

相關關係

資料來源：《誤區---思維中常犯的6個基本錯誤》

Thomas Kida 原著，張雲譯，2011年，人民郵電出版社

如果兩個變量之間確實有關聯，衡量關聯程度的最佳方法是什麼？統計學家發展出了一種指標稱為「相關係數」，其範圍從 $-1 \sim +1$ 。相關係數為 0，意味著彼此沒有關係，如果一個變量上升另一個也上升，相關係數就是正，一個上升時另一個下降，相關係數就是負，當係數愈接近 $+1$ 或 -1 ，就代表相關程度越高。

相關係數為我們提供了一個判斷兩個變數是否相關的最佳指標，但是在解釋相關性時我們必須牢記：「有相關關係並不表示有因果關係」。我們千萬不要誤以為，如果兩個變量是相關的，那麼就代表一個變量導致了另一個。例如，一家公司在廣告上投入的經費和產生的銷量之間存在相關性，就認為是廣告導致了銷量，其實也有可能是因為產品的品質提升才使銷量增加，只是廣告活動恰巧也在同一期間進行。即使一個變量與另一個變量確實存在某種相關性，也並不表示我們可以說是哪個導致哪個。天生尋找原因的傾向，很容易導致我們從相關性得出因果的推論，這是我們必須堅決抵制的誘因！

相關性有時還潛藏有欺騙的性質。兩個變量的確彼此相關，但不是因為它們有直接的因果關係，而是因為這二個變量都與另一個變量(即第三變量)相關。例如人均煙草消費量高的國家，肺癌比例也高，是正相關，人們就說抽煙會導致肺癌。但人均煙草消費量愈高，霍亂的病例就愈低，呈負相關，難道是抽煙可以預防霍亂？當然不是！想想看，真正的原因是什麼呢？

即使實踐證明兩件事情確實彼此相關，我們仍然要審慎評價，認清它們關聯性的數據是如何計算出來的。政客和特殊利益團體一直企圖說服民眾，他們在某個問題上的立場是正確的，他們經常使用統計數據來支持自己的觀點。如果不明白那些統計數據是如何計算出來的，就很容易被誤導去相信某些並非事實的事情。例如疫苗與效價、利率與通漲、失業率與經濟景氣、溫室氣體排放與氣候暖化、股票線形走勢與未來股價、股市漲跌與政治局勢……等等，實在太多了！