行政院國家科學委員會專題研究計畫　成果報告

市場調查分析的新方法(II):模糊統計檢定分析與決策模式

計畫類別：個別型計畫
計畫編號：NSC94-2416-H-004-013-
執行期間：94年08月01日至95年07月31日
執行單位：國立政治大學應用數學學系

計畫主持人：吳柏林

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 95年9月18日

行政院國家科學委員會補助專題研究計畫 ■成果報告
□期中進度報

市場調查分析的新方法(II):模糊統計檢定分析與決策模式

計畫類別：■ 個別型計畫　　□ 整合型計畫
計畫編號：NSC 94-2416-H-004-013
執行期間：94 年 8 月 1 日至 95 年 7 月 31 日

計畫主持人：吳柏林
共同主持人：
計畫參與人員：

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
■出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，■一年□二年後可公開查詢

執行單位：國立政治大學應用數學學系

中華民國 95 年 8 月 20 日

1. 基本理論與性質

統計分析在各個領域中皆被廣泛利用。一些基本的敘述統計參數，如期望值、中位數及眾數等，並不因其方法簡單而失其重要性。在分析資料時，這些參數能夠簡單且快速地描述資料的基本結構，其中又以期望值最常被利用。在知識經濟之社會，多元思維逐漸取代傳統二元邏輯的思考與分析方法。過去使用單一數值的樣本來計算期望值的方法，已漸不符合現今複雜多變的智慧科技時代之需求。尤其是在具有多變性、不確定性、與訊息不完整性的財金與經濟環境下，過分強調對於數值之運算與數學假設的前提，反而更容易造成與現實環境及條件的背離、甚至是脫節。故在進行財金與經濟方面問題的研究時，利用軟計算方法與模糊統計的分析將會是一種較爲進步的測度方法。

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。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的參考。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。本研究我們提出一些關於模糊統計量的檢定問題與法則。

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.

模糊母體均數檢定

令 $F\mu$ 爲模糊樣本母體均數，我們欲檢定在模糊檢定水準 α 下，是否接受 H_0： $F\mu = F\mu_0$ 之假設，其中 $F\mu_0$ 爲模糊母體均數。

離散型模糊母體均數檢定方法

1. 樣本:設 Ω 為一論域，令 $\{L_j, j=1,...,k\}$ 為佈於論域 Ω 上的 k 個語言變數，$\{x_i, i=1,...,n\}$ 為一組模糊隨機樣本，且對每個樣本 x_i 對應語言變數 L_j 有一標準化之隸屬度 m_{ij}. 令 $L_{\max} - L_{\min}$ 表示相對於語言變數之全距(以 5 等第為例，即為 5-1=4)
2. 統計假設: H_0 ： $F\mu = F\mu_0$ vs. H_1 ： $F\mu \neq F\mu_0$
3. 統計量: 求此組模糊樣本 $\{x_i, i=1,...,n\}$ 之模糊樣本均數 $F\overline{X}$. 計算樣本均數與母體均數的反模糊化值 $\overline{X}_F$ 與 μ_0.
4. 決策:在模糊檢定水準 α 下，若 ｜ $\overline{X}_F$ - μ_0 ｜ ＞ $\alpha(L_{\max} - L_{\min})$ ，則拒絕 H_0.

Note: 對於左尾檢定 H_0 ： $\mu \leq \mu_0$ vs. H_1 ： $\mu > \mu_0$ 在模糊檢定水準 α 下，決策法則為:若 $\overline{X}_F$ - μ_0 ＞ $\alpha(L_{\max} - L_{\min})$ ，則拒絕 H_0. 右尾檢定亦同。

關於連續型模糊樣本，爲了使符號一致，我們直接說 $F\mu$ 等於某區間，而不一定將它反模糊化。

連續型模糊母體均數檢定方法

1.樣本:設 Ω 為一具有模糊均數$[a,b]$之論域，令$\{ x_i = [x_{li}, x_{ui}], i = 1,...,n \}$為一組模糊區間隨機樣本。

2 統計假設 H_0 ：$F\mu =_F [a,b]$ vs. H_1 ：$F\mu \neq_F [a,b]$

3 統計量:計算 $F\overline{X}$ =[$\bar{x}_l$, $\bar{x}_u$]

4 決策準則：在模糊檢定水準 α 下，計算 k=α $(b\text{-}a)$，若 $|\bar{x}_l - a| > k$ 或 $|\bar{x}_u - b| > k$，則拒絕 H_0 .

進行區間相等的檢定時，有時會遇到統計量落於事先假設的區間內的情況，但是可能由於區間相對過小，得到拒絕等於之結論。因此我們必須考慮模糊屬於檢定，以符合實際需要。

有界樣本的模糊屬於檢定

1 統計假設: H_0 ：$F\mu \in_F F\mu_0$ vs. H_1 ：$F\mu \notin_F F\mu_0$

2 統計量：隨機抽取一組模糊樣本$\{ x_i, i = 1,...,n \}$，計算 $F\overline{X}$ =[$\bar{x}_l$, $\bar{x}_u$]

3 決策：在模糊檢定水準 α 下，計算 k=α（b−a），若 $\bar{x}_l$ <a-k 或 $\bar{x}_u$ >b+k 時，拒絕 H_0

無下界樣本(*sample with no lower bound*)的模糊屬於檢定

1 統計假設: H_0 ：$F\mu \in_F (-\infty, b]$ vs. H_1 ：$F\mu \notin_F (-\infty, b]$

2.統計量: 隨機抽樣一組模糊樣本$\{x_i, i = 1,...,n\}$，計算 $F\overline{X}$ =(-∞, $\bar{x}_u$]

3 決策：在模糊檢定水準 α 下，令 k=α b，若 x_u >b+k 時，拒絕 H_0

無上界樣本(*sample with no upper bound*)的模糊屬於檢定

1 統計假設: H_0 ：$F\mu \in_F [a, \infty)$ vs. H_1 ：$F\mu \notin_F [a, \infty)$

2. 隨機抽樣一組模糊樣本$\{x_i, i = 1,...,n\}$，計算 $\overline{FX}$ =[$\bar{x}_l$,∞)

3 決策：在模糊檢定水準 α ,令 k=α $\boldsymbol{a}$，當 x_l <a-k 時拒絕 H_0

例 1　某農場主人想引進新品種的雞，作為炸雞用途。只要大部分的人試吃後的平均評價在普通以上，他就引進大量繁殖。於是他隨機找了 5 位顧客試吃，然後依語言變數給予隸屬度，得下表 1

表 1 農場主人想引進品種 P 的雞的滿意隸屬度函數

試吃者	1=很不滿意	2=不滿意	3=普通	4=滿意	5=很滿意
A	0	0	0	0.3	0.7
B	0	0	0	0	1
C	0	0.4	0.6	0	0
D	0	0	0	0.8	0.2
E	0.1	0.9	0	0	0

我們將此問題化成假設檢定 H_0 ： $\mu_f \le 3$ vs. H_1 ： $\mu_f > 3$。

經計算後可得 $\overline{X}_f$ =3.68. 在模糊檢定水準 α =0.1 下, $k = 0.1 \cdot (5-1)$. 因為 $\overline{X}_f$ - $\mu_f = 0.68 > 0.4$，故拒絕 H_0 ，因此他決定引進該新品種的雞。

例 2. 某人想開服裝店分店，準備只賣 12 種款式衣服。 因為是區域性，故他只找住在想要開店的地方附近的人士市場調查。他分別就 12 種款式衣服，隨機各找 10 人，給予 1 至 5 等第的滿意度隸屬度評分。他決定整體平均分數若大於 2.5 則開店，否則放棄開店。經過統計後得 $\overline{A}_f$ =3, $\overline{B}_f$ =2, $\overline{C}_f$ =3.1, $\overline{D}_f$ =2.5, $\overline{E}_f$ =2.6, $\overline{F}_f$ =4, $\overline{G}_f$ =1.2, $\overline{H}_f$ =2, $\overline{I}_f$ =1.8, $\overline{J}_f$ =2.9, 求出

$$\overline{X}_f = \frac{1}{10}(3+2+3.1+2.5+2.6+4+1.2+2+1.8+2.9+3+2)=2.51.$$

在模糊檢定水準 α =0.1 下，k= $0.1 \cdot 4$ =0.4. 因 2.51-2.5=0.01，而 0.01<0.4, 因此他決定不開店。

例 6.3 人力資源部提出現今 20 歲至 26 歲的年輕人要求平均待遇為 2 萬至 4 萬元。主計單位想要檢定此報告是否屬實，於是隨機找 6 位 20 至 26 歲的年輕人調查得到他們要求的待遇分別為

[3,4]，[1.8,2]，[2,3]，[4,6]，[2,2.5], [2.5,3].

統計假設 H_0 ： $F\mu =_F [2,4]$ vs. H_1 ： $F\mu \neq_F [2,4]$

將 1.8 看成[1.8,1.8]，根據模糊樣本均數定義可得

$F\overline{X} = [\,\overline{x}_l, \overline{x}_u\,] = [\,\dfrac{3+1.8+2+4+2+2.5}{6}, \dfrac{4+2+3+6+2.5+3}{6}\,] = [2.55, 3.42]$。在模糊檢定水準 α =0.1 下，計算 k= $0.1 \cdot (4-2) = 0.2$. 因為 2.55＞2 ± 0.2, 3.42＜4 ± 0.2. 故拒絕人力資源部平均待遇為 2 萬至 4 萬元的說法。

但是[2.46,3.42]確實落於[2,4]區間， $F\mu \in_F [2,4]$，也就是說 20 歲至 26 歲的年輕人要求平均待遇區間屬於 2 萬至 4 萬元區間。

3 模糊類別資料之卡方 χ^2 齊一性檢定

卡方齊一性檢定用來決定兩個或兩個以上的母體中，各類別的比例是否齊一之統計檢定方法。我們使用的卡方齊一性檢定統計量爲

$$\chi^2 = \sum_{i=A,B} \sum_{j=1}^{c} \frac{(n_{ij}-e_{ij})^2}{e_{ij}},$$

其中 n_{ij} 表示第 i 組樣本，第 j 類別的觀察次數， e_{ij} 表示在 H_0 下，第 i 組樣本，第 j 類別的期望次數。當 n 夠大時，卡方齊一性檢定統計量 χ^2 會漸近於自由度爲(c-1)之 χ^2 分配。因此在顯著水準 α 下，可查得 χ^2 分配之臨界值 $\chi^2_\alpha(c-1)$ ，若檢定統計量 χ^2 值大於 $\chi^2_\alpha(c-1)$ ，則拒絕 H_0 。

若從選項隸屬度來考慮，類別資料的單位是可再分割的。例如當我們被問到對政府某項施政滿意度時，可能是 0.6 滿意 0.4 非常滿意的模糊樣本。此時傳統 χ^2 檢定便無法處理此類類別資料問題。爲了解決此問題，因此我們提出模糊類別資料之卡方 χ^2 檢定過程如下:

離散型模糊母體均數齊一性檢定

1.樣本：設 Ω 為一論域，令 $\{L_j, j=1,...,k\}$ 為佈於論域 Ω 上的 k 個語言變數，$\{a_1,a_2,\cdots,a_m\}$ 與 $\{b_1,b_2,\cdots,b_n\}$ 來自兩不同模糊母群體 A,B 之兩組模糊隨機樣本。且對每個隨機樣本對應語言變數 L_j 均有一標準化之隸屬度 mA_{ij}, mB_{ij} 。$MA_j=\sum_{i=1}^{m} mA_{ij}$, $MB_j=\sum_{i=1}^{n} mB_{ij}$ ，為樣本對語言變數 L_j 之隸屬度總和。

2.事先假設: $H_0: F\mu_A =_F F\mu_B$, A, B 兩母體有相同之分配比率。

$$F\mu_A = \frac{\frac{MA_1}{m}}{L_1} + \frac{\frac{MA_2}{m}}{L_2} + ... + \frac{\frac{MA_k}{m}}{L_k} \text{ vs. } F\mu_B = \frac{\frac{MB_1}{n}}{L_1} + \frac{\frac{MB_2}{n}}{L_2} + ... + \frac{\frac{MB_k}{n}}{L_k}.$$

3.統計量：$\chi^2 = \sum_{i=A,B}\sum_{j=1}^{k}\frac{([Mi_j]-e_{ij})^2}{e_{ij}}$ （ e_{ij} 為期望次數，為了符合軟體計算 χ^2 檢定要求，我們表格之各細胞隸屬度總和用 4 捨 5 入以取得整數值。對於樣本數大於 25 個之模糊樣本其結果對決策影響並不大)。

4.決策法則：在 α 顯著水準下，若 $\chi^2 > \chi^2_\alpha(k-1)$ ，則拒絕 H_0

區間型模糊母體均數齊一性檢定

1. 樣本：設 Ω 為一論域，令 $\{L_j, j=1,...,k\}$ 為佈於論域 Ω 上的 k 個語言變數，$\{a_1,a_2,\cdots,a_m\}$ 與 $\{b_1,b_2,\cdots,b_n\}$ 來自兩不同模糊母群體 A,B 之兩組模糊區間隨機樣本。且對每個隨機樣本對應語言變數 L_j 均有一標準化之隸屬度 mA_{ij}, mB_{ij} 。$MA_j=\sum_{i=1}^{m} mA_{ij}$, $MB_j=\sum_{i=1}^{n} mB_{ij}$ ，為樣本對語言變數 L_j 之隸屬度總和。

2. 事先假設: $H_0: F\mu_A =_F F\mu_B$; A, B 兩母體有相同之分配比率。

3. 統計量：$\chi^2 = \sum_{i=A,B}\sum_{j=1}^{k}\frac{([Mi_j]-e_{ij})^2}{e_{ij}}$. (e_j 為期望次數，為了符合軟體計算 χ^2 檢定要求，我們表格之各細胞隸屬度總和用 4 捨 5 入以取得整數值。)

4. 決策法則：在 α 顯著水準下，若 $\chi^2 > \chi^2_\alpha(k-1)$ ，則拒絕 H_0.

例 3 某政黨競選總部在屏東縣長選舉期間分析選情, 想了解選民的性別對政黨政營的支持度比率是否相同, 於是委託選舉研究中心探討選民的性別對政黨的支持度比率是否有差異, 進行兩種問卷調查:1.傳統勾選一項，2.模糊隸屬度選項。結果如下表 3：

表 3 選民的性別對政黨政營的支持度

	政黨的支持度			χ^2 齊一性檢定	政黨的模糊支持度			χ^2 齊一性檢定
類別	泛綠	泛藍	其他	$\chi^2=13.43$	泛綠	泛藍	其他	$\chi^2=2.41$
男性	220	280	100	> 5.99	256.3	229.4	114.2	<5.99
女性	180	140	80	$=\chi^2_{0.05}(2)$	151.7	161,2	87.1	$=\chi^2_{0.05}(2)$

統計假設 H_0 ：選民的性別對政黨的支持度比率相同。H_1:選民的性別對政黨的支持度比率不相同。

在 $\alpha=0.05$ 顯著水準下，可以看出選民性別對政黨的支持度比率結果之差異。若應用傳統回答法，則 χ^2 齊一性檢定結果為拒絕 H_0 。而若用模糊隸屬度回答，則 χ^2 齊一性檢定結果為接受 H_0 。可以觀察出來，用累加模糊隸屬度回答過程中，將隱性隸屬度考慮進去，會造成總隸屬度值與僅頭一票方法之差異。

四完成之工作項目及成果。

本研究計畫提出應用軟計算方法與模糊統計檢定於市場調查研究，期望能對社會一些公共議題，或商業行爲的表達作更合理統計檢定與統計決策的分析。經由初步實證研究發現，應用模糊統計分析於解決區間樣本與離散模糊樣本上，有實際意義與貢獻。並可讓我們更瞭解市場狀況、更能掌握等待時間、當然也能更清楚地掌握市場趨勢。

完成之工作項目及成果如下:

(1) 人類的思維主要是來自於對自然現象和社會現象的主觀意識，而人類的知識語言也會因本身的主觀意識、時間、環境和研判事情的角度不同而具備模糊性。模糊檢定設計的應用，讓模糊統計學更進一發輝功能。

(2) 建構模糊綜合指標，模糊檢定與決策分析，及探討出一些有趣的性質並與嚴謹數學證明。

(3) 從過去講求嚴謹精密原則的觀點來看，模糊理論也許會讓人覺得是在開倒車，甚至有人誤認說這是一門冒牌的馬虎的學問。甚至也有人主張，既然人類所有的情感皆是模糊的，就根本不該拿來當作研究對象。但是模糊理論並不只是與人類本身有關而已，以往那些嚴密的技術之所以會陷入僵局，是因爲若要爲該複雜的系統下嚴密的定義與敘述，需要耗費龐大的時間與精力來使其程式化，但語句總有詞窮，定義也有詭論時，或者辛苦寫出的程式，也常因爲太過複雜而無法求出答案。讓人類思維得以更合理且完善地比較出來，研究分析，應用模糊邏輯將是一優選擇。

本研究研究計劃中，可能遇到在一些待解決的問題，如

(1) 連續型模糊檢定量，僅討論隸屬度函數型態屬於均勻分配與單峰對稱兩種的情況，對於其它函數型態，如 S-函數、Z-函數、梯形函數與高斯函數等，尚須多加探討。

(2) 在基本的回歸參數檢定統計方面，尚可針對模糊期望值、與模糊變異數等，深入研究。但若能應用電腦軟計算將會精準許多。

(3) 在基本的模糊回歸模式參數檢定統計方面，尚可深入研究。