



「IRT試題分析的理論與實作」研習班

陳怡婷／國立教育研究院籌備處約聘幹事

在教育研究中，常需應用測驗來瞭解研究對象的特質與屬性。而根據解釋測驗結果測驗分數時所用的數學模式和理論的不同，測驗理論可以分為古典測驗理論和現代測驗理論（又稱試題反應理論）。此次很榮幸邀請到目前於教育部台灣省中等學校教師研習會擔任助理研究員的曾建銘博士，於8月9日在本處傳習苑202電腦教室，為本處同仁介紹試題反應理論（IRT），總計30位同仁參與課程。

上課內容共分為兩大部分，上午課程主要講述測驗理論發展的趨勢、古典測驗理論與試題反應理論的基本概念與兩者各自的優缺點、試題反應模式及其相關特性、試題參數與能力的估計、訊息函數、題庫的建立等主題。下午課程則使用BILOG-MG軟體進行IRT試題分析的實作。課程中所介紹的試題特徵曲線（ICC）是用來描述測驗所欲測量的潛在特質（能力），與其在試題上正確反應機率間的一種數學關係。常見的試題反應模式有三種：（1）一個參數對數形模式：主要是估計試題難度參數 b 值；（2）兩個參數對數形模式：主要是估計試題難度參數 b 值、試題鑑別度參數 a 值；（3）三個參數對數形模式：主要是估計試題難度參數 b 值、試題鑑別度參數 a 值、試題猜測度參數 c 值。透過試題分析除了可以提高測驗的品質外，尚有一項重要的價值，那就是可以選擇優良的題目來建立題庫，以供日後進行適性化之測量。電腦化適性測驗（Computerized Adaptive Testing, CAT）就是一個有意義的應用，受試者在接受電腦測驗時，電腦會先從題庫中選取一個試題進行測驗，如果答對，就選取一個較難的題目進行測驗；如果答錯，就選取一個較簡單的題目進行測驗，如此持續測驗的過程直到受試者的能力被精確的估計出來為止。

雖然試題反應理論相較於古典測驗理論擁有許多優點，例如：試題反應理論所採用的試題參數是一種不受樣本種類影響的指標、試題反應理論所估計的考生能力值也不受測驗的影響，對不同測驗及考生之測驗分數可以進行有意義的比較等。但由於試題反應理論的應用需要有複雜的數學運算和相當龐大的樣本，因此古典測驗理論目前還是被普遍所採用的。最後，非常感謝曾建銘老師此次精采的授課內容，著實讓本處同仁受益良多。



95.08.09 本處辦理IRT技術分析法之電腦講座。(施培元攝影)