

# 統計概念的啓蒙和發展

蘇國樑  
國立空中大學 商學系

## 摘要

本文是透過研究兒童處理具體資料集的活動，企圖歸納出兒童之統計概念的基礎發展模式依序為：(1)輪廓運思期；(2)排序運思期；(3)造類運思期；(4)子類內造數運思期；(5)子類和量合併運思期及(6)子類和量序對運思期。藉此活動模式發現統計概念的獲得和呈現可以不必透過造數活動，而且是主體透過資料處理活動後所抽象出來的能夠表現出資料集內在的結構關係，或稱為資料結構。因此，教學過程中對資料集的量化、圖表化和機率分配模型化都應以學習和獲得資料結構的統計知識為導向。

關鍵詞：統計概念、統計結構、統計認識。

## 前言

H. E. Daniel(1975)曾在英國皇家統計學會的演說提到：「統計不是數學，也不是應用數學。」<sup>1</sup>D. Moore(1988)<sup>2</sup>和 D. J. Hand(1998)<sup>3</sup>亦曾經提出類似的說法。然而，D. J. Hand 卻提到：「數學是統計的基礎(lie at the root of)，即透過數學才能將統計概念或其他科學的概念形式化。」<sup>3</sup>所以，幾位學者的意思應該是“統計概念不同於數學概念”，因為一句“統計不是數學”是過於粗糙的和籠統的。這裡不禁要問“統計概念和數學概念的差異是什麼？”或是問“統計概念是什麼？”因而本文要探討的是：統計認識和統計概念；為何研究兒童的統計概念；兒童統計概念的發展；統計概念對統計教育的意義。

## 1. 統計認識和統計概念

L. J. Savage 曾說：「(1)統計包含兩件事：了解資料及獲得更容易理解的資料；(2)統計是關於不確定性.....機率分析似乎是一種工具，用來處理不確定性的工具。」<sup>4</sup>David Moore 曾寫道：「統計思維(statistical thinking)的要素是：1、過程中變異是隨處存在的；2、過程中資料的需求性；3、心中應存有會產生變異之資料的想法；4、變異的量化；5、變異的解釋。」<sup>4</sup>這兩位統計學者認為統計或統計思維的對象是資料中的變異性和不確定性，因為變異性和不確定性的存在是自然現象。因此，統計是處理變異性和不確定性的工具或思維模式，C. R. Rao 更是將他的書名的副題訂為“怎樣運用偶然性”<sup>5</sup>。然而，統計規律是蘊涵

著迴歸到或趨向平均數(regression to the average)的現象，是一種類似決定論的理想化現象，卻造成在使用統計結果時，容易忽略了變異性和不確定性存在的自然現象。但是，變異性和不確定性是意味著什麼呢？如何處理和思考變異性和不確定性才能彰顯出統計的功能？

Colin Mallows 因而提出：「統計思維是考慮關於現實問題之量化資料的關係，經常是表現在變異性和不確定性。它企圖精確化和解釋資料對興趣的問題所應做的說明。」<sup>4</sup> 也就是說，統計思維的對象是量化資料的關係，而變異性和不確定性只是表現量化資料關係的結果之一，因此，統計的功能在精簡化和呈現資料的關係。然而，資料的關係是建立在主體對資料集及其相關現象所做的認識活動和結果<sup>6</sup>。其次，統計思維是對資料集和其相關的現象所進行的認識思維活動，而統計認識活動的結果是能夠說明、描述和再表現資料的統計概念。因此，統計概念和統計認識資料的關係是有關聯性的，而且統計概念是統計認識過程的結果和更深入的統計認識的對象。

另外，Garfield 認為統計教育學家對統計思維模式(pattern)比較少研究<sup>7</sup>，而比較重視統計認識的最後結果或是最後結果的記錄。這不但表示對統計概念是比較少研究，也表示統計教育學者對統計思維的認識是比較靜態的，因為統計概念是動態統計思維過程的結果和對象。因此，有必要探討資料的統計認識或資料的統計概念。由於個體發展過程和種群演化過程是類似的，故先從探討兒童的統計概念著手。文章亦參考了國小數學新課程標準中，有關統計教材綱要主要包含：資料整理和統計圖表<sup>8</sup>。

## 2. 為何研究兒童的統計概念

為了尋求統計概念發生和發展的源頭，因此，以研究兒童統計概念的發展和過程做為基礎，且研究兒童的概念也比較能夠深化和具體化。另外，各年齡層的統計教材和教學的畫分是傳統的作法，也是為了讓學生獲得統計概念，而不是獲得所謂的兒童的、中學生的、大學生的或成人的統計概念。統計概念並沒有分成兒童的、中學生的、大學生的或成人的統計概念。所以，以研究兒童的統計概念發展模式是一個也意義的和可行的方向。

統計學的傳統看法是：「蒐集資料、整理資料、分析資料，再由分析資料的結果來做推論的科學方法。」<sup>9</sup> 因而譚寧君亦提出 4 個統計過程的重點「(1)蒐集資料、(2)整理資料、(3)呈現資料、及(4)解釋資料。」<sup>10</sup> 這裡看到是高水平的統計方法和過程，且強調的是資料處理活動的方法而非統計概念。然而，概念是經由對活動經驗的反思抽象(reflective abstraction)而轉化成的，因此，可以透過這 4 個資料處理活動的過程從而掌握統計概念。蒐集資料的重點在資料的觀察、度量、取得和記錄，雖然這些活動對資料的合理性具有相

當的影響，但本文的重點在統計概念的獲得和發展。所以，譚寧君所提之後 3 個活動可以做為分析和掌握統計概念的主要過程。而對統計概念的研究應從低水平的統計認識開始，因為高水平的統計分析方法和其結果，雖然可以使資料集的表現方式更一般化和更簡化一個資料集，但卻也和真實事物的資料集相差太遠。例如，平均數的點估計和區間估計的計算式雖然很簡單，但其所隱含的統計概念的合理解釋卻尚無定論。因此，本文的方法就是掌握兒童的後 3 個資料處理活動，並參照拙著《國小兒童統計概念分析之研究(I)》<sup>11</sup>，從而歸納出統計概念的發展模式。

朱建正提出：「在教學中，我們發現學生很容易就了解別人的策略，要模仿也非常容易。」<sup>12</sup>這裡隱含著：能夠模仿策略，並不表示了解引發該策略的動機，或由該策略所獲得的有關資料集的概念。也就是說，知其然，不一定知其所以然，或是俗稱的假學習。然而，策略活動的研究卻是獲得兒童是否具有概念的方法之一，因為能自發性地執行策略活動，則表示能將具體活動抽象化，並將其抽象概念轉化、進行具體化或表徵化的活動。因此，研究概念發展的方法之一是研究其表徵的策略活動。

另外，從研究觀察兒童對單一屬性資料集的統計活動，再對照大、中學生對於數量化資料集的統計活動可以發現：兒童身上發生的一些基本的統計活動現象，在大、中學生的身上也可以找到類似的現象。例如，大、中學生對數量化資料集(如身高或體重的資料)的分類也會有困難，原因之一是缺乏共同組距(class interval)單位。之二是，缺乏共識的組數(number of classes)，因為組距和組數是需要“自創”的，通常教學時只做原則性的規範。之三是，缺乏對量化資料的質的認識，例如，量化的身高資料不僅表現身高的量，並且是表現身高的差異。也就是說，以研究兒童的統計概念做為研究大、中學生統計概念的基礎是合理的。

### 3. 兒童統計概念的發展

為了研究資料的統計概念，考慮藉由研究“一次呈現”具體資料的認識開始，省去規劃蒐集資料的程序和方法；其次，蒐集資料和登錄資料的原因和目的之一在，希望將所觀察的現象能夠於同一時空內一齊呈現出來，其結果和效果亦是在一次呈現其所記錄之資料。因此，將研究目標集中在資料集的處理活動和資料集的次數(frequency)關係上，而暫不考慮資料集內元素出現的時間順序。另外，為了研究統計概念本文只考慮對單一屬性<sup>11</sup>(single characteristic)之具體資料的處理，而非多屬性之具體資料或數量化資料的處理上，故有「其它後續活動則透過已溝通的統計圖表格式進行讀表與填表的活動，此時表格的讀取均仍只在單向度的考量上。」<sup>13</sup>這樣的考慮是因為低年級的兒童尚無法處理多屬性之具體資料或

數量化的資料，如同兒童尚無法了解乘法對加法的分配律的性質，例如，兒童可以掌握  $2 \times 3 + 2 \times 5 = 2 \times (3 + 5)$  是因為  $2 \times 3 + 2 \times 5 = 16$  且  $2 \times (3 + 5) = 16$ 。

考慮譚寧君所提之後 3 個資料處理活動應包含有：資料範圍的界定、分類與點計<sup>11</sup>，而這些活動和測量活動或計數活動有關。例如，測量活動先要確定被測量對象是要有界的，故資料處理的活動需要借用到測量活動來界定被測量的資料集的範圍；而計數活動的結果可以獲得被測量的量和測量單位量之間的數值關係，故計數活動用以確定資料集內部間之質和量的關係。也就是說，資料集內部之統計關係的獲得和統計結構的表現方式均需利用到測量活動或計數活動。而數概念和測量活動與計數活動有密切關係<sup>14</sup>，故資料集內在關係的基礎之一是建立在數概念和其相關的活動上。因此，統計概念和數概念甚至與數的運算概念是有高度的關聯性。

然而，這裡應釐清(distinguish)統計概念和數概念的差異。數概念是透過測量活動與計數活動而獲得關於活動對象的數值關係，因此，數概念不刻意保留對象的質的內容。相反地，統計概念卻強調和保留統計對象的質的內容和數值的關係。原因是數概念是造數活動的結果，而統計概念是認識資料集活動(或稱造群活動)的結果。因此，若統計概念中缺乏統計對象的質的內容，則無法對整個資料集做出正確的認識。所以，可以視數概念活動對統計概念的獲得具有輔助性的功能。

另外，造數活動中的分類活動和統計活動中的分類活動亦有所不同。甯自強在兒童的整數詞有關的概念活動中，提出數概念模式依序為：「(1)合成運思；(2)累進性合成運思；(3)部分-全體運思；(4)測量運思。」<sup>15</sup>而此 4 個運思是建立在，建構出一系列的聚集單位(a composite unit)的概念上：新的聚集單位是以舊的聚集單位為基礎；全體為一聚集單位是以部分為一聚集單位為基礎。因此，這樣所構成之“類”和“類的類(classes of a class)”的標準和關係是建立在：“類的類”是小(或是舊)的聚集單位包含於較之為大(或是新)的聚集單位“類”之內，而非建立在元素的特性或屬性上。然而，資料處理活動所建立的統計分類標準卻是建立在元素的某些特性或屬性上，這也是資料集的內在關係中之質的內容。因此，資料集可以分類成若干個不互屬的子類(subclasses)，而統計概念中的聚集活動亦是考慮以各子類為基礎，然後在各子類內再進行數概念的聚集活動。也可以說，在掌握認識對象後，數概念活動是由內往外的，而統計概念活動則是由外往內的。

綜合上述數概念與統計概念的差異和前節的 3 個統計過程重點，並參照拙著《國小兒童統計概念分析之研究(I)》和其中的實例，可以歸納出兒童之資料的統計概念的發展模式(pattern)依序為：(1)輪廓運思期；(2)排序運思期；(3)造類運思期；(4)子類內造數運思期；(5)子類和量合併運思期及(6)子類和量序對運思期。筆者須先行說明的是，若遇有模式無法

解釋的新案例時，則應考慮更新或加強模式。茲將前述模式依序說明如下：

(1)資料集所佔的空間輪廓為一個聚集單位，元素的外在特徵和數量都是內嵌(embedded)於資料集內的。例如，資料集的呈現、說明或重製是注重在資料集所呈現的輪廓。在這個時期中兒童雖然已有量的保留概念，卻仍然受到形的保留概念的影響，進而影響到對資料集的認知。這也符合許多統計教育的研究結果：強調視覺化(visualization)對資料處理的重要性。此時兒童尚未產生資料量化的概念，因此，可以暫時界定兒童尚處於輪廓運思期(stage of configuration operation)的統計前置概念階段。

(2)資料集內每一個元素和其空間位置為一個聚集單位。此時整個資料集是一個類(class)、一個被測量的對象，資料集內每個元素是獨立的。兒童並未注意到元素的外在特徵，只注意元素的對應位置。因此，兒童視資料集為一個以元素和其對應之空間位置為一聚集單位所組成的一個群。故資料集的呈現、說明或重製，是以空間位置為序(order)做出一個接著一個的方式進行重現。這裡已顯示出一個統計活動特徵：重造一個資料集不見得會用到數概念和數的運算概念。

(3)以資料集內構成的元素的外在特徵如顏色、形狀或大小等為一個聚集單位，即以資料集為一整群或全類做為基礎，進行分類活動。例如，一組積木有 X、Y、Z、W 等顏色，共 t 個；因此，資料集的說明、呈現或重製的方式是，採一次或若干次抓取包含有 X、Y、Z、W 等顏色的若干個積木。元素的特徵脫嵌出資料集而成為一個獨立的單位，各別元素尚內嵌於資料集內。

(4)以資料集內的每一個子類為一個新的群或是新的測量對象，在各子類內進行重造表徵物活動或是數概念活動。例如，一組積木的呈現、說明或重製可以是以符號表之，譬如，X 色是 x 個相同大小的“○”、Y 色是 y 個相同大小的“○”、Z 色是 z 個相同大小的“○”、W 色是 w 個相同大小的“○”等；或採依序抓取或說明有 X 色有 x 個、Y 色有 y 個、Z 色有 z 個、W 色有 w 個等。量脫嵌於子類但尚內嵌於資料集之內，子類和量亦同時內嵌於資料集的輪廓內。這裡又再次顯示出前一階段的統計活動特徵：重造一個資料集不見得會用到數概念和數的運算概念。

(5)子類和子類對應之量兩者間的關係為一個聚集單位，而此關係尚內嵌於具體資料集內。對資料的說明、重製或比較的方式是注意類和量的對應關係的呈現，且需要具體物和相關活動來確認此一對應關係。例如，由一組彩色的積木所做出的“自用”的次數分配表，兒童認為此分配表將不適用於具有同樣色彩的花片的重製。也就是說，X 色有 x 個、Y 色有 y 個、Z 色有 z 個、W 色有 w 個等是用來表示該組積木的，不可以用來表示花片的。簡言之，製作具有同樣色彩的積木和花片的統計表是不相同的，兒童尚處於有逆運思

(inversible operation)時期。

這裡要提出的是，可以進行分類與記錄活動和利用分類與記錄活動來了解一個資料集是兩回事。

(6)子類和子類的量(或數)合成一個聚集單位，形成一個已脫嵌出資料集的序對(ordered order)關係，故可將此一關係記錄成“自用”的格式。因此，整組資料的說明、呈現或重製是根據類和其量的序對關係為基礎的。例如，將一組有 X 色有  $x$  個、Y 色有  $y$  個、Z 色有  $z$  個、W 色有  $w$  個等的積木記錄成“自用”的表，說明或重製時以此一“自用”表所顯示出的序對關係為重製或說明的依據。亦可以此“自用”的表為基礎製作出一組具有相同序對關係的花片，而不必拘泥於具體資料集的外在的形式特徵。也就是說，兒童已經進入所謂可互逆運思(reversible operation)時期。

統計概念是資料處理活動後主體所抽象出來的能夠表現出資料集內在的關係，而此關係所顯示的功能是，不但能夠以更簡單的、精簡的和系統化的方式來了解、呈現或說明原來的資料集，而且是做為更進一步之統計思維的對象。這也使得統計對象獨立於具體對象之外，而做出更為客觀的、普遍的描述和研究，而不需要局限於某個特定的或是具體化的資料集。由於統計對象是以獨立於具體資料對象的統計概念為對象，因此，統計的概念和統計學的發展變得更為抽象，但也因之而更具有客觀化和普遍化的性質。

#### 4. 統計概念對統計教育的意義

資料集和其內在關係之表徵方式的差異是因人而異的。前面所提的資料集內在關係的“自用”記錄格式是一種“自創”的統計關係格式，這是一種可以用來自我溝通和與他人溝通的格式，但並非約定成俗的統計關係格式。藉由教學指導和教室的社會性互動的溝通，不斷地交互辯證可以讓學生學習到，符合統計文化中所規範的(normative)或約定成俗的資料集統計關係的格式或是統計結構的形式，也就是次數分配表(frequency distribution tables)。這也說明了格式化的次數分配表非自發性的產物。

統計表的概念可以說是資料處理的活動經驗的產物，雖然它們在掌握和認識資料比起直接面對原始資料集來得相對地更精簡和系統化。但是次數分配表所表現的是資料集內在關係的序對關係，這是一種抽象的質和量關係。因此，這種資料集內在序對的質化和量化關係，可以再轉變成為視覺化的關係，即製作成統計圖。這是對資料集的抽象描述和說明再進一步的具體化，或是一種由雜亂無序的具體物資料進化到系統化的質和量之序對關係的具體化。這樣的視覺化和具體化不僅有利於呈現資料集的統計概念，更有助於掌握資料的整體概念和認識。

因此，對於小學的統計圖表教學可以藉由具體物資料的處理的示範，比較原始資料和

資料處理的結果，建立兒童認識統計關係與統計概念，和認識資料處理的模式與功能。中學以上則可以介紹量化資料集的處理模式，建立量化資料的序對關係，並以此序對關係為對象介紹簡化和代表此序對關係的量數或量關係，如平均數、中位數或眾數。然而，需強化和說明這些更簡化的量關係和量化資料之序對關係之間的關係，而非只是強調操作和演算這些更簡化的量關係的公式。

大學的基礎統計課程則可注重統計關係的進一步函數化，並說明這些統計關係的演進，譬如，隨機變數是進一步地將資料集內在質的關係的量化。因此，原先用來呈現資料集內在結構之質和量的序對關係則可進一步地轉化成量和量的序對關係。這樣的量和量的序對關係不僅將統計概念進一步地推升至更簡化或一般化的量和量的序對關係，因為可以擺脫對特定或具體資料集對象的序對關係；並且可以此量和量的序對關係為對象建立數學化的統計模型，即機率分配模型。由此可以讓學生了解這樣的量和量的序對關係及其機率模型，是可以直接地或間接地和資料集產生了聯絡。所以，透過對這些簡化或一般化的量序對關係和機率模型所做的運算、分析、說明和推論，譬如估計和檢定(test)，是可以演繹出對資料集所做的描述、說明或分析。

以上的分析是建立在統計思維的發展模式(pattern)為基礎之上，即「從序對化關係的建立過渡到單值化的整體概括概念，再由單值化的概括概念過渡到整體結構模型的建立。」<sup>16</sup>這也符合了人類認識的發展過程和規律，即由具體認識上升到抽象認識，再由抽象認識上升到具體認識。

## 5. 結論

因為統計概念的獲得和呈現可以不必透過造數活動，正如 C. R. Rao 所說的「有關統計學的最早記錄，大約可追溯到遠古，甚至在算術出現以前，原始人在樹上刻痕作為計算家畜及其它財產的一種方法。」<sup>17</sup>因此，一個資料集的統計結構或關係的呈現可以不需要利用到數值關係。然而，一個資料集的統計結構或關係被視成一物時，則圖表的呈現方式就需要被再簡化，以求做進一步的分析和發展。但統計關係如果太過簡化，如統計量數：平均數或中位數，則不易掌握資料集的整體概念，而應該發展出一個能夠描述資料整體概念的函數關係模型或方程式。數學中的機率分配模型滿足了這個需求，即可以利用機率分配模型去和統計資料集的關係進行匹配，這其中是需要經過多道轉化的手續，本文不擬深入探討。然而，以機率分配模型來配合統計結構，並不表示統計關係可以機率化，而是機率分配模型的統計化，才能為統計所用。也就是說，機率分配模型是被借用來表示統計對象的整體質和量的序對關係，而不僅僅是將統計對象的質和量的序對關係轉化成機率分配模型的機會關係。這樣機率分配模型才可以被用來做為對資料集或統計對象的整體結構分

析，而不僅僅是對資料集或統計對象做出現機會的分析。

### 註釋

1. H. E. Daniel: 〈Statistics in universities—a personal view〉, 《J. R. Statist. Soc. A》, 1975年, Vol. 138, Part 1, 第1頁。
2. D. Moore: 〈Comment in Harold Hotelling: The teaching of Statistics〉, 《Statistical Science》, 1988年, Vol. 3, No. 1, 第63-108頁。
3. D. J. Hand: 〈Breaking misconceptions—statistics and its relationship to mathematics〉, 《The Statistician》, 1998年, Vol. 47, Part 2, 第245-250頁。
4. Colin Mallows: 〈The zeroth problem〉, 《The American Statisticians》, 1998年, Vol. 52, No. 1, 第1-9頁。
5. C. R. Rao: 《統計與真理——怎樣運用偶然性》(Statistics and Truth: Putting Chance to Work), 石堅、李竹渝譯, 九章出版社, 1998年版。
6. 蘇國樑: 〈做為整體認識的統計學〉, 見《數學教育學報》(中國), 第8卷, 第1期, 1999年2月。
7. J. Garfield: 〈How students learn statistics〉, 《International Statistical Review》, 1995年, Vol. 63, No. 1, 第25-34頁。
8. 教育部: 《國民小學課程標準》, 教育部印, 中華民國82年, 第93-94、110頁。
9. 蘇國樑: 《統計學》, 國立空中大學, 1994年。
10. 譚寧君: 〈國民小學數學新課程統計圖表教材分析〉, 《國民小學數學科新課程概說》, 台灣省國民教師研習會, 1994年, 第127頁。
11. 蘇國樑: 《國小兒童統計概念分析之研究(I)》(NSC83-011-S-180-002N), 1994年。
12. 朱建正: 〈從具體解題活動出發的國小統計教材設計的理念與實際〉, 《國民小學數學科新課程概說》, 台灣省國民學校教師研習會, 1997年, 第220頁。
13. 譚寧君: 〈國民小學數學新課程統計圖表教材分析〉, 《國民小學數學科新課程概說》, 台灣省國民教師研習會, 1994年, 第133頁。
14. 甯自強: 《國小數學教育論文集(I)》, 自印, 1997年, 第161-167頁。
15. 同上, 第181-184頁。
16. 蘇國樑: 〈統計對象的抽象原則〉, 手稿, 1999年。
17. C. R. Rao: 《統計與真理》(Statistics and Truth: Putting Chance to Work), 石堅、李竹渝譯, 九章出版社, 1998年, 第28頁。