湯鴻翥編

教育統計學

大華書局出版

編輯旨趣

這一套教科書，是供給簡易師範學校，鄉村師範學校，縣立師範學校，及師範講習科等而編輯的，師範課程中，應有的科目，大致具備。

本局同人鑒於各書肆所出的師範教科書，大都僅合於高中師範，而沒有注意到上列的各項學校，即有一二種出版的，也已成爲陳跡，不適於現時的應用，所以本局特邀請各專家，編輯此項教科書，以應時需。

現今教育界，又注重師範教育了，可是爲國民教育之母的上列各項學校，所用的教科書，或勉強應用高中師範課本，或以別種教育參考書雜湊之，實爲教育界的矛盾現象，因此本局特編輯此項書籍，以貢獻於師範教育界。

本局立意編輯這套教科書，爲時已久，務使各書的內容，材豐而辭約，文淺而義深；俾教者有發揮之餘地，學者得應用之機會。但粗率之處，或所難免，還望教育界明達之士，進而教之，則本局幸甚，教育界幸甚。

大華書局編輯部謹啓

中華民國二十二年五月

教育統計學

編輯大意

一　本書取材,除照部頒高中師範科教育測驗與統計課程暫行標準中的統計部分所規定之教材外,另加數章,以資補充.凡師範學校,高中師範科,簡易師範,或其他各種師範學校,均可採作教本之用.其他如小學教師,研究教育測驗者,以及各種教育機關辦事人員,均可用爲參考書.

一　至另行加入的數章,除參照前東大附中擬定的教育統計學綱要中所定的標準外,再參加編者數年來從事於是項實際工作及講授之經驗,融合而成.舉凡普通實用的統計方法,可謂已儘量搜羅.

一　各學校如將測驗與統計分作兩學程教授,或學程名雖爲一,而分作上下兩學期教授的,則採用此書作教本,極爲相宜.因本書之內容一貫,能自成系統.如係混合教授的,則測驗部分,同時可採用本局出版拙編之教育測驗一書,因該書亦遵照部頒之標準而編輯.故此兩書能分能合,均有相關之處.

28220

一 合教時,如因時間怱促的關係,則可將第一第二第九三章略而不教.簡易師範,或初中程度的師範採作教本時,則除第一第三第九三章略去外,再可將第七章中的27,28兩節,及第八章中34節之後半節,與35,36兩節,亦可省去不教.

一 第三章表列法,第四章圖示法,亦可放在最後教授.因其方法與其他各章,並無何種連續性,故可隨教者之便利而定以先後.

一 本書內容,編輯時,文字力求淺顯.如遇意義上,術理上,略覺複雜處,均多舉例證,反復說明,俾便教授及閱讀者能簡易明曉.至書中所舉之例,均與教育測驗有聯帶的關係.蓋取其時能與測驗聯絡,而相互印證也.

一 書中所列各種方法之特性,及如何應用,均詳爲敍述.俾學者有所比較,而易於瞭解.

一 教育統計的學程,不能專習理論,或敍述方法之來源及步驟.最緊要者,在於實地之練習.據編者自來教學之經驗,如專重講述,而不實地練習,則正如耳邊過風,卽學亦等於不學.深望讀者,除本書練習章後所附之習題外,多多採入相類的,加以演習,以資熟練.

一 本書除參考Rugg, Thurstone, Odell, Otis, Holzinger, Yule, McCall, Kelley, Rietz, Fisher, Jones, Pearson, 諸家之著作

外,在中文方面,復採取於下列諸先生之著作者甚多.故特申明,以表謝忱而便參考.

1	朱君毅	教育統計學	商務版
2	周調陽	教育統計學	中華版
3	薛鴻志	教育統計學大綱	北高師出版部
4	艾　偉	教育統計學講義	中大講義股
5	王仲武	統計學原理及應用	商務版
6	俞子夷	測驗統計法概要	商務版
7	趙文銳	統計學原理	商務版
8	邰爽秋	教育圖示法	

一　編者學識譾陋,且編輯時間甚爲傯促,遺漏及不週之處,在所難免,深望讀者鑒諒,並有以教正!

民國二十二年六月,編者謹識.

教育統計學

目　　錄

第一章　緒論

第二章　教育事實之徵集

第三章　表列法

第四章 圖示法

第五章 全體數的統計

第六章 集中數量

第七章 差異數量

第八章　相關數量

第九章　可靠性

附錄

表　　次

圖次

教育統計學

第一章 緒論

1 統計學之史略

統計二字,在英文爲 Statistics. 究其字源之所出,則一般學者,多云由拉丁字 Status 而來.考拉丁字 Status 之原義,則係表示國家之情勢,及其所處之實際地位.是則統計事實之由來,自部落時代進而至國家組織,即已具其雛形矣.蓋一國之君主,都要知道他所管區域內的各種情形,例如要知道他國裏的財富,以決定國庫之收入,就不得不從事於土地或戶口之調查;欲知道他的戰鬪力之大小,又不得不從事於武人或兵士之調查.是以統計之應用,及其事實之發現,實至早且古.言其實例,則從歷史上看,歷代皆有,似難歷歷枚舉.惟最顯著而發現最早者,則首推紀元前 3050 國王之建築金字塔.當埃王之建金字塔時,會將全口及財富之現狀,調查一次,藉以分擔建築的工程.前 1605 年,羅馬氏第二(Romeses II)復作埃及土地之

調查,按人民之數額,重行分配,使各得其所,此皆就歐洲上古時期統計事實之最早而尤最著者而言.論及我國,則首推紀元前2200年夏禹王治水時之禹貢九州𥳑之編製.當時禹王本其治水之經歷,區分全國爲九州,並詳列其土質、河山、田賦、產物、運輸及其貢賦之次第等等.雖然照現代的眼光看起來,其組織未見如何的完善,然在當時,頗具有統計之作用矣.故禹王實爲中國統計學之鼻祖,在西國亦皆傳誦之.厥後在商湯則有井田之制度,即土地戶籍之調查;在春秋則有千乘、萬乘之邦,即兵車調查之比較;他如史記之年表,漢書之表志,則尤爲我國統計中之傑作.此後如歐洲大陸各國,無不用統計之方法,去考察其國家政治之情形,而尤以十六世紀西歐之商業日漸興起的時代,邦人尤多提倡而樂用之.惟當時僅爲文言之記載,未知應用數字,且缺乏專家之研究,故未成爲純正而有系統之科學.直至十八世紀中葉,始漸有簡明數字之記載.至十九世紀,國家之情形益複雜,政治之材料愈豐富,數字之記載法,亦愈趨進步,於是應用之範圍,因之亦日加擴大.而昔日僅用以記載國家政治情形者,至今則凡蒐集數字材料之科學,如物理、化學、天文學、氣象學、生物學、遺傳學、心理學、教育學、社會學、人類學等等,莫不借重統計之方法,而作發明原理,表明狀況,或各種研究之工具及行徑矣.故就目前情形而論,則

統計學之自身,不惟成爲純正之科學,卽謂之科學之科學,亦不爲過.

2 統計學之定義

統計名辭之由拉丁文 Status 而來吾人已言之如上故統計學最初之定義,實不過拉丁文所謂考察國家情勢之學術 (Inquiry into the condition of a state) 而已,厥後此學逐漸發達,應用日廣,內容亦隨之而有變更,故其意義,亦未免因此而有過狹之嫌.加以各學者對於研究之目的及其見解有廣狹之別,故所立之定義,亦頗難一致.如綜計各家所立之定義則不下數十種,惟據現代的眼光看來,不是簡而無當,就是偏而不全,殊難表示目前所謂統計學之眞正意義.玆將足以代表之各家所下的定義,略述五六如下:—

I 橫山亞男之定義 氏謂統計學者,乃觀察社會與國家動靜現象之學術也.

II 鮑來之定義 鮑氏 (A. L. Bowley) 稱統計爲平均數之學,或計算之學.

III 尤爾之定義 尤氏 (G. U. Yule) 則謂凡可量之事實,而受無窮原因之影響者,皆爲統計學之所當研究.

IV 金氏之定義 金氏 (W. L. King) 則曰統計學者,乃一種用分析某種枚舉 (An enumeration) 或某種聚

集的事物 (Collection) 所得之結果,以判斷 (Judging) 自然的或社會的現象之學術也.

V 帥克利斯之定義 帥氏 (H. Secrist) 則謂統計學者,乃爲由多數原因所深切造成者的種種事實之彙集,(Aggregates) 用有系統之方法搜集,而求達到預定之目的,並求其彼此間之關係,依次排列,按照相當精確之標準,用數字敍述,枚舉,或估計之,以闡明各種單獨或相關之現象之學術也.

VI 桑戴克之定義 桑氏 (E. L. Thorndike) 則謂統計學爲測量宇宙間萬事萬物之狀況,及其差異變化,以及其相關之方法.

上列六家之定義,橫山亞男之所言,似專偏於考察國家之情勢.鮑氏之所定,則專言統計學一部分之方法,似欠周到.尤氏則未見詳明,頗難使初學者領會.金氏則僅言分析的方法,而未提及綜合之作用.帥桑二氏之定義,可稱完全,吾人當咀嚼而細味之.

VII 本書之定義 吾人綜合上列各家之所述,並截長補短,可暫定之如下,非敢云定義,實歸納各家之敍述耳.

統計學者,乃根據科學的方法,搜集或測量宇宙間萬事萬物之狀況,用分析,綜合,比較,或枚舉,估計等各種的方法,而將其所得之結果,以推求其間的差異,變化,相關,及確度

之學術也.

VIII　何謂教育統計學　所謂教育統計學者,按朱君毅氏之所說,*即以統計學之方法,應用於教育問題之上,其與他種統計學出入之處,即在材料之不同,及方法注重點之互異.

* 見朱君毅教育統計學第三頁.

3　統計學之大別

近代統計學,大別之可分為二:即(1)統計法,(2)應用統計學.前者常以發明原則,定理,方法,公式等,為其職志.猶如純粹科學家之以發明原理,原則為其目的相同.至於應用方面,則留待他人,非彼等之所欲過問.故有時亦得稱之為純粹統計學.(Pure statistics)至於應用統計學,(Applied statistics)則係就前者所發明之定理,原則,方法,及公式等,應用之於具體事實.猶如應用科學家之就純粹科學家所發明的,去應用於實際者相同.教育統計學屬後者,而不屬前者.

4　教育統計學之內容

教育統計學之內容,概言之,得分(1)表列法;(2)圖示法;(3)分析法.表列法是將散漫的事實,整理排列

出來,使之簡單明瞭.圖示法是將欲表示之事實,用圖形作最清楚,最顯著之表顯,使閱者一目瞭然.分析法乃爲將簡約之數字,顯示所蒐集之材料之內容,通常多用集中,差異,相關,確度四種數量,而顯示之.此三者,皆爲顯示統計材料之重要法則,各有其相當功用.而本書卽欲將此三者,在以後各章中,略加討論及敍述.

5 教育統計學之需要及其應用

I 教育統計學之需要 *「自科學的精神及方法,應用到教育問題上來,統計學就成爲考研教育的重要工具.有了此種工具,教育學者,才能從複雜的事實中,尋出條理來;從偶然的情境中,找出原則來;從枝枝節節的現象中,理出些因果或連帶的關係來.蓋統計之功用,如同顯微鏡,望遠鏡,能補足我們肉眼的不足,牠能將種種現象縮起影來,使我們看得清楚,又能幫助我們找眞理.故欲在教育方面發古人所未發,明今人所未明,都非得直接或間接借重統計法不可.至於辦教育行政的,或當教員的,要想知道自己所用方法的效率之如何,也非得懂些統計法不可.」故時至今日,凡是研究教育或從事於教育的人們,幾乎一刻不能離開此種工具.此教育統計之所以爲必要也.

本節係轉錄陶知行先生代薛鴻志氏所著教育統計學大綱書中之序言二中的一段,特此聲明,並深表謝意.

II　教育統計學之應用　教育統計學為研究教育者一刻不能離開的工具,吾人已略如上述.如果再把牠直截了當的來說,那末就是教育上各種問題之研究,都有應用教育統計學之必要.若簡單言之,得分以下之各方面:——

1　關於教育行政方面者　如行政機關的組織、行政效率的研究、行政人員的資格與俸給、任免法規與職務之分擔以及經費之歲入、歲出、來源、用途,及學款之支配等等.

2　關於教訓方面者　如課程分量、課本內容、考績方法、兒童升降辦法、教師之學識、經驗、俸給、學生之年齡、年級、退學、轉學、懲戒、獎勵之章程及學生之學業、操行、體格測量、疾病檢查、畢業生之趨向、家屬之職業以及地方之風俗習慣與社會生活之需要等等.

3　關於教育學理方面者　如各種問題之實驗及研究、方法取舍之決擇、各種學科之價值、學制之採用或試驗等等.

4　關於學生學習方面者　如學生學習進步之速率、分量、限度、疲勞、形式之陶冶、性向之好尚、努力與偷懶

等等.

5 關於教育測驗方面者 教育測驗,無論就實施方面,或編造方面來講,均離不了教育統計.如實施時而離了教育統計,則實施後決無結果之可言.蓋測驗之本身,是一種工具,其意義及效用,全賴統計方法去表顯或解釋.如編造時而離了教育統計,則測驗之標準及常模,或測得之成績,根本無從而訂定或整理.故教育測驗之能建樹於科學基礎上,未始不有賴於教育統計學.

以上所述,僅舉舉之大者.要之,教育統計學之應用,不僅如教育官署,或學校中所編繪之教育統計之圖表.而其眞正的任務,非僅止敍述往事之記載或報告,而在注重於分析,綜合,比較,以推求其間所呈顯之常理,而作考研各種學理的唯一的工具,藉以作改進的法則.

練習問題

1 試略述統計事實之發生及史略.

2 試略述各家對於統計學所立之定義,並批評其長短.

3 統計法與應用統計學之區別何在?

4 教育統計學與其他統計學之區別何在?

5 試略述教育統計學之內容.

6 試略述教育統計學之需要及應用.

第二章　教育事實之徵集

6　教育事實之來源

統計方法,雖爲研究各種學科之工具,但無事實以作研究之材料,則雖有良好之工具,亦無從入手.關於統計事實之徵集,及其事實之來源,則每因學科之不同,而略有差異.本書因限於篇幅,未能詳加討論,故僅就教育事實之來源,作一簡單之敍述.

I　來自教育行政團體之報告者　例如國家所頒布關於教育之法令,教育廳,局,所製定之各種學校規程及命令,以及教育部彙編之教育法規,及全國教育概況等等.

II　來自學校規程及教育文件者　各校除遵用一定的教育法規外,尙有其特殊之規程,如組織系統,會議規程,課外活動,暫行學則,經濟之出納,學生之升降及進退,教職員之聘任及解約,成績考查,課表之支配,學程之增設,教學之進行,作業之程序等等,均可搜集而作爲統計之資料.他

如教育上重要文件之往返,亦均應在搜集之列.

III 來自公共教育團體之報告者 公共教育團體之報告,可謂最富於研究之資料.例如各省,縣之教育年鑑或年報,以及統計圖表,與夫各學校的一覽,概況,紀念冊,計畫書,週刊或季刊等等,均富於實際之材料.他如各種公私所組織之會社,(例如中華職業教育社,平民教育促進會等等.)亦時常能搜集到各種實際之材料.

IV 來自各種教育研究及測驗者 至於各種教育研究或實驗,以及各種教育上與心理上之測量,則尤富於實際而難能可貴之材料,吾人均當搜集之.

V 來自各種教育調查及考察之記錄者 例如學校調查,省縣督學之視察報告,公私團體之參觀記錄,或私人之考察日記等等,亦均可搜集而作為資料.

VI 來自其他者 此外如關於教育之著作,與雜誌及報章上,常見有實際之材料,亦當隨時留意而搜集之.

以上之所述,僅就其主要之來源,略加指明.其實凡是與教育問題有關係之材料,均當時加留意,而隨時搜集之.

7 教育事實徵集之方法

教育事實之來源,吾人已在上節中,略事敘述,茲將進而研究徵集之方法.關於徵集之方法,如細分之則甚多,茲為

簡括計,僅舉最重要的幾種,略加討論.

I　表格訪問法　此法在西文曰 Questionnaire methods, 在我國則譯之而曰"表格訪問法,"間亦有譯作"條問法"或"問答法"者.此法普通用之者甚多,蓋取其輕而易舉也.至於此法之實行及製作,亦簡而且便.祇要將自己所需之材料,發爲問題,印寄各地,請人塡答,答就後,請其寄回原處,以備統計.不過答塡所發之問題,非他人應盡之義務,有時往往隨意塡寫,確度甚小;或竟有付之字紙簍,置之不理,以致收回者寥寥無幾,實有得不償失之效果.故對於這一點,一般用之者,都說是此法之最大缺點.惟據有經驗者之報告,訪問表格之優劣,與其收回之多寡,成密切之正比.通常表格優良者,恆能收回二分之一,或四分之三.表格過劣者,則常求半數而不可得.是則剛才所說的最大缺點,似不全在方法之本身,而在所發訪問表格之優劣.欲求訪問表格製作之優良,當注意下列各點:——

1　問語務求簡單明晰,並須直接,而便於答復,不可含糊迂迴,需人思索,或竟而發生誤會.如遇有專門名詞,須下一簡單之註釋.

2　問題不可過少或過多,過少則不足應用,過多則徒增手續之紛煩,及塡答人之惹厭.故確定問題之種類及多寡,祇求能將自己所要之材料,全體包括無遺爲準

則.

3 訪問事項之主旨,均當詳細載明,使答填者明瞭其用意,而不致僞報或不覆.

4 所設問題前後均須銜接.一切問題,尤以有關於研究主旨者爲限.

5 所設問題,須合於論理的; (Logical) 行格之寬狹,尤應適於填寫.

6 每問所要之答案不可過長,最好在十個字左右,或只用一二數字,及"是""否"等單字作答之地位,問卷上須預爲留出.

7 凡計算上所有總數,百分數或平均數等各事項,切不宜使填答者任之.

8 訪問表格擬好後,須先用油印二三十分,先行發填作初試,看收回後,有無須修改之處.

9 表格經修改後,最好用較硬之紙鉛印,其大小及式樣,以合於填寫,便於攜帶,或將來整理收藏時之利便爲標準.

10 每一訪問表格發出時,應附上郵票,或粘有郵票之信封,以備填答者寄回之方便.

11 發出訪問表格之數目,須倍於自己所希望欲收回之數目,否則將來或有材料不足之虞.

12 塡答者之姓名住址等,非必要時,可隱而不塡,或聽其自便.

II 親自調查法 用表格訪問法作調查,吾人已略述之如上,雖能以較小之經濟,及較省之時間或人才等,而收得較多之材料.然不確不盡之處,在所難免.其可靠之度,往往甚小.因此親自調查法,才應此而起.所謂親自調查法者,即親身實地去調查或搜集材料之謂也.由此法所得之材料,其可靠度雖能增加,然時間及精力之所耗,卻較多,故在範圍狹小,現象簡略時,或可一行.至若欲得心理或敎育測驗及各種敎育問題實驗硏究之材料,則總以親身躬行之,比較可靠.

III 其他 除上述兩法以外,尙有所謂抽樣調查法,及用原有表册法者.所謂抽樣調查法者,即對於所欲調查之事實,抽其能代表者而調查之.此法之當注意者,即在抽樣時之必須爲隨機 (At raudom) 之選擇,或多抽能代表之人數,以減少其機誤. (Probable error) —— 請參閱下列第九章38節.——所謂用原有表册法者,即將被調查之處所,原有的簿籍表册上之記錄,拿來作統計之謂也.例如欲調查某城之小學兒童之實足年齡,則即將各小學學生之學籍簿,拿來逐一謄錄,然後計算之.在此法之所最宜注意者,即在謄寫之問題.蓋各校原有之簿籍表册,勢不能久假不歸,

任吾人作長時期之統計,故必須親自謄錄,謄畢即行送還.然原有之記錄格式,每不一致,謄寫時,勢必彙而類之,此時最易蹈遺漏,妄增,誤入及顚倒諸弊.故非細心者,恐難勝任,此亦爲吾人所應注意者.

搜集事實之方法,不僅上述的幾種,惟這幾種比較普徧而常用,故吾人特提出而略述其梗概.至其他各法,仍宜隨時活用,不必拘泥.

練習問題

1 試言敎育事實之來源有幾種?並略述其梗概.

2 試略述敎育事實徵集的幾種方法.

3 表格訪問法之缺點爲何?優點爲何?

4 試述訪問表格製作時之注意點.

5 抽樣調查法之應注意之點爲何?

6 用原有表册法去搜集統計材料之注意點爲何?

第三章 表列法

8 表列之意義及功用

I 表列之意義 表列者,類別排列之謂也.係將散漫之事實,分門別類,列成系統,俾易研究或考察.蓋不有表列之法,則漫無系統之事實,必無從整理,既不便用精細繁複之統計法,以作精深之研究,又不便用淺顯易知之圖示法,——詳述第四章——以供大衆之瞭解.故表列法實爲統計事實之縮影,不僅爲表顯統計材料重要方法之一,且亦爲一切表顯法則之初步,亦卽爲一切表顯法則之根本.

II 表列之功用 約而言之,得分下列六點:——

1 能發現其一定之規律 事實之現象,往往雜揉紛錯,漫無系統之可尋,如若排列有序,自不難發現其規律之狀態.

2 便於比較 材料未經表列,則彼此間相互之關係,或其同異之點,必難以覺察.若分類而排列之,則其間

之關係,一目暸然,比較自易.

3 易於記憶 凡統計材料一經表列,卽極爲整齊.因其得有論理的次序,聯念的幫助,較之淩亂無序之材料,自然易於記憶.

4 宜於總計 若欲知事實之總數,則表列更不可少.雖說是總數一項,亦可用他法以計算之,然因全體事實,列在一縱行或一橫行上,排列整齊,固不若散亂時之計算綦煩也.

5 免避重複 如用普通書文報告統計的事實,則往往有同一事實,而反覆提及至數次者.若用表列,則同一事實,祇要用一個標目去概括一切,故能免去重複.

6 易得具體之概念 散漫之材料,讀之不特難得其要領,抑且淆亂不清.如一經表列,則繁賾之事實,畢呈於一紙,且排列有序,故易得具體之概念.

9 表之種類

表之種類,若細行分析,實在太多,一時殊難枚舉.玆僅就所包事實之多寡,及式樣構造之繁簡,作分類之標準,如下列的四種:——

I 單項表 單項表亦稱一重的表列,表內僅包括一種事實,故甚爲簡單.玆舉例如下:——

表一　某中學某年度各級學生數

年　級	學　生　數
初中一年級	48
初中二年級	42
初中三年級	30
高中一年級	44
高中二年級	40
高中三年級	28
總　計	232

II　二項表　二項表亦稱二重的表列,表內包括兩種的事實,其例如下:——

*表二　上海市二十年度公私立幼稚園數及兒童數

類　別	園　數	兒　童　數
省立…………	1	38
市立…………	13	639
私立已立案	5	336
私立未立案	20	843
總　計	39	1,856

III 三項表 三項表亦稱三重的表列,表內包括三種事實.如下列表三,能知校數,學生數,及教職員數三種事實其例如下:——

* 表三 上海市二十年度公私立師範學校數學生數及教職員數

類別	校數	學生數	教職員數
市立………	1	117	24
私立已立案	4	943	121
私立未立案	4	326	42
總計	9	1,386	187

IV 四項表 四項表亦稱四重的表列,表內包括四種事實.如下列表四,除知校數,學生數,教職員數三種事實外,並亦能知經費數.其例如下:——

* 表四 上海市二十年度公私立專科學校數學生數教職員數及經費數

類別	校數	學生數	教職員數	經費數（以元爲單位）
國立…………	3	539	47	4,440
省立…………	1	928	77	112,000
市立…………	4	1,153	137	78,890

公共租界工部局立	5	859	59	396,222
法租界公董局立…	1	508	23	136,316
私立已立案	49	14,127	1,299	1,451,171
私立未立案	49	5,649	610	783,345
總　計	112	23,763	2,252	2,961,884

* 上列表二,三,四,中之材料,係錄自二十二年二月二十二日時事新報第二張第四版之轉載.

表之項數,本無限制,如五項六項,甚而至於七,八,十餘項,照理論上言,固無不可,全在製表者酌量之採用.不過與其合多項紛繁的事實,作成一極複雜之表,則不若抽其重要之事實,各舉比較簡明的表列以便於閱讀,而易於記憶.至三項及四項表,則亦並無不便於閱讀或記憶之處,吾人宜多用之.

10　作表之規則

欲使事實表列明晰,必須所作之表,合乎種種規則.否則表列之後,仍不能使閱者明瞭,就失去了表列的功用.作表之法,在一般人的眼光看起來,似甚簡單,一若無討論之價值.然若作者任意爲之,則每不倫不類,乖謬百出.卽近今一般統計學專家,對於此項工作,亦復認爲窮思竭慮,方能入

般,何況一無經驗者.雖云作表之規則,並非固定不變,可以隨時更改,然下列之規則,均由經驗而成,不無參考之價値.茲申述之:

1 表數及表題,均須寫在表之上端,其故由於表之排列,係自上而下,以上部爲起點.此適與圖相反,蓋圖之繪法,係自下向上,故圖題應寫在下面.

2 表題須簡賅明晰完善,以無庸另註解釋而能了解表中所列的事實爲宜.

3 表題中所舉之各項目,須與表中所列各項目之次第相同.(參閱表一至表四)

4 所表之事實,如有空間或時間的關係時,應在表題中註明.否則他人引用時,必感困難.此普通人每多遺漏,學者極宜注意.(參閱表一至表四)

5 有時表中之重要數字,或重要項目,須指出時,宜列入名稱之內,俾引起讀者之注意.

6 表的項目的讀法,應自左而右,自上而下.每見有許多的表,其項目自右而左,閱讀極形不便.

7 項目多的時候,可分大項目及小項目兩種.惟小項目應列在大項目的下面,並應向內稍移,(參閱第五表)大項目可用粗畫字表出之,藉醒眉目.

8 各項目之次序,可依照位置之先後,等級之高低

時間之早晚,及筆畫之繁簡去排列.若爲事實所許,則可依照數目的大小去排列.

9 有時對於某重要項目,如欲特別表顯以資比較時,則所欲特別表顯之項目,及該項目之數字,可用粗畫字或紅顏色表示之.

10 總數宜置於表之極右端或最下層,最好能用較粗大的字體表出之.

11 如長表須分作數頁或數表排列時,則以後各頁或各表,必須重註項目,以免閱者發生不便.如橫幅甚長之表,左部所註項目,不便通覽有上下看錯之感時,則於右端可照樣添入.

12 如表中行數太多的時候,則每五行可隔一空行,或以粗線表出之,以便閱覽.

13 如各項目之文字較長,有分寫於兩行或以上之必要時,則皆宜橫寫,以免閱表者空費辨別之時間.如上列表四的"公共租界工部局立"如有分寫於兩行之必要時,宜寫作:

公共租界
工部局立; 不應寫作 · 公租工局
共界部立.

14 如各項目之文字,有的較長,有的較短時,則較短項目之文字的右邊,可用點線向右連接.例如表二之"省立"及"市立"較短,"私立未立案"較長,故將省

立及市立之右旁,用點線連接,以導引之.

15 如能用雙式分配表,則尤為便利.此類表之功用,能在同一表內,排列兩種材料.一為自左而右橫列,一為自上而下豎列.蓋取其能在一表之中,可並列二表之材料也.(參閱第六表)

16 表中每一豎行,須以直線畫分之.表中排列數字之處,通常不用橫線.蓋一用橫線,反足以使表不清楚,不美觀,而不易觀察.惟上部所列項目處之下,及下面所列總數處之上,均須用橫線區分之.

17 細項目之間,用一直線已足.如大項目則宜用雙直線或粗直線.

18 表之頂端及底邊,須用雙直線或粗直線,通常多用雙直線.

19 表之左右兩邊,不宜用線圍住.若圍住之,則足以使此表形似箱篋,殊不美觀.

20 表中遇到沒有數目可塡的地方,應該塡一虛線或短實線,不可塡零字或空缺.蓋塡零字,能使閱者眼倦或眼花,空而不塡,恐閱者意為遺漏.

21 表中排列之數字,如有特別意義時,可用粗畫字表示出來,使閱者易於察覺.

22 表中數字,須一律用阿拉伯字橫寫,漢字則極不

適宜.國內各教育機關之各種學務報告表,至今仍有沿用漢文數字而直寫者,似非速加改革不可.

23　數字單位應該上下相對,不得參差,遇有小數時,則小數之位數,尤須排齊.否則加減及比較,極感困難.有時數目多至四位以上,則宜用分斷點,每三位分作一段

24　表中數目字,字體不得太小,否則易使閱者眼倦.

25　表的說明及附註,應在表的附近,不應相隔太遠.普通表的說明,都列在表的下面.(參閱第五表)

26　如事實中斷,宜留出固有地位,並宜作虛線以表出之.(參閱第七表)

*表五　十九年度江蘇全省各種學校概數表

(根據二十一年十一月出版江蘇教育概覽)

類別	校數
初等教育………	**8,346**
幼稚園………	86
初級小學………	7,356
完全小學………	898
其他小學………	6
中等教育………	**186**
初級中學………	87

完全中學………	36
師範學校………	45
鄉村師範………	12
職業學校………	6
高等教育………	9
總計	8,541

附註：表中之其他小學係指省立鄉村師範實驗小學而言

*表六　以德爾滿算學四則測驗試驗京市九個小學之中高各級男女生總數

學校	三年級		四年級		五年級		六年級		總計
	男	女	男	女	男	女	男	女	
A	40	34	34	29	42	30	38	30	277
B	32	28	26	20	36	26	32	27	227
C	21	20	27	32	28	20	26	32	206
D	30	40	25	26	38	36	29	34	258
E	38	32	24	29	28	28	48	63	290
F	23	23	25	21	30	26	32	28	208
G	12	14	11	12	14	15	12	18	108

H	14	9	15	12	19	14	10	9	102
I	13	11	14	16	18	13	16	10	I11
總計	223	211	201	197	253	208	243	251	1,787

*本表轉錄拙著:小學算學四則能力之實驗研究文中之第一表,原文見中華教育界十九卷七期.

*表七　紐約城學校用吳狄(Woody)之算術量表測驗所得之中分數(1917年5月)

測驗	年級及學期							
	III		IV		V		VI	
	A	B	A	B	A	B	A	B
加法	12.0	12.6	13.3	14.5	......	......	17.5	18.1
減法	8.0	9.1	10.0	10.7	......	......	13.7	14.2
乘法	7.5	8.2	9.4	1.3	......	......	16.0	17.1
除法	2.1	6.0	7.2	8.0	......	......	11.3	12.4
總計	29.6	35.9	39.9	34.5	......	......	58.5	61.8

*本表轉錄周調陽氏著:教育統計學書中之第十表.

練習問題

1　表列法在教育統計學上,何以為一切表顯法則之

初步或根本?

2 試述表列之意義及功用.

3 試舉一及四項表或四重的表列的實例.(除書上所舉者外)

4 表題爲何須寫在表之上端?

5 所表之事實,如有空間及時間的關係時,爲何應在表題中註明?

6 雙式分配表,有何特長之點?試舉一例.(除書上所舉者外)

7 表之左右兩端,爲何不宜用線圍住?

8 表中遇到沒有數目填的地方,爲何應該填一虛線,而不該填零字或空缺?

9 數字單位或有小數時,爲何要上下相對,不能參差?

10 如事實中斷時,爲何宜留出固有的地位,並作虛線以表出之?

11 有時表中之重要數字或重要項目,須指出時,宜用何法,方使人注意?

12 如橫幅甚長之表,左部所註項目不便閱覽時,宜用何法以補救之?

第四章 圖示法

11 圖示之意義及功用

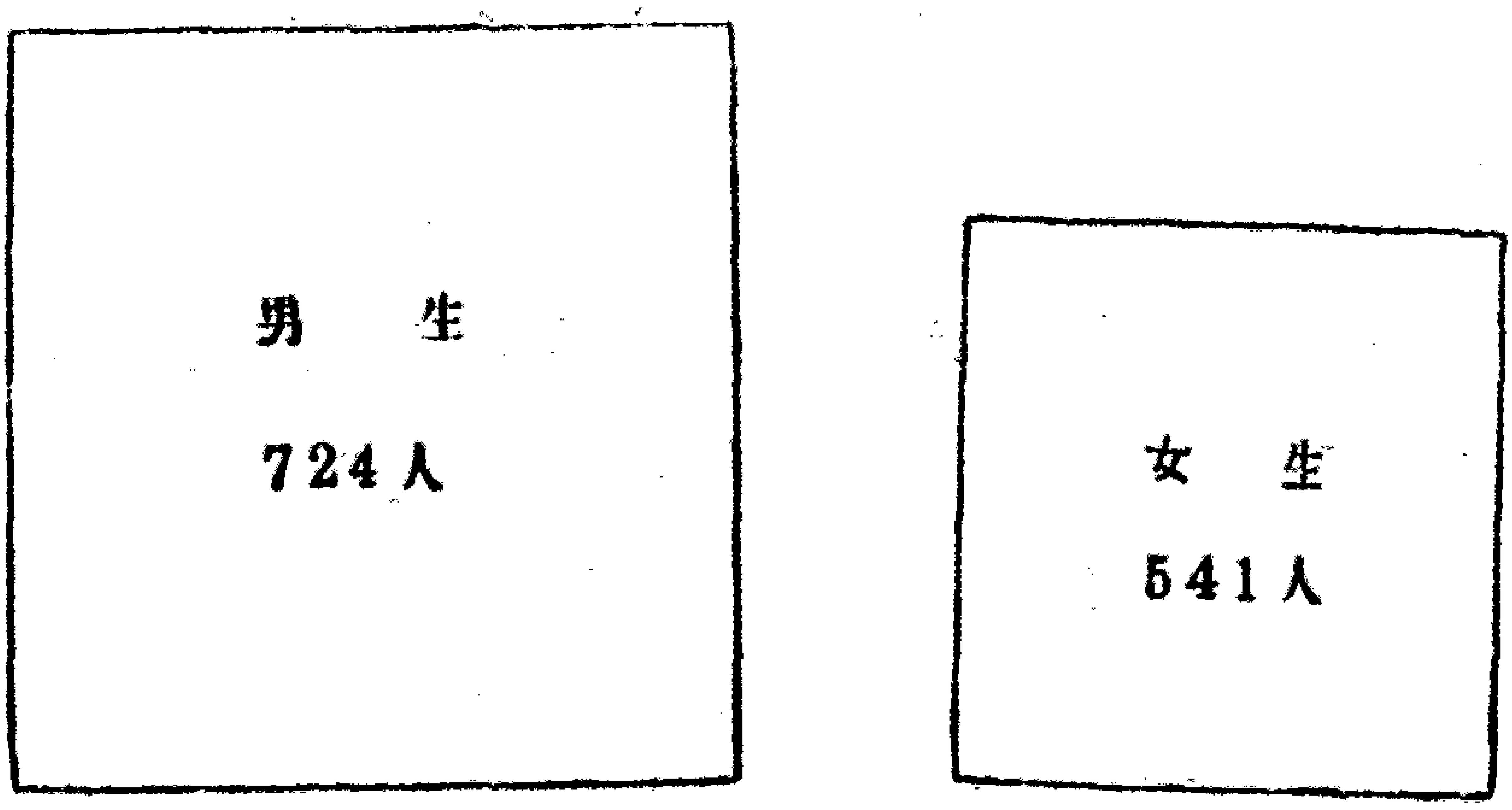
圖示者,將表列所得之事實,用適當的指事,象形諸法以呈現其結果也.表顯統計事實之工具,通常可分爲文字,表列及圖示三種,此三種工具,各有其相當之功用.惟比較言之,如用文字表顯,不若用表列之簡明.蓋用文字說明事實時,非惟繁冗,且亦不便於比較,而易滋誤解.如用表列以說明事實,雖覺簡賅正確,然數字滿紙,乾燥乏味,恐非受有相當之統計訓練者,多不願細加披閱,或詳行考察及比較.惟用圖形表顯之,則事實間相互之關係,一望卽知.雖屬庸常之人,亦可了解;且能於極簡短之篇幅中,最少量之時間內,作充分之比較.如再進而言之,則圖形仍能利用文字及數字說明各種要點.故如製造合法,則三者之長,均能兼有.是以圖示法爲表顯法中之最優者.

12 圖形之種類

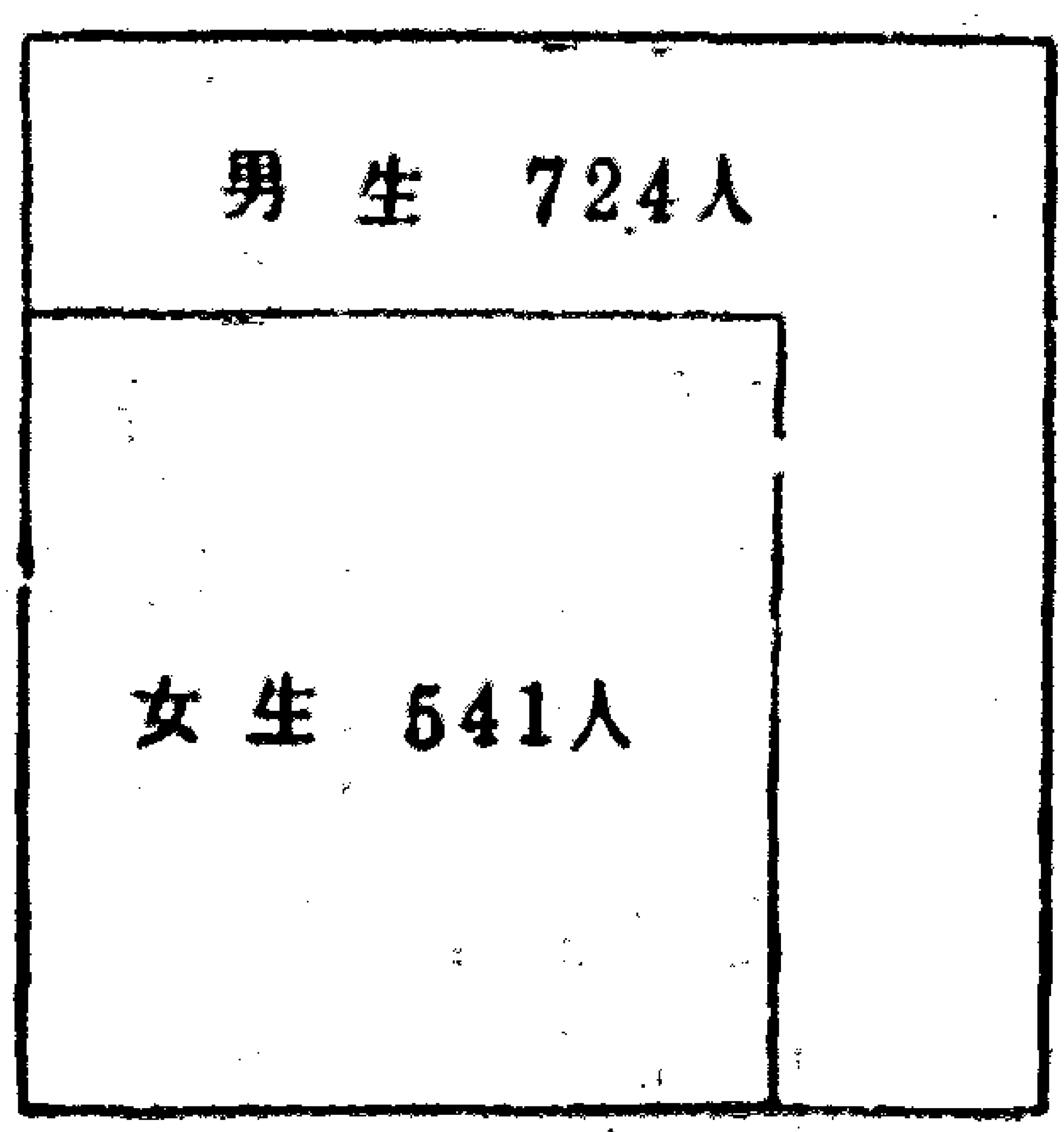
圖形之種類甚多,勢難一一敘述,且其中多有不適用者,故吾人實無盡量討論之必要.下列各種圖形,係選擇最常用者,並將每種圖形,略述其優劣,俾便應用時之選擇.

I 方形圖 此類圖形,係以一方形代表一項目,由各方形面積之大小,作比較之基礎.通常所用者,多為並列方形與疊置方形,間亦有用立體者.茲舉例如下列圖一、二、三.

(顯示並列方形之畫法)

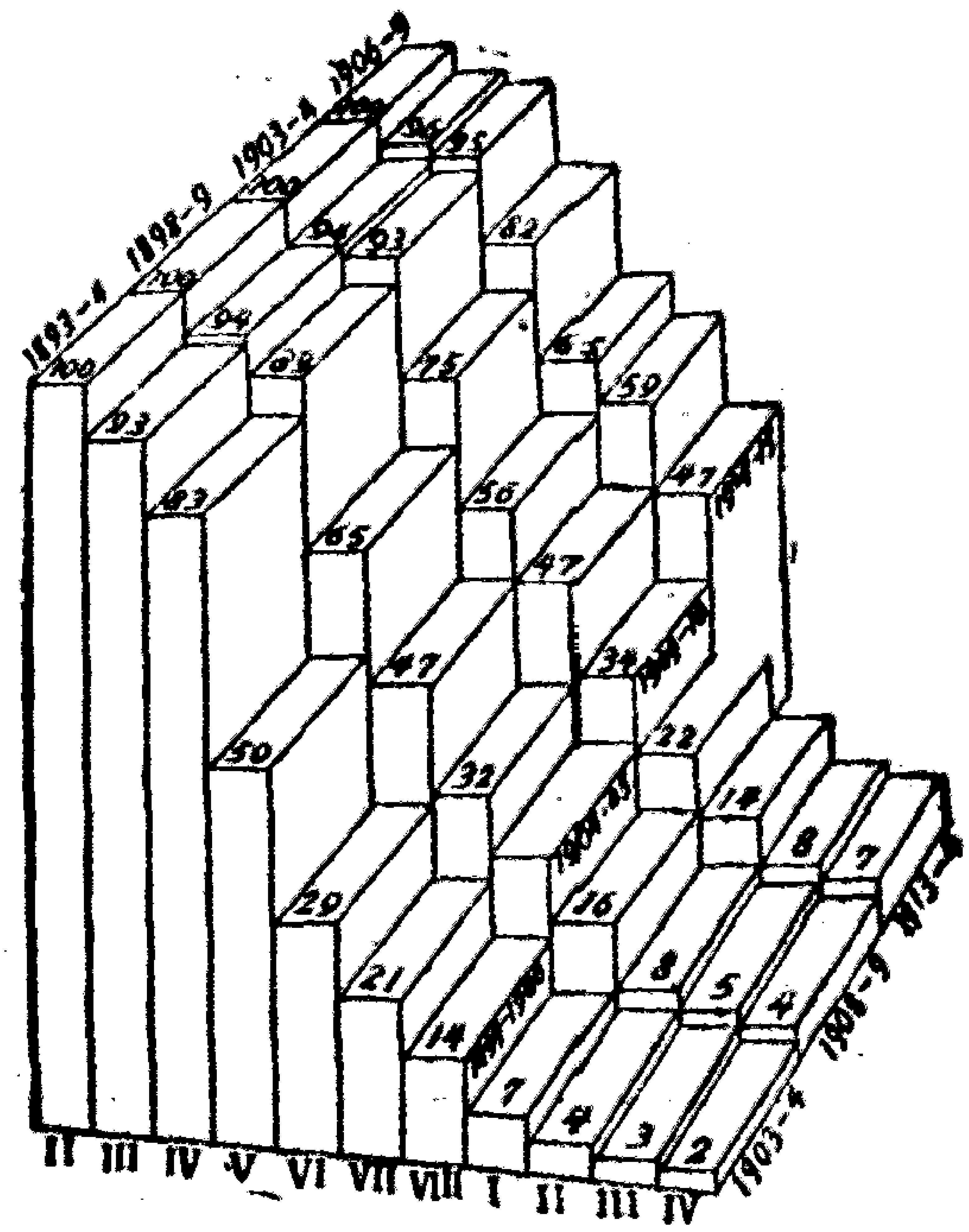


圖一 某小學某年度上學期在校男女學生比較圖



圖二 用圖一之材料顯示疊置方形圖之畫法

（顯示用立體之畫法）



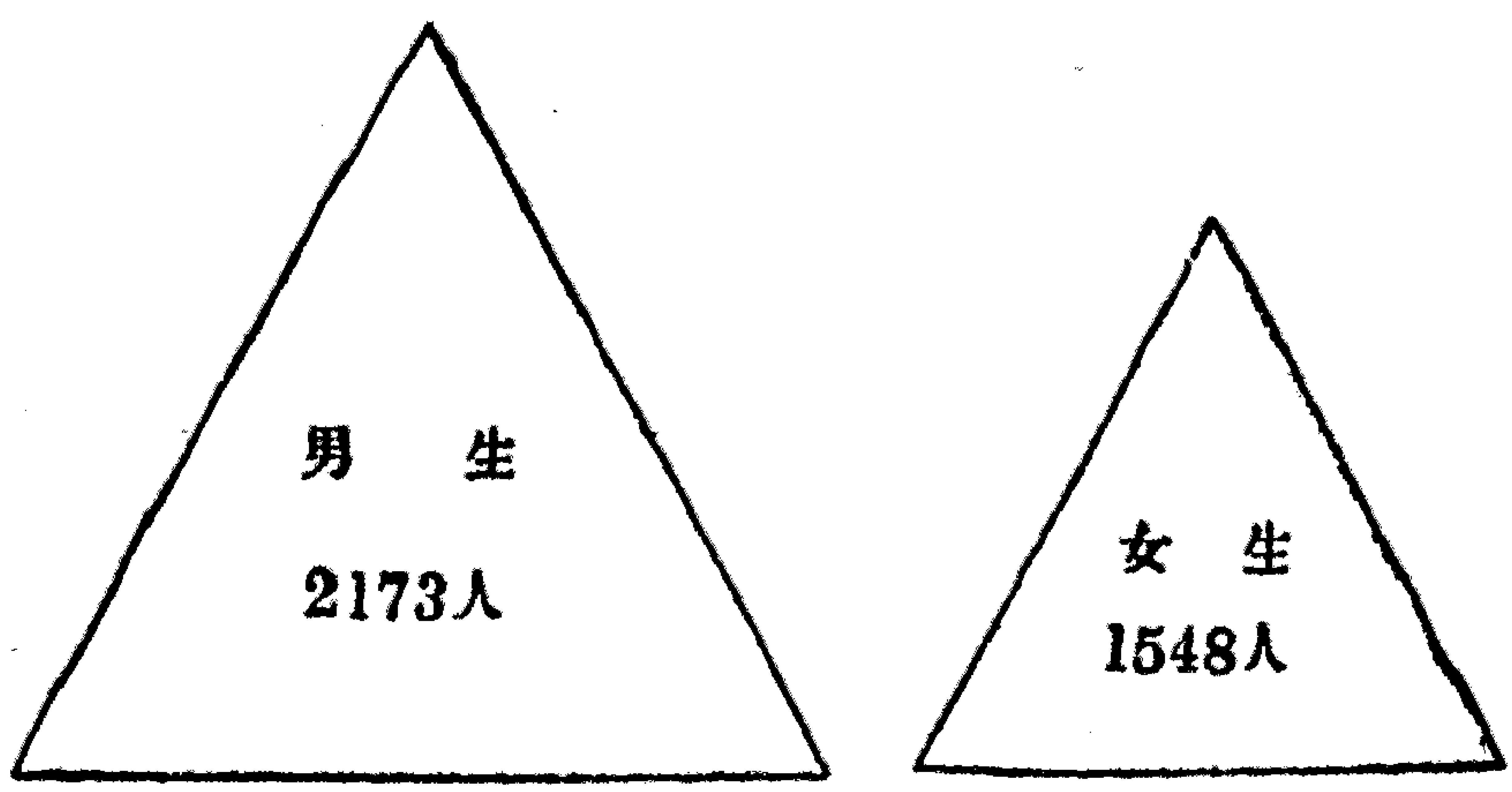
初級　　高級

＊圖三　美國聖路易學校 (The St. Louis School) 學生人數比較圖

＊本圖轉錄自 H. O. Rugg. Statistical methods applied to Education 書上 341PP. 第61圖.

此種圖形之缺點,爲各方形之面積,甚難正確,比較似覺不便,苟無數字以說明,幾令人不知其間之相差,故以少用爲宜.

IV 三角形圖 此類圖形,以一三角形式表一項目,亦以其面積之大小爲比較之基礎.其比較之不正確一若方形圖,故亦以少用爲宜.(參閱下列第四圖)

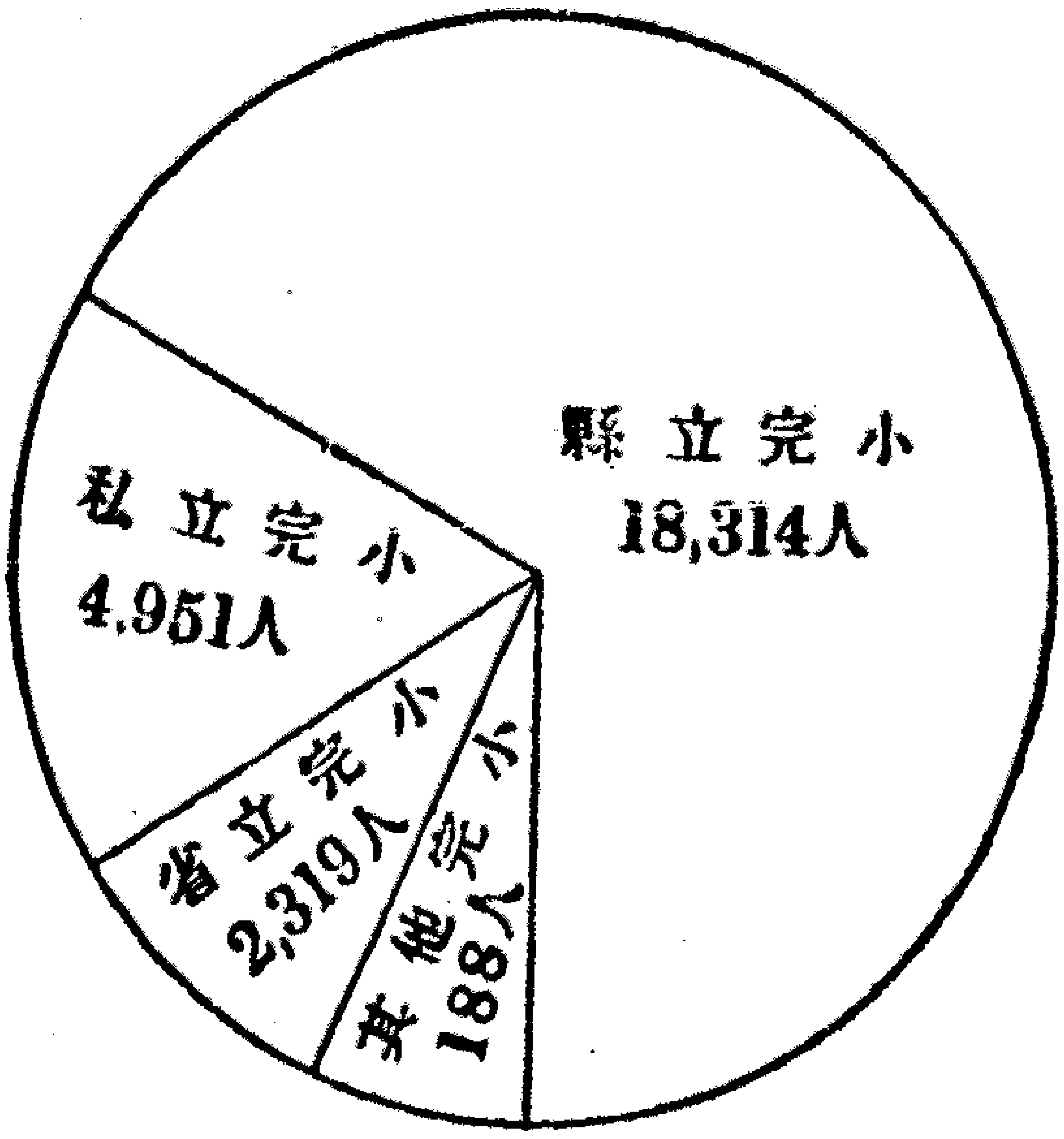


圖四 十九年度江蘇全省男女幼稚生概數比較圖

(本圖亦可將此兩三角形疊置如圖二之畫法)

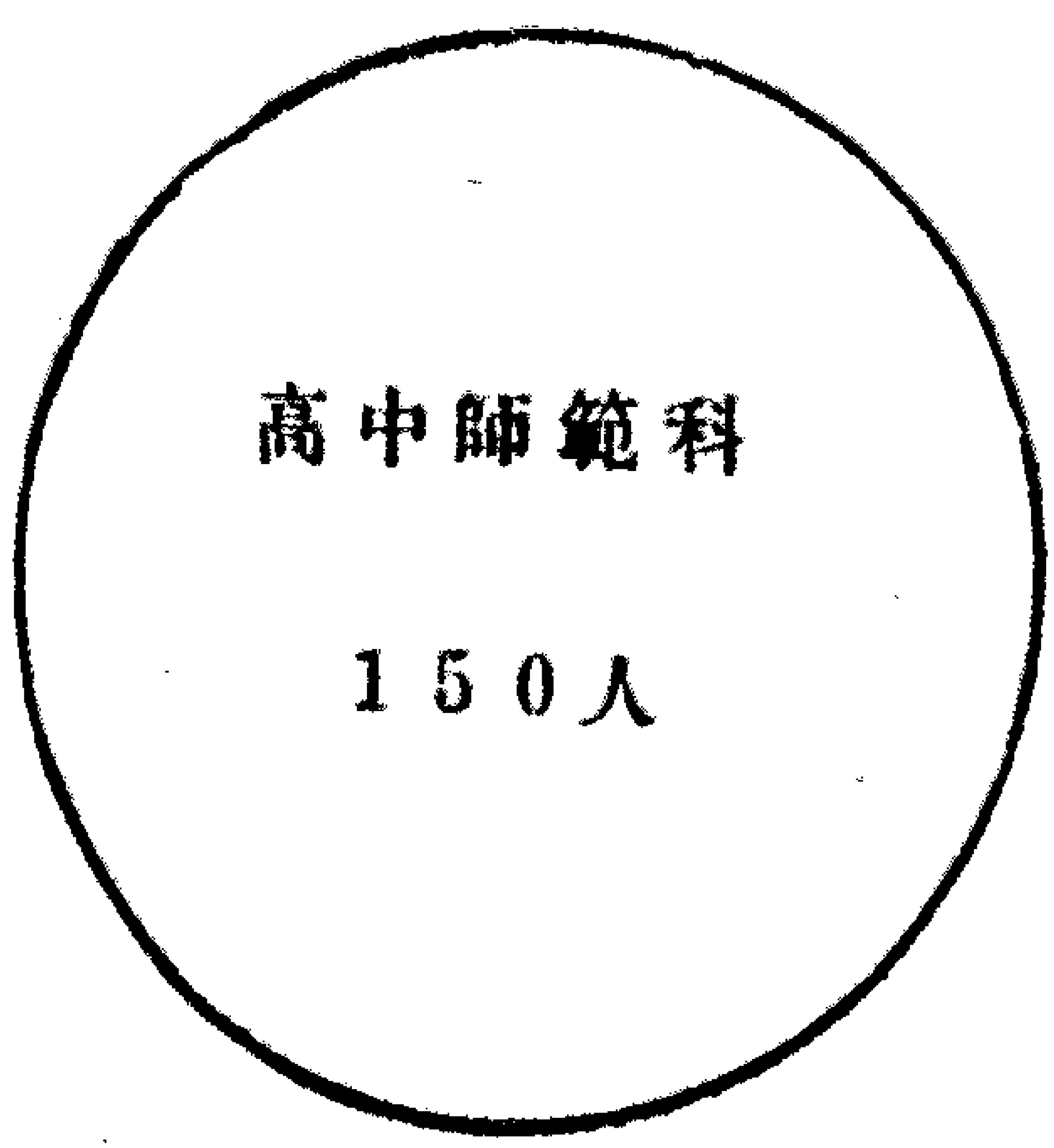
V 圓形圖 此類圖形,通常所用者可分(1)摺扇圓形,(2)單圓形,(3)疊圓形,今分述如下:——

1 摺扇圓形 係以全圖代表某一事實之全體,再按各項事實所佔之百分比數,將全圓分作若干扇形,以比較之.繪此項圓形時,每1%應佔3.6度,蓋因一圓周,爲360度也.舉例如圖五:

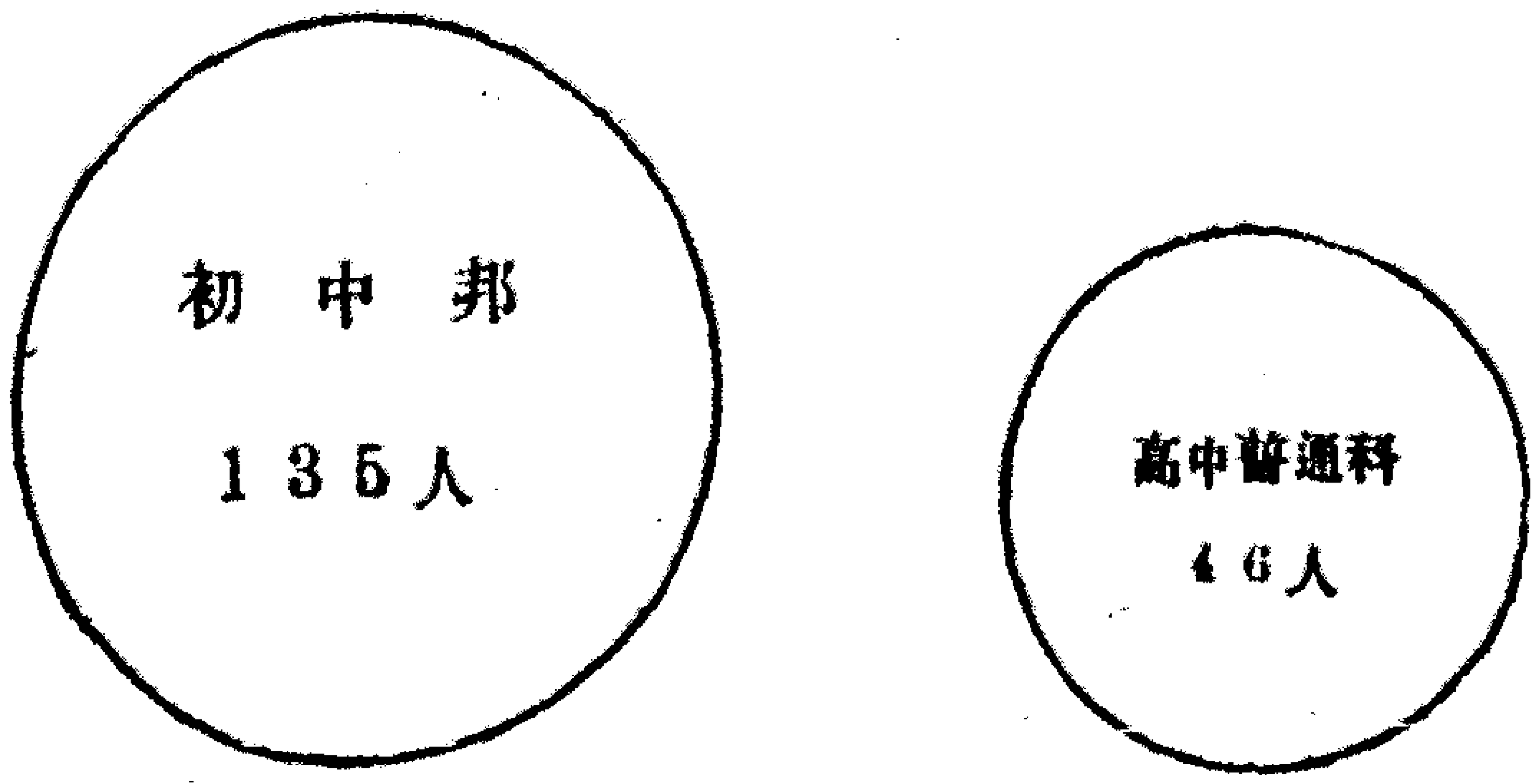


圖五　十九年度江蘇全省完全小學畢業生概數比較圖

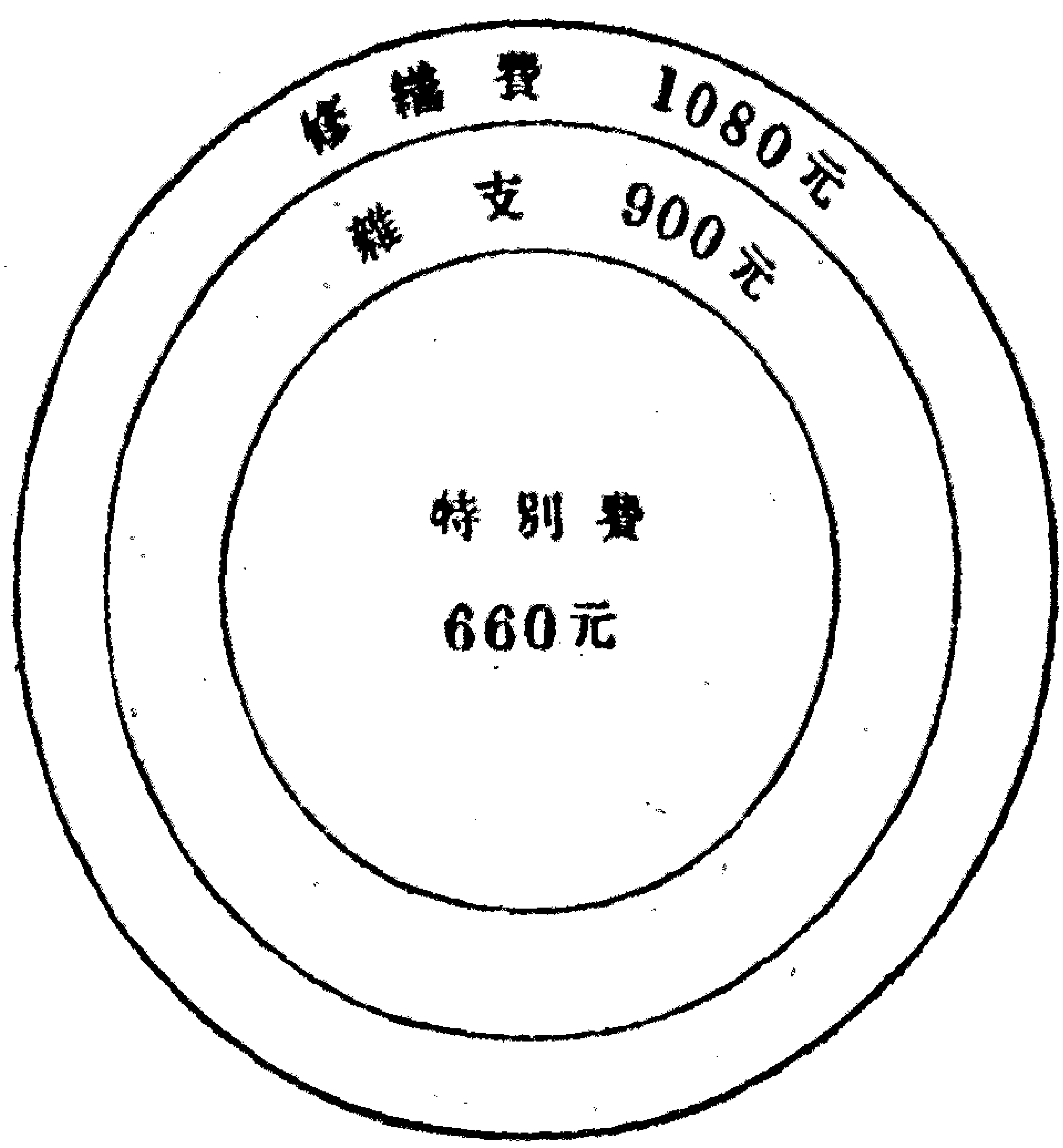
2　單圓形　係以若干不等之圓形,代表若干不等之事實,大者示大量,小者示小量.其例如圖六:



圖六　江蘇省立太倉師範學校二十一年度下學期在校學生人數比較圖



3 疊圓形 係以同一圓心作成若干之圓周,代表若干不等之事實或部分,各圓僅代表每圓之一部分,若以代表全部,則大誤矣:其實卽將上列(2)之單圓形,層層疊置也.舉例如圖七:



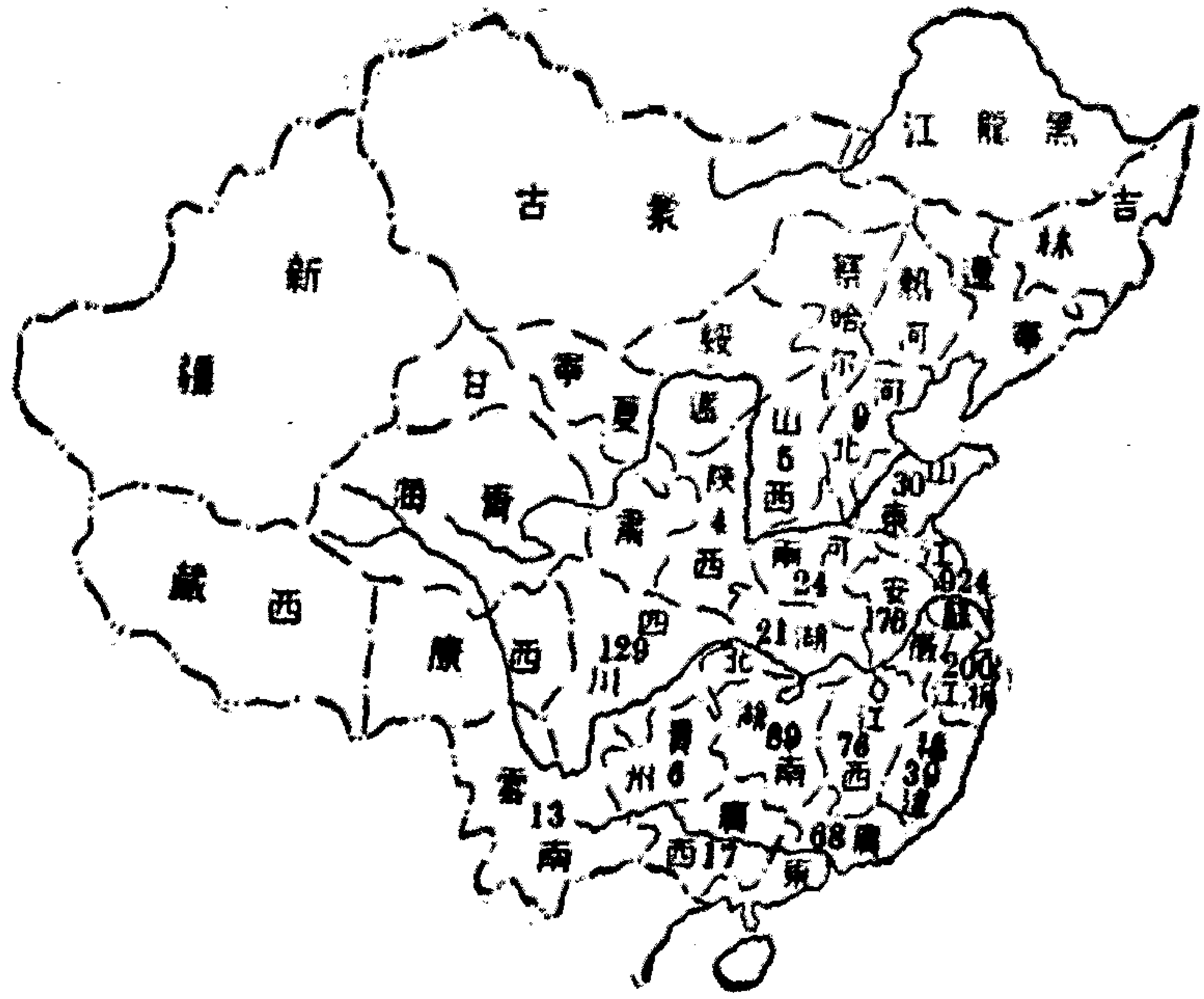
圖七 江蘇省立太倉師範學校二十年度各項經費預算比較圖

以圓形表示事實,本不甚妥,惟以其簡單易繪,且觀衆頗易了解,式樣亦尙美觀,故採用者甚多.惟鄙意亦以爲少用爲宜.玆將其缺點略述如下,希閱者加以注意:

a　摺扇圓形之缺點　(甲)假使其中有過小之扇形時,甚難註出該項目之文字及數目,即能勉强註出,亦模糊不清.(乙)各扇形中所寫文字及數目,其方向甚難一致,閱時殊覺不便.(丙)如各扇形相差之度甚小時,甚難辨別;即相差之度較大,若不註明數字時,亦頗難知其究竟相差爲若干.

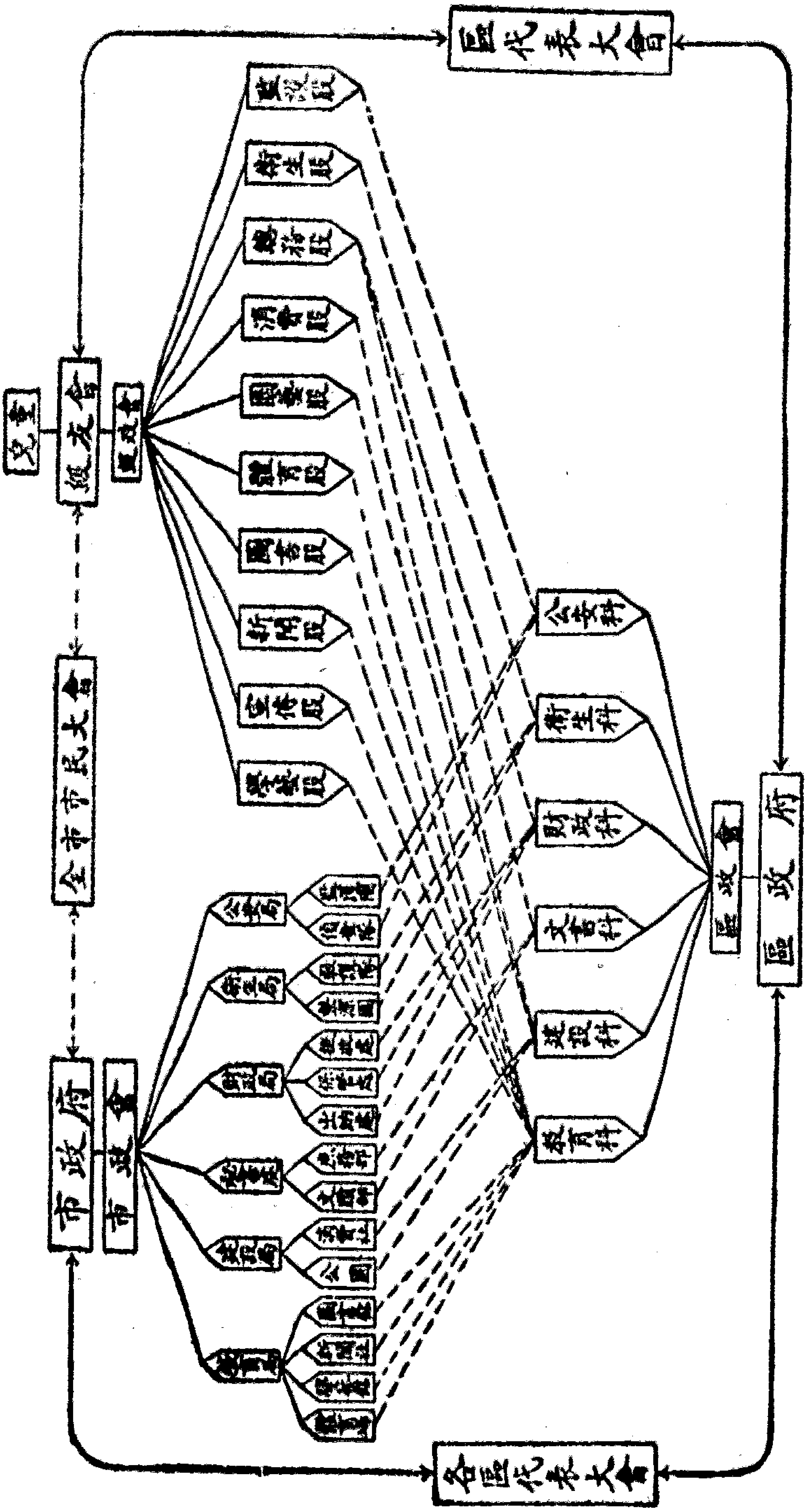
b　單圓形及疊圓形之缺點　(甲)不能看出各圓形所代表之面積之大小.(乙)閱者之視線,常爲圓形所混淆.近人多有將各圓形中,(尤以摺扇圓形爲甚)畫上許多五花八門或鮮豔奪目的圖案,意爲非如此,不足表其美觀或繪工之精緻.其實如此一畫,閱者之注意,不集中於所代表之事實,乃集中於所畫之精緻與否.此所謂枉費心血,而不切實際者,大可不必.

Ⅳ　地圖　此係以地圖代表事實,藉以顯明地方與事實之關係者也.惟其繪法與通常地圖不同.蓋通常之地圖,關於地理之特點,繪製頗詳.至圖示法中之地圖,則可不必,即不能與眞圖逼似,亦屬無妨.此項圖形,簡易明曉,閱者一目瞭然,且易於比較,吾人可採用之.舉例如第八圖:



圖八　某大學某年度各省學生人數分布圖

V　組織系統圖　此項圖形,係表示某種機關或某種組織之系統,藉以顯明彼此間之類屬性,及彼此間之關係者也.各敎育行政機關,或各種學校組織,類多採用之.茲舉例如第九圖:



圖九 上海市市立萬竹小學兒童活動組織系統圖

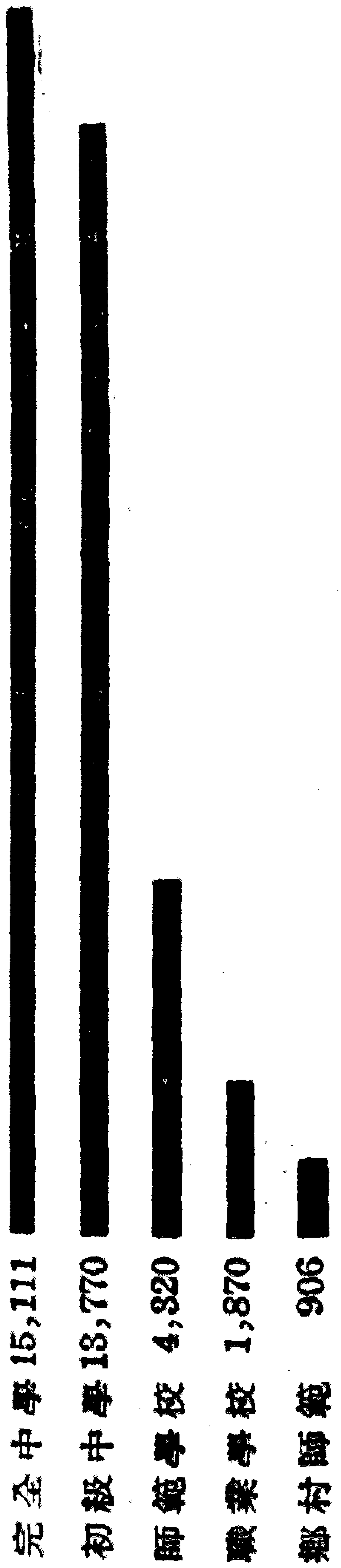
（摘錄二十一年一月該校出版之兒童活動中第九面上之圖）

VI　條形圖　此類圖形,係用平行之長方條形(或直線)製成,以各條之長短,比較各項事實之多寡.通常有橫條圖與直條圖之分.橫條圖之各長方條,係從一直立之基線,(普通繪畫時,此線可不必畫出.)自左向右伸長,(如遇兩面展開時,亦有自右向左伸長者.)該線代表零點,各條之價值,卽以從該線伸長之長度表示之.至縱條圖之各條,則皆直立作柱形,其下爲一水平底線,(等於橫條圖直立之基線)代表零點,各條之價值,卽以由此點向上伸長之長度表示之.如再從此橫條圖與縱條圖中再細分之,則又可別爲(1)單式條形圖,(2)複式條形圖,(3)分段條形圖.此三種式樣,橫直條均可.玆僅舉橫條之例,直條可依此類推.

1　單式條形圖　係以一條代表某事物之一項目,舉例如圖十.作此項圖形時,手續雖簡單,但有數點亦須加以注意:卽(a)條形宜稍寬,便易觀察.(b)各條相隔之距離,不可太狹.(c)各條以用黑色爲原則,若注重某一項目時,則該被注重之項目,得特別用他種顏色表示之.(d)如係橫條圖時,則圖之左端,宜先列出各該項目之文字,次在文字之右,再列各項目之數字,在數字之右,則畫各條形.若係直條圖時,得依此類推.惟換其方向及位置而已.(e)各條所代表各該項目之實

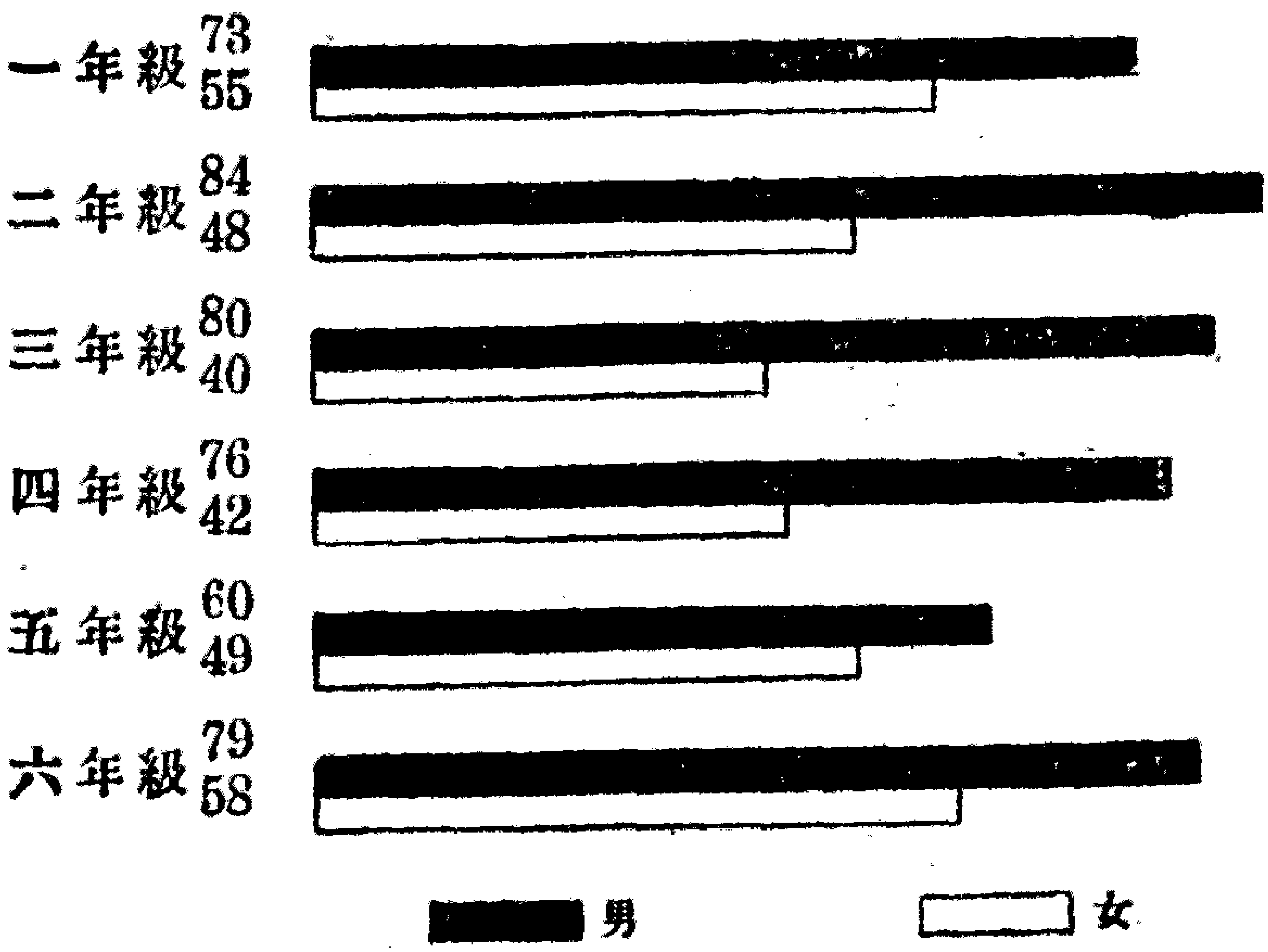
在數目,均宜詳細列出,以便作精確之比較,且其數字之位置,亦宜上下相對,與表列時相同.

2　複式條形圖　此類圖形,係以兩道或兩道以上之條形,並列爲一組,以比較兩種或兩種以上之事實.對於並列之各條形宜用顏色或影線(最好不用顏色)表明,以示區別.且應在圖下註明各條所表示的事實,例如圖十一之下,表明男女,及圖十二之下,表明操行之甲,乙,丙者是.

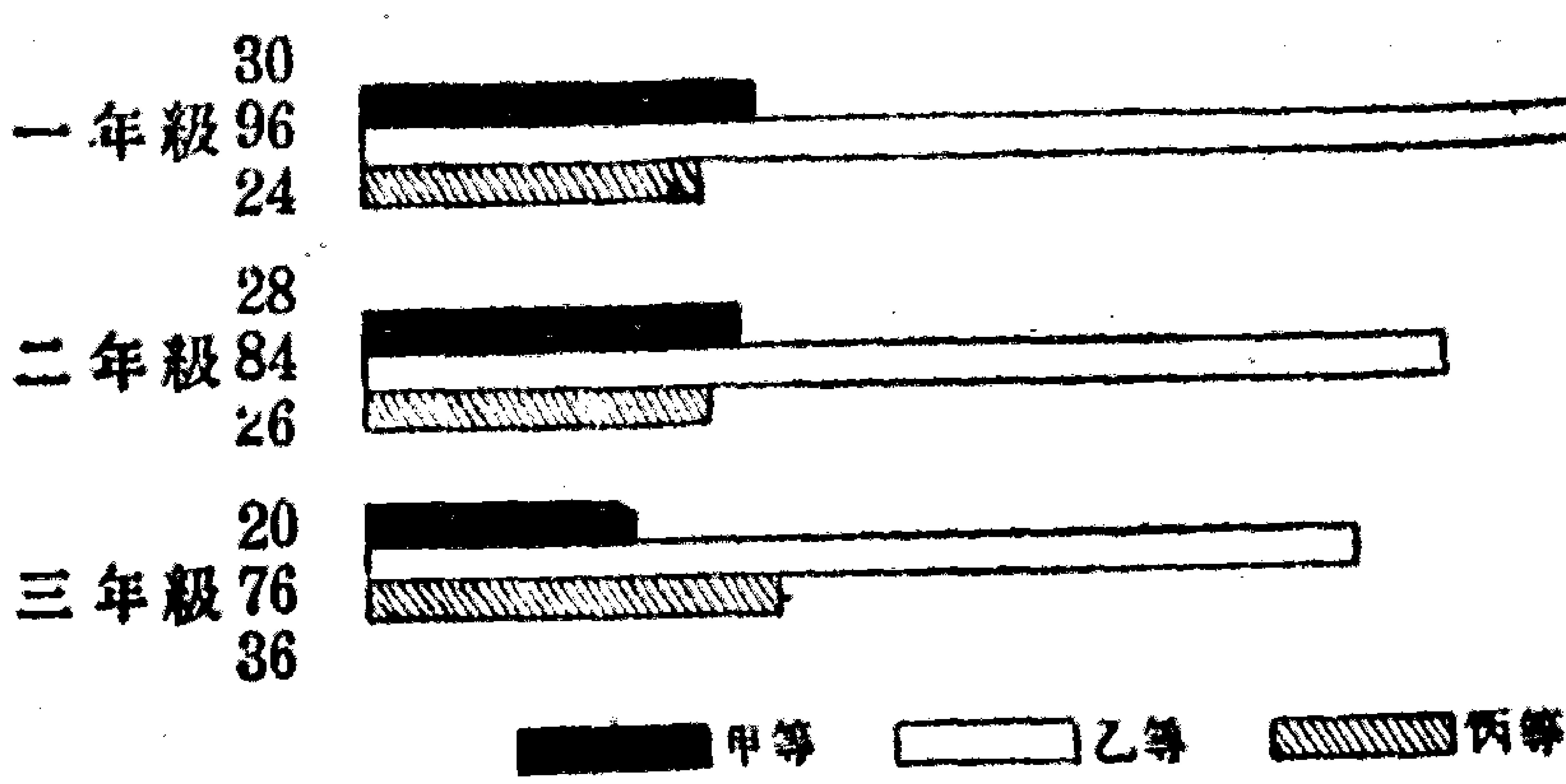


圖十　十九年度江蘇全省中等學校學生概數比較圖

（數字根據二十二年二月該校所出概況第二面上之表）



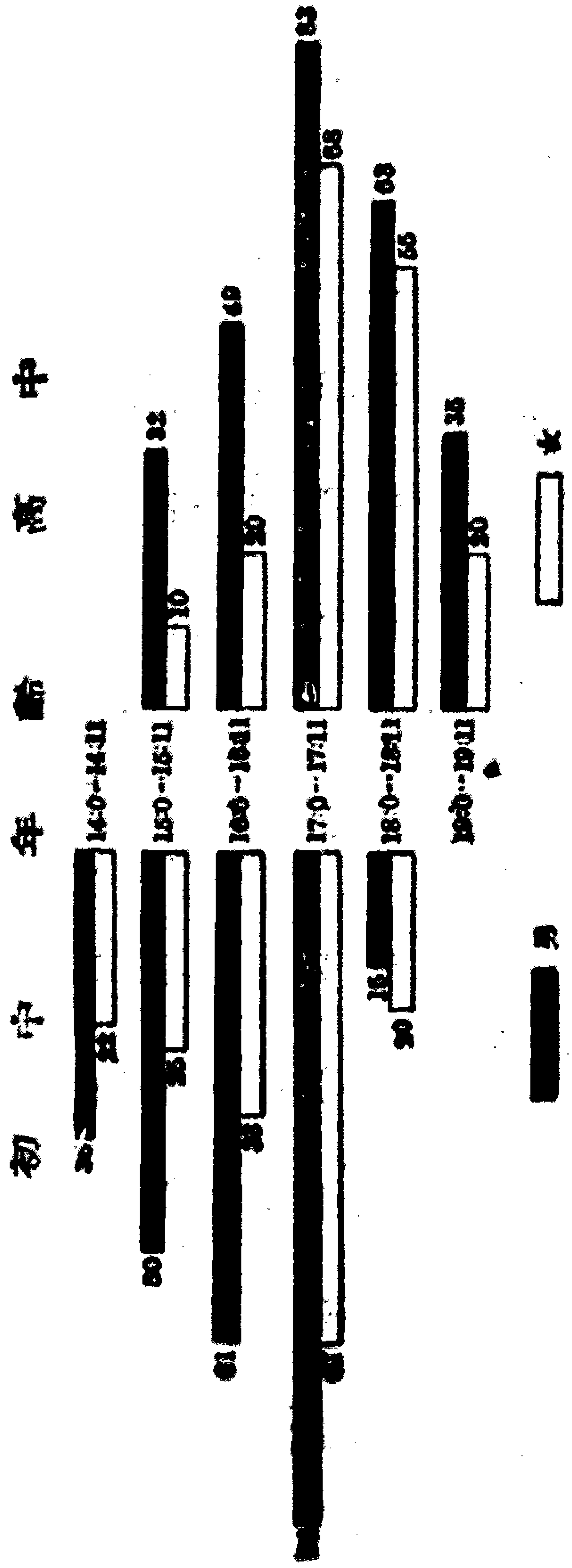
圖十一　上海市市立萬竹小學二十一年度下學期各級（上）男女生比較圖



圖十二　某初級中學某年度各級學生操行等第比較圖

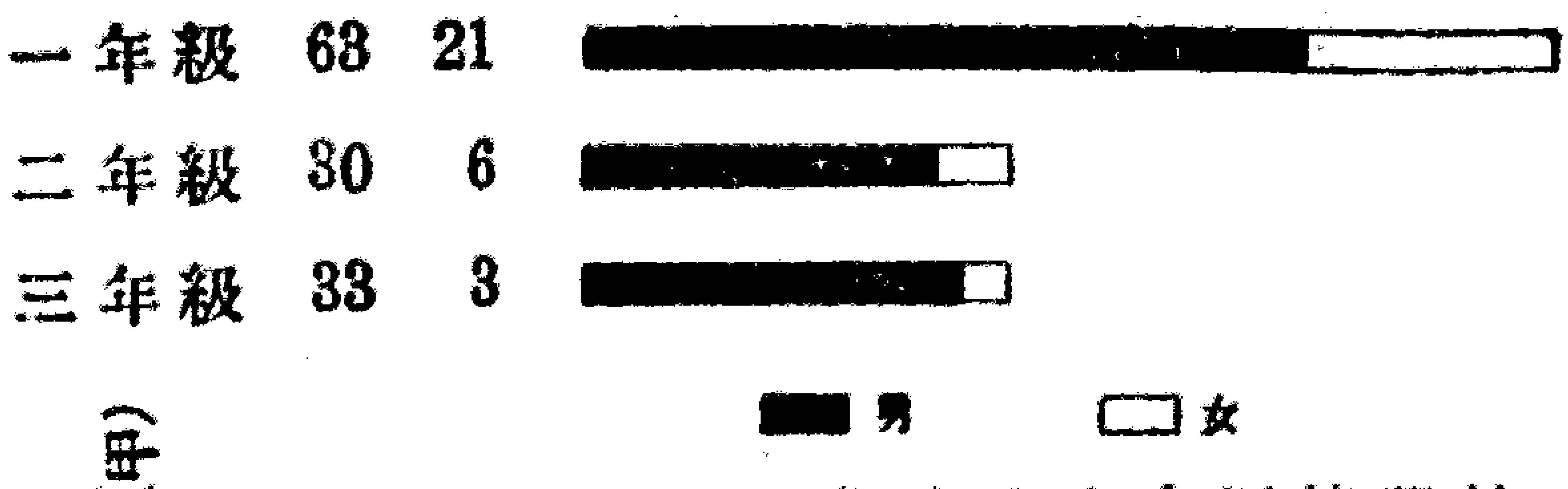
上列圖十一,係兩道條形圖.第十二圖,係三道條形圖.通常亦有四道條形圖,或四道以上者.惟在同組內條形太多,恐各條繪畫之區別及觀察時,兩皆感到不甚方便.故項數亦不宜過多.

上述之單式條形圖.及複式條形圖,均為自左向右展開者,普通亦有左右兩方展開者,名之曰對向條形圖.茲特舉例如左列圖十三,以便有時採用.



圖十三 某中學某年度在校各生之年齡比較圖

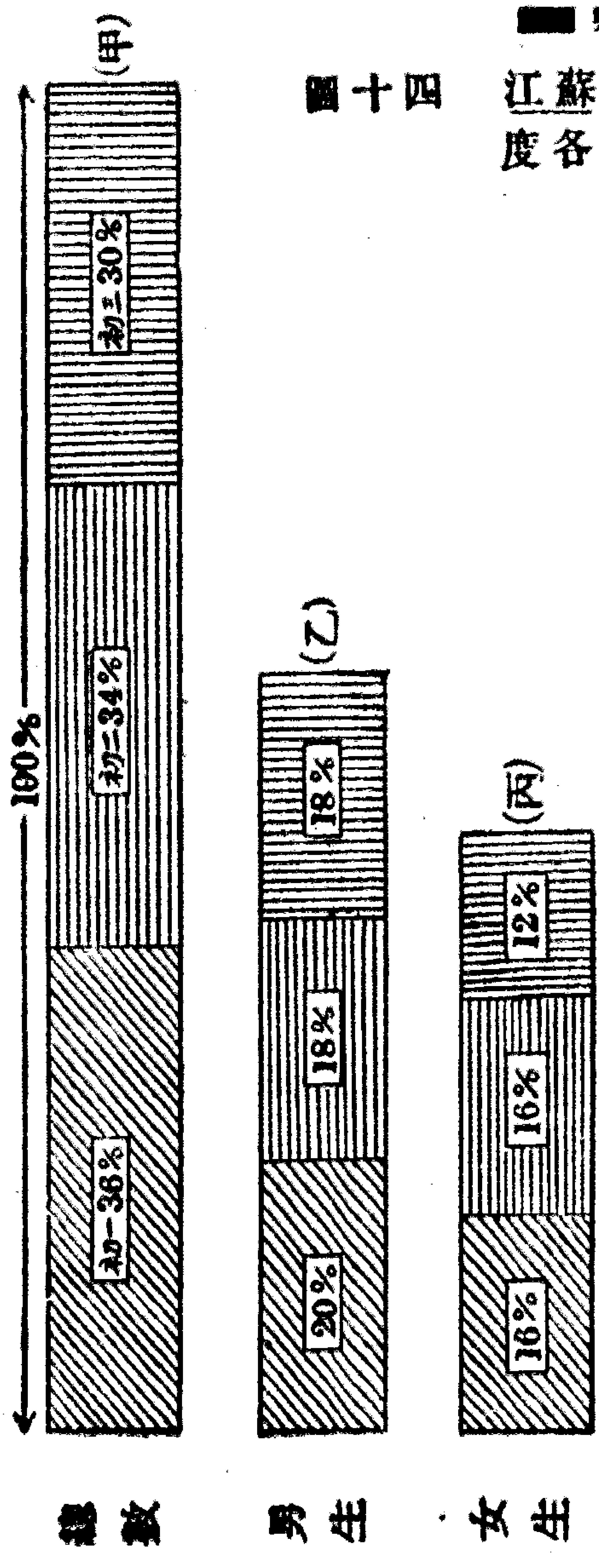
複式條形圖除上舉各例外,通常亦有如圖十四之畫法者.茲亦舉例如下列十四圖,俾便採用



圖十四 江蘇省立太倉師範學校二十年度各級男女生比較圖

3 分段條形圖

此類圖形,係以一條形,代表某事物之全部,再按全部之各部數值的比例,分此全條爲若干段,而每段則代表一部.此種圖形,又可分爲二類:(A)不分類的分段條形圖,(B)分類的分段條形圖.如左列圖十五之甲部,則屬於不分類者.如甲,



圖十五 某中學初中部某年度各級學生比較圖

乙,丙三部並用,即爲分類之條形圖.

Ⅷ　曲線圖　直線圖(或條形圖)表示事實,雖甚明瞭,但欲比較各事實增減變遷之趨勢,則直線圖又不若曲線圖矣.曲線圖爲研究各種科學不可少之良好工具,吾人宜熟習而常用之.

此種圖形,在教育上及心理上所最常用者,大概可分爲次數多邊圖,及次數面積圖.茲分述之如下:——

1　次數多邊圖　此類圖形,係以次數所由生之量數爲主體,用橫豎二量尺,分別表明量數之價值,及各種量數所發現之次數,然後按各量數之價值,以觀察其次數變遷之規律.茲舉例如第十六圖:

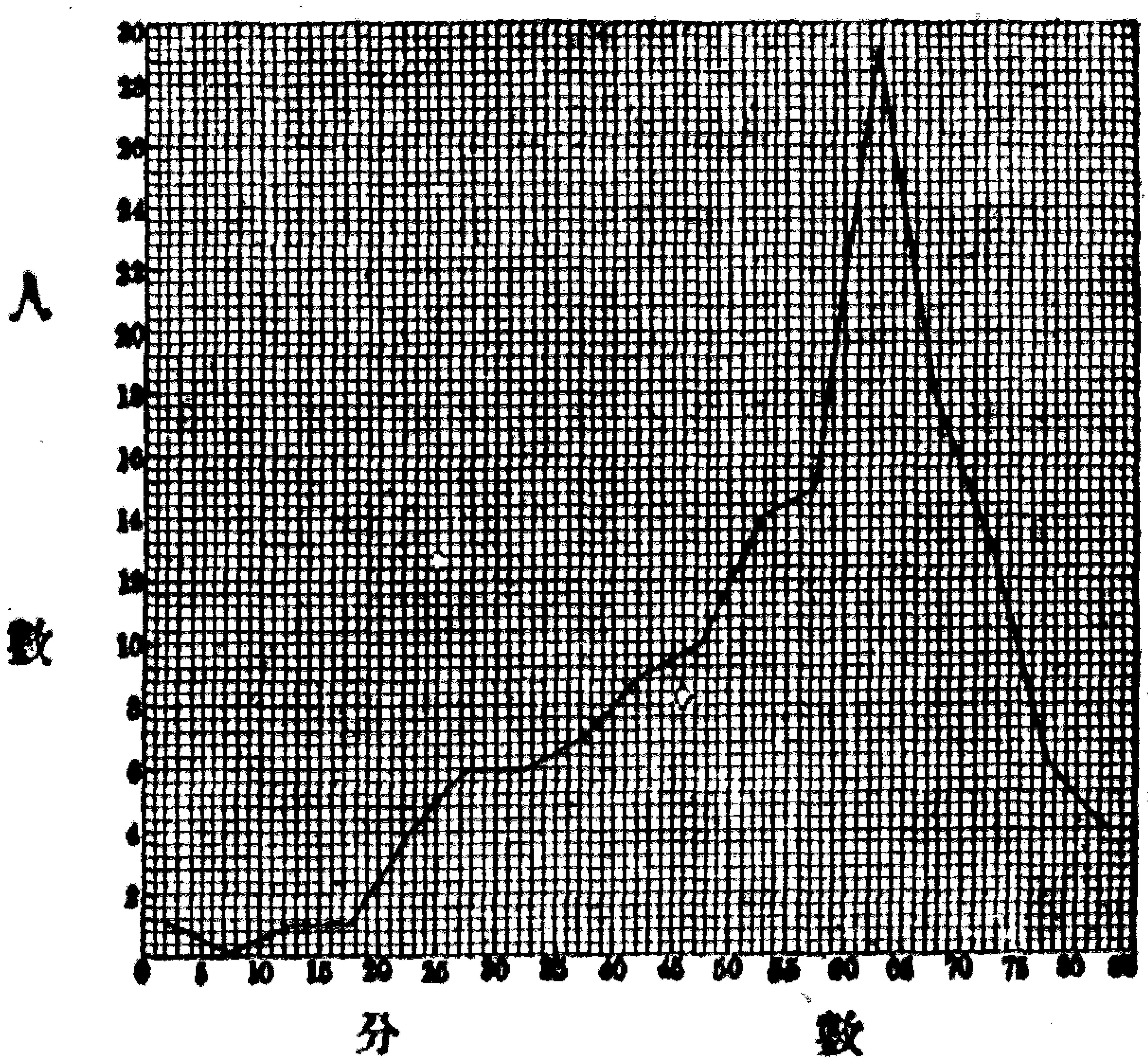
言其作法,則可分作五個步驟,茲分述之:

a　用精細之方格紙,而酌用其全面或一部.此種格紙,可由書坊或文具店購得.

b　作一橫坐標(或X軸,)分爲若干相等之單位,代表X變量,(如圖十六則代表分數)並在此坐標之下,記明各單位之數字,由小而大.

c　作一縱坐標(或Y軸,)亦分爲若干相等之單位,代表Y變量(如圖十六則代表人數.)惟此縱坐標,須與橫坐標在零點相交.此點通常稱之爲原始點(Origen.)其各單位數字之記法,亦由小而大,惟須

在坐標之左.



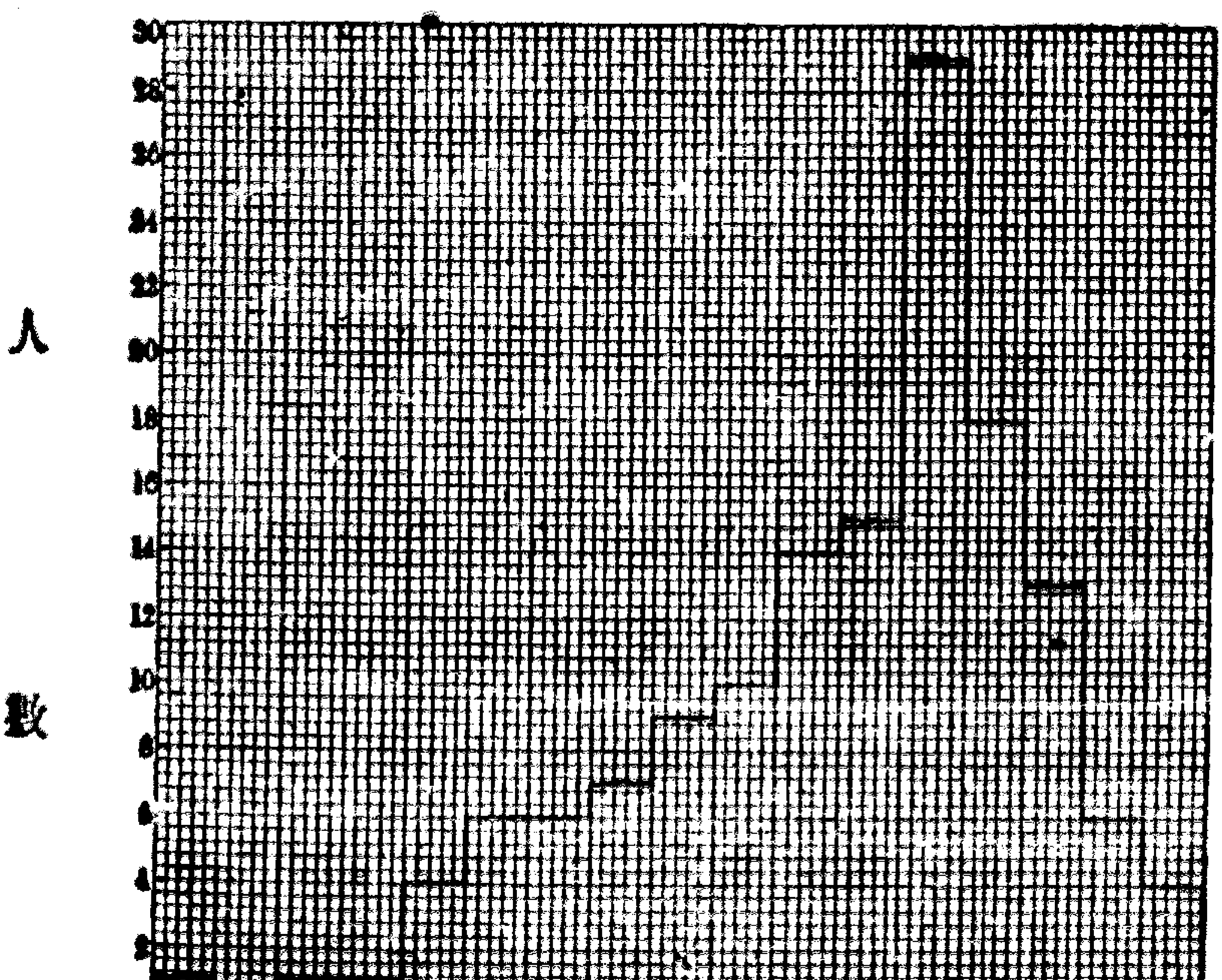
圖十六　某私立中學某年度入學試驗錄取新生算術分數之次數分配圖

d　在橫坐標 0—5 之組中,取其中點 (2·5),讀上至第一格,(代表次數 1,) 卽 0 分與 5 分之間,有一人,作一「 • 」爲記.餘皆依此作「 • 」,直至 80—85 的一組止.

e　將各點用線連之卽得.

作曲線圖之注意點及規則頗多,均詳下列作圖之規則中.有時在同一曲線圖內,亦有表示兩種以上的項目者,則可用虛線或其他斷線以分別之.

2　次數面積圖　此類圖形,亦稱直方圖.係用矩形之面積,表明次數之分配.矩形之底邊,與所分各組距離(即橫坐標上之各單位)之長度相等.矩形之高,與在該組間數量之次數相等.用此類圖形表顯事實時,往往在量數之距離長,次數之數目小,甚至或有間斷而不相連續之時.蓋取其較爲精確而明顯也.言其作法,則 a, b, c 均與上述次數多邊圖相同,d, e, 兩步,則不取其各組之中點,而將每點延長爲一橫線,其長度以所分之一距離爲限,至最後,再將各橫線用縱線連接之卽得.其例如第十七圖:



圖十七　根據十六圖之材料顯示次數面積圖之畫法

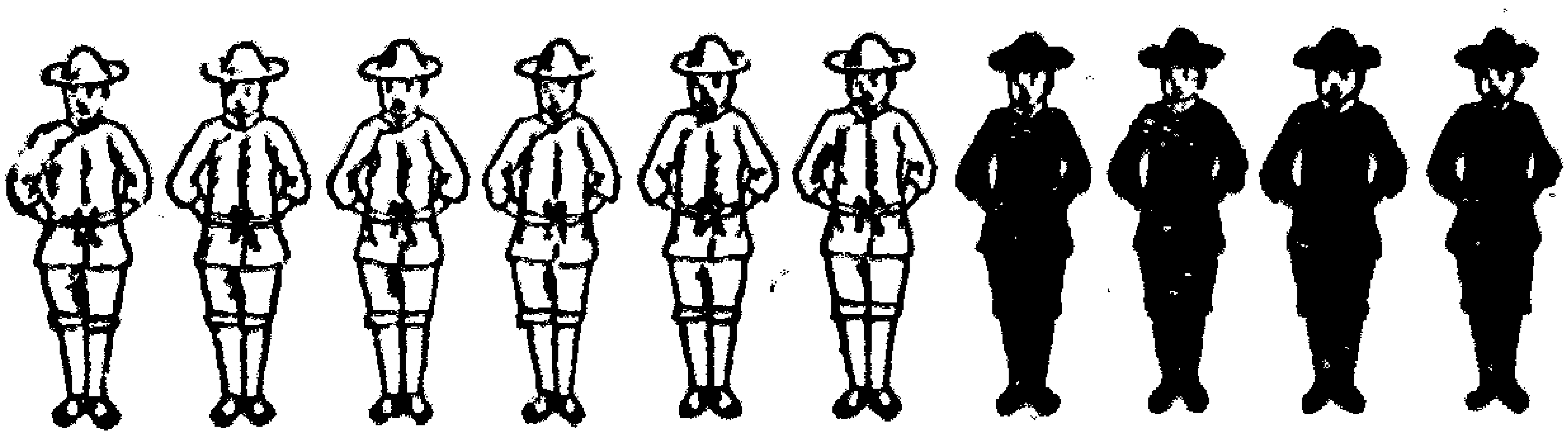
VIII 形像圖　形像圖者,即用圖畫或實物表顯某種觀念,或作某種比較.此類圖形,頗適用於通俗.茲舉例如下列十八及十九兩圖.

民國二十一年　　民國二十二年

183人　　262人

圖十八　某小學學生增加人數之比較

用上圖,以比較事實,不甚正確,究無正當統計之價值.故應用時,應加審慎,最好能避而不用.



農夫識字者 60 %

農婦識字者40％

圖十九　某鄉村農夫與農婦識字人數之比較

上圖黑色人形,係代表不識字者,白色人形,係代表識字者.上排係表示農夫,故畫男人像.下排係表示農婦,故畫女人像.此種圖形,極易了解,且亦無不正確之弊.故如欲用形像圖時,此圖可供採用.

凡是可以用形像圖表示之事實,均可以用直條圖表示.且其結果,有時亦不如用直條圖所表示者之顯明與精確.不過比較通俗,故亦有採用之價值.惟究無正當統計價值之可言,故採用時,宜詳加審慎.

上述圖形,共有八類之多.比較起來,以條形圖及曲線圖,最有價值,吾人宜常用之.至其他各類,並非全無價值之可言.惟宜審察情形,少用爲是.

13　作圖之規則

作圖之規則綦繁,似難一一詳述無遺.且亦依事實而變更,只好隨機應用.下列各條,係就白林東氏(Brinton, W.C.)所著之圖示事實法(Graphic methods for Preseting Facts)

書中所載圖形表顯法之規則,及美國圖示標準委員會(Joiut Committee in Standard for Graphic Presentation)所審定之標準,並參加最近各專家之意見,彙合而成.茲條述之如下:——

1 圖形之通常排法,宜自左至右.

2 能用線或條形代表事實之數量,爲最妥善.因用體積或面積,最容易發生誤解.

3 曲線圖上,最好能將零度線畫出.

4 倘若零度線不能照常畫出時,則零度線與其他橫線之間,應作波線以斷之.

5 波線不應將圖內表示數目之部分截斷.

6 若用兩條以上之曲線,或直條比較時,其零點務須相合,或在一條線上.

7 橫量尺上之數字,宜列於橫坐標之下.豎量尺上之數字,宜列於縱坐標之左.遇必要時,橫量尺亦得重見於圖之上端.豎量尺亦得重見於圖之右端.

8 若數字不能列於圖內之時,得列一表附於圖側.

9 若曲線圖以百分爲標準者,則凡百分線宜較闊大,其他用以比較之線,亦宜較闊.蓋皆所以示區別者也.

10 若圖表示年月者,則兩旁界線,不宜粗大,蓋因時間之起止,不能加以限制.

11　圖中縱橫引線,除必要外,不宜太多.

12　圖中曲線,宜與他線不同,以示區別.

13　圖上有時應載明所代表之數目或方程式.

14　圖題宜詳備明晰,如遇必要時,不妨多加說明.

15　圖題應放在圖之下端.

16　選擇圖式時,宜注意讀此圖者爲何種人,若讀此圖者,爲一般民衆,則當選用形像等通俗方式.

17　圖上之字,其位置及方向,應便於閱讀.

18　有時須標出圖例時,其位置應在圖中空白之處,否則應在圖下

19　圖例之大小,應與圖中所畫者一律,其次序亦應與圖內所指部分之次序符合.

20　若有多項事實在圖上表顯以作比較時,其中最重要之一種,得用紅色或粗直條,粗曲線及實線等表出之.

21　理想之圖,應能以最低限度之文字與數字,明白的表顯圖中事實之意義.

22　凡與圖形本身無關之美術化的邊緣,或其他不需之彩色,一概不用.

練習問題

1　試言圖示之意義及功用.

2 教育統計上表顯事實之良好工具,何以是圖示法?
3 圖形之種類,可分為幾種?並各舉一例以明之.
4 三角形圖、方形圖、圓形圖之缺點何在?
5 條形圖與曲線圖何以為圖示法中最良好之圖形?
6 組織圖與系統圖之用途若何?並何處宜用何者?
7 在何種情形之下,形像圖有採用之必要?
8 次數分配圖與次數面積圖(或直方圖)作法之不同點何在?
9 圖題為何應放在圖之下面?
10 圖中縱橫引線,何以不宜太多?
11 假使用曲線,表示圖中兩種以上的項目時,宜用何法以區別之?
12 若有多項事實在圖上表顯以作比較時,其中最重要之一種,須用何法將其特別顯出?
13 美術化的邊線,或其他之彩色,何以可一概不用?
14 圖中曲線,何以須比較粗闊?

第五章　全體數的統計

14　全體數的統計之重要

如把事實之全部,用統計方法來研究,來表示,並能使人一見而知全體之大概情形,則非得借重於全體數的統計不可.假使不先做這一步統計,那末以後一切別種統計上的計算,均無從進行.此項工作,爲統計上之第一步手續,無論如何,不得省略.全體數的統計之最重要者,得分（1）次數表,（2）分配圖.茲分述之:

15　次數表

吾人如先將有數目之事實,集合而統計之,則其第一步工作,當爲分類.次數表 (Frequency table) 卽係一種分類法在統計學上此爲初步.下列第八表中所列的數字,係某某私立中學二十年度入學試驗錄取的 144 個新生的算術分數.

表八　某私立中學二十年度入學試驗錄取的 144 個新生之算術分數.

48	64	64	60	41	48
61	71	80	74	56	68
70	26	36	74	60	49
71	52	57	60	40	67
48	63	64	53	58	66
60	72	84	28	64	68
45	3	36	27	44	49
12	34	57	54	58	67
18	31	60	60	65	68
50	56	83	74	55	40
61	63	38	75	43	41
72	79	57	63	65	46
20	78	63	52	42	45
52	64	39	30	65	84
62	57	56	32	67	74
73	32	61	53	66	80
23	35	25	33	43	73
53	58	53	54	66	76
63	64	62	63	68	67

24	77	74	77	47	51
54	64	61	62	68	59
64	59	53	55	45	53
70	39	29	40	69	60
22	37	28	56	68	69

表九　某私立中學二十年度入學試驗錄取的 144 個新生之算術分數之次數分配表

組　距	劃　　記	次數
0—4.9	/	1
5—9.9		0
10—14.9	/	1
15—19.9	/	1
20—24.9	////	4
25—29.9	卌 /	6
30—34.9	卌 /	6
35—39.9	卌 //	7
40—44.9	卌 ////	9
45—49.9	卌 卌	10
50—54.9	卌 卌 ////	14
55—59.9	卌 卌 卌	15
60—64.9	卌 卌 卌 卌 卌 ////	29
65—69.9	卌 卌 卌 ///	18
70—74.9	卌 卌 ///	13
75—79.9	卌 /	6
80—84.9	////	4
總	計	144

假使我們用次數表給他列成如第九表,則大體的情形一望卽知.例如:

1 分數在40與44·9之間者有9人.

2 分數在45與49·9之間者有10人.

3 分數在75與79·9之間者有6人.(餘皆可類推)

4 分數最少者在0—4·9之間計1人,最多者在80—84·9之間計4人.

5 最多人的分數,在60—64·9之間.

6 分數在55分以上者,共有85人,佔全數之59%.

7 分數在30分以下者,僅13人,佔全數之9%.

次數表之功用,吾人已略述如上,今進而言其作法.

I 作表 先作一表,並分表爲三行,第二行須特別放寬,以便畫記.第一行寫組距,或以i代之.(i爲class interval之縮寫.)第二行寫畫記,第三行寫次數,或以f代之.(f爲frequency之縮寫.)

II 定組距 組距者,卽將全量先分爲若干相等部分後之一部.蓋因全量之等級太多,計算費時.如上述之例,最少者之分數爲3分,最多者爲84分若無組距,則雖有歸納之方法,徒爲形式而已.欲定組距,須先求兩極差(Range).所謂兩極差者,卽各量數中,最大量與最小量之差.例如上例之兩極差,卽爲84—3=81.兩極差求得後

卽可進而决定組距.在决定組距的時候,還有幾點,吾人必須注意.玆分述之如下:——

（1） 組距之大小,可隨意决定.以5可,（上例卽以5）以3,以10,以20,以50,……均無不可.

（2） 决定大小時,有二原則必須遵守.此二原則爲何?第一:卽組距不能太多,亦不能太少.蓋太多,計算麻煩;太少,精確之程度亦隨之而減低.照盧克（H. O, Rugg）之主張,以爲最好在10至20之間.照其他學者之主張,以爲最好在15至25之間.反正相差不遠,吾人儘可隨意擇定.至其第二原則,則爲將斷定時,先須檢閱各量數之情形之如何.以各量數之與各組距之中點,愈近愈好.蓋採用組距的辦法,有一根本的假定,卽假定任何組距中所有之價値,均集中於其組距之中點,而以此中點作代表.例如上例自0—4·9,自5—5·9各爲一組距.此兩組距,前者可以其中點2·5作代表,而後者可以7·5作代表.因此,假使有七個分數如4,5,6,4,5,4,6,則以0—9·9（假使以10作組距）爲適宜,而以4—13·9爲不適宜.蓋前者與中點5相差最大者僅1,而後者與中點9,相差最大者爲5也.

明瞭上述二個原則後,則組距之大小,不難定奪.普通

定奪之方法,即用一個想定的數目,去除兩極差,使除得之商數,在10至20之間.(認定在10至20之間爲最好.)例如兩極差爲90,用想定之數目5去除,得商數18,此18適在10至20之間,甚爲適宜,故即以5爲組距.假定兩極差爲120,用想定之數目15去除,其商爲8,此8不在10至20之間,故以15作組距,爲不適宜.

還有一點,亦爲吾人須當注意.即組距之大小,應各組一律,不宜或大或小.

貳　組限　組限者,即組距之界限.定奪組距後,即須注意於組限.例如上舉之例,自0至84·9,以5爲組距,共計17組.假使吾人之寫法,如0—5,5—10,10—15,15—20,……等等,則對於不是逢5或逢10的數目之歸入,可絲毫不成問題,然若遇到5,10,15,20……等,則不免有猶豫之困難.例如甲生爲15分,將歸入10—15之組距內乎?抑將歸入15—20之組距內乎?欲避免此項含糊不決之困難,或覺而重複之弊,則組限之不可不注意矣.假使吾人將0—5寫作0—4·9,5—10寫作5—9·9,則上述之猶豫或重複,可全避免.其實此4·9及9·9,爲無限個·9,若再加·1,即爲5或10.故除非深有經驗者,決不可寫0—5,或5—10,以免錯誤.若嚴刻的講起來,則雖深有經驗者,亦不可輕於嘗試.

Ⅳ　晝記　定妥了組距及組限後,然後才可將各個量數,依次歸入.每歸入一量數時,最好在此量數上作一「✓」記號,以免重入或遺漏.並在畫記項下作「/」記號,至歸滿五個時,則作「𝍸」(參閱第九表,)以便總計.其實此𝍸,卽等於國人通常所寫之「正」字者一樣.惟統計學上多用𝍸,而不用「正」,蓋從習慣也.

Ⅴ　總計　表上之第三行爲次數.所謂次數者,卽各該組距內所發現之數目也.例如上述在 0—4·9 者爲 1,卽得 0—4·9 分有一人,在 80—84·9 者爲 4,卽得 80—84·9 分者有四人.將各組之次數相加,卽得總計,記於最後之一列.如上例爲 144. 做完了此五個步驟,則次數表成矣.

16　次數分配圖

有了次數表,再可作一次數分配圖.次數分配圖,通常得別爲次數多邊圖,與次數直方圖兩種.其實有了次數表,已能將散漫之事實,使吾人一望而知.不過僅有次數表,比較起來,仍不免有些麻煩.且對於統計無訓練者,大都不肯細心觀察.假使有了分配圖,卽諳於統計者,比較起來,亦能省時.且各種特性,不難於短時間內,一望而知,此卽所謂次數分配圖之功用.例如上列第十六圖,或第十七圖,(此二圖係根據上舉之次數表中之材料而晝成)吾人一望而知

爲:——

a　知分數之最小者在0—5之間.最高者在80—85之間.如以中點作代表,則前者爲2·5,後者爲82·5.

b　知最多人的分數在60—65之間.

c　知兩極差爲50.(85—0=80,或82·5—2·5=80.)

d　知此分配圖爲右偏,可推知試題太易,或此種私立中學招生如招兵,太不嚴格.蓋試題太易,則得高分數之人,一定較得低分數者爲多.(在此例內得50分以上者已可算得高分數.)如得高分者多,則圖之右端,當然是高.反之,若試題太難,則得低分者多.若得低分者多,則圖之左端,當然較高.若再就招生太不嚴格方面而言,則得0—5者,尙能錄取,未免太過濫收.此項推測,每甚精確,此可謂次數分配圖最著功效的一點.

e　知各個分數之變量,在5—10中間忽斷,以後則繼續不斷.

f　知此圖祇有一處高起,可名之曰單峯形.且能決定其粗略衆數(詳見本書第六章第20節中)爲62·5.

至於次數分配圖之作法,則吾人已詳述如上第四章第12節之Ⅷ曲線圖中,此處固無贅述之必要.

除此而外,尙有一點,亦爲吾人所當注意者.即凡是次數分配的情形,大概總是中間較多,兩邊較少.假使兩邊能完

全對稱,則卽名之曰常態分配.不過在事實上,要完全對稱者,絕難得到.至少總有些微之偏斜,故祇能稱之近似的常態而已.

練習問題

1 全體數的統計,爲何是必要?

2 作次數表時,爲何一定要採用組距之辦法?

3 組距之大小,如何定奪?並有否必須遵守之原則?

4 組距之界限,爲何一定要畫分清楚?

5 組距之根本假定爲何?

6 次數分配圖,有何最顯著之功效?

7 試以下列材料,作一次數表.作成後,並畫一次數多邊圖.

某中學 149 人的歷史分數

58	73	85	82	96	72
32	76	82	83	94	47
54	82	27	84	49	96
60	83	78	44	69	63
56	81	29	42	36	86
60	83	89	63	68	99
28	85	76	41	43	37
58	88	93	86	99	38

42	90	89	99	65	49
30	90	90	34	60	76
48	72	90	56	75	82
54	74	26	43	80	83
60	56	85	46	64	69
72	76	88	49	94	89
60	78	37	98	49	82
72	31	84	89	46	78
79	25	88	94	49	82
85	84	86	76	67	98
34	81	88	96	69	44
88	82	39	96	69	83
73	85	79	96	96	78
83	89	31	81	76	82
44	30	44	93	83	36
28	88	89	33	32	29
76	90	84	91	90	

第六章　集中數量

17　集中數量之意義

譬如施行同一種測驗於兩班不同的學生,欲知其成績之孰優孰劣,以及其程度之孰爲整齊,而作一精確的比較,則必須有一個可以代表此兩班全體成績之數量,然後才能比較其優劣,及整齊與參差.雖說是前章所述之次數表中,已能告訴我們最多的幾分,最少的幾分,或大部分人是幾分.然像這樣的告訴我們,我們倒反不十分清楚.假使能告訴我們這一班的平均數是多少,其中的差異是多少,那便易使我們徹底的瞭解,而且還覺得有一個比較一定的意義.在統計學上,表明兩組成績(或兩種以上不同的測驗)的優劣,類用集中數量.表明兩組程度的整齊與參差,常用差異數量.差異數量,將在下章另行詳述,本章則專論集中數量.

現在吾人所急欲知道的問題,就是集中數量的意義.欲

明瞭集中數量的意義,吾人須先知凡是任何事實,若集爲一羣時,必有多少及大小之不同.此種多少及大小,如在分配圖上看起來,是集中於某處.此種集中之現象,在統計學上稱之爲集中之趨勢.(Central tendency) 若用一最適中之數量,以表示其趨勢,或用一最適當之簡單數目,(Single number) 代表其全部之分配,則此數量,卽稱爲集中數量.故集中數量,在量尺上(Scale) 爲一點,(Point) 而此一點,卽能表明各量數在量尺中分配之位置.因此亦有稱集中數量爲點數量者.

通常用的集中數量,計有四種,卽:(1)算術平均數,(2)中數,(3)衆數,(4)四分點.在四分點之中,復可分爲上四分點及下四分點.玆分述之如下:——

18 算術平均數

I 何謂算術平均數 算術平均數在英文曰 Arithmetic mean, 卽我們通常所用之平均數,簡易明曉,人皆知之.卽把各個數量相加,得一總數,然後用個數(卽次數之總數)去除,其除得之商,卽爲算術平均數.然何以名之曰算術平均數,而不簡名之曰平均數呢?此則因統計學上表明集中趨勢之數量,亦有總稱之曰平均數者.在此種情形之下,則平均數爲廣義的平均數,包括中數,衆數,四分點等,而

非吾人通常所言之狹義平均數.算術平均數,爲表明集中趨勢所用之平均數之一種,且亦爲狹義平均數之一種.蓋狹義的平均數中,又可分爲算術平均數,倒數平均數,(Harmonic mean)幾何平均數,(Geometric mean)等三種.是則狹之又狹矣.至狹義平均數之後二者,在教育上,通常用之者甚少,且其算法亦不易瞭解,故初學者大多略而不提.茲將算術平均數討論之如下:——

II　算術平均數之求法　計分四法,分述之如下:——

1　第一法、量數之未歸類者　此法極形簡單,常人所用者,類多如是,固無舉實例之必要.如以公式表明之,則爲:——

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

式中之M,爲算術平均數.(Mean之縮寫)

Σ為總和之記號,讀Sum.

N爲各個之總數,或次數之總和.

X爲各量數之價值.

2　第二法　量數已歸類惟不分組距者　上述第一法,爲量數之未歸類者之算法.如量數一多,則一個一個的相加,而以其總次數相除,卽覺麻煩.設或相加時,有一數遺漏,或一數加錯,則非惟結果不正確,卽欲覆算時,亦必須從頭再來.且統計學上之第一步,每將次數歸類,

而作成次數表.(參閱第五章)故用第一法以求平均數,在次數多的時候,鮮有用之者.茲將次數已歸類者之算法,略述如下,例如下列第十表,已將量數歸類,但並未分有組距.

表十 江蘇省立太倉師範一年級學生教育心理臨時試驗分數之次數分配

分數 x	次數 f	次數×分數 fX
50	1	50
54	1	54
60	3	180
63	2	126
68	3	204
70	10	700
75	12	900
84	8	672
88	4	352
90	2	180
	N=46	3,418

$$M=\frac{3,418}{46}=74.30\text{强}$$

如將上列之算法,以公式表明之,則為:——

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

式中之M爲算術平均數.

f 爲次數,或人數.

x 爲量數,或分數.

Σfx 爲人數與分數相乘之積之總和.

N 爲總人數,或次數之總數.

3　第三法　量數已歸類且亦分組距者　平常所作之次數表,除非全距或兩極差甚小者,其組距單位每在 2 以上者,茲舉一實例,並述其計算法如下:——

表十一　顯示量數已歸類且亦分組距者之求算術平均數之算法

分數 i	中值 Sm	次數 f	次數×中位 fSm
30—34·9	32·5	2	65·0
35—39·9	37·5	4	150·0
40—44·9	42·5	11	467·5
45—49·9	47·5	16	760·0
50—54·9	52·5	20	1,050·0
55—59·9	57·5	27	1,552·5
60—64·9	62·5	28	1,750·0
65—69·9	67·5	24	1,620·0

70—74·9	72·5	18	1,305·0
75—79·9	77·5	13	1,007·5
80—84·9	82·5	12	990·0
85—89·9	87·5	5	437·5
		N=180	11,155·0

$$M=\frac{11155\cdot0}{180}=61\cdot97\ 强$$

上列之算法,如以公式表明之,則爲:——

$$M=\frac{\Sigma f Sm}{N}$$

在此式內,M爲算術平均數. Sm 爲各組分數之中值,亦卽組距之中點.蓋在上述第15節中,已論及統計學上對於組距之辦法,有一根本假定也.N爲次數之總數.fSm爲次數乘中值之積.其實此fSm,與上述第二法中之 fx 的意義,完全相同.蓋前者無組距之可言,故無中點或中值之假定.後者其組距單位爲5,故假定32·5,37·5,42·5,……等等,而以此32·5,37·5,42·5,代表30—34·9,35—39·9,40—44·9,……也.

4 第四法 用假定平均數以求得者 用上述第一法,求算術平均數之不便利,吾人已在第二法中論及之.然用第二第三兩法,亦未見如何便利,或者他的麻煩亦並不一定減於第一法.蓋因次數一多,或數目一大,要

逐一去乘,而且乘後還要相加,其手續已不見如何便利,假使其中次數甚大,或分數之數目亦甚多,或組距之中值帶有小數,則三四位的乘法,已覺麻煩.(例如組距爲120—124·9,次數爲285,組距中值爲122·5,以122·5×285=34,932·5.)若無計算機以代勞,則非特麻煩,抑且容易錯誤.故第一第二第三,三種方法,均不見有何種便利可言.本法爲求算術平均數之簡捷法,通常多用之.茲先述其公式如下:——

$$M = A.M. + \frac{\Sigma f d}{N} i$$

式中之M,爲眞正的算術平均數.A.M.爲假設的算術平均數.(assumed mean) f 爲次數.d (deviation) 爲各量數與假設平均數之差.i 爲組距之單位.N爲次數之總和.其算法如下列第十二表.

表十二　顯示用第四法(簡捷法)求算術平均數

組距 i	次數 f	差數 d	次數×差數 f d
30—34·9	2	—6	—12
35—39·9	4	—5	—20
40—44·9	11	—4	—44
45—49·9	16	—3	—48
50—59·4	20	—2	—40

55—59·9	27	—1	—27
60—64·9	28	0	—191
65—69·9	24	1	24
70—74·9	18	2	36
75—79·9	13	3	39
80—84·9	12	4	48
85—89·9	5	5	25
	N=180		+172

$$M=A.M.+\frac{\Sigma fd}{N}i$$

$$A.M.=62\cdot5$$

$$\Sigma fd=\Sigma fd_{正}+\Sigma fd_{負}=172+(-191)=-19$$

$$N=180$$

$$\frac{\Sigma fd}{N}i=\frac{-19}{180}\times5=-.52778\text{或}-.53$$

$$\therefore M=62\cdot5+(-\cdot53)=61\cdot97$$

試一比較第四法與第一、二、三，三法，則何簡何繁，不難一望而知.在第四法中，fd之積數，最大者爲48，此類可由心算而得.在第三法中，fsm之積，最大者爲1,750·0，其餘1,620·0，1,552·5等，亦復不鮮，且非用算具不可.再試比較Σfd及Σfsm之和數，則前者爲—19，而後者爲11,155·0，相差有六百倍之多.故第四法之簡而且便，似可予吾人之深信.茲再述

其求算之步驟如下:——

作　表

1　先作次數表,並求次數之總和.

2　假定假設平均數所在之組距,並書○於其組距上.如上例假定在 60—64·9 之組距,此組距之中點爲62·5,即假設平均數.其實此○,可任置於任何組距中.因反正有 Σfd 以校正之也.不過如果置○於最前之組距,或最後之組距,或不居中之組距,則前者 Σfd 正太大,而後者 Σfd 負亦太大.如 Σfd 太大,與 f 相乘時,及以N 去除時,還是麻煩.故通常均觀察大勢,每置○於最中之組距,或就次數集積最多的一組.(上下幾組亦無妨.)蓋取其可用代數方法,以正負相對消.

3　各組距無論其單位之大小,如上例雖以 5 爲單位,然均作「1」計算,蓋公式中 $M=A.M.+\frac{\Sigma fd}{N}i$ 之「i」,卽爲還原之意.

4　如明瞭上列 2,3,兩步以後,則 d 項之各數不難求得.如上例假定 60—64·9 之組距爲○,則 55—59·9 爲 −1,50—54·9 爲 −2,…… 65—69·9 爲 +1,70—74·9 爲 +2,……餘類推.蓋 55—54·9 之中值 57·5,較 60—64·9 之中值 62·5 小五單位,50—54·9 之中值 52·5,較 60—64·9 之中值 62·5 小十單位.上述第三步,無論其各組距單位之

大小,均作 1 計算.今本例既以 5 爲單位,則小五單位者,必爲 —1, 小十單位者,必爲 —2, 小十五單位者,必爲—3.(+1+2+3……等可類推.)

5 求fd的一項,簡而且易,即將次數乘差數(d)惟正負號必須記下,以免錯誤.

6 將$fd_{正}$及$fd_{負}$,分別相加,得$\Sigma fd_{正}$及$\Sigma fd_{負}$.再將此Σfd之正負相消,即得Σfd.本例$\Sigma fd=-19$.

7 以次數之總數(N),除—19,得—.10555.再以5(即組距之單位因假定此5單位作1,故在此時須還原.)乘之,得—•52778,或—•53.此—•53,通常稱之爲校正數,或更正數(Correction.)按此—•53之所由得,乃從公式中之$\frac{\Sigma fd}{N}i$而來,($\frac{\Sigma fd}{N}i=\frac{-19}{180}\times 5=-\cdot 53$.)故通常亦有將算術平均數之公式,寫作$M=A.M.+C$者.因$C=\frac{\Sigma fd}{N}i$,故$M=A.M.+\frac{\Sigma fd}{N}i$,亦可寫作$M=A.M.+C$.

8 最後之一步,即將校正數加於假設平均數之上.本例校正數爲—•53,假設平均數爲62•5.以$62\cdot 5+(-.53)=61\cdot 97$.*

* 在本例中,所求得之M爲 61•97.適與上列第三法者完全相同.平常用第三第四兩法求M,類能完全相同.惟有時亦略有不同者,不過相差不至影響而及於整數

1或2以上者.

19 中數

I 中數之意義 中數在英文曰Median,亦爲集中數量之一.言其意義,則卽爲若干量數,先經由小而大,依次排列後,其最中間的一數,卽謂中數.申言之,亦卽在此一點(集中數量在量尺上爲一點)之兩邊,各有全量之半數,或二分之一,亦卽百分之五十.例如下列第十三表之材料,其中數爲61·79.在此數上下之量數,各爲90.卽各有$\frac{1}{2}$的量數,大於或小於此中數.故中數之位置,恰在正中的一點.

II 中數之求法

第一法 從未分組距的量數中求中數 以此法求中數,甚爲簡單.祇要將各個量數由小而大,依次排列.排成後,擇其最中的一個卽得.例如下列各分數,係某小學三年級19人之算術T分數.

分數	70,	61,	42,	62,	40,	38,	59,	64,	72,	41,	48,	58,	54,	44,	39,	43,	67,	63,	54,
次數	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

今將各分數由小而大,排列如下:——

分數	38,	39,	40,	41,	42,	43,	44,	48,	50,	54,	55,	59,	61,	62,	63,	64,	67,	70,	72,
次數	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

上舉之例,其中數爲54.蓋在此由小而大之數列中,54

爲其最中的一數,比他大的有九個數目,比他小的亦有九個數目.如以公式表明之則爲:——

$$Md之位置=\frac{N+1}{2}$$

式中之Md爲中數.N=次數之總數.如上例次數之總數爲19,以 19+1, 再用2除,得10.此卽十九個量數中,從任何一端數起,數至第十個數目,就是中數所在的位置.在本例第十個量數之分數爲54,故此54,卽爲中數.但初學者切勿以此10爲中數之價值.

上擧之例,其數適爲奇數,而非偶數.若數目爲偶數時,則中數卽在兩個最中數目之中間.例如將上擧之數例,在各端最加一個學生的分數,假定爲73.則其中數卽爲$\frac{54+55}{2}=54\cdot5$. 這一點與公式 Md之位置$=\frac{N+1}{2}$亦甚符合.蓋加添一個,其N爲20,以$\frac{20+1}{N}=10\cdot5$.卽無論將任何一端數起,數至第十個及第十一個數目的中間,就是中數所在的位置.本例從左數右,(或自右數左亦可,)數至第十個數目時爲54分,第十一個數目爲55分.(若自右數左則第十數爲55分,第十一數爲54分.)今中數的位置,既在第十數與第十一數的中間,故中數卽爲$\frac{54+55}{2}=54\cdot5$.

第二法　從已分組距中之量數求中數　上列第一法,僅能適用於最少量之數目中.若量數之數目加多,則吾人將各量數由小而大的排列,即覺麻煩.且平常在未進行求算各集中數量之先,必列一次數表.茲根據上述第十二表之材料,求算之,如下列第十三表.

中數之位置,在量尺上正中的一點,吾人已言之如上既爲正中的一點,故求算時,自上求下可,自下求上亦可.茲先述其公式如下:——

自上求下者爲:——

$$Md = L + \frac{\frac{N}{2} - f}{f} i$$

自下求上者爲:——

$$Md = U - \frac{\frac{N}{2} - f}{f} i$$

式中之 Md 爲中數. L 爲含有中數之組距之下限, U 爲含有中數之組距之上限. N 爲次數之總和. f 爲含有中數之組距之次數. F 爲含有中數之組距以上或以下之次數之總和. i 爲組距之單位.其算法如下列第十三表.

表十三　顯示求中數之計算法

組距 i	次數 f	
30—34·9	2	
35—39·9	4	
40—44·9	11	
45—49·9	16	
50—54·9	20	
55—59·9	27	
60—64·9	28	←在此組距內含有中數 $\begin{cases}\text{有10個量數較中數低}\\\text{有18個量數較中數高}\end{cases}$
65—69·9	24	
70—74·9	18	
75—75·9	13	
80—84·9	12	
85—89·9	5	
	N=180	

自上算下

$$\frac{N}{2}=\frac{180}{2}=90$$

$$90-(2+4+11+16+20+27)=90-80=10$$

$$Md=60+\frac{10}{28}\times5=60+1\cdot79=61\cdot79$$

$$\frac{N}{2}=\frac{180}{2}=90$$

$$90-(5+12+13+18+24)=90-72=18$$

$$Md=65-\frac{18}{28}\times5=65-3\cdot21=61\cdot79$$

試觀上表,則知中數之求算,甚爲簡單.茲述其步驟如下:——

1　先作次數表.

2　求次數之總和(N),求得後,以2除之,得90.此卽總量數之二分之一,或一半.

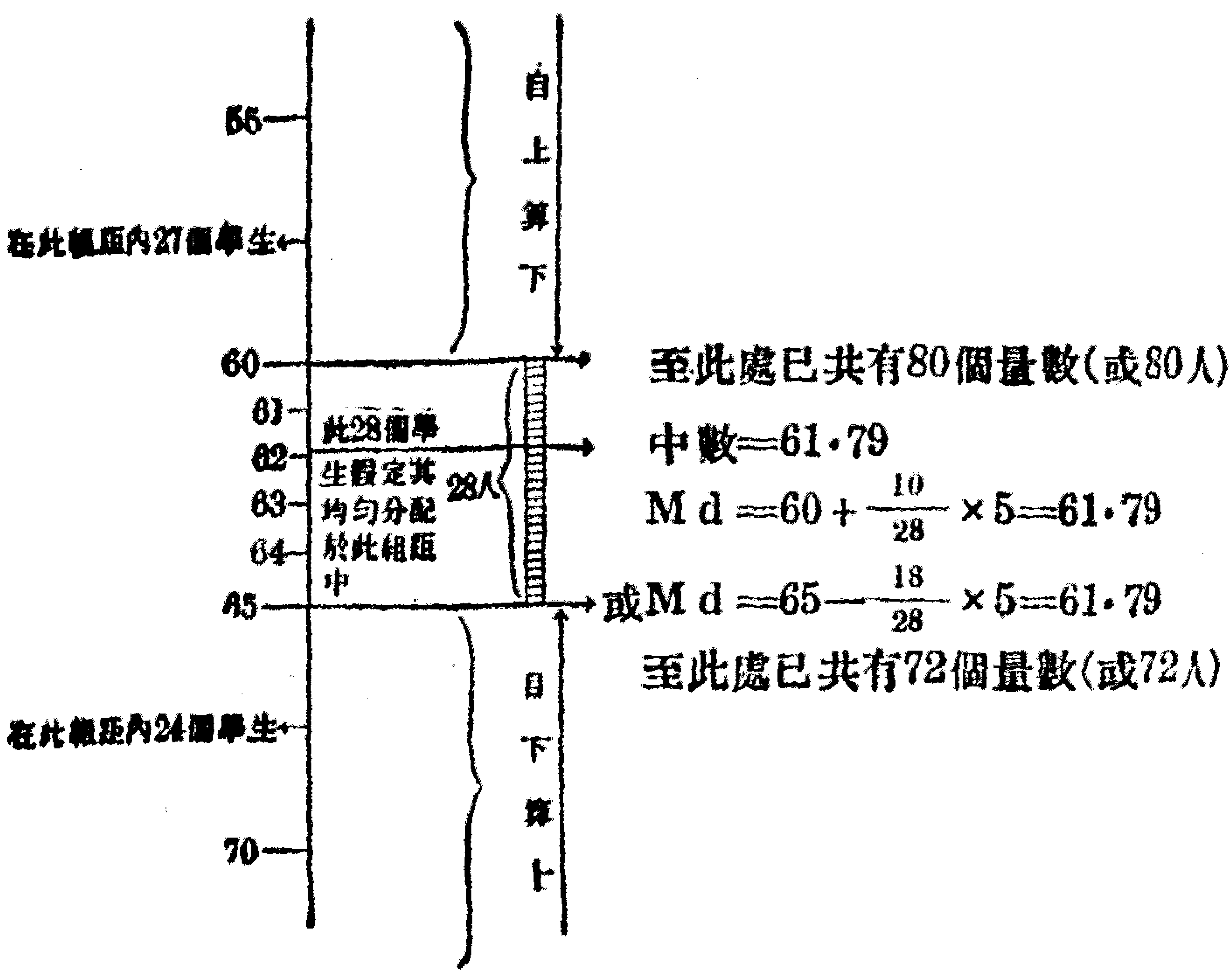
3　將次數之上端,(或從下端起亦可.)向下累加,(如從下端起,則向上累加.)至含有中數之一組距爲止.蓋中數爲最中間的一數,如累加時,超過$\frac{N}{2}$,則其位置卽非中數之位置,故累加時切不能超過$\frac{N}{2}$.本例不超過$\frac{N}{2}$的累加次數,得80,(自上算下)以F代之.

4　從$\frac{N}{2}$中,減去F,卽$\frac{N}{2}-f$.本例則爲90—80=10.

5　以60—64·9組距中之次數28,除10,再以組距單位5乘之,得$\frac{10}{28}\times5=1\cdot79$.蓋吾人所需要求的,是第九十個學生所得的分數.(卽第九十個量數之數値.)今至分數59·9之處,已有80個學生,則其餘的十個學生,必在60—64·9之組距中.在此組距中之次數(卽學生數)共28,則卽28人中應有10人,歸到中數之下端,以補其不足.假定此28個學生,平均分配於60—64 9之距度中,則此第十個人所在的位置,應爲$\frac{10}{28}$

$\times 5 = 1 \cdot 79$.（請參閱下列第二十圖）

6　將此1·79,加於60之上,得61·79,此即所求之中數.然爲什麽要加於60之上呢?蓋因至59·9分處,已爲第八十人所在之位置,則第八十一,八十二,八十三,……直至第九十人所在之位置必爲從60·179,60·358,60·537,而直至61·79.（如從下端向上算,則須從65減,其理亦可由此以類推.）



圖二十　說明中數之計算法

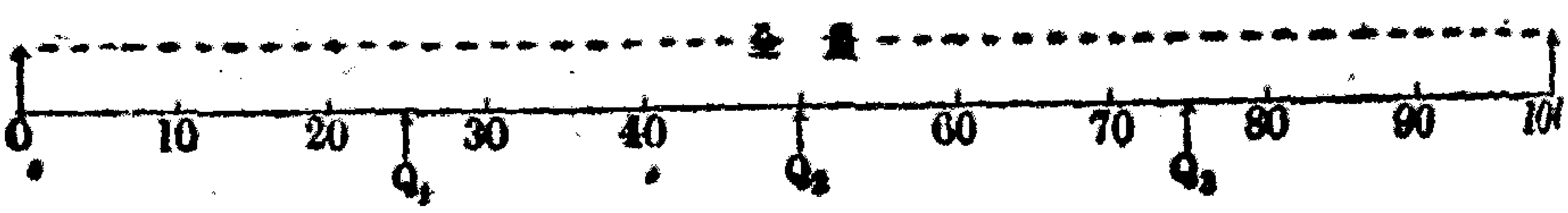
還有一點,吾人倘須加以注意.即自上算下,或自下算上,學者可隨意選擇.蓋無論從何端算起,其所得之

中數,必相一致,而無絲毫之差異.否則必陷於謬誤.在實際求算時,如既從此端求算,卽可不必再從他端去求算,藉省時間.學者如時間充足,或為精確而反證起見,則亦可先從此端求算,後再從他端去求算,以相互印證,知其是否有求錯之處.

20　上四分點及下四分點

I　上四分點及下四分點之意義　上四分點,亦稱七十五分點,又稱上二十五分點,在英文曰 Upper quatile, 普通以 Q_3 為符號.下四分點,亦稱二十五分點,又稱下二十五分點,在英文曰 Lower quatile, 普通以 Q_1 為符號.言其意義則不難從上述中數的一節中,推想而知.蓋吾人曾在前面說過,在中數之上下各有全體量數之一半,或二分之一,亦卽百分之五十.今上四分點既為七十五分點,則其位置必落在中數以上之量數之$\frac{1}{2}$的一點.故上四分點,在量尺上所居之位置,有全量之四分之一,在他以上,有全量之四分之三,在他以下.反之,則下四分點之位置,必落在中數以下之量數之$\frac{1}{2}$的一點.其在量尺上之位置,卽有全量之四分之一,在他以下,有全量之四分之三,在他以上.換言之,則卽將全量分為相等之四段,從最小的量數或最劣的成績數起,數至第一段終了處之量數之價值,卽為下四分點之

價值.再數至第二段終了處之量數之價值,即爲中數之價值.又數至第三段終了處之量數之價值,即爲上四分點之價值.因此有時亦稱中數爲五十分點,以 Q_2 代表之.而上四分點亦卽爲中數以上的全體量數之中數,下四分點亦卽中數以下之全體量數之中數.試觀下圖,卽益見其明顯矣.



(md)

圖二十一 顯示 Q_1 Q_2 及 Q_3 在量尺上所落之位置

II 上四分點及下四分點之計算法 言其求法,則與求中數之理相同.茲爲簡便計,僅述及一計算之實例如十四表,並順便述其公式.至其計算時之手續及步驟,則從略.

表十四 顯示求上四分點及下四分點之計算法

組距 i	次數 f
30—34·9	2
35—39·9	4
40—44·9	11
45—49·9	16

50—54·9	20→在此組距內含有Q_1	{有12個量數較Q_1低 有 8 個量數較Q_1高
55—59·9	27	
60—64·9	28	
65—69·9	24	
70—74·9	18→在此組距內含有Q_3	{有 3 個量數較Q2低 有15個量數較Q_3高
75—79·9	13	
80—84·9	12	
85—89·9	5	
	N=180	

$$\frac{N}{4}=\frac{180}{4}=45 \quad Q_1=\frac{1}{4}\times180=45 \quad Q_3=\frac{3}{4}\times180=135$$

$$45-(2+4+11+16)=45-33=12$$

$$Q_1=50+\frac{12}{20}\times5=50+3=53$$

$$135-(2+4+11+16+20+27+28+24)=135-132=3$$

$$Q_3=70+\frac{3}{18}\times5=70+\cdot83=70\cdot83$$

試觀上表,則其求算方法,全與中數相同.茲再述其公式如下:——

$$Q_1=L+\frac{\frac{N}{4}-f}{f}i$$

$$Q_3=L+\frac{\frac{3N}{4}-f}{f}i$$

在此式內,L為含有下四分點及上四分點之組距之下限.N為次數之總和.f為含有下四分點及上四分點之組距之次數.F為由起端(即自上算下)算至含有下四分點或上四分點之組距之次數之和.i為組距之單位.至其計算步驟則與算中數相同,並亦可自下算上.學者可由中數之求法中類推之,不贅分述.

21 衆數

I 衆數之意義 衆數亦曰範數,在英文曰 Mode. 言其意義,則亦簡易明曉.即量表中次數最多的一個量數之價值,或量表中次數最密積之處之量數之價值.若將次數分配畫成次數曲線圖或直方圖,則在圖上最長縱線或最高豎線所在之處之數量,即為衆數.其實衆數即時式之意,蓋凡物之時式者,用之者或好之者必多,用之或好之者多,其數必衆.故其意義,本甚簡明.

II 衆數之求法 平常分衆數為兩種:一為粗略衆數(Crude mode),二為理論衆數(Theoretical mode).前者亦稱觀察衆數,後者亦稱真確衆數.因其有此兩種的區別,故其求法,亦得分作兩種.

1 粗略衆數之求法 求粗略衆數,十分簡單.可毋須計算,祇要一看次數表裏或分配圖裏次數最多或次

數最密積之處的數量卽得.例如上列表十得75分者爲12人,卽75分發現十二次,其餘皆無如此之多,故此75分卽爲衆數.但初學者,不能誤此12爲衆數.如此之求法,爲量數之未分組距者.如已將量數分作組距者.則其求法亦相同.也祇要一看次數最多或次數最密積的那一個組距是什麽,然後再將此組距之中值,作爲衆數卽得.假使有時在同一量表上次數最多或次數密積的組距有兩個,或兩個以上,(請閱十五表之材料)那末祇要將此兩組距之中值相加,用2除之卽得.(如十五表之材料之衆數卽爲$\frac{67.5+72.5}{2}=70$).但此必須此兩組相互毗連者,如此兩組距相隔較遠,(請觀十六表之材料,)那末祇能將此兩組距之中值,各各寫出,而名之曰雙衆數.如第十六表之材料,一爲57·5,一爲77·5.但教育事實上很少發現雙衆數者.

表十五　九十四個中學生的歷史分數

分數	人數
40—44·9	3
45—49·9	2
50—54·9	4
55—59·9	8
60—64·9	12

65—69·9	18
70—74·9	18
75—79·9	14
80—84·9	6
85—89·9	4
90—94·9	3
95—99·9	2

表十六　七十八個中學生的算術分數

分數	人數
40—44·9	2
45—49·9	3
50—54·9	5
55—59·9	16
60—64·9	6
65—69·9	7
70—74·9	10
75—79·9	16
80—84·9	4
85—89·9	4
90—94·9	3

95—99.9	2

2　理論衆數之求法　求理論衆數,遠不如粗略衆數之簡易.其算法之繁複,恐非初學者所能涉及.對此統計學大家皮爾生(Karl Pearson)曾發明一個從經驗而得的公式.他的公式如下:——

$$mo=m-3(m-md)$$

在此式內,mo=理論衆數.m=算術平均數.md=中數.不過由此公式而求得之理論衆數,亦並非眞正的理論衆數,祇能稱之爲切近的或近似的理論衆數.且亦僅能適用於次數分配略有偏斜之時.如偏斜過甚,則此公式,即不能應用.蓋算術平均數,中數,衆數三者彼此之間,有其一定之關係也.教育事實之分配,雖不能絕對常態,然大多數卻能近乎常態,僅有稍帶偏態者.故皮氏之公式,吾人大可採用.祇要平均數及中數求出以後,一代即得,其簡易實不亞於粗略衆數之求得.茲將上列表十一之材料,用皮氏之公式,求之如下:——

十一表之平均數=61·97　　中數=61·79

$$mo=61·97-3(61·97-61·79)=61·97-3\times·18=61·97-·54=61·43$$

吾人試綜觀上列第十一表之材料,則算術平均數爲61.

97,中數為61.79,衆數為61.43,此三種集中數量,相差無幾.然第十一表之材料,其分配並不常態,故如分配常態,或完全對稱,則此三種數量,可完全集中於量尺上之同一點上.反過來講,那末如果分配不對稱,則此三種數量之各項,須逐一求出,以資比較,其理由及用意,亦可想見矣.而此三種數量之各有其特性,或各有其用處,亦可想見矣.

22 各種集中數量之性質及功用

吾人在上節中已述及,如果次數分配不對稱時,則各種集中數量,各有其特性及功用,必須逐一求出,以資比較之必要.茲將各集中數量之性質及功用,分述之如下:——

I 算術平均數

1 算術平均數,根據全部量數而求出.故比較中數及衆數,來得精確.

2 算術平均數之計算,甚為簡易,祇要用加法及除法卽可求得.若過於繁複時,亦祇要將各量數歸類(卽分組距)可用代數方法以計算,甚稱便利,並不若求中數或衆數時必須先排成次序,或畫分配圖等纔能求出.

3 有時要與兩極端量數作比較,或欲兩極端量

數對之發生影響,則惟算術平均數,有此功能.

4　算術平均數其意義簡易明曉,人所共知.應用時可毋待解釋.至於中數及衆數,則恐不能如此.

5　祇要知道量數的總數,與次數的總數,就是不知道各量數之次數之多少,亦能求出算術平均數.若中數及衆數,在此情形之下,恐亦無法計算.

6　若將所得之算術平均數,乘次數之總數,卽能得出各量數之總和.此點在中數與衆數,則不能.

7　在次數表上或次數分配圖上,或已依其大小排列成行的數列上,不能立卽指出算術平均數所在之地位.若中數與衆數,則較易指出,此爲算術平均數之不及者.

8　算術平均數恆易受兩極端量數之影響.如極端量數有變動,則算術平均數,立卽變動.其實大多數的情形之下,極端量數,儘可不必注重.故此亦爲算術平均數之不及中數及衆數者.

9　算術平均數,在不連續的量數上,常與事理不合,或常爲實際所沒有的數目.譬如計算火車輛數的平均數爲9.87,此9.87輛的一個數目,在實際上是不可能的.

II　中數

1 中數受極端量數之影響甚微,故固定而不受取樣之變動.假使極端量數不近情理之時,以中數作代表量數,最為適宜.

2 中數之指出量表上之位置,較平均數為易.

3 即不知道極端量數之價值,而只要知道極端量數之次數,亦能求出中數,此亦較算術平均數為優者.

4 所分組距之大小,影響於中數甚微,不若衆數之易受此種情形,而生變動

5 若量數愈多,或次數分配愈近常態之時,則中數之價值愈可靠.若分配完全對稱,則與算術平均數及衆數相同.

6 中數並非根據全部量數之價值而求出,其各數量之價值,祇有間接之關係,故其精確不及算術平均數.

7 若事實繁複,或量數過多時,雖於分組之後,亦不能用代數法計算,且其計算,亦不若算術平均數之簡易明瞭.

8 若欲與極端量數作比較,或欲對極端量數發生影響時,則中數殊不可能,此為不及算術平均數者.

9 以中數乘次數之總數,不能得各量數之總和

10　中數在不連續之量數上,亦常與事理不合,或亦常爲實際上所沒有的數目.譬如某城之學校,中數爲 48·34. 此 48·34 的數目,在實際上是不可能的.

Ⅲ　衆數

1　如欲免除極端量數之影響時,則以衆數作代表數,最爲適宜.蓋極端量數不能影響於衆數.譬如測量學生之智力,少數之天才或白癡之成績,能影響於算術平均數,但於衆數則無影響之可言.

2　粗略衆數,計算簡便,顯而易見.不若算術平均數與中數之必須用算學方法者.

3　求衆數可毋須向全部量數去求,既可不必知道極端量數之價值,又可不必顧到極端量數之次數.譬如欲求該班學生智力程度之衆數,可不必知道天才之成績,亦不必測量白癡之記錄.

4　衆數極易使人明曉,其意義.譬如某城小學教員之月薪爲四十元,大家都很容易懂得.若說月薪之算術平均數或中數是 41·23 元,恐聽者便不容易瞭解.因實際並沒有得月薪 41·23 元之教員也.至衆數之所代表者,亦爲最普通之情形.例如該班學生中有幾個天才生,其某種測驗之分數,都是 100 分.其餘的都是低能兒,測驗所得的分數,都是二三十分.假使用

平均數作代表,則該班之成績,皆屬中等.然若用衆數作代表,則卽能見得一班兒童之成績之情形.此可謂他種集中數量之所不能及者.

5　如欲與極端量數發生影響時,則以衆數作代表數,全不適用.

6　衆數極易隨所分組距之大小而變動,故其精確度不及算術平均數及中數.

7　理論衆數,計算煩難,不能使常人應用.

8　用少數齊一的量數,決定大多數不同的量數,此於理不通.例如某班學生今日用俞子夷算術應用題測驗第一類,得T分數90分者有五人,在次數爲最多.其餘所得之分數,人數皆各不相同,或相同亦無達五人者,又如明日用俞氏測驗第二類得0分者有六人,其餘亦皆不相同,而不達六人者.則在前者之粗略衆數爲90分,而後者爲0分.假使我們卽據此而曰某班用俞氏測驗第一類之分數爲90分,而用俞氏測驗第二類則爲0分,此豈非笑話.

9　衆數亦並不根據全部量數而求得,且其各量數之價值,連到間接的關係總很少,故其精確度,不及算術平均數與中數.

10　以衆數乘次數之總數,其總和不能與各量數

之總和相符合.

IV　上四分點及下四分點

1　上四分點及下四分點之特性及功用,大致與中數相仿.蓋中數爲百分之五十點,而下四分點則爲中數左半各量數中(或較中數小的一半的量數中)之中數.上四分點則爲中數右半之各量數中(或較中數大的一半的量數中)之中數.故就全分配觀之,則中數能代表二分之一之量數所佔之地位,而上下四分點則能代表四分之三及四分之一之量數所佔之地位.假使在常態的次數分配中,則中數已能作代表.然在不對稱或偏態之次數分配中,如僅用中數作代表數,恐必令人發生誤解.此種情形,可觀下列第十七表,卽能瞭然矣.

表十七　甲乙兩小學的六年級生默讀T分數之次數分配

分數	學生數	
	甲校	乙校
10—19·9	15	2
20—29·9	2	1
30—39·9	1	3

40—49·9	3	15
50—59·9	12	12
60—69·9	3	25
70—79·9	1	3
80—89·9	2	1
90—99·9	25	2
	N=64	N=64

甲校之 Q_1,md, 及 Q_3

$$Q_1=20+\frac{1}{2}\times 10=25\cdot 00$$

$$md=50+\frac{11}{12}\times 10=59\cdot 17$$

$$Q_3=90+\frac{9}{25}\times 10=93\cdot 6$$

乙校之 Q_1,md, 及 Q_3

$$Q_1=40+\frac{10}{15}\times 10=46\cdot 67$$

$$md=50+\frac{11}{12}\times 10=59\cdot 17$$

$$Q_3=60+\frac{15}{25}\times 10=66\cdot 00$$

試觀上列表十七之材料,則甲乙兩校之 md, 同等於 59·17.然就 Q_1 觀之,則甲校爲 25·00,乙校爲 46·67,

乙校高於甲校幾及一半.如再就 Q_3 觀之,則甲校爲93·60,乙校爲66·00.而乙校反低於甲校三分之一.如吾人僅以中數表示此兩校之默讀程度,則閱者必以爲此兩校完全相等.故在非常態分配中,除求中數外,應再就 Q_1 及 Q_3,以資比較.

練習問題

1　何謂集中趨勢?何謂集中數量?

2　通常用的集中數量有幾種?試列舉其名,並略述其意義.

3　試言算術平均數,中數,衆數及 Q_1,Q_3 之優點與劣點.

4　假使次數分配爲常態時,則算術平均數,中數,衆數三者之地位,是否集中於一點?

5　試以下列材料求 m, md, mo,（用 Pearson 之公式）及 Q_1 與 Q_3.——求算時各用最簡捷之方法.

i	f
20—24·9	2
25—29·9	1
30—34·9	3
35—39·9	6
40—44·9	8

45—49·9	10
50—54·9	14
55—59·9	20
60—64·9	28
65—69·9	18
70—74·9	16
75—79·9	14
80—84·9	9
85—89·9	3
90—94·9	6
95—99·9	2

第七章　差異數量

23　差異數量之意義及功用

若施行同一測驗於兩班不同的學生而欲知其成績之孰優孰劣,則吾人可用集中數量去表示,此已於前章述之.然集中數量在量表上爲一點,僅能表示其所居之位置之高下,(或成績之優劣)而不能表明其程度之整齊與參差,此可謂集中數量之所不及者.欲知程度之整齊與參差,在統計學上常用差異數量以表示.此種情形,吾人不難於下舉例中,一觀便知.

表十八　某小學三年級甲乙兩組學生之算術分數

甲組 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100.

乙組 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55.

	算術平均數	兩極差
甲組	55	100
乙組	55	10

試觀上列表十八之材料,則甲乙兩組之算術平均數,(集中數量之一)皆爲55,完全相等.而其兩極差,(離中數量之一)則甲組爲100,乙組爲10,相差適爲十倍.故如全部分配,僅以集中數量表示之,則必誤斷此兩組之優劣,完全相同.然實際上此兩組非但不相等,抑且大相逕庭.換言之,亦卽甲組良莠不齊,其成績之佳者滿格(Full mark),劣者得零(Zero mark).而乙組則頗皆整齊,既無極優者,又無極劣者.此種參差之情形,在統計學上稱之爲離中趨勢.若用一簡單數目以表明其趨勢或代表其全部之參差度,則卽謂之差異數量,或離中數量.故差異數量之重要,原不亞於集中數量也.

集中數量之所表示者,乃量數在量尺上分配之位置,或若干量數集中的性質.差異量數之所表示者乃量數在量尺上之距離,或其離中之程度.故集中數量在量尺上爲一點,而差異數量則爲一距離(Distance).假使各量數之分配愈集中,則卽同於或近於各集中數量之量數亦愈多,而集中數量之所代表者亦愈合.假使各量數之分配愈不集中,則卽異於集中數量之量數亦愈多,而集中數量之所代表者,亦愈不稱職.故如僅用集中數量表示全部分配,而不進求差異數量,或離中數量,則其斷定之陷於謬誤可必.因此差異數量不僅可以比較兩測量分配之情形,且亦能考

證集中數量所代表之價值.

通常用之差異數量,計分四種:(1)兩極差,(2)四分差,(3)平均差,(4)標準差.除上述四種以外,尚有比較的差異數量,亦爲普通所常用,且亦甚重要.茲將一併分述之如下:——

24 兩極差

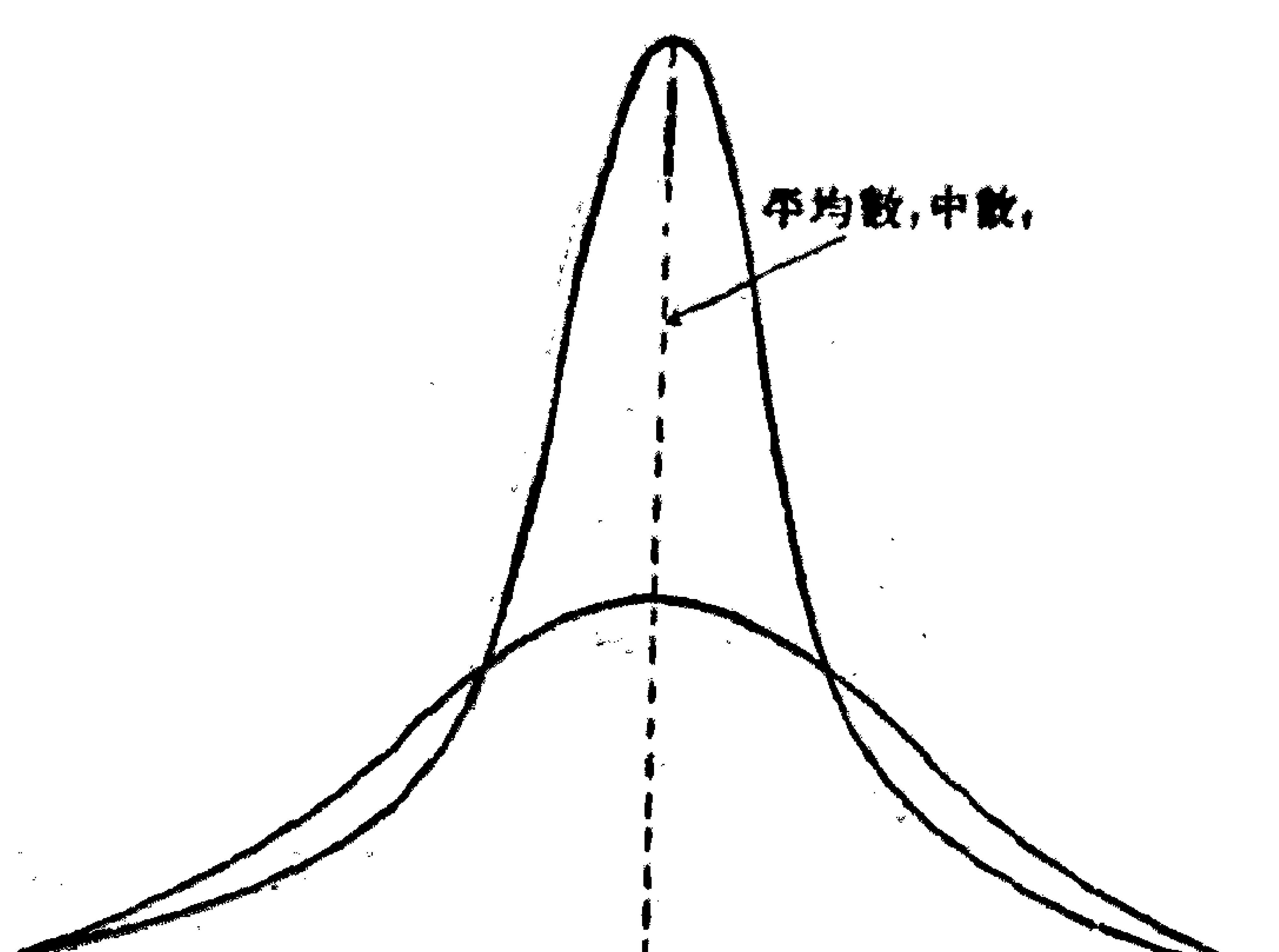
何謂兩極差 兩極差(Range)亦稱全距離,即全量中兩極端量數之差數,亦即次數分配中最大量數與最小量數之差數.言其所包含之量數,則佔全體量數百分之一百至其求法,則甚簡單,祇要將測量中之最大數減最小數即得.例如下列表十九之材料爲100—0=100.

兩極差雖爲表明離中趨勢之一種最簡單的數量,然用之者卻甚少.如嚴密的言之,則幾無功用之可言.至多僅能察其大概,作參考或檢查之用,亦正如集中數量中之粗略衆數.蓋每一測量中之所謂最大量數與最小量數所發現之次數,常有多寡之變動,如一變動,則兩極差之大小,亦必隨之而不同.例如下舉第十九表之材料,兩極差本爲100,(100—0=100)若將左端之數截去,則一變而爲60.(100—40=60)若將右端之數截去,則又變而爲70,(70—0=70).假使以左右兩端之數各各截去,則即變而爲30,(70—40=30).

其距離之變動既如此之大,故以少用爲宜.抑尤有進者,設有兩種測量,其每測量之各次數分配,除非偶而湊巧外,必不能完全相等.設或其兩極差相等,然其一測量中之各量數,集中之傾向較大,而他一測量中之各量數比較散漫,或集中之傾向較小.(請參閱第二十二圖)如此情形,在事實上常可遇到.假使在此種情形之下,如僅用兩極差以表示,則其內部分配之實情,必無從表出.故以兩極差表明離中趨勢爲最不精確的方法.自以少用爲宜,或亦可棄而不用.

表十九　某小學十八個三年級生之默字成績

0, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 100



圖二十二　顯示集中數量相同差異數量不同之兩常態面積

25　四分差

I　何謂四分差　四分差 (Quartile deviation) 亦稱二十五分差.非量數與集中數量之差,不過表示離中趨勢的一種數量而已.言其性質,則卽分配中下四分點 (Q_1) 與上四分點 (Q_3) 之間之距離之半數.而在 Q_1 以下,及 Q_3 以上之極端量數,置之不問.假使次數分配爲對稱時,則 md—Q_1,或 Q_3—md 之數量,卽可用作四分差之數量.且 md—Q_1,必等於 Q_3—md.此種情形,試觀下列第二十表之計算,當不難瞭然於胸.但實際之測量,極難有對稱的次數分配.故 md—Q_1,未必等於 Q_3—md.是以普通求四分差時,必須將 Q_3—Q_1 之數量折中.因其必須將 Q_3—Q_1 之數折中,故其公式卽爲:

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

在此公式中,Q 爲四分差之符號.Q_3 爲上四分點,Q_1 則爲下四分點.

表二十　某小學64人作法分數之次數分配

i	f
20—24·9	1
25—29·9	2
30—34·9	3

35—39·9	4
40—44·9	5
45—49·9	6
50—54·9	7
55—59·9	8
60—64·9	7
65—69·9	6
70—74·9	5
75—79·9	4
80—84·9	3
85—89·9	2
90—94·9	1
	N=64

$$Q_1 = 45 + \frac{1}{6} \times 5 = 45 \cdot 83$$

$$Q_2 \text{ or md} = 55 + \frac{4}{8} \times 5 = 57 \cdot 50$$

$$Q_3 = 65 + \frac{5}{6} \times 5 = 69 \cdot 17$$

$$Q_3 - md = 69 \cdot 17 - 57 \cdot 50 = 11 \cdot 67$$

$$md - Q_1 = 57 \cdot 50 - 45 \cdot 83 = 11 \cdot 67$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{69 \cdot 17 - 45 \cdot 83}{2} = 11 \cdot 67$$

II　四分差之求法　四分差之求法,甚爲簡單.蓋其公式中之Q_3及Q_1吾人在第六章第19節中,已詳述其求法.假使對於求Q_3及Q_1不生問題,則Q之求算,亦卽在其中.

III　四分差之優點及劣點

1　優點　四分差之優點,在乎意義明晰,計算簡便.卽常人亦多能瞭解.且在次數分配爲常態時,其值適與統計學中一極重要之單位名曰"機誤"(Probable error)者相等.至機誤之意義與用法,則將在下列第九章中另述之.惟機誤專用以表明常態分配之差度.至若敎育方面之測量,很難有常態完全對稱的分配.故用Q以表明四分度,最爲適宜.

2　劣點　四分差僅包含全體量數中間之一半,(卽Q_1至Q_3之間)對於Q_1以下及Q_3以上之兩極端量數則置之不問.故僅能適用於一般毋須用精細數目代表之測量.

26　平均差

I　何謂平均差　在23,24,兩節上所述的兩極差及四分差,其實都不是差異數量.祇不過用一種間接的方法,以觀察其密集之程度而已.蓋差異數量,亦名離中數量.所謂離中者,卽對於集中數量之距離或差異.上述兩極差及四

分差,其求算並未根據於任何的集中數量,其實都算不得集中數量.眞正根據集中數量,而測定其差異程度,通常用者,爲平均差與標準差.標準差之意義及求算,將於下節另詳之.本節則先述平均差.平均差在英文曰: Mean deviation, 簡寫爲M.D.,言其意義,則甚簡明.卽分配中各量數與集中數量(算術平均數中數或衆數)之差數之算術平均數.故此數對於全分配中之各量數,均負有直接之關係,不若兩極差與四分差之間接或局部的可比.惟各差數相加時,不能計其正負號,此爲本法中根本須注意的一點,亦爲本法中對於數理上說不過去的一點,或亦可謂本法的缺點.蓋若各差數相加而計其正負號,則正負相銷,結果爲零或不爲零,其所餘亦無幾.假使各差數相加等於零,則不能求差數之平均數,根本無所謂平均差.假使各差數因正負相銷而所餘無幾,則不能得正確之差數平均數.故各差數相加之時必須不計其正負號.計算平均差時,如按理論上講,用算術平均數、中數、衆數去求諸無不可.假使次數分配爲常態時,則均、中、衆三數皆集中於一點,當無問題之可言.不過在教育測量上,極難得有常態的分配,故用衆數去求平均差者很少.蓋因衆數多變動而不精確也.至於用平均數及中數去求者,則屢見不鮮.惟就術理而言,則由中點數求出之平均差數,其量最小.(此點艾偉先生曾做過一個實

驗,其原文發表在國立中央大學教育學院教育季刊第一卷第三號.題爲統計學上均差之研究.讀者如欲得其詳,可參閱其原文.）故平常多用中數去求.假使讀者一定要用算術平均數去求,亦未始不可.且其所求得之數,與用中數而求得者,相差亦甚微,惟務須註明用何數去求.不過在研究工作上,要用到平均差時,最好事前擇定一個.（或m或md）一致的用下去,切不宜此亦用,彼亦用,藉以減少比較上之困難.

II　平均差之求法　計分三種,今分述之如下:

第一法　由未分組距之量數求平均差　由未分組距之量數求平均差,甚爲簡單,祇要先求得各量數之算術平均數或中數,（本書之求平均差,均以中數爲根據）.然後將各量數減中數,得差數d (deveation,) 將差數再相加,不計正負號,用總次數N除之,即得.例如下列表二十一.

表二十一　顯示由未分組距之量數求平均差

量數	38,39,40,41,42,43,44,48,50,54,55,59,61,62,63,64,67,70,74.
差數 (d)	16,15,14,13,12,11,10, 6, 4, 0, 1, 5, 7, 8, 9, 10,13,16,20.

md=54	Σd=190
N=19	$M.D.=\frac{190}{19}=10$

上舉例中md爲54,N爲19,其差數(d),即由各量數與md之相減而得.(不計正負號)例如38—54=16,39—54=15,……72—54=20等等.其各差數相加,得190,以190用19(N)除,得10.如以公式表示之,則爲:

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

在此公式中,M.D.爲平均差之符號.Σd爲各差數之總和,N爲次數之總和.

第二法　由已分組距之量數求平均差　舉例如下列第二十二表.(根據上述第十一表之材料)言其公式則爲:

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

在此公式中除多一"f"外,完全與第一法之公式相同.其f則代表次數.

表二十二　顯示用第十一表之材料求平均差

i	m	f	d	f d
30—34·9	32·5	2	29·29	58·58
35—39·9	37·5	4	24·29	117·16

40—44·9	42·5	11	19·29	212·19
45—49·9	47·5	16	14·29	228·64
50—54·9	52·5	20	9·29	185·80
55—59·9	57·5	27	4·29	115·83
60—64·9	62·5	28	·71	19·88
65—69·9	67·5	24	5·71	137·04
70—74·9	72·5	18	10·71	192·78
75—79·9	77·5	13	15·71	204·23
80—84·9	82·5	12	20·71	248·52
85—89·9	87·5	5	25·71	128·55
		N=180		1849·20=Σfd

md=61·79

N=180

Σfd=1849·20

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

$$\therefore M.D.=\frac{1849\cdot20}{180}=10\cdot27$$

茲述其求算之步驟如下:——

1　先作次數表,並將組距之中值列出.

2　求次數之總和(N),得180.

3　照常法求中數,得61·79

4 求各組距中値與中數 61·97 之差數(d),得 29·29,24·29,……25·71,等等.

5 將各差數與各該次數相乘,得 fd.如 58·58,117·16……128·55,等等.

6 不計正負號,求差數之總和,得 Σfd.本例爲 1849·20.

7 以次數之總和 (N) 190,除 Σfd,卽 $\frac{1849\cdot 20}{180}=10\cdot 27$.此卽所求之平均差.

第三法 用簡捷法求平均差 用上列二法求平均差,意義雖甚簡明,然計算非常麻煩;尤其是在數目太多而太大的時候.故通常求算時,用之者甚少.本法比較簡單,猶如求算平均數的簡捷法一樣,用之者頗多.茲先述其公式如下:──

$$M.D.=\frac{\Sigma fd+c(N_b-N_a)}{N}\times i$$

式中 M.D. 爲平均差.Σfd 爲次數乘差數(與假設中數之差數)之總和.c 爲校正數 $=\frac{T.md-A.md}{i}$.卽由眞正中數 (True median) 減去假設中數 (Assumed median),以組距單位除.N_b 爲比眞正中數小的各次數之和;或眞正中數以下之各次數之和.N_a 爲比眞正中數大的各次數之和;或眞正中數以上之各次數之和.i 爲組距單位.N 爲次數之總和.其算法如下列第三十二表.

表二十三　顯示用簡捷法求平均差

i	f		d	fd
30—34·9	2	$N_b=80$	—6	—12
35—39·9	4		—5	—20
40—44·9	11		—4	—44
45—49·9	16		—3	—48
50—54·9	20		—2	—40
55—59·9	27		—1	—27
60—64·9	28	$N_a=100$	0	0
65—69·9	24		1	24
70—74·9	18		2	36
75—79·9	13		3	39
80—84·9	12		4	48
85—89·9	5		5	25
	N=180			$363=\Sigma fd$

眞正中數 (T.md.)=61·79

假設中數 (A.md.)=62·50

$$c=\frac{\text{T.md}-\text{A.md}}{i}=\frac{61\cdot79-62\cdot50}{5}=\frac{-\cdot71}{5}=-\cdot142$$

$(N_b-N_a)=80-100=-20$

$$\text{M.D.}=\frac{\Sigma fd+c(N_b-N_a)}{N}\times i$$

$$\therefore \text{M.D.} = \frac{363+(-\cdot142)(-20)}{180}\times5$$

$$= \frac{(363+2\cdot84)\times5}{180} = \frac{1829\cdot2}{180} = 10\cdot16$$

茲再述其求算之步驟如下:——

1° 先作次數表.

2 求次數之總和(N),得180.

3 依18節上求中數之方法,求眞正中數,得61·79.

4 用含有眞正中數之組距之中值,爲假設中數(簡寫爲 A.md.). 本例爲 62·50.

5 假設各組量數相差爲 1,求假設中數與各量數之差,得 d 欄之各數.

6 將 d 欄之各數,與其相對之次數相乘,得 fd 欄之各數.不計正負號相加,得 $\Sigma fd=363$.

7 求校正數(c),卽求眞正中數與假設中數相差之數.本例眞正中數爲 61·79,假設中數爲 62·50,故其相差之數爲61·79—62·50=—·71.然此差數,係按組距=5 計算.今本例中計算之單位,假設組距=1,故此 —·71 必須以 5 除之.以 —·71÷5=—·142.此—·142,卽爲每一量數之校正數.故校正數之公式,爲:

$$c = \frac{\text{T.md}-\text{A.md.}}{i}.$$

8　求量數之較眞正中數大者(N_a).本例爲5+12+13+18+24+28=100.最後之28亦算入,因該組之28個量數,均假設集中於此組距之中點(此本爲組距辦法之根本假定),其價値=62·50,較眞正中數61·79爲大也.

9　再求量數之較眞正中數小者(N_b).即 $N—N_a$,或180—100=80.

10　將$N_b—N_a$,即將80—100=—20再將從第七步中所求出之校正數—·142,與此差數相乘爲(—20)×(—·142)=28·4,此即爲總校正數.

11　將第十步中所求出之總校正數,加於差數之總數上,得363+2·84=365·84.此層計算之理,當甚簡明.蓋363爲由假設中數所求出之差數之總和,與由眞正中數所求出之差數相差爲+(—·142×80)—(—·142×100).此—·142爲校正數,80爲比眞正中數小的各次數之和,或眞正中數以下之各次數之和.100爲比眞正中數大的各次數之和,或眞正中數以上之各次數之和.今假設中數既與實得中數相差爲—·142,則在眞正中數以下之80個量數,比眞正中數小—·142,故必須+(—·142×80).又在眞正中數以上之100個量數,比眞正中數大—·142,故必須—(—·142×100).

若先由80—100,然後再與—·142相乘亦可.即 363+[(80—100)×(—·142)]=363+(—20×—·142)=363+2·84=365·84.

12 再將365·84用組距單位5乘之,(因此數係假定組距之數為1,故必須還原.)再用N除之,即得平均差$=\frac{365\cdot84\times5}{180}=\frac{1829\cdot2}{180}=10\cdot16.$

III 平均差之優點與劣點

1 優點

a 平均差係根據全體量數而求出,其對於各量數均能發生直接之影響.

b 平均差之意義,極易使人了解.

c 平均差之精確度,較優於兩極差及四分差.

2 劣點

a 求平均差時將正負符號,一概不問,此種計算,於數理上不免牽强.

b 在實際上求差異量數而應用平均差者較少,且其計算亦不見容易.

27 標準差

I 何謂標準差 標準差原名Standard deviation,普通常以希臘字母"σ"(讀若 Sigma)代之.言其意義,則即

各差數之平方之算術平均數之方根.故有時亦有譯作均方差者.在前節求平均差時,正負符號,一概不問,此種計算,不免牽强,吾人亦已言之如上.今標準差將差數平方,則負號自消,故在數理上毫無缺憾.惟初既平方,故結果不得不開方,以相抵銷,此固爲簡而易曉者.如再就計算手續而論,則標準差能用代數式而計算,非常簡捷.因其在數理上無欠缺,計算又簡捷,故用之者甚多.求標準差時,如照理論言,則凡集中數量,無論其爲算術平均數,中數,衆數,均可作爲根據.惟以數學上之理論而言,則欲求精密之差數時,須使差數之平方和爲最小,而求最小之差數平方和,須由算術平均數求各量數之差數.故通常計算標準差時,不用中數,衆數,而皆用算術平均數.

II　求法　標準差之求法,亦可分爲三種.茲特分述之如下:——

第一法　由未分組距之量數求標準差　由未分組距之量數求標準差,甚爲簡單.祇要先求各量數之算術平均數,然後將各量數與算術平均數相減,得差數 d,再將 d 開平方後,再加,得 Σfd^2,以總次數 N 除之,即得.例如下列第二十四表:——

表二十四　顯示由未分組距之量數求標準差

x	d	d^2
8	—19	361
12	—15	225
18	— 9	81
20	— 7	49
24	— 3	9
26	— 1	1
32	5	25
36	9	81
44	17	289
50	23	529
		$1650=\Sigma d^2$

$$m=\frac{270}{10}=27$$

$$\sigma=\sqrt{\frac{1650}{10}}=12.85$$

上舉之例,算法簡明,似無列舉其步驟之必要.如以公式表示之,則爲:

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

式中之"σ"爲標準差之符號.d 爲各量數與算術平均數之差數.d^2爲差數之平方.Σd^2 爲差數之平方之總和.N爲次數之總和.

第二法　由已分組距之量數求標準差　舉例如下列第二十五表.(根據上列第十一表之材料)言其公式,則爲:

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

在此式中,除多一"f"外,完全與第一法之公式相同,其 f 則爲次數.

表二十五　顯示由已分組距之量數求標準差

i	m	f	d	fd^2
30—34·9	32·5	2	29·47	1736·96
35—39·9	37·5	4	24·47	2395·12
40—44·9	42·5	11	19·47	4169·88
45—49·9	47·5	16	14·47	3350·08
50—54·9	52·5	20	9·47	1793·60
55—59·9	57·5	27	4·47	539·46
60—64·9	62·5	28	·53	7·84
65—69·9	67·5	24	5·53	733·92
70—74·9	72·5	18	10·53	1995·84
75—79·9	77·5	13	15·53	3135·34
80—84·9	82·5	12	20·53	5057·76

85—89·9	87·5	5	25·53	3258·90
				28179·70=Σfd²

m=61·97

$$\sigma=\sqrt{\frac{28179\cdot70}{180}}=12\cdot51$$

玆再述其求算之步驟如下:——

1　先作次數表,並將組距之中值列出.

2　求次數之總和 (N),得180.

3　照常法求算術平均數,得61·97.

4　求各組距中值與算術平均數之差數 (d),得—24·47,……25·53 等等.

5　將各差數 (d) 自乘,得d^2.再將d^2欄之各數,與其相等之次數相乘,得fd^2.爲1736·96,2395·12,……3258·90,等等.

6　求fd^2欄各數之和,得Σfd^2.本例Σfd^2=28179·70.

7　將$\Sigma fd^2\div N$,再開方卽得標準差.本例爲$\sqrt{\frac{28179\cdot70}{180}}=12\cdot52$.

第三法　用簡捷法求標準差　上述一二兩法,意義雖甚簡明,然遇數目大或次數多的時候,則計算異常麻煩,而尤其在沒有乘方表可檢查的時候.故實際上鮮有人用之者.通常之求算,類多用簡捷法.玆特述之如下:—

用簡捷法求標準差之公式爲:

$$\sigma=\sqrt{\frac{\Sigma fd^2}{N}-c^2}\times i$$

式中之σ爲標準差. Σfd^2 爲各量數與假設平均數(A.m.)之差數之平方之總和. c 爲校正數. i 爲組距單位. N 爲次數之總和. 茲用上列第十一表之材料演算如下列第二十六表:

表二十六　顯示用簡捷法求標準差

i	f	d	fd	fd^2
30—34·9	2	—6	—12	72
35—39·9	4	—5	—20	100
40—44·9	11	—4	—44	176
45—49·9	16	—3	—48	144
50—54·9	20	—2	—40	80
55—59·9	27	—1	—27	27
60—64·9	28	0	—191	0
65—69·9	24	1	24	24
70—74·9	18	2	36	72
75—79·9	13	3	39	117
80—84·9	12	4	48	192

85—89·9	5	5	25	125
	N=180		+172	$180\overline{)1129}=\Sigma fd^2$
				$6·272=S^2$

$N=180 \quad \Sigma fd=-191+172=-19 \quad \Sigma fd^2=1129$

$$c=\frac{\Sigma fd}{N}=\frac{-19}{180}=(-·1056)$$

$$c^2=(-·1056)^2=·011$$

$$\sigma=\sqrt{\frac{\Sigma fd^2}{N}-c^2}\times i$$

$$\therefore \sigma=\sqrt{\frac{1129}{180}-(-·1056)^2}\times 5$$

$$=\sqrt{6·272-·011}\times 5$$

$$=\sqrt{6·261}\times 5$$

$$=2·502\times 5=12·51$$

一觀上例,則知用簡捷法求標準差,其fd²一項的數目,非常之小.如再就Σfd²一項而論,則比用第二法演算者,要小到二十五倍之多.且用第二法以演算者,帶有小數,而用本法演算者,則絕無小數.故其簡易與麻煩,殆不可同日而語.至於他所以簡易的理由,則與求算術平均數時一樣,似可不再加解釋.茲將其求算之步驟,分述之如下:——

1 先作次數表,並求出N=180.

2　擇定一組距,將此組距之中值爲假設平均數,並置〇點於其上.本例則擇定 60—64·9 的一組,其中值則爲 62·5.關於擇定組距時,照理論上言,最好能在含有眞正平均數的一組距內.其所以如此之理由,則與求平均差時之必須擇定含有眞正中數的一組距者相同.但實際演算時,眞正平均數並不需用.若僅爲擇定組距起見,而另求一眞正平均數則殊不值得,故普通每選擇量數集聚最多的一組.不過有一點須特別注意,就是不能像求平均數時可任置〇點於何處,而不礙結果.以愈近於眞正平均數者爲愈佳.若相差太遠,則每影響於所求之結果之不正確.

3　假定組距之單位爲 1,求假設平均數與各組距中值之差數 d.如本例中之 $-1, -2, -3, \cdots\cdots +1, +2, +3, \cdots\cdots$ 等等.

4　將各次數與其相對之差數相乘,得 fd.如本例中之 $2\times-6=-12$, $4\times-5=-20$, $\cdots\cdots 5\times5=25$, 等等.

5　用代數法相加,求次數與差數相乘之積之和,得 Σfd. 如本例中 $\Sigma fd+=172$, $\Sigma fd-=191$, $+172-191=-19$.

6　將次數總數 N 除 Σfd, 得校正數 c.如本例 $N=180$, $\Sigma fd=-19$, $c=\frac{\Sigma fd}{N}=-\cdot1056$.

7　將每一fd各以其相對之d乘之,得 fd^2. 如例中之第一列fd為—12,其相對之d為 —6,得72.其實卽將d乘方,然後以f乘之.如本例中第一列d為—6,方之得36,以36×2=72.

8　求 fd^2 之總和.本例 $\Sigma fd^2=1129$.

9　用次數總數N除 Σfd^2. 本例則為 $\frac{1129}{180}=6{\cdot}272$.此6·272以 S^2 代之.此 S^2 為由假設平均數所求出之標準差之平方.但由假設平均數所求出之差數之平均數,除非偶而湊巧,必不能與由眞正平均數所求出之差數之平均數相同或無誤.其不同之數,卽為正負兩類相差之數之平均數,故必須以c數校正之.若假設平均數適與眞正平均數相等,則正負兩類相差之數必為○,卽無須校正.然本例之校正數為—·1056,故由假設平均數所求出之標準差之平方,與由眞正平均數所求出之標準差之平方,相差為 $(-{\cdot}1056)^2={\cdot}011$.此·011,亦卽所謂 c^2 也.

10　由假設平均數所求出之標準差之平方,減去校正數 c^2,得 σ^2.故 σ^2 卽等 S^2-c^2. 故標準差之公式亦可寫作 $\sigma=\sqrt{S^2-c^2}\times i$.在本例 S^2-c^2 卽等於 $6{\cdot}272-{\cdot}011=6{\cdot}261$.

11　將 σ^2 開方,卽得 σ.本例 $\sigma=\sqrt{6{\cdot}272-{\cdot}011}=$

2·502. 然此係假設組距之單位為 1,而計算者,故必須用組距之實際單位相乘而還原.

12　用原來之組距單位"5",乘所求得之標準差,得實際之標準差.本例則為 2·502×5=12·51.

III　標準差之特點　標準差為各種差異數量中之最佳者.在教育統計上用之者甚多.蓋以其精確之程度,較大於其他的差異數量,而其計算又不甚難之故也.言其特點,約可分為下列數點:——

1　根據全體量數而求得,與各量數皆有直接之關係.

2　可用代數法計算,故簡而且易.

3　求算時,不若平均差之對數理上有所欠缺.

4　受取樣之變動之影響甚小.

5　標準差之計算方法,對於求皮爾生 (Pearson) 之相關係數時,幫助甚多.且計算可靠度時,亦須應用之.(請參閱下列第九章)標準測驗中之T分數,(Total ability) 及通常小學適用之「S」分數,亦均從標準差值 (S. D. Value) 而求得.故標準差無論在統計上,測驗上,應用者均甚多.

28　比較的差異數量

I 何謂比較的差異數量 比較的差異數量,亦稱相對的或相關的差異數量.在英文則曰relative Vairability.言其意義,則卽求各測量或各事實之差數之比率,而比較其彼此間相互之差度.吾人在以前所敘述過的差異數量,(如四分差,平均差,標準差等.)都是絕對的,而非相對的.假使對於這兩種測量或兩種事實,如僅就各測量或事實之絕對差數,而直接比較,則其所得之結果,必難可靠.此因各測量之單位有所不同,及其所得之集中數量之大小,亦不能盡等之故.例如將學生之操行成績之標準差,與其身長之標準差相比較,則其不能比較或無意義之可言,顯然而知.此無他,單位不同之故也.又如以一百個大學敎授之薪金之標準差,與一百個小學敎員之薪金之標準差相比較,則其不能比較,或無甚意義之可言,亦顯然而知.此無他,其集中數量大小懸殊之故也.故欲比較兩種測量之離中趨勢,有時不能用其絕對的差數,而須用其相對的差數,方稱正確而可靠.對此皮爾生(Pearson)曾立定一個公式,以求其比較的差異數量.此種比較的差異數量,通常亦稱之爲差異係數(Cofficient of variation).玆轉錄其公式如下:——

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

式中之V,卽代表差異係數,M爲平均數,σ爲標準差.

II 功用 言其功用,則可舉一實例說明之.例如某中

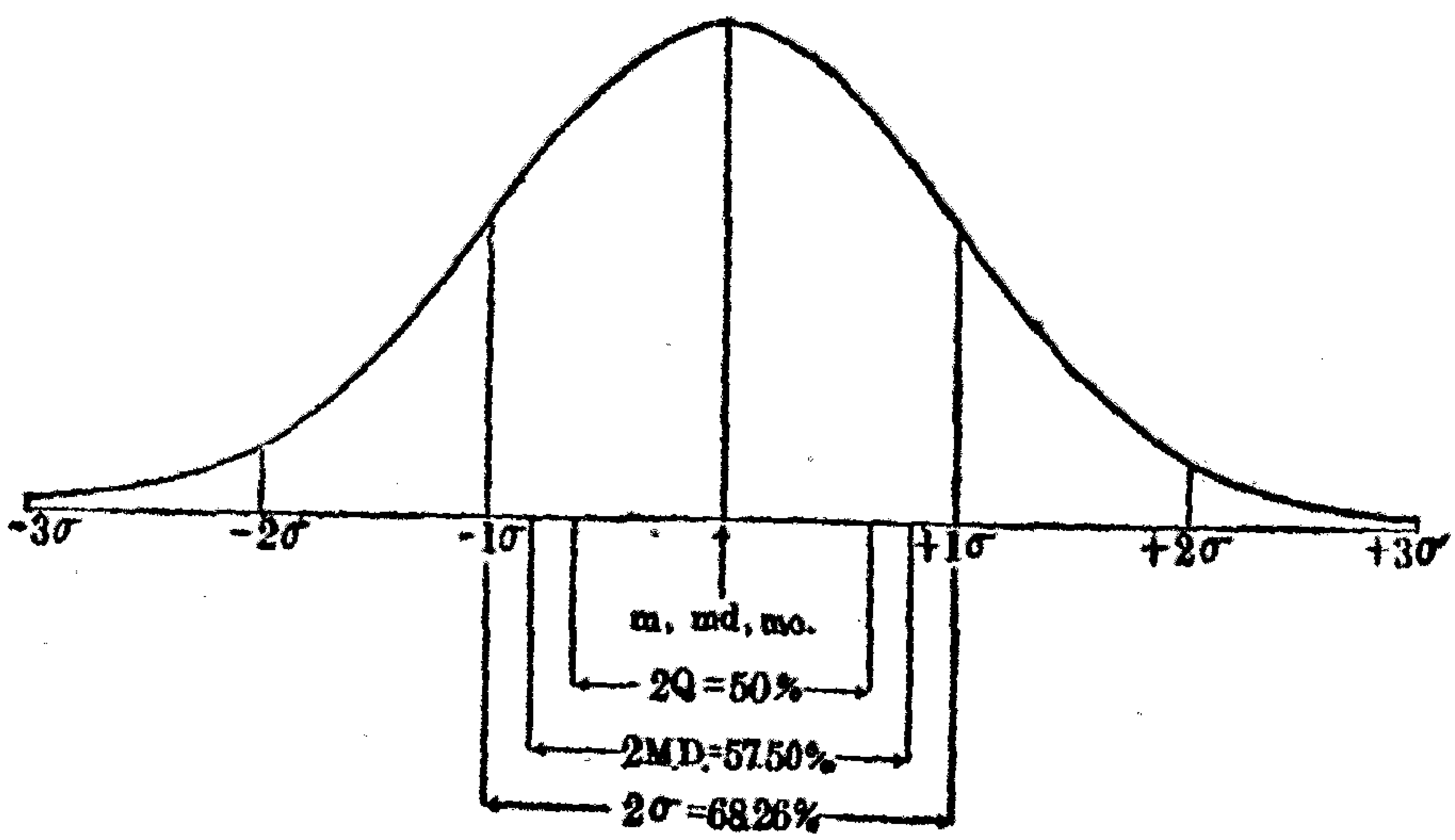
學招考時,其錄取各生之國文成績之算術平均數爲 72·14, 標準差爲 12·46. 其投考各生之國文算術平均數爲 48·20, 標準差爲 16·21. 若以皮氏之公式求其比較的差數,則前者爲 $V=\frac{100\times12\cdot46}{72\cdot14}=17\cdot27$. 而後者則爲 $V=\frac{100\times16\cdot21}{48\cdot20}=33\cdot63$. 有了此種係數,吾人可知投考各生之差數,實幾二倍於錄取各生之差數.（相差 16·36 或等於50％）換言之,亦卽錄取各生之國文程度,較投考各生之國文程度爲整齊二倍.假如像此種事實,僅比較其絕對的差數,則前者之標準差 12·46, 後者之標準差 16·21, 相差僅3·55, 或等於 10·5%, 然其實則二者相差甚遠.

又如某校女生每年之小食費,平均數爲$12·14, 標準差爲 $4·94. 每年之書籍,文具用品費平均數爲 $48·34, 標準差爲 $19·89. 在此例中,如就平均數而言,一則不過十二元,一則有四十八元之多.驟視之,似不能比較,且其事實又各不同.但是我們如果拿比較的差數來算,則前者爲 $V=\frac{100\times4\cdot94}{12\cdot14}=40\cdot70$. 後者爲 $V=\frac{100\times19\cdot89}{48\cdot34}=41\cdot15$. 可知此兩種之相差,僅些微之數,可以拿來比較.故在此種事實上,比較的差異數量的應用,非常重要,此卽所謂比較差數之功用.不過此種數量,是一個純粹的數目,並無單位.蓋因此數之表示,僅不過是差異的相關.故有時亦稱作相關的差數,

意卽在此.

29 各種差異數量之關係及其換算之公式

吾人在上述各節中,已將幾種重要的差異數量,敍述一過,茲特略言其間之關係.由經驗之所知,若次數分配爲常態或略形偏態時,則平均差,約當標準差五分之四,而四分差則約當標準差三分之二.又就經驗所得之結果,標準差之六倍大約包括全體量數百分之 99•73. 故若由此關係以推算,則七倍半平均差之距離,或九倍四分差之距離亦包含全體量數百分之 99•73. 惟以上各種之關係,僅能在數量較多時符合之,若數量過少之時,則不能相符,此爲吾人所當注意者.如再就常態曲線上所佔之面積而言,則正負各一個四分差,等於全面積50•%.正負各一個平均差,等於全面積57•50%.正負各一個標準差,等於 68•26 %. 至其正負第二個Q或M.D.或σ,因距離圖上之高峯較遠,故所佔的面積,當亦較少.又第三個或第四個 Qm.D.或σ,距離高峯又遠,故其所佔之面積,亦當更少也.(請參閱下列第二十三圖)



圖二十三 顯示正負各一個σ,M.D.,及Q,在常態曲線下所佔之面積

至若次數分配,完全對稱時,則其間之理論關係如下:—

$\sigma=1{\cdot}2533$M.D.　　$\sigma=1{\cdot}4826$Q 或 P.E.*

M.D.$=0{\cdot}7979\sigma$　　M.D.$=1{\cdot}1483$Q 或 P.E.

Q 或 P.E.$=\cdot 6745\sigma$　　Q或 P.E.$=\cdot 84535$M.D.

* 在次數分配爲常態時,Q與P.E.完全相等,故Q亦可稱P.E.,此吾人已於前述24節中言及,請複閱可也.

有了以上各種的關係,吾人有時若須用多種差異數量時,即可依此而換算,所謂換算者,即祇要知其一數,而他數即可從此而推算,毋須一一去另求.惟須注意者,就是要在次數分配爲常態的時候,方能符合.

練習問題

1 何謂離中趨勢?

2 何謂差異數量或離中數量?

3 差異數量與集中數量之不同點何在?

4 差異數量之功用何在?

5 各種差異數量,何者最適用?何者最不適用?其故何在?

6 平均差與標準之不同點何在?

7 試略言兩極差,四分差,平均差,標準差之優點與劣點.

8 比較的差異數量之功用何在?

9 用上述第五章上練習問題(7)之材料,求兩極差,四分差,平均差,標準差,求算時,各用最簡捷之方法.

10 將下列材料,用皮爾生之公式,求差異係數.

A 某城每女子,每年所需之化粧費.

M=\$9·94 σ=\$3·42

B 某城每女子,每年所需之教育費.

M=\$42·43 σ=\$19·27

11 求標準差時,通常何以用算術平均數作根據?求平均差時,何以用中數作根據?試言其理由安在?

第八章　相關數量

30　相關及相關數量之意義

I 相關之意義　吾人在上列第六第七兩章中之所敍述者,都是一種測量,或一種事實,或一個變數的幾種特性.例如各個量數的分配,是否集中;(卽集中數量)各個量數彼此間的相差,是近是遠;(卽差異數量)而並未論到兩種事實或兩個變數的相關.宇宙間之一切事物,雖是千變萬化,彼此錯綜,然每有某種關係之存在,或每每互生關係.如就敎育事實或心理事實來舉例:(1)智力高者,學業成績是否也高?(2)國文能力優良者,學習其他智識學科是否亦優良?(3)物理成績優良者,化學成績是否亦優良?(4)體重者,是否體育成績也好?諸如此類,比比皆是,不勝枚舉,然皆不出於彼此間相關之問題.故所謂相關者,卽兩種事實彼此間互相爲因,或所發生之關係也.換言之,亦卽甲種事實發生變動時,其乙種事實,是否亦隨之而

變動?或甲種事實增長時,乙種是否亦隨之而增長?或甲種減少時,乙種是否亦隨之而減少?抑兩者之增減,互相為反;或毫無影響.此中相因之有無,或相因之程度,是即所謂相關之意義.

II 相關數量之意義 吾人如已明瞭相關之意義,則相關數量之意義,亦不難隨之而瞭解.所謂相關數量者,即以一簡單之數字,表明各種事實相關之程度.通常多稱此數曰"相關係數",(Coefficient of correction)而以"r"代表之.

31 相關之類別

如依相關之性質而分,普通可別而為三類:(1)正相關,(2)負相關,(3)無相關.玆分述之如下:——

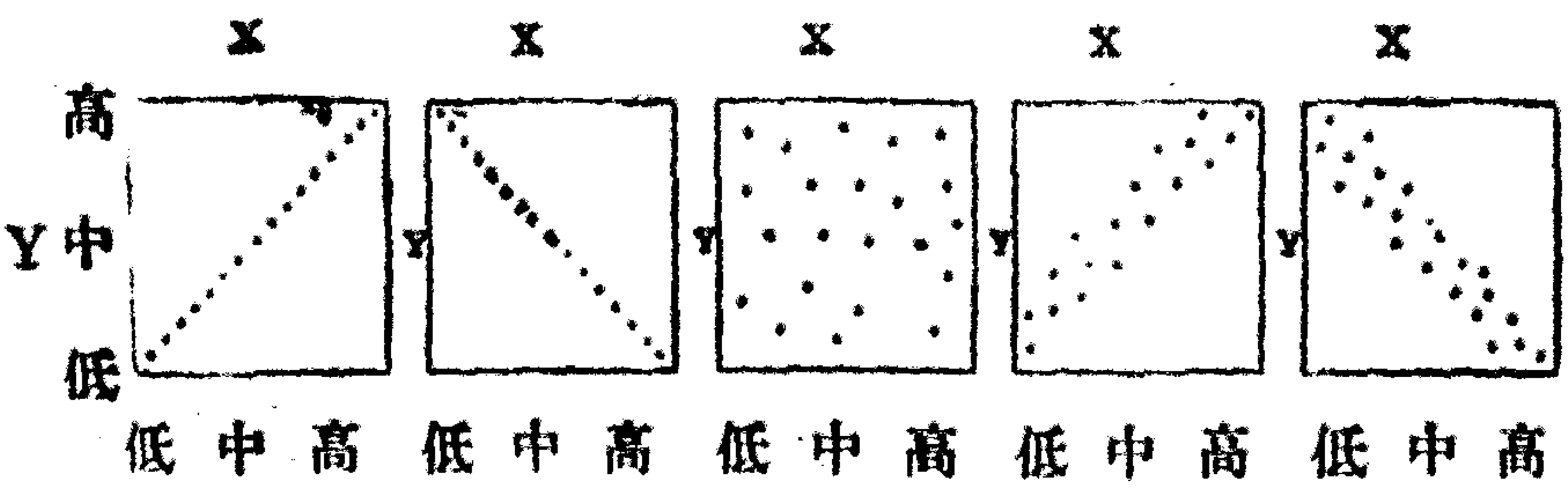
I 正相關 正相關亦名積極的相關.即兩種事實,同相增減之謂也.換言之,亦即甲種事實之量數增加時,乙種亦隨之而增加;或甲量減少時,乙量亦隨之而減少.如舉一實例,則如熱度之增高,與鐵條之漲大是.蓋熱度愈高,則鐵條愈大也.

II 負相關 負相關適與上述之正相關相反.亦稱消極的相關,即兩種事實相反之謂也.換言之,亦即甲量增加時,乙量反減少;甲量減少時,乙量反增加.如舉一實

例,則如壓力與容量是.蓋壓力愈大,則容量愈小也.

III　無相關　無相關亦曰零相關,適居於正相關與負相關之間.即兩種事實增減時,各不受影響之謂也.換言之,亦即甲量增加或減少時,乙量並不隨之發生變動.爾爲爾,我爲我,休戚各不相關.如舉一實例,則如鞋子之大小,與智力之高低固絲毫無關係也.

以上各種的相關,如用圖形來表示,則如下列圖二十四之A,B,C,D,E.抑且一目瞭然.



圖二十四A (顯示完全正相關)　圖二十四B (顯示完全負相關)　圖二十四C (顯示無相關)　圖二十四D (顯示正相關而不完全)　圖二十四E (顯示負相關而不完全)

上列圖二十四A,其x量漸增時,y量亦隨之而漸增,故爲完全正相關.圖二十四B,其x量漸增時,y量反漸減,故爲完全負相關.圖二十四C,其x量變動時,y量毫不受影響,故爲無相關.圖二十四D,其x量漸增時,y量亦漸增.但y量之漸增,並不與x量成嚴密之正比,(Strictly proportional to the increase in x)故爲正相關而不完全者.圖二十四

E,其 x 量漸增時,y 量反漸減.但此二量之增減,並不成嚴密之正比,故爲負相關而不完全者.

32 r 之數值及其高低之解釋

相關之意義與種別,吾人已言之如上.茲將進而言 r 之數值.r 乃代表相關係數,此吾人已敍述於上.故此處之所謂 r 之數值者,即相關之數值,亦即兩種量數相關之指數之值.惟此種數值,必先待解釋而後有意義.例如得 r 數值 •44,究以此表示兩量之相關程度高乎?抑以此表示兩量之相關程度低乎?假使以此表示高的,那末高到何種程度?表示低的,低到何種程度?諸如此類之問題,必先待解釋,而後意義才顯.否則其數值雖求得,亦徒等於形式,曾有何意義之可言.在通常凡是表明完全正關係者,其 r 之數值爲「1」.表明完全負關係者,其 r 之數值爲「—1」.表明完全無關係者,其 r 之數值爲「○」.如 r 由 +1 逐漸減小,則其關係密切之程度,亦逐漸減低;迨減至○時,則其關係之程度,完全消失.反之,如由 —1 逐漸增加,則其相反之關係,即逐漸減低;如增至○時,則其相反之關係,完全消失.故 r 數值之範圍,爲由 +1 經過○,以至 —1 者.而由○至 +1 之間之各數,均表示正相關.由○至 —1 之間之各數均表示負相關.而○之所表示者,則爲無相關.然此僅就 r

數值之範圍,加以解釋.究竟何高何低或高到何種程度,低到何種程度,仍未涉及.今將再進而討論之.關*於此種高低之解釋,本無一定之界限.而一般統計學者,亦每因其觀點之不同,意見亦常不一致.例如:——

A. R. Crathorne 主張	W. A. Mccall 主張	H. O. Rugg 主張
自 0 至±.25為最低	自 0 至±.40為低	自 0 至±.20為無足重輕
自±.25至±.40為低	自±.40至±.70為切實	自±.20至±.40為有相關惟甚低
自±.40至±.55為中數	自±.70至±1.00為高	自±.40至±.70為相關明顯
自±.55至±.70為高		自±.70至±1.00為高
自±.70至±1.00為最高		

綜觀上列三家之主張,其對於高低之解釋,並不一致.艾偉先生曾參照各專家之意見,作一訂正相關係數之量尺.茲附錄如下,以供參考:

最低	低	切實	高
0 — ±.20	±.20 — ±.40	±.40 — ±.70	±.70 — ±1.00

係節錄艾先生之著作,原文載中華教育界十七卷六期數學成績與其他學科成績之相關研究一文中.

上列艾氏之訂正量尺,似折衷各專家之意見.蓋如A. R. Crathorne之主張,似略嫌繁複.而W. A. Mccall之主張,則似略嫌簡單.至H. O. Rugg之主張,則艾氏與其相同.惟Rugg之名詞繁雜,且意義亦不甚明曉,故艾氏之所訂正者,或足為吾人解釋r數值高低之標準也.

33 相關係數之功用及其制限

I 功用 相關係數之功用,吾人在上述29節中已略能窺及其一部分.茲為詳盡而加重注意計,特再在此提及之;然亦僅能就教育事實或心理事實上來舉幾個比較重要的例子.例如下列幾個問題,在求得相關係數後,統能回答:(1)凡是編造一種教育測驗或心理測驗,究竟有多少可靠?(2)假使一種測驗的量表加長,或套數加多,其可靠度是否亦增加?(3)同樣兩套的測驗,是否能測得同一的能力?或其難易是否相等?(4)學生這一科成績優良的,是否那一科成績也優良?(5)測驗的成績和教員評判的成績,是否能相等?(6)平常舊式的考試究竟有多少可靠?(7)學生智力優良的,學業成績是否亦優

良?(8)學校經費之多寡,對於學校效率究竟有多少關係?諸如此類,誠難枚舉;以上所言者,僅不過爲極小的一部分.總之,凡是要研究兩件事實有多少的關係,都可以用相關係數來解答.

II 制限 欲研究兩件事實有多少的關係,吾人固可用相關係數來解答.但是欲研究兩件事實爲什麼有關係?則相關係數卻無從而解釋.下列諸問題恐非相關係數所能解答者.例如:(1)有 ·75 的相關係數,是否有全體的百分之七十五的成績,完全相同或相對.(Prefest Correspondence)(2)相關係數是正,並且較大的時候,假使甲方分數較大時,是否乙方分數也完全較大.(3)兩件事實有很大的相關時,是否此兩件事實有因果之關係.類此之問題,亦不一而足;以上所言者,亦僅不過爲極小的一部分.常人每誤認相關係數爲研究兩件事實種種關係的萬能者,其實則有未然.總之,相關係數只能告訴我們兩件事實有多少的關係,而不能解釋爲什麼有關係.此爲吾人所當注意,而不能誤解者.

34 相關係數之求法

相關係數之求法,每因其相關情形之各異,而有所不同.相關之情形,普通得分直線相關,與曲線相關兩種.此二者,

均屬由變量的事實求相關.除此而外,尚有所謂由品質的事實求相關者.惟後者之相關,不如前者之重要,一般初級統計學書中之所述者,類皆爲由變量的事實而求之者.不過在教育之事實上,類多屬於直線性的相關.故本書之所討論者,亦以屬於直線性者爲限.如在直線性相關中再細別之,則又可分爲二數相關,與多數相關等等.多數相關求法比較複雜,恐非初學者所易瞭解.故本書亦均從略,而專述其直線性的二數相關.求直線性的二數相關之最常用之方法,得分(1)乘積率法,(2)等級相關法,(3)異號相關法.玆分述於下.

35 用乘積率法求相關係數

乘積率法,(Product-moment)又稱均方相關法.爲求直線性相關最完美之方法.其公式爲皮爾生(Pearson)所發明,故有時亦稱皮爾生法(Pearson's method).用此法求相關係數時,可視量數之繁簡,分爲兩種不同的計算法.若量數繁多時,則須另製"相關表".(Correlation table)若量數簡單時,則毋須另製相關表.玆特分述之如下:——

工 不列相關表之求法 先引用皮爾生之公式如下:——

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

如欲再求簡明，則亦可將此公式變成 $r=\frac{\Sigma xy}{\sqrt{\Sigma x^2\Sigma y^2}}$，蓋因 $N\sigma_x\sigma_y=N\sqrt{\frac{\Sigma x^2}{N}\cdot\frac{\Sigma y^2}{N}}=N\frac{\sqrt{\Sigma x^2\Sigma y^2}}{N}=\sqrt{\Sigma x\,\Sigma y^2}$ 之故也。在此式中，r＝皮爾生相關係數之符號。x 爲甲乙兩測量中之任何測量之各量數與其算術平均數之差數。y 爲他測量中之各量數與其算術平均數之差數。（設前者爲甲測量，則後者爲乙測量；設甲測量爲 x，則乙測量爲 y；設甲測量爲 y，則乙測量爲 x。）Σxy 爲甲乙兩測量中各相當差數之乘積之和 σ_x 爲其一測量之標準差。σ_y 爲其他測量之標準差。N 爲每量之次數總數。至其公式中之所以用 σ_x 與 σ_y，則亦顯而易明。茲爲易於瞭解計，特節錄趙文銳氏所譯統計學原理中的一段大意，以解釋之如下：——「若結果實在增加之長度爲 2·6，而同時增加之平均寬度爲 0·73，則此二者之比，卽爲吾人所欲求之相關係數。然若一量之變化，都不過一寸或幾分，而他量之變化，輒以一丈或十丈計，則此一寸之變化，在甲量已爲巨大之變化，而在乙量則可視爲毫無變化也。故吾人欲比較二量之變化，不能以實在之數字相比較，當先求此等量數之標準差，於是此二者之單位均爲標準差，然後可以從事比較矣。」此固爲吾人之所顯而易明者。茲舉例如下列第二十七表，以說

明其算法.

表二十七 顯示用皮爾生之公式不列相關表而求相關係數之計算法

學生號數	X 甲測量之量數	Y 乙測量之量數	x 甲測量之量數與其平均數之差	y 乙測量之各量數與其平均數之差	x^2	y^2	Σxy +	Σxy —
1	85	100	8	25	64	625	200	
2	83	70	6	—5	36	25		30
3	92	72	15	17	225	289	255	
4	91	85	14	10	196	100	140	
5	73	46	—4	—29	16	841		
6	68	84	—9	9	81	81		116
7	82	77	5	2	25	4	10	81
8	36	70	—41	—5	1681	25	205	
9	76	36	—1	—39	1	1521	39	
10	87	88	10	13	100	169	180	
11	75	72	2	—3	4	9	6	
12	76	80	1	5	1	25		5
	m=77	m=75			$\Sigma x^2=2430$	$\Sigma y^2=3714$	$\Sigma xy=753$	

$$r=\frac{\Sigma xy}{\sqrt{\Sigma x^2\,\Sigma y^2}}=\frac{753}{\sqrt{2430\times 3714}}=\frac{753}{49\cdot 295\times 60\cdot 88}$$

$$=\frac{753}{3010\cdot 79}=\cdot 250$$

茲將其演算之步驟,分述之如下:——

1　表上第一行之1,2,3,4,……等,代表學生.

2　表上第二行之X,代表甲測量之量數.第三行之Y,代表乙測量之量數.

3　求甲測量各量數之算術平均數,得77.將此77與其各量數相減,即得差數,如表上第四行之8,6,……—1,等等.其量數較平均數大者,記正號;小者,記負號.此種正負號,切不能遺漏,以免錯誤.但正號亦得不記.求畢後再用同法求乙測量,即得表上第五行之各數.

4　將每一差數平方,即得x^2及y^2,如表上之第六行及第七行.再將x^2及y^2相加,即得Σx^2及Σy^2.本例$\Sigma x^2=2430$,$\Sigma y^2=3714$.

5　將甲測量之每一差數,與其相對之乙測量之每一差數相乘,即得Σxy.如表上之第八行.但須注意正負號,例如學生1在甲測量之差數爲+8,乙測量之差數爲25,以$8\times25=200$.又如學生12,甲測量之差數爲—1,乙測量之差數爲+5,以$-1\times5=-5$.乘畢後,以代數法相加,即得Σxy.本例則爲753.

6　將各數代入公式而求算之,即得$r=\cdot250$,求公式中之$\sqrt{\Sigma x^2\cdot\Sigma y^2}$時,先將$\Sigma x^2$及$\Sigma y^2$,分別開方而後相

乘,或先相乘,而後總開方,可隨各人之便.因其無礙於結果.

不過像上列表二十七之求法,祇能在量數較少時用之.如量數繁多時,則計算綦繁.設或甲乙兩量之算術平均數帶有小數,則計算差數及差數自乘時,尤覺麻煩,此其不便者一.至其第二不便者,則在不能將各個量數分配之位置,作一明顯的表示.因為有此二點的不便,故平常在量數較多之時,多不喜用.據經驗之所得,大概量數在三四十左右,則用上列表二十七之方法,尚稱簡便.如超過三四十以上,則以用相關表之求法為宜.茲特列有相關表之求法,分述如下:——

II 列有相關表之求法 用相關表求相關係數,亦有稱之為用分布圖以求相關係數者.其所應用之公式,與前不列相關表者相同.惟必須先製相關表.茲先將製表之步驟,略行敍述,然後再進而討論其求法.

1 製相關表之步驟

1 應用前述第五章上15節中所述之方法,決定組距之多寡大小及界限.

2 指定一測量之量為 x 量,又一測量之量為 y 量.指定後,再定 ox 及 oy 二量尺之直線,使成直角,相交於零.通常皆使 oy 線為豎線,列於左方,ox 線為

橫線,列於下方.

3　將 x 量中所定之各組距,循 ox 線自左而右的排列.又將 y 量中所定之各組距,循 0y 線自下而上的排列.凡小量數皆由零點起,故 x 量之組距,自左而右.y 量之組距,自下而上.x 量之各組距,本宜寫於底端之橫線,但通常恆寫於頂端之橫線.此種列法,可任意爲之,惟通常皆認爲如此較爲方便.如能用方格紙排列,則各量數歸入時,尤覺方便.

4　組距及量尺既定之後,再將各量數依次歸入.此歸入之方法,亦與前述第五章上15節中所述者相同.惟前者僅止一種測量之量數之歸入,而此則須將兩種測量之各量數配合成對而歸入之也.其歸入之方法,例如第二十八表上之材料,以地理分數爲 x 量,歷史分數爲 y 量.在 x 量中第一人所得之分數爲52,與 y 量中之58分成對.列入表中時,先循 ox 線察得量數52所在之組距 (50—54·9), 再循 oy 線察得量數58所在之組距 (55—59·9), 然後沿 50—54·9 之組距欄向右橫移,至與其相對之量數 55—59·9 之組距之豎欄內止,卽在其相交之方格中,畫一豎線「 / 」,依此類推,將其他各量數一一記入於其相交之方格內,卽成第二十九表,如將

表內之記號線,用數字書出之,則卽成第三十表,所謂相關表,至此卽製成矣.再觀第三十表,卽可察知此兩種量數相關之大概情形矣.

表二十八 某中學九十人的地理分數與歷史分數

地理	歷史	地理	歷史	地理	歷史
52	58	85	82	87	86
50	54	83	85	86	87
58	60	83	82	87	87
60	56	83	78	88	87
56	60	87	78	88	88
60	58	86	76	86	89
55	54	85	93	89	88
57	60	84	89	88	89
55	72	85	90	86	88
51	60	83	90	87	88
64	72	82	85	88	86
67	79	81	84	86	87
68	85	82	84	88	89
68	88	82	88	87	87
72	73	82	86	88	88
73	76	83	88	88	91
78	82	86	79	87	92
76	83	86	83	88	93
76	81	87	82	81	78
80	83	87	88	92	92

80	85	86	82	94	96
79	88	87	83	93	98
79	89	87	84	98	96
80	90	86	82	94	97
81	72	86	84	95	96
71	74	87	83	92	98
82	76	88	81	94	99
84	78	87	83	95	96
81	84	88	85	98	99
83	81	86	86	99	94

表二十九　顯示歸入第二十八表之材料

地理分數(x)

歷史分數(y)

	50-54.9	55-	60-	65-	70-	75-	80-	85-	90-	95-99.9
95-99.9									/	////
90-						/			/	/////
85-						///	///// ///// //	///// ///// ///// //	/////	
80-					/	///	///// /	///// //	//	
75-							////	//		
70-					//	/				
65-						/		//		
60-		//			/					
55-	/		///		/					
50-54.9	/	/	/							

表三十 顯示列有相關表之相關係數計算法

x

y	50-54.9	55-	60-	65-	70-	75-	80-	85-	90-	95-99.9	f_y	d_y	fd_y	fd_y^2	$\Sigma x'y'$ +	−
95-99.9									1^{9}	4^{12}	5	3	15	45	57	
90-						1			1^{6}	5^{8}	7	2	14	28	46	
85-						3	2^{1}	17^{2}	5^{3}		37	1	37	37	61	
80-					1	3	6	7	2		19	0				
75-							3^{-1}	2^{-2}			5	−1	−5	5		7
70-					2^{2}	1					3	−2	−6	12	4	
65-						1		2^{-6}			3	−3	−9	27		12
60-		2^{16}			1^{4}						3	−4	−12	48	36	
55-	1^{25}		3^{15}		1^{5}						5	−5	−25	125	75	
50-54.9	1^{30}	1^{24}	1^{18}								3	−6	−18	108	72	
f_x	2	3	4	0	5	9	21	28	9	9	90					
d_x	−5	−	−3	−2	−1	0	1	2	3							
fd_x	−10	−12	−12	−0	−5		21	56	27	36						
fd_x^2	50	48	36	0	5		21	112	81	144						

$\Sigma fd_x + = 140$

$\Sigma fd_x - = 39$

$\Sigma fd_x = 140 - 39 = 101$

$\Sigma fd_y + = 66$

$\Sigma fd_y - = 75$

$\Sigma fd_y = 66 - 75 = -9$

$\Sigma fd^2_x = 497$

$\Sigma fd^2_y = 435$

$\Sigma x'y' + = 351$

$\Sigma x'y' - = 19$

$\Sigma x'y' = 351 - 19 = 332$

$$c_x = \frac{\Sigma fd_x}{N} = \frac{101}{90} = 1 \cdot 12$$

$$c^2_x = (1 \cdot 12)^2 = 1 \cdot 2544$$

$$c_y = \frac{\Sigma fd_y}{N} = \frac{-9}{90} = -\cdot 01$$

$$c^2_y = (-\cdot 01)^2 = \cdot 0001$$

$$c_x c_y = 1 \cdot 12 \times (-\cdot 01) = -\cdot 0112$$

$$\frac{\Sigma x'y'}{N} = \frac{332}{90} = 3 \cdot 6889 \qquad \sigma y = \sqrt{\frac{\Sigma fd^2_y}{N} - c^2_y}$$

$$\sigma_x=\sqrt{\frac{\Sigma fd'_x}{N}-cx_2}\qquad\qquad \sigma_y=\sqrt{\frac{\Sigma fd^2_y}{N}-cy^2}$$

$$=\sqrt{\frac{497}{90}-1\cdot2544}\qquad\qquad =\sqrt{\frac{435}{90}-\cdot0001}$$

$$=\sqrt{5\cdot5222-1\cdot2544}\qquad\qquad =\sqrt{4\cdot8333-\cdot0001}$$

$$=\sqrt{4\cdot2678}=2\cdot066\qquad\qquad =\sqrt{4\cdot8332}=2\cdot198$$

$$r=\frac{\frac{\Sigma x'y'}{N}-c_xc_y}{\sigma_x\sigma_y}=\frac{3\cdot6889-(-\cdot0112)}{2\cdot066\times2\cdot198}=\frac{3\cdot7001}{4\cdot5411}=\cdot8148$$

or $=\cdot815$

2　算法　相關表製成後,其第二步手續,卽爲求算 r.求算 r 時,與上述不用相關表之求算 r 之公式一樣.亦卽引用皮爾生之公式.

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

不過在實際求算時,吾人常將公式變之而爲

$r=\dfrac{\frac{\Sigma x'y'}{N}-c_xc_y}{\sigma_x\sigma_y}$.此式與上述 $r=\dfrac{\Sigma xy}{N\sigma_x\sigma_y}$ 之公式,並無差異,特形式上之不同已耳.在此式中 r 爲相關係數,N 爲每量中之次數總數. σ_x 爲 x 量之標準差. σ_y 爲 y 量之標準差.x' 爲 x 量各量數與其假設平均數之差數.y' 爲 y 量各量數與其假設平均數之差數.因其差數爲由假設平均數而求得,故必須各有校正數,

如 c_x c_y 以更正之.換言之,亦即必須有 $c_x c_y$ 以校正之,方能得實在之 Σxy. 每一差數 x',必須校正數 c_x,每一差數 y',必須校正數 c_y,故每一 $x'y'$,須有校正數 $c_x c_y$.按全表計之,卽 $\frac{\Sigma x'y'}{N}-c_x c_y$,亦卽 $r=\frac{\Sigma xy-Nc_x c_y}{N\sigma_x\sigma_y}$,或 $r=\frac{\frac{\Sigma x'y'}{N}-c_x c_y}{\sigma_x\sigma_y}$. 今將由 $r=\frac{\Sigma xy}{N\sigma_x\sigma_y}$ 變作 $r=\frac{\frac{\Sigma x'y'}{N}-c_x c_y}{N\sigma_x\sigma_y}$ * 之公式,證明之如下:——

設 Σx 與 Σy 爲二假設平均數,

c_x 與 c_y 爲二校正數,

M_x 與 M_y 爲二眞正平均數,

則 $M_x=\Sigma_x+c_x$,　　$M_y=\Sigma_y+c_y$.

設 x 與 y 爲由二眞正平均數 M_x 與 M_y 所得之差數,

x' 與 y' 爲由二假設平均數 Σx 與 Σy 所得之差數,

則 $x'=x+c_x$,　　$y'=y+c_y$.

因此,所以 $\Sigma x'y'=\Sigma(x+c_x)(y+c_y)=\Sigma xy+c_y\Sigma x+c_x\Sigma y+\Sigma c_x c_y$.

現在因爲 Σx 及 Σy,爲由二眞正平均數所求出之差數其和均等於○.

故 $\Sigma x'y'=\Sigma xy+\Sigma c_x c_y$ 或 $\Sigma xy=\Sigma x'y'-\Sigma c_x c_y$.

以此代入皮氏原公式 $r=\frac{\Sigma xy}{N\sigma_x\sigma_y}$,

或 $r=\frac{\Sigma x'y'-Nc_xc_y}{N\sigma_x\sigma_y}$，同以N除之，

則卽得 $r=\frac{\frac{\Sigma x'y'}{N}-c_xc_y}{\sigma_x\sigma_y}$.

* 見 H. O. Rugg:—— Statistical methods applied to Education. Page 269.

此公式之改變，既已明瞭，則吾人卽可求出其公式中所代表之各數，然後再求 r，固已易如反掌。茲將上列第三十表之計算法，再述其求得各數之步驟如下：——

1　將表上各橫行之次數相加，得 f_y 欄內之各數。將各豎行之次數相加，得 f_x 欄內之各數。並分別書其和數於 f_y f_x 相交之方格中，卽得N。本例則N爲90。惟須注意者，卽 f_y 之總數，必須與 f_x 之總數相等，否則必陷於謬誤。

2　擇定兩量中含有平均數之組距，作爲假設平均數之所在。如本例則 x 量擇定 75—79·9 的一組距，y 量則擇定 80—84·9 的一組距。選擇時，上下幾個組距，無大妨礙，只求其不過遠，以免不正確或計算時之麻煩。擇定後，再分別書二粗線，使其特別顯著，以示區別。

3　求 x y 兩量之各量數，與其各該估計平均

數之差數,如表中d_x及d_y欄內之$-1,-2,\cdots\cdots+1,+2,\cdots\cdots$等.求算時,以每一組距視作1單位.

4 將x,y兩量之次數,與其相當之差數相乘,得fd_x及fd_y欄內之各數.如本例fd_y欄內之各數,爲$5\times3=15, 7\times2=14 \cdots\cdots$等等.但須注意正負號.

5 將fd_y及fd_x欄內之各數,用代數法相加,並求其總和卽得Σfd_y,及Σfd_x.在本例$\Sigma fd_y=66-75=-9$. $\Sigma fd_x=140-39=101$.

6 用量數總數N除Σfd_y及Σfd_x,卽得校正數c_y及c_x.在本例$c_y=\frac{-9}{90}=-{\cdot}01$. $c_x=\frac{101}{90}=1{\cdot}12$.

7 將c_y及c_x方之,卽得c^2_y及c^2_x.本例$c^2_y=(-{\cdot}01)^2={\cdot}0001$. $c^2_x=(1{\cdot}12)^2=1{\cdot}2544$.

8 將d欄之各數,乘fd_y及fd_x欄內之各數,卽得fd^2_y,及fd^2_x欄內之各數.

9 將fd^2_y及fd^2_x欄內之各數相加,卽得Σfd^2_y及Σfd^2_x.在本例$\Sigma fd^2_y=435$. $\Sigma fd^2_x=497$.

10 用量數總數N,除Σfd^2_y,再減去c^2_y,然後開方,卽得σ_y.在本例$\sigma_y=\sqrt{\frac{435}{90}-{\cdot}0001}=\sqrt{4{\cdot}833-{\cdot}0001}=\sqrt{4{\cdot}8332}=2{\cdot}198$.求$\sigma_x$,可依此類推.在本例$\sigma_x=\sqrt{\frac{497}{40}-1{\cdot}2544}=\sqrt{5{\cdot}5222-1{\cdot}2544}=$

$\sqrt{4.2678}=2.066$. 此與上述第六章26節求標準差之簡捷法之理相同,讀者可參閱.惟在此有一注意之點,卽求 σ_x 及 σ_y 時,皆以組距爲 1 單位而計算,然無須用組距單位相乘,而使其還原.此因求 $\Sigma x'y'$ 時,尙須用 1 作單位也.若將 σ_x 及 σ_y 乘原來單位,使其還原,則以後求 $\Sigma x'y'$ 時,亦必使之還原爲原來單位.反正在數理上,同乘同除,其値不變,故此等手續可省.

11　σ_x 及 σ_y 求得後,卽可求每一方格中之 $x'y'$. 求 $x'y'$ 時,最須注意者爲:對於各差數之正負號.蓋正負號一錯,對於計算之結果,有極大之影響.求算時,先須看其 x 與 y 兩量交切處之次數.例如本例內 x 量中 55—59.9 之組,與 y 量 60—64.9 之組相交切處有量數 2 個.此 2 個量數,與 y 量之假設平均數相差爲 —4, 與 x 量之假設平均數相差亦爲 —4, 以 $-4\times-4=16$, 再以次數乘之,卽得 $16\times2=32$. 最好能將此16書於該方格內之右上方,則比較的能一目瞭然.又如 x 量中 90—94.9 之組,與 y 量85—89.9 之組之交切處有量數 5 個.此 5 個量數與 y 量之假設平均數相差爲 1,與 x 量之假設平均數相差爲 3,以 $1\times3=3$, 再以 5 乘之得 $5\times3=15$. 如

將此 3 書於該方格上之右上方,則亦能一目瞭然.餘均可依此類推,將所有方格內之量數之 x'y',全行算出.

定 Σx'y' 之正負號,其法極易,其理亦淺,祇要將 x y 兩量之假設平均數所在之組距,各畫一粗線,分全表爲四部分.如所列之組距,俱從小量數起,則在左邊者,(即 y 量)由下而上,在頂面者,(即 x 量)自左至右.如是,則表之右上部及左下部之各 Σx'y',皆爲正數.在左上部及右下部之各 Σx'y',皆爲負數.在兩假設平均數所在之各組中之各 Σx'y',皆爲零,可不必計算.茲爲便於明瞭計,另附圖如下列第二十五圖.

左上部	x=— y=+ —×+=— xy=—	x=+ y=+ +×+=+ xy=+	右上部
左下部	x=— y=— —×—=+ xy=+	x=+ y=— +×—=— xy=—	右下部

圖二十五　顯示相關表中 xy 之正負

12 每方格中之 $x'y'$ 求得後,即可將每一橫行之各 $x'y'$,依代數法相加,加得後,將其和數列於表右之 $\Sigma x'y'$ 欄下,此卽爲每橫行之 $\Sigma x'y'$. 其正負各數,最好能分列兩行,以醒眉目.

13 將 $\Sigma x'y'$ 欄下之各正數及負數相加,使正負相消,卽得 $\Sigma x'y'$ 之總數.如本例則爲 $351-19=332$.

14 因爲 $\Sigma x'y'$ 係由兩量之假設平均數求得,故必須有校正數以校正之,方能得實在的 Σxy. 故每一差數 x',必須有 c_x 校正之,每一差數 y' 必須有 c_y 校正之.同理每一 $x'y'$,亦必須有 c_xc_y 以校正之.

15 以 c_x 乘 c_y.在本例則爲 $1{\cdot}12\times-{\cdot}01=-{\cdot}0112$.

16 以N除 $\Sigma x'y'$ 求 $\dfrac{\Sigma x'y'}{N}$ 在本例則爲 $\dfrac{332}{90}=3{\cdot}6889$.

17 求 $\dfrac{\Sigma x'y'}{N}-c_xc_y$ 在本例則爲 $\dfrac{332}{90}-1{\cdot}12\times-{\cdot}01=3{\cdot}6889-(-{\cdot}0112)=3{\cdot}7001$.

18 將上列各數代入公式 $r=\dfrac{\dfrac{\Sigma x'y'}{N}-c_xc_y}{\sigma_x\ \sigma_y}$

$$=\frac{\dfrac{332}{90}-(1{\cdot}12\times-{\cdot}01)}{2{\cdot}066\times2{\cdot}198}=\frac{3{\cdot}6889-(-{\cdot}0112)}{4{\cdot}5411}=\frac{3{\cdot}7001}{4{\cdot}5411}$$

$=\cdot8148$,或 $\cdot815$.此卽爲所求之 r.

Ⅲ　塞斯頓相關表之計算法　上列之計算法,係普通所常用者.最近塞斯頓氏 (L.L. Thurstone) 在其所著之 The Fundamentals of Statistics 一書中之第二十二章 202 面上,(該書朱君毅氏有譯本,名曰教育統計學綱要商務出版)亦列有一計算法.其計算時,所應用之公式與上述皮氏之公式完全相同.其計算手續,亦與其相同.其所不同者,僅計算 Σxy 時,略有出入.塞氏之法,頗稱便當.茲特將塞氏之原表,介紹之如下:

塞氏相關表之計算手續,除 Σxy 時略有出入外,其餘與前述者完全相同,故無贅述之必要.茲僅將求 Σxy 時之手續,略加敍述:

1　先求每一橫行各量數與 x 量之假設平均數之差數和,得 $\Sigma x+$.例如最低端的一橫行裏,(即組距 102—106 的一行)其右邊第一方格之量數爲 1,與 x 量之假設平均數之差數爲 1,以 $1\times1=1$,此即 $\Sigma x+$.其左邊第一方格之量數爲 3,與 x 量之假設平均數之差數爲 -1,以 $3\times-1=-3$.其第二方格之量數爲 2,與 x 量之假設平均數之差數爲 -2,以 $2\times-2=$

表三十一　塞斯頓相關表

本表轉錄自 Thurstone: The Fundamentals of Statistics. PP. 202

	1	2	3	4						5					6	7	8	9	10	11	12	13	14	15	16	17
A	Y		61	62	63	64	65	66	67	68	69	70	71	72	73	74	f	y	fy	fy²	Σx		Σx		Σxy	
B			619	629	639	649	657	669	679	689	699	709	719	729	739	749										
C	X		7	6	5	4	3	2	1	0	1	2	3	4	5	6					−	+	−	+	−	+
D	187	191									3	2	2	1			9	10	90	900		17		17		170
E	182	186							1						1		2	9	18	160	1	5		4		36
	177	181										1	1				3	8	24	192		11		11		88
	172	176									1	1	6	2			11	7	77	539		29		29		209
	167	171									2	6	1	2	1		13	6	78	468		30		30		180
	162	166	1						2		3	8	2	2	2		22	5	110	550	9	43		34		170
	157	161						4	1		7	5	7	1			31	4	124	496	9	42		33		132
F	152	156					2	2	3		10	12	11		1	1	56	3	168	504	13	78		65		195
	147	151					2	3	7		12	9	8	3			49	2	98	196	19	66		47		94
	142	146					7	12	10		17	8	15	5	2		93	1	93	93	55	108		53		53
S	137	141															110	0								
	132	136		2		4	9	18	18		8	8	4	3	1	1	93	1	93	93	109	59	50			50
	127	131	1	1	1	7	7	11	15		18	9	5	2			93	2	186	372	104	59	45			90
	122	126		1	4	2	12	17	16		4	5			1		76	3	228	684	120	19	101			303
	117	121		2	2	10	9	6	6		2	2	2				48	4	192	768	107	12	95			380
	112	116			2	7	3	3	3				2				20	5	100	500	56	6	50			250
H	107	111		3	1	5	2	1	1								13	6	78	486	52		52			312
K	102	106			1			2	3		1						8	7	56	392	12	1	11			77
L	f		2	10	11	38	57	93	106	126	109	87	75	23	9	4										
M	x		7	6	5	4	3	2	1	0	1	2	3	4	5	6										
N	fx		14	60	55	152	171	186	106		109	174	225	92	45	24										
P	fx²		98	360	275	608	513	372	106		109	348	675	368	225	144										

$\Sigma x_+ = 669$	$\Sigma y_+ = 880$	$\Sigma xy_+ = 2783$
$\Sigma x_- = 744$	$\Sigma y_- = 933$	$\Sigma xy_- = 0$
$\Sigma = -75$	$\Sigma y = -53$	$\Sigma xy = +2783$
$\Sigma x^2 = 4201$	$\Sigma y^2 = 7377$	$N = 750$

$d_x = \frac{\Sigma x}{N} = \frac{-75}{750}$　　$d_y = \frac{\Sigma y}{N} = \frac{-53}{750}$

$d_x = -.10\quad d_x^2 = .01$　　$d_y = -.071\quad d_y^2 = .005$

$\sigma_x^2 = \frac{\Sigma x^2}{N} - d_x^2 = \frac{4201}{750} - .01 = 5.60 - .01 = 5.59$

$\sigma_x = 2.36$

$\sigma_y^2 = \frac{\Sigma y^2}{N} - d_y^2 = \frac{7377}{750} - .005 = 9.835 - .005 = 9.83$

$\sigma_y = 3.14$

$$r = \frac{\Sigma xy/N - d_x d_y}{\sigma_x \sigma_y} = \frac{2783/750 - (-.10 \times .07)}{} = \frac{3.71 - .007}{2.36 \times 3.14}$$

$r = +.50$

X 變量＝體高吋數

Y 變量＝體重吋數

材料來源　俄海阿省立大學學生

統計者　審斯碼　日期 1922，一月17.

相關紙

衛斯特心理系高等工業學校教職員本書文庫 1373號

—4. 其第五方格之量數爲1,與x量之假設平均數之差數爲—5,以 $1\times-5=-5$. 將 —3,—4,—5, 和之得—12, 此卽 $\Sigma x-$. 依此類推,將各橫行之 $\Sigma x-$ 或 $\Sigma x+$, 一一求出,記入表右之 $\Sigma x-$,及 $\Sigma x+$ 的一欄裏.

2 各橫行之 $\Sigma x-$ 及 $\Sigma x+$ 求得後,再用代數法相加,卽得 $\Sigma x+$,或 $\Sigma x-$.例如最低端的一橫行裏, $\Sigma x-=12$, $\Sigma x+=1$, 相加得 $\Sigma x-=11$. 其實此欄可有可無,惟有了此欄,則眉目較淸.

3 再求與各該橫行相當之直行組距與y量之假設平均數之差數y,以此去乘 Σx, 卽得各該橫行之 Σxy. 例如最低端的一橫行裏, $\Sigma x=-11$, 與相當之y量之假設平均數之差數爲—7,以此 $-7\times-11=77$. 此卽該橫行之 Σxy. 其餘均可照樣求得.

4 各橫行之 Σxy 求畢後,再用代數法相加,卽得 Σxy 之總數.此與前述者相同.

塞氏相關表之計算手續,與前述者之不同,僅止於此.據編者平常求算相關係數之經驗,以用塞氏之方法求 Σxy, 較爲便利.且不容易錯誤,學者常可採用之.

36 用等級法求相關係數

吾人上述的以乘積率法求相關係數,係從兩測量中之各量數之絕對價值或原值 (Original Values or Scores) 而求得者.如此所求得之結果,固屬精確.但有時卻亦不免有困難之處,或不便之點.其困難之處,即在有時教師的評定成績,不用分數而以成績之優劣排列等級.如第一,第二,第三,等等.其不便之點,即在量數較少的時候,若用乘積率法以求算,不免費時力較多.因此二點,故求算相關係數時,除用乘積率法外,再另有用等級法者. (Correlation by rank) 所謂用等級法求相關者,即求算時,僅憑兩測量中各量數所在之位置,而不問其各量數之絕對數值之大小.由等級法求相關之方法有二:一爲斯披門 (Spearman) 之等級法,一爲斯披門之等級法之簡捷法.茲分述之如下:——

I　斯披門之等級法　斯氏由其經驗之所得,發明一公式如下:——

$$\rho=1-\frac{6\,\Sigma(V_x-V_y)^2}{N(N^2-1)}\text{或}\rho=\frac{6\,\Sigma D^2}{N(N^2-1)}$$

式中之 ρ (讀如rho),爲斯披門之等級法求出之相關係數之符號.V_x 與 V_y,爲 x 與 y 兩量數之等級.N 爲每量次數之總和.由此法求出之 ρ,較由乘積率法求出之 r 之價值稍小.惟皮爾生曾發明一個公式,可將 ρ 轉化爲 r,其公式爲:

$$r=2\,\mathrm{Siu}\left(\frac{\pi}{6}\rho\right)$$

其式中之 ρ，卽等 $1-\frac{6\Sigma D^2}{N(N^2-1)}$，故我們於求得 ρ 後，可直接改爲 r 之價值，無須實行去求.（由 ρ 求 r 之價值，見附表一）茲舉一例如下列第三十二表，以說明求 ρ 之方法.

表三十二 顯示斯披門等級相關之求法

學生號數	手工分數 x	圖畫分數 y	x之等級	y之等級	x, y,等級之差數 D	x, y,等級之差數方 D^2
A	94	93	1	1	0	0
B	90	92	2	2·5	—·5	·25
C	86	92	3·5	2·5	1	1
D	86	70	3·5	7	—3·5	12·25
E	72	82	5	4	1	1
F	70	76	6	5·5	—·5	·25
G	68	65	7	9	—2	4
H	66	76	8	5·5	2·5	6·25
I	64	68	9	8	1	1
J	61	60	10	10	0	0
N=10						62=ΣD^2

$$\rho=1-\frac{6\Sigma D^2}{N(N^2-1)}=1-\frac{6\times 26}{10(10^2-1)}=1-\frac{156}{990}$$

$=1-·1576=·84$ 若 $\rho=·84$ 則 $r=·85$（由附表一查得）

玆再述其求算之步驟如下:——

1　表之第一行,係學生號數.第二行係手工分數,以 x 代之.第三行係圖畫分數.以 y 代之.

2　求 x 欄各分數所居之等級.求等級時,以最高分數爲第一可,以最低分數爲第一亦可.但 x y 兩量,則必須一致.在本例則以最高分數爲第一.例如 A 生得94分爲最高,居第一.B 生得90分爲第二.其中如果兩個分數相同者,則可平均一下.例如學生 C 及 D 均爲86分,應佔之等級爲第三及第四.然誰爲第三,誰爲第四,很難斷定.故卽將兩數相加,以二除,各得3.5.此法叫做"中級法",(The mid-rank method) 用之者頗多.如遇三個分數或三個以上的分數相同時,亦可將其應佔的幾個等級相加而後除之.

3　x 欄之等級求畢後照樣求 y 欄內之等級.

4　從 x 欄之各等級,減去相對之 y 欄之各等級,卽得 D 行之各數.正負號須分別記出,如不記出亦無妨,反正自乘以後,正負號取消也.

5　將 D 行各數自乘,卽得 D^2 行內之各數.

6　將 D^2 行各數相加,卽得 ΣD^2.

7　代入公式 $\rho = 1 - \frac{6\,\Sigma D^2}{N(N^2-1)} = 1 - \frac{6\times 26}{10(10^2-1)}$

$= 1 - \frac{156}{990} = .84.$

8 以 ·84 査附表一即得 r=·85

II 斯披門簡捷法 斯氏除前法外,尙有一簡捷法.其公式爲.

$$R=1-\frac{6\Sigma G}{N^2-1}$$

式中之R,爲由簡捷法求出之相關係數之符號.G爲y量之等級大於x量之等級或x量之等級大於y量之等級.N爲每量次數之總和.此法與前法相同,惟比較簡易.蓋前法須求各等級差數之平方和,而此法則僅須求G之和.由此法求出之R之值,亦小於由乘積率法求出之r之值.惟皮爾生亦曾發明一種公式,可將R轉化爲r.其公式爲:——

$$r=2\text{Cos}\frac{\pi}{3}(1-R)-1$$

其式中之R,即等於$1-\frac{6\Sigma G}{N^2-1}$.故我們於求得R後,可直接改爲r之價値,而無須實行去求.(由R求r之價値見附表二)茲亦舉一例如下列第三十三表,以說明求R之方法.

表三十三 顯示斯披門等級相關之簡捷法

學生號數	手工分數 x	圖畫分數 y	x之等級	y之等級	G (x大於y)
A	94	93	1	1	
B	90	92	2	2·5	

C	86	92	3·5	2·5	1
D	86	70	3·5	7	
E	72	82	5	4	1
F	70	76	6	5·5	·5
G	68	65	7	9	
H	66	76	8	5·5	2·5
I	64	68	9	8	1
J	61	60	10	10	
N=10					6=ΣG

$$R=1-\frac{6\,\Sigma G}{N^2-1}=1-\frac{6\times 6}{99}=1-\frac{36}{99}=\cdot 6354$$

$R=\cdot 6354$ $r=\cdot 55$（由附表二查得）

茲再述其求算之步驟如下:——

1 表之第一,二,三,四五,共五行,與上述之由非簡捷法而求 ρ 者,完全相同,無贅述之必要.

2 從 x 量之等級減 y 量之等級,祇記 x 大於 y 者,而不記 y 大於 x 者,卽得 G 行內之各數.(如欲記 y 大於 x 者亦可,惟始終必須一致,亦卽假使記 y 大於 x,則不記 x 大於 y.)此「G」原名爲 Gains, 卽「勝於」之意.

3 將 G 之各數相加卽得 ΣG.本例之 ΣG 爲 6.

4　代入公式 $R=1-\frac{6\Sigma G}{N^2-1}=1-\frac{6\times6}{99}=1-\frac{36}{99}$ $=\cdot6354.$

5　以$\cdot6354$查附表二卽得$r\cdot=\cdot58.$

37　用異號法求相關係數

求相關係數之最簡捷而最精密之方法,當推異號法.(Method of unlike signed pairs) 此法之公式,爲薛伯特 (Sheppard) 所發明.茲舉之如下:——

$$U=\frac{u+\left(\frac{\frac{u}{u+\ell}+\frac{1}{2}}{2}\right)d}{N}$$

在此公式內,U爲薛伯特異號相關係數之符號.ℓ爲x y兩量之各量數與各平均數相差之數之符號相同者.在此法內,其各量數與其平均數相差之數祇記正負號而不記其數值.凡量數之較其平均數大者,記正號,小者,記負號.所謂差數符號之相同者,卽x量中之某量爲正差數,其相對y量中之某量亦爲正差數,例如十十.或x量中某量爲負差數,其相對y量中之某量亦爲負差數,例如一一.u爲差數符號之相異者,適與上述之ℓ相反.例如十一,或一十.d爲x y兩量中之差數符號有一爲〇,或兩個全爲〇者.例如〇〇,十〇,〇十,一〇,等.N爲每量之次數總數,亦卽$\ell+$

u+d之和數.茲舉例如下列第三十四表,以說明其算法.

表三十四　顯示薛伯特異號相關之求法

學生號數	音樂分數 x	體育分數 y	音樂分數之差數符號 x'	體育分數之差數符號 y'
A	50	81	—	+
B	58	54	—	—
C	63	50	—	—
D	67	73	—	0
E	72	74	—	+
F	74	74	0	+
G	84	73	+	+
H	88	80	+	+
I	90	84	+	+
J	94	87	+	+
N=10	m=74	m=73		

$N=10 \qquad l=6$

$u=2 \qquad d=2$

$$U=\frac{u+\left(\dfrac{\dfrac{u}{u+l}+\dfrac{1}{2}}{2}\right)d}{N}=\frac{2+\left(\dfrac{\dfrac{2}{2+6}+\dfrac{1}{2}}{2}\right)2}{10}$$

$$=\frac{2+(\cdot375\times2)}{10}=\frac{2\cdot75}{10}=\cdot275$$

若 $U=\cdot275$　則 $r=\cdot6495$（由附表三查得）

茲再述其求算之步驟如下:——

1 第一第二第三共三行與前述求 ρ 之法相同.

2 將 x y 兩量分別求算術平均數.本例 x 量爲 74. y 量爲 73.

3 將 x 量之各量數分別減去求得之算術平均數.如該量數大於算術平均數,則記 + 號;若小於算術平均數,則記 — 號.求畢 x 量之各數後照樣求 y 量之各數,則卽得第四及第五兩行之各數.

4 細察四五兩行之各對符號如相同者則爲 ℓ,相異者則爲 u,遇有〇者,則爲 d,並分別計算 u ℓ 及 d 之數目,本例 ℓ 爲 6.u 爲 2.d 亦爲 2.

5 將各數代入公式,卽得 U=·275.

6 將 U 值 ·275,查附表三得 r=·6495.

於此吾人有一可注意之點,卽用此求相關係數時,假設 x y 兩量如差數無〇時,(卽 d)則 ℓ+u,卽爲 N.由此以觀,則 u 必與其 N 有一相當之比例.故若兩量中之差數無〇時,則由 $\frac{u}{N}$ 之數,卽可照附表三查得相當之 r.但事實上 x y 兩量中總有〇的差數,故必須引用上列之公式.換言之,亦卽無〇的差數時,則就是不用上列的公式,亦能從附表三中查出相當之 r.

以上所述的三種求 r 之方法,(卽皮爾生之乘積率法,

斯披門之等級相關法,薛伯特之異號相關法.)苟 x y 之各量數其分配爲對稱或偏斜不甚,且其間之關係線爲直線性時,則由此三法所求得之 r,其數無多大之差異.否則,以用皮氏之乘積率法爲最妥.蓋皮氏之法,其 x y 各量數之價值,皆直接顯出計算上之效力及關係,不若其他二法之間接.蓋其他二法,有一根本的假定.此根本之假定,卽凡是各等級之相差,都是相等的.(例如90,89,70,50四個量數,排起等級來,爲 1,2,3,4.如以實在數值而言,則前二數與後二數之相差,適爲二十倍.但就等級而言,則前後相等.如以薛氏之異號法言之,則前兩數均爲＋號,後兩數均爲－號.然同爲正號及負號,其間差異卻甚大.此種假定,於事實頗形不合.所以假使在分配偏斜較大的時候,如用此二法,恐不能精確.不過在量數較少的時候,若用皮氏之方法,未免過嫌麻煩,不若用其他二法之簡便.據一般經驗之所得,大約 N 不到30之時,可用其他二法.惟由此兩法所求得之相關係數,僅能用之以表明相關之有無,而不能用之以斷定其間關係密切之程度.故吾人用此兩法,以及解釋其求得之結果時,尙須加以注意也.

練習問題

1　何謂相關?何謂相關數量?

2　何謂正相關,負相關,無相關?

3 試以艾氏之訂正相關量尺爲標準,解釋下列相關係數之高低:—— +·72,—·64,+·18,+·46,+·32,—·56,+·89.

4 試言相關係數之功用及其制限.

5 試用下列材料,列成相關表,用皮爾生之乘積率法,求相關係數.求$\Sigma x'y'$時,須採用 Thurstone 的方法.

數學	化學	數學	化學	數學	化學
85	95	75	70	65	70
60	80	70	75	50	80
65	55	70	60	60	85
35	70	60	80	60	65
90	95	35	80	60	65
90	95	90	80	80	85
40	70	90	80	60	90
90	60	75	60	85	90
75	85	90	65	45	65
70	90	90	95	85	60
90	85	65	90	35	60
90	90	45	55	90	90
85	75	65	80	50	55
85	95	60	65	80	65

85	95	65	90	75	40
60	95	75	75	30	90
85	95	85	70	35	95
80	90	85	60	45	70
85	90	35	55	55	80
85	75	90	85	75	95
35	75	90	90	75	80
60	70	90	85	75	75
85	85	50	75	65	40
60	65	90	85	35	90
35	60	90	95	40	90
45	65	45	55	30	40
65	60	80	75	45	80
70	90	65	55		
75	85	50	90		
80	75	75	80		

6　試用下列材料,用 $r=\frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$ 之公式求 r

物理	化學	物理	化學
72	78	64	68
78	90	48	42
84	84	56	56

65	70	84	82
70	65	75	75
45	50	80	82
58	62	90	90
77	80	48	50
84	92	92	94
82	81	88	88
		92	98

7 試用下列材料,用斯披門等級相關簡捷法,及薛伯特之異號相關法求相關係數.並將其求得之係數,轉化爲r

英文	手工	英文	手工
20	80	58	48
32	50	32	68
80	50	75	90
38	88	44	67
72	24	27	42
64	64	46	64
71	76	67	38
50	20	24	56

8 假使N爲48,且其分配爲偏斜,用斯薛二氏之公式求相關係數,是否可靠?何故?

第九章 可靠性

38 何謂可靠性

可靠性亦稱之曰可靠度.(Reliability or unliability) 乃應用機率 (Probability) 之原理,以斷定統計結果之確度者也.吾人在前幾章上所求得之集中數量,差異數量,相關數量等,若不計算其可靠性,則究竟有幾多確度,吾人尙不得而知.例如吾人已測得某城的二百兒童之算學四則能力之成績,並亦求得其平均數.假使吾人欲用此平均數而證明(或下一判斷)某城兒童的算學四則能力之如何,則其先決之問題,就是不能不知此已求得之平均數可靠之程度.假使此平均數之可靠性甚大,那末所證明的或所判斷的,也是正確的.假使爲不可靠的,那末所證明或所下之判斷,也是不可靠的.故可靠性之計算,極爲重要.教育上之統計事業,被視爲科學的,與可信的,亦以其能計算可靠性之故.蓋吾人測量任何事實,必不能得其事實之全量,而僅爲

全量中之一部分之取樣.例如吾人欲編造一個小學算術四則測驗的常模 (Norms),照理想或最完美之辦法,自然最好將全國所有小學的兒童,一一加以測驗,然後求其平均數.(當然不僅止平均數,惟此處僅以平均數設例.)然在事實上,每因限於時間,經濟,人才,或其他種種之關係,決不能毫無缺漏,故勢不得不從事於抽選若干兒童,加以測驗,來做代表.此種抽選在統計學上,名之曰抽樣或取樣(Sampling).吾人苟對此能有一相當的瞭解,則吾人在無限量之事物中,取其有限的若干量作樣本,而從此樣本中所求得之平均數,(此處亦僅以平均數舉例,其他差異量數,相關量數,當然爲同樣的情形,以下均仿此.)苟非出於偶然,決不能與其無限量數中所求出之眞正的平均數相同.其所以不能相同之理由,當亦甚淺顯.蓋因有限的若干量數中,若加多或減少一個量數,則其平均數亦必多少隨之而變更.譬如測量某級學生之算學能力,若第一次被測時爲40人,其平均數爲82;若第二次測驗時,適有甲生因病未到,不能參加,則其平均數除非偶而湊巧,決不能適等於82.又如第三次測驗時,若加添一丁生,則其平均數,亦決難等於第一次之82,或第二次之所試驗者.因此吾人如欲得一眞正的平均數,必須就其事物之所有一切的量數中求之.然此種眞正的平均數,在事實上決不可能.故平常所求得

之平均數,除非偶而湊巧,可與眞正的平均數相等外,其餘者僅能爲其近似值,或近似平均數.此種近似平均數,平常稱之曰實得平均數.

現在吾人所急欲知道的問題,就是平常實得之平均數,究與眞正者相去有多遠,或相差若干?夫實得平均數既與眞正平均數有相差,則吾人必須求其相差之度,以證明實得平均數變動之大小,即不將全量一一來測量,亦可推知眞正平均數所發現之界限.既省逐一測驗的時間,又省各箇統計之勞力,此即所謂可靠性之意義,及可靠性之計算之應用.而教育之統計事業之被視爲科學的與可信的,亦即在此.至其可靠與不可靠之分辨,(或可靠性之大小)則即視實得數與眞正數之相差之大小.如相差愈小,則此實得數之可靠性愈大.反之,則其可靠性愈小,亦即不可靠性愈大.故其間之差度之大小,與可靠性成反比例,而與不可靠性成正比例.至如何能使此二者之差度愈小,則全視所取以代表全體之樣本之多少,及各樣本彼此間之σ之大小而定.如所取之樣本愈多,σ愈小,則差度亦愈小.反之,則亦愈大.若樣本增至無限量,或覺與理想之全體相近,則其差度亦愈近,或覺至不相差.故欲減小差度,全賴增加所測之量數及彼此間之σ之減小.

39　取樣時之注意點

取樣的意義,或何謂取樣,吾人已在前節中略述一過.因取樣一道,極爲重要,故特在此言其注意點.在取樣時之最應注意的事,就是應儘量的用隨機的,或儘量的用機會的方法選擇.所謂隨機的選擇者,即當選擇時,須絲毫不存故意之念.換言之,亦即"無黨無偏."例如著手測量時,必須各種學校,各種兒童,各種年級都有.決不能單測量某種兒童或某種學校某種年級.苟如此,則即所謂故意.如略存故意之念則所取之樣本雖多,直等於無用.夫取樣一道,在商業上亦常有之.例如中國販菜子到外國的時候,分裝數千袋或數萬袋.在事實上當然不能一袋一袋的打開來查驗,祇能請有經驗者,任取數袋,加以觀察.假使觀察者以爲不錯或很好,則其餘的當然也不錯或很好.假使有一二奸商,存心取巧,把面上的若干袋,或已裝入的若干袋,裝著好的菜子,把下層的若干袋,或未裝入的若干袋混以細沙,(即存故意之念)則其結果樣子不能代表全體,外商因此扣付貨款,或全體將原物退還,要求賠償一切的損失,甚至其他外商,因此生懼而裹足,從此銷路減少,或竟至斷絕,營業完全失敗,此可謂咎由自取.在測驗上之選擇樣子之不能代表全體,雖不若商業上失敗之鉅,然而損鉅款,費時間,傷精神,徒勞無益,亦殊不值得.故取樣之絕對不應存故意之念,而應隨機選擇,實爲最應注意之一事也.

40　可靠性之用標準差表示者

欲知實得量數之可靠性之大小,應知實得數與眞正數之差數之大小,此吾人已略述如上列第37節中.表示差數之大小,平常可有兩種方法:一爲標準差(σ),一爲機誤(P. E.).茲先將用標準差表示之公式,分述之於本節中.論及公式之來源,若用數理引申出來,則頗冗長,且亦深奧而不便於初學.不過明瞭以上37節中的意義,則應用公式,亦並不見困難.茲錄其各種公式如下——

1　求算術平均數之可靠性之用標準差表示之公式(即求實得算術平均數與眞正算術平均數之差數之標準差

$$\sigma_m = \frac{\sigma_{dis}}{\sqrt{N}}$$

在此公式內,σm 爲實得算術平均數與眞正算術平均數之差數之標準差.σdis (σdistribution) 爲全分配之標準差.N爲次數之總和.茲試用下例以說明此公式之用法及意義.

譬如欲調查一千個學生的算術四則成績,分成十組測驗,其各組結果如下:——

表三十五　一千個學生分成十組測驗的算術四則成績

組別	算術平均數	標準差
1	35	10
2	33	11
3	34	13
4	37	12
5	35	14
6	36	13
7	34	13
8	36	12
9	35	13
10	35	9

平均數之平均數=35

標準差之平均數=12

平均數標準差=1·095

標準差的標準差=1·48

上列第三十五表上之平均數標準差,是十組平均數的標準差,可以表示十組平均數的可靠性.標準差的標準差,是十組標準差的標準差,可以表示十組標準差的可靠性.同理,也可以求中數的標準差,四分差的標準差,相關係數的標準差,用以表示中數,四分差,相關係數的

可靠性.

假使有了求算術平均數之可靠性之用標準差表示之公式,那末不必把全體一千人分成十組測驗,祇要有一組的測驗求得其平均數及標準差以後,卽可用公式求平均數標準差.譬如將上表的各組測驗,隨意用第一組的結果,代入公式,則爲:——

$$\sigma_m=\frac{\sigma \text{dis}}{\sqrt{N}}=\frac{10}{\sqrt{100}}=1\cdot 0$$

從此公式而求得之平均數標準差,是1·0.此1·0,比實在測驗十組後而求得之平均數標準差 1·095,相差不過爲·095.然而可以省去九組的測驗,和九組的統計的勞力了.玆再將上述第六章第17節第十二表上的算術平均數,來計算其可靠性,並解釋之如下:——

第十二表上之$m=61\cdot 97$, $\sigma=12\cdot 51$, $N=180$.

依公式求算,則得:——

$$\sigma_m=\frac{12\cdot 51}{\sqrt{180}}=\frac{12\cdot 51}{13\cdot 416}=\cdot 93$$

求得此$\sigma_m=\cdot 93$以後,此算術平均數61·97,究竟有多少可靠,或可靠性有多大,此爲吾人所急欲知道者.在統計學上之習慣用法,多以±3σ,爲確實限度.因從 −3σ至 + 3σ,其中含有全體量數百分之 99·73.故平常均以±3σ算的確.(中數之確實限度,爲±3σ_{md},標準差之確實限度爲±3σ_σ,相關係數的確實限度爲

±3 σ_r.）我們既知 σ_m 爲·93,則此實得算術平均數61·97,和眞正算術平均數的出入,就是61·97±3(·93).換言之,亦卽眞正的算術平均數在(61·97—2·79)至(61·97+2·79)之間,或卽不出於59·18至64·76之間.

不過還有一點要請讀者注意,就是所說的眞正的算術平均數,並不是指被測驗的180個學生說的,乃是指180個學生所代表的許多學生而言.蓋此180個學生之算術平均數,已求出爲61·97也.假使此180個學生之選法,最合於隨機之選擇,則其算術平均數61·97有與其餘的許多學生所得之算術平均數適合之可能.惟在實際上,極難得到有此種結果.故就此已得之算術平均數定其餘的許多學生所得的眞正的算術平均數,在59·18至64·76之間,或者說小到59·18,大到64·76.雖云在59·18之下,64·76之上,尙有眞正的算術平均數發現之機會,然爲數卻極少,在一萬次中,僅三次而已.如照理論上講,我們若用±1 σ,那末眞正平均數逃出推算的範圍很多,（其中僅含全量百分之68·268,其機遇爲2·15與1之比.）若用±2 σ,也還不少,（其中僅含全量百分之95·55,機遇爲21與1之比.）用了±3 σ,則機會比較的要算少,（含全量百分之99·73,機遇爲369與1之比.）所以平常多用±3 σ,就是這個意思.其實能用到

±4 σ,（含全量百分之99·994,機遇爲 12819 與 1 之比.）或±5 σ,（含全量百分99·99994,機遇爲 174398 與 1 之比.）當然還要好.此均可於下列第41節第三十七表見之.

以上之所敍述者,全是關於算術平均數一方面的,惟可靠性之公式,不僅限於算術平均數,如中數,四分差,標準差,相關係數,以及兩數之差數等等,統可用來計算.茲特逐一錄其公式如下.至其公式之用法及意義,則均可由上述算術平均數之公式而類推,恕不贅述.

2　求中數之可靠性之用標準差表示之公式（卽求實得中數與眞正中數之差數之標準差）

$$\sigma_{md}=\frac{1\frac{1}{4}\sigma_{dis}}{\sqrt{N}}\text{或}=\frac{1\cdot25331\sigma_{dis}}{\sqrt{N}}.$$

3　求四分差之可靠性之用標準差表示之公式（卽求實得四分差與眞正四分差之差數之標準差）

$$\sigma_{Q}=\frac{1\cdot11\sigma_{dis}}{\sqrt{2N}}$$

4　求標準差之可靠性之用標準差表示之公式（卽求實得標準差與眞正標準差之差數之標準差）

$$\sigma_{\sigma}=\frac{\sigma_{dis}}{\sqrt{2N}}$$

5　求相關係數之可靠性之用標準差表示之公式（卽求實得相關係數與眞正相關係數之差數之標準

差）

a 由皮爾生之公式$\frac{\Sigma xy}{N\sigma_x\sigma_y}$求得之 r 之公式爲:—

$$\sigma_r=\frac{1-r^2}{\sqrt{N}}$$

b 由斯披門等級相關法求得之 ρ 並由 ρ 轉化爲 r 之公式爲:——

$$\sigma_r=\frac{1.05(1-r^2)}{\sqrt{N}}$$

c 由斯披門等級相關之簡捷法求得之 R，並由 R 轉化爲 r 之公式爲:——

$$\sigma_r=\frac{.638}{\sqrt{N}}$$

d 由薛伯特之異號相關法求得之 U，並由 U 轉爲 r 之公式爲:——

$$\sigma_r=\frac{1.68}{\sqrt{N}}$$

6 求兩數之差數之可靠性之用標準差表示之公式爲:——（即求實得兩數之差數與眞正兩數之差數之標準差）

$$\sigma_d=\sqrt{(\sigma_{甲})^2+(\sigma_{乙})^2}\text{ 或}$$

$$=\sqrt{(\sigma\text{measure I})^2+(\sigma\text{measure II})^2}$$

在此公式中，σ_d 爲兩數之差數之標準差. $\sigma_{甲}$ 爲甲測量的所用的數量的標準差.例如所用的爲算術平均數，即爲算術平均數之標準差.所用的爲中數，即爲中數之標準差等等.對此公式，吾人似須先加以解釋.且此公式

亦爲表示可靠性之各種公式中之最重要,而最有用者之一個.在進行教育上之實驗研究時,爲不可少,今特先言其意義.

假使我們欲比較兩種實驗的結果,看此兩種結果之相差,究竟因爲方法不同之關係呢?或是因爲偶然的關係呢?碰到此種情形,那末上述之公式,大有助於我們之解答.例如吾人欲決定兒童初學寫字時之進步之多寡,是用映寫好呢?或是用臨寫好呢?欲決定這個問題,則非實驗不爲功.實驗時,必先將同程度同能力之學生,(最好所有情形皆相等)分爲兩組練習.假定甲組用臨寫練習,名之曰甲法組.乙組用映寫練習,名之曰乙法組.練習了幾個月後,再去測驗他們的進步,視何組較優,即斷定何種方法爲良.今假定兩組的成績如下:——

甲組	$m=81 \cdot 97$	乙組	$m=79 \cdot 97$
	$\sigma= 6 \cdot 66$		$\sigma= 5 \cdot 80$
	$N=81$		$N=100$

一觀上例兩組之成績,則二者進步平均數之差數,爲$81 \cdot 97-79 \cdot 97=2 \cdot 00$.然此差數$2 \cdot 00$,是否可靠.究竟是否甲法優於乙法,或是由於其他偶然的原因,此爲吾人所急欲知道者.又假使此兩組再繼續的實驗下去.甲組所優勝之$2 \cdot 00$,是否要變「0」,或者乙組的進步,反要勝

過甲組.欲解決此種問題,則我們可以先求各組的平均數的標準差:——

$$甲組\ \sigma_m=\frac{6\cdot66}{\sqrt{81}}=\cdot74 \qquad 乙組\ \sigma_m=\frac{5\cdot80}{\sqrt{100}}=\cdot58$$

求得後,再用上列公式求兩數相差之標準差:——

$$\sigma_d=\sqrt{(\cdot74)^2+(\cdot58)^2}=\sqrt{\cdot884}=\cdot94$$

在求得此 ·94 後,我們卽可由此斷定此兩組進步數之眞正差數,乃在實得差數 2·00 ± 3 (·94) 之間,亦卽在 —·84 至 4·82 之間.換言之,亦卽多到 4·82,少到 —·84.也就是甲組或者還要不如乙組 —·84, 但是乙組優勝的機會,卻不及甲組來得多.

上面所述兩數之差數之可靠性的公式,也不僅限於算術平均數,亦可用之於中數.四分差.標準差.相關係數等.不過如用中數時,要把各組的中數的標準差代入.用四分差時,就要把各組的四分差的標準差代入.用標準差及相關係數時,均依此類推.

此種差數之可靠性,在實驗研究方面,爲用最廣,歷來多用上述公式求算,以判斷其確度.但其方法似稍嫌陳舊,且初學者亦甚難瞭解其意義.故麥柯 (W. A. Mc Call) 另創一試驗係數.(Experemental Coefficient)此種試驗係數之公式,比上式更其簡明,求出甚易,且亦可自然的表示任何差數謬誤之程度.茲將用上式所求得之差數之標

準差,代入麥氏之公式並解釋其意義如下:——

麥氏之公式如下:——

$$試驗係數=\frac{差數}{2\cdot78\sigma_d}$$

前例之差數,爲$2\cdot00$ σ_d爲$\cdot94$,以之代入公式,則爲:——

$$試驗係數=\frac{2}{2\cdot78\times\cdot94}=\cdot76$$

這試驗係數$\cdot76$,係由差數$2\cdot00$求得,我們尚不能斷定甲法確優於乙法.因爲照第三十六表上查得實驗係數$\cdot76$,其相近之機會,僅爲75比1.(將$\cdot76$入爲$\cdot8$,則其機會爲75:1.)假使甲組勝乙組之差數,由2進而爲$2\cdot61$,那末我們可以的確說,甲法優於乙法了.因爲由差數$2\cdot61$所求得之試驗係數,爲:$\frac{2\cdot61}{2\cdot78\times\cdot94}=1\cdot0$.試驗係數爲$1\cdot0$時,乃表明適合之確定程度.如查第三十六表,則其相近之機會爲369:1,適與$\pm3\sigma$之機會相等.

所以吾人若欲知道所得之差數有多少機會要變0,或爲負數,那末我們於求得試驗係數後,可按下列第三十六表直接查得之,且亦可直接斷定此差數究竟有多少可靠.第三十六表之數,非僅可以適用於兩個算術平均數之差數,且可適用於任何兩個數量之差數,此爲吾人所須明瞭者.

*表三十六 用機遇法則說明試驗係數

試驗係數	相近之機會
·1	1·6比1
·2	2·5比1
·3	3·9比1
·4	6·5比1
·5	11 比1
·6	20 比1
·7	38 比1
·8	75 比1
·9	160 比1
1·0	369 比1
1·1	930 比1
1·2	2350 比1
1·3	6700 比1
1·4	20000 比1
1·5	67000 比1

*本表轉錄自 W. A. Mc Call: — How to measure in Education, PP. 405, 47表.

41 可靠性之用機誤表示者

I　機誤之意義　機誤在英文曰Probable Error,簡寫爲P. E..在次數分配爲常態時,其值適與四分差相等.惟其用途不同.蓋四分差被採爲差數之用,而機誤則用以度量可靠性,此吾人已於前述第六章第24節中言及之.據塞斯頓 (L. L. Thurstone) 在其統計學綱要書中所言,以爲:*「假使實際上有一個分配,我們去量他的離中趨勢, (Variability) 則此離中趨勢,卽稱之爲四分差;若僅估計其"可有的"或"或有的" (Might-be) 離中趨勢,則卽稱之謂機誤.」總之,此二種概念本來相同,其數值亦相等,特其用法之不同已耳.

* 見 Thurstone:-The Fundmentals of stat'stics,PP.178上

II　機誤與標準差之關係　在前述第六章第28節中,吾人亦已言及當次數分配爲常態時,P. E. 或 $Q.=.6745\sigma$.在習慣上多用P. E.表示可靠性,而用σ者比較爲少.今既知$P.E.=\cdot6745\sigma$,則化σ爲P. E. 的方法,祇要將$\cdot6745$乘σ卽可.因此上列可靠性之用標準差表示之公式,亦可改爲:——

$$(1)\quad P.E._{m}=\frac{\cdot6745\,\sigma_{dis}}{\sqrt{N}}$$

$$(2)\quad P.E._{md}=\frac{\cdot84535\,\sigma_{dis}}{\sqrt{N}}$$

$$(3)\quad P.E._{Q}=\frac{\cdot74868\,\sigma_{dis}}{\sqrt{N}}$$

$$(4)\quad P.E._{\sigma}=\frac{.6745\ \sigma_{dis}}{\sqrt{2N}}$$

$$(5)\quad P.E._{r}=\frac{.6745(1-r^2)}{\sqrt{N}}$$

$$(6)\quad P.E._{d}=.6745\sqrt{(\sigma_{甲})^2+(\sigma_{乙})^2}$$

若用斯披門之等級相關法求ρ,簡捷法求R,及由薛伯特之異號相關法求U,並將ρ,R,U,轉爲r後,如欲算P.E.時,亦可依上列各式之改算法而求得之.惟算ρ,R,U時,往往在人數不出30時用之.照理論上而言,若N小於30時,卽求算其機誤,亦決難有可靠之結果.換言之,就是求出了以後,亦不能遽下結論.故本書將此公式從略.如學者於求得後欲試算之,則可用上列各式之改算法,而求得之.

42 用機遇法則解釋σ及P.E.所表示之可靠度

機遇在英文曰Chances.例如吾人擲銅元一枚,當其落下之時,不是龍面,就是字面,故其得龍與得字之機遇,各爲一半,亦卽爲1與1之比.再如袋內裝紅球三百,綠球一百,掩目而任取之,則得紅球之機遇爲3,得綠球之機遇爲1,而其比例則爲3與1.又如袋內藏紅球一千,綠球一個,則得紅球之機遇極多,爲999,而得綠球之機遇極少,僅不過1.在統計學上,常用此種機遇,以表示事實之可有與否.通

常得有二法,一以標準差表示之,一以機誤表示之.如在一常態分配圖內,在±1 σ之間,共有百分之68·268人數,尙餘百分之31·732人數,則在±1 σ內之事實,其可有之機遇爲2·15,其不可有之機遇爲1,此爲2·15與1之機遇.$\left(\frac{68 \cdot 268}{31 \cdot 732}=2 \cdot 151\right.$,以$68 \cdot 268 : 31 \cdot 732 = 2 \cdot 151 : x$, $x = \left.\frac{31 \cdot 732 \times 2 \cdot 151}{68 \cdot 268}=1.\right)$又如在±2 σ之間,共有百分之95·55人數,尙餘百分之4·45人數,則在±2 σ內之事實,其可有之機遇爲21,其不可有之機遇爲1,此爲21與1之機遇$\left(\frac{95 \cdot 55}{4 \cdot 45}=21 \cdot 4.\right)$再如在±1 P.E.之間,共有百分之50人數,尙餘百分之50人數,則在±1 P.E.內之事實,其可有之機遇爲1,其不可有之機遇亦爲1,此爲1與1之機遇$\left(\frac{50}{50}=1.\right)$又如±2 P.E.之間,共有百分之82·26人數,尙餘百分之17·74人數,則在±2 P.E.內之事實,其可有之機遇爲4·6,其不可有之機遇爲1,此爲4·6與1之機遇$\left.\frac{82 \cdot 26}{17 \cdot 74}=4 \cdot 6.\right)$茲將此二種計算機遇之方式,列表如下列第三十七及第三十八兩表,藉資檢查,以驗其所表示之確度.

表三十七　機遇之用標準差表示者

範　　圍	機　　遇

土 1 σ	2·15——1
土 2 σ	21——1
土 3 σ	369——1
土 4 σ	12819——1
土 5 σ	174398——1

表三十八　機遇之用機誤表示者

範圍	機遇
土 1 P.E.	1——1
土 2 P.E.	4·5——1
土 3 P.E.	21——1
土 4 P.E.	142——1
土 5 P.E.	1310——1
土 6 P.E.	19200——1
土 7 P.E.	420000——1
土 8 P.E.	17000000——1
土 9 P.E.	1000000000——1

練習問題

1　何謂可靠性?

2　取樣時所應注意的爲何事?

3　設有一統計結果，$N=169$，$m=72\cdot46$，$md=72\cdot32$，$Q=7\cdot2$，$\sigma=12\cdot48$，$r=\cdot69$，（用皮爾生之乘積率法求得）試求各數量之P.E..

4　設有A B兩測量；A量之　$md=71\cdot24$，$\sigma=14\cdot32$，$N=112$，B量之　$md=74\cdot36$，$\sigma=16\cdot32$，$N=144$，試依照麥氏試驗係數之公式，求試驗係數。

附 表 一

由 ρ 之價值求 r

$$r=2\sin\left(\frac{\pi}{6}\rho\right) \qquad \rho=1-\frac{6\,\Sigma D^2}{N(N^2-1)}$$

ρ	r	ρ	r	ρ	r	ρ	r
·01	·0105	·26	·2714	·51	·5277	·76	·7750
·02	·0209	·27	·2818	·52	·5378	·77	·7847
·03	·0314	·28	·2922	·53	·5479	·78	·7943
·04	·0419	·29	·3025	·54	·5580	·79	·8039
·05	·0524	·30	·3129	·55	·5680	·80	·8135
·06	·0628	·31	·3232	·56	·5781	·81	·8230
·07	·0733	·32	·3335	·57	·5881	·82	·8325
·08	·0838	·33	·3439	·58	·5981	·83	·8421
·09	·0942	·34	·3542	·59	·6081	·84	·8516
·10	·1047	·35	·3645	·60	·6180	·85	·8610
·11	·1151	·36	·3748	·61	·6280	·86	·8705
·12	·1256	·37	·3850	·62	·6379	·87	·8799
·13	·1360	·38	·3935	·63	·6478	·88	·8893
·14	·1465	·39	·4065	·64	·6577	·89	·8986
·15	·1569	·40	·4158	·65	·6676	·90	·9080
·16	·1674	·41	·4261	·66	·6775	·91	·9173
·17	·1778	·42	·4363	·67	·6873	·92	·9269
·18	·1882	·43	·4465	·68	·6971	·93	·9359
·19	·1986	·44	·4567	·69	·7069	·94	·9451
·20	·2091	·45	·4669	·70	·7167	·95	·9543
·21	·2195	·46	·4771	·71	·7265	·96	·9635
·22	·2299	·47	·4872	·72	·7363	·97	·9727
·23	·2403	·48	·4973	·73	·7460	·98	·9818
·24	·2507	·49	·5075	·74	·7557	·99	·9909
·25	·2611	·50	·5176	·75	·7654	1·00	1·0000

附表二

由R之價值求r

$r=2\cos\frac{\pi}{3}(1-R)-1$ $R=1-\frac{6\,\Sigma G}{N^2-1}$

R	r	R	r	R	r	R	r
·00	·000	·26	·429	·51	·742	·76	·937
·01	·018	·27	·444	·52	·753	·77	·942
·02	·036	·28	·458	·53	·763	·78	·947
·03	·054	·29	·472	·54	·772	·79	·952
·04	·071	·30	·486	·55	·782	·80	·956
·05	·089	·31	·500	·56	·791	·81	·961
·06	·107	·32	·514	·57	·801	·82	·965
·07	·124	·33	·528	·58	·810	·83	·968
·08	·141	·34	·541	·59	·818	·84	·972
·09	·158	·35	·554	·60	·827	·85	·975
·10	·176	·36	·567	·61	·836	·86	·979
·11	·192	·37	·580	·62	·844	·87	·981
·12	·209	·38	·593	·63	·852	·88	·984
·13	·226	·39	·606	·64	·860	·89	·987
·14	·242	·40	·618	·65	·867	·90	·989
·15	·259	·41	·630	·66	·875	·91	·991
·16	·275	·42	·642	·67	·882	·92	·993
·17	·291	·43	·654	·68	·889	·93	·995
·18	·307	·44	·666	·69	·896	·94	·996
·19	·323	·45	·677	·70	·902	·95	·997
·20	·338	·46	·689	·71	·908	·96	·998
·21	·354	·47	·700	·72	·915	·97	·999
·22	·369	·48	·711	·73	·921	·98	·9996
·23	·384	·49	·721	·74	·926	·99	·9999
·24	·399	·50	·732	·75	·932	1·00	1·0000
·25	·414						

附 表 三

由U之價值求r.

U爲異號各對差數總數之百分比例數.此表所列U之價值,係從·00起至·50止,而與r之正數各相當價值並列.若U過·50至1·00時,則r爲負數,可由相反之關係求r之價值.

U	r	U	r	U	r	U	r
·00	1·0000	·13	·9174	·26	·6848	·39	·3387
·01	·9996	·14	·9044	·27	·6615	·40	·3089
·02	·9982	·15	·8905	·28	·6375	·41	·2788
·03	·9958	·16	·8757	·29	·6129	·42	·2485
·04	·9924	·17	·8602	·30	·5877	·43	·2180
·05	·9880	·18	·8439	·31	·5620	·44	·1873
·06	·9826	·19	·8268	·32	·5358	·45	·1564
·07	·9762	·20	·8089	·33	·5091	·46	·1253
·08	·9688	·21	·7902	·34	·4819	·47	·0941
·09	·9604	·22	·7707	·35	·4542	·48	·0628
·10	[illegible]	·23	·7504	·36	·4260	·49	·0314
·11	·9407	·24	·7293	·37	·3973	·50	·0000
·12	·9295	·25	·7074	·38	·3682		

中華民國廿四年拾月四日　收到

書名	教育統計學
編著者	湯鴻翥
出版者	上海公平路三十四號 大華書局
印刷者	上海南成都路新大沽路口 國光印書局
出版日期	中華民國二十三年 二 月初版 中華民國二十　年　月　版
裝訂冊數	平裝一冊
定價	大洋六角
總發行所	上海公平路三十四號 大華書局
分發行所	全國各大書局
本書編號	153