

教育講叢書

陳選善主編

教育統計學

沈有乾著

世界書局印行

書叢話講育教

編主善選陳

話講學計統育教

著編乾有沈

行印局書界世

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

版權所有
不准翻印

編著者 沈有乾

主編者 陳選善

發行人 李煜瀾

出版者 世界書局

發行所 世界書局

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，定名為教育講話叢書。
- 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，以為專門研究之基礎。
- 三 本叢書內容着重在歷史背景之敘述，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，俾啟發讀者思想，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為對象。

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』。本書寫作的動機即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尚未提到過的：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	一
第二講	統計的材料與方法	一二
第三講	觀察與取樣	二一
第四講	從混亂中建設秩序	三一
第五講	三種常用的中心量數	三九
第六講	參差量數	五〇
第七講	個體在團體中的地位	六二
第八講	次數分配的形式——偏斜度與峻峭度	七二
第九講	二項分配與常態分配	八〇
第十講	取樣誤差的估計與重要性的考驗	九〇
第十一講	方差分析法	九九
第十二講	相關量數	一一〇
第十三講	相關問題續論	一二二

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，更為重要階級。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……者之數目為低。」讀者如不能在前段推論中發現任何破綻，也不足怪。因為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，其所引為理由的事實，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，真是一大恥辱。

從前面所舉例子，可見極簡單的統計數字尚且會被人誤用，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尚欠努力。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？即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，除非採取不關痛癢的態度，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 ($Q = \frac{2q}{n}$) 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。第一題因為難易適中，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，他們認為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，他們認為普通以最高分數的百分之六十為界限，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。因為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，普通一定認為試題太難，或學生太差，成為很特殊的結果。

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，他們往往並不因為某生超過了六十分才給他及格，或因為某生還不夠六十分就給他不及格。他們的辦法是倒過來的，因為某生應當及格才給他六十以上的分數，或因為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・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，所以在百分中佔二十分，一百人共得一百分，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，平均還不到五十分，參差愈大，則平均分數愈低。爲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，即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，一個因為常用，所以識的人多，另一個因為不常用，所以識的人少。又如兩件相似的史地事實，一件因為教師認為重要而認真教學，所以了解與記憶的人多，另一件因為教師認為次要而約略提及，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。

第二種難易並不是真的難易。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。

第一種難易是真的難易，因為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。茲以加法位數為例，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。』

以上的議論是不错的。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，能謹慎反省，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，即使勉強學得統計方法的皮毛，結果也只是自誤誤人而已。