

# 多向度測驗等化及試題參數估計

李信宏

國立彰化師範大學數學系

## 摘要

單向性(unidimensionality)是試題反應理論(item response theory, IRT)的重要假設之一，然而大多數的測驗評量中，試題可能包括了幾個不同向度的能力或者特質，意即測驗是多向度(multidimensionality)的。這些多向度測驗資料如果以 IRT 模式及 BILOG 等軟體來分析，明顯並不符合單向性假設，其可能造成的估計偏誤是值得探討的。本研究以二向度測驗模式為主，結合不同的等化與連結方法，在模擬生成兩份題本的作答資料後，透過定錨題設計(common-item design)等化，再觀察估計這些二向度測驗試題參數的結果。模擬研究中以不同的試題結構設計，分別生成一種單向度與四種二向度的試題作答資料，再利用 BILOG 和 NOHARM 兩種軟體來估計多向度鑑別度參數 MDISC 和多向度難度參數 MDIFF。最後，以參數估計值的偏誤(bias)與根均方差(RMSE)的大小作為評價準則。

關鍵詞：試題反應理論、多向度、等化