教育講話叢書
陳選善主編

教育統計學

沈有乾著

世界書局印行

教育講話叢書

陳選善主編

教育統計學講話

沈有乾編著

世界書局印行

中華民國三十五年七月再版

教育講話叢書

教育統計學講話

實價國幣
外加運費匯費

編著者 沈有乾
主編者 陳選善

發行人 李煜瀛
出版者 世界書局
發行所 世界書局

版權所有
不准翻印

教育講話叢書編輯凡例

一　本叢書就教育各部門，約請專家，分別編著，定名爲教育講話叢書。

二　本叢書編輯主旨在就教育各部門，給予讀者一種綜合的鳥瞰，以爲專門研究之基礎。

三　本叢書內容着重在歷史背境之敘述，發展趨向之指示，各家學說各派主張之介紹，參考材料之提供，研究方法之指導，俾啓發讀者思想，引起讀者興趣。

四　本叢書文字力求流利生動，深入淺出，俾便閱讀而利了解。

五　本叢書讀者以師範生，大學教育系學生，中小學教師，一般從事教育工作人員及對於教育有興趣者爲對象。

教育統計學講話序

大概一種學術或一種事業愈進步則愈不能不利用數字，統計學可以幫助我們了解數字與處理數字，所以在現代生活上一天比一天重要起來。曾有外國作家譏評中國人，說我們『不肯數』，所以『不足數』本書寫作的動機卽是希望中國教育界人士漸漸能數，肯數，一洗『不足數』的恥辱！

本書所講，關於計算的方法與公式比較簡略，這樣可以多留些篇幅討論統計數字與術語的意義。而且一部份讀者或者但求了解人家的統計報告，並不希望自己做任何統計工作，關於計算的詳細討論對於他們非但並無用處，甚且會減少興趣。至於有志學習計算的讀者，如能細閱本書的簡要說明，固不難無師自通，但這種讀者的大部份可以在學校得教師面授，也不全賴書本上的說明。

本書限於篇幅，當然未能完備。但有一部份材料卻是多數教育統計書本所尙未提到過的：如平均數與標準差的簡捷計算法，t分配與均方差比率分配，以及方差分析法，都是值得推廣應用的。

本書原稿雖經一再修改，不明白與不正確之處在所難免，希望讀者不吝指教！

三十三年十月　沈有乾

教育統計學講話目錄

第一講　教育與統計數字

中國傳統教育向以文字爲主，對於數字的意義與運用太不注意。但數字有其尊嚴，絕不容人輕忽或侮慢。假設有人不記收支而經營商業，必至蝕本破產。又如有人誤算壓力而建築工事，必至傾毀傷人。那就是獲罪於數字的殘酷報復。教育事業的成敗當然也和統計的精疏有密切的關係。但教育家的責任每不如工商家的責任明顯確定，所以教育家迄未能如工商家尊重數字，良可慨歎！

教育家輕忽數字，濫引數字，誤解數字的例子實在太多了！現在我們提出幾件來討論，並非故意攻訐個人，實在因爲這種不好的風氣再難坐視不予矯正。

照民國十九年教育部統計，全國中小學學生人數（萬以下不計）如下：

初小	九百一十四萬
高小	一百三十九萬
初中	四十萬
高中	十萬

現在假定有人根據以上統計數字，發表下面的結論：

「學生能由初小入高小者不過七分之一而強，能由高小入初中者三分之一而弱，能由初中入高

中者又僅四分之一。……普通觀念，以小學中學分段，事實上初小與高小之間，初中與高中之間，更爲重要階級。因由初小升入高小，及由初中升入高中者，反較高小升入初中……者之數目爲低。」

讀者如不能在前段推論中發現任何破綻，也不足怪。因爲原文是二十三年八月十九日申報上所載（曾轉載於教育雜誌第二十四卷第二號）教育界名流聯名建議修正教育制度的理由之一部份。這議案曾引起全國教育家的注意，贊成和反對的議論在各種雜誌報紙上發表極多，有些刊物特徵求專家意見，發行專號，例如教育雜誌第二十五卷第一號裏就有三十多位教育家討論這問題。但這麼多篇文章裏始終並無一處對於前面所引推論提出任何疑問，也始終並無隻字涉及那些統計數字！

這麼重要的一個建議，一定經過一番很精密的研究，這麼重要的一篇文章，一定經過一番很鄭重的審查，而且聯名提案的是幾位全國聞名的大人物，發表意見的有三十多位教育家，其所引爲理由的事實，應當不至於有任何問題。但普通觀念確以小學中學分段，照我們直接間接所得印像，學生由小學升入中學，似乎比由初小升入高小，或由初中升入高中更加困難。這種零星經驗當然是不可靠的，絕不可與全國教育統計相提並論。不過從教育統計推論所得的結果既與普通印像不符，我們對於這推論的結果似乎未便輕予接受。

高小學生一百三十九萬，比初小學生九百十四萬，確不過七分之一而強。初中學生四十萬，比高小學生一百三十九萬，確是三分之一而弱。高中學生十萬，比初中學生四十萬，確僅四分之一。但讀者如略加反

省，不難想到初小有四年，高小僅二年，初高中各三年，而不論中小學高年級的學生普通總不會比低年級多。所以：

初小四年級生最多不過　二百二十八萬五千
高小一年級生最少也有　六十九萬五千
高小二年級生最多不過　六十九萬五千
初中一年級生最少也有　十三萬三千
初中三年級生最多不過　十三萬三千
高中一年級生最少也有　三萬三千

因此：

學生能由初小入高小者約佔百分之三十
　能由高小入初中者約佔百分之一十九
　能由初中入高中者約佔百分之二十五

可見由初小升入高小，或由初中升入高中者，終究比由高小升入初中者的百分比高。普通印像還與事實符合。

一經指出之後，那篇提案裏的差誤是非常簡單而明顯的。但起稿的那位先生竟會以四年的初小，與

二年的高小，三年的初中或高中直接比較，其餘聯名提案的幾位先生也未察覺，至於參加討論的各位，恐怕根本對於統計數字未加注意。如果那段推論是修正教育制度的建議根據，那提案的價值也就可想而知了。教育家輕忽數字竟到這樣程度，眞是一大恥辱。

從前面所舉例子，可見極簡單的統計數字尙且會被人誤用，稍爲曲折一些的材料，當然更不容易得到正確的解釋了。教師們接觸最多的數字大概要推學生的成績分數，可是多少教師對於自己或別人所批的分數能夠不作歪曲的推論，或不接受別人所作歪曲的推論？在每年會考成績發表的時候，教育部或教育廳或教育局的發言人常常會有以下一類的話：

『這次會考成績，國語最好，平均七十二分・六四，算術最差，平均六十三分・二七。可見各校對於算術的教學尙欠努力。而且去年的算術會考成績平均是六十四分・三八，所以一年來算術的教學非但未有進步，簡直是退步了。』

試問今年的成績怎樣和去年的比較？試題是相同的嗎？如果不是相同的，怎麼知道今年的試題不比去年的難？卽使試題經過嚴密的實驗，難度可以保證相等，六十三分・二七與六十四分・三八相差只一分・一一，怎麼知道不是因爲少數考生偶然疏忽，看錯試題或寫錯答數的原故？更進一步，試問國語與算術的成績怎樣可以比較？怎麼可以知道算術的試題不比國語的試題難？怎麼可以知道算術的批分不比國語的批分嚴？怎麼可以知道算術的課程標準不比國語的課程標準高？

一部份教師當然會想到上面列舉的問題，而感覺發言人的議論是很可笑的。但大多數教師與社會上其他知識份子，除非採取不關痛痒的態度，恐怕都不免接受那結論。否則那種缺乏常識的人又怎會代表教育局或教育廳或教育部，發表那種可笑的談話，而毫不引起反感？

其實，對於統計數字作正當的解釋，不一定需要引用高深的術語，背誦冗長的定義，記憶繁複的公式，或執行長篇的計算。只要對於數字不輕忽，不厭惡，能細辨其意味，作簡單的推演而加以反省，別說關於改革學制和會考成績的那種簡單笑話是可免的，即使遇到比較複雜的問題也不至於不能辨別其是非曲直。例如教育雜誌第二十五卷第十一號所載『記分法新公式的介紹』，多數讀者看了但覺其可厭或可怕，只得置之不理，或盲目採用。但如平心靜氣，細讀一遍，略作簡單推演，不待統計專家，即可加以適當的判斷。下面的評論大部份採自作者在教育雜誌第二十六卷第二號發表關於該記分公式的『商榷。』

普通的記分辦法，以百分為最高分數，以六十分為及格最低分數。至於試題的難度，並無嚴格規定，每題所佔分數，更不是因為難度不同而有多少的。現在我們所要討論的記分公式的發明者認為最適合學生程度的試題應當被受試者百分之五十所答對。每次各個試題當然難易不同，但答對各題的人數平均應當是百分之五十。他們——公式是兩人聯名發明的——採用普通及格標準，認為凡得到最高分數的百分之六十就能及格。他們又認為每一試題所佔分數應當和難度成正比例，即應當和答對的人數成反比例。他們所提出的公式，就是用以計算每一試題應佔分數的符號，略加省略後，形式很簡單，意義也明白：

設q是某一試題答錯者與全體受試者的百分比，

而n是試題總數，

則以n除q的兩倍而得的Q $\left(Q=\frac{2q}{n}\right)$ 便是答對某題者應得的分數。

原文中舉有六個例子，茲鈔錄其一，以見公式的應用。受試者一百人，試題共有五個，各題的難度與應佔分數如下：

題次	答對人數	答錯人數	q	應佔分數
一	五〇	五〇	五〇	二〇・〇
二	七〇	三〇	三〇	一二・〇
三	三〇	七〇	七〇	二八・〇
四	九七	三	三	一・二
五	三	九七	九七	三八・八

普通的記分辦法，五題不分輕重，各佔二十分。照那公式，試題愈難則所佔分數愈多。第一題因爲難易適中，佔二十分。第三第五兩題比較難，所以所佔分數比較多，第三題佔二十八分，第五題佔三十八分有餘。第二第四兩題比較容易，所以所佔分數比較少，第二題佔十二分，第四題佔一分有餘。這是那公式的主要

『貢獻。』此外，公式的發明人又加上兩個附帶的條件：一條是關於試題的平均難度，他們認爲適合學生程度的試題應當平均被百分之五十所答對。一條是關於及格標準，他們認爲普通以最高分數的百分之六十爲界限，在他們的記分法也可適用。我們先把這兩個附帶的條件加以研究，然後再分析那公式的要點。

太容易的試題只能考驗少數壞學生，太難的試題只能考驗少數好學生，惟有百分之五十能解答的試題，其辨別力最大。這是編造測驗者的常識，是有學理根據的。但是這個條件若要應用於學校中的考試，就不免有問題了。因爲考試所問應當是每一學科的重要材料，重要材料當然不應當只有一半學生學會。普通的試題大概平均總有百分之七八十被答對。如平均只有百分之五十答對，普通一定認爲試題太難，或學生太差，成爲很特殊的結果。

至於及格標準，普通雖規定六十分，但因記分方法不盡客觀，容許種種伸縮，六十分的意義並無嚴格解釋。實際上教師往往並不把六十分做及格的定義，卻把及格當做六十分的定義。換句話講，他們往往並不因爲某生超過了六十分才給他及格，或因爲某生還不夠六十分就給他不及格。他們的辦法是倒過來的，因爲某生應當及格才給他六十以上的分數，或因爲某生不應當及格就給他六十以下的分數。不過我們所討論的公式是只適用於客觀的，嚴格的，正誤可以明確決定的記分法，如算術那種科目，或其他科目採用新式試題的時候。公式的發明者既主張試題應當平均被百分之五十所答對，恐怕多數受試者得不

到六十分吧！

我們就用前面的例子，來推算受試者平均可得幾分。我們已經知道每題答對人數與應佔分數，二者相乘即是全體受試者在該一試題上所得分數，所以：

題次	答對人數	應佔分數	全體受試者所得分數
一	五〇	二〇・〇	一〇〇〇・〇
二	七〇	一二・〇	八四〇・〇
三	三〇	二八・〇	八四〇・〇
四	九七	一・二	一一六・四
五	三	三八・八	一一六・四
總計		一〇〇・〇	二九一二・八

受試者共計一百人，合得二千九百餘分，平均每人僅二十九分！

讀者或者不免懷疑，試題既是平均被百分之五十答對，爲甚麼受試者不是每人平均得五十分？這層初看確乎有些奇怪，但細看前表也不難明白。五題中只有第一題恰被百分之五十所答對，所以在一百分中佔二十分，一百人共得一千分，所以每人平均得十分。如五個試題都是這般難度適中的，受試者確會平均得五十分。但其餘四題或難或易，在第二或第三題上，一百人共得八百四十分，平均每人僅得八分，在第

四或第五題上，一百人共得一百十六分，平均每人僅得一分。這是因為分數佔得愈多的試題答對的人數愈少。

所以，應用那公式後，如每一試題都是被百分之五十所答對，則受試者平均每人可得五十分。試題的難度如有參差，參差愈大，則平均分數愈低。為使這一點更加明白起見，不妨假設一種極端的情形：試題僅有兩個，一個全體答對，一個全體答錯，平均仍是百分之五十答對。照那公式被全體答對的試題不能佔分數（因為無人答錯，q是零，Q也是零，）所以在該題上全體受試者共得零分。全體答錯的試題獨佔一百分（因為q是一百，n是二，q約二倍被n除仍是一百，）但既無人答對，在該題上全體也得零分。結果當然全體都是零分了！

以上的分析已夠證明那『記分法新公式』是不適實用的。現在我們可以更進一步，討論該公式的基本假定，卽難題應多佔分數之說，判斷其是否合理。

一個試題的難易，若憑答對人數多少以決定，可以由於兩種不同的原因。第一種難易由於所需能力的高低與經驗的多少，例如能識字的人比能造句的人多，能造句的人比能作文的人多，因為不能識字或未學識字的人一定不能造句，不能造句或未學造句的人一定不能作文，而識字的人也未必能造句，能造句的人也未必能作文。同理，加減比乘除容易，整數比分數容易，算術比代數容易。若試題的難易是指這一類難易，答對較難的試題者一定能夠答對較易的試題，而答對較易的試題者未必能夠答對較難的試題。

第二種難易則由於偶然經驗的不同，例如兩個形義複雜程度相等的字，一個因爲常用，所以識的人多，另一個因爲不常用，所以識的人少。又如兩件相似的史地事實，一件因爲教師認爲重要而認眞教學，所以了解與記憶的人多，另一件因爲教師認爲次要而約略提及，所以了解與記憶的人少。試題的難易若指這一類難易，答對較難的試題者有時反而不能答對較易的試題。

第二種難易並不是眞的難易。所謂容易的試題是關於習見的，常用的，和多數人所注意的一類知識技能。所謂難的試題是關於偏僻的，罕用的，和多數人所忽略的一類知識技能。所以愈難的試題實際愈不重要，若說難的試題應比容易的多佔分數，是等於說不重要的試題應比重要的多佔分數，顯然不合情理。

第一種難易是眞的難易，因爲答對難的試題者一定能答對容易的試題，而不能答對容易的試題者一定不能答對難的試題。茲以加法位數爲例，假定一位加法的難度等於一，兩位的難度等於二，餘依次類推。主張難的試題應當多佔分數者大概根據以下這樣的理由：

『兩位的加法不是比一位的難麼？做對兩位的加法不是比做對一位的表示更高的能力麼？若做對一位的給一分，做對兩位的也給一分，不是把兩種顯然有高低的能力同樣待遇麼？若一位加法的難度是一，而兩位的難度是二，當然做對一位的只應得一分，而做對兩位的應得二分。』

以上的議論是不錯的。假設能做兩位加法的人不另試做一位的，或所做一位加法不另給分，則做對兩位加法當然應得兩分。但實際通行的記分法和適用那公式的記分法都是每題給分的。若能做一位加

法的給一分，能做兩位的給二分，則能做兩位的因爲也會做一位的，可得三分。同理，能做三位加法的不僅得三分而可得（三加二加一）六分，能做四位加法的不僅得四分而可得（四加三加二加一）十分。能力是一二三四之比，而所得分數成爲一三六十之比了。若不問試題難易，能做一位加法者給一分，能做兩位者加給一分，能做三位者再加一分，則所得分數恰與所有能力相稱。於此可見難題應多佔分數之說實屬似是而非，公式的發明者未免自尋煩惱了。

照前面的討論，那公式的不合理，不適用，是很容易明白的。做教師的都應當有充份的統計常識，判明其不可採用，編雜誌的也應當有充份的統計常識，決斷其不必予以介紹。但教育界人士大都不習於數字及公式，見了不敢自加思考，以爲創造公式的人總有其理由，不是普通人所可非難的。這種情形，好比領款或付款者不能自己點數，而任人發給或索取，勢非受人欺騙不可。

前段所謂統計常識，其主要部份是一種態度一種精神不輕忽數字不以數字爲可厭或可怕能欣賞數字，能加以推演而發現其含蘊的意義能謹愼反省自己不輕下結論也不輕易接受別人的結論有了這種精神，未學統計方法者不難無師自通。缺了這種精神，卽使勉強學得統計方法的皮毛，結果也只是自誤誤人而已。

第二講　統計的材料與方法

『統計』兩字在第一講中已經見過面，但尚未正式向讀者介紹。究竟統計是甚麼？

留意教育狀況的人一定會問，『現在中國各省市各種學校有多少？各校學生有多少？各處教育經費有多少？』留意教育發展的人一定會問，『歷年各種學校，各校學生，教育經費，是否年年有增加？』留意兒童身心發展的人一定會問『兒童的身長、體重、知識、學業等等怎樣跟着年齡進步？』要回答這類問題，必需直接或間接觀察事實，觀察到的事實必需利用數字來表達。根據直接或間接觀察，用數字表達的事實，便是統計。

所以統計的主要特性有兩點：第一，統計所代表的是直接或間接觀察所得的事實。第二，統計是用數字表達的。

考西文中關於統計的『statist』、『statistics』、『statistical』等字都是從拉丁字『status』來的，有政治的『國家』的意義。照一百五十年前歐洲著作家的用法，『statistik』或『statistics』是國家重要情形的表明，而表明的方式，在那時既缺乏精確的記載，幾乎全是用文字的。所以在近百五十年間這名詞的意義曾經過顯著的演化，一方面漸漸地專門適用於數字表達的事實，一方面從國家的政治經濟推廣應用於社會、教育、生物、心理、氣象等等每種科學與每種事業。換句話講，統計的內容日趨精深，統計的應用日

趣廣博。

根據直接或間接觀察，用數字表達的事實也稱做『統計材料』因為『統計』兩字有時當做『統計方法』或『統計學』用。本書所要談的『教育統計』並不是教育統計材料的本身，而是解釋與處理教育統計材料的方法。

但是方法也隨着材料而不同。所以材料的性質也值得研究一番。我們現在先舉多少統計材料的例子於下，再討論怎樣歸類。

1. 民國二十一年全國各省中等學校學生人數。(表一)
2. 清代進士及第之省區的分佈。(表一)
3. 現代中國人物之籍貫分佈。(表一)
4. 民國元年至二十一年全國專科以上學校學生、教員、及經費之比較。
5. 民國元年至二十一年浙江省初等學校及中等學校歷年之學校數、學生數、及經費數。(表二)
6. 民國二十一年度全國專科以上學校各科學生之人數及百分比。(表三)
7. 民國二十五年國立浙江大學入學考試考生智力測驗之成績等級。(表四)
8. 民國二十五年國立浙江大學入學考試考生智力測驗之成績分數。(表五)
9. 七千五百二十四常用字之筆畫。(表六)

10. 八千五百八十五人之身長。（表七）

這十個例子的大部份有表附在本書的後面。很明顯的，表一的省區與表三的學科是並無一定排列次序的。山東可以排在四川與河南之間，河南也可以排在山東與四川之間。工程可以排在農林與醫藥之間，農林也可以排在工程與醫藥之間。表二的年度與表四的等級卻不是可以任意排列的。二年必需排於元年與三年之間，『良』等必需排於『優』等與『常』等之間。至於表五的分數，表六的筆畫數，表七的尺寸，也都非依照次序不可，非由大而小，即由小而大，絕不容第三種排列法。所以統計材料可以分爲兩大類，一種是無秩序的，一種是有秩序的。

表四與表五的成績雖然同是有秩序的，但表四的等級是用文字形容的，表五的分數是用數字形容的。所以有秩序的統計材料又可以分爲兩類，一種是用文字形容的，一種是用數字形容的。

表六與表七的材料雖然同是用數字形容的，但也有一點根本不同的地方。表六的筆畫一定是整數的，三畫與四畫之間別無其他可能的筆畫。表七的長度的性質完全不同，用尺不能分別的時候可用寸，用寸不能分別的時候可用分，並無一個絕對不能再小的單位。不論兩個長度相差怎樣少，介乎其間的另一長度總是可能的。所以用數字形容的統計材料又可以再分爲兩類，一種是間斷的，一種是綿續的。

以上所講的統計材料分類法完全根據材料的形式，不問其內容。若從內容講，表一的材料是地理的或空間的，表二的材料是時間的或歷史的，表三與表四的材料是品質的，表五表六表七的材料都是數量

的。

再進一步，品質的材料又可以分爲無秩序的與有秩序的，數量的材料又可以分爲間斷的與綿續的。

因爲地理的或空間的材料是無秩序的，時間的或歷史的材料是綿續的，所以形式分類與內容分類的關係是很簡單的。現在把兩種分類法對列如下：

形式分類			例子	內容分類	
無秩序的			(表一)	地理的或空間的	
			(表三)	無秩序的	品質的
有秩序的	文字形容的		(表四)	有秩序的	
	數字形容的	間斷的	(表五)(表六)	間斷的	數量的
		綿續的	(表七)	綿續的	
			(表二)	時間的或歷史的	

這兩種分類法不但適用於教育統計材料，也適用於任何其他統計材料。

表一至表七所陳列的統計材料當然限於某一階段，一方面已經經過一番整理，不能說是粗糙的原料，但也不是完工的分析結果，因爲更進一步的研究還等着進行。例如表五所列事實，其初步材料應該是這樣的。

姓名	測驗分數	本校錄取否	他校錄取者校名
趙甲	二五八	不取	
錢乙	五一六	取	交通大學
孫丙	三七二	取	中央大學
李丁	四二四	取	
……	……	……	

如要知道某人得到幾分，非查看這初步材料不可。列成表五之後，個別事實便不能顧到了。而且表五以二十五個不同的分數併入一組，所以全體的最高分數與最低分數也不能確定，所知道的只是最高分數不出乎五二五分至五四九分的範圍，最低分數不出乎一〇〇分至一二四分的範圍。但表五可以顯示測驗分數的分佈情形，不像初步材料那般紛繁零亂，看了茫無頭緒。因爲注意了大體，勢必忽略細節。如目的在欣賞隊伍的排列操演，當然不能細辨各人的身材面貌了。

在某種情形下，表五所列的事實固嫌不夠詳盡，但在別種情形下，那樣的事實或許又嫌不夠簡約了。表五所列的事實可以簡約成以下九個數目：

受試者	中位分數	下四分位	上四分位
投考者全體	三一二・一	二六一・一	三六六・八

本校錄取者	三七二・二	三二三・六	四二四・一
亦被他校錄取者	三八八・八	三三九・五	四四六・七

或以下六個數目：

受試者	平均分數	標準差
投考者全體	三一五・七	七五・九
本校錄取者	三七四・六	七〇・〇
亦被他校錄取者	三九二・八	六六・九

以上十五個數目不能算是統計的『原料』，而是經過統計方法造成的『出品』了。至於『中位』『下四分位』『上四分位』『平均』『標準差』這些『出品』的意義與算法，正是我們以後所要詳細說明的，現在先略談統計方法的大概。

統計方法的應用不外乎三方面的工作：（一）搜集材料，（二）敍述事實，（三）考驗假設。

關於搜集材料，各種學科與各種事業因爲各有特性，各有其特殊的方法，統計學不能一一加以研究。但不論搜集何種材料，如不能絕對正確，材料的誤差便是統計學的研究對象。其實『絕對正確』在任何材料是不可能的，不過物質科學的實驗可以施行比較嚴密的控制，減低誤差，所以需要統計方法的時候比較少。生物科學雖也可以應用實驗方法，但有多數因子是不能嚴密控制的，所以實驗與統計常常聯合

應用。至於社會科學，實驗方法幾何是完全不適用的，所以統計方法更加重要。教育統計的材料大部份是社會現象，小部份是生物（心理）實驗的結果。我們預備在第三講專門討論統計材料的誤差，包括其來源與性質。

材料既經搜集之後，怎樣加以整理，怎樣加以分析，怎樣作扼要的敍述，這是統計方法的主要任務。從紛繁零星的個別事實，歸類分組，排列成表一至表七那樣有意義有組織的方式，只是工作的第一步，尙有進一步分析的需要。例如從表五所列事實，雖然可以知道錄取者的成績比投考者全體的成績好，又可以知道錄取者的成績不若投考者全體的成績參差不齊，但怎樣對於這兩點作更扼要的敍述，就非應用統計方法再加分析不可。又如考生錢乙得五一六分，這五一六分究竟有甚麼意義，應當怎樣解釋，在投考者全體中佔甚麼地位，在錄取的考生中佔甚麼地位，這類問題也非應用統計方法不能回答。

再如從表一所列數字，可見民國二十一年的中等學校學生，清代的進士及第，與鮑伍二氏人名錄中的現代名人，都是江蘇一省最多，雲貴陝甘等省最少，三者似乎有密切的關係。但其他各省的情形又不很一致，例如廣東與四川兩省的中等學校學生佔第二三位，而其進士及第與現代名人比較少，浙江省的進士及第與現代名人都佔第二位，而其中等學校學生僅佔第八位。這三種人物的省區分佈情形究竟有多少關係？這種關係應當怎樣敍述？再從表五所列事實，讀者或也不免想及類似的問題：入學考試的成績與入學後的學業成績有多少關係？能否根據前者以推測後者？怎樣推算法？推算的正確程度又怎樣？這類問

題更非應用統計方法不能回答。

統計方法的一大部份就是敍述事實的方法，普通統計書籍的大部份篇幅也都用於說明這種方法。但若應用統計方法的研究者計畫以所得事實與某種假設互相比較，俾決斷其是否符合，卽感敍述法不夠用。例如某省有小學生四十二萬九千人，其中二十二萬一千是男生，二十萬八千是女生，我們不可遽然斷定該省的教育重男輕女，因爲卽使男女絕對平等，雙方數目也會有偶然的出入，譬如投擲四十二枚硬幣的結果，未必正反兩面各得二十一枚。如果投擲四十二萬九千枚硬幣的結果，正反面相差至一萬三千是相當容易的，男女生人數的差別也只得認爲是偶然的，是無關緊要的了。所以解決這個問題的關鍵有兩點：第一，必需推演投擲四十二萬九千枚硬幣的所有可能結果。第二，必需決定怎樣算做『相當容易』。這類就是考驗法的任務。普通統計書籍對於考驗法只有極簡單的說明，本書也未能破例。但近年來統計學在這方面突飛猛進，從事實驗者尤需要充份加以利用。作者所著實驗設計與統計方法一書（中華書局發行，印刷中）是專門討論考驗法的。

以上所講是統計方法的三種應用。至於我們研究統計方法，也有三種可能的目標：

一、明了普通的統計術語，對於統計數字能作確切的解釋，審辨統計的歧途，不至接受錯誤的結論。

二、應用統計方法於其所努力的學問或事業，將調查或實驗所得結果作適當的整理和考驗。

三、成爲統計學家，對於統計方法本身有所貢獻。

多數讀者大概並無第三種野心，作者也並不計畫鼓勵每一讀者成爲統計專家，因爲社會上還有很多事情值得讀者努力。而且理論統計學的研究需要高深的數學，也不是每一讀者所能感覺與趣的。

第二種目標是讀者可以普遍採用的。不論所做的調查或實驗是怎樣簡單，其結果一定需要應用統計方法來整理和考驗。卽以普通教師而論，至少必需時時處理學生的學業成績，而學業成績的處理是必需應用統計方法的。此外如行政人員有所考核或計畫，也非根據事實作統計的分析不可。至於各種問題的研究人員，要從紛繁的材料歸納到扼要的結論，統計方法更是惟一適用的工具。

一部份讀者或許並不希望自己做任何統計的工作，但求了解人家所報告的結果。換句話講，他們並不計畫自己使用統計的工具，卻願意學些統計的常識。這好比有志讀書而無意寫作，那就是採取最低限度的第一種目標了。統計數字既是現代生活中常常遇到的，對於統計數字作確切的解釋，應當是受過普通教育者的常識。缺乏統計常識的人隨時會遇到困難，隨時會遺誤事機，隨時會鬧出笑話，第一講中已經舉過例子。所以第一種目標應當是普通教育的目標之一，不若第二三種目標是屬於高等教育與專門教育的。

第三講 觀察與取樣

現在我們假定要調查上海市內初中學生的國文程度，討論進行的方法。首先應當決定的問題是：甚麼是國文程度？國文包括那幾方面？程度憑甚麼單位查考？應當怎樣查考？上海市內初中學生有多少？能否一一加以查考？可否抽查一部份？用甚麼方法選擇？前四問題是關於觀察的，後四問題是關於取樣的。

觀察的第一步是確定觀察的目標，使所觀察的事實即是所要觀察的。調查學生國文程度的時候，如果注重寫作而忽略閱讀，或注重閱讀的速度而忽略理解，或注重單字的意義而忽略成語的用法，結果只是調查了國文程度的一方面。當然，我們可以縮小觀察的目標，專門研究國文程度的某一方面，那也是一種辦法。要緊的是不要像盲人摸巨象那般誤認一小部份做全體，引起嚴重的誤差。

觀察的目標既經確定之後，其次是規定觀察的單位。如果所要觀察的是具體的物件，問題比較簡單，但也必需於進行觀察之前下一定義，否則臨時便會發生困難。例如調查教師的時候，不在學校供職的家庭教師，或另有其他職業的兼任教師，是否計算在內，應當預先決定。

關於學業成績的單位，理論上是非常複雜而尚未得到完滿解決的問題。但實際上惟有採用測驗做對題數或所得分數，雖然不能與物質測量所用單位相比，至少比普通教師日常所批分數有意義。如調查者並無適用的現成標準測驗，必需自行編造，其工作當然非常費事，或竟成為全部研究的主要貢獻。但測

驗的編造已經自成一種學科，另有專書討論，本章只得從略。

除了確定目標與規定單位以外，觀察的方法應當儘量客觀化，觀察的情境應當加以適當的控制。客觀觀察的好處是結果不因觀察的人而不同。例如新式試題用塡充、彙選、是非、配偶、重列等法，答案的正誤可以明確斷定，可以預先做好批分的標準，不容異議，所以結果正確。反之，論說題的答案，評判常有困難，結果因人而異，所以這種分數是不客觀，不正確的。如同一事實可以採用客觀的方法觀察，主觀的方法應當摒棄不用。但有些問題不能全用客觀的方法解決，卽如作文的評判，或任何藝術創作的估量，不能不以個人的意見爲最後的依據。除非我們根本放棄這類問題，不加硏究，便不得不採用主觀的觀察法。不過主觀觀察的結果旣因人而異，同一作品應當由兩人或更多人觀察，才可以比較其結果，而估計其誤差。

不論客觀或主觀的觀察，控制觀察的情境也可以使觀察更加正確。例如考試限定時間，考生不許發問，考場附近不許高聲談笑，就是情境的控制。有了這種控制，考生才受公平的待遇。否則甲生可以多寫幾分鐘，乙生可以從教師探得暗示，丙生或因外面聲音而不能專心，便引起觀察的重要誤差了。有嚴密控制的觀察，其誤差可以減至最低限度，所以比普通觀察正確。但控制是相對的，不是絕對的。有一部份因子，例如考生的隔夜睡眠情形，顯然會影響第二天的考試成績，但極不容易加以控制。因爲控制不能完善，考試成績不能十分正確。

非但控制不完善的觀察方法，就是完全未加控制的觀察方法，並不是完全不可採用。不過採用的時

候應當研究其誤差的性質與大小。最緊要的問題是：未加控制的因子對於個別事件是否有一致的影響？例如前段所講睡眠情形會影響考試成績，只要考生的睡眠情形是個別不同的，個別的成績雖然可以發生重要的誤差，但這種誤差因爲並不一致，在全體仍是不重要的。反之，如果多數考生寄宿一處，而宿舍於考試的前夜發生火警，或其他使人不能安睡的事故，這些考生的成績便因之而一致降低。這種誤差非但在個別成績是重要的，在全體成績也不容忽視。假說我們已經知道有這類特別事故發生，當然可以改變考期，以免受其影響。可慮的是有些我們從未想到的因子，或者也會引起出乎意料之外的重要誤差。查考這類誤差的惟一辦法是用預先經過試驗證明難度相等的試題，在同樣（不加控制）的情形下再行考試一次，以兩次結果互相比較。這種辦法雖然對於觀察的時間與費用必需加倍，但因其能供給估計誤差的資料，是鄭重其事的研究者所樂於採用的。

假設調查國文程度的人非但對於考試時間與其他情境不加規定，根本並不編造或採用一套題材做觀察的工具，而只是向各校徵集學生的國文作品與試卷，那就成爲完全不加控制的觀察方法了。從這種材料，研究者可以分析學生所用的單字、成語、句法等等，也可以得到很有價值的結論。但這是觀察現成的事實，當然不能像規定考題的實驗那般可以自己計劃預冀的結果。

以上所講的觀察，不論是否完全客觀，不論是否有嚴密的控制，都是直接的觀察。但我們有時也採用間接的觀察方法。間接觀察對於調查初中學生的國文程度是大不相宜的，但也不是絕對不適用的。尤其

在初步研究期間，訪問初中學校的國文教師，探詢其平時觀察的結果，也是有相當價值的。因爲教師與學生天天接觸，所知應當很詳細，所以訪問法或者還可以得到直接觀察所不能得的材料。不過訪問者必需牢記，訪問所得不一定是事實，或者只是教師的意見。因之，教師的心理也有加以研究的必要。也許那位教師是一個悲觀者，所以他認爲學生的程度一天不如一天。也許那位教師自己不能引起學生的興趣，因之他認爲學生對於這基本學科的態度不好。總之，研究者切不可誤信意見爲事實，訪問的結果才有意義。

訪問雖然不大適用於眼前的問題，卻是搜集統計材料者所常常採用的方法。若所要調查的即是被訪問者的意見，這方法當然成爲直接的觀察方法，並無間接觀察的缺點了。但多數調查者的目的仍在事實，所以採用訪問或其他間接觀察方法的時候應當非常小心。重要的問題有兩個：第一，被訪問者是否能夠講出你所要知道的事實？第二，即使能夠，他是否願意眞實地奉告？例如有人訪問主婦，調查家庭費用，恐怕大多數主婦可以歸爲三類：有的自己並不清楚，隨口胡說一個數目了事。有的自知浪費太大，以爲有隱瞞一部份的必要。有的以爲節儉是可恥的，不免故意誇張幾分。無經驗的訪問者或者不能想像人家會毫無理由地欺騙他，但這種不必要的說謊似乎是很通行的。

比訪問法更加流行的間接觀察是問卷法。問卷與訪問不同的地方是以文字代語言，以郵信代奔走，所達到的範圍可以比較廣大，所搜集的答案必需比較簡短。發出問卷的人應當預料有一大部份永遠不會寄回，否則必至大失所望。寄回問卷者與不寄回者的心理有甚麼差別，也是值得考慮的。如果這兩種人

對於當前的問題有顯然不同的態度，收回的答案只能代表一方面的見解，這類誤差是很嚴重的。

凡是自己搜集的材料，不論所用的觀察方法是直接或間接的，都稱爲初級材料。採用別人的初級材料稱爲二級材料。採用別人的二級材料應當稱爲三級材料，但利用現成材料者應當追究初級材料，才是正當辦法，因爲轉載一次即多一次錯誤的機會，除非初級材料已經無法獲得，研究者不應當採用二級材料爲三級材料。

採用別人的材料，必需注意當初搜集材料的目的，觀察的方法，正確的程度，取樣的範圍。甲乙兩問題雖然同樣涉及某種事實，但爲研究甲問題而搜集的材料是否適用於乙問題，往往有考慮的必要，不宜貿然引用。如在同一來源的材料發現互相矛盾的數字，或在兩種來源的材料發現不能符合的事實，應當詳細研究，尋出錯誤的原因，再決定取捨。

前面我們一再講到減低觀察的誤差，即增高材料的正確，似乎數字總是愈正確愈好，但研究者的目標既不是完全免除誤差，也不是盡可能方法減誤差至最低限度。研究者應當首先認清觀察的誤差有兩種：一種是不偏的誤差，是變化無定的，是偶然的，是正負不一，可以互相彌補的。一種是偏的誤差是恆定不變的，是有系統的，是正負一致，積聚爲患的。事實經過觀察手續彷彿物體經過鏡子反照，不偏的誤差彷彿玻璃毛糙不平使影像模糊不清，偏的誤差彷彿鏡子彎曲凹凸，使影像失卻本形。這個差別在前面討論考生睡眠情形的時候已經分辨明白，應當不必再費筆墨。所以偏的誤差是應當盡可能方法予以免除的。至

於不偏的誤差，非但不能完全免除，也不一定應當減至最低限度。只須不妨礙研究的目的，耗費過分的時間與金錢，以求過分的正確，是不必的。

而且誤差的來源不止一處。如果某一方面的較大誤差無法減小，努力於減低另一來源的較小誤差是無用的。這種舍本務末的企圖，當然發生於『察秋毫之末而不見車薪』的眼光，和下段所述莫須有學生的『正確』同樣可笑：

地質學教授於一九三〇年發表其研究結果，斷定某處巖石的年齡大約是三千萬。三年之後，一位過分正確的學生在論文裏講到該處巖石，引用教授研究的結果，說『所以其現在的年齡是三千萬零三』

前面講誤差可以按照其性質分為不偏的與偏的兩種，這分類非但適用於觀察的誤差，也適用於取樣的誤差。取樣也稱為抽樣，即是抽取一部分以代表全體『取樣』這名詞固然是統計專家的術語，取樣的辦法是未學統計方法者所常採用的。例如買西瓜者不能斷定其甜否，惟有嘗試的一法。所嘗試的限於少數，根據嘗過的少數，推測全部，就是取樣。全部西瓜的甜味當然不能與嘗過的完全相等，所以有取樣誤差。假設所嘗的瓜是買者未經選擇，隨便指定的，其他的瓜有的甜些，有的差些，則取樣誤差是不偏的。但如果所嘗的瓜是賣者憑其經驗用心選出的，結果一定是比其餘的瓜特別甜，這取樣誤差便是偏的了。取樣與觀察一樣，免除偏的誤差比減低不偏的誤差更加重要。

不適當的取樣會引起偏的誤差。例如調查全國人民身長體重者限於江浙兩省以內取樣，北方高大的人都被遺漏，結果只量到了比較矮小的，當然不能代表全國。又如測驗十二歲兒童智力者限於小學四年級以內取樣，比較聰明的早已升入高級，受試者都是比較愚笨的，顯然與普通的十二歲兒童不同。調查初中學生國文程度的時候，如徵求學校自動參加，願意合作的大概是程度較好的，不合作而遺漏的大概是程度較差的。又如任各校校長和教師指定一部份學生參加，參加者一定又是程度較好的。可能引起偏的誤差的機會很多，研究者應當謹愼從事。規模大的學校或者比規模小的學校程度高，市立的學校或者比私立的學校程度高，住宅區的學校或者比工廠區的學校程度高，研究者如只顧接洽便利，地點適中，而疏忽學校的性質，取樣便有不公正的危險。

取樣的方法可以分為(一)隨機取樣，(二)立意取樣與(三)混合取樣。

如全體中每一單位被抽取的機會是相等的，這取樣方法便是隨機取樣。假設上海有初級中學（完全中學當然也在內）三百所，研究者可以用相等的卡片三百張，上面寫着學校的名稱或號數，混合排亂之後，隨意抽取一部份。如抽取的數目比較多，取樣誤差比較低，抽取的各校可以與全體無大差別。如抽取的數目少，碰巧也可以做全體的適當代表，但不巧可以成為全體的例外。所以隨機取樣是否最好的取樣方法，一大部份決定於取樣的廣狹。

如取樣不廣，隨機取樣所得結果往往不足以代表全體。在這種時候有人主張採用立意取樣，卽是故

意選擇程度適中的學校。但研究者認爲程度適中的學校其實際程度未必適中，這是一種缺點。即使確是程度適中，這些學校的學生或者又太齊整，不能代表全體的程度參差情形，這是第二種缺點。不先詳細調查多數學校而企圖避免這兩種缺點，是不可能的。如果已經調查多數學校，又何必從中抽取一小部份所以立意取樣普通也不是很好的方法。

如果研究者先按照可能影響學生國文程度的因子，如學校規模的大小，所在地點的性質，市立或私立等，分所有學校爲幾類，每類之中再用隨機取樣方法抽取一定的比例，這是混合取樣，即混合採用隨機取樣與立意取樣兩種方法。因爲先把全體分成若干層級，這混合取樣稱爲層級化取樣。大體上講，層級化取樣可以減少兩種單純取樣的缺點，因而減低取樣誤差。

嚴格從理論上講，調查學生的國文程度，取樣應當以學生爲單位，隨機取樣應當按全體學生用抽籤一類方法指定其一部份，立意取樣應當從每一學校選擇程度適中的學生，層級化取樣應當從每一學校用隨機取樣方法抽取一定的比例。但爲實際便利起見，不得不放棄從每一學校抽取一部份學生的方法，而採用抽取一部份學校用其全體學生的方法。以學校爲單位的取樣方法雖然誤差比較大，但實際上經濟得多。

所以調查上海市初中學生的國文程度，不得不以學校爲單位，採用隨機取樣或層級化取樣的方法。但究竟應當抽取多少學校這問題可不簡單，有三個先決問題：第一，計畫中的研究可以允許多少取樣誤

差?誤差的大小必需不妨礙結論的正確。第二,初中學生的國文程度參差到甚麼地步?這是事實問題,必需調查工作實地進行之後才可以知道,不容憑空猜測。第三,抽取學生或學校的多少與取樣誤差的大小成甚麼關係?因爲取樣誤差與參差程度的估計方法尙待以後討論,此刻不能作很淸楚的說明。但簡要地講,取樣誤差的大小一方面與全體學生的參差程度成正比例,一方面與抽取的學生人數成反比例。

前一點是很容易明白的。齊整畫一的東西隨意抽取一二件,就足以代表全體。反之,參差不齊的東西中抽取少數代表,其勢不能與全體的各種情形都符合。例如嬰兒身體的大小比較一致,量其一二後裁製衣服,幾乎每人可穿。成人就參差得多,照高的量了對於矮的不合適,照瘦的量了對於胖的不合適。所以學生的程度如果齊整,取樣誤差就小,學生的程度愈參差,取樣誤差愈大。

至於取樣誤差怎樣與抽取人數成反比例,大體上也不難舉例說明。爲使情形極端簡單起見,假設全體學生對於某一試題的成績是半數做對半數做錯,如照該一試題給分,半數得一百分,半數得零分,平均五十分。現在用隨機取樣的方法,抽取少數學生做全體的代表。最低限度當然是抽取一人,非一百分卽零分,離平均五十分極遠。如抽取二人爲一組,這二人的樣組有四種可能方式,卽(一)甲乙俱得一百分,平均一百分,(二)甲乙俱得零分,平均零分,(三)甲一百分,乙零分,或(四)甲零分,乙一百分,平均五十分。所以二人一組的取樣誤差只有一半時候與單人一組的同樣大,其餘一半時候的取樣誤差等於零。如抽取四人爲一組,樣組有十六種可能方式:

(一)甲乙丙丁都得零分，平均零分。

(二)甲一百分，其餘三人零分，(三)乙一百分，其餘三人零分，(四)丙一百分，其餘三人零分，(五)丁一百分，其餘三人零分，平均都是二十五分。

(六)甲乙零分，丙丁一百分，(七)甲丙零分，乙丁一百分，(八)甲丁零分，乙丙一百分，(九)乙丙零分，甲丁一百分，(十)乙丁零分，甲丙一百分，(十一)丙丁零分，甲乙一百分，平均都是五十分。

(十二)甲零分，餘一百分，(十三)乙零分，餘一百分，(十四)丙零分，餘一百分，(十五)丁零分，餘一百分，平均都是七十五分。

(十六)甲乙丙丁都得一百分，平均一百分。

所以四人的樣組只有十六分之一的時候平均零分或一百分，有四分之一的時候平均二十五分或七十五分，有八分之三的時候平均五十分。這樣的取樣誤差當然又比二人的樣組小。

以上所講，和五人的樣組，十人的樣組的平均數分配情形都列於後面表八。如以各種樣組平均數的均方差（意義見後）互相比較可見其恰與樣組的大小成反比例。

第四講　從混亂中建設秩序

統計材料既經搜集之後，第一步是歸類工作。歸類以前，事實是個別的，混亂的，無組織的，歸類即是把個別事實加以組織，從混亂中建設秩序。歸類的基本條件有二：

一、各類必需互相排斥，即歸入某一類的不能再歸入另一類，所以任何事件只有一類可以歸入，決不會有重複。

二、各類必需共同盡舉，即不歸入某一類的一定可以歸入另一類，所以任何事件必有一類可以歸入，決不會有遺漏。

可以絕對保證不違背以上條件的歸類法是二分法例如學生分爲男的與非男的（女的，）或寄宿的與非寄宿的（通學的，）或及格的與不及格的。這種分類法只能適用於極簡單的事實，否則就嫌太粗陋了。但二分法如繼續應用，結果可以得到不論怎樣多的類別。例如現代名人可以分爲

江蘇的與非江蘇的，後者再分爲

廣東的與非廣東的，後者再分爲

四川的與非四川的，後者再分爲

河北的與非河北的，……

這樣繼續下去，不難分爲十八省區。非但表一的省區，任何合理的分類，如表三的學科，表四的等級，表五、的分數，都可以算做二分法繼續應用的結果，都合乎前面的基本條件。

除了不重複，不遺漏而外，歸類工作當然還有別的問題應當考慮。類別的大小和總數是否適中？各類是否平等？

分類太少或每類範圍太大的缺點是把很多不同的事實歸爲一類，不顧它們的差別。分類太多或每類範圍太狹的缺點是過分注意了細節，不能表現全體的大勢。至於怎樣可稱適中，一方面當然決定於研究的目標，一方面也決定於材料的範圍與數量。表一把民國二十一年的中等學校學生，清代的進士及第，與鮑吳二氏人名錄中的現代名人按照省區分爲十八類，以見教育、人才，與地理的關係，大體上是很適當的。倘若研究者要做更進一步的分析，要知道一省以內——尤其是人才特別多的省區以內——各縣的分別，當然可以再照縣分類。但初步分析總以省爲單位，可以表現全國的大勢。表三以專科以上學校的學生按照學科分爲九類，也是因爲研究的目標只在比較文理法商農工等科的重大分別。如更進一步，法政可以再分爲政治、經濟、法律，文藝可以再分爲文學、美術、音樂，照此類推，一共分成二三十類也未必太多。表四的測驗成績但分六類，比了表五的十八類當然太不正確。表六的常用字分爲三十三類，略嫌太多，如以兩類併爲一類，也無不可。

地理的材料與品質的材料，歸類的時候，比數量的材料不費斟酌；因爲地理與品質的可能區分方法

是極有限的，而數量可以隨處劃分，第一個問題是分爲幾類，第二個問題是分在甚麼地方，都是需要考慮的。

數量材料的歸類普通稱爲分組。類別的多少問題就是組距的大小問題。如無別的應當考慮的地方，分組以不比十五組少不比二十五組多爲最適當但也看全體材料的範圍與數量。如範圍廣而數量大，不妨多分幾組，如範圍狹而數量小，不妨少分幾組。但少到十組以內，或多到三十組以外，普通總是不相宜的。更有一點，間斷材料的最自然的組距是一，除非全體的限域太廣，不得不用一的倍數做組距，綿續材料的組距則必需斟酌規定。

組距既經確定之後，其次的問題是規定組限，即兩組之間的界限。組限的規定應當顧到兩點：第一，各組以內個別事實應當均勻分佈，如有集中一處的趨勢，這集中處應當在一組的中央，切不可一致偏近上限或下限。第二，兩組之間劃分應當絕對淸楚，歸類的時候不至於有重複或遺漏。

關於第一點，最應當注意的是教師所批的分數。因爲他們往往特別多用五與十，所以從五至九，從十至十四，或從一至五，從六至十的分組法是不妥當的。在這種情形下，惟一適當的分組法是從三至七，從八至十二，使五與十都在一組的中央。分組的時候如不注意這一點，根據分組結果的計算便有偏的誤差了。因爲一經分組之後，原有的確實分數即置之不管，全組都用該組的中點做代表。假設每一分數都是五的倍數，而分組的時候又用五的倍數做每組的下限或上限，算出的平均數便會差二分半之多。如做下限，會

太高二分半，如做上限，會太低二分半。

表五、表六、表七的分數、字數、人數，都是中央幾組多，兩端幾組少。在這種對稱的或略帶偏斜的分配情形下，兩端各組以內的分配情形是相反的，一部份是組內平均偏向上限，一部份是組內平均偏向下限。這是無法避免的。根據這種分組法算出的平均數也不至於有偏的誤差。

再看表十的單字應用次數，是極端的偏斜分配，應用次數少的字很多，應用次數愈多，字數愈少。所以每組之內的分配情形大概也是偏斜的，近下限的字數多，近上限的字數少，試以表十的第一組與表十一的首十組比較，可知實際情形確是這樣的。這是極端偏斜分配的必然特性，不是改換組限所可改善的，算出的平均數也必然太大。補救的方法容後討論。

至於避免重複與遺漏，在間斷的材料是很容易的。例如表五以一〇〇——一二四為第一組，一二五——一四九為第二組，一二四與較小的分數歸入第一組，一二五與較大的分數歸入第二組，一二四與一二五之間並無帶小數的分數存在，所以毫無問題。但假設分數是綿續的材料，或可以帶小數的，一二四與一二五之間有空隙，便不適用了。要免除這缺點，必需下一組的上限即是上一組的下限，第一組寫做一〇〇——一二五，第二組寫做一二五——一五〇，這分組法固然可避免遺漏，卻又引起重複的問題。一二五分究竟歸入第一組或第二組？解決這問題有兩種辦法：或者把一二五分分為兩半，分別歸入第一第二兩組。這是表九所採用的辦法，表九只寫了組中點，未寫上下限，如改寫組限，第一組是五二——五三，第二組

是五三——五四，因爲有一個女人恰是五十三吋所以半個歸入第一組，半個歸入第二組。這辦法實際是正確的，不過用小數計人，或別的完整的東西，好像又不大妥當，因之也不大通行。另外一種辦法是規定一種共同的了解，卽是把下限的數值算做一組以內，而把上限的數值算做一組以外的。所以一〇〇——一二五不包括一二五，而一二五——一五〇是包括一二五的，所以一二五不歸入第一組而歸入第二組，其餘在組限上的數值可以照此類推。表七每組只寫下限，不寫上限，是採用這辦法的。爲使規定的了解不需另作說明，也不至臨時忘卻起見，多數書籍採用以下的寫法：一〇〇——一二四·九九，一二五——一四九·九九，或五七——五七·九九，五八——五八·九九等。

後一辦法的問題是：各組的中點究竟在甚麼地方？這不但是寫法問題，是實際問題，必需查考觀察與登記時所用的單位。如表七的身長，當初所用的單位是八分之一吋，所以第一組中最矮的是五十七吋，最高的是五十七吋又八分之七，中點是五十七吋又十六分之七。若以五十七吋爲下限，五十八吋爲上限，五十七吋半爲中點，便有等於十六分之一吋的誤差了。

所以組限的規定必需參照觀察與登記所用的單位。組限的末位最好採用這單位的半數，下一組的上限就是上一組的下限，既無遺漏，也不會重複。按照這條規則，表五的組限應當改爲九九·五——一二四·五，一二四·五——一四九·五等，表七的組限應當改爲五六又十六分之十五——五七又十六分之十五，五七又十六分之十五——五八又十六分之十五等。

不論是地理材料或品質材料的分類，或數量材料的分組，類別應當平等組距應當一律。表一的省區是平等的，倘省與縣夾雜一起，便不平等了。表三的學科也是平等的，倘法政、文藝等學院的大類別中夾着土木工程、電機工程、物理化學等學系的小類別，便不平等了。表四的等級原是按照下列分數限域規定的：

優等	四七五分至五四九分
良等	四〇〇分至四七四分
常等	三二五分至三九九分
可等	二五〇分至三二四分
劣等	一七五分至二四九分
不及格	一〇〇分至一七四分

每等的限域都是七十五分，所以是平等的，倘限域大小不一，便不平等了。至於表五的分數，每組的組距都是二十五，表六的筆畫，每組的組距都是一，所以也都是平等的。

組距應當相等，這是理由很顯然，不必說明的。但書報刊物中的統計表常有不遵守這條規則的。例如表六的材料，三十畫以上的字一共不過四個，似乎分為三組太費篇幅，這樣想的人便會合併三組為一組，標做『三一——三三』，甚且含糊地寫着『超過三十畫』了事。這種不合理的分組是不必要而可以避免的。但有些材料用等距法分組確乎是困難的。例如表十與表十一所列的單字應用次數，如一律以十為

組距分組，必需分成近千組，卽以百爲組距分組，也需分成近百組。而應用次數同在第一百的字數參差又非常之大，應用一次的字有六百零一個之多，應用一百次的字大概不過五個。所以應用次數少的一端，以百爲組距已經太大，而應用次數多的一端，以百爲組距又太小。所以表十的組距從十擴大至百，再從百擴大至千。表十一更有一部份用一爲組距。大概極端偏斜的分配，事實上不得不用組距不等的分組法。採用這種分組法的時候同時應當設法使各組可以互相比較，如表十除列絕對字數外另列相當於一律以十爲組距的字數，或如表十一於每次組距擴大後仍將小組各數歸併大組與其後的大組並列。這種辦法，好比照了全身的相，因爲面部特別重要，另照半身的相掛在旁邊，毫無不相稱的地方。否則組距不等的分組法相當於頭大脚小的諷刺畫，不足以登大雅之堂。

另外一種處理極端偏斜分配的方法，是用等距的對數分組，如表十二。表十二照每字應用次數的對數分組，每組的組距都是十分之四，但如查看每字應用次數的本數，組距仍是不等的，其與表十或表十一不同的地方是組距不僅有三四種，而是每組逐漸擴大，各組的組距完全不同。原本極端偏斜的分配，改用對數分組之後，偏斜的程度會大減的。

表五的考生人數或測驗卷數，表六的字數，表七、表九的人數，表十、表十一、表十二的字數，統稱次數，指得到這樣分數的卷子發現過多少次，這樣高矮的人發現過多少次。所以這些表都稱爲次數分配表。次數分配表可以畫成兩種圖，一種稱直方圖，一種稱次數多邊形。直方圖的好處是圖中每組所佔面積可以正

確地代表次數，缺點是每組範圍以內的水平線與兩組之間的垂直升降線並不符合實際情形。次數多邊形恰好相反，兩組中點之間的聯絡用斜線表明其逐漸升降，但線下的面積不能正確地代表次數。另一不同之點是兩個或更多的多邊形可以畫在一起以資比較，而幾個直方圖是極不容易清楚地畫在一起的。圖一是從表十一與表十二畫的直方圖，圖二是從表五畫的次數多邊形。

次數分配可以按照其形狀分爲三大類。大多數的分配是中間次數多而兩端次數少，或者對稱或者略帶偏斜，這種稱爲『i』字形分配。極端偏斜的分配，次數最多的地方在一端，另一端次數極少而拖延極長，這種稱爲『j』字形分配。另有一種不大遇見的分配，次數在兩端多而在中部少，稱爲『u』字形分配。表五與圖二的『本校錄取者』成雙峯分配，『亦被他校錄取者』成三峯分配，都是因爲次數少而分組多的原故。三組合併一組後，如表四所示，多峯即消失而成單峯分配。如兩種不同的東西混在一處，當做他們一種，也會成爲雙峯分配，例如表九所列的身長，男女各是單峯分配，混合後成爲雙峯分配。

第五講 三種常用的中心量數

次數分配表列成後第一步分析工作是尋找一個可以代表全體的量數，稱爲中心量數。讀者未學統計方法即已了解而能應用的平均數便是中心量數的一種。但平均並不是惟一的中心量數。常用的中心量數至少有三種，除平均而外，有中位和型範。

我們假定京滬鐵路沿綫的上海、蘇州、無錫、常州、丹陽、鎮江、南京每兩站之間的距離是相等的，這七個地方便可以用一至七的數字代表。再假定某學會的會員散居這七個地方人數如下表：

地點	次序	人數
上海	一	六
蘇州	二	三
無錫	三	一
常州	四	一
丹陽	五	一
鎮江	六	〇
南京	七	三

這張表便是一張次數分配表。最後，再假定這學會要選擇一個中心地點舉行年會，那就等於從這張次數分配表中尋找一個中心量數了。

讀者如按照會員十五人的地點次序計算其平均數，結果得三，所以無錫是會員的平均地點。全體會員集中無錫，需要六人西行二站，三人西行一站，一人東行一站，一人東行二站，三人東行四站。如以一人行一站作爲一個單位，則東行西行各十五個單位，總計三十個單位。平均地點有兩個特點：第一，各會員從各處到平均地點的總行程可以分爲兩半，方向是相反的。第二，如計算各人行程的平方總和到平均地點比到任何其他地點的結果爲小。這兩點在這例子裏並未顯露任何優點。

另一個辦法是選擇七個地點中的第四個，即常州，可以稱爲中央地點。中央地點的好處是把最遠的路程減至最低限度。會員從任何地點出發，不論東行或西行，最多不過行三站路程。所以從南京出發的三位，到常州可以比到無錫每人減少一站路程，似乎是一種優點。但常州是但從地點中選擇的，絕未顧到會員的分佈情形。如計算各會員的行程，東行三站的三人，東行一站的一人，西行一站的一人，西行二站的三人，西行三站的六人。總計東行十個單位，西行二十五個單位，共三十五個單位，比了到無錫多了十個單位。所以中央地點更無甚麼可取之處。

會員住在上海的比任何其他地方多，他們當然希望在上海集會，如果表決照最多數通過，上海就會被選。上海的好處是有六個會員可以完全不必旅行，缺點是從南京去的比較遠，需行六站的路程。這種地

點稱爲型範地點。型範地點的好處便是不必旅行的人數可以增至最高限度，但同時必需增加少數人的路程，比到任何其他中心地點遠。所以型範地點的利弊是與中央地點相反的。

再試試蘇州看，各會員到蘇州需要六人東行，六人西行，東行共二十一個單位，西行共六個單位，總計二十七個單位，比集中於任何其他地點的總行程短。所以蘇州是中位地點。中位地點有兩個特點：第一，相反行程的人數是相等的。第二，總行程減至最低限度。

照前面的討論，中位地點與型範地點似乎多比平均地點好。我們如看重旅行人數，而看輕每人行程，則型範地點是最好的中心地點。假設旅費是不管遠近一律的，則全體集中於型範地點的費用最省。但普通旅費是按照遠近比例計算的，所以應當以總行程的長短爲惟一標準，最經濟的地點是中位地點。

爲顯露型範、中位、與平均三者的差別起見，前面所用的次數分配是很偏斜的。如用完全對稱的次數分配，三種中心量數會集中一處，毫無分別。但其定義與算法仍是不同的。

型範普通稱爲衆數，也稱爲範數。但型範的概念不但適用於數量的材料，也適用於地理的和品質的材料，非但次數最多的數量是型範，次數最多的地點或品質也是型範。所以表一的型範省區是江蘇，表三的型範學科是法政，表四裏投考者全體的型範等級是『可』等，錄取者的型範等級是『常』等。至於數量材料，普通以次數最多一組的中值爲型範。所以表七的型範長度是六十七吋又十六分之七，表九的男子型範長度是六十八吋半，女子型範長度是六十三吋半。

照以上的例子，型範似乎是很容易決定的。但數量材料的分組既缺乏獨一無二的法則，型範就會因之而不同。例如表十與表十一所列同是三種小學國語教科書中四二六九字的應用次數，這四二六九字的型範應用次數照表十是五，照表十一是一，相差很遠。再看表五所列錄取者的測驗分數，除兩端次數很少，無關緊要的各組外，分配表的中部是這樣的：

組限	組中值	次數
二七四·五——二九九·五	二八七	二九
二九九·五——三二四·五	三一二	四〇
三二四·五——三四九·五	三三七	四七
三四九·五——三七四·五	三六二	四三
三七四·五——三九九·五	三八七	五〇
三九九·五——四二四·五	四一二	三四
四二四·五——四四九·五	四三七	三一

各組中次數最多的是五〇，該組中值是三八七，就是型範分數。但如果把前表兩組合併一組，有兩種合併法：

組限	組中值	次數

二七四・五——三二四・五	二九九・五	六九
三二四・五——三七四・五	三四九・五	九〇
三七四・五——四二四・五	三九九・五	八四

或者

二九九・五——三四九・五	三二四・五	八七
三四九・五——三九九・五	三七四・五	九三
三九九・五——四四九・五	四二四・五	六五

型範分數照後法是三七四・五，照前法是三四九・五，與前面的三八七是各不相同的。

因爲表十一的分組法比表十精確，所以我們可以斷定那四二六九字的型範應用次數是一而不是五。但表五的錄取者成績分配是多峯分配，兩組合併一組是合理的，兩種合併的方法卻不能分別優劣，因此正確的型範分數無從斷定，這是型範的一大缺點。所以型範的應用限於不求正確而要迅速得一約略印像的時候。

中位不能適用於無秩序的材料，但是除了數量材料外，也適用於有秩序的品質材料。查表四，投考者全體總計二〇〇八人，半數是一〇〇四人。這二〇〇八人或二〇〇八個成績，按照次序排列後，第一〇〇四與一〇〇五人之間的等級是『可』等，就是投考者全體的中位等級。在同表中，『本校錄取者』與『亦

被他校錄取者』的中位等級都是『常』等。因爲表四是近乎對稱的 i 字形分配，所以中位等級和型範等級是相等的。

數量材料的中位比品質材料不容易計算，因爲品質材料的一類是假定一律的，而數量材料的一組只是一個限域，可以再畫分的。前面已經講過表四的『可』等實際上是二四九·五——三二四·五分的一組，現在如將表四的等級改爲分數，原來的中位『可』等不能卽改爲二八七分，因爲第一〇〇四與一〇〇五人並不在該組的中央。詳細情形是這樣的：

可等以下	卽	二四九·五分以下	四〇一次
可等	卽	二四九·五——三二四·五分	七三三次
可等以上	卽	三二四·五分以上	八七四次

我們現在假定二四九·五——三二四·五分這範圍內的七三三次是均勻分配的，目的在從這範圍內尋找一個一〇〇四次在其下，一〇〇四次在其上的中位。先將『可』等以下的四〇一次與『可』等以上的八七四次除外，『可』等以內的七三三次應當有六〇三次在中位以下，一三〇次在中位以上。所以二四九·五——三二四·五這限域應當分爲七三三段，其六〇三段在中位以下，一三〇段在中位以上。所以中位是三一一·二分。

讀者或者要責問從表五可知二四九·五——三二四·五分範圍以內的七三三次並不是均勻分

配的，現在假定它們均勻分配，不會引起誤差麼？當然，除非偶然的巧合，這假定不能與事實完全符合，但在正確事實不可得的時候，這假定比任何其他假定合理。現在既有表五，分組較細，據以計算的中位應當比根據表四正確。這計算見表十三，求得的中位是三一二·一分。因為表五是近乎對稱的次數分配，中位附近的次數雖非絕對均勻分配，也不是極不均勻，所以假定其為均勻，尚不至引起嚴重誤差。試看一組分為三組後，求得的中位僅從三一一·二變為三一二·一，相差極少，不像型範那般因分組不同而大有出入。但在顯著偏斜的分配，中位附近的次數決不會很均勻，若從寬廣組距的次數分配表求算中位，結果一定有嚴重的誤差。例如表十一所列與圖二所示極端偏斜的次數分配，中位附近的次數顯然有峻峭的升降，所以除非組距很小，求出的中位一定不正確，而且從偏斜的方向可知其一定太大。

表十一的總次數是四二六九，半數是二一三四·五，所以中位在一一——二〇或一〇·五——二〇·五那一組範圍以內，如用該組計算，所得中位是十一。這中位一定是正確的，因為十一是該組十數中最小的，決不會太大。但如不從以十為組距的『一一——二〇』組算，而從以百為組距的『一——一〇〇』組算，結果得五十七，就太大了四十六。如從以千為組距的『一——一〇〇〇』組算，結果得五百零五，太大了四百九十四。所以從極端偏斜的次數分配求算中位，應當特別小心，必需採用最小的組距。

平均是只能適用於數量材料的，但實際上卻是應用最廣的中心量數，因其意義與算法是未學統計方法者所家諭戶曉的。平均是全體數量總和被總次數除後所得商數，隨着任何單個數量的變動而增減

所以比了中位與型範的計算有兩點不同的地方：第一，如有全體數量的總和與總次數，雖缺少分配情形平均也能算出，而中位與型範是無從決定的。例如我們曉得某省有小學一萬二千所，有小學生六十萬，但並不知道這一萬二千所小學每校的學生人數，型範與中位是顯然無從計算的，而平均每校學生五十人，用一次除法就夠了。第二，如次數分配表不完全，一端只寫不及若干而並不註明下限，一端只寫超過若干而並不註明上限，型範與中位往往仍可計算，而平均卻無從決定了。例如表五所列投考者全體成績分數，除中間一組標明是從二九九・五到三二四・五外，其餘各組可概括於『不及二九九・五』與『超過三二四・五』二句之下。這種殘缺不全的事實無法可以求算其平均，而中位仍可照樣求出。如『二九九・五——三二四・五』的前後列出兩組，當然仍不足以計算平均，而型範卻就可以決定了。

用次數分配表計算平均，每組以中值代表全體。如次數分配是對稱或略帶偏斜，其誤差是不偏的，所以是不關緊要的；但在極端偏斜的J字形次數分配，每組的次數都是偏向一面的，所以組中值不能代表全體，平均的計算非常困難，所以不如中位適用。

從次數分配表計算平均數，當然可以用每組的中值乘該組的次數，再將乘出的積數加起來，最後以總次數除積數的和數即得。但這方法很費事，而且錯誤的機會也多，不如表十三所示的方法簡捷。讀者試先照前法將平均算出，再與表十三比較，以見結果相等而繁簡大有不同。

除了計算平均而外，表十三的一部份是為計算中位與幾種參差量數（見後）用的。如果只要計算

平均，表十三的第一列『組限』可省，第五列『二次累計』更用不着，只需『組中值』『次數』與『累計』三列就够了。『組中值』與『次數』不需任何說明，『累計』是次數的累計。求算平均數的步驟現在說明如後：

一、排列次數分配表，各組應當標明『組中值』『組限』可省。

二、觀察全表的形勢，選擇一個適中的組中值做出發點。這出發點愈近平均愈便利，但平均正在求算中，無從先知其確實結果，只能憑全表形勢猜測，好在卽使所用出發點離平均很遠，計算的結果依然正確，只是所用的數字比較大些，略爲不便。如次數分配是對稱或近乎對稱的，選擇最中心的組中值最妥當。表十三有十八組，所以採用第九組。如次數分配是顯然偏斜的，應當在最中心組與型範所在組之間的各組中選定一組。如同時計算中位，出發點不應當預先決定，等計算『累計』的時候再看。

三、累計是從分配表的兩端向出發點算的。從上而下，第一數與次數一樣是二，其次的十六是二與第二組次數十四的和數，第三是四十，是這十六與第三組次數二十四的和數，這樣加下去，到出發點的前一組爲止。從下而上也是這樣，先是三，三加十七是二十，二十加二十五是四十五，依次加上去，加到八七四爲止。如出發點並未預先決定，不論從上而下或從下而上的累計以不超過總次數的半數爲度。表十三也符合這條件，從上而下到八七二，從下而上到八七四，這兩數的任何一個如再

加二六二便要超過總次數的半數了。

四、累計算出之後，可以核對一下。從上而下的最後一數，即八七二，從下而上的最後一數，即八七四，再加上出發組的次數，即二六二，這三數的和數應當等於總次數即二〇〇八。如結果未能符合，必有差誤，應當尋出改正。

五、將所有從上而下的累計加起來，得二二八一。所有從下而上的累計也加起來，得二五七八。再從二五七八減去二二八一，得二九七。此處應當注意，因爲次數分配表上各組是小的在上大的在下，所以才將從上而下累計的總和從從下而上的累計總和減去。這是最便利的寫法。假設各組是大的在上小的在下，被減數必需寫在減數之下了。

六、前一步的結果（二九七）被總次數（二〇〇八）除。以所得商數（・一四八）乘組距（二五），加於出發組中值（三一二），即得平均數（三一五・七〇）。如前一步中減數反比被減數大，所得差數當然成爲負數，此處最後一步的加法應當改爲減法。

現在將這三種中心量數的差別列在後面，這些差別大半都已討論過了。

	平均	中位	型範
有時稱爲	「平均數」或「均數」	「中數」	「衆數」或「範數」
英文原名	Mean	Median	Mode

常用符號M	Mdn	Mo
從該處出發至全體各數其距離的平方的總和最小	其距離的總和最小	距離等於零的次數最多
次數分配圖上重心點所在	平分面積爲二的界綫所在	最高峯所在
可以正確決定	可以正確決定	往往不能
受全體任何數值的影響	有時不受影響	有時不受影響
僅有數值總和與總次數亦可算出	不能	不能
從各部份的平均可以計算全體的平均	不能	不能
不能從不完全的材料算出	有時能	有時能
僅適用於數量材料，不適用於品質材料或無秩序材料	亦適用於有秩序的品質材料	亦適用於無秩序的材料

第六講　參差量數

按照表八樣組平均數的分配情形視樣組之大小而不同但不論每組是一或二或四或五或十，一〇二四個平均數的平均或中位總是五十，其不同處全在參差的程度。樣組小至一的時候，平均數（其實是用不着平均的本數）非零卽一參差達最高度樣組逐漸擴大平均數逐漸集中。樣組大至十的時候，一〇二四個平均數中有九百十二個集中於三十與七十之間，三十以下和七十以上的只有一百十二個。所以僅僅知道一個次數分配的中心量數是不够的，同時需要另一種量數代表參差的程度。

假設我們要調查上海初中學生的國文程度，中心量數固然可以供給我們很重要的消息，但參差程度也是必需知道的。究竟各校之間是否大有高低？每校各個學生之間是否大有優劣？這些問題如不能回答，徒有一個平均或中位或型範，非但不足以代表全部事實，並且還有引起誤會的危險。我們對於次數分配應當按照次序提出以下的問題：

一、這是甚麼東西的次數分配？
二、這分配包括多少次觀察？（總次數是多少？）
三、這分配的中心量數是甚麼？
四、這分配的參差情形怎樣？

五、這分配的偏斜程度怎樣？
六、這分配的峻峭程度怎樣？

中心量數已經討論過了。偏斜量數與峻峭量數留待以後再講。現在我們討論參差量數。最容易想到的可能參差量數是全體的限域，即最大數與最小數的差別。限域是最簡單最容易算，但可惜不是最可靠的參差量數。因為次數分配兩極端的次數普通總是很少的，偶然的出入是不免的，所以最大數與最小數是非常不穩定，而限域也就不可靠了。例如表七所列身長八千多人中最矮的五十七吋，有二人，最高的七十七吋，也有二人。如果取樣的時候這四人並未遇到，限域就會小二吋。或者另外遇到一二人比五十七吋更矮，或比七十七吋更高，限域就會擴大。此外，有些事實的限域原是固定的，但限域以內的分佈情形可以大有出入。例如表八所列五個次數分配的限域都是從零到一百，但它們的參差程度顯然大不相同，如果採用限域，它們的參差量數不得不認為是相等的了！所以限域不是正確的參差量數，其惟一用處是供給一個迅速而約略的印像。

全體的限域既以兩極端欠穩定而不能成為可靠的參差量數，採用一部份的限域應當可以避免不穩定的缺點。有的統計學家曾經建議在次數分配的兩極端各除去總次數的十分之一，採用中部十分之八的限域。這限域在多數普通的次數分配已經被證明比了全體次數任何其他部份的限域固定可靠，但迄今尚未通行。現在通行的四分差實際上是中部半數限域的一半，即在分配的兩端各除去總次數的四

分之一，然後計算中部半數的限域而取其一半。

計算四分差的第一步是計算上下四分位。上四分位就是中位以上半個次數分配的中位下四分位就是中位以下半個次數分配的中位。四分位是將全體次數分成相等四部的界限位置。一個次數分配有三個四分位，除上下兩個外，中位當然就是中四分位。假設一百人照高矮次序排列，最高二十五人的高度超過上四分位，第二十六人至第五十人的高度在上四分位與中位之間，第五十一人至第七十五人的高度在中位與下四分位之間，最矮二十五人的高度則不及下四分位。

計算上下四分位的方法可從計算中位的方法類推。表十三因同時計算各種中心量數及參差量數，所以有五列。如只需計算上下四分位與四分差，『組中值』與『二次累計』兩列可省。表中的Q_1代表下四分位，Q_3代表上四分位。Q代表四分差，以二除Q_3減去Q_1的差數即得。

表八的五個次數分配是考驗參差量數的很好材料。這五個分配的參差程度顯然大不相同，前面因爲它們的限域是相等的，所以我們斷定限域不是可靠的參差量數。現在再看它們的四分差是否隨着樣組的擴大而縮小。表八的每組次數是完全集中於組中值的，所以計算四分位的方法很簡單。總次數一〇二四的四分之一是二五六，只需從分配的兩端數起，決定第二五六與二五七兩數的平均即得。照這方法計算，五個分配的上下四分位與四分差很快就可以求得：

樣組大小	下四分位	上四分位	四分差

一	〇〇	一〇〇	五〇
二	二五	七五	二五
四	二五	七五	二五
五	四〇	六〇	一〇
十	四〇	六〇	一〇

五個分配的四分差雖然不像全體限域那般完全一律，但也不是五個五樣的，樣組從二擴大到四，或從五擴大到十，從分配表可見參差程度顯然減小，而四分差竟站定於二五或一〇而絲毫不動，所以四分差雖然勝過限域，仍不是很正確的參差量數。

全體限域只根據兩極端，四分差只根據上下四分位，而次數分配的參差情形並不完全決定於兩極端或上下四分位。所以限域與四分差都不是最好的參差量數比四分差更加正確的參差量數應當根據全體的情形，不僅決定於兩點。如果我們以全體的每一數量與平均或中位比較，求其差數，取其平均，其結果應當是很好的參差量數，實際上這參差量數早已被人採用，稱為平均差

平均差有兩種算法，差數可以從平均或中位算起。前面曾講中位地點的好處，說從各處到中位地點的總路程比到任何其他地點為近，所以平均差從中位算最小，應當從中位算，不過實際上也有人從平均算，最好報告平均差的時候標明『離中平均差』或『離均平均差』，以免誤會。本書的平均差除非說明

係離均平均差一律是離中平均差。

從次數分配表計算平均差，每組以組中值爲代表，所以第一步是以各組中值與中位（或平均）比較，求其差數。所有差數當然都一律待遇，不論其爲中位減去組中值或組中值減去中位的結果。第二步是以這些差數與各組次數相乘。第三步是以乘得的積數相加。第四步是以總次數除前一步的結果。這方法的每步理由雖然很容易明白，實際的工作——尤其是第二步——頗費事，遠不如表十三所用的簡捷法。讀者可以就表十三的材料，先照前法將平均差算出，再研究簡捷法，證明其非但方法上比較簡捷，結果也同樣正確。

計算平均差的簡捷法，一部份與計算平均數的簡捷法相同。但計算累計的時候應當採用中位或平均所在的一組做出發點，視所算係離中或離均平均差而決定。中位所在組在計算累計時即可決定，而平均所在組非待累計一列的總和算出不能先知道。這是平均差應當從中位計算的另一理由。表十三的中位是三一二・一〇，所以累計列以『三一二』組爲出發點，自上而下加至八七二，自下而上加至八七四。經過核對後，再將自上而下的累計加起來，得二三八一，自下而上的累計也加起來，得二五七八。計算平均數的時候是從二五七八減去二三八一。現在計算平均差，以二三八一與二五七八相加，得四八五九。這四八五九乘了組距二五之後是全體各數量距離三一二的總差數，如再以總次數二〇〇八除之，所得結果是距離三一二的平均差。但中位不是三一二，而是三一二・一〇，所以尚需加以校正。總次數二〇〇八中有

八七四個的差數，每個應當減去・一〇，其餘一一三四差數應當每個加・一〇，一一三四減八七四得二六〇，所以表十三的M.D.是四八五九乘二五之後再加・一〇乘二六〇，然後再以二〇〇八除得六〇。五一，即是所求的離中平均差。

如所求的是離均平均差，因為平均與中位同是比三一一二略大，所需要的更改是很少的。平均是三一五・七〇，比三一一二大三・一五，所以只需以『三・一五』替代『・一〇』其他完全照舊，算出的結果便是離均平均差了。

平均差必需顧到每一數量，任何數量如有變動，平均差就隨着變動，所以參差程度顯然不同的次數分配決不至有相等的平均差。但平均差有一個缺點，就是計算的時候必需不顧差數的正負符號，假設每一差數的符號都留着，一部份差數是正的，一部份差數是負的，離均平均數一定等於零，離中平均數也並無任何意義了。所以不顧差數的符號是應當的。不過不顧符號之後，這些差數有了一種特殊的地位，不能用普通的數學方法處理，很不便利。類似這樣的缺點是四分差與限域所共有的，中心量數中的中位與型範也不能免。

另外一種計算參差程度的方法是將每一差數自乘，用其平方，因為正負數的平方都是正的，所以不會互相抵消而等於零，而平方是極容易處理的，所以可以避免平均差的缺點。差數自乘之後再加以平均稱為均方差。均方差從平均數計算比從任何其他數量計算為小，因之總是這樣算法。從平均數計算的均

方差的方根稱爲標準差。均方差因爲經過平方的手續，與原本所用的單位不一致，標準差則仍舊回到原本的單位。例如原本材料是長度，用寸爲單位，均方差的單位變爲方寸，標準差再回到寸。

均方差與標準差的計算如不採用表十三的簡捷法是很費事的。每一組中值減去平均數後，加以乘方，再乘各該組的次數，再求其總和，以總次數除之，得均方差。每一組必需做兩次乘法，而所乘的數普通總不免帶小數，因爲平均數往往帶着小數。讀者不妨用表十三的材料先照前法將標準差算出，更可以見得簡捷法的好處。

計算標準差必需先計算平均數。現在假定表十三的平均數已經照以前所講的方法算出，此後的步驟如下：

一、計算『二次累計』。『二次累計』與『累計』一樣算法，『累計』是『次數』的『累計』，『二次累計』是『累計』的累計。不過『二次累計』列比『累計』列短，從上而下和從下而上各少最後一數。所以從上而下二加十六得十八，再加四十得五十八，依次加下去，得一四〇九爲止，『累計』列的最後一數八七二不加進。從下而上也是這樣，得一七〇四爲止，八七四不加進。

二、『二次累計』各數算出後，應當核對一下。『二次累計』與『累計』兩列從上而下的最後一數（一四〇九與八七二）相加，應當等於『累計』從上而下的總和（二二八一）。兩列從下而上的最後一數（一七〇四與八七四）相加，應當等於『累計』從下而上的總和（二五七八）如

有不符合處，應當尋出差誤更正。

三、求『二次累計』列的和數，得六八五四，乘二，得一三七〇八，再加上『累計』全列的和數四八五九，得一八五六七。

四、以總次數二〇〇八除前一步的結果，再減去·一四八的平方。這·一四八是計算平均數的時候求出的，是『累計』列從上而下與從下而上兩半列分別計算的和數的差別（二九七）被總次數（二〇〇八）除所得商數。這步的結果是九·二二四六，是以組距的平方爲單位的均方差。

五、求前步所得九·二二四六的方根，再乘組距二五，得七五·九三，即是標準差，是用原有單位的。

以上所講的計算標準差簡捷法，與前面所講的計算平均數簡捷法，都與普通統計方法書籍中所講的略有不同，但更加簡捷。至於這方法的數理根據，其實並無甚麼奧妙，讀者略加思索與試驗，不難了然。作者曾在別處（教育雜誌第二十六卷第四號）列表說明，必要時亦可供讀者參考。

均方差應當以平均數計算，前面已經講過。凡從任何其他出發點計算的均方差，必然比從平均數計算的大，其差別卽等於該出發點與平均數的差別的平方。這層關係可以用代數方法證明，讀者如有這種興趣，這練習是很值得做的。前面所講計算方法的第四步，有『減去·一四八的平方』，卽根據這層關係。因爲這層關係，採用均方差或標準差有種種便利，例如次數分配經過合倂、分割、或擴大之後，其均方差不必完全重算，原有分配的均方差可加以利用。

所以均方差或標準差既正確，又便利，是應用最廣的參差量數。平均差雖然正確，很不便利。四分差的好處是計算比較省事，且像中位那樣，可以從不完全的次數分配計算。

計算平均差與標準差的時候，各組次數假定其集中於組中值，計算上下四分位以求四分差的時候，四分位所在組的次數假定其均勻分配於組限之內。在對稱與近乎對稱的次數分配這種假定會引起偏的誤差。如組距很大，這種誤差不容忽視。表十四用表十三的材料，先以兩組倂一組，再以三組倂一組，計算標準差，平均差，與四分差，可見所得結果跟着組距擴大。凡近乎對稱的 i 字形分配，組距愈大則所得參差量數亦愈大。如次數分配屬 u 字形，情形正相反，即組距愈大則求得的參差量數愈小。這種誤差的原因讀者不難推想，似不必再加說明。

如次數分配屬 i 字形，分配的兩端次數是逐漸減少的，求得的均方差可加以校正，即減去組距的十二分之一。這校正法是謝伯 (Sheppard) 所貢獻的，稱爲謝伯校正法。現在將表十三與表十四的均方差與標準差加以校正與比較如下：

組距	校正前均方差	校正後均方差	校正前標準差	校正後標準差
二五	五七六五	五七一三	七五・九三	七五・五八
五〇	五八九九	五六九一	七六・八一	七五・四四
七五	六〇八九	五六二一	七八・〇三	七四・九七

在這例子，謝伯的校正似乎過份些，但校正後比校正前正確多了。平均差與四分差不適用這校正法。

現在將三種參差量數的差別表列於後：

標準差	平均差	四分差
英文名 standard deviation	mean deviation	quartile deviation
常用符號 σ	M. D.	Q
從平均計算最小	從中位計算最小	
受全體數值的影響	同	否
應與平均合用	與平均或中位合用	與中位合用
不能從不完全的材料算出	不能	能
便於處理	不便	不便
適用謝伯校正法	不適用	不適用

標準差、平均差、與四分差都是用原有單位的，並不是抽象的數量，因之凡採用不同單位的事實，其參差量數便不能直接比較。這裏所謂單位不同，有兩種可能的情形：一種是可以折合計算的，例如公分與英吋雖然不同，但同屬長度，改用一致的單位後，參差量數就可以直接比較了。一種是事實的性質不同，不能折合計算，例如長度與重量，那是根本無從比較的。

此外，有時兩個次數分配雖然所用單位相同，如直接比較其參差量數，必需十分小心，才可以避免不合理的結論。例如各省政府的全部支出與教育經費當然用同一單位。但全部支出比較大，所以各省之間的參差也大，教育經費只是全部支出的一部份，所以各省之間的參差也小。假定每省的教育經費都佔全部支出的十分之一，教育經費的標準差必然等於全部支出的標準差的十分之一。這『十分之一』是很難解釋的，如果籠統說『各省教育經費的參差程度遠不如其全部支出的參差程度，僅及其十分之一，』那就很容易引起誤會了。爲避免這種誤會起見，應當另作兩種比較：一方面計算每省教育經費佔全部支出的百分比，如一律是百分之十，其標準差或別種參差量數當然都等於零。另一方面計算教育經費的標準差佔平均數的百分比，以與全部支出的標準差佔平均數的百分比互相比較。這百分比稱爲『參差係數』如各省教育經費在全部支出所佔比例是相等的，教育經費的參差係數與全部支出的參差係數當然也是相等的。一定要教育經費的參差係數小於全部支出的參差係數，才可以說『教育經費不如全部支出那樣參差不齊。』

凡以下一類問題的解答都應當採用參差係數：互象與蚊子的重量何者比較參差不齊？從嬰兒發展到成人，身長體重的個別差異是逐漸增加抑逐漸減少？心理方面的個性差別故意未列入前面的問題，因爲心理測量尙未有眞的零點，普通分數只有相對的意義，參差係數是不適用的。

因爲參差係數是抽象的數量，不論原來是尺寸或斤兩，都是一律的，在相當條件下是可以比較的。前

面雖講長度與重量不能比較，其參差係數既然同是百分比，互相比較未必全無意義。因之假設有人說『這班學生的身長似乎比其體重更加參差不齊，』他的印像是否正確，是可以考驗的。

第七講　個體在團體中的地位

凡大小、長短、輕重、高低、多少、優劣、強弱、老幼、快慢一類字的意義是相對的，不是絕對的。我們說小學生年幼，是說他們比成人或比大學生中學生年幼，實際上他們比嬰兒年長得多。我們說某一位運動家跑得快，是說他比普通人或普通運動家跑得快，實際上他決不如馬跑得快。所以這類相對的事實但用文字形容是不够的，應當用數字形容。

數字雖然可以比文字詳細而正確，但其意義未必明了，甚且可以不如文字有意義。例如說『某人跑得快，』確乎不如說『某人跑百公尺需十一秒鐘，』後者比前者詳細而正確多了。但除非我們對於運動有相當常識，那『十一秒鐘』是不很有意義的。對於運動記錄不注意，對於奔跑速度無常識的人，但曉得在十一秒鐘裏跑百公尺，卻不知道這是極容易抑極難，甚且不能決定其爲『快』或『慢』。作者記得有一位生理學教師的分數非常之寬，全班每一學生都在九十九分以上，最優的與最劣的但有小數點以下的分別，如但看分數，誰會想到得到九十九分以上的人竟是全班的倒數第一名？

普通人以爲學生的成績分數是有確定意義的，但因爲試題有難易的不同，批分有寬嚴的差別，『七十五分』究竟算『優』？算『劣，』抑算『平常，』這是因學校，因學科，因教師而異的。百分制度下的分數尙且如此，並無固定制度的測驗分數當然更加不容易解釋。當然，測驗的分數是客觀的，比教師批的主觀

分數正確，但正確與意義是兩件事，趙甲所得的智力測驗分數『二五八』是十分正確的，不過這『二五八』究竟是好，是壞，是怎樣好，是怎樣壞，尙待加以解釋。

所以但說趙甲的成績『並不好』固然不正確，說他的成績是『二五八分』仍乏意義。有意義而正確的形容方法必需說明趙甲在全體所佔的地位。這『全體』的性質當然也必需附帶說明，如『高中畢業生』或『投考某大學的高中畢業生。』前面所講的『在十一秒鐘裏跑百公尺』如能說明其在全世界或全國或全省或全校運動會的地位，也就有意義了。

要知道趙甲成績在投考浙江大學的高中畢業生中所佔地位，當然必需知道全體的成績分配情形，如表五或表十三所示。從表十三，我們知道二五八分在平均以下，或在中位以下。這是很有意義的形容，但尙不够正確，因爲從一百分到三百分都是『在平均以下。』

更進一步的形容是說二五八分『在下四分位以下，』或『靠近下四分位。』這比了但說『在中位以下』當然正確，因爲範圍縮小了。但下四分位以下有全體的四分之一，從一百分到二百六十一分的五百零二人都在這範圍以內。所以仍嫌太籠統。

另一種方法是很容易想到，也很簡單的，即是說明二五八分在全體所佔的名次，例如第一千五百三十三名。名次是很正確詳盡的，不過必需說明全體人數，才有意義。所以趙甲的二五八分應當這樣形容：在全體二〇〇八人中佔第一五三三名。這方法的惟一缺點是必需用兩串數字，一串說明全體人數，一串說

明趙甲名次略嫌不很方便。但這缺點是很容易補救的，用一個百分比就行了。我們不說『在實有的二〇〇八人中佔第一五三三名』，而說『在假定的一百人中佔第七六名』。這確是一種很好的方法，但與實際通行的方法有小小的一點不同。

回到前面採用四分位的方法，『在下四分位以下』已被批評為太籠統，『靠近下四分位』當然也不妥當，因為並未說明靠得怎樣近。根本的缺點是四分位相距太遠，單位太大，不夠細密。但我們既可以用四分位將全體分為相等的四部份，為甚麼不可以用『五分位』『十分位』『百分位』甚至『千分位』分全體為相等的五部份，十部份，百部份，甚至千部份？的確，這些『……分位』非但是可能的，而且是早經採用的，其中百分位是很通行的。

有些一知半解的統計者，以為百分位有一百個，一百人中最小的分數是第一個百分位，最大的分數是第百個百分位。但百分位與名次有根本的差別，非但名次由大而小，百分位由小而大。百分比是將全體分為一百部份的界限，只有九十九個，並無一百個，正像四分位只有三個並無四個一樣。第一個百分位並不是一百個分數中最低的一個，而是最低兩個的分界。同理，第九十九個百分位並不是一百個分數中的最高分數，而是介乎最高次高兩個分數之間。一百個分數中最高分數的百分位不是第一百而是第九十九個半，因為百分位是不佔面積的分界，一個分數可以半個在其上而半個在其下。總之百分位可以這樣定義：凡第幾個百分位之下恰有全體的百分之幾，其餘則在其上。為使讀者對於百分位的意義更加清楚

起見，將二十人，五十人，一百人，二百人中第一名與第十名的百分位等級列下：

全體人數	第一名的百分位等級	第十名的百分位等級
二〇	九七・五	五二・五
五〇	九九	八一
一〇〇	九九・五	九〇・五
二〇〇	九九・七五	九五・二五

名次與百分位等級的關係可以用下列公式表明，式中 x 是名次，N 是總次數或全體人數，R 是百分位等級：

$$R. = \frac{100(N - x + .5)}{N}$$

這裏『百分位等級』與『百分位』的分別必需加以說明。我們知道下四分位即是第二十五個百分位，中位即是第五十個百分位，上四分位即是第七十五個百分位。所以從表十三可以知道：

第二十五個百分位是二六一・一四分

第五十個百分位是三一二・一〇分

第七十五個百分位是三六六・八二分

倒過來，可以說

二六一・一四分是第二五個百分位

三一二・一〇分是第五〇個百分位

三六六・八二分是第七五個百分位

那就等於說

二六一・一四分的百分位等級是二五

三一二・一〇分的百分位等級是五〇

三六六・八二分的百分位等級是七五

根據表十三，我們可以指定求算某一個百分位，求出的是成績分數。我們也可以指定求算某一個成績分數的百分位等級，求出的是百分數。研究整個次數分配的時候需要計算百分位，研究個別成績在全體所佔地位的時候才需要計算百分位等級。計算百分位的方法可以從計算中位與四分位的方法類推，可以寫成下列公式：

$$P_p = l + \frac{pN - F}{f} i$$

式中的P_p是所求的百分位，如所求的是第二五個百分袪則p是・二五，l是所求百分位所在組的下限，

F是該組下限即l以下的總次數，f是該組以內的次數，i是該組的組距，N是全體的總次數。

計算百分位等級的方法只需將計算百分位的方法倒轉過來。現在假定要從表十三計算二五八分的百分等級，即計算全體的百分之幾在二五八分以下。在算出這百分比以前，必需先行計算二五八分以下的次數共有多少。二五八分所在組的下限是二四九・五，從表十三的累計列知道二四九・五分以下的成績共有四〇一次，從次數列知道在二四九・五與二七四・五分之間的成績有二一七次。這二一七次假定其均勻分配於全組之內，一部份在二五八分以下，一部份在其上。因為組距是二五，二五八分減去下限二四九・五分得八・五，所以用二五除八・五，再乘二一七，所得結果七三・七八，即是該一組內在二五八分以下的次數。以這次數加於該組以下的次數四〇一，得四七四・七八，即是成績在二五八分以下的總次數。再算四七四・七八在全體總次數二〇〇八的百分比，得二三・六四，即是二五八分的百分位等級。這計算方法可以寫成如下的公式。

$$R_x = \frac{100}{N}\left(F + \frac{x-l}{i} f\right)$$

式中x是原有分數，其百分等級是R_x，即所需要計算者，N是全體總次數，l是x所在組的下限，F是l以下的總次數，f是x所在組以內的次數，i是該組的組距。

如有多數成績的百分位等級需要計算，或者計劃以每五分或每十分的百分位等級一律求出，以為

解釋任何分數的參考，最便利的辦法是將各組限的百分位等級先行算出，因爲那只需把累計的次數折合全體總次數的百分比就夠了。至於其他分數的百分位等級，按照其所在組內距離上下限的比例插入，也是很簡單的。表十三各組限的百分位等級列於表十五。

從表十五可見百分位等級的增減與原有成績分數的增減是並無固定比例的。成績分數每行一律增加二十五分，百分位等級則每行增加多少不一，少的不到一級，多的有十餘級。例如成績分數從二四・五增加到一四九・五，百分位等級從・一〇增加到・八〇，所加僅・七〇，而成績分數從二九九・五增加到三二四・五，百分位等級從四三・四三增加到五六・四七，所加有一三・〇四之多。多數的次數分配既是近乎對稱的i字形分配，百分位等級的增減大概總在中央快而在兩端慢。所以一個學生的成績從第四五進步到第五〇個百分位是比較容易的，而從第五進步到第一〇個百分位，或從第九〇到第九五個百分位是非常之不容易的。同理，一個在第五一個百分位的學生比了在第四九個百分位的學生是所優極微的，另外一個在第九九個百分位的學生比了在第九七個百分位的卻好得多了。

前面的討論有兩種應用。第一，一切競賽得勝的分數或獎品普通往往並不公平合理。例如第一名得三分，第二名得二分，第三名得一分，或第一名得五分，第二名得三分，第三名得一分，都假定第一第二兩名之間的差別與第二第三兩名之間的差別是相等的。實際上前者所差比後者所差爲多，所以第一名所得分數應當增加。

第二，百分位等級間的距離既不相等，百分位等級不可用以加減或計算平均。例如國語成績的百分位等級是九九，算術成績的百分位等級是六九，二者的平均是八四，但這八四是無意義的，實際上兩種成績平均的百分位等級決不止八四，或者是九〇或九一，也未可知，應當從原有成績分數求其平均，再從平均數的次數分配計算百分位等級。總之，平均數的百分位等級並不是百分位等級的平均數。

百分位等級的意義因爲與名次相彷，所以大衆容易了解，應用很廣。但等級之間的距離不一律，不能加減，不能計算平均，這是一大缺點。下面所講標準分數的好處就是可以免除這缺點。

前面已經講過，但說某人的成績在平均以下或平均以上是不夠正確的。若說平均以上或以下多少分，因爲分數的意義不確定，全體的參差程度未說明，仍不能表達其在全體的正確地位。必須不用原有分數而用參差量數，才是最正確而有意義的方法。從原有分數減去平均分數，再以標準差除之，結果便是標準分數。『標準分數二』等於『在平均以上二個標準差』。因爲全體的限域普通不過六個標準差，卽從平均以下三個標準差到平均以上三個標準差，所以標準分數的單位是很大的，非算到小數不夠正確。標準分數非但常帶小數，並且有一半時候是負的，負的標準分數當然比平均分數小。

標準分數既是從原有分數減平均分數再被標準差除的結果，全體標準分數的平均必然等於零，其標準差必然等於一。但標準分數常有小數與負數，不很簡潔，且不易推行，所以應用的時候不妨略加變通。一種普通的變通辦法是採用標準分數的十倍再加五十。這種分數的全體平均當然等於五十，其標準差

等於十，從零到一百相當於平均以下五個標準差到平均以上五個標準差，因爲與常用的百分記分法類似，易於推行。但這種分數有時仍不免用小數，因之有人採用標準分數的百倍，再加五百。

標準分數與百分位等級都是參照分配情形以說明個體在團體中所佔地位的方法，對於學生的成績分數最有用處。如要比較兩次考試或兩種學科的個別成績分數，採用標準分數或百分位等級，可以不受試題難易或批分寬嚴或單位不同的影響。

標準分數之間的單位是一律的，所以比百分位等級易於比較，且不妨計算平均。不過標準分數的平均數也不一定等於平均數的標準分數，這兩種算法的結果如有差別，標準分數的平均比平均的標準分數合理。因爲標準分數總以標準差爲單位，計算其平均是毫無問題的，普通分數的參差程度未必相等，計算平均的時候，參差程度高的分數太佔勢力，算出的平均往往是不公平的。

假設我們要計算國語與算術的平均分數，規定兩種學科不分輕重。但國語教師所批分數的參差量數低，算術教師所批分數的參差量數高一倍。下列學生的國語與算術的成績恰好相反，如果兩種學科不分輕重，他們的兩科平均分數應當是相等的，但結果並不這樣，平均分數的次序完全與算術分數的次序一樣，可見平均分數中算術分數所佔勢力比國語成績大。

學生	國語分數	算術分數	平均分數
甲	九〇	七四	八二

乙、	八八	七八	八三
丙	八六	八二	八四
丁	八四	八六	八五
戊	八二	九〇	八六

如果先算國語與算術的標準分數，因爲標準分數的參差程度是一律的，兩種學科在平均時所佔勢力才眞是相等。現在假定國語分數的平均是八四，標準差是四，算術分數的平均是七八，標準差是八，兩種學科的標準分數如下：

學生	國語標準分數	算術標準分數	平均標準分數
甲	一·五	負·五	·五
乙	一·〇	·〇	·五
丙	·五	·五	·五
丁	·〇	一·〇	·五
戊	負·五	一·五	·五

第八講　次數分配的形式——偏斜度與峻峭度

前面講次數分配有i字形、j字形、u字形三大類，但它們究竟怎樣分別？是否全憑視覺印像？三大類是否夠精細？次數分配的形狀除了分為幾類，用文字敍述其不同外可否用數字敍述其差別？

中心量數是決定次數分配圖的位置的，中心量數如有增減次數分配圖應當向右或向左移動，毫不牽涉次數分配的形式。參差量數是決定次數分配圖左右展開的程度的，參差量數如有增減，次數分配圖應當放寬或收狹，總次數是決定次數分配圖所佔面積的，如總次數有增減次數分配圖應當增高或減低。放寬或收狹增高或減低，只是改變縱橫的比例，並不眞正影響次數分配圖的形式。所以辨別分配形式的量數尚待另外尋找。

我們曾經一再提到次數分配的偏斜程度。有的次數分配是絕對對稱的（例如表八，）有的是略帶偏斜而近乎對稱的（例如表十三，）有的是顯然偏斜的（例如表六表十二，）有的是極端偏斜的（例如表十一。）偏斜程度應當可以被綿續不斷的數量所形容，每一次數分配應當可以算出其偏斜量數。

有一種偏斜量數是根據四分位計算的，以上四分位與中位間的距離，和中位與下四分位間的距離比較，如二者相等偏斜量數等於零，如前者大於後者，所得偏斜量數是正的，如後者大於前者所得偏斜量數是負的。所以正的偏斜是型範偏左右尾較長，負的偏斜是型範偏右左尾較長，但兩距離的差數尚不可

直接用做偏斜量數，因爲差數的大小並不完全決定於偏斜的程度，一部份決定於參差的程度，應當用兩距離的和數除其差數，所得商數才是適當的參差量數。這參差量數必然在負一與正一的限域之內，是抽象的數量。

茲以表十三的材料爲例，計算偏斜量數如下：

$$Q_3 = 366.82 \qquad\qquad Mdn = 312.10$$
$$Mdn = 312.10 \qquad\qquad Q_1 = 261.14$$
$$Q_3 - Mdn = 54.72 \qquad\qquad Mdn - Q_1 = 50.96$$

$$\frac{54.72 - 50.96}{54.72 + 50.96} = \frac{3.76}{105.68} = .036$$

讀者如照這方法計算表六材料的偏斜量數，可得・〇四九，表十二的偏斜程度似乎不在表六之下，但其偏斜量數僅・〇〇八，反比表十三爲小。表十一的偏斜量數是・五七一，可見・五所代表的偏斜程度是很顯著的。

另外一種計算偏斜程度的方法是根據平均、型範、與標準差，即以平均減型範，再以標準差除之。如用表十三的材料計算：

$$M = 315.70 \qquad Mo = 312.00 \qquad \sigma = 75.93$$
$$315.70 - 312.00 = 3.70 \qquad 3.70 \div 75.93 = .049$$

大概用這方法計算的偏斜量數比前面所講的偏斜量數大，所以這兩種偏斜量數的意義是不同的。為便利辨別起見，前者不妨稱為『四分偏斜度』後者不妨稱為『標準偏斜度』。茲將表八表十三，表十二表六，與表十一的兩種偏斜量數列下，讀者可以同時參閱各表。

表次	四分偏斜度	標準偏斜度
表八	.○○○	.○○○
表十三	.○三六	.○四九
表十二	.○○八	.一三三
表六	.○四九	.一八八
表十一	.五七一	.二四五

憑我們對於各表的視覺印像，表十二的四分偏斜度似乎太小，不如標準偏斜度足以充分代表事實。除了表十一之外，標準偏斜度總比四分偏斜度大。按照四分偏斜度，表十一比表六的偏斜度不止十倍，但照標準偏斜度則前者不及後者兩倍。考其原因，大概因為表十一的標準差非常之大。茲將該表標準偏斜度的計算根據列下，以供參考：

平均	七○．七六
型範	一．○○

標準差　二八五·

表十二的平均數與標準差當然不可以完全用組距千計算；上列數字是照原表用一、十、百、千四種組距分段計算後再合併的。為讀者便於練習時核對起見，茲將每段的『次數乘平均』與『次數乘（從本段平均計算的）均方差』列左：

段	組距	次數	次數乘平均	次數乘均方差
一	一	二一〇五	七，七五五	一五，八八一
二	十	一六一〇	五七，七四五	八四九，〇八四
三	百	五一四	一四二，五五七	一七，六四九，五〇〇
四	千	四〇	九四，〇二〇	八七，〇〇〇，〇〇〇

計算偏斜量數的需要並不像計算中心量數與參差量數那樣多，因為普通對於中心所在與參差程度必需正確數字，而對於偏斜情形直接從次數分配表或圖得一約略印像就夠了。惟一需要計算正確偏斜量數的時候是從實得次數分配材料配合一個可以用數學方程式正確表達的理論次數分配曲線。配合次數分配曲線的理論與方法很是複雜，本書不能討論。這裏所要說明的是，前面所講的四分偏斜度與標準偏斜度都不是很好的偏斜量數，不能據以配合次數分配曲線。

配合次數分配曲線時偏斜度普通根據第三矩計算。在說明計算方法以前，我們應當先把這『矩』

字解釋清楚。先規定一個起點，計算全體每一數值減這起點的差數，正負符號留着，求差數的總和，以總次數除之，結果便是第一矩。如用零做起點，任何次數分配的第一矩必然等於平均數，如用次數分配的平均數做起點，其第一矩必然等於零。以平均數做起點的第二矩就是均方差，因為第二矩與第一矩不同的地方就是不用每一差數而用其二次方。所以第一矩可以當做中心量數，第二矩可以當做參差量數。第三矩卽是各差數的三次方的平均，以平均做起點的三次方是有正有負的，平均的正負卽代表偏斜的方向，如次數分配完全對稱，各差數的三次方必然正負抵消，恰等於零。

除第一矩應當以零為起點外，較高各矩一概以平均數為起點，所以第二、三、四矩凡起點不加說明時卽知其以平均數為起點。第三矩可根據以計算偏斜度，但不可直接採用，因為第三矩的單位是統計材料原有單位的三次方，偏斜度決不應當用立方尺一類的具體東西表達，去除原有單位三次方的方法如下：

$$\beta_1=\frac{\mu_3^2}{\mu_2^3}=\frac{\left(\frac{\Sigma x^3}{N}\right)^2}{\left(\frac{\Sigma x^2}{N}\right)^3}\quad 或\quad \gamma_1=\frac{\mu_3}{\sqrt{\mu_2^3}}=\frac{\frac{\Sigma x^3}{N}}{\sqrt{\left(\frac{\Sigma x^2}{N}\right)^3}}=\frac{\frac{\Sigma x^3}{N}}{\sigma^3}$$

式中x是全體每一數值減平均的差數，Σ指示求總和，N是總次數，μ_3是第三矩，μ_2是第二矩。β_1與γ_1都是偏斜量數，γ_1卽是用標準差的三次方除第三矩的結果，β_1卽是γ_1的平方，所以二者實際並無甚麼差別，但γ_1有正負之分，非但指示偏斜的程度，也指示偏斜的方向，所以比β_1有用。第三矩的計算雖然並不困難，卻

相當費事，而且需要也不多，所以本書從略了。

爲甚麼有些次數分配是對稱的，而有些是偏斜的？學生的成績分數是否應當成對稱分配？如果試題難易適中，學生也未經選擇，而各試題不是互相關連的，學生的成績分數大概會成爲對稱分配。如果試題太難，有好些學生得零分，或試題太易，有好些學生得足分，成績的分配勢必偏斜。如果學生曾經甄別，或者優等生已經升學，或者劣等生已經退學，成績的分配也會偏斜。又如試題互相關連，有一對則都對，一誤則都誤的趨勢，總分數的分配也非偏斜不可。所以發現學生成績分數成爲偏斜分配的時候，應當研究其原因，或許是應有的現象，不可武斷對稱分配爲合理，偏斜分配爲不合理。

次數分配的形式並不完全決定於偏斜度，因爲偏斜程度相等的次數分配未必就是形式相等的。表八所列的五個次數分配雖然一律對稱，其形式大不相同。它們相差的地方統計學家稱爲峻峭度。峻峭度高的次數分配有陡的高峯與一端或兩端的長尾，峻峭度低的次數分配是比較平頂的，甚且是u字形的。峻峭量數有大同小異的兩種，都根據第四矩計算：

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{\frac{\Sigma x^4}{N}}{\left(\frac{\Sigma x^2}{N}\right)^2} \qquad \gamma_2=\beta_2-3$$

計算峻峭量數的需要也不多，所以本書對於計算方法從略。但表八所列的分配都是很齊整的，所以

極容易計算，結果如下：

樣組	第二矩	第四矩	β_2
一	一二五〇〇	六二五〇〇〇〇	一・〇〇
二	一二五〇	三一二五〇〇〇	二・〇〇
四	六二五	九七六五六二・五	二・五〇
五	五〇〇	六五〇〇〇〇	二・六〇
十	二五〇	一七五〇〇〇	二・八〇

如算γ_2，都得負數。γ_2以β_2減三，初看似無理由。但γ_2等於零（即β_2等於三）的次數分配稱爲有「適中的峻峭度」，正的γ_2（即超過三的β_2）代表「陡峻」的分配，負的γ_2（即不及三的β_2）代表「平頂」的分配。照這定義，表八所列五個次數分配都是「平頂」的。

學生成績有一定的最低分數與最高分數，不能超出這限域，而且試題有限，普通雖用百分法，實際並不能分出一百等級。這種次數分配大概是「平頂」的。試題多而限域廣的測驗分數則成爲峻峭度比較適中的次數分配。至於「陡峻」的次數分配往往有一端（或兩端）是無限制的。極端偏斜的分配大概同時是「陡峻」的。

「陡峻」的次數分配往往代表一種不穩定的情形。例如非常時期的物價指數，疲倦時期的反應時

間，其次數分配中常有少數指數或時間遠在中心量數以上。這種情形一方面增高次數分配的峻峭度，一方面指示材料有不穩定的趨勢。因為非常時期有一部份物品非但漲價遠在平均以上，更有市上絕跡無從購買之勢，疲倦時期的反應非但遲緩，更有注意力分散未能察覺號令之勢。無從購買的物品，其物價指數可以說是無限的，不察號令而未有反應，其反應時間也可以說是無限的。無限的數值當然不會在實得次數分配中直接遇到，但不穩定的情形可以從峻峭度間接表現。

偏斜度與峻峭度確定之後，次數分配的形式也就決定了。前面所講配合次數分配曲線，就是根據偏斜度與峻峭度選定曲線的函數形式，再按照中心量數，參差量數與總次數推算曲線方程式中的常數，即決定曲線的位置，廣狹，與高低。

如總次數很小，算出的偏斜度與峻峭度是很不正確的。總次數小到三的次數分配，其β_2總是等於一又二分之一。總次數小到二的次數分配總是對稱的，所以總次數等於三的分配，其峻峭度是無意義的，總次數等於二的分配，其偏斜度是無意義的。換句話講，總次數至少有三，才可以計算偏斜度，至少有四，才可以計算峻峭度。這種情形正像單個數值無從計算參差量數一樣。

第九講　二項分配與常態分配

我們以前所引用的次數分配，除了表八所列的幾個，都是實際觀察所得的材料。但表八是完全根據理論推算的：假定全體學生一半得零分，一半得百分，則樣組的平均分數應當成爲表八所列的次數分配。其實表八所列的分配，就是現在要講的二項分配。

假設投擲硬幣得到正面或反面的機會各佔一半，如投擲幾次，每次都一樣，累積的可能結果如下：

一次之後　正　或　反

二次之後　正正　或　正反　或　反正　或　反反

三次之後　正正正　或　正正反　或　正反正　或　正反反　或　反正正　或　反正反　或　反反正　或　反反反

這樣下去，投擲一次有兩種可能的結果，兩次有四種可能的結果，三次有八種可能的結果，以至四次五次，有十六種三十二種的可能結果。這些可能結果的發生機會是一律相等的。

如果只問正反面的次數，不問發生的次序，例如『正反』與『反正』同是一正一反，可併在一起，可能的結果如下：

一次之後　『正』佔二分之一，『反』佔二分之一。

二次之後『正正』佔四分之一，『正反』佔二分之一，『反反』佔四分之一。

三次之後，『正正正』佔八分之一，『正正反』佔八分之三，『正反反』佔八分之三，『反反反』佔八分之一。

但正面與反面不必並計，如只計正面可能的結果可以列成簡單的次數分配表如下：

投擲次數	零次得正	一次得正	二次得正	三次得正
一次	二分之一	二分之一		
二次	四分之一	四分之二	四分之一	
三次	八分之一	八分之三	八分之三	八分之一

假設所擲不是硬幣而是骰子，每次有六種可能的結果，二次有三十六種，三次有二百十六種，太複雜了。但我們可以只問是否得六點，得六點算成功，其餘都算失敗，這樣每次成功的機會是六分之一，失敗的機會是六分之五。如只計成功的次數，可能的結果如下：

投擲次數	零次成功	一次成功	二次成功	三次成功
一次	六分之五	六分之一		
二次	卅六分之廿五	卅六分之十	卅六分之一	
三次	二一六分之一二五	二一六分之七五	二一六分之一五	二一六分之一

讀者如果記得初等代數裏的二項展開公式，我們可以用p代表成功的機會，q代表失敗的機會，前面所講的可能結果都可以按照公式推算，即：

$$(q+p)^1=q+p=\frac{1}{2}+\frac{1}{2} \text{ 或 } \frac{5}{6}+\frac{1}{6}$$

$$(q+p)^2=q^2+2qp+p^2=\frac{1}{4}+\frac{2}{4}+\frac{1}{4} \text{ 或 } \frac{25}{36}+\frac{10}{36}+\frac{1}{36}$$

$$(q+p)^3=q^3+3q^2p+3qp^2+p^3=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8} \text{ 或 } \frac{125}{216}+\frac{75}{216}+\frac{15}{216}+\frac{1}{216}$$

或者更概括些，可以用n代表投擲次數，寫成普遍適用的公式如下：

$$(q+p)^n=q^n+nq^{n-1}p+\frac{n(n-1)}{2}q^{n-2}p^2+\frac{n(n-1)(n-2)}{2\cdot3}q^{n-3}p^3+\cdots\cdots$$

如n超過三，以後各項可以用類推的方法寫下去，應當並無困難。這公式非但包括前面用做例子的六個分配，更包括無數個可能的分配。凡是可以用前面公式推算的無數個可能分配都是二項分配，所以二項分配是有很多可能變化的，視公式中的p，q與n而定。

照前面的公式，每一項的數值是一個比率，總和是一，所以分配的面積倒是固定的。

二項分配的平均等於n與p的積，所以一方面跟着n增減，一方面跟着p消長。這是當然的：每次的

成功機會愈高，平均的成功次數愈多。總共嘗試（投擲）的次數愈多，平均的成功次數愈多。

二項分配的均方差等於npq的積。均方差與n成正比例是容易了解的，因爲嘗試的次數愈多，分配的限域當然愈廣，任何參差量數都就增加了。至於p與q，因爲二者的和數必然是一，p大則q小，p小則q大。但p與q相等，即各等於二分之一的時候，積數最大，相差愈多，則積數愈小。

如二項分配的p與q相等，分配是對稱的。可見凡n相同的二項分配，偏斜愈甚則參差愈小。至於偏斜度，一方面跟着p與q的差別增減，一方面與n也有關係。因爲二項分配的偏斜度是照下式計算的：

$$\gamma_1=\frac{q-p}{\sigma}=\frac{q-p}{\sqrt{npq}}$$

所以即使p與q的差別很大，只要n增加，偏斜度也會減低。茲以·○二爲p，以·九八爲q，算出γ_1如下：

n	一	一○○	四○○	九○○	二五○○
γ_1	六·九	·六九	·三四	·二三	·一三七

所以不論p與q怎樣不等，二項分配是跟着n的增加而趨向對稱的。

二項分配的峻峭度可以用下式計算：

$$\gamma_2=\frac{1-6pq}{npq}$$

於此可見凡p與q的積等於六分之一的二項分配是峻峭度適中的。積數較大，即比較近乎對稱的二項分配是『平頂』的。積數較小，即更加偏斜的二項分配是『陡峻』的。但n逐漸增加的時候，『平頂』的和『陡峻』的二項分配都逐漸趨向於適中的峭峻度。試看下列γ_2：

n	pq相等	pq相差·六	pq相差·九六
一	負二	·二五	四五·〇二
二	負一	·一二五	二二·五一
十	負·二	·〇二五	四·五〇二
五十	負·〇四	·〇〇五	·九〇〇四
百	負·〇二	·〇〇二五	·四五〇二
千	負·〇〇二	·〇〇〇二五	·〇四五〇二

照前面的討論，可見二項分配雖然有很多可能的變化，但也有固定的趨向。n逐漸增加，二項分配逐漸趨向對稱與適中的峻峭度，即γ_1與γ_2都趨向零。所以對稱與適中的峻峭度似乎是二項分配的普遍形式，假定n大到相當的程度。

我們在第二講曾談到某省有小學生四十二萬九千人，其中二十二萬一千是男生，二十萬八千是女生，說如要考驗男女生的差別究竟是否偶然的，抑是重男輕女的表現，必需推演投擲硬幣四十二萬九千

次或四十二萬九千枚的可能結果看正反面相差一萬三千的結果是否容易得到。這推演在理論上並無困難，只需按照二項展開的公式將

$$\left(\frac{1}{2}+\frac{1}{2}\right)^{429000}$$

的每項寫下就行了。但實際上所需要的工作非常繁重，所以二項分配除非可以簡單化是很不便於應用的。

這問題數學家已替我們解決了。他們從前面的二項展開公式出發，假設n很大，可以得到所謂常態分配的公式如下：

$$y=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$$

式中的π是圓周率三・一四一六，e是自然對數之底二・七一八二八，σ是分配的標準差，x是從平均出發的距離，y是分配曲線的高度。讀者有了這公式可以推算常態分配曲線每一點的高度，但我們用不着自己推算，因為早已有人詳細算好了。表十六列示常態曲線每距標準差十分之一的高度，與曲線下左右面積。這表以平均（亦即中位與型範）為出發點，所列是分配曲線的右半，因為常態曲線是對稱的，左半即是右半的對面。如將距離從向右改為向左，第一列每行加一負號，第二列高度完全照舊，再將第三第四兩列面積對換，即可適用於曲線的左半了。

根據表十六我們可以畫成圖三。但圖三同時將 $(.5+.5)^4$ 的二項分配也畫出，以資比較。這二項分配的多邊形是從七點連接畫成的，這七點的高度與常態曲線相差如下：

從平均出發的距離	常態曲線高度	$(.5+.5)^4$多邊形高度
零	.三九八九四	.三七五〇
向左或右一標準差	.二四一九七	.二五〇〇
向左或右二標準差	.〇五三九九	.〇六二五
向左或右三標準差	.〇〇四四三	.〇〇〇〇

常態分配是平滑曲線，兩端是無限的但在四個標準差以外離底線太近，無法在圖上畫出。其實在三個標準差以外的面積，每端只有全體的萬分之二十七，已經不很重要，四個標準差以外的面積每端只有全體的十萬分之六，更加不足輕重了。因為實得的次數分配有很多近似常態分配，所以普通不妨假定次數分配的全體限域約佔六個標準差。

常態曲線的主要特性是對稱的形式，所以平均、中位、與型範都相等。不謹慎的作家有時將『常態分配』與『對稱分配』兩個名詞互相代用，忘卻了峻峭度與偏斜度是同樣重要的。常態分配雖然必是對稱的，而對稱分配未必是常態的。任何平頂的或陡峻的分配可以是對稱的，但它們決不是常態的。常態分配必需有適中的峻峭度，其 γ_2 與 γ_1 同樣必需是零。

在常態分配，三種參差量數有一定的關係如下：

$$Q = .84535\ M.D. = .67449\,\sigma$$

$$M.D. = 1.1829\,Q = .79788\,\sigma$$

$$\sigma = 1.4826\,Q = 1.2533\ M.D.$$

因爲很多次數分配近似常態分配，以上的比例也約略適用於多數分配。

因爲很多次數分配近似常態分配，任何分配如果缺乏不是常態分配的證據，往往就被假定爲近似常態。這種假定當然不能萬無一失，但總比任何其他假定爲合理。根據這假定，前面所講的標準分數與百分位等級可以按照表十六的常態分配面積互相轉化。例如標準分數一相當於百分位等級八四，標準分數負二相當於百分位等級二，百分位等級一相當於標準分數負二．三，百分位等級九〇相當於標準分數一．三。

如果百分位等級可以轉化爲標準分數，凡是有秩序的品質材料都可以改成數量材料。因爲有秩序的品質可以計算百分位等級，而百分位等級可以轉化爲標準分數。表四的成績等級原是從表五的成績分數改的。假設我們不知其原有分數，而有改成分數的需要，從表四各等級的百分比即可以求得百分位等級，再查表十六，即可以改成標準分數。投考者全體的成績比較近乎常態分配，這方法不妨一試。結果如下。

等級	百分位等級	標準分數	不用負數及小數
優	九八·九	二·三	七三
良	九一·九	一·四	六四
常	七一·三	·六	五六
可	三八·二	負·三	四七
劣	一一·〇	負一·二	三八
不及格	一·〇	負二·三	二七

至於錄取的成績，因爲已經經過一番淘汰，不能假定爲常態分配，而且並無不及格者，所以不能用以計算各等級的標準分數。

有人主張學生成績規定各等級的百分比，也是根據常態分配。最普通的比率似乎是以下兩種：

	甲等	乙等	丙等	丁等	戊等
第一種百分比	六	二二	四四	二二	六
第二種百分比	五	二〇	五〇	二〇	五

這種方法只能適於人數多而中常的班級，如學生不多或特別優秀而努力，非大加變通不可。

美國人麥柯爾（W. McCall）所創用而曾在中國風行一時的T分數也完全根據常態分配。照麥

氏的理論，寧信常態分配而不可信測驗的原有分數。所以原有分數摒棄不用，只據以決定百分位等級，然後按照常態分配，將百分位等級改爲標準分數。

除了前面所講的應用，常態分配尙有一種更加不成問題的用處。那就是在估計誤差，考驗重要性的時候。

第十講　取樣誤差的估計與重要性的考驗

在第二講與第九講曾提到過一個問題：某省有小學生四十二萬九千人，其中二十二萬一千是男生，二十萬八千是女生，女生比男生少一萬三千。有人說這是重男輕女的表現，有人說這是不足輕重的偶然差別。這爭論不能解決，請教於統計學家，問他投擲硬幣四十二萬九千次，所得正面與反面會不會相差一萬三千次。統計學家說結果可以從推演 $(.5+.5)^{429,000}$ 得到，推演並不費事，因爲這四十多萬次的二項分配實際上不能與常態分配分別，只需根據常態分配計算，非常簡單。但有一個先決問題：投擲硬幣四十二萬九千枚當然有很多很多（寫起來必需用十二萬個以上的零）可能結果，這很多可能結果中，有一種是四十二萬九千枚全得正面，有一種是全得反面，所以正反面相差一萬三千次當然是可能的。但我們決不能因其可能就承認或接受偶然之說，因爲全是正面或全是反面也屬可能，假設那四十二萬九千個全是男生，難道我們也說那是偶然的結果嗎？

讀者一定要說，四十二萬九千枚硬幣全得正面或全得反面雖屬可能，卻非常不容易，如果實際得到這種結果，我們也會懷疑硬幣的兩面輕重不等，不能接受偶然之說。

所以我們必需規定範圍，如實得結果在這範圍以內，我們承認其與偶然之說並無重要的不符合處，可以接受偶然之說，如所得結果超出這範圍，不得不認爲事實與假設有重要的不符合處，只得拒絕假設

了。

以實得結果與偶然之說的可能結果比較，在統計術語稱爲『重要性的考驗』。如實得結果在偶然之說下雖屬可能而發生的機會很小，事實與理論便有『重要的不符合處』，或者簡單些，即說實得結果是『重要的』，足夠據以拒絕偶然之說或別種假設的理論。現在的問題是，實得結果在偶然之說下發生的機會必需小到甚麼程度才可以稱爲『重要』？這問題的答案多少是武斷地決定的，大概總在二十分之一與千分之一之間。這比率稱爲『重要性層級』，二十分之一是最低的重要性層級，千分之一是最高的重要性層級，普通採用百分之一。

我們現在假定採用百分之一的重要性層級，來考驗事實（即四十二萬九千小學生中有二十二萬一千是男生，二十萬八千是女生）與偶然之說是否有重要的不符合處。偶然之說下的可能結果是$(.5+.5)^{429,000}$的二項分配，也就是平均等於二十一萬四千五百，標準差等於三二七·五的常態分配。實得結果與理論的平均相差六千五百，等於標準差的一九·八倍，在表十六已經查不到，但從表十六可知距離平均四·四個標準差以外，常態曲線每端的面積不過十萬分之一，所以早已超過百分之一的重要性層級，偶然之說不得不拒絕了。

前面所講是重要性考驗的基本理論與方法。第一步是從假設的理論（如偶然之說）推演所有的可能結果。常態曲線是最常用以代表這種可能結果的，但有時必需採用別種分配。第二步是預先規定重

要性層級，即規定實得結果如超越某種範圍即認爲與假設的理論有重要的不符合處。規定重要性層級即是在常態分配或別種理論分配的兩端劃定界限，界限以內的實得結果是不重要的，以外的是重要的。第三步是以實得結果與假設的可能結果比較，觀察其是否在規定的界限以內。第四步是決定接受或拒絕假設的理論。

爲使重要性考驗的方法更加清楚起見，我們更舉一個簡單的例子。甲乙兩校賽球八次，都是甲勝乙負。但乙校學生不肯承認甲校球隊能力確比乙校高超，說那是偶然的運氣。甲校學生不服，要把偶然運氣的理論考驗一下。兩校學生公決採用百分之一的重要性層級。照偶然運氣的理論，每次勝負的機會應當相等，八次的勝負可能結果可從 $(.5+.5)^8$ 推演如下：

$$(.5+.5)^8=\frac{1}{256}+\frac{8}{256}+\frac{28}{256}+\frac{56}{256}+\frac{70}{256}+\frac{56}{256}+\frac{28}{256}+\frac{8}{256}+\frac{1}{256}$$

即八次全勝的機會佔二五六分之一，八次全負的機會也佔二五六分之一，所以八次勝負完全一致的機會是一二八分之一，是小於百分之一。所以這結果是重要的，乙校學生的偶然運氣之說就被推翻了。

這例子所需的二項分配只有八次方，所以不難推演，不必借用常態分配。而且這二項分配是「平頂」的，γ_2 等於負·二五，形式與常態分配略有不同，借用常態分配恐怕不能得很正確的結果。如借用常態分配，其平均是四，標準差是二的方根，即一·四一四，實得結果零或八，因爲本身必需算做半個不足半個超

過，應當照·五或七·五算，與平均相差三·五，等於標準差的二·四八倍。查表十六，二·五個標準差以外的面積每端有千分之六兩端有千分之一二，大於百分之一。所以如照常態分配計算甲校學生只得接受乙校學生的偶然運氣之說了。

照前面的例子，常態分配是不大適用的，但二項分配只有八次方的時候並不難算，根本不必借用常態分配。如二項分配計算太難，必需借用常態分配，二項分配的n一定比較大，形式就近似常態分配，借用也就毫無問題了。

前面兩個例子裏我們所考驗的是實得次數（男女的次數，正反面的次數，勝負的次數）與理論次數的差別。理論的可能結果分配的標準差也稱爲標準誤差，在這兩個例子是次數的標準誤差。次數的標準誤差算法與二項分配的標準差一樣的，所以

$$\sigma_f=\sqrt{npq}$$

因爲n大到相等程度後二項分配與常態分配差別很少，所以普通借用常態分配的時候很多。表十六所列面積並無二十分之一，百分之一，與千分之一，現在另從詳細的表上查出相當於這三種重要性層級的標準差距離如下：

重要性層級	離平均的標準差數
·〇五	一·九六〇

.〇一	二.五七六
.〇〇一	三.二九一

以上的關係既是固定的，考驗重要性的時候可以直接採用從平均的距離做標準，如實得結果與理論平均相差超過二個標準差，即可稱爲『重要』如超過三個標準差，可稱爲『很重要』。在適用常態分配的問題，所考驗的即是事實與理論相差是否有標準誤差的兩倍或三倍。

除了次數以外，平均數，均方差，偏斜度，峻峭度都可以考驗，而平均數的考驗需要最多。平均數的標準誤差是這樣算的：

$$\sigma_M=\frac{\sigma}{\sqrt{N}}$$

所以平均數的標準誤差一方面與次數分配的標準差成正比例，一方面與總次數的方根成反比例。這公式的正確性可以從表八看出。全體學生一半得零分，一半得百分，其正確的平均是五十分，標準差也是五十分。現在從全體學生取樣，如每次抽取一人，或得零分，或得百分，機會相等，所以取樣的標準誤差即等於原來全體的標準差。如每次抽取四人，有五種可能的平均數：得五十分的機會佔八分之三，得二十五分或七十五分的機會各佔四分之一，得零分或百分的機會各佔十六分之一，所以平均數的標準誤差是二十五分。樣組增加四倍，取樣誤差減小一半。至於每次抽取五人十人或更多人的結果，讀者可以按照二項展

關公式自行推算，這裏不多講了。

在這個例子，可能的平均數成爲二項分配，所以樣組很小的時候不宜借用常態分配。但我們應當注意，在這個例子，所從取樣的宇宙（卽全體學生）只分零與百兩種，這是因爲我們要便於計算，採用了一個例外地簡單的例子。實際上平均數一定從變化很多的數量材料計算的，決不至但有兩種數値。而且所從取樣的宇宙卽成近乎常態的分配，所以卽使樣組很小，平均數的分配也不至與常態分配相差太遠，採用常態分配以考驗重要性是不會引起嚴重誤差的。

照表十三的計算，二〇〇八個投考者的測驗成績平均三一五・七〇分，標準差七五・九三分。現在如果每次抽取九人，九人一組的可能平均數列成次數分配，其平均仍是三一五・七〇分，其標準差爲二五・三一分。這二五・三一卽是九個分數的平均數的標準誤差。假定有九人的平均分數是四百分，四百比三一五・七勝八四・三分，相差有標準誤差的三・三倍，所以可以斷定這差別是重要的，那九人不能認爲是原有二〇〇八人中隨意抽取的，或者已經經過一番選擇，或者不但是高中畢業而是大學畢業生。

因爲原來二〇〇八人的成績分配近乎常態，九人平均成績的分配一定更近乎常態，所以根據常態分配以考驗重要性倒並無問題。普通考驗平均數的重要性，其困難在理論的可能平均數不能完全推演出，只知其平均而不知其標準差。例如有人做心理實驗，比較兩種學習方法，受試者兩人，結果都是甲法勝乙法，一人相差三五分，一人相差四八分，平均相差四一・五分。現在要考驗這四一・五與零是否有重要

的差別，卽考驗甲乙兩法的差別是否偶然的。照偶然之說，兩法的差別平均是零，但其標準差應當是多少，就不容易決定了。普通卽採用樣組的標準差以替代理論宇宙的標準差，如樣組相當的大，所引起的誤差倘不重要。如樣組很小，如目下的例子，這方法是大有問題的。

實得結果是三五與四八，平均四一・五，標準差六・五，標準誤差是四・六一。四一・五與零相差有四・六一的九倍，照常態分配這結果應是非常重要的，我們應當拒絕偶然之說，承認甲法確比乙法優勝。

但前面的結論是不可靠的。因爲樣組太小，只有二，有兩點不可靠的地方：第一，從樣組的標準差算出的標準誤差是非常不可靠的，第二，可能的結果並不成爲常態分配。要免除這兩點困難，我們必需放棄常態分配而採用所謂t分配，而t的定義是：

$$t=\frac{m}{\frac{\sigma}{\sqrt{n-1}}}$$

式中m是樣組的平均，卽例子裏的四一・五，σ是樣組的標準差，卽例子裏的六・五，n是樣組的大小，卽例子裏的二，所以t等於六・三八。相當於各種重要性層級的t列於表十七。但t分配不是一個分配，有無數個，視其自由度而定。這例子的n雖是二而自由度只有一，因爲從兩個分數已經算出一個平均數，有了這平均數兩個分數已失去一部份自由，換句話講，平均數決定之後，兩個分數不能各自消長，其一增則

其二必減，所以共有一個自由度。所以在這種例子裏，自由度數是 n 減一。

查表十七中一度自由的 t 分配，相當於重要性層級千分之一的 t 是六三六有餘，相當於重要性層級百分之五的 t 是一二有餘，而實得的 t 只有六・三八，是不重要的。

讀者細看表十七，可見自由度數漸增，t 分配漸近常態分配，自由度數超過三十之後，差別就很小。但必需自由度數等於無限，t 分配才完全等於常態分配。

以上是考驗平均數與零的的差別，這種需要大概還不如考驗兩個平均數的差別更加普通。例如表九所列的男女身長，平均當然男比女長，但差別是否重要，尚有考驗的必要。平均數差數的均方誤差等於兩平均數的均方誤差的和數。如樣組很大，如表九所列，可計算平均數的差數有其標準誤差（均方誤差的方根）的幾倍，參照常態曲線的面積，判決差數是否重要。但這方法不如計算 t 與參照 t 分配可以同樣適用於大小樣組。計算 t 的公式如下：

$$t=\frac{M_1-M_2}{\sqrt{\frac{n_1\sigma_1^2+n_2\sigma_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

式中兩個 M 是平均數，兩個 σ 是標準差，兩個 n 是兩組的次數。所求出的 t 有 n_1+n_2-2 度自由，查表十七的時候必需注意。

假設現在有男女各二人，體高如左：

男	六尺二寸	五尺四寸	平均五尺八寸
女	四尺九寸	四尺三寸	平均四尺六寸

$$t=\frac{58-46}{\sqrt{\frac{32+18}{2+2-2}}\sqrt{\frac{1}{2}+\frac{1}{2}}}=\frac{12}{\sqrt{25}}=\frac{12}{5}=2.4$$

查表十七，二度自由的t二·四尚未達到百分之五的重要性層級，所以根據這四人的體高我們不能斷定男女有重要的差別。

我們所講的重要性考驗是指統計的重要性，並不是實際的重要性。統計的重要性一方面決定於差別的大小，一方面決定於樣組的大小。有時差別雖大，而樣組太小，如前例，差別便不可靠，所以是統計上不重要的。有時差別雖小而樣組很大，差別便是統計上重要的。例如據生物學家研究，嬰兒的男女比率並不是各佔百分之五十，是男多於女，這差別曾經考驗，在統計上是非常重要的。這結論的成立是由於調查廣泛，因之差別雖小，（約百分之一）卻是十分可靠的。但這種可靠的差別在實際上有甚麼意義，是另一問題。

第十一講　方差分析法

標準差的均方誤差是平均數的均方誤差的一半，兩個標準差的差數的均方誤差是各個標準差的均方誤差的和數：

$$\sigma^2_\sigma = \frac{\sigma^2}{2N} \qquad \sigma^2_{\sigma_1-\sigma_2} = \sigma^2_{\sigma_1}+\sigma^2_{\sigma_2} = \frac{\sigma_1^2}{2N_1}+\frac{\sigma_2^2}{2N_2}$$

但標準差的取樣分配並非常態，所以除非樣組很大，均方誤差或標準誤差不是考驗參差量數的最好工具。考驗參差量數最好用均方差的比率，這種比率的分配情形視均方差的自由度數而異。表十八甲乙分列相當於·○五與·○○一重要性層級的均方差比率。應用表十八的時候，實得均方差的算法與普通略有不同，普通用總次數除方差總和，現在必需用自由度數除方差總和。自由度數的意義前面已經講過，如果均方差是從實得平均計算的，自由度數是總次數減一。至於理論的均方差，算法照以前所講的公式，但其自由度數是無限的，因爲它不是從有限的樣組算出的。查表十八的時候還要注意，均方差比率不可小於一。換句話講，較大的均方差應當做分子，較小的做分母，表中 n_1 是較大而居於分子地位的均方差的自由度數，n_2 是較小而居於分母地位的均方差的自由度數。

現在舉兩個例子：一個考驗實得均方差是否與理論的均方差有重要的不符合處，另一個考驗兩個

實得均方差是否有重要的差別。

表一列着民國二十七年中等學校學生，清代進士及第，與現代名人在十八省的分佈情形。如果這三種人是均勻分佈於十八省的，每省應當佔十八分之一，但各省當然不能完全一律，總有參差照二項分配，十八省之間的均方差應當是：

$$\sigma^2 = npq = n\cdot\frac{1}{18}\cdot\frac{17}{18} = \frac{17n}{324}$$

因為進士及第的人數最少，計算比較簡單，我們以進士及第做考驗的材料。照前面公式，十八省進士及第的人數如果只是偶然有多少，不關環境或遺傳的，人數的參差應當是：

$$\sigma^2 = \frac{17n}{324} = \frac{17\times336}{324} = 17.63$$

而實得的均方差是　$16244\div17=955.53$

實得均方差與理論均方差的比率是　$955.53\div17.63=54.2$

實得均方差有十七度自由，理論均方差有無限度自由，但表十八的n_1，從十二跳到二十四。查表十八乙，相當於·〇〇一重要性層級的均方差比率是從二·七四跳到二·一三，總比所得的比率小，所以實得的與理論的均方差有非常重要的差別，我們不得不拒絕偶然之說而承認十八省之間確有重要因素使人

才集中於少數省區。

因爲這個例子的實得均方差與理論的相差非常之大，已經超過千分之一的重要性，表十八也不能指出所得均方差比率相當於分配面積的確實比率，所以不能與普通的考驗方法比較其結果。照普通方法，

理論的標準差是　$\sqrt{17.63}=4.20$

實得的標準差是　$\sqrt{16244\div 18}=30.04$　$30.04-4.20=25.84$

而標準差的標準誤差是　$\sqrt{17.63\div 648}=.165$

所以相差有標準誤差的一五七倍，在普通的常態曲線表也查不出確實的面積，但知其非常重要而已。

讀者應當注意，如所考驗的是實得標準差與理論標準差的差別，所用的標準誤差是理論標準差的標準誤差。如所考驗的是兩個實得標準差的差別，所用的標準誤差是標準差差數的標準誤差。

在富利門的個性差別（Freeman, Individual Differences）一書裏，有聖約翰（St. John）測驗男女兒童智力的結果，男性的智商標準差比女性大些，應當加以考驗。因爲男女樣組都相當的大，可以用標準誤差與常態曲線。結果如下：

男童的標準差是　13.4　　人數是　503

女童的標準差是 11.5 人數是 455

男女標準差相差 1.9

這差數的標準誤差是 $\sqrt{\frac{13.4^2}{2\times503}+\frac{11.5^2}{2\times455}}=\sqrt{\frac{179.56}{1006}+\frac{132.25}{910}}=.569$

所以差數有其標準誤差的 $\frac{1.9}{.569}=3.34$ 倍，顯已超過千分之一的重要性層級。如用均方差比率考驗，自由度數是五〇二與四五四，所以

男童的均方差是 $\frac{503\times13.4^2}{502}=\frac{503\times179.56}{502}=179.92$

女童的均方差是 $\frac{455\times11.5^2}{454}=\frac{455\times132.25}{454}=132.54$

兩個均方差的比率是 $\frac{179.92}{132.54}=1.358$

但分子自由度數等於五〇二，分母自由度數四五四，在表十八不備，而均方差比率又不適用插入法。讀者如對於前面用標準誤差的考驗不放心，惟有參考費學與夜子合編的生物農業與醫藥研究適用的統計表 (Statistical Tables for Biological, Agricultural and Medical Research)，採用均方差的自然對數，再照公式計算。詳細的步驟現在只得從略。

均方差比率的考驗法是非常有用的，因為所考驗的不限於兩個次數分配的參差量數是否有重要的差別，任何均方差都可加以考驗，而複雜統計材料往往可以將其方差分析為若干部份，歸於若干因子，而比較其輕重。這就是方差分析法，其應用非常之廣，現在不能詳盡討論，只可以舉一個例子。

假定表十九所列是教師五人對於三本考卷先後兩次所批的分數。從這三十個分數我們可以研究許多問題：

一、三本考卷之間究竟是否有高低優劣？
二、五個閱卷者的標準是否寬嚴不同？
三、時間對於所批分數有否影響？第一次與第二次有否分別？
四、閱卷者對於考卷是否各有好惡？
五、時間對於各個閱卷者是否有不同的影響？是否有些閱卷者先寬後嚴，有些相反？
六、時間對於各本考卷是否有不同的影響？是否有些考卷的分數先少後多，有些相反？

表十九所列三十個分數的總平均是五〇，從每個分數減去五〇之後，所得差數如表十九甲所列。這三十個差數的平方總和是九九六六，可以分成以下七部份，考驗前面六個問題。

一、求出前後兩次所批分數的每次平均差數，得三與負三，各十五個，如表十九乙所列，方差的總和是二七〇，共有一度自由，代表時間的影響。

二、求出每本試卷的平均差數，得二〇，〇，與負二〇，各十個，如表十九丙所列，方差的總和是八〇〇〇，共有二度自由，代表三本試卷之間的參差。

三、求出每一閱卷者所批分數的平均差數，得一〇，五，〇，負五與負一〇，各六個，如表十九丁所列，方差的總和是一五〇〇，共有四度自由，代表閱卷者五人之間的參差。

將表十九乙丙丁所列三種平均差數從表十九甲所列總差數除去後，殘餘差數如表十九戊所列。

四、表十九己所列各數代表試卷與時間的交互作用，試卷一似乎第二次批分的時候比第一次退步，而試卷三則進步。這種差數的平方的總和是八〇，有二度自由。

五、表十九庚所列各數代表閱卷者與時間的交互作用，閱卷者乙似乎最有先寬後嚴的傾向，甲次之，戊則兩次完全一致，而丙與丁則有先嚴後寬的傾向。這些差數的平方總和是六〇，有四度自由。

六、表十九辛所列各數代表閱卷者與試卷之間的交互作用，閱卷者丙似乎好試卷一而惡試卷三，丁反之，甲似乎好試卷一而惡試卷二，乙反之。這些方差總和是四〇，有八度自由。

七、從表十九戊所列各數，減去表十九己庚辛所列各數，得最後殘餘差數，如表十九壬所列。方差的和數是一六，有八度自由，代表試卷，閱卷者，與時間的第三級交互作用。這部份方差可用以考驗其他各部份方差的重要性。

茲將七部份方差並列於後：

參差原因	方差和數	自由度	均方差
試卷	八〇〇〇	二	四〇〇〇
閱卷者	一五〇〇	四	三七五
時間	二七〇	一	二七〇
試卷與時間的交互作用	八〇	二	四〇
閱卷者與時間的交互作用	六〇	四	一五
試卷與閱卷者的交互作用	四〇	八	五
誤	一六	八	二

以上七個均方差可以互相比較，計算其比率，考驗其差別的重要性。但表十九所列數字是爲便利說明而假造的，並無考驗的價值。

表二十甲乙所列是眞實試驗結果。取十五本原批分數相等的歷史考卷，請十四人批閱，每人批兩次，共得四百二十個分數。如照前述方法分析，得結果如下：（小數從略）

參差原因	方差和數	自由度	均方差
考卷	一五八七五	一四	一一三四
閱卷者	八五七四	一三	六六〇

時間	五三九	一	五三九
閱卷者與時間	一〇〇四	一三	七七
考卷與閱卷者	九四七三	一八二	五二
考卷與時間	六四八	一四	四六
誤差	三三九〇	一八二	一八·六

我們用一八二度自由的第三級交互作用均方差做考驗的根據，其他六個均方差如確乎比這個大，其所代表的參差便不是無意義的。考驗的結果如下：

考卷　考卷之間的均方差一一三四有均方誤差一八·六的六〇倍，查表十八乙，分子一四度自由不備，但有一二度與二四度，分母一八二度自由不備，只有一二〇度與無限度。但四個均方差比率中最大者不過三·〇二，遠在實得比率之下，所以考卷之間的參差是非常重要的。那十五本考卷的原批分數雖然相等，但照十四人平均的意見，確有高低優劣，平均分數的限域從三一分到五一分，絕非偶然的參差。

閱卷者　閱卷者之間的均方差六六〇有均方誤差一八·六的三五倍，照前段所述表十八乙所列的比率，也超過千分之一的重要性層級很遠。所以十四人批分標準確乎寬嚴不齊，每人平均所批分數自三二分到四七分，比考卷的參差只是略少。

時間　批分先後的均方差五三九有均方誤差一八·六的二九倍。查表十九乙，分子一度自由，分母

如一二〇度自由均方差比率不過一一·三八，也超過千分之一的重要性層級很遠。所以第二次批分平均比第一次所少雖不足三分，這差別卻不是偶然的。

閱卷者與時間的交互作用　交互作用的均方差都比較小，但仍是重要的。閱卷者與時間的交互作用有一三度，自由千分之一的重要性最多只需比率三·〇二，實得均方差七七，是均方誤差一八·六的四倍，可見時間對於閱卷者的影響並不一致。照前段結論，閱卷者全體有先寬後嚴的趨勢，第二次批分平均比第一次減低二分又一五分之四，現在知道各人之間的參差也不是偶然的。查閱卷者B第二次平均比第一次多批三分以上，閱卷者D第二次平均比第一次少批六分以上。

閱卷者與考卷的交互作用　這一八二度的均方差是均方誤差一八·六的二·八倍。分母分子的自由度都是一八二，表十八乙不備，只得查最近的四個比率（即分子二四度與無限度，分母一二〇度與無限度）這四個比率中最大的不過二·四〇，所以實得比率是重要的，可見閱卷者對於考卷各有好惡。例如閱卷者B對於第十本與第十一本考卷的態度顯然是不同的。如閱卷者十四人所批每本考卷的平均分數可以代表其眞正價值，從每人所批分數減去十四人平均所批分數後的差數便是每人的誤差。這誤差可以分爲兩部份，一部份是對於各本卷子共同的誤差，即個人太寬或太嚴的傾向，一部份是對於每本卷子特有的誤差，即個人特殊的好惡。前者的方差和數是八五七四，後者的方差和數是九四七三，按參差的大小，後者略超過前者，但因自由度數不同，後者不如前者重要。

考卷與時間的交互作用　這裏實得均方差四六是均方誤差一八・六的二・五倍。表十八乙上自由度數相差最近的四個比率兩個較大兩個較小，所以不能確定其是否達到千分之一的重要性。但表十八甲相當於分子一二度分母一二〇度的比率不過一・八三，所以實得比率至少已經超過百分之五的重要性。可見時間對於考卷的影響並不一致，有些考卷第二次看似乎比第一次好（例如第一本），有些第二次不如第一次（例如第十五本）。

所以非但考卷，閱卷者，與時間都是決定所批分數的因子，這三個因子的三種第二級交互作用也都是重要的。至於第三級交互作用，是顯然無意義的，均方差是誤差，用以考驗其他各部份的均方差。

各部份方差的計算方法，有種種變化與可以簡省的地方，未能盡述，但都不外乎應用以下原則：

$$\frac{\Sigma X^2}{N}=\frac{\Sigma(x+d)^2}{N}-d^2$$

即凡從平均以外的出發點計算均方差，必然比從平均計算的結果大，其差別即等於該出發點與平均的差數的平方。這層在討論參差量數的時候已經講過。但讀者應當注意平均不必指四二〇個分數的總平均，凡Σ號代表一部份而非全體方差和數的時候，平均當然也是一部份的平均而非全體的總平均。

關於各部份方差的自由度數，計算的規則是這樣的：如以a代表考卷的本數或任何第一因子的變化，以b代表閱卷者人數或任何第二因子的變化，以c代表批閱次數或任何第三因子的變化，則由於

第一因子的參差有自由度數　$a-1$

第二因子的參差有自由度數　$b-1$

第三因子的參差有自由度數　$c-1$

一二因子交互作用的參差有自由度數　$(a-1)(b-1)$

一三因子交互作用的參差有自由度數　$(a-1)(c-1)$

二三因子交互作用的參差有自由度數　$(b-1)(c-1)$

一二三因子交互作用的參差有自由度數　$(a-1)(b-1)(c-1)$

總計有自由度數　$(a-1)+(b-1)+(c-1)+(ab-a-b+1)+(ac-a-c+1)$

$+(bc-b-c)+(abc-ab-ac-bc+a+b+c-1)=abc-1$

更加複雜的方差分析不難照此類推。

第十二講 相關量數

我們對於任何事物，都可以作幾方面的觀察。各方面可以分別研究，也可以共同研究。如共同研究，除了每一方面的分配情形而外，還有各方面的互相關係問題，比了分別研究更有意義。分別研究的方法前面已經講過大概，相關的研究是現在所要講的。

假設我們研究某校學生入學考試成績與進校一學期後學業成績的關係，所要解決的是以下一類的問題：

各個學生的這兩種成績是不是有一致升降的傾向？入學考試成績好的學生，是不是進校一學期後的成績也好？入學考試成績差的學生，是不是進校一學期後的成績也差？這種傾向怎樣可以加以測量？

我們能不能根據一個學生的入學考試成績，推測他進校一學期後的成績？怎樣推測才是最有效的辦法？推測的公式是不是跟着相關程度的高低而不同？

有根據的推測與憑空瞎猜有甚麼不同？推測的誤差應當怎樣計算？是不是可以預先估計？是不是跟着相關的高低而有大小？

相關的研究是最有興趣，最有意義，也最可以應用的。相關的材料更是隨時隨地可以遇着的。除了前

面所引入學考試成績與進校一學期後學業成績的關係外，再舉若干例題於下：各種學科的成績之間的關係，小學成績與中學或大學成績的關係，在校成績與將來事業的關係，身長與體重的關係，父母與子女能力的關係，家境與求學努力的關係，米價與學費的關係，學費與教師薪水的關係，班級大小與教學效率的關係。

處理相關統計材料的第一步是編列一張相關表，如表二十一甲至表二十一己那樣。相關表與次數分配表約略相似而比較複雜，次數分配表中的次數是根據一個變量而歸類的，相關表中的次數是同時根據兩個變量而歸類的。所以如果根據一個變量可以把次數分爲十組，根據兩個變量便可以把次數分爲一百組。不過這一百組中往往有好些組是備而無用的，尤其在相關比較高的時候。

在未經確實計算之前，從相關表中次數分配的大概情形，已可約略看出相關的高低程度。相關最高的時候，雙方完全一致消長，所有的次數都集中於一條斜線。相關漸低，次數逐漸分散，直到毫無相關的時候，次數也毫無集中斜線的傾向。表二十一甲至表二十一己代表逐漸增高的相關情形。每表左上角的字母r，即是相關係數。相關係數最低是零，即毫無相關，所以表二十一甲的次數是非常散漫的，X大的時候，Y既是大小不一，X小的時候，Y仍是大小都有。相關係數最高是一，即完全的相關，所以表二十一己的次數都在一條斜綫上，X與Y完全一致消長，毫無參差。介乎這兩種極端情形之間，相關係數便有介乎零與一之間的數值。讀者如將表二十一甲至己比較觀察，可以對於高低不同的相關係數有些具體的觀念。

但相關係數的極限並不是零與一，卻是正一與負一。負的相關也不是罕見的，例如同一年級的兒童，其年齡與智商的相關大概是負的，即年齡愈大者智商愈低。我們如把表二十一甲至己的『X』改爲『十減X，』所有的相關係數都變成負的，但其數值的高低一律仍舊。所以相關係數如果等於負一，相關表中的次數也完全集中於一條斜線，係數數值漸低，則次數漸散漫。

現在我們進一步研究相關係數應當怎樣計算。我們的目的要正一代表完全一致的相關，負一代表完全相反的關係，零代表缺乏任何關係。我們如從第七講裏所講的標準分數着想，這問題是相當簡單的標準分數是一律從平均出發，以標準差爲單位的。假設入學考試成績與進校一學期後成績有完全一致的相關，這兩種標準分數應當完全一致消長，每一學生的兩種標準分數應當恰好相等，因爲每一學生的兩種成績應當同在平均以上（或以下）一個標準差的幾倍（或百分之幾）。如用Z_1與Z_2代表兩種標準分數，完全一致相關的條件是：

$$Z_1 = Z_2 \quad \text{或} \quad Z_1Z_2 = Z_1^2 = Z_2^2 \quad \text{所以} \quad \frac{\Sigma Z_1Z_2}{N} = \frac{\Sigma Z_1^2}{N} = \frac{\Sigma Z_2^2}{N}$$

如果關係是完全相反的，兩種標準分數應當有相反的符號，而數值相等。所以相關係數等於負一的條件是：

$$z_1 = -z_2 \quad \text{或} \quad z_1z_2 = -z_1^2 = -z_2^2 \quad \text{所以} \quad \frac{\Sigma z_1z_2}{N} = \frac{-\Sigma z_1^2}{N} = \frac{-\Sigma z_2^2}{N}$$

但 $\frac{\Sigma z^2}{N}$ 即是標準分數的標準差，一定等於一，所以 $\frac{\Sigma z_1 z_2}{N}$ 就可以當做相關係數，實際上也確乎就是相關係數。相關係數普通用字母 r 代表，所以相關係數的公式是：

$$r=\frac{\Sigma z_1 z_2}{N}$$

照這公式，我們如要計算入學考試成績與進校後一學期成績的相關係數，只需先將每一學生的兩種分數合成標準分數，求其積數，再求全體學生的這種積數的平均即得。

標準分數的定義是：

$$z=\frac{X-M}{\sigma}=\frac{x}{\sigma}$$

所以前面的公式也可以寫成：

$$r=\frac{\Sigma\left(\frac{x}{\sigma_1}\cdot\frac{y}{\sigma_2}\right)}{N}$$

但是為每一學生分別做兩次除法當然不如為全體學生合併做一次除法簡單，實際上不必求出標準分數，可以應用下面的公式，方便多了：

$$r=\frac{\Sigma xy}{N\sigma_1\sigma_2}$$

這公式是最常見的相關係數公式。但實際計算的時候，如用每一分數與平均數的差別，還不是最便利的方法。因爲平均數往往是帶着小數的，差數因之也帶着小數，求積數比較費事。我們可以依照計算平均數與標準差的方法，先用一近似平均數做出發點，計算每一分數與這出發點的差別，用這種差數求積數，最後加以校正。如用M′代表這種出發點，用x′與y′代表從這種出發點計算的差數，即：

$$x' = X - M'_1 \qquad y' = Y - M'_2$$

校正數便是x′或y′的平均，也就是平均數與出發點的差數：

$$c_1 = \frac{\Sigma x'}{N} = M_1 - M'_1 \qquad c_2 = \frac{\Sigma y'}{N} = M_2 - M'_2$$

照以上這些定義，相關係數的公式可以寫成：

$$r = \frac{\frac{\Sigma x'y'}{N} - c_1c_2}{\sigma_1\sigma_2}$$

相關係數的另一個公式是根據下面的關係：

$$s^2 = (x' + y')^2 = x'^2 + 2x'y' + y'^2$$

$$d^2 = (x' - y')^2 = x'^2 - 2x'y' + y'^2$$

$$s^2-d^2=4x'y' \qquad \frac{s^2-d^2}{4}=x'y'$$

所以

$$r=\frac{\dfrac{\Sigma s^2-\Sigma d^2}{4N}-c_1c_2}{\sigma_1\sigma_2}$$

最後這兩個公式是實際計算最適用的。最通行的相關係數計算表大都兩式並用，互相核對，以免錯誤，更是萬全的方法。

以上所舉四個公式，前二個可以說是相關係數的最簡單定義，後二個最便於直接應用。此外公式的變化還多得很。據調查各種統計書籍的結果，相關係數的公式總共有五十二種不同的寫法。這些寫法都有其特殊的好處，但都可以用簡單代數證明相等，現在限於篇幅，不能列舉。

相關係數是表示兩件事情有否一致升降傾向的一個數量。如要知道入學考試成績好的學生是不是進校一學期後的成績也好，入學考試成績差的學生是不是進校一學期後的成績也差，只需計算兩種分數的相關係數。結果愈近正一，愈可以作肯定的回答。負的係數指示相反的傾向，零指示兩種分數各自升降，毫無一致或相反的關係。

假設兩人各以硬幣若干枚投擲，每次以所得正面計算分數。投擲多次之後，每人的分數約略成二項分配。但兩人的硬幣與投擲方法既無關係，兩人的分數當然各自升降，並無一致（或相反）的傾向，所以

相關係數應當是零。但假設兩人的分數決定於同一組硬幣，則兩人的分數當然每次相等，完全一致升降，相關係數一定等於一。這是非常明顯的。

假設兩人以一部份硬幣共同投擲，一部份分別投擲，則每次分數並不相等，卻也不完全是各自升降。比方硬幣總共十枚，五枚共同投擲，五枚分別投擲，兩人每次至多相差五分，不像十枚完全分別投擲那樣，可以相差到十分之多。所以在這種情形下，相關的高低就決定於共同的硬幣與個別的硬幣的比例。其實共同的硬幣在全體硬幣所佔的百分比可以直接當做相關係數。實地試驗可以證實這一點，不過實驗結果有取樣誤差，除非樣組很大，即投擲很多次，結果不免略有出入。

表二十一甲至己所列次數是根據二項分配算出的，代表投擲十枚硬幣中有零、二、四、六、八、十枚是共同的。例如表二十一丙，代表四枚硬幣共同投擲，六枚分別投擲的理想結果，表中各數是一個四次的二項分配與一個六次的二項分配的合併結果。

所以如果相關係數所代表的是投擲硬幣的結果，係數可以直接解釋做共同因素的百分比。這是十分正確的解釋。但這種解釋不能普遍適用，例如身長與體重的相關係數·五〇很難認爲足以證明決定身長與體重的因素有百分之五十是共同的。

相關係數最普遍適用的解釋是係數的推測或估計的能力，與估計的誤差的減少。下面我們討論推測或估計。

我們首先考慮憑空猜測的方法與誤差。在第五講與第六講已經討論過，平均差從中數計算，均方差從平均數計算，比從任何其他數值計算小。所以憑空猜測的最妥當方法是採用平均數或中數，因爲既無根據可以推想每一個體在團體中的地位，惟有採用平均數或中數作爲每一數值的估計，才不至於與實際情形相差太遠。如採用中數，平均差即是估計的平均誤差。如採用平均數，均方差即是估計的均方誤差。因爲平均數與均方差比中數與平均差便於處理，我們採用平均數與均方誤差。

假設甲乙兩種成績的相關係數是零，知道了一個學生的甲種分數，絲毫無補於他的乙種分數的估計。不論其甲種分數怎樣好或怎樣壞，其乙種分數的最妥當估計還是平均數。這種估計的均方誤差等於乙種分數的均方差。所以結果與憑空猜測完全一樣。

甲種分數眞完全無法利用嗎？我們可以試試看：假定用甲種成績的標準分數當做乙種成績的標準分數，這樣估計的均方誤差可以算出如下：

$$\frac{\Sigma(z_1-z_2)^2}{N}=\frac{\Sigma z_1^2}{N}-\frac{2\Sigma z_1 z_2}{N}+\frac{\Sigma z_2^2}{N}=1-2r+1=2-2r=2(1-r)$$

這均方誤差顯然隨着相關的高低而不同。相關高到一的時候，估計的均方誤差低到零。相關低到零的時候，估計的均方誤差高到二。

如用甲種分數的平均數當做乙種成績的估計，即用零當做乙種標準分數，估計的均方誤差等於乙

種標準分數的均方差，卽是一：

$$\frac{\Sigma(0-z_2)^2}{N}=\frac{\Sigma z_2^2}{N}=1$$

這均方誤差是固定的，並不因相關高低而不同。

所以如果相關係數是零，我們應當採用零當做任何學生的乙種標準分數。這種估計的均方誤差等於一，與憑空猜測的結果一樣。但相關既是零，這是惟一合理的估計方法，無法可以減低估計的均方誤差。如果相關係數是一，我們應當採用每一學生的甲種標準分數當做他的乙種標準分數。這種估計當然完全正確，均方誤差等於零。

現在的問題是，如果相關係數高於零而低於一，我們應當採用甚麼方法估計？按照前面的推算，如用甲種成績的標準分數當做乙種成績的標準分數，估計的均方誤差等於 2(1－r)，如一概用零當做乙種成績的標準分數，估計的均方誤差等於一。比較這兩種估計方法的均方誤差，可見相關係數如大於·五，前者比後者小，相關係數如小於·五，前者比後者大。根據這個比較，似乎相關係數大於·五的時候應當採用甲種成績的標準分數當做乙種成績的標準分數，而相關係數小於·五的時候應當一概採用零當做乙種成績的標準分數。這是假定估計乙種成績，必需將甲種成績完全利用或完全不顧。

但實際上有一種介乎二者之間的折中辦法，更加合理。估計乙種成績的方法，除了絕對不用甲種成

績，或完全採用甲種成績而外，爲甚麽不按照相關係數的高低，利用甲種成績而加以保留？如相關係數等於・五，以甲種成績標準分數的一半當做乙種成績的標準分數，不是比前面所講的兩種極端方法更加合理麽？完全可信的消息當然應當全信，完全不可信的消息當然應當置之不理。但若某種消息既非完全可信，又非完全不可信，則全信之固是盲從，全不置信亦屬可惜。最合理的態度是疑信參半。所以甲乙兩種成績的相關係數如等於・五，而某生的甲種成績在平均以上一個標準差，其乙種成績的最合理估計，既非與甲種成績同樣在平均以上一個標準差，亦非與憑空猜測那般等於平均數，而是平均以上半個標準差。

假設我們用相關係數乘甲種成績的標準分數，當做乙種成績的標準分數，這種估計的均方誤差應當如下面的推算：

$$\frac{\Sigma(rz_1-z_2)^2}{N}=\frac{r^2\Sigma z_1{}^2}{N}-\frac{2r\Sigma z_1z_2}{N}+\frac{\Sigma z_2{}^2}{N}=r^2-2r^2+1=1-r^2$$

這均方誤差的最高數值是一，最低數值是零，不會超過一，也不會超過 $2(1-r)$，因爲 $1-r^2=(1+r)(1-r)$，而 $1+r$ 決不會超過二。所以這種估計方法的結果不會不如前面兩種方法的結果。實際上這方法已經包括那兩種方法，同時適用於不論怎樣高與怎樣低的相關係數。

我們採用上面加有一畫的字母來代表估計的分數，以別於實得分數，前段所講的估計方法可以寫

成如下的方程式：

$$\bar{z}_2 = rz_1$$

如不用標準分數而用分數與平均的差數，方程式成為：

$$\frac{\bar{y}}{\sigma_2} = r\frac{x}{\sigma_1} \quad 或 \quad \bar{y} = r\frac{\sigma_2}{\sigma_1}x$$

如不用差數而用原有分數，方程式成為：

$$\bar{Y} - M_2 = r\frac{\sigma_2}{\sigma_1}(X - M_1) \quad 或 \quad \bar{Y} = M_2 - r\frac{\sigma_2}{\sigma_1}M_1 + r\frac{\sigma_2}{\sigma_1}X$$

這三個適用於估計的方程式稱做消長方程式或回歸方程式，$r\sigma_2/\sigma_1$ 稱做消長係數或回歸係數，也可以用單個字母 b_{21} 代表。如根據乙種分數估計甲種分數，這些方程式中的 x 與 y 或 X 與 Y，以及各字母的下標 1 與 2 應當互換。更有一點必需注意的，前面所推演出的估計均方誤差 $1-r^2$ 只適用於標準分數。如所估計的是原有分數或其與平均數的差數，估計的均方誤差應當用所估計的分數的均方差乘過。換句話講，如從甲種分數估計乙種分數，估計的均方誤差是 $\sigma_2^2(1-r^2)$，而 $\sigma_1^2(1-r^2)$ 是從乙種分數估計甲種分數的均方誤差。

每一張相關表上，可以畫出兩條消長線或回歸線，相當於前面的消長方程式或回歸方程式。求出相

關表中每行與每列（即縱行）的平均數的一條直綫相當於從x估計y的回歸方程式，通過每行平均數的一條直線相當於從y估計x的回歸方程式。如相關係數是零，這兩條回歸線即是縱橫軸線，互成直角。相關漸高，兩線漸傾斜。如相關是一，兩線合併成一線。讀者可以用表二十一甲至己自行試驗。

所以，從x估計y，即是取相當於指定的x的一列y的平均數，而估計的均方誤差即是各個y距離各列平均數（不是全體平均數）的均方差。如相關是零，各列平均數並無高低，都等於全體平均數，所以估計的均方誤差等於全體的均方差。如相關是一，每列之內並無參差，所以估計的均方誤差是零。

除非相關係數等於一，估計的結果比了所從估計的分數總是退向平均數。最初研究相關問題的高耳頓（Sir Francis Galton）早就發現相當於父母每一高度的子女的平均高度總比父母的高度近於平均數。這是相關係數低於一的必然情形。一知半解者因此以爲極高與極矮的人將一代一代的減少，那是不合事實的可笑結論。試將父母與子女的地位互換，退向平均的現象同樣存在。相當於子女每一高度的父母的平均高度也總比子女的高度近於平均數，難道這又指示我們的祖先都是並無高矮的嗎？

讀者必需注意，相關只是數值上的關係，其惟一普遍適用的解釋是，相關愈高愈能減小估計的誤差。我們絕對不可從相關係數直接推論因果關係。

第十三講　相關問題續論

相關的範圍很廣，問題很複雜。第十二講所處理的只是相關的一小部份，其他各部份本講也不能詳盡地討論，只能略舉一二，以見問題的存在。

前次所講的相關係數是皮耳生(Karl Pearson)的積矩相關。如果相關的兩方面不用數量而用等第或名次代表，計算的時候可以採用士比門(C. Spearman)的等差相關方法，更加簡便。等差相關的公式是

$$\rho=1-\frac{6\Sigma d^2}{n(n^2-1)}$$

式中的n代表全體的次數，d代表兩種名次的差數。茲舉一簡單例子計算如下：

甲種成績名次	乙種成績名次	d	d^2
一	二	1	1
二	一	1	1
三	三	0	0
四	五	1	1
五	四	1	1

$$p = 1 - \frac{6 \times 4}{5 \times 24} = 1 - \frac{1}{5} = .80$$

讀者如以名次當做數量而用積矩公式計算，結果是完全一樣的。其實等差公式可以用代數方法從積矩公式推演而得：

$$r = \frac{\Sigma xy}{N\sigma_1\sigma_2} = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma(x-y)^2}{2N\sigma_1\sigma_2} = \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2} - \frac{\Sigma(x-y)^2}{2N\sigma_1\sigma_2}$$

但N個相鄰數的均方差是：

$$\sigma_1^2 = \sigma_2^2 = \sigma_1\sigma_2 = \frac{n^2-1}{12}$$

而

$$x - y = d$$

所以

$$r = 1 - \frac{6\Sigma d^2}{n(n^2-1)} = p$$

因等差方法比積矩方法簡便，原屬數量的材料有時也被折合等第計算。但次數分配大都不是均勻的，所以前面的演算不能適用。算出的結果當然也不能與積矩方法相等。如不相等，等差方法只是一種簡短的約計，不如積矩方法正確。

前次所講兩種成績的相關，並不顧及他種成績或任何第三變量，如身長、體重、年齡、年級之類。那種相關稱爲全相關與單相關。與全相關對立的有偏相關，與單相關對立的有複相關。現在略述偏相關與複相

關的大意。

假設有人報告他的調查結果，說兒童的身長與體重相關高至.六五，這報告有一個很大的缺點，就是絕未提及兒童的年齡。因爲身長與體重是身體發育的兩大表現，而年齡是發育的最重要因素。如所調查的兒童屬於同一年齡，則身長與體重的參差限域很小，相關也因之而低。如所指兒童包羅十二歲以內所有年齡，則年長者大概總比年幼者又高又重，相關勢必很高。所以身長與體重的相關係數視年齡的限域而不同，如不說明年齡的限域，相關係數便缺乏意義。如要比較年齡限域不同的身長與體重相關，我們必需假定兒童的年齡完全相等，計算其身長與體重的偏相關係數。當然我們可以把兒童按照年齡分組，分別求算每組的相關係數。另一方法是應用以下公式推算：

$$r_{12\cdot3}=\frac{r_{12}-r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$$

式中的 r_{12} 是第一第二兩變量（身長與體重）的全相關係數，r_{13} 與 r_{23} 是第一三兩變量與第二三兩變量（身長與年齡、及體重與年齡）的全相關係數，$r_{12\cdot3}$ 是假定第三變量（年齡）固定時第一二兩變量（身長體重）的偏相關係數。這樣求出的偏相關係數等於分組計算所得全相關係數的平均。

$r_{12\cdot3}$ 是第一級偏相關係數。如應用前面公式，以 $r_{12\cdot4}$ 代 r_{12}，以 $r_{13\cdot4}$ 代 r_{13}，以 $r_{23\cdot4}$ 代 r_{23}，所得爲 $r_{12\cdot34}$，是第二級偏相關係數。第三級或更高級的偏相關係數可依此類推。全相關係數也稱爲零級相

關係數。

$r_{12\cdot3}$ 大都比 r_{12} 小，普通往往認爲這是因爲第三變量是第一第二兩變量的共同因子的原故。這種解釋雖然有時可以偶與事實符合，但並不普遍正確。例如以下的相關係數，

$$r_{12}=r_{13}=r_{23}=\frac{1}{2} \qquad r_{12\cdot3}=r_{13\cdot2}=r_{23\cdot1}=\frac{1}{3}$$

至少有兩種可能的解釋：(一)三個變量有一共同因子，每一變量是這共同因子與各別的特殊因子的和數。(二)三個變量並無完全共同的因子，存在於第一第二兩變量的因子甲不存在於第三變量，存在於第一第三兩變量的因子乙不存在於第二變量，存在於第二第三兩變量的因子丙不存在於第一變量，另外每一變量也各有其特殊因子。這兩種可能解釋究竟孰正孰誤，或二者都誤，不是相關方法所能解決的。

前次所講的估計，是根據一種成績估計另一種成績。但我們也可以根據兩種成績估計第三種成績。這種估計可以應用下面的多項消長方程式：

$$\bar{z}_3=\frac{r_{13}-r_{12}\,r_{23}}{1-r^2_{12}}\,z_1+\frac{r_{23}-r_{12}\,r_{13}}{1-r^2_{12}}\,z_2 \quad 或$$

$$\bar{x}_3=\frac{r_{13}-r_{12}\,r_{23}}{1-r^2_{12}}\,\frac{\sigma_3}{\sigma_1}\,z_1+\frac{r_{23}-r_{12}\,r_{13}}{1-r^2_{12}}\,\frac{\sigma_3}{\sigma_2}\,z_2$$

這種估計普通總比根據一種成績的估計可以正確，其均方誤差是：

$$\frac{\sigma^2_3(1-r^2_{12}-r^2_{13}-r^2_{23}+2r_{12}\,r_{13}\,r_{23})}{1-r^2_{12}}$$

估計結果與實得分數的相關係數是複相關係數，可照以下公式求得：

$$r_{3\cdot12}=\sqrt{\frac{r^2_{13}+r^2_{23}-2r_{12}\,r_{13}\,r_{23}}{1-r^2_{12}}}$$

從估計的均方誤差或複相關的公式，可見根據以作估計的兩種成績應當互相相關低，而各與所要估計的成績相關高，則估計可以正確，誤差可以低而複相關可以高。

根據以估計的變量當然不限於二，且並無任何限制。但變量加多之後，消長方程式與複相關公式更加複雜，本書不便討論。照數理講，變量愈多，估計愈正確，複相關愈高。但實際上採用四五種變量以後，再加新的變量往往並無新的貢獻，因其所供給的知識已經包容在已有的變量以內，複相關係數因此每在距離一尚遠的時候，達到不容易再進步的限度。

更有一點，變量愈多則消長係數的取樣誤差愈大。所以從一組材料求出的多項消長方程式，如應用於另一組材料，其複相關係數總有降低的趨勢。所以增加變量必需同時擴大研究的範圍，否則不能增加收獲。

複相關與多項消長方程式在職業選擇與職業指導非常有用。我們的希望是學者能發現幾種變量，

或發明幾種測驗方法，可用以估計人的各種能力，或在各種事業的成功機會，得到近於一的複相關係數，可以減少許多浪費，增加無限幸福。

前次所講的相關，如表二十一甲至己所示，各行或各列的平均成一直線，所以是直線相關。與直線相關對立的尙有曲線相關，如表二十二所示。表二十二各列的平均數都等於全體的平均數，所以成一橫線。各行的平均數則先由小而大，再由大而小，成一半圓形。如照前次所講直線相關的方法，從表二十二的材料計算相關係數，結果得零。但表中次數既集中成一弧形，X與Y決非毫無關係。其實表中的X是標準分數的估計標準誤差，而Y是相關係數，$X=\sqrt{1-Y^2}$，X與Y互成函數，關係可謂密切。於此可見相關係數不適用於這一類的問題，應當求消長曲線的方程式與相關比率。

消長曲線方程式的求算方法本複雜，本書不擬討論。表二十二的消長曲線可用 $Y^2=1-X^2$ 或 $Y=\pm\sqrt{1-X^2}$ 代表，讀者試以各數值代X而求Y，所得結果可與表二十二完全適合。

相關比率有y依x的相關比率與x依y的相關比率，二者不同，求法可應用前面講過的方差分析法，並不很困難。如求y依x的相關比率，先計算全部y分配從總平均的方差總和，再求每列的平均數，計算各列平均離總平均的方差總和，另外可以計算各列內部（卽每一y離其所在列平均的）方差總和，以資核對，因各列之間的方差與各列之內的方差和數應當等於全部方差總和。相關比率是以全部方差總和除各列之間的方差總和，而取其商數的方根。

如計算x依y的相關比率，所要計算的是全部x分配的方差總和，各行之間的方差和數，與各行以內的方差和數。

照這方法，用表二十二的材料，計算x依y的相關比率，得·九四五，再算y依x的相關比率，得零。實際上x既是y的函數，從y計算（不僅是估計）x，可以絕對正確，毫無誤差，相關比率應是一。只因表二十二的組距太大，相關不免減低，例如y的·九五那組，包容y自·九〇至一·〇〇，所以相當於y的x自·〇〇直至·四四。如y恰是·九五，則相當於y的x是·三一二二，不會像表二十二上那樣散漫。至於y依x的相關比率，因相當於每一x的y有正負兩種數值，平均是零，即每列的平均y都是零，所以各列之間的方差是零，而各列以內的方差總和等於全部分配的方差總和，所得結果就不得不等於零。如以表二十二的上下兩半分開，y依x的相關比率便和x依y的相關比率一樣等於·九四五了。這樣分開之後，相關係數也不是零，而在上半等於負·八八一，在下半等於正·八八一。

相關比率普通用希臘字母η（讀作「依它」）做符號，η_{12}是第一變量依第二變量的相關比率，η_{21}是第二變量依第一變量的相關比率。η_{12}與η_{21}普通不相等，總是正的，總不會比r_{12}（即是r_{21}）低。

以前所講的相關限於數量材料的關係，但品質之間的關係或質與量之間的關係也應當考慮。例如籍貫與身長的關係，職業與智力的關係，都是質與量之間的關係。而籍貫與職業的關係便是品質之間的關係了。

相關比率的方法是適用於質量關係的。質依量的相關比率固屬不可能，量依質的相關比率卻不難計算。照前面所講的方法，如要計算智力依職業的相關比率，只需以全體的智力方差總和作分母，以各職業間的平均智力方差和數作分子，取其商數的方根。

計算品質之間的相關，以前所講的方法都不能適用，必需完全採取新的方法。現在以表二十三甲至丁作例，說明方法的大概。表二十三甲所示是一千人的本人國籍與其所最愛好的音樂國別的分配，如二百個英國人中有三二人最愛好英國音樂，一六人最愛好法國音樂，七五人最愛好德國音樂等等。假設本人國籍與其所最愛好的音樂的國別毫無關係，各國人所最愛好的每一國音樂應當佔相等的比例。所謂相等的比例並不是說表二十三甲的二十五個次數應當完全相等，因爲愛好德國音樂的人多，愛好英國音樂的人少，這是不能決定國籍與所好音樂的相關高低的。問題是各國人愛好德國音樂的是否同樣多，愛好英國音樂的是否同樣少。

所以我們第一步應當根據國籍與所好音樂毫無相關的假設算出二十五個理論的次數。這只需用每國人數乘每國音樂被愛好的總數比率，即用每行的總數乘每列的總數，再用全體的總數除，結果是表二十三乙所示的分配。因爲各國人數相等，所以每列五個次數都是相等的。這當然並不是假設次數分配的必然狀態。

以表二十三甲與表二十三乙比較，可見出入很多，其差數列示於表二十三丙，各國人愛好本國音樂

的特別多，所以從左上至右下角斜綫上五個差數是每列最大的，其餘大都是負數。實得次數與假設次數的差別，如表二十三丙所示，頗足以指示相關的高低。但這種差數總是正負相抵，總和得零。

以實得次數與假設次數的差數自乘，再除以假設次數，所得結果如表三十三丁所示。這些數值的總和稱爲χ。χ^2是考驗實得次數與假設次數差別重要性的極有用方法，相當於各重要性層級與各自由度數的χ^2列示於表二十四。

在求算相關量數以前，我們先考驗表二十三甲乙的差別是否堪稱重要。照表二十三丙，$\chi^{2\prime}$等於一九五·六。但在應用表二十四之前，尙需決定這χ^2的自由度數。這χ^2是從二十五個次數算出的，但這二十五個次數並不完全互相獨立，每行或每列的總數在計算假設次數時必需應用，五個次數的四個確定以後，其餘一個也就失卻自由。所以每行或每列只有四度自由，總共有十六度自由。如以s代表行數，t代表列數，自由度數總是等於s－1與t－1的積數。

查表二十四，相當於重要性層級千分之一與十六度自由的χ^2是四·〇一五，現在實得χ^2是一九五·六，超過很多。所以差別非常重要；假設必需拒絕，各國人民對於各國音樂各有所偏愛。但僅從χ^2不能知悉偏愛的情形，特別愛好本國的音樂，或特別厭惡本國的音樂，可以是同一χ^2的原因。我們必需參照表二十三丙，才可以辨別偏愛的方向。

χ^2所指示的是實際與理論的差別是否重要。相關不存在的理論雖已拒絕，相關的高低尙待計算。從

表二十三甲可見各國人民雖然都有偏愛本國音樂的傾向，這傾向雖然已被χ^2的考驗證明爲非常重要，但各國人民愛好異國音樂的仍有不少，所以相關不會很高，必需每例的次數都集中於一格，相關才會高至一。

計算品質相關最適用的是朱普羅的係數(Tschuprow's coefficient)，公式如下：

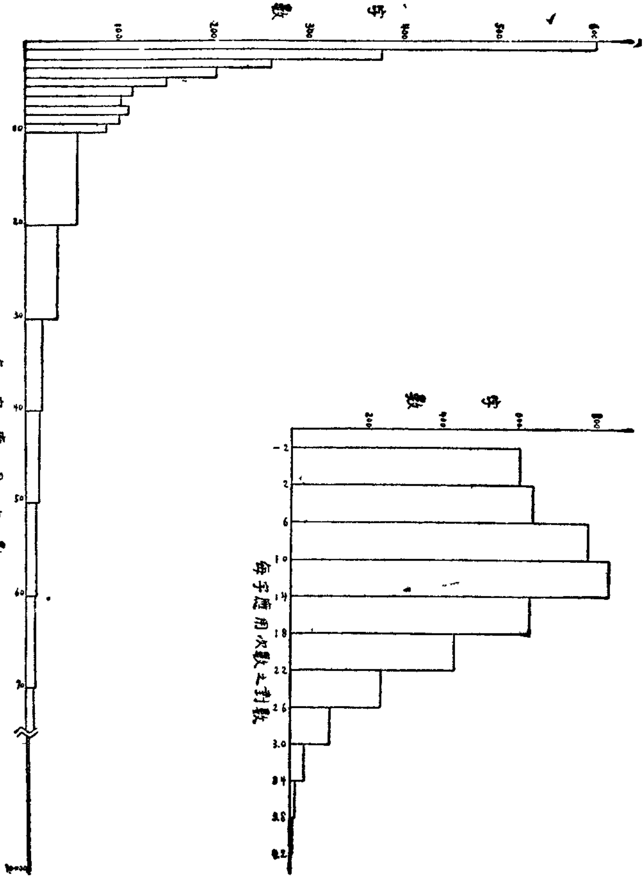
$$T^2=\frac{\chi^2}{N\sqrt{s-1}\sqrt{t-1}}$$

式中的T是朱氏係數，χ^2是各個實得次數與假設次數的差數的乘方，被假設次數除後的總和；N是總次數，s與t是行數與列數。從前面所得χ^2一九五·六，可求朱氏係數如下：

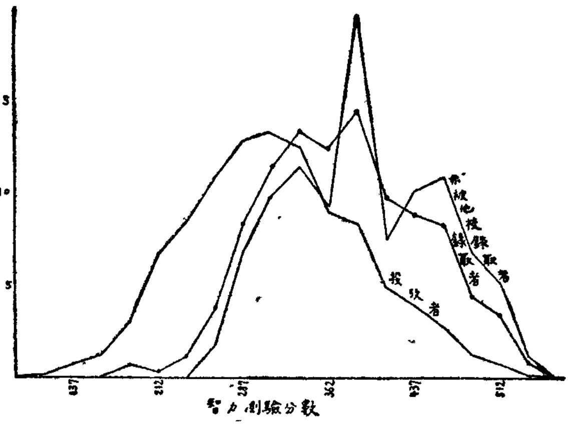
$$T^2=\frac{195.6}{1000\sqrt{5-1}\sqrt{5-1}}=\frac{.1956}{4}=.0489 \qquad T=.22$$

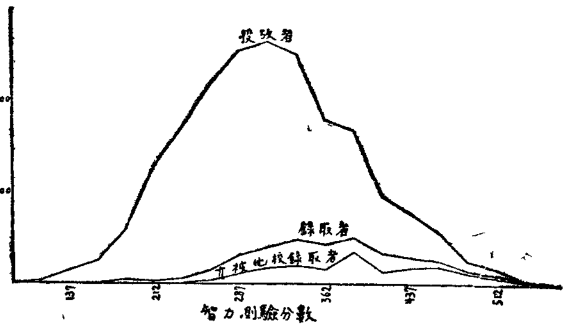
品質的相關係數不像數量的相關係數可用以作估計，所以計算的需要比較少。很多時候用χ^2考驗事實與假設的差別就夠了。

本講提出士比門的等差相關方法以補充皮耳生的積矩相關方法，提出偏相關以補充全相關，提出復相關以補充單相關，提出曲線消長與相關比率以補充直線消長與相關係數，提出質與量的相關及品質之間的相關以補充數量之間的相關，各種相關方法總算都備一格了。但因限於篇幅，提到的方法未能詳細討論，更有許多方法根本未能提到。更詳細更完備的研究，只得請讀者另求諸別的著作了。

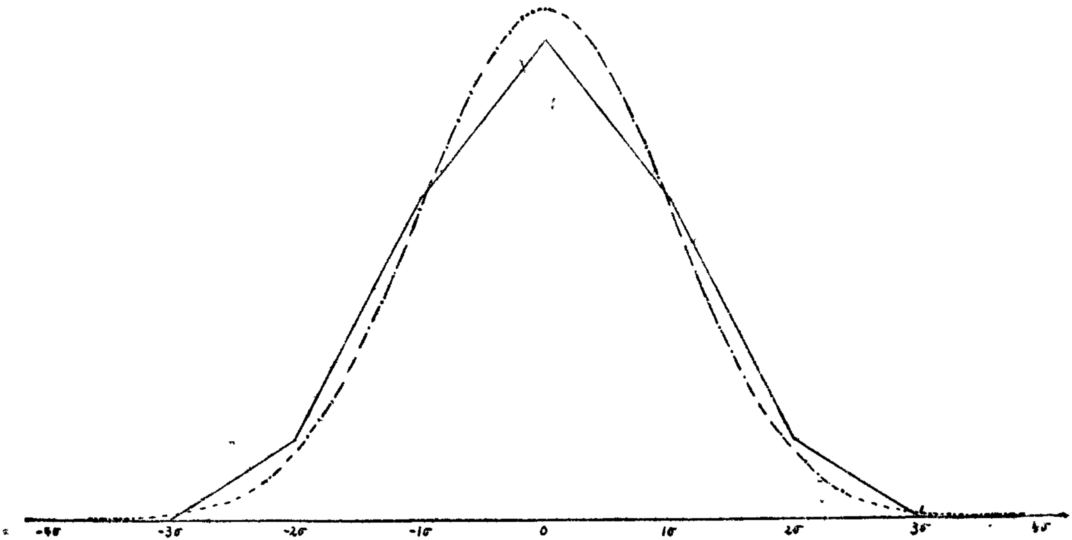


圖一 三種小學國語教科書中4269字之應用次數





圖二 二十五年浙江大學入學投考智力測驗成績



圖三　常態分配ト二項分配$(S+S)^4$

插表

表一　民國二十一年中等學校學生,清代進士及第,及鮑吳二氏人名錄中現代名人在十八省之分佈

省　區	中等學校學生	進士及第	現代名人
江　蘇	76,056	119	133
廣　東	69,802	10	95
四　川	55,801	5	23
河　北	54,919	17	107
湖　南	36,045	13	19
河　南	31,324	4	18
山　東	28,940	13	28
浙　江	22,785	81	108
廣　西	20,576	4	9
湖　北	17,291	14	29
福　建	15,519	10	48
安　徽	13,893	18	42
江　西	13,411	19	23
山　西	13,278	4	9
雲　南	11,745	0	15
貴　州	8,515	3	5
陝　西	6,064	2	6
甘　肅	2,784	0	2
總　計	498,748	336	719

中等學校學生根據教育部發表之統計，上海南京兩市併入江蘇省，北平市併入河北省，青島威海衞兩市併入山東省。

進士及第根據張耀翔氏研究，見心理雜誌選存,頁330—346,中華書局二十一年出版。

現代名人根據朱君毅氏研究，見心理雜誌選存,頁 319—330

凡三者不共有之省區均從略。

表二　民國元年至二十一年浙江省初等學校及中等學校歷年之學校數，學生數及經費數

年度	初等學校			中等學校		
	學校數	學生數	經費數	學校數	學生數	經費數
1	6,103	272,295	1,942,740	50	6,814	434,746
2	6,629	287,902	2,145,246	68	8,652	565,624
3	6,898	297,587	1,992,196	78	8,743	595,121
4	7,429	322,732	2,148,749	67	8,915	694,666
5	8,233	329,607	2,194,449	59	9,255	694,163
6	8,937	393,238	2,510,899	57	8,787	719,534
7	9,712	409,556	2,564,537	54	9,462	751,554
8	9,824	419,545	2,622,381	55	9,819	842,937
9	10,360	448,376	2,620,397	52	9,507	802,479
10	10,688	4 1,603	2,577,588	60	10,289	866,909
11	10,932	489,476	2,935,274	63	10,987	882,464
12	11,303	512,110	3,021,347	86	12,808	1,003,555
13	11,252	516,803	3,018,946	81	12,903	1,004,649
14	11,147	514,171	2,940,961	75	12,485	1,032,044
15	11,245	516,705	2,938,833	80	11,778	1,119,069
16	10,960	506,561	3,328,622	82	12,806	1,444,356
17	11,738	577,118	2,972,483	90	14,891	1,592,742
18	11,957	605,748	4,262,430	92	16,735	1,869,373
19	12,472	659,725	4,702,033	107	20,607	2,247,973
20	12,932	683,601	5,033,545	122	21,965	2,226,566
21	13,415	713,048	5,184,008	129	22,550	2,370,015

從浙江省教育廳發表之教育統計

表三　民國二十一年度全國專科以上學校之各科學生人數及百分比

學科	學生人數	百分比
法政	14,523	34.0
文藝	9,312	21.8
工程	4,439	10.4
理科	4,159	9.7
教育	3,368	7.9
商業	2,867	6.7
醫藥	1,852	4.3
農林	1,557	3.6
未分科	633	1.5
各科總計	42,710	100.0

從教育部發表之二十一年度教育統計

表四　民國二十五年國立浙江大學入學考試考生智力測驗之成績等級

等級	投考者全體		本校錄取者		亦被他校錄取者	
	人數	百分比	人數	百分比	人數	百分比
優	45	2.2	30	8.6	23	13.1
良	235	11.7	91	26.0	50	28.4
常	594	29.6	140	40.0	71	40.3
可	733	36.5	82	23.4	32	18.3
劣	361	18.0	7	2.0	—	—
不及格	40	2.0	—	—	—	—
總計	2008	100.0	350	100.0	176	100.0

表五　民國二十五年國立浙江大學入學考試考生智力測驗之成績分數

測驗分數	投考者全體	本校錄取者	亦被他校錄取者
100—124	2		
125—149	14		
150—174	24		
175—199	60	2	
200—224	132	1	
225—249	169	4	
250—274	217	13	3
275—299	254	29	12
300—324	262	40	17
325—349	248	47	20
350—374	179	43	16
375—399	167	50	35
400—424	99	34	13
425—449	79	31	18
450—474	57	26	19
475—499	25	15	12
500—524	17	12	9
525—549	3	3	2
	2008	350	176

補表

表六 7524 常用字之筆畫

筆畫	字數
1	5
2	30
3	58
4	101
5	140
6	195
7	324
8	479
9	498
10	608
11	621
12	661
13	649
14	555
15	562
16	451
17	393
18	292
19	254
20	180
21	144
22	110
23	79
24	53
25	35
26	15
27	17
28	4
29	4
30	3
31	1
32	1
33	2
	7524

所謂常用字指列入商務印書館國音學生字彙者

表七 8585 人之身長

身長(吋)	人數
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
	8585

從Yule, G. U., and Kendall, M. G., An Introduction to the Theory of Statistics.

表八　1024樣組之平均數分配

樣組之平均	每組1	每組2	每組4	每組5	每組10
00	512	256	64	32	1
05					
10					10
15					
20				160	45
25			256		
30					120
35					
40				320	210
45					
50		512	384		252
55					
60				320	210
65					
70					120
75			256		
80				160	45
85					
90					10
95					
100	512	256	64	32	1
樣組總數	1024	1024	1024	1024	1024
平均之平均	50	50	50	50	50
平均之均方差	2500	1250	625	500	250

注意樣組逐漸擴大時其平均數之參差程度逐漸減小，二者恰成反比例。

表九　1164男人與1456女人之身長

身長（吋）	男	女	男女合計
52.5		.5	.5
53.5		.5	.5
54.5		0	0
55.5		1	1
56.5		5	5
57.5		15	15
58.5		15.5	15.5
59.5	1	52	53
60.5	2.5	101	103.5
61.5	1.5	150	151.5
62.5	9.5	199	208.5
63.5	31	223	254
64.5	56	215	271
65.5	78.5	169.5	248
66.5	127	151.5	278.5
67.5	178.5	81.5	260
68.5	189	40.5	229.5
69.5	137	19.5	156.5
70.5	137	10	147
71.5	93	5	98
72.5	52.5	0	52.5
73.5	39	1	40
74.5	17		17
75.5	6.5		6.5
76.5	3.5		3.5
77.5	1		1
78.5	2		2
79.5	1		1
	1164	1456	2620

從 Fisher, R. A., Statistical Methods for Research Workers.
注意男女合計後成雙峯分配。

表十　三種小學國語教科書中4269字之應用次數

每字應用次數	字　　數	以10為組距的字數
1—　10	2105	2105
11—　20	559	559
21—　30	350	350
31—　40	180	180
41—　50	140	140
51—　60	100	100
61—　70	99	99
71—　80	73	73
81—　90	56	56
91—　100	53	53
101—　200	271	27.1
201—　300	83	8.3
301—　400	60	6.0
401—　500	38	3.8
501—　600	21	2.1
601—　700	17	1.7
701—　800	10	1.0
801—　900	6	.6
901—　1000	8	.8
1001—　2000	24	.24
2001—　3000	9	.09
3001—　4000	2	.02
4001—　5000	3	.03
5001—　6000	1	.01
6001—　7000	0	.00
7001—　8000	0	.00
8001—　9000	0	.00
9001—10000	1	.01
	4269	

根據王文新的小學分級字彙研究。

注意組距逐漸擴大，使各組之絕對字數難以比較，比較需用以 10 為組距的字數。

表十一 三種小學國語教科書中4269字之應用次數

每字應用次數	字數	每字應用次數	字數
1	601	1— 100	3715
2	377	101— 200	271
3	259	201— 300	83
4	203	301— 400	60
5	148	401— 500	38
6	117	501— 600	21
7	103	601— 700	17
8	111	701— 800	10
9	100	801— 900	6
10	86	901— 1000	8
1— 10	2105	1— 1000	4229
11— 20	559	1001— 2000	24
21— 30	350	2001— 3000	9
31— 40	180	3001— 4000	2
41— 50	140	4001— 5000	3
51— 60	100	5001— 6000	1
61— 70	99	6001— 7000	0
71— 80	73	7001— 8000	0
81— 90	56	8001— 9000	0
91—100	53	9001—10000	1
	3715		4269

注意組距屢次擴大後小組各數仍歸併大組並列。

表十二　三種小學國語教科書中4269字之應用次數

每字應用次數	應用次數之對數	字　數
1	—.2— .2	601
2— 3	.2— .6	636
4— 9	.6—1.0	782
9— 25	1.0—1.4	829
26— 63	1.4—1.8	623
64— 158	1.8—2.2	426
159— 398	2.2—2.6	232
399— 999	2.6—3.0	100
1000— 2511	3.0—3.4	30
2512— 6309	3.4—3.8	9
6310—15800	3.8—4.2	1
		4269

此表取材與表十表十一同。表十與表十一呈極端偏斜之分配，此表改用對數分組，偏斜度大減。若按對數求算術平均數，再查其逆對數，則所得爲各字應用次數之幾何平均數。凡極端偏斜之分配，幾何平均數比算術平均數爲合理而適用。

插表

表十三　各種中心量數及參差量數之計算

組限	中價	次數	累計	二次累計
99.5				
	112	2	2	2
124.5				
	137	14	16	18
149.5				
	162	24	40	58
174.5				
	187	60	100	158
199.5				
	212	132	232	390
224.5				
	237	169	401	791
249.5				
	262	217	618	1409
274.5				
	287	254	872	
299.5				
	312	262	2281	
324.5				
	337	248	874	
349.5				
	362	179	626	1704
374.5				
	387	167	447	1078
399.5				
	412	99	280	631
424.5				
	437	79	181	351
449.5				
	462	57	102	170
474.5				
	487	25	45	68
499.5				
	512	17	20	23
524.5				
	537	3	3	3
549.5				
		2008	2578	6854
			2281	2
			297	13708
			4859	4859
				18567

$$M=312+\frac{297}{2008}\times 25$$

$$=312+.148\times 25=315.70$$

$$Mdn=299.5+\frac{1004-872}{262}\times 25$$

$$=299.5+12.60=312.10$$

$$Q_1=249.5+\frac{502-401}{217}\times 25$$

$$=249.5+11.64=261.14$$

$$Q_3=374.5-\frac{502-447}{179}\times 25$$

$$=374.5-7.68=366.82$$

$$Q=\frac{366.82-261.14}{2}=52.84$$

$$M.D.=\frac{4859\times 25+.10\times 260}{2008}$$

$$=60.51$$

$$\sigma=\sqrt{\frac{18567}{2008}-.148^2}\times 25$$

$$=\sqrt{9.2246}\times 25=75.93$$

表十四　組距擴大後參差量數之膨脹（取材與表十三同）

組限	組中值	次數	累計	二次累計
99.5				
	124.5	16	16	16
149.5				
	174.5	84	100	116
199.5				
	224.5	301	401	517
249.5				
	274.5	471	872	
299.5				
	324.5	510	1389	
349.5				
	374.5	346	626	
399.5				
	424.5	178	280	402
449.5				
	474.5	82	102	122
499.5				
	524.5	20	20	20
549.5				
		2008	1028	1193
				2
			1389	2386
			−361	2417
			2417	4803

$$M=324.5-.1798\times 50$$
$$=315.51$$
$$Mdn=299.5+\frac{1004-872}{510}\times 50$$
$$=312.44$$
$$Q_1=249.5+\frac{502-401}{471}\times 50$$
$$=260.32$$
$$Q_3=399.5-\frac{502-280}{346}\times 50$$
$$=367.42$$
$$Q=\frac{367.42-260.22}{2}=53.60$$
$$M.D.=\frac{2417\times 50+12.06\times 264}{2008}$$
$$=61.27$$
$$\sigma=\sqrt{\frac{4803}{2008}-.1798^2}\times 50$$
$$=76.81$$

組限	組中值	次數	累計	二次累計
99.5				
	137	40	40	40
174.5				
	212	361	401	
249.5				
	287	733	441	
324.5				
	362	594	874	
399.5				
	437	235	280	325
474.5				
	512	45	45	45
549.5				
		2008	1199	410
			−441	2
			758	820
			1640	1640
				2460

$$M=287+.3775\times 75$$
$$=315.31$$
$$Mdn=311.20$$
$$Q_1=259.83$$
$$Q_3=371.47$$
$$Q=55.82$$
$$M.D.=64.39$$
$$\sigma=\sqrt{1.0826}\times 75=78.03$$

表十五　百分位等級之計算（取材與表十三同）

測驗分數	次　數	累　計	百分位等級
99.5			
	2	2	
124.5			.10
	14	16	
149.5			.80
	24	40	
174.5			1.99
	60	100	
199.5			4.98
	132	232	
224.5			11.55
	169	401	
249.5			19.97
	217	618	
274.5			30.78
	254	872	
299.5			43.43
	262	1134	
324.5			56.47
	248	1382	
349.5			68.82
	179	1561	
374.5			77.74
	167	1728	
399.5			86.06
	99	1827	
424.5			90.99
	79	1906	
449.5			94.92
	57	1963	
474.5			97.76
	25	1988	
499.5			99.00
	17	2005	
524.5			99.85
	3	2008	
549.5			

表十六　常態曲線之高度及其下面積

距離 x/σ	高度 y	左方面積 p	右方面積 q
0.0	.39894	.50000	.50000
0.1	.39695	.53983	.46017
0.2	.39104	.57926	.42074
0.3	.38139	.61791	.38209
0.4	.36827	.65542	.34458
0.5	.35207	.69146	.30854
0.6	.33322	.72575	.27425
0.7	.31225	.75804	.24196
0.8	.28969	.78814	.21186
0.9	.26609	.81594	.18406
1.0	.24197	.84134	.15866
1.1	.21785	.86433	.13567
1.2	.19419	.88493	.11507
1.3	.17137	.90320	.09680
1.4	.14973	.91924	.08076
1.5	.12952	.93319	.06681
1.6	.11092	.94520	.05480
1.7	.09405	.95543	.04457
1.8	.07895	.96407	.03593
1.9	.06562	.97128	.02872
2.0	.05399	.97725	.02275
2.1	.04398	.98214	.01786
2.2	.03547	.98610	.01390
2.3	.02833	.98928	.01072
2.4	.02239	.99180	.00820
2.5	.01753	.99379	.00621
2.6	.01358	.99534	.00466
2.7	.01042	.99653	.00347
2.8	.00792	.99744	.00256
2.9	.00595	.99813	.00187
3.0	.00443	.99865	.00135
3.2	.00238	.99931	.00069
3.4	.00123	.99966	.00034
3.6	.00061	.99984	.00016
3.8	.00029	.99993	.00007
4.0	.00013	.99997	.00003
4.2	.00006	.99999	.00001
4.4	.00002	.99999	.00001

表十七　相當於各重要性層級的t

自由度數	重要性層級 .05	.01	.001
1	12.706	63.657	636.619
2	4.303	9 925	31 598
3	3 182	5 841	12 941
4	2.776	4.604	8.610
5	2.571	4.032	6.859
6	2.447	3.707	5.959
7	2.365	3 499	5.405
8	2.306	3.355	5.041
9	2.262	3.250	4.781
10	2.228	3.169	4.587
11	2.201	3.106	4.437
12	2.179	3.055	4.318
13	2.160	3 012	4 221
14	2.145	2 977	4.140
15	2.131	2.947	4.073
16	2 120	2 921	4.015
17	2 110	2 898	3.965
18	2 101	2 878	3.922
19	2.093	2 861	3 883
20	2.086	2.845	3.850
21	2 080	2 831	3.819
22	2.074	2 819	3 792
23	2.069	2.807	3.767
24	2.064	2.797	3.745
25	2.060	2.787	3.725
30	2.042	2.750	3.646
40	2.021	2.704	3.551
60	2.000	2.660	3 460
120	1.980	2.617	3 373
∞	1.960	2.576	3.291

表十八甲　相當於 .05 重要性層級的均方差比率

n_2	n_1 1	2	4	6	12	24	∞
1	161.4	199.5	224.6	234.0	243.9	249.0	254.3
2	18.51	19.00	19.25	19.33	19.41	19.45	19.50
3	10.13	9.55	9.12	8.94	8.74	8.64	8.53
4	7.71	6.94	6.39	6.16	5.91	5.77	5.63
5	6.61	5.79	5.19	4.95	4.68	4.53	4.36
6	5.99	5.14	4.53	4.28	4.00	3.84	3.67
7	5.59	4.74	4.12	3.87	3.57	3.41	3.23
8	5.32	4.46	3.84	3.58	3.28	3.12	2.93
9	5.12	4.26	3.63	3.37	3.07	2.90	2.71
10	4.96	4.10	3.48	3.22	2.91	2.74	2.54
11	4.84	3.98	3.36	3.09	2.79	2.61	2.40
12	4.75	3.88	3.26	3.00	2.69	2.50	2.30
13	4.67	3.80	3.18	2.92	2.60	2.42	2.21
14	4.60	3.74	3.11	2.85	2.53	2.35	2.13
15	4.54	3.68	3.06	2.79	2.48	2.29	2.07
16	4.49	3.63	3.01	2.74	2.42	2.24	2.01
17	4.45	3.59	2.96	2.70	2.38	2.19	1.96
18	4.41	3.55	2.93	2.66	2.34	2.15	1.92
19	4.38	3.52	2.90	2.63	2.31	2.11	1.88
20	4.35	3.49	2.87	2.60	2.28	2.08	1.84
21	4.32	3.47	2.84	2.57	2.25	2.05	1.81
22	4.30	3.44	2.82	2.55	2.23	2.03	1.78
23	4.28	3.42	2.80	2.53	2.20	2.00	1.76
24	4.26	3.40	2.78	2.51	2.18	1.98	1.73
25	4.24	3.38	2.76	2.49	2.16	1.96	1.71
30	4.17	3.32	2.69	2.42	2.09	1.89	1.62
40	4.08	3.23	2.61	2.34	2.00	1.79	1.51
60	4.00	3.15	2.52	2.25	1.92	1.70	1.39
120	3.92	3.07	2.45	2.17	1.83	1.61	1.25
∞	3.84	2.99	2.37	2.09	1.75	1.52	1.00

n_1 爲較大均方差（居分子地位）的自由度數。

n_2 爲較小均方差（居分母地位）的自由度數。

表十八乙 相當於.001重要性層級的均方差比率

n_2 \ n_1	1	2	4	6	12	24	∞
1	405284	500000	562500	585937	610667	623497	636619
2	998.5	999.0	999.2	999.3	999.4	999.5	999.5
3	167.5	148.5	137.1	132.8	128.3	125.9	123.5
4	74.14	61.25	53.44	50.53	47.41	45.77	44.05
5	47.04	36.61	31.09	28.84	26.42	25.14	23.78
6	35.51	27.00	21.90	20.03	17.99	16.89	15.75
7	29 22	21.69	17.19	15.52	13.71	12.73	11.69
8	25.42	18.49	14.39	12.86	11.19	10.30	9.34
9	22 86	16.39	12.56	11.13	9.57	8.72	7.81
10	21.04	14.61	11.28	9.92	8.45	7.64	6.76
11	19.69	13.81	10.35	9.05	7.63	6.85	6.00
12	18.64	12.97	9.63	8.38	7.00	6.25	5.42
13	17.81	12.31	9.07	7.86	6.52	5.78	4.97
14	17.14	11.78	8.62	7.43	6.13	5.41	4.60
15	16.59	11.34	8.25	7.09	5.81	5.10	4.31
16	16.12	10.97	7.94	6.81	5.55	4.85	4.06
17	15.72	10.66	7.68	6.56	5.32	4.63	3.85
18	15.38	10.39	7.46	6.35	5.13	4.45	3.67
19	15.08	10.16	7.26	6.18	4.97	4.29	3.52
20	14.82	9.95	7.10	6.02	4.82	4.15	3.38
21	14.59	9.77	6.95	5.88	4.70	4.03	3.26
22	14.38	9.61	6.81	5.76	4.58	3.92	3.15
22	14.19	9.47	6.69	5.65	4.48	3.82	3.05
24	14.03	9.34	6.59	5.55	4.39	3.74	2.97
25	13.88	9.22	6.49	5.46	4.31	3.66	2.89
30	13.29	8.77	6.12	5.12	4.00	3.36	2.59
40	12.61	8.25	5.70	4.73	3.64	3.01	2.23
60	11.97	7.76	5.31	4.37	3 31	2.69	1.90
120	11.38	7.31	4.95	4.04	3.02	2.40	1.56
∞	10.83	6.91	4.62	3.74	2.74	2.13	1.00

n_1 爲較大均方差（居分子地位）的自由度數。

n_2 爲較小均方差（居分母地位）的自由度數。

表十九　五人兩次所批三本試卷的分數

試卷	第一次						第二次					
	閱卷者						閱卷者					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	87	81	76	67	64	75	75	67	68	59	56	65
2	64	60	52	45	44	53	54	52	48	45	36	47
3	41	39	28	26	21	31	39	31	28	28	19	29
平均	64	60	52	46	43	53	56	50	48	44	37	47

表十九甲　除去總平均(50)後的差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	37	31	26	17	14	25	25	17	18	9	6	15
2	14	10	2	—5	—6	3	4	2	—2	—5	-14	—3
3	—9	-11	-22	-24	-29	-19	-11	-19	-22	-22	-31	-21
平均	14	10	2	—4	—7	3	6	0	—2	—6	-13	—3

方差和數＝9966　　自由度數＝29

表十九乙　代表時間影響的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	3	3	3	3	3	3	—3	—3	—3	—3	—3	—3
2	3	3	3	3	3	3	—3	—3	—3	—3	—3	—3
3	3	3	3	3	3	3	—3	—3	—3	—3	—3	—3
平均	3	3	3	3	3	3	—3	—3	—3	—3	—3	—3

方差和數＝270　　自由度數＝1

表十九丙　代表試卷間參差的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	20	20	20	20	20	20	20	20	20	20	20	20
2	0	0	0	0	0	0	0	0	0	0	0	0
3	-20	20	-20	-20	-20	-20	-20	-20	-20	-20	-20	-20
平均	0	0	0	0	0	0	0	0	0	0	0	0

方差和數=8000　　自由度數=2

表十九丁　代表閱卷者間參差的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	10	5	0	—5	10	0	10	5	0	—5	-10	0
2	10	5	0	—5	-10	0	10	5	0	—5	-10	0
3	10	5	0	—5	-10	0	10	5	0	—5	-10	0
平均	10	5	0	—5	-10	0	10	5	0	—5	-10	0

方差和數=1500　　自由度數=4

表十九戊　代表時間，試卷，與閱卷者之間各種交互作用的差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	4	3	3	—1	1	2	—2	—5	1	—3	—1	—2
2	1	2	—1	—3	—1	0	—3	0	1	3	—1	0
3	—2	1	—5	—2	—2	—2	2	—1	1	6	2	2
平均	1	2	—1	—2	0	0	—1	—2	1	2	0	0

方差和數=196　　自由度數=22

表十九己　代表時間與試卷交互作用的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	2	2	2	2	2	2	—2	—2	—2	—2	—2	—2
2	0	0	0	0	0	0	0	0	0	0	0	0
3	—2	—2	—2	—2	—2	—2	2	2	2	2	2	2
平均	0	0	0	0	0	0	0	0	0	0	0	0

方差和數＝80　　自由度數＝2

表十九庚　代表時間與閱卷者交互作用的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	1	2	—1	—2	0	0	—1	—2	1	2	0	0
2	1	2	—1	—2	0	0	—1	—2	1	2	0	0
3	1	2	—1	—2	0	0	—1	—2	1	2	0	0
平均	1	2	—1	—2	0	0	—1	—2	1	2	0	0

方差和數＝60　　自由度數＝4

表十九辛　代表試卷與閱卷者交互作用的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	1	−1	2	−2	0	0	1	−1	2	−2	0	0
2	−1	1	0	0	0	0	−1	1	0	0	0	0
3	0	0	−2	2	0	0	0	0	−2	2	0	0
平均	0	0	0	0	0	0	0	0	0	0	0	0

方差和數＝40　　　自由度數＝8

表十九壬　代表第三級交互作用的一部份差數

	第一次						第二次					
	甲	乙	丙	丁	戊	平均	甲	乙	丙	丁	戊	平均
1	0	0	0	1	−1	0	0	0	0	−1	1	0
2	1	−1	0	−1	1	0	−1	1	0	1	−1	0
3	−1	1	0	0	0	0	1	−1	0	0	0	0
平均	0	0	0	0	0	0	0	0	0	0	0	0

方差和數＝16　　　自由度數＝8

表二十甲　十四人第一次所批十五本考卷之分數

考卷	閱卷者													
	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
1	31	43	33	30	34	44	46	31	48	43	42	31	41	50
2	26	37	32	41	37	29	43	31	32	34	41	32	31	42
3	29	55	35	40	43	41	49	43	51	40	46	36	45	46
4	38	52	36	63	61	48	52	52	53	40	52	50	39	43
5	32	41	33	42	47	40	45	47	48	42	39	32	38	48
6	35	37	48	53	53	62	51	41	46	38	49	48	50	47
7	46	53	36	57	48	60	58	58	53	46	45	53	50	51
8	21	45	27	37	41	39	36	36	39	35	41	25	38	45
9	26	37	30	27	29	38	44	43	38	29	40	24	24	37
10	39	39	35	39	40	50	55	57	52	42	47	45	38	40
11	28	50	37	48	48	41	49	55	49	36	46	38	41	52
12	38	68	35	60	48	49	61	70	69	43	53	44	57	62
13	24	35	30	34	31	32	40	38	41	37	43	38	40	45
14	39	30	33	40	44	58	53	47	44	43	41	34	27	41
15	42	55	47	45	70	60	65	53	58	51	53	52	49	53

表二十乙　十四人第二次所批十五本考卷之分數

考卷	閱卷者													
	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
1	29	43	32	55	40	43	53	61	40	43	44	36	41	40
2	27	34	23	43	34	41	35	35	29	29	40	34	32	36
3	31	48	29	36	39	49	46	45	48	39	35	34	49	44
4	45	49	34	57	51	45	38	39	46	33	43	49	40	40
5	27	45	24	49	40	52	48	45	42	38	45	40	39	50
6	48	27	31	49	51	66	56	50	44	43	52	40	53	45
7	43	54	37	56	43	71	55	53	44	42	50	48	57	50
8	34	33	25	47	36	31	35	31	34	30	37	30	39	48
9	30	30	21	20	22	34	43	35	32	25	35	28	24	27
10	41	46	32	40	37	44	42	37	49	39	39	42	43	43
11	29	43	27	38	40	45	48	39	44	30	43	36	40	42
12	40	61	24	63	41	46	57	62	63	28	45	53	53	54
13	29	32	16	32	25	42	41	33	38	29	40	40	37	38
14	40	35	28	29	33	61	46	48	42	34	40	43	26	28
15	49	50	51	56	54	52	48	50	51	43	48	41	48	37

從 Hartog, Sir Philip, and Rhodes, E. C., The Marks of Examiners.

表二十一甲　相關表之例一

r=.00		X											
		0	1	2	3	4	5	6	7	8	9	10	合計
Y	0	1	10	45	120	210	252	210	120	45	10	1	1024
	1	10	100	450	1200	2100	2520	2100	1200	450	100	10	10240
	2	45	450	2025	5400	9450	11340	9450	5400	2025	450	45	46080
	3	120	1200	5400	14400	25200	30240	25200	14400	5400	1200	120	122880
	4	210	2100	9450	25200	44100	52920	44100	25200	9450	2100	210	215040
	5	252	2520	11340	30240	52920	63504	52920	30240	11340	2520	252	258048
	6	210	2100	9450	25200	44100	52920	44100	25200	9450	2100	210	215040
	7	120	1200	5400	14400	25200	30240	25200	14400	5400	1200	120	122880
	8	45	450	2025	5400	9450	11340	9450	5400	2025	450	45	46080
	9	10	100	450	1200	2100	2520	2100	1200	450	100	10	10240
	10	1	10	45	120	210	252	210	120	45	10	1	1024
	合計	1024	10240	46080	122880	215040	258048	215040	122880	46080	10240	1024	1048576

表二十一乙　相關表之例二

r＝.20		X 0	1	2	3	4	5	6	7	8	9	10	合計
Y	0	4	32	112	224	280	224	112	32	4			1024
	1	32	264	960	2016	2688	2352	1344	480	96	8		10240
	2	112	960	3652	8096	11536	10976	7000	2912	736	96	4	46080
	3	224	2016	8096	19072	29120	60016	21056	9856	2912	480	32	122880
	4	280	2688	11536	29120	47824	53312	40768	21056	7000	1344	112	215040
	5	224	2352	10976	30016	53312	64288	53312	30016	10976	2352	224	258048
	6	112	1344	7000	21056	40768	53312	47824	29120	11536	2688	280	215040
	7	32	480	2912	9856	21056	30016	29120	19072	8096	2016	224	122880
	8	4	96	736	2912	7000	10976	11536	8096	3652	960	112	46080
	9		8	96	480	1344	2352	2688	2016	960	264	32	10240
	10			4	32	112	224	280	224	112	32	4	1024
	合計	1024	10240	46080	122880	215040	258048	215040	122880	46080	10240	1024	1048576

表二十一丙　相關表之例三

r = .40		X											
		0	1	2	3	4	5	6	7	8	9	10	合計
Y	0	16	96	240	320	240	96	16					1024
	1	96	640	1824	2880	2720	1536	480	64				10240
	2	240	1824	6000	11136	12720	9120	3984	960	96			46080
	3	320	2880	11136	24320	33024	28800	16000	5376	960	64		122880
	4	240	2720	12720	33024	53120	55296	37440	16000	3984	480	16	215040
	5	96	1536	9120	28800	55296	68352	55296	28800	9120	1536	96	258048
	6	16	480	3984	16000	37440	55296	53120	33024	12720	2720	240	215040
	7		64	960	5376	16000	28800	33024	24320	11136	2880	320	122880
	8			96	960	3984	9120	12720	11136	6000	1824	240	46080
	9				64	480	1536	2720	2880	1824	640	96	10240
	10					16	96	240	320	240	96	16	1024
	合計	1024	10240	46080	122880	215040	258048	215040	122880	46080	10240	1024	1048576

表二十一丁　相關表之例四

r=.60		X 0	1	2	3	4	5	6	7	8	9	10	合計
Y	0	64	256	384	256	64							1024
	1	256	1408	3072	3328	1792	384						10240
	2	384	3072	9408	14592	12288	5376	960					46080
	3	256	3328	14592	31488	37632	25344	8960	1280				122880
	4	64	1792	12288	37632	62208	59136	32000	8960	960			215040
	5		384	5376	25344	59136	77568	59136	25344	5376	384		258040
	6			960	8960	32000	59136	62208	37632	12288	1792	64	215040
	7				1280	8960	25344	37632	31488	14592	3328	256	122880
	8					960	5376	12288	14592	9408	3072	384	46080
	9						384	1792	3328	3072	1408	256	10240
	10							64	256	384	256	64	1024
	合計	1024	10240	46080	122880	215040	258040	215040	122880	46080	10240	1024	1048576

表二十一戊　相關表之例五

r=.80		X 0	1	2	3	4	5	6	7	8	9	10	合計
Y	0	256	512	256									1024
	1	512	3072	4608	2048								10240
	2	256	4608	15616	18432	7168							46080
	3		2048	18432	45056	43008	14336						122880
	4			7168	43008	82432	64512	17920					215040
	5				14336	64512	100352	64512	14336				258040
	6					17920	64512	82432	43008	7168			215040
	7						14336	43008	45056	18432	2048		122880
	8							7168	18432	15616	4608	256	46080
	9								2048	4608	3072	512	10240
	10									256	512	256	1024
	合計	1024	10240	46080	122880	215040	258040	215040	122880	46080	10240	1024	1048576

表二十一己　相關表之例六

r=1.00		X											
		0	1	2	3	4	5	6	7	8	9	10	合計
Y	0	1024											1024
	1		10240										10240
	2			46080									46080
	3				122880								122880
	4					215040							215040
	5						258040						258040
	6							215040					215040
	7								122880				122880
	8									46080			46080
	9										10240		10240
	10											1024	1024
	合計	1024	10240	46080	122880	215040	258040	215040	122880	46080	10240	1024	1048576

表二十二　曲線相關之一例

Y	X .05	.15	.25	.35	.45	.55	.65	.75	.85	.95	合計
.95	105	115	126	137	53						536
.85					98	166					264
.75							185	30			215
.65								185			185
.55									166		166
.45									98	53	151
.35										137	137
.25										126	126
.15										115	115
.05										105	105
-.05										105	105
-.15										115	115
-.25										126	126
-.35										137	137
-.45									98	53	151
-.55									166		166
-.65								185			185
-.75							185	30			215
-.85					98	166					264
-.95	105	115	126	137	53						536
合計	210	230	252	274	302	332	270	430	528	1072	4000

表二十三甲　一千人按照國籍與所好音樂國別的分配

國籍	所好音樂國別					
	英	法	德	意	西	合計
英	32	16	75	47	30	200
法	10	67	42	41	40	200
德	12	23	107	36	22	200
意	16	20	44	76	44	200
西	8	53	30	43	66	200
合計	78	179	298	243	202	1000

表二十三乙　根據機遇假設的分配

國籍	所好音樂國別					
	英	法	德	意	西	合計
英	15.6	35.8	59.6	48.6	40.4	200
法	15 6	35.8	59.6	48.6	40.4	200
德	15.6	35.8	59.6	48.6	40.4	200
意	15.6	35.8	59.6	48.6	40.4	200
西	15.6	35.8	59.6	48.6	40.4	200
合計	78	179	298	243	202	1000

表二十三丙　實得次數與假設次數的差別

國籍	所好音樂國別					
	英	法	德	意	西	合計
英	16.4	—19.8	15.4	— 1.6	—10.4	0
法	— 5.6	31.2	—17.6	— 7.6	— .4	0
德	— 3.6	—12.8	47.4	—12.6	—18.4	0
意	.4	—15.8	—15.6	27.4	3.6	0
西	— 7.6	17.2	—29.6	— 5.6	25.6	0
合計	0	0	0	0	0	0

表二十三丁　X^2的計算

國籍	所好音樂國別					
	英	法	德	意	西	合計
英	17.24	10.95	3.98	.05	2.68	34.90
法	2.01	27.18	5.20	1.19	.00	35.58
德	.83	4.58	37.70	3.27	8.38	54.76
意	.01	6.97	4.08	15.45	.32	26.83
西	3.70	8.26	14.70	.65	16.22	43.53
合計	23.79	57.94	65.66	20.61	27.60	195.60

表二十四　相當於各重要性層級的X^2

自由度數	重要性層級		
	.05	.01	.001
1	3.841	6.635	10.827
2	5 991	9.210	13 815
3	7.815	11.341	16.268
4	9.488	13.277	18.465
5	11.070	15.086	20.517
6	12.592	16.812	22.457
7	14.067	18 475	24.322
8	15.507	20 090	26.125
9	16 919	21.666	27 877
10	18.307	23.209	29.588
11	19.675	24.725	31.264
12	21.026	26.217	32.909
13	22.362	27.688	34.528
14	23.685	29 141	36.123
15	24.996	30.578	37.697
16	26 296	32.000	39.252
17	27.587	33 409	40.790
18	28 869	34 805	42 312
19	30.144	36.191	43.820
20	31.410	37.566	45.315
21	32.671	38.932	46.797
22	33.924	40.289	48.268
23	35.172	41.638	49.728
24	36.415	42.980	51.179
25	37.652	44.314	52.620
26	38.885	45.642	54.052
27	40.113	46.963	55.476
28	41.337	48.278	56.893
29	42.557	49.588	58.302
30	43.773	50.892	59.703

中西名詞對照

Analysis of variance 方差分析
average 平均
average deviation 平均差

Biased error 偏的誤差
bimodal distribution 雙峯分配
binomial distribution 二項分配

Central tendency, measures of 中心量數
chance 機會，偶然
chance error 偶然的誤差
class interval 組距
class limits 組限
continuous data 綿續的材料
correlation coefficient 相關係數
correlation ratio 相關比率
curvilinear correlation 曲線相關
curvilinear regression 曲線消長

Dichotomy 二分法
decile 十分位
discrete data 間斷的材料
dispersion, measures of 參差量數
distribution 分配

Error 誤差
error variance 均方誤差
estimate 估計

Freedom, degrees of 自由度
frequency 次數
frequency distribution 次數分配
frequency polygon 次數多邊形

Galton, Sir Francis 高耳頓

Histogram 直方圖

Interaction 交互作用
interview 訪問

Kurtosis, measures of 峻峭量數

Leptokurtic distribution 陡峻分配
level of significance 重要性層級
lower quartile 下四分位

Mean 平均
mean deviation 平均差
mean square deviation 均方差
median 中位
mesokurtic distribution 峻峭度適中的分配

mode 型範
moment 矩
multiple correlation 複相關
multiple regression 多項消長

Normal distribution 常態分配

Ordered data 有秩序的材料

Partial correlation 偏相關
Pearson, Karl 皮耳生
percentile 百分位
percentile rank 百分位等級
platykurtic distribution 平頂分配
primary data 初級材料
product-moment 積矩
purposive sampling 立意取樣

Qualitative data 品質的材料
quantitative data 數量的材料
quartile 四分位
quartile deviation 四分差
questionnaire 問卷
quintile 五分位

Random sampling 隨機取樣
range 限域
rank-difference 等差
rectilinear correlation 直線相關
rectilinear regression 直線消長
regression 消長,回歸

Sample 樣組
sampling 取樣
secondary data 二級材料
Sheppard's correction 謝伯的校正
significance, test of 重要性考驗
significant discrepancy 重要的不符合處
simple correlation 單相關
skewness, measures of 偏斜量數
Spearman, C. 士比門
standard deviation 標準差
standard error 標準誤差
standard score 標準分數
statistical data 統計材料
statistical method 統計方法
statistics 統計
stratified sampling 層級化取樣
systematic error 系統的誤差

Total correlation 全相關
Tschuprow's coefficient 朱普羅的係數,朱氏係數

Unbiased error 不偏的誤差
unordered data 無秩序的材料
upper quartile 上四分位

Variability, measures of 參差量數
variance (從平均的 均方差
variation, coefficient of 參差係數