須藉圖形以資比較。下列數字卽屬於此類資料而可適用此種圖形者。

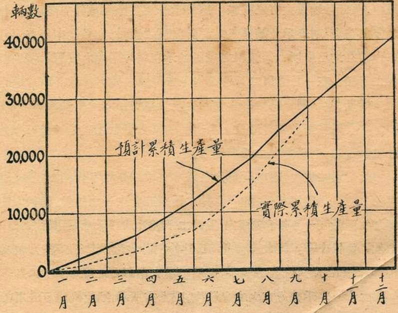
表 五

1924年美國斯彼德韋爾汽車製造公司預計累積生產量與實際累積生產量之比較

月份	預計生產量（輛數）	預計累積生產量（輛數）	實際生產量（輛數）	實際累積生產量（輛數）
一月	2,000	2,000	750	750
二月	2,000	4,000	1,250	2,000
三月	2,000	6,000	1,250	3,250
四月	3,000	9,000	2,000	5,250
五月	3,000	12,000	1,500	6,750
六月	3,750	15,750	3,750	10,500
七月	3,750	19,500	4,250	14,750
八月	5,000	24,500	6,000	20,750
九月	4,000	28,500	5,750	26,500
十月	4,000	32,500		
十一月	4,000	36,500		
十二月	4,000	40,500		

上表係假定爲表示九月末之生產狀態者。

圖十九中之兩曲綫，係累積次數曲綫 (cumulative frequency



圖十九. 1924年美國斯彼德韋爾汽車製造公司預計累積生產量與實際累積生產量之比較

curve)。觀圖可知每月月末之實際生產量與預計生產量之關係，而實際生產量低於預計生產量者約爲若干，亦可自尺度上推定之．按統計圖表相依爲用，故兩者恆同時並列，閱者欲知差異之確實數量，可參閱附表中數字。此種累積次數圖應用頗廣，容於下章內說明之．

甘悌氏進行圖

前表內之統計資料，亦可應用甘悌氏（H. L. Gantt）所發明之圖形以表示之。關於此圖之構造及其種種用途，以本書篇幅無多，未能詳述，玆僅將其特點略加說明於下。

生產量預計表一經編竣，卽可用甘悌氏進行圖（Gantt progress chart）以核對實際生產數量與預計生產數量之差別。例如表五內所示之生產量預計表擬就後，可將每月與每年之預計數量錄入如圖二十之圖形中。在每月地位左方之登錄，係表示各該月份之預計生產數量；在其右方之登錄，則表示各該月末之預計累積生產量。在該圖中年首兩月之工作結果均已表明。其較粗之綫係表示在該時期內之實際累積生產量爲二千輛。在一月份與二月份欄內上方較細之綫係表示各該月份之實際生產量。倘該兩月份之每月實際生產量適與預計生產量相等，則其較細之綫當延長至各該月地位之全部。如某月實際生產量超過該月之預計生產量時，則用細綫二條表示之。

於此所應注意者，圖中所劃分之每月地位，雖代表相等之時間，然所代表之生產量則不相同。如一月份全距離五分之一之地位，係代表四百輛之生產量（一月份預計生產量爲二千輛）；八月份全距離五分之一之地位，係代表一千輛之生產量（八月份預計生產量爲五千輛）。閱此圖時，如欲知生產量之實際數字，則須參看表中每月之數額。

圖二十一表示九月末之生產狀態。圖上方所示之箭形，表示工作已完之時日。此時之實際累積生產量較預計生產量落後半閱月，此可由圖

圖二十. 預計生產量與實際生產量之比較：甘梯氏進行圖（表示二月二十八日爲止之生產狀態）

圖二十一. 預計生產量與實際生產量之比較：甘梯氏進行圖（表示九月三十日爲止之生產狀態）

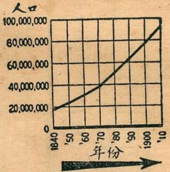
中箭形與粗綫之比較而得之。七、八、九、三個月生產量之細綫，係表明各該月實際生產量已超過預計生產量之情形。

甘悌氏圖在政府機關與商業組織方面爲用頗廣。用此種圖形能將各地或各部分之進展狀況併列於一圖，故地位頗爲經濟。此圖爲表示工作進行狀態及比較實際工作與預計工作時最簡單最有效之圖示法，故應用時頗可促進行政管理之效率焉。

製圖之標準規則

在自然科學與社會科學方面以及企業界中，圖示法之應用已極普遍，惟因其用途廣泛，作圖之例，極不一致。爲補救此項缺憾起見，在美國嘗有標準圖示法討論會之召集，對此論題有關之各團體均派代表出席。該會卒編就一報告書，介紹各種標準圖示法以供採用。是項報告係一種簡明圖示之範本，嘗發表於美國各種刊物內，惜未能廣爲流傳耳。該會所提出之建議如下(註)：

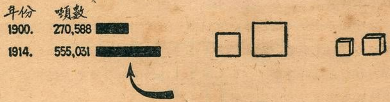
1 圖之普通排列應自左而右。



圖例一

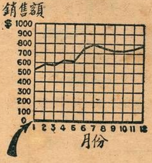
(註)此項報告書係 The Joint Committee on Standards for Graphic Representation 所編製者，該會主席爲 Willard C. Brinton氏。全文曾刊載於一九一五年出版之 "Quarterly Publications of the American Statistical Association" 第14卷第790—797頁。本書之轉載，事前嘗得 Brinton 氏之同意。紐約 American Society of Mechanical Engineers 備有印本，供各界參考之用。

2. 數量宜用綫之長度表示，因面積圖與體積圖最易引起誤解也。



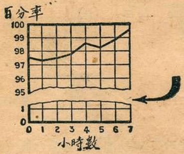
圖例二

3. 在選定綫圖之縱尺度時，宜將零綫繪入圖中。



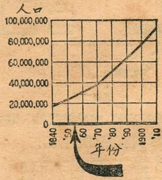
圖例三

4. 如綫圖中縱尺度上之零綫，不易繪入，可將該圖繪成橫截狀，使零綫仍能繪入圖中。

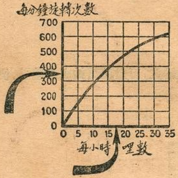


圖例四

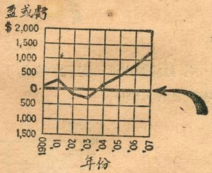
5. 尺度上之零綫須較行格為粗。



圖例五(甲)

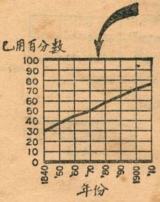


圖例五(乙)

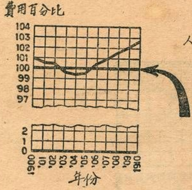


圖例五(丙)

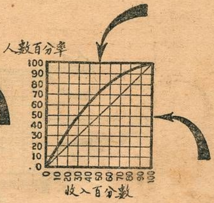
6. 綫圖尺度如表示百分數，則代表百分之綫或其他用作比較標準之綫，例須較粗，俾便識別。



圖例六(甲)

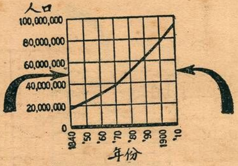


圖例六(乙)



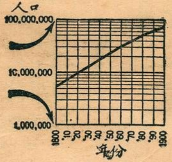
圖例六(丙)

7. 如圖之尺度表示時期，但所表示之時期並不完全時，則該尺度之起訖兩邊綫，不必加粗，因此兩邊綫實際並不表示時間之起訖也.



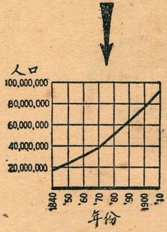
圖例七

8. 曲綫圖紙如用對數尺度，則其上下邊綫應取10之自乘數。

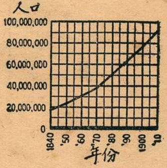


圖例八

9. 圖中行格之疎密，須求適當，不宜過密，以免眉目不清。

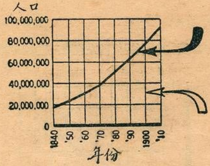


圖例九(甲)



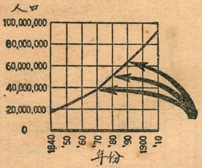
圖例九(乙)

10. 圖中曲綫應較粗於行格。

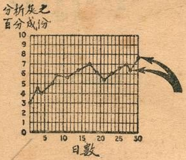


圖例十

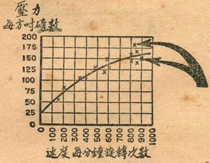
11. 如一曲綫係表示數個觀察數値時，應將圖中代表每個觀察之各點分別標明之。



圖例十一(甲)

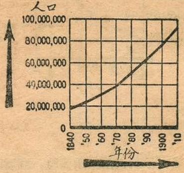


圖例十一(乙)



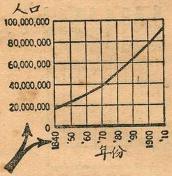
圖例十一(丙)

12. 綫圖之橫尺度，其讀法應自左而右，縱尺度應自下而上。

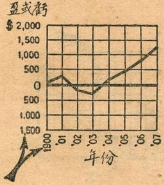


圖例十二

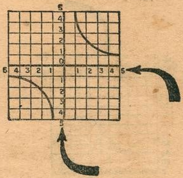
13. 圖中表示尺度之數字,須置於圖左及圖下,或沿縱橫軸排列。



圖例十三(甲)

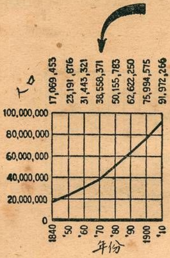


圖例十三(乙)

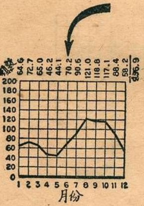


圖例十三(丙)

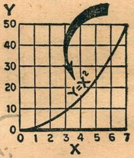
14. 曲綫所代表之數字資料或其公式,有時亦須列入圖中。



圖例十四(甲)

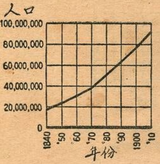


圖例十四(乙)



圖例十四(丙)

15. 如圖中不便列入數字時，應將此項數字製爲表式，與圖並列。

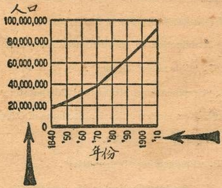


年份	人口
1840	17,069,453
1850	23,191,876
1860	31,443,321
1870	38,558,371
1880	50,155,783
1890	62,622,250
1900	75,994,575
1910	91,972,265

圖例十五

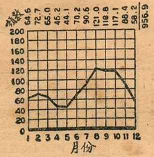
16．圖上所有一切標識及數字，務須置於適當地位，使閱者以圖之基綫作底，或以圖之右邊綫作底閱看，不感困難。

左



圖例十六

17．圖之標題應力求簡明完備，遇必要時，可用副標題或附註以補充其意義。



1914年各月第二廠之鋁塊產量

產量以公噸計，

碎塊鋁之銷售額不在內。

圖例十七

參考書

1. 關於數學概念者

關於圖示法與對數用法之基本數學概念，在代數學與解析幾何學之任何種標準教科書中均有述及。在下列各書內論述頗詳，足供研究統計學者之參考。

Griffin, F. L. "Introduction to Mathematical Analysis.'

Karsten, Karl G. "Charts and Graphs".

Lipka, Joseph. "Graphical and Mechanical Computation".

Mellor, J. W. "Higher Mathematics for Students of Chemistry and Physics".

Schultze, Arthur. "Graphic Algebra".

Steinmetz, C. P. "Engineering Mathematics".

Whitehead, A. N. ' An Introduction to Mathematics".

2. 關於圖示法者

Brinton, W. C. "Graphic Methods for Presenting Facts".

Clark, Wallace. "The Gantt Chart".

參考書(續)

Field, J. H. "Some Advantages of the Logarithmic Scale in Statistical Diagrams". Journal of Political Economy, October, 1917.

Fisher, Irving. "The 'Ratio' Chart". Quarterly Publications of the American Statistical Association, June, 1917.

Maskell, A. C. "Graphic Charts in Business".

Maskell, A. C. "How to Make and Use Graphic Charts".

Karsten, Karl. "Charts and Graphs".

(上列各書之出版書局及年月，詳載於書末之圖書目錄中。)

第三章　統計資料之整理：次數分配

研究商業與經濟之統計學家，其所任之工作，為整理、分析與解釋關於商業事務與經濟狀況之數量資料。此外尚須加以蒐集資料之工作，惟通常所用之資料每由已經採集之原始資料(primary data)或次級資料(secondary data)編製而得者。

惟時間數列（time series）及其他與時間無關之資料，其所用分析方法並不相同。研究時間數列之主要目的，在於測度與分析變量之數值在時間上所發生之變動；例如吾人欲研究若干年內銷售額之變動，白煤產量之變動，或一般物價水準之變動。但若研究某時期內收入之分配，則研究之手續與前者迥異。此時吾人所欲知者，為各種收入階級中所包含之人數，故此處之整理問題，在於斷定變量之各值所重覆發現之次數及其分配狀態。此種資料一經整理後，卽成一次數數列（frequency series），而與時間數列迥不相同。研究此兩種數列所用之分析方法互異，故有分述之必要。本章所論述者，為不因時間而有變動之各種資料之整理與其初步分析之方法。

未經整理之資料

吾人所處理之數量資料，在未經整理以前，其原始狀態率皆淩亂，而乏整齊之形式與結構。此種資料或為企業方面所具備之生產或銷售之紀錄，或為搜集所得之各種物價。卽在應用別種機關所搜集之資料時，此項資料或已列成表式，但其排列與格式每不合於考查人（investigator）研究之用。是以統計之初步工作，在於將數字編成一適當形式以求合用，俾與性質相似之資料比較，且可為進一步分析之準備。科學方法包括觀察、推演與證實三步驟，前已言之；數字資料乃係觀察之所得，必須先行整理，使成為整齊之形式，始可進而推演。

下列數字係代表某製造廠工人210人在某週內按件計工之工資數，此爲資料之未經整理者之一例。

工人 210 人在某週內每人所得工資數

$26.25	$28.70	$24.15	$29.75	$29.20	$30.60	$23.40	$24.75
26.70	24.35	25.75	27.20	28.30	25.25	27.75	27.60
28.20	27.30	27.80	26.35	27.40	28.30	26.60	25.75
27.70	28.60	25.30	27.80	26·40	27.30	28.35	27.00
24.30	27.80	27.60	26.30	27.40	23.50	29.60	27.80
27.60	25.35	27.55	29.00	24.10	27.00	24.50	27.25
26.15	29.30	23.10	27.10	28.50	27.45	26.15	28.35
27.95	25.55	27.55	26.60	24.25	30.00	28.55	28.00
27.30	27.90	25.25	24.10	27.45	24.55	26.55	27.55
26.75	31.00	24.00	25.35	26.50	28.30	27.95	25.55
30.25	28.55	26.75	24.60	25.75	26.55	27.80	28.90
29.55	30.00	24.60	25.75	26.30	27.00	28.25	25.25
25.75	26.25	26.30	26.75	27.90	28.30	25.70	26.30
26.60	27.00	30.75	28.60	28.10	23.50	24.75	25.15
26.30	27.25	28.15	29.10	30.10	29.90	28.55	27.30
26.55	27.55	23.00	24.50	22.85	26.55	27.55	28.10
30.70	28.60	27.90	26.80	24.10	25.25	26.30	27.90
26.90	25.30	25.80	28.85	27.55	27.30	25.00	26.00
26.55	27.80	28.60	30.55	29.50	24.10	25.15	27.15
28.10	26.30	27.10	24.60	27.80	26.30	27.90	29.80
24.10	25.15	27.50	24.25	25.70	26.80	30.15	29.30
28.15	28.65	24.55	25.85	26.10	27.00	26.80	27.50
29.00	23.00	28.60	29.30	28.55	28.80	27.55	23.60
26.10	27.15	25.75	26.80	27.15	26.30	28.55	25.80
24.55	25.80	26.75	27.30	27.55	28.25	25.60	26.30
26.85	27.30	28.10	32.00	28.15	26.30	27.75	26.25
28.60	26.00						

整列

前項數字如依大小次序排列，即可成一整齊連貫之形式，其所包含

之最高數與最低數間之全距(range),及在此全距內各項數字分配之槪況,可因排列而易於明瞭,作為進一步整理之準備。此項數字依大小次序所排成之整列(array)如下:

整列:工人 210 人在某週內每人所得工資數

$22.85	$25.15	$26.15	$26.75	$27.45	$27.95	$28.60
23.00	25.15	26.15	26.75	27.45	27.95	28.65
23.00	25.15	26.25	26.80	27.50	28.00	28.70
23.10	25.25	26.25	26.80	27.55	28.10	28.80
23.40	25.25	26.25	26.80	27.55	28.10	28.85
23.50	25.25	26.30	26.80	27.55	28.10	28.90
23.50	25.25	26.30	26.85	27.55	28.10	29.00
23.60	25.30	26.30	26.90	27.55	28.15	29.00
24.00	25.30	26.30	27.00	27.55	28.15	29.10
24.10	25.35	26.30	27.00	27.55	28.15	29.20
24.10	25.35	26.30	27.00	27.55	28.20	29.30
24.10	25.55	26.30	27.00	27.55	28.25	29.30
24.10	25.55	26.30	27.00	27.60	28.25	29.30
24.10	25.60	26.30	27.10	27.60	28.30	29.50
24.15	25.70	26.30	27.10	27.60	28.30	29.55
24.25	25.70	26.30	27.15	27.70	28.30	29.60
24.25	25.75	26.35	27.15	27.75	28.30	29.75
24.30	25.75	26.40	27.15	27.75	28.35	29.80
24.35	25.75	26.50	27.20	27.80	28.35	29.90
24.50	25.75	26.55	27.25	27.80	28.50	30.00
24.50	25.75	26.55	27.25	27.80	28.55	30.00
24.55	25.75	26.55	27.30	27.80	28.55	30.10
24.55	25.80	26.55	27.30	27.80	28.55	30.15
24.55	25.80	26.55	27.30	27.80	28.55	30.25
24.60	25.80	26.60	27.30	27.80	28.55	30.55
24.60	25.85	26.60	27.30	27.90	28.60	30.60
24.60	26.00	26.60	27.30	27.90	28.60	30.70
24.75	26.00	26.70	27.30	27.90	28.60	30.75
24.75	26.10	26.75	27.40	27.90	28.60	31.00
25.00	26.10	26.75	27.40	27.90	28.60	32.00

次數表

上項整列所表現數字之形式，雖較未整理前整齊明瞭，但欲使讀者對於該項資料所包含之意義完全領會，猶須作進一步之整理。工廠經理雖可由此整列而知某週內工人所得最低工資數爲 $22.85，最高工資數爲 $32.00，及大部分工人所得工資數在 $25.00 與 $29.00 之間，但所得僅此，而對於資料之眞義，猶有失之含混之處。如用分組手續，將工資數分爲數組 (class)，而以各項數字分別歸納於各組中，即可得一簡潔之工資分配表。茲將分組後所得之表列下。表中每組之距離爲美金二元，此距離謂之組距 (class-interval)。

表　六

按某週工資分組之工人次數分配(組距＝$2)

某週工資數	獲得左列工資數之人數(次數)
$22.00 *to* $23.99	8
24.00 *to* 25.99	48
26.00 *to* 27.99	96
28.00 *to* 29.99	47
30.00 *to* 31.99	10
32.00 *to* 33.99	1
	210

此表爲原來數字之縮影，不特明示吾人以工資之全距，且表述該全距內工人 210 人所得工資數之分配狀況。此表已將該資料所有繁瑣之事實略去，吾人所可得者爲在某週內工資數之在 $24.00 至 $25.99 之間者爲 48 人，但此 48 人在該組內之分配情形，則無從知悉。就表中數字觀之，此 48 人所得之工資數或皆爲 $24.00。但欲將資料分組簡縮，則失却繁瑣與詳盡之事實，要所不免。

如將組距縮小，則失却詳盡之程度可以減少，然同時組數必增多，而表之形式亦必趨於繁瑣，反覺眉目不淸。下列各表係用同一資料，以

一元，五角，二角五分爲組距而分組所得之結果。

按某週工資分組之工人次數分配

表七 (組距＝$1)		表八 (組距＝$0.50)		表九 (組距＝$0.25)	
某週工資數	次數	某週工資數	次數	某週工資數	次數
$22.00—22.99	1	$22.50—22.99	1	$22.75—22.99	1
23.00—23.99	7	23.00—23.49	4	23.00—23.24	3
24.00—24.99	21	23.50—23.99	3	23.25—23.49	1
25.00—25.99	27	24.00—24.49	11	23.50—23.74	3
26.00—26.99	42	24.50—24.99	10	23.75—23.99	0
27.00—27.99	54	25.00—25.49	12	24.00—24.24	7
28.00—28.99	34	25.50—25.99	15	24.25—24.49	4
29.00—29.99	13	26.00—26.49	22	24.50—24.74	8
30.00—30.99	9	26.50—26.99	20	24.75—24.99	2
31.00—31.99	1	27.00—27.49	24	25.00—25.24	4
32.00—32.99	1	27.50—27.99	30	25.25—25.49	8
		28.00—28.49	17	25.50—25.74	5
	210	28.50—28.99	17	25.75—25.99	10
		29.00—29.49	7	26.00—26.24	6
		29.50—29.99	6	26.25—26.49	16
		30.00—30.49	5	26.50—26.74	10
		30.50—30.99	4	26.75—26.99	10
		31.00—31.49	1	27.00—27.24	11
		31.50—31.99	0	27.25—27.49	13
		32.00—32.49	1	27.50—27.74	14
				27.75—27.99	16
			210	28.00—28.24	9
				28.25—28.49	8
				28.50—28.74	14
				28.75—28.99	3
				29.00—29.24	4
				29.25—29.49	3
				29.50—29.74	3
				29.75—29.99	3
				30.00—30.24	4
				30.25—30.49	1
				30.50—30.74	3
				30.75—30.99	1
				31.00—31.24	1
				31.25—31.49	0
				31.50—31.74	0
				31.75—31.99	0
				32.00—32.24	1
					210

上列四表(表六至表九)係根據同一資料，依四種簡縮之程度整理而成者。表六、表七及表八具有一共同之特點：卽在首尾各組內次數(frequency)較少，愈近中間則各組次數愈見增多。組數愈多，次數分配不規則之程度亦愈著。表九之組距爲二角五分，計有 38 組，在其全距內

次數之分配，極無規則，毫不合於對稱(symmetry)狀態；其他各表之結構，較為整齊有序，與對稱狀態亦較接近。各表所示者皆為工資資料之簡縮形式，閱表卽可知該工廠工人在某週內所得工資之大小及其分配之情形，比較未經整理以前紛亂無章之數字明瞭多矣。蒐集之資料經此整理後，謂之次數分配 (frequency distribution)，其目的如其定義所述，在於用簡縮之形式表示一變量之各數值在其全距內之分配情形。編製次數分配表為整理與分析此種數量資料之初步手續。

編製次數表之步驟

次數表之編製方法，前節已述其概要，惟關於編製方面之重要各點，尙未述及，此處應作簡略之論述。上例中第一步手續係將各項數字依大小排成整列，惟實際製表時，此項排列手續並非必要。吾人可先就資料查明其上下兩極限，並決定分組之數目，將各組距塡入適當之空白紙上，而後將各項數值一一歸納於各組內。各項數值既已歸入各組，卽可進而計算各組之次數，而後製成上列表式。此項手續雖屬簡單，然牽涉數要點，不可不分別討論之。

組距之大小

決定組距之大小時(亦卽決定組數時)，所應注意之基本要點，在使每組內項目之分配，勿與勻整分配相差過遠，蓋在解釋次數表與以後根據此表從事計算時，例須用每組之中點 (mid-value) 以代表該組內各項目之數值也。例如根據表八計算時，在 \$26.00—\$26.50 一組內 22 項之數值，均須用該組之中點 \$26.25 代表之。此係根據各組內項目分配勻整之一種假定，與實際情形固不能全相符合，試閱表八之原來數值，卽知此項假定之不準確，然欲求絕對準確，則非將原來資料中每個不同之數值各成一組不可，但事實上不得不將資料簡縮，故簡縮時務須在不

違反其他條件下，使此種差誤儘量減少。如表六所定之組距卽嫌過大，與該資料不合。

欲求適合上項假定，必須將組數增多，已如上述，然分組之第二條件常與此抵觸。此第二條件爲分組時各組次數之分配必須整齊有序。如組數過多，組距過小，則其次數分配必欠整齊而乏結構，表九卽不合於此項條件；且若組數較少，處理資料當必較易，而其意義亦易於領會。通常組距之決定，宜使組數在 10 與 25 之間。至於某表之組數應爲若干，仍須視資料之性質而定。前例表八中，組距定爲美金五角，似與上述兩條件最能適合。

組限之位置

組限 (class limit) 位置之決定亦頗關緊要，如選擇得宜，則製表手續較爲簡單，而此後計算亦可省便。如所定組限與組距皆爲整數，製表時尤爲便利。每組中點如爲整數，則計算平均數與其他統計量數時可減省手續。欲選擇適當之組限與中點，莫如將組距之數値定爲 5 或其倍數，惟此係就一般而言，在應用時仍須斟酌資料之性質而定也。

數字資料之分配，常有集中於若干數値之傾向。茲以下列數字爲例，將此點加以說明。下表爲一九二一年會員銀行向聯邦準備銀行貼現所用商業票據之件數，依貼現率之高下分配而成之次數表之一部。

貼現率(百分率)	票據件數
6	18,970
$6\frac{1}{4}$	697
$6\frac{1}{2}$	4,616
$6\frac{3}{4}$	135
7	17,362
$7\frac{1}{4}$	10

上表之次數頗有密集於整數之趨勢，其次則含有 $\frac{1}{2}$ % 之各個百分率，其集中之次數亦多，而介乎鄰接兩個百分率間，則表中並一項而無

之。如遇此種資料，分組時各組之中點應選取項數密集之數值，選定組距時，亦應注意及此，蓋根據次數表計算時每組內之各項數值均以各該組之中點作爲代表也。上述資料之組距如定爲 $\frac{1}{2}\%$，則各組之排列應爲 $5\frac{3}{4}-6\frac{1}{4}$，$6\frac{1}{4}-6\frac{3}{4}$，……，而非 $6-6\frac{1}{2}$，$6\frac{1}{2}-7$……矣。

組限與觀察之確度

製表時組限之定義務須明確表示，使每組之全距無含混之弊，而各項數值歸入各組時，亦可不致發生困難。如下列之表式，爲吾人所偶見。

組距	次數
0—10	3
10—20	8
20—30	15
30—40	6
40—50	2

上表內之組限如不附以解釋，則遇某項數值適爲10時，其應歸入第一組抑第二組，不免發生問題。故每組之組限務必確切標明如下式，使其定義無含混不清之弊。

組距	次數
0— 9.9	3
10—19.9	8
20—29.9	15
30—39.9	6
40—49.9	2

惟組限之定義，與觀察之確度有關，如觀察之確度至十分之一單位爲止者，則組限依上式標明後，前項困難當可避免；但若觀察之確度僅至單位爲止者（如10之數值係介於 9.5 與 10.5 之間），則遇有與上項組限相同之數值，歸組時仍有困難。在此情形下，可將該項數值之次數，平分爲二，各以半數列入鄰接兩組之中。

關於組限問題，猶爾氏（Yule）(註)曾定一法則，謂組限之數值應較

(註)見 "An Introduction to the Theory of Statistics" 第 81 頁。

紀錄之數値展後一位小數。試以前述數字爲例，如觀察之確度至十分之一單位爲止者，則紀錄時作爲 9.9 之一項數値，其實際數値係介乎 9.85 與 9.95 之間，故組距應定爲0—9.95，9.95—19.95……，各組始有準確之意義。依此分組，則各組之中點當爲 4.95，14.95……。編製與應用次數表時，皆應注意及此。無論何時，組限之數値應視觀察之確度而定。

製表時如將各組之上下兩限併列表內，則製表工作較爲簡單；如僅列各組之下限 (lower limit)，或各組之中點，則歸組時每易發生錯誤。又若用表中數字計算時，更須於表內加列一欄，以記明各組中點之數値。

其他要件

表中各組組距之大小，必須一致，以便彼此比較。然組距大小不一之表式，亦偶有所見。表中之組距如不一致，或爲 5，或爲 10，則因組距之大小各異，以致各組不能相互比較。惟此種表式或因注重某距離內次數之分配情形，故有此需要，但遇此情形可製兩表，其主要表中組距仍須一致，而另附一輔助表以說明某距離內次數分配之情形可矣。

表中組距除須一致外，其意義更須肯定。前例按件計算工資之工人，其工資數在三十元及以上者，如歸入於標明爲"$30及以上"之一組內，則該組之上限 (upper limit) 究爲幾何，無從斷定；且根據此表計算時，此種錯誤關係尤鉅。惟若表中包含數項極端之數値，以致組距不能不如此標明時，則須於表末另加附註，載明此數項之實際數値，俾讀者得以了解。

前兩節內所述之差誤，可舉例(參看表十)說明於下：

在此例中，其最小組下限不明，最大組上限不明，故其組距無由斷定，而中間各組之組距，又不一致，組距爲 $1.00 者兩組，$1.50 者一組，$2.50 者一組。此種表式殊無價値之可言。

表十

按每日所得工資數分組之 *ABC* 公司工人之次數分配

工資組	每組工人數（次數）
$3.00 以下	15
3.00—3.99	30
4.00—4.99	40
5.00—6.49	30
6.50—8.99	10
9.00 及以上	5
	130

統計表之結構

上節所述，為關於編製次數表時所發生技術方面之種種問題，對於全表之格式，縱行(column)、橫行(row)之排列，表之標題(title)及其標識(notation)等均未涉及，排列表式之各種普通原則亦未論述。列表細則多不勝舉，此處所欲論述者，僅為關於統計表組織方面所應考慮之各點。

統計表為以簡略之形式表述一羣數量資料之一種方法，故其形式必須簡潔明瞭而有意義，且須易於解釋，否則卽不能具備製表與分類之功用。雜亂無章之統計表與缺乏條理之論文無異。編製統計表必有其目的，排列表式時應將此目的明白表現。排列之方法，雖因資料之性質而異，然對於標準規律務須嚴格遵守。下列各項普通原則，可為選擇統計表之形式與排列時之一助。

1.表之標題應以明瞭簡潔之文字，將表內所列資料詳細載明。

2.縱行橫行之標目務求簡明，不宜含糊。

3.變量之尺度應自左而右或自下而上逐漸增加。

4.縱行橫行可編列號次，以便引證。

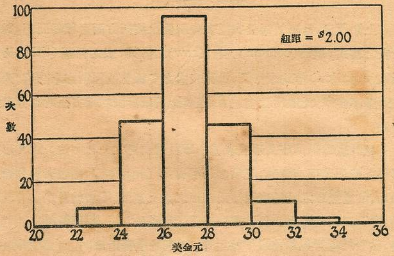
5. 表中所用之度量衡單位務宜明白詮釋。

6. 資料之來源務須一一註明。

7. 每一表式應自成一單位，所列資料之意義，須不費解釋而自明。如有解釋之必要時，應將註釋列入表內，或於表末加註。

次數分配之圖示

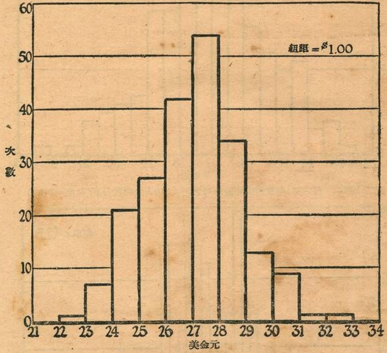
前節所述之次數分配，既將資料用簡明之形式表示，且可作進一步研究之準備，故在統計上實具有極大之功用。此種分配不僅可以表式顯示，且可用第二章內所述縱橫坐標制之圖解表示之。次數分配之各項特性，用圖解表示，最為明顯。



圖二十二． 直方圖：按某週工資數分組之工人 210 人之次數分配(組距＝$2.00)

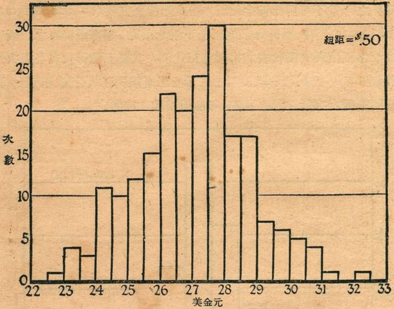
表六所列為工人 210 人某週所得之工資數，組距為美金二元；圖二十二即為該項資料之圖解。圖中組距沿 x 軸量之，次數沿 y 軸量之，兩軸各用相當之尺度表明；惟須注意者，橫坐標之尺度不自零點起，而自 $20起。為圖示便利起見，自 0 至 $20 之部分未列入圖中，讀者對此兩變

量關係之印象不可因此發生錯誤。繪製此種直方圖（column diagram or histogram）時，應將每個組距上下限之兩點，用橫綫接連之。圖中各條形之面積與其所代表之次數成正比例，其總面積係代表各組次數之和，即 210。此種圖解頗能明白表示各工資組內工人分配之人數，吾人觀圖則對於次數分配之狀況，可獲極明確之印象。

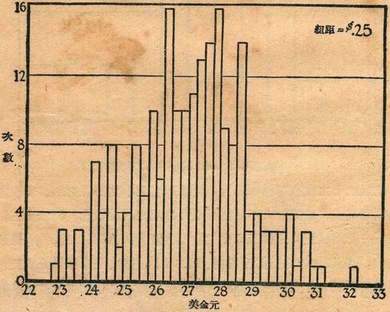


圖二十三．　直方圖：按某週工資數分組之工人210人之次數分配（組距＝$1.00）

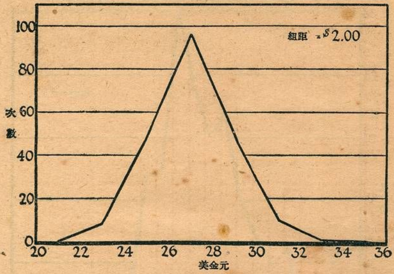
惟圖二十二之組距猶嫌過大，以致頗有與事實不符之處。表中將詳情刪略過多，足使吾人不易獲得項目分配之眞確觀念。圖二十三係以美金一元爲組距時所得工人分配狀況之直方圖。該圖組距較小，條形亦較狹，所得圖解較爲整齊而對稱。圖二十四亦然，是圖係以美金五角爲組距所得工人分配狀況之圖解。圖二十五所示之分配狀況，係以美



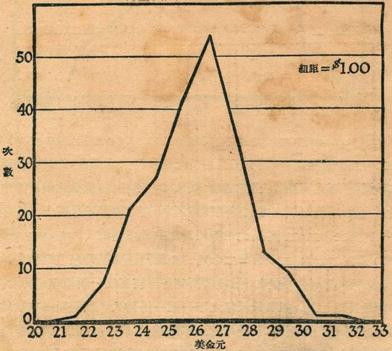
圖二十四． 直方圖：按某週工資數分組之工人210人之次數分配（組距＝$.50）



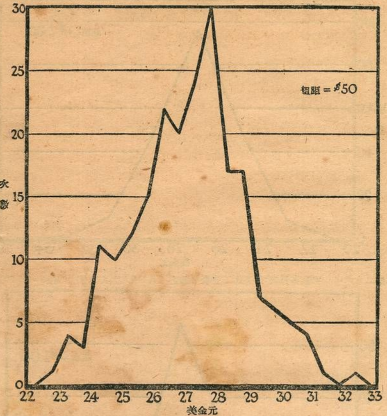
圖二十五． 直方圖：按某週工資數分組之工人210人之次數分配（組距＝$.25）



圖二十六．　次數多邊圖：按某週工資數分組之工人210人之次數分配（組距＝$2.00）



圖二十七．　次數多邊圖：按某週工資數分組之工人210人之次數分配（組距＝$1.00）



圖二十八． 次數多邊圖：按某週工資數分組之工人210人之次數分配（組距＝$.50）

金二角五分為組距，惟該項組距用於此種資料失之過狹，故該圖之結構亦不規則（吾人應注意上列四圖之縱尺度並不相同，故欲比較各圖中各組之次數，不能逕以直方之高度作為標準，而必須參考尺度上之數字）。

與圖二十二、二十三、二十四相對之次數多邊圖 (frequency polygon) 為圖二十六、二十七及二十八。此種圖解係以每組中點沿橫軸量之，每組次數沿縱軸量之，然後將所定各點接連成一斷續綫而得者。圖中將最高組之上一組及最低組之下一組一併繪入，該兩組之次數均為零，故此多邊圖之兩端，在該兩組之中點處，得與基綫相接。次數多邊圖中曲綫下之全部面積，表示次數之總數，但各組距內所包含之面積，並

不與各組之次數成正比例，因每組內次數分配並不勻整也。在各組中點處所有縱坐標之高度，則係代表各組之次數。

曲綫之修勻

變更組距之大小，又可得下述之結果。當組距逐漸縮小至某程度時，其所得直方圖與多邊圖亦漸見勻滑而整齊；逾此程度，則勻整之狀態漸見消失而呈現間斷之處，蓋在組距較大時所得各組次數有規則之變化，至此已爲次數變化不規則之各組所破壞。在圖二十五中卽可見次數變化不規則之一斑。此種不規則狀態，顯與次數分配之普通法則相悖；按普通法則，當工資數逐漸增加時，各組內工人數亦逐漸增加，至$27.50附近而達最高點，由此點起，工人數復逐漸減少，至最後上限爲$32之一組時，工人數僅得一人。此210工人所擔任之工作相同，而其收入之大小係取決於各個工人工作之速率與技術之高下，故各工資組內次數之增減自應表現一有規則之狀態。設吾人所用工資資料非一週之數字，而爲52週之數字，用此210人每人在全年內52週之平均工資，則雖用較小之組距，而所得次數分配必較前此爲勻整，蓋各週之特殊變動，經此平均後業已消除也。又設吾人採用10,920工人(52×210)一週內之工資數，則所得結果必與用一年內之平均數字者相同。由是知吾人倘欲得整齊勻滑之狀態，不僅須縮小組距，且須添加項數，使少數之意外變動，得以消除；否則如僅將組距縮小，則所得圖形必將如圖二十五所示。雖然，增加項目常爲事實所不許，吾人自不得不採用修勻曲綫之方法，以求組距極小、項目極多時次數分配之勻整狀態。採用此法吾人可得一勻滑而有規則之次數曲綫。

此種修勻後之次數曲綫 (smooth frequency curve)，可以代表資料所應有之分配狀態。前節嘗謂次數多邊圖中，因每組內次數分配之不勻整，故各組距內所包含之面積並不與各組之次數成正比例；惟在修勻

表十一

1918 年美國個人收入之次數分配

(包括收入之在 $4000 以下者)

收入組(註一)	人數(註二)
$ 0 *to* $ 100	62,809
100 *to* 200	103,704
200 *to* 300	209,087
300 *to* 400	489,963
400 *to* 500	961,991
500 *to* 600	1,549,974
600 *to* 700	2,154,474
700 *to* 800	2,668,466
800 *to* 900	3,013,034
900 *to* 1,000	3,144,722
1,000 *to* 1,100	3,074,351
1,100 *to* 1,200	2,850,526
1,200 *to* 1,300	2,535,285
1,300 *to* 1,400	2,205,728
1,400 *to* 1,500	1,832,230
1,500 *to* 1,600	1,512,649
1,600 *to* 1,700	1,234,397
1,700 *to* 1,800	999,996
1,800 *to* 1,900	811,236
1,900 *to* 2,000	663,788
2,000 *to* 2,100	549,787
2,100 *to* 2,200	463,222
2,200 *to* 2,300	395,115
2,300 *to* 2,400	340,141
2,400 *to* 2,500	295,490
2,500 *to* 2,600	258,650
2,600 *to* 2,700	227,731
2,700 *to* 2,800	201,488
2,800 *to* 2,900	178,901
2,900 *to* 3,000	154,499
3,000 *to* 3,100	142,802
3,100 *to* 3,200	128,217
3,200 *to* 3,300	115,583
3,300 *to* 3,400	104,504
3,400 *to* 3,500	94,803
3,500 *to* 3,600	86,405
3,600 *to* 3,700	79,023
3,700 *to* 3,800	72,562
3,800 *to* 3,900	66,900
3,900 *to* 4,000	61,894

(註一)各組之定義應釋為 '$0 至不滿 $100', '$100 至不滿 $200' 等是,故個人收入適為 $100 者,應歸入第二組內。

(註二)據美國經濟研究所報告,謂所列人數,其末位數字未見全無差誤,此種數字所表示之確度,僅屬形式而已。

之次數曲綫圖中，此種不規則之狀態業已消除，故在橫尺度某兩點上所豎立之縱綫間之面積，應與該兩數值間所含之理想次數成正比例。此外又因該綫呈現勻滑之趨勢，故可用插補法(interpolation)以斷定表中未載之數值(註一)。

表十一所列資料(註二)爲一九一八年個人收入在美金四千元以下者之分配，可用以說明曲綫之修勻方法。

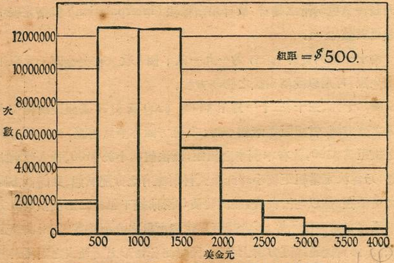
圖二十九、三十、三十一爲用$500，$200及$100爲組距所得個人收入之直方圖。當組距逐漸縮小時，其直方圖亦漸見勻整，惟因原有資料之組距爲$100，故在本例中之組距，無法使其小於$100。吾人當前之問題，乃爲決定當組距縮小時此項資料所應有之分配狀態。如將直方圖中之斷續綫，易以勻滑曲綫，而同時使勻滑曲綫下包含之總面積與直方圖中原有之總面積相等，則此綫可以代表所求之分配形態。勻滑曲綫之繪法，係使曲綫通過直方圖上各點，同時使各條形上所切去之面積與其所圈入之面積大略相等，則曲綫下所包含之全面積，自必與直方圖中原有之全面積相等。繪勻滑次數曲綫時，綫下包含之全面積，須與原有面積相等，實屬必要之條件。至於各條形之面積，雖因有不規則變化之存在，每不能與此法則相合，然此仍不失爲繪圖時之一種普通原則(用曲綫配合數字資料之精密方法，後文當再論述之。前節所述僅憑觀察以修勻曲綫之方法，其所得結果常與所求曲綫甚爲接近)。

圖三十二係將圖三十一之收入分配直方圖修勻後所得之圖形。該圖已將各組間不勻滑之輪廓加以修勻，故所得之圖成一依極微數量而增減之勻滑曲綫，可以代表收入分配之實際狀態，因其包含人數至數千

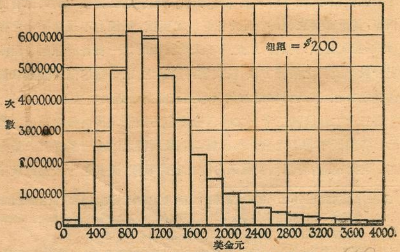
(註一)在統計實際工作上，每爲事實所限，所搜集之資料常有間斷之處，故各變量之數值往往不能連續。今用插補法可以估計兩已知數值間之各值，或估定曲綫上兩點間之各點。插補所得之數值以與已知各值最相合者爲最準確。

(註二)錄自一九二一年National Bureau of Economic Research出版，紐約Harcourt, Brace & Co.發行之"Income in the United States"第一卷第132—133頁。

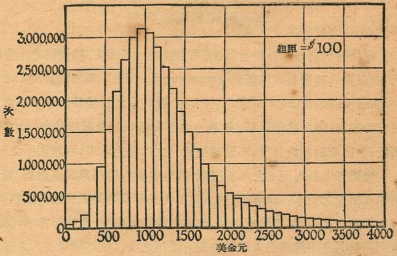
萬人時，其分配狀態固應如此也。此時吾人已得一可以代表基本分配之曲綫，其由分組而得之不規則變動及間斷之處均經消除矣。



圖二十九． 直方圖：1918年美國個人收入之次數分配
包括收入之在 $4000 以下者（組距 = $500）

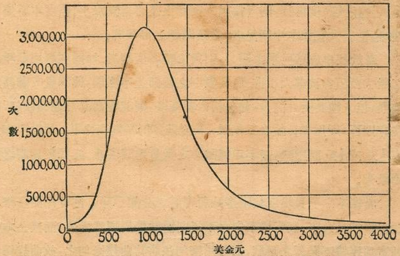


圖三十． 直方圖：1918年美國個人收入之次數分配
包括收入之在 $4000 以下者（組距 = $200）



圖三十一．　直方圖：1918年美國個人收入之次數分配

包括收入之在 $4000 以下者（組距＝$100）



圖三十二．　次數曲綫圖：1918年美國個人收入之次數分配

包括收入之在 $4000 以下者（根據組距爲$100之直方圖繪成）

連續數列與間斷數列

曲綫修勻之是否合於邏輯，須視資料之性質而定。就資料之性質而言，次數數列可分為連續數列 (continuous series) 與間斷數列 (non-continuous or discrete series) 兩大類。數列中之自變量依無限小之差量而增減者，謂之連續數列；數列中之自變量依定量而增減者，謂之間斷數列。間斷數列之曲綫，不若連續數列之曲綫之勻滑，而呈凹凸之狀。

世間可供吾人研究之凡百現象，其性質實有連續與間斷之別。吾人所用以研究之樣本(sample)，無論其所代表者為連續數列抑間斷數列，其自變量之數值，每呈間斷之狀，即在連續數列中，此種情形亦所不免，因測量器具之確度與吾人理解力均有限制之故。例如測量百人之高度所得之次數數列，其自變量(高度)必依定量而增減。在測量時，或以一英寸為最小單位，或以一英寸之八分之一為最小單位，但若將萬人或千萬人各按其高度排列，則鄰接兩人高度之差必為無限小。高度本為連續變量，惟吾人在樣本中所見者，常呈間斷之形狀耳。

至於利率或貼現率之變動，則與人體高度迥異。如將利率或貼現率之數字百項各按其大小排列，其數字之變動並不連續，與測量人體高度無異，但人體高度之測量，如包括人數極多，則其變動必呈連續狀態，貼現率則否，所用項目縱多至無限，其變動必仍呈間斷之狀，蓋貼現率乃以 $\frac{1}{2}\%$ 或 $\frac{1}{4}\%$ 而增減，而不以無限小之數量增減者也。此種數列謂之間斷數列。

當樣本之項目增至無限多時所應有之分配形態，亦可用曲綫修勻法以推求其大概。假定樣本中之不規則變動係屬偶然性質，而其基本變動實為連續不間斷之形狀，則可用修勻法將其曲綫修勻之，故曲綫修勻法實際僅能適用於連續數列。如根據樣本所得人羣高度之直方圖，即可依此法修勻，以求全體人羣高度分配之近似狀態，且可用作插補數值之

根據。此法如用於間斷數列，則殊不合邏輯。吾人設欲繪製一貼現率分配之勻滑曲綫，以求貼現率在 4.3675% 之一點上之理想次數，則無意義矣。雖然，在統計實務方面，每有以間斷數列視為連續數列，以助問題之解決者，此時曲綫修勻法亦可適用；惟在解釋與應用修勻曲綫時，此連續與間斷變動在邏輯上之區分，殊不可忽視耳(註一)。

統計資料之累積排列法

為欲達到某種目的，統計資料有須累積排列(cumulative arrangement)，而不宜採用各組次數分列之形式者。下列數字可說明此種排列法之優點。

此項數字為研究電桿木之耐用時期(註二)所得之結果。

表十二

按耐用時期長短分組之電桿木 248,707 根之次數分配

耐用時期 (年數)	電桿木根數 (次數)
0— 0.9	1,150
1— 1.9	4,221
2— 2.9	10,692
3— 3.9	13,966
4— 4.9	16,633
5— 5.9	18,211
6— 6.9	19,011
7— 7.9	19,260
8— 8.9	20,909
9— 9.9	19,879

(註一)讀者倘對於本節各點欲得更詳盡之論述，可參閱 G. H. Knibbs 氏著 The Theory and Justification of Curve Smoothing 一文，原文見 H. Secrist 氏著 "Readings and Problems in Statistical Methods" 第 278—282 頁，一九二〇年紐約 Macmillan 書館出版。

(註二)參看 Edwin Kurtz 氏著 Replacement Insurance，原文見一九二一年七月出版"administration" 第41—69頁。

表 十 二 (續)

耐用時期 (年 數)	電桿木根數 (次 數)
10—10.9	20,764
11—11.9	15,454
12—12.9	14,237
13—13.9	13,779
14—14.9	9,764
15—15.9	8,534
16—16.9	7,659
17—17.9	6,918
18—18.9	4,591
19—19.9	1,798
20—20.9	815
21—21.9	313
22—22.9	102
23—23.9	47

觀表可知電桿木在使用之第一年內卽行損壞者計有 1,150 根,使用一年而未滿二年卽行損壞者計有 4,221 根,餘類推;此爲次數表之普通形式。在此例中如將次數排作累積形式,則其意義必更明瞭,可由下表見其一斑。

表 十 三

按耐用時期長短分組之電桿木 248,707 根之累積次數分配

(向上累積式)

耐用時期	電桿木根數 (次 數)
不滿1年者	1,150
不滿2年者	5,371
不滿3年者	16,063
不滿4年者	30,029
不滿5年者	46,662
不滿6年者	64,873
不滿7年者	83,884
不滿8年者	103,144

表 十 三（續）

耐用時期	電桿木根數（次數）
不滿 9 年者	124,053
不滿10年者	143,932
不滿11年者	164,696
不滿12年者	180,150
不滿13年者	194,387
不滿14年者	208,166
不滿15年者	217,930
不滿16年者	226,464
不滿17年者	234,123
不滿18年者	241,041
不滿19年者	245,632
不滿20年者	247,430
不滿21年者	248,245
不滿22年者	248,558
不滿23年者	248,660
不滿24年者	248,707

次數數列之累積形式可分向上累積（cumulated upward）及向下累積（cumulated downward）兩種。上表係向上累積式，由此表吾人可推知耐用不滿某年數之電桿木根數；若將累積之方向顛倒，則欲推知耐用滿某年數以上之電桿木根數，常較便利。試仍用前項電桿木之數字依向下累積式排列之，可得下表。

表 十 四

按耐用時期長短分組之電桿木 248,707 根之累積次數分配

（向下累積式）

(1) 耐用時期	(2) 電桿木根數（次數）	(3) 百分比
0年及以上	248,707	100.0
1年及以上	247,557	99.5
2年及以上	243,336	97.8
3年及以上	232,644	93.6

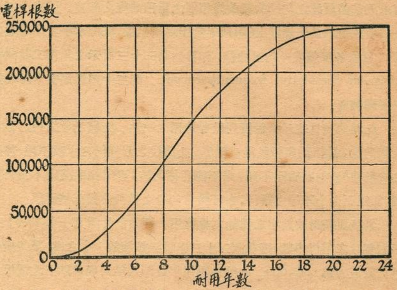
表 十 四（續）

耐用時期	電桿木根數（次　數）	百分比
4年及以上	218,678	88.0
5年及以上	202,045	81.2
6年及以上	183,834	73.8
7年及以上	164,823	66.3
8年及以上	145,563	58.5
9年及以上	124,654	50.1
10年及以上	104,775	42.1
11年及以上	84,011	33.8
12年及以上	68,557	27.6
13年及以上	54,320	21.8
14年及以上	40,541	16.3
15年及以上	30,777	12.4
16年及以上	22,243	8.9
17年及以上	14,584	5.9
18年及以上	7,666	3.1
19年及以上	3,075	1.2
20年及以上	1,277	0.5
21年及以上	462	0.2
22年及以上	149	0.06
23年及以上	47	0.02
24年及以上	0	0.00

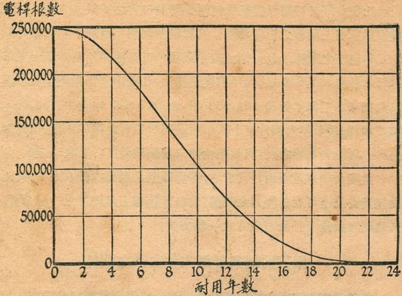
前項累積次數表在處理各種資料時，每有顯著之利便。吾人倘欲將折舊問題作科學上之研究時，常須爲各種器物分別製就精密之生命表(mortality table)，而此表以採取累積形式最爲合宜。製此表時，每有將次數化爲百分比之必要，如表十四第(3)欄所示。此種百分比之意義須視其所根據之原來數字而定。

累積次數曲綫

簡縮資料必須分組，而累積資料之功用，每因分組而受限制，倘吾人不用數學插補法，則所得之累積數字，僅限於表中所列各點，故吾人



圖三十三· 累積次數曲綫圖：按耐用時期長短分組之電桿木分配（向上累積式）



圖三十四 · 累積次數曲綫圖：按耐用時期長短分組之電桿木分配（向下累積式）

有求與前節所述之勻滑次數曲綫相似之累積曲綫之必要。將表十三內之各項數值繪於坐標圖上(耐用年數作爲橫坐標,電桿木根數作爲縱坐標),更依各點繪成一勻滑曲綫,卽得如圖三十三所示之累積次數曲綫(ogive or cumulative frequency curve)。圖三十四係用表十四之資料所繪製者。

此種曲綫爲表示次數數列最有效用之一法。此綫已將組距之界限大部消除,故組距與組數縱有變更,此曲綫之形態仍可大致不變。普通之次數曲綫,如其組距不同,每不能互相比較,而累積曲綫則不受此限制,卽組距大小不等之資料,亦可繪爲勻滑之累積曲綫而無錯誤。

累積曲綫對於數字之插補尤爲適用。例如欲知耐用年數不滿十五年六個月之電桿木數,可在圖三十三中橫坐標等於 $15\frac{1}{2}$ 之處,讀其縱坐標之數值約爲 222,000。如欲知耐用年數滿八年六個月以上之電桿木數,則可自圖三十四中依同法測定之,此數約爲135,000。

此種曲綫又可用以插補任何兩點間所包含之次數。此兩點之距離,並不以原表上所載之組距爲限。例如欲知耐用年數在十年六個月以上十五年以下之電桿木數,則由表中或圖中可得耐用年數不滿十五年之電桿木數爲 217,930 根;更用前述插補法,在圖中求得耐用年數不滿十年六個月之電桿木數爲 154,000 根,由 217,930 中減去 154,000 得 63,930,卽爲十年六個月至十五年兩點間所包含之電桿木數。惟此數不過爲一近似於實值之數,與其他由修勻曲綫插補所得之數字無異。

累積曲綫亦可逕由整列求之,而毋須經過編製次數表之手續。此種曲綫實質上僅可視爲整列之圖解,爲統計整理上最簡單之一種形式,而又爲處理數量資料最有效力之一種方法。

累積曲綫與普通次數曲綫之關係

累積曲綫與普通次數曲綫乃係用同一資料所繪製兩種排列不同之

形態，而各有其特殊之優點。如吾人能瞭解此兩種曲綫結構上之關係，則對於兩種排列之特點更易明瞭。此種關係可以圖三十五表示之。

此圖係根據下列次數表繪製，表示按照每千呎木料所需勞工費用分組所得之美國鋸木廠分配(註)。

表十五

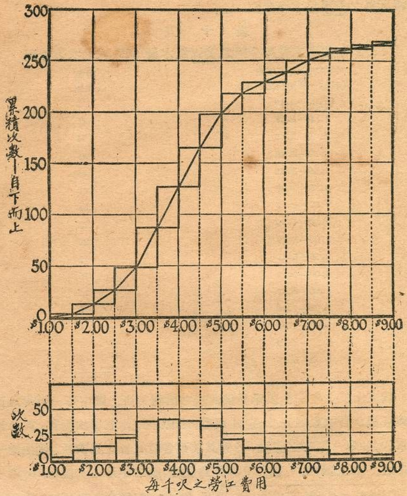
1921 年依勞工費用分組之美國鋸木廠 269 家之次數分配

每千呎之勞工費用（包括一切雇員之薪工）	廠數（次數）
$1.00—1.49	3
1.50—1.99	10
2.00—2.49	14
2.50—2.99	22
3.00—3.49	38
3.50—3.99	40
4.00—4.49	38
4.50—4.99	33
5.00—5.49	20
5.50—5.99	11
6.00—6.49	10
6.50—6.99	11
7.00—7.49	8
7.50—7.99	4
8.00—8.49	4
8.50—8.99	3
	269

圖三十五上半部係表示累積曲綫之構造方法。圖中每一矩形之面積與各組所含次數成正比例，此與直方圖相同；惟就作法而言，此係累積形式，即每一矩形之底綫，各為其前列各組次數之累積數。例如第一矩形之 y 值(次數)為 3，係以 0 線作為底綫；第二矩形之 y 值為 10，則不

(註)摘錄一九二三年一月"Monthly Labor Review"第14頁所載 Ethelbert Stewart 氏著 Labor Efficiency and Productiveness in Sawmills 一文。$9.00 以上之勞工費用，因其數字散漫，均未列入圖表中。

復以 0 線爲底綫，而以 3 爲底綫，餘類推。接連各矩形所成之曲綫，其斜度 (slope) 初以各組次數較小，頗爲平坦，繼以各組次數漸大而漸陡峭，終以近上限處各組次數又漸減少而復轉平坦，此卽所謂累積曲綫也。



圖三十五． 1920年按勞工費用分組之美國鋸木廠次數分配

說明累積曲綫與普通次數曲綫構造上之區別

假定各個 x 之數值不變，將代表各組次數之各個矩形，皆以零綫作爲共同之基綫排列之，卽得前文所述之直方圖，由此圖形卽可繪成次數多邊圖或修勻次數曲綫。

參考書

Bowley, A. L. "Elements of Statistics" (52—81).

Chaddock, R. E. "Principles and Methods of Statistics" (Chaps. IV, V).

Day, Edmund E. "Standardization of the Construction of Statistical Tables". Quarterly Publications, American Statistical Association, March, 1920.

Jones, D. C. "A First Course in Statistics" (5—21).

Kelley, Truman L. "Statistical Method" (1—37).

King, W. I. "Elements of Statistical Method" (83—107).

Pearl, Raymond. "Medical Biometry and Statistics" (74—104),

Rugg, H. O. "Statistical Methods Applied to Education" (57—94).

Secrist, Horace. "Introduction to Statistical Methods" (116—157).

Secrist, Horace. "Readings and Problems in Statistical Methods" (242—271).

Yule, G. Udney, "An Introduction to the Theory of Statistics" (75—105).

第四章　次數分配之表述：平均數

數量資料之分組與次數分配之編製，爲統計整理與分析之重要步驟。分組而後可表明資料之基本結構，揭示大量資料之劃一性質，惟此爲統計分析之第一步，吾人尙須應用其他方法將資料全部之重要特性，爲精密之測度與表述。次數分配之本身須簡縮之，使其特性可用三四有意義之數字表明之。

如每一次數分配含有特殊之性質，而各隨一獨具之法則，則研究與表述此種次數分配，誠屬困難，所幸事實並不如此。數量資料彼此性質縱不相同，然一經整理編爲次數分配後，每表現共有之特性，而合於普通法則，故在某方面所獲得之經驗，即可用爲他方面之南鍼。大量資料之變化具有劃一之性質，故由各種科學研究所得之數量資料，可用一普通方法整理之、分析之而比較之。

次數分配之比較

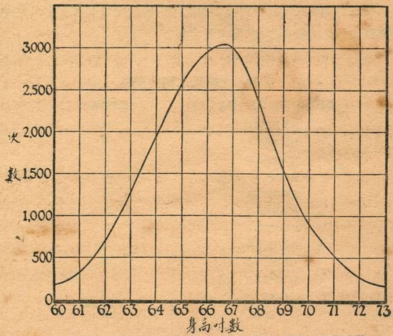
世間各種現象之數量排列，每遵循一普通法則，此可由比較各種資料之次數分配以發見之。吾人可觀察下列各種次數分配與次數曲綫之特性，並比較其分配狀態。

圖三十六係用下列某軍隊中兵士 18,780 人之身材高度分配所繪製之次數曲綫(註)。

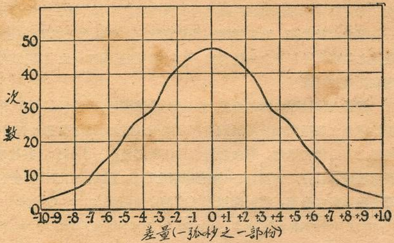
表　十　六

身高吋數	兵士數	身高吋數	兵士數
60+	197	67+	3,017
61+	317	68+	2,287
62+	692	69+	1,599
63+	1,289	70+	878
64+	1,961	71+	520
65+	2,613	72+	262
66+	2,974	73+	174
		總　計	18,780

(註)錄自 G. C. Whipple 氏著 "Vital Statistics" 第 377 頁。



圖三十六．　次數曲綫圖：按身材高度分組之兵士18,780人之次數分配



圖三十七．　次數曲綫圖：天文觀察差誤之分配

圖三十七係用卜拉德賚氏(Bradley)測量星球四百七十次所得觀察差誤(註)之次數曲綫。其原來數字如下:

表 十 七

某種天文測量之觀察差誤

差誤數量(一弧秒之一部分)	差誤次數
自0.0至0.1	94
自0.1至0.2	88
自0.2至0.3	78
自0.3至0.4	58
自0.4至0.5	51
自0.5至0.6	36
自0.6至0.7	26
自0.7至0.8	14
自0.8至0.9	10
自0.9至1.0	7
1.0以上	8
總計	470

前項資料因未將差誤分別正負,故不知其差誤爲高於眞實數值抑低於眞實數值,若據以製圖,祗可得次數曲綫之半部。圖三十七所示爲一完整之次數曲綫,乃係根據高於眞實數值之差誤次數必與低於眞實數值之差誤次數相等之一種假定,將其他半部補繪而成之曲綫。該項假定在此處頗可適用。

倘吾人移動礮口向某目標(即一點)精確瞄準,而後發射百次,則此百粒彈丸之碰撞點必分佈於該目標之四周。無論此礮之構造如何精密,瞄準之方法如何準確,彈丸之適中目標者必僅占一小部分,其他碰撞點將散佈於目標之四周,成一極有規則之形態。試繪一矩形圖,使包括彈丸碰撞點之全部,復將此矩形(或稱分佈區域 zone of dispersion)

(註)見 Mellor 氏著 "Higher Mathematics for Students of Chemistry and Physics" 第 514 頁。

分爲八個相等部分，則彈丸碰撞點在此八個部分間之分佈狀態，將如下圖所示（實際上之分佈或與下圖略有參差，惟普通必具此分佈狀態）：

2	7	16	25	25	16	7	2

圖三十八．礮彈發射時彈丸百粒之分佈區域；表示彈丸之理想百分分配

此項普通法則可驗諸各種槍礮而不爽。槍礮之構造愈精密，則分佈區域之範圍亦愈狹隘，然在該區域內碰撞點分佈之狀態，在理論上應相類似。槍礮瞄準所用之法則卽基此而成。

玆將實地射擊之結果與理想分配一比較之。下表係礮位與一靜止之目標相距二百碼處發射千彈所得之紀錄（註）。彈丸分佈區域係用平行綫等分爲十一部。

表十八

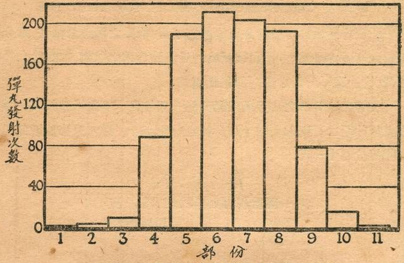
一礮發射彈丸千枚之次數分配

部分	彈丸碰撞次數
1(頂)	1
2	4
3	10
4	89
5	190
6	212
7	204
8	193
9	79
10	16
11(底)	2
總計	1,000

（註）此項試驗見 "The Report of the Chief of Ordnance", 1878, Appendix S. Wiley 氏所著 "The Method of Least Squares" 第14頁中曾引用之。該書係1897年紐約 Mansfield Merriman 公司出版。

此項彈丸分佈狀態如圖三十九所示。

彈丸分佈區域係分爲十一個部分，在說明上述之理想分配時，係分爲八個部分，故兩者不能直接比較，但其分配狀態，却與前述各例之普通形態相同。彈丸分配之較爲集中於目標下半部之原因，當係放置礮位時，未能完全顧及地心吸力所致。



圖三十九·　直方圖：一礮發射彈丸千枚之次數分配

茲更以擲錢爲例，假定錢幣落地時，正面與反面向上之次數分配，均係隨機遇而定。試以錢幣十枚同時拋擲百次，其每次正面向上之枚數如下表所示（在此項試驗中，每次拋擲正面向上之枚數，最多爲 10 枚，最少爲 0 枚）：

表　十　九

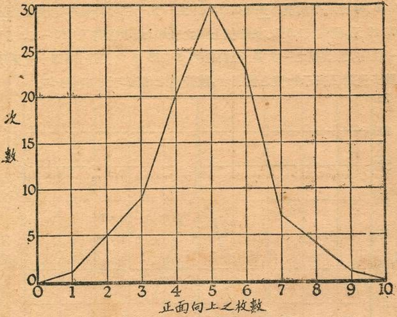
錢幣十枚同時拋擲百次所得之次數分配

正面向上之枚數	次數
10	0
9	1
8	4
7	8

表十九（續）

正面向上之枚數	次數
6	23
5	30
4	20
3	9
2	5
1	1
0	0
	100

圖四十表示上項次數分配之狀態。



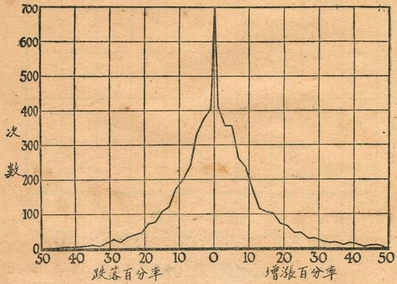
圖四十.　次數多邊圖：擲錢試驗中錢幣正面向上枚數之次數分配

經濟資料之分配

在上述四種絕不相同之現象中，吾人嘗發現數量資料排列之普通法則，惟此四項實例皆非代表經濟事實，現當研究經濟資料是否亦具有相似之特徵。參考第三章內所舉各例，而與上述四例比較之。該章內所

舉各例，爲工人某週工資之分配，電桿木之耐用年數，鋸木廠之勞工費用，及美國人民每年收入在四千元以下者之收入分配(此項收入分配，如將四千元以上之收入亦一併列入，則其曲綫將向右方極度延長，而成長尾狀)。茲更就經濟方面加舉數例於下。

圖四十一表示物價變動之分配狀態。圖中所用資料爲密哲爾氏(W. C. Mitchell)研究所得某年5,578種物品之批發價格與前一年比較之結果(註)。例如一九一二年紐約高地米特令棉花(middling upland cotton)每磅之平均價爲美金0.115元，至一九一三年每磅平均價增至美金0.128元，計增11.5%。在物價上漲之表中，此項卽歸入“10—11.9%”之一組內，爲該組之一次數。全表包括5,578次之紀錄。圖四十一係用次數多邊圖以表示物價之變動者，該綫並未用修勻法將其修勻。



圖四十一　次數多邊圖：某年批發物價比前一年漲落百分率5540項之次數分配

(註)摘錄Bulletin 284, U. S. Bureau of Labor Statistics, Part I, "The Making and Using of Index Numbers" 第18頁。此項數字係表示自跌落51%起至上漲51%止之物價變動。此外物價跌落55%者計1項，上漲52%至104%者計37項，圖中均未繪入。

表二十係表示一八八二年至一九一三年間倫敦對紐約匯價（一鎊合美金元之數）之分配。該表係將三十二年內共384個月之匯兌率依數字大小所組成之次數分配（註）。

表　二　十

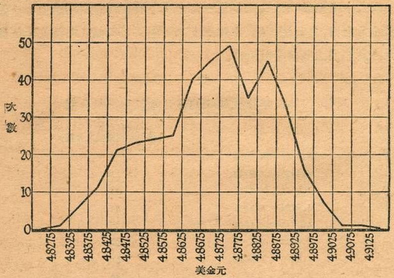
1882—1913 年倫敦對紐約每月匯價之次數分配

組　距	次數（發現左列匯價之月數）
$4.8275—$4.8324	1
4.8325— 4.8374	6
4.8375— 4.8424	11
4.8425— 4.8474	21
4.8475— 4.8524	23
4.8525— 4.8574	24
4.8575— 4.8624	25
4.8625— 4.8674	40
4.8675— 4.8724	45
4.8725— 4.8774	49
4.8775— 4.8824	35
4.8825— 4.8874	45
4.8875— 4.8924	33
4.8925— 4.8974	16
4.8975— 4.9024	8
4.9025— 4.9074	1
4.9075— 4.9124	1
	384

此項資料如圖四十二所示。

於此所應注意者，凡根據自然界資料繪製之曲綫，恆呈對稱(symmetry)與勻整之形態，而根據經濟資料繪製之次數曲綫及直方圖，則

（註）此項資料見 E. G. Peake 氏著 'An Academic Study of Some Money Market and Other Statistics" 附錄 I，1923 年倫敦 P. S. King 公司出版。原表註明'此項數字係根據 "The Economist" 每月初發表之匯價所計算之每月平均數。自 1886 年起係電匯價，以前則係短期匯價'。



圖四十二． 次數多邊圖：倫敦對紐約匯價之次數分配（384個月之每月平均匯價）

未必皆然，有偏斜（skew）於右方或左方者，亦有因鄰接各組間次數增減之過驟，其曲綫呈不規則之形態者。惟兩種資料雖有此種區別，然不論其資料之屬於經濟、天文、人體測量、射擊或機遇方面者，其分配之形態恆有相似之處。此項次數分配之一般特徵當於下節論述之。

次數分配之一般特徵

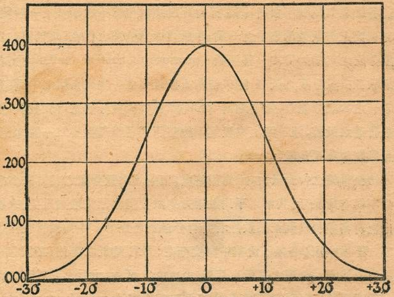
每種測度其數值必具有差異（variation）之特徵：人體之高矮不同，星球測量所得之結果不同，礮彈之發射卽在同一情形之下，其碰撞點之地位不同，個人收入有多寡之分，匯兌市價有高低之別，凡此種種觀察數值，莫不變動於極高極低之兩限度間。

各種觀察之數值，其分配之形態，沿橫軸尺度自左而右量之，爲由一極低之值，漸及於極高之值，而依次移動觀之，其在各點發現之次數（卽連續各組之次數），漸次增多，而至一次數極多之點，由此乃逐漸減

少，而至次數極少之點，其增減之狀態頗有規律，而其次數每呈現集中之趨勢(central tendency)。集中趨勢者，次數密集於橫軸尺度上某數點之謂也。此集中趨勢又爲各種次數分配所同具之第二特徵。

倘自最大集中點沿 x 尺度量其左右各組與此點之差異，則可發現差額較小各組所含之次數，必遠較差額較大各組者爲多，且極端之差額頗不多見，而在集中點左右兩旁之分配，在自然科學與純粹機遇之各例中幾完全相等，而在經濟方面之各例中，亦近乎相等(惟此亦常有例外，個人收入之分配，卽爲其最著之一例，其最大集中點左右兩邊之分配，頗不相等也)。

圖四十三所示之曲綫，謂之機率曲綫(probability curve)，或稱差誤常態曲綫(normal curve of error)。其特徵當俟後文詳論之，此處所以先行提示者，因其爲次數分配之基本形式。前述各例中，其分配形態有與之極相類似者，亦有表現顯著之差別者。與此形式不同之分配，亦



圖四十三．差誤常態曲綫

所常見，惟就其基本形式而論，此種常態曲綫在統計學上極關重要。即與此形式極不相同之分配，亦不無相似之點，故仍可用一普通方法以表述該種次數分配。數量資料之分配必有不同之處，其彼此間之差異，及其與某種標準形式差異之處，固含有重要之意義，但不論其差異程度若何，彼此間終必有類似之處，故每一次數分配，决非一隔離之現象，而為全體之一部分，而可採取他種分配所用之方法以表述之、分析之。

既知此種普通之形式，又將如何表述一種次數分配或區別其與他種分配之差異乎？由前文所述各點，即可推知各種表述與區別之方法，請分論之。

表述次數分配之方法

各個觀察所得之數值，恆必分配於 x 尺度之各點上，已如前述。在此各點之中，吾人可擇一足以代表其全體之單獨數值以表述其次數分配。各點之次數旣不相同，則選擇時理應選取包含次數最多之一數值，亦即尺度上次數最為集中之一點，作為該次數分配之代表。此數值即為測度分配之集中趨勢之一種量數 (measure of central tendency)。在個人收入次數分配之例中，應以包括人數最多之一個收入組之中點(在表十一之次數分配中，此中點為 \$950)作為該次數分配之代表數值。惟此最普通之數值，僅為測度次數集中趨勢所用各種量數中之一種。此各種量數概稱為平均數(average)。

此種單純之代表數值，雖為用甚廣，但關於次數分配之其他許多事實尙不能有所表示。其中以平均數兩旁次數分配之性質最關重要。次數之分配有密集於尺度上某數點者，有分散於各點者，平均數代表性之高下，須視其兩旁各數值之集中程度而定；故平均數之應用，必須藉一測量中心值兩旁分配之離散度量數(measure of variation)以補充之。

研究次數分配時，更應顧及分配之對稱程度(degree of symmetry)。

最大集中點之兩旁，其次數分配是否勻稱，或其次數曲綫是否如前述個人收入分配之例，有偏斜於一邊之形態，皆應考究明瞭。如其曲綫並不對稱，則當確知其不對稱(asymmetry)之程度，於是有測度分配之偏斜度量數(measure of skewness)。

最後，吾人可用差誤常態曲綫爲標準，以測度次數曲綫之峻峭程度。如圖四十一所示物價變動之次數多邊圖，一經修勻，則此曲綫必較常態曲綫爲峻峭，而此項目集中於中心數值之事實，頗具重大之意義。此種特徵謂之峯度(kurtosis)。測量峯度爲表述次數分配之最後步驟。

前述各種量數既已求得，則統計分析工作可謂完成其大半。此項工作開始之時，係先將凌亂之資料一一分組，製成一次數表之形式；次將該表所呈現之重要事實，歸納爲三個或四個有意義之量數。經此分析手續，則不僅可揭露一次數分配之特性，且可將該分配與性質相似之其他次數分配比較之。例如欲將數千萬未經整理之美國人民個人收入之數字與同樣之英國數字作比較，必不可能；但若由此兩類數字各求其平均數值，或最足代表之收入額，並將此中心值兩旁收入之分配狀態予以表述，則比較時可得一確切之標準。處理與分析大量資料時，不論其研究之目的若何，應藉表述次數分配之各種量數，將資料儘量簡縮、綜合而比較之。

下節當將表述次數分配之一種量數，卽關於測度集中趨勢者，加以討論。此點討論以後，當進而論述測量離散度與偏斜度之方法。

平均數

吾人已知平均數——單獨之典型數值——所以能代表次數分配者，乃因大量數字有密集於中心值附近之趨勢，與各個數值散佈之狀態又多少具有規則與勻整形式之故，是以此項代表數值之有無意義，當全視次數之是否密集於尺度上之中心點而定。當平均數之值確爲一典型數

值時，方可用此數以代表分配之全體。如次數分配中各數值差異頗大，且無集中之趨勢，則决無一單獨數值可代表次數分配之全體者。例如 3，125，1000 三個數字之算術平均數爲 376，但此 376 之一數值絕不能用以代表其原有之三個數值。由是知次數分配必具備集中於中心數值之趨勢之基本條件，其平均數始有代表全體之價值。

吾人苟回憶次數分配之一般特徵，卽可得一合理之平均數。前文中嘗提及在 x 尺度上有次數最爲集中之一點，卽包含次數最多之一數值，此值當可視爲分配全體之典型數值。此典型數值謂之衆數 (mode)，衆數所在之一組，謂之衆數組 (modal group)。試繪一次數曲綫以代表某一次數分配，此衆數卽爲與最高縱坐標相對之 x 值（註）。最高縱坐標乃係表示該衆數組之次數者，學者在决定衆數時，對於最高縱坐標之 x 值與最高縱坐標之值，每易混淆。衆數之值並非 y 軸上之距離，乃 x 軸上之距離也。至於縱坐標則用爲測度各組之次數，而並不測度各次數所歸納之數值者。

除衆數外，吾人又可在 x 值之尺度上選擇左右兩邊各占全體項數之半之一點，作爲次數分配之典型數值。此值當爲大於全體項數之半而同時亦爲小於全體項數之其他一半之數值，故謂之中位數 (median)。例如有人估計一九一八年美國人民個人收入之中位數之值爲美金 1,140 元，卽謂享有收入者之 37,000,000 人之中，其半數之個人收入皆少於此數，其他半數之個人收入皆多於此數。如將次數分配繪成一次數曲綫，則在 x 軸上中位數值之一點上所豎立之縱綫，適劃分曲綫下全面積爲兩個相等部分，此由於中位數本身之定義，同時亦由於次數曲綫下之面積，係代表該次數分配之全體項數故也。

算術平均數 (arithmetic mean) 爲代表次數分配所用之第三種平均數，係由計算而得之一種平均數，故受次數分配中任何一項數值之影

（註）嚴格言之，衆數爲配合於某一次數分配上之理想次數曲綫中、相對於最高縱坐標之 x 值。

響，此爲與衆數、中位數不同之點，蓋後兩者均視各項目在次數表中之位置而定，而不受個別項目之數值之影響也。算術平均數實爲次數分配之重心，如曲綫能繪成立體形，則此曲綫平衡支點處之 x 值，即爲算術平均數。

平均數中尙有所謂幾何平均數 (geometric mean) 與倒數平均數 (harmonic mean)兩種，其性質當俟後文討論之。

前述各種平均數之計算，在包括項數較多時，頗費手續，惟若運用適當方法，則可減省工作不少。玆爲便於說明各種方法起見，引用下列各種符號：

M　算術平均數。

Mo　衆數。

Md　中位數。

m　個別觀察所得之數值；在次數分配中，此數值爲各組中點之值。

f　次數分配中每組之項數。

N　一個數列或次數分配中之總項數。

Σ　(sigma)表示各項相加之符號，其意義爲'……之和'。

算術平均數之計算

應用前項符號，則算術平均數公式爲

$$M=\frac{\Sigma m}{N}$$

例如 2, 5, 6, 7 四個數值之算術平均數，等於此四數之和除以 4，即 $\frac{20}{4}=5$ 是也。如各個數值係根據其實際數值，則算術平均數之計算，僅須費簡單之加法與除法手續。第三章內嘗表列工廠工人 210 人之某週工資數，如將此種數字一一相加，所得和數，再以 210 除之，則得每人某週平均工資爲美金 26.982 元。惟如此計算，將 210 項數字一一相

加，工作已覺煩悶，假如吾人欲將美國 37,000,000 人個人收入之數字亦依此一一相加，則其計算工作幾屬不可能；故爲便於實際工作起見，通常係根據次數分配表計算平均數，而不由未經分組之原有資料計算。茲以北達科多小麥產量爲例，以說明平均數之計算手續。美國農業經濟局嘗記載一九一一年至一九二一年各年北達科多州五十三郡農田每英畝平均所產小麥之數量。茲將此項數字簡列於下：

表 二 十 一

北達科多州各郡小麥產量(註)算術平均數之計算

組距 (每英畝蒲式耳數)	中點 m	次數 f	fm
0 — 1.9	1	3	3
2 — 3.9	3	26	78
4 — 5.9	5	78	390
6 — 7.9	7	101 107	749
8 — 9.9	9	113	1,017
10 —11.9	11	65	715
12 —13.9	13	40	520
14 —15.9	15	22	330
16 —17.9	17	45	765
18 —19.9	19	41	779
20 —21.9	21	21	441
22 —23.9	23	8	184
		569	5,971

$$M=\frac{\Sigma(fm)}{N}=\frac{5\ 971}{569}=10.49\ \text{蒲式耳}$$

在討論數值分組時，嘗述及應需注意之各點，在本例中更可闡明其重要程度。組距大小之確定，應與每組內項目分配勻整之假定，不甚抵觸爲主，此項原則嘗於前文論及之；蓋若每組內項目分配勻整，則其中點始可代表該組內之各項，但若項目分配欠勻，則該組中點即不能代表各項。算術平均數可根據次數分配計算者，係基於每組內項目分配

(註)該時期內之各郡平均產量見 Bulletin 165, Agricultural Experiment Station, North Dakota Agricultural College, 1922 及"Cost of Production and Farm Organization on 126 Farms in North Dakota", 1921,第120—121頁。

匀整之一個假設上，故以每組中點乘該組項數時，係假定其乘積略與該組內各個別數值相加之和數相等。算術平均數公式因之變為

$$M=\frac{\Sigma(fm)}{N}.$$

表二十一係舉示應用此式計算之手續。

依此手續計算所得之數值，亦稱加權算術平均數(weighted arithmetic mean)。吾人計算時，實際上係求 m 欄內十二個數字之算術平均數，惟此平均數非十二個數字之簡單平均數，而係將此十二個組中點各按其組距內所包含之項數加權後計算而得。設有五人於此，其中兩人每年收入各 \$2,000，另三人各 \$3,000，則計算五人平均收入之手續，正與此相同。計算時並不僅將 \$2,000 及 \$3,000 相加而除以 2，而先將此 \$2,000 之數以權數2乘之，\$3,000 之數以權數 3 乘之，而後求其和數得 \$13,000，再以5除之而得。根據次數分配計算算術平均數之手續，雖具有此加權之形式，惟加權平均數(weighted average)之名詞尚有其較狹之意義，容後申論之，故此名詞通常不適用於按照次數分配計算之平均數。

計算算術平均數之簡捷法

通常算術平均數如由次數表計算，似較由未經分組之資料計算為易，但若包含項數過多，如照前法由次數表計算，手續仍頗煩瑣，惟此手續亦可設法化簡。

由算術平均數之計算法中可推知各數值對其算術平均數之正負離中差(deviation)之總和為 0，此點可以代數式證明之。今設 $m_1, m_2, m_3 \cdots\cdots m_n$ 代表一個數列中之各數值，M 代表其算術平均數，$d_1, d_2, d_3, \cdots\cdots d_n$ 代表各個數值對其算術平均數之離中差，則

$$M=\frac{m_1+m_2+m_3+\cdots\cdots m_n}{N}\cdots\cdots\cdots\cdots(1)$$

$$NM=m_1+m_2+m_3+\cdots\cdots+m_n\cdots\cdots\cdots\cdots(2)$$

今 m 之項數爲 N，故

$$(m_1-M)+(m_2-M)+(m_3-M)+\cdots\cdots+(m_n-M)=0 \qquad (3)$$

但 $m_1-M=d_1$，$m_2-M=d_2$，………故方程式 (3) 可寫作

$$\Sigma d=0$$

吾人既知各數值對其算術平均數之正負離中差之和爲 0，則可先擇一假定算術平均數(arbitrary origin or assumed mean)，再求各數值對此假定平均數之正負離中差之和，更由此數值以斷定假定平均數與眞實平均數(true mean)之差數。此項差數亦卽各數值對其假定平均數之離中差之平均也。今設 M' 代表假定平均數，$c=M-M'$，又設 d'_1，d'_2，d'_3…………d'_n 代表各數值對 M' 之離中差（卽 $d'_1=m_1-M'$，$d'_2=m_2-M'$，餘類推)，則

$$d'_1=d_1+c,\ d'_2=d_2+c,\ d'_3=d_3+c,\cdots\cdots d'_n=d_n+c$$

$$\Sigma d'=\Sigma d+Nc$$

但 $$\Sigma d=o$$

故 $$\Sigma d'=Nc$$

$$c=\frac{\Sigma d'}{N}$$

已知 M' 及 c 兩者之數值後，則眞實平均數亦可求得，因 $M=M'+c$。其計算手續可用一簡例說明之。

表二十二

算術平均數之計算(簡捷法)

(未經分組之資料)

m	f	d'	
5	1	-15	$M'=20$
15	1	-5	$c=\frac{\Sigma d'}{N}=\frac{+25}{5}=5$
25	1	$+5$	
35	1	$+15$	
45	1	$+25$	$M=M'+c=20+5=25$
	5	$+25$	

以實數為單位

如以各數值對其眞實平均數之離中差作爲標準，則依照假定平均數 20 以計算各數值之離中差時，每一離中差必含有一固定之差誤，此固定差誤卽爲眞實平均數與假定平均數之差數。故離中差之總和中，實含有N個固定差誤，因每項計算時均含此差誤而共有N項也．離中差之和以N除之，卽得此固定差誤之數值，而眞實平均數亦可依此求得矣。

數量資料如依次數分配之形式排列，計算離中差時又以組距爲單位，則平均數之計算手續更將便利不少。在校正差誤時，此眞實平均數與假定平均數之差數仍可化爲原來單位。茲仍以說明算術平均數普通計算法所用之小麥產量數字，說明此簡捷法之應用。

表二十三

1911—1921年北達科多州各郡小麥產量算術平均數之計算

（簡捷法）

組距（每英畝蒲式耳數）	中點 m	次數 f	d'（以組距爲單位）	fd' −	fd' +	M之計算 $M'=9$
						1.各數值對M'之離中差之和：
0— 1.9	1	3	—4	12	…………	$+778$
2— 3.9	3	26	—3	78	…………	-353
						$+425$
4— 5.9	5	78	—2	156	…………	2.c 之計算（以組距爲單位）：
6— 7.9	7	107	—1	107	…………	$c=\frac{425}{569}=.7469$
8— 9.9	9	113	0	…………	…………	3 化 c 爲原來單位：
10—11.9	11	65	1	…………	65	組距＝2
12—13.9	13	40	2	…………	80	c（按原來單位計算）
14—15.9	15	22	3	…………	66	$=.7469\times 2$
16—17.9	17	45	4	…………	180	$=1.4938$
18—19.9	19	41	5	…………	205	4.M之決定：
20—21.9	21	21	6	…………	126	$M=M'+c$
22—23.9	23	8	7	…………	56	$M=9+1.4938$
		569		-353	$+778$	$M=10.4938$蒲式耳.

用簡捷法計算算術平均數之步驟，略如下述：

1. 將數字資料列成次數分配之形式。
2. 就靠近次數分配之中心處選定一組之中點作爲假定平均數。
3. 在次數表中闢一欄，以便將每組各項對於假定平均數之離中差(d')列入，此離中差係以組距作爲單位。在假定平均數所在組內各項之離中差爲 0，其下一組內各項之離中差爲 -1，其上一組內各項之離中差爲 $+1$，餘類推。
4. 將每組之離中差乘以該組之次數，並注意其正負號，其乘積列入 fd' 欄內。
5. 就 fd' 欄內各項求其總和。
6. 將此總和除以次數之和 (N)，所得商數即爲以組距作爲單位之校正數 (c)。
7. 校正數(c)以組距乘之，所得乘積即爲以原來單位表示之校正數。
8. 以此校正數與假定平均數(M')相加，其和數即爲眞實平均數(M)

中位數之決定：根據未經分組之資料

所謂中位數(median)者，在 x 數值各依其大小排列時，其地位最爲居中之一數值也。此值之地位係劃分項數爲兩個相等部分，其中一部分之各數值大於中位數，另一部分小於中位數。此種中位數在許多次數分配中頗爲適用。

當資料未經整理成爲次數分配之形式時，中位數之決定至爲簡易，祇須將資料依其數值之大小順序排列，然後由尺度之一端起依次點算項數，至劃分全部項數爲兩個相等部分處之數值，卽爲中位數之值。茲以簡單數字說明中位數決定之方法。下列數字係七人之每年收入金額。

\$750 \$975 \$1128 \$1450 \$1475 \$1825 \$1950

前項數字之尺度，係自 \$750 起至 \$1950 止，依此尺度排列者共有

七項。在 $1000 之數值上，其一邊有 2 項，另一邊有 5 項，故非中位數。此處中位數之值應爲 $1450，卽七人中一人之收入金額。此數值之兩邊各有 3 項，假定將此項分割爲二，則此點之兩邊各有 $3\frac{1}{2}$ 項。此可用圖四十四表示之。由該圖可見中位數之位置適將項數等分爲二。

當項數爲偶數時，則中位數之決定方法稍異，試用下表數字說明之。表內所列係一九二一年美國二十二州鋸木廠工人之平均工資率。

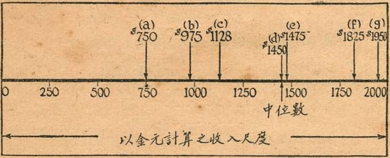
表二十四

1921年美國22州鋸木廠工人之平均工資率(註)

州別	每工人工作一小時之平均工資率
弗基尼阿	$0.586
賓夕法尼亞	.621
泰內西	.627
北卡羅來那	.647
西弗基尼阿	.658
緬因	.686
南卡羅來那	.721
威斯康辛	.729
密雪干	.730
阿拉巴瑪	.749
明尼蘇達	.755
泰克薩斯	.761
佐治亞	.779
阿康薩	.785
密士失必	.798
路易斯安那	.824
佛羅里達	.825
愛達荷	.833
加利福尼亞	.864
蒙大拿	.904
華盛頓	1.045
俄勒岡	1.106

(註)錄自 1923 年一月出版 "Monthly Labor Review"，第 9 頁。

本表之中位數當係兩邊各為十一州之數值，故大於 $0.755(卽明尼蘇達之平均工資率)與小於 $.761(卽泰克薩斯之平均工資率)之間之任何一數值，均可適合於中位數之定義。在此情形下，中位數實無法確定，但通常係將此兩個數值之平均數作為中位數。故在本表中 22 項數字之中位數應為 $0.758。



圖四十四. 根據未經分組之資料決定中位數位置之圖解(七人之個人收入金額)

本例中中位數之值非為任何一州之平均工資率，此種情形在項數為偶數時所習見。又吾人所處理之資料如為間斷數列，則所得之中位數，不特為任何一項之數值所無，且為事實上所不致發現者。例如一九〇〇年至一九二二年紐約市場六十日至九十日期之商業票據貼現率，其中位數當為 4.865%，但此中位數純係假作之貼現率，而為實際市場上所絕無者。

中位數之決定：根據已經分組之資料

中位數之決定，在資料已分組而成次數分配之形式時，其手續亦大致相同，惟資料既已分組，卽未能推知各個項目之實際數值，故決定中位數之手續較為複雜。茲就下列 276 個連續月份內紐約市場 60 日至90日期之商業票據每月平均貼現率，以說明決定中位數之方法。

表二十五

紐約票據貼現率中位數之決定

(1900—1922年六十日至九十日期之雙記名商業票據)

組　距 (以百分率表示之貼現率)	中點 m	次數 f
2.75—3.24	3.00	8
3.25—3.74	3.50	31
3.75—4.24	4.00	52
4.25—4.74	4.50	38
4.75—5.24	5.00	39
5.25—5.74	5.50	42
5.75—6.24	6.00	35
6.25—6.74	6.50	13
6.75—7.24	7.00	5
7.25—7.74	7.50	4
7.75—8.24	8.00	9
		276

$$\frac{N}{2}=\frac{276}{2}=138$$

$$Md=4.75+\left(\frac{9}{39}\times .50\right)$$

$$=4.75+.115$$

$$=4.865$$

在本例中決定中位數，卽為決定其上下兩邊各占 138 項之一項數值。吾人試從尺度下端起逐漸向上移動以點算各組之次數，至第一組(卽包含 2.75 至 3.25 各值之一組)上限時，則已點算 8 項，其餘 268 項尙未點算；至第二組上限時，已點算 39 項；至第四組上限時，已點算 129 項；在第五組上限處則已點至 168 項。因中位數兩邊之項數須各為 138 項，故中位數必在第五組上下兩限之間，然自該組下限 4.75 起移動至何處時，始為中位數之所在點乎？

吾人當可回憶為計算平均數起見，嘗假定每組內各項目有勻整之分配。在點算各組項數至第五組下限時，已點算 129 項，則在該組內 39 項中尙須點算 9 項以補足 138 項之數。吾人既已假定每組內項目分配勻整，則該組內第 9 項之數值，當為橫尺度上相對於該組組距之 $\frac{9}{39}$ 之一點。該組組距為 .50，故 .50 之 $\frac{9}{39}$ 等於 .115。吾人在尺度上自下端起逐漸移動已至 4.75 處，現仍須繼續向上移動 .115 之距離。此點在尺度

上之數值爲4.865,即係使兩邊各占138項之劃分點,亦即中位數之值也.

此項中位數之計算手續,見上列次數表之右方。茲更將中位數之計算步驟簡述於下:

1. 將資料整理成一次數分配之形式。
2. 將總項數以 2 除之,即爲中位數兩邊每邊所應有之項數。
3. 自尺度下端起,將各連續組之項數一一相加,至中位數所在組之下限爲止。
4. 決定該組內所應補湊之項數。此項數與前節已求得之項數相加時須適等於$\frac{N}{2}$項之數。
5. 將所應補湊之項數,除以中位數所在組之項數,此即補湊項數所占該組全距之分數。
6. 以所得分數與組距相乘。
7. 更將此乘積與中位數所在組之下限相加,即得中位數之值。

前述末三項步驟,即係簡單之插補法。

在點算項數時,亦可變其方向,自尺度上端起逐漸向下點算,惟此時前述第 7 項步驟,非加法而爲減法,係從中位數所在組之上限,減去分數與該組組距相乘之積。

如N爲奇數,則$\frac{N}{2}$將爲帶分數,至於中位數之決定手續,仍與前述相同。

四分位數與十分位數

除中位數以外,吾人又常須在 x 尺度上決定將次數分配之全體項數劃分爲其他相等部分之界點。中位數係劃分全體項數爲兩個相等部分,四分位數(quartile)、十分位數 (decile) 以及百分位數 (percentile) 之意義亦與此類似。所謂四分位數者,依其字義而言,當爲尺度上劃分全體項數爲四個相等部分之界點,十分位數爲劃分全體項數爲十個相

等部分之界點，百分位數為劃分全體項數為百個相等部分之界點。故第一四分位數為將尺度上全體項數分為$\frac{1}{4}$與$\frac{3}{4}$之比例之界點，第二四分位數與中位數之數值相同，第三十分位數係將尺度上全體項數分為$\frac{3}{10}$與$\frac{7}{10}$之比例之界點。以上所述，在點算項數時，皆自尺度之下端起。

例一．　第一四分位數(Q_1)之計算：貼現率(根據表二十五)

$$\frac{N}{4}=69$$

$$Q_1=3.75+(\frac{30}{52}\times.50)=4.038$$

例二．　第四十分位數(D_4)之計算：貼現率(根據表二十五)

$$\frac{N}{10}=27.6 \qquad D_4=4.25+(\frac{19.4}{38}\times.50)$$

$$\frac{4N}{10}=110.4 \qquad =4.505$$

依圖決定中位數、四分位數、十分位數與百分位數之方法，當於後文論述之。

衆數之決定

衆數者，為一次數曲綫上相對於最高縱坐標之 x 變量之數值也。衆數值之概念，極易悟解，因其為最普通之值。例如最普通之工資，最普通之收入，最普通之身材高度等。此值為次數最為密集之一點，故德人費希奈爾氏(Fechner)嘗稱此種平均數為dichtester wert，英譯為thickest value，意謂次數最為密集之數值也。據此定義，則衆數之特質已能確切表現，但有時真實之數值頗不易決定。通常從事統計工作時，欲求此衆數，祇能得其近似值，蓋在應用方面此近似值通常已够準確也(註)。

決定衆數近似值之方法，可就下列次數分配說明之。

(註)欲求更精確之衆數數值，其方法詳見第十五章末之附註。

表二十六

五釐債券之次數分配

(此表係根據 1923 年十月三十一日紐約證券交易所所開債息五釐鐵路及實業債券之行市編製)

價格組距	中點 m	次數 f
$71.5 以下		30
71.5—73.49	72.5	2
73.5—75.49	74.5	5
75.5—77.49	76.5	5
77.5—79.49	78.5	5
79.5—81.49	80.5	12
81.5—83.49	82.5	13
83.5—85.49	84.5	15
85.5—87.49	86.5	11
87.5—89.49	88.5	17
89.5—91.49	90.5	17
91.5—93.49	92.5	33
93.5—95.49	94.5	45
95.5—97.49	96.5	51
97.5—99.49	98.5	59
99.5—101.49	100.5	23
101.5—103.49	102.5	9
103.5—105.49	104.5	1
		353

表內 71.5 以下之 30 項，其數值有極大之離散度，且該組祇有上限而無下限，故根據此表事實上無法計算其算術平均數。遇此情形，平均數以採用衆數爲宜。

在各組中，以 97.5—99.49 一組所包含之項數爲最多，該組似爲衆數組，該組之中點 98.5 似可作爲衆數之近似值；但若分組變更，則可得絕不相同之衆數值。試將原有債券市價，依照大小不同之數個組距分別製表，可得如下之結果（此處僅舉示中間各組之次數，已足證明衆數值

隨組距大小而有變更，似無須將全表各項一併列入)。

(a)

組距＝1

組距	f
93.5— 94.49	17
94.5— 95.49	28
95.5— 96.49	26
96.5— 97.49	25
97.5— 98.49	25
98.5— 99.49	34
99.5—100.49	18
100.5—101.49	5

(b)

組距＝2

組距	f
94.5— 96.49	54
96.5— 98.49	50
98.5—100.49	52

(c)

組距＝3

組距	f
91.5— 94.49	50
94.5— 97.49	79
97.5—100.49	77

以 1 爲組距，得衆數值爲 99，以 2 爲組距，並使各組組限稍異於表二十六所列，得衆數值爲 95.5，以 3 爲組距，得衆數值爲 96。如分組再予變更，則仍可得其他不同之衆數值。由此觀之，衆數似爲一種不可捉摸之平均數，蓋根據同一資料，衆數值常因組距之大小與組限位置之變更而有差異。

惟此困難之發生，率由於事實上取樣不能過多所致。吾人果能將取樣之範圍爲無限制之擴大，而將包含之項數爲無限制之增加，則必有極大之次數密集於某一數值上，而眞實之衆數乃可由此確定，蓋若包含之項數較多，則組距縮小時，衆數之近似值漸與眞實衆數接近。組距過大，詳盡之事實每爲所掩蔽；反之，組距縮小時，所表現之事實必較詳盡，而次數分配之眞相乃可明白顯示；但事實上取樣不能過多，故組數增加時，各組間之次數每呈間斷與不規則之狀態，因之分配遂欠對稱與勻整，而最大集中點之實際地位，乃亦難於決定。前述根據債券市價所製各表，卽爲其一例。

如應用數學方法，則項數不爲無限制之增加，亦可求得眞實衆數之值。關於曲綫修勻之手續，前嘗述其梗概。修勻曲綫之一法，可依該曲綫所根據之次數分配資料，配以適當之理想次數曲綫。此曲綫在理論上係表示組距無限縮小與項數無限增加時所應有之次數分配。在此理想次數曲綫上所得最高縱坐標之 x 值，卽爲眞實衆數。

在實際應用方面，僅求衆數之近似值，亦已適用，而可採用更簡單之各種方法以求此近似值。前節所述以次數最大一組之中點作爲衆數，雖係一粗略之方法，但若分組時果能注意前述各項分組原則，則依此方法所得結果，大體尚不致發生重大差誤。

如次數分配尙屬勻整，可就衆數組內用插補法以推定其衆數值，由此所得之值較以該組中點作爲衆數值者，爲更與眞實衆數相近似。茲仍以表二十六債券價格爲例以說明之。按該表內衆數組兩邊之次數分配並不對稱，衆數組之中點爲 98.5，較衆數組略小一組之中點爲 96.5，包含價格 51 項，較衆數組略大一組之中點爲 100.5，包含價格僅 23 項；而次於衆數組上下鄰接兩組之各組，其兩邊次數分配亦欠平衡，有偏向於小於衆數組之一邊之趨勢。以前計算算術平均數與中位數時，嘗假定每組內各項數值在該組上下兩限間分配勻整，但本例中衆數組內各數值之分配並不勻整，由衆數組以外各組之次數分配觀之，似可推定衆數組內各項之次數有偏向於該組下半部卽 97.5 至 98.5 間之可能，故此衆數大抵在中點 98.5 以下，而未見適在中點。吾人試用加權法以確定衆數在該組內之位置，並假定此衆數被拉向尺度下端之力爲 51（卽係較衆數組略小一組之次數），被拉向尺度上端之力爲23（卽係較衆數略大一組之次數）。此可用下列符號，以公式表示之。

l ＝衆數組之下限。

f_1＝較衆數組略小一組之次數。

f_2＝較衆數組略大一組之次數。

i ＝組距。

插補公式爲

$$Mo=l+\frac{f_2}{f_2+f_1}\times i$$

茲應用此公式以計算表二十六債券價格之衆數值，則得

$f_2=23+9+1=33$

$$Mo=97.5+\frac{23}{74}\times 2=97.5+.62=98.12$$

如權數（以f_1及f_2代表者）根據大於衆數組與小於衆數組之鄰接各兩組或各三組之總次數，則所得衆數值更爲近似。在本例中如根據衆數組兩邊各三組之總次數計算，則得債券價格之衆數值爲 97.91。

衆數之位置，有時因次數並不如前例之密集於一點，而分散於各點，以致不易確定者。如表九所示以 \$.25 爲組距之工資分配，具有兩個明確之衆數點。此種分配謂之雙峯分配 (bi-modal distribution)，因圖示時其次數曲綫呈現兩個高峯也。如全部資料確屬同一性質，則此種欠勻整之分配，係由於所取資料不足與所採分組方法欠當所致。如所取樣本包含項數不多，而組距失之過小，則往往有得兩個衆數點之可能。遇此情形，吾人可變更組限與增加組距，使次數分配內祗存一個衆數組，而由此決定衆數之近似值。此與包含項數極多時求眞實衆數之手續正相反。當項數爲無限多時，則組距可縮至無限小，以便求得眞實衆數之值；然在包含項數無多時，則欲決定次數最爲密集之一點，必須將組距增大，以消除因所取樣本不足而發生次數分配不規則之狀態。

如變更組距與組限，次數分配仍具雙峯形態，此或由於資料性質不一，誤將兩個性質不同之分配，混爲一個所致。此種錯誤在研究生物測量中所習見，例如誤認兩種不同之動物爲一種，因之所得分配常現雙峯狀態，而表示其中實包含兩種不同之分配。如資料性質互異，則次數分配失其整個之意義，不僅在研究經濟統計時爲然，即研究其他統計時，亦莫不皆然。

由算術平均數及中位數決定衆數之方法

此外求衆數近似值之方法，亦有根據算術平均數、中位數及衆數三者間之關係以計算者。次數分配完全對稱時，算術平均數、中位數及衆數三者之值全相等；次數分配如不對稱，此三者之值各不相同；如次數

分配不對稱之程度不甚顯著,則三者之間恆有一定之關係:在次數曲綫圖中,衆數與算術平均數之相距最遠,而中位數與算術平均數之距離,約等於算術平均數與衆數距離之三分之一。但若次數分配極不對稱,則三者之間無此關係。在次數分配之不對稱程度不甚顯著時,三者之中如已知任何兩種平均數之值,即可由三者之關係以求得第三種平均數之值。惟此法在實際應用上僅可適用於衆數值之決定,因其餘兩種平均數之值,可採用其他更精確之計算方法也。即在求衆數值時,亦必在其他更精確之計算方法不能應用時,始可依此法計算。

下列公式係根據算術平均數、中位數及衆數三者之關係而得。

$$Mo = Mean - 3(Mean - Md)$$

茲將表十二電桿木之資料代入公式,得下列結果:

$$Mo = 9.33 - 3(9.33 - 9.015) = 8.385$$

此值略小於衆數組之中點 8.5,而同時亦小於按照加權法在衆數組內插補所得 8.49 之數值(係用衆數組左右鄰接各四組之次數作爲權數者)。

吾人有須注意者,上列各個衆數值孰爲準確,殊無標準可言。前述關於決定衆數之各種方法,亦僅能求其近似值,此點在解釋或應用其結果時,切宜注意及之。

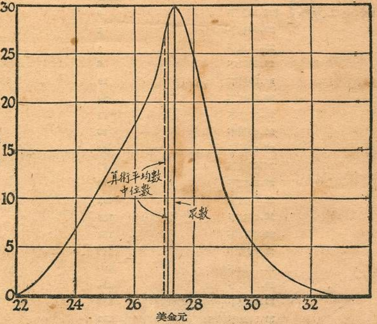
按圖決定衆數中位數四分位數與十分位數之方法

吾人如將按圖確定前述各種量數之方法,略加討論,則對於次數曲綫與累積次數曲綫可得更確切之瞭解。

在普通次數曲綫圖中,衆數之值頗易確定,蓋依衆數之定義,其值當爲 x 尺度上相對於曲綫最高縱坐標之一點。在次數多邊圖中,亦可得粗略之衆數值,此值即爲次數最多一組之中點。再如欲得更迫近於眞實衆數之值,則須用觀察法或數學方法,在修勻之曲綫圖中求之。圖四

十五係說明按圖決定衆數之方法。圖中曲綫(係用觀察法修勻者)係根據表八之工資資料繪成者。在橫尺度上與曲綫最高縱坐標相對之值為$27.50，此僅係衆數之近似值，可與按照加權法求得之值 $27.69 及由算術平均數與中位數計算所得之值$27.3470比較之。

中位數與算術平均數之位置，已在該曲綫上記明。前嘗論及次數分配不對稱(偏斜)程度不甚顯著時，算術平均數、中位數及衆數三者間恆有一定之關係：中位數係在算術平均數與衆數之間，其與算術平均數之距離，約等於算術平均數與衆數兩者間距離之三分之一。證諸本例，此種關係確見存在，故在此修勻之曲綫圖中，可求得衆數之近似值，惟因原有資料欠整齊，故用觀察法以修勻曲綫，自不免稍覺武斷耳。



圖四十五．　按某週工資分組之工人次數分配之修勻曲綫
表示算術平均數、中位數與衆數三者關係之圖解

圖四十六係根據同一資料，按照表二十七之分組所繪成之累積次

數曲綫。此累積次數曲綫在某距離內之峭度 (steepness), 係隨該距離內之項數多寡而定。該曲綫其始緩漸上升，繼而上升較速，最後在曲綫之上端上升又轉緩漸，故其衆數之値，當爲橫尺度上相對於峭度最高之一點，在此點上次數增加最速，亦卽次數分配中項數最爲集中之一點也。在修勻次數曲綫中，此衆數値可就曲綫上斜度最高一點 (卽轉向點 point of inflection) 之 x 値決定之，而在本例累積次數曲綫中，亦可用此法以推定其衆數之値約爲 \$27.50。

表二十七

某製造廠工人之累積次數分配 何以下

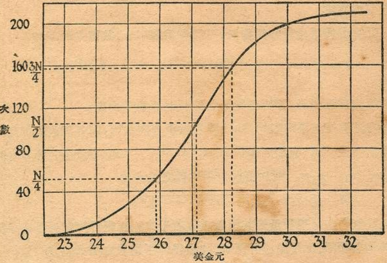
(按照某週工資數分組)

不滿22.50者

某週工資數	工人數(次數)
\$22.50以下	0
23.00以下	1
23.50以下	5
24.00以下	8
24.50以下	19
25.00以下	29
25.50以下	41
26.00以下	56
26.50以下	78
27.00以下	98
27.50以下	122
28.00以下	152
28.50以下	169
29.00以下	186
29.50以下	193
30.00以下	199
30.50以下	204
31.00以下	208
31.50以下	209
32.00以下	209
32.50以下	210

至於中位數、四分位數及十分位數之値，亦可在累積次數曲綫圖中

得。此種曲綫一經修勻，則應用插補法可得良好之結果。又若圖之尺度够大時，用插補法且可得精確之數字。在縱尺度上（卽表示累積次數之尺度）確定自底綫起約等於$\frac{N}{2}$距離之一點，由此點起繪一水平綫，與累積次數曲綫相交，此交點之橫坐標，卽為中位數之數值。在決定此值時，可由交點繪一縱綫，垂直於 x 軸上量其橫坐標卽得。圖四十六說明此法之應用。用此法求得中位數之數值為 $27.125，如逕用插補法則得 $27.1458。至於四分位數亦可依照此法求得，其縱尺度應劃作四等分，更由此分點各繪水平綫交於累積次數曲綫可矣。



圖四十六．　根據某週工資之累積次數曲綫決定中位數及四分位數之圖解

關於平均數之選擇，在某種情形之下，尤其在求百分率 (rates) 或比率 (ratios) 之平均，而非實際數量之平均時，前述各種平均數皆不適用，惟幾何平均數及倒數平均數則較為適宜，故有研究之必要。

幾何平均數

幾何平均數係 N 項數字相乘之積之 N 方根，其值以公式表示之如

下:

$$M_g=\sqrt[n]{a_1\cdot a_2\cdot a_3\cdots\cdots a_n}$$

例如2,4,8三個數字之幾何平均數爲

$$M_g=\sqrt[3]{2\times4\times8}$$

$$=\sqrt[3]{64}$$

$$=4$$

依上式計算,可知一個數列中如有一項數值爲零時,其幾何平均數之數值亦爲零。

計算幾何平均數如用對數,可節省手續不少。其算式如下:

$$Log\,M_g=\frac{log\,a_1+log\,a_2+log\,a_3+\cdots\cdots+log\,a_n}{N}$$

幾何平均數之對數,等於各數量之對數之算術平均數。

在計算幾何平均數時,如欲將各數量分別加權,則各個權數即爲各數量之指數。今設權數之和爲 N,a_1,a_2,a_3,……a_n 各數量之權數各爲 w_1,w_2,w_3,……w_n,則加權幾何平均數之公式爲

$$M_g=\sqrt[n]{a_1{}^{w_1}\cdot a_2{}^{w_2}\cdot a3^{w_3}\cdot\cdots\cdots a_n{}^{w_n}}$$

按此式無異將每項數量依其權數重複若干次,正與計算加權算術平均數時情形相似。如用對數計算,則加權幾何平均數之公式當爲

$$Log\,M_g=\frac{w_1log\,a_1+w_2log\,a_2+w_3log\,a_3\cdots\cdots+w_nlog\,a_n}{N}$$

茲就下表數字以說明幾何平均數之計算方法。此表係紐約證券交易所所開一一五種紅利優先股票價格之分配。表內價格均係一九二二年八月二十六日之收盤價。

表二十八

優先股票價格幾何平均數之計算

組距	中點 m	次數 f	組距中點之對數 $\log m$	組距中點之對數×次數 $f\log m$
\$35—\$44.9	\$ 40	1	1.602060	1.602060
45— 54.9	50	6	1.698970	10.193820
55— 64.9	60	8	1.778151	14.225208
65— 74.9	70	5	1.845098	9.225490
75— 84.9	80	14	1.903090	26.643260
85— 94.9	90	22	1.954243	42.993346
95—104.9	100	27	2.000000	54.000000
105—114.9	110	18	2.041393	36.745074
115—124.9	120	14	2.079181	29.108534
		115		224.736792

$$Log\ M_g=\frac{224.736792}{115}\qquad Log\ M_g=1.954233$$

$$M_g=\$\ 90.00$$

幾何平均數之特質

幾何平均數之特質若何，可就該平均數與各數量間之關係推求而得。

在一個數列中，其算術平均數恆可替代各數量，而使各數量之原有總和不變。例如2,4,8三個數字之和爲14，其算術平均數爲$4\frac{2}{3}$，如將原有三個數字各以此平均數替代之，則其總和仍爲14，此爲算術平均數之特質。至於幾何平均數之特質則爲在一個數列中，其幾何平均數恆可替代各數量，而使各數量之原有乘積不變。例如2,4,8三個數字之乘積爲64，其幾何平均數爲4，如將原有三個數字各以此平均數替代之，則其乘積仍爲64。

復次，算術平均數之特質，爲大於平均數各項離中差之和，適等於小於平均數各項離中差之和（不計其正負號）；換言之，卽大於平均數各項對於平均數差額之和，與小於平均數各項對於平均數差額之和，彼此相等也。幾何平均數之特質，則爲大於幾何平均數與小於幾何平均數兩

邊各項對於該平均數所得比率之乘積彼此相等;換言之,卽幾何平均數對於較小各項所得比率之乘積,等於較大各項對於幾何平均數所得比率之乘積。例如3,6,8,9之幾何平均數爲6,可得下列等式:

$$\frac{6}{3}\times\frac{6}{6}=\frac{8}{6}\times\frac{9}{6}$$

上述第二點,尤爲幾何平均數最主要之特質。此係求平均比率之一種方法,其在經濟統計方面之主要用途,係供編製物價指數之用,因指數所注重者爲物價變動之百分率。物價自50漲至100,其重要程度正與100漲至200相等,惟此種相等非算術平均數所能表示。如用算術平均數計算,則前者所增加之絕對數爲50,而後者所增達100,是不啻予後者以倍重之權數。吾人習常所舉之例爲兩項物價之變動,其一自100漲至1000,計漲十倍,其另一項自100跌至10,僅及原價之十分之一。1000與10之算術平均數爲505,其幾何平均數爲$\sqrt{1000\times10}$,卽100。由此幾何平均數可知一漲一跌程度相等,故相抵銷,而算術平均數乃高至505,顯屬謬誤,其不適於求平均比率之用,當可想見。此點當於第六章物價指數內再爲詳述。

前章內所述用對數尺度製圖對於某種目的所具有之優點,實與幾何平均數之用途有關。幾何平均數亦有稱爲對數平均數者,則以幾何平均數之對數,卽係各數量對數之算術平均數之故。在計算變動百分率之平均時,所注重者爲比率而非絕對差數,故幾何平均數最爲適用。

在計算本金按照複利率累積之本利和時,尤非採用幾何平均數不可。今設P_o爲本金,P_n爲若干時期後之本利和,r爲利率,n爲若干時期內之年數,則n年後按照每年複利一次,P_o累積所得之本利和,當如下列方程式所示:

$$P_n=P_o(1+r)^n$$

自上式化爲

$$r=\sqrt[n]{\frac{P_n}{P_o}}-1$$

例如本金1000元按複利計算，十二年後之本利和爲1600元，較本金增60%，如按算術平均數計算，則利率爲5%，但此非本利累積之百分率，其實際利率應爲：

$$r=\sqrt[12]{\frac{1600}{1000}}-1$$

$$=\sqrt[12]{1.60}-1$$

$$=1.04-1$$

$$=\ .04\text{卽}4\%$$

此外凡欲求增減百分率之平均時，當發生同樣問題，如用算術平均數必得錯誤之結果。

幾何平均數表示集中趨勢之討論

或問何種次數分配，其集中趨勢應以幾何平均數代表之乎？此一問題如加以討論，頗饒興趣。大抵當絕對數值繪於算術尺度上，其次數分配頗具對稱狀態時，則算術平均數恆優於幾何平均數；但若絕對數值繪於算術尺度上，其曲綫並不對稱，而繪於幾何尺度上可減除其不對稱狀態，而成一對稱曲綫時，則幾何平均數優於算術平均數。在後一種分配中，呈對稱形態者，爲集中趨勢兩旁之相對離中差 (relative deviation)，卽以比率表示之離中差，而非絕對離中差(absolute deviation)。此種次數分配如欲確切表示其集中趨勢，應採用各數量對數之算術平均數（此值卽爲絕對數量之幾何平均數之對數，前已述及）以代表之。由各對數所得之曲綫，當在幾何平均數之對數之兩邊成對稱狀態。例如各項物價漲落百分率之對數所繪成之次數曲綫，必在其幾何平均數之對數之兩邊呈對稱狀態。但此種漲落百分率，其眞數之分配狀態並不對稱，大於算術平均數各項之全距，必遠較小於算術平均數各項之全距爲

大(註一),此因物價上漲時,可漲至1,000%或1,000%以上,而在物價跌落時,則不能跌去100%以上也。關於此點,在第六章物價指數內當有更詳盡之討論(註二)。

此種根據各個數值之對數而不根據其眞數繪製之次數曲綫,對於個人收入分配之資料,尤爲適用。此項資料如按眞數繪製,則以收入分配之全距過大,每使曲綫各部份之特性不能在圖中表現,但若繪於雙對數圖中(即 x 及 y 皆用對數繪製者),即能避免此種困難,而可揭示全部分配之眞象及各部分間相互之關係,同時並可表明爲眞數尺度所不能顯示之某種要點。此種雙對數圖似可使曲綫中大於衆數之部分,化爲平滑而成直綫,故倡用此雙對數圖以表述個人收入資料之巴拉圖氏(Vilfredo Pareto)嘗賴以發明關於收入分配之巴拉圖法則(Pareto's Law)。當美國經濟研究所精密研究美國人民收入分配之結果後,對於根據巴拉圖法則所推演而得之各種結論雖置懷疑,而對於雙對數尺度之用於收入分配之資料,固仍承認其具有價值也。關於此種次數曲綫,在美國經濟研究所所編美國人民收入("Income in the United States")報告書中嘗作具有興趣之論述。

(註一)參看圖四十九及圖五十。

(註二)美國 C. M. Walsh 氏所著 "The Problem of Estimation" 第35頁對於平均數之選擇,嘗定有下列標準:

(a) 某數列各項中其最大值或最小值之限度不可估量或不能確定時,以採用算術平均數爲宜。

(b) 某數列各項中之最小值可確定爲0或大於0,而其最大值之限度不可估量或不能確定時,以採用幾何平均數爲宜。物價變動之百分率卽其一例,故 Walsh 氏認爲編製物價指數,以採用幾何平均數最爲適當。

(c) 在實際應用方面,或由於資料之性質關係,最大值與最小值之限度皆存在,而不能應用前項標準時,則必須先行研究資料之實際離散度。如其衆數與算術平均數較爲相近,則應採用算術平均數,如其衆數與幾何平均數較爲相近,則應採用幾何平均數。

倒數平均數

倒數平均數(harmonic mean)為平均數中用途較為狹隘者，惟在某種資料中如採用倒數平均數，則可避免差誤之發生。在計算時間速率之平均時，非用倒數平均數不可，而對於他種特殊資料，亦常有應用之必要。下述之例係說明此種平均數之應用方法：

例如某汽車行駛4英里之距離，其速率為每小時20英里，嗣又續駛4英里，其速率為每小時30英里，欲求該汽車之平均速率。按上述兩項速率之算術平均數為每小時25英里，但照此計算係屬錯誤。該汽車係先按每小時20英里之速率行駛12分鐘，嗣按每小時30英里之速率續駛8分鐘，是兩次共計行駛8英里之距離，費時共20分鐘，故其平均速率當為每小時24英里。此實即20及30之加權平均數，前者之權數為12，後者權數為8也。但若按照兩個速率計算其倒數平均數，可得相同之結果。若干數值之倒數平均數，即為各數值倒數之算術平均數之倒數。今設r_1 r_2 ……r_n代表所欲平均之各個速率，則求倒數平均數(H)之公式為

$$\frac{1}{H}=\frac{\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3}+\cdots\cdots\frac{1}{r_n}}{N}$$

用上項數字代入公式，得

$$\frac{1}{H}=\frac{\frac{1}{20}+\frac{1}{30}}{2}$$

$$=\frac{5}{120}$$

$$H=24$$

如欲求a,b兩數之倒數平均數，則由公式$H=\frac{2ab}{a+b}$計算，手續更為簡便。如有a,b,c三個數字，則可用公式$H=\frac{3abc}{ab+ac+bc}$。此兩公式與前述第一公式相等，但若數列中項數較多，則計算倒數平均數

時，如用製就之倒數表，手續可便利不少(註)。

倒數平均數在分析經濟資料方面之用途，可以物價爲例以說明之。例如物品論價有按"每元若干件"者，欲求此種價格之平均數，應用算術平均數每得錯誤之結果。今設有高下不同之物價三項，一爲"每元四件"，一爲"每元五件"，一爲"每元二十件"，問平均每元售若干件(卽平均價)？根據上述數字(4,5,20)求得算術平均數爲$9\frac{2}{3}$，驟視之似爲平均每元所售件數，卽每件平均價爲10.34分；但按諸實際，原有之三項價格，亦卽每件25分、每件20分及每件5分之謂，其算術平均數爲$16\frac{2}{3}$分，如易爲與原價同一形式之價格，此平均數卽爲每元6件。如欲予每個物價以相等之權數，則後一平均數係準確之答案。蓋按照"每元售若干件"之價格所求得之算術平均數，乃係加權平均數，對於每元所售件數較多之物價，不啻卽予以較大之權數。上項準確之結果，亦可直接按照原有物價用倒數平均數求之，蓋4,5,20三個數字之倒數平均數爲6，卽平均每元所售之件數也。

各種平均數之相互關係

如根據一個數列同時求得各種平均數，則各平均數間恆有各種關係之存在。

1. 在完全對稱之次數分配圖中，算術平均數、中位數與衆數三者同在一點。
2. 在偏態不甚顯著之次數分配圖中，中位數之位置處於算術平均數與衆數之間，中位數與算術平均數之距離約等於算術平均數與衆數距離之三分之一。故此種次數分配，其三種平均數間之關係約如下式所示：

(註)一九一九年紐約 Spar and Chamberlain 公司出版之 "Barlow's Tables of Squares, Cubes, Square Roots, Reciprocals" 一書頗切實用。

$$Mo = M - 3(M - Md)$$

3.任何數列之算術平均數，必大於其幾何平均數。

4.任何數列之幾何平均數，必大於其倒數平均數；惟若該數列中所有數量各各相等，則算術平均數、幾何平均數與倒數平均數三者亦必相等。

5.任何兩數值之幾何平均數，等於倒數平均數與算術平均數兩者之幾何平均數。例如2與8之倒數平均數爲$3\frac{1}{5}$，算術平均數爲5，幾何平均數爲4，而$3\frac{1}{5}$與5之幾何平均數亦爲4也。但數列包括項數在兩項以上者，不在此列。

6.統計資料之離散度(dispersion)如合乎算術定律，則衆數及中位數兩者之值，與算術平均數相差較近，而與幾何平均數相差較遠；但若其離散度合乎幾何定律，則衆數及中位數兩者之值，與幾何平均數相差較近，而與算術平均數相差較遠。

各種主要平均數之特質

算術平均數：

1.算術平均數之值係受數列中每個數量之影響，故遇有極端數量存在時，頗嫌其所受影響過鉅，而不適於某種用途。

2.算術平均數易於計算，且在任何數列中其值皆可確定。

3.算術平均數係由計算而得，故可用代數方法處理之。

中位數：

1.中位數之值不受極端數量之影響。

2.統計事項之不能以數量表示者，仍可求得其中位數。

3.凡統計資料不齊備，而能確知其各項之總項數與其大概之位置以及次數分配中間各項數量之大小時，其中位數亦可決定。

4.中位數非若算術平均數、幾何平均數與倒數平均數三者之可用代數方法處理者。

衆數:

1.衆數之值不受極端數量之影響。

2.衆數之近似值易於確定,但欲決定眞實衆數,則計算手續頗繁。

3.次數分配如包含項數無多,集中趨勢又不顯著,則其衆數無意義。

4.衆數之位置居於次數分配之密集點,故允爲分配中之典型數值。

5.衆數不能用代數方法處理。

幾何平均數:

1.幾何平均數所受極端數量之影響,較算術平均數爲小。

2.在任何數列中,幾何平均數之值皆可確定。

3.幾何平均數對於等比之變動視爲同等重要,故欲求數量變動百分率之平均,或比率之平均者,非用幾何平均數不可,而用以求物價變動之比率,尤爲適當。

4.幾何平均數可用代數方法處理。

倒數平均數:

1 倒數平均數適用於平均時間速率與類似性質之事物。在經濟統計方面,亦有用以計算物價指數者。

2.倒數平均數因計算費時與意義不通俗,故在普通統計分析方面每棄而不用。

3.倒數平均數可用代數方法處理。

上述各點係欲說明每種平均數各有其特殊用途。爲適應某種目的或在某種情形下,一種平均數常較優於其他各種平均數,吾人應先明瞭每種平均數之特質及其缺點,始可知所取捨。惟欲將次數分配作詳盡之表述,常須同時斷定二三種主要平均數及其他各種統計量數之數值。算術平均數或爲最切實用之一種平均數,因其計算簡便,且可適用代數方法處理,其意義旣確切,又通俗,故極合於統計上之應用,惟其用途並不普遍,應在適當情形之下始可採用之耳。幾何平均數之優點漸爲一般所

瞭解,故在統計方面之應用亦日廣。平均數之選擇對於統計分析頗關重要,學者不可不注意也。

參 考 書

Bowley, A. L. "Elements of Statistics"(82—109,138—139).

Chaddock, R. E. "Principles and Methods of Statistics"(Chaps. VI, VII, VIII).

Jones, D. C. "A First Course in Statistics"(22—41).

Kelley, Truman L. "Statistical Method"(44—68).

King, W. I. "Elements of Statistical Method"(121—140).

Pearl, Raymond. "Medical Biometry and Statistics"(264—272).

Rugg, H. O. "Statistical Methods Applied to Education"(97—148).

Secrist, Horace. "Introduction to Statistical Methods"(234—292).

Walsh, C. M. "The Problem of Estimation"(1—67).

Yule, G. U. "An Introduction to the Theory of Statistics"(106—132).

Zizek, Franz. "Statistical Averages"(82—109).

第五章　次數分配之表述：離散度與偏斜度之量數

在前數章內吾人首先論及大量資料之簡縮方法，使化為簡單形式，俾全部資料之特性可以易於確定；次又論及表述資料特性之方法。當資料已組成次數分配之形式，則前項目的業已達到；迨已求得一單純數值即平均數以代表次數分配之集中趨勢後，則後項目的亦已達到一部分。但任何一種平均數不能單獨表述次數分配之全部情形，如欲測度次數分配之主要特性，或與另一次數分配相比較，尚須另求三種數值而後可。其第一種為測度次數分配中各項與其中心值之相離程度者，即離散度（"scatter", variation or dispersion）之量數是也；第二種為測度中心值兩邊之分配狀態為平衡或不平衡者，即測度次數分配對稱程度（degree of symmetry）之量數是也；第三種為測度在衆數值上次數之密集程度者，即測度次數分配之峯度(kurtosis)之量數是也。本章所欲論述者，為測度離散度與偏斜度之各種方法，至於測量峯度之方法，容後討論。

離散度之性質與意義

數量資料之呈現離散現象，已在前文提及，此種現象對於統計分析上之關係，亦已約略述及。吾人實際所蒐集之各種數量資料，無論其屬於社會、生物或經濟方面，皆具有離散之特質，即各個數量大小不同是也。此各個數量不同之特質實與各個數量類似之特質同一重要。在生物方面，變動之現象為進化過程之一要素。今若吾人欲測驗某種族生理方面之特質，例如體高，不能以測量各人之體高為已足，而又須測度彼等之平均離中差。又如一國人民之平均收入額固屬重要，但社會各階級間收入差異之情形尤關重要。又如物價之變動，每足阻害經濟制度之正常運行，致物價制度下各組織份子因其所受變動之影響不同，而有備受困

苦者，亦有不勞而獲者，此非由於物價水準之一般的變動，而實由於各個物價變動程度不一所致也。

平均數之本身無甚意義，吾人必須同時求得次數分配之離散度以補充之。如次數分配之離散度甚大，且無顯著之集中趨勢，則其平均數毫無價值；離散度愈小，平均數愈有意義。故欲表述某一次數分配，或與另一次數分配比較，除測度集中趨勢外，又須測度其離散度以補充之。

絕對離散度之量數

離散度量數有用原來資料所用之數量單位表示者，亦有用抽象數字而與原來單位絕無關係之百分率表示者。用原來單位所測度者爲絕對離散度(absolute variability)；用抽象數字表示者爲相對離散度(relative variability)。在比較兩種次數分配之離散度時，後者較前者更爲適用。茲先討論絕對離散度之量數。

全距

測量離散度之粗略方法，係以全距(range)爲標準。全距者，即次數分配中最小一項數值與最大一項數值之絕對差數也。表三十所列係一八八二年至一九一三年倫敦對紐約之每月匯價。表內所包含最小一項之原來匯價爲$4.83，最大一項爲$4.908，故該次數分配之全距爲 $4.908－$4.83，即$.078。在尺度上此$.078 之距離內包含次數分配中之一切項數。如未知該兩項之原來數字，則其全距亦可自次數表內估計而得，即以分配中最小一組之下限與最大一組之上限相減之差數作爲全距。依此所得表三十之全距爲 $.085。

由是知全距之大小，全視兩極端之數值而定。如遇一項數值極大或極小，則全距之數值必大受影響。全距之數值因其太不穩定，且不能代表各項之實際分配狀態，故統計方面不常採用。全距恆有用以測度股票

市價之變動者，但是否合於此用，不無疑問耳。

平均差

測量平均數兩邊各項離散度比較準確之量數，尚有所謂平均差（mean deviation）者。平均差之計算，係先求各項對於平均數之差數（即離中差），而後求其平均差數。其計算方法可就下列簡例說明之：

表二十九
平均差之計算

m	f	d
3	1	6
6	1	3
9	1	0
12	1	3
15	1	6

$M=9$

$M.D.=\frac{18}{5}=3.6$

前項次數分配之平均數（在本例中算術平均數與中位數之值相等）爲9。先將各項之離中差，不問其符號爲正爲負，一一相加，再將所得和數除以項數，卽得該分配之平均差。其計算手續可用下式表示之：

$$M.D.=\frac{\Sigma d}{N}$$

如用文字說明，則一個數列之平均差卽係該數列內各項對於平均數（或爲算術平均數，或爲中位數）之離中差之算術平均數也。在各項離中差相加及平均時，其符號之爲正爲負，均所不問。至於計算各項離中差所根據之平均數，或爲算術平均數，或爲中位數，實際上無大出入，惟理論上似以後者爲宜，則因平均差之由中位數計算者，其數值爲最小故也。

表三十係說明統計資料已組成次數分配表時，計算平均差之方法。吾人猶憶以前根據次數分配表計算算術平均數時，嘗假定每組內各項數值密集於其組中點，換言之，卽每組內各項之數值，皆可視爲等於各

該組之中點；今於平均差之計算，亦作同樣之假定。

表三十

1882—1913年倫敦對紐約匯價平均差之計算

（本表數字係根據每月月初之匯價）

組　距	組中點 m	次數 f	對於假定平均數之離中差 d'	fd'	
$4.8275—$4.8324	$4.830	1	$.040	$.040	假定平均數＝$4.8700
4.8325— 4.8374	4.835	6	.035	.210	中位數＝ 4.8721
4.8375— 4.8424	4.840	11	.030	.330	差　數＝$ .0021
4.8425— 4.8474	4.845	21	.025	.525	各項離中差須補足$.0021
4.8475— 4.8524	4.850	23	.020	.460	者196項
4.8525— 4.8574	4.855	24	.015	.360	各項離中差須減去$.0021
4.8575— 4.8624	4.860	25	.010	.250	者188項
4.8625— 4.8674	4.865	40	.005	.200	淨計：離中差須補足$.0021
4.8675— 4.8724	4.870	45	....	....	者8項
4.8725— 4.8774	4.875	49	.005	.245	各項對於假定平均數之離
4.8775— 4.8824	4.880	35	.010	.350	中差之和＝$5.020
4.8825— 4.8874	4.885	45	.015	.675	校正數(8×$.0021)＝
4.8875— 4.8924	4.890	33	.020	.660	$.0168
4.8925— 4.8974	4.895	16	.025	.400	各項對於中位數之離中差
4.8975— 4.9024	4.900	8	.030	.240	＝$5.020＋.0168
4.9025— 4.9074	4.905	1	.035	.035	＝$5.0368
4.9075— 4.9124	4.910	1	.040	.040	$M.D.=\frac{\$5.0368}{384}$
		384		$5.020	＝$.01312

由此表計算平均差，亦可採取表二十九之計算方法，先求每組中點對於中位數之差數，乘以各該組之次數，然後一一相加，再以項數除之。例如本表匯價之中位數為$4.8721，第一組內1項對於中位數之差

數爲$.0421;第二組內共6項,其每項對於中位數之差數爲 $.0371,餘類推。惟照此一一計算其眞實離中差,因有小數,手續頗爲繁瑣,爲求簡捷起見,可從假定中位數(或假定算術平均數)以計算離中差,再將各項離中差相加,而後再按校正數校正之。此校正數卽爲計算離中差時不根據眞實平均數,而根據假定平均數所發生之差誤也。表三十卽係說明此種簡捷計算法。

吾人如根據算術平均數計算平均差,亦可用此簡捷法。用簡捷法計算時,如以組距爲單位,則平均差之計算手續更可簡化,與前章內計算算術平均數之簡捷法相似。

在表三十內計算各項離中差時,係根據假定平均數 $4.870 而並不根據眞實中位數$4.8721計算者。各項對此假定平均數之差數,不問其符號之爲正爲負,一一相加,其總和爲 $5.020,然此總和究與根據中位數計算之各項離中差之和相差若干乎? 在以組中點作爲假定平均數之一組內共包含45項,每項之d'(卽對於假定平均數之離中差)皆爲0,但實際該組內每項對於中位數之離中差應爲$4.8721－$4.870,卽$.0021,故該組內每項由假定平均數所得之離中差須各補足 $.0021 之數而後可。在計算小於該組之各組內每項之離中差時,亦發生同樣之差誤;總計發生此項差誤者共196項,每項均應補足 .0021之數。在大於假定平均數所在組之各組內每項之離中差則發生異方向之差誤。例如該組之上一組共包含49項,每項對於假定平均數之離中差爲 $.005,而每項對於中位數之離中差應爲 $.0029,故應由各項離中差減去 $.0021;總計發生此項差誤者共188項。盈虧相抵,淨計尙須補足 $.0021 者8項,故各項離中差之總和$5.020尙須補足8×$.0021,卽$.0168之數,由此求得各項對於中位數之離中差之和爲$5.0368,其平均差爲$.01312。

此計算平均差之簡捷法可化爲下列公式:

$$M.D.=\frac{\Sigma(fd')+(N_s-N_l)c}{N}$$

公式中N_s＝對於假定平均數之離中差比較對於中位數（或算術平均數）之離中差爲小之各項項數。

N_l＝對於假定平均數之離中差比較對於中位數（或算術平均數）之離中差爲大之各項項數。

c＝假定平均數與中位數（或算術平均數）之差數。

在應用此項公式時，假定平均數與中位數（或算術平均數）須在同一組內(註)。

茲更列計算平均差之程序於下：

1. 求中位數之數值。

2. 以中位數所在組之中點作爲假定平均數，求每組各項對於假定平均數之離中差，更將此差數乘以各該組之次數，而後求乘積之總和，其乘積之符號爲正爲負，可不置論。

3. 點算對於假定平均數之離中差比較對於中位數之離中差爲大之各項項數，及對於假定平均數之離中差比較對於中位數之離中差爲小之各項項數。將此兩類項數之差乘中位數與假定平均數之差數，卽得各項對於假定平均數離中差之和與各項對於中位數離中差之和之差額。應用上項差額以校正各項對於假定平均數所得離中差之和。

4. 將此已校正之離中差之和除以總項數，卽得根據中位數計算之平均差。

（根據算術平均數計算平均差，手續與此相同）。

（註）在前例計算時，係假定中位數所在組內之各項數值密集於該組之中點者，但若假定該組各項在其組距內分配勻整，而將計算方法依此略予變更，則所得結果之確度當可稍高（關於此種計算法之說明，見 H.L.Rietz 氏所編"Handbook of Mathematical Statistics"第29—31頁）。

標準差

平均差之計算，將正負符號概置不問，在學理上不免牽強。標準差(standard deviation)之計算，則無此弊病，且在數學意義上亦較準確。統計學上通常以σ作為標準差之符號，即希臘字sigma也。

標準差之計算方法，係先將根據算術平均數求得之各項離中差一一自乘後相加，除以項數，再開平方根即得。故標準差者實即各項離中差自乘後之算術平均數之平方根也。標準差亦稱均方根差(root-mean-square deviation)，為由計算方法而得之名詞。標準差之數值，以從算術平均數計算者為最小，故求各項離中差恆以算術平均數作為中心值。今以簡單之例說明標準差之計算方法於下：

表三十一

標準差之計算

m	f	d	d^2	
3	1	-6	36	$M=9$
6	1	-3	9	$\sigma=\sqrt{\frac{90}{5}}$
9	1	0	0	
12	1	$+3$	9	$=\sqrt{18}$
15	1	$+6$	36	$\sigma=4.24$
	5		90	

統計事項如未組成次數表，則計算標準差之公式為

$$\sigma=\sqrt{\frac{\Sigma d^2}{N}}$$

統計事項如已組成次數表，則標準差之計算手續稍覺繁瑣，為求簡捷起見，各項離中差之計算恆以假定平均數為中心值。由已組成次數分配表之統計事項求標準差之普通公式為

$$\sigma=\sqrt{\frac{\Sigma fd^2}{N}}$$

上式中f表示每組次數，d表示每組中點依算術平均數計算之離中差，N表示總項數。由上式得

$$\sigma^2=\frac{\Sigma fd^2}{N}$$

如以d'表示依假定平均數計算之每項離中差，s表示依假定平均數求得之標準差，則得

$$s^2=\frac{\Sigma f(d')^2}{N}$$

以算術平均數為中心值所求得之標準差(σ)，恆較以尺度上其他任何一點為中心值所求得之標準差為小，故s^2恆大於σ^2。今設以c表示眞實平均數與假定平均數之差，可得下列公式(註)：

$$\sigma^2=s^2-c^2$$

由是知標準差之數值甚易求得，祇須計算s^2及c^2即可。表三十二係說明標準差之計算手續。表內所列係384個月倫敦對巴黎之匯價。

(註)因 $\sigma^2=\frac{\Sigma d^2}{N}$，　但 $\Sigma d=0$，

$s^2=\frac{\Sigma(d')^2}{N}$，　故 $\Sigma(d')^2=\Sigma d^2+Nc^2$，

$d'=d+c$，　$\frac{\Sigma(d')^2}{N}=\frac{\Sigma d^2}{N}+c^2$，

$(d')^2=d^2+2cd+c^2$，　$s^2=\sigma^2+c^2$，

$\Sigma(d')^2=\Sigma d^2+2c\Sigma d+Nc^2$，　$\sigma^2=s^2-c^2$。

參看 Yule 氏著 "Introduction to the Theory of Statistics" 第134頁。

表三十二

1882—1913年倫敦對巴黎匯價標準差之計算

（表內數字根據 The Economist 所載每月月初之匯價）

(1) 組距 (法郎)	(2) 組中點 (法郎) m	(3) 次數 f	(4) 對於假定平均數之離中差 d'	(5) fd'	(6) $f(d')^2$	(7) $(d'+1)^2$	(8) $f(d'+1)^2$
25.07—25.089	25.08	1	—8	— 8	64	49	49
25.09—25.109	25.10	4	—7	— 28	196	36	144
25.11—25.129	25.12	14	—6	— 84	504	25	350
25.13—25.149	25.14	20	—5	—100	500	16	320
25.15—25.169	25·16	45	—4	—180	720	9	405
25.17—25.189	25.18	60	—3	—180	540	4	240
25.19—25.209	25.20	40	—2	— 80	160	1	40
25.21—25.229	25.22	43	—1	— 43	43	0	
25.23—25.249	25.24	42	0			1	42
25.25—25.269	25.26	32	1	32	32	4	128
25.27—25.289	25.28	26	2	52	104	9	234
25.29—25.309	25.30	21	3	63	189	16	336
25.31—25.329	25.32	20	4	80	320	25	500
25.33—25.349	25.34	4	5	20	100	36	144
25.35—25.369	25.36	6	6	36	216	49	294
25.37—25.389	25.38	2	7	14	96	64	128
25.39—25.409	25.40	2	8	16	128	81	162
25.41—25.429	25.42	2	9	18	162	100	200
		384		—372	4,076		3,716

$N=384$

組距＝.02法郎

c（以組距爲單位）$=\dfrac{-372}{384}=-.969$

c^2（以組距爲單位）$=.9390$

s^2（以組距爲單位）$=\dfrac{\Sigma f(d')^2}{N}=\dfrac{4{,}076}{384}=10.6146$

σ^2(以組距爲單位)$=s^2-c^2=10.6146-.9390=9.6756$

σ (以組距爲單位)$=3.11$

σ (原來單位)　$=3.11\times.02=.0622$

於此所當注意者，上表中逐步計算，皆以組距爲單位，故最後須將求得之結果化爲原來單位。眞實平均數與假定平均數之差(c)之計算，係將各項離中差之和除以總項數求得。如欲求算術平均數之值，可先將 c 化爲原來單位，再將此數值依代數方式與假定平均數相加，惟此項手續並不需要，因標準差之計算，可不必求眞實平均數之值也。計算標準差時，在計算方面有無錯誤，可採用表三十二第(7)(8)兩欄所示之校勘法(查里歐校勘法 the Charlier check)(註)。在前例中求標準差時，各項離中差均係按照假定平均數計算，今若改按假定平均數所在組之下一組之中點計算，則所得各項離中差之值等於 $d'+1$，其平方後之數值當如第(7)欄所示。再將第(7)欄數字乘以各組次數得第(8)欄數字，其和爲3,716。此和數與計算標準差時求得之各項數字有固定之關係，因

$$\Sigma f(d'+1)^2=\Sigma f[(d')^2+2d'+1]$$
$$=\Sigma f(d')^2+2\Sigma fd'+\Sigma f$$

即　$$\Sigma f(d'+1)^2=\Sigma f(d')^2+2\Sigma fd'+N$$

上式如以表三十二內所求得之各數值代入，可校核標準差之計算有無錯誤。其式如下：

$$3,716=4,076+2(-372)+384$$
$$=3,716$$

統計資料如已分組成爲次數分配，則標準差之計算手續，可簡述如下：

1.選定靠近次數分配中心之一組之中點，作爲假定平均數

(註)參看C. V. L. Charlier氏編"Vorlesungen Uber Die Grundzüge Der Mathematischen Statistik" 第19頁。

2.就上下各組，以組距爲單位，分別計算各組中點對於假定平均數之離中差；次將各組離中差乘以各該組之次數。

3.將各項離中差與次數相乘之積一一相加而除以 N，即得以組距爲單位之 c，再求 c^2。

4.將各組離中差自乘，再乘以各該組之次數。

5.以N除前項乘積之和，即得以組距爲單位之 s^2。

6.依照公式$\sigma^2=s^2-c^2$求σ^2，再將所得之值開平方，即爲以組距爲單位之σ。

7.依此求得σ之數值，以組距乘之，即得以原來單位表示之σ。

標準差之各種特質及其與他種離散度量數之關係，當於下節論述之(註)。

四分位差

在討論平均數之一章內，嘗述及四分位數與十分位數之確定方法。四分位數者，係當次數分配內各數量依次排列於價值尺度上時，將全部項數截分爲四等分之分點也；十分位數者，爲將全部項數截分爲十等分之分點也。故四分位數與十分位數一經確定，次數分配之離散度已不啻明白表示，惟此種量數雖亦可表示次數分配之離散狀態，但欲爲概括之敍述，或供比較之用，則較遜於平均差與標準差。蓋平均差或標準差祇有一個數值，其意義較爲扼要，而四分位數與十分位數恆包括數個彼此相關之數值，其意義頗難領悟也。惟用一個單純數值表示離散度之方法，亦可自四分位數蛻化而得，即所謂四分位差(quartile deviation)是也。四分位差計算手續之簡便與意義之易於明瞭，均爲他法所不可及。

在尺度上第一四分位數與第三四分位數之間，包含全體項數之半。

(註)標準差有時須加校正(舍巴德氏校正法 Sheppard's correction)，其校正法詳見第十五章。

凡次數分配愈集中，離散度愈小，則此兩者之距離愈短，故欲根據此兩個四分位數之關係，以測度次數分配之離散度，亦可得相當之準確性。四分位差實即在尺度上第一四分位數與第三四分位數間之距離之半(semi-interquartile range)。令$Q.D.$表示四分位差，Q_1表示第一四分位數，Q_3表示第三四分位數，則以公式表示之四分位差如下：

$$Q.D.=\frac{Q_3-Q_1}{2}$$

如將尺度上第一四分位數與第三四分位數間之中點以K表示之，則次數分配全體項數之半當在$K\pm Q.D.$之距離內。茲仍以前述倫敦對巴黎匯價之數字為例，說明於下：

$$Q_3=25.262$$

$$Q_1=25.174$$

$$Q.D.=\frac{25.262-25.174}{2}$$

$$=.044$$

$$K=25.174+.044$$

$$=25.218$$

由是知該次數分配全體項數之半當在$25.218\pm.044$之距離內。此點在表述次數分配之意義上，與匯價平均數值（算術平均數、中位數或衆數）同等重要。次數分配如完全對稱，則K之數值與中位數之數值合而為一（即中位數適在尺度上第一四分位數與第三四分位數間之中點）。今匯價之分配不甚對稱，故中位數之值為25.214，而K之數值則為25.218，兩者乃微有出入。

機誤

吾人如將關於天文或其他形體的測量之結果予以考究，則可知對

於同一事物，各人測量所得之數值互異，惟此種不同數值之分配狀態常循一定法則，如繪為曲綫當與差誤常態曲綫相似。於此吾人遂亟需有一種測定差誤離散度之量數，作為測量結果可靠程度之指標。測量結果無論為各人所測量者，或為一人測量多次者，如數值上差異甚大，則此項測量結果即不能認為可靠；反之，如差異不鉅，則測量結果之可靠程度亦必較高。通常用以測定此離散度之量數，謂之機誤(probable error)。機誤為一數值，即當各個觀察值之差誤超過此數值者適占全體項數之半也。各個觀察所得之數值，以接近其算術平均數之機會為最多，故機誤之測度恆以算術平均數為原點。按機誤一名詞之由來，係因每次觀察所得數值與其算術平均數之程度，超過機誤之機率適為$\frac{1}{2}$也，故當觀察所得數值已組成次數分配時，在其算術平均數左右兩邊各取等於機誤之距離，在此距離內應包含全體項數之半。

此種離散度之量數原為測定觀察差誤之用，但亦有供作與機誤原有意義不同之其他用途者，此時機誤可視為機差(probable deviation)，以免誤解。此機差當為算術平均數兩邊所應取之距離，使各項離中差之大於此距離者，適為全體項數之半也。

機誤用作測定離散度之量數，必須在次數分配合乎差誤常態法則時，始有明確之意義，如次數分配呈現偏斜狀態即不可用。在常態次數分配中，四分位差之數值與機誤相等，惟在表述偏斜之次數分配時，則前者之意義反較後者更為直接，此時以採用四分位差為宜。關於用機誤測定統計結果可靠程度之方法，容後申論之。

在常態次數分配中，機誤與標準差兩者之數值恆有一定之關係，故前者之數值可從後者求得。以公式表示之為$P.E.=0.6745\sigma$。

各種離散度量數之相互關係

茲為便於瞭解前述各種離散度量數之意義起見，再將其相互之關

係作簡略之說明並加以比較於下：

1.全距者，乃尺度上一定距離，全體項數盡在此距離之內者也。

2.四分位差者，爲尺度上另一距離。就第一四分位數與第三四分位數間之中點，在其左右兩邊各取等於四分位差之距離，則在此距離間當包含全體項數之半。

3.在完全對稱或不甚偏斜之次數分配中，依照算術平均數計算之平均差，約等於標準差$\frac{4}{5}$。以算術平均數爲中心，取平均差$7\frac{1}{2}$倍之距離，約可包含全體項數99%。

4.在完全對稱或不甚偏斜之次數分配中，若從算術平均數左右兩邊各取相等於標準差之距離，則在此距離內約包含全體項數三分之二（在常態次數分配中，則此距離內當包含全體項數 68.26%）。若從算術平均數左右兩邊各取兩倍於標準差之距離，則此距離內約包含全體項數 95%(在常態次數分配中，當包含全體項數95.46%)。若從算術平均數左右兩邊各取三倍於標準差之距離，則此距離內約包含全體項數 99%（在常態次數分配中，當包含全體項數 99.73%）。故吾人求標準差時，可以算術平均數爲中心，試取六倍於標準差之距離，以觀此距離內是否約包含全體項數99%，則對於標準差之計算是否有誤，即可知其大略矣。

試參看圖四十三，則對於常態次數分配之標準差之意義，當更明瞭。

5.在常態次數分配中，機誤等於0.6745 σ。以算術平均數爲中心，在兩倍於機誤之距離內應包含全體項數50%。以算術平均數爲中心，在八倍於機誤之距離內，應包含全體項數99%。

各種離散度量數之特質

全距

1.全距之計算,手續簡單,而其意義亦易瞭解。如吾人僅欲測定離散度之大概,可採用之。

2.全距之數值係依兩極端數值之大小而定,故極不穩定,常因一項數量之去留而受重大影響。

3.全距對於兩極端數值間之次數分配狀態,全無提示。

四分位差

1.四分位差之計算,手續簡單,意義亦易於瞭解,為測定離散度之一種粗略量數。就準確程度而言,四分位差較勝於全距。

2.四分位差並不依照任何平均數計算各項數量之離中差。

3.四分位差之數值不受第一四分位數與第三四分位數間各項數量分配狀態之影響,亦不受該距離以外各項數量分配狀態之影響。如兩個次數分配絕不相似,而四分位數偶然相同,則其四分位差亦必相等。四分位差既不受各項數量離中差之影響,故不能視為測定離散度之精確量數。

4.四分位差不能用代數方法處理。

平均差

1.平均差受每項數量之影響,且因係從各項單獨數量對其中位數(或算術平均數)之離中差平均而得,故其意義頗為確切。

2.平均差所受極端離中差之影響,比標準差為小。

3.依數學理論而言,平均差不及標準差之合理與合用。

標準差

1.標準差受每項數量之影響。

2.標準差之計算,將各項數量之離中差一一自乘,以消除其正負

號，而後相加，在代數學上絕無抵觸。

3.標準差在數學上具有確切之意義，用代數方法處理絕對適當。

4.標準差所受取樣差誤之影響，大致較測量離散度之其他量數爲小。

5.分析差誤常態曲綫，常以標準差爲單位，頗足增廣標準差之用途。

機誤

1.機誤之用於常態次數分配，具有確切之意義，而在偏斜之次數分配則否，故不可用以表述後一種次數分配。

2.在適用機誤之次數分配，機誤極有用途，其最要者卽可用作測量數量可靠程度之指標。

3.在常態次數分配中，機誤與標準差間有一定之關係，故機誤之數值不難確定。

前述測量離散度之各種量數，各有其特殊用途，惟大體言之，以標準差爲最優，而欲使離散度之測量結果十分準確，尤不能不採用標準差。至於機誤實際上祇係標準差數值之分數，惟其用途範圍較爲狹隘。

相對離散度之測度

在前節內已將絕對離散度之測度方法加以討論。用上述各種方法所求得離散度之數值，皆係用原來單位表示者。例如倫敦對巴黎匯價之標準差係以法郎爲單位，生鐵產量之標準差係以噸爲單位等是。如吾人之目的僅爲表述一單純之次數分配，則測量離散度自以採用原來單位較爲相宜，但如欲比較兩個不同之次數分配，卽不免發生困難。在兩個次數分配之原來單位彼此互異時，困難之發生固屬顯見，卽原來單位彼此相同，此種困難仍所難免。例如測量犬馬體重之離散度，雖同以磅爲單位，但當馬體重量之標準差大於犬體重量之標準差時，吾人不能遽謂前者之離散度必大於後者。測量離散度所得之絕對數值，必須與計算各

項離中差所根據之平均數比較，始有意義。如與平均數脫離，意義卽不明瞭，故在比較時，離散度之數值須化成比率之形式。最簡便之手續係將該數值除以計算各項離中差所根據之平均數，而得該值與平均數之百分率。此值遂成抽象數字，而爲測量次數分配之相對離散度之量數，乃可用以與其他次數分配之相對離散度之量數彼此比較矣。

離散度係數

在測定相對離散度之各種量數中，應用最爲普遍者，當推披爾遜氏(Pearson)所創之離散度係數(coefficient of variation)此離散度係數係以V表示，實卽以算術平均數除標準差所得之百分率也。其公式如下：

$$V=\frac{\sigma}{M}\times 100$$

以前述倫敦對巴黎匯價一例中求得之結果代入上式，得

$$V=\frac{.0622}{25.2206}\times 100$$

$$=.25\%$$

同時期倫敦對紐約匯價之離散度可從表三十之資料求得。其離散度係數爲.33%，由是知倫敦對紐約匯價之變動，顯較同時期倫敦對巴黎匯價爲劇烈。

至於其他測量離散度之絕對量數，亦可依照計算此係數之方法，以其離中差所根據之平均數除之，以求離散度之指數，惟披爾遜氏之離散度係數，應用最爲普遍耳。

偏斜度之量數

關於表述次數分配之集中趨勢與測定次數分配中心值兩邊之離散度之方法，已如上述。吾人尙需要一種量數以測定次數分配之偏斜度

(skewness or asymmetry)，蓋對於次數分配中心值兩邊各數量之分配之爲對稱，抑欠勻整而呈偏態，亦理應研究之。此項量數旣已求得，則吾人可有三種簡單之量數——平均數、離散度量數及偏斜度量數——以簡述次數分配之特質。測定偏斜度之量數可有數種，茲分述於下。

次數分配如係完全對稱，則其算術平均數、中位數及衆數均在一點。當次數分配漸趨偏斜，此三者之數值亦漸分離，而以算術平均數與衆數之差異爲最大，此項差數卽可用作測度偏斜程度之量數。此測定偏斜度之量數亦可仿照求相對離散度之方法，化爲抽象形式之係數，俾可與另一次數分配之偏斜度係數互爲比較。披爾遜氏主張以次數分配之標準差除算術平均數與衆數之絕對差數作爲偏斜度量數。其公式如下：

$$sk(skewness)=\frac{M-Mo}{\sigma}$$

在對稱之次數分配中，算術平均數與衆數合一，故此時偏斜度量數之數值等於0。偏斜度量數之數值或爲正數，或爲負數，胥視尺度上此兩種平均數位置之關係而定。

如次數分配不甚偏斜，則用下式測定偏斜度尤爲便利。

$$sk=\frac{3(M-Md)}{\sigma}$$

用此式求得之結果，與用前述第一式所求得者相近似，此因次數分配不甚偏斜時，中位數之位置介乎算術平均數與衆數之間，而中位數與算術平均數之距離，約等於算術平均數與衆數距離之三分之一也。

衆數之值不易確定，故披爾遜氏公式常不適用，而須採用計算較易之其他方法。鮑萊氏 (Bowley) 主張用第一四分位數、第三四分位數與中位數三者關係以爲計算之方法。蓋如次數分配對稱，此兩個四分位數與中位數之距離應各相等；如次數分配不對稱，則兩個距離不相等。今以q_2表示第三四分位數與中位數之差額，q_1表示中位數與第一四分位數之差額，吾人卽可用下式測定偏斜度。

$$sk=\frac{q_2-q_1}{q_2+q_1}$$

依上式求得之數值，當在0與±1之間，因次數分配如完全對稱，則$q_2=q_1$，即偏斜度量數之數值爲0；如次數分配極不對稱，以致中位數與第一四分位數或與第三四分位數合而爲一，則q_2或q_1等於0，偏斜度量數之數值當爲 +1或−1。鮑萊氏嘗解釋謂計算結果等於.1時，即次數分配略帶偏斜之表示，等於.3時，即次數分配呈現顯著偏斜之表示。

用此法求得偏斜度量數之數值不能與用披爾遜氏公式所求得者相比較。

峯度

關於表述次數曲綫第四種特質之方法，前嘗提及，此乃以差誤常態曲綫爲標準，測定某一曲綫峻峭程度之方法。在統計學上此種次數分配之特質謂之峯度(kurtosis)。其測度方法當在第十五章內論述之。

參考書

Bowley, A. L. "Elements of Statistics" (110—117).

Chaddock, R. E. "Principles and Methods of Statistics" (Chap. IX).

Davies, G. R. "Introduction to Economic Statistics" (29—46).

Jones, D. C. "A First Course in Statistics" (42—51).

Kelley, Truman L. "Statistical Method" (70—82).

King, W. I. Elements of Statistical Method" (141—158).

Pearl, Raymond. "Medical Biometry and Statistics" (272—278).

Rietz, H. L. (editor). "Handbook of Mathematical Statistics" (27—33)

Rugg, H. O. "Statistical Methods Applied to Education" (149—179).

Secrist, Horace. "An Introduction to Statistical Methods" (377—424).

West, Carl J. "Introduction to Mathematical Statistics" (45—58).

Yule, G. U. "Introduction to the Theory of Statistics" (133—156).

第六章　物價指數

指數之性質

指數 (index number) 之一名詞常用以指統計分析上數種相類之方法而言。指數之用於物價變動之研究者爲最廣，惟其他各種用途亦應加討論，使讀者得明瞭其重要之性質。時間數列所化成之百分比，亦常被稱爲指數，爲指數中形式最簡單之一種。下表所列美國棉花消費量之指數，卽爲其一例。

表三十三

1913—1923年美國國產棉花之消費量

(1913年消費量＝100)

年份	棉花消費量 (單位：袋)	棉花消費量百分比
1913	5,583,468	100.0
1914	5,448,760	97.6
1915	6,008,984	107.5
1916	6,620,415	118.6
1917	6,815,811	122.0
1918	6,176,547	110.5
1919	5,919,520	106.0
1920	5,843,200	104.5
1921	5,406,721	96.8
1922	6,087,520	109.0
1923	6,513,696	116.7

物品價格亦可採用同法以百分比表示，而以某日或某時期之價格作爲基數。

表三十四

1913年及1919—1923年美國米乃波立斯頭號北方春麥之平均市價

(1913年平均價＝100)

年份	每蒲式耳平均價	比價
1913	$0.8735	100
1919	2.5660	294
1920	2.5581	293
1921	1.4660	168
1922	1.3450	154
1923	1.1810	135

時間數列中各項，如以一固定基數爲 100，以百分比表示之，不但可使各期之數值便於比較，而對於該數列之變動趨勢，亦較原有形式更易明瞭。在比較數個數列之趨勢時，用百分比之形式亦較爲便利。

指數之名詞雖可用於此種百分比，但吾人應保留此名詞，以用於代表由數個數列綜合所得之數字較爲適當。所綜合之各個數列，或關於物價，或關於物品之生產量，消費量，工資，貿易額，以及隨時間變化之任何事物。編製此項特種指數常發生極複雜之問題，惟其最後目的爲欲求一單純與簡要之數字，以表示各個數列所受各種勢力之淨結果，則各指數皆同。

試根據美國煤與石油之產量，編製一種簡單指數，以表示產量之增減變化。在編製此種指數時，須將烟煤、白煤與石油之產量綜合計算。表三十五係一九一〇年至一九二三年烟煤、白煤、石油三個數列之產量及其百分比。

吾人設欲根據此三個數列編製一種粗略之燃料生產指數，則因物品數量單位之不同，不能逕將原有之產量數字相加，但若採用百分比，則可免除此種困難。將每年之三個百分比相加，而求其簡單平均數，卽得所求之指數。依此求得之各年指數見表三十六：

表三十五

1910—1923年美國烟煤白煤及石油產量

(1910年產量＝100)

年份	烟煤產量 百萬公噸	百分比	白煤產量 百萬公噸	百分比	石油產量 百萬蒲式耳	百分比
1910	417.1	100	84.49	100	209.6	110
1911	405.9	97	90.46	107	220.4	105
1912	450.1	108	84.36	100	222.9	106
1913	478.4	115	91.52	108	248.4	118
1914	422.7	101	90.82	107	265.8	127
1915	442.6	106	89.00	105	281.1	134
1916	502.5	120	87.58	103	300.8	143
1917	551.8	132	99.61	118	335.3	160
1918	579.4	139	98.83	117	355.9	170
1919	465.9	112	88.09	104	378.4	181
1920	568.7	136	89.60	106	442.9	211
1921	415.9	100	90.44	107	472.2	225
1922	404.5	97	52.90	63	557.5	266
1923	545.3	131	95.20	113	725.7	346

表三十六

1910—1923年美國煤與石油之生產指數

(1910年產量＝100)

年份	指數	年份	指數
1910	100	1917	137
1911	103	1918	142
1912	105	1919	132
1913	114	1920	151
1914	112	1921	144
1915	115	1922	142
1916	122	1923	196

計算前項指數係將每年三個百分比相加，以三除之而得，是各個百分比所用之權數相等。此種根據權數相等之百分比所求得之指數，謂之未加權指數(unweighted index)，但此名詞易滋誤解，蓋實際權數固已採用，惟在本例中各個百分比之權數相等而已。此種指數因係根據相等之權數編製，故其指數不能準確反映三個數列之綜合趨勢，蓋若各個數列予以相等之權數，即謂各個數列之重要程度相等，此點實與事實相

左。下列數字係一九二一年烟煤、白煤、石油批發市場之成交價值，以示三種數列之比重程度(註)。

礦產	1921年批發市場之成交價值
烟煤	$1,948,000,000
白煤	731,000,000
石油	712,000,000

上列數字表示三種物品在批發市場之重要程度約為 3,1,1 之比。吾人可將此項比重資料作爲權數，加於各該數列，以計算每年指數。茲將根據前述三個數列計算一九一〇年及一九二三年指數之手續列舉於下，以說明其計算方法。

表三十七

煤與石油生產量加權指數之計算

礦產	1910年生產量之百分比	權數	權數×百分比	1923年生產量之百分比	權數	權數×百分比
烟煤	100	3	300	131	3	393
白煤	100	1	100	113	1	113
石油	100	1	100	346	1	346
		5	500		5	852

1910年燃料生產指數＝500÷5＝100

1923年燃料生產指數＝852÷5＝170

按照上法求得十四年間之每年指數如下：

表三十八

1910—1923年美國煤與石油生產量之加權指數

年份	指數	年份	指數
1910	100	1917	135
1911	101	1918	141
1912	106	1919	124
1913	114	1920	145
1914	107	1921	126
1915	111	1922	124
1916	121	1923	170

(註) 此項數字係美國勞工統計局所編製。

表三十六及表三十八所列兩種指數，試加比較，則有顯著之差異。後一種指數因其所用權數較爲合理，故所得結果自較準確，尤足以代表三個數列所受各種勢力之綜合結果。

此外尙有一種指數，其算法係將每期各項數值一一相加，得其綜合數值，作爲全體數值之代表，而不求其平均數者。惟欲計算此種指數，各數值之單位必先一致而後可。此種指數常用以測度物價水準之變動，係將某時期若干物品之綜合價格，與另一時期各該物品之綜合價格比較而得。下列數字卽爲其一例。

表三十九

1913—1923年美國勃蘭特斯脫(Bradstreet)雜誌社之批發物價指數

年份	指數	年份	指數
1913	$9.21	1919	$18.66
1914	8.90	1920	18.81
1915	9.85	1921	11.37
1916	11.82	1922	12.12
1917	15.64	1923	13.40
1918	18.71		

上列每年綜合價格係各該年九十六種物品之平均批發市價之總和，惟在價格相加之先，各項價格均曾折合爲每磅價格，俾各項價格間有一比較之標準。此種指數亦可化爲百分比，以任何一年作爲基期，將其他各年之綜合價格折算爲基年(base year)綜合價格之百分率。

由前例可見指數之形式繁簡不一，自簡單之百分比以至各項百分比之平均數，絕對數或百分比之綜合數，概可稱爲指數。顧指數之形式雖異，但其目的爲測量某時期內之數值變動，或表示時間數列內各項數值之變動則一也。除前述各例外，指數尙有更廣泛之意義。例如吾人欲測驗售貨員之能力，可就決定此能力之各因素一一考查，評定其績分，再求其平均數，卽得售貨員能力之指數。又如欲測驗企業組織下各部分之效率，亦可編製一種指數。惟是任何指數之編製，應將各個不同之因素化爲可供比較之數值，而後根據此各個數值求一單純數字，以代表全

體之數值。故指數可用以比較一時期內某項事物之變動，或同時期內各項事物之變化。可知指數之各種形式(除第一例含義較狹僅指簡單百分比之數字以外)，實即統計上之一種平均數，故指數之計算及其用途，遂不免受平均數各項原則及缺點之拘束。

本章所欲論述者，祇爲應用於時間數列之指數編製方法。指數之應用於各種時間數列時所涉及之原則及慣例並不一致，故須將各種數列分別討論。請先討論批發物價指數。

物價之變動

吾人如將物價之變動詳加研討，則錯綜紛紜，漫無準則，幾無顯著之趨勢可尋。試任取物價數項列表於下，以示物價變動凌亂紛歧之一斑。

表四十

批發物價(註)

物品	單位	批發價格 1923年10月	批發價格 1923年11月
磚，普通建築用，各產地之平均價	1000	$14.752	$14.746
生鐵，原料	噸(毛重)	23.500	20.875
水泥，Portland，各廠平均價	桶	1.893	1.842
胡蔴子油，未製者，紐約	加侖	.943	.910
鋼條，匹茲堡，培塞麥	噸(毛重)	40.000	40.000
錫版，匹茲堡	百磅	5.500	5.500
電化紫銅，提鍊廠出品	磅	.126	.128
鉛，紐約	磅	.069	.069
鋅，紐約	磅	.067	.067
白煤，紐約，徹斯納特	噸(毛重)	11.471	11.478
烟煤，支加哥	噸(淨重)	4.600	4.525
生石油，油井售價，本薛文尼	桶	2.500	2.388
汽發油，馬達用，紐約	加侖	.185	.170
棉花，米特令，新奧雷安斯	磅	.292	.339
小麥，二號赤色冬麥，支加哥	蒲式耳	1.097	1.061
糖，顆粒形，紐約	磅	.090	.087

(註)表內物價係美國勞工統計局所編。

表內十六種物品中，其一九二三年十一月與十月價格之比較，無變動者四種，漲者三種，跌者九種。其中某數種物品價格漲跌無多，而某數種物品之價格變動則甚劇烈。此一小部分物價之變動，已足象徵市場上全部物品價格變動之情況。各種物品價格之變動趨勢不一，有上漲者，有下跌者，亦有絕無變動者。在成千累萬種物品之中，每種物品各因在國內市場或國外市場所受影響不同，其價格變動因而各異，惟各種物品之價格，互有關聯，一物價格之漲跌，足以影響他物之價格，而同時亦受他物價格變動之影響。凡此皆係指每種物品本身所受之勢力而言，此外尙有範圍較廣之勢力，足以影響整個物價制度下全體物品之價格者。經濟統計學者之職責，即在從各時期紛亂之物價變動中求其規律，從許多繁瑣之變動中窺測經濟變動之大勢。

吾人所欲研究之物價變動，其所受勢力種類繁多，惟就大體言之，可歸納爲兩大類。其一爲關於各個物品本身之生產消費狀況者。各個物品生產消費狀況之變更，常可直接影響於其價格。例如新市場之發現，產地之擴展，生產技術之改進，風尙之變更，需求之轉移，以及季節發生之變化等，皆足使物價隨時變動。平時此種勢力最爲明顯，商人與消費者均能見及。此種勢力亦能影響物價之水準，但不能使整個物價制度呈一致上漲或下跌之現象。

至於可使物價普遍漲落之勢力，則屬於另一類，其範圍較爲廣泛。例如一般生產技術之進步，足以增進勞工之生產能力，其結果直接增加物品之供給數量，間接影響於其價格。又如貨幣制度之變更，黃金供給數量之增減，皆足以影響貨幣之流通數量，而可使物價直接受其影響。依此類推，則凡銀行與信用制度之改革以及商業習慣之變更，足以影響信用票據之使用範圍及貨幣與信用之流通速率者，對於物價亦有同樣之作用。凡此種種勢力，皆可影響一般物價，惟其所及之範圍，不若直接影響個別物品之各種勢力之確切耳。

普通批發物價指數之目的

前節所述各種勢力恆同時存在，而發生聯合作用，物價種種變動即爲此聯合作用之淨結果，故吾人殊無法將各種勢力一一隔絕，而分別估量其輕重。研究此種種變動，可從各種不同之觀點着手。吾人可研究物價制度內各部分所發生之變動，以決定其變動之性質及程度。此種研究對於物價變動之狀況及各個物價之相互關係，可獲明確之認識，惟吾人當前之問題則爲決定物價所受一切勢力之淨結果。各個物價之變動是否互相抵銷，使上漲之物價與下跌之物價適相平衡，以致物價之水準不生變動乎？抑在一時期內各個物價之漲勢與跌勢緩急不同，致使物價之水準爲之增高或降落乎？此種趨勢是否存在？如其存在，究爲何種趨勢？又將如何測度之乎？前所論述之各種統計方法可用以解答本問題乎？

此項研究之第一步，應先解答上述最後之一問題。前嘗論述簡縮數字資料之方法，但此種方法僅能在適合某種情形時始可應用。就平均數而言，統計資料必須性質一致，而呈現顯著之集中趨勢時，平均數始有意義。平均數之選擇又必須視統計資料之分配狀態而定。如對於統計原始資料之分配，事前絕未研究，則所選取之平均數或他種統計量數必欠恰當。吾人欲解答此問題，應先決定物價資料之性質，並研究該項資料整理後次數分配之形態。

設吾人研究物價變動之目的，爲普通決定兩個時期內批發物價水準之變動。決定物價水準之變動，實即測度兩個時期內貨幣價值之變動也。研究之資料爲兩時期內若干項物品之價格，每品之兩項價格可測度其在此兩時期內因受外界種種勢力而發生之變動。如將許多物價彙集一處，則此項集合資料可代表各種勢力之相互作用。此各種勢力中，有特殊者，有普通者；特殊勢力僅影響一二物品之價格，普通之勢力則可影響多種物品或全體物品之價格。吾人所欲決定者爲此各種勢力所引

起物價變動之淨結果，故必須用一種量數以測度足使各個物價漲落之一切勢力之混合結果。此量數卽爲批發物價指數。

此時吾人所用資料之單位，爲每一個物價之變動。前所論述之統計方法，能否應用於此項資料之整理與分析，須視各項單位之動態而定。下述一例係舉示此項資料分組後之次數分配。

表四十一

1914年物品346種之比價分配(註)

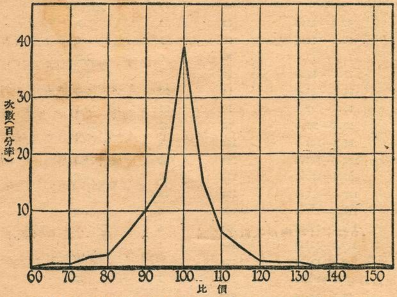
(1913年平均價=100)

比價	中點 m	項數 f	占總項數之百分率
62.5— 67.4	65	1	.3
67.5— 72.4	70	1	.3
72.5— 77.4	75	5	1.5
77.5— 82.4	80	7	2.0
82.5— 87.4	85	20	5.6
87.5— 92.4	90	35	10.0
92.5— 97.4	95	51	14.5
97.5—102.4	100	134	39.0
102.5—107.4	105	50	14.5
107.5—112.4	110	21	6.0
112.5—117.4	115	12	3.5
117.5—122.4	120	3	1.0
122.5—127.4	125	2	.6
127.5—132.4	130	2	.6
132.5—137.4	135		
137.5—142.4	140	1	.3
142.5—147.4	145		
147.5—152.4	150	1	.3
		346	100.0

(註)表內所包含之346種物品，係美國勞工統計局編製批發物價指數所用者。其原來價格及各品之百分比見該局叢書第269種 "Wholesale Prices, 1890—1919".

物價比率之次數分配

每項物價之變動可用一比率表示，卽某一時期一物價格與另一時期同物價格之比率也。此比率可依本章前此所舉指數之各例，化爲百分比形式，藉便比較。今以表四十內各對價格中之一對爲例以說明之。如生鐵一九二三年十一月份價格與一九二三年十月份價格之比爲 \$20.875:\$23.500，化爲百分比爲 88.8:100。上列次數表(表四十一)係根據一九一四年三百四十六種物品之價格，各以一九一三年各該物品之價格作爲基數計算而得之百分比編製而成。此三百四十六項百分比數字之分配如表四十一。



圖四十七· 次數多邊圖：1914年物品346種之比價分配(1913年平均價=100)

表示此項分配之次數多邊圖如圖四十七。圖中次數係按照各組項數占總項數之百分率繪製，俾可與其他性質相似之次數分配比較。此次數分配與前此所示標準形式頗有相似之處：其一，爲顯著之集中趨勢。本例中物價變動頗有穩定之趨勢，各項物價比較其基年價格漲落在 2.5%

以內者，達總項數之 39%。其二，在此分配中大於衆數各項之全距雖比較小於衆數各項之全距爲大，而此點又頗關緊要，但中心值兩旁之分配，大體呈對稱狀態。吾人暫不研究此次數分配應採用何種平均數以表示其集中趨勢，然平均數之可適用於此種次數分配，則屬顯見也。

上述之例係比較前一年物價與後一年物價之變動。在此兩時期間物價之水準幾無變動。密哲爾氏(W. C. Mitchell)嘗舉示更爲詳備之一例，係將一八九〇年至一九一三年間5578種物品價格之逐年變動，採用與上例相同之分組法，製爲次數分配表。表中大於衆數各項之全距所超過小於衆數各項之全距之情勢似更顯著，其原因由於在此二十三年中，各年物價上漲者居多所致。該次數分配如圖四十一所示。

在研究物價逐年變動時，物價之惰性最爲明顯。吾人應再研究較長時期內物價變動之狀況，藉以明瞭是否具有與逐年變動相同之分配狀態。試舉兩例研究之：一爲十年間之物價變動，其期末之物價水準與期首約略相等者，一爲五年間之物價變動，其逐年物價之增漲頗速者。表四十二係一九〇〇年物品 222 項之價格，以一八九〇年爲基期之變動百分率。在此時期內，一八九〇年至一八九六年間物價之水準下跌，而一八九六年至一九〇〇年間物價之水準上漲，跌後回漲，故一九〇〇年之物價水準較一八九〇年僅低1%。

以此項數字資料繪爲次數多邊圖如圖四十八。圖中次數係按照各組項數占總項數之百分率繪製(圖四十七與圖四十八之縱尺度大小不同，比較時宜注意)。

圖四十七與圖四十八所表示之兩種次數分配迥不相同。圖四十八中之全距顯較圖四十七中之全距爲大，此由於包含時期較長，比價分散之所致。再則圖四十八中之集中趨勢雖仍明顯，但密集於衆數組之次數之百分率遠較前一種爲低。兩種分配繪於算術尺度上，皆呈對稱狀態(吾人應注意兩種分配中用算術平均數所測度前後兩時期之物價水準

幾全相等)，惟前一種分配之集中趨勢遠較後一種爲明顯，而各個比價對於其算術平均數之離中差亦較小。此種分配頗與極準確之物體測量或由極準確之礮發出彈丸所得之分配相似。後一種曲綫則與不甚準確之物體測量或由陳舊不準確之礮發出彈丸所得之分配相似，其衆數值發現之次數不多，而各個比價之離中差亦較大。由是知計算物價變動包

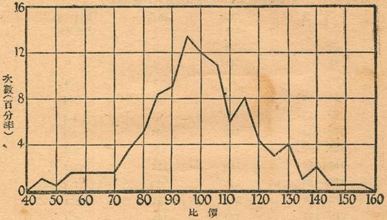
表四十二

1900年物品222種之比價分配

(1890年平均價＝100)

比價	中點 m	項數 f	占總項數之百分率
42.5— 47.4	45	2	1.0
47.5— 52.4	50	1	.5
52.5— 57.4	55	3	1.5
57.5— 62.4	60	3	1.5
62.5— 67.4	65	3	1.5
67.5— 72.4	70	3	1.5
72.5— 77.4	75	8	3.5
77.5— 82.4	80	12	5.0
82.5— 87.4	85	19	8.5
87.5— 92.4	90	20	9.0
92.5— 97.4	95	30	13.5
97.5—102.4	100	27	12.0
102.5—107.4	105	24	11.0
107.5—112.4	110	14	6.0
112.5—117.4	115	18	8.0
117.5—122.4	120	10	4.5
122.5—127.4	125	7	3.0
127.5—132.4	130	9	4.0
132.5—137.4	135	2	1.0
137.5—142.4	140	4	2.0
142.5—147.4	145	1	.5
147.5—152.4	150	1	.5
152.5—157.4	155	1	.2
		222	100.0

含時期愈長，則後一種分配所呈現之散漫趨勢愈顯著，其最高縱坐標之數值漸減小，分配之全距漸增大，而所得曲綫之形態愈趨平坦，曲綫之展佈亦愈廣泛，以致任何平均數之代表性爲之減弱。蓋若次數分配不集中，則依此求得之任何一種平均數僅爲一抽象數字而無切實意義也。



圖四十八．　次數多邊圖：1900年物品222種之比價分配(1890年平均價＝100)

吾人於此可得一實驗之結論，即物價變動可用統計方法測度，又物價所包含之時期如不過長，其變動可用平均數表示。至於吾人應以若干年認爲最長時期，在此時期內之物價變動方可測度，事實上殊不易決定。物價指數如欲求其準確而具有意義，則須根據短時期內之物價編製，尤以根據物價逐年變動之指數最爲準確。如物價指數之編製目的僅在表示物價之大體趨勢，則包含時期自須稍長，惟編製與引用此種指數者須明瞭其缺點之所在耳(註)。

以上所舉兩例皆係選擇前後兩時期之物價用算術平均指數所測度之水準大致相等者，此係特殊之例，故吾人尙須研究物價水準變動較烈之兩個時期內各項價格變動之分配。下表所列係一九一八年 1437 種物

(註)此項結論大半根據 W.C.Mitchell 氏研究之結果。參看美國勞工統計局叢書(物價叢書)第284種 "The Making and Using of Index Numbers".

品比價之分配，以一九一三年七月至一九一四年六月之平均物價作爲基數者(註)。

表四十三

1918年物品1437種之比價分配

(1913年七月至1914年六月之平均價＝100)

比價	中點 m	項數 f	占總項數之百分率
36	36	1	*
49	49	1	*
50—69	60	4	.3
70—89	80	17	1.2
90—109	100	61	4.3
110—129	120	64	4.5
130—149	140	130	9.0
150—169	160	212	14.7
170—189	180	219	15.2
190—209	200	164	11.4
210—229	220	135	9.4
230—249	240	104	7.2
250—269	260	76	5.3
270—289	280	54	3.8
290—309	300	42	3.0
310—329	320	[illegible]0	2.1
330—349	340	31	2.1
350—369	360	16	1.1
370—389	380	13	.9
390—409	400	7	.5
410—429	420	7	.5
430—449	440	8	.6
450—469	460	4	.3
470—489	480	4	.3
490—509	500	4	.3
510—529	520	5	.4

(註)此項物品比價之分配係美國戰時工業局物價組(Price Section of the War Industries Board)所編，轉載於美國勞工統計局叢書第284種卷一第70頁。

表四十三（續）

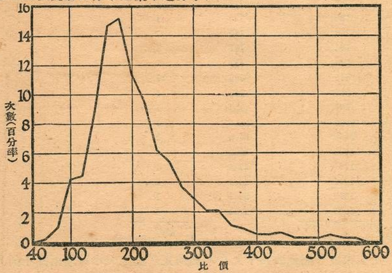
比價	中點 m	項數	占總項數之百分率
530—549	540	3	.3
550—569	560	4	.3
587	587	1	*
627	627	1	*
727	727	1	*
730	730	1	*
743	740	1	*
761	761	1	*
784	784	1	*
826	826	1	*
848	848	1	*
900	900	1	*
1165	1165	1	*
1356	1356	1	*
1585	1585	1	*
1764	1764	1	*
2049	2049	1	*
2863	2863	1	*
3009	3009	1	*
		1437	

*百分率在.1%以下.

上項次數分配見圖四十九（本圖所定尺度之大小與前兩圖不同）。

該次數分配細加研究，可獲與前述兩例相同之結論，卽該分配之集中趨勢甚爲明顯，故可用平均數以代表分配之全體。且因其衆數組爲包含中點180之一組，故知項數密集之原因，並不由於物價之惰性，而由於足以影響整個物價制度之各種勢力之存在。惟此分配與前述兩種次數分配猶有顯著之差別，卽分配之偏斜傾向，在前述第一種分配中已略呈端倪，而在此分配中則更較明顯。圖中繪於算術尺度上之曲綫，頗呈不對稱之狀態，項數最密集之點係迫近於尺度之下端，而曲綫向右延長，漸見尖細，越出本圖尺度範圍以外而成長尾狀。此種情形可參看上表而益明瞭。例如以100作爲基數，其最小之比價爲36，在尺度上小於基數64點；其最大之比價爲3009，大於基數2909點。在一般物價水準劇漲之時，

此項比價分配偏斜之狀態亦所常有(按戰時工業局指數一九一八年之物價比較基期物價高 94%)。此種情形雖可解釋衆數組位置偏低之原因,但未說明此分配之性質。在說明此次數曲綫之形態時,吾人應解答關於編製指數之基本問題,即吾人應採用何種平均數,求一單純數字以表示物價變動百分率之集中趨勢乎?



圖四十九. 次數多邊圖:1918年物品1437種之比價分配

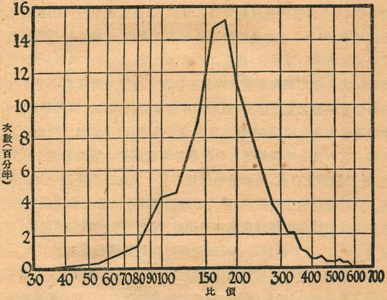
(1913年七月至1914年六月之平均價=100)

物價變動率之平均問題

以百分率表示之物價增高率,其高度並無限制。物價增高100%,500%,1000%,或1000%以上,事實上均有可能。美國戰時工業局研究戰時物價嘗發現最大之增高率為 acetiphenetidin 一物之4981%。但物價跌落百分率最多不過100%,此時物價格蓋已跌落至0矣。圖中曲綫之呈現偏斜狀態者,即由於此。如比價項數極多,則製表後繪於算術尺度上之次數曲綫,即呈現此種特性,而當物價上漲時,此種特性尤為明顯。

次數分配之呈現偏斜狀態與平均數之選擇有關,此點在討論平均

數時已有論及。如所欲平均之各項數值爲比率而非絕對數，以採用幾何平均數爲宜。試根據前述之次數分配計算各種平均數，互爲比較，可證此說之不謬。表中比價1437項之算術平均數爲217，衆數爲180，故算術平均數與衆數之距離相差頗遠。任何平均數如其與衆數相差之距離過遠，則此平均數顯已失去其代表性。根據同一資料求得之幾何平均數爲194，此數值與衆數及中位數（191）均極接近。其他姑勿置論，即以此各種平均數之關係而言，此種分配自以採用幾何平均數而捨棄算術平均數爲宜也。



圖五十．　次數多邊圖：繪於對數尺度上之1918年物品1437種之比價分配
（1913年七月至1914年六月之平均價＝100）

吾人在論述平均數時，嘗論及凡次數分配繪於對數尺度上之曲綫比較繪於算術尺度上之曲綫更與常態曲綫相近時，即表示該分配之變動比較合乎幾何法則，而不合於算術法則。圖五十係繪於對數尺度上，表示以一九一三年七月至一九一四年六月爲基期所得一九一八年1437

項比價之分配。前在算術尺度上繪成曲綫所呈現顯著之不對稱狀態，今在對數尺度上已不復存在，而其曲綫與常態曲綫極爲相似。

物價比率變動之合乎幾何法則，雖可視爲平均此種比率應採用幾何平均數之強有力之理由，但不能遽謂除幾何平均數以外，其餘各種平均數卽不宜供作編製物價指數之用。就大體言之，幾何平均數固不失爲一種合理之平均數，惟優良之指數儘有採用他種平均數者，此點容後再行討論。

茲將前數頁內所述及之理論，再作簡略之說明。在確定指數編製方法以前，吾人首應研究統計資料之性質，以及此種資料所組成之次數分配之狀態，俾便決定普通統計方法之是否適用。此處之統計資料爲以比率表示之各個物價變動，彙合許多物價比率卽可得一次數分配。吾人已知此種次數分配頗與合乎差誤常態法則之各種統計資料之分配相似，而常含有顯著之集中趨勢，故可用平均數以代表之；惟若時期愈長，則集中趨勢亦漸欠明顯，離散之程度愈見分明，而平均數之代表性乃愈見減弱。吾人又知物價變動之分配常呈偏斜狀態，而在物價上漲時爲尤甚。此種偏態之存在，乃由於物價比率有一定之下限而無一定之上限所致。在前此討論平均數時，吾人已知物價比率之分佈合乎幾何法則，故欲平均此項變動率應採用幾何平均數。惟此項結論雖無可非議，但在某種情形之下，編製指數常有採用他種平均數較爲相宜者。

編製指數之應採用何種平均數之問題，以及聯帶發生之加權方法以何者爲最優之問題，皆爲討論指數編製法之中心。假定批發物價指數之編製目的係爲測度一般物價之平均變動率，則應採用何種平均數，上節已論及之，惟因加權問題乃引起新因素，故須將加權問題併入討論。

編製指數採用加權方法之主要理由，前嘗述及。指數不論其爲測度生產數量之增減，貿易額之消長，或物價之漲落，其所根據之各項數量及比率，重要程度並不相等。例如編製生產指數，煤在生產方面之地位

比鉛爲重要；編製批發物價指數時，小麥價格比較胡蔴子價格更爲重要。欲測定指數中所包含各品之比重程度，方法不一而足，大抵權數之選擇，須視指數編製之目的及搜集權數資料有無困難而定。下節將舉示權數之實例數種，然後根據此種實例再作詳盡之討論。

編製指數之各種方法

現時編製批發物價指數所用方法不一，其所由不同之原因頗多。其一，編製指數究以何種方法爲最優，在理論上迄無定論。其二，從事統計工作之各機關，其財力人力均不相同，故實際上所用方法不能一致。其三，編製指數之目的不同，故所用方法亦不能相同。

編製指數方法既不相同，結果勢難一致。吾人可就同一資料，應用各種方法，將其結果互相比較，則對其不同之處可易於明瞭。表四十四所示一九一〇年至一九二三年每年十二月一日十二種主要農畜品之農場價格，爲說明各種方法時所用者。

符號之意義

說明各種指數計算法所用之符號，具有下述之意義：

p_0'：基期("0")第一種物品之價格。

q_0'：基期同一物品之數量。

p_1'：計算時期("1")同一物品之價格。

q_1'：計算時期同一物品之數量。

p_0''：基期第二種物品之價格。

q_0''：基期第二種物品之數量。

p_1''：計算時期第二種物品之價格。

q_1''：計算時期第二種物品之數量。

$\frac{p_1'}{p_0'}$：比價(計算時期第一種物品之價格與基期同一物品價格之百

表四十四

1910—1923年每年十二月一日主要農產品十二種之農場價格(註一)

農產品	單位	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923
玉蜀黍	蒲式耳	$ .480	$ .618	$ .487	$ .691	$ .644	$ .575	$ .889	$.1279	$1.365	$1.347	$ .677	$0.423	$ .657	$ .727
棉花	磅	.141	.088	.119	.122	.068	.113	.196	.277	.276	.356	.140	.162	.238	.310
乾草	公噸	12.14	14.29	11.79	12.43	11.12	10.63	11.22	17.09	20.13	19.55	16.72	12.11	12.59	14.07
小麥	蒲式耳	.883	.874	.760	.799	.986	.919	1.603	2.008	2.042	2.151	1.443	.926	1.009	.928
燕麥	蒲式耳	.344	.450	.319	.392	.438	.361	.524	.666	.709	.715	.472	.302	.394	.415
白番芋	蒲式耳	.557	.799	.505	.687	.487	.617	1.461	1.228	1.193	1.606	1.164	1.101	.582	.823
糖(註二)	磅	.0393	.0494	.0405	.0354	.0392	.0520	.0569	.0672	.0728	.0728	.0577	.0370	.0570	.0747
大麥	蒲式耳	.578	.869	.505	.537	.543	.516	.881	1.137	.917	1.210	.707	.419	.525	.540
菸草	磅	.093	.094	.108	.128	.098	.091	.147	.240	.280	.390	.211	.199	.231	.203
胡蘆子	蒲式耳	2.317	1.821	1.147	1.199	1.260	1.740	2.486	2.966	3.401	4.383	1.766	1.451	2.114	2.108
黑麥	蒲式耳	.715	.832	.663	.634	.865	.834	1.221	1.660	1.516	1.345	1.278	.697	.692	.647
米	蒲式耳	.678	.797	.935	.858	.924	.906	.889	1.896	1.918	2.668	1.189	.952	.934	1.103

(註一)表內1910—1920年之物價係 Warren M. Persons 氏所搜集,參看"Review of Economic Statistics"初卷第三號第105頁中"Fisher's Formula for Index Numbers". 1921—1923年之物價係根據"Weather, Crops and Markets".

(註二)各年糖價係每年十二月份粗製糖之批發價格,該品之農場價格無從搜集.

分比)。

$\frac{q_1'}{q_0'}$:比量。

P_0:基期之物價水準。

P_1:計算時期之物價水準。

簡單物價指數

費暄氏 (Irving Fisher) 分析各種指數計算法時(註),嘗分別為六種基本方法:即綜合法 the aggregative(或稱實價綜合法 price aggregative)、算術平均法、倒數平均法、幾何平均法、中位數法及衆數法。其中衆數法事實上絕少採用,可勿具論,其餘五種方法之特性,將以最簡單之方式說明之,然後再研討比較繁複之方式。

(一) 實價綜合法

簡單綜合法指數之編製,係將某時期之各項物價一一相加,依此求得各時期物價之總和,互為比較,即可測定一般物價之變動。用前述符號表示之,得

$$\frac{P_1}{P_0}=\frac{\Sigma p_1}{\Sigma p_0}$$

茲根據表四十四之數字,計算簡單綜合法指數如表四十五。表內第(2)欄為實價之綜合數,為便於比較起見,將此項數字一律化為百分比,而以一九一〇年之綜合價格作為基數,如第(3)欄。

用此法計算物價指數所得之結果,將與根據同一資料用他法計算所得之結果互為比較。實價綜合所得之指數既非不加權指數,又非權數相等之指數,因每一物品影響於指數之勢力大小,全視每單位之價格而定,而各品單位之大小,乃隨市場習慣為轉移,此為實價綜合法之主要

(註)見"The Making of Index Numbers",一九二二年 Houghton Mifflin Co. 出版.

缺點。在本指數中，乾草係以噸計價，故其權數較其餘十一品之總權數猶大，胡蔴子之重要程度次之。用實價相加所求得之指數，因加權方法全不合理，故不能表示該項農產物價之變動。

表四十五

農產品物價指數

(實價綜合法)

(1) 年份	(2) 指數 (實價綜合數)	(3) 指數，百分比 (1910=100)
1910	$18.9653	100
1911	21.5814	114
1912	17.3785	92
1913	18.5124	98
1914	17.4722	92
1915	17.3540	91
1916	21.5739	113
1917	30.5142	161
1918	33.8198	178
1919	35.7938	189
1920	25.8247	136
1921	18.7790	99
1922	20.0230	106
1923	21.9437	116

為欲避免因物品買賣單位不同所發生權數輕重不等之弊病起見，於是有將各品價格折算為同一單位之辦法，例如乾草、米、玉蜀黍、棉花以及其他物品一律折合為每磅之價格，然後將此項價格一一相加，以求指數者，勃蘭特斯脫指數(Bradstreet's index)即係採用此法計算。實則此法不過將前述不合理之加權方法，易以同樣不合理之另一種加權方法，且此法原欲使各品之權數相等，而事實上固未嘗予各品以相等之權數。例如一九一〇年乾草每磅值$.00607，棉花每磅$.141，米每磅$.015，

故在各品每磅價格之綜合數內，棉花之權數九倍於米，二十三倍於乾草矣。

二　比價之算術平均數

編製指數所用之另一方法，係將每品之價格，各以該品基期價格作為100化成百分比，然後根據各個百分比，採取通常所用之平均數方法，求其平均，卽得所求之指數。下表係就兩年之資料，以一九一〇年作為基期，說明此項計算手續之初步。

表四十六

編製指數所用比價之計算

(1) 物品	(2) 單位	(3) 1910年價格	(4) 比價	(5) 1911年價格	(6) 比價
玉蜀黍	蒲式耳	$ .480	100	$ .618	128.8
棉花	磅	.141	100	.088	62.4
乾草	公噸	12.14	100	14.29	117.7
小麥	蒲式耳	.883	100	.874	99.0
燕麥	蒲式耳	.344	100	.450	130.9
白番芋	蒲式耳	.557	100	.799	143.5
糖	磅	.0393	100	.0494	125.6
大麥	蒲式耳	.578	100	.869	150.2
菸草	磅	.093	100	.094	101.1
胡蔴子	蒲式耳	2.317	100	1.821	78.7
黑麥	蒲式耳	.715	100	.832	116.2
米	蒲式耳	.678	100	.797	117.5
			1200		1371.6

由上表數字可計算兩個年份比價之算術平均數。每一比價之公式為$\frac{p_1}{p_0}$，如其項數為N，則求計算時期之指數所用之公式為

$$\frac{\Sigma\left(\frac{p_1}{p_0}\right)}{N}$$

茲以表內數字為例，計算指數如下：

$$\text{指數(1910年)}=\frac{1200}{12}=100$$

$$指數(1911年)=\frac{1371.6}{12}=114.3$$

依此計算一九一〇年至一九二三年之各年指數見表四十九第(3)欄。

比價之簡單平均數含有加權之意味

前節所述指數計算法，通常稱爲比價之"未加權"指數，但按諸實際仍含有加權之意味，與前此所舉實價綜合法之兩例相似。此處各品用作權數之數量，卽爲該品在基年得價$100所售出之數量。在前例中用作權數之各品數量如下：

玉蜀黍	208.3	蒲式耳
棉花	710.0	磅
乾草	8.24	公噸
小麥	113.3	蒲式耳
燕麥	291.0	蒲式耳
白番薯	180.0	蒲式耳
糖	2650.0	磅
大麥	173.2	蒲式耳
菸草	1076.0	磅
胡蔴子	43.2	蒲式耳
黑麥	140.0	蒲式耳
米	147.7	蒲式耳

是以比價簡單平均之計算，實不啻依照上列各數量，決定在十一年間每年各項物品所可售得之總價值。按照一九一〇年之價格，上列數量之各項物品，均可售得$100，故其總價值爲$1200；按照一九一一年之價格，上列數量之物品共可售得$1371.60。此兩年之總價值各以12除之，得表四十九第(3)欄內之指數：一九一〇年指數爲100，一九一一年爲114(114.3)等等。由是知"比價之未加權指數"實際卽係實價之加權綜合數，所謂各品權數相等者，僅可謂一九一〇年基期內用作權數之數量，其價值各爲$100而已(註)。

(註)此項簡單平均比價之特性，爲 F. R. Macauly 氏所發現，見一九一五年十二月"American Economic Review"第928頁 Macauly 氏著作中。

比價之中位數

計算每年平均比價，亦有捨算術平均數而用中位數者。其計算法係先將表四十六第(6)欄內之比價，依大小次序排列，得下列之分配：

62.4	117.7
78.7	125.6
99.0	128.8
101.1	130.9
116.2	143.5
117.5	150.2

上列數字分配內，最小之比價爲 62.4，最大之比價爲 150.2，中位數爲 117.6。此中位數之值卽一九一一年之指數也。依比價之中位數求得各年之指數見表四十九第(4)欄。

比價之幾何平均數

各年比價亦可用幾何平均數求其平均，以與前述各例所得之結果相比較。如以 $\frac{p_1'}{p_0'}$ 表示一個單純比價，則求比價N項之幾何平均數之公式爲

$$Mg=\sqrt[N]{\frac{p_1'}{p_0'}\times\frac{p_1''}{p_0''}\times\frac{p_1'''}{p_0'''}\times\cdots\cdots}$$

幾何平均數通常係用對數計算，故上式可變爲

$$Log\ Mg=\frac{log\left(\frac{p_1'}{p_0'}\right)+log\left(\frac{p_1''}{p_0''}\right)+log\left(\frac{p_1'''}{p_0'''}\right)+\cdots}{N}$$

茲說明一九一〇年及一九一一年比價之幾何平均數之計算方法如下。表內各品之比價係根據表四十六重行列入者。

表四十七

比價幾何平均數之計算

(1) 物　品	(2) 1910年比價	(3) 第(2)欄數字之對數	(4) 1911年比價	(5) 第(4)欄數字之對數
玉蜀黍	100	2.0	128.8	2.10992
棉　花	100	2.0	62.4	1.79518
乾　草	100	2.0	117.7	2.07078
小　麥	100	2.0	99.0	1.99564
燕　麥	100	2.0	130.9	2.11694
白番芋	100	2.0	143.5	2.15685
糖	100	2.0	125.6	2.09899
大　麥	100	2.0	150.2	2.17667
菸　草	100	2.0	101.1	2.00475
胡麻子	100	2.0	78.7	1.89597
黑　麥	100	2.0	116.2	2.06521
米	100	2.0	117.5	2.07004
		24.0		24.55694

$$Log\ Mg(1910年)=\frac{24}{12}=2$$

$$Mg=anti\text{-}log\ 2=100$$

$$Log\ Mg\ (1911年)=\frac{24.55694}{12}=2.04641$$

$$Mg=anti\text{-}log\ 2.04641=111.3$$

此 111.3 之數值卽一九一一年之指數也。依此求得之各年指數，見表四十九第(5)欄。

(五)　比價之倒數平均數

倒數平均數之特質，已在第四章內有所論述。吾人猶憶倒數平均數之倒數，卽係各數值倒數之算術平均數。本例各數值卽爲以$\frac{p_1'}{p_0'}$表示之比價，每一比價之倒數爲$\frac{p_0'}{p_1'}$，故求比價N項之倒數平均數之公式爲

$$\frac{1}{H}=\frac{\frac{p_0'}{p_1'}+\frac{p_0''}{p_1''}+\frac{p_0'''}{p_1'''}+\cdots\cdots}{N}$$

或

$$H=\frac{N}{\Sigma\left(\frac{p_0}{p_1}\right)}$$

其計算方法如下表所示：

表四十八
比價倒數平均數之計算

(1) 物品	(2) 1910年比價	(3) 第(2)欄數字之倒數	(4) 1911年比價	(5) 第(4)欄數字之倒數
玉蜀黍	100	.01	128.8	.007763975
棉花	100	.01	62.4	.016025640
乾草	100	.01	117.7	.008496177
小麥	100	.01	99.0	.010101010
燕麥	100	.01	130.9	.007639419
白番芋	100	.01	143.5	.006968641
糖	100	.01	125.6	.007961783
大麥	100	.01	150.2	.006657790
菸草	100	.01	101.1	.009891197
胡麻子	100	.01	78.7	.012706480
黑麥	100	.01	116.2	.008605852
米	100	.01	117.5	.008510638
		.12		.111328602

$$H(1910年)=\frac{12}{.12}=100$$

$$H(1911年)=\frac{12}{.111328602}=107.8$$

依此求得之各年指數，見表四十九第(6)欄。

表四十九
1910—1923年農產物價指數
(1910年=100)

(1) 年份	(2) 實價綜合數(以百分比表示)	(3) 比價之算術平均數	(4) 比價之中位數	(5) 比價之幾何平均數	(6) 比價之倒數平均數
1910	100	100	100	100	100
1911	114	114	118	111	108
1912	92	95	93	92	90
1913	98	104	98	100	97
1914	92	101	102	97	91
1915	91	104	104	102	101
1916	113	156	152	151	147
1917	161	208	208	204	198
1918	178	215	209	210	205
1919	189	252	226	241	231
1920	136	151	143	145	139
1921	99	115	99	107	101
1922	106	129	114	124	119
1923	116	142	134	135	129

上述五種指數之計算，未嘗採用合理之加權制度，故該五種指數皆稱爲"未加權"指數。實則此名詞易滋誤解，蓋第一種根據實價綜合法求

得之指數，實爲加權甚重之指數，惟權數不合理耳；其後四種指數，實際亦係加權指數，每品所用權數爲一九一〇年$100所可購得該品之數量。茲將此五種指數併列一表(表四十九)，互爲比較。各年指數皆取整數，整數後一位之小數，四捨五入。

該表內各指數繪爲曲綫，如圖五十一。

各種簡單指數之比較

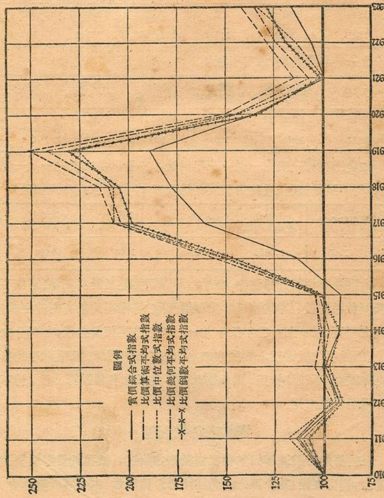
用比價平均法求得之四種指數，彼此極爲相似，惟與實價綜合式指數相比，則其相似程度較淺。實價綜合法之缺點，前嘗論及，故用此法測度物價變動殊不可靠。至於其餘四種指數中，算術平均數、幾何平均數及倒數平均數之指數，由於該平均數本身之性質，恆有一定之關係，即除基年以外，幾何平均指數恆較算術平均指數爲小，而倒數平均指數又恆較幾何平均指數爲小，且各項比價愈分散，此差數亦愈增大。中位數指數則因每期所平均者祇有十二項數字，故其數值不穩定，與其他三種平均指數之關係，遂難於確定。

吾人對於此種數值各不相同之指數，將何所取捨乎？在各種"未加權"指數中實無一足稱完備者，因其所採用之權數，並不能表示指數內所包含各品之相對的重要程度也。茲姑將權數問題暫勿置論，吾人亦可設法測驗物價變動所用各種測度方法之適當否乎？

時間顚倒測驗法

費暄氏(Irving Fisher)嘗用所謂時間顚倒測驗法 (time reversal test)以測驗指數計算方法之優劣。此法實即測驗應用每種方法，計算前後兩時期之指數，經互換基期後所得之結果是否合理耳。例如糖每磅價格自一九一〇年之 $.04漲至一九一一年之 $.08，故一九一一年之價格，應爲一九一〇年之200%，一九一〇年之價格，應爲一九一一年之50%。一數應爲另一數之倒數，而其乘積(2.00×.50)應等於一。據此原則，如

用某法計算指數，而得某年一般物價之水準爲其前一年水準之200%，則時間顚倒後，第一年之物價水準應爲第二年水準之 50%。故如根據任何兩年之物價，採用同一種指數計算法，將基期互換所求得之兩個指數，應互爲倒數，而其乘積應等於一，否則此指數計算法必有偏誤之處。



圖五十一．　1910—1923年農產物價五種簡單指數之比較(1910年＝100)

此項測驗可仍就一九一〇年及一九一一年之物價資料，以求前述各種指數計算法之是否合於上項測驗。以一九一〇年爲基期，求得指數如下：

年 份	實價綜合數（以百分比表示）	比價之算術平均數	比價之中位數	比價之幾何平均數	比價之倒數平均數
1910	100	100	100	100	100
1911	113.79414	114.3	117.6308	111.3	107.8

以一九一一年爲基期，求得指數如下：

年 份	實價綜合數（以百分比表示）	比價之算術平均數	比價之中位數	比價之幾何平均數	比價之倒數平均數
1910	87.87799	92.76	85.0117	89.85	87.47
1911	100	100	100	100	100

如將上列第一表內一九一一年指數，各乘以第二表內一九一〇年指數，得下列數值（相乘時指數須化爲比率，並非即以百分比相乘）：

實價綜合數	比價之算術平均數	比價之中位數	比價之幾何平均數	比價之倒數平均數
1.00	1.0602	1.00	1.00	.9429

上列各種指數計算法中，合乎時間顚倒測驗者計有三種。算術平均數及倒數平均數皆不能適合時間顚倒測驗。算術平均數有偏高之弊，其偏高之程度，以一九一〇年及一九一一年之差誤相混合後，在百分之六以上。倒數平均數偏誤之程度亦如之，惟其方向相反。此兩種平均數所有之偏誤，如不設法校正，則根據此項平均數之各種計算方法，皆不宜供作編製指數之用。

指數之加權問題

在前節內已將五種簡單物價指數略加論述，惟指數常因權數問題之羼入，變化複雜，其計算公式之種類乃大爲增加。本書所欲論述者，則爲比較重要之數種。

吾人欲求測度物價變動之準確指數，必須採用合理之權數，此種權數必須確能表示指數內各項物品之相對的重要程度。如權數問題未加注意，即難免發生偶然的加權或權數不合理之弊。

茲仍以前述各例所用之物價爲例，以說明加權方法及指數所受權數變更之影響。編製此項農產物價指數所用權數，或爲農產物之生產數量，或爲其生產價值，全視所選指數方式而定。各項物品一九一〇年至一九二三年之生產數量如表五十所示(註)：

實價加權綜合法

由實價綜合法所得指數，其不合理之處，前已論述。惟在實價相加之先，如各以適當之權數乘之，則可免不合理之結果。如所用權數爲基年(即"0"年)之生產數量，則實價加權綜合公式爲

$$\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$$

此法爲美國勞工統計局所採用者，惟該局所用權數非基年之數量，而爲另一年之數量而已。加權綜合式指數之計算手續，將在表五十一中說明之。

將表中第(5)及第(8)欄內之數值分別相加，以求其總價值。以兩年中任何一年作爲基年，而以另一年之總價值化爲其百分比，即得百分比之指數。例如以一九一〇年之總價值作爲基數，則一九一一年之指數爲111.5。依此求得其他各年之指數見表五十四第(2)欄。

加權綜合價值之計算，亦有不採用基期之權數，而用後一時期之權數者，則屬於另一種之加權綜合式指數。例如吾人可用q_1（第一計算年"1"之數量)爲權數，以比較第一計算年"1"與基年"0"之物價，但亦可用

(註)一九一〇年至一九一九年之生產數量係根據合氏(E. E. Day)所著"An Index of the Physical Volume of Production" 第8頁，一九二一年出版(係根據一九二〇年九月 "Review of Economic Statistics" 所載原文翻印者)。一九二〇年至一九二三年之生產數量，係根據 "Weather, Crops and Markets".

表五十

1910—1923年農產品十二種之每年產量

年份	玉蜀黍(百萬蒲式耳)	棉花(百萬包)(註一)	乾草(百萬公噸)	小麥(百萬蒲式耳)	燕麥(百萬蒲式耳)	白番芋(百萬蒲式耳)	糖(百萬磅)(註二)	大麥(百萬蒲式耳)	菸草(百萬磅)	胡蔴子(百萬蒲式耳)	黑麥(百萬蒲式耳)	米(百萬蒲式耳)
1910	2886	11.61	69.38	635.1	1186	249.0	1618	173.8	1108	12.72	34.90	24.51
1911	2531	15.69	54.92	621.3	922	292.7	4296	160.2	905	19.37	33.12	22.93
1912	3125	13.70	72.69	730.3	1418	420.6	3764	223.8	963	28.07	35.66	25.05
1913	2447	14.16	64.12	763.4	1122	331.5	3941	178.2	854	17.85	41.38	25.74
1914	2673	16.14	70.07	891.0	1141	409.9	4134	195.0	1035	13.75	42.78	23.65
1915	2995	11.19	85.92	1025.8	1549	359.7	4230	228.8	1062	14.03	54.05	28.95
1916	2567	11.45	91.19	636.3	1252	287.0	4671	182.3	1153	14.30	48.86	40.86
1917	3065	11.30	83.31	636.7	1593	442.1	3950	211.8	1249	9.16	62.93	37.74
1918	2503	12.04	76.66	921.4	1538	411.9	4220	256.2	1439	13.37	91.04	38.61
1919	2917	11.42	91.33	941.0	1248	357.9	3634	165.7	1390	8.92	88.48	41.06
1920	3209	13.44	87.85	833.0	1497	403.3	4665	189.3	1582	10.77	60.5	52.07
1921	3069	7.95	82.38	814.9	1078	361.7	....	154.9	1070	8.03	61.7	37.61
1922	2906	9.76	95.88	867.6	1216	453.4	....	182.1	1247	10.37	103.4	41.40
1923	3054	10.08	98.90	785.7	1230	412.4	....	198.2	1475	17.43	63.0	33.26

(註一)每包毛重500磅.

(註二)糖之產量數字,係各該收穫年美國大陸所產甜菜糖及蔗糖之數量,再加上該收穫年七月一日起之會計年度內自美國殖民地輸入之數量之總額。計算1921年至1923年指數時,係根據估計之數量。

q_2(第二計算年"2"之數量)爲權數,以比較第二計算年"2"與基年"0"之物價。今以代數式表示之,則求第一計算年"1"指數之公式爲

$$\frac{\Sigma p_1\ q_1}{\Sigma p_0\ q_1}$$

用此式計算指數之手續,與前例完全相同,惟所用權數則須每年更換一次。依此求得之各年指數見表五十四第(3)欄。

表五十一

實價加權綜合式指數之計算

(1) 物品	(2) 單位	(3) 1910年價格	(4) 權數(1910年生產數量,單位百萬)	(5) 價格×權數(p_0 q_0)	(6) 1911年價格	(7) 權數(1910年生產數量,單位百萬)	(8) 價格×權數(p_1 q_0)
玉蜀黍	蒲式耳	$ .480	2,886	$1,385,280,000	$ .618	2,886	$1,783,548,000
棉花	磅	.141	5,805	818,505,000	.088	5,805	510,840,000
乾草	公噸	12.14	69.38	842,273,200	14.29	69.38	991,440,200
小麥	蒲式耳	.883	635.1	560,793,300	.874	635.1	555,077,400
燕麥	蒲式耳	.344	1,186	407,984,000	.450	1,186	533,700,000
白番芋	蒲式耳	.557	349.0	194,393,000	.799	349.0	278,851,000
糖	磅	.0393	3,618	142,187,400	.0494	3,618	178,729,200
大麥	蒲式耳	.578	173.8	100,456,400	.869	173.8	151,032,200
菸草	磅	.093	1,103	102,579,000	.094	1,103	103,682,000
胡蔴子	蒲式耳	2.317	12.72	29,472,240	1.821	12.72	23,163,120
黑麥	蒲式耳	.715	34.90	24,953,500	.832	34.90	29,036,800
米	蒲式耳	.678	24.51	16,617,780	.797	24.51	19,534,470
				$4,625,494,820			$5,158,634,390

前述兩例中之權數皆係數量,因價格乘數量後得價格之總數;但若加權於比價時,卽不能用數量作權數。此種抽象比價必須用價值作爲權數,則所得乘積方可比較,因價值有一共同之貨幣單位,而數量單位則各品互異故也。用作權數之價值,可用各種方式求得之。

費暄氏(註)嘗列擧計算價值之四種方式如下,其中第二第三種係屬混合式。

(註)見"The Making of Index Numbers"第54頁。

I. 每品之權數＝基年價格×基年數量(p_0q_0)

II. 每品之權數＝基年價格×計算年數量(p_0q_1)

*III.*每品之權數＝計算年價格×基年數量(p_1q_0)

IV. 每品之權數＝計算年價格×計算年數量(p_1q_1)

上列各種價值權數皆含有所謂權數之偏誤(weight bias)(以數量作爲權數，即無偏誤發生)，正與某數種平均數之含有偏誤相似。第一第二兩種權數(係按基年價格求得之價值)常偏低，第三第四兩種權數(係按計算年價格求得之價值)常偏高，此點既可以數理證明之(註)，亦可以數字證實之。

茲爲便於說明起見，下述各例皆以一九一〇年基年各品生產量之價值作爲權數。此項價值見表五十二第(3)欄。各品權數均以百萬爲單位，所以便於計算也。

比價之加權算術平均數

此種指數之計算，係將每品之比價各與其權數相乘，而後以此乘積

(註)用第三種權數求得之指數，必較用第一種權數求得之指數爲高。每品之比價如以第三種權數乘之，得

$$\frac{p_1}{p_0}\times p_1q_0$$

如以第一種權數乘之，得

$$\frac{p_1}{p_0}\times p_0q_0$$

倘p_1大於p_0(即比價大於100)，則第三種權數(p_1q_0)必大於第一種權數(p_0q_0)，故比價在100以上之各品，如用第三種權數，其所得重量必較用第一種爲大；倘 p_1小於p_0，則第三種權數(p_1q_0)必小於第一種權數(p_0q_0)，故比價在100以下之各品，如用第三種權數，其所得重量必較用第一種爲小。今以價格高漲之物品所得重量過大，價格跌落之物品所得重量過小，其結果按照第三種權數求得之指數，恆較按照第一種權數求得之指數爲大。第四種權數與第二種權數比較，所得情形亦同。至於第一種與第四種權數之間，並無一定之關係，但就大體而論，用第四種權數求得之指數，當較用第一種權數求得者爲大，蓋基年之權數恆偏低，而計算年之權數恆偏高也(關於權數偏誤問題，費暄氏嘗作詳盡之討論，見 "The Making of Index Numbers" 第五章第384－387頁)。

相加，除以權數之和。其計算手續如下表所示：

表五十二　比價加權算術平均數之計算

(1) 物品	(2) 1910年 比價	(3) 權數	(4) 比價×權數	(5) 1911年 比價	(6) 權數	(7) 比價×權數
玉蜀黍……	100	$1,385	$138,500	128.8	$1,385	$178,388.0
棉　花……	100	819	81,900	62.4	819	51,105.6
乾　草……	100	842	84,200	117.7	842	99,103.4
小　麥……	100	561	56,100	99.0	561	55,539.0
燕　麥……	100	408	40,800	130.0	408	53,407.2
白番芋……	100	194	19,400	143.5	194	27,839.0
糖……	100	142	14,200	125.6	142	17,835.2
大　麥……	100	100	10,000	150.2	100	15,020.0
菸　草……	100	103	10,300	101.1	103	10,413.3
胡蔴子……	100	29	2,900	78.7	29	2,282.3
黑　麥……	100	25	2,500	116.2	25	2,905.0
米……	100	17	1,700	117.5	17	1,997.5
		$4,625	$462,500		$4,625	$515,835.5

（表內權數係1910年基年各品生產數量之價值，以百萬元為單位。）

$$\text{加權算術平均數(1910年)} = \frac{\$462,500}{\$\ \ 4,625} = 100.$$

$$\text{加權算術平均數(1911年)} = \frac{\$515,835.5}{\$\ \ 4,625} = 111.5.$$

於此所應注意者，此處一九一一年之指數，與根據表五十一計算所得之指數完全相同。按後者係用實價加權綜合法計算者，其權數為基年之生產數量。比價算術平均式之指數，如用基年價值作為權數，則恆與根據此種綜合價值所化成百分比之指數相等(註)。

(註)此點可以代數式證明之。今設各品之基年價值為 p_0q_0，其第二年之比價為 $\frac{p_1}{p_0}$，則此種比價之加權算術平均數等於

$$\frac{\frac{p_1'}{p_0'} \times p_0'q_0' + \frac{p_1''}{p_0''} \times p_0''q_0'' + \frac{p_1'''}{p_0'''} \times p_0'''q_0''' + \cdots\cdots}{p_0'q_0' + p_0''q_0'' + p_0'''q_0''' + \cdots\cdots}$$

由上式化為

$$\frac{\Sigma p_1q_0}{\Sigma p_0q_0}$$

此式即前述實價加權綜合法之公式。

比價倒數平均數之公式，如其權數為第二年之數量，可化為

$$\frac{\Sigma p_1q_1}{\Sigma p_0q_1}$$

此式亦即以第二年數量作為權數之實價加權綜合法之公式也。

比價之加權幾何平均數

加權幾何平均數之計算手續，與未加權之幾何平均數相同，惟前者每一比價之對數，須與權數相乘，再以此項已加權之對數之和，除以權數之和，其結果卽爲所求指數之對數(註)。其計算法在下表內說明之。

表五十三

1911年比價加權幾何平均數之計算

(1910年＝100)

物　品	1911年比價	比價之對數	權數(1910年生產價值)	比價之對數×權數
玉蜀黍……	128.8	2.10992	1,385	2922.23920
棉花……	62.4	1.79518	819	1470.25242
乾草……	117.7	2.07078	842	1743.59676
小麥……	99.0	1.99564	516	1119.55404
燕麥……	130.9	2.11694	408	863.71152
白番芋……	143.5	2.15685	194	418.42890
糖……	125.6	2.09899	142	298.05658
大麥……	150.2	2.17667	100	217.66700
菸草……	101.1	2.00475	103	206.48925
胡蔴子……	78.7	1.89597	29	54.98313
黑麥……	116.2	2.06521	25	51.63025
米……	117.5	2.07004	17	35.19068
			4,625	9401.79973

$$Log\ Mg=\frac{9401.79973}{4625}=2.032822$$

$$Mg=107.9$$

如表所示，以一九一〇年爲基期之一九一一年指數爲 107.9。依此求得其他各年之指數見表五十四第(5)欄，與其他各種加權指數並列一表。

然則此三種加權指數，吾人將用何法區別其優劣乎？吾人可先用前次測驗五種簡單指數所用之時間顚倒測驗法，作同樣之測驗。該三種加權指數皆不合於時間顚倒測驗，其中加權幾何平均式指數亦非例外。

(註)加權幾何平均公式見第114頁。

按簡單幾何平均式指數雖合於此項測驗，但一經加權，指數卽含有偏誤。以時間顚倒測驗而言，三種加權指數中無一能適應者。吾人可更以費暄氏所發現之第二種基本測驗法，卽所謂因子顚倒測驗法（factor reversal test）測驗之。

因子顚倒測驗法

每一物品任何一年之總價值，本係該品是年之生產數量與每單位之價格相乘之積；以代數式表示之爲 $p_0'q_0'$。某年總價值與前一年總價值之比爲 $\frac{p_1'q_1'}{p_0'q_0'}$。今若某年之價格及數量各增一倍，則其價格之百分比爲200，數量之百分比亦爲200，其價值之百分比應爲400，故第二年之總價值應爲第一年總價值之四倍。價值之百分比爲價格與數量兩個百分比相乘之積，此項關係在個別物品單獨觀察時，頗爲明顯。

如吾人根據多種物品編製某年與後一年間價格變動之指數，及某年與後一年間數量變動之指數，則此兩種指數之乘積，應等於第二年總價值對於第一年總價值之比率。如此項乘積並不與價值之比率相等，則可斷定價格及數量兩種指數中，必有一種含有差誤，或兩者皆含有差誤。

今吾人以前述第一種加權綜合式指數（公式爲 $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$）試驗其是否合於時間顚倒測驗。依照此項公式，可將 p 及 q 之地位互易，以求物量指數。其公式爲

$$\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$$

上式之分子分母所含之物價因子，數值相同，因吾人現欲測度者，爲數量變動之結果也。以前述十二種農產品之數字代入公式，得

$$1911年之物量指數\ (1910年=100)=\frac{\$4,446,264,630}{\$4,625,494,820}=.96125$$

以一九一〇年爲基期之一九一一年生產數量指數化爲百分率之形式得96.125，其加權綜合式之物價指數爲111.5，兩者之乘積爲

$$.96125 \times 1.115 = 1.0718$$

上式之意義，謂若物價增高11.5%，同時物量減低3.875%，則物值應增7.18%也。

吾人如直接由兩年之價值以求一九一一年價值之比率，得

$$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} = \frac{\$4,748,718,320}{\$4,625,494,820} = 1.02664$$

兩項結果相差$4\frac{1}{2}$%，可知加權綜合公式不合於因子顚倒測驗，故此式不能認爲完善。

吾人再以第二種加權綜合式指數試驗其是否合於此種試驗。按以一九一〇年爲基期之一九一一年指數如下：

$$物價指數 = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} = 106.8$$

$$物量指數 = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} = 92.05$$

$$乘積 = .9205 \times 1.068 = .9831$$

（在物價及物量指數兩者相乘時，係按照比率計算，並不按照百分比計算）。

依此求得之結果與物值比率相比較，仍相差4%以上，惟其差異之方向與前相反耳。

至於加權幾何平均式指數，亦不合於因子顚倒測驗。按幾何平均式與綜合式之簡單指數本無偏誤，而因權數之採用，以致指數含有差誤，惟欲使指數確切表現事實，則權數又不能捨棄不用。事實上無論簡單指數或加權指數，無一能同時合於此兩項基本測驗者。費暄教授嘗試驗公式四十六種，其中合於時間顚倒測驗者祇有四種（即簡單之幾何平均式、中位數式、衆數式及綜合式），而無一能合於因子顚倒測驗者。

理想公式

吾人爲矯正公式之偏誤，於是有配合法之採用。其法係將差誤方向相異之公式，求其幾何平均數，以消除其偏高偏低之差誤。費暄敎授嘗用此法，將所有計算公式一一加以配合，以試驗公式配合後，能否合於兩種基本測驗。其結果乃發現配合之公式十三種，皆能合於該兩項基本測驗，並由費暄氏選定其中準確可靠與計算簡便之一種，作爲理想公式("ideal" formula)。用此公式求得之指數，實卽前述兩種綜合式指數之幾何平均數。其公式(註)爲

$$\sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}$$

茲爲舉例說明起見，可將前已求得之兩個單純指數代入公式，以求理想指數。其一九一一年之指數爲

$$\text{理想指數} = \sqrt{111.5 \times 106.8}$$
$$= 109.14$$

此理想指數能合於時間顚倒及因子顚倒兩種測驗。茲先證明其合於第一種測驗如下：

$$1911\text{年物價指數}(1910\text{年}=100) = 109.14$$
$$1910\text{年物價指數}(1911\text{年}=100) = 91.63$$
$$1.0914 \times .9163 = 1.00$$

次證明其合於因子顚倒測驗。以一九一〇年爲基期，按照理想公式求得之一九一一年指數爲

$$\text{物價指數} = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} = 109.14$$

$$\text{物量指數} = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} = 94.07$$

(註)此式亦爲Bowley, Pigou, Walsh, Young 諸氏所各自發現者。見"The Making of Index Numbers" 第十五章240—242頁。

$$物值指數 = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} = 102.664$$

物價指數及物量指數之乘積 $= 1.0914 \times .9407 = 1.02668$

茲將一九一〇年至一九二三年之理想指數、理想公式所由配合之兩種加權綜合式指數、以及用基年價值作爲權數之幾何平均式指數，併列下表。此四種指數繪成曲綫，如圖五十二。

表五十四

1910—1923年農產物價加權指數之比較(註)

(1910年=100)

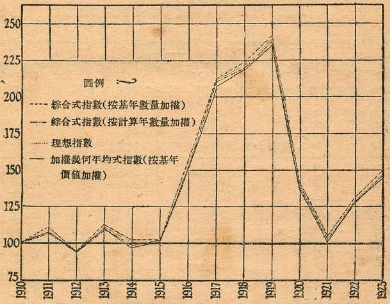
(1) 年份	(2) 綜合式指數 (按基年數量加權)	(3) 綜合式指數 (按計算年數量加權)	(4) 理想指數 (第(2)(3)欄指數之幾何平均數)	(5) 加權幾何平均式指數 (按基年價值加權)
	$\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$	$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$		
1910	100.0	100.0	100.0	100.0
1911	111.5	106.8	109.1	107.9
1912	94.4	93.7	94.0	94.0
1913	112.4	109.4	110.9	109.9
1914	102.8	98.9	100.8	97.0
1915	102.3	102.3	102.3	101.0
1916	156.8	151.3	154.0	151.2
1917	212.6	210.9	211.7	207.2
1918	224.2	220.3	222.2	219.3
1919	241.3	238.2	239.7	235.1
1920	139.9	137.9	138.9	137.2
1921	104.3	103.6	104.0	101.6
1922	131.6	127.9	129.7	128.9
1923	149.3	145.2	147.2	144.1

此四種加權指數互爲比較，其相差程度比各該式之簡單指數相互間之差度，已大見減小。各年指數中雖尙不免略有出入之處，惟前此簡單指數中所呈現奇特之差異，確已不復存在矣。

上述測度一九一〇年與後此各年物價平均變動之四種指數中，自以理想指數爲最優。於此所應注意者，此項指數係爲測度兩個指定時

(註) 1910—1919 年之兩種加權綜合式指數係 W. M. Persons 氏所編製者 (見 "Fisher's Formula for Index Numbers"，原文載 "Review of Economic Statistics" 初卷第三號第107頁)。

期間之變動，不能與中間各期指數直接比較。例如一九二三年之指數，係按照一九一〇年及一九二三年之價格及數量之關係求得，其間含有兩重權數，而權數又各年不同。如欲以一九二三年與一九二二年比較，則須編一新指數，而以一九二三年及一九二二年之價格及數量爲根據，其他各年之價格及數量不與焉。如根據上表所列該兩年之理想指數，直接比較，則因採用加權制度，每易發生差誤。



圖五十二．　1910—1923年農產物價四種加權指數之比較(1910年＝100)

幾何平均式指數，如其權數固定不變，則每年指數不特可與基年指數直接比較，且可與其他任何一年之指數直接比較，此爲幾何平均式指數之優點。其基期可由指數直接轉換，所得結果與根據原有物價更換基期重行計算者完全相同。理想指數如以同法轉換基期，雖不致發生重大

差誤，然不能得極準確之結果(註)。

理想公式之未能普遍採用，尚有其他重要原因在，蓋若採用此式，則用作權數之數量資料須每年或每月搜集一次，其間不無困難之處，而計算手續又頗費時間也。惟指數果欲力求準確，則計算手續之困難，要所不免。費暄氏以為指數如不採用理想公式，亦可採用下式以替代之，而計算手續之簡捷則遠過之。其式如下：

$$\frac{\Sigma(q_0+q_1)p_1}{\Sigma(q_0+q_1)p_0}$$

此式由形式簡單、計算便捷、結果準確及循環差誤徵渺之四點觀之，費暄氏認為最優良之公式，亦為愛奇渥斯氏(Edgeworth)及馬渉爾氏(Marshall)所最倡導者。採用此式所得結果，大抵與根據理想公式求得者，相差不及 $\frac{1}{4}$ %。下表係採用一九一〇年及一九一一年之資料，以說明其計算方法。

表五十五

按聯合數量加權之綜合指數之計算

(1) 物品	(2) 單位	(3) 1910年價格	(4) 1910年數量＋1911年數量(單位：百萬)	(5) 1910年價格×數量之總和 (3)×(4)	(6) 1911年價格	(7) 1911年價格×數量之總和 (6)×(4)
玉蜀黍	蒲式耳	.480	5,417	$2,600,160,000	$ .618	$3,347,706,100
棉花	磅	.141	13,650	1,924,650,000	.088	1,201,200,000
乾草	公噸	12.14	124.30	1,509,002,000	14.29	1,776,247,000
小麥	蒲式耳	.883	1,256.4	1,109,401,200	.874	1,098,093,600
燕麥	蒲式耳	.344	2,108	725,152,000	.450	948,600,000
白番芋	蒲式耳	.557	641.7	357,426,900	.799	512,718,300
糖	磅	.0393	7,914	311,020,200	.0494	390,915,600
大麥	蒲式耳	.578	334	193,052,000	.869	290,246,000
菸草	磅	.093	2,008	186,744,000	.094	188,752,000
胡蔴子	蒲式耳	2.317	32.09	74,352,530	1.821	58,435,890
黑麥	蒲式耳	.715	68.02	48,634,300	.832	56,592,640
米	蒲式耳	.678	47.44	32,164,320	.797	37,809,680
				$9,071,759,450		$9,907,352,710

(註)如指數注重於逐年之比較，則編製理想指數可採用連環基期制 (chain system)，先以每年之物價作為基價而求次年指數，得各年之環比指數 (link index number)，然後任擇一年為基期，將各年環比指數連續相乘，卽得以該年為基年之連環指數 (chain index number)。Warren M. Persons 氏嘗發現連環指數包含累積性之差誤，如連環之時間過長，常有發生重大差誤之虞。

$$\frac{\Sigma(q_0+q_1)p_1}{\Sigma(q_0+q_1)p_0}=\frac{\$9,907,352,710}{\$9,071,759,450}$$

$=109.2$(以1910年爲基期之1911年指數)

(此指數係以百分比表示者。)

此項公式所需數字資料，與理想公式相同，惟其中數量資料頗不易搜集，非在舉行普查時期，無從求得，故未值普查各年，不得不採用固定權數，此時或以採用加權綜合公式

$$\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$$

最爲相宜。加權幾何平均公式雖有種種優點，但受權數偏誤之影響。如權數資料不能蒐集，甚或不能約計，不得不採用簡單公式時，則比價之簡單幾何平均數及簡單中位數遠較其他各法爲優，而兩者之中，尤以幾何平均數更切實用。

各種指數之可靠性

物價指數皆係根據物價之樣本所編製，而求得之指數，則常視爲全體物價之代表，故吾人需要一種方法以判定各種指數之可靠性，並以試驗根據各種樣本求得之指數，是否有穩定性。就吾人經驗而言，根據各種樣本編製指數，每得不同之結果，然則究以何種指數在取樣上之差誤爲最小乎(註一)?

凱萊氏(Truman L. Kelley)(註二)嘗試算各種指數之機誤，並依機誤之大小分級評定各種指數之優劣。加權幾何平均式及加權中位數式兩種指數最爲可靠，所受取樣差誤之影響爲最小，故列居首位。費暄氏之理想指數位次較低，但仍較比價之加權算術平均式及倒數式之位次爲高。至於簡單未加權之算術平均式指數，則列居末位。

(註一)種樣問題與統計量數之可靠性有關之處，後文有更詳盡之討論。

(註二)見 Kelley 氏著 "Statistical Method' 第334—346頁。

凱萊氏認加權幾何平均指數爲最可靠,最有伸縮,而最優良之物價指數。此外用選擇之固定權數(未必定與交易量或消費量之數字完全相等者)之綜合式指數

$$\frac{\Sigma(p_1 w)}{\Sigma(p_0 w)}$$

因其具備優良指數之各要素,亦認爲最優之公式,其位次與加權幾何平均式相埒,而較理想公式爲高。此式中所用權數,不定爲實際數量,故解決加權問題,較有伸縮之餘地。

關於編製物價指數之其他問題

前文已將關於編製物價指數所用平均方法等技術問題,加以討論。在各種方法中,有不宜採用者,亦有適合於某種目的者。應用指數者務須透切瞭解該指數所採用之編製方法,俾可確知該指數之作用及其可靠程度,而不致誤解。

編製指數者所遇之問題,猶不限於前述數種,應用指數者所應注意者,亦非僅此數種。其與平均方法及加權問題同等重要者,厥爲關於選取樣本之實際問題。吾人如欲準確測度物價水準之變動,必須決定計算期內貨幣(包括信用)之流通數量與市場上一切物品實際交易數量之比率而後可,故欲測度兩時期間一般物價之變動,亦必須詳悉該兩時期內之該兩種要素。惟此項資料自不易蒐集,故物價指數之編製,不得不採用抽樣法。編製物價指數之最要問題,乃爲決定指數內所應包含之物品項數及其性質。

指數包含物品之項數問題

指數編製方法常隨目的而異,前已言之,故指數內所應包含之物品多寡及其種類,亦隨編製之目的爲轉移。指數如爲測度一般物價水準之

變動，則指數內應包含之物品項數問題，解答甚易，即樣本愈大，指數之代表性可以愈高，因樣本較大，其次數多邊圖必較樣本小者爲更能合於全體物價變動之理想曲綫也。美國勞工統計局指數係根據物價404項編製，勃蘭特斯脫指數係根據物價96項編製。後者雖有其特殊之優點，然就測度一般物價變動之標準而言，兩者比較，當以勞工統計局指數更爲可信，顧包含物品項數較少之指數，未必皆無價值而可委棄者。密哲爾氏嘗將物價項數多寡懸殊之數種指數，詳加比較。此項研究，足使吾人對於物價制度及指數之性質，可獲明確之認識。密氏試驗結果，發現各種指數變動之相似，頗爲意料所不及。根據少數物價編製之指數，所表示物價變動之趨勢，與根據數百項物價編製之指數大致相同；惟若細加觀察，則物價之漲落變動，不無些微差異，而此種差異常足使吾人對於計算時期之物價變動不能確定。遇此情形，則包含品目較多之指數，如物品選擇適當，確能代表物價制度內之各原素時，應爲最足代表一般物價之變動者也。

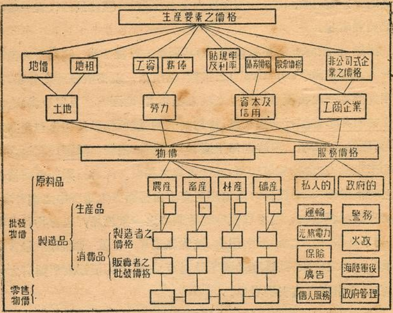
雖然，指數亦有爲某種目的，物品項數不妨從少者。例如吾人欲編製一種感應靈敏（"sensitive"）之指數，以預測一般物價變動之趨勢，而非用爲測度一般物價水準變動之精確標準者，則選取品目以少爲宜。美國紐約聯邦準備銀行所編指數，即爲其一例。該項指數最初僅包含基本物品（原料品）12項，現時則係根據22品之價格編製。潘遜氏所編之批發物價指數，包含物品僅10項，其性質與之相似。此種指數即係商情循環之物價指數，其編製目的係爲測度商情由盛而衰之過程中之趨勢者。欲使指數感應靈敏，應選取價格變動劇烈之少數物品，包含物品之項數，切不可多，惟此種指數之用途，較爲狹隘耳。蓋就一般而論，包含物品項數較多之指數，其感應性雖不免遲鈍，然此乃物價制度所固有之特性，在指數中應忠實表現者也。

討論指數內所應包含物品之項數問題，不能不與物品之性質問題

同時討論之。指數代表性之高下，雖視物價項數之多寡而定，但各項物品之性質若何，尤關重要。在選取物價時所應注意者，爲各類物價之動態，各有其特殊之異點。此各類物價，其相互之關係、其動態、其與經濟制度運行之關係、及其與商情盛衰變動之關係等，皆爲經濟學家及企業家所極端重視之事實。

物價制度之原素

吾人如欲將物價制度內各原素間之複雜關係，用圖示法表明之，自屬不可能。圖五十三僅係表示此項關係之大概，圖中所列之各原素，亦祇係比較重要者。其重要綱目爲各生產要素之價格、批發物價、零售物價及私人因營利而服務與政府服務之價格。私人服務之價格，其一部份係歸納於批發物價項下，另一部份因係爲消費者所服之勞役，故歸納於



圖五十三． 物價制度內各原素間關係之圖解

零售物價項下。在每一項內，各包含價格相聯之重要類目。在生產要素價格之第一門內所列地租及地價、俸薪及工資、貼現率及利率、債券及股票價格，在物價高漲或跌落時期，各有其特殊之動態。各種物品之價格，其動態互殊，公私團體服務之價格，其變動亦各有其特殊之點。

物價制度內各單位及各類目相互之關係，欲用圖解以關係綫一一表明之，殊不可能，即屬可能，則每一單位及每一類目皆可繪製無數關係綫，與其他單位及類目相連結，以各單位間莫不均有關聯也。此種關聯或爲直接而又重要者，或頗微渺而在事實上難以根究者。

物價制度之用平面圖表示，事實上亦欠恰當，上圖僅能表明一轉瞬間之狀態。物價前期與後期必有關聯，現時與將來亦必有關聯。例如今日之製造品價格，必受有以往之原料品及勞力費用等等之影響也。

凡能準確測度貨幣購買力或一般物價水準變動之指數，應包含物價制度中大小各類目內分別選出之代表物品。現時所發表之指數，尚未有能照此辦理者，每見一種批發物價指數，編製者自信其爲一般物價變動之測度標準，實則未必盡然也。

吾人現時所欲討論者，爲批發物價指數之編製方法。前述種種，原爲解釋此項指數之眞實範圍及其意義，並以表示其所測度者，僅係物價制度內最重要之一種原素而已。其次尚須考慮者，則爲此種原素內所包含之各各不同而又互相關聯之各類物價，以研究其與編製指數之關係。

批發物價之價格分類

批發物價指數既係根據物價樣本編製，則所選樣本必須確能代表全體物價，必須包含物價制度內各項原素中之典型價格。劃分原素之方法，必須根據各類物品價格變動之特性而定。依此原則分類，其最顯見者，爲各類之工業品。如紡織品與鋼鐵之價格，皮革與化學品之價格，係受不同之勢力所支配。各種工業所受商情盛衰之影響，時期既先後，程

度亦有深淺，故編製批發物價指數務必在各種重要工業分類中一一選取物品以代表之。倘指數中某類物品特占優勢，則指數即失去若干之代表性。例如勃蘭特斯脫指數所加於棉品、生熟牛皮、燻醃肉等品之權數，各大於其實際貿易上原有之重要程度，因之非但指數之用途有所限制，且足減損其批發物價指數之代表價值。

各種工業品價格變動不同之範圍，可就一九一三年以來農產品及家用品之批發價格指數加以比較，即可得其大略。

欲使指數具有代表性，則所取樣本不僅須就各類工業中選取適當項數之代表品，即原料與製造品價格之變動亦各不同，故兩類中均須選取物品以代表之。大抵原料之價格，對於商情盛衰之感覺更爲敏銳，其變動先於製造品，變動之程度亦較劇烈。此其原因有二：其一，原料之銷售於市場，或供消費，或供製造之用，在商情由衰而盛之時期，競爭之製造商人鑒於消費者之需要增加(或爲預測需要之增加)，致使遂不惜重價競購原料，故需要增加之壓力首及於原料市場，而因製造商之競購，遂使原料價格在他種物價受有影響以前，先已騰漲(惟此非不變之定律，在一九一九年商情轉旺期間之物價變動，即爲其例外)；反之，商情之恐慌，每始於製造品之積存過剩，故當商情衰落初現端倪之時，原料需要首先激減。簡單純粹之商業勢力，在原料市場之作用，恆較製造品市場爲自由，故吾人可由原料價格之變動，以預測其他物價變動之趨勢焉。

製造品價格較爲穩定之第二原因，由於製造品價格內所含變動比較穩定之各種成本費用如雜項開支及工資等，恆占較大之成分所致。工資、利率及地租之變動，常較物品價格遲鈍而和緩。物價之包含此種原素者，其價格必較爲穩定，故當物品之漸由原料以達製造品之階段，其價格內所包含此種穩定之原素漸多，故其變動亦漸趨穩定(註)。各種製

(註)參看密哲爾氏著 "The Making and Using of Index Numbers" 一文中所舉之例，見美國勞工統計局刊物第 284 種第44—45頁。

造品價格變動之所以分歧者，即由於此。

在上述各類中，每類又包含小類若干，其價格變動亦互不相同。原料品類中之農產、畜產、林產及鑛產之價格變動，常呈顯著之差異。農產品價格之變動，不僅受商情盛衰之影響，且隨氣候及收穫豐歉爲轉移，其變動程度雖未見和緩，然其反映商情之功用，較鑛產品爲低(註一)。畜產及林產之價格變動，反映商情之程度，似介乎農產及鑛產之間。在選取原料品作爲編製指數之物價樣本時，亦應使此四小類之物品有適當之比例(註二)。

復次，各種製造品之價格變動亦不盡一致，故各種製造品未見能合成一類而不可再分爲若干小類者。製造品之中有可供再生產或再加工之用者，其價格變動受製造商出價高下之影響，頗與原料品相似，因之變動亦較劇烈；製造品如僅供消費者之需求，則其價格受商情之影響較淺，變動自較穩定。此與前所述及之理論相吻合，即凡物價內所包含穩定之原素，如工資及雜項開支之成分愈大時，則物價變動愈乎穩是也。由此觀之，批發物價指數所根據之物價樣本，應包含製造者之物品、消費者之物品、及製造各階段所需之物品在內，自不待言(註三)。

物價分類雖尚有其他方法，惟依吾人觀察，前述分類最稱重要。欲使批發物價指數具有代表性，應在所有各類中選取物價，而各類之權數須與各該類中全部物品在貿易上所佔地位之輕重成正比例。

(註一)惟吾人不據以斷定農業生產量及農產價格兩者與一般商情毫無關聯。吾人可有充分之理由，足以徵信農業及商業狀況具有直接之因果關係，惟此關係發生之初，常呈現異方向之變動而已。

(註二)參看"The Making of Index Numbers"第47頁。

(註三)見前書第45及49頁。

美國之批發物價指數(註)

美國勞工統計局指數

美國現編之批發物價指數中，當以勞工統計局 (Bureau of Labor Statistics) 指數之聲譽爲最著。此項指數始編於一九〇二年。在一九一三年以前之指數，爲比價之未加權平均數，每項比價之基數，爲一八九〇年至一八九九年之十年平均價格；迨一九一四年，據該局出版之報告所載指數，其編製方法頗多變更，指數之基期亦同時更換。現時此項指數任何一期之數字(月指數或年指數)，皆係實價之加權綜合數，惟爲便於比較起見，此綜合數均已化爲一九一三年基期之百分比。

此項指數現時根據物價 404 項計算(每種物品或採用數項物價，其物品等級及調查價格之城邑，詳見該局之報告中。例如棉花一品，係用紐奧雷安斯州米特令及紐約州高地米特令 [Middling New Orleans 及 Middling Upland, New York]兩項價格)。每品之價格，各與其固定權數卽一九一九年之交易數量相乘。

茲以棉花一品爲例，說明其計算方法於下：

物 品	1919年平均價(每 磅)	1919年交易數量	1919年平均價×1919年交易數量
棉花，紐奧雷安斯州米特令	$.319	3,806,921,000磅	$1,214,407,799
棉花，紐約州高地米特令	.325	1,903,461,000磅	618,624,825

全部 404 項物價皆依上法計算後，將末一欄內各品價值一一相加，卽得計算時期(在本例中爲一九一九年)之指數。在發表指數時，此綜合

(註)各指數之詳細說明及其比較，具載於美國勞工統計局刊物第 284 種 "Index Numbers of Wholesale Prices in the United States and Foreign Countries". 現時該局出版之各期 "Wholesale Prices"，對於該局批發物價指數之編製內容，均有詳盡之說明。

價值各化爲百分比，而以一九一三年之綜合價值作爲100。該項指數因係採用加權綜合法計算，故可轉換爲任何一年或一月之基期，僅須將求得之百分比，化爲以該年或該月爲基期之百分比可矣。

由此觀之，該項指數實係根據一批貨物之批發價格計算者。各期貨品並無變更，惟因各品價格之變動，總價值乃隨之增減，故指數實係測度個別物價變動所給予總價值之影響之淨結果。全部物品分爲九類，各有其綜合價值，故除包含全部物品之總指數外，尙有各分類指數。物品之分類如下(註)：

農產品	建築材料
食　物	化學品及藥材
衣　着	家用品
燃料及燈光	雜　類
金屬及其製品	

聯邦準備局指數

美國聯邦準備局(The Federal Reserve Board)編有美國批發物價指數兩種。其一係根據勞工統計局所調查之物價編製者，惟爲研究物價制度內各原素起見，將物品重行分類。該指數所採用之權數及計算方法，與勞工統計局指數相同。該局編製分類指數之類名如下：

原料品
　農　產
　畜　產
　林　產
　礦　產
生產品
消費品

此種分類之用途，最初係由密哲爾氏所闡明，其意義在前數頁內已有解釋。各類指數最初在一九一三年十月聯邦準備局月刊("Federal Reserve Bulletin)內發表，該刊並載有各類之品目表。

(註) 各類所包含之品名表，見美國勞工統計局出版最近一期之"Wholesale Prices'.

聯邦準備局為欲使國際物價便於比較起見，另編有世界主要商業國之批發物價指數。此項指數之編製，乃鑒於各國所編物價指數，其編製方法既不相同，而所包含之物品種類又復各殊，礙難比較之故。現時該局按期發表之國際物價指數，計有美、加、英、法、日五國。

此項指數之計算方法，與勞工統計局指數所採用者相同，係將每品之價格與其權數相乘，得每品之價值，而後將各品之價值一一相加，得綜合價值。此綜合價值係以百分比表示，而以一九一三年之綜合價值作為100者。此項指數所以異於勞工統計局指數者，一為所包含物品項數之多寡不同，一為所採分類標準與權數之不同。

每一國際物價指數所包含之物價自90項至100項不等，按物品計，約70種。物品分類係採用兩種標準：其第一種標準，前已述及，即係將全部物品分為原料品、生產品及消費品；其第二種標準，則將全部物品另分為國產品及輸入品。此外又另編一輸出品指數，該指數內所包含各品，係由上述各類中選取之重要輸出品。前述五國之國際物價指數，皆按此六類發表分類指數及總指數；此外又發表按照各國物價折合為金物價後計算之金物價指數。

編製前項分類指數所用權數，為一九一三年之國內生產、輸入或輸出數量。求總指數時，係將國內產品之綜合價值與輸入品之綜合價值兩項相加，而後計算百分比。

此項指數已編製者，為一九一三年之指數及自一九一九年一月至最近止之月指數(法國指數係自一九二〇年一月起)(註)。

戰時工業局指數

美國所有之批發物價指數中最稱詳備者，當推戰時工業局物價股

(註)關於此項指數之詳細說明，見一九二四年聯邦準備局出版"Prices in the United States and Abroad".

(the Price Section of War Industries Board)爲研究歐戰期內物價所編之指數。此項研究所包含之時期，係自一九一三年起迄一九一八年止。研究結果詳見戰時工業局出版歐戰期內之物價（"History of Prices During the War"）一書內，其節略具載於該局出版物價叢書第一種。

各年物價計有1474品，皆由該局職員親自搜集，惟其中供編製指數之用者計1366品。物品按性質分類者計有五〇類，按工業分類者計有七類。計算指數除總指數外，並分別計算此五〇類、七大類及許多小類之指數。

此項指數皆係以百分比表示之實價加權綜合數，與勞工統計局指數相似。權數資料係根據一九一七年全年美國生產及輸入之總數量。每類各品之某月價格，先與其權數一一相乘，而後相加，卽得該類是月份之綜合價值。當合併各小類之綜合價值以求七大類指數及總指數時，每類各按其重要程度，分別加權，其方法與加權於各個物品時所用者相同。

綜合價值之折合爲百分比，係選取一九一三年七月至一九一四年六月之十二個月作爲基期，因其適當歐戰發生之前一年也，惟此基期可任意轉換而不損及指數之準確性。

除前述各政府機關所編之指數外，尙有私家編製之批發物價指數數種，其最著者爲勃蘭特斯脫及滕氏所編之指數。

勃蘭特斯脫批發物價指數

勃蘭特斯脫(Bradstreet)指數係以實價綜合數發表者。此項指數之編製，係將商業上重要物品96種之價格，一律折合爲"每磅"之價格計算，並不採用加權制。其月指數係根據每月一日之價格計算，非若勞工統計局指數之根據每月平均價計算者。又因價格無須折算爲百分比，故無基期，惟引用指數者可根據發表之綜合價格，換算爲以任何一月或任何一年爲基期之百分比。此項指數所包含之時期，係自一八九二年起以

迄於今。

前嘗論及任何指數，無論其爲有意的或無意的，均含有加權之意味。勃蘭特斯脫指數雖屬簡單指數，實則對於某數種物品加權特重。據一八九七年七月十日出版之勃蘭特斯脫物價表中，焦煤之每磅價格爲$0.0007，火酒每磅價爲$0.34，羊毛每磅價爲$0.50。在各品之每磅價格相加時，顯見無意中所加於後兩種物品之權數，遠較前一種物品爲重。潘遜氏(Warren M. Persons)嘗就一九二一年九月一日勃蘭特斯脫指數(註)製成下表，以示各類物品之權數分配。

表五十六

1921年九月一日勃蘭特斯脫指數之權數分配

物品分類	權數(百分分配)
紡織品及其原料:	
棉織品	16.7
羊毛	5.6
其他	3.3
食物:	
牛乳、雞蛋、奶油、乳酪	8.6
燻醃豬肉及牛肉、豬油	6.9
其他食物	12.4
生熟皮	13.8
建築材料	1.5
煤及焦煤:	
煤	*
焦煤	*
金屬品:	
鋼鐵	.7
其他金屬品	4.3
化學品及藥品	9.6
畜產	3.6
油	4.0
水果	3.2
麵包原料	1.0
船用品	.9
雜項	4.2

*不滿$\frac{1}{10}$%。

(註)此項指數見"Review of Economic Statistics"初卷第三號第365—366頁。

上述權數之分配，足使棉織品、生熟皮、化學品及藥品、燻醃豬牛肉等，在指數中占有重大勢力，因之指數有偏重於各該項物品之特質；且此項權數並不固定，常因價格變動而有變更。例如一九二〇年二月一日紡織品之權數占總權數之 34.46%，而在十八閱月以後，因紡織品價格之跌落遠較一般物價為劇烈，故至一九二一年九月一日紡織品之權數遂減為總權數之25.58%矣。

勃蘭特斯脫指數雖含有不合理之加權意義，亦自有其優點，倘為某種用途如供測度商情變化之用者，此為批發物價指數中最適當之一種。

滕氏批發物價指數

滕氏公司 (R. G. Dun & Co.) 之商務代理所所發表之批發物價指數，按月具載於滕氏月刊 ("Dun's Review")。該項指數以實價綜合數發表而無基期，此與勃蘭特斯脫指數相同，惟指數之編製方法，權數之分配，以及包含之時期，則兩種指數迥不相同。

滕氏指數實際祇係在一年內個人所需數項大宗物品之數量，以美金表示之費用總額而已。求月指數時，係先將指數內各品之價格，分別與其每人每年之平均消費量相乘。例如一九二一年一月一日之指數為 $198.600，此即指數內包含之各品在一年內個人所需數量，按照一九二一年一月一日之價格求得之費用總額也。指數中之價格為每月一日之價格。指數所用之物品約 300 種，詳細品名表及其實際項數均無發表，惟知所包含者皆為生活必需品，奢侈品則概在擯棄之列。權數資料雖係根據每人平均消費量，但指數非為測度生活費用之變動，而為測度一般批發物價水準變動之用。主要物品如生鐵、煤、建築材料等非供個人直接消費之品亦皆採入。此項指數起編於一九一〇年，而計算指數則追溯至一八六〇年。

就理論上言之，滕氏指數所用每項物價按照每人平均消費量加權

之方法，固屬妥善，但此法在實際編製時究竟如何運用，則不無疑問之處，蓋美國現時關於物品之消費額，除少數物品外，尚無適當可靠之資料可供採用也。且就發表之數字觀之，食物之權數幾及總權數之半，按諸事實，似失之過大。滕氏指數之每呈現特殊之變動，而常與其他一般批發物價指數之趨勢未能一致者，亦即由於食物之權數過大之所致。

費暄氏每週批發物價指數

費暄氏(Irving Fisher)嘗於一九二三年一月間編製一種批發物價指數，每週發表一次。物價指數之按週編製而比較詳備者，當以此為嚆矢。

此項指數係以百分比數字發表，而以一九一三年為基期，惟此種百分比係由實價加權綜合數求得。其權數採用固定制，係根據戶口財產普查局(The Census Bureau)所紀載之一九一九年物品成交數量。其計算公式係前述加權綜合式之第一種，即

$$\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$$

惟此處所用權數為一九一九年之數量，而非基期之數量，其辦法與勞工統計局所採用者相同。費暄氏認為就現有可供採用之資料而論，依此法計算指數所得之結果，最能與"理想"指數相接近。

此項指數所用加權方法，頗能別具心裁。指數係根據滕氏月刊所載物品200種之價格編製，其最重要之各品，係取一九一九年之實際數量為權數，惟在確定類權數時稍予變更而已；至於其他次要各品之權數，則為求計算上便利起見，採用下列各項數字：1000，100，10，1，.1，即以其中與每項物品實際數量之絕對數(或幾何比例)最為相近之一數作為該品之權數。費暄氏採取此項辦法，因鑒於次要各品採用整齊之權數，在指數上所發生可能之差誤，常不及百分之一，無足輕重之故。

費喧氏編製一九二三年每週指數時所用之基期綜合價值，實乃一九二二年十一月之價值，惟此項指數已與勞工統計局指數相接合。其接合方法如下：例如一九二二年十一月之勞工統計局指數為156，計算一九二三年每週指數時，在求得每週綜合價值與基期（一九二二年十一月）綜合價值之比率後，尚須乘以156，以換算為一九一三年之基期。現時該項指數之基期綜合價值已改為一九二三年十一月之價值，惟仍以同樣之接合手續，將指數折算為一九一三年基期之百分比。

每週之指數例在次星期一之日報內發表，故此項指數不特為現有一般批發物價指數中最切實用之一種，復以發表迅速，尤為增色不少。

潘遜氏測度商情循環之物價指數

前述各種指數皆以測度一般物價水準之變動為目的，惟實際僅以表示批發物價為限。此外尚有為特種目的編製之指數，此種特殊目的之指數之編製，頗有日見增多之傾向。在已編之指數中，潘遜氏（Warren M. Persons）與考伊爾氏（Eunice S. Coyle）合編之"商情循環之物價指數"（commodity price index of business cycles）為最值注意（註）。

此項指數編製之目的，係為測度一般商情由興旺而入衰落與由衰落而入興旺時期內之商情變動者也；換言之，此項指數係由選配各種物品之批發物價而成之商情循環之指數也。所謂商情循環，實質上原係一種物價現象。商情之變化雖尚有其他重要因素，然此種種因素必先影響於物價而後有所感覺，而商人、受雇者以及消費者所感受商情循環之影響，亦必有見於物價之變動而後發現。各項物價固皆受商情變化之影響，但以證券價格與批發物價所受影響最為直接而亦最劇烈。潘遜教授乃編製一種批發物價指數以反映及測度與商情循環有直接關係之物價變動。因其編製之目的與普通目的（"general purpose"）之指數適不相

（註）參看"Review of Economic Statistics"初卷第三號第353—369頁。

同，故所用編製方法亦隨之而異。

在決定此項指數之組織時，潘遜氏嘗將按照各工業分類之物價指數及若干物品之個別價格加以試驗，藉為選取各項物價之標準。凡其變動與一九〇三年至一九一四年間之物價循環變動及一般商情相吻合者，皆在選取之列。最初試驗之結果，發現勞工統計局及勃蘭特斯脫所編物價指數及各種非物價數列之循環變動，在上述時期內，大致相似，其波動綫之頂點，皆在一九〇七、一九一〇及一九一二年。嗣又將各指數所共同呈現之波浪式變動作為標準，用以試驗各種物價數列是否與此波動趨勢相符，而將其不相符者擯棄之。

在試驗工業分類之物價指數時，係先按照各分類編製指數，而後用上法試驗。其結果發現鑛產、鋼鐵、皮革、化學品及紡織品等類之物價指數，皆呈現上項之循環變動，且其趨勢與商情循環亦屬相似。至於農產、食品製造業以及石、粘土、玻璃製造業各類之物價指數，並無上項循環變動，其趨勢與一般物價及一般商情之變動，亦不相侔。

其次用同一方法，為個別物價變動之試驗。凡各種物價並未呈現循環變動，或其變動無伸縮性者，因其不合指數之編製目的，皆所不取。依此法試驗後，發現在未呈現循環變動之分類指數中，亦含有若干物價，其個別之變動呈現上述循環變動者；又發現個別物品之價格變動，比較分類指數更為劇烈而敏銳，於是決定商情循環之物價指數，按照個別物價，而不按分類指數編製。

最後決定選入指數之物品僅十種，為棉子油、焦煤、鋅塊、生鐵、鐵條、燻醃豬肉、生牛皮、印花布、粗布、絨綫。此數項物品之價格變動甚為劇烈，且於商情變動之感應性亦極敏銳，其所代表之各工業，亦皆佔重要之地位。

此十種物品自一八九〇年以後之價格資料，皆已蒐集，而編成商情循環之指數。所有每月及每年之價格，俱以百分比表示，而以一八九〇

至一八九九年爲基期（基期現已改爲一九一九年）。其指數爲此項比價之未加權幾何平均數。

其他批發物價指數

紐約聯邦準備銀行現編有包含基本物品20種之批發物價指數，其性質略與潘遜氏指數相似，惟其編製目的則未明言爲測度商情循環之變動。實際上此種指數係一種測度基本物品市價變動而感應敏銳之指數。此項指數刊載於紐約聯邦準備銀行出版之月刊"Monthly Review"，係以百分比表示，而以一九一三年爲基期者。指數起編於一九一三年，以迄於今。

安納利斯脫（"The Annalist"）之批發食物價格指數，係按週發表者。此項指數所根據者，爲主要食物25種之價格。其計算方法係先將此項價格折算爲百分比，而以一八九〇至一八九九年之平均價作爲基價，而後求其未加權之算術平均數。此項指數之編製，僅爲測度家用食物帳所應包含各種食物之批發價格之變動而已。

此外紐約奇勃孫公司（Thomas Gibson）亦編有批發食物價格指數一種，在該公司出版之每週商情報告內發表。該項指數係根據主要食物22種之價格編製，至於所採何種加權制度，初未明言，僅謂該指數係採用滕氏指數之方法而已。該項指數爲比價之加權平均數，計算每項比價所用之基價，爲一八九〇至一八九九年之平均價，惟所得結果尙須與一常數相乘，使此項指數與滕氏指數相啣接，蓋其目的欲以接續滕氏指數也。

各指數編製方法之差異

吾人如欲將前述各種批發物價指數作長時期之比較（註），其價格

（註）密哲爾氏在其所著"The Making and Using of Index Numbers"一文中，嘗將三種主要之美國批發物價指數作長時期之比較。原文具載於美國勞工統計局叢書第284種卷一第94—112頁。

變動大體雖極相似，但細加審察，要不無變動互異之處。例如一種指數所表示物價水準之變動，本月較上月增高，而同期之另一種指數並無變動，甚有表示跌落者。此種矛盾情形，在價格變動劇烈時，如一九一六年至一九二一年之時期內，尤所常見。其變動不同之原因，在前此討論指數編製方法時，嘗有述及。任何矛盾之變動，必不外乎下述之原因：

a. 物品項數多寡之不同。

b. 採入指數之物品在性質上之不同。

c. 物價蒐集方法之不同（或為合同價，或為市價；或採用每月一日之價格，或為每月之平均價）。

d. 權數之不同。

e. 平均方法之不同。

概略：批發物價指數

茲將各種指數之主要特質及其用途，就其相異各點，概括論述之。

1. 美國現編各種物價指數中，其編製最稱完善者，首推勞工統計局所編之一般批發物價指數。該指數所測度者，非物價變動之平均比率，而為購置一定數量之物品所需費用之增減。
2. 在一九一三年至一九一八年間所編之指數中，以戰時工業局所編之一般批發物價指數最為優良。
3. 滕氏指數測度批發物價之變動，尚屬完美，惜於食物製品加權過重，指數所包含之各品及其實際權數又未予發表，致指數之價值不免減損。
4. 勃蘭特斯脫指數於測度批發物價之變動，亦頗優良，惟於原料品過於重視，頗足減損指數之代表價值。其所含權數輕重之不合理，亦為一大弱點，然因其偏重於原料品之故，以之測驗商情盛

衰，反較勞工統計局及滕氏指數為勝任愉快。

5.潘遜氏指數非為測度一般物價之變動者。其編製之初，即擬以測驗商情循環為目的，故測度此種現象自頗切當。

6.紐約聯邦準備銀行所編20種物品之指數，對於測度基本原料之價格變動，為其特長。

7.聯邦準備局為比較國際物價所編之指數，可使各國之物價變動易於比較，故頗切實用。其所編各國之指數對於測度各該國之一般物價變動，亦頗精確。

8.費暄氏之每週批發物價指數，理解準確，編製精審，為用甚著。該項指數發表迅速，尤足增進其效用。

9.安納利斯脫所編之批發食物價格指數，用途較狹。其編製方法因係採用比價之未加權平均數，故所得結果在價值方面每致發生懷疑，且其所採基期為一八九〇年至一九九九年，亦失之過遠。

其他物價指數

應用指數以測度物價之變動，並不以批發物價為限，亦有用以測度他種物價之變動者。此處略將其他指數之性質及其編製方法說明之。

零售物價指數

美國勞工統計局按期發表零售食物價格指數一種。其編製方法與該局用以計算批發物價指數者，大致相同，所稍異者乃由於採用之資料性質不同之故。

編製指數所用食物43種(註)之實際零售價格，例於每月十五日向零售商店直接調查。調查區域包括美國主要城市51處，間有各處習慣不盡相同者，在可能範圍以內，務使各處價格能相互比較。計算指數時，係先

(註)此項指數自一九二一年一月起，係根據食物43種編製，以前係根據食物22項。

求每品價格之總數，再以報告該品價格之商店數除此總數，即得每品之全國平均價格；次將各品平均價，按照該局家計調查所得一九一八年之每家平均消費數量，分別加權，以求實價加權綜合數，其法與編製批發物價指數所用者相同；最後將零售物價之綜合數，以一九一三年為基期，換算為百分比。此項百分比之指數在勞工月刊 "Monthly Labor Review"內按期發表。

同時此51城市亦分別編製零售食物價格指數，其編製手續與全國指數相同。全國之平均物價及各城市之物價，均在零售物價刊物內按期發表。此種辦法頗足增進該局工作之實際價值。

編製指數以測度批發物價變動之種種困難，前嘗一再論及，而編製零售物價指數，其困難則猶有過之，蓋除編製批發物價指數所發生之一切理論問題仍須一一解決外，在實際方面，如蒐集適當之權數資料，調查準確之物價數字，以及判別價格之能否相互比較等，在在發生困難。復以物品之未能標準化，及商業習慣與各地風尚之不同，尤足使各地價格不易比較。因此種種原因，現時編行之零售物價指數，尚無一能受人信仰如優良之批發物價指數者。

生活費指數

若謂編製零售物價指數不易，則編製生活費指數之困難，當更倍蓰，蓋欲將家計調查所得每家消費之各項目，如食物價格、房租、零售衣着價格、燃料及燈光費用、零售傢具價格以及其他雜項物品之價格，一一求其平均數，以供編製此項價格指數之用，每發生統計上之種種困難，而須設法解決之；且除關於平均數及加權方法發生學理上之問題外，他如關於蒐集準確詳備之數字，亦必發生更困難之實際問題。在現時所編生活費指數尚未能確切表現其可靠程度以前，應用此項指數者，務宜特別注意。

現時已編之生活費指數有二：一爲勞工統計局所編，一爲屬於雇主公會之全國實業會 (National Conference Industrial Board) 所編，前者在勞工月刊 ("Monthly Labor Review") 發表，後者在該會出版之刊物內發表。上項指數內所有家庭消費之主要項目，俱按照家用賬簿所示之各品比重程度，分別加權，其綜合價值亦皆以一九一三年爲基期，折算爲比價後發表者(全國實業會指數係以一九一四年七月爲基期)。

農產品之價格指數及其購買力指數

美國農業部編有指數一組，頗切實用。其中一種係爲測度主要農產品農場價格之按月變動者，另一種係爲測度畜產品農場價格之變動者，兩者復合編爲第三種指數，以測度一切農產品之平均價格變動。上述三種指數俱以百分比表示，而以一九一三年爲基期。

該部爲欲使農產品價格指數變動之意義更爲明顯起見，另計算第四種指數，以表示農產品購買力之變動。以貨幣表示之農產物價指數，因貨幣之購買力常在變動之中，其本身無甚意義，故吾人需要一種測度農民所購各種物品價格變動之指數，以供計算農產品購買力指數之用。農業部編製此種指數，係將勞工統計局之批發物價指數，剔除農產品及食物後，改編而成。茲以一九二三年一月份指數爲例，以說明農產品購買力指數之計算方法。按是月農產品價格指數爲 116 (一九一三年=100)，剔除農產品及食物後之批發物價指數爲 170。其意謂一九二三年一月農民於出售其產品時所獲得之代價，較一九一三年增 16%，同時一般批發物價較一九一三年增 70%，故農民之購買力，以一九一三年爲基期，應爲$\frac{116}{170}\times 100$，即68.2%也。

茲將本節所述各種指數之一九一三年各月數字，列表於下：

表五十七

物價指數

1913年各月農產品價格及其購買力之指數

(一九一三年＝100)

月份	農場價格			躉售物價*	農產品之購買力§
	農產品(每月十五日)	畜產品(每月十五日)	農產品及畜產品合計		
一　月	126	106	116	170	68
二　月	130	107	118	172	69
三　月	134	106	120	175	69
四　月	139	107	123	176	70
五　月	140	105	123	172	71
六　月	139	100	120	168	71
七　月	136	102	119	165	72
八　月	136	102	119	163	73
九　月	138	109	123	164	75
十　月	139	103	121	161	75
十一月	137	97	117	160	73
十二月	137	94	116	158	73

*農產品及食物除外。　§以購買農產品及食物以外之各種物品表示者。

編製上項購買力指數時，嘗有一重要之假設，即謂農民所購物品之價格變動，可以校正後之勞工統計局批發物價指數測度之是也。此項假設不無商榷之餘地，蓋一則農民購買物品，通常並不躉批購入；再則該指數所包含者，亦不盡爲農民所購買之物品也。雖然，在此尙無比較適當之指數可資採用以前，原係權宜之計，惟其中所含可能之差誤，在引用此項指數時應加注意耳。

前項指數自一九一三年起以迄於今，皆按月編製，在美國農業部出版之農產品及其市場("Crops and Markets")內發表。

貨幣工資及眞實工資之指數

近年工業糾紛，風起雲湧，工資問題遂爲一般人士所密切注意，此雖由於時代關係，然此問題實爲一永久性之重要問題。工資指數已由政府機關、雇主公會、工會以及私人編製者，爲數頗夥。各指數之優劣，此處不擬評論，本文所欲陳述者，爲關於編製及解釋此種指數方面之重要數點而已。

編製貨幣工資指數所涉及之問題，其最要者不過爲平均物價之問題，惟此處之物價，係指勞力而非商品耳。平均時所發生之困難，在於紀錄勞力價格方法之不同。吾人計算工資指數時，應按照每小時工資率計算乎？抑按照每週應得之工資總數（即依普通每小時工資率，核算全週應得之工資數）計算乎？抑按照每週實得工資數計算乎？如按每小時工資率計算，則每週實際工作之時間、失業、半失業、加工以及額外獎金等之變動，均未能顧及。如按每週應得工資數計算，則每週工作小時數之增減，雖可反映於其指數，但仍有其他缺點。如按每週實得工資數計算，倘能包括忙閒季節，自屬最切實用，惜此種資料殊不易獲得，因在事實上諸多窒礙也。紐約州勞工局（The New York State Department of Labor），嘗按照每週實得工資數編有此種指數，其工資資料乃得自該州重要各工廠之報告。此項指數係以貨幣單位表示之平均工資數，惟可任擇一基期，換算爲百分比表示之指數。

除工資資料外，尚有關於加權及平均方法之問題，此處可無須討論。其更重要者，則有"代表性"（"representativeness"）之一問題，此與編製批發物價指數時所發生選取物品之種類問題相似。如吾人僅取熟練工人之工資，求其平均工資，則其結果不能用作代表普通工資率之指數；如以工頭及薪資甚厚之技術工人之工資，選入普通工資指數，則須各按其應有之權數，分別加權，且須明白解釋其結果。故遇此種情形，當仿照批發物價指數編製分類指數之辦法，分編各級工人之分類指數，並將分類指數用適當權數加權後，彙編全體工人工資之加權平均指數。吾人如將工資分類研究，編成分類指數，則對工資變動之性質，必較各類彙合編成之總指數易於明瞭，此點正與研究物價變動時之情形相似。

編製工資指數因有上述種種不同之點，故現時已編之各種工資指數，其所用編製之方法迥不相同。除其中嘗將編製方法及所用資料之性質作詳盡之說明，可供審擇者外，吾人對於各該種指數所表示工資變動

之經過情形，尙未可貿然置信也。

貨幣工資指數所表示者，必在生活費用無變動時，始有意義。蓋當物價與生活費用增高時，工資指數所表示工資之增高，或屬虛妄；反之，當物價與生活費用跌落時，工資指數所表示工資之跌落，亦不正確，故吾人如欲測度工人實際享受之豐嗇，應顧及生活費用之增減變動，將貨幣工資指數折算爲眞實工資指數，始有準確之意義。

茲就下表說明眞實工資指數之編製手續。表內所列之工資指數，係根據紐約州工人475,000人之平均每週工資額計算，而以一九一四年十二月爲基期者。生活費指數係美國勞工統計局所編包括全國主要城市三十二處之指數，亦以一九一四年爲100(註)。

表五十八

1914—1923年紐約州眞實工資指數

(1914年十二月＝100)

月 份	紐約工人每週工資之百分比	生活費指數	眞實工資指數(卽貨幣工資之購買力)
1914年十二月	100	100	100
1915年十二月	107	102	105
1916年十二月	123	115	107
1917年十二月	140	138	101
1918年十二月	185	169	109
1919年十二月	209	193	108
1920年十二月	225	195	115
1921年十二月	198	169	117
1922年十二月	210	165	127
1923年十二月	222	168	132

求眞實工資指數時，係將工資指數除以同期之生活費指數卽得。如僅以工資指數測度工人之收入狀況，所得結果之不正確，自屬顯見。校正以後之工資指數所表示眞實工資數雖仍有增加，但遠不若貨幣工資數增加之劇烈。雖然，上例中各項數字，至多亦僅能近乎準確，蓋用以計算眞實工資之兩種指數，在編製方面均有發生重大差誤之可能，故由此

(註)本例中紐約州之工資指數，如能以紐約一州之生活費指數校正之，自更適當，惟依此計算所得之結果，相差恐不致過大。

求得之眞實工資指數，亦僅能視為測度紐約工人眞實工資變動之一種粗略指數而已。

參考書

Bowley, A. L. "Elements of Statistics" (196—213).

Chaddock, R. E. "Principles and Methods of Statistics" (Chap. X).

Fisher, Irving. "Revision of the Weekly Index Number". Journal of the American Statistical Association, Sept., 1924.

Fisher, Irving. "The Making of Index Numbers".

Flux, A. W. "The Measurement of Price Changes". Journal of the Royal Statistical Society, March, 1921(167—215).

Kelley, Truman L. "Statistical Method" (331—347).

Mitchell, W. C. "The Making and Using of Index Numbers". Part I, Bulletin 284, U. S. Bureau of Labor Statistics, Oct., 1921.

Persons, Warren M. and Coyle, Eunice. "A Commodity Price Index of Business Cycles". Review of Economic Statistics, Prel. Vol. III(353—369).

Persons, Warren M. "Fisher's Formula for Index Numbers". Review of Economic Statistics, Prel. Vol. III(103—113).

Walsh, C. M. "The Measurement of General Exchange Value".

Walsh, C. M. "The Problem of Estimation".

Young, Allyn A. "The Measurement of Changes in the General Price Level". Quarterly Journal of Economics, Aug., 1921(557—573).

Young, Allyn A. "Index Numbers' (in Rietz, H. L., "Handbook of Mathematical Statistics", 181—194).

第七章　時間數列之分析：長期趨勢之測度

以上各章均係討論次數數列以及關於整理及表述此項數列之各問題，吾人現時所欲討論者，為因時間而變動之資料以及分析此項變動之方法。在經濟統計中，時間數列極為重要，蓋各種經濟及商業資料，如銀行票據交換額、鋼鐵產量、銷貨數量等等大都含有時間性之變動也。時間數列在其他統計方面，則不甚重要，故分析時間數列之方法發展亦較緩，迨近數年統計方法之應用於經濟界方面漸見普遍後，始漸發展耳。

企業家在日常內部管理方面，以及經濟學家分析一般經濟狀況之變動時，每涉及時間數列之問題。例如銷貨、進貨、損益等內部管理方面之問題，股票價格、利率、企業倒閉等一般經濟問題，莫不與時間變動有關。在分析此種數列時，吾人應斷定其變動之性質及程度，並須將週期變動(periodic fluctuations)及意外變動 (accidental fluctuations) 之因素，予以隔離 (isolate)，以便研究。在企業方面，公司經理所欲知者，為銷貨數量之增減，其變動之時間與原因，以及銷貨數量與生產量之比較；在經濟方面，經濟學家所欲知者，為各種物價變動之趨勢，以及一般物價水準升降之詳細情形。倘欲決定商業政策或預測未來之經濟情勢，必先研究已往之變動與趨勢，比較互有關聯之各種時間數列之變動以為根據；他如商業循環之研究，亦必應用此種分析方法而後可。吾人現時所欲討論者，即為分析時間數列之方法也。

時間數列之初步整理

時間數列之初步整理，常較次數數列為簡易，蓋後者須將原來資料編為次數分配，而前者不論其為原始資料或次級資料，往往可就原有形式着手分析，而無須另加編制也。惟有數點應加注意者，茲分述之。

資料所載數字之日期，其意義若何，應辨別清楚，並分別註明之。例如月數字之資料，或爲每月一日之數字（如勃蘭特斯脫之批發物價指數），或爲每月之平均數（如勞工統計局之物價指數），或爲每月之總數（如棉花之消費數量）。此種數字又或爲逐月之累積數，表示某年內至各該月終之總數，如煤產量之資料往往以累積數表示之。又數字如爲每月或每年之平均數，則須查明求得此項平均數之方法。

時間數列中各時期之資料，其性質務必相同，庶可比較。如資料性質不一致(not homogeneous)，則分析所得之結果必生錯誤而不適用。吾人日常所見發表之統計資料，其性質每不一致，其例不勝枚舉。如政府機關或商業團體所發表商品之生產量或消費量，因係彙集各機關各社團之報告編製而成，故其日期、單位及性質往往不盡一致。如物價數列常因物品之單位或品質稍有參差，或調查地點之不同，以致各時期之物價難作比較。如戶口調查每因分類方法之有變更，各期調查所得之數字亦有礙難比較之慨。又如曩時進出口貿易之數字，其編製方法前後各期頗不一致，故比較時發生錯誤，常所難免。又如銷貨員之銷貨數量，或因派往地點有變更而受重大影響。又如美國鋼鐵公司所發表用以表示其債務性質之所謂'未辦噸量'之數字，其意義亦時有變更：如在一九二二年七月此項噸量係指實際定貨約期交貨之數量，而在一年以前則僅指載於定貨合同，並未限期交貨，而隨時可以取消定貨之總數量。以上所舉爲時間數列中所常見之錯誤，故在分析數列之前，必先辨別資料是否準確可靠，性質是否一致，以定取捨焉。

時間數列之圖示

研究時間數列之初步工作，須將資料作圖，俾便觀察，並作進一步之分析，蓋時間數列之特性及其趨勢，每可由圖而得其梗概。數列之資料有可繪於普通算術尺度之圖上者，有可繪於半對數或雙對數尺度之

圖上者。用對數尺度作圖之意義及其功用，前已詳言之矣。尺度之選擇常視資料之性質及研究之目的而定：倘研究之目的在於表示銷貨數額、物價或生鐵產量等數量變動之絕對值，或比較各數列間數量相差之絕對值，則以用普通算術尺度為宜；如研究之目的在於表示數量變動之百分率，或比較相對變動之差異，則以用半對數尺度為宜矣。大體言之 如吾人熟諳半對數尺度之圖解，自以採用該種圖形較為適當。經濟資料之屬於時間數列者，如以時間按照算術尺度，以另一變量按照對數尺度製圖，尤可得明確之意義，而欲比較兩種時間數列之關係，亦以此種半對數圖最為清晰。

時間數列之研究工作，有時在製成統計圖以後已告完成，蓋所有變動之大勢以及季節變動與其他週期性之變動，或各期變動及其趨勢之比較，皆可由圖中測知其大概。惟吾人須知由圖觀察，並不能獲精確之印象，即所得之比較亦僅屬試驗性質；雖然，圖中所示之變動趨勢及各種關係，已較原來數字易於領會，倘吾人僅欲粗知其大概，則此種圖解雖未能予吾人以確切之印象，亦已盡其能事。但若吾人認為獲得此項印象尚未滿足，則非作更精確之測度與更詳密之分析不可矣。以下所述，即為此種測度及分析之方法。

時間數列所受各種勢力之影響

分析時間數列之普通目的，厥為隔離（isolate）數列中所受一種或多種勢力之影響。吾人倘欲研究一數列已往之變動情形以預測其未來之變動，或與另一種或多種數列互為比較，則須將數列中所受各種勢力之影響予以隔離，以便研究。惟事實上欲隔離各個勢力必難絕對準確，有時即相當準確之程度亦不易得，但若吾人所用之數列包含長時期之資料，則其所受各種勢力之影響，並非不可測度而得一相當準確之結果者。現時經濟資料之分析，對於此種測度方法尚未能充分利用之。

然則時間數列係受何種勢力之影響乎？按各種勢力之存在與否每隨數列之性質而定，某種勢力或爲某種數列所特有，惟大致可分爲數種，下文當分別討論之。

長期趨勢

在經濟統計中，各種時間數列大抵含有顯著之趨勢（trend）。此種趨勢或有一定之方向，或雖方向無定，而增減之比率大致相同者，亦有因新因素之隨時發生，其變動方向與比率時有表現驟然之變更者。例如物品之產量，商店之銷貨額，在一長時期內，其數字之變動，常有逐漸向上之傾向。此外，人口之變動，礦物產量之消長，以及汽車登記輛數之增減等等，皆其例也。在某種經濟現象中，此種傾向亦有屬於負方向，而呈現逐漸向下之傾向者，如前半世紀中美國利率之變動即爲其一例。故吾人所謂長期趨勢（secular trend）（即長時期內之傾向）實包含逐漸向上或逐漸向下之變動傾向也。

分析時間數列所得某一時間之長期趨勢值（trend value）即作爲該時期之常態數值（normal value）。常態數值者，乃謂一數列在剔除長期趨勢以外所受各種意外及複雜勢力之影響後，而在正常狀態下，呈逐漸向上或逐漸向下之傾向時所應得之數值也。常態數值之觀念實爲分析經濟資料之基本要點，因求得各期之常態數值後，即可以此項數值作爲基數，以斷定其他各種勢力之影響。故斯尼竇氏（Carl Snyder）嘗有'吾人處置經濟資料，欲求準確而迅捷，除此法以外，實無其他方法可資採用'之語，亦具見此項概念重要之一斑矣。

於此應注意所謂長期趨勢者，乃統計數列所表現平滑而有規則之長期變動也，故若絕對數值或比率之增減變動，倏忽無定，則無長期趨勢之可言。一種數列有因新因素之發生，或舊因素之消滅，亦可釀成偶然的意外變動者，但若將數列所包含之時間分爲數階段而分別求其變

動之趨勢，則與長期趨勢之定義所謂在長期內逐漸變動之趨勢相抵觸；故一數列變動之因素，如常有極端變更，則該數列可認爲無長期趨勢之存在。

以上所論，非謂任何時間數列皆有一向上或向下之顯著趨勢。例如氣壓之變化，雖經過一相當時間，而常沿一固定之水平綫升降，絕無向上或向下之顯著趨勢者也。

週期變動

吾人觀察時間數列之圖解時，雖可由曲線之向上或向下傾向，而發現其長期趨勢，但常因其他種種變動之混入，不能僅憑觀察以作精確之斷定。此種種變動或有規則，或無規則，或劇烈，或和緩，或簡單，或複雜，故時間數列內各時期之數值，常爲彙合長期趨勢及其他種種勢力所得之淨結果。此種種勢力之存在，足使各時期之數值不循長期趨勢綫而變動，以致該數列所表現之長期趨勢爲之晦而不彰，如無此項勢力以擾亂長期趨勢，則該數列必可呈現緩漸與正常之變動。

此種種勢力大致可分爲數種。經濟統計之資料，如採用月數字，大抵含有季節變動 (seasonal variations)。例如商品之消費量及生產量、利率、銀行票據交換額、鐵路運貨數量以及其他各種資料，莫不含有週期性之季節變動。此種變動週期性顯著，而週期之長短亦固定不變，常以十二月爲一週。此外週期性不甚顯著，而其變動極有規則者，則爲循環變動(cyclical fluctuations)，係因受經濟與商業狀況循環變動之影響而發生者。例如物價、工資、工業生產量、股票交易額以及其他屬於各個企業組織活動方面之各種統計數列，莫不受商業興旺與商業衰落循環變動之影響。循環變動週期之長短或不盡同，但其起伏變動常整齊有序，可供研究。

意外變動

時間數列除受上述各種勢力而表現較有規則之變動外，復受另一種勢力而含有不規則與意外之變動（random fluctuations）。如三藩市之地震、戰爭、水災、火災以及其他各種意外事件，皆足引起不規則而比較和緩之變動。此種變動雖與季節變動及循環變動有別，但其足使數列中一變量之數値離開其常態値，則一也。由是知各時期之實際數値與常態値相差之數，即爲所受此種種擾亂因素彙合而成之淨結果。

時間數列之分析，即在可能範圍內，隔離上述各種勢力之影響也。吾人之問題，或在研究其中單獨一種因素，或在分別研究各種因素。如吾人所用之資料爲每年數字，則季節變動之因素，自不存在。以下討論分析方法時所採用之資料均係年數字，故季節問題可暫置勿論。

長期趨勢之測度

下表所列之數字(註)可作爲研究上述各種問題之資料。表中所列數値，爲求計算手續便利起見，故以十萬萬美金元爲單位。

表五十九

1860—1923年紐約市各銀行票據交換額

（單位：十萬萬美金元）

1860	7.2	1876	$ 21.6	1892	$ 36.7	1908	$ 79.3
1861	5.9	1877	23.3	1893	31.2	1909	103.6
1862	6.9	1878	19.9	1894	24.4	1910	97.3
1863	14.9	1879	29.2	1895	29.9	1911	92.4
1864	24.1	1880	38.6	1896	28.8	1912	100.7
1865	26.0	1881	49.4	1897	33.4	1913	94.6
1866	28.7	1882	46.9	1898	42.0	1914	83.0
1867	28.7	1883	37.4	1899	60.8	1915	110.6
1868	28.5	1884	31.0	1900	52.7	1916	159.6
1869	37.4	1885	28.2	1901	79.4	1917	177.4
1870	27.8	1886	33.7	1902	76.3	1918	178.6
1871	29.3	1887	33.4	1903	66.0	1919	235.8
1872	33.8	1888	31.1	1904	68.6	1920	243.2
1873	35.5	1889	35.9	1905	93.8	1921	194.4
1874	22.9	1890	37.4	1906	104.7	1922	217.9
1875	25.1	1891	33.7	1907	87.2	1923	214.0

（註）此項數列應先平減價値（deflate），而後分析。其手續下文將有說明。

表中任何一年之數字，例如一九一二年之交換額 $100.7（單位十萬萬美金元），係彙合長期趨勢、週期變動及意外變動等種種勢力而成之淨結果，此點前已言之矣。吾人當前之問題，卽爲測定該數列所受第一種勢力卽長期趨勢之影響。

圖五十四所示，卽一八六〇年至一九二三年紐約市各銀行票據交換額之圖解。此圖表現一顯著之長期趨勢，同時又顯示脫離長期趨勢之其他各種較有規則之變動。測度長期趨勢之方法有數種：一曰移動平均數法 (moving averages)，其法係根據各期數值求移動平均數，以消除

圖五十四. 1860—1923年紐約市各銀行票據交換額及其移動平均數

長期趨勢以外其他各種勢力之影響，而使求得之各期數值表示逐漸增高或逐漸減低之長期變動；二曰配合曲綫法，此法係假定時間與另一變量有函數關係之存在，而以另一變量作爲時間之函數，應用數學方法以求與資料最相配稱之曲綫，卽以此綫代表其長期趨勢；三曰隨手畫法，其法係憑觀察之經驗，就圖中表示原來資料之綫，詳察其變動之趨勢，隨手畫一平滑之趨勢綫。三者之中，後者最無標準，而含有試驗與揣測之性質。更有以一種統計數列作爲另一種數列之基綫或趨勢綫之方法

者，惟此法必須在後種數列之資料性質一致時，始可採用之。

移動平均數法

用移動平均數以決定長期趨勢之方法，係在全時期內，求若干年（或若干月或若干週）之移動平均數，作爲中間一年之趨勢值 (trend value) 或常態値 (normal value)。下表所列爲一九〇〇年至一九二三年紐約市各銀行票據交換額之三年、五年、七年及九年移動平均數。

表六十

1900—1923年紐約市各銀行票據交換額之三年、五年、七年及九年之移動平均數

（單位：十萬萬美金元）

年份	原來資料	三年移動平均數	五年移動平均數	七年移動平均數	九年移動平均數
1900	$ 52.7				
1901	79.4	$ 69.5			
1902	76.3	73.9	$ 68.6		
1903	66.0	70.3	76.8	$ 77.4	
1904	68.6	76.1	81.9	82.3	$ 78.7
1905	93.8	89.0	84.1	82.3	84.3
1906	104.7	95.2	86.7	86.2	86.3
1907	87.2	90.4	93.7	90.6	88.1
1908	79.3	90.0	94.4	94.0	92.0
1909	103.6	93.4	92.0	95.0	94.8
1910	97.3	97.8	94.7	93.6	93.6
1911	92.4	96.8	97.7	93.0	94.3
1912	100.7	95.9	93.6	97.5	102.3
1913	94.6	92.8	96.3	105.5	113.2
1914	83.0	96.1	109.7	116.9	121.6
1915	110.6	117.7	125.0	129.2	137.0
1916	159.6	149.2	141.8	148.5	153.7
1917	177.4	171.9	172.4	169.7	164.1
1918	178.6	197.3	198.9	185.7	177.8
1919	235.8	219.2	205.9	201.0	192.4
1920	243.2	224.5	214.0	208.8	
1921	194.4	218.5	221.1		
1922	217.9	208.8			
1923	214.0				

表內一九〇四年之三年移動平均數爲自一九〇三年至一九〇五年三年間之平均數，其五年之移動平均數爲自一九〇二至一九〇六年五年間之平均數，其餘各年之移動平均數亦用同法計算。每一平均數均集

中於時間之中點，而以平均數代表該點之常態數值。移動平均數之年數如爲奇數，則集中手續 (centering process)較易，即以移動平均數作爲中間一年之常態值；惟所定年數不必定爲奇數，如年數爲偶數時，則於求得各年移動平均數後，再求其兩年移動平均數，而使此項平均數集中於一年之中間。表中五年與九年之移動平均數及原來資料均繪於圖五十四內。

移動平均數之特質

設有一數列，其所包含之長期趨勢成一水平綫，或斜度不變而成一傾斜之直綫，而其所包含之循環變動，其週期之長短 (length of cycle) 有定，每一週期內起伏變動之幅度(amplitude)亦有一定時，則用循環時期之年數或其倍數以求移動平均數，可得一直綫，恰與原來之長期趨勢綫相合。在同一情形之下，如計算移動平均數之年數，較週期之年數或多或少時，則移動平均數所得之綫並非直綫，而爲表現另一種循環變動之曲綫。此新循環週期之長短與原來週期相同，惟其變動之幅度較小，且該綫所示循環之最高點及其最低點之位置亦未必與原來循環相吻合。大體言之，移動平均數所用之時期愈長，則由此所得新循環之幅度亦愈小(註)。

表六十一所列，爲任意假設之數字，用以引證前述各點者。表內(1)行內每五個數字成一週期之資料，其長期趨勢爲一水平綫。

由表觀之，(2)(3)兩行內所列之移動平均數均成一水平綫，而已將循環變動完全消除；但當平均數之項數與循環週期之長短或其倍數不相等時，則循環變動仍不能消除，如(4)(5)兩行所示。

如一數列之長期趨勢綫爲一傾斜之直綫，而其循環變動有一固定之週期，及一固定之幅度，則用移動平均數所得之結果正與上例相同。

(註)幅度之減小並無一定規則，惟循環變動則仍存在。

表六十一

說明移動平均數之應用

(1) 循環資料	(2) 五項移動平均數	(3) 十項移動平均數（集中數字）	(4) 三項移動平均數	(5) 八項移動平均數（集中數字）
2				
6			$5\frac{1}{3}$	
8	$6\frac{1}{5}$		8	
10	$6\frac{1}{5}$		$7\frac{2}{3}$	
5	$6\frac{1}{5}$		$5\frac{2}{3}$	$6\frac{3}{8}$
2	$6\frac{1}{5}$	$6\frac{1}{5}$	$4\frac{1}{3}$	$6\frac{13}{16}$
6	$6\frac{1}{5}$	$6\frac{1}{5}$	$5\frac{1}{3}$	$6\frac{3}{8}$
8	$6\frac{1}{5}$	$6\frac{1}{5}$	8	$5\frac{6}{8}$
10	$6\frac{1}{5}$	$6\frac{1}{5}$	$7\frac{2}{3}$	$5\frac{11}{16}$
5	$6\frac{1}{5}$	$6\frac{1}{5}$	$5\frac{2}{3}$	$6\frac{3}{8}$
2	$6\frac{1}{5}$	$6\frac{1}{5}$	$4\frac{1}{3}$	$6\frac{13}{16}$
6	$6\frac{1}{5}$	$6\frac{1}{5}$	$5\frac{1}{3}$	$6\frac{3}{8}$
8	$6\frac{1}{5}$	$6\frac{1}{5}$	8	$5\frac{6}{8}$
10	$6\frac{1}{5}$	$6\frac{1}{5}$	$7\frac{2}{3}$	$5\frac{11}{16}$
5	$6\frac{1}{5}$	$6\frac{1}{5}$	$5\frac{2}{3}$	$6\frac{3}{8}$
2	$6\frac{1}{5}$		$4\frac{1}{3}$	$6\frac{13}{16}$
6	$6\frac{1}{5}$		$5\frac{1}{3}$	
8	$6\frac{1}{5}$		8	
10			$7\frac{2}{3}$	
5				

〔(3)(5)兩行所列平均數均係集中數字，卽在求得十項或八項之移動平均數後，再求兩項移動平均數集中而得者。〕

表六十二所列(1)行內之數字，係由上表內之資料以一定增量(constant increment) +3 遞加而得者。此與將上項循環變動之數字加於斜度爲+3之直綫上時所得之結果相同。

表六十二內數字如採用與循環週期相等或與其倍數相等之項數求移動平均數，則所得之長期趨勢綫，爲一不含循環變動，斜度爲+3之直綫；但若所用項數與週期或其倍數不相等時，則循環變動依然存在，週期與前相同，惟其幅度較小。此可於(4)(5)兩行之數字見之。

上述兩例所假設之數字，係較爲單純而屬於理想之資料，蓋謂循環變動之週期及幅度皆有一定也。倘此簡單情形並不存在，則準確可信之

移動平均數即難揣測，而其解釋亦較困難。當循環變動之週期無定時，欲決定移動平均數之年數，殊爲困難，惟就大體而論，此時似以選取與各週期平均年數相等或較多之年數，更爲相宜；否則如所選年數較少，所得移動平均數仍必含有顯著之循環變動，故選較長之時期者，所以消除循環變動之勢力也。

表六十二

移動平均數應用於直綫趨勢之數列

(1) 循環資料	(2) 五項移動平均數	(3) 十項移動平均數 (集中數字)	(4) 三項移動平均數	(5) 八項移動平均數 (集中數字)
2				
9			$8\frac{1}{3}$	
14	$12\frac{1}{5}$		14	
19	$15\frac{1}{5}$		$16\frac{2}{3}$	
17	$18\frac{1}{5}$		$17\frac{2}{3}$	$18\frac{3}{8}$
17	$21\frac{1}{5}$	$21\frac{1}{5}$	$19\frac{1}{3}$	$21\frac{13}{16}$
24	$24\frac{1}{5}$	$24\frac{1}{5}$	$23\frac{1}{3}$	$24\frac{3}{8}$
29	$27\frac{1}{5}$	$27\frac{1}{5}$	29	$26\frac{6}{8}$
34	$30\frac{1}{5}$	$30\frac{1}{5}$	$31\frac{2}{3}$	$29\frac{11}{16}$
32	$33\frac{1}{5}$	$33\frac{1}{5}$	$32\frac{2}{3}$	$33\frac{3}{8}$
32	$36\frac{1}{5}$	$36\frac{1}{5}$	$34\frac{1}{3}$	$36\frac{13}{16}$
39	$39\frac{1}{5}$	$39\frac{1}{5}$	$38\frac{1}{3}$	$39\frac{3}{8}$
44	$42\frac{1}{5}$	$42\frac{1}{5}$	44	$41\frac{6}{8}$
49	$45\frac{1}{5}$	$45\frac{1}{5}$	$46\frac{2}{3}$	$44\frac{11}{16}$
47	$48\frac{1}{5}$	$48\frac{1}{5}$	$47\frac{2}{3}$	$48\frac{3}{8}$
47	$51\frac{1}{5}$		$49\frac{1}{3}$	$51\frac{13}{16}$
54	$54\frac{1}{5}$		$53\frac{1}{3}$	
59	$57\frac{1}{5}$		59	
64			$61\frac{2}{3}$	
62				

〔(3)(5)兩行所列平均數均係集中數字，即在求得十項或八項之移動平均數後，再求兩項移動平均數集中而得者〕.

又設循環變動之週期有定，而各週期內變動之幅度不等時，如採用與週期相等之年數計算移動平均數，則所得之長期趨勢綫，仍含有較小之循環變動。此種次級循環 (secondary cycles) 變動之幅度，在移動平均數之年數與週期相等或爲其倍數時爲最小。又若循環變動之週期無

定，而各週期內變動之幅度又不相同時，則移動平均數之年數，以選取與各週期之平均年數或其倍數相等者爲宜，因由此所得之移動平均數用以表示長期趨勢，最爲恰當也。

長期趨勢之形態，原不盡爲直綫，如爲曲綫，則問題更見複雜。倘長期趨勢綫呈向上內凹狀 (concave upward)，則由移動平均法求得之數值必大於實際趨勢值；倘長期趨勢綫呈向上外凸狀 (convex upward)，則由移動平均法求得之數值必小於實際趨勢值。

表六十三

移動平均數應用於非直綫趨勢之數列(遞加增加率)

(1) x	(2) x^2	(3) 循環資料	(4) (2)+(3)	(5) 五項移動平均數	(6) 實際趨勢值 $(x^2+6.2)$
0	0	2	2		
1	1	6	7		
2	4	8	12	12.2	10.2
3	9	10	19	17.2	15.2
4	16	5	21	24.2	22.2
5	25	2	27	33.2	31.2
6	36	6	42	44.2	42.2
7	49	8	57	57.2	55.2
8	64	10	74	72.2	70.2
9	81	5	86	89.2	87.2
10	100	2	102	108.3	106.2
11	121	6	127	129.2	127.2
12	144	8	152	152.2	150.2
13	169	10	179	177.2	175.2
14	196	5	201	204.2	202.2
15	225	2	227	233.2	231.2
16	256	6	262	264.2	262.2
17	289	8	297	297.2	295.2
18	324	10	334		
19	361	5	366		

此種情形可舉例說明之。表六十三所載之數字，其趨勢綫呈向上內凹狀，此綫係以週期有定幅度相同之循環變動數字，加於向上內凹狀之長期趨勢綫而成。其長期趨勢係依遞加之增加率增加 (increasing at a constantly increasing rate)，故成向上內凹狀。如求得之移動平均數已將循環變動之影響完全消除，則所得之各項數值應爲循環資料五

項平均數($6\frac{1}{5}$)與函數關係$y=x^2$內y數值兩項之和。

然由表六十三觀之，移動平均數之數值，常大於實際趨勢值。凡用移動平均數以求此種數列之長期趨勢綫，皆不免發生此項錯誤。

表六十四

移動平均數應用於非直綫趨勢之數列(遞減增加率)

(1) x	(2) $\sqrt{x}$	(3) 循環資料	(4) (2)+(3)	(5) 五項移動平均數	(6) 實際趨勢值 ($\sqrt{x}+6.2$)
0	0	2	2.00		
1	1.00	6	7.00		
2	1.41	8	9.41	7.428	7.61
3	1.73	10	11.73	7.876	7.93
4	2.00	5	7.00	8.166	8.20
5	2.24	2	4.24	8.414	8.44
6	2.45	6	8.45	8.634	8.65
7	2.65	8	10.65	8.834	8.85
8	2.83	10	12.83	9.018	9.03
9	3.00	5	8.00	9.192	9.20
10	3.16	2	5.16	9.354	9.36
11	3.32	6	9.32	9.510	9.52
12	3.46	8	11.46	9.658	9.66
13	3.61	10	13.61	9.800	9.81
14	3.74	5	8.74	9.936	9.94
15	3.87	2	5.87	10.068	10.07
16	4.00	6	10.00	10.194	10.20
17	4.12	8	12.12	10.318	10.32
18	4.24	10	14.24		
19	4.36	5	9.36		

表六十四所載之數字，其趨勢綫呈向上外凸狀，此[illegible]同一循環變動之數字，加於向上外凸狀之長期趨勢綫而成。其長期趨勢係依遞減之增加率增加 (increasing at a constantly decreasing rate)，故成向上外凸狀。此時如採用優良方法以求長期趨勢，而能完全消除循環變動之影響，則所得之各項數值應爲循環資料五項平均數($6\frac{1}{5}$)與函數關係$y=\sqrt{x}$內y數值兩項之和。

但由表六十四觀之，移動平均數之數值常小於實際趨勢值。此兩項數值之差異，在x數值較小時，尤爲顯著，蓋x數值小，則受遞減增加率之影響較大也。

吾人於此可得如下之結論，即在趨勢綫爲直綫時，移動平均數之年數至少應等於週期之年數，而在數列之變動無甚規則時，尤以採用週期年數之倍數爲宜，蓋移動平均數之時期愈長，則平均數之數值亦愈穩定；但若趨勢綫之形式非直綫，而呈向上內凹狀或向上外凸狀，則移動平均數之時期愈長，發生之差誤亦愈大，故用移動平均數法以測度此種數列之長期趨勢時，所取時期以短爲佳，而應等於各個週期之平均年數。

以上所論，均係根據假設之資料，討論移動平均數應用之方法，然實際上一個數列之變動常受種種複雜情形之影響，故欲決定移動平均數之年數，並不如前述之簡單。如一數列循環變動之週期與其幅度均不固定時，移動平均數之時期自以稍長者爲佳，但一般數列之趨勢往往爲非直綫形，則平均數之時期又以短者爲宜。在實際應用上所當考慮者，移動平均數之數值絕不能包括全時期，平均數之時期愈短，則遺漏之年數亦愈少。吾人必先參酌資料之性質，並注意前述各點，然後始可擇定平均數之時期。

移動平均數法之應用，並不限於測度長期趨勢，而亦可用以消除不規則變動之影響，此時移動平均數之時期以選擇小於循環週期之年數者爲宜。

吾人現時仍以紐約市各銀行票據交換額爲例，以討論移動平均數之年數問題。由圖五十四內觀之，兩種移動平均數所繪成之曲綫頗多相異之處。五年移動平均數之綫與原來資料之綫最爲接近；九年移動平均數綫雖最平滑，但與原來資料相差亦最鉅，尤以一八九三年至一八九八年及一九一一年至一九一五年兩時期爲尤甚，因在此兩時期內票據交換額變動最爲劇烈也。除此兩時期外，似以九年移動平均數最能代表其長期趨勢。

在確定移動平均數之年數時，吾人可參酌該時期內商情變遷之狀況以爲準繩。紐約市各銀行票據交換額爲一感應敏銳之一般商情指標，

對於投機交易及實業狀況之變遷，均具有特別靈敏之感應性，故在該時期內，無論大小之商情循環變動，一一反映於該數列。在一八七〇年至一九二〇年之以往之半世紀中，吾人已知所經歷商業循環變動之週期數，故可用此實例以為根據，以斷定以上所舉兩種移動平均數中，究以何種最能代表該時期內銀行票據交換額之長期趨勢，而用以測度循環變動之程度。惟依此研究，乃就已知之結論，作步驟相反之推考，要為平時工作所罕見耳。

假定以商情之由衰落以迄恢復常態之前一年為每一週期之起點，則在該全時期內，可得以下之各週期(註)：

1870—1878年	1904—1908年
1879—1885年	1908—1911年
1886—1897年	1911—1914年
1898—1904年	1915— 年

吾人現時可計算數列內各年數值對於其同一年份移動平均數之差，以試求循環變動之週期數。由三年移動平均數求得循環變動之週期數，自嫌過多，蓋該項變動實乃種種意外而又細微之變動，不能視為循環變動。由數列各項對於其七年及九年移動平均數之差，可得下列循環變動之各週期：

由七年移動平均數所得循環變動之各週期	由九年移動平均數所得循環變動之各週期
1871—1878年	1871—1878年
1879—1884年	1879—1887年
1885—1888年	1888—1897年
1888—1897年	1898—1900年
1898—1900年	1900—1903年
1900—1903年	1904—1907年
1904—1907年	1908—1914年
1908—1911年	1915— 年
1911—1914年	
1915—1918年	

(註)此項循環變動之週期，係根據 W. F. Ogburm 氏暨 Dorothy Thomas 氏所編之指數而得。參看一九二二年九月出版 "Quarterly Publications of the American Statistical Association" 第324—340頁。

上列兩種移動平均數所得循環變動之週期，可與前此所列之週期相比較，其中不同之處，一部分自由於所用資料迥異之故，惟有若干不同之點頗堪注意。例如由七年移動平均數所得之週期數較其餘兩種各多三週期。如七年移動平均數內一八八五年至一八八八年及一九一五年至一九一八年之兩週期均為其餘兩種所無，後者在此兩期內僅見微小之波動。一八九八年至一九〇〇年之週期，雖同見於兩種移動平均數，而為用作比較根據之指標所無。一九一一年至一九一四年之週期雖同見於指標，而不見於九年移動平均數。歐戰期間票據交換額劇增，足使九年移動平均數受其影響，以致最後一週期因平均作用而為之消滅矣。

由上推知，欲由票據交換額之資料，求其長期趨勢，則移動平均數之時期，不可少於七年。九年移動平均數較為合用，惟在期末交換額之增加率劇變時則為例外。七年之時期似覺過短，因其移動平均數過於偏重微小變動也。

就一般而論，移動平均數有伸縮性 (flexibility)，此為其優點。用數學方法配合曲綫以代表長期趨勢，常須將全時期分作數階段，而各配一曲綫。此在情形發生變化，或數列之增加率或減少率劇變時，用數學方法配合長期趨勢綫，則不能不將時期分為數段。此時倘用移動平均數，則因其具有伸縮性，而能迎合變遷之情況，不令趨勢猝然中斷，常較計算費時之數學趨勢綫尤為適用，此即移動平均數法之優點也。

除前述之簡單移動平均數外，又有所謂加權移動平均數者。加權方法，乃在計算每一移動平均數時，將各年數值分別加權，其數值愈在中間者，其所加權數亦愈大。此種加權移動平均數稱為遞進平均數 (progressive mean)，係以兩項展開式中各項之係數用作移動平均數各項之權數。如上例採用五項移動平均數時，其由第一項至第五項之權數，應依次為 1, 4, 6, 4, 1。遞進平均數雖可用於某種數列之分析，但欲用以消除時間數列中之週期變動或不規則變動，似不適宜耳。

配合數學曲綫法

時間數列之資料中，頗有可用數學曲綫以代表其長期趨勢者。倘數列中各數值之增加(或減少)為一固定之增量(increments)(或減量 decrements)，則用一直綫表示其長期趨勢，頗為恰當；倘各數值之增加為一固定之百分率，如依照複利法則，本利和之金額係按固定之利率而增加，則可用一數學曲綫以表示其長期趨勢。經濟統計之資料，頗有遵循一定法則而增減者，故此種資料之分析與解釋以及預測未來之趨勢，常可求得一數學方程式以表示該項法則，作為根據。任何資料雖常有不依一定法則，而在長期趨勢綫之上下變動者，但在此種變動之中，仍有規則存在，故方程式之價值，並不因此而減損也。

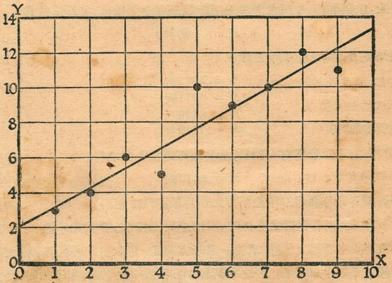
數學曲綫及移動平均數雖皆用以代表長期趨勢，但有一根本不同之點，應予注意。移動平均數並不假定資料之必遵循一定法則，而全以實際資料為根據，如長期趨勢有所變更，則移動平均數亦必隨此新趨勢而變動，故此平均數能適應情況之變遷，而伸縮自如，其目的僅在表示數列變動之大體趨向。數學曲綫配合之目的，雖亦僅以表示其大體趨勢為限，但其意義稍有不同。此法係假定數列之各種變動，在資料之表面上，無論其為意外變動或他種變動，均遵循一定之法則。其所假定者固屬不能證實之經驗法則 (empirical law)，但經數學方程式表示後，則成一固定劃一之趨勢綫。故欲使此種假定確實而有效，則在此法則所存在之時期內，一切經濟狀況務必相當穩定，而無劇烈變化足以影響研究之時間數列而後可。設吾人已求得一黃金產量趨勢綫之方程式，而適在此時採礦方法根本變更，則黃金產量之趨勢亦必立受鉅大之影響，已得之方程式即不能適用，蓋在採礦方法變更前後之兩時期內，因資料性質並不一致，自不當用同一之曲綫方程式，以表示兩時期內之長期趨勢也。

數學曲綫之形式甚多，故在實際工作時，欲測定長期趨勢，首須擇

定一相當之曲綫形式。此雖屬初步工作，而實亦最難者，因選擇時須憑個人之判斷力，絕無一定標準可循。關於此點，當俟各種主要曲綫之特性及其配合方法討論後，再行論述。茲先就第二章內所載各種曲綫或其類似之形式中擇定一種，以研究其配合方法。

直綫之配合：最小平方法

吾人將資料作圖後，如見圖中各點之趨勢能用直綫表示者，則配合之工作僅在決定$y=a+bx$方程式中之a,b兩常數。所求a,b之數值，應以由此求得之直綫最能適合資料之趨勢者為限。茲可舉一簡單之實例，以說明求a,b之方法。圖五十五中共有九點，即1,3; 2,4; 3,6; 4,5; 5,10; 6,9; 7,10; 8,12; 9,11是也。試就圖中各點配合一直綫。



圖五十五．　在九點上配合直綫之圖解

此處若用觀察法亦可決定a,b之近似值。如手執一綫之兩端，在各點左右移動，使此綫愈能適合各點之趨勢愈佳，即可由此求得一最適當

之位置，然後決定此綫之斜度及其y截部，則該綫方程式亦可確定矣。惟此法漫無標準，直綫之位置隨人而異，因之所得方程式亦無一定，而在理論上，圖中各點僅有一綫最為適合。此綫常數之值，可用最小平方法求之。

最小平方法之定理，此處不必詳細討論，可簡述於下：設有一數量經測量數次後，得若干不同之觀察值(observation value)，而欲求此量最可能之數值 (the most probable value)。吾人可證明該數值應為剩餘數(residual) 平方之和為最小時之值 (剩餘數係指某一數值與各觀察值之差)，在理論上卽為各觀察值之算術平均數也。如有多人測量一固定之距離，則各人所量得之結果未必一致，但最可能之數值，應為各人所量尺度之算術平均數。計算算術平均數之方法包含下列步驟，述之於此，以便解釋。今所求者為上項距離最可能之數值，其形式應為

$$M=(\text{一常數})$$

如測量所得M之值共有三項如下：

$$M=5672\text{呎}$$
$$M=5671\text{呎}$$
$$M=5676\text{呎}$$

各項相加，得 $3M=17019\text{呎}$

此式中之未知數僅有一個M，故其值可逕由上式求得

$$M=5673\text{呎}$$

此值卽為與各觀察值相差平方之和為最小時之值。

當測度兩變量間之關係時，其問題與上相似。吾人此時之目的，在於根據圖中各點之位置，求得一直綫方程式，以表示此種關係，惟由此所得之直綫不一而足，因之在其各個方程式中之常數，亦不相同，換言之，卽各點並不在同一直綫之上也。各點既不在同一直綫上，然則以何綫為最合於各點分佈之情勢，亦卽各方程式中之常數以何者為最可能之值乎？此其答案與測量一單獨數量時正復相同。直綫方程式之常數應

使該直綫繪於圖中時，其與各點相距平方之和爲最小數。假定由觀察所得之每對變量均係表示兩變量眞實關係之近似程度，吾人欲由此確定一最可能之相關程度，而以直綫表示之，則該綫正爲用最小平方法求得之綫，蓋該綫與各點相距平方之和較其他任何直綫爲最小也(註)。

吾人於此有x與y之九對數值，將各對數值代入直綫之普遍方程式$y=a+bx$中，可得下列九個觀察方程式(observation equations)：

$$3=a+1b$$
$$4=a+2b$$
$$6=a+3b$$
$$5=a+4b$$
$$10=a+5b$$
$$9=a+6b$$
$$10=a+7b$$
$$12=a+8b$$
$$11=a+9b$$

上列九式中，任何兩式均可應用聯立方程式之解法，得a,b之數值，但每對方程式所得a,b之數值，不能適合其他各個方程式，故須聯合此九個觀察方程式以求兩個常態方程式 (normal equations)，更由此兩式以求a,b之最可能值，代入直綫之普遍方程式中，即得所求之直綫。以每個觀察方程式中a之係數乘該式中各項，由此求得之九個方程式一一相加，即得第一個常態方程式。此處各個觀察方程式中之係數均爲1，相乘後各式不變，故即可用此原來方程式相加。次以每個觀察方程式中b之係數乘該式中各項，由此求得之九個方程式一一相加，即得第二個常態方程式。此處各個觀察方程式中，第一式b之係數爲1，故以1乘該式中各項；第二式b之係數爲2，故以2乘該式中各項；餘類推。茲將本例中求常態方程式之計算手續舉示於下：

(註)關於最小平方法之詳細討論及其計算方面校核之方法，可參考附錄A.

表六十五

由觀察方程式求常態方程式之方法

$3 = a + 1b$	$3 = 1a + 1b$
$4 = a + 2b$	$8 = 2a + 4b$
$6 = a + 3b$	$18 = 3a + 9b$
$5 = a + 4b$	$20 = 4a + 16b$
$10 = a + 5b$	$50 = 5a + 25b$
$9 = a + 6b$	$54 = 6a + 36b$
$10 = a + 7b$	$70 = 7a + 49b$
$12 = a + 8b$	$96 = 8a + 64b$
$11 = a + 9b$	$99 = 9a + 81b$
$70 = 9a + 45b$	$418 = 45a + 285b$

由此求得之兩個常態方程式爲

$$70 = 9a + 45b$$

$$418 = 45a + 285b$$

至此僅須求兩式中a, b之數值。以5乘第一式各項，由第二式中減去之，即可消除a而得b之數值爲$\frac{68}{60}$即1.133；再將此值代入兩式中任何一式，即得a之數值爲2.111，故本例中所求之直綫方程式爲

$$y = 2.111 + 1.133x$$

實際計算時，無須將所乘之方程式一一紀錄而後相加之，僅須將x及y之相當數值代入下列兩個典型方程式 (type equations)即可(註)。

$$\Sigma(y) = na + b\Sigma(x)$$
$$\Sigma(xy) = a\Sigma(x) + b\Sigma(x^2)$$

式中所用各符號之意義，解釋於下：

$\Sigma(y)$: y各數值之和。

$\Sigma(x)$: x各數值之和。

$\Sigma(xy)$: x與y之每對數值相乘之積之和。

(註)常態方程式作成之法，見附錄A。

$\Sigma(x^2)$: x各數值平方之和。

n: x與y數值相配之對數，亦即圖上之點數。

應用下列表式，則計算手續可較省便。

表六十六

配合直綫所需之各數值

x	y	xy	x^2	
1	3	3	1	$n=9$
2	4	8	4	$\Sigma(x)=45$
3	6	18	9	$\Sigma(y)=70$
4	5	20	16	$\Sigma(x^2)=285$
5	10	50	25	$\Sigma(xy)=418$
6	9	54	36	
7	10	70	49	
8	12	96	64	
9	11	99	81	
45	70	418	285	

將表內所列五個數值代入前述典型方程式中，即得所求之兩個常態方程式。此兩式與前所求得者完全相同。

直綫方程式既已求得，即可將x各數值代入式中，而分別求y各數值，再將所得各值與y之各觀察值(即原來數值)一一比較，而求其差。下表所示，即求差數所用之行列也。

表六十七

一變量之觀察值與其計算值之比較(註)

	y (觀察值)	y (計算值)	d	d^2	xd
1	3	$3.2\frac{4}{9}$	$-.2\frac{4}{9}$	.0597	$-.2\frac{4}{9}$
2	4	$4.3\frac{7}{9}$	$-.3\frac{7}{9}$	.1427	$-.7\frac{5}{9}$
3	6	$5.5\frac{1}{9}$	$+.4\frac{8}{9}$	.2390	$+1.4\frac{6}{9}$
4	5	$6.6\frac{4}{9}$	$-1.6\frac{4}{9}$	2.7041	$-6.5\frac{7}{9}$
5	10	$7.7\frac{7}{9}$	$+2.2\frac{2}{9}$	4.9381	$+11.1\frac{1}{9}$
6	9	$8.9\frac{1}{9}$	$+.0\frac{8}{9}$	.0079	$+.5\frac{3}{9}$
7	10	$10.0\frac{4}{9}$	$-.0\frac{4}{9}$	.0020	$-.3\frac{1}{9}$
8	12	$11.1\frac{7}{9}$	$+.8\frac{2}{9}$	.6760	$+6.5\frac{7}{9}$
9	11	$12.3\frac{1}{9}$	$-1.3\frac{1}{9}$	1.7190	-11.8
合計			0.0	10.4885	0.0

(註)d及xd行中小數點後仍用帶分數者，欲使各項差數之和適等於0，無稍差異也。

由上表可測知圖五十五中各點與直綫縱距離之和爲0,以x各數値乘各項差數而後相加,所得之和亦爲0,故計算之準確與否,可據此校核之。各項差數平方之和(10.4885)爲最小之數值。如將直綫方程式中a或b之數値予以變更,則所得直綫與各點縱距離之平方之和,必大於10.4885。

直綫之配合:特例

在任何情形下,吾人均可應用聯立方程式之解法,由兩個常態方程式求得a,b之最可能數值,惟在特殊情形下,計算手續尙可簡略。此種情形爲經濟資料中所常見。設x各數値爲連續數字,如時間數列中之連續年份然,則在圖中可用x之中數爲原點。若x之項數爲奇數,則此原點卽爲中間一項。此時$\Sigma(x)$之値爲0,而兩個常態方程式變爲

$$\Sigma(y)=na$$

$$\Sigma(xy)=b\Sigma(x^2)$$

例如一數列包含之時期,係自一九〇〇年起至一九二〇年爲止,吾人可取一九一〇年爲原點,如是則一九〇九年x之數値爲-1,一九一一年x之數値爲$+1$,餘類推,依此求a,b之數値,手續上自可便利不少。如數列之年數爲偶數,亦可應用此項簡便方法,將中間兩年之中點作爲原點,而以半年爲單位,計算各年x之數値,則$\Sigma(x)$亦必等於0。

又如x各數値爲由0點起之連續正數時,$\Sigma(x)$及$\Sigma(x^2)$之數値亦易斷定。自1起至第n項止之連續正數之和,可用公式$\frac{n(n+1)}{2}$求之。例如自1至10各數之和爲$\frac{10(10+1)}{2}=55$,卽常態方程式中$\Sigma(x)$之値也。自1起至第n項止之連續正數平方之和,可用公式$\frac{2n^3+3n^2+n}{6}$求之。例如自1至5各數平方之和爲$\frac{250+75+5}{6}=55$,卽常態方程式中$\Sigma(x^2)$之値也。故常態方程式亦可改爲

$$\Sigma(y)=na+b\left(\frac{n(n+1)}{2}\right)$$

$$\Sigma(xy)=a\left(\frac{n(n+1)}{2}\right)+b\left(\frac{2n^3+3n^2+n}{6}\right)$$

用此式求a,b之數值，往往較原來方程式爲簡易。凡時間數列均可應用此式，以第一年x之數值編列爲1，將以後各年x之數值順次編列連續數字，可節省計算手續。

遞升冪級數式曲綫之配合

以上所論，均限於直綫之長期趨勢，但時間數列之長期趨勢亦常有不宜用直綫表示，或雖用直綫表示，而須將時間分作數段，分別配合直綫者。此種分段辦法，如因全時間內情勢驟變，前後兩期資料之性質不一致時，未嘗不可引用；但如全時期內情況大致相同，資料性質亦尚一致時，則此種分段辦法，實與長期趨勢之意義相背。數列之不宜以直綫配合者，吾人常可配合遞升冪級數式之曲綫，以表示其長期趨勢。配合此種曲綫之計算手續，可簡述於下。

遞升冪級數方程式之普遍形式爲 $y=a+bx+cx^2+dx^3\cdots\cdots$。此種方程式之曲綫雖未必爲拋物綫式之曲綫，但通常稱爲拋物綫。式中x之最高冪數爲二次冪時，稱爲二次拋物綫；爲三次冪時，稱爲三次拋物綫；餘類推。普通應用時，式中之冪數以不過二次或三次爲宜。如x之最高冪數爲二次冪時，則式中含有三個未知數，而須用三個常態方程式以求未知數之數值矣。

求常態方程式之方法與配合直綫時所用者相同，即以每個觀察方程式中第一個未知數之係數乘該式中各項，由此求得之各個方程式一一相加，即得第一個常態方程式；以每個觀察方程式中第二個及第三個未知數之係數分別乘各該式中各項，由此求得兩組方程式各別相加，即得第二個及第三個常態方程式。由此三個常態方程式即可求得a,b,c三

個未知數之值。此三值為 a,b,c 三個常數最可能之數值。三個常態方程式之普遍形式如下：

$$\Sigma(y)=na+b\Sigma(x)+c\Sigma(x^2)$$

$$\Sigma(xy)=a\Sigma(x)+b\Sigma(x^2)+c\Sigma(x^3)$$

$$\Sigma(x^2y)=a\Sigma(x^2)+b\Sigma(x^3)+c\Sigma(x^4)$$

茲以1,2; 2,6; 3,7; 4,8; 5,10; 6,11; 7,11; 8,10; 9,9九點為例，說明配合二次拋物綫之計算手續。吾人無論在配合曲綫，或從事其他繁重之計算工作時，均須有條不紊，按步進行，使每一步驟與前後各步互相關聯，而在可用校核法時，又應儘量採用，蓋計算時縱極端審慎，錯誤或仍所難免也。計算之步驟應排列成表，使每步之計算，及其所得之結果，俱可詳列於表內。

本例所用之資料及其計算步驟，列如下表。

表六十八

配合二次拋物綫所需各項數值之計算

x	y	xy	x^2	x^2y	
1	2	2	1	2	$n=9$
2	6	12	4	24	$\Sigma(x)=45$
3	7	21	9	63	$\Sigma(x^2)=285$
4	8	32	16	128	$\Sigma(x^3)=2,025$
5	10	50	25	250	$\Sigma(x^4)=15,333$
6	11	66	36	396	$\Sigma(y)=74$
7	11	77	49	539	$\Sigma(xy)=421$
8	10	80	64	640	$\Sigma(x^2y)=2,771$
9	9	81	81	729	
45	74	421	285	2,771	

本例中 x 各數值為自 1 起之連續整數，故 $\Sigma(x)$，$\Sigma(x^2)$，$\Sigma(x^3)$ 及 $\Sigma(x^4)$ 之數值又可自預先製就之表中求得之(註)。

(註)參考 Pearson 氏所編"Tables for Statisticians and Biometricians"第28表，一九一四年劍橋大學出版。

將表內各數值代入前述三式中，可得下列常態方程式：

$$74=9a+45b+285c$$

$$421=45a+285b+2,025c$$

$$2,771=285a+2,025b+15,333c$$

由此聯立方程式可得a,b,c三個常數之值如下：

$$a=-.929$$

$$b=+3.523$$

$$c=-.267$$

所求之二次拋物綫方程式爲

$$y=-.929+3.523x-.267x^2$$

此綫及原來之九點均繪於圖五十六內。

倘x各數值爲連續數字，如本例中所用者，亦可將x之中位數作爲原點，以節省計算工作，因此時 $\Sigma(x)$ 及 $\Sigma(x^3)$俱等於0，而常態方程式可變爲：

$$\Sigma(y)=na+c\Sigma(x^2)$$

$$\Sigma(xy)=b\Sigma(x^2)$$

$$\Sigma(x^2y)=a\Sigma(x^2)+c\Sigma(x^4)$$

在配合 $y=a+bx+cx^2+dx^3$ 形式之三次拋物綫時，吾人須斷定式中四個常數之值，而需用四個常態方程式。此四式之形式如下：

$$\Sigma(y)=na+b\Sigma(x)+c\Sigma(x^2)+d\Sigma(x^3)$$

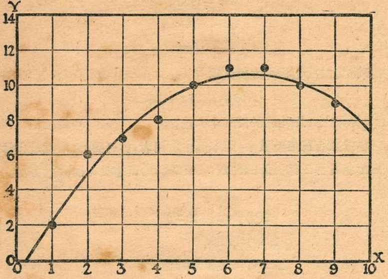
$$\Sigma(xy)=a\Sigma(x)+b\Sigma(x^2)+c\Sigma(x^3)+d\Sigma(x^4)$$

$$\Sigma(x^2y)=a\Sigma(x^2)+b\Sigma(x^3)+c\Sigma(x^4)+d\Sigma(x^5)$$

$$\Sigma(x^3y)=a\Sigma(x^3)+b\Sigma(x^4)+c\Sigma(x^5)+d\Sigma(x^6)$$

欲求四個或四個以上之常數數值，不僅所需計算工作頗爲繁重，且此種曲綫在理論上是否值得採用以表示數列之長期趨勢，亦成問題，蓋若方程式中常數增多，雖可使資料之轉變在曲綫上一一表現，然此種曲

綫斷非表示長期趨勢之綫也(註)。在經濟資料中，果有一眞實趨勢之存在，其趨勢綫無論其爲直綫或爲曲綫，應爲一簡單匀滑之綫，而實際數值在所配曲綫之兩旁雖不免有微小之起伏變動，但極端之起伏變動應所罕見也。如綫之兩旁遇有極端之變動，是必由於數列之本質有所變更，決不能適用一綫以代表其長期趨勢，而須將時期截爲兩段或數段，



圖五十六. 在九點上配合二次拋物綫之圖解

(註)據 Steinmetz 氏之意見，在用上述遞升冪級數式曲綫以代表經驗曲綫時，須合下列條件，方可採用。

1. 如 a, b, c（連續係數）之數值逐項減低，且減低之速度甚大，以致在觀察範圍內，x 高次冪之數迅速變小，而成爲式中之次要項目者。

2. 如連續係數係依隨一定法則，而自成一收斂級數 (convergent series)，如指數之級數，或三角函數之級數等等者。

3. 如各係數中，僅有數個係數之數值較大，其餘均甚小，但欲偏重於數個較大之係數，不必計及較小之係數者(參看Steinmetz氏著"Engineering Mathematics"第214—215頁)。

分別配以趨勢綫。史丹曼茲(Steinmetz)氏嘗謂"在觀察範圍之內，自然界情形不變，始可用一單獨方程式以代表經驗曲綫(empirical curve)"，此言雖指自然科學方面之資料而言，然亦可適用於經濟資料也。

測定商業倒閉家數長期趨勢之實例

前數節所述爲用簡單之資料以配合直綫或拋物綫之手續，在討論其他曲綫以前，請舉實例以說明之。茲就一八九七年至一九二一年間美國商業倒閉家數之資料，分別配合直綫、二次拋物綫及三次拋物綫。吾人用同一資料配合三種長期趨勢綫後，可比較其結果，而注意其不同之點，以供研究。

爲計算便利起見，可用中間一年，卽一九〇九年，作爲原點。常態方程式中所需各數值可由表六十九得之。表內x值代表時間，y值代表商業倒閉之家數。

表六十九

1897—1921年美國商業倒閉家數

配合長期趨勢綫所需各項數值之計算

(1)	(2)	(3)	(4)	(5)	(6)
年份	x	y (商業倒閉家數)	xy	x^2y	x^3y
1897	—12	13,083	—156,996	1,883,952	—22,607,424
1898	—11	11,615	—127,765	1,405,415	—15,459,565
1899	—10	9,642	—96,420	964,200	—9,642,000
1900	—9	9,912	—89,208	802,872	—7,225,848
1901	—8	10,648	—85,184	681,472	—5,451,776
1902	—7	9,973	—69,811	488,677	—3,420,739
1903	—6	9,775	—58,650	351,900	—2,111,400
1904	—5	10,417	—52,085	260,425	—1,302,125
1905	—4	9,967	—39,868	159,472	—637,888
1906	—3	9,385	—28,155	84,465	—253,395
1907	—2	10,274	—20,548	41,096	—82,192
1908	—1	14,066	—14,066	14,066	—14,066

表六十九(續)

年份	x	y	xy	x^2y	x^3y
1909	0	11,872	…………	…………	…………
1910	1	11,588	11,588	11,588	11,588
1911	2	12,679	25,358	50,716	101,432
1912	3	13,832	41,496	124,488	373,464
1913	4	14,553	58,212	232,848	931,392
1914	5	16,780	83,900	419,500	2,097,500
1915	6	19,035	114,210	685,260	4,111,560
1916	7	16,498	115,486	808,402	5,658,814
1917	8	13,073	104,584	836,672	6,693,376
1918	9	9,331	83,979	755,811	6,802,299
1919	10	5,515	55,150	551,500	5,515,000
1920	11	8,463	93,093	1,024,023	11,264,253
1921	12	19,982	239,784	2,877,408	34,528,896
合計	0	301,958	+188,084	15,516,228	+9,881,156

$n=25$

$\Sigma(x)=0$

$\Sigma(y)=301,958$

$\Sigma(xy)=188,084$

$\Sigma(x^2)=1,300$

$\Sigma(x^2y)=15,516,228$

$\Sigma(x^3)=0$

$\Sigma(x^3y)=9,881,156$

$\Sigma(x^4)=121,420$

$\Sigma(x^5)=0$

$\Sigma(x^6)=13,471,900$

配合直綫之演算

x之原點既在時期之中點,則求a,b兩常數所用之直綫常態方程式爲

$$\Sigma(y)=na$$
$$\Sigma(xy)=b\Sigma(x^2)$$

將表六十九中求得之數值代入上式,則

$$301,958=25a$$
$$188,084=1,300b$$

由此兩式可得a,b之數值爲

$$a=12,078$$
$$b=144.7$$

故直綫方程式爲

$$y=12,078+144.7x$$

配合二次拋物綫之演算

配合二次拋物綫時所用之常態方程式有三。其形式如下：

$$\Sigma(y)=na+c\Sigma(x^2)$$

$$\Sigma(xy)=b\Sigma(x^2)$$

$$\Sigma(x^2y)=a\Sigma(x^2)+c\Sigma(x^4)$$

將相當之數值代入式中，則

$$301,958=25a+1,300c$$

$$188,084=1,300b$$

$$15,516,228=1,300a+121,420c$$

解式中之常數，可得

$$a=+12,258$$

$$b=+144.7$$

$$c=-3.45$$

故二次拋物綫之方程式爲：

$$y=12,258+144.7x-3.45x^2$$

配合三次拋物綫之演算

配合三次拋物綫時，須用四個常態方程式以求四個常數之數值。其形式如下：

$$\Sigma(y)=na+c\Sigma(x^2)$$

$$\Sigma(xy)=b\Sigma(x^2)+d\Sigma(x^4)$$

$$\Sigma(x^2y)=a\Sigma(x^2)+c\Sigma(x^4)$$

$$\Sigma(x^3y)=b\Sigma(x^4)+d\Sigma(x^6)$$

由此所得之四個聯立方程式為

$$301,958 = 25a + 1,300c$$

$$188,084 = 1,300b + 121,420d$$

$$15,516,228 = 1,300a + 121,420c$$

$$9,881,156 = 121,420b + 13,471,900d$$

解四式中常數之數值，可得

$$a = +12,258$$

$$b = +482$$

$$c = -3.45$$

$$d = -3.61$$

故三次拋物綫之方程式為

$$y = 12,258 + 482x - 3.45x^2 - 3.61x^3$$

表 七 十

1897—1921年美國商業倒閉家數

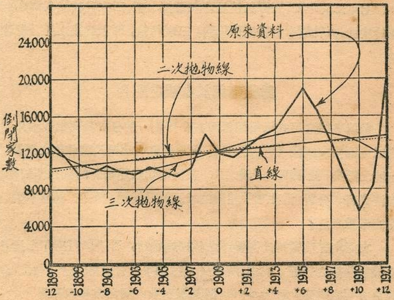
實際數值、由三種長期趨勢綫求得之常態數值、及實際數值對常態數值之百分差

年份	實際數值	直綫之常態數值	二次拋物綫之常態數值	三次拋物綫之常態數值	實際數值對各趨勢綫常態數值之百分差		
					直 綫	二次拋物綫	三次拋物綫
1897	13,083	10,342	10,025	12,215	+26.5%	+30.5%	+7.0%
1898	11,615	10,486	10,249	11,344	+10.8	+13.2	+2.3
1899	9,642	10,631	10,466	10,703	−9.3	−8.0	−10.0
1900	9,912	10,776	10,676	10,272	−8.0	−7.0	−3.8
1901	10,648	10,920	10,880	10,030	−2.5	−2.0	+6.5
1902	9,973	11,065	11,076	9,953	−9.9	−10.0	+0.2
1903	9,775	11,210	11,266	10,022	−12.8	−13.2	−2.5
1904	10,417	11,354	11,448	10,213	−8.3	−9.0	+2.1
1905	9,967	11,499	11,624	10,506	−13.3	−14.3	−5.1
1906	9,385	11,644	11,793	10,878	−19.4	−20.4	−13.7
1907	10,274	11,789	11,955	11,313	−12.9	−14.1	−9.2
1908	14,066	11,933	12,110	11,776	+17.8	+16.0	+19.4
1909	11,872	12,078	12,258	12,258	−1.7	−3.2	−3.2
1910	11,588	12,223	12,400	12,733	−5.2	−6.6	−9.1
1911	12,679	12,367	12,533	13,179	+2.5	+1.0	−3.8
1912	13,832	12,512	12,660	13,575	+10.5	+9.2	+1.9
1913	14,553	12,657	12,781	13,900	+15.0	+13.9	+4.6
1914	16,780	12,801	12,895	14,131	+31.0	+30.0	+18.6
1915	19,035	12,946	13,002	14,246	+47.0	+46.2	+33.5
1916	16,498	13,091	13,102	14,225	+26.0	+25.9	+15.9
1917	13,073	13,236	13,194	14,045	−1.2	−0.9	−7.0
1918	9,331	13,380	13,281	13,685	−30.2	−29.8	−31.8
1919	5,515	13,525	13,330	13,123	−59.1	−58.7	−57.8
1920	8,463	13,670	13,432	12,338	−38.0	−37.0	−31.4
1921	19,982	13,814	13,498	11,307	+44.6	+48.0	+76.8

商業倒閉家數之原來資料及由前述三個方程式所得之綫，均繪於圖五十七中。茲將y之實際數值（商業倒閉家數）、及由各方程式計算所得之常態數值、以及實際數值對常態數值之百分差，列於表七十內。

各種長期趨勢綫之比較

由以上討論及圖五十七之圖解，可見用各種曲綫代表長期趨勢所得結果頗有差異。在本例中，直綫與二次拋物綫極爲接近，但三次拋物綫所表現之長期趨勢則與兩者迥不相同。吾人選取此項資料，卽欲極度顯示此種不同之點。各年之常態數值，係由趨勢綫上各該年之縱坐標測度而得，故若所用趨勢綫不同，則視作常態之標準自亦不同，而由實際數值對常態數值之差所表示之循環變動，亦必隨所選曲綫之種類而異。此點可由表七十所列實際數值對常態數值之百分差見之。



圖五十七．　1897—1921年美國商業倒閉家數及其三種長期趨勢綫

實際數值對常態數值之百分差，由直綫求得與由二次拋物綫求得者極爲相似，但由三次拋物綫求得者，即顯有不同之處。表內百分差有正負之別。由三次拋物綫求得之差與由其他兩種趨勢綫求得之差彼此比較，正負號不同者共有四年，其餘各年正負號雖同，而百分數頗有相差甚鉅者。吾人常有根據已往資料所配合之趨勢綫延長至資料時期以外，以推測未來之可能趨勢者；若就此點而論，則三次拋物綫延長後所得結果，必不合理。趨勢綫之延長，無論其爲直綫或曲綫，雖皆含有揣測之意味，然延長直綫及二次拋物綫，則所得結果皆比較可靠。

關於選擇趨勢綫之原則，下文當更有詳盡之討論，此處可就數點作簡要之說明。如在長時期內，資料之實際數值始終高於趨勢綫或低於趨勢綫時，則此綫之配合恐不適當。在商業倒閉家數之一例中，直綫及二次拋物綫均不能緊隨實際資料之綫，而相距甚遠；尤以一八九九年至一九〇七年之九年間，實際數值始終在兩綫之下。三次拋物綫與實際資料頗爲接近，故就原有資料範圍以內，三者相較，似以此綫之配合最爲貼切。

對數應用於曲綫之配合

前述各綫之形式俱屬簡單而甚合用者，惟分析時間數列如用半對數式之直綫或曲綫，則尤爲適用。用半對數或比例尺度作圖之優點，前已論之詳矣，此種圖形最優之點，在能確切表示相對的或比例的變動。經濟資料之分析，常注重於相對的或比例的變動，故選擇趨勢綫時，此點應考慮及之。

在測度經濟資料比例變動之長期趨勢時，吾人仍可應用上述各種直綫或曲綫之普遍方程式，所稍異者，式中y一項須用$\log y$代替之，故直綫之方程式應爲 $\log y=a+bx$，遞升冪級數曲綫之方程式應爲 $\log y=a+bx+cx^2+dx^3+\cdots$。作圖時，如繪於算術尺度之圖上，可以$x$之眞數

與y之對數繪製；如繪於半對數尺度之圖上，則縱軸用對數尺度，橫軸用算術尺度，而x及y俱以眞數繪製。兩者比較，當以後者較爲簡便。

茲將配合 $\log y=a+bx+cx^2$ 式之曲綫時所需計算之步驟說明之。設吾人欲決定一九〇八年至一九二二年美國石油產量之趨勢綫，則須將相當數值代入常態方程式，以求a,b,c各常數。其所需數值列表於下：

表七十一

1908—1922年美國石油生產量

配合趨勢綫時所需各數值之計算

年份	x	y 產量 (百萬桶)	$\log y$	$x\cdot\log y$	$x^2\cdot\log y$
1908	−7	178.5	2.25164	−15.76148	110.33036
1909	−6	183.2	2.26269	−13.57614	81.45684
1910	−5	209.6	2.32139	−11.60695	58.03475
1911	−4	220.4	2.34321	−9.37284	37.49136
1912	−3	222.9	2.34811	−7.04433	21.13299
1913	−2	248.4	2.39515	−4.79030	9.58060
1914	−1	265.8	2.42456	−2.42456	2.42456
1915	0	281.1	2.44886	....	....
1916	1	300.8	2.47828	2.47828	2.47828
1917	2	335.3	2.52543	5.05086	10.10172
1918	3	355.9	2.55133	7.65399	22.96197
1919	4	378.4	2.57795	10.31180	41.24720
1920	5	442.9	2.64631	13.23155	66.15775
1921	6	472.2	2.67413	16.04478	96.26868
1922	7	557.5	2.74624	19.22368	134.56576
			36.99528	9.41834	694.23282

$n=15$　　　　$\Sigma(x)=0$

$\Sigma(\log y)=36.99528$　　　　$\Sigma(x^2)=280$

$\Sigma(x\cdot\log y)=9.41834$　　　　$\Sigma(x^3)=0$

$\Sigma(x^2\cdot\log y)=694.23282$　　　　$\Sigma(x^4)=9352$

此處所欲求解之三個常態方式程之形式如下：

$$\Sigma(\log y)=na+b\Sigma x+c\Sigma x^2$$
$$\Sigma(x\cdot\log y)=a\Sigma x+b\Sigma x^2+c\Sigma x^3$$
$$\Sigma(x^2\cdot\log y)=a\Sigma x^2+b\Sigma x^3+c\Sigma x^4$$

將求得之數值代入上式，則

$$36.99528=15a+280c$$
$$9.41834=280b$$
$$694.23282=280a+9352c$$

解式中之常數得

$$a=+2.450508$$
$$b=+\ .033637$$
$$c=+\ .0008488$$

故所求趨勢綫之方程式，爲

$$\log y=2.450508+.033637x+.0008488x^2,$$

此式之原點在一九一五年。

表七十二

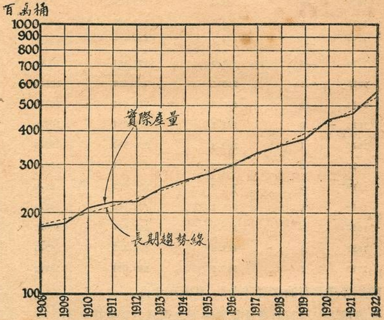
美國石油產量之長期趨勢及其實際值與趨勢值之比較

（趨勢綫係配合於產量之對數上者）

年份	x	y(實際值)產量（單位:百萬桶）	趨勢值之對數	y(計算值)趨勢值（單位:百萬桶）	實際值對趨勢值之百分比
1908	—7	178.5	2.256639	180.6	98.8
1909	—6	183.2	2.279242	190.2	96.3
1910	—5	209.6	2.303543	201.2	104.2
1911	—4	220.4	2.329540	213.6	103.2
1912	—3	222.9	2.357236	227.6	97.9
1913	—2	248.4	2.386629	243.6	102.0
1914	—1	265.8	2.417720	261.6	101.6
1915	0	281.1	2.450508	282.2	99.6
1916	1	300.8	2.484994	305.5	98.5
1917	2	335.3	2.521177	332.0	101.0
1918	3	355.9	2.559058	362.3	98.2
1919	4	378.4	2.598636	396.9	95.3
1920	5	442.9	2.639913	436.4	101.5
1921	6	472.2	2.682886	481.8	98.0
1922	7	557.5	2.727557	534.0	104.4

將代表某年之 x 值代入式中，即可求得該年趨勢值或常態值之對數，更由此對數以求其眞數。各年常態值及實際值對趨勢值之百分比，見表七十二。

各年實際產量之數字及其趨勢綫見圖五十八。由前項方程式所得之綫，極爲配稱，頗能表示石油產量之長期趨勢，可由圖中窺見一斑。



圖五十八．　1908—1922年美國石油產量及用對數配合之長期趨勢綫

試以此綫與根據產量之眞數配合所得相類之曲綫(二次拋物綫)互爲比較，即可立見其優點。該拋物綫常數之數值用最小平方法求得後，代入該綫之普遍方程式中，即得下式：

$$y = 279.387 + 24.262x + 1.650x^2$$

式中原點亦在一九一五年。石油之實際產量、及由上式求得之趨勢值、以及兩者之百分比，均列入下表內。

表七十三

美國石油產量之長期趨勢及其實際值與趨勢值之比較

（趨勢綫係配合於產量之眞數上者）

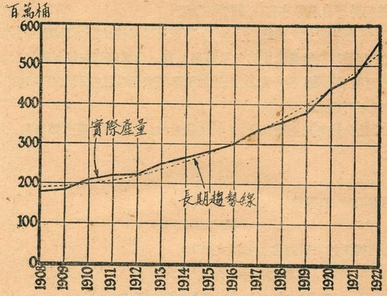
年份	x	y（實際值）產量（單位：百萬桶）	y（計算值）趨勢值（單位：百萬桶）	實際值對趨勢值之百分比
1908	—7	178.5	190.403	93.7
1909	—6	183.2	193.215	94.8
1910	—5	209.6	199.327	105.2
1911	—4	220.4	208.739	105.6
1912	—3	222.9	221.451	100.7
1913	—2	248.4	237.463	104.6
1914	—1	265.8	256.775	103.5
1915	0	281.1	279.387	100.6
1916	1	300.8	305.299	98.5
1917	2	335.3	334.511	100.2
1918	3	355.9	367.023	97.0
1919	4	378.4	402.835	93.9
1920	5	442.9	441.947	100.2
1921	6	472.2	484.359	97.5
1922	7	557.5	530.071	105.2

該趨勢綫及石油之實際產量，見圖五十九。

此處用以表示石油產量長期趨勢之兩種方程式中，均含有三個常數，故其結果可互相比較。由圖中觀之，配合於對數上之曲綫似較配合於眞數上者爲優，但實際須用一準確之量數以測度之，方可斷定。吾人可用均方根差作爲測度之標準。均方根差之計算與前此用於次數分配時之算法相同，所稍異者，各項差數係實際值對趨勢值之差，而非實際值對算術平均數之差。

配合於眞數之趨勢綫所求得之均方根差（卽標準差）爲 12.46，配合於對數之趨勢綫所求得之均方根差爲 9.44（兩者之單位均爲百萬桶）。

於此可見對數趨勢綫兩旁散佈之程度遠較眞數趨勢綫爲小，故在本例中，應用對數曲綫以表示長期趨勢較爲適當(註)。



圖五十九．　1908—1922年美國石油產量及用眞數配合之長期趨勢綫

前述兩類曲綫在經濟統計方面，最爲適用。時間數列之長期趨勢大都可用遞升冪級數之曲綫表示之。此項曲綫或用資料之眞數配合，或用資料之對數配合(其區別前者係用y值之眞數配合，後者則用y之對數配

(註)上列兩標準差之數值，係由單位百萬桶之實際差數求得者。如用百分差計算，則對數趨勢綫之優點當更顯著。由百分差求得眞數趨勢綫之標準差爲 4.01，對數趨勢綫之標準差爲2.69。

於此所應注意者，在比較兩綫之標準差時，該兩綫之方程式中所含常數之數目必須相同，因所含常數或多或少與標準差數值之大小有關。若所含常數甚多，而與所繪點數相同，則由此所配合之趨勢綫可通過各點，此時標準差之值將等於 0，但此綫已失去趨勢綫之意義，斷不可用以測度長期趨勢也。

合;至於代表時間之x值,則均用其真數)。此兩類曲綫均含有伸縮性,爲在曲綫中應用甚廣者。此外尙有數種曲綫,在時間數列方面雖爲用較狹,然有時所得結果亦甚佳,請分述於下。

普通拋物綫式之曲綫($y=ax^b$)不適用於經濟統計之時間數列,因吾人不應將時間變量依幾何級數處理之也。此種曲綫在雙對數尺度之圖中,可變爲直綫,前嘗言之。此綫如能準確表示某種數列之長期趨勢,亦未始不可採用耳。

此種曲綫可用對數配合,使成直綫形式,其配合手續亦頗簡易。方程式

$$y=ax^b$$

應用對數時 卽變爲

$$\log y=\log a+b\log x$$

配合此綫所用之兩個常態方程式如下:

$$\Sigma(\log y)=n\log a+b\Sigma(\log x)$$

$$\Sigma(\log x\cdot\log y)=\log a\Sigma(\log x)+b\Sigma(\log x)^2$$

根據實際資料計算上兩式中所需之各項數值,代入式中,卽可求得$\log a$及b之數值,其手續與配合普通直綫時相同(註)。

簡單之指數曲綫$y=ab^x$,用於分析時間數列,頗爲普遍。其對數方程式爲

$$\log y=\log a+(\log b)x$$

配合此種曲綫所用之兩個常態方程式爲

$$\Sigma(\log y)=n\log a+\log b\Sigma(x)$$

$$\Sigma(x\cdot\log y)=\log a\Sigma(x)+\log b\Sigma(x^2)$$

(註)Raymond Pearl 氏所著 "Medical Biometry and Statistics" 一書附錄內(一九二三年非勒特爾菲亞 Saunders 出版)載有一表,將眞數1起至100止之對數之和一一列入,頗切實用。

此綫與以前所述半對數曲綫之直綫形式相類似。配合曲綫時，先解常態方程式中兩個常數之對數，更由此對數求眞數，卽可將前項對數方程式還原爲指數曲綫之形式。

統計學家亦有用康氏曲綫(Gompertz curve)以解釋經濟統計之資料者。此綫之方程式原爲保險學家計算保險率時所發現。其形式爲

$$y=ab^{c^x}$$

提倡應用此式以分析經濟統計者，謂人口之增加常遵循一滋長之普通法則，而此法則在工業生產方面亦當存在，則以工業生產量爲人口滋長之函數故也(註一)。

布爾氏(Raymond Pearl)暨李特氏(Lowell J. Reed)曾用類似之滋長曲綫以預測人口之蕃殖(註二)，嗣經發現該綫亦可用以表示某種經濟資料之長期趨勢。其方程式如下：

$$y=d+\frac{b}{e^{-ax}+c}$$

配合曲綫之簡捷法：平均數法

曲綫之配合，欲求十分準確，自應採用最小平方法；但若配合之目的，僅欲求其近似值，則有時可採用他種方法。

吾人若用平均數法配合直綫，可將一數列分爲時期長短相等或近乎相等之兩部分，由兩部分所含x及y之數值，可求得如下式之兩個形式相同之方程式：

(註一)關於此綫之特性，在 Raymond B. Prescott 氏所著 "Law of Growth in Forecasting Demand" 一文內曾有敍述。原文見一九二二年十二月出版 "Journal of the American Statistical Association" 第471—479頁。

(註二)參看一九二三年 Committee of Plan of New York and Its Environs 發表Raymond Pearl 暨 Lowell J. Reed 兩氏所著 "Predicted Growth of Population of New York and Its Environs".