

應用灰關聯分析於電腦選題之選題策略研究

A Study on Adaptive Grey Relation Analysis for Computer Examination Items Selection

黃永發 顏碧玲 林坤緯 許淑玲
朝陽科技大學 朝陽科技大學 朝陽科技大學 朝陽科技大學
資訊與通訊系 資訊工程系 資訊工程系 資訊工程系
yfahuang@cyut.edu.tw beling7617@gmail.com kwlin@cyut.edu.tw grebeyen@ms4.kn
tech.com.tw

摘要

本研究根據測驗理論之訊息量 (Information)，應用灰色系統理論 (Grey system theory) 之灰關聯分析 (Grey relational analysis) 作選題策略，以五個能力等級訂出目標訊息量，選出符合施測者之施測目標訊息量的測驗題目，由選題之模擬結果可知，本研究提出之修正型灰關聯分析法之訊息均方差 (Mean square error, MSE) 甚低，比起目標訊息平均法 (Average target information, ATI) 及隨機選題法，灰關聯分析法均較好。

關鍵詞：灰色系統理論，灰關聯分析，目標訊息量，選題策略，均方差。

Abstract

In this study, we based on test theory information to select test items by measuring the amount of target tests questions. This study is based on the Grey relational analysis in Grey system theory to perform tests selection strategy. There are five ability levels being set out the target information. From simulation results, it is shown that the proposed modified Grey relational analysis method outperforms both the average target information (ATI) and random methods in mean square error (MSE).

Keywords: Grey system theory, Grey relational analysis, Target information, Mean square error, Item selection strategies

1. 前言

不管在任何之學習機構的教學過程中，為了確認學生的學習成效，須根據適當的選題及

量化分析，發展出一份能適合各個程度的測驗題目，來評估學習成效，是非常重要的第一步，尤其在中小學之基礎課程之學習中，除了升學時之學力測驗，以作為其基礎學科能力之學習成效評估外，每個學期或每個學季也須以測驗題目評量學生之學習成效，而測驗題目的設計則是達成正確評估的重要因素之一 [4-7, 10]。因此，為使學生在學科測驗時，能快速有效地由題庫中找到適當之測驗題目，由電腦來做自動選題是數位科技時代之趨勢，因此有關電腦選題之相關研究不論在數量及質量均大量成長 [4-6]，其中有部分結合網路技術，應用於個人電腦化適性測驗 [5]。然而，目前國內之各項考試中，團體性之紙筆測驗仍然是被大量採用，例如：多元入學之學力測驗或國中會考即是發展測驗型題庫之實例，因此如何挑選考試測驗之試題來符合施測者之特殊目的，是個很重要研究課題。

灰色系統理論 (Grey system theory) 是由大陸學者鄧聚龍於 1982 年所提出 [1,2]，發展應用至今，已有許多不錯的研究成果 [3,4]；其中灰關聯分析 (Grey relational analysis) 主要是透過參數間關聯性，由部分已知不明確條件找出所需要之訊息，進而明瞭參數間之互動關係。而灰關聯分析主要就是從少資訊，即是少數數據且不確定性出發，以多維角度來分析量化，以及序列化之種種關係，灰關聯分析是有參考系之整體比較，而一般一對一之比較，或是兩者之間的比較，都是無參考系之比較，無參考系之比較其缺點為：由於沒有參考對象，因而忽略了比較環境，容易產生錯覺，產生誤解。而灰關聯分析之主要用於找出邊界予以畫清，鑑別效能，辨認模式，確認同構，分析主次。而灰關聯分析著重數量化之整體比較，也就是有測度之比較，相較於距離空間之特性為

有測度但是無整體性。

因此，本研究應用灰關聯分析作試題參數選題策略來，其能快速地並有效地降低誤差量，找出適當之測驗題目，使測驗題目更能符合施測者。

傳統常用之選題法有：(1)標準法(Standard)，(2)任意選取法(Random)，(3)中間難度法(Middle difficulty)，(4)上下法(Up and down)，(5)最大訊息法(Maximum Information)等[6]，其共同之優點為執行效率相當高，而缺點為選題方式簡單而缺乏彈性，以致於所選出之試題測驗訊息函數與目標訊息函數之間的誤差均偏高。傳統選題方法為任意選取法[6]：使用亂數表，由題庫中任意選取 30 題。近年來由於電腦選題法可以有效減少誤差，其中主要的研究結果如孫光天教授等[9]於 1999 所發表之使用貪婪演算法作為選題策略，每加入一題，就計算其修正量是否為負值，如為負值表示測驗訊息函數逐漸逼近目標訊息函數。其優點為與傳統選題方法比較，其測驗訊息函數與目標訊息函數之誤差值小，但其計算量太大。為了改善傳統選題誤差量過大之缺點。本研究提出灰關聯分析，做為選題策略，期能有效減少誤差。

2. 灰關聯分析

灰關聯分析乃是灰色系統理論，對於離散數列之間相關程度的一種測量方式，它與以往因素分析所採取之方法不同，傳統之因素分析如統計迴歸分析(Regression) 需要大量數據，而且要求數據之分佈必須為典型分佈如：線性的、對數的、或是常態分佈(Normal distribution)等。而灰關聯分析具有少數數據及多因素分析之特性，恰可以彌補統計迴歸分析之缺點。

在序列可比較之狀況下，為達到灰關聯分析之目的，必須做數據處理，此一處理過程稱為灰關聯生成，而灰關聯生成之方式可以分為下列幾項：

1. 原始數據已滿足要求。
2. 利用原始數據做數據之正規化。
3. 利用灰色理論方法其方法如下所述：

(1)望大型：希望目標愈大愈好。

$$x_i^*(k) = \frac{x_i^{(0)}(k) - \min x_i^{(0)}(k)}{\max x_i^{(0)}(k) - \min x_i^{(0)}(k)} \quad (1)$$

(2)望小型：希望目標愈小愈好。

$$x_i^*(k) = \frac{\max x_i^{(0)}(k) - x_i^{(0)}(k)}{\max x_i^{(0)}(k) - \min x_i^{(0)}(k)}, \quad (2)$$

(3)望目之型式：

$$x_i^*(k) = 1 - \frac{|x_i^{(0)}(k) - OB|}{\max\{\max x_i^{(0)}(k) - OB, OB - \min x_i^{(0)}(k)\}}, \quad (3)$$

其中 $\max x_i^{(0)}(k)$ 為 $x_i^{(0)}(k)$ 中之最大值， $\min x_i^{(0)}(k)$ 為 $x_i^{(0)}(k)$ 中之最小值， OB 為 $x_i^{(0)}(k)$ 中選定之目標值。

在灰關聯空間中量化的測度公式，稱為灰關聯度。而在求灰關聯度時，如果只取一個序列 $x_0(k)$ 為參考序列時，我們稱為局部灰關聯度。如果其中任何一個序列 $x_i(k)$ 均可為參考序列時，我們稱為整體性灰關聯度，而灰關聯度的定義是表示兩個序列間的關聯度。現在我們將其相關理論敘述如下：在灰關聯空間中有一序列

$$x_i(k) = (x_i(1), x_i(2), \dots, x_i(k), \dots, x_i(K)) \quad (4)$$

其中 $i=0, \dots, M, k=1, \dots, K$ ，亦即

$$x_0(k) = (x_0(1), x_0(2), \dots, x_0(k), \dots, x_0(K)) \quad (5)$$

$$x_1(k) = (x_1(1), x_1(2), \dots, x_1(k), \dots, x_1(K))$$

...

$$x_M(k) = (x_M(1), x_M(2), \dots, x_M(k), \dots, x_M(K))$$

灰關聯係數為

$$\gamma_{0i} = \frac{\Delta_{\min} + \zeta \cdot \Delta_{\max}}{\Delta_{0i}(k) + \zeta \cdot \Delta_{\max}} \quad (6)$$

其中(1) $i=1, \dots, M, k=1, \dots, K, j \neq i$ 。

(2) $x_0(k)$ 為參考序列， $x_i(k)$ 為一特定之比較序列。

(3) $\Delta_{0i}(k) = |x_0(k) - x_i(k)|$ 為第 i 列數列與參考數列之差的絕對值。

(4) $\Delta_{\min} = \min_{\forall j \in i} |x_0(k) - x_j(k)|$ 為所有相減數列之中最小值。

(5) $\Delta_{\max} = \max_{\forall j \in i} |x_0(k) - x_j(k)|$ 為所有相減數列之中最大值。

(6) ζ : 辨識係數 $\zeta \in (0, 1]$ ，其值可依實際需要做調整。

辨識係數 ζ 的大小由實際的數學驗證中可以知道只會改變相對數值之大小，而對其灰關聯的排序沒有影響[11,12]。因此本研究將 ζ 的值設定為 0.5。

當取得灰關聯係數以後，由於灰關聯係數

的數很多，訊息過於分散，不便於比較，因此求關聯係數之和或是平均值，稱之為灰關聯度，表示為

$$\Gamma_i = \sum_{k=1}^K \gamma_{0i}(k) \quad (7)$$

修正型灰關聯係數為

$$\gamma_{0i}(k) = \frac{\Delta_{\min} + \zeta \cdot \Delta_{\max}}{[\Delta_{0i}(k)]^\lambda + \zeta \cdot \Delta_{\max}} \quad (8)$$

其中 λ 為調整因子。

3. 試題反應理論

關於測驗理論大概可以分為古典測驗理論(Classical test theory)以及試題反應理論(item response theory) [6]。古典測驗理論是所有測驗理論中，最早被使用來建立資料間實證關係的測驗理論。其主要是在估計一個測驗之實得分數之信度，也就是說它想要評估真實分數與實得分數之間的關聯程度。

試題反應理論針對古典測驗理論之缺失，而有以下之修正與特點：

- (1) 當代測驗理論試題參數之估計，不受測試者之樣本而影響，也就是與受測者無關。
- (2) 對於能力不同之人，給予不同之測量誤差指標，因此能夠精確估計受測者之能力值。
- (3) 對於能力不同之人，除了能夠提供較精確估計值之外，針對原始得分相同之測試者，也提供不同之能力估計。
- (4) 當代測驗理論對於同質性之試題，能夠測量估計出受試者之能力值，而不受測驗之影響。
- (5) 當代測驗理論以試題訊息(item information)及試卷訊息量，作為評定整體試卷，及試題之準確性。

試題反應理論中用以達成施測者施測之目試題編製，其最重要之工具為試題訊息量(item information)，以下將對訊息函數做介紹。

訊息函數為試題反應理論中提出用來描述試題在不同能力值下所產生之訊息量，藉此訊息量來挑選測驗試題以及比較測驗的相對效能，並作為建立分析與診斷測驗的主要參考依據，試題訊息函數的定義如下：

$$I_i(\theta) = \frac{(1.7a_i)^2(1-a_i)}{\left[c_i + e^{1.7a_i(\theta-b_i)} \right] \left[1 + e^{-1.7a_i(\theta-b_i)} \right]^2} \quad (9)$$

$I_i(\theta)$:代表試題*i*對具能力值 θ 之受試者所提供之

訊息。

a_i :鑑別度，範圍介於0與2之間。

b_i :難易度，通常只取 ± 2 之間。

c_i :猜測度。

由上式可知，當難易度 b_i 與能力值 θ 愈接近時，訊息量會越大。當鑑別度 a_i 較高時訊息量也會較大，當猜測度 c_i 接近0時，訊息量會增加[9,10]。

4. 測驗題目選題問題

以試題反應理論為基礎所建構的測驗中，如何由題庫進行選題組成一份符合測驗設計者施測目的之試卷，是測驗理論中相當重要一項。本研究利用訊息函數，來選取施測者所希望達成之測驗目的之試題。其編製過程之步驟由學者Lord於1977提出[8]：

步驟1：為設定施測者所要的訊息函數之分配，稱之為目標訊息函數(Target information function)。

步驟2：由題庫中選出一試題，使所選出試題之訊息量總和。能夠盡量符合目標訊息函數。

步驟3：每選入一試題，要重新計算所選入試題之訊息量。

步驟4：重複上述之步驟直到測驗訊息函數，及目標訊息函數之誤差量達到施測者設定之目標為止。

以訊息函數而言，每一題目對不同能力值 θ 均反應出不同之訊息量 $I_i(\theta)$ ，而此訊息量將可用來做選題時之參考，當訊息量越大，表示此題目對受試者能力估計越具參考價值，而一測驗之訊息量，為測驗中每一題目累加求得，因此針對不同施測需求，可以由題庫中選取不同題目，藉以達成測驗設計者的目標。例如以傳統之選題方法來選30題[7]，建構一目標函數曲線，最大峰值在能力為0。

不管用任何方法，我們均對每一能力等級 θ ，訂出一目標訊息量 D_θ ，當以某方法選出題目後，可計算出選出題目所形成測驗之訊息量為 O_θ ，兩者將有一誤差，我們將此均方差(Mean square error, MSE)定義為

$$MSE = \sum_{\theta=1}^K (D_\theta - O_\theta)^2 \quad (10)$$

其中 K 為能力總等級數。

4.1 試題選題策略

設題庫 \mathbf{T} 為試題 t_i 之集合， $\mathbf{T} = \{t_1, t_2, \dots, t_M\}$ ， M 為題庫之試題數，假設我們有一目標試卷

$\mathbf{S}, \mathbf{S} \subseteq \mathbf{T}, \mathbf{S} = \{s_1, s_2, \dots, s_i, \dots, s_N\}$, s_i 為被選入目標試卷的試題, N 為目標試卷的試題數。因此對於題庫, 試題 t_i 在每一能力等級 θ 的訊息值表示為

$$t_i(\theta) = \frac{(1.7a_i)^2(1-c_i)}{\left[c_i + e^{1.7a_i(\theta-b_i)} \right] \left[1 + e^{-1.7a_i(\theta-b_i)} \right]}, \quad (11)$$

其中 $1 \leq i \leq M$, $-r \leq \theta \leq r$, r 為一正數, $t_i \in \mathbf{T}$ 。而對於目標試卷, 第 i 次選入目標試卷之試題 s_i 的訊息值為

$$s_i(\theta) = \frac{(1.7a_i)^2(1-c_i)}{\left[c_i + e^{1.7a_i(\theta-b_i)} \right] \left[1 + e^{-1.7a_i(\theta-b_i)} \right]}, \quad (12)$$

其中 $1 \leq i \leq N$, $s_i \in \mathbf{S}$, $-r \leq \theta \leq r$, $\theta \in Z$ 。

本研究中, 先蒐集題庫集合 \mathbf{T} , 訂定各能力等級之目標訊息值 D_θ ($-r \leq \theta \leq r$), 並設定總選題數為 N 。選試題數, 設初始狀態為 $N' = N$, $i = 1$, 目標試卷題數為零, 即 $\mathbf{S} = \{\}$ 。

4.2 目標訊息值平均法(Average Target Information, ATI)

本選題法乃以逐題選出之方式, 即每一題之選出均經由下列 4 個步驟選出, 先定義選到第 i 題後之平均訊息量為

$$d_\theta^i = \frac{\left| D_\theta - \sum_{j=1}^i s_j(\theta) \right|}{N'}, \quad (13)$$

其中 $1 \leq i \leq N$, $-r \leq \theta \leq r$, 因此, 當進行第 i 題之選題時, 各步驟之進行程序如下之說明:
步驟 1: 求選到第 i 題後之平均目標訊息量 d_θ^i , (13), 當要選第 1 題時, 其平均目標訊息量為

$$d_\theta^1 = \frac{D_\theta}{N}, \quad (14)$$

其中 $-r \leq \theta \leq r$, θ 為學生能力等級, r 為正數, θ 為整數。

步驟 2: 求題庫中每一試題與平均目標訊息值在各能力值差之絕對值和, 及

$$\Delta E_j = \sum_{\theta=-r}^r \left| d_\theta^1 - t_j(\theta) \right|, \quad (15)$$

其中 $1 \leq j \leq M$, 選取(15)中最小值之試題, 即

$$l = \underset{j}{\operatorname{arg\,min}} (\Delta E_j) \quad (16)$$

步驟 3: 進行選取動作, 將(16)所選取之試題選

入目標試卷, 即

$$s_i = t_l, \quad (17)$$

再將選出之試題移出題庫, 並設定 i 增加 1, 及 N' 減少 1。

步驟 4: 判斷是否停止選取。如果 $N' > 0$, 則回到步驟 1。如果 $N' = 0$ 則停止, 並輸出選出之試題 \mathbf{S} 。

4.3 隨機選取法

本選題法也以逐題選出之方式, 即每一題之選出均經由下列 3 個步驟選出, 但本方法並不考慮其試題訊息量為何, 是最簡單之方式, 但其選取之試題總訊息量將趨近於題庫 \mathbf{T} 之平均訊息量, 即

$$O_\theta^R = \frac{\sum_{j=1}^M t_j(\theta)}{M}. \quad (18)$$

開始選題前, 先選定題庫集合 \mathbf{T} , 再設定總選題數為 N , $i = 1$, 當選取第 i 題時, 進行下列步驟:

步驟 1: 取系統時間, 當亂數種子, 以均勻(Uniform)亂數選取到第 l 題, $1 \leq l \leq M$ 。將所選取之試題選入目標試卷, 選取之試題選入目標試卷, 即(17)之 $s_i = t_l$, 再將選出之試題移出題庫, 並設定 i 增加 1。

步驟 2: 若未達到選題數, $i < N$, 則執行步驟 1, 否則停止。

步驟 3: 輸出目標試卷 \mathbf{S} 。

4.4 灰關聯試題選題法

將目標訊息值設 D_θ 為參考序列 $x_0(k) = (x_0(1), x_0(2), \dots, x_0(k), \dots, x_0(K))$, 題庫訊息值 $\mathbf{T}(\theta) = \{t_1(\theta), t_2(\theta), \dots, t_M(\theta)\}$ 設為其他序列 $\mathbf{X}(k) = \{x_i(k)\}, i = 1, \dots, M$ 。

進行第 i 題選題步驟如下:

步驟 1: 求灰關聯係數之大小, 由(8)之修正灰關聯係數得第 i 題之灰關聯係數為

$$\gamma_i = \frac{\Delta_{\min} + \zeta \cdot \Delta_{\max}}{[\Delta_i(k)]^{1.5} + \zeta \cdot \Delta_{\max}}, \quad (19)$$

其中 $i = 1, \dots, M, k = 1, \dots, K$, 辨識係數 $\zeta = 0.5$, $\Delta_i(k) = |x_0(k) - x_i(k)|$, $\Delta_{\min} = \min_{\forall j \in I} |x_0(k) - x_j(k)|$,

$\Delta_{\max} = \max_{\forall j \in I} |x_0(k) - x_j(k)|$ 。

步驟 2: 對試題 t_i 之訊息數列 x_i , 求其灰關聯

係數之和，得各試題 t_i 之灰關聯度

$$\Gamma_i = \sum_{k=1}^K \gamma_i(k), 1 \leq i \leq M, \quad (20)$$

選取灰關聯度最大者為選題試卷中之試題，即

$$l = \underset{j}{\operatorname{arg\,max}} (\Gamma_j) \quad (21)$$

步驟 3：進行選取動作，將(21)所選取之試題選入目標試卷，即

$$s_i = t_l, \quad (22)$$

再將選出之試題移出題庫，題庫總題數 M 減 1，並設定 i 增加 1。

步驟 4：第 i 次灰關聯選題後，若 $i < N$ ，則設定新的參考數列，我們以目標值和選題之試卷訊息值在各個能力等級之差值，定義新的目標訊息量為

$$D_\theta^i = D_\theta - s_i(\theta) \quad (23)$$

其中 s_i 為選取之第 i 題，在選取第 $i+1$ 題時，參考序列 $x_0(k) = \{x_0(1), x_0(2), \dots, x_0(k), \dots, x_0(K)\}$ 更新為 D_θ^i ，再執行步驟 1 之第 $i+1$ 題選題。

若 $i=N$ ，則停止，輸出目標試卷 S 。

5. 模擬結果

本實驗之題庫與孫光天教授於 1999 年所發表之選題策略研究相同[9]，為國民小學自然科電腦化適性測驗系統之題庫，題庫之總題數為 320 題。為了評估驗證灰色關聯之選題策略之效能，我們從 320 題中選取不同題數 $N=10, 15, 20, 25, 30$ 題，並訂定在不同能力值(θ)的目

標訊息量 $D_\theta (D_{-2}, D_{-1}, D_0, D_1, D_2)$ 來探討，目標訊息值 $D_\theta (-2 \leq \theta \leq 2)$ 訂為表 1 所示。

表 1 不同題數之目標訊息值 D_θ

目標值 D_θ 題數	能力等級(θ)				
	-2	-1	0	1	2
十題	1	2	3	3	2
十五題	1	3	4	4	2
二十題	2	4	6	5	3
二十五題	2	4	7	6	3
三十題	2	5	8	7	4

我們用 Matlab 程式語言模擬電腦選題的步驟，以比較灰關聯度、目標訊息值平均法與隨機選取法之選題效能。

圖 1 中顯示選取不同題數 $N=10, 20, 30$ 題時，三種方法所選出試題與在不同能力值(θ)的目標訊息量 D_θ 之比較圖，由圖 1 可看出，隨機選取法(Random)選出試題之訊息量大多是最底的，而灰關聯度(GR)法的選題結果，則雖然有時超過目標訊息量，但與目標訊息值平均法(ATI)、隨機選取法比較，大多較接近目標訊息值，因此灰關聯度(GR)在不同能力值都有較好之效果。

我們進一步計算(10)之 MSE，得圖 2 所示之三種方法在 $N=10-30$ 題之 MSE 比較，由圖 2 可看出，除了 $N=15$ 時，灰關聯度(GR)法與 ATI 相等，在其他 N 值，灰關聯度(GR)法均優於 ATI 法，而且，灰關聯度(GR)法與 ATI 法應當比隨機選取法好，不過在 $N=25$ 時，ATI 法比隨機選取法差，這應該是此題庫之特殊情形造成。

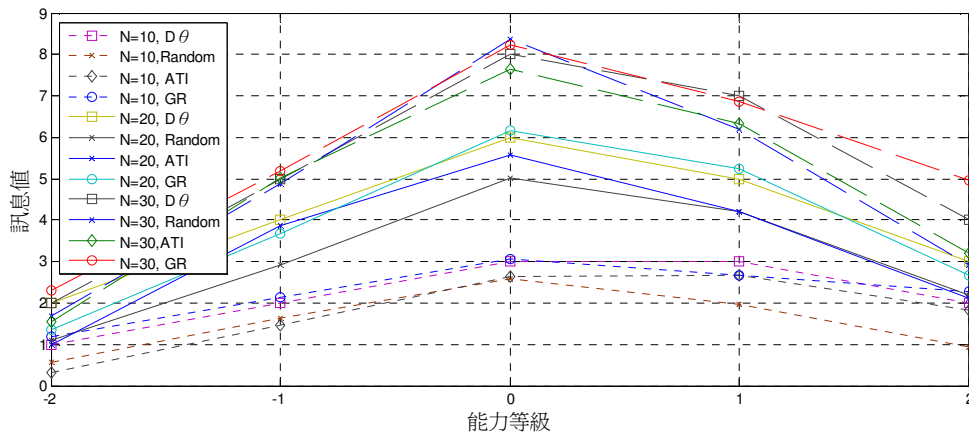


圖 1 不同選題策略在各能力等級之訊息量比較

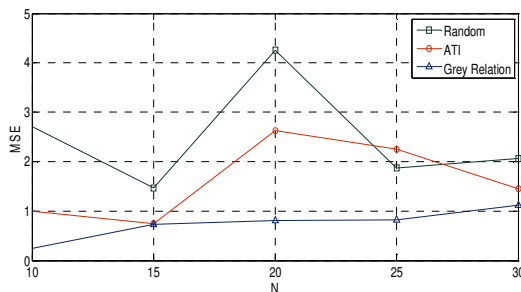


圖 2 不同選題策略在 MSE 訊息量誤差之比較

6. 結論

本研究根據測驗理論之訊息量 (Information)，選出符合施測者之施測目標訊息量的測驗題目，本研究在 320 題之題庫中，分別選出 10,15,20,25,30 題，並以五個能力等級訂出目標訊息量，應用灰關聯分析 (Grey relational analysis) 作選題策略，由選題之模擬結果可知，本研究提出之修正型灰關聯分析法之訊息均方差 (Mean square error, MSE) 甚低，比起目標訊息平均法 (Average target information, ATI) 及隨機選題法，灰關聯分析法均較好。

誌謝

本研究成果為國科會所贊助，計畫編號 NSC 101-2221-E-324-024-及 NSC 99-2632-E-324-001-MY3。

參考文獻

- [1] 鄧聚龍，*灰色系統基本方法*，華中理工大學出版社，1996。
- [2] 鄧聚龍，*灰色系統理論教程*，華中理工大學出版社，1992。
- [3] 江金山、吳佩玲、蔣祥第、張廷政、詹福賜、張軒庭、溫坤禮合著，*灰色理論入門*，高立圖書有限公司，1998。
- [4] 何榮桂、杜玲均、莊謙本，”改良式之灰色預測電腦化適性測驗選題測驗”，*第七屆國際電腦輔助教學研討會*，pp. 393~400，1998。
- [5] 杜淑芬，黃國禎，”網路測驗及評估系統試題配置最佳化之研究”，*第七屆國際電腦輔助教學研討會*，pp. 337~343，1998。
- [6] 余民寧，”電腦化適性測驗”，*研習資訊 10 卷(5 期)*，5-9 頁教育測驗出版社。
- [7] 林秋先、林億芳，”利用灰色預測於網路適

性教學時程研究”，*第四屆灰色系統理論與應用研討會*，pp.170-173，1999。

- [8] Lord, F. M. *Applications of item response theory to practical testing problems*, Hillsdale, HJ: Erlbaum, 1980.
- [9] 孫光天，陳岳宏，賴膺守，謝凱隆，陳新豐，”使用貪婪演算法辦作為一有效益之選題策略”，*八十八年全國計算機會議*，pp. A379-386，淡江大學，1999。
- [10] 黃家輝，”利用灰色理論於選題策略之研究”，*義守大學，碩士論文*，高雄，2001 年。
- [11] Wen, K.-L., Wu J. H., “On Distinguishing Coefficient & Relational Grade,” *The Journal of Grey System*, Vol. 8, No. 1, pp. 11-18, 1996.
- [12] Wong, C.C. and Lai, H.R., “A New Grey Relational Measurement,” *The Journal of Grey System*, Vol. 12, No. 4, pp. 341-346, 2000.