

數據素養

黃文璋

國立高雄大學統計學研究所

1 前言

大約從兩、三年前起,“大數據”(Big data)在台灣開始被廣泛重視。眾裡尋他千百度,政府更有如發現治國良方,緊抓不放。

以行政院國家發展委員會為例,於民國103年6月18日委託舉辦“網入民意—大數據下的輿情探索論壇”;又於民國104年1月30日,先是訂出“政學合作運用巨量資料分析協助政府施政之應用領域規劃”,接著召開分組工作會議,在“規劃”中,便說是依據“104年1月13日‘行政院國家資訊通信發展推動小組—研商運用大數據分析進行前瞻施政規劃會議’的決議辦理”。從政府、學術界到民間,可說此起彼落,即使在你意想不到的地方,也不時會出現“大數據”三字。像是在“泛科學”網站上,有篇文章的標題便取為“大數據與人道援助—群眾標記應用”。要談人道援助,也得先來個大數據。標題加上此三字,說不定可讓點擊次數提高不少。演變至今,獲知日本愛媛縣為了促婚,搬出大數據(見民國104年3月17日中國時報“大數據牽紅線”一文),也就不足為奇了。另外,有次在報上看到一則“國立臺灣圖書館”2014年讀者借閱統計之新聞,其中提到最熱

門的圖書為“偷書賊”(The Book Thief),最受歡迎的影片則為“冰雪奇緣”(Frozen)。覺得有趣,遂轉寄給在某大學圖書館服務的一個朋友參考,還順便問他們館裡是否有類似的統計?兩星期後收到回信,信中先表示他們原本對系統統計功能不是很熟,花了一些時間研究,以索書號進行分析,終於搜尋出去年(2014年)的借閱結果。我不過轉寄一新聞,他們便認真地去探究,真是相當用心。信裡又說,系統似乎也可將索書號與讀者的類型做交叉分析,只是他們的電腦跑太久後便當掉了,因此一時無法得到報表。至於熱門借閱和預約的排行榜,過去並沒有做到這麼詳細,但未來會找到處理的方法。接著說,“在現今big data應用的趨勢下,我們的確應該更加關注讀者的借閱行為”。你看!關注讀者的借閱行為,也要帶上是與大數據有關。

大數據如此重要,善於掌握先機者,當然趨之若鶩。好些學術機構,遂陸續成立“大數據中心”。有關大數據的研究,已成為顯學。而眾人開口閉口大數據,言談間,大數據彷彿成為有如瑞士刀(Schweizer Messer)的萬能工具。似乎只要祭起大數據,任何疑難雜症,都能迎刃而解。不過大數據如此被推崇,倒也非人人認同。如中國時報民國104年2月21日的社論標題,便為“令人憂心的大數據治國”。

大數據又稱巨量資料、海量資料,或大資料。像是東吳大學已成立一校級的“海量資料分析研究中心”,且“巨量資料管理學院學士學位學程”也於104學年度起開始招生。雖有不同的稱呼,但都不如大數據響亮。上個世紀的九十年代開始,興起“資料探勘”(Data Mining)的研究。面對資料,如何從中挖掘出有用的資訊?這自然是門重要的學問。資料探勘相當於大數據的前身,前者雖也曾引起一些風潮,卻遠遠比不上後者的全民化。實在是大數據三字太簡明易懂了,不像“探勘”二字,隱含著科學或工程,讓人覺得其中有專業存在,因而不敢輕易一探究竟。至於大數據,不過是“大”而已,多麼可親!因此迅即人人琅琅上口了。

何謂大數據?首先,今日的數據,並不是只指數字而已。舉凡文

字、數字、符號、圖形、聲音，及影像等，都是數據，其中都含有資訊。依“維基百科”的說明，“大數據幾乎無法使用大多數的資料庫管理系統處理，而必須使用在數十、數百甚至數千台伺服器上，同時平行運行的軟體”。又說“大數據由巨型資料集組成，這些資料集大小常超出人類在可接受時間下的收集、度用(curation, 源自於拉丁文curare, 照料的意思)、管理和處理能力。”簡單講，大數據就是一台機器存放不下的資料量。當然隨著科技進步，大數據的大小，也會隨之改變。甚至大數據的界定，也與持有資料的機構之能力有關。有人以為天大地大的數據，他人說不定完全未放眼裡。至於數據大小的衡量，至少包含量(volume, 指資料大小)、速度(velocity, 指資料輸入輸出的速度), 及多樣化(variety, 指涉及的變數)等三方面。

不論資料探勘或大數據，所做的工作，大抵就是數據分析。數據分析很難嗎？只要將數據輸入電腦，有這麼多現成的軟體，不就立即得到各個面相的統計結果，何難之有？的確沒錯，很輕易便可得到要多少有多少的圖表，再發揮一些想像力加以說明，便是一份有模有樣的報告。只是如果不具備數據素養，就根本不知在分析些什麼。所得到的結果，精確度令人存疑不說，甚至可能扭曲一些現象，導致連參考價值都失去。因此處理數據者，若即使基本的數據素養都缺乏，如何侈言大數據？

大學入學考試中心(簡稱大考中心)，在其網頁上寫著，成立的目的是“研究改進大學入學制度與技術並辦理大學入學考試”。大考中心每年辦理學科能力測驗(簡稱學測)、指定科目考試(簡稱指考)，及高中英語聽力測驗(簡稱英聽)，掌握高中生入大學的管道。想進心目中理想大學的高中生，三更燈火五更雞，然後以大考中心提供的試題來驗收成效，再依大考中心提供的規則換算成績。一切都是大考中心說了算，大考中心可說相當威權，但夠不夠權威呢？

每次考試結束，大考中心便獲得一筆數據。不過十餘萬考生，且數據單純，僅是數字而已，因此怎麼看都不能算是大數據。雖不是大數據，那

大考中心處理的如何？頗令人好奇。大考中心歷次考試的統計，皆刊登於其網站，資料公開，容易取來檢視。底下我們便以大考中心舉辦的英聽及學測為例，來看具備基本的數據素養，究竟是否容易？

2 104學年度第一次英聽

英聽正式考試是從102學年度起，每年兩次，每次皆分上午及下午兩個場次。各場次皆包含40題選擇題，成績的表示採等級制，依序為A、B、C、及F，共4個等級。大考中心對各等級之說明如下：

- A
 - 1 能幾乎完全聽懂以高中階段詞彙及句型構成之敘述或問句。
 - 2 能幾乎完全理解訊息豐富、主題多樣化的對話內容及相關訊息，並作出適當推論。
 - 3 能幾乎完全理解長篇、語意間接且具引申含意之陳述內容，並作出適當推論。
- B
 - 1 能大致聽懂以高中階段詞彙及句型構成之敘述或問句。
 - 2 能大致理解訊息單純、主題生活化的對話內容及相關訊息，並作出簡單推論。
 - 3 能理解簡短、語意直接之陳述內容，並作出簡單推論。
- C
 - 1 能約略聽懂以高中階段詞彙及句型構成之敘述或問句。
 - 2 能約略理解訊息單純且明確、語意直接之日常生活對話內容。

3 能約略理解簡短、語意直接之陳述內容的主旨。

F 1 僅能聽懂少部分高中階段詞彙及句型構成之敘述或問句。

2 僅能聽懂少部分日常生活對話內容。

3 僅能聽懂少部分簡短陳述。

其中B的第3點，漏掉“大致”二字，且是從102學年度起便漏掉至今。會不會是特地不寫？不像是，因沒什麼道理只有該處缺一形容詞，大考中心似不夠嚴謹。由大考中心的說明顯示，在聽力及理解力方面，A、B、C及F，分別代表幾乎完全、大致、約略，及少部分等4個層次。數學裡有“幾乎到處”(almost everywhere)的概念。在實數軸上，一函數若幾乎到處為0，表不等於0的地方之“測度”(不妨想成長度)為0。此時若視該函數為0，並不至於有何太大的不妥。在3年來已舉辦的6次英聽中，其中5次至少有10%以上的考生得A等級(102學年度的第一次甚至有超過20%)，比率不低。另一方面，每年還有不算低比率的考生，當年第一次得A等級，第二次的等級卻下降，甚至降到F。給人的感覺是，拿到A等級，離什麼幾乎完全聽懂，及幾乎完全理解，都還有段不小的距離。大考中心對語意的掌握，與一般的習慣，似乎差異不小。

104學年度，首次有大學採計英聽為入學門檻，自此高中生更辛苦了。大考中心於民國103年11月1日及104年2月3日，分別舉辦104學年度第一次英聽及第二次英聽。大考中心說，成績在計算上包括兩部分。首先，透過測驗統計中的等化(Equating)程序，將原始分數轉換至歷次各場次考試相同的量尺上。其次，依據事先設定的各等級切截分數(Cutoff Score)，計算考生成績等級。我們對大考中心分數的轉換，以及等級的切截，一向很不安，常擔心其思考邏輯，底下將陸續說明。

美國教育測驗服務社(Education Testing Service, 縮寫ETS)舉辦的托福考試(Test of English as a Foreign Language, 縮寫TOEFL),

歷經幾次變革，目前採網路化測驗(Internet-Based Test, 縮寫iBT), 測驗項目包含閱讀(reading)、聽力(listening)、口說(speaking)及寫作(writing)。每項30分，總分120分。托福考試的聽力可有30分，即分成30級，大考中心的英聽卻只分4等級。又同是大考中心所舉辦，學測各科皆分15級，看不出英聽只分少少的4等級之原因。難道是因執著於幾乎完全、大致、約略，及少部分等4個形容詞嗎？如果是，學測的15級又如何說明？

在104學年度英聽第一次考試的“成績統計報告”中，指出共有128,576人報考，實際到考人數為127,014人，到考率為98.79%。將127,014除以128,576，得到0.98785...，顯然大考中心的百分比，乃採四捨五入至小數第2位。在該報告中，提供表1至表4等4個統計表。對數字若有些敏感度，看到表1中第3行3.21、36.34、43.42，及17.03，那4個百分比之和恰為100%，便會覺得有些可疑。要知數字太吻合，反易讓人產生警覺心。因4個值皆經過四捨五入，但和卻仍為1，是有點巧。檢查一下。A等級的人數21,624除以127,014，得0.17024...，因此應為17.02%才對，而非17.03%。原來為了使和為1，大考中心動了些手腳，造成報告不實。以為神不知鬼不覺，不料禁不起被人拿計算機敲一下的檢驗。會不會是筆誤？當然不是。因表1及表3中，皆有數處百分比不正確，即都被大考中心調整過。如同在表1中，最右側累積百分比那行，由於累積至等級B，人數共105,390，除以127,014，得0.82975...，因此應為82.98%才對，而非82.97%。大考中心應是為了使此百分比，等於B(43.42%)、C(36.34)，及F(3.21)，3個百分比的和。眾所皆知，當分數化為小數時，常會有捨入誤差(rounding error)。給一例子。假設有90人參加英聽，且A、B、C，及F，4個等級的人數分別為30、30、30，及0。如大考中心百分比的表示，便應為33.33%、33.33%、33.33%，及0%。豈可為了使和為100%，而硬生生將A等級改為33.34%？一刀切下去，可差到1個等級，對考生影響甚大。大考中心對於計算，豈可如此隨意？由此可看出，大考中心的數據素

養相當不足。

表4以各等級人數的多寡，來比較104學年度與103學年度英聽的第一次考試。這是6份報告中，唯一一次做此比較，卻讓人感到有些可笑，要知畫蛇添足向來是不好的。表4顯示，104學年度各等級的人數皆比103學年度多，但這能說明什麼？小明對小華炫耀，我衣服比你多，球鞋比你多，...，小華一路吃驚，終於靈光一閃，說“我被當的科目比你多”。兩次考試到考人數不同，104學年度較103學年度，增加17,738人，約增加16.23%，不算少的比率，因此各等級人數隨之增加，不過是合理而已。不妨來計算百分比。104學年度的A、B、C、及F，4等級所占百分比，各約為17.02%、43.42%、36.34%，及3.21%(合計99.99%)；103學年度的A、B、C、及F，4等級的百分比，各約為19.53%、45.14%、33.03%，及2.29%(合計99.99%)。對比較好的等級A及B，104學年度皆較少，而比較差的等級C及F，104學年度皆較多。此顯示104學年度的成績乃遜於103學年度，豈可用各等級的人數皆較多來混淆？至於有這樣的差異，究竟是由於誤差，還是夠“顯著”？可經由統計檢定來判別。

但無論如何，這項比較的意義是什麼？事實上，大考中心雖在每份成績統計報告中，皆強調上、下午場的“測驗信度”(test reliability, 指測驗分數的一致性或穩定性)良好，但不同次呢？以得到A等級的比率為例，3年下來的6次英聽，分別約有23.53%、12.58%、19.53%、8.12%、17.02%，及13.02%，波動可說極大。原因為何？由報告中看不出來。托福成績可不必敘明是那一次考的，也就是測驗信度極高。對於英聽的測驗信度，是否真如大考中心所說的很高，由數據並無法讓人產生信心。因此完全不理解大考中心提供表4的目的。

3 104學年度第二次英聽

104學年度第二次英聽的“成績統計報告”裡，亦給出4個統計表，其中

表1及表2與第一次類似,表3及表4則第一次沒有。我們給出表1、表3,及表4,且重新命名為表5、表6,及表7。

再度,表5中的累積百分比那行,44.66%應為44.65%。

比較表5與表1知,得A等級的比率,第二次較第一次少,從約17.02%(注意,我們已指出表1中的17.03%有誤)降至約13.02%;B等級從約43.42%略降至約42.32%;C等級從約36.34%增至約41.75%;F等級則從約3.21%略降至約2.91%。表面上看起來,第二次的成績大致稍差。是成績退步嗎?不盡然。不能只是這樣比,這點猜想大考中心亦知道,所以他們才提供表6及表7,以比較兩次考試考生的表現。為什麼不能只是這樣比各等級比率?大考中心指出,第二次實際到考的66,260人中,有56,790人第一次亦到考,約占85.98%;且約占第一次到考127,014人的44.85%。不考第一次卻考第二次的9,290(= 66,260 - 56,970)位考生,合理的猜測,程度較可能偏差,因第一次考的時候,或許覺得準備尚不夠周全,於是先略過。又依規定,兩次英聽的成績可擇優採計。所以第一次已考等級A者,當然較少會想考第二次,因毫無必要。由此知在第二次考試,考生中可能較缺乏好手。由於兩次考試,考生程度的分布,可能差異不小,所以整體成績有差並不奇怪,不表示考生考不好。

本來不考白不考,但考試畢竟還是花錢且花時間,如果大考中心只給出表6及表7,便不算提供太多有價值的資訊。不過表7仍顯示一值得留意的現象。兩次考試僅間隔短短的3個月,便有92人可由C等級進步兩等級至A;有44人可由F等級進步兩等級至B進,甚至有6人進步三等級至A。若參加托福考試,恐怕很難如此大躍進,每次最多都只進步一點點。英聽的測驗效度(test validity,指能正確測量到所欲測量的特質或功能之程度)到底多高?未免令人懷疑。另外,由於等級分的太粗略,雖表6給出兩次皆到考者,有不低的約28.88%進步比率,但究竟進步者原始成績是落在該等級的上層、中層或下層?此一對考生應屬有用的資訊,由“成績統計報告”中,並無法獲知。

扣掉重覆考，兩次考試共136,304(= 127,014 + 66,260 - 56,970)人到考。因成績擇優採計，由表1及表7，經過簡單的計算(讀者不妨自行嘗試)，可得兩次共136,304位考生中，有30,228人得A等級，56,848人得B等級，45,432人得C等級，及3,796人得F等級。A、B、C及F，4等級分別約占22.18%，41.71%，33.33%，及2.78%。13萬多的考生，最終各等級所占比率，大考中心是否認為合理？由於大考中心未做此統計，“成績統計報告”中，除給出4個表外，說明也不是很多，因此大考中心對4等級分佈的預先設定為何，並不得而知。說不定他們沒什麼特別的想法，只是承辦考試，考完加減乘除一番，得到一些數值，便完成一份成績統計報告。

曾有某大學教務處招生組，於該校研究所碩士班招生考試報名截止後，向全校各系所寄出報名人數統計表。其中錄取率那項，有些研究所超過100%。不論那種考試，除非最後有罕見的增額錄取等情事，否則錄取率最多只能達到100%，豈會超過100%？何況尚未放榜，根本不知是否會有增額錄取。原來他們是將各研究所的招生名額除以實際報名人數。由於有些研究所報名情況不佳，當報名人數少於招生名額時，除出來便大於100%了。該比值在大於1時，或許仍可有某些意義，但就是不能稱之為錄取率。卻求錄取率，只知盲目相除，正是典型的缺乏數據素養。只是從事統計分析時，採這類作法的人，顯然不少。

4 關於學測的探討(一)

目前高中生想進大學，學測成績相當關鍵。學測是從民國83年開始舉行的，其中別出心裁的花招不少，常把人弄得昏頭轉向。

學測共有國文、英文、數學、社會及自然等5科。每科原始滿分都至少是100；依序是108、100、100、144，及128。除英文及數學外，其他3科的總分，都不是常見的考試滿分。108、144，及128，這3個數都有好些有趣的性質。如正五邊形每個內角是108度，又36天罡星及72地煞

星，梁山泊共108位好漢；144為12的平方，且為一費波那契數(Fibonacci number)；至於128則是2的7次方。驚嘆數學之美後，難免讓人以為，那3科的滿分，乃經過精心設計。

學測的題型，以電腦可讀為主。國文包含單選題及多選題(不提示有幾個選項)；數學除單選題及多選題(不提示有幾個選項)外，還有電腦可讀型的填充題；英文與社會皆只有單選題；自然單選題及多選題皆有(有提示有幾個選項，亦有不提示有幾個選項)。國文與英文，尚有需人工閱卷的非選擇題。各科答錯均不倒扣。自然考科的試題分兩部分；第一部分每題計分，共80分；第二部分共56分，但得分超過48分便以48分計(即最多可答錯或不答8分)。除了前述5科外，大考中心從102學年度起還增考英聽，且自104學年度起，英聽成為必考科目之一。雖然各科滿分有所不同，但考試時間，除國文是120分鐘外，其餘4科皆為100鐘。

各科考題及計分方式，存在不小差異，看起來各科的自主性不小。但此自主性其實是虛的，因大考中心最後會將各科成績皆轉換成0至15的“級分”。各科看似深思熟慮產生的原始滿分，為何要一致換算成15滿級分？大考中心的說法是，為了不要讓考生為了0.1分的差距而計較。這真是個很可笑的理由。考完收到成績後，為孩子挑選校系的家長，常會懊惱地說，如果多1級分就比較好選了。本來競賽就是為了一分高下，豈有要求不在乎分數？

有人可能會替大考中心緩頰，說有些大學(尤其在歐美)，教師成績評定以A、B、C、D，及F的方式，那正是級分，因此使用級分制不奇怪。這當然沒錯。但眾所皆知，教師給成績，本來就相當主觀。因此一門課，若任課教師不同，甚至同一教師在不同年度開設，則所給的成績，乃難以相比。至於學測，來自各地的十餘萬位考生，要能一起比。這跟教師給多少A多少B，有很大的裁量權，乃完全不同。且之所以採以電腦閱卷為主，不就是為了消除由於主觀因素，所造成給分的差異？大考中心轉換級分可能產生的問題，我們陸續會說明。無論如何，大考中心負責辦理考

試，該在意的是測驗的信度及效度，以提高考試之公平性，而非考生是否斤斤計較分數。

級分如何得出？各科取前1%考生成績的平均，除以15，即為各級分之級距。舉例來說明。數學滿分為100，如果到考人數為150,327人，則1%便是1,503.27，取1,504(採無條件進位至整數)。最高分的前1,504人原始成績之平均，假設是93.748...分(此時先不四捨五入)，除以15後，得6.249...。再經四捨五入，取至小數第二位，得6.25，則6.25分即為級距。於是原始成績0分為0級分，0.01至6.25為1級分，6.26至12.50為2級分，...，81.26至87.50為14級分，最後87.51至100為15級分。對104學年度的學測，表8給出各科原始分數與級分之對照，表9給出各科各級分的人數及所占百分比，資料皆取自大考中心的網站。

為何換算級分時，會先取前1%考生成績的平均？猜想大考中心的考量是，考科若題目過難，將導致獲高分者很少，如此該科原本的上界(即滿分)，效用便不太大。他們的解決辦法，是將各科頂部1%分數之平均值，當做該科成績的“屋頂”，然後往下等分15格。這樣一來，除了15級分的原始成績之範圍最大外，屬於各級分的原始成績之範圍“等長”。只是既然想到調整屋頂，何以不調整“地板”？即取最低1%考生成績的平均，當做該科成績的“地板”。然後屋頂與地板間，等分15格。如此除了15及1級分外，屬於各級分的原始成績之範圍均等長。這樣底部級分就較不會“浪費”了。舉個例子來看。表9-4顯示，104學年度的社會科，由0至3那4個級分，才共217(= 7 + 2 + 17 + 191)人，約占社會科到考人數144,009的0.0015%。當低分人數不多時，底部級分分這麼細，並無太大意義，相當於浪費級分了。

該對大考中心調整屋頂的的巧思致敬嗎？事實上，大考中心的作法，效果不見得有多大。以數學為例。104學年度的數學考題，被不少教師視為難的不得了。但由表9知，前1%考生的平均成績，仍高達97.95。如果將各科前1%考生的平均成績，分別除以各科滿分，則數學

的0.9795為最高(不妨猜測第二高是那一科)。事實上, 5科中數學滿分最多, 共502人。其次依序是自然115人、社會5人, 及英文1人。至於國文, 無人得滿分108, 最高分98, 且僅1人。對於數學這科, 就是有一些頂尖高手, 不論面對多難的題目, 他們都似如魚得水, 不會受挫。

由表9可看出, 各科級分採原始成績等分(除15級分外), 將造成各級分人數所占百分比, 有可能相差很大。如檢視國文科, 12級分的人數最多, 約占19.04%, 即將近5分之1, 且由10至13級分的人數, 便約占63.46%, 至於1級分的人數才約占0.02%。何以不依到考人數等分? 如此各科的每一級分, 其人數所占百分比便很接近。

我們不理解大考中心的思維, 不知目前的設計, 背後是否有什麼大道理。表9還顯示, 各科同一級分的意義, 可能大不相同。如果有人問, 考12級分到底好不好? 這問題可不能立即回答好或不好。就以104學年度為例。數學若考到12級分, 便約只輸8.34%的考生, 且名列前12.98%, 算是不錯。至於社會, 若得12級分, 便約輸30.11%的考生, 且名列前48.14%, 不過是中等而已。如果各科級分是依到考人數等分, 則各科的同一級分, 所反映考生的表現, 便很接近。但依照大考中心目前採用的方式, 不只各科的同一級分, 所顯示考生表現之優劣相差很多, 且同一科在同一級分, 於相異年度, 也無法對照。大考中心所引進的級分制, 實令人難以理解其間設計的邏輯。

學測除了級分外, 各科及總級分, 又都依百分位數(percentile), 訂出頂標、前標、均標、後標, 及底標等五項標準。在計算百分位數時, 各科均不含缺考生, 總級分則不含五科皆缺考者。

在此頂標乃成績位於第88百分位數之考生級分, 前標乃成績位於第75百分位數之考生級分, 均標乃成績位於第50百分位數之考生級分, 後標乃成績位於第25百分位數之考生級分, 底標乃成績位於第12百分位數之考生級分。方向一轉, 不依成績, 改採以百分比來分群了。之前我們指出, 以成績等分來換算級分不宜。如今以百分位數來分割, 得到5標, 相當

於以人數來分群，該能獲得肯定了吧！並不盡然。

有一木製圍棋枱，靠牆斜立著，久未使用。一年後赫然發現，由於長期經日光照射，棋枱中間凹了下去，不堪使用。如果將不平坦的棋枱翻面，繼續讓日光照射，你覺得會因此變回平坦嗎？恐怕很難。先是15級分，再來5標，成績經二度扭曲，所代表的意義已很含混了。要知數據素養若不夠，有時看似面面俱到的作法，潛在的問題，將層出不窮。

5 關於學測的探討(二)

百分位數是什麼？在九年一貫數學課程綱要(底下簡稱九年課綱)裡，此題材被置於九年級(即國中三年級)。屬於國中的數學，總不該太難吧！即使沒學過的讀者，顧名思義，大約便能想出百分位數“大概的”意思。只是不曉得算不算知易行難，當學生實際拿到一組數字，想求百分位數時，卻常落入左支右絀的窘境，最後開始懷疑書上的定義是否恰當。由於屢使師生感到困惑，因此會有學者建議將此題材移至高中，也就不足為奇了。但倒也不必以鄰為壑，中小學數學課程，畢竟篇幅很有限，凡難以交待清楚的概念，釜底抽薪，寧可都不碰，應才是較佳的作法。

九年課綱在附錄一“分年細目銓釋”(底下簡稱細目)中，於9-d-04認識“百分位數”的概念項下，有5點說明，其中首2點為：

- (1) 百分位數和中位數、四分位數一樣，可以表示某資料組在總資料中的相對位置。學生應能自資料之相對累積次數分配表求出百分位數。
- (2) 知道百分位數通常用於分析總次數多的資料，避免在資料數少的例子中，做百分位數的教學。

細目中對於“中位數”，在9-d-02的說明中，有：

中位數是將資料排序後，前後各切一半的中間位置資料值。…中位數會使落在兩邊的資料呈現出某種“平衡”狀態。…中位數則是個數的平衡。

本來資料(或說數據)不一定只包含數字，但在談百分位數及中位數時，涉及的資料都須全是數字。又在課綱的附錄四“標準用詞與解釋”(底下簡稱用詞解釋)，在“百分位數”項下是：

各筆或各組資料的相對位置，表示有百分之多少的資料比該筆或該組資料的數要小。

於“中位數”項下是：

第50百分位數，通常表示比這筆或這組數大和比這筆或這組數小的資料各佔一半。

看到這裡，你可能以為，跟你所想的差不多，百分位數的意思本該如此，豈會有困惑？

首先，百分位數可否不落在所給的數據中？由細目(1)中說“某資料”，以及用詞解釋中說“該筆或該組資料”，可見不行。即任一百分位數，皆須為所給數據中的某一個。又可否有不為整數的百分位數？譬如，能說第37.8百分位數嗎？因細目中既然指出百分位數，可“表示某資料組在總資料中的相對位置”，難免令人好奇相對位置可否表示的細一點？九年課綱的定義中沒說不行，而也的確可以。

依用詞解釋，假設有一組數據1, 2, …, 100，則分別為第0百分位數，第1百分位數，…，第99百分位數。即數字 k 為第 $k - 1$ 百分位數， $k = 1, 2, \dots, 100$ ，共有100個百分位數。至於中位數，依用詞解釋，因是第50百分位數，即為51。只是比51小的數有50個，果真占一半；但比51大的數有49個，占49%，並不到一半，與細目及用詞解釋中，所述均不合。怎

會如此？事實上，對任何一組取離散值的數據，數據中“永遠不存在”（而非“通常表示…”）一個數，以使“比此數大和小的數各占一半”。

例如，對數據1, 2, 3, 4, 5，一般人會認為3是中位數。但依用詞解釋，1, 2, 3, 4, 5，分別為第0, 20, 40, 60, 及80百分位數。這筆數據，總共就這5個百分位數，其他的一個都不存在。數據的中位數不存在，可能違反不少人的直觀。當數據的個數為奇數且全相異，則由小至大排列後，我們總以為正中間那個為中位數。另外，對偶數個全相異的數據，如1, 2, 3, 4, 5, 6，或視中間那兩個，即3及4（3與4看起來“地位”相同），或視中間兩個的平均，即 $3.5(= (3 + 4)/2)$ ，為中位數，可能是不少人的認知。但由於第50百分位數為4，因此只有4為中位數，3與3.5皆不是。能有的百分位數，我們全給出來，1, 2, 3, 4, 5, 6，分別為第0, 16.66..., 33.33..., 50, 66.66..., 及83.33...百分位數。只是被判定為中位數的4，當然不會使落在它兩邊的數（分別有3個及2個），呈現個數的平衡。

再看一例。假設有1, 1, 1, 1, 1, 1, 1, 1, 1, 及2等10個數，即9個1與1個2。則第0百分位數為1，第90百分位數為2。就此2個百分位數，其餘百分位數皆不存在。因此當然也沒有中位數。

定義不周詳、違反直觀，且常有若干被認為該存在的百分位數不存在，是九年課綱所給百分位數之定義，被認為不太妥的幾個原因。不過依九年課綱的定義，雖有些重要的百分位數可能不存在，但一旦存在便唯一。這算是其定義的優點。附帶一提，由於在九年課綱中，百分位數必唯一，所以能有的百分位數之個數，最多等於數據的個數。

底下給一個九年課綱之外，常被採用之百分位數的定義。

對屬於區間 $[0, 100]$ 的任一實數 k ，第 k 個百分位數，以 p_k 表之，表數據中，至少有 $k\%$ 個小於或等於 p_k ，且至少有 $(100 - k)\%$ 大於或等於 p_k ，而中位數即 p_{50} 。依此定義，不但允許百分位數不落在所給數據中，也允許非整數的百分位數。

底下我們重新來檢視之前的幾個例子。

先看數據 $1, 2, \dots, 100$ 。區間 $[k, k + 1]$ 中的任一實數皆為 p_k , $k = 1, \dots, 99$ 。至於 $(-\infty, 1]$ 中的任一實數皆為 p_0 , $[100, \infty)$ 中的任一實數皆為 p_{100} 。共有101個百分位數, 各百分位數皆不唯一, 且整數 k 同時為 p_{k-1} 及 p_k , $k = 1, \dots, 100$ 。又 $[50, 51]$ 中的任一實數皆為中位數。值得注意的是, 依九年課綱所得到的第 k 百分位數, 為我們的第 $k + 1$ 百分位數。

其次看數據 $1, 2, 3, 4, 5$ 。 $(-\infty, 1]$ 中的任一實數皆為 p_0 , 且 p_1, p_2, \dots , 及 p_{19} 皆為1; $[1, 2]$ 中的任一實數皆為 p_{20} , 且 p_{21}, p_{22}, \dots , 及 p_{39} 皆為2; $[2, 3]$ 中的任一實數皆為 p_{40} , 且 p_{41}, p_{42}, \dots , 及 p_{59} 皆為3; $[3, 4]$ 中的任一實數皆為 p_{60} , 且 p_{61}, p_{62}, \dots , 及 p_{79} 皆為4; $[4, 5]$ 中的任一實數皆為 p_{80} , 且 p_{81}, p_{82}, \dots , 及 p_{99} 皆為5; $[5, \infty)$ 中的任一實數皆為 p_{100} 。所有第 $0, 1, \dots, 100$ 的百分位數皆存在, 至於中位數不但唯一, 且就是為3, 此符合一般人的直觀。在此例中, 有些百分位數唯一, 有些不唯一, 且一數可同時是幾個不同的百分位數。至於對 $1, 2, 3, 4, 5, 6$, 經類似的討論, 可得 $[3, 4]$ 中的任一實數皆為中位數。這結果尚令人滿意。至於其餘百分位數, 就留給讀者自行求出。

最後來看數據 $1, 1, 1, 1, 1, 1, 1, 1, 1, 2$ 一例。 $(-\infty, 1]$ 中的任一實數皆為 p_0 , 且 p_1, p_2, \dots , 及 p_{89} 皆為1。又 $[1, 2]$ 中的任一實數皆為 p_{90} ; p_{91}, p_{92}, \dots , 及 p_{99} 皆為2, 且 $[2, \infty)$ 中的任一實數皆為 p_{100} 。由0至100, 101個百分位數皆存在, 且中位數 p_{50} 為1, 仍符合一般人的直觀。

我們給的百分位數之定義, 應算是比較恰當的, 至少定義沒有不合邏輯處, 且會存在的百分位數較多。當然, 也可略修改我們給的定義, 即限制百分位數皆須屬於原來那筆數據中。只在原數據中決定百分位數, 這想法有其道理。現就以 $1, 2, 3, 4, 5$ 那筆數據為例。此時 p_0, p_1, \dots 及 p_{19} 皆為1; p_{20} 可為1或2; p_{21}, p_{22}, \dots , 及 p_{39} 皆為2; p_{40} 可為2或3; p_{41}, p_{42}, \dots , 及 p_{59} 皆為3; p_{60} 可為3或4; p_{61}, p_{62}, \dots , 及 p_{79} 皆為4; p_{80} 可為4或5; $p_{81}, p_{82}, \dots, p_{99}$, 及 p_{100} 皆為5。

讀者不難查到更多與上述兩種之一有些類似、不太一樣，或看不太懂在寫些什麼之百分位數的定義。例如，維基百科上寫的是：“百分位數，統計學術語，如果將一組數據從小到大排序，並計算相應的累計百分位，則某一百分位所對應數據的值，就稱為這一百分位的百分位數。”如果不懂百分位，就不知百分位數是什麼了。

至此，你可能會覺得百分位數的概念，並不像原先以為的簡單了。的確如此。對連續型的數據(如區間[150, 220])較無問題，只要依九年課綱細目裡，找出“前後各切一半的中間位置”185，便得到中位數了。只是中學數學裡，在討論百分位數時，通常面對的是離散型的數據，並無法“前後各切一半”。當初將百分位數此一題材，放進國中數學的眾學者，可能並未深思。遺憾的是，雖有點複雜，且意義不是那麼明確，但大考中心或許不以為意，才會依百分位數，將學測的總級分及各科的級分，皆訂出5標。因此其後會衍生出若干問題，也就不奇怪了。大考中心的百分位數之定義為何？我們稍後再說明。

有些讀者可能會以為，我們太吹毛求疵，因所舉的例子，數據的量都不夠大。九年課綱在前述細目(2)中，不早就強調“百分位數通常用於分析總次數多的資料，避免在資料數少的例子中，做百分位數的教學。”至於學測各科的到考生，都有10萬人以上，屬於“總次數多的資料”，因此百分位數，大可放心地使用。真的是這樣嗎？

6 關於學測的探討(三)

學測成績，能用於繁星推薦，及個人申請等進入大學的管道。各大學校系在第一階段篩選時，可對學測的總級分及各科級分，依5標訂出檢定標準，做為篩選門檻。就以國立台灣大學為例。在104學年度，生物產業機電工程學系，只對數學1科訂出均標；牙醫學系訂出國文前標、英文頂標、數學頂標、社會均標，及自然頂標，5科中涵蓋3標；數學系訂出國文

均標、英文均標、數學均標、社會後標，及自然均標，各科門檻均不高；化學工程學系，則國文、英文、數學及自然，4科皆須達到頂標，門檻不低；獸醫學系對國文、英文、數學、自然，及總級分，皆訂出均標，是少數門檻中包含總級分者。由此大致可看出，即使同一所大學，各學系於訂定檢定標準時，可能有各種考量，差異有時很大。而於訂門檻時，所能依據的，並非原始成績或級分，而是總級分及各科級分的5標。5標既然在篩選過程中，扮演重要的角色，因此大考中心提供的5標，意義應該要很清楚才行。只是並非如此。

如前所述，頂標、前標、均標、後標，及底標等5標，乃分別依到考生之第88、75、50、25、12等百分位數之級分而定。看到這裡，不少人腦海中可能浮現12、25、50、75，及88等百分比，分別為達到各標的人數之百分比。可惜大考中心並無此邏輯。這裡涉及的百分位數如何定義？大考中心在學測的簡章有說明。其計算很容易，舉例來看。假設數學到考生有150,327人。由 $150,327 \cdot 0.88 = 132,287.76$ ，採無條件進位至整數，即取132,288，則從最低分往上數之第132,288位到考生的級分，便為頂標； $150,327 \cdot 0.75 = 112,745.25$ ，則從最低分往上數之第112,746位到考生的級分，便為前標； $150,327 \cdot 0.50 = 75,163.5$ ，則從最低分往上數之第75,164位到考生的級分，便為均標； $150,327 \cdot 0.25 = 37,581.75$ ，則從最低分往上數之第37,582位到考生的級分，便為後標； $150,327 \cdot 0.12 = 18,039.24$ ，則從最低分往上數之第18,040位到考生的級分，便為底標。

底下以上例來對頂標略做解析。因“至多”有132,287位考生，其數學的原始分數，比第132,288位考生的低，不到88% ($132,287/150,327 = 0.8799\dots$)，所以大考中心並非採九年課綱中百分位數之定義。又為何我們說“至多”呢？數學滿分100，而在表8中，分數區間端點表示至小數第2位。由0.00至100，共有10,001個不同的分數，15萬多的到考生，每一分數平均約有15人。所以對第132,288位考生，很可能有好幾位考生，數學與他同分，此將造成分數比他低的，少於132,287位。事實上，依104學

年度試卷中各類题目的計分方式，數學之原始分數，並不會有小數，因而每一分數之同分者會更多。如此一來，這位被大考中心挑出來，當第88百分位數指標的考生，數學原始分數低於他的考生，將更不到88%了。

大考中心所採百分位數的定義，應相當接近在上一節中，我們所給的。此因分數少於或等於第132,288位考生的，至少占88%；而分數多於或等於第132,288位考生的，由於至少有18,040(= 150,327 - 132,287)位，而 $18,040/150,327 = 0.12\dots$ ，即至少占12(= 100 - 88)%。但要注意的是，如同在上一節裡，所舉的那幾個例子，這第132,288位考生，其數學的原始分數，雖被稱做是第88百分位數，卻可能同時為好多個不同的百分位數。

靈敏的讀者可能已警覺到了，假設數學的分數只有整數，則雖多達十餘萬位考生，但分數僅101筆，因此根本不算“總次數多的資料”。尤有進者，就算分數真能計算到小數第2位，1萬多個不同的分數，乍看不少，到頭來卻仍非“總次數多的資料”。因原始分數只是過客，被大考中心束之高閣，只剩下級分。15級分連同0級分，相當於總共才區區16筆相異數據。看看在上一節中，我們所給的那幾個數據個數不多的例子，不但一數可能同時為好幾個百分位數，一百分位數也可能會是很多個不同的值。此現象即使當數據不算太少時，都可能發生，何況僅16筆相異數據？因此看似精準的百分之88、75、50、25，及12，後來完全不是那麼一回事，乃屬必然。

我們之前曾質疑同一科在不同年，及同一年的各科，相同級分皆無法相互對照。如今各科級分及總級分的5標，都是同一百分位數，總可以對照了吧！非也。若這樣想，便被大考中心誤導了。

表10給出104學年度的學測，總級分及各科的5標之級分，資料取自大考中心網站。至於各標級分右側括號內之百分比，為達到該級分之考生，其人數占到該科到考生之百分比。這是我們利用表9，及大考中心網站上“五科總級分人數百分比累計表”，經簡單計算後，所求出的。這項資

料，其實對大學校系、考生與家長，都是頗具參考價值之數據。可了解究竟有多少百分比的考生，達到某大學校系之篩選門檻。只是不知大考中心何以不附上？難道是因擔心提供後，會讓人對5標的內涵產生懷疑？才乾脆模糊些，所謂民可使由之，不可使知之。

由表10可看出，對同一標，各科達到的考生之百分比，相差不小。以頂標為例。本來既然是第88百分位數，一般人難免以為，不論再怎麼進位，或有多少人同分，能達到的考生之百分比，都該接近12%。結果最低的是英文，百分比約12.18，還算符合頂標的意義。但最高的國文，達到頂標的考生，是預定12%的兩倍多，多達28.50%，這樣也算頂尖？大學校系以為訂了個高門檻來篩選考生，豈料此門檻一點都不高，真是嚴重被大考中心誤導了。檢視5標，各科都有某些標，達到的考生所占百分比，偏離設定的百分比很大，只有總級分較好些。這是合理的，因由0至75，總級分共有76筆相異數據，比各科級分的16多了不少。即使如此，總級分達到均標之考生，約占52.08%，與50%有相當明顯的差異。更引人注意的是，國文頂標及前標都是13級分，怎會這樣？該訂頂標或前標？大學校系說不定曾仔細分析利弊，討論再三。結果大考中心底牌掀開，告訴你白花心思了，二標沒有差別。

回頭看上一節所給的那組數據1, 1, 1, 1, 1, 1, 1, 1, 1, 2之例，且採我們對百分位數之定義，並在數據中找。則 p_0, p_1, \dots ，及 p_{89} 皆為1； $p_{90}, p_{91}, \dots, p_{99}$ ，及 p_{100} 皆為2。由於 $p_{12}, p_{25}, p_{50}, p_{75}$ ，及 p_{88} 皆為1，故從頂標至底標，5個標皆為1。5標居然為同一值！不必訝異，主要原因是，雖數據個數為10，但相異的卻只有2個。不要忘記學測各科的十餘萬到考生，也只有16個相異的級分。一旦有某一級分的人數，占了過高的百分比，便可能導致上下二標是同一級分。有些讀者可能會說等等，要依大考中心的算法求5標。結果其實是一樣的。因由 $10 \cdot 0.12 = 1.2$ 、 $10 \cdot 0.25 = 2.5$ 、 $10 \cdot 0.50 = 5$ 、 $10 \cdot 0.75 = 7.5$ ，及 $10 \cdot 0.88 = 8.8$ ，得知要找第2、3、5、8，及9個數，而這5個數都是1。

附帶一提，表10中各科的5標，利用表9中各科級分人數百分比之累計即可得。印證我們之前隱含指出的，換算成級分後，就相當於有16筆相異數據，而非有十餘萬筆數據。以國文為例。由表9，人數百分比之累計，至8級分時，首度超過12，故底標為8；至10級分時，首度超過25，故底標為10；至11級分時，首度超過50，故均標為11；至13級分時，首度超過75，故前標為13；也是至13級分時，首度超過88，故頂標仍為13。本來人數百分比之累計，至7級分時，約為11.69，離12並不太遠，約只差0.31。但因“固守”要大於或等於12的原則，因此再多1級，造成人數百分比之累計約為16.92，與12差了約4.92。捨0.31就4.92，使誤差增大不少。又人數百分比之累計，至9級分時，約為24.55，離25約只差0.45。但再度得再多1級，於是人數百分比之累計約達36.19，與25差了約11.19。再度捨小小的0.45就11.19。再來看人數百分比之累計，至11級分時，約52.47，而至10級分時，僅約36.19，距50太遠了。所以合理地取11，這是均標。至於人數百分比之累計，至12級分時，約71.50，不到75，得再多1級。但13級分的人數，所占百分比不少，約16.52。而 $71.50 + 16.52 = 88.02$ ，不但遠大於75，且略超過88，遂使頂標亦為13級。

事實上，因各科各級分的考生人數，常都占相當份量，因此各科達到某標人數之百分比，通常比 $(100 - \text{該標之百分位數})\%$ 大，甚至還可能大不少，這有點像一個數兩邊算的後果。以均標為例，這是依50百分位數得到的。但表10顯示，各科達到均標人數之百分比，分別為63.81、55.73、51.27、60.90，及50.30。其中國文的63.81，及社會的60.90，其百分比都較50大了10以上。5標實際的意義，與原本宣稱的百分位數既然可能差那麼多、各科5標的真正意義常差很大，且不能相互對照，大考中心何必苦守由12至88那5個百分比？依級分累計人數的百分比，斟酌訂定5標不就好了？甚至又何必先級分後5標？各科按原始分數的累計人數，約略等分成15級不就好了。如此每一級分人數的百分比，便將“接近”（因有同分及除不盡等因素，各級分人數的百分比不易剛好相

等)。不必5標,各校系篩選門檻,以級分來訂定即可。這樣既簡單,且更透澈,因達到某一級分人數的百分比,各科以及各年,是很接近的。

歸根結柢,原始分數不就可以了?大考中心何須迂迴曲折,棄原始分數,改採級分,再搭配5標?

7 關於學測的探討(四)

考完試,看分數普遍太低,教師有時會做一些調整,也就是將分數給個轉換。以開根號乘以10的方式來調整,在早期很常見。即將 x 分,換成 $10\sqrt{x}$ 分。此公式的優點是,0分仍是0分,100分仍是100分,所有成績都仍介於0到100間,且高低分順序維持不變。當然也有缺點,就是每人加的分數不同。考25分者變成50分,加的最多。而本來考1分與0分,差別極小,但換算後,前者成為10分,後者仍是0分。又雖每一科的高低順序維持不變,但若幾個老師都同樣採此調整成績的方式,則總分便有差了。例如,假設有A、B二生,考國英數3科。A生分別得81、64及36分,總分181;B生分別得100、100及4分,總分204。原本B生總分較A生高出23分,遙遙領先。但3科都經開根號乘以10的調整後,A生成績分別為90、80及60分,總分230;B生則分別為100、100及20分,總分220,反而落後A生10分。這種例子極多,讀者可自行嘗試舉一些。

如果想知葡萄跟芒果何者較營養,則該分析新鮮的葡萄及芒果之成分,而非比較被加工過的葡萄乾與芒果乾,此為眾所皆知的道理。大學校系在辦理甄選時,可依自己學系的屬性,或招生策略的考量,對學測的某些科目不採計、依成績採計,或加重計分,看起來很有自主性。只是各大學校系拿到的,並非考生各科原本的分數,而是被加工過的成績,即級分,以及意義不明卻用來當門檻的5標。為何不讓各大學校系自行對原始分數做判斷?非得依被轉換過的成績?為何不能自行訂定門檻?而只能依5標?大考中心未免太越俎代庖了!本來考題及配分都是大考中心百分

之一百決定的，大考中心自己卻覺得成績不能用來評比，必須轉換。轉一次成爲級分，再換一次弄出5標，背後究竟有何大道理來支持其作法？頗令人存疑。

雖我們已指出級分及5標設計的若干缺失，底下各再給一個。

大考中心由於依分數等距，而非依人數百分比，將原始分數轉換成級分，造成考生成績排名可能超乎想像的大翻轉。假設有A、B二生，皆參加104學年度的學測。國文、英文、數學、社會及自然等5科，A生原始分數各得82、88、100、126及116，總分512；B生原始分數各得83、89、92、127，及117，總分508，A生優於B生。由表8，A生5科級分各爲14、14、15、14，及14，總級分71；B生5科級分各爲15、15、15、15，及15，總級分75。A生原始總分比B生多4分，其中有一科多8分，另4科各少1分。高的那科多很多，低的那4科都才少些許，原本A生不但不會被視爲比B生差一截，甚至會覺得A生略好些。但換算成級分後，B生興高采烈，因他得到滿級分。至於原始總分較高的A生，反而比B生少了4級分，能選填的大學校系差很多了，有點氣餒。不是爲了讓學生不要斤斤計較分數，才採級分制，現在看來，怎能不計較呢？總分較高，總級分卻低了4級，還不是最離譜的。假設C生5科原始分數各得98(國文最高分)、69、71、144及128，總分510。5科中有3科考到最高分，其中有兩科還滿分，相當不容易。但由表8，C生5科級分各爲15、11、11、15，及15，總級分67。C生原始總分較B生高，但卻足足少了8級分，能選填的大學校系與B生完全不一樣了，黯然神傷。怎會這樣？要知有些學系，沒有75級分，是通不過申請入學第一階段的篩選。只是能通過的，果真就是最頂尖？67級分一定遠不如75級分？現在看起來都未必。

大考中心依百分位數來訂出5標的級分，其百分位數是從下往上數，但其實從上往下數應較恰當。本來引進百分位數以訂定5標，如前所述，已屬匪夷所思了。採從下往上數，更加重其不合理性。什麼意思？

假設自然的到考生有150,250位。因 $150,250 \cdot 0.88 = 132,220$ ，故依現行作法，從最低分往上數之第132,220位到考生的級分便為頂標。如果採從上往下數呢？即因 $150,250 \cdot 0.12 = 18,030$ ，故將從最高分往下數之第18,030位到考生的級分訂為頂標。依第一種算法，往上數之第132,220位到考生，為從最高分往下數之第18,031(= $150,250 - 132,220 + 1$)位。兩種作法，所得頂標那位指標學生，有一點小差別，前後差一位。不過這是因乘出來是整數，若有進位算出來便會一樣。舉例來看。如前假設數學的到考生有150,327位。 $150,327 \cdot 0.88 = 132,287.76$ ，進位得132,288。此為從最高分往下數之第18,040(= $150,327 - 132,288 + 1$)位，又因 $150,327 \cdot 0.12 = 18,039.24$ ，進位後再度得到18,040。即此時從下往上數，或從上往下數，會得到同一位頂標的指標考生。無論如何，你可能覺得反正差異不大，固定一種算法來訂5標即可。但底下將說明，採從上往下數該是較合理的。

對於學測，大考中心雖宣稱“考生五科都必須應考”，但實際執行時，各科都會有些缺考。現假設自然的150,250位到考生中，15、14及13級分，各有6,001、12,029位，及13,243位。得到15及14級分者，合計18,030位，恰占12%。另一方面，由0至13級分共132,220(= $150,250 - 18,030$)位，占了88%。則依目前算法，自然頂標為13級分。可看出此時頂標的那位指標考生，為13級分的“頂”，至於自然達到頂標，為拿到13、14及15級分者，其人數共有 $6,001 + 12,029 + 13,243 = 31,273$ ，約占自然到考生的20.81%(= $31,273/150,250$)，遠高於預定的12%。現假設有位自然一向成績很不好的考生，於第二天上午考完第4科英文後，已快耗竭了，心想最後一科自然，打算申請的大學校系都不採計，於是把心一橫，不管老師及爸媽的叮嚀，就不考了。如此一來，自然的到考生成為150,249位。而 $150,249 \cdot 0.88 = 132,219.12$ ，進位得132,220。因少了1位到考生，由0至13級分也較原本少掉1位，即共132,219位。於是第88百分位數的那位指標考生，往上跳一位，變成14級分的“底”。

各科的缺考者，通常是該科較差的學生。如今因多了一位屬於底層的缺考生，竟然影響到上層，使頂標提高1級分。原先達到頂標的考生有20.81%，如今只有12%。這一來，將造成很多學生因自然未達頂標，而通不過某些校系的門檻了。某位考生一念之間的決定，後續影響竟然如此大，寧非怪事？難道是蝴蝶效應(butterfly effect)？如果各標的計算，是從最高分往下分別數12%、25%、…，就不會有這種大起大落的現象了。

在統計裡，對收集到的資料，一直攜帶著常覺累贅，而資料太多後，也不易了解其內涵。如何能達到資料減縮(data reduction)的目的，並使有用的資訊毫無損失？這便引進了充分統計量(sufficient statistic)的概念。例如，欲估計某銅板出現正面的機率 p 。一個辦法就是反覆投擲 n 次，各次的結果分別以 X_1, X_2, \dots, X_n 表之，其中 $X_i = 1$ ，表第 i 次得正面， $X_i = 0$ ，表第 i 次得反面， $i = 1, 2, \dots, n$ 。於是有一串0, 1的數字。對估計 p ，直觀上只要知道總共之正面數 $S_n = X_1 + X_2 + \dots + X_n$ 便夠了，而不必知道究竟那些次得正面，及那些次得反面。 S_n 便為一個充分統計量。即有了 S_n 後，就不用再知道 X_1, X_2, \dots, X_n 各是了多少了，資料遂大幅減縮。但大考中心提供的級分，卻對擬了解學生程度，失去很多重要的資訊。他們為何如此做呢？

8 關於學測的探討(五)

爲了打破一試定江山，自民國83年起，台灣高中生進大學，開始有了多元入學的方式。二十餘年發展下來，各大學提撥的繁星推薦，及申請入學等招生管道之名額，逐漸增加。這些管道的通行證，均仰賴學測成績。對於欲進大學的高中生，學測遂扮演著極關鍵的角色。

當初學測的開創者，爲有別於過去的聯考，將學測成績，定位在充當各大學校系檢定學生的門檻。設計此考試者，可能認爲不過是用來當門

檻而已，學生及大學校系，均不必太在意。對通過篩選的考生，各大學校系，再依自己特色及需求，訂出下一步的選才標準，而這一關應是比較重要的。遂有意將學測成績弄得模糊些。將原始分數轉換成級分不夠，又訂出5標，做為篩選的初步門檻。如此一來，挑人的一方，及被挑的一方，更有如在打迷糊戰了。對有興趣就讀的大學校系，若未達門檻，考生當然不必嘗試去申請。但達到門檻去申請後呢？各大學校系，在第二階段的篩選，可能有審查、面試或筆試等，工作不少。考慮作業量，並無法來者不拒。因此各大學校系，皆會在第一階段，訂出篩選倍率，因而拉高能通過申請的第一道門檻。於是門檻便不是低標準，而是一嚴格的關卡，級分遂相當重要。只要級分不夠，一切免談，其他表現再突出都沒用。

以國立台灣大學數學系為例。該系訂的檢定標準不高，各科不是均標就是後標。但由於依數學及自然，皆以倍率3來篩選，結果通過篩選的標準，103學年度(招生名額23)，數學加自然的級分為30，即兩科皆須拿到15級分；104學年度(招生名額24)則為29，即可以有1科14級分。換句話說，一開始的低門檻，徒具形式。這當然不只是台大才如此，不妨看國立中山大學財務管理學系。104學年度招生名額23，檢定標準是國文、英文，及數學皆均標(依序為11、9，及7級分)，也是不高。至於篩選倍率為英文10倍、數學6倍，及國文3倍。結果篩選順序一英文、二數學，及三國文，通過篩選的最低級分，依序為13、11，及14，都比均標高多了，數學且比頂標還高。以為只是國英文數均標即可，便貿然來申請者，就上當了，將輕易浪費一個申請學系。

大考中心篩選的順序，是從倍率高的開始。至於倍率相同者，假設有兩科，則依該二科的級分和來篩選。大學校系在訂定倍率時，不知是否真了解，這種篩選方式，會造成什麼影響？以國立台灣大學地質科學系為例。該系104學年度招生名額22，檢定標準是國文、英文、數學，及自然皆均標，篩選倍率為國文10倍、英文8倍、數學6倍，及自然3倍。結果篩選順序一國文、二英文、三數學，及四自然，通過篩選的

最低級分，依序為12、13、13，及14。如果有位考生，前述4科的級分分別為11、15、15，及15，雖該學系看似較重視的3科都考到滿級分，但因第一科國文不到12級分，當下就被刷掉了。篩選倍率最低，應比較像是該學系最在乎的一科，才會訂一較緊的關卡。實際上，卻是篩選倍率最高的，被大考中心採為第一道關卡。過不了此關，便沒有下一步。各大學校系，真的都了解這點嗎？國立台灣大學物理學系的篩選倍率訂為數學3倍，及自然5倍，而化學系訂為英文3倍、數學4倍，及自然5倍。此二學系最在乎的應是自然，篩選倍率卻訂的最高，雖與很多大學校系的作法相反，但可看出其背後的想法。當然如何訂篩選倍率才較適合，也與該大學校系招收學生的競爭力有關，倒也不能一概而論。另外，在第二階段的“學測成績採計方式”那項，對第一階段篩選倍率較高的科目，有些大學校系訂出較高的加權，有些訂出較低的加權。加權高低與篩選倍率高低間的關連，是否符合邏輯？相信有不少大學校系，在填寫各科倍率及權重時，並非很了解大考中心的操作過程。因此自己斟酌再三，才謹慎填上的那些數字，究竟會讓自己的學系，收到更想要，還是沒那麼想要的學生，可能不十分清楚。

有關篩選倍率的問題還有不少(像是篩選倍率相同，便以級分和來篩選，就很值得討論)，但限於篇幅，我們就此打住。

由於參加個人申請的每位考生，以申請6個校系為上限，考量能通過篩選的機會，因此考生會參考去年的資料。於是有可能因前一年通過篩選的標準較高，導致下一年的考生卻步。結果因少人申請，使通過篩選的標準降低，讓不少未申請者扼腕。除級分很高及級分很低的考生外，大部分中間的考生，決定申請校系時，都是一場賭。

設計出毫無邏輯的級分制，及問題重重的5標，不但未能達到讓學生不太在意成績的初衷，反而降低學測的效度。再加上只能申請6個校系，學測造成對擬進大學的高中生而言，機運成分占很大。因不足以依賴，讓學生不再考試至上，願意花時間去做些其他的事，可能是學測的最大功

能。只是這一套莫名所以的制度，何以能通行無阻多年？

由於藉助一些數學來制定級分及5標，篩選倍率當然也與數學有關，讓普遍對數學畏懼的社會大眾，不敢置喙。設計者不是很聰明，就是認為其他人不夠聰明。

大家都聽過“朝三暮四”的故事，出自“列子”的“黃帝篇”，“莊子”的“齊物論”亦有引用。有位愛猴者，養了一大群猴子。他自己縮衣節食，卻儘量供養猴子。但後來實在撐不下去了，不得不減少猴子的口糧。他跟猴子說“從明天起，每天早上給你們3顆果子，晚上給你們4顆果子，可以嗎？”猴子一聽說早上的食糧減少了，個個呲牙咧嘴地站了起來，很是生氣。老翁趕緊改口說“這樣好了，每天早上還是給你們4顆果子，晚上又給你們3顆，可以嗎？”猴子一聽早上增加了，且晚上還有，便高興起來了。

如果連朝三暮四及朝四暮三，都可以把猴子搞糊塗，那最高1%考生成績的平均，及百分位數等，當然更令人莫測高深了，以為其中必有什麼了不起的專業。不是有人以“專業問題，專業解決”回覆提問者，就能讓人閉嘴。且這樣的回答，常也能獲得不少推崇。因此自認非專業者，豈敢對學測的各項“專業設計”，說三道四？

總之，本文藉大考中心舉辦的英聽及學測，以說明數據素養具備之不易。因而想用大數據治國，不是企圖唬人，就是緣木求魚罷了。

表1 104學年度高中英語聽力測驗第一次考試到考學生成績等級人數及百分比

等級	人數	百分比(%)	自F級往A級累積	
			累積人數	累積百分比(%)
A	21,624	17.03	127,014	100.00
B	55,149	43.42	105,390	82.97
C	46,159	36.34	50,241	39.55
F	4,082	3.21	4,082	3.21

表2 104學年度高中英語聽力測驗第二次考試到考學生性別
人數及百分比

性別	人數	百分比(%)	各等級人數百分比(%)			
			A	B	C	F
男生	61,361	48.31	15.88	41.93	38.19	4.00
女生	65,653	51.69	18.10	44.81	34.61	2.48

表3 104學年度高中英語聽力測驗第一次考試到考學生區域
人數及百分比

區域	人數	百分比(%)	各等級人數百分比(%)			
			A	B	C	F
北區	62,475	49.18	19.77	43.83	33.58	2.82
中區	28,256	22.25	15.65	43.94	37.08	3.33
南區	32,373	25.48	13.84	42.99	39.88	3.29
東區	2,548	2.01	9.97	38.23	44.86	6.94
離島	732	0.58	5.46	37.30	51.37	5.87
其他	630	0.50	11.75	29.21	43.17	15.87

註：1. 區域別以到考學生就讀學校所在地區分。

2. 其他區域包括進修學校、大陸臺商子弟學校、大陸及港澳地區學校及國外學校等。

表4 104學年度與103學年度高中英語聽力測驗第一次考試
各等級人數

等級	104學年度 第一次考試(1)	103學年度 第一次考試(2)	差異(1)-(2)
A	21,624	21,346	278
B	55,149	49,330	5,819
C	46,159	36,098	10,061
F	4,082	2,502	1,580
合計	127,014	109,276	17,738

表5 104學年度高中英語聽力測驗第二次考試到考學生成績
等級人數及百分比

等級	人數	百分比(%)	自F級往A級累積	
			累積人數	累積百分比(%)
A	8,630	13.02	66,260	100.00
B	28,043	42.32	57,630	86.98
C	27,661	41.75	29,587	44.66
F	1,926	2.91	1,926	2.91

表6 104學年度高中英語聽力測驗兩次考試共同到考生等級
改變的人數與百分比

等級改變情形	人數	百分比(%)
提升	16,451	28.88
維持不變	37,702	66.18
下降	2,817	4.94
合計	56,970	100.00

表7 104學年度高中英語聽力測驗兩次考試共同到考生等級改變情形

		第一次考試等級				合計
		A	B	C	F	
第二次考試等級	A	26	7,711	92	6	7,835
	B	14	18,619	7,261	44	25,938
	C	0	2,310	18,725	1,337	22,372
	F	0	4	489	332	825
合計		40	28,644	26,567	1,719	56,970

表8 104學年度學科能力測驗原始分數與級分對照表

表8-1 國文、英文及數學

科目	國文	英文	數學
級距	5.90	6.32	6.53
前1%	88.5	94.8	97.95
滿分	108	100	100
級分	分數區間		
15	82.61 – 108.00	88.49 – 100.00	91.43 – 100.00
14	76.71 – 82.60	82.17 – 88.48	84.90 – 91.42
13	70.81 – 76.70	75.85 – 82.16	78.37 – 84.89
12	64.91 – 70.80	69.53 – 75.84	71.84 – 78.36
11	59.01 – 64.90	63.21 – 69.52	65.31 – 71.83
10	53.11 – 59.00	56.89 – 63.20	58.78 – 65.30
9	47.21 – 53.10	50.57 – 56.88	52.25 – 58.77
8	41.31 – 47.20	44.25 – 50.56	45.72 – 52.24
7	35.41 – 41.30	37.93 – 44.24	39.19 – 45.71
6	29.51 – 35.40	31.61 – 37.92	32.66 – 39.18
5	23.61 – 29.50	25.29 – 31.60	26.13 – 32.65
4	17.71 – 23.60	18.97 – 25.28	19.60 – 26.12
3	11.81 – 17.70	12.65 – 18.96	13.07 – 19.59
2	5.91 – 11.80	6.33 – 12.64	6.54 – 13.06
1	0.01 – 5.90	0.01 – 6.32	0.01 – 6.53
0	0.00 – 0.00	0.00 – 0.00	0.00 – 0.00

註. 上表第三列, 乃各科原始分數前1%之平均, 為我們利用各科級距乘上15所估計。

表8-2 社會及自然

科目	社會	自然
級距	9.06	8.33
前1%	135.9	124.95
滿分	144	128
級分	分數區間	
15	126.85 – 144.00	116.63 – 128.00
14	117.79 – 126.84	108.30 – 116.62
13	108.73 – 117.78	99.97 – 108.29
12	99.67 – 108.72	91.64 – 99.96
11	90.61 – 99.66	83.31 – 91.63
10	81.55 – 90.60	74.98 – 83.30
9	72.49 – 81.54	66.65 – 74.97
8	63.43 – 72.48	58.32 – 66.64
7	54.37 – 63.42	49.99 – 58.31
6	45.31 – 54.36	41.66 – 49.98
5	36.25 – 45.30	33.33 – 41.65
4	27.19 – 36.24	25.00 – 33.32
3	18.13 – 27.18	16.67 – 24.99
2	9.07 – 18.12	8.34 – 16.66
1	0.01 – 9.06	0.01 – 8.33
0	0.00 – 0.00	0.00 – 0.00

註. 上表第三列, 乃各科原始分數前1%之平均, 為我們利用各科級距乘上15所估計。

表9 104學年度學科能力測驗各科級分人數百分比累計表

表9-1 國文

級分	人數	百分比	累計人數	累計百分比
15	4,402	3.05	144,250	100.00
14	12,882	8.93	139,848	96.95
13	23,824	16.52	126,966	88.02
12	27,459	19.04	103,142	71.50
11	23,486	16.28	75,683	52.47
10	16,778	11.63	52,197	36.19
9	11,015	7.64	35,419	24.55
8	7,548	5.23	24,404	16.92
7	5,475	3.80	16,856	11.69
6	4,185	2.90	11,381	7.89
5	3,165	2.19	7,196	4.99
4	2,223	1.54	4,031	2.79
3	1,356	0.94	1,808	1.25
2	417	0.29	452	0.31
1	30	0.02	35	0.02
0	5	0.00	5	0.00

表9-2 英文

級分	人數	百分比	累計人數	累計百分比
15	6,850	4.76	143,883	100.00
14	10,673	7.42	137,033	95.24
13	13,148	9.14	126,360	87.82
12	12,699	8.83	113,212	78.68
11	13,784	9.58	100,513	69.86
10	12,564	8.73	86,729	60.28
9	10,463	7.27	74,165	51.55
8	10,089	7.01	63,702	44.27
7	9,571	6.65	53,613	37.26
6	8,775	6.10	44,042	30.61
5	10,431	7.25	35,267	24.51
4	13,243	9.20	24,836	17.26
3	9,451	6.57	11,593	8.06
2	2,081	1.45	2,142	1.49
1	55	0.04	61	0.04
0	6	0.00	6	0.00

表9-3 數學

級分	人數	百分比	累計人數	累計百分比
15	2,984	2.07	144,061	100.00
14	4,140	2.87	141,077	97.93
13	4,893	3.40	136,937	95.05
12	6,682	4.64	132,044	91.66
11	7,618	5.29	125,362	87.02
10	10,544	7.32	117,744	81.73
9	10,409	7.23	107,200	74.41
8	13,671	9.49	96,791	67.19
7	12,918	8.97	83,120	57.70
6	15,738	10.92	70,202	48.73
5	13,520	9.38	54,464	37.81
4	15,059	10.45	40,944	28.42
3	11,572	8.03	25,885	17.97
2	10,588	7.35	14,313	9.94
1	3,643	2.53	3,725	2.59
0	82	0.06	82	0.06

表9-4 社會

級分	人數	百分比	累計人數	累計百分比
15	6,496	4.51	144,009	100.00
14	17,450	12.12	137,513	95.49
13	19,416	13.48	120,063	83.37
12	25,958	18.03	100,647	69.89
11	18,380	12.76	74,689	51.86
10	17,873	12.41	56,309	39.10
9	10,487	7.28	38,436	26.69
8	9,884	6.86	27,949	19.41
7	6,205	4.31	18,065	12.54
6	6,383	4.43	11,860	8.24
5	3,548	2.46	5,477	3.80
4	1,712	1.19	1,929	1.34
3	191	0.13	217	0.15
2	17	0.01	26	0.02
1	2	0.00	9	0.01
0	7	0.00	7	0.00

表9-5 自然

級分	人數	百分比	累計人數	累計百分比
15	6,004	4.18	143,718	100.00
14	7,504	5.22	137,714	95.82
13	9,370	6.52	130,210	90.60
12	10,223	7.11	120,840	84.08
11	11,994	8.35	110,617	76.97
10	12,952	9.01	98,623	68.62
9	14,249	9.91	85,671	59.61
8	14,988	10.43	71,422	49.70
7	14,916	10.38	56,434	39.27
6	13,853	9.64	41,518	28.89
5	13,854	9.64	27,665	19.25
4	10,267	7.14	13,811	9.61
3	3,286	2.29	3,544	2.47
2	238	0.17	258	0.18
1	10	0.01	20	0.01
0	10	0.01	10	0.01

表10 104學年度學科能力測驗總級分與各科成績標準一覽表

標準 項目	頂標	前標	均標	後標	底標
國文	13(28.50%)	13(28.50%)	11(63.81%)	10(75.45%)	8(88.31%)
英文	14(12.18%)	12(30.14%)	9(55.73%)	6(75.49%)	4(91.94%)
數學	12(12.98%)	10(25.59%)	7(51.27%)	4(82.03%)	3(90.06%)
社會	14(16.63%)	13(30.11%)	11(60.90%)	9(80.59%)	7(91.76%)
自然	13(15.92%)	11(31.38%)	9(50.30%)	6(80.75%)	5(90.39%)
總級分	63(13.27%)	57(25.86%)	47(52.08%)	36(76.75%)	28(88.24%)

註. 各級分括號中之百分比, 為達到該級分之累積考生所占百分比。