

明新科技大學 103 校內專題研究計畫成果報告

模糊子集合衡量在祛除資料異常值影響之應用探討

Survey on Application of Fuzzy Subsethood Measure in
Elimination of Outliers Data

計畫類別：任務型計畫 整合型計畫 個人計畫

計畫編號：MUST-103-工管-7

執行期間：103 年 1 月 1 日至 103 年 9 月 30 日

計畫主持人：洪振創

共同主持人：

計畫參與人員：邱志瑋

蔣居洲

處理方式：公開於校網頁

執行單位：明新科技大學工業工程與管理系

中 華 民 國 103 年 10 月 31 日

公開授權書

(提供本校辦理紙本與電子全文授權管理用)

本授權書為明新科技大學校內專題研究計畫成果報告授權人：洪振創

在明新科技大學 管理 學院 工業工程與管理 系所

103 年度校內專題研究計畫。

研究計畫編號：MUST-103-工管-7

研究計畫名稱：模糊子集合衡量在剔除資料異常值影響之應用探討

計畫類型：個人計畫

執行期限：103 年 1 月 1 日 至 103 年 9 月 30 日

茲同意將授權人擁有研究之上列成果報告；紙本授權全文公開陳列於本校圖書館，為學術研究之目的以各種方法重製，或為上述目的再授權他人以各種方法重製，不限地域與時間，惟每人以一份為限；成果報告之電子檔（含摘要），本校圖書館保留以供文獻典藏使用，但可依使用權限授權於網路公開，提供讀者非營利性質之免費線上檢索、閱覽、下載或列印。

成果報告之電子檔案使用權限授權，請勾選下列一項：

- 校內外立即公開全文(含摘要)
- 校內外立即公開摘要，校內立即公開全文，一年後校外公開全文
- 校內外僅於公開摘要，校內立即公開全文，校外永不公開全文

授權人：洪振創

(親筆正楷簽名)

E-Mail：echon@must.edu.tw

中 華 民 國 103 年 10 月 31 日

摘要

傳統績效評估模式，往往是在確定的決策環境下進行，在此條件下所得到的衡量資料，均是可量化的、確定不變的；然而在實際評估一個組織或個人績效時，所面對的決策環境卻往往是主觀的、不可量化的、不確定的、群體決策參與。為了避免在績效評估過程中產生偏差及造成獨斷，集合多位專家以群體決策的方式取得共識，是績效評估過程中最佳的選擇方案之一。

傳統的問卷調查經常透過李克特式量表，以整數級距（1、2、3、4、5）作為定性衡量等第（例如：非常不同意、同意、普通、同意、非常同意）的計分方式，再將個別衡量結果以普通平均方式加以計算，所得到的平均值即為最後的共同結果。為解決傳統問卷調查量表的缺點，利用三角模糊數或梯形模糊數以表示衡量語意措詞之隸屬程度，以李克特式五點量表為例，例如以(3, 4, 5) 三角模糊數代表“同意”的衡量結果，以(3, 4, 5, 5) 梯形模糊數代表介於“同意”與“非常同意”之間的衡量結果。

本研究計畫探討明確值、模糊值等不同的資料衡量方式，利用問卷調查進行資料收集，以作為個別專家準則績效衡量資料之描述；在進行專家績效衡量之集成運算前，對於明確值衡量資料之極端值而言，可以藉由統計方法加以偵測，對於模糊值衡量資料而言，則以模糊子集合衡量方法，作為決定決策者權重設定的方法，此方法將以降低潛在極端值的影響權重，達到剔除極端值的相同（相近）效果。

本研究計畫一年為期，主要的研究重點方向為：群體決策方法，明確值、模糊值的績效衡量方式，群體決策方法在績效衡量上的應用，明確值、模糊值之衡量資料的表示方式與集成方法，應用李克特尺度量表之問卷調查表設計，並以問卷調查資料分析為案例，以驗證系統的正确性與可行性。

關鍵詞：績效衡量，群體決策，模糊子集合衡量，模糊資料集成，極端值。

Abstract

In the situations of performance measurement, decision makers (DM) usually set up their performance rating under the circumstances (1) subjective judgment (2) qualitative measurement (3) uncertainty (4) group decision making. In order to consolidate each individual performance rating into summarized results, data aggregation method should be well selected both in respects of crisp data as well as fuzzy data.

Likert Scale questionnaire is the common tool used for collection of individual judgment, and the calculated arithmetic average is assumed as the summarized results of investigated group. For crisp performance measuring data, outlier elimination methods should be applied as the filters of “bias data” to enhance the reliability of performance measurement. This research will propose Fuzzy Subsethood Measure as the “Outlier Elimination” method for fuzzy data analysis, which can reduce the impact weighting of “Outlier” , thus adjust the calculated results escaping from the bias.

The main purpose of this research is going to survey theory of group decision making, and data aggregation methods applied in scopes of performance measurement. The measuring data for criteria rating is collected in format of either crisp data or fuzzy number (expressed in types of triangular or trapezoidal fuzzy number). Data aggregation methods surveyed in this project include, simple arithmetic mean and geometric mean used in crisp data aggregation, fuzzy weighted average and fuzzy subsethood measure processed in fuzzy number.

This research is a one-year’s survey project, the major contents will focus on group decision making, subjective and objective group consensus methods applied in criteria rating, and data aggregation for both crisp data and fuzzy number. Also an empirical performance measurement system, collected with Likert Scale questionnaires, would be just the real case study for purpose of system validation and accuracy assured..

Keywords: Performance Measurement, Group Decision Making, Fuzzy Subsethood Measure, Fuzzy Data Aggregation, Outlier.

目錄

摘要.....	I
Abstract.....	II
第一章前言.....	1
1.1 研究目的.....	2
1.2 研究架構.....	2
第二章文獻探討.....	4
2.1 績效評估.....	4
2.2 群體決策.....	5
2.3 李克特態度量表.....	7
2.4 李克特統計分析方法介紹.....	9
第三章 研究方法.....	11
3.1 資料衡量方式.....	11
3.2 模糊理論介紹.....	12
3.3 李克特模糊尺寸分析.....	17
3.4 模糊子集合衡量方法.....	18
3.3.1 模糊子集合.....	18
3.3.2 模糊子集合衡量彙整法.....	19
第四章 案例應用.....	22
第五章 結論.....	25
參考文獻.....	27
計劃成果自評表.....	29
運用於教學成果記錄表.....	30

第一章 前言

為了避免在績效評估過程中產生偏差及造成獨斷，集合多位專家以群體決策的方式取得共識，是績效評估過程中最佳的選擇方案之一；Herrera et al.(1996)對群體決策(Group Decision Making)的定義：(1)二位以上成員參與決策，每人均有各自的認知、態度、動機與人格特質；(2)存在一個共同認知的問題；(3)期望達成共同的決策。群體決策是由一群人透過某種模式共同作成決策，一般而言群體可以是由委員會、任務小組、評議會、研究小組等方式形成，而如何達成群體共識(Group Consensus)則是群體決策問題中的一項重要課題；在決策的過程之中，需要考量到決策的整體性，同時要彙整各個決策成員的意見，需要同時考慮到整體性和個別意見。

決策可以區分為定性與定量決策兩大類型，Forman 與 Peniwati (1998)認為主要是透過(1)彼此間的討論獲得共識；(2)表達出個人的偏好這兩種方式。而決策又可分確定狀況下之決策問題(decision under certainty)、風險下之決策問題(decision with risk)、完全不確定狀況下之決策問題(decision under strict certainty)三種。在這種情形之下，我們對於個人的決策判斷，必須去加以整合，而存在的整合方法也有很多種。在最常見的整合方法中，主要可以大致區分為兩種：(1)集成專家個人的判斷(Aggregation of Individual Judgments)，(2)集成專家個人的權重值(Aggregation of Individual Priorities)，這兩種方法各有其適用之範圍，同時也有使用上之限制；採用 AIJ 或是 AIP 的意見集成方法，是依據決策群體的成員究竟是集合在一處進行決策，或是分散於各處做出個人之決策。

問卷調查的資料分析是屬於 AIP 的意見集成方法，傳統的問卷調查經常透過李克特式量表，以整數級距(1、2、3、4、5)作為定性衡量等第(例如：非常不同意、同意、普通、同意、非常同意)的計分方式，再將個別衡量結果以普通平均方式加以計算，所得到的平均值即為最後的共同結果，此種衡量與計算方式有其簡單、方便的優點。

李克特式量表與語意差別量表計分皆為次序、等距整數方式，有其便利明確之優點。但是填答者因受限於只能在數個回答項中勾選出一個答案，容易迫使填答者扭曲自己感受而強迫回答；且不同填答者雖選擇相同的語意措辭，因認知感受的尺度不同，其量化計分卻不一定相等。應用李克特式量表的傳統問卷調查，填答者只能在數個事先設定好的項目中勾選出一個答案，有時會迫使填答者忽視自己的真實感受而採取扭曲的回應，而且不同填答者雖選擇相同的語意措辭，也會因認知感受的尺度不同，限定採取強制性單一等第的選擇，往往會造成判斷上的偏差。

為解決傳統問卷調查量表的缺點，利用三角模糊數(T.F.N.: Triangular Fuzzy Number)或梯形模糊數(Tr.F.N.: Trapezoidal Fuzzy Number)以表示衡量語意措詞之隸屬程度，以李

克特式五點量表為例，下圖所示為不同衡量等第的語意變數 所對應的三角模糊數；當專家進行績效衡量時，可以依據個別專家對問卷題目的感受程度，例如以(3, 4, 5) 三角模糊數代表“同意”的衡量結果，以(3, 4, 5, 5) 梯形模糊數代表介於“同意”與“非常同意”之間的衡量結果。

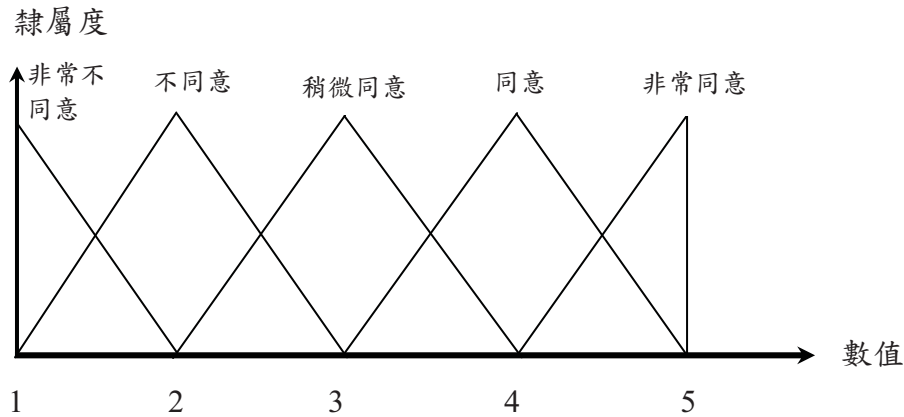


圖 1.1 對稱三角模糊數

1.1 研究目的

在統計學中我們最常使用的就是常態分佈，例如，我們的身高、體重、考試分佈、家庭年平均收入....等等。但統計了那麼多數據，該如何解析？當出現一些異常值或極端值要如何來評估及截取呢？在專家績效衡量整合計分上，最常被使用的方法為算術平均法，然而利用算數平均法計算所有專家的共同衡量結果，容易受到極端值(Outlier)的影響，而如何剔除極端值(Outlier Elimination)。對於明確值衡量資料而言，可以藉由統計方法加以偵測，應用在剔除極端值的常見方法有標準化分數法、Hampel identifier 法、截尾平均值法、Winsorized 平均值法、盒狀圖及 Dixon 考驗法；對於模糊值衡量資料而言，提出以模糊子集合衡量 (Fuzzy Subsethood Measure) 方法，以作為決定決策者權重設定的方法，此方法將以降低潛在極端值的影響權重，達到剔除極端值的相同 (相近) 效果。

1.2 研究架構

本研究計畫針對圖1-2所示之的研究架構，主要的研究重點 方向為：探討群體決策的概念與方法，介紹明確值、模糊值的績效衡量方式，探討不同群體決策方法在績效衡量上的應用；探討明確值、模糊值等不同衡量資料的表示方式，對於明確值資料，將探討適用的統計分析方法加以偵測，對於模糊值資料，應用模糊子集合衡量法剔除極端值；並以「K 公司生產品質統計資料分析」為案例，以驗證系統的正确性與可行性。

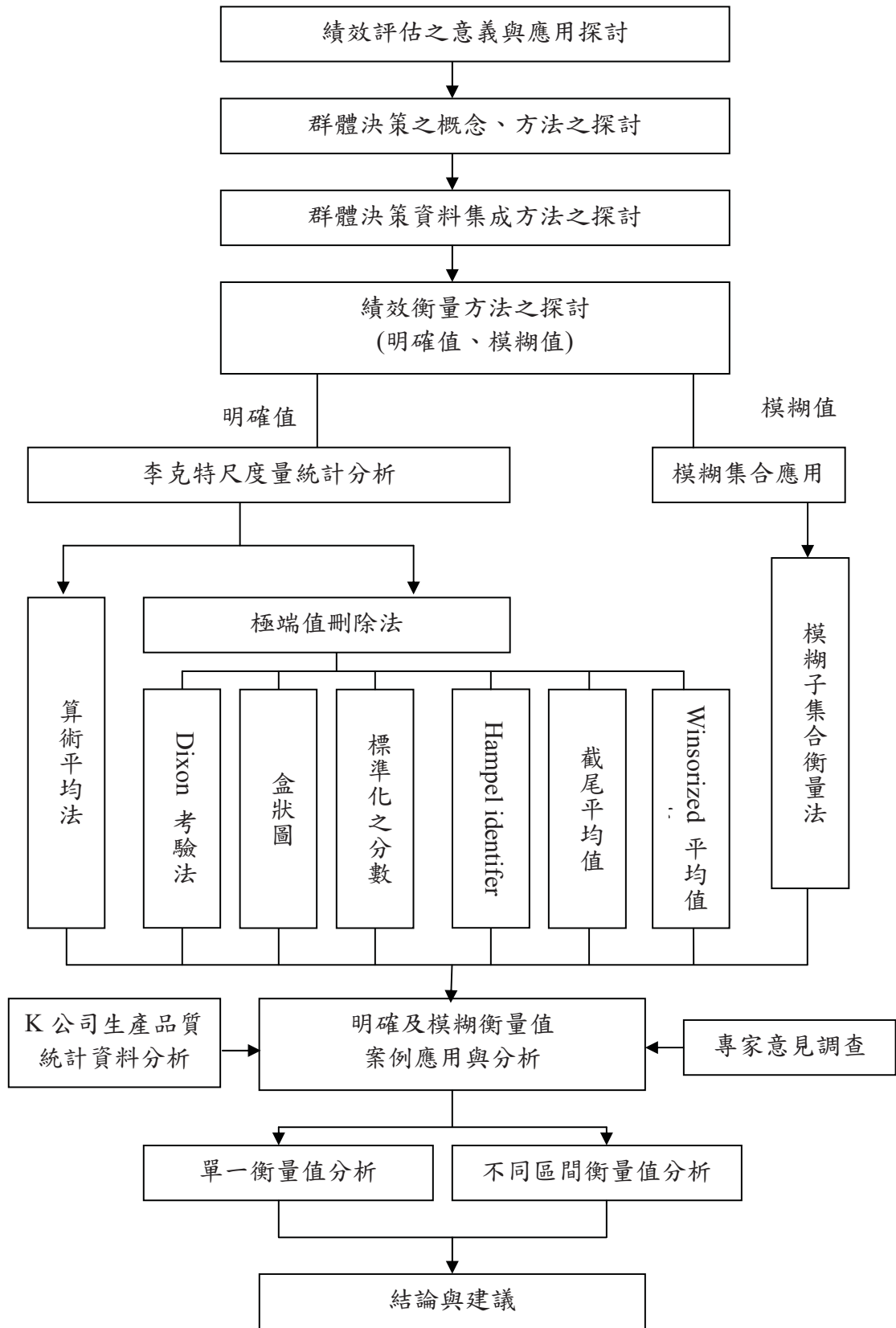


圖1.2 研究流程圖

第二章 文獻探討

績效評估決策環境往往是主觀的、不可量化的、不確定的、群體決策參與，為了避免在績效評估過程中產生偏差及造成獨斷，集合多位專家以群體決策的方式取得共識，是績效評估過程中的主要考量因素，本章將針對績效評估、群體決策、李克特態度量表分別加以探討。

2.1 績效評估

企業管理大師彼得杜拉克(Peter F. Drucker)認為，績效評估(Performance Evaluation)或稱績效衡量(Performance Measurement)是用來衡量組織及員工的工作表現，並將結果用來了解企業現有的人力資源提供員工工作回饋、激勵員工、協助員工生涯發展及組織未來的策略方向等功能。Evans et al. (1996) 認為績效評估是管理中之控制功能，而績效評估及績效管理更有助於公司能更有效的管理資源，以及衡量並控制目標，所以績效評估制度是一個結合獎酬制度的衡量方式，短期的衡量應具有日常作業產能的控制系統及目標修正的功能，長期之下則為策略管理、規劃及達成之工具。

張火燦(2000)認為績效評估對企業經營具有正面的意義，若能善用績效評估的工具及手法，正確有效地評估企業績效，並做適當的回饋，則企業的管理制度就有好的根基，也才能發展組織功能，進而掌握企業核心競爭能力，達成企業經營目標。Tesoro 與Tootson (2000)認為績效評估可協助企業對於欲完成的目標，衡量其成功程度；對組織改造，提供建議或修正措施；提供回饋機制給管理者；評估內部的輸入與產出。

許士軍(2000)認為績效評估在管理活動中具有控制的功能，具有消極和積極之意義，消極意義係了解規劃之執行進度與狀況，如有差異並達到一定程度時，即應採取修正的因應對策；積極意義則是希望藉由績效評估制度之建立，能在事前或活動進行中，對於行動者之決策與行為產生影響或引導作用，使其個人目標和組織目標能趨於一致，即為「目標一致化」(Goal Congruence)的作用。

歸納專家學者對績效評估的看法，績效評估可分為兩個層次，包括組織的績效評估與員工的績效評估。前者是建立組織的管理系統，後者是建立對員工的回饋系統，兩者必須相輔相成以發揮功能。於本研究所指的績效評估制度，意指員工的績效評估，以教師的績效評估為主。

何永福、楊國安(1995)認為有四種不同的考評途徑：

(一) 絕對標準法 (Absolute Standards)：絕對標準乃是對員工工作訂定一固定標準，以

員工績效與所訂定標準相比較之下而予以評定等級，所以考核基礎是依員工的績效表現為主。使用此方法時，員工相互之間是不做比較的，常見的絕對標準有三種：特徵評核法（Trait-rating Scale）、行為定向評核法（Behaviorally-anchored Rating Scale）及行為觀察評核法（Behavior Observation Scale）。

（二）相對標準法（Relative Standard）：相對標準是針對員工績效與其他員工相互比較而評定何人表現較佳，何人表現較差，所以考核基礎是以員工之間的比較為主。

（三）目標擬定法（Objective-based Approach）：目標考評顧名思義就是評估者用一套設定好的目標對員工進行考評，所以考核基礎是以結果導向的考評方法。

（四）直接指標法（Direct Indexes Approach）：根據一些重要的指標如生產效率、出缺席狀況為根據，判斷員工的工作表現。

2.2 群體決策

在現實生活中，決策通常不是由一個人所進行的，而是由一個群體一起進行。而群體決策之優點是可集眾人之力，集思廣益，收集各方意見，經投票或表決，將風險降至最低，可謂納百川匯集為一流。而群體決策大多是由專案團隊、任務團隊、委員會、或其他類似的團體，以會議或面對面的討論方式來共同完成決策案。一般我們是在幾個特定的或多個方案之中，去挑選出一個最合適的方案，而且這個方案是可以達成我們的特定目標。在傳統上，主要由部屬或群體先收集資料，最後會以一個具有較高職位或較大權力的人來做出決策，而其他則只是擔任輔助決策的角色。但是，在面對大型且複雜的問題時，特別是在組織決策的環境之下，之中可能牽扯到資源的分配與效率的考量等眾多因素，單一決策者要做出決策的困難度便大幅的增加了。

Robbins (1994)認為相對於單一個體的決策過程，群體決策過程具備以下顯著的優點：

- (1) 能夠集思廣益，提供更多的資訊：透過群體互動方式可以產生更多有效的方案，以作為解決問題的參考，藉以提高決策績效，提昇決策品質。
- (2) 促進群體共識：由於決策程式涉及相關群體決策結果也具有共識性，故決策本身較不易被質疑或推翻，在理論上的合法性與實務上的接受度都可提昇。
- (3) 能夠增加決策的合法性與接受度：符合民主的潮流，透過民主程序制定的決策，其正當性高於獨自決策。
- (4) 能夠避免個人決策容易受自私、人情和個人權力等因素的影響。

洪振創與蔡千姿(1998)提到，組織決策通常不是單純的由某一個人所完成的，常見的方式是藉由群體決策的方式，來做出組織決策。群體決策指的是由若干人透過某種模式共同作成決定，一般而言群體可以是由委員會、任務小組、評議會、研究小組等方式形成，

群體決策可以讓個別決策者之間意見進行交流，在群體連續的互動過程中不斷地改變與調整個別的意見，經由群體間的意見交流過程而改善決策結果，而非單純屬於個人意見的集合體。

大部分的組織決策是要做出一個整體考量的，不論決策的過程是公開或不公開的、決策範圍是只限於內部或包含外部的，我們都要在決策的過程之中，考量到決策的整體性，正由於這樣一個需要整體考量的結果，通常我們很難去達到群體共識，另一方面，要去匯總各個決策成員的意見，也變得困難許多。

接下來介紹三種常見的群體決策的方法：

1.腦力激盪法(Brainstorming)

由Alex Osborn(1938)在其著作”How to Think u”一書中所提出，其特色為腦力激盪法為一種激發創意的群體決策模式，並鼓勵奇想、歡迎搭便車，參與人數以 5-12 人較佳，時間也不宜太長，且主持人不能有任何指導或限制之言論，以免壓抑了決策者之創意。

2.名目群體技術(Nominal Group Technique,NGT)

是由Delbacq 和 Van de Van 在 1968 年從社會心理、工業工程與社會工作等研究中與問題規劃、設計的相關領域中所發展出來的群體決策技術(Delbacq et al., 1975)。主要的目的是用來解決組織內長期會議成效不彰的問題。

名目群體技術是一種結合腦力書寫、腦力激盪與投票(voting)技術所發展出來的方法，可以匯集群體意見並決定意見權重的方法，每一參與者先經由完全獨立的思考方式寫下自己對問題的看法，然後以輪流循環(round-robin)方式向主持人提出不同於他人的意見與看法，就在他人提出意見與說明的時候，又可以將自己的新見解寫在紙上，如此循環直到所有的意見發表完，然後由所有參與者依照個人看法，自由地選取固定個數的意見(如 6 個)，並賦與不同的權重(如 6,5,4,3,2,1)，經匯總後則可以作為方案選擇之依據。名目群體技術並不試圖去達成一致的共識 (Greenberg and Baron 1995)。名目群體技術的優點在於它讓群體正式集會，但卻不像傳統的互動群體一樣，限制了個人獨立思考的能力。Dunham(1991)曾整理名目群體技術的優缺點如下表：

表 2-1：名目群體技術的優缺點

優點	缺點
1. 成員間平等參與 2. 個別成員的影響力相等 3. 比開放討論的群體產生更多且更有創意的意見 4. 與會者有更高的滿意度 5. 減少一般面對面會議所面臨的社會	1. 需要額外的事前準備 2. 傾向侷限在單一主題；單一目的的會議。很難在會議中更換議題。 3. 需要所有人都同意使用名義群體技術(可能有人會持反對意見)

影響 6. 促進與會者聚焦在解決問題上，而非人身攻擊 7. 更有會議結束即達成決議的感覺	
--	--

資料來源：Dunham, 1991

3. 德爾菲法(Delphi Method)

德爾菲法是 1950 年代美國蘭德公司(Rand Corporation)由Dalkey & Halmer et al.發展設計而來(Dalkey & Halmer,1963)。透過問卷調查的方式徵詢專家意見，以取得專家之共同意見，並以統計方法予以呈現。設置專門找碴之角色，即異議份子(devil's advocate)，組織為避免在群體決策中，出現了少數壟斷及群體思考之負面作用，於是在組織中設計了一個提供不同觀點、創意性方案、批判性思考、挑戰現況的角色。優點為可提升群體決策之品質，但卻會降低群體決策之速度，因郵件之往返及整理需耗費過多時日。

決策問題是專家學者持續一直在研究的課題，以往之研究只針對單準則形態問題研究，而在現今環境快速變動之環境下多，決策者面對的的決策問題是複雜多變的者，往往是無法運用單一準則就可以解決，而是要將同一個決策問題所屬多種評估準則都納入考量，並依此做出最適當的決策，在現今是。也因此，多準則決策方法(Multiple Criteria Decision Making ; MCDM)成為現今常被決策者使用的方法。

2.3 李克特態度量表

在本研究中利用李克特態度量表來作為問卷調查統計；態度量表是用來測量受訪者對一情境或議題多種面向的態度，並提供結合對不同面向的態度成一全盤性指標的技術。此方法降低了受訪者僅受其對該情境議題之一或二面向意見的影響，而做出之意見表達風險（胡龍騰等，2000）。態度量表包括瑟斯（Thurstone type）量表、李克特（Likert type）量表、古特曼（Guttman scale）量表等方式。

李克特量表又稱為總加量表法，此法是由一組句子所構成，這組句子是以圍繞所要測量的問題集到的眾多句子，採用項目分析方法篩選出辨別力較強的句子組成，根據受測者對這組句子的各項回答，使用總和計分方式，以判明其態度的強弱，而其評分方法是假設每項題目比重相等，受測者按照其同意與不同意的程度選擇，若以高分表示喜好，正向題回答「非常同意」將可得5分，回答「非常不同意」則只得1分，而且負向題必須逆轉計分。其編製程序如下：

步驟一 組織或建構反映問題中主要議題之態度敘述。敘述應以文字表達對議題反映正向與負向之態度，亦即贊成或反對該議題。（若你的量表為單向度，則只需正向敘述）。

確定所有敘述皆與主要議題具有邏輯性的聯結。亦需決定期望受訪者以類別或數字量表回答該問題。

步驟二 針對小群體進行這些敘述之測試。

步驟三 針對回收之答案，以加權方式分析答案，量化價值。量化價值依正向與負向敘述有所不同。對正向敘述而言，答案具有贊成態度者則賦予最高分數。

步驟四 依受訪者對每項敘述所勾選之答案，以步驟三賦予量化價值，計算每一受訪者之態度分數。

步驟五 比較所有受訪者每一項目的分數，以界定無區別項目。當幾乎每一為受訪者以相同的答案回答相同的問題，則無區別敘述無助於區分受訪者的。

李克特式量表的結構共包括兩部分：第一部份是對測量主題所編擬的許多積極和消極的態度敘述句；第二部分則是用來表示看法的評分表，分成「非常同意」、「同意」、「稍微同意」、「不同意」、「非常不同意」五等級，並依序給予分數。例如，在教學評量問卷上，表達偏好評比使用李克特尺度可表示如下：

表2.2 李克特量表範例

問題	反應尺度				
	非常不同意	不同意	稍微同意	同意	非常同意
教師上課的態度認真，具教育熱忱。	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
教師上課很少遲到、早退或中途離席。	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
教師對學生考核及評分公正合理。	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
...					

在眾多測量工具中，李克特氏總加量表是較為人所熟知的且常為研究者所採用（王文科，1996）。此量表自從1932年Likert 發表以來，因為有著尺度建構方便明確及問卷編製簡單的優點而被廣泛使用至今（葛樹人，1988）。

李克特式量表評估等級有很多，有最少的兩點評量，甚至到十點以上的李克特式量表都有，不過這是後來發展出來的，最原始的李克特式量表只有三點和五點的評量等級，其中的五點評量等級更是廣為大家所使用。李克特式量表五點量表法，Nunally(1978, p595) 第二版心裡計量 (Psychometrics) 一書中，提及選項數的多寡會影響信度的大小，當選項數從2點增加到7點的時候，信度增加最快，11點後信度增加非常的小(王朝正，2002)；根

據Berdie(1994) 研究指出「在大多數的情況下，五點量表正好可以表示溫和意見與強烈意見間之區別是最為可靠的，選項超過五點，一般人難有足夠的辨別力。」當量表的選項愈多，答案分佈的情形就會愈廣，變異數也會變得愈大，進而影響可信度；另外，較大的選答變異數，表示也會有較大的抽樣誤差，就統計資料分析而言，是不該有太大的變異數(彭仁信，1994)。同時，彭仁信(1994)的研究結果也指出，五點量表的態度反應是明確的。

2.4 李克特統計分析方法介紹

在傳統上都是採用算術平均數來彙整專家的意見，但唯一的缺點就是容易受到極端值的影響，而削弱平均數的代表性，也使得用算術平均數來彙整變得不客觀。

極端值又稱為離群值，一般是指某一個觀察值與其他觀察值的數值呈現很大的差異，所以懷疑該觀察值與其他觀察值並不是經由同一機制所產生(Stevens, 1990)，因此該觀察值的可信度有待驗證。

本研究介紹三種以敘述性統計為基礎的極端值判斷方法，包括標準化分數法、截尾平均值法及Winsorized平均值法，分別敘述如下：

(一) 標準化分數法

標準化分數法是假設某變數資料符合常態分佈的情形，是將資料轉成標準化分數或Z分數來進行判斷；換言之，先算出每筆資料離開平均值的距離，即所謂的離均差分數(deviation score)，再除以該變數的標準差。根據常態分布的性質，約有99%資料的分數會落在平均值 ± 3 個標準差之內，因此有一些文獻會將Z分數大於3或小於3的數據視為極端值(例如Shiffler, 1988; Stevens,1990)，此外，也有學者以Z分數大於4 或小於4，或者是其他數據作為切割點(Shiffler, 1988)。以Z分數大於某一個切割點作為極端值判斷的方式，其優點是辦法快捷、道理易懂，缺點則是此法比較適合應用在符合常態分布的變數上。

標準化分數方法的優點是方便及簡單明瞭，但缺點是相當不可靠，其中一個原因是，當資料中有極端值存在時，其標準差的值可能會很大，所以相對地，Z分數會顯得小，沒有超過 ± 3 這個門檻，所以檢測不出極端值。

(二) 截尾平均值

如果資料之中含有極端值，則該組資料的平均值很可能會偏小或趨大，此時通常建議使用是中位數可能比平均值更有代表性。然而，平均值的計算亦可用作判別資料中是否有極端值的存在，其判斷的方式可以透過將所有資料的平均值與截尾平均值(trimmed mean)做一比較。截尾平均值的基本概念是，希望能夠將極端值先刪除，再用剩餘的數據來估算

出資料的平均值，該值可能更能代表該組資料的平均表現。此種將最大值或最少值捨去的處理方式，生活中可能有不少實際應用的層面，例如奧運會的花式溜冰比賽，就是採用截尾平均法去除最高及最低的成績後，再計算平均得分。

當計算截尾平均值時，需要先把所有的數據資料依大小作排序動作，假設有極端值存在時，極端值會出現在序列的兩末端部份，接著需要刪除兩個末端某一比例的資料，再計算餘下數據的平均值。假如所有資料的平均值與截尾平均值差異很大，即反映出原資料可能存在極端值的問題。

截尾平均值法最關鍵的問題是決定需要捨去多少比例的資料。設 n 代表樣本數，且 $G=[\gamma n]$ ，其中 γ 是介於0至0.5之間的一個常數，括號則代表高斯函數，取 γ 乘 n 積之整數部份。截尾平均值是指刪除最小的 G 個及最大的 G 個後，再求其餘數據之平均值。例如當 $\gamma=0.1$ 時，指的是求刪除前後段各10%資料的截尾平均值。例 $n=21$ ，故 $G=[(0.1)(21)]=[2.1]=2$ ，即刪除最大2個值與最小的2個值，即保留18筆資料求平均值。

(三) Winsorized平均值法

有鑑於截尾平均值的計算非常浪費辛苦蒐集所得的資料，因此提出Winsorized平均值的方法，此法是在兩端各刪除 G 筆數據之後，以剩餘數據的最大值及最小值各補上 G 次，使原來的變數仍維持 n 筆數據，再求其平均值。

第三章 研究方法

專家績效衡量整合計分上，最常被使用的方法為算術平均法，然而利用算數平均法計算所有專家的共同衡量結果，容易受到極端值(Outlier)的影響，對於明確值衡量資料而言，可以藉由統計方法加以偵測，應用在剔除極端值的常見方法有標準化分數法、Hampel identifier 法、截尾平均值法、Winsorized 平均值法、盒狀圖及 Dixon 考驗法；對於模糊值衡量資料而言，提出以模糊子集合衡量 (Fuzzy Subsethood Measure) 方法。

3.1 資料衡量方式：明確值、模糊值

在一般生活之中，許多時候我們所做出的判斷都是以明確值的方式來表現，例如：飲料容量700cc、及格分數60分、薪水30000元...等，這類的描述是以”單一明確值”的方式呈現；但是，在日常實務中有些時候，往往無法以明確值來表示，或是以明確值的表示方式並不足夠，例如：根據以往的產量記錄推估預測今天的產能最多、平均、至少產量值...等，此資料呈現是模糊值的方式。

在群體決策之中，專家所做出之判斷值，常常會有明確值與模糊值的分別，主因在於有時候人類的思考上，並不能完全的對一件事做出精準的判斷，其中存在著一個思考與想像的空間，而這個空間則可利用模糊的概念來描述。本研究中，我們將探討專家在進行李克特式五點量表時，應用明確值、模糊值進行判斷的情形，例如：當選擇為非常不同意時，明確值時可能值為1，而模糊值此時為 [1, 1, 2] 。

由於有以上明確值、模糊值二種資料衡量方式，所以我們在進行專家意見集合時，也需進一步去計算與整理所獲得之資料。以下將以簡圖呈現二種衡量值在模糊集合中的形態：

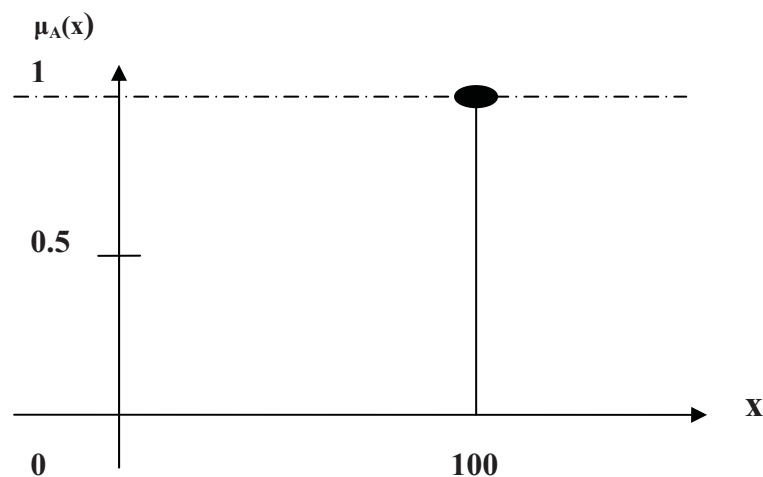


圖 3.1 明確值之歸屬函數值

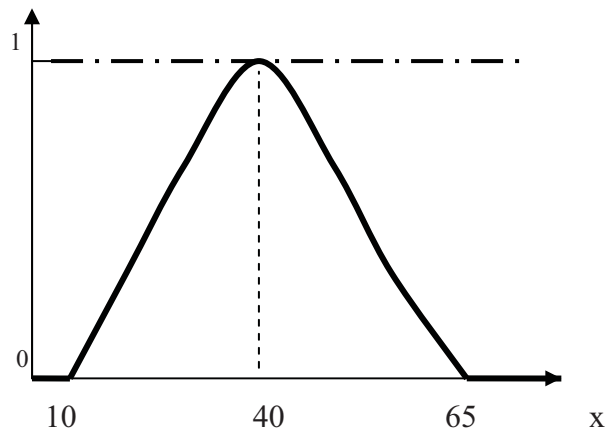


圖 3.2 模糊值之歸屬函數值

從以上二個圖可以看出，其二種資料表示方式的特性，明確值為 x 在 $\mu_A(x)$ 為 1 時之值，以圖 3.1 為例，該值為 100。對模糊值來說， x 在 $\mu_A(x)$ 為 1 時之值，以圖 3.2 為例，該值為 40，但是小於 10 和大於 65 時 $\mu_A(x)$ 為 0，在 10~40 之間，其 $\mu_A(x)$ 值會隨之上升，到 40 時， $\mu_A(x)$ 值為 1，40~60 其 $\mu_A(x)$ 值會隨之下降到 0；其由於明確值為模糊數之特化的性質。

3.2 模糊理論介紹

氣象主播在天氣預報時，通常只告訴我們明天北部地區氣溫是 18 度到 23 度間，屬於陰時多雲短暫雨的天氣，出門記得加件外套及攜帶雨具。在日常生活裡存在很多的不確定性 (Uncertainty) 與模糊性 (Fuzziness) 的資料，然而我們願意接受模糊的資料。

模糊理論是為了解決真實世界中普遍存在的模糊現象而發展的一門學問。早在 1920 年羅素為模糊邏輯立下邏輯基礎，卻沒有繼續發展，直到 1965 年美國加州大學柏克萊分校 (U. C. Berkeley) 的札德 (L. A. Zadah) 教授，在資訊與控制 (Information and Control) 學術期刊上，發表模糊集合 (Fuzzy Sets) 的文章，模糊理論蘊育而生。模糊理論以模糊集合為基礎，主要內容包含了模糊邏輯 (Fuzzy Logics)、模糊推理 (Fuzzy Inference)、模糊控制 (Fuzzy Control) ... 等。模糊理論應用的領域很廣泛，如專家系統、績效評量、影像識別、語言或指紋辨識 ... 等等。

在 Kaufmann and Gupta (1988) 及汪明生與汪仲祥 (2003) 的研究之中整理了對模糊理論的相關概念，以下擷取與本研究有關之部份做一介紹：

(1) Fuzzy Set (模糊集合)

模糊集合其定義為：令 U 為被討論的全體對象，稱為論域 (Universe of Discourse)；論域中的每個對象，稱做元素，以 u 表示； U 上的一個模糊子集 A ，是指：對於任意 $x \in U$ ，

都指定了一個實數 $u_a(x) \in [0,1]$ ，稱為 x 隸屬於 A 的程度。即 $u_a(x)$ 為一映射 (Mapping)：

$$u_a(x): U \rightarrow [0,1]$$

稱為 a 的隸屬函數 (Membership Function)。當 A 值域 = $\{0,1\}$ 時， $u_a(x)$ 蜕化成一個普通子集的特徵函數， a 便成一個普通子集。模糊集合的高度 (Height) 是指最大的隸屬程度 (Degree of Membership)，以 $hgt A$ 表示。至少有一元素之隸屬程度為 1 的模糊集合，稱為正規化 (Normalization) 的模糊集合。在處理實際問題時，把普通集合的絕對隸屬關係加以擴充，使元素對集合的隸屬關係度由如上取 0 或 1「非此即彼」之特性，推廣至可以取單位區間 $(0,1)$ 中的任意一數值，進而實現定量刻畫不確定性問題之模糊性質。

(2) Fuzzy Number (模糊數)

模糊數乃實數 (Real Numbers) 的模糊子集 (Fuzzy Subset)，而且它是代表信賴區間 (Confidence Interval) 觀念的一種擴充，結合可能性分析之 α 水準 (level α Presumption) 與 α 水準信賴區間之性質。模糊數為一不精確值 (Imprecision Number)，與機率論中之隨機變數 (Random Number) 是不同的。數學上具正規化且為凸集合，以具有區段性連續之隸屬函數之模糊集合，稱之為模糊數，亦即模糊數須滿足下列條件：

1. 凸性的模糊子集合 (Convex Fuzzy Subset)
2. 正規化的模糊子集合 (Normality of A Fuzzy Subset)
3. 區段連續 (Piecewise Continuous)

模糊數常用為三角形模糊數 (Triangular Fuzzy Numbers；T.F.N) 與梯形模糊數 (Trapezoidal Fuzzy Numbers；Tr.F.N)。

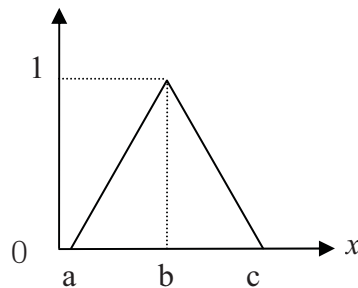


圖 3.3 三角模糊數

若模糊數 A 是三角模糊數，則歸屬函數 $\mu_A(x)$ 為

$$\mu_A(x) = \begin{cases} (x-a)/(b-a), & a \leq x \leq b \\ (x-c)/(b-c), & b \leq x \leq c \\ 0, & \text{otherwise} \end{cases}$$

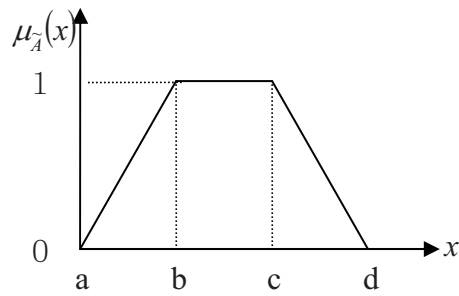


圖 3.4 梯形模糊數

若模糊數 A 是梯形模糊數，則歸屬函數 $\mu_A(x)$ 為

$$\mu_A(x) = \begin{cases} (x-a)/(b-a), & a \leq x \leq b \\ 1 & , b \leq x \leq c \\ (x-d)/(c-d), & b \leq x \leq c \\ 0 & , otherwise \end{cases}$$

由於以圖形表示模糊數的特性相當不方便，因此須改以數學式子表示，任何梯形模糊數 \tilde{A} ，可用 $\tilde{A} = (a | b, c | d) L-R$ 表示， \tilde{A} 定義為：(1) $a \leq b \leq c \leq d$ ；(2) 當 $X \in (a, d)$ ， $\mu(X | \tilde{A}) = 0$ ；當 $X \in (b, c)$ ， $\mu(X | \tilde{A}) = 1$ ；(3) 當 $X \in (b, c)$ ， $\mu(X | \tilde{A})$ 是連續從 0 到 1 呈線性單調遞增；當 $X \in (c, d)$ ， $\mu(X | \tilde{A})$ 是連續且從 0 到 1 呈線性單調遞減。

而三角模糊數 $\tilde{A} = (a, b, d)$ ，即當 $b = c$ 時， \tilde{A} 為梯形模糊數，當 $a = b = c = d$ 時，則 \tilde{A} 為明確值，此外， (b, c) 的區間值為 \tilde{A} 最有可能出現的數值，當決策者所有的資訊愈少，此區間距離也愈大，亦即愈模糊。

(3) α -cut (α 截集)

α 截集是將模糊數轉為明確值的方法，亦即當模糊隸屬度為 α 時， T 所包含的區間集合， $T_\alpha = [(b-a)\alpha + c - (c-b)\alpha]$ ，其中， $0 \leq \alpha \leq 1$ 如圖 3.5

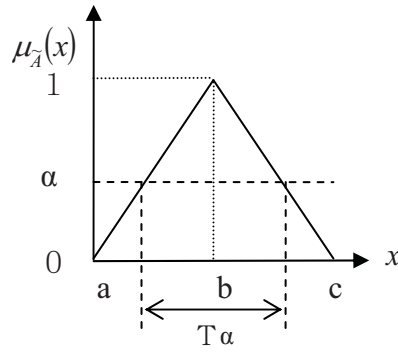


圖 3.5 三角模糊數之 α -截集

(4) 模糊數之基本運算

假設兩正三角模糊數 $A=[a_1, a_2, a_3]$ ， $B=[b_1, b_2, b_3]$ ， $a_1, a_2, a_3, b_1, b_2, b_3$ 皆屬於正實數 R^+ ，其模糊數運算的基本運算式如下：

1. 加法 (+) (Fuzzy Addition)

$$[a_1, a_2, a_3] (+) [b_1, b_2, b_3] = [a_1 + b_1, a_2 + b_2, a_3 + b_3]$$

2. 減法 (-) (Fuzzy Subtraction)

$$[a_1, a_2, a_3] (-) [b_1, b_2, b_3] = [a_1 - b_1, a_2 - b_2, a_3 - b_3]$$

3. 乘法 (\times) (Fuzzy Multiplication)

$$[a_1, a_2, a_3] (\times) [b_1, b_2, b_3] = [a_1 \times b_1, a_2 \times b_2, a_3 \times b_3]$$

4. 除法 ($:$) (Fuzzy Division)

$$[a_1, a_2, a_3] (:) [b_1, b_2, b_3] = \left[\frac{a_1}{b_3}, \frac{a_2}{b_2}, \frac{a_3}{b_1} \right]$$

5. 倒數 (Fuzzy Inverse)

$$\tilde{A}^{-1} = [a_1, a_2, a_3]^{-1} = \left[\frac{1}{a_3}, \frac{1}{a_2}, \frac{1}{a_1} \right]$$

(5) 語意變數 (Linguistic Variables)

在模糊理論之中，對於某些事物或問題，無法單純的以數值來表示時；或是為了讓人們容易理解，有時候可以利用類似：非常有錢、很有錢、有錢；很冷、有點冷、適中等這類描述性的字眼來做為語意尺度，之後利用對這些語意尺度的描述，據以轉化成各個尺度的模糊數，而得到數值化的結果。

語意尺度的決定可為等距的區間尺度，亦可為非等距的區間尺度，端看應用問題的特性而定。Chen & Hwang(1992)提出非等距的區間尺度，利用八種模糊語意變數表，提出一個簡單的方法，將模糊資料用模糊語意詞句加以表示，使能將決策者所給予的語意變數值，

轉換成相關的模糊數。如表3.1所示。亦可以統計方法透過調查推求詞彙的隸屬函數(汪培庄，1990)。

表3.1 八種模糊語意變數表

語意尺度	1	2	3	4	5	6	7	8
語意數	2	3	5	5	6	7	8	11
極低								◎
很低			◎		◎	◎	◎	◎
低至很低							◎	◎
低		◎	◎	◎	◎	◎	◎	◎
稍低				◎	◎		◎	◎
些微低						◎		
中等	◎	◎	◎	◎		◎	◎	◎
些微高						◎		
稍高				◎	◎		◎	◎
高	◎	◎	◎	◎	◎	◎	◎	◎
高至很高							◎	◎
很高			◎		◎	◎	◎	◎
極高								◎

◎：表示選取的語意

資料來源：Chen & Hwang(1992)

本研究採用三角模糊數均勻劃分語意變數範圍的函數圖形，將各影響因素權重之重要度以五個語意尺度來衡量，如表 3.2 所示語意尺度之分佈範圍與其對應之三角模糊數。

表 3.2 模糊尺度與相對應之語意變數定義

模糊尺度 (Intensity of fuzzy scale)	語意變數之定義 (Definition of linguistic variables)
$\tilde{1}=(1, 1, 2)$	非常不同意
$\tilde{2}=(1, 2, 3)$	不同意
$3=(2, 3, 4)$	稍微同意
$4=(3, 4, 4)$	同意
$5=(4, 5, 5)$	非常同意

(6) 解模糊化的方法

解模糊化就是將模糊資料轉為明確的資料，主要是為了方便模糊排序過程中所使用的

工具。解模糊化並無單一方法，要使用哪一種方法，端看問題的特性而定。常用解模糊的方法有下列七種：

一、重心法 (Center of Area Defuzzification)

其觀念就是以模糊集合之「中心值」來代表整個模糊集合，其計算公式如下：

$$F = \frac{\sum g(x_i) * u_a(x_i)}{\sum u_a(x_i)}$$

$g(x_i)$ ：對隸屬度之一個重要性測量權數

$u_a(x_i)$ ：為模糊集合隸屬函數

F：代表模糊集合之重心

對三角模糊數可寫為如下公式所示：

$$F = \frac{[(UR_i - LR_i) + (MR_i - LR_i)]}{3} + LR_i$$

UR_i ：三角模糊數的上限值

二、形心法 (Center of Sum Defuzzification)

三、平均最大隸屬法 (Mean of Maximal Defuzzification)

四、第一個最大值法(First of Maxima Defuzzification)

五、最後一個最大值法(Last of Maxima Defuzzification)

六、最大值之平均值解模糊化法(Middle of Maxima Defuzzification)

七、中心平均值解模糊化法(Center Average Defuzzification)

解模糊化之方法眾多，在過往的研究之中，也有許多不同的手法解模糊。而如何去選擇何種解模糊化方法目前皆無論，端視使用者的運用狀況而定。

3.3 李克特模糊尺度分析

問卷調查的資料分析是屬於 AIP 的意見集成方法，傳統的問卷調查經常透過李克特式量表，以整數級距 (1、2、3、4、5) 作為定性衡量等第 (例如：非常不同意、同意、普通、同意、非常同意) 的計分方式，再將個別衡量結果以普通平均方式加以計算，所得到的平均值即為最後的共同結果，此種衡量與計算方式有其簡單、方便的優點，典型的李克特式量表應用如下：

常見的傳統量表，可分為李克特式量表與語意差別量表兩種型態，其計分方式舉例如下：

李克特式量表					
語意措詞	非常不滿意	不滿意	稍微滿意	滿意	非常滿意
對應計分值	1	2	3	4	5

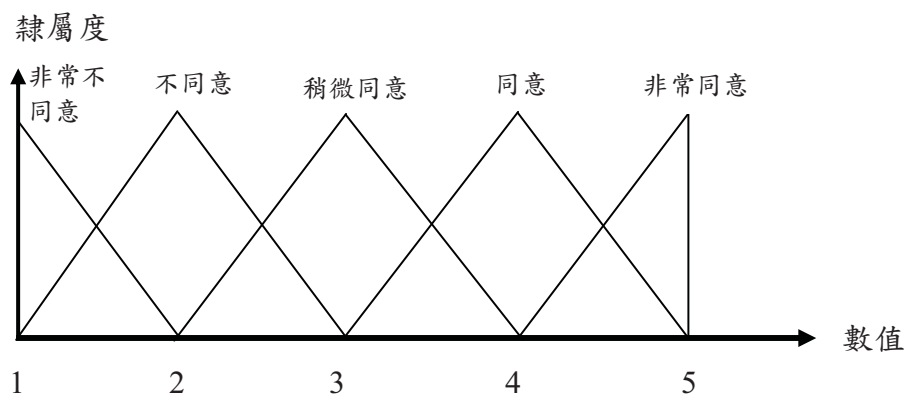
語意差別量表										
語意措詞	非常不滿意	←-----→								非常滿意
對應計分值	1	2	3	4	5	6	7	8	9	10

李克特式量表與語意差別量表計分皆為次序、等距整數方式，有其便利明確之優點。但是填答者因受限於只能在數個回答項中勾選出一個答案，容易迫使填答者扭曲自己感受而強迫回答；且不同填答者雖選擇相同的語意措辭，因認知感受的尺度不同，其量化計分卻不一定相等。因此，相關文獻指出傳統量表的計分尚存有若干缺點（吳柏林，民 83）：

- (一) 人類的思考與行為本來充滿著模糊過程，傳統問卷常被過度的解釋。
- (二) 為迎合數字的精確要求，實驗資料常有被過度使用之嫌。
- (三) 為簡化或降低數學模式的複雜性，卻將實際狀況間之相關與動態特質忽略。

應用李克特式量表的傳統問卷調查，填答者只能在數個事先設定好的項目中勾選出一個答案，有時會迫使填答者忽視自己的真實感受而採取扭曲的回應，而且不同填答者雖選擇相同的語意措辭，也會因認知感受的尺度不同，限定採取強制性單一等第的選擇，往往會造成判斷上的偏差。

為解決傳統問卷調查量表的缺點，利用三角模糊數（T.F.N.: Triangular Fuzzy Number）或梯形模糊數（Tr.F.N.: Trapezoidal Fuzzy Number）以表示衡量語意措詞之隸屬程度，以李克特式五點量表為例，下圖所示為不同衡量等第的語意變數所對應的三角模糊數；當專家進行績效衡量時，可以依據個別專家對問卷题目的感受程度，例如以(3, 4, 5) 三角模糊數代表“同意”的衡量結果，以(3, 4, 5, 5) 梯形模糊數代表介於“同意”與“非常同意”之間的衡量結果。



3.4 模糊子集合衡量方法

Forman 與 Peniwati (1998) 認為，在進行群體決策時，主要是透過(1)彼此間的討論獲得共識；(2)表達出個人的偏好這兩種方式。在這種情形之下，對於個人的決策判斷，必須去加以彙整，而存在的彙整方法也有很多種。在最常見的彙整方法中，主要可以大致區分為兩種：

- (1)集成專家個人的判斷 (Aggregation of individual judgments, AIJ)：屬於定性決策
- (2)集成專家個人的權重值 (Aggregation of individual priorities, AIP)：屬於定量決策

3.4.1 模糊子集合

傳統的集合理論(Set Theory)中，若集合 A 中的每一個元素都屬於集合 B，則稱 A 是 B 的子集，亦可讀作 A 包含於 B，並以 $A \subset B$ 記之。

模糊子集合在模糊集合上是很重要的概念。Zadeh 對模糊集合的「包含」定義，當 $m_A(x)$ ， $m_B(x)$ 分別為 A，B 的隸屬函數且 A，B 為定義於 X 上的模糊集合，對所有屬於 X 的 x 而言只有在 $m_A(x) \leq m_B(x)$ 時，則稱模糊集合 A 包含於模糊集合 B 中 ($A \subset B$)，且程度介於 0 到 1 之間。

Kosko 則主張當 $m_A(x) \leq m_B(x)$ 不等式只在某些 x 不滿足時我們仍可以考慮模糊集合 A 是模糊集合 B 的子集合的程度，以 $S(A,B)$ 表示，稱之為模糊子集合衡量。Kosko 定義如下：

$$S(A,B) = \text{degree}(A \subseteq B) = \begin{cases} 1 & A = 0 \\ \frac{\sum_{x \in X} (m_A(x) \cap m_B(x))}{\sum_{x \in X} m_A(x)} = \frac{M(A \cap B)}{M(A)} & A \neq 0 \end{cases}$$

在Young(1996)、Fan, Xie 和Pei(1999)以及楊宗仁(2002)研究之中整理了對子集合的相關概念，以下擷取與本研究有關之部份做一介紹：

$S(A,B)$ 表示 A 是 B 的子集合的程度，且 S 為一個定義於 $f(X) \times f(X)$ 的測度函數， $f(X)$ 表示定義於 X 上的所有模糊集合的集合。探討如何界定子集合測度，考慮以下的因素大約可粗略的分成兩個觀點：

- (1).子集合的衡量必須與一般集合中的包含關係一致，即當對象是一般的集合時。 $S(A,B) \in \{0,1\}$ 其中 A，B 為隸屬函數非 0 即 1 的一般的普通集合。
- (2).子集合的衡量是一集合包含於另一集合的冪集合的程度。如前所述，從一般集合的觀點，一集合 A 為另一集合 B 的子集合，則 A 會屬於 B 的冪集合之中。由此可立即地推廣到模糊的概念上，A 是 B 的子集合的程度視為 A 屬於 B 的冪集合的程度。

Young 針對Zadeh對子集合的衡量的定義、Kosko 對子集合的衡量與熵之間關聯的探討，以及模糊集中包含時的相對應關係作為子集合的衡量的條件，提出三個子集合的衡量的公設。並連結模糊熵、機率和模糊邏輯。

Fan, Xie 和Pei 對子集合的衡量提出新的定義，以空集合與宇集合的關係及模糊集中包含時的關係，訂定子集合的衡量的公設。並且探討在此公設下的子集度的一些運算是否具有封閉性。最後再運用於聚類方法的評價上。

子集合的衡量的形式並無標準樣式，針對兩集合包含的關係，依據不同的需要，可以定出不同的形式出來。為了確定所定的形式是合理的，符合公設是基本要求。而所得出的數值可以視為是一種資料形式的轉換，將兩個集合之間的包含關係轉變成介於〔0,1〕的量化數值。

3.4.2 模糊子集合衡量彙整法(Subsethood Aggregation Method, SbAM)

本研究針對Kosko(1986)提出兩模糊集合之交集的觀點所定義的子集合衡量為應用的對象。從Kosko的(1986)定義可知，當兩位專家意見交集相同時，當專家的意見範圍越大者，所得的得子集合的程度就會越小，反之當專家的意見範圍越小者，所得的得子集合的程度就會越大，不同於Zwick et al(1987)提出相似程度量測，不論專家的意見範圍大或小，卻有相同的相似性估計值。且可利用模糊子集合衡量的觀念，將專家意見彙整結果與五等第(很不同意，不同意，尚可同意，同意，非常同意)做程度匹配(Match Degree)，可得到全體專家之語意判定。

本研究提出模糊子集合衡量彙整法(Subsethood Aggregation Method, SbAM)的操作程序下：

1. 計算任兩位專家間之模糊子集合相似程度 $S(\tilde{R}_i, \tilde{R}_j)$ ：

將專家個別之模糊評估值進行兩兩配對，且求出決策者彼此間之模糊子集合相似程度估計值。兩模糊評估值之模糊子集合相似程度計算公式如下：

$$S(A, B) = \text{degree}(A \subseteq B) = \frac{\sum_{x \in X} (\mu_A(x) \cap \mu_B(x))}{\sum_{x \in X} \mu_A(x)} = \frac{\mu(A \cap B)}{\mu(A)}$$

模糊子集合相似程度估計值的意義在於，當兩位專家之間的意見重疊部分相同時，專家意見範圍越大，則模糊子集合相似程度估計值越小，反之，專家意見範圍越小，則模糊子集合相似程度估計值越大。

2. 建構認同矩陣(Agreement Matrix, AM)：

假如專家人數為 n ，則可建立 $n \times n$ 之認同矩陣，矩陣內的值即代表任意二位專家相似程度。認同矩陣定義如下：

$$AM = [S_{ij}]_{n \times n}, i = 1, 2, \dots, n, j = 1, 2, \dots, n$$

$$AM = \begin{bmatrix} 1 & S_{12} & \cdots & S_{1j} & \cdots & S_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ S_{i1} & S_{i2} & \cdots & S_{ij} & \cdots & S_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ S_{n1} & S_{n2} & \cdots & S_{nj} & \cdots & 1 \end{bmatrix}$$

3. 計算每位專家的平均認同程度(AE_i)：

利用步驟二認同矩陣，計算每位專家E_i的，i=1,2,...,n，與其他專家的平均認同程度(Average Agreement of Expert)，計算公式如下：

$$A(E_i) = \frac{1}{n-1} \sum_{j=1}^n S_{ij} \quad j \neq i$$

4. 計算每位專家的相對認同程度(Relative Agreement Degree, RAD)：

進一步計算出專家E_i的相對認同程度，即稱為正規化，以RAD_i表示，計算公式如下：

$$RAD_i = \frac{A(E_i)}{\sum_{i=1}^n A(E_i)} \quad i = 1, 2, \dots, n$$

5. 計算每位專家的共識程度係數(Consensus Degree Coefficient, CDC)：

如不考慮專家各別的重要程度情況下，可設定β=0。所求得之相對認同程度(RAD)即等同於各專家之共識程度係數(CDC)。共識程度係數計算如下：

$$CDC_i = \beta \times w_i + (1 - \beta) \times RAD_i$$

其中 0 ≤ β ≤ 1

6. 彙整全體專家之模糊評估值：

利用每位專家所得之共識程度係數與該專家之模糊評估值予以彙整，其值為最後的彙整(Aggregation)的結果。公式如下：

$$\tilde{R} = \sum_{i=1}^n CDC_i \otimes \tilde{R}_i$$

\tilde{R} ：彙整後之模糊評估值

\tilde{R}_i ：第i位專家之模糊評估值，i=1,2,...,n

7. 計算專家意見彙整之模糊評估值與五等第(非常不同意、不同意、稍微不同意、同意、非常同意)各別進行兩兩配對之模糊子集合相似程度 $S(\tilde{R}_i, \tilde{R}_j)$ 。

第四章 案例應用

Linguistic variables can be used as the expression to match with natural language, term and its related modifier (hedge) formed completed terms for full descriptions on the real case scenario (Yen and Langari1999). Fuzzy numbers, especially Triangular Fuzzy Number (T.F.N.), are adopted for imprecise or vague judgments, and their arithmetic calculation were provided by Kaufmann & Gupta (1991) as well as by Zimmermann (1991). Likert's type questionnaire is the common tool used for collection of individual judgment, and the calculated arithmetic average is assumed as the summarized results of investigated group generally. We will demonstrate an empirical example in details as followings.

Table 4-1 shows the linguistic terms and their corresponding T.F.N. for five criteria rating, they are (1, 1, 2) for Very Low(VL), (1, 2, 3) for Low(L), (2, 3, 4) for Medium(M), (3, 4, 5) for High(H), and (4, 5, 5) for Very High(VH) in respective. Figure 4-1 lists T.F.N. membership diagrams for these five criteria rating.

Table 4-1: Linguistic terms and their T.F.N. for five criteria rating

Linguistic terms	Corresponding fuzzy number
Very Low (VL)	(1, 1, 2)
Low (L)	(1, 2, 3)
Medium (M)	(2, 3, 4)
High (H)	(3, 4, 5)
Very High (VH)	(4, 5, 5)

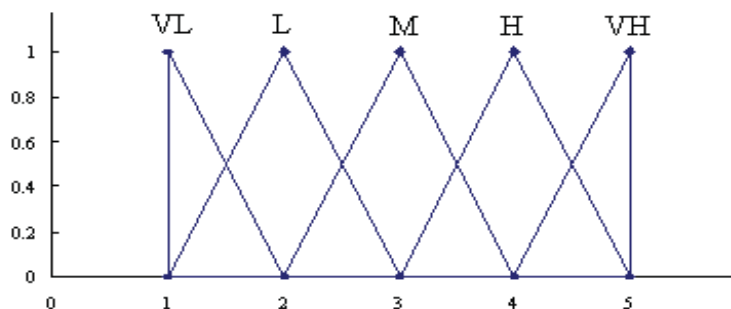


Figure 4-1: Membership diagram for five criteria rating

According to calculation procedures of SbAM discussed in previous section. Degree of 20

paired fuzzy subsethood measure are calculated, where $S_{ii} = 1$, $S_{12} = 0.5$, $S_{21} = 0.25$

$$(s(\tilde{R}_1, \tilde{R}_2) = \frac{|\tilde{R}_1 \cap \tilde{R}_2|}{|\tilde{R}_1|} = 0.50 \quad s(\tilde{R}_2, \tilde{R}_1) = \frac{|\tilde{R}_1 \cap \tilde{R}_2|}{|\tilde{R}_2|} = 0.25) \text{ and etc.. A } 5 \times 5 \text{ agreement matrix (AM) is}$$

constructed and listed in Table 4-2.

Table 4-2: Agreement matrix for five criteria rating

Criteria Rating	VL	L	M	H	VH
VL	1	.5	0	0	0
L	.25	1	.25	0	0
M	0	.25	1	.25	0
H	0	0	.25	1	.25
VH	0	0	0	.5	1

The general formula for each summarized degree of subsethood measure S_{ij} is listed in Table 4-3. For example, every questionnaire in group of Very Low(VL), would get $s_{11} = (n_1 - 1)$ and $s_{12} = (.5 * n_2)$, then the summarized degree of subsethood measure (multiply n_1) $S_{11} = n_1 * (n_1 - 1)$ and $S_{12} = n_1 * (.5 * n_2)$. If the number of questionnaire corrected is 50, and the number of questionnaire corresponding to each criteria rating are $n_1=2$ for Very Low(VL), $n_2=5$ for Low(L), $n_3=14$ for Medium(M), $n_4=25$ for High(H), and $n_5=4$ for Very High(VH) in respective. Table 4-4 list the calculation results while setting $n_1=2$, $n_2=5$, $n_3=14$, $n_4=25$ and $n_5=4$ to corresponding contents of table 4-3. Then sum of each row, listed in column “Sum”, can be got easily, and the criteria weighting derived from SbAM method, listed in column “SbAM Wt.”, can be calculated via. individual row-sum divided by total row-sum.

Table 4-3: Agreement matrix for five criteria rating with corresponding frequency

Term (Fre.)	VL (n_1)	L (n_2)	M (n_3)	H (n_4)	VH (n_5)
VL (n_1)	$n_1 * (n_1 - 1)$	$n_1 * (.5 * n_2)$	0	0	0
L (n_2)	$n_2 * (.25 * n_1)$	$n_2 * (n_2 - 1)$	$n_2 * (.25 * n_3)$	0	0
M (n_3)	0	$n_3 * (.25 * n_2)$	$n_3 * (n_3 - 1)$	$n_3 * (.25 * n_4)$	0
H (n_4)	0	0	$n_4 * (.25 * n_3)$	$n_4 * (n_4 - 1)$	$n_4 * (.25 * n_5)$
VH (n_5)	0	0	0	$n_5 * (.5 * n_4)$	$n_5 * (n_5 - 1)$

Table 4-4 : Example for criteria weighting calculation (SbAM vs. Equal Wt.)

Term (Fre.)	VL (2)	L (5)	M (14)	H (25)	VH (4)	Sum	SbAM Wt.	Equal Wt.
VL (2)	2	5	0	0	0	7	.006	.040
L (5)	2.5	20	17.5	0	0	40	.036	.100
M (14)	0	17.5	182	87.5	0	297	.266	.280
H (25)	0	0	87.5	600	25	712.5	.637	.500
VH (4)	0	0	0	50	12	62	.055	.080

Likert's questionnaire is the common tool used for collection of individual judgment, and the calculated arithmetic average is assumed as the summarized results of investigated group generally. As comparison the results of criteria weighting calculated through SbAM and arithmetic average. Based on analysis data listed in table 4-4, we can get the conclusion that arithmetic average method set every DM in equal weighting (0.02), and the total weighting for every criterion is then calculated as $0.02 * n_i$. While applying SbAM method to determine weighting of each DM, viewing on the results in table 4-4, criterion High (H) with 25 DMs get higher weighting from 0.5 to 0.637, but criterion Very Low (VL) with 2 DMs just get less weighting from 0.04 to 0.006. Then we got the conclusion that, the more the concentration for DMs will get the larger impacts, and vise-versa.

第五章 結論

Methodologies for weights setting of DM under group decision making have been discussed in previous sections. We provide three sets of test data, each one contains three measurements in forms of T.F.N. arranged by three experts in respective. Table 5-1 lists these test data, weights setting via various approaches, and aggregated fuzzy measurement. $\tilde{R}_1 \sim \tilde{R}_3$ represents individual fuzzy measurement, \tilde{S} represents aggregated fuzzy measurement, data in the same row corresponding to $\tilde{R}_1 \sim \tilde{R}_3$ represent Weights Setting via assigned approach. Each decision maker's weighting is set equally and all measurements are aggregated via arithmetic average in traditional performance evaluation model. Approach of arithmetic average neglects mutual interactions among DM, but the others include SAM, SbAM, LSDM, DLSM and OAM are all approaches of mutual interactions.

Both SAM and SbAM are based on degree of “superposition” for each paired measurements. If “span” of all measurements are equal, then same conclusions will be got for both SAM and SbAM. Test data set 1 listed in table 5-1 demonstrates the conclusion. According to test data set 2, if “span” of a certain measurements is varied, e.g. DM 2 revised its measurement from (2, 3, 4) to (1, 3, 5), and assumed the others were left unchanged. Based on test data set 2, we observed that results referred to SAM kept unchanged, while referred to SbAM, weighting of DM 2 has been reduced from 0.5 to 0.333 due to its measurement with wide “span” (less precise). It's just fit to the practical conclusion “the wider the decision range is, the smaller the decision effect should be”.

Table 5-1: Comparisons for SAM vs. SbAM Methods

Methods Test Data	Arithmetic Average	SAM	SbAM
$\tilde{R}_1(1,2,3)$	0.333	0.250	0.250
$\tilde{R}_2(2,3,4)$	0.333	0.500	0.500
$\tilde{R}_3(3,4,5)$	0.333	0.250	0.250
\tilde{S}	(2,3,4)	(2,3,4)	(2,3,4)
$\tilde{R}_1(1,2,3)$	0.333	0.250	0.333
$\tilde{R}_2(1,3,5)$	0.333	0.500	0.333
$\tilde{R}_3(3,4,5)$	0.333	0.250	0.333
\tilde{S}	(1.67,3,4.3)	(1.5,2.5,3.5)	(1.67,3,4.33)
$\tilde{R}_1(1,2,3)$	0.333	0.50	0.600
$\tilde{R}_2(2,3,5)$	0.333	0.50	0.400
$\tilde{R}_3(7,8,9)$	0.333	0	0
\tilde{S}	(3.3,4.3,5.7)	(1.5,2.5,4)	(1.4,2.4,3.8)

Continue to review on test data set 3, measurement of DM 3 has no interaction to both DM1 and DM2. Based on degree of “superposition”, there is no doubt about the fact that weighting of DM3 is set to zero. If weighting of a DM is set to zero, there is no effect on final conclusion, which can be looked as the filter to outliers. Both SAM and SbAM support functionality to sift normal data from outliers. Let’s review test data set 3 in deep, the measurement range of DM1 is less than that of DM2, in other words, the judgment from DM1 is more precise than that of DM2. It is obvious that weighting of DM1 would be larger than that of DM2. SbAM made it sense, but SAM seemed to have a bias. The bias existed in the instance that the one with more precise judgment has same effect to that with loose.

The decision weighting of each DM is set equally in traditional performance evaluation model, and arithmetic average is applied as the method for group data aggregation. Actually it is not fair and not reasonable, especially in case of outliers are existed. Once bias has been caused by outliers, it is far apart from the actual group consensus. Even then the calculation results will be varied from the various methods applied, but all methods discussed are all dependent on group interaction. One word to say, decision weighting of each DM is impacted by total decision group.

The solution procedures of SbAM are similar to that of SAM provided by Hsu and Chen (1996). Both of them are based on degree of “superposition”, the only difference existed is that SbAM is derived from fuzzy subsethood measure, but SAM is derived from fuzzy similarity measure. Through the discussion in previous section, it is obvious that SbAM is prior to SAM while applying in group data aggregation.

參考文獻

1. 王文俊，認識Fuzzy，全華科技圖書公司，2001。
2. 汪明生、汪仲祥，”模糊層級分析法應用於IC 產業政策選取之研究”，國立中山大學公共事務管理研究所碩士論文，2003。
3. 洪振創，王國明、顧志遠，”群體決策下模糊績效評估模式之建構與應用”，元智工學院工業工程研究所博士論文，1996。
4. 洪振創、蔡千姿，”具回饋功能之模糊績效評估資料收集系統—在網際網路上之設計與應用”，行政院國科會(NSC -87-2213-E-159-008)，1997。
5. 塗兆輝、洪振創，”群體決策績效衡量模式在物流中心作業之應用探討-明確值、區間值與模糊值之資料收集與分析”，明新科技大學工程管理研究所碩士論文，2006。
6. 馮國臣、任麗偉著，趙忠賢、張宏志、溫坤禮編修，模糊理論基礎與應用，新文京開發出版股份有限公司，2007。
7. 楊宗仁，模糊核保定期壽險上之探討，真理大學數理科學研究所碩士論文，2002。
8. 薛蕙俐、洪振創，應用專家群體決策方法建構物流中心倉儲理貨之績效評估架構，明新科技大學工程管理研究所碩士學位論文，2005。
9. 潘煜奇、洪振創，資料集成運算方法之探討及其在績效評估系統之應用—以教師自我評量為例，明新科技大學工程管理研究所碩士學位論文，2007。
10. 吳攻澧、洪振創，模糊子集合衡量在集體決策下準則權重設定與績效衡量之應用探討—以學生對老師之教學評量為例，明新科技大學工程管理研究所碩士學位論文，2009。
11. 盧盛忠、余凱成、徐旭、錢冰鴻著、蘇哲仁效訂，組織行為學—理論與實務，五南圖書出版公司，2000。
12. Bart Kosko著，林基興譯，模糊思考 (Fuzzy Thinking)，全華科技圖書公司，1994。
13. Chua-Chin Wang, Hon-Son Don, A modified measure for fuzzy subsethood, Information Sciences, Volume 79, Issues 3-4, July 1994, Pages 223-232
14. Dennis, A. R., Nunamaker, J.R. and Vogel, D.R., “A Comparison of Laboratory and Field Research in the Study of Electronic Meeting System.” Journal of Management Information System, Vol. 7, No.3, pp.107-135, 1991.
15. Delbecq, A. L., Van de Ven, A. H. and Gustafson, D. H. “Group Technique for Program Planning, a Guide to Norminal Group Technique and Delphi Process.” Scott Foreman, USA, 1975
16. Dong, W.M. and Wong, F.S., Park Fuzzy weighted average and implementation of the extension principle", fuzzy sets and systems 21 (1987) 183-199.

17. Guu, S.M., " Fuzzy weighted averages revisited", Fuzzy Sets and Systems 126 (2002) 411-414
18. Guh, Y.Y. Hon, C.C. Wang, K.M. and Lee,E.S. "Fuzzy weighted average: a max-min paired elimination method",Comput. Math. Appl.32 (1996) 115 -123.
19. Guh, Y.Y. Hon,C.C. and Lee,E.S., " Fuzzy weighted average:the linear programming approach via charnes and cooper's rule", Fuzzy Sets and Systems 117 (2001) 157-160.
20. Herrear, F., Herrera-Viedma, E., Verdegay, J.L., Diret approach in group decision making using linguistic OWA operators, Fuzzy Set and Systems, 79,175-190,1996
21. Jiulun Fan, Weixin Xie and Jihong Pei, Subsethood measure: new definitions, Fuzzy Sets and Systems, Volume 106, Issue 2, 1 September 1999, Pages 201-209
22. Kao,C. and Liu, S.-T. "Fractional programming approach to fuzzy weighted average", Fuzzy Sets and Systems 120 (2001) 435– 444.
23. Kaufmann, Arnold and Gupta, Madan M., Fuzzy mathematical models in engineering and management science, Elsevier Science Publishers Inc., New York, 1988
24. Kaufmann, Arnold and Gupta, Madan M., "Introduction to Fuzzy Arithmetic: Theory and Applications", Van Nostrand Reinhold, N.Y. (1991).
25. Lee,D.H.and Park, D., "An efficient algorithm for fuzzy weighted average", Fuzzy Sets and Systems 87 (1997) 39-45.
26. Liou, Tian-Shy and Wang, Mao-Jiun, "Fuzzy weighted average : An improved algorithm", fuzzy sets and systems, 49 (1992) 307-315.
27. P. Sevastjanov and P. Figat, Aggregation of aggregating modes in MCDM: Synthesis of Type 2 and Level 2 fuzzy sets, Omega, Volume 35, Issue 5, October 2007, Pages 505-523
28. Robbins, Stephen P., Management, Prentice-Hall, pp. 168-173, 1994.
29. Virginia R. Young, Fuzzy subsethood, Fuzzy Sets and Systems, Volume 77, Issue 3, 12 February 1996, Pages 371-384
30. Zimmermann, H.J., Fuzzy set theory and its application, 2nd Ed., Kluwer Academic Publishers Boston, (1991).

明新科技大學 103 年度校內專題研究計畫 運用於教學成果記錄表

計畫類型	◎個人型□整合型□任務導向型		計畫編號	MUST-103-工管-7	
計畫名稱	模糊子集合衡量在祛除資料異常值影響之應用探討				
計畫主持人資料	姓名	洪振創		職稱	副教授
	學院	管理學院		系所	工業工程與管理系
聘用助理	系科班級	學號	姓名	聘僱起訖時間	工作內容
	無				
融入課程	開課班級	課程名稱		修課人數	課程內容概述
	碩士班	績效管理		6	問卷調查資料分析方法
指導專題或碩士論文	指導班級	專題(論文)名稱		分組人數	專題(論文)內容概述
	碩二	模糊子集合衡量與統計分析方法在祛除資料異常值之比較(預計)		1	探討模糊子集合衡量與統計分析方法之特性與演算法則，針對特定領域相關應用資料之比較分析，驗證模糊子集合衡量應用的可行性與效益。
指導學生參與活動或競賽	活動或競賽名稱			參與人數	活動或競賽成果概述
	無				
製作教材與教具	教材與教具名稱			教材與教具概述	
	應用 EXCEL 進行模糊子集合衡量之應用分析			只需輸入原始數據，藉由設計 EXCEL 分析工具，得到即時計算的效果	
其他促進教學之成果說明	無				