

## 壹、前言

一般社會科學研究的依變項大都為連續量尺，但有愈來愈多的研究其變項不屬於連續變項，例如：銀行信用風險的違約有無、組織管理某種能力的擅長與否、教育測驗的通過失敗、高血壓的診斷等，都屬於非連續變項。而這種非連續變項的迴歸模式稱為非線性模式，不管是名義變項（nominal）、次序變項（ordinal），甚至二分變項（dichotomous），在進行迴歸分析之前均必須透過連結函數（link function）做適當的轉換（王濟川、郭志剛，2003；Agresti, 2007；McCullagh & Nelder, 1989），方能與解釋變項進行迴歸分析。其中若以二分變項為依變項進行迴歸分析，是透過logit轉換以邏吉斯迴歸（logistic regression）來進行（或常態機率轉換以機率迴歸（probit regression）來分析），以多類別（multinomial）或次序變項為依變項進行迴歸分析時，也是以邏吉斯迴歸為基礎，例如，多類別邏吉斯迴歸（multinomial logistic regression）、次序邏吉斯迴歸（ordinal logistic regression）（Agresti, 2007）。因此有關非連續依變項的迴歸分析，以二元邏吉斯迴歸最為重要，可以延伸到次序與多類別依變項。

除了研究的依變項不屬於連續量尺外，也有愈來愈多的資料存在著階層（hierarchical）或多層次（multilevel）結構，或稱具有巢套或內屬（nested）的特性（Goldstein, 2011；Hox, 2010；Raudenbush & Bryk, 2002；Snijders & Bosker, 1999），亦即學生巢套於班級、班級又巢套於學校、或是重複觀測資料內屬於學生、學生內屬於學校等這樣的特徵。這樣的巢套結構資料嚴重違反了傳統迴歸分析所要求資料必須來自獨立且同質的母體，造成迴歸係數估計值的標準誤被低估，獲致容易拒絕虛無假設型 I 誤差膨脹的結果。除此之外，過去只探討個體層次解釋變項對依變項影響的研究，可能忽略重要的總體層次變項，例如，班級學生的共享氛圍、教師效能或學校資源等脈絡（contextual）變項對個體層次依變項的影響，容易獲得不精確的偏誤估計（Goldstein, 2011；Hox, 2010；Raudenbush & Bryk, 2002；Snijders & Bosker, 1999）。因此，多層次的研究也愈來愈受到社會科學研究者的重視，除了所探討的問題現象具有多層次理論外，統計方法的正確使用與實證資料的階層特徵，更凸顯多層次分析（multilevel analysis）或是階層線性模式（hierarchical linear modeling, HLM）的必要。

除了上述巢套結構外，很多實徵觀察資料並不屬於這種階層巢套，例如臺灣有許多相同社區的學生到不同學校就學以及不同社區的學生到相同學校就讀，這種資料結構稱為「交叉分類」（cross classified）（Beretvas, 2008；Goldstein, 2011；

Hox, 2010; Raudenbush & Bryk, 2002)。換言之，最底層的分析單位學生同時隸屬於兩個不同的上層結構，例如學校與社區，但這兩個不同的上層結構並沒有任何的巢套關係，在本研究稱為非典型巢套資料。Raudenbush與Bryk（2002）即示範說明社區與學校對學生成績表現的影響，和學生來自初中學校與現在高中學校對其現在成績的影響，就是屬於這種交叉分類或非典型巢套的分析。此外，有關於年齡一年代一世代的研究（age-period-cohort, APC）（Luo & Hodges, 2012），亦可以利用交叉分類或非典型巢套模型來進行。

事實上，我們所生活的環境並非完全是階層巢套關係，反倒像交叉分類。例如每天上班的工作環境即左右我們，但下班後我們所參與的社團組織亦同時影響我們對事情的看法，公司與社團組織就是兩個無嵌套的上層結構。這種資料結構不像典型的階層巢套，以本研究為例，個體層次學生內屬於第一層總體層次科系，這個科系又內屬於第二層總體層次的學校，而非典型巢套是個體層次學生同時內屬於兩種不同類別的總體層次，但這兩種類別的總體層次並不互相嵌套，例如科系與學校類型。傳統的階層巢套主要來自多階段的集群隨機抽樣，即先隨機抽取學校，再由學校中隨機抽取班級，然後由所抽取的班級中再隨機抽取學生。而非典型巢套基本上是以一個總體層次類別為抽樣架構，伴隨著重要的控制變項：另一個總體層次類別，但這兩個總體層次沒有隸屬關係（Beretvas, 2011; Hox, 2010）。交叉分類資料在多層次研究的重要性，在於同一位學生受到環境或脈絡的影響不再只有學校而已，除了學校的資源或班級老師的教學態度、自我效能、班級氛圍等因素外，連帶學生所隸屬的另一個脈絡變項：社區的平均社經地位、社區資源、社區環境氛圍等亦影響學生學業表現或是塑造人格態度養成的另一來源。本研究目的即介紹階層巢套與交叉分類邏輯迴歸分析方法，並以台灣高等教育整合資料庫現有資料進行示範分析比較。

## 貳、台灣高等教育整合資料庫的資料結構

本研究分析的資料來自國立臺灣師範大學教育研究與評鑑中心（以下簡稱台灣高等教育庫中心）於2005年對全臺灣大學一年級（以下簡稱大一）新生所做的抽樣調查，其資料庫稱為台灣高等教育整合資料庫，本研究資料庫中可以分析的資料是個體層次46,961位大一新生，但該中心為避免因學校與科系的資訊被揭露，造成資料分析結果在學校與科系間被比較，因此該中心釋放了有限資訊的科系與學校類型編碼。若根據該中心的抽樣架構，是先學校再科系、然後是學生，這樣的抽樣架構稱為階層巢套結構，其結構如圖1所示（Beretvas, 2008; Rasbash &