# 行政院國家科學委員會專題研究計畫 成果報告

# 基於 GHMM 之核平滑化無參數 IRT 教育測驗分析模型之研究（I）

## 研究成果報告(精簡版)

計 畫 類 別 ： 個別型
計 畫 編 號 ： NSC 95-2413-H-468-003-
執 行 期 間 ： 95 年 08 月 01 日至 96 年 07 月 31 日
執 行 單 位 ： 亞洲大學生物科技學系

計畫主持人： 劉湘川
共同主持人： 郭伯臣
計畫參與人員： 碩士班研究生-兼任助理：黃文俊、楊智為、張雅媛
計畫參與人員：歐順德、林奎光、蔡伯仁、林文質

報 告 附 件 ： 出席國際會議研究心得報告及發表論文

處 理 方 式 ： 本計畫涉及專利或其他智慧財產權，1 年後可公開查詢

中 華 民 國 96 年 10 月 31 日

# 行政院國家科學委員會補助專題研究計畫

基於 GHMM 之核平滑化無參數 IRT 教育測驗分析模型之研究(I)

計畫類別： 個別型計畫

計畫編號：NSC－95－2413－H－468－003

執行期間：95 年 8 月 1 日至 96 年 7 月 31 日

執行單純：亞洲大學生物科技與生物資訊學系

計畫主持人：劉湘川

共同主持人：郭伯臣

計畫參與人員：黃文俊、歐順德、林奎光、蔡伯仁、林文質

報告類型： 精簡報告
報告附件： 出席國際會議研究心得報告及發表論文
處理方式： 本計畫可公開查詢

中 華 民 國 96 年10 月31 日

**「基於 GHMM 之核平滑化無參數 IRT 教育測驗分析模型之研究」**

**主持人：劉湘川　　　　共同主持人：郭伯臣**

# 壹、研究動機與計畫概述

隨著資訊科技快速進步、測驗形式的改變及需求量的快速增加，以試題反應理論（IRT）為核心的現代測驗理論廣泛應用於教育測驗的編製、施測、計分、分析與解釋上。但眾所周知廣為應用之三參數洛基 IRT 模式，在教育測驗的實際應用上仍同時具有下列四種不利之限制：

一、　三參數洛基 IRT 模式，有測驗試題局部獨立之限制，不適用於像全民英檢類之時間序列測驗，當其前後試題間有狀態轉移序列相關之測驗分析

二、　三參數洛基 IRT 模式，有測驗試題局部獨立之限制，亦不適用於非時間序列測驗，當試題間有關聯之測驗分析，且無法與試題關聯結構分析法或試題順序理論分析法整合應用。

三、　常用之三參數洛基 IRT 模式並未考慮受試者遺漏作答或未予作答情況，不完全符合實際教育測驗之應用。

四、　常用之三參數洛基 IRT 模式是極大樣本之估計模式，至少需上千受試者之樣本，才可得穩定可靠之參數估計值，只適用於大型入學測驗等，不適用於個別學校或班級之教學測驗。

針對第一種及第二種限制，亦即：當一種教育測驗之前後試題間，若有狀態轉移序列相關時，我們可考慮採用具有該特性之「隱藏式馬可夫模型(HMM)」(劉湘川 2004)，或不受該限制之「無參數試題反應理論模式(KN-IRT)」，至於原有之「隱藏式馬可夫模型(HMM)」由於只能分析所有試題之選項分配之亂度相等之測驗，因而，劉湘川於 2004 年將其修正擴張為「廣義隱藏式馬可夫模型(GHMM)」，提出 GHMM 專有適用之三種參數最佳化解演算法，獲得國科會資助進而發展獲得「基於廣義隱藏式馬可夫模型之教育測驗分析模型」，該模式不僅可分析如全民英檢類所有試題選項分配可為任意亂度之狀態轉移序列相關之測驗試題，同時可分析每一試題包含「認知作答」、「猜測作答」、及「未予作答」之三種作答機率，亦即該模式兼具改進第三種限制之優點，唯該研究僅屬古典測驗理論具有上述優點之改進模式，不能精確分析受試者之個別能力，固有必要進一步結合可分析受試者能力之有關試題反應理論模式，整合改進為不具上述所有限制之完備教育測驗分析模式。

為同時克服第一、第二，第四種之限制，宜考慮可分析受試者能力之「無參數試題反應理論模式(NP-IRT)」不僅均無試題局部獨立之假設條件，不具第一、二兩種限制，且可適用受試者千人以下較小之樣本，亦即，不具第四種限制。較

為盛行者為 Ramsay 於 1991 年提出之「核平滑化無參數試題反應理論模式(KN-IRT)」，劉湘川以相關鑑別度及高階相關鑑別度等替代 Ramsay 所提之高低鑑別度進行受試者能力之改進估計，分別於 2000 年、2001 年、2002 年、2003 年獲得系列「改進之核平滑化無參數試題反應理論模式(KN-IRT)」及進而與試題關聯結構分析法整合之應用模式，但可惜的是；該改進之系列無參數試題反應理論模式與一般無參數試題反應理論模式一樣，不具分析猜測之功能，固此模式有必要進一步結合可同時處裡未答及猜測之「廣義隱藏式馬可夫模型(GHMM)」，整合改進為不具上述所有限制之完備教育測驗分析模式。

針對第三種限制之改進，劉湘川與劉新梧於 2001 年以引進隨機未作答虛擬選項方式提出「不完全資料之多元計分三參數試題選項分析擴充模式」，這是三參數洛基 IRT 模式兼顧未答情況之改進模式，可具受試者遺漏作答或未予作答之分析功能，優於既有之參數型試題反應理論模式，但卻仍未能改善第一、二、四之不利限制，類似地，劉湘川亦於 2001 年以引進隨機未作答虛擬選項方式提出「核平滑化無參數試題選項分析擴充模式」，則可具受試者遺漏作答或未予作答之分析功能，亦即，改進之「核平滑化無參數試題選項分析模式」，可同時克服上述四種限制，唯仍不具分析猜測之功能，仍有必要進一步結合可同時處裡未答及猜測之「廣義隱藏式馬可夫模型(GHMM)」，以整合發展成不具上述所有限制之完備教育測驗分析模式。

綜合以上所述，得悉；三參數洛基 IRT 模式及其改進模式其估計雖較精確，卻不適用於較小樣本之學校及班級教學測驗，亦不適用於狀態轉移序列相關之測驗分析，更不能與試題關聯結構分析法或試題順序理論分析法整合應用，而劉湘川所提之「改進之核平滑化無參數試題選項反應理論模式(KN-IRT)」正好克服了上述模式之缺失，只有未解決之猜測問題，則可考慮藉由與「廣義隱藏式馬可夫模型(GHMM)」整合而解決之，這正是本研究計畫之背景與動機。

若只考慮應用於大型入學測驗，且無需與試題關聯結構分析法或試題順序理論分析法整合應用，即無第二種限制及第四種限制之考慮時，則可結合參數型現代測驗理論，發展出較精確之「基於廣義隱藏式馬可夫模型兼顧未作答之參數型 IRT 教育測驗分析模型」，若兼顧小型教學測驗及大型入學測驗，且考量與試題關聯結構分析法或試題順序理論分析法整合應用時，即須兼顧四種不利限制之考慮時，則應結合改進之無參數型現代測驗理論，發展出較實用之「基於廣義隱藏式馬可夫模型兼顧猜測及未作答之無參數型 IRT 教育測驗分析模型」。

本研究計畫主要目的即擬在作者多年持續研究發展之既有成果基礎下，結合以作者先前所發展之「基於廣義隱藏式馬可夫模型之教育測驗分析模型」，與作者已提出之「改進之核平滑化無參數試題反應理論模式」之兩種模式，發展出兼顧上述四種限制，可彈性適用各種教育測驗之整合新模型；「基於廣義隱藏式馬可夫模型兼顧猜測及未作答之核平滑化無參數試題反應理論模式」及其便利使用之電腦程式，並以蒙地卡羅模擬研究法，進行所擬發展之完備整合模型；「基於 GHMM 之核平滑化無參數 IRT 教育測驗分析模型」，與既有常用之「三參數洛基 IRT 分析模型」、及「改進之核平滑化無參數試題反應理論分析模式(KN-IRT)」

**大量資料之模擬實驗比較，以驗證作者所擬發展之該整合新模型之優越性，進而提出其與「IRS」或「OT」之整合應用模式。**

## 貳、本研究本年度為第一年度、主要工作為：

一. 探討HMM、**GHMM**、IRT、NP-IRT 相關文獻

二. 研發結合**GHMM與**IRT**之整合模式**，GHMM**與**NP-IRT**之整合模式的**相關演算法

三. 蒙地卡羅模擬資料試驗

四. 實際教育測驗資料試驗

**五.** 撰寫研究報告

原定計畫進度規劃如下

| 月份<br>研究步驟 | 95.8 | 95.9 | 95.10 | 95.11 | 95.12 | 96.1 | 96.2 | 96.3 | 96.4 | 96.5 | 96.6 | 96.7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 蒐集文獻資料 | ■ | ■ | ■ | ■ | ■ | | | | | | | |
| GHMM+PIRT 數學理論探究 | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | |
| GHMM+KN-IRT 數學理論探究 | | | ■ | ■ | ■ | ■ | ■ | ■ | | | | |
| GHMM+KN-IRT 電腦分析系統設計 | | | | ■ | ■ | ■ | ■ | ■ | ■ | | | |
| 進行模擬資料實驗 | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | |
| 進行實際資料實驗 | | | | | | | | | ■ | ■ | ■ | |
| 撰寫研究報告 | | | | | | | | | | | ■ | ■ |

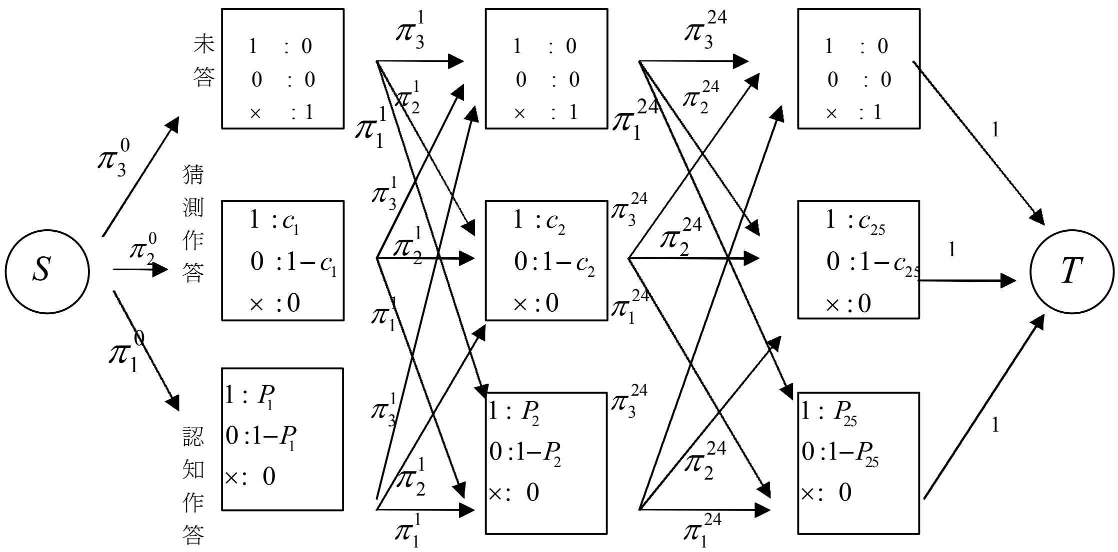
| 月份<br>研究步驟 | 96.8 | 96.9 | 96.10 | 96.11 | 96.12 | 97.1 | 97.2 | 97.3 | 97.4 | 97.5 | 97.6 | 97.7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 蒐集文獻資料 | ■ | ■ | ■ | ■ | ■ | | | | | | | |
| GHMM+KN-IRT+ OT 數學理論探究 | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | |
| GHMM+KN-IRT+ IRS 數學理論探究 | | | ■ | ■ | ■ | ■ | ■ | ■ | | | | |
| GHMM+KN-IRT+ SS 數學理論探究 | | | | ■ | ■ | ■ | ■ | ■ | ■ | | | |
| 進行模擬資料實驗 | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | |
| 進行實際資料實驗 | | | | | | | | | ■ | ■ | ■ | |
| 撰寫研究報告 | | | | | | | | | | | ■ | ■ |

## 參、 研究成果

### 一、提出 GHMM 與 IRT 及 KN-IRT 整合模式演算法

#### （一）GHMM 與 IRT 及 KN-IRT 整合模式

【說明簡例】 下圖顯示 W 個受試者作答 25 選擇題、每人在作答每一試題時，可自 3 種作答策略：「認知作答」、「猜測作答」、「遺漏作答」中依作答能力選取其一作答，任一作答結果同樣有「答對」、「答錯」、及「未答」三種可能，分別以符號 $+,-,\times$ 表示，只是發生之機率有別，其中作答第 t 題時，採取「遺漏作答」時之「未答機率為 1」，採取「猜測作答」時之「猜對機率為 $c_t$ 」採取「認知作答」時之「作對機率為 $P_t=\overline{P_t(\theta)}=\frac{1}{W}\sum_{w=1}^{W}P_t(\theta)$ 」，其中 $P_t(\theta)$ 為分離猜測及未答後之第 t 題受試能力 $\theta$ 之無參數試題特徵曲線，開始選答第一題時採取 3 種作答策略之機率分別為狀態轉移機率 $\pi_1^0,\pi_2^0,\pi_3^0$ 選答完第 t 題-在作答第(t+1)題時，採取 3 種作答策略之機率分別為狀態轉移機率 $\pi_1^t,\pi_2^t,\pi_3^t$ ，上述整個過程即為一 GHMM 與 KN-IRT 整合模式應用之例




$$S \longrightarrow \underset{s_1=2}{1} \longrightarrow \underset{s_2=3}{\times} \longrightarrow \underset{s_{25}=1}{0} \longrightarrow \mathbf{T}$$

圖 1、25 選擇題 3 作答策略 3 作答符號、變動狀態轉移機率、測驗分析示意圖、

其中：$\pi_1^t, \pi_2^t, \pi_3^t, c_t \in (0,1)$，$\pi_1^t + \pi_2^t + \pi_3^t = 1$，$a_t > 0$，$b_t \in R$

$$P_t = \overline{P_t(\theta)} = \frac{1}{W}\sum_{w=1}^{W} P_t(\theta), \quad t = 1,2,...,25,$$

**1.** 若爲 GHMM 與 IRT 之整合模式時 $P_t(\theta) = \dfrac{1}{1+\exp\left[-a_t(\theta - b_t)\right]}$

**2.** 若爲 GHMM 與 KN-IRT 之整合模式時

$P_t(\theta)$ 爲分離猜測及未答後之第 t 題受試能力 $\theta$ 之無參數試題特徵曲線

$$P_t(\theta) = \frac{\sum_{w=1}^{W} \exp\left[-\frac{W^{\frac{2}{5}}\left(\hat{\theta}_w - \theta\right)^2}{2.42}\right] x_{wt}}{\sum_{w=1}^{W} \exp\left[-\frac{W^{\frac{2}{5}}\left(\hat{\theta}_w - \theta\right)^2}{2.42}\right]} \quad , t = 1,2,...,T, w = 1,2,...,W$$

$\overline{P_t(\theta)} = \frac{1}{W}\sum_{w=1}^{W} P_t(\theta)$ 爲分離猜測及未答後之第 t 題全部受試平均答對機率

$$P_t(\theta) = \frac{1}{1+\exp\left[-a_t(\theta - b_t)\right]}$$

**（二）GHMM 與 IRT 及 KN-IRT 整合模式估計演算法簡介**

採用 **GHMM 與 IRT 及 KN-IRT 之二段估計演算法，**

1、第一階段：基於固定轉移機率 GHMM 估計

(1). 初始估計值之設定

$$因\ D_t = \begin{bmatrix} d_{11}^t & d_{12}^t & d_{13}^t \\ d_{21}^t & d_{22}^t & d_{23}^t \\ d_{31}^t & d_{32}^t & d_{33}^t \end{bmatrix} = \begin{bmatrix} P_t & (1-P_t) & 0 \\ c_t & (1-c_t) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad t = 1,2,...,T$$

，其中 $d_{13}^t = d_{23}^t = d_{31}^t = d_{32}^t = 0$, $d_{33}^t = 1$ ，$t = 1,2,...,T$ 恒為常數，不需估計，只要考慮 $P_t, c_t$ 之估計初始值，通常選擇型測驗為四或五選一，可令猜測機率估計初始值為 $c_t = 0.2$ ，$t = 1,2,...,T$ ，平均答對機率估計初始值為

$P_t = 0.5$ , $t = 1,2,...,T$ 。

因 $y_{wt3}$ 為受試者 w 作答試題是否採取未答策略 3 之可觀察指示變數，可令

$\pi_3^t = \pi_3 = \frac{1}{W}\sum_{w=1}^{W} y_{w3t}$ ，並令 $\pi_1^t = \pi_1 = 0.9\left(1-\pi_3\right)$ 、 $\pi_2^t = \pi_2 = 0.1\left(1-\pi_3\right)$

為對應參數估計之起始值。

(2). 進行 GHMM 估計法

利用上述設定之初始估計值可有效地估得 $\left(\pi_1^t, \pi_2^t, \pi_3^t\right) = \left(\pi_1, \pi_2, \pi_3\right)$ 、$c_t$ 、

$P_t = \overline{P_t(\theta)}$ ， $t = 1,2,...,T$ 。

2、第二階段：IRT 混合模式之 MLE-EM 估計

(1) 參數起始估計值之設定

i. 試題難度參數起始估計值 $b_t^{(0)}$ 之設定

由第一階段 GHMM 估得之 $P_t = \overline{P_t(\theta)}$ ，實質上為第 t 題不含猜對之平均答對率，亦即古典測驗理論第 t 題之試題難度，仿 Fisher 之 Z 轉換，可轉換成 IRT 混合模式第 t 題之試題難度起始估計值如下：

$$b_t^{(0)} = \frac{1}{2}\log\sqrt{\frac{1-p_t}{p_t}} \quad , \quad t = 1,2,...,T$$

ii. 試題鑑別度參數起始估計值 $a_t^{(0)}$ 之設定

考慮以古典測驗理論第 t 題之試題相關鑑別度，轉換成 IRT 混合模式第 t 題之試題鑑別度起始估計值如下：

因 $x_{wt}$ 為受試者 $w$ 是否答對第 $t$ 題指示函數，且受試者 $w$ 對第 $t$ 題採取認知作答之轉移策略之機率為 $\pi_1^{t-1}$ 其後答對第 $t$ 題為 $p_t$，可得受試者 $w$ 以

認知作答 T 題全部答對之以 $\frac{\pi_1^{t-1} p_t}{\pi_1^{t-1} p_t + \pi_2^{t-1} c_t}$ 為權值之加權總分為

$$X_w = \sum_{t=1}^{T} x_{wt} \cdot \frac{\pi_1^{t-1} p_t}{\pi_1^{t-1} p_t + \pi_2^{t-1} c_t} \quad , \quad w = 1, 2, ..., W$$

以所有受試者 $w$ 之認知作答加權總分為 $X_w$ 與其是否答對第 $t$ 題指示函數 $x_{wt}$ 所得之點二系列相關係數 $r_t^{(0)}$，即為第 $t$ 題之古典測驗理論試題鑑別度如下：

$$r_t^{(0)} = \frac{\sum_{w=1}^{W} \left(X_w - \bar{X}\right)\left(x_{wt} - \bar{x}_t\right)}{\sqrt{\sum_{w=1}^{W} \left(X_w - \bar{X}\right)^2} \sqrt{\sum_{w=1}^{W} \left(x_{wt} - \bar{x}_t\right)^2}} \quad , \quad t = 1, 2, ..., T$$

其中：$\bar{X} = \frac{1}{W} \sum_{w=1}^{W} X_w \quad , \quad \bar{x}_t = \frac{1}{W} \sum_{w=1}^{W} x_{wt}$

參照 Fisher 之 Z 轉換，可轉換成 IRT 混合模式第 $t$ 題之試題鑑別度起始估計值如右：$a_t^{(0)} = \log \sqrt{\frac{1 + r_t^{(0)}}{1 - r_t^{(0)}}}$

iii.受試者 $w$ 能力參數起始估計值 $\theta_w^{(0)}$ 之設定

令 $S_X = \left[ \frac{1}{W-1} \sum_{w=1}^{W} \left(X_w - \bar{X}\right)^2 \right]^{\frac{1}{2}}$，得能力參數起始估計值如下：

$$\theta_w^{(0)} = \frac{X_w - \bar{X}}{S_X} \quad , \quad w = 1, 2, ..., W$$

iv. 受試者 w 認知作答試題 t 之 IRT 起始估計式 $P_t^{(0)}\left(\theta_w^{(0)}\right)$ 如下：

$$P_t^{(0)}\left(\theta_w^{(0)}\right) = \frac{1}{1 + \exp\left[-a_t^{(0)}\left(\theta_w^{(0)} - b_t^{(0)}\right)\right]}$$

v. 受試者 w 作答試題 t 採取作答策略 j 之不可觀察指示變數 $y_{wtj}^{(0)}$ 起始估計值如下：

$$y_{wt1}^{(0)}=\frac{\pi_1\left[P_t^{(0)}\left(\theta_w^{(0)}\right)\right]^{x_{wt}}\left[1-P_t^{(0)}\left(\theta_w^{(0)}\right)\right]^{1-x_{wt}}}{\pi_1\left[P_t^{(0)}\left(\theta_w^{(0)}\right)\right]^{x_{wt}}\left[1-P_t^{(0)}\left(\theta_w^{(0)}\right)\right]^{1-x_{wt}}+\pi_2 c_t^{x_{wt}}\left(1-c_t\right)^{1-x_{wt}}}$$

$$y_{wt2}^{(0)}=\frac{\pi_2 c_t^{x_{wt}}\left(1-c_t\right)^{1-x_{wt}}}{\pi_1\left[P_t^{(0)}\left(\theta_w^{(0)}\right)\right]^{x_{wt}}\left[1-P_t^{(0)}\left(\theta_w^{(0)}\right)\right]^{1-x_{wt}}+\pi_2 c_t^{x_{wt}}\left(1-c_t\right)^{1-x_{wt}}}$$

假使 $y_{wt3}^{(0)}=0$ 則 $y_{wt1}^{(0)}+y_{wt2}^{(0)}=1$，否則 $y_{wt3}^{(0)}=1$。

(2) 估算起始概似函數 $l^{(0)}$。

將 $\left(\pi_1\ ,\pi_2,\pi_3\right)$ 、$c_t$ 、$P_t^{(0)}\left(\theta_w^{(0)}\right)$ 、$y_{wtj}^{(0)}$ 及受試 w 作答試題 t 答對與否之指示變數 $x_{wt}$ 代入估算，則由試題間局部獨立，受試間獨立，且受試間與試題間獨立可得概似函數如下：

$$l=l_1+l_2+l_3$$

$$=\ln\left[\prod_{w=1}^{W}\prod_{t=1}^{T}\left(\left[\pi_1 P_t^{x_{wt}}\left(\theta_w\right)\left(1-P_t\left(\theta_w\right)\right)^{1-x_{wt}}\right]^{y_{wt1}}\left[\pi_2 c_t^{x_{wt}}\left(1-c_t\right)^{1-x_{wt}}\right]^{y_{wt2}}\left[\pi_3\right]^{y_{wt3}}\right)\right]$$

其中：$P_t\left(\theta_w\right)=\dfrac{1}{1+\exp\left[-a_t\left(\theta_w-b_t\right)\right]}$

$$l_1=\sum_{w=1}^{W}\sum_{t=1}^{T}\left(y_{wt1}\ln\pi_1+y_{wt2}\ln\pi_2+y_{wt3}\ln\pi_3\right)$$

$$l_2=\sum_{w=1}^{W}\sum_{t=1}^{T}y_{wt1}\left[x_{wt}\ln P_t\left(\theta_w\right)+\left(1-x_{wt}\right)\ln\left(1-P_t\left(\theta_w\right)\right)\right]$$

$$l_3=\sum_{w=1}^{W}\sum_{t=1}^{T}y_{wt2}\left[x_{it}\ln c_t+\left(1-x_{wt}\right)\ln\left(1-c_t\right)\right]$$

(3) 估出 $a_t^{(1)}$ 、$b_t^{(1)}$ 和 $\theta_w^{(1)}$ 值，可以求出新的 $P_t^{(1)}\left(\theta_w^{(1)}\right)$，再算出新的實際採取作答策略 j 與否之指示變數 $y_{wtj}^{(1)}$。估計說明如下：

i. 試題 t 之鑑別度參數 $a_t^{(1)}$ 與難度參數 $b_t^{(1)}$ 之估計：

$$\begin{bmatrix} a_t^{(1)} \\ b_t^{(1)} \end{bmatrix} = \begin{bmatrix} a_t^{(0)} \\ b_t^{(0)} \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 l}{\partial a_t^2} & \frac{\partial^2 l}{\partial a_t \partial b_t} \\ \frac{\partial^2 l}{\partial b_t \partial a_t} & \frac{\partial^2 l}{\partial b_t^2} \end{bmatrix}_0^{-1} \begin{bmatrix} \frac{\partial l}{\partial a_t} \\ \frac{\partial l}{\partial b_t} \end{bmatrix}_0$$

其中：$\left[\frac{\partial l}{\partial a_t}\right]_0 = \sum_{w=1}^{W} y_{wt1}^{(0)} \left(\theta_w^{(0)} - b_t^{(0)}\right)\left[x_{wt} - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]$

$$\left[\frac{\partial l}{\partial b_t}\right]_0 = -\sum_{w=1}^{W} y_{wt1}^{(0)} a_t^{(0)} \left[x_{wt} - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]$$

$$\left[\frac{\partial^2 l}{\partial a_t^2}\right]_0 = -\sum_{w=1}^{W} y_{wt1}^{(0)} \left(\theta_w^{(0)} - b_t^{(0)}\right)^2 P_t^{(0)}\left(\theta_w^{(0)}\right)\left[1 - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]$$

$$\left[\frac{\partial^2 l}{\partial b_t^2}\right]_0 = -\sum_{w=1}^{W} y_{wt1}^{(0)} \left[a_t^{(0)}\right]^2 P_t^{(0)}\left(\theta_w^{(0)}\right)\left[1 - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]$$

$$\left[\frac{\partial^2 l}{\partial a_t \partial b_t}\right]_0 = \left[\frac{\partial^2 l}{\partial b_t \partial a_t}\right]_0$$

$$= \sum_{w=1}^{W} y_{wt1}^{(0)} \left[a_t^{(0)}\left(\theta_w^{(0)} - b_t^{(0)}\right) P_t^{(0)}\left(\theta_w^{(0)}\right)\left[1 - P_t^{(0)}\left(\theta_w^{(0)}\right)\right] - \left[x_{wt} - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]\right]$$

ii. 受試者 w 能力參數 $\theta_w^{(1)}$ 之估計：

$$\theta_w^{(1)} = \theta_w^{(0)} - \left[-\sum_{t=1}^{T} y_{wt1}^{(0)}\left[a_t^{(0)}\right]^2 P_t^{(0)}\left(\theta_w^{(0)}\right)\left[1 - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]\right]^{-1}\left[\sum_{t=1}^{T} y_{wt1}^{(0)} a_t^{(0)}\left[x_{wt} - P_t^{(0)}\left(\theta_w^{(0)}\right)\right]\right]$$

iii. 受試者 w 作答試題 t 採取作答策略 j 之不可觀察指示變數 $y_{wtj}^{(1)}$ 之估計：

$$y_{wt1}^{(1)} = \frac{\pi_1\left[P_t^{(1)}\left(\theta_w^{(1)}\right)\right]^{x_{wt}}\left[1 - P_t^{(1)}\left(\theta_w^{(1)}\right)\right]^{1-x_{wt}}}{\pi_1\left[P_t^{(1)}\left(\theta_w^{(1)}\right)\right]^{x_{wt}}\left[1 - P_t^{(1)}\left(\theta_w^{(1)}\right)\right]^{1-x_{wt}} + \pi_2 c_t^{x_{wt}}\left(1 - c_t\right)^{1-x_{wt}}}$$

$$y_{wt2}^{(1)} = \frac{\pi_2 c_t^{x_{wt}} \left(1-c_t\right)^{1-x_{wt}}}{\pi_1 \left[P_t^{(1)}\left(\theta_w^{(1)}\right)\right]^{x_{wt}} \left[1-P_t^{(1)}\left(\theta_w^{(1)}\right)\right]^{1-x_{wt}} + \pi_2 c_t^{x_{wt}} \left(1-c_t\right)^{1-x_{wt}}}$$

假使 $y_{wt3}^{(1)}=0$ 則 $y_{wt1}^{(1)}+y_{wt2}^{(1)}=1$，否則 $y_{wt3}^{(1)}=1$

iv. 將 $\left(\pi_1, \pi_2, \pi_3\right)$ 、$c_t$ 、$x_{wt}$ 與步驟 3 估出的新值 $P_t^{(1)}\left(\theta_w^{(1)}\right)$ 、$y_{wtj}^{(1)}$ 代入估算可得概似函數 $l^{(1)}$ 。

v. 計算 $|l^{(1)}-l^{(0)}|$ 是否小於設定門檻值，若不是則新參數值等於舊參數值，反覆遞迴步驟 3、4 求出新的概似函數，直到 $|l^{(k+1)}-l^{(k)}|$ 小於設定門檻值，終止參數迭代估計。

經基於固定轉移機率 GHMM 之 IRT 混合模式之兩階段估計，可得 $\hat{\pi}_1=\pi_1, \hat{\pi}_2=\pi_2, \hat{\pi}_3=\pi_3$ 、 $\hat{c}_t=c_t$ 、 $\hat{a}_t=a_t^{(k)}, \hat{b}_t=b_t^{(k)}$ 、 $\hat{\theta}_w=\theta_w^{(k)}$ 、 $\hat{y}_{wt2}=y_{wt2}^{(k)}$ $\hat{y}_{wt1}=y_{wt1}^{(k)}$，、 $\hat{P}_t\left(\hat{\theta}_w\right)=\dfrac{1}{1+\exp\left[-a_t^{(k)}\left(\theta_w^{(k)}-b_t^{(k)}\right)\right]}$

**3.** 若為 GHMM 與 NP- IRT 混合模式之估計

只須以劉湘川之改進之核平滑 NP- IRT 之估計法替代參數型 IRT 之 MLE 估計法即可

## 肆、蒙地卡羅模擬實驗成果

一、蒙地卡羅模擬研究步驟

1. 以廣義隱藏式馬可夫模型與核平滑化無參數試題反應理論結合模式為基礎，利用程式軟體 MATLAB 撰寫估計試題特徵曲線精確度程式。
2. 比較參數試題反應理論、廣義隱藏式馬可夫模型與參數試題反應理論混和模式、核平滑化無參數試題反應理論與廣義隱藏式馬可夫模型與核平滑化無參數試題反應理論整合模式四種估計試題特徵曲線精確度。
3. 比較不同受試者人數對估計試題特徵曲線精確度的影響。

二、模擬實驗執行流程

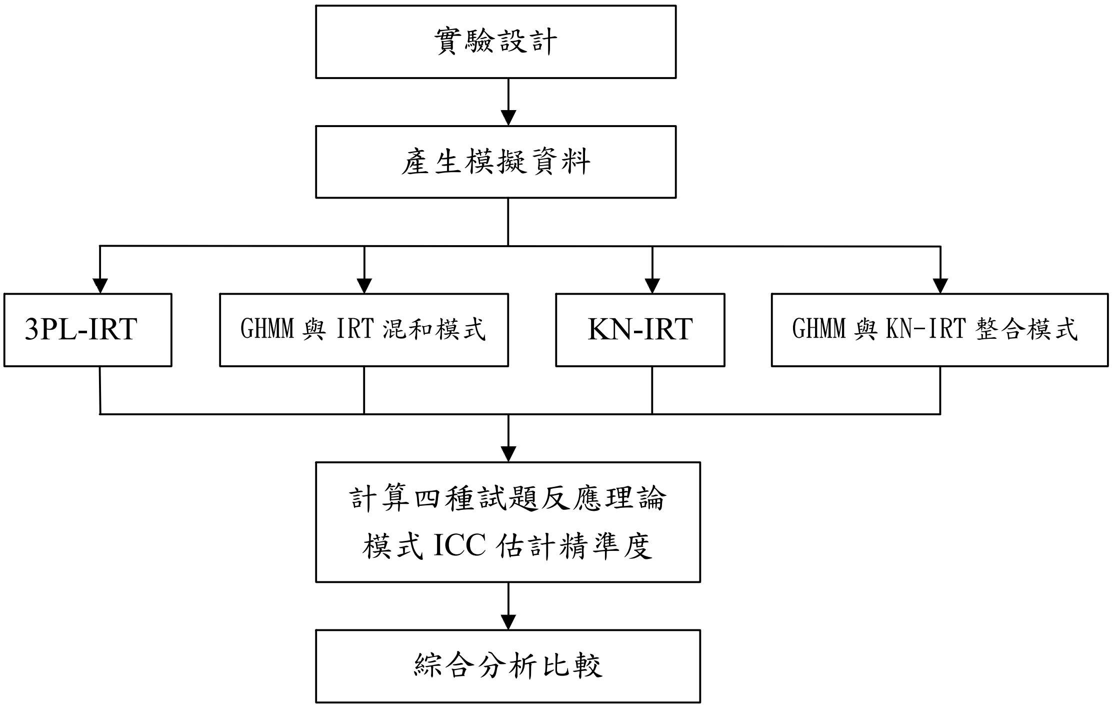



圖 2 模擬實驗執行流程圖

三、模擬實驗結果果

根據先前提到的研究方法與研究架構，進行模擬實驗。在試題數 25 題，受試者人數分別為 100、200、500、1000、1500、2000 人六種情形，估計出四種試題反應理論模式之 ICC 精準度，共有 24 種組合方式加以了解。且每組情形都模擬 50 次。以下每一種組合皆為模擬實驗 50 次之平均結果，表 1 為模式與人數不同組合下 ICC 估計值與真值之 *RMSE* 之平均值，結果顯示參數型 GHMM+2PL-IRT 整合模式優於參數型 3PL-IRT 模式，且無參數型 GHMM+KNIRT 整合模式優於無參數型 KN-IRT

表 1 模式與人數組合下 ICC 估計值與真值之 *RMSE* 之平均值

| 模式＼人數 | 100 | 200 | 500 | 1000 | 1500 | 2000 |
|---|---|---|---|---|---|---|
| 3PL-IRT | 0.1910 | 0.1601 | 0.1630 | 0.1450 | 0.1281 | 0.1254 |
| GHMM+2PL-IRT | 0.1488 | 0.1279 | 0.1246 | 0.1052 | 0.1090 | 0.1071 |
| KN-IRT | 0.1834 | 0.1703 | 0.1598 | 0.1444 | 0.1297 | 0.1193 |
| GHMM+KNIRT | 0.1482 | 0.1431 | 0.1155 | 0.0947 | 0.0847 | 0.0792 |

**伍、教育測實驗結果**簡介

**本研究編製**概念題及應用題併列之題組型時間序列診斷測驗內容如下，試題如附錄，經 GHMM+NP-IRT 實驗結果可估得受試學生能力，試題特徵曲線，認知作答、猜測作答、等參數值，有利於個別受試學生之診斷教學

一、時間序列診斷測驗

（一）試題來源：國小康軒版之數學教科書、教學指引

（二）試題範圍：六年級上下學期共 11 個單元

| | 六年級上學期 | 六年級下學期 |
|---|---|---|
| 單元名稱 | 因數、倍數 | 百分率 |
| | 速率 | 表面積 |
| | 數的四則 | 分數四則 |
| | 面積 | 柱體的體積 |
| | 分數的加減 | 容量 |
| | | 大單位的換算 |
| | | 時間的計算 |

（三）試題數量：十個題組，每組兩題，共 20 題

（四）試題編製原則：

1.依雙向細目表命題：

根據雙向細目表及參考試題編製應有的命題原則與技巧，由各單元中挑選一題基本概念題及一題應用或理解之較難問題，逐一設計、撰寫。

2.編製時間序列題組型測驗

為達成研究目的，將試題編成簡報格式，同一單元的兩題呈現在一頁中，並限定呈現時間及設定題組轉換時的提醒聲音，編製成兩題為一題組的時間序列測驗。

## 陸、計畫結果自評

1. 本研究完成 GHMM 之 IRT 教育測驗分析整合模型之下列相關事宜

   （1）基於 GHMM 之參數型 IRT 教育測驗分析模型之估計方法

   （2）基於 GHMM 之核平滑化無參數型 IRT 教育測驗分析模型之估計方法

   （3）基於 GHMM 之參數型 IRT 教育測驗分析模型之電腦應用系統之設計

   (4)基於 GHMM 之核平滑化無參數型 IRT 教育測驗分析模型之電腦應用系統之設計

   （5）進行大量蒙地卡羅模擬資料實驗

   （6）實際應用資料實驗

2. 對於學術研究、國家發展及其他應用方面預期之貢獻

   學術研究方面，本計畫發展新的教育測驗分析模型理論分析技術，發表相關論文 2 篇；在國家發展方面，本研究發展新的教育測驗分析模型理論分析技術，，可供在各類國家教學研究分析用。

3. 參與之工作人員，預期可獲之訓練如下：

   （1）瞭解基於二點計分 GHMM 之核平滑化無參數型 IRT 教育測驗分析模型

   （2）瞭解時間序列與非時間序列試題間有關聯之教育測驗分析模式

   （3）如何產生時間序列教育測驗統計分析模型模擬資料進行實驗

   (4) 如何撰寫期刊論文

# 參考文獻


劉湘川(2000)：點二系列相關試題鑑別指數之值譜分析及其在 IRT 上之應用。測驗統計年刊第八輯。1-20 頁。台中市：國立台中師範學院。

劉湘川(2001a)：相關加權核平滑化無參數試題選項特徵曲線估計法及其 IORS 整合模式。第五屆華人社會心理與教育測驗學術研討會。C5.1 ，1-10 頁。台北市：中國測驗學會、台灣師範大學。

劉湘川(2001b)：核平滑化試題選項特徵曲線與選項關聯結構整合擴充模式。測驗統計年刊第九輯。1-18 頁。台中市：國立台中師範學院。

劉湘川、劉新梧(2001)：不完全資料之多元計分三參數試題選項分析擴充模式。測驗統計年刊第九輯。19-45 頁。台中市：國立台中師範學院。

劉湘川(2003a)：高階相關比累進加權核平滑化試題選項分析綜合模式。測驗統計年刊第十輯。197-218 頁。台中市：國立台中師範學院。

劉湘川(2003b)混合型語義結構分析之研究。測驗統計年刊第十一輯。1-12 頁。台中市：國立台中師範學院。

劉湘川(2003c)：核平滑化試題與選項分析模式之條件最大概似數值估計。測驗統計年刊第十一輯。17-40 頁。台中市：國立台中師範學院。

劉湘川、楊志良(2003)態度問題關聯結構分析方法之發展-以健保態度問題為例。第六屆工程科技與中西醫學應用研討會。台中縣：台中健康暨管理學院。

劉湘川、劉東昇（2003）態度問題關聯結構分析方法之研究。中華心理學會。台北市：輔仁大學。
劉湘川、簡茂發（2004.）混合型態度問題關聯結構分析。*第六屆兩岸心理與教育測驗學術研討會*。中國測驗學會。陝西師範大學。
劉湘川(2004)：廣義隱藏式馬可夫模型應用於測驗分析之研究。
測驗統計年刊第十二輯。1-22 頁。台中市：國立台中師範學院。
竹谷誠（1987）。評定尺度デ ータの意味分析法。日本行動計量學會誌，14，2，10-17。

Chia-Lin Shen, Kuei-Jen Lee, Hsiang-Chuan Liu..（2004.）Ordering Analysis of Gene Expression Dynamics. The 21st Workshop on Combinatorial Mathematics and Computation Theory
Hsiang-Chuan Liu, Ten-Wei Hsiuh & Bor-Chen Kuo, (2003.). ：Item Ordering Theories Based on Nonparametric Item Response Theory, IMPS-2003 International Meeting of the Psychometric Society, Sardinia, Italy.

MacDonald, I. L., & Zucchini, W. (1997). Hidden Markov and Other Models for Discrete-valued Time Series. London: Chapman&Hall.
Makoto Takeya（1999）Structure analysis methods for instruction, Takushoku University Press, Hachioji, Tokyo, Japan
Ramsay,J.O. (1991). Kernel smoothing approaches to nonparametric item characteristic curve estimation. *Psychometrika, 56,* 611-630.

附錄：時間序列診斷測驗試題

康軒版數學第十一、十二冊時間序列電腦測驗共 20 題
六年級數學領域　　性別：□男/ □女
_______縣(市)_______國小　六年_____班 姓名：_______________座號：_______

| 六年級數學領域 GHMM 題組型序列測驗<br><br>國立台中教育大學測驗統計研究所<br>指導教授：劉湘川　博士 | 注意事項<br>一、試題共 20 題。<br>二、每次依序顯示相關的 2 題在螢幕上，會以「打字聲」告知題目的呈現，每組顯示 1.5 分鐘，滿 1.5 分鐘時停止停止顯示。<br>三、請在答案紙上依題號作答。 |
|---|---|

| | |
|---|---|
| 研究生：林奎光 | 四、交卷時，請檢查你是否已在答案紙上的基本資料欄上填妥你的基本資料？ |
| 測驗開始 | 1.（　）下列何者是 27 和 33 的公因數？<br>(A)2 (B)3 (C)5 (D)9<br><br>2.（　）下列何者是 27 和 33 的公倍數？<br>(A)207 (B)237 (C)267 (D)297 |
| 3.（　）$2\frac{1}{2}\times1+2\frac{1}{2}\times2=$？<br>(A) $2\frac{1}{2}$ (B)5 (C) $7\frac{1}{2}$ (D)10<br><br>4.（　）$2\frac{1}{2}\times1\times1+2\frac{1}{2}\times2\times2+2\frac{1}{2}\times3\times3=$？<br>(A)40 (B)35 (C)30 (D)25 | 5.（　）算式「32÷(4-2)x(3+1)」的答案是多少？<br>(A)64 (B)32 (C)16 (D)3<br><br>6.（　）算式「32÷4-2x3+1」的答案是多少？<br>(A)64 (B)32 (C)16 (D)3 |
| 7.（　）有一長方形牆壁，用邊長 30 公分或 40 公分的正方形瓷磚，都正好鋪滿，牆壁長的大約 4 公尺，這面牆的長度是多少公分？<br>(A) 460 (B) 480<br>(C) 520 (D) 540<br>8.（　）若牆壁的長是寬的 2 倍，則牆壁的面積是多少平方公尺？<br>(A) 10.8 (B) 11.2<br>(C) 11.5 (D) 11.52 | 9.（　）哥哥在加油站打工每 2 小時賺 170 元，每天工作 8 小時，則每天可賺多少元？<br>(A)630 (B)680 (C)730 (D)780<br>10.（　）承上題，哥哥要工作幾天才能賺到 20400 元？<br>(A)18 (B)22 (C)26 (D)30 |
| 11.（　）精緻餅乾禮盒，曉華吃了 $\frac{1}{3}$ 盒，則還剩下幾盒？<br>(A) $\frac{1}{4}$ (B) $\frac{1}{3}$ (C) $\frac{2}{3}$ (D) $\frac{3}{4}$<br><br>12.（　）同上題，若偉偉又吃了 $\frac{1}{4}$ 盒， | 13.（　）爸爸要買一隻原價 3000 元的手錶，爸爸付了 2400 元，是原價的多少？<br>(A)80% (B) $\frac{3}{4}$ (C)0.7 (D)百分之 65<br><br>14.（　）原價 2500 元的項鍊，媽媽以 1625 元買到，這項鍊打了幾折？<br>(A)6 折 (B)65 折 (C)7 折 (D)75 折 |

| | |
|---|---|
| 小明又吃了$\frac{1}{6}$盒，則禮盒還剩下幾盒餅乾？<br>(A)$\frac{1}{4}$ (B)$\frac{1}{3}$ (C)$\frac{2}{3}$ (D)$\frac{3}{4}$ | |
| 15.（　）實心圓柱體，底面的圓半徑為 5 公分則底面的圓面積是多少平方公分？<br>(A)78.5 (B)62.8 (C)31.4 (D)25<br><br>16.(　)同上題，若圓柱體的高為 5 公分，則圓柱體表面積為多少平方公分？<br>(A)78.5 (B)157<br>(C)314 (D)471 | 17.（　）每瓶汽水容量 1.25 公升，小萍買了 40 瓶，共有幾公升的汽水？<br>(A)5000 (B)500 (C)50 (D)5<br><br>18.(　)承上題，小萍共買了幾公秉的汽水？<br>(A)0.005 (B)0.05 (C)0.5 (D)5 |
| 19.（　）5 月有 31 天，共是幾個禮拜又幾天？<br>(A)6 個禮拜又 1 天 (B)5 個禮拜又 3 天(C)4 個禮拜又 5 天 (D)4 個禮拜又 3 天<br><br>20.（　）西元 2007 年 5 月 1 日星期二，西元2007年的6月1日是星期幾？<br>(A)一 (B)三 (C)五 (D)日 | 測驗結束<br><br>感謝您的作答，別忘了答案卷上的基本資料要填寫完整再交卷唷！ |

# 參加國際會議心得報告

報告人姓名：劉湘川

會議名稱：INFORMATION SCIENCES 2007, the 10th joint conferences,

會議期間及地點：July. 21-24, 2007, Salt Lake City, Utah, USA.

發表論文題目：

1. **Hsiang-Chuan Liu**, Wen-Chih Lin, Wei-Sheng Weng (2007). A Choquet Integral Regression Model Based on a New Fuzzy Measure, INFORMATION SCIENCES 2007, Proceedings of the 10th joint conferences, Salt Lake City, Utah, USA. 18-24 July, 2007, Page(s): 1349-1355 (EI 級論文)

2. Hsiang-Chuan Liu, Jeng-Ming Yih, Shin-Wu Liu (2007) Fuzzy c-mean algorithm based on Mahalanobis distances and better initial values, INFORMATION SCIENCES 2007, Proceedings of the 10th joint conferences, Salt Lake City, Utah, USA. 18-24 July, 2007, Page(s): 1398-1404 (EI 級論文)

3. Hsiang-Chuan Liu, Der--Bang Wu, Hsiu-lan Ma (2007,7), Fuzzy clustering with new separable criterion, INFORMATION SCIENCES 2007, Proceedings of the 10th joint conferences, Salt Lake City, Utah, USA. 18-24 July, 2007, Page(s): 1405-1411 (EI 級論文)

與會心得：

感謝國科會補助參加此次會議，主辦單很用心在籌劃這次的會議，充分讓人感覺到他們的熱情。

此次會議個人計發表3篇論文。均為EI級論文，所發展之三種理論模式均以教育測驗資料為應用實例‧ 且在會議結束前，獲得JOURNAL OF MULTIPLE-VALUED LOGIC AND SOFT COMPUTING“Fuzzy Theory and Technology” Guest-edited by Prof. Cengiz Kahraman之邀稿，此期刊為SCI-E級，個人已投出改進擴張論文：「 **Hsiang-Chuan Liu**, Jeng-Ming Yih, Shin-Wu Liu, Der--Bang Wu (2007) Fuzzy c-mean algorithm based on adaptive Mahalanobis distances」。

參加此會議 7 天，與世界各國研究與技術開發專業人才進行交流，並從展覽會中收集到一些與研究相關資料，獲益良多，亦看到了他國學者的在會議中互動的方式和技巧，也了解到自己不足及可以改進的地方。

攜回資料名稱及內容：

會議論文初稿全集 CD，會議議程冊，及未來相關國際會議的 Call for Paper 海報。