

模糊統計在社會調查分析的應用*

吳柏林** 楊文山***

壹、前言

社會科學中的實地調查方法，常以受訪者填答問卷，或是經由面對面、電話訪談，而由訪員填寫問卷，做為收集社會科學經驗資料的一種重要工具(Sudman and Bradburn, 1982)。但是，在社會科學理論中所提出的假設，定義的範圍(concepts)，及其發展成可以測量的經驗指標(empirical indicators)，由於遣辭用句、字眼順序以及問題型式迥異，問卷的設計可能會產生相當多的歧異，並進而影響到研究結果的信度與效度(Schuman and Presser, 1981；楊文山與蔡瑤玲，1995)。另

* 本論文所使用之部份資料係採自行政院國家科學委員會所贊助之「臺灣地區社會意向基本調查」(計劃編號 NSC81-0301-H-001-46-B1) 81年6月經濟組專題調查大臺北地區之預試問卷。該計劃係由中央研究院中山人文社會科學研究所執行。在此感謝參與此次研究計劃的諸位教授；中央研究院社科所林忠正教授對模糊問卷的設計所提出的寶貴意見，並提供測試此一問卷的機會，在此特別表示感謝之意。而二位匿名評審教授對於本文之數據以及文章所提供的寶貴修正意見，也使文章內容更加出色並精確。在此也感謝政治大學統計所研究生黃麗鳳同學在資料處理以及運算過程中所付出的時間與努力。社會意向研究助理對於收集資料所付出的辛勞，謹此一併致謝。本文之內容與解釋由作者負責。

** 政治大學應用數學系教授

*** 中央研究院中山人文社會科學研究所研究員

外，與自然科學研究對象不同的是，社會科學中的經驗指標，與經驗事實的相對關係並非一成不變，社會現象會因時、空、情境的不同，而產生很大的變化及差誤，若未能以更客觀的調查研究方法，很容易誤導研究結果。

社會調查過程中，社會科學家總是使用某些固定或清晰的回答模式測度喜好程度，並以整數所衍生的名義尺度記錄此評量和行為結果間的統計相關性 (Manton, Woodbury and Tolley, 1994)。然而，社會學中所涉及的概念常常是模糊 (fuzzy)，而非確定、清楚的。比方，在一個典型的調查中，受訪者被問及去年一年中的總收入 (total income) 是多少；又比方，受訪者被問及「理想子女數」，「支持政黨程度」，以及「老人的年齡標準和身份」等問題時，這些問題並沒有清楚的界定，受訪者在面對這些問題時可能在「贊成」與「不贊成」間產生許多不同程度的「贊成」與「不贊成」的答案。

但是，當社會學家從事社會經驗性研究時，以上的許多問題，常常以某些固定的認定做為回答問題的依據 (Li, 1989)。老人的年齡和身份認定就是一個相當有趣且值得探索的問題。65歲以上是否就可以算定是接受老人年金的年齡，或是更年輕一點，或是更年長一些，或是在60歲到70歲之間，但依老人身體及家庭收入狀況而定，就是一個相當有趣的經驗性政策問題。可是，當社會科學家設計問卷時，很容易把此一本質模糊的問題設計成「你贊不贊成政府開辦老年年金保險制度？」(臺灣地區社會意向調查，83年2月定期調查執行報告)，或「你是否贊成／不贊成65歲以上老人接受老人年金」或「不知道」做為可能的答項。所以，當問題本身的經驗指標可能是模糊，或不一定的狀況時，如果用結構性的問卷設計，並且延用清楚 (crisp) 的數學模式做為分析的依據時，就可能導致誤差，產生信度、效度的偏誤，影響研究的可能結果。

在這篇論文中，我們介紹一個模糊統計調查，與分析模糊資料的方法。並將此一模糊統計與分析方法實際地應用在社會調查中，從社會與經濟問題來分析人類思考與行為模式。用以驗證此一模式在測量人類模糊思維的可能性及可行性。這新的方法稱為模糊邏輯模式理念 (the fuzzy logic model of beliefs)。模糊邏輯的應用，主張個人喜好程度不需非常清晰或是有條理。所以，模糊邏輯與布林 (Boole) 模式理念大相逕庭，布林模式中 A 和 B 的比較只有三種結果： $A > B$ ， $A < B$ ， $A \sim B$ 。但人的思維是複雜的，具有不明確的喜好。因此能對真實狀況做全盤描繪的構想，早已成為某些邏輯學家的目標。換句話說，對人類而言，模糊模式比直接指定單一物體一個值，較合適於評估物體間的相關特性。由於其他特性有助於意見的評定，因此使用模糊邏輯時，必須對其他特性也加以說明，以便將人們的喜好程度轉換成便於計算的效用函數 (utility function)。

對那些人類在思考認知不易表達完善的問題，藉隸屬度函數與模糊統計分析似乎可提供更明確表達。本文亦架構幾個重要的模糊統計量的定義，如模糊期望值，模糊眾數，模糊秩階…等。最後我們做幾個實證調查，希望藉模糊問卷調查及模糊統計分析，對人類在思想認知不易表達之問題，能有更合理的測度與解釋。

模糊邏輯彌補了布林邏輯的不足，但並沒有取代布林邏輯，兩個方法可以互存。其關鍵在於每一理論各描述不同的理念。而模糊子集理論 (fuzzy set theory) 的主要特質，在於對於無界的可能偏好，能做出無限的解說。使用模糊邏輯來測度，必需要使用模糊統計代替傳統統計。而模糊邏輯測度主要根基於模糊集合的基本理論。由於模糊集合的基本理論，相當佔據篇幅，所以在下面的模糊測量中將做一簡要的說明，在此不做詳細介紹，有興趣的讀者可參考 Klir & Folger (1988), Li (1989), Zimrmann (1991) 或 Manton, Woodbury and Tolley (1994)。

貳、模糊邏輯與測度

首先，我們簡述有關模糊集合理論在社會科學研究上的應用。基本上，模糊邏輯模式理念的發展，源起於加州柏克來系統控制學家 L. A. Zadeh (1965) 所提出的模糊集合理論，對於測量模糊概念所做的重要貢獻 (McNeill and Fredberger, 1993)。Zadeh 以為過去數學理論中的集合論 (set theory) 和人思維的最大不同在於數學的精準性及排它性。在傳統集合理論中，若 x 為 A 集合中的任一元素，則 $x \in A$ ，而 x 具有某種性質，若 $x \notin A$ ，則 x 不為 A 集合中的元素，且不具有此一性質，其它性質則不可能同時存在。但集合論忽略了一項重要的事實，那就是 X 可能存在於 A 或 \bar{A} 的範圍中，而不固定於一點上，而是一項程度上的隸屬性。

模糊理論乃擴展傳統二值邏輯的觀念與運算方法，利用隸屬度函數 (membership function) 來表示事物的模糊現象，並對於不同的考量賦予特別的運算。對社會科學家而言，有時應用模糊數學方法與模式來表示可能較直接給定單一的特定值來得適切些，並可能較合於用來評估變數與變數之間的相關性。此外，由於其他因素往往也會影響人們對某一事物的評定，因此使用模糊邏輯時，必須對所謂的「其他因素」加以說明，以便將人們的喜好程度轉換成便於計算的隸屬度函數。

要對模糊的現象予以適當的集合描述，必須應用新的邏輯假設，即論域上的對象由屬於某一集合到不屬於某一集合是漸進而非突變的。因此，隸屬度函數的設定在模糊理論的測度與分類上，扮演相當重要的角色，其可定義如下：設 A 是 X 論域 (或定義域) 上的一個模糊子集 (論域指被討論之事物全體的變數，有時亦稱為空間)。 A 的隸屬度函數以 μ_A 表示，乃指滿足 $\mu_A: X \rightarrow [0,1]$ 的對應關係。此 μ_A 的值反映了 X 中的元素 x 對於 A 模糊子集的隸屬程度。 $\mu_A(x)$ 的值接近於 1，表示 x 隸屬於 A 的程度很高。 $\mu_A(x)$ 的值接近於 0，表示

x 隸屬於 A 的程度很低。利用這些介於 $[0,1]$ 的實數，即可標記人們對被討論之事物的看法。

在模糊集合中的元素，可能只有部份屬於某集合。例如，「年輕人」這一名詞，幾歲以上不再年輕，並無一個確切的界限來區分年輕和不再年輕。而模糊集合的定義，可以顯示出一個20歲的人是「90%的年輕」，當某一個60歲的人相對的只有「30%的年輕」。在模糊集合中，隸屬度的全距從 0到 1。因此，每個語言變項，比方「年輕」，代表一個可能性分布。此分布的均值表示人類真實評定是「年輕人」的數值。而關於分布的評定，可能因人而異，且不需為常態。

在社會學及經濟學的研究中，研究對象的答案很少是確切的真或偽。若要試著去分析人類理念，必會遭遇到行為的不確定性。而連續區間值具有能處理亦真亦偽狀況的能力，因此應用區間值的模糊特性所做的分析調查，可使研究者處理不確定性，確實是符合實際狀況的測量工具。

2.1 模糊隨機變數

在討論模糊分析調查之前，我們先給予模糊隨機變數的定義及基本觀念：

定義 2.1. 令 $Q(\mathfrak{R})$ 為實數軸 \mathfrak{R} 上所有模糊集合類的標準集合表象 (normal set representation)。若 \mathfrak{R} 的子集之一家族系統 $\{A_\alpha(\omega) \mid \omega \in \Omega, \alpha \in (0,1)\}$ 具有下列性質：(i) 對所有 $\omega \in \Omega, \{A_\alpha(\omega) \mid \alpha \in (0,1)\}$ 為 $X_\omega \stackrel{def}{=} X(\omega)$ 的標準集合表象。(ii) 對所有的 $\alpha \in (0,1)$ ，映射 A_α 和 \bar{A}_α 是 $u \rightarrow \bar{B}_1$ - 可測的。(μ 是 σ -代數； \bar{B}_1 是 $\mathfrak{R} \cup \{-\infty, +\infty\}$ 上的柏瑞爾 σ -代數 (Borel σ - algebra)；此處我們定義對所有的 $\alpha \in (0,1)$ 和 $\omega \in \Omega$ ， $\underline{A}_\alpha(\omega) \stackrel{def}{=} \inf A_\alpha(\omega)$ ； $\bar{A}_\alpha \stackrel{def}{=} \sup A_\alpha(\omega)$ 。則稱映射 $X: \Omega \rightarrow Q(\mathfrak{R})$ 稱為模糊隨機變數 (fuzzy random variable)。

欲應用傳統的數學邏輯觀念，明確定義一模糊隨機變數並不容易。此處僅提供一比較合理與符合經驗法則參考。目前對於模糊隨機變數、模糊統計量之定義，文獻上尚無一致的共識，有興趣之讀者可參考 Kwakernaak (1978)、Kruse & Meyer (1987)、或 Maton, Woodbury & Tolley (1994)。

2.2 模糊假設檢定

古典的統計檢定必須陳列明確的假設。比方，當我們想檢定兩母體平均數是否有差異時，虛無假設是「兩個平均數確實相等」。然而，有時我們想要知道的只是兩平均值是否非常逼近，此時傳統的檢定方法並不適用於這種包含不確定性的假設檢定。

Casals *et al.* (1986, 1989) 曾以模糊事件表徵的模式為基礎來討論統計假設檢定；Saade & Schwarzlander (1990) 針對混合資料 (hybrid data) 提出模糊假設檢定；Miyamoto (1990: 240) 也曾論及模糊假設檢定。但他們對模糊假設檢定方法的構思不但基於資料的隨機不確定性，也包含了資料中的非隨機不確定性。而「兩母體平均數幾乎相等」的假設檢定，只討論包含於假設中的非隨機不確定性，這樣的不確定性不但與資料本身無關，且無法假設於資料上。因此 Casals *et al.* (1986, 1989)，Saade & Schwarzlander (1990)，甚或貝氏近似所提供的架構都無法處理上述的假設檢定。

Watanabe & Imaizumi (1993) 將 Neyman & Pearson 設立的古典假設檢定擴展，提出了上述模糊假設的模糊統計檢定 (the fuzzy statistical test)，而用此方法檢定所得的推斷結論也是模糊的。在第三節中，我們試著利用模糊語言系統和模糊統計從事問卷調查，來分析和說明人類的理念模式。

參、模糊統計分析

將模糊統計分析運用於調查領域來測度理念，並不受限於特定理論。Sugeno (1977) 提議當客觀性伴隨著主觀性時，系統變得模糊。他並強調判斷一系統是否為模糊。也就是說，判斷系統是由一確定值或一模糊區間所構成，有其必要性。

通常利用確定數值所做的傳統抽樣調查分析的缺點包括：

- (1) 人類的思考與行為本來充滿著模糊過程，傳統問卷的數字常被過度解釋。
- (2) 為迎合數字的精確需求，實驗資料常有被過度使用之嫌。
- (3) 為了簡化或降低數學模式的複雜性，卻將實際狀況之相關與動態特質忽略。

而以模糊區間估計人類喜好，是一種描繪人們對字句的思維或感覺的方法。因此凡是牽涉到人的思慮的問題，用模糊觀念處理較合適。本論文即對抽樣調查所獲得的傳統和模糊統計樣本做進一步的統計分析與探討。

3.1 模糊統計

一般人在日常生活中常對可能發生的事進行推論。例如，一個決策的過程中，某些重要因素的考量，都可能是下決策的關鍵。所以當我們考慮的因素愈多愈詳密，愈能確信最後所做的決策。模糊統計資料就是考量了人們擁有多重喜好，而以模糊區間和加權模糊區間來進行討論。但因個人認知的喜好程度不同，而具有不定等長的期望區間。下面我們提出一些關於模糊統計量的定義：

定義3.1. 令 $F_i = (a_i, b_i)$ 是樣本模糊區間， $i = 1, 2, \dots, n$ ； n 為樣本數。則樣本 $\{F_i\}$ 的期望模糊區間 (the expected fuzzy interval) 上下

界分別是樣本模糊區間上下界值的個別加權和。換言之，期望模糊區間 (a, b) 中， $a = \frac{\sum a_i}{n}$ (期望最小值)， $b = \frac{\sum b_i}{n}$ (期望最大值)。

定義3.2. 樣本 $\{F_i\}$ 期望模糊區間的平均數 (average of the expected fuzzy interval) 定義為 $uF = \frac{a+b}{2}$

定義3.3. 樣本 $\{F_i\}$ 模糊區間的中位數 (median of the fuzzy interval) 定義為 $median_F = (m_l, m_u)$ ，其中 m_l 為 $\{a_i\}$ 的中位數， m_u 為 $\{b_i\}$ 的中位數。

定義3.4. 樣本 $\{F_i\}$ 模糊區間的眾數 (mode of the fuzzy interval) 定義為 $mod e_F = (m_L, m_U)$ ，其中 m_L 是 $\{a_i\}$ 的眾數， m_U 是 $\{b_i\}$ 的眾數。

定義3.5. (模糊排序) 樣本 $\{F_i\}$ 模糊區間的等級定義為 $Rank(F_i) \leq Rank(F_j)$ ，如果 $\frac{a_i+b_i}{2} \leq \frac{a_j+b_j}{2}$ 。當 $\frac{a_i+b_i}{2} = \frac{a_j+b_j}{2}$ 時，我們稱 $Rank(F_i) = Rank(F_j)$ ，即 F_i 與 F_j 形成相持 (tie)。

3.2 模糊問卷調查表

在社會調查中，若欲由喜好程度的資訊得知個人的理念，可以分別以傳統的調查表得到受訪者行為意圖的點估計，另由模糊問題方式得到模糊統計估計，以便兩者加以比較。例如，探討的問題為：

「在現代社會中(如臺灣)你預期男性和女性應受幾年教育?(請給一個確定的年數)」

要求受訪者回答一個確定的數字，此為傳統抽樣調查的問題形式。而調查表中另一問題詢問：

「在現代社會中 (如臺灣) 你預期男性和女性應受幾年教育? (請給一個範圍)」

這樣的問題要求受訪者給定一個區間來回答模糊形式的問題。

以上所介紹的是獲得模糊屬量資料以及傳統資料的問卷形式。至於屬性資料的搜集，傳統的問法是要要求受訪者回答一個確定的等距尺度，屬單一邏輯。例如評量「對生活是否有枯燥不快樂的感受」，問卷形式可以為

1.沒有感受 -----> 5.感受很強

「覺得生活枯燥不快樂 _____ . _____ . _____ . _____ . _____ .」

而應用模糊理論，語意與模糊統計所設計的問卷，運用可複選的形式符合實際狀況，可能可以得到較精確的結果。其問答方式為

「覺得生活枯燥不快樂 沒有感受 _____ . 感受很強」

要求受訪者在 _____ 線段上劃下感受範圍 (區間)，比方：稍為感受可記為 //////；有時很強有時普通 //////。用如此的回答形式來獲得模糊感受程度。若對分配程度較具概念者，我們亦可引導他們以對事件感覺之隸屬度值，分別填入。例如：

等級	1	2	3	4	5	模糊感受程度
覺得生活枯燥不快樂		0.6	0.4			稍微感受

等級	1	2	3	4	5	模糊感受程度
覺得生活枯燥不快樂				0.2	0.8	感受很強

等級	1	2	3	4	5	模糊感受程度
覺得生活枯燥不快樂	0.5				0.5	有時感受很強有時沒有感受

以下，我們將以實際問卷調查所獲得的資料，比較傳統和模糊問卷形式所提供之資訊。其中實例研究一至三，乃截取中央研究院中山人文社會科學研究所1992年6月以隨機抽樣法，抽調100位大台北都會地區(包括台北市與台北縣)的民眾，所做的「臺灣社會意向調查」研究之預試問卷調查資料。其中有98份完整的回收調查表，成功率为98%；實例研究四，乃截取北部某商業專科學校學生輔導中心於1993年，以配額抽樣法對北部地區五所專科學校所做的寂寞感受抽樣調查。562位受訪者中，共回收497份完整的調查表，有效問卷回收率約88%。

3.3 實例分析

實例研究一

問題：在大台北地區，一個四口之家一個月的生活費用大概需要多少？(生活費用的定義：包括食、衣、住、行等)

表 3-1 為「台灣地區社會意向」預試調查搜集所得的資料。其中「上限」、「下限」是指將模糊區間資料的上、下界值各別加以彙總的次數統計資料；「確定值」資料，是指傳統要求受訪者回答確定數字，所彙總的次數統計資料。

表 3-1：人們對日常生活花費需求的看法

下限 (次數)	8 1	10 15	15 4	20 12	25 9	30 25	35 7	38 1	40 18	45 1	50 14	55 1	60 2	70 1	50 1
上限 (次數)	10 1	20 5	25 2	30 17	35 3	40 25	45 4	50 17	55 2	60 12	70 4	80 2	100 3	200 1	
確定值 (次數)	10 1	15 1	17 1	20 4	24 1	25 8	27 1	30 27	35 6	37 1	38 2	40 8	45 6	50 23	55 1
確定值 (次數)	60 2	70 1	75 1	90 1	100 1	150 1									

(單位千元)

其隸屬度函數為：

$$\begin{aligned}
 u(x) = & \frac{1}{98} I_{[8,10]}(x) + \frac{1}{98} I_{[10,15]}(x) + \frac{5}{98} I_{[15,20]}(x) + \frac{13}{98} I_{[20,25]}(x) + \frac{20}{98} I_{[25,30]}(x) \\
 & + \frac{28}{98} I_{[30,35]}(x) + \frac{32}{98} I_{[35,40]}(x) + \frac{25}{98} I_{[40,45]}(x) + \frac{23}{98} I_{[45,50]}(x) + \frac{20}{98} I_{[50,55]}(x) \\
 & + \frac{19}{98} I_{[55,60]}(x) + \frac{8}{98} I_{[60,65]}(x) + \frac{8}{98} I_{[65,70]}(x) + \frac{5}{98} I_{[70,75]}(x) + \frac{5}{98} I_{[75,80]}(x) \\
 & + \frac{3}{98} I_{[80,85]}(x) + \frac{3}{98} I_{[85,90]}(x) + \frac{3}{98} I_{[90,95]}(x) + \frac{3}{98} I_{[95,100]}(x) + \frac{1}{98} I_{[100,105]}(x)
 \end{aligned}$$

其中 $I_{[c,d]}(x)$ 為指標函數 (indicator function) ; 即 $I_{[c,d]}(x) =$

$$\begin{cases} 1 & \text{if } x \in [c,d]; \\ 0 & \text{if } x \notin [c,d]. \end{cases}$$

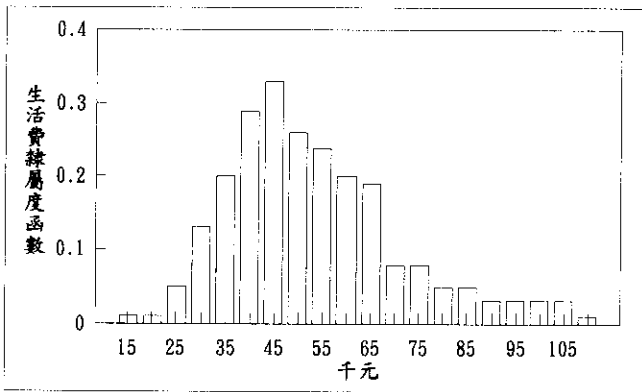


圖 3-1 生活費隸屬度函數

表 3-2 為根據模糊區間資料重新再統計彙總而得的相對頻數。

表 3-2：分組計算生活需求的相對頻數

需求分組(單位千元)	頻 數	相 對 頻 數
8--10	1	0.01
10--15	1	0.01
15--20	5	0.05
20--25	13	0.13
25--30	20	0.20
30--35	28	0.29
35--40	32	0.33
40--45	25	0.26
45--50	23	0.24
50--55	20	0.20
55--60	19	0.19
60--65	8	0.08
65--70	8	0.08
70--75	5	0.05
75--80	5	0.05
80--85	3	0.03
85--90	3	0.03
90--95	3	0.03
95--100	3	0.03
150--200	1	0.01

以「相對頻數」為各數值對於「生活費用」的隸屬度。

藉隸屬度函數，我們可以對本身是模糊的語言作定量的描述、劃分。35,000元對於「生活費用」模糊集而言，具有最高隸屬度，是最多民眾認可的生活所需成本。表 3-3將模糊調查和傳統調查的度量結果做一比較。

表 3-3：模糊統計與傳統統計分析結果之比較

模糊分析	$\mu_F = 40.9$	$F = (35.1, 46.7)$	$median_F = (30, 40)$	$mode_F = (30, 40)$
傳統分析	$\bar{x} = 39.9$	$S = 18.2$	$median = 35$	$mode = 30$

實例研究二

問題：在現代化的社會中，一個男生應該受多少年的教育才夠？

表 3-4 為「台灣地區社會意向」預試調查搜集所得的資料。其中「上限」、「下限」是指將模糊區間資料的上、下界值各別加以彙總的次數統計資料；「確定值」資料，是指傳統要求受訪者回答確定數字，所彙總的次數統計資料。為抽樣調查搜集的資料。

表 3-4：人們對男生教育年限的看法

下限 (次數)	9 5	12 36	14 18	15 3	16 34	18 1	21 1		
上限 (次數)	12 4	14 2	16 47	18 22	20 8	21 1	22 8	24 4	25 2
確定值 (次數)	12 18	14 8	15 4	16 56	18 11	21 1			

表 3-5 為根據模糊區間資料統計而得的相對頻數。

表 3-5：男生教育年限的相對頻數

教育年限	頻 數	相 對 頻 數
9	5	0.05
10	5	0.05
11	5	0.05
12	41	0.42
13	37	0.38
14	55	0.56
15	56	0.57
16	90	0.92
17	43	0.44
18	44	0.45
19	22	0.22
20	22	0.22
21	15	0.15
22	14	0.14
23	6	0.06
24	6	0.06
25	2	0.02

以「相對頻數」做為各數值對於「男生教育年限」的隸屬度。其隸屬度函數為

$$\begin{aligned}
 u(x) = & \frac{5}{98} I_9(x) + \frac{5}{98} I_{10}(x) + \frac{5}{98} I_{11}(x) + \frac{41}{98} I_{12}(x) + \frac{37}{98} I_{13}(x) + \frac{55}{98} I_{14}(x) + \frac{56}{98} I_{15}(x) \\
 & + \frac{90}{98} I_{16}(x) + \frac{43}{98} I_{17}(x) + \frac{44}{98} I_{18}(x) + \frac{22}{98} I_{19}(x) + \frac{15}{98} I_{20}(x) + \frac{15}{98} I_{21}(x) + \frac{14}{98} I_{22}(x) \\
 & + \frac{6}{98} I_{23}(x) + \frac{6}{98} I_{24}(x) + \frac{2}{98} I_{25}(x)
 \end{aligned}$$

其中 $I_c(x)$ 為指標函數；即 $I_c(x) = \begin{cases} 1 & \text{if } x = c; \\ 0 & \text{if } x \neq c. \end{cases}$

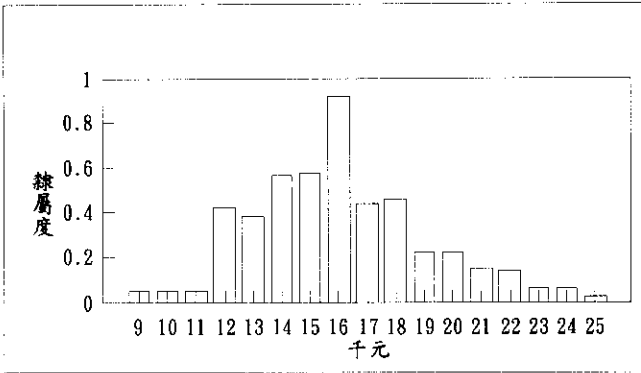


圖 3-2 男生教育年限隸屬度函數

16年對於「男生教育年限」模糊集而言，具有最高隸屬度，是最多民眾認為男生應受的教育年限。表 3-6將模糊調查和傳統調查的度量結果做一比較。

表 3-6：模糊統計與傳統統計分析結果之比較

模糊分析	$uF = 15.7$	$F = (13.8, 17.6)$	$median_F = (14, 16)$	$mode_F = (12, 16)$
傳統分析	$\bar{x} = 15.3$	$S = 1.9$	$median = 16$	$mode = 16$

實例研究三

問題：在現代化的社會中，一個女生應該受多少年的教育才夠？

表 3-7為「台灣地區社會意向」預試調查搜集所得的資料。其中「上限」、「下限」是指將模糊區間資料的上、下界值各別加以彙總的次數統計資料；「確定值」資料，是指傳統要求受訪者回答確定數字，所彙總的次數統計資料。為抽樣調查搜集的資料。

表 3-7：人們對女生教育年限的看法

下限 (次數)	6 1	9 5	12 37	14 19	15 3	16 31	18 1	21 1	
上限 (次數)	9 1	12 4	14 3	16 47	18 20	20 9	22 8	24 4	25 2
確定值 (次數)	9 1	12 19	14 10	15 5	16 56	18 6	21 1		

其隸屬度函數為

$$\begin{aligned}
 u(x) = & \frac{1}{98} I_6(x) + \frac{1}{98} I_7(x) + \frac{1}{98} I_8(x) + \frac{6}{98} I_9(x) + \frac{5}{98} I_{10}(x) + \frac{5}{98} I_{12}(x) + \frac{42}{98} I_{12}(x) \\
 & + \frac{38}{98} I_{13}(x) + \frac{57}{98} I_{14}(x) + \frac{57}{98} I_{15}(x) + \frac{88}{98} I_{16}(x) + \frac{41}{98} I_{17}(x) + \frac{42}{98} I_{18}(x) + \frac{22}{98} I_{19}(x) \\
 & + \frac{22}{98} I_{20}(x) + \frac{14}{98} I_{21}(x) + \frac{14}{98} I_{22}(x) + \frac{6}{98} I_{23}(x) + \frac{6}{98} I_{24}(x) + \frac{2}{98} I_{25}(x)
 \end{aligned}$$

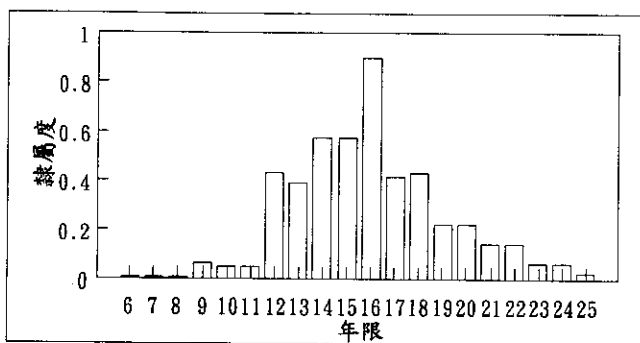


圖 3-3 女生教育年限隸屬度函數

表 3-8：女生教育年限的相對頻數

教育年限	頻 數	相 對 頻 數
6	1	0.01
7	1	0.01
8	1	0.01
9	6	0.6
10	5	0.5
11	5	0.5
12	42	0.43
13	38	0.39
14	57	0.58
15	57	0.58
16	88	0.90
17	41	0.42
18	42	0.43
19	22	0.22
20	22	0.22
21	14	0.14
22	14	0.14
23	6	0.06
24	6	0.06
25	2	0.02

表 3-8 為根據模糊區間資料統計而得的相對頻數。以「相對頻數」做為各數值對於「女生教育年限」的隸屬度。

16 年對於「女生教育年限」模糊集而言，具有最高隸屬度，是最多民眾認為女生應受的教育年限。表 3-9 將模糊調查和傳統調查的度量結果做一比較。

表 3-9：模糊統計與傳統統計分析結果之比較

模糊分析	$uF = 15.6$	$F = (13.7, 17.5)$	$median_F = (14, 16)$	$mode_F = (12, 16)$
傳統分析	$\bar{x} = 15.1$	$s = 1.9$	$median = 16$	$mode = 16$

實例研究四

由於快速的社會變遷，個人與個人間的關係逐漸由個人與團體間的關係所取代。因此，寂寞在現代社會中是一個相當普遍而惱人的問題。Russel et al. (1978) 爲了在寂寞的實證研究上有所突破，尋找建立一個可廣泛使用之單一而整體性的寂寞量表，經過一再地修訂，發展出 UCLA 寂寞量表。而後 Hays & DiMatteo (1987) 以因素分析等統計方法，由原 UCLA 量表中，抽出八題建立了 ULS-8 簡式寂寞量表。研究中針對寂寞感受程度的測量，此研究乃根據 UCLA 寂寞量表，配合臺灣的特殊背景，經過統計相關分析與模糊集合檢驗加以綜合、刪減後，發展而得精簡的8題寂寞量表。

本研究在探討如何應用模糊統計方法，對寂寞量表建立一套更有效率的測量準則。並希望藉這套測量準則，能事先預測出那些具有極端或雙重性格之學生，而施予事前輔導。首先我們針對北部地區五所專科學校作抽樣調查，表 3-10 爲傳統抽樣調查搜集的各題反應統計資料。

由表 3-10，我們發現對寂寞的綜合感受程度，呈平穩然後遞減的分配。但是此量表對於那些情緒起伏甚大者或過於內向者，可能無法得到一良好的度量。對寂寞所衍生出的各種副作用，如自殺，破壞性行爲亦無法提出較明確的預警作用。

表 3-10：傳統抽調寂寞反應頻率統計表

題目\感受程度	沒有感受 → 感受很深					合 計
	1	2	3	4	5	
1. 生活枯燥無趣味 (百分比)	96 (19.3)	132 (26.6)	152 (30.6)	76 (15.3)	41 (8.2)	497 (100)
2. 覺得生命失去目標	150 (30.2)	141 (28.4)	113 (22.7)	60 (12.1)	33 (6.6)	497 (100)
3. 無法肯定自我成就	67 (13.5)	118 (23.7)	142 (28.6)	102 (20.5)	68 (13.7)	497 (100)
4. 目前所學非所興趣	109 (21.9)	119 (23.9)	139 (28.0)	65 (13.1)	65 (13.1)	497 (100)
5. 在團體中被忽視	149 (30.0)	173 (34.8)	114 (22.9)	46 (9.3)	15 (3.0)	497 (100)
6. 沒有人真正瞭解我	92 (18.5)	127 (25.6)	153 (30.8)	76 (15.3)	49 (9.9)	497 (100)
7. 覺得缺少親近知己	143 (28.8)	132 (26.6)	109 (21.9)	65 (13.1)	48 (9.7)	497 (100)
8. 覺得需要異性朋友	132 (26.6)	107 (21.5)	158 (31.8)	55 (11.1)	45 (9.1)	497 (100)
合 計	938 (23.6)	1049 (26.4)	1080 (27.2)	545 (13.7)	364 (9.1)	3976 (100)

事實上寂寞是一個相當模糊的名詞，任何人都會有寂寞的感覺。而感覺強度的指標，傳統的測量工具卻很難訂定出來。主要困難有

- (1) 母體寂寞表的分配難以確立。
- (2) 寂寞強度的動態性、週期性無法測知。
- (3) 寂寞強度的兩極性或多極性無法區分。

基於以上原因，我們考慮用模糊問卷的方式，經由統計分析與模糊量度理論，建立一套模糊寂寞量度。希望能藉此一新的量表，更實際的測量出寂寞的感受與反應。表 3-11 為抽樣調查搜集的各題模糊反應統計資料。

表 3-11：模糊抽調資料寂寞反應頻率統計表

題目\感受程度	沒有感受 → 感受很深					合計
	2	3	4	5		
1. 生活枯燥無趣味 (百分比)	236.91 (47.7)	115.82 (23.3)	75.49 (15.2)	38.33 (7.7)	30.41 (6.1)	496.96 (100)
2. 覺得生命失去目標	260.98 (52.5)	117.72 (23.7)	65.72 (13.2)	31.40 (6.3)	21.15 (4.3)	496.96 (100)
3. 無法肯定自我成就	219.21 (44.1)	120.20 (24.2)	83.70 (16.8)	46.54 (9.4)	27.30 (5.5)	496.96 (100)
4. 目前所學非所興趣	54.36 (51.2)	115.52 (23.2)	68.35 (13.8)	29.20 (5.9)	29.53 (5.9)	496.96 (100)
5. 在團體中被忽視	261.76 (52.7)	120.68 (24.3)	76.34 (15.4)	24.68 (5.0)	13.52 (2.7)	496.96 (100)
6. 沒有人真正瞭解我	226.84 (45.6)	118.84 (23.9)	72.67 (14.6)	47.68 (9.6)	30.93 (6.2)	496.96 (100)
7. 覺得缺少親近知己	246.02 (49.5)	107.52 (21.6)	75.19 (15.1)	42.36 (8.5)	25.87 (5.2)	496.96 (100)
8. 覺得需要異性朋友	262.39 (52.8)	111.14 (22.4)	69.64 (14.0)	30.65 (6.2)	23.15 (4.7)	496.96 (100)
合計	1968.47 (50)	927.44 (23)	587.10 (15)	290.84 (7)	201.86 (5)	3975.71 (100)

根據表 3-11，得到二專生的模糊綜合寂寞隸屬度函數為

$$u(x) = 0.05 \times I_1(x) + 0.23 \times I_2(x) + 0.15 \times I_3(x) + 0.07 \times I_4(x) + 0.05 \times I_5(x)$$

我們訂定以下的寂寞量表標準：

定義 3.6. 單項模糊寂寞量表分配，SFDL (fuzzy distribution of loneliness for single item)，通常以抽樣結果得到。令 m 為感受程度之等級量表； P_{ij} 為個人在寂寞量表中，第 i 題感受程度第 j 等級的隸屬度。則單項模糊寂寞指數 I_{SFL} 定為

$$I_{SFL} = \frac{1}{m} \left(\sum_{j=1}^n jP_{ij} + \frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m jP_{ij} \right| \times P_{ij}}{k} \right) \tag{3.1}$$

其中 k 為感受程度等級中， P_{ij} 不為零的個數。

式 (3.1) 中，分母的第一項即為傳統的度量方法，第二則可視為模糊樣本的加權計分。單項模糊寂寞量表的分配函數，通常以抽樣結果得到。

以下我們以幾個特殊情況，說明定義 3.6 中對模糊感受的加權情形。以較常用之 5 等第量表為例 ($m = 5$)：

1. 當樣本選 5 時，此道題感受很強，個人寂寞量表感受百分比為 (0,0,0,0,1)，我們定此樣本之感受程度為清晰集合「5」， $I_{SFL} = 1$ 。
2. 當樣本選 3 時，此道題感受普通，個人寂寞量表感受百分比為 (0,0,1,0,0)，我們定此樣本之感受程度為清晰集合「3」， $I_{SFL} = 0.6$ 。
3. 當樣本為兩極化時，例如：個人寂寞量表感受百分比為 (0.5,0,0,0,0.5)，表示感受程度有時很強烈，有時沒有感受，此時樣本之模糊集合「 $0.5I_1 + 0.5I_5$ 」。感受程度為 $4 (= 3+1)$ ， $I_{SFL} = 0.8$ 。

4. 當樣本感受程度均等分佈時，例如：寂寞量表感受百分比為 (0.2,0.2,0.2,0.2,0.2)，此時樣本之模糊集合「 $0.2I_1+0.2I_2+0.2I_3+0.2I_4+0.2I_5$ 」。感受程度為 3.48 (3+0.45)， $I_{SFL} = 6.96$ 。

定義 3.7. 設 FDL 為模糊綜合寂寞量表分配。令 n 為模糊問卷題數，則綜合模糊寂寞指數 I_{SFL} 定為

$$I_{FL} = \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{m} \left(\sum_{j=1}^m j p_{ij} + \frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{k} \right) \right] \quad (3.2)$$

式 (3.2) 亦可寫成

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{m} \sum_{j=1}^m j p_{ij} + \frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{m \times k} \right) \\ &= \frac{\sum_{j=1}^m j \left(\sum_{i=1}^n p_{ij} \right)}{n \times m} + \sum_{i=1}^n \left(\frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{n \times m \times k} \right) \\ &= \frac{\sum_{j=1}^m j s_j}{n \times m} + \sum_{i=1}^n \left(\frac{\sum_{j=1}^m \left| j - \sum_{j=1}^m j p_{ij} \right| \times p_{ij}}{n \times m \times k} \right) \end{aligned}$$

其中最後一個等號第一項內的 s_j 表示為第 j 級感受程度的總合； $j=1, \dots, m$ 。

根據定義，我們得知一般二專生的單項模糊寂寞指數分別為：

表 3-12：二專生模糊抽調資料的單項模糊寂寞指數

題號	1	2	3	4	5	6	7	8
寂寞指數	0.5	0.47	0.51	0.48	0.46	0.51	0.49	0.47

綜合模糊寂寞指數為0.49

3.4 結果分析

通常，若有較多的樣本資料居於中間位置，在統計學上我們稱為集中趨勢 (central tendency)。一般以眾數或平均數用來代表此集中趨勢的量數。在實例一至三的研究，表 3-3、3-6、3-9 中，我們以統計上常用以代表群體的三種平均數——算數平均數、中位數、眾數，來比較傳統和模糊調查資料度量結果的異同。

在傳統的統計分析中，平均數都是用點來表示，所以又稱為點量數。但在模糊統計中，由於分析的是模糊區間資料，因此亦以區間形式來表示資料的集中趨勢，而其運算理念的構思，仍源於傳統統計的相關定義。另一方面，同一問題對於同一受訪者而言，其依傳統問卷形式所回答的確定數值答案，必包含於依模糊問題形式回答的模糊區間中，因此這兩種型式的資料，本身即具有某一程度的相關性，所以表 3-3、3-6、3-9，兩種資料分析的結果也有某種程度的相似性。

模糊邏輯考量了人類思維的複雜性及行為的不確定性，其允許模糊性現象的存在，較符合實際狀況，此一特點彌補了布林邏輯的不足，但並沒有取代布林邏輯。將模糊邏輯的觀念運用於問卷調查之分

析上，提供了一個新的搜集及分析資料的理念，和不同於往常的統計結果展示方法。

此外問卷調查時，對有多重感受的受訪者而言，傳統問法的單一邏輯，可能使其無所適從，且當尋問的問題涉及個人隱私時，受訪者很可能拒答。因此實例研究四中，允許人們擁有多重感受的模糊理論觀念，不但可減少廢卷率，還可幫助我們偵測出那些具有雙重性或危險傾向的學生，這是傳統的問卷型式所無法達成的目標。

肆、結 論

自從 Zadeh 在 1965 年提出了「模糊集合」的概念做為測試不明確隸屬度的工具以來，許多的模糊研究就嚐試以此方法設定理論架構，並嚐試以此一理論做為建構更為先進模糊理論的基礎。但模糊理論在社會科學中的應用迄今還是相當缺乏，而一些應用的範疇大部份均限定在計算機科學中的專家系統應用研究中 (Costas et al., 1994)。近年來社會科學中重要的研究，包括 Manton 和他的同仁所做有關模糊隸屬度和老人健康的一系列研究，這些研究的成果已受到相當廣泛的重視和應用 (Manton, Woodbury and Tolley, 1994)。雖然在心理計量學有一些模糊量表的設計，做為測量態度的依據 (Hesketh et al., 1988; Costas et al., 1994)，以及某些態度與偏好，如社會學中對於理想子女數以及男女教育年限的應用及測量 (Li, 1989)。但是大部份的研究對象均侷限於大學主修社會科學的學生的樣本中，這些研究幾乎不會將 Zadeh 所發展隸屬度的概念在一般人口的樣本中加以測試。本研究的重要貢獻在於利用「模糊集合」概念所發展出來的隸屬度概念應用於一般大型實地調查的樣本中，並且嚐試以模糊邏輯所發展出來的模糊測量量表，更進一步以模糊統計做為分析資料的依據。

在社會科學研究的測量理念上，模糊統計和模糊相關性的使用可說是數字模式的推廣。如果利用假性的「精確值」，可能誤導模型的

建構，也可能擴大研究結果和實際狀態之間的誤差。因此，具有模糊性的數值資料，在計量方法的演算過程中，是否符合人類的邏輯推論與歸納的原則，值得吾人深思。數值模糊的特性，在社會和經濟的實證研究過程中，應予以慎重考慮。

近年來，在人文社會科學的測度理念裡，模糊統計和模糊相關性日漸受到重視，這應是複雜的人文社會科學現象無法以傳統數值模型充分合理解釋的一種自然發展結果。Manski (1990) 曾指陳數字資料有需求過度及過度解釋的危險，如果應用模糊的數值資料，較能避免這樣的危險發生。有鑑於此，本文乃嚐試應用區間值的模糊特性所做的分析調查，可使研究者在處理不確定情境時能較符合實際狀況的測量工具。

本研究所獲得的結果，和 Manski (1990) 所發現的結果有些不同，即表現程度常發生在不是傾向於零，就是傾向於中立的狀態。因此我們建議對心理行為模式的建構與分析，可倚賴如語言法則的使用，而這些由許多語言造成的相關和傳統的積分、微分、代數和 ARMAX 方程間存在著直覺上的相似點，見 Sun & Wu (1993)。

最後，Kosko (1992) 建議神經網路和模糊語言系統共同應用做為函數估計量。希望此研究的方向將提供模糊統計一個有用的工具。然而，為了得到合適的信念準確度，期待神經計算學將成為值得花費時間的技巧，且將在社會科學研究方面，激勵更多未來模糊統計方法的實驗工作。

參考資料

楊文山與蔡瑤玲

- 1995 〈實地調查中複查資料的結構模型分析：以台灣地區社會意向調查為例〉，見章英華、傅仰止、瞿海源(編)，《社會調查與分析——社會科學研究方法檢討與前瞻之一》，頁83-110。臺北：中央研究院民族學研究所。

中央研究院中山人文社會科學研究所

- 1994.7 《台灣地區社會意向調查》，83年2月定期調查執行報告。

Casals, M. R., M. A. Gil, and P. Gil

- 1986 "On the Use of Zadeh's Probabilistic Definition for Testing Statistical Hypotheses from Fuzzy Information," *Fuzzy Sets and Systems*. 20: 175-190.

Casals, M. R. and M. A. Gil

- 1989 "A Note on the Operativeness of Neyman-Pearson Tests with Fuzzy Information," *Fuzzy Sets and Systems* 30: 215-210.

Costas, Concepcion San Luis and Pedro Prieto Maranon

- 1994 "Application of Diffuse Measurement to the Evaluation of Psychological Structures," *Quality and Quantity* 28: 305-313.

Hays, R. D. And M. R. DiMatteo

- 1987 "A Short-form Measurement of Loneliness," *Journal of Personality Assessment* 51 (1): 69-81.

Hesketh, B., R. Pryor, M. Gleitzman, and T. Hesketh

- 1988 "Practical Applications and Psychometric Evaluation of a Computerised Fuzzy Graphic Ration Scale," in Zeteny (ed.) *Fuzzy Sets in Psychology*. Elsevier Science Pub. North-Holland.

Klir, George, J., and Tina A. Folger

- 1988 *Fuzzy Sets, Uncertainty, and Information*. Englewood Cliffs, NJ: Prentice- Hall.

Kosko, B.

1992 *Neural Networks and Fuzzy System*. Englewood Cliffs, NJ: Prentice Hall.

Kruse R. and K. D. Meyer

1987 *Statistics with Vague Data*. D. Reidel Publishing Company.

Kwakernaak, H.

1978 "Fuzzy Random Variables. Part I: Definitions and Theorems," *Information Science*. 15: 1-15.

Li, Shaomin

1989 "Measuring the Fuzziness of Human Thoughts: An Application of Fuzzy Sets to Sociological Research," *Journal of Mathematical Sociology* 14: 67-84.

Manski, C.

1990 "The Use of Intention Data to Predict Behavior: A Best-case Analysis," *Journal of the American Statistical Associatio*. vol. 85, No 412: 934-940.

Manton, Kenneth G., Max A. Woodbury, H. Dennis Tolley

1994 *Statistical Applications Using Fuzzy Sets*. New York: John Wiley & Sons, Inc.

McNeill, Daniel and Paul Fredberger

1993 *Fuzzy Logic*. New York: Simon and Schuster.

Miyamoto, S.

1990 *Fuzzy Sets in Information Retrieval and Cluster Analysis*. Dordrecht: Kluwer Academic Publishers.

Russel, D., L. A. Peplau, and C. E. Cutrona

1978 "The Revised UCLA Loneliness Scale: Concurrent and Discriminant Validity Evidence," *Journal of Personality and Social Psychology* 39(3): 472-480.

Saade, J. J. and H. Schwarzlander

1990 "Fuzzy Hypotheses Testing with Hybrid data," *Fuzzy Sets and Systems* 35: 197-212.

Schuman, Howard, and Stanley Presser

1981 *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Sudman, S. and N. M. Bradburn

1982 *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.

Sugeno, M.

1977 "On Structured Set of Systems. Report No. 5," *The Working Group on Fuzzy Systems*. Tokyo.

Sun, C. And B. Wu

1993 "On Fuzzy Analyzing and Modeling in Linguistic Semantics," *1993 Conference of Fuzzy Theory and Applications*. Hsinchu, Taiwan.

Watanabe N. and T. Imaizumi

1993 "A Fuzzy Statistical Test of Fuzzy Hypotheses," *Fuzzy Sets and Systems* 53: 167-178.

Zadeh, L.

1965 "Fuzzy Sets," *Information and Control* 8(3): 338-53.

Zimmermann, H. J.

1991 *Fuzzy Set Theory and Its Applications*. Boston: Kluwer Academic Publishers.