# 統計學新論

王思立編著

立信會計圖書用品社發行

# 統計學新論

王思立編著



立信會計圖書用品社發行

# 自　　序

統計方法與統計科學之重要，由來雖久，關於二者之著作，無論中西，求其博約兼顧，條理淸晰，切合大學課本之條件者，尙不多見。思立先後在交大黔校，中央大學，交通大學及立信會計等校主講統計學一科，迄今七載。鑒於學子之習斯科者參攷乏書，爰於授課之餘，從事編集。七年之中，講授近二十次。事前事後每將講稿加以修正補充，近且經數月之埋頭整理，晨夕無間，卒成此書。付梓之後，非但個人多年來艱苦工作粗告段落，莘莘學子之初習斯科者，亦可得一可用之課本。此書編著之另一動機，在供給統計工作人員及工商實業界人士作爲參攷。

本書共約二十萬言，分爲六編二十三章，足供大學或專科學校一年中講授之用。各章之末，每附有統計資料及問題，俾學生課外練習時，有所參照。惟以有系統之數字資料獲得匪易，教師如有更適當之資料，足供應用，自不必爲書中例舉者所限。

本書初稿大半由交通大學財務管理系同學王君朝典代爲謄寫。此外，航空工程系同學方君天佐於繕稿及繪圖賜助亦多，均誌謝忱。又本書倉卒付印，錯誤之處勢所難免，幸祈先進專家與讀者諸君發現時，不吝賜教，俾來日再版時加以更正。

民國三十六年十月

王思立序於上海徐家匯國立交通大學

# 統計學新論

## 目錄

### 第一編 統計資料

## 第四編 物價指數

## 第五編　時間數列之分析

# 第 一 編

# 統 計 資 料

# 第一章　統計學之性質

## 第一節　統計之意義

統計（Statistics）一詞，古今之意義不同，古代統計學之對象爲國家。無論對外或對內，政府在決定政策之前，對於國內政治情況，諸如國家之疆土，人口 財富等，必須加以調查或估計，然後方能依據實際情形，確定政策。統計學即依文辭敘述之方法，記載國家各種政治情況之學。此種統計學實係政治學、經濟學、與地理學三科中一切數量材料之總匯。古代如此，直至近代初期仍如此。十八世紀中年德國統計學頗爲發展，當時英國學者亦稱道之，並主張其國人宜急起直追。彼等所稱之統計學依其對象及方法觀之，仍未脫離古代統計學之範圍，即以文辭敘述之方法，以說明國家之政治情況者。

十九世紀中葉，由於政治、經濟、文化各方面劇變之結果 人民生活及各種科學之內容，無不與之俱變，統計學中變更尤大。經最近百年以來之演化，至今日統計一名詞之意義，可分爲下列三層：

一、統計資料：所謂統計資料（Statistical Data），已不限於國家政治方面之事實，而指社會現象及自然現象中任何用數字表示之材料。日常所謂統計，每即指統計資料而言，如人口統計、資源統計、教育統計、工商統計、物價統計、金融統計、進口貿易統計、出口貿易統計、無一不指以數字表示之事實而言 企業單位、工商團體，及政府機關之報告中所記載者，多半爲統計資料，良以如此方足以見出過去及現在之情況，並顯示將來之趨勢。此種統計資料之特點，在其自身係由許多項因素造成之結果；

此種結果，又時時在變化中，決非只受一種因素支配或一成了變。

二、統計方法：所謂統計方法（Statistical Method）即搜集分析及表示數字資料之方法，亦稱之為統計技術 Statistical Technique）。社會愈進步，吾人對於生活及工作上所遇之問題，愈欲求確實具體之解答，此種解答必須依賴數字事實。故社會愈進步，學術愈昌明，數字材料之積累亦愈多，處理方法，亦愈益重要。科學方法原分有演繹法與歸納法兩種，而歸納法又分實驗法（Experimental Method）與統計法（Statistical Method）兩種。統計方法係處理大量數字資料之唯一方法。數字資料儘可分門別類，但處理之工具為統計方法則同，統計方法之普遍性與數學及語文無異。

三、統計科學 所謂統計科學（Science of Statistics）或統計學，(Statistics) 亦稱之為統計理論，（Theory of Statistics），學者所舉之定義不同．英國統計學者包萊教授（Prof. Bowley）在所著之統計學概論（Elements of Statistics）中，初謂統計學為計數之學（Science of Counting），繼謂統計學為平均數之學(Science of averages)統計資料既為大量事實，統計方法既為處理此大量事實之方法，則計算數目及平均數在統計學中自極重要，但如謂統計學即為計數之學或平均數之學，亦實過份偏狹。統計方法為科學方法，統計學則為科學方法論，係研究科學方法中統計方法之學。至于計數與平均數則僅為統計方法中之重要節目而已。

## 第二節 統計方法之程序

統計資料門類雖至紛繁，但自應用統計技術處理過程觀之，仍有同一之程序 其步驟如下：

一、搜集資料：統計方法之對象既為數字資料，是以第一步工作自應

爲資料之搜集。統計資料可分爲原級資料（Primary Data）及次級資料(Secondary Data)兩種：前者指研究者親自搜集之資料。亦稱之爲原始資料（Original Data）。後者則指他人或其他機關已搜集發表之資料，亦稱之爲轉得資料（Derived Data）。所謂搜集資料，嚴格言之，只指原級資料之如何取得，擴而言之，兼亦包括次級資料之採集與甄別。統計資料，雖有原級及次級之分，但搜集後爲分析研究之張本則初無二致。

二、整理資料：無論原級或次級資料，每嫌凌亂，須需整理工作，方合於分析之用。整理之方法爲分類。分類之前各個事實依其原有面目而存在，稱之爲統計項目（Statistical Items）。既經分類之後，以歸併化簡之結果，個別項目已不復存在，全部資料組成爲有系統有秩序之統計數列（StatisticalSeries）。此種統計數列每用表格表出，簡單清晰，避免文字上。之重複說明，事半功倍。此項表格稱之爲統計表（Statistical Tables）分類爲整理材料之方法，而製造統計表則又爲分類之手段。在整理資料中至關重要。

三、表示資料：統計資料經過分類化簡，組成數列之後，須設法表示或發表。發表之目的或爲供他人閱覽，或爲供自己作進一步分析之根據。表示之工具約有三種方式：

1. 文辭方式：即用文辭平鋪直敘。此種方法在初期統計學中每採用之，近百年來統計技術進步，此種粗淺之方法已不適用。

2. 表格方式：所謂表格方式即將統計資料，依合理之體系，列入縱橫欄格之內，以期便於比較或分析，使資料各部份間之關係，充分現露。無論供他人閱覽或爲自己分析，統計表格俱爲極重要之工具。

3. 圖形方式：統計表之要旨在以最經濟之空間，最具體之方法將全部數列表出，但如吾人注意之點，不在詳細數字，而僅在其梗概或趨勢，則

有待於統計圖。蓋圖形係利用點，綫與面積，將統計資料繪出一輪廓，活現于吾人之目前。但圖形不能離開統計表格而獨立存在。因圖形所表示者爲材料之縮寫，而表格所表示者則爲資料之細目。縮寫或大綱，究不宜離開其細目而獨存。

四、分析資料：以上三步工作，即搜集資料整理資料與發表資料，俱可視爲統計研究中之準備工作。分析資料方爲統計研究之本體、或其中最重要之部份，統計資料經分類整理後所成之統計數列，最要者有二種，即次數數列（Frequency Series）與時間數列（Time Series）。故統計分析每以次數與時間兩種數列爲限。

五、統計推理：統計分析每根據一部份資料，稱之爲統計樣本（Statistical Sample）最後以樣本研究之結果，擴而充之，用以說明全部資料，稱之爲統計全體（Statistical Population）。如依大學男生百人身體高度而得之平均數，充爲全國男生身高之平均數，依上海市中五十家銀樓金價之平均數，充爲全市金價之平均數。此種過程稱統計推理（Statistical Inference）或統計概括（Statistical Generalization）或逕稱之爲統計歸納（Statistical Induction）。依樣本求得之結果，與依全部資料所得之結果，如完全符合，則以局部代表全體，自正確無誤。如依樣本所得之結果與依全體而得之結果中間，雖有出入或差額，但此差額之限度，可以求出，則以局部代表全部，雖稍有不脗合處，但吾人已可確定樣本結果可靠之程度。此種統計推理作用，對于統計研究關係重大，其理由不難想像。

總之，統計方法之程序分爲五步驟：即搜集資料，整理資料，發表資料，分析資料，及批判結果。與一般科學方法之程序，初無大異。因科學方法之步驟，不外觀察分析與驗證，統計方法實亦如此。

## 第三節　統計方法之功用

統計學之對象爲統計方法。統計方法與實驗方法，同爲科學方法。但實驗方法應用于自然科學，而統計方法則應用于社會科學。因自然科學之對象較爲單純，可在實驗室中聽研究者自由支配，反覆試驗，以發現其共相，由共相以成立原理原則，舉例言之，以氫和氧之爲水，水遇熱化爲汽，遇冷化爲冰，研究者可在實驗室中屢驗不爽。反之，社會科學之對象爲社會現象，在本質上與此絕不相同，不能用人力加以控制，而只能按其實際發生之情形，加以觀察歸納，以發現其共相，最後得到原理原則。諸如人口之增減，物價之漲落，以及任何天災人禍之發生，任何社會現象之變動，俱只能隨時或就地觀察，歸納，即只能用統計方法，而不能自由加以支配，即不能用實驗方法。

社會科學中固不能應用實驗方法，即自然科學之對象，亦有不能加以控制實驗者，舉凡不能由人力控制施以實驗者，雖爲自然現象，隸屬于自然科學範圍之內，但仍須依重於統計方法。在天文學，氣象學，生物學，水力學中，均遇有此種情形，是均須利用統計方法。即最純粹之自然科學，如物理學者，有時亦有因上述情形而須利用統計方法之處。是以統計學非社會科學，統計方法亦非社會科學所專有，更非少數社會科學，如經濟學及教育學所獨有。在此二種科學中，數字資料及數量方法最爲重要，但吾人不能以統計方法在經濟學及教育學中特別重要，即抹殺統計方法之普遍性也。

# 第二章　統計資料之搜集

## 第一節　統計單位之確定

統計方法之對象爲具體事實，絕非抽象之數字，具體事實必有其單位，是爲統計單位（Statisticalunit)。在算學中之數目爲100，在統計學中100 之後必附以單位，如100人，100元，100斤，100 家公司，100 所工廠，100 次罷工，100 件犯罪。單位如爲人，貨幣單位之圓或度量衡單位如斤、尺、等，意義簡明，無須解釋，只要搜集資料過程中，計算數目不發生錯誤，所得之資料，當無何問題，但計數之單位如爲公司或工廠，則情形不同。以公司而論，依公司法規定，企業單位之稱爲公司者必須有一定之條件，並曾履行一定法律程序，但實際上所自稱爲公司者，未必于此相合，吾人所稱之公司，究依法律之觀點，或另有觀點，須加規定。又如計數之單位爲意外傷害，何種情形方構成傷害，極有出入，以工廠傷害而言，須工人受傷至何種程度，影響工作時間至何種期限方稱爲傷害。又如以犯罪而論，何謂犯罪，如不加以統一之規定，並由調查員一致遵守，則所得之資料，必彼此不同，而缺乏純一性質（Homogeneity），不純一之資料，根本不能加以分析與比較，强行分析或比較，所得之結論必不正確。

美國統計學家塞克力特（Secrist）在所著統計方法引論（Introduction Of Statistical Methods）中，區分統計單位爲三類，茲參照塞氏之意見，加以補充，而將統計單位分類如下：

一、單純單位（Simple units）：凡單位只表示事物之一種情況，即事物之種類相同者，爲單純單位。如統計單位爲人，自與其他動物不同。爲

公司自與其他企業組織不同。此種單位以其性質單純，如經明確統一之規定，加以調查人之一致遵守，不易發生流弊，即不致造成不純一之資料。

二、複合單位（Composite unit）原爲單純單位，表示一種情況，但兩個單純單位可聯合爲一單位，表示事物之兩種情況，如噸（ton）爲表示重量之單位，哩（Mile）爲表示距離之單位，各爲單純單位，而噸哩（Ton-mile）則成爲複合單位，此外呎磅（Foot-pound）亦是。複合單位所異于單純單位者，只在後者表示一種情况，前者則表示二種情況，其餘相同，均爲搜集資料時計算數目之單位，故合稱爲列舉之單位（Units of Enumeration）。

三、比或係數（Ratios or Coefficients）計算一種事物之絕對多寡，意義不顯明，如能在單位之中包括有相對之比較，則予人之印像必較爲顯明而深刻，是爲係數，係數或比爲單位，可有下列三種情形：

1. 關于時間之係數，如工人每日或每月之工資若干。

2. 關于地域之係數：如每一畝地出產稻米之數量，每一省份人口之數目。

3. 關於情況之係數：如一歲嬰兒死亡率爲多少，以及各個年齡之死亡率爲多少，概以千分之幾表示之。此種單位不論其關於時間，空間或情況，概稱爲比較與解釋之單位（Units of Comparison and Interpretation）

## 第二節　搜集原級資料之方法

統計資料自其獲得之方法觀之，可分下列二類：

1. 原級資料（Primary Data）

2. 次級資料（Secondary Data）

原級資料卽調查者自己搜集之資料，次級資料則爲採用他人搜集發

表之資料。二者之區別並非絕對不移，如甲所搜集之資料，對甲爲原級資料，乙採用之則爲次級資料。所謂統計資料之搜集，可兼指原級與次級兩種資料之獲得，但嚴格言之，則只指原級資料之獲得而言。其方法有下列三種：

一、採訪法（Interview Method）可分兩種情形：

1. 個人採訪法（Personal Investigation）即研究者親自實地調查法國學者勒普萊（Le Play）之調查工人家計（Family Budgets）最爲出名。氏爲調查工人生活，必親至工人家中居住，詳細調查其收支情況，一家完竣，再轉入另一家中同樣調查，經長期積累，最後得有系統之資料。此法之優點：一爲直接自來源採訪資料，所得者較爲眞實可靠。二爲由一人自始至終負責採訪，統計單位能劃一規定，前後一致所得資料合于純一性（Homogeneity），純一之資料加以比較與分析，方有價值。其缺點：一爲可有主觀成分介入，且以一人始終負責，主觀成分亦一致存在，不能相互抵消，結果資料殊欠公允。二爲一人之時間精力有限，此法所能調查之範圍有限。範圍較小之問題尚可應用此法，如範圍大，或限期完成，則根本不能用此法，因完成調查之期限太長，現象已變，前後搜集之資料已不合于純一性。

2. 調查員採訪法（Interview by Enumerators），一人之時間精力有限不能勝任，乃派出多數調查員同時採訪，可以補救上法之不足，調查範圍雖大，以調查員人數衆多，仍可限期完成，允稱良法。惟調查員既多，對于統計單位劃一之規定更爲重要，必須事先使各個調查員對于統計單位有一致之認識，並於調查期間確實遵循無渝，所得結果方可公允純正，否則各調查員對單位之見解不一，所搜集之資料亦必偏畸複雜，不能加以綜匯而供分析之用。

二、塡表調查法：（Questionaire Method）調查員採訪法之困難，爲所需人員及費用浩大，非私人及一般機關所能担負。乃有通信調查法，其法卽由調查機關印製表格，內載調查事項，寄由被調查人塡報寄還，被調查者供給資料，稱爲報告人（Informant）。此法經濟迅速，公私統計調查，每常用之。但其缺點則：一爲不完全：所發出之調查表，由報告人塡報寄還者每只佔一小部份，大部份調查表則寄出後杳無回音。以此少數答案而代表全部資料自嫌不充分。二爲不正確：所得之答案根本不多，在此無多之答案中，眞實可靠者固然有，但亦有漫不經心，或有意虛報者。

通信調查法所得答案之多少眞僞，于表格中所載之問題關係甚大，英統計學者包萊氏在所著統計學概論中，對擬定問題曾提出四項標準：

1. 收件人答覆之能力（Ability to answer）統計調查之範圍較廣，收到調查表者教育程度高低不同，表中所列問題必使教育程度低者亦能答覆，如問題太專門化，則只教育程度高者能答覆，其餘或根本不明瞭所詢事項或已遺忘，無從回答。

2. 問題及措辭宜簡明具體（Definiteness）所問旣簡明具體無模糊不淸或模稜兩可之處，則答案方能確實，言之有物。依此標準最好所問爲是或非，一簡單數字，一確定日期，一確定地點，或類此之簡明問題。所得答案確切，方能根據以製表。否則敷衍了事之答案，雖收回亦無用。

3. 收件人不厭于答覆（Reluctance to answer）問題之內容與措辭必使收件人樂于報告而不感覺厭煩，甚至恐懼。如調查各人之財產，收入、年齡、婚姻、職業、信仰、心身缺陷等，收信人每不樂于回答，深懼答覆後引起不利之結果，或自認爲私事無告人之必要。

4. 眞實無僞（Truthfulness）卽所詢者不但收件人不厭倦回答，其答案確爲眞實而不虛僞，無參以主觀意見使與實際不合之必要，如調查各

個人之社會地位，過去經歷等。收件人于囘答時，常有意虛報，使結果不可靠。

三、估計法（Estimation）當實際調查不可能時，亦可用估計法以得資料。統計之對象爲大量數字資料，其甄別資料時所訂確實合用之標準，與算學不同，估計法用之得體，所得材料每亦甚爲可靠，農業統計中對于收穫統計（Crop Statistics）每用估計法以求之，初依過去之經驗與現時各地穀物生長之情形，估定秋收數量，向負責調查機關報告並可隨時依雨量，氣候情形加以修正，直至實際收穫時爲止，故所報告之資料之可靠性與日俱增。

## 第三節　次級資料之來源

次級資料即由他人搜集，整理、分析，發表、之資料，其來源爲數至多，實難一一列舉，歸納言之，次級資料之源可有四類：

一、政府機關之公報：各級政府機關，每有統計部門專司調查統計資料，並發表其研究之結果，供給行政及一般之參攷，以吾國政府統計機關爲例，國民政府有主計處統計局。各部會有統計處，各省市以至縣政府有統計室，各有其經常或臨時公報，發表統計資料，以供行政之參攷。

二、工商協會：各種工商行業，每各有協會或公會之組織，其任務之一，即爲搜集于各個會員有關之資料加以整理與分析，刊行發給各個會員供給業務上之參攷。故遇于某一行業有關之問題須有數字資料時，可查閱其公會之報告。

三、學術機關：各大學、法商學院，研究所或其他研究機關，亦常有統計調查，以其研究之結果公佈于世，如戰前南開大學經濟研究所曾編有華北物價指數按期發表即例。

四、刊物及報張：各統計期刊，或經濟、財政、金融類之刊物，多載有統計資料，如中央銀行月刊、中國農民銀行月刊、財政評論等，俱載有行情及其他種經濟資料，足供參攷，此外各種報紙亦均載有統計資料。

# 第三章　統計表格

## 第一節　統計表之意義及種類

無論原級資料或次級資料，均須經過適當整理，然後方能分析，整理之方法爲分類，表示分類之方法爲製表（Tabulation）。所謂製表，即循一定之原則，將統計資料列入縱橫空間之內，以便于比較與分析。此項表格稱爲統計表（Statistical Tables）統計表可別爲二大類：即總表（General Tables）與簡表（Summary Tables）總表中所載爲關於某種現象之全部數字資料，按其本來面目，加以彙合登記，而未加任何改變，故亦稱爲原級表（Primary Tables）就其用途而論，總表中所載係供給各方面作參攷之用，故亦稱爲一般用途表（General-Purpose Tables）舉例言之，如調查上海全市工廠，已收得答案後，即按照所有權、組織、資本總額、生產品種類及數量、工人人數、工資制度及金額等項，製定總表，將各個答案中所載之項目逐一登記于總表之內，除化零爲整外，對于各答案之原有面目，毫無更改。

與總表對待者爲簡表，意即將總表中所載之資料，加以分類化簡，以期便於特種需要，亦稱爲特殊用途表（Special-Purpose Tables）或轉得表（Derivative Tables）亦可稱之爲次級表（Secondery Tables）。

## 第二節　統計數列之種類

原級資料多爲個別分立之事實，稱爲統計項目（Statistical Items）。雖經登入總表，仍爲臚舉全部事實，未經整理，自無法分析。整理之方法爲

分類。分類之後，個別項目已不存在，而成爲統計數列(Statistical Series)分類時所根據之標準有三，統計數列亦隨之而分爲三種，每種數列，均以簡表表出之：

一、品質數列

所謂品質數列（Categorical Series）即根據各個項目共同具備之一種性質而分類。例如中央政府稅收，根據來源分爲關稅、貨物稅、所得稅等。中央政府支出，根據用途分爲政務費、國防費、建設費、教育費等。又以同學性別、籍貫、年級、院系等爲標準，結果均爲品質數列，記載品質數列之表格稱爲品質表（Categorical Tables）試以下列第1,2,3表中交大三十五年度第一學期註冊學生人數統計資料爲例，以見品質表之性質。

(第1表)　國立交通大學各院學生人數統計表

| 院　系 | 人　數 |
|---|---|
| 管理學院 | 539 |
| 理學院 | 170 |
| 工學院 | 2,159 |
| 專修科 | 127 |
| 先修班 | 187 |
| 電信研究所 | 14 |
| 總　計 | 3,246 |

(第2表)　交大各年級學生人數統計表

| 年　級 | 人　數 |
|---|---|
| 一 | 937 |
| 二 | 899 |
| 三 | 921 |
| 四 | 489 |
| 總　計 | 3,246 |

（第3表） 交大學生性別統計表

| 性別 | 人數 |
| --- | --- |
| 男 | 3,013 |
| 女 | 233 |
| 總計 | 3,246 |

在上列三品質表中，分類之標準不拘爲院科，年級或性別，均爲各項目俱備之一種性質，因稱之爲單項表格（Single Tabulation）。但分類時亦可根據統計項目之二種或三種性質，結果所成之表格爲雙項表（Double Tabulation）及三項表（Treble Tabultion）仍以交大學生統計資料爲例，以見雙項表（第4表）與三項表（第5表）之製法：

（第4表） 雙項表格——交人各院各年學生人數

| 院別 ╲ 年級 | 一 | 二 | 三 | 四 | 總計 |
| --- | --- | --- | --- | --- | --- |
| 管理學院 | 188 | 161 | 136 | 101 | 589 |
| 理學院 | 38 | 58 | 61 | 13 | 170 |
| 工學院 | 494 | 607 | 688 | 370 | 2,159 |
| 專修科 | 28 | 61 | 36 | 2 | 127 |
| 先修班 | 187 | | | | 187 |
| 電信研究所 | 2 | 12 | | | 14 |
| 總計 | 937 | 899 | 921 | 489 | 3,246 |

二、時間數列

社會現象每隨時間變動，無論消費、生產、貿易、價格、工資、僱用、出生、死亡、俱因日、因月、因年不同。按各個統計項目發生之先後次序，加以分類，製成數列時，稱之爲時間數列（Time Series）或歷史數列（Historical Series）記載此項歷史數列之表格，稱爲歷史表（Hi t:rical Tables）。同爲歷史表，如只記載各年或各月之總數，則爲單項表，如將各年

各月之總數分為二部或三部份則為二項或三項表。現以中國經濟研究所所編十年來之上海批發物價指數為例，以見時間數列之性質:

(第5表) 三項表格——交大各院各年男女生人數統計表

| 院別 \ 性別 \ 年級 | 一 | | 二 | | 三 | | 四 | | 共計 | | 總計 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 男 | 女 | 男 | 女 | 男 | 女 | 男 | 女 | 男 | 女 | |
| 管理學院 | 164 | 24 | 127 | 34 | 102 | 34 | 74 | 30 | 467 | 122 | 589 |
| 理學院 | 33 | 5 | 37 | 21 | 44 | 17 | 10 | 3 | 124 | 46 | 170 |
| 工學院 | 485 | 9 | 592 | 15 | 674 | 14 | 365 | 5 | 2,116 | 43 | 2,159 |
| 專修科 | 27 | 1 | 59 | 2 | 36 | 0 | 2 | 0 | 124 | 3 | 127 |
| 先修班 | 170 | 17 | | | | | | | 170 | 17 | 187 |
| 電信研究所 | 2 | 0 | 10 | 2 | | | | | 12 | 2 | 14 |
| 總計 | 881 | 56 | 825 | 74 | 856 | 65 | 451 | 38 | 3,013 | 233 | 3,246 |

(第6表) 上海批發物價指數表

簡單幾何平均數　　民國二十五年=1

| 年份 | 指數 |
|---|---|
| 二十六年 | 1.166 |
| 二十七年 | 1.426 |
| 二十八年 | 2.320 |
| 二十九年 | 5.057 |
| 三十年 | 10.990 |
| 三十一年 | 34.530 |
| 三十二年 | 143.600 |
| 三十三年 | 1,007 |
| 三十四年 | 39,667 |
| 三十四年 | 1,190 |
| 三十五年 | 5,199 |

註：上表轉錄自方顯廷主編經濟評論第一卷第一期，就中三十一年四月至三十四年十月係按偽幣市價編製。

三、次數分配

如分類時所根據之標準，非為時間地域或各項目所俱備之性質，而係根據各項目數值之大小，分類結果所形成之數列為次數數列（Frequecy Series）或稱為次數分配（Frequency Distribution）。記載次數分配之表格稱為次數表。次數表（Frequency Tables）為整理統計資料最重要之利器，在統計技術中之地位至關重要，茲將製造次數表之程序，舉例說明之。

（第7表） 大學男生二十人高度表

單位：公分

| 號數 | 高度 | 號數 | 高度 |
|---|---|---|---|
| 1 | 167.8 | 11 | 166.3 |
| 2 | 160.2 | 12 | 160.0 |
| 3 | 163.5 | 13 | 158.5 |
| 4 | 171.3 | 14 | 116.6 |
| 5 | 166.5 | 15 | 170.4 |
| 6 | 167.0 | 16 | 163.9 |
| 7 | 171.0 | 17 | 172.5 |
| 8 | 169.2 | 18 | 156.6 |
| 9 | 171.6 | 19 | 166.6 |
| 10 | 162.5 | 20 | 166.4 |

根據上列大學男生二十人之高度為例，製造次數分配或次數表之程序，可分為下列幾步驟：

1. 決定組距之多少與大小　上列之大學男生二十人之高度，如認為個別項目過多，須加整理歸併，製為次數表時，最初須決定最大項目及最小項目之數值，以及兩者間相差異之範圍。求最大最小項目之程序，應將所有項目依其數值之大小加以排列，稱為統計序列（Statistical Array）

序列之第一項爲最小項目，最末一項爲最大項目。但不經製成序列，亦可直接依觀察法以確定最大與最小兩項目。由上列資料中可見最小項目爲第十八項，其值爲 156.6 公分，稱爲下限，以 L 代表之。最大項目爲第十七項，其值爲 172.5 公分，稱爲上限，以 U 代表之。上限減下限，卽 (172.5 −156.6)＝15.9 公分，爲此二十人高度相差之範圍 (Range)。如將此範圍分爲五個相等之小範圍 (Subranges) 或組距 Class Intervals) 並取其整數時，則組距之大小爲四公分。如將組距加大，則分成組距之數卽減少。反之如將組距再改小，則分成之組距必更加多。組距多少原無一定標準，但自原則言之，如分成之組距過多，則結果未能充分化簡，計算不便，失去分組之意義。反之，分成之組距過少，則歸併化簡過份，使資料之特點消失。在普通情形下數百項之數列中，所包括組距之數目每以十組爲準則。現在男生二十人之高度資料中，姑且分爲五組，其組距爲 4 公分。

2. 選定組限及中值 上列二十人高度中，已定組距爲 4 公分，組數爲五個。現規定各組距上下限 (Class Limits) 及介于上下組限間之組中值 (Class Mid-values) 如下(第 8 表)：

(第 8 表) 二十男生身高分類組距及中值表

| 高度組距 | 中值 |
|---|---|
| 154—158 | 156 |
| 158—162 | 160 |
| 162—166 | 164 |
| 166—170 | 168 |
| 170—174 | 172 |

確定組限之標準有二：一卽使組中值能充分代表一組距之內各項目之原值。每組距所包括之項目甚多 計算不便。現率以中值代表之，並非各個項目之原值 俱等於中值，以中值爲標準而觀其項目時，有大于中值

者，有小於中值者。中值爲一折中數值，使大小於中值者從中抵消。結果以中值乘次數恰等於各項原值之總和。換言之卽中值等於各項目數值之平均數。二卽便於計算，蓋分類製表爲統計分析之開始，將來種種計算，不再根據原始資料，而根據次數表。所謂根據次數表者卽根據各組之中值及次數。中值選擇得當對於計算關係至大。中值如能爲整數或五之倍數則計算簡易。反之，中值爲小數則計算繁難。總之，組限之選定，卽以使中值能爲整數並能充分代表各項原值爲準。

組限之表示方法亦值一述。統計資料自其代表一變數 (Variable) 觀之，可分爲二種，卽連續變數 (Continuous variable) 及間斷變數 (Discrete variable)。如以 x 代表變數，所謂連續變數卽在 a b 範圍內 x 可爲任何數值，吾人現在舉例中之高度及重量均屬此。以高度而論，如衡量之工具精密，則高度自可有小數。反之，所謂間斷變數卽 x 之值只能由 a 躍而爲b，a．b之間不容有其他數值，如計算人數或爲十人，或爲十一人，不容爲 10，5 人。計算年齡亦然。變數有此區別 組限表示之方法亦異。連續變數如現在之高度組限，表示之方法如下：

(第 9 表)

| 組距 | 中值 |
|---|---|
| 154—158 | 156 |
| 158—162 | 160 |
| 162—166 | 164 |
| 166—170 | 168 |
| 170—174 | 172 |

(第 10 表)

| 組距 | 中值 |
|---|---|
| 154—157.99 | 156 |
| 158—161.99 | 160 |
| 162—165.99 | 164 |
| 166—169.99 | 168 |
| 170—173.99 | 172 |

上列第 9 第 10 兩表中表示方法相同，但第 10 表中所用之表示方法可免去分類登記時發生誤會。如某一高度爲 158，由第一種方法則分入第一組或第二組，未經解釋不易斷定。但由第二種辦法，則 158 自歸入第二

組內。第一組之組距仍爲 154—158，其上限仍爲 158，但爲清楚計，寫爲 157.99。卽使用第二種寫法，158 亦當入第二組內。凡數值小於上限者，列入本組內；其等於上限者，則列入較高之一組中。至於間斷變數，如年齡，則當採用下列第 11 表中表明組限之辦法。

(第11表)　間斷變數分類組限表

| 組　距 | 中　値 |
|---|---|
| 1—10 | 5 |
| 11—20 | 15 |
| 21—30 | 25 |
| 31—40 | 35 |
| 41—50 | 45 |
| 51—60 | 55 |
| 61—70 | 65 |
| 71—80 | 75 |

3.　製造登記表

所謂登記表或稱記號表(Tally Sheet) 卽在一表格中標明各組組距，然後卽攷察每一項目或變項 (Variate) 之數值，應屬於何組距範圍內，卽在對該組之空格中作一記號，如此逐項記入表中。以大學男生二十人之高度爲例，可作成下列之登記。

(第12表)　大學男生廿人高度登記表

單位：公分

| 組　距 | 人　數 |
|---|---|
| 155—157.99 | / |
| 158—161.99 | /// |
| 162—165.99 | /// |
| 166—169.99 | /// |
| 170—173.99 | |

4.　製造次數表

登記表為整理資料最基本工作，逐項登記時必特別用心，蓋至易發生錯誤。登記表完成後，製造次數表則至為簡易。即製成表格，列有組距，中值及次數三欄，在次數一欄中將各組之記號，譯為數字即成。

(第13表) 大學男生廿人高度次數表

單位：公分

| 高度組距 | 中值 | 次數 |
|---|---|---|
| 154--158 | 156 | 1 |
| 158--162 | 160 | 3 |
| 162--166 | 164 | 3 |
| 166--170 | 168 | 8 |
| 170--174 | 172 | 5 |
| 總計 | | 20 |

## 第三節 變相次數表

以上係普通次數表製造之程序。此外尚有二種統計表，係由次數表變化而來者，亦在此說明，一為累積次數表，二為相關表。

一 累積次數表

累積次數表即將次數表中之次數逐項加起，以得新次數。有向上累積 (Upward Cumulative)及向下累積 (Downward Cumulative)二種辦法，茲依上列次數分配用第14，15兩表說明如下：

(第14表) 大學男生二十人高度分配向上累積次數表

| 組距 | 次數 | 向上累積次數 |
|---|---|---|
| 154—158 | 1 | 1 |
| 158—162 | 3 | 4 |
| 162—166 | 3 | 7 |
| 166—170 | 8 | 15 |
| 170—174 | 5 | 20 |
| 總計 | 20 | |

(第15表) 向下累積次數表

| 距　組 | 次　數 | 向下累積次數 |
|---|---|---|
| 154—158 | 1 | 20 |
| 158—162 | 3 | 19 |
| 162—166 | 3 | 16 |
| 166—170 | 8 | 13 |
| 170—174 | 5 | 5 |
| 總　計 | 20 | |

向上累積次數，係以每組之上限爲標準，而計算在此限度以下之次數。如在二十人高度分配中，在158公分以下者，計一人，在162公分以下者，計四人，在166公分以下者計七人，在170分以下者，計十五人，全部二十人之高度俱在174公分以下。是以向上累積次數表亦稱爲"較小"累積次數表("Less than" Cumulative Frequency Table)。至於向下累積次數表與此恰相反，係以每組之下限爲標準，而計算在此限度以上之次數。如二十人高度分配中，在170公分以上者計五人，在166公分以上者十三人，在162公分以上者計十六人，在158公分以上者計十九人，全體二十人之高度在154公分以上。是以向下累積次數表，亦稱爲"較大"累積次數表("More than" Cumulative Frequency Table)

二　相關表

次數表之第二種變相形式爲相關表(Correlation Tables)。兩種現象彼此有關係之實例甚多。如學生之高度與重量有關，學生之勤學與成績有關，貨幣數量與物價有關，美鈔價格與黃金價格有關。凡兩種有關之資料，如所包括之項目過多時，須加以分類化簡。相關資料分類時所用之登記表及分類結果而製成之次數表，除在一表以內包括甲乙二種資料之外，一切程序與整理一種統計資料，製造次數表相同。現以民國卅六年元月份上海

黃金與美鈔市價資料爲例，在登記表中，一方爲黃金價格之分類，一方爲美鈔價格之分類，因稱爲複式分類，其登記表爲複式登記表 (Double-entry Tally Sheet)，其次數表爲複式次數表 (Double-entry Frequency Table) 或稱爲相關表(Correlation Table)。現錄卅六年元月份黃金與美鈔價格如下，(第16表) 並依之而製成複式登記表(第17表) 與相關表(第18表)。表中以黃金市價爲自變數，以 x 代表之，在複式登記表中橫列，自左而右；以美鈔市價爲因變數，以 y 代表之，在登記表中，縱列自下而上：

(第16表) 民國卅六年元月上海黃金與美鈔市價表

單位：黃金(恰赤)十市兩合國幣百萬元

美鈔一元合國幣千元

| 日期 | 黃金 | 美鈔 | 日期 | 黃金 | 美鈔 |
|---|---|---|---|---|---|
| 1 | 新年假日 | —— | 17 | 3.705 | 6.45 |
| 2 | 新年假日 | —— | 18 | 3.830 | 6.60 |
| 3 | 新年假日 | —— | 19 | 星期日 | —— |
| 4 | 3.570 | 6.50 | 20 | 4.010 | 7.10 |
| 5 | 星期日 | —— | 21 | 4.280 | 7.10 |
| 6 | 3.520 | 6.30 | 22 | 舊年假日 | —— |
| 7 | 3.575 | 6.35 | 23 | 舊年假日 | —— |
| 8 | 3.690 | 6.50 | 24 | 舊年假日 | —— |
| 9 | 3.785 | 6.95 | 25 | 舊年假日 | —— |
| 10 | 3.700 | 6.60 | 26 | 星期日 | —— |
| 11 | 3.653 | 6.50 | 27 | 3.910 | 6.90 |
| 12 | 星期日 | —— | 28 | 3.920 | 6.90 |
| 13 | 3.715 | 6.60 | 29 | 3.930 | 6.90 |
| 14 | 3.695 | 6.60 | 30 | 3.945 | 7.20 |
| 15 | 3.715 | 6.63 | 31 | 4.120 | 7.83 |
| 16 | 3.700 | 6.50 | | | |

（第17表）　上海黃金與美鈔市價複式登記表

ix：黃金十市兩價格組距（單位：國幣百萬元）

iy：美鈔一元價格組距（單位：國幣千元）

| iy＼ix | 3.500—3.700 | 3.700—3.900 | 3.900—4.100 | 4.100—4.300 |
|---|---|---|---|---|
| 7.50—7.20 | | | | 1 |
| 7.10—7.50 | | | 11 | 1 |
| 6.70—7.10 | | 1 | 111 | |
| 6.30—6.70 | 1 | | | |

（第18表）　上海黃金與美鈔市價相關表

x：黃金十市兩價格（單位國幣百萬元）

y：美鈔一元價格（單位國幣千元）

| iy＼ix | | | 3.500—3.700 | 3.700—3.900 | 3.900—4.100 | 4.100—4.300 | 總計 |
|---|---|---|---|---|---|---|---|
| | y＼x | | 3.6 | 3.8 | 4.0 | 4.2 | |
| | | fy＼fx | 6 | 7 | 5 | 2 | 20 |
| 7.50—7.90 | 7.70 | 1 | | | | 1 | |
| 7.10—7.50 | 7.3 | 3 | | | 2 | 1 | |
| 6.70—7.10 | 6.90 | 4 | | 1 | 3 | | |
| 6.30—6.70 | 6.50 | 12 | 6 | 6 | | | |
| 總計 | | 20 | | | | | |

# 附統計練習資料

第19表　240個大學男生身高與體重表

身高單位：公分

體重單位：公斤

| 學號 | 體重 | 身高 | 學號 | 體重 | 身高 |
|---|---|---|---|---|---|
| #1 | 55.0 | 169.6 | 41 | 48.8 | 167.5 |
| 2 | 45.7 | 162.0 | 42 | 85.2 | 174.3 |
| 3 | 41.6 | 163.3 | 43 | 42.9 | 161.8 |
| 4 | 55.2 | 173.8 | 44 | 50.5 | 164.5 |
| 5 | 53.6 | 169.9 | 45 | 51.6 | 169.2 |
| 6 | 56.4 | 179.6 | 46 | 50.5 | 172.2 |
| 7 | 52.3 | 165.6 | 47 | 45.7 | 157.5 |
| 8 | 55.5 | 166.0 | 48 | 49.1 | 163.9 |
| 9 | 55.5 | 169.9 | 49 | 50.0 | 167.7 |
| 10 | 52.9 | 166.6 | 50 | 45.2 | 163.4 |
| 11 | 53.4 | 163.0 | 51 | 54.5 | 172.8 |
| 12 | 55.9 | 164.3 | 52 | 49.3 | 170.4 |
| 13 | 55.0 | 165.8 | 53 | 47.5 | 160.2 |
| 14 | 48.6 | 16[illegible].0 | 54 | 53.4 | 167.1 |
| 15 | 69.1 | 159.0 | 55 | 72.3 | 167.4 |
| 16 | 60.5 | 171.0 | 56 | 47.5 | 164.5 |
| 17 | 47.3 | 163.3 | 57 | 52.0 | 179.9 |
| 18 | 47.7 | 166.8 | 58 | 43.2 | 153.7 |
| 19 | 77.7 | 171 5 | 59 | 56.4 | 155.7 |
| 20 | 48.2 | 163.5 | 60 | 46.8 | 167.1 |
| 21 | 54.7 | 171.3 | 61 | 39.7 | 159.1 |
| 22 | 49.5 | 167.0 | 62 | 54.7 | 172.6 |
| 23 | 39.7 | 162.4 | 63 | 54.5 | 170.0 |
| 24 | 56.1 | 170.4 | 64 | 55.0 | 177 2 |
| 25 | 44.1 | 158.3 | 65 | 59.5 | 169.7 |
| 26 | 57.0 | 165.2 | 66 | 46.8 | 176.6 |
| 27 | 49.1 | 164.3 | 67 | 54.3 | 164.3 |
| 28 | 44.5 | 154.2 | 68 | 46.6 | 163.1 |
| 29 | 57.0 | 174.4 | 69 | 45.2 | 161.3 |
| 30 | 59.1 | 168.3 | 70 | 48.2 | 167.6 |
| 31 | 44.7 | 167.3 | 71 | 49.5 | 165.8 |
| 32 | 41.8 | 157.2 | 72 | 56.6 | 167.7 |
| 33 | 51.4 | 161.3 | 73 | 52.7 | 172.4 |
| 34 | 45.5 | 163.3 | 74 | 45.2 | 154 3 |
| 35 | 47.3 | 172.4 | 75 | 43.2 | 160.5 |
| 36 | 48.0 | 162.4 | 76 | 48.8 | 163.5 |
| 37 | 63.6 | 179.2 | 77 | 47.5 | 166.3 |
| 38 | 62.7 | 183.5 | 78 | 50.5 | 177.6 |
| 39 | 55.0 | 172.4 | 79 | 52.0 | 167.4 |
| 40 | 40.9 | 154.1 | 80 | 59.3 | 172.6 |

| 學號 | 體重 | 身高 | 學號 | 體重 | 身高 |
|---|---|---|---|---|---|
| 81 | 50.5 | 161.4 | 121 | 51.4 | 166.1 |
| 82 | 61.4 | 171.2 | 122 | 55.9 | 169.6 |
| 83 | 49.5 | 165.6 | 123 | 63.6 | 172.6 |
| 84 | 95.5 | 172.6 | 124 | 55.2 | 169.2 |
| 85 | 54.1 | 171.0 | 125 | 59.1 | 168.3 |
| 86 | 41.8 | 169.9 | 126 | 53.6 | 174.0 |
| 87 | 44.8 | 163.2 | 127 | 54.7 | 167.7 |
| 88 | 43.2 | 167.4 | 128 | 55.5 | 174.9 |
| 89 | 45.5 | 165.8 | 129 | 60.9 | 176.4 |
| 90 | 53.2 | 169.9 | 130 | 52.7 | 172.8 |
| 91 | 63.4 | 177.1 | 131 | 53.2 | 165.9 |
| 92 | 50.0 | 155.4 | 132 | 62.0 | 178.3 |
| 93 | 46.8 | 165.8 | 133 | 57.3 | 130.2 |
| 94 | 49.5 | 164.0 | 134 | 46.8 | 162.6 |
| 95 | 49.1 | 159.9 | 135 | 54.1 | 167.6 |
| 96 | 49.5 | 168.3 | 136 | 55.9 | 167.7 |
| 97 | 49.5 | 167.3 | 137 | 56.6 | 167.3 |
| 98 | 56.8 | 171.4 | 138 | 55.0 | 165.5 |
| 99 | 53.2 | 169.0 | 139 | 52.3 | 166.7 |
| 100 | 53.2 | 167.0 | 140 | 64.1 | 165.6 |
| 101 | 51.4 | 166.0 | 141 | 50.0 | 170.9 |
| 102 | 50.5 | 168.6 | 142 | 77.0 | 173.7 |
| 103 | 50.7 | 177.0 | 143 | 57.7 | 172.4 |
| 104 | 50.7 | 167.3 | 144 | 50.0 | 170.5 |
| 105 | 48.2 | 164.0 | 145 | 44.1 | 166.6 |
| 106 | 58.2 | 166.0 | 146 | 51.1 | 165.1 |
| 107 | 45.5 | 164.9 | 147 | 54.5 | 172.6 |
| 108 | 50.9 | 170.2 | 148 | 50.7 | 164.5 |
| 109 | 53.2 | 167.0 | 149 | 52.3 | 176.8 |
| 110 | 58.2 | 168.5 | 150 | 57.0 | 155.6 |
| 111 | 50.0 | 172.0 | 151 | 58.2 | 171.8 |
| 112 | 51.8 | 169.6 | 152 | 53.4 | 167.7 |
| 113 | 50.9 | 166.4 | 153 | 46.4 | 169.2 |
| 114 | 52.0 | 163.4 | 154 | 41.8 | 190.4 |
| 115 | 48.2 | 169.2 | 155 | 41.4 | 160.5 |
| 116 | 52.3 | 171.4 | 156 | 52.3 | 169.4 |
| 117 | 43.8 | 167.4 | 157 | 45.0 | 168.8 |
| 118 | 46.6 | 171.6 | 158 | 48.6 | 169.3 |
| 119 | 54.1 | 173.4 | 159 | 56.0 | 171.5 |
| 120 | 46.8 | 164.3 | 160 | 49.1 | 165.1 |

| 學號 | 體重 | 身高 | 學號 | 體重 | 身高 |
|---|---|---|---|---|---|
| 161 | 46.8 | 162.8 | 201 | 48.4 | 154.1 |
| 162 | 50.2 | 164.4 | 202 | 52.0 | 153.0 |
| 163 | 45.0 | 166.0 | 203 | 48.4 | 154.2 |
| 164 | 57.3 | 164.0 | 204 | 53.2 | 166.8 |
| 165 | 56.4 | 167.0 | 205 | 57.7 | 170.7 |
| 166 | [illegible]9.5 | 165.9 | 206 | 57.0 | 174.1 |
| 167 | 55.7 | 158.4 | 207 | 45.9 | 163.9 |
| 168 | 47.7 | 166.1 | 208 | 40.9 | 153.8 |
| 169 | 45.5 | 156.2 | 209 | 52.5 | 162.1 |
| 170 | 45.9 | 165.6 | 210 | 54.7 | 169.1 |
| 171 | 67.3 | 184.2 | 211 | 46.1 | 162.5 |
| 172 | 53.8 | 1[illegible]1.7 | 212 | 53.6 | 165.8 |
| 173 | 45.5 | 162.1 | 213 | 46.8 | 164.7 |
| 174 | 44.6 | 166.6 | 214 | 42.3 | 158.7 |
| 175 | 47.3 | 161.1 | 215 | 52.3 | 166.8 |
| 176 | 45.5 | 164.9 | 216 | 59.5 | 159.1 |
| 177 | 60.7 | 166.3 | 217 | 68.4 | 172.3 |
| 178 | 50.5 | 168.3 | 218 | 58.2 | 172.1 |
| 179 | 51.5 | 169.4 | 219 | 48.2 | 159.2 |
| 180 | 56.1 | 172.0 | 220 | 49.5 | 178.5 |
| 181 | 50.9 | 172.2 | 221 | 57.3 | 175.1 |
| 182 | 48.2 | 168.6 | 222 | 60.7 | 167.8 |
| 183 | 59.5 | 172.0 | 223 | 53.6 | 160.2 |
| 184 | 54.5 | 167.4 | 224 | 49.3 | 163.7 |
| 185 | 52.0 | 171.8 | 225 | 50.9 | 171.5 |
| 186 | 59.8 | 161.8 | 226 | 49.1 | 172.5 |
| 187 | 64.5 | 171.8 | 227 | 58.4 | 166.5 |
| 188 | 52.0 | 174.7 | 228 | 53.2 | 167.0 |
| 189 | 62.3 | 176.4 | 229 | 53.4 | 171.0 |
| 190 | 56.1 | 1[illegible]6.6 | 230 | 55.9 | 169.2 |
| 191 | 55.5 | 167.8 | 231 | 60.9 | 171.6 |
| 192 | 42.5 | 159.[illegible] | 232 | 48.6 | 162.5 |
| 193 | 52.7 | 169.5 | 233 | 56.6 | 166.3 |
| 194 | 54.5 | 175.3 | 234 | 53.8 | 160.0 |
| 195 | 61.8 | 178.0 | 235 | 45.3 | 169.8 |
| 196 | 52.0 | 162.2 | 236 | 50.0 | 158.5 |
| 197 | 48.8 | 165.5 | 237 | 53.8 | 166.6 |
| 198 | 52.4 | 166.7 | 238 | 61.4 | 181.2 |
| 199 | 45.0 | 167.4 | 239 | 56.4 | 170.4 |
| 200 | 50.7 | 161.5 | 240 | 52.0 | 163.9 |

註: 轉錄自清華大學一九三七級年刊

## 習　題

1. 試由上列第19表中所載 200 名大學男生身高與體重資料中任擇一部份，製成 100 名大學男生身高與體重複式登記表。
2. 試由上題中所得之複式登記表製成 100 名大學男生身高與體重相關表。
3. 試由 (2) 題所得相關表製成 100 名大學男生身高次數分配表。
4. 試由 (2) 題所得相關表製成 100 名大學男生體重次數分配表。
5. 試由 (3) 題所得之次數表製向上與向下累積次數表各一。

# 第四章　統計圖形

## 第一節　品質圖

統計表中所記載者爲統計資料之全部數值，統計圖（Statistical Charts）則係利用點、線、面積或體積以表示統計資料中之要點，是以統計圖與統計表之相輔而行，欲窺資料之詳情，須讀統計表；欲見資料之梗概，則須閱統計圖，表中所示爲資料之本質，圖中所示則爲資料之縮影，統計圖可根據其形式，分爲綫圖，面積圖及體積圖，但統計圖之功用既在表現統計資料；圖形之分類最好根據統計數列與統計表格。換言之，即分爲品質圖（Categorical Charts），歷史圖（Historical Charts 與次數圖（Frequency Charts）三種，分別表示品質時間及次數三種數列。

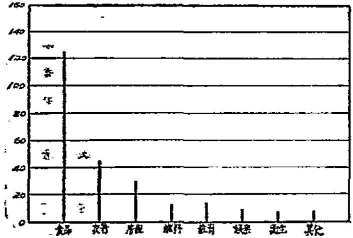
品質數列、係根據材料中各個項目具備之一種特質而分類，如家庭費用按其用途分類，製成一品質數列（第20表）將此數列用圖形表示之，則爲品質圖現分別舉例如下：

（第20表）某家庭每月平均費用表

| 類　　別 | 數　　額 |
|---|---|
| 食品 | $ 125,0 0 |
| 衣著 | 45,000 |
| 房租 | 30,000 |
| 燃料及燈火 | 12,500 |
| 敎育 | 13,750 |
| 娛樂 | 8 750 |
| 衛生 | 7,500 |
| 其他 | 7,500 |
| 總計 | $ 250,000 |

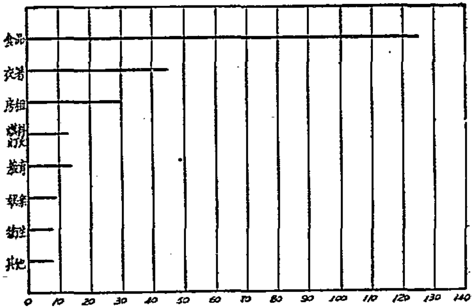
上列之品質圖可用不同之圖形以表示之:

一、 綫圖。



(第1圖) 家庭費用直綫圖(縱列式)

單位: 1,000元



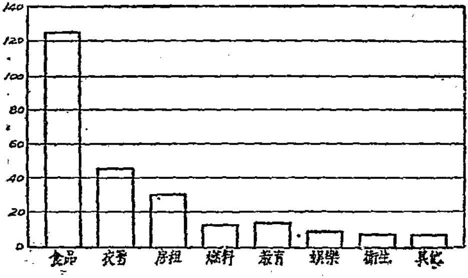
(第2圖) 家庭費用直綫圖(橫列式)

單位: 1,000元

所謂綫圖（Line Diagrams）係以一基綫爲共同出發點，作長短不同之直綫，用以代表各項目數值之大小。此種平行直綫。可有縱橫兩種排列方式：分別用上列第1，2兩圖表明之：

二、長條圖。

上列家庭費用資料，於綫圖之外，又可用并列長條形之高度以表示之，稱爲長條圖（Bar Diagrams）各長條仍自共同基綫開始，寬度相同，其長條形之長短或高低，即代表各數值之大小，亦可分縱列與橫列二種方式，其作法除以長條形代替直綫形以外，一切與直綫圖相同，現作一縱列之長條圖，以資說明：




（第3圖） 縱列式長條圖——家庭費用圖

單位：1,000元

上列第3圖係以單純之長條形，表示各項目數值之大小，但如每一項目中又可分爲幾部份，如同爲食品類，又可分爲米麵，菜蔬，肉類等，其他每一項目均可分爲幾部份。現將每一類費用用一長條形表示，此一長條形

又分爲數段落，分別代表各種食品，此種圖形，稱爲分段長條圖 (Component part Bar Diagrams)

三、圓形圖。

在下列第21表中仍以上列家庭費用爲例，以總費用$250,000作爲百分之百，其中每一類費用，俱以總費用爲共同分母而化爲百分數。總額在百分數中以100表示之，以圓形面積表示時，則爲360°是每百分之一，在圓周上佔3.6°，各部份之百分數既可求出，各部份所佔之度數，亦可求出。如食品一類費用爲$125,000，總費用爲$250,000，食品類費用之百分數爲50，食品類費用所佔之度數爲$50\times3.6^{\circ}=180^{\circ}$。其他各類費用所佔之度數，均可依同法求出。各類費用所佔之度數求出後，自可依之以繪出圓形圖(The Circular Diagram) 如下：

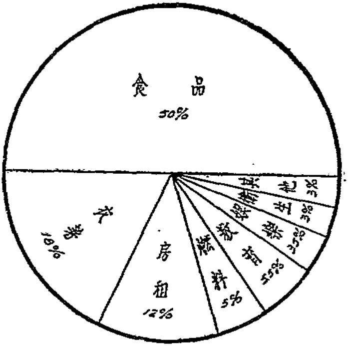
(第21表) 家庭費用表

| 類別 | 款額 | 百分數 | 圓周度數 |
|---|---|---|---|
| 食品 | $ 125,000 | 50.0 | 180.0 |
| 衣著 | 45,000 | 18.0 | 64.8 |
| 房租 | 30,000 | 12.0 | 43.2 |
| 燃料 | 12,500 | 5.0 | 18.0 |
| 教育 | 13,750 | 5.5 | 19.8 |
| 娛樂 | 8 750 | 3.5 | 12.6 |
| 衛生 | 7,500 | 3.0 | 10.8 |
| 其他 | 7,500 | 3.0 | 10.8 |
| 總計 | $ 250,000 | 100.0 | 360.0 |

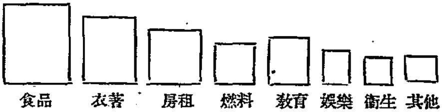
四、面積圖及體積圖。

上列之品質數列，亦可用面積圖(Area Diagrams)及體積圖(Volume Diagrams)以表示之。如以八正方形面積代表上列八種家庭費用，以面積

之大小代表費用之多少。




（第4圖）　家庭費用圓形圖




上列八正方形，係以圖形面積之大小，代表家庭費用之多少，面積圖之缺點，爲各面積缺乏共同基礎，無法彼此比較。由八圖形中殊難見出八種費用相對之情形。此種現象又可以八立方形或圓柱形表示之，但結果犯同一流弊。此外用兵士圖形之大小，表示各國陸軍之多少，以兵艦之大小

表示各國海軍之多少，以飛機之大小表示各國空軍之多少，雖均可予以生動概括之印象，但無法作精確之比較也。

五、 統計地圖。

如所示之資料爲地理數列 (Geographical Series)，則每在空白地圖中將此數列用符號表示之，稱爲統計地圖 (Statistical Maps)，依在地圖中所用之符號不同，統計地圖可分三種：

1. 加點地點 在空白地圖中加點 (Dots) 以表示現象之變異者，稱爲加點地圖 (Dotted Maps)，在地圖中加點之方法有二：一爲各點之大小相等，每點俱代表相同之數值，以點之多少代表統計事實變異之情形。如作交大同學籍貫分配圖，每點代表一人，甲省十人，卽在圖中作十點，乙省五人，在圖中作五點，丙省一人，在圖中作一點。二爲每省一點，各點之大小，卽代表統計事實變異之情形，如丙省有同學一人，在圖中作一小點，乙省有同學五人，在圖中作一點，其面積爲上列一點之五倍，甲省有同學十人，其點之面積爲小點之十倍。

2. 交叉線地圖 所謂交叉線地圖 (Cross-hatched Maps) 卽利用交叉線之形狀或密度以表示統計事實在地域上之變異。舉凡人口、氣候、雨量、各種農林礦產品等，無一不因地而異，此種地理變數固可用加點地圖表示，亦可藉交叉線之形狀或密度不同以表示之。

3. 顏色地圖 在普通地圖中以顏色之種類或深淺程度不同，以表示現象之變異，稱爲顏色地圖 (Colored Maps)。

## 第二節 歷史曲綫地圖

凡統計事實隨時間而變者稱爲時間數列或歷史數列，表示歷史數列之圖形，稱爲歷史圖(Historical Charts) 歷史數列每用曲線圖以表示之，

此種圖形係爲歷史曲線圖 (Historical Curves) 同爲歷史曲線圖，以作圖時所採取之標尺 (Scale) 不同因分爲三種:

算術圖。

算術圖 (Arithmetical Charts) 卽在圖中作x,y兩軸，x軸代表時間，不作原點，卽以第一年或第一月在x軸之最左端開始，依此由左而右，以代表時間單位之變動，y軸代表統計事實之變，依算術卽等差標尺作圖；現以美國人口資料爲例作圖卽明(第5圖)

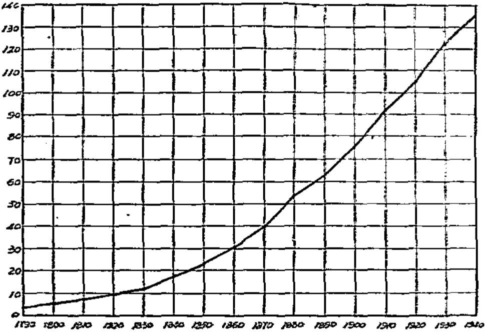
(第22表) 美國人口 (單位：百萬)

| 年份 | 人口 | 年份 | 人口 |
|---|---|---|---|
| 1790 | 3.9 | 1870 | 38.6 |
| 1800 | 5.3 | 1880 | 50.2 |
| 1810 | 7.2 | 1890 | 62.9 |
| 1820 | 9.9 | 1900 | 76.0 |
| 1830 | 12.9 | 1910 | 92.0 |
| 1840 | 17.1 | 1920 | 105.7 |
| 1850 | 23.2 | 1930 | 122.8 |
| 1860 | 31.4 | 1940 | 131.7 |

二、 百分數圖。

時間數列除依原始數值作曲線圖之外，亦可由時間中選擇其一爲基期 (Base Period)，使基期之數值爲百分之百，其他各時間單位之數值以基期數值爲共同分母化爲百分數，然後再以百分數作圖，x 仍代表時間，與上無異，y軸代表統計事實之變，與上所異者，過去 y 軸代表原始數值，或絕對數值，現在百分數圖中，(Percentage Charts)中，y軸代表百分數之變動，所成之曲線形狀與上圖中相同，但化爲百分數之特點，在可將數種資料作入同一歷史圖中，以比較彼此百分數不同之變動，現仍以美國

人口爲例 在第23表中使1790年之人口爲百分之百或100，其他各年人口俱爲1790年之倍數，如1790年爲100，1940年爲3377。



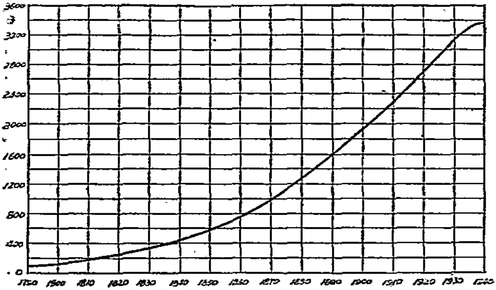

單位:百萬人

(第5圖) 美國人口算術圖

單位：百萬人

(第23表) 美國人口百分數變動表

| 年 份 | 人 口 | 百分數 | 年 份 | 人 口 | 百分數 |
|---|---|---|---|---|---|
| 1790 | 3.9 | 100 | 1870 | 38.6 | 990 |
| 1800 | 2.3 | 136 | 1880 | 50.2 | 1287 |
| 1810 | 7.2 | 185 | 1890 | 62.9 | 1613 |
| 1820 | 7.9 | 246 | 1900 | 76.0 | 1842 |
| 1830 | 12.9 | 331 | 1910 | 72.0 | 2359 |
| 18 0 | 17.1 | 438 | 1920 | 105.7 | 2710 |
| 1850 | 23.2 | 595 | 1930 | 122.8 | 3148 |
| 1860 | 31.4 | 805 | 1940 | 1 1.7 | 3377 |



（第6圖） 美國人口百分數圖

三、 對數圖。

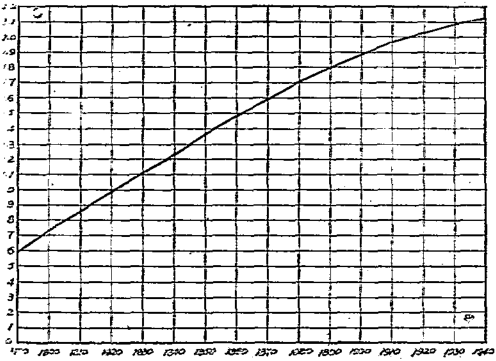
對數圖(Logarithmic Charts)卽x,y兩軸俱採用對數或等比標尺。但在歷史圖中，x軸係代表時間，時間之變動係依算術進度，故x軸仍採用算術或等差標尺，而y軸代表統計事實之變動，現以對數或等比標尺以表示之。因二軸之一採用對數標尺，故稱半對數圖(Semi-logarithmic Charts)亦稱比例圖(Ratio Charts)其作法有二：

1. 根據對數作圖 卽將時間數列中各年各月之原始數值查出其對數(見第24表)然後依對數作圖。

2. 根據原始數值作圖 上列根據對數作圖之缺點有二：一卽須查對數，如統計資料甚長，則查對數之工作甚大；二卽圖中所表出者爲對數，而非原始數值，資料之原來面目無法見出。補救之道，爲先製作半對數標尺，然後將資料原始數值作入半對數標尺中，作半對數標尺，只根據一至十之對數(見第25表)卽可，工作簡易。

（第24表）　美國人口原值及其 數表

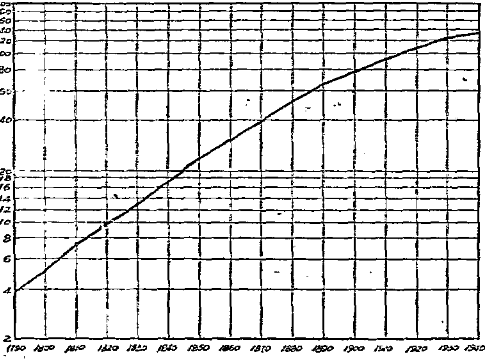
| 年　份 | 人　口 | 對　數 | 年　份 | 人　口 | 對　數 |
|---|---|---|---|---|---|
| 1790 | 3.9 | 0.591 | 1870 | 38.6 | 1.587 |
| 1800 | 5.3 | 0.724 | 1880 | 50.2 | 1.701 |
| 1810 | 7.2 | 0.857 | 1890 | 62.9 | 1.799 |
| 1820 | 9.6 | 0.982 | 1900 | 76.0 | 1.881 |
| 1830 | 12.6 | 1.111 | 1910 | 72.0 | 1.964 |
| 1840 | .7.1 | 1.233 | 1920 | 105.7 | 2.024 |
| 1850 | 23.2 | 1.365 | 1930 | 122.8 | 2.089 |
| 1860 | 31.4 | 1.497 | 1940 | 131.7 | 2.120 |



（第7圖）　美國人口年對數圖之一

（第25表） 一至十之對數表

| 數 | 對　數 |
|---|---|
| 1 | 0 |
| 2 | .301 |
| 3 | .477 |
| 4 | .802 |
| 5 | .699 |
| 6 | .778 |
| 7 | .845 |
| 8 | .903 |
| 9 | .954 |
| 10 | 1.000 |
| 100 | 2.000 |
| 1,00 | 3.000 |




（第8圖） 美國人口年對數圖之二

半對數圖在時間數列中至關重要，其原因有二：一爲各個變項數值懸殊之數列，可在半對數圖中表示，而在算術圖及百分數圖中，此種數列則無法表示，因最大項目與最小項目相差既大，在同一圖中實無法兼顧。如由絕對數值化爲相對數值或對數時，此種困難，完全解除，蓋十之對數爲1，百之對數爲二，千之對數爲三，萬之對數爲四，各變項之絕對值爲十，百，千，萬。實無法作圖。化爲對數則只爲一，二，三，四，作圖決無困難。二爲表示相對變動或變動率(Rate of Change)。在時間數列中，吾人所欲知者，非爲絕對變動，而爲相對變動，或變動速率，如人口及生產力每年增加百分之幾，物價每年每月上漲百分之幾，此種百分率或相對變動，至關重要。在半對數圖中，曲線坡度之高低，代表變動速率之大小，凡曲線之坡度高者，表示其增加之速率高；反之，凡曲線之坡度低者，表示其增加率低；凡曲線之坡度相同時，則表示其變動率或相對變動相同，故各時期相對變動之情形，由曲線之坡度即可見其梗概。

## 第三節 次數圖

次數表及累積次數表及相關表俱可利用圖形表示之，稱之爲次數圖(Frequency Diagrams)。現用上章中所得之五十男生高度分配表爲例，以說明次數圖及累積次數圖之作法，用黃金與美鈔市價相關表爲例，以說明分散圖 (ScatterDiagrams) 之作法。

一， 多邊形次數圖。

作x，y兩軸相交于原點，x軸代表統計資料，y軸代表次數。在橫軸x上標明各組距之中值，依此各中值豎立縱綫，以代表各組之次數，每二次數縱綫以直線聯接之，結果成爲多邊形次數圖 (Frequency Polygons)。最大最小兩中值以外各加一個中值，使其次數爲零，聯接之結果，使多邊形

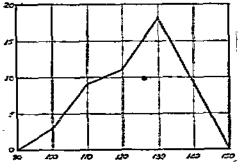
次數曲線之兩端，落於x軸上。

（第26表）　五十男生高度分配表

單位：吋

| 高度組距 | 中　值 | 次　數 |
|---|---|---|
| 95--105 | 100 | 3 |
| 105—115 | 110 | 9 |
| 115—125 | 120 | 11 |
| 125—135 | 130 | 18 |
| 135—145 | 140 | 9 |
| 總　計 | | 50 |

（第9圖）　五十男生高度多邊形次數圖

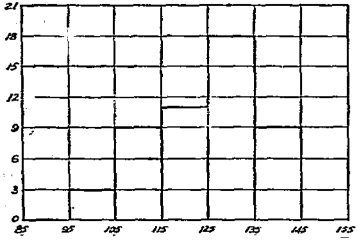


二，　柱形次數圖。

仍作x，y兩軸相交於原點。在x軸上標明資料各組之組限，y軸則代表次數，各組之組距既相等，是軸上每一組上下兩組限所包括之範圍亦相等，現依據此各組距向上豎立長方形，各長方形之寬度相等，其高度則代表各組次數之多少，此種次數圖，依其形式稱爲柱形圖（Column

Diagrams) 或稱爲直方圖。現仍依上表所載五十男生之高度分配爲例，而作柱形圖如下：

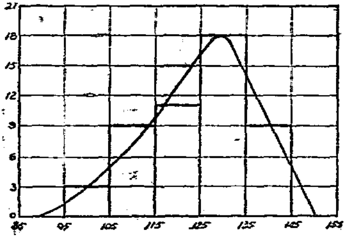
(第10圖)　五十男生柱形次數圖




三、　修勻次數曲線。

所謂修勻次數曲線(Smoothed Frequency Curves)。即根據上列之多邊形次數圖與柱形次數圖而去其稜角作成一平滑曲線，以代表各組次數。

(第11圖)



修匀之要點，在以平滑曲線代表不規則之多邊形曲線。但平滑曲線下之面積代表全部次數，與多邊形曲線及柱形圖之代表全部次數相同，是平滑曲線下之面積宜儘量與柱形圖之面積相等。由此可見修匀過程中包括有主觀之成分。至於修匀在理論上之根據，則因統計研究係根據選樣所得之局部資料。多邊形次數圖所以呈現棱角等不規則形狀者，以分類時所定之組距過大，組距數目過少，抽查之項目尚未充分，不能代表全部資料，如將組距縮小，組距之數目加多，並擴大樣本，增加抽查之項目，則次數曲線之棱角，將逐漸消失，最後之包括全部資料時，次數曲線之形狀爲平滑而無不規則之處，故修匀曲線係代表加大樣本或全部資料之理想次數曲線(Ideal Frequency Curves)

四、累積次數曲線圖

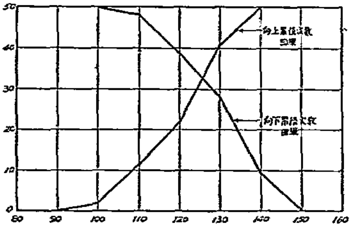
累積次數表，不論其爲向上累積或向下累積，均可製成曲線圖，稱爲累積次數圖(Ogive)。現以上列之五十男生高度分配爲例，製成累積次數表并將向上向下兩種累積次數作入同一圖形卽第12圖中：

(第27表) 五十男生高度分配累積次數表

單位：吋

| 組 距 | 中值 | 次數 | 向上累積次數 | 向下累積次數 |
|---|---|---|---|---|
| 95—105 | 100 | 3 | 3 | 50 |
| 105—115 | 110 | 9 | 12 | 47 |
| 115—125 | 120 | 11 | 23 | 38 |
| 125—135 | 130 | 18 | 41 | 27 |
| 135—145 | 140 | 9 | 50 | 9 |
| | | 50 | | |

（第12圖）　五十男生高度累積次數曲線圖




五、　分散圖。

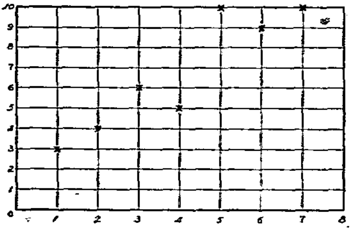
x, y 兩種資料間之相關，無論其為分組之相關表及未分組之原料俱可作成圖形，以表示相關之有無與正負。稱為分散圖(Scatter Diagrams)

1.　未分組資料

（第28表）　未分組相關資料

| x | y |
|---|---|
| 1 | 3 |
| 2 | 4 |
| 3 | 6 |
| 4 | 5 |
| 5 | 10 |
| 6 | 9 |
| 7 | 10 |

（第13圖） 分散圖



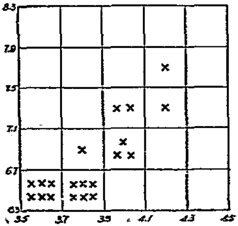

2. 分組資料

（第29表） 民國三十六年元月上海黃金美鈔市價相關表

x: 黃金十市兩價格（單位：國幣百萬元）

y: 美鈔一元價格 （單位：國幣千元）

| $ix$ \ $iy$ | | | 3.50—3.70 | 3.70—3.90 | 3.9 —4.10 | 4.10 —4.30 | |
|---|---|---|---|---|---|---|---|
| | x \ y | | 3.6 | 3.8 | 4.0 | 4.2 | |
| | | $fx$ \ $fy$ | 6 | 7 | 5 | 2 | 20 |
| 7.50—7.90 | 7.7 | 1 | | | | 1 | |
| 7.10—7.50 | 7.3 | 3 | | | 2 | 1 | |
| 6.70—7.10 | 6.9 | 4 | | 1 | 3 | | |
| 6.30—6.70 | 6.5 | 12 | 6 | 6 | | | |
| | | 20 | | | | | |




（第14圖）　分散圖

資料：三十六年元月，上海黃金與美鈔市價。

## 習　題

1. 試根據第一章習題中所作之100名大學男生身高分配繪一多邊形次數圖。
2. 試根據全上資料繪一柱形次數圖。
3. 試依上題所得之柱形次數圖繪一修勻次數曲線圖。
4. 試覓一機關支出分配資料，繪一圓形圖。
5. 試覓一等差變動之時間數列繪一算術圖。
6. 試覓一等比變動之時間數列繪一半對數圖。

# 第二編

## 次數分配之分析

# 第五章　集中趨勢量數

## 第一節　算術平均數

統計方法之程序，首在收集資料，繼爲整理資料。統計資料經過整理手續，由繁化簡，製成次數表與次數圖之後，初步工作告一段落。在次數表內，中央一組距次數最多，上下於此之各組距次數漸减，至最大最小兩極端組距之次數最少。此種現象稱之爲次數分配之集中趨勢（The Central Tendency）或簡稱爲集勢。集中之點容或不在中央，而向上向下稍偏，但其集中趨勢之存在，則初無不同。在次數曲線圖中，此種集中趨勢尤爲顯著。代表各組次數分配之曲線，在中央有一高峯，兩側逐漸下降，最後在左右兩極端，此曲線落於$X$軸上。此即表示次數愈在中央而愈集中；愈在兩端愈分散。集中趨勢之意義，即在次數分配所包括之各個項目，雖彼此大小不同，但其大小出入有一定之範圍，不能出於最大最小兩邊界範圍之外，即在此限度以內，其出入小之機會亦比出入大之機會爲多。以一百男生之身高或一百工人之工資而言，雖彼此可不同，但相差無多之機會多，而相差懸殊之機會少。次數分配有集中趨勢，吾人方能求平均數（Averages）以代表各個項目之值。故平均數可稱之爲集中量數（Measures of Central Tendency），或稱之爲集勢量數。最重要者約有下列五種：

1　算術平均數　　4　中位數

2　幾何平均數　　5　衆數

3　倒數平均數

一 未分組資料

算術平均數（The Arithmetic Mean）或簡稱之爲平均數（The Mean; The Average）即統計資料中各個項目之和,除以項目之數。如統計資料未經分類,而視爲一變數（Variable）以$X$表示之。其中所包括之項目,視爲變項（Variates）,以$X_1,X_2,X_3,\cdots X_n$表示之,各個變項之總和以$\Sigma X$表示之,變項之項數以$N$表示之,算術平均數以$\bar{X}$表示之,則算術平均數之公式爲:

$$\bar{X}=\frac{\Sigma X}{N}$$

如學生五人之年齡分別爲 20, 22, 23, 25, 28, 則五人之平均年齡依算術平均數公式求之:

$$\bar{X}=\frac{\Sigma X}{N}=\frac{20+22+23+25+28}{5}=\frac{118}{5}=23.6 \text{ 歲}$$

二 分組資料

1 普通法

但如統計資料之變項過多,先經分組製爲次數分配,再由次數分配以求算術平均數時,則變數$X$所代表者已非各個項目之原始數值,而爲各組之中值。各個項目之原有數值,不復存在,而各化爲次數,以所在組距之中值代表之,是每一組中,各變項之總和,俱爲$fX$,即項數與中值之乘積,而全部次數分配中所有變項之總和爲$\Sigma fX$,變項之總數則爲次數總數,以$\Sigma f$表示之,是由次數分配中,求算術平均數之公式爲:

$$\bar{X}=\frac{\Sigma fX}{\Sigma f}$$

依此公式以求算術平均數,稱之爲普通法（The Long Method）

須將次數表加以擴充，以便計算變項總和即各組中值與次數乘積之總和之用。在下列第30表中係根據五十男生之體重次數分配，以示範算術平均數計算之方法：

（第30表）五十男生體重算術平均數計算表

單位：磅

| $i$ | $X$ | $f$ | $fX$ |
|---|---|---|---|
| 95—105 | 100 | 3 | 300 |
| 105—115 | 110 | 9 | 990 |
| 115—125 | 120 | 11 | 1320 |
| 125—135 | 130 | 18 | 2340 |
| 135—145 | 140 | 9 | 1260 |
| 總　計 | | 50 | 6210 |

$$\bar{X}=\frac{\Sigma fX}{\Sigma f}=\frac{6210}{50}=124.2\text{磅}$$

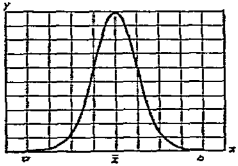
2　簡捷法之一——單位離差法

由次數分配中以求算術平均數，已較自未分組資料中計算時爲簡單。但根據各組距之中值及次數計算，仍嫌工作繁重，須另謀簡捷方法（The Short Method）。簡捷法自離差（Deviation）一觀念出發。離差可釋爲差異（Difference）或誤差（Error），但非錯誤（Mistake）。在天文學中以最精密之儀器測定一天體之位置，在物理學中以最精密之儀器測定一物體之長度或重量時，連續試驗一百次，而一百次所得之結果各不相同，是究以何一數值以代表此一百次試驗之結果，其解決之方法爲求算術平均數，算術平均數爲最合於機率之值（The Most Probable Value）亦稱之爲標準值（The Normal Value），爲一百次試驗所得結果之重心。此一百次結果俱稱之爲觀察值（Observed Values）。以平均數爲準則，

每次試驗所得之結果與平均數不同，此種不同稱之爲離差，或誤差，以符號表示之，$X$ 代表各變項，即各次試驗之結果或觀察值，$\bar{X}$ 代表平均數或標準值，$x$ 代表離差，則 $x=X-\bar{X}$，$x$ 即離差。其存在受下列三原則之支配，並可以圖表示之：

1. 小額離差多於大額離差。 2. 正負離差大致機會均等。

3. 極大離差根本少見。




上列三定律可藉次數曲線圖說明：$X$ 軸係代表各觀察值，其彼此不同，以 $a$，$b$ 爲範圍，$Y$ 軸代表次數之多少，$\bar{x}$ 代表算術平均數所估標之點。由此點所豎立之次數縱線最高，即表示觀察值與標準值完全脗合，根本無離差存在時次數最多。在此點以左及以右，曲線逐漸下落，即表示觀察值與標準值間離差小時次數多，而離差大時次數少；換言之，即小額離差多於大額離差。次以最高次數縱線爲準則，此線恰好平分曲線下之面積，左右兩半之面積完全相等，其構造又完全相同，即表示正負離差相等。最後次數曲線在最左最右 $a$，$b$ 兩點落於 $X$ 軸上，即表示在此兩點已無次數，或最大離差，無論正負均不存在。

以上三定律所代表之意義，無非各觀察值有集中之趨勢。此項誤差原理在次數分配之分析中至關重要。分類歸併之結果，由未分組資料化爲分

組資料.在次數分配中，以組中值代表一組之內各個變項，非各個變項之原值俱等於組中值，但以中值爲最合於機率之數值，以中值爲標準值時，各個變項雖有離差，但正負離差彼此抵消，以中值乘次數所得之總值，恰等於各個變項原值之總和。同理，在次數分配中，以算術平均數爲標準時，非各個變項俱等於平均數，但以正負離差機會均等，從中抵消關係，使平均數與全部次數之乘積恰等於各個變項原有數值之總和，正負離差均等抵消，亦即由算術平均數所計離差之總和等於零。因離差總和爲零，則離差平方之總和，方能爲最小二乘方，離差平方總和爲最小二乘方時，算術平均數方成爲最合於機率之數值。算術平均數離差之和爲零，可以符號表示之，即 $\Sigma fx=\Sigma f(X-\bar{X})=0$.此種情形可以公式證明之：

$X_1, X_2, X_3, \cdots\cdots, X_n$＝各組中值。

$f_1, f_2, f_3, \cdots\quad, f_n$＝各組次數

$x_1, x_2, x_3\cdots\cdots,, x_n$＝各組離差（由平均數計算）

$$
\begin{array}{ll}
x_1=X_1-\bar{X}; & f_1x_1=f_1(X_1-\bar{X}) \\
x_2=X_2-\bar{X}; & f_2x_2=f_2(X_2-\bar{X}) \\
x_3=X_3-\bar{X}; & f_3x_3=f_3(X_3-X) \\
\cdots\cdots\cdots\cdots; & \cdots\cdots\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots; & \cdots\cdots\cdots\cdots\cdots\cdots \\
\hline
 & \Sigma fx=\Sigma f(X-\bar{X})=0
\end{array}
$$

上列離差總和爲零，即 $\Sigma fx=0$，可證明如下：

(1) $\bar{X}=\dfrac{f_1X_1+f_2X_2+f_3X_3+\cdots\cdots+f_nX_n}{f_1+f_2+f_3+\cdots\cdots+f_n}$

(2) 上式等號前後各以 $(f_1+f_2+f_3++f_n)$ 乘之，

$$\bar{X}(f_1+f_2+f_3+\cdots+f_n)=f_1X_1+f_2X_2+f_3X_3+\cdots+f_nX_n$$

(3) 由上式等號前後各減平均數$\bar{X}$計$\Sigma f$次，

$$0=f_1(X_1-\bar{X})+f_2(X_2-\bar{X})+f_3(X_3-\bar{X})+\cdots+f_n(X_n-\bar{X})$$

$$\therefore\quad \Sigma f(X-\bar{X})=\Sigma fx=0$$

求算術平均數之簡捷法即根據此理而來。在次數分配中，先以一組中值爲原點 (Origin) 或假定平均數 (Assumed Mean)，以$\bar{Z}$表示之；假定平均數與眞正平均數 (True Mean) 未必符合，其間之差額稱爲改正數 (Correction) 以$c$表示之。三者之關係如下：

$$\bar{X}=\bar{Z}+c$$

$\bar{Z}$即假定平均數爲已知數，係一組之中值，故設法求出改正數$c$時，則眞正平均數$\bar{X}$即可得到。求改正數$c$之方法爲依假定平均數$\bar{Z}$計算離差。此種離差以$d$表示之，則$d=X-\bar{Z}$，開展之後，可得公 如下：

$$\begin{array}{ll}
f_1(X_1-\bar{Z})=f_1d_1; & f_1X_1-f\bar{Z}=f_1d_1 \\
f_2(X_2-\bar{Z})=f_2d_2; & f_2X_2-f\bar{Z}=f_2d_2 \\
f_3(X_3-\bar{Z})=f_3d_3; & f_3X_3-f\bar{Z}=f_3d_3 \\
\cdots\cdots\cdots\cdots; & \cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots; & \cdots\cdots\cdots\cdots \\
\hline
\Sigma f(X-\bar{Z})=\Sigma fd & \Sigma fX-\Sigma f\bar{Z}=\Sigma fd
\end{array}$$

$$\therefore\quad \Sigma fX=\Sigma f\bar{Z}+\Sigma fd$$

上式各項均以$\Sigma f$除之：

$$\frac{\Sigma fX}{\Sigma f}=\frac{\Sigma f\bar{Z}}{\Sigma f}+\frac{\Sigma fd}{\Sigma f}\quad\therefore\quad \bar{X}=\bar{Z}+\frac{\Sigma fd}{\Sigma f}$$

可見吾人所求之改正數$c$爲由假定平均數所得之平均離差(Average Deviation about the Assumed Mean)。已證明此點後，則根據簡法求平

均數可計算如下：

第31表 五十男生體重分配算術平均數簡捷法計算表

$\bar{Z}=120$

| $i$ | $X$ | $f$ | $d$ | $fd$ |
|---|---|---|---|---|
| 95—105 | 100 | 3 | —2ı | —60 |
| 105—115 | 110 | 9 | —10 | —90 |
| 115—125 | 120 | 11 | 0 | 0 |
| 125—1.5 | 1°0 | 18 | 10 | 180 |
| 1 5—145 | 140 | 9 | 20 | 180 |
| 總 計 | | 5ȷ | | 210 |

$$\bar{X}=\bar{Z}+\frac{\Sigma fd}{\Sigma f}=120+\frac{210}{50}=120+4.2=12ł.2\text{ 磅}$$

3. 簡捷法之二——組距離差法

上列之簡捷法，係根據各中值與假定平均數間之離差，稱爲單位離差法（Unit-deviation Method）。但各組距既完全相同 組距 $i$ 爲 10 磅，此一常數未計算之前可以提出，卽以 $i$ 除之。最後結果再以組距乘之，經過此種提出組距之程序 計算手緒，更爲簡化：

（第32表） 五十男生體重算術平均數簡捷法計算表

$\bar{Z}=120$

| $i$ | $X$ | $f$ | $d_i$ | $fd_i$ |
|---|---|---|---|---|
| 95—105 | 100 | 3 | —2 | —6 |
| 105—115 | 110 | 9 | —1 | —9 |
| 115—125 | 120 | 11 | 0 | 0 |
| 125—135 | 130 | 18 | 1 | 18 |
| 135—.45 | 140 | 9 | 2 | 18 |
| 總 計 | | 5ȷ | | 2 |

$$\bar{X}=\bar{Z}+\frac{\Sigma fd_i}{\Sigma f}(i)=1\ 0+\frac{21}{50}\times 10=120+4.2=124.2\text{ 磅}$$

## 第二節　幾何平均數

在一組統計資料中，先求各變項之連乘積，次開此連乘積之方根，所得卽爲幾何平均數(The Geometric Mean)。統計資料爲一變數，以 $X$ 表示之；其各變項以 $X_1, X_2, X_3, \cdots X_n$ 表示之。幾何平均數以 $G.M.$ 表示之。資料中如只包括2變項，卽 $X_1$ 與 $X_2$，則幾何平均數 $G.M. = \sqrt{X_1 \cdot X_2}$。如包括三變項，則 $G.M. = \sqrt[3]{X_1 \cdot X_2 \cdot X_3}$。如包括 $N$ 個變項時，則幾何平均數之公式爲 $G.M. = \sqrt[N]{X_1 \cdot X_2 \cdot X_3 \cdots\cdots X_n}$。

解此公式必根據對數 (Logarithms)，公式等號前後各取對數：

$$Log.\,G.M. = \frac{logX_1 + logX_2 + \cdots\cdots + logX_n}{N} = \frac{\Sigma logX}{N}$$

幾何平均數之對數旣爲 $\frac{\Sigma logX}{N}$，則幾何平均數自然爲 $\frac{\Sigma logX}{N}$ 之逆對數(Antilogarithm)。幾何平均數之公式，原甚複雜，但利用對數之結果，已將其過程化爲算術平均數之計算，上式中所謂幾何平均數之對數，實卽爲各變項對數之平均數。求得此對數平均數後再查得其逆對數，卽爲所求之幾何平均數。

以上係未分組資料求幾何平均數之方法，以手續簡單，除說明外，無待以實際數字計算證明。如資料爲次數分配，則 $X$ 所代表者爲各組中值，$f$ 所代表者爲各組次數，則幾何平均數 ($G.M.$) 及幾何平均數之對數 ($log\ G.M.$) 與此對數之逆對數 (Antilog of log G. M.) 之公式如下：

$$(1)\quad G.M. = \sqrt[\Sigma f]{X_1{}^{f_1} \cdot X_2{}^{f_2} \cdots\cdots X_n{}^{f_n}}$$

$$(2)\quad log\ G.M. = \frac{f_1 logX_1 + f_2 logX_2 + \cdots + f_n logX_n}{\Sigma f}$$

$$= \frac{\Sigma f logX}{\Sigma f}$$

(3) $G.M.=Antilog.\ of\ \frac{\Sigma f\log X}{\Sigma f}$

第33表 五十男生體重幾何平均數計算表

| i | X | f | logX | flogX |
|---|---|---|---|---|
| 95—105 | 100 | 3 | 2. | 6.0.03 |
| 105—115 | 110 | 9 | 2.0414 | 18.3726 |
| 115—125 | 120 | 11 | 2 0792 | 22.8712 |
| 125—135 | 130 | 18 | 2.1139 | 38.05 2 |
| 135—145 | 140 | 9 | 2.1461 | 19.3149 |
| 總 計 | | 50 | | 104.6389 |

$$\log G.M.=\frac{\Sigma f\log X}{\Sigma f}=\frac{104.6089}{50}=2.09218$$

$G.M.=2.09218$ 之逆對數，此數之值爲何，須根據對數表計算。由對數表中查得：

| 逆對數 | 對 數 | 對 數 |
|---|---|---|
| 123.7 | 2.09237 | |
| ? | | 2.09218 |
| — 123.6 | 2.09202 | 2.09202 |
| .1 | 35 | 16 |

$$35:16=.1:x$$

$$x=\frac{16\times0.1}{35}=\frac{1.6}{35}=0.046$$

∴ 幾何平均數 $G.M.=123.646$ 磅

## 第三節 倒數平均數

倒數平均數(The Harmonic Mean)之意義，即先求 $X_1, X_2\ X_3, \cdots\cdots$

$X_n$各變項之倒數 (Reciprocals) 以 $\frac{1}{X_1}, \frac{1}{X_2}, \cdots\cdots \frac{1}{X_n}$ 表示之。繼求各倒數之平均數，全部程序可以符號表示如下：

(1) 各變項：$X_1, X_2, \cdots\cdots X_n$

(2) 各變項之倒數：$\frac{1}{X_1}, \frac{1}{X_2}, \cdots\cdots \frac{1}{X_n}$

(3) 各變項倒數之平均數：$\frac{1}{H.M.} = \frac{\Sigma \frac{1}{X}}{N}$

(4) 倒數平均數：$H.M. = \frac{N}{\Sigma \frac{1}{X}}$

如只有兩變項，$X_1=80$，$X_2=85$，$\frac{1}{X_1}=0.01250$，$\frac{1}{X_2}=0.01177$，

$$\Sigma \frac{1}{X}=0.02427,\ N=2$$

$$H.M. = \frac{2}{0.02427} = 82.406$$

以上係由未分組資料求倒數平均數之方法，如依次數分配，則計算之程序如下：

(1) 各變項或組中值：$X_1, X_2, \cdots, X_n$

(2) 各變項之倒數：$\frac{1}{X_1}, \frac{1}{X_2}, \cdots, \frac{1}{X_n}$

(3) 各組次數：$f_1, f_2, \cdots\cdots, f_n$

(4) 各組次數與倒數乘積：$\frac{f_1}{X_1}, \frac{f_2}{X_2}, \cdots\cdots, \frac{f_n}{X_n}$

(5) 倒數平均數公式：

$$H.M. = \frac{\Sigma f}{\Sigma \frac{f}{X}}$$

（第34表）　五十男生體重倒數平均數計算表

| $i$ | $X$ | $f$ | $\frac{1}{X}$ | $\frac{f}{X}$ |
|---|---|---|---|---|
| 95—105 | 100 | 3 | .010000 | .030000 |
| 105—115 | 110 | 9 | .009091 | .081819 |
| 115—125 | 120 | 11 | .008333 | .091663 |
| 125—135 | 130 | 18 | .007692 | .138456 |
| 135—145 | 140 | 9 | .007143 | .064287 |
| | | 50 | | .406225 |

$$H.M.=\frac{\Sigma f}{\Sigma\frac{f}{X}}=\frac{50.000000}{0.406225}=123.084\text{ 磅}$$

## 第四節　中位數及其他分割數值

中位數（Median）即統計資料中各變項依其大小排成序列時所產生中央變項之值。可見中位數係由位置而決定者，稱爲位置平均數（Position Averages）與上三節中所述之算術、幾何及倒數三種平均數之根據各個變項所計算者，稱爲計算平均數（Computed Averages）不同。中位數在位置上所佔據之點，恰將全部變項等分爲兩半，故亦可稱分割值（Partition Value）。分割值中除中位數之外，尙有四分位數，十分位數，百分位數等，玆分述之：

一、中位數　在未分組資料中求中位數時，須將各變項排爲序列，居於序列中央之項目稱爲中項（The Middle Item）或稱中位項（The Median Item）。中項之值卽爲中位數，以$M$表示之。但中項及中位數之決定，於變項總計爲偶數或奇數又不相同。如序列中共計三項：80，90，100，則中項爲第二項，中位數$M$爲90；如序列中共計四項70，80，90，100，則中項平

分第二項與第三項之間，中位數可以兩中央變項之平均數充之，即中位數 $M=(80+90)\div 2=85$。但平分項目者不必85，此外81，82，以至88，89等，無一不平分四項爲兩半。

以上由未分組資料之序列中求中位數，所用爲觀察法(The Inspection Method)。由次數分配求中位數之方法則爲插補法(The Interpolation Method)。在次數分配中，中位數仍爲中央項目之值，此值在位置上所佔據之點，恰將各變項分爲兩半，換言之，即大於中位數及小於中位數者各佔次數之半。次數總計以 $\Sigma f$ 表示之，次數之半自以 $\frac{\Sigma f}{2}$ 表示之。在未分組資料中，中位數之位置尙因項目總數爲奇數或偶數而不同。在次數分配中，此種分別不復存在，不論次數總數爲奇數或偶數，其中位變項，或簡稱中項，概指次數總計之半即 $\frac{\Sigma f}{2}$ 所決定之變項而言，但中位數非即指此變項之中值而係指此中項之極限(Limit)而言。此理須藉累積次數表說明之，無論由向上累積次數表或由向下累積次數表所得之結果相同：

1 由向上累積次數表求中位數

(第35表) 五十男生體重向上累積次數表

| 組距 | 中值 | 次數 | 向上累積次數 |
|---|---|---|---|
| 95—105 | 100 | 3 | 3 |
| 105—115 | 110 | 9 | 12 |
| 115—125 | 120 | 11 | 21 |
| 125—135 | 130 | 18 | 41 |
| 135—145 | 140 | 9 | 50 |
| 總計 | | 50 | |

由上列第35表可知次數總計 $\Sigma f$ 爲50，而次數之半 $\frac{\Sigma f}{2}$ 爲25，換言之，中項爲第25項，此項之上限即爲中位數。第25項之上限所以爲中位數

吾人循組距之方向由小而大，自各組累積次數逐一觀察即明。計算中位數時係假定各組距中之次數平均分配(Uniform Distribution)。舉例言之，在第一組距95—105內，共計三項，其組距爲10磅，平均分配即將組距化爲三等分，每項各佔其三分之一，是第一項所佔之範圍爲$95-98\frac{1}{3}$，第二項所佔之範圍爲$98\frac{1}{3}-101\frac{2}{3}$；第三項所佔之範圍爲$101\frac{2}{3}-105$，換言之，即第三項之下限爲$101\frac{2}{3}$，上限爲105。同理，第十二變項之上限爲115，第二十三變項之上限爲125，第四十一變項之上限爲135，吾人所求者爲第二十五變項之上限，其值介乎125與135之間。決定此值所用者爲插補法：

第23項之上限……………………125

第25項之上限（中位數）………?

第41項之上限……………………135

第23項落在115—125一組中，自第24項起，即落在125—135一組中，此組距可稱爲中位數組距(Median Class)。在125—135一組中，組距$i$爲10磅。項數爲18，即18變項所佔之範圍爲125—135，假定此18項平均分配，各佔一等分，全組之範圍爲10磅時，每一變項所佔之範圍自爲全組範圍十八分之一，已知第二十三項之上限或第二十四項之下限爲125磅，第二十四項之上限或第二十五項之下限爲$(125+\frac{10}{18})=125.56$磅，第二十五項之上限或第二十六項下限爲$(125+\frac{10}{18}\times 2)=(125+1.11)=126.11$磅，此值實即中位數。因中位數之意義，即爲一點，恰分資料爲相等之兩半，換言之，即大於中位數及小於中位數者各佔一半。現126.11磅爲第二十五項之上限，亦自爲第二十六項之下限，可見此值恰將資料中之五十變項分爲相等之兩半。由上列分析，可得求中位數之公式如下：

$$M=L+\frac{\frac{\Sigma f}{2}-F}{f_m}i$$

$M$＝中位數

$L$＝中位數組之下限，現為125

$\Sigma f/2$＝次數總計之半數，現為50/2＝25

$F$＝中位數組以下各組之累積次數，現為23

$f_m$＝中位數組中之次數，現為18

$i$＝組距，現為10

$$\therefore M=125+\frac{25-23}{18}\times 10$$

$$=125+\frac{20}{18}=125+1.11=126.11\text{磅}$$

2 由向下累積次數表求中位數

(第36表 五十男生體重向下累積次數表

| 組距 | 中值 | 次數 | 向下累積次數 |
|---|---|---|---|
| 95—105 | 100 | 3 | 50 |
| 105—115 | 110 | 9 | 47 |
| 115—125 | 120 | 11 | 38 |
| 125—135 | 130 | 18 | 27 |
| 135—145 | 140 | 9 | 9 |
| 總計 | | 50 | |

根據向上累積次數表求中位數時，係求第二十五項之上限。現依向下累積次數表時，則係循組距之方向由大而小，自各組累積次數逐一觀察。仍假定每組中所包括之各變項平均分配，則第九項之下限為135磅，第二十七項下限為125磅。吾人所要求者為第二十五項之下限，此一下限即為中位數，因其值一方為第二十五項之下限，一方為第二十六項之上限，恰

爲五十項資料之等分線。此一數值仍依插補法求出，其方法之由來說明如下，並可由說明而得一公式：

第九項之下限……………………135

第二十五項之下限……………………？

第二十七項之下限……………………125

第九變項落在135—145一組中，自第十變項起即落入125—135一組中，其組距$i$等於10磅項數爲18，即18變項所佔之範圍爲125—135。假定平均分配，各佔一等分。全組之範圍爲10磅時，每一變項所佔之範圍自爲全組範圍十八分之一。已知第九項之下限或第十項之上限爲135磅，則第十項之下限或第十一項之上限自爲 $135-\frac{10}{18}=135-0.56=134.44$，第十一項之下限或第十二項之上限爲 $135-\frac{10}{18}\times 2=135-1.11=133.89$ 磅，以此推之，第二十五項之下限或第二十六項之上限爲 $135-\frac{10}{18}(25-9)=135-\frac{160}{18}=135-8.89=126.11$ 磅。此即爲所求之中位數。因此值恰爲分界線，一方爲第二十五項之下限，一方爲第二十六項之上限也。由此可得公式如下：

$$M=U-\frac{\frac{\Sigma f}{2}-F}{f_m}i$$

$M$＝中位數

$U$＝中位數組之上限，現爲135

$\Sigma f/2$＝項數總計之半數，次數總計爲50，其半數爲25

$f_m$＝中位數組中之次數，現爲18

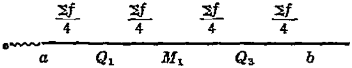
$i$＝組距，現爲10

$$\therefore\ M=135-\frac{25-9}{18}\times 10$$

$$=135-\frac{160}{18}=135-8.89=126.11\text{ 磅。}$$

二、下四分位數

以上所謂之中位數亦可稱之爲二分位數，因中位數爲分割值，恰將資料割爲兩等分。現將此義稍加引伸，即得四分位數(The Quartiles)，四分位數可分爲下四分位數(The Lower Quartile)及上四分位數(The Upper Quartile)。下四分位數爲分割值，全部變項之中，在此分割值以下者四分之一；在此值以上者四分之三，故亦稱爲第一四分位數(The First Quartile)，用$Q_1$表示之。上四分數與此相反，亦爲分割值，全部變項之中，分在此值以下者四分之三，在其上者四分之一。故亦稱爲上四分位數，以$Q_3$表示之。$Q_1$，$M_1$及$Q_3$三項分割值之關係可用圖形證明如下：




上圖所示，次數分配中所包括之變數，以$a$.$b$爲範圍，最小項之下限爲$a$，最大項之上限爲$b$，在$a$至$b$全部距離範圍之內，以下四分位數，中位數及上四分位數三分割值爲界線，恰將次數分爲四等分，使全部距離之四段中各包括四分之一次數。但此四段距離彼此則可相等亦可不相等，且十九不相等。所謂分爲四等分者爲全部變項或次數總計，絕非由$a$至$b$之全部距離之數值也。

下四分位數$Q_1$仍可由插補法，依向上與向下兩種累積次數表以求之：

1 由向上累積次數表求下四分位數

（第37表） 五十男生體重向上累積次數表

| 組 距 | 中 值 | 次 數 | 向上累積次數表 |
|---|---|---|---|
| 95—105 | 100 | 3 | 3 |
| 105—115 | 110 | 9 | 12 |
| 115—125 | 120 | 11 | 23 |
| 125—135 | 130 | 18 | 41 |
| 135—145 | 140 | 9 | 50 |
| 總 計 | | 50 | |

根據向上累積次數表求下四分位數，即係求全部次數四分之一所在變項之上限。在上列第37表中全部次數 $\Sigma f$ 為 50，其四分之一即 $\Sigma f/4=50/4=12.5$ 項，所求之下四分位數即第 12.5 變項之上限，第十二項之上限在 105—115 一組之內，四分位數必在 115—125 一組之內，此組稱為下四分位數組。關於用插補法求下四分位數之一切推理與求中位數相同，吾人可逕以公式求之：

$$Q_1=L+\frac{\frac{\Sigma f}{4}-F}{f_{Q1}}i$$

$Q_1$＝下四分位數

$L$＝下四分位數組之下限，現下四分位數組為 115—125，其下限為 115，即 $L=115$.

$\Sigma f/4$＝次數總額四分之一，現次數總額 $\Sigma f=50$，其四分之一為 12.5。

$F$＝四分位數組以下各組之累積次數，現為 12。

$f_{Q1}$＝下四分數組內之次數，現為 11。

$i$＝組距，現為 10。

$$\therefore\quad Q_1=115+\frac{12.5-12}{11}\times 10$$

$$=115+\frac{0.5}{11}\times 10=115+\frac{5}{11}=115+0.45=115.45\text{ 磅}$$

2. 由向下累積次數表求下四分位數

（第 8 表） 五十男生體重向下累積次數表

| 組　距 | 中　值 | 次　數 | 向下累積次數 |
|---|---|---|---|
| 95—105 | 100 | 3 | 50 |
| 105—115 | 110 | 9 | 47 |
| 115—125 | 120 | 11 | 38 |
| 125—135 | 130 | 18 | 27 |
| 135—145 | 140 | 9 | 9 |
| 總　計 | | 50 | |

根據向下累積次數表求下四分位數時，係求次數總數四分之三所決定變項之下限，現次數總額爲50。其四分之三爲37.5，即下四分位數在變項中之次序爲第37.5項，下四分位數實爲第37.5項之下限，因小于此值者佔全部次數四分之一，大于此值者佔全部次數四分之三。依插補法可得一公式如下：

$$Q_1=U-\frac{\frac{3\Sigma f}{4}-F}{f_{Q1}}i$$

$Q_1$＝下四分位數

$U$＝下四分位數組距之上限，現爲125磅

$\Sigma f$＝次數總額，現爲50項，其四分之三爲37.5項

$F$＝下四分位數組以上各組之累積次數，現爲27項

$f_{Q1}$＝下四分位數組中之次數，現爲11項

$i$＝組距，現爲10磅

$$Q_1 = 125 - \frac{37.5-27}{11} \times 10$$

$$= 125 - \frac{105}{11} = 125 - 9.55 = 115.45 \text{ 磅。}$$

三、上四分位數

下四分位數係一分割値，全部變項之中，小于此者佔四分之一，大于此者佔四分之三。上四分位數$Q_3$與此相反，即分割後，全部變項之中，小于此者計四分之三，大于此者計四分之一。由向上累積次數表求上四分位數$Q_3$時，係求次數總額四分之三即 $3\Sigma f/4$ 所定變項之上限；由向下累積次數表求上四分位數 $Q_3$ 時，則係求次數總額四分之一所定變項之下限。現分别依插補法公式求之：

1. 由向上累積次數表求上四分位數

（第 39 表） 五十男生體重向上累積次數表

| 組　距 | 中　値 | 次　數 | 向上累積次數 |
|---|---|---|---|
| 95—105 | 100 | 3 | 3 |
| 105—115 | 110 | 9 | 12 |
| 115—125 | 120 | 11 | 23 |
| 125—135 | 130 | 18 | 41 |
| 135—145 | 140 | 9 | 50 |
| 總　計 | | 50 | |

由上列第 39 表可知上四分位數項次爲 $3\Sigma f/4$，即第 37.5 項，此項包括在 125—135 一組中，故此組可稱爲上四分位數組（The Upper Quartile Class），其下限爲 125 磅，上限爲 135 磅。上四分位數公式：

$$Q_3 = L + \frac{\frac{3\Sigma f}{4} - F}{f_{Q3}} i$$

$Q_3$＝上四分位數

$L$＝上四分位數組距之下限，現爲125磅

$\frac{3\Sigma f}{4}$＝次數總額四分之三，現次數總額爲50項，其四分之三爲37.5項

$F$＝上四分數組以下各組累積次數，現爲23項

$f_{Q_3}$＝上四分數組內之次數，現爲18

$i$＝組距，現爲10磅

$$\therefore\quad Q_3=125+\frac{37.5-23}{18}\times 10$$

$$=125+\frac{145}{18}=125+8.06=133.06\text{ 磅}$$

2. 由向下累積次數表求上四分位數

(第40表) 五十男生體重向下累積次數表

| 組距 | 中値 | 次數 | 向下累積次數 |
|---|---|---|---|
| 95—105 | 100 | 3 | 50 |
| 105—115 | 110 | 9 | 47 |
| 115—125 | 120 | 11 | 38 |
| 125—135 | 130 | 18 | 27 |
| 1 5—145 | 140 | 9 | 9 |
| 總計 | | 50 | |

在向下累積次數表中，上四分位數$Q_3$爲第12.5項之下限，此點自在125－135一組中。此組爲上四分位數組距；其下限爲125磅，上限爲135磅。此組以下之累積次數爲9項，此組以內之次數爲18項，組距$i$爲10磅，可得上四分位數公式如下：

$$Q_3=U-\frac{\frac{\Sigma f}{4}-F}{f_{Q3}}i$$

$$=135-\frac{12.5-9}{18}\times 10$$

$$=135-\frac{35}{18}=135-1.94=133.06\text{磅}$$

由上文中已知：下四分位數 $Q_1$ 爲 115.45 磅，中位數 $M$ 爲 126.1 磅；上四分位數 $Q_3$ 爲 133.06 磅，三者之關係可用線圖表明之：

| $a$ | $Q_1$ | $M$ | $Q_3$ | $b$ |
|---|---|---|---|---|
| 95 | 115.45 | 126.11 | 133.06 | 145 |

上列線圖所表示者，即五十人體重之差異，以 95 磅至 145 磅爲範圍。在全距範圍之內，介乎 95 磅至 115.45 磅之間者計 12.5 項，介乎 115.45 磅至 126.11 磅之間者計 12.5 項；介乎 126.11 磅至 133.06 磅之間者計 12.5 項，最後介乎 133.06 磅至 145 磅之間者計 12.5 項。此三項分割值1求出後，吾人對於次數分配內部結構情形之認識，自比由單一平均數所認識者爲清晰確實。

四、十分位數與百分位數

由上文中可知中位數係將資料分割爲兩等分，四分位數係將資料分割爲四等分。由此推之，如將資料分割爲十等分，則須求出十分位數（The Deciles）；分爲百等分，則須求出百分位數（The Percentiles）。根據向上累積次數表或向下累積次數表計算均可，但須次數分配中所包括之次數較多時，計算十分位數及百分位數方有意義。茲列舉求十分位數之公式，至於百分位數之求法自可推出：

1 由向上累積次數表計算：

$$\text{第一十分位數}\,D_1=L+\frac{\frac{\Sigma f}{10}-F}{f_{D1}}$$

$$\text{第二十分位數}\,D_1=L+\frac{\frac{2\Sigma f}{10}-F}{f_{D2}}i$$

……………………………………

$$\text{第九十分位數}\,D_9=L+\frac{\frac{9\Sigma f}{10}-F}{f_{D9}}i$$

2 由向下累積次數表計算:

$$\text{第一十分位數}\,D_1=U-\frac{\frac{9\Sigma f}{10}-F}{f_{D1}}i$$

$$\text{第二十分位數}\,D_2=U-\frac{\frac{8\Sigma f}{10}-F}{f_{D2}}i$$

……………………………………

$$\text{第九十分位數}\,D_9=U-\frac{\frac{\Sigma f}{10}-F}{f_{D9}}$$

## 第五節 衆 數

所謂衆數 (Mode) 卽在一統計資料中，最常出現之變項或最常見之觀察值，如五十男生之年齡中，以廿二歲者最多，高度中以六十五吋者為最多，重量中以一百三十磅者爲最多。故衆數亦稱爲最普通值 (The Most Common Value)，最顯著值 (The Most Prominent Value)或最標準值 (The Most Typical Value) 亦係由位置决定，爲位置平均數，而非根據各個項目計算者。茲分述决定衆數之方法:

一 觀察法

在未分組之原料中將各個變項排成序列，其中最常見之變項即爲衆數。由次數分配中觀察衆數之法則，可依五十男生重量次數表爲例以明之:

(第41表) 五十男生體重分配表

單位：磅

| 組 距 | 中 值 | 次 數 |
|---|---|---|
| 95—105 | 100 | 3 |
| 105—115 | 110 | 9 |
| 115—125 | 120 | 11 |
| 125—135 | 130 | 18 |
| 135—145 | 140 | 9 |
| 總 計 | | 50 |

由上列第41表中可見125—135一組中次數最多，計18項，此組稱爲衆數組 (The Modal Class)，衆數組之中值，即爲衆數，或稱爲衆數值(The Modal Value)。由觀察法估定，粗而不精，因稱爲概約衆數 (The Crude Mode)或近似衆數(The Approximate Mode)。

二 移動平均數法

如次數分配中之次數有集中趨勢，在一組之中次數最多，自易于由觀察法決定衆數組及衆數。此種分配可稱爲單衆數分配(Mono-modal Distributions)。如次數分配雖亦具有集中趨勢，但次數最多者不只一組，而爲兩組，是爲雙衆數分配(Bi-modal Distributions)。此種次數分配中既包括有兩組次數均最多，甚至此兩組之次數彼此相等，自不能用觀察法確定衆數。但可用移動平均數法(Method of Moving Average)將此種困難解除，最後見出單一衆數組，並由此而確定衆數。現以下列分配爲例:

(第42表) 某機關薪級分配表

| 組距 | 中值 | 次數 | 兩組移動均數 |
|---|---|---|---|
| | | 0 | |
| | | | 9 |
| $100—150 | $125 | 18 | |
| | | | 24 |
| 150—200 | 175 | 30 | |
| | | | 28 |
| 200—250 | 225 | 26 | |
| | | | 20 |
| 250—300 | 275 | 14 | |
| | | | 12 |
| 300—350 | 325 | 10 | |
| | | | 15 |
| 350—400 | 375 | 20 | |
| | | | 22 |
| 400—450 | 425 | 24 | |
| | | | 27 |
| 450—500 | 475 | 30 | |
| | | | 22.5 |
| 500—550 | 525 | 15 | |
| | | | 16.5 |
| 550—600 | 575 | 18 | |
| | | | 11 |
| 600—650 | 625 | 4 | |
| | | | 2.5 |
| 650—700 | 675 | 1 | |
| | | | .5 |
| | | 0 | |
| 總計 | | 210 | 210 |

上列第42表為一標準薪級分配表，就中薪級低者有一組次數最多，薪級高者又一組次數最多，自無法由其原有次數中確定衆數組及衆數值。將每兩鄰組之次數加以平均，作為新次數，即現出175－225兩中值間次數最多，計28項，此兩中值所造成之間隔實為衆數組，此組之中值200元，實即衆數 由上可見所謂移動平均數法，實為觀察法之補充。原有次數中有兩峯情形，無法直接使用觀察，經過移動平均之後，此種困難已經解除，由觀察而確定衆數，與上無異。至求移動均數時，最上最下各加一組，假定其次數各為零。移動平均結果所得之次數總計，與原有總計相等。此種雙峯分配，雖可利用移動平均數取消一峯，最後求得衆數，但此項衆數之代表性並不特別顯著。

三　連續歸併法

上列薪級分配中，以兩組之中次數均最多，不能直接確定衆數，須先設法取消其中之一冢，除移動平均數外，又可用連續歸併法 (Method of Successive Groupings)，即將其組距逐漸擴大，將其次數逐漸歸併，最後再依歸併之結果，以確定衆數。現仍依上例說明之：

（第43表）　某機關薪級分配表

| 組距 $i=50$ | 原組次數 | 兩組歸併次數 | 移動一組次數 | 三組歸併次數 | 移動一組次數 |
|---|---|---|---|---|---|
| 100—150 | 18 | 48 | 略去18 | 74 | 略去18 |
| 150—200 | 30 | | 56 | | 70 |
| 200—250 | 26 | 40 | | | |
| 250—300 | 14 | | 24 | 44 | |
| 300—350 | 10 | 30 | | | 54 |
| 350—400 | 20 | | 44 | | |
| 400—450 | 24 | 54 | | 69 | |
| 450—500 | 30 | | 45 | | 53 |
| 500—550 | 15 | 23 | | | |
| 550—600 | 8 | | 12 | 13 | |
| 600—650 | 4 | 5 | | | 略去 4 |
| 506—700 | 1 | | 略去 1 | | 略去 1 |
| 總計 | 210 | 210 | 210 | 210 | 210 |

連續歸併所得之衆數組距：

1. 100—200　　2. 150—250

3. 100—250　　4. 150—300

原有分配中衆數組距有二，即 150—200 及 450—500，但經連續歸併後，可斷定 150—200 為唯一衆數組，蓋以擴大並移動組距將次數歸併所得四種結果中，無一不包括 150—200 一組，可斷定衆數組為150—200；衆數值為 175 元。

四　動差法

以上三法均在先確定衆數所在之組距，次以此組之中值爲衆數值。此種以衆數組距中值爲衆數之方法，係假定衆數組內之次數合于常態分配，次數向中值集中，其大于中值或小于中值之變項雖有離差，但以正負抵消，所有變項仍以中值爲代表值，即衆數值。如衆數組中之次數非常態分配，則中值只爲概約衆數，而眞正衆數(The True Mode)並不等於中值。試以五十男生體重分配爲例。

(第44表)　五十男生體重分配表

單位：磅

| 組　距 | 中　值 | 次　數 |
|---|---|---|
| 95—105 | 100 | 3 |
| 105—115 | 110 | 9 |
| 115—125 | 120 | $11=f_b$ |
| 125—135 | 130 | $18=f_{mo}$ |
| 135—145 | 140 | $9=f_a$ |
| 總　計 | | 50 |

由上表中可見衆數組爲125—135，此組之次數最多爲18項，以$f_{mo}$表示之，概約衆數爲此組距之中值，即130磅。但此組中之次數不合乎常態分配，可由其上下兩鄰組中之次數見出。衆數組以上之組距爲135—145，其次數爲9，以$f_a$代表示之，以下之組距爲115—125，其次數爲11，以$f_b$表示之，如$f_a$等於$f_b$吾人可進而斷定衆數組內之次數合乎常態分配，全組次數向中值集中，並以中值爲界限分爲兩半，此時之眞正衆數，自爲組中值。但現$f_a=9$，而$f_b=11$，可斷定其眞正衆數必小於衆數組距之中值，即小于130。因以$f_a$及$f_b$代表兩種離心力時，如二者平衡則衆數即爲組中值，現二者不平衡，$f_b$之力大于$f_a$，則衆數自亦傾向于$f_b$所據之方向，而小於衆數組之中值。由此可得依動差法(Method of mo-

ment-of-force）求眞正衆數之公式如下：

$$M_O = l + \frac{f_a}{f_a + f_b} i$$

$M_O$ ＝衆數

$l$＝衆數組距之下限，現爲 125 磅

$f_a$＝衆數組以上一組之次數，現爲 9

$f_b$＝衆數組以下一組之次數，現爲 11

$i$＝組距，現爲 10 磅

$$\therefore \quad M_O = 125 + \frac{9}{9+11} \times 10 = 125 + \frac{90}{20}$$

$$= 125 + 4.5 = 129.5 \text{ 磅}$$

五 皮爾森氏經驗法

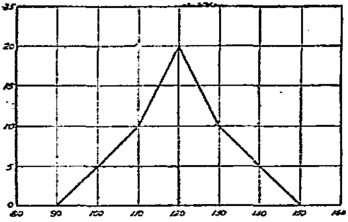
次數分配依其結構之形狀分之，不外對稱分配(Symmetrical Distribution)，左偏分配(Left-skewed Distribution)及右偏分配(Right-skewed Distribution)三種。無論在何種分配中，算術平均數，中位數與衆數三種集中量數間均有一定之關係。英統計學家皮爾森氏(Karl Pearson)根據此種關係求出衆數，稱之皮爾森經驗法，現以五十男生體重次數表爲例，將對稱，左偏及右偏三種分配分別說明之：

1. 對稱分配

（第45表） 五十男生體重分配表

| 組距 | 中值 | 次數 |
|---|---|---|
| 95—105 | 100 | 5 |
| 105—115 | 110 | 10 |
| 115—125 | 120 | 20 |
| 125—135 | 130 | 10 |
| 135—145 | 140 | 5 |
| 總計 | | 50 |

第1圖 五十男生體重次數圖




$\bar{X}=M=M_o=120$

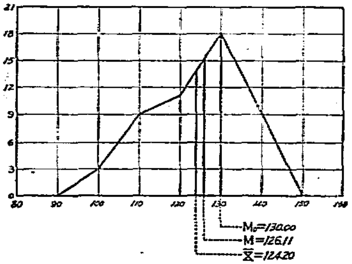
由上列第16圖中，可見其次數最多之組距即衆數組距爲115—125，衆數即爲此一組距之中值，爲120磅，但此值又恰爲全部次數分配之重心，故算術平均數亦爲120磅。最後由120磅估據一點所豎立之最高縱綫(The Maximum Ordinate)，恰平分次數曲綫下之面積，即此點左右兩邊各包括次數之一半，故中位數亦爲120磅。總之，在對稱分配中，平均數，中位數及衆數三種集勢量數相重合于一點，現爲120磅。

2. 左偏分配

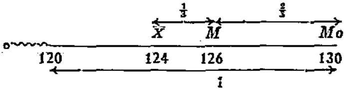
(第46表) 五十男生體重分配表

| 組　距 | 中　值 | 次　數 |
|---|---|---|
| 95—105 | 100 | 3 |
| 105—115 | 110 | 9 |
| 115—125 | 120 | 11 |
| 125—135 | 130 | 18 |
| 135—145 | 140 | 9 |
| 總計 | | 50 |

第17表　五十男生體重次數圖




由上圖可見次數曲綫向左偏傾時，爲左偏分配，至其所以左偏，則由于最高縱綫以下各組中次數較稀，以上各組次數較密。在此分配中，已知算術平均數$\bar{X}$爲124.20磅，中位數爲126.11磅，概約衆數爲130磅。以上三種集勢量數，如均去其小數，取其整數時，則平均數$\bar{X}$爲124，中位數$M$爲126，衆數$Mo$爲130，三者間之關係，可用下列直綫以表明之：




由上圖可知：如自平均數$\bar{X}$至衆數$Mo$之距離爲一時，則中位數介乎兩者之間，佔此距離三分之一，皮爾森氏根據此種關係，推出下列公式：

(1)　由原點$o$至衆數$Mo$之距離減掉由原點$o$至平均數$X$之距離等于平均數$\bar{X}$至衆數$Mo$之距離，以符表示之，$\bar{X}$至$Mo$間之距離$=Mo-\bar{X}$。

(2) 由原點 $o$ 至中位數 $M$ 之距離等于由原點 $o$ 至平均數 $\bar{X}$ 之距離另加由平均數 $\bar{X}$ 至衆數 $Mo$ 間距離三分之一，以符號表示之：

$$M=\bar{X}+\frac{Mo-\bar{X}}{3}$$

(3) $3M=3\bar{X}+Mo-\bar{X}$

$Mo=\bar{X}+3M-3\bar{X}$

$\therefore\quad Mo=\bar{X}-3(\bar{X}-M)$

已知 $\bar{X}=124.20$, $M=126.11$

$\therefore\quad Mo=124.20-3(124.20-126.11)$

$=124.20-3(-1.91)=124.20+5.73=129.93$ 磅。

3. 右偏分配

(第47表) 五十男生體重分配表

| 組　距 | 中　值 | 次　數 | $fX$ | 向上累積次數 |
|---|---|---|---|---|
| 95—105 | 100 | 9 | 900 | 9 |
| 105—115 | 110 | 18 | 1980 | 27 |
| 115—125 | 120 | 11 | 1320 | 38 |
| 125—135 | 130 | 9 | 1170 | 47 |
| 135—145 | 140 | 3 | 420 | 50 |
| 總　計 | | 50 | 5790 | |

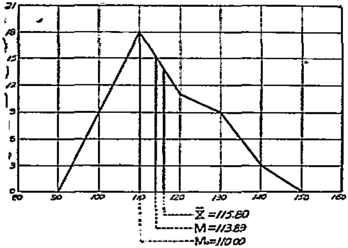
$$\bar{X}=\frac{\Sigma fX}{\Sigma f}=\frac{5790}{50}=115.8\text{ 磅}$$

$$M=L+\frac{\frac{\Sigma f}{2}-F}{f_m}i=105+\frac{25-9}{18}\times 10$$

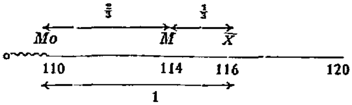
$$=105+\frac{160}{18}=105+8.89=113.89\text{ 磅}$$

概約衆數 $Mo=110$ 磅

第18圖 五十男生體重次數圖




上列爲一右偏分配。平均數 $\bar{X}$，中位數 $M$，及衆數 $Mo$ 已求出，並標入次數圖中，仍去其小數，" 取整數，則 $\bar{X}$ 等於116，$M$ 等于 111，$Mo$ 等于 110，三者之關係亦可用直綫表明之：




在右偏分配中，$\bar{X}$，$M$，$Mo$，三種集勢量數之位置雖與左偏分配中相反，但中位數 $M$ 所佔之據點仍爲由 $\bar{X}$ 至 $Mo$ 全距三分之一，由此仍可得三式如下：

（1） 由原點 $o$ 至平均數 $\bar{X}$ 之距離減掉由原點 $o$ 至衆數 $Mo$ 之距離，等于由衆數 $Mo$ 至平均數 $\bar{X}$ 之距離，以符號表之，$Mo$ 至 $\bar{X}$ 間之距離 $=\bar{X}-Mo$。

(2) 由原點 $o$ 至中位數 $M$ 之距離等于由原點 $o$ 至平均數 $\overline{X}$ 之距離減掉由衆數 $Mo$ 至平均數 $\overline{X}$ 間距離三分之一，以符號表之：

$$M=\overline{X}-\frac{\overline{X}-Mo}{3}$$

(3) $3M=3\overline{X}-\overline{X}+Mo$

$$Mo=\overline{X}-3\overline{X}+3M.$$

$$\therefore\ Mo=\overline{X}-3(\overline{X}-M)$$

已知 $\overline{X}=115.80,\quad M=113.89$

$$\therefore\ Mo=115.80-3(115.80-113.89)$$

$$=115.80-3\times 1.91$$

$$=115.80-5.73=110.07\text{ 磅}$$

## 第六節 平均數之特點

本章所述集中趨勢量數共分五種，即算術平均數、幾何平均數、倒數平均數、中位數及衆數。前三者係根據統計資料中各個變項算出，稱爲計算平均數(Computed Averages)，後二者則由統計資料中一部份變項推出，稱爲位置平均數 (Position Averages)。兩類或五種平均數中，以何者爲優，何者爲劣，未可一概而論。一優良平均數，究應具備何種條件或特點，英統計學家尤爾(Yule, G. U.)美統計學家克拉姆(Crum, William L.)及德氏(Day, Edmund E.)等均有所論述。綜合言之，可分下列各點：

一， 平均數之數值應確定無疑。其意即在任何一組統計資料中，平均數之值均可求出，而無模糊難以確定之虞。此項特點每稱之爲明確之嚴格性(Rigidity of Definition)。三種計算平均數均合于此項標準，在任何一種統計數列中，無論求算術、幾何或倒數平均數，俱可獲得確定不移之結果，

毫無問題。位置平均數之情形則不同。以中位數而論，在未分組資料中，如變項之項數爲奇數，則中位數之決定尚無困難。以甲乙丙三工人之工資爲例，三人每日工資分別爲 100 元，120 元，140 元，則中位數自爲 120 元。但如工人爲甲乙丙丁四人，每日工資分別爲 100 元，120 元，140 元，160 元，則中位數須求兩中項之平均數，即中位數之值爲 130 元。但所謂中位數係將所有變項劃爲兩等分之分割值，而劃爲兩等分者不必一定爲兩中項之平均數。換言之，130 元固將四工人之工資劃爲兩等分，此外 125 元或 135 元，亦均將各變項劃爲兩等分。任一數值，只要比第一中項爲大，比第二中項爲小時，俱合乎中位數之條件，不必一定爲兩中項之折中數也。中位數之值已不易確定，衆數之情形尚不及此。因衆數在未分組資料中爲最常見之數值，在分組資料中爲次數最多之組中值，如統計資料缺乏集中趨勢時，根本無任何常見之變項或次數密集之組距，則無法求概約衆數。

二、平均數之計算應根據所有變項，而不只根據一部份變項。平均數旣係代表全部資料中所有變項之數值，其計算自須根據一切變項，使各個變項，不拘大小，俱在平均數中表現其重要性。依此標準，則計算平均數較位置平均數爲優。因計算平均數顧名思義，係根據一切變項計算而來，而位置平均數則只依一部份變項推出。但此一特點亦有時成爲位置平均數之優點者：第一，如次數分配具有空端(Open Ends)，則只能求位置平均數，而無法求計算平均數。試以下列之空端分配爲例：

下列 100 學生某科成績分配中，最上最下兩組距均爲開端。中間各組距雖上下組限俱全，可以確定組中值，但最上一組之下限爲 90 而上限則未標明，90以上不必爲 100，是組中值絕不能即認定係 95。同理最下一組之上限雖爲40，但下限並未標明，40以下，可爲30 亦可爲20，10，0不等。下限既不定，此組之中值自無法確定。各組中值既不完全，當然無法計算算

(第48表) 100學生某科成績次數分配表

| 組距 | 人數 |
|---|---|
| 40 以下 | 5 |
| 40—50 | 10 |
| 50—60 | 25 |
| 60—70 | 35 |
| 70—80 | 15 |
| 80—90 | 8 |
| 90 以上 | 2 |
| 總計 | 100 |

術、幾何、倒數三種平均數，但概約衆數一望即知其為 65，中位數亦不難由向上或向下累積次數表中求出。

三、平均數之值應富于穩定性(Stability)，換言之，即平均數之值，不宜受樣本中少數項目變動之影響。以此標準而論，則計算平均數似不如位置平均數。因計算平均數中算術平均數之值，不論最大或最小極項(Extreme Variates)，均可影響。換言之，次數分配中有少數極大變項，算術平均數即隨之偏大，有少數極小變項，算術平均數即隨之偏小。至於幾何及倒數兩種平均數則均受最小變項之影響。無論受最大變項之影響，使平均數之值偏大，或受最小變項之影響，使其值偏小，俱使平均數之價值減低。反之，位置平均數中，衆數既為次數分配中次數最密集一組之中值，則少數最大最小項目之變動，自絲毫不發生影響。中位數于項目變更後是否受影響，須視位置是否變動而定，如最大最小之變項同等增加，于中位數毫無影響；如只最大或最小一種變項變動，則雖對中位數稍有影響，實微而不著。因少數變動之加減，對于位置影響至少。

四、平均數應能用代數法處理。換言之，即其公式可以用代數方法證明，

並可作進一步之分析。因平均數之計算僅爲統計分析工作之開始，將來各種離中趨勢及相關等量數之計算，無一不以平均數爲出發點，是以平均數與將來統計分析之工作不可分。統計分析既爲算學分析，則平均數之公式必爲確定之算學公式，方足以爲統計分析之出發點。依此而論，則位置平均數不如計算平均數。後者中尤以算術平均數之易于以代數方法處理，爲統計分析中最重要之量數。

五、平均數應易于計算並易于了解。平均數既爲統計分析之初步，其計算簡易，自極重要。一種平均數在學理上之根據無論如何穩固，但如計算之手續繁難，其重要性不免爲之減殺。依此而論，算術平均數爲最優者，在理論上算術平均數爲全部資料之重心，在計算之手續上，亦至爲簡易。至于幾何平均數須利用對數，倒數平均數須利用倒數，縱使有表可查，但項目太多時，仍感手續冗繁。至于中位數及衆數手續雖簡，但學理上之根據不如計算平均數之穩固。

最後，平均數係計算之結果，發表給讀者閱覽者。讀者未必人人有算學訓練，是必以平均數之意義明顯，公式簡易方能爲讀者所了解。以此而論，計算平均數中之算術平均數及位置平均數中之衆數，均最易爲人了解；而幾何平均數，倒數平均數及中位數三種比較難于理解。總之，五種中之中，似以算術平均數爲最優，但有時以目的特殊，算術平均數，亦非十全十美者。

## 習　題

1. 試根據第三章習題中所作之100名大學男生體重分配表，求算術平均數：

　(1) 用普通法

（2） 用假定原點組距離差簡捷法

2. 根據同一資料求幾何平均數

3. 根據同一資料求倒數平均數

4. 根據同一資料求下列各分割值：

（1） 中位數

（2） 下四分位數

（3） 上四分位數

5. 根據下列各法同上各題資料求乘數：

（1） 觀察法

（2） 動差法

（3） 皮爾森經驗公

# 第六章　離中趨勢量數

## 第一節　絕對離勢量數

統計資料之第一特點爲集中趨勢，因求各種集勢量數即平均數以代表之。其第二特點爲離中趨勢（Dispersion）。所謂離中趨勢，即各個變項與平均數間之差異，此種差異亦可求出適當數值以代表之，稱爲離勢量數（Measures of Dispersion）。離勢量數之大小與平均數之是否可靠，關係至大。離勢量數愈小，即各個變項與平均數間之差異愈小，則平均數代表各個變項之程度愈高，最後限度至離勢量數爲零，即每一變項俱與平均數之值相同，毫無離差，則平均數可靠之程度達於極點。反之，離勢量數愈大，即各個變項與平均數間之差異愈大，則平均數代表各個變項之程度愈低或平均數可靠之程度愈淺。故離勢量數直接代表離中趨勢之大小，間接代表平均數可靠性之深淺。離中量數可分爲絕對與相對兩類，絕對離勢量數（Absolute Measures of Dispersion）係其量數之單位與統計資料之單位同一，而相對離勢量數（Relative Measures of Dispersion）則指將絕對量數之單位，設法消除，使成爲抽象之小數或百分數。茲先述各種絕對離勢量數：

一、全距

離勢量數中最簡易者爲全距（Range），以 $R$ 表示之，即各個變項中，最大變項減去最小變項所剩餘之值。此值所代表者爲全部資料中各個變項相差之範圍。統計資料如未經分類，則最大與最小變項可由原料中指出，二者相減所得之值，即爲全距。如五十男生重量中，最重者爲144磅，最

輕者爲 95 磅，則全距爲 144 磅減 95 磅，等於 49 磅，即爲五十人體重出入之範圍。在次數分配中，全距則由全分配之上限與下限而定，以符號表示：

$$R=U-L$$

仍以五十人之體重爲例，其上限 $U$ 爲 145 磅，下限 $L$ 爲 95 磅，則全距 $R=145-95=50$ 磅

由上文可見全距 $R$ 之意義顯明，求法簡單。但其缺點亦顯而易見，由公式 $R=U-L$，可見全距只由最大最小兩變項確定，非根據各個變項計算，是上限或下限一變，全距即隨之而變。反之，在上下限以內之全部分配結構，無論如何變更，俱於全距毫無影響，凡上下限相同之次數分配，其全距即同，而內部結構則可完全不同。

二、四分差

四分差 (Quartile Deviation) 簡號爲 $Q$ $D.$，仍由全距變化而來，全距之計算，係根據上下兩極限或最大最小兩變項，四分差之計算則根據上下兩四分位數。全部資料或次數以 $\Sigma f$ 代表之，由 $Q_1$，$M$，及 $Q_3$ 即下四分位數、中位數、及上四分位數三種分割值數化爲四等分，就中在 $Q_1$ 即下四分位數及 $Q_3$ 即上四分位數兩分割值範圍以內者，恰佔全部變項之半，即 $\Sigma f/2$。此一半變項比較合乎常態，其離中趨勢較小。反之，在上下兩四分位數以外之一半變項，則爲最大與最小之極端變項，其離中趨勢較大，現以上下兩四分位數爲標準，位於此兩分割值以外之一半變項，置之不顧，而只求介乎 $Q_1$ 與 $Q_3$ 兩分割值以內半部變項中最大項目與最小項目差異之範圍，稱爲內四分位距 (Interquartile Range) 以 $Q.R.$ 表示之，其公式爲：

$$Q.R.=Q_3-Q_1$$

以 2 除之，稱爲半內四分位距 (Semi-interquartile Range)，或稱爲

四分差 (Quartile Deviation)，以 $Q.D.$ 表示之，四分差之公式如下：

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

以上章中所用五十男生體重資料爲例，已知 $Q_1=115.45$ 磅，$Q_3=133.06$ 磅，依此兩四分位數可得四分差如下：

$$Q.D.=\frac{Q_3-Q_1}{2}=\frac{133.06-115.45}{2}=\frac{17.61}{2}=8.805\text{ 磅}$$

四分差 $Q.D.$ 之意，卽五十人體重中，在上下兩四分位數範圍以內二十五人體重之全距爲 17.61 磅，其半距卽四分差爲 8.805 磅，實爲一二·五人體重差異之範圍也，可見四分差與全距相同，均爲位置量數(Position Measures)。所異者卽全距由次數分配之上限與下限決定，而四分差則由次數分配上下四分位數決定，其他變項一槪未曾顧及。舉凡次數分配之上下限相同者，其全距必同。同理，舉凡次數分配之上下四分位數相同者，其四分差亦必同。實則無論介乎上下兩四分位數之間，或處於兩者以外之次數分配可完全不同，故四分差之不足以代表次數分配之離中趨勢也，與全距相同。

三、平均差

所謂平均差 (Mean Deviation) 卽以算術平均數 $\bar{X}$ 或中位數 $M$ 爲準，先計算各變項之離差，次求各離差之平均數。計算各變項離差時，槪求絕對值，而不計其正負號，因如計算正負號則各個變項與算術平均數間之離差總和爲零，根本無法求平均差也。平均差之符號爲 $M.D.$，現說明計算方法：

1. 未分組資料

(第49表) 平均差計算表

| $X$ | $\|x\|=\|X-\bar{X}\|$ |
|---|---|
| 1 | 3 |
| 2 | 2 |
| 3 | 1 |
| 4 | 0 |
| 5 | 1 |
| 6 | 2 |
| 7 | 3 |
| 28 | 12 |

$$X=\frac{\Sigma X}{N}=28/7=4$$

$$M.D.=\frac{\Sigma|x|}{N}=12/7=1.714$$

2. 次數分配

(第50表 五十男生體重平均差計算表($\bar{X}=124.2$磅)

| $i$ | $X$ | $f$ | $\|x\|=X-\bar{X}$ | $f\|x\|$ |
|---|---|---|---|---|
| 95—105 | 100 | 3 | 24.2 | 72.6 |
| 105—115 | 110 | 9 | 14.2 | 127.8 |
| 115—125 | 120 | 11 | 4.2 | 46.2 |
| 125—135 | 130 | 18 | 5.8 | 104.4 |
| 135—145 | 140 | 9 | 15.8 | 142.2 |
| 總計 | | 50 | | 493.2 |

$$M.D.=\frac{\Sigma f|x|}{\Sigma f}=\frac{493.2}{50}=9.86\text{磅}$$

在上列第50表中係根據算術平均數計算平均差。如以中位數為標準而求平均數，除以中位數$M$代替平均數$\bar{X}$之外，其他一切手續相同，但所得之結果必小於上列根據算術平均數所得之平均差。茲在下列第50表中，

依五十男生體重分配以求之：

(第51表)　五十男生體重平均差計算表

$M=126.11$ 磅

| $i$ | $X$ | $f$ | $\|X-M\|$ | $f\|X-M\|$ |
|---|---|---|---|---|
| 95—105 | 100 | 3 | 26.11 | 78.33 |
| 105—115 | 110 | 9 | 16.11 | 144.99 |
| 115—125 | 120 | 11 | 6.11 | 67.21 |
| 125—135 | 130 | 18 | 3.89 | 70.02 |
| 135—145 | 140 | 9 | 13.89 | 125.01 |
| 總　計 | | 50 | | 485.56 |

$$M.D.=\frac{\Sigma f\,|X-M|}{\Sigma f}=\frac{485.56}{50.}=9.711\text{ 磅}$$

四、標準差

以上所得之平均差，無論根據平均數或中位數，俱由次數分配中一切項目計算而來，爲計算離勢量數(Computed Measures of Dispersion)，非復如全距及四分差之根據最大與最小兩變項之爲位置量數。根據各個變項計算固爲平均差之優點，但在計算過程中，不計正負號，亦殊不[illegible]理。補救之道，即利用各離差自乘方而將其正負號取消，結果有標準差(Standard Deviation)，以 $\sigma$ (=Sigma)表示之。標準差之定義爲：方根－平均數－平方－離差 (Root-Mean-Square-deviation)。分析言之，已得算術平均數 $\bar{X}$ 之後，依之以求標準差 $\sigma$，其步驟有四：

1. 求各變項之離差
2. 求各離差之平方
3. 求各離差平方之平均數
4. 求此平均數之方根，即爲標準差。

標準差之計算可根據未分組資料與次數分配，根據每種資料又有普通與簡捷兩種方法，茲分述之：

1. 未分組資料

甲 普通法 所謂普通法 (The Long Method) 即根據算術平均數以計算標準差，現以 $X$ 代表變數，以 $\bar{X}$ 代表平均數，以 $x$ 代表各變項之離差，則求標準差之步驟，即如上列在定義中所舉者：

1. 求各變項之離差 $x=X-\bar{X}$
2. 求各離差之平方 $x^2=(X-\bar{X})^2$
3. 求各離差平方之平均數 $\sigma^2=\frac{\Sigma x^2}{N}$
4. 求離差平方平均數之方根，即得標準差，其公式為：

$$\sigma=\sqrt{\frac{\Sigma x^2}{N}}=\sqrt{\frac{\Sigma(X-\bar{X})^2}{N}}$$

上列四步驟可用計算表以求之，手續簡易。

(第 52 表) 標準差計算表

| $X$ | $x=X-\bar{X}$ | $x^2$ |
|---|---|---|
| 1 | −3 | 9 |
| 2 | −2 | 4 |
| 3 | −1 | 1 |
| 4 | 0 | 0 |
| 5 | 1 | 1 |
| 6 | 2 | 4 |
| 7 | 3 | 9 |
| 28 | 0 | 28 |

$$\bar{X}=\frac{\Sigma X}{N}=\frac{28}{7}=4$$

$$\sigma=\sqrt{\frac{\Sigma X^2}{N}}=\sqrt{\frac{28}{7}}=\sqrt{4}=2$$

乙 簡捷法 在普通法中係根據眞正平均數$\bar{X}$求標準差。如平均數有小數位時，則計算各離差自乘方之手續繁重，乃有簡捷法。在未分組資料中求標準差之簡捷法，係令假定平均數$\bar{Z}$等于零，由是又推出公式如下：

$\bar{X}$＝眞正平均數

$\bar{Z}$＝假定平均數＝0

$x$＝由眞正平均數所計之離差， $x=X-\bar{X}$

$d$＝由假定平均數所計之離差， $d=X-\bar{Z}$

$\because\ \bar{Z}=0$， $\therefore\ d=X \qquad \Sigma d^2=\Sigma X^2$.

$c$＝眞正與假定兩平均數間之差異，

$\bar{X}=\bar{Z}+c$， 或 $c=\bar{X}-\bar{Z}$

$$\because\ \bar{Z}=0 \qquad \therefore\ c=\bar{X}=\frac{\Sigma X}{N}$$

$d-c=X-\bar{X}=x$， 或 $d=c+x$，

$\Sigma d^2=\Sigma(c+x)^2=Nc^2+2c\Sigma x+\Sigma x^2$

$\because$ (1) $2c\Sigma x=0$

(2) $c=\bar{X}$， $Nc^2=N\bar{X}^2$

(3) $\Sigma d^2=\Sigma X^2$

$\therefore\ \Sigma X^2=N\bar{X}^2+\Sigma x^2$

$\therefore\ \Sigma x^2=\Sigma X^2-N\bar{X}^2$，各項以$N$除之，

$\frac{\Sigma x^2}{N}=\frac{\Sigma X^2}{N}-\bar{X}^2$，等號前後各開平方

$$\therefore \quad \sigma=\sqrt{\frac{\Sigma x^2}{N}}=\sqrt{\frac{\Sigma X^2}{N}-\bar{X}^2}=\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}$$

(第53表) 標準差計算表

| $X$ | $X^2$ |
|---|---|
| 1 | 1 |
| 2 | 4 |
| 3 | 9 |
| 4 | 16 |
| 5 | 25 |
| 6 | 36 |
| 7 | 49 |
| 23 | 140 |

$$\sigma=\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}=\sqrt{\frac{140}{7}-\left(\frac{28}{7}\right)^2}$$

$$=\sqrt{20-4^2} \qquad =\sqrt{20-16}$$

$$=\sqrt{4} \qquad =2$$

2. 次數分配

甲 普通法 在分組資料或次數分配中,依普通法求標準差,除增加次數$f$一因素外,其他一切手續及意義與在未分組資料中相同,茲以五十男生體重爲例:

(第54表) 五十男生體重標準差計算表

$\bar{X}=124.2$磅

| $t$ | $X$ | $f$ | $x=X-\bar{X}$ | $fx$ + | $fx$ − | $fx^2$ |
|---|---|---|---|---|---|---|
| 95—105 | 100 | 3 | −24.2 | | 72.6 | 1766.92 |
| 105—115 | 110 | 9 | −14.2 | | 127.8 | 1814.76 |
| 115—125 | 120 | 11 | −4.2 | | 46.2 | 194.04 |
| 125—135 | 130 | 18 | 5.8 | 104.4 | | 605.52 |
| 135—145 | 140 | 9 | 15.8 | 142.2 | | 2246.76 |
| 總計 | | 50 | | 246.6 | 246.6 | 6628.00 |

$$\sigma=\sqrt{\frac{\Sigma fx^2}{\Sigma f}}=\sqrt{\frac{6628}{50}}=\sqrt{123.36}=11.5\text{磅}$$

乙 簡捷法之一——單位離差法 在次數分配中所用之簡捷法與在未分組資料中所用者相同。惟一異點，為在未分組資料中所用之假定平均數$\overline{Z}$為零，現在次數分配中所用之假定平均數$\overline{Z}$為一組之中值，每以用次數最多或居于中央一組之中值為便，推演公式之步驟與上文中大致相同：

$\overline{X}$＝真正平均數

$\overline{Z}$＝假定平均數＝中央組距之中值。

$x=X-\overline{X}$

$d=X-\overline{Z}$

$\overline{X}=\overline{Z}+c,\ c=\overline{X}-\overline{Z}$， 或 $\overline{Z}=\overline{X}-c$

$d=X-\overline{Z}=X-(\overline{X}-c)=X-\overline{X}+c=x+c$

$d^2=(x+c)^2=x^2+2cx+c^2$

$\Sigma fd^2=\Sigma f(x+c)^2=\Sigma fx^2+2c\Sigma x+\Sigma fc^2$

$\Sigma fd^2=\Sigma fx^2+\Sigma fc^2$

$$\frac{\Sigma fd^2}{\Sigma f}=\frac{\Sigma fx^2}{\Sigma f}+c^2,\ \text{或}\ \frac{\Sigma fx^2}{\Sigma f}=\frac{\Sigma fd^2}{\Sigma f}-c^2$$

在算術平均數簡捷法中已證明：

$c=\overline{X}-\overline{Z}=\frac{\Sigma fd}{\Sigma f}$， 代入上式

$$\frac{\Sigma fd^2}{\Sigma f}=\frac{\Sigma fx^2}{\Sigma f}-\left(\frac{\Sigma fd}{\Sigma f}\right)^2$$

等號前後各開平方：

$$\sigma=\sqrt{\frac{\Sigma fx^2}{\Sigma f}}=\sqrt{\frac{\Sigma fd^2}{\Sigma f}-\left(\frac{\Sigma fd}{\Sigma f}\right)^2}$$

（第55表）五十男生體重標準差計算表

| $i$ | $X$ | $f$ | $d$ | $fd$ | $fd^2$ |
|---|---|---|---|---|---|
| 95—105 | 100 | 3 | —20 | —60 | 1200 |
| 105—115 | 110 | 9 | —10 | —90 | 900 |
| 115—125 | 120 | 11 | 0 | 0 | 0 |
| 125—135 | 130 | 18 | 10 | 180 | 1800 |
| 135—145 | 140 | 9 | 20 | 180 | 3600 |
| 總 計 | | 50 | | +210 | 7500 |

$$\sigma=\sqrt{\frac{\Sigma fd^2}{\Sigma f}-\left(\frac{\Sigma fd}{\Sigma f}\right)^2}=\sqrt{\frac{7500}{50}-\left(\frac{210}{50}\right)^2}$$

$$=\sqrt{150-(4.2)^2} \qquad =\sqrt{150-17.64}$$

$$=\sqrt{136.36} \qquad = 11.50\text{磅}$$

丙　簡捷法之二——組距離差法　在上列簡捷法中，所用爲假定平均數之單位離差，但各組中離差之組距既相同，自可將此共同因素取消之，即除以組距 $i$，所得結果爲組距離差，以 $d_i$ 表示之，則計算手稜更爲化簡，最後所得結果，再使之還原，以組距 $i$ 乘之。

（第56表）五十男生標準差計算表

| $i$ | $X$ | $f$ | $d_i$ | $fd_i$ | $fd_i^2$ |
|---|---|---|---|---|---|
| 95—105 | 100 | 3 | —2 | —6 | 12 |
| 105—115 | 110 | 9 | —1 | —9 | 9 |
| 115—125 | 120 | 11 | 0 | 0 | 0 |
| 125—135 | 130 | 18 | 1 | 18 | 18 |
| 135—145 | 140 | 9 | 2 | 18 | 36 |
| 總 計 | | | | +21 | 75 |

由上表己知 $\Sigma fd_i=21$，$\Sigma fd_i^2=75$，以之代入簡捷法公式，自可得標準差。但在未代入之前，可利用薛立愛覆驗法（Charlier Check）以證明計算

過程中是否有誤。此法即設法於計算表中求得 $\Sigma f(d_i+1)^2$ 之值，幷檢驗是否與其展開式相等，其展開式自爲：$\Sigma fd_i^2+2\Sigma fd_i+\Sigma f$. 根據上表可作一覆驗法計算表如下：

(第57表) 薛立愛覆驗法計算表

| $i$ | $X$ | $f$ | $d_i$ | $d_i+1$ | $f(d_i+1)$ | $f(d_i+1)^2$ |
|---|---|---|---|---|---|---|
| 95—105 | 100 | 3 | −2 | −1 | −3 | 3 |
| 105—115 | 110 | 9 | −1 | 0 | 0 | 0 |
| 115—125 | 120 | 11 | 0 | 1 | 11 | 11 |
| 125—135 | 130 | 18 | 1 | 2 | 36 | 72 |
| 135—145 | 140 | 9 | 2 | 3 | 27 | 81 |
| 總計 | | 50 | | | | 167 |

$\Sigma f(d_i+1)^2=167$

$\Sigma fd_i^2+2\Sigma fd_i+\Sigma f=75+2\times21+50=167$

由上例覆驗，可知計算過程正確無誤，即可依簡捷法公式以求標準差，但所得爲組距單位之標準差，再以組距乘之，即爲最後所要求之結果：

$$\sigma_i=\sqrt{\frac{\Sigma fd_i^2}{\Sigma f}-\left(\frac{\Sigma fd_i}{\Sigma f}\right)^2}=\sqrt{\frac{75}{50}-\left(\frac{21}{50}\right)^2}$$

$$=\sqrt{1.5-(0.42)^2}=\sqrt{1.5-0.1764}$$

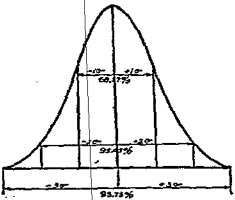
$$=\sqrt{1.3235}=1.15$$

$$\sigma=i\times\sigma_i=10\times1.15=11.5\text{ 磅}$$

以上各種離勢量數之中，以標準差一項爲最優，在次數分配中，即以平均數 $\bar{X}$ 代表各變項之集中趨勢，而以標準差 $\sigma$ 代表離中趨勢. 各個變項雖與平均數不同，但其出入每有一定之範圍. 此種範圍可利用標準差表示之，稱爲估計界限 (Zones of Estimate) 如爲常態分配，則各變項離中趨勢之範圍如下：

(第58表)　常態分配估計界限表

| 自平均數$\bar{X}$之距離(依標準差$\sigma$計) | 常態分配中所包括次數百分比 |
|---|---|
| $\bar{X}\pm 0.6745\sigma$ | 50.00 |
| $\bar{X}\pm 1\sigma$ | 68.27 |
| $\bar{X}\pm 2\sigma$ | 95.45 |
| $\bar{X}\pm 3\sigma$ | 99.73 |



第19圖　常態分配估計界限圖

上表內容之意，卽在常態分配中，各個變項與平均數$\bar{X}$雖不相同，但其差異有一定之範圍，以平均數$\bar{X}$爲準時，減去0.6745倍標準差$\sigma$及加上0.6745倍標準差$\sigma$所造成之範圍中，恰包括變項之半數，卽次數之50%；平均數加減一倍標準差所造成之範圍，恰包括次數之68.27%；加減二倍標準差所造成之範圍，恰包括次數之95.45%；加減三倍標準差所造成之範圍，恰包括次數之99.73%，已近乎全部次數卽100%。足徵各個變項雖與平均數不同，但其出入之範圍，以平均數加減三倍標準差爲最大限度，百項之中出此範圍之外者，不過0.27項而已。

以上估計限界只限於常態分配。如實際之分配不合乎常態時，則估計界限之百分比未必完全如上表所載者。如以五十男生體重分配爲例，一

望而知其非爲常態分配，故以標準差而定估計範圍時，只爲近似值，而非能絕對正確也。

(第 59 表)　五十男生體重估計界限計算表

$\overline{X}=124.2$;　$\sigma=11.5$

| 自平均數之距離 | | 估計界限 | 常態分配中應包括之次數 |
|---|---|---|---|
| 依標準差計 | 依 $\overline{X}$ 原單位計 | | |
| $\overline{X}\pm0.6745\sigma$ | $124.2\pm7.8$ | 116.4—132.0 | 50.00 |
| $\overline{X}\pm1\sigma$ | $124.2\pm11.5$ | 112.7—135.7 | 68.27 |
| $\overline{X}\pm2\sigma$ | $124.2\pm23.0$ | 101.2—147.2 | 95.45 |
| $\overline{X}\pm3\sigma$ | $124.2\pm34.5$ | 89.7—160.7 | 93.73 |

由上表可知如五十男生之體重合乎常態分配，則五十變項之中，介乎116.4至132磅之間者，應爲二十五項，全部變項即五十人之體重，儘可與平均數不同，但無一出乎89.7磅至160.7磅之外者。

## 第二節　相對離勢量數

上節所得四種絕對離勢量數，即全距，四分差，平均差及標準差，用以描寫次數分配中之離中趨勢，自無問題，用以比較次數分配間之離中趨勢時則遭遇困難：

第一，兩次數分配之單位不同，根本上無法彼此比較。如身高分配之單位爲公分，體重分配之單位爲公斤，雖各求出平均數及標準差，但吾人無法即根據兩者之標準差以斷定其離中趨勢之大小。解決之道，須將絕對離勢量數化爲相對離勢數量，取消身高與體重兩標準差之單位，使各成爲百分數。既同爲百分數，則可彼此比較，其離勢百分比高者，即離中趨勢大，反之，其離勢百分比低者，即離中趨勢小。

第二，甲、乙兩標準差之單位儘同，但其大小所代表之意義不同，亦不

能卽根據標準差以斷定兩分配離勢之大小。如三十六年元月份上海黃金市價平均數爲每十市兩價值法幣三百八十三萬元，標準差爲一十九萬三千元，以$\sigma_x$代表之，美鈔市價平均數爲美鈔每圓價值法幣六千七百六十元，標準差爲三百六十四元，以$\sigma_y$代表之，吾人如以 $\sigma_x=\$193,000$ 及 $\sigma_y=\$364$，而卽認爲黃金市價之離勢大，而美鈔市價之離勢小，自屬謬誤。蓋黃金市價之標準差大，其平均數亦大，自標準差之絕對值卽十九萬三千元觀之，其數誠大，但如化爲百分比時，其值未必卽大。反之，美鈔市價之標準差小，其平均數亦小，自標準差之絕對值卽三百六十四元觀之，其數誠小，但如化爲百分比時，其值未必卽小。

相對離勢量數，卽以百分比表示離中趨勢之大小，其意義及重要性已如上述，其公式則有下列三種：

一、標準差係數

所謂標準差係數（Coefficient of Standard Deviation）亦稱爲離勢係數（Coefficient of Variation），卽以算術平均數$X$除標準差$\sigma$，將商數用百分比形式表出，以$V\sigma$代表之，如五十男生重量平均數$X=124.2$磅，標準差$\sigma=11.5$磅。標準差係數爲：

$$V\sigma=\frac{\sigma}{X}=\frac{11.5}{124.2}=9.26\%$$

二、平均差係數

所謂平均差係數（Coefficient of Mean Deviation）係以平均數$X$或中位數$M$除平均差$M.D.$所得之商數用百分比形式表出，以$V_{M.D.}$代表之。如五十男生體重分配中算術平均數$X=124.2$磅，由平均數計算之平均差$M.D.=9.86$磅，則平均差係數爲：

$$V_{M.D.}=\frac{9.86}{124.20}=7.94\%$$

三、四分差係數

所謂四分差係數(Coefficient of Quartile Deviation),即以上下兩四分位數之平均數除四分差，所得之商數用百分比表出;以 $V_Q$ 代表之。如五十男生體重分配中,上四分位數 $Q_3=133.06$ 磅，下四分位數 $Q_1=115.45$ 磅,則四分差係數爲:

$$V_Q=\frac{\frac{Q_3-Q_1}{2}}{\frac{Q_3+Q_1}{2}}=\frac{Q_3-Q_1}{Q_3+Q_1}$$

$$=\frac{133.06-115.45}{133.06+115.45}=\frac{17.61}{248.51}$$

$$=7.08\%$$

## 習　　題

1. 試根據第三章習題所作之100名大學男生體重分配表求下列各絕對離勢量數:

　(1) 全距

　(2) 四分差

　(3) 平均差:

　　i 依算術平均數計算

　　ii 依中位數計算

　(4) 標準差:

　　i 普通法

　　ii 假定原點組距離差法(附用薛立愛氏覆驗法)

試依上題所得各結果,求下列各相對離勢量數:

（1） 標準差係數

（2） 平均差係數

（3） 四分差係數

# 第七章　偏態與峯度

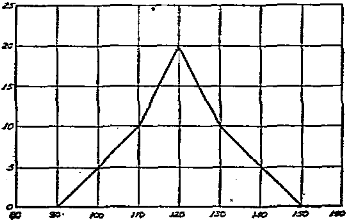
## 第一節　偏態

次數分配中各變項有離中趨勢，但其大于平均數各變項之離中趨勢與小于平均數各變項之離中趨勢亦不相同，由於大小兩半變項離勢不同，乃造成次數分配之偏態(Skewness)。試依五十男生體重分配表格及圖形爲例以說明之：

一　對稱分配　　(第60表)　五十男生體重分配表

| 組　距 | 中　值 | 次　數 |
|---|---|---|
| 95—105 | 100 | 5 |
| 105—115 | 110 | 10 |
| 115—125 | 120 | 20 |
| 125—135 | 130 | 10 |
| 135—145 | 140 | 5 |
| 總　計 | | 50 |

第20圖　五十男生體重次數圖



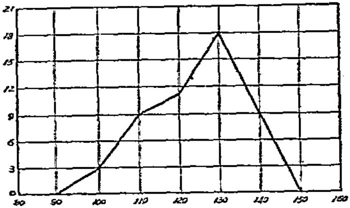
由上列第 20 圖可見此五十男生體重資料所造成者爲一對稱分配(Symmetric Distribution)。在次數曲綫圖中以最高縱綫爲分界綫，此綫恰將曲綫下之面積劃爲兩等分；左右兩半之構造完全相同。在此項對稱分配中 $\bar{X}$，$M$，及$Mo$ 即平均數，中位數及衆數三值重合於一點，俱爲最高縱綫所定出之點：即三種量數均爲 120 磅。

二 左偏分配

(第 61 表) 五十男生體重分配表

| 組 距 | 中 值 | 次 數 |
|---|---|---|
| 95—105 | 100 | 3 |
| 105—115 | 110 | 9 |
| 115—125 | 120 | 11 |
| 125—135 | 130 | 18 |
| 135—145 | 140 | 9 |
| 總 計 | | 50 |

第 21 圖 五十男生體重次數圖




由上列第 21 圖形，可見出此五十男生體重資料所造成爲左偏分配(Left Shewed Distribution)。試以衆數 130 磅爲標準而向左右兩方向觀察之，在衆數所在點即最高縱綫之右，次數曲綫急轉直下，由 18 而 9，再

由 9 而零；反之，在最高縱綫以左次數曲綫乃緩和下降，其次數由 18 而 11，由 11 而 9，由 9 而 3，最後由 3 而零。凡在最高縱綫之一側面，次數曲綫緩和下降終至于零，則偏態方向卽在于此，如上圖次數曲綫在最高縱綫左側次數曲綫下降緩和，卽造成左偏分配，亦稱爲負偏分配 (Negative-skewed Distribution)

三 右偏分配

(第 62 表) 五十男生體重分配表

| 組距 | 中值 | 次數 |
|---|---|---|
| 95—105 | 100 | 9 |
| 105—115 | 110 | 18 |
| 115—125 | 120 | 11 |
| 125—135 | 130 | 9 |
| 135—145 | 140 | 3 |
| 總計 | | 50 |

第 22 圖 五十男生體重次數圖


上圖所示爲右偏分配 (Right-Skewed Distribution)。以最高縱綫卽衆數所在之點 110 磅爲準，在此綫以左之次數較爲密集，其曲綫下降急峻；而在最高縱綫以右者較爲疏散，次數曲綫下降緩和，因造成右偏分配，亦稱爲正偏分配 (Positive-skewed Distribution)。偏態之意義既明，現說

明計算偏態係數之公式:

一 皮爾森氏偏態係數公式之一

皮爾森氏 (Karl Pearson) 根據$\bar{X}$與$Mo$即平均數與衆數之關係而得一偏態係數公式:

$$Sk=\frac{\bar{X}-Mo}{\sigma}$$

1 對稱分配

在對稱分配中$\bar{X}$與$Mo$既相等,在圖中重合于於一點,則二者之差爲零,不拘標準差爲何值,偏態係數$Sk$必爲零。既爲對稱分配根本無偏態,其係數自爲零也。

2 左偏分配

平均數$\bar{X}$之特質爲易受極端變項之影響,各變項之中有少數最小極項時,則平均數即小,有少數最大極項時,則平均數即大。在左偏分配中,因有最小極項存在,故平均數$\bar{X}$較小。次數最密集之點即衆數較平均數爲大.根據平均數與衆數所得之偏態係數自然爲負。以五十男生體重左偏分配爲例,平均數$\bar{X}=124.2$磅,衆數$Mo=130$磅,標準差$\sigma=11.5$磅,以三值代入公式,則得偏態係數之值如下:

$$Sk=\frac{\bar{X}-Mo}{\sigma}=\frac{124.2-130}{11.5}=\frac{-5.8}{11.5}=-0.544$$

3 右偏分配

在右偏分配中,因有最大極項存在,故平均數較大,次數密集之點即衆數較平均數爲小,根據平均數與衆數,所得之偏態係數自然爲正。以五十男生右偏分配爲例,平均數$\bar{X}=115.8$磅,衆數$Mo=110$磅,標準差$\sigma=11.5$磅。以三值代入公式,則得偏態係數之值如下:

$$Sk=\frac{\bar{X}-Mo}{\sigma}=\frac{115.8-110}{11.5}=\frac{5.8}{11.5}=0.544$$

二 皮爾森氏偏態係數公式之二

在講述衆數時，曾舉出皮爾森氏經驗公式，其形式爲：

$Mo=\bar{X}-3(\bar{X}-M)$。由此可以推出一偏態係數公式：

$$Sk=\frac{3(\bar{X}-M)}{\sigma}$$

1 對稱分配

在 $Mo=\bar{X}-3(\bar{X}-M)$ 一式中，減號後 $3(\bar{X}-M)$ 即代表偏態者。如爲對稱分配，則 $Mo$，$\bar{X}$，$M$. 三種均數重合于一點，即三者彼此相等，則 $3(\bar{X}-M)$ 爲零，偏態係數全式 $3(\bar{X}-M)/\sigma$ 自亦爲零，根本無偏態存在。

2 左偏分配

在左偏分配中，因受最小極項之影響，平均數 $\bar{X}$ 偏小，而中位數 $M$ 係平分全部次數爲兩等分之值，自較平均數爲大，依平均數及中位數所得之偏態係數自爲負。以五十男生體重左偏分配爲例，$\bar{X}=124.2$ 磅 $M=126.11$ 磅，$\sigma=11.5$ 磅。將此三值代入公式則得偏態係數之值：

$$Sk=\frac{3(\bar{X}-M)}{\sigma}=\frac{3(124.2-126.11)}{11.5}$$

$$=\frac{-5.73}{1.15}=-0.498$$

3 右偏分配

在右偏分配中，因受最大極項之影響，平均數 $\bar{X}$ 偏大，而中位數 $M$ 爲平分全部次數爲兩等分之值，自比平均數 $\bar{X}$ 爲小，是依平均數及中位數所得之偏態係數爲正。以五十男生右偏分配爲例。$\bar{X}=115.8$ 磅，$M=113.89$ 磅，$\sigma=11.5$ 磅。將此三值代入公式，即得偏態係數之值：

$$Sk=\frac{3(115.8-113.89)}{11.5}=\frac{5.73}{11.5}=0.448$$

三　包萊氏偏態係數公式

包萊教授（Prof. Bowley）根據 $M$，$Q_1$，與 $Q_3$ 卽中位數及上下四分位數而計算偏態之有無。三者旣各爲四分位數，將全部次數劃爲四等分，則在 $Q_1$ 與 $M$ 間包括次數四分之一，在 $M$ 及 $Q_3$ 間亦包括次數四分之一。在此二距離間所包括之次數雖相等，但此二距離自身之長短則不必相等。中位數至上四分位數間之距離以 $(Q_3-M)$ 表示之，下四分位數至中位數間之距離以 $(M-Q_1)$ 表示之。根據此兩段距離是否相等而得一偏態係數公式如下：

$$Sk=\frac{(Q_3-M)-(M-Q_1)}{Q.D.}$$

$$=\frac{Q_3+Q_1-2M}{\frac{Q_3-Q_1}{2}}$$

1　對稱分配

在對稱分配中，由下四分位數 $Q_1$ 至中位數 $M$ 間之距離，與中位數 $M$ 至上四分位數 $Q_3$ 間之距離恰相等，卽 $(M-Q_1)$ 與 $(Q_3-M)$ 相等，則偏態係數公式自等於零，表示根本無偏態。

2　左偏分配

在左偏分配中，中位數固爲次數之平分界限，但小于中位數之一半次數較爲分散，大于中位數之一半次數較爲集中。後者旣較集中，則自中位數 $M$ 至上四分位數 $Q_3$ 間所佔之距離較短，前者旣較爲分散，則自下四分位數 $Q_1$ 至中位數 $M$ 間所佔之距離較長，換言之，卽 $(Q_3-M$ 小于 $(M-Q_1)$，依此二者所求得之偏態係數自然爲負。以五十男生體重左偏分配爲

例，$Q_1=115.45$ 磅，$M=126.11$ 磅，$Q_3=133.06$ 磅。將此三值代入公式，則得偏態係數如下：

$$Sk=\frac{Q_3+Q_1-2M}{\frac{Q_3-Q_1}{2}}=\frac{133.06+115.45-2\times126.11}{\frac{133.06-115.45}{2}}$$

$$=\frac{248.51-252.22}{\frac{17.61}{2}}=\frac{-3.71}{8.805}=-0.421$$

3　右偏分配

在右偏分配中，以中位數爲標準時，小于中位數之一半次數較爲集中，大于中位數之一半次數較爲分散。後者既較爲分散，則自中位數 $M$ 至上四分位數 $Q_3$ 間所佔之距離較長；前者既較爲集中，則由下四分位數至中位數 $M$ 間所佔之距離自然較短，換言之，卽 $(Q_3-M)$ 大于 $(M-Q_1)$，依此二者所得之偏態係數自必爲正。以五十男生體重右偏分配爲例，$Q_1=106.94$ 磅，$M=113.89$ 磅，$Q_3=124.55$ 磅。將此三值代入公式，卽偏態係數公式如下：

$$Sk=\frac{Q_3+Q_1-2M}{\frac{Q_3-Q_1}{2}}=\frac{124.55+106.94-2\times113.89}{\frac{124.55-106.94}{2}}$$

$$=\frac{231.49-227.78}{\frac{17.61}{2}}=\frac{3.71}{8.805}=0.421$$

## 第二節　峯度

次數分配之第一特點爲集中趨勢，因求各種平均數以代表之；第二特點爲離中趨勢，因求各種離勢量數以代表之；第三特點爲偏態，因求各種偏態係數以代表之，偏態實爲離勢問題之一部。次數分配之第四特點爲峯

度(Kurtosis)。所謂峯度直接指次數分配作爲次數圖時，其次數曲綫在最高縱綫(Maximum Ordinate)附近向上升起之限度，間接則指以衆數爲標準各個變項向中心密集之程度。如次數曲綫升起之峯度高，則表示其次數向中心密集之程度高，由此項分配中所求出平均數之代表性高，平均數一值足以充分代表各變項。反之，次數曲綫之峯度低，則表示各次數向中心密集之程度不高，由此項分配中所求平均數之代表性亦不高，平均數一值未必能充分代表各變項。代表峯度高低之數值爲峯度量數(Measure of kurtosis)求峯度量數之工具則爲動差(Moments)，動差之計算可分爲下列三層：

一 概約動差

所謂概約動差(Crude Moments)係指由假定平均數$Z$所求之動差而言，其符號爲$V(=\nu_1)$計算時概用組距單位，現以五十男生體重分配爲例，以求各級概約動差。

第63表 五十男生體重概約動差計算表

| $i$ | $X$ | $f$ | $d=\frac{X-Z}{i}$ | $fd$ | $fd^2$ | $fd^3$ | $fd^4$ |
|---|---|---|---|---|---|---|---|
| 95—105 | 100 | 3 | −2 | −6 | 12 | −24 | 48 |
| 105—115 | 110 | 9 | −1 | −9 | 9 | −9 | 9 |
| 115—125 | 120 | 11 | 0 | 0 | 0 | 0 | 0 |
| 125—135 | 130 | 18 | 1 | 18 | 18 | 18 | 18 |
| 135—145 | 140 | 9 | 2 | 18 | 36 | 72 | 144 |
| | | 50 | | 21 | 75 | 57 | 219 |

假定平均數 $Z=120$ 磅

第一概約動差 $V_1=\frac{\Sigma fd}{\Sigma f}=\frac{21}{50}=0.42$

第二概約動差 $V_2=\frac{\Sigma fd^2}{\Sigma f}=\frac{75}{50}=1.50$

第三概約動差 $V_3=\frac{\Sigma fd^3}{\Sigma f}=\frac{57}{50}=1.14$

第四概約動差 $V_4=\frac{\Sigma fd^4}{\Sigma f}=\frac{219}{50}=4.38$

## 二 主要動差

所謂主要動差(Principal Moments)即指由眞正平均數$\bar{X}$所求出之動差。其符號爲$\pi$(Pi)。計算之方法有二，即普通法與簡捷法。根據普通方法以$\bar{X}$爲中心，而求其各級離差之平均數，依組距單位表出，即爲各級主要動差。但以眞正平均數每有小數，乘方次數增多，小數位計算困難，須另覓捷徑。所謂簡捷法係由概約動差推求主要動差。茲證明其方法之來源，並以五十男生體重分配爲例而求其主要動差之值：

$$x=X-\bar{X}$$

$$d=X-\bar{Z}$$

$$\bar{X}=\bar{Z}+c$$

$$x=X-\bar{X}=X-(\bar{Z}+c)=X-\bar{Z}-c=d-c$$

$$c=\frac{\Sigma fd}{\Sigma f}=V_1$$

$$\therefore\ x=d-V_1$$

第一主要動差 $\pi_1=\frac{\Sigma fx}{\Sigma f}=0$

第二主要動差 $\pi_2=\frac{\Sigma fx^2}{\Sigma f}=\sigma^2=\frac{\Sigma fd^2}{\Sigma f}-\left(\frac{\Sigma fd}{\Sigma f}\right)^2$

$$\because\ V_1=\frac{\Sigma fd}{\Sigma f}\ ,\qquad V_2=\frac{\Sigma fd^2}{\Sigma f}$$

$$\therefore\ \pi_2=V_2-V_1^2$$

$$\because\ V_1=0.42\ ;\qquad V_2=1.5$$

$$\therefore\ \pi_2=1.5-(0.42^2)=1.5-0.1764=1.3236$$

第三主要動差 $\pi_3=\dfrac{\Sigma fx^3}{\Sigma f}=\dfrac{\Sigma f(d-V_1)^3}{\Sigma f}$

$$=\frac{\Sigma f(d^3-3V_1d^2+3V_1^2d-V_1^3)}{\Sigma f}$$

$$=\frac{\Sigma fd^3}{\Sigma f}-\frac{3V_1\Sigma fd^2}{\Sigma f}+\frac{3V_1^2\Sigma fd}{\Sigma f}-V_1^3$$

$$=V_3-3V_1V_2+2V_1^3$$

$$\because\ V_1=0.42,\ V_2=1.5,\ V_3=1.14$$

$$\therefore\ \pi_3=1.14-3\times0.42\times1.5+2\times(0.42)^3$$

$$=1.14-1.89+0.1482=-0.6018$$

第四主要動差 $\pi_4=\dfrac{\Sigma fd^4}{\Sigma f}=\dfrac{\Sigma f(d-V_1)^4}{\Sigma f}$

$$=\frac{\Sigma f(d^4-4V_1d^3+6V_1^2d^2-4V_1^3d+V_1^4)}{\Sigma f}$$

$$=\frac{\Sigma fd^4}{\Sigma f}-4V_1\frac{\Sigma fd^3}{\Sigma f}+6V_1^2\frac{\Sigma fd^2}{\Sigma f}-4V_1^3\frac{\Sigma fd}{\Sigma f}$$

$$+V_1^4$$

$$=V_4-4V_1V_3+6V_1^2V_2-3V_1^4$$

$$\because\ V_1=0.42,\ V_2=1.5,\ V_3=1.14,\ V_4=4.38$$

$$\therefore\ \pi_4=4.38-4\times0.42\times1.14+6\times(0.42)^2\times1.5$$

$$-3(0.42)^4$$

$$=4.38-1.9152+1.5876-0.09336$$

$$=3.95904$$

三 改正動差

以上計算次數分配之各級動差時，概係中值求之，係假定每一組中所包括之變項，一律等於其組中值，此種假定實含有一錯誤，須設法改正之，

統計學家薛伯氏Sheppard提出下列改正數，稱爲薛氏改正數(Sheppard corrections)，主要動差經此改正後稱爲改正動差(Cor ected Moments)，其符號爲$\mu$(mu)：

第一改正動差 $\mu_1=\pi_1=0$

第二改正動差 $\mu_2=\pi_2-\frac{1}{12}$

$\because\quad \pi_2=1.3236,\ \frac{1}{12}=0.08333$

$\therefore\quad \mu_2=1.3236-0.08333=1.24027$

第三改正動差 $\mu_3=\pi_3=-0.6018$

第四改正動差 $\mu_4=\pi_4-\frac{1}{2}\pi_2+\frac{7}{240}$

$\because\quad \pi_2=1.3236,\quad \pi_4=3.95904,$

$\frac{7}{240}=0.029167$

$\therefore\quad \mu_4=3.95904-\frac{1}{2}\times 1.3236+0.029167$

$=3.326407$

計算動差之程序在由概約動差而主要動差，由主要動差而改正動差；計算之目的則在由動差而求出代表峯度之量數。上列所謂之改正動差係用組距單位，現設法將其單位消除之，成爲係數，以$\beta$代表之，其公式如下：

$$\beta_1=\frac{\mu_3^2}{\mu_2^3}=\frac{(-.6018)^2}{(1.2403)^3}=\frac{0.3622}{1.9067}=0.19$$

$$\beta_2=\frac{\mu_4^2}{\mu_2^2}=\frac{3.3264}{(1.2403)^2}=\frac{3.3264}{1.5383}=2.16$$

$$K_2=\frac{\beta_1(\beta_2-3)^2}{4(4\beta_2-2\beta_1)(2\beta_2-3\beta_1-\ )}$$

$$=\frac{0.19(2.16-3)^2}{4(4\times 2.16-3\times 0.19)(2\times 2.16-3\times 0.19-6)}$$

$$=\frac{0.19\times 0.7056}{4\times 8.07\times(-2.55)}=\frac{0.134064}{72.314}=0.0019$$

在常態分配中，$B_1=0$，$B_2=3$，$K_2=0$，就中 $B_1$ 係代表曲綫之偏態，$B_2=3$ 係代表曲綫之峯度者，卽次數分配如合於常態時，其次數曲綫之偏態等於零，峯度等於3。如一次數分配之峯度量數爲 3 時，則稱其次數曲綫爲正常峯度 (Mesokurtic)；如所得之量數 $\beta_2$ 小于 3 時，則稱其次數曲綫爲低闊峯度 (Platykurtic)；如所得之量數 $\beta_2$ 大于 3 時，則稱其次數曲綫爲高狹峯度 (Leptokurtic)。在吾人所分析之五十男生體重分配中，其峯度量數 $\beta_2$ 等于 2.16，是小于 3，爲低闊峯度，但低闊之程度並不過劇，去正常峯度不太遠。換言之，卽其次數之集中趨勢尙相當顯著，平均數足以代表各個變數也。

## 習　題

1. 試根據 100 名大學男生體重分配資料而求下列各偏態係數：
   (1) 皮爾森氏公式(依平均數與中位數)
   (2) 包萊氏公式
2. 試根據同一資料而計算其偏態係數 $\beta_1$ 及峯度量數 $\beta_2$ 之值。

# 第三編
# 相關

# 第八章　相關之性質

## 第一節　因果關係與相互關係

自然現象中間，彼此有關係，社會現象中間，彼此之關係更多。自然現象間之關係每屬於因果關係（Causation）。所謂因果關係，即甲乙兩現象中一爲前因，一爲後果。甲種現象存在，乙種現象必隨之而發生，例如：水遇冷化爲冰，遇熱化爲氣，俱係因果分明，有原因在前存在，必有結果隨後發生，是甲乙兩現象間有因果關係。

社會現象或生物現象間亦有關係存在，有者爲因果關係，有者則不易以因果關係解釋之。因爲社會現象係許多複雜因素交織造成之結果，每一因素俱隨時在變動之中，其結果亦隨時在變動之中，變動之程度又彼此不同，吾人不能只根據其中一項原因，即斷定必產生某種結果，或必達到某種程度。以吾人之高度與重量而論，凡身材高大者，亦多半較重；身材低小者，亦多半較輕。但此僅就大多數人或平均情形而論；其中亦有例外者，如某甲之身材雖高，但以清瘦而體輕；某乙之身材雖矮，但以肥胖而體重。即其身高而重，身低而輕者，其高度與重量增加或減少之程度，亦至爲參差不一。身高與體重中間有關係存在，自甚顯明，但不能解爲因果關係。蓋前已分析，嚴格而論，甲乙兩現象，如甲爲因，乙爲果，則甲已存在，乙必發生，其變化在程度上且有一定之比例。今高度與重量間之關係，既於此不合，不能稱爲因果關係，只能稱爲互相關係，簡稱爲相關（Correlation），或稱之爲聯繫（Association）。他如父子之身材，夫妻之年齡，亦均只能以相關說明之。以父子之身材而論，凡父親身材高大者，子女亦大半高大，可見

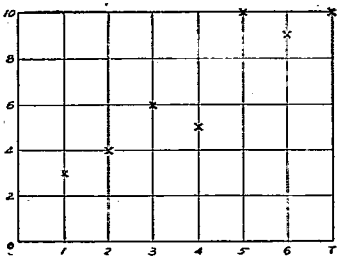
父子身材之間有關係，但亦有父親高大，而子女矮小，或父親矮小，而子女高大者。以夫妻年齡而論，丈夫年齡大者，妻子年齡每亦大；丈夫年齡小者，妻子年齡每亦小，足徵夫妻年齡中間有關係。但亦有丈夫年齡大而妻子年齡小，或丈夫年齡小而妻子年齡大者。舉經濟事實為例：三十五年下半年中，上海市場中美鈔及黃金價格，曾發生狂烈之波動，美鈔上漲，黃金亦隨上漲，二者中間有關係存在，顯而易見，但吾人不宜即此斷定美鈔上漲為因，黃金上漲為果。美鈔上漲雖引起黃金上漲，但二者上漲之程度大不相同，八月中旬美鈔上漲之程度劇烈，黃金上漲亦甚劇烈。九月下旬美鈔上漲雖仍狂烈，而金價波動之範圍至小，程度至微。最後年底十一、二月間，金價繼續猛漲，由二百萬至三百萬，由三百萬而四百萬元，美鈔雖亦漲，但較為緩和，在五千至七千之間，足徵二者間非係前因後果關係，而為相關。又如通貨膨脹與物價上漲，物價上漲與生活費用之提高，生活費用之提高與公教人員待遇之調整，其間之關係，亦均為相關或聯繫。

## 第二節　相關之種類

只就兩種現象，或兩組統計材料，或兩個變數（Variables），而分析其中之關係，稱為簡單相關（Simple Correlation）。如身高與體重間之相關即是。為便於分析計，吾人可就兩變數中，指定其一為自動變數（Independent Variable），或稱為自變數，以 $X$ 表示之，所餘之一為隨動變數（Dependent Variable），或稱為因變數，以 $Y$ 表示之。$X$ 變動，$Y$ 即隨之而變，因造成相關，以 $X$，$Y$ 二變數中各對變項製入一圖，如 $X$，$Y$ 兩變數之各項，大致分佈在一條直線上，稱之直線相關（Linear Correlation）。例如依下列 $X$，$Y$ 兩變數各值製圖，

（第 64 表）　未分組相關資料

| X | Y |
|---|---|
| 1 | 3 |
| 2 | 4 |
| 3 | 6 |
| 4 | 5 |
| 5 | 10 |
| 6 | 9 |
| 7 | 10 |

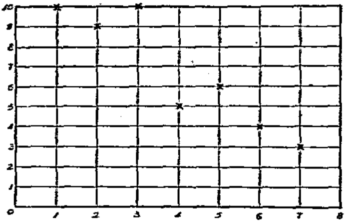



（第 73 圖）　分散圖

依 $X$，$Y$ 兩變數每一對變項在圖中作一點，七對數字共作七點，七點大致循一直綫分佈，因稱為直綫相關（Linear Correlation）。圖中代表材料分佈所循直綫之方向，自左下方向右上方，係表示 $X$ 加大時，$Y$ 亦加大，$X$ 減小時，$Y$ 亦減小。$X$，$Y$ 變動之方向相同，稱為正相關（Positive Correlation）；反之，如 $X$ 加大時，$Y$ 減小；$X$ 減小時，$Y$ 加大。雖亦可造成直綫相關，但其方向則為負相關 (Negative Correlaton)。如依下列數值作一分散圖即明。

(第65表)

| X | Y |
|---|---|
| 1 | 10 |
| 2 | 9 |
| 3 | 10 |
| 4 | 5 |
| 5 | 6 |
| 6 | 4 |
| 7 | 3 |




(第24圖) 分散圖

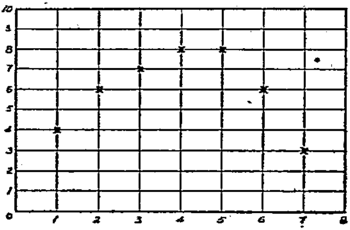
自上列第23，24兩圖觀之，各點雖循一直綫之方向斜列，但并非完全位於一直綫上，而係離開直綫，分散於一條面積之內，此圖形既表示相關材料分散之情形，因稱之爲分散圖（Scatter Diagram）。如各點完全落於直綫上，不拘其直綫之方面爲正爲負，俱爲完全相關（Perfect Correlation），如以百分數表示其程度時，爲百分之百。反之，各點之分佈毫無規律，絕不遵循任何直綫或狹長地帶，則表示二現象或二變數中間根本無關係，以百分數表之，其值爲零。一般而論，兩種現象間相關之程度，每介乎零與百分之百兩極端之間，即現象間非絕對分立，亦非絕對相依。兩變數

各位所製成之各點，落於一長條地帶之內，卽表示相關之折中情形。各點愈近于直綫卽長條地帶愈狹，表示相關之程度愈深。反之，各點愈分散，長條地帶愈寬，則表示相關之程度愈淺。

以上係兩項相關各循直綫分佈者，稱爲直綫相關。但有時資料間之相關並非如此簡單，卽非能以直綫代表者。甲乙兩現象，甲繼續加大，乙則始而加大，繼而不變，最後減小，或同爲增加或減少，其加減之速率不同，結果所形成之相關表現于分散圖中時，非爲直綫乃爲曲綫相關（Curvilinear Correlation）。依下列第66，67兩表中所列之資料循次作成第25，26兩圖

（第66表）

| X | Y |
|---|---|
| 1 | 4 |
| 2 | 6 |
| 3 | 7 |
| 4 | 8 |
| 5 | 8 |
| 6 | 6 |
| 7 | 3 |

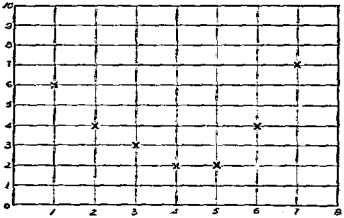



（第25圖）分散圖

即可見曲綫相關之一般：

（第67表）

| X | Y |
|---|---|
| 1 | 6 |
| 2 | 4 |
| 3 | 3 |
| 4 | 2 |
| 5 | 2 |
| 6 | 4 |
| 7 | 7 |




（第26圖） 分散圖

曲綫相關之實例甚多，如農業中肥料與農產品，或投資與生產報酬，生產數量與生產成本均是。以投資及其報酬而言，投資加百分之十，生產結果所加者，最初可多於百分之十，但以生產報酬遞減關係，投資又加百分之十，生產品亦加百分之十；投資再加百分之十，生產品所加者，則少於百分之十，試以 X 代表投資總額，爲自變數；Y 代表生產品總數爲因變數而製成分散圖，其材料之分佈，即循曲綫。上列所述，只係兩種現象間之相

關，名爲簡單相關。但相關初不限於兩種現象間，社會現象至爲複雜，許多現象中間每有相互之關係，亦即許多現象均與某一現象有聯繫。以吾人身體發育而論，年齡，高度，重量及體力四種因子中間有關係，亦可言三者中任何一項與所餘一因子有關係。以農產品而論：温度，雨量，肥料三者與之各有關係。現有甲、乙、丙、丁四組材料，亦即四個變數，如吾人假定甲、乙、丙三者俱爲自變數，丁爲因變數，進而研究三項自變數依集體之方式而與丁項因變數間之相關，稱爲多項相關(Multiple Correlation)，亦稱爲多相關或複相關。在多項相關中，因變數仍爲一項，其自變數雖有數項之多，在想像中仍可視爲一項。所異於簡單相關者，在彼中之自變數，可稱爲簡單變數 (Simple Variable)。現在多項相關中之自變數，可稱爲複合變數 (Composite Variable)，意即其中包括有幾個變數也。

同有甲、乙、丙三自變數，丁爲因變數，但吾人所研究者，非甲乙丙三者集合體對於丁之相關，而係考察甲，乙、丙三者中每一因子對丁之相關。研究時，自變數雖有三，但吾人可依次先提出一自動因子，單獨考察其與因變數間之關係。其他自變數，均可依次研究，因稱爲偏相關 (Partial Correlation)，或純相關(Net Correlation)。

吾人分析相關時，先分析簡單相關，次及於多相關及偏相關。後二者之方法雖較繁，但仍可視爲簡單相關之引伸，基本原理及技術相同，只以包括之因子加多，手續乃繁，因子愈多，手續自愈繁，實際上之價值亦愈少。

## 第三節 相關表

上節所舉之例中 $X$，$Y$ 兩變數只包括七對變項，爲數不多，不須分類化簡，即可直接作各種分析。但如變項加多至數百數千，則必須先行分類化簡，製成相關表(Cor ela ion Table)，然後方能加以分析。相關表之作

法已於以前統計表一章中舉例說明，現仍用此表，以爲分析之根據。

（第 68 表）　　民國三十六年一月

上海金條與美鈔市價相關表

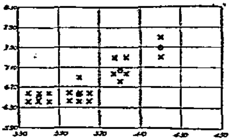
$X$：金條十市兩價格（單位國幣百萬元）

$Y$：美鈔一圓價格（單位國幣千元）

| $i_x=0.20$ ╲ $i_y=0.40$ | | 3.500–3.700 | 3.700–3.900 | 3.900–4.100 | 4.100–4.300 | 總計 |
|---|---|---|---|---|---|---|
| | $Y$ ╲ $X$ | 3.600 | 3.800 | 4.000 | 4.200 | |
| 7.50–7.90 | 7.70 | | | | 1 | 1 |
| 7.10–7.50 | 7.30 | | | 2 | 1 | 3 |
| 6.70–7.10 | 6.90 | | 1 | 3 | | 4 |
| 6.30–6.70 | 6.50 | 6 | 6 | | | 12 |
| 總　計 | | 6 | 7 | 5 | 2 | |
| 行列平均數 | | 6.50 | 6.56 | 7.06 | 7.50 | |

上列之相關表，即利用表格方式，將相關材料表出，表中 $X$ 各項橫列，自左而右；$Y$ 各項縱列，自下而上，與作圖時 $X$，$Y$ 各變項之表示法相同，自表中吾人即可見出相關材料各項目係循一對角綫分佈，自左下方而右上方，表示金條與美鈔間之關係屬於正直綫相關。反之，如材料之分佈雖爲直綫，但係自右下方而左上方，則爲負直綫相關。如材料在表中分佈係循曲綫而非循任何對角綫，則爲曲綫相關。此時即不能應用直綫相關技術來分析，因如此則不足以發現相關之確實程度。

相關之完全與否及其種類，自相關表中已可見其大概。在分組材料中，亦利用分散圖。或在分組材料中，可以 $X$，$Y$ 兩變數各上下組限爲標準，而作分散圖。



（第27圖）　黃金與美鈔市價相關分散圖

$X$: 黃金市價，單位國幣百萬元

$Y$: 美鈔市價，單位國幣千元

除直接依相關表中各小方格之次數作圖外，並可間接作分散圖。相關表中，不但有自變數 $X$ 即金條市價次數分配，及因變數 $Y$ 即美鈔市價次數分配。此外表中每一欄每一列各成一分配。試以自變數 $X$ 各中值爲準，則每一欄因變數各成一分配。例如 $X_1=3.600$，第一欄 $Y$ 之次數分配如下，并可依下列第69表而求平均數：

（第69表）

| $Y$ | $f$ | $fY$ |
|---|---|---|
| 6.50 | 6 | 39.00 |
| 6.90 | 0 | 0 |
| 7.30 | 0 | 0 |
| 7.70 | 0 | 0 |
| 總計 | 6 | 39.00 |

$$\overline{Y}=\frac{\Sigma fY}{\Sigma f}=\frac{39.00}{6}=6.500$$

$X_2=3,800$　　$\overline{Y}_2=6,557$

$X_3=4,000$　　$\overline{Y}_3=7,060$

$X_4=4,200$　　$\overline{Y}_4=7,500$

上列平均數稱爲行列平均數 (Array Means)，用$\bar{Y}_i$表示之，可展開爲$\bar{Y}_1$，$\bar{Y}_2$，$\bar{Y}_3$，$\bar{Y}_4$，求得後記入相關表中。以 $X$ 各變項中值爲標準，分別書於各欄之最下一方格中，如 $X=3.600$ 時，$\bar{Y}_1=6.50$，卽言以金價三百六十萬元爲準時，共計六次，其美鈔平均價格爲六千伍百元。$X_2=3.800$，$\bar{Y}_2=6.557$，卽言以金價爲三百八十萬元爲準時，共計七次，美鈔平均價格爲六千五百五十七元。同理金價四百萬元時，美鈔平均市價爲七千元；金價爲四百二十萬元時，美鈔平均市價爲七千五百元。此項序列平均數，亦可以 $X$ 各中值爲準，而書入上列之分散圖中，以 • 表示之。自代表序列平均數之圖觀之，二變數間所存在者，仍爲正的直綫相關。

分析直綫相關，計算直綫相關量數(Measures of Linear Correlation)之方法共有二種，卽最小二乘方法 (Method of Least Squares)，及積差法(Method of Product-deviaton)。以下二章分別論述之。

## 習　題

1　試依一百男生身高與體重相關表中各別次數製一分散圖。

2.　試依上題資料求因變數各序列平均數並作入上圖中。

# 第九章　最小二乘方法

## 第一節　未分組資料

分析兩組材料間之直綫相關，在分析程序上可分三步驟。換言之，即直綫相關之量數有三:

一　估計綫(Line of Estimation)或回應綫(Line of Regression)。

二　估計標準誤差(Standard Error of Estimate)。

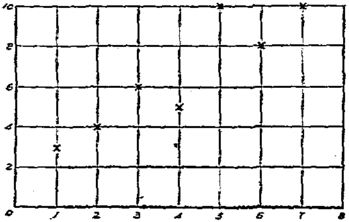
三　相關係數(Coefficient of Correlation)。

統計材料有未分組及分組之別。根據未分組材料，即根據 $X$，$Y$ 兩變數之各對變項原有數值，計算上列三種量數時，手續比較簡單。所謂分組材料，即原有相關材料項目過多，先經分類化簡，製成相關表，然後根據相關表計算上列三種相關量數，現就未分組材料舉例說明計算直綫相關三量數之程序。

一、估計直綫

(第70表)

| $X$ | $Y$ |
|---|---|
| 1 | 3 |
| 2 | 4 |
| 3 | 6 |
| 4 | 5 |
| 5 | 10 |
| 6 | 9 |
| 7 | 10 |




（第28圖）　分散圖

自上列之分散圖中，可見材料之分佈，係自左下方向右上方循對角綫分佈於一狹長之地區，因而可斷定二數列間之關係爲直綫相關，並進行求出一直綫以代表二者間之平均關係.求此直綫係根據直綫方程式，$Y=a+bX$，依此公式求得之直綫,稱爲回應綫，公式中之$X$爲自變數，$Y$爲因變數，即$Y$隨$X$而變。$X$，$Y$各變項原有值，可稱爲觀察值(Observed Valu-s)。依直綫式而得$Y$之新值稱之爲標準值 (Normal Values) 或估計值 (Estimated Values)，其意即根據$X$之原值而估計者。$a$，$b$爲常數，$a$代表相關直綫相切於$Y$軸時之高度，$b$代表相關直綫之斜度，$a$值愈大，相關綫切於$Y$軸之點愈高，$b$值愈大，則相關綫之斜度愈深，利用此直綫方程式求回應綫時，此式稱之爲回應方程式 (Equation of Regression)，或稱爲估計方程式(Equation of Estimation)。依此公式而得因變數$Y$之值，作圖時爲一直綫，稱爲回應綫或估計綫 (Line of Estimation) 或逕稱爲相關綫。

最小二乘方法，係指求直綫式中$a$，$b$兩常數之方法而言。在圖中確定各觀察點後，本可由觀察而作一直綫，以表示$X$，$Y$間之關係，但觀察欠

情，人各不同，不能得一致之結果，乃借重最小二乘方法。在一變數 $X$ 中，如已知各變項為 $X_1$，$X_2$，$X_3$ …… $X_N$ 時，現擬求得單一數值以代表各變項，此單一代表數值之條件有二：一即各變項與此代表值間正負離差之總和等於零，二即各變項與此代表值間之離差平方和為一最小值；合於此二條件時，此單一數值之代表性方最高。算術平均數合於此二條件，故在單一數列中，平均數為各變項之代表值，亦稱為最合於機率之值（The Most Probable Value）。現研究二組材料間之相關時，亦擬求出此最合於機率之數值。此值即為依最小二乘方法所求得之相關綫，此綫最足以代表二者平均關係，此直綫即以自變數 $X$ 之已知值為標準，以定出因變數 $Y$ 之值。由直綫式而定 $Y$ 之各值，稱為 $Y$ 之估計值或標準值，可用 $Y_c$ 以表示之，相當於一變數中之算術平均數，係代表集中趨勢者。至因變數 $Y$ 之原值，則為觀察值，以 $Y_o$ 表示之。以 $Y_c$ 即估計值為標準時，$Y_o$ 即各觀察值未必完全符合，即 $Y_o$ 各點不在 $Y_c$ 直綫上，是有離差或離中趨勢，此離差以 $y$ 表示之，即 $y=Y_o-Y_c$。依最小二乘方法所定出之相關綫，即所謂 $Y$ 之各估計值，亦合於上述算數平均數之二條件：其一為離差之和等于零，即 $\Sigma y=0$，其二為離差之平方和為一最小值，即 $\Sigma y^2<\Sigma d^2$，$d$ 係依相關綫外任一直綫以計算之離差。

以上為最小二乘方法之基本意義，現進而說明在 $Y=a+bX$ 公式中，求 $a$，$b$ 兩常數之方法。$Y=a+bX$，係總方程式，以兩項相關材料各變項代入之，則得許多觀察方程式（Observed Equations）：

| $Y$ | $=$ | $a$ | $+$ | $b$ | $X$ |
|---|---|---|---|---|---|
| 3 | $=$ | $a$ | $+$ | $b$ | $\cdot 1$ |
| 4 | $=$ | $a$ | $+$ | $b$ | $\cdot 2$ |
| 6 | $=$ | $a$ | $+$ | $b$ | $\cdot 3$ |

$$
\begin{aligned}
5 &= a + b.4 \\
10 &= a + b.5 \\
9 &= a + b.6 \\
10 &= a + b.7 \\
\hline
47 &= 7a + b.28
\end{aligned}
$$

上列七觀察式中，取任二者解之，可求得 $a$，$b$ 之值，但僅適合于該二式，而不能適合於其他各觀察式，必另設法以求之。上列共計七觀察式，現欲聯合之而得出二標準式 (Normal Equations)，解此二聯立標準方程式，即得 $a$，$b$ 最合於機率之值，適合於任一觀察式，此二標準式之求法如下：

第一標準式，係就各觀察式中以 $a$ 之係數乘式中之每一項，各個觀察式均經此手續之後，再求各式之和，即為第一標準式 (The First Normal Equation)。吾人一觀各觀察式中，$a$ 之係數均為 1，是無須以 1 乘各項，可逕求各式之和，自上表之最底下總數一行中即得之，為 $47=7a+28b$。此式中等號左邊者，實為因變數 $Y$ 之各變項之和，即 $Y_1+Y_2+Y_3+\cdots\cdots+Y_n=47$。等號右邊第一項俱為 $a$，其和為 $Na$，$N$ 為變項或觀察式之數，現在變項之數為 7，即 $N$ 等於 7，$Na=7a$。等號右邊第二項為 $b$ 乘自變數各變項之和，即 $bX_1+bX_2+bX_3+\cdots\cdots+bX_n$，即 $b\Sigma X$，等於 $28b$，即 $\Sigma X$ 等於 28。由是而得第一標準式之形式為 $\Sigma Y=Na+b\Sigma X$，依現在材料，該式為 $47=7a+28b$。

第二標準式，係就各觀察式中以 $b$ 之係數乘式中各項。各個觀察式均經此手續之後，再求各式之和，即為第二標準式 (The Second Normal Equation)。其要旨係以各式中之係數乘式中各項，然後求其總和。茲將第二兩標準式計算程序，表列如下：

| $XY$ | $=$ | $aX$ | $+$ | $bX^2$ |
|---|---|---|---|---|
| 3 | = | 1a | + | b.1 |
| 8 | = | 2a | + | b.4 |
| 18 | = | 3a | + | b.9 |
| 20 | = | 4a | + | b.16 |
| 50 | = | 5a | + | b.25 |
| 54 | = | 6a | + | b.36 |
| 70 | = | 7a | + | b.49 |
| 223 | = | 28a | + | b.140 |

上述第二標準式,係以$b$之係數乘各項而後再求各式之和,但$b$之係數爲自變數$X$之各變項,即$X_1$, $X_2$, $X_3\cdots$ $X_n$, 就第一觀察式而言,乘以$b$之係數即$X_1$, 實爲$X_1Y_1=aX_1+bX_1^2$, 其他各式以此類推。在所得之總和式中,等號左邊者實爲$X_1Y_1+X_2Y_2+X_3Y_3+\cdots+X_nY_n=\Sigma XY$, 現$\Sigma XY=223$。等號右邊之第一項爲$aX_1+aX_2+aX_3+\cdots+aX_n=a\Sigma X$, 現$a\Sigma X=28a$.等號右邊之第二項爲$bX_1^2+bX_2^2+bX_3+$

(第71表) 標準式計算表

| $X$ | $Y$ | $XY$ | $X^2$ |
|---|---|---|---|
| 1 | 3 | 3 | 1 |
| 2 | 4 | 8 | 4 |
| 3 | 6 | 18 | 9 |
| 4 | 5 | 20 | 16 |
| 5 | 10 | 50 | 25 |
| 6 | 9 | 54 | 36 |
| 7 | 10 | 70 | 49 |
| 28 | 47 | 223 | 140 |

$N$=成對變項之數=7

$\cdots\cdots+bX^2_n=b\Sigma X^2$，現 $b\Sigma X^2=140b$。由是而得第二標準式之形式為 $\Sigma XY=a\Sigma X+b\Sigma X^2$，依現在材料，該式為 $223=28a+140b$。

上文之要旨在解釋如何自許多觀察式以得二標準方程式。至吾人實際計算時只作一計算表即可，手續至簡。

第一標準式：$\Sigma Y=Na+b\Sigma X$

第二標準式：$\Sigma XY=a\Sigma X+b\Sigma X^2$

以上列第71表中所得各總數代入

$47=7a+28b$ (1)；$223=28a+140b\cdots(2)$

4(1)： $188=28a+112b\cdots(3)$

(2)−(3)： $35=\qquad 28b$

$$\therefore \qquad b=\frac{35}{28}=1.25$$

以 $b$ 之值代入 (1)：

$$47=7a+28\times1.25$$

即 $7a=47-35=12$

$\therefore \quad a=1.714$

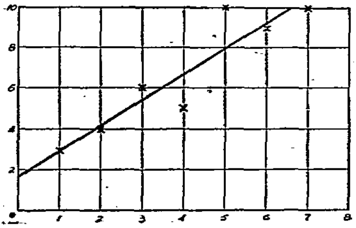
求得 $a$, $b$ 二常數之值，代入總方程式 $Y=a+bX$ 中，卽得直線方程式：$Y_c=1.714+1.25X$。此式卽表示 $Y$ 隨 $X$ 而變，每當 $X$ 變一單位時，$Y$ 變 1.25 單位，現有依此直綫式而計算 $Y$ 之標準值，并依此標準值而作圖表明由此標準值造成之相關直綫.其計算表如下：

## 二 估計標準誤差及估計界限

在一組材料中，平均數係代表集中趨勢之數值，而標準差則代表離中趨勢之數值。在兩組相關之材料中，相關直綫所代表者亦為集中趨勢或平均關係，但 $Y$ 之觀察值并不完全落在直綫上，卽觀察值與估計值之間有離

（第72表）　標準值計算表

| $X_o$ | $Y_o$ | $Y_c=1.714+1.25X$ |
|---|---|---|
| 1 | 3 | 2.964 |
| 2 | 4 | 4.214 |
| 3 | 6 | 5.464 |
| 4 | 5 | 6.714 |
| 5 | 10 | 7.964 |
| 6 | 9 | 9.214 |
| 7 | 10 | 10.464 |




（第29圖）

差，是亦可求出標準差以代表之，實卽表示估計直綫可靠之程度。在一組材料中，其標準差以$\sigma$代表之。現在兩組相關材料中，標準差之符號爲$S_y$，名稱爲估計標準誤差（Standard Error of Estimate）。以示與標準差有別，實則兩者之意義，算法與功用全同，所異者標準差係在一組材料之中，代表各項目之離中趨勢情形，亦卽表示平均數可靠之程度者。而估計標準誤差，則在相關材料中，代表各項目之離中趨勢，表示相關直綫可靠之程度。求估計標準誤差，一如求標準差，其過程爲離差一平方一平均

數一平方根:此四階段可用符號表示之而得公式:

1. 離差$=y=Y_o-Y_c$

2. 離差平方$=y^2=(Y_o-Y_c)^2$

3. 離差平方平均數$=\frac{\Sigma y^2}{N}=\frac{\Sigma(Y_o-Y_c)^2}{N}$

4. 離差平方平均數之根 $=S_p=\sqrt{\frac{\Sigma y^2}{N}}=\sqrt{\frac{\Sigma(Y_o-Y_c)^2}{N}}$

(第73表) 估計標準誤差計算表

| $X_o$ | $Y_o$ | $Y_c$ | $y=Y_o-Y_c$ | | $(Y_o-Y_c)^2$ |
|---|---|---|---|---|---|
| | | | + | − | |
| 1 | 3 | 2.964 | 0.036 | | 0.001296 |
| 2 | 4 | 4.412 | | 0.214 | 0.045796 |
| 3 | 6 | 5.464 | 0.536 | | 0.287296 |
| 4 | 5 | 6.714 | | 1.714 | 2.937796 |
| 5 | 10 | 7.964 | 2.036 | | 1.145296 |
| 6 | 9 | 9.214 | | 0.214 | 0.045796 |
| 7 | 10 | 10.464 | | 0.464 | 0.215296 |
| 總計 | | | 2.608 | 2.608 | 7.678572 |

已知$\Sigma(Y_o-Y_c)^2=7.678572$, $N=7$

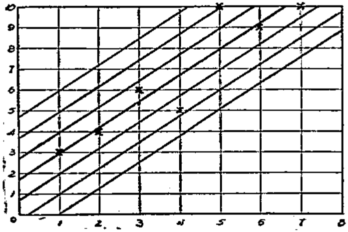
$$\therefore\quad S_p=\sqrt{\frac{\Sigma(Y_o-Y_c)^2}{N}}=\sqrt{\frac{7.678572}{7}}=1.047$$

求得估計標準誤差之功用,與標準差相同,係用以代表材料之集中與分散之情形者。$S_p$之值愈大,材料愈不集中;反之,$S_p$之值愈小,則材料愈集中;如$S_p$爲零時,則材料集中達於最高度,卽$Y_o$與$Y_c$完全相符合;而無任何離差,估計值直綫以外無任何觀察值據點。實際上材料有集中趨勢;同時亦有離中趨勢。如材料尙合乎或近於常態分配,則原值所定之各觀察點,雖以離中趨勢而不完全在估計直綫上,但其離中情形受一定限

度即所謂估計界限(Zones of Estimate)。此界限計算之方法，係根據估計直綫與估計標準誤差，現列計算表并製形圖如下：

（第 74 表） 估計界限計算表

| 自估計直線之距離 | | 完全常態分配 | 現在實際分配 | |
|---|---|---|---|---|
| 依 $S_y$ 計算 | 依因變數 $Y$ 單位計算 | 百分數 | 次數 | 百分數 |
| $Y_c\pm 0.6745S_y$ | $\pm 0.706$ | 50.00% | 5 | 71.43% |
| $Y_c\pm 1S_y$ | $\pm 1.047$ | 68.27% | 5 | 71.43% |
| $Y_c\pm 2S_y$ | $\pm 2.094$ | 95.45% | 6 | 85.56% |
| $Y_c\pm 3S_y$ | $\pm 3.141$ | 99.73% | 7 | 100.00% |




（第 30 圖 估計界限圖

上列第 74 表中所列完全常態分配之百分數，係自機率表中查得者。其意爲自估計直綫上加及下減 0.6745 倍估計標準誤差時，包括材料項目 50%；上加及下減一倍估計標準誤差時，包括材料項目 68%；上加及下減二倍估計標準誤差時，包括材料項目 95%；上加下減三倍估計標準誤差時，包括材料之99.73%。換言之，任一觀察值與標準值相較，雖有離差，但相差有一定之範圍，不能出乎回應直綫加減三倍標準誤差所得範圍之外。

三 相關係數

上文中已得二種結果，即估計直線及估計標準誤差。估計直線係代表$X$，$Y$二現象或二變數間之平均關係者。自變數$X$變動，因變數$Y$亦變動，估計標準誤差則表示因變數$Y$之觀察值與估計值間離差之情形；亦即各變項離開集中數值之程度。估計標準誤差，與標準差相同，其單位俱同於統計材料原有單位，俱爲絕對離中數值。標準差既爲絕對離中量數，用以表示一組材料之離中趨勢時，自無問題，但如吾人之目的在比較二組材料之離中趨勢，則有困難，解決之辦法爲將絕對數化爲相對數；即以一組材料之算術平均數除其標準差，結果產生一相對數值，稱爲離差係數(Coefficient of Variation)，已不受原有單位之限制，可與其他任何離差係數相比較。

在分析相關中，亦遇此問題。以估計標準誤差代表一對數列間之相關，雖無問題，但以之與另一對相關數列之估計標準誤差相比較時則有困難。現亦設法將絕對數值化爲相對數值，擺脫原有單位之限制，而成爲一種抽象數字，以代表相關之程度，英國統計學家皮生森(Karl Pearson)於此發明相關係數(Coefficient of Correlation)。其法之要旨，即以因度數$Y$之標準差$\sigma_y$除其估計標準誤差$S_y$。換言之，即以$S_y/\sigma_y$表示離中之程度，可稱爲離中係數。一極端情形，爲$S_y$之值等於零，即各項觀察值與估計值完全脗合，估計直線以外無任何觀察值所造成之點。估計標準誤差$S_y$既爲零，則其係數$S_y/\sigma_y$亦自爲零，即可斷言爲完全相關，或言相關之程度達於百分之百。反之，另一極端爲$S_y$之值等於$\sigma_y$之值，即估計標準誤差等於標準差，係表示估計直線上各點俱與因變數之平均數相等，即$Y_c$成爲一常數等於$\bar{Y}$，估計直線爲經由平均數$\bar{Y}$之直線，與$X$軸平行者。此即表示無論$X$之值如何變，$Y_c$之值不變，俱等於$\bar{Y}$。足證$X$，$Y$二變數

間無相關，故估計標準誤差等於標準差時，則離中係數等於 1，相關之程度等於零，即二數列間毫無聯繫。

以上係由離中係數 $S_p/\sigma_p$ 之值而推斷相關程度，其法似嫌迂迴。此係數之形式稍加改變，即可直接斷定相關之程度 改變之方法係將上公式改為：$\sqrt{1-\frac{S^2_p}{\sigma^2_p}}$。如為直綫相關時，此式結果稱之為相關係數，以 r 表示之：

$$r=\sqrt{1-\frac{S^2_p}{\sigma^2_p}}$$

一切結論與上文相同，但較為直接簡便。相關係數 r 之值，由估計標準誤差 $S_p$ 及標準差 $\sigma_p$ 二值而定。如 $S_p$ 等於零時，則 r 等於 1，或百分之百，表示為完全相關。反之如 $S_p$ 等於 $\sigma_p$ 時，則 r 等於零，表示根本無相關。是 $S_p$ 之極限為零與 $\sigma_p$，而 r 之極限為 1 與零。實際上二現象既非絕對孤立，亦非完全相依。相關係數 r 之值亦每介乎零與 1 之間，常為一小數。r 之值愈大，即表示相關程度愈深，估計直綫愈足以代表二數列間之平均關係。反之 r 之值愈小，即表示相關程度愈淺，估計直綫不足以代表二數列間之平均關係。此外，r 之值可為正，亦可為負。如自分散圖中已可斷定二變數之關係為正相關，則 r 之值亦為正。反之，如自分散圖即已斷定二數列之關係為負相關，則 r 之值亦自為負。

依上二段所舉之數字材料為例，已知 $S_p$ 等於 1.047，$S_p^2$ 等於 1.096939，$\sigma$ 及 $\sigma^2$ 可依簡便法以求之：

$$\sigma=\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}$$

$$\therefore\quad \sigma^2=\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2=\frac{367}{7}-\left(\frac{47}{7}\right)^2=\frac{2569-2209}{49}$$

$=\frac{360}{49}==7.3469$

$\therefore \quad r=\sqrt{1-\frac{S^2_{e}}{\sigma^2_{y}}}=\sqrt{1-\frac{1.096939}{7.3469}}=\sqrt{\frac{6.249961}{7.3469}}$

$=\sqrt{0.8507}=0.922$。

(第75表) 求$\sigma$之計算表——簡便法

| $Y$ | $Y^2$ |
|---|---|
| 3 | 9 |
| 4 | 16 |
| 6 | 36 |
| 5 | 25 |
| 10 | 100 |
| 9 | 81 |
| 10 | 100 |
| 47 | 367 |

## 第二節 分組資料

上節所述係利用最小二乘方法，直接就未分組材料，以求直綫相關之三種量數。前已言之，如材料中所包括成對變項之數目多及數百以至數千，直接求直綫相關量數，非但工作繁重，且易發生錯誤，是以須先加以分類歸併，製成相關表，然後計算相關量數。根據相關表求相關量數，有普通與簡捷二種方法，均可得到三種相關量數。

一 普通法

1. 回歸直綫 直綫方程式之形式為$Y=a+bX$，求解$a,b$二常數時，須先得二標準方程式。在未分組材料中，第一標準式為$\Sigma Y=Na+b\Sigma X$，第二標準式為$\Sigma XY=a\Sigma X+b\Sigma X^2$。在分類資料或相關表中，原有項目已

化爲次數，故在二標準方程式中，亦須將次數$f$加入補充，其形式如下：

第一標準式： $\Sigma fY = a\Sigma f + b\Sigma fX$

第二標準式： $\Sigma fXY = a\Sigma fX + b\Sigma fX^2$

現以民國三十六年元月上海黃金與金鈔市價相關表爲例．以說明二標準方程式及直綫總方程式之計算方法：

（第76表） 金條與美鈔市價相關表

$X$：金條十市兩價格：單位百萬元　　$Y$：美鈔一圓價格：單位千元

| $i_x$ / $i_y$ | | | 3.50—3.70 | 3.70—3.90 | 3.90—4.10 | 4.10—4.30 | 總計 |
|---|---|---|---|---|---|---|---|
| | $Y$ \ $X$ | | 3.60 | 3.80 | 4.00 | 4.20 | |
| | | $f_x$ / $f_y$ | 6 | 7 | 5 | 2 | 20 |
| 7.50—7.90 | 7.70 | 1 | | | | 1 | |
| 7.10—7.50 | 7.30 | 3 | | | 2 | 1 | |
| 6.70—7.10 | 6.90 | 4 | | 1 | 3 | | |
| 6.30—6.70 | 6.50 | 12 | 6 | 6 | | | |
| 總計 | | 20 | | | | | |

求解$a$，$b$二常數時，須依上列相關表另製一計算表如下：

（第77表） 標準方程式計算表

| $X$ | $Y$ | $f$ | $fX$ | $fY$ | $fX^2$ | $fXY$ |
|---|---|---|---|---|---|---|
| 3.6 | 6.5 | 6 | 21.6 | 39.0 | 77.76 | 140.40 |
| 3.8 | 6.5 | 6 | 22.8 | 39.0 | 86.64 | 148.20 |
| 3.8 | 6.9 | 1 | 3.8 | 6.9 | 14.44 | 26.22 |
| 4.0 | 6.9 | 3 | 12.0 | 20.7 | 48.00 | 82.80 |
| 4.0 | 7.3 | 2 | 8.0 | 14.6 | 32.00 | 58.40 |
| 4.2 | 7.3 | 1 | 4.2 | 7.3 | 17.64 | 30.66 |
| 4.2 | 7.7 | 1 | 4.2 | 7.7 | 17.64 | 32.34 |
| 總計 | | 20 | 76.6 | 135.2 | 234.12 | 519.02 |

上列第77表作法，即在相關表中，以自變數$X$各值爲標準，$X_1=3.6$，因變數$Y$與之有聯繫者計一項，即$Y=6.5$，次數$=1$，即將此三値分別寫入計算表$X$，$Y$，$f$三欄內；$X_2=3.8$時，$Y$與之發生聯繫者計二項，即$Y_1=6.5$，$Y_2=6.9$，次數爲6與1，亦列入計算表中；$X_3=4.0$時，$Y$與之發生聯繫者計二項，即$Y_2=6.9$及$Y_3=7.3$，次數爲3與2；$X_4=4.2$時，$Y$與之發生聯繫者計二項，即$Y_3=7.3$及$Y_4=7.7$，而次數1與1。$X$，$Y$與$f$三欄中之各項數字，俱由相關表中轉得而來，其他各欄則爲此三欄之乘積，不難計算，最後將各欄內之各項目分別相加，所得總數，記于各欄最底一列內，即可依之而解二標準方程式，求得常數$a$，$b$之値：

第一標準式： $\Sigma fY=a\Sigma f+b\Sigma fX$ ………………………… (1)

第二標準式： $\Sigma fXY=a\Sigma fX+b\Sigma fX^2$ ………………………… (2)

以上表中各總數代入二標準式，可得：

(1) $135.20=20.0a+76.60b$

(2) $519.02=76.6a+294.12b$

(1) $\div 4$ $33.80=5.0a+19.15b$ ………………… (3)

(2) $\div 2$ $259.51=28.3a+147.06b$ ………………… (4)

(4) $\times 5$ $1297.55=191.5a+735.300b$ ………………… (5)

(3) $\times 38.3$ $1294.54=191.5a+733.445b$ ………………… (6)

(5) $-$(6) $3.01=1.855b$

$\therefore\quad b=3.01/1.855=1.6226$

以$b=1.6226$代入(3)：

$33.8=5a+19.15\times 1.6526=5a+31.07279$

即 $5a=33.8-31.07279$

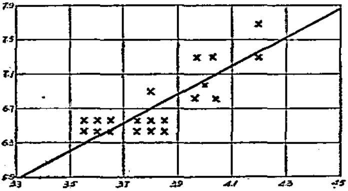
$\therefore\quad a=2.72721/5=0.54544$

以所得 $a$, $b$ 之值代入，則得回應直綫總方程式如下：

$$Y_c=0.54544+1.6226X$$

（第78表） 回應直綫計算表

| $X_o$ | $Y_c=0.54544+1.6226X$ |
|---|---|
| 3.6 | 6.38680 |
| 3.8 | 6.71132 |
| 4.0 | 7.03584 |
| 4.2 | 7.36036 |




（第31圖） 回應直綫配合圖

$X$: 金價　　$Y$: 美鈔價

由上列第31圖可見各觀察點之分佈，係由左下方向右上方，循一對角綫增進，爲正直綫相關。根據最小二乘方法所得之回應直綫，配合觀察值所據之各點，至爲恰當，是此直綫最足以代表 $X$, $Y$ 兩變數，即民國三十六年元月上海金條與美鈔行情間之關係。

2. 估計標準誤差　上表所得之回應直綫雖足以代表 $X$, $Y$ 間之平均關係或集中趨勢，但觀察值所據之各點并不完全落在直綫上，即各觀察值 $Y_o$ 與各標準值 $Y_c$ 中間有離差，此種離差須求出估計標準誤差以代表之：

（第79表）　美鈔市價估計標準誤差計算表

| $X_o$ | $Y_o$ | $Y_c$ | $f$ | $Y_o-Y_c$ | $f(Y_o-Y_c)$ + | $f(Y_o-Y_c)$ − | $f(Y_o-Y_c)^2$ |
|---|---|---|---|---|---|---|---|
| 3.6 | 6.5 | 6.387 | 6 | 0.113 | 0.678 | | 0.076614 |
| 3.8 | 6.5 | 6.711 | 6 | −.211 | | 1.266 | 0.267126 |
| 3.8 | 6.9 | 6.711 | 1 | 0.189 | 0.189 | | 0.035721 |
| 4.0 | 6.9 | 7.036 | 3 | −.136 | | 0.408 | 0.055488 |
| 4.0 | 7.3 | 7.036 | 2 | 0.264 | 0.528 | | 0.139392 |
| 4.2 | 7.3 | 7.360 | 1 | −.060 | | 0.060 | 0.003630 |
| 4.2 | 7.7 | 7.360 | 1 | 0.340 | 0.340 | | 0.115630 |
| 總計 | | | 20 | | 1.735 | 1.734 | 0.693541 |

$$S_y=\sqrt{\frac{\Sigma f(Y_o-Y_c)^2}{\Sigma f}}=\sqrt{\frac{0.693541}{20.00}}$$

$$=\sqrt{0.03467705}=0.18622$$　（單位千元）

上列所得之回應直線係代表 $X$，$Y$ 兩變數間之平均關係或集中趨勢者，而估計標準誤差，則係代表離中趨勢，即觀察值 $Y_o$ 與標準值 $Y_c$ 間之離差者。吾人可利用此估計標準誤差之值，以確定估計界限 (Zones of Estimate)。即以自變數 $X$ 或金條每一變項爲準時，而考察因變數 $Y$ 或美鈔價格變動之範圍。茲依下列表格計算之：

（第80表）　美鈔市價變動估計界限計算表

$S_y=0.18622$

| 自估計直線之距離：依估計標準誤差計 | 自估計直線之距離：依因變數計 | 常態分配應包括次數百分比 |
|---|---|---|
| $Y_c\pm0.6745S_y$ | $Y_c\pm0.1259$ | 50.00 |
| $Y_c\pm1S_y$ | $Y_c\pm0.1862$ | 68.27 |
| $Y_c\pm2S_y$ | $Y_c\pm0.3724$ | 95.45 |
| $Y_c\pm3S_y$ | $Y_c\pm0.5586$ | 99.73 |

（第 81 表） 美鈔市價變動範圍計算表

（以 $X$ 爲標準 依 $Y_c \pm 3S_p$ 爲範圍）

$3S_p = 0.5586$

| $X_o$ | $Y_c$ | $Y_c \pm 3S_p$ |
|---|---|---|
| 3.6 | 6.3868 | 5.8282—6.9454 |
| 3.8 | 6.7113 | 6.1527—7.2699 |
| 4.0 | 7.0358 | 6.4772—7.5944 |
| 4.2 | 7.3604 | 6.8018—7.9190 |

由上列第 81 表可知美鈔市價變動之範圍，以自變數 $X$ 即金價任一變項爲準時，美鈔市價之變動，均有一定之範圍，以 $(Y_c \pm 3S_p)$ 即回應直線加減三倍估計標準誤差表示之，如金價爲 3.6 時，美鈔市價變動之範圍爲5.8282—6.9454；金價爲3.8時，美鈔市價變動之範圍爲6.1527—7.2699；金價爲 4.0 時，美鈔市價變動之範圍爲 6.4772—7.5944；金價爲 4.2 時，美鈔市價變動之範圍爲 6.8018—7.9190。至第 79表即估計界限表中所列之百分數，係指常態分配而言，現在之分配並非常態，故所得之範圍，僅係概約數值。

3. 相關係數 相關係數 r 係由估計標準誤差 $S_p$ 及標準差 $\sigma_p$ 兩值求出，已知 $S_p = 0.1862$；$S^2_p = 0.034677$，標準差則可由下表求出。

（第 82 表） 標準差 $\sigma_p$ 計算表

$(\overline{Y} = 6.76)$

| $y$ | $f$ | $Y = Y - \overline{Y}$ | $f(Y - \overline{Y})$ | $f(Y - \overline{Y})^2$ |
|---|---|---|---|---|
| 6.5 | 12 | −.26 | −3.12 | 0.8112 |
| 6.9 | 4 | 0.14 | 0.56 | 0.0784 |
| 7.3 | 3 | 0.56 | 1.62 | 0.8748 |
| 7.7 | 1 | 0.94 | 0.94 | 0.8836 |
| | 20 | | 0 | 2.6480 |

$$\sigma_y=\sqrt{\frac{\Sigma f(Y-\overline{Y})^2}{\Sigma f}} \quad =\sqrt{\frac{2.648}{20.00}}$$

$$=\sqrt{0.1324} \qquad =0.364$$

$$r=\sqrt{1-\frac{S^2_y}{\sigma^2_y}} \quad =\sqrt{1-\frac{0.034677}{0.132400}}$$

$$=\sqrt{0.7381} \qquad =0.859$$

二 簡捷法

1. 回應直線 依普通法解二聯立標準式及估計方程式時，係根據自變數 $X$ 及因變數 $Y$ 各個原有中值計算，數字太大，手續繁重，因有簡捷法以補救之。其法之要旨，即在計算過程中，不用二變數之原有中值 而代之以自假定原點之組距離差 (Group-deviation from Assumed Origins)，現列出相關表并標明 $X$，$Y$ 二變數之組距離差，以便根據計算：

(第 83 表) 金條與美鈔市價相關表

| $i_y$ \ $i_x$ | | | | 3.50 3.70 | 3.70 3.90 | 3.9) 4.0 | 4.10 4.30 | 總計 |
|---|---|---|---|---|---|---|---|---|
| | $Y$ \ $X$ | | | 3.60 | 3.80 | 4.00 | 4.20 | |
| | | $f_y$ \ $f_x$ | | 6 | 7 | 5 | 2 | 20 |
| | | | $d'x$ \ $d'y$ | 0 | 1 | 2 | 3 | |
| 7.50 - 7.90 | 7.70 | 1 | 3 | | | | 1 | |
| 7.0 7.50 | 7.30 | 3 | 2 | | | 2 | | |
| 6.70 7.10 | 6.90 | 4 | 1 | | 1 | 3 | | |
| 6.30 6.70 | 6.50 | 12 | 0 | 6 | 6 | | | |
| 總計 | | 20 | | | | | | |

上列第 83 表係以 $X_1$ 即 3.6 為自變數 $X$ 之假定平均數，以 $Y_1$ 即 6.5 為因變數 $Y$ 之假定平均數. 自變數之假定平均數既為 $X_1$，則依此平均數

或原點而計算之離差自為 $(X-X_1)$，以 $d_x$ 表示之，即 $d_x=(X-X_1)$，其四變項應為 0, 0.20, 0.40, 0.60，稱為單位離差。(Deviations in Original Unit)。為化簡計，將各單位離差除以組距 $i_x$，並以 $d'_x$ 表示之，$d'_x=(X-X_1)/i_x$，組距 $i_x=0.2$，故自變數組距離差四變項為 0, 1, 2, 3。同理，因變數之假定原點為 $Y_1$ 即 6.5，依此原點即計算之離差為 0, 0.40, 0.80, 1.20，稱之為單位離差，以 $d_p$ 表示之。為化簡計，單位離差除以因變數之組距 $i_p$，並以 $d'_p$ 表示此種結果，則 $d'_p=(Y-Y_1)/i_p$，組距 $i_p=0.40$，故除後所得之四項組距離差，亦為 0, 1, 2, 3。

在普通法中，直綫總方程式及二標準方程式，俱根據 $X$, $Y$ 兩變數之各中值，即 $X_1$; $X_2$……及 $Y_1$, $Y_2$……等計算。現在簡捷法中，已不依兩變數之各中值，而改依二者之組距離差即 $d'_x$ 及 $d'_p$ 之各變項，此總方程式及二標準方程式之形式宜加以修正。

直綫回應總方程式：　$d'_p=a+bd'_x$

第一標準方程式：　$\Sigma fd'_p=a\Sigma f+b\Sigma fd'_x$

第二標準方程式：　$\Sigma fd'_xd'_p=a\Sigma fd'_x+b\Sigma f(d'_x)^2$

（第84表）　標準方程式計算表

| $d_x$ | $d'_p$ | $f$ | $fd'_x$ | $fd'_p$ | $fd'_xd'_p$ | $f(d'_x)^2$ |
|---|---|---|---|---|---|---|
| 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| 1 | 0 | 6 | 6 | 0 | 0 | 6 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 3 | 6 | 3 | 6 | 12 |
| 2 | 2 | 2 | 4 | 4 | 8 | 8 |
| 3 | 2 | 1 | 3 | 2 | 6 | 9 |
| 3 | 3 | 1 | 3 | 3 | 9 | 9 |
| 總計 | | 20 | 23 | 13 | 30 | 45 |

上列計算表(第84表)中,自變數組距離差,因變數組距離差及次數三欄,即 $d'_x$, $d'_y$, $f$ 三欄中之各項係由相關表中得來，如以 $d'_x$ 之第一變項,即 $d'_x=0$ 爲準時, $d_y$ 之第一變項即 $d_y=0$ 與之發生聯繫,二者公有之次數爲1,即以0, 0, 1三值記入表中第一列爲首三欄中，其他各列之數字亦均依此原則自相關表中轉錄;有此三欄數字後,其他四欄中之數字,則爲前三欄之乘積,自不難算出。各欄之最底一格乃其總數,計爲:

$\Sigma f \quad =20;$ $\qquad\qquad$ $\Sigma f d'_x d'_y=30$

$\Sigma f d'_x=23;$ $\qquad\qquad$ $\Sigma f(d'_x)^2=45$

$\Sigma f d'_y=13;$

以上列五總數代入原式,則得二標準方程式如下:

(1) $\quad 13=20a+23b$

(2) $\quad 30=23a+45b$

(1)×23: $\quad 299=460a+529b$ ……………………(3)

(2)×20: $\quad 600=460a+900b$ ……………………(4)

(4)−(3): $\quad 301= \qquad 371b$

$\therefore \qquad b=301/271=0.8113$

以 $b=0.8113$ 代入(1):

$$13=20a+23\times0.8113$$
$$=20a+18.6599$$
$$20a=13-18.6599=-5.6599$$
$$\therefore \quad a=-\frac{0.56599}{2}=-0.283$$

以 $a$, $b$ 之值代入，可得出直綫迴應總方程式,幷可依總方程式而求得因變數 $d'_y$ 之標準值 (Normal Values) 或估計值(Estimated Values,

總方程式: $\quad d'_y=-0.283+0.8113d'_x$

（第85表） 標準值計算表

| $d'_x$ | $d'_{pc}=0.283+0.8113d'_x$ |
|---|---|
| 0 | $-.2830$ |
| 1 | 0.5283 |
| 2 | 1.3396 |
| 3 | 2.1509 |

2. 估計標準誤差 以上已依 $X$，$Y$ 兩變數之假定原點組距離差而求得迴歸式及標準值，以此標準值即 $d'_{pc}$ 作圖所得爲一直綫，係代表 $d'_x$ 及 $d'_y$ 二變數之平均關係者。但 $d'_{po}$ 雖循直綫分佈，實并不在迴歸直綫上。即 $d'_{po}$ 與 $d'_{pc}$ 中間有離差，現求估計標準誤差以代表二者間平均離差情形，估計標準誤差之公式原爲： $S_y=\sqrt{\Sigma f(Y_o-Y_c)^2/\Sigma f}$，現係依假定原點組距離差計算，其公式自應爲：

$$S'_y=\sqrt{\frac{\Sigma f(d'_{po}-d'_{pc})^2}{\Sigma f}}$$

（第86表） 估計標準誤差 $S'_y$ 計算表

| $d'_x$ | $d'_{po}$ | $f$ | $d'_{pc}$ | $d'_{po}-d'_{pc}$ | $f(d'_{po}-d'_{pc})$ + | $f(d'_{po}-d'_{pc})$ − | $f(d'_{po}-d'_{pc})^2$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 6 | $-.2830$ | 0.2830 | 1.6980 | | 0.480534 |
| 1 | 0 | 6 | 0.5283 | $-.5283$ | | 3.1698 | 1.674605 |
| 1 | 1 | 1 | 0.5283 | 0.4717 | 0.4717 | | 0.222501 |
| 2 | 1 | 3 | 1.3396 | $-.3396$ | | 1.0188 | 0.345984 |
| 2 | 2 | 2 | 1.3396 | 0.6604 | 1.3208 | | 0.872256 |
| 3 | 2 | 1 | 2.1509 | $-.1509$ | | 0.1509 | 0.022771 |
| 3 | 3 | 1 | 2.1509 | 0.8491 | 0.8491 | | 0.720971 |
| | | | | | 4.3396 | 4.3395 | 4.339622 |

$$S'_y=\sqrt{\frac{\Sigma f(d'_{po}-d'_{pc})^2}{\Sigma f}}=\sqrt{\frac{4.339622}{20}}$$

$$=\sqrt{0.2169811} \qquad =0.4658$$

3. 相關係數　在分組材料中，相關係數之意義及公式與在未分組材料中相同。在分組材料中，依普通法及簡捷法所求相關係數 r 之公式亦同，所異者只爲在普通法中係根據因變數之原有單位，其公式爲：

$$r=\sqrt{1-\frac{S^2_p}{\sigma^2_p}}$$

在簡捷法中，係根據因變數之組距離差計算，其公式爲：

$$r=\sqrt{1-\frac{(S'_p)^2}{(\sigma'_p)^2}}$$

已知　$S'_p=0.4658$;　$(S'_p)^2=0.216981$

$(\sigma'_p)^2$ 可依下列計算表求出：

(第 87 表)

| $t_p$ | $Y$ | $f$ | $d'_p$ | $fd'_p$ | $f(d'_p)^2$ |
|---|---|---|---|---|---|
| 6.30—6.70 | 6.5 | 12 | 0 | 0 | 0 |
| 6.70—7.10 | 6.9 | 4 | 1 | 4 | 4 |
| 7.10—7.50 | 7.3 | 3 | 2 | 6 | 12 |
| 7.50—7.9 | 7.7 | 1 | 3 | 3 | 9 |
| 總計 | | 20 | | 13 | 25 |

$$\sigma'_p=\sqrt{\frac{\Sigma f(d'_p)^2}{\Sigma f}-\left(\frac{\Sigma fd'_p}{\Sigma f}\right)^2} \qquad =\sqrt{\frac{25}{20}-\left(\frac{13}{20}\right)^2}$$

$$=\sqrt{1.25-.4225} \qquad =\sqrt{0.8275}$$

$$=0.9097$$

已知　$(S'_p)^2=0.216981$;　$(\sigma'_p)^2=0.8275$

$$\therefore \quad r=\sqrt{1-\frac{0.216981}{0.827500}} \qquad =\sqrt{1-0.262213}$$

$=\sqrt{0.737\text{'}87} \qquad =0.859$

以上直綫相關三種係數，自回應直綫方程式起始，俱依簡捷法求出。但亦可利用還原過程，將依簡捷法所得之回應式化成依普通法之回應式，即依 $X$ 原有中值之回應式。還原之程序分爲二步驟：

（甲） 由組距離差改爲單位離差 單位離差係以 $d_x$ $d_y$ 表示之，而組距離差則用 $d'_x$，$d'_y$ 表示之，上列所得之回應式：$d'_y=-.283+0.811$ $d'_x$，係依組距離差求得者。$d'_x$ 係代表自變數之組距離差，自變數組距 $i_x$ 之値等於 0.20，簡捷法計算中由單位離差化爲組距離差時，係以組距 $i_x$ 除單位離差，即 $d'_x=d_x/i_x$；同理，$d_y$ 代表單位離差，$d'_y$ 代表除以因變數組距後之組距離差，則 $d'_y=d_y/i_y$，自變數之組距 $i_x$ 等於 0.20，因變數之組距 $i_y=0.40$，現以 $d'_x=d_x/i_x$ 及 $d'_y=d_y/i_y$ 代入已得之組距離差回應式，可得一新方程式如下：

$$\frac{d_y}{i_y}=-.283+0.8113\frac{d_x}{i_x}$$

以 $i_x=0.2$ 及 $i_y=0.4$ 代入則得：

$$\frac{d_y}{0.4}=-.283+0.8113\frac{d_x}{0.2}$$

式中各項俱乘以 0.4 則得：

$$d_y=-.1132+1.6226d_x$$

（乙） 由假定原點化爲 $X$，$Y$ 兩變數之中値 上列已將回應式由組距離差化爲單位離差，$d_x$，$d_y$ 即係代表兩變數之單位離差者。但單位離差俱由假定原點計算，現擬由假定原點改爲中値。已知自變數 $X$ 之假定原點爲 $X_1$ 即 $X=3.6$，依此原點計出之單位離差爲 $d_x=X-3.6$，因變數 $Y$ 之假定原點爲 $Y_1$ 即 6.5，依此原點計出之單位離差爲 $d_y=Y-6.5$，以之代入上式，則又得一新方程式如下：

$Y-6.5=-0.1132+1.6226(X-3.6)$

$Y=6.5-0.1132+1.6226X-5.8136$

$Y_c=0.5451+1.6226X$

上列為由還原而得之回應式最後之形式，係依自變數 $X$ 之各中值以估計因變數 $Y$ 之標準值 $Y_c$ 者。其數值與依普通法所得者全同，可以相互對照。

## 習　題

一、　(第 88 表)　民國三十六年四月上半月上海市米麵市價表

米：上等白粳每石價值法幣(千)元

麵：利朋洋粉每袋價值法幣(千元

| 日期 | 米 | 麵 | 日期 | 米 | 麵 |
|---|---|---|---|---|---|
| 1 | 107 | 74.0 | 9 | 14) | 92.3 |
| 2 | 108 | 75.0 | 10 | 144 | 94.5 |
| 3 | 110 | 76.0 | 11 | 140 | 93.4 |
| 4 | 1 4 | 78.0 | 12 | 114 | 94.0 |
| 5 | …… | …… | 13 | 星期日 | 星期日 |
| 6 | 星期日 | 星期日 | 14 | 144 | 97.0 |
| 7 | 128 | 83 | 15 | 153 | 100.0 |
| 8 | 135 | 93 | | | |

(一)試根據上列第 88 表所載三十六年四月上半月上海米麵市價資料，依最小二乘方法求下列三種直線相關量數：

1. 回應直線方程式及標準值，並繪出觀察值分散圖及標準值配合圖。

2. 估計標準誤差及估計界限，並以所得之各界限繪入上題所作之圖中。

3. 相關係數

二、試根據100名大學男生身高與體重相關表，依最小二乘法中普通與簡捷兩種方法分別計算上題中所列三種直線相關常數。

# 第十章　積差法

## 第一節　未分組資料

分析兩組資料間直綫相關之另一方法爲積差法，(The product-deviation Method)。此法之要旨在根據 $X,Y$ 二變數之離差乘積以求相關量數。根據積差法所求之結果，仍同于最小二乘方法所得之三種量數，兩法可謂殊途同歸，所異者只在依兩法求三種相關量數時，先後次序稍有不同耳。現仍依次根據未分組及分組材料分別舉例說明之，材料與上章所用者相同，藉資比較二法所得之結果。本節先根據未分組資料依次計算下列三種相關量數：

1. 相關係數

2. 回應直綫

3. 估計標準誤差

一　相關係數　相關係數仍以 r 表示之，求 r 有普通及簡捷二法：

1. 普通法

$$r=\frac{\Sigma xy}{N\sigma_x\sigma_y}$$

上列相關係數公式中各項目之意義，大致如下：$X$ 爲自變數，其平均數爲 $\bar{X}$；$Y$ 爲因變數，其平均數爲 $\bar{Y}$，$x=X-\bar{X}$，爲自變數離差，$y=Y-\bar{Y}$ 爲因變數之離差。$xy$ 爲離差乘積。$X$ 與 $Y$ 各爲變數，其離差 $x$ 與 $y$ 自亦各爲變數，離差乘積可展開爲變項 $x_1y_1, x_2y_2, x_3y_3, \cdots\cdots x_ny_n$ 離差乘積之和爲 $\Sigma xy$ $N$ 爲變項之總數，亦即 $X,Y$ 或 $x,y$ 成對之數，$\frac{\Sigma xy}{N}$ 爲離

差乘積之平均數，或簡稱爲積差均數，(Average Product-deviation)，可以代表 $X$，$Y$ 變數間相關之情形。但 $x$ 與 $y$ 各有其單位，須設法除掉，其法即二變數標準差之乘積即 $\sigma_x\sigma_y$ 除之，其形式爲：$\frac{\Sigma xy}{N\sigma_x\sigma_y}$ 上列結果爲一抽象數值，即稱爲相關係數(Coefficient of Correlation)，代二表變數間相關之程度，不再受自變數或因變數單位之限制。依此公式求相關係數之程序，可分下列各步驟：

(1) 求自變數 $X$ 及因變數 $Y$ 之平均數，即求 $\overline{X}$ 及 $\overline{Y}$

(2) 求 $X$，$Y$ 各變項之離差，離差等於各變項減平均數。換言之，即求

$$x_1,\ x_3,\ x_3,\cdots\cdots x_n \text{ 及 } y_1,\ y_2,\ y_3,\cdots\cdots y_n$$

(3) 求各對離差之乘積，即 $x_1y_1$，$x_2y_2$，$x_3y_3$，$\cdots\cdots x_ny_n$

(4) 求離差積之平均數，即 $\frac{\Sigma xy}{N}$

(5) 求 $X$，$Y$ 之標準差，即求 $\sigma_x$ 及 $\sigma_y$

(6) 以上列二標準差之乘積，除離差積之平均數。結果即爲相關係數，其公式如下：

$$r=\frac{\Sigma xy}{N\sigma_x\sigma_y}$$

計算時可製表格如下：

(第 89 表) 相關係數計算表

| $X$ | $Y$ | $x=X-\overline{X}$ | $y=Y-\overline{Y}$ | $x^2$ | $y^2$ | $xy$ |
|---|---|---|---|---|---|---|
| 1 | 3 | −3 | −3.714 | 9 | 13.793796 | 11.142 |
| 2 | 4 | −2 | −7.714 | 4 | 7.365796 | 5.428 |
| 3 | 6 | −1 | −.714 | 1 | 0.509796 | .714 |
| 4 | 5 | 0 | −1.714 | 0 | 2.937196 | 0 |
| 5 | 10 | 1 | 3.286 | 1 | 10.19.796 | 3.286 |
| 6 | 9 | 2 | 2.286 | 4 | 5.225796 | 4.572 |
| 7 | 10 | 3 | 3.286 | 9 | 10.297796 | 9.858 |
| 28 | 47 | 0 | 0 | 28 | 51.428572 | 35.000 |

$$\bar{X}=\frac{\Sigma X}{N}=\frac{28}{7}=4$$

$$\bar{Y}=\frac{\Sigma Y}{N}=\frac{47}{7}=6.714$$

$$\sigma_x=\sqrt{\frac{\Sigma x^2}{N}}=\sqrt{\frac{28}{7}}=\sqrt{4}=2$$

$$\sigma_y=\sqrt{\frac{\Sigma y^2}{N}}=\sqrt{\frac{51.428572}{7}}=2.7105$$

$$\therefore \quad r=\frac{\Sigma y}{N\sigma_x\sigma_y}=\frac{35}{7\times2\times2.7105}=0.922$$

2. 簡捷法 應用普通法求相關係數，係利用$X$，$Y$二變數與正平均數之離差，以求離差乘積，因離差每有小數位，計算之手續繁重，乃有簡便法，此法係假定$X$，$Y$二數列之平均數各爲零，依零而求二組中各變項之離差；再求各對離差之乘積 名爲離差乘積，實即直接求$X$，$Y$二組中各對變項原值之乘積，手續自簡，茲證其方法如下：

$X$：自變數 $\bar{X}$：爲自變數平均數

$Y$：因變數 $\bar{Y}$：爲因變數平均數

$x=X-\bar{X}$，爲自變數自平均數$\bar{X}$而計之離差。可展爲：

$x_1=X_1-\bar{X}$， $x_2=X_2-\bar{X}$， $x_3=X_3-\bar{X}$…………

$y=Y-\bar{Y}$：爲因變數自平均數$\bar{Y}$而計之離差，可展爲：

$y_1=Y_1-\bar{Y}$， $y_2=Y_2-\bar{Y}$， $y_3=Y_3-\bar{Y}$…………

$X'=0$，爲自變數之假定平均數

$Y'=0$，爲因變數之假定平均數

$d_x=X-X'$，爲自變數自假定平均數$X'$而計之離差，

$\because$ $X'=0$， $\therefore$ $d_x=X$，可展開爲：

$d_{x1}=X_1$， $d_{x2}=X_2$， $d_{x3}=X_3$，…… …$d_{xn}=X_n$

$d_y=Y-\overline{Y}'$，爲因變數自假定平均數 $\overline{Y}'$而計之離差。

$\because\ Y'=0\ \ \therefore\ d_y=Y$，可展開爲：

$d_{y1}=Y_1,\ \ d_{y2}=Y_2,\ \ d_{y3}=Y_3,\cdots\cdots\ \ \cdots d_{yn}=Y_n$

$C_x=$自變數中眞正平均數及假定平均數間之差額。

$C_x=\overline{X}-\overline{X}'=\overline{X}$

$C_y=$因變數中眞正平均數及假定平均數間之差額。

$C_y=\overline{Y}-\overline{Y}'=\overline{Y}$

$d_x-C_x=X-\overline{X}=x$

$\therefore\ d_x=C_x+x$，可展開爲：

$d_{x1}=C_{x1}+x_1,\ d_{x2}=C_{x2}+x_2,\ d_{x3}=C_{x3}+x_3\cdots\cdots\ \cdots\cdots$

$d_{xn}=C_{xn}+x_n$

$d_y-C_y=Y-\overline{Y}=y$

$\therefore\ d_y=C_y+y$，可展開爲：

$d_{y1}=C_{y1}+y_1,\ d_{y2}=C_{y2}+y_2,\ d_{y}=C_{y3}+y_3\cdots\cdots\cdots\cdots$

$d_{yn}=C_{xn}+x_n$

$d_xd_y=(x+C_x)(y+C_y)$ 可展爲：

$$d_{x1}d_{y1}=(x_1+C_x)(y_1+C_y)=x_1y_1+C_xy_1+C_yx_1+C_xC_y$$
$$d_{x2}d_{y2}=(x_2+C_x)(y_2+C_y)=x_2y_2+C_xy_2+C_yx_2+C_xC_y$$
$$d_{x3}d_{y3}=(x_3+C_x)(y_3+C_y)=x_3y_3+C_xy_3+C_yx_3+C_xC_y$$
$$\vdots\qquad\qquad\vdots\qquad\qquad\vdots\qquad\qquad\vdots\qquad\qquad\vdots$$
$$d_{xn}d_{yn}=(x_n+C_x)(y_n+C_y)=x_ny_n+C_xy_n+C_yx_n+C_xC_y$$

$$\Sigma d_xd_y=\qquad\Sigma xy+C_x\Sigma y+C_y\Sigma x+NC_xC_y$$

$\because\ \Sigma x=0,\qquad \Sigma y=0$

$\therefore\ \Sigma d_xd_y=\Sigma x+NC_xC_y$

$$\Sigma xy=\Sigma d_x d_y-NC_xC_y$$

式中每項俱以$N\sigma_x\sigma_y$除之:

$$\therefore\quad \frac{\Sigma xy}{N\sigma_x\sigma_y}=\frac{\Sigma d_x d_y}{N\sigma_x\sigma_y}-\frac{Nc_yc_x}{N\sigma_x\sigma_y}$$

$$=\frac{\frac{\Sigma d_x d_y}{N}-C_xC_y}{\sigma_x\sigma_y}$$

$$\therefore\quad r=\frac{\Sigma xy}{N\sigma_x\sigma_y}=\frac{\frac{\Sigma d_x d_y}{N}-C_xC_y}{\sigma_x\sigma_y}$$

$\because \overline{X}'=0$, $\therefore d_x=X-\overline{X}'=X$, 展開爲

$d_{x1}=X_1$, $d_{x2}=X_2$, $d_{x3}=X_3$, ………… $d_{xn}=X_n$,

$\Sigma d_x=\Sigma X$,

$\because \overline{Y}'=0$, $\therefore d_y=Y-\overline{Y}'=Y$, 展開爲:

$d_{y1}=Y_1$, $dY_2=Y_2$, $d_{y3}=Y_3$ ………… $d_{yn}=Y_n$,

$\Sigma d_x=\Sigma X$,

各變項離差乘積:

$$d_{x1}d_{y1}=X_1Y_1$$
$$d_{x2}d_{y2}=X_2Y_2$$
$$d_{x3}d_{y3}=X_3Y_3$$
$$\vdots\qquad\vdots$$
$$d_{xn}d_{yn}=X_nY_n$$
$$\Sigma d_x d_y=\Sigma XY$$

$$C_x=\overline{X}-\overline{X}'=\overline{X}=\frac{\Sigma X}{N}$$

$$C_y=\overline{Y}-\overline{Y}'=\overline{Y}=\frac{\Sigma Y}{N}$$

$$\sigma_x=\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}$$

$$\sigma_y=\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}$$

$$\therefore \quad r=\frac{\frac{\Sigma d_x d_y}{N}-C_x C_y}{\sigma_x \sigma_y}$$ 一式可改書爲:

$$r=\frac{\frac{\Sigma XY}{N}-\frac{\Sigma X}{N}\cdot\frac{\Sigma Y}{N}}{\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}\cdot\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}}$$

有此離差乘積簡便法公式後，仍根據普通法中所用之實例計算，以見此法之簡捷情形。

(第 90 表) 相關係數計算表

| $X$ | $Y$ | $XY$ | $X^2$ | $Y^2$ |
|---|---|---|---|---|
| 1 | 3 | 3 | 1 | 9 |
| 2 | 4 | 8 | 4 | 16 |
| 3 | 6 | 18 | 9 | 36 |
| 4 | 5 | 20 | 16 | 25 |
| 5 | 10 | 50 | 25 | 100 |
| 6 | 9 | 54 | 36 | 81 |
| 7 | 10 | 70 | 49 | 100 |
| 28 | 47 | 223 | 140 | 367 |

$$r=\frac{\frac{\Sigma XY}{N}-\frac{\Sigma X}{N}\cdot\frac{\Sigma Y}{N}}{\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}}=\frac{\frac{223}{7}-\frac{28}{7}\cdot\frac{47}{7}}{\sqrt{\frac{140}{7}-\left(\frac{28}{7}\right)^2}\cdot\sqrt{\frac{367}{7}-\left(\frac{47}{7}\right)^2}}$$

$$=\frac{\frac{1561-1316}{49}}{\sqrt{20-16}\sqrt{\frac{2569-2209}{49}}}=\frac{\frac{245}{49}}{\sqrt{4}\cdot\sqrt{\frac{360}{49}}}=\frac{\frac{245}{49}}{2\sqrt{7.346939}}$$

$$=\frac{5}{2\times 2.7105}=\frac{5}{5.421}=0.922$$

二　回應直綫

估計直綫仍依回應方程式作出。在最小二乘法中，回應方程式係由二聯立標準式解 $a$，$b$ 兩常數求出。現在積差法中，回應方程式係由比較二變數之離差求出，即比較 $x$，$y$。但 $X$，$Y$ 二變數各有單位，是 $x$，$y$ 兩離差亦各有單位，須各以標準差除之，化爲相對數值，不受原有單位之限制，方可比較。即求出 $\frac{x}{\sigma_x}$ 及 $\frac{y}{\sigma_y}$，如 $X$ 爲自變數，$Y$ 爲因變數，則 $Y$ 依 $X$ 而變之回應式爲：

$$\frac{y}{\sigma_y}=r\frac{x}{\sigma_x},$$

等號前後各以 $\sigma_y$ 乘之，則回應式爲：

$$y=r\frac{\sigma_y}{\sigma_x}$$

上列公式與最小二乘方法之回應式形式不同而實質相同。形式上不同之點爲最小二乘方法之回應方程式爲 $Y=a+bX$，現在積差法中之回應方程式爲 $y=r\frac{\sigma_y}{\sigma_x}x$，所以在此形式上之不同，因現在積差法中係根據二變數之離差計算，是計算之原點，爲 $X$ $Y$ 兩平均數所在點，則 $a$ 自爲零而不存在，此點可就最小二乘方法之公式以證明之。在最小二乘方法中，回應式爲 $Y=a+bx$，兩標準式爲：

1. $\Sigma Y=Na+b\Sigma X$

2. $\Sigma XY = a\Sigma X + b\Sigma X^2$

如計算之原點爲$\overline{X}$及$\overline{Y}$所在之點，換言之，卽依離差$x$, $y$計算時，則回應式爲，$y=a+bx$，兩標準式爲：

1. $\Sigma y = Na + b\Sigma x$

2. $\Sigma xy = a\Sigma x + b\Sigma x^2$

因$\Sigma x$及$\Sigma y$均等於零，自(1) $Na=0$亦卽$a=0$。自(2) $\Sigma xy = b\Sigma x^2$，$\therefore\ b=\frac{\Sigma xy}{\Sigma x^2}$。以上不過證明在積差法回應式中$a$何以不存在。二法另一點不同者，卽在最小二乘方法回應總式中$x$之係數爲$b$，在積差法中，$x$之係數改爲$r\frac{\sigma_p}{\sigma_x}$，此點至易證明，二者形異而實同。

已知在最小二乘方法中，如根據眞正平均數之離差計算$b=\frac{\Sigma xy}{\Sigma x^2}$，回應總式爲：$y=\frac{\Sigma xy}{\Sigma x^2}x$。積差法之回應總式則爲$y=r\frac{\sigma_p}{\sigma_x}x$，所謂兩者形異而實同，必須設法證明：

$$b=\frac{\Sigma xy}{\Sigma x^2}=r\frac{\sigma_p}{\sigma_x}$$

證明：1. $r\frac{\sigma_p}{\sigma_x}=\frac{\Sigma xy}{N\sigma_x\sigma_p}\cdot\frac{\sigma_p}{\sigma_x}=\frac{\Sigma xy}{N\sigma_x\sigma_x}$

$$2.\quad \frac{\Sigma xy}{N\sigma_x\sigma_x}=\frac{\Sigma xy}{N\sqrt{\frac{\Sigma x^2}{N}}\sqrt{\frac{\Sigma x^2}{N}}}=\frac{\Sigma xy}{N\cdot\frac{\Sigma x^2}{N}}=\frac{\Sigma xy}{\Sigma x^2}$$

$$\therefore\ r\frac{\sigma_p}{\sigma_x}=\frac{\Sigma xy}{\Sigma x^2}=b$$

故$y=bx$與$y=r\frac{\sigma_p}{\sigma_x}x$完全相同。$x=X-\overline{X}$, $y=Y-\overline{Y}$，以之代入積差法回應式中，則得新回應式如下：

$$Y-\overline{Y}=r\frac{\sigma_p}{\sigma_x}(X-\overline{X})$$

$$Y_c = \bar{Y} + r\frac{\sigma_y}{\sigma_x}(X - \bar{X})$$

$\bar{X}=4, \bar{Y}=6.714$, $r=0.922$, $\sigma_x=2$, $\sigma_y=2.7105$,

$$\therefore \quad Y_c = 6.714 + 0.922\ \frac{2.7105}{2}(X-4)$$

$$Y_c = 6.714 + 1.25(X-4)$$

$$Y_c = 6.714 + 125X - 5.00$$

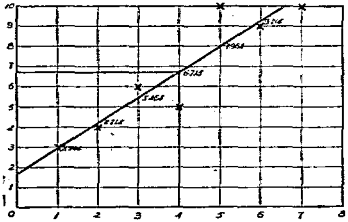
$$\therefore \quad Y_c = 1.714 + 1.25X$$

上式與在最小二乘方法中所得之結果，完全相同，可以對照，足見最小二乘方法與積差法完全相同。現依上列迴應總式，求出因變數標準值，並作一迴應直綫或估計直綫圖，藉以說明何以依平均數之離差計算時，常數 $a$ 即不存在：

（第91表） 標準值計算表

| $X$ | $Y_c=1.714+1.25X$ | $x$ | $Y_c=1.25x$ |
|---|---|---|---|
| 1 | 2.964 | −3 | −3.75 |
| 2 | 4.214 | −2 | −2.50 |
| 3 | 5.464 | −1 | −1.25 |
| 4 | 6.7.4 | 0 | 0 |
| 5 | 7.964 | 1 | 1.25 |
| 6 | 9.214 | 2 | 2.50 |
| 7 | 10.464 | 3 | 3.75 |

下圖中所表示者係 $Y$ 隨 $X$ 而變，由 $X$ 之觀察值以估計 $Y$ 之標準值，$X$ 觀察值每增加一單位時，$Y$ 之標準值則增加1.25單位。絕對變動相同，乃造成直綫。但吾人可换一觀點，即由兩平數 $\bar{X}$ 與 $\bar{Y}$ 之點觀之。$\bar{X}=4, \bar{Y}=6.714$，依此兩值為坐標而作一點，依 $X, Y$ 之原值觀之，此點固代表(4, 5,714)，但如依離差觀之，此點既由兩平均數所定，則 $x=X-\bar{X}=0, y=$




第32圖 回應直線配合圖

觀察值 $Yo$：叉號

標準值 $Y_c$：直線

$Y-\overline{Y}=0$，故自離差觀之，此點則代表(0，0)。同理，循直綫至左下方之各點及至右上方之各點，亦一方代表$X$,$Y$之原值，一方代表離差。如(3，5.464)所據之點亦可標爲(−1，−1.25)；(2，4.214)一點，可標爲(−2，−2.50)；(1，2.964)一點，亦可標爲(−3，−3.75)；(5，7.964)所據之點，亦可標爲(1，1.25)；(6，9.214)所據之點，亦可標爲(2，2.50)；(7，10.464)所據之點，亦可標爲(3，3.75)。可見：$Y_c=1.714+1.25X$及$y=1.25x$兩回應式所代表者係同一直綫。

三 估計標準誤差

依最小二乘方法求相關係數之公式爲：

$$r=\sqrt{1-\frac{S_y^2}{\sigma_y^2}}$$

此公式稍加變化，即可得積差法中求估計標準誤差之公式：

$$1.\quad r=\sqrt{1-\frac{S_y^2}{\sigma_y^2}}\;;\qquad 2.\quad r^2=1-\frac{S_y^2}{\sigma_y^2}\;;$$

3. $\sigma^2_p r^2 = \sigma_p^2 - S_p^2$　　4. $S_p^2 = \sigma_p^2 - \sigma_p^2 r^2$

5. $S_p^2 = \sigma_p^2(1-r^2)$　　6. $\therefore\ S_p = \sigma_p\sqrt{1-r^2}$

已知：$\sigma_p = 2.7105$； $r = 0.922$； $r^2 = 0.8507$

$\therefore\ S_p = 2.7105\ \sqrt{1-0.8507} = 2.7105\sqrt{0.1493}$

$= 2.7105 \times 0.3864 \quad = 1.047$

## 第二節　分組資料

如相關資料所包括之項目過多時，直接計算相關量數、手續太繁，乃先整理分類，製成相關表，然後再由相關表用積差法，以求三種相關量數，現仍以民國三十六年元月份上海金條與美鈔市價相關表爲例，以資與最小二乘方法比照。

一　普通法

1.相關係數

依普通法以求相關係數時，其公式爲：

$$r = \frac{\Sigma fxy}{\Sigma f\sigma_x\sigma_y}$$

上式之形式及意義，與未分組材料中相同。所異者，現係分類資料，爲複式次數表，各個項目，已化爲次數，故以$\Sigma f$代替$N$，即以次數總合代項數總合。在計算過程中，須將相關表加以補充，并須另有計算表分述如下：

在下列相關表（第92表)中，縱橫兩面俱加以補充。在橫列自變數卽金價次數分配中，自最頂一行起，表示標題者共計七行，依次序爲金價組距$i_x$，中值 $X$，次數 $f_x$，中值與次數乘積$fX$，其正離差$x$，離差與次數之乘積$fx$，及離差平方與次數乘積$fx^2$。在縱列因變數卽美鈔市價分配中，亦有相同之七欄，計爲鈔價組距$i_y$，中值$Y$，次數$f_y$，次數與中值乘

（第92表）金條與美鈔市價相關表

$X$：黃金價格組距中值

$Y$：美鈔價格組距中值

| $Y \backslash X$ | | | | | | 3.6 | 3.8 | 4.0 | 4.2 | 總計 |
|---|---|---|---|---|---|---|---|---|---|---|
| | $fy \backslash fx$ | | | | | 6 | 7 | 5 | 2 | 20 |
| | | $fY \backslash fX$ | | | | 21.6 | 26.6 | 20.0 | 8.4 | 76.6 |
| | | | $y \backslash x$ | | | −.23 | −.03 | 0.17 | 0.37 | |
| | | | | $fy \backslash fx$ | | −1.38 | −.21 | 0.85 | 0.74 | 0 |
| | | | | | $fy^2 \backslash fx^2$ | .3174 | .0063 | .1445 | .2138 | .742 |
| 7.7 | 1 | 7.7 | .9 | 0.94 | .8836 | | | | 1 | |
| 7.3 | 3 | 21.9 | .54 | 1.62 | .8748 | | | 2 | 1 | |
| 6.9 | 4 | 27.6 | .14 | 0.56 | .0784 | | 1 | 3 | | |
| 6.5 | 12 | 78.0 | −.26 | −3.12 | .8112 | 6 | 6 | | | |
| 總計 | | 135.2 | | 0 | 2.6280 | | | | | |

將$fY$與正離差$y$，次數與離差乘積$fy$及次數與離差平方乘積$fy^2$。相關係數公式中所需$X$與$Y$之平均數及標準差，即$\bar{X}$，$\bar{Y}$，$\sigma_x$及$\sigma_y$四項，與以前計算方法相同，由相關表補充行列中即可求出：

$$\bar{X}=\frac{\Sigma fX}{\Sigma f}=\frac{76.6}{20}=3.83$$

$$\bar{Y}=\frac{\Sigma fY}{\Sigma f}=\frac{135.2}{20}=6.76$$

$$\sigma_x=\sqrt{\frac{\Sigma f(X-\bar{Y})^2}{\Sigma f}}=\sqrt{\frac{0.742}{20.0}}=\sqrt{0.371}=0.1926$$

$$\sigma_y=\sqrt{\frac{\Sigma f(Y-\bar{Y})^2}{\Sigma f}}=\sqrt{\frac{2.648}{20}}=\sqrt{0.1324}=0.3638$$

依積差法求相關係數之公式爲一分數，其分子爲 $\Sigma fxy$，卽爲離差乘積總合。可直接利用相關表內各次數小方格中求之，凡一方格中有次數時，必有與此次數相當之離差 $x$ 及 $y$，如 $x_1=-.23$，$y_1=-.26$，$f=6$，則 $x_1y_1=-.23\times(-.26)=0.0598$，$fx_1y_1=6\times0.0598=0.3588$。在此次數方格中計記三數字，最上爲次數卽 6，其次爲離差乘積，卽 0.0598，最後爲次數乘離差乘積，卽0.3588。凡有次數之方格，均得求其乘積 $fxy$，最後再求此次數與離差乘積之總和，卽爲 $\Sigma fxy$。求此總和時，可縱橫均計算之，卽將橫列每行中之 $fxy$ 求一總數，記入最右總數欄內，最後求其總和，并將縱列每欄中之 $fxy$ 求一總數記入最下總數行內，最後求其總和，二法所

第 93 表 積差總和計算表

| $x$ | $y$ | $f$ | $xy$ + | $xy$ − | $fxy$ + | $fxy$ − |
|---|---|---|---|---|---|---|
| −.23 | −.26 | 6 | 0.0598 | | .3588 | |
| −.03 | −.26 | 6 | 0.0078 | | .0468 | |
| −.03 | 0.14 | 1 | | −.0042 | | −.0042 |
| 0.17 | 0.14 | 3 | 0.0238 | | .0714 | |
| 0.17 | 0.54 | 2 | 0.0918 | | .1836 | |
| 0.37 | 0.54 | 1 | 0.1998 | | .1998 | |
| 0.37 | 0.94 | 1 | 0.3478 | | .3478 | |
| | | | | | 1.2082 | −.0042 |

得結果必相同，否則計算有誤。

又可由相關表中離差 $x$，$y$ 及次數 $f$ 三者之配合而另列一計算表，反較清醒。由上列第 92 表可知離差乘積總和，

$$\Sigma fxy=1.2082-0.0042=1.204$$

$$r=\frac{\Sigma fxy}{\Sigma f\sigma_x\sigma_y}=\frac{1.204}{20\times0.1926\times0.364}=\frac{1.20400}{1.40025}=0.859$$

2.回應直線

在上節未分組材料中，已舉出積差法之回應直線方程式爲 $y=r\frac{\sigma_x}{\sigma_y}x$，或 $Y=\overline{Y}+r\frac{\sigma_y}{\sigma_x}(X-\overline{X})$，現在分組材料中，此公式無改，已知：

$r=0.859$； $\sigma_x=0.1926$； $\sigma_y=0.364$

$\overline{X}=3.83$； $\overline{Y}=6.76$；

∴ 回應直線方程式：

$$Y=6.76+0.859\cdot\frac{0.3638}{0.1926}(X-3.83)$$

$$Y_c=6.76+1.6226(X-3.83)=6.76+1.6226X-6.21456$$

$$\therefore\ Y_c=0.544+1.6226X$$

上式與最小二乘方法中所得者，完全相同，足見無誤。

3.估計標準誤差

依積差法求估計標準誤差 $S_y$ 之公式爲：

$$S_y=\sigma_y\sqrt{1-r^2}\qquad \text{已知}$$

$\sigma_y=0.364$； $r=0.859$； $r^2=0.737787$

$$\therefore\ S_y=0.364\sqrt{1-0.737787}=0.364\sqrt{0.62213}$$

$$=0.364\times0.512$$

$$=0.18637$$

## 二 簡捷法

### 1. 相關係數

現在分組材料或相關表中，依簡捷法以求相關係數時，仍與在未分組材料中相同，係根據 $X$, $Y$ 兩變數之假定平均數及自此假定平均數而得之組距離差以求之。在相關係數公式中，除補充之以次數 $f$ 一項因素外，餘皆與未分組資料中相同，其形式如下：

(第 94 表) 金條與美鈔市價相關表

$X$：金條價格組距中值

$Y$：美鈔價格組距中值

| $Y$ \ $X$ | | | | | 3.6 | 3.9 | 4.0 | 4.2 | 總計 |
|---|---|---|---|---|---|---|---|---|---|
| | $f_y$ \ $f_x$ | | | | 6 | 7 | 5 | 2 | 20 |
| | | $d_y$ \ $d_x$ | | | 0 | 1 | 2 | 3 | |
| | | | $fd_y$ \ $fd_x$ | | 0 | 7 | 10 | 6 | 23 |
| | | | | $fd_y^2$ \ $fd_x^2$ | 0 | 7 | 20 | 18 | 45 |
| 7.7 | 1 | 3 | 3 | 9 | | | | 1 | |
| 7.3 | 3 | 2 | 4 | 12 | | | 2 | 1 | |
| 6.9 | 4 | 1 | 4 | 4 | | 1 | 3 | | |
| 6.5 | 12 | 0 | 0 | 0 | 6 | 6 | | | |
| 總計 | 20 | | 13 | 25 | | | | | |

$$r=\frac{\frac{\Sigma f d_x d_y}{\Sigma f}-C_x C_y}{\sigma_x \sigma_y}$$ 亦可書爲:

$$r=\frac{\frac{\Sigma f d_x d_y}{\Sigma f}-\frac{\Sigma f d_x}{\Sigma f}\cdot\frac{\Sigma f d_y}{\Sigma f}}{\sqrt{\frac{\Sigma f d_x^2}{\Sigma f}-\left(\frac{\Sigma f d_x}{\Sigma f}\right)^2}\cdot\sqrt{\frac{\Sigma f d_y^2}{\Sigma f}-\left(\frac{\Sigma f d_y}{\Sigma f}\right)^2}}$$

依上列公式求相關係數時，其先決各項如 $\Sigma f$，$\Sigma f d_x$，$\Sigma f d_y$，$\Sigma f d_y^2$、$\Sigma f d_x^2$，等，均可自相關表之補充欄內求出 手續同前，較爲困難者，仍爲 $\Sigma f d_x d_y$，須由相關表另列表以求之。

(第95表) 積差總額和 $\Sigma f d_x d_y$ 計算表

| $d_x$ | $d_y$ | $f$ | $d_x d_y$ | $f d_x d_y$ |
|---|---|---|---|---|
| 0 | 0 | 6 | 0 | 0 |
| 1 | 0 | 6 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 3 | 2 | 6 |
| 2 | 2 | 2 | 4 | 8 |
| 3 | 2 | 1 | 6 | 6 |
| 3 | 3 | 1 | 9 | 9 |
| | | 20 | | 30 |

由上列第 94，95 兩表：已知各總數如下:

$\Sigma f=20$， $\Sigma f d_x=23$， $\Sigma f d_y=13$

$\Sigma f d_x^2=45$， $\Sigma f d_y^2=25$， $\Sigma f d_x d_y=30$，

$$\therefore r=\frac{\frac{30}{20}-\frac{23}{20}\cdot\frac{13}{20}}{\sqrt{\frac{45}{20}-\left(\frac{23}{20}\right)^2}\cdot\sqrt{\frac{25}{20}-\left(\frac{13}{20}\right)^2}}=\frac{1.5-1.15\times 0.65}{\sqrt{2.25-(1.15)^2}\sqrt{1.25-(0.65)^2}}$$

$$=\frac{1.5-0.7475}{\sqrt{2.25-1.3225}\sqrt{1.25-0.4225}}=\frac{0.7525}{\sqrt{.9275}\sqrt{.8275}}$$

$$=\frac{0.7525}{0.963\times 0.9097}=\frac{0.75250}{0.87604}$$

$$=0.859$$

2. 回應直線方程式

積差法中之回應直線方程式，在上節中業已說明，係由 $X,Y$ 兩變數之眞正平均數單位離差即 $x$，$y$ 求出，雖在簡捷法中，亦不能直接由 $X,Y$ 兩變數之假定平均數組距離差計算。換言之，即在簡捷法中，亦不能根據 $d_x$ 及 $d_y$，而須根據 $x$ 與 $y$ 以求直綫回應方程式也：

$$y=r\frac{\sigma_y'\cdot i_y}{\sigma_x'\cdot i_x}x$$

已知：$r=0.859$，$\sigma_x'=0.963$，$i_x=0.2$，$\sigma_y'=0.9097$，$i_y=0.4$；

$$\sigma_x=i_x\sigma_x'=0.2\times 0.963=0.1926$$

$$\sigma_y=i_y\sigma_y'=0.4\times 0.9097=0.3638$$

$$y=r\frac{\sigma_y}{\sigma_x}x$$

$$y=0.859\frac{0.3538}{0.1926}x=\frac{0.3125}{0.1926}x=1.6226\,x$$

$\bar{X}=3.83$；$\bar{Y}=6.76$；$x=X-3.83$；$y=Y-6.76$

$\therefore$ $y=1.6226\,x$，可改書爲：

$$Y-6.76=1.6226(X-3.83)$$

$Y_c=6.76+1.6226X-6.21456$

$Y_c=0.54544+1.6226X$

3. 估計標準誤差

$S_y'=\sigma_y'\sqrt{1-r^2}$，已知：

$\sigma_y'=0.9097$，$r=0859$，$r^2=0.737787$，

$\therefore S_y'=0.9097\sqrt{1-0.737787}$

$=0.9097\sqrt{0.262213}=0.9097\times0.512=0.4658$

## 習 題

一、試根據上章習題中所載三十六年四月上半月上海市米麵市價未分組資料，依積差法中之普通與簡捷兩種方法，分別求下列三種直綫相關量數：

(一) 相關係數

(二) 回應直綫

(三) 估計標準誤差

二、試根據100名大學男生身高與體重相關表，依積差法中之普通法求上題所列三種直綫相關量數，

三、試依上題相關表依簡捷法求相關係數。

# 第十一章　曲綫相關

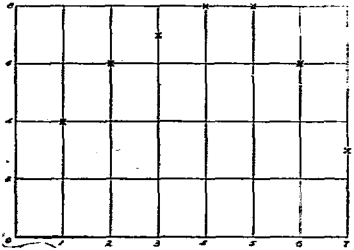
## 第一節　未分組資料

以上二章所分析者為直綫相關。依兩項相關中之各對變項製成分散圖，即可斷定兩變數間之關係是否為直綫相關。如自分散圖中吾人可斷定兩種資料之關係為直綫相關，自可根據以上兩章中所述之方法以分析之，但如在分散圖各對變項所佔各點之分佈，非循直綫，乃循一曲綫，是為曲綫相關（Curvilinear Correlation），不宜根據直綫相關之方法，以求直綫相關量數，而須另闢途徑，以求曲綫相關量數（Measures of Curvilinear Correlation）。曲綫相關之存在，由下列之資料表及分散圖即可見出：

（第96表）

| $X$ | $Y$ |
|---|---|
| 1 | 4 |
| 2 | 6 |
| 3 | 7 |
| 4 | 8 |
| 5 | 8 |
| 6 | 6 |
| 7 | 3 |

由下列分散圖中之各點觀之，可見$X$，$Y$兩變數間之關係為曲綫相關，而非直綫相關。無論在自然現象或社會現象中，曲綫相關之實例均甚多。以經濟學中效用遞減定律言之，試以$X$代表消費單位，以$Y$代表效用，消費增加，則效用自隨之增加，即$X$加，$Y$隨之而加。但消費單位繼續增加，而效用總和$Y$雖亦增加，但所增加之邊際效用則逐漸減少，即$X$




（第33圖） 分散圖

加十分之一，$Y$ 所加者不及十分之一，是卽邊際效用或最末效用漸減。消費單位再繼續增加，最後因已超過飽和點，慾望已完全滿足，增加消費，非但無益而且有害，非但無邊際效用，而且有反作用，使總效用逐漸減少，卽 $X$ 增加，$Y$ 減少。

凡遇上述情形，卽 $X$ 繼續增加，$Y$ 則始而增加，繼而減少，或始而減少，繼而增加，均爲曲綫相關。在曲綫相關之資料中，仍可計算三種直綫相關量數，但所得之迴應直綫，非爲最近似數值，換言之，卽標準值 $Y_c$ 與觀察值 $Y$，中間不十分配合。二者旣不配合，則估計標準誤差必大，而相關係數必小。但吾人不能由於相關係數之小卽斷定 $X$，$Y$ 兩變數間相關程度之淺。蓋兩者間之關係爲曲綫相關，根本不宜計算直綫相關量數，相關之程度，根本上不能由相關係數以代表也。

曲綫相關所計算之量數仍有三種，與直綫相關大同而小異：

一、二次拋物式迴應曲綫

二、估計標準誤差

三、相關指數

以上三種量數，可就未分組資料求之，亦可自相關表以求之，但俱只能依最小二乘方法計算，而不能用積差法。現依上列之未分組資料而計算三種曲綫相關量數如下：

一、估計方程式

現在之回應綫或估計綫既爲二次拋物曲綫，總方程式爲 $Y=a+bX+cX^2$，因有三常數，自須有三標準式方可解出。三者之形式如下：

1. $\Sigma Y=Na+b\Sigma X+c\Sigma X^2$

2. $\Sigma XY=a\Sigma X+b\Sigma X^2+c\Sigma X^3$

3. $\Sigma X^2Y=a\Sigma X^2+b\Sigma X^3+c\Sigma X^4$

（第97表） 計算表

| $X$ | $Y$ | $XY$ | $Y^2$ | $X^2Y$ | $X^3$ | $X^4$ |
|---|---|---|---|---|---|---|
| 1 | 4 | 4 | 1 | 4 | 1 | 1 |
| 2 | 6 | 12 | 4 | 24 | 8 | 16 |
| 3 | 7 | 21 | 9 | 63 | 27 | 81 |
| 4 | 8 | 32 | 16 | 128 | 64 | 256 |
| 5 | 8 | 40 | 25 | 200 | 125 | 625 |
| 6 | 6 | 36 | 36 | 216 | 216 | 1296 |
| 7 | 3 | 21 | 49 | 147 | 343 | 2401 |
| 28 | 42 | 166 | 140 | 782 | 784 | 4676 |

由上列第97表中已知：

$N=7$，$\Sigma X=28$，$\Sigma Y=42$，$\Sigma XY=166$

$\Sigma X^2=140$，$\Sigma X^2Y=782$；$\Sigma X^3=784$，$\Sigma X^4=4676$

以上列各總數代入，即得三聯立方程式如下：

(1) $42=7a+28\quad+140c$

$$(2)\quad 166 = 28a + 140b + 784c$$

$$(3)\quad 782 = 140a + 784b + 4676c$$

$$4(1):\quad 168 = 28a + 112b + 560c \cdots\cdots(4)$$

$$(2)-(4):\quad -2 = 28b + 224c \cdots\cdots(A)$$

$$20(1):\quad 840 = 140a + 560b + 2800c \cdots\cdots(5)$$

$$(3)-(5):\quad -58 = 224b + 1876c \cdots\cdots(B)$$

$$8(A):\quad -16 = 224b + 1792c \cdots\cdots(C)$$

$$(B)-(C):\quad -42 = 84c$$

$$\therefore\quad c = -\frac{42}{84} = -.5$$

以 $c$ 之值代入 $(A)$：

$$-2 = 28b - 224 \times 0.5 = 28b - 112$$

$$28b = 112 - 2 = 110$$

$$\therefore\quad b = \frac{110}{28} = 3.93$$

以 $b, c$ 之值代入 (1)：

$$42 = 7a + 28 \times 3.93 - 140 \times 0.5$$

$$42 = 7a + 110.04 - 70 = 7a + 40.04$$

$$7a = 42 - 40.04 = 1.96$$

$$\therefore\quad a = \frac{1.96}{7} = 0.28$$

以 $a, b, c$ 三常數之值代入，卽得吾人所求之二次拋物估計總方程式如下：

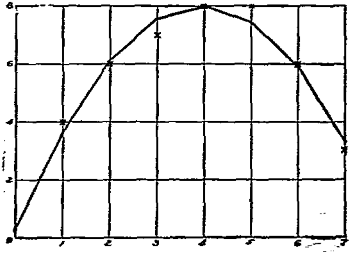
$$Y_c = 0.28 + 3.93X - 0.5X^2$$

依上列總方程式可求標準值如下（第 98 表），並可作入分散圖中，以見由

公式所得之標準值與觀察值間配合情形：

（第 98 表）　標準值計算表

| $X_o$ | $Y_o$ | $Y_c=0.28+3.93X-0.5X^2$ |
|---|---|---|
| 1 | 4 | 3.71 |
| 2 | 6 | 6.[illegible]4 |
| 3 | 7 | 7.57 |
| 4 | 8 | 8.00 |
| 5 | 8 | 7.43 |
| 6 | 6 | 5.86 |
| 7 | 3 | 3.29 |




（第 34 圖）　估計值與觀察值比較圖

由上圖中可見代表標準值之曲綫與代表原值之各點，在圖中至爲接近，各觀察值或在迴歸曲綫上，或離曲綫極近，足徵 $X$，$Y$ 間之相關用二次拋物綫代表之最爲恰當。

二、估計標準誤差

估計標準誤差計算之方法與直綫相關中完全相同，其功用亦同，直接

代表觀察値與估計値之離差情形，間接表示$X$，$Y$各對變項間之平均關係者，故估計標準誤差之値愈小，則估計曲綫配合各觀察値之程度亦愈高。茲根據下列計算表求之：

(第99表)　估計標準誤差計算表

| $X_o$ | $Y_o$ | $Y_c$ | $Y_o-Y_c$ + | $Y_o-Y_c$ − | $(Y_o-Y_c)^2$ |
|---|---|---|---|---|---|
| 1 | 4 | 3.71 | .29 | | .0841 |
| 2 | 6 | 6.14 | | .14 | .0196 |
| 3 | 7 | 7.57 | | .57 | .3249 |
| 4 | 8 | 8.00 | 0 | | 0 |
| 5 | 8 | 7.43 | .57 | | .3249 |
| 6 | 6 | 5.86 | .14 | | .0196 |
| 7 | 3 | 3.29 | | .29 | .0841 |
| 總計 | | | 1.00 | 1.00 | 0.8572 |

$$S_p=\sqrt{\frac{\Sigma(Y_o-Y_c)^2}{N}}=\sqrt{\frac{0.8572}{7}}=\sqrt{0.122457}$$

$$=0.35$$

三、相關指數

在直綫相關中，以相關係數r代表$X$，$Y$兩變數間相關之程度，現在曲綫相關中，則以相關指數(Index of Correlation)代表$X$，$Y$間相關之程度。相關指數之符號爲$\rho$(rho)。在直綫相關中，自變數爲$X$，求相關係數時，係將估計標準誤差之絕對値除以標準差而化爲相對値。現在曲綫相關中相關指數仍係由估計標準誤差$S_p$及標準差$\sigma_p$二量數以求之，即相關指數之公式如下：

$$\rho_{yx}=\sqrt{1-\frac{S_p^2}{\sigma_p^2}}$$

已知$S_p=0.35$，$S_p^2=0.122457$，至于標準差及其平方即$\sigma_p$及

$\sigma_y^2$ 則可由下表計算之:

(第100表) 因變數之標準差計算表

| $Y$ | $Y^2$ |
|---|---|
| 4 | 16 |
| 6 | 36 |
| 7 | 49 |
| 8 | 64 |
| 8 | 64 |
| 6 | 36 |
| 3 | 9 |
| 42 | 274 |

$$\sigma_y=\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}=\sqrt{\frac{274}{7}-\left(\frac{42}{7}\right)^2}$$

$$=\sqrt{\frac{274}{7}-36}=\sqrt{\frac{274-252}{7}}$$

$$=\sqrt{\frac{22}{7}}=\sqrt{3.142857}=1.772$$

$$\sigma_y^2=3.142857,\quad S_y^2=0.122457$$

$$\rho_{xy}=\sqrt{1-\frac{0.122457}{3.142857}}=\sqrt{\frac{3.020400}{3.142857}}$$

$$=\sqrt{0.961036}=0.98$$

## 第二節　分組資料

根據分組相關資料即相關表,而求曲線相關量數時,方法與上節所用者相同。其惟一異點,即分組後有次數一因素,須加以補充,同時以資料較繁,計算三種相關量數時,分有普通及簡捷二種方法:

方法：

一、普通法

在曲線相關中，係用二次拋物式曲線以代表 $X, Y$ 兩變數間之平均關係，此曲線之總方程式爲：$Y=a+bX+cX^2$，式中既有 $a, b, c$ 三常數，求解時自須有三標準方程式：

1. $\Sigma fY=a\Sigma f+b\Sigma fX+c\Sigma fX^2$
2. $\Sigma fXY=a\Sigma fX+b\Sigma fX^2+c\Sigma fX^3$
3. $\Sigma fX^2Y=a\Sigma fX^2+b\Sigma fX^3+c\Sigma fX^4$

（第100表）施用肥料與產麥數量相關表

$X$：每畝土地施用肥料數量　　單位：磅

$Y$：每畝土地出產小麥數量。　單位：蒲耳

| $f_x$ \ $f_y$ | | | 0—40 | 40—80 | 80—120 | 120—160 | 總計 |
|---|---|---|---|---|---|---|---|
| | $X$ \ $Y$ | | 20 | 60 | 100 | 140 | |
| | | $f_x$ \ $f_y$ | 4 | 3 | 6 | 3 | 15 |
| 24—32 | 28 | 3 | | 1 | 2 | | |
| 16—24 | 20 | 7 | | 2 | 3 | 2 | |
| 8—16 | 12 | 3 | 2 | | | 1 | |
| 0—8 | 4 | 2 | 2 | | | | |
| 總計 | | 15 | | | | | |

由上所列每畝土地施用肥料及出產小麥數量相關表觀之，可知二者間之關係爲曲綫相關，小麥產量雖隨肥料數量而變，但肥料繼續增加，小麥產量始而增加，終而減少，在相關表中即可見出次數配係循一弧形曲綫，其平均關係自爲曲綫式，現以二次拋物曲綫代表之，結果必至恰當，但依普通法計算時，手續過於繁重，故用之者較少。

二、簡捷法：

1. 回應曲綫：

求曲綫相關量數時，如用普通法，則計算手續至繁，尤以求回應曲綫時爲最繁，因而簡捷法於此至要。在普通法中，係根據$X$，$Y$兩變數原有各中值計算，現在簡捷法中，則改而根據兩變數之假定平均數組距離差計算，所用之數字簡單，手續自較簡易。現仍依普通法中所用每畝土地施肥及產麥相關表以說明之。

（第101表） 施肥及產麥相關表

$X$：肥料 單位：磅。 $Y$：小麥 單位：蒲耳。

| $i_x$ \ $i_y$ | | | | 0—40 | 40—80 | 80—120 | 120—160 | 總計 |
|---|---|---|---|---|---|---|---|---|
| | $X$ \ $Y$ | | | 20 | 60 | 100 | 140 | |
| | | $f_x$ \ $f_y$ | | 4 | 3 | 5 | 3 | 15 |
| | | | $d'_x$ \ $d'_y$ | 0 | 1 | 2 | 3 | |
| 24—32 | 28 | 3 | 3 | | 1 | 2 | | |
| 16—24 | 20 | 7 | 2 | | 2 | 3 | 2 | |
| 8—16 | 12 | 3 | 1 | 2 | | | 1 | |
| 0—8 | 4 | 2 | 0 | 2 | | | | |
| 總計 | | 15 | | | | | | |

（第102表） 回應曲綫方程式計算表

總方程式：$d'_{yc}=a+bd'_x+cd'^2_x$

| $d'_x$ | $d'_y$ | $f$ | $fd'_x$ | $fd'_y$ | $fd'_x d'_y$ | $fd'^2_x$ | $fd^2_x d'_y$ | $fd'^3_x$ | $fd^4_x$ |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| 1 | 2 | 2 | 2 | 4 | 4 | 2 | 4 | 2 | 2 |
| 1 | 3 | 1 | 1 | 3 | 3 | 1 | 3 | 1 | 1 |
| 2 | 2 | 3 | 6 | 6 | 12 | 12 | 24 | 24 | 48 |
| 2 | 3 | 2 | 4 | 6 | 12 | 8 | 24 | 16 | 32 |
| 3 | 1 | 1 | 3 | 1 | 3 | 9 | 9 | 27 | 81 |
| 3 | 2 | 2 | 6 | 4 | 12 | 18 | 36 | 54 | 162 |
| 總計 | | 15 | 22 | 26 | 46 | 50 | 100 | 121 | 326 |

標準方程式:

1. $\Sigma fd'_y = a\Sigma f + b\Sigma fd'_x + c\Sigma fd'^2_x$

2. $\Sigma fd'_x d'_y = a\Sigma fd'_x + b\Sigma fd'^2_x + c\Sigma fd'^3_x$

3. $\Sigma fd'^2_x d_y = a\Sigma fd'^2_x + b\Sigma fd'^3_x + c\Sigma fd'^4_x$

以上列第102表中各總數代入,則標準方程式爲:

(1) $26 = 15a + 20b + 50c$

(2) $46 = 22a + 50b + 124c$

(3) $100 = 50a + 124b + 326c$

(2)÷2 $23 = 11a + 25b + 62c$ ······························ (4)

(3)÷10 $10 = 5a + 12.4b + 32.6c$ ······························ (5)

(5)×3: $30 = 15a + 37.2b + 97.8c$ ······························ (6)

(6)−(1): $4 = 15.2b + 47.8c$ ······························ (7)

(7)÷4: $1 = 3.8b + 11.95c$ ······························ (A)

(4)×5: $115 = 55a + 125b + 310c$ ······························ (8)

(5)×11: $110 = 55a + 136.4b + 358.6c$ ······························ (9)

(9)−(8): $-5 = 11.4b + 48.6c$ ······························ (10)

(10)÷2: $-2.5 = 5.7b + 24.3c$ ······························ (B)

(A)×5.7 $5.7 = 21.66b + 68.115c$ ······························ (C)

(B)×4.8 $-9.5 = 21.66b + 92.340c$ ······························ (D)

(D)−(C): $-15.2 = 24.225c$

∴ $c = -15.2/24.225 = -.6274$

以 $c$ 之值代入(A)式;得:

$$1 = 3.8b - 11.95 \times 0.6274 = 3.8b - 7.49743$$

$$3.8b = 8.49743$$

∴　$b==8.49743/3.8=2.2362$

以 $b, c$ 之值代入(5)式，得：

$$10=5a+12.4\times2.2362-32.6\times0.6274$$

$$10=5a+27.72888-20.45324$$

$$5a=10-27.72888+20.45324=2.72436$$

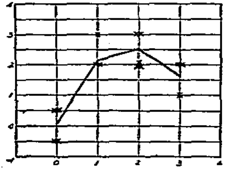
∴　$a=2.72436/5=0.544872$

回應曲綫總方程式：$d'_{yc}=0.54487+2.2362d'_x-0.6274d'^2_x$

（第103表）回應曲綫標準值計算表

| $d'_x$ | $d'_{yc}=0.54487+2.2362d'_x-0.627\ d'^2_x$ |
|---|---|
| 0 | 0.54 87 |
| 1 | 2.15367 |
| 2 | 2.50767 |
| 3 | 1.606?7 |

上列第103表中所得之回應曲準標綫值 $d'_{yc}$ 可與原有觀察值 $d'_y$，合併作入同一分散圖中，以見所得之曲綫是否配合於各觀察值。



（第35圖）回應曲綫標準值配合圖

2. 估計標準誤差　由上圖中可見二次拋物回應綫對於肥料與產量間之相關，至爲配合，如吾人不取此種回應曲綫而取直綫公式以求標準值

時，則所得之值，配合於原值必不如此恰當。回應曲綫配合雖當，但準標值與觀察值間仍有相當差異。在直綫相關中，以回應直綫代表 $X$，$Y$ 兩變數間之平均關係；而以標準誤差 $S_y$ 代表二者間差離之情形。現在曲綫相關中，仍以回應曲綫代表 $X$，$Y$ 兩變數間之平均關係，而以標準誤差直接代表觀察值 $Y_o$ 與標準值 $Y_c$ 間之離差，間接代表 $X$，$Y$ 間之平均關係。曲綫相關中標準誤差之意義與以前相同，計算方法亦與以前在直綫相關中完全相同，可根據上列所得之假定原點組距離差 $d'_y$，及此每標準值 $d'_{yc}$ 以求之，是爲簡捷法。亦可根據因變數中值 $Y_o$ 及單位標準值 $Y_c$ 以求之，是爲普通法，現仍利用簡捷法，依組距離差以計算之，較爲簡便。

(第 104 表)　估計標準誤差 $\sigma'_y$ 計算表

| $d'_x$ | $d'_{yo}$ | $d'_{yc}$ | $f$ | $d_{yo}-d_{yc}$ | $f(d_{yo}-d_{yc})$ + | $f(d_{yo}-d_{yc})$ − | $f(d'_y-d'_{yc})$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0.54487 | 2 | −.54487 | | 1.08974 | 0.5933 |
| 0 | 1 | 0.54487 | 2 | 0.45513 | .91026 | | 0.4142 |
| 1 | 2 | 2.15367 | 2 | −.15367 | | .30734 | 0.0473 |
| 1 | 3 | 2.15367 | 1 | 0.84633 | .84633 | | 0.7160 |
| 2 | 2 | 2.50767 | 3 | −.50767 | | 1.52301 | 0.7736 |
| 2 | 3 | 2.50767 | 2 | 0.49233 | .98466 | | 0.4845 |
| 3 | 1 | 1.60637 | 1 | −.60637 | | .60637 | 0.3653 |
| 3 | 2 | 1.60637 | 2 | 0.39313 | .78626 | | 0.3090 |
| 總計 | | | 15 | | 3.52751 | 3.526 6 | 3.7068 |

$$S'_y=\sqrt{\frac{\Sigma f(d'_y-d'_{yc})^2}{\Sigma f}}=\sqrt{\frac{3.7068}{15.00}}$$

$$=\sqrt{0.24712}\qquad=0.497$$

上列所得者係組距單位標準誤差，以 $S'_y$ 表示之，因變數 $Y$ 之組距爲 8 蒲耳。上列所得之結果以 8 乘之，卽得單位標準誤差。

$S_y=i_y.S'_y=8\times0.97=3.976$ 蒲耳。

標準誤差 $S_p$ 求出之後，自可利用之以定估計界限（Zones of Estimate），卽以迴應曲綫加減三倍標準誤差爲離中趨勢之範圍，并以此種情形表示于配合圖中。

3. 相關指數　在直綫相關中，相關係數係由估計標準誤差及標準差求出；現在曲綫相關中之相關指數，亦依標準誤差及標準差兩種量數求出。已知組距單位標準差誤爲 $S'_p=0.497$，其平方 $S'^2_p=0.24$ı12，至組距單位之標準差 $\sigma'_p$ 則可依下列第105表求出。

（第105表）　組距單位標準差計算表

| $Y$ | $f$ | $d'_p$ | $fd'_p$ | $fd'^2_p$ |
|---|---|---|---|---|
| 4 | 2 | −2 | −4 | 8 |
| 12 | 3 | −1 | −3 | 3 |
| 20 | 7 | 0 | 0 | 0 |
| 28 | 3 | 1 | 3 | 3 |
| 總計 | 15 | | −4 | 14 |

$$\sigma'y=\sqrt{\frac{\Sigma fd'^2_p}{\Sigma f}-\left(\frac{\Sigma f'd_p}{\Sigma f}\right)^2}=\sqrt{\frac{14}{15}-\left(\frac{-4}{15}\right)^2}$$

$$=\sqrt{\frac{210-16}{225}}\qquad=\sqrt{\frac{194}{225}}$$

$$=\sqrt{0.862222}\qquad=0.9285$$

$$\rho_{yx}=\sqrt{1-\frac{S'^2_p}{\sigma'^2_p}}\qquad=\sqrt{1-\frac{0.24712}{0.86222}}$$

$$=\sqrt{\frac{0.61510}{0.86222}}\qquad=\sqrt{0.713391}$$

$$=0.8446$$

以上所得之迴應曲綫標準值 $d'_{..}$，係根據 $X$，$Y$ 兩變數之假定平均數

組距離差計算，亦可設法還原，化爲根據自變數 $X$ 各中值所定之標準值，與依普通法所得之標準值 $Y_c$ 相同。其還原之方法與直線相關中相同：

1. 由組距離差改爲單位離差　上列回應曲綫方程式 $d'_{yc}=0.54487+2.2362d'_x-0.6274d'^2_x$，係依組距離求得者。卽爲簡單計，曾將單位離差除以組距，化爲組距離差，$d'_x=d_x/i_x$，$d_y'=d_y/i_y$，自變數組距 $i_x=40$，故 $d'_x=d_x/40$；因變數組距 $i_y=8$，故 $d'_y=d_y/8$，以此二者代入上式，則可得一新方程式如下：

$$\frac{d_y}{8}=0.54487+2.2362\left(\frac{d_x}{40}\right)-0.6274\left(\frac{d_x}{40}\right)^2$$

$$\frac{d_y}{8}=0.54487+0.055905d_x-0.000392d^2_x$$

$$d_y=4.35896+0.44724d_x-0.003136d^2_x$$

2. 由單位離差改爲 $X$, $Y$ 兩變數之中值　上式中 $d_x$ 及 $d_y$ 均爲自假定平均數 $X_1$ 及 $Y_1$ 所計算之單位離差，現須自離差改爲中值，自變數之假定平均數 $X_1=20$，$d_x=X-20$；因變數之假定平均數 $Y_1=4$，$d_y=Y-4$，以此二者代入上式，則可得最後之回應曲綫方程式如下：

$$Y-4=4.35896+0.44724(X-20)-0.003136(X-20)^2$$

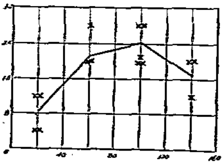
$$Y=4+4.35896+0.44724X-8.94480-0.003136(X^2-40X+400)$$

$$Y=-0.58584+0.44724X-0.003136X^2+0.125440X-1.25440$$

$$Y_c=-1.84024+0.57268X-0.003136X^2$$

（第106表）回應曲綫標準值計算表

| $X_c$ | $Y_c=-1.84024+0.57268X-0.003136X^2$ |
|---|---|
| 20 | 8.35896 |
| 60 | 21.23096 |
| 100 | 24.06776 |
| 140 | 16.86936 |



(第36圖) 回應曲綫配合圖

## 習　題

(第107表) 施肥與產麥相關表

*X*: 肥料，單位(磅)

*Y*: 小麥，單位(蒲耳)

| $i_y$ \ $I_x$ | | | 0—20 | 20—40 | 40—60 | 60—80 | 80—100 | 100—120 | 120—140 | 140—160 | 160—180 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Y \ X | | 10 | 30 | 50 | 70 | 90 | 110 | 130 | 150 | 170 |
| | | $f_y$ \ $f_x$ | | | | | | | | | |
| 32—36 | 34 | | | | | 6 | 15 | 10 | 4 | 6 | 2 |
| 28—32 | 30 | | | | 1 | 11 | 20 | 9 | 5 | 1 | |
| 24—28 | 26 | | | 1 | 15 | 20 | 3 | | | | |
| 20—24 | 22 | | | 2 | 12 | | | | | | |
| 16—20 | 18 | | | 10 | 2 | | | | | | |
| 12—16 | 14 | | | 8 | | | | | | | |
| 8—12 | 10 | | 4 | 4 | | | | | | | |
| 4—8 | 6 | | 10 | | | | | | | | |
| 0—4 | 2 | | 6 | | | | | | | | |

一、試根據上列第107表 用簡捷法求曲綫相關量數:

1. 回應曲綫總式及標準值·並繪一分散圖與配合圖。
2. 估計標準誤差,並將估計界限繪入上圖中。

3. 相關指數

二、試自覓適當未分組資料求三種曲綫相關量數。

# 第十二章　相關比與等級相關係數

## 第一節　相關比

以上已得有相關係數及相關指數。前者係一抽象數值用以代表二種材料間直綫關係之程度者。後者亦係一抽象數值用以代表二種材料間曲綫關係之程度者。此外尚可求一抽象數值，亦係代表相關之程度者，即相關比 (Correlation Ratio)，以 $\eta$ (eta) 表示之。如 $X$ 代表自變數，$Y$ 代表因變數，$Y$ 值隨 $X$ 之值而變，則相關比之符號為 $\eta_{yx}$。反之，如 $Y$ 為自變數，$X$ 為因變數時，則相關比之符號為 $\eta_{xy}$。此種數值只能根據分組材料即相關表求之，不能根據未分組材料計算，相關表中之材料，無論為直綫或曲綫相關，均可求相關比。如係直綫相關，則所求之 $\eta_{yx}$，$\eta_{xy}$ 及 r 三值相等，如為曲綫相關，則此三值彼此不相等。但 $\eta_{yx}$ 及 $\eta_{xy}$ 均大於 r，故英統計學家皮爾森(Karl Pearson)即根據 $\eta$ 及 r 間之關係，以斷定統計材料間之關係是否係直綫相關，稱為直綫性之測驗(Test of Linearity)，其符號為 $\zeta$(Zeta)，$\zeta=\eta^2-r^2$，如相關比之平方減去相關係數之平方，所餘之值 $\zeta$ 等於 0，則表示二種材料間之關係為直綫相關。反之，如二者之差大於零，即 $\eta$ 大於 r，則表示為曲綫相關，二者之差額愈大，即 $\zeta$ 之值愈大時，則曲綫之程度愈高。

依相關表而求相關比之方法有二；現仍以上章所用之金條與美鈔市價相關表為例，將求相關比之二法，分別說明之。

一、行列平均數離差法：

所謂行列平均數離差法 (Method of Deviation of Array Means

(第108表 金條與美鈔市價相關表

| $l_x$ / $l_p$ | | | 3.50—3.70 | 3.70—3.90 | 3.90—4.10 | 4.10—4.30 | |
|---|---|---|---|---|---|---|---|
| | $X$ / $Y$ | | 3.60 | 3.80 | 4.00 | 4.20 | |
| | | $f_x$ / $f_p$ | 6 | 7 | 5 | 2 | 20 |
| 7.50—7.90 | 7.70 | 1 | | | | 1 | |
| 7.10—7.50 | 7.30 | 3 | | | 2 | 1 | |
| 6.70—7.10 | 6.90 | 4 | | 1 | 3 | | |
| 6.30—6.70 | 6.50 | 12 | 6 | 6 | | | |
| $Y$—序列平均數 | | 20 | 6.50 | 6.56 | 7.06 | 7.50 | |

from the General Mean) 即根據行列平均數與總平均數間之離差以求相關比,其程序可分以下數步驟:

1. 求因變數分配之總平均數(The General Mean)及其各行列平均數(Array Means),相關表中因變數 $Y$ 自成一次數分配,自可求出平均數如下:

(第109表) 民國三十六年元月上海美鈔平均價格計算表

單位:美鈔一元對法幣千元,

| $l_p$ | $Y$ | $f$ | $fY$ |
|---|---|---|---|
| 6.30—6.70 | 6.50 | 12 | 78.0 |
| 6.70—7.10 | 6.90 | 4 | 27.6 |
| 7.10—7.50 | 7.30 | 3 | 21.9 |
| 7.50—7.90 | 7.70 | 1 | 7.7 |
| | | 20 | 135.2 |

$$\bar{Y}=\frac{\Sigma fY}{\Sigma f}=\frac{135.2}{20.0}=\frac{13.52}{2.00}=6.76$$

除因變數總平均數以外,在相關表中以自變數各變項卽中值 $X_1$,$X_2$,$X_3$,及 $X_4$ 爲標準時,每個中值所針對一欄內之 $Y$,成一次數分配,既爲

次數分配，自可求一平均數，稱爲行列平均數，以$\bar{Y}_i$表示之。現計算各行列平均數，并於求得後記入相關表中最底一格內。

$$X_1=3.6;\ \bar{Y}_1=\frac{\Sigma fY}{\Sigma f}=\frac{6.5\times 6}{6}=6.5$$

$$X_2=3.8;\ \bar{Y}_2=\frac{6.5\times 6+6.9\times 1}{7}=\frac{45.9}{7}=6.5\dot{6}$$

$$X_3=4.0;\ \bar{Y}_3=\frac{6.9\times 3+7.3\times 2}{5}=\frac{35.3}{5}=7.06$$

$$X_4=4.2;\ \bar{Y}_4=\frac{7.3\times 1+7.7\times 1}{2}=\frac{15.0}{2.0}=7.50$$

2. 求行列平均數之標準差及總標準差　吾人如將所求得之行列平均數$\bar{Y}_i$視爲一變數，并以因變數總平均數$\bar{Y}$爲一共同標準，則各行列平均數$\bar{Y}_i$與總平均數$\bar{Y}$間，可求出離差及離差之平方，分別以$(\bar{Y}_i-\bar{Y})$及$(\bar{Y}_i-\bar{Y})^2$表示之，乘以各組次數$f$，則離差平方之公式爲$f(\bar{Y}_i-\bar{Y})^2$此係變數，可展開爲$f_1(\bar{Y}_1-\bar{Y})^2$，$f_2(\bar{Y}_2-\bar{Y})^2$，……，其總和爲：$\Sigma f(\bar{Y}_i-\bar{Y})^2$其次數總和則爲$\Sigma f$，依此兩項總和，可求行列平均數之標準差，以$\sigma_{my}$表示之則：

$$\sigma_{my}=\sqrt{\frac{\Sigma f(\bar{Y}_i-\bar{Y})^2}{\Sigma f}}$$

此公式中所需要之兩項總和，可由下列計算表以得之：

(第110表)　行列平均數標準差計算表

| $\bar{Y}_i$ | $\bar{Y}$ | $f_x$ | $\bar{Y}_i-\bar{Y}$ | $f(\bar{Y}_i-\bar{Y})$ | $f(\bar{Y}_i-\bar{Y})^2$ |
|---|---|---|---|---|---|
| 6.50 | 6.76 | 6 | —.26 | −1.56 | 0.40 6 |
| 6.56 | 6.76 | 7 | —.20 | −1.40 | 0.2800 |
| 7.06 | 6.76 | 5 | —.30 | −1.50 | 0.4500 |
| 7.50 | 6.76 | 2 | 0.74 | 1.48 | 1.0952 |
| 總計 | | 20 | | | 2.2308 |

$$\sigma_{mp}=\sqrt{\frac{\Sigma f(\overline{Y}_i-\overline{Y})^2}{\Sigma f}}=\sqrt{\frac{2.2308}{20.00}}$$

$$=\sqrt{0.11154}\qquad=0.334$$

除行平均數標準差以外，尚須求因變數 $Y$ 之總標準差即 $\sigma_p$，可依下列計算表中得之：

(第111表) 標準差 $\sigma_p$ 計算表 $\overline{Y}=6.76$

| $Y$ | $f$ | $y=Y-\overline{Y}$ | $fy$ | $fy^2$ |
|---|---|---|---|---|
| 6.50 | 12 | -.26 | -3.12 | 0.8112 |
| 6.90 | 4 | 0.14 | 0.56 | 0.0784 |
| 7.30 | 3 | 0.54 | 1.62 | 0.8748 |
| 7.70 | 1 | 0.94 | 0.94 | 0.8836 |
| 總計 | 20 | | 0 | 2.6480 |

$$\sigma_p=\sqrt{\frac{\Sigma f(Y-\overline{Y})^2}{\Sigma f}}=\sqrt{\frac{2.648}{20.00}}=\sqrt{0.1324}=0.364$$

3. 求相關比 所謂相關比 即將上列所得之行列平均數標準差由絕對數化為相對數，用以表示相關程度，其方法即係以因變數之總標準差 $\sigma_p$ 除其行列平均數之標準差。此種結果、即為相關比以 $\eta_{px}$ 表示之

$$\eta_{xp}=\frac{\sigma_{mp}}{\sigma_p}=\frac{0.334}{0.364}=0.918$$

二 原有中值離差法 所謂原有中值離差法 (Method of Deviation of Orginal Mid-Values from Array Means) 即以因變數各行列平均數為標準而求因變數 $Y$ 各中值之離差，由此離差求標準差，進而再由此標準差求相關比，其步驟如下：

1. 求各中值以行列平均數為準之標準差 因變數各中值為 $Y$，其行列平均數為 $\overline{Y}_p$ 以行列平均數 $\overline{Y}_i$ 為準之離差為 $(Y-\overline{Y}_i)$，依此離差而求

之標準差，以$\sigma_{cy}$表示之，則$\sigma_{cy}$之公式如下：

$$\sigma_{cy}=\sqrt{\frac{\Sigma f(Y-\bar{Y}_t)^2}{\Sigma f}}$$

求解此標準差須根據相關表（第112表）及分欄計算表（第113表）

（第112表） 上海黃金美鈔市價相關表

| $i_y$ \ $i_x$ | | | 3.50−3.70 | 3.70−3.90 | 3.90−4.10 | 4.10−4.30 | |
|---|---|---|---|---|---|---|---|
| | $Y$ \ $X$ | | 3.60 | 3.80 | 4.00 | 4.20 | |
| | | $f_y$ \ $f_x$ | 6 | 7 | 5 | 2 | 20 |
| 7.50−7.90 | 7.70 | 1 | | | | 1 | |
| 7.10−7.50 | 7.30 | 3 | | | 2 | 1 | |
| 6.70−7.10 | 6.90 | 4 | | 1 | 3 | | |
| 6.30−6.70 | 6.50 | 12 | 6 | 6 | | | |
| 行列平均數 | | 20 | 6.50 | 6.56 | 7.06 | 7.50 | |

（第113表） 標準差$\sigma_{cy}$計算表

| $Y$ | $\bar{Y}_t$ | $f$ | $Y-\bar{Y}_t$ | $f(Y-\bar{Y}_t)$ | $f(Y-\bar{Y}_t)^2$ |
|---|---|---|---|---|---|
| 6.50 | 6.50 | 6 | 0 | 0 | 0 |
| 6.50 | 6.56 | 6 | −.06 | −.36 | .0216 |
| 6.90 | 6.56 | 1 | 0.34 | 0.34 | .1156 |
| 6.90 | 7.06 | 3 | −.16 | −.48 | .0768 |
| 7.30 | 7.06 | 2 | 0.24 | 0.48 | .1152 |
| 7.30 | 7.50 | 1 | −.20 | −.20 | .0400 |
| 7.70 | 7.50 | 1 | 0.20 | 0.20 | .0400 |
| 總計 | | 20 | | | 0.4092 |

$$\sigma_{cy}=\sqrt{\frac{\Sigma f(Y-Y_t)^2}{\Sigma f}}=\sqrt{\frac{0.4092}{20.00}}=\sqrt{0.02046}=0.143$$

2． 求相關比 依此法求相關比，與依最小二乘方法求相關係數時所用之公式形式相同，

$$\eta_{px}=\sqrt{1-\frac{\sigma^2_{ap}}{\sigma^2_p}}\qquad=\sqrt{1-\frac{0.02046}{0.13246}}$$

$$=\sqrt{\frac{0.11194}{0.13240}}\qquad=\sqrt{0.8456}$$

$$=0.918$$

## 第二節　等級相關係數

等級相關係數，與上節所述之相關比相反，只能用以分析未分組資料間之相關。其法係將 $X$, $Y$ 兩變數所包括之各變項，依其數值之大小排爲序列，然後依此序列之等級而求相關係數，稱之爲等級相關係數 (Coefficient of Rank Correlation) 以 $\rho_r$ 表示之，用以代表二變數間相關之程度，現以學生七人在兩門課程中考試之分數爲例，以見等級相關係數之計算方法：

(第114表)　等級相關係數計算表

| 學生號數 | 甲種課程 分數 | 甲種課程 等級:$X$ | 乙種課程 分數 | 乙種課程 等級:$Y$ | 等級離差 $d=X-Y$ | 等級離差平方 $d^2$ |
|---|---|---|---|---|---|---|
| 1 | 100 | 1 | 90 | 1 | 0 | 0 |
| 2 | 95 | 2 | 80 | 2 | −2 | 4 |
| 3 | 90 | 3 | 85 | 3 | 0 | 0 |
| 4 | 85 | 4 | 70 | 5 | −1 | 1 |
| 5 | 80 | 5 | 88 | 2 | 3 | 9 |
| 6 | 75 | 6 | 60 | 6 | 0 | 0 |
| 7 | 70 | 7 | 50 | 7 | 0 | 0 |
| 總計 | | | | | | 14 |

$$N=7,\qquad \Sigma d^2=14$$

$$\rho_r=1-\frac{6\Sigma d^2}{n^3-n}\qquad=1-\frac{6\times 14}{7^3-7}=1-\frac{84}{336}$$

$$=252/336\qquad=0.75$$

利用上列等級相關係數之公式，以測驗二現象間有無相關或相關程度之高低，手續至爲簡易。此公式係由直綫相關中依積差法求相關係數之公式變化而來，換言之，$\rho r$ 係由 $r$ 變來，茲證明之：

以 $X, Y$ 代表二組等級之各變項，并以 $\bar{X}, \bar{Y}$ 代表平均數，以 $x, y$ 代表離差，則可得：

$$x=X-\bar{X}, \qquad X=x+\bar{X};$$
$$y=Y-\bar{Y}, \qquad Y=y+\bar{Y}。$$

$X, Y$ 旣名爲代表等級之變數，是其變項各爲 1,2,3,4,5,6,7，各爲第一自然變數（First Natural Numbers）；$X, Y$ 二變數之各個變項相同，均爲 1,2,3,4,5,6,7，則其平均數自彼此相同，即 $\bar{X}=\bar{Y}$。

以 $d$ 代表 $X, Y$ 間之差別，即 $d=X-Y$，已知 $X=x+\bar{X}$，$Y=y+\bar{Y}$ 故 $d=x+\bar{X}-(y+\bar{Y})$，即 $d=x+\bar{X}-y'-\bar{X}$，即 $d=x-y$，$d^2=(x-y)^2=x^2-2xy+y^2$，但 $d, x, y$ 各爲變數，是：$\Sigma d^2=\Sigma(x-y)^2=\Sigma(x^2-2xy+y^2)=\Sigma x^2-2\Sigma xy+\Sigma y^2$

$$\therefore \quad 2\Sigma xy=\Sigma x^2+\Sigma y^2-\Sigma d^2 \cdots\cdots\cdots\cdots\cdots\cdots(1)$$

已知 $x=(X-\bar{X})$，$x^2=(X-\bar{X})^2$，$\Sigma x^2=\Sigma(X-\bar{X})^2$。但在第一自然數中，平均數 $\bar{X}=(n+1)/2$。以之代入：

$$\Sigma x^2=\Sigma\left(X-\frac{n+1}{2}\right)^2=\Sigma X^2-2\left(\frac{n+1}{2}\right)\Sigma X+\frac{(n+1)^2}{4}n$$
$$=\Sigma X^2-(n+1)\Sigma X+\frac{n(n+1)^2}{4}。$$

但在第一個自然數中：

(a)　$\Sigma X=1+2+3+\cdots\cdots\cdots+n=\dfrac{n(n+1)}{2}$

(b)　$\Sigma X^2=1^2+2^2+3^2+\cdots\cdots+n^2=\dfrac{n(n+1)(2n+1)}{6}$

$$\therefore\ x^2=\frac{n(n+1)(2n+1)}{6}-\frac{(n+1)(n+1)n}{2}+\frac{n(n+1)^2}{4}$$

$$=\frac{n(n+1)[4n+2-6n-6+3n+3]}{12}$$

$$=\frac{n(n+1)(n+1)}{12}$$

$$=\frac{n^2-n}{12}$$

$\Sigma x^2=\dfrac{n^3-n}{12}$，　$\Sigma y^2=\dfrac{n^3-n}{12}$，以之代入(1)：

$$\therefore\ 2\Sigma xy=\frac{n^3-n}{12}+\frac{n^3-n}{12}-\Sigma d^2$$

$$2\Sigma xy=\frac{2(n^3-n)}{12}-\Sigma d^2$$

$$\Sigma xy=\frac{n^3-n}{12}-\frac{\Sigma d^2}{2}=\frac{1}{2}\left(\frac{n^3-n}{6}-\Sigma d^2\right)\cdots\cdots\cdots\cdots(2)$$

依積差法以求等級相關係數之公式爲：

$$P_r=\frac{\Sigma xy}{N\sigma_x\sigma_y}=\frac{\Sigma xy}{N\sqrt{\dfrac{\Sigma x^2}{N}}\sqrt{\dfrac{\Sigma y^2}{N}}}=\frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$$

已知：$\Sigma xy=\dfrac{1}{2}\left(\dfrac{n^3-n}{6}-\Sigma d^2\right)$；

$$\Sigma x^2=\frac{n^3-n}{12},\ \sqrt{\Sigma x^2}=\sqrt{\frac{n^3-n}{12}};$$

$$\Sigma y^2=\frac{n^3-n}{12},\ \sqrt{\Sigma y^2}=\sqrt{\frac{n^3-n}{12}}$$

$$\therefore\ \rho_r=\frac{\dfrac{1}{2}\left(\dfrac{n^3-n}{6}-\Sigma d^2\right)}{\sqrt{\dfrac{n^3-n}{12}}\sqrt{\dfrac{n^3-n}{12}}}=\frac{\dfrac{1}{2}\left(\dfrac{n^3-n}{6}-\Sigma d^2\right)}{n^3-n}$$

$$=\frac{1}{2}\left(\frac{n^3-n}{6}-\Sigma d^2\right)\times\frac{12}{n^3-n}=\frac{6\left(\frac{n^3-n}{6}-\Sigma d^2\right)}{n^3-n}$$

$$=\frac{n^3-n-6\Sigma d^2}{n^2-n}=1-\frac{6\Sigma d^2}{n^3-n}$$

## 習　題

一、試根據100名大學男生身高與體重相關表求相關比。

三、試覓適當未分組相關 料求等級相關係數。

# 第十三章　多項相關與偏相關

## 第一節　多項相關

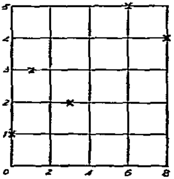
一、多項回應方程式。

吾人在以上數章中所分析者，係 $X, Y$ 兩變數間之相關，稱為簡單相關，或稱為兩項相關。但自然現象或社會現象間有連帶關係者，初不只兩項，各種現象中間每均有相互關係。以人類身體發育而論，身高與體重間固有關係，身高、體重與體力間，實彼此各有關係。以農業經營言論，收穫與肥料間固有關係，但肥料之外，氣候與雨量對於收穫亦各有關係。以物價而論，黃金與美鈔兩種價格間，固有關係，但米價、布價、以及其他多種價格，無一不與之息息相關。此等許多現象間之連帶關係，稱之為多項相關，或複相關（Multiple Correlation）。多項相關分析之方法，實只為簡單相關技術之引伸。在簡單相關中，無論在身高與體重間，收穫與肥料間，黃金與美鈔間，均只有一自變數 $X$ 及一因變數 $Y$。在多項相關中，無論資料中所包括之變數為三個，四個，或更多，因變數仍只一個，以 $Y$ 代表之，其他均為自變數，以 $X_1, X_2, X_3$ 等代表之。多項相關所研究者，即各組自變數聯合對于因變數所發生之總相關。多項回應方程式（Multiple Equation of Regression），即代表此種相關之方程式，換言之，即代表因變數 $Y$ 如何隨自變數 $X_1, X_2, X_3$ 等變動之方程式。此方程式一經求得，自可如在簡單相關中，依之以求因變數之標準值。現舉實例說明之：

在多項相關回應方程式中，亦有未分組資料，與分組資料，直線與曲線之別，須視資料之實際情形而定。下表所列資料為三項相關，自變數為

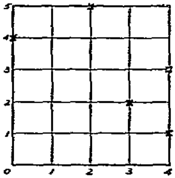
（第115表） 三項相關資料

| $X_1$ | $X_2$ | $Y$ |
|---|---|---|
| 0 | 4 | 1 |
| 1 | 4 | 3 |
| 3 | 3 | 2 |
| 6 | 2 | 5 |
| 8 | 0 | 4 |




（第37 a 圖） 分散圖

$(X_1, Y)$




（第37 b 圖） 分散圖

$(X_2, Y)$

$X_1$及$X_2$，因變數爲$Y$。因三變數各只包括五變項，自不分類，而依各項原有數值，直接計算多項相關係數。在計算之前，可作分散圖。自第一分散圖即$X_1$及$Y$各變項之分散圖觀之，可見二者爲正直綫相關。自第二分散圖即$X_2$及$Y$各變項之分散圖中，可見二者爲負直綫相關。綜觀兩圖既俱爲直綫相關，自可依直綫相關技術以求多項相關係數。最初即爲求多項回應程式，用以代表三變數間之平均關係。在簡單相關中，自變數爲$X$，因變數爲$Y$，回應方程式爲$Y=a+bX$，聯立標準式爲：

1. $\Sigma Y=Na+b\Sigma X$

2. $\Sigma XY=a\Sigma X+b\Sigma X^2$

現在多項相關中，自變數有二，即 $X_1$ 及 $X_2$，因變數仍爲 $Y$，則囘應方程式爲：$Y=b_0+b_1X_1+b_2X_2$，此方程式中旣有 $b_0$, $b_1$ 及 $b_2$ 三常數，自須有三聯立方程式，方能解決，三者之形式如下：

1. $\Sigma Y=Nb_0+b_1\Sigma X_1+b_2\Sigma X_2$
2. $\Sigma X_1Y=b_0\Sigma X_1+b_1\Sigma X_1^2+b_2\Sigma X_1X_2$
3. $\Sigma X_2Y=b_0\Sigma X_2+b_1\Sigma X_1X_2+b_2\Sigma X_2^2$

（第116表）　多項囘應標準方程式計算表

| $X_1$ | $X_2$ | $Y$ | $X_1X_2$ | $X_1^2$ | $X_2^2$ | $X_1Y$ | $X_2Y$ |
|---|---|---|---|---|---|---|---|
| 0 | 4 | 1 | 0 | 0 | 16 | 0 | 4 |
| 1 | 4 | 3 | 4 | 1 | 16 | 3 | 12 |
| 3 | 3 | 2 | 9 | 9 | 9 | 6 | 6 |
| 6 | 2 | 5 | 12 | 36 | 4 | 30 | 10 |
| 8 | 0 | 4 | 0 | 64 | 0 | 32 | 0 |
| 18 | 13 | 15 | 25 | 110 | 45 | 71 | 32 |

三項相關囘應方程式：$Y=b_0+b_1X_1+b_2X_2$

三聯立標準方程式之形式，已見上文。現以上列計算表總數行中所得各值代入，則得三聯立方程式如下：

(1)　$5b_0+18b_1+13b_2=15$

(2)　$18b_0+110b_1+25b_2=71$

(3)　$13b_0+25b_1+45b_2=32$

18(1)：　$90b_0+324b_1+234b_2=270$ ……………………(4)

5(2)：　$90b_0+550b_1+125b_2=355$ ……………………(5)

(4)—(5)：　$-226b_1+109b_2=-85$ ……………………(A)

13(1)：　$65b_0+234b_1+169b_2=195$ ……………………(6)

5(3)：　$65b_0+125b_1+225b_2=160$ ……………………(7)

(6)-(7):　　$109b_1-56b_2=35$ ……………………$(B)$

$226(B)$:　$24,634b_1-12,656b_2=7910$ ……………………$(C)$

$109(A$ :　$-24,634b_1+11,881b_2=-9265$ ………………$(D)$

$(C)+(D)$:　　$-775b_2=-1355$

$\therefore\ b_2=1355/775=1.7484$

$(B)$:　$109b_1-56\times1.7484=35$

$109b_1=97.9104+35=132.9104$

$\therefore\ b_1=132.9104/109=1.2194$

(1):　$5b_0+18\times1.2194+13\times1.7484=15$

$5b_0=15-21.9492-22.7292=-29.6784$

$\therefore\ b_0=-29.6784/5=-5.9357$

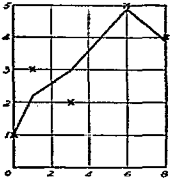
$b_0$, $b_1$, $b_2$ 三常數既已求出，代入則得多項回應線方程式式：

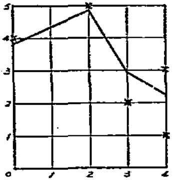
$Y=-5.9357+1.2194X_1+1.7484X_2$，並可依此線以求 $Y$ 之標準值。

(第 117 表)　標準值 $Y_c$ 計算表

| $X_1$ | $X_2$ | $Y_c$ | $Y_c=-5.9357+1.2194X_1+1.7484X_2$ |
|---|---|---|---|
| 0 | 4 | 1 | 1.05.9 |
| 1 | 4 | 3 | 2.2773 |
| 3 | 3 | 2 | 2.9677 |
| 6 | 2 | 5 | 4.8775 |
| 8 | 0 | 4 | 3.8195 |

由下列兩圖中，可見依最小二乘方式所得之回應式及標準值，配合原有資料甚爲恰當；即所得之回應線方程式足以代表三變數間之平均關係。上列之多項相關回應式及附帶之聯立標準式，亦有用不同符號以表示者。如回應線方程式爲：





標準值配合圖

(第38圖) (叉號)代表 $Y_o$, 曲綫代表 $Y_c$　　(第39圖) (叉號)代表 $Y_o$ 曲綫代表 $Y_c$

$X_1=a+b_{12.3}X_2+b_{13.2}X_3$

三聯立標準方程式爲:

1. $\Sigma X_1=Na+b_{12.3}\Sigma X_2+b_{13.2}\Sigma X_3$

2. $\Sigma X_1X_2=a\Sigma X_2+b_{12.3}\Sigma X_2^2+b_{13.2}\Sigma X_2X_3$

3. $\Sigma X_1X_3=a\Sigma X_3+b_{12.3}\Sigma X_2X_3+b_{13.2}\Sigma X_3^2$

上列表示之方法，與吾人所用者形式上雖稍不同，實質上則完全相同，惟似不如吾人于上列所用者爲簡明。三項相關回應總方程式及聯立標準方程式明瞭後，相關變數再加多，均可循理類推，可得一般形式如下:

多項回應總方程式: $Y_c=b_0+b_1X_1+b_2X_2+\cdots\cdots+b_k$

聯立標準方程式:

1. $\Sigma Y=Nb_0+b_1\Sigma X_1+b_2\Sigma X_2+\cdots\cdots+b_k\Sigma X_k$

2. $\Sigma X_1Y=b_0\Sigma X_1+b_1\Sigma X_1^2+b_2\Sigma X_1X_2+\cdots\cdots+b_k\Sigma X_1X_k$

3. $\Sigma X_2Y=b_0\Sigma X_2+b_1\Sigma X_1X_2+b_2\Sigma X_2^2+\cdots\cdots+b_k\Sigma X_2X_k$

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

k. $\Sigma X_kY=b_0\Sigma X_k+b_1\Sigma X_1X_k+b_2\Sigma X_2X_k+\cdots\cdots+b_k\Sigma X_k^2$

## 二、多項回應式估計標準誤差

在簡單相關中，以回應直綫代表 $X, Y$ 兩變數間之平均關係，以估計標準誤差代表觀察值 $Y_o$ 與標準值 $Y_c$ 間之平均離差。現在多項相關中仍以依多項回應式所定之標準值代表自變數 $X_1$ $X_2$ 與因變數 $Y$ 間之平均關係，而以估計標準誤差代表觀察值 $Y_o$ 與標準值 $Y_c$ 中間之離差，此標準值旣依多項回應式所決定，故此中之標準誤差稱爲多項回應式估計標準誤差（The Standard Error of Estimate of the Equation of Multiple Regression）或簡稱之爲多項標準誤差。如相關之變數爲三項，則估計標準誤差之符號爲 $S_{Y\cdot X_1X_2}$。其計算方法及所得數值之意義與在簡單相關中完全相同，現依上節實例計算之。

（第118表）　多項估計標準誤差計算表

| $Y_o$ | $Y_c$ | $Y_o-Y_c$ + | $Y_o-Y_c$ − | $(Y_o-Y_c)^2$ |
|---|---|---|---|---|
| 1 | 1.0579 | | 0.0579 | 0.00335241 |
| 3 | 2.2773 | 0.7227 | | 0.522 95 9 |
| 2 | 2.9677 | | 0.9677 | 0.9 614329 |
| 5 | 4.8775 | 0.1225 | | 0.01500625 |
| 4 | 3.8195 | | 0.1805 | 0.032[illegible]8025 |
| | | 1.0257 | 1.0256 | 1.5.967749 |

$$S_{Y\cdot X_1X_2}=\sqrt{\frac{\Sigma(Y_o-Y_c)^2}{N}}=\sqrt{\frac{1.5\ 967749}{5.0}}$$

$$=\sqrt{0.3019355}=0.5495$$

## 三、多項相關係數

在簡單相關中，如爲直綫相關，卽用相關係數 r 以代表相關之程度。現在多項相關中，仍以相關係數代表此種相關程度，稱之爲多項相關係數（Coefficient of Multiple correlation），$R$ 在上例中自變數有二，卽 $X_1$ 及

$X_2$，因變數為 $Y$，則多項相關係數之符號為 $R_{Y\cdot X_1X_2}$ 其公式之形式及計算之方法與以前在簡單相關中相同。

$$R_{Y\cdot X_1X_2}=\sqrt{1-\frac{S^2_{Y\cdot X_1X_2}}{\sigma_y^2}}$$

今已知 $S^2_{Y\cdot X_1X_2}=0.3019355$，$\sigma_y^2$ 則可由下列計算表中得之。

（第119表）因變數標準差計算表

| $Y_0$ | $Y_0^2$ |
|---|---|
| 1 | 1 |
| 3 | 9 |
| 2 | 4 |
| 5 | 25 |
| 4 | 16 |
| 15 | 55 |

$$\sigma_y=\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}=\sqrt{\frac{55}{5}-\left(\frac{15}{5}\right)^2}=\sqrt{11-9}=\sqrt{2}$$

$$R_{Y\cdot X_1X_2}=\sqrt{1-\frac{0.3019355}{2}}=\sqrt{\frac{1.6980645}{2}}$$

$$=\sqrt{0.84903225}\quad=0.9214$$

上列所得之多項相關係數，代表第一自變數 $X_1$ 及第二自變數 $X_2$ 聯合與因變數 $Y$ 中間相關之程度者。計算方法與簡單相關中之相關係數相同。惟一異點，即多項相關係數已不附正負號。蓋在多項相關中，$X_1$ 與 $Y$ 間之相關可為正，而 $X_2$ 與 $Y$ 間之相關則可為負。同理，在多項回應式中，自因變數 $Y$ 隨第一自變數 $X_1$，或隨第二自變數 $X_2$ 觀之，各為直綫相關，但以正負之方向不同。最後由多項回應方程式所得之標準值，則不再為直綫矣。以上係未分組材料間之多項直綫式相關，至于多項變數分組資

料及曲綫相關，則可依過去數章中所述之方法，加以引伸即可，只手續更加繁重耳。

## 第二節 偏相關

在多項相關所包括之變數中，一爲因變數，其餘均爲自變數。各自變數旣已視爲一複合自變數，故多項相關之要點，仍爲一因變與一自變數間之相關。在偏相關中，因變數仍爲一項，而自變數則可分爲二項或多項。吾人所分析者，已非各個自變數聯合對於因變數之聯繫，而係由各自變數中提出一項，假定其他各項不變，單獨分析此一自變數與因變數間之相關，每一自變數均如此單獨提出，分析其與因變數中間之相關。此種每一自變數與因變數之相關，稱之爲偏相關 (Partial Correlation) 或純相關 (Net Correlation)。偏相關與多項相關之區別已見上文，偏相關與簡單相關之區別，在前者承認各個自變數同時存在，爲分析偏相關乃提出其中之一項，而假定其他各項不變。在簡單相關中，只分析一自變數與因變數間之關係，根本不考慮其他自變數之存在也。

偏相關分析之程序，係先求簡單相關係數，再進而由此簡單相關係數以求偏相關係數。現仍利用上節多項相關中之實例以說明之：

(第120表) 三簡單相關係數計算表

| $X_1$ | $X_2$ | $X_3$ | $X_1^2$ | $X_2^2$ | $X_3^2$ | $X_1X_2$ | $X_1X_3$ | $X_2X_3$ |
|---|---|---|---|---|---|---|---|---|
| 0 | 4 | 1 | 0 | 16 | 1 | 0 | 0 | 4 |
| 1 | 4 | 3 | 1 | 16 | 9 | 4 | 3 | 12 |
| 3 | 3 | 2 | 9 | 9 | 4 | 9 | 6 | 6 |
| 6 | 2 | 5 | 36 | 4 | 25 | 12 | 30 | 10 |
| 8 | 0 | 4 | 64 | 0 | 16 | 0 | 32 | 0 |
| 18 | 13 | 15 | 110 | 45 | 55 | 25 | 71 | 32 |

上列第 120 表所列材料，與複相關所用者相同，只將複相關中所用之 $Y$ 代入以 $X_3$，餘均無改。由未分組材料求相關係數簡便法之公式爲：

$$r=\frac{\frac{\Sigma XY}{N}-\frac{\Sigma X}{N}\cdot\frac{\Sigma Y}{N}}{\sqrt{\frac{\Sigma X^2}{N}-\left(\frac{\Sigma X}{N}\right)^2}\cdot\sqrt{\frac{\Sigma Y^2}{N}-\left(\frac{\Sigma Y}{N}\right)^2}}$$

依此公式稍加補充，可求出三個簡單相關係數（$N=5$）

1. $$r_{12}=\frac{\frac{\Sigma X_1X_2}{N}-\frac{\Sigma X_1}{N}\cdot\frac{\Sigma X_2}{N}}{\sqrt{\frac{\Sigma X_1^2}{N}-\left(\frac{\Sigma X_1}{N}\right)^2}\sqrt{\frac{\Sigma X_2^2}{N}-\left(\frac{\Sigma X_2}{N}\right)^2}}=\frac{\frac{25}{5}-\frac{18}{5}\cdot\frac{13}{5}}{\sqrt{\frac{110}{5}-\left(\frac{18}{5}\right)^2}\sqrt{\frac{45}{5}-\left(\frac{13}{5}\right)^2}}$$

$$=\frac{5-9.36}{\sqrt{22-(3.6)^2}\sqrt{9-(2.6)^2}}\qquad=\frac{-4.36}{\sqrt{22-12.96}\sqrt{9-6.76}}$$

$$=\frac{-4.36}{\sqrt{9.04}\sqrt{2.24}}\qquad=\frac{-4.36}{3.006\times1.496}$$

$$=-4.36/4.496976\qquad=-0.9689$$

2. $$r_{13}=\frac{\frac{\Sigma X_1X_3}{N}-\frac{\Sigma X_1}{N}\cdot\frac{\Sigma X_3}{N}}{\sqrt{\frac{\Sigma X_1^2}{N}-\left(\frac{\Sigma X_1}{N}\right)^2}\sqrt{\frac{\Sigma X_3^2}{N}-\left(\frac{\Sigma X_3}{N}\right)^2}}=\frac{\frac{71}{5}-\frac{18}{5}\cdot\frac{15}{5}}{\sqrt{\frac{110}{5}-\left(\frac{18}{5}\right)^2}\sqrt{\frac{55}{5}-\left(\frac{15}{5}\right)^2}}$$

$$=\frac{14.2-3.6\times3}{\sqrt{22-(3.6)^2}\sqrt{11-(3)^2}}\qquad=\frac{14.2-10.8}{\sqrt{22-12.96}\sqrt{11-9}}$$

$$=\frac{3.4}{\sqrt{9.04}\sqrt{2}}\qquad=\frac{3.4}{3.006\times1.414}$$

$$=3.4/4.250484\qquad=0.7996$$

$$3.\quad r_{23}=\frac{\frac{\Sigma X_2X_3}{N}-\frac{\Sigma X_2}{N}\cdot\frac{\Sigma X_3}{N}}{\sqrt{\frac{\Sigma X_2^2}{N}-\left(\frac{\Sigma X_2}{N}\right)^2}\sqrt{\frac{\Sigma X_3^2}{N}-\left(\frac{\Sigma X_3}{N}\right)^2}}=\frac{\frac{32}{5}-\frac{13}{5}\times\frac{15}{5}}{\sqrt{\frac{45}{5}-\left(\frac{13}{5}\right)^2}\sqrt{\frac{55}{5}-\left(\frac{15}{5}\right)^2}}$$

$$=\frac{6.4-2.6\times 3}{\sqrt{9-(2.6)^2}\sqrt{11-3^2}}=\frac{6.4-7.8}{\sqrt{9-6.76}\sqrt{11-9}}$$

$$=\frac{-1.4}{\sqrt{2.24}\sqrt{2}}=\frac{-1.4}{1.496\times 1.414}$$

$$=-1.4/2.115344=-0.6614$$

以上係根據未分組材料 $X_1, X_2, X_3$ 三變數以求相關係數之法，與二項相關中求簡單相關係數完全相同。如爲分組材料則材料之形式爲三相關表。由三相關表計算三簡單相關係數時，又與二項相關相同，求得三簡單相關係數之後，即依而求偏相關係數，其公式如下：

1．$X_1$ 與 $X_2$ 間之偏相關係數，假定 $X_3$ 固定不變。

$$r_{12\cdot 3}=\frac{r_{12}-r_{13}\cdot r_{23}}{\sqrt{1-r^2_{23}}\sqrt{1-r^2_{23}}}=\frac{-.9689-.7996\times(-.6614)}{\sqrt{1-(.7996)^2}\sqrt{1-(-.6614)^2}}$$

$$=-0.44/\sqrt{.360619\times 5625}=-4.4/0.45038$$

$$=-0.977$$

2．$X_1$ 與 $X_3$ 間之偏相關係數，假定 $X_2$ 固定不變。

$$r_{13\cdot 2}=\frac{r_{13}-r_{12}\cdot r_{23}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{23}}}=\frac{0.7996-(-.9689)(-.6614)}{\sqrt{1-(-.9689)^2}\sqrt{1-(-.6614)^2}}$$

$$=\frac{0.7996-0.6408}{\sqrt{1-.9388}\sqrt{1-.3374}}=\frac{0.1588}{\sqrt{.0612}\sqrt{.5625}}$$

$$=0.1588/0.1855=0.8561$$

3. $X_2$ 與 $X_3$ 間之偏相關係數，假定 $X_1$ 固定不變，

$$r_{23\cdot 1}=\frac{r_{23}-r_{12}\cdot r_{13}}{\sqrt{1-r^2_{12}}\sqrt{1-r_{13}^2}}=\frac{-.6611-(-.9689)\times.7996}{\sqrt{1-(-.9689)^2}\sqrt{1-(.7995)^2}}$$

$$=\frac{-.6614+.7747}{\sqrt{1-.9388}\sqrt{1-.63938}}=\frac{0.1133}{\sqrt{0.0612}\sqrt{0.36062}}$$

$$=\frac{0.1133}{\sqrt{0.02206944}}\qquad=\frac{0.1133}{0.1485}$$

$$=0.7629$$

同理可求三估計直綫及各估計綫之估計標準誤差

# 第 四 編

# 物 價 指 數

# 第十四章　物價指數之性質

## 第一節　價比之意義及種類

指數（Index Numbers）係指百分數；用以表示統計數列中之相對變異者。統計數列有屬性數列，次數分配與時間數列三種，自理論言之：三者中所包括之變異，俱可用指數表示之，但實際上指數之主要功用在表示時間數列中所包括之變異；特別是物價中之變異。以指數表示物價之變動，乃稱之爲物價指數（Index Numbers of Prices）。所謂物價指數，即係利用價格百分比（Price Relatives），或簡稱價比，以表示價格變動者。

同一商品之價格因時而有不同，自其因時不同乃成爲價格數列（Price Series）。比較各時期價格時，每將實際價格化爲相對之百分比，即價比。所謂價比，即一時期價格與另一時期價格間之比例，此比例每以100乘之，使成爲百分數形式。求價比時，須定有基期（Base Period），基期價格稱爲基價（Base Price），以$P_o$表示之；基期之價比爲100，由$\frac{P_o}{P_o}\times 100$而來，其他各時期稱爲計算期（Given Periods），其價格爲計價（Given Prices），以$P_t$表示之，其價比爲$\frac{P_t}{P_o}\times 100$，基期制度有三，價比亦因可分三種：

一、固定基期價比（Fixed Base Relatives）

所謂固定基期價比，係就各時期中選擇其一爲基期，固定不變，使其價比爲100，其他各時期均爲計算期，俱以基期價格爲共同分母，將各計算期價格化爲百分數，其公式爲$\frac{P_t}{P_o}\times 100$，現以1936年至1946年每年一月十五日美棉平均價格爲例，以見價比之計算方法：

(第121表) 美棉價比計算表。

| 年　份(一月五日) | 每磅平均價格 | 固定基期價比 | 連環價比 | 連鎖價比 |
|---|---|---|---|---|
| 40 | 10.12¢ | 100.00 | —— | 100.000 |
| 41 | 9.37 | 92.59 | 92.59 | 92.590 |
| 42 | 17.82 | 176.08 | 190.18 | 176.088 |
| 43 | 19.74 | 195.06 | 110.77 | 195.053 |
| 44 | 20.15 | 199.11 | 102.08 | 199.110 |
| 45 | 20.20 | 199.60 | 100.25 | 199.860 |
| 46 | 22.36 | 220.95 | 110.69 | 220.946 |

*註：表中資料見於1947年世界年鑑(World Almanac for 1947)第三八六頁

上列第121表中之固定基期價比，係以1940年一月十五日之平均價格爲基價，其價比爲100，但基期價格可爲一日，一月，一年或數日，數月或數年之平均價格，殊無定則，惟其中以數月或數年之平均價格爲較佳，此數月或數年又宜選擇其經濟情況正常，物價穩定之時間。在社會不安物價狂烈波動之時間，固不宜爲基期。卽在經濟安定期間，一日或一月之價格，亦容有偏高或偏低，由數月以至數年平均之結果，則可高低抵銷，得到折中之基價。以此基價爲標準所得之指數，方最宜於代表物價相對變動之程度。選擇基期又須注意者卽基期去計算期不宜過遠。如太遠時可設法換爲較近之時期，如自1913改爲1926，自1926改爲1936。

二、連環價比 (Link Relatives)

固定基期價比之缺點，卽年份甚多，各年之價格變動過大，化爲價比時，彼此懸殊，不便比較。補救方法爲不用固定基期，改爲移動基期。其辦法爲以上一年爲本年之基期，以本年再爲下一年之基期，輪流基期以求價比，稱爲連環價比，或簡稱環比，以$L$表示之。如1940年棉價爲10.12¢，1941年棉價爲9.37¢，以1940年爲基期，則1941之環比，卽$L_{1941}=\frac{9.37}{10.12}\times$

100=92.59，1942 年之棉價爲 17.82ƒ，以 1941 年爲基期，則 1942 年之環比，即 $L_{1942}=\frac{17.82}{9.37}\times 100=190.18$，其他各年之環比均可依同法求出。此種環比之優點，即在將本年與上年或任何兩鄰年之關係表出。如以1940年爲準，則 1941 年之棉價下跌 7.41%；以 1941 年爲準，1942 年上漲 90.18%；同理以 1942 年爲準，1943 年上漲 10.77%；以 1943 年爲準，則 1944 年上漲 2.08%；以 1944 年爲準，則 1945 年上漲 0.25%；以 1945 年爲準，則 1946 年上漲 10.69%。此種環比與固定基期之價比相較時，所予吾人之印象自更爲深刻，尤以在物價狂動時間，價格月與月不同，週與週不同，則每月或每週求出環比，以見短期中物價之變，其功用更大。

三、連鎖價比 (Chain Relatives)

所謂連鎖價比，或簡稱爲鎖比，以 $C$ 表示之，係由環比變化而來，或言爲環比還原於固定基期價比之方法。第一年之固定價比旣爲 100，是第一年之鎖比，即 $C_{1940}$ 亦爲 100。第二年之鎖比等於上年之鎖比乘本年之環比，以符號表示之，$C_{1941}=C_{1940}\times L_{1941}=100\times 92.59=92.59$。同理，$C_{1942}=C_{1941}\times L_{1942}=92.59\times 190.18=176.088$。其他各年之鎖比均依同法求出，并將固定，連環與連鎖三種價比列入同一表中，可資比較。由第121表中可見上年鎖比與本年環比相乘之結果，恰等於本年之固定基期價比。以1940 與 1942 兩年之符號表示之：$C_{1940}\times L_{1942}=R_{1042}$。此種還原關係只能適合單一商品之價比。如商品之種類不只一種，物價指數係由各種價比求平均數而來，則此種還原關係即不復存在。

## 第二節　物價指數之意義及種類

所謂物價指數係利用價格百分比即價比，以表示物價之變動者，是以物價指數實係由價比中引伸而來，詳言之，物價指數可具三層意義：

一、單一商品之價比

吾人比較一種商品價格在不同時期中之起伏消漲，每將實際價格化爲相對價格即價比。求價比時或用固定基期，以見各年各月之價格較基期上漲或下落若干。或用移動基期，以見年與年，月與月或週與週短時間此種價格上漲或下落之限度。物價以外，其他時間數列如消費，生產，交易以及人口之出生，死亡與移動等，無一不可藉百分比以見其在各年各月中相對變動之情形。

二、多組價比之平均數

物價指數最簡單之形式，雖可爲單一商品之價比，但其基本意義，則指由多種物價數列中所求得之平均數而言。經濟生活中商品種類至夥，時間一過，甲種物價可上升，乙種物價可下落，同爲上升，其程度各異，同爲下落，其程度亦異。吾人於此紛亂情形中，欲斷定一般物價水準之究爲上升或下落必須由各種價格中求平均數方可。個別物價猶如飛蜂亂飛，而物價指數所代表之水準，則猶如蜂羣之總動向。惟商品之稀有性及單位各不相同，故各商品單位價格，亦彼此相去懸殊，必須先將各商品之實際價格化爲相對價比之後，再由各價比求平均數，所得結果爲物價指數，代表平均價格之相對變動，不能卽直接由各實際價格逕求平均數也。

三、加權價比之平均數

上列所言物價指數係指實價化爲價比，再由價比求一平均數，以代表物價之平均變動。所注意者爲物價之本身，而未及其數量，因稱爲簡單物價指數(Simple Index-Numbers)。惟各種商品之消費，生產或交易數量，大不相同，亦卽各種商品在吾人經濟生活中所佔之重要性不同。在簡單物價指數中，旣只由各種商品之價比而求平均數，是將各種價比一視同仁，實係認爲各種商品同等重要。揆之實情，確不相同，吾人于編製物價指數

時，除注意價格一因素外，對於各種商品之數量亦須注意，不但使各種物價之變動表現於物價指數之中，各種商品之數量多寡，亦使其反映於其中，所得之結果爲加權指數（Weighted Index Numbers）。

# 第十五章　簡單物價指數

## 第一節　簡單總合指數

編製物價指數時，可以包括之因素有二，即價格與數量。如只依價格編製，不顧數量，則所得之指數稱爲簡單物價指數(Simple Index Numbers of Prices)，或稱爲未加權物價指數(Unweighted Index Numbers of Prices)。簡單指數中最易于編製且易于了解者爲簡單綜合指數(Simple Aggregative Index of Actual Prices)。現以 1940 以來美國三種糧食價格爲例，以說明綜合指數編製之程序。

（第 122 表）　美國三種糧價簡單綜合指數計算表

單位：每蒲耳價值美圓

| 年份（一月十五日） | 小麥價格 $P_i'$ | 包穀價格 $P_i''$ | 燕麥價 $P_i'''$ | 綜合價格 $\Sigma P_i$ | 綜合指數 $\frac{\Sigma P_i}{\Sigma P_0}$ |
|---|---|---|---|---|---|
| 1940 | $0.85 | $0.5 | $0.36 | $1.74 | 100.0[illegible] |
| 1[illegible]4 | 1.06 | 0.73 | 0.50 | 2.30 | 131.61 |
| 1944 | 1.46 | 1.13 | 0.78 | 3.37 | 1[illegible]3.68 |
| 1946 | 1.54 | 1.10 | 0.72 | 3.36 | 1[illegible]3.10 |

註：上表資料見1947年度世界年鑑第三八六頁

上表所載係三種糧食在四年度中一月十五日之平均價格。第一種商品即小麥之價格以 $Pi'$ 代表之，第二種商品即包穀之價格以 $Pi''$ 代表之，第三種商品即燕麥之價格以 $Pi'''$ 代表之。三者因時而異，自各爲一變數，或各爲一價格數列，現欲由三價格數列求一新數列，以表示三種糧價之平均變動，最簡單之方法爲將同一年中三項單位價格相加，所得之總

數，係小麥，包穀及燕麥各一蒲耳在一年中價格之總和，稱爲綜合價格(Aggregates of Actual Prices)，以$\Sigma P_a$表示之，各年中商品之種類及數量未變，而綜合價格不同，如1940年之綜合價格爲1.74圓，1942年爲2.30圓，1944年爲3.37圓，1946年爲3.36圓，卽足以表示三種糧食在各年中綜合或平均變動情形。惟實際價格不便比較，因以1940年爲基期年，使其綜合價格爲100.00，將其他各年之綜合價格俱化爲1940年卽基年價格之百分數，結果爲簡單綜合指數，1942年之指數爲131.61，1944年爲193.68，，1946年爲193.10。

簡單綜合指數之優點爲易于編製，並易爲讀者所理解。其缺點則在名義上未加權數，實際上因各種商品之單位不同或單位價格不同，無形中包括有偶然之權數。以吾人所舉例而言，三種糧食之單位同爲蒲耳，但三者之單位價格高低不同，就中小麥之價格最高，包穀次之，燕麥最低，其單位價格高者在綜合價格中所佔之比例卽大，反之，單位價格低者在綜合價格中所佔之比例卽小。如1940年之綜合價格爲$1.74，以之爲100%時，則三種單位價格所佔之百分比如下：

(第123表)

| 糧食種類 | 單價 | 百分比 |
|---|---|---|
| 小麥 | $0.85 | 48.85 |
| 包穀 | 0.53 | 30.45 |
| 燕麥 | 0.36 | 20.70 |
| 總計 | $1.74 | 100.00 |

由上表可知小麥價格在總合價格中所佔之百分比最高，爲48.85%，包穀價格次之，佔30.45%，燕麥最低，佔20.70%。凡單位價格在綜合價格中佔分比高者，則對綜合指數之影響卽大；反之，其佔百分比低者，對綜

合指數影響卽小。是以各種商品單價之高低無形中構成權數。此外，各種商品單位不同時所構成之權數尤爲顯著。如吾人以鐵銅銀三種金屬單位價格編製綜合指數，鐵之單位爲噸，銅之單位爲磅，銀之單位爲兩，如某年中鐵每噸之價爲美幣二十元，銅每磅之價爲一角，銀每兩之價爲五角，則此三種金屬之綜合指數必受鐵價變動之支配，因鐵價在綜合價格中佔極高之百分比，是鐵價上漲，則指數隨之上漲，鐵價下落，指數隨之下落。其他兩種金屬之價格以單位小，單價低，在指數中幾完全埋沒。但商品之單位不同，由于各種商品數量之多少及市場中交易之習慣，絕不足以代表其相對重要性之高下。總之，在綜合指數中，無論單位不同，或單價不同，但足以形成不合理之權數，使所得之指數，不足以代表實際情形。

## 第二節 簡單算術平均數指數

此種指數編製之程序有二步驟：

1.各種商品之實際價格俱化爲價比。

2.由同一年之各價比中，求出算術平均數。現仍依三種糧價爲例，以說明編製之程序。

（第124表） 三種糧價簡單算術平均指數計算表

| 年份<br>（一月十五日） | 小麥<br>$p_1'$ | 小麥<br>$\frac{p_1''}{p_0''}$ | 玉米<br>$p_1''$ | 玉米<br>$\frac{p_1''}{p_0''}$ | 燕麥<br>$p_1'''$ | 燕麥<br>$\frac{p_1'''}{p_0'''}$ | $\Sigma\frac{p_1}{p_0}$ | $I=\frac{\Sigma\frac{p_1}{p_0}}{N}$ |
|---|---|---|---|---|---|---|---|---|
| 1[illegible]0 | [illegible]0.85 | 100 0[illegible] | [illegible]0.5[illegible] | 100.00 | [illegible]0.3[illegible] | 100.00 | 300.00 | 100.00 |
| 1942 | 1.0[illegible] | 1[illegible]4.[illegible]1 | 0.7[illegible] | 13[illegible].74 | 0.[illegible]0 | 138.89 | [illegible]01.3[illegible] | 133.78 |
| 1944 | [illegible] | 1[illegible]1.7[illegible] | [illegible].13 | 13[illegible]1 | 0.78 | 216.67 | [illegible]01.6[illegible] | [illegible]00.55 |
| 194[illegible] | 1.[illegible] | [illegible]8.1[illegible] | 1.10 | [illegible] | 0.72 | 200.00 | [illegible]86.54 | 196.21 |

現在求三種商品之價比與單一商品求價比之手續相同，以 $p_1'/p_0'$，$p_1''/p_0''$，$p_1'''/p_0'''$，等符號表示之，求得各組價比數列後，每年由各

價比中求一平均數，即爲簡單算術平均數指數 在次數分配中，變數爲$X$，平均數公式爲$\bar{X}=\frac{\Sigma X}{N}$，現在之變數爲$R=\frac{p_i}{p_o}$，平均數即指數公式爲$I=\Sigma\frac{p_i}{p_o}/N$。依此公式所得之平均數，列入計算表中最右一欄中，即爲所求之指數。

簡單算術式指數之優點爲易于計算，易于了解。其缺點即受極大極小變項之影響，使指數之值過高或過低，此其一；此外算術式指數對於物價之上漲比對物價之下落較爲加重，試舉例言之：

(第12 表)

| 年份 | 甲種商品 | | 乙種商品 | | $\Sigma\frac{p_i}{p_o}$ | $\frac{\Sigma\frac{p_i}{p_o}}{N}$ |
|---|---|---|---|---|---|---|
| | $p_i{}'$ | $\frac{p_i{}'}{p_o{}'}$ | $p_i{}''$ | $\frac{p_i{}''}{p_o{}''}$ | | |
| 1940 | $1.00 | 100.0 | $2.00 | 100.0 | 200.0 | 100.0 |
| 1946 | 2.00 | 200.0 | 1.00 | 50.0 | 250.0 | 12[illegible].0 |

在第二年中甲種商品自1元上漲爲2元，其價比爲200，乙種商品價格自2元下落爲1元，其價比爲50，平均之結果，計算年之指數則爲125，足證簡單算術式對物價之上漲特別加重也。

## 第三節　簡單中位數指數

簡單價比中位數指數 (Simple Medians of Prices Relatives) 編製之程序可分爲二步：

1.求各種商品之價比，與上法同。

2.就中一年各價比中求出中位數，即爲本年之指數，用以表示本年價格之平均變動，如一年中之價比項目甚多，必依其大小排成序列(Array)，再由序列中求出中位項，其值爲中數。反之，如同一年中價比項目無多，則

一望卽得中位數，下表所列只三種商品價格，卽每一年中只有三項價比，是可用觀察法確定中位指數。

(第 126 表) 三種糧價簡單中位數指數計算表

| 年份 (一月十五日) | 小麥 | | 玉米 | | 燕麥 | | L |
|---|---|---|---|---|---|---|---|
| | $p_1'$ | $\frac{p_1'}{p_0'}$ | $p_1''$ | $\frac{p_1''}{p_0''}$ | $p_1'''$ | $\frac{p_1'''}{p_0'''}$ | |
| 19.0 | $ 0.8 | .00.00 | $ 0.53 | 100.00 | $ .36 | 100 00 | 100.00 |
| 19 2 | 1.06 | 124.71 | 0.73 | 137.74 | . 0 | 138.89 | 137.74 |
| 1944 | 1.46 | 171.76 | 1.13 | 213.21 | .78 | 216.67 | 213.21 |
| 1946 | 1.54 | 181.18 | 1.10 | 207.55 | .72 | 200.00 | 200 00 |

上表最右一欄所得爲中位數指數，與簡單算術式指數比較時，其顯著優點有三：

1. 易于計算易于了解。

2. 指數不受極端項目之影響，中位數之意，卽介於各極端項目中間項目之值。

3. 不如算術平均式之過份加重價格之上漲。

中位數之缺點，卽同一年中價比項目過少時，中位數不足以代表各價比或平均價格之變動，但如項目多時，則中位式較其他任何簡單指數公式爲優。

## 第四節 幾何平均式指數

幾何平均數指數 (Simple Ge metric Average of Price Relatives) 計算方法，一如普通幾何平均數，須利用對數表，其計算之程序如下：

1. 求各商品之價比，卽 $\frac{p_1'}{p_0'}$，$\frac{p_1''}{p_0''}$，$\frac{p_1'''}{p_0'''}$，……

2. 查各價比之對數，$log\left(\frac{p_1'}{p_0'}\right)$，$log\left(\frac{p_1''}{p_0''}\right)$，$log\left(\frac{p_1'''}{p_0'''}\right)$，……

3.求各價比對數之平均數，其公式爲 $\Sigma log\left(\frac{p_1}{p_0}\right)/N$

4.查上項所得平均數之逆對數，即爲簡單幾何式指數，其公式爲：

$$I = log^{-1}\frac{\Sigma log\left(\frac{p_1}{p_0}\right)}{N}$$

(第127表) 三種糧價簡單幾何式指數計算表

| 年份（一月十五日） | 小麥 | | | 玉米 | | | 燕麥 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $p_1{}'$ | $\frac{p_1{}'}{p_0{}'}$ | $log\frac{p_1{}'}{p_0{}'}$ | $p_1{}''$ | $\frac{p_1{}''}{p_0{}''}$ | $log\frac{p_1{}''}{p_0{}''}$ | $p_1{}'''$ | $\frac{p_1{}'''}{p_0{}'''}$ | $log\frac{p_1{}'''}{p_0{}'''}$ |
| 1940 | $ 0.8[illegible] | 100.00 | 2.0000 | $ 0.53 | 100.00 | 2.0000 | $ 0.36 | 100.00 | 200.00 |
| 1942 | 1.06 | 124.71 | 2.0959 | 0.73 | 137.74 | 2.1391 | 0.50 | 138.89 | 2.1427 |
| 1944 | 1.46 | 171.76 | 2.2349 | 1.13 | 213.21 | 2.3288 | 0.78 | 216.67 | 2.3358 |
| 1946 | 1. 4 | 181.1[illegible] | 2.2581 | 1.10 | [illegible]07.55 | 2.317[illegible] | 0.72 | 200.00 | 2. 010 |

續上表

| 年份（一月十五日） | $\Sigma log\frac{p_1}{p_0}$ | $\frac{\Sigma log\frac{p_1}{p_0}}{N}$ | I |
|---|---|---|---|
| 1940 | 6.0000 | 2.0000 | 100.00 |
| 1942 | 6.3777 | 2.1259 | 133.63 |
| 1944 | 6.8995 | 2.2998 | 199.54 |
| 1946 | 6.8762 | 2.2921 | 195.39 |

簡單幾何式指數之優點有二:

1.對于各價比之上漲或下落,所加之重量相等,非如算術式之畸重物價之上漲,可自下表見之:

(第128表)

| 年份 | 甲種商品 | | | 乙種商品 | | | $\Sigma\log\frac{p_1}{p_0}$ | $\frac{\Sigma\log\frac{p_1}{p_0}}{N}$ | Antilog of $\Sigma\log\frac{p_1}{p_0}/N$ |
|---|---|---|---|---|---|---|---|---|---|
| | $p_1'$ | $\frac{p_1'}{p_0'}$ | $\log\frac{p_1'}{p_0'}$ | $p_1''$ | $\frac{p_1''}{p_0''}$ | $\log\frac{p_1''}{p_0''}$ | | | |
| 1940 | $ 1.00 | 100.0 | 2.000 | $ 2.00 | 100.0 | 2.000 | 4.000 | 2.030 | 100.0 |
| 1946 | 2.00 | 200.0 | 2.301 | 1.00 | 50.0 | 1.699 | 4.030 | 2.000 | 100.0 |

上表中最右一欄所載之指數俱爲 100.0,卽表示幾何式之對於物價上漲與下落,無畸重畸輕之弊。

2.幾何指數更換基期之手續簡易:編製指數時選擇基期之標準有二:一爲基期係經濟情况正常物價穩定之時期,二爲基期不易過遠,過遠則與現在比較不易,據此觀之,基期爲1913不如1926,1926又不如1936,但1937以後之時期不宜作基期之用,因1937蘆溝橋事變爆發,中日戰事開始後,世界大戰不久隨踵而起,經濟脫節,物價狂漲,不合乎基期條件。基期不宜太遠,由遠改近之方法有普通與簡捷之別:普通法卽另選較近之年份爲基年,使其價比爲100,進而求各價比,由各價比求指數。此種方法手續繁重。另有簡捷法卽以新定基年之指數爲100,其他各年之計算年,其指數俱除以基年之原指數,結果之指數已更換爲新基年指數,如原基年爲1926,現擬改基年爲1936時,卽以1936指數爲共同分母而除各年之指數,則所得爲以1936年爲基年之指數,此種簡捷改變基期方法,只有簡單幾何式可以適用,依普通及簡捷方法,更換基期所得之指數大致相同,其他指數公式則不合於此,故幾何式計算手續雖稍爲繁重,但仍爲最流行之指數公式。

## 第五節　倒數平均數指數

倒數平均數指數（Harmonic Average of Prices Relatives）計算之程序與倒數平均數相同，其步驟如下：

1.求各價比：$\frac{p_i'}{p_o'}$，$\frac{p_i''}{p_o''}$，$\frac{p_i'''}{p_o'''}$，……

2.求各價比之倒數：$1/\frac{p_i'}{p_o'}$，$1/\frac{p_i''}{p_o''}$，$1/\frac{p_i'''}{p_o'''}$，……，各價比倒數亦可書爲：$\frac{p_o'}{p_i'}$，$\frac{p_o''}{p_i''}$，$\frac{p_o'''}{p_i'''}$，……

3.求各倒數之平均數，其公式爲：$\Sigma\frac{p_o}{p_i}/N$

4.求上列所得平均數之倒數，卽爲倒數平均數指數，其公式如下：

$$I=\frac{N}{\Sigma\frac{p_o}{p_i}}$$

（第129表）　三種糧價倒數平均數指數計算表

| 年份（1月15日） | 小麥 | | | 玉米 | | | 燕麥 | | | $\Sigma\frac{p_o}{p_i}$ | $I=\frac{N}{\Sigma\frac{p_o}{p_i}}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_i'$ | $\frac{p_i'}{p_o'}$ | $\frac{p_o'}{p_i'}$ | $p_i''$ | $\frac{p_i''}{p_o''}$ | $\frac{p_o''}{p_i''}$ | $p_i'''$ | $\frac{p_i'''}{p_o'''}$ | $\frac{p_o'''}{p_i'''}$ | | |
| 1940 | $ 0.85 | 100.00 | 010000 | $ 0.53 | 100.00 | 010000 | $ 0.36 | 100.00 | 010000 | .030000 | 100.00 |
| 1942 | 1.06 | 124.71 | 005019 | 0.73 | 137.74 | 007210 | 0.50 | 138.89 | 007200 | 0.022579 | 132.87 |
| 1944 | 1.46 | 171.76 | 005822 | 1.13 | 213.21 | 004630 | 0.78 | 216.67 | 004615 | 0.015127 | 198.32 |
| 1946 | 1.54 | 181.16 | 009519 | 1.10 | 207.55 | 004818 | 0.72 | 200.00 | 005000 | 0.015337 | 195.61 |

（第130表）　四種簡單指數比較表

| 年份 | 綜合式 | 算術式 | 中位式 | 幾何式 |
|---|---|---|---|---|
| 1940 | 100.00 | 100.00 | 100.00 | 100.00 |
| 1942 | 131.61 | 133.78 | 137.74 | 133.63 |
| 1944 | 193.68 | 200.55 | 213.21 | 199.45 |
| 1946 | 193.10 | 196.24 | 200.00 | 195.93 |

# 習　題

（第131表）　三種金屬批發價格與生產數量表

價格單位：美元

| | 價格 | | | 生產量 | | |
|---|---|---|---|---|---|---|
| | 鐵（單位：噸） | 鋼（單位：噸） | 銅（單位：磅） | 鐵（000噸） | 鋼（000噸） | 銅（000磅） |
| 1926 | $21.32 | $43.00 | $ .138 | 38,181 | 3,218 | 2,323 |
| 27 | 20.44 | 43.00 | .130 | 34,867 | 2,806 | 2,[illegible]26 |
| 28 | 19.21 | 43.00 | .146 | 39,304 | 2,647 | 2,468 |
| 29 | 20.54 | 43.00 | .181 | 41,549 | 2,722 | 2,740 |
| 30 | 20.26 | 43.00 | .129 | 29,905 | 1,873 | 2,157 |
| 31 | 18.69 | 41.00 | .08[illegible] | 17,813 | 1,158 | 1,501 |
| 32 | 17.12 | 42.38 | .056 | 8,518 | 403 | 681 |
| 33 | 18.26 | 39.33 | .070 | 14,353 | 416 | 7.2 |
| 34 | 20.45 | 36.38 | .084 | 15,626 | 1,010 | 891 |
| 35 | 20.01 | 36.38 | .086 | 21,178 | 712 | 1,178 |
| 36 | 20.91 | 36.63 | .095 | 30,619 | 1,207 | 1,615 |

註：上表資料見：Frederick E. Croxton, Workbook in Applied General Statistics, 1937, P. 105.

一、試根據上表中價格資料，編製一三種金屬簡單綜合式指數。

二、試根據上表中價格資料，編製一三種金屬簡單算術平均式指數。

三、試根據上表中價格資料，編製一三種金屬簡單中位式指數。

四、試根據上表中價格資料，編製一三種金屬簡單幾何式指數。

五、試根據上表中價格資料，編製一三種金屬簡單倒數式指數。

# 第十六章　加權物價指數

## 第一節　加權之意義

上節所述五種指數，只根據商品之價格編製，而未顧及其數量，乃稱之爲簡單物價指數，或稱爲未加權物價指數。其中綜合式指數，名義上雖未加權數，但指數之編製，係根據多種商品之單位價格，無形中在實際上已加有權數。因商品之單位大者，價格亦高，在綜合式指數中所發生之作用亦大；反之，商品之單位小者，其價格亦小，在綜合式指數中所發生之作用亦小。由單位造成之權數，純出偶然，極不合理。在四種價比平均數指數中，爲避免商品單位所產生之不合理權數計，乃將實際價格化爲價比，再由價比而計算平均數。此種辦法，實際上，係將各種物價視爲同等重要，加以相等之權數。惟各種商品之生產，交易或消費之數量，既各不同，是在人類經濟生活中各種商品之重要性亦各不相同，故吾人編製指數時，於價格之外，必須顧及各商品之數量，所得即爲加權物價指數（Weighted Index Numbers of Prices）。

加權制度有二，即內含權數（Implicit Weight）及外加權數（Explicit Weight），所謂內含權數制度，即各種商品重要性不同，搜集物價資料時，對於重要商品，將其各地市場及各種牌號之價格多加收容；反之，對不重要之商品，則所搜集之價格數列減少，結果所造成之指數，表面上未曾加權，實已有權數在其中矣。所謂外加權數制度，即根據各種商品實際消費、生產、販賣之數量，加爲權數，所加者可爲實際數量（Physical Quantities）。以 $q$ 表示之，亦可爲價值（Values），即單位價格與實際數量之乘積，以

$pq$ 表示之。無論所加之權數爲數量或價值，各年所加者可爲常數，亦可爲變數。就數量而言，如所加爲基年生產數量($q_o$)，則爲常數。同一種商品各計算年度之數量，俱以基年生產量代表之。如所加爲各計算年之產量，以 $q_1$ 代表之，是每年之價格變，數量亦隨之而變，兩者均爲變數。但又可有二種折中情形：其一卽以普查年(Census Year)之數量爲權數，商品產量於每五年或十年普查時方能知之，非如價格資料之易於調查。故兩次普查年中間爲固定權數，每逢普查年份一改，又爲變動權數。其二卽爲以聯合數量爲權數。以 $q_o$ 卽基年數量爲權數犯下偏(Downward Bias)之弊，以 $q_1$ 卽各計算年數量爲權數犯上偏(Upward Bias)之弊，俱使所得之指數稍欠公允。爲調合上下偏計，乃以基年與每一計算年數量之和爲權數，以 $(q_o+q_1)$ 表示之。上下偏誤可以彼此抵銷，使結果所得之指數較爲中正可靠，現仍依上章所用三種糧價及下表中此三種糧食供給數量爲例，以說明加權指數編製之程序。

(第132表)　美國小麥玉米及燕麥三種糧食數量表

單位：百萬蒲耳

| 年份 | 小　麥 | 玉　米 | 燕　麥 |
|---|---|---|---|
| 1940 | 174(九月28日) | 63(十二月28日) | 10 (一月6日) |
| 1942 | 222(十月17日) | 61 (二月16日) | 11(十一月7日) |
| 1944 | 117 (一月1日) | 21 (二月26日) | 17 (十月14日) |
| 1946 | 144(九月22日) | 21 (三月3日) | 46 (十月20日) |

## 第二節　加權綜合指數

簡單綜合指數之公式爲 $I=\frac{\Sigma p_1}{\Sigma p_o}$。在加權綜合指數中，所加之權數爲實際數量，但有基年數量與計年數量之別，分別以 $q_o$ 及 $q_1$ 表示之。如用基

年數量，則加權綜合式爲：$I=\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$，稱之爲拉斯皮爾公式(Laspeyres' Formula)。如用計年數量則加權綜合式爲：$I=\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$，稱之爲巴賽公式(Paasche's Formula)，現依上章所用三種糧食之價格及上節所列三者之數量，以說明加權綜合指數編製之步驟，但價格資料之搜集較易，而數量資料之搜集則難，上章中所用之價格資料爲各年中一月十五日之平均價格，上節中所舉之數量則爲一年中全美供給最多一日之數量，就中一九四六年尚缺，姑且以一九四五年之數量代充之，相去當不太遠。

一、基年數量加權綜合指數計算表

(第133表)

數量單位：百萬蒲耳

| 年份 | 小麥 | | | 玉米 | | | 燕麥 | | | $\Sigma p_1 q_0$ | $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $q_0$ | $p_1 q_0$ | $p_1$ | $q_0$ | $p_1 q_0$ | $p_1$ | $q_0$ | $p_1 q_0$ | | |
| 1940 | \$0.85 | 174 | 147.90 | \$ 0.53 | 63 | 33.39 | \$ 0.36 | 10 | 3.60 | 184.89 | 100.00 |
| 1942 | 1.06 | 174 | 184.44 | 0.73 | 63 | 45.99 | 0.50 | 10 | 5.00 | 235.43 | 127.34 |
| 1944 | 1.46 | 174 | 254.04 | 1.13 | 63 | 71.19 | 0.78 | 10 | 7.80 | 333.03 | 180.12 |
| 1946 | 1.54 | 174 | 267.96 | 1.10 | 63 | 69.30 | 0.72 | 10 | 7.20 | 344.46 | 186.31 |

二、計年數量加權綜合指數計算表

(第134表)

數量單位：百萬蒲耳

| 年份 | 小麥 | | | | | 玉米 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $p_1 q_1$ | $p_0 q_1$ | $p_1$ | $p_0$ | $q_1$ | $p_1 q_1$ | $p_0 q_1$ |
| 194 | \$ 0.8 | \$ 0.85 | 174 | 147.90 | 147.90 | \$ 0.53 | \$ 0.53 | 63 | 33.39 | 33.39 |
| 1912 | 1.06 | 0.85 | 222 | 235.32 | 188.70 | 0.73 | 0.53 | 61 | 44.53 | 32.33 |
| 1944 | 1.46 | 0.85 | 117 | 170.82 | 99.45 | 1.13 | 0.53 | 21 | 23.73 | 11.13 |
| 1946 | 1.54 | 0.85 | 144 | 221.76 | 122.4 | 1.10 | 0.53 | 21 | 23.10 | 11.13 |

續　上　表

| 年份 | 燕 | | 麥 | | | $\Sigma p_1 q_1$ | $\Sigma p_0 q_1$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$ |
|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $p_1 q_1$ | $p_0 q_1$ | | | |
| 1940 | \$ 0.36 | \$ 0.36 | 10 | 3.60 | 3.60 | 184.89 | 184.89 | 100.00 |
| 1942 | 0.50 | 0.36 | 11 | 5.50 | 3.96 | 285.35 | 224.99 | 126.83 |
| 1944 | 0.78 | 0.36 | 17 | 13.26 | 6.12 | 207.81 | 116.70 | 178.07 |
| 1946 | 0.72 | 0.86 | 46 | 33.12 | 16.56 | 277.98 | 150.09 | 185.21 |

## 第三節　加權算術平均數指數

在綜合指數中，所加之權數爲各種商品之實際數量；在各種價比平均式指數中，則不能以數量爲權數，因各種商品之單位不同，不能直接以之與價比相乘。必先將各種不同之單位設法化爲相同之單位，其法卽以單位價格與數量相乘，其乘積爲商品之價值，以 $pq$ 表示之，價値旣俱以貨幣爲單位，自可以之與各價比相乘，然後求平均數。故綜合式中所加之權數爲數量 $q$，而各平均式所加之權數則爲價値 $pq$。但單價及數量各有基年與計算年之區別，基年價格爲 $p_o$，數量爲 $q_o$，計算年之價格爲 $p_1$，數量爲 $q_1$，此四因子交配結果可得四種價値，卽：$p_oq_o$，$p_oq_1$，$p_1q_o$ 及 $p_1q_1$。編製加權算術式指數，不能用第一種價値卽 $p_oq_o$ 爲權數，其理至易證明。以 $p_oq_o$ 爲權數之加權算術式爲：

$$\frac{\frac{p_1'}{p_o'}p_o'q_o'+\frac{p_1''}{p_o''}p_o''q_o''+\frac{p_1'''}{p_o'''}p_o'''q_o'''}{p_o'q_o'+p_o''q_o''+p_o'''q_o'''}$$

分子化簡後，全式變爲 $\frac{p_1'q_o'+p_1''q_o''+p_1'''q_o'''}{p_o'q_o'+p_o''q_o''+p_o'''q_o'''}=\frac{\Sigma p_1q_o}{\Sigma p_oq_o}$，是爲基年數量加權綜合指數公式。故編製加權算術式指數時，只能用其他三種價値爲權數，而不能用 $p_oq_o$。同理編製加權倒數指數時，不能用 $p_1q_1$ 爲權數，蓋

用 $p_1q_1$ 爲權數則加權倒數式爲：

$$\frac{p_1'q_1'+p_1''q_1''+p_1'''q_1'''}{\frac{p_o}{q_1}p_1q_1+\frac{p_o}{q_1}p_1q_1+\frac{p_o}{p_1}p_1q_1}$$

分母化簡後公式變爲 $\frac{p_1'q_1'+p_1''q_1''+p_1'''q_1'''}{p_o'q_1'+p_o''q_1''+p_o'''q_1'''}=\frac{\Sigma p_1q_1}{\Sigma p_oq_1}$，是爲計算年數量加權綜合指數公式。編製加權算術式既只能用 $p_1q_o$，$p_oq_1$，$p_1q_1$ 三種價值爲權數，就中 $p_1q_o$ 及 $p_oq_1$ 二種價值中，一爲基年數值，一爲計算年數值，基年數值犯下偏，計算年數值犯上偏，此上下偏誤相反之數值配合結果，使偏誤從中抵消，實爲優良之加權制度。二者相比，$p_1q_o$ 較 $p_oq_1$ 爲佳，蓋 $p_1q_o$ 中所需爲基年數量，各年不變，而 $p_oq_1$ 中爲各計年數量，搜集不易，至第四種價值即 $p_1q_o$ 一方以俱爲計算年之數值，同犯上偏，一方以資料不易搜集，不宜採用。

現以 $p_1q_o$ 爲權數而編製三種物價加權算術式指數，其公式爲

$$\frac{\Sigma\left(\frac{p_1}{p_o}p_1q_o\right)}{\Sigma p_1q_o}$$，計算表如下：

（第 135 表）　三種糧價加權算術平均數指數計算表

| 年份 | 小 | | | | 麥 | 包 | | | | 穀 |
|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $\frac{p_1}{p_o}$ | $q_o$ | $p_1q_o$ | $\frac{p_1}{p_o}p_1q_o$ | $p_1$ | $\frac{p_1}{p_o}$ | $q_o$ | $p_1q_o$ | $\frac{p_1}{p_o}p_1q_o$ |
| 1940 | \$ 0.85 | 100.00 | 174 | 147.90 | 14790.00 | \$ 0.53 | 100.00 | 63 | 33.39 | 3339.00 |
| 1942 | 1.06 | 124.71 | 174 | 184.44 | 23001.51 | 0.73 | 137.74 | 63 | 45.99 | 6334.66 |
| 1944 | 1.46 | 171.76 | 174 | 254.04 | 43333.91 | 1.13 | 213.21 | 63 | 71.19 | 15178.42 |
| 1946 | 1.54 | 181.18 | 174 | 267.96 | 48548.99 | 1.10 | 207.55 | 63 | 69.30 | 14383.22 |

續 上 表

| 年份 | 燕 | | | | 麥 | $\Sigma p_1 q_0$ | $\Sigma \frac{p_1}{p_0} p_1 q_0$ | $I=\frac{\Sigma \frac{p_1}{p_0} p_1 q_0}{\Sigma p_1 q_0}$ |
|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $\frac{p_1}{p_0}$ | $q_0$ | $p_1 q_0$ | $\frac{p_1}{p_0} q_1 q_0$ | | | |
| 1940 | $ 0.36 | 100.00 | 10 | 3.6 | 360.00 | 184.89 | 18 489.00 | 1,0.00 |
| 1942 | 0 50 | 138.89 | 10 | 5.0 | 694.45 | 230 43 | 30,030.62 | 127.56 |
| 1944 | 0.78 | 216.67 | 10 | 7.8 | 1190.03 | 333.03 | 60,502.36 | 181.67 |
| 1946 | 0.72 | 200.00 | 10 | 7.2 | 1440.00 | 344.46 | 64,372.21 | 1 6.68 |

## 第四節　加權幾何平均數指數

簡單幾何平均數指數公式爲：

$$I=\log^{-1}\frac{\Sigma\log\frac{p_1}{p_0}}{N}$$

現在可加之權數亦爲 $p_0q_0$，$p_1q_0$，$p_0q_1$ 及 $p_1q_1$ 四種價值，就中仍以第三種價值卽 $p_1q_0$ 爲佳，依此價值爲權數，則加權幾何平均數指數公式爲：

$$I=\log^{-1}\frac{\Sigma\log\frac{p_1}{p_0}p_1q_0}{\Sigma p_1q_0}$$

茲依上章中之價比對數及上節中之價值權數，加以計算，以資證明。至于加權倒數平均數指數，學者可以推想而知之，此中不加例證。

（第 13a 表） 三種糧價加權幾何平均數指數計算表

| 年份 | 小 | | | | | 麥 | 包 | | | | | 穀 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $\frac{p_1}{p_0}$ | $\log\frac{p_1}{p_0}$ | $q_0$ | $p_1 q_0$ | $\log\frac{p_1}{p_0} p_1 q_0$ | $p_1$ | $\frac{p_1}{p_0}$ | $\log\frac{p_1}{p_0}$ | $q_0$ | $p_1 q_0$ | $\log\frac{p_1}{p_0} p_1 q$ |
| 1940 | $ 0.85 | 100.0 | 2.0000 | 174 | 147.90 | 295.8000 | $ 0.53 | 100.00 | 2.0000 | 63 | 33.39 | 66.76 |
| 1942 | 1.06 | 124.71 | 2. 959 | 174 | 1 4.44 | 386.5678 | 0.73 | 137.74 | 2.1 91 | 63 | 34.99 | 98.377 |
| 1944 | 1.46 | 1 1.76 | 2.2349 | 174 | 254.04 | 567.7540 | 1.13 | 13.21 | 2.3288 | 63 | 71.19 | 165.787 |
| 1946 | 1.54 | 1 1.18 | 2.25 1 | 174 | 267.96 | 605.0805 | 1.10 | 207.55 | 2.3171 | 63 | 69.30 | 160.575 |

續　上　表

| 年份 | 燕 $p_1$ | $\frac{p_1}{p_0}$ | 麥 $\log\frac{p_1}{p_0}$ | $q_0$ | $p_1 q_0$ | $\log\frac{p_1}{p_0} p_1 q_0$ | $\Sigma p_1 q_0$ | $\Sigma \log\frac{p_1}{p_0} q_1 p_0$ | $\frac{\Sigma \log\frac{p_1}{p_0} q_1 p_0}{\Sigma p_1 q_0}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1940 | $ 0.36 | 100.00 | 2.0000 | 10 | 3.6 | 7.2000 | 184.89 | 369.8800 | 2.0000 |
| 1942 | 0.50 | 138.89 | 2.1427 | 1 | 5.0 | 10.7135 | 23[illegible].43 | 495.6585 | 2.10[illegible]3 |
| 1944 | 0.78 | 216.67 | 2.3358 | 10 | 7.8 | 18.2192 | 333.03 | 751.7605 | 2.2573 |
| 1946 | 0.72 | 200.00 | 2.3010 | 1[illegible] | 7.2 | 16.5672 | 344.[illegible]6 | 782.2227 | 2.27[illegible]9 |

上　續　表

| 年份 | $\log^{-1}\frac{\Sigma \log\frac{p_1}{p_0} p_1 q_0}{\Sigma q_1 p}$ |
|---|---|
| 1940 | 100.00 |
| 1942 | 127.44 |
| 1944 | 180.84 |
| 1946 | 186.61 |

## 第五節　費暄教授理想指數

在加權綜合指數公式中，所加之權數爲實際數量。就中有基年數量 $q_0$ 與計期數量 $q_1$ 之別。依基年數量所編之指數，即拉斯皮爾公式犯下偏，依計年數量所編之指數即巴裴公式犯上偏，美國經濟統計學家費暄教授（Professor Irving Fisher）在其所著物價指數編製論（The Making of Index Numbers）一書中，乃配合此兩種綜合指數公式，使其上偏與下偏從中抵銷。其綜合之法，即先依上列兩種綜合公式各求一指數，再將此兩種指數求出其幾何平均數，稱爲理想指數（The Ideal Index）現將其公式列舉於下，并依上文中已得之兩種加權綜合指數編製理想指數。

$$I=\sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}\cdot\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}$$

（第137表）　三種物價理想指數計算表

| 年份 | $\frac{\Sigma p_1 q_o}{\Sigma p_o q_o}$ | $\frac{\Sigma p_1 q_1}{\Sigma p_o q_1}$ | $\frac{\Sigma p_1 q_o}{\Sigma p_o q_o}\cdot\frac{\Sigma p_1 q_1}{\Sigma q_o p_o}$ | $\sqrt{\frac{\Sigma p_1 q_o}{\Sigma p_o q_o}\cdot\frac{\Sigma p_1 q_1}{\Sigma p_o q_1}}$ |
|---|---|---|---|---|
| 1940 | 100.00 | 100.00 | 10,000.0000 | 100.00 |
| 1942 | 127.34 | 126.83 | 16,150.5322 | 127.08 |
| 1944 | 180.12 | 178.07 | 32,073.9684 | 179.09 |
| 1946 | 186.31 | 185.21 | 34,506.4751 | 185.76 |

## 第六節　聯合數量加權綜合指數

理想式之特點，在利用求兩加權綜合指數幾何平均數之方法，將權數之偏誤抵銷。但其計算之手續繁雜，且不易爲讀者所了解，費暄教授因又提出一代替公式，彼稱之爲算術交配加權綜合公式 (Arithmetical Crossed Weight Aggregative)，其公式之形式爲：

$$I=\frac{\Sigma p_1(q_o+q_1)}{\Sigma p_o(q_o+q_1)}$$

仍爲加權綜合式，所用之權數爲基年與計年之聯合數量，故亦可稱爲聯合數量加權綜合公式，費暄教授所稱算術交配，係對其理想式中之幾何交配而言，此兩公式所需全部價格與全部數量之資料相同，交配之目的，在於取消偏誤，但幾何交配之手續繁雜，而算術交配之手續簡易，兩種交配之結果，幾完全相同，算術交配加權綜合式所得之指數與幾何交配加權綜合式或即理想式所差不逾0.25%，可見前者，實又優於後者，英國經濟學家哀基渥斯(Edgeworth) 及馬歇爾 (Alfred Marshall) 亦極力贊揚此式，故費暄教授特稱之爲哀基渥斯－馬歇爾公式。現計算如下：

（第138表） 聯合數量加權綜合指數計算表

| 年份 | 小 | | | | | | 麥 |
|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $q_0$ | $q_0+q_1$ | $p_1(q_0+q_1)$ | $p_0(q_0+q_1)$ |
| 1940 | 0.85 | 0.85 | 171 | 174 | 346 | 29[illegible].10 | 291.10 |
| 1942 | 1.06 | 0.85 | 222 | 174 | 395 | 418.70 | 335.75 |
| 1944 | 1.46 | 0.85 | 117 | 174 | 293 | 423.40 | 246.50 |
| 1946 | 1.54 | 0.85 | 144 | 174 | 317 | 488.18 | 269.75 |

續上表

| 年份 | 包 | | | | | | 類 |
|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $q_0$ | $q_0+q_1$ | $p_1(q_0+q_1)$ | $p_0(q_0+q_1)$ |
| 1940 | 0.53 | 0.53 | 63 | 63 | 126 | 66.78 | 66.78 |
| 19.2 | 0.13 | 0.53 | 61 | 63 | 124 | 90.52 | 65.72 |
| 1944 | 1.13 | 0.53 | 21 | 63 | 84 | 94.92 | 44.52 |
| 1946 | 1.10 | 0.53 | 21 | 63 | 84 | 92.40 | 44.52 |

續上表

| 年份 | 燕 | | | | | | 麥 |
|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $q_0$ | $q_0+q_1$ | $p_1(q_0+q_1)$ | $p_0(q_0+q_1)$ |
| 1940 | 0.36 | 0.36 | 10 | 10 | 20 | 7.20 | 7.20 |
| 1942 | 0.50 | 0.36 | 11 | 10 | 21 | 10.50 | 7.56 |
| 1944 | 0.78 | 0.36 | 17 | 10 | 27 | 21.06 | 9.72 |
| 1946 | 0.72 | 0.36 | 46 | 10 | 56 | 40.32 | 10.16 |

續上表

| 年份 | $\Sigma p_1(q_0+q_1)$ | $\Sigma p_0(q_0+q_1)$ | $\dfrac{\Sigma p_1(q_0+q_1)}{\Sigma p_0(q_0+q_1)}$ |
|---|---|---|---|
| 1940 | 358.08 | 368.08 | 100.00 |
| 1942 | 519.72 | 409.03 | 127.06 |
| 1944 | 539.38 | 300.74 | 179.33 |
| 1946 | 620.90 | 334.13 | 185.83 |

（第139表） 六種加權指數比較表

| 年份 | (1) 拉斯皮爾氏 | (2) 巴賽氏 | (3) 算術平均式 | (4) 幾何平均式 | (5) 理想式 | (6) 聯合數量式 | (6)—(5) |
|---|---|---|---|---|---|---|---|
| 1940 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 0 |
| 1942 | 127.34 | 126.83 | 127.56 | 127.44 | 127.08 | 127.06 | —.02 |
| 1944 | 180.12 | 178.07 | 181.67 | 181.84 | 179.09 | 179.35 | 0.26 |
| 1946 | 187.31 | 185.21 | 186.88 | 186.61 | 185.76 | 185.83 | 0.07 |

## 習 題

一、試根據上章習題中所載之三種金屬批發價格與生產數量，編製一加權算術式物價指數。

二、試根據同一資料，編製一加權幾何式物價指數。

三、試根據同一資料，編製一理想式物價指數。

四、試根據同一資料，編製一聯合數量加權綜合式物價指數。

五、試將上列所得四種指數連同前章所得各種簡單物價指數作一總比較表。

六、試繪一各種指數比較圖。

# 第十七章　物價指數之測驗

## 第一節　時間互換測驗

物價指數之公式甚多 在簡單指數公式中以何者為優?在加權指數公式中又以何者為優? 統計學家每舉出試驗之標準, 稱為物價指數之測驗 (Index Number Tests)。試驗簡單物價指數之方法為時間互換測驗 (Time Reversal Test), 或稱為基期互換測驗。吾人於比較前後兩時期之價格變動時,可以二者中之任一時期為基期,以另一時期為計算期, 如以前一期為基期,以後一期為計期,所得之指數稱為前進指數 (Forward Index)。反之,如以後一期為基期,以前一期為計期,則所得之指數為後退指數 (Backward Index)。求得前進指數與後退指數 後,可進而求二者之乘積,幷觀察其乘積是否等於一,如等於一, 則稱其指數公式合於時間互換測驗;反之,如其乘積不等於一,則稱其公式不合於時間互換測驗。分析一種商品之價格時, 以前後二期互為基期,作出前進與後退兩種指數,二者互為倒數,其乘積為一,理至顯明。如三十六年三月一日鷄蛋每枚價格為二百元,五月一日鷄蛋每枚之價格為四百元,先以三月為基期, 其價比為 1, 以五月為計期,其前進價比為 2,再以五月為基期,其價比為 1,以三月為計期,其後退價比為 0.5, 前進價比與後退價比之乘積為: $2\times0.5=1$。如商品不只一種, 則依兩時期之價格資料分別求出前進與後退兩指數,二種指數之乘積則不必為 1, 其乘積不等於 1 者, 即不合於時間互換測驗,現以簡單物價指數章中 1940 與 1946 兩年資料為例,分別試驗各種簡單指數公式是否合乎時間互換測驗.

一、簡單綜合指數時間互換測驗計算表（第140表）

| 年份 | 小麥 $p_t$ | 包穀 $p_t$ | 燕麥 $p_t$ | $\Sigma p_t$ | $\frac{\Sigma p_t}{\Sigma p_o}$ | $\frac{\Sigma p_o}{\Sigma p_t}$ |
|---|---|---|---|---|---|---|
| 1940 | $ .85 | $ 0.53 | $ 0.36 | $ 1.74 | 1.0000 | 0.5179 |
| 1946 | 1.54 | 1.10 | 0.72 | 3.36 | 1.9313 | 1.0000 |

時間互換測驗：

$$T=\frac{\Sigma p_t}{\Sigma p_o}\cdot\frac{\Sigma p_o}{\Sigma p_t}=1.9310\times0.5179=1.0000$$

二、簡單算術平均數指數時間互換測驗計算表（第141表）

| 年份 | 小麥 | | 包穀 | | 燕麥 | | $\Sigma\frac{p_t}{p_o}$ | $\frac{\Sigma\frac{p_t}{p_o}}{N}$ |
|---|---|---|---|---|---|---|---|---|
| | $p_t$ | $\frac{p_t}{p_o}$ | $p_t$ | $\frac{p_t}{p_o}$ | $p_t$ | $\frac{p_t}{p_o}$ | | |
| 1940 | $ 0.8 | 1.0000 | $ 0.53 | 1.0000 | $ 0.36 | 1.0000 | 3.0000 | 1.0000 |
| 1946 | 1.54 | 1.8118 | 1.10 | 2.0755 | 0.72 | 2.0000 | 5.8654 | 1.9624 |
| 1940 | 0.85 | .5519 | 0.5 | .4818 | 0.36 | 1.0000 | 1.5337 | 0.5112 |
| 1946 | 1.54 | .0000 | 1. 0 | 10000 | 0.72 | 2.0000 | 3.0000 | 1.000 |

時間互換測驗：

$$T=\Sigma\frac{p_t}{p_o}/N\cdot\Sigma\frac{p_o}{p_t}N=1.9624\times0.5112=1.0032$$

三、簡單中位數指數時間互換測驗計算表（第142表）

| 年份 | 小麥 | | 包穀 | | 燕麥 | | I |
|---|---|---|---|---|---|---|---|
| | $p_t$ | $\frac{p_t}{p_o}$ | $p_t$ | $\frac{p_t}{p_o}$ | $p_t$ | $\frac{p_t}{p_o}$ | |
| 1940 | $ 0.85 | 1.0000 | $ 0.53 | 1.0000 | $ 0.36 | 1.0000 | |
| 1946 | 1.54 | 1.8118 | 1.10 | 2.0755 | 0.72 | 2.0000 | 2.0000 |
| 1940 | 0.85 | 0.5519 | 0.53 | 0.4818 | 0.36 | 0.5100 | 0.5000 |
| 1946 | 1.54 | 1.0000 | 1.10 | 1.0000 | 0.72 | 1.0000 | |

時間互換測驗：

$$T=2.0000\times0.5000=1.00$$

四、簡單幾何平均數指數時間互換測驗計算表

（第143表）

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $\frac{p_1}{p_0}$ | $\log\frac{p_1}{p_0}$ | $p_1$ | $\frac{p_1}{p_0}$ | $\log\frac{p_1}{p_0}$ | $p_1$ | $\frac{p_1}{p_0}$ | $\log\frac{p_1}{p_0}$ |
| 1940 | \$ 0.85 | 00.00 | 2 0000 | \$ 0.53 | 100.00 | 2.0000 | \$ 0.36 | 100.0 | 2.0000 |
| 1946 | 1.54 | 81.18 | 2.2581 | 1.10 | 2.7.55 | 2.317[illegible] | 0.72 | 200.00 | 2.3010 |
| 1940 | 0.85 | 55.19 | 1.7419 | 0.53 | 48.18 | '.68267 | 0.36 | 50.00 | .69897 |
| 1946 | 1.54 | 100.00 | 2 0000 | 1.10 | .00.00 | 2.0000 | 0.72 | 100.00 | 2.0000 |

續 上 表

| 年份 | $\Sigma\log\frac{p_1}{p_0}$ | $\frac{\Sigma\log\frac{p_1}{p_0}}{N}$ | $\log^{-1}\frac{\Sigma\log\frac{p_1}{p_0}}{N}$ |
|---|---|---|---|
| 1940 | 6.00[illegible]0 | 2.00000 | 00.00 |
| 1942 | 6.87620 | 2.29236 | 195.95 |
| 1944 | 5.'2374 | 1.7 791 | 5 .04 |
| 19[illegible]6 | 6.0000 | 2.0 000 | 100.03 |

時間互換測驗：

$$T=0.19595\times0.5104=1.00$$

上列各種簡單指數公式中，合乎時間互換測驗者，計有綜合式，中位

式及幾何式三種，就中以簡單幾何平均數公式爲較優。蓋前已言之，簡單綜合公式中包含有不合理之權數，簡單中位數公式每以項目過少而不可靠。至於算數平均數公式犯上偏，倒數平均數公式犯下偏，兩者之偏誤俱由於平均數型態關係，總稱之爲型偏（Type Bias），故在簡單指數公式之中，以幾何平均數一公式爲最佳，但如指數編製中包括之價比甚多，則中位數又優于幾何式。

## 第二節　因子互換測驗

試驗各加權物價指數公式之工具爲因子互換測驗（The Factor Reversal Test）。分析一種商品在前後兩時期中價格之變動，前一期爲基期，後一期爲計期，基期之價格爲基價，以 $p_o$ 表示之，基期之數量爲基量，以 $q_o$ 表示之；計期之價格爲計價，以 $p_t$ 表示之，計算期之數量爲計量，以 $q_t$ 表示之，求計期價值指數（Value Index）之方法有直接與間接兩種：

1.直接法：即依基年價值 $p_oq_o$ 及計年價值 $p_tq_t$ 求之，計期之價值指數爲 $I=\frac{p_tq_t}{p_oq_o}$

2.間接法：由價格指數 $\frac{p_t}{p_o}$ 及數量指數 $\frac{q_t}{q_o}$ 以求之，即價值指數等於價格指數與數量指數之乘積，以公式表示之即 $I=\frac{p_t}{p_o}\cdot\frac{q_t}{q_o}$

在一種商品時，無論直接或間接所得之計期價值指數必彼此相等，至易證明。例如：某種商品在基年中之單位價格爲 5 元，交易數量爲 4 單位，在計年中之價格爲 10 元，數量爲 8 單位，則由直接法所求之計期價值指數爲：$I=\frac{10\times8}{5\times4}\times100=400$，亦可依間接法即由兩因子指數而得價值

指數。價格指數爲 $\frac{10}{5}=2$，數量指數爲 $\frac{8}{4}=2$，價值指數爲兩者之乘積，即 $=\frac{p_1}{p_o}\cdot\frac{q_1}{q_o}=\frac{10}{5}\times\frac{8}{4}\times100=400$。足見只有一種商品時，先分別求得價格指數及數量指數，再以二者之乘積爲價值指數，與直接求得之價值指數相同。如所研究者爲多種商品，依各種公式分別編製加權價格指數及加權數量指數，然後再求兩種指數之乘積，是爲間接由因子而求得之價值指數。同時，亦可直接求價值指數，如直接與間接兩法所得之價值指數彼此相同，則稱其公式合乎因子互換測驗；反之，如兩法所得之結果不相同，則斷定其不合乎因子互換測驗。現就各加權指數公式別試驗之：

一、基年數量加權綜合指數因子互換測驗。

以基年數量爲權數，則加權綜合價格指數公式爲 $\frac{\Sigma p_1q_o}{\Sigma p_oq_o}$，至於加權綜合數量指數公式，只將式中之兩因子 $p$ 與 $q$ 之位置互換即得，其形式爲：$\frac{\Sigma q_1p_o}{\Sigma q_op_o}$，此式中係以基期價格爲權數，固定不變，只數量 $q_1$ 一因子變動，求得加權綜合式之價格指數與數量指數後，再求二者之乘積即爲價值指數，以符號表示之，則價值指數爲：$\frac{\Sigma p_1q_o}{\Sigma p_oq_o}\cdot\frac{\Sigma q_1p_o}{\Sigma q_op_o}$，此係間接由因子而得之價值指數。此外亦可直接求價值指數，其公式爲：$\frac{\Sigma p_1q_1}{\Sigma p_oq_o}$，或書爲 $\frac{\Sigma v_1}{\Sigma v_o}$，現根據上章所用 1940 及 1946 兩年之資料，依間接與直接之次序求出價值指數，最後觀察間接與直接兩法所得是否相同。

（第 144 表）　基年數量加權綜合價格指數計算表

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | | $\Sigma p_1q_o$ | $\frac{\Sigma p_1q_o}{\Sigma p_oq_o}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $q_o$ | $p_1q_o$ | $p_1$ | $q_o$ | $p_1q_o$ | $p_1$ | $q_o$ | $p_1q_o$ | | |
| 1940 | $ 0.85 | 174 | 147.90 | $ 0.53 | 63 | 33.39 | $ 0.36 | 10 | 3.60 | 184.89 | 100.00 |
| 1946 | 1.54 | 174 | 267.96 | 1.10 | 63 | 69.30 | 0.72 | 10 | 7.20 | 344.46 | 186.31 |

（第 145 表） 基年價格加權綜合數量指數計算表

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | | $\Sigma q_1 p_0$ | $\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $q_1$ | $p_0$ | $q_1 p_0$ | $q_1$ | $p_0$ | $q_1 p_0$ | $q_1$ | $p_0$ | $q_1 p_0$ | | |
| 1940 | 174 | $ 0.85 | 147.90 | 63 | $ 0.53 | 33.39 | 10 | $ 0.36 | 3.60 | 184.89 | 100.00 |
| 1946 | 174 | 0.85 | 122.40 | 21 | 0.53 | 11.13 | 46 | 0.35 | 16.56 | 150.09 | 81.18 |

（第 146 表　綜合價值指數計算表

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | | $\Sigma p_1 q_1$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $q_1$ | $p_1 q_1$ | $q_1$ | $p_1$ | $q_1 p_1$ | $q_1$ | $p_1$ | $q_1 p_1$ | | |
| 1940 | $ 0.85 | 174 | 147.90 | $ 0.53 | 63 | 33.39 | $ 0.36 | 10 | 3.60 | 184.89 | 100.00 |
| 1946 | 1.54 | 144 | 221.78 | 1.10 | 21 | 23.10 | 0.72 | 46 | 33.12 | 277.98 | 185.21 |

（第 147 表） 因子互換測驗計算表

| 年份 | $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$ | $\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$ | $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \cdot \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$ |
|---|---|---|---|---|
| 1940 | 100.00 | 100.00 | 100.00 | 100.00 |
| 1946 | 186·31 | 81.16 | 151.25 | 185.21 |

上列因子互換測驗計算表內第一欄中係價格指數，第二欄中係數量指數，第三欄中爲價格與數量兩種指數之乘積，卽由因子而求得之價值指數，第四欄中則爲依直接法所得之價值指數，兩法所得之價值指數不同，卽加權綜合式不合於因子互換測驗。

二、費暄教授理想指數公式之測驗。

費暄教授以爲各加權指數公式無一合乎因子互換測驗者，其不合之原因或由于權數，或由於平均數之公式，或二者兼而有之。以權數而言，有基年與計年之別，基年數量犯下偏，計年數量犯上偏，俱爲權偏（Weight Bias)。加權綜合指數式及加權幾何指數式所犯者俱爲權偏，以平均數公

式兩言，簡單算術平均數指數式犯上偏，簡單倒數平均數指數式犯下偏，俱爲型偏(Type Bias)，而加權算術及加權倒數兩種指數式，則各兼犯型偏與權偏，加權綜合指數式所犯者既只爲權偏，基年數量犯下偏，計年數量犯上偏，理想指數公式即係將此權數偏誤相反之兩種加權綜合指數公式加以配合，使其上偏與下偏從中抵消；其配合之方法爲求兩種加權綜合指數之幾何平均數，即爲理想指數式。所謂理想指數式者，即其公式既合乎時間互換測驗，復合乎因子互換測驗，仍依三種糧食在1940與1946兩年份之資料證明之。

1. 理想指數式基期互換測驗

甲　基期數指加權綜合(Laspeyres)前進後退指數計算表

(第148表)

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | | $\Sigma p_1 q_0$ | $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $q_0$ | $p_1 q_0$ | $p_1$ | $q_0$ | $p_1 q_0$ | $p_1$ | $q_0$ | $p_1 q_0$ | | |
| 1940 | $ 0.85 | 174 | 147.90 | $ 0.53 | 63 | 33.39 | $ 0.36 | 10 | 3.60 | 184.89 | 100.00 |
| 1946 | 1.54 | 174 | 276.96 | 1.10 | 63 | 69.30 | 0.72 | 10 | 7.20 | 344.46 | 186.31 |
| 1940 | 0.85 | 144 | 122.40 | 0.53 | 21 | 11.13 | 0.36 | 46 | 66.56 | 150.09 | 53.99 |
| 1946 | 1.54 | 144 | 221.76 | 1.10 | 21 | 23.10 | 0.72 | 46 | 33.12 | 277.98 | 100.00 |

乙　計期數量加權綜合(Pasche)前進與後退指數計算表

(第149表)

| 年份 | 小麥 | | | | | 包穀 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $p_1 q_1$ | $p_0 q_1$ | $p_1$ | $p_0$ | $q_1$ | $p_1 q_1$ | $p_0 q_1$ |
| 1940 | $ 0.85 | $ 0.85 | 174 | 147.90 | 147.90 | $ 0.53 | $ 0.53 | 63 | 33.39 | 33.39 |
| 1946 | 1.54 | 0.85 | 174 | 221.76 | 122.40 | 1.10 | 0.53 | 21 | 23.10 | 11.13 |
| 1940 | 0.85 | 1.54 | 174 | 147.90 | 267.96 | 0.53 | 1.10 | 63 | 33.39 | 69.30 |
| 1946 | 1.54 | 1.54 | 174 | 221.76 | 221.76 | 1.10 | 1.10 | 21 | 23.10 | 23.10 |

續　上　表

| 年份 | 蒜 | | | 麥 | | $\Sigma p_1 q_1$ | $\Sigma p_0 q_1$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$ |
|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $p_0$ | $q_1$ | $p_1 q_1$ | $p_0 q_1$ | | | |
| 1940 | $0.36 | $0.36 | 10 | 3.60 | 3.60 | 184.89 | 184.8[illegible] | 100.00 |
| 1946 | 0.72 | 0.3 | 46 | 33.12 | 16.56 | 277.98 | 150.[illegible]9 | 1[illegible]5.21 |
| 1940 | 0.36 | 0.72 | 10 | 3.60 | 7.20 | 154.8[illegible] | 344.46 | 53.73 |
| 1946 | 0.72 | 0.72 | 46 | 33.12 | 33.12 | 277.[illegible]8 | 277.98 | 100.00 |

丙　理想公式前進與後退指數計算表

（第150表）

| 年份 | $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$ | $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}\cdot\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$ | $\sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}\cdot\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}$ |
|---|---|---|---|---|
| 1940 | 10[illegible].00 | 1[illegible]0.00 | 10,000.0[illegible]00 | 100.00 |
| 1946 | 136.31 | 1[illegible]5.21 | 34,506.4751 | 185.75 |
| 1940 | 53.9 | 53.[illegible]3 | 2900.8827 | 53.86 |
| 1946 | 10[illegible].00 | 1[illegible]0.00 | 10 000.0000 | 100.00 |

時間互換測驗：前進指數乘後退指數＝1.8575×0.5386

＝1.0004

上列研究之程序係以兩年中之價格及數量資料為根據，先編製基期數量加權綜合前進指數與後退指數及計期數量加權綜合前進指數與後退指數，然後將所得兩組指數配合，即求二者之幾何平均數，所得之結果即為理想公式之前進指數與後退指數，最後再求此前進與後退兩指數之乘積，結果為一，足見理想公式合乎時間互換測驗。

2. 理想指數式因子互換測驗。因子互換之要旨在各種商品之價格、數量與價值之變動俱可用指數表出，價值之變動可直接求各年之綜合價值，然後再求價值指數以表示之，其公式為：$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$，亦可根據理想公式先

求物價指數及物量指數，然後再求兩因子指數之乘積，所得結果爲價值指數；應與上述直接法所得者相同，此種關係可用指數公式符號表出：

理想式物價指數：$\sqrt{\dfrac{\Sigma p_1q_0}{\Sigma p_0q_0}\cdot\dfrac{\Sigma p_1q_1}{\Sigma p_0q_1}}$

理想式物量指數：$\sqrt{\dfrac{\Sigma q_1p_0}{\Sigma q_0p_0}\cdot\dfrac{\Sigma q_1p_1}{\Sigma q_0p_1}}$

因子互換測驗：$\sqrt{\dfrac{\Sigma p_1q_0}{\Sigma p_0q_0}\cdot\dfrac{\Sigma p_1q_1}{\Sigma p_0q_1}}\cdot\sqrt{\dfrac{\Sigma q_1p_0}{\Sigma q_0p_0}\cdot\dfrac{\Sigma q_1p_1}{\Sigma q_0p_1}}=\dfrac{\Sigma p_1q_1}{\Sigma p_0q_0}$

理想公式之價格指數業已求出，可資利用，現求其數量指數：

甲　基年價格加權綜合式數量指數計算表　（第151表）

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | | $\Sigma q_1p_0$ | $\frac{\Sigma q_1p_0}{\Sigma q_0p_0}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $q_1$ | $p_0$ | $q_1p_0$ | $q_1$ | $p_0$ | $q_1p_0$ | $q_1$ | $p_0$ | $q_1p_0$ | | |
| 1940 | 174 | 0.85 | 147.90 | 63 | 0.53 | 33.39 | 10 | 0.3 | 3.60 | 184.89 | 100.03 |
| 1946 | 174 | 0.85 | 122.40 | 21 | 0.53 | 11.13 | 46 | 0.36 | 16.56 | 159.0 | 81.13 |

乙　計年價格加權綜合式數量指數計算表　（第152表）

| 年份 | 小麥 | | | | | 包穀 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $q_1$ | $q_0$ | $p_1$ | $q_1p_1$ | $q_0p_1$ | $q_1$ | $q_0$ | $p_1$ | $q_1p_1$ | $q_0p_1$ |
| 1940 | 174 | 174 | 0.85 | 147.90 | 147.90 | 63 | 63 | 0.53 | 33.3 | 33.39 |
| 1946 | 144 | 174 | 1.54 | 221.75 | 267.96 | 21 | 3 | 1.1 | 23.1 | 19.30 |

續上表

| 年份 | 燕麥 | | | | | $\Sigma q_1p_1$ | $\Sigma q_0p_1$ | $\frac{\Sigma q_1p_1}{\Sigma q_0p_1}$ |
|---|---|---|---|---|---|---|---|---|
| | $q_1$ | $q_0$ | $p_1$ | $q_1p_1$ | $q_0p_1$ | | | |
| 1940 | 10 | 10 | 0.3 | 3.60 | 0.60 | 184.89 | 184.89 | 100.00 |
| 1946 | 46 | 10 | 0.72 | 33.12 | 7.20 | 277.98 | 344.46 | 80.70 |

丙 理想公式數指數計算表 （第153表）

| 年份 | $\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}$ | $\frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}$ | $\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \cdot \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}$ | $\sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \cdot \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$ |
|---|---|---|---|---|
| 1940 | 100.00 | 100.00 | 10,000.0000 | 100.00 |
| 1946 | 81.13 | 80.70 | 6,551.2260 | 80.94 |

丁 價值指數計算表 （第154表）

| 年份 | 小麥 | | | 包穀 | | | 燕麥 | | | $\Sigma p_1 q_1$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $p_1$ | $q_1$ | $p_1 q_1$ | $p_1$ | $q_1$ | $p_1 q_1$ | $p_1$ | $q_1$ | $p_1 q_1$ | | |
| 1940 | \$0.85 | 174 | 147.90 | 0.53 | 63 | 33.39 | 0.36 | 10 | 3.60 | 184.89 | 100.00 |
| 1946 | 1.54 | 144 | 221.76 | 1.10 | 21 | 23.10 | 0.72 | 46 | 33.12 | 277.98 | 150.35 |

戊 理想公式因子互換測驗計算表 （第155表）

公式：價格指數×數量指數＝價值指數

| 年份 | $\sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \cdot \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}}$ | $\sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \cdot \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$ | $\sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \cdot \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \cdot \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \cdot \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}$ | $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_0}$ |
|---|---|---|---|---|
| 1940 | 100.00 | 100.00 | 100.00 | 100.00 |
| 1946 | 185.76 | 80.94 | 160.35 | 150.35 |

## 第三節 循環測驗

上二節中所述之時間互換及因子互換兩種測驗，俱係用以分析二時期之指數者，循環變動(The Circular Test)則係用以分析長於二時期之指數者。前已言之，一種商品之價比有固定、連環與連鎖三種，各年份之固定基期價比，與連環價比再加以連鎖所得之價比，應彼此相同，連鎖價比實爲連環價比之還原，可稱爲循環測驗。單一商品之價比，合乎此種測驗，在價比中業已證明。在簡單物價指數公式中，幾何平均數合乎此種測驗；

在加權指數公式中，固定權數之綜合式及幾何平均數式亦合乎此種測驗，其他各簡單與加權指數公式均於此不合，就中理想公式所差者不甚大，此種循環差額 (Circular Discrepancy) 之產生，由于理想式中之權數與時俱變，而循環測驗之權數則須固定不變，是以理想公式之不合於循環測驗不足爲病。

## 第四節　各種指數公式之可靠性

理想公式既合於時間互換測驗與因子互換測驗，於循環測驗雖稍有未合處，又非即可據而斷定爲其公式之缺點，則理想公式允爲物價指數公式中之最優者。惟可批評之點有二：

1、在根據理想公式所編之指數中，不但價格變，作爲權數之數量亦變，搜集資料困難，既得全備無缺之價格與數量資料，計算手續亦頗繁重。

2、理想指數中所包括之價格與數量二因子，既俱與時俱變，吾人對於指數之變動，無法斷定係由於價格之變或由於數量之變，惟編製指數之目的在於以指數表示價格之變動者。

由於上述二種原因，費暄教授另提出一代替之指數公式，即聯合數量加權價格指數式，其形式爲 $\frac{\Sigma p_1(q_o+q_1)}{\Sigma p_o(q_o+q_1)}$，式中所需之價格，自與年俱變，與理想指數或其他指數公式中所用者相同，至所需之權數，則爲基年與各計年數量之和，計年之數量自亦與年俱變。基年之數量爲權數時犯下偏，計年之數量犯上偏，兩者聯合之結果，上下偏即可抵消 編製之手續比較簡易，所得之指數與理想式相差極微。費暄教授以爲無論就正確 迅速，簡易，或循環差額最小觀點而論，聯合數量綜合價格指數實爲最優之公式，英國學者哀基渥斯(Edgeworth)及馬歇爾 (Alfred Marshall) 亦均

綜道之，費暄教授因稱此指數式爲哀基渥斯—馬歇爾公式。但如計年數量無法得致時，則基年數量加權價格指數，實又爲最適用之公式，倘基年數量亦無法獲得，則簡單幾何平均數是爲最優良之公式，如所包括之價比數列爲數尙多，則中位數指數或尙較幾何平均數爲優。

（第156表）　美國歷年出超總值調節計算表

出超數値單位：百萬美圓

| 年份 | 出超値 | 美國勞工統計局批發物價指數 | 改正出超値 |
|---|---|---|---|
| 1926 | 378 | 100.0 | 378 |
| 27 | 681 | 95.4 | 714 |
| 28 | 1,037 | 97.7 | 1,061 |
| 29 | 842 | 95.3 | 884 |
| 30 | 782 | 86.4 | 905 |
| 31 | 334 | 73.0 | 458 |
| 32 | 288 | 64.8 | 444 |
| 33 | 225 | 65.9 | 341 |
| 34 | 478 | 74.9 | 638 |
| 35 | 235 | 80.0 | 294 |
| 36 | 33 | 86.3 | 41 |
| 37 | 265 | 78.6 | 307 |
| 38 | 1,134 | 77.1 | 1,443 |
| 39 | 859 | 78.6 | 1,114 |
| 40 | 1,396 | 87.3 | 1,776 |
| 41 | 1,802 | 98.8 | 2,061 |
| 42 | 5,335 | 98.8 | 5,400 |
| 43 | 9,583 | 103.3 | 9,304 |
| 44 | 10,337 | 104.0 | 9,939 |
| 45 | 5,667 | 105.8 | 5,256 |

註：最近數年出超猛加，係包括租借法案出口物資。

## 第五節　統計調節

凡時間數列之單位爲貨幣時，稱爲價值數列。此種價值數列之形成旣由於數量與價格兩因子，則價值數列之變動，可由於實際數量之變，亦可由至貨幣價值或物價水準之變。舉例言之，出口或出超總值與年俱增時，如物價指數未變，則總值之增必由於實際數量之增。反之，如物價上漲，則實際數量未加甚或減少，仍可以物價上漲而使總值膨脹，欲使總值之變十足代表實際數量之變，則必自總值之中設法將價格變動一因素消除，其法爲以物價指數除原有之價值數列，稱爲統計調節(Statistical Deflation)。節調以後，所得之改正數列，仍爲價值數列，但已不受物價波動之影響，而純粹代表數量之變者，試以美國歷年來出超數值，以見調節之意義及算法：

（參見上頁第156表）

# 第五編

# 時間數列之分析

# 第十八章　直綫長期趨勢

## 第一節　時間列數之性質

經濟現象如消費、生產、交易及價格，各與時俱變，社會現象如人口之出生死亡，結婚及離婚，亦均與時俱變，稱爲時間數列。驟視之，時間數列似無何規律，分析言之，則每一時間數列中，每均包括有四種成分，可個別加以研究者：

1. 長期趨勢 (Secular Trend)
2. 季節變動 (Seasonal Variation)
3. 循環變動 (Cyclical Fluctuation)
4. 意外變動 (Irregular Movement)

所謂長期趨勢，即社會或經濟現象，在數十年中一致上升或下降之變動，此種一致上升或下降，如循一直綫發展，則稱爲直綫長期趨勢或簡稱爲直綫趨勢 (Straight line Trend)，可求出一直綫以代表之，此直綫即稱爲趨勢直綫(Trend Line)。如趨勢直綫上升，即表示此一現象之數值與年俱增。反之，如趨勢直綫下降，則表示此一現象之數值與年俱減。至于社會現象長期趨勢之所以上升或下降，則以人口增加，生產技術進步，生產數量增加，生活標準提高，醫藥與衛生改進，生育率及死亡率之降低等。凡此所舉各端，其發生之作用綜和而一致，因而造成長期中一致上升或下降，卽長期趨勢。

季節變動，卽一年四季或十二月中社會經濟現象之波動起伏。如一年之中，農業品在秋季收穫上場時，其價格最低，冬季稍高無多，入春則農產品之價格上漲，至夏季因青黃不接，其價更上漲而達於極點，直至秋收時

再下跌。經濟活動如此，人口之出生與死亡，在一年之中亦與月俱變。季節變動之所以發生，可以氣候之轉變，產業之旺淡，商業習慣及社會風俗等。以產業旺淡而論，如春耕季節農業工具之需求加大，秋收季節農產品之供給加多，其他工業品之產銷亦有類似情形。以商業習慣而論，如工商業中逢年過節之結賬收款，公私債款之歸還，亦有一定期限。以社會風俗而論，如年節以及其他例假日期。至於氣候轉變所造成季節現象尤爲顯著。總之，其原因爲一季一月之所特有，而爲其他季節之所無者。

循環變動指社會經濟現象在每數年中之上下起伏。因其變動周而復始，富有節奏性，故稱之爲循環變動，或稱產業循環(Business Cycle)。全部循環作用，可分四階段：

1. 恐慌 (Panic)　　3. 復興 (Recovery)

2. 蕭條 (Depression)　　4. 繁榮 (Prosperity)

經濟情況初尚平靜繁榮，突然發生恐慌。短期恐慌之後，繼之以較長期之蕭條。蕭條之中漸孕復興之象。終而重入繁榮之境。此種週期性之上下波動，關係於工商業及整個社會機構者至大。經濟及統計學家對於產業循環問題細心研究成爲著述者亦至多。

意外變動，每由臨時爆發事項所造成，如地震，戰爭，罷工，水火災害，均可使經濟機構突然脫節。綜合時間數列之成分中，季節、循環及意外三種變動，俱爲短期變動，與之對待者則爲長期趨勢。每一時間數列雖均包括上述四項因子，但以資料不同，所受各因子影響之輕重亦不同，故兩時間數列，於未加分析其成分之前，不能逕加比較以得結論。

## 第二節　決定趨勢直綫之方法

分析時間數列必自作圖入手，將各年各月資料繪成曲綫圖後，時間數

列中四因子之各點每一一呈現，以長期發展而論，其形式爲直綫或曲綫，

(第157表) 美國波斯頓城各月結婚數目表

(本表資料見 Hexter, M. B., Social Consequence of Busin-ss Cy-les, P. f6,)

| 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1900一 | 523 | 1902一 | 482 | 1904一 | 527 | .906一 | 553 | 1908一 | 5 8 | 1910一 | 539 |
| 二 | 485 | 二 | 44 | 二 | 476 | 二 | 554 | 二 | 506 | 二 | 4 2 |
| 三 | 276 | 三 | 334 | 三 | 275 | 三 | 335 | 三 | 42 | 三 | 577 |
| 四 | 538 | 四 | 592 | 四 | 622 | 四 | 648 | 四 | 522 | 四 | 573 |
| 五 | 531 | 五 | 370 | 五 | 403 | 五 | 412 | 五 | 433 | 五 | 3 7 |
| 六 | 711 | 六 | 802 | 六 | 960 | 六 | 952 | 六 | 894 | 六 | 95 |
| 七 | 431 | 七 | 422 | 七 | 469 | 七 | 392 | 七 | 299 | 七 | 498 |
| 八 | 426 | 八 | 414 | 八 | 471 | 八 | 573 | 八 | 479 | 八 | 487 |
| 九 | 493 | 九 | 581 | 九 | 601 | 九 | 739 | 九 | 7 2 | 九 | 633 |
| 十 | 731 | 十 | 657 | 十 | 788 | 十 | 830 | 十 | 773 | 十 | C67 |
| 十一 | 688 | 十一 | 649 | 十一 | 728 | 十一 | 790 | 十一 | 740 | 十一 | 649 |
| 十二 | 416 | 十二 | 428 | 十二 | 416 | 十二 | 496 | 十二 | 45 | 十二 | 395 |
| 總計 | 6 249 | 總計 | 6,172 | 總計 | 6,736 | 總計 | 7,374 | 總計 | 6,9 3 | 總計 | 6,617 |
| 平均 | 520.75 | 平均 | 5'4.35 | 平均 | 591.33 | 平均 | 614.50 | 平均 | 582.75 | 平均 | 551.42 |
| 1901一 | 501 | 1903一 | 500 | 1905一 | 480 | 1907一 | 631 | 1909一 | 534 | 1911一 | 640 |
| 二 | 468 | 二 | 490 | 二 | 462 | 二 | 5·7 | 二 | 531 | 二 | 628 |
| 三 | 277 | 三 | 2" | 三 | 357 | 三 | 384 | 三 | 353 | 三 | 393 |
| 四 | 585 | 四 | 636 | 四 | 497 | 四 | 666 | 四 | 583 | 四 | 622 |
| 五 | 382 | 五 | 385 | 五 | 421 | 五 | 483 | 五 | 448 | 五 | 44) |
| 六 | 80 | 六 | 88 | 六 | 906 | 六 | 1,052 | 六 | 958 | 六 | 1,080 |
| 七 | 427 | 七 | 436 | 七 | 508 | 七 | 594 | 七 | 477 | 七 | 596 |
| 八 | 447 | 八 | 500 | 八 | 471 | 八 | 559 | 八 | 54 | 八 | 585 |
| 九 | 549 | 九 | 647 | 九 | 659 | 九 | 745 | 九 | 692 | 九 | 781 |
| 十 | 721 | 十 | 712 | 十 | 766 | 十 | 842 | 十 | 736 | 十 | 892 |
| 十一 | 716 | 十一 | 7 8 | 十一 | 734 | 十一 | 808 | 十一 | 779 | 十一 | 856 |
| 十二 | 4'0 | 十二 | 456 | 十二 | 474 | 十二 | 439 | 十二 | 506 | 十二 | 598 |
| 總計 | 6,310 | 總計 | 6, 4 | 總計 | 6 775 | 總計 | 7,711 | 總計 | 7,176 | 總計 | 8,111 |
| 平均 | 525.83 | 平均 | 553.33 | 平均 | 564.53 | 平均 | 642.59 | 平均 | 598 | 平均 | 675.95 |

## 續上表

| 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 | 年月 | 結婚數 |
|---|---|---|---|---|---|---|---|---|---|
| 1912一 | 5[illegible]0 | 1914一 | 6 4 | 1916一 | 610 | 19 8一 | 5 8 | 1920一 | 622 |
| 二 | 545 | 二 | 677 | 二 | 644 | 二 | 495 | 二 | 576 |
| 三 | 315 | 三 | [illegible]52 | 三 | 471 | 三 | 41[illegible] | 三 | 352 |
| 四 | 658 | 四 | 702 | 四 | 614 | 四 | 678 | 四 | 782 |
| 五 | 467 | 五 | 611 | 五 | 541 | 五 | 57[illegible] | 五 | 520 |
| 六 | 931 | 六 | 1,168 | 六 | 1,174 | 六 | 1.059 | 六 | 1,329 |
| 七 | 52 | 七 | 570 | 七 | 676 | 七 | 599 | 七 | 592 |
| 八 | 59[illegible] | 八 | 840 | 八 | 562 | 八 | 7[illegible]2 | 八 | 609 |
| 九 | 092 | 九 | 756 | 九 | 836 | 九 | 720 | 九 | 556 |
| 十 | 975 | 十 | 888 | 十 | 929 | 十 | 568 | 十 | 941 |
| 十一 | 830 | 十一 | 849 | 十一 | 870 | 十一 | 515 | 十一 | 816 |
| 十二 | 681 | 十二 | 557 | 十二 | 634 | 十二 | 426 | 十二 | 550 |
| 總計 | 7 885 | 總計 | 8,72 | 總計 | 8,551 | 總計 | 7,359 | 總計 | 8 54[illegible] |
| 平均 | 657.08 | 平均 | 727 | 平均 | 712.58 | 平均 | 6 6.52 | 平均 | 7 2.08 |
| 19'3一 | 5 6 | 915一 | 664 | 19 7一 | 640 | 1919一 | 476 | | |
| 二 | 525 | 二 | 581 | 二 | 620 | 二 | 509 | | |
| 三 | 479 | 三 | 357 | 三 | 398 | 三 | 4[illegible]2 | | |
| 四 | 638 | 四 | 632 | 四 | 894 | 四 | 585 | | |
| 五 | 537 | 五 | 626 | 五 | 695 | 五 | 514 | | |
| 六 | 1,110 | 六 | 1,078 | 六 | 1,300 | 六 | 1,16[illegible] | | |
| 七 | 575 | 七 | 525 | 七 | 76[illegible] | 七 | 571 | | |
| 八 | 637 | 八 | 637 | 八 | 684 | 八 | 691 | | |
| 九 | 78[illegible] | 九 | 758 | 九 | 744 | 九 | 743 | | |
| 十 | 854 | 十 | [illegible]2 | 十 | 869 | 十 | 922 | | |
| 十一 | 836 | 十一 | 7[illegible] | 十一 | 881 | 十一 | 883 | | |
| 十二 | 503 | 十二 | 509 | 十二 | 694 | 十二 | 603 | | |
| 總計 | 8,120 | 總計 | 8,131 | 總計 | 9,191 | 總計 | 8,122 | | |
| 平均 | 676.67 | 平均 | 677.58 | 平均 | 765.92 | 平均 | 676.83 | | |

其方向爲上升或下降，俱可由圖上現出。現以美國波斯頓城各月份結婚數目資料爲例，以昇直綫趨勢分析之方法：

（第40圖）　波斯頓城結婚數目算術圖

由上列第40圖中曲綫觀之，可見1900至1920共計二十一年間，波斯頓城結婚數目雖在每數年之中有一循環變動，但自全期觀之，仍有一致上升之趨勢。此種趨勢可以一直綫代表之。決定趨勢直綫之方法，約有下列三種：

一、半平均數法：

所謂半平均數法（Method of Semi-averages）即將時間數列分爲前後兩半，各求一簡單平均數，以此兩平均數標入圖中，定出兩點，並經此兩點作一直綫，即爲趨勢直綫。各年各月中直綫軌跡所定之值，即爲趨勢數值（Trend Ordinates），係代表各年各月長期發展者。亦稱爲標準值（Normal Values）現有資料二十一年，由前十年中求一平均數，後十一年中求出一平均數如下：

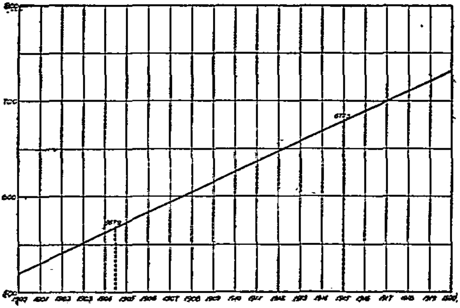
(第158表) 前十年之平均數

| 年份 | 各月平均結婚數 |
|---|---|
| 1900 | 521 |
| 1901 | 526 |
| 1902 | 514 |
| 1903 | 553 |
| 1904 | 561 |
| 1905 | 565 |
| 1906 | 615 |
| 1907 | 643 |
| 1908 | 583 |
| 1909 | 598 |
| 總計 | 5,679 |
| 平均數 | 567.9 |

(第159表) 後十一年之平均數

| 年份 | 各月平均結婚數 |
|---|---|
| 1910 | 551 |
| 1911 | 676 |
| 1912 | 657 |
| 1913 | 677 |
| 1914 | 727 |
| 1915 | 678 |
| 1916 | 713 |
| 1917 | 766 |
| 1918 | 6 6 |
| 1919 | 677 |
| 1920 | 7.2 |
| 總計 | 7.450 |
| 平均數 | 677.3 |

(第41圖) 波斯頓城結婚數目趨勢直綫圖



依上列第158,159兩表計算結果，前十年之平均數爲567.9，所計算之年份既爲1900至1909十年，則所得平均數係代表第五年終或第六年初，即1904年底或1905年初之趨勢值。後十年之平均數爲677.3，所計算之年份既爲1910至1920十一年，則所得平均數係代表居中央一年即1915年七月一日之趨勢值。現依此兩日期之趨勢值在圖中作一直綫，由直綫之軌跡即可確定各年之趨勢數值。如將趨勢直綫繪入原有資料曲綫圖中，由趨勢直綫與原值曲綫對照，則一方可見長期中或理論上應有之發

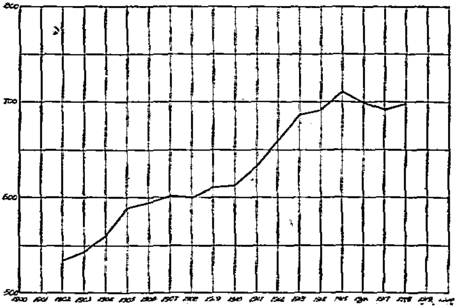
(第160表)　波斯頓城每年各月平均結婚數目五年移動平均數計算表

| 年份 | 每年均數 | 移動總數 | 移動平均數 |
|---|---|---|---|
| 1900 | 521 | | |
| 1901 | 526 | | |
| 19 2 | 514 | 2,675 | 535 |
| 1903 | 553 | 2,719 | 544 |
| 1904 | 561 | 2,808 | 562 |
| 1905 | 565 | 2,9 7 | 587 |
| 1906 | 615 | 2,9 7 | 593 |
| 1907 | 643 | 3,004 | 601 |
| 19 8 | 563 | 2,990 | 591 |
| 1909 | 598 | 3,051 | 610 |
| 1910 | 551 | 3,065 | 613 |
| 1911 | 676 | 3.159 | 632 |
| 19 2 | 657 | 3,238 | 658 |
| 1913 | 677 | 3,4 5 | 683 |
| 1914 | 727 | 3,452 | 690 |
| 1915 | 678 | 3,561 | 712 |
| 1916 | 713 | 3,509 | 700 |
| 1917 | 766 | 3,450 | 630 |
| 1918 | 616 | 3,484 | 697 |
| 1919 | 677 | | |
| 1920 | 712 | | |

展，一方可見每年實際發展與長期趨勢出入之情況。

二、移動平均數法：

在原始資料曲綫圖中所以有波浪起伏，係由於循環變動存在。如能概約判斷每一循環所佔之時期為五年或七年，則求五年或七年移動平均數之結果，可將循環變動所發生之作用抵消，其移動平均數卽為代表長期發展之趨勢數值。在上列第160表中，係依波斯頓城結婚數目求五年移動平數均藉資說明：




（第42圖） 波斯頓城結婚數目趨勢曲綫圖

移動平均數法已較半平均數法為優。蓋在半平均數法中，趨勢直線之決定只根據兩點。而在移動平均數法則趨勢綫係由各個年份計算，自較為合理。但此法亦非無缺點：第一，移動平均之結果，雖可將各年中由循環作用所造成之波瀾起伏消除，但所得仍未能為一直線。第二，首尾兩端之年份，無法求移動平均數。如用三年移動制，則所得之平均數首尾各缺一年；

五年移動制，則首尾各缺二年；如用七年移動制，則首尾各缺三年。究竟用三年、五年或七年移動制，亦殊無一定準則。所可言者，只爲求得三年、五年或七年之移動平均數時，可作圖以斷定究以何者爲較優，最足以代表長期趨勢一因素，而不再受其他因素所紛擾。

三、最小二乘方法：

甲 普通法

求趨勢直線最重要者爲最小二乘方法。依最小二乘方法求回應直線，在相關論中業已詳爲闡述。回應線係表示有自變數與因變數各一時，因變數隨自變數變化之情形。現在時間數列中，時間爲自變數，以 $X$ 代表之，而社會經濟現象則爲因變數，以 $Y$ 代表之，故可用同法求直線以表示社會經濟現象隨時間變動之情形。此直線在相關論中稱爲回應線，在時間數列中則稱爲趨勢直線。趨勢直線與回應線相同，係代表最合於機率之數值。各原值與趨勢值間離差之和爲零，離差平方之和爲一最小值，如此方成爲最足以代表各年中實際發展之直線。茲依同一資料以計算之，有普通及簡捷兩法，現分別依例說明之。依最小二乘方法之直線趨勢總方程式爲：

$$Y=a+bX$$

解此方程式所需之兩聯立標準方程式爲：

$$(1)\quad \Sigma Y=Na+b\Sigma X$$

$$(2)\quad \Sigma XY=a\Sigma X+b\Sigma X^2$$

解此兩聯立方程式，則須根據下列計算表，由表中可知：

$N=21\quad \Sigma X=210,\ \Sigma Y=13129$

$\Sigma XY=139119,\ \Sigma X^2=2870$

以上列各總數代入，可得下列兩聯立方程式：

(第161表) 波斯頓城結婚數目趨勢直線計算表

| 年份 | $X$ | $Y$ | $XY$ | $X^2$ | $Y_c=523.5155+10.1675X$ |
|---|---|---|---|---|---|
| 1900 | 0 | 521 | 0 | 0 | 523.5155 |
| 1901 | 1 | 526 | 526 | 1 | 533.6830 |
| 1902 | 2 | 514 | 1028 | 4 | 543.8505 |
| 1903 | 3 | 553 | 1659 | 9 | 554.0180 |
| 1904 | 4 | 561 | 2244 | 16 | 564.1855 |
| 1905 | 5 | 565 | 2825 | 25 | 574.3530 |
| 1906 | 6 | 615 | 3690 | 36 | 584.5205 |
| 1907 | 7 | 643 | 4501 | 49 | 594.6880 |
| 1908 | 8 | 583 | 4664 | 64 | 604.8555 |
| 1909 | 9 | 598 | 5382 | 81 | 615.0230 |
| 1910 | 10 | 551 | 5510 | 100 | 625.1905 |
| 1911 | 11 | 576 | 7436 | 121 | 635.3580 |
| 1912 | 12 | 657 | 7884 | 144 | 645.5255 |
| 1913 | 13 | 677 | 8801 | 169 | 655.6930 |
| 1914 | 14 | 727 | 10178 | 196 | 665.8605 |
| 1915 | 15 | 678 | 10170 | 225 | 676.0280 |
| 1916 | 16 | 713 | 11408 | 256 | 686.1955 |
| 1917 | 17 | 766 | 13022 | 289 | 696.3630 |
| 1918 | 18 | 616 | 11088 | 324 | 706.5305 |
| 1919 | 19 | 677 | 12863 | 361 | 716.6980 |
| 1920 | 20 | 712 | 14240 | 400 | 726.8655 |
| | 210 | 13,129 | 139,119 | 2,870 | |

(1) $13129=21a+210b$

(2) $139119=210a+2870b$

10(1), $131290=210a+2100b$ ………(3)

(2)−(3): $7829=\qquad 770b$

$$\therefore\quad b=\frac{7829}{770}=10.1675$$

以$b$之值代入(1):

$$13129=21a+210\times10.1675$$

$$13129=21a+2135.175$$

$$21a=13129-2135.175=10993.825$$

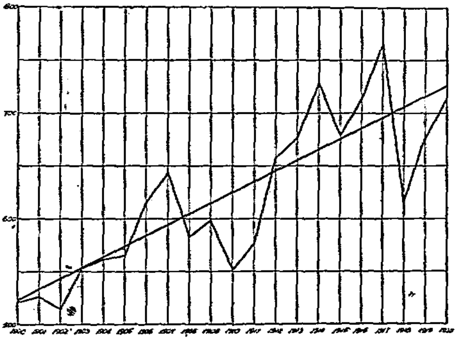
$$\therefore \quad a=\frac{10993.825}{21}=523.5155$$

直綫趨勢方程式:

$$Y_c=523.5155+10.1675X$$

$X$:代表一年: 原點:1900年七月一日

依上列直綫方程式,自可求得各年之趨勢值。記入上列計算表中最右一欄中,用$Y_c$表之,以示與原值$Y_o$有別。



(第43圖) 波斯頓城結婚數目趨勢值與實際值比較圖

乙簡捷法：

（第162表）　波斯頓城結婚數目趨勢直綫計算表

| 年份 | $X$ | $Y$ | $XY$ | $X^2$ | $Y_c=625.1905+10.1675X$ |
|---|---|---|---|---|---|
| 1900 | −10 | 521 | −5210 | 100 | 523.5155 |
| 1901 | −9 | 526 | −4734 | 81 | 533.6830 |
| 1902 | −8 | 514 | −4112 | 54 | 543.8505 |
| 1903 | −7 | 553 | −3871 | 49 | 554.0180 |
| 1904 | −6 | 561 | −3366 | 36 | 564 1855 |
| 1905 | −5 | 56[illegible] | −282 | 25 | 574 3530 |
| 1906 | −4 | 61[illegible] | −246 | 16 | 58[illegible].5205 |
| 1907 | −3 | 613 | −1 29 | 9 | 594 6880 |
| 1908 | −2 | [illegible]83 | −1166 | 4 | 604.8555 |
| 1909 | −1 | 598 | −598 | 1 | 615.0230 |
| 1910 | 0 | 5 1 | 0 | 0 | 625.1905 |
| 1911 | 1 | 676 | 676 | 1 | 635.3580 |
| 1912 | 2 | 657 | 1314 | 4 | 645.5255 |
| 1913 | 3 | 677 | 2 3 | 9 | 655.6930 |
| 1914 | 4 | 727 | 29 8 | 16 | 665.8605 |
| 1915 | 5 | 678 | 3390 | 25 | 676.0280 |
| 1916 | 6 | 713 | 4278 | 36 | 586.1955 |
| 1917 | 7 | 766 | 5362 | 49 | 696.3630 |
| 1918 | 8 | 6 6 | 4958 | 6 | 706.5305 |
| 1919 | 9 | 677 | 6093 | 8 | 716.6980 |
| =20 | 10 | 712 | 7120 | 10 | 726.8555 |
| | 0 | 13120 | 38100<br>−30271<br>7829 | 770 | |

$N=21$，　$\Sigma X=0$，　$\Sigma Y=13129$，　$\Sigma XY=7829$，　$\Sigma X^2=770$

依普通法求趨勢直綫之總方程式爲：$Y=a+bX$，兩聯立標準方程式爲 (1) $\Sigma Y=Na+b\Sigma X$ 及 (2) $\Sigma XY=a\Sigma X+b\Sigma X^2$，現以 $\Sigma X=0$，故簡捷法中之兩聯立標準方程式如下：

(1)　$\Sigma Y=Na$

(2) $\Sigma XY = b\Sigma X^2$

以上列各總數代入：

(1) $13129=21a$, $\therefore$ $a=\frac{13129}{21}=625.1905$

(2) $7829=770b$, $\therefore$ $b=\frac{7829}{660}=10.1675$

直綫趨勢總方程式：

$$Y_c=625.1905+10.1675X$$

$X$ 代表一年原點：1910 年七月一日。在普通法中所得之結果爲：

$a=523.5155$, $b=10.1675$

總方程式爲：$Y_c=523.5155+10.1675X$

$b$ 係每年增加數，(Annual Increment)。無論依普通法或依簡捷法所得 $b$ 之數值相同，至兩法所得 $a$ 之值所以不同，以原點不同。在普通法中原點爲 1900 年七月一日。在簡捷法中之原點則爲 1910 年七月一日。所謂原點即 $X$ 等於零所在之點。$X$ 既爲零，則 $Y_c=a$，即 1900 年七月一日之趨勢值爲 523.5155，1910 年七月一日之趨勢值爲 625.1905。普通法與簡捷法所得之 $b$ 既相同，$a$ 又可化爲相同，則依兩直綫方程式所得之趨勢值必完全相同。

上列兩方程式，俱爲年份趨勢方程式(Annual Trend Equation)，可化爲半年份趨勢方程式(Semi-annual Trend Equation)，及月份趨勢方程式 (Monthly Trend Equation)。在年份方程式中，每年增加數 $b$ 既爲 10.1675，則半年增加數爲：$\frac{b}{2}=\frac{10.1675}{2}=5.08375$，每月增加數爲 $b_m=\frac{b}{12}=\frac{10.0675}{12}=0.8473$，半月增加數爲 $\frac{b_m}{2}=\frac{0.8473}{2}=0.4237$。上列簡捷法年份趨勢方程式 $Y=625.1905+10.1675X$ 中，1910 年七月一日之趨勢值爲 625.1905，半月增加數既爲 0.4237，則1910 年七月十五

之趨勢值自為 $Y_c=625.1905+0.4237=625.6142$ 以 $x$ 代表一月，以1910年七月十五日為原點，則可得一月份趨勢方程式如下：

$$Y_c=625.6142+0.8237x.$$

依此月份方程式可求各月份趨勢值 (Monthly Trend Ordinates) 如下：

(第163表) 美國波斯頓城結婚數目月份趨勢值計算表

月份趨勢方程式：$Y_c$ 625.614+0.824$x$　　原點：1910年7月15日

| 年 月 | $x$ | $Y_o$ | $Y_c$ | 年 月 | $x$ | $Y_o$ | $Y_c$ |
|---|---|---|---|---|---|---|---|
| 19 0一 | −126 | 523 | 521.790 | 1902一 | −102 | 482 | 541.566 |
| 二 | −125 | 485 | 522.614 | 二 | −10 | 441 | 542.390 |
| 三 | −124 | 276 | 523.438 | 三 | −100 | 334 | 543.214 |
| 四 | −123 | 538 | 524,262 | 四 | − 99 | 592 | 544.033 |
| 五 | −122 | 531 | 525,086 | 五 | − 98 | 370 | 544.862 |
| 六 | −121 | 711 | 525.9 0 | 六 | − 97 | 802 | 545,686 |
| 七 | −120 | 431 | 526.734 | 七 | − 9 | 422 | 546.510 |
| 八 | −119 | 426 | 527.558 | 八 | − 95 | 414 | 547.334 |
| 九 | −118 | 493 | 528 382 | 九 | − 94 | 81 | 548.158 |
| 十 | −117 | 731 | 529 206 | 十 | − 93 | 657 | 549.982 |
| 十一 | −116 | 688 | 530.030 | 十一 | − 92 | 619 | 549.806 |
| 十二 | −115 | 416 | 530 854 | 十二 | − 91 | 428 | 550.630 |
| 1901一 | −11 | 501 | 531.678 | 1903一 | −90 | 500 | 551.454 |
| 二 | −113 | 468 | 532. 02 | 二 | −89 | 490 | 552 278 |
| 三 | −112 | 277 | 533.326 | 三 | −88 | 227 | 553.102 |
| 四 | −111 | 585 | 534 1.0 | 四 | −87 | 638 | 553.926 |
| 五 | −110 | 382 | 534.974 | 五 | −86 | 385 | 554.750 |
| 六 | −109 | 807 | 535.798 | 六 | −85 | 881 | 555.574 |
| 七 | −108 | 427 | 536.622 | 七 | −84 | 436 | 556.398 |
| 八 | −107 | 447 | 537.446 | 八 | −83 | 500 | 557.222 |
| 九 | −106 | 5 9 | 538 270 | 九 | −82 | 647 | 559.046 |
| 十 | −105 | 721 | 539.094 | 十 | −81 | 7.2 | 559.876 |
| 十一 | −104 | 716 | 539.918 | 十一 | − 80 | 718 | 559.694 |
| 十二 | −103 | 430 | 540.742 | 十二 | −79 | 465 | 560.518 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1904一 | -78 | 527 | 561.342 | 1907一 | -42 | 631 | 591.006 |
| 二 | -77 | 476 | 562.166 | 二 | -41 | 517 | 591.830 |
| 三 | -76 | 275 | 562.990 | 三 | -40 | 384 | 592.654 |
| 四 | -75 | 622 | 563.814 | 四 | -39 | 668 | 593.478 |
| 五 | -74 | 403 | 564,638 | 五 | -38 | 483 | 594.302 |
| 六 | -7 | 960 | 565.462 | 六 | -37 | 1,052 | 595.126 |
| 七 | -72 | 469 | 566 286 | 七 | -36 | 594 | 595.950 |
| 八 | -71 | 471 | 567.110 | 八 | -35 | 554 | 596.774 |
| 九 | -70 | 601 | 567.934 | 九 | -34 | 745 | 597.528 |
| 十 | -69 | 768 | 568.758 | 十 | -33 | 842 | 598.425 |
| 十一 | -68 | 723 | 569.582 | 十一 | -32 | 808 | 599.246 |
| 十二 | -67 | 416 | 570.406 | 十二 | -3, | 433 | 600.070 |
| 1905一 | -66 | 480 | 571.230 | 1908一 | -30 | 568 | 600.894 |
| 二 | -65 | 462 | 572.054 | 二 | -29 | 506 | 601.7,8 |
| 三 | -64 | 387 | 572.878 | 三 | -28 | 421 | 602.542 |
| 四 | -63 | 497 | 573.702 | 四 | -27 | 322 | 603.366 |
| 五 | -62 | 421 | 574.526 | 五 | -26 | 433 | 604.190 |
| 六 | -61 | 916 | 575.350 | 六 | -2, | 694 | 605.014 |
| 七 | -60 | 508 | 576.174 | 七 | -24 | 499 | 605.838 |
| 八 | -59 | 471 | 576.998 | 八 | -23 | 479 | 606.662 |
| 九 | -58 | 659 | 577.822 | 九 | -22 | 702 | 607.486 |
| 十 | -57 | 766 | 578.646 | 十 | -21 | 773 | 108.310 |
| 十一 | -56 | 734 | 579.470 | 十一 | -20 | 740 | 609.134 |
| 十二 | -55 | 474 | 580.294 | 十二 | -19 | 654 | 609.958 |
| 1906一 | -54 | 553 | 581.118 | 1903一 | -18 | 534 | 610.782 |
| 二 | -53 | 554 | 58 .942 | 二 | -17 | 531 | 611.606 |
| 三 | -52 | 235 | 582.766 | 三 | -16 | 353 | 612.430 |
| 四 | -51 | 648 | 583.590 | 四 | -15 | 533 | 613.254 |
| 五 | -50 | 412 | 584.414 | 五 | -14 | 448 | 614 078 |
| 六 | -49 | 952 | 585.238 | 六 | -13 | 953 | 614.902 |
| 七 | -48 | 492 | 586.062 | 七 | -12 | 477 | 615.726 |
| 八 | -47 | 573 | 586.866 | 八 | -11 | 546 | 616.550 |
| 九 | -46 | 739 | 587.710 | 九 | -10 | 692 | 617.374 |
| 十 | -45 | 831 | 588.534 | 十 | - 9 | 769 | 618.198 |
| 十一 | -44 | 790 | 589.359 | 十一 | - 8 | 779 | 6,9.072 |
| 十二 | -43 | 496 | 590.182 | 十二 | - 7 | 518 | 619.846 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1910— | −6 | 5 9 | 620.670 | 1913— | 30 | 596 | 650.334 |
| 二 | −5 | 492 | 6 2.4 4 | 二 | 31 | 525 | 65 .153 |
| 三 | −4 | 377 | 622. 18 | 三 | 32 | 479 | 651.982 |
| 四 | −3 | 578 | 623.142 | 四 | 33 | 688 | 652. 05 |
| 五 | −2 | 97 | 6.3.966 | 五 | 34 | 537 | 653.630 |
| 六 | −1 | 953 | 624.793 | 六 | 35 | 1,110 | 654.454 |
| 七 | 0 | 498 | 625.6.4 | 七 | 36 | 5'5 | 655 273 |
| 八 | 1 | 487 | 626.438 | 八 | 37 | 637 | 656.102 |
| 九 | 2 | 683 | 627.262 | 九 | 38 | 780 | 656.926 |
| 十 | 3 | 667 | 628.086 | 十 | 39 | 854 | 657.753 |
| 十一 | 4 | 649 | 628.910 | 十一 | 40 | 836 | 65 .574 |
| 十二 | 5 | 595 | 629.734 | 十二 | 41 | 50[illegible] | 653.393 |
| 1911— | 6 | 640 | 630.558 | 1914— | 42 | 634 | 66 .222 |
| 二 | 7 | 628 | 631..9 | 二 | 43 | 677 | 6 1.046 |
| 三 | 8 | 393 | 632.[illegible]03 | 三 | 44 | 452 | 6 1.870 |
| 四 | 9 | 422 | 633.026 | 四 | 45 | 702 | 662.694 |
| 五 | 10 | 440 | 633.851 | 五 | 46 | 611 | 663.518 |
| 六 | 11 | 1,050 | 6 4.678 | 六 | 47 | 1,188 | 664.342 |
| 七 | 12 | 596 | 6 5.502 | 七 | 48 | 570 | 665 166 |
| 八 | 13 | 585 | 636.326 | 八 | 49 | 840 | 665.990 |
| 九 | 14 | 781 | 637.150 | 九 | 50 | 756 | 666.814 |
| 十 | 15 | 529 | 637.974 | 十 | 51 | 888 | 667.638 |
| 十一 | 16 | 856 | 636.7.8 | 十一 | 52 | 849 | 668.462 |
| 十二 | 17 | 598 | 639.622 | 十二 | 53 | 557 | 669.286 |
| 1912— | 18 | 530 | 640.446 | 1915— | 54 | 664 | 670.110 |
| 二 | 19 | 545 | 641.270 | 二 | 55 | 53 | 670.934 |
| 三 | 20 | 3 5 | 642.036 | 三 | 56 | 357 | 671.758 |
| 四 | 21 | 658 | 642.918 | 四 | 57 | 682 | 672.5.2 |
| 五 | 22 | 467 | 643.742 | 五 | 58 | 1 626 | 673.'06 |
| 六 | 23 | 981 | 644.566 | 六 | 59 | ,0'8 | 674.230 |
| 七 | 24 | 524 | 645.330 | 七 | 60 | 525 | 675.05 |
| 八 | 25 | 530 | 646.214 | 八 | 61 | 37 | 675.878 |
| 九 | 26 | 792 | 647.038 | 九 | 62 | 753 | 676.702 |
| 十 | 27 | 972 | 647.862 | 十 | 63 | 924 | 677.526 |
| 十一 | 28 | 870 | 648.686 | 十一 | 64 | 7.9 | 678 350 |
| 十二 | 29 | 681 | 649.510 | 十二 | 65 | 509 | 679.174 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1916一 | 66 | 610 | 679.938 | 1919一 | 102 | 476 | 709.662 |
| 二 | 67 | 644 | 680.822 | 二 | 103 | 509 | 710.4[illegible]6 |
| 三 | 68 | 471 | 681.6 6 | 三 | 104 | 432 | 711.310 |
| 四 | 69 | 614 | 682.470 | 四 | 105 | 5[illegible]5 | 712.134 |
| 五 | 70 | 54[illegible] | 683.294 | 五 | 106 | 544 | 712.958 |
| 六 | 71 | 1,147 | 684.118 | 六 | 107 | 1163 | 713.782 |
| 七 | 72 | 6[illegible]6 | 684.942 | 七 | 108 | 571 | 714.606 |
| 八 | 73 | 562 | 685.766 | 八 | 109 | 691 | 715.430 |
| 九 | 74 | 836 | 686.590 | 九 | 110 | 743 | 716.254 |
| 十 | 75 | 929 | 687.414 | 十 | 111 | 922 | 717.078 |
| 十一 | 76 | 670 | 688.238 | 十一 | 112 | 883 | 717.902 |
| 十二 | 77 | 624 | 689.062 | 十二 | 113 | 6[illegible]3 | 718.726 |
| 1917一 | 78 | 6[illegible]0 | 689.886 | 1920一 | 114 | 622 | 719.550 |
| 二 | 79 | 620 | 690.710 | 二 | 115 | 576 | 720.374 |
| 三 | 80 | 393 | 691.534 | 三 | 116 | 352 | 7[illegible]1.198 |
| 四 | 81 | [illegible]94 | 692.358 | 四 | 117 | 782 | 722.022 |
| 五 | 82 | 695 | 693.182 | 五 | 118 | 520 | 722.846 |
| 六 | 83 | 1300 | 694.0[illegible]6 | 六 | 119 | 1329 | 7[illegible]3.670 |
| 七 | 84 | [illegible]69 | 694.830 | 七 | 1[illegible]0 | 592 | 724.494 |
| 八 | 85 | 684 | 6[illegible]5.654 | 八 | 121 | 609 | 725.318 |
| 九 | 86 | 744 | 696.[illegible]78 | 九 | 122 | 856 | 726.142 |
| 十 | 87 | 869 | 697.3 2 | 十 | 123 | 941 | 726.[illegible]66 |
| 十一 | 8 | 884 | 698.126 | 十一 | 124 | 816 | 727.7[illegible]6 |
| 十二 | 89 | 694 | 698.950 | 十二 | 125 | 550 | 728.614 |
| 1918一 | 90 | 598 | 699.774 | | | | |
| 二 | 91 | 495 | 700.598 | | | | |
| 三 | 92 | 415 | 701.422 | | | | |
| 四 | 93 | 678 | 7[illegible]2.246 | | | | |
| 五 | 94 | 570 | 703.070 | | | | |
| 六 | 95 | 1059 | 703.894 | | | | |
| 七 | 96 | 599 | 704.718 | | | | |
| 八 | 97 | 712 | 705.542 | | | | |
| 九 | 98 | 730 | 706.366 | | | | |
| 十 | 93 | 568 | 707.190 | | | | |
| 十一 | 10[illegible] | 645 | 708.014 | | | | |
| 十二 | 101 | 426 | 708.838 | | | | |

# 習題

(一) 美國汽油及火油生產數量表 (第164表)

單位：百萬桶每桶42加侖

| 年份 | 汽油 | 火油 | 年份 | 汽油 | 火油 |
|---|---|---|---|---|---|
| 1916 | 41.0 | 2.5 | 1927 | 330.4 | 30.1 |
| 17 | 67.9 | 5.2 | 28 | 376.9 | 43.2 |
| 18 | 85.0 | 6.7 | 29 | 435.1 | 53.2 |
| 19 | 94.2 | 8.4 | 30 | 432.2 | 52.6 |
| 20 | 116.3 | 9.2 | 31 | 431.5 | 43.6 |
| 21 | 112.7 | 10.7 | 32 | 3 2.6 | 36.3 |
| 22 | 147.7 | 12.0 | 33 | 61.6 | 33.8 |
| 23 | 171.9 | 19.4 | 34 | 416.9 | 36.6 |
| 24 | 213.3 | 22.2 | 35 | 457.8 | 30.3 |
| 25 | 250.6 | 26.8 | 36 | 504.7 | 42.0 |
| 26 | 299.7 | 32.5 | | | |

註：上表資料轉錄自 Croxton's Workbook in Applied General Statistics, P. 15

一、試將上表中之汽油產量與火油產量分別繪一算術標尺曲線圖。

二、試依上表兩種資料分別計算直線長期趨勢值，並以所得之趨勢值作入(一)題之圖中。

(二) 美國生鉄各年各月每日生產數量表(單位：千噸)(第165表)

| 年月 | 產量 | 年月 | 產量 | 年月 | 產量 | 年月 | 產量 |
|---|---|---|---|---|---|---|---|
| 1919—1 | 106.9 | 1920—1 | 97.3 | 1921—1 | 78.0 | 1922—1 | 3.15 |
| 2 | 105.0 | 2 | 102.7 | 2 | 79.2 | 2 | 58.2 |
| 3 | 99.7 | 3 | 189.9 | 3 | 51.5 | 3 | 65.7 |
| 4 | 82.6 | 4 | 91.3 | 4 | 39.9 | 4 | 69.1 |
| 5 | 68.0 | 5 | 96.3 | 5 | 39.4 | 5 | 74.4 |
| 6 | 70.5 | 6 | 101.5 | 6 | 35.5 | 6 | 78.7 |
| 7 | 78.3 | 7 | 98.9 | 7 | 27.9 | 7 | 77.6 |
| 8 | 88.5 | 8 | 101.5 | 8 | 30.8 | 8 | 58.6 |
| 9 | 82.9 | 9 | 104.3 | 9 | 32.9 | 9 | 67.8 |
| 10 | 80.1 | 10 | 106.2 | 10 | 40.2 | 10 | 85.1 |
| 11 | 79.8 | 11 | 97.8 | 11 | 47.2 | 11 | 99.0 |
| 全年均數12 | 84.9 | 全年均數12 | 67.2 | 全年均數12 | 59.2 | 全年均數12 | 99.6 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1923—1 | 104.2 | 1926—1 | 107.0 | 1929—1 | 111.0 | 1932—1 | 31.4 |
| 2 | 104.9 | 2 | 104.4 | 2 | 114.5 | 2 | 33.3 |
| 3 | 113.7 | 3 | 111.0 | 3 | 119.8 | 3 | 31.2 |
| 4 | 118.3 | 4 | 116.0 | 4 | 122.1 | 4 | 28.4 |
| 5 | 124.8 | 5 | 112.3 | 5 | 125.7 | 5 | 25.3 |
| 6 | 122.6 | 6 | 107.8 | 6 | 123.9 | 6 | 23.9 |
| 7 | 116.7 | 7 | 104.0 | 7 | 122.1 | 7 | 18.5 |
| 8 | 111.3 | 8 | 103.2 | 8 | 121.2 | 8 | 17.1 |
| 9 | 104.2 | 9 | 104.5 | 9 | 116.6 | 9 | 19.8 |
| 10 | 101.6 | 10 | 107.6 | 10 | 115.7 | 10 | 20.8 |
| 11 | 96.5 | 11 | 107.9 | 11 | 106.1 | 11 | 21.0 |
| 全年均數12 | 94.2 | 全年均數12 | 98.7 | 全年均數12 | 91.5 | 全年均數12 | 17.9 |
| 1924—1 | 97.4 | 1927—1 | 100.1 | 1930—1 | 91.2 | 1933—1 | 18.4 |
| 2 | 106.0 | 2 | 105.0 | 2 | 101.4 | 2 | 19.8 |
| 3 | 111.6 | 3 | 112.4 | 3 | 104.7 | 3 | 17.5 |
| 4 | 107.8 | 4 | 114.1 | 4 | 106.1 | 4 | 20.8 |
| 5 | 84.4 | 5 | 109.4 | 5 | 104.3 | 5 | 23.6 |
| 6 | 67.5 | 6 | 103.0 | 6 | 97.8 | 6 | 42.2 |
| 7 | 47.6 | 7 | 95.2 | 7 | 85.2 | 7 | 57.8 |
| 8 | 60.9 | 8 | 95.1 | 8 | 81.4 | 8 | 59.1 |
| 9 | 68.4 | 9 | 92.5 | 9 | 75.8 | 9 | 50.7 |
| 10 | 79.9 | 10 | 89.8 | 10 | 69.8 | 10 | 43.8 |
| 11 | 83.7 | 11 | 88.3 | 11 | 62.2 | 11 | 36.2 |
| 全年均數21 | 95.5 | 全年均數12 | 87.0 | 全年均數12 | 53.7 | 全年均數12 | 38.1 |
| 1925—1 | 168.7 | 1928—1 | 9[illegible].6 | 1931—1 | 55.3 | 1934—1 | 39.2 |
| 2 | 114.8 | 2 | 100.0 | 2 | 61.0 | 2 | 5.1 |
| 3 | 115.0 | 3 | 103.2 | 3 | 65.6 | 3 | 52.2 |
| 4 | 108.5 | 4 | 106.2 | 4 | 67.3 | 4 | 57.6 |
| 5 | 94.5 | 5 | 105.9 | 5 | 64.3 | 5 | 65.9 |
| 6 | 89.1 | 6 | 102.7 | 6 | 54.6 | 6 | 64.3 |
| 7 | 85.9 | 7 | 99.1 | 7 | 47.2 | 7 | 39.5 |
| 8 | 87.2 | 8 | 101.2 | 8 | 41.3 | 8 | 34.0 |
| 9 | 9[illegible].9 | 9 | 102.1 | 9 | 39.0 | 9 | 29.9 |
| 10 | 97.5 | 10 | 108.6 | 10 | 37.9 | 10 | 30.7 |
| 11 | 100.8 | 11 | 110.1 | 11 | 36.5 | 11 | 31.9 |
| 全年均數12 | 104.9 | 全年均數12 | 108.7 | 全年均數12 | 31.6 | 全年均數12 | 33.2 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 935—1 | 47.7 | 1931—1 | 65.4 | 1937—1 | 103.6 | 1938—1 | 46.1 |
| 2 | 57.5 | 2 | 62.9 | 2 | 107.1 | 2 | 46.4 |
| 3 | 27.1 | 3 | 65.8 | 3 | 111.6 | 3 | 46.9 |
| 4 | 55.5 | 4 | 80.1 | 4 | 113.1 | | |
| 5 | 55.7 | 5 | 85.4 | 5 | 114.1 | | |
| 6 | 51.8 | 6 | 86.2 | 6 | 103.6 | | |
| 7 | 49.6 | 7 | 83.7 | 7 | 112.9 | | |
| 8 | 56.8 | 8 | 87.5 | 8 | 116.3 | | |
| 9 | 59.2 | 9 | 91.0 | 9 | 113.7 | | |
| 10 | 63.8 | 10 | 96.5 | 10 | 93.3 | | |
| 11 | 68.9 | 11 | 98.3 | 11 | 66.6 | | |
| 全年均數12 | 63.0 | 全年均數12 | 100.5 | 全年均數12 | 48.1 | | |

註：上表所載資料係轉錄自 Crum & Others, Introduction to Economic Statistics, p.p. 354 & 336.

(一) 試根據上表所載一九一九至一九三七年各月份美國生鐵產量求各年平均數

(二) 試依(一)題所得各年平均產量，求直線長期趨勢總式並求各年之趨勢值。

(三) 試以各年之實際平均產量及趨勢值繪一算術曲線圖。

(四) 試求月份長期趨勢總式。

(三) 美國靴，鞋及拖鞋各月總產量表(單位：百萬雙.(第166表)

| 年月 | 產量 | 年月 | 產量 | 年月 | 產量 | 年月 | 產量 |
|---|---|---|---|---|---|---|---|
| 1922—1 | 25.12 | 1923—1 | 30.74 | 1924—1 | 26.50 | 1925—1 | 26.56 |
| 2 | 24.55 | 2 | 31.30 | 2 | 26.83 | 2 | 26.45 |
| 3 | 29.35 | 3 | 35.84 | 3 | 28.86 | 3 | 29.89 |
| 4 | 26.65 | 4 | 31.87 | 4 | 28.00 | 4 | 29.41 |
| 5 | 26.23 | 5 | 30.93 | 5 | 25.24 | 5 | 25.11 |
| 6 | 24.83 | 6 | 28.97 | 6 | 22.26 | 6 | 23.45 |
| 7 | 22.69 | 7 | 25.27 | 7 | 21.39 | 7 | 24.76 |
| 8 | 27.68 | 8 | 30.03 | 8 | 25.47 | 8 | 28.49 |
| 9 | 28.[illegible]9 | 9 | 27.55 | 9 | 27.72 | 9 | 29.77 |
| 10 | 30.37 | 10 | 30.70 | 10 | 29.83 | 10 | 31.06 |
| 11 | 30.08 | 11 | 26.[illegible]5 | 11 | 55.32 | 11 | 24.63 |
| 12 | 27.85 | 12 | 22.68 | 12 | 24.60 | 12 | 24.49 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1926—1 | 23.87 | 1929—1 | 27.25 | 1932—1 | 21.23 | 19 5—1 | 29.56 |
| 2 | 25.70 | 2 | 27.71 | 2 | 25.66 | 2 | 30.[illegible]7 |
| 3 | 29.93 | 3 | 30.90 | 3 | 30.68 | 3 | 34.23 |
| 4 | 26.64 | 4 | 29.38 | 4 | 25.95 | 4 | 34.36 |
| 5 | 23.13 | 5 | 29.16 | 5 | 22.50 | 5 | 31.26 |
| 6 | 25.04 | 6 | 28.12 | 6 | 23.50 | 6 | 27.23 |
| 7 | 55.05 | 7 | 33.22 | 7 | 20.44 | 7 | 32.27 |
| 8 | 29.65 | 8 | 36.44 | 8 | 30.78 | 8 | 37.24 |
| 9 | 31.67 | 9 | 34.83 | 9 | 33.78 | 9 | 33.91 |
| 10 | 31.66 | 10 | 37.19 | 10 | 33.07 | 10 | 35.95 |
| 11 | 26.76 | 11 | 27.72 | 11 | 25 15 | 11 | 27.71 |
| 12 | 25.42 | 12 | 22.48 | 12 | 20.10 | 1 | 8.95 |
| 1927—1 | 24.99 | 1930—1 | 26.83 | 1933—1 | [illegible]2.74 | 1936—1 | 33.36 |
| 2 | 27.21 | 2 | 25.90 | 2 | [illegible]6.38 | 2 | 33.05 |
| 3 | 31.28 | 3 | 28.63 | 3 | 28.38 | 3 | 36.63 |
| 4 | 28.39 | 4 | 29.00 | 4 | 27.63 | 4 | 33.40 |
| 5 | 25.63 | 5 | 24.51 | 5 | 32.97 | 5 | [illegible]0.26 |
| 6 | 27.50 | 6 | 23.90 | 6 | 34.[illegible]6 | 6 | 29.[illegible]7 |
| 7 | 27.78 | 7 | 24.12 | 7 | 33.75 | 7 | [illegible]5.69 |
| 8 | 35.06 | 8 | 28.43 | 8 | 37.0[illegible] | 8 | 40.67 |
| 9 | 33.93 | 9 | 29.33 | 9 | 31 [illegible]3 | 9 | 40.97 |
| 10 | 32.27 | 10 | 27.73 | 10 | 31.46 | 10 | 39.92 |
| 11 | 25.97 | 11 | 1[illegible].54 | 11 | 2[illegible].69 | 11 | [illegible]0.24 |
| 12 | 23.52 | 12 | 17.54 | 12 | [illegible]0 01 | 12 | 33.[illegible]8 |
| 1928—1 | 26.21 | 1931—1 | 19.89 | 1934—1 | [illegible]6.04 | 19 7—1 | 37.15 |
| 2 | 29.63 | 2 | 23 97 | 2 | 30.5 | 2 | 39.58 |
| 3 | 32 30 | 3 | [illegible]9.36 | 3 | [illegible]5.55 | 3 | 46.12 |
| 4 | 26.63 | 4 | 29.89 | 4 | 34.4 | 4 | 40.30 |
| 5 | 26 43 | 5 | 28.45 | 5 | 24.06 | 5 | [illegible]5.41 |
| 6 | 27.28 | 6 | 27.84 | 6 | 28.54 | 6 | [illegible]4.45 |
| 7 | 28.15 | 7 | 28.61 | 7 | 28.3 | 7 | 24.86 |
| 8 | 34.97 | 8 | 33.4 | 8 | 35.62 | 8 | 38.66 |
| 9 | 31.00 | 9 | 31.29 | 9 | 26.18 | 9 | [illegible]4.08 |
| 10 | 33.3[illegible] | 10 | 25.3 | 10 | 28.71 | 10 | 29.09 |
| 11 | 26.44 | 11 | 18.5 | 11 | 23.65 | 11 | 21.29 |
| 12 | 21.91 | 12 | 19 59 | 12 | 23.20 | 12 | 21.05 |

（一）試根據上表所載一九二二至一九三七年各月美國鞋類生產數量求各年平均數。

（二）試依上題所得之各年平均數繪一算術標尺曲線圖。

（三）試依上題曲線圖形斷定其長期趨勢受何定律支配，並求其總公式及各年之趨勢值。

# 第十九章　曲綫長期趨勢

## 第一節　二次拋物式趨勢曲綫

凡時間數列在各年各月中之增加或減少大致循等差級數定律者，俱用直綫以代表其長期發展，依直綫方程式以求各年各月之趨勢值。蓋變數旣按等差級數發展，每年每月所增加或減少之數量相同，作成算術曲綫圖時，其資料係循直綫分佈，代表此長期趨勢自以直綫爲最宜。但如資料並非按等差定律發展，作成算術圖時非循一直綫乃循一曲綫，自不能再根據直綫方程式以求各年之趨勢值。在算術圖中，由資料所構成之曲綫如呈下凹形式(Concave Downward)係表示時間演進，而資料之數量始而增加，繼而數量雖增但增加之比例漸減，終而數量可能不增加甚且減少。在分析兩項相關時，曾舉消費效用漸減及生產報酬漸減兩定律爲例，並用二次拋物方程式以求兩項間之相關量數。在時間數列中，如遇時間變數之發展亦循此定律者，自亦用二次拋物方程式以求一趨勢曲綫(Trend Curve)。現以美國石油生產數量爲例，說明應用最小二乘方法求二次拋物趨勢曲綫之程序：

二次拋物趨勢曲綫總方程式：

$$Y_c=a+bX+cX^2$$

三聯立標準方程式：

$$(1)\quad \Sigma Y=Na+b\Sigma X+c\Sigma X^2$$

$$(2)\quad \Sigma XY=a\Sigma X+b\Sigma X^2+c\Sigma X^3$$

$$(3)\quad \Sigma X^2Y=a\Sigma X^2+b\Sigma X^3+c\Sigma X^4$$

（第167表）　美國石油生產數量二次拋物趨勢曲綫計算表

單位：百萬桶　　每桶：4'加侖

| 年份 | $Y$ | $X$ | $XY(+)$ | $XY(-)$ | $X^2$ | $X^2Y$ | $X^3$ | $X^4$ | $Y_c$ |
|---|---|---|---|---|---|---|---|---|---|
| 1916 | 3 1 | −10 | | 3010 | 190 | 301 0 | −1000 | 10000 | 240.518 |
| 1917 | 335 | −9 | | 3.15 | 81 | 27135 | −7.9 | 6561 | 309.919 |
| 1918 | 356 | −8 | | 2848 | 64 | 22784 | −512 | 4036 | 375.044 |
| 1919 | 37 | −7 | | 2646 | 49 | 18522 | −3:3 | 2001 | 436.893 |
| 1920 | 443 | −6 | | 2558 | 36 | 15948 | −216 | 1296 | 495.466 |
| 1921 | 472 | −5 | | 2650 | 25 | 11800 | −1.5 | 6.5 | 550.763 |
| 1922 | 558 | −4 | | 2232 | 16 | 8928 | −64 | 256 | 602.784 |
| 1923 | 732 | −3 | | 2196 | 9 | 6588 | −27 | 81 | 651.529 |
| 1924 | 714 | −2 | | 1428 | 4 | 28 6 | −8 | 16 | 696.998 |
| 1925 | 764 | −1 | | 764 | 1 | 764 | −1 | 1 | 739.101 |
| 1926 | 771 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 778.108 |
| 1927 | 901 | 1 | 901 | | 1 | 901 | 1 | 1 | 811.749 |
| 19 8 | 902 | 2 | 1804 | | 4 | 3608 | 8 | 16 | 846.114 |
| 19.9 | 1007 | 3 | 3 21 | | 9 | 9063 | 27 | 81 | 875.203 |
| 1930 | 898 | 4 | 3592 | | 16 | 14368 | 64 | 256 | 901.016 |
| 19 1 | 851 | 5 | 4255 | | 25 | 21 75 | 125 | 625 | 923.553 |
| 19 2 | 785 | 6 | 4710 | | 36 | 28260 | 216 | 1293 | 942.814 |
| 1933 | 9 6 | 7 | 634 | | 49 | 44393 | 343 | 2401 | 958.793 |
| 1934 | 938 | 8 | 7264 | | 64 | 58112 | 512 | 4096 | 971.508 |
| 1935 | 997 | 9 | 8973 | | 81 | 80757 | 729 | 6561 | 980.941 |
| 1936 | 1100 | 10 | 11000 | | 100 | 11000 | 1000 | 10000 | 987.098 |
| 總計 | 15079 | 0 | 51862 | 23 57 | 770 | 516163 | 0 | 50666 | |

由上表已得：$\Sigma Y=15079$，$N=21$，$\Sigma X=0$

$$\Sigma XY=51862-23157=28705$$

$$\Sigma X^2=770，\quad \Sigma X^2Y=516163.$$

$$\Sigma X^3=0，\quad \Sigma X^4=50666.$$

以各總數代入上列三聯立方程式：

$$(1)\quad 15079=21a+770c$$

$$(2)\quad 28705=770b$$

$$(3)\quad 516163=770a+50666c$$

由(2)：$\quad b=\dfrac{28705}{770}=37.279$

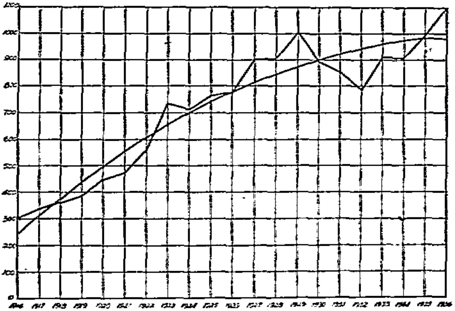
110(1)：$\quad 1,658,690=2310a+\ 84,700c\cdots\cdots\cdots(4)$

3(3)：$\quad 1,548,489=2310a+151,998c\cdots\cdots\cdots(5)$

(4)—(5)：$\quad 110,201=\qquad -67,298c$

$$\therefore\quad C=-110,201/67,298=-1.638$$

代入(1)：$\quad 15079=21a+770\times(-1.638)$



(第44圖)　美國石油生產量實際值及趨勢值比較圖

單位：百萬桶

$$21a=15079+1261.26=16,340.26$$

$$\therefore a=16340.26/21=778.108$$

二次拋物趨勢曲綫總方程式:

$$Y_c=778.108+37.279X-1.638X^2$$

依上式可求得各年趨勢值列入上列計算表中最右一欄內，並以各年之實際值 $Y_o$ 及趨勢値 $Y_c$ 一同作入上頁第44圖中，以見二次拋物曲綫配合之情形。

## 第二節　冪數式趨勢曲綫

時間數列在算術圖中如循直綫定律發展，則求直綫以代表其長期趨勢。如循下凹式曲綫發展，則求二次拋物曲綫以代表其長期趨勢。但如在算術圖中循曲綫發展其曲綫爲上凹形式(Concave Upward)，作成半對數圖時則改成循直綫分佈，是表示在長期中按等比或幾何級數發展，即每年或每月變更之比率相同，須依冪數曲綫式 (Exponential Curve) 以求其長期趨勢。其方程式之形式爲:

$$Y_c=ab^x$$

等號前後各取對數，則方程式化爲:

$$\log Y_c=\log a+X\log b,$$

式中除 $a$, $b$, $Y_c$ 三者均由眞數改爲對數外，其一般形式仍爲直綫方程式，可按上章所用之方法舉出兩聯立標準方程式，以求兩常數 $\log a$ 與 $\log b$，及兩者之逆對數，即 $a$ 與 $b$:

$$(1)\quad \Sigma\log Y=N\log a+\log b\Sigma X$$

$$(2)\quad \Sigma X\log Y=\log a\Sigma X+\log b\Sigma X^2$$

茲以美國人口變數爲例，以說明冪數式趨勢曲綫計算之程序:

(第168表) 美國人口冪數趨勢曲綫計算表

單位:百萬人

| 年份 | $X$ | $Y$ | $\log Y$ | $X\log Y(+)$ | $X\log Y(-)$ | $X^2$ |
|---|---|---|---|---|---|---|
| 1790 | −15 | 3.93 | 0.5944 | | 8.9160 | 225 |
| 1800 | −13 | 5.31 | 0.7251 | | 9.4263 | 169 |
| 1810 | −11 | 7.24 | 0.8587 | | 9.4567 | 121 |
| 1820 | −9 | 9.64 | 0.9841 | | 8.8569 | 81 |
| 1830 | −7 | 12.87 | 1.1096 | | 7.7672 | 49 |
| 1840 | −5 | 17.07 | 1.2322 | | 6.1610 | 25 |
| 1850 | −4 | 23.19 | 1.3563 | | 4.0959 | 9 |
| 1860 | −1 | 31.44 | 1.4375 | | 1.4975 | 1 |
| | 0 | | | | | |
| 1870 | 1 | 38.56 | 1.5861 | 1.5861 | | 1 |
| 1880 | 3 | 50.16 | 1.7004 | 5.3012 | | 9 |
| 1890 | 5 | 62.95 | 1.7990 | 8.9950 | | 25 |
| 1900 | 7 | 75.99 | 1.8858 | 13.1656 | | 49 |
| 1910 | 9 | 91.97 | 1.9636 | 17.6724 | | 81 |
| 1920 | 11 | 100.71 | 2.0241 | 22.2657 | | 121 |
| 1930 | 13 | 122.78 | 2.0791 | 27.1553 | | 169 |
| 1940 | 15 | 131.70 | 2.1196 | 31.7940 | | 225 |
| | 0 | 793.51 | 23.5306 | 127.7377 | 56.1775 | 1,360 |

註: 上表所載美國人口資料係轉錄自一九四七年世界年鑑第一八五頁,書中對最近六年中每年七月一日全美人口估計如下:(第169表)

| 年份(七月一日) | 人口(百萬) |
|---|---|
| 1941 | 133.20 |
| 1942 | 134.66 |
| 1943 | 136.50 |
| 1944 | 138.08 |
| 1945 | 139.62 |
| 1946 | 140.39 |

由上表計算中已知：

$N=16$, $\Sigma Y=790.51$; $\Sigma \log Y=23.5306$; $\Sigma X=0$

$\Sigma X\log Y=127.7377-56.1775=71.5602$, $\Sigma X^2=1,360$

以上列各總數代入，可得二聯立標準式，並得 $\log a$ 及 $\log b$ 如下：

(1) $23.5306=16\log a$

$\therefore \log a=23.5306/16=1.47066$

(2) $71.5602=1360\log b$

$\therefore \log b=71.5602/1360=0.052618.$

$a$ 之對數既爲 1.47066，則 $a$ 爲 1.47066 之逆對數。現係普通對數，10 之對數爲 1，100 之對數爲 2，$a$ 之對數爲 1.47066，則 $a$ 之值必介乎 10 與 100 之間 欲知 $a$ 之值只須查對數中小數 0.47066 之逆對數。至于對數中之整數 1，則表示 $a$ 之值介乎 10 與 100 之間。依對數表求 $a$ 之方法如下：

| 眞數 | 對數 | 對數 |
|---|---|---|
| 29.56 | .470704 | |
| $a$ | | .470660 |
| 29.55 | .470557 | .470557 |
| .01 | .000147 | .000103 |

$.01 : x = .000147 : .000103$

$$x=\frac{.000103\times .01}{.000147}=.007$$

$\therefore a=29.557$

以上係表示 $a$ 之值必介乎 29.55至29.56 之間 對數相差爲 .000147 時，眞數相差 0.01，現對數相差 0.000103 時，眞數差額按比例法求得爲 0.007，可知 $a$ 之值爲 29.557。

同理 $b$ 之對數爲0.052618，$b$ 爲.052618之逆對數，1之對數爲0，10之對數爲1，$b$ 之對數 既爲0.052618，足見 $b$ 之值介乎1與10之間，實爲何值，亦可按對數表依上法求出：

| 眞 數 | 對 數 | 對 數 |
|---|---|---|
| 1.129 | .052694 | |
| $b$ | | .052618 |
| 1.128 | .052309 | .052309 |
| .001 | .000385 | .000308 |

$$.001:x=.000385:.000308$$

$$x=\frac{.000308\times.001}{.000385}=.0008$$

$$\therefore\ b=1.1288.$$

$$\log a=1.47066 \qquad a=29.557$$

$$\log b=0.052618, \qquad b=1.1288$$

以之代入總方程式，則可得下列二種形式：

1. 冪數趨勢方程式：

$$Y_c=29.557(1.1288)^X$$

2. 半對數趨勢方程式：

$$\log Y_c=1.47066+0.052618X$$

在兩式中 $X$ 所代表者其爲五年，原點爲1865年底或1866年初。趨勢方程式既有兩種形式，求趨勢值之方法自亦有兩種：一即直接由冪數公式求之；二爲間接由半對數公式求之，初得 $\log Y_c$，即趨勢值之對數，次查其逆對數即爲趨勢值。如時間較短，即 $N$ 爲數不大，則由直接法或間接法求趨勢值均可。但如時間長，即 $N$ 爲數甚大，則由直接法求趨勢值計算工作繁重，以依間接法即半對數公式求之爲尚。茲分別計算如下：

## 1. 美國人口趨勢值計算表

（第1[illegible]0表）

趨勢方程式 $Y_c = ab^x$

$Y_c = 29.557(1.1288)^x$

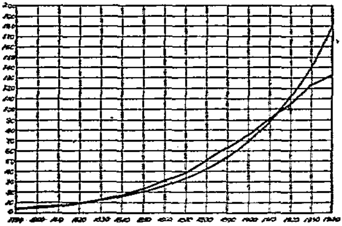
| 年份 | $X$ | $1.1288^x$ | $Y_c = 29.557(1.1288)^x$ |
|---|---|---|---|
| 1790 | −15 | $\frac{1}{1.1288^{15}} = .1625$ | 4.8030 |
| 1800 | −13 | $\frac{1}{1.1288^{13}} = .2070$ | 6.1182 |
| 1810 | −11 | $\frac{1}{1.1288^{11}} = .2638$ | 7.7971 |
| 1820 | −9 | $\frac{1}{1.1288^{9}} = .3360$ | 9.93[illegible]1 |
| 1830 | −7 | $\frac{1}{1.1288^{7}} = .4282$ | 12.6563 |
| 1840 | −5 | $\frac{1}{1.1288^{5}} = .5456$ | 16.1263 |
| 1850 | −3 | $\frac{1}{1.1288^{3}} = .6953$ | 20.5510 |
| 1860 | −1 | $\frac{1}{1.1288^{1}} = .8860$ | 26.1875 |
| 1870 | 1 | $1.1288^1 = 1.1288$ | 33.3639 |
| 1880 | 3 | $1.1288^3 = 1.4382$ | 42.5118 |
| 1890 | 5 | $1.1288^5 = 1.8327$ | 54.1691 |
| 1900 | 7 | $1.1288^7 = 2.3352$ | 69.0215 |
| 1910 | 9 | $1.1288^9 = 2.9755$ | 87.9468 |
| 1920 | 11 | $1.1288^{11} = 3.7914$ | 112.0624 |
| 1930 | 13 | $1.1288^{13} = 4.83[illegible]0$ | 142.7898 |
| 1940 | 15 | $1.1288^{15} = 6.1556$ | 181.94[illegible]1 |

## 2. 美國人口趨勢值計算表 (第171表)

趨勢方程式：$\log Y=\log a+X\log b$

$$\log Y=1.47006+.052618Y$$

| 年份 | $X$ | $.05\ 62X$ | $\log Y$ | $Y_c=\log^{-1}\log Y$ |
|---|---|---|---|---|
| 1790 | −15 | −.78930 | 0.68138 | 4.80 |
| 1800 | −13 | −.68406 | 0.78660 | 6.12 |
| 1810 | −11 | −.57882 | 0.89164 | 7.79 |
| 1820 | −9 | −.47358 | 0.99708 | 9.93 |
| 1830 | −7 | −.36834 | 1.10232 | 12.66 |
| 1840 | −5 | −.26310 | 1.20756 | 16.13 |
| 1850 | −3 | −.15786 | 1.31280 | 20.55 |
| 1860 | −1 | −.05262 | 1.41602 | 26.19 |
| 1870 | 1 | 0.05262 | 1.52328 | 33.36 |
| 1880 | 3 | 0.15786 | 1.62[illegible]32 | 42.51 |
| 1890 | 5 | 0.26310 | 1.73376 | 57.17 |
| 1900 | 7 | 0.36834 | 1.83900 | 69.03 |
| 1910 | 9 | 0.47357 | 1.94424 | 87.93 |
| 1920 | 11 | 0.57882 | 2.04943 | 112.07 |
| 1930 | 13 | 0.68406 | 2.15472 | 142.60 |
| 1940 | 15 | 0.78930 | 2.25996 | 181.95 |



第45圖 美國人口實際值與趨勢值比較圖

本節分析之要點，即時間數列原始資料在算術圖中呈示上凹形狀或在半對數圖中呈示直綫時，則依冪數方程式以求其長期趨勢。此項冪數曲綫亦稱爲複利曲綫(Compound Interest Curve)，蓋複利曲綫之公式爲：

$$Y=P(1+r)^x$$

前後各取對數，

$$logY=logP+Xlog(1+r),$$

上文中之冪數公式化爲對數形式時爲：

$$logY=loga+Xlogb$$

兩對數式相比，$logb$ 即 $log(1+r)$，已知冪數式爲：

$$logY=1.47066+.052618X,$$

化爲複利式時即：$logP=1.47066$，$log\ (1+r)=.052618$，變爲真數時 $P=29.557$，$b$ 或 $(1+r)=1.1288$，$r=1.1288-1=0.1288$。即增加率爲 12.88%，$X$ 代表五年，意即美國人口按長期趨勢一因素每五年中增加率(Rate of Increase)爲 12.88%，每五年增加數爲 1.1288，則每十年增加之數自爲 $(1.1288)^2=1.2742$。以 $B$ 或 $(1+R)$ 代表之，即：

$$B=(1+R)=1.2742$$

$$R=1.2742-1=0.2742=27.42\%$$

意即美國人口按長期趨勢一因素每十年之增加率爲27.42，證之以各年之趨勢值，完全符合，舉凡社會經濟資料，在半對數圖中循直綫分佈者，不拘其直綫爲上升或下降，俱可用冪數式或複利公式以求長期趨勢，並可求出其變動之速率。各時間數列之長期趨勢原不能彼此比較，但求得各數列之變動率 r，則可相比，如人口之增加率，生產數量之增加率，國民所得之增加率等，均可彼此比較，由比較而對各種經濟現象發展之過程及今後之動向，均可藉以明瞭。足見複利曲綫式在經濟統計分析中有莫大之功用。

## 第三節 其他趨勢曲綫

凡在算術圖中循直綫定律發展之現象，悉依直綫方程式以求其長期趨勢。在算術圖中，循下凹曲綫發展者，依二次拋物式以求其長期趨勢。在算術圖中循上凹曲綫發展即在半對數圖中則循直綫者，則用羃數式或複利式以求其長期趨勢。此外，在算術圖中呈上凹綫改爲半對數圖時仍呈上凹姿態，僅上凹之程度較爲緩和，係表示凡種現象非但其絕對數値與年俱增，即其增加之速率亦與年俱增，複利式中代表變更速度之 r 俱固定不變，自不適合，須採用半對數二次拋物曲綫公式，其總方程式及各聯立標準方程式如下：

$$logY = loga \quad Xlogb + X^2logc$$

$$(1) \quad \Sigma logY = Nloga + logb\Sigma X + logc\Sigma X^2$$

$$(2 \quad \Sigma XlogY = loga\Sigma X + logb\Sigma X^2 + logc\Sigma X^3$$

$$(3)\Sigma X^2logY = log\_\Sigma X^2 + logb\Sigma X^3 + logc\Sigma X^4$$

此趨勢總式及各聯立標準式中與過去所述者不同之點，在于過去在二次拋物綫中所用者爲原數，現在所用者則爲對數，至于計算手續除利用對數外，一切與前無殊，所得之趨勢値作入半對數圖中，仍爲曲綫。

此外，尚有一曲綫公式，亦于此稍加說明。即羅吉斯曲綫（Logistic Curve），此曲綫公式之形式爲：

$$\frac{1}{Y} = a + bc^x$$

此式實爲上述羃數式之變相，假定以 $Y'$ 代替 $\frac{1}{Y}$，$Y' = bc^x$ 實爲羃數式。而 $a$ 爲外加 常數 依 $Y' = a + bc^x$ 所得之値各求其倒數，即爲所求之趨勢値。$Y' = \frac{1}{Y} = a + lc^x$，則 $Y_c = \frac{1}{a+bc^x}$。除加一常數及化爲倒數外，

其他與線數式無異。

## 習　題

### (一) 美國石油每年生產數量表

(第172表)

單位：百萬桶，每桶42加侖

| 年份 | 產　量 | 年份 | 產　量 | 年份 | 產　量 |
|---|---|---|---|---|---|
| 1900 | 63.6 | 1913 | 248.4 | 1926 | 770.9 |
| 1901 | 66.4 | 1914 | 265.8 | 1927 | 901.1 |
| 1902 | 88.8 | 1915 | 281.1 | 1928 | 901.5 |
| 1903 | 100.5 | 1916 | 300.8 | 1929 | 1,007.3 |
| 1904 | 117.1 | 1917 | 335.3 | 1930 | 898.0 |
| 1905 | 134.7 | 1918 | 355.9 | 1931 | 851.1 |
| 1906 | 126.5 | 1919 | 378.4 | 1932 | 785.2 |
| 1907 | 166.1 | 1920 | 442.9 | 1933 | 905.7 |
| 1908 | 178.5 | 1921 | 472.2 | 1934 | 908.1 |
| 1909 | 183.2 | 1922 | 557.5 | 1935 | 996.6 |
| 1910 | 209.6 | 1923 | 732.4 | 1936 | 1,099.7 |
| 1911 | 220.4 | 1924 | 713.9 | 1937 | 1,277.7 |
| 1912 | 222.9 | 1925 | 763.7 | 1938 | |

註：上表資料轉錄自克拉好著程廣統計引論，第 353 頁。

(一) 試根據上表所載美國石油生產數量，求二次拋物曲綫長期趨勢綫方程式並計算各年之趨勢值。

(二) 試繪一美國石油生產數量圖，分別以曲綫代表實際產量及趨勢值。

（二）美國哈特福特電力公司（Hartford Elec. Light Co.）
各月平均就地產銷電力數量表　（第173表）

單位：百萬瓩小時（Million Kilowatt Hours）

| 年份 | 產銷量 | 年份 | 產銷量 | 年份 | 產銷量 |
|---|---|---|---|---|---|
| 1913 | 3.35 | 1923 | 8.20 | 1933 | 16.48 |
| 1[illegible]14 | 3.47 | 1924 | 8.92 | 1934 | 17.71 |
| 1915 | 4.28 | 1925 | 10.59 | 1935 | 19.65 |
| 1916 | 5.19 | 1926 | 11.40 | 1936 | 2[illegible].31 |
| 1917 | 5.[illegible]9 | 1927 | 12.10 | 1937 | 23.10 |
| 1918 | 6.60 | 1928 | 14.00 | | |
| 1919 | 6.42 | 1929 | 18.32 | | |
| 1920 | 6.94 | 1930 | 17.82 | | |
| 192[illegible] | 6.57 | 1831 | 17.36 | | |
| 1922 | 7.61 | 1932 | 16.22 | | |

（一）試根據上表所載美國哈特福特電力公司產銷電力數量，求複利曲綫長期趨勢總公式，並計算各年之趨勢值。

（二）試繪一曲綫圖以表，複利曲綫配合實際產銷數量之情形。

# 第二十章　季節變動

## 第一節　簡單月份平均數法

時間數列中第一種可以分別分析之因素爲長期趨勢，第二種可以個別分析之因素則爲季節變動。分析長期趨勢之前，須將統計資料製成曲綫圖，分析季節變動時亦須自作圖入手。所異者在分析季節變動中，其資料必爲月份資料，非復爲年份資料，將各年各月之資料逐月作成曲綫之際，首先入于吾人之視綫者爲循環變動及長期趨勢。如詳細觀察，則可見各個年份中，不拘其在循環變動中之爲上升或下降，各年中之某一月或某一季，每較其他各月或各季爲高漲或低落，足見有季節變動存在，可求季節變動指數(Indices of Seasonal Variation)以代表之。此項指數可逕簡稱爲季節指數 (Seasonal Indices)。倘分年逐月製圖尚不能斷定有季節變動存在時，可將每年十二月份資料繪入同一段代表十二月份資料之圖形中，每年十二月之變化以一曲綫表示之，如各綫在各月份中每共同起伏，則必有季節變動無疑。反之，如此起彼伏，毫無規律，則殆無季節變動。此種資料如經分析，所得各月份之季節指數彼此相同，如均爲100，則證明其毫無季節變動。如各月指數不同，則係表示有季節變動存在。吾人于計算季節指數時所採取之資料，在時間上自宜與長期趨勢分析中所用者相同。如以爲過長時，似亦可摘取其一部份資料。各年之季節變動有彼此相同之趨勢，則摘取一部份而計算之，所得結果與依據全部資料計算所得者雖不同，亦相去不太遠。但摘取之資料最短亦宜爲一循環變動之全部，

計算季節指數最簡易方法爲簡單月份平均數法 (Method of Simple

Monthly Averges)，或簡稱爲簡單月均法。現就直綫趨勢章中所有之波斯頓城結婚數目資料中摘取 1911 至 1920 共計十年一百二十月份之資料爲例，以說明季節指數計算之程序:

(第 174 表)　波斯頓城結婚數目季節指數計算表

| 年＼月 | 一 | 二 | 三 | 四 | 五 | 六 | 七 |
|---|---|---|---|---|---|---|---|
| 1911 | 640 | 628 | 593 | 622 | 440 | 1080 | 596 |
| 1912 | 530 | 545 | 315 | 658 | 467 | 381 | 524 |
| 1913 | 596 | 525 | 479 | 688 | 537 | 1110 | 575 |
| 1914 | 631 | 677 | 452 | 702 | 611 | 1188 | 570 |
| 1915 | 664 | 581 | 357 | 682 | 626 | 1078 | 525 |
| 1916 | 610 | 644 | 471 | 914 | 541 | 1174 | 676 |
| 1916 | 640 | 620 | 268 | 894 | 095 | 1400 | 749 |
| 1918 | 5·8 | 495 | 415 | 678 | 570 | 1 59 | 599 |
| 1919 | 476 | 509 | 432 | 585 | 544 | 1163 | 571 |
| 1920 | 622 | 576 | 352 | 782 | 520 | 1329 | 592 |
| 各月總計 | 6,010 | 5,800 | 4,034 | 6,905 | 5551 | 11462 | 5997 |
| 各月平均 | 601 | 5 0 | 406 | 691 | 555 | 1146 | 660 |
| 季節指數 | 87.1 | 84.1 | 58.8 | 100.0 | 80.5 | 165.1 | 87.0 |

續　上　表

| 年＼月 | 八 | 九 | 十 | 十一 | 十二 | 各年總計 | 各年平均 |
|---|---|---|---|---|---|---|---|
| 1911 | 585 | 781 | 8 2 | 856 | 593 | 8,111 | 676 |
| 1912 | 590 | 792 | 972 | 8 0 | 681 | 7,883 | 657 |
| 1913 | 637 | 780 | 854 | 835 | 503 | 8,120 | 677 |
| 1914 | 840 | 756 | 888 | 849 | 557 | 8,724 | 727 |
| 1016 | 637 | 758 | 924 | 790 | 509 | 8,131 | 678 |
| 1917 | 552 | 836 | 929 | 870 | 624 | 8,551 | 713 |
| 1918 | 684 | 744 | 869 | 884 | 694 | 9 191 | 766 |
| 1920 | 712 | 730 | 568 | 545 | 426 | 7,393 | 616 |
| 1921 | 691 | 743 | 922 | 883 | 603 | 8,122 | 677 |
| 1922 | 603 | 856 | 9 1 | 188 | 5·0 | 8,545 | 712 |
| 每月總計 | 6547 | 7778 | 87 9 | 8139 | 5745 | 合 | 6899 |
| 每月平均 | 655 | 778 | 876 | 816 | 575 | 8279 | 690 |
| 季節指數 | 94.9 | 112.8 | 1 7.0 | 118.3 | 83.3 | | 100.00 |

。上列第174表中每橫列一排中爲同一年內十二月份每月結婚數目，每縱一欄中爲各年同一月份之結婚數目。就各排而言，同一年之十二月份數值可相加，由總數而求出年份平均數 (Annual Averages)，記入計算表中最右一欄內。吾人在直綫趨勢章中計算長期趨勢時所用者卽係此項年份平均數。自縱列之各欄而言，則可施行下列各計算之步驟。

一、求各月份平均數：同一欄內所載旣爲各年同一月份之資料，自亦可由月份總和(Monthly Totals 進而求得月份平均數(Monthly Averages)，列入計算表最末一年各月資料以下一行內。月份平均數卽十年之內同一月份之平均數值。

二、求月份總平均數：總平均數 (The General Averages) 可自縱橫兩種平均數中求之。自縱行十個年份平均數中可得總平均數爲總月平均數；自橫列十二月份平均數中可得其總平均數爲總年均數。所根據之資料旣相同，無論縱橫求之，其總平均數必彼此相等。

三、求季節指數：由上表所得之各月份平均數 可見出結婚數目各月不同，現再以總平均數爲共同分母，將各月之平均數化爲百分數，列入最下一行中，卽爲季節指數。由指數中吾人可見波斯頓城結婚數目以六月之指數爲最高；九，十，十一等月份次之，而以三月爲最低，五月次之，十二，一，二等月份又次之。

## 第二節　改正月份平均數法

季節指數所表示者應只爲季節變動，而不受其他因素之紛擾。所謂其他因素自係指長期趨勢，循環變動及意外變動而言。就中意外變動，來自突然，且不常發生，可假定其不存在，或於其存在時加以說明。至於循環變動，如資料所包括之時間較長，起訖恰包括循環作用之全部，則第一年元

月份在產業繁榮時間，第二年元月份可在恐慌期間，第三年元月份可在蕭條期間，第四年元月份可在復興期間；依此四元月份資料求平均數時，則循環變動所產生之作用可從中抵消，所餘者只長期趨勢一因素須設法取消之，使季節指數純粹代表季節變動，不受其他任何因素之紛擾。此項指數稱為改正季節指數 (Corrected Seasonal Indices)，所用之方法為改正月份平均數法 (Method of Corrected Monthly Averages)，或稱為月份平均改正趨勢法(Method of Monthly Averages Corrected for Trend)茲分述此法之程序：

一、求各月份平均數：此與上法完全相同無須重述。

二、減消趨勢改正數：在直綫趨勢章中，依1900至1920共計二十一年波斯頓城結婚數目資料已得年份趨勢方程式為：

$$Y=625.1905+10.1675X$$

原點為1910年七月一日。月份趨勢方程式為：

$$Y_c=625.614+0.824X$$

原點為1910年七月十五日。即1910年七月十五日之趨勢值為625.614，以後每進一月，趨勢值增加0.824；以前每退一月，趨勢值減少0.824。在上節所得各月份平均數中，如以七月為原點時，則八月份平均數因趨勢作用而增加者為0.824，九月份平均數因趨勢作用而增加者0.824×2，十月份平均數因趨勢作用所加者為0.824×3，十一月份平均數因趨勢作用者增加者為0.824×4，十二月份平均數因趨勢作用而增加者為0.824×5，凡此均須自各月份平均數中減去，結果方代表純粹季節變動。再由七月份起後退觀之，七月份平均數既為趨勢作用而較六月份增加0.824，則六月份平均數須加上0.824，五月份平均數須加上0.824×2，四月份平均數須加上.824×3，三月份平均數須加上0.824×4，二月份平均數須加上0.824×5，

一月份平均數須加上0.824×6，所得結果方代表純粹季節變動。

（第175表） 改正月份平均數及季節指數計算表

| 月份 | 簡單月份數 | 改正數 | 改正月均數 | 季節指數 |
|---|---|---|---|---|
| 一 | 87.1 | +0.824×6 | 92.014 | 91.47 |
| 二 | 81.1 | +0.824×5 | 88.220 | 67.67 |
| 三 | 53.8 | +0.824×4 | 62.096 | 61.85 |
| 四 | 100.0 | +0.824×3 | 102 472 | 103 06 |
| 五 | 80.5 | +0.824×2 | 82.1 8 | 81.82 |
| 六 | 166.1 | +0.824×1 | 166.924 | 166.25 |
| 七 | 87.0 | 0 | 87 000 | 86.65 |
| 八 | 94.8 | −0 824×1 | 91.076 | 98.70 |
| 九 | 112.8 | −0 824×2 | 111.152 | 110.70 |
| 十 | 127.0 | −0 824×3 | 124.528 | 124.03 |
| 十一 | 118.3 | −0 824×4 | 115.004 | 114 54 |
| 十二 | 83.3 | −0 824×5 | 79.180 | 78.86 |
| 總計 | | | 1,204.844 | |
| 均平數 | | | 100. 04 | 100.00 |

三、 求改正季節指數：改正月份平均數求得後，已可見出季節變更大致情形。現再由十二改正月份平均數求出總平均數，並以之爲共同分母，將各月份改正平均數俱化爲百分數，即爲改正季節指數。

## 第三節 十二月移動平均數法

在長期趨勢章中已說明何爲移動平均數法。如所分析者爲年份資料並假定每一循環爲期五年，則五年移動平均數結果已將循環作用取消，而爲長期趨勢。同理，現所根據者爲月份資料，所分析者爲季節變動與長期趨勢，求十二月移動平均數(Twelve Month Moving Averages)之結果，自將季節變動取消，而爲長期趨勢數値。求得此月份趨勢値之後，再由原

始資料中設法消除之，吾人所分析者既爲季節變動與長期趨勢兩因素，長期趨勢消除之後，所餘自爲純粹季節變動，上係綜論十二月移動平均數法之要旨，現分述其程序：

一、　求十二月移動平均數：上表中每年每月俱有四行數字：第一行爲各

(第176表)　十二月移動平均數趨勢值計算表

| 年＼月 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 八 | 九 | 十 | 十一 | 十二 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1911 | 640 | 628 | 393 | 622 | 440 | 1080 | 596 | 585 | 781 | 892 | 856 | 593 |
| | | | | | | | 8111 | 8001 | 7918 | 7840 | 7876 | 7903 |
| | | | | | | | 676 | 667 | 660 | 653 | 656 | 659 |
| | | | | | | | 672 | 664 | 657 | 645 | 658 | 655 |
| 1912 | 530 | 545 | 315 | 658 | 467 | 981 | 524 | 580 | 792 | 972 | 830 | 681 |
| | 7804 | 7732 | 7737 | 7648 | 7828 | 7802 | 7885 | 7951 | 7931 | 8045 | 8125 | 8195 |
| | 650 | 644 | 635 | 646 | 652 | 650 | 657 | [illegible] | 661 | 675 | 677 | 683 |
| | 647 | 645 | 646 | 649 | 651 | 654 | 660 | 662 | 668 | 676 | 680 | 689 |
| 1913 | 596 | 525 | 479 | 688 | 537 | 1110 | 575 | 637 | 780 | 854 | 836 | 503 |
| | 8324 | 8375 | 8422 | 8410 | 8292 | 8298 | 8120 | 8158 | 8310 | 8283 | 8237 | 8371 |
| | 694 | 698 | 702 | 701 | 691 | 692 | [illegible] | 680 | 693 | 690 | 691 | 693 |
| | 696 | 700 | 702 | 696 | [illegible] | 685 | 679 | 687 | 692 | 691 | 695 | 701 |
| 1914 | 634 | 977 | 452 | 702 | 611 | 1188 | 570 | 840 | 756 | 888 | 849 | 557 |
| | 8449 | 8441 | 8647 | 8623 | 8657 | 8670 | 8724 | 8754 | 8658 | 8563 | 8543 | 8558 |
| | 704 | 704 | 721 | 719 | 721 | 723 | 727 | 730 | 722 | 714 | 712 | 713 |
| | 704 | 713 | 720 | 720 | 722 | 725 | 729 | 726 | [illegible] | 713 | 713 | 709 |
| 1915 | 664 | 581 | 357 | 682 | 626 | 1078 | 525 | [illegible] | 758 | 924 | 790 | [illegible] |
| | 8448 | [illegible] | 8200 | 820 | 8238 | 8179 | 8131 | 8077 | 8140 | 8254 | 8188 | [illegible] |
| | 704 | 700 | 681 | 684 | 687 | 682 | 678 | 673 | 678 | 688 | 632 | 675 |
| | 702 | 692 | 684 | 686 | 685 | 680 | 676 | 676 | 683 | 660 | 654 | 674 |
| 1916 | [illegible] | [illegible] | 471 | 614 | 541 | 1174 | 676 | 562 | 836 | 929 | 870 | 624 |
| | 8197 | 8348 | 8273 | 8315 | 8356 | 8436 | 8551 | 8581 | 8557 | 8484 | 8764 | 8918 |
| | 673 | 696 | 689 | 683 | 696 | 703 | 713 | 715 | 713 | 707 | 730 | 743 |
| | 685 | 993 | 686 | 890 | 700 | 708 | 714 | 714 | [illegible] | 719 | 737 | 748 |
| 1917 | 640 | 620 | 398 | 894 | 695 | 1300 | 769 | 684 | 744 | 869 | 884 | 694 |
| | 9644 | 9137 | 9259 | 9167 | 9107 | 9121 | 9191 | [illegible] | 9024 | 9041 | 8825 | [illegible] |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 754 | 761 | 772 | 764 | 729 | 720 | 768 | 762 | 742 | 753 | 735 | 725 |
| | 758 | 767 | 798 | 747 | 725 | 743 | 764 | 752 | 748 | 744 | 730 | 715 |
| 1918 | 593 | 495 | 415 | 678 | 570 | 059 | 599 | 712 | 730 | 568 | 545 | 426 |
| | 8459 | 8289 | 8317 | 8303 | 8002 | 7663 | 7395 | 7273 | 7287 | 7304 | 72 1 | 7185 |
| | 705 | 631 | 683 | 692 | 667 | 6 9 | 6 6 | 606 | 607 | 609 | 601 | 5 9 |
| | 693 | 982 | 688 | 680 | 653 | 628 | 611 | 607 | 608 | 605 | 600 | 6 3 |
| 1919 | 470 | 503 | 432 | 585 | 544 | 1163 | 571 | 691 | 743 | 922 | 883 | 6 3 |
| | 7289 | 7261 | 7240 | 7253 | 7607 | 7345 | 8122 | 8268 | 8335 | 8255 | 8452 | 8428 |
| | 607 | 605 | 603 | 604 | 634 | 662 | 677 | 689 | 695 | 638 | 704 | 7 2 |
| | 666 | 601 | 604 | 616 | 648 | 670 | 683 | 6 2 | 6 2 | 696 | 703 | 709 |
| 1920 | 622 | 576 | 352 | 782 | 520 | 1329 | 592 | 609 | 856 | 911 | 816 | 553 |
| | 8594 | 8615 | 8583 | 8646 | 8565 | 8598 | 8545 | | | | | |
| | 716 | 718 | 711 | 721 | 722 | 717 | 712 | | | | | |
| | 717 | 715 | 716 | 722 | 720 | 715 | | | | | | |

續 上 表

月實際數值，第二行爲十二月移動總數，第三行爲十二月移動平均數，卽上列總數除以 12。全部資料起於1911年一月，止於1920年十二月，1911年內十二月平均之結果所得爲1911年七月一日之數值。1911年上半年中六月之趨勢值自無法求得；同理最末一年卽 1920 下半年六月趨勢值亦無法求得。第四行爲改正十二月平均數。所以需要改正者，因第一行實際數值係代表每月十五日之數值，而第三行之平均數則係代表每月一日之數值，兩者相差半月，無法比較。將已得之十二月移動平均數再求二月移動平均數卽爲改正。蓋已得七月一日及八月一日平均數後，再求此二月之平均數、自爲七月十五日之數值。其他各月平均數俱按此法改正之，所得爲最後十二月移動平均數，係代表每月月中之趨勢值者。

二、 取消趨勢因素：分析季節變動之要點在自時間數列中將趨勢一因素取消。上列所得十二月移動平均數，卽係月份趨勢值。以每月之趨勢值除同月份實際值所得之百分比中，卽已將趨勢一因素取消，稱爲實際對趨勢

比例法(Actual-to-trend Method)，以符號表之：$R=A/T$此種月份百分比$R$中既已不包括長期趨勢，自係純粹代表月與月間之變動者：

(第177表)　純粹月與月間變動百分比計算表　$R=A/T$

| 年＼月 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 八 | 九 | 十 | 十一 | 十二 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1911 | 640 | 628 | 393 | 622 | 440 | 1080 | 596 | 585 | 781 | 892 | 856 | 598 |
| | | | | | | | 672 | 664 | 657 | 645 | 658 | 655 |
| | | | | | | | 89 | 88 | 119 | 138 | 130 | 91 |
| 1912 | 530 | 545 | 315 | 653 | 467 | 981 | 524 | 590 | 792 | 97[illegible] | 830 | 681 |
| | 647 | 645 | 646 | 643 | 651 | 654 | 680 | 662 | 668 | 676 | 680 | 689 |
| | 82 | 84 | 49 | 101 | 72 | 150 | 70 | 89 | 119 | 144 | 122 | 99 |
| 1913 | 596 | 525 | 473 | 688 | 537 | 1110 | 575 | 637 | 780 | 854 | 839 | 503 |
| | 696 | 700 | 702 | 696 | 692 | 685 | 679 | 687 | 692 | 691 | 695 | 701 |
| | 86 | 75 | 68 | 99 | 78 | 162 | 85 | 93 | 113 | 124 | 120 | 72 |
| 1914 | 634 | 677 | 452 | 702 | 611 | 1188 | 570 | 840 | 756 | 888 | 849 | 557 |
| | 704 | 713 | 720 | 720 | 722 | 725 | 720 | 726 | 718 | 713 | 713 | 709 |
| | 90 | 95 | 63 | 98 | 85 | 164 | 78 | 116 | 105 | 125 | 119 | 79 |
| 1915 | 664 | 581 | 357 | 682 | 626 | 1078 | 525 | 637 | 758 | 924 | 790 | 519 |
| | 702 | 692 | 684 | 686 | 685 | 680 | 676 | 676 | 683 | 660 | 654 | 674 |
| | 95 | 84 | 52 | 99 | 91 | 151 | 78 | 94 | 111 | 140 | 121 | 76 |
| 1916 | 610 | 644 | 471 | 614 | 541 | 1174 | 676 | 562 | 836 | 923 | 870 | 624 |
| | 685 | 693 | 686 | 690 | 700 | 708 | 714 | 714 | 710 | 719 | 737 | 749 |
| | 89 | 93 | 69 | 89 | 77 | 166 | 95 | 79 | 118 | 129 | 118 | 83 |
| 1917 | 640 | 620 | 398 | 894 | 695 | 1300 | 769 | 684 | 744 | 869 | 884 | 694 |
| | 758 | 767 | 768 | 747 | 725 | 743 | 784 | 752 | 748 | 744 | 730 | 715 |
| | 84 | 81 | 52 | 120 | 96 | 175 | 101 | 91 | 99 | 117 | 121 | 97 |
| 1918 | 598 | 495 | 415 | 678 | 570 | 1059 | 599 | 712 | 730 | 568 | 545 | 426 |
| | 693 | 682 | 688 | 680 | 653 | 628 | 611 | 607 | 608 | 605 | 610 | 603 |
| | 86 | 73 | 60 | 100 | 87 | 169 | 96 | 117 | 120 | 94 | 91 | 71 |
| 1919 | 476 | 509 | 432 | 585 | 544 | 1163 | 571 | 691 | 743 | 922 | 883 | 603 |
| | 606 | 604 | 604 | 619 | 648 | 670 | 6[illegible]3 | 692 | 692 | 696 | 703 | 709 |
| | 79 | 84 | 72 | 95 | 84 | 174 | 84 | 100 | 107 | 132 | 126 | 85 |
| 1920 | 622 | 576 | 352 | 782 | 520 | 1[illegible]29 | 592 | 609 | 856 | 941 | 816 | 550 |
| | 717 | 715 | 716 | 722 | 720 | 715 | | | | | | |
| | 87 | 81 | 49 | 108 | 72 | 186 | | | | | | |

上表中每月份有三行數字:第一行爲實際值,第二行爲改正十二月移動平均數,卽趨勢值,第三行則爲實際對趨勢之百分比。此百分比卽係純粹代表月與月間之季節變動者,不包括其他因素。

三、 求各月份百分比之中位數:上表所得百分比爲代表月與月間之變動或季節變動者。在各年度中同一月份之百分比固高低不同,但概括觀之,仍有一致上升或一致下降之趨勢。如將同一月份各百分比作成次數分配,顯然有集中趨勢存在。每一月份中各百分比均可求出中位數以代表各月份一般變動,稱爲月份百分比中位數(Medians of Monthly Per Cent Relatves)。自未分組資料中,求中位數時須先排製序列,如包括之變項爲奇數,則居於中央之變項卽爲中位數。如變項爲偶數,則中央兩變項之平均數爲中位數。現同一月份中旣各有九個百分比,變項爲奇數,是以中央變項卽爲中位數。

(第178表) 波斯頓城結婚對數月份百分比中位數計算表

| 序列＼月份 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 八 | 九 | 十 | 十一 | 十二 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 79 | 73 | 49 | 89 | 72 | 159 | 78 | 79 | 93 | 94 | 91 | 71 |
| 2 | 82 | 75 | 49 | 95 | 72 | 159 | 78 | 88 | 105 | 117 | 118 | 72 |
| 3 | 84 | 81 | 52 | 93 | 77 | 162 | 79 | 89 | 107 | 124 | 119 | 76 |
| 4 | 89 | 81 | 52 | 99 | 78 | 164 | 84 | 91 | 111 | 125 | 120 | 79 |
| 5 | 86 | 84 | 60 | 9[illegible] | 84 | 166 | 85 | 93 | 113 | 121 | 121 | 83 |
| 6 | 87 | 84 | 63 | 100 | 85 | 169 | 89 | 94 | 118 | 132 | 121 | 85 |
| 7 | 89 | 84 | 68 | 101 | 87 | 174 | 95 | 100 | 119 | 138 | 122 | 91 |
| 8 | 90 | 93 | 69 | 108 | 91 | 175 | 98 | 116 | 116 | 140 | 126 | 97 |
| 9 | 95 | 95 | 72 | 120 | 98 | 186 | 101 | 117 | 120 | 144 | 130 | 99 |
| 中位數 | 86 | 84 | 60 | 99 | 84 | 166 | 85 | 93 | 11[illegible] | 1[illegible]9 | 121 | 8[illegible] |

四、 求季節指數:上表所得十二月份百分比中位數已爲代表十年之內各月份季節變動之量數。爲便於與其他方法所得結果比較起見,復求十二月

份中位數之總平均數，並以此總平均數爲共同分母，將各月之百分比中位數化爲此總平均數之百分比，卽爲最後之季節指數。

（第179表） 季節指數計算表

| 月份 | 中位數 | 季節指數 |
|---|---|---|
| 一 | 86 | 85.79 |
| 二 | 84 | 83.79 |
| 三 | 60 | 59.85 |
| 四 | 69 | 98.75 |
| 五 | 84 | 83.79 |
| 六 | 166 | 165.59 |
| 七 | 85 | 84.79 |
| 八 | 93 | 92.77 |
| 九 | 113 | 112.72 |
| 十 | 129 | 128.68 |
| 十一 | 121 | 120.70 |
| 十二 | 83 | 82.79 |
| 總計 | 1203 | |
| 平均數 | 100.25 | 100.00 |

## 第四節 環比中位數法

環比中位數法(Method of Median Link Relatives)由美國經濟統計學家波森斯教授(Prof. Warren M. Persons)所提倡，學者共認係求季節變動指數最優良之方法，其計算之步驟大致如下：

一、 求月份環比 所謂月份環比(Monthly Link Relatives)卽以上月實際數值，除本月實際數值所得之百分比，現以波斯頓城結婚數目，計算各月環比如下：

（第180表）　波斯頓城結婚數目各月環比計算表

環比 $(L)$ = 本月/上月

| 年＼月 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 八 | 九 | 十 | 十一 | 十二 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1911 | 640 | 6 8 | 393 | 622 | 440 | 1080 | 596 | 585 | 781 | 892 | 856 | 598 |
| |  | 98 | 63 | 158 | 71 | 245 | 55 | 98 | 134 | 114 | 96 | 70 |
| 1912 | 530 | 545 | 315 | 658 | 467 | 981 | 524 | 590 | 792 | 972 | 8 0 | 681 |
| | 89 | 103 | 58 | 209 | 71 | 210 | 53 | 113 | 134 | 123 | 85 | 82 |
| 1913 | 596 | 5[illegible]5 | 479 | 688 | 537 | 1110 | 57[illegible] | 637 | 780 | 854 | 836 | 503 |
| | 88 | 88 | 91 | 144 | 78 | 207 | 52 | 111 | 122 | 109 | [illegible]8 | 60 |
| 1914 | 634 | 677 | 452 | 702 | 611 | 1184 | 570 | 810 | 756 | 986 | 849 | 557 |
| | 126 | 107 | 67 | 155 | 87 | 194 | 48 | 147 | 90 | 117 | [illegible]6 | 66 |
| 1915 | 664 | 581 | 357 | 682 | 626 | 1075 | 525 | 637 | 758 | 924 | 790 | 503 |
| | 119 | 88 | 73 | 191 | 92 | 172 | 49 | 121 | 119 | 122 | 85 | 64 |
| 1916 | 610 | 644 | 471 | 614 | 541 | 1174 | 676 | 562 | 876 | 929 | 570 | 624 |
| | 120 | 106 | 73 | 130 | 88 | 217 | 58 | 83 | 149 | 1[illegible]1 | 94 | 72 |
| 1917 | 640 | 623 | 398 | 894 | 695 | 1300 | 769 | 684 | 744 | 869 | 681 | 694 |
| | 103 | 97 | 64 | 225 | 78 | 187 | 59 | 89 | 109 | 78 | 102 | 79 |
| 1918 | 98 | 495 | 415 | 678 | 570 | 1059 | 599 | 712 | 730 | 568 | 545 | 426 |
| | 86 | 83 | 84 | 133 | 84 | 186 | 57 | 119 | 10[illegible] | 78 | 96 | 78 |
| 1919 | 476 | 509 | 432 | 585 | 544 | 1163 | 571 | 691 | 741 | 922 | 883 | 603 |
| | 112 | 107 | 85 | 135 | 93 | 214 | 49 | 121 | 108 | 124 | 96 | 68 |
| 1920 | 622 | 576 | 352 | 782 | 521 | 1329 | 592 | 609 | 856 | 941 | 516 | 550 |
| | 103 | 93 | 61 | 222 | 66 | 256 | 45 | 103 | 141 | 110 | 87 | 67 |

二、 求各月份環比之中位數：上表中每一月份有二行數字，第一行爲實際數值，第二行上月除本月所得之連環百分比。此種百分比係代表月與月間之變動者。現由同一月份各環比中求出其中位數，以代表各月份之平均變動，稱爲環比中位數（Medians of Link Relatives）仍須如上節中排成序列。除第一月以外，均爲雙數變項，自須由位居中央兩變項求平均數方得中位數。

(第181表)　環比中位數計算表

| 序另＼月份 | 一 | 二 | 三 | 四 | 五 | 六 | 七 | 八 | 九 | 十 | 十一 | 十二 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 8 | 58 | 130 | 66 | 172 | 45 | 63 | 90 | 78 | 85 | 60 |
| 2 | 86 | 88 | 61 | 135 | 71 | 186 | 48 | 89 | 103 | 78 | 8 | 84 |
| 3 | 88 | 89 | 63 | 144 | 71 | 187 | 49 | 98 | 108 | 109 | 87 | 66 |
| 4 | 89 | 93 | 61 | 155 | 78 | 194 | 49 | 103 | 109 | 110 | 94 | 67 |
| 5 | 103 | 97 | 67 | 158 | 78 | 207 | 52 | 111 | 119 | 111 | 96 | 68 |
| 6 | 103 | 98 | 73 | 163 | 94 | 210 | 53 | 113 | 122 | 114 | 96 | 70 |
| 7 | 112 | 103 | 73 | 191 | 87 | 214 | 55 | 119 | 134 | 117 | 96 | 72 |
| 8 | 119 | 106 | 84 | 209 | 88 | 217 | 57 | 121 | 134 | 122 | 96 | 78 |
| 9 | 120 | 107 | 65 | 222 | 92 | 245 | 58 | 121 | 141 | 123 | 98 | 79 |
| 10 | 126 | 107 | 91 | 225 | 93 | 256 | 59 | 147 | 149 | 124 | 102 | 82 |
| 中位數 | 103 | 98 | 70 | 161 | 81 | 209 | 53 | 112 | 121 | 113 | 96 | 69 |

三、　求連鎖百分比：上表最底下一行所載十二月份環比中位數係代表十年之中各月份平均季節變動者。但其中尚包括有長期趨勢，須設法消除之。其方法爲求連鎖百分比(Chain Relatives)或簡稱鎖比，以 $C$ 代表之。除第一元月份之鎖比假定爲 100 外，其他各月份之鎖比槪爲上月份鎖比與本月份環比之乘積。以二月份鎖比爲例，$C_2 = C_1 \times L_2$，各月鎖比依下表計算之。

(第182表)　十二月份鎖比計算表

| 月份 | 乘積 | 鎖比 |
|---|---|---|
| 元 | | 100.00(假定值) |
| 二 | 100.00×98 | 98.00 |
| 三 | 98.00×70 | 68.60 |
| 四 | 68.60×161 | 110.45 |
| 五 | 110.45×81 | 89.46 |
| 六 | 89.46×209 | 186.97 |
| 七 | 186.97×53 | 99.09 |

續　上　表

| | | |
|---|---|---|
| 八 | 99.09×112 | 110.98 |
| 九 | 110.98×121 | 134.29 |
| 十 | 134.29×113 | 1 1.75 |
| 十一 | 1 1.75×96 | 1 5.68 |
| 十二 | 145.68×69 | 100.52 |
| 元 | 100.52×103 | 103.54(計算値) |

四、取消趨勢因素：上列鏈比中第一元月份之鏈比假定爲100，第二元月份由計算而得之鏈比則爲103.54，二者間之差額卽代表由長期趨勢作用所致，須設法取消之。在十二月中由於趨勢因素所產生之數值爲103.54－100＝3.54，係一正數，卽表示長期趨勢係向上。在十二月中趨勢因素所造成者旣爲3.54，則月與月間由趨勢因素所致之差數自爲 3.54/12＝

（第183表）改正鏈比計算表

| 月　份 | 原有鏈比 | 改正數 | 改正鏈比 | 季節指數 |
|---|---|---|---|---|
| 元 | 100.00 | 0 | 100.000 | 87.17 |
| 二 | 98.20 | 0.295 | 97.905 | 85.34 |
| 三 | 68.60 | 0.590 | 68.100 | 59.36 |
| 四 | 110.48 | 0.885 | 109.565 | 95.51 |
| 五 | 89.46 | 1.180 | 88.280 | 76.95 |
| 六 | 186.37 | 1.475 | 185.495 | 161.70 |
| 七 | 99.09 | 1.770 | 97.320 | 82.83 |
| 八 | 110.98 | 2.065 | 108.915 | 94.94 |
| 九 | 134.29 | 2.360 | 131.930 | 115.00 |
| 十 | 151.75 | 2.655 | 149.095 | 1 9.97 |
| 十一 | 145.68 | 2.950 | 142.790 | 124.42 |
| 十二 | 100.52 | 3.245 | 97.275 | 84.79 |
| 元 | 103.54 | 3.540 | 100.090 | |
| 總計 | | | 1,376.610 | |
| 平均數 | | | 11 .718 | 100.00 |

0.295。以第一元月份爲標準，每兩月間之差數既爲 0.295，則逐月將此差數累積減去之後，第一元月與第二元月之鎖比必彼此相同，可稱爲改正鎖比 Chain Relatives Corrected for Trend)。

五、　求季節指數：上表所得改正鎖比已爲代表月與月間純粹季節變動之百分比，爲求與他法所得結果易於比較起見，仍求十二月份改正鎖比之總平均數，

$$總平均數=\frac{十二月改正鎖比之和}{12}=\frac{1376.61}{12}=114.718$$

求得此總平均數後，即以之爲共同分母，將各月份之改正鎖比化爲此總平均數之百分比，即爲最後季節指數，列入上表中最右一欄內。

## 第五節　各種季節指數之比較

依以上四種不同方法所得之季節指數可列入同一表中互相比較，並可作圖以明之：

（第 184 表）　波斯頓城結婚數目季節指數比較表

| 類別 / 月份 | 簡單月均數 | 改正月均數 | 十二月移動平均數 | 環比中位數 |
|---|---|---|---|---|
| 一 | 87.1 | 91.67 | 85.79 | 87.17 |
| 二 | 84.1 | 87.87 | 83.79 | 85.34 |
| 三 | 53.8 | 61 86 | 59.85 | 59 36 |
| 四 | 100.0 | 102.06 | 98.75 | 95.51 |
| 五 | 80.5 | 81.82 | 83.79 | 76.93 |
| 六 | 166.1 | 166.25 | 165·59 | 161. 0 |
| 七 | 87.0 | 86.65 | 84.79 | 82.83 |
| 八 | 91.9 | 93.70 | 92.77 | 94.94 |
| 九 | 112.8 | 110.70 | 11 .72 | 115.00 |
| 十 | 127.0 | 124.03 | 1 8 68 | 1 9.97 |
| 十一 | 118.3 | 114.54 | 1 0.70 | 1 4.42 |
| 十二 | 83.3 | 78 86 | 80.79 | 84. 9 |

由上列第184表，可見四種方法所得之季節指數大致相同，回顧四法，可作結語如下：

第一法，即簡單月份平均數法手續簡便，且易於了解。是否合用，須視統計資料受長期趨勢影響之程度而定。如資料受長期趨勢之影響不大，則長期趨勢一因素可以不顧，而用此法。同時須循環變動在全期中分配相當勻稱，以各年中同一月份資料平均之結果，循環作用於中沖消。此種時間數列雖不受趨勢及循環變動之影響，但同一月份中之資料作爲次數分配每有偏態，是算術平均數不免受極大與極小變項之影響，而使所得之季節指數不可靠。

第二法，即改正月份平均數法，已能確定趨勢作用並設法取消之，是已比第一法爲進步，但其缺點仍爲算術平均數公式之受極端變項之影響。

第三法，即十二月移動平均數法之優點：第一，係於依十二月移動平均數法求得月份趨勢值之後，再用實際對趨勢法取消趨勢作用，使下餘之百分數純粹代表月與月間之季節變動。第二，係以中位數決定指數，免受極端變項之影響。其缺點則爲首尾兩端各有六月份無法求出趨勢值。

第四法，即環比中位數法之缺點爲計算手續該繁；其優點則爲所得結果比較可靠。

## 習　　題

一、 試根據上章所載一九一九至一九三七年美國生鐵各月產量，按簡單月均數法求季節指數。

二、 試根據同一資料按改正月均數法求季節指數。

三、 試根據十二月移動平均數法求季節指數。

四、 試根據環比中位數法求季節指數。

五、 試將四法所得結果作一總表,並繪一圖，分別以一綫曲代表每種季節指數。

# 第二十一章　循環變動

## 第一節　減消法

時間數列與次數分配兩種資料在分析之程序上頗相類似，在次數分配中，首先求平均數以代表資料之常態或集中趨勢，稱為集勢量數，次求標準差以代表資料之變態或離中趨勢，稱為離勢量數。現在時間數列中，首先求長期趨勢數值及季節指數，以代表資料之常態或集中趨勢，次則自原始資料中將趨勢與季節兩因素所發生之作用取消，使剩餘之差數代表離中趨勢或循環與意外兩種變動之總和。但意外變動不常發生，故時間數列中之離差實即代表循環變動者。在次數分配中，離勢量數有絕對與相對兩類，在時間數列中之循環數值亦有絕對與相對兩種，計算方法大致與在次數分配中相同，稱為減消法(The Deduction Method)。

一、　年份資料　時間數列如為份資料，則各年份實際數值(Actual Values)，以$A$代表之，各年份趨勢數值(Annual Trend Urdinates)則以$T$代表之。循環變動既為離差，可以 $Y$ 代表之，$y=(A-T)$，即循環數值等于由趨勢值所得之離差。由此離差亦可進而求標準差，並以標準差為共同分母，將各年之循環數值普遍以標準差除之。未除以標準差之前，$y$ 所代表者為絕對循環數值。既除之後，$y/\sigma$ 所代表者為相對循環數值。現以波斯頓城結婚資料為例，以說明由減消法求循環數值之步驟：

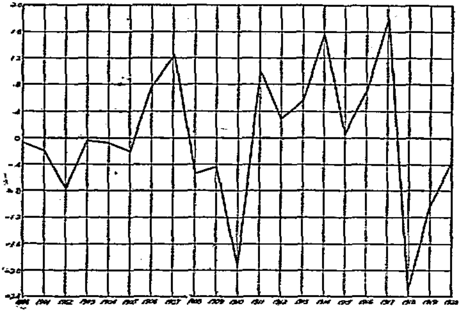
二、　月份資料　如所分析之數列為月份資料，除仍以 $A$ 代表各月份實際數值，以 $T$ 代表各月份趨勢數值外，並以 $S$ 代表各月份之季節指數，趨勢與季節兩因素之乘積為 $TS$，係代表資料之常態或集中趨勢。由此乘

(第185表) 波斯頓城結婚數目各年份循環值計算表

$$\sigma=\sqrt{\frac{\sum(A-T)^2}{N}}=\sqrt{\frac{32220}{21}}=39.2$$

| 年份 | 實際數值 $A$ | 趨勢數值 $T$ | 絕對循環值 $y=A-T$ | $y^2$ | 相對循環值 $y/\sigma$ |
|---|---|---|---|---|---|
| 1900 | 521 | 524 | −3 | 9 | −.08 |
| 1901 | 526 | 534 | −8 | 64 | −.20 |
| 1902 | 514 | 544 | −30 | 900 | −.77 |
| 1903 | 553 | 554 | −1 | 1 | −.03 |
| 1904 | 561 | 564 | −3 | 9 | −.08 |
| 1905 | 565 | 574 | −9 | 81 | −.22 |
| 1906 | 615 | 585 | +30 | 900 | +.77 |
| 1907 | 643 | 595 | +48 | 2304 | +1.22 |
| 1908 | 583 | 605 | −22 | 484 | −.56 |
| 1909 | 598 | 615 | −17 | 289 | −.43 |
| 1910 | 551 | 625 | −74 | 5476 | −1.89 |
| 1911 | 676 | 635 | +41 | 1681 | +1.05 |
| 1912 | 657 | 646 | +11 | 121 | +.28 |
| 1913 | 677 | 656 | +21 | 441 | +.54 |
| 1914 | 727 | 666 | +61 | 3721 | +1.56 |
| 1915 | 678 | 676 | +2 | 4 | +.05 |
| 1916 | 713 | 686 | +27 | 729 | +.69 |
| 1917 | 766 | 696 | +70 | 4900 | +1.79 |
| 1918 | 616 | 707 | −91 | 8281 | −2.32 |
| 1919 | 677 | 717 | −40 | 1600 | −1.02 |
| 1920 | 712 | 727 | −15 | 225 | −.38 |
| | | | | 32220 | |

積即常態值所求之離差，即係循環數值。就中亦有絕對與相對之別。現以波斯頓城各月結婚數目為例，製成第186表如下，最後並作一圖：



第46圖 波斯頓結婚數目年份循環數值圖

(第186表) 波斯頓城結婚數目月份循環值計算表

$$\sigma=\sqrt{200992/252}=89.3$$

| 年 月 | A | T | S | TS | y=A-TS | $y^2$ | $y/\sigma$ |
|---|---|---|---|---|---|---|---|
| 19.0一 | 523 | 521.8 | 87.2 | 454 | + 69 | 4761 | + .77 |
| 二 | 485 | 522.6 | 85.3 | 445 | + 40 | 1600 | + .45 |
| 三 | 276 | 523.4 | 59.4 | 309 | − 33 | 1089 | − .37 |
| 四 | 538 | 524.3 | 95.5 | 503 | + 35 | 1225 | + .39 |
| 五 | 531 | 525.1 | 77.0 | 404 | +127 | 16129 | +1.42 |
| 六 | 711 | 525.9 | 161.7 | 852 | −141 | 719881 | −1.58 |
| 七 | 431 | 526.7 | 82.8 | 437 | − 6 | 36 | − .07 |
| 八 | 426 | 527.6 | 94.9 | 501 | − 75 | 5625 | − .84 |
| 九 | 493 | 528.4 | 115.0 | 607 | −114 | 12996 | −1.28 |
| 十 | 731 | 529.2 | 130.0 | 688 | + 43 | 1849 | + .48 |
| 十一 | 688 | 530.0 | 124.4 | 657 | + 31 | 961 | + .35 |
| 十二 | 416 | 530.9 | 84.8 | 451 | − 35 | 12250 | − .39 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1901一 | 501 | 531.7 | 87.2 | 463 | +38 | 1440 | −.43 |
| 二 | 468 | 532.5 | 85.3 | 453 | +15 | 225 | +.17 |
| 三 | 277 | 533.3 | 59.4 | 314 | −37 | 1369 | −.41 |
| 四 | 585 | 534.2 | 95.5 | 513 | +72 | 5184 | +.81 |
| 五 | 382 | 535.0 | 77.0 | 412 | −30 | 900 | −.34 |
| 六 | 807 | 535.8 | 161.7 | 868 | −61 | 3721 | −.68 |
| 七 | 427 | 536.6 | 82.8 | 446 | −19 | 361 | −.21 |
| 八 | 447 | 537.4 | 93.9 | 510 | −63 | 3969 | −.71 |
| 九 | 549 | 538.3 | 115.0 | 619 | −70 | 4900 | −.78 |
| 十 | 721 | 539.1 | 130.0 | 701 | +20 | 400 | +.22 |
| 十一 | 716 | 540.0 | 124.4 | 670 | +46 | 2116 | +.52 |
| 十二 | 430 | 540.7 | 84.8 | 458 | −28 | 784 | −.31 |
| 1902一 | 482 | 541.6 | 87.2 | 471 | +11 | 121 | +.12 |
| 二 | 441 | 542.4 | 85.3 | 461 | −20 | 400 | −.22 |
| 三 | 334 | 543.2 | 59.4 | 320 | +14 | 196 | +.16 |
| 四 | 592 | 544.0 | 95.5 | 522 | +70 | 4900 | +.78 |
| 五 | 370 | 544.9 | 77.0 | 420 | −50 | 2500 | −.56 |
| 六 | 802 | 545.7 | 161.7 | 885 | −83 | 6889 | −.93 |
| 七 | 422 | 546.5 | 82.8 | 453 | −31 | 961 | −.35 |
| 八 | 414 | 547.3 | 94.9 | 520 | −106 | 11236 | −1.19 |
| 九 | 581 | 548.2 | 115.0 | 630 | −49 | 2401 | −.55 |
| 十 | 657 | 549.0 | 130.0 | 714 | −57 | 3249 | −.64 |
| 十一 | 649 | 549.8 | 124.4 | 682 | −33 | 1089 | −.37 |
| 十二 | 428 | 550.6 | 84.8 | 468 | −40 | 1600 | −.45 |
| 1903一 | 500 | 551.5 | 87.2 | 480 | +20 | 400 | +.22 |
| 二 | 490 | 552.3 | 85.3 | 469 | +21 | 441 | +.21 |
| 三 | 277 | 553.1 | 59.4 | 326 | −49 | 2401 | −.55 |
| 四 | 638 | 553.9 | 95.5 | 531 | +107 | 11449 | +1.20 |
| 五 | 385 | 554.8 | 77.0 | 427 | −42 | 1764 | −.47 |
| 六 | 881 | 555.6 | 161.7 | 900 | −19 | 361 | −.21 |
| 七 | 436 | 556.4 | 82.8 | 461 | −25 | 625 | −.28 |
| 八 | 500 | 557.2 | 94.9 | 529 | −29 | 841 | −.32 |
| 九 | 647 | 558.0 | 115.0 | 642 | +5 | 25 | +.06 |
| 十 | 712 | 558.9 | 130.0 | 727 | −15 | 225 | −.17 |
| 十一 | 718 | 560.0 | 124.4 | 706 | +12 | 144 | +.13 |
| 十二 | 456 | 560.5 | 84.8 | 476 | −20 | 400 | −.22 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1904一 | 527 | 561.3 | 87.2 | 488 | + 39 | 1 21 | + .44 |
| 二 | 476 | 562.2 | 85.3 | 478 | − 2 | 4 | − .02 |
| 三 | 275 | 593.0 | 59.4 | 332 | − 57 | 32 9 | − .61 |
| 四 | 622 | 563.8 | 95.5 | 541 | +121 | 14641 | +1.35 |
| 五 | 4 3 | 564.4 | 77.0 | 435 | − 2 | 1024 | − .36 |
| 六 | 960 | 565.5 | 161.7 | 917 | + 43 | 1849 | + .48 |
| 七 | 469 | 566 3 | 82.8 | 470 | − 1 | 1 | − .01 |
| 八 | 471 | 567.1 | 94.9 | 539 | − 68 | 4624 | − .76 |
| 九 | 601 | 567.9 | 115.0 | 653 | − 52 | 2704 | − .58 |
| 十 | 788 | 568.8 | 130.0 | 740 | + 48 | 2304 | + .54 |
| 十一 | 728 | 569.6 | 124.4 | 707 | + 21 | 441 | + .24 |
| 十二 | 416 | 570.4 | 84.8 | 485 | − 69 | 4761 | − .77 |
| 1905一 | 480 | 571. | 87.2 | 497 | − 17 | 289 | − .19 |
| 二 | 462 | 572.1 | 85.3 | 486 | − 24 | 579 | − .27 |
| 三 | 387 | 572.9 | 59.4 | 338 | + 49 | 2401 | − .55 |
| 四 | 497 | 573.7 | 95.5 | 551 | − 54 | 2916 | − .60 |
| 五 | 421 | 574.5 | 77.0 | 443 | − 22 | 484 | − .25 |
| 六 | 916 | 575.4 | 161.7 | 932 | − 16 | 256 | − .18 |
| 七 | 508 | 576.2 | 82.8 | 478 | + 30 | 900 | + .34 |
| 八 | 471 | 577.0 | 94.9 | 548 | − 77 | 5929 | − .86 |
| 九 | 659 | 577.8 | 113.0 | 665 | − 6 | 36 | − .07 |
| 十 | 766 | 578.6 | 130.0 | 753 | + 13 | 169 | + .15 |
| 十一 | 734 | 579 5 | 124.4 | 710 | + 15 | 225 | + .17 |
| 十二 | 474 | 580.3 | 84.8 | 493 | − 19 | 361 | − .21 |
| 19 6一 | 553 | 581.1 | 87.2 | 505 | + 48 | 2304 | + .54 |
| 二 | 554 | 58 .9 | 85.3 | 495 | + 59 | 3481 | + .66 |
| 三 | 335 | 582.8 | 59.4 | 344 | − 9 | 81 | − .10 |
| 四 | 6 8 | 583.6 | 95.5 | 560 | + 88 | 7744 | + .99 |
| 五 | 412 | 584.4 | 77.0 | 450 | − 38 | 1444 | − .43 |
| 六 | 952 | 585.2 | 161.7 | 948 | + 4 | 16 | + .04 |
| 七 | 492 | 585.1 | 82.8 | 466 | + 6 | 36 | + .07 |
| 八 | 573 | 586.9 | 94.9 | 499 | + 74 | 5496 | + .83 |
| 九 | 739 | 587.7 | 115.0 | 676 | + 63 | 3969 | + .70 |
| 十 | 830 | 588.5 | 130.0 | 765 | + 65 | 4225 | + .73 |
| 十一 | 790 | 589.4 | 124.4 | 730 | + 60 | 3600 | + .67 |
| 十二 | 498 | 590.2 | 84.8 | 502 | − 6 | 36 | − .07 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1907一 | 630 | 591.0 | 87.2 | 514 | +116 | 13456 | +1.30 |
| 二 | 517 | 591.8 | 85.3 | 503 | + 14 | .96 | + .16 |
| 三 | 384 | 592.7 | 59 4 | 350 | + 34 | 1156 | + .38 |
| 四 | 668 | 593.5 | 95.5 | 569 | + 99 | 9801 | +1.11 |
| 五 | 483 | 594.3 | 77.0 | 457 | + 26 | 676 | + .29 |
| 六 | 1052 | 595.1 | 161.7 | 964 | + 88 | 7744 | + .99 |
| 七 | 594 | 596.0 | 82.8 | 495 | + 99 | 9801 | +1 11 |
| 八 | 554 | 596.8 | 94.9 | 567 | — 13 | 169 | — .15 |
| 九 | 745 | 597.6 | 115.0 | 688 | + 57 | 32.9 | + .64 |
| 十 | 842 | 558.4 | 130.0 | 777 | + 65 | 4225 | + .73 |
| 十一 | 8 8 | 599.2 | 124.4 | 743 | + 65 | 4225 | + .73 |
| 十二 | 433 | 600.1 | 84.8 | 510 | — 77 | 5929 | — .86 |
| 19 8一 | 568 | 600.9 | 87.2 | 523 | + 45 | 2025 | + .50 |
| 二 | 506 | 601.7 | 85.3 | 512 | — 6 | 3 | — .07 |
| 三 | 421 | 602.5 | 59.4 | 356 | + 65 | 4225 | + .73 |
| 四 | 522 | 603.4 | 95.5 | 579 | — 57 | 32'9 | — .64 |
| 五 | 433 | 604 2 | 77.0 | 465 | — 32 | 102 | — .36 |
| 六 | 894 | 605.0 | 161.7 | 930 | — 86 | 7369 | — .96 |
| 七 | 499 | 605.8 | 82.8 | 503 | — 4 | 16 | — .04 |
| 八 | 479 | 606.7 | 94.9 | 577 | — 98 | 9604 | —1 10 |
| 九 | 702 | 607.5 | 115.0 | 699 | + 3 | 9 | + 03 |
| 十 | 773 | 608 3 | 130.0 | 790 | — 17 | 289 | — .19 |
| 十一 | 740 | 609.1 | 121.4 | 755 | — 15 | 223 | — .17 |
| 十二 | 456 | 610.0 | 84.8 | 619 | —163 | 26569 | —1.83 |
| 1909一 | 534 | 610.8 | 87.2 | 532 | + 2 | 4 | + .02 |
| 二 | 531 | 611.6 | 85.3 | 520 | + 1' | 121 | + .12 |
| 三 | 353 | 612.4 | 59.4 | 361 | — 8 | 64 | — .09 |
| 四 | 583 | 613.3 | 95.5 | 538 | — 5 | 25 | — .06 |
| 五 | 448 | 614.1 | 77.0 | 473 | — 25 | 625 | — 28 |
| 六 | 958 | 614.9 | 161 7 | 996 | — 38 | 1444 | — .43 |
| 七 | 477 | 615.7 | 82.8 | 511 | — 34 | 115 | — .33 |
| 八 | 546 | 616.6 | 94 9 | 586 | — 40 | 1600 | — .45 |
| 九 | 692 | 617.4 | 115.0 | 710 | — 18 | 324 | — .20 |
| 十 | 769 | 618.2 | 130.0 | 803 | — 34 | 1156 | — .38 |
| 十一 | 779 | 619.0 | 124.4 | 768 | + 11 | 121 | + .12 |
| 十二 | 516 | 619.8 | 84.8 | 52. | — 21 | 441 | — .24 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1910一 | 539 | 620.7 | 87.2 | 540 | − 1 | 1 | − .01 |
| 二 | 402 | 621.5 | 85.3 | 529 | −127 | 16129 | −1.42 |
| 三 | 377 | 622.3 | 59.4 | 367 | + 10 | 100 | + .11 |
| 四 | 573 | 623.1 | 95.5 | 598 | − 25 | 625 | − .28 |
| 五 | 397 | 624.0 | 77.0 | 480 | − 83 | 6889 | − .93 |
| 六 | 950 | 624.8 | 161.7 | 1013 | − 63 | 3969 | − .70 |
| 七 | 498 | 625.6 | 82.8 | 520 | − 22 | 484 | − .25 |
| 八 | 487 | 626.4 | 94.9 | 595 | −108 | 11664 | −1.21 |
| 九 | 683 | 627.3 | 115.0 | 721 | − 38 | 1444 | − .43 |
| 十 | 667 | 628.1 | 130.0 | 816 | −149 | 22201 | −1.67 |
| 十一 | 649 | 628.9 | 124.4 | 780 | −131 | 17161 | −1.47 |
| 十二 | 395 | 629.7 | 84.8 | 536 | −141 | 19881 | −1.58 |
| 1911一 | 640 | 630.6 | 87.2 | 549 | + 91 | 8281 | −1.02 |
| 二 | 628 | 631.4 | 85.3 | 536 | + 92 | 8464 | +1.03 |
| 三 | 393 | 632.2 | 59.4 | 373 | + 20 | 400 | + .22 |
| 四 | 622 | 633.0 | 95.5 | 608 | + 14 | 196 | + .16 |
| 五 | 440 | 633.9 | 77.0 | 488 | − 48 | 2304 | − .54 |
| 六 | 1080 | 634.7 | 161.7 | 1029 | + 51 | 2601 | + .57 |
| 七 | 596 | 635.5 | 82.8 | 528 | + 68 | 4624 | + .76 |
| 八 | 585 | 636.3 | 94.9 | 604 | − 19 | 361 | − .21 |
| 九 | 781 | 637.2 | 115.0 | 733 | + 48 | 2304 | + .54 |
| 十 | 892 | 638.0 | 130.0 | 829 | + 63 | 3969 | + .70 |
| 十一 | 856 | 638.8 | 124.4 | 792 | + 64 | 4096 | + .72 |
| 十二 | 598 | 639.6 | 84.8 | 544 | + 54 | 2961 | + .60 |
| 1912一 | 530 | 640.4 | 87.2 | 557 | − 27 | 729 | − .30 |
| 二 | 545 | 641.3 | 85.3 | 545 | 0 | 0 | 0 |
| 三 | 315 | 642.1 | 59.4 | 379 | − 64 | 4096 | − .72 |
| 四 | 658 | 642.9 | 95.5 | 617 | + 41 | 1681 | + .46 |
| 五 | 467 | 643.7 | 77.0 | 496 | − 29 | 841 | − .32 |
| 六 | 981 | 644.6 | 161.7 | 1045 | − 64 | 4096 | − .72 |
| 七 | 524 | 645.4 | 82.8 | 535 | − 11 | 121 | − .12 |
| 八 | 590 | 646.2 | 94.9 | 614 | − 24 | 576 | − .27 |
| 九 | 792 | 647.0 | 115.0 | 744 | + 48 | 2304 | + .54 |
| 十 | 972 | 647.9 | 130.0 | 842 | +130 | 16900 | +1.46 |
| 十一 | 830 | 648.7 | 124.4 | 805 | + 25 | 625 | + .28 |
| 十二 | 681 | 649.5 | 84.8 | 553 | +128 | 16384 | +1.43 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1913一 | 596 | 650.3 | 87.2 | 566 | + 30 | 900 | + .34 |
| 二 | 525 | 651.2 | 85.3 | 553 | − 28 | 784 | − .31 |
| 三 | 479 | 652.0 | 59.4 | 385 | + 94 | 8836 | +1.05 |
| 四 | 988 | 652.8 | 95.5 | 627 | + 61 | 3721 | + .68 |
| 五 | 537 | 653.6 | 77.0 | 504 | + 33 | 1089 | + .37 |
| 六 | 1110 | 654.5 | 161.7 | 1061 | + 49 | 2401 | + .55 |
| 七 | 575 | 655.3 | 82.6 | 544 | + 31 | 961 | + .35 |
| 八 | 637 | 656.1 | 94.9 | 623 | + 14 | 196 | + .16 |
| 九 | 780 | 656.9 | 115.0 | 756 | + 24 | 576 | + .27 |
| 十 | 854 | 657.8 | 130.0 | 855 | − 1 | 1 | − .01 |
| 十一 | 836 | 658.6 | 124.4 | 817 | + 19 | 361 | + .21 |
| 十二 | 503 | 659.4 | 84.8 | 560 | − 57 | 3249 | − .64 |
| 1914一 | 634 | 660.2 | 87.2 | 574 | + 60 | 3600 | + .67 |
| 二 | 677 | 661.0 | 85.2 | 5 2 | +115 | 13225 | +1.30 |
| 三 | 452 | 661.9 | 59.4 | 391 | + 61 | 3721 | + .68 |
| 四 | 702 | 662.7 | 95.5 | 636 | + 66 | 4356 | + .74 |
| 五 | 611 | 663.5 | 77.6 | 511 | +100 | 10000 | +1.12 |
| 六 | 1188 | 664.3 | 161.7 | 1079 | +112 | 12544 | +1.29 |
| 七 | 570 | 665.2 | 82.8 | 52 | + 18 | 324 | + .20 |
| 八 | 840 | 666.0 | 94.9 | 633 | +207 | 42849 | +2.32 |
| 九 | 756 | 666.8 | 115.0 | 76 | − 11 | 121 | − .12 |
| 十 | 888 | 667.6 | 130.0 | 868 | + 20 | 400 | + .22 |
| 十一 | 849 | 668.5 | 124.4 | 830 | + 19 | 361 | + .21 |
| 十二 | 557 | 669.3 | 84.8 | 569 | − 12 | 144 | − .13 |
| 1915一 | 664 | 670.1 | 87.2 | 583 | + 81 | 6561 | + .91 |
| 二 | 581 | 670.9 | 85.3 | 570 | + 11 | 121 | + .12 |
| 三 | 357 | 671.8 | 59.4 | 396 | − 39 | 1521 | − .44 |
| 四 | 682 | 672.6 | 95.5 | 646 | + 36 | 1296 | + .30 |
| 五 | 626 | 673.4 | 77.0 | 518 | +108 | 11664 | +1.21 |
| 六 | 1078 | 674.2 | 161.7 | 1092 | − 14 | 196 | − .16 |
| 七 | 525 | 675.1 | 82.8 | 560 | − 35 | 1225 | − .39 |
| 八 | 637 | 675.9 | 94.9 | 642 | − 5 | 25 | − .06 |
| 九 | 758 | 676.7 | 115.0 | 779 | − 21 | 441 | − .24 |
| 十 | 924 | 677.5 | 130.0 | 881 | + 43 | 1849 | + .48 |
| 十一 | 790 | 678.4 | 124.4 | 841 | − 51 | 2601 | − .57 |
| 十二 | 509 | 679.2 | 84.8 | 577 | − 68 | 4624 | − .76 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1916一 | 610 | 68 .0 | 87.2 | 592 | +18 | 324 | +.20 |
| 二 | 644 | 681.8 | 85.3 | 579 | +65 | 4225 | +.73 |
| 三 | 471 | 681.6 | 59.4 | 402 | +69 | 4761 | +.77 |
| 四 | 614 | 682.5 | 95.5 | 656 | −42 | 1764 | −.47 |
| 五 | 541 | 683.3 | 77.0 | 526 | +15 | 225 | +.17 |
| 六 | 1174 | 684.1 | 161.7 | 1108 | +66 | 4356 | +.74 |
| 七 | 676 | 684.9 | 82.8 | 569 | +107 | 11449 | +1.20 |
| 八 | 562 | 685.8 | 94.9 | 652 | −90 | 8100 | −1.01 |
| 九 | 836 | 689.6 | 115.0 | 790 | +46 | 2116 | +.52 |
| 十 | 929 | 687.4 | 130.0 | 893 | +36 | 1296 | +.40 |
| 十一 | 870 | 688.2 | 124.4 | 853 | +17 | 289 | +.19 |
| 十二 | 624 | 689.1 | 84.8 | 586 | +38 | 1444 | +.43 |
| 1917一 | 640 | 689.9 | 87.2 | 600 | +40 | 1600 | +.45 |
| 二 | 620 | 690.7 | 85.3 | 587 | +33 | 1089 | +.37 |
| 三 | 398 | 691.5 | 59.4 | 408 | −10 | 100 | −.11 |
| 四 | 894 | 69?.4 | 95.5 | 664 | +230 | 52900 | +2.58 |
| 五 | 695 | 693.2 | 77.0 | 534 | +161 | 25921 | +1.80 |
| 六 | 1300 | 694.0 | 161.7 | 1124 | +176 | 30276 | +1.97 |
| 七 | 769 | 694.8 | 82.8 | 577 | +192 | 36864 | +2.15 |
| 八 | 684 | 695.7 | 94.9 | 661 | +23 | 529 | +.26 |
| 九 | 744 | 696.5 | 115.0 | 802 | −58 | 3364 | −.95 |
| 十 | 869 | 697.3 | 130.0 | 906 | −37 | 1369 | −.41 |
| 十一 | 884 | 698.1 | 124.4 | 866 | +18 | 324 | +.20 |
| 十二 | 694 | 699.0 | 84.8 | 594 | +100 | 10000 | +1.12 |
| 1918一 | 598 | 699.8 | 87.2 | 609 | −11 | 121 | −.12 |
| 二 | 495 | 700.6 | 85.3 | 596 | −101 | 10201 | −1.13 |
| 三 | 415 | 01.4 | 59.4 | 414 | +1 | 1 | +.01 |
| 四 | 678 | 702.2 | 95.5 | 674 | +4 | 16 | +.04 |
| 五 | 570 | 703.1 | 77.0 | 541 | +29 | 841 | +.32 |
| 六 | 1059 | 703.9 | 161.7 | 1140 | −81 | 6561 | −.91 |
| 七 | 599 | 704.7 | 82.8 | 585 | +14 | 196 | +.16 |
| 八 | 712 | 705.5 | 94.9 | 671 | +41 | 1681 | +.46 |
| 九 | 730 | .06.4 | 115.0 | 812 | −82 | 19 4 | −.92 |
| 十 | 568 | 707.1 | 130.0 | 9.9 | −351 | 123201 | −3.83 |
| 十一 | 545 | 708.0 | 1 4.4 | 878 | −333 | 110889 | −3.73 |
| 十二 | 426 | 708.8 | 84.8 | 693 | −117 | 31329 | −1.98 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1919— | 476 | 709.7 | 87.2 | 618 | −142 | 20164 | −1.59 |
| 二 | 509 | 710.5 | 85.3 | 604 | − 95 | 9025 | −1.06 |
| 三 | 432 | 711.3 | 59.4 | 419 | + 13 | 169 | + .15 |
| 四 | 585 | 712.1 | 95.5 | 684 | − 99 | 9801 | −1.11 |
| 五 | 544 | 713.0 | 77.0 | 549 | − 5 | 25 | − .06 |
| 六 | 1163 | 713.8 | 161.7 | 1157 | + 6 | 36 | + .07 |
| 七 | 571 | 714.6 | 82.8 | 593 | − 22 | 484 | − .25 |
| 八 | 691 | 715.4 | 94.9 | 679 | + 12 | 144 | − .13 |
| 九 | 743 | 716.3 | 115.0 | 823 | − 80 | 6400 | − .90 |
| 十 | 922 | 717.1 | 130.0 | 932 | − 10 | 100 | − .11 |
| 十一 | 883 | 717.9 | 124.4 | 890 | − 7 | 49 | − .08 |
| 十二 | 603 | 718.7 | 84.8 | 611 | − 8 | 64 | − .09 |
| 1920— | 622 | 719.6 | 87.2 | 626 | − 4 | 16 | − .04 |
| 二 | 576 | 720.4 | 85.3 | 612 | − 36 | 1293 | − .40 |
| 三 | 352 | 721.2 | 59.4 | 425 | − 73 | 5329 | − .82 |
| 四 | 782 | 722.0 | 95.5 | 693 | + 89 | 7921 | +1.00 |
| 五 | 520 | 722.8 | 77.0 | 557 | − 37 | 1369 | − .41 |
| 六 | 1329 | 723.7 | 161.7 | 1173 | +156 | 24336 | +1.74 |
| 七 | 592 | 724.5 | 82.8 | 602 | − 10 | 100 | − .11 |
| 八 | 609 | 725.3 | 94.9 | 689 | − 80 | 6400 | − .90 |
| 九 | 856 | 726.1 | 115.0 | 835 | + 21 | 441 | + .24 |
| 十 | 941 | 727.0 | 130.0 | 945 | − 4 | 16 | − .04 |
| 十一 | 816 | 727.8 | 124.4 | 903 | − 87 | 7569 | − .97 |
| 十二 | 550 | 728.6 | 84.8 | 620 | − 70 | 4900 | − .78 |
| | | | | | | 2,009,9.0 | |

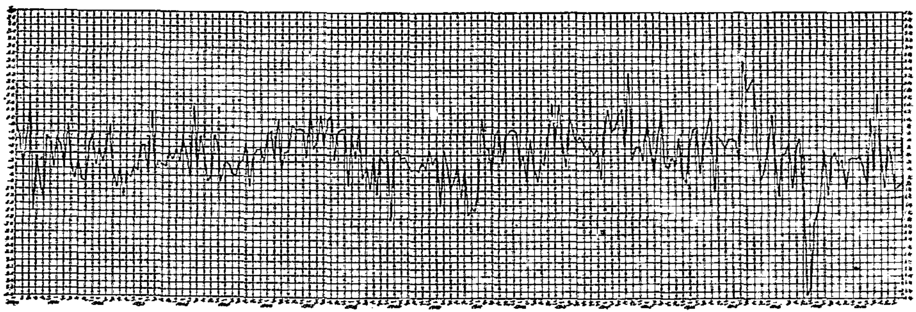
## 第二節 除消法

由時間數列中取消常態數值以發現循環變動之另一方法爲以趨勢數值除實際數值，稱爲除消法 (The Division Method)，或稱爲實際對趨勢法 (The Actual-to-Trend Method)，仍依年份資料與月份資料分別言之。

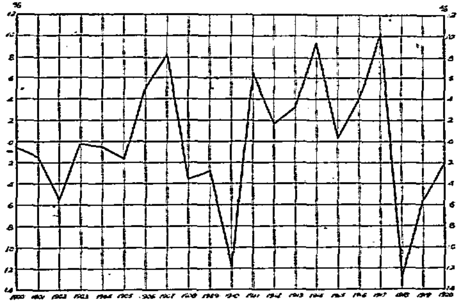
一、 年份資料 時間數列如係年份資料，以趨勢除實際所得之結果，用百分比表示之，並再以100爲標準而求百分比離差(Per Cent Deviations)，此百分比離差卽係循環數值，現以上章所用年份資料爲例：

(第187表) 波斯頓城結婚數目循環值計算表

| 年份 | $A$ | $T$ | $R=\frac{A}{T}$ | $y=R-100$ |
|---|---|---|---|---|
| 1900 | 521 | 524 | 99.4 | − 0.6 |
| 1901 | 526 | 534 | 98.5 | − 1.5 |
| 1902 | 514 | 544 | 94.5 | − 5.5 |
| 1903 | 553 | 554 | 99.8 | − 0.2 |
| 1904 | 561 | 564 | 99.5 | − 0.5 |
| 1905 | 565 | 574 | 98.4 | − 1.6 |
| 1906 | 615 | 585 | 105.1 | + 5.1 |
| 1907 | 643 | 595 | 108.1 | + 8.1 |
| 1908 | 583 | 605 | 96.4 | − 3.6 |
| 1909 | 598 | 615 | 97.2 | − 2.8 |
| 1910 | 551 | 625 | 88.2 | −11.8 |
| 1911 | 676 | 635 | 106.5 | + 6.5 |
| 1912 | 657 | 646 | 101.7 | + 1.7 |
| 1913 | 677 | 656 | 103.2 | + 3.2 |
| 1914 | 727 | 666 | 109.2 | + 9.2 |
| 1915 | 678 | 676 | 100.3 | + 0.3 |
| 1916 | 713 | 686 | 103.9 | + 3.9 |
| 1917 | 766 | 696 | 110.1 | +10.1 |
| 1918 | 616 | 707 | 87.1 | −12.9 |
| 1919 | 677 | 717 | 94.4 | − 5.6 |
| 1920 | 712 | 727 | 97.9 | − 2.1 |



(第47圖)、波斯頓城結婚數目月份循環圖



(第48圖)　波斯頓城結婚數目年份循環數值圖

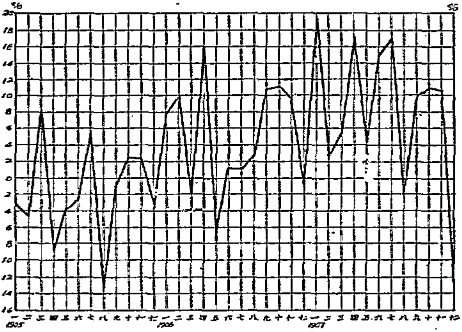
二、　月份資料　所分析者如爲月份資料，其實際數仍以 $A$ 代表之，趨勢數值仍以 $T$ 代表之，至于各月份季節指數則以 $S$ 代表之。循環數值計算方法，係以各月份趨勢值除同月份實際數，並以所得之商數百分比由同月份之季節指數求離差。此項離差代表由時間數列中消除長期趨勢及季節變動作用後之剩餘數值，即代表純粹循環變動者。現由上數章所分析之波斯頓城結婚數目月份資料中摘取 1905－1907 三年計三十六月份之資料爲例，藉資示範：

（第188表） 波斯頓城結婚數目循環數值計算表

| 年份 | A | T | $\frac{A}{T}\times 100$ | S | R−S |
|---|---|---|---|---|---|
| 1905 一 | 480 | 571.2 | 84.03 | 87.2 | − 3.17 |
| 二 | 462 | 572.1 | 80.76 | 85.3 | − 4.54 |
| 三 | 387 | 572.9 | 67.55 | 59.4 | + 8.15 |
| 四 | 497 | 573.7 | 86.63 | 95.5 | − 8.87 |
| 五 | 421 | 574.5 | 73.28 | 77.0 | − 3.72 |
| 六 | 916 | 575.4 | 159.19 | 161.7 | − 2.51 |
| 七 | 508 | 576.2 | 88.16 | 82.8 | + 5.36 |
| 八 | 471 | 577.0 | 81.63 | 94.9 | −13.27 |
| 九 | 659 | 577.8 | 114.05 | 115.0 | − 0.95 |
| 十 | 766 | 578.6 | 132.39 | 130.0 | + 2.39 |
| 十一 | 734 | 579.5 | 126.66 | 124.4 | + 2.26 |
| 十二 | 474 | 580.3 | 81.68 | 84.8 | − 3.12 |
| 1906 一 | 553 | 581.1 | 95.16 | 87.2 | + 7.96 |
| 二 | 554 | 581.9 | 95.21 | 85.3 | + 9.9 |
| 三 | 335 | 582.8 | 57.48 | 59.4 | − 1.92 |
| 四 | 648 | 583.6 | 111.03 | 95.5 | + 5.53 |
| 五 | 412 | 584.4 | 70.50 | 77.0 | − 6.50 |
| 六 | 954 | 585.2 | 163.02 | 161.7 | + 1.32 |
| 七 | 492 | 586.1 | 83.94 | 82.8 | + 1.14 |
| 八 | 573 | 586.9 | 97.63 | 94.9 | + 2.73 |
| 九 | 739 | 587.7 | 125.74 | 115.0 | +10.74 |
| 十 | 830 | 588.5 | 141.04 | 130.0 | + 1.04 |
| 十一 | 790 | 589.4 | 131.03 | 124.4 | + 9.63 |
| 十二 | 496 | 590.2 | 84.04 | 84.8 | − 0.74 |
| 1907 一 | 630 | 591.0 | 106.60 | 87.2 | +19.40 |
| 二 | 517 | 591.8 | 87.36 | 85.3 | + 2.06 |
| 三 | 384 | 592.7 | 64.79 | 59.4 | + 5.39 |
| 四 | 668 | 593.5 | 112.55 | 95.5 | +17.05 |
| 五 | 483 | 594.3 | 81.27 | 77.0 | + 4.27 |
| 六 | 1052 | 595.1 | 176.78 | 161.7 | + 5.08 |
| 七 | 594 | 596.0 | 99.66 | 82.8 | +16.86 |

續　上　表

| | | | | | |
|---|---|---|---|---|---|
| 八 | 554 | 596.8 | 92.83 | 94.9 | －2.07 |
| 九 | 745 | 597.6 | 124.67 | 115.0 | ＋9.67 |
| 十 | 842 | 598.4 | 140.71 | 130.0 | ＋10.71 |
| 十一 | 808 | 599.2 | 134.85 | 124.4 | ＋10.45 |
| 十二 | 433 | 600.1 | 72.15 | 84.8 | －12.65 |



(第49圖)　波斯頓城結婚數目循環值圖

## 第三節　循環變動間之相關

在次數分配中所用以分析相關之技術，亦可用之于時間數列。惟在時間數列中相關問題較在次數分配中爲複雜。次數分配之內容比較單純，資料分類以數量大小爲唯一標準。在時間數列中，因包括長期趨勢，季節變動及循環變動中之三因素，分析甲乙兩時間數列間相關以前，必先將每一

數列中之三因素分別獨立加以研究。就中兩數列之長期趨勢間或季節變動間之共同上升或共同下降不能卽認爲其有相關存在。在時間數列中之相關，乃專指兩者循環變動間之相關而言，稱爲循環相關 Cyclical Correlation)。因分析之目的不同 循環相關又分二種：

一、 相互關係 與在次數分配中所分析去相同。循環變動間有因果關係與相互關係者甚多，可應用相關技術以分析之。試舉數列：在時間數列中，結婚或生育人數與就業人數有關 就業人數與物價高低有關，出口入口與出超入超有關，甲乙兩種物價之變動有關，甲乙兩種商品之生產或消費數量有關。凡彼此有連帶關係之數列，均可于求得長期趨勢之後，再求離差或循環變動，進而由離差乘積求相關係數，以代表循環變動間相關之程度。以年份資料爲例，有甲乙兩時間數列，分別以 $X$，$Y$ 代表之。均可求得趨勢數值，並以趨勢數值爲標準以求絕對循環數值 $x$ 與 $y$，及相對循環數值 $x/\sigma_x$ 與 $y/\sigma_y$。得到兩者之相對循環數值各變項後，卽可求各對變項之乘積及其總和，最後以變項數目除相對離差乘積總和，卽得相關係數，以符號表示之，公式如下：

$$r=\frac{\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}}{N}$$

依上列公式所得相關係數 $r$，仍爲一抽象之小數，代表兩數列循環變動間相關之程度者。與次數分配中相關係數唯一不同之點，卽在次數分配中係以算術平均數爲標準以計算離差，進而由離差乘積以求相關係數。現在時間數列中，則以長期趨勢爲標準以計算離差，進而再由離差乘積求相關係數。長期趨勢究採用直綫或曲綫公式，頗有主觀成分介乎其間。非如在次數分配中求相關係數時所根據者，爲算術平均數之絕不因人而異。

二、 前後關係 上列所分析者爲循環變動間之相互關係，與在次數分配

中所分析者相同。相關係數之另一功用，在依以決定兩種現象循環變動發生之順序。甲乙兩現象之循環變動，在時間上可完全符合，亦可有先後，甲種變動發生在前，乙種變動發生隨後。其變動在前者稱爲引前 (Time Lead)，其變動在後者稱爲落後 (Time Lag)。此種引前與落後之關係，稱爲時間相關 (Temporal Correlation)，或稱爲落後相關 (Lagging Correlation)。落後相關是否存在，取甲乙兩組月份資料作入同一圖形中加以觀察，即可發現，但此法粗而不精，只作爲初步之導，欲確定甲乙兩現象循環變動在時間上究相差幾月，則須根據相關係數。此中所用相關係數之公式，與上文中相同：

$$r=\frac{\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}}{N}$$

假定甲數列爲二十年計二百四十月份股票價格，乙數列爲同期二十年二百四十月份工人就業數目。由圖形中可見股票價格引前而就業人數落後，兩者前後相距究爲幾月，則須用相關係數之數值以定之。初可假定股票行情及就業人數之循環變動，同時發生，而無前後，則求離差乘積時自以同月份之兩離差相乘，其積差爲 $\frac{x_1}{\sigma_x}\cdot\frac{y_1}{\sigma_y}$，$\frac{x_2}{\sigma_x}\cdot\frac{y_2}{\sigma_y}$，$\frac{x_3}{\sigma_x}\cdot\frac{y_3}{\sigma_y}$ ……，積差總和爲 $\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}$，成對變項之數即月數爲 240，故相關係數公式：

$$r=\frac{\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}}{240}$$

假定股票價格引前一月，或工人就業落後一月，則求積差時，係以上一月股票價格離差與本月工人就業數目離差相乘，以本月股票離差與下

月就業離差相乘。以符號表示之，各積差為：$\frac{x_1}{\sigma_x}\cdot\frac{y_2}{\sigma_y}$，$\frac{x_2}{\sigma_x}\cdot\frac{y_3}{\sigma_y}$，$\frac{x_3}{\sigma_x}\cdot\frac{y_4}{\sigma_y}$，……積差總和為 $\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}$，但已比上式中減少一項，即月數為 239，相關係數公式為：

$$r=\frac{\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}}{239}$$

同理，如工人就業落後兩月，則積差為：$\frac{x_1}{\sigma_x}\cdot\frac{y_3}{\sigma_y}$，$\frac{x_2}{\sigma_x}\cdot\frac{y_4}{\sigma_y}$，$\frac{x_3}{\sigma_x}\cdot\frac{y_5}{\sigma_y}$，……，月份數目減少二項，即 238，相關係數公式：

$$r=\frac{\Sigma\frac{x}{\sigma_x}\cdot\frac{y}{\sigma_y}}{238}$$

前後相差一月，二月之相關係數，可如上法求出，相差三月，四月，五月之相關係數，自仍可同法求出。由所得各結果觀之，如依同時變動所得之係數最小，依相差一月所得之係數加大，依相差兩月所得之係數更大，而依相差三月所得之係數則減小；依相差四月、五月所得之係數更小。因可斷定此兩現象循環變動在時間上相差二月；股票價格上升或下落在前，兩月後工人就業則隨之而增加或減少。由股票價格之變動以推測工人就業之變動，稱為商情預測(Business Forecasting)。此種分析技術在經濟統計中之重要自甚明顯。商情預測所以可能及可靠，在乎各種經濟現象中間有基本關係，假如此基本關係一變，商情預測亦必隨而修正也。

## 習　題

一、 試就一九一九至一九三七年各月份美國生鐵產量資料，由最小二乘

方法所得之月份趨勢值及環比中位數法所得之季節指數，求各月標準值。

二、 試由各月生鐵實際產量減掉其標準值，以求絕對循環變動數值。

三、 試求標準差並以之爲各月公約數，以求相對循環變動數值。

四、 試依上所得結果作一循環數值圖。

# 第六編

# 選樣問題

# 第二十二章　常態曲綫

## 第一節　機率

統計資料類別至多，但各種資料共同特點之一，即其次數分配，每合乎誤差常態定律(Normal Law of Errors)。製成次數圖時，其次數曲綫之形式，每近于常態曲綫(Normal Curve)。昔日學而認爲常態曲綫係所有統計資料分配之定律，但經近年研究之結果，統計資料之分配雖以常態分配爲最基未定律，但其不合于常態之資料爲數亦不少。統計分析所得之結果，如平均數、標準差、迴歸綫及估計標準誤差，無一不根據常態分配之假定而存在，足見常態分配在統計方法中之重要。

常態分配之研究自機率論(Probability)開始。任一事件(Event)，可發生 $n$ 種結果。就中以 $a$ 代表成功或合意之結果(Success or Faverable Outcome)，以 $b$ 代表失敗或不合意之結果(Failure or Untaverable Outcome)，則 $a+b=n$，即成功之方式加失敗之方式，等于所有方式。成功之機率以 $p$ 代表之，則 $p=a/n$。敗之機率以 $q$ 代表之，則 $q=b/n$，

$$\therefore \quad p+q=\frac{a}{n}+\frac{b}{n}=1,$$

"1"代表必然性。此事件之發生祇有成敗兩種結果，非成功卽失敗，故成敗機率之和必爲性。由上可知，機率爲一分數，其分母代表一切結果之總數，其分子則代表成功或失敗之結果之總數。試舉例言之，如投擲銅錢一枚，投擲之結果，不外正面(Head)及反面(Tail)二種結果，卽 $n=2$，$a=1$ $b=1$，$a+b=n$，成功卽正面之機率，以 $p$ 表示之，$p=a/n=1/2$。失敗

即反面之機率以$q$表示之，$q=b/n=1/2$，成功與失敗，或正面與反面機率之和爲$p+q$爲$\frac{1}{2}+\frac{1}{2}=1$，爲必然性，表示投擲銅錢之結果，非正面卽反面，絕無第三種可能性。此種結論，自不限于成功與失敗機會均等之事件。如以投擲骰子爲例，骰子有六面，分別爲1，2，3，4，5，6各點，則可能發生之結果爲$n$，$n=6$，發生1，2，3，4，5，6各點之機率各爲$\frac{1}{6}$。如認定發生1點爲成功，其他各點俱爲失敗，則成功之機率爲$\frac{1}{6}$失敗之機率爲$\frac{5}{6}$，成功與失敗機率之和爲：

$$p+q=\frac{1}{6}+\frac{5}{6}=1。$$

又以撲克牌52張爲例，共有黑桃(Spade)，紅心(Heart)，紅方(Diamond)，及梅花(Club)四種。每種計13張。其中之一張爲A牌(Ace)。如以抽得任一張黑桃爲成功，抽得其他爲失敗，則成功之機率爲：$p=\frac{13}{42}=\frac{1}{4}$；失敗之機率爲$q=\frac{39}{52}=\frac{3}{4}$。如只認定抽得黑桃中之A牌爲成功，其他均爲失敗，則成功之機率爲$\frac{1}{52}$，失敗之機率爲$=\frac{51}{52}$。最後以買獎券爲例，如獎券共計10,000張，各張得獎之機會均等，則買得一張中彩之機率爲$p=1/10,000$，失敗之機是爲：$q=9,999/10,000$。

以上爲機率之基本觀念。有時須相加或相乘。先就機率相加而言，以投擲骰子爲例，如認定1，2點兩面均爲成功，其他3，4，5，6點四面爲失敗，則成功之機率爲$\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$。以抽得撲克牌爲例，如認定抽得黑桃及紅心俱爲成功，抽得其他二種爲失敗，則成功之機率爲$p=\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$。如以抽得黑桃中之A牌或紅心中之A牌爲成功，其他均爲失敗，則成功

之機率爲：

$$p=\frac{1}{52}+\frac{1}{52}=\frac{1}{26}$$

以在壹萬張彩票中買五張爲例，中彩之機率爲 $p=5/10,000$。

至于機率之所以相乘，則可依二獨立事件(Two Independent Events) 之機率而言。以銅錢二枚，同時投擲一次，稱爲複合事件(Compound Event)，就中甲銅錢之發生正面或反面，與乙銅錢毫無關係，絕對彼此獨立，而不相依。甲乙二銅錢同時發生正面之機率實爲二事件成功機率之乘積，$p_1 \cdot p_2=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$。即以投擲甲乙二骰子爲例，如認定甲，乙俱發生 1 點爲成功，則成功之機率爲$p_1\cdot p_2=\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$。但有時須相乘與相加同時並用，方可得一事件之機率，以同時投擲骰子二枚爲例，如認定二者聯合點數爲 5 時係成功，則成功之機率算 如下：

| | 甲骰 | 乙骰 | 聯合機率 |
|---|---|---|---|
| | 1 | 4 | $\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$ |
| | 2 | 3 | $\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$ |
| | 3 | 2 | $\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$ |
| + | 4 | 1 | $\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}$ |

甲乙二骰聯合發生 5 點之機率爲：$p=\frac{4}{36}=\frac{1}{9}$

## 第二節　二項式分配

上節已言及欲知一事件之機率有時須相加，有時須相乘，有時則乘與加並用。現擬求出公式以代表此理。仍以投擲銅錢爲例而觀察其成功與失敗之一切機率。如投擲銅元一枚，可能發生之結果不外二種，即反面與正面。發生反面以$T$表示之，發生正面以$H$表示之，則$H$與$T$之機率各爲$\frac{1}{2}$，即$p=q=\frac{1}{2}$，$q+p=1$。

如投擲$a$，$b$銅元二枚，則失敗與成功，或反面與正面一切結果之機率如下：

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *a* | *b* | *a* | *b* | *a* | *b* | *b* | *a* |
| T | T | T | H | H | T | H | H |

一枚銅錢失敗與成功之機率，爲二者機率之和，即$q+p=\frac{1}{2}+\frac{1}{2}=1$，二枚銅錢失敗與成功之機率，則爲上列機率和之乘積，其公式爲：

$$(q+p)(q+p)=(q+p)^2=\left(\frac{1}{2}+\frac{1}{2}\right)\left(\frac{1}{2}+\frac{1}{2}\right)=\left(\frac{1}{2}+\frac{1}{2}\right)^2$$

$$=\frac{1}{4}+\frac{1}{2}+\frac{1}{4},$$

其意即投擲$a$,$b$二銅錢時，可能發生之結果計三種：二者俱出反面之機率爲$\frac{1}{4}$，一反一正之機率爲$\frac{1}{2}$，二者俱出正面之機率爲$\frac{1}{4}$。如投擲$a$,$b$,$c$三枚銅錢，則一切可能之結果如下：

| | | | |
|---|---|---|---|
| *a b c* | *a b c* | *a b c* | *a b c* |
| TTT | TTH | THH | HHH |

一切可能結果之機率公式爲：

$$(q+p)^3=\left(\frac{1}{2}+\frac{1}{2}\right)^3=\left(\frac{1}{2}\right)^3+3\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)+3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^3=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8}。$$

其意卽投擲 $a,b,c$ 三枚錢時，一切可能之結果共計四種，卽三枚俱出反面之機率爲 $\frac{1}{8}$，二枚出反面一枚出正面之機率爲 $\frac{3}{8}$，一枚出反面，二枚出正面之機率爲 $\frac{3}{8}$，三枚同時出正面之機率爲 $\frac{1}{8}$。

如投擲四枚銅錢，則一切可 結果如下：

| | | | | |
|---|---|---|---|---|
| *a b c d* | *a b c d* | *a b c d* | *a b c d* | *a b c d* |
| TTTT | TTTH | TTHH | THHH | HHHH |

一切可能之結果機率公式爲：

$$(q+p)^4=\left(\frac{1}{2}+\frac{1}{2}\right)^4=\left(\frac{1}{2}\right)^4+4\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)+6\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2+4\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3+\left(\frac{1}{2}\right)^4=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}+\frac{1}{16},$$

其意義卽投擲四枚銅錢，一切結果之機率，共計五項：四枚俱出反面之機率爲 $\frac{1}{16}$，三枚出反面一枚出正面之機率爲 $\frac{4}{16}$，二枚出反面，二枚出正面之機率爲 $\frac{6}{16}$，一枚出反面三枚出正面之機率爲 $\frac{4}{16}$，四枚俱出正面之機率爲 $\frac{1}{16}$。

以上四例中，如投擲銅錢枚數爲 1 時，一切結果機率之公式爲 $(q+p)^1$，展開式之項數爲 2。銅元之數爲 2 時，機率公式爲 $(q+p)^2$，展開式之項數爲 3。銅錢之數爲 3 時，機率公式爲 $(q+p)^3$，展開式之項數爲 4。

銅錢之數爲 4 時，機率公式爲$(q+p)^4$，展開式之項數爲 5。銅錢之數爲 $n$ 時，則機率之公式爲$(q+p)^n$，$(q+p)^n$ 之展開式稱爲二項展開式 (Binomial Expansion)，其項數爲 $n+1$，其形式如下：

$$(q+p)^n=q^np^0+nq^{n-1}p^1+\frac{n(n-1)}{1.2}q^{n-2}q^2+\frac{n(n-1)(n-2)}{1\cdot2\cdot3}\cdot$$

$$q^{n-3}p^2+\cdots\cdots\cdots\cdots+q^0\cdot p^n$$

上式卽代表 $n$ 個銅錢投擲一次各種結果之機率。投擲銅錢之數爲 $n$，一切可能結果之數或卽稱爲項數(Number of terms)爲 $n+1$。如投擲銅錢四枚，則代表一切結果之機率公式爲 $(q+p)^4$，展開此公式時共得五項，卽$(n+1)$項。以 $q=\frac{1}{2}$，$p=\frac{1}{2}$，$n=4$，$n+1=5$，四値代入公式並展開之，其各項目如下：

$$\frac{1}{2}+\frac{1}{2}\qquad\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0+4\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1+\frac{4\cdot3}{1\cdot2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2$$

$$+\frac{4\cdot3\cdot2}{1\cdot2\cdot3}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3+\frac{4\cdot3\cdot2\cdot1}{1\cdot2\cdot3\cdot4}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4$$

$$=\frac{1}{16}+\frac{4}{16}+\frac{6}{16}+\frac{4}{16}+\frac{1}{16}$$

上列二項展開式中共計五項，係代表銅錢四枚投擲或試驗一次時，所可能發生各種結果之機率。如以發生正面 H 爲成功，並以 $X$ 代表發生正面或成功之數目。則一切可能發生正面或成功之機率，可製表如下：

發生正面(成功)數目：$X$

銅錢數目，$n=4$

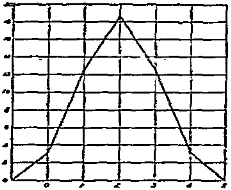
一切結果項數，$n+1=5$

機率：$P$

(第189表) 投擲銅錢發生正面(成功)機率表

| X | P | (1) $f_c$ | (2) $f_c$ |
|---|---|---|---|
| 0 | 1/16 | 1 | 3.125 |
| 1 | 4/16 | 4 | 12.500 |
| 2 | 6/16 | 6 | 18.750 |
| 3 | 4/16 | 4 | 12.500 |
| 4 | 1/16 | 1 | 3.125 |
| 總計 | 1 | 16 | 50.000 |

表中以$P$表示之一欄係指銅錢四枚，試驗一次時各種成功結果之機率。如連續試驗十六次，自須以16乘各種結果之機率。即在十六次試驗之中：四枚銅錢全出反面無一枚出正面者計一次，只一枚出正面者計四次，二枚出正面者計六次，三枚出正面者計四次，四枚俱出正面無一出反面者計一次，合計試驗十六次。試驗一次時各種結果之機率爲分數，試驗十六次時各種結果之機率已化爲次數，稱爲理論次數(Theoretical Frequencies)。以$f_c$表示之，亦列入上表中。就中$X$及$f_c$二欄卽構成二項分配(Binomial Distribution)或常態分配(Normal Distribution)。以銅錢四枚連投十六次之各種結果，現已求出，連投五十次之各種結果自可依同法求出，並依之而作常態曲綫圖如下：



(第50圖) 投擲銅錢四枚計五十次發生正面理論次數常態曲綫圖

上列之理論次數曲綫已爲左右對稱之曲綫。曲綫上所以仍有不規則處，係獨立事件（Independent Events），即銅錢之數太少，即 $n$ 之值太小。如 $n$ 加大，即銅錢加多，則 $(n+1)$ 即組距加多，理論次數分配，必表現爲平滑曲綫(Smooth Curve)。此時之二項式分配方爲與正常態分配。但在多數次數分配中，組距爲數無多，故所得之理論次數曲綫，雖係左右對稱，而不規則之處勢所難免，故二項式分配只構成近似之常態分配耳。

由二項式分配中求算術平均數及標準差，可另推出簡便方法。已知 $n$＝獨立事件數，$q$＝失敗機率，$p$＝成功機率，$q+p=1$，即失敗與成功機率之和爲必然性。二項式 $(q+p)^n$ 係求一切成功或失敗之機率公式。二項展開式各項，

$$q^n p^0+\frac{n}{1}q^{n-1}p^1+\frac{n(n-1)}{1.2}q^{n-2}p^2+\cdots\cdots\cdots\cdots\cdots+q^0p^n,$$

係代表所有各種成功或發生正面之相對次數，其總和仍等于 1，可製一次數表如下：

(第 190 表)二項式分配次數表

| $X$ | $f$ | $fX$ |
|---|---|---|
| 0 | $q^n p^0$ | $0$ |
| 1 | $\frac{n}{1}q^{n-1}p^1$ | $nq^{n-1}p^1$ |
| 2 | $\frac{n(n-1)}{1.2}q^{n-2}p^2$ | $n(n-1)q^{n-2}p^2$ |
| 3 | $\frac{n(n-1)(n-2)}{1.2.3}q^{n-3}p^3$ | $\frac{n(n-1)(n-2)}{1.2}q^{n-3}p^3$ |
| 4 | $\frac{n(n-1)(n-2)(n-3)}{1.2.3.4}q^{n-4}p^4$ | $\frac{n(n-1)(n-2)(n-3)}{1.2.3}q^{n-4}p^4$ |
| ⋮ | ⋮ | ⋮ |
| $n$ | $q^0 p^n$ | $np^n$ |
| 總計 | 1 | $np$ |

續 上 表

| $X$ | $fX^2$ |
|---|---|
| 0 | $0$ |
| 1 | $nq^{n-1}p^1$ |
| 2 | $2n(n-1)q^{n-2}p^2$ |
| 3 | $\frac{3n(n-1)(n-2)}{1.2}q^{n-3}p^3$ |
| 4 | $\frac{4n(n-1)(n-2)(n-3)}{1.2.3}q^{n-4}p^4$ |
| ⋮ | ⋮ |
| $n$ | $n^2p^n$ |
| 總計 | |

由上表 $fX$ 一欄中各項可得

$$\Sigma fX=np\ [q^{n-1}+(n-1)q^{n-2}p^1+\frac{(n-1)(n-2)}{1.2}q^{n-3}p^2$$

$$+\frac{(n-1)(n-2)(n-3)}{1.2.3}q^{n-4}p^3+\cdots\cdots+p^{n-1}]$$

$$=np(q+p)^{n-1}$$

$$\because (q+p)^{n-1}=1$$

$$\therefore \Sigma fX=np(q+p)^{n-1}=np,$$

$$\overline{X}=\frac{np}{\Sigma f}\qquad \because \Sigma f=1,\qquad \therefore \overline{X}=np$$

依上例 $n=4$, $p=\frac{1}{2}$, $\therefore \overline{X}=np=4\frac{1}{2}=2$

卽投擲銅錢四枚時，發生正面或成功之平均數爲 2，卽二枚發正面者代表集中趨勢。

由上表 $fX^2$ 一欄各項可得：

$$\Sigma fX^2=np\,[q^{n-1}+2(n-1)\,q^{n-2}p^1+\frac{3(n-1)(n-2)}{1.2}q^{n-3}p^2$$

$$+\frac{4(n-1)(n-2)(n-3)}{1.2.3}q^{n-4}p^3+\cdots\cdots+np^{n-1}].$$

$$=np[\{q^{n-1}+(n-1)q^{n-2}p^1+\frac{(n-1)(n-2)}{1.2}q^{n-3}p^3$$

$$+\frac{(n-1)(n-2)(n-3)}{1.2.3}q^{n-4}p^3+\cdots+p^{n-1}\}+\{n-1\cdot q^{n-2}p^1$$

$$+\frac{(n-1)(n-2)}{1}q^{n-3}p^2-\frac{(n-1)(n-2)(n-3)}{1.2}q^{n-4}p^3$$

$$+\cdots\cdots+(n-1)p^{n-1}\}]$$

$$=np[(q+p)^{n-1}+p(n-1)(q+p)^{n-2}]$$

$$\because (q+p)^{n-1}=1,\qquad (q+p)^{n-2}=1$$

$$\therefore \Sigma fX^2=np[1+p(n-1)]$$

已知假定原點爲零時，求標準差之簡便法公式如下：

$$\sigma=\sqrt{\frac{\Sigma fX^2}{\Sigma f}-\left(\frac{\Sigma fX}{\Sigma f}\right)^2}$$

$$\Sigma fX^2=np[1+p(n-1)]$$

$$\Sigma fX=np,\qquad \Sigma f=1$$

$$\therefore\quad \sigma^2=np[1+p(n-1)]-(np)^2$$

$$=np+n^2p^2-np^2-n^2p^2$$

$$=np-np^2=np(1-p)=npq$$

$$\therefore\quad \sigma=\sqrt{npq}$$

依上例：　$n=4,\qquad p=q=\frac{1}{2}$

$$\therefore\quad \sigma=\sqrt{4\cdot\tfrac{1}{2}\cdot\tfrac{1}{2}}=\sqrt{1}=1$$

## 第三節　常態曲綫配合法

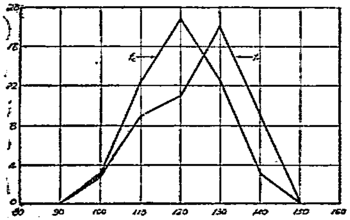
以曲綫圖表示次數分配稱爲次數曲綫圖，次數曲綫所代表者可爲實際次數，在本書之首業已言及，亦可求理論次數。代表理論次數之曲綫稱爲常態曲綫。常態曲綫所代表各理論次數，可由上節所述二項展開式求得。上節所述投擲銅錢試驗中，係以獨立事項或銅錢枚數爲$n$，各種結果數目爲$(n+1)$，試驗次數爲$N$。現以實際次數分配中組距數目爲$(n+1)$，則$n$所代表者自爲組距數目減1，次數總計爲$N$，或$\Sigma f$。如以第一，二編中所用五十男生體重分配爲例，共計5組，次數總計爲50項 $(n+1)=5$，$n=4$，$N=50$。是依二項展開式自可求得各組之理論次數，並可與實際次數作入同一圖中彼此比較：

（第191表）　五十男生體重分配理論次數計算表

| 組　距 | 中值 | 實際次數 | 機率 | 理論次數 |
|---|---|---|---|---|
| 95—105 | 100 | 3 | 1/16 | 3.125 |
| 105—115 | 110 | 9 | 4/16 | 12.500 |
| 115—125 | 120 | 11 | 6/16 | 18.750 |
| 125—135 | 130 | 18 | 4/16 | 12.500 |
| 135—145 | 140 | 9 | 1/16 | 3.125 |
| 總　計 | | 50 | 1 | 50.000 |

上表中之理論次數係依二項式$N(q+p)^n$求得者。$n$代表組距數目減1，次數分配中共計5組，故$n$爲4，$N$代表次數總計，現爲50。故理論次數之公式爲：

$$f_c=50(\tfrac{1}{2}+\tfrac{1}{2})^4$$



(第51圖) 五十男生體重分配次數圖

以上係根據二項式求理論次數與常態曲綫之方法。但如次數分配之組距過多，展開式之項數自隨之加多，計算手續甚繁。理論次數之計算，乃另有簡捷途徑，即由機率公式(Probability Equation)及機率表(Probability Table)以求之。根據機率表以求理論次數為最常用之方法，至由二項展開式所求者，在於解釋理論次數之由來。

$$Y=\frac{N}{\sigma_i\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$$

上列即機率曲綫公式，亦稱為常態曲綫公式。依此公式所求得者為各組距理論次數在常態曲綫上之高度，故稱為縱綫法(The Ordinate Method)。依縱綫法求常態曲綫之步驟有二：

一、 求最高縱綫 所謂最高縱綫(Maximum Ordinate)亦稱為中央縱綫(Central Ordinate)，以$Y_o$代表之。即在次數曲綫圖中，自算術平均數$\overline{X}$所在據點豎起之次數縱綫，此據點之值既為平均數$\overline{X}$，其離差$x$必等于零：即$x=X-\overline{X}=0$，回顧上列機率曲綫全式，如離差$x$為零，則表示e之指數全部必為零。即$-\frac{x^2}{2\sigma^2}=0$，$e^0=1$，故最高縱綫之公式為：

$$Y_o=\frac{N}{\sigma_i\sqrt{2\pi}}$$

在五十男生體重分配中已知：$N=50$，$\overline{X}=124.2$，$\sigma=11.5$，$i=10$，$\sigma_i=1.15$，$\pi=3.1416$，$\sqrt{2\pi}=\sqrt{2\times3.1416}=2.506628$，以各值代入上式，則得最高縱綫如下：

$$Y_o=\frac{50}{1.15\times2.506628}=\frac{50}{2.8826222}=17.34532$$

二、求各組距之次數縱綫　已知最高縱綫 $Y_o$ 爲 17.34532，現化爲相對次數，使 $Y_o$ 等于 1，左右兩側各組距之次數既均少于最高縱綫所代表者，其相對次數用 $Y/Y_o$ 表示之，必均小于 1，卽低于最高縱綫。以 $\overline{X}$ 所在之據點卽以最高縱綫所在之據點爲標準，在 $X$ 軸上任何據點次數縱綫之相對數值，俱爲最高縱綫之小數或百分數。此小數可由常態曲綫縱綫表 (Normal Curve Ordinate Table) 查出：

（第192表）　常態曲綫縱綫表

（各綫高度：$Y/Y_o$）

| $\frac{x}{\sigma}$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 1.00000 | .99995 | .99980 | .99955 | .99920 | .99875 | .9982[illegible] | .99755 | .99685 | .99596 |
| 0.1 | .99501 | .99396 | .99283 | .991 8 | .99025 | .98881 | .98728 | .98565 | .98393 | .98211 |
| 0.2 | .98 20 | .97819 | .97609 | .9739[illegible] | .97161 | .96923 | .96676 | .96420 | .96156 | .95882 |
| 0.3 | .95600 | .95309 | .95010 | .94702 | .94387 | .9 065 | .93723 | .93382 | .93024 | .92677 |
| 0.4 | .92312 | .91399 | .91558 | .91161 | .90774 | .9 371 | .89961 | .89543 | .89119 | .88688 |
| 0.5 | .88250 | .87805 | .87353 | .86336 | .86132 | .85962 | .85488 | .85006 | .84519 | .84060 |
| 0.6 | .83527 | .83023 | .83514 | .82010 | .81 81 | .20957 | .80429 | .79896 | .79459 | .78817 |
| 0.7 | .78270 | .77721 | .77167 | .76610 | .76 48 | .75484 | .74916 | .74342 | .73769 | .73193 |
| 0.8 | .72615 | .72033 | .71448 | .70861 | .70272 | .69621 | .69087 | .68493 | .67896 | .67298 |
| 0.9 | .66689 | .66097 | .65494 | .64891 | .64287 | .63683 | .63077 | .62472 | .61865 | .61259 |

| $\frac{x}{\sigma}$ | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .0 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.0 | .60653 | .60047 | .59440 | .58834 | .58228 | .57623 | .57017 | .56414 | .55810 | .55209 |
| 1.1 | .54607 | .54007 | .53409 | .52812 | .52214 | 51620 | .51027 | .50437 | .49848 | .49260 |
| 1.2 | .48675 | .48092 | .47511 | .46933 | .46357 | .45793 | 45212 | .44644 | .44078 | .43516 |
| 1.3 | .42956 | .42399 | .41845 | .41294 | .40747 | 40202 | .39661 | .39123 | .38580 | .38058 |
| 1.4 | .37531 | .37007 | 36487 | .35971 | 35459 | .34950 | .34445 | .33944 | .33441 | .32954 |
| 1.5 | .32465 | .31980 | .31500 | .31023 | .30550 | .30082 | .29618 | .29158 | .28702 | .28251 |
| 1.6 | .27804 | .27361 | .26923 | .26489 | .26059 | .25634 | .25213 | .24797 | .24385 | .23978 |
| 1.7 | .23575 | .23176 | .22782 | .22392 | .22008 | .21627 | .21251 | .20879 | .20511 | .20148 |
| 1.8 | .19790 | .19436 | .19086 | .18741 | .18400 | .18 64 | .17732 | .17904 | .17081 | .16762 |
| 1.9 | .16448 | 16137 | .15831 | .15530 | .15237 | .14939 | 14650 | .14364 | .14083 | .13806 |
| 2.0 | .13534 | .13265 | .13000 | .12740 | .12483 | .1223 | 11981 | .11737 | .11496 | .11259 |
| 2.1 | .11025 | .10795 | .10570 | .10347 | .10129 | .09914 | .09702 | .09495 | 09290 | .09090 |
| 2.2 | .08892 | .08698 | .08509 | .08320 | .08136 | .07656 | .07778 | .07604 | .07433 | .07265 |
| 2.3 | .07100 | 06939 | .06780 | .06624 | .06471 | .06321 | 06174 | .06029 | .05892 | .05750 |
| 2.4 | .05614 | .05481 | .05350 | .05222 | .05096 | .04973 | .04852 | .04737 | .04618 | .04505 |
| 2.5 | .04394 | .04285 | .04179 | .04074 | .03972 | .03873 | .03775 | .03680 | .03586 | .03494 |
| 2.6 | .03405 | .03317 | .03232 | .03148 | .03066 | .02986 | .02908 | .02831 | .02757 | .02684 |
| 2.7 | .02612 | .02542 | .02474 | 02408 | .02343 | .02280 | .02218 | .02157 | .02098 | .02040 |
| 2.8 | .01984 | .01929 | .01876 | .01823 | .01772 | .01723 | .01674 | .01627 | .01581 | .01536 |
| 2.9 | .01642 | .01449 | .01408 | .01367 | 01328 | .01288 | .01252 | .01215 | .01179 | .01145 |
| 3.0 | .01111 | .00348 | .00598 | .00432 | .003 9 | .00219 | .00153 | .00106 | .00073 | .00050 |
| 4.0 | 00034 | .00022 | .0001 | 00010 | .00006 | .00004 | .00003 | .00002 | .00001 | .00001 |

上列第 192 表中所載係以平均數$\overline{X}$爲標準，左右任何一據點縱綫之相對高度 最左一欄所列則爲表示距離平均數$\overline{X}$所在據點之遠近者。距離遠近自然以離差 $x$ 代表之；而離差則以標準差爲單位，卽除以標準差，所得結果$x/\sigma$用代表之，稱爲標準離差(Normal Deviates)，實卽相對離差以之表示各點離$\overline{X}$之遠近，由表中可摘舉幾項重要數值如下，並加以說明：

（第193表）

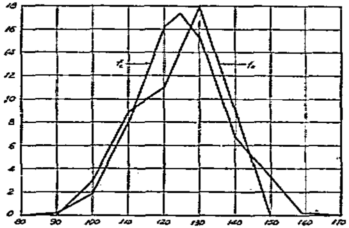
| 標準離差：$x/\sigma$ | 縱綫相對高度：$Y/Y_0$ |
|---|---|
| 0.00 | 1.00000 |
| 0.01 | .99995 |
| 0.10 | .995 1 |
| 1.00 | .60653 |
| 2.00 | .13534 |
| 3.00 | .01111 |
| 4.00 | .00034 |
| 4.09 | .00001 |

上列第193表所列各對數值之意，卽任$X$軸上任何一點，無論在平均數$\bar{X}$所佔據點左或右，其離差爲標準差之零倍時 其縱綫之高度爲1，離差爲標準差百分之一時，其縱綫高度爲 .99995；離差爲標準差十分之一時，其縱綫高度爲 0.99501，離差爲標準差一倍時，其縱綫 度爲 0.60653；離差爲標準二倍時，其縱綫高度爲0.13534；離差爲標準差三倍時，其縱綫高度爲0.01111；離差爲標準差四倍時，其縱綫高度爲0.00034；離差爲標準差四•〇九倍時，其縱綫高度爲0.00001，已近于零，回顧上列五十男生體重次數分配中，平均數$\bar{X}$等于124.2磅，標準差$\sigma$等于11.5磅，共包括五組距，以$\bar{X}$爲原點，每一組值之離差爲 準差之若干倍，均可求出。此項數值求得後，卽可依之查縱綫表，以得各組距次數縱綫之相對高度。最後再設法將此相對高度化爲絕對高度，卽爲各組距之理論次數，可依之作常態曲綫圖 玆將全部過程表列如下：

(第194表) 五十男生體重分配常態曲綫各縱綫計算表

$\overline{X}=124.2$； $\sigma=11.5$

| $i$ | $X$ | $f_o$ | $x=X-\overline{X}$ | $x/\sigma$ | $Y/Y_o$ | $f_c=Y_o\cdot Y/Y_o$ |
|---|---|---|---|---|---|---|
| | 89.7 | 0 | −34.5 | −3.00 | 0.01111 | 0.193 |
| 95—105 | 100 | 3 | −24.2 | −2.10 | 0.11[illegible]23 | 1.[illegible]12 |
| 105—115 | 110 | 9 | −14.2 | −1.23 | 0.46[illegible]33 | 8.1[illegible]1 |
| 115—125 | 120 | 11 | − 4.2 | −0.37 | 0.93382 | 16.197 |
| | 124.2 | | O | O | 1.000 | 17.345 |
| 125—135 | 130 | 18 | 5.8 | + .50 | 0.8825[illegible] | 15.307 |
| 135—145 | 140 | 9 | 15.8 | +1.37 | 0.3[illegible]123 | 6.786 |
| | 158.7 | 0 | 34.5 | +3.00 | 0.01111 | 0.193 |
| 總計 | | 5 | | | | |




(第52圖) 五十男生體重次數曲綫圖

上列曲綫圖中一爲實際次數曲綫，一爲常態曲綫，常態曲線所代表者爲理論次數。在上列計算表中已知平均數爲124.2磅，此點之離差$x$等于零，次數縱綫之相對高度等于1，絕對高度等于17.345，各組距次數縱綫之相對高度則由常態曲線表中查出，塡入計算表中相對高度$Y/Y_o$一欄

內。各組距之次數既為最高縱綫之百分數，已知最高縱綫及各組距次數之百分數，則各組距次數縱綫自不難求出。已知最高縱綫即 $X$ 等于平均數 124.2 磅之縱綫為 17.345，則 $X$ 等于 100 磅時之縱綫自為：$17.345\times0.11025=1.912$；$X$ 等于 110 磅時之縱綫為 $17.345\times0.46973=8.147$；$X$ 等于 120 磅之縱綫為 $17.345\times0.93382=16.197$；$X$ 等于 130 磅時之縱綫為 $17.345\times0.8825=15.307$；$X$ 等於 147 磅時之縱綫為 $17.345\times0.39123=6.786$。用曲綫連此五點即成一常態曲綫。由常態曲綫縱綫表中知：離差為標準差正負三倍時，其次數縱綫之相對高度各為 0.011111，其絕對高度各為 $7.345\times0.01111=0.193$；離差為標準差正負四倍時，其次數縱綫之相對高度各為 0.00034，可認為零，其絕對高度各為：$17.345\times0=0$，左右各增加此兩點後，所作常態曲綫之兩端方各落於 $X$ 軸上。至於曲綫上仍呈有不規則形狀處，則由於組距過少所致，如組距數目增加，組距之大小減小，最後必能得一平滑之常態曲綫。

## 第四節　計算理論次數之方法

上節之要點在根據常態曲綫縱綫表，以求次數分配中各組距次數縱綫之高度，並依各縱綫之高度製圖，所得為常態曲綫圖。因所得各縱綫之目的在用於繪理論次數曲綫圖，稱為縱綫法(The Ordinate Method)。至於各組理論次數，每根據常態曲綫面積表 (The Normal Curve Area Table)以求之，因稱為面積法 (The Area Method)。常態曲綫下之面積，係代表全部次數，假定此全部面積等於 1，而以最高縱綫即自平均數所在點豎立之縱綫，平分為左右相等之兩半。全部面積代表次數總數既等於 1，或百分之百；以最高縱綫為準，在 $X$ 軸上，無論由此點向左或向右，至任何一縱綫間所包括之面積，究佔全部面積百分之幾，均可由面積表中查

出。面積係代表次數者；所佔面積百分比既可查得，則所包括次數之百分比或相對次數亦可算出。相對次數求得後，各組之絕對自不難計算。

面積表中仍係由平均數 $\overline{X}$ 所在據點，至左右任何縱綫間所包括全部面積之百分比。最左一欄所列者代表各縱綫據點離開平均數之遠近。此種距離仍與上節縱綫法中相同，係以常態離差 $x/\sigma$ 爲單位，現由本書之末所載面積表中摘出幾項重要數值，加以解釋，藉示全表之意義及用途：

（第 195 表） 常態曲綫面積簡表

| $x/\sigma$ | $A$ | $2A$ |
|---|---|---|
| 0.00 | O | O |
| 0.01 | 0.003989 | 0.007978 |
| 0.1 | 0.039828 | 0.079656 |
| 0.6745 | 0.250000 | 0.500000 |
| 1.00 | 0.341345 | 0.682690 |
| 2.00 | 0.477250 | 0.954500 |
| 3.00 | 0.498650 | 0.997300 |
| 3.99 | 0.499967 | 0.999934 |

上表標題中之 $x/\sigma$，即常態離差 $A$ 即面積，$2A$ 即面積 $A$ 之二倍。全表之意義，即在 $X$ 軸任何一點，無論在平均數 $\overline{X}$ 之左或右，其離差爲標準差之零倍時，所佔之面積爲零，其離差爲標準差之百分之一倍時，所包括之面積爲 0.004；其離差爲標準差之十分之一倍時，所包括之面積爲 0.0398；其離差爲標準差之 0.6745 倍時，所包括之面積爲 0.25；其離差爲標準差之一倍時，所包括之面積爲 0.34134；其離差爲標準差之二倍時，所包括之面積爲 0.47725；其離差爲標準差之三倍時，所包括之面積爲 0.49865；其離差爲標準差之三·九九倍時，所包括之面積爲 0.50 或 $\frac{1}{2}$。既爲常態曲綫，左右必完全對稱。故自平均數 $\overline{X}$ 所佔之據點向左右各至一縱綫，其離差各爲標準差之 0·六七四五倍時，此兩縱綫間之面積恰

爲 50/100；其離差各爲標準差之一倍時，此兩綫間之面積爲 68.27%；其離差各爲標準差之二倍時 此兩縱綫間之面積爲 95.45%；其離差各爲標準差之三倍時，此兩縱綫間之面積爲 99.73%；其離差各爲標準差之三・九九倍時，此兩縱綫間之面積爲其全部卽 1 或 100%。

回顧以前所用之五十男生體重分配中，已知算術平均數爲 124.2 磅，標準差爲 11.5 磅，各組距之標準離差 $x/\sigma$ 自可使平均數 $\overline{X}$ 及標準差 $\sigma$ 求出。相對離差求得後，每一組距之內所包括之面積均可由常態曲綫面積表內查出。已知各組距之面積，各組距內所包括之理論次數卽可算出。現作下列計算表以說明之：

（第 196 表）　五十男生體重分配理論次數計算表

$\overline{X}=124.2$ 磅；　$\sigma=11.5$ 磅

| 組距 $i$ | 實際次數 $f_o$ | 組限 下限 | 組限 上限 | 離差 $x=(組限-\overline{X})$ | 常態離差 $x/\sigma$ | $\overline{X}$與組限間之面積 | 組距以內之面積 | 理論次數 $f_c$ |
|---|---|---|---|---|---|---|---|---|
| — | 0 | 83.7 | | −34.5 | 3.0000 | 0.498650 | 0.005543 | 0.2772 |
| 95—105 | 3 | 95 | | −29.2 | 2.53913 | 0.494457 | 0.041917 | 2.0958 |
| 105—115 | 9 | 105 | | −19.2 | 1.66957 | 0.452540 | 0.164395 | 8.2198 |
| 115 | | 115 | | − 9.2 | 0.80000 | 0.288145 | | |
| ～ | 11 | | | | | | 0.316048 | 15.8024 |
| 125 | | | 125 | ＋ 0.8 | 0.06957 | 0.027963 | | |
| 125—135 | 18 | | 135 | ＋10.2 | 0.63913 | 0.326391 | 0.298488 | 14.9244 |
| 135—145 | 9 | | 145 | ＋20.8 | 1.80870 | 0.464852 | 0.138461 | 6.9230 |
| — | 0 | | 158.7 | ＋34.5 | 3.0000 | 0.498651 | 0.035148 | 1.7574 |
| 總計 | 50 | | | | | | 1.000000 | 50.0000 |

求理論次數時所用之離差，仍係以算術平均數 $\overline{X}$ 爲標準，但所求者非各組距中值 $X$ 之離差，乃組限之離差。在小於平均數各組中，係求下限之離差；在大於平均數各組中，係求上限之離差，仍均用 $x$ 代表之。各組距之絕對離差 $x$ 求得後，再各除以標準差 $\sigma$，所得爲標準離差 $x/\sigma$，其手

積與在上節中相同。得到標準離差後即可查面積表，以求平均數與各組限間之面積。既爲常態曲綫，則平均數所佔據點或最高次數縱綫左右兩側之面積完全相等，而且對稱。故自平均數 $\overline{X}$ 所佔據點出發，無論向左或向右，只要代表距離之標準離差相同，則左右兩邊所包括之面積亦必相同。如標準離差爲 0.07，即左右兩側之縱綫去最高縱綫之距離爲 0.07 時，所包括之面積同爲 0.027903；其距離爲 0.8 時，左右兩側所包括之面積同爲 0.288145。總之，不論在最高縱綫以左或以右，凡距離相等，則面積必同。由最高縱線至各組限間所包括之面積既可由表中查出，則每一組距以內之面積均可算出：

(1) 115—124.2 之面積爲 0.288145，124.2—125 之面積爲 0.027903，則 115—125 組距內之面積自爲上述兩面積之和，即 $A=0.288145+0.027903=0.316048$。

(2) 105—124.2 之面積爲 0.452540，而 115—124.2 之面積爲 0.288145，則 105—115 組距內之面積自爲上述兩面積之差，即 $A=0.452540-0.288145=0.164395$。

(3) 95—124.2 之面積爲 0.494457，而 105—124.2 之面積爲 0.452540，則 95—105 組距內之面積自爲上述兩面積之差，即 $A=0.494457-452540=0.041917$。

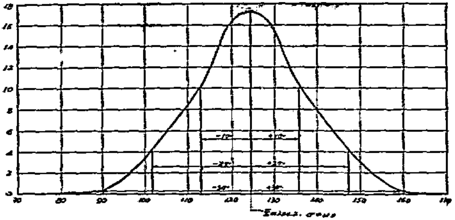
(4) 89.7—124.2 之面積爲 0.50000（實爲 0.49965）而 95—124.2 之面積爲 0.494457，則 89.7—95 之面積自爲上述兩面積之差，即 $A=0.500000-0.494457=0.005543$。

(5) 124.2—135 之面積爲 0.326391，而 124.2—125 之面積爲 0.027903，則 125—135 組距內之面積自爲上述二面積之差，即 $A=0.326391-0.027903=0.298488$。

(6)　124.2—145 之面積爲 0.464852，而 124.2—135 之面積爲 0.326391，則 135—145 組距內之面積自爲上述兩面積之差，即 $A=0.464852-0.326391=0.138461$。

(7)　124.2—158.7 之面積爲 0.50000，而 124.2—145 之面積爲 0.464852，則 145—158.7 之面積自爲上述兩面積之差，即 $A=0.50000-0.464852=0.035148$。

以上已將每一組距上下組限內所包括之面積求出。常態曲綫以下之全部面積既等於 1，最高縱綫左右之面積各爲 ½，或 0.50000，在實際分配中，自下限至平均數即在 95—124.2 間之離差僅爲標準差之 2.53913 倍，面積僅爲 0.494457，同理自平均數至上限即在 124.2—145 間之離差僅爲 1.8087 倍，面積僅爲 0.464852。足見實際分配之上下限與常態曲綫之極限即正負離差各爲標準差之三倍，左右面積各爲 0.50 尙未盡合，故依此標準在下限 95 以下補出面積 0.005543，在上限 145 以上補出面積 0.0 5148，經此補充後，面積一欄內之總和方等於 1，合乎常態曲綫面積之條件。




(第 53 圖)　常態曲綫面積及估計區限圖

各組距所包括之面積既已求得，再求其理論次數時，手續至爲簡易。全部面積等於1，係代表次數總計即$\Sigma f$者。現全部次數等於50。表中面積一欄實係代表各組距理論次數之百分比或相對次數者。相對次數總和爲1或100%，絕對次數總和爲50，各組次數之百分比既已知，由百分比或相對次數化爲絕對次數手續甚易，以各組之百分比乘次數總和即得。現依理論次數作一常態曲綫圖，並在圖中標明估計區界如上：

## 第五節　常態曲綫配合適度之測驗

過去統計學中認爲常態曲綫足以代表所有統計資料之分配。但經近年研究之結果，雖仍公認常態曲綫爲代表次數分配最基本之定律，間亦有統計資料不合於常態分配者。現擬設法研究常態曲綫配合於實際次數分配，是否適當。此種判斷，自可從上節次數圖形中觀察理論及實際二次數曲綫而得一概括之印象。但此法欠精，且無客觀之標準。現另有測驗配合是否適當之方法，其基本觀念自次數離差 (Deviations of Frequencies) 即$(f_o-f_c)$而來。如常態曲綫配合達於最高度，則理論與實際二次數曲綫，完全重合，各組次數無任何離差，即每組之$(f_o-f_c)=0$。但實際上，兩條次數曲線並未重合。即有次數離差存在，現即以各組次數離差平方之和爲批判之標準。此數值以$\chi^2$代表之，故稱爲$\chi^2$測驗法 (The $\chi^2$ Test)，實即常態曲綫配合適度之量數 (Measure of Goodness of Fit of the Normal Curve)。現以五十男生體重爲例以說明之：

下表中所列理論次數上下兩端歸併之目的在使兩端之次數離差平方之值不過份膨脹。如理論次數$f_c$太小時，以之除絕對次數離差平方即$(f_o-f_c)^2$時，即使此絕對離差不大，但以$f_c$太小，除得之結果仍甚大。兩端歸併初無一定標準，以使歸併後次數之值大於10爲度。在上例中以組距過

(第1?7表) 五十男生體重分配常態曲綫配合適度測驗計算表

$f_o$:實際次數; $f_c$:依面積法而得之理論次數

| 組距 | $f_o$ | $f_c$ | $f_o-f_c$ | $(f_o-f_c)^2$ | $\frac{(f_o-f_c)^2}{f_c}$ |
|---|---|---|---|---|---|
| — | 0 | 0.28 } 2.38 | | | |
| 95—105 | 3 | 2.10 | +0.62 | 0.3844 | 0.1615 |
| 105—115 | 9 | 8.22 | +0.78 | 0.6084 | 0.0740 |
| 115—125 | 11 | 15.80 | −4.80 | 23.0400 | 1.4582 |
| 125—135 | 18 | 14.92 | +3.86 | 14.8996 | 0.9986 |
| 1 5—145 | 9 | 6.92 } 8.68 | +0.32 | 0.1 24 | 0.0118 |
| — | — | 1.76 | | | |
| 總計 | 50 | 50.00 | | | 2.7041 |

少,雖亦加以歸併,但不能達此標準。

得到$\chi^2$之值等於2.7041後,即可判斷常態曲綫配合適當之程度。如理論次數與實際次數完全配合,則兩種次數中間絕無離差,$\chi^2$之值必等於零 反之,如兩種次數間有離差存在,即理論次數與實際次數不完全符合,$\chi^2$有數值存在,而不等於零。此項不配合之原因有二: 一爲選樣不充分,統計分析根據局部資料或樣本(Sample),而非根據全部資料(Population)。如根據全部資料,則理論次數與實際次數可以配合,即可以常態曲綫代表全部資料之次數分配。但根據局部材料則二種次數間有相當誤差,此種誤差非由於常態曲綫之不合,乃由於實際分配中所包括之項目過少,不足以充分代表全部之資料之眞面目 此種誤差可稱爲選樣誤差 Sampling Error)。如繼續擴大樣本,則選樣誤差可逐漸減少,最後至常態曲綫充分配合之境界爲止。二爲理論與實際兩種次數根本不相符合;換言之,在本質上卽不能以常態曲綫代表現有之次數分配,局部資料固不合於常態定律,全部資料亦於此不合。上例中$\chi^2$之值等於2.7041,常態曲綫是否配

合，須查 $\chi^2$ 數值表方能決定：

（第198表） 配合適度量數 $\chi^2$ 數值表

| $n$ | $p$=.99 | .95 | .50 | .10 | .05 | .02 | .01 |
|---|---|---|---|---|---|---|---|
| 1 | .000157 | .00393 | .455 | 2.706 | 3.841 | 5.412 | 6.635 |
| 2 | .0201 | .103 | 1.386 | 4.605 | 5.991 | 7.824 | 9.210 |
| 3 | .115 | .352 | 2.366 | 6.251 | 7.815 | 9.837 | 11.341 |
| 4 | .297 | .711 | 3.357 | 7.779 | 9.488 | 11.668 | 13.277 |
| 5 | .554 | 1.145 | 4.351 | 9.236 | 11.070 | 13.388 | 15.086 |
| 6 | .872 | 1.635 | 5.348 | 10.645 | 12.592 | 15.033 | 16.812 |
| 7 | 1.239 | 2.167 | 6.346 | 12.017 | 14.067 | 16.622 | 18.475 |
| 8 | 1.646 | 2.733 | 7.344 | 13.362 | 15.507 | 18.168 | 20.090 |
| 9 | 2.088 | 3.325 | 8.343 | 14.684 | 16.919 | 19.679 | 21.666 |
| 10 | 2.559 | 3.940 | 9.342 | 15.987 | 18.307 | 21.161 | 23.209 |
| 11 | 3.053 | 4.575 | 10.341 | 17.275 | 19.675 | 22.618 | 24.725 |
| 12 | 3.571 | 5.226 | 11.340 | 18.549 | 21.029 | 24.054 | 26.217 |
| 13 | 4.107 | 5.892 | 12.340 | 19.812 | 22.362 | 25.472 | 27.688 |
| 14 | 4.660 | 6.571 | 13.339 | 21.064 | 23.685 | 26.873 | 29.141 |
| 15 | 5.229 | 7.261 | 14.339 | 22.307 | 24.996 | 28.259 | 30.578 |
| 16 | 5.812 | 7.962 | 15.338 | 23.542 | 26.296 | 29.633 | 32.000 |
| 17 | 6.408 | 8.672 | 16.338 | 24.769 | 27.587 | 30.995 | 33.409 |
| 18 | 7.015 | 9.390 | 17.338 | 25.989 | 28.869 | 32.346 | 34.805 |
| 19 | 7.633 | 10.117 | 18.338 | 27.204 | 30.144 | 33.687 | 36.191 |
| 20 | 8.260 | 10.851 | 19.337 | 28.412 | 31.410 | 35.020 | 37.566 |
| 21 | 8.897 | 11.591 | 20.337 | 29.615 | 32.671 | 36.343 | 38.923 |
| 22 | 9.542 | 12.338 | 21.337 | 30.813 | 33.924 | 37.659 | 40.289 |
| 23 | 10.196 | 13.091 | 22.337 | 32.007 | 35.172 | 38.968 | 41.638 |
| 24 | 10.856 | 13.848 | 23.337 | 33.196 | 36.415 | 40.270 | 42.980 |
| 25 | 11.524 | 14.611 | 24.337 | 34.382 | 37.652 | 41.566 | 44.314 |
| 26 | 12.198 | 15.379 | 25.336 | 35.563 | 38.885 | 42.856 | 45.642 |
| 27 | 12.879 | 16.151 | 26.336 | 36.741 | 40.113 | 44.140 | 46.963 |
| 28 | 13.565 | 16.928 | 27.336 | 37.916 | 41.337 | 45.419 | 48.218 |
| 29 | 14.256 | 17.708 | 28.336 | 39.087 | 42.557 | 46.693 | 49.588 |
| 30 | 14.953 | 18.493 | 29.336 | 40.256 | 43.773 | 47.962 | 50.892 |

上表中 $n$ 係代表組距數目 $p$ 係代表機率，表格內部則爲 $\chi^2$ 數值，即相對次數離差平方和。組距數目多少，對於 $\chi^2$ 數值之大小有關，上例中共計五組，即 $n=5$，則 $\chi^2$ 各種數值之機率如下：

(第199表) $\chi^2$ 各種數值機率表

$n=5$

| $\chi^2$ | $p$ |
|---|---|
| 0.554 | .99 |
| 1.145 | .95 |
| 4.351 | .50 |
| 9.236 | .10 |
| 11.070 | .05 |
| 13.388 | .02 |
| 15.086 | .01 |

根據上表所列 $\chi^2$，各種數值之機率，判斷常態曲綫配合是否適宜，至爲簡易。假定常態曲綫十足代表全部資料之分配，理論與實際兩種次數即 $f_c$ 與 $f_o$ 完全配合，則無次數離差，$\chi^2$ 之值等於零。現以根據局部資料，分配中僅包括五十項，由於選樣不充分可以產生誤差。組距數目爲5時，由於 樣誤差所造成 $\chi^2$ 數值等於 0.554 之機率爲 99%。同理，由於選樣誤差所致 $\chi^2$ 數值等於 1.145 之機率爲 95%，所致 $\chi^2$ 數值等於 4.351 之機率爲 50%，所致 $\chi^2$ 數值等於 9.236 之機率爲 10%，所致 $\chi^2$ 等於 11.007 之機率爲 5%，所致 $\chi^2$ 等於 13.388 之機率爲 2%，所致 $\chi^2$ 等於 15.086 之機率爲 1%。

由上文可見 $\chi^2$ 之值之大小於 $n$ 有關，至於 $p$ 又隨 $\chi^2$ 而變。換言之，$n$ 之數值已定時，則 $\chi^2$ 之值愈大，機率愈小，即相對離差平方總和之值愈大，其機率愈小。現以機率等於百分之五爲顯著性之標準 (Standard of

Significance ，凡 $n$ 等於五、機率等於百分之五時，$\chi^2$之數值爲 11.07。任何次數分配，其組距數目爲五，其 $\chi^2$之值小於或等於 11.07 時，即可斷定常態曲綫於實際統計資料配合適當；理論與實際次數須相配合，而無離差存在，即 $\chi^2$等於零，但以選樣不充分所致 $\chi^2$之值爲 11.07 之機率爲 5%。在吾人所分析之五十男生體重分配中，$n$ 等於 5，機率等於 5% 時，$\chi^2$之數值僅爲 2.7041，旣小於標準數值，足見常態曲綫足以配合於體重分配也。惟此種結論又與顯著性之標準有關。由同一次數分配所得 $\chi^2$之數值，以機率等於百分之五爲標準時，認爲常態曲綫可以配合，如改以機率等於百分之十，或百分之五十爲標準時，則常態曲綫即不配合，如改以機率等百分之九十五以至百分之九十九時，常態曲綫配合於統計資料之機會至少。

## 習　題

一、 試以第一編中所作 100 名大學男生高度次數分配爲例，依二項式定理求理論次數。

二、 試依同一資料由常態曲綫縱綫法求各組理論次數曲綫之高度，並繪一理論次數常態曲綫圖。

三、 試依同一資料由常態曲綫面積法求各組之理論次數。

四、 試根據三題中所得之理論次數，求常態曲綫配合量數並評定其值之是否顯著。

# 第二十三章　統計推理

## 第一節　統計敘述與統計歸納

分析統計資料之技術，至此已告一段落。以分析所得之各種結果，代表所包括各個變項之特點，稱爲統計敘述 (Statistical Description)。吾人所研究者爲一百男生身高或體重，以所得之平均數代表此一百男生身高或體重之集中趨勢，以所得之標準差代表此一百男生身高或體重之離中趨勢，以所得之相關係數代表一百男生身高與體重間之平均關係，均爲統計敘述之實例。自一百項事實開始，吾人之分析工作，由搜集原料而製成次數分配，再由次數分配而求出各種結果，以此結果代表此一百項事實之特性，自絕無問題。但統計研究之工作，每欲更進一步。即由一百男生之身高或體重，進而推及全國男生之身高與體重。換言之統計分析之程序，最初自全部統計資料 (Statistical Population) 中，抽查一部分資料，稱爲統計樣本 (Statistical Sample)；繼根據此樣本加以分析，求出各種統計量數 (Statistical Measures)，並以此量數代表統計樣本之特點；最後則應用由樣本所得之量數以說明全部統計資料之通性。其最初之由全部中抽得樣本，稱爲統計選樣 (Statistical Sampling)。其最後利用樣本結果以說明全部資料，則稱爲統計概括 (Statistical Generalization)，統計推理 (Statistical Inference) 或統計歸納 (Statistical Induction)。

統計歸納所以存在，自以選樣在統計研究中之不能免。統計資料中，可以搜集鑑類加以分析說明者，雖亦存在，但大多數統計資料，每無法窺其全豹，全部起料既不克獲得，唯一解決之道，在以分析樣本所得之結果，

代表全部統計資料之通 ˊ，卽實行統計推理。統計推理存在之基石，一爲代表全部之樣本；須爲具備代表性之樣本 (Representative Smple)：二爲確定樣本結果與全體結果間差異之限度， 或稱爲樣本結果可靠性 (Reliability) 之限度。有代表性之樣本，宜具備之條件如下：

1 樣本之選擇未介入偏見或成見。
2 樣本內所包括之成分，彼此完全獨立不倚。
3 樣本資料所由來之地域，須無根本上之差別。
4 組成樣本之各個項目所具備之情形，須彼此相同。

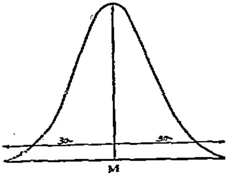
## 第二節 算術平均數之可靠性量數

由富于代表性之樣本所得到之結果，仍不能與由全部材料所得之結果完全相同；是統計推理中，不免包括有錯誤不相脗合之成分，現擬求出適當之量數，直接代表樣本結果與全體結果間差別之限度，間接代表樣本結果可靠性之程度，此類量數概括言之，稱爲可靠性量數 (Measures of Reliability)，分析言之，稱爲樣本量數之標準誤差(Standard Errors of Sample Measures)。某一樣本量數之標準誤差愈小，卽表示此一樣本量數與全體量數間之差別愈小，則此一樣本量數之可靠性愈高，以之代表全體，說明全體時亦較穩定無誤。反之，某一樣本之標準誤差愈大，則表示此一樣本量數與全體量數間之差別愈大，是卽此一樣本量數之可靠性愈低；在此種情况之下，以樣本結果代表全體，說明全體時，所包括錯誤之成分甚多，所作之結論動搖而不穩固。各種樣本量數，均可計算其標準誤差，以代表此種量數之可靠性，現分述之：

一、 平均數之標準誤差

假定以米價爲例。設有一千個 查員，同時出動，每人各向一百家米

店調查品級相同之米價。每人所得一百項米價，構成一米價樣本，共獲得一千個米價樣本，每一樣本內包括一百項價目。每一樣本均可製成次數分配，由分配中求出其平均米價，即樣本平均數。共計有一千份樣本，自有一千個平均米價，或樣本平均數，以$\overline{X}_1$，$\overline{X}_2$，……，$\overline{X}_{1000}$表示之。得到此一千一樣本平均數之後，又可將此一千個樣本平均數製成一次數分配，稱爲樣本平均數次數分配(Frequency Distribution of Sample Means)。個別樣本中各變項所造成之次數分配，可爲常態亦可爲偏態；統計全體所造成之次數分配，亦可爲常態，亦可爲偏態，均殊無定規；但樣本平均數所形成之次數分配是，每爲常態，統計資料之種類儘可不同，此種常態分配之特質，則初無二致。可用下列曲線圖以代表之。




(第54圖) 由同一全體中隨機抽出大量樣本各樣本平均數之理論分配圖
M=真正平均數

一千樣本平均數所造成者既爲一次數分配，自亦有其集中趨勢與離中趨勢兩種特質，分別求出其平均數及標準差以代表之。樣本平均數之平均數(Average of Sample Means)稱爲總平均數(The General Mean)以$M$表示之，樣本平均數之標準差(The Standard deviation of Sample

Means about the General Mean)稱為樣本平均數分配之標準差，或簡稱為平均數之標準誤差 (Standard Error of the Mean)，以 $\sigma_{\bar{x}}$ 一符號表示之，樣本平均數所形成分配既為常態分配，則各個估計界限(Zones of Estimate)中所包括總次數之百分比，自如常態曲線面積表中所表示者：

(第200表) 估計界限表

| 總平均數加減標準差倍數 | 包括次數之百分比 |
|---|---|
| $M\pm0.6745\sigma_{\bar{x}}$ | 50.00 |
| $M\pm1\sigma_{\bar{x}}$ | 68.27 |
| $M\pm2\sigma_{\bar{x}}$ | 95.45 |
| $M\pm3\sigma_{\bar{x}}$ | 99.73 |

全部材料無法搜數搜集，是全部資料之平均數或全體平均數 (Population Mean)無法求出。全體平均數稱為真正平均數 (The True Mean) 真正平均數雖無法求得，但自全體中連續抽樣，所得樣本平均數之總平均數，必與真正平均數相去不遠。故可即以總平均數 $M$ 權作為真正平均數，並標入上列之理論分配曲綫圖中。吾人以此總平均數為標準，進而考察各個樣本平均數 $\bar{X}_i$ 時，則各樣本平均數與總平均數中間均有離差，離差有正有負，有大有小。但其離差小時次數多；離差大時次數少；離差最大最小均有一定之限度，不能超越 $a$.$b$ 之範圍。由總平均數加減 $0.6745\sigma_{\bar{x}}$ 時，恰包括個別樣本平均數之半。換言之，一千個樣本平均數之中，介乎總平均數上加下減 0.6745 倍標準誤差範圍之內者恰五百項，此值稱為機率誤差 (The Probable Error) 或簡稱為機誤，以 $P.E.$ 表示之；即 $P.E.=0.6745\sigma_{\bar{x}}$，機誤一值，美國統計學家用之甚廣，而歐洲統計學家率用標準誤差。實際言之，機誤只為標準誤差之一小數，所代表之意義為：一千個樣本平均數中，在總平均數加減機誤 $P.E.$ 之內者恰佔五百項。換言之，

任一樣本平均數介乎總平均數加減機誤範圍以內之機會與落在此範圍以外之機會恰好相等。樣本平均數與總平均數雖不相同，但其差別最大者亦不過爲總平均數加減三倍標準誤差。蓋百項之中在此範圍以內者已爲99.73項；其在此範圍以外者，百項之中不過0.27項。換言之，任一樣本平均數雖與眞正平均數不同，但其値在眞正平均數加減0.6745倍標誤差範圍以內之機會爲1對1。以眞正平均數爲標準，樣本平均數之離差爲0.6745倍標準誤差時，其一對一之機會既可推出，離差爲標準誤差之其他倍數時，其發現機會之比例亦不難推出，現表明如下：

（第201表） 離差爲標準誤差各種倍數時發現機會表

| 標準誤差倍數 | 在指定標準誤差倍數範圍以內之機率 | 在指定標準誤差倍數範圍以外之機率 | 在指定標準誤差倍數範圍以內及以外機會之比例 |
|---|---|---|---|
| 0.6745 | 50.00 | 50.00 | 1.00:1 |
| 0.7 | 51.61 | 48.39 | 1.07:1 |
| 0.8 | 57.62 | 42.37 | 1.36:1 |
| 0.9 | 63.19 | 36.81 | 1.72:1 |
| 1.0 | 68.27 | 31.73 | 2.15:1 |
| 2.0 | 95.45 | 4.55 | 20.98:1 |
| 3.0 | 99.73 | 0.27 | 369.40:1 |
| 4.0 | 99.99366 | 0.00634 | 15,770.0 :1 |

樣本平均數分配之標準誤差 $\sigma_{\bar{x}}$ 之公式爲：

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{x}}$$

上式中分子 $\sigma$ 指全體標準差而言，但全體標準差若不可得，乃以樣本標準差代替，以 $S$ 表示樣本標準差，則可得平均數標準差公式如下：

$$\sigma_{\bar{x}}=\frac{S}{\sqrt{N}} \quad 或 \quad \sigma_{\bar{x}}=\frac{S}{\sqrt{N-1}}$$

已知五十男生之平均體重 $\bar{X}$ 爲124.2磅，標準差 $S$ 爲11.5磅，樣本

平均數之標準差誤差爲：

$$\sigma_{\bar{x}}=\frac{S}{\sqrt{N-1}}=\frac{11.5}{\sqrt{50-1}}=\frac{11.5}{\sqrt{49}}=\frac{11.5}{7}=1.643\text{ 磅}$$

眞正平均數，即全國大學生平均體重 $M$ 雖不可知，但由五十人體重所得之樣本平均數 $\bar{X}$，或 124.2 磅，可視爲最近於此之量數。眞正平均數 $M$ 雖不等於 124.2 磅，但眞正平均數與樣本平均數 $\bar{X}$ 差別之範圍可以求出。蓋眞正平均數加減三倍標準誤差範圍以內旣包括 99.73% 個樣本平均數，是任一樣本平均數均在眞正平均數加減三倍標準誤差範圍之內。眞正平均數雖不知，但其與樣本平均數間最大離差亦不過加減三倍標準誤差。故樣本平均數加減三倍標準誤差可視數眞正平均數之範圍。

$$M=\bar{X}\pm3\sigma_{\bar{x}}$$

$$\because\quad \bar{X}=124.2;\quad \sigma_{\bar{x}}=1.643$$

$$\therefore\quad M=124.2\pm1.643\times3=124.2\pm4.929\text{ 磅}$$

換言之，眞正平均數之值雖不可知，但其數值所佔據之範圍則可由樣本平均數推出。其最低限度爲樣本平均數減去三倍標準誤差，即眞正平均數 $M$ 最小之值爲 124.2 減 3 倍 1.643，等於 119.276 磅。其最高限度爲樣本平均數加上三倍標準誤差，即眞正平均數 $M$ 最大之值數 124.2 加 3 倍 1.643，等於 129.129 磅。吾人在統計敍述中向以平均數爲一常數。現在統計推理中平均數改爲一變數，其範圍在樣本平均數加減三倍標準誤之間。此種範圍亦可改用平均數之機率誤差（$P.E._{\bar{x}}$）表示之，其公式如下：

$$P.E._{\bar{x}}=0.6745\frac{\sigma}{\sqrt{N}}$$

$$=0.6745\times1.645=1.1082\text{ 磅}$$

$$M=\bar{X}\pm P.E._{\bar{x}}$$

$=124.2\pm1.1082$

$M$ 之最低限度值 $=\overline{X}-P.E._{\overline{x}}$；

$=124.2-1.1082=123.0918$ 磅

$M$ 之最高限度值 $=\overline{X}+P.E._{\overline{x}}$：

$=124.2+1.1082=125.3082$ 磅。

## 第三節　其他統計量數之可靠度量數

在上節中已闡明統計技術中最重要量數即算術平均數 $\overline{X}$ 之可靠度·

（第 202 表）　標差誤差與機率誤差公式表

| 統計量數 | 標　準　誤　差 | 機　率　誤　差 |
|---|---|---|
| 平均數 | $\sigma_{\overline{x}}=\frac{\sigma}{\sqrt{N}}$ | $P.E._{\overline{x}}=0.6745\frac{\sigma}{\sqrt{N}}$ |
| 中位數 | $\sigma_{m}=1.25\frac{\sigma}{\sqrt{N}}$ | $P.E._{m}=0.84535\frac{\sigma}{\sqrt{N}}$ |
| 平均差 | $\sigma_{M.D.}=0.6028\frac{\sigma}{\sqrt{N}}$ | $P.E._{M.D.}=0.4066\frac{\sigma}{\sqrt{N}}$ |
| 標準差 | $\sigma_{\sigma}=\frac{\sigma}{\sqrt{2N}}$ | $P.E._{\sigma}=0.6745\frac{\sigma}{\sqrt{2N}}$ |
| 標準差係數 | $\sigma_{v}=\frac{v}{\sqrt{2N}}\sqrt{1+2v^2}$ | $P.E._{v}=0.6745\frac{v}{\sqrt{2N}}\sqrt{1+2v^2}$ |
| 相關係數 | $\sigma_{r}=\frac{1-r^2}{\sqrt{N}}$ | $P.E._{r}=0.6745\frac{1-r^2}{\sqrt{N}}$ |
| 等級相關係數 | $\sigma_{\rho}=\frac{1-\rho^2}{\sqrt{N}}(1+0.086\rho^2+0.013\rho^4+0.002\rho^6)$ | $P.E._{\rho}=0.6745\frac{1-\rho^2}{\sqrt{N}}(1+0.086\rho^2+0.013\rho^4+0.002\rho^6)$ |
| 多項相關係數 | $\sigma_{R_{1\cdot23\cdots n}}=\frac{1-R^2_{1\cdot23\cdots n}}{\sqrt{N}}$ | $P.E._{R_{1\cdot23\cdots n}}=0.6745\frac{1-R^2_{1\cdot23\cdots n}}{\sqrt{N}}$ |
| 淨相關係數 | $\sigma_{r_{12\cdot34\cdots n}}=\frac{1-r^2_{12\cdot34\cdots n}}{\sqrt{N}}$ | $P.E._{r_{12\cdot34\cdots n}}=0.6745\frac{1-r^2_{12\cdot34\cdots n}}{\sqrt{N}}$ |

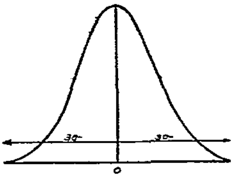
分別用平均數之標準誤差 $\sigma_{\bar{x}}$ 及機率誤差 $P.E._{\bar{x}}$ 以計算樣本平均數可靠之程度。其他統計量數，不拘爲集中趨勢，離中趨勢或相關等量數，亦可利用標準誤差及機率誤差作爲可靠性之量數。已分舉其要者之公式如上。

## 第四節　兩平均數間差數之顯著性

由統計全體中抽出兩樣本，各求其平均數，兩樣本平均數間必有差別，現分析此種差別是否顯著（Significant）。所謂顯著者，卽能證明兩平均數間之差別，係由兩樣本根本非來自同一全體；所謂不顯著則係證明兩樣本原來自同一全體，而兩平均數所以不同，純由於選擇項目時，偶然間機會之所致。統計全體雖可純一，但以兩樣中所包括之項目不盡相同，根据各項目所得之平均數自亦不能毫無差別。第以此種差別，有一定之範圍，苟超出範圍之外，則不能認爲係出於選樣之誤（Sampling Error）。此種平均數間之差別否是顯著，亦可利用上節中可靠性量數方法以分析之。

設由一統計全體中，連續抽出大量成雙之樣本，每雙樣本俱各求出兩平均數，每兩平均數間俱求其差數，并將此大量平均數間之差數，組成次數分配，是爲平均數間差數之次數分配（Frequency Distribution of Differences between two sample Means），此項分配必合乎常態。平均數間差數旣爲常態分配，卽表示平均數間之差數，有正有負，有大有小，此差數分配之平均數將爲零，卽平均數間差數分配之平均數或眞正差數（True Difference between two Means）之值數零，其個別差數或正或負，卽大於零或小於零，但其最大限度不能出於加減此分配三倍標準誤差範圍之外。此種關係可圖示如下：

下圖所示者數從同全體之中，連續抽出大量成雙之樣本，各求其平均數。每兩樣本平均數間雖有差數，但其差數有大有小，有正有負，恰造成一




(第 55 圖) 由同一全體隨機抽出大量成對
樣本平均數間差數之理論次數圖

常態分配。此分配之平均數即平均數間眞正差數必爲零。其個別差數由于偶然之機會可正可負即大于零或小于零，但其值最大限度不得超過此分配標準差三倍之範圍。成對樣本平均數間之差數以 $D$ 表示之。$D$ 係代表一變數，其平均數爲零，其標準差即樣本平均數間差數分配之標準差以 $\sigma_D$ 表示之，稱爲樣本平均數差數之標準誤差 (Standard Error of difference between Means)。樣本平均數間差數標準誤差之公式，係由兩樣本平均數各自之標準誤差而來，即 $\sigma_D$ 由 $\sigma_{\bar{x}}$ 而來：

$$\sigma_D=\sqrt{\sigma_{\bar{x}_1}^2+\sigma_{\bar{x}_2}^2}$$

甲樣本平均數之標準誤差：$\sigma_{\bar{x}}=\dfrac{\sigma_1}{\sqrt{N_1}}$

乙樣本平均數之標準誤差：$\sigma_{\bar{x}_2}=\dfrac{\sigma_2}{\sqrt{N_2}}$

$$\therefore\quad \sigma_D=\sqrt{\frac{\sigma_1^2}{N_1}+\frac{\sigma_2^2}{N_2}}$$

現以全國大學生體重爲一統計全體，抽出甲乙兩樣本，各包括 50 項，即各包括五十人之體重(兩樣本所包括之項目自不必彼此相等)，甲樣本平均數 $\bar{X}_1=124.2$ 磅，其標準差 $\sigma_1=11.5$ 磅，乙樣本之平均數 $\bar{X}_2=130$

磅，其標準差 $\sigma_2=12.5$ 磅。甲乙兩平均數不相同，其差數是否顯著依上列公式分析之：

$$\sigma_1=11.5, \qquad \sigma_1{}^2=132.25$$

$$\sigma_2=12.5, \qquad \sigma_2{}^2=156.25$$

$$N_1=50, \qquad N_2=50$$

$$\therefore \quad \sigma_D=\sqrt{\frac{\sigma_1{}^2}{N_1}+\frac{\sigma_2{}^2}{N}}$$

$$=\sqrt{\frac{132.25}{50}+\frac{156.25}{50}}=\sqrt{2.645+3.125}$$

$$=\sqrt{5.770} \qquad =2.404 \text{ 磅}$$

$$3\sigma_D=3\times 2.404=7.212 \text{ 磅}$$

假定全國大學生體重爲一同種之統計全體（Homrgeneous Statistical Population），由此全體中連續抽樣，大量樣本平均數間差數之平均數或眞正差數（True Difference between Means）爲零。個別兩樣本平均數間之差數以選樣偶然機會關係，其差數不必爲零。但其值必不出乎眞正差數加減三倍標準誤差範圍之外。百項差數之中，在此範圍以內者爲 99.73 項，而出乎此範圍以外者僅 0.27 項耳。故樣本平均數之眞正差數原爲零，實際差數雖不爲零，而未超過三倍標準誤差範圍以外者，則此項差數可視爲選樣之結果，而非以材料之來源根本不同類。現甲樣本平均數爲124.2磅，乙樣本平均數爲 130 磅，兩平均數間之差數爲：

$$D=X_1-X_2=124.2-130=-5.8 \text{ 磅}$$

而個別差數離開眞正差數爲零之範圍爲正負三倍標準誤差，即爲 $3\sigma_D=\pm 7.212$ 磅。現在兩樣本平均數間之差數爲 $-5.8$ 磅，尙未超過三倍標準誤差範圍之外，足證此項差數由於選樣時偶然機會所致，而非甲乙兩樣本之來源基本不同也。

# 立信會計叢書目錄

## ★簿記類

- 簿記初階　李文杰編
- 商業簿記　甘允壽編
- 初級商業簿記教科書　陳文麟 施仁夫編
- 高級商業簿記教科書　潘序倫著
- 英文高級簿記會計　潘序倫著
- 高級商業簿記實習題附屬文件

## ★會計學類

- 會計學（一—四冊）　潘序倫著
- 會計學　錢素君 夏治濤編
- 初級會計學　王逢辛編著
- 會計學概要　李鴻壽編
- 會計學教科書　潘序倫 王澹如編著
- 會計問題（上下冊）　施仁夫 唐文瑞編著

## ★銀行會計類

- 銀行會計　陳穎安編著
- 銀行會計　顧準 陳穎安著
- 暫行銀行統一會計制度

## ★成本會計類

- 成本會計　陳文麟譯
- 陀氏成本會計（上下冊）　施仁夫譯
- 勞氏成本會計　潘序倫譯
- 勞氏成本會計習題　潘序倫譯
- 棉紡織廠成本會計　陳文麟著

## ★政府會計類

- 政府會計　張蕙生 王成杰編
- 中國政府會計制度　潘序倫 顧準編著
- 公有營業會計　余肇池編著
- 政府會計人員手冊　汪元錚編
- 政府會計制度一致規定

## ★審計學類

- 審計學　顧詢 唐文瑞編
- 審計學　潘序倫 顧詢著
- 政府審計原理　蔣明祺著
- 政府審計實務　蔣明祺著
- 銀行內部審計　陳成權著
- 審計問題　錢迺澂編
- 查帳報告書及工作底稿　顧詢 錢迺澂編
- 審計問題答解　錢迺澂編
- 中國現行審計制度　許祖烈著

## ★其他會計類

- 股份有限公司會計上下冊　潘序倫著
- 鐵道會計　張心澂著
- 電業會計　楊湧編著
- 各業會計制度（一二三集）　潘序倫編
- 倉庫實務會計　卡宗遜著
- 會計名詞彙譯　潘序倫 莫啟歐編譯
- 會計數學　李鴻壽 莫啟歐編
- 會計數學用表　李鴻壽
- 決算表之分析及解釋　潘錦甲譯
- 決算表之分析　黃組方著
- 決算表之編製及內容　黃組方編著
- 無形資產論　施仁夫譯

▲各種簿記會計書籍均有習題詳解專供各校教員參攷之用須憑學校證明文件方可照售

# 立信商業叢書目錄

## ★商業類

商極常識 陳文 張英閣編
商業概論(上下册) 陳文著
商業應用文作法 龐翔勛編著
財政學概論 王延超著
國家經濟學原理 林和成譯
貨幣學 陳紹武著
銀行學 金天錫 朱榮岩 陳嶺光著
銀行實務概要 王濟如著
廣告學 丁馨伯著
投資學 任福履著
公司實務 鄒世賢編著
珠算匯宗 李鴻鏞著

## ★法規類

公司法
新公司法解釋 張蘂元編
活頁直接稅法規
活頁工商法規
銀行法
鑛業法規
保險業法規
工商業奬勵法規
政府會計審計法規
工商業同業公會及人民團體組織法規

## ★統計類

統計學 褚一飛編著
統計學新論 王思立編著
調查統計 蕭承祿編著

立信會計師事務所設計：

# 立信帳簿表單

印製精良　格式完備

種類

帳簿、帳夾、傳票、表單、機關用各式帳簿表單、學生簿記練習紙

（詳細目錄索閱即行寄奉）



立信會計圖書用品社股份有限公司

印製發行

立信會計叢書

# 統計學新論

全一冊



每冊基價國幣二元
外埠酌加郵費運費

編著者　王　思　立

發行人　顧　詢

發行所　立信會計圖書用品社
上海河南路三三九號
南京中山路二一三號
重慶小什字立信大樓

印刷者　周順記印刷所
上海惠民路三一八號

中華民國三十六年十月初版
中華民國三十七年一月再版　(滬)