

抽樣及非抽樣誤差

抽樣的目的是以有限的樣本來替代龐大的母群體，並以樣本指標(統計量)來代替母群體指標(參數)。可是在進行抽樣時，不可避免的會產生誤差，因為從母群體中隨機抽出來的樣本，當然不可能和母群體完全一致。這種誤差又分為二種，一種是由於主觀因素破壞了隨機原則而產生的誤差，稱為系統誤差

(**System error**)；另一種是由於抽樣的隨機性引起偶然的代表性誤差，一般稱為抽樣誤差(**Sampling error**)。雖然抽樣誤差不可避免，但是可以運用大數定律精確的計算，確定它具體的數量界限，並可通過抽樣設計加以控制(例如選擇隨機樣本的大小)。所以隨機抽樣誤差是隨機抽樣這個動作造成的誤差，是樣本統計量和母體參數之間的差距，也就是信賴敘述中所稱的誤差界限。

大部份的抽樣調查也會有隨機誤差以外的誤差，甚至連普查都逃不過的差錯，稱為非抽樣誤差。非抽樣誤差包括處理誤差(**Processing error**)，就是指人們在資訊處理過程中犯的錯誤，例如書寫錯誤、計算錯誤、鍵入錯誤……等，只要足夠小心注意和電腦應用的普及均有助於減少這類誤差。

一般在作訪問調查時還有一種非抽樣誤差來源就是無回應(**nonresponse**)，這是指隨機樣本中因為種種原因而無法接近而得不到資料。例如電話訪問中有經常不在家的、不接電話的、電話只是答錄機的、電話未獲轉接的、接了電話拒絕合作的……等等。人員訪問中也有豪宅根本無法接近的、社區保全不放行的、以及不願接受訪問的……等等。尤其現在詐騙電話猖獗、有線電話面臨淘汰、手機來電篩選……等的社會環境變遷下，使得這種無回應的困擾愈來愈嚴重。

受訪者就算回應了還是會有另一種誤差，我們稱為回應誤差(**response error**)，這也是一種非抽樣誤差，是指受訪對象給了不正確的答案。例如受訪者回報的年齡或收入不實在、誤解了或未真正理解問題在問什麼、自己記不得了就隨便答、給的答案只是亂猜的、對敏感議題有意給一個政治正確或訪問者希望的答案……等等。

當然還有一項對抽樣調查結果有影響的就是問卷的設計。問卷中的問題是怎麼問的、問題的措詞、先後順序的安排、主辦的機構……等等都可以影響受訪者的回答。然而各種團體使用具有引導傾向的問題及自發性回應樣本的情況卻有增無減。要做一個客觀公正的調查其實並不容易，他們必需要認真準備抽樣底冊、用心的設計問卷、減少無回應、使用技巧的統計分析，那麼調查結果還是可以提供精確且有用資訊的。

以後我們看到任何的民意調查報告時，一定要注意：

- 誰做的？
- 什麼時候做的？
- 母群體是什麼？
- 樣本多大？
- 樣本怎麼選取的？
- 用什麼方式做的調查（郵寄、電話、人員、網路……）？
- 有效樣本率是多少？
- 調查時使用的問卷？（注意問題是怎麼問的、措詞、順序……等）
- 使用什麼統計分析方法？
- 下結論的依據是什麼？

以上事項如果沒有說清楚也查不到，那麼我們最好對這種調查結果抱持懷疑態度，因為不是有意隱藏就是水準不夠！