

抽樣調查的方法

(資料來源：龍騰教師手冊)

抽樣調查是很重要的一種觀測研究，西方有句諺語「你不必吃完整條牛才知道肉是老的」。抽樣的精髓正是：從檢查一部分來得知全體。但是正如瞎子摸象的寓言，抽到的樣本是不是有代表性是我們所關切的重點。我們先看看下面兩個案例：

案例 A：一名研究員抽樣調查各監獄中因偷竊而入獄者的智商，作出如下結論：「小偷的平均智商低於一般人。」你覺得這個調查可不可信？

案例 B：一個電視節目中，以現場 call-in 方式統計民眾是否曾經接到過詐騙電話，節目結束時的總結「有高達八成的民眾曾經接到過詐騙電話」。這個結論合理嗎？

案例 A 中的研究員顯然直接選擇容易取得的樣本做觀測，這種樣本可稱為方便性樣本。問題是這些方便做觀測的小偷，或許原來就因為智商較低而被捉入獄，與逍遙在外的偷兒們有相當的差距，或許他們不願意認真作答，得到的結論當然不可信。案例 B 也是壞樣本的例子，我們稱之為自發性回應樣本。打電話 call-in 的人通常是上過當或有強烈的負面感覺，比較會不嫌麻煩的去做回應，這個結論的比例顯然偏高。「自發性回應樣本」是由人們自行決定要不要回應；而「方便樣本」則是由訪談者來決定，兩者都由於人為因素易造成偏差。統計學家於是利用機率來挑樣本，藉著不牽涉人為選擇的機率來消除這些偏差。

常用的抽樣方法有簡單隨機抽樣、系統抽樣、分層抽樣及部落抽樣四種。除簡單隨機抽樣外，其餘三種抽樣方法依序介紹如下：

(1) 系統抽樣

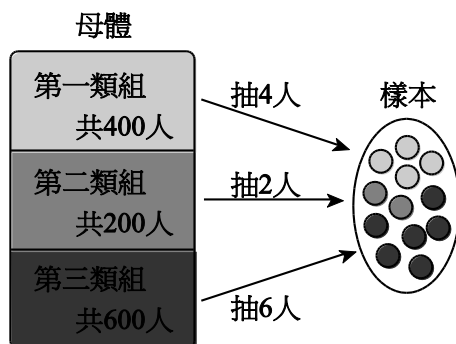
系統抽樣基本上是只做一次簡單隨機抽樣後，就採取依固定間隔數抽出樣本。例如在 50 人的母體中欲抽出 5 個樣本，我們將母體分成 5 區（每一區有 10 人），先在 1 到 10 號中抽出一個號碼，比如 4，就依間隔 10 的順序叫出下列號碼：

(1... 4 ...10), (11... 14 ...20), (21... 24 ...30), (31... 34 ...40), (41... 44 ...50)

系統抽樣的優點是只需抽出一個號碼，其他號碼就跟著決定了，在欲抽出的樣本很多時，顯然較簡單隨機抽樣方便。值得注意的是當資料有某種特性或規律時，抽出的樣本可能無法代表母體。例如百貨公司的顧客人數，每逢週末假日的人數就比平常日多，像這樣母體有循環性時，選擇系統抽樣比較容易形成偏差。

(2) 分層抽樣

當我們對母體的資料掌握了更多訊息時，一個更詳細的抽樣方法是分層抽樣。這個抽樣方法是將母體分成數層，然後從每一層抽出簡單隨機樣本合併起來。分層的原則是：同一層的個體有較接近的特質。例如調查全校高二同學每週的平均補習時數，將高二同學依所讀類組分作第一類組，第二類組，第三類組三層，如圖 1 所示。若想從中抽取 12 人，則依各層人數比例抽取如下：

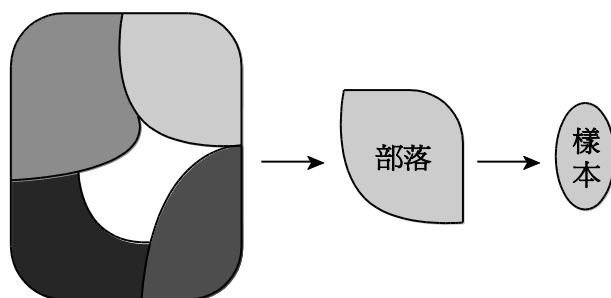


⬆ 圖 1

可以看到，抽出的樣本涵蓋了母體各層，而樣本數的比即該層在母體中的比，同時因為同一層的個體相似性高，抽出的樣本較具代表性，如此更能充分反映母體的情況，在實際調查中是被廣泛應用的一種抽樣方法。

(3) 部落抽樣

部落抽樣是將母體按某種標準分成數個差異甚小的組，每組稱為一個部落，然後從中隨機抽取若干部落，再對這些部落做普查或抽查，如圖 2 所示。區分部落的原則與分層方法剛好相反，分層時要求「物以類聚」，分部落則希望同一部落中包含較多有差異性的個體，而部落與部落間的差異不大。換句話說，讓部落成為母體的縮影，如此便能充分反映母體。例如欲調查全校高二同學的身高時，因為分班情形與身高無關，我們可將各班級視為一個部落，任意抽出一個班的同學，再以簡單隨機抽樣從中抽出樣本即可。



⬆ 圖 2

要驗證部落間的差異是否很小有時相當困難，不過部落抽樣的好處是：當母體非常龐大以致其他抽樣法不易實施，部落抽樣顯然經濟得多。

上述的抽樣方法各有其優缺點與實施上的限制，所以在實際的抽樣調查中，往往會視情況需要，結合數種方法來進行抽樣。