

統計與生活

任教科別：數學科 作者：張雅淳

摘要

在經驗科學裡，觀察和實驗是求知最重要的初步工作。觀察和實驗的紀錄通常是一大堆原始資料或數據，如何從這大量的資料中尋求其潛藏的意義，統計學是不可或缺的分析利器。有人戲稱統計是「混水摸魚術」，在混濁的水中想捉到魚並非純然靠運氣，而是必須具備某種訣竅。依據這種訣竅，在混水中能對魚的所在「洞若觀火」，一清二楚，自然手到擒來。對一大堆數據來說，統計方法正是顯示意義的訣竅。現代統計學大致分成「記述」統計和「推測」統計兩大類，前者如上面所述為分析及解釋資料的意義；後者為依據樣本結果推測群體的狀況。

壹、前言

每個人一定都看過氣象報告吧!「明天或一週溫度預測」「降雨機率」「豪大雨特報」「低溫寒流特報」...，這些報導會影響你明天帶傘或穿衣多寡嗎?新聞報導「失業率」「經濟成長率」「離婚率」的高低會影響你工作的意願或是心情嗎?選舉前各政黨或候選人「支持度」「得票率」的預測你又相信多少?

企業或市場宣佈「投資報酬率」時，是否影響你的投資決策?其實統計數字無時無刻充斥在你的生活周遭。統計學其實是數學的其中一門課程，是用來搜集，分析，演繹以及呈現數據的。它被廣泛的應用在各門學科之上，從物理和社會科學到人文科學，甚至被用來工商業及政府的情報決策

貳、正文

一、統計學的意義

給定一組數據，統計學可以摘要並且描述這份數據。這個用法稱作為描述統計學。另外，觀察者以數據的形態建立出一個用以解釋其隨機性和不確定性的數學模型，以之來推論研究中的步驟及母體，這種用法被稱做推論統計學。這兩種用法都可以被稱作為應用統計學。另外也有一個叫做數理統計學的學科專門用來討論這門科目背後的理論基礎。

二、統計學的歷史

統計學的英文 statistics 最早是源於現代拉丁文 statisticum collegium (國會)以及義大利文 statista (國民或政治家)。德文 Statistik, 最早是由 Gottfried Achenwall(1749)所使用，代表對國家的資料進行分析的學問，也就是「研究國家的科學」。在十九世紀統計學在廣泛的數據以及資料中探究其意義，並且由 John Sinclair 引進到英語世界。因此，統計學的初衷是作為政府(通常是中央政府)以及管理階層的工具。它大量透過國家以及國際統計服務蒐集國家以及本土的資料。另依方面，普查則提供關母體的資訊。統計背後牽涉到更多數學導向的領域，如機率，或是從經驗科學(特別在天文學)中獲得的經驗證據設定估計參數。在今日的世界裡統計已經被使用在不僅僅是國家或政府的事務，更延伸到商業，自然以及社會科學，醫療等甚至更多方面。因為統計學擁有深厚的歷史以及廣泛的應用性，統計學通常不只被認為是數學所處理的對象，而是與數學本身的哲學定義與意義有密切的關聯。許多知名的大學擁有獨立的數理統計學系。統計學也在如心理學，教育以及公共衛生學系中被視為是一門主科。

三、統計學的觀念

為了將統計學應用到科學，工業以及社會問題上，我們由研究母體開始。這可能是一個國家的人民，石頭中的水晶，或者是某家特定工廠所生產的商品。一個母體甚至可能由許多次同樣的觀察程序所組成；由這種資料蒐集所組成的母體我們稱它叫時間序列。

為了實際的理由，我們選擇研究母體的子集代替研究母體的每一筆資料，這個子集稱做樣本。以某種經驗設計實驗所蒐集的樣本叫做資料。資料是統計分析的對象，並且被用做兩種相關的用途：描述和推論。

描述統計學處理有關敘述的問題：資料是否可以被有效的摘要，不論是以數學或是圖片表現，以用來代表母體的性質？基礎的數學描述包括了平均數和標準差。圖像的摘要則包含了許多種的表和圖。推論統計學被用來將資料中的數據模型化，計算它的機率並且做出對於母體的推論。這個推論可能以對/錯問題的答案所呈現(假設檢定)，對於數字特徵量的估計(估計)，對於未來觀察的預測，關聯性的預測(相關性)，或是將關係模型化(迴歸)。其他的模型化技術包括變異數分析(ANOVA)，時間序列，以及資料採礦。

相關的觀念特別值得被拿出來討論。對於資料集合的統計分析可能顯示兩個變數(母體中的兩種性質)傾向於一起變動，好像它們是相連的一樣。舉例來說，對於人收入和死亡年齡的研究期刊可能會發現窮人比起富人平均來說傾向擁有較短的生命。這兩個變數被稱做相關的。但是實際上，我們不能直接推論這兩個變數中有因果關係；參見相關性推論因果關係(邏輯謬誤)。

如果樣本足以代表母體的，那麼由樣本所做的推論和結論可以被引申到整個母體之上。最大的問題在於決定樣本是否足以代表整個母體。統計學提供了許多方法來估計和修正樣本和蒐集資料過程中的隨機性(誤差)，如同上面所提到的透過經驗所設計的實驗。

要了解隨機性或是機率必須具備基本的數學觀念。數理統計(通常又叫做統計理論)是應用數學的分支，它使用機率論來分析並且驗證統計的理論基礎。任何統計方法是有效的只有當這個系統或是所討論的母體滿足方法論的基本假設。誤用統計學可能會導致描述面或是推論面嚴重的錯誤，這個錯誤可能會影響社會政策，醫療實踐以及橋樑或是核能發電計畫結構的可靠性。即使統計學被正確的應用，結果對於不是專家的人來說可能會難以陳述。舉例來說，統計資料中顯著的改變可能是由樣本的隨機變量所導致，但是這個顯著性可能與大眾的直覺相悖。人們需要一些統計的技巧(或懷疑)以面對每天日常生活中透過引用統計數據所獲得的資訊。

四、統計方法

1. 測量的尺度

統計學一共有四種測量的尺度或是四種測量的方式。這四種測量(名目，順序，等距，等比)在統計過程中具有不等的實用性。等比尺度(Ratio measurements)擁有零值及資料間的距離是相等被定義的，等距尺度(Interval

measurements) 資料間的距離是相等被定義的但是它的零值並非絕對的無而是自行定義的(如智力或溫度的測量)。(Ordinal measurements) 順序尺度的意義並非表現在其值而是在其順序之上。名目尺度(Nominal measurements)的測量值則不具量的意義。

2. 統計技術

以下列出一些有名的統計檢定方法以及可供驗證實驗數據的程序

- 變異數分析(ANOVA)
- 費雪最小顯著差異法(Fisher's Least Significant Difference test)
- 學生 t 檢驗(Student's t-test)
- 曼-惠特尼 U 檢定(Mann-Whitney U)
- 回歸分析(regression analysis)
- 相關性(correlation)
- 皮爾森積矩相關係數(Pearson product-moment correlation coefficient)
- 史匹曼等級相關係數(Spearman's rank correlation coefficient)
- 卡方分配(chi-square)

五、延伸學科

有些科學廣泛的應用統計的方法使得他們擁有各自的統計術語,這些學科包括:

- 農業科學
- 生物統計
- 商用統計
- 資料採礦(應用統計學以及圖形從資料中獲取知識)
- 經濟統計學
- 電機統計
- 統計物理學
- 人口統計
- 心理統計學
- 教育統計學
- 社會統計(包括所有的社會科學)
- 文獻統計分析
- 化學與程序分析(所有有關化學的資料分析與化工科學)
- 運動統計學,特別是棒球以及曲棍球

叁、結語

近年來,由於資訊傳播迅速,資料的蒐集亦大量增多,一大堆的數據更是叫人難以消化。那麼,我們該如何整理、綜合和呈現數據,並從資料中獲得資訊呢?此時彙總資料最有效的方式:就是利用統計圖和統計表。統計圖表的優點,就是以簡易的方式顯現出資料的特性,讓我們可以很容易地從圖形看出資料的「長相」。

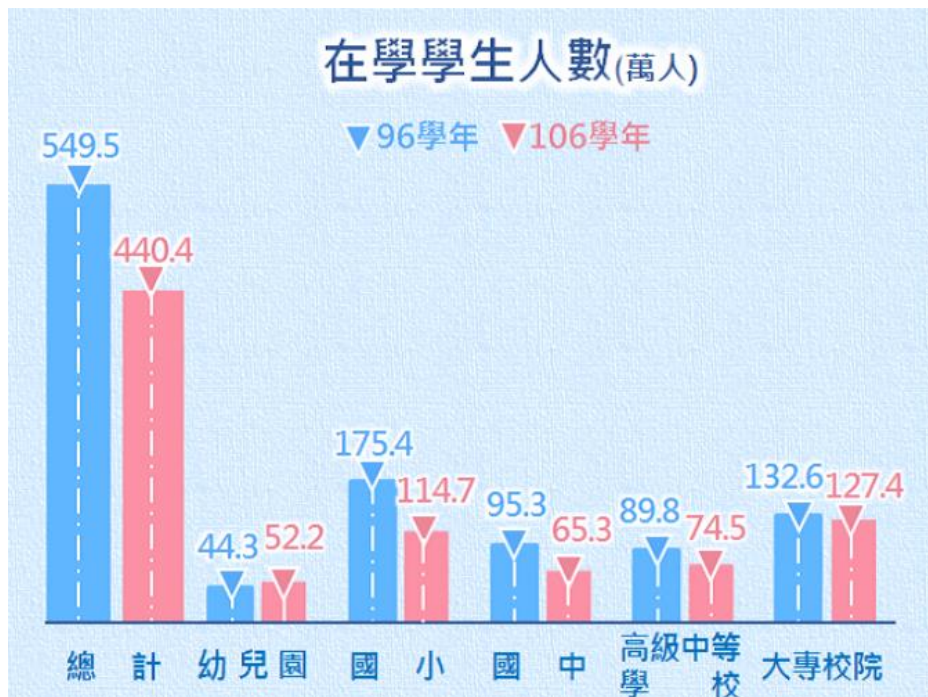
這是文字和數字遠不及圖表所能表達的喔！

生活中的統計圖表

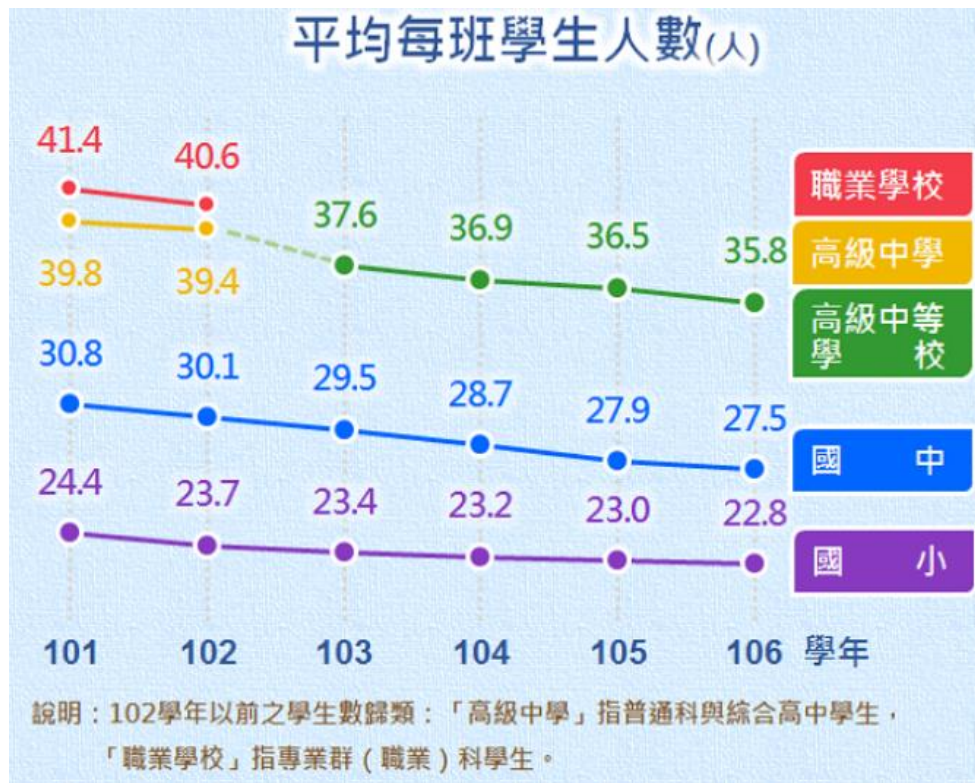
各教育階段學生實際數與推估數									
學年度	國民小學		國民中學			高級中等教育		大專校院	
	1年級		7年級	畢業生		1年級	畢業生	1年級	(中推估)
91	191.8	32.1	95.7	32.6	31.4	94.2	32.1	31.1	26.9
100	145.7	20.9	87.3	27.2	30.9	95.4	33.4	29.6	27.6
101	137.3	20.2	84.5	28.6	28.5	94.8	32.8	29.5	27.8
102	129.7	19.9	83.2	27.5	26.8	91.7	30.2	29.4	27.1
103	125.3	19.8	80.3	24.3	27.7	87.3	28.4	29.0	27.2
104	121.4	19.2	74.8	23.0	27.2	84.6	29.4	27.1	27.0
105	117.5	17.5	68.7	21.5	24.0	83.0	28.7	25.5	25.5
106	114.9	18.3	65.3	20.9	22.7	79.7	25.3	26.3	24.1
107	116.2	21.5	62.5	20.2	21.2	74.5	24.0	25.6	24.8
108	117.6	21.2	60.9	19.9	20.6	68.5	22.4	22.6	24.2
109	117.6	19.9	59.9	19.8	19.9	65.1	21.8	21.4	21.4
110	119.3	20.8	58.9	19.2	19.6	62.3	21.0	20.0	20.4
111	122.8	21.0	56.5	17.6	19.6	60.8	20.7	19.5	19.1
112	125.1	20.7	55.1	18.4	18.9	59.7	20.6	18.8	18.6
113	124.0	20.3	57.5	21.6	17.3	58.7	20.0	18.5	18.0
114	122.7	20.0	61.2	21.3	18.1	56.3	18.3	18.5	17.8
115	122.6	19.7	62.7	19.9	21.3	54.9	19.1	17.8	17.7
116	121.3	19.5	62.0	20.8	21.0	57.5	22.5	16.4	17.2
117	119.6	19.4	61.8	21.0	19.7	61.1	22.2	17.1	15.8
118	118.2	19.2	62.6	20.7	20.6	62.5	20.7	20.1	16.5
119	117.0	19.1	62.1	20.4	20.8	61.9	21.7	19.8	19.2
120	116.0	19.0	61.1	20.0	20.4	61.7	21.9	18.6	18.9

說明：虛線以下為推估數。

長條圖



折線圖



圓餅圖



肆、參考文獻

資料來源：統計網路學習館

資料來源：統計案例分析－統計思考是什麼

（國立成功大學統計系）吳鐵肩、李婉甄

資料來源：生活中的統計圖表(魔數學堂)