統計學

中華經濟統計研究所

叢書之一

統計學

褚一飛編著

立信會計圖書用品社發行

中華經濟統計研究所
叢書之一

統　　計　　學

全一冊

每冊基價國幣一元四角五分
外埠酌加郵費運費

編著者　褚　一　飛

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

中華民國三十二年十一月初版
中華民國三十六年六月五版　（滬）

卷頭語

邇來吾國統計學教科書已有不少佳著出版，故於任教時雖常有同學以余所授頗多爲各書所未及，屢請將講稿付印，然竊意終覺無此必要。蓋憑個人經驗，認爲筆記講授方法較有效益。但自抗戰以還，統計功用日益顯著，統計教育日見發達，兼以青年學子羣集後方，各校學生倍增，程度參差不一，遂使專憑筆記講授漸感困難。蓋整理筆記既需充分時間，且須有相當書籍以供參考。顧戰時學生課外活動較多，時間頗感缺乏，參考書籍亦因交通關係而殊感稀少，故各校同學皆紛紛要求編印講義，竟使余不得不勉爲其難。柰事實上因同時兼任中央政治學校重慶大學復旦大學華西工商專科學校四校課務，每週往來於南溫泉沙坪壩北碚江北之間，竟毫無餘暇從事寫作。今春適聞劉坤闓俞壽榮兩同志於軍需學校及中央政治學校普通科講授統計學時，曾編有講稿數章，因即商得劉俞兩君同意將該項講稿稍加補充，完成本書，以應目前迫切之需要。全書凡二十章，分訂上下兩册。劉君撰寫上册第一章緒論，第二章統計資料之搜集與整理，第三章表列法

，第四章圖示法，第五章平均數，第六章相差度，第七章偏斜度與峯度，第八章確度。俞君撰寫上册第九章比例，第十章指數，下册第五章時間數列之分析，第十章抽樣概論。余草擬下册第一章機率，第二章相關度，第三章相關函數，第四章偏相關與複相關，第六章相變度，第七章插補法，第八章配合概論，第九章常態曲線與常態曲面，以及附錄各篇。全書於短短數月內完成，掛漏之處，自所難免，尚望海內賢達不吝指教，俾再版時得以改正，幸甚幸甚！

又戰時印刷困難異常，此次承文化建設印刷公司諸位先生極力協助，始克儘速出版，並承內子郁韻漪女士親任抄寫，雖夏日炎炎，空襲頻傳，而工作迄未間斷，此皆所應感謝焉。

三十二年暑期

飛寫於四川萬縣

統計學概論目錄

統計學概論

第一章　緒論

第一節　統計之意義

統計一詞，含有相輔相成之二義：其一爲資料，卽表出事實之數列；其二爲方法，卽搜集、整理、表列、圖示上述資料並加以提綱挈領，比較研究，以便述敍或說明事實之眞象之方法。譬如鍊鋼，統計資料、鋼鐵也；統計方法，採鑛、提鍊、以及如何應用鋼鐵之方法也。統計資料、通常簡稱統計，如資源統計、農業統計、社會統計、人口統計、經濟統計、財政統計、政治統計、軍事統計等是。研究或敍述統計方法之學術，謂之統計學，如農業統計學、人口統計學、經濟統計學、教育統計學等是。本書簡述一般統計方法，是爲普通統計學。

統計資料乃原因複雜而以數字表出之多項事實的總合，各項事實係依照預定目的，有系統的方法及合理的確度標準、計算或估計而得，且其排列次第可以表現相互關係俾便比較者。逐層解釋如次：

1.統計恆爲多項事實的總合，單項事實不得稱爲統計。構成統計之事實，或爲某地某現象在各時間之變動；或爲某現象在同時同地之各種情形，或爲某現象在異時異地之狀況。一人之死生、一事之偶發、一次交易、一日天雨、一官、一兵、一槍、一彈、皆不足以構成統計。而多數人之死亡，多數同類事件之發生，屢次交易，歷年雨量，各部隊之官兵械彈則均可編成統計。何者？以其爲多項事件之總合，可爲依時間、地域或屬性分析研究之資也。

2.統計所表現之事實，其原因甚爲複雜。各項事實係從錯綜複雜之對象內測量而得，常與他種測量密切關聯，其所從產生之環境互異且其變化多端，日新月異。

3.統計乃以數字表出事實。統計所注重者爲數量之多寡，而非品質之異同。凡有差別，概以數目區分之。例如某省歷年水田每市畝平均出產稻穀若干市石，統計也；同一現象，若以豐富、優良、平常、不佳、甚嗇等語句表明之，除非各語句均代表相當之數量，則不得謂爲統計。

4.統計必須依照合理的確度標準實際計算或設法估計。夫如是，方可以相加、平均、提要以爲行政上或學理上之依

據。且在搜集事實之過程中，確度標準，必須一致。至於如何正確，方爲合理，則視統計之目的而定，無絕對之標準，可資遵循。若用以爲某項設施之依據，則需要較高之確度；若祗需得到一般之印象，則較大之差誤，或亦無關重要。

5.數量的觀察之不失爲名實相符之統計，必需根據一定目的用有系統的方法爲之。計算或估計事物之目的不同，則搜集資料之單位及確度標準亦隨之而異。目的變更，資料或仍可繼續獲得，而所得之資料，未必與目的相符，因而不足以爲統計分析之資，而作有效論證之依據。

6.統計必須能依次排列，以表現各事實之相互關係。至其排列次第之標準，或依時間，或依地域，或依屬性。欲事物可以比較，必須彼此於品質上有共同之點。散漫零亂之數字，道聽塗說之消息，漠不相關之資料，來源不明，確度無稽，苟亦以統計名之，未免魚目混珠。必須品質純一，堪資比較，而爲事實之總合，乃得謂之統計。

統計學或統計方法者就計算或估計所得之資料，用科學方法分析綜合，由是以說明社會現象或自然現象，並從而推論其靜態或動態之規律，以爲學理之印證或人類活動之指針

之學術也、略加闡明如後：

1. 統計方法之對象，為前述之資料。統計方法亦卽搜集整理、陳示並分析研究資料之方法。

2 統計方法為科學方法，卽觀察事實，認識現象，由各項事例以求共同原則(分析)或由共同之原則以求現象之印證(綜合)之方法。

3. 統計方法之應用，不限於社會現象，自然現象亦可應用。例如生物學上進化論遺傳論均以生物統計為基礎，近世數理統計之進步，亦多源於生物學之研究。氣象學上測量氣溫、氣壓、濕度、風力、在若干次觀察中取其代表，與統計方法，殊無二致，歷年之平均氣溫，正與人口統計上之出生、死亡、婚姻相似。倘不用統計方法，不足以自歷年之平均數，獲得若何結論；藉統計方法之研究，可知歷年之變動究為偶然事件，抑有特殊意義；究為長期升降，抑屬週期變動；究竟有無規律可尋；為來日現象之預測，上述問題，均極重要。統計方法與天文學相接近者有兩端，其一為最小二乘法之引用，導源於一天文學者。渠因欲在對某恆星位置之幾種觀察結果中，選擇最可能之估計，乃藉最小二乘法應用差

誤律以決定之。其次爲確度之漸次增高，此點與地質學及其他應用科學亦極相同。科學之測量，最初大抵取一極粗率之觀察結果，如太陽之距離、岩層之厚度、某原素之原子重量、某物體之比重等是；俟材料日漸累積，儀器之確度增高，計算之方法改良，乃使測量逐漸可靠。此種進步，殊堪注意。在目前人類智識之狀態下，若干統計上之測量，以資料之缺乏，未能臻於精確之境地，因而批評家不免指摘初步之估計殊無價值。而自科學之立場觀之，此種批評，殊爲錯誤，因合乎論理原則之差誤的測量，倘能說明可能差誤之限度究勝至於無所知，且由繼續努力可望確度與日俱增。至於社會科學方面，尤多可述者。尋常往往以爲統計方法限於人口學之範圍，頗不盡然。倘使人口學不限於研究人口數，出生、死亡、婚姻、年齡分配、性比例、地域分佈等，亦包括其他經濟現象及社會動態、如職業分配、所得、工資、物價、生產、貿易、交通等，則吾人已擴大人口學之範圍，使之包羅現代統計工作之泰半，爲社會學家經濟學家所直接從事者。至少對於上述學者，統計學可爲下列之界說：統計學者，視社會機構爲一整體，自各方面以測量社會機構之科學也。Le

Play 之於單獨家庭從事個案研究也：家庭分子之職業、收入以及其經濟地位，均一一詳加紀錄，此種研究雖自有其價值，究不能稱爲統計的研究。在人口學上，吾人研究多羣家庭，例如從事某項實業之戶數，其平均收入、支出、儲蓄等，必如是然後爲統計。於個案研究，個體極重要；而在統計方法，個體不足輕重。欲測量一羣事物，各個個體之特性，全不注意；必須多數人具有同種特性，然後此種特性引起吾人注意。例如調查人民之所得，高等技師月入萬金，低能工役，月入數元，於平均所得，影響至微，並不特別提出；惟有技術工人每月平均收入五百元，臨時勞役每日平均八元，方值得分別注意。個人之變態不居，而集體之變動，則甚爲遲鈍；欲知每個原子之運動，絕不可能；而欲說明固體物運動之定律，則較爲容易。在社會現象之測量上，大數及從大數求得之平均數，往往賦有極大之惰性。例如，全國人口、國民所得、出生率、死亡率、平均工資等變動極微，而某家之人口、所得、死生、收支則變動極鉅且速。統計測量之所以可能，正由於大數之有惰性；故統計方法之應用，乃偏重大數之觀察。統計方法與經濟學之關係，尤爲顯明，馬夏爾

有言曰：統計資料泥土也，我輩經濟學者運用泥土以製造甎瓦。蓋因統計人員供給經濟學者以事實，經濟學者用以印證其學理，用以爲學理之基礎。因經濟學家以研究社會集體之現象爲主，視個人不過爲社會之一份子，而統計學乃大數之科學平均之科學，故經濟學家引領以待統計界供給其材料。苟非統計陳示事實，則國民經濟上關於貿易額貨幣購買力等項問題，限於純粹理論，無異空中樓閣。統計人員有如實驗室從事實驗之化學家；經濟學者有若研究室從事鑽研之化學家。因而統計人員，必如小心翼翼之實驗家，蒐集、整理、敍述事實，而不推理立論；縱或從事因果之分析，亦但提證據而不爲結論。不過統計人員亦同時得爲經濟學者，正如化學家旣能實驗亦能從事純理研究。實際上，若理論化學家而不嫻悉實驗方法及其困難，則殊少權威之可言；以此類彼，倘經濟學者而不明暸統計方法及其困難，亦不知何處可得正確之數字，亦不能評議資料之價值，亦不自知其理論究竟立於何等可靠之基礎上，則殊未能自命有充分之素養。

4.統計方法對於一切現象，有時作靜態之觀察，有時作動態之觀察。靜態者一時存在之狀態也，動態者繼續發生之

狀態也 靜態爲動態之橫斷面，動態爲靜態之延長線。例如人口數、年齡分配、性比例、人口密度等須在某一時點觀察之，此時點之前後，其狀況未必相同；而出生、死亡、婚姻等，則須在一時期內爲經常之紀錄，期末之累積事實，始爲研究之對象。

5.統計方法之目的在發現規律以資印證學理或爲學理之依據，或作人類活動之指南針。例如貨幣數量學說，其始也，依物價統計，貿易統計，以及金融統計之研究，獲得物貨交易額，貨幣暨信用之數量及流通速度與物價之關係，並決定貨幣爲主動，物價被動。欲印證此說之是或反證此說之非，唯有上述各項統計能勝其任。至統計在人類活動上之應用，實爲統計發生之初因。古代以國家之徵兵徵稅，而對人口，財產爲詳細之記載，統計(Statistics)之字根爲國家(State)，或由於此。今日政府權力擴大，行政設施，在在須依據事實，故統計事業乃隨政治清明而日見發達。近世紀來以經濟循環之研究，爲商情預測之張本，更爲統計方法之充分運用，雖則目前此項學說日見式微，而其方法則永不至磨滅，而將日見應用推廣而發揚光大也。

由上所述，可知統計學之內容約如下述：

1.選擇對象，蒐集資材。(取材)

2.依事實之共同徵性，分類整理。(整理)

3.陳示已經分類之資料。(陳示)

4.計算平均數，相差度等，以資提綱挈領。(分析)

5,考訂資料之確度。(審核)

6.測定資料間之關係或其變動趨勢。(應用)

因統計方法包括以上各種手續，故平常凡調查、登記、分類、整理、列表、繪圖、計算等，均以統計代之，是時統計係用作動詞。又統計工作常將觀察之事實，彙成總數，故統計亦間常爲人用作總計之意。

第二節　統計之功用

由統計意義之說明，可見其功用甚廣，無論自然現象或社會現象，靜態或動態，靡不可藉統計方法，研究其資料，以解答其全部或部分之問題。

自統計學之分類觀之：

1.純理統計學，以數學爲基礎，探求統計方法之原理原則，以爲從事統計研究或辦理統計事務之圭臬。數十年來，

統計方法日見謹嚴，實由數理統計研討之進步。欲能從事數理統計之研究，必須有較深之數學造詣，至少須嫻熟初等微積分。一般學者對於標準差，次數曲線，尤其是相關係數，應用常不免太濫，推理因而不得謹嚴，或由數學根基欠佳所致。

2.應用統計學乃應用統計方法於各種現象，以求其一般法則，以表明各現象之眞象，藉以構成學理，指示行動。其在生物學、氣象學、天文學、人口學、經濟學上之應用，已於統計學界說中，言其梗概，茲不贅述。

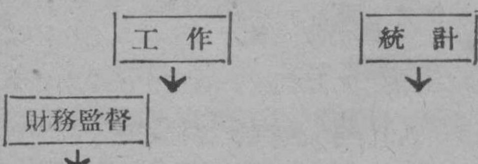
就某項現象之應用統計，其功用可概括言之：

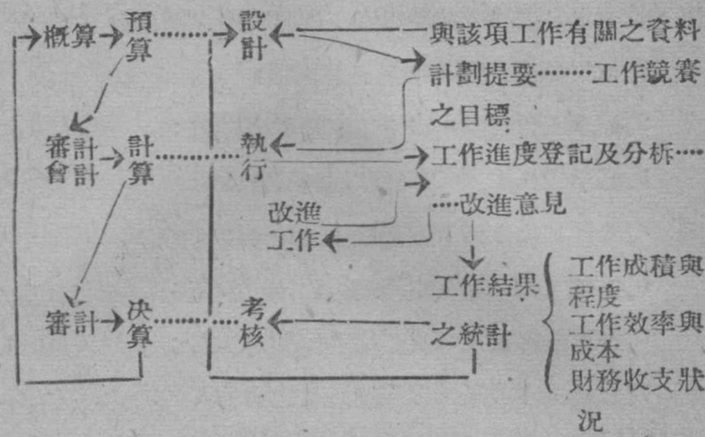
1.綜合分析，表明眞象，以觀現象全體之分佈；

2.提綱挈領，以簡馭繁，以察現象內幕之情況；

3.尋求關係，推斷趨勢，以測現象未來之變化。

就某項工作之過程言，統計之功用如下：





當一件工作之決定實施也，先應將各方面有關之參考資料，充分蒐集，以爲擬訂工作計劃之張本；計劃旣定，內容複雜，非主管其事者，無須知其詳；一部分工作者，無庸知其全，應提示要點，以爲各工作單位競賽之目標。各部分依工作要點邁步前進，日就月將，應將工作進行狀況及進行程度，詳爲記錄，並加以分析，何以未能按原定計劃實施，追求其癥結之所在，以便改進執行方法。及工作完成，依逐步登記之記錄，編製工作結果統計，工作成績與標準是否與原定計劃相符；工作效率及成本之高低如何；經費收支之狀況如何；一一考核，以爲獎懲工作人員，改進來期工作，宜

傳工作效果之依據。故統計在工作過程中之功用如下：

1.蒐集各項資料，確定工作計劃；

2.提示工作要點，以為工作競賽之目標；

3.表示工作進度，以便改進執行方法；

4.考核工作成績，藉定獎懲之標準；

5.計算工作成本，以便尋求經濟之途徑；

6.表現財務收入，以定事業擴縮之方針；

7.宣揚工作效果，以籲獲社會之同情。

現代社會分工日細，合作愈見需要，大規模之工作，大多需要繁複之組織，衆多之人員，鉅量之財物，方能舉行；而立機關，用人員，耗財物，無非求某項計劃之集體推行；至行政工作之推行，傳遞命令、報告工作、概以文書出之。為提高工作效率，降低工作成本起見，有不斷調整機構，改進人事，革新財務制度，改善文書辦法之必要；而如何調整改革，應否裁員減政，無不需要正確之統計，以供參攷。非然者憑一二人之私見，輕事更張，勞民傷財，無裨實益。此所以凡百公私事業，均應逐步發展統計工作，以求舉措之有事實根據也。

爲三節　統計之限制與誤用

吾人旣知統計之功用甚廣，猶不可不知統計之限制。統計方法並非萬能，在不能以數字測量之現象，統計卽失其功用。例如蓄意是否忠誠，秉性是否聰敏，持身之廉潔，待人之信用等，於人類活動均極重要，而均不能直接用數量測量。充其量統計所能表現者，僅爲現象可測量之一面，而通常所測者，並非欲測之現象本身，僅其有關之數量而已。例如吾人欲知貧窮之程度，及其增減情形，但貧窮之意義甚不易確定，其程度亦無法測量，甚至貧民之數目亦無由獲得。吾人所能爲力者，僅列出官方發表之貧民數目，或再加上私人之估計；然此於人民貧窮程度之探究，並無線索可循也。再如欲得健康統計，而主要之數字，不外死亡率，平均壽命，流行病症等，與建康本身大不相牟。是以統計者對於各類問題之供獻，僅爲客觀的測量，而此每居各項資料中較不重要之地位；然而客觀的測量於問題之解答，其需要之般，殊無異正確測量之於工程之建築也。

統計事業之所以不發達，各方面對之了解不深，應用不當乃至信心缺乏，是其主要原因。有人責備統計：『任何事

情，均可以統計證明之。』此何故也？蓋因人民需求資料，報紙雜誌任意供給之，殊不知統計材料之充實，胥賴統計工作之建全，而由於種種原因：如統計機構之缺乏，工作人員之不力，調查登記之未臻完備等，以致完備之資料，殊不易得。資料之供給者，乃不得不訴諸未必正確之估計，而大衆又無判別之智識，不知何種資料方屬可用，何種數目始有統計價值，再者局部之資料，每被視爲通例，目的全異之估計，引爲完善之資料。諸如此類不可靠之條件下，作理論之依據，由於巧妙之選擇材料。勢必致任何歪曲之學說，均可以統計證明之。然而，統計人員實不任其過咎。通常根據統計，發表歪曲理論者，不外引用數字，而割裂其上下文，並略去關於不完備之註脚；或者應用數字於全不相干之現象；或以局部之估計作全體之眞象；或單引於所持之理有利之資料，而掩蔽其他；或由果推因，不加深思。凡此所述，皆論理（邏輯）上之基本錯誤也，而其責任往往諉諸統計。統計豈任其咎耶？

惟應用統計之時，實不可不愼重將事。故意引用統計以證明歪曲之理論者，較諸不懂統計者，其罪尤無可逭。統計

以確定之數字出現，加之以整齊之表，美觀之圖，每每令人神往。致歪曲之理論，亦每每爲人所信，而釀成錯誤之言論。統計爲學者之工具，正如自衞之武器，旣可以防禦危險，亦可任意傷人，甚至操刀自損。統計可爲正確理論之根據，亦可以爲歪曲理論之註解，惟在用者之善自選擇耳。忠誠之學者，應注意下列各端：

1.切勿預存成見，以爲統計數字必能證明何事。

2.於願望相舛忤之數字，切勿輕易拋棄之。

3.某事之各種可能原因，均應重加考慮，一一記述，切勿歸於某一因素——事實爲若干原因之綜合。

4.切勿比較並無共同之點之資料。

此外善意之學者，每每於統計方法不甚嫻悉，致有下列之傾向，必須引爲鑑戒：

1.任意應用手邊之資料，而忽略其他可能之資料。

2.集中注意於統計數字，而忽視其來源與單位。

3.不注意所用資料之原來目的及其確度。

4.不明統計資料之是否有連續性，而統計資料之內容往往因編製機關及人員之變動而隨之變動。

5.摘取數字，而忽略重要之詳細情形。例如應用總數，而不注意總數所由獲得之細數。

學者能深省統計方法之限制，誤用統計資料之流弊，從而慎重將事，則庶乎其可矣。

第四節　辦理統計之程序

辦理統計之程序，須視所辦統計之種類、範圍、性質而異，一般之程序，大約如下；

1.確定計劃或方案　在統計工作進行之前，應將研究事實之範圍與性質，精密研討，以求所須解決問題之所在，並審查問題之因素，由是確定統計計劃或方案，庶其日後工作有一定步驟。

2.選定單位　依據所須研究之問題與其因素，以及應行搜集資料之性質和用途，選定適稱之單位，以爲搜羅資料之標準。各種單位應加以明確界說，使無誤會之可能。

- 個體單位
  - 自然單位……如一人，一牛。
  - 人爲單位……如一桌，一公司，一師。
- 事態單位……如火災，交易，犯罪，自殺，戰爭。
- 測量單位
  - 度量衡單位……如一尺，一升，一斤。
  - 金錢單位……如一元，一鎊，一法郎。

簡單單位……如一人，一尺，一桌。
組合單位……如家庭，工廠，農場，學校。
表述單位……如百兩，萬噸，百分數，等成年男子。

3.搜集資料　按照既定單位；根據一定方法，或爲調查，或爲登記，搜集必需之材料，以供分析綜合。

4.分類整理　將搜羅所得零星雜亂之材料，用人工或機器整理之。

5.陳示資料　用表列法及圖示法陳示已經整理之資料，並加文字之說明以供研究，或作施政之參考。

6.勘校結果　統計結果可靠與否，應詳加審核勘校。其已知確度者，應明白書明之

普通事業機關辦理統計，大致作竣上列步驟，編成統計報告，即已竣事。至於詳細之分析，研究，多係學術機關之工作。但欲使資料正確可靠，各項統計人員，不可不明統計分析方法。

第五節　統計報告之編製

統計報告以其目的之不同，約可分爲兩類：

1.呈報上級機關者　如分公司對總公司，下級機關對上

級機關，會員對於其會社之報告等。此類報告又可分爲兩種

甲、一次報告　此類報告乃因特種目的而舉辦統計所編製之報告。應針對目的，解答問題，並詳述調查登記之經過，提供翔實可行之意見。

乙、經常報告　此種報告，均有一定時期和一定格式，爲有歷史性並供一般事務改進之參考用者。應按期依式呈報，以使上級機關如時彙編範圍更大之報告。切不可遲延時期，或自創一格，以貽「一馬不行百馬憂」之譏。

2.向社會發表者　政府統計結果，除祕密類外，其餘均應公告或公開，公告者按期公布，俾衆週知；公開者，隨時可任人民借閱或抄錄。此類報告，應斟酌對方程度，務求通俗易曉，免生疑竇。學術機關、團體或個人亦往往以其研究所得，編成統計報告，公諸社會。其目的不外提倡某項改革工作，徵求社會公意；或發現某項原理，要求學術界之認識。前者最須注意體例，務須途途是道，引人入勝；後者則須確有高見，言人之所未言，方有學術價值。

以上不過略述一般編製統計報告之原則，普通統計報告，大抵包括下列四項內容：

1.統計表

2.統計圖

3.圖表的說明

4.將來的推測或改進意見。

流行的報告，太多千篇一律的圖表，徒事裝璜，至於圖表所代表的事實，全不加以說明，以致材料之來源及其確度，不得而知，殊失統計的眞意。除專供查閱之統計圖表以外，似均應附以詳細說明，敍述編製經過以及編者之感想或意見，使統計報告成爲生動而有用之文字與數字。

第六節　統計人員之修養

所謂統計人員包括（一）根據一定目的擬具統計計劃或方案之人員；（二）搜集材料之調查或登記人員；（三）整理材料繪製圖表之人員；（四）分析資料編製報告之人員。各項人員自須有各項專門之學養，不可一槪而論：

1.擬具計劃與編製報告之人員應對於所辦統計之事業對象，有深切之研究，例如辦理物價統計者，應精研經濟學及經濟現狀；辦理交通統計者，應深切了解交通理論與技術；辦理軍事統計者，應明白國防原理及軍事學等是。至於對統

計方法須有根底，自不待論。

2.調查人員須有豐富之常識與機變之智能，有時且須能說調查區域之方言。尤須有刻苦耐勞之精神，和藹熱誠之態度與乎不怕風霜之體格，才能勝任愉快。

3.室內審核材料及整理材料之人員，其工作頗爲單調，最須要耐煩、細心、有恆、注意。並須對於整理之材料有相當常識；於算盤計算機之運用嫻熟。如須兼繪圖表，更應有相當之訓練，始克勝任。繪圖人員於略有統計常識之外，兼精美術與幾何畫，則尙矣。

至於一般統計人員均應具備之條件如下：

1.毫無成見　統計人員若預存成見，則其結果，不免人爲而不合統計之規律者，美諺云『數字不能撒謊，惟撒謊者好造數字耳』誠哉斯言！故統計人員，從事工作之先，不宜預存成見；工作之際，尤不可有利害偏見。如是則統計結果，庶乎其可矣。

2.能守祕密　統計工作之有關國家要政個人利害者，切應嚴守祕密，不可絲毫洩漏。設使此類消息一經傳出，國家個人或統計人員，必蒙重大損害。被害之團體或個人，日後

必不樂於供給材料，或給予虛僞材料，則統計工作，阻礙多端矣。故嚴守祕密，實統計人員，必備之條件。（軍事時期，有關國防祕密者，更應特別注意）

3.秉性精細　統計上之事實，不外以數字代表其現象，往往差之毫厘，謬以千里。苟其頭緒紛繁，手續複雜，要非有精細耐煩之腦筋，何以克稱厥職？

此外，從事任何工作，腦筋淸楚爲第一要件，文字通順次之，語言流利亦甚重要，三者雖半由天賦，而勤能補拙，祇須努力學習，自能左右逢源，應付裕如也。

第二章　統計資料之搜集與整理

第一節　初級資料與次級資料

統計資料，種類緊多，自資料之已否整理觀之，可分爲兩類：

1.初級資料　此類資料，係方從調查或登記搜集而來，未經整理彙編者，亦稱粗料。其來源乃爲資料原始取得之處，謂之原始來源。原始材料之優點在於來源直接，眞實可靠，惟整理較爲麻煩耳。若材料性質差異，填寫方法不齊，則整理之時，更感困難。惟目的不同，則搜集資料之標準各異，辦理統計者，切不可存心避難就易，採取他人現成之資料，濫竽充數，致獲得錯誤之結論。

2.次級資料　此項資料乃搜集已經他人整理彙編之資料，亦稱成料。其來源謂之次級來源，或爲政府機關，或爲學術機關，或爲個人。搜集現成資料，只須運用得當，自屬事半功倍，惟切宜注意：搜集之機關，是否可靠；主持之人，能力如何；刊布資料爲一時的抑爲繼續的。若貿然採用他人杜撰之資料，不免貽誤大事，不可不再三留意焉。除上述各端而外，尚須注意下列各點：

甲、原始來源　次級資料，以愈接近原始來源愈爲可靠。因爲轉輾謄寫翻印，差誤在所不免；而且次級來源，往往將原編者的註釋省略，刊布之資料，或許祇有原來之一部分。雖則次級資料，較原始資料，格外整齊，吾人採用時，總宜格外注意。

乙、原始單位　利用次級資料，必須注意其單位。例如作工人調查，所謂工人，是指加入工會之工人，仰包括未入會者？係指工廠工人，抑包括一般勞役而言？又例如軍隊統計，所謂部隊，專指國防軍隊，抑包括憲警及保安隊等而言？凡此種種，均應加以研究。

丙、原來目的　搜集資料，往往各有特定目的，并非盲目從事。引用次級資料，應當追問原來搜集資料之目的，是否與吾人之宗旨相脗合？若不相同，則應如何調整分類，以應吾人之需要。

丁、原來方法　目的不同，搜集方法固難期一致；即目的相同，搜集方法亦每每各異。調査或登記的時期及區域之範圍如何？係抽査，抑爲普査？係直接搜集，抑估計而得？如係估計，其根據及方法如何？諸如此類，均須審愼辨別

，判斷其方法是否合理。

戊、數字確度 統計數字，無絕對正確的可能，故刊布之資料，常註明其確度，以供他人研究。吾人在採用時，當特別注意其確度，是否有偏誤及蒙蔽揑造之處？如爲抽查，其標樣之大小如何？代表性如何？均須詳加審察。

以上五端，均經一一審核，方可決定此項資料是否可用。有時對於同一問題，有兩種以上之次級資料，因來源不同而內容互異，此時亟須追求不同之原因安在，若均尙可靠，則折衷應用；如其不然，則擇善而從之。

第二節 靜態資料之調查

無論自然現象或社會現象，均可從靜態或動態兩方面觀察之，一時存在之現象，謂之靜態現象；繼續發生之現象，謂之動態現象。前者爲某時點存在之狀況，如民國二十九年六月三十日重慶市之人口數，某某學校之學員生人數，我國之軍隊人數，某公司之資產及負債數等是；動態現象爲某期間發生之事件，如民國二十九年上半年重慶市之出生死亡結婚人數，某某學校入學、畢業、請假學生人數，我國軍隊之死傷補充人數，某公司之進貨銷貨及盈虧情形等是。代表靜

態現象之統計資料，吾人名之曰靜態資料，代表動態現象之統計資料，名之曰動態資料。一時存在之現象，同時呈現於目前，故可一次同時調查之；繼續發生之現象，出現無常，未可預卜，故必須經常不斷登記之。本節先述靜態資料之調查方法。

現在警察調查戶口，形似調查而實爲登記，因所查之事項爲遷移、出生、死亡、婚嫁等偶然發生之現象，非同時呈現者，須憑被查者之記憶作答。此類事項本應由人民自行申請登記，惟因我國人民向無此種習慣，不得不由官廳派遣警察請人民登記耳，實不得謂之調查也。

一、普查與抽查

將調查範圍以內全部分子，一一加以調查，謂之普查，亦稱全體調查，在調查範圍以內，選擇一部分分子，加以調查，據此以估計全體之眞象，謂之抽查，亦稱抽樣調查。前者如人口普查、農業普查、工業普查等是；後者如物價調查、生活費調查等是。

普查除國家之人口資源或範圍較狹之調查外，極少行之。理論上普查雖最完善，惟有時限於人員、經費、時間種種

原因，勢不可能。且若有足夠之標樣，統計人員又有正確判斷能力，亦無普查之必要。

統計人員恆敢於萬千對象中，選取一部分之現象，以推斷全體現象者，蓋恃下列三項基礎定律：

1.統計常性定律 無論何種事物，在其全部內選擇一部分觀察之，就平均言，此一部分必具有全部之普通徵性。如在我國人口中，選擇若干為代表，以量其身長體重，其平均長度重量，必與一一測量全國人口之結果相近。

2.大數量性定律 （第一定律系1）無論何種現象，就大量觀察之，非有特別情形為之因，則變動極小。如家庭人口與收支變動頗大，而就全國觀之，苟非有戰爭饑饉等為之因，則人口及財政收入變動極小。

3.小數恆性定律 （第一定律系2）無論何種事物，其罕見之現象，或具有非常之性質者，雖為少數而常有之。都市之火災，船舶之失事，均屬此例，牙科耳科醫生，古玩字畫商人，殆賴此定律以謀生者。

抽查雖有上列三項定律為基礎，倘須遵守下列三要件，方能期其結果正確：

1.抽樣不可避難就易，或預存成見；

2.每一分子，被抽之機會均等；

3.每一分子，不受其他分子當選或落選之影響。

抽查之方法有三：

1.廣泛抽樣　採取標樣愈多愈好，不問標樣之性質如何。此法之弊，在容易避難就易。卽調查者不存成見，資料亦常缺乏代表性。現在用此法者甚少。

2.隨機抽樣　完全根據機遇法則，採取標樣，使全體分子皆有均等之當選機會。其優點在使選樣者，不能有主觀成見參雜其間。其危險在標樣往往不多，未能適量；有時仍難免感情作用，惟此法只須謹依上述三項要件，仍不失爲善良之抽樣方法。

3.代表抽樣　在取樣之前，將調查對象，作一初步之觀察與切實之研究，然後規定一確切之標準，以選擇標樣，作爲全體之代表。此法今日用之最廣，其危險在標準若失之主觀，調查者可以有意或無意採取其所希望之標樣，以達到預期之結果，惟事前若果有詳密之觀察，力求標準之客觀化，並採取足量的標樣，則結果當較前二法爲佳。

二、原始調查方法

調查方法可大別爲兩類：直接搜集原始資料，謂之原始調查，搜集他人現成資料(卽次級資料)，謂之繼起調查，關於利用他人現成資料所應注意之點，已於前節說明，茲專論原始調查方法。

原始調查方法有四：

1.親自調查　卽研究者個人親自往原始來源一一調查。此種調查，適於範圍較小之精深研究，同時研究者對於所探討之問題，須有特殊興趣，否則不免半途而廢。其缺點在標樣太少，難於代表全體，且研究者易雜入主觀直覺的見解，影響研究之結論。法人 Le play 用此法研究歐洲工人家庭預算，英人 Booth 用之以研究倫敦貧民生活，均窮年累月，方竟厥功。

2.派員調查　任用專門調查員，加以訓練，使任調查工作。範圍校廣之調查，多須用此種方法。美國人口普查，卽用此法。此法耗費太大，私人調查，往往以限於財力，難以舉辦。

3.通信或頒表調查　將調查表發給被調查者，由其自行

塡寫，應用此法者，常假定對方有一定程度之智識。結果之成敗，全恃對方對此事之興趣而定。若非有法律制裁，被調查者，常不願置答；所答亦不易完全與正確。爲使被調查者不致因疑竇而生頑忽，應將調查之目的，主辨機關與主管人員，註明於調查表上。惟此種調查，節省經費，故公私機關作範圍較廣之調查，常採用之。

4.估計法　若干事項，有時不能正確切實測量或計算，祇能用估計法來補救。其法應將應行調查的事實，規定適當問題，特約專人估計報告或派員至各地徵集。美國農業部許多農產收穫統計，我國農業實驗所的農情報告，均係採用此法。各種估計，雖未必十分可靠，但若估計者忠實可信，則其結果之錯誤，常能互相補償抵消。此法優點，在於簡易而經濟。所宜注意者，估計並非揑造，估計方法須有確切根據，乃有價値。

在不背地方情形與習慣以及適應調查之性質兩原則下，有時可酌採各種方法合併應用。例如法德戶口普查，先用通信調查，令人民塡表，於一定日期，再派員校正，是兼用通信調查與派員調查；美國勞工統計局編製工人生活費指數時

，食品類用通信法，服用類則派員調查。

三、調查前之準備

實地調查前的重要準備有兩項：

1.確定目的　調查之前，必有一預定之目的，擬具計劃者對於所研究問題的各方面，均須有相當認識，正確概念，才能用一定標準，選擇適當方法，以搜集準確合用的資料。例如作工資調查，如係研究生產成本，當調查貨幣工資率；如欲研究工人生活狀況，則當注重實際所得，以及貨幣購買力，又例如編製物價指數之先，必須決定，是測量貨幣購買力之升降，是研究進出口物價，抑是研究某階級之生活費變動，才能決定要調查之物價，爲躉售物價，零售物價，或某地、某時、某階級、某種類之物價，將來編製之方法，亦大相逕庭、若干人或因缺乏從事事業之經驗，或因不諳統計方法，或因主觀成見太深，往往由不完備不適當之資料，得到不正確之結論，此非統計之過也，乃辦統計者之過也。

2.擬訂表格　調查表格乃搜集資料之工具，其製作之良否，與調查是否順利，結果是否優良，關係至切。多數調查、，常因表格不良，以致失敗。惟製表技術，全恃統計人員之學識與經驗而定，殊難罄述，茲僅就表格的內容與形式，略

述一般原則如下：

甲、表格的內容

(1)措詞須懇切簡明，務使被調查者樂於答覆或填寫，不致敷衍塞責。

(2)內容應不涉私隱私權，不損私人人格，以免引起反感與厭惡。

(3)有時宜採旁證法補救，如問青年男子之婚姻，與其問「結婚否」，不如問「妻子在何處」；問女子的年齡，最好由其兄弟姊妹之年齡來旁證。

(4)問題須具體易答，最好能用「是，否」，記號(如✓，一)或數字作答，或用確切不移之字句作答。

(5)問題數量，以足夠需要爲度，不可貪多。國家之人口普查，多設一問題，甚至須多花若干萬元。惟法律上規定有答復義務之問題，則不妨稍多。

(6)問題應依重要次序排列，重要者列在前面，使首先答復，其後之問題，答復不全，亦無關緊要；通信調查，此點尤屬重要。

(7)問題應避用專門術語或抽象名詞，如問工人「

每月實際所得幾何」，工人將瞠目不知所對。

乙、表格的形式

(1)表之大小寬狹，宜便於攜帶與收藏。美國標準調查表有2×5吋，5×8吋，$8\frac{1}{2}\times11\frac{1}{2}$吋三種。

(2)表之質料，以堅硬爲宜，紙面須平滑，使調查者可用手持之填寫。

(3)若干種性質不同之調查表，最好以顏色區別之，既便調查，又便整理與保管。

(4)宜用大小不同之文字與粗細不同之線條，表示重要性不同之問題，與關係密切與否之事項。

(5)問題之排列，不可太擁擠，當疏密得宜，並須留足填寫答案之空白。

(6)表上之標題，務求簡明醒目，使一望而知其意義。

(7)有時應寫明調查機關，調查目的；並留餘地填寫調查員或填表人姓名及調查或填表之年月日。

第三節 動態資料之登記

靜態資料與動態資料之意義，以及前者可以同時調查，

後者必須經常登記之理由，已於前節約略述之。

自然界與人類無日不在活動中，欲在此萬狀紛陳之現象中，求得各種規律與法則，以為吾人來日活動之鑑戒，自應將各種各色發生之事態，加以紀錄，彙類整理，分析綜合，自萬變中以求不變之原理，自異象中以求共象。如何紀錄，最為經濟，最為正確，此統計學上所當研究之登記方法也。普通將登記法，視為調查方法之一種，殊屬不當，實則調查與登記，乃搜集資料之兩種不同方法。人類有文字而後有記載，由記載產生歷史，廣義的歷史，包括一切山川文物，典章制度，學術思想，民族國家之沿革變動的系統紀錄，有歷史而後今人可將前人之一切經驗據為已有，今日之文化文明，乃人類歷史上活動之總和。如何將此寶貴經驗，更有效的記載，以為短期內之參考，並以傳之於後人，此乃登記方法之所以亟須切實講求也。

我國現行統計法第三條規定，政府應辦之統計為下列五種：

1. 基本國勢調查之統計——國家之人民、土地、資源及政治、社會、經濟、文化等同在一時期內舉行之普查之統

計（施行細則第十條）

2. 各機關職務上應用之統計——舉凡可供各該機關擬訂其施政計劃與行使其職務所需要之參政材料。（施行細則第十四條）

3. 各機關所辦公務之統計——各該機關執行職務經過與結果之紀錄。凡屬下列性質者，均應有詳確之紀錄：

甲、可以表現施政計劃推行之成績與程度者；

乙、可以表現工作之效率與每單位之費用者；

丙、可以表現財政收支之狀况者。（施行細者第十七條）

4. 公務人員及其工作之統計——公務人員之統計包括各機關及所屬各部分之組織編制、員額、與公務員之資歷、等級、薪給以及進退、遷調、獎懲等之記載，（施行細則第二十條）。公務人員工作之統計包括各個公務人員辦理各項職務之成績、效率與所費以及日常辦公到退、請假、出勤等之情形。（施行細則第二十一條）

5. 各機關認為應辦之其他統計。

上列五種統計，除第一種純屬靜態資料外，其餘動靜各

半。而最重要者爲第三種，簡稱所謂「公務統計」。廣義言之，政府所辦之統計，概可稱之爲公務統計。現代國家權力日加擴充，人民之一舉一動，幾無不受國家之督導，而政府之一切措施，又莫不與人民有關，故推而廣之，公務統計殆可包括一切統計而言。而公務統計動態部分殆十居八九，過去在政府辦理統計者，祇知有「調查統計」，而忽視「登記統計」殊屬錯誤，吾人亟宜矯正之。

調查須有調查計劃，登記亦須有一定方案，統計法所稱統計方案，雖包括靜態統計而言，但仍以動態統計爲其主要部分。統計方案應明定之事項爲(一)統計之機關單位及其分級；(二)統計區域；(三)分期進行之統計計劃；(四)統計科目；(五)統計單位；(六)統計表册格式；(七)調查登記及編製之方法；(八)統計之公開程度；(九)統計報告印行範圍；(十)其他應行明定之事項。(統計法第六條)

統計方案之擬訂最關重要，方案一經訂定，則日後之登記、整理、編製報表等工作，均可按期依式爲之，正如會計制度一經確定，則一切記帳，過錄，編製報表均成機械工作也。惟會計以財務爲對象，金錢爲單位，簡而易行；統計之

對象，千差萬別，其單位又雜沓紛陳，非於統計對象之事業，有精深研究，統計方案實無由擬制，加以相沿成習，會計關係經費之收支，每爲一般人所重視；而統計之良窳，似無足輕重，致爲人所看輕，其實統計之重要，並不亞於會計，讀者詳研第一章第一二兩節，自能明之，毋庸贅述。

統計方案之擬訂，一如調查計劃之編製，須擬訂之人員於其統計對象之學理，有深刻研究，於其技術，有相當經驗。例如擬訂教育統計方案者，須知教育學，曾從事教育工作者尤佳；而教育範圍甚廣，擬訂小學教育統計方案者，更須懂小學教育。其他各種統計方案，無不皆然。此外於方案擬訂之前，更須注意下列三端：

1.研究有關法規

甲、組織法規　規定各機關及其部分組織或人員之組織編制及職掌。

乙、服務規程　規定各機關及其部分組織或人員執行職務之程序及服務之規律。

丙、單行法規　規定每一事項之辦理方法與應行報告之內容。

由組織規程及服務規程，分析各機關組織系統及其職掌，以確定(1)統計之機關單位及其分級(2)統計區域。由單行法規，分析各項公務，以確定(1)統計科目，(2)統計單位，(3)統計表册格式，

2.研究過去習慣，卽檢討公務實況　習慣爲不成文法，凡統計方案之無法規可資依據者，應研究過去辦理此項統計之辦法，參酌學理，以定取舍依違。所當注意者，一般人之惰性頗大，大都樂於墨守成規，而不願學習新方法，一有更張，輒加反對。當擬訂方案之初，宜一本大公，擇善而從。其於法於理，均無依據；於實際工作，又無效用之可言者，宜大刀闊斧，毅然革除，絕不可姑息，有所顧忌。應行檢討之實況如下：

甲、各項職務之名稱。

乙、各項職務中包括之事項。

丙、各項職務之辦理情形暨有關事實與數字之逐級報告程序。

丁、規定前項逐級報告程序之法令條文。

戊、前項逐級報告之表册。

己，負責審核並登記前項表册之部分組織。

庚，登記表册之法定或習用格式。

辛，負責彙齊、整理、陳示、呈報或發表前項登記册中事實與數字之部分組織。

壬，整理之表册及手續。

癸、彙報或發表時之法定或習用之報表格式。

3.研究有關學理　新辦之公務或事業，統計方法，既無法規可資遵循，亦無習慣以供參考，則惟有根據學理以爲擬訂統計方案之準則。卽有成規可循者，亦應詳研學理，以資印證。而一事所涉之書報，往往浩如烟海，故非素有涵蓄者，每難執其要以概其餘。

至統計方案之內容，約分下列各項：

1.逐級登記、整理、報告之程序　依行政或工作系統，確定此項程序。例如戶口變動統計，依現行辦法，由戶報甲，卽由戶長向甲長申請登記、由甲、保、鄉或鎮、縣或市、省、逐級整理彙報至內政部，由內政部統計處彙報國民政府主計處。

2.應用表册之目錄　原始書表、登記表册、整理表、報

告表、其名稱應一一列舉之，

3.各套應用表册之關係　由原始書表至最後報告表，中間經過極繁複之手續，離原始書表愈近，內容愈繁；離最後報告表愈近，則內容愈簡，自複雜之原始表至簡單之報告表，其間經過之整理程序及各表之關係，應繪圖表明之。

4.原始書表之格式　即申請書、登記表、發票、證書等，其格式應一一詳加規定，以便取得材料之基本來源，有所遵循。

5.登記底册及登記卡片　最基本之取得材料機關，應預備詳盡之登記底册，將原始書表上有統計價值（有歷史意義及可供行政之參考者）與留備查考之事項，一一登記之，此爲登記方法中最重要之步驟，材料是否詳確可靠，胥視登記底册而定，故須特別愼重以擬定其格式。某事項之主管機關爲便於管理查閱，及整理材料起見。並常須將重要事項製成卡片，例如經濟部工業司之於工廠登記是也。

6.整理表之格式　下層組織應如何整理材料，必須規定其格式，以期方法一致。

7.報告表之格式　下級報告表，均須詳訂格式，至於最

後報告表，則可預先擬訂雛形，詳情容後决定。因最後編製統計報告之機關，其人員大都嫻熟統計技術，且最後報告，常因各項目的，而有詳略異同，未可拘於定式也。

登記底册，整理表及報告表之下，均應規定(1)辦理機關(2)材料來源(3)辦理程序(4)說明（關於辦理方法，時期，塡表方法，排列次序等，應詳加說明。）。

第四節　資料之審核

塡就之調查表，登記底册，以及下級機關呈報之報告表，在收到之後，整理之前，應詳加審核，務使原始資料達到下列四點：

1.正確　在原則上，調查表、登記表或報告表上之文字及數字，每個均有存在的權利，審核者不能隨意改正。若發現兩個或兩個以上不同之答案，而不能按照機遇法則或一般原則以决定何一答案較爲正確時，必須覆查或退還重塡，或寫明某表某處派員或通信查詢，務求正確。如拋棄一表，於調查無甚關係時，可竟棄之，惟限於抽查。

2.相容　如有不相容之處，應設法改正之。例如細數相加不等於總數，則必有一誤。數字最易犯前後兩字顚到之誤

，若相差爲9之倍數、多半是犯此誤。又例如某處之戶數大於口數，必有錯誤，或係戶數多寫一位；或係口數少寫一位，或有其他錯誤。再如某人之年齡爲8歲，而其妻爲26歲，則8歲或爲28歲之誤。

3.一致 同一事項用幾個不同名詞塡列，如「理髮匠」與「剃頭匠」，「未婚」與「已經訂婚」是也；又同一物品用不同單位塡列，如「斤」，「市斤」，「公斤」是也。當設法使其一致，以便分類及整理，

4.完備 一項問題務求答案完備，有空白未塡之處，校核者可從其他答案中，找到相當之答語或數字補充之，不可任意視爲「〇」或否定的答案。如未塡之處，甚爲重要，則惟有發還重塡，或通信詢明之。

資料旣經審核，如認爲已無問題時，應由審核人簽名或蓋章，表示業經負責審核之意。並宜逐表登記，以便查閱資料已彙齊與否，或已收到何時期之資料。更可以考核調查員或下級機關是否守時，是否認眞，檢閱自己整理資料之進度。例如：

二十九年度各航政局進出口貨物月報表登記

月別	漢口航政局			上海航政局			天津航政局			…………
	收到	審核	彙編	收到	審核	彙編	收到	審核	彙編	
1	｜	ㄥ	○	｜	ㄥ	○	｜	ㄥ	○	
2	｜	ㄥ	○	｜	ㄥ	○	｜	ㄥ	○	
3	｜	ㄥ	○	｜	ㄥ	○	｜	ㄥ	○	
4	｜	ㄥ		｜	ㄥ		｜	ㄥ		
5	｜			｜	ㄥ			ㄥ		
6				｜						

某某縣第一督查區戶口調查登記

調查區	調查員	調查表		覆查表		再覆查表	
		收到	發還覆查	收到	發還再覆查	收到	審核
1	趙　某	102	4	4	—	—	—
2	錢　某	110	7	7	1	1	∠
3	孫　某	125	3	3	1		—
4	李　某	95	—	—	—	—	—
⋮							

第五節　資料之整理

整理資料乃將復雜散漫之粗料，依照預定標準，分門別類，作成有系統有條理之資料，以便陳示而資分析。科學分類及列表之要義且俟下章述之，至於整理或歸類方法有人工整理與機械整理兩種：

1. 人工整理　小規模之統計工作，大半均用人工整理。人工整理又有兩種方法：

甲、記號法　由原始調查表或登記表整理之。或先將個別之調查表或登記表，彙錄（彙錄法）在一總表上，然後用記號整理之。

中央各機關公務員調查彙錄表

某某機關第1頁

調查表	姓名	性別	年齡	職稱	職掌	月薪	到職年	月
1	趙某	+	36	科長	………	260	26	9
2	錢某	+	27	科員	………	140	27	4
3	孫某	−	23	辦事員	………	70	25	5
⋮								

就原始表或彙錄表上之某項材料，分類（如性別，職稱、職掌）或分組（如年齡、月薪、在職期間）後將各個情形、分別記入各相當類組，記號隨人規定，普通以五個為一組如一、丅、下、疋、正；丨、𠂆、冂、口、囗；丨、‖、Ⅲ、Ⅲ、卌等是。

某機關職員年齡整理表

歲數分組	記　號	人　數
總計	——	
16——20	下	3
21——25	正正	9
26——30	正正正正正正正正一	41
31——35	正正正正正正正	34
⋮		

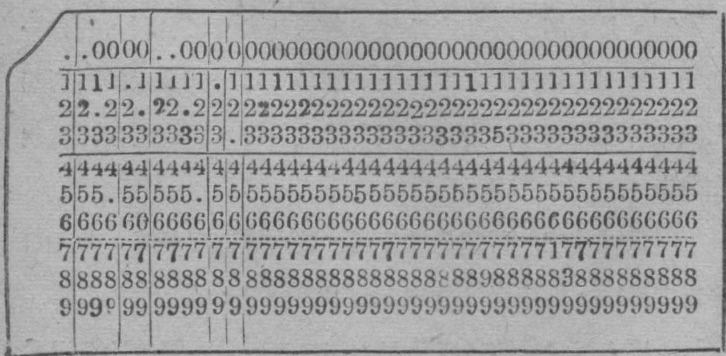
乙、卡片法　用堅硬紙質，每一單位事實（如一人，一校，一工廠，一機關等）謄一卡片，然後某項目（如年齡、職稱、月薪等）分類或分組。將各類或組之名稱，各書一紙條，排列在相當大之桌台上，將各卡片分別放置於其所屬之類組，放置完畢，計算各類各組之數目，填寫於預先製就之表上。

2.機械整理　科學先進各國，較大規模之統計，如人口普查之整理，均用機械整理。其手續如下：

甲、編號　將品質分類，都化爲數量。例如性別以「

○」表男，「1」表女；教育程度以「○」表不識字；「1」表小學程度「2」表中學程度，「3」表大學程度；婚姻狀況以「○」表未曾結婚，「1」表現有配偶，「2」表鰥寡，「3」表離婚等是。

乙、打孔　在特製編號卡片上，規定各事項所佔的行數，如第一行表性別，第二至四行表表年齡，第五及六行表省市藉，第七至十行表縣市籍，第十一行表婚姻狀況，第十二行表教育程度等是。例如某人爲男性，25歲，四川巴縣人，現有配偶，大學畢業乃將卡片置於打孔機上，打孔如下：



丙、歸類　將已打孔卡片置於歸類機內，通以電流，則同類同組(孔之地位相同)之現象，卽各歸一類。每分鐘約可整理250 張。

第六節　資料之保管

原始資料，旣經整理以後，似無再行保存之價值。實則不然，因硏究目的不同，可用各種觀點，各種標準作不同的整理，以符合分析方法。故原始資料，應妥爲保存。通常且宜謄寫複本，分別保管，以防一份萬一損失時，不致影響其餘一份。他若業經整理之成料，圖書刊物，檔卷文書等，均應分門別類，加以保管，以便檢查。至於保管，應有一定次序，自不必論。某種調查之調查表應視其多寡與重要性，訂定妥善保管辦法。玆假定以交通統計爲例，擬定分類方法如後，以供參考。共分十類。

000　交通行政及交通統計行政

100　交通行政統計

200　鐵道統計

300　公路統計

400　航政統計

500 航空統計

600 電政統計

700 郵遞統計

800 郵政儲匯及簡易壽險統計

900 其 他

每類之下分綱，例如鐵道統計，分

210 路線及局站

229 新路工程

230 鐵道工務

240 鐵道機務

$50 鉄道運輸

260 鐵道行政

每綱下再分項，例如鐵道運輸綱分

251 貨運

252 客運

253 軍運

255 公務運輸

255 營業收支

項下分目，目下再分細目，可用小數。此種分類方法，係仿杜威氏圖書分類法，其優點在於簡明，第一數字表類，第二數字表綱，第三數字表項…………，如保管熟練以後，『252』一望而知爲鐵道類、運輸綱、客運項，於歸檔調卷，異常方便。

第三章 表列法

第一節 統計資料之分類

統計資料之整理，乃將繁複之事項依一定標準分類，並列表以陳示之。關於歸類之手續，前章第五節已述其大要，本章再論分類之標準，及列表之方法。

一、分類之要義 分類者將事物之某項性質相同或相近者歸入一組之謂也。如學生之分男女，是將性別相同者歸為一組：成績之甲乙，是將考試分數之相近者歸入一組：人之分嬰孩、兒童、少年、青年、壯丁、老年等，乃將年齡之相近者，歸入一組是也。宇宙萬象紛陳，倘不分門別類，何以觀其條理，執簡以馭繁？惟分類若不妥當，則又足以掩沒真象，攪亂思路，不可不慎也。

分類有兩項步驟，先將每一事項，按各種標準而標明其特質，例如某生之性別為男，年齡為二十歲，籍貫為四川巴縣，教育程度為中學畢業，婚姻狀況為未曾結婚，家庭職業為農業等是，是之謂離心工作，然後以一項或數項標準為主，將特質相同或相近之各事項歸入一組，如凡言教育程度不

能讀　國父遺囑者爲不識字，凡在小學肄業或畢業或非正式學校修業纔滿六年者爲小學程度，凡在中學肄業或畢業或非正式學校修業六年以上或剛滿十二年者爲中學程度，凡在大學畢業或肄業或非正式學校修業十二年以上者爲大學程度等是。是之謂向心工作。

分類之原則一曰周延，即不遺漏，卽某標準下之各組能包括應行分類之一切事項，而無有事項無組可歸者；一曰相斥，卽不重覆，卽凡能歸入某組者，僅能歸入該組，絕無歸入他組之可能。此兩項原則，雖甚簡單，而行之頗非易易。例如職業分類，疾病分類，世界上尚無統一方法，以其內容太繁，欲求一不遺漏不重覆之分類，殊非易事也。

二、分類之標準　統計事項之分類，不外下列三種標準：

1.根據事物之屬性分類，例如三十二年六月三十日重慶市之人口，按職業分類是也。

2.根據所在地域分類，例如三十二年六月三十日我國之人口按行政區域分類是也。

3.根據發現之時間分類，例如我國歷年之人口數是也。

根據屬性分類，更有下列各個不同的標準：

種類之不同…………如種族、職業、性別。

程度之差別
- 品質之差別……如年齡分老、壯、幼；成績分甲、乙、丙、丁、戊。
- 數量之差別
  - 不等級的‥如「18歲以下」「19歲—45歲」，「46歲以上」。
  - 等級的‥‥每組的大小（組距）相等，此種分類之結果，爲標準的次數分配

用一個標準分類，稱爲單純的分類，用兩個或兩個以上的標準分類，稱爲複雜的分類，用兩個標準如身長與體重，出身與職別，通常可以表示兩個因素的關係，故所列成之表，謂之相關表。

第二節 表列之功用

表列者將分類整理後之資料。依某種次序排列之之謂也。將散漫之資料，列成有系統之表，容易使人領悟其要旨，旣易於審查其錯誤，又便於分析其結果。且有時不待更深之研究，卽足以表現相互關係因果或趨勢。再則因排列有

一定次序，可借聯念作用，與人以深刻之印象，容易紀憶。至於因數字之大小，線格之粗細，排列之先後，符號及顏色等之應用，以代表文字之說明，省却繁複之敍述，節省篇幅，猶其餘事也。

第三節　統計表之種類

統計表可以依種種標準分類：

一、依統計事項分類的標準分類

1. 屬性數列表
2. 時間數列表　詳見第五節
3. 空間數列表

二、依統計表之用途分類

1. 普通表　將固有之事實，應有盡有，許加記載，保存其眞象，以備各人依其不同之目的作詳細研究時之資料。政府按期發表之統計報告多屬此類。

2. 專門表　乃學者爲特定目的，將普通表中之某項事實提出，作精細之研究，以發現原理原則而證明其學說之合理者。

三、依統計表之形式分類

1. 開式表　印刷較便，形式美觀，曉近報告表多採用

此種形式。

2.封式表　　晚近除調查表及登記表外採用者較少

四、依內容之繁簡分類

1.單項式　　表示依一個標準分類的表，如學生人數依所屬學院分：

學　　　　院	學　生　人　數
總　計………	
文學院……	
理學院………	
法學院………	
商學院……	

2.雙項式　表示依兩個標準分類的表，如學生人數依(1)所屬學院(2)學籍分：

學院	學生人數		
	總計	正式生	特別生
總計……			
文學院……			
理學院……			
法學院……			
商學院……			

3.三項式　表示依三個標準分類的表，如學生人數依(1)所屬學院　(2)學籍　(3)性別分：

學院	學生人數								
	總計			正式生			特別生		
	合計	男	女	合計	男	女	合計	男	女
總計……									
文學院……									
理學院……									
法學院……									
商學院……									

4.多項式　表示依四個或四個以上標準分類的表，如畢業學生人數依(1) 所屬學院 (2)學籍 (3) 性別 (4)年度 分：

年度	學院	學生人數								
		總計			正式生			特別生		
		合計	男	女	合計	男	女	合計	男	女
總計	合計………									
	文學院………									
	理學院………									
	法學院……									
	商學院………									
二十七年度	合計………									
	文學院………									
	理學院………									
	法學院………									
	商學院………									
二十八年度	合計………									
	文學院………									
	理學院………									
	法學院………									
	商學院………									

第四節　製表法則

如何製表，雖無一定不移之規律，惟依據經驗，可得一般之法則如下：

一、關於表題者

1.表題宜簡明。應將表內現象所有之徵性扼要提示。普通應將事實、時間、地點三者表明之。如民國三十年我國各省區之小麥產量。

2.表題之意義應與標目一致。

3.表題置於表上，寫法自左至右。

4.一表如不只一頁，各頁均應將表題寫出並註明頁次。

二、關於標目者

1.標目之地位，以便於比較爲原則。舉例如下

第一位：同行比較，如年齡；

第二位：同列比較，如性別；

第三位：同行隔列比較，如婚姻狀況；

第四位：同列隔行比較，如國籍。

性別及國籍 年齡及婚姻	總計			華人			外人		
	合計	男	女	合計	男	女	合計	男	女
	I	II	III	IV	V	VI	VII	VIII	IX
總　計………(1)									
滿二十歲者………(2)									
未滿二十歲者………(3)									
已婚人數………(4)									
滿二十歲者………(5)									
未滿二十歲者………(6)									
未婚人數………(7)									
滿二十歲者………(8)									
未滿二十歲者………(9)									

2.標目之次序，依下列標準，由上至下，由左至右排列之。

(一)重要程度　　(五)地域位置

(二)等級高低　　(六)筆劃多少

(三)時間先後　　(七)字母前後

(四)數量大小

3.標目之用語(字眼)，須能表現全項事實，自身能充分說明而無須附註解釋者。若能避免專門術語而能使一般人了解者更佳。

4.比較重要之事項、其標目可用較粗大之文字表示之。

5.大項下分小目、小目後分細節，各低一字書寫。

6.表如太長或太寬，標目可於右端重寫一遍或寫標目之次序。

7.各項目之前，視情形不同列「總計」或「平均」「百分比」等名稱。

8.為便於檢查起見，標目可用數字或字母標明其次序。

三、關於數字者

1.表中數字，以用亞刺伯字為宜。

2.表中數字，應對齊單位；其單位名稱應在標目下註明。

3.小數點（56.34）及分段點（450,783,421），必須標明並對齊。

4.特別重要之數字，可用不同顏色（如紅色表實支數超過預算數），較粗字體（如總數），不同字體（如意大利體）表示之。

5.無數字之空格，不可輒作爲「0」，應分別情形表示之，例如以「.00」表示計至二位小數止之百分數，而無有效數字者；以「—」表示實際上無數字者；以「……」表示本有數字而未據報者，以「××」表示無從獲得之數字（如淪陷區之數字）等是。

四、關於線格者

1.標目與數字之間、用較粗之線隔之。

2.縱橫各線，視其性質以定其距離及粗細。大項之間用粗線，距離較大；小目之間，用次粗線，距離次大；細節之間用細線，距離最小。

3.直行之線，均須劃出；橫列細節或小目之線，不須劃

出。

4.不相統屬之事實，用極粗線或完全隔離之雙線，貫通全表，以示隔絕。

5.表之上下邊線，用與表中粗細不同或形式不同之線劃之。或用最細線（表示事實之與他表相連續者，或時間上未終止者），或用最粗線，或用一組一細之雙線，（＝，＝）。

6.左右邊線，非必要時，不必劃出。（因通常均用開式表）

五、關於註脚者

1.關於表題、標目、數字、線格未能表明之意義，應加註脚，以說明之。

2.爲便於閱覽，標目之說明，可置於標目之附近。

3.每行每列之說明，應置於表之最下端或最右端，標以「備註」、全表表題，或數字之說明，應置表之下，標以「附註」，或說明甚長者，卽標以「說明」。

4,表題，標目或數字，或某文字，須加說明者，應於其右上角標以相當於「附註」或「說明」之符號，如☆，×，

米，半，(1)，(2)等是。

5.說明應特別簡潔，不失表列之功用。

第五節 統計數列

一、統計數列之意義

統計數列乃統計事項經分類整理後依次排列之一串數字，一變量決定一度量，他一變量亦隨之而決定一度量。換言之。表示一變量爲他一變量之函數，變量之任意變動者，謂之自變量，其隨自變數之變動而變動者，謂之倚變量。

一個現象之徵性甚多，通常須固定其餘徵性，以觀某個徵性之變化時，其現象之變化。例如研究人口現象，若固定空間，以觀各時期之人口數 (如民國成立以來我國歷年之人口總數)，則時間爲自變量，人口爲倚變量；若固定時間，以觀各區域之人口數(如民國三十二年六月三十日我國各行政區之人口總數)，則空間爲自變量，人口數爲倚變量；若固定時間及空間，以觀各業之人口數，(如民國三十二年六月三十日我國各業之人口數)，則職業爲自變量，人口爲倚變量。

統計數列與數學級數(數列與級數，英文均爲 Series)

不同之點。在於後者表示絕對不移之關係，如$y=a+bx$，當x決定一數值後，y必有一定之數值，即x與y有數學上之絕對關係，而統計數列則自變量與倚變量之間，只有經驗的關係，不過在時間數列和次數數列，通常亦可求出其近似之數學函數，以表示某種意義。（如長期趨勢之求法，曲線之配合等是。）

二、統計數列之種類

統計表製成之後，普通均表示一統計數列。統計數列可依統計事項分類標準之不同或謂依數列中自變量性質之不同，分為下列各類：

- 統計數列
  - 1. 時間數列（歷史數列）
    - 簡單時間數列
    - 累積時間數列
      - 以前累積時間數列
      - 以後累積時間數列
  - 2. 地理數列（空間數列）
  - 3. 屬性數列
    - 質量數列（依種類或品質分類者）
    - 數量數列（依數量分組者）
      - 枚舉數列（嚴格言之不得稱為數列）
      - 次數數列
        - 簡單次數數列
        - 分組次數數列
        - 累積次數數列
          - 以上累積次數數列
          - 以下累積次數數列

簡單時間數列

某校歷年畢業人數

年 份	人 數
總 計……	502
民國二十五年……	100
民國二十六年……	103
民國二十七年……	151
民國二十八年……	148

以前累積時間數列

某校歷年累積畢業人數

年 份	人 數
民國二十五年……	100
民國二十六年……	203
民國二十七年……	354
民國二十八年……	502

以後累積時間數列

某校歷年畢業人數

（以後累積）

年 份	人 數
民國二十五年……	502
民國二十六年……	402
民國二十七年……	299
民國二十八年……	148

地理數列

某年月日我國各註防區士兵人數

註 防 區	人 數
總 計……	
江 蘇……	
浙 江……	
⋮	

質量數列（依種類分）

某年月日我國各兵種士兵人數

兵種	人數
總計……	
陸軍……	
海軍……	
空軍……	

質量數列（依品質分）

某年月日某製革廠各級工人人數

等級	人數
總計……	
精工……	
半精工……	
粗工……	

枚舉數列

某年月日某班學生年齡一覽

學號	歲數
1	21
2	23
3	23
4	28
⋮	⋮

簡單次數數列

某年月日某班學生各年齡之人數

歲數	人數
總計	121
18	1
19	2
20	13
⋮	⋮

某年月日某班學生各組年齡之人數

年齡分組	人數		
	分組次數數列	以下累積次數數列	以上累積次數數列
總計	124	(121)	(121)
18—20	3	3	121
20—22	30	33	118
22—24	43	76	88
24—26	21	97	45
26—28	17	114	24
28—30	5	119	7
30—32	1	120	2
32—34	—	120	1
34—36	1	121	1

統計表之名稱，亦可以統計數列之名稱名之，如時間數列表，地理數列表，次數數列表（通稱次數表）等是。

三、統計數列之連續性

統計數列亦可依其自變量之連續與否，分為下列三種：

1. 自變量為連續的，稱為連續數列。如人口之年齡分配、身長分配、體重分配等是。

2.自變量爲不連續的，稱爲不連續數列。如書之售價，工人之工資，銀行之利率等是。

3.自變量爲可數的，稱爲可數數列。如書之頁數，字之筆數，人之口數等是。若可數之數有限，則數列之性質與不連續數列相似；若可數之數無限或相當大（如一國之人口）則又與連續數列相似。故通常謂統計數列可分爲連續的與不連續的兩種。

所謂自變量爲連續的，乃測量或計算之結果，爲近似值，而非絕對眞確之數，如言布一尺非眞一尺而毫厘不爽也。一尺與二尺之間，可分爲無限段，一寸與二寸之間，一分與二分之間，均可有無限段，任測量之技術，如何精確，均不能達絕對精確之目的。所謂不連續者，乃以人爲的方法，決定測量或計算的單位，如物價計至分止，工資計至元止，利率計至毫止是也。至於可數者，乃以自然或人爲的個數，爲計算之單位也，如人、馬（自然的）、書、公司是也。

無論自變量性質如何，其倚變量可爲連續的，不連續的，或可數的　如時間恆爲連續的，各飛機製造廠每年出產之飛機數則爲可數的；各石油礦每月出產之油量爲連續的；各工廠每日發放之工資數則爲不連續的。空間亦爲連續的，而

在各單位空間中觀察之結果，亦可爲連續的或不連續的。

連續與否，在統計上究竟有何意義？因連續變量的測量與乎不連續變量的計算，往往發生種種人爲的差誤，如年齡之報告，常集中在5,10之倍數上，薪資亦然；測量天然物如樹葉之長度，亦不免常得一分，半分之尾數。因此，吾人在設計調查或登記之時，固須注意，在整理及列表之時，尤應使數列近於理論的分配，而不呈特異之現象。

第六節　次數表之作法

時間數列，空間數列，質量數列（品質之分類，頗爲困難，須加以注意）之作法（卽分類、歸類及列表），只須將資料依前述方法分類排列，略別計算，卽已竣事。惟次數表之編製方法，頗有再加討論之必要。

枚舉數列，卽每數量發現一次，起下一次，實不得謂之數列，爲便於以後之討論起見，姑存此名。例如某年月日某校某班學生之年齡（已過生日之歲數）卽一枚舉數列也。

21	21	23	21	22	21	29	25	23
23	20	23	26	22	21	22	23	21
20	20	24	22	24	23	27	27	25

23	23	25	21	24	21	21	29	23
21	22	25	35	24	20	24	22	20
26	20	23	20	20	26	30	22	24
26	24	24	21	27	23	18	21	23
25	19	23	26	21	24	22	19	20
22	21	25	24	27	26	27	23	23
22	24	22	22	23	26	22	23	—
23	23	20	26	25	26	22	23	—
24	22	22	20	20	28	21	20	—
24	22	26	22	23	23	24	23	—
28	20	21	28	23	22	27	24	—

簡單次數表之作法，亦殊簡易，卽將所有量數(measure)之數值（ value ），由小至大排列，分別求其次數可也：

年　　齡	記　　號	人　　數
18	一	1
19	丅	2
20	正正下	13
21	正正正丅	17
22	正正正正	19

23	正正正正正	24
24	正正正	14
25	正丅	7
26	正正一	11
27	正一	6
28	下	3
29	丅	2
30	一	1
31		—
32		—
33		—
34		—
35	一	1

至於累積次數表（累積時間數列表同）之作法，亦甚簡易，將各組或各項次數，依次累積卽得。第一組照錄，第一二組次數相加得第二組之累積次數，第二組之累積次數加第

三組之次數，得第三組之累積次數，第三組之累積次數加第四組之次數，得第四組之累積次數，依此類推。

茲須詳加討論者，厥爲分組次數表。分組次數表的作法：(一)先從各量數中找出最大和最小的量數，以最大量數減去最小量數，其差謂之全距。(二)決定組數，即將全距分爲距離相等或不等的若干組。(三)決定組限，最小組限應小於或等於最小量數，最大組限應大於或等於最大量數。每組的距離（即上限與下限之差），謂之組距。(四)用記號法計算各組的次數。

以上述學生年齡爲例：

(一)最大量數爲35，最小量數爲18，

$$全距 = 35 - 18 = 17$$

(二)決定以兩歲（組距）爲一組，則

$$17 \div 2 = 9\text{（組數須取整數）}$$

(三)決定第一組爲18－19歲，第二組爲20－21歲（說明見下）

(四)求各組次數（即人數）

年齡分組	記　號	次　數
總計	×	121
18，19	下	3
20，21	正正正正正正	30
22，23	正正正正正正正正下	43
24，25	正正正正一	21
26，27	正正正丅	17
28，29	正	5
30，31	一	1
32，33		—
34，35	一	1

所當注意者，此項年齡曾經申明（見前）係已過生日之歲數，所謂十八歲係指已滿十八歲未及十九歲之意，同樣所謂十九歲，乃已滿十九歲未滿二十歲之意，其餘類推。上表寫法，乃便於整理時歸組起見耳，作分析工作時應改寫如下：（亦可謂調查及整理時將18看爲一距離，代表滿18歲未滿

19歲之一年，分析時，則將18變成滿18歲時之一點。)

某校某班學生各組年齡之人數

(某　年　月　日)

年齡組	人數
總計	121
18——20	3
20——22	30
22——24	43
24——26	21
26——28	17
28——30	5
30——32	1
32——34	—
34——36	1

分組次數表之編製，由上觀之，似甚簡易，但若遇較複雜之資料。則亦頗爲不易。玆將組距，組限決定之原則概述如下：

1.組距之大小　組距以相等爲原則，各組最好明白表示出來，非不得已時，不用「以上」或「以下」表示，此種組距，使統計分析甚爲困難。組數不可太多，亦不宜過少，以

10至25組爲宜。數列作成後，須視次數分配的情形，决定應否再分組。務使成合理之狀態。而無斷裂分立之情形。（參考第五章第四節繼續併組法）

2.組限之數值　分組求次數，乃假定次數在各組間均勻分配，而組中值可以代表該組各次數之量數者。故組中值，以在次數集中點上或十數系統（如10,20,30,100,1000等是）上爲佳。

3.組限之表示　組限之表示，整理表與報告表或計算表不同。整理表以便於歸組爲宜，計算表則以便於分析爲尚。舉例如下：

某項手續費計至分止

整理表	報告表
0——9.99	0——10
10——19.99	10——20
20——29.99	20——30
30——39.00	30——40
40——49.99	40——50
⋮	⋮

某項工資計至角止

整理表	計算表
5——9.9	5——10

10——14.9	10——15
15——19.9	15——20
20——24.9	20——25
25——29.9	25——30
⋮	⋮

可數數列之組限，在整理時不成問題，在分析時則恆須假定每一個體佔駐一單位距離。

例如字之劃數

整理時	分析時
1——3	.5——3.5
4——6	3.5——6.5
7——9	6.5——9.5
⋮	⋮

又如人數或元數相當大量時，則亦視爲連續數列如下：

整理時	分析時
1——1,000	0——1,000
1,001——2,000	1,000——2,000
2,001——3,000	2,000——3,000
⋮	⋮

總之統計分析之時，大抵視數列爲連續的，總以不失眞

象爲宜。例如書之頁數1—3，包含三個單位，.5—3.5 仍包括三個單位；1—3頁之組中值爲2, .5—3.5 之組中值仍爲2 也。至如上例 1—1,000本應寫爲 .5—1,000.5，1.001—2,000 應寫爲 1000.5—2,000.5，惟大量中，相差半單位，殊不足輕重，曷如寫作 0—1000，1,000—2.000等之簡便也。

第四章　圖示法

第一節　統計圖之功用

統計資料的陳示，除前章所述表列法之外，尙有圖示法。表列法雖可使數字井然有條，但仍不甚便於比較，亦不易於記憶；且以統計表說明事實，常嫌艱深，非有統計知識者不能完全領悟。一般人對於數字，總覺枯燥乏味，欲其發生興趣，非藉圖示方法不可。因為事實一經圖解，其全體分佈情形、徵性、以及其各部分之關係，均能同時呈現於目前，非但具有一般常識者一目了然，卽粗識文字之輩，亦能明其大意。

玆將統計圖之功用，分述如次：

1. 節省記憶，便於比較　統計資料，常有多種相互關係存乎其間，單憑理解記憶，並加比較，洵非易事。比較一二事物，其數量較小者，尙不甚難，倘數目太大或非整數，記憶比較，煞費腦力，若以圖形顯之，則差別立見，關係亦明，旣易記憶，又便比較。

2. 引起美感，增加興趣　圖乃有藝術意味之作品，觀圖

譬如賞花，能給人以視覺上之愉快而使發生美感。各種事業之宣傳或展覽，常用圖畫數幀，以表明其事業內容及發展情況，實因文字數字，常失之冗長而抽象，不若圖示之簡明而具體也。

3.表現集體現象　統計資料乃事實之總合，前已言之。但用文字數字表現，只能逐漸領會，而圖形則可將集體事實，整個呈於目前，不復支離破碎，使人顧此失彼。

4.集中觀者注意　以文字數字說明事實，使人倦於思索減少注意，難收預期之功效，若以圖形示之，則使觀者集中注意力，而增加傳達思想的效力。

為顯明統計圖之功用起見，更將圖與表之作用比較如下：

1.統計表用以排列及分析資料；統計圖則可表明已經排列或分析之結果，檢閱圖形，省時節力而可得明確之概念。

2.表之作用，在於陳示詳細之資料，以供專家之研究；圖之作用在於顯明簡潔之關係，以供民衆之閱覽。

3.表乃用數字表明確切而抽象之概念，圖則用符號綫條等顯示概略而具體之意義。

4.表可以表示各個事件之特質，圖則可以比較集體事實之徵性。

5.列表爲圖示前必有之步驟，故列表工作在前，繪圖工作在後。

鮑萊氏嘗謂平均法與圖示法乃初級統計學中之主要方法，凡須處理數字之學者或官吏，均應深切了解。大量而繁複之數字，無論表列如何明晰，吾人總難於理解其整個徵性。任何數列，如各城市之人口數，歷年之死亡率，若干工人之工資，歷年之進口貨等等，數列愈長，愈不易領悟。十年之數列頗易理解，二十年或尚不甚難；若有百年之時間數列陳於目前，一覽之餘，甚至毫無印象，爲山太大，見其樹而不見其林也。可否應用平均方法，須視平均數是否眞能代表，而使吾人最易理解。倘若可用平均法，吾人儘可選擇三個，四個甚至十個適當之數目（廣義的平均數）以表示數列之主要特徵。圖示法之主要目的亦然，無非欲陳示大數，使其整個徵性，顯而易見，應否用圖示法，亦僅視所繪之圖形能否與統計數列以最明白之顯示，使其特性一望而知。圖示法有一功用，爲平均法所無者，即時間數列之連續性，僅圖示法

可以充分表示之；但實際上平均數仍優於圖形，因平均數能單獨存在，與其所從計算之量數脫離，而代表各量數；欲與其他數列比較，亦非借助於平均數不可。自學理上觀之，圖形乃輔助工具而非主要工具，僅借以幫助眼力，節省時間；不用圖形，統計分析仍可按步進行也。（見Bowley：Elements Of Statistics, ch. VII）

第二節　統計圖之種類

統計圖之種類繁多，應用之機會以及圖式之變化，存乎一心，未可一概而論。依繪製之目的，可分爲陳示圖、說明圖、及計算圖三種；依其用途，可分爲書圖、掛圖及桌圖三種；依其所代表之統計數列，可分爲時間數列圖、地理數列圖、屬性數列圖（品質分配圖及次數分配圖），依其形式，可分爲統計形圖、統計線圖、統計地圖、其他統計圖四種。茲依最後一種分類方法，臚列如次：

單式長條圖
複式長條圖
分段長條圖
混合長條圖
對稱長條圖

- 統計形圖
  - 長條圖
    - 並列長條圖
    - 全距長條圖
    - 曲線長條圖
    - 百分長條圖
    - 延長長條圖
  - 平面圖
    - 圓形圖
      - 並列多圓圖
      - 疊置多圓圖
      - 帶扇圓形圖
      - 放射式圓圖
      - 百分比圓圖
    - 三角形圖
      - 並列三角形圖
      - 疊置三角形圖
      - 百分比三角形圖
    - 方形圖
      - 並列方形圖
      - 疊置方形圖
      - 百分比方形圖
  - 象形圖
    - 條式象形圖
    - 比例象形圖
    - 會意象形圖
  - 立體圖
- 次數線圖
  - 次數直方圖
  - 次數多邊圖
  - 修勻次數曲線圖
  - 累積次數圖

- 統計圖
  - 統計線圖
    - 歷史線圖
      - 簡單歷史線圖
      - 修勻歷史線圖
      - 累積歷史線圖
      - 距限歷史線圖
      - 帶紋歷史線圖
      - 山狀歷史線圖
      - 分歧歷史線圖
    - 對數線圖
      - 單對數線圖
      - 雙對數線圖
        - 歷史對數線圖
        - 次數對數線圖
  - 統計地圖
    - 數字地圖
    - 加點地圖
    - 影線地圖
    - 顏色地圖
    - 象形地圖
    - 標針地圖
    - 模型地圖
  - 其他統計圖
    - 組織系統圖
      - 上行組織系統圖
      - 下行組織系統圖
      - 集合組織系統圖
    - 工作程序圖
    - 工作進度圖

統計圖之種類，以上所舉，似乎應有盡有，實則變化無

窮，豈可罄述。應如何圖示，方可期達到理想之目的，要在乎各人領悟統計方法之原理，加之以幾何畫之技術，並以美術法則潤澤其間，則庶乎其可矣。

第三節　製圖法則

一　製圖之步驟

1.選擇圖式　視材料性質之不同，選擇適當之圖式

2.確定圖域　視圖式之不同，確定此圖形置於紙之何一部分，最爲完美妥貼。

4.置定値軸　須用値軸之圖，如條圖、線圖，値軸應置於適當地位。

4.劃分表尺　視數字之性質，於値軸分成等距或不等距之節段，是爲表尺，表尺之數値，務須能包括最大之量數。

5.繪製章底　用鉛筆起草。

6.審核圖稿　審核圖材、圖式、方法等有無錯誤或遺漏。

7.上墨着色　上墨注意線之粗細，墨之濃淡，着色注意色之深淺。

8.揩拭圖面　於墨色全乾後，用橡皮揩拭鉛痕及汚點，

其不能揩去者，用刀片刮去之。然後用指甲、象牙刀、或其他堅平之物，輕輕磨擦，使之平整。

9.校對全圖　材料有無錯誤，圖形已否完備，比較方法是否適當，表尺正確否，零點適宜否，零線、百分線及表示事實之線夠粗否，圖題、說明、附表等完全否，鉛筆痕跡已揩淨否，應留空地適宜否，其地一切均明晰簡要否，均應詳加核對。

二　製圖規律

1.圖式　為避免錯覺，通常圖式以線圖、條圖、地圖為主。

(一)表現時移事遷之材料，宜用歷史線圖。

(二)次數分配，宜用次數線圖。

(三)地理分佈，宜用統計地圖。

(四)表示比例關係，須用對數線圖。

(五)表示以上以下或以前以後之數列，須用累積線圖。

(六)表示組織及人員或工作程序，須用系統圖。

(七)宣傳或廣告用圖，宜着鮮明顏色，或劃粗大線條

或製象形圖。

（八）陳示一處，供人閱覽之圖，須視閱者之程度而定。

2.圖題　圖題以簡明扼要爲原則，但爲確能表示圖中所列之事實起見，亦不妨稍微繁複，分行排列。其位置以在圖之上下兩方爲宜、但有時亦可斟酌圖式而位置之。書法應自左至右。

3.基線及表示事實之線　線圖、條圖以縱軸，橫軸或中軸爲基線，在圓形圖則爲圓周，三角形圖爲三邊，方形圖爲四邊，其餘類推，基線應用粗線表示。表示事實之線，須用最粗線或着顏色。

4.表尺　自變量表尺之分度，應劃於各分組或各單位時間之交界處；倚變量表尺之分度應有適當之間隔，間隔太大，則顯示升降太甚，反之反是。爲便於閱覽起見，應作若干細線，垂直於各表尺，但線不宜過多，多則刺目。對數線圖上，對數表尺，應於自然數1,2,……10及10之任何倍數處作指導線，並於10的任何次方處作較粗線以醒眉目。除非有特殊困難，倚變量表尺上之零線（通常爲橫軸）必須繪出，始

便於比較。對數尺度上，若仍書自然數，則似無零線，因自然數1其對數爲零，實則關於對數仍有零線。零線既必須繪，如遇特殊情形，表示事實之圖形或線，距離零線太高，難於繪出，得作二齒線，割去全圖之一段，但非不得已時不宜應用齒線。

5.差異之顯示　在圖上各事物之差異程度，除以實地，條紋、花紋、格子、斑點、虛地等顯明之外，更可以着色與否，不同顏色、顏色濃淡等表示之。事物之重要者，多着深色、上紅色、繪實地，次要者多着淺色，繪虛地。

6.數字　數目排列應由下而上，由左而右，有時上下倘可顛倒，而左右則不宜變更。如數字太長時得略去末後若干位數或零位，而於附註中說明之。

7.說明及附表　統計圖應與有關之說明及附表栢接近，以便參照閱讀。

8.其他　圖材來源，繪製日期及繪製機關等，均應分別書明。圖材來源對於研究者之參考，至關重要。機關年月以書於圖之右下角爲妥。

（附錄一）1915年美國圖示標準法則聯合委員會規定圖示

標準法則十七條，玆附錄於後，以備參考。

1.圖的排列，須從左而右。

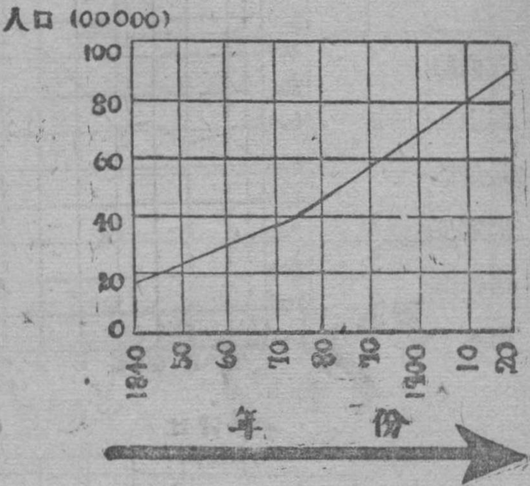


圖1.

2.能用直條代表數量最好，因爲面積與體積均易引起誤解。

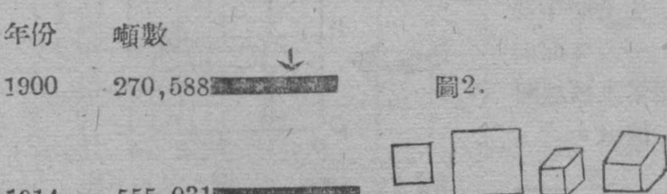
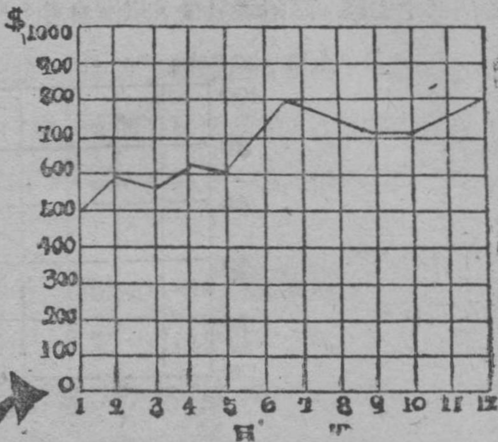


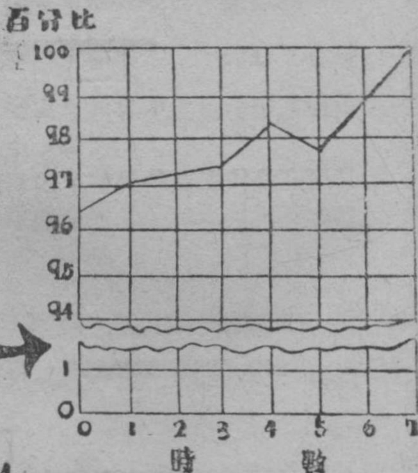
圖2.

銷售數



3.在曲線圖上，最好能將零度橫線繪出。

圖3.



4.如零度橫線不能照常表示在圖上，即當用一橫斷的裂痕截去全圖中間一段，而將零度橫線圖繪出。

—— 圖4.

5.零度線應與其餘各線有顯明的區別。

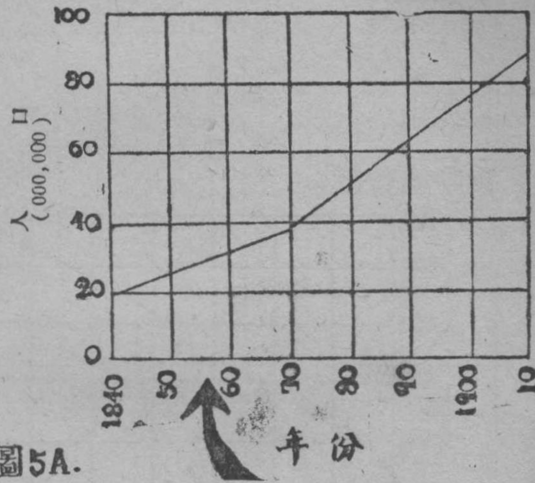


圖5A.

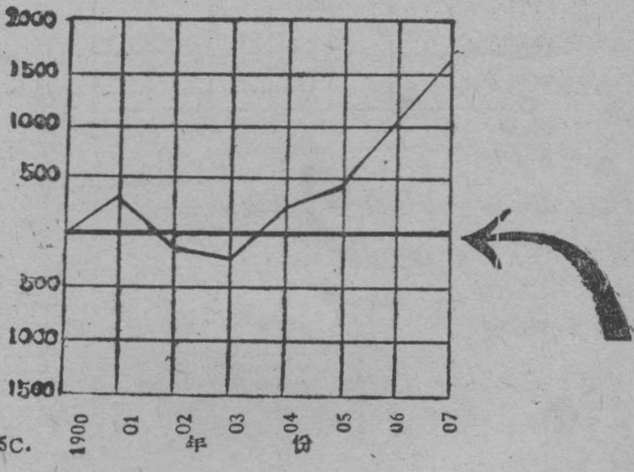


圖5C.

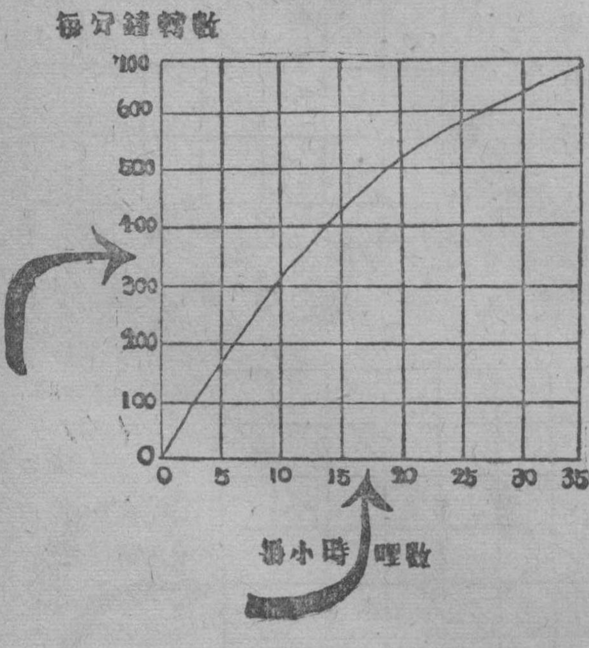


圖5B.

6.如曲線之表尺用百分比表示凡百分線宜較顯明，其他用作比較之線亦宜較爲粗大。

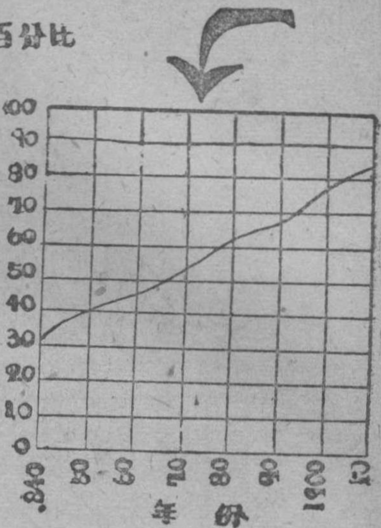


圖6A.

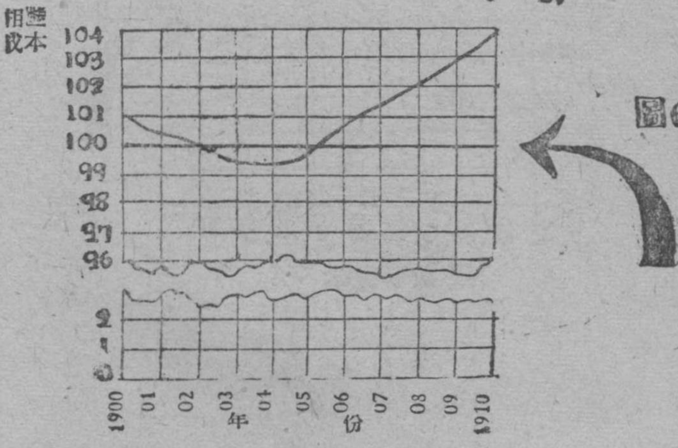
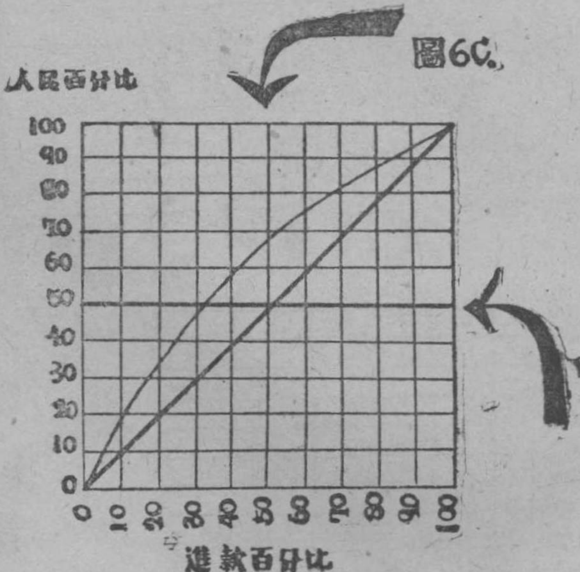


圖6B.

圖60.


人民百分比
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
90
100
進款百分比

7.表示時間數列之圖，其所表示之時間不完全者、首線及末線不必特別區別因此二線並不表示時代之始終。

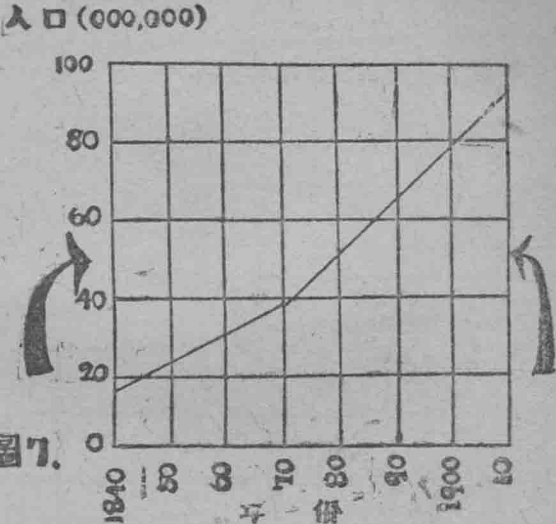
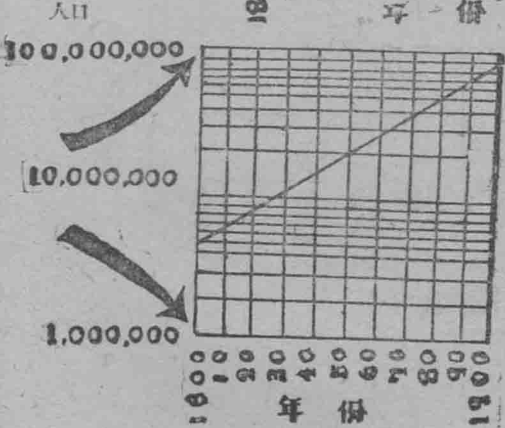


圖7.



8.對數圖上，上下沿線，應在對數表尺上關於自然數爲十乘冪處。

圖8.

9.縱橫綫除爲醒目之必要外，不宜太多。

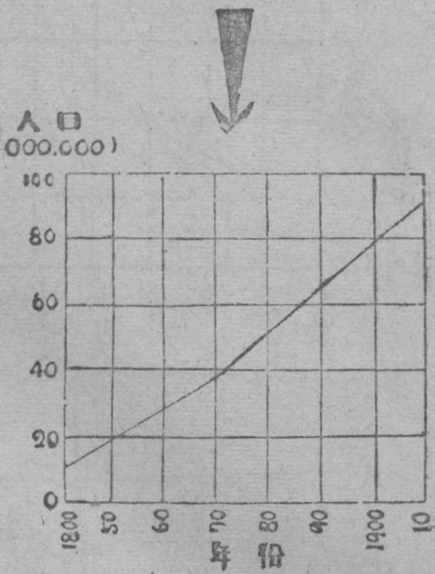


圖9A,

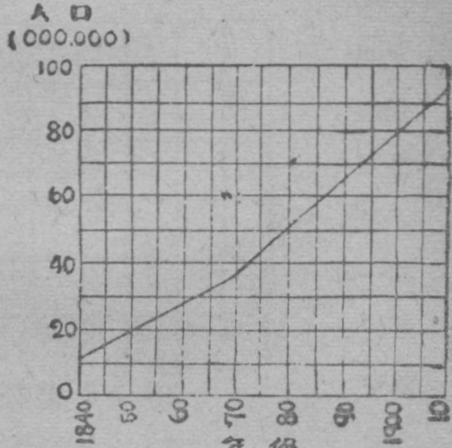


圖 9 B.

10. 圖上曲綫與行格應有顯明的區別。

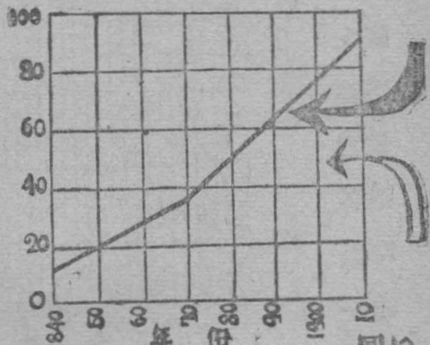


圖 10.

11. 如果曲線是表示連續一系的觀察，最好由圖上各點切實將各種觀察分別表明。

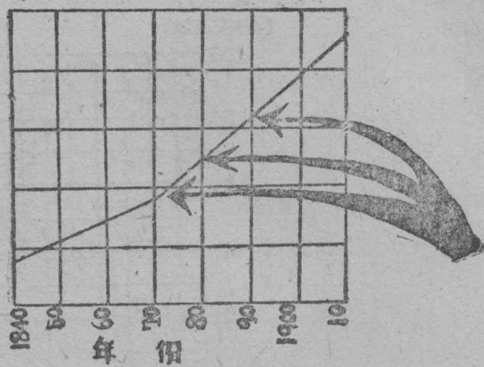


圖11A.

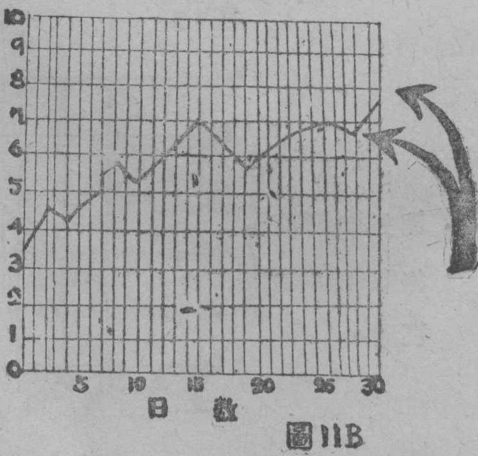


圖11B

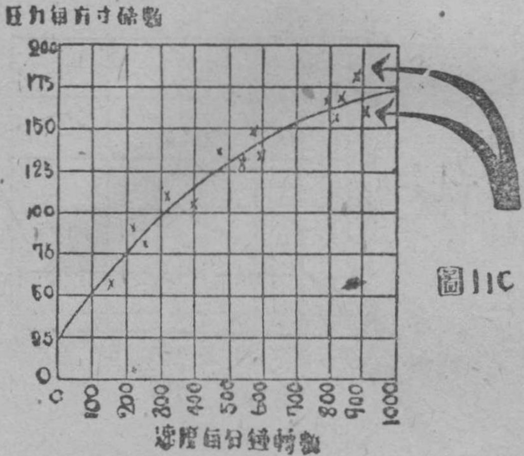


圖 11C

12，橫表尺須自左至右讀之，縱表尺須由下而上讀之。

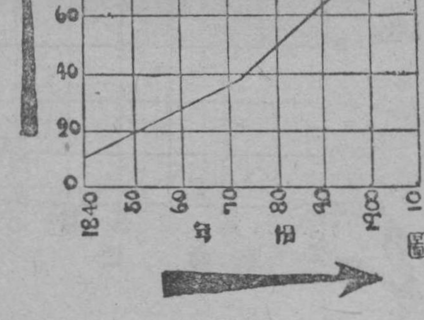


圖 12.

13. 表尺上之數字須於縱軸之左及橫軸之下或沿列於有關係之軸上。

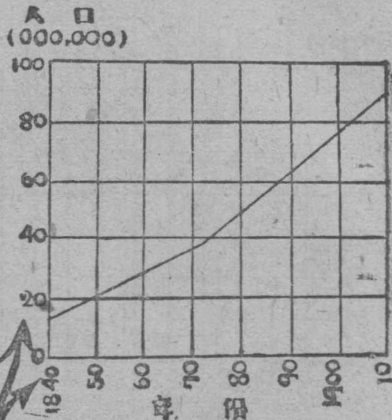


圖13A

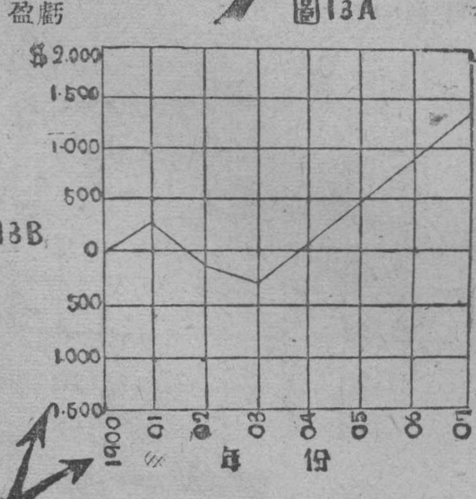


圖13B

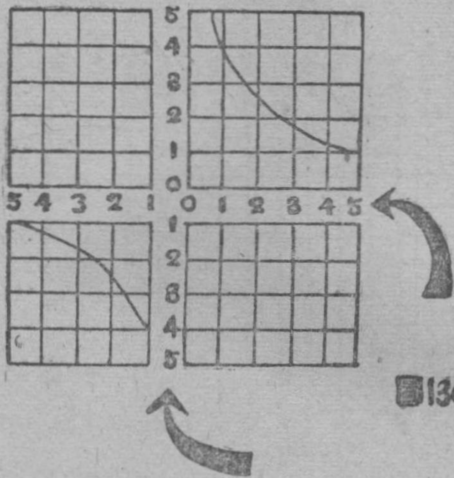


圖13C

14. 圖線所代表之數量或公式常須列入圖中。

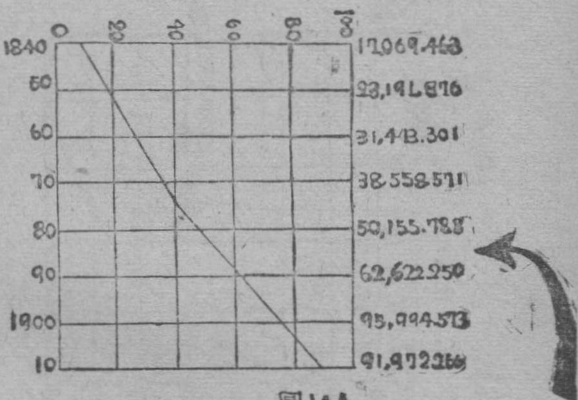
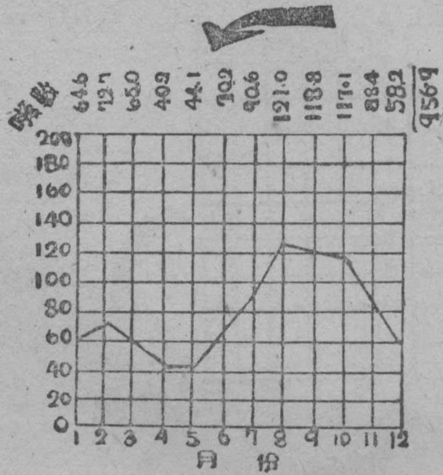


圖14A.

圖14C


64.5
72.7
65.0
40.2
44.1
70.2
90.6
121.0
118.8
117.1
88.4
58.2
956.9
200
180
160
140
120
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12
月 份

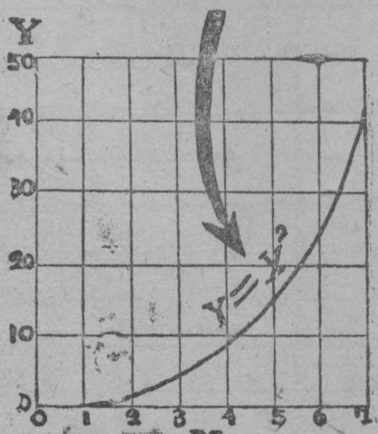
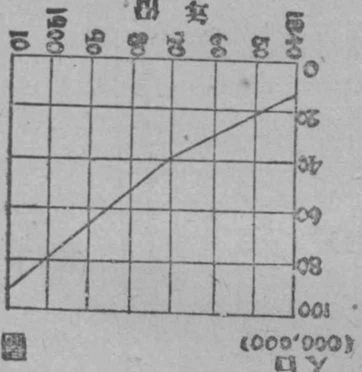

Y
50
40
30
20
10
0
0 1 2 3 4 5 6 7
X

圖14C

15. 設數字材料不便列入圖中，可另作一表附列之。



年份	人口
840	17,069,453
50	23,191,876
60	31,443,321
70	38 558,371
80	50,155,783
90	62,626,250
1900	75,994,573
10	91,972,266

圖15.

16, 圖上所有文字及數字，均宜置於圖下或圖左，以便閱讀。

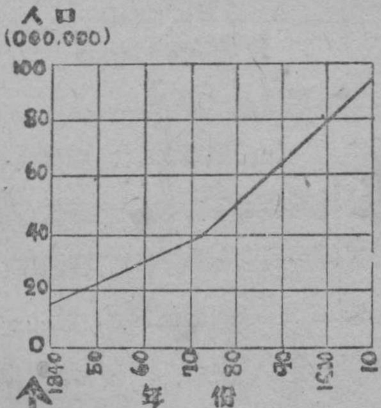


圖16.

17. 圖之標題宜清晰完備，如遇必要，不妨加以標目及說明，務使讀者一目了然。

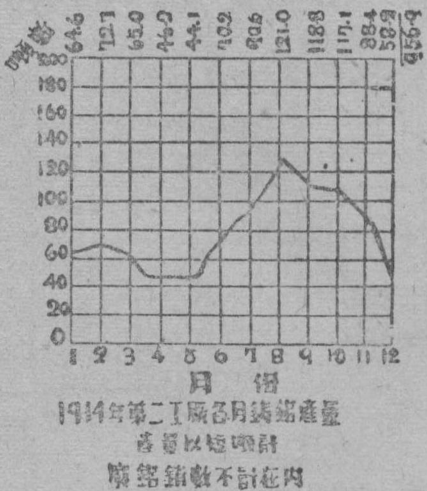


圖17.

〔附錄二〕製圖儀器及一般用品

1.製圖儀器　(一)鴨嘴筆(二)兩脚規(三)分劃規(四)三角板(五)雙綫鴨嘴(六)丁字尺(七)圖板(八)分度器(九)縮放器(十)虛線規(十一)曲線規(十二)齊邊器(十三)直尺

2.一般用品　(一)鉛筆(2H至 6H)(二)鋼筆(三)毛筆(四)圖畫紙(五)墨水(六)色彩(七)橡皮(八)一面口剃刀(九)法蘭絨(十)麵包(十一)圖釘(十二)小刷(十三)筆池(十四)刷帚

第四節　統計形圖

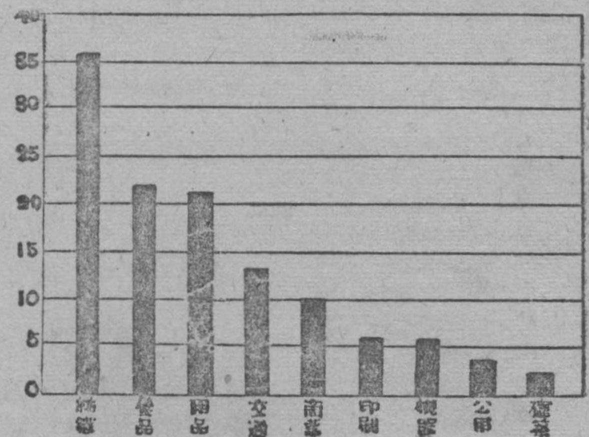
以幾何圖形(點，綫條，面積、體積)或實物圖形(人物)陳示統計資料，謂之統計形圖。目力對距離之長短差別之判斷力，較強於對面積與體積大小之判斷力，故長條圖實優於平面圖與立體圖；至象形圖乃以實物之形狀，吸引人之注意力，若欲表現較爲正確之大小觀念，則又遜於立體圖矣，故除在展覽會中，宣傳品上以及廣告上外，甚少見之。茲將統計形圖略舉數例如後以供參考：

一、長條圖

1.單式長條圖　依量數之大小，比例作成長條，依次排列之，縱列橫陳，均無不可。下圖爲縱列之單式長條圖。

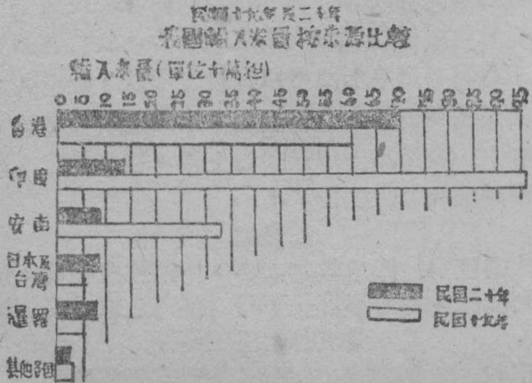
某年某市各業罷工案件

罷工次數　　　　　　　　單式長條圖

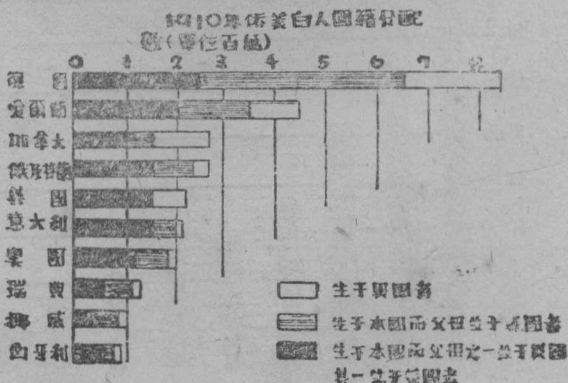


2.複式長條圖　用兩條或兩條以上之長條，合為一組，若干組並列之，是為複式長條圖：

複式長條圖

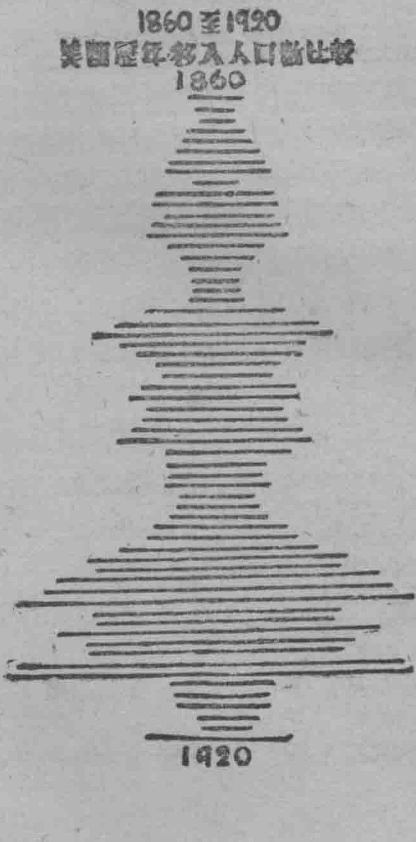


3.分段長條圖　單式長條圖，若將每條構成之情形，分段表明之，是為分段長條圖。



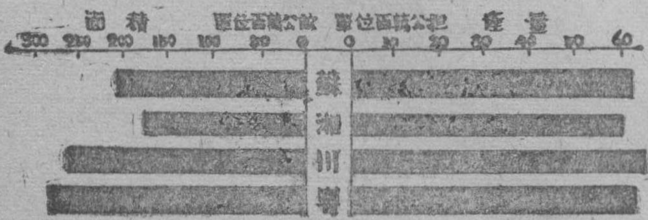
4.混合長條　即分段之複式長條圖

5.對稱長條圖　將各長條作成對稱形式。此種圖形宜於表現長期間之變化，亦可分段。



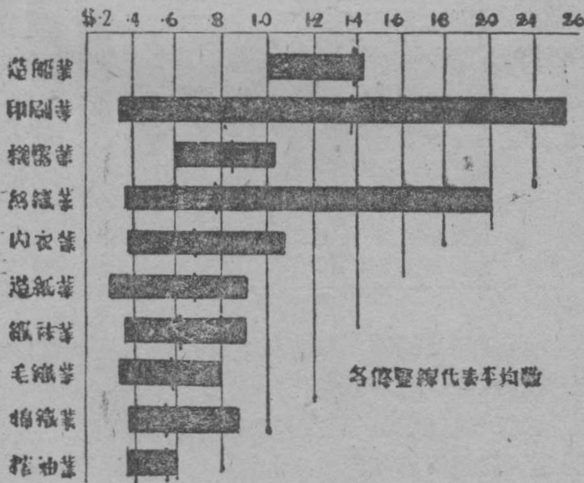
6.並列長條圖　將兩個單式長條圖並列成「非」字形。如男女年齡分配及歷年出入口貿易額，可用此種圖式。

民國二十五年蘇湘川粵四省棉產面積及產量圖



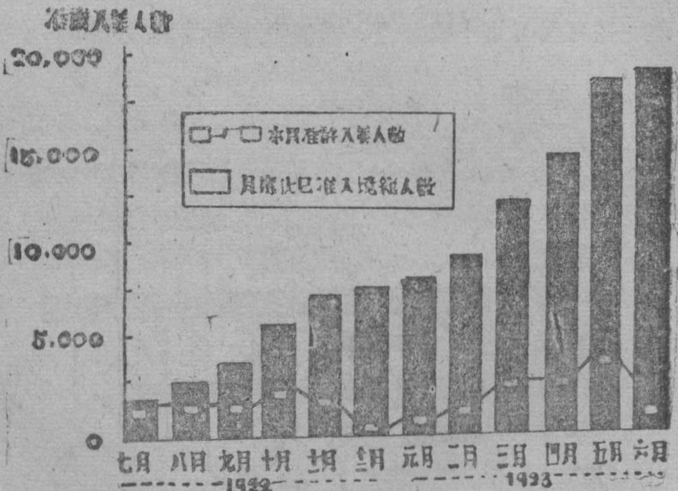
7.全距長條圖　即以長條之長短表示每一變量全距大小之圖形，此種圖形不僅可表示變量之最高值，最低價，平均值。並可示其集中或散漫，偏低或偏高。

民國二十四年上海市各業工人每日工資率比較圖



8.曲綫長條圖　即同項目之長條間，更以曲綫聯綴之圖形也，亦有稱條綫混合圖者。此種圖形多半在時間數列欲表示某項目之升降變遷時用之。

瑞典人民准許入美人數按月比較圖(1922年7月—1923年6月)



9.百分比長條圖　即等長之分段長條圖，每條各段分別代表其構成分子之百分比。

10.延長長條圖　單式長條圖中有特別長之一二條折斷延長之，與用齒綫法之意義相同。

二、圓形圖

1.並列圓形圖　以兩個或兩個以上之圖面積表示大小不同之事實，各圓半徑與各量數平方根成比例　　$A=\pi R^2$

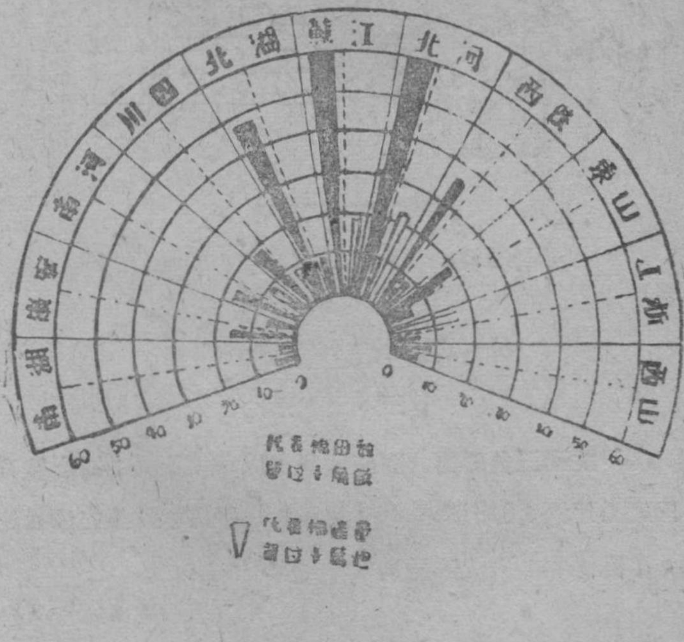
2.疊置圓形圖　同心疊置或同點疊置，面積相疊易使視覺混亂，引起誤會，不宜多用。

3.帶扇圓形圖　依事實之分項實數，比例圓周度數，分別表現之。多圓並列亦可。

4.放射式圓形圖　利用鐘表形，表示十二個月或廿四小時之內變化。

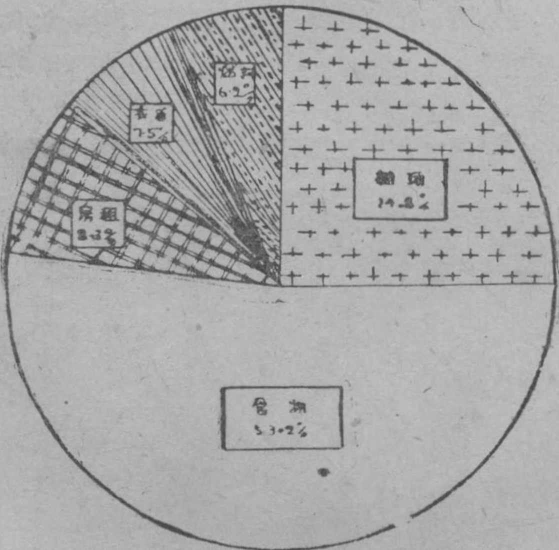
5.扇形圖　即以圓形一部之面積代表數量大小之圖形。此種圓形或爲並列之若干扇形，或爲一大圓形中劃分若干小扇形。

民國二十四年我國各省棉田及棉產比較圖



6.百分比圓形圖　將圓周分為二十等分，作成百分尺度將事實構成之百分比分別表示之。

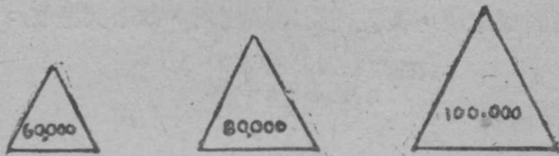
民國十八年四月至十九年三月
上海工人家庭生活費分配



三、三角形圖

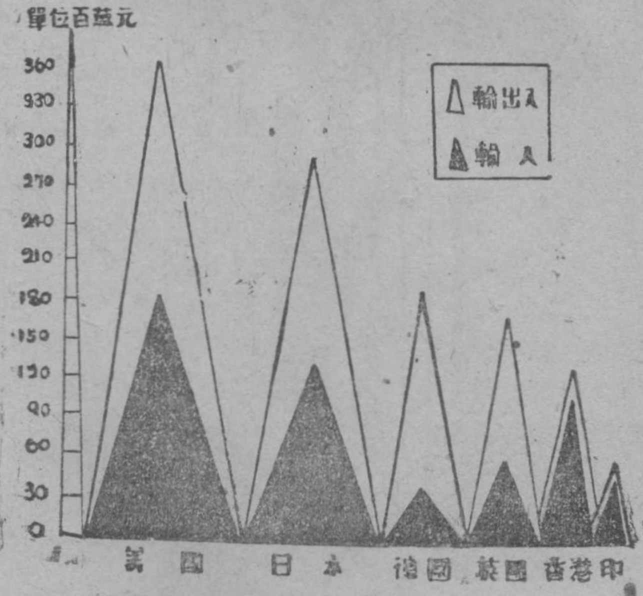
1.並列三角形圖　同底並列，以高為比例；同高並列，以底為比例；相似形並列，以量數之平方根為邊之邊為比例。下圖係相似三角形並列。$A=\frac{1}{2}hb$

甲乙丙三城居民人數比較



2.疊置三角形圖　同底、同高或同角圖疊置均可，其缺點同疊置圓形圖。

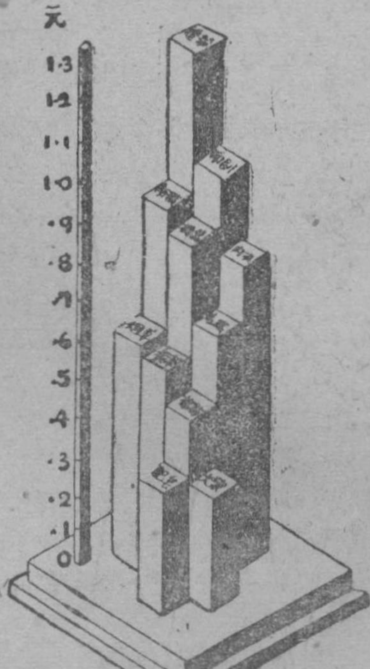
民國二十五年我國輸出入淨值國別圖



五、立體圖

平面兩度，尚不易辨明其大小差別，立體三度，自更難察覺。惟同底比較，與長條圖功用相同，如下圖是也。

民國二十三年上海市主要
平均每日工資率比較



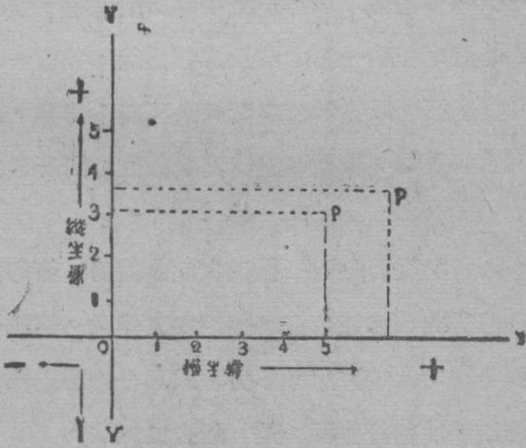
六、象形圖

以實物形之大小，代表數量，雖易引人入勝，但不易表示眞確數量。

第五節　統計綫圖

統計綫圖亦名坐標圖。

二直線垂直相交，橫綫X'X 謂之x 軸，縱線 YY'謂之Y 軸，其交點O 謂之原點。自平面上任一點P



，作線垂直於x 軸，其長度謂之P 之縱坐標，再自P 作線垂直於y 軸，其長度謂之P 之橫坐標。任一對坐標値，決定一點。各點連綴卽成坐標綫。等組距分組次數數列及時間數列用統計線圖表示，最爲相宜。

一、次數線圖

次數綫圖乃用以以表示等組距之分組次數數列。以次數所

由生的變量爲自變量，次數爲倚變量，自變量以x 軸爲表尺，倚變量以y 軸爲表尺。

1.次數直方圖　在每組上下組限上樹立直線，其長度等於該組之次數，每組連綴各直綫之頂端，卽成次數直方圖。

2.次數多邊圖　以每組組中點爲橫坐標，該組之次數爲縱坐標並假定在第一組之前及末組之後，各有一組，各以組中點爲橫坐標，以○爲縱坐標。求平面上各點，連綴各點，卽成次數多邊圖，次數多邊圖與次數直方圖所包含之面積相等。

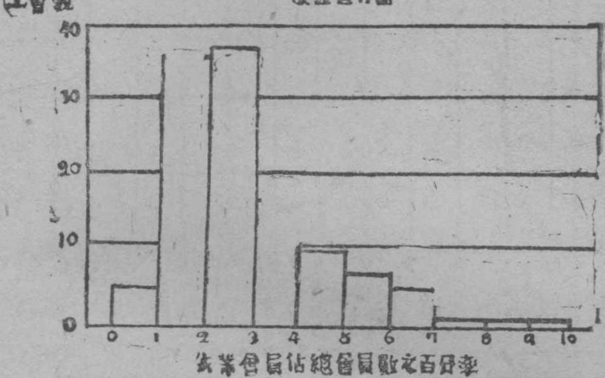
3.修匀次數曲線圖　簡稱次數曲線圖。修匀方法有二：(1)在直方圖或多邊圖上，隨手去其稜角，使成平滑曲線，謂之隨手修匀法(2)用數學方程式配合理論坐標值，根據理論坐標值用曲線規作圖，是爲配合曲線法。

4.累積次數圖　根據累積次數作圖，亦有直方，多邊，修匀三種，修匀一種爲用較廣，卽所謂累積次數曲線圖。視次數之累積爲以下或以上，又可分爲以下累積次數圖，與以上累積次數圖兩種。

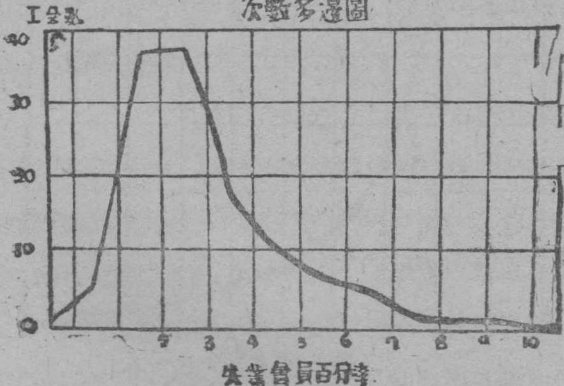
玆將各種次數線圖舉例如下：

一百二十家工會失業會員百分率之工會數分配

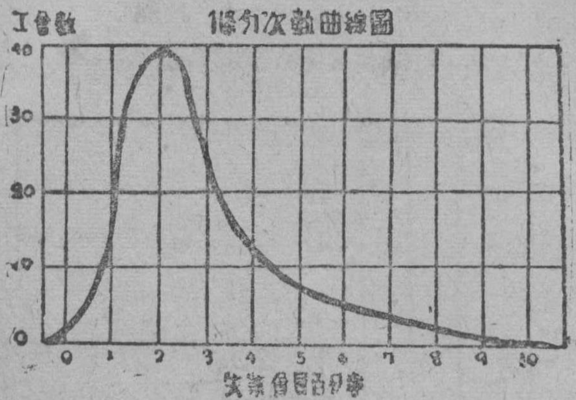
次數直方圖



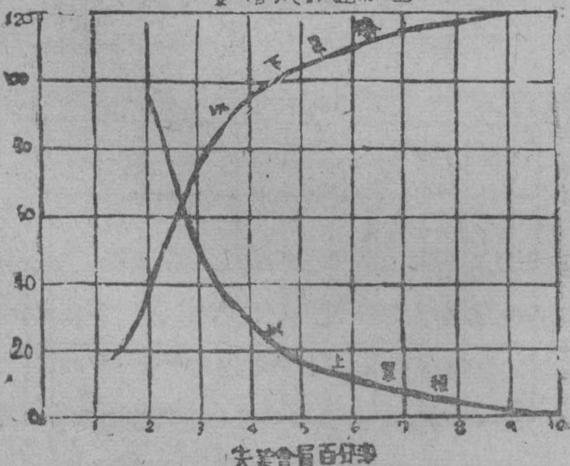
次數多邊圖



修勻次數曲線圖

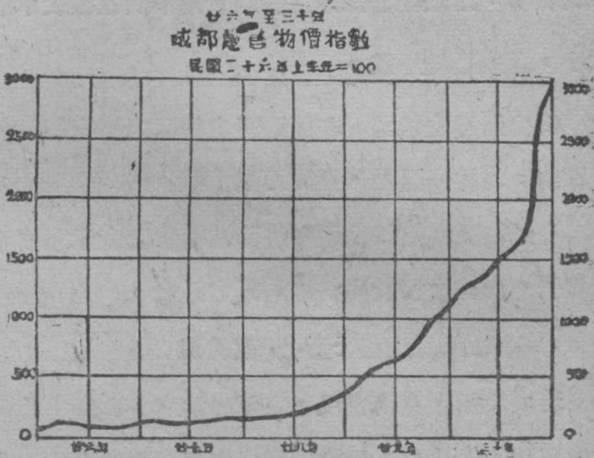


累積次數曲線圖



二、歷史線圖

歷史線圖乃用以表示時間數列者，通常以橫坐標表時間，縱坐標表各時間之量數。作法與次數線圖同。惟左右兩端均不用粗線，以示數列所代表之時間，僅在悠長時間中截取一段，而非時間之始末也。



材料來源　金陵大學農業經濟系經濟週訊

1.簡單歷史線圖　以每一時期之中點爲橫坐標，以該時期之量數爲縱坐標，在平面上取定各點，依次連綴各點爲折線，即爲簡單歷史線圖，與次數多邊圖相似。

2.修勻歷史曲線圖　相當於修勻次數曲線圖。修勻方法三有(1)隨手修勻法(2)移動平均法(3)配合曲線法。

3.累積歷史線圖　相當於累積次數線圖，

4.距限歷史線圖　以距限表示兩時間數列之差別。

5.帶紋歷史線圖　將某時間數列各個量數依其構成部分分別表明，連綴相當部分之各點，用不同花紋表示各部分之大小，遂成帶紋形狀。

6.山狀歷史線圖　在簡單歷史線圖之事實線與各邊線之範圍內，均塗以黑墨，全圖乃成山狀。

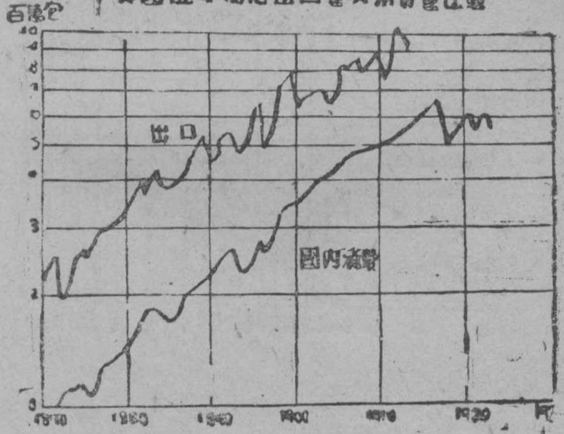
7.分歧歷史線圖　用多邊圖或曲線圖表示事實正負相反之狀況，如物價指數與貨幣購買力是也。

三、對數線圖

若表尺採用算術表尺，表示事實之絕對數（自然數），只能表現等差變化；對數線圖之表尺依對數分度，是爲對數表尺，可以表現事實之比率變化。對數圖表尺之分度雖依對數，但分度點上仍寫自然數。凡有比率關係之次數數列或時間數列，均宜用對數線圖。

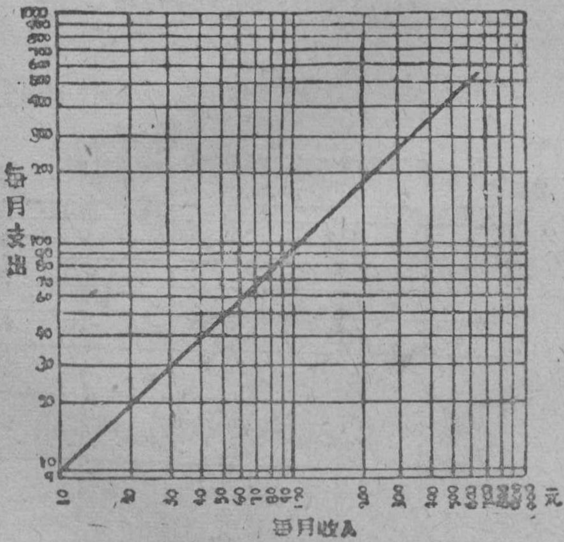
1.單對數線圖　以X軸（自變量）爲算術表尺，y軸爲對數表尺。此種對數圖應用最廣。

美國歷年棉花出口量與消費量比較



2.雙對數線圖　縱橫坐標均用對數。

某地人民平均每月收入及支出分配



第六節　統計地圖

統計地圖乃表示地理數列者，其優點可使事實發生之地域，量數在地域上的分配均能表出之。

1.數字地圖　將地理數列之量數分別寫在各地域上。

2.加點地圖　以大小不同之點，或大小一律而帶扇形之點，或大小一律數目不同之點表示各地域之量數，第三種方法最好，下圖是其例也。

我國省區人口比較

3.影線地圖　以線之稀密表示分佈之稀密，最宜於顯示事實之密度。

4.顏色地圖　用不同之顏色，代表量數之大小，形象美觀，展覽宣傳最爲有效。但印刷麻煩，通常以影線圖代之。

5.象形地圖　以事物之具體形象繪於各地域上，如全國農產物之分佈是也。

6.標針地圖　用插針方法表示統計事實，如以針之密度表示分佈，以針連線代表某種區域是也。常用於大掛圖上，

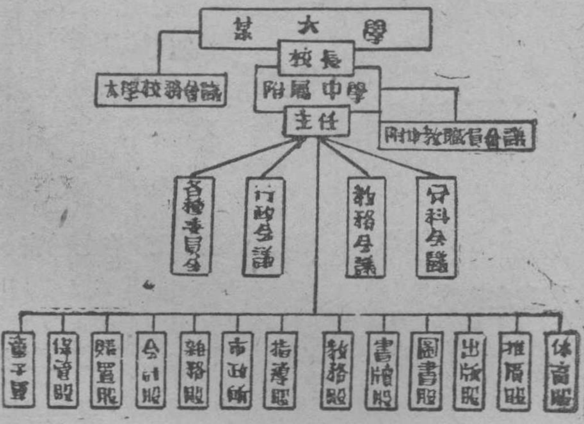
其優點在於可以隨時更改。

7.模型地圖　用模型表示山川之形勢，物產之分佈。展覽會中多用之。

第七節　其他統計圖

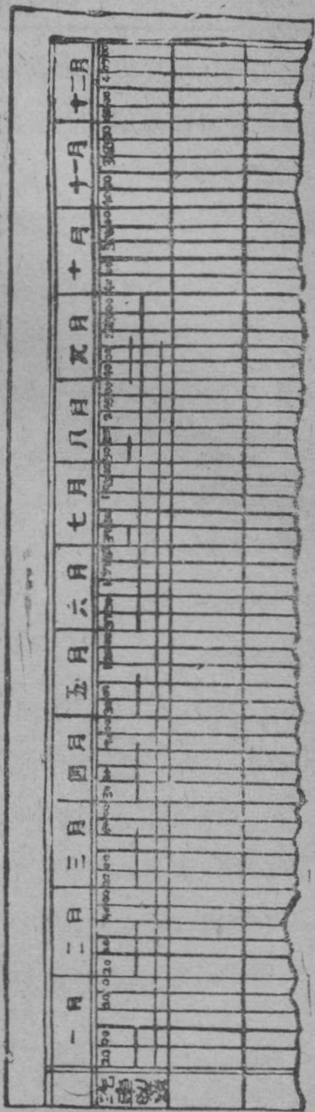
根據法規表示機關團體組織之系統或工作之程序，是爲系統圖。最宜注意等級高低及先後次序，茲舉下行組織系統圖一例，以供參考。

某大學附屬中學之組織系統



至於工作進度可用條圖、線圖、乃至地圖表示之，凡工作計劃一經決定，卽可用此圖，按照工作計劃所規定者，以核對實際之工作。此種圖形之式樣甚多，其普通者爲剛德氏工作進程圖，如下圖所示。該圖於每月之左首登記該月份

預計完成之工作，右首登記預計完成工作之累積數；而後以細線表示每月之實際工作數，粗線表示該期中實際工作之累積數，若某月之實際工作俱等於每月之預計工作，則細線全越各月之空格，若某月之實際工作超過該月之預計工作，則該月空格內當有兩條細線。惟必須注意者每月所分格數雖同，而每月所代表之數量隨該月預計工作而異，如正月份預計工作爲製造汽車二千輛，則每格代表四千輛，八月份預計工作爲製造汽車五千輛，則每格代表一萬輛。工作進程圖爲管理工作進行最有功



某公司汽車製造進程圖

用之工具，可以隨時比較實際完成與預計之工作，故工商企業組織暨行政機關用之者甚多。

第五章　平均數

引　論

一、以前各章提要

以前各章自成一段落，第一章敍述何謂統計，何以要辦統計，統計如何辦法，以及辦理之人員應具備若何條件。第二章提示如何從原始來源或次級來源搜集資料，並應如何加以審核，整理，登記與保管。第三章說明統計資料之科學分類方法，分類之後應如何列表，使繁復之事實，井井有條，第四章敍述已經列表之資料，如何用圖形陳示。舉凡第二章至第四章所敍統計資料之搜集、整理、表列、圖示諸方法，可以「統計技術」概括之。後此所論各種方法，則均可稱之爲「統計分析」　政府之一般統計工作，只須了解統計技術，卽差足應付；惟欲技術之應用，不致鑄成大錯，則不可不略知分析方法；欲求更進一步之研究，則又非本書所能詳述者也。

二、統計分析之目的

統計技術討論如何取得及排列統計資料之程序，統計分

析則計算統計常數以描寫此繁複之資料；統計技術說明統計數列如何建立，統計分析則尋求數列內部之關係或其將來之趨勢；技術所以取材，不厭其備而詳；分析乃在成器，務求其精而簡。

統計資料、經分類整理之後，可分成時間數列，地理數列以及屬性數列，前已言之。各種數列所代表之自然或社會現象，靜態或動態，類皆因果複雜，關係幾微。欲充分顯明其特性，尋求其規律，推斷其趨勢，使複雜者綱舉目張，幾微者條分理析，必有特種方法加以研究，良非表列圖示所能盡其功效也。

第一節　平均數之意義與種類

一、平均數之意義

平均數者，代替變量之統計常數也。或謂統計學乃平均數的科學，或謂統計學乃研究大數的科學，而應用平均數可以取一二簡單數字，代表繁複之數羣或大量之數目，則上述二定義合而為一矣。

此統計數列或彼統計數列，欲其可以比較，非假此均加以撮要不可。吾人凡遇統計事項，往往希望得一平均數：物

價千萬，平均物價如何？學生一班，平均成績如何？薪給有高低，平均薪給如何？農事有忙閒，平均狀況如何？農產有豐歉，平均產量如何？諸如此類，不勝枚舉。

平均數之觀念，雖則至爲普遍，而對於平均數有明確之認識者太少，致常不免於誤用，若干人往往因對於某項事實，無深切了解，輒憑藉一二特例，概括言之，曰平均而論，大抵如此如彼。如云某機關平均待遇甚低，某校平均成績甚優，其實未嘗平均，就其直覺所得而言，如非其頭腦不科學，即其人必有某種偏見也。吾人應用平均數，切宜愼之。

二、平均數之功用

1.用簡單數字，表示複雜數列。如林中之樹，若逐一舉其高度，吾人不能明確認識，若謂此林之樹，平均高度爲幾丈幾尺，則言者易言，聽者反易於領悟矣。

2.使此羣事項可與其他事項比較。兩個或兩個以上之機關，欲知其待遇究竟孰高孰低，不能單視規定標準，亦不能任取一二人比較，必須分組比較其平均薪給。

3.使抽查結果可以代表事實眞象。欲問我國人民與日本人民孰高孰矮，豈必將萬萬千千之人一一量之，只須確定標

準，擇取百數之樣本，加以測量，求其平均數即可比較矣。

4.用明確數字代替模糊觀念。平常但知此林之樹，高於彼林；此機關之待遇，高於彼機關；中國人高於日本人；但欲問究竟相差之程度如何，非借助於平均數不可。

三、平均數之種類

平均數之種類繁多,通用者有算術平均數,中位數,衆數,幾何平均數四種，倒數平均數有時亦用之。玆將其定義述之如下，至其特性，用途與計算方法，分別敍述於以後各節。

1.算術平均數　以量數的項數除量數數值的總和，其商謂之算術平均數。以M表之。

$$\text{算術平均數} = \frac{\text{量數的總和}}{\text{項數}}, \quad M = \frac{\Sigma X}{N}$$

2.中位數　全體量數順大小次第排列，恰居正中位置之量數，謂之中位數。以Md表之

$$\begin{array}{ccccccc} 3 & 5 & 6 & 8 & 10 & 12 & 15 \\ & & & \uparrow & & & \\ & & & M_d & & & \end{array}$$

3.衆數　發現次數最衆的數值，謂之衆數，以Mo表之。

家庭人口	家　數
總　計	100
1	2
2	3
3	5
4	10
5	25
Mo→6	38
7	12
8	5

4.幾何平均數　設N爲量數的項數，各量數乘積之N次方根，謂之幾何平均數，以G表之。

$$\text{幾何平均數}=\sqrt[N]{\text{第一量數}\times\text{第二量數}\times\cdots\cdots\text{第N量數}}$$

$$G=\sqrt{X_1\cdot X_2\cdot\cdots\cdots X_z}$$

5.倒數平均數　各量數的倒數的算術平均數的倒數，謂之倒數平均數，以H表之。

$$\text{倒數平均數}=\cfrac{1}{\cfrac{\cfrac{1}{\text{第一量數}}+\cfrac{1}{\text{第二量數}}+\cdots+\cfrac{1}{\text{第N量數}}}{\text{項數}}}$$

$$H=\cfrac{1}{\cfrac{\cfrac{1}{X_1}+\cfrac{1}{X_2}+\cdots\cdots+\cfrac{1}{X_z}}{N}}=\cfrac{1}{\cfrac{\Sigma\frac{1}{X}}{N}}=\frac{N}{\Sigma\frac{1}{X}}$$

算術平均數，幾何平均數與倒數平均數，如上所述，其定義均可以數學公式表示之，故統稱有數學函數關係的平均數；而中位數與衆數，均不能用數學公式表示其定義，故稱無數學函數關係的平均數。

平均數種類旣多，應當用何種平均數最爲相宜，須視(一)計算平均數的目的（二）統計數列的性質（三）應用者關於各種平均數之認識（四）應用者對於統計對象之精通程度而定，未可一概而論也。

第二節 算術平均數

一、算術平均數之用途

算術平均數乃最普通之平均數，尋常所用平均數，均係指算術平均數而言。應用算術平均數之目的不外兩種：

1.避免大數 如謂某學生八門功課之成績共計 600 分，不如謂其平均成績爲75分，因75分與給分標準可以比較也。又如欲比較兩時期某種商品之出口總値，與其謂民國元年至民國十年爲100,000,000元，民國十一年至民國二十年爲110,000,000元，不如謂前者平均每年爲10.000,000元，後者平均每年爲11,000,000元。

2.劃一單位　十年與十年此較，尚可應用總數，若爲期長短同，不更非平均不可。如民國元年至民國十年爲100,00,000元，而民國十一年至民國二十一年爲132,000,000元，倘不各以年數除之。兩期相差程度，全無所知。迨分別平均以後，乃知兩期每年爲10,000,000元與12,000,000元。機械學上，此種意義之算術平均數，應用最廣，如言每平方吋平均壓力，某機器每分鐘平均工作，某列車每小時平均速率等是。每年平均之單利率，亦此類也。

百分數之應用，其目的亦在劃一單位，以便比較。譬如比較人口之增加，或商業之繁榮，若但舉總數，殊無意義。重慶增加50,000人，遠不如南溫泉增加10,000人之重要，若謂重慶增加百分之十，南溫泉增加百分之二十，則意義甚顯矣。百分數者每百單位之平均數也。通常出生、死亡、婚姻等均以比率表之，而謂每千人若干；在此種情況下，且係雙重平均數，以其表示每年每1000人中若干也。

又尋常所云「每人」若干，亦係劃一單位之一法。如欲比較民國元年與民國二十九年某市人民對於酒之消耗，僅言民國元年爲十萬斤而民國二十九年爲十一萬斤，並不足以證

明前者少而後者多，若謂每人爲3.45斤與1.26斤，則彰彰明矣。

二，枚舉數列之算術平均數

由前述定義：

$$算術平均數=\frac{量數的總和}{項數}，即M=\frac{\Sigma X}{N}；$$

或作：算術平均數×項數＝量數的總和，即$M\cdot N=\Sigma X$

其中M表示算術平均數N表示項數，X表示各量數，Σ讀「細格馬」（Sigma）表示總和，ΣX表示各量數的總和之意。

ΣX 乃 $\sum_{k=1}^{k=N} X_k$ 之省寫，因X共有N個，即

$X_1, X_2, \cdots\cdots X_n$

而以X_k爲普通項，$\sum_{k=1}^{k=N} X_k$表示「X_k的總和，而k從1到N」之意。

例如有學生5人，（N＝5），其歲數爲20，22，23，24，27。則五人之平均年齡（M）如下：

$$M=\frac{\Sigma X}{N}=\frac{X_1+X_2+X_3+X_4+X_5}{5}$$

$$=\frac{20+22+23+24+27}{5}=\frac{116}{5}=23.2\text{歲}$$

此種計算方法最爲簡單，毋庸贅述。上例爲枚舉數量數列，在簡單時間數列，如言每年平均產量，每月平均支出均可仿照應用。品質數列則通常不能用此法，而多須採用百分數。如謂中國五族，平均每族若干人，了無意義，若謂漢族百分之幾，滿族百分之幾，蒙族百分之幾，回族百分之幾，藏族百分之幾，則概念分明矣。又如陸，海，空三軍平均每軍若干人，亦屬無謂之詞，必以百分數表示之然後可。地理數列應用平均數亦宜謹愼，如世界各國平均每國人口若干，未嘗有人言之，而我國每縣平均人口若干，或各省每縣平均人口之比較，又未嘗不可。因國家乃由歷史造成，殊難視爲同一單位，而縣區則由國家權力劃分之行政單位也。

三、加權算術平均數

由平均數再計算平均數，必須加以更進一步之考慮。如甲乙丙三國立中學某期畢業考試之平均成績爲68,74,與77,

最簡單方法，可知三國立中學之平均成績爲

$$M=\frac{68+74+77}{3}=\frac{219}{3}=73$$

吾人須問所謂甲國立中學平均成績爲68，乙74，丙77，究竟如何計算而得？若謂甲校有普通科與藝術科，乙校尙有師範科，丙校尙有師範科與商科，則或係

$$M_{甲}=\frac{\text{兩科各平均分數之和}}{2}=68$$

$$M_{乙}=\frac{\text{三科各平均分數之和}}{3}=74$$

$$M_{丙}=\frac{\text{四科各平均分數之和}}{4}=77$$

其實甲校普通科畢業80人，藝術科畢業20人，前者平均成績爲67，後者爲78，故

$$M_{甲}=\frac{80\times 67+72\times 20}{100}=\frac{6800}{100}=68$$

同理乙校共畢業200人，丙校 300人，M乙及M丙亦如M甲之方法計算而得，由是三國立中學之平均畢業成績如下：

$$M=\frac{M_{甲}\times\text{甲校畢業生數}+M_{乙}\times\text{乙校畢業生數}+M_{丙}}{\text{甲校畢業生數}+\text{乙校畢業生數}+\text{丙校畢}}$$

$$\frac{\times\text{丙校畢業生數}}{\text{業生數}}$$

$$=\frac{68\times100+74\times200+77\times300}{100+200+300}$$

$$\frac{44700}{600}=74.5$$

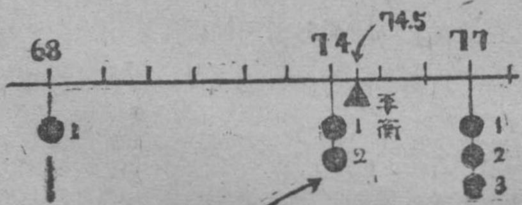
此法與將各校各生之分數相加，除以三校畢業生總數，毫無二致也。

上式吾人若將分子分母，同除以100，即

$$M=\frac{1\times68+2\times74+3\times77}{1+2+3}=\frac{447}{6}=74.5$$

$$Mw=\frac{\Sigma MX}{\Sigma M}$$

結果完全相同。猶如謂甲乙丙三校爲1,2,3,人。可見只須有三校畢業生人數之比例，即可求得三校之平均數，而無須知其總數也。各量數因其重要性之不同，而與以大小不同之比例，謂之加權，其表示重要性之數謂之權數，所得之平均數謂加權算術平均數。權者稱錘也，平均數者，衡器平衡之點也，如下面所示，可以悟加權之意矣。



$$1\times6.5+2\times5=7.5=3\times2,$$

以比例爲權數有二優點，

1.若不知畢業生確數，但知畢業生與學生總數略成比例，即可以各校學生總數爲權數。

2.比例不必爲絕對正確之數，假定三校畢業生爲112,209,298,即可以逕以1,2,3,代之，其結果相差，至爲有限，而無損於統計數字之價值也。

平均數應否加權，須視統計資料之性質而定，大體言之，項數愈多，則權數之影響愈小，例如由全國各縣之平均戶量（每戶平均戶數）以求全國之平均戶量，加權與否，無甚關係。但若各項包含之基本單位數目太相懸殊，如陸海空軍，則又非加權不可也。

四、簡單次數數列之算術平均數

簡單次數數列之算術平均數與加權算術平均數，兩者之計算方法相同，意義亦相似，惟前者係以實際次數乘量數，而後者得以抽象之比例爲權數，乘其量數耳。若加權算術平均數亦以實際之比重爲權數，則與簡單次數數列殊無二致也。

簡單次數數列算術平均數之公式爲

$$- M=\frac{\Sigma fX}{\Sigma f}=\frac{\sum_{k=1}^{k=t} fk \quad Xk}{\sum_{k=1}^{k=t} fk} \qquad \left(試與\frac{\Sigma wX}{\Sigma w}比較\right)$$

$$=\frac{f_1X_1+f_2X_2+\cdots\cdots+ftXt}{rf+\cdots\cdots f_1+f_2+f_3}$$

其中M代表算術平均數，X代表量數，f 代表次數，fX代表各量數與其相應對的次數之乘積，Σ fX代表量數與次數乘積之總和，Σ f 代表次數之總和。茲舉例如下：

簡單次數數列算術平均數計算表

年　齡(量數) X (1)	學生人數（次數） f (2)	量數與次數之乘積 fX (3)
18.5	1	18.5
19.5	2	39.0
20.5	13	266.5
21.5	17	365.5
22.5	19	427.5
23.5	24	564.0
24.5	14	343.0
25.5	7	178.5

26.5	11	291.5
27.5	6	165.0
28.5	3	85.5
29.5	2	59.0
30.5	1	30.5
31.5	—	—
32.5	—	—
33.5	—	—
34.5	—	—
35.5	1	35.5
總計	121(Σ f)	2,869.5(Σ fX)

〔附註〕此項年齡，在調查時已說明為已過生日之歲數，則調查結果之二十三歲，有二十四人，必有剛滿二十三歲者，有在二十三歲與二十四歲之間者，以及將近滿二十四歲者，故以二十三歲半代表二十四人之年齡，其餘仿此。

上表 $\Sigma f=121$，$\Sigma fX=2,869.5$。故

$$M=\frac{\Sigma fX}{\Sigma f}=\frac{2,869.5}{121}=23.7$$

由是，可以23.7歲代表某校某班學生之年齡。人告我以

某校某班學生平均年齡爲23.7歲，余無須親見其學生，而立可想像其必爲一羣活潑潑之青年也。

五、分組次數數列算術平均數之直接計算法

通常因材料太多，若將量數及其次數一一列舉（簡單次數數列），尙嫌麻煩，乃將量數分組而作成次數分配表（分組次數數列）。由分組次數數列計算算術平均數，量數既經分組，則無確定之量數，吾人必須假定以各組距中值代表各該組之量數。各組次數雖並不皆集中於組之中點，而偏向於次數分配之中心，但次數分配中心之兩端，差誤互相抵消，結果仍可相同或大致相近也。

分組次數數列算術平均數之直接計算法。與簡單次數數列算術平均數之計算法，公式完全相同，惟以組中值代替量數耳。

$$M=\frac{\Sigma fX}{\Sigma f}=\frac{\sum\limits_{k=f}^{k=t} fk\ Xk}{\sum\limits_{k=i}^{k=t} fk}$$

$$=\frac{f_1X_1+f_2X_2+\cdots\cdots+ftXt}{f_1+f_2+\cdots\cdots+ft}$$

其中M代表算術平均數，X代表組中值，f代表次數，fX代表各組中值與其相應對的次數之乘積，Σ fX代表所有乘積之總和，Σ f 代表次數之總和。玆舉例如下：

分組次數數列算術平均數直接計算法

（某校某班學生之年齡分配，見前）

年齡分組 (1)	組中值 X (2)	學生人數（次數）f (3)	組中值與次數之乘積 fX (4)
18—20	19	3	57
20—22	21	30	630
22—24	23	43	989
24—26	25	21	525
26—28	27	17	459
28—30	29	5	145
30—32	31	1	31
32—34	33	—	—
34—36	35	1	35
總計		121(Σ f)	2,871(Σ fX)

上表Σ f=121，Σ fX=2,871故

$$M=\frac{\Sigma fX}{\Sigma f}=\frac{2,871}{121}=23.7$$

可見同一資料，在此例中，由簡單次數數列與由分組次數數列計算算術平均數計至小數一位為止，結果完全相同。但此乃偶然之符合耳，非必盡能相等也。

六、等組距分組次數數列算術平均數之簡捷計算法

今有20歲者六人，25歲者八人，30者五人，求此十九人之平均年齡，依簡單次數數列算術平均數之計算法計算如下：

$$M=\frac{\sum fX}{\sum f}=\frac{f_1X_1+f_2X_2+f_3X_3}{f_1+f_2+f_3}$$

$$=\frac{6\times20+8\times25+5\times30}{6+8+5}=\frac{120+200+150}{19}=\frac{470}{19}=24.7$$

吾人若變更原點及單位，計算較為簡捷。

$$M=\frac{6\times20+8\times25+5\times30}{19}$$

$$=\frac{6\times(20-25+25)+8\times(25-25+25)+5\times(30-25+25)}{19}$$

$$=\frac{6\times[(20-25)+25]+8\times[(25-25+25)]+5\times[(30-25)+25]}{19}$$

$$=\frac{(6\times252+8\times25+5\times25)+[6\times(20-25)+8\times(25-25)+5\times(30-25)}{19}$$

$$=\frac{19\times25}{19}+\frac{6\times(-5)+3\times0+5\times5}{19}$$

$$=25+\frac{6\times\left[\left(\frac{-5}{5}\right)\times5\right]+8+\left[\frac{0}{5}\times5\right]+5\times\left[\frac{5}{5}\times5\right]}{19}$$

$$=25+\frac{[6\times(-1)]\times5+[8\times0]\times5+[5\times1]\times5}{19}$$

$$=25+\frac{6\times(-1)+8\times0+5\times1}{19}\times5$$

$$=25+\frac{-6+5}{19}\times5$$

$$=25+\frac{-1}{19}\times5$$

$$=25+(-.05)\times5$$

$$=25-.3=24.7$$

上式謂之以25爲新原點，5爲新單位。新原點普通謂之假定算術平均數，而在等組距之次數分配表上新單位即組距也。

再以代數方法述之如下：

$$M=\frac{f_1X_1+f_2X_2+f_3X_3}{\Sigma f}$$

$$=\frac{f_1(X_1-A+A)+f_2(X_2-A+A)+f_3(X_3-A+A)}{\sum f}$$

$$=\frac{f_1A+f_2A+f_3A}{\sum f}+\frac{f_1(X_1-A)+f_2(X_2-A)+f_3(X_3-A)}{\sum f}$$

$$=\frac{A\sum f}{\sum f}+\frac{f_1\cdot\frac{X_1-A}{i}\cdot i+f_2\cdot\frac{X_2-A}{i}\cdot i+f_3\cdot\frac{X_3-A}{i}\cdot i}{\sum f}$$

$$=A+\frac{f_1\cdot\frac{X_1-A}{i}+f_2\cdot\frac{X_2-A}{i}+f_3\cdot\frac{X_3-A}{i}}{\sum f}\times i$$

若以 x 代 $\frac{X-A}{i}$，即以 x_1 代 $\frac{X_1-A}{i}$，x_2 代 $\frac{X_2-A}{i}$，x_3 代 $\frac{X_3-A}{i}$ 則

$$M=A+\frac{f_1x_1+fx_2+f_3x_3}{\sum f}\times i$$

即 $M=A+\frac{\sum fx}{\sum f}\times i$

此即由等組距分組次數數列計算算術平均數之簡捷法的公式也。其中 M 代表算術平均數，A 代表假定算術平均數，應取次數最多或位置適中之組的組中值爲假定算術平均數，i 爲組距，x 爲各組中值對 A 之離差而以 i 爲單位即 $x=\frac{X-A}{i}$，茲舉例如下：

等組距分組次數數列算術平均數之簡捷計算法

（某校某班學生之年齡分配，同前）

年齡分組 (i=2) (1)	組中值 X (2)	學生人數(次數) f (3)	$\frac{X-A}{i}$ x (4)	fx + (5)	fx — (6)
18—20	19	3	—2		6
20—22	21	30	—1		30
22—24	23(A)	43	0		
24—26	25	21	1	21	
26—28	27	17	2	34	
28—30	29	5	3	15	
30—32	31	1	4	4	
32—34	33	—	5	—	
43—36	35	1	6	6	
總 計		121(Σf)	×	80	—36
				44(Σfx)	

計算步驟：

(1) 確定各組中值，卽16,21,……35，見第二行。

(2) 確定假定算術平均數，本例第三組次數最多，其組中値爲23，卽以23爲假定算術平均數，A=23、

(3) 求以A為原點，以i(=2)為單位之組中值：

$$x_1=\frac{19-23}{2}=-2,\ x_2=\frac{21-23}{2}=-1=,\ x_3=\frac{23-23}{2}=0,\cdots$$

……………(見第四行。)

(4)求次數(f)與新組中值(x)之乘積fx，見第五、六兩行。

(5)求次數之總和，$\Sigma f=121$見第三行總計欄。

(6)求fx之總和，$\Sigma fx=44$，見第五、六行總計欄。

(7)代入公式求算術平均數。

$$M=A+\frac{\Sigma fx}{\Sigma f}\times i$$

$$=23+\frac{44}{121}\times 2$$

$$=23+\frac{88}{121}$$

$$=23+.7$$。

此法可避免繁複之乘積，使運算便利，而無須借助於計算機或算盤，即可迅速得出結果，是以謂之簡捷也。惟其應用限於等組距之分組次數數列耳。不等組距之數列亦可單用變更原點法計算之。

七、統計係數

統計係數乃平均數之一種特別應用，係用表示某種現象發生之比率者。若以C代表統計係數，N代表某事物之總數，Q代表發生之事件，則其公式如下；

$$C=\frac{Q}{N}$$

例如某次戰役，參加作戰將士共10,000人，陣亡150人，負傷200人，則

陣亡係數：$C=\frac{Q}{N}=\frac{150}{10,000}=.015$

負傷係數：$C=\frac{Q}{N}=\frac{200}{10,000}=.02$

又如某地某年年中人口爲10,000人，出生280人，死亡210人，則

出生係數：$C=\frac{Q}{N}=\frac{280}{10,000}=.028$

死亡係數：$C=\frac{Q}{N}=\frac{210}{10,000}=.021$

係數在尋常統計中佔一重要地位，在人口學中，尤其饒有興

味。人口總數或增或減，而某種係數，則變動甚微；在常態之時期，常歷久而不變。因係數之應用，各國之統計資料可以比較；除非有特別事變，如戰爭者，未來若干年之情況，恆可加以預測，且有驚人之正確性。不同地域之出產，各組年齡，各類職業，各種疾病之死亡，各歲年齡之婚姻，以及自殺，犯罪，火災，某種商品之消費等均可作成係數，以資比較。他若名勝之旅客、火車上之遺傘、汽車之失事，等等，若能獲得材料，無不可以作成係數，統計工作進步之國家，種種重要係數，常刊諸報告，從事統計研究者，不可不略知其梗概也。

係數有精粗之分，在粗係數中，N包含衆多之因素，而在精係數中，某種結果(Ω)，應與其直接原因(N)相比較，例如計算某戰區某月之陣亡係數，其精粗之比較如下：

$$粗係數=\frac{某月陣亡人數}{某月月中將士總額}$$

$$精係數=\frac{某月陣亡將士人數}{某月參加作戰之將士總額}$$

再如計算結婚係數，　吾人應問何種人爲適婚人，其答案乃成年之未婚男女以及鰥夫寡婦，其總數即婚姻精係數之分母

(N)。

但粗係數亦有其用途，亦可代表某種現象之發生率。且若所含各種因素中，只有一個爲變數，其餘爲常數，則粗係數，仍可表示一個因素之變動，例N 爲某戰區將士總額，n 爲參加作戰之將士，M 爲陣亡將士，則粗係數爲

$$C=\frac{M}{N}=\frac{M}{n}\times\frac{n}{N}$$

若參加作戰之將士，與將士總額常成一定比率即$\frac{n}{N}=r$，則

$$C=\frac{M}{n}\times r \qquad C'=\frac{C}{r}=\frac{M}{n}$$

是C或C'仍爲精係數也。

婚姻率亦然，若N爲人口總數，n爲適婚人數，M爲結婚件數，則粗係數

$$C=\frac{M}{N}=\frac{M}{n}\times\frac{n}{N}$$

若適婚人數與人口總數，常成一定比例，即 $\frac{n}{N}=r$，則

$$C=\frac{M}{n}\times r, \qquad C'=\frac{C}{r}=\frac{M}{n}$$

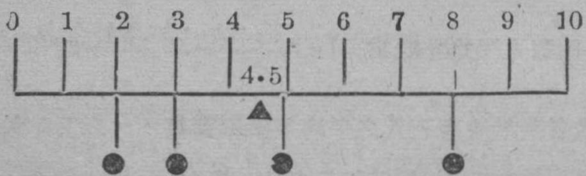
是C之大小仍與$\frac{M}{n}$成比例也。

八、算術平均數之特質

算術平均數之意義，用途及其計算方法，已略如前述，由其計算方法，吾人可將其重要特質，歸納言之如下：

1.算術平均數係從全體量數中計算而出，離平均數愈遠之量數，其影響愈大。例如有三人焉，一年20,一年23,一年24，則三人平均年齡為$M=\frac{20+23+24}{3}=22.3$。若加入一人年為23,則四人平均年齡為 $M=\frac{20+22+23+24}{4}=22.2$，可見此22歲者對於原有三人之平均年齡影響甚小，以其年齡與平均年齡距離甚近也。倘若加入者非一青年,而為一幼童，其年9歲,則四人平均為$M=\frac{9+20+23+24}{4}=19$或加入者非一青年而為一長者,其年61歲,則四人平均為$M=\frac{20+23+24+61}{4}=32$，兩者均以平均數以重大之變動,$(22.3-19=3.3, 32-22.3=9.7)$，以其年齡與平均年齡距離太遠也。

2.若將數列分爲兩部分，甲部分之量數均大於算術平均數，乙部分之量數，均小於算術平均數，則甲部分各量數之所過者，洽等於乙部分之所不足者。簡言之，數列各量數對算術平均數之離差的代數和等於零，故曰算術平均數乃數列之重心。如有西瓜四個，其重量爲2斤、3斤、5斤、8斤其平均重量爲$M=\frac{2+3+5+8}{4}=4.5$斤，若在一無重量之棒上，刻以等距離的度數，在相當於各西瓜重量之刻度處，繫以重量相等之錘，則必平衡於4.5斤之處，因4.5斤乃重心之所在也。



但　$(2-4.5)+(3-4.5)+(5-4.5)+(8-4.5)=(-2.5)$
$+(-1.5)+.5+3.5=-4+4=0$

通常寫作　$\Sigma(X-M)=0$　，以d代X－M，則$\Sigma d=0$.

3.一數列分成兩組或數組時，只須知各組量數之項數而無須知各量數之數值，該數列之算術平均數，卽可用加權方法由各組算術平均數求得之。如計政班分會計組與統計組，

會計組30人平均成積75分，統計組40人，平均成積78分，則計政班平均成積爲

$$M=\frac{W\quad X}{W}=\frac{30\times75+40\times78}{30+40}=\frac{537}{7}=767\text{。}$$

惟須注意須各組項數相等，始可直接平均，如會計組與統計組均爲35人，則35：35＝1：1故

$$M=\frac{75+78}{2}=76.5\text{。}$$

九、算術平均數之優點

1.算術平均數，通俗易曉，以之代表一統計數列，其意義幾盡人皆能了解。

2.其計算方法，亦甚簡易，普通只須加法及除法，卽可算出，較繁者，亦只須嫻習加減乘除，卽足應用。無須按序排列或用圖解法。

3.係從全體量數中求出，每個量數均給以同等重要性，只須量數均屬同質，則算術平均數，富有代表之意義。

4.有時欲與極端量數以相當地位者，可以算術平均數代表之。（如社會享受不均，求總消費値之每人平均數，可以

表示社會平均分配時，每人之享受。）

5.只須知量數總値及其項數，而無庸知各量數之數値，即可求得算術平均數。（如知某市某年之人口數及米之消費量即可求出每人之消費量，並不須挨戶訪問，調查其一年中之食米數量也。）

十、算術平均數之缺點

1,算術平均數往往爲抽象的數目，而非具體之事實，如言平均戶量爲5.34人，實際上絕無家庭，其人口爲五個又小數點幾者。

2.每個量數與以相等的地位，結果太大或太小之極端量數影響甚大，故不同質之數列，恆不能用算術平均數代表之。如火車一列可載2,000 人，汽車一輛可載30人。遂謂火車與汽車之平均運量爲1,015人，實屬荒誕之至。

3,算術平均數旣從全體量數中算出，任何量數一有差誤，算術平均數卽受其影響而發生差誤，極端量數有差誤，影響尤大。

4.依品質分類或分等級之數列，不能求算術平均數．如工人之分精粗，年齡之分少壯老，成績之分甲乙丙丁戊等均

無法計算算術平均數。

5.次數數列雖經製成次數線圖，亦不能利用之以決定算術平均數之數值。

6.數列兩端之量數不定，如用「以上」或「以下」表示者，不能計算算術平均數。

第三節　中位數

一、中位數之意義

當量數按大小次第排列，其恰居正中位置之量數，謂之中位數，因是中位數之前後，各有量數之一半。例如但知有200人每月工資不及20元，有一人每月20元，有200人每月多於20元，則20元即爲其中位數。凡當各個量數未能一一知其數值，或各量數欲予以同等地位者應求中位數以爲其代表。

二、枚舉數列之中位數

五人依高矮次第排列，第三人之高度爲其中位數。

七人依高矮次第排列，第四人之高度爲其中位數。

百零一人依高矮次第排列，第五十一人之高度爲其中位數，n個人依高矮次第排列，則第$\frac{n+r}{2}$人之高度爲其

中位數。

故當量數之項數n爲奇數時，則

$$Md = 第\frac{n+1}{2}個量數之數值$$

例如11人之年齡如下

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
18	19	21	23	23	23	24	25	27
(10)	(11)							
27	28							

$$Md = 第\frac{11+1}{2}個人之年齡$$

$$= 第六個人之年齡$$

$$= 23歲$$

六人依高矮排隊，中位數爲第三人之高度，或第四人高度，或第三、第四兩人之平均高度

五十人依高矮排隊，中位數爲第二十五人之高度，或第二十六人之高度，或第二十五、二十六兩人之平均高度。

故量數之項數n爲偶數時，則

$$Md = 第\frac{n}{2}個量數之數值$$

$$或第\frac{n}{2}+1個量數之數值$$

或第$\frac{n}{2}$個，第$\frac{n}{2}+1$個兩個量數之平均數值。

通常為求一律起見，取第$\frac{n}{2}$個與第$\frac{n}{2}+1$個兩個量數之平均數值。

例如上列八數中，若尚有第十二人為29歲，則中位數為23或24均無不可，但通常取

$$Md=\frac{23+24}{2}=23.5。$$

時間數列，地理數列以及品質數列均可依此法以求中位數，但其中位數不必為數量值，而為代表之時期，地域或品質。

三、簡單次數數列之中位數

簡單次數數列之中位數，無論次數為奇數，抑為偶數，總依下列公式求之：

$$Md=第\frac{n+1}{2}個量數之數值。$$

如遇n為偶數，即取第$\frac{n}{2}$個與第$\frac{n}{2}+1$個量數平均之。

	年歲	人數(次數)	累積次數	
	18	1	1	
	19	3	4	
	21	6	10	
	23	11	21	
Md——→	24	14	35	$\leftarrow -\frac{57+1}{2}$
	25	10	45	
	27	9	54	
	30	3	57	

Md＝第$\frac{57+1}{2}$人(卽第28人)之年齡

＝24歲。

上表若再增加一人爲35歲，則中位歲應爲第29（$=\frac{58}{2}$）人與第30（$=\frac{58}{2}+1$）人之平均年齡，但第29人及第30人均爲24歲，故中位數仍爲24歲，可見此35歲之人，對於中位數，毫無影響。

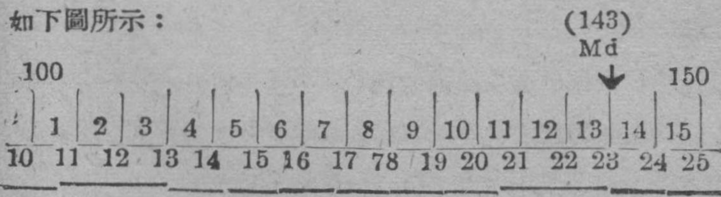
四、分組次數數列之中位數

分組次數數列，無論其組距相等與否，均假定中位數所在組內各次數平均分配，而用插補法求之。

今有某機闗薪給分配表如下：

薪給(元)		人數	累積次數 以下	累積次數 以上
(1)	400以上	2	46	2
(2)	300—400	5	44	7
(3)	200—300	3	39	10
(4)	150—200	11	36	21
Md→(5)	100—150	15	25←	36← $\frac{45}{2}=23$
(6)	75—100	6	10	42
(7)	50— 75	3	4	45
(8)	50以下	1	1	46

人數一共46人，中位數應爲第23(．$=\frac{46}{2}$本應取$\frac{46+1}{2}=$ 23.5，爲簡便起見，卽取$\frac{46}{2}$)人之薪給。今100元以上者36人，150元以下者25人，可見中位數在100—150之間，卽在第5組內。若第5 組之15人分配均勻，比例散布在本組各點，則有如下圖所示：



$$\text{中位數}=M2=100+\frac{150-100}{15}\times(23-10)$$

$$=700+\frac{50}{15}\times 13$$

$$=100+43$$ （若不取小數）

$$=143\text{元}$$

$$\text{或}Md=150-\frac{150-100}{15}\times(23-21)$$

$$=150-\frac{50}{15}\times 2$$

$$=150-7$$

$$=143.\text{元}$$

寫成普通公式，則為

$$Md=L+\frac{i}{f}\left(\frac{N}{2}-F_1\right)$$

式中Md代表中位數，L代表中位數所在組之下限，i代表該組之組距，f代表該組內之次數，N代表總次數，F_1代表中位數所在組以下（量數小於中位數組之下限）各組之次數總和。由此公式求中位數，其步驟如下：

1.求以下累積次數；

2.確定中位數所在組，卽累積次數剛巧大於$\frac{n}{2}$之組，(上表第5組)$\frac{N}{2}=\frac{46}{2}=23$；

3.求中位數所在組之組距，$i=150-100=50$；

4.觀察中位數所在組之下限及次數．$L=100, f=15$.

5.觀察剛巧小於$\frac{n}{2}$之累積次數，$F_1=10$

6.代入公式，求中位數

$$Md=L+\frac{i}{f}\left(\frac{N}{2}-F_1\right)$$

$$=100+\frac{50}{15}\times(23-10)$$

$$=100+\frac{10}{3}\times 13$$

$$=100+43$$

$$=143.\text{元}$$

改用以上累積次數求之，則公式如下：

$$Md=U-\frac{i}{f}\left(\frac{u}{2}-F_2\right)$$

其中U代表中位數所在組之上限，F_2代表中位數所在組以上

（量數大於中位數所在組之上限）各組之次數總和。其餘同前。

五、四分位數、十分位數、百分位數

中位數將量數項數分爲二等分，故可稱爲二分位數。同理，四分位數將量數分爲四等分，十分位數將量數分爲十等分，百分位數將量數分爲百等分。所有中位數，四分位數，十分位數，百分位數統稱分割數。將全體量數，由小而大，順序排列，分割成若干等分，分界上各項之數値，卽分割數。

分割數之求法，中位數已述之於前，其他分割數可仿照中位數求之。

今設 n爲任何正整數，卽 $n=1,2,3,\ldots\ldots$，則枚舉數列與簡單次數數列之下四分位數（第一個四分位數 Q_1 ）與上四分位數（第三個四分位數，Q_3）求法如下：

量數的項數	下四分位數，Q_1	上四分位數，Q_3
4n	$\frac{1}{2}$（第n項＋第$\overline{n+1}$項）	$\frac{1}{2}$（第 n 項＋第$\overline{3n+1}$項）
4n+1	$\frac{1}{4}$第 n 項＋$\frac{3}{4}$第$\overline{n+1}$項	$\frac{3}{4}$第3n+1項＋$\frac{1}{4}$第$\overline{3n+2}$項

4n+2	第$\overline{n+1}$項	第$\overline{3+12}$項
4n+3	$\frac{3}{4}$第$\overline{n+1}$項$+\frac{1}{4}$第$\overline{n+2}$項	$\frac{1}{4}$第$\overline{3n+2}$項$+\frac{3}{4}$第$\overline{3n+3}$項

十分位數與百分位數可依同法，列成一表。

至於分組次數數列，求分割數之公式，可列之如下：

(1)中位數

$$Md=L+\frac{i}{f}\left(\frac{N}{2}-F_1\right)$$

(2)四分位數

$$Q=L+\frac{i}{f}\left(\frac{N}{4}-F_1\right)$$

$$Q3=L+\frac{i}{f}\left(\frac{3N}{4}-F_3\right)$$

	薪給(元)	人數	以下累積次數	
	50 以下	1	1	
	50——75	3	4	
	75——100	6	10	
$Q_1\rightarrow$	100——150	15	25 $\leftarrow$	$\frac{46}{4}=11.5$
$Q_2\rightarrow$	150——200	11	36 $\leftarrow$	$\frac{3\times46}{4}=34.5$
	200——300	3	39	
	300——400	5	44	
	400以上	2	46	

$$Q_1=100+\frac{50}{15}\left(\frac{46}{4}-10\right)$$

$$=100+\frac{50}{15}\times(11.5-10)$$

$$=100+\frac{10}{3}\times 1.5$$

$$=105$$

$$Q_3=150+\frac{50}{11}\left(\frac{3\times 46}{4}-25\right)$$

$$=150+\frac{50}{11}(34.5-25)$$

$$=150+\frac{50}{11}\times 9$$

$$=1933$$

(3)十分位數

$$Dr=L+\frac{i}{f}\left(\frac{rN}{10}-Fr\right)$$

$$D_3=L+\frac{i}{f}\left(\frac{3N}{10}-F_3\right)$$

薪給（元）	人數	以下累積次數
50以下	1	1
50——75	3	4
$P_{15}\rightarrow$ 75——100	6	10 $\leftarrow \frac{15\times46}{100}=6.9$
$D_3\rightarrow$ 100——150	15	25 $\leftarrow \frac{3\times46}{10}=13.8$
150——200	11	36
200——300	3	39
300——400	5	44
400以上	2	46

$$D_3=L+\frac{i}{f}\left(\frac{3\times46}{10}-10\right)$$

$$=100+\frac{50}{15}(13.8-10)$$

$$=100+\frac{10}{3}\times3.8$$

$$=113$$

(4)百分位數

$$Pr=L+\frac{i}{f}\left(\frac{rN}{100}-Fr\right)$$

$$P_{15}=L+\frac{i}{f}\left(\frac{15}{100}-F_{15}\right)$$

$$=75+\frac{25}{6}\left(\frac{15\times46}{100}-4\right)$$

$$=75+\frac{25}{6}(6.9-4)$$

$$=75+\frac{25}{6}\times2.9$$

$$=87$$

計算分割數之普通公式：

1 枚舉數列及簡單次數數列之分割數：

第 r 個S 分位數＝第 $\left(\frac{rN}{S}+\frac{1}{2}\right)$ 個量數

例如中位數爲第 1 個二分位數，故

$$Md=\text{第}\left(\frac{N}{2}+\frac{1}{2}\right)\text{個量數}$$

2. 分組次數數列之分割數：

第 r 個S 分位數 $=L+\frac{i}{f}\left(\frac{N}{S}-Fr\right)$

例如下四分位數爲第 1 個四分位數，故

$$Q_1=L+\frac{i}{f}\left(\frac{N}{4}-F_1\right)$$

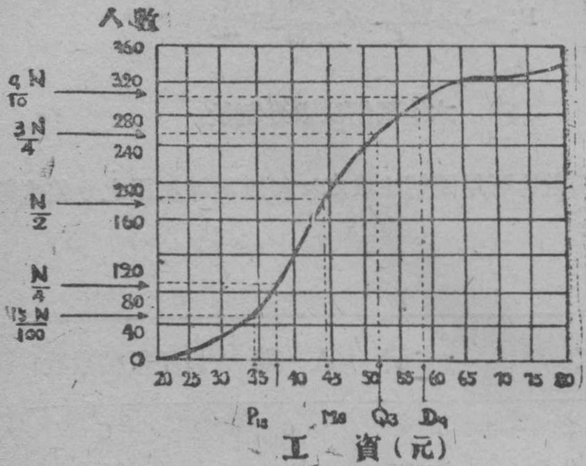
又第23個百分位數爲

$$P_{23}=L+\frac{i}{f}\left(\frac{23N}{100}-F_{23}\right)$$

六、圖解法求分割數

分割數均可用圖解法求之，但限於等組距之分組次數數列。

工資（元）	人數	以下累積數數
2.0——2.5	6	6
2.5——3.0	16	22
3.0——3.5	34	56
3.5——4.0	61	117
4.0——4.5	66	183
4.5——5.0	57	240
5.0——5.5	37	277
5.5——6.0	28	305
6.0——6.5	9	314
6.5——7.0	6	320
7.0——7.5	8	328
7.5——8.0	8	336



步驟

(1) 作修勻累積次數曲線圖。

(2) 求分割數之次數地位

中位數之地位：$\frac{N}{2}=168$

下四分位數之地位：$\frac{N}{4}=84$

上四分位數之地位：$\frac{3N}{4}=252$

第九個十分位數之地位$\frac{9N}{10}=302.4$

第十五個百分位數之地位$\frac{15N}{100}=50.4$

(3) 在縱表尺上覔取分割數之次數地位，作線與 x 軸平行，與修勻累積次數曲線相交。

(4) 在各交點上作線垂直於x軸，與x軸之交點，卽所求之分割數。

觀圖：　$Md=4.4$

$Q_1=3.7$

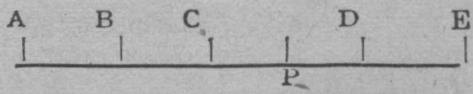
$Q_2=5.2$

$D_9=5.9$

$D_{15}=3.4$

七、中位數之特質

中位數最重要之特質，爲與各量數的距離之和爲最小者。今有A、B、C、D、E 五個兵站，若將電話交換機設在中位數處，則所需之電話線爲最小者，證之如下



中位數在C站。

電話線之和＝AC＋BC＋CD＋CE

若電話交換機改設在P處，則

電話線之和＝AP＋BP＋CP＋PD＋PE

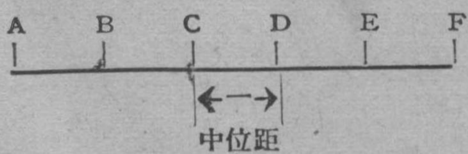
＝(AC＋CP)＋(BC＋CP)＋CP＋(CD－CP)

＋(CE－CP)

＝AC＋BC＋CD＋CE＋CP

故電話線較前多費CP之長度。

若兵站爲偶數（量數之項數爲偶數）則交換機設在當中之二站上或二站之間，所需線料，均屬最小限度。此二



站之間，謂之中位距。學者試自證在中位距上任意一點，與各兵站之距離爲最小。

八、中位數之優點與缺點。

優點：

1.中位數不受極端量數之影響，用中位數表示社會平均財富，不因有萬貫家財者而使平均財富增加。

2.中位數確定方法簡易，只須計算項數，即可求得。

3.極端量數不確定者，欲求平均數以爲代表，最好求中位數。

4.不可測量之統計事項，如品質資料，亦可求中位數。

缺點：

1.中位數有時不能代表一數列，例如調查所得兩羣人數相等之工人，一羣月入工資在10元與20元之間，一羣則在30元與40元之間，中位數，必在20元與30元之間，但事實上20元與30元之間，全無一人。故中位數之應用，限於量數相當集中之數列，少數極端量數無甚影響者。

2.欲與極端量數以較大之比重時，中位數不適用。

3.中位數有時亦爲抽象數目，與實際不符。

4.計算方法雖簡，而無一定數學公式可資遵循。

第四節 衆數

一、衆數之意義

衆數亦稱範數。在分組次數表或簡單次數表上，次數發現最多之量數或數值，卽爲衆數。或謂衆數乃最密集的地位，或謂衆數乃最顯著的數值。若將次數分配表繪諸圖上，以修勻之連續曲線表示次數分配之情形，則與最大縱坐標相應對之橫坐標，卽衆數也。

例如戶口調查，若戶量爲四人之戶數最多，則四人卽戶量分配之衆數，而此等家庭謂之代表家庭。又衆數不必爲數值，現象中發現最常之情形亦爲衆數。如某處死亡病因以痢疾爲最多，則痢疾亦爲衆數，謂之流行症；如某時期以盛錫福之呢帽最通行，則盛錫福呢帽爲衆數，或稱最時行的呢帽。

吾人尋常好言「普通甚麼」，如普通人，普通公務員，普通工人，普通食物，普通香煙等等。普通者，最常見之意也，而最常見之情形卽衆數也。就知識程度言，普通人均爲小學程度；就薪俸言，普通公務員均在百元左右；就工資言，普通工人均在二十元左右；就種類言，普通食物爲米；就

牌號言，普通香煙爲白金龍；則凡此小學程度，百元二十元、米，白金龍均衆數也。

二、決定衆數之方法

甲、觀察法

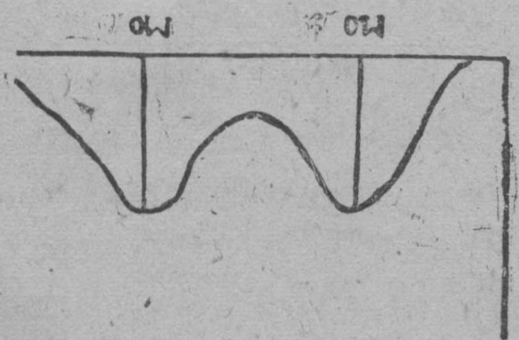
衆數既爲最常見之量數或情形，吾人只須將資料分類分組，作成次數分配表，即可由觀察而得衆數。

年齡	學生人數
總計	109
12	1
13	3
14	15
15	30
16	21
17	14
18	9
19	8
20	4
21	2
22	2

由上表觀之、由12歲起人數漸增，至15歲，登峯造極，得30

入，16歲以上之人數又漸減，則15歲即衆數也。

統計資料如不同質，則其數列常呈多衆數之數列，而其曲線則爲多峯曲綫，如技術截然不同之兩羣工人，混合調查，最普通之熟練工人工資爲50元，而最普通之粗工工資爲15元，則50元與15元均爲衆數，其曲綫成下圖形式：



又如某大學有附中附小，三者之學生，並非按步升級而係限制年齡各自招考者，則其學生之混合年齡分配，或將成爲三峯曲線，則呈下圖形式：

可見不同質之統計資料，可以有多個衆數，分別代表其羣性，反之，若統計數列有多個衆數，亦可反證其資料大概爲不同質。

乙、繼續併組法

資料縱爲同質，有時因組距不夠大，而呈多峯曲綫，此時應採用繼續併組法以決定衆數。

觀下表，自第二行向下讀之，可見人數漸增，至685 乃達最高點，自此以下，雖時多時少，但無有過685 者。但在此行共有14個最大值，數目之增減，殊無規律。是則有14個衆數，（最大縱坐標相應之橫坐標）惟在 1.15——1.24 組之衆數最爲顯著耳。若吾人擴大組距。繼續併組，則可以求得較爲正確之衆數。

「歸併階組決定眾數法」

工資率	工人人數							
	一個組	二個組		三個組			五個組	
.25 — .34	1							
.35 — .44	15	16		93				
.45 — .54	99		24		154		310	
.55 — .64	85	144				307		
.65 — .74	129		242	255				
.75 — .84	113	270			439			725
.85 — .94	169		282			483		
.95 — 1.04	201	370		674			1477	
1.05 — 1.14	304		505		1390			
1.15 — 1.24	685	989				1088		
1.25 — 1.34	48		784	262				2092
1.35 — 1.44	458	559			1023			
1.45 — 1.54	466		395			496	1397	
1.55 — 1.64	79	533		740				
1.65 — 1.74	202		974		603			
1.75 — 1.84	379	581				587		934
1.85 — 1.94	60		389	660				
1.95 — 2.04	319	331			396		970	
2.05 — 2.14	45		412			583		
2.15 — 2.24	265	310		343				
2.25 — 2.34	33		298		349			640
2.35 — 2.44	101	134				330		
2.45 — 2.54	106		247	310			806	
2.55 — 2.64	13	204			372			
2.65 — 2.74	163		176			78		
2.75 — 2.84	2	165		180				322
2.85 — 2.94	45		47		146			
2.95 — 3.04	139	144				147	188	
3.05 — 3.14	5		134	181				
3.15 — 3.24	47	52			64			
3.25 — 3.34	10		59			57		85
3.35 — 3.44	0	19		233				
3.45 — 3.54	221		231		226		254	
3.55 — 3.64	5	226				242		
3.65 — 3.74	16		21	32				
3.75 — 3.84	11	27			27			140
3.85 — 3.94	0		11			93		
3.95 — 4.04	88	22		82			96	
4.05 — 4.14	0		32		85			
4.15 — 4.24	3	3				1		
4.25 — 4.34	0		3	3				6
4.35 — 4.44	0	0			3			
4.45 — 4.54	3		3			4	4	
4.55 — 4.64	1	4		4				
4.65 — 4.74	0		1		1			
4.75 — 4.84	0	0				0		9
4.85 — 4.94	0		0	8				
4.95 — 5.04	8	8			6		8	
5.05 — 5.14	0		8			8		
5.15 — 5.24	0	0	1	1				
5.25 — 5.34	1	1						

見 BOWLEY ELEMENTS OF STATISTICS P.93

在一角組，衆數甚不確定，已如前述。今從.25—.44組起，作二角分組，則人數爲16,144,270,370,989,557,538,531,等等，在1·05—1·24組間之989爲最大者：若二角併組從·35—·54起，則人數爲74,242,282,505,784,924,274等等,最大人數924在1·35—1·54組，是故二角組不能決定衆數之所在。今再作三角併組，若從.55—.84起、則有355,674,1242(1.15—1.44),740等；若從·65—·94起,則有439,1190(·95—1·24),1003 等若從·75—1·04 起，則有433,1088(1·05—1·34),996等,在此併組程序中，所有衆數之所在組均包括1·15—1·24一組，故可想定衆數爲1.20或近於此數，雖不中不遠矣。

繼續併組法，可撮述如下：將資料之表列，繼續擴大組距，直至有一定規律或成單峯數列爲止；觀察最顯著之衆數所有組爲何；若組限變更，而衆數移動(如上述之二角組)，則組距仍不夠大；若組限變更，則衆數並不移動，則衆數即在各大組(如上述之1.15—1·44,·95—1·24,1·05—1·34)所包含之最小組(如上述之1·15—1·24)中。

丙、插補法

有時雖則資料同質，分配甚有規律，亦不易精密決定衆數。試觀下表：

年　　齡	人　　數
18——20	3
20——22	30
22——24	43
24——26	21
26——28	17
28——30	5
30——32	1
32——34	——
34——36	1

雖則知衆數必在22—24—組，但究在何處，無由決定。在此種情形，通常用插補法決定之。卽取衆數所在組之上下兩組的次數，以爲決定之因素。假定上下兩組，對於衆數之牽引力，與其次數成比例，次數多者牽引力大。故上表之衆數爲

$$Mo=22+2\times\frac{21}{30+21}$$

$$=22+\frac{42}{51}$$

$$=22.8$$

或 $$Mo=24-2\times\frac{30}{30+21}$$

$$=24-\frac{90}{51}$$

$$=22.8$$

寫成公式卽爲 $$Mo=L+i\times\frac{fz}{f_1+f_2}$$

或 $$Mo=U-i\times\frac{f_2}{f_1+f_2}$$

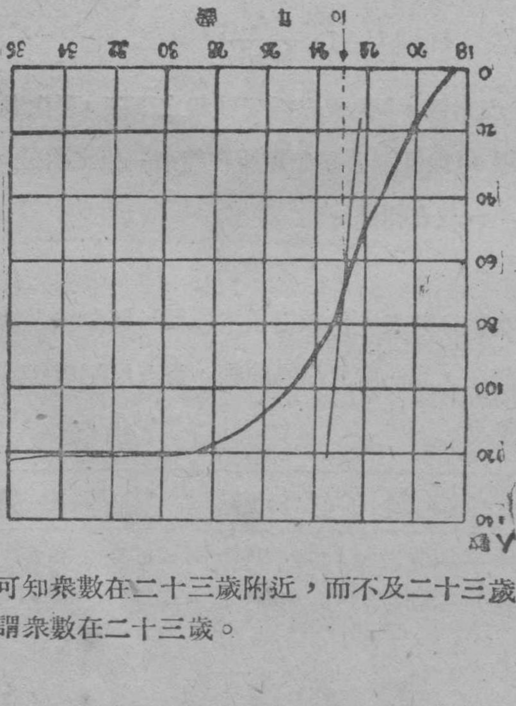
其中Mo代表衆數，L代表衆數所在組之下限，U代表衆數所在組之上限，i爲組距，f 爲衆數所在組下一組之次數，f_2爲衆數所在組上一組之組限。

丁、圖解法

若將次數分配表，作成累積次數圖，則曲線之轉點，卽爲衆數之所在，用一直尺以觸曲線，當直尺剛巧切斷曲線之點，卽轉點也。

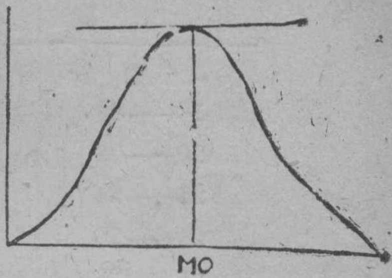
年　　齡	人　數	累積次數
18——20	3	3
20——22	30	33

22——24	43	76
24——26	21	97
26——28	17	114
28——30	5	119
30——32	1	120
32——34		120
34——36	1	121



由上圖可知衆數在二十三歲附近，而不及二十三歲，吾人卽可大略謂衆數在二十三歲。

或作修勻次數分配圖，其最大縱坐標相應之橫坐標，亦即衆數之所在。



戊、皮爾生估計法

依皮爾生式之經驗，在「略不對稱」的次數曲線上，算術平均數與衆數的距離，大約等於算術平均數與中位數的距離之三倍，即

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

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

所以假定（一）統計資料是略不對稱的次數分配，並且(二)算術平均數和中位數早已求得，即可用上列公式，估計衆數之數值。但不具備上列第一條件，不可用此公式；不具備第二條件，大可不必求得算術平均數與中位數再來求衆數。

年齡	人數
總計	121
18——20	3
20——22	30
22——24	43
24——26	21
26——28	17
28——30	5
30——32	1
39——34	—
34——36	1

上表假定已經求得算術平均數(M)爲23.7,中位數(Md)爲23.3,則衆數可估計如下：

Mo= ×M—3(M—Md)

=23.7—3(23.7—23.3)

2=3.7—1.2

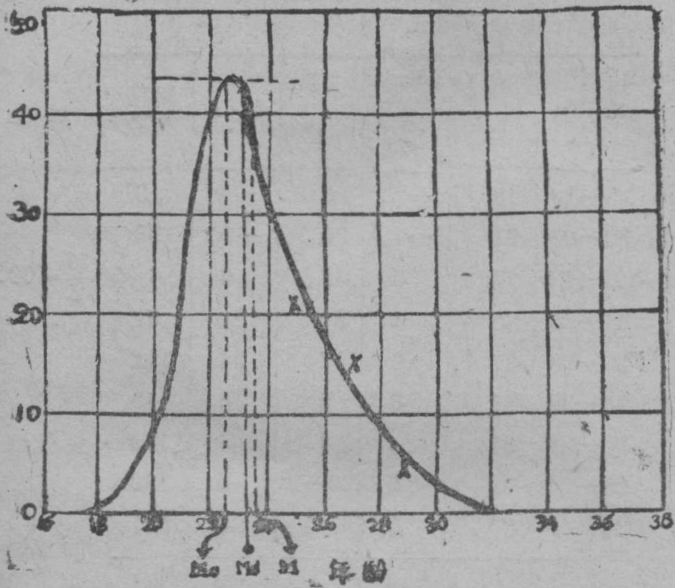
=22.5

「附」算術平均數與中位數

年 齡	組中值 X	$\frac{X-23}{2}$ x	人數 f	f x	累積次數
18—20	19	—2	3	—6	3
30—22	21	—1	30	—30	33(F)
Md→22—24	23(A)	0	43 (f)	0	76← $\frac{11}{2}=60.5$
24—26	25	1	21	21	97
26—28	27	2	17	34	114
28—30	29	3	5	15	119
30—32	31	4	1	4	120
32—34	33	5	—	0	120
34—36	35	6	1	6	121
總 計	×	×	121	44	×

$$M = A + \frac{\Sigma fx}{\Sigma f} \times i = 23 + \frac{44}{121} \times 2 = 23 + .7 = 23.7$$

$$M = L + \frac{\frac{N}{2} - F}{f} \times i = 22 + \frac{\frac{121}{2} - 33}{43} \times = 22 + 1.3 = 23.3,$$



年齡

三、衆數之重要性

吾人常聞「普通人」如何，「一般人」如何，「大多數人」如何，究竟作何解釋？如言普通工人，究竟指工資收入爲算術平均數者而言，抑係指其工資爲衆數者而言？尋常的意義，乃指工資收入爲衆數者。普通公務員，並非謂其某種可量品質（如薪給，年齡）爲全體之算術平均數，而係謂此種

品質可以找得最大多數之相同者。列如普通公務員，公餘之暇，多閱覽報紙，而不讀古文；多觀電影而不聽音樂；月薪多爲百元之譜；住宅多在所在機關附近等。但所謂「普通人「仍非眞有其人，僅就某種徵性，舉其共同之點而言耳。平均數之所以能代表現象，蓋因其具有概括事物之抽象觀念。

在衆數上，吾人可以找出大多數之人，所謂謀大多數人之幸福，卽此等人之幸福。而算術平均數與中位數，則不免遠離實際，僅爲一數字觀念。惟有衆數方可找得最大多數之實例。衆數表示最常見之事物，獲得最多之例證，其應用亦極其普遍。成衣發售之衣商，須知身材之衆數，而非算術平均數；某處將設立一郵局，須知郵件之衆數，某人擬開設日用品商店，須知日用品之衆數，而非其他任何平均數。諸如此類，衆數均異常明確，而毫無決定之煩者。

四、衆數之優點與缺點

衆數之最大特點，亦其最大優點，卽絕不受極端量數之影響。某生成績爲零分，與算術平均數以極大之影響，而與衆數無涉。某國萬貫家財者少數，而衣食不周者千千萬萬，又一國人民均屬小康之家，兩者之平均財富，就算術平均而

言，或者大致相同，但就衆數而言，則不免大相逕庭。研究歷年之國民生活，在算術平均數，吾人僅可抽象說明一般生活程度之提高；而用衆數，則可指出大多數人生活之改善程度。Booth 之 ''London ,, 一書，即充分運用衆數者。每項職業，指出年齡分配之衆數；各業工貲，指出工貲之衆數。其描述現代城市之產業工人，即以衆數爲準繩。總之，衆數之最大優點，在於忽視特異之事項，而取其普通事項。此外，衆數之决定，無庸知道極大極小之量數，只須知道最大多數之量數，即可决定；且謂平均戶量爲每家4.23人，遠不如謂大多數之家庭爲四口之家；又不可測量之品質，亦可取衆數爲其代表。

衆數之缺點，即若干統計數列不可應用。若統計數列，並無典型可尋，而爲一羣不規則之數目，則衆數無由决定。衆數之爲用，原在指示典型，離此典型之數字，則可視爲另有原因，例如工資衆數爲二十元，其過與不及者，或由於機會有優劣，技術有精粗。苟統計數列所表現之事實，果有典型存在，則衆數必可指出之，惟衆數對於代表統計事實，只能指出典型之所在，欲表現其他徵性，自須藉助於其他統計

常數也。

第五節　幾何平均數

一、簡單幾何平均

設有n個量數

$$X_1, X_2, \cdots\cdots \cdots X_n,$$

則其幾何平均數G之公式如下：

$$G = \sqrt[n]{X_1 \quad X_2 \cdots\cdots\cdots\cdots X_n}$$

惟開n次方，事實上不可能，須利用對數求之：

$$\log G = \frac{1}{n}(\log X_1 + \log X_2 \cdots\cdots + \log X_n) = \frac{1}{n}\sum \log X$$

幾何平均數恆小於算術平均數，而大於倒數平均數。

若平均數着重於量數間的比率關係，而非注重其間之絕對差數，則應選擇幾何平均數。例如若8與13的相差，其重要性恰和13與18的相差相同，則8與18的平均數，應爲

$$M = \frac{8+18}{2} = 13$$

但若8與12的比率，如同12與18的比率一樣重要，則8與18的平均數應爲

$$G=\sqrt{8\times18}=\sqrt{144}=12。$$

今設有五個量數，X_1, X_2, X_3, X_4, X_5, 其中X_1與X_2小於算術平均數M，亦小於幾何平均數G，而X_3, X_4, X_5，大於M，亦大於G，則

$$M=\frac{X_1+X_2+X_3+X_4+X_5}{5}$$

或 $$5M=X_1+X_2+X_3+X_4+X_5$$

$$\underline{(M-X_1)+(M-X_2)=(X_3-M)+(X_4-M)+(X_5-M)\,;}$$

$$G=\sqrt{X_1\times X_2\times X_3\times X_4\times X_5}$$

或 $$G_5=X_1\times X_2\times X_3\times X_4\times X_5$$

$$\frac{G}{X_1}\times\frac{G}{X_2}=\frac{X_3}{G}\times\frac{X_4}{G}\times\frac{X_5}{G}$$

可見在算術平均數，各較小量數之所不及平均數者，恰等於各較大量數之所過者；而在幾何平均數，則平均數與各較小量數之比的乘積，恰等於各較大量數與平均數之比的乘積。一重差數，一重比率，彰彰明甚。

二、加權幾何平均

若各量數並非同等重要，則當分別重輕，各與以大小不

同之權數，設量數及其權數（或次數）如下：

量數	X_1	X_2	……	X_n
權數	W_1	W_2	……	W_n
（次數	f_1	f_2	……	f_n）

則加權幾何平均數Gw之公式如下：

$$G_W = \sqrt[\Sigma W]{X_1^{W_1} \quad X_1^{W_2} \cdots\cdots X_n^{W_n}}$$

或　$$\log G_W = \frac{1}{\Sigma W}(W_1 \log X_1 + W_2 \log X_2 + W_n \log X_n$$

$$= \frac{1}{\Sigma W} \Sigma W \log X$$

$$[\log G_W = \frac{1}{\Sigma f}(f_1 \log X_1 + f_2 \log X_2 + \cdots\cdots + f_n \log X_n)$$

$$= \frac{1}{\Sigma f} \Sigma f \log X$$

幾何平均數除兩項之簡單幾何平均（開平方）外必須運用對數計算。

三、幾何平均數之應用

甲、計算物價指數

幾何平均數最重要之用途，厥爲計算物價指數。因爲物

價指數從100漲到120，其增漲程度恰與從120漲到144樣同，

$$\frac{120}{100}=1.2=\frac{144}{120}$$

而並不與120漲到140一樣。Jcvons最初引用幾何平均數，以計算物價指數，蓋由此也。

譬如米每石由50元漲到55元，吾人並不甚感覺，而木炭每担由5元漲至10元，則以爲物價飛騰，同是漲5元，何以感覺不同，蓋因前漲者10%，而後者則增100%也。同理，肉價每斤由5角漲至6角，遠不若由2角漲至3角之重要，雖均爲增漲一角而程度則大不相同也。

某地簡易零售物價指數計算表

（二十九年八月）

民國二十六年＝100

商品	單位	基期物價 p_0	計算期物價 p_1	價比 $\frac{p_1}{p_0}\times100$	價比之對數 $\mathrm{Log}(\frac{p_1}{p_0}\times100)$
米	市升	.12	.40	333	2.5224
麵粉	市斤	.12	.55	458	2.6609
豆油	市斤	.20	.30	400	2.6021
食鹽	市斤	.15	.45	300	2.4771
猪肉	市斤	.22	.80	361	2.5611

木炭	市斤	.35	1.80	514	2.7110
白布	市尺	.14	90	643	2.3082
總計	×	×	×	×	18.3448

【註】基期物價爲民國二十六年之平均物價，計算期物價爲二十九年八月之平均物價

$$\log G=\frac{1}{n}\Sigma\log X=\frac{1}{n}M\log\left(\frac{p_1}{p_0}\times 100\right)$$

$$=\frac{1}{7}\times 18.3448=2.6204$$

G＝417（二十九年八月之簡易零售物價指數）

乙、其他用途

凡數量之增減，依比率升降者，求其平均數或增減率均應取幾何平均。假定我國人口之增殖率與日本相同，則我國增四千七百萬，日本只能增加七百萬。馬爾薩斯謂人口之增加依幾何級數，在太平盛世，若國民不用人爲方法節制生育，此說確爲不移之論。此外複利率及學生學力之計算，均須用幾何平均。

若以P_0表示本數，n 表示時期，P_n表示本數與增加數

之和,r表示增加率，則

$$Pn=Po(1+x)n$$ (由Po,n,r求Pn之公式)

$$Po=\frac{Pn}{(1+r)n}$$ (由Pn,n,r求Pn之公式)

$$n=\frac{\log Pn-\log Po}{\log(1+r)}$$ (由Po,Pn,r求n之公式

$$r=n\sqrt{\frac{Pn}{Po}}-1$$ (由Po,Pn,n求r之公式)

茲舉例如下：

(1)求Pn 某地民國元年年中有十萬人，設其增殖率爲每年1%,問民國二十九年年中之人口數爲若干？

$Po=100,000$，$n=28$，$r=.01$

$\underline{Pn=Po(1+r)n}$

$$\log Pn=\log Po+n\log(1+r)$$

$$=\log 100,000+28\times\log(1.01)$$

$$=5+28\times.0043$$

$$=5.1204$$

$\therefore Pn=132,000$(民國二十九年之人口)

(2)求Po 某地民國二十九年年中之人口數爲132,000

設已知其增殖率爲1%,問民國九年中之人口爲若干？

$$Pn=132,000,\ n=20,\ r=.01$$

$$Po=\frac{Pn}{(1+r)n}$$

$$\log Po=\log Pn-n\log(1+r)$$

$$=\log 132,000-20\log(1.01)$$

$$=5.1206-20\times .0043$$

$$=5.0344$$

$\therefore$ Po=108,250(民國九年的人口)

(8)求 n 某地民國十年之人口數爲 109,330,設其增殖率爲1%.問幾年後之人口將增至120,700？

$$Po=109,330,Po=120,700,r=.01$$

$$n=\frac{\log Pn-\log Po}{\log(1+r)}$$

$$=\frac{\log 120.700-\log 109,330}{\log(1\cdot 01)}$$

$$=\frac{5.0817-5.0387}{\cdot 0043}$$

$$=\frac{.0430}{\cdot 0043}$$

$= 10$ (民國二十年)

(4)求 r 某地民國六年之人口數爲105,070,至民國二十六年增至128,100,問其平均增殖率爲幾何?

$$Po = 105,070, \quad Pn = 128,100, \quad n = 20$$

$$r = \sqrt[n]{\frac{Pn}{Po}} - 1$$

$$1 + r = \sqrt[n]{\frac{Pn}{Po}}$$

$$\log(1+r) = \frac{1}{n}(\log Pn - \log Po)$$

$$= \frac{1}{20} \times (\log 128,100 - \log 105,070)$$

$$= \frac{1}{20}(5.1077 - 5.0215)$$

$$= \frac{1}{20} \times .0860$$

$$= .0043$$

$$\therefore 1 + r = 1.01$$

$$\therefore r = .01 = 1\%$$

x複利率計算法可仿以上各例,Po爲本金,Pn爲本利和

，r 為複利率， n 為年數。

第六節　倒數平均數

倒數平均數　H為各量數的倒數的算術平均數的倒數。

各量數：　$X_1, X_2, \cdots\cdots X_n$

各量數的倒數：　$\frac{1}{X_1}, \frac{1}{X_2}, \cdots\cdots \frac{1}{X_n}$

各數量的倒數的算術平均數 $\frac{1}{n}\left(\frac{1}{X_1}+\frac{1}{X_2}+\cdots+\frac{1}{X_n}\right)=\frac{1}{n}\Sigma\frac{1}{X}$

各量數的倒數的算術平均數的倒數，即倒數平均數

$$H=\frac{1}{\frac{1}{n}\Sigma\frac{1}{X}}=\frac{n}{\Sigma\frac{1}{X}}$$

若各量數各有其不同之次數或權數 $f_1, f_2 \cdots\cdots\cdots\cdots f_n$　則

$$H=\frac{1}{\frac{1}{\Sigma f}\Sigma\frac{f}{X}}=\frac{\Sigma f}{\Sigma\frac{f}{X}}=\frac{n}{\frac{f_1}{X_1}+\frac{f_2}{X_2}+\cdots+\frac{1}{X_n}}$$

倒數關係之統計事項不多，故倒數平均數之應用有限通常用以求平均速率及有倒數關係之平均價格。舉例如下：

(1)某輪船溯江而上，先以每小時30里行60里，後以每

小時20里行60里(注意所行里數相同)，問其平均速率如何？

$$H=\frac{n}{\frac{1}{X_1}+\frac{1}{X_2}}=\frac{2}{\frac{1}{30}+\frac{1}{20}}=\frac{2}{\frac{2+3}{60}}=\frac{120}{5}=24\text{里}$$

蓋因前後共費 5小時，行120里，故平均每小時行 24里。若用算術平均數 $M\frac{30+20}{2}=25$里，則錯矣。

若該輪船先以每小時30里行200 里，再以每小時20里行100 里，最後以每小時15里行50里，始達目的地，(注意各種速率所行里數不同)。問此段行程中平均速率幾何？

$$H=\frac{f_1+f_2+f_3}{\frac{f_1}{X_1}+\frac{f_2}{X_2}+\frac{f_3}{X_3}}=\frac{200\times100+50}{\frac{200}{30}+\frac{100}{20}+\frac{50}{15}}$$

$$=\frac{4+2+1}{\frac{4}{30}+\frac{2}{20}+\frac{1}{15}}=\frac{7}{\frac{8+6+4}{60}}=\frac{70}{3}=23\frac{1}{3}\text{里}$$

蓋因每小時30里行200里，費$6\frac{2}{3}$小時，每小時20里行100里，費5小時；每小時15里行50里，費$3\frac{1}{3}$時，前後費15小時，共行350 里，每小時速率自爲$\frac{350}{15}=23\frac{1}{3}$里。

(2) 物價以「每元若干」表示者，平均價格應取倒數平均

數，因每元若干係為倒數關係，其價格實際上應為每件若干元（小數）。例如蛋價六月為每元15個，七月為每元12個，八月為每元10個，假定三個月交易量大致相同（權數或次數相同）問三個月之平均蛋價如何？

$$H=\frac{n}{\frac{1}{X_1}+\frac{1}{X_2}+\frac{1}{X_3}}=\frac{3}{\frac{1}{15}+\frac{1}{12}+\frac{1}{10}}$$

$$=\frac{3}{\frac{4+5+6}{60}}=\frac{180}{15}=12$$

蓋因每元15個，則每個 $6\frac{2}{3}$分；每元12個，則每個 $8\frac{1}{3}$分；每元10個，則每個10分；平均價格自為

$$\frac{1}{2}\left(6\frac{2}{3}+8\frac{1}{3}+10\right)=\frac{25}{3}=8\frac{1}{3}\text{分}$$

每元當為 $100\div8\frac{1}{3}=\frac{300}{25}=12$個

若用算術平均謂平均每元為$\frac{15+12+10}{3}=\frac{37}{3}=12\frac{1}{3}$個，則錯矣。

今有某甲買肥皂三種，每元5 塊者10塊，每元四塊者12

塊，每元 3塊者15塊問平均每元買肥皂幾塊？

$$H=\frac{f_1+f_2+f_3}{\frac{f_1}{X_1}+\frac{f_2}{X_2}+\frac{f_3}{X_3}}=\frac{10+12+15}{\frac{10}{5}+\frac{12}{4}+\frac{15}{3}}$$

$$=\frac{37}{2+3+5}=\frac{37}{10}=3\frac{7}{10}\text{塊。}$$

因某甲費去10元，買得肥皂37塊，平均每元自爲$3\frac{7}{10}$塊。

第七節　各種平均數之比較

一、如何選擇適用之平均數

由以上各節之討論，平均數之功用當甚明顯：即在用少數之簡單數字，以表明複雜之數列。萬千量數，吾人之腦筋一時不能理解，必須加以歸類、化簡、平均。至於應選擇何種平均數，殊難一概而論：總以能表現統計事項之顯著特色與主要性質爲宜。而不同之數列往往須用不同之平均數表示之，大抵優良適當之平均數具備下列諸條件：

1.倘有典型存在，必能表現之：

2.與極端量數以應有的影響：

3.不因計算方法不同而變更其數值，或移動其位置；

4.計算簡易。

各種平均數，均有長處與缺點，吾人當依分析之目的以及材料之性質，以決定究竟應擇取何種平均數：

1.若重視各量數之差數，欲取長補短，則用算術平均數。算術平均應否加權，須視各量數之重要性而定；又可否計算百分數或統計係數，亦須視材料而定。

2.為避免極端量數之影響，宜用中位數及衆數，若資料確屬同質，量數相當集中而目的在表示大多數之實際典型，則尤以衆數最為適合。

3.為比較多數複雜之數列，資料頗不同質，亦不甚集中，則可求十分位數，百分位數，十分級之算術平均數等以比較之。各組之百分數亦為良好之比較標準。

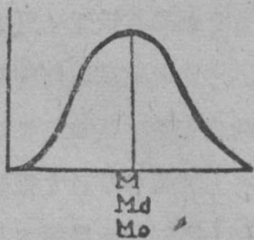
4.若注重量數間之比率，如物價之漲跌，人口之增殖，複利率之計算，學力之增進等，均宜用幾何平均數。

5.若各量數係以倒數關係表出者，如速率之言每小時若干里，物價之言每元若干件者，則當求倒數平均數。

二、各種平均數之關係

1.由算術平均數、中位數、衆數三者之地位，可略知統

計數列分配之梗概：若成對稱分配，則三者，合而為一；

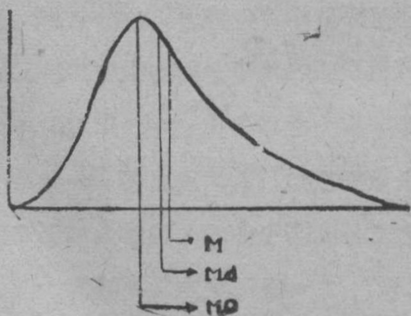


若量數集中於較小數值，而呈略不對稱之分配，則中位數大於衆數而小於算術平均數，且大略有

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

或 $$Mo=M-3(M-Md)$$

之關係；

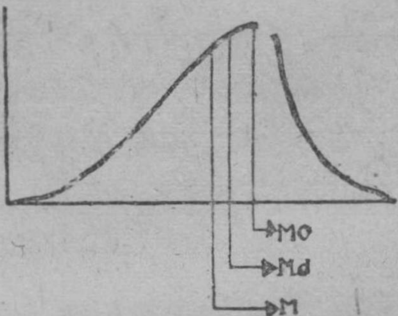


若量數集中於較大數值，而呈略不對稱之分配，則中位數小於衆數而大於算術平均數亦大略有

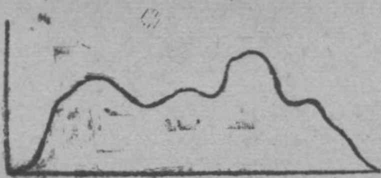
$Mo-M=3(Md-M)$　　　　（試與上面比較）

或　$Mo=M-3(M-Md)$

之關係；



若量數不同質，則衆數不確定，而成多峯分配。



2.算術平均數，幾何平均數，倒數平均數，可由數學公

式證明成下列關係

$$M>G>H.$$

證明方法頗煩，茲從略。

兩個量數之幾何平均數等於算術平均數與倒數平均數之乘積的幾何平均數，因

$$M=\frac{a+b}{2}\qquad H=\frac{2}{\frac{1}{a}+\frac{1}{b}}=\frac{2ab}{a+b}$$

故 $$MH=\frac{a+b}{2}\times\frac{2ab}{a+b}=ab,$$

$$\sqrt{MH}=\sqrt{ab}=G$$

例如4與9兩量數，

$$M=\frac{4+9}{2}=\frac{13}{2},H=\frac{2}{\frac{1}{4}+\frac{1}{9}}=\frac{72}{13},G=\sqrt{4\times9}=6$$

而 $$6=\sqrt{\frac{13}{2}\times\frac{72}{13}}$$

故 $$G=\sqrt{MH}$$

且 $$M=\frac{13}{2}>\frac{12}{2}=6=G,H=\frac{72}{13}<\frac{72}{12}=6=G$$

故 $M>G>H.$

第六章　相差度

引　言

前章所論平均數，吾人專注意統計數列之中心位置或集中趨勢。所謂統計數列，乃指一羣人或事物具有若干徵性（如所屬之地域，發生之時間，賦有之屬性等），吾人按照其某一徵性（變量，如年齡），加以分類歸組，而按預定次序排列之。欲陳示此種數列，或則列之於表，或則繪之於圖，但為化繁為簡，並便於與其他數列比較起見，吾人須要確定或計算幾個數值，以顯明數列之徵性，此種數值吾人稱之曰表徵數，通常所取之表徵數如下：

1. 取一平均數，以定中心位置或表集中趨勢，此前章之目的也。
2. 取一相差度，以表示量數之分佈情形或離中趨勢，本章專述此點。
3. 取一偏斜度，以表示分配不對稱之程度。此項留待下章討論。

第一節　相差度之意義與種類

一、相差度之意義

甲乙兩國人民之平均財富相等，未必兩國之經濟生活相同，因甲國少數人腰横萬貫，大衆則地無立錐；而乙國或皆爲中產之家，平均數只能表現取長補短，將全國私人財富集中以後再行平均分配與全體人民之情形。甲乙兩班學生平均成績相等，未必兩班學生之程度一模一樣，因甲班優秀者多而低劣者亦不少，乙班則均中庸發憤之流，扯長補短，兩班成績仍可相等也。今欲知甲乙二國貧富懸殊與分配均勻之情形，甲乙二班成績是否參差，惟有取另一表徵數以資測量，卽相差度是也。故相差度者表示各量數分布情形或離中程度之數量也。

二、相差度之種類

測量離中程度之相差度，通常用者有下列幾種：

1.全距　卽最大量數與最小量數之相差數。

$$Rg = (\text{最大量數}) - (\text{最小量數})$$

2.四分位差　卽上下四分位數之距離之一半。

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

3.平均差　即各量數與平均數(M或Md,通常取Md)之離差之絕對值之算術平均數：

$$M.D.=\frac{\Sigma|d|}{N}\ (d=X-Md,)$$

4.標準差　即各量數對算術平均數之離差之平方之算術平均數的平方根：

$$Sd=\sqrt{\frac{\Sigma d^2}{N}}\quad (d=X-M)$$

5.均互差　即各量數相互之差數之算術平均數：

$$g=\frac{S}{\frac{n(n-1)}{2}}=\frac{2S}{n(n-1)}$$,S爲相互差數之和。

三、相對相差度與離散係數

上述各種相差度，均用具體單位（如歲、斤、尺、升）表示，謂之絕對相差度。欲與性質不同之數列比較，則惟有引用相對相差度，即用一比率以消滅具體單位。俗語謂「衣不長寸，鞋不長分」，蓋因衣本長，鞋本短，相差不得超

過某種比例也。若謂無論縫衣製帽作鞋，均不得有百分之三之差誤，則可以概括而言，無庸分別言其分寸之單位矣。相對相差度，可取下列幾種：

1.四分差度$=\dfrac{Q_3-Q_1}{Q_3+Q_1}$ $\left(\dfrac{\text{四分位差}}{\text{上下四分位數之平均數}}\right)$

2.平均差度$=\dfrac{\text{M. D.}}{\text{Md}}$ $\left(\dfrac{\text{平均差}}{\text{中位數}}\right)$

3.標準差度$=\dfrac{\text{Sd}}{\text{M}}$ $\left(\dfrac{\text{標準差}}{\text{算術平均數}}\right)$

惟通常所用之相對相差度，乃標準差對算術平均數之百分數，稱為離散係數。

$$\text{C.V.}=\frac{100\text{Sd}}{\text{M}}$$

第二節　全距

全距為表示差異情形最簡單之方法，亦為最粗略之方法。在枚舉數列或簡單次數數列即最大量數與最小量數之差數。例如某班學生年齡最長者為35歲，最輕者為19歲則全距為

$\text{Rg}=35-19=16$歲；

又如某班統計學成績最優者爲95分，最次者爲35分，則

$$Rg = 95 - 35 = 60\text{分}。$$

在分組次數數列，全距爲最大組限與最小組限之差數，例如某班學生年齡分配最大組限爲 36 歲，最小組限爲18歲，則

$$Rg = 36 - 18 = 18\text{歲}。$$

全距僅取兩極端量數以決定其數值，對於中間量數完全漠視，殊屬不當。例如某班學生僅有一生因害病甚久，致成績爲35分，其餘均在60分以上，遽謂該班成績參差不齊，甯有是理。又如某機關僅有主管人月薪 600 元，一二錄事月薪25元、其餘均在50—400元之間，乃謂該機關薪給全距爲575元，待遇懸殊，亦不足以表明眞象也。

全距之應用，通常在表示金融、證券市場，每日價格之相差情形，讀者閱報卽可見某日外匯最高幾何，最低若干，兩者相減，卽可得是日價格變動之全距。至於稍稍精密之統計分析，未有用全距以作相差度者。

第三節　四分位差(附十分位差，百分位差)

上下四分位數之差數以及其他配對之分割數（如第一個

與第九個十分位數，第三個與第七個十分位數，第二十五個與第七十五個百分位數等）之差數，自與離中程度有關。惟無論數列分配情形如何，上下四分位數之間及其外之分配情形如何，只須四分位數相等，則所得四分位差相等，是其缺點。是故此種相差度亦頗不靈敏，惟在粗略之分析上，仍不失爲簡單明瞭之相差度，故用之者仍甚多。

四分位差乃上下四分位數相差數之半，約略與概差相等。

四分位差之求法，只須求得四分位數，代入公式卽得

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

四分位數之求法，見第五章第三節。

	年齡	人數	累積次數	
	18－20	3	32F₁2	
$Q_1 \longrightarrow$	20－22	30	33	$\longrightarrow \frac{121}{4}=30\frac{1}{4}$
	22－24	43	76, F_3,	
$Q_3 \longrightarrow$	24－26	21	97	$\longrightarrow \frac{3\times121}{4}=90\frac{3}{4}$
	24－28	17	114	
	23－30	5	119	

30－32	1	120
32－34	－	120
34－36	1	121

$$Q_1=L+\frac{\frac{N}{4}-F_1}{f}\times i=20+\frac{30\frac{1}{4}-3}{30}\times 2=21.82$$

$$Q_3=L+\frac{\frac{3N}{4}-F_3}{f}\times i=24+\frac{90\frac{3}{4}-76}{21}\times 2=25.490$$

$$Q.D.=\frac{Q_3-Q_1}{2}=\frac{25.40-21.82}{2}=1.79$$

$$\text{四分差度}=\frac{Q_3-Q_1}{Q_3+Q_1}=\frac{25.40-21.82}{2}=1.079$$

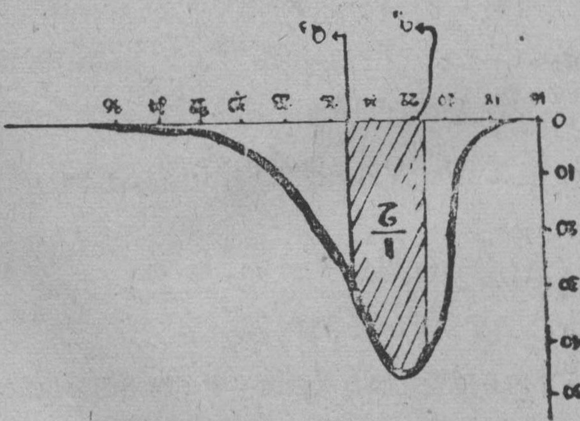
上下四分位數之間有次數之一半，而兩四分位數之中點為½（25.4+21.8）=23,6則上下兩四分位數必為

23.6±1.8

故吾人描述上表統計數列，可一言以蔽之曰：某班學生齡年之中位數為23.3歲（或算術平均數為23.7歲亦可）並有

一半學生在23.6±1.8歲之間。集中及離中情形均已表現，遂可與同類性質之數列相比較矣。若欲與不同性質之數列比較，則謂四分差度爲.076或四分差離散係數爲7.6%可也。

在對稱分配中，M.Md與M。三者合一，而四分位差Q.D.與槪差P.E.相等。



第四節　平均差

統計數列各量數對某平均數或其他定點之差數謂之離差

。離差之絕對值(|d|)之算術平均數即平均差。

一、枚舉數列之平均差

今有五人其年齡如下： 21,22,23,26,28

Md＝第三人之年齡＝223歲

$$M = \frac{21+22+25+26+28}{5} = 24$$

離中位數之平均差：

$$M.D. = \frac{\sum|d|}{N}$$

$$= \frac{|23-223|+|22-23|+|23-23|+|26-23|+|28-23|}{5}$$

$$= \frac{2+1+0+3+5}{5} = 2.2$$

$$\text{平均差度} = \frac{M.\ D.}{Md} = \frac{2.2}{23} = .096$$

離算術平均數之平均差

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

$$=\frac{|21-24|+|22-24|+|23-24|+|26-24|+2|8-24|}{5}$$

$$=\frac{3+2+1+2+4}{5}=2.4$$

吾人遂謂此五人年齡之中位數爲23歲（或算術平均數爲24歲）平均差爲2.2歲（或2.4歲）。

二、簡單次數數列之平均差

（分組次數數列以組中值代表量數X同此）

	工　資 (X)	工　人 (f)	離　差 $\lvert d=x-Md\rvert$	$f\lvert d\rvert$
	14	25	1	25
Md——→	15	20	0	0
	16	15	1	15
	18	10	3	30
	20	6	5	30
	24	3	9	27
	28	1	13	13
	總　計	80	×	140

$$\text{M d} = \text{第}\ \frac{N+1}{2}\ \text{個工人之工資}$$

$$= \text{第}\ 40\frac{1}{2}\ \text{個工人之工資}$$

$$= 15\text{元}$$

$$\text{M.D.} = \frac{\Sigma f|d|}{\Sigma f} = \frac{140}{80} = 1.75\text{元}$$

$$\text{平均差度} = \frac{\text{M.D.}}{\text{Md}} = \frac{1.75}{15} = .117$$

吾人遂謂此羣工人工資之中位數爲15元，平均差爲1.75元。

同法可求算術平均數及對算術平均數之平均差。

三、等組距分組次數數列平均差之簡捷計算法

$$\text{M.D.} = \frac{\text{Mf}|d'| + C(\text{Nb} - \text{Na})}{\text{Mf}} \times i$$

M,d代表假定中位數，卽中位數所在組的組中値

d，代表各組中値與假定中位數之離差而以組距爲單位，$d' = \dfrac{X - M'd}{i}$

C 代表眞正中位數 Md 與假定中位數M,d 之差數而以

組距爲單位，卽 $C = \frac{Md - M'd}{i}$

Na 代表比 Md 大之組中値相應對之次數之和

Nb 代表比 Md 小之組中値相應對之次數之和

年 齡	學生人數	組中値 x	$\lvert d'\rvert = \frac{\lvert X-M'd\rvert}{i}$	$f'\lvert d'\rvert$
18－20	3 Nb	19	2	6
20－22	30 Nb	21	1	30
Md→22－24	43 Nb	23 M'd.	0	0
24－26	21 Na	25	1	21
26－28	17 Na	27	2	34
28－30	5 Na	29	3	15
30－32	1 Na	31	4	4
32－34	－ Na	33	5	0
34－36	1 Na	35	6	6
總 計	121	×	×	116

$Md = 23.3$（見前）中在 22－24 一組，故以23爲假定中位數。即 $M'd = 23$

$$C = \frac{Md - M'd}{i} = \frac{23.3 - 23}{2} = .15$$ （C有時爲負數）

19,21,23 均小於 Md, 故 $Nd = 3 + 30 + 43 = 76$

其餘各組中值均大於Md,故 $Nd = 121 - 76 = 45$

$\sum f|d'| = 116, \sum f = 121, i = 2$

代入公式

$$M.D. = \frac{\sum f|d'| + C(Nb - Na)}{\sum f} \times i$$

$$= \frac{116 + .15(76 - 45)}{121} \times 2$$

$$= \frac{120.65 \times 2}{121} = 1.99$$

$$= \frac{120.65 \times 2}{121} = 1.99$$

$$平均差度 = \frac{M.D.}{Md} = \frac{1.99}{23 \cdot 3} = .086$$

吾人遂謂此班學生年齡中位數爲23.3歲，平均差爲2，若與不同性質之數列相比較X，則謂平均差度爲 8.6%

同法可求對算術平均數(23.7歲)之平均差。

(附)等組距分組次數數列平均差簡捷計算法公式之證明

分組	組中值	次數
a_1 — b_1	X_1	f_1
a_2 — b_2	X_2	f_2
⋮	⋮	⋮
a_{j-1} — b_{j-1}	X_{j-1}	f_{j-1}
a_j — b_j	$X_j \leftarrow Md$	f_j
a_{j+1} — b_{j+1}	X_{j+1}	f_{j+1}
⋮	⋮	⋮
a_{t-1} — b_{t-1}	X_{t-1}	f_{t-1}
a_t — b_t	X_t	f_t

設Md在a_j-b_j組，該組之組中值爲X_j，而$Md \angle X_j$。(Md ⅂X_j，證法同。)

$$\sum f|X-Md|$$

$$=f_1(Md-X_1)+f_2(Md-X_2)+\cdots+f_{j-1}(Md-X_{j-1})$$

$$+ fj(Xj - Md) + fj+1(Xj+1 - Md) + \cdots + ft-1(Xt-1 - Md) + ft(Xt - Md)$$

$$= \sum_{k=1}^{j-1} fk(Md - Xk) + \sum_{k=j}^{t} fk(Xk - Md)$$

$$= \sum_{k=1}^{j-1} fk(\overline{Md - Xj} + \overline{Xj - Xk}) + \sum_{k=j}^{t} fk(\overline{Xk - Xj} + \overline{Xj - Md})$$

$$= \sum f|X - Xj| + (Md - Xj)\left(\sum_{k=1}^{j-1} fk - \sum_{k=j}^{t} fk\right)$$

并　　令　$M'd = Xj$

$Nb = \sum_{k=1}^{j-1} fk$（比Md小之各組中值相應對之次數之和）

$Na = \sum_{k=j}^{t} fk$（比Md大之各組中值相應對之次數之和）

則　$\sum f|X - Md| = \sum f|X - M'd| + (Md - M'd)(Nb - Na)$

故 $$M.D.=\frac{\sum f|X-M'd|+(Md-M'd)(Nb-Na)}{\sum f}$$

$$=\frac{\sum f\left[\frac{|X-M'd|}{i}+\frac{Mb-M'd}{i}(Nb-Na)\right]}{\sum f}\times i$$

令 $$|d'|=\frac{|X-M'd|}{i},\ C=\frac{Md-M'd}{i}$$

則 $$M.D.=\frac{\sum f|d'|+C(Nb-Na)}{\sum f}\times i$$

四、平均差之特質

平均差係各量數對中位數或算術平均數（有時亦可用衆數）之離差絕對値之算術平均數，意義甚爲明顯，且係由全體次數計算而得，每一次數均有影響，不若全距之僅取兩極端數，亦不若四分位差之僅取兩四分點以爲計算之根據，中位數與各次數距離之和爲最小，在論中位數時已加證明。故平均差以取對中位數之離差較爲相宜。

但平均差有一缺點，卽須取離差之絕對値。因離差有正

有負故也。為避免取絕對值起見，乃取對算術平均數之離差平方為計算相差度之根據，平方皆為正數也。（且在機遇理論上，離差之重要性視其平方而定）。

第五節　標準差

為避免取絕對值之煩，以各量數對算術平均數之離差平方之算術平均數的平方根為相差度，即標準差是也。

一、枚舉數列之標準差

年　齡 X	d = X −	b²
21	−3	9
22	−2	4
23	−1	1
26	2	4
29	4	16
120($\sum X$)	0($\sum d$)	34($\sum d^2$)

$$= \frac{\sum X}{N}\ \frac{120}{5} = 24$$

$$Sd = \sqrt{\frac{\sum d^2}{N}} \quad \left[即 Sd = \sqrt{\frac{\sum (X-M)^2}{N}}\right]$$

$$= \sqrt{\frac{34}{5}} = \sqrt{6.8} = 2.6歲$$

$$標準差度 = \frac{Sd}{M} = \frac{2.6}{24} = .108$$

$$離散係數：\quad C.V. = \frac{100\,Sd}{M} = 10.8$$

吾人乃可簡述此五人之年齡：N=24，Sd=2.6，標準差度爲.108，或離散係數爲10.80

二、簡單次數數列之標準差

$$公式：\quad Sd = \sqrt{\frac{\sum f(X-M)^2}{Mf}} = \sqrt{\frac{\sum fd^2}{N}}$$

（分組次數數列以組中值代表量數X，亦用此公式）

工資 (X)	工人 (f)	fX	$d=X-M$	d^2	fd^2
14	25	350	−2.13	4.5369	113.4225
15	20	300	−1.13	1.2769	25.5380
16	15	240	−.13	0199	.2535
18	10	190	1.87	3.4969	34.9690
20	6	120	3.87	14.9769	89.8614
24	3	72	7.87	61.9369	185.8107
28	1	28	11.87	140.8969	140.8969
總計	80($\sum f$)	1290 ($\sum fX$)	×	×	590.7520($\sum fd^2$)

$$M=\frac{\sum fX}{\sum f}=\frac{1290}{80}=16.13\text{元}，\sum fd^2=590.7520,$$

$$Sd=\sqrt{\frac{\sum fd^2}{\sum f}}=\sqrt{\frac{590.7520}{80}}$$

$$=\sqrt{7.3844}=2.72\text{元}$$

$$\text{標準差度}=\frac{Sd}{M}=\frac{2.72}{16.13}=.168$$

$$\text{離散係數}\quad C.V.=\frac{100Sd}{M}=16.8\text{。}$$

吾人可簡述此80個工人工資分配如下：平均每人每月16.13元，標準差爲2.72元，離散係數爲16.8%。

三、簡單次數數列與不等組距分組次數數列之標準差的簡捷計算法

由上述公式 $Sd=\sqrt{\frac{\sum f(X-M)^2}{\sum f}}$ 計算標準差，乘方

與乘法，殊嫌太繁，若用變更原點方法，則計算較簡。

$$\sum f(X-M)^2=\sum f(X-A+A-M)^2$$

$$=\sum f[(X-A)-(M-A)]^2$$

令 $x=X-A$，　$C=M-A$，則

$$\sum f(X-M)^2=\sum f(x-C)^2$$

$$=\sum f(x^2-2Cx+C^2)$$

$$=\sum fx^2-2C\sum fx+C^2\sum f$$

但 $\sum fx=\sum f(X-A)$

$$=\sum fX-A\sum f$$

$$=M\sum f-A\sum f$$

$$=(M-A)\sum f$$

$$=C\sum f$$

故 $\sum f(X-M)^2=\quad fx^2-2C(C\sum f)+C^2\sum f$

$$=\sum fx^2-2C^2\sum f+C^2\sum f$$

$$=\sum fx^2-C^2\sum f$$

故 $Sd=\sqrt{\dfrac{\sum f(X-M)^2}{\sum f}}$

$$=\sqrt{\frac{\Sigma fx^2-C^2\Sigma f}{\Sigma f}}$$

$$=\sqrt{\frac{\Sigma fx^2}{\Sigma f}-C^2}\left[=\sqrt{\frac{\Sigma f(X-A)^2}{\Sigma f}-(M-A)^2}\right]$$

此卽變更原點後，計算標準差之公式也。式中：

X 代表各量數與假定算術平均數之離差，$x=X-A$。

C 代表算術平均數與假定算術平均數之差數，$C=M-A$.

工資 X	工人 (f)	$x=X-A$	fx	x^2	fx^2
14	25	−2	−50	5	100
15	20	−1	−20	1	20
16(A)	15	0	0	0	0
18	10	2	20	4	40
20	6	4	24	16	96
24	3	8	24	64	192
28	1	12	12	144	144
總計	80(Σf)	×	10(Σfx)	×	592(Σfx)

$$\sum = A + \frac{\Sigma fx}{\Sigma f} = 16 + \frac{10}{80} = 16.13\text{元}，A = 16$$

$$C = M - A = 16.13 - 16 = .13$$

$$\Sigma f = 80，\Sigma fx^2 = 592$$

代入公式：

$$Sd = \sqrt{\frac{\Sigma f2^2}{\Sigma f} - C^2}$$

$$= \sqrt{\frac{592}{80} - .13^2}$$

$$= \sqrt{7.4 - 0169} = \sqrt{7.3831}$$

$$= 2.72\text{元}$$

四、等組距分組量數數列標準差之簡捷計算法

若分組量數數列之組距相等，則標準差之計算方法，可用變更單位方法而更簡單化。由上公式

$$Sd = \sqrt{\frac{\Sigma f(X-A)^2}{\Sigma f} - (M-A)^2}$$

$$= \sqrt{i^2 \times \left[\frac{\frac{\Sigma f(X-A)^2}{i^2}}{\Sigma f} - \frac{(M-A^2)}{i^2} \right]}$$

$$= \sqrt{i^2} \times \sqrt{\frac{\Sigma f\left(\frac{X-A}{i}\right)^2}{\Sigma f} - \left(\frac{M-A}{i}\right)^2}$$

令 $x = \frac{X-A}{i}$，$C = \frac{M-A}{i}$，則

$$Sd = i \times \sqrt{\frac{\Sigma fx^2}{\Sigma f} - C^2}$$

此即變更原點與單位後標準差之計算公式也，式中

i 代表組距

x 代表各量數與假定算術平均數之離差，而以組距為單位

即 $x = \frac{X-A}{i}$

C 代表算術平均數與假定算術平均數之差數，而以組距爲單位，即

$$C=\frac{M'-A}{i}。$$

茲舉例如下：

年　齡	學生人數 f	組中值 X	$x=\frac{X-A}{i}$	fx	x^2	fx^2
18－20	3	19	－2	－6	4	12
20－22	30	21	－1	－30	1	30
22－24	43	23(A)	0	0	0	0
24－26	21	25	1	21	1	21
26－28	17	27	2	34	4	68
28－30	5	29	3	15	9	45
30－32	1	31	4	4	16	16
32－34	—	33	5	0	25	0
34－36	1	35	6	6	36	36
總　計	121 (Σf)	×	×	44 (Σfx)	×	228 (Σfx^2)

$A=23, i=2, \Sigma f=121, \Sigma fx=44, \Sigma fx^2=228$

$$M = A + \frac{\Sigma fx}{\Sigma f} \times i = 23 + \frac{44}{121} \times 2 = 23.7$$

$$C = \frac{M - A}{i} = \frac{23.7 - 23}{2} = .35$$

$$Sd = i \times \sqrt{\frac{\Sigma fx^2}{\Sigma f} - C^2}$$

$$= 2 \times \sqrt{\frac{228}{121} - .35^2}$$

$$= 2 \times \sqrt{1.8843 - .1225}$$

$$= 2 \times \sqrt{1.7618}$$

$$= 2 \times 1.33$$

$= 2.66$歲.

$$\text{標準差度} = \frac{Sd}{M} = \frac{2.66}{23.7} = .112$$

離散係數　$C.V. = \frac{100Sd}{M} = 11.2$

五、標準差之特質

標準差係根據全體量數對算術平均數之離差求得，是各量數均與以應有之影響，此優點與平均差共有之。至其計算，取離差平方而非其絕對值，合乎數學原理，處理便利，此又勝於平均差一籌者。且再求進一步之分析，作機遇理論上之研究，惟有標準差能勝任之。求相關係數卽須利用標準差。故 YUle 氏主張，宜用標準差以表示相差度，惟簡明易解，則不及四分位差及平均差矣。

第六節　均互差

均互差，乃將各量數相互之差數相加，而以差數之項數除之。設量數有 n 個，則差數有 $\frac{n(n-1)}{2}$ 項，設差數之和爲 S，則均互差 g（因均互差係意大利教授 Gini 所首倡，故用 g 表之）之公式如下：

$$g = \frac{S}{\frac{n(n-1)}{2}} = \frac{2S}{n(n-1)}$$

一、枚舉數列之均互差

例如甲乙丙三人之年齡爲23,21,則

$$S=(25-22)+(25-21)+(22-21)$$

$$=3+4+1=8$$

$$g=\frac{2S}{n(n-1)}=\frac{2\times 8}{3\times 2}=2.67$$

今設有五個量數，按小大次第排列如下：

$$X_1,\ X_2,\ X_3,\ X_4,\ 5_5,$$

則 $$S=(X_5-X_4)+(X_5-X_3)+(X_5-X_2)+(X_5-X_1)+$$
$$(X_4-X_3)+(X_4-X_2)+(X_4-X_1)+$$
$$(X_3-X_2)+(X_3-X_1)+$$
$$(X_2-X_1)$$

$$=4X_5+2X_4+0X_3+(-2)X_2+(-4)X_1$$

若寫成與項數 5 之關係，則爲

$$S=(5-1)X_5+(5-3)X_4+(5-5)X_3+(3=5)X_2+$$
$$(1-5)X_1$$

再設有六個量數按數小次第排列如下：

$X_1, X_2, X_3, X_4, X_5, X_6,$

$$S = (X_6-X_5)+(X_6-X_4)+(X_6-X_3)+(XX_6-x_2)+$$
$$(X_6-X_1)$$
$$(X_5-X_4)+(X_5-X_3)+(X_5-X_2)+(X_5-X_1)+$$
$$(X_4-X_3)+(X_4-X_2)+(X_4-X_1)+$$
$$(X_3-X_2)+(X_3-X_1)+$$
$$(X_2-X_1)$$
$$= 5X_6+3X_5+1X_4+(-1)X_3+(-3)X_2+(-5)X_1$$
$$= (6-1)X_6+(6-3)X_5+(6-5)X_4$$
$$(1-6)X_1+(3-6)X_2+(5-6)X_3$$

由是可知若量數有 n 個，則

$$S = (n-1)Xn+(n-3)Xn-1+(n-5)Xn-_2+\cdots\cdots+$$
$$(1-n)X_1+(3-n)X_2+(5-n)X_3+\cdots\cdots$$
$$=(n-1)Xn+(n-3)Xn-1+(n-5)Xn-_2+\cdots\cdots$$
$$-(n-1)X_1-(n-3)X_3-(b-_5)X_3-\cdots\cdots$$
$$=(n-1)(Xn-X1)+(n-3)(Xn-1-X_2)+(n-5)$$
$$(Xn-_2-X_3)+\cdots\cdots。$$

X'	X''	$X'-X''$	p	$p(X'-X'')$
X_n	X_1	X_n-X_1	$n-1$	$(n-1)(X_n-X_1)$
X_{n-1}	X_2	$X_{n-1}-X_2$	$n-3$	$(n-3)(X_{n-1}-X_2)$
X_{n-2}	X_3	$X_{n-2}-X_3$	$n-5$	$(n-5)(X_{n-2}-X_3)$
⋮	⋮	⋮	⋮	
$\Sigma X'$	$\Sigma X''$	×	×	$\Sigma p(X'-X'')$

$$M=\frac{\Sigma X}{n}=\frac{\Sigma X'+MX''}{n}$$

$$S=Mp(X'-X'')$$

$$g=\frac{2S}{n(n-1)}\quad\frac{2\Sigma p(X'-X'')}{n(n-1)}$$

此即枚舉數列均互差之計算公式也，其中X'代表較大之一半量數（若項數爲奇數，則多一個）按由大而小之次第排列，X''代表較小之一半量數，由小而大排列之。

相對均互差，可取 $\frac{g}{M}$ 或 $\frac{g}{Md}$ 。

茲舉例如下：

學生分數		$X'-X''$	p	$p(X'-X'')$
X'	X''			
94	56	38	20	760
90	63	27	18	486
87	65	22	16	352
85	69	16	14	224
84	71	13	12	156
83	72	11	10	110
81	74	7	8	56
80	76	4	6	24
79	77	2	4	8
79	77	2	2	4
78	–	–	0	0
929	700	×	×	2180

$$\Sigma X'=920，\Sigma X''=700，\Sigma p(X'-X'')=2180$$

$$M=\frac{\Sigma X'+\Sigma X''}{n}=\frac{1620}{21}=77\text{分}$$

$$g=\frac{2S}{n(n-1)}=\frac{2\Sigma p(X'-X'')}{n(n-1)}$$

$$=\frac{2\times2180}{21\times20}=\frac{218}{21}=10.4\text{分}$$

$$\text{均互差度}=\frac{g}{M}=\frac{10.4}{77}=.135$$

故此班學生平均成績爲77分，均互差爲10.4分，均互差度爲13.5%，

二、次數數列之均互差

在次數數列，設$\Sigma f=n$，各量數（或組中值）及其次數如下：

$X_1, X_2, X_3, \cdots X_{i}, X_{i+1}, X_{i+2}, \cdots X_{k-2}, X_{k-1}, X_k$

$f_1, f_2, f_3, \cdots f_i, f_{i+1}, f_{i+2}, \cdots f_{k-2}, f_{k-1}, f_k$

則相互差數之和爲

$S=f_k X_k(n-f_k)+$

$f_{k-1}X_{k-1}(n-2f_{k}-f_{k-1})+$

$f_{k-2}X_{k-2}(n-2f_k-2f_{k-1}-f_{k-2})+$

$\cdots\cdots\cdots\cdots+$

$f_{i+1}X_{i+1}(n-2f_k-2f_{k-1}-\cdots-2f_{i+2}-f_{i+1})-$

$f_1X_1(n-f_1)-$

$f_2X_2(n-2f_1-f_2)-$

$f_3X_3(n-2f_1-2f_2-f_3)-$

$\cdots\cdots\cdots\cdots-$

$f_iX_i(n-2f_2-\cdots\cdots-2f_{i-1}-f_l)$.

年齡	組中値 X	人數 f	fX	p	pfX −	pfX +
18－20	19	3	57	－118	－6,726	－
20－22	21	30	630	－85	－53,550	－
22－24	23	43	989	－12	－11,868	－
24－26	25	21	525	52	－	27,300
26－28	27	17	459	90	－	41,310
28－30	29	5	145	112	－	16,240
30－32	31	1	31	118	－	3,658
32－34	33	－	－	119	－	0
34－36	35	1	35	120	－	4,200
總計	×	121	2871	×	72,144	92,708

$$M = \frac{MfX}{\Sigma f} = \frac{2871}{125} = 23.7\text{歲}$$

$$g = \frac{2S}{u(u-1)} = \frac{2\Sigma pfX}{n(n-1)}$$

$$= \frac{(92,708-72,144)}{121\times ...}$$

$$= \frac{20564}{8260}$$

$$= 2.83\text{歲}$$

$$\text{均互差度} = \frac{g}{M} = \frac{2.83}{23.7} = \frac{2.83}{23.7} = .119 = 11.9\%$$

第七節　各種相差度之比較

一、相差度之選擇

相差度之功用，在於測量各量數之分佈情形或離中程度，已如前述。故欲知量數之集中情形或欲知代表數值，應求一適當之平均數，而欲知量數之參差程度，則不可再求一適當之相差度。應求何種相差度，其原則可略述如下：

1.全距太粗，不足以表示離中程度。但有時亦可用以觀察懸殊情況之大概。

7.四分位差表示居中一半量數，最大(上四分位數)與最小(下四分位數)二者之距離之一半、較全距爲精，但仍未取全體量數爲計算之根據。較粗之分析，可取中位數（或算術平均數）及四分位差表示數列之徵性。

3.平均差係根據全體量數對平均數（$\sum d$或$\sum$或$\sum O$）之離差計算，惟採取絕對值，數學上之處理，殊不便利。通常可取中位數及平均差代表一數列。

4.標準差根據全體量數對算術平均數之離差平方計算而得，無取絕對值之弊。較精之分析，均取算術平均數與標準差以表示集中及離中情形。

5.均互差觀念簡明，極合論理，以其將各量數一一互相比較也。惟因計算麻煩，今日用之者，尚不多觀。

6.性質不同之數列（如某大學之學生年齡，與某小學學生之年齡）或單位不同數列（如美國公務員年俸以美元計，我國公務員月薪以我國國幣計）比較離中情形，應取相對相差度或離散係數。

二、相差度與次數分配之關係

1.全距包括全體次數所佔之距離。

2.上下四分位數之中點左右各取一四分位差之距離，即$\frac{Q_3+Q_1}{2}\pm Q.D.$，包括次數之一半。九倍四分位差約包括全體次數99%。

3.在對稱次數分配或略不對稱次數分配上，平均差之七倍半，約包括全體次數之99%。

4.在對稱次數分配或稍微偏斜之分配中，在算術平均數之左右各取⅔Sd（P.E.＝,6745Sd）約包括次數之半，各取 Sd 約包括次數之⅔（常態分配下，包括68.26%），若各取 2Sd，約包括95%，若各取 3Sd，則包括 99%矣。

三、各種相差度之相互關係

據 Chaddock 之計算，在完全對稱分配中，Sd. M.D. 與 Q.D.相互之間，有確定之關係如下：

$$Sd = 1.2533\,M.D. \qquad (\text{或} Sd = \tfrac{5}{4} M.D)$$

$$Sd = 1.4825\,Q.D. \qquad (\text{或} Sd = \tfrac{3}{2} Q.D)$$

$$M.D. = 7979\,Sd \qquad (\text{或} M.D. = \tfrac{4}{5} Sd)$$

$$M.D.=1.1843Q.D. \quad (\text{或}M.D.=\tfrac{6}{5}Q.D.)$$

$$(P.E.=)Q.D.=.6745Sd \quad (\text{或}Q.D.=\tfrac{2}{3}Sd)$$

$$Q.D.=.8453M.D. \quad (\text{或}Q.D.=\tfrac{5}{6}M.D.)$$

本章所舉某班 121 人年齡分配之例，係略不對稱之次數分配，試就以上關係比較如下：

$$sd=2.66，M.D.=1.99，Q.D.=1.79$$

1.由Sd估計 M.D. 及 Q.D.

$$M.D.=\tfrac{4}{5}Sd=\tfrac{4}{5}\times 2.66=21.3(1.99)$$

$$Q.D.=\tfrac{2}{3}Sd=\tfrac{2}{3}\times 2.66=1.77(1.79)$$

2.由M.D. 估計 Sd 及 M.D.

$$Sd=\tfrac{5}{4}M.D.=\tfrac{5}{4}\times 1.99=2.49(2.66)$$

$$Q.D.=\tfrac{5}{6}M.D.=\tfrac{5}{6}\times 1.99=1.66(1.79)$$

3.由Q.D. 估計 Sd 及 M.D.

$$Sd=\tfrac{3}{2}Q.D.=\tfrac{3}{2}\times 1.79=2.69(2.66)$$

$$M.D.=\tfrac{6}{5}Q.D,=\tfrac{6}{5}\times 1.79=2.15(1.99)$$

上述之 M.D. = 1.99 係對中位數之平均差，若用對算術平均數之平均差，估計結果當更為接近也。

第七章　偏斜度與峯度

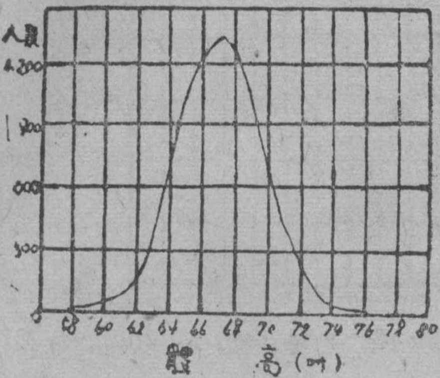
第一節　偏斜度之意義

平均數用以示數列之代表，相差度用以示數列之離散程度。而對於代表數值之左右兩側，依如何形式以行分配則尙未談及，其編斜是否對稱？若不對稱其偏斜之程度如何？偏斜之方向如何？實均有研究之必要。顯示次數分配之偏斜程度與方向者稱爲偏斜度 (Skewness) 亦表徵數之一種也。

多數數列之次數分配不能完全對稱，或左或右總有幾許偏斜，甚或有完全不對稱者，依于爾氏 (G. U. Yule) 之研究，次數分配有下列數種基本形式：

(一)對稱之次數分配(The symmetrical distribution) 此種分配在任何情形下均甚爲稀少，尤以經濟統計爲然，但在人類或生物測量之統計上較爲多見，例如下圖所示英國8855人之體高統計，爲近似於對稱之次數分配。

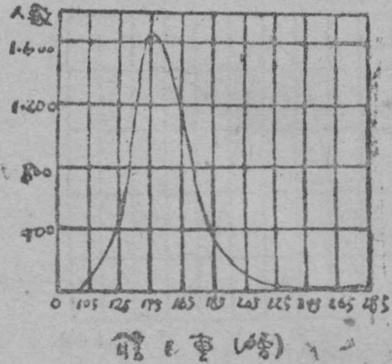
圖二十三　對稱之次數分配



（資料見Yulex:An Introduction to the Theory Of Statistlcsxxp.88)

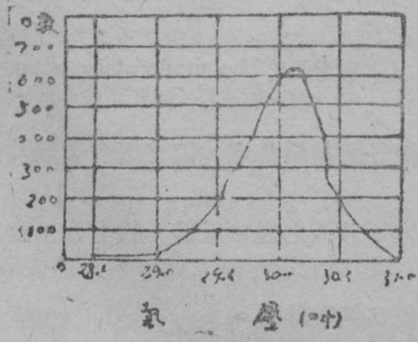
(二)略不對稱之次數分配(The moderately asymmetrcai distribution）此種分配甚爲普遍，但有偏右與偏左兩種，前者如下圖所示英國7749人之體重統計；後者如下圖所示英國1878-1880年之每日氣壓統計。

圖二十四　略不對稱之次數分配(偏右)



(資料見 Yulex:x„An Iatroduction to the Tdeory Of Satistics„ P.95)

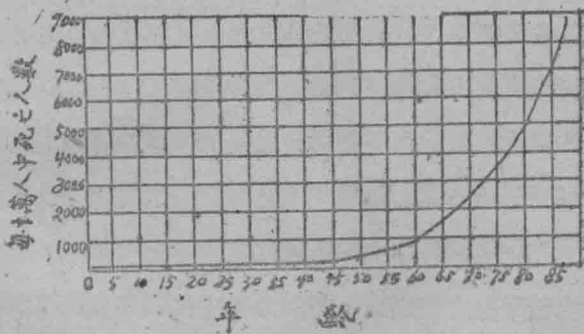
圖二十五　略不對稱之次數分配(偏左)



（資料見 Yule:An Intoduction to the Theory of Statistics p.96）

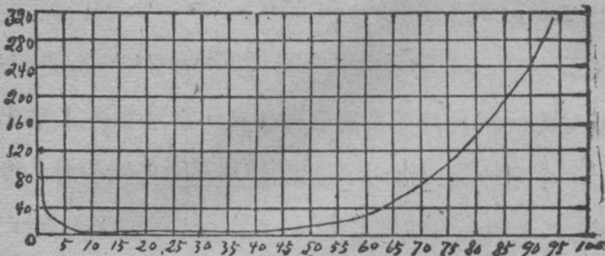
（三）完全不對稱之次數分配（The extemely asdmmetricaldistridution）此種分配有各種形式，常見者爲 J 字形，其爲 U 字形，前者如下圖所示1915—24年紐約男人每年患心病死亡人數依照年齡之分配，後者如下圖所示 1923 年紐約人口死亡率依照各種年齡之分配。

圖二十六　J字形之次數分配



（資料見紐約腺病心病學社所出版之紐約心病死亡統計）

圖二十七 U字形之次數分配



（資料見紐約衛生局第四十四次年報）

第二節 偏斜度之測定方法

偏斜度係用以表示次數分配之偏斜程度與方向者，而所謂偏斜程度與方向均對對稱之次數分配而言，故偏斜度之測定方法為比較對稱與不對稱時之情形，而加以測定之法。比較時可有三種不同之觀點，因此有三種不同之方法。

(一)就各種平均數觀之

次數分配之對稱與否對於衆數中位數與算術平均數之影響不同，在對稱之次數分配三者合而為一，不對稱之次數分配，三者分而為三。若次數曲線向右偏斜，則算術平均數因受極端數值之影響甚大，故向右移動甚多，中位數祇受次數

多少之影響而不受各項數値大小之影響，故雖亦向右移動，但其移動之程度較算術平均數爲微；反之若次數曲線向左偏斜，則算術平均數與中位數亦均向左移動，其移動之程度，算術平均數亦較甚於中位數。至於衆數則不論次數曲線之偏右或偏左，均能維持其原有之位置。故算術平均數與衆數之距離卽可用以測定偏斜程度，惟此距離含有測量單位，常以標準差除之，以求得其偏斜度，設Sk示偏斜度，M示算術平均數，M_o示衆數，Sd示標準差，則得

$$Sk=\frac{M-M_o}{Sd}$$

在次數分配對稱時，Sk=0，對稱而偏右時 Sk 爲正數；對稱而偏左時Sk爲負數。

但衆數不易確定，故在略不對稱之次數分配，又可用算術平均數與中位數（M_d）之距離之三倍示之卽得。

$$Sk=\frac{3(M-M_d)}{Sd}$$

（二）就四分位數之位置觀之

次數分配之對稱與否影響第一個四分位數與第三個四分

位數之位置，在對稱之次數分配，中位數與第一個四分位數(Q_1)及第三個四分位數(Q_3)之距離相等，即$Q_3-M_d=M_d-Q_1$；在不對稱時，$Q_3-M_d \neq M_d-Q_1$。因此觀察次數分配是否對稱，可由下式觀之：

$$(Q_3-M_d)-(M_d-Q_1)\begin{cases}=0\text{爲對稱。}\\ \neq 0\text{爲不對稱。}\end{cases}$$

而測量次數分配之偏斜程度，即可由$(Q_3-M_d)-(M_d-Q_1)$之差量示之。但此差量含有測量單位，故常以Q_3-Q_1除之，而得偏斜度(Sk)之公式如下：

$$Sk=\frac{Q_3+Q_1-2M_d}{Q_3-Q_1}$$

此法所求得之Sk，當次數分配對稱時，其數值亦爲零，其最大可能值爲1，最小可能值爲－1。若次數分配不對稱而偏右，Sk爲正數，偏左爲負數，故亦與上法所求得之Sk相同。

惟上列公式之除數若改用$Q.D.=\frac{Q_3-Q_1}{2}$者，則其所求得之Sk爲上列公式所求得之Sk之二倍。

(三)就各種動差之數值觀之

皮爾生氏(K. Pearson)規定n次動差(nth moment)之公式如次

$$\mu_n = \frac{1}{\sum f_t} \sum f_t (X_t - M)^n$$

(X_t示各數量，f_t示各次數，M示算術平均數)

在對稱之次數分配，各級單次動差等於0，即

$$\mu_{2p+1} = \frac{1}{\sum ft} \sum ft(Xt - M)^{2p+1} = 0$$

(p為正整數)

反之，在不對稱之次數分配，故各級單次動差不等於0，即

$$\mu_{2p+1} \neq 0。$$

因此次數分配之對稱與否？可以各級單次動差之數值示之，但一次動差恆等於0，故常用三次動差(μ_3)示之，然k_3含有測量單位，必須以同單位之差量或離散度除之，鮑萊氏以Sb^3爲除數，所得公式特稱x公式，以X示偏斜度；皮爾生

氏(K.Pearso)以k_2^3爲除數，所得公式特稱皮爾生公式，以K示偏斜度。兩公式如次：

$$K_1 = \frac{\mu_3}{Sd^3}$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

鮑萊公式所求得之K_1爲皮爾生公所求得之β之方根通常用鮑萊公式。

上述各種方法，第一法與第二法均較簡單。若已知算術平均數，衆數與標準差，即可由公式求得偏斜度；同樣，若已知中位數，第一個四分位數與第三個四分位數，亦可由公式見(224)頁求得偏斜度。故一般求偏斜度時，均用此二公式，但因衆數不易確定，中位數等屬於分割數系統，究不如第三法之精確可靠，因此第三法雖計算繁複，而在作精確之分析時必須用之，尤以配合曲線時爲然。

茲以吾人常舉之某班學生年齡分配爲例以明偏斜度之各種求法：

求偏斜度實例

(1) X	(2) f	(3) X－M	(4) (X－∑)³	(5) f(X－M)³
總　計	30			305.034
19	2	－3.4	－39.304	－78.608
20	5	－2.4	－13.824	－69.120
21	4	－1.4	－2.744	－10.976
22	7	＝0.4	－0.064	－0.448
23	4	0.6	0.216	0.864
24	2	1.6	4.096	8.192
25	4	2.6	17.576	70.303
27	1	4.6	97.336	97.336
29	1	6.6	287.496	287.496

該項數列之算術平均數(M)＝22.4，衆數M_o＝22，中位數Md＝22，第一個四分位數Q_1＝21，第三個四分位數Q_3＝24，標準差Sd＝2.30，故可用第一法，第二法求得偏斜度如下：

(一)第一法　$Sk=\frac{Md-M_o}{Sd}=\frac{22.4-22}{2.30}=\frac{0.3}{2.4}$

$=0.17$

(二)第二法　$Sk=\frac{Q_3+Q_1-2Md}{Q_3-Q_1}=\frac{24+21-2\times 22}{24-21}$

$$=\frac{1}{3}=0.33$$

至於第三法，須先求得μ_3，由表上得

$$\mu_3=\frac{\sum ft(Xt-M)^3}{\sum ft}=\frac{305.036}{30}=10.1679,$$

而後代入鮑萊公式得　$K=\frac{\mu_3}{Sd^3}=\frac{10,1679}{12.167}=0.83$，故

(三)第三法　　$Sk=X_1=0.83$，

各種測定方法所依據之觀點不同，所求得之偏斜度自不能相等，故比較數個數列之偏斜度時，必須應用同一方法，斯讀者不可不注意者也。

第三節　峯度之意義及測定方法

峯度(Kurtosis)亦爲表徵數之一種，然其意義不甚顯著，且亦必於次數數列時用之，故通常應用峯度者甚少，而一般研究統計者亦常忽略之也。

平均數爲數列之代表數値，離散度用以示數列之離散程度，偏斜度用以示次數分配之偏斜度，峯度則表示次數曲線峯形之相對高低程度者。測量次數曲線峯形之相對高低程度係以差誤常態曲線(Normal Curve of error)爲標準，正如以不對稱分配與對稱分配相比較，而得偏斜度相似，曲線之峯形高於常態曲線之峯形者，稱爲高峯態(Leptokurtic)，其意卽示集中於代表數値附近之次數較多：反之曲線之峯形低於常態曲線之峯形者，稱爲低峯態，(Platykurtic)其意卽示集中於代表數値附近之次數較少。故欲完全描述一次數數列尤以配合曲線時，實有研究峯度之必要。

至於測定峯度之方法，亦隨觀點之不同而異，通常有三種方法。

(一)分割數法——

次數愈集中於平均數時，曲線之峯形愈高，而四分位差必愈小：反之，若次數愈不集中於平均數附近時，曲線之峯

註：差誤常態曲線亦稱高斯曲綫(Gaussian Curve)

$$fx=\sqrt{\frac{1}{2\pi}}\,e-\frac{x^2}{2}$$

形愈低，而四分位差必愈大。故峯形之高低恰與四分位差之大小成相反之變化、乃可以四分位差作為測量峯度之標準，而以第九個十分位數與第一個十分位數之差量除之。設以K示峯度，則得公式如次：

$$K=\frac{Q.D.}{D_9-D_1}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(62)$$

當曲線為正峯態(Mesokurtic)時：$Q.D.=0.6744898Sd$，$D_1=1.2815516Sd$ $D_9=1.2815516Sd$，故得

$$K=\frac{0.6744898Sd}{2.5631032Sd}=0.26315$$

從而可知當 $K=0.26315$時，其曲線為正峯態；$K<0.26315$時，其曲線為高峯態；$K>0.26315$ 時，其曲線為低峯態。

(二)動差法

數列之各級單次動差，在次數分配對稱時均為零，故不能以單次動差示曲線峯形之高低，至於各級偶次動差，不管次數

分配之對稱與否，均有其數值，設 μ_{2p} 示各級偶次動差，x 示各數值對算術平均數之差量，y 爲次數函數，則

$$\mu_{2p}=\int_{-\infty}^{\infty}x^{2p}y\,dx$$

在常態曲線時。

$$\mu_{2p}=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}x^{2p}e^{-\frac{x^2}{2\sigma^2}}dx$$

$$=\left[-\frac{\sigma}{\sqrt{2\pi}}x^{2p-1}e^{-\frac{x^2}{2\sigma^2}}\right]_{-\infty}^{+\infty}$$

$$+\frac{(2p-1)\sigma^2}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\overline{x}^{2p-2}e^{-\frac{x^2}{2\sigma^2}}dx$$

$$=0+(2p-1)\sigma^2\int_{-\infty}^{+\infty}\overline{x}^{2p-2}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\overline{x}^2}{2\sigma^2}}dx$$

$$=(2p-1)\sigma^2\mu_{2p-2}\text{。}$$

若令 p=1，則所得爲二級動差，實即標準差之平方數，

自能用以示曲線之峯形，故通常以四級動差爲測量峯形之標準，如上式以p=2，則

$\mu_4 - 3\sigma^2\mu_2 - 3\sigma^4$，

但四級動差含有測量單位，故鮑萊氏主張以σ^4除之，並以K_2示峯度，其公式如下：

$$K_2 = \frac{\mu_4}{\sigma^4}$$

皮爾生氏則用μ_2^2爲除數，並以β_2示峯度，其公式如下：

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

實則兩公式完全一樣，固毫無區別者也。而由上述公式即可求得常態曲線之峯度爲$K_2 = \beta_2 = 3$。從而知峯度K_2或β_2等於3時，即爲正峯態，K_2或β_2大於3時爲高峯態，K_2或β_2小於3時爲低峯態。

至四級動差之積分形式雖如上述，而計算時當用下列公式：

$$\mu_4 = \frac{\sum ft(Xt-M)^4}{\sum ft} \quad\cdots\cdots(65)$$

(Xt示各量數，ft示各次數，M示算術平均數。

一次數曲線之峯形，雖常受四分位差之影響，然其感應不甚靈敏，而第九個十分位數與第一個十分位數常可因一二數値起極大之變化，實際峯態固未必能有同樣之變化也，故分割數法殊嫌粗略，在精密之統計分析，則以使用動差法爲宜。

玆以前節所舉之某班學生年齡分配爲例，以明峯度之求法。

求峯度

(1) X	(2) f	(3) X－M	(3) (X－M)^4	(5) f(X－M)^4
總　計	30			2990,3360
19	2	－3.4	133.6336	267.2672
20	5	－2.4	33.1776	165.8880
21	4	－1.4	3.8416	16.3664
22	7	－0.4	0.0256	0.1792
23	4	0.6	0.1296	0.5184
24	2	1.6	6.5536	13.1972
25	4	2.6	45.6976	182.7904
27	1	4.6	447.7456	447.7456
29	1	6.6	1897.4736	1897.4736

(一)分割數法

$$Q.D.=\frac{Q_3-Q_1}{2}=\frac{24-21}{2}=\frac{3}{2}=1.5$$

$D_9=25 \qquad D_1=20$

$$\therefore \quad K=\frac{1.5}{25-20}=\frac{1.5}{5}=0.3 \quad >0.26315$$

(二)動差法　由上表求得$\sum ft(Xt-M)^4$

$=2990.3360$，故得

$$\mu_4=\frac{\sum f(Xt-M)^4}{\sum ft}=\frac{2990.3360}{30}=96.3445$$

$$\therefore \quad X_2=\beta_2=\frac{\mu_4}{\sigma^4}=\frac{96.3445}{(2.30)^4}=\frac{96.3445}{27.9821}$$

$$=3.4428>3$$

依分割數法所得結果，可謂此分配爲低峯態曲線，但依動差法所得結果，此分配確爲略高於正峯態之曲線（依 $E'=\frac{K_2-3}{8}=0.055$，約高百分之5.5），其所以如此者，實由於分割數法不甚精密之故耳。

第八章　確　度

第一節　一般概念

一、測量之本質　無論自然現象或社會現象，均無絕對精密之測量；猶如世界上無絕對直之直線或絕對純粹之液體。試觀自然現象：欲衡量一物之重量正確至一兩，乃易事也；由於衡器之進步，吾人漸能稱至一分，一厘，甚至一毫，一絲，但終必有一限度。再如角度，肉眼所能辨者至爲有限，而應用六分儀可以測量一弧至十五秒之程度，若格靈威治天文家則能觀察至一秒之百分之一，然亦終必至一限度，踰此再求精確不可能矣。

以上云云，其結果吾人遂謂正確至一厘，或一秒；同理，吾人估計金錢數目，謂正確至一元或一分。

二、自然測量與統計測量，太陽之視差，乃用以決定日球與地球之距離者，視差之測量與統計上之估計頗爲相似。十八世紀之天文家估計視差爲10秒，相當於九千六百萬英里；因觀測方法以及儀器之改進，天文家遂漸同意視差爲 8 秒，但小數仍各不相同。1865 年以來未嘗有人估計較8.8秒倘

爲精確者，而近世紀來之觀察，均同意視差在8.76秒與8.78秒之間，是故日球與地球距離，其差誤已不出四百分之一矣。$\left(\frac{8.78-8.76}{8.76}=\frac{1}{438}\right)$由此可見：1.最先之估計，必有待於校正；2.正確程度與時俱進；3.絕對正確仍未達到，亦永不能達到。統計測量亦然，所有生命預測，國富估計，物價漲落，工資升降均祇能有某種程度之正確性。

三、可得之正確程度：，在自然測量上，吾人常能達到一極高之確度，例如每立方之水，毫無疑問，其重量可正確至百萬分之一；但有時只須知道其正確達到十分之一，即於願已足，例如與地球最近之恆星之距離，約爲三千四百萬至三千七百萬哩是也。在統計上亦往往如是，譬如卅年四川糧食產量爲八千萬石至一萬萬石，重慶市普通勞役每日工貲爲五元至八元等是。上述例子之缺點爲當吾人估計一數，明知並不甚精確，但吾人不得而知其差誤之限度。吾人不易謂「估計數爲25.4元，其差誤約在上下三角之間，絕不致有角月以上之差誤」；但在自然測量上，吾人常能視所用測量儀器之精密，正確至其最小之分度止，如秤之錢，尺

之分是也。

四、實際上需要之確度吾人雖不能獲得精確之數字，但常能估計至實際上需要之正確程度；因通常祇需要某種習慣之確度。譬如耕地面積以畝表示，而無庸正確至若干方丈；目前之物價最小單位爲一分；出生時間祇需記其日期，而不必記淸何時何分；火車時刻表祇標明幾點幾分，而無秒數；大洋輪船行期以時爲準，而非分數；身高僅量至生的米突爲止；百米賽跑，只記至十分之幾秒。同理，在統計估計上，吾人罕需正確至千分之一，甚至無須正確至百分之一。每週工作時間之千分之一，不過三分鐘；每日工資之千分之一，不過五厘錢；重慶人口多一百少一百，國家支出多一萬元少一萬元，吾人全不介意，凡此限度內之確度吾人恆能獲得之。

第二節　差誤之意義

一、相對差誤之界說　確度之反面爲差誤，差誤愈小，則確度愈高。眞實數值減估計數值，謂之絕對差誤；絕對差誤與估計數值之比率謂之相對差誤。若估計數值高於眞實數值，差誤爲負，若估計數值低於眞實數值，差誤爲正。

例如農業工資每月平均眞實數值爲14元，而估計所得爲13元，則差誤爲 $\frac{14-13}{13}=\frac{1}{13}$ 或7.7%；若估計所得爲15元，則差誤爲 $\frac{14-15}{15}=-\frac{1}{15}$ 或−6.6%。

以記號表示：設以u表示測量所得之數（估計數值），而其眞實數值爲u^1，並以e表示相對差誤，則

$$e=\frac{u^1-u}{u}$$

而

$$u_1=u(1+e)$$

$$ue=u^1-u$$

ue 卽絕對差誤。

二、差誤之表示方法　當研究差誤之時，所謂眞實數值無從知道，所能得知者，最多不過其可能之限度。例如某次戰役，吾人估計負傷人數爲 4.5 %，而由可能獲得之材料（如從參戰部隊之名册或傷兵收容所之名册觀之），得知此4.5 %，上下.5 % 必近於事實，卽謂差誤似不致高於

$\frac{.5}{4.5}=\frac{1}{9}$ 或11％，其絕對差誤爲.5%。同時吾人亦能㩀差誤之確定限度。負傷人數之百分比，必在0與100之間，若確知有1%負傷，而92% 均已歸隊，則負傷人數百分比無論如何在1%與8%之間，則絕對差誤最多不過±3.5%（因4.5-1=3.5， 4.5-8=-3.5)，然則相對差誤之最高限度爲

$\frac{\pm 3.5}{4.5}=\pm\frac{7}{9}$ 或±73%. 此限度雖則太大，但究竟比含混言之「負傷4.5%，確度不得而知」正確多多。若再加調查，或許能使差誤範圍更可縮小，並能决定負傷比率必在3.5%與4.5%之間；則應曰「此次戰役負傷4%，正確至1%」，正同「某物重15斤3兩，正確至一兩」也。

有時吾人雖則不得確知差誤限度，但在總數內，往往某幾位數字絕對正確，其餘絕不可靠。例如英國根據1891年人口普查，及1881—91年人口增加率。估計1896年人口爲39,124,496. 吾人可以斷言其最後之兩位或三位數字。無異猜測；而其最前之兩位或三位數字，絕對可靠。其報告應爲39,100,000 或謂 39,124,000 ± 5,000 或其他表示方法；

此種書法總比39,124,496正確多矣。

三、微細數字之省略 一般統計報告常將數字寫至一分一厘，此在政府機關發表數字，或無可評，因政府統計機關，其職責在接受下級機關之報告而彙集表列之，說明搜集之方法以及填報之時期，至於確度之決定，則留待經濟學者或統計學者之探討。但在綜合陳述社會現象時或為科學的估計時，列舉細數（一則因其未必可靠，二則因其於理論上不關重要，於讀者了無意義），非但不必要；且甚非正確之表示。曩昔為避免不正確之數字起見，往往簡述圓整數字（例如謂地球之直徑為8,000哩），若確有理由，能得更正確之數字（例如比較赤道直徑與較小之兩極半徑）則應將較精確之數字寫出，同時標明其確度。

第三節 計算差誤之定律

一、和數之差誤 各量數之和數的差誤等於各量數之差誤之和，但各量數之差誤均須乘以各該量數對其和數之比率。

設測量n個數量，其估計數值為$u_1u_2,\cdots\cdots un$，其差誤為$e_1, e_2,\cdots\cdots en$；又估計數值之和數為u，和數之差誤為e，

則和數之眞實數值爲

$u(l+e)$，而各量之眞實數值爲 $u_1(l+e_1)$，$u_2(l+e_2)$，

……$u_n(l+e_n)$原故

$$u(l+e)=u_1(l+e_1)+u_2(l+e_2)+\cdots\cdots u_n(l+e_n)$$

但 $$u=u_1+u_2+\cdots\cdots+u_n$$

相減，得 $$ue=u_1e_1+u_2e_2+\cdots\cdots+u_ne_n$$

故 $$e=e_1\times\frac{u_1}{u}+e_2\times\frac{u_2}{u}+\cdots\cdots+e_n\frac{u_n}{u}$$

當若干項爲負數時，此式亦可應用。

例如某年勞工階級對於膳食、衣着、居住之支出估計平均約爲25元，5.5元，6.5元，而眞實平均數爲27元，4.5元，6元，則差誤爲 $\frac{2}{25}$，$-\frac{2}{11}$，$-\frac{1}{13}$ 故和數之差誤爲

$(n=25+5.5+6.5=37)$

$$e=\frac{2}{25}\times\frac{25}{37}-\frac{2}{11}\times\frac{5\cdot5}{37}-\frac{1}{13}\times\frac{6\cdot5}{37}$$

$$=.054-.027-.0135$$

$$= + .0135 \text{ 或 } + 1\frac{1}{3}\%$$

此項定律有一重要之用途，當某一總數，其大部分之確度甚高，而其餘部分則不得而知，可應用此定律以求總數之確度。例如參戰部隊之大部分已報告負傷人數爲3,365，並知其差誤最多不過1%，而小部分則未據報告：吾人估計約有100人受傷，並假定其差誤甚大，爲$\frac{2}{3}$或67%，則總數之差誤爲

$$\frac{1}{100} \times \frac{3365}{3465} + \frac{2}{3} \times \frac{100}{3465} = .029 \text{或不過}$$

3%.是則其差誤極近於大部分之差誤。

二、簡單算術平均數之差誤　各量數之算術平均數的差誤等於各量數之差數之和，但各量數之差誤，均須乘以各該量數對其和數之比率。

設 $m_1, m, \cdots\cdots\cdots mn$ 爲 n 個估計數值，其眞實數值爲 $m_1(1+e_1), m_2(1+en), \cdots\cdots mn(1+e_n)$，則估計平均數爲

$$\frac{m_1+m_2++}{n}$$

眞實平均數爲 $\frac{m_1(1+e_1)+m_2(1+e_2)++}{n}$

算術平均數之差誤爲

$$e=\frac{\frac{m_1(1+e_1)+m_2(1+e_2)++}{n}-\frac{m_1+m_2++}{n}}{\frac{m_1+m_2++}{n}}$$

$$=\frac{e_1m_1+e_2m_2++}{m_1+m_2++}$$

$$=e_1\times\frac{m_1}{\sum m}+e_2\times\frac{m_2}{\sum m}+\cdots\cdots\cdots\cdots,$$

其中 $\sum m$ 表示各估計數值之總和。

由公式觀之，可知若各項差誤相差不大並且均爲正或均爲負每項差誤對平均數之差誤，影響均不甚大，且平均數之差誤大致與任何一項差誤相近；倘若在一般情形中，各項差誤有正有負，則平均差誤之誤勢將減少甚多，卽確度增大。故平均法可以削弱差誤。

三、加權算術平均數之差誤加權算術平均數之差誤爲下列二種差誤之和：(1)估計量數之差誤，此種差誤與簡單算術平均數之差誤相仿，(2)權數之差誤，若各量數大抵相等，則此種差誤爲至微細。

令$w_1, w_2, \cdots\cdots w_n$爲n個估計量數$M_1, M_2, \cdots\cdots M_n$之估計權數，再令$w_1(l+d_1), w_2(l_1+d_2)\cdots\cdots$爲眞實權數，$M_1(l+e_1), M_2(l+e_3)\cdots\cdots$爲眞實量數。

設Mw表估計加權算術平均數，而$Mw(1+E)$表其眞實數值，則

$$Mw=\frac{\sum WM}{\sum W}, \sum WM = Mw.\sum W$$

$$Mw(1+E)=\frac{[\sum W(l+d)M(l+e)]}{\sum W(l+d)}$$

$$MwE=\frac{\sum[W(l+d)M(l+e)]}{\sum[W(l+d)]}-\frac{\sum WM}{\sum W}$$

$$=\frac{\sum W.\sum[W(l+d)M(l+e)]-\sum WM.\sum W[l+d)]}{\sum W\cdot\sum[W(l+d)]}$$

$$E.\sum WM.\sum(1+d)=\sum W.\sum WM+\sum W.\sum WMe+\sum W.\sum WMd$$

$$+\sum W.\sum WMed-\sum WM.\sum W-\sum WM\sum Wd$$

$$=\sum W.\sum WMe+\sum W.\sum WM(d+ed)-\sum WM\sum Wd$$

茲設E,e,d均不過.1,則其乘積不過.01，可略之，

$$E.\sum WM.\sum W=\sum W.\sum WMe+\sum W.\sum WMd-\sum WM\sum Wd$$

$$=\sum W.\sum WMe+\sum[Wd(M.\sum W)]-\sum[Wd(\sum WM)]$$

$$=\sum W.\sum WMe+\sum W([M.\sum W-\sum WM)d]$$

$$\therefore E=\frac{\sum WMe}{\sum WM}+\frac{\sum[W(M.\sum W-\sum WM)d]}{\sum WM.\sum W}$$

若W_1M_1寫作m_1，則含量數之差誤e之各項與定律二相同。即$\frac{\sum me}{\sum m}$至d之係數需再加分析：

因$\sum WM=MW.\sum w$,

$$Mt\sum W-\sum MW=\sum WMt-\sum WMw$$

$$=\sum W.(Mt-Mw)=mt\sum W$$

其中mt表示某一估計量數對其加權平均數之差，

$$\therefore E = \frac{\sum WMe}{\sum WM} + \frac{\sum W.\sum Wmd}{\sum WM.\sum W}$$

$$= \sum \frac{WM}{\sum WM}.e + \sum \frac{Wm}{\sum WM}.d$$

故估計量數之差誤（e）各項包含M，M_2等，而權離差之差誤(d)各項僅包含各量數對其加權平均數之差離（m）；倘各量數對其平均數之相差度相當小，則各個離差均相當小，且wm之總和0，因$\sum Wm = \sum WM - Mw\sum W = 0$，故倘使權數的差誤均相等，則平均數內由於權數之差誤爲零；倘使權數之差誤與各個離差(m)不均同號，而且大差數不剛巧均遇着大權數，則勢必微小。

由是觀之，除非差誤之大小，量數之大小，以及權數之大小有一種特別連帶之關係，權數之差誤非但如同量數之差誤，有自行減低之效，而且有其係數，足使差誤互相冲消。倘項數頗多，必須權數差誤甚大，然後可與平均數以足堪重視之差誤。事實上，量數之差誤，其影響既遠在權數差誤之上

，只須權數之估定頗爲合理，而各量數又非甚懸殊，則權數之差誤常可忽視不計。

四、乘積之差誤，乘積之差誤約略等於各因子之差誤之代數和。

設 $f_1, f_2, \cdots\cdots fn$ 爲估計因子，其眞實數値爲 $f_1(1+e_1)$，$f_2(1+e_2)\cdots\cdots$，則乘積之差誤爲

$$E=\frac{f_1(1+e_1).f_2(1+e_2)\cdots\cdots-f.\ f_2\cdots\cdots}{f_1.f_2\cdots\cdots}$$

$$=(1+e_2).(1+e_2)\cdots\cdots-1$$

若略去兩個及兩個以上 e 之乘積，則

$$E=e_1+e_2+\cdots\cdots+e_n$$

各個 e 之爲正抑爲負，其機會相同。若一對一對之 e 不同號，則差誤有互相冲消之勢。倘各個差誤之符號相同，卽均爲正或均爲負，縱然各個差誤甚小，乘積之差誤必大。

例如估計100人每人平均收入25元，而實際上爲105人每人平均收入26元（差誤均爲正），則總收入之差誤，依上述公式計算爲 $\frac{5}{100}+\overline{\frac{1}{25}}=.09$。

倘若實際上爲105人（正差誤）每人平均收入24元（負差誤），則總收入之差誤將爲 $\frac{5}{100}-\frac{1}{25}=.01$。

五、商數之差誤，商數之差誤約略等於分子分母差誤之代數差。

設 u_1, u_2 爲估計之分子分母，其眞實數值爲 $u_1(1+e_1)$，及 $u_2(1+e_2)$，則商數之差誤

$$E=\frac{\frac{u_1(1+e_1)}{u_2(1+e_2)}-\frac{u_1}{u_2}}{\frac{u_1}{u_2}}=\frac{1+e_1}{1+e_2}-1=\frac{e_1-e_2}{1+e_2}$$

$$=(e_1-e_2)(1-e_2+e_2^2-e_2^3+-)$$

若略去兩個及兩個以上 e 之乘積，則

$$E=e_1-e_2$$

倘使分子分母之差誤均爲正數或均爲負數，則勢將互相冲消；倘使二者約略相等，則商數之差誤極小。

吾人可應用定律五以比較某現象兩個時間數列之平均數。

記號之命定與定律二、三同，以m, e, d爲第一時期之估

計數値計數値，m^1, e^1, d^1 爲另一時期之估計數値，則 $m^1{}_1, m^1{}_2$，………之簡單算術平均數對 m_1, m_2，………之簡單算術平均數的比率（商數）之差誤爲

$$E=\sum\left(e^1\frac{m^1}{\sum m^1}\right)-\sum\left(e\frac{m}{\sum m}\right)$$

$$=\left(e^1{}_1\frac{m^1{}_1}{\sum m^1}-e_1\cdot\frac{m_1}{\sum m}\right)+\left(e^1{}_2\cdot\frac{m^1{}_2}{\sum m^1}-e_2\cdot\frac{m_2}{\sum m}\right)++$$

倘若各量數（m）在兩個時期，並無劇烈變動，則 $\frac{m^1}{\sum m^1}$ 與 $\frac{m}{\sum m}$ 必相差甚微。忽略此種微末之差別，則簡單算術平均數之比率的差誤

$$E=\sum\left(\frac{m_1{}^1}{\sum m^1}\ (e_1{}^1-e_1)\right)$$

假如兩時期之估計，環境之條件大致相同，卽差誤之機會相同，測 e_1 與 e 非但同號而已，抑且大致相等。

令 V_1, V_2，…………代替（$e_1{}^1-e_1$），（$e_2{}^1-e_2$）…………則

$$E=\sum\left(V_1\ \frac{m^1}{\sum m^1}\right)$$，其中各個V大致甚小.

同理亦可分析加權平均數之比率的差誤。依原則：權數之差誤遠不及量數差誤之重要，吾人亦可應用上式以資估計加權平均數之比率的差誤之近似數。此公式可以文字表示如下：

六、同種數列之兩個算術平均數（觀察之時期不同）之差誤，約略等於各相應對之項目的差誤之差的和數，各差誤之差均須乘以第二時期各該量數對其總和之比率。

此項定律非常重要，有更爲舉例說明之必要：

設在兩年中，吾人之估計，一部分較另一部分爲可靠，乃可應用下列公式。

	第一年	第二年
估計之人數或權數	w：差誤 d	w^1；差誤d^1
估計之平均收入或量數	m_1；差誤e_1	$m^1{}_1$；差誤$e^1{}_1$
估計之人數，(較不正確者)	rw；r之差誤p	r^1w^1；r^1之差誤p^1
估計之收入，(較不正確者)	m_2；差誤e_2	$m^1{}_2$；差誤$e^1{}_2$

依假設$e_1 \angle e_2$, $e_1{}^1 \angle e^1{}_2$

第一年平均數之差誤爲 E_1 =

$$\frac{\frac{w(1+d).m_1(1+e_1)+r(1+p).w(1+d).m_2(1+e_2)}{w(1+d)+r(1+p)w(1+d)}-\frac{wm_1+rwwm_2}{W+rW}}{\frac{wm_1+rwM_2}{w+rw}}$$

$$=\frac{\frac{m_1(1+e_1)+r(1+p)m_2(1+e_2)}{1+r(1+p)}-\frac{m_1+rm_2}{1+r}}{\frac{m_1+rm_2}{1+r}}$$

$$=\frac{m_1(1+e_1)+r(1+p)\ m_2(1+e_2)}{1+r(1+p)}\cdot\frac{1+r}{m_1+rm_2}-1$$

略去 e 與 p 之乘積

$$=\frac{m_1+m_1e_1+rm_2+rm_2re_2m_2+lrm_1+rm_1e_1+r^2m_2+r^2pm_2+r^2e_2m_2}{\{1+r(1+p)\}\ (m_1+rm_2)}$$

$$-\frac{m_1+rm_1+rpm_1+rm_2+r^2m_2+r^2pm_2}{\{1+r(1+p)\}\ (m_1+rm_2)}$$

$$=\frac{m_1e_1+rpm_2+re_2+re(M)_2rm_1e_1+r^2e_2m_2-rpm_1}{\{1+r(1+p)\}\ (m_1+rm^2)}$$

$$=\frac{e_1m_1\{1+r(1+p)\}+e_2rm_2\{1+r(1+p)\}+rp(m_2-m_1)\underbrace{-rpe_1m_1-r^2pe_2m_2}_{\text{略　去}}}{\{1+r(1+p)\}\ (m_1+rm_2)}$$

$$=e_1\frac{m_1}{m_1+rm_2}+e_2\frac{rm_2}{m_1+rm_2}+p\frac{r}{1+r(1+p)}\cdot\frac{m_2-m_1}{m_1+rm_2}$$

略去$\frac{r}{1+r(1+p)}$中之 rp，則

$$\sum_1=e_1\frac{m_1}{m_1+m_2}+e_2\frac{rm_2}{m_1+rm_2}+p\frac{r}{1+r}\frac{m_2-m_1}{m_2+rm_2}$$

由是觀之，e_2與 p（較不正確部分之差誤）均與 r（較不正確部分之權數與比較正確部分之權數的比率）相乘，p 並與m_2-m_1相乘，而m_2-m_1常甚微小，同時依假定e_1甚為微小。

為說明簡便起見，假定在兩時期較不正確之部分對全體之比率（r）不變，同時假定兩部分平均收入比率$\left(\frac{m_1}{m_1+rm_2}\right)$亦不變，乃得兩年算術平均數之比率的的差誤

$$E=(e_1'-e_1)\frac{m_1}{m_1+m_2}+(e_2'-e_2)\frac{rm_2}{m_1+rm_2}+(p'-p)\frac{r}{1+r}\cdot\frac{m_2-m_1}{m_1+rm_2}$$

例：蘇格蘭農業工人平均工資變動

	1867		1892	
	人數	平均收入(鎊)	人數	平均收入(鎊)
已婚農夫 估計數值	1,000 (w)	36 (m_1)	1,200 (w^1)	49 (m_1')
已婚農夫 假設眞實數值	1,010	35	1,220	48
其他僱傭 估計數值	200 (rw)	34 (m_2)	240 (r^1w^1)	41.25 (m'_2)
其他僱傭 假設眞實數值	220	37	240	47

$$r = \frac{200}{1,000} = \frac{1}{5}, \quad r^1 = \frac{240}{1,200} = \frac{1}{5}$$

$$d = \frac{1,010 - 1,000}{1,000} = \frac{1}{100}, \quad d^1 = \frac{1,220 - 1,200}{1,200} = \frac{1}{60}$$

$$e_1 = \frac{35 - 36}{36} = -\frac{1}{36} \quad e_1' = \frac{48 - 49}{49} = -\frac{1}{49}$$

$$e_2 = \frac{37 - 34}{34} = \frac{3}{34} \quad e_1' = \frac{47 - 41.25}{41.25} = \frac{23}{165}$$

$$p = \left(\frac{220}{1010} - \frac{200}{1000}\right) \div \frac{200}{1000} = \frac{9}{101}$$

$$p^{1}=\left(\frac{240}{1220}-\frac{240}{1200}\right)\div\frac{240}{1200}=-\frac{1}{61}$$

上表假定已婚農夫之收入估計過高，而其他僱傭則過低。蓋因事實上其他僱傭之供宿及他項供應之實際收入不易估定。且兩種農人之人數比例亦不能正確知之。

代入前列公式：

(1)由於已婚農夫收入估計之差誤

$$E(e_1)\doteq(e_1'-e)\frac{m_1}{m_1+rm_2}$$

$$=\left(-\frac{1}{49}+\frac{1}{36}\right)\times\frac{36}{36+\frac{1}{5}\times 34}=.0062$$

(2)由於其他僱傭收入估計之差誤

$$E(e_2)=(e'_2-e_2)\frac{rm_2}{m_1+rm_2}$$

$$=\left(\frac{23}{165}-\frac{3}{34}\right)\times\frac{\frac{1}{5}\times 34}{36+\frac{1}{5}\times 34}=.0081$$

(3)由於兩種農人人數比例估計之差誤

$$F(p)=(p^1-p)\frac{r}{1+r}\cdot\frac{m_2-m_1}{m_1+rm_2}$$

$$=\left(-\frac{1}{61}-\frac{9}{101}\right)\times\frac{\frac{1}{5}}{1+\frac{1}{5}}\times\frac{34-36}{36+\frac{1}{5}\times 34}=.0008$$

第三種差誤，即由於權數之差誤，至爲做小。

兩年平均收入比率之差誤爲

$$E=E(e_1)+E(e_2)+E(p)$$

$$=.0062+.0081+.0008$$

$$=.0151$$

試由實際數字觀之：

$$\text{估計比率}=\frac{1892\text{ 之 估計平均收入}}{1867\text{ 之 估計平均收入}}$$

$$=\frac{\dfrac{1200\times 49+240\times 41.25}{1200+240}}{\dfrac{1000\times 36+200\times 34}{1000+200}}$$

$$=\frac{47.708}{35.667}=1.3376$$

$$\text{假定眞實比率}=\frac{\text{1892 之 假定眞實平均收入}}{\text{1857 之 假定眞實平均收入}}$$

$$=\frac{\dfrac{1220\times 48+240\times 47}{1220+240}}{\dfrac{1010\times 35+220\times 37}{1010+220}}$$

$$=\frac{47.836}{35.358}=1.3529$$

$$\text{比率之差誤}=\frac{1.3529-1.3376}{1.3376}$$

$$=\frac{.0153}{1.3376}=.0114$$

,0151 與.0114 之差別乃由於計算方法之略去較不重要項目故也。

所當注意者實際數量比率之差誤與增加率之差誤不同，在上例中

估計增加率＝1.338－1＝.338　　或 33.8%

假設眞實增加率＝1.353－1＝.353　　或 35.3%

$$\text{則增加率之差誤}=\frac{35.3-33.8}{33.8}=\frac{1.5}{33.8}=0.45$$

第四節　偏誤與非偏誤

一、偏誤與非偏誤之意義。無論平均數或比率之差誤，亟應區別其爲偏誤抑爲非偏誤。其分別可舉例說明如次：設派遣若干人員赴各地調查工業狀況，其目的在證明工資高，工作情況合乎衞生，勞資協調等，則彼等或將僅考察經營最善之廠家，並僅取精工及長工之工資，因而所得各地平均工資不免太高，此種差誤，謂之偏誤，其方向相同，均有增高平均數之勢，總平均數之差誤，約等於每處之差誤。反之，若調查人員胸無成見，調查至爲公平，因而各視環境之不同，若干地點太高，其他則太低，此種差誤，謂之非偏誤，過高與過低之機會相等，調查之地點愈多則總平均數之差誤愈小。試觀下表，可見兩種差誤對平均數之影響：

	事實	有成見之調查	無成見之調查
甲區平均工資	24元	25元	24元
乙區平均工資	23	25	25
丙區平均工資	26	27	25
丁區平均工資	27	28	28
戊區平均工資	28	30	27
總平均	25.6	27	25.8
差誤	—	5.2%	1%

設乘一脚踏車，以行程自動測量器測量公路之距離，則將發現各個里程標石間之距離，未必相等，但其不足一公里或大於一公里之機會相等，行程愈遠，則總差誤愈小。是爲非差誤。倘使測量器不甚準確，往往行一公里之遙，紀錄950公尺，則其差誤爲偏誤。是故偏誤常能測定而剔除之，但非偏誤則只能任其自行冲消。尋常偏誤大抵由於測度工具之不精而產生；非偏誤則發生於測量技術上之過失。

二、偏誤與非偏誤之比較重要性　當調查人口時，婦女所報告之年齡，常較眞實年齡爲輕，足使平均年齡減低；而一般人則常報告圓整數目如五十歲、六十五歲之類，但就平均而論，並無甚影響。前者爲偏誤，後者爲非偏誤。總之、在單獨估計中，非偏誤不若偏誤之重要；但在兩個同性質之估計的比率中，則偏誤常能自行冲消。

令 $e_1, e_2 \cdots$ 表各個量數之非偏誤，$d_1, d_2 \cdots$ 表其偏誤，則由定律二可知平均數之差誤爲

$$E = \sum\left(e \cdot \frac{m}{\sum m}\right) + \sum\left(d \cdot \frac{m}{\sum m}\right)$$

第一項之各個非偏誤 e，有正有負，足以互相冲消。設以 a 表示 $e_1, e_2, \cdots\cdots\cdots\cdots e_n$ 之平均數，則上式第一項之總和

其一次近似值爲 $\dfrac{2_2}{3\sqrt{n}}$

例如有百次測量，每次非偏誤約爲 $\frac{1}{10}$，則平均數中由於非偏誤之差誤不過

$$\frac{2a}{3\sqrt{n}} = \frac{2}{3} \times \frac{1}{10} \div \sqrt{100} = \frac{1}{150}$$

但偏誤則無互相冲消之效，若每次高$\frac{1}{10}$，則平均數亦將高$\frac{1}{10}$。

三、偏誤之重要。倘目的在求確度，則當權宜輕重，注意關係之重要，而忽略非偏誤。斤斤於測量技術改良，減少非偏誤；而由測量工具所生之偏誤，反置之不問，良屬無謂。當調查之時，常有偏誤存在其間，苟吾人不知其有無，自無法可想；設或知之，則縱用最粗疏之方法以校正之，較諸置若罔聞，尤爲正確。因當用無成見之方法校正有成見之差誤（偏誤），則使偏誤變爲非偏誤，由是平均數所包含之項數愈多，其差誤愈小。例如一萬農人調查之結果，得每月平均工資十三元，並有理由足以證明各個非偏誤約爲一元，則非偏誤對平均工資之影響不過

$$\frac{2a}{3\sqrt{n}} = \frac{2}{3} \cdot \frac{1\text{元}}{\sqrt{000}} = \frac{1\text{元}}{150} = ,007\text{元}$$

但此平均數不免爲一妄誕之平均數；因農人之實際所得

例如農忙時期之加工，計件工作之工資。膳宿之供給等，未必能有精確之計算。若捨此而不計算，則平均數之差誤甚至有 5 元之多。倘對各個農人之上述實際所得，與以估計（化偏誤爲非偏誤），縱然每次估計可有 5 元之差誤（非偏誤），而無其他偏誤之存在，則平均數之差誤不過

$$\frac{2a}{3\sqrt{n}}=\frac{2}{3}\ \frac{5\text{元}}{\sqrt{10000}}=\frac{1\text{元}}{30}=.033\text{元}$$

是則全部差誤不過4分(.007+.033)而非5元。在考定一般報告之數字的確度時，此項原則必須銘諸座右，注意有無偏誤之存在。

四、比率之確度　若研究比率之確度，情形完全不同。依定律五比率之差誤約等分子、分母差誤之差、設以 q_r，q 表分子、分母之偏誤，e^1，表其非偏誤，則比率之差誤：

$$E=(d^1-d)\quad+(e^1-e)$$

其中非偏誤（e^1-e）約與 e^1 或 e 相等，因若以 E 表 e 或 e^1 之概差，其（s^1-e）之概e，爲 E $\sqrt{2}$ 。倘e或e^1

不致大於$\frac{2}{5}$，則(e^1-e)似不致大於$\frac{4}{5}\sqrt{\quad 2}$即$\frac{3}{5}$。

但偏誤(d^1-d')則不然，若分子分母之偏誤同號(均為正，或均為負)則比率之偏誤必較分子分母之偏誤為小。祗須分子分母之估計方法確實相同，卽調查同樣對象所問之問題相同，關於問題之內容取舍一致，則分子分母之偏誤，必無二致。再以前舉農業工資為例，設兩次調查，除調查每月平均工資之外，均有忽略任何其他收入之顯然錯誤，則兩次比率之差誤之由於於此項偏誤者僅視工資率對其他收入之比例有無變動耳，而在短期間內，此項比例之變動，必甚微末。若兩次均僅調查夏季工資以代表全年平均工資，則比率之差誤，亦僅視夏季工資對平均工資之比例如何耳。故變動遲緩之現象，兩期估計(方法相同)之比率的差誤，當較各期之差誤為小；何者？因非偏誤雖則略有增加，而最重要之偏誤則大形減低矣。此處吾人無容知道有無偏誤之存在，偏誤將自行消滅。若知其確有偏誤，並有方法足以作更正確之估計，自非無益；但若在一期內剔除其偏誤，而在另一期內則置之不問，以求兩期之比率，則不免鑄成大錯。為比較起見，使後

期之估計較前期爲正確，非徒無益，且常困難重重。因非偏誤雖可略微減低（設E與E_1爲兩期之非偏誤，則比率之非偏誤爲$\sqrt{E^2+E_1^2}$。此項偏誤視 E 或 E_1之減小而隨之減小），但比較重要之偏誤反而有增無減。

五、連續發表之統計其確度宜一致　政府統計人員或其他人員編製統計，不免兩難：一方面爲使每年之報告，漸入正確起見，必須力求方法之改善，必須注視情况之變化，而調整其編製及表列之方法；另一方面爲使歷年數字便於比較起見，則宜絕對保守，明知過去之錯誤而不加改革，但十分小心，不再有新錯誤或新遺漏。此種困難，通常之避免方法如次：當起初引用改良之方法時，表列時除用新之分類外，另爲一舊式表式，以資與以前比較。當方法之變動所引起之差別，已經明瞭後，則往時之數字，乃可以調整之，使其確度與後期相若。　例如英國商業部1898以後，編製出口總值統計，包含隨貨出售與外人之船舶，而以前則否，故表列兩列兩期數字如下：

	1899	1898
英貨出口總值（不計隨貨出售與外人之船舶）……	225,465,000鎊	233,359,000鎊
英貨與英屬商品由英轉運出口總值……	65,020,000	60,655,000
總　計……	320,485,000	294,014,000
隨貨出售與外人之船舶總值	9,195,000	未　詳
新　總　計……	329,680,000	

在搜集資料及表列資料時，輕於些小之更動，常爲統計上錯誤之原因，此不可不注意也。

第五節　結論

本章討論之結果：可概述如次。有兩種方法足以增高確度：平均法可以減小非偏誤，而比率法可以減小偏誤。權數之差誤，遠不及其他差誤之重要。差誤雖不易計算，但可以所含各項之差誤表明之；雖不能得到絕對之正確，而可用上述方法以消減差誤，並用數學方法以測量消減之程度。其在同種數列或同法估計之數列的加權算術平均數的比率原則，有差誤必可大大消滅，此所以指數有其特殊之功用也。

第九章 比例

第一節 比例之意義

研習經濟學或社會科學書常用及各種比例(Ratios)或比率(Rates)，如死亡率(Death rates)，出生率(Birfh rates)，結婚率(marriage rates)，稅率(Tax rates)，勞工轉移率(Labor trunover rates)，股票轉移率(Stock turnover rates)等等均是。蓋由大量觀察而來之數字為絕對數值，而僅此絕對數值有時難以明瞭二者間之關係。例如甲市之人口為2,000,000人，某年共死亡20,000人；乙市之人口為500,000人，某年共死亡15,000人。吾人若謂甲市死亡人數較乙市死亡人數多5,000人，殊屬毫無意義，蓋甲乙二市人口並非相等，其人口多者之死亡人數當然應比人口少者為多，故必須將死亡人數與總人口對照後始能互相比較，亦即求得死亡人數對總人口之比例後始能比較，故有時欲比較各種現象間之關係時，非求出其比例不可。

比例為一分數，係用以比較某現象與其結合現象者，而某現象為分子，其結合現象為分母。例如某公司民國三十年之銷貨對平均存貨之比例為3.8，其意卽為民國三十年之銷

貨爲分子，平均存貨爲分母，所得分數之數值爲 3.8 • 惟比例有用百分數表示者，卽將原分數乘以 100 示之，此亦不可注意者也。

比率通常爲以時間爲單位之比例〔註〕例如死亡率爲每年每一千人中之死亡人數，利率爲每年每一元之利息，此均以一年爲單位之比例也。

由於此種關係，吾人可注意絕對增加，相對增加，絕對增加率，與相對增加率之區別。設若某地民國二十年之人口爲1,000人，民國三十年增至1,500人，則絕對增加爲 500 人，而相對增加爲增加人數對民國二十年人口之比例卽50%，但此兩種數字均未顧及時間之因素。絕對增加率爲每年之增加人數卽50人，相對增加率爲每年增加人數對民國二十年人口之比例，卽每年增加 5 %，此兩種數字均係以時間爲單位者。比率通常爲相對比率。且在表示人口之增加率時，常非如上例所舉之以某一年爲基年之簡單比例，而爲幾何平均之比例。

註有時亦有例外，例如外匯率並非以時間爲單位，而指某一時期之外匯價格•

第二節　比例之種類

統計學上所用之比例，按性質之不同，可分五種：(一)變之比率(Rates of Change)，(二)分配之比例(Disatributlon satos)·(三)類別間之比例(Inter-class ratios)，(四)不同事物之比例(Hybrid ratios)，(五)特種統計係數。茲分述於下。

(一)變動之比率　變動之比率爲解釋社會生活與經濟生活變動之重要工具。設吾人欲比較國民財富之增加，或價格之變動，均須用變動比率。下表所示雞蛋每月價格之環比，即變動比率，係示各月蛋價對上月蛋價之變動，若爲100無變動，超過100之數，爲增高之百分數，不及100之數，減低之百分數。

雞蛋每月價格之環比

	1917	1918	1919	1920	1921
一 月	99.0	106.9	104.0	104.7	94.0
二 月	95.0	106.7	84.4	87.8	81.2
三 月	94.4	81.8	68.5	81.9	58.9
四 月	70.0	77.2	103.6	83.3	69.9
五 月	115.8	93.4	107.3	90.4	99.0
六 月	108.7	96.1	104.9	98.9	96.0
七 月	91.0	103.0	95.3	99.2	113.4
八 月	105.3	112.1	106.8	109.0	120.9
九 月	111.4	105.8	104.3	110.5	114.3
十 月	122.6	114.3	109.9	113.3	112.5
十一月	105.3	113.5	120.8	113.6	129.2
十二月	109.9	119.5	114.6	114.2	115.6

（資料來源： 林和成著實用工商統計254頁）

(二)分配之比例　分配之比例係示一部分對其全體之數字關係，常用百分數示之。下表第三列所示卽分配之比例，係示各種族所佔全人口之百分比，如白種佔全人口之89.7%，黑種佔全人口之9.9%等等。

(三)類別間之比例　類別間之比例係示某一部分與另一部分之關係，或某一種類與另一種類之關係。下表第六

列所示性比例卽類別間之比例，係示男性對女性之比例，例如總人口之性比例爲104.0，卽示總人口中男性對女性之比例爲104比100。

（四）不同事物之比例　不同事物之比例係用以比較不同事物者，故分子分母非同一種類之事物，此與上述數種比例所不同者也。所有「每人若干」之比例關係均屬此類，如每人所得，每人費消量，每人財富等是。同樣，鉄路上每噸或每人之收費亦不同事物之比例，其餘可依此類推。因此種比例係不同事物之較，故必須用平均數表示，而不能如上述比例之以百分數表示之。

（五）特種統計係數　特種統計係數爲特種比例，係示同一數列中某項目對另一項目或兩數列有關項目間之關係。前者如以前所述離散係數爲同一數列之標準差對算術平均數之比例，用以示該數列之離散程度者；後者如各種相關係數(Coefficients of Correlation)係示兩數列相互間之關係者。

美國1920年各種族人口比較表

(1) 種族別	(2) 人口數	(3) 各族佔全人口之百分比	男	女	性比例
總人口	105,710,623	100.0	53,900,431	51,810,189	104.0
白種	94,820,915	89.7	48,430,655	46,390,260	104,4
黑種	10,463,131	9.9	5,209,436	5,253,695	99,2
印地安人	244.437	0.2	125,068	119,369	104,8
中國人	61,639	1.0	53,891	7,748	695,5
日本人	111,010	0.1	72,707	38,303	189.8
菲律賓人	5,603	(a)	5,232	371	1,410.2
北印度人	2,507	(a)	2,409	98	(b)
高麗人	1,224	(a)	923	301	306.6
夏威夷人	110	(a)	75	35	(b)
其他種族	44	(a)	35	9	(b)

(a)小於0.1%，故未列入。

(b)女性人口少於100，故未計算性比例。

（資料來源；　美國1920年第十四次普查，第三編人口，第十一頁）

第三節　比例之計算與應用

比例之計算甚爲容易，祇須根據上述各種比例之意義，

將比較之事物一爲分子一爲分母卽可，惟爲使比例數字顯著起見，常乘以10之乘方數，最通用者爲乘以100，卽百分數。他如乘1,000，乘100,000亦常用之；乘1,000者如死亡率，其意卽每1,000人中死亡之人數；乘100,000者如特種死亡率，其意卽每100,000人中因某種病因死亡之人數。若比例數欲求準確至一位小數，則必須計算至兩位，而後略去一位。此爲計算與應用時一般注意之點，其尤應遵守之原則如下：

(一)比例之分母小於100，不宜用百分數，蓋易致錯誤之結果也。

(二)求比例之平均數時須注意加權。設有學生以其獎學金額之50%及家庭所給費用之10%購書，則除非兩種收入相等。否則購書費用決非兩種收入之30%。同樣，各省人口增加率之平均數，決非全國人口之增加率，除非各省之人口數相若。故求比例之平均數時必須注意加權，其權數爲各原比例之分母。

(三)分配之比例僅當總數同質時始能適用。在求分配比例之前必須考慮總數與其各部分是否同一性質，例如美國農業普查所得農場牲畜之總數卽非完全同質，若用以求馬、騾

、驢、牛、山羊、豬、綿羊、家禽等佔總牲畜之百分數，即不甚適用。

(四)增加率或減少率僅當增加或減少數量與基數有顯著關係時始為重要。例如一地人口之增加係基於自然之原因，而文盲之增加或減少並非基於自然之原因，若文盲之變動率用全人口為基數，則比較此種變動率時必須加以極大之注意，蓋正確之比較固當以文盲人數與文盲減少人數相比較也。

(五)比例之良否賴於分子分母之適當選擇。通常一種事實必須與其所由出之事實比較，例如比例之分子若為戰爭之傷亡人數，則分母必須為參加戰爭之人數；若分子為結婚人數，則分母應為可婚人數，雖然該項原則常被混亂，但需要遵守亦至屬明顯者。

(六)比例之比較僅當其分子與分母均屬可以比較時始有意義。比較兩種或數種比例時常易發生錯誤，縱使此兩種比例甚為一致或甚相似，而經深密之考慮後，常亦有不能作嚴格之比較者，故若比較比例或百分數時，必須注意其分子與分母能否互相比較。此一原則於邏輯比較上甚為重要，且常不易被遵守，蓋資料常係不同時間或不同機關所搜集，自不易相比較也。但多數不能遵守此原則之原因係基於疏忽不能比較分子之或分母，下節所述經濟現象中之數項特種比例即

其例也。

第四節　特種比例

在表示商業情況之各種比例中，信用率 (Credit ratios) 卽爲有特殊用處之一種比率，故有稱之爲信用測度計者。華爾氏(Alexander wale) 曾對此種商業比率作特殊之研究，此種比率對於信用之决定有重大之幫助，尤以『靜態比率』(Static ratios)與『動態比率』(velocity ratios) 之分別甚爲重要，前者測定某一期之財政狀況，後者係銷貨與資產負債中某一特定項目之比較。『靜態比率』分四種：(1)流動率(Current ratic) 卽流動資產對流動負債之比率；(2)應收款項對商品之比率；(3)債款對純值之比率；(4)純值對固定資產之比率。『動態比率』係銷貨對(1)應收款項，(2)商品，(3)純值，與(4)固定資產之比率。習慣上常以流動率爲標準而甚少注意其他各種信用之比率，但華爾氏以爲較合邏輯之分析，當按各種商業定其標準，各種比例均可似流動率之一樣應用也。

多種商業或經濟上所用之比例，均可加以適常之討論，但於此所欲討論者爲易致錯誤之數種。

其一爲勞工移轉率。表示勞工移轉之方法甚多，但最

普通之一種爲一年中辭退人數除以工資賬之平均人數，即如下式：

$$\frac{\text{一年中辭退人數}}{\text{工資賬之平均人數}}\times 100=\text{勞工移轉率}$$

然有以更換人數代替辭退人數爲分子者，亦有以平均在工人數代替工資賬之平均人數爲分母者，吾人自無法一一列舉各種可能之方法；但若一用辭退人數，一用更換人數，則比較時必得錯誤之結果。

其二爲利率，(Interest rates)，即一年所付利息對本金之百分比例。對於借款時期之長短並無關係，例如短期借款雖不滿一年，其利率仍表示一年所付之利息，同樣長期借款，雖至三年五年，其利率亦表示一年所付之利息，如此始可比較也。

其三爲稅率。(Tax rates)。所得稅率常指對於所得應課之百分數。財產稅率在美國指財產估價每元之千分數，但應加區別者爲名義稅率(Nominal tax rate)與估計之眞實稅率(Estimated true tax rate)，後者爲稅額除以財產之市價。例如財產稅率爲千分之23.5，若有一汽車估價1,000元，則應納

稅款爲23.5元，而若該輛汽車實值僅 500 元，則眞實稅率爲千分之47.0，若該輛汽車實值 2,000 元，則眞實稅率爲千分之 11.75 ，故比較兩城之名義稅率實毫無價值，蓋分母性質不同不能比較也。例如美國1919年芝加哥(Chicago)之名義稅率爲59.29，尼伯拉斯加州之峨瑪哈 (Omaha, Nebraska) 爲100.73,威斯康州加瑪的孫(Madison, Wisconsin)爲13.50；但其售價之眞實稅率爲14.77，18.53與13.50。

上述爲商業與經濟上習見之比例，在生命統計之範圍內，比例之應用亦甚爲廣泛，其最主要者爲死亡率s，出生率，結婚率，離婚率，與疾病率。但須注意者此諸比率常以求得之比例乘以 1,000 。

死亡數爲某一時期(常爲一年)內死亡人數對於該時期中之生存人數比例，因是常用一年內死亡人數與該年七月一日估計人數之比例示之。但吾人必須注意普通死亡率，特種死亡率，與標準死亡率之區別。若某城民國二十五年之死亡人數爲 20,000，該年七月一日之估計人口數爲2,000,000，則普通死亡率爲10.0卽1,000人中死亡10人。若死亡之人數中

，有10.000人爲死於腸熱病者，則即可據以計算腸熱病死亡率，特種死亡率即分別就各種死亡原因，性別，年齡組成或其他特種分類所計算之死亡率，在分析比較上，特種死亡率較普通死亡率有價值。標準死亡率係將普通死亡率就性別與年齡之組成加以調整後之比率。

另有所謂嬰兒死亡率者，係根據特種方法計算，通常以某年嬰兒死亡人數對嬰兒出生人數之比例示之。

出生率較爲簡單，即某一時期（常爲一年）內出生人數對於該時期中點生存人數之比例。自然增加率爲出生率與死亡率之差數，若無遷入或徙出，則出生超過死亡之數，應與普查時人口增加之數相同。

結婚率普通指一年內結婚人數對於該年七月一號估計人數之比例而言，但用此法比較殊不合理，因各地人口之年齡組合不同，所能結婚之人數亦有差異，以不同條件之事項相比，顯然不合比較之原則。故合理之結婚率應以可婚人數爲比例之基礎，即爲一年內結婚人數除以該年七月一號之可婚

人數所得之比例數。

離婚率之計算方法與結婚率同樣，普通係以一年內離婚人數除以該年七月一日之估計人數，但合理之離婚率應爲一年內離婚人數對該年七月一號已結婚人數之比例。

疾病率係一年內某種病件對該年七月一號估計人數之比例。此種比率對於流行病學之研究，極有用處。

在生命統計範圍內，比例之應用固廣，而其錯誤亦最易發生。例如普通死亡率（亦稱粗死亡率或總死亡率）係以全體人口爲比例之基礎，包括兩性，一切年齡，國別，一切職業，及一切死亡原因；而吾人知年齡，國別，兩性諸因素對於死亡率能發生影響，故不管人口之組成僅比較普通死亡率實易生極大之錯誤。卽就特殊死亡率言之，若不細加分析，而妄爲比較，亦易生錯誤之結果：例如按職業之種類求男子之死亡率，結果，銀行行長之死亡率，定比賣報男孩之死亡率爲高．此決非職業之不同，使銀行行長之死亡率高於賣報男孩之死亡率，實因年齡不同之影響。總之，吾人使用比例時必須細加分析並注意上節所述之應用原則也。

第十章 指數

第一節 指數之意義

指數者，用簡單之數字表示現象在空間或時間之變化者也。就狹義之指數定義言之，其所示之現象爲綜合複雜不能直接測量者，所得指數即大陸派所稱之綜合指數，（國人常譯 Aggregatve index 爲綜合指數實欠妥）；就廣義之指數定義言之，不管現象之單獨或綜合，簡單或複雜，凡表示現象在空間或時間之變化者均爲指數，即包括大陸派所稱之個體指數與綜合指數。但個體指數實即上章所述之比例數，故本章所述限於綜合指數，而一般所謂指數，亦指綜合指數而言也。

例如物價，其變化甚爲複雜，世間物品不止一種，其變化之趨勢亦不一律，或上漲，或下落，或相差甚大，或變動甚微，苟無簡單之數字以示一般物價之變化，則異地異時之物價將無由比較。更就生產而論，煤鉄之生產以噸計，米麥之生產以擔計，布綢之生產以疋計，發電機之生產以馬力計，併此性質迥異單位不同之產量而欲比較其在不同時間或空

間所生之變化，非先將複雜之數量化成簡單之數字不可，此簡單之數字卽指數也。

第二節　指數之種類

指數之種類甚多，若依比較之基性及所用資料之不同，可分下列數種：

(一)依比較之基性分

(1)時間性指數　所謂時間性指數，卽以時間爲基性，表示現象在時間上之變化者也。例如以民國二十五年之平均物價爲基數，求各時間之指數卽是。惟以基期（Base Pe-riod）之不同，又可分爲定基，環比及鎖比等三種。

(2)地域性指數　以地域爲基性表示現象在空間之變化者謂之地域性指數，例如以南京之平均工資爲基數求各地之工資指數卽是。

(二)依所用資料之不同分　指數如依所用資料之不同分，種類甚多，茲舉其最著者述之如次：

(1)物價指數　在現代經濟制度之下，吾人之經濟行爲莫不受物價之支配，整個經濟機構莫不受物價之影響，故經濟學者必研求物價現象也，而物價現象錯綜複雜難以捉摸

，故有物價指數之編製，物價指數即用以示物價現象之變化者也。物價指數又以所用物價之不同，可分為下列三種：

甲、批發物價指數——批發物價為商人大宗買賣物品時所議定之價格，以此價格編成指數，即批發物價指數(Index numbers of wholesale prices)。其效用乃在測定市場上一般物價之漲跌，及商業循環之眞相者也。

乙、輸出入物價指數——輸出入物價，為國際貿易商人所宣佈之價格，或進出口商在貨物通過海關時所報，或海關所估之價格，以此價格編製之指數，即輸出入物價指數 (Index numbers of export and import prices)。編製此種指數之目的，乃在表示國際市場上，一國人民實際所付價格之變動，並可藉以推測國際貿易盛衰之趨勢。

丙、零售物價指數——零售物價，係消費者直接購買物品所付之價格，以此價格編成指數，即所謂零售物價指數(Index numbers of retail prices)是也。編製此種指數之目的，乃在測定人民生活費用度之高低及貨幣購買力之強弱，研究社會經濟者，俱以此種指數為立論之根據。

(2)生活費指數　生活費乃維持人類生存之費用，如衣

食住及燃料等費用是。生活費指數 (Index numbers of the cost of living) 乃卽測量生活費之變遷者也。

(3)工資指數　工資指數(Index numbers of Wages)乃以工人工作之報酬，編成之指數，其效用在測定工人工作酬勞之變遷，並可藉以解決勞資糾紛問題。

(4)外匯指數　外匯指數(Inbex numbers of foreign exchange)者測量一國貨幣對外匯率變動而編製之指數也。其在用金國與用金國之間，可不必有基期而以平價爲基數，惟在用銀國與用金國之間，則須有基期方可比較。

(5)國外貿易指數　國外貿易指數(Index numbers of foreign trade)乃以一國對外貿易額編成，用以表示對外貿易之狀況者也。此種指數宜用對外貿易額編成，否則卽不足以確示對外貿易之實際趨勢，蓋貿易值之增加或減少，不能認爲貿易額亦增加或減少也。

(6)證券指數　證券指數(Index numbers of securities)乃以公債券，公司股票及其他有價證券之價格或買賣數量等編製之指數。其效用，乃在窺測債券股票等發行業務之盛衰

及投資利益之大小者也。

指數除上述六種外，又有所謂生產量指數及消費量指數，用以測定供求之是否相應；成本指數用以預測營業之盈虧，據爲定價之標準；種類之多，殊不勝枚舉。

但指數雖依基性之不同，可分時間性指數與地域性指數兩種，而一般使用者均爲時間性指數，故以下各節以時間性指數爲討論之對象，而地域性指數不難由此想像得知也。

第三節　指數之計算方法

一、總和法(Aggregative methcd)

總和法係將各統計數列之屬於同一時期者分別總和之，而後分別求各期總和對基期總和之比例，此比例乘以 100 即爲指數。設

基期各數值之總和爲：$p_o' + p_o'' + p_o''' + \cdots\cdots + p_o^{n}$

$$=\sum p_o,$$

計算期各數值之總和爲：$p_1' + p_2'' + p_1''' + \cdots\cdots + p_1^{(n)}$

$$=\sum p_1$$

則總和式指數之公式為

$$A = \frac{\sum p_1}{\sum p_0} \times 100$$

例重慶市民國二十九年一月至六月八種物品之零售價格如下表：

重慶市民國二十九年一月至六月八種物品之零售價格表

（單位二元）

物　品	一月	二月	三月	四月	五月	六月
米	1.480	1.000	2.050	2.540	2.820	3.600
麵　粉	0.180	0.190	0.240	0.300	0.300	0.400
猪　肉	0.600	0.600	0.600	0.600	0.720	0.800
菜　油	0.850	0.870	0.800	0.800	1.000	1.120
豆　腐	0.020	0.020	0.020	0.020	0.020	0.030
白土布	0.170	0.520	0.090	0.900	0.300	0.920
嵐　炭	6.750	7.700	7.860	7.800	8.700	12.050
肥　皂	0.400	0.450	0.570	0.580	0.8[illegible]0	0.900
共　計	10.750	11.950	12.830	13.600	1[illegible].0[illegible]0	19.820

先分別計算各月八種物價之總和，計一月為 \$10,750 二月為 11.900 三月為 12.830 四月為13.600五月為 15.400 六

月為 19.820若以一月為基期，則即可根據公式得各月之指數如下：

$$一月之指數=\frac{10.750}{10.750}\times100=100.00$$

$$二月之指數=\frac{11.950}{10.750}\times100=111.16$$

$$三月之指數=\frac{12.830}{10.750}\times100=119.35$$

$$四月之指數=\frac{13.600}{10.750}\times100-126.51$$

$$五月之指數=\frac{15.400}{10.820}\times100=143.26$$

$$六月之指數=\frac{19.820}{19.750}\times100=184.57$$

用總和法所求得之指數稱總和式指數•

二、比率平均法

比率平均法者，先求兩時期相應對數值之比率，而後再依平均法以求指數之法也。依所取平均數之不同，而有下述五種：

(1)算術平均　求各比率之算術平均數，而後乘以 100，所得指數稱爲算術平均式指數。

設計算期之各項數值爲p_1'，p_1''，$p_1'''$$p_1(n)$；基期之各項數值爲$p_o'$，$p_o''$，$p_o'''$，........$p_o{}^{n}$　；兩時期相應對數值之比例爲$\frac{p_1'}{p_o'}$，$\frac{p_1''}{q_o''}$，$\frac{p_1'''}{p_o'''}$，........$\frac{p_1(n)}{p_o n}$；項數爲　n；則算術平均式指數(I_M)之公式爲

$$I_M = \frac{I}{n}\left(\frac{p_1'}{p_o'} + \frac{p_1''}{p_o''} + \frac{p_1'''}{p_o'''} + \cdots\cdots\cdots + \frac{p_1(u)}{p_o n}\right) \times 100$$

$$= \frac{I}{n}\Sigma\frac{p_1}{p_o} \times 100$$

例如前表之資料，求算數平均式指數，可先求各項物品各計算期價格對基期價格之比例。設以一月爲基期，求各項物品其各月價格對一月價格之比例如下表：

重慶市民國二十九年一月至六月八種物品之零售價格比率表 （一月爲基期）

物品	一月	二月	三月	四月	五月	六月
米	1.0000	1.0871	1.3851	1.7162	1.0054	2.4324
麵粉	1.0000	1.0556	1.3333	1.6697	1.6667	2.2222
猪肉	1.0000	1.0000	1.0000	1.0000	1.2000	1.3333
菜油	1.0000	1.0235	1.9412	0.9412	1.2471	1.3176
豆腐	1.0000	1.0000	1,0000	1.0000	1.0000	1.6000
白土布	1.0000	1.1064	1.4681	1.9140	1.9574	1.9574
嵐炭	1.0000	1.1407	1.1644	1.1644	1.2889	1.7852
肥皂	1.0000	1.1250	1.4250	1.4500	2.1500	2.2500
共計	8.0000	8.4323	9.7171	10.8534	12.4155	14.7981
指數	100.00	106.65	121.46	135.67	155.49	184.98

而後按照公式，求各月各項比例之算術平均數乘以 100，卽得各月之算術平均數，如上表指數欄所示。

（2）幾何平均　求各比率之幾何平均數，而後乘以 100，所得指數稱爲幾何平均式指數，設計算期與基期相應對數値之比率爲 $\frac{p_1'}{p_0'}$，$\frac{p''}{p_0''}$，$\frac{p_1'''}{p'''}$ ………………… $\frac{p_0(n)}{p_1(n)}$；項數爲 n

；則幾何平均式指數(IG)之公式爲

$$I_G = u\sqrt{\frac{p'}{p_o'}, \frac{p''}{q_o''}, \frac{p_1'''}{p_o'''} \cdots\cdots\cdots\cdots \frac{p_1u}{p_ou}} \times$$

實際計算時，必須利用對數，即採用下式

$$\log I_G = \frac{I}{N}\left(\log\frac{p_1'}{p_o'} + \log\frac{p_1''}{p_o''} + \log\frac{p_1'''}{p_o'''} + \cdots\cdots\cdots\cdots + \log\right.$$

$$\frac{p_1(n)}{p_o n} + \quad 2 \quad = \frac{I}{N}\sum\log\frac{p_1}{p_o} + 2 \cdots\cdots$$

仍用以上所列之資料，求幾何平均式指數，除求各項物品價格之比率如上列所列外，並求各項比率之對數，如下表：

重慶市民國二十九年一月至六月八種物品零售價格比率之對數　（一月爲基期）

物品	一月	二月	三月	四月	五月	六月
米	0,00000	0,03887	0.11148	0.23457	0.27998	0.36966
麵粉	0.00000	0.02350	0.12493	0.22186	0.22180	0.34678
猪肉	0,00000	0.00000	0,00000	0.00000	0.07918	0.12493
菜油	0,00000	0 01009	1.97368	1.97368	0.09578	0.11979
豆腐	0.00000	0.00000	1,00000	0.00000	0.00000	0.17609
白土布	0,00000	0.04392	0.16676	0.28215	0.29167	0.29168
嵐炭	0.00000	0.05717	0. 6610	0.06610	0.11022	0.25169
肥皂	0,00000	0.05115	0.15381	0.16137	0,33244	0.35218
共計	0.00000	0.21970	1.62670	1.93973	1 41123	2.03280
平均	0.00.00	0.02746	0.2033	0,24247	0.17640	0.25410
指數	100.00	106.53	119.77	131.06	150.11	179.52

而後按照公式，求各對數之算術平均數加 2 即求得指數之對數，復由指數之對數求得指數，如上表指數欄所示。

（3）倒數平均　求各比率之倒數平均數，而後乘以 100，所得指數稱爲倒數平均式指數。設計算期與基期相應對數值之比率爲 $\frac{p_1'}{d_0'}$，$\frac{p_1''}{p_0''}$，$\frac{p'''}{p_0'''}$……………$\frac{p\,u}{p_0u}$；項數爲u；則倒數平均數之公式爲

$$\frac{1}{I_H}=\frac{1}{N}\left(\frac{1}{\frac{p_1'}{p_0'}}+\frac{1}{\frac{p''}{p_0''}}+\frac{1}{\frac{p'''}{p_0'''}}+\cdots\cdots+\frac{1}{\frac{pu}{p_0u}}\right)\times 100$$

$$=\frac{1}{N}\left(\frac{p_0'}{p_1'}+\frac{p_0''}{p_1''}+\frac{p_0'''}{p_1'''}+\cdots\cdots+\frac{p_0u}{p_1u}\right)\times 100$$

$$=\frac{1}{N}\sum\frac{p_0}{p_1}\times 100$$

倒數平均數指數係各比率之倒數平均數，故僅於比率有倒數關係時用之，一般甚少採用也。

(4)中位數　求各比率之中位數，而後乘以 100，所得指數稱為中位數指數。本法之計算，無需資料之全部，僅就其一部分卽足以決定，故主張使用比率之中位數為物價指數者相當衆多，概括彼等之理由有三：甲、計算簡單，乙、計算物價指數時，所用以平均之商品並非全體，僅擇其一部份以概一切，故所得平均數與眞正確值之誤差，以中位數平均法為最小；丙、幾何平均數與算術平均數均受極端項之影響甚大，而中位數則無此缺點．但中位數旣僅使用中間部分之若干資料，則此部分之資料發生變化，卽足影響中位數

所求得之指數，是以此指數實爲不安定之數值，且用一部分資料，亦不足以眞確表示全體，故該法雖爲愛奇渥斯(Edgeworth)等所竭力提倡，而仍未爲世人所樂用。

(5)衆數　求各比率之衆數，而後乘以 100，所得指數稱爲衆數式指數，因衆數本身爲一不安定數值，且甚難確定，故尙無學者積極主張之．

第四節　指數之加權

構成指數上所使用種種數值之本身，其重要性各有不同，卽以物價指數而言，所使用之商品在經濟上之地位大不相同，因此其價格之變動，對於社會經濟之影響亦相差甚巨，如米價之變動當較茶價之變動爲重要，若編製指數時予米價與茶價以同樣地位，則所得指數必不甚眞確，故計算指數時，各數值有加權之必要也．

指數之加權，視所採計算方法而不同，玆分別述之如次。

一、總和法之加權

總和法係將統計數列之屬於同一時期者分別總和之，而後求各期總和對基期總和之比率，其所根據者爲各實際數值

，因是加權所取之權數，應爲各該數值之次數，例如用總和法編製物價指數係根據各項物品之實際價格，則加權時之權數，應爲各該物品之數量，或爲生產量，或爲消費量，或爲交易量，蓋如是求得乘積之總和爲其總金額，就此總金額求其指數，自爲合理之方法。用總和法加權所得之指數爲加權總和式指數。

設計算期之各項數值爲 p_1'，p_1''，p_1'''............$p_1^{\,n}$；基期之各項數值爲 p_0'，p_0''，p_0'''............$p_0^{(u)}$；權數爲 q'，q''，q'''..............$q^{\,n}$；則加權總和式指數（Iw.A.）之公式如次：

$$\text{Iw.A.}=\frac{p'q'+p''q''+p'''q'''+\cdots\cdots p^u\ q^u}{p_0'q'+p_0''q''+p_0'''q'''+\cdots\cdots p_0^u\ q^u}\times 100$$

$$=\frac{\sum pq}{\sum p_0q}\times 100$$

生活費指數常爲加權總和式指數，因生活費指數係示生活費用之昇降者，生活費用爲各項生活必需品之價格與其消費量之乘積，故若以消費量爲權數，用加權總和法所編成之

指數，爲最足以表示生活費用之昇降者。

例如社會部統計處所編生活費指數均爲加權綜合法，茲就其所編重慶市職業工人生活費指數爲例，說明其計算方法如次：

重慶市職業工人生活費指數所取之物品及其權數見下表所列(1)(3)兩欄：設有二十九年一月及二十六年上半年各項物品之價格如(4)(6)兩欄所列，則計算以二十六年上半年爲基期之二十九年一月之指數如下：

用加權綜合法計算重慶市二十九年一月工人生活指數

物品名稱 (1)	單位 (2)	q (3)	二十六年上半年		二十九年一月	
			q_o (4)	p_oq (5)	p (6)	p q (7)
米	市斗	6.8	1.208	8.214	1.480	10.064
麵粉	市斤	1.4	0.113	0.158	0.180	0.252
切麵	市斤	1.3	0.136	0.178	0.200	0.260
蠶豆	市斗	0.7	2.107	1.517	0.200	0.840
猪肉	市斤	3.9	0.257	1.002	0.600	2.340
牛肉	市斤	1.1	0.175	0.193	0.300	0.330
猪油	市斤	1.8	0.365	0.157	0.960	1.920

桊油	市斤	0.9	0.300	0.270	0.850	0.765
蔴油	市斤	0.5	0.300	0.150	1.200	0.600
鷄蛋	個	7.5	0.017	0.128	0.060	0.450
醬油	市斤	0.8	0.148	0.118	0.350	0.280
鹽	市斤	2.1	0.130	0.273	0.200	0.420
紅糖	市斤	0.5	0.080	0.040	0.600	0.300
黃豆芽	市斤	2.4	0.024	0.058	0.100	0.240
豆腐	塊	21.4	0.012	0.259	0.920	0.432
榨桊	市斤	1.7	0.080	0.141	0.256	0.435
粉條	市斤	0.9	0.200	0.218	0.300	0.270
水	挑	12.7	0.050	0.635	0.300	3.810
白土布	市尺	1.2	0.100	0.120	0.470	0.564
藍土布	市尺	1.9	0.072	0.137	0.600	1.140
陰丹士林布	市尺	0.6	0.158	0.195	0.820	0.492
草鞋	雙	6.0	0.020	0.120	0.200	1.200
布鞋	雙	0.4	0.400	0.160	1.800	0.720
碎煤	挑	0.8	0.400	0.320	3.000	2.400
劈柴	捆	3.3	0.163	0.538	6.810	2.673
桊油	市斤	0.5	0.300	0.150	0.850	0.425
土鹽燭	市斤	0.4	0.300	0.120	1.200	0.480
火柴	小匣	2.8	0.007	0.020	0.050	0.140
肥皂	塊	2.4	0.102	0.245	0.400	0.960
毛巾	條	0.3	0.232	0.085	1.000	0.300

煙	絲	市兩	4.5	0.150	0.675	0.300	1.350
乾	酒	市斤	0.5	0.167	0.084	0.400	0.200
草	紙	百張	0.9	0.100	0.090	0.200	0.180
平	房	市方	1.7	1.646	2.798	5.925	10.073
總	計	丈			19.966		47.005

由$\sum p_0q=19.966$(第5欄總計)，$\sum p_1q=47.005$(第7欄總計)代入公式卽得

$$Iw.A.=\frac{47,005}{19,966}\times100=235.43$$

二、比率平均法之加權

比率平均法係先求兩時期相應對數値之比率，而後求諸比率之平均數以求得指數，其所根據者爲各數値之比率，因是所取權數應爲無單位之比値或爲同一單位之數。例如物價指數用比率平均法編製時，若以各物品之生產量或貿易量或消費量爲權數顯屬錯誤，而必須用生產總値或貿易總値或消費總値等爲權數，若此總金額有同一公約數，則約之爲簡單數値作爲權數亦可。但比率平均法有種種，除中位數與衆數兩種不用加權外，其餘三種平均之加權公式如次：

(1)加權算術平均　加權算術平均係各比率乘其相應對

之權數而後總加之，再除以權數總和之謂，所得指數稱加權算術平均數指數。

設計算期之各項數値爲p_1'，q_1''……p_1^n；基期之各項數值爲p_0'，p_0''；p'''……p_0^n 兩時期相應對數値之比率爲

$\frac{p_1'}{p_0'}$，$\frac{p_1''}{p_0''}$，$\frac{p_1'''}{p_0'''}$…………$\frac{p_1 u}{p_0 u}$；p'之權數爲 W'.，p"之權數爲W"，依此類推；則加權算術平均數指數(Iw. M.)之公式如次：

$$I_{w.M.} = \frac{\frac{p_1'}{p_0'}\cdot w' + \frac{p_1''}{p_0''}\cdot w'' + \cdots\cdots + }{w' + w' + \cdots\cdots}$$

$$= \frac{\sum p_1 w}{\sum w} \times 100$$

仍用以前之資料，以各物品之每家平均消費値爲權數【註】則用加權算術平均法計算廿九年二月份指數（以一月爲基期）如下表：

註：係根據社會部所編公務員及工人生活費指數中採用之權數，決定各物品之每家平均消費量而後乘以一月份之物價所得之乘積。

用加權算術平均法計算重慶市二十九年二月零售物價指數

(1) 物品	(2) W（每家平均消費值）	(3) $\frac{p_1}{p_1}$（29年2月對1月之比價）	(4) $\frac{p}{p_0}\cdot W$
米	10.212	1.0811	11.040
麵粉	0.270	1.0556	0.285
豬肉	1.280	1.0000	1.280
菜油	0.935	1.0235	0.957
豆腐	0.446	1.0000	0.446
白土布	2.054	1.1064	2.273
嵐炭	5.400	1.1407	6.160
肥皂	1.080	1.12q0	1.215
共計	21.677		23.656
指數			109.13

指數109.13卽係$\sum\frac{p_1}{p_0}W$　與$\sum w$之數值代入公式求得。

（2）加權幾何平均　加權幾何平均係各比率以其相應對之權數爲方冪，而後連乘之，再開以權數總和次方之謂。所得指數稱加權幾何平均數指數。

設計算期與基期相應對數值之比率爲$\frac{p_1'}{p_0'}$，$\frac{p_1''}{p_0''}$，$\frac{p_1'''}{p_0'''}$

……………$\frac{p_1(n)}{p_0(n)}$；p'之權數爲w'，p''之權數爲w''，依此類

推；則加權幾何平均數指數（Iw.G.）之公式如次：

$$Iw.G.=\sqrt[\sum w]{\left(\frac{p_1'}{p_0'}\right)^{w'}\left(\frac{p_1''}{p_0'}\right)^{w''}\cdots\cdots}$$

$$\times 100$$

$$=\sqrt[\sum w]{\Pi\left(\frac{p_1}{p_0}\right)^{w}}\times 100$$

實際計算時，必須利用對數，即採用下列公式：

$$\log Iw.G.=\frac{w'\log\frac{p_1'}{p_0'}+w''\log\frac{p_1''}{p_0''}+\cdots\cdots}{w'+w''+\cdots\cdots\cdots\cdots}$$

$$+12$$

$$=\frac{\sum\left(w.\log\frac{p_1}{p_0}\right)}{\sum w}+2$$

例如上表之資料，用加權幾何平均法計算重慶市二十九

年二月之零售物價指數(以一月爲基期)，可先求各比率之對數，而後如下表之計算：

用加權幾何平均法計算重慶市二十九年二月零售物價指數

(1) 物品	(2) w	(3) $\log\frac{p_1}{p_0}$	(4) $w.\log\frac{p_1}{p_0}$
米	10.212	0.0338659	0.3458386
麵紛	0.270	0.0234994	0.0063448
肉猪	1.286	0.0000000	0.0009000
菜油	0.935	0.0100878	0.0094321
豆腐	0.446	0.0000000	0.0000000
白土布	2.054	0.0439122	0.0901957
嵐炭	5.400	0.0571714	0.3087256
皂肥	1.080	0.0511525	0.0552448
共計	21.677		0.8157815
log I			2.0376335

由上表$\sum w=21.677$，$\sum\left(w.\log\frac{p_1}{p_0}\right)=0.8157815$，代入公式(74)，求得$\log I=2.0376335$。求logI之反對數，卽得指數爲109.05。

(3)加權倒數平均　加權倒數平均係將各比率之倒數以其相應對之權數乘之，而後求其總和除以權數總和，再求其倒數之謂。所得指數稱加權倒數平均式指數。

設計算期與基期相應對數值之比率爲 $\frac{p_1'}{p_o'}$，$\frac{p_1''}{p_o''}$，$\frac{p'''}{p_o'''}$ ………… $\frac{p_1u}{p_ou}$；p' 之權數爲w'，p"之權數爲 w "，依此類推；則加權倒數平均式指數(Iw.H)之公式如次：

$$\frac{1}{\mathrm{I\,wH.}} \quad \frac{\dfrac{w'}{\dfrac{p_o'}{p_1'}} + \dfrac{w''}{\dfrac{p_1''}{p_o''}} + \cdots\cdots\cdots\cdots}{w' + w''}$$

$$= \frac{\sum w \dfrac{1}{\dfrac{p}{p}}}{\sum w} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(7\ )$$

或 $\mathrm{\cdot I\,w\,.H.} = \dfrac{\sum W}{\sum W \dfrac{1}{\dfrac{p_1}{p_o}}} \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(76)$

加權倒數平均式指數，僅於比率有倒數關係且有加權必

要時始用之，一般甚少使用也。

由上所述，可知綜和法之加權須用數量，而比率平均法須爲無單位之數值或同一單位之數。以數量爲權數，問題比較簡單。或取基期之數量（Laspeyres法）或取計算期之數量(Paasche法)或取兩期之平均數量（Edgeworth與Marsheli)甚至取若干期之平均數量均可；若取無單位之數值或同一單位之數爲權數，則有一加研究之必要。以物價指數言之，通常之交易值，消會值或生產值爲權數，而此諸價值等於數量之物價，其大小常隨物價而變動，以是指數之結果往往隨所用權數爲基期價值或計算期價值而有一種偏誤、所謂加權之偏誤（ Biasmthe weighting)是也。基期某種物價高（假定數量無劇烈之變動）其價值必較鉅，而所得之價比小；某種物價低，其價值必較小，而所得之價比大。以較大之權數乘小的價比，或以較小之權乘大的價比，無形中均足使指數偏低，反之，如以計算期之價值爲權數，則結果適得其反，因計算期某種物價高，其價值必較鉅，而所得之價小亦大，某種物價低，其價值必較小，而所得之價比亦小，以較小之權數乘小的價比，或以較大之權數乘大的價比，無形中均足使

指數高　。故權數爲基期價值，則指數偏低；爲計算期價值，則指數高　。除基期價值與計算期價值外，亦可以基期數量值計算期物價・或計算期數量乘基期物價所得之價值爲權數。茲以 p_0與p_0 分別代表基期之物價與數量；　q與p　分別代表計算期之物價與數量，則以價值爲權數之四種基本方法如下：

(1)以某期數量乘基期價格所得之價值(p_0p_0)爲權數　。

(2)以計算期數量乘基期價格所得之價值(p_0 pr)爲權數。

(3)以基期數量乘計算期價格所得之價值($p_1\ p_0$)爲權數。

(4)以計算期數量乘加算期價格所得之價值($p_1\ p_1$)爲權數。

據費雪教授之報告，按第二法加權，指數亦有偏低之弊；按第三法加權，指數亦有偏高之弊。吾人究應何去何從？當視所用之公式而定。蓋一方面爲加權的偏誤，一方面尙有公式的偏誤（Type bias），如公式本身有偏高之弊，再用三、四兩法加權，則指數將有兩重偏誤，愈形偏高；如公式

本身有偏低之弊再用一、二兩法加權，則指將有兩重偏誤、愈形偏低。比例平均法均可用上述四法加權，共得二十種加權之公式，除中位數及衆數爲任性之公式外，茲將其餘十二式偏誤之有無與方向，及其爲一重偏誤或二重偏誤如下：

二重偏高	加權算術平均第三式 加權算術平均第四式
一重偏高	加權幾何平均第三式 加權幾何平均第四式
無偏誤	加權算術平均第一式 加權算術平均第二式 加權倒數平均第三式 加權倒數平均第四式
一重偏落	加權幾何平均第一式 加權幾何平均第二式
二重偏低	加權幾數平均第一式 加權幾數平均第二式

所應注意者，簡單幾何平均式原無偏誤，而加權結果則生偏誤；簡單算術平均式原有偏高之弊，但用第一第二兩法加權則無偏誤；簡單倒數平均式原有偏低之弊，但用第三第四兩法加權則無偏誤，可見權數選取之重要也。惟所謂無偏

誤非絕對準確之謂，不過公式之偏誤可與相反的加權之偏誤相消，但是否能相消無餘，未可定也。

第五節　指數之基期

依第三節所述，指數之計算必須以某一時期爲基準，此一時期稱爲指數之基期。指數之基期依其決定之方法可分爲二種；一曰固定基期(Fixed base) 二曰連環基期。固定基期係以某一時期之數値爲基準，此基準在未修改以前恆固定不變。連環基期則其作爲基準之時期隨時變換，如上年之物價作爲本年之基價，本年之物價作爲明年之基價是，如此各以上一年爲基期所得之指數稱爲環比指數，環比指數之各環相乘，而得鎖比指數，環比指數與鎖比指數則爲根據連環基期而得之指數。

連環基期之利在於比較時期之距離較短，其變化之測定較容易而準確，例如所比較者爲前後鄰接兩時期之物價，則環比指數確較固定基期指數爲明確，惟吾人研究一現象當作長期之觀察。而各期環比指數因小數四捨五入之關係，不免有微小之差誤，此項差誤愈積愈大，經過長時期，每每與事實相去甚遠且若基期時常變換，計算必較繁複，故一般指數

均樂用固定基期也。

然則基期應爲一星期乎？一月乎？一年乎？抑五年十年之平均乎？更應以何年，何月，何時期爲基期？均有一加討論之必要。

對於第一問題之解决，則與所取之平均法有關。如用算術平均數，則基期宜長，蓋若基期中之數値如有一二項極漲或極跌，則所得比値勢必異常之低或異常之高，而極高之比値大有左右算術平均之能力。卽有極低之數値，亦不能與之抵銷（比値之上昇無限，其下落則以零爲極限），故欲消除此種變態之影響，不可不用較長之時期；反之若用幾何平均數，則與基期之選擇無關，若用中位數而數値項數又甚多者，則基期之影響亦甚微。惟測量經濟現象之指數，不管其平均方法爲何，一年常較一月爲佳，而數年之平均又較一年爲佳，蓋一年之平均數爲基數可免除季節變動之影響，若更數年（常爲五年）至十年之平均數値爲基數，則更可免除循環變動之影響。

至於指數基期究應選擇何年？何月？何時期？概括言之，乃愈近指數時期愈好，蓋指數基期與指數時期相距愈遠，

則比值之趨勢愈形散漫，而變動升降之迹，愈不明顯。但尤應注意者，卽所選基期中不應有特殊事變，一般選擇基期之條件如次：

(一)社會經濟界之穩定——基期宜定於一般經濟及社會狀況穩定之時，不宜在經濟及社會狀况劇變之時。

(二)比值離散度之大小——以物價指數言之，宜採用物價平穩之時爲基期，惟物價之平穩與否，可以該時期比值離散度之大小測定之，蓋物價平穩之時，各比價之離散度及離散係數必小；非然者必大。

(三)基數供給之詳確——基數爲指數計算之標準，若其數字不能詳盡準確，則指數亦必不能確實，故基期必爲一時期其能供給十分詳盡準確之數字

第六節　指數公式

指數公式多至百數十，何者適用？何者不適用？甚難斷論，法國統計學家馬耳克(March)以爲指數公式在形式上所應滿足之標準凡五：

(一)敏感性——所據以編製指數之各數值中，有一值發生變化，其指數卽現出其變化之影響者，

(二)單位共通性——指數之數值不依數量單位之取法而異其結果者。

(三)比例性——各個物品價値依同一比例變化時，指數亦發生同一變化者。

(四)循環性——雖變更其基期，但任意二時期指數値之相對的變化比例仍不變更者。

(五)單純性——計算方法簡單者。

美國統計學家費暄氏更進而作實際之研究，以同種貨品，同種物價，以一九一三年爲基期，一九一八年爲計算期，用一百三十餘種不同之公式計算指數，結果最大之指數爲244,最小之指數爲178,兩者相差達66.據渠之研究，測驗公式良否之方法有二；(一)時間互換測驗法（Time reversal test），(二)因子互換測驗法（Eactor reversal rest），茲即以物價指數爲例說明如次：

(一)時間互換測驗法　物價指數係以測量物價之相對的變動爲目的，故必須選取某一時期之物價爲標準以與另一時期之物價相比較，永得其變動之程度。依理甲乙兩時期互爲基期所得之指數應互爲倒數，亦即兩者相乘之積應等於一；

依此原則而測驗公式是否優良卽所謂時間互換測驗，諸公式中適合時間測驗者僅簡單幾何平均，簡單中位數，簡單衆數及簡單總和式，餘均不合。玆就合與不合兩類公式中選擇其一舉例說明如次：

(1)簡單算術平均不合於時間互換測驗之例。　由表四十八，二十六年上半年米每市斗 1.208 元，陰丹士林布每市尺 0.158 元；二十九年一月米每市斗 1.480 元陰丹士林布每市尺 0.820 元。以二十六年上半年爲基期，二十九年一月之基數爲 320.3，以二十九年一月爲基期，二十六年上半年之基數爲50.2。兩者之乘積爲 1.620 與 1 之差達 62.0％；此62.0％爲向前指數與向後指數之聯合差誤，吾人雖無法察知差誤之由於向前指數者若干，由於向後指數者若干，然簡單算術平均式指數無論爲向前或向後，均有偏高之弊，固爲吾人所熟知也。而此種偏高卽可由時間互換測驗之。

(2)簡單幾何平均合於時間互換測驗之例，　設仍取上例用簡單幾何平均法求定指數，得二十九年一月之指數爲212.2，二十六年上半年之指數如39.7。兩者之乘積爲 1.003 較 1 僅差0.1％，此蓋由小數四捨五入之故。

(二)因子互換測驗法　此種測驗適用於加權指數，蓋加權指數公式中p代表物價，q代表交易量，若以p q 互換，則其求得之指數非物價指數而爲物量指數，因子互換測驗者卽物價指數與物量指數相乘之積是否與兩時期交易值之比率相等之測驗也。若商品祇有一種則其物價指數與其物量指數相乘之積等於計算期交之易值與基期之交易值之比也　若商品不止一種，則根據上述諸公式求得之指數均不能滿足此條件，費暄教授之所謂「理想公式」則與此條件相合，其公式如下：

$$I = \sqrt{\frac{\sum(P_1q_0)}{\sum(P_0q_0)} \times \frac{\sum(P q_1)}{\sum(P_0q_1)}} \cdots\cdots\cdots\cdots (77)$$

若其中p與p互換，卽得物量指數如下：

$$\sqrt{\frac{\sum(p P_0)}{\sum(p_0P_0)} \times \frac{\sum(q P)}{\sum(q_0P_1)}}$$

兩者相乘，則得

$$\frac{\sum p q_1}{\sum p_0q_0}$$

此卽計算期之交易值與基期之交易值之比也。

除上述兩種測驗方法外，尚有一種測驗法謂之循環測驗(Circular test)，適用於兩個以上時期之比較　設以第一年爲基年，計算第二第三兩年之指數，而後以第二年之指數除第三年之指數，視其結果是否等於以第二年爲基期年者計算第三年指數，若相等卽爲合於循環測驗。但關於此項測驗，各家意見略有不同：潘蓀氏謂此乃時間互換測驗之合理的推論，蓋兩個時期旣可互換而無矛盾，則三個時期自當亦無矛盾，但適合此測驗必先有一條件，加權數須爲常數是也。費喧氏則以爲此項測驗可以不必，蓋祇有權數爲常數之時方可滿足此測驗，而各時代物品之重要性不同，權數不能固定不變也　通常各公式均不能適合此測驗，卽費喧氏所最稱道之理想公式亦不適合也。

再就時間互換測驗言之，時間互換測驗固爲試驗公式優劣之方法，但優良之公式未必盡合於時間互換測驗，合於時間互換測驗者亦未必卽爲優良之公式。如加權總和公式雖不合於時間互換測驗，然仍不失爲優良之公式；反之簡單中位數及簡單衆數式指數雖合於時間互換測驗，但依然無補於其任性之弱點，此吾人所應注意者也。