

給職能治療研究者的一點建議（II）：方法篇

姚開屏

摘要

作者延續前一篇對職能治療研究者提出綜合性的建議（見給職能治療研究者的一點建議（I）：綜論篇），本篇繼續針對數項常見於研究上，容易被誤解或疏忽的統計分析及研究方法，予以討論並提出具體的解決之道，期望藉由此方面的溝通，協助職能治療的研究更上一層樓。（職能治療學會雜誌 1996; 14: 103-113）

關鍵語：職能治療、研究方法、統計方法

筆者前一篇是對職能治療研究者在研究上提出整體性的建議，⁽¹⁾本篇的目的則是就統計方法或研究觀念上常容易被研究者忽略或誤解之處做詳細的討論，並提出具體解決之道。以下分為十二點來一一討論。

一、取樣方法及推論

取樣的方式有分成機率取樣 (probability sampling) 與非機率取樣 (non-probability sampling) 兩大類。其中機率取樣又可分為簡單隨機取樣 (simple random sampling)、系統隨機取樣 (systematic random sampling)、分層隨機取樣 (stratified random

sampling) 及聚落隨機取樣 (cluster random sampling)。而非隨機取樣又可分為偶發取樣 (accidental sampling)、立意取樣 (purposive sampling) 及配額取樣 (quota sampling)。每一種取樣的方式都有其特點及適用情形，對於各種取樣方法的介紹，讀者可以參見市面上許多的書籍。在此我想特別提到，嚴格的說起來，只有機率取樣所收集到的資料才適用於做統計分析，⁽²⁾因為計算平均數的抽樣分配 (sampling distribution of the mean) 的標準差（此種標準差又稱為標準誤差，standard error）的條件之一是樣本必須是被隨機取樣的，也就是說每一個研究母群中各份子被選取的機率是相同的，因

台灣大學心理系副教授

受文日期：86年2月1日

修正日期：86年2月15日

接受刊載日期：86年2月15日

索取抽印本聯絡人：姚開屏，台灣大學心理系

台北市羅斯福路四段一號

著此種條件，估計參數的信賴區間(confidence interval)、抽樣誤差(sampling error)等才得以建立。

然而在臨床醫學研究中，可能因為受限於時間及人力資源或患某種病症的人數實在太少，而沒有辦法使用機率取樣的方式抽樣。目前臨床研究中以偶發取樣或立意取樣等非隨機取樣的方式被使用得最多。筆者建議研究者應該盡可能在能力所及之處進行適當的隨機取樣，特別是當自己的研究結果要推論回廣大的母群時。例如：若要建立台灣地區學齡兒童的視覺動作整合能力(visual-motor integration)常模(norm)，則選擇適當的隨機取樣方式就非常的重要。研究者的研究若無法使用隨機取樣，則至少研究者在研究過程中(例如：將樣本分組實驗時)能採隨機指派(random assignment)的方式分組。⁽³⁾另外，因為研究的最終目的多半是想將研究樣本的結論推論回去至母群，若研究者受限於某些因素而不能做隨機取樣，但他若能證明自己樣本重要的特質與欲推論回去的母群的重要特質相近，則可較為放心的做推論，否則研究者最好在做推論時做某種程度的保留。例如：對台大醫院的精神病人給予某種「訓練課程」，研究結果結論此「訓練課程」對所有的精神病人有效。這個推論出現的問題是台大醫院(一級教學單位及國家醫學中心)的精神病人是否具有代表性，能代表全台灣的所有精神病人？會來台大住院的病人在教育水準、家庭經濟狀況等皆可能與台灣大多數的精神病人不同，因此若研究的結果可能受到除「訓練課程」以外的因素影響，我們就不能推論研究結果至廣大的母群，除非研究者能證明這些「訓練課程」以外的因素並不足以影響研究結果。

二、樣本人數

通常我們會以為樣本數越大越好，其實這也未必正確。樣本貴在具代表性而非大量性。研究若是使用一大堆卻不具代表性的樣本，倒不如使用少數卻具代表性的樣本來得正確。另外，相對比較起來，實驗性質的研究使用的樣本數可較少，而屬於觀察或相關性質的研究，就必須有較大的樣本才好，所使用的經驗法則是樣本數至少大於變項數的五至十倍。⁽⁴⁾大量的樣本數的影響為即使研究結果呈現些微的差異，也很容易因為樣本數很大而檢驗結果達顯著，這樣的顯著結果只有「統計上」的意義，並沒有臨床實際及理論上的意義。⁽⁴⁾因此研究者應考慮在有限的資源內，收集適量且具代表性的樣本。

三、量化的分析需配合文獻的佐證

有人誤以為量化的研究決定了一切，統計分析的顯著結果具有不可違抗的性質，這種觀念是錯誤的，因為量化分析只是較為客觀的研究「工具」，量化研究若沒有文獻的佐證與支持，所得的顯著結果也不見得有太多的意義，可能只是「垃圾進、垃圾出」(garbage in, garbage out)罷了。⁽⁴⁾

⁽⁴⁾ 一個無研究目的的人隨意收集周圍的許多變項來做相關，也可發現許多無法做任何解釋的變項間高相關，如果此人又用了大量的樣本，則可能各種無法解釋的各種顯著關係都將出現，因此研究者切忌被統計的結果牽著鼻子跑。一個好的研究者在研究上需居於主導的地位，研究設計一開始時需參考足夠且正確的文獻，從中發現研究問題，並對研究問題做適當的假設，再作適當的資料收集與分析，獲得結果之後研究者需去瞭解為什麼使用量化分析所得到的結果(不)會支持自己原先的假設。有一則笑話，一個研究人員的研

究資料經統計分析後，發現結論非常荒謬，完全離自己的假設有一大斷的距離，因為統計的結果確定是如此，他沒辦法只好將就地編出許多不合理的討論與推論，可是後來才發現他只是輸錯一筆影響重大的資料罷了。在實際的環境中，這種事情也可能發生，因此研究者應以研究文獻理論為基礎，而視量化分析方法為輔助研究的工具，小心地解釋分析所獲得的結果。

反觀之，有些人視統計分析為「垃圾進、垃圾出」的方法，這可能的原因並非統計本身的錯誤，而是使用統計分析的人誤用或濫用了各種統計方法並導致錯誤的結論。統計方法的產生有深厚的理論研究為基礎（例如：在許多假設條件之下而推導出方法），對統計方法一知半解的人很容易誤用或濫用了統計，因此，有人視統計為「垃圾進、垃圾出」，實非統計之罪，乃使用者之過！

四、測量工具的建立

醫學的研究難免會牽涉到測量，有些測量（例如：血壓、體重、血紅素）可用很客觀的工具來計算出，但有些測量則不是那麼容易。因為醫療效果必需使用一個良好的測量工具才能正確的反映出，因此測量工具的建立是非常重要的，這裡所謂「良好」是指工具需要有好的信度、效度及標準化等特質。關於測量工具發展之基本觀念及步驟之介紹，讀者可參見筆者發表在職能治療學會雜誌上的文章⁽⁵⁾。若讀者想更深入的瞭解這些步驟背後的原理及實際執行步驟時該注意的事項，可與我討論。建立一個好的測量工具實際上不如想像中的簡單，需經過對研究主題有經驗的人來設計測量的內容及形式，並收集適當的樣本資料，而後作各題目的分析，刪除或修正不適當

的題目，之後再次收集資料，如此反覆多次，並做信、效度分析，最後才能得到最後版本的測量工具。在設計的過程中有許多方面的細節該注意，例如：題目的形式(format)，度量化的方法(scaling methods)，語句/內容的描述方式，題目的排列位置等都有許多原則該注意。編製測量工具是一個非常繁瑣的過程，研究者只有經由不斷的練習及思考，才可設計出好的測量工具。最後我想強調，要證明職能治療的效果，從一開始（若有需要的話）一個好的篩檢工具及一個好的測量受試者表現的工具是非常必要的，這些工具的發展應該先於任何對治療效果的研究。

五、度量化方法

在編製量表時，幾乎不可避免的要考慮到用何種度量化方法(scaling methods)來收集資料。度量化的方法有許多種，各種度量化方法的形式也相差很多，我在此僅介紹一種目前最常被用到的度量化方法--李克氏量尺(Likert scale)，並討論用此方法設計量表的原則。

李克氏量尺是Likert於1932年發展出來的，⁽⁶⁾這種量尺法以古典測驗理論(classical test theory)為基礎，通常可分為三大類：

(a)評估同意度(agreement)，例如：李克氏六點量表

完全不同意 不同意 有點不同意
有點同意 同意 完全同意

(b)評鑑程度(evaluation)，例如：李克氏五點量表

很好 好 普通 差 很差

(c)評估頻率(frequency)，例如：李克氏五點量表

總是 經常 偶而 很少 從未

有時這些文字也伴隨著數字（如：1--2--3--4--5--6或百分比100%--75%--50%--25%--0%）。總之，

李克氏量尺法就是用排列順序大小的等級尺度來收集資料。接下來我要談幾個大家對此類型量尺所常提到的話題：

A.如何分析此類型的資料？

使用李克氏量尺的方式收集資料需有至少兩方面的假設，即假設各點間距離（如：1-2, 2-3, 5-6）是相等的（equal interval scales），另外也假設所有人對各點詞的定義是共通的（interpersonal comparability）。⁽⁷⁾如此在做資料分析時，才可以各點的實際數字或順序來作運算分析。事實上這兩個假設可能並不太合理，因為各點乃人為硬性訂出的，若要視其數字間等距以做為運算的基礎是有些牽強。另外，每個人對同一詞句（例如：總是、經常、偶而）的定義會不同，如何將每個人的反應結果混為一談呢？針對這個問題，我過去曾提出一個變通的方法來解決，⁽⁸⁾不過這個方法是經由數學模型的方式推演而來的，對於一般沒經驗於寫電腦程式的研究者來說，要自己應用此模型較為困難，因此目前仍多是用傳統的方式來分析以李克氏量尺所收集到的資料。但研究者仍應該瞭解使用李克氏量尺的一些假設及限制，在作資料分析及結果解釋時需要特別小心。傳統的分析方式多是將資料視為等距變項，而後用一般的母數統計方法分析，但也有人主張用無母數統計法。由過去的研究經驗得知，前面的第一個假設（點間等距）所產生的問題可能較小，較大的問題出現在第二個假設（共通定義）。如果研究者能將各詞做比較清楚的定義，例如：對「總是」、「經常」、「偶而」等詞另外加註實際頻率值的範圍，使得受測者有較一致的判斷標準，則資料分析時用母數統計法應可被接受。另外，Comrey在討論量表使用的點數時，也主張（至少五點）最好是使用七點量尺，則用母數統計分析

法來分析資料時才較恰當。⁽⁹⁾最近我的同事針對上百個中文量尺詞作分析，以找出每個量尺詞的尺度數值(scale values)。她研究結果的應用，或可解決點間不等距的困擾。⁽¹⁰⁾

B.使用點數越多越好，以增加精確度？

有人認為為了增加測量的準確程度，應該用的點數越多越好。這是錯誤的，因為若填量表者的心智程度無法區分到如此細微的差別，反而會減低了測量的準確度、可信度。那是不是使用愈少點數愈好呢？少的點數或可增加測量的一致性，但卻不能反映出受測者真正的程度。因此決定測量的點數需依照研究的問題、目的及研究的對象而定。要考慮填量表者的心智能力，以找到最適合且最能反映出受測者特質的點數為最佳，因此回答決定點數之多寡這個問題並沒有一定的答案。試舉一例，欲測量老年人或教育程度較低的人對某事件的看法，則可能使用的量尺點數不可太多，但對其他類型的人則可能會用不同的點數。

C.奇數點或偶數點？中間點(neutral point)存在的必要性？

這個問題如同前一個問題一樣沒有一定的答案。有些人主張為了防止填答者（特別是中國人）循「中庸」之道圈中間之點，而主張一定要用偶數點。這個論點其實也不盡然，因為受測者未必有此傾向，況且中間點也有其存在的價值。決定點數之多寡應該看研究的內容及研究的對象而定。若研究主題期望區分受測者的差異，而題目有容易使受測者走「中庸」之道的傾向，則可能用偶數點尺度較佳，以強迫受測者區分心中的差別，以做出抉擇。

六、建立量表效度所使用的因素分析法

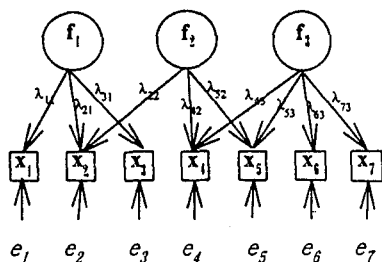
要想瞭解本節所談的「因素分析」的基本概念，讀者可參考相關此主題的專書或多變分析 (multivariate analysis) 書籍。本節則假設各位讀者已有因素分析的基本知識及分析經驗，而來討論常被用來作為建立量表效度的一種因素分析抽取因素法--主成份分析法 (principal component analysis)。若讀者完全沒有任何因素分析的知識或經驗，可略過此節。

發展測量工具時，「因素分析」是一種常被用來作為選題或建立效度的一種方法。讀者會發現很多的論文在使用因素分析建立測量工具的效度時，都是用主成份分析法來抽取因素，並且用最大變異法 (varimax) 來旋轉主軸以得最好的解。但依照較近代的因素分析學者的看法，在探討測量工具的效度時，用主成份分析法來抽取因素並不恰當，而用最大正交轉軸法來旋轉主軸有時也不見得合適，⁽¹¹⁾理由是主成份分析的目的及考量，與我們所要以建立工具效度的邏輯不符。

假設某測量工具有 k 個變項 (x_1, x_2, \dots, x_k)，若這 k 個變項實則是由 m 個潛在因素 (f_1, f_2, \dots, f_m) 所決定 ($m < k$)，則用因素分析法以建立工具效度，即是藉由所觀察的變項間的關係，而找出那些因素。我們可用以下的數學模式及圖形表示：(注意圖中箭頭的方向)

$$x_i = \lambda_{i1}f_1 + \lambda_{i2}f_2 + \dots + \lambda_{im}f_m + e_i \quad (i=1, 2, \dots, k),$$

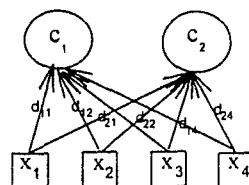
λ_i 為因素負荷量， e_i 為測量誤差。



此數學模式告訴我們，變項是潛在因素的線性組合，而測量誤差是存在的。因素分析的目的是在尋找潛在因素以解釋變項間的關係，它只分析變項間受共同因素影響的部份，即共通值 (communality)。其分析的過程是藉由尋找適當 (小於變項個數) 數目的潛在因素，以去估計變項間的相關矩陣，使得估計的相關矩陣 (estimated correlation matrix) 與觀察的相關矩陣 (observed correlation matrix) 差距在可接受的範圍內，用此以決定因素的數目。用主因子法 (principal factor) 或最大概似法 (maximum likelihood) 來抽取因素，則符合了此邏輯。⁽¹¹⁾

但以主成份分析來抽取因素的方法其本質卻是：找出原變項的線性組合，稱為「主成份」 (components, 有人也稱其為「因素」⁽¹²⁾)，使得受試者在這些成份上的變異量為最大，因此就可用少數幾個成份來解釋個體間的差異。另外，主成份分析不考慮測量誤差的存在，而所找出的 m 個主成份 (c_j) 彼此間必是相互獨立的。在資料分析時，此方法著眼點在於變項的共變量而非變項間的相關。我們可用以下的數學模式及圖形表示：(注意圖中箭頭的方向)

$$c_j = d_{j1}x_1 + d_{j2}x_2 + \dots + d_{jk}x_k \quad (j=1, 2, \dots, m), d_j \text{ 為主成份係數}$$



由以上粗略的介紹，我們可以知道用主成份分析法或其他因素分析法所做的因素抽取方式在本質上是不一樣的。主成份分析法若被用來建立測量工具的效度，則在本質上並不符合建立效度

的邏輯。另外，因為主成份分析法不考慮誤差的存在，而可能傾向選取較多的因素，並且各變項在第一因素上的負荷量會較大。因此當考慮使用主成份分析法或非主成份分析法時，即使二者分析的結果差別（特別是在當測量的變項很多時）可能不會太大，⁽¹²⁾研究者仍主張使用邏輯上比較正確的方法去抽取因素，而不用主成份分析法。⁽¹¹⁾另外，因素分析有因素未定性 (indeterminacy) 而可考慮做轉軸以得最好的解。在決定用正交 (orthogonal) 或用斜交 (oblique) 轉軸時，研究者主要應該以理論為根據，而不是以資料為導向。⁽¹¹⁾也就是說研究者在做轉軸之前，就由研究所根據的理論中知道因素間彼此的關係，而已決定是使用正交或斜交轉軸法。

筆者在此節中用很簡略的方式討論建立量表效度所使用的因素分析法，包括做因素抽取時所用的方法及欲得最好解所用的轉軸法，讀者可參見文獻13至26，以對上述的討論得到更完整的瞭解。

七、相關性 vs. 因果關係

兩變項相關並不表示變項間有因果關係的存在。雖然這個觀念大家都能接受，但仍然在一些研究論文中可以看出作者有將變項之相關性推論成因果關係的傾向。例如：抽煙與肺癌相關性高，若就因此結論說抽煙引起肺癌，則此推論並不恰當。要推論變項間的因果關係常常是不容易的，就單一獨變項而言，最常使用的方法是實驗法，也就是研究者控制所有其他的變項，而只單獨操弄一個獨立變項，看此獨立變項影響依變項的情形來瞭解因果關係。若想瞭解許多變項間的因果關係，則可根據理論而提出因果關係的假說模型，再藉由具有解釋因果關係的分析法，如：結構方

程模型法(structural equation modeling)等方法來瞭解變項間的因果關係。⁽²⁷⁻²⁸⁾

八、量化研究 vs. 質性研究

將所研究的現象予以「量化」一直是科學研究的基本要求，任何測量結果皆嘗試著用數字來表示，並且用發展已上百年的計量方法（例如統計法）來分析，這種量化分析有很深厚的理論為基礎，因此其方法能為大眾所接受。反觀「質性」研究的崛起乃最近數十年的事，目前在人類、社會學或教育界的領域裡使用質性研究的較多。主張量化研究或主張質性研究的兩派人士時有彼此不服，互相攻訐的情形發生。研究者不禁要問到底那一種方法好呢？我自己的學習背景雖是以量化為主，但我個人認為兩種方法皆有其優缺點，使用時需視研究主題而定。在通常情況下，我會主張如果量化的資料收集並不困難時，應該盡量使用量化的方法，這是因為量化的發展有良好而久遠的歷史，只要能收集適當的資料，並且使用適當的量化方法，就可得到適當的結論。在一些情況下，我們若非得用質性研究時，研究者需注意使用何種質性研究方式。目前質性研究仍有許多派別，由於質性研究的發展歷史尚淺，因此方法仍不很穩定，難免有各說各話，卻又各自漏洞百出的情形出現。研究者不論使用何種方法，最好都能對各種方法有深入的瞭解，包括其理論背景、過程步驟、資料收集及結論探討等。

九、直接測量 vs. 間接測量

醫療界所收集到的資料通常為直接觀察到的資料，這些資料可能是研究對象的人口學資料、實驗室的分析資料或醫療測量所反映出的資料等，這些資料若有客觀的測量工具，通常都能得到正

確的結果。反觀心理學的研究，則多半是用間接的方式來收集有興趣的研究資料，例如：態度 (attitude)、個性 (personality)、智力 (intelligence)、價值觀 (values)、成就 (achievement) 等，這些變項我們稱做「潛在變項」(latent variables)，這是因為這些變項無法被直接觀察收集到，他們的存在是因為根據某些學說，而後被驗證推導出來的。至於他們如何被收集而成為具體的資料，則需經由有經驗的心理學者，根據一套嚴格的過程而設計量表收集資料而得。因此在設計測量此類型的變項時，需經由設計良好的心理計量步驟 (psychometric procedure) 以減少測量上的誤差。

一些從事醫療研究的人會堅決的認為間接測量方式所收集到的資料是不可信且不客觀的，因此反對使用任何量表、問卷，他們寧可只相信實驗室精密儀器下所分析的資料。的確，精密測量的結果可信度極高，但是實驗室的方法並不能解決所有的研究課題，至少我前面所提到的各種潛在變項資料，若僅用目前的醫學研究法還無法收集得到。這些醫療人員的反對可能原因之一是，他們不瞭解量表編製的過程實在並不如他們所想像的一樣簡單，毫無依據，完全不客觀及標準化。編製一個好的測量量表其實需由受過相當訓練的人，依據嚴格的編製標準及理論根據，並經過不斷的修正才得以完成。職能治療相關的研究也可能去探究研究對象的潛在變項 (例如：態度、想法、認知能力等)，因此這些研究人員最好也能有基本的量表編製知識，或與心理計量學者合作共同設計客觀且標準化的量表。只有嚴謹設計出的量表，才能確實的反映出受試者的潛在特質。

十、選擇適合的統計方法

一個好的研究需要使用「適當」的統計方法

來分析資料，所謂「適當」不一定要用複雜的統計方法。若資料能以最簡單的統計方法分析，而得到所需要的結果，就是用了最適當的統計方法。另外，若資料不符合母數統計法所需的假設，則使用無母數統計法可能比使用母數統計法來的更合適些。因此認為「母數統計法較無母數統計法重要或有力(power)」的觀念是不正確的，統計方法的選用應該依資料的性質而定，研究者需瞭解資料的特性，以找出最適合的方法來分析資料。

十一、資料分析的順序

我常接到朋友的電話或學生直接來辦公室問我關於資料分析的問題，在他們非常簡單的介紹他們的研究後，就直接問我他們的研究資料該用何種統計方法分析。我通常不直接給予答案，而是反問他們他們所收集到資料的各變項是用何種測量尺度 (measurement scale) 所測量的，各變項的基本分配是什麼樣子，各變項間的關係如何。如果我沒有這些訊息，我根本無法正確的回答他們的問題，因為這些訊息直接決定了資料分析的方法。我希望研究者養成一個習慣，也就是當研究者收齊資料時，他的下一步資料分析不是直接想應該用什麼母數或無母數假設檢定法來檢定資料，而是該先盡可能的將所有變項的「描述性」資料先求出，這些描述性資料包括了變項的集中趨勢指標 (如：平均數、中數、眾數)，離散趨勢指標 (如：標準差、變異數、偏度、峰度) 等，以及變項之間的相關係數。因為這些描述性的資料可提供給研究者相當豐富的訊息，導引研究者下一步的研究方向。有時光從這些描述性的資料，就可反映出許多有趣的研究結果，或激發出更多的研究課題。而非常重要的一點是決定使用何種假設檢定法，必須隨變項的測量尺度屬性及資料分配

情形而定。例如：通常而言，只有用等距 (interval) 或等比 (ratio) 的測量尺度所得之資料才可考慮用母數統計法分析，然而這些資料若其分佈為極偏態或樣本人數相當的少，則還需考慮使用無母數統計法來分析資料可能更適合，因為多半的母數統計法（如：T檢定、ANOVA、迴歸分析等）皆需有資料是從常態分配的母群中出來的假設，如果資料不符合這種要求，就不能用母數統計分析法。最後，再次提醒：認為「推論性統計資料較描述性資料重要」的觀念是需被修正的。

十二、統計方法的假設

在前一篇相關的文章給「職能治療研究者」的一點建議 (I) : 綜論篇"中我已提過每一種統計分析法的背後都有假設 (assumptions)，在這些假設之下，才能導引出這些統計方法。因此這些統計方法之所以可被使用，資料必須先符合這些統計方法的假設，若資料不符合就不能使用這些統計方法，前文中我曾以 T 檢定為例，說明統計方法的假設。不過在此我得提出一個觀念——「強韌性」(robustness)。統計方法的「強韌性」是指當資料的性質離假設所需的條件（如：常態分配、變異數同質性等）的要求有一點距離時，其檢驗結果仍不受影響的程度。仍以二樣本獨立 T 檢定為例，若資料的偏態 (skewness) 不是太大時，也就是說所算出的偏態值在 ± 1 之間（常態分配偏態值為 0），則資料都可被視為常態分配。就二獨立樣本的變異數同質性的假設來看，若二組樣本人數相同，則變異數即使「不」同質，對檢定結果的影響也很小。^{6, 29)} 以下用一點篇幅談談該如何檢驗資料是否符合常見的假設特質。

如何知道資料的分配是否是常態分配 (normality)？那就非得由資料的描述性統計而得知。例

如：從資料平均數、中數、眾數之間的相對關係就可粗略的反映出資料是否為偏態，是否與常態分配的對稱性質不同。另外，資料的偏度 (skewness) 與峰度 (Kurtosis) 值也可告訴我們資料的分配情形。若用現成的統計軟體（如：SPSS、SAS）可求得各種集中趨勢及離散趨勢指標，在 SPSS 中可用 DESCRIPTIVES 或 FREQUENCIES 及其附帶的指示，在 SAS 則可用 PROC MEANS 或 PROC UNIVARIATE 及其附帶的指示，來獲得變項的各種描述性資料。另外，這兩個常用的統計軟體也可直接提供對資料作常態分配假設檢定（ H_0 : 資料為常態分配 vs. H_1 : 資料為非常態分配）的結果，因此可使用：

SPSS: NPAR TESTS /K-S (normal)

SAS: PROC UNIVARIATE NORMAL;

二者皆提供 Kolmogorov-Smirnov 常態分配假設檢定的檢驗值及 p 值。

如何判定二獨立組資料有相近的變異數 (homogeneity of variances)？在用 SPSS (T-TEST) 或 SAS (PROC TTEST) 對資料作統計檢驗分析時，報表就直接提供對二組樣本變異數是否同質的檢驗結果。研究者需先看這部份的結果，再依此決定要看 T 檢驗哪些部份的結果。若讀者做 ANOVA 分析，則多組獨立資料的變異數是否同質的假設檢定也可直接由報表中讀出。

另一個常見的假設為資料間彼此獨立 (independence)，以二組獨立樣本為例，理論上說起來，二組獨立樣本的來源需做到絕對的隨機 (random)，此乃意味著每組內的每個樣本從母群體中被抽取到的機會需均等，並且每個樣本被選取時，並不受其他樣本是否被選取的影響。在許多實際的研究上，真正的隨機取樣 (random sampling) 是不太可能達成的，然而研究者需至少對樣本做到隨

機指派 (random assignment)，以減少因樣本不獨立或不隨機而產生檢驗結果的誤差。(3)在做複迴歸分析 (multiple regression) 或相關分析 (correlation analysis) 時，資料彼此間的獨立性也是一項非常重要的假設，目前仍沒有非常好的方式來檢驗此特性，比較常用的方法是使用時間序列 (time series) 統計法中求資料間相關性的方法，例如：將迴歸分析後的殘差 (residuals) 與時間 (time) 之變化作相對應之圖，可以看出是否資料值的變化隨時間的改變做有系統的 (如：上昇、下降) 變化。另外，迴歸分析中可用 Durbin-Watson 統計值來知道資料間相關性的強度，在 SPSS 或 SAS 皆可使用迴歸分析及其附帶的指示而可得此檢驗值：

SPSS: REGRESSION /VARIABLES = Y X1 X2
/DEPENDENT = Y
/RESIDUALS = DURBIN

SAS : PROC REG; MODEL Y = X1 X2 / DW;

以上我討論了十二點研究者在研究過程中所可能遭遇的問題，並對一些研究上較容易混淆的觀念予以澄清，希望我的討論能提供給各位在未來的研究上有所幫助，也祝福職能治療，無論是在研究方面或在臨床治療方面都有很大的進展。

參考文獻

1. 姚開屏：給「職能治療研究者」的一點建議 (I)：綜論篇。職能治療學會雜誌1996; 14:96-102。
2. Dane FC: Research Methods. California: Brooks/Cole, 1990.
3. Cohen BH: Explaining Psychological Statistics. California: Brooks/Cole, 1996.
4. 余民寧：垃圾進、垃圾出--談計量方法的正用與誤用。政大教育研究所主編：教育研究方法論文集。台北：台灣書店，1994，101-112。
5. 姚開屏：從心理計量的觀點看測量工具的發展。職能治療學會雜誌1996; 14:v-xxi.
6. Likert R: A technique for the measurement of attitudes. *Arch Psychol* 1932; 140: 44-53.
7. Brady HE: Factor and ideal point analysis interpersonally incomplete data. *Psychometrika* 1989; 54: 181-202.
8. Yao KPG: Bayesian estimation of Thurstonian ranking models based on the Gibbs sampler (dissertation). Champaign, Illinois: University of Illinois, 1995. 163p.
9. Comrey AL: Factor-analytic methods of scale development in personality and clinical psychology. *J Consul Clin Psychol* 1988; 56: 754-761.
10. 翁儷禎：頻率及同意度詞之心理量尺值。中華心理學刊 (審稿中)。
11. 翁儷禎：因素分析應用之一覽。章英華、傅仰止、瞿海源編：社會調查與分析。台北：中央研究院民族學研究所，1995: 245-259。
12. Kline P: An Easy Guide to Factor Analysis. London: Routledge, 1996.
13. Bentler PM, & Kano Y: On the equivalence of factors and components. *Multivariate Behavioral Research* 1990; 25: 67-74.
14. Catell RB, Vogelman S: A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research* 1977; 12: 289-325.
15. Cliff N: The eigenvalues-greater-than-one rule and the reliability of components. *Psychological*

- Bulletin 1988; 103: 276-279.
16. Cliff N: Derivations of the reliability of components. *Psychological Report* 1992; 71: 667-670.
 17. Comrey AL: Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology* 1978; 46: 648-659.
 18. Jensen AR, & Weng LJ: What is a good g? *Intelligence* 1994; 18: 231-258.
 19. Joreskog KG: Basic ideas of factor and component analysis. In: Joreskog KG, Sorbom D eds: Advances in Factor Analysis and Structural Equation Models. Cambridge MA: Abt Books, 1979, 5-20.
 20. Lee HB, & Comrey AL: Distortions in a commonly used factor analytic procedure. *Multivariate Behavioral Research* 1979; 14: 301-321.
 21. Mulaik SA: Factor analysis and *Psychometrika*: Major developments. *Psychometrika* 1986; 51: 23-33.
 22. Thurstone LL: Current misuse of the factorial methods. *Psychometrika* 1937; 2: 73-76.
 23. Tzeng OC: On reliability and number of principal components: Joinder with Cliff and Kaiser. *Perceptual and Motor Skills* 1992; 75: 929-930.
 24. Velicer WF, Peacock AC, Jackson DN: A comparison of component and factor patterns: A Monte Carlo approach. *Multivariate Behavioral Research* 1982; 17: 371-388.
 25. Wilson EB, & Worcester J: Note on factor analysis. *Psychometrika* 1939; 4: 133-148.
 26. Zwick WR, & Velicer WF: Comparison of five rules for determining the number of components to retain. *Psychological Bulletin* 1986; 99: 432-442.
 27. 姚開屏：結構方程模型的基本原則及符號。職能治療學會雜誌，1993; 11: 79-93。
 28. 姚開屏：結構方程模型法在職能治療研究上之應用。職能治療學會雜誌，1993; 11: 65-78。
 29. Howell DC: Statistical Methods for Psychology. 3rd edition, Boston, MA: PWS-Kent, 1992.



Some Suggestions to the Researchers of Occupational Therapy, Part II: Methodology

Kaiping Grace Yao

Abstract

The author discusses some points of statistical methods and research methodology that are misunderstood easily in the research of occupational therapy. The purpose of this paper is to assist the researchers of occupational therapy to do better studies. (J Occup Ther ROC 1996;14:103-113)

Key words: occupational therapy, research methodology, statistical methods

Kaiping Grace Yao, Ph.D.
Associate Professor
Department of Psychology,
National Taiwan University,
Taipei, Taiwan, ROC.

Received: Feb. 1, 1997

Revised: Feb. 15, 1997

Accepted for publication: Feb. 15, 1997

Address reprint requests to: Kaiping Grace Yao

Department of Psychology, National Taiwan
University, Taipei, Taiwan, ROC.

kaiping@ccms.ntu.edu.tw

