

## 隨機抽樣的附帶問題

隨機抽樣的方法和注意事項已有所討論，現在讓我們再來補充說明一些真正實施隨機抽樣時相關的附帶問題：

### 一、樣本大小

隨機抽樣的樣本數應該要抽多少？這就是樣本大小的問題。一般來說，樣本數愈大，則以樣本推估母群體時，抽樣誤差會愈小，但相對的時間和經費要花的愈多。相對的，樣本數愈少，可以節省資源和時間，但抽樣誤差就比較大。誤差如果太大，所得結果就會失去實用價值，一個不必要的大樣本只是浪費資源而已。因此，適當的樣本數應依需要或實務上的限制來決定。

假設一個 50 位男同學的班級真正平均身高是 170.2 公分，由班上隨機抽出 5 位同學作為樣本，那麼這 5 位同學的平均身高可能大於或小於全班同學真正的平均身高，其間的差異即為抽樣誤差。雖然每次抽出來的 5 個樣本不會一樣，但我們可以利用統計理論知道這個誤差會在什麼範圍。如果我們將樣本數增加到 49 位，那麼不管抽到的是哪 49 位同學，他們的平均身高和真值相差一定是非常的有限，由此我們可以了解樣本數愈大抽樣誤差愈小的涵意。

選舉前的民調也是這樣，隨機抽樣得到結果的誤差只要在可接受的範圍(例如  $\pm 3\%$ )內即可，抽樣的樣本數目所需的時間和預算又可承擔，那麼就是可行的了。同學們，你能夠依這些要求把所需的樣本數計算出來嗎？然後再比較看看民調機構實際上抽了多少樣本。

### 二、放回還是不放回

以上述由班上隨機抽 5 個同學為例，如果我們是用亂數表來抽學生座號，同樣號碼第二次出現時就把它刪除，這其實相當於做 50 支籤再從中抽 5 支一樣。如果是做籤，一次只抽一支，被抽到的就不再放回籤箱中或是一次就從箱中抽出 5 支，那麼就不會發生重號的現象。像這樣抽出後不放回籤箱的抽法，因為母群體一共有 50 個，籤就會愈抽愈少，這種情況樣本就是來自有限母群體。

如果抽出一支後記錄號碼，再將籤放回抽籤箱中的話，同一支籤

就有再度被抽中的可能，雖然母群體只有 50 個，但永遠也抽不完，這時就稱為樣本是來自無限母群體。在統計理論中，由有限母群體和由無限母群體中抽樣所適用的公式是不同的，由無限母群體中抽樣比較簡單。在實際運用時(如摸彩)，不放回的方式比較常見，在母群體相對遠大於樣本數時，放回與不放回的差別並不大，所以我們通常都是以無限母群體的公式來作計算。

### 三、多段式隨機抽樣

當母群體相當大時，為了配合實際需要，我們有時必需混合利用不同的隨機抽樣方法。例如，要了解學生的視力，可以先用簡單隨機抽樣抽出若干學校，再從被抽出的學校以年級來分層、以班級為集群的方式，隨機抽出若干班作全班的調查。這樣實施起來就方便得多，這種混合方式的隨機抽樣也可稱為「多段抽樣法」。

### 四、很難編號時

有的情況很難或根本無法對母群體的全部個體加以編號，例如某一養殖池中的魚類或某一林區中的所有樹木等，我們可以在地圖上以劃定區域的方式，然後挑選接近被抽到區域位置中的個體作為樣本。例如對於一群會游動的魚，我們可以指定在那些區域最先發現的若干條作為樣本，以作為實務上的變通。

有時，我們會對樣本毫無選擇的餘地。例如一位醫師要研究某種疾病的治療方法和效果，他不可能將全國患有這種疾病的病人隨機抽到他的醫院作為他的病人，只能以自己找他診治的病人作為樣本，然後推敲這些樣本(非隨機)可能歸屬的群體，所以他的研究結果也只能推論到這個群體而非所有患者。