



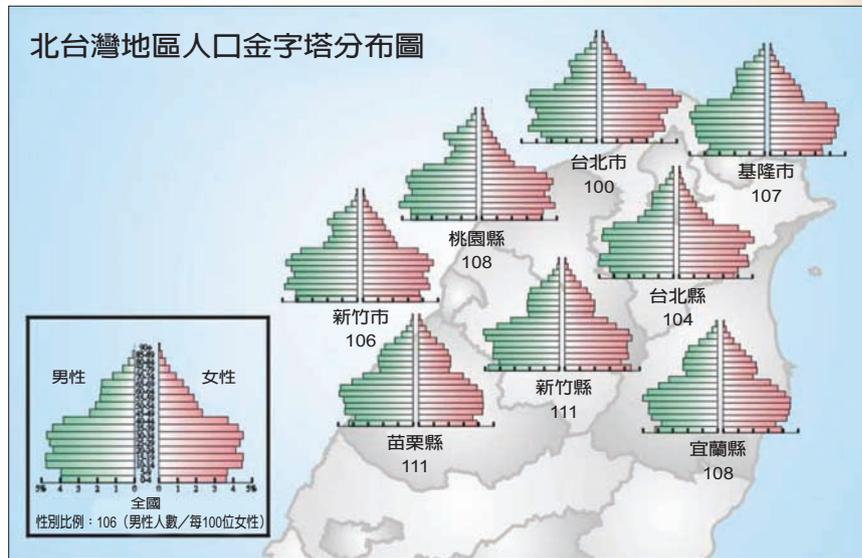
讓數字說實話 或謊話

若想了解上帝在想什麼，
就必須學統計，因為統計學就是在測量祂的旨意。

—南丁格爾

■陸偉明

在日常生活中會碰到很多資料或數字，你可以比大小、比多寡、或是比兩者的關聯性，都要靠統計。統計學 (statistics) 是我們了解這個世界的科學方法之一。她是來自 state, status 這個字源，最初做為描述人口、政府普查的資料。後來



人口金字塔

天文學家為了處理資料的變異性，才發展了初步的統計方法。在今天的國中課本中，都還有人口金字塔的圖，類似莖葉圖的開展形式，以表現出不同年齡層的人口數。

以下簡單介紹幾個和統計學有關的小故事，以及統計學的兩大支：描述統計和推論統計。

19世紀為了測量個別差異及生物的族群變異，統計技巧又更進一步。高爾登 (Galton, 英國人, 1822-1911) 是達爾文的表弟，他把父親的遺產用來旅行，曾到過非洲南部。最為人所知的是他發現了人類有指紋，他也丈量人類的身高，以了解遺傳的特性。1884年，他成立了人體測量實驗室。他不僅對身體上的特徵有興趣，也認為這



指紋 達爾文的表弟高爾登發現了人類有指紋，他也丈量人類的身高，以了解遺傳的相關性。

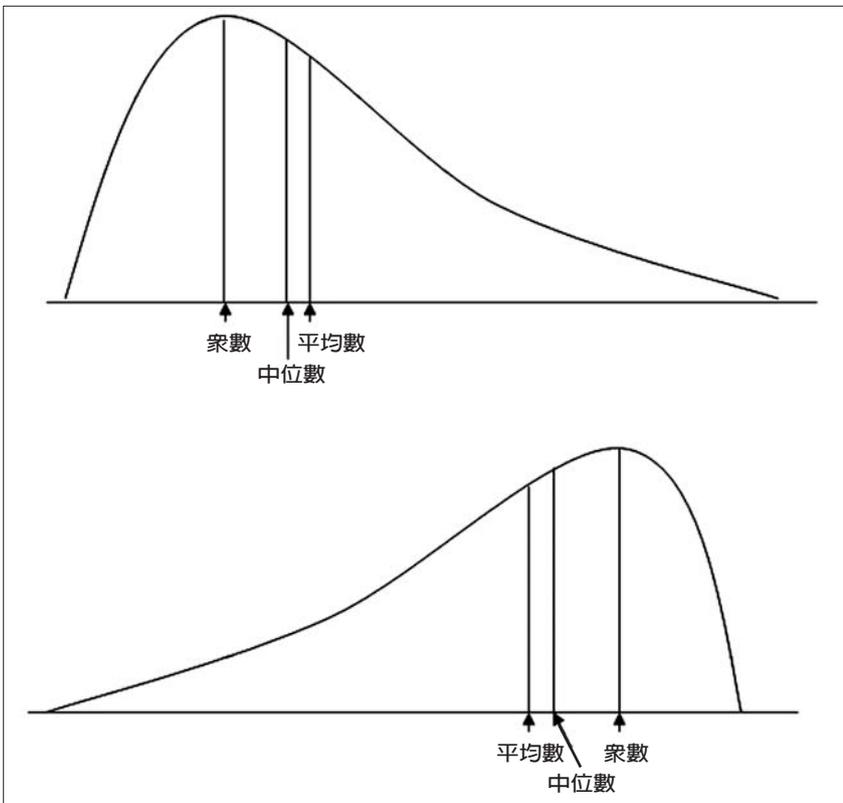
個世代的智力是與上個世代的智力「相關的」，是經由遺傳而來的。

高爾登雖然沒有發展出公式去計算今天我們所熟知的相關係數，但他發現非常高的父母所生的孩子，往往會比父母矮些，反而不是最高的，而非非常矮的父母所生的孩子，則往往比父母高，因而在1885年發現了迴歸到平均 (平凡) (regression toward the means (mediocrity)) 的現象。我們把 regression 翻譯成迴歸，後來成為非常重要的統計分析方法之一。

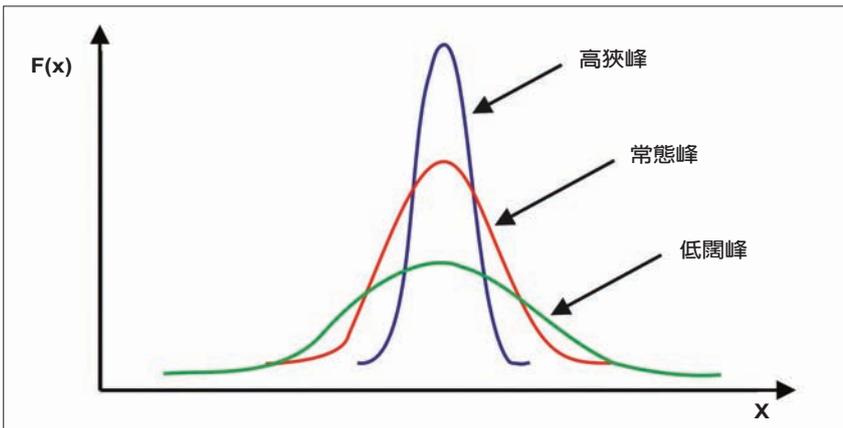
我們雖然知道生物變異是生命的基本特性，也是達爾文天擇理論的基礎。然而皮爾森 (Karl Pearson, 英國人, 1857-1936) 卻最先了解到變異是統計裡的重要本質。皮爾森描述分布函數的4個參數是平均數、標準差、對稱與峰度 (有分高狹峰與低闊峰兩種)。

皮爾森並把高爾登所發現的相關用更完整的形式表達出來，就是相關係數。1900年他還發明了適合度檢定，來檢定人類頭蓋骨的容量分布，這是最古老的統計檢定。他曾經丟銅板 24,000 次，結果正面出現 12,012 次！發現了隨機 (或機率) 現象，就是銅板正面出現的機率是 0.5005，他的耐力真是驚人。

另一個和統計學有關的是孟德爾修士 (Gregor Mendel) 用豌豆連續進行 8 年的雜交實



正偏（上圖）與負偏（下圖）的分布



高狹峰與低闊峰分布

驗。選擇實驗的材料是很重要的，豌豆是自花授粉，都是純種，且品種狀態差異性非常明顯，便於雜交實驗。他挑出了7對豌豆特性做為研究對象：高/矮莖、圓滑/皺縮種子、黃/綠子葉、葉腋/莖頂著衣、灰/白種皮、飽滿/皺縮豆莢、以及黃/綠未熟豆莢。8年內總共進行了225次交配實驗，並用統計方法處理了12,980個雜種！

他注意到豌豆經過雜交後，不同特徵的植株比例變化，會有一個數值關係。例如兩個親本雜交後的第一代，會完全表現出其中一個親本特性（如高莖），而雜交第二代的結果高莖有787株，矮莖有277株，兩種性狀的比例大約是3：1。1866年他發表〈植物雜種的研究〉，但論文在塵封34年後，才於1900年由另外3位科學家重新發現，世人公認遺

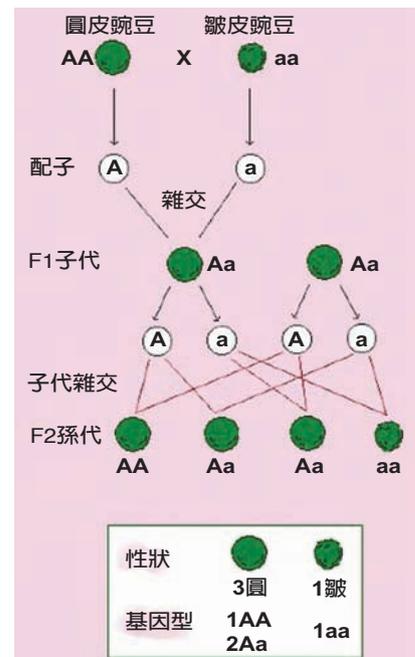


繪圖者：張雅鈴

孟德爾修士用豌豆連續進行8年的雜交實驗，他注意到豌豆經過雜交後，不同特徵的植株比例變化，會有一個數值關係。

傳學誕生了。

還有一個和統計學有關的生物發現，是盧瑞亞（Luria）發現細菌可以自然突變，這是他從「吃角子老虎」中獎機率是隨機的而領悟到的。因為若細菌會自然突變，它們的子代會有其特性，且會集中在其家族裡，而不是隨機分布在各家族裡。他在1969



孟德爾的豌豆實驗也和統計學有關

年獲得諾貝爾生理醫學獎，所指導的第一個研究生就是發現DNA雙螺旋結構的華森。

「尋找系統、登高望遠」是科學研究的法門之一，希望以上的小故事能提供讀者一些對統計學的興味。以下接著簡介統計學的基本原理和應用的幾個方向。在這裡先說一個故事。

曾任教育部部長、現任中央研究院副院長的曾志朗院士，在他的《人人都是科學人》一書中對「安太歲」這件事有生動的描述。他曾經爲了不想拂逆長輩的好意，私下進行了一個研究，比較這12年來每年沖到太歲和沒有沖到的人，在當年度的厄運是否有所不同。

【定義變項】 首先，界定什麼是厄運，如失業、失竊、交通意外、生病、考不好等。

【抽取樣本】 再根據不同職業篩選15、25、45及65歲4個年齡層抽取一些樣本。

【蒐集資料】 再根據大家的生肖分成沖和不沖太歲兩大類別，並計算厄運發生的次數。

【統計檢定】 最後統計結果顯示，年齡、行業和沖不沖太歲3個變項都沒有達到顯著差異。

【結論】 結論是，有沒有沖到太歲，和該年所遭遇到的厄運，兩者之間並無法證明有顯著的關係。

這個故事其實是在描述一個統計推論的過程，很生活化但富有科學教育的意義。而更精彩的還在後頭。

長輩認爲統計分析的結果雖然如此，但並不能說一定沒有這回事（寧可信其有，即使碰到的機率很低），或者是沖到的人都已經安過太歲了，也就是樣本大概受到污染了（偏差樣本）。

曾院士回去後再把已安過太歲的樣本剔除，重新做了統計檢定，結果兩組還是沒有顯著差異。

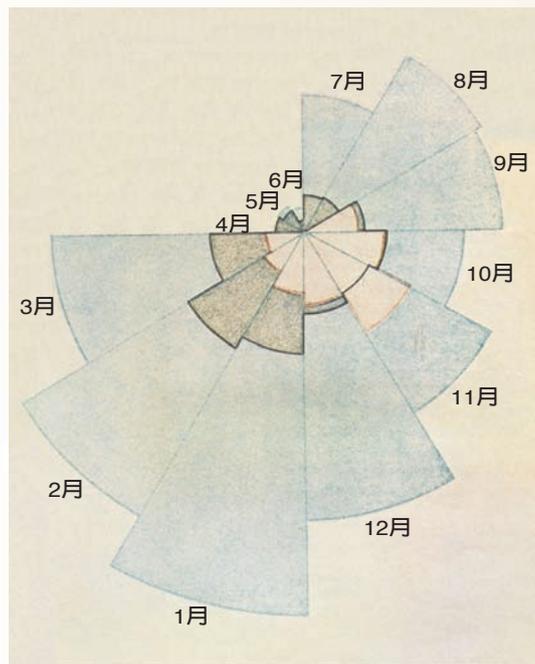
統計學是從數據中獲取資訊的科學，分爲兩大支，一支是描述統計，一支是推論統計。前者是我們在小學高年級時就開始接觸的平均數、圓餅圖、直方圖、長條圖等，主要用來描述資料的性質。而後者是要從現有資料中導出結論。

圓餅圖是南丁格爾（Florence Nightingale, 1820-1910）爲了要求英軍在戰地開設醫院所設計出來的報告方式，藉以強調她在護理上的成果。若是要注意到觀察值的分布（正／負偏）、集中的情形、離散點等特性，John Tukey（1915-2000）所設計的莖葉圖則有很好的效果。頻率分配及直方圖可以看出連續資料的分布情形，是進行資料



繪圖者：張雅鈴

南丁格爾



圓餅圖是南丁格爾爲了要求英軍在戰地開設醫院所設計出來的報告方式，藉以強調她在護理上的成果。

分析前很重要的步驟，不可省略。兩個連續變項的關係，則可以使用散布圖作圖。

許多報章雜誌常有很多精美的圖表，可以達到一圖勝千字的效果，但是別忘了統計會讓數字說話，也會讓數字說謊話。如長條圖會因為座標尺度的大小而在視覺上有不同的效果，因此在製作時要特別注意。

推論統計則是植基於機率論，是經過了抽取樣本、考驗假設的過程。我們對有興趣的現象，可以經由觀察一些變項，並記錄下資料，這些都是從樣本中得來的。樣本是一組觀察值由母群體抽出，具有母群體的代表性。母群體有其異於其他群體的特徵，如中小學老師，指的是在中小學而不是在大學教書的老師；如癌症病人，就與精神科病人不同。

樣本要能代表母群體特性，就必須有代表性，要能真正代表母群體，指的是在母群體中的每一單位都要有相同抽到的機會，這就叫做隨機抽樣。而廣播電視常見的call in，是一種偏差樣本，因為它並不是隨機抽取而是自願表達意見的，所以欠缺代表性。

機率是要了解母群體中抽出某項的機會為何，而統計則是基於樣本的狀況去推論母群體為何。也就是說，統計使用的是演繹法的精神。

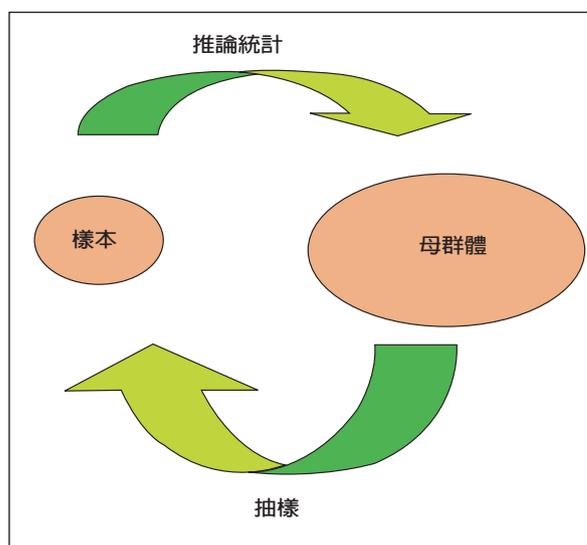
前提（假設）：樣本的狀況是如此如此。

結論：母群體的状况。

這種推理方式，對沒有受過邏輯訓練的人來說實在是很彀扭、不直觀，但這就是科學論證、推理的方法。有些學者認為中國雖然發明過許多精巧的技術，但因為沒有發展出推理、演繹的思維形式，所以一直到近代還沒能產生真正的科學！

母群體的母數，在希臘文是幾乎測量的意思，我們只能從樣本中去估計母數。17世紀開始，從賭博的問題發展了機率論，成為統計一個很重要的基礎。

機率分布是長期的頻率分布，使我們對隨機樣本有更好的預測。如最

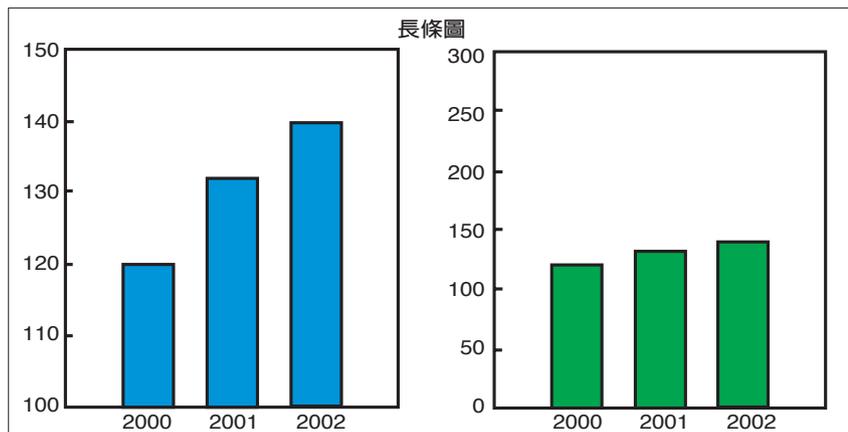


統計和機率的關係

常被使用的常態分布，又稱做高斯分布。是為了紀念偉大的數學家高斯，他在18歲時發明最小平方法，並曾用常態分布研究天文觀察中誤差分布的情形。這個分布是鐘形曲線、兩邊對稱的分布。大樣本觀察值的分布，很多是鐘型曲線。

常態分布的觀念在心理測驗中被大量運用，以制訂常模並決定受試者的相對位置（高下）。頗負盛名甚至可說是惡名昭彰的智力IQ，是以100分為平均數、15分為標準差的常模來「決定」學生的聰明才智。其實要發展智力，必須累積知識與活動，若能自動自發做一個好奇寶寶，我們大腦的智力會不斷發展！

至於統計分布則有Z分布、t分布等。當蒐集資料後，就可以計算出統計檢定量，並和某種特定統計分布比較，以考驗（檢定）我們的科學問題。波柏（Karl Raimund Popper, 1902-1994）從邏輯的觀點對「科學性」下了一個定律，即科學性的問題必須有否定的形式。波柏認為要正面證明科學的定律或研究者假設（ H_0 , alterna-



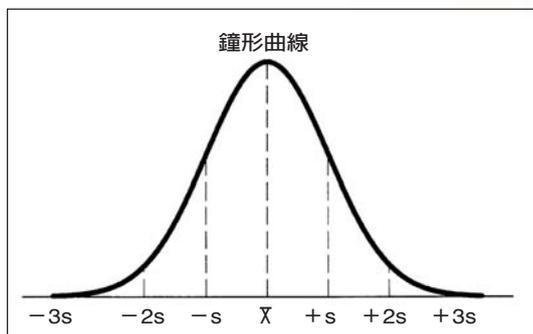
左圖描述的是這3年來的犯罪數，您會不會感到節節上升而恐懼呢？但若是右圖，就感覺好像沒有那麼嚴重，而實際上是相同的資料呈現在不同的座標尺度上，因此不要被眼睛給騙了。

tive hypothesis) 很難, 但可以經否定虛無假設 H_0 (null hypothesis) 來間接支持 H_a , 這就是所謂的 Null Hypothesis Statistical Testing, NHST, 是推論統計最重要的精神與途徑。

通常在實驗設計時會有實驗組及控制組 (對照組), 以比較給予不同處理後的效果。但有時受到研究倫理的限制, 或者是處理隨機指派不能做到時, 必須建立一個模型來描述各種變項間的關係, 這稱為觀測研究。相關分析是最常見的, 但須注意高相關並不代表具有因果關係。比如說夏天的犯罪率和汽水的銷售量有高相關, 並不意味著汽水喝愈多就愈容易犯罪。迴歸分析可說是用來解釋現象裡變項間的關係最常用的統計方法, 變異數分析則是它的特例。

統計學所處理的就是從樣本中觀察及加計無所不在的誤差, 並把這不確定性納入統計模型, 做更佳的預測與解釋。比如說以一個描述學生成就的徑路分析模型而言, 前面兩個變項可能分別是父母的教育程度和收入, 接下來會影響到學生對自己的期望, 進而才影響到學業成就。而期望與成就都有誤差項, 顯示還有其他因素會影響到這兩個變項, 只是在這個模型中, 沒有考慮到的因素就歸為誤差。

機率論始於賭博, 但今天從決定論典範轉移到模糊 (fuzzy) 典範, 從品質管理、流行病學到氣象水文、入口網站 Google 的搜尋引擎、人壽保險費率、新藥有效性的比較、民意調查等, 這些都有充分的統計意義, 而非偶然的結果。統計真



有 68% 的觀察值會落在平均數上下加減 1 個標準差內, 而有 95% 的觀察值落在平均數上下加減 2 個標準差內, 99.7% 的觀察值落在平均數上下加減 3 個標準差內。其中 \bar{x} 是平均數, s 是標準差。

的改變了我們所認知的世界。

準實驗設計和實驗設計最大的不同在於沒有隨機化。其實隨機抽樣和隨機分派會使組內變異一致, 以利於實驗處理效果的比較。證明因果關係並不是那麼容易, 最有名的例子之一, 就是早年統計學大師費雪對抽煙導致肺癌的反駁。因為許多研究並非隨機化設計, 或發表的論文有一面倒的傾向, 事實就不容易浮現。

比如說追溯性的研究, 是由疾病開始去回溯哪些條件和疾病有關。要找到另一組對照組和本實驗組一致, 並不容易。另一種是追蹤性研究, 看看數年後的變化, 但能處理的也是很小的特定樣本, 有推論母群體的困難。不過雖然每個研究或許有些許瑕疵, 但若很多研究都趨向相同的結論時, 證據就明顯了。統合分析可以幫我們這個忙。

美國統計學學者杜奇 (Tukey) 有個名言: 「對正確的問題有近似的答案, 勝過對錯的問題有精確的答案。」這是很值得研究者深思的。「研究」(re-search), 顧名思義不就是從過往的文獻中再加入新的觀點或解決之道嗎?

學習統計其實就是學習科學思維、態度和處理問題的方法。有一種詞彙分析, 還可以把文學作品中的詞彙統計出來, 其規律性可以反映出作者的風格, 對文獻考古與版本確

定有極大的幫助, 甚至連是男性的口氣或女性的口氣都可以分析出來。現在, 你可以就本文約 5 千字裡所呈現的詞彙, 進行統計分析。假設是: 作者是男的還是女的? □

延伸閱讀

鄭惟厚譯 (2002) 《統計學的世界》, 台北: 天下文化。

鄭惟厚譯 (2003) 《看漫畫, 學統計》, 台北: 天下文化。

陸偉明

成功大學教育研究所

機率論始於賭博, 但今天從品質管理、流行病學到氣象水文、入口網站 Google 的搜尋引擎、人壽保險費率、新藥有效性的比較、民意調查等, 這些都有充分的統計意義, 而非偶然的結果。統計真的改變了我們所認知的世界。