序　一

闡發事理之要端曰文曰數我國向富於文的觀念而貧於數的觀念故文字之傳播汗牛充棟求一用數理闡明事物者除天文律呂及簡單之戶口册外殆絕無之近日各衙署雖頒有統計報告然極遲緩且不盡眞確始予亦頗不謂然乃從政教育部以來首提議徵各省普通教育統計雖經部通行各省其遲緩且過於向所經歷此固由於政治之關係然省責之縣縣責之學校均含有比較公文書略輕之感念可斷言也然教育之實際非有精確之統計殆無由改良而進步需要與普通之感念相差若此則教育行政之隱憂自亟且深懼也薛子乃乘此時機首編教育統計學大綱以應教育社會之需要若從事教育者能人手一編罩感於數的觀念之重要斯誠教育界大幸也夫

中華民國十一年十二月上旬陳寶泉謹序

序　二

自科學的精神和方法應用到教育的問題上來，統計法就成爲研究教育的重要工具。有了這種工具，教育學者，才能從複雜的事實中尋出個條理來；從偶然的景象中找出個原則來；從枝枝節節的現象中理出個因果或連帶的關係來。研究教育的人，幾乎一刻不能離開這個工具。要想在教育方面發古人所未發，明今人所未明，非直接或間接借重統計法不可。就是辦教育行政或當教員的，要想知道自己所用方法的效力如何，也得曉得些統計法。統計法能使眼界小的變大些，眼界近的變遠些，眼界偏的變完全些。他的功用，如同顯微鏡，千里鏡，無非是要補足我們肉眼的缺點。他把教育界的種種現象縮起影來，使我們看得清楚。他幫助我們找眞理，但祇是一種工具。有時人把統計太看重了，做起文章來，滿篇的數目字；做起書來，十分九的篇幅都爲數目字佔去。他們拿工具當作目的看待，未免喧賓奪主了。這是我們研究統計的人應當注意的。教育統計法輸入中國，到現在祇有四年。在高等專門大學中曾經担任這門功課的人，有張

耀翔，張見安，朱君毅和我四個人。社會上對於這門學問也漸漸的覺得他的需要了。我們因為事務很忙，幾年來還沒有得到機會，做本適用的書來報答社會。今薛君遠舉用了二年的時候，採集Rugg, Yule, Thorndike, Brown, McCall, Whipple諸家的著作的菁華，輯成此書。他的毅力，眞可令人佩服。薛君在校時最喜歡研究統計；畢業後，應中華教育改進社之聘，掌理統計事宜，將近一年。此書是他的學問經驗融合的出產品。讀者當能得到好多的益處。這是我所深信的。

民國十一年十二月十四日　陶知行

自 叙

統計的方法,可該括的分爲表記(tabular methods),圖示 (graphic methods), 分析 (analytic methods) 三種。表是序列分散的事實, 圖是便利事實的表顯。這二種方法的用處很廣 。但是這本書未將他們列入僅敍述分析方法的緣故 , 是因爲分析法的材料容易搜集,而其法則已有一定不易的成規了。至於表圖的作法,雖然亦有規則, 但是變動的地步較大些; 且在中國此種學術幼稚的時期,更不易得到相當的材料,作應用方面的標準。所以當我把分析法草成的時候,表圖二種方法,尙未研究有滿意的結果。我希望此書再版的時候,此二部分,或者可以加入。

這本書的名稱 , 雖然名爲「教育統計學」, 但若嚴格說起來 , 其意義卽是「應用於教育上的統計法(Statistical Methods Applied to Education)」。因爲統計法早已作一種研究學術的普通工具 (參看第一章緒論), 現在又作研究教育的利器了。所以這本書裏的方法,雖然是偏重教育方面的應用,而研究他種應用統計方法的科學,也可以用作參考。

這本書的篇章，是本李湘宸張見安兩先生課堂講授的次第，並參考下列各書編成的：

Statistical Methods Applied to Education, By Harold O. Rugg.

An Introduction to the Theory of Statistics. By G. Udny Yule.

The Introduction to the Theory of Mental and Social Measurements. By Edward L. Thorndike

The Essentials of Mental Measurements. By William Brown.

How to Measure in Education. By william A. McCall.

Manual of Mental and Physical Tests. By Guy Montrose Whipple.

最小二乘法　無錫顧澄先生譯著

讀者諸君，若想作探本窮源的研究，可參考上列各書，或他種關於統計的著作。

這本書中說明的材料，多半由上列各書選擇來的。選來的材料，多半是於教育有關係的；並且注重應用方面。關於公式的用法，則多作解釋；關於公式的來

源,重要的和淺近的則加以說明,非重要的和深奧的則概從簡略。這是希望粗通數理的教育家,也可利用這些法則,作研究的工具。

這本書中的名辭,是採取上列各書中所通用的。中文的譯語,多由各數學書及各雜誌中採取來的。無從採取的名詞,則由著者創譯。故皆附以原文,預備讀者對勘。不妥的地方,必不能免。讀者諸君若能不吝教正,是無任歡迎的。

這本書的成功,藉重課業教師及同學友人的力量很多:李湘宸,張見安,張耀翔諸先生的指示校訂;同學王卓然的激勵勸勉,方永蒸,夏宇衆的商榷訂正;吳兪三的剖析數理;友人蕭偉一,尹彤墀,楊可大的抄對稿件,核算例題,皆是此書作成的原素。是以特誌書端,藉申謝意。

中華民國十一年十二月十七日

薛鴻志叙於北京中華教育改進社

目 錄

第一章 緒論

頁數

第二章 教育測量

第三章 統計材料之分類法

第八章 常則概率曲線之應用

第九章 證確數量 (reliability measures)

第一章
緒論

統計術語之變遷 西文統計(Statistics)一語，係由拉丁字 Status 而來。按拉丁字 Status 之意，係表示國家之政治情形，及其所處之實際地位。是統計之學，原為考察國家情況之學也。然其初記載國家情況之特點，僅以文言。自十八世紀中期，統計上所公布之材料，始漸有簡明數字之記載。至十九世紀以後，政治之材料益豐，數字記載之法益進步，於是數字統計方法，遂代文言以興。文言既變為數字，統計一語，遂含以數字表明分量之意義矣。意義既變，其功用之範圍，亦因之推廣。昔日之統計專論國情，至是則凡搜集數字材料之科學，皆須借重於統計。如化學，氣象學，生物學，心理學，遺傳學等，皆以統計之方法為研究之行徑矣。

統計學之應用 顧何以性質極不相同之各科學，而能適用此同一的統計之方法乎？是蓋因各種科學有一公同之性質在。其所具公同之性質為何？即被測量之各種事實，其結果為變異的，非確定的也。

今試就物質科學言之。化學之實驗，據吾人之常識

所知,應有定量不移之結果矣。然舉行實驗之時,如測酸素與水素之原子量,各次測量之結果,雖不甚差,必不能完全盡同。所謂酸素一原子量十六倍於水素者,蓋言其極相近似之値耳。此種差異結果之起因,或由於觀察之人;或由於觀察之器具;或由於溫度,濕度之變動;壓力,引力,及震動力之影響:其原因之複雜,雖極精於實驗之人,亦不能操縱自如,使之全行消滅;或僅使其中某一原因存在,以求理想中之結果。故必藉助統計方法,將各次實驗之結果,綜合而考察之,方可得概然之定律也。

又如治遺傳學者高爾登(Galton)主張「遺傳趨常」之說。用統計方法,研究英人身體之高度。研究之結果,體高六英尺之英人,其子之體高平均不過五英尺十寸三分;體高五英尺六寸者,其子體高平均却有五英尺八寸三分。因英國人平均之體高,在五英尺六寸和六英尺之間也。此類研究,不能就少數人之身體高度,卽下斷定。必須統多量之人數而觀察之,始可得此普徧之規律。

要而言之,凡關於人事的各種科學,其所函之現象,至爲繁複。一果之生,或由於衆因結合之影響,或由

於衆因中任意一因之勢力。造果之因無常，則所生之果，亦不能確定不變。若僅就諸結果中任一結果以爲斷言，而能得其眞者，苟非出於偶然，决無其事。故須統合各個結果分析之，比較之，馭之以概然之數，以推求其間所呈顯之常理。統計學之所以見重於研究科學之士者，有由然矣。

教育統計學之需要　每一科學之建設，必根據以下三事：(1)理論，(2)實驗，(3)結果之試驗。既有理論，必須施諸實驗，以備證明理論之當否。既施諸實驗，必須考察實驗所生之結果，與理論有無乖違，以斷定理論之是否正確。教育之有理論，自古已然。若以理論施諸實驗，用標準試驗 (Standard tests) 以測量其效果，則爲近世所僅有。然研究之方法既變，則改進之速率驟增。故今日之教育，浸浸然有成爲一種科學之勢矣。顧實驗結果之測量，與標準試驗之製作法，皆非統計法則，不能竟其功也。

用科學的方法研究教育，固不僅恃實驗已也；調查事實，亦極重要。如各地方之教育實況，各學校之教學成績，此等固有之材料，皆極有研究之價值。若能搜集各種事實，用統計方法，類列以整理之，分析而考察之；雖各種事實，紛紜錯雜，視之若茫無頭緒；然

提綱絜領，亦莫不順事成章，顯其固有之規律。本乎此類規律，以解決各種教育問題，罕有不當於實際者也。應行調查各問題，舉其大者，可分下列數種：

(1) 關於學生方面者；如年齡，年級之記錄，升級，留級之方法，退學，轉學之章程等事。

(2) 關於課程方面者；如教科書之內容，地方之風俗習慣，及社會生活之需要等事。

(3) 關於教師方面者；如教師之學識，經驗，委任方法，薪俸數目，及教師家庭中之職業等事。

(4) 關於教授方面者；如學生進步之速率，分量，限度，及形式之陶冶等事。

(5) 關於教育行政方面者；如行政機關之組織，人員之資格及薪俸，用人法則，職務情形等事。

(6) 關於經費方面者；如歲入，歲出，來源，用途，及學款之支配等事。

此類教育調查，亦爲近世教育趨勢中重要之一。然在吾國舉行之者甚少。材料旣形缺乏，難爲寔例之說明。故此書姑先略去此篇，僅就數量(measure)計算之法則，於第三章後分別解釋之。第二章就教育測量方法，敘其大意。製作標準試驗之法則，將於第八章28節略述之。

第二章
教育測量

1 測量之應用

測量一事,爲吾人日常生活不可少之方法。如言輕重,長短等等,皆所以測量某事物之分量也。然苟試問所言之輕重究有若干,非以稱及天秤等器具測之,難得輕重之確數也。苟試問所言之長短究有若干,非以尺度測之,亦難得長短之確數也。測量長短輕重等之器具,吾人用之最多,習以爲常,若不覺其最爲便利者。然使無此種器具,而任各個人主觀之判斷,以計物之輕重長短,其相差之度,必不可以道里計也。

2 主觀之測量

物質方面,能用客觀之標準,定多寡之分量。精神方面亦可用此同一方法,定測度分量之標準。或謂精神方面之特質,隱微奧秘,不若物質方面之事物,顯然可見,用分量之方法表出之,殊不易易。不知吾人日常對於此種隱微之特質,亦有分量之估計。如言此童之聰明强於彼童。甲生之品行優於乙生等類。聰明,品行等,既有比較,有差異,即可用分量表明之,特

所定之分量無客觀之標準耳。對於聰明，品行等優劣之批評，各人之意見，必難相合；甲以爲優，乙或反以爲劣。因各人自有其主觀之標準，故所定之分量，亦因之差異。由是言之，精神方面之性質非不可測，特乏精確之測量法則耳。

教育方面，憑主觀以評定優劣之事，不一而足。其判斷相差之大，有爲吾人所料想不到者。果用試驗方法證明之，即可明其眞象矣。吾師張耀翔先生爲北京高師心理學教授，曾研究憑主觀定分數之差異。用北京高師附小一年級某生之國文一篇，原文一字未改，令某高師學生一百十五人評定。以零分爲最低，以十分爲最高。玆錄其題目及原文如下：

題目：諸生自述在家溫課之狀況。

原作：人生於世，若通達文字，則讀書而已。然讀書而不溫，與不讀異乎。故吾每日歸家之後，四時至六時，即吾溫課之時，無一不加謹愼者也。

此一百十五人所評定之分數，列於第一表。

就第一表所列可見各評定者之意見不一。給一分者有三人，給二分者有五人；而給八分者有二十二人，約佔全體人數百分之十九；給九分者尙有二十一

人,約佔全體數百分之十八。就各人所給之分數。平均計算,得平均數(mean) 6.6,再就各人所給分數差異之程度求之,得標準差數(Standard deviation) 2.06

第一表 一百十五人評定某一高小學生國文之分數

分數	人數	人數百分比	
10	0	—	
9	21	18.26	平均數=6.6
8	22	19.13	
7	27	23.48	標準差數=2.06
6	21	18.26	
5	8	6.95	
4	2	1.74	
3	6	5.22	
2	5	4.35	
1	3	2.61	
0	0	—	
總數	115		

上述之試驗,猶可說國文之價值,分別細微,不易評定,故有若此之差異程度;若評定算術之分數,當不至此。實則不然。張先生亦曾擬一算術試驗,以證明給算術分數之差異程度。茲將先生試驗之方法,錄之如下:「余(張先生自稱)故意造一極簡單之數學題,使評論員人人理會,人人自已能算,故人人能察覺他人算之對與不對。乃按題作一算式,將算法完全作

對，不過將答作錯。猶恐評論員有疏忽，乃另將正答書出，加以叮嚀，使評論員共知錯處所在，然後再定分數。」算題與算式列下：

算題：今有銀洋五萬九千六百七十七元五角六分，四百八十九人分之，問每人應得若干？

算式

```
489)59677.56(122.4
    489
    1077
     978
      997
      978
      1956
      1956
```

正答＝122.04

上列算式曾經某高師學生一百十四人評定分數。以十分爲最高，以零分爲最低。此一百十四人所給之分數，列於第二表：

第二表　一百十四人評定算術之分數

分數	人數	人數百分比
10	1	.87
9	8	7.01
8	13	11.43
7	13	11.43
6	21	18.41
5	27	23.66
4	13	11.43
3	8	7.01
2	8	7.01
1	1	.87
0	1	.87
總數	114	

平均數＝5.535
標準差數＝2.024

就第二表所列，各人所給之分數，其不同之範圍更大。給零分者，尚有一人；而給十分者，亦有一人。平均數爲 5.535，標準差數爲 2.026。再就第一表和第二表差異程度之大小比較之，其結果如第三表：

第三表 比較第一表和第二表差異之程度

	第一表	第二表
平均數 M	6.6	5.535
標準差數 σ	2.06	2.026
$\frac{\sigma}{M}100$	31.2	36.6
差數	5.4	

就第三表言之，第一表差異之程度爲百分之31.2，第二表差異之程度爲百分之 36.6。第二表差異之程度，較第一表大百分之 5.4。由此可見憑主觀評定算術分數，亦不一致。就此二表之材料言之，其精密確塞與否姑不論；評定算術分數之差異，是否必大於評定國文分數，亦姑且不論。所可斷言者，卽凡任個人之意見，評定算術國文分數時，同有紛歧之主張。並可推知憑主觀評定他種科學分數或各類事體時，各人所下之判斷，彼此亦極難相同。是蓋因各人所定之

標準不同，故根據此標準所定之價值，亦因之紛歧。

3 客觀之測量

(1)測量表(Scale) 主觀之測量，既少公同之標準，卽難得相同之意見。藉或有相同之時，亦僅出於偶然。故欲求各個人之意見劃一，得正確測量之結果，必須以客觀之標準爲標準。如測量長度，世人皆以巴黎博物院中所保存之米達(Meter)尺爲標準，因此可以知二米達爲若干長；三，四，五米達各爲若干長。又如測量溫度，以華氏寒暑表之度數爲標準，可以知溫度 32 度爲若干，75 度爲若干。有公認之標準，遂無紛歧之意見。長短，寒暖等測量器具發明較早者，因各科學家之思想見解，易趨一致故也。若智力，道德，行爲，以及社會情狀等等，用客觀之標準測量較難，故其測量之工具，發明亦較晚。近世關於心理，教育等測量表之發明尙少。茲試舉其一二：測量智慧之表，如皮奈西門之智慧測量表 (Binet - Simon intelligence scale)，分人之智力爲若干年齡，以測人之聰明高低。測量算術能力之表，如吳德(Woody)，考提斯(Coutis)（皆美國人）所作之算術測量表，按算題之難易，列成等級，訂定分數，以考驗學生之算

術能力。此類標準測量表，在中國教育界中，尙未有可靠之製作也。

(2)單位(Unit)　物質方面計算某物之多寡，必先以同類之量定其標準，然後以此標準計其物之分量。此先定之標準謂之單位。如：

計算長度，以尺爲單位時，可言某物長若干尺。

以丈爲單位時，可言某物長若干丈。

計算容量，以升爲單位時，可言某物有若干升。

以斗爲單位時，可言某物有若干斗。

精神方面之測量，亦須有一定之單位，以計算所測量之分量。如試驗學生識字能力，以單字爲單位，卽按識字之數目定能力之高低。又如試驗算術能力，以算題之數目爲單位，卽按作對題數之多寡，定能力之優劣。凡學校中通常考試時，所用單位之標準，大抵類此。然此種單位極不正確。如單字之難易，各個不同。若以每一單字爲一單位，則此單位之量，已不相同，焉能得正確測量之結果。算術題目，亦復如是。其難易之程度，參差不齊。如下列二題：

(a) 13)65065(,　(b) .003).0936(

孰難孰易，抑其難易相等。若不將其難易之程度，先

行定準，僅以一題為一單位；而謂能作二題之兒童，其能力卽倍於作一題者之能力，其不可信，彰然明矣。

是以精神方面測量所用之單位，其價値必須相等，否則必不應用。如前節所舉之各測量表，皆用等値之單位製成者（見後28節）。惟各表所用之單位，亦彼此不同。有以標準差為單位者，有以概誤差(Probabl error)為單位者，其不劃一之情形，亦猶物質方面各種測量器具，所用之單位，彼此互異。如法國每1米達，為華尺3.125尺；每1竍(liter)，當華之305.1758立方寸；其他如碼，哩，磅，噸，先令，辨士等量，其單位皆甚歧異，學算術諸等法者，概能言其制度之複雜矣。

(3)標準點(Reference Point)　測量之必須有標準點，亦猶其必須有單位也。無論何種測量皆須有起始之點，作分別等級之標準。如地理家定北京之位置，苟無一定之標準，則無從說明。若用經過格林威池(Greenwich)之經度線及與赤道相重之緯度線為標準，言北京在東經度若干度，在北緯度若干度，則北京之位置，可以確定。如測身體之高低，則以赤足之底或露頂之頭為標準。又如測陸地之高低，則以海洋

面爲標準。特科學家所定之標準各有不同。如溫度測量表，攝氏以結冰之點爲零點，華氏則以冰點下32度爲零點，而列氏復以冰點爲零點。蓋當一種科學進步之始，各科學家所用之方法，多憑主觀之意見。迄愈臻進步，各主觀之意見，愈趨於同，則公共的客觀標準，於是立焉。客觀之標準既立，則可適用於一般。故以攝氏寒暑表測量溫度，以英尺測陸地之高時，若言若干度和若干尺，雖異時異地之人，皆能了解此計算之意義矣。

精神方面之測量法，現方在幼稚時期，製作標準試驗者所用之標準，紛歧不一，自屬當然。如(a)某種試驗不分難易等級時，即以無分爲標準點。(b)某種測量表以作者所推測之零度爲標準點。(c)有以平均數下3或2.5標準差爲標準點者。(d)亦有以最低之分數爲標準點者。其在吾國，尙乏通行之標準試驗。研究之士，正可於此種試驗製作之始，特加注意，以求各種測量表之劃一。否則亦將如吾國錢幣，度，量，衡等制度，駁雜紛繁，致學術進步之研究，發生困難。而考究各種測量之材料，亦宜注意及此，然後可不致誤謬。例如甲童比60磅重10磅，乙童比60磅重2磅，若不計其60磅之標準，僅言甲童比乙童重五倍，稍有常識之人，即知其不合於事實矣。

第 三 章

統計材料之分類法

4 次數分配(distribution of frequency)

有統計之材料而欲整理說明之，其第一步，卽爲列表；將參差不齊之材料，作有系統的分類。然後卽可按表研究其材料之內容。試舉例說明之：

設有一人受暫時記憶之試驗，於一定之時間內，使之記字十二個，共試驗四十次，每次所能記之字數多寡不等。其各次之數目如第四表中所列：

第四表 一人於一定時間內所能記之字數，共記四十次

7	10	6	6	6	10	1	6	6
9	4	8	7	11	7	7	6	9
5	4	7	5	8	10	6	6	7
7	8	3	6	10	4	6	2	8
8	8	8	9					

此四十次中每次所記之字數，有相同者，有不相同者。相同之數，如記4字共三次；記5字共二次；記6字共十次等等。不相同之數，如記1字僅有一次；記2字僅有一次；記3,11字亦各僅有一次。欲將此四十次所記之字數，作有系統的排列，卽將各類數量，由

小而大，自上而下（自下而上亦可），由1起至11止，依次一一列於左方數量（measure，簡寫爲m，）欄之下，如第五表：

第五表　排列第四表之材料

數量 m	排列	次數 f
1	丨	1
2	丨	1
3	丨	1
4	丨丨丨	3
5	丨丨	2
6	卌　卌	10
7	卌　丨丨	7
8	卌　丨丨	7
9	丨丨丨	3
10	丨丨丨丨	4
11	丨	1
總數		40

然後循第四表之數字，依次檢查。該表起始之數字爲7，卽在數量欄中7字之右方畫一豎線；第二數字爲10，卽在10字右方畫一豎線；第三，第四，第五三數字皆爲6，卽在6字右方畫三豎線。每一數量畫四個豎線之後，若再須畫線時，則在已畫之四豎線上，畫一斜線，使每組之線數爲五，以便總計。依此法排列至表終爲止，其排列結果如第五表中央一欄所表明

者。然後再將所畫之豎線核算，以數字表出之，列於畫線右方次數(frequency，簡寫為f)欄之下。由是可見記一字之次數為1，記二字之次數亦為1；記三字，四字之次數各為1及3等等。再將所列之次數相加，其和為40，與原四十次之數相符。惟考察所排列有無錯誤，則頗覺困難。若將各個量數，書於方格紙塊之上，以一方格代一數量，然後將同類之紙塊集成一組，數各組紙塊之數目，以核對排列之有無錯誤，較易易也。

前舉之例，亦可由最小之數量起，按各個數量之數值，一一單列，排成一次序分配 (order distribution)，如；1, 2, 3, 4, 4, 4, 5, 5, 6,6, 6, 6, 6, 6, 6, 6, 6, 6, 等等至11為止。此種排列亦較第四表，更為清楚。第當數量較少之時用之則可；若如第六表所列之一百二十三名中學學生之英文分數，其數量總數共有一百二十三；如將各個分數順次一一列排，則表式太長，不宜於用。故須用分組方法，使數量之相近者列為一組，將全體數量分為若干組，化繁為簡，以利推算。此種分組之手續，約有下列四步。明此四步法則，則一切材料皆可整理就緒，作推算之基礎：(1)檢查最小數量與最大數量中間之距離；(2) 定每組組距 (classinterval)之大小;(3)定每組組距之界限;(4)排列每組中數量之次數。由此四種方法求得之結果,即爲次數分配。各步之方法分述如下:

(1)檢查最小數量與最大數量中間之距離 試以第六表所載之材料說明之。一百二十三名學生之英文分數,多寡不等,最少之分數爲20,最多之分數爲95,中間相差之數爲75。此75即爲最小數量與最大數量中間之距離。既求得距離之後,即依第(2)法定組距之大小。

第六表 一百二十三名中學學生之英文分數

80	57	45	74	95	80	73	87	59	80	57	52
75	75	63	75	84	50	77	76	63	90	79	80
58	71	60	85	76	76	72	73	56	75	84	80
87	85	69	85	40	66	78	79	73	86	88	75
80	79	80	60	87	80	78	82	52	75	67	80
77	80	66	74	73	79	60	66	57	74	76	70
55	87	87	72	73	68	87	81	60	75	35	73
75	67	78	86	73	79	40	82	55	65	80	86
79	65	73	56	71	73	80	67	78	62	79	79
81	77	82	78	93	78	70	72	79	45	81	75
20	80	30									

(2)定組距之大小 設全部數量中間之距離,單位

甚少,如僅有10單位,15單位,或多至20單位,則無須再求簡捷,將單位歸併。如前第四表之材料,其全部單位之距離由1起至11止,僅有十一個單位,即用此自然之單位,分類排列,無庸另行規定組距。然如第六表之材料,單位之數有75。若仍用其自然之單位,則計算頗繁。自然之單位,既不應用,勢必須另定組距,以求簡捷。規定組距之時:第一,各組組距之大小必須相等。第二,須不失精確之程度;每一組距所含之數量,其平均之價值,須與該組中值(midvalue)之數,無大差異。而後據此分組之材料,求平均數,差異數等,所得之數,方足以代表原有之材料;其事實之眞象,方不致消失。蓋此種分組法則,係假定每一組距之中,各個數量之價值,皆與該組之中值相等。如第七表,45.0-49.99組距之間,所含之二數量,其平均數與中值47.5相差無多;50.0-54.99組距之間,所含之三數量,其平均數與中值相差更少。此種假定,若當每組數量增多時,更可與事實相符。第三,須力求簡捷。凡組距愈大,則所分之組數愈少,計算時亦愈省力。故但求不背精確之程度,能使組距之度愈長,愈利於計算。通常適當之組數,大概在二十五與十之

間。若組數過三十，則不免繁瑣；少於十，則易失精確。就二十五與十中間，擇一相宜之數，分全距離之單位，即可得每一組距單位之約數。然後用其相近之整數，作組距之數值。如第六表，以15分全距離75，則每一組距之數得5；即以5爲組距數値，分75單位爲15組。若將最大之數量95另分一組，則分爲16組，看第一圖及第二圖自明。

(3)定組距之界限 定組距之界限時，須注意下列二要點：(a)清楚：如排列第六表，以5爲組距時，先由最小之數量20起，依次排列爲20，25，30等等。然20與25組距間之界線須詳細書出；如第七表，組距欄內之20-24.99,25-29.99等等，而24.99之意義，即小數之後有無限個9，若再於小數下無限個零之後加1時，即爲25。然亦有僅用中値之數表[10]組距之界限者；如22.5,27.5等等。此法於排列數量之時，中値左右之數量，容易混淆不淸，使排列發生錯誤。非若前法正確明晰，前組最後之數量與後組最前之數量，截然分開，毫無含混之弊。如第七表，凡20-24.99以內之數量，皆列在此組中。若數量比24.99大者，即列在另一組中，餘可依此類推。

(b)精細： 組距之界限，必須細分，而後與各數量之價值比較，始能明顯；如第七表，定 20-25,25-30 等之界限，為便利計，本無不可。若細分為20-24.99,25-29.99等等，則各數量之價值與此組距之界限更易比較，不難定其應歸何組也。若各數量之價值有二位小數之時，可將組距之界限，推至小數下三位，以求類列各數量之便利。

第七表 一百二十三名中學學生英文分數之次數分配；組距為5單位

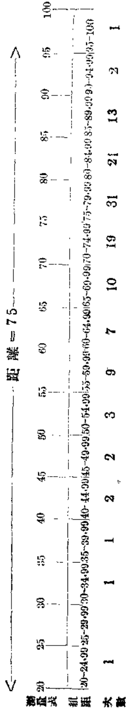
組距	中值 m	排列	次數 f
20.0-24.99	22.5	|	1
25.0-29.99	27.5		
30.0-34.99	32.5	|	1
35.0-39.99	37.5	|	1
40.0-44.99	42.5	||	2
45.0-49.99	47.5	||	2
50.0-54.99	52.5	|||	3
55.0-59.99	57.5	卌 ||||	9
60.0-64.99	62.5	卌 ||	7
65.0-69.99	67.5	卌 卌	10
70.0-74.99	72.5	卌 卌 卌 ||||	19
75.0-79.99	77.5	卌 卌 卌 卌 卌 卌 |	31
80.0-84.99	82.5	卌 卌 卌 卌 |	21
85.0-89.99	87.5	卌 卌 |||	13
90.0-94.99	92.5	||	2
95.0-100	97.5	|	1
總數			123

欲用各組距之中值，與各該組中各數量之次數相乘，以求平均數或差異數等之時，以整數之中值爲最方便。故定組距之界限時，苟計算上需要中值，可使其數爲整數；然後標準中值，列出每組之界限數字。如以20爲中值，組距爲5時，則中值下界限之數字爲17.5，中值上界限之數字爲22.99。其組距之列法即17.5-22.49,22.5-27.49，等等；如第八表：

第八表 係第六表之材料，中值爲整數，組距爲5單位

組距 m	中值	排列	次數 f
17.5-22.49	20	|	1
22.5-27.49	25		
27.5-32.49	30	|	1
32.5-37.49	35	|	1
37.5-42.49	40	||	2
42.5-47.49	45	||	2
47.5-52.49	50	|||	3
52.5-57.49	55	卌 |	6
57.5-62.49	60	卌 |||	8
62.5-67.49	65	卌 卌	10
67.5-72.49	70	卌 ||||	9
72.5-77.49	75	卌 卌 卌 卌 卌 |||	28
77.5-82.49	80	卌 卌 卌 卌 卌 卌 ||||	34
82.5-87.49	85	卌 卌 ||||	14
87.5-92.49	90	||	2
92.5-97.49	95	||	2
總數			123

第五，第七，第八各表，表明數量價値之等級，悉用數字。若用圖形表明之，其意義更可明晰。試就第七表說明之，以一直線代表其全距離，將此直線分成16等分，每一等分代表一組距，組距之價値用數字注明之，每一組距之單位爲5，如第一圖及第二圖。此種直線的表明數量分配之法則，於說明上最爲便利，學者若遇於數字測量表有模糊時，可用此直線法分析之，當無不可解釋者矣。



第一圖 表明測量表，單位，組距之用法

組距	測量表	次數
	20	
20-24.99		1
	25	
25-29.99		
	30	
30-34.99		1
	35	
35-39.99		1
	40	
40-44.99		2
	45	
45-49.99		2
	50	
50-54.99		3
	55	
55-59.99		9
	60	
60-64.99		7
	65	
65-69.99		10
	70	
70-74.99		19
	75	
75-79.99		31
	80	
80-84.99		21
	85	
85-89.99		13
	90	
90-94.99		2
	95	
95-100		1
總數	100	123

第二圖 表明測表，單位，及組距之用法

5 數量之性質

數量可按其性質分為連續數量與不連續數量。不連續數量，乃計數量之單位為自然獨立者，不可再行細分。如統計某學校各班學生之數目，某班學生20人，某班學生21人等等；20及21二數量中間有自然之罅隙，無可再行細分之數量；如某班學生之數目為20.5,21.5,或21.99等等，皆不合於事實。故不連續數量佔測量表中之一點。若將各個不連續數量分組排列，則每一組之數量，即佔測量表中之距離矣。

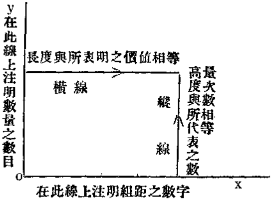
連續數量佔測量表中兩點間之距離，且可細分至無盡數。如測量兒童體高，體重，智力等等，其數量皆爲連續的。如言某兒童體高60寸，此不過言其最相近之高度而已，實則此60寸之高度，不一定恰在60寸之一點，或在59.5與60.5距離之中間；再細分之，或在59.49與60.49距離之中間。若用極精密之尺度測量之，更可分爲再細微之數量。又如第四表之記字試驗，似乎爲不連續數量，實在亦爲連續的。例如某次記字3個，是就其最相近之整數言之，若用較精密之方法測之，或能記3.6個3.9個。第七表之英文分數亦復如此。如言20分，是亦就其最相近之整數分數言之。此等表明數量之數字，在測量表中，以距離言，約有二種解釋：(a)如數字3，係佔2.5與3.5間之距離；或(b)佔3與4間之距離。由(a)之意義，是以3爲最近之數量；其所佔之距離，係由其數與前一數中間距離之半數起，至其數與後一數中間距離之半數止。由(b)之意義，是以3爲最低之限度；凡能及此限度，而不能及其次限度之數量，卽列於此組之中。如第七表中20等數，乃代表20至24.99等中間之距離，卽從(b)之意義也。如第八表中20等數，則代表17.5至24.99等

中間之距離，是從(a)之意義也。

6 圖形表明法

如第四節所述之列表方法，用以整理所蒐集之材料，使之由繁而簡，可以一目了然，誠爲緊要之統計方法。然此外尙可畫成圖形，將材料列於圖內，使其意義之表明，格外淸楚。此類圖形，概有二種：(1)次數多邊形(Frequency Polygon)，如後第四圖及第五圖；(2)柱形圖(Column Diagram)，如第六圖。玆將畫圖方法之大概，說明如下：

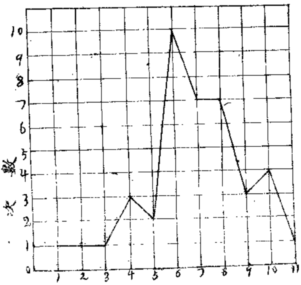
作圖之法，先畫標準線二個。一爲橫線，或底線，通常名 ox 線。此線係代表測量表；組距之數字，即沿此直綫自左而右注明之，如第一圖及第三圖。一爲豎線，或邊線，通常名爲 oy 線。此線係代表數量之次數。表明數量次數之數字，卽沿此線自下而上注明之。與 ox 平行之線，表明數量之價值，名爲橫線；與 oy 線平行之線，表明數量之次數，名爲縱線，如第三圖。



第三圖 說明畫圖之法則

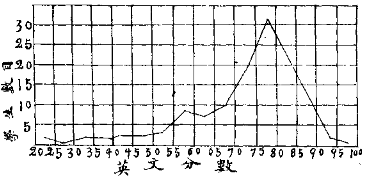
畫ox線時須注意之點: (a)計算次數分配之距離，共有單位若干。(b)排列距離之單位於ox線上時，其距離之遠近以畫紙之大小為準。不可使單位相距過大,致一紙之中不能容一圖之全部;亦不可使之相距太小,使所畫之曲線(Curve)不足以表出特異之點。適宜之度,在畫者自己決定之。(c)在每一組距之中央點上畫一直線。此直線之長即以所代表之數量之次數為準。將各豎線之頂端連以直線,即成一曲線圖形。通常之曲線圖,即就各橫線與各縱線相交之點以直線連接之而成。

畫oy線時應注意之點: 直線上注明數量次數之單位，其大小亦無一定，可隨畫者自己之意思決定之。但:(a)單位之大小,以能使次數分配之全部繪於一圖之中為宜,此項可視紙之大小而定;(b)單位之大小以能明顯表出次數分配之特點為宜，卽各橫線與各縱線相交時所成之次數多邊形不可過低,過低則不足以表明次數分配中之變動。如表明進步速率之圖形，所作之多邊形，若過於扁平，尤不足以顯速率之變化。表明此種材料之圖，最不宜用狹小之單位,



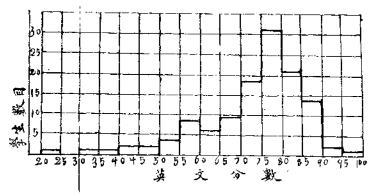
第四圖　次數多邊形(表明第五表之材料)

(1)次數多邊形　此類圖形,係用由底邊向上引出直綫之長度,表明次數之分配。第四圖係表明一個單位之數量;第五圖係表明分組之數量,假定凡每組間所有數量之價值,各與其組間之中值相等。而每組間所有數量之合,即與其組距中點上垂線之高度相等。全體數量即與各組間各中點上垂線之總長相等。



第五圖　次數多邊形(表明第七表之材料)

(2)柱形圖　此類圖形，係用矩形之面積，表明次數之分配。矩形之底邊與組距之長相等。矩形之高與在該組間數量之次數相等。此種表明數量之方法，係假定在每組中所有數量之價値完全相等。如第六圖。



第六圖　柱形圖(表明第七表之材料)

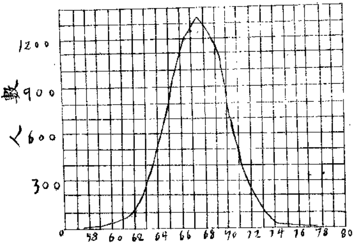
上述之次數多邊形，若用以表明數量少而距離長之次數分配，則甚不精確。此時若用柱形圖之面積表示數量之數目，則較爲顯明。故數量較少之測量時，可以用柱形圖表示所測之結果。

7 次數分配之種類

就第四，第五二圖，可見其所表明之數量分配，甚不整齊。然假設使其組距愈小，所測之數量，亦按配合之數愈增，增至使每組之次數，不至再生變動時爲止；則此等不整齊之分配，愈可趨向於整齊。如第九表之材料，其數量八千有餘，以次數多邊形表明之，如第七圖。其曲線比四，五二圖，整齊多矣。實際上測量之數量，其分配之不整齊者極多；能與理論的分配相合者，幾不可見。然撮其概要，皆可以理論的分配代表之。此類理論的次數分配，最普通者，約有下列二種：(1)對稱的分配(Symmetrical distributon)，(2)偏斜的分配(Asymmetrical distribution)

第九表 英國成年男子8585人體高之次數分配

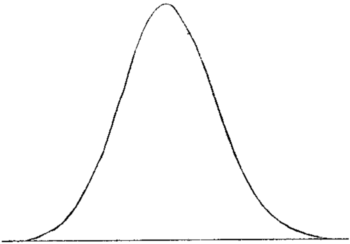
英寸 m	人數 f
57–	2
58–	4
59–	14
60–	41
61–	83
62–	169
63–	394
64–	669
65–	990
66–	1223
67–	1329
68–	1230
69–	1063
70–	646
71–	392
72–	202
73–	79
74–	32
75–	16
76–	5
77–	2
總數	8585



體高按英寸計算

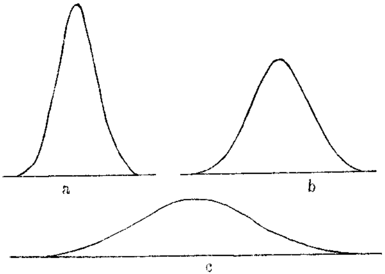
第七圖　英國成年男子8585人體高之次數分配

(1)對稱的分配　對稱的分配，即中央一數量之次數最多，其兩邊之數量，分配相稱，且依次減少；起始減少遲緩，次則較速，再次復較遲；其形如第八圖。



第八圖　理論的對稱次數分配

然此種分配雖甚規則；顧其種類，亦極不一。如第九圖之a,b,c三種圖形，雖皆係對稱；然其高低及距離各有差別。其次數分配之規律固同；但其分配之地位亦不一也。

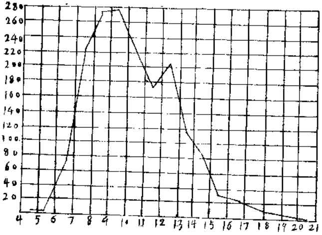


第九圖 理論的對稱次數分配

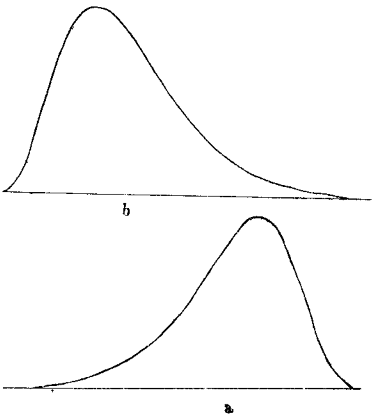
(2)偏斜的分配　偏斜的分配，即在一次數最多之數量之一邊，數量數目驟行減少；而在他一邊之數量數目則減少較遲。然偏斜之分配亦有偏斜輕重之別。偏斜重之分配，於教育測量中尚不多見；最常見者，則爲偏斜輕之分配。如第十表，以年歲較幼之學童居多數，故次數分配偏向測量表中之低端；以圖表明之，如第十圖；其理論的分配如第十一圖a形。

第十表 某縣國民及高小學校1743名學生年齡之次數分配表

年齡	人數
m	f
4	2
5	3
6	74
7	220
8	277
9	278
10	224
11	176
12	204
13	119
14	86
15	31
16	21
17	15
18	6
19	3
20	1
總數	1743



第十圖 某縣國民及高小學校1743名學生年齡之次數分配



第十一圖 理論的偏斜次數分配：a形係偏向低端者，b形係偏向上端者

第十表之分配，係偏向低端者；亦有次數分配，偏向上端者，如第五圖。因英文分數較多之學生居多數，故分配偏向上端；其理論的分配如第十一圖b形。

上述之二種圖形，皆係表明常態的次數分配。若實際上之測量，所得之數量分配，頗多畸特之形。苟能多排列次數分配表，多畫次數曲綫，卽可得此種經驗矣。

習　　題

1 試將下例二問題，以5為組距：A，排列次數分配表；B，作次數多邊圖形：

a 國文分數

50 65 65 45 60 90 78 75 78 90 78 70

70 60 60 64 70 50 80 70 60 48 75 60

60 78 78 56 70 45 75 68 70 45 66 48

65 58 50 36 55 50 58 55 75 56 40 78

40 58 55 40 60 75 50 70 70 50 50 55

55 60

b 數學分數

75 95 75 100 65 75 75 100 75 75 60 65 88

100 75 89 75 75 65 75 85 60 90 70 72 85

65 50 75 50 50 80 70 65 70 63 75 60 55

75 60 99 100 100 88 70 95 95 90 100 100

95 88 90 75 85 90 100 75 92 90 70

第四章

位置數量 (point measures)

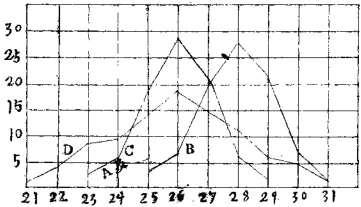
8 次數分配之說明法

次數分配表之作法,已於前章敘其大要;原有材料之情形,頗可於此類表中見之。然若就二種測量之材料,比較其異同,以考究其異同之原因,僅恃次數分配表,則材料散漫,不能作賅括的說明。是以必須有一數量,以代表各個測量之特性,然後比較此代表數量,以求其異同之因果。此類代表數量共有二種:(1)位置數量,(2)差異數量。位置數量,係表明數量在測量表中分配之位置;差異數量,係表明數量相差之距離。用位置數量,可以比較二種測量中各數量所集合之地位;用差異數量,可以比較二測量中各數量分配之情形,以考證位置數量之價值。如第十一表所列之四種測量,若僅就位置數量比較C,D,必認為C與D可相等;而實際上二者之數量分配不同;是非再求各數量分配之距離,不足以有確實之比較。然若僅比較其差異之數量,則A與B二測量亦無差異;實則其位置數量亦不相同。差異數量於第五章敘述之;此章則專論位置數量。

第十一表 A,B,C,D四人每點鐘所得之工資；以分計算

分 m	A 鐘	B 點	C 數	D 目
21				2
22				4
23	3		3	8
24	7		7	9
25	20	3	20	15
26	28	7	28	18
27	22	20	22	15
28	7	28	7	11
29	1	22	1	7
30		7		5
31		1		1

上表中A及B之所得，其數量之位置不等，而其差異之度相同；C及D差異之度不等，而其數量之分配相似。看第十二圖自明。



第十二圖 A,B,C,D四人每點鐘所得之工資；以分計算。

D所佔之距離最大，其數量之位置與A及C相似；B與A及C所佔之距離相等，然其數量之位置不同。

9 位置數量之種類

最常用之位置數量，約有四種：(1)範數(Mode簡寫爲Mo)，(2)平均數(Mean,簡寫爲M)，(3)中點數(Median,簡寫爲Md)，(4)百分點(Percentile)。前三種亦稱集中數量(Measures of central tendency)或簡稱中量。試分述之於下：

(1)範數　範數卽測量表中次數最多之數量，如第七表一百二十三名學生之英文分數，以數量77.5之次數爲最多，共有31人。得其他任一分數之學生數，皆無31人。故77.5爲此次數分配之範數。然此種範數可謂之爲視察範數(Inspection Mode)，與理論範數(theoretical mode)有別，區分之如下：

(a)視察範數，顯而易見，最易求得。如第七表之77.5；及第五圖(表明第七表中之材料)中所畫之曲線，亦以77.5之次數爲最高。故視察範數，按表與圖皆易察出。然必須次數分配對稱(如第八圖)之時，視察範數，方有作代表數量之價値；因此種次數分配，其範數，中點數，平均數，三者皆相重於一點。若次數分配偏斜(如第十一圖)，則此種範數，甚不可靠。且此種範數，常隨組距之大小及數量增加之情形變更，

其價值非若理論範數之精確也。

(b)理論範數,雖爲眞確之範數,然不如視察範數之易求。凡實際上之測量,所測之數量,其數不能甚多;所用之測量方法,不能甚精;故所得之結果,亦不能與理論之現象符合。欲求理論之範數,必須整理不規則之次數曲線,以理論之曲線切合之,然後可求一近似之理論範數。此法之數理高深,非此淺近之書所能備載。然就範數,中點數,及平均數近似之關係,有一求範數之公式:卽,範數=平均數－3(平均數－中點數)。此公式於次數分配無大偏斜之時,可以應用;因此時範數,中點數,及平均數雖不相同;而其彼此之關係,中點數常佔平均數與範數之距離之三分之一。用此公式,卽可求出近似之範數,與眞確範數相差極微。若先求得平均數及中點數,則此範數卽易算出。如第七表,中點數=76.05;平均數=73.11,則範數=73.11－3(73.11－76.05)=73.11+8.82=81.93卽爲近似範數。

(2)平均數　平均數之意義最爲明瞭,爲吾人最常用之算法。茲將其求法,分下列兩種說明之。

(a)**通常算法**　普通求平均數之方法,卽將各個

數量之價値相加，所得之和，以數量之數目除之卽得。試以公式表明之：設M爲平均數，N爲數量之總數，Σm爲各數量價値之和（Σ爲相加之記號），其式爲：

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

若各數量之次數不祇一次，如第十二表；則須將各次數與其相當之數量相乘，乘後再相加，以求各數量價値之總數，其式爲：$M=\frac{\Sigma fm}{N}$ 式中之f=次數，fm=次數與數量相乘之積

第十二表　137名學生答對代數因數問題數目之次數分配

答對題數（數量）	學生數（次數）	次數與數量相乘之積
m	f	fm
13	1	13
12	3	36
11	8	88
10	14	140
9	29	261
8	35	280
7	21	147
6	16	96
5	5	25
4	3	12
3	1	3
2	1	2
	N=137	Σfm=1103

$$M=\frac{1103}{137}=8.05$$ 平均每人約答對八題

第十二表之例，數量相差之單位爲1，算法較爲簡明；再就第十三表分組之材料，說明其算法如下：

1. 求每組組距之中値（假定此中値，足以代表該組中各個數量）

2. 以次數與中値相乘（如中値爲整數，相乘之時甚便利；前4，(3)，(b)節言使中値爲整數者，即以此也）。

第十三表（引用第八表之材料，將中値變爲整數）

組距	中値 m	次數 f	次數與數量之乘積 fm
92.5-97.49	95	2	190
87.5-92.49	90	2	180
82.5-87.49	85	14	1190
77.5-82.49	80	34	2720
72.5-77.49	75	28	2100
67.5-72.49	70	9	630
62.5-67.49	65	10	650
57.5-62.49	60	8	480
52.5-57.49	55	6	330
47.5-52.49	50	3	150
42.5-47.49	45	2	90
47.5-42.49	40	2	80
32.5-37.49	35	1	35
27.5-32.49	30	1	30
22.5-27.49	25		
17.5-22.49	20	1	20
		N=123	8875

$$M=\frac{8875}{123}=72.15$$

(b)簡法　上述之計算方法，層次明顯，理亦易通。特所計算之數過多，不甚省力。明下列之簡捷方法，則前法可以棄置不用。

設有54,52,64,56,50等五個數目；如求其平均數時，先估計此五數之平均數，約有若干。例如估計56爲此五數之平均數，然後以56與此五數一一比較；64與56相差爲8；56與56相差爲零；54,52,50與56相差爲-2,-4,-6。其結果如下：

64	8
56	0
54	-2
52	-4
50	-6
	$\Sigma d=-4$

以d(deviation)代各個差數

Σd＝差數之和

然後將比56多之差數與比56少之差數相減，得-4；再以此五個數目平均之，得-.8；此卽估計平均數與實得平均數相差之數。然後由估計平均數56減去-.8(56-8)，得55.2；卽爲實得平均數。若用通常算法求之，亦得55.2。以公式表明之：　則M=E.M.+C　式中之M=實得平均數(Obtained mean，或簡寫爲O.M.)，E.M.=估計平均數(Estimated mean)，C=更正數(Correcton)

若所估記之平均數與實得平均數巧相符合，則在估計數上各差數之和與估計數下各差之和必相等；正負差數相加，等於零；如前例，假使所估計之數爲55•2，則在估計數上之差數爲8•8，•8，共9•6；在估計數下之差數爲-1•2，-3•2，-5•2，共-9•6；兩種差數相加得9•6+(-9•6)=0此時更正數亦爲零。但通常所估計之平均數，决難與實得平均數絲毫不爽；既有差數，則必須將此差數更正，方能與實得平均數相符。故求C之數目，爲必不可少。既求得C數，再應用代數算法，將C與E. M. 相加；如前例E. M. + C=56+(-•8)=56-•8=55•2。

應用上述簡捷算法之理，求分組材料之平均數，最稱便利。雖作法稍有不同，亦易了解。觀下列各條自明：

1. **先定估計之平均數。** 定估計平均數之標準，可先計算組之數目，由次數分配之一端起，約數至組數二分之一，選擇一組，假定其含有估計平均數之數量；以該組之中値，作估計平均數之價値。或就次數分配中多數數量聚集之處，擇定一組，以其中値作估計平均數。學者若能多演習題，則於平均數之估計，雖不必中，當不遠矣。如第十四表，由小數一端起，數至95--99.99組，已達組數二

分之一；以估定此組含有平均數之數量。其中值爲 97.5，即以此數代表估計平均數。而含有次數較多各組，如100–104.99,90–94.99,85–89.99 與此估計平均數相距較近；以 f 與 d 相乘之時，數目較小，計算上甚爲便利。

第十四表 某學校教員每月之薪俸數目

m	f	d	fd
60 - 64.99	2	-7	-14
65 - 69.99	4	-6	-24
70 - 74.99	5	-5	-25
75 - 79.99	15	-4	-60
80 - 84.99	15	-3	-45
85 - 89.99	22	-2	-44
90 - 94.99	16	-1	-16
95 - 99.99	8	0	-228
100 - 104.99	26	1	26
105 - 109.99	7	2	14
110 - 114.99	10	3	30
115 - 119.99	2	4	8
120 - 124.99	4	5	20
125 - 129.99	4	6	24
130 - 134.99	3	7	21
135 - 140.00	1	8	8
	N=144		151

$$C = \frac{151-228}{144} \times 5$$

$$M = 97.5 + \frac{151-228}{144} \times 5 = 97.5 + \frac{-77}{144} \times 5$$

$$= 97.5 - 2.67 = 94.83$$

2. 估計數量既定之後，再求各數量與估計數量相差之數。假定各組距之單位相差均為 1，無論各組距之價值，實在相差若干，皆按其差數為 1 計算；例如第十四表，以估計平均數 97.5 與所測之數量 97.5 比較，其差為零；與數量 92.5 相差為 -1；與 87.5 相差為-2；與 82.5 相差為-3；與數量 102.5 相差為 1；與數量 107.5 相差為2；等等，如第十四表d欄。既估定平均數，則按0,-1,-2…；1,2,3…等順次直書，即得各 d 數。

3. 將各 d 數用相當之次數乘之，得 fd；如第十四表之fd欄。相乘之時，須注意負號。

4. 將各正負fd之數相加，求代數式之和 Σfd；如第十四表，負號差數為-228,正號差數為151；151+(-228) = -77

5. 用次數除 Σfd，並以組距單位之數乘之，求更正數C；如第十四表，N=144，組距之數為 5，其式為 $C=\frac{-77}{144}\times 5=\frac{-385}{144}=-2.67$，即估計平均數與實得平均數相差之數。

6. 將C(=-2.67)與估計平均數相加，得97.5+(-2.67)=97.5-2.67=94.83，即實得平均數。

(3)中點數　中點數之意義，即依數量價值之次第，在此數之每邊各有數量總數之二分之一。如第十六表之材料，共有數量 123，其中點數為 76.05；在此數上下之數量，各為 61.5，即有$\frac{1}{2}$數量等於或大於此中點數；有$\frac{1}{2}$數量等於或小於此中點數。求連續數量之中點數，與求不連續數量之中點數，方法相同。

惟其中亦有應行區別之點，特分述之如下：

(a)求不連續數量之中點數 數量之數目(N)，可為偶數，亦可為奇數。按前中點數之定義，求偶數數量之中點數時，則以第$\frac{N}{2}$與第($\frac{N}{2}+1$)二數量之和之平均數為中點數；遇有不合理之數值時，則取其數之最近整數。求奇數數量之中點數時，則以含有第$\frac{N}{2}$數量之數量價值，為中點數。如第十五表；有家庭十六，即N=16，$\frac{N}{2}=8$；由小數一端向下數，至第8個家庭有兒童7人，至第九個($\frac{N}{2}+1$)家庭，有兒童10人，由是中點數＝$\frac{7+10}{2}=8.5$；然因.5個兒童於理不合，故取其最近之整數9為中點數。若將此表有12個兒童之家庭除去時，則N為奇數(=15)。求其中點數時，以2除15＝7.5；由表之任一端數起，至含有第7.5數量之第8個家庭，其兒童數為7，是以中點數＝7

第十五表

十六個家庭之學齡兒童數目

兒童 m	家庭 f
3	1
4	1
5	1
5	1
6	1
7	1
7	1
7	1
10	1
10	1
10	1
11	1
11	1
12	1
13	1
13	1
N =	16

$\frac{N}{2}=8$, $\frac{N}{2}+1=9$

第8個數量為7，第9個數量為10，是以中點數＝$\frac{7+10}{2}=8.5$，然因此數不合理，故取其最近之整數9為中點整數。

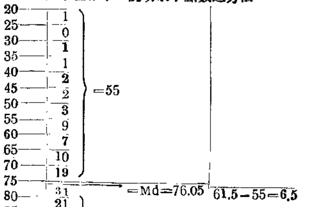
(b) 求連續數量之中點數　玆將其求法列下，並將前第七表之分組材料列於第十六表以說明之。

第十六表　一百二十三名學生之英文分數

m		f	
20.0 - 24.99		1	
25.0 - 29.99			
30.0 - 34.99		1	
35.0 - 39.99		1	
40.0 - 44.99		2	
45.0 - 49.99		2	
50.0 - 54.99		3	
55.0 - 59.99		9	
60.0 - 64.99		7	
65.0 - 69.99		10	
70.0 - 74.99		19	F = 55
75.0 - 79.99	此組含有中點數	31	f = 31
80.0 - 84.99		21	
85.0 - 89.99		13	
90.0 - 94.99		2	
95.0 - 100		1	
		N = 123	

$$\frac{N}{2} = 61.5$$

第十三圖　說明求中點數之方法



$$Md = 75 + \frac{61.5-55}{31} \times 5$$

$$= 75 + \frac{6.5}{31} \times 5 = 76.05$$

算法： 1. 先將各數量依次排列，作次數分配表。

2. 將各次數相加求出數量之總數，及總數量之二分之一；以N代數量之總數，則第十六表，$N=123$, $\frac{N}{2}=61.5$

3. 定含有中點數之組，以f代該組數量之次數；如第十六表，含有中點數之組爲75—79.99，其次數$f=31$。

4. 從次數分配任一端起，計算含有中點數組以外各組數量次數之和，以F代之；如第十六表，由小數量之一端起，計算數量之次數，至含有中點數組爲止，則$F=55$。

5. 由$\frac{1}{2}$數量減去F,即$\frac{N}{2}-F$；如第十六表$\frac{N}{2}-F=61.5-55=6.5$。所餘之數量6.5，即爲在該組中達到中點數之數量數目。參看第十三圖。

6. 用含有中點數組之次數分所餘之次數，其公式爲$\frac{\frac{N}{2}-F}{f}$。如第十六表，$\frac{\frac{N}{2}-F}{f}=\frac{6.5}{31}$，此商數即爲計算中點數之比例率。

7. 將所得之比例率以組距數乘之；以i代組距數；如第十六表，$i=5$, $\frac{6.5}{31}\times 5=1.05$

8. 將1.05與含中點數組低界限之價值相加，如第十六表，$75+1.05=76.05$，即所求之中點數。表明組距高低限界之數字，須用整數，方利於計算。

再以代數式表明之，其式如下： 設V＝含有中點數組低界限之價值，則

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

由數量低端算起，可適用上式；若由數量上端算起，則前式中之 V= 含有中點數組高界限之價值，其式如下：$Md=V-\frac{\frac{N}{2}-F}{f}i$ 依第十六表，則

$$Md=80-\frac{61.5-37}{31}\times5=80-3.95=76.05$$

故無論由何端算起皆可，而所得之中點數則一。

(4)百分點(Percentile) 在中點數之上下，各有總數量之$\frac{50}{100}$，故亦稱中點數爲50百分點(50 percentile)。此外尙有$\frac{25}{100}$點，可簡稱爲25百分點，或下百分點，以Q_1代之；及$\frac{75}{100}$點，可簡稱爲75百分點，或上百分點，以Q_3代之。在下百分點以上之數量，有以下數量之三倍。在上百分點以上之數量，有以下之數量之三分之一。求下百分點與上百分點之法，與求中點數相似，試以公式表明之：

設 p = 百分點　　P_p = 百分點之價值

V_p = 含有百分點組低界限之價值

F_p = 含百分點組以外次數之和　N = 次數

f_p = 含百分點組中之數量次數　i_p = 組距數

則$P_p=V_p+\frac{\frac{p}{100}N-F_p}{f_p}i_p$　如第十六表，求25百分點時，其式爲：

$$P_p=65+\frac{\frac{25}{100}\times123-26}{10}\times5=65+\frac{4.75}{10}\times5=65+2.375=67.375$$

求75百分點時，則 $P_p=80+\frac{\frac{75}{100}\times123-86}{21}\times5=80+\frac{6.25}{21}\times5=81.49$

百分點之用法，與中點數相倣。求得下百分點與上百分點，卽可知百分之二十五，百分之五十，與百分之七十五之數量，所佔之地位矣，（參看第五章，11，(2)節)。

10 各集中數量之性質及功用

(1)範數

(a)範數非由所測之全體數量求出，乃由其中集合於一組之小部分數量求出。

(b)範數顯而易見，因其由次數最大之組規定而成。

(c)範數不受極端數量（最大數量及最小數量）之影響。理論範數根據高深切合曲線(Curve-fitting)之理論，最難求出，故不常用。

(d)範數之變動最大，故精確之度不及平均數和中點數。其變動之情形隨分組之方法而異；如第六表之

材料,以3為組距時,則範數=80.5;以5為組距時,則=77.5;以10為組距時,則=75,(讀者可就該表一一求之)。又因所取之實例(Samples)不同,其價值亦隨之而異。是以作精密之統計時,範數不適於用。

(c)範數最易使人了解其意義。所代表者,為最普通之情形。如言美國普通公民,即指所最常見之公民而言;又如設某地方有十人擁有資財百萬,其他之公民皆窮匱不支,若用平均法計算該地方之經濟狀況,每人皆屬小康,實則一班公民真經濟窮困。若用範數代表之,即可見一班公民經濟窮窘之狀況矣。由此可見範數之特點,能代表普通之現象,為他種集中數量所不可及。

(2)中點數

(a)中點數亦非由全體數量之價值求出,各數量之價值僅有間接之關係;即求中點數之始,先就各數量價值之大小列成次第,然後依此次第之價值推算中點數。

(b)中點數較平均數容易計算,然精確之度不及平均數。

(c)中點數之價值受極端數量之影響很小,故每不

足以作代表數量。然若極端之數量不近情理之時，欲減少其誤謬之效力，以中點數作代表數量最爲相宜。

(d)分組之時，組距之位置及大小，影響於中點數很小。如第六表之材料以3爲組距時，中點數=75·64；以5爲組距時，中點數=76·05；以10爲組距時，中點數=75.1，

(e)數量愈多，次數分配愈近於對稱之時，中點之價值愈可靠，與平均數之價值愈相近。

(3)平均數

(a)平均數由所測之全體數量求出，故比範數及中點數精確。

(b)用簡捷方法，平均數亦極易求。

(c)平均數易受極端數量或不合理數量之影響；就此點言，其功用不及範數及中點數。

(d)平均數容易使人誤會其所代表各數量之價值，皆與平均數相近似。

問題

2 求問題1之 Q_1, Md, Q_3, Mo

3 求下列二表之 Q_1, Md, Q_3, Mo

a

分數	人數
50-51.9	4
52-53.9	3
54-55.9	5
56-57.9	5
58-59.9	4
60-61.9	7
62-63.9	4
64-65.9	7
66-67.9	1
68-69.9	2
70-71.9	1
72-73.9	3
74-75.9	1
76-77.9	
78-79.9	1
總數	48

b

分數	人數	分數	人數
32-33.9	1	60-61.9	3
34-35.9		62-63.9	3
36-37.9		64-65.9	1
38-39.9		66-67.9	7
40-41.9		68-69.9	3
42-43.9		70-71.9	4
44-45.9	1	72-73.9	2
46-47.9		74-75.9	3
48-49.9	2	76-77.9	3
50-51.9	1	78-79.9	1
52-53.9	3	80-81.9	
54-55.9	3	82-83.9	1
56-57.9	1	84-85.9	
58-59.9	4	86-87.9	1
		總數	48

第五章

差異數量（variability measures）

差異數量可分爲絕對的差異數量與比較的差異數量二種。

11 絕對的差異數量（measures of absolute variability）

絕對的差異數量可分四種：（1）距離，（2）四分差數，（3）平均差數，（4）標準差數。

（1）距離（Range） 距離爲次數分配中最大數量與最小數量之差，包含所測之數量全體，乃表明差異最簡單之方法，亦爲表明差異最不精確之方法。每一測量中之最大量與最小量，並無眞正之界限，特其數量發現之次數較少而已。大數量與小數量發現之次數，恒有多寡之變動。故距離之範圍亦隨之生大小之變動。如第七表，距離之數爲95－20＝75，使由小數量之一端去一數量，則其距離爲95－30＝65；而減此一數量，卽將原距離約減百分之13。若原距離爲65，增此一數量，卽將原距離約增百分之15。故距離之變動最大。且同一距離中間兩種測量之數量分配，必不能盡同；其一測量之各數量或與中量相距較近；而他一

測量之各數量或散漫於距離之中，集中之程度甚小。是以測量中數量分配之形式，用距離之法則，不能表而出之。故距離作代表差異之數量，最不適宜。普通多以下列三種爲差異數量之代表。

(2)四分差數 (Quartile deviation,簡寫爲Q) 四分差數爲表明次數分配形式之簡便方法，其算法之公式如下：

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

式中之Q代四分差數，Q_1 代下百分點，Q_3 代上百分點。算 Q_1 及 Q_3 之法，參看前第四章，9，(4) 節求百分點之方法。求出 Q_1 及 Q_3 即可求Q。如前第十六表，$Q_1=67.375$，$Q_3=81.49$，則$Q=\frac{81.49-67.375}{2}=\frac{14.115}{2}=7.06$

Q_1 之下有全體數量 25%；Q_3 之上有全體數量 25%；Q_3-Q_1 包含全體數量 50%。Q 爲全體數量中央之50% 所佔距離之半數；若將此距離列於平均數之左右，則此距離可含全體數量$\frac{50}{100}$。Q_1，Q_3 與 Md三點分全體數量爲四等分。設使次數分配爲對稱，則$Md-Q_1=Q_3-Md$，卽可用此差數作四分差數。但實際之測量，極難得有對稱的次數分配，則 $Md-Q_1$ 不一定等於 Q_3-Md。故必須取 Q_3-Q_1 之折中數，表明四分差異

之程度。

當次數分配對稱之時，Q與概誤差（Probable error，簡寫爲P.E.）完全相等。惟概誤差須專用以表明常則分配（Normal distributim）之差度（見後第九章，30節）。若教育方面實際之測量，多得偏斜的分配，故以用Q表明四分差度爲宜。

四分差數，簡便明晰，容易計算。凡測量之無須用精確數目代表時，此法頗適於用。

(3)平均差數（Mean deviation，簡寫爲M.D.）　平均差數爲各數量與中量之差數之平均數。各差數相加之時，不按正負記號，若按正負號總計各差數時，則由平均數所求出之各差數，其正負之和爲零，不得求差數之平均數；而由中點數所求出之各差數，其正負之和，亦必很小，不能得正確之差數平均數。故差數相加時，必須不論正負之記號。由平均數或中點數，固皆可求各數量之差數；但中點數較平均數容易計算，且就術理上之關係，由中點數求出之平均差數，其量最小。故普通皆用中點數求平均之差數。

平均差數之求法甚簡單，分述於下：

算法　(a)通常算法　其作法之次第如下：

1. 求出中點數。

2. 求出每一數量與中點數之差數。

3. 求各差數之和。

4. 求各差數之平均數：用次數除各差數之和，即得平均差數。

(以上四步算法甚簡單，故勿庸舉例說明)

(b) 簡捷算法　若用通常法求分組數量之平均差數，其計算步驟，與前(a)節之4層無異。惟計數甚繁，頗費時間，故從略。下所列者乃簡捷方法，計算簡便，理亦易明。就第十七表說明之如下：

第十七表　289名學生之拉丁文分數

		f	d	fd	
95.0–100		22	3	66	= 141 個數量
90.0– 94.99		68	2	136	
85.0– 89.99	O.Md = 84.375	51	1	51	
80.0– 84.99	A.Md = 82.5	28	0		= 120 + 28 = 148 個數量
75.0– 79.99		47	1	47	= 120 個數量
70.0– 74.99		33	2	66	
65.0– 69.99		21	3	63	
60.0– 64.99		9	4	36	
55.0– 59.99		6	5	30	
50.0– 54.99		2	6	12	
45.0– 49.99		1	7	7	
40.0– 44.99		1	8	8	
		N = 289		$\sum fd = 522$	

實得中點數 = 84.375

假定中點數 = 82.5

更正數 = 1.875/5 = .38

$$M.D. = \frac{522 + (.38 \times 148) - (.38 \times 141)}{289} 5$$

$$= \frac{522 + 7 \times .38}{289} 5 = \frac{524.66}{289} 5$$

$$= 1.816 \times 5 = 9.08$$

1. 將各數量分組列成次數分配表，求次數之和；如表N=289

2. 按前章求中點數之方法，求中點數；如表Md=84.375

3. 用含有中點數之組（80.0−84.99）之中值，爲假定之中點數（Assumed median 簡寫爲A.Md）；如表A.Md=82.5

4.假定各組數量相差爲1，求各數量與假定中點數之差數d；再以次數乘之，求fd；然後求差數之總數；得 Σfd=522

5. 求更正數(c)；即求實得中點數與假定中點數相差之數。此例實得中點數爲84.375，與假定中點數82.5相差爲84.375−82.5=1.875；此差數係按組距=5計算；若設組距=1，即設5=1，則$1.875=\frac{1.875}{5}=.38$ .38爲每一數量之更正數，總計更正數爲+(.38×148)−(.38×141)

6. 求4節與5節之和，即522+(.38×148)−(.38×141)=522+(148−141).38=522+2.66=524.66

（此層計算之理甚簡明，522爲由假定中點數所求之差數之和，與由實得中點數所求出之差數相差爲+(.38×148)−(.38×141).38爲更正數，148爲實得中點數以下之次數，141爲實得中點數以上之次數。假定中點數既與實得中點數相差爲.38，則在實得中點數以下共有148個數量（120+28）比實得中點數小.38，故必須+(.38×148)。又在實得中點數以上之數量，共有141比實得中點數大.38，故必須−(.38×141)。若先將148減去141，然後再與.38相乘，即522+(148−141).38=522+7×.38亦可）。

7. 再將524.66用N除之，即得平均差數：$\frac{524.66}{289}=1.816$.
此數係假定組距之數爲1求出之者；故必：—

8. 以5乘之；$1.816\times5=9.08$

此簡捷法甚緊要，再舉第十八表之例以說明之。試以公式代各數値；設M. D.爲平均差數，O.Md.爲實得中點數，A. Md.爲假定中點數，N_a爲實得中點數以上之數量，N_b爲實得中點數以下之數量，i爲

第十八表　一百二十三名學生之英文分數

	f	d	fd	
20.0 – 24.99	1	−11	−11	
25.0 – 29.99	0	−10		
30.0 – 34.99	1	−9	−9	
35.0 – 39.99	1	−8	−8	
40.0 – 44.99	2	−7	−14	
45.0 – 49.99	2	−6	−12	= 55 個數量
50.0 – 54.99	3	−5	−15	
55.0 – 59.99	9	−4	−36	
60.0 – 64.99	7	−3	−21	
65.0 – 69.99	10	−2	−20	
70.0 – 74.99	19	−1	−19	O. Md = 76.05
75.0 – 79.99	31	0		A. Md = 77.5
80.0 – 84.99	21	1	21	c = −1.45 5 = −.29
85.0 – 89.99	13	2	26	=68個數量
90.0 – 94.99	2	3	6	
95.0 – 100	1	4	4	
	N=123		$\sum d=222$	

組距之單位數，C爲更正數$=\dfrac{O.Md-A.Md}{i}$；則M. D.

$=\dfrac{\sum fd+c(Nb-Na)}{N}$　　按第十八表，$C=\dfrac{76.05-77.5}{5}=\dfrac{-1.45}{5}$

$=-.29$，$M.D.=\dfrac{222+\{-.29(55-68)\}}{123}=\dfrac{222+3.77}{123}=\dfrac{225.77}{123}$

$=1.836$。再以組距之單位數5乘之，得$1.836\times5=9.180$

(4)標準差數 (Standard deviatian,簡寫爲S.D.)

標準差數爲各差數(各數量與中量相差之數)之平方平均數之方根。標準差數與平均差數之區別即在於此。算平均差數時,若將正負之差數相消,於理不合;是以規定不論符號之方法。然符號仍然存在,不過棄之不用耳。要想將符號取消,使宜於代數式之計算,最簡單之方法,卽是將各差數平方,以去其負數記號。平方之後相加,求各平方差數之總數。然後將總數用次數平均,得平方差數之平均數。再將此平均數開方,即得標準差數。

按術學之理論,求精密之差數時,須使差數之平方和爲最小。求最小之差數平方和,須由平均數求各數量之差數。故普通求各差數不用中點數,皆用平均數。

標準差數之符號爲 S.D.,普通常亦以 σ 表明之。

求標準差數之公式爲:

$$\sigma=\sqrt{\frac{\Sigma d^2}{N}} \qquad (1)$$

式中之d爲任一數量與平均數之差,d^2爲差數之平方,N爲數量之總數。

若差數須與次數相乘之時,其公式爲:

$$\sigma=\sqrt{\frac{\sum fd^2}{N}} \qquad (2)$$

公式中之 f 爲數量之次數。

算法 (a)通常法 按(1),(2)兩公式其算法如下:—

1. 求次數分配之平均數。(參考第四章 9,(2) 節)。

2. 求各數量與平均數之差數 d。(分組之數量,以每組之中値爲代表數,與平均數相較,求 d)

3. 將各差數平方,求 d^2

4. 將各平方差數與相當之次數相乘,求 fd^2。(若各個數量單列,則用公式(1),即不用此層)

5. 將各平方差數相加,求出總數 $\sum d^2$,或 $\sum fd^2$。

6. 用次數 N除平方差數之和,即 $\frac{\sum d^2}{N}$ 或 $\frac{\sum fd^2}{N}$,求出平方差數之平均數。

7. 將平方差數之平均數開方,即 $\sqrt{\frac{\sum d^2}{N}}$ 或 $\sqrt{\frac{\sum fd^2}{N}}$;求出標準差數。

上述之通常算法,計數太繁,故無庸舉例說明,另有簡捷方法如下:—

(b)簡捷法 就第十九表說明之。

第十九表　前第九表之材料

m	f	d	fd	fd²
57 -	2	-10	20	200
58 -	4	-9	36	324
59 -	14	-8	112	896
60 -	41	-7	287	2,009
61 -	83	-6	498	2,988
62 -	169	-5	845	4,225
63 -	394	-4	1576	6,304
64 -	669	-3	2007	6,021
65 -	990	-2	1980	3,960
66 -	1223	-1	1223	1,223
67 -	1329	0	-8584	——
68 -	1230	1	1230	1,230
69 -	1063	2	2126	4,252
70 -	646	3	1938	5,814
71 -	392	4	1568	6,272
72 -	202	5	1010	5,050
73 -	79	6	474	2,844
74 -	32	7	224	1,568
75 -	16	8	128	1,024
76 -	5	9	45	405
77 -	2	10	20	200
	N=8585	—	8763	Σfd²=56,809

$$C=\frac{8763-8584}{8585}=\frac{179}{8585}=.0209, C^2=.0004$$

$$\sigma=\sqrt{\frac{56809}{8585}-.0004}=\sqrt{6.6168}=2.57$$

A.D.=2.041

Q　=1.75

1. 排列次數分配表。

2. 擇假定之平均數。於數量集合最多之處，擇定一數，作假定平均數；如表，得67

3. 求各數量與假定平均數之差數；如表第三欄d之下,1,2,3……等；-1,-2,-3……等。

4. 用次數與各差數相乘，求 fd；如表第4欄之 $2\times-10=-20$，$4\times-9=-36$ 等等。

5. 將 fd 用代數法相加求其和數；$\Sigma fd=8763-8584=179$

6. 用N除 Σfd,求更正數 C；$\frac{179}{8585}=.0209$

(以上六層與求平均數之簡法相同，可參考第四章 9,(2),(b)節。

7. 用 d 乘每一fd，求差數之平方 fd^2；如表第五欄之 $-10\times-20=200$；$-9\times-36=324$ 等等。

8. 求 fd^2 之和數；$\Sigma fd^2=56,809$

9. 用N平均 Σfd^2；$\frac{\Sigma fd^2}{N}=\frac{56809}{8585}=6.6172$，以 S^2 代之。

S^2 爲由假定平均數求出之標準差數之平方。設使假定平均數與實得平均數相等，即無庸更正數。然此例之更正數爲 .0209，故此假定之標準差數之平方與實得標準差數之平方相差爲 $(.0209)^2=.0004$。

10. 由假定標準差數之平方減去更正數 C^2，求 σ^2；$\sigma^2=S^2-C^2=6.6172-.0004=6.6168$

11. 將 σ^2 開方，求 σ；$\sigma=\sqrt{S^2-C^2}=\sqrt{6.6168}=2.57$ 英寸。此數係假定組距之單位爲1計算，故必須與組距之實在單位相乘還原。故：—

12. 將所求之標準差數用原組距數乘之，得實在之標準差數。

此例組距之數爲 1,故無庸再乘。然如下例第二十表，則必須用原組距數 5 乘所求之標準差數，始得實得的標準差數。再將此簡捷算法以公式表明之如下：— $\sigma=\sqrt{\frac{\Sigma fd^2}{N}-C^2}\times i$

第二十表　第十四表之材料

m	f	d	fd	fd²
60- 64.99	2	-7	-14	98
65- 69.99	4	-6	-24	144
70- 74.99	5	-5	-25	125
75- 79.99	15	-4	-60	240
80- 84.99	15	-3	-45	135
85- 89.99	22	-2	-44	88
90- 94.99	16	-1	-16	16
95- 99.99	8	0	-228	
100-104.99	26	1	26	26
105-109.99	7	2	14	28
110-114.99	10	3	30	90
115-119.99	2	4	8	32
120-124.99	4	5	20	100
125-129.99	4	6	24	144
130-134.99	3	7	21	147
135-140.00	1	8	8	64
	N = 144		151	∑fd-1477

如第二十表 $\sigma=\sqrt{\frac{1477}{144}-(.534)^2}\times 5=\sqrt{10.257-.285}\times 5=15.8$

標準差數之功用　心理試驗學者及統計學者多以標準差數爲測量差異之數量；蓋因其精確之程度比平均差數較大，而其計算亦不甚難之故也。就經驗所得之結果，凡對稱或偏斜輕之次數分配，其標準差數之六倍，大約包括全體數量百分之99。此種經驗規則，可以証明計算是否錯誤；並使標準差數之意義，有一確切與具體之解釋。如第十九表，6倍標準差數

2.57為15.42;而由60至75.4距離之間,8585數量中除去37個數量(2+4+14+2+5+10)外,其餘皆包含在此15.42距離之內;即8585−37=8548,約為總數99.6%。然數量過少之測量,則不能與此定理相符。

標準差數與平均差數之關係　凡對稱或偏斜輕之次數分配,平均差數約當標準差數五分之四。如第十九表 $\frac{A.D.}{S.D.}=\frac{2.041}{2.57}=.80$。前節言全體數量99%約占標準差數6倍之距離;由此關係推算,則全體數量99%占平均差數之距離,約$7\frac{1}{2}$倍。然數量過少之時,則與此定理不符。若次數分配對稱之時,二者理論之關係如下:—

$$\sigma=1.2533\ \mathrm{M.D.}\quad 或\ \mathrm{M.D.}=0.7979\sigma$$

標準差數與四分差數之關係　若對稱或偏斜輕之次數分配,則Q約當σ三分之二。如第十九表,$\frac{Q}{\sigma}=\frac{1.75}{2.57}=.68$。然數量過少之時,亦不能與此相符。按全體數量百分之99約占6倍標準差數之距離,若就上式之關係推算,約占9倍Q之距離。

12　求比較的差異數量(measures of relative variability)

欲比較二種測量差異之程度,不能就各測量之絕

對的差異數量直接比較;蓋因各測量所用之單位,與所得之集中數量,其數之大小不能盡等;因之由中量所得之絕對差異數量,亦參差不齊。若僅就所得之絕對差數。比較各測量差異之大小,必不正確。是以必須用比較的差異數量作比較。蓋各測量之差數,對於其中量,有其相當之比率(如標準差數與平均數;平均差數,四分差數與中點數,各有相當之比率數)。若就此比率數考其差異之大小,自當較為正確。如以平均數除標準差數,用皐爾生(Pearson)之公式表明之,則$V=100\frac{\sigma}{M}$。此比率數名為差異係數(Coefficient of variation);即標準差數與平均數之百分比率數。如第二十一表中有四種測量,每種之平均數與標準差數皆已求出;其單位及平均數各異,而其標準差數亦不等。若比較各測量差異之大小,用以上之公式,則A之$V=100\frac{1.4}{3.9}$;$B,=100\frac{19.3}{77.8}$;$C,=100\frac{2.9}{12.4}$;$D,=100\frac{11.3}{76.9}$;表中最低之行,即所求之各比率數。原單位之種類及大小,與此比率數毫無影響。故此數僅為表明差異程度之數目而已。

第二十一表　四種測量

	A 每秒鐘所讀之字數	B 圖畫分數按百分算	C 每分鐘所答之算題數	D 算學分數
平均數M.	3.9	77.8	12.4	76.9
標準差數 σ	1.4	19.3	2.9	11.3
比率數V.	36	25	23	15

由上表中比率數，可以知差異最大者爲 A；最小者爲 D；B，D 兩種其單位相同，其平均數相差甚小；故亦可直接比較其差異程度爲 B 較 D 大也。

13　求次數分配偏斜之方法(Measures of shewness of distributions)

次數分配對稱時，範數，中點數，平均數三者合而爲一，前章已言之矣。若次數分配偏斜之時，三中量亦恒有相當之關係；卽中點數之位置，占平均數與範數之距離之三分之一(見第四章，9，(1)，(b)節)。故欲推求次數分配偏斜之程度，須就此中量之關係，用差異數量求出比率之數。此數與前節差異係數之性質相同，僅爲測量偏斜之數量，與原單位之大小及種類毫無關係。其式如下：

(1) 偏斜之度 $= \frac{\text{Mean} - \text{Mode}}{\sigma}$，因眞確範數不易計

算，故將上式改爲：偏斜之度 $= \frac{3(\text{Mean}-\text{Median})}{\sigma}$，因 Mode = Mean − 3 (Mean − Median)

(2)偏斜之度 $= \frac{(Q_3-\text{Md})-(\text{Md}-Q_1)}{Q} = \frac{Q_1+Q_3-2\text{Md}}{Q}$

(1)式比(2)式較精密，故引用最多。

上式之理，亦甚簡明。如 Mean − Median，$(Q_3-\text{Md})-(\text{Md}-Q_1)$爲零時，則次數之分配對稱，不偏不依；若 Median，$(\text{Md}-Q_1)$之數小於 Mean與$(Q_3-\text{Md})$之時，則低端之次數多，次數分配向低端偏斜，而所得之比率數爲正。若 Mean 與 $(\text{Md}-Q_1)$ 之數大時，則上端之次數多，次數分配向上端偏斜，所求之比率數爲負。故按比率數之正負，即可知次數分配之偏斜向下或向上之程度也。

(1)式之最小值爲零，卽當次數分配爲對稱之時也；其最大值爲2；但實際上測量時，所得之數，能過±1者極少。

(2)式之最小值亦爲零，卽次數分配之形爲對稱；其最大值在此公式中不易見出，但若當次數分配之形較偏斜時，如第十一圖，可以確定其比率數，不能過於1。

問 題

4 求 2，3 兩題之 Q 及 S.D.

5 求下列四種測量之差異係數：

	A	B	C	D
M.	57	70	60.29	65.58
S.D.	11.14	10.72	6.8	10.36

第六章
關係數量

(Relationship messures: Correlation)

14 關係數量之意義

關係數量，係表明二種或數種測量之量彼此相關之程度也。譬如某人之記憶力強，則此人之推理力與其強健之記憶力有何關係？其人記憶力之量大，是否其推理力亦大；其人之記憶力弱，是否其推理力亦弱？設使記憶力強之人，其推理力亦強；記憶力弱，其推理之力亦弱，且二者強弱之分量相等；然其強弱相等之分量如何？若其記憶力強推理力弱；推理力弱，而記憶力強之時；其強弱不同之分量又如何？又如英文程度優良之學生，其國文之程度，是否優良？其優劣之分量若何？教育方面與心理方面類此之測量甚多，欲知其相關之分量，或大或小，為正為負，必求一簡單之數量表明之，亦猶用集中數量及差異數量，表明測量之集中及差異之程度也。此簡單之數量稱為關係係數(Coefficient of Correlation)，通常以r代之。

15 關係之種類

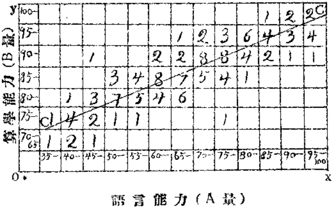
關係之種類有三：(1)正關係或積極關係。二種測量顯正關係時，則二量之性質相近；即其一量增

加時,他一量亦必隨之增加。

(2)負關係或消極關係。 二種測量表現負關係時,則二量之性質相反;即其一量增加時,他一量隨之遞減。

(3)無關係。 二種測量表現無關係時,則其一量之增減,與他一量無關。

以上各種關係,可用曲線表明之。如第十四圖,設oy為豎線,代表B量;設ox為橫線,代表A量,將A,B

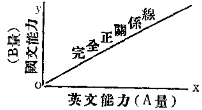


第十四圖 一百三十名學生語言及算學相關能力之分配

之各數量分為若干組,依法(詳後16,(1),(c)節)排列。然後將每一A組相當之各B數量之價值相加,求其平均數;如A量35-40組中各B數量為67.5+72.5=140,以n(=2)除之,得平均數70;如A量40-45組中各B

數量爲 (67.5×2) + (72.5×4) + 77.5=502.5，以 n (=2+4+1=7)除之，得平均數71.8；又如A量95－100組中各B數量之價值，爲87.5+(92.5×4)+(97.5×2)，以7除之=93.2。次將各平均數之點以CC′線連接之，則此直線卽爲表明 A,B二量之關係線 (Correlation line)。惟此圖所作之線，係一直線；若就各平均點連以直線，則其線必上下變動，屈曲不直；然假使所測量之數量愈多，測量之方法愈精密，則此曲線可與直線愈相近。由是二量之關係線，卽可以直線表明之。此章所述之關係，均假定其爲直線的；曲線的關係及算法從畧。

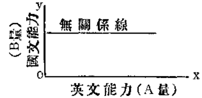
以直線表明正關係時，如第十五圖。設英文能力以 ox 線表明之；國文能力以 oy 線表明之，此圖表明



第十五圖

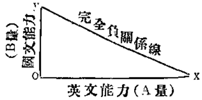
A B 二量顯完全正關係；卽 A 之大量數量與 B 之大量數量相連，A之小量數量與 B 之小量數量相連。如

甲學生之英文程度第一，其國文之程度亦第一；乙生之英文程度第二，其國文之程度亦第二。若 A,B 表明無關係時，如第十六圖；即A之增減不影響於B。如英



第十六圖

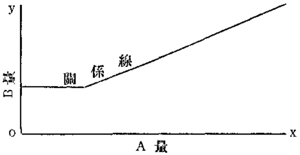
文能力增加，而國文能力不生變動。若 A,B 表明負關係時，如第十七圖；即 A 之大量數量與 B 之小量數量相連；A 之小量數量與 B 之大量數量相連。如甲生之



第十七圖

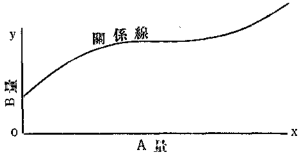
英文程度第一，其國文之程度則爲末第一；乙生之英文程度第二，其國文之程度爲末第二等等。

前三種關係圖形，爲 A, B 關係之大概；他種 A,B 之關係線亦甚多。如第十八圖，A 量增至若干時，



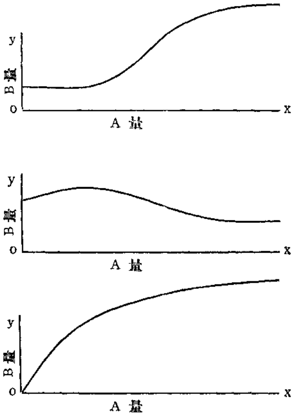
第十八圖 此關係線表明在A之若干量以下時，B之平均量不生變動

B之平均量無變動；自該點以上，則A量增加，B量亦隨之增加。又如第十九圖，在A之低端及上端，A



第十九圖 此關係線表明在A量若干限度以內，B量不變動

增加，B亦增加；但在A之中段，則B無變動。其他各種關係線之變動，如第二十圖各形。



第二十圖　各種關係線之形狀

r之數值　表明完全正關係r之數值爲1。若r由1次第減小,如.90,.75,.42,.01等等,則關係密切之程度,亦漸減少;迨r減至零時,則關係之程度爲零。表明完全負關係r之數值爲-1;r由-1次第增加,則此相反之關係次第減少;迨r增至零時,則此相反之關係消失。實際上之測量,其關係之程度,大抵在1與

－1之間。而γ表明正關係時，多為1與0間之小數；表明負關係時，多為0與－1間之小數。γ之值在.60或.70以上時，其關係為最高；由.35或.40至.50或.60時，可以表明關係之顯著；由.15或.20至.35或.40，表明關係之低小；不及.15或.20時，則其關係無足重輕。至γ為1,0,－1之理由，就以後求γ各公式中求之，即可以知矣。

16 關係數量之求法

求關係數量之方法，通常所用者，約有下列四種：

(1)乘積率(product－moment)法 (a)求關係數之公式：$\gamma=\frac{\Sigma xy}{N\sigma_x\sigma_y}$ 再求簡捷，可將此式變成 $\gamma=\frac{\Sigma xy}{\sqrt{\Sigma x^2\Sigma y^2}}$ 因$N\sigma_x\sigma_y=N\sqrt{\frac{\Sigma x^2}{N}\cdot\frac{\Sigma y^2}{N}}=N\frac{\sqrt{\Sigma x^2\cdot\Sigma y^2}}{N}=\sqrt{\Sigma x^2\cdot\Sigma y^2}$ 如第二十二表，係十人之觀念聯絡及注

第二十二表 十人之觀念聯絡及注意力

個人	觀念聯絡 A量	注意力 B量	A之各數量與平均數之差數 x	B之各數量與平均數之差數 y	x^2	y^2	xy
1	85	100	15	37	225	1369	555
2	85	70	15	7	225	49	105
3	83	90	13	27	169	729	351
4	78	80	8	17	64	289	136
5	76	30	6	－33	36	1089	－198
6	73	40	3	－23	9	529	－69
7	70	30	0	－33	0	1089	0
8	66	80	－4	17	16	289	－68
9	48	40	－22	－23	484	529	506
10	36	70	－34	7	1156	49	－238
平均數	70	63			2384	6010	1080

$\Sigma x^2=2384$ $\Sigma xy=1653-573=1080$

$\Sigma y^2=6010$

$$\gamma=\frac{1080}{\sqrt{(2384\times6010)}}=\frac{1080}{\sqrt{14327840}}=\frac{1080}{3785.21}=.29$$

意力之測量。其作法之次第：—

1. 將每人之觀念聯絡及注意力之分數，分別列在A量及B量之欄內。

2. 將A量之各數量相加，用人數平均，求平均數70；次將此平均數與各數量相減，將各差數列入x欄內，用正負號標明之。

3. 以同一方法求B量欄內之平均數及差數，將此差數列入y欄內。

4. 將每一差數自乘，求x^2及y^2；將x^2及y^2相加，求Σx^2及Σy^2：

5. 將相對數量之差數x,y相乘，求xy(注意正負號)，列入xy欄內；再將xy用代數法相加，求xy之和Σxy。

6. 將各數代入公式$\gamma=\frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$，即得關係數$\gamma$。

既求得γ，即可表明二量關係之程度。

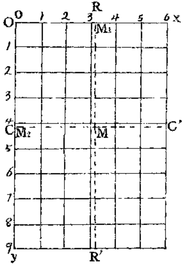
惟如第二十二表所列之方法，當數量較少時用之較為簡捷。然其功用甚不完備：第一，若當數量加多之時，則計算繁雜；第二，不能將各個數量分配之位置，表示明顯；第三，僅能表明二量關係之有無及關係之大小，未能表明A,B互相關係之程度。蓋A,B二量中其一量有單位之變更時，他一量隨之有若干之變更，亦須闡明。如後第二十五表之材料，係測六十一人之想像力及記憶力。若就二者之關係言，則其關係數為·52。然想像力影響於記憶力較大，其影響之

程度爲 ·69。反之，記憶力影響於想像力較小，其影響之程度僅爲 ·39。卽前者想像力有單位之變更時，則有 ·69 之記憶力隨之變更；後者記憶有單位之變更時，則隨其變更之想像力爲 ·39，其量較微。

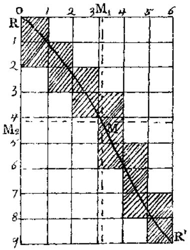
以上三者皆爲關係法中重要之點。解決第一，第二兩問題，須製關係表（Correlation table）；解決第三問題，須求消長之係數（regression coefficient）。分別說明之於下：

(b)求消長係數之公式：$b_1=r\frac{\sigma_x}{\sigma_y}$，$b_2=r\frac{\sigma_y}{\sigma_x}$ 前第十四圖中之 CC' 線，係與各豎行各中點最相近之直線；而各橫行各中點，亦有相近之直線。畫此等直線之時，僅就觀察判斷之力，必不正確；若當數量分配星散而各平均點參差過甚之時，更非易易。故必須有一定之公式，作求此直線之標準，如第二十一圖，以 RR' 線連接橫行之各平均點；以 CC' 線連接豎行各平均點。ox，oy 爲A,B二量之測量表；01,12,23 等爲組距。M_1爲 ox 之平均點；M_1 M爲 ox 之平均線。M_2 爲 oy 之平均點；M_2M爲 oy 之平均線。假使A,B 二量各自獨立，毫不相關，則各豎行及各橫行之次數分配相似；橫行平均點之直線 RR'，可與M_1M相合；豎行平均點

之直線CC',可與 M_2M 相合。此爲極端無關係時之情形。反之,極端有關係之情形時,則各個數量之分配密集於一線之左右,佔每行方格之數很少,而橫行與豎行平均點之各直線,幾立於一線上,如第二十二圖

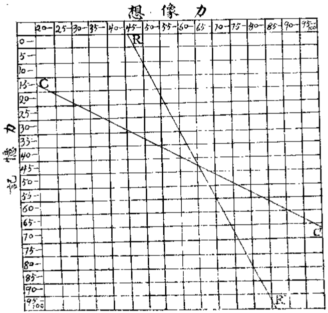


第二十一圖

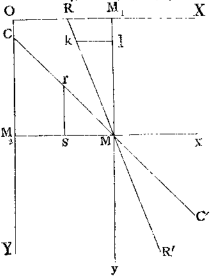


第二十二圖

之RR'線。但就實際言之,此類極端之情形甚少;CC'與RR'雖成直角,亦難相重合,而成銳角之時居多,如二十三圖。求RR'直線之公式爲 $x=r\frac{\sigma_x}{\sigma_y}y$,求CC'



第二十三圖



第二十四圖

直線之公式爲$y=\gamma\frac{\sigma_y}{\sigma_x}x$。二式中之$\gamma\frac{\sigma_x}{\sigma_y}$，$\gamma\frac{\sigma_y}{\sigma_x}$，皆爲消長係數。再將求此二切合直線公式之由來及消長係數之意義，次第說明於下：

就第二十四圖言之，設橫行各平均點皆立於RR'線上。設M_2爲oy之平均點，M_2x爲oy之平均線；再作RR'線，使之與M_2x線相交於M點。然後經過M點向ox畫一垂直線，則此垂直線必交於ox之平均點M_1。因設RR'線與M_1M相交所成之角度爲b_1，則b_1爲$\angle M_1MR$之正切$(=\frac{kl}{lM})$所含之角度。再設各數量與My之差數爲x，與Mx之差數爲y，則$\frac{x}{y}=b_1$，$x=b_1y$。若每一橫行y之數量有n個時，則$\Sigma(x)=n.b_1y$。若按全表計算，則$\Sigma(ny)$爲由M_2平均點各差數之和$=o$，由是$\Sigma(x)=b_1.\Sigma(ny)=o$，故$M_1$必爲ox之平均點，M可名爲全表之平均點。既知RR'線必過全表之平均點M，即可求b_1。設P爲相關之各x，y乘積之和之平均數；即$P=\frac{\Sigma(xy)}{N}$，則$NP=\Sigma(xy)$；求每一橫行之$\Sigma(xy)$時，因$\Sigma(x)=n.b_1y$，則$\Sigma(xy)=y.\Sigma(x)=nb_1y^2$。按全表計算，$\Sigma(xy)=b_1.\Sigma(ny^2)=Nb_1.\sigma_y^2$。因$\Sigma(xy)=NP=Nb_1\sigma_y^2$，故$b_1=\frac{P}{\sigma_y^2}$(1)；同理，設豎行之平均點皆立於CC'直線上，則CC'與橫線M_2M所成之

角度爲 b_2；由是，$b_2=\frac{P}{\sigma_x^2}$ (2)。(1)，(2) 二式常用下式表明之。因 $\gamma=\frac{\Sigma xy}{N\sigma_x\sigma_y}=\frac{P}{\sigma_x\sigma_y}$，$P=\gamma\sigma_x\sigma_y$；則 $b_1=\gamma\frac{\sigma_x}{\sigma_y}$，$b_2=\gamma\frac{\sigma_y}{\sigma_x}$。即 $\frac{x}{y}=\gamma\frac{\sigma_x}{\sigma_y}$，$x=\gamma\frac{\sigma_x}{\sigma_y}y$ (3)；$\frac{y}{x}=\gamma\frac{\sigma_y}{\sigma_x}$，$y=\gamma\frac{\sigma_y}{\sigma_x}x$ (4)。由 (3) 公式可求 RR' 線，由 (4) 可求 CC' 線。用 A，B 二量之絕對價值 X，Y 代差數 x，y 亦可。

$b_1=\gamma\frac{\sigma_x}{\sigma_y}$，$b_2=\gamma\frac{\sigma_y}{\sigma_x}$ 之稱爲消長係數者，因 b_1 爲 x 依 y 爲消長之係數，即按平均計算，y 有一單位之變化時，x 隨之變化若干，b_2 爲 y 依 x 爲消長之係數，即按平均計算，x 有一單位之變化時，y 隨之生變化若干。

(C)關係表之製法　在第三章曾說明次數分配列表之方法。但該種分配表，係排列一種測量。若將二種測量列入一表時，則其法與前稍有不同。劃分組數之方法，及定組距之界限，與前無異。其次則先定 ox 及 oy 二測量表之直線，使成直角相交於 o；通常皆使 oy 線爲豎線，列於左方；ox 爲橫線，列於上方。小數量皆由 o 點起，循 oy 線自上而下排列 B 量；循 ox 線自左而右排列 A 量（凡此皆任意定之，但此種列法較利

於說明)看第二十五圖即可明瞭。用方格紙排表最爲方便。

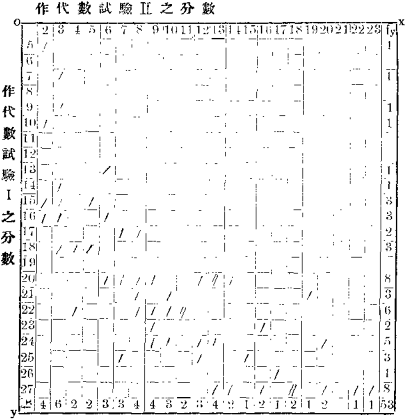
第二十三表　五十三名學生所作之代數試驗Ⅰ與試驗Ⅱ之分數

試驗Ⅰ	27	27	27	16	27	18	27	9	15	15	21	20	26	10	22	24	16	13	23
試驗Ⅱ	20	18	14	3	13	3	16	3	3	7	8	8	17	2	9	20	2	6	9

試驗Ⅰ	15	22	20	17	21	20	20	15	22	23	27	16	25	22	14	17	25	22	5
試驗Ⅱ	2	11	6	8	19	7	9	5	8	13	18	6	12	11	3	7	17	4	2

試驗Ⅰ	25	18	27	20	18	27	24	24	24	22	21	20	20	20	24
試驗Ⅱ	15	4	23	12	5	22	10	9	12	10	10	13	13	14	13

測量表既定之後,再按所測之每對數量之價值,依次排列。如第二十三表,以代數試驗II爲 A 量,循 ox 線排列;以代數試驗 I 爲B量,循 oy 線排列。試驗I所得之數量 27,與試驗II所得之數量20爲一對;數量 27 與 18 又爲一對。列入關係表中之時,循 oy 線得數量 27,循 ox 線查得數量 20;然後沿 27 欄向右橫數,至與其相對之數量 20 之竪欄內爲止,即在該方格之中畫一竪線;依此方法將其他每對數量記入各相當方格之內,如第二十五圖。再另作一表,將前表



第二十五圖 排列第二十三表之材料

每方格所畫之豎線，用數字書出之，列入此表各相當方格之中，如第二十四表。而每方格中數量之價值，假定其等於每方格之中點。就此表之數字，即可考察

二量關係之大概；卽 A 之大數量，多與 B 之大數量相連；A 之小數量，多與 B 之小數量相連，其關係之傾向，由上左方斜向下右方；可知其關係數概爲正數；若所記入之數量由上右方斜向下左方；

第二十四表　排列第二十三表之材料

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	fy
5	1																						1
6																							
7		1																					1
8																							
9		1																					1
10	1																						1
11																							
12																							
13					1																		1
14		1																					1
15	1	1		1																			3
16	1	1			1																		3
17						1	1																2
18		1	1	1																			3
19																							
20					1	1	1	1			1	2	1										8
21							1		1									1					3
22			1				1	1	1	2													6
23								1							1								2
24								1	1		1	1							1				5
25						1					1			1									3
26																1							1
27												1	1		1		2		1		1	1	8
28	4	6	2	2	3	3	4	4	3	2	3	4	2	1	2	1	2	1	2		1	1	53

卽A之大數量與B之小數量相連，A之小數量與B之大數量相連，其關係數必爲負數。至特別之數量如第二十四圖2與5，3與7二對數量與一班數量相差過遠，亦可由此表查出。

（d）就關係表求r，$\sigma_x\sigma_y$之方法　列表旣成之後，須將公式$r=\frac{\Sigma xy}{N\sigma_x\sigma_y}$中各部分之數值一一求出：卽先求數量總數N；A數量之標準差數σ_x；B數量之標準差數σ_y；及A，B各相關數量差數乘積之和Σxy。然後將各部分得數代入公式，卽可求出r之數值矣。試就第二十五表說明之如下：

1. 將各橫行之次數相加列入f_y欄，將各豎行之次數相加列入f_x欄；再將f_y，f_x欄各次數相加（f_y之數目與f_x相等），書其和數於f_y，f_x相交之方格中；如表，N=61

2. 次求σ_x及σ_y；用求標準差數之簡法（參看前章11，(4)，(b)節）。

a. 估計B量（記憶力）含有平均數之組爲40—44.9，其估計平均數=42.5；A量（想像力）含有平均數之組爲60—64.9，其估計平均數爲62.5。即在此二組之上下各畫粗線二條，以示區別。

b. 求A，B各數量與估計平均數相差之單位，如表中d欄。

c. 用A，B各次數與d相乘，求fd，如表中fd欄（注意正負號）

d. 將fd用代數法相加，求其和數；$\Sigma fd_y=63$　$\Sigma fd_x=26$

e. 用N（=61）除Σfd_y，Σfd_x求更正數；$C_y=\frac{63}{61}=1.03$，

$C_x = \frac{26}{61} = .42$

f. 將 C_y 及 C_x 自乘；$C_y^2 = .13$，$C_x^2 = .18$

g. 將 fd 用 d 乘之，求 Σfd^2，如表中 fd^2 欄；$\Sigma fd_y^2 = 1305$ $\Sigma fd_x^2 = 740$

h. 用 N 除 Σfd_y^2 及 Σfd_x^2，求 S_y^2 及 S_x^2；$S_y^2 = 21.4$，$S_x^2 = 12.13$

i. 由 $S_y^2 - C_y^2$ 求 σ_y^2，由 $S_x^2 - C_x^2$ 求 σ_x^2；$\sigma_y^2 = 21.4 - .13 = 21.27$，$\sigma_x^2 = 12.13 - .18 = 11.95$

j. 將 σ_y^2 及 σ_x^2 開方，求 σ_y 及 σ_x，$\sigma_y = \sqrt{21.27} = 4.61$，$\sigma_x = \sqrt{11.95} = 3.46$

σ_y 及 σ_x 皆以組距為 1 記算，然無須用原組距單位乘之，使還原；因尚須用此 1 作單位求 $\Sigma x'y'$ 故也。

3. 求 $x'y'$ 之和；由表中二估計平均數欄推算最易。

a. 求每一方格中之 $\Sigma x'y'$，然後將此數書於該格右上方；如為負數，須用負號注明之。由 A 及 B 小數一端算起，B 之 0－4.9 組中與 A 之 55－59.9 組中有數量一個，此一數量與 B 之估計平均數相差為－8，與 A 之估計平均數相差為－1；$8 \times -1 = 8$；再用次數乘之，$8 \times 1 = 8$，即此方格 $\Sigma x'y' = 8$，將此數 8 書於方格之右上方。又如 B 之 10－14.9 組中與 A 之 40－44.9 組中有數量一個，此一數量與 B 之估計平均數相差為－6，與 A 之估計平均數相差為－4；故此方格之 $\Sigma x'y' = -6 \times -4 \times 1 = 24$，即該方格中右上方之數。又如 B 之 50－54.9 組中與 A 之 50－54.9 組中有數量二個，此二個數量與 B 之估計平均數相差為 2，與 A 之估計平均數相差為－2，故此方格

之$\sum x'y' = 2 \times -2 \times 2 = -8$，卽該方格右上方之數。依此類推，將凡有數量方格中之$\sum x'y'$全行算出。定各　$x'y'$之正負號最易。在估計平均數上下畫四粗線，將表分成四部分，如下圖：

左上部	$- \times -$ $= +$	$+ \times -$ $= -$	右上部
左下部	$- \times +$ $= -$	$+ \times +$ $= +$	右下部

就此圖可見表之左上部，及右下部各$\sum x'y'$皆爲正數；右上部及左下部各$\sum x'y'$皆爲負數。在A，B平均數之組中各$\sum x'y'$皆爲零，故不計算。

b. 將每一橫行之各$\sum x'y'$相加（有正負數則相消），將此和數列於表右方$\sum x'y'$欄之下，正負各數分列兩行最好。

c. 將$\sum x'y'$欄下各正數及負數相加，使正負數相消，求$\sum x'y'$之總數，如表得$583 - 57 = 526$

4. 求r。前節所得之$\sum x'y'$，旣係由估計平均數推算，因之亦必須有更正數，方能得實在的$\sum xy$。每一差數x'必須更正數C_x；每一次差數y'必須更正數C_y；故每一$x'y'$須有更正數$C_x C_y$；按全表計之，即$\dfrac{\sum x'y'}{N} - C_x C_y$。　$r = \dfrac{\sum x'y' - NC_x C_y}{N\sigma_x \sigma_y}$ 或

$$\frac{\dfrac{\sum x'y'}{N} - C_x C_y}{\sigma_x \sigma_y} \; *$$

如前表$\sum x'y' = 526$　$C_x C_y = .42 \times 1.03$

$N\sigma_x \sigma_y = 61 \times 3.46 \times 4.61$　故$r = \dfrac{526 - (.42 \times 1.03 \times 61)}{61 \times 3.46 \times 4.61} = \dfrac{526 - 26.39}{61 \times 3.46 \times 4.61} = .52$

*此式之證明如下：

「設 M_x, M_y 為 A,B 之實得平均數

E_x, E_y 為 A,B 之估計平均數

x,y 為 A,B 之各數量與 M_x, M_y 相差之數

x',y' 為 A,B 之各數量與 E_x, E_y 相差之數

則 $x'=x+C_x,\ y'=y+C_y$

$\Sigma x'y'=\Sigma(x+C_x)(y+C_y)=\Sigma xy+C_y\Sigma x+C_x\Sigma y+\Sigma C_xC_y$

而 Σx 及 Σy 為由實得平均數求出之差數，其和等於零。

是以 $\Sigma x'y'=\Sigma xy+\Sigma C_xC_y$　即 $\Sigma xy=\Sigma x'y'-\Sigma C_xC_y$

以之代入公式 $\gamma=\dfrac{\Sigma xy}{N\sigma_x\sigma_y}$ ，

則 $\gamma=\dfrac{\Sigma x'y'-NC_xC_y}{N\sigma_x\sigma_y}$　或　$\dfrac{\dfrac{\Sigma x'y'}{N}-C_xC_y}{\sigma_x\sigma_y}$　」

5. 求消長係數 $b_1=\gamma\dfrac{\sigma_x}{\sigma_y}$，$b_2=\gamma\dfrac{\sigma_y}{\sigma_x}$。將所求之 $\gamma, \sigma_x, \sigma_y$ 之數代入即得；此例 $b_1=.52\dfrac{3.46}{4.61}=.39$　$b_2=.52\dfrac{4.61}{3.46}=.69$

6. 算RR'，CC' 二關係線之公式，$x=\gamma\dfrac{\sigma_x}{\sigma_y}y$，$y=\gamma\dfrac{\sigma_y}{\sigma_x}x$；此例 $x=.39y$，$y=.69x$

7. 求 RR'，CC' 二關係線。用 X，Y 之絕對價值代 x,y，較為簡便。此例 $M_x=62.5+.42\times5=64.6$　$M_y=42.5+1.03\times5=47.65$

設　X 為A量中任一數量，Y 為B量中任一數量

則　$x=X-64.6$，　$y=Y-47.65$

因　$x=.39y$，則 $X-64.6=.39(Y-47.65)$，$X=46.02+.39Y$

若　$Y=0$，則 $X=46.02$；$Y=100$，$X=85.02$

因 $y=.69x$，則 $Y-47.65=.69(x-64.6)$，$Y=3.08+.69x$

若 $x=20$，則 $Y=16.88$；$x=100$，$Y=72.08$

如前第二十三圖，按此四數值畫四點，連以RR'，CC'二直線，則RR'，CC'相交於表之中央點M；$M_x=64.6$，$M_y=47.65$。RR'與各橫行平均點之實在線接近；CC'與各豎行平均點之實在線接近。就此二線之位置，亦可觀A,B二量關係之情形也。

(2)等級差異法(method of Rank-differences)

(a)由 ρ 求 r 法　用乘積率法求 r 之時，係將各個數量與平均數之差數d，全行計算。然若不注重各數量差數之大小，僅就相關各數量所居之等級，亦可求A,B二量之關係。

就A,B各數量所居之等級求 r 時，其公式爲：

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

式中之 ΣD^2 爲表明A,B各相關數量等級之數字差數之平方和。N爲相關數量之對數。如第二十六表，A,B各有相關之數量十對，則N=10。先按A之各數量之大小排列1, 2, 3, 4, 5, 6, 7, 8, 9, 10之等級，如表之第四欄。再排列B之各數量之等級，如表之第五欄。然後求各對數量等級之差數；如A大於B，則其差爲

第二十六表　用等級法則求γ之實例

個人	A量 觀念聯絡	B量 注意力	A之等級	B之等級	A,B等級之差數D +	0	−	ΣD^2
1	85	100	1.5	1	.5			.25
2	85	70	1.5	5.5			−4	16
3	83	90	3	2	1			1
4	78	80	4	3.5	.5			.25
5	76	30	5	9.5			−4.5	20.25
6	73	40	6	7.5			−1.5	2.25
7	70	30	7	9.5			−2.5	6.25
8	66	80	8	3.5	4.5			20.25
9	48	40	9	7.5	1.5			2.25
10	36	70	10	5.5	4.5			20.25
總數					12.5		−12.5	89.00

正，列入第六欄；如A,B相等則其差爲0，列入第七欄；如A小於B，則其差爲負，列入第八欄。將各差數自乘，列入第九欄，$\Sigma D^2=89$。求γ時須先求ρ。此例N=10，依公式$\rho=1-\frac{6\times 89}{10\times 99}=1-.54=.46$。下列之第二十七表，係分列$\rho$及與$\rho$相當之$\gamma$價値，既求得$\rho$，再按表求$\gamma$甚易。此例$\rho=.46$，$\gamma=.477$。

第二十七表 由 ρ 之價值求 γ 表

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$$

ρ	γ	ρ	γ	ρ	γ	ρ	γ
.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	.8121
.09	.0942	.34	.3542	.59	.6081	.84	.8516
.10	.1047	.35	3645	.60	.6180	.85	.8610
.11	.1151	.36	.3748	.61	.6250	.86	.8705
.12	.1256	.37	.2850	.62	.6379	.87	.8799
.13	.1360	.38	.3935	.63	.6478	.88	.8893
.14	.1465	.39	.4056	.64	.6577	.89	.8936
.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	.9009
.25	.2611	.50	.5176	.75	.7654	1.00	1.0000

排列 A,B 各數量等級之時，遇有等級相同之數量，則各數量須佔其位置之平均數。如第二十六表中 A 之 85 與 85 二數量應居 1,2 兩位置；然因其價值相同，故佔其等級之平均數 1·5 及 1·5。又如 B 之 70 兩數量，應佔等級 5 與 6，而平均佔 5·5 及 5·5。由是A,B各最大之等級數，可恰與數量之對數相等。

(b)由R求γ法　$\gamma=2\text{cosine}\frac{\pi}{3}(1-R)-1,\ R=1-\frac{6\Sigma g}{N^2-1}$

此法與前法相似,而更簡易。前法用等級差數之平方和,此則僅用正差數之和。求γ時先求R;既得R,則按第二十八表,察得相當γ之值。如第二十六表,正差數之和爲12•5, N = 10, $R=1-\frac{6\times 12.5}{10^2-1}=1-\frac{25}{33}$ $=1-.76=\cdot 24,\ \gamma=\cdot 399$

第二十八表　由R之價値求γ表

$$R=1-\frac{6\Sigma g}{N^2-1},\ \gamma=2\text{cosine}\frac{\pi}{3}(1-R)-1$$

R	γ	R	γ	R	γ	R	γ
.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						

(3)異號差數對數法 (method of Unlike Signed Pairs) $\gamma = \text{cosine}\ \pi\ U$

由 U 求 γ 時,先求 A,B 之各數量與各平均數相差之數,用正負號注明之。A 之十差數,與相當 B 之十差數,及 A 之一差數與相當 B 之一差數,皆爲同號之對數,以 l 表明之。十一 與 一十 爲異號之對數,以 u 表明之。$l+u$=N(對數總數)。故異號對數 u 與總對數有相當之百分比例數。若 A,B之差數無 0 時,由 $\frac{u}{N}$ 之數,照二十九表 即可得相當之 γ。若A,B之差

第二十九表 由 U 之百分比例數求 γ 表

U 爲異號差數對數與對數總數之百分比例數。若以 U 爲 l 之百分比例數,則 γ 之值爲負數。是以當 U 過.50時,可按 l 之百分比例數,求 γ 之負數。

U	γ	U	γ	U	γ	U	γ
.00	1.0000	.13	.9174	.26	.6848	.38	.3682
.01	.9996	.14	.9044	.27	.6615	.39	.3387
.02	.9982	.15	.8905	.28	.6375	.40	.3089
.03	.9958	.16	.8757	.29	.6129	.41	.2788
.04	.9924	.17	.8602	.30	.5877	.42	.2485
.05	.9880	.18	.8439	.31	.5620	.43	.2180
.06	.9826	.19	.8268	.32	.5358	.44	.1873
.07	.9762	.20	.8089	.33	.5091	.45	.1564
.08	.9688	.21	.7902	.34	.4819	.46	.1253
.09	.9604	.22	.7707	.35	.4542	.47	.0941
.10	.9510	.23	.7504	.36	.4260	.48	.0628
.11	.9407	.24	.7293	.37	.3973	.49	.0314
.12	.9295	.25	.7071			.50	.0000

數有0時，如 00,+0,0+,-0,0-各對數量，可用下式求U之數値，然後求γ。$U=\dfrac{u+\left(\dfrac{\frac{u}{u+l}+\frac{1}{2}}{2}\right)d}{N}$

式中d代表有零數之差數對數，即將所有d之對數，變爲l及u；其比例卽按$\frac{1}{2}$與$u+l$之比例之半數。其如此支配d數之理：第一，凡含有零數之對數可爲l，亦可爲u；第二，若精密區分之，含零之對數爲l及u之比例，與l及u佔其他對數之比例，易成相同之正比，不易成相同之反比。譬如$l=90$，$u=10$，$d=10$，若用精密之區分法，則此10零對差數爲l之數，不能與爲u之數相同；且易爲$9l$，$1u$；不易爲$1l$，$9u$。如第二十二表，$l=5$，$l=4$，$d=1$，則

$$U=\frac{4+\left(\dfrac{\frac{4}{5+4}+\frac{1}{2}}{2}\right)\times 1}{10}=\frac{161}{360}=\cdot 45$$ 按二十九表，得$\gamma=\cdot 156$

以上所述之四種求γ之法（$\gamma=\frac{\Sigma xy}{N\sigma_x\sigma_y}$，$\gamma=2\,\text{sine}\left(\frac{\pi}{6}\rho\right)$，$\gamma=2\,\text{cosine}\,\frac{\pi}{3}(1-R)-1$，$\gamma=\text{cosine}\,\pi U$），苟A,B各數量之分配爲對稱或偏斜輕微，其間之關係線爲直線的之時，則由此四法所求之γ，其數無大差異。否則若當N過於30之時，以用第一法爲最好；

因A,B各個數量之價值，皆直接顯計算上之效力。且用關係表排列各數量之分配，可使其分配之眞象明瞭。就其分配之情形，可以推論所求之結果。若當N不及30之時；可用等級差異法求r，以明A,B間關係之有無；然不可就此卽斷定其間關係密切之程度。總之，數量較少之時，所得之r甚難近於精確；而用各法求得之r，亦彼此互異。欲就r以下結論，必須格外加意也。

17 求實得r之眞確價值法

前節所述測量A,B之關係；係假定測得其間之眞確關係。求得r，卽決定其關係之情形。然按之實際，此r之價值，難表A,B間之眞確關係。因測量之時，於每一相關之數量中必有若干之誤差(error)。此種誤差可分二種:(a)變性誤差(variable error)或簡稱變差；(b)恒性誤差(constant error)，或簡稱恒差。凡求A,B之r時，往往有此種誤差混入，使r之價值減少或加大，失其眞確之價值。下列各法卽就實得之r，以求眞確之r。

(1)更正r受變差之方法　變差卽偶然之誤差。此種誤差，恒由於測量工具之變動，及測量者觀察

之精粗等等而生。如甲生之記憶力爲54分，是僅就此一次所測之結果而言；若再測量一次或數次時，其結果不能毫無變動，終有多少之差異。此種偶然之誤差在求中量之時，若N愈增加，其間所含之大小誤差可以兩相抵消，使所求之結果平衡，去眞確之程度愈近；然在求關係數時，僅將數量增加，斷不能使其無害於結果。A,B各數量之間，若有此偶然之誤差，必使其關係數減小；無論其關係數爲正爲負，皆與零數接近。若各數量全出於偶然，則其關係數必爲零。故每得一小量之關係數時，可表明二種意義：(a) A,B實在之關係缺乏；(b) 偶然誤差之發生過大。r受變差之影響減小時，可名曰微弱(Attenuation)

欲更正r微弱之量，求其眞確之價值時，至少必須將A,B各測量二次，求兩種獨立之數量。得此兩種獨立之數量，卽可就下列之方法求眞確之r：

設A,B爲相關之兩種事實

設p爲A之眞確數量

設q爲B之眞確數量

設r_{pq}爲A,B之眞確關係數

設p_1及p_2爲A之二種獨立數量

設 q_1 及 q_2 爲 B 之二種獨立數量

設 $r_{p_1q_1}$ 爲 A 之第一種數量與 B 之第一種數量之關係數

設 $r_{p_1q_2}$ 爲 A 之第一種數量與 B 之第二種數量之關係數

設 $r_{p_2q_1}$ 爲 A 之第二種數量與 B 之第一種數量之關係數

設 $r_{p_2q_2}$ 爲 A 之第二種數量與 B 之第二種數量之關係數

設 $r_{p_1p_2}$ 爲 A 之兩種數量之關係數

設 $r_{q_1q_2}$ 爲 B 之兩種數量之關係數

$$\text{則 } r_{pq}=\frac{\sqrt[4]{(r_{p_1q_1})(r_{p_1q_2})(r_{p_2q_1})(r_{p_2q_2})}}{\sqrt{(r_{p_1p_2})(r_{q_1q_2})}}$$

若用 $r_{pq}=\dfrac{\sqrt{(r_{p_1q_2})(r_{p_2q_1})}}{\sqrt{(r_{p_1p_2})(r_{q_1q_2})}}$ 更正 r，更可精當，且易計算。

就上式雖可將眞確之關係數求出，然其繁複之弱點仍不能免。若測量之時，能於所取實例慎加選擇，則事半功倍矣。

(2)更正 r 受恒差之方法

變差之起，各自獨立，彼此不相繫屬；卽一數量所受變差之影響，與他一數量無關。恒差則不然，其性質

爲不易的，其影響及於A或B列數量之全體；或A,B二列數量並受其影響。然無論其影響一列或二列數量，皆足使所要考察之A,B關係，不得顯其眞象。

恒差之影響於二列數量者；例如測量兒童心理能力之關係時，A,B二種心理能力，可隨兒童之年齡發達。若就不同年齡之兒童，測A,B能力之關係，卽A與B當同一年齡時全無關係；然因年齡之影響，其間亦必有關係實現。

由某一原因所生之恒差，影響於任一列數量使 r 之値減少時，可名曰縮減(Constriction)。反之，遺漏其關係，使r增加時，可名曰脹大(Dilation)。若恒差影響於所比較之二列數量時，可名曰偏僻(Distortion)。

更正 r 縮減之量時，其式爲：$r_{pq}=\dfrac{r_{pq}'}{\sqrt{1-r_{pv}^2}}$

式中之 r_{pq}' 爲A,B之實得關係數，

r_{pv} 爲A，B中之一量與混入之第三量之關係數。

r_{pq} 爲A,B之眞確關係數

恒差之影響於一列數量者；Spearman曾舉一例，如考察音度之識別力與學業等級之關係數爲.49。然同時査知所測量之兒童，多半受過音樂之訓練。受

過音樂訓練之兒童，其分辨音度之能力與未受此種訓練之兒童，自不相同，必使 r 減少。故須將此關係除去。又求得此二者之關係數爲 .61。按前式，則：

$$r_{pq}=\frac{.49}{\sqrt{1-.61^2}}=.62$$

設再有不合宜之 r_{pw} 混入時，其式爲：

$$r_{pq}=\frac{r_{pq}'}{\sqrt{1-r_{pv}^2-r_{pw}^2}}$$

更正 r 脹大之量時，其式與前正相反：

$$r_{pq}=r_{pq}'\cdot\sqrt{1-r_{pv}^2-r_{pw}^2}$$

更正 r 偏僻之量時，其式爲：

$$r_{pq}=\frac{r_{pq}'-(r_{pv})(r_{qv})}{\sqrt{(1-r_{pv}^2)(1-r_{qv}^2)}}$$

式中 r_{pq}' 爲A,B之實得關係數

r_{pv}, r_{qv} 爲 A,B 各與混入之第三量之關係數。

r_{pq} 爲 A,B 之眞確關係數。

設使應有第三量之關係而未使之混入時，求其眞確 r 之式爲：

$$r_{pq}=r_{pq}'\cdot\sqrt{(1-r_{pv}^2)(1-r_{qv}^2)}+(r_{pv})(r_{qv})$$

問 題

6 求第二十四表之關係數及消長係數。

以代數試驗 I 之分數爲 B，以試驗 II 之分數爲 A；以 21 爲 B之估計平均數，以 10 爲A之估計平均數

第七章

概率曲線 (Probability curve)

18 概率之原理

吾人日常推測某事時，慣用「大概如此」及「大概不如此」等語。所謂某事大概如此之意，卽謂某事易於如此實現也。所謂某事大概不如此之意，卽謂某事易於如此失敗也。而所以能下此概然之判斷者，蓋就昔日之經驗，常見某事於此相似情形之中，易於如此或不易於如此而已。

欲明此概然之分量，須以數學之方法定之。此概然之分量可名曰概率。以概率之大小，定某事成敗得失之程度。概率之數可爲 1 或 0；或 1 與 0 間之分數。例如一大組事物中任一事物或得或失之數，與此組事物得失之總數成相當之比例；卽總數得失之概率爲1，而其中任一事得失之概率，則爲小於1之某數。

先取最簡單之例說明之：如擲銅幣一枚，當其落下之時，正面向上之機會有 $\frac{1}{2}$；背面向上之機會亦有 $\frac{1}{2}$。若以正面向上爲得，則得之機會有 $\frac{1}{2}$，失之機會亦有 $\frac{1}{2}$。或得或失之概率爲 $\frac{1}{2}$，而得失概率之總數爲 $\frac{1}{2}+\frac{1}{2}=1$

又如以骰子一枚擲下一次之時，得幺之機會僅有 $\frac{1}{6}$，得其他各點亦各有 $\frac{1}{6}$ 之機會。得失之總數爲6，得任一面之概率爲 $\frac{1}{6}$，得失概率之總數爲 $\frac{1}{6}+\frac{5}{6}=1$

若一事物之得法有幾種，則得此事物之概率，等於得各種之概率之和。

例如一袋之中，盛有綠球12個，紅球18個，黑球19個。若由袋中取出任意一球之時，能得綠球之概率爲 $\frac{12}{49}$（綠球共12個，球之總數49個），能得紅球之概率爲 $\frac{18}{49}$，能得黑球之概率爲 $\frac{19}{49}$；而能得或綠或黑球之概率爲 $\frac{12}{49}+\frac{19}{49}=\frac{31}{49}$，能得或黑或紅球之概率爲 $\frac{19}{49}+\frac{18}{49}=\frac{37}{49}$；能得或綠或紅或黑球之概率爲 $\frac{12}{49}+\frac{18}{49}+\frac{19}{49}=1$

若由各不相關之若干單事合成一複事（compound event）時，則得此複事之概率，等於得各單事之概率之積。

例如有袋二個，甲袋中有黑球8個，紅球9個；乙袋中有黑球3個，紅球11個。在甲袋中取得黑球（一單事）之概率爲 $\frac{8}{17}$，在乙袋中取得黑球（一單事）之概率爲 $\frac{3}{14}$。若從二袋中同時各取出一球時，則得二球盡黑（二單事合成之複事）之概率爲 $\frac{8}{17}\times\frac{3}{14}$。因

甲乙二袋之球兩兩相合之法，共有 17×14 個；而甲袋 8 個黑球與乙袋 3 個黑球兩兩相合之法，共有 8×3 個。故得複合黑球之概率爲 $\frac{8\times3}{17\times14}$，得複合紅球之概率爲 $\frac{9\times11}{17\times14}$，得黑球與紅球複合之概率爲 $\frac{8\times11}{17\times14}+\frac{3\times9}{17\times14}$，三者概率之和 $\frac{8\times3}{17\times14}+\frac{8\times11}{17\times14}+\frac{3\times9}{17\times14}+\frac{9\times11}{17\times14}=1$ *

19　二項式之分配 (binomial distribution)

今設得作一事一次之概率爲 p，失之概率爲 q，則 $p+q=1, q=1-p$。試用代數法則說明得失之概率如下：

設有一事於此，試作一次時，失之概率可得 q，得之概率可得 p。設將 a，b 二事合作一次時，失 a 之概率可與失 b 之概率兩兩相合，則失之複合概率爲 $q\times q=q^2$；得 a 之概率可與得 b 之概率兩兩相合，則得之複合概率爲 $p\times p=p^2$；又得 a 之概率可與失 b 之概率相合，失 a 之概率可與得 b 之概率兩兩相合，則得失之複合概率爲 $p\times q+q\times p=2pq$，如第二十六圖第三行。設將 a, b, c 三事合作一次時，則失 a, b, c

* 以上之說明，多引用無錫顧澄先生所譯著之最小二乘法中"概率之原則"一段所述者；欲究其詳，請參考先先原著。

三者之概率可相合，其失之複合概率爲 $q \times q \times q = q^3$；得 a, b, c 三者之概率可相合，其得之複合概率爲 $p \times p \times p = p^3$；得 a, b, c 之任一事而失其餘二事之概率爲 $3 \times p \times q^2 = 3pq^2$；得 a, b, c 之任二事而失其餘一事之概率爲 $3 \times p^2 \times q = 3p^2q$，如圖第五行。設作 a, b, c, d 四事時，則失四者之概率爲 q^4，得四者之概率爲 p^4，得其任一事爲失其餘三事之概率爲 $4 \times p \times q^3 = 4pq^3$，得其任二事而失其他二事之概率爲 $6 \times p^2 \times q^2 = 6p^2q^2$，失其任一事而得其餘三事之概率爲 $4p^3q$，如圖第七行。

得之數目		0	1	2	3	4
1	一事	q	p			
2		q^2	$pq + pq$	p^2		
3	二事	q^2	$2pq$	p^2		
4		q^3	$pq^2 + 2pq^2$	$2p^2q + p^2q$	p^3	
5	三事	q^3	$3pq^2$	$3p^2q$	p^3	
6		q^4	$pq^3 + 3pq^3$	$3p^2q^2 + 3p^2q^2$	$3p^3q + p^3q$	p^4
7	四事	q^4	$4pq^3$	$6p^2q^2$	$4p^3q$	p^4

第二十六圖

由是試作某事時，其得失各項之概率，可以二項式展開表明之，即：

試作一事時，用 $(q+p)$ 二項展開式

試作二事時，用 $(q+p)^2$ 二項展開式

試作三事時，用 $(q+p)^3$ 二項展開式

試作四事時，用 $(q+p)^4$ 二項展開式

若將 n 個事物合作 N 次時，其二項式為 $N(q+p)^n$，其連續各項為 $N\left\{q^n+nq^{n-1}p+\frac{n(n-1)}{1.2}q^{n-2}p^2+\frac{n(n-1)(n-2)}{1.2.3}q^{n-3}p^3\cdots\cdots+p^n\right\}$ 此由理論方面表明次數分配之形式也。就前式之次數分配，可求其平均數及標準差數；其算法與前求 M 及 S.D. 相同：(1)列連續各項次數於第一欄；(2)列各項相差之數於第二欄，而以得之數 0 (0 successes) 為平均數；(3)以次數乘差數，列於第三欄；(4)以次數乘差數平方，列於第四欄；N 為通用之數，為便利計，可以省略。其式如下：

(1)	(2)	(3)	(4)
f	d	fd	fd^2
q^n	0	—	—
$nq^{n-1}p$	1	$nq^{n-1}p$	$nq^{n-1}p$
$\frac{n(n-1)}{1.2}q^{n-2}p^2$	2	$n(n-1)q^{n-2}p^2$	$2n(n-1)q^{n-2}p^2$
$\frac{n(n-1)(n-2)}{1.2.3}q^{n-3}p^3$	3	$\frac{n(n-1)(n-2)}{1.2}q^{n-3}p^3$	$\frac{3n(n-1)(n-2)}{1.2}q^{n-3}p^3$
⋮	⋮	⋮	⋮

設N為一單位(unity)之數,則第一欄各項之和自必為1;因此將第三欄各項相加,即得平均數。而第三欄之和為$np\left\{q^{n-1}+(n-1)q^{n-2}p+\frac{(n-1)(n-2)}{1.2}q^{n-3}p^2+\cdots\right\}$

$$=np(q+p)^{n-1}=np\text{,即平均數 }M=np$$

標準差數之平方,為第四欄各項之和減平均數之平方n^2p^2,即:$\sigma^2=np\left\{q^{n-1}+2(n-1)q^{n-2}p+3\frac{(n-1)(n-2)}{1.2}q^{n-3}p^2+\cdots\cdots\right\}-n^2p^2$

括弧中各項用二項式法解之(以$(1-p)$代q,去其相同之項數),得$(n-1)p+1$。故$\sigma^2=np\left\{(n-1)p+1\right\}-n^2p^2=np(np-p+1)-n^2p^2$

$$=n^2p^2-np^2+np-n^2p^2=np-np^2=np(1-p)=npq$$

就以上所述,用二項式之法則,不僅能求出某種次數分配之平均數及標準差數,並能表明次數分配之全體形式,此就理論方面而言;若從實際方面比較,與此相差亦不甚遠。Yule曾舉擲骰之例三個,其實得之分配與二項式展開之次數比較如下:

(1)(W.F.R. Weldon)擲骰十二個,共擲4096次,每次所擲之骰以見4,5或6各點為得,以1,2,3為失。依前節之理論,$p=q=.5$;理論平均數$M=np=12\times.5=6$;理論標準差數$\sigma=\sqrt{npq}=\sqrt{.5\times.5\times12}=1.732$。共得4,5或

6之次數如下:

得之數	次數	二項式各項之次數	得之數	次數	二項式各項之次數
0	–	1	7	847	792
1	7	12	8	536	495
2	60	66	9	257	220
3	198	220	10	71	66
4	430	495	11	11	12
5	731	792	12	–	1
6	948	924			
			總數	4096	4096

前表之平均數 $M=6.139$;標準差數 $\sigma=1.712$;平均得之數爲 $6.139\div12=0.512$,與理論之得數 .5 相差爲 .012

(2) (W.F.R. Weldon) 擲骰十二個,共擲4096次,而每次僅以 6 爲得數。理論之 $p=\frac{1}{6}$,$q=\frac{5}{6}$;理論平均數 $M=12\cdot\frac{1}{6}=2$;理論標準差數 $\sigma=\sqrt{\frac{1}{6}\times\frac{5}{6}\times12}=1.291$;共得 6 之次數如下:

得之數	次數	得數	次數
0	447	5	115
1	1145	6	24
2	1181	7	7
3	796	8	1
4	380	總數	4096

前表之M=2.00；σ = 1.296；平均得之數 = 2.00÷12=0.1667，與理論之得數 $\frac{1}{6}$=.1667 至第四位小數尙相同。

(3) (G.U.Yule)擲骰三個，共擲648次，以5或6爲得數。結果得次數如下：

得數	次數	二項式各項之次數
0	179	192
1	298	288
2	141	144
3	30	24
總數	648	648

前表之 M = 1.034；σ = 0.823；平均得數 = 0.345；理論之 $p=\frac{1}{3}$；$q=\frac{2}{3}$；理論M=1；理論σ=0.816

就前舉之三例而言，實際與數理方面頗相近似。而例(1)及(3)之二項式展開之次數分配，與實得之數目亦相差幾微。故用二項式法則求出之平均數，標準差數，及分配之形式等，與實際所得極相近似。其不相同之處，不過因所選之實例，稍有變動，致生差異耳。至其能得近似之結果，必須有與理論相合之條件。例如擲骰之時，自始至終，必須用同一之骰，不得中間改用他種。如有更換，則q，p之機會，前後不能相

等。非但每次所用之骰，須限於一種，即此一組中之各個骰，彼此亦須相同。否則此輕彼重，此偏利於得么，彼偏利於得2，則各個骰所生q,p之機會，亦不能相等。若各骰之六面不勻，所刻之點有深淺，卽一骰中各點之q，p機會，亦不相等。Weldon對於擲骰之法，曾提議用有綢紋之紙，上製小溝，將紙斜立，使骰滾過斜紋，則發生q,p之機會，當更可以均等矣。

前舉三例之中，有最可注意之點者，卽爲q,p之價值。例(1)之$p=q=\frac{1}{2}$，$n=12$，則項數$=13$；其概率最大之項爲中央之第七項。此中央項之左右各項之次數分配兩兩相近似（二項式則兩兩相等）。其曲線之形狀對稱。例(2)之$p=\frac{1}{6}$，$q=\frac{5}{6}$，$n=12$；最大概率之項爲第三項。例(3)之$p=\frac{1}{3}$，$q=\frac{2}{3}$，$n=3$，最大概率之項爲第二項。p,q之價值，在例(2)，(3)皆不相等。因之各項之次數分配，偏向於一方，不如例(1)之中正。故次數分配之形狀，與p,q及n之價值，皆有關係。若p,q相等，則二項式展開各項之次數分配必對稱；因任何項之p,q價值，可互換，而不影響於該項之價值。故距兩端相等各項之價值亦相等。若

p,q 之價值不等，則各項之分配偏斜。若 n 之價值不變，則分配愈偏向者，q,p 之機會愈不均等。下例之第三十表，係設 n=20，N=10,000，p 之價值按0.1遞增，用二項式推算其由 .1至.5之次數分配。p=.1，則得 2 之次數最大，得 1 之次數亦幾相等。而得 9 之

第三十表 10,000 $(q+p)^{20}$ 之二項式各項次數，p 之値由.1 至.5

（用最近之整數）

得數	p = .1 q = .9	p = .2 q = .8	p = .3 q = .7	p = .4 q = .6	p = .5 q = .5
0	1216	115	8	—	—
1	2702	576	68	5	—
2	2852	1369	278	31	2
3	1901	2054	716	123	11
4	898	2182	1304	350	46
5	319	1746	1789	746	148
6	89	1091	1916	1244	370
7	20	545	1643	1659	739
8	4	222	1144	1797	1201
9	1	74	654	1597	1602
10	—	20	308	1171	1762
11	—	5	120	710	1602
12	—	1	39	355	1201
13	—	—	10	146	739
14	—	—	2	49	370
15	—	—	—	13	148
16	—	—	—	3	46
17	—	—	—	—	11
18	—	—	—	—	2
19	—	—	—	—	—
20	—	—	—	—	—

次數在此10,000中亦有一次。及 p 之價值增加,而最大次數之位置亦次漸前進。次數分配兩末端之長度亦漸次相等。及 p=.5時,則分配爲對稱。p=.6時分配

第三十一表　$10,000(.9+.1)^{100}$之二項式各項次數。(用最近之整數)

得數	次數	得數	次數
0	— —	12	988
1	3	13	743
2	16	14	523
3	59	15	327
4	159	16	193
5	339	17	106
6	596	18	54
7	889	19	26
8	1148	20	12
9	1304	21	5
10	1319	22	2
11	1199	23	1

之形狀與 q=.6 時相似,惟顚倒其兩端而已。若p,q之價值不變,雖其價值不等,將 n 之價值增加,卽可減少分配偏斜之度。故 n 之價值愈增,則偏斜之度愈減少。前表之第一次數分配,係 p=.1,q=.9,n = 20;若 n=100,則其分配如第三十一表,得 9 及10爲最大之次數。而分配之兩端漸進於等長。若 n=2,則分配爲:

得數	次數
0	8100
1	1800
2	100

其最大之次數，偏在一端。故無論p,q之價值若何，苟n之量增加至充分時，即可認其分配爲對稱。因 $p-q$ 之價值，與標準差數 $\sqrt{npq}$ 相較，爲數甚微，其偏斜之度不顯著故也。

20 常則分配之推理

常則的分配，即對稱的分配。如前節所述，若q,p之機會均等，則最大概率之項居中，各項之分配爲對稱的。若n增加至充分之程度時，雖q,p之機會不等，而最大概率之項亦占中央，各項分配亦成對稱之形式。然此不過就代數之法則，以說明概率分配之形式耳。若本此理而演譯之，以推論心理與教育方面測量之結果，何以多能與此概率之分配相合，亦可得類似之推理。蓋每一測量之結果，恆由種種之原因(factors)複合而成；各個原因彼此獨立，不生相互之影響。而其造成此結果之機會，亦彼此均等(p=q)。其影響之勢力，或顯或隱，純出於機會公例，而無偏

彼偏此之情形。以故測量之結果，亦得循諸常軌，成對稱的分配。

再推而言之，若造成複合量之原因數目（=n）愈增，雖其機會不等（即 $p-q$ 與 $\sqrt{npq}$ 相較爲數甚小），而此諸原因分配之形式，亦可近於對稱；因之由此諸原因所合成之分配形式，亦必近於對稱。

反之，若原因之數目有限，而其中有一原因或數原因，特顯其影響之勢力者；則諸原因之分配，必構成偏斜之形；因之由此諸原因所合成之複合分配，其形式亦必偏斜。

多數學者率認以上之假定，合乎自然，據之以解釋對稱或偏斜分配之理。例如第九表體高之測量，所以能成規則之分配者，即因每一複合之量，由多種原因複合而成：如頭骨之大小，脊椎骨之高低，腿之長短，連接軟骨之厚薄，脊柱曲度之大小等等。而此各個原因，又由他種原因複合而成。如脊柱屈曲所以有大小之差者，又有他種原因爲之構成。因諸種原因之關係不甚密切，故可各顯其獨立之勢力，以造成複合之量，使各複合量之分配，成常則之形。又如兒童算術之能力，亦有種種原因集成此複合之量：如遺傳之

能力，身體發達之情況，家庭之環境，學校之訓練，社會之接觸等等，皆為構成算術能力之根本原因。而此種原因，亦恒彼此不相關係；其分配之形狀，近於對稱。故結果各人之算術能力，亦成對稱的分配。然若複合量原因之中，有特顯勢力之原因，則複合量之分配，不能呈對稱之形：如各年齡之死亡率，其分配不顯對稱，而偏斜於老年及嬰兒之方面。又如個人每日之費用，亦常不呈對稱之分配。經常費用，每日大抵相同。然意外之担負，賓客之酬酢，及珍品之選購等等，次數雖少，而所費甚多。此類分配，卽由於諸原因之勢力，無均等實現之機會，致使其分配之形，不能對稱。

21 概率曲線之公式

就二項式開展各項，可以與實際測量之分配相比較，前已略言之矣。但二項式展開各項之分配，其性質為不連續的。若以圖解表明之，則其各次數相連之線為多邊形，非為連續之曲線。且求二項式之展開，計算太繁。是以須另用簡法，使次數分配所成之多邊形，與連續曲線形相切合，以替代二項式分配之法則。而此曲線所函之各個縱線，其和可與測量所得之

各個縱線之和相近似；且任二縱線 y_1,y_2 間所函之面積,可與其相當之量x_1,x_2中間所函之數量數目相等。

求連續曲線之公式，以 $p = q = \frac{1}{2}$ 時爲最簡單；此時曲線之形狀,爲對稱的。其公式爲 $y = y_0 e^{-\frac{x^2}{2\sigma^2}}$。式中之 e 爲常數 2.71828，卽納氏對數之底（the base of the Napierian logarithm)；x 爲在曲線之底線上由平均數至某一點之距離；y爲由 x 點向上引出縱線之高度；σ爲標準差數（求法見前第六章，11，(4) 節)。在該章曾言σ爲在ox線上之單位距離,用以測量各數量差異之程度。而在此公式中之 σ，卽由前法所求之 σ，用以測量 x 之差異；y_0爲曲線下最高之縱線，卽 x 之値發現最多之概然次數。設y_0爲整數,則y爲y_0之分數部分。若x=o,e=1，則 $y=y_0$。曲線在x =o 點之左右爲對稱的；於是平均數，中點，範數，皆重合於 x = o 之點上。y_0 可由公式 $y_0 = \frac{N}{\sigma\sqrt{2\pi}}$ 求得。式中 N 爲總數量，σ 爲標準差數，π 爲常數(=3.1416)。故求常則曲線之完全公式爲 $y = \frac{N}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$（欲知此公式之說明，可參考Yule:Theory of Statistics書中第十五章)。由此公式求出之對稱曲線，爲用甚廣。以下二章，卽就其用途之大要，分別說明之。

第八章

概率曲線之應用

22 概率曲線下之面積

前章21節，曾言連續曲線下之縱線y，係代表x之次數；各y線相加之和，與所測之數量總數相近似。然若以函x之兩縱綫y_1，y_2中間之面積代表其數量之數目亦可，且較爲精確。

如第三十二表，卽用積分法則，將曲綫下面積之全部，按σ之距離，分爲若干部分。設全面積爲整數(=10,000)，而分爲各部分時，用：(1)曲線，(2)底線，(3)中量(x=o)上之縱線，(4)與距中量某點上之縱線爲界限。(1)，(2)，(3)爲各部分面積通用之界限；(4)之求法，以σ爲單位，求中量與某點x之距離。如表中第三欄第一行之.01，卽$\frac{x}{\sigma}$=.01，則x=.01σ；而在中量與.01σ間所函之數量爲40／10,000，或全數量.4／100；在中量與.02σ間所函之數量，爲80／10,000，或全數量.8／100；在中量與.1σ間所函之數量爲398／10,000，或全數量3.98／100。與中量相距5.0σ時，僅有3／10,000,000之數量，在此界限以外。而在中量與1.0σ間之數量爲34.13／100；在中量與2.0σ間之

第三十二表　常則概率曲線下各 $\frac{x}{\sigma}$ 之各部分面積之價值表

假定全面積為10,00）　100＝百分之一

x＝與平均數相差之數　　σ＝標準差數

$\frac{x}{\sigma}$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	△
0,0	0000	0040	0080	0120	0160	0199	0239	0279	0319	0359	39,8
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753	39,4
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141	38,6
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517	37,5
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879	36,0
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224	34,2
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549	32,2
0,7	2580	2611	2642	2673	2704	2734	2764	2794	2823	2852	30,0
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133	27,7
0,9	3159	3186	3212	3238	3264	3290	3315	3340	3365	3389	25,3
1,0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621	22,9
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830	20,5
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015	18,2
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177	15,9
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319	13,8
1,5	4332	4345	4357	4370	4383	4394	4406	4418	4429	4441	11,9
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545	10,1
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633	8,6
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706	7,1
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767	5,9
2,0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817	4,8
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857	3,9
2,2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890	3,1
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916	2,5
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936	2,0
2,5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952	1,5
2,6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964	1,2
2,7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974	1,0
2,8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981	0,7
2,9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986	0,5
3,0	4986,5	4986,9	4987,4	4987,8	4988,2	4988,6	4988,9	4989,3	4989,7	4990,0	0,37
3,1	4990,3	4990,6	4991,0	4991,3	4991,6	4991,8	4992,1	4992,4	4992,6	4992,9	0,27
3,2	4993,129										0,17
3,3	4995,166										0,12
3,4	4996,631										0,10
3,5	4997,674										0,07
3,6	4998,409										0,05
3,7	4998,922										0,03
3,8	4999,277										0,02
3,9	4999,519										
4,0	4999,683										
4,5	4999,966										
5,0	4999,997133										

數量爲47.72／100;在中量與3.0σ間之數量爲$\frac{49.865}{100}$；在此界限以外之數量，僅有$\frac{135}{100,000}$或$\frac{.135}{100}$。因曲線既爲常則的，而中量上下等距離所函之面積相等，故±1σ函有數量$\frac{68.26}{100}$；±3σ函有數量$\frac{99.73}{100}$；僅有$\frac{.27}{100}$之數量在此±3σ界限以外。在第五章中謂標準差數之六倍，大概包含全體數量$\frac{99}{100}$之定律，即本諸此。

前表中記面積部分之數，由.01σ起按每一.01σ記至3.20σ止；再按每一.1σ記3.3σ，3.4σ，3.5σ，至4.0σ止；其次則僅記4.5σ及5.0σ

在前表之最末一欄△下之數，係表中兩鄰欄之大概差數。由此可推算與中量相距.036σ，.057σ等等之值。如求中量與+1.464σ間之數量，按表可查$1.46\sigma=4279$而$13.8\times\frac{.004}{.01}=13.8\times.4=5.52$

$4279+5.52=4285$

23 實際的與理論的次數分配之比較

若實際的次數分配爲常則的形式，其平均數及差異數爲已知，則按前表可求其理論的次數分配，以考究二者切合之程度(參看下節)

24 若次數分配爲常則的可由已知之中量及差異數作次數分配表

例如某一測量之平均數爲24.0，標準差數爲4.0，其次數之分配爲常則的，則按前表可求其次數分配。如第三十三表，由平均數 24 與 25，卽平均數（=24）與 $+.25\sigma(\frac{25-24}{4}=.25\sigma)$中間所函之次數，爲全數百分之 9.87；25 與 26，即 $.25\sigma$ 與 $.5\sigma$ $(\frac{26-24}{4}=.5\sigma)$ 中間所函之次數爲百分之9.28（即$.5\sigma$之次數19.15-9.87 = 9.28）。依此方法，則在平均數上下各差數中間所函之次數，皆可算出。應用此法，可將實際的次數分配，與其理論的次數分配相比較，以考究二者切合之程度。作此種比較之時，先定實際的次數分配之平均

第三十三表 某一測量之理論的次數次分配

平均數=24.0 $\sigma=4.0$

數量	次數	數量	次數
<11	0.06	24－25	9.87
11－12	0.07	25－26	9.28
12－13	0.17	26－27	8.19
13－14	0.32	27－28	6.80
14－15	0.60	28－29	6.30
15－16	1.05	29－30	3.88
16－17	1.73	30－31	2.68
17－18	2.68	31－32	1.73
18－19	3.88	32－33	1.05
19－20	6.30	33－34	0.60
20－21	6.80	34－35	0.32
21－22	8.19	35－36	0.17
22－23	9.28	36－37	0.07
23－24	9.87	>37	0.06

數，卽以此平均數作理論的次數分配之中量，然後標準此中量，求與實際次數相當的理論次數。

25 由已知之任何距離求其中間所含數量之百分數

若一測量之中量及差異數爲已知，則按前第三十二表卽可推算任何距離間之數量。例如某測量之平均數 $=10$，$\sigma=2.4$，求其 12.4 及 12.6 二距離中間所函之數量。12.4與平均數$(=10)$相距爲$1\sigma(=2.4)$，12.6與平均數相距爲 1.0833σ；平均數與 1σ 及 1.08σ 中間所函之百分數量爲 34.14 及 36.00。故 1σ 及 1.08σ 中間之百分數量爲 $36.00-34.14=1.86$；若再求精密，將 $.0033\sigma$所含之數量加入，則將前表中 1.08σ及1.09σ之二面積(3599及3621)之差數$(=22)$與 .0033 相乘，得百分之.07，與1.86相加爲1.93，卽12.4與12.6中間所函之百分數量。欲熟習前第三十二表之用法，可練習下列數題：

7 $M=10$, $\sigma=3$, 7與13間之百分數量若干？

8 $M=22$, $\sigma=4.4$, 18與20間之百分數量若干？

9 $M=15.5$, $\sigma=2.1$, 22以上之百分數量若干？

10 $M=15.5$, $\sigma=2.1$, 13以下之百分數量幾何？

11 M=14.86,σ=3.00, 12與13間之百分數量幾何?

12 M=14.86,σ=3.00, 14與16間之百分數量幾何?

26 就底線上之標準點求函已知之百分數量之距離

反之,若某測量之中量及差異數爲已知,則按前第三十二表,即可求函任何百分數量之距離。例如平均數 = 8.0 , σ=2.0, 求平均數以上函總數四分之一(即百分之二十五)之距離。按前表,函百分之 24.86之距離爲.67σ;函25.17之距離爲.68σ;則函百分 25 數量之距離爲 $.67\sigma+\frac{(25.00-24.86)\times.01\sigma}{25.17-24.86}=.6745\sigma$, $.6745\times2$ =1.35。故由平均數 8.00數至 9.35 (=8.00+1.35), 即爲所求之距離。又如求函總數量百分之80之最近距離時,按常則曲線之形式,數量集合愈密之處,愈與平均數相近。由是,題中函數量百分之80之距離,即在平均數上下各函數量百分之40之距離。按前表,此距離各爲+1.28σ及 −1.28σ;再求精密,爲 +1.2817σ及 −1.2817σ。由已知之平均數及σ求之,8.00±(1.2817×2)=8.00±2.5634,即5.4366及10.5634之距離。若由二極端向中央計算之時,求±3σ之距離即可,不必顧及理論曲線之二極端之價值,因其數爲無窮的也。試

作下列各題:

13　M = 10，σ = 2；a，在平均數上函數量百分之30之距離，其界限若何？b,在平均數下函百分之20之界限若何？c，又函中央三分之二之數量，其界限若何？

14　M = 17.24，σ = 4.0；a，函中央四分之三之數量其界限若何?b，又函最低百分之10之數量，其界限若c，又由頂端算起，函數量六分之一之第二組，其界限若何?何?

27　標準試驗之製作法

(1)　標準試驗之需要　每一問題之內容，各有其難易之程度，學校之考試題目亦然（參看第二章,3,(2)節)。然向來學校舉行試驗時所用之題目,悉以主觀評定其難易程度。既憑主觀之標準,自失比較之根據。以故學生程度之高低，不能有精確之考查。例如某生於中學一年級時,其國文之分數平均爲80；至第三年級時亦得80分。修學期間，增加二年，而其進步之程度,不能藉記分之法則以顯。又如同一學校同一年級之學生,第一級得平均分數70分之學生,其學力恒優於第二級得平均分數90分之學生。故舊制

記分法則，非但個人先後之進步，無從比較；卽與他人之程度比較，亦無確實之根據。再推而言之，甲級與乙級，甲校與乙校，甲省與乙省等等，其成績之優劣，亦皆無比較可言。是以欲考察學生之確實程度，必須就各個問題釐定難易之等級。然後以此種問題爲標準，舉行試驗，則優劣之區別，斯可眞確。其分別問題難易等級之法則見下節。

(2)分別問題難易之法則　編製標準試驗方法甚繁，應加注意之點亦甚多；此節所述，僅爲釐定問題等級之大綱。能應用此法，實際製作，自可從經驗中多闢新理。欲知其詳細作法，可參考M. Call所著之How to Measure in Education第二篇。其製作之步驟，可分下列數條：

(a)將一試驗中各問題，按先易後難之次第，排列其難易等級，由作者之意見或評判人多數之意見決定之。

(b)令各學校各年級之學生，答此試驗中各問題。學生之數目愈多，愈可得常則的次數分配。

(c)評定各問題之答案，或是或非；問題中如有不適宜者，卽將其取消。

(d)計算答對每一問題之學生總數，並以受試驗之學生全數除之，求答對每一問題之學生百分數。

(e)由學生之百分數，求各問題標準差(S.D.)之價值。此層即應用前第三十二表。以該表中全面積之各部分代作對每一問題之學生百分數。以各σ(即ox線之距離)之值，代各問題難易之等級。於是就已知之面積(或學生之百分數)可求相當之σ(或問題難易之地位)。惟三十二表之標準點，係在ox線上之中央(即x=o)。依百分記分法，此點可作為50分。若以此為標準，推算答對每一問題之學生百分數及其相當之距離，必須另行加減，頗不便於應用。故有教育統計學者，將前表改造，如第三十四表，即由三十二表改造而成。以三十二表之零下5σ為標準點，定全距離為10σ。每1σ又分為二十等分。總計分全距離為200部分。然後將各部分之相當面積(或數量)求出，並將每1σ用10乘之，使與通用之百分記分法適合。能求出答對每一問題之學生百分數，即可按表查得該問題難易之地位，亦即可定該問題之分數。例如答對某一問題之學生數占百分之50，則該問題難易之地位，在距離中所占之部分為50，其分數亦定為50。若答某一問題之

第三十四表　表明零點上各百分率之標準差之距離。零點乃在平均數以下5σ，每一標準差用10乘之，以消除其小數。

標準差之價值	百分率	標準差之價值	百分率	標準差之價值	百分率	標準差之價值	百分率
0	99.999971	25	99.38	50	50.00	75	0.62
0.5	99.999963	25.5	99.29	50.5	48.01	75.5	0.54
1	99.999952	26	99.18	51	46.02	76	0.47
1.5	99.999938	26.5	99.06	51.5	44.04	76.5	0.40
2	99.99992	27	98.93	52	42.07	77	0.35
2.5	99.99990	27.5	98.78	52.5	40.13	77.5	0.30
3	99.99987	28	98.61	53	38.21	78	0.26
3.5	99.99983	28.5	98.42	53.5	36.32	78.5	0.22
4	99.99979	29	98.21	54	34.46	79	0.19
4.5	99.99973	29.5	97.98	54.5	32.64	79.5	0.16
5	99.99966	30	97.72	55	30.85	80	0.13
5.5	99.99957	30.5	97.44	55.5	29.12	80.5	0.11
6	99.99946	31	97.13	56	27.43	81	0.097
6.5	99.99932	31.5	96.78	56.5	25.78	81.5	0.082
7	99.99915	32	96.41	57	24.20	82	0.069
7.5	99.9989	32.5	95.99	57.5	22.66	82.5	0.058
8	99.9987	33	95.54	58	21.19	83	0.048
8.5	99.9983	33.5	95.05	58.5	19.77	83.5	0.040
9	99.9979	34	94.52	59	18.41	84	0.034
9.5	99.9974	34.5	93.94	59.5	17.11	84.5	0.028
10	99.9968	35	93.32	60	15.87	85	0.023
10.5	99.9961	35.5	92.65	60.5	14.69	85.5	0.019
11	99.9952	36	91.92	61	13.57	86	0.016
11.5	99.9941	36.5	91.15	61.5	12.51	86.5	0.013
12	99.9928	37	90.32	62	11.51	87	0.011
12.5	99.9812	37.5	89.44	62.5	10.56	87.5	0.009
13	99.989	38	88.49	63	9.68	88	0.007
13.5	99.987	38.5	87.49	63.5	8.85	88.5	0.0059
14	99.984	39	86.43	64	8.08	89	0.0048
14.5	99.971	39.5	85.31	64.5	7.35	89.5	0.0039
15	99.977	40	84.13	65	6.68	90	0.0032
15.5	99.972	40.5	82.89	65.5	6.06	90.5	0.0026
16	99.966	41	81.59	66	5.48	91	0.0021
16.5	99.960	41.5	80.23	66.5	4.95	91.5	0.0017
17	99.952	42	78.81	67	4.46	92	0.0013
17.5	99.942	42.5	77.34	67.5	4.01	92.5	0.0011
18	99.931	43	75.80	68	3.59	93	0.0009
18.5	99.918	43.5	74.22	68.5	3.22	93.5	0.0007
19	99.903	44	72.57	69	2.87	94	0.0005
19.5	99.886	44.5	70.88	69.5	2.56	94.5	0.00043
20	99.865	45	69.15	70	2.28	95	0.00034
20.5	99.84	45.5	67.36	70.5	2.02	95.5	0.00027
21	99.81	46	65.54	71	1.79	96	0.00021
21.5	99.78	46.5	63.68	71.5	1.58	96.5	0.00017
22	69.74	47	61.79	72	1.39	97	0.00013
22.5	99.70	47.5	59.87	72.5	1.22	97.5	0.00010
23	99.65	48	57.93	73	1.07	98	0.00008
23.5	99.60	48.5	55.96	73.5	0.94	98.5	0.000062
24	99.53	49	53.98	74	0.82	99	0.000048
24.5	99.46	49.5	51.99	74.5	0.71	99.5	0.000037
						100	0.000029

學生數目爲百分之99.38，在該問題之地位及分數爲25。若某問題僅有百分之.000029人不能答對，則其題最易，卽定其價値爲0。若僅有百分之.000029人能答對之問題，則其題爲最難，卽定其價値爲100。此10σ距離之上下，尙有無限之百分數，因常則曲線兩端之數，爲無窮的也。故爲應用方面計，可將此距離斷自中點上下之5σ；雖此距離以外，尙有數目，惟其數甚微，棄之亦不關緊要。改造此表，亦有採中點上下3σ及2.5σ之距離者，然與三十四表相較，既見簡略；且其價値亦各不同。例如以5σ爲全距離，分爲一百等分時，則25分當1.25σ之距離，以6σ爲全距離分爲一百等分時，則25分當1.5σ之距離，與以10σ爲全距離分爲一百等分，其25分當2.5σ之距離，各不相同。故若各種試驗所採用之標準不同，雖其中記分之數字偶有符合，亦不可卽認其有相等之價値，

(f)既求得各問題確實難易之等級，卽按其等級依次排列，再依法試驗之，以定其最終之等級。蓋各問題之次第先後更換之時，其難易之程度，往往亦受影響。恆有某一問題之意義，能引起答後一問題之思想者。故當各問題次序變更之時，欲使其難易之程度精密，必須作最終之訂正也。

第九章

證確數量（Reliability measures）

28 應用常則曲線斷定統計結果之確度

測量任何事實，必不能得其事實之全體數量；所得者，僅其全體數量中之一部分耳。而由此一部分數量所求出之集中數量，苟非出於偶然，斷不能爲眞確之中量。眞確中量，必須由某一事實之全體數量求之。而由一部分有限之數量求得之中量，爲實得中量。譬如測量某一學生之國文能力，測量20次時，得平均數75分；若再多測二次，將所得之二數量與前之20個數量相加，以求其平均數，則此平均數必難仍爲75分。若再加二數量時，其平均數之價值，又將變動。若欲得此學生之國文能力眞確平均數，必須就其萬有之數量求之。測量一羣學生之國文能力亦然。人數有增減，則所得之平均數即有變動。其眞確之平均數，必須就此同類學生之全數求之。故實得中量，除非偶然可與眞確中量相等，僅能爲其近似之值。此專就集中數量言之也。而次數分配，差異數量，關係數量，及差數數量（measures of difference）等，亦有實得與眞確之分。實得之數量，與眞確之數量，非出偶

然，不能有相等之數值。

實得數量旣與眞確數量有差，則必須求其相差之度，以證明實得數量變動之大小，然後可知眞確數量發現之界限。此證明法則即應用前概率曲線之原理。

假定在某一城中有學生二萬，欲知其拼音能力之平均數，將此二萬學生共分爲一百組，每組之人數爲二百名。此全體學生之次數分配，與常則概率曲線相合。而所分之各組學生，乃盡量用機會之方法選擇者；其次數分配，亦皆與常則曲線相合。再假定將此一百組之學生，一一試驗，並求得各組之平均數，及各組平均數之平均數。此各組平均數之平均數，即此二萬學生之眞確平均數。惟各組之一百個平均數，不能皆彼此相等，亦不能皆與眞確平均數相等；其中有與眞確平均數相差甚多者，有與之相差甚少者，亦有與之相等者。若將各組之平均數與眞確平均數相較，以求其差數；然後將各個差數排列一次數分配表，則此一百個差數分配之形式，亦與常則之曲線相合，而誤差之零點正當眞確之平均數。設所求出各組之平均數，及各組平均數與眞確平均數之差數，如第三十五表所列：

第三十五表　二萬學生拼音能力各平均數之分配

（此係假定的）

各組平均數之分類	中值	誤差	次數（即組數之百分數）
72.1－72.3	72.2	－1.0	1
72.3－72.5	72.4	－.8	2
72.5－72.7	72.6	－.6	3
72.7－72.9	72.8	－.4	10
72.9－73.1	73.0	－.2	21
73.1－73.3	73.2	0	25
73.3－73.5	73.4	.2	21
73.5－73.7	73.6	.4	11
73.7－73.9	73.8	.6	3
73.9－74.1	74.0	.8	2
74.1－74.3	74.2	1.0	1

就前表所列，可以推知實得平均數與眞確平均數相差之限度。此一百個數量中有 67 個數量列於 73.5 及72.9 以內，有33個數量列於此界限以外；有94個數量列於73.9及72.5之價値以內。換言之，即眞確平均數與實得平均數相差不出73.5及72.9之界限時，一百個數量中，約有67個數量；相差不出73.9及 72.5 之界限時，一百個數量中，約有94個數量。以機會言，概爲2與1(67:33)之比，眞確平均數與實得平均數相差不能出於73.5及72.9之界限以外；16 與 1 (94:6) 之比眞確平均數與實得平均數相差不能出於 73.9及 72.5 界限以外。再換言之，即眞確平均數列於73.5 及 72.9 中間之機會，概爲 2 與 1 之比；列於 73.9 及 72.5中間

之機會，概爲16與1之比。而眞確平均數無發現之機會，按諸前表，亦可確定；卽在74.3及72.1之界限以外，無其發現之機會矣。

由是，若眞確數量與實得數量之差數，其次數分配爲常則的，則將此差數之差異數量求出，卽可證明包含眞確數量之限度。若代表全體之實例愈多，則眞確數量愈可與實得數量相近，卽二者之差度愈小。若各個實例彼此相差之度愈微，則眞確數量與實得數量之差異數量亦愈小。故實得數量確實程度之增加，卽在所測之數量數目（N）增加，或各數量之差異程度（σ）減少；由下29節各證誤公式可以見之。

求實得數量與眞確數量差度之差異數量有二：(1)標準差，(2)概誤差。分述之於29及30兩節。

29 用標準差求證誤數(unreliability)之方法

(1)求平均數之證誤數之公式爲：$\sigma_M=\frac{\sigma_{dis.}}{\sqrt{N}}$

式中之N＝所測之數量數目

$\sigma_{dis.}$＝各數量之標準差數

σ_M＝實得平均數與眞確平均數之差度之標準差

試用下例，以說明此公式之用法及意義：

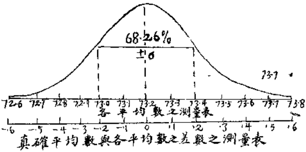
設$N=900$，$\sigma_{dis.}=6$，$M=73.2$ 則$\sigma_M=\frac{6}{\sqrt{900}}=.2$

前例平均數之證誤數 σ_M 爲.2，其意義可用下列之方法說明之。

習慣所用，常以$\pm3\sigma$爲確實限度。故平均數之確實限度爲$\pm3\sigma_M$；中點數之確實限度爲$\pm3\sigma_{Md}$；標準差數之確實限度爲$\pm3\sigma_\sigma$；關係數之確實限度爲$\pm3\sigma_r$等等。

就前例平均數之證誤數.2，可以確定眞確平均數乃在$73.2\pm3(.2)$，卽73.8及72.6之間也。但在73.8及72.6間之眞確平均數，並非此$N=900$之平均數，因其平均數已求出爲73.2；是乃爲此900數量所附屬之較大團體之眞確平均數。假使此900數量之選法，最爲精當，則其平均數73.2可與其較大團體之眞確平均數適合。然按之實際，極難得此美滿之結果。於是就此已得之平均數，定此較大團體之眞確平均數，在73.8及72.6中間某點。雖73.8之上及72.6之下，尙有眞確數量發現之機會（因常則曲線之兩端爲無限的），然其機會極少，10,000次中僅能有3次而已。而在$73.2\pm1(.2)$，$73.2\pm2(.2)$，及$73.2\pm3(.2)$中間發現之機會，與前第三十五表所舉之例大概相同；卽眞確

數量在73.2±1(.2)中間之機會爲2.15比1(68.26:31.74),在73.2±2(.2)中間之機會爲21比1(95.44:4.56),在73.2±3(.2)中間之機會爲369比1(9972:27)。參看第二十七圖自明:



第二十七圖　眞確平均數與各組平均數之差數之常則的次數分配

上述之解釋,可以適用於求平均數之謬誤數量;亦卽可以適用於任何謬誤數量。

(2)求中點數之謬誤數之公式爲: $\sigma_{Md}=\frac{1.25331\sigma_{dis.}}{\sqrt{N}}$

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

由此式可以確定眞確中點數,乃在實得中點數±3σ_{Md}之間。

(3)求Q之證誤數之公式為：$\sigma_Q=\frac{1.11\sigma_{dis.}}{\sqrt{2N}}$

由此式可以確定眞確之四分差數，乃在實得四分差數±3 σ_Q 之間。

(4)求標準差數之證誤數之公式為：$\sigma_\sigma=\frac{\sigma_{dis.}}{\sqrt{2N}}$

由此式可以確定眞確之標準差數，乃在實得標準差數±$3\sigma_\sigma$ 之間。

(5)求r之證誤數之公式為：

(a)由 $\frac{\Sigma xy}{N\sigma_x\sigma_y}$ 求得之r，則用證誤公式$\sigma_r=\frac{1-r^2}{\sqrt{N}}$

(b)由$2\sin\left(\frac{\pi}{6}\rho\right)$求得之$r$，則用證誤公式$\sigma_r=\frac{1.05(1-r^2)}{\sqrt{N}}$

(c)由$\cos\pi U$求得之r，則用證誤公式$\sigma_r=\frac{1.63}{\sqrt{N}}$

(d)由$2\cos\frac{\pi}{3}(1-R)-1$求得之$r$，則用證誤公式$\sigma_r=\frac{.638}{\sqrt{N}}$

(6)差數之證誤數　此證誤數，亦為證誤數量中最有用者之一。若當實驗教法時，有實驗式與限制式兩班，此數量之用途尤為緊要。因由此二班之平均數，中點數，或他種數量，即可判斷所實驗之結果如何。

此種差數之證誤數，可以確定斷案之價值。向來此類斷案，često用下列之公式求出之；但其方法，稍嫌陳

舊：$\sigma=\sqrt{(\sigma 數量\mathrm{I})^2+(\sigma 數量\mathrm{II})^2}$

假設實驗式之第一班25名學生，用新法教授，其進步之平均數為18，進步之標準差數為4；而限制式之第二班36名學生，其進步平均數為16，進步之標準差數為3；則二者平均之進步數之差數為 18－16＝2。此差數是否確實，是否可以斷定假使再就此兩班學生繼續實驗時，此第一班優勝之差數，不致為零，或反為第二班優勝之差數。在求差數之證誤數以先，須求與此差數有關各數量之證誤數。其算法如下：

$$\sigma_{\mathrm{MI}}=\frac{4}{\sqrt{25}}=.8 \qquad \sigma_{\mathrm{MII}}=\frac{3}{\sqrt{36}}=.5$$

$$\sigma_{\mathrm{d}}=\sqrt{(.8)^2+(.5)^2}=.94+$$

式中之σd＝差數之標準差

由此可以確定此二進步平均數之真確差數，乃在實得差數2±3(.94)之間；即在－.82與4.82之間。可見真確差數亦有為零或在零下之機會。若真確差數在零之下，則將使此種實驗之效果，有利於第二組之可能。雖然，此真確差數在零上而利於第一組及新式教法之機會甚大。至定此機會較大之程度究竟如何

，詳見後列方法。

任何二個數量，皆可求其差數之證誤數；惟最常計算者，爲平均數之差數。故此公式 σ_d 亦可以適用於 σ, γ 等等各數之差數。求平均數之 σ_d 時，先用求平均數證誤數之公式，求 σ_{MI} 及 σ_{MII}；然後將此二數代入求 σ_d 之公式中；而求 Q，σ，γ 等數之 σ_d 亦然，須先用求 γ 及Q等證誤數之公式，求出 $\sigma_{\gamma I}$ 及 $\sigma_{\gamma II}$ 或 σ_{QI} 及 σ_{QII}，然後代入求差數證誤數之公式。

麥柯 (McCall)氏因差數之證誤數，在實驗方面爲用最多；又因本機會之法則，甚難思索其意義；於是創一實驗係數。此實驗係數甚易求出，且可自然的表示任何差數謬誤之程度。在前節求出差數 2 之證誤數爲 .94，其實驗係數卽可由下式求出：

實驗係數 $= \dfrac{\text{差數}}{2.78\sigma_d}$ 因差數爲2，σ_d 爲.94，

故實驗係數 $= \dfrac{2}{2.78 \times .94} = .76$

此實驗差數2，僅有其應有之數.76倍。由是可以確定新法教授，實在較限制式之方法爲優。設使差數非爲2，而爲2.61時，則其實驗係數如下：

$$\text{實驗係數} = \frac{2.61}{2.78 \times .94} = 1.0$$

實驗係數爲1.0時，乃表明適合之確定程度；實驗係

數爲.5時，表明$\frac{1}{2}$之確定程度；實驗係數爲2.0時，表明二倍之確定程度等等。

若欲根據機會法則，以說明眞確差數爲零或小於零，即眞確差數利於相反之斷言時，所求出實驗係數，即可由下表直接得之。下表之數，非但可以適用於二平均數之差數；且可適用於任何二數量之差數。

第三十六表 用機會法則說明實驗係數表

實驗係數	相近之機會
.1	1.6 比 1
.2	2.5 „ 1
.3	3.9 „ 1
.4	6.5 „ 1
.5	11 „ 1
.6	20 „ 1
.7	38 „ 1
.8	75 „ 1
.9	160 „ 1
1.0	369 „ 1
1.1	930 „ 1
1.2	2350 „ 1
1.3	6700 „ 1
1.4	20000 „ 1
1.5	67000 „ 1

30 化σ數量爲P.E.數量之方法

在前第五章11節曾述及當次數分配對稱之時，則P.E.，Q.皆相等。而P.E.等於.6745σ。習慣上多用P.E.表明證誤數，而用σ表明之者較少。P.E.數之求法，用.6745乘σ數量即可。試用上列公式之二，說明

原书缺页

問　　題

15 求下表中各數量之證誤數量

分數	次數	求得之數
2－3	1	M＝7.0
3－4	1	Md＝7.0
4－5	2	Q＝1.4
5－6	4	S.D.＝2.22
6－7	4	γ ＝ .303(用乘積率法求出)
7－8	5	
8－9	3	
9－10	2	
10－11	1	
11－12	0	
12－13	1	
N	24	

第三十七表　由 ρ 之價值求 γ

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

ρ	γ	ρ	γ	ρ	γ	ρ	γ
.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	.0833	.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	.4056	.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	.7368	.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

第三十八表 由R之價値求γ

$$\gamma = 2\,\mathrm{cosine}\frac{\pi}{3}(1-R)-1,\ R=1-\frac{6\Sigma g}{N^2-1}$$

R	γ	R	γ	R	γ	R	γ
.00	.030	.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	.656
.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	.806	.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	.96	.9999
.24	.399	.50	.732	.75	.932	1,00	1.0000
.25	.414						

第三十九表　由U之百分比例數求γ

$\gamma = \text{cosine}\,\pi U$

U爲異號差數對數與對數總數之百分比例數。若以U爲1之百分比例數，則γ之值爲負數。是以當U過.50時，可按1百分比例數，求γ之負數。

U	γ	U	γ	U	γ	U	γ
.00	1.0000	.13	.9174	.26	.6848	.38	.3682
.01	.9996	.14	.9044	.27	.6615	.39	.3387
.02	.9982	.15	.8905	.28	.6375	.40	.3089
.03	.9958	.16	.8757	.29	.6129	.41	.2788
.04	.9924	.17	.8602	.30	.5877	.42	.2485
.05	.9880	.18	.8439	.31	.5620	.43	.2180
.06	.9826	.19	.8268	.32	.5358	.44	.1873
.07	.9762	.20	.8089	.33	.5091	.45	.1564
.08	.9688	.21	.7902	.34	.4819	.46	.1253
.09	.9604	.22	.7707	.35	.4542	.47	.0941
.10	.9510	.23	.7501	.36	.4260	.48	.0628
.11	.9407	.24	.7293	.37	.3973	.49	.0314
.12	.9295	.25	.7074			.50	.0000

第四十表 常則槪率曲線下各σ之各部分面積之價值表。

假定全面積爲10,000 100=百分之一

x=與平均數相差之數 σ=標準差數

$\frac{x}{\sigma}$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	△
0,0	0000	0040	0080	0120	0160	0199	0239	0279	0319	0359	39,8
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753	39,4
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141	38,6
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517	37,5
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879	36,0
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224	34,2
0,6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549	32,2
0,7	2580	2611	2642	2673	2704	2734	2764	2794	2823	2852	30,0
0,8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133	27,7
0,9	3159	3186	3212	3238	3264	3290	3315	3340	3365	3389	25,3
1,0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621	22,9
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830	20,5
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015	18,2
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177	15,9
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319	13,8
1,5	4332	4345	4357	4370	4383	4394	4406	4418	4429	4441	11,9
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545	10,1
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633	8,6
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706	7,1
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767	5,9
2,0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817	4,8
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857	3,9
2,2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890	3,1
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916	2,5
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936	2,0
2,5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952	1,5
2,6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964	1,2
2,7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974	1,0
2,8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981	0,7
2,9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986	0,5
3,0	4986,5	4986,9	4987,4	4987,8	4988,2	4988,6	4988,9	4989,3	4989,7	4990,0	0,37
3,1	4990,3	4990,6	4991,0	4991,3	4991,6	4991,8	4992,1	4992,4	4992,6	4992,9	0,27
3,2	4993,129										
3,3	4995,166										0,17
3,4	4996,631										0,12
3,5	4997,674										0,10
3,6	4998,400										0,07
3,7	4998,922										0,05
3,8	4999,277										0,03
3,9	4999,519										0,02
4,0	4999,683										
4,5	4999,966										
5,0	4999,997133										

第四十一表　表明零點上各百分率之標準差之距離。零點乃在平均數以下5σ，每一標準差用10乘之，以消除其小數。

標準差之價值	百分率	標準差之價值	百分率	標準差之價值	百分率	標準差之價值	百分率
0	99.999971	25	99,38	50	50,00	75	0,62
0.5	99.999963	25,5	99,29	50,5	48,01	75,5	0,54
1	99.999952	26	99,18	51	46,02	76	0,47
1,5	99.999938	26,5	99,06	51,5	44,04	76,5	0,40
2	99.99992	27	98,93	52	42,07	77	0,35
2,5	99.99990	27,5	98,78	52,5	40,13	77,5	0,30
3	99.99987	28	98,61	53	38,21	78	0,26
3,5	99.99983	28,5	98,42	53,5	36,32	78,5	0,22
4	99.99979	29	98,21	54	34,46	79	0,19
4.5	99.99973	29,5	97,98	54,5	32,64	79,5	0,16
5	99.99966	30	97,72	55	30,85	80	0,13
5,5	99.99957	30,5	97,44	55,5	29,12	80,5	0,11
6	99.99946	31	97,13	56	27,43	81	0,097
6,5	99.99932	31,5	96,78	56,5	25,78	81,5	0,082
7	99.99915	32	97,41	57	24,20	82	0,069
7.5	99.9989	32,5	95,99	57,5	22,66	82,5	0,058
8	99.9987	33	95,54	58	21,19	83	0,048
8.5	99.9983	33,5	95,05	58,5	19,77	83,5	0,040
9	99.9979	34	94,52	59	18,41	84	0,034
9,5	99.9974	34,5	93,94	59,5	17,11	84,5	0,028
10	99.9968	35	93,32	60	15,87	85	0,023
10,5	99.9961	35,5	92,65	60,5	14,69	85,5	0,019
11	99.9952	36	91,92	61	13,57	86	0,016
11,5	99.9941	36,5	91,15	61,5	12,51	86,5	0,013
12	99.9928	37	90,32	62	11,51	87	0,011
12,5	99.9912	37,5	89,44	62,5	10,56	87,5	0,009
13	99.989	38	88,49	63	9,68	88	0,007
13,5	99.987	38,5	87,49	63,5	8,85	88,5	0,0059
14	99.981	39	86,43	64	8,08	89	0,0048
14,5	99.971	39,5	85,31	64,5	7,35	89,5	0,0039
15	99.977	40	84,13	65	6,68	90	0,0032
15,5	99.972	40,5	82,89	65,5	6,06	90,5	0,0026
16	99.966	41	81,59	66	5,48	91	0,0021
16,5	99.960	41,5	80,23	66,5	4,95	91,5	0,0017
17	99.952	42	78,81	67	4,46	92	0,0013
17,5	99.942	42,5	77,34	67,5	4,01	92,5	0,0011
18	99.931	43	75,80	68	3,59	93	0,0009
18,5	99.918	43,5	74,22	68,5	3,22	93,5	0,0007
19	99.903	44	72,57	69	2,87	94	0,0005
19,5	99.886	44,5	70,88	69,5	2,56	94,5	0,00043
20	99.865	45	69,15	70	2,28	95	0,00034
20,5	99.84	45,5	67,36	70,5	2,02	95,5	0,00027
21	99.81	46	65,54	71	1,79	96	0,00021
21,5	99.78	46,5	63,68	71,5	1,58	96,5	0,00017
22	99.74	47	61,79	72	1,39	97	0,00013
22,5	99.70	47,5	59,87	72,5	1,22	97,5	0,00010
23	99.65	48	57,93	73	1,07	98	0,00008
23,5	99.60	48,5	55,96	73,5	0,94	98,5	0,000062
24	99.53	49	53,98	74	0,82	99	0,000048
24,5	99.46	49,5	51,99	74,5	0,71	99,5	0,000037
						100	0,000023

附錄II 本書所用之符號及公式

1 符　號

A. Md = assumed Median 假定中點數

b_1 代表 x 依 y 爲消長之係數

b_2 代表 y 依 x 爲消長之係數

C = correction 更正數

d = deviation 差數

E.M. = estimated mean 估計平均數

f = frequency of measures 數量次數

i = class- interval 組距

M = mean 平均數

m = measure 數量

Md = median 中點數

M.D. = mean deviation 平均差數

Mo. = mode 範數

N = 數量總數

O.M. = obtained mean 實得平均數

P_p = 百分點之價値

P.E. = probable error 概誤差

Q_1 = 下百分點

Q_3 = 上百分點

Q = quartile deviation 四分差數

γ = coefficient of correlation 關係係數

S.D. = standard deviation 標準差數，亦用 σ = sigma 代之。

σ_{dis} = 各數量之標準差

σ_M = 實得平均數與真確平均數之差度之標準差

σ_{Md} = 實得中點數與真確中點數之差度之標準差

σ_Q = 實得四分差數與真確四分差數之差度之標準差

σ_σ = 實得標準差數與真確標準差數之差度之標準差

σ_γ = 實得關係係數與真確關係係數之差度之標準差

σ_d = 差數之標準差

V = 差異係數；或含有中點數組低，或高界限之價值

X = A 列數量中之任一數量

x = A列數量中之任一數量與其平均數之差數

x' = A列數量中之任一數量與其估計平均數之差數

Y = B列數量中之任一數量

y = B列數量中之任一數量與其平均數之差數

y' = B列數量中之任一數量與其估計平均數之差數

Σ = 相加之符號

2　公　　式

頁數

(1)求近似範數之公式：$Mo. = M. - 3(M. - Md.)$ ……………………38

(2)求平均數之公式：$M. = \frac{\Sigma fm}{N}$，其簡法之公式為：

$M. = E.M. + C$，或 $M = E.M. + \frac{\Sigma fd}{N} i$ ……… 41

(3)求中點數之公式：$Md = V \pm \frac{\frac{N}{2} - F}{f} i$ ……………………47-48

(4)求百分點之公式：$P_p = V_p + \frac{\frac{p}{100}N - F_p}{f_p} i_p$ ………………48

(5)求四分差數之公式：$Q = \frac{Q_3 - Q_1}{2}$ ………………………………51

(6)用簡法求平均差數之公式：$M.D. = \frac{\Sigma fd + C(N_b - N_a)}{N}$，

$C = \frac{O.Md - A.Md}{i}$ ………………58

(7)用簡法求標準差數之公式：$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - C^2} \times i$ ……………62

(8)求單獨生差異係數之公式：$V = 100\frac{\sigma}{M.}$，或$100\frac{M.D.}{Md.}$ ……… 65

(9)求偏斜度之公式：偏斜之度$= \frac{3(\text{mean} - \text{median})}{\sigma}$，

或，偏斜之度$= \frac{Q_1 + Q_3 - 2Md}{Q}$ ………………67

(10)求消長係數之公式：$b_1 = r\frac{\sigma_x}{\sigma_y}$，$b_2 = r\frac{\sigma_y}{\sigma_x}$ …………… 73,87

(11)用乘積率法求關係係數之公式：$r = \frac{\Sigma xy}{N\sigma_x\sigma_y}$，

或$= \frac{\Sigma xy}{\sqrt{\Sigma x^2 \cdot \Sigma y^2}}$ ……… 74-75

其簡法之公式：$r = \frac{\Sigma x'y' - NC_xC_y}{N\sigma'_x\sigma'_y} = \frac{\frac{\Sigma x'y'}{N} - C_xC_y}{\sigma_x\sigma_y}$ ………87

(12)用等級差異法求r之公式：$r = 2\sin\left(\frac{\pi}{6}\rho\right)$，$\rho = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$ 88

或$r = 2\text{cosine}\frac{\pi}{3}(1-R) - 1$，$R = 1 - \frac{6\Sigma g}{N^2-1}$ 91

$\sigma_\gamma = \frac{1.05(1-\gamma^2)}{\sqrt{N}}$，$\gamma$由$2\sin\left(\frac{\pi}{6}\rho\right)$求得，

$\sigma_\gamma = \frac{.638}{\sqrt{N}}$，$\gamma$由$2\cos\frac{\pi}{3}(1-R)-1$求得；

$\sigma_\gamma = \frac{1.63}{\sqrt{N}}$，由cos[illegible]求得···31

(23)求實驗係數之公式：實驗係數 $= \frac{\text{差數}}{2.78\,\sigma_d}$ ·····················133

(24)化σ為P.E.之公式：P.E. $= .6745\sigma$，用此式中之.6745乘以[illegible]

由(18)至22各公式之右方，即得求各P.E.之公式··134-153

附錄III 問題解答

1

a,

數量 m	次數 f
35 - 39.99	1
40 - 44.99	3
45 - 49.99	5
50 - 54.99	7
55 - 59.99	10
60 - 64.99	9
65 - 69.99	5
70 - 74.99	8
75 - 79.99	11
80 - 84.99	1
85 - 90	2
總數	62

b,

數量 m	次數 f
50 - 54.9	3
55 - 59.9	1
60 - 64.9	5
65 - 69.9	5
70 - 74.9	6
75 - 79.9	15
80 - 84.9	1
85 - 89.9	7
90 - 94.9	6
95 - 100	13
總數	62

2

a, $Q_1 = 54.6$
$Md = 62.8$
$Q_3 = 71.06$
$Mo. = 77.5$

b, $Q_1 = 71.25$
$Md = 78.66$
$Q_3 = 92.9$
$Mo. = 77.5$

3

a, $Q_1 = 56$
$Md = 60.86$
$Q_3 = 65.14$
$Mo. = 61$，及65

b, $Q_1 = 58$
$Md = 66.29$
$Q_3 = 71.5$
$Mo. = 67$

4　問題2：a之Q＝9.73
a之S.D.＝11.95
b之Q＝10.83
b之S.D.＝13.25

問題3：a之Q＝4.57
a之S.D.＝6.8
b之Q＝6.75
b之S.D.＝10.36

5　A＝19.55
B＝15.31
C＝11.28
D＝15.79

6　$\gamma = .77$
$b_1 = .71$，即 $x = .71y$
$b_2 = .83$，即 $y = .83x$

7　百分之68.3

8　百分之14.3

9　百分之 0.1

10　百分之11.7

11　百分之 9.7

12　百分之26.1

13　a，10及11.68
b，10及 8.95
c，8.07及11.93

14　a，12.64及21.84
b，12.11及分配之低界限
c，18.96及21.11

15　$\sigma_M = .45$，$\sigma_{Md} = .57$，$\sigma_Q = .36$
$\sigma_\sigma = .32$，$\sigma_\gamma = .13$

附錄IV 勘誤表

頁數	行數	字數	誤字	正字
4	1	9	事	理
8	第二表右方		標準差數=2.024	標準差數=2.026
12	6	末	Probabl	Probable
33			(上圖下之字)b	a
,,			(下圖下之字)a	b
42	四末	20	佔	估
,,	末三	24	直	值
51	8	末	中點之	中點數之
87	8	末	$\Sigma x'y' \quad \Sigma C_x C_y$	$=\Sigma x'y' - \Sigma C_x C_y$
88	末七		$\gamma=2\sin\left(\frac{\pi}{6}\rho\right)$	$\gamma=2\sin\left(\frac{\pi}{6}\rho\right)$
120	8	1	——	(添一)何?
,,	9	5	何?	(去掉)
124	13	10	25σ	2.5σ
130	3	末		27)