科學教育學刊 Chinese Journal of Science Education
2000, 第八卷第二期,99- 2000, 8(2), 199-13

統計對象的抽象

蘇國樑
國立空中大學 商學系

(投稿日期：民國88年9月8日, 修訂日期：89年6月9日, 接受日期：89年6月21日)

摘要：本文主要在探討統計對象的幾個特有層面的抽象性質：一、統計對象的內容；二、統計對象的特殊內容；三、統計對象的特殊高度；四、統計對象的抽象方法。這四方面的抽象不僅關係著統計的內容、形式符號、意義和所指(reference), 亦關係著統計的教學活動和相關的發展。透過探討統計對象的抽象性, 不僅能更深刻地認識到統計分配的意義, 並且了解到統計分配是構成統計的理想元素或是理想的單位"數"。這樣的結果不僅可以反省和反映出對統計學的學科目標和統計形式的意義, 並提供統計和其相關的研究發展的目標和方向；而且也指出統計教學的具體目標和方向, 不僅是在傳承一些統計中的數學方法和技巧, 並且還要是傳授對分配和其相關的性質或規律的統計知識。

關鍵詞：統計對象、統計結構、整體認識、統計同構、理想元素。

目的、重要性和方法

一、目的和重要性

面對統計教學時就必然要對統計知識的看法有所釐清, 因為教師對統計知識的看法, 將會影響其對統計教學的看法, 從而影響其對統計的教學方式。例如, 金鈐和林福來(1998) 兩位教授就是了解到, 教師對數學教學的看法會影響到其數學教學方式, 因而進行〈準數學教師學習教學之前的教學關念及其緣起〉的研究；然而, 教師對數學教學的看法是受到其對數學知識的看法的影響。一般統計課程的教學內容充滿了數學化的統計公式和定理, 然而, 這些數學化的統計公式和定理是統計概念和統計知識的物化或符號化的結果。而統計教學的直觀對象就是這些數學化的公式和定理的傳授, 學生學習統計時的直觀統計對象也是這些數學化的公式和定理, 但是, 這些教和學的對象卻是符號化的、形式化的和抽象的。

另外, 一般將統計學界定為一門實用性很高的學科, 正因為此, 統計的抽象性反倒容易被忽略了。從而統計形式知識就成了實用性的代名詞, 故統計的形式知識就被等價於可以很容易地被具體化和概念化的統計知識。因此, 統計計算的方法和技術便成為統計教學的目標, 從而統計教學方式就是以這些抽象的統計公式和定理的推導為主。這也是一般所認為

的，學會了統計的計算方法和技術自然而然地就獲得統計知識和概念。

然而，提出統計對象的抽象性並不意味著強調統計是抽象的，而和其實用性產生對立。恰恰相反，筆者正是希望借著提出統計對象的抽象而了解到統計教學中所傳授的形式化統計知識是抽象的，因而學生不會自然而然地就接受和吸收這種形式化的統計知識。所以，教學中必須特別注意到統計形式知識、統計概念和客觀具體的統計對象之間的對應關係，或者是說應注意到統計知識的源頭和所指，從而在抽象的統計對象和相應的具體對象之間取得協調統一。這樣才能落實統計知識的實用性的說法，因為在教學中由教師單向地強調其實用性，如舉例並重複地計算出一些統計量數或方程式，並不能化解統計的教學對象的抽象性。故其重要性是不言而喻的。

二、研究方法

在統計教學的過程中除了需要指出數學化的計算和操作活動外，還需要注意到統計解釋和分析，而統計解釋和分析的基礎就在於對統計分配的深刻認識。本文是根據筆者長期觀察統計教學方式、統計教材編撰方式和有關統計教學的研究論述等所呈現的一些現象所做出的論述。這樣的研究方法類似於選樣晤談法（金鈐和林福來，1998），因為統計教師對統計和統計教學的看法常常會反映到其教材編撰或教學研究的論述上。然而，有關統計的教學方式、教材編撰或教學研究等，大多注重在統計中的數學計算或操作的研究，譬如，套公式求出計算結果；或者注重假設演繹的數學化統計定理的推導，譬如，關心統計定理中的假設條件和結論及可能的推廣等。這些作法雖然對統計技術的傳承是重要的，但是對統計概念或知識的獲得的效果是有限的。因此，有必要將關於統計的教學、教材和教學研究的論述做一個重新審視。

本文企圖指出統計教學中統計對象的抽象性，以便在統計教學時隨時提示或輔助以具體的統計對象或目標，這不僅可以彌補統計教學時學生的實踐活動的不足，亦可以釐清統計的學科目標和教學目標。下節將針對幾個教學或教材中的統計對象的抽象性，依序地進行探討：一是、統計對象的內容的抽象，二是、統計對象的特殊內容的抽象，三是、統計對象的特殊高度的抽象，四是 統計對象的抽象方法。這裡的特殊內容是指相對次數分配。一是、因為分配雖然可以以“量和相對量”的序對形式呈現，但是這種序對形式並不是單純地反映出如函數的對應關係，而是用來表示出資料集或群體的內部組成關係；二是、因為相對次數分配雖然無法直接顯示出實際的組成數量關係，但可以反映出相對的或概括的組成關係。如此我們就可以透過相對次數分配而對資料集或群體產生整體的認識，而這種以整體認識為基礎再進一步地認識和其相關的性質就是所謂的特殊高度。

對目前統計文化和內容的切入分析

一、統計對象的內容

統計學的主要功能是在於找出資料集的組成內容和組成關係，或稱為資料集的統計結構，因此，在以數量來表現組成內容和組成關係時，就不能完全捨棄質的考慮（蘇國樑，1999）。也就是說，必須同時考慮組成的質的內容和內容彼此之間的數量關係。具體的說，就是“有什麼東西”和“各是多少”。譬如，如果在“一群人”、“一堆產品”等客觀事物群中只抽象出該群人或物的量，這是對該群人或物的數學概念；然而，對一群事物的統計關係或概念，就不僅是這一群事物有“多少”個人或物，還應

該指出這一群事物內有“什麼樣”的人或物。又譬如，每年癌症死亡人數是數學對象的量，但是統計對象則包含了癌症總死亡人數內的年齡別、病因別、性別等等和癌症的量關係。這就是考慮了組成事物群的質的內容，然後在各個質的內容之下再進行量的考慮，即需要考慮“質和量”的特性。這樣才獲得一群事物的組成內容和組成關係，有時簡稱為統計結構，從而透過這樣的組成關係或結構進行認識、重建或重現此一群事物。人類就是透過重建或重現一群事物的方式，從而對更大範圍的(統計學上稱之為未知的或不確定的)情境做出整體理解和整體認識(統計學上稱之為推論或預測)。

這裡的組成內容和組成關係或結構就是一種包含“質和量”的序對關係的抽象概念。如果放棄質的關係而只保留量的概念，則所得到的就只是數學中量的概念而已。因此，即使是將質的關係轉換成量的形式，就是說將“質和量”的序對關係轉化成“量和量”的序對關係，從而成為一種看起來是新的、純粹量化的統計關係。這樣的轉換成“量和量”的序對關係是為了方便進行數學操作、運算和表示(presentation)，而不是捨棄和統計概念有關的質的關係而只保留純數量概念。

接下來是將統計操作對象的內容畫分為具體物資料、數據資料和序對資料時所進行的抽象化活動，為方便起見本節只舉出幾個簡單的例子做說明，因為本節主要是在說明雖然這些抽象活動(如分類活動)並不困難，但是其結果仍然是抽象的。

(一)內容是具體物

對具體物資料集抽象出一組“質和量”或“類和量”的序對關係，再根據此一組序對關係建立出統計圖表，因此，統計圖表是用來顯示出具體物資料集的內在“質和量”或“類和量”的序對關係。因而主體可以透過這些內在的序對關係，認識到或虛擬地重建出具體物資料集。然而，這些內在“質和量”或“類和量”的序對關係，或是統計圖表已經不再是原始的、具體的資料集，而是具體物資料集的抽象表示形式。這就是一種統計抽象。

[例 1]考慮資料集Ω = {□△△○□△□△○△}，可以利用 R = {(□，ˇ ˇ ˇ)，(△，ˇ ˇ ˇ ˇ ˇ)，(○，ˇ ˇ)}，或 R_0 ={(符號，個數)} = {(□，3)，(△，5)，(○，2)}的“關係”來表示Ω，因為主體可以利用關係 R 或 R_0 來重現或重造Ω的內容。但是，顯然地 R ≠ Ω且 R_0 ≠ Ω。

即使是將Ω重新排列而得出一個較為有序的(organized)新集合 Ω_0 = {□□□，△△△△△，○○}，雖然這裡的 Ω_0 可以更簡潔地、明顯地表現出Ω的內容，但是，仍然 Ω_0 ≠ Ω。因為 Ω_0 是將Ω內部元素進行重排和並置的結果。也就是說，不同的統計分類活動類型會得出不同的統計分類結果，因為影響資料出現的時空因素發生變動。統計的主要功能是，在同一時空之下“一次呈現”資料集的內在組成關係，或是說捨棄時空因素而只呈現資料集的組成關係或結構。

(二)內容是數據

對數據資料集的統計抽象是以數據資料集為統計對象，所建立的“量和量”的序對關係。這裡的“量和量”的序對仍然是包含了資料集的組成內容和其相應關係的數量化，所以，基本上仍然是一種“質和量”的序對關係。這樣的抽象方式是有其方便之處，譬如，可以進一步地利用純粹數量、數學方程式或數學模型來表現資料集的內容。因而它們看起來就更數學化、更抽象化，從而使主體忽略了其所隱含之統計的實質意義。

[例 2]考慮有一數據資料集Ω = {4, 3, 4, 4, 2, 3, 3, 4, 3, 2, 4, 2, 4, 1, 4, 3}，則絕對次數的序對關係是 R = {(數值元素，次數)} = {(1, 1)，(2, 3)，(3, 5)，(4, 7)}，或次數的序對關係符號化成 S = {(x, 2x − 1)；x = 1, 2, 3, 4}，或次數關係化成數學模型為 y = f (x) = 2x−1，x = 1、2、3、4。這裡可以看出 R ≠ Ω、S ≠ Ω和 f (x) ≠ Ω，甚至 R、S、f (x)和Ω的表示形式均互為不相同。然而，統計上卻說成 R、S 和 f (x)都是用來表示或呈現Ω；也就是從統計的角度來看，R、S 和 f (x)都是可以表示Ω的組成內容和組成關係。也就是說，R、S 和 f (x)就是對Ω的組成內容和組成關係的一種抽象的表現方式。

這裡很清楚地看出是以純粹的數量方式或數學模型來表現統計關係，因為資料集的內容是藉著量的形式而呈現，其本質卻是一種質的形式，譬如，關係 R 所表示的是"元素 1"有"1"個、"元素 2"有"3"個、...、"元素 4"有"7"個，這是一種 "存在性" 的表現方式。另外，不僅是表示Ω的方式是多樣的，就是單以Ω的內容和其組合的表現形式也是多樣性的，譬如，可以是{1, 2, 4, 2, 3, 3, 4, 3, 4, 4, 4, 4, 2, 3, 3, 4}或{1, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4}等等。這樣就可以很清楚地了解到主體對統計對象的認識方式也是多樣的，從而主體對統計對象的思考方式或操作方式都是多樣的。因此，統計教育的工作重點就可能是：一、統計教育工作者不應自我設限於所謂的標準的、制式的或唯一的統計操作和活動方式，二、如何讓學習主體建立或認識到規範的或約定俗成的統計操作和活動方式。

(三)內容是序對數據集

想要了解或描述一個群體內組成份子可以透過對每個成份的各種有關特徵的解釋和描述，譬如，想要說明中國人可以從中國人的膚色、身高、髮色、偏好或衣食住行等特徵進行描述。這裡暫時只考慮含有兩個特徵變數的數據集，也就是只考慮成雙的序對數據集或是雙向的矩陣表。面對成序對的數據集時，不僅可以考慮建立其分配關係，而且能透過這樣的序對數據資料找出兩個特徵變數之間的關係或是相關性。

[例3]考慮序對數據集{(1, −1), (−1, −5), (2, 1), (0, −3), (4, 6),(−2, −2), (4, 4), (-2, −5), (5, 4), (1, 0)}。首先，對序對中第一個觀察值設定為第一個特徵變數 X 的觀察值；其次，設定序對中第二個觀察值為第二個特徵變數 Y 的觀察值，這樣的設定方式並不是唯一的；再次，X 和 Y 的觀察值的範圍做分類和點計。建立一個雙向的分配表如下：

	$-2 \le x \le 0$	$1 \le x \le 3$	$3 \le x \le 5$	總數
$-6 \le y \le -3$	ˇ ˇ ˇ (3)	(0)	(0)	3
$-2 \le y \le 2$	ˇ (1)	ˇ ˇ ˇ (3)	(0)	4
$3 \le y \le 6$	(0)	(0)	ˇ ˇ ˇ (3)	3
總數	4	3	3	10

上表中可以概括地抽象出：當 X 增加時 Y 隨之增加。

另外，也可以將群體內關於組成份子的特徵變數的範圍界定在幾個"類別" (catalog)，譬如，有病和無病、贊成和反對等。因此，資料集可以直接顯示各個類別的元素個數，譬如，考慮調查吸煙和肺病的資料集如下：

	吸煙	不吸煙	總數
肺病	55	10	65
無肺病	5	30	35
總數	60	40	100

就這個資料表可以直接讀出 4 個交互情況的組成次數；也可以讀出"吸煙與否"和"得肺病與否"兩個特徵變數之間存在某種關聯性，就是對雙向表的數量內容做出某種質的抽象。這裡的兩個例子的前提假設是每個觀察值或每個人都是可抽象出一個數字"1"。

二、統計對象的特殊內容

(一)次數分配的意義

本文一開始就指出，統計的目標在透過對資料集的統計活動和其結果，從而認識或重建資料集或資料集所代表的原始群體。也就是說，上一節所提的"質和量"或"量和量"的序對關係基本上是提供事物或資料集的一種"存在性"或"存在狀態"。統計上將這種表示事物或資料集的存在性或存在狀態的"質和量"或"量和量"的序對關係稱作"次數分配"。因此，統計對象的特殊內容就是"次數分配"(frequency distribution)或是和次數分配有關的一些統計量數或形式。這些和次數分配有關的一些統計量數或形式基本上亦是為了提供對次數分配的理解。

次數分配是對一群客觀事物抽象出對應的"類和量"、"質和量"或"量和量"的序對概念，這種序對概念的數量化就是次數分配的原型(prototype)。因而統計對象便從原始的、具體的或客觀的資料集轉換成此一序對關係，或是說統計對象轉換成資料集的次數分配，而不在乎原來的對象究竟是物的組合群、人的組合群或其他事件的組合群。例如，可以從身高、體重或IQ商數等資料集，抽象出相應的統計分配的概念模型為單峰對稱的"直方圖"，從而歸納出單峰對稱的經驗分配。這裡關心的統計對象資料集內部組成元素的序對關係的具體形式。

(二)相對次數分配的意義

次數分配的表現形式是豐富的，譬如，可以是表格化、序對化或圖形化的。其次，次數分配是資料集內在的組成元素和組成關係的數量化，即統計結構，故次數分配具有一種結構的概念。另外，統計活動所追求的目標不僅是對象(即次數分配)是否是統計結構化，而且是這種結構是否具有"一般性"或"普遍性"。

[例4]擲3個公正錢幣的所有可能的、想像的現象集合$\Omega = \{TTT, TTH, ..., HHH\}$，若將T和H分別視為0和1，則原始的所有可能的現象將轉化成為$\Omega_0 = \{000, 001, ..., 111\}$。令$R_1$表示$\Omega$的內容數值關係，則$R_1 = \{(0, 1), (1, 3), (2, 3), (3, 1)\}$。然而，這樣的次數分配關係是一種理想的或想像的情境，在實踐中並無法得出更精簡化或普遍化的關係形式。譬如，對一次投擲3個錢幣的實驗實際進行100次，其所有的結果將可能會是$R_2 = \{(0, 12), (1, 40), (2, 37), (3, 11)\}$。這種是建立在絕對次數所構成的組成關係，雖然可以讓主體直接獲得"質和量"或"類和量"的序對關係。然而，R_1和R_2之間的關係是什麼呢？也就是說，如何在R_1和R_2之間做比較？充其量將R_1和R_2分別看成是，一次投擲3個錢幣的實驗分別是進行了8次和100次，其所有可能的結果的組成結構。R_1和R_2兩者都是一次投擲3個錢幣的實驗的一個特例，也就是上述的並不是一種更精簡化或普遍化的關係形式。正因為如此，其應用的範圍就只能限定於某種特定的集合或資料集。

既然統計的目標是對非特定的一個資料集或群體的整體認識，那麼，整體的資料集或群體的構成元素的總個數或總次數就必須加以考慮。因此，若令R_3表示Ω的內容相對的數值關係，則$R_3 = \{(0, 1/8), (1, 3/8), (2, 3/8), (3, 1/8)\}$。而一次投擲3個錢幣的實驗所實際進行的100次實驗，其可能結果的相對的數值關係則是$R_4 = \{(0, 12/100), (1, 40/100), (2, 37/100), (3, 11/100)\}$。這種利用相對的數值關係所呈現的組成"量和量"的序對關係不再只是簡單的"類和量"或"質和量"的形式，而是"類和相對量"或"質和相對量"的形式，為避免混淆可以將其書為"類和密度"、"質和密度"或"量和密度"的序對形式。

若將R_3和R_4化成另一種等價的表示形式：$R_3 = \{(0, 0.125), (1, 0.375), (2, 0.375), (3, 0.125)\}$和$R_4 = \{(0, 0.12), (1, 0.40), (2, 0.37), (3, 0.11)\}$，則$R_3$和$R_4$將可以進比較，譬如，可以看成$R_3 \approx R_4$。因為$R_3$和$R_4$的內容元素是相

同的，而其組成的關係結構是建立在相同的基礎上，即都是建立在介於 0 到 1 的分數或小數所表現出的相對次數之上。因此，機率密度函數的定義是「$f(x) \ge 0$，x 是樣本空間中任意元素；且$\sum f(x) = 1$ 或 $\int f(x)\,dx = 1$。」(Hogg & Craig, 1974, p23-25)如果只考慮“絕對次數關係”則 $R_1 \neq R_2$，因為R_1和R_2的組成內容元素雖然相同，但組成的關係結構是建立在不同的基礎上，即都是建立在自然數所表現出的絕對次數之上。

同理，[例 1]中的關係可以改成相對次數的形式{(□, 3/10)，(△, 5/10)，(○, 2/10)}；[例 2]中的關係可以改成相對次數的形式{(1, 1/16)，(2, 3/16)，(3, 5/16)，(4, 7/16)}；[例 3]中的雙向關係可以改成相對次數的形式

	$-2 \le x \le 0$	$1 \le x \le 3$	$3 \le x \le 5$	總數
$-6 \le y \le -3$	3/10	0/10	0/10	3/10
$-2 \le y \le 2$	1/10	3/10	0/10	4/10
$3 \le y \le 6$	0/10	0/10	3/10	3/10
總數	4/10	3/10	3/10	1

和

	吸煙	不吸煙	總數
肺病	55/100	10/100	65/100
無肺病	5/100	30/100	35/100
總數	60/100	40/100	1

這裡獲得一個觀念就是，不僅是絕對的次數可以表現出資料集的統計結構，而且“相對次數”也可以表現出資料集的統計結構。統計就是透過這個可以表現出統計結構的相對次數的比率或比值（稱之為頻率或密度），來表示資料集或群體的統計結構，這個比率化的統計結構一般稱之為機率分配。因此，機率分配是用來表示資料集或群體內的元素在“整體”中所佔之“相對”的規模、程度或範圍。因此，在教學上應該說明，統計上的比率形式的相對次數分配是用來表現出整體組成的統計結構，而不是測量機會的程度，雖然它常常被稱為是機率分配。

(三)和相對次數分配有關的量數或形式

在考慮整體總個數下，所得出之相對次數分配不僅保留內容組成的關係結構，而且是一種較為精簡化或普遍化的相對數量關係形式。正因為這樣的相對基礎上，其應用的範圍才不至於受限於某種特定數量的集合或資料集。進一步地可以將[例 2]中的相對次數的形式數學化成 $f(x) = (2x-1)/16$，$x = 1, 2, 3, 4$；[例 4]中 R_3 的關係數學化成為模型 $f(x) = \binom{3}{x}(1/2)^x(1/2)^{3-x}$，$x = 0, 1, 2, 3$。因此，前述的相對次數的比率關係或這裡的數學模型都是以客觀事實或理想情境為基礎所建立的整體組成結構，或稱為整體的統計結構。這種統計結構的普遍性不僅適用在一次擲 3 個硬幣的實驗上，亦適用在 3 人中對某一特定議題的贊成或反對的所有現象的描述，而且不會受到實際試行的次數的限制。這就是具有普遍性。

另外，在[例 4]中考慮比較 R_3 和 R_4 時，也可以以其各自的平均數或集中量數 $m_3 = 3/2 = 1.5$ 和 $m_4 = 1.47$ 進行比較。這裡表面上進行的是兩個數量的比較，但事實上是以兩個不同的相對次數分配或群體為基礎的比較活動。這也是統計分析和數學分析（如前述之兩個平均數的比較）的差異。

還有，在考慮測量兩個特徵變數（X, Y）所得的序對數據集$\{(x_1, y_1), \ldots, (x_n, y_n)\}$，並建立兩個變數的分配時，先對 x 和 y 的數值的範圍做分類和點計。這裡隱含著每一對數據資料的相對範圍、程度或重要性是相等的，因而給定每一對數據的比值為 1/n，或是說將每一對數據抽象為一個相對量“1/n”。而找出兩個特徵

變數之間的關係或是相關性，可以利用相關係數“$r = \sum(x - \bar{x})(y - \bar{y})/ns_x s_y$”所顯示的數量程度來表示 X 和 Y 之間的線性關係，即序對數據集的實際散佈情況是否可以用一個想像的直線表示之。很顯然地，1/n 是每一個各別觀察值和全體觀察值之間的關係的比率形式，r 是表示 X 和 Y 之間的線性關係的數量形式。所以，教學時就需要提醒，這些表列形式、數量形式或比率形式就是用來提供主體認識、想像、重建或虛擬出由這兩個特徵變數所構成的整體和兩個特徵變數彼此間的關係。

三、統計的實用性

由於統計研究的對象是“質和密度”或“量和密度”的序對所構成的相對次數的分配關係，其應用就不再限定於某種特定個數所構成的集合或資料集。因此，集合或資料集的範圍大小或總個數的多寡是可以依照實際需要而做增減的，譬如，需要考慮成本因素、時效因素或破壞性等。

另外，以整體現象的相對組成關係為對象時，就不需要做無限度或對整體中每個組成份子做觀察或實驗，而可以在分散的觀察或實驗的時間和地點，取得有限的觀察值或實驗值，從而進行統計分析，譬如，民意調查或市場調查等。因此，[例 1]中的關係{(□, 3/10), (△, 5/10), (○, 2/10)}，可以看成是整個市場中對 3 個不同產品的“購買率”的概括理解，或全體選民中對 3 個不同候選人的“支持率”的概括理解。這就是為什麼統計是可以不計較個別性影響之下對宇宙現象進行整體性認識。

再次，統計的實用性就是建立在對自然或人文現象的整體認識性。因為統計的整體認識是透過統計分配從而了解一群有限資料集的內在關係，進而以此有限資料集所建立的內在統計結構去解釋、認識或預測更大範圍的原始對象。這是建立在樣本和母群體之間具有相同的或相似的組成比率結構的基礎上，也就是說，將兩者之間理想化成具有統計的同構性(statistical isomorphism)。因此，近代許多自然科學、實驗科學甚至社會科學都需要藉助統計結構概念以找出其研究對象的內在關係，進而解釋、認識或預測其研究對象。這裡要說的是，統計同構是統計推論之所以能夠成立的前提或基礎概念。譬如，I. J. Good (1988)亦曾說過機率是一種證據的權值(weight of evidence)，雖然他是為建立歸納邏輯的推論的準確程度，而不是為了建立統計的分配概念。但證據和推論卻說明了，其中的理論是隱含著實際的樣本分配結構是相似於或甚至是相同於理論的或群體的分配結構，如此才可能進行歸納推理並說明其推論的準確程度。

四、統計對象的特殊高度

統計研究的對象是資料集或母群體的相對次數分配關係或組成的比率關係。然而，這種相對次數分配的概念是源自於對資料集的具體實踐活動的結果，譬如，相對次數分配表；或是對其活動結果的再抽象，譬如，直方圖或數學模型。因此，統計對象有時和具體的或原始的資料集相去甚遠，甚至是一種創造出的想像物或理想物。正因為如此，所以很多時候理想化的理論相對次數分配並無法和實際客觀的資料集的相對次數分配完全相同(identical)，甚至有些令人不滿意的差異。既然統計研究的對象從具體的資料集的相應的相對次數分配提升至抽象的、理論的相對次數分配，那麼關於理論的相對次數分配或機率分配的一些其它量數關係或性質便進一步地成為統計研究的對象。譬如，各種集中量數、離散量數或分配的數學組合或轉換，如機率分配 f(x)的特徵函數$\phi(t) = \int e^{itx} f(x)dx$。這裡要提的是，此後所指的分配就是指相對次數分配、機率分配或統計結構。

[例 5] 特徵函數或動差母函數的引入雖然

在數學上是一種分別是對分配所做的 Fourier 轉換或 Lapalace 轉換，但對統計而言這是對分配的進一步抽象，譬如，常態分配 $N(\mu, \sigma^2)$和 Poisson (λ)分配就可以分別地、進一步地被轉換成 $\exp(i\mu t - \sigma^2 t^2/2)$和 $\exp[\lambda(e^{it} - 1)]$等特徵函數模型，這些函數模型並不能看成真正的"分配模型"，但它們和其對應的分配模型不僅存在 1-1 的對應關係，而且更凸顯出分配和其參數間的關係。這裡的特徵函數和其性質就是一種統計對象。譬如，以機率分配為元素和卷積(convolution)為運算子所構成的結構(F，*)，雖然是一種具有數學美的半群數學結構(Feller, 1971, p293)，但這裡的(F，*)所構成的結構就是一個統計對象。

統計中除了以 x、y、z 表示統計對象內的組成元素外，還需以 X、Y、Z 表示一種測量活動和其結果的隨機變數，和以 F、G、H 表示其機率分配函數或理論的相對次數分配函數。在數學中"量+量"等於另一個量，即可以找到一個量與之對應，譬如，$x + y = z$。而統計中"隨機變數+(或 –)隨機變數"等於另一個隨機變數，譬如，$X + Y = Z$。Z 不僅表示一個新的隨機變數，而且表示一個具有新分配的新整體，譬如，兩個獨立的隨機變數 X 和 Y，分別具有自由度為 r 的卡方分配 $X \sim \chi^2(r)$、$Y \sim \chi^2(r)$，則 $Z = X + Y \sim \chi^2(2r)$。

上述是將隨機變數和其相應的分配當作"數"來運算和操作時，並不表示可以忽略它們所代表的一個新群體的整體變數和統計結構。因此，包含隨機變數的一些代數操作和其相關的極限理論便成為新的統計對象，因為隨機變數的"代數操作"與"極限理論"分別表現出"群體和群體的關係"與"群體相對次數分配的極限分配"。又，因為任意隨機變數的代數組合所得出之新隨機變數和新分配都是一種特殊變數和分配，故尋找一個普遍性的變數和分配或和其有關的方法便具有方法論的意義。如同[例 5]所述，分配和特徵函數之間存有 1-1 對應的關係，因此，可以利用特徵函數法獲得隨機變數的代數操作後所對應的分配和極限分配。

[例 6]當隨機樣本$(X_1, ..., X_n)$來自常態分配 $N(\mu, \sigma^2)$，則 $\overline{X}$ 和 S^2 具有下列三性質：一、$\overline{X}$ 也具有常態分配 $N(\mu, \sigma^2/n)$，二、nS^2/σ^2 具有卡方分配 $\chi^2(n-1)$，三、$\overline{X}$ 和 S^2 兩者是隨機獨立的(stochastically independent)。(Hogg & Craig, 1974, p172-175) 因為隨機變數 $nS^2 = \sum(X_i - \overline{X})^2$ 在給定 $\overline{X} = \overline{x}$ 時的特徵函數是 $E[\exp(itnS^2/\sigma^2) \mid \overline{x}] = (1 - 2it)^{-(n-1)/2}$，而特徵函數是$(1 - 2it)^{-(n-1)/2}$ 所對應的機率分配是 $\chi^2(n - 1)$。這裡很明顯地特徵函數$(1 - 2it)^{-(n-1)/2}$ 和 $\overline{x}$ 無關，因此，$\overline{X}$ 的分配和 nS^2/σ^2 的分配是機率獨立的，也就是說，"$\overline{X}$ 的分配"和"S^2 的分配"是機率獨立的。這裡的統計對象是特徵函數和其相應的機率分配，並透過機率分配所對應的特徵函數來分辨相關或獨立與否；而不是從 $\overline{X}$ 和 $S^2 = \sum(X_i - \overline{X})^2/n$ 兩者的一些形式符號的代數組合來分辨相關或獨立與否。如果只從形式符號的代數組合來看 $\overline{X}$ 和 $S^2 = \sum(X_i - \overline{X})^2/n$，兩者不可能是無關的，因為 $\overline{X}$ 和 S^2 都是由$(X_1, ..., X_n)$的代數組合而得的。因此，抽象的統計"分配"或機率"分配"就可以看成是統計中的"理想元素"或是理想的單位"數"，這也是在進行統計推論的教學前必須先探討理論機率分配和其模型的原因。

統計中分配雖然牽涉到序對關係中的量，但這些的量並非直接地抽象自資料集的個別元素，而是建立在子類為基礎的量化的序對關係。也就是說，先從資料集的元素之間的比較、分類活動開始，再對各子類內的事物抽象出量的概念，即進行造數活動或點計活動。所以，統計分配中的"質和量"或"量和量"的序對關係與真實世界的整體現象可能有些距離，不僅是想像中的概念產物而且是以抽象的形式表現出來。另外，既然統計分配是理想的單位"數"，

那麼對於統計分配的各種代數操作組合就必須考慮到“組合的合適性”，譬如，IQ 的分配和身高分配的組合就是不合適的，不僅是單位不同而且是對象的性質也是不同的。

直方圖、相對次數分配表或經驗分配(empirical distributions)的分配概念或統計結構概念，是建立在對具體事物或數據資料集的接抽象上。另一些概念則是建立在較為間接的抽象上，譬如，由一組平均數和標準差(μ，σ)所建立的虛擬(virtual)分配概念。這種由主體所虛擬出的統計分配和真實世界的“群體”不僅是有些距離的，而且個別主體所虛擬出的分配概念也是不相同的。如果再將直接活動的結果分配，如直方圖、相對次數分配表或經驗分配，進一步地理想化成為理論的參數分配模型，譬如，以常態分配的曲線近似單峰對稱的直方圖。則雖然理想分配和經驗分配之間存在差異，但是理論統計就是以這些理想化的分配為對象或基礎，進一步在其間找出或建立數學關係或數學結構的科學。因此，這些都是理想中的產物，也就是說，是人類想像下的創造物。例如，單峰對稱的經驗分配可以和常態分配相匹配(fit)，而每次試行只有兩種可能的結果的實驗可以和二項分配相匹配等等。所以，在教學或研究上可以專注於常態分配或二項分配等各種理想分配的各樣性質的說明和開發。

這裡需要提的是，理論統計雖然是以抽象的、理想的分配為研究對象，但這裡的理想分配歸根究底是可以聯繫到某一具體的資料集或群體的經驗分配。因此，所謂抽象的、理想的統計分配是需要能夠反映出客觀的事實，而非任意地、無限制地憑空臆測和抽象的。這也是適合度測驗(goodness of fit test)的基礎。因此，統計關係或結構的建立和加工雖然存在主觀的多樣性，但也存在著相對的客觀準則。這就是統計的客觀性。同理，迴歸模型或時間數列模型的選擇(model selection)也是以能夠反映出成對資料集的實際散佈情況為基礎。所以，E. L. Lehmann (1990)才會提出「模型的選擇基礎是，能夠使下一個觀察值和模型的最佳預測值所形成的期望誤差為最小。」

五、統計對象的抽象方法

統計學是一門具有特定學科目標的科學，自然會發展出幾個特有的抽象概念和方法，其中約略包含了 1、整體抽象，2、理想化抽象，3、估計的抽象，4、效率的抽象，將依序說明如下。了解這些統計中特有的抽象概念和方法，不僅是對建立統計理論具有重要意義，而且在教學中的重要意義是使學生不會模糊掉統計對象的焦點或統計目標。

(一)整體抽象

統計的整體抽象指的是由一些對象資料集（具體物或數據集）抽象出其中共同的內部組成關係。譬如，由一組數據資料集抽象出一個“相對次數分配”，或是由一組數據資料集抽象出一個集中量數和一個離散量數等等。因此，分配相同的不同整體視為具有相同的整體抽象概念。例如，若以公分為測量單位的身高分配可能和 IQ 分數的分配是相似的或相同的，這是因為實際的身高分配或 IQ 分配分別可以和某個特殊的常態分配相匹配。但這只表示兩者的內部統計結構相同，並不表示兩者是相同的整體。統計學就是以這種具有一般性、普遍性且整體性的分配為研究對象。這就隱含有所謂的“等價分配的集合”的觀念，即屬於同一等價集合中的兩個分配是具有相同機率分配者，譬如，常態分配族或 Poisson 分配族等。

另外一種統計等價概念是建立在“平均數和變異數”都相同的分配，也就是說，具有相同平均數和變異數的兩個群體是屬於相同的等價集合，即可視同具有相同的群體表現者，或是說具有相同的分配。這就是統計邏輯中特殊的同一律；也是變異數分析的基礎概念，譬如，

在變異數分析中行效果、列效果和交互效果均不顯著，則接受虛無假設，這就表示幾個群體所具有的平均數和變異數是無差異的，從而這些群體在某項特徵（譬如分配）上的表現是無差異的。至於迴歸模型和時間數列模型更是以整個成對資料集為基礎，所建立的兩個或多個特徵變數之間的數學化的關係模型。

(二)理想化抽象

這裡所說的理想化抽象是指由一群實際資料集或現象觀察值集合中所引出的抽象，進而將抽象的概念再加以簡化和完善化，因而所得出的統計概念和現實的整體原型就會有所差異。例如，任何真實事物集合的分配很難嚴格地說是具有常態分配；另外，在現實世界中更無法找到具有我們熟悉的統計分配狀態的整體現象。因此，相應的理論分配必然是一種簡化和完善化的結果。

簡化和完善化的理論統計分配對統計的發展是十分重要的，而且這種理想化的數學的統計模型對與之相應的事實對象也是十分重要的，因為我們才能不必針對某個特定的、具體的統計對象進行分析，而且可以透過數學理論或操作進行分析各種不同的實際問題。

[例 7] 關於成對的數據資料集所呈現的幾何散佈圖是不容易看出兩個變數間的關係。可以以相關係數“$r = \sum(x-\overline{x})(y-\overline{y})/ns_x s_y$”表示兩個變數之間的線性關係的程度，或甚至以最小平方直線模型$y = a + bx$具體化出兩者之間的線性關係。這樣不僅是找出線性關係(r)，也有可能找出的是因果關係(y)，進而可以對更大的範圍的整體做出解釋或預測。

(三)估計的抽象

統計的功能在透過樣本資料集的分配對統計對象提出說明、解釋或預測，其中包括統計對象的理論分配與和理論分配有關的群體量數(如μ和σ)、分配模型的參數(如 b(n, p)的 p 或 Poisson(λ)的λ)或關係方程式(如$Y = \alpha + \beta x + \varepsilon$的$\alpha$和$\beta$)的參數的選擇和估計。其抽象性依次敘述如下。

一、對群體量數的估計主要是以相應的樣本量數取代群體量數，譬如$\overline{X} \to \mu$和$S \to \sigma$；並透過這些樣本量數虛擬出群體的分配$F_{\overline{X},S}$，從而說明、解釋或預測群體的各種可能的性質。這種對群體分配的虛擬就是一種抽象。

二、以樣本資料集為基礎所建立的經驗分配來估計群體的分配，這是無母數統計的一種；或是假設群體的分配為一個具有參數的數學機率模型，則可以利用樣本資料集建立對參數的估計值，從而建立出群體分配的近似模型。假設的群體分配就是建立在想像中，這是抽象之一；經驗分配是主體透過統計分類建立的數量關係並不存在於樣本中，這是抽象之二；假設群體的分配是具有參數的機率模型，這種參數的存在假設就是抽象之三。

三、對於兩個特徵變數的資料集可以透過資料集建立出關於兩個特徵變數的關係數值 r，這是用來估計以群體而言兩個特徵變數的相關係數ρ。r 是無法說明群體的分配，但可以用來表示變數 X 和 Y 之間的線性關係之程度的測量值，從而用來說明群體成員中 X 和 Y 兩個特徵之間的關係和其測量數值 ρ。甚至將群體成員的散佈情形理想化成為一個具有線性關係的數學模型$y = \alpha + \beta x$，從而找出 α 和 β 的估計值 a 和 b，故$y = a + bx$是用來表示變數 X 和 Y 之間關係的估計模型。以一個簡單的數值 r 表示實際的關係程度是抽象之一；將實際散佈情形模型化成線性模型$y = a + bx$，而成為可以認識或思維的具體對象，這是主體創造出來的想像物並將其符號化的結果，這是抽象之二；所有的 r 值和所有的$y = a + bx$值亦可以組成各自的分配，這是抽象之三。

其實，還有許多和理論分配有關之技術性的統計性質無法在本文一一列出，因為這裡所開發的是在統計教學時對估計統計對象的抽象

性質的說明，而不是做統計理論或技術的探討或開發。也就是說，這裡針對的是統計知識在傳承上的意義。

(四)效率的抽象

在對上述的量數或模型的估計是需要考慮到選擇或估計的有效性(efficiency)，在統計上就是考慮對統計對象的估計量數或估計模型是否具有"不偏"和"最小變異"兩個性質。在"不偏"和"最小變異"的要求下，對理想化之統計對象理論參數機率模型中的參數估計，發展出充分統計量和最大概似估計量等有效的估計方法(principles)。就方法論而言這是將抽象方法的基礎和程序進行系統化和理論化，而不是單憑直觀的素樸猜測或碰運氣。不偏和最小變異就隱含著"平衡"、"對稱"和"穩定"的觀念，因為不偏就是指著估計量的分配的平均數等於所要估計的參數，而且此一估計量的變異數小於其他估計量的變異數。

為了符合點估計中的效率要求——不偏和最小變異，進一步地將參數化的機率分配歸納地建立了兩種族類："指數族"(exponential family)和"群族"(group family)。從而在這兩個分配族中分別建立可以符合不偏和最小變異之原則的一致最小變異的不偏估計量(uniformly minimum variance and unbiased estimator)和最小風險的等價估計量(minimum risk equivariant estimator)(Lehmann, 1983)等的方法、概念和理論。這裡的不偏、一致最小變異和最小風險等性質都是為了直接反映出對參數估計量的有效性的要求，即"平衡"、"對稱"和"穩定"。由於不偏和最小變異隱含著對估計量分配的依賴性，因而是對統計估計量特有的抽象效率。這裡又再次說明分配在統計中的重要性，也是統計教學所必須指出的中心議題。

另外，關於理論分配和其相關的參數的統計檢定和區間估計。這裡的統計對象是和建立統計檢定與區間估計有關的估計量和其分配。對於檢定和區間估計的要求是，分別建立出統計對象的一致最強力的不偏檢定(uniformly most powerful unbiased test)和一致最強力的不變檢定(invariant test)，從而分別建立出不偏信賴集(unbaised confidence sets)和一致最精確的等價信賴集(uniformly most accurate equivariant confidence sets) 等的區間估計(Lehmann, 1986)。這裡的不偏性、一致最強的和一致最精確的都是在反映對統計檢定和信賴集的有效性的要求。這些有效性是和統計對象的分配的預設條件或性質有關的，如單調概度比(monotone likelihood ratio)、對稱性(symmetry)和不變性(invariance)等統計分配特有的抽象性質。因此，更說明了統計效率的抽象性，而這些抽象性都是和分配有著密切的關係。

教學上的建議

David S. Moore（鄭惟厚譯，1998）在《統計，讓數字說話》中說到「統計的目的就是從數據中找訊息」，又提到如何從數據中找訊息有三個部分：「如何產生數據、如何整合數據和如何從數據中得出結論」。說到底，一般認為統計只是一種從數據中找訊息的技術和方法。筆者不禁要問：從數據中找"什麼"訊息？對這樣的"訊息"的統計意義是什麼？

首先，由前一節關於統計對象的抽象的討論可以看出，統計的目標是要找出資料集或群體的分配和其相關之性質。這在統計教學上是具有指導作用的，因為統計教學中通常充滿數學的操作和計算活動，教學過程中如果缺乏對統計目標，分配，的說明和探討，那麼統計教學就成為純粹數學的操作和計算活動的教學。分配不僅是統計活動的目標也是統計教學的對象，因此，分配是構成統計的理想的，基本的"單位"。譬如，研究特徵函數也是因為特徵函數和分配之間是存在著 1-1 對應的函數關係。至於其

他借用統計方法的科學研究或教學也都應該提醒，統計分析是以分配為基礎和目標，從而在蒐集資料和解釋、說明資料集上才能有探討和開發的方向。

第二，教學時須說明分配本身就是一個抽象的概念。雖然分配可以用來說明資料集或群體的組成結構，但是卻無法說明資料集或群體的實際散佈情況。這是因為分配是用來說明資料集或群集的"存在性"，但無法說明是以何種方式而存在的真實情形。另外，分配的功能是在同一時空中將組成結構呈現出來，放棄考慮對資料集或群集有影響的時空因素。因此，教學時就必須說明分配所提供的組成結構是以共時空的特殊性質為基礎，故需要注意到在超越這個有限的時空範圍的解釋將會是一種擴大解釋。譬如，選舉的民意調查就需要注意所適用的時空範圍。

第三，教學或研究時還需注意到統計推論的前提假設是統計的同構性，這是指樣本的相對次數分配等價於理論的或群體的相對次數分配。雖然統計同構性的假設並不表示統計同構性的事實是存在的，然而，這至少反映出統計理論和抽象是建立在可以具體實踐的客觀標準的基礎上。所以，統計認識是既抽象又具體，既微觀又宏觀。譬如，直方圖雖然是一個簡單又基本的統計圖，但是它卻可以提供研究時對樣本分配的一個整體認識，透過統計同構的假設而想像出群體或理論的分配，進而引用適合該理論分配的一些統計理論或技術。嚴格而言，卡方檢定也是在統計同構的前提假設之下所進行的檢定。簡言之，透過統計同構的假設，才能認識真正的群體或是未知的群體。

第四，至於統計量數（如集中量數或離散量數）的統計分析不僅是將其計算得出，並且是要以此統計量數為依據建構出相應的虛擬分配，從而整體認識一個資料集或群體，也就是說，提供一個整體性的思考方向。譬如，求出"$\overline{X}$和S^2"或"μ和s^2"的目的是利用其來虛擬出一個虛擬分配。雖然虛擬分配只提供部分的或抽象的現象而不是具體地反映出整個客觀事實，但是至少勝過毫無依據的胡亂猜測。又譬如，變異數分析不僅是對各種不同的平均數進行分析，但其主要的功能是對其後面所隱藏的各種分配進行統計分析。因為幾個群體要進行整體的數學比較是有其技術上的困難，請參見拙著〈統計對象的抽象原則〉(蘇國樑, 2000)。

第五，次數分配、統計量數及和兩者相關的統計對象的處理過程是適用於各種數據資料集，因此，對它們的思維方式也適用於各種數據資料集。既然是一種思維過程或方式，那麼便會產生極大的想像空間，以至於會有個體在重新建構整體時發生的多樣性或差異性。雖然這樣，但在統計教學中要強調的是，這些多樣性或差異性是需要以具體資料集或其相關的量數為客觀的基礎，而不是任意的想像。譬如，每個學生以一個平均數所虛擬出的虛擬分配可能都是不相同的，但這些不相同的虛擬分配卻必須有一個共同的客觀基礎，那就是平均數是相同的。又譬如，不同個體透過不同的分類標準對一組資料集進行分類將得出不同的資料分配，但是這些不同的資料分配的客觀基礎就是該組資料集。

第六，既然已經了解到分配就是資料集或群體的組成結構，則樣本的相對次數分配就是客觀的分配，而理論機率分配就是理想的分配。因此，樣本的相對次數分配就是分配的認識論的觀點，而機率分配是分配的本體論的觀點。這也可以對統計教學的二元論的對立觀點：統計教學就是純粹資料的探索分析的教學(Shaughness, Garfield, and Dormolen, 1996)和統計教學是以機率理論為基礎的處理偶然性知識的教學，進行調解而達到協調統一。因為兩者雖然對分配的看法和立場不同，但是對分配的認識則具有相輔相成的作用。

綜觀上述六點，還可以將統計課程的演進和數學課程的演進做一個比較。本節一開始就提到分配就是構成統計學科的基本單位。那麼，具體客觀的統計單位就是資料集的分配，由資料集分配的理想化所抽象出的機率分配就是理想的統計單位。譬如，人類是無法在自然界中找到常態分配的，只能由單峰對稱的資料集分配抽象出常態分配和其方程式；如同我們無法在自然界中看到“1”、“2”或“3”，我們只能從一個人、兩個蘋果或三棵樹中分別地抽象出“1”、“2”或“3”。既然數學的基本單位的“數”是可以被操作或計算的，那麼，統計的單位“分配”也應該是可以被操作或計算的，就是機率分配的代數運算。統計上只考慮機率分配的代數運算，因為具體的資料集分配的代數運算是不可行的，代之而起的是以資料集的描述統計量數來行之，其理由可以參考拙著〈統計對象的抽象原則〉(蘇國樑, 2000)。另外，數學由具體數的代數運算，除了提升為變數的代數運算，並進一步地演進出函數。因此，機率分配的函數(如特徵函數)也成為統計教學或研究的目標。將兩者課程的演進歷程以圖示之，由統計和數學的類比關係和互補關係，可以約略地看出統計教學的具體目標和方向，也可看出分配在統計科學中的地位。

結　　論

從分類活動開始的一系列統計活動是非自發性的、不自然的（蘇國樑, 1999）和抽象的，都需要藉助一連串的現象的觀察和對觀察的轉化才能獲得統計知識和計算技術。所以，會喪失某些客觀現實的性質。這就是所謂的抽象，但也是因為抽象才會有普遍性。也會產生類似於數學中只考慮對象的量的關係，而不考慮數學對象的質的關係時，而喪失了對象的某些客觀事實和性質。譬如，5 蘋果 + 10 蘋果 = 15 蘋果，是可以表現出蘋果的數量的存在性，但卻無法了解這些蘋果的顏色、品種、大小或甜度等等。這也讓我們了解到統計教育不能只是統計工作者對日常統計教學的簡單反省的結論，因為統計教學是具有一系列的抽象活動和對一系列的抽象對象的運思活動和認識活動。而要獲得統計教育和其本質意義，是可以透過歷史發展的角度、概念發展的角度、課程發展的角度或課程設計的角度等各種不同的角度，來探討統計教育的哲學意義。所以，統計教育是一門結合統計學、認知科學、資訊科學和哲學等跨學科的新學科，故統計教育是屬於知識性的統計學(knowledge statistics)，而不應簡化成屬於技術性的統計學(technique statistics)。

數學課程 :	具體數	→	數和其運算	→	變數和其運算	→	函數和其運算
統計課程 :	資料分配	→	統計量數和其運算	→	機率分配和其運算	→	特徵函數和其運算

這裡的機率分配和其運算還包含了分配模型或隨機模型的建立和其參數的估計、檢定等。

了解統計對象的抽象性對統計教學是有重要的意義的，因為統計雖然使用大量的數學符號、方程式和方法，但這些都是反映著具體的統計資料集或相應的具體分配，而不是一堆數學計算式的堆砌。因此，對這些數學式的抽象符號和統計意義之間做出適當的聯絡是具有統計教學的意義和目的的。也就是說，對統計而言任何分配和其有關的性質是可以反映出或對應到某個具體的資料集，至於對分配和其性質的進一步研究也是為了獲得更進一步地了解相應資料集的性質和結構。因此，分配是統計活動中一種相對獨立的對象，從而不需要只研究特定的具體資料集。如果將機率分配當做教學的直接對象，如集合函數、映射關係等，而忽略分配的具體意義，如(F, *)所構成的代數半群和相關的性質，這就是純粹的數學活動。也可以說是，統計學受到數學化的具體表現。但是，如果統計教育工作者對分配擁有深刻的認識，則這樣的純數學活動在某種程度上對統計教學或發展才會具有正面意義的。

參考資料

1. 金鈐和林福來（1998）：準數學教師學習教學之前的教學關念及其緣起。科學教育學刊，6(3), 219-254。
2. 蘇國樑（1999）：統計概念的啟蒙和發展。科學教育月刊，220, 9-16。
3. 蘇國樑（2000）：統計對象的抽象原則。見科學教育月刊。
4. 鄭惟厚譯（1998）：統計，讓數字說話 (*Statistics: Concepts and controversies* by David S. Moore)。台北市：天下遠見出版公司。
5. Feller, W. (1971). *An introduction to probability theory and its applications, Vol. 2* (2nd ed.). New York: John Wiley & Sons.
6. Good, I. J. (1988). The interface between statistics and philosophy of science. *Statistical Science*, 3(4), 386-412.
7. Hogg, R. V., & Craig, A. T. (1974). *Introduction to mathematical statistics* (4th ed.) (pp.172-175). London: MacMillan Company.
8. Lehmann, E. L. (1990). Model specification: The views of Fisher and Neyman, and later developments. *Statistical Science*, 5(2), 160-168.
9. Lehmann, E. L. (1983). *Theory of point estimation*. New York: John Wiley & Sons, Inc.
10. Lehmann, E. L. (1986). *Testing Statistical hypotheses* (2nd ed.). New York: John Wiley & Sons, Inc.
11. Shaughness, J. M., Garfield, J., & van Dormolen, J. (1996). Data Handling. In A. J. Bishop et al. (Ed.), *International handbook of mathematics education* (pp.205-238). Netherlands: Kluwer Academic Publishers.

The Conception on Statistical Objects

Kuo-Liang Su
Department of Business, National Open University
Lu Chow, Taipei County

Abstract

This article concerned with four aspects of statistical objects: 1. the content of statistical objects; 2. the particular content of statistical objects; 3. the particular height of statistical objects; 4. the conceptual fundations objects. The abstractions of this 4 aspects are connected to the content, formal symbols, meanings, and references. In addition, they are also related to the development of understanding during instruction. The study of conceptions of statistical objects can offer not only knowledge about statistcal objects, but also that a distribution is the ideal basic element or unit. This results not only reflect the objectives of statistics and the meanings of statistical forms, but can provide statistical researchers direction for future research. The results also have strong implications for statistics instruction. The purpose of statistics instruction is not only in transferring methods and skills to students, but also to transfer statistical knowledge about the nature of distributions to students.

Key words: statistical object, statistical structure, holistic knowing, statistical isomorphism, and ideal element.