

## 敘(描)述統計

主要資料來源：《赤裸裸的統計學》，Charles Wheelan 著(美)，曹 檳譯，中信出版社，2013。

國內生產毛額（英語：Gross Domestic Product，縮寫：GDP），亦稱國內生產總額、國內生產總值，在描述地區性生產時稱地區生產毛額，是指一定時期內（一個季度或一年）一個區域的經濟活動中所生產出全部最終成果（產品和勞務）的市場價值（market value）。國內生產毛額是國民經濟的核心指標，在衡量一個國家或地區經濟狀況和發展水準時，亦有相當的重要性。GDP 除以人口數量便可以得到人均 GDP，一般用作衡量不同國家或地區，經濟狀況與生活水平的指標。

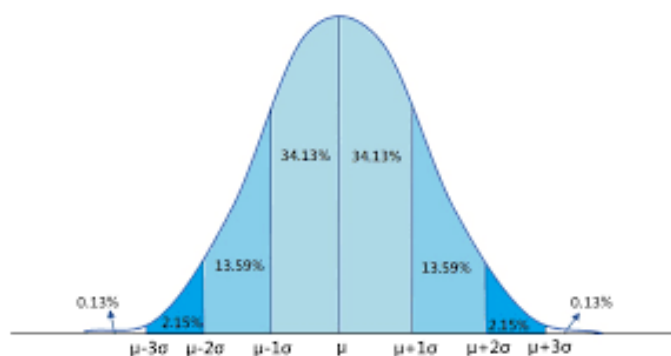
然而，人均 GDP 並沒有考慮通貨膨脹的因素，還有，我們需要知道的是普通人的收入，而不是泛泛的人均收入。因為如果貧富不均嚴重，處於收入排行榜中頂端 1% 富人收入的爆炸性增長，自然能夠拉動人均收入提升，這和較均衡的普遍性收入增長，二者是有本質上區別的。這點告訴我們，這些描述性的數據雖然為我們提供了一個針對某一現象的可操作、有意義的概括，但任何的簡化也會面臨誤導和被濫用的危險。

數據分佈的「中間位置」（平均值）是有一些問題的，它容易受到遠離中心區域「異常值」的干擾而出現失真。因此，我們不應該用人均收入來衡量中產階級的經濟狀況。好在統計學中還有一個數據也可用來表示分配的「中間位置」，那就是「中位數」，它是指依順序排列以後，位在 1/2 位置的數值（如果是偶數，就是位列中間二個數的平均）。如果一組數據分佈中沒有特別離譜的異常值，那麼它的中位數和平均數會是差不多的。中位數也有一些親戚，把數據分成 4 個部份，處於底部的 25% 就稱為第一四分位數，以此類推。同樣的，也可以分為 10 個部份，這時就稱為「十分位數」，分成 100 個部分當然稱就為百分位數，第 99 百分位數就是指收入最高的那 1%。

如果一位學生某科目考了 83 分，這是一個「絕對數字」，我們不需要知道其他學生考多少分，就可以對他的表現給出一個評價。如果我們說這位同學是排第 9 名，這就是一個「相對數字」，它處於一個更大背景中時才有意義，告訴我們有 8 位同學表現比他出色。例如，在台灣的大學學科能力測驗有五個科目，考生的每個科目有 0 至 15 的級分，得到相同總級分的考生還真是不少，但如果參考了百分位數（相對分數）就會對填寫志願幫助較多。

標準差是另一個能夠幫助我們，在一大堆雜亂無章的數字中，發現真理的指標，可以用來衡量數據相對於平均值的分散程度。在統計學裡最重要、最有

用、最常見的數據分佈稱為「常態分配(佈)」。亦即在一般情況下，數據的分佈都是呈對稱的，以平均數為中軸如「鐘」的形狀，而且有 68.2% 的數值位於平均數前後一個標準差的範圍之內、有 95.4% 在二個標準差之內、有 99.7% 在三個標準差內，聽起來很白痴，但事實上就是如此，應用範圍可大了！



常態分配圖

敘(描)述統計經常需要比較二個數據或數量，例如張三比李四高 5 公分、今天比昨天降  $5^{\circ}\text{C}$  等，由於我們對其中所包含的數量單位並不陌生，所以很容易理解。但如果說 A 牌的麥片比 B 牌的納含量高 31 毫克，除非你內行，否則並不能給你帶來多少具體的訊息。

如果你知道賣場將某商品價格上調 20%，然後特價打 8 折，請問就是原價嗎？這種百分數的變動表示的是某個數字相對於其他事物的變化值，關鍵是我們要弄清楚所謂的其他事物到底是什麼。假設國防預算由 3% 上調為 5%，我們可說上調了 2%，(即  $5\% - 3\%$ )，也可以說上調了 67%，即  $(5-3)/3 = 0.67$ ，其實這是百分差和百分率的不同，二者當然不可以混為一談。但是不同的評論者，就會加以利用，提出不同的分析和解說。

將一系列複雜的訊息濃縮成一個數字，這是所有指數都具備的優點，然後可以對原先無法展開簡單比較的事物進行排名，例如比較大學的優劣、奧運會各國的表現、國家的貧富、甚至主要國家軍力的比較。當然，用各種各樣的方式來濃縮訊息，也同樣有缺點，因為採用不同的方式，就有可能導致不同的結果。所有指數都是取其構成的描述性數據以及它們的權重，任何一點的些微變化，都有可能引起結果的改變。因此，這些指數有可能是情況雖不完善但有現實意義，也有可能是完全偏頗的。

還有一個必須分清楚的概念就是在評估經濟狀況的過程中，我們要特別注意，千萬不要把「收入」和「工資」搞混，這二者是不相同的，工資是我們付出固定金額的勞動所得，如時薪、月薪；收入則是全部所得的總和，可能有多

種來源。如果一個人有另一份兼職工作或者常常要加班，那他的收入會增加，但工資並沒有變。也就是說，一個人可能收入會增多，但工資沒有發生變化；或者工資下降了，但加班增多了，他的收入依然有可能上升。因此，如果一個人不得不付出更多的勞動力來取得更多的收入，我們就很難評價他的生活質量是更好還是更壞。好在我們還有其他指標如：中位數工資、第 90 百分位數…等，分析結論就不致於共同指向一個唯一「正確」的答案。