

壹、緒論

國內兩個大型資料庫「台灣高等教育資料庫」與「台灣教育長期追蹤資料庫」(Taiwan Education Panel Survey, TEPS)，最主要的特徵是有非常完善的抽樣架構(中央研究院人社中心調查研究專題中心，2014；張苙雲，2003)，分別是橫斷面的學校、科系與學生，縱貫面的學校、學生與不同波次的調查，這樣的抽樣架構反映出階層資料結構的特色，也造成資料結構最底層的資料學生與各波次的調查是巢套於上層結構科系與學生內，這樣的巢套特性使得傳統的統計分析方法，例如迴歸分析與因素分析等建立在資料來自於獨立的假設被違反，而必須改採階層線性模式(hierarchical linear modeling, HLM)或是多層次模式(multilevel modeling, MLM)的統計方法(溫福星、邱皓政，2009)，使得教育科學的研究範疇逐漸往多層次模式發展，不管是橫斷面還是縱貫面的實徵研究，MLM的統計方法持續受到教育科學界的高度重視。

過去教育與心理研究的追蹤資料，大都以重複觀測(repeated measures)的變異數分析取向進行資料分析，但是這種重複觀測資料在傳統的變異數分析中有兩個限制：第一，必須是完整的或是平衡的資料結構設計(balanced design)，也就是每位受試者都要有相同的重複觀測個數(Kenny, Bolger, & Kashy, 2002, p. 1)；第二，雖然允許組內的解釋變項進行共變數分析，但無法納入組間解釋變項，如受試者的特徵或屬性(Heck & Thomas, 2009, p. 35)，去解釋重複觀測資料的變化趨勢。因此，晚近HLM或混合模式的成長模式(growth model)即克服了上述變異數分析的限制，可以將不相等個數或不同間距的重複觀測資料與受試者的特徵或屬性同時放在一個統計模式中進行分析(Wallace & Green, 2002)。不僅如此，更可以擴展到受試者所屬的環境(context)、組別(group)或實驗處理(treatment)，以檢視環境、組別或實驗處理的因素對受試者追蹤資料變化的影響(Singer & Willett, 2003)。由於教育與心理研究的重複觀測波次不多，不像財務與經濟的時間數列資料(time series data)經常成百上千，因此以重複觀測量數描述較為適當，但考慮大部分國內、外研究仍以追蹤資料稱之，因此本文重複觀測與追蹤資料是指相同的數據，同時這兩個名詞也交互於本文中。

國內有關於追蹤資料的資料庫不少，除了TEPS外，臺灣青少年成長歷程研究(Taiwan Youth Project, TYP)亦蒐集臺北市、新北市與宜蘭縣國中生國中3年的固定樣本追蹤資料，使得國內教育科學研究者得以有追蹤資料可以進行分析，探討這些樣本資料背後國中生母體的成長學習相關的歷程變化。此外，政府在1997年修訂《特殊教育法》後，接受特殊教育的身心障礙學生明顯增加，因此，研究小組針對國內身心障礙重要議題進行研究，蒐集各教育階層身心障礙學生與其家庭、就讀學校以及所居住縣市之相關資料，探討身心障礙學生受教情形與其學業成就表現的影響因素，從2007年起開始建構特殊教育長期追蹤資料庫(Special Needs Education Longitudinal Study, SNELS)。

在橫斷面的 MLM 中，很重要的兩個議題是組織層次截距項方程式是否有意義，以及組間迴歸係數與組內迴歸係數的問題。首先，如果個體層次的所有解釋變項沒有包含 0 的範圍，則個體層次方程式的截距項就沒有意義，連帶在組織層次的截距項方程式也就沒有解釋上的意義，這連帶影響到截距項方程式誤差項變異數的解釋。其次，MLM 如同變異數分析存在組內與組間資料結構，也因此是個體層次解釋變項對個體層次結果變項的影響，可以區分出組內迴歸係數與組間迴歸係數的差異，如果不進行對個體層次解釋變項適當的處理，則多層次分析所估計出的迴歸係數可能是一個偏誤的結果。鑑於上述這兩個理由，不少學者例如 Edwards 與 Lambert (2007)、Hofmann 與 Gavin (1998)、Kreft、de Leeuw 與 Aiken (1995)、Raudenbush 與 Bryk (2002) 都在探討 MLM 下個體層次解釋變項的平減問題，不同的平減方法 (centering)，例如組平均數平減、總平均數平減與不平減在解釋上的不同意義，以及這些不同平減方法在什麼條件下可以是等價關係。

相對在追蹤資料的 MLM 分析中，國內研究都顯少觸及這個平減議題，不是不平減就是採取總平減策略 (巫博瀚、陸偉明、賴英娟，2012；李靜芳、溫福星，2008；周玉慧，2011；楊惠卿，2010；蕭佳純，2011；蕭佳純、董旭英，2011；謝雨生、周玉慧，2012)。重複觀測資料或是追蹤資料與橫斷面資料最大的不同，在於存在同時具有橫斷面與縱貫面資料的特色，因此財務經濟學者稱為縱橫或是混合時橫資料 (panel date)，採用 MLM 分析重複觀測資料，如同橫斷面分析一樣，隨時間變動的 (time-varying) 解釋變項對結果變項的影響可以拆解為組內 (within person) 迴歸係數與組間 (between person) 迴歸係數。Allison (2005)、Bolger 與 Laurenceau (2013)、Kenny 等 (2002) 及 Twisk (2003) 都建議要對追蹤資料隨時間變動解釋變項進行適當的組平均數平減建議，以避免混淆組內與組間迴歸係數。

鑑於國內學者大都採用不平減與總平減策略來處理追蹤資料的隨時間解釋變項，由於在統計模式上這兩個方法是等價模式，所以這兩種方法估計的結果其斜率的意義可能存在偏誤，再加上上述國外學者鮮少介紹各種平減方法其背後的原理，因此本研究目的即在追蹤資料在 MLM 下，以統計模式探討有關隨時間變動解釋變項在各種平減方法下對斜率估計的原理及其所代表的涵義，並以實證資料演示范例說明各種平減策略，以及置回組平均數對組間與組內迴歸係數的影響，包含隨機截距模式、引進時間變項與隨機係數模式三種情況。

貳、統計模式

為了簡化模式的複雜度與介紹的清晰度，本研究僅以線性關係的 MLM 來示範說明。在這個 MLM 中，個體層次資料分別是受試者在各波次調查的依變項 Y 、解釋變項 X ，在組織層次是受試者的性別 Z ，在隨機效果設定中以隨機截距模型為說明重點。解釋變項 X 即為隨時間變動的解釋變項，它也可以包含時間變項，而 Z 則為非隨時間變動的解釋變項，其 MLM 方程式