中 華 心 理 學 刊
民93，46卷，4期，283-292

Chinese Journal of Psychology
2004, Vol. 46, N0. 4, 283-292

潛在變項組型混合模型之結構方程模型估計：隨機EM算則

鄭中平　翁儷禎

國立台灣大學心理學系

論文編號：03012；初稿收件：2003年5月15日；完成修正：2003年8月19日；正式接受：2003年8月30日
通訊作者：翁儷禎　台北市羅斯福路四段一號台灣大學心理系　（E-mail: ljweng@ntu.edu.tw）

實徵研究經常遭遇資料遺漏，本研究即在探討資料遺漏機制為潛在變項組型混合模型時，結構方程模型的最大概似估計，並利用模擬資料為範例比較其與常用遺漏值處理法的差異。潛在變項組型混合模型為組型混合模型的延伸，此模型假設觀察變項的遺漏組型反映潛在變項之類別，而非外顯類別，且各類別可有相異的結構方程模型。本研究以隨機EM算則估計當結構方程模型的資料遺漏機制符合此模式時之參數，並以模擬資料瞭解其與不同遺漏值處理法表現之差異。結果顯示，本研究建議的方法對因素負載量與潛在類別比率等參數之估計良好。

關鍵詞：非隨機遺漏、組型混合模型、結構方程模型、最大概似法、潛在變項

結構方程模型（structural equation models，簡稱SEM）爲社會科學常用分析方法之一，適用於多個測量變項間的相互關係之探討。結構方程模型包括研究者難以直接觀察到的潛在變項，此類潛在變項包括連續潛在變項以及晚近拓展之潛在類別變項。雖然潛在變項爲結構方程模型之重要特色（如Muthen, 2001a, b; Quintana & Maxwell, 1999），但其相關的遺漏值處理法中，卻鮮見將潛在變項對資料遺漏的影響納入考量。本研究即在探討當潛在類別影響資料遺漏組型時，結構方程模型的參數估計，此資料遺漏模型在文中稱之爲「潛在變項組型混合模型」（latent variable pattern mixture model）。遺漏資料可能影響樣本的代表性與研究推論的有效性，依據資料遺漏機制選擇適當的遺漏值處理法方能避免此等影響（King, King, Bachrach & McArdle, 2001）。

本文分爲四部份，首先說明問題緣起及相關估計方法，其次具體描述潛在變項組型混合模型並導出估計程序。第三部份以模擬資料爲例，瞭解相較於其他遺漏值處理法，本研究建議估計程序之表現；最後則總結本研究，並討論研究相關假設。以下將先介紹組型混合模型（pattern mixture model），再說明遺漏機制涵蓋潛在類別之可能性，以及處理具潛在類別與遺漏資料時，常用之參數估計算則。

組型混合模型。遺漏機制的討論中，Little與Rubin（1987）以及Little（1993, 1994, 1995）提出組型混合模型，考量外顯變項的類別與資料遺漏型態間的關係。當觀察變項的關係可能因各種遺漏組型不同而異時，此遺漏機制稱爲組型混合模型。亦即，遺漏組型即爲一外顯類別變項，不同遺漏組型表示不同類的受訪者組群，且觀察變項的分配在相異組群可能不相同。組型混合模型所以得名係因不同的遺漏組型下形成不同分配，整個資料的邊際分配爲有限混合模型（finite mixture model），因此稱爲「組型混合模型」（Little, 1993, 1995）。

Birmingham與Fitzmaurice（2002）將組型混合模型應用在學童心臟病危險因子縱貫研究的遺漏資料處理上。Birmingham與Fitzmaurice針對五個同齡群體（cohort）測量三次體重，每次測量間隔兩年，以分析不同年齡與性別學童是否有不同之肥胖症比率。根據三次測量資料的遺漏型態，可以將資料分

作者感謝兩位審查人，以及楊志堅教授的指正。

爲七種遺漏組型，包括三次皆到測的完整資料，以及到測兩次與到測一次之各三種遺漏組型。資料遺漏的直接原因通常爲父母未簽署學童受測同意書或學童於測量日缺席，然而Birmingham與Fitzmaurice認爲遺漏可能與過胖有關。例如，較胖學童可能因感困窘而刻意不到校，使得其資料遺漏率較高。依此推理，這些資料遺漏較多之受訪學童，肥胖率應高於資料未遺漏之學童，但卻因資料遺漏而無法在分析中直接顯示。因此，如未依遺漏組型區分資料，進行分析將可能會低估肥胖率。由於不同的遺漏組型可能有不同之肥胖率，應以遺漏組型分組，分別估計肥胖率後再行推估整體之肥胖率，即採組型混合模型處理資料遺漏。

在臨床心理學上，Niaura, Spring, Borrelli, Goldstein, Keuthen等人（2002）探討百憂解（Prozac）在戒煙上的療效，將受試分爲兩個不同劑量組與安慰劑組三組，各組同時接受爲期八週的認知行爲治療，並以第一、二、四、六、八週時受試者是否仍抽煙作爲依變項評估療效。在989名受試者中，僅有570名接受整個療程，Niaura等人考量三組受試中完成療程者比率未必相同，而完成療程與未完成療程者可能對百憂解有不同反應，乃建議將三組受試再分別細分爲「完成組」與「未完成組」，分別估計兩組效果後再估計實驗之療效。Niaura等人之作法即將受試者區分爲完成與未完成兩遺漏組型，兩類遺漏組型有其不同之觀察變項分配，這樣的資料遺漏機制即屬組型混合模型。

許多常見之遺漏值處理方法，例如列刪除法與對刪除法，需假設資料爲完全隨機遺漏（missing completely at random），不適用於組型混合模型下之結構方程模型。Hedeker與Gibbons（1997）將多樣本結構方程模型（multi-sample SEM）處理遺漏值之分析方式視爲組型混合模型遺漏情形下的結構方程模型（例如Lee,1986; Muthen, Kaplan & Hollis, 1987）。組型混合模型假設遺漏組型反應不同的受訪者類別，而不同類別可能有相異的結構方程模型，因此不同遺漏組型者之資料應分開處理。以多樣本結構方程模型處理遺漏值，乃將每個遺漏組型視爲一個群體進行分析，因此可處理組型混合模型遺漏情形下的結構方程模型，然以此方法處理組型混合模型時，組型個數不能太多。

遺漏機制中的潛在類別變項。組型混合模型乃針對外顯類別變項討論資料遺漏機制，但此等討論對心理學研究而言可能不夠充足，而需進一步引進潛在類別變項。我們可從兩個角度瞭解將潛在類別引入遺漏機制模型之必要性。

首先，可以從「理論構念」的角度引入潛在變項。心理學研究常因理論構念無法直接測量，而引進潛在變項以表達理論構念（如Everitt, 1984; Joreskog & Sorbom, 1993）。在遺漏機制的討論中，也可以考量引進潛在變項以表達理論構念。例如，在較長的政治議題問卷上，受訪者可區分爲「敏感受訪者」與「疲勞受訪者」兩類，其中「敏感受訪者」的特點是對敏感議題較易有拒答的傾向，而「疲勞受訪者」產生遺漏資料的原因則爲因題目增多產生疲勞以致未填答。兩類受訪者的資料遺漏組型可能恰巧相同，但其資料遺漏的原因卻相異。因此，兩類受訪者無需對應特定的資料遺漏組型，遺漏組型僅反應受訪者隸屬於某類型的可能性。如果一受訪者在敏感題目上遺漏的機率較高，其他題目遺漏的機率較低，則此受訪者爲「敏感受訪者」的機率較高。另一方面，若受訪者在各題的遺漏機率隨著題序增加而升高，則其可能爲「疲勞受訪者」。此情形中，受訪者類別爲潛在類別，反映研究者之理論概念，遺漏組型不用以界定潛在類別，而視作潛在類別的指標變項（indicators）。

其次，可以從測量誤差的角度引入潛在變項。當外顯變項有測量誤差時，可引進潛在變項，將外顯變項視爲潛在變項與測量誤差之和（例如Griliches, 1974），遺漏機制模型中的潛在變項同樣也可以從這個角度來討論。以前述「敏感受訪者」與「疲勞受訪者」爲例，兩者在問卷各題產生資料遺漏的機率不同，但仍可能產生相同的遺漏組型。例如一個疲勞受訪者填答問卷時，可能恰巧都漏答敏感題目，因此遺漏組型未必完全決定受訪者之類別，而可能有誤差。將受訪者類別視爲潛在分類變項，遺漏組型作爲潛在分類的指標變項，則可以將誤差考慮在內。

引進潛在類別變項對遺漏資料特性的影響，亦可以前述之Niaura等人（2002）戒煙研究爲例，說明引進潛在類別變項的理論意義及對誤差之考量。Niaura等人將受訪者區分爲「完成者」與「未完成者」兩類分別估計療效，隱含兩者之療效不同。此假設可有其理論基礎，例如研究者可以假設完成者與未完成者在「毅力」人格特質上相異，毅力低者戒煙較不容易，毅力高者戒煙則較容易，兩者對戒煙治療之敏感性不同，因此需個別進行分析。研究者也可能假設完成者與未完成者在其他特質上有差異，因而影響療效。但無論研究者認爲完成者或未完成者有何特質，特質常作爲一理論概念，直接將理論概念（「毅力」）與外顯變項（「遺漏組型」）連結，忽略了兩者之差異。未完成原因很多，未必全部反應毅力特質，可能有受訪者因不可抗力因素而未繼續接受療程，但因其毅力高而有較佳療效，將其歸於毅力低的未完成組，將使療效的估計產生偏誤。引進潛在類別則能考量前述問題，研究者可以假設有「毅力高」與「毅力低」兩類受試者，唯此類別爲潛在類別變項，遺漏組型爲潛在類別之指標變項。由於以遺漏組型作爲潛在類別之指標變項，因此也需要描述潛在類別變項與資料遺漏間關係的

遺漏機制模型。

包含潛在類別的模型常被稱爲有限混合模型，考量潛在類別對資料遺漏影響之模型亦爲有限混合模型。有限混合模型與具遺漏資料之模型常以EM算則（Expectation-Maximization algorithm）或其變形估計模型參數，本研究亦採取類似作法，估計當潛在類別影響資料遺漏時結構方程模型之參數，底下即簡介EM及相關算則在遺漏資料與有限混合模型上之使用。

EM算則。有鑑於完整資料相關估計方法之長足發展，處理具遺漏資料時，如果能將遺漏資料填補爲完整資料，常能增進資料處理上的簡易性。EM算則即善用完整資料與遺漏資料之關係以簡化運算，用來處理資料遺漏時的模型參數估計（Jamshidian & Bentler, 1999）。

EM算則是一個疊代（iterative）程序，每次疊代分成E步驟（expectation step）與M步驟（maximization step）。E步驟推導給定觀察外顯變項下遺漏變項之條件機率密度函數，再計算目前參數估計值下對數概似函數之期望值。M步驟則根據E步驟所得資料求參數之最大概似估計值，相當於完整資料之最大概似估計（Dempster, Laird & Rubin, 1977; McLachlan & Krishnan, 1997）。類似的概念早期即已出現，唯Dempster等人將其形式化命名爲EM算則，並證明其收斂性質。Dempster等人討論之遺漏資料並不限於觀察值，亦包括潛在變項或參數，因此能運用在許多不同問題上，遂使EM算則成爲廣爲使用的估計方法之一（Meng & van Dyk, 1997）。

有限混合模型之參數估計亦常使用EM算則，此乃因將各觀察值隸屬於潛在類別的訊息視爲遺漏資料。我們引進潛在類別的隸屬函數變項向量 C_i，用以標示第i個觀察值隸屬之潛在類別。令 C_i 爲 $n_i \times 1$ 的向量，其元素 $C_{i,g}$ 表示資料隸屬於潛在類別g之機率，若該筆資料隸屬於潛在類別 g，$C_{i,g}=1$，對其他潛在類別 $t(t \neq g)$，則。可視爲假設遺漏資料（hypothetic missing data），而適用EM算則，如同Dempster等人（1977）處理常態有限混合模型，McCutcheon（1987）處理潛在類別模型，Bartholomew與Knott（1999）處理混合潛在類別模型（mixed latent class model）等。

雖然EM算則在計算上較簡易，但採用EM算則分析有限混合模型問題仍有一些缺點。首先，當起始值不佳時，可能會掉入局部極值。其次，潛在類別組數需先設定，此對探索性較強的議題可能相當困難。再者，資料中潛在類別間區分不佳時（poorly separated），亦會影響EM算則的表現（Celeux & Diebolt, 1985; Diebolt & Celeux, 1993）。針對前述問題，隨機EM算則（stochastic EM algorithm）爲可能解答之一。

隨機EM算則。隨機EM算則是EM算則的變形，爲處理遺漏資料常用方法之一。隨機EM算則與EM算則差異僅在以S步驟取代EM算則之E步驟，即在求出條件機率密度函數後，據以產生模擬資料插補遺漏值，形成完整資料以簡化運算，而非如E步驟求取對數概似函數之期望值。Celeux與Diebolt（1985）利用隨機EM算則處理有限混合模型之問題，以隸屬函數變項爲需增補（augmented）資料，利用貝氏定理求出目前觀察變項下，隸屬函數變項之多項式分配，據以產生模擬之隸屬函數變項，將每一觀察值分配在各潛在類別，有限混合模型問題即可轉化爲多個樣本之統計分析問題，估計上較EM算則容易。除了計算上較簡單外，隨機EM算則引入隨機因素，可以避免當起始值不佳時掉入局部極值問題，同時無須設定正確之潛在類別數，而僅需設定類別數上限即可。部份研究也顯示，潛在類別區分不佳時，隨機EM算則表現優於EM算則（Celeux & Diebolt, 1985; Diebolt & Celeux, 1993）。如果有多重解，隨機EM算則亦可呈現出解答的可能區域（例如，Diebolt & Ip, 1994）。

隨機EM算則爲馬可夫鏈蒙地卡羅法（Markov Chain Monte Carlo）的一種，亦爲MCEM算則（Monte Carlo EM algorithm）的特例（Diebolt & Celeux, 1993; McLachlan & Krishnan, 1997; Nielsen, 2000）。一般常以隨機EM算則求取參數的後驗分配（例如，Diebolt & Ip, 1996），在特定情形下，隨著隨機EM算則的疊代，估計參數會在待估計參數空間形成遍歷的（ergodic）馬可夫鏈，並收斂至此馬可夫鏈之靜態分配（stationary distribution）。此靜態分配平均值接近最大概似估計值，而變異數反應遺漏資料產生的訊息流失（information loss），由隨機EM算則得出之參數估計值，可以視爲此靜態分配的一個樣本（Diebolt & Ip, 1996; Van Zwet, 2001）。用在點估計時，隨機EM算則與MCEM算則類似，當估計值隨疊代次數增加而在定值附近隨機變動時，即表示估計值收斂（Tanner, 1996; Wei & Tanner, 1990），此時可再疊代若干次，取估計數值平均作爲最終之估計值，以增加估計精確度（例如，Bockenholt & Tsai, 2001; Diebolt & Ip, 1996; Nielsen, 2000; 鄭中平與翁儷禎，2002）。

本研究之目的即在漏機制中考量潛在類別變項，以表達潛在類別對資料遺漏組型的影響，稱之爲「潛在變項組型混合模型」，並以隨機EM算則求取此遺漏機制之結構方程模型參數的最大概似估計值，最後以模擬資料作爲範例，初步評估本研究建議方法之表現。

潛在變項組型混合模型

本節討論考慮潛在類別對資料遺漏影響之潛在變項組型混合模型，先描述此模型之形式，再討論

該模型遺漏機制之特性以及估計方法。

模型描述。遺漏機制的討論常引入遺漏指標變項，以指示外顯變項是否遺漏。假設有p個外顯變項，每一外顯變項對應一遺漏指標變項。令M為p×1的遺漏指標變項向量，M之變項皆為二元變項，值0表示資料遺漏，值1表示資料完整，M之數值即反應了外顯變項的遺漏組型。

$$M_j = \begin{cases} 0 & \text{若 } Y_j \text{ 遺漏} \\ 1 & \text{若 } Y_j \text{ 未遺漏} \end{cases} \qquad j = 1, 2, \ldots\ldots, p \tag{1}$$

潛在變項組型混合模型假設遺漏組型反映的是受訪者隸屬各潛在類別的機率而非外顯類別，不同潛在類別可有個別的結構方程模型，而外顯變項與遺漏指標變項皆反映潛在類別。因此，外顯連續變項與潛在類別變項形成有限混合結構方程模型（finite mixtures SEM），而遺漏指標變項M與潛在類別變項則形成潛在類別模型（潛在類別分析，latent class analysis）。

有限混合結構方程模型有許多種描述法（例如，Jedidi, Jagpal & Desarbo, 1997a,b），本研究以Dolan與van der Maas（1998）之模型描述：

$$\mathrm{Y} \mid g = \mathrm{v}_g + \Lambda_g \eta_g + \varepsilon_g \tag{2}$$

$$\eta_g = \alpha_g + \mathrm{B}_g \eta_g + \zeta_g \tag{3}$$

式2表示潛在類別為g時外顯連續變項與潛在變項的關係，其中$\mathrm{Y} \mid g$為p×1的外顯變項向量，v_g為p×1的截距向量，η_g為k×1的潛在變項向量，Λ_g為p×k的係數矩陣，ε_g為p×1的殘差向量。式3表示潛在類別為g時潛在變項間的關係，其中α_g為k×1的截距向量，B_g為k×k的係數矩陣，ζ_g為k×1的殘差向量。

假設在給定潛在類別下，外顯變項的條件分配為多元常態分配，則外顯連續變項分配為多元常態有限混合分配。潛在變項組型混合模型的遺漏機制在形式上相當於潛在類別模型，遺漏指標變項間的關係為潛在類別所造成，若給定潛在類別，則遺漏指標變項間彼此獨立，稱為局部獨立（local independence, McCutcheon, 1987）。在此假設下，當潛在類別為g時，M組型為m的機率為

$$p(M=m \mid g) = \prod_{j=1}^{p} p(M_j = m_j \mid g) = \prod_{j=1}^{p} \lambda_{jg}^{mj} (1-\lambda_{jg})^{1-mj} \tag{4}$$

其中與分別表示在給定潛在類別為第g類下，M組型恰為m與第j個二元變項恰為的機率，為p×1的向量，表示潛在類別為第g類下各題未遺漏機率，其元素為第j題數值為1（未遺漏）的機率。各潛在類別下都可能觀察到M為m，因此當有n_c個潛在類別時，M組型為m的機率為

$$p(M=m) = \sum_{t=1}^{n_c} \pi_t{}^* p(M=m \mid t) = \sum_{t=1}^{n_c} \pi_t{}^* \prod_{j=1}^{p} \lambda_{jg}^{mj} (1-\lambda_{jt})^{1-mj} \tag{5}$$

其中π_t為第t個潛在類別佔母體比率。假設潛在類別數已知（恰為），待估計參數包括與，前者為p×n_c的機率矩陣，表示每個潛在類別在每一題的遺漏機率，而π則為n_c×1維向量，表示各潛在類別比率，其元素和為1。

潛在變項組型混合模型下之結構方程模型可以式2、3與5描述。如果僅考慮外顯連續變項（Y），包括式2與3，則模型為有限混合結構方程模型。若只考慮外顯二元變項（遺漏指標變項M，式5），則為潛在類別模型。

遺漏機制的性質。潛在變項組型混合模型包含潛在類別，為有限混合模型之一，在有限混合模型中，若各潛在類別內之遺漏情形符合隨機遺漏，稱為潛在可忽略遺漏（latent ignorability），然此情形並不符合隨機遺漏，不宜僅就觀察資料進行估計，而應納入遺漏機制分析（Frangakis & Rubin, 1999; Hills, 2001）。潛在變項組型混合模型中，研究者興趣所在之外顯連續變項Y為有限混合結構方程模型，且各潛在類別之遺漏情形由於局部獨立假設乃為完全隨機遺漏，屬於潛在可忽略遺漏，因此應將遺漏機制納入分析，即分析$W=(Y', M')'$，而非僅分析Y。W符合有限混合模型，分配為

$$f(W) = \sum_{t=1}^{n_c} \pi_t{}^* f_{1,t}(Y; \mu, \Sigma)^* f_{2,t}(M; \lambda) \tag{6}$$

其中$f_{1,t}(\cdot)$表示潛在類別為t時之多元常態分配機率密度函數，$f_{2,t}(\cdot)$是多元白努利分配機率函數。待估計參數可以分成潛在類別比率、結構方程模型參數與遺漏機制參數三部分，潛在類別比率為$\pi = (\pi_1, \pi_2, \ldots, \pi_{n_c})$，結構方程模型參數包括$\nu = (\nu_1, \nu_2, \ldots, \nu_{n_c})$、$\Lambda = (\Lambda_1, \Lambda_2, \ldots, \Lambda_{n_c})$、$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{n_c})$與$\mathrm{B} = (\mathrm{B}_1, \mathrm{B}_2, \ldots, \mathrm{B}_{n_c})$，以及$\zeta$與$\varepsilon$的共變數矩陣$\Psi = (\Psi_1, \Psi_2, \ldots, \Psi_{n_c})$與$\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_{n_c})$，遺漏機制參數則為$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{n_c})$。

估計。本研究採隨機EM算則推導當遺漏機制為潛在變項組型混合模型時結構方程模型之最大概似估計，以LVPM-ML（latent variable pattern mixture model maximum likelihood estimation）稱之，並利用隨機EM算則，視隸屬函數變項及遺漏變項為待增補之遺漏資料。對於第i筆資料，假設其未遺漏變項、遺漏變項、遺漏指標變項與隸屬函數變項分別為$Y_{i,obs}$、$Y_{i,miss}$、M_i與C_i，在給定觀察值，即$Y_{i,obs} = y_{i,obs}$與$M_i = m_i$下，則C_i之第g個元素$C_{i,g}$與$Y_{i,miss}$之分佈如下（推導步驟請見附錄一）

$$\begin{aligned} & f(C_{i,g}, Y_{i,miss} \mid Y_{i,obs} = y_{i,obs}, M_i = m_i) \\ & = f(C_{i,g}, Y_{i,obs} = y_{i,obs}, M_i = m_i)\, f_g(Y_{i,miss} = y_{i,obs}) \end{aligned} \tag{7}$$

若已知隸屬函數變項C_i之條件機率分配，可據此機率分配隨機決定該筆資料隸屬之類別。進一步給定隸屬函數類別g後，遺漏變項則呈多元常態分

配 $N(\hat{\mu}^{g}_{mis|obs}, \hat{\Sigma}^{g}_{mis|obs})$，可依此多元常態機率密度函數產生隨機向量，填補成為完整資料。由於隸屬函數變項已知，並已填補遺漏變項，可以式8估計潛在類別比率，結構方程模型參數以多樣本結構方程模型分析求得（例如，Joreskog & Sorbom, 1993），遺漏機制參數部份則以式9求得。

$$\pi_g = \frac{\sum_{i=1}^{N} C_{i,g}}{N} \tag{8}$$

$$\lambda_{jg} = \frac{\sum_{i=1}^{N} C_{i,g} M_{ij}}{\sum_{i=1}^{N} C_{i,g}} \tag{9}$$

綜言之，本研究建議之隨機EM算則，包含如下步驟：

1. 將資料隨機分組，決定隸屬函數變項，以式8、式9與結構方程模型分析求取參數作為初始值。
2. 執行S步驟，在目前參數估計值下求取隸屬函數變項與遺漏變項之條件分佈，填補隸屬函數變項與遺漏變項而形成多樣本之完整資料。
3. 執行M步驟，針對步驟2所得之各組完整資料，以式8、式9與多樣本結構方程模型分析求得潛在類別比率、遺漏機制參數與結構方程模型參數估計值。
4. 重複步驟2與3若干次，並以估計參數對疊代次數作圖，當參數估計趨向於定值附近跳動時，可視為收斂，此時再行疊代，以最後數次估計值之平均作為隨機EM算則估計值。

範例

本研究產生一組模擬資料，以初步比較數種遺漏值處理法於潛在變項組型混合模型下之表現。此範例之目的乃在探討如果資料遺漏機制為潛在變項組型混合模型時，本研究推導之估計法的表現是否較其他遺漏值處理法為佳。探討的遺漏值處理法包括列刪除法、平均值插補法、迴歸插補法與本研究建議之LVPM-ML。由於本研究焦點在遺漏資料的處理，故乃假設外顯連續變項的結構方程模型設定正確，即為有限混合結構方程模型。進行有限混合結構方程模型分析時，需以各個觀察值的個別資料進行分析，因此，無法產生完整個別資料之遺漏值處理法皆不適用，此類方法包括對刪除法與全訊息最大概似估計（full information ML, Arbuckle, 1996; Enders, 2001），多樣本結構方程模型遺漏值處理法則因本例遺漏組型過多而不適用。

本範例產生完整資料時，設定樣本數人為1000，並假定有「疲勞受訪者」與「敏感受訪者」兩個潛在類別，各佔全樣本之66.7%與33.3%。兩類受訪者的結構方程模型皆為四變項之斜交二因素模型，各因素之平均數為0，標準差為1，因素間相關則為 .5，前二變項在第一個因素上有非零負載量，後二變項則在第二因素上有非零負載量，其餘負載量為0。非零之因素負載量與變項平均數則依潛在類別不同而異，「疲勞受訪者」分別設定為 .8與0，「敏感受訪者」則為 .6與2。在模擬遺漏資料上，為使遺漏機制符合潛在變項組型混合模型，乃假設疲勞類別受訪者各題遺漏機率隨題序而增加，分別為 .05、.15、.25與 .35，敏感類別受訪者對前二題較敏感，因此有較高遺漏機率（.35），對後二題遺漏機率則較低（.05），產生模擬資料之參數數值整理於表一模型理論值一欄。

產生遺漏資料後，即以不同遺漏值處理法分析，並估計結構方程模型之參數，同時亦以兩種方式分析完整資料。分析完整資料時，首先假設受訪者所屬類別已知，進行多樣本結構方程模型分析，其次假設所屬類別未知，僅知存在二個潛在類別，以有限混合結構方程模型分析之。第一種分析結果旨在瞭解隨機抽樣過程造成之誤差，第二種分析結果之估計值，則除受抽樣誤差影響外，亦考慮未知受訪者類別造成之變動。本研究著眼之情境中，受訪者類別為潛在變項，因此以第二種分析結果作為考量遺漏值處理法優劣參考。列刪除法、平均值插補法與迴歸插補法則先處理遺漏值，再以最大概似法進行有限混合結構方程模型分析。前述之多樣本結構方程模型與遺漏值處理法之有限混合結構方程模型分析，皆以Mplus 2.02（Muthen & Muthen, 1998）進行[1]。LVPM-ML則同時處理遺漏值並進行有限混合結構方程模型分析，由研究者以SAS / IML程式撰寫。本範例先疊代800次，隨著疊代次數的增加，參數估計值趨於穩定，為使所得之參數估計值較穩定，乃再疊代200次，共計疊代1000次，以最後200次參數估計值之平均作為LVPM-ML的估計值。

各種遺漏值處理法中除列刪除法採410筆完全未遺漏資料分析外，其餘各方法皆處理1000筆資料，完整資料與各種遺漏值處理法之參數估計值列於表一，並以各類參數估計值與母群參數之RMSD（root mean squared difference）為各方法參數估計之評估標準。完整資料的分析中，先假設受訪者所屬類別已知，將之區分為疲勞類與敏感類，進行多樣本結構方程模型分析，其在因素負載量、因素間相關、變項平均值與類別比率估計值的RMSD分別為 .142、.121、.044與 .003，此部份反應抽樣誤差與樣本人數對參數估計之影響。本範例雖然整體樣本數達1000，但因分為兩群體分析，且敏感類受試者人數較少（336），致該組之參數估計表現較差，因此綜合兩樣本估計值之RMSD亦較大。若受訪者

所屬類別未知，有限混合結構方程模型估計值之RMSD分別爲 .134、.130、.048與 .011，與多樣本結構方程模型分析類似，未知受訪者潛在類別造成之參數估計變動不大，此可能由於本範例中兩潛在類別之變項平均數差異相當大，如果潛在類別之分配差異較小，則未知受訪者潛在類別造成之參數估計變動可能變高。

四種遺漏值處理法中，因素負載量估計以本研究建議之LVPM-ML最佳，RMSD爲 .088，優於列刪除法與迴歸插補法（RMSD爲 .227與 .244），平均值插補法最差（.523）。因素間相關以LVPM-ML與列刪除法估計最佳（RMSD爲 .082與 .105），迴歸插補法與平均值插補法最差（.173與 .207）。變項的平均數以列刪除法與LVPM-ML估計最佳（.088與 .134），迴歸插補法次之（.323），仍以平均值插補法最差（.406）。潛在類別比率估計值以列刪除法與LVPM-ML估計較準（RMSD爲 .013及 .036），其他方法表現不佳。四種遺漏值處理法中，LVPM-ML亦能估計不同潛在類別各變項之遺漏比率，由表一可看出遺漏組型如假設模式，疲勞類別之遺漏比率依題序增加而升高，敏感類別則在前二題有較高遺漏比率，LVPM-ML對各題遺漏機率估計之RMSD爲 .019。

綜合看來，本範例中，LVPM-ML表現最佳，除變項平均數外，對其餘參數的估計都優於其他方法，對因素負載量與因素間相關的估計甚至較完整資料之有限混合結構方程模型爲佳，同時亦提供各類別資料遺漏之訊息。列刪除法表現次之，特別對

表一

潛在變項組型混合模型遺漏機制之不同遺漏值處理法參數估計值

潛在類別	參數	模型理論值	完整資料 組別已知	完整資料 組別未知	列刪除法	平均值插補法	迴歸插補法	LVPM-ML
疲勞類	LY 1 1	.8	.762	.849	.888	.795	.914	.868
	LY 2 1	.8	.860	.725	.633	.890	1.093	.726
	LY 3 2	.8	.763	.725	.744	.905	.845	.745
	LY 4 2	.8	.805	.776	.765	.889	1.047	.863
	PH 2 1	.5	.608	.617	.553	.709	.743	.564
	ME(Y1)	.0	.032	.037	-.044	.290	.263	.015
	ME(Y2)	.0	.020	-.010	-.022	.428	.371	.024
	ME(Y3)	.0	.051	.028	.000	.676	.339	.117
	ME(Y4)	.0	.088	.069	-.017	.718	.416	.124
	P(M1=0)	.05						.018
	P(M2=0)	.15						.165
	P(M3=0)	.25						.244
	P(M4=0)	.35						.367
	π	.667	.664	.656	.654	.907	.805	.631
敏感類	LY 1 1	.6	.916	.910	1.036	.890	.639	.731
	LY 2 1	.6	.410	.496	.367	-.803	.473	.578
	LY 3 2	.6	.586	.629	.595	.494	.207	.680
	LY 4 2	.6	.736	.746	.958	.917	.977	.746
	PH 2 1	.5	.367	.358	.362	.295	.467	.403
	ME(Y1)	2.0	1.963	1.908	1.784	1.984	1.921	1.693
	ME(Y2)	2.0	2.037	2.046	1.941	2.030	1.869	1.866
	ME(Y3)	2.0	1.996	1.992	1.904	2.032	2.464	1.958
	ME(Y4)	2.0	2.036	2.026	2.014	2.276	2.319	1.972
	P(M1=0)	.35						.376
	P(M2=0)	.35						.343
	P(M3=0)	.05						.031
	P(M4=0)	.05						.040
	π	.333	.336	.344	.346	.093	.195	.369

註：參數LY 1 1至LY 4 2分別表示四個題目之非零因素負載量，PH 2 1爲因素間相關，ME(Y1)至ME(Y4)表示四個題目之平均值，P(M1=0)至P(M4=0)表示四個題目遺漏比率，π表示潛在類別佔全體比率，模型理論值爲資料產生時之參數。參數估計時設定因素平均數爲0，變異數爲1。

因素間相關的估計較準確，平均值插補法與迴歸插補法的表現都相當差。

討論

結構方程模型分析需要實徵資料以驗證研究者假設的理論模型，而資料發生遺漏是其收集過程經常會遭遇到的情形。Little（1993）提出的選擇模型（selection model）與組型混合模型對心理學研究而言有其不足之處，因為此兩類模型均僅著眼於可觀察變項與資料遺漏間的關係，資料遺漏也可能與潛在變項有關。Muthen等人（1987）將潛在連續變項加入選擇模型，鄭中平與翁儷禎（2002）則推導此時的最大概似估計。本研究則將潛在類別變項加入組型混合模型，稱之為潛在變項組型混合模型。潛在變項組型混合模型假設觀察變項的遺漏組型並非受訪者的分類，而是潛在類別的指標變項，此模型為Little組型混合模型擴充至潛在變項層次之延伸。簡言之，潛在變項組型混合模型假設觀察變項的遺漏組型反映潛在變項之類別，且各類別可有其結構方程模型。

本研究建議以隨機EM算則估計資料遺漏機制為潛在變項組型混合模型時結構方程模型的參數，並以模擬資料為例，比較數種遺漏值處理法的表現。結果發現LVPM-ML最能回復資料產生歷程的因素負載量，對於不同潛在類別之遺漏機率與潛在類別比率估計亦表現良好，唯因素間相關估計略差。列刪除法對因素結構的回復表現次於LVPM-ML，但優於平均值插補法與迴歸插補法。

本研究建議之LVPM-ML表現優良，相較於完整資料，LVPM-ML估計時除利用連續資料訊息外，遺漏指標亦作為潛在類別指標變項，可能因此使受訪者所屬潛在類別之估計較準確，進而影響各類別因素負載量之估計。唯本範例假設之遺漏比率較高，前述結論是否能運用至一般情形，宜以設計完整之模擬研究探究。

LVPM-ML假設資料遺漏組型受潛在類別影響，如違反任一假設都可能使估計程序表現不如預期理想，未來如能發展檢定假設是否成立的方法，或瞭解假設未成立時估計方法之強韌性，均對LVPM-ML的實際運用有所助益。雖則初步研究發現推導之隨機EM算則表現良好，唯此研究為一初探性研究，有關該模式假設與適用性之相關議題尚待後續研究進一步探討瞭解。

註釋

1：常用之結構方程模型軟體中，LISREL 8.5、EQS 5.7b與AMOS 5.0皆未提供有限混合結構方程模型分析。

參考書目

鄭中平與翁儷禎（2002）。潛在變項選擇模型結構方程模型之最大概似估計。「調查研究」，*12*，*5-27*。

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243-277). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bartholomew, D. J. & Knott, M. (1999). *Latent variable models and factor analysis.* London: Arnold.

Birmingham, J. & Fitzmaurice, G. M. (2002). A pattern-mixture model for longitudinal binary responses with nonignorable nonresponse. *Biometrics, 58,* 989-996.

Bockenholt, U., & Tsai, R. C. (2001). Individual differences in paired comparison data. *British Journal of Mathematical and Statistical Psychology, 54,* 265-277.

Celeux, G., & Diebolt, J. (1985). The SEM algorithm: A probabilistic teacher algorithm derived from the EM algorithm for mixture problem. *Computational Statistics Quarterly, 2,* 73-82.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39,* 1-38.

Diebolt, J., & Celeux, G. (1993). Asymptotic properties of a stochastic EM algorithm for estimating mixing proportions. *Communications in Statistics, Series B: Stochastic Models, 9,* 599-613.

Diebolt, J., & Ip, E. H. S. (1996). Stochastic EM: method and application. In W. R. Gilks, S. Richardson & D. J. Spiegelhalter (Eds.), *Markov chain Monte Carlo in practice* (pp. 259-273). London: Chapman & Hall.

Dolan, V. C., & van der Maas, H. L. J. (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika, 63,* 227-253.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data, *Structural Equation Modeling, 8*, 128-141.

Everitt, B. S. (1984). *An introduction to latent variable models.* New York: Chapman and Hall.

Finkbeiner, C. (1979). Estimation for the multiple factor model when data are missing. *Psychometrika, 44,* 409-420.

Frangakis, C. E. & Rubin, D. B. (1999). Addressing complications of intent-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika, 86,* 365-

379.

Griliches, Z. (1974). Errors in variables and other unobservables. *Econometrika, 42,* 971-998. (Reprinted in Aigner, D. J. , & Goldberger, A. S. (Eds), (1977). *Latent variables in social-economic models.* Amsterdam: North-Holland.)

Hedeker, D., & Gibbons, R. D. (1997). Application of random-effect pattern-mixture models for missing data in longitudinal studies, *Psychological Methods, 2,* 64-78.

Hill, J. L. (2001). Accommodating missing data in mixture models for classification by opinion-changing behavior. *Journal of Educational and Behavioral Statistics, 26,* 233-268.

Jamshidian, M. & Bentler, P. M. (1999). ML estimation of mean and covariance structures with missing data using complete data routines. *Journal of Educational and Behavioral Statistics, 24,* 21-41.

Jedidi, K., Jagpal, H. S., & Desarbo, W. S. (1997a). STEMM:A general finite mixture structural equation model. *Journal of Classification, 14*, 23-50.

Jedidi, K., Jagpal, H. S., & Desarbo, W. S. (1997b). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science, 16*, 39-59.

Johnson, N. L., & Kotz, S. (1972) *Distributions in statistics: continuous multivariate distributions.* New York: John Wiley & Sons.

Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS command language.* Mooresville, IN: Scientific Software, Inc.

King, D. W., King, L. A., Bachrach, P. S., & McArdle, J. J. (2001). Contemporary approaches to missing dada: The Glass is really half full. *PTSD Research Quarterly*, 12, 1-8.

Lee, S. Y. (1986). Estimation for structural equation models with missing data. *Psychometrika, 51*, 93-99.

Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association, 88*, 125-134.

Little, R. J. A. (1994). A class of pattern-mixture models for normal incomplete data. *Biometrika, 81*, 471-483.

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association, 90*, 1112-1121.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: John Wiley & Sons.

McCutcheon, A. L. (1987). *Latent class analysis.* Sage University Paper series on Quantitative Applications in the social Science, series no.07-064. Newbury Park, CA: Sage.

McLachlan, G. J., & Krishnan, T. (1997). *The EM algorithm and extensions.* New York: John Wiley & Sons.

Meng, X. L. & van Dyk, D. (1997). The EM algorithm – an old folk-song sung to a fast new tune. *Journal of the Royal Statistical Society, Series B, 59*, 511-567.

Muthen, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52* , 431-462.

Muthen, L. K. & Muthen, B. (1998). *Mplus user's guide.* Los Angeles: Muthen & Muthen.

Muthen, B. (2001a). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling.* (pp. 1-33). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Muthen, B. (2001b). Second-generation structural equation modeling with a combination of categorical and continuous latent variables. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change.* (pp. 291-322). Washington, DC: American Psychological Association.

Niaura, R., Spring, B., Borrelli, B., Hedeker, D., Goldstein, M. G., Keuthen, N., DePue, J., Kristeller, J., Ockene, J., Prochazka, A., Chiles, J. A., & Abrams, D. B. (2002). Multicenter trial of fluoxetine as an adjunct to behavioral smoking cessation treatment. *Journal of Consulting and Clinical Psychology, 70,* 887-896.

Nielsen, S. F. (2000). The stochastic EM algorithm: Estimation and asymptotic results. *Bernoulli, 6*, 457-489.

Quintana, S. M. & Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *Counseling Psychologist, 27,* 485-527.

Tanner, M. A. (1996). *Tools for statistical inference.* NY: Springer.

Van Zwet, E. (2001). Perfect stochastic EM. In M. de Gunst, C. Klaassen & A. van der Vaart (Eds.), *State of the art in probability and statistics* (pp. 607-616). Ohio: Institute of Mathematical Statistics.

Wei, G. C. G., & Tanner, M. A. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithm. *Journal of the American Statistical Association, 85*, 699-704.

附錄 潛在變項組型混合模型之遺漏變項條件機率密度函數

令 Y_i 爲第 i 筆資料的外顯變項向量，維度爲 p×1，其中 $Y_{i,miss}$ 與 $Y_{i,obs}$ 分別爲 Y_i 中遺漏變項向量及未遺漏變項向量。令 M 爲遺漏指標變項向量，現欲在給定隨機EM算則M步驟估計之第 g 個潛在類別 Y 平均數向量 $\hat{\mu}^g$ 和共變數矩陣 $\hat{\Sigma}^g$ 下，求得固定 $Y_{i,obs}=y_{i,obs}$ 與 $M_i=m_i$ 時，隸屬變項 $C_{i,g}$ 與 $Y_{i,miss}$ 之條件機率密度函數。

爲方便表達，乃引進選取變項並重排的運算 w_{uv}，w_{uv} 爲 a×p 矩陣，a 爲選取變項數（Finkbeiner, 1979）。每筆資料的連續變項可分成二類：遺漏變項與未遺漏變項，w_{uv} 之下標 u 與 v 依序標示是否選取遺漏變項與未遺漏變項，0 表示不選取該類變項，1 表示選取該類變項；例如，w_{01} 表示不選取遺漏變項，但選取未遺漏變項。以 Y_i'= [2 . 5 . . 3]爲例（.表示遺漏值），$(w_{01}Y_i)'$ = [2 5 3]，對應之 w_{01} 如下所示。

$$w_{01}=\begin{bmatrix}1&0&0&0&0&0\\0&0&1&0&0&0\\0&0&0&0&0&1\end{bmatrix}$$

同樣的，w_{11} 則相當於先選取遺漏變項，再選取未遺漏變項，$(w_{11}Y_i)'$= [. . . 2 5 3]，對應之 w_{11} 如下。

$$w_{01}=\begin{bmatrix}0&1&0&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\1&0&0&0&0&0\\0&0&1&0&0&0\\0&0&0&0&0&1\end{bmatrix}$$

在給定潛在類別 g 下，Y_i 的機率密度函數爲

$$f_g(Y_{i,miss}, Y_{i,obs}) \sim N(\hat{\mu}^g_{miss,obs}, \hat{\Sigma}^g_{miss,obs}) \qquad \text{(A1)}$$

其中 $\hat{\mu}^g_{miss,obs} = w_{11}\hat{\mu}^g$，$\hat{\Sigma}^g_{miss,obs} = w_{11}\hat{\Sigma}^g w'_{11}$。可觀察之未遺漏變項亦符合多元常態分配：

$$f_g(Y_{i,obs}) \sim N(\hat{\mu}^g_{obs}, \hat{\Sigma}^g_{obs}) \qquad \text{(A2)}$$

其中 $\hat{\mu}^g_{obs} = w_{01}\hat{\mu}^g$，$\hat{\Sigma}^g_{obs} = w_{01}\hat{\Sigma}^g w'_{01}$。遺漏變項 $Y_{i,miss}$ 在給定 $Y_{i,obs}$ 時，條件機率密度函數仍爲多元常態分配（Johnson & Kotz, 1972）：

$$f_g(Y_{i,miss}|Y_{i,obs}) \sim N(\hat{\mu}^g_{miss|obs}, \hat{\Sigma}^g_{miss|obs}) \qquad \text{(A3)}$$

其中

$$\hat{\mu}^g_{miss|obs} = \hat{\Sigma}^g_{miss,obs}\hat{\Sigma}^{g^{-1}}_{obs}(Y_{i,obs} - \hat{\mu}^g_{obs}) + \hat{\mu}^g_{miss}$$

$$\hat{\Sigma}^g_{miss|obs} = \hat{\Sigma}^g_{miss} - \hat{\Sigma}^g_{miss,obs}\hat{\Sigma}^{g^{-1}}_{obs}\hat{\Sigma}^g_{obs,miss}$$

$$\hat{\Sigma}^g_{miss} = w_{10}\hat{\Sigma}^g w'_{10}，\ \hat{\mu}^g_{miss} = w_{10}\hat{\mu}^g。$$

隸屬變項 $C_{i,g}$ 與 $Y_{i,miss}$ 之條件機率密度函數則爲

$$\begin{aligned}&f(C_{i,g}, Y_{i,miss}|Y_{i,obs}=y_{i,obs}, M_i=m_i)\\&=f(C_{i,g}|Y_{i,obs}=y_{i,obs}, M_i=m_i)\,f(Y_{i,miss}|C_{i,g}=1, Y_{i,obs}, M_i=m_i)\\&=f(C_{i,g}|Y_{i,obs}=y_{i,obs}, M_i=m_i)\,f(Y_{i,miss}|C_{i,g}=1, Y_{i,obs}=y_{i,obs})\\&=f(C_{i,g}|Y_{i,obs}=y_{i,obs}, M_i=m_i)\,f_g(Y_{i,miss}|Y_{i,obs}=y_{i,obs})\end{aligned} \qquad \text{(A4)}$$

前者爲白努利分配，其參數（即 $C_{i,g}$ 爲 1 之機率）可以貝氏定理求得

$$p(C_{i,g}=1|Y_{i,obs}, M_i) = \frac{\pi_g{}^* f_g(Y_{i,\text{obs}})^* f_g(M_i)}{\sum_{t=1}^{n_c}\pi_t{}^* f_t(Y_{i,\text{obs}})^* f_t(M_i)} \qquad \text{(A5)}$$

其中 $f_g(M_i)=\prod_{j=1}^{p}p(M_{ij}=M_{ij}|g)=\prod_{j=1}^{p}\lambda_{jg}{}^{mij}(1-\lambda_{jg})^{1-mij}$；

後者 $f_g(Y_{i,miss}|Y_{i,obs}=y_{i,obs})$ 爲第 g 個潛在類別時，給定觀察變項後遺漏變項之條件機率密度函數，如式 A3 所示，爲多元常態分配。

Estimation of Structural Equation Models with Latent Variable Pattern Mixture Model via Stochastic EM Algorithm

Chung-Ping Cheng and Li-Jen Weng

Department of Psychology, National Taiwan University

The maximum likelihood estimation method using the stochastic EM algorithm was developed for structural equation models (SEM) with latent variable pattern mixture model. Latent variable pattern mixture model is an extension of pattern mixture models with measurement errors and theoretical constructs considered. The patterns of missing were assumed to reflect latent classes rather than categories of manifest variables. Each latent class was allowed to have distinct structural equation model. The results of this simulation study indicated that the proposed estimation method via stochastic EM algorithm performed well compared to other missing data treatment methods and yielded satisfactory parameter estimates.

Keywords: *nonignorable missingness, pattern mixture model, structural equation models, maximum likelihood method, latent variable*