

## 卡方檢定

卡方檢定(  $\chi^2$  test )是應用最廣泛的統計方法之一，尤其在檢定兩個類別之間是否具有關聯性 ( association ) 時更為有用。要注意的是它只適用於諸如人數或次數等的間斷變數資料(由類別變數點計而來)，不適用於處理連續變數的資料。

在醫學研究或某些研究領域裡，我們常會遇到需要比較「實際觀察次數」(observed frequency)與「理論次數」(theoretical frequency)之間是否有差異的問題，所謂「觀察次數」是指在實驗、調查或實際觀察中所得到的實際次數，而「理論次數」則是指根據機率理論、某種函數或某種理論曲線得到的應有次數。例如，假設擬研究某疾病的 A、B 二種治療方法的治療效果是否不同，根據實際採用這二種治療方法得到的結果如下表：

治療效果	治療法 A	治療法 B	合計
有效	80 (74)	105(111)	185
無效	20 (26)	45( 39)	65
合計	100	150	250

因為「治療方法」和「治療效果」均為類別變數，而表格內沒括弧的數字即為實際的觀察次數。亦即：在 100 位採用治療法 A 的患者中實際上有 80 位有效，20 位無效，有效率為 80%，而治療法 B 的有效比例為 70%(105/150)，看起來當然是不一樣，但要確定此差異是不是抽樣變異造成的，我們就可以作卡方檢定。首先，我們先假設這兩種治療方法並無差異，那麼理論上的有效率應為  $185/250 = 0.74$ ，所以 A 療法的 100 名患者理論上有效的就應該是 74 人，B 療法 150 名患者理論上有效的就應該有 111 人，如括弧中的數字所示，無效者亦同。

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$$

然後我們可以利用皮爾森(Pearson)卡方檢定公式來作檢定，即：

式中 O 代表觀察次數(Observed Frequency)，E 代表期望次數(Expect Frequency)，自由度  $df = (r-1)(c-1)$ ，亦即「行數減 1」乘以「列數減 1」。

此時，

$$\begin{aligned}\chi^2 &= (80 - 74)^2/74 + (105 - 111)^2/111 + (20 - 26)^2/26 + (45 - 39)^2/39 \\ &= 0.662 + 0.324 + 1.385 + 0.923 \\ &= 3.294 \\ df &= (2-1) \times (2-1) = 1\end{aligned}$$

然後，我們可以在任一本統計學教科書的附表中查卡方分配表，在  $df=1$  時，卡方值大於 3.84146 的機率  $p$  為 0.050，而本例卡方值小於此，表示機率是大於 0.05 的。若我們的顯著水準  $\alpha$  設定為 0.05，就應該接受虛無假設，亦即 A、B 兩種療法之間療效並無差異。

皮爾森卡方檢定是一個近似的檢定，所以只有大樣本才適用此法，一般而言，必需注意以下二點：

### 1. 所需樣本大小

每個細格(cell)內的期望次數(E)均須大於 1，且至少有 80%以上的期望次數等於或大於 5 以上，否則不宜使用。不宜使用的原因為：樣本數可能太少，此時最好的方法就是再增大樣本數或者將鄰近的幾個細格予以合併。或者是雖然總樣本數不小，但某事件發生的機率卻太小，此時增大樣本數也無濟於事，最好是改用費雪恰當檢定(Fisher's Exact Test)。

### 2. 連續性校正

如果自由度為 1 時，必須使用，也就是一般所謂的「連續性校正」，因為理論上  $\chi^2$  分配表中的  $\chi^2$  值是近似值，當樣本數夠大時沒有問題，可是當自由度為 1 或樣本太小時，近似起來就要比其他自由度時為差，所以要加以校正，校正的方法就是，在計算每一小格  $\chi^2$  值時，把分子的公式由  $(O - E)^2$  改成  $(|O - E| - 0.5)^2$ ，這樣算出來的  $\chi^2$  值會略微減小。以上例來說，自由度雖為 1，我們檢定的結果是接受虛無假設，如作了連續性校正，那算出來的  $\chi^2$  值還會再小一些，不會影響接受虛無假設的結果。

卡方檢定適用於做關聯性檢定( Test of Association)、適合度檢定(Test of Goodness-of fit)、獨立性檢定(Test of Independence)和同質性檢定(Test of homogeneity)。

