

評估以參數方式建構兩種診斷試劑準確度檢定

巫品秀*、黃怡婷、汪群超
國立臺北大學統計學系

摘要

預防醫學常會使用診斷試劑來了解受測者疾病或患病狀況，良好準確的診斷試劑能及早發現疾病的發生，讓病人獲得即時的治療。一般透過比較受測者實際患病狀況及診斷試劑判定結果的一致性來評估試劑的準確度，然而，依資料型態的不同，其型態可分為二元、順序及連續三種型態，評估試劑準確度的評估方式會有不同。當判定結果為二元時，敏感度和專一度是兩個衡量準確度的基本準則。當判定結果為連續型或順序型時，則透過臨界值的選擇來設定患病與否的標準；隨著臨界值不同，檢驗試劑的判定標準也會不同，此時，接收操作者特徵曲線(Receiver Operating Characteristic Curve，簡稱ROC 曲線)便是一個常用於評估診斷試劑準確度的工具。本論文利用參數化方法來建構ROC 曲線，假設患病和未患病受測者的判定結果服從某一母數分配族，其分配的特徵由位置參數和尺度參數決定，且具備對稱性；因此，可透過標準化的過程，將分配平移至位置參數為 0、尺度參數為 1 的標準化分配，並以 F_0 表示標準化後的分配函數。假設患病受試者和未患病受試者的試驗結果分別服從分配函數 F_D 和 F_N ，在給定偽陽性率為 t 之下，ROC 曲線可表示為

$$\text{ROC}(t) = \overline{F}_D(\overline{F}_N^{-1}(t)),$$

建構出 ROC 曲線後，便可利用 ROC 曲線下面積(Area Under the ROC Curve，簡稱 AUC)來作為診斷試劑準確度的總結指標。在參數化假設下，ROC 曲線會變成可由 a 、 b 兩參數來描述，其中 a 、 b 為位置參數及尺度參數的函數，對此兩個參數的估計，可透過 Green 和 Swet (1966)所設定的信號診斷理論之架構來建構資料的概似函數，進而求得最大概似估計式。本論文推導出兩個參數的估計式，及其大樣本分配。再者，當一種病症有兩種診斷試劑可以做為判定患病與否的依據時，即可透過總結準確度的指標直接比較兩種診斷試劑的優劣。本論文將就不同資料型態，設定比較兩種診斷試劑的準確度的檢定與其表現。

關鍵詞：AUC、Binormal 模型、ROC 曲線、敏感度、專一度、診斷試劑