

# 電腦與統計方法在教育研究上之配合運用

簡 茂 發

主席、各位教育界的前輩和同仁：現在我僅就有關「電腦在教育研究上之運用」第二部份：「電腦與統計方法在教育研究上之配合運用」加以說明。關於我要報告的大綱已印在研討會手冊第八頁。另外請各位從資料袋裡拿出一份資料，這是兩個教育研究的實例，用以說明「多元迴歸分析法」之運用途徑。

## 一、電腦提供教育研究資料處理之服務

現代教育強調以科學的方法，推動研究發展工作。因此，相當重視「實徵性研究」( empirical approach )。進行「實徵性研究」時，通常是用調查、測驗或實驗的科學方法，來搜集有關教育活動數據資料。這些數據的資料，我們可以應用統計的方法加以分析，而把複雜的問題和教育現象化成有系統、有意義的研究結果。如果教育研究所涉及的變數相當複雜，或接受調查、測驗、實驗的人數眾多，那麼資料處理( data processing )是一件非常繁重的工作。這時如果我們能夠利用電腦這種科學的工具進行統計分析，那麼電腦可以提供很有效的服務，使統計工作得以順利進行。

至於談到利用電腦來作教育研究資料的統計分析，我們首先要說明電腦在分析資料方面的一些優點。一般而言，利用電腦處理資料的優點有三個：快速( speed )、正確( accuracy )、經濟( economy )。電腦具有四個特點，現在分別說明如下：

1. 敏捷之操作：使用電腦來處理資料，電腦的操作效率很高，它可以非常敏捷地進行分類、比較、演算等工作。
2. 廣大之儲量：在電腦裡，我們可以把研究所收集的各種資料儲存進去，蓋因電腦有廣大的儲存量。把資料儲存進去之後，需要使用時，可利用各種指令方式，調出資料加以統計分析。
3. 高度之可靠性：電腦有高度的可靠性，電腦固然也有可能發生錯誤，不過通常是由於人為的疏忽所致。我們在使用電腦進行資料分析時，發生錯誤的機會應當比較少。在我們設計電腦程式的時候，如作適當的安排，那麼在處理資料的過程中，有一些不合理的資料或不合理現象產生，電腦有時候會停止操作，有時會把一些可能錯誤的地方告訴我們，提醒我們必須再查對這些資料，探究什麼地方有了問題而加以改進。
4. 邏輯及數學之運用：電腦是根據人腦的原理加以設計的。使用電腦時，只要在程式上給予適當指令，它可以作邏輯及數學上的運用和推演，或按照電腦程式的指示，進行分類、比較、轉移等工作，或者擔負非常繁複之統計分析的任務。

另外可以發現電腦和人類不一樣的地方，即是電腦不會厭煩、疲倦，經常保持最高的、最佳的工作狀態，提供最忠實的服務。研究者如果能用電腦來分析資料，通常可使我們所收集的資料，作最充分的

運用，進行深入的統計分析，以獲致最大的研究結果。這是使用電腦分析資料的好處。

不過，我們也應該瞭解電腦有它的限制。電腦畢竟是沒有生命的機器，只能仗機械式的操作，完全聽命於人，聽命於我們給它的各種程式上面的指示。因此，我們用電腦分析資料，當然必須把研究資料化成電腦語言的形式，後才能按照我們的指令行事。另外，如果遭到一些問題，電腦不像人腦能夠隨機應變或作具有創造性的調整；電腦本身也不能夠發展自己、改變自己。電腦有優點，也有缺點，因此我們人腦與電腦在研究上應該作適當的配合運用，兩者就如同主僕的關係一樣，如果能夠適當配合，那麼在研究上一定可以相得益彰。

## 二、電腦促進統計方法之革新及發展

由於電腦的發明和推廣應用，使統計學的內容和方法不斷地革新進步，對於教育研究的設計與實施產生不少影響。在這方面可歸納成數點，加以簡單的說明。

由於現在有了電腦作為我們處理資料的科學工具，因此在統計方法上，我們比較強調下列的幾個發展趨勢：

1. 擴大樣本的人數，再就樣本性質細分組別，以減少抽樣誤差，增加統計推論的正確性：在教育研究上電腦可以使我們增加樣本的人數，增加統計分析結果應用在推論方面的正確性。譬如研究學生的課業負擔是否太重，過去沒有電腦，也許我們不管大學生、中學生或小學生，通通包括在研究對象中，以探討其課業負擔的情形。而現在我有了電腦，研究人數可以多一些，同時根據大、中、小學生予以分組，另外可以再按性別、年級……等許多分類因素，把整個樣本配合研究上的需要予以細分，因為這樣細分成比較小的組別，然後再行抽樣，也許誤差就會比較少些。
2. 建立多變數分析的數理模式，發展各種多元變數的統計方法：顯然我們可以看到現在統計上有一個重要的發展趨勢，強調多變數分析方法的運用，譬如我們在測驗統計或研究上用到的因素分析、多元迴歸分析、典型相關分析、辨別分析、多元變異數分析、多元共變數分析等皆是。這些統計分析的方法，由於牽涉到很多繁複的計算過程，不容易用人工方法加以處理；若用人工處理，要花費很多時間、人力，所以在過去很難辦到。現在有電腦可用以處理繁複的資料，所以多變數分析乃被採用，可以說是相當普遍。
3. 採用「運算公式」替代「定義公式」，以節省電腦作業時間：一般統計的公式有兩種，一種是「定義公式」，另一種是「運算公式」。定義公式可以使我們知道某一種統計量數的意義是什麼，不過，定義公式在我們一般的計算上，如果要一把原始數據資料代進去，可能是相當的麻煩。如今利用電腦作為我們統計分析的工具，為了要節省電腦操作的時間，通常有一種趨勢，就是利用「運算公式」來代替「定義公式」。不過在此順便提及：如果電腦的運作單位準確的範圍有限的话，替代之後可能成一種削截的誤差，而減低了統計結果的準確性。在未能調整電腦運作單位的準確範圍之前，一個變通的方法就是採用「倍式準確性的浮點情況」以代替「單式準確性的浮點情況」，但由於電腦作業過程比較繁複，而相對地增加了電腦的時間。如果我們能將電腦運作單位準確性的範圍加以擴大，那麼這個問題就可以解決了。
4. 不經由傳統的歸類途徑而直接對原始數據資料進行統計處理，以避免分組誤差的影響：我們傳統的統計方法，在從前沒有計算機之時，要用人工加以演算，所以如果我們所得到的資料限多，通常須先把資料加以分整理，化繁為簡，變成次數分配表的形式，再做統計，這樣可能會比較簡單一些。不過把資料加以分類化成次數分配表，會造成分組上的誤差。現在由於使用電腦作為資料分析的工具，我們採取了直接對原始分數進行統計處理的方法，這樣可以避免分組誤差的影響

。但是這並不是說，次數分配表完全失去價值，如果我們想要看出數據資料分配的情形，也可以在電腦程式上要求把它打出來。例如：我們用直線相關的方法去分析雙變量資料，也許會得到相關等於零的結果，但是相關等於零，並不一定說它沒有相關，因為它可能不是直線相關而是曲線相關。至於如何發覺這項事實呢？也許我們可以把相關的兩個變項，用散佈圖的方式，把它呈現出來。

5. 重視報導各種統計假設檢定結果所犯第一種錯誤的機率，而更能推斷其顯著水準：在過去談到統計推論時，我們常常求出一些考驗顯著性的量數出來之後，再查某種抽樣分配的機率表，可以得到犯第一種錯誤的誤差是否小於 0.05 或小於 0.01。利用電腦分析，可以直接求出我們在作統計推論時，可能會犯錯的機率是多少，這是電腦在統計運用上一個很重要的發展。
6. 統計學教學之重點在於統計方法之選用和統計結果的解釋分析，而非計算過程之訓練：由於電腦已成為我們應用的一工具，我們並不過份強調繁複的統計運算的訓練，而注重教導學生如合適當地選用統計方法，以及如何解釋統計分析結果的意義。因此，現在統計學的教學，通常要和電腦密切配合，尤其是高級統計學牽涉到比較繁複的統計資料分析的過程，更必須藉助於電腦。我們在教學上的重點，並不是於計算的過程，而是強調統計方法的性質之瞭解及其選用的時機，以及統計分析結果之解釋與應用。

以上是說明電腦促進統計方法之革新及發展的情形。

### 三、利用電腦進行統計分析之途徑

我們知道電腦可以提供服務，幫助我們進行資料的分析。不過，這裡牽涉到一個實際的問題：首先考慮有無必要利用電腦進行統計分析，然後再考慮利用腦進行統計分析的可能性。此等問題又涉及：1. 變項的多少，2. 樣本的大小，3. 統計分析的可能性，4. 能否獲得電腦程式，5. 經費等因素。

1. 變項的多少：如果一項教育上的研究所包括的變項很多，則其資料統計分析的過程就比較繁複，就需使用電腦進行分析；如果研究所包含的變項有限，用人力即可算出，就沒有使用電腦分析的必要。
2. 樣本的大小：如果樣本有限而且變項又少，就不必用電腦分析；要是變數多，樣本又眾多，就需要利用電腦分析研究資料。
3. 統計分析的複雜性：如果我們做統計分析只是求次數或百分比，那麼只要用人工方法就可以做到，不必用電腦來分析；但是如果進行多變數分析，雖然人力上可以做到，但不可能在短時間內做到，也很容易發生錯誤，在這種情形下，當然就必須用電腦作為我們統計分析的科學工具。
4. 能否獲得電腦程式：現在有一些電腦程式是現成的，不過由於各種電腦機種有所不同，未必能適用，有另行設計的必要。因此，我們要考慮到有沒有現成的電腦程式，或有沒有人員具備設計電腦程式的能力。
5. 經費：「錢」是個很實際的問題，我們必須考慮到。從整個節省人力、時間的觀點來看，利用電腦分析教育研究資料，花一筆錢還是值得的。我們要從經濟的原則來考慮這個問題，不要說任何統計分析都要用電腦，例如次數或百分比之類的簡單統計，如果也用電腦分析，就顯得大才小用了。

接著，談到利用電腦進行統計分析的程序，茲扼要說明如下：

1. 擬定研究計劃時，即須顧及電腦處理資料有關的事項：因電腦只能處理可量化的資料，不能量化的資料，事實上無法作統計分析。在進行研究設計之時，要把樣本加以適當的分組，編錄便於分

類的識別號碼，並考慮決定採用何種統計分析方法。

2. 資料的編碼、登錄、打卡，將資料轉化成電腦可以閱讀的形式：電腦的程式設計師或操作人員，應該懂得統計分析的方法，教育研究者也要瞭解電腦的性能，具備電腦的知識，如此兩者易於溝通，才能把我們的研究資料，用最適當的方法，加以最有效的分析。
  3. 決定分析的變項以及採用的統計方法，寫成分析指引，將統計分析步驟、統計公式及所需表格的形式，逐項明確列出，以便設計電腦程式或根據需要從現有的電腦程式中加以選擇。
  4. 試驗電腦程式之適用性：在電腦人員設計好某一電腦程式以後，要先試驗處理部份資料，看看電腦程式之適用性，核對其結果是否正確，必要時加以修改。
  5. 正式進行資料處理工作：希望能夠由電腦的協助，以獲致最佳且最要的統計分析之結果。
- 以上所述，是有關利用電腦進行統計分析之途徑。

#### 四、教育研究實例：多元迴歸分析法之應用

##### (一)「教育研究多元迴歸分析實例(一)」(見附件一)

這是從教育研究所林清江所長主持的最近完成之一份研究報告中有關多元迴歸分析的部份摘錄而來。林清江所長主持的專案是「國中畢業生升學與就業意願之影響因素」，其中用了很多統計方法進行資料之分析。現在我想提出來說明的是我們想利用研究上所收集到的各種資料，來預測影響學生升學或就業意願之最主要因素是什麼。在資料第一頁，分別把研究上所涉及的各種變項作了簡單說明。在這些變項中，是以所謂的第七個變項做為我們的標準變項，其他的都是當作預測變項。

在這裡，我們以「升學就業意願」為依變項，也叫做「標準變項」，分成一、二、三、四等四個層次分數，分數愈高表示就業的意願愈強，也就是分數愈低表示升學意願愈強。自變項又叫做「預測變項」，以  $X_1$  為例加以說明，這是表示學生的身份，離校者用「0」表示，在校者用「1」表示。所有的變項都必須用適當的方式加以量化，唯有量化資料，始能進行統計上的分析。至於這些變項之量化情形，可參看文字的說明。

第四頁提到逐步多元迴歸的方程式，由迴歸方式可以求出各個變項的迴歸係數。茲將逐步迴歸分析的方法及其基本原理說明如下：

首先求出研究上所涉及的這些變項之間的相關矩陣，第六頁所示便是根據調查結果所求出的相關矩陣，另外註明各變項分數的平均數和標準差。多元迴歸分析就是以相關分析為基礎的一種多變數分析的應用。

接著翻到下一頁，在第一個步驟中，是以一個最重要的預測變項和標準變項求出相關。最重要的影響因素就是第八個變項——「父親的教育程度」，其與標準變項的相關係數是 .31，相關係數的平方是 .097。第二步用迴歸的方法把第二個重要的變項找出來，發現第二個變項和第八個變項結合在一起，與標準變項的多元相關係數是 .41，它能夠解釋升學或就業意願總變異量的百分之十七。然後再尋找第三個因素，現在應加進去的是「母親的教育期望」，加進去之後，我們發現多元相關係數提高了，決定係數也顯著地提，高至於是否有顯著的提高，可以由「F」值看得出來。如此逐步找出重要因素加進去，一直加到第十五個因素「父母管教方法」，我們可以發現到第十五個加進去之後，決定係數有顯著的增加；加上第十六個因素之後，發現多元相關係數雖然也有所增加，但未達顯著水準，所以多元迴歸分析到此為止。

由上述實例，我們可以得知：就全體樣本來說，如果我們想推論它所代表的國中學生，其升學或就業意願會受到那些因素的影響，根據分析的結果找出來十五個因素，對國中學生的升學或就業意願具有

決定性的或重要的影響，其多元相關係數是 .56，決定係數是 .31，也就是我們根據這十五個因素可以解釋國中學生升學或就業意願的總變異量裡面的百分之卅一。

### (二)「北科大外國學生學業成就有關因素之研究」(見附件二)

第二個例子是我個人所作的研究。因為當時我在北科大進修，利用外國學生的一些檔案資料，使用多元迴歸分析法加以處理。各位可以看到這個實例和剛才第一個實例所用的方法原理相同，但處理程序剛好相反。剛才我們是將最重要的因素找出來，然後其餘的因素再一個個加進去，看看它相關係數有沒有顯著的增加。而現在第二個實例，各位可以看到它包括 18 個變數，第一個變數是外國學生的 GPR，即所謂「學業平均成績」，在這個研究中作為標準變項；另外從第 2 到 18 個變項作為預測變項。關於這些變項可按照邏輯性質加以分類，分成兩大類：一類是「智能的變數」，另一類是「非智能的變數」。「智能的變數」包括由第 2 個至第 12 個，其中又分成兩種：一種是利用標準化測驗測出來的一些心智的能力；另一種是有關課業的表現情形。至於「非智能的變數」也分成兩類：一種是身份的基本資料；另一種是所處環境的因素。

這個研究最主要的目的，就是想要找出影響外國學生的學業成績之最主要因素。分析的方法，首先是求出預測變項和標準變項的相關矩陣，然後以此為基礎加以分析。在第十頁，把多元迴歸分析的體系列成一個表，最上面有一個 Full Model (完全模式)，就是把所有的預測變項結合在一起，求其與第一個變項(即標準變項)之間的多元相關係數，這裡所列出的是它的決定係數為 .8730，亦即在外國學生學業成績的總變異量中，可以用上述十七個因素解釋其中的百分之八七.三。

其次，第二個步驟是應用「把某些變項去掉」的方式。請看左邊 Full Model 底下括號中減去第二至第十二變項，也就是把十七個變項裡的第 2 個到第 12 個變項去掉之後，看看其決定係數有沒有降下來。這 2 到 12 就是我們所謂的「智能因素」。根據研究的結果，由 .8730 降到 .4865，經統計顯著性的考驗，發現有顯著下降。再看右邊 Full Model 減掉第 13 至第 18 變項(即非智能因素)其決定係數下降得非常有限。由此可知：智能因素比較重要，所以要想知道影響外國學生學業成績的重要因素，就要從智能因素中去找。右邊畫了一條虛線，表示不必再往下找，應該由左邊去找才是。

左邊在智能因素底下，在 Full Model 中把標準化所測出來的能力因素(從 2 到 6)去掉。發現決定係數降得很大，從 .8730 降到 .6862。由此可知：在智能因素中，學校的課業比較重要。那麼，其次我們就應該由學校課業中的因素再去找最重要的影響因素。最後研究獲致的總結是：大學學業成績對這些外國學生在研究所進修的學業成績具有最重要的影響力。

以下兩種實例的說明，其所根據的原理和方法相同，但統計處理的程序不一樣。實例(一)在找出那些因素對國中學生升學或就業意願產生綜合的最有力之影響；實例(二)在於從許多的變項中找出那一個因素對北科大外國學生的學業成績具有最重要的影響力。

#### 附件：教育研究「多元迴歸分析」之實例

- 一、「國中畢業生升學與就業意願之影響因素」專題研究報告中之第四章第六節：影響國中學生升學與就業意願的綜合因素之探討。
- 二、「北科大外國學生學業成就有關因素之研究」。

附件一：

教育研究「多元迴歸分析」之實例（一）

「影響國中學生升學與就業意願之綜合因素探討」

上述各節曾分別探討影響國中學生升學與就業意願之個人因素、學校因素、與社會因素。所用的統計方法為卡方與 t 值檢定。為了更深入探討各項因素對於升學與就業意願之綜合影響，乃進一步使用「逐步多元迴歸分析」( Stepwise Multiple Regression Analysis ) 方法，以瞭解諸因素所能解釋及預測升學與就業意願之總變異的多寡，並進一步尋求足以影響國中學生升學與就業意願的一些重要因素，作為實施國中學生升學與就業輔導之參考。

本項統計分析在未進入電腦作業前，首先將每一自變項與依變項，加以數量化。茲將本研究之依變項與自變項之代號與量化情形說明如下：

1. x 7：升學與就業意願為依變項，分 1. 2. 3. 4. 四個層次分數，分數愈高表示就業意願愈強。
2. x 1：在校或離校，0 為離校，1 為在校。以下皆為自變項。
3. x 4：性別，0 為女，1 為男。
4. x 6：城鄉，0 為鄉村，1 為城市。
5. x 8：父親的職業，由 1 至 9 分，分數愈大代表教育程度愈高。
6. x 9：父親的職業，由 1 至 6 分，分數愈大代表職業等級愈高。
7. x 10：父親教育態度，由 1 至 21 分，分數愈大表示教育態度愈關心與積極。
8. x 11：父親的教育期望，由 1 至 9 分，分數愈大表示教育期望愈高。
9. x 12：父親的管教方式，由 1 至 6 分，分數愈大表示管教方式愈嚴格。
10. x 14：母親的教育態度，說明如  $\times 10$ 。
11. x 15：母親的教育期望，說明如  $\times 11$ 。
12. x 16：母親的管教方式，說明如  $\times 12$ 。
13. x 18：教師期望，由 1 至 4 分，分數愈大表示教師的期望愈高。
14. x 19：同輩團體，由 1 至 3 分，分數愈大表示愈受同學影響。
15. x 20：升學輔導，由 1 至 12 分，分數愈大表示學校行政措施愈注重升學輔導。
16. x 21：就業輔導，由 1 至 15 分，分數愈大表示學校行政措施愈注重就業輔導。
17. x 22：社會價值觀念，由 1 至 32 分，分數愈大表示升學取向的價值觀念愈濃厚。
18. x 24：成就動機，由 1 至 50 分，分數愈大表示成就動機愈強烈。
19. x 25：對自己身體特質的態度，由 1 至 20 分，分數愈大表示對自己身體特質愈感覺滿意。
20. x 26：對己能力的態度，由 1 至 20 分，分數愈大表示對自己能力愈有信心。
21. x 27：對自己人格特質的看法，由 1 至 20 分，分數愈大表示自己愈有良好的人格特質。
22. x 28：對外界環境接納的態度，由 1 至 20 分，分數愈大表示與他人環境之適應愈良好。
23. x 29：對自己價值與信念的態度，由 1 至 20 分，分數愈大表示愈相信努力會成功，好人有好報。
24. x 31：學業成就，分數愈大表示國中三年級學業總平均成績愈高。
25. x 32：智力，分數愈大表示智力愈高。
26. x 33：文字理解，分數愈大表示文字理解能力愈高。
27. x 34：知覺速度，分數愈大表示知覺速度愈快。

- 28. x 35：數字理解，分數愈大表示數字理解愈強。
- 29. x 36：空間知覺，分數愈大表示空間知覺能力愈高。

本研究所用的逐步多元迴歸方程式為：

$$Y = a + b_1x_1 + \dots + b_nx_n + e$$

Y 為依變項，即升學與就業意願

a 為常數

b1 ..... bn 為各項迴歸係數

x1 ..... xn 為自變項

e 為誤差

(三)全體樣本

表一是鄉村與城市樣本合併後之全體樣本各變項間之相關矩陣、平均數、與標準差。

表一中左邊第一行為依變項 ( X7 ) 與諸自變項的相關係數，其值介於 .06 與 - .31 間，大部份為負值，負值表示該變項分數愈高，則其升學意願也愈強。

茲將影響「升學與就業意願」之逐步多元迴歸分析結因列如表二。

表二顯示：在逐步多元迴歸分析過程中，第一步驟先投入 X8 ( 父親教育程度 )，其相關係數為 .31，決定係數為 .10，達到顯著水準。迨至十五步驟投入十五個變項後，其多元相關係數為 .56，決定係數為 .31 ( F = 5.23, P < .05 )。此後再投入其他變項，其多元相關係數已無顯著的增大趨勢 ( P < .05 )。因此本研究乃認定此十五個變項為影響國中學生升學與就業意願的重要因素。按其相對重要性依次為父親教育程度、社會價值觀念、母親教育期望、母親教育態度、教師期望、成就動機、對自己能力的態度、父親的職業、學校就業輔導、在校或離校、學校升學輔導、父親教育態度、同輩團體、性別、父親管教方式等。此等十五個變項足以解釋與預測國中學生升學與就業意願之總變異量的百分之卅一。





表二 影響國中學生全體樣本 ( N = 2670 ) 升學與就業意願  
之因素多元迴歸分析結果

步驟	投入變項順序	多元相關係數	決定係數	F	值
1	X8 ( 父親教育程度 )	.31196	.09732	287.64791**	
2	X22 ( 社會價值觀念 )	.40642	.16518	216.77301**	
3	X15 ( 母親教育期望 )	.46291	.21429	166.63477**	
4	X14 ( 母親教育態度 )	.49505	.24507	108.67458**	
5	X18 ( 教師期望 )	.51308	.26325	65.72968**	
6	X24 ( 成就動機 )	.52121	.27166	30.76468**	
7	X26 ( 對自己能力的態度 )	.53055	.28148	36.36952**	
8	X9 ( 父親的職業 )	.53657	.28790	24.01021**	
9	X21 ( 學校就業輔導 )	.54138	.29309	19.51503**	
10	X1 ( 在校或離校 )	.54623	.29836	19.98286**	
11	X20 ( 學校升學輔導 )	.55062	.30318	18.37837**	
12	X10 ( 父親教育態度 )	.55353	.30639	12.31014**	
13	X19 ( 同輩團體 )	.55562	.30871	8.89728**	
14	X4 ( 性別 )	.55771	.31104	8.99233**	
15	X12 ( 父親管教方式 )	.55893	.31240	5.23250**	

\* P &lt; .05

\*\* P &lt; .01

附件二：

教育研究「多元迴歸分析」之實例（二）

「北科大外國學生學業成就有關因素之研究」

FACTORS RELATED TO SCHOLASTIC  
ACHIEVEMENTS OF FOREIGN  
STUDENTS AT UNC, USA

Maw-fa Chien

I .Statement of the problem

It is generally expected that foreign students will encounter more difficulties than American students in adjusting to academic programs of the university in the United States. The purpose of this study was to determine whether the presently filed data about the individual foreign students at the University of Northern Colorado ( UNC ) were useful in predicting their scholastic achievements at the graduate level.

II .Description of the sample

The Ss in this study were 52 graduate students who came from foreign countries and were enrolled for the winter quarter of 1972 at UNC.

III .Collection of the data

The data about the Ss used in this study were obtained from Graduate School Office, The Office of International Education, and Counseling and Testing Center at UNC.

## IV. Descriptions of the variables

Table 1 : List of Variables

Number	Variable	Abbr.
1.	Cumulative Grade-Point Average ( Criterion )	GPAc
2.	WAIS Verbal IQ	IQv
3.	WAIS Performance IQ	IQp
4.	GRE Verbal Score	GREv
5.	GRE Quantitative Score	GREq
6.	TOEFL Score	TOEFL
7.	High School Grade-Point Average	GPAh
8.	Undergraduate Grade-Point Average	GPAu
9.	Cumulative Credits	CC
10.	Academic Load ( in average )	AL
11.	Graduate Academic Program	GAP
12.	Major	Maj
13.	Sex	Sex
14.	Age	Age
15.	Marital Status	MS
16.	Residential Pattern	RP
17.	Participation in Student-Club Activities	PSCA
18.	Means of Financial Support	MFS

For variables 11 through 18 except 14, the Ss were binary coded as follows :

GAP(11) : 1 for an S who was working for doctoral program ;  
0 for an S who was working for master's program.

Maj(12) : 1 for an S whose major was in the field of education ; 0 otherwise.

Sex(13) : 1 for a male ; 0 for a female.

MS(15) : 1 for the married ; 0 for the unmarried.

RP(16) : 1 for an S who lived in the dormitory on campus ; 0 otherwise.

PSCA(17) : 1 for an S who often participated in studentclub activities ;  
0 otherwise.

MFS(18) : 1 for an S who was supported by the home government or U.S. educational institution; 0 for an S who was self-supporting.

## V .Results and Discussion

### 1. Means and Standard Deviations of Variables :

Table 2 : Means and Standard Deviations  
( N = 52 )

Variable	Mean	S.D.
1. GPAc	3.50	0.22
2. IQv	114.71	8.04
3. IQp	113.50	8.38
4. GREv	326.84	64.61
5. GREq	577.37	76.53
6. TOEFL	493.16	25.43
7. GPAh	3.46	0.26
8. GPAu	3.53	0.23
9. CC	36.82	17.50
10. AL	12.58	1.93
11. GAP	0.34	0.47
12. Maj	0.61	0.49
13. Sex	0.50	0.50
14. Age	27.61	5.17
15. MS	0.53	0.50
16. RP	0.35	0.46
17. PSCA	0.42	0.47
18. MFS	0.39	0.49

## 2. Correlation Analysis

Table 3:

Table 3: Correlation Matrix

Variable	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	.67	.22	.61	.64	.49	.67	.82	.67	.10	.36	.31	.21	-.04	.26	-.29	-.12	.64
2		.21	.82	.81	.61	.38	.55	.38	-.14	.11	.45	.13	-.17	.21	-.26	.02	.39
3			.31	.48	.33	-.16	.03	.04	.21	.10	.22	.28	-.24	-.26	.03	-.13	.08
4				.85	.77	.29	.52	.29	.09	.10	.44	.07	-.23	.12	-.23	.07	.30
5					.64	.29	.51	.35	.07	.24	.37	.06	-.20	.11	-.22	-.17	.34
6						.15	.31	.11	.04	-.02	.29	.06	-.30	-.21	.08	.17	.17
7							.78	.60	.08	.19	.06	.20	.18	.33	-.29	.04	.50
8								.66	-.04	.36	.32	.31	-.01	.35	-.39	.00	.69
9									.18	.60	.19	.22	.21	.44	-.44	-.22	.50
10										.19	-.15	.08	-.21	-.21	.13	.00	-.10
11											.01	.17	.21	.46	-.40	-.50	.33
12												.16	-.29	.10	-.10	.05	.10
13													-.01	.00	-.06	.05	.27
14														.46	-.35	-.28	.19
15															-.76	-.42	.23
16																.39	-.36
17																	-.05

2. Correlation Analysis:

The intercorrelation coefficients ( Table 3 ) among 18 variables are productmoment coefficients.  $1/\sqrt{N} = 0.14$  may be used as the standard error for inferring if a relationship different from zero. Variables 2 through 12 except 3 and 18 were found to have significant positive relationship to the criterion. The variable 18 had the same predictive relationship to the criterion, too. However, the variable 16 had significant negative relationship to the criterion.

### 3. Regression Analysis :

For the purpose of using multiple linear regression the determine the unique contribution of proper sets of the predictor variables to the prediction of the criterion, 17 predictors were grouped into logical sets, subsets, sub-sub-sets, and so forth down to the individual variables as shown in Table 4.

As shown in Chart 1, the regression of the criterion, cumulative grade-point average, on all of the predictors, 2-18, had an RSQ equal to 0.8730. The square root of this value, 0.93, is the correlation coefficient between the criterion and the best weighted composite of the predictors.

Table 4 : Hierarchy of Variables

Intellectual variables :	
Mental abilities measured by standardized tests :	
WAIS : Verbal IQ	(2)
Performance IQ	(3)
GRE : Verbal score	(4)
Quantitative score	(5)
TOEFL score	(6)
School work :	
Past record :	
High school GPA	(7)
Undergraduate GPA	(8)
Current record :	
Cumulative credits	(9)
Academic load	(10)
Graduate academic program	(11)
Major	(12)
Non-intellectual variables :	
Identification data :	
Sex	(13)
Age	(14)
Marital status	(15)
Environmental conditions :	
Residential pattern	(16)
Participation in student-club activities	(17)

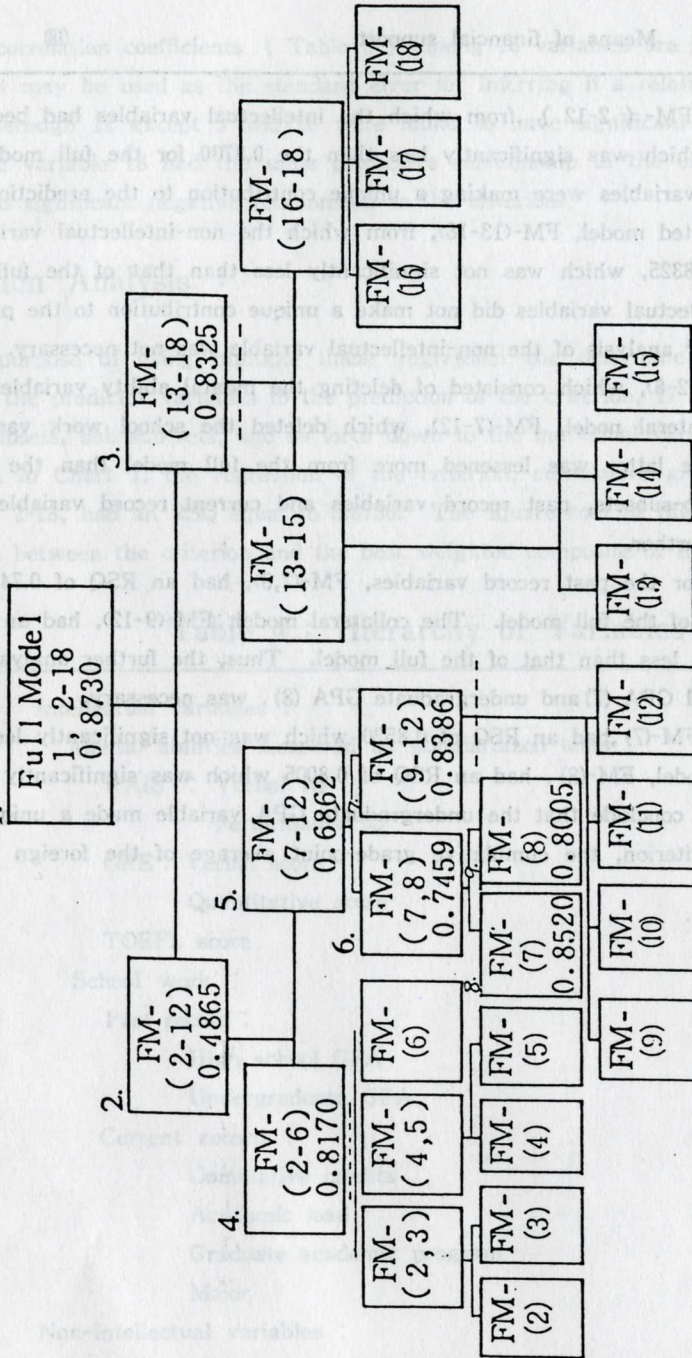
The restricted model, FM- ( 2-12 ) ,from which the intellectual variables had been deleted, had an RSQ equal to 0.4865, which was significantly less than the 0.8730 for the full model. This indicated that the intellectual variables were making a unique contribution to the prediction of the criterion. However, the restricted model, FM-(13-18), from which the non-intellectual variables had been deleted, had an RSQ of 0.8325, which was not significantly less than that of the full model. This indicated that the non-intellectual variables did not make a unique contribution to the prediction of the criterion. Thus, the further analysis of the non-intellectual variable was not necessary.

The next subset, FM-(2-6), which consisted of deleting the mental ability variables, had an RSQ equal to 0.8170 ; the collateral model, FM-(7-12), which deleted the school work variables, had an RSQ equal to 0.6862. The latter was lessened more from the full model than the former. This suggested that the two sub-subsets, past record variables and current record variables about school work ,should be analyzed further.

The restricted model for the past record variables, FM-(7,8), had an RSQ of 0.7459, which was significantly less than that of the full model. The collateral model, FM-(9-12), had an RSQ of 0.8386 which was not significantly less than that of the full model. Thus, the further analysis of two individual variables, high school GPA (7) and undergraduate GPA (8) , was necessary.

The restricted model, FM-(7) ,had an RSQ of 0.8520 which was not significantly less than that of the FM. The collateral model, FM-(8) , had an RSQ of 0.8005 which was significantly less than that of the FM. Thus, we can conclude that the undergraduate GPA variable made a unique contribution to the prediction of the criterion, the cumulative grade-point average of the foreign student at the graduate level.

CHART 1 SCHEMATIC FOR REGRESSION MODELS\*



\* With exception of Block 1, numbers in parentheses in second line indicate variables omitted from FM. Third line in block shows RSQ for the model. The dotted line shows that testing of further subsets was terminated at this point.



## VI. Summary

Of the 17 predictors used in this study, the undergraduate GPA variable had the highest relationship to the criterion. The regression analysis indicated that this variable made a unique contribution of the prediction of the foreign student's cumulative GPA at the graduate level.

## REFERENCES

- Schmid, J., and Reed, S.R. Factors in retention of residence hall freshmen. Journal of Experimental Education, 1966, 35 (1), 28-35.
- Schmid J., and Huston, S.R.(Eds.) Topics in human factors research. New York : MSS Information Corporation, 1972.
- Kelly, F.J., et al. Research design in the behavioral sciences : Multiple regression approach. Carbon-dale and Edwardsville : Southern Illinois University Press, 1969.
- Walton, B.J. Research on foreign graduate students. International, Educational and Cultural Exchange, 1971, 6 (3), 17-29.

## 中文摘要

本研究旨在以「多元迴歸方法」( Multiple Regression Approach ) 分析美國北科羅拉多大學外國學生的學業成就與其他各種因素之關係，藉以找出決定其學業成績的重要預測因素，作為進行外國學生學業輔導的參考。

本研究係就北科大五十二位外國研究生現有檔案中個人資料加以歸類整理，以累積的學業平均分數 ( cumulative grade-point average ) 為標準變項 ( criterion variable )，以其他十七種有關因素為預測變項 ( predictor variables )，進行多元迴歸分析。為了統計分析的方便，將預測變項分為心智的因素、非心智的因素兩大類別。前者包括 WAIS, GRE, TOEFL 等測驗分數與高中、大學各項學業記錄資料；後者包括性別、年齡、婚姻、住宿、參與學生社團活動以及求學生活經費來源等情況。

本研究發現下列四項主要的結果：

- (一) 累積的學業平均分數與十七種預測變項之複相關係數高達 0.93，亦即後者足以解釋前者變異量的 87%。
- (二) 心智的因素對累積的學業平均分數之預測有決定性的影響；但非心智因素則否。
- (三) 在心智的因素中，學業記錄資料比標準測驗分數更具有預測累積的學業平均分數之效力。
- (四) 在十七種預測變項中，大學部的學業平均分數與累積的學業平均分數之間有最密切的關係。換言之，前者是預測後者一項重要因素。