國 立 政 治 大 學 應 用 數 學 系

數 學 教 學 碩 士 在 職 專 班

碩士學位論文

模糊資料相關係數及在數學教育之應用

Correlation of Fuzzy Data and its Applications in Mathematical Education

碩專班學生：林立夫 撰

指導教授：吳柏林 博士

中華民國　一〇〇 年　六 月　十三 日

摘要

兩變數之間是否相關，以及相關的程度與方向是統計研究學者所關注的一項課題。傳統上使用皮爾森相關係數(Pearson's Correlation Coefficient)來表達兩實數變數間線性關係的強度與方向。然而，對於反映人類思維不確定性的模糊資料而言，傳統的相關分析方法卻有不足與不適用之缺失。

本論文的主要目的在於尋求一個合理、適用的區間模糊資料相關係數，提供研究者簡單且容易計算的模糊相關係數求法，用以了解區間模糊資料間的相關程度。接著利用轉換離散型模糊數成爲區間模糊數的方式，處理離散型模糊資料間的相關係數。最後，以國中數學教學現場所調查的資料做實例應用。

關鍵詞：模糊統計、區間模糊數、模糊相關係數。

Abstract

In statistical studies, the correlation between two variables and its strength and direction are always concerned. Traditionally, the Pearson's Correlation Coefficient is used to convey the linear relationship between two variables. However, the traditional correlation analysis is not applicable to the fuzzy data which are able to reflect more appropriately the uncertainty of human thinking.

The main purpose of the study is to find a reasonable and usable correlation coefficient of interval fuzzy data which provides researchers a simple and easy way to calculate and find the fuzzy correlation coefficient. Meanwhile, it can help us understand the correlation of interval fuzzy data. Moreover, we use the process of transforming discrete fuzzy number into the interval fuzzy number to deal with the correlation coefficient of discrete fuzzy data. Finally, we utilize the data from mathematics teaching in junior high school for application.

Keywords: Fuzzy statistics; Interval fuzzy number; Fuzzy correlation coefficient.

目次

圖目次

表目次

1. 前言

自 Zadeh (1965)提出模糊理論以來，解決了許多人類思維方式對環境和研判事情角度的多元複雜與曖昧、不確定現象。統計學中，相關係數用來表示兩個隨機變數之間線性關係的強度與方向，然而卻少有研究在討論當樣本爲模糊樣本時，樣本之間的關係。我們希望能夠了解當樣本爲模糊樣本時，樣本之間的關係如何表示。例如：當我們詢問某人的理想薪資與希望工作時數時，少有人會回答一個切確的數字，因爲人類思維的複雜與不確定性，大多數我們獲得的答案是一個範圍。而我們就是希望能夠知道這樣兩組資料之間的關係。

近來對於模糊理論應用的相關研究也愈來愈多。Dubois 與 Prade (1987)、Heilpern (1992)、Carlsson 與 Fuller (2001)都曾針對模糊數的平均數、變異數進行相關研究。Korner (1997) 提出模糊變異數的概念。Gebhardt et al. (1998) 將相關的研究作了有條理的整理。Wu (1999)則針對模糊隨機變數討論其可能的機率密度函數。Galvo 與 Mesiar (2001)將模糊理論用於中位數之推廣。Liu 與 Kao (2002)定義了模糊數的相關係數。Hung 與 Wu (2001,2002)應用期望區間及α-截集(α-cuts)的方法來衡量模糊數之間的相關性。Buckley (2003, 2004)則將模糊理論與傳統的機率與統計推論相結合，並作出有系統的歸納整理。

區間相關是目前研究相當熱門的一個課題。區間資料在統計分析上也越來越受到重視，像每日氣溫、薪資收益、股市每日高點與低點、理想的子女數、結婚年齡等。但是對於區間的統計量目前還沒有一致的完善定義。例如，對於區間資料的眾數、中位數、平均數文獻也少見；見 Hung and Wu (2006)。江彥聖 (2008)提出模糊相關係數定義，但運算十分繁複。本研究針對模糊區間樣本，定義更爲便捷且符合實際的方式做深入的探討。

本研究從教育統計的角度來看，運用有關教育的區間資料來進行統計分析，希望能夠得到一些有意義且異於傳統的結果，並藉著區間資料的相關分析探討人類思維的模糊性與有助於教育心理研究與度量的進一步發展。

2. 研究方法

2.1 傳統實數樣本之線性相關係數

如果我們想要了解 X 與 Y 兩個現象之間的相關程度，最直接的方法就是將 (X,Y) 的資料散佈圖畫出來。由散佈圖的呈現，我們便可以約略看出兩組資料之間的相關性(吳柏林, 2005)。在傳統的實數樣本相關分析中，我們使用皮爾森相關係數(Pearson's Correlation Coefficient)來表達兩變數間線性關係的強度。它的定義為：

$$\text{相關係數}\quad \rho = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

當 $\rho > 0$ 時，我們稱 X 與 Y 之間為正相關；當 $\rho < 0$ 時，稱 X 與 Y 之間為負相關；當 $\rho = 0$ 時，則稱 X 與 Y 之間沒有關係存在，或說統計無關。然而在實務應用上，母體的變異數 σ_X^2、σ_Y^2 以及它們之間的共變異數 $Cov(X,Y)$ 並不容易得到，所以，我們用樣本相關係數 γ 來估計 ρ，即

$$\gamma = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \qquad (2.1)$$

其中 (x_i, y_i) 為第 i 對樣本值，$i = 1,2,...,n$；$\overline{x}$ 與 $\overline{y}$ 分別為其樣本平均數。

以下為傳統實數樣本相關係數的一些特性：

假設有 n 組資料 (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n)，而 (x, y) 之相關係數為 γ，則：

1. γ 的值必介於 -1 與 $+1$ 之間。
2. γ 的大小表示線性關係的強度，正、負號表示相關的方向。詳言之：

$\gamma > 0$，若(x, y)值的圖形為一帶狀且從左下方至右上方。

$\gamma < 0$，若(x, y)值的圖形為一帶狀且從左上方至右下方。

$\gamma = +1$，若所有(x, y)值均剛好落於一直線上，且具正斜率（完全線性正相關）。

$\gamma = -1$，若所有(x, y)值均剛好落於一直線上，且具負斜率（完全線性負相關）。

$|\gamma|$值越大（亦即γ愈接近$+1$或-1），則表示線性關係之強度愈大。

3. γ的值愈接近0，則意味著線性相關愈弱。

2.2 模糊統計量

模糊理論中，隸屬度函數是最基本的概念，它是從傳統集合中的特徵函數(characteristic function)所衍生出來的。隸屬度函數區分爲離散型(discrete type)與連續型(continuous type)兩種，用來表示元素對模糊集合的隸屬度(membership grade)，其範圍介於 0 到 1 之間。本小節簡單介紹模糊數，以及離散型模糊數的樣本平均數與變異數，並提出例子供參考。同時爲了計算離散型模糊數的相關係數，本研究必須先將離散型模糊數反模糊化，再利用離散型模糊數的標準差將離散型模糊數轉換成模糊區間以求得其相關係數。

定義 2.1： 模糊數（*Fuzzy Number*） (吳柏林，2005)

設U為一論域，令$\{L_1,\ L_2,\ \ldots,\ L_k\}$為論域$U$的因子集。$\mu$為一對應到[0, 1]間的實數函數，即$\mu: U \to [0,1]$。假若佈於論域$U$之一述句$X$其相對於因子集的隸屬度函數以$\{\mu_{L_1}(X), \mu_{L_2}(X), \cdots, \mu_{L_k}(X)\}$表示，則在離散的情形下，述句$X$的模糊數可表示成：

$$\mu_U(X) = \frac{\mu_1(X)}{L_1} + \frac{\mu_2(X)}{L_2} + \cdots + \frac{\mu_k(X)}{L_k}$$

其中「+」是或的意思，$\dfrac{\mu_i(X)}{L_i}$表示述句 X 隸屬於因子集L_i的程度，而且$\sum_{i=1}^{k} \mu_{L_i}(X) = 1$。當$U$為連續時，述句$X$的模糊數可表示成：$\mu_U(X) = \int_{x \in X} \dfrac{\mu_i(X)}{L_i}$。

例 2.1：教師對學生課堂表現評比的模糊數表示

某班級五位任課教師對班上某位學生課堂表現進行評比如表 2.1：

表 2.1 五位教師對某位學生課堂表現的隸屬度

教師	L_1=非常不專心	L_2=不專心	L_3=普通	L_4=專心	L_5=非常專心
A	0.7	0.2	0.1		
B			0.2	0.5	0.3
C				0.3	0.7
D	0.1		0.1	0.8	
E		0.4	0.4	0.2	

假設 X = 學生課堂表現，以模糊數表示為 $\mu_U(X)$。

假設論域 $U = \{ L_1 =$ 非常不專心, $L_2 =$ 不專心, $L_3 =$ 普通, $L_4 =$ 專心, $L_5 =$ 非常專心$\}$。其中教師 B 對於該生課堂中的表現隸屬度函數為

$$\{ \mu_1(X) = 0,\ \mu_2(X) = 0,\ \mu_3(X) = 0.2,\ \mu_4(X) = 0.5,\ \mu_5(X) = 0.3 \}$$

亦可以模糊數表示為

$$\mu_U(X) = \frac{0}{L_1} + \frac{0}{L_2} + \frac{0.2}{L_3} + \frac{0.5}{L_4} + \frac{0.3}{L_5}$$

定義 2.2: 離散型模糊樣本平均數 *(Fuzzy mean for discrete type)* (吳柏林，2005)

設 U 為一論域，令 $L = \{L_1, L_2, \cdots, L_k\}$ 為佈於論域 U 上的 k 個語言變數，$\{x_i = \frac{m_{i1}}{L_1} + \frac{m_{i2}}{L_2} + \ldots + \frac{m_{ik}}{L_k}, i = 1,2,\ldots,n\}$ 為一組模糊樣本($\sum_{j=1}^{k} m_{ij} = 1$)。則模糊樣本平均數為

$$F_{\bar{x}} = \frac{\frac{1}{n}\sum_{i=1}^{n} m_{i1}}{L_1} + \frac{\frac{1}{n}\sum_{i=1}^{n} m_{i2}}{L_2} + \cdots + \frac{\frac{1}{n}\sum_{i=1}^{n} m_{ik}}{L_k}$$

其中 m_{ij} 為第 i 個樣本相對於語言變數 L_j 之隸屬度。

例 2.2: 教師對學生課堂表現評比的模糊平均數

承例 2.2，某班級五位任課教師對班上某學生課堂表現進行評比如表 2.1。則此五位教師對於學生課堂表現之模糊平均數為

$$F_{\bar{x}} = \frac{\frac{1}{5}(0.7+0+0+0.1+0)}{L_1} + \frac{\frac{1}{5}(0.2+0+0+0+0.4)}{L_2} + \frac{\frac{1}{5}(0.1+0.2+0+0.1+0.4)}{L_3}$$

$$+\frac{\frac{1}{5}(0+0.5+0.3+0.8+0.2)}{L_4} + \frac{\frac{1}{5}(0+0.3+0.7+0+0)}{L_5}$$

$$= \frac{0.16}{L_1} + \frac{0.12}{L_2} + \frac{0.16}{L_3} + \frac{0.36}{L_4} + \frac{0.2}{L_5}$$

顯示在於 L_4=專心的隸屬度是較大的。意謂著該生課堂表現呈現專心狀態是多於其它語言變數的。

定義 2.3: 離散型模糊樣本變異數 *(Fuzzy variance for discrete type)* (吳柏林，2005)

設 U 為一論域，令 $L = \{L_1, L_2, \cdots, L_k\}$ 為佈於論域 U 上的 k 個有序變數，而 $\{x_i = \frac{m_{i1}}{L_1} + \frac{m_{i2}}{L_2} + \ldots + \frac{m_{ik}}{L_k}, i = 1,2,\ldots,n\}$，$\sum_{j=1}^{k} m_{ij} = 1$ 為論域中抽出的一組樣本觀測值，若其樣本平均數為

$$F_{\overline{x}} = \frac{\widehat{\mu}_1}{L_1} + \frac{\widehat{\mu}_2}{L_2} + \cdots + \frac{\widehat{\mu}_k}{L_k}$$

則其模糊樣本變異數 *(Fuzzy variance for discrete type)*為

$$FS^2 = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(m_{i1}-\widehat{\mu}_1)^2}{L_1} + \frac{\frac{1}{n-1}\sum_{i=1}^{n}(m_{i2}-\widehat{\mu}_2)^2}{L_2} + \ldots + \frac{\frac{1}{n-1}\sum_{i=1}^{n}(m_{ik}-\widehat{\mu}_k)^2}{L_k}$$

例 2.3: 教師對學生課堂表現評比的模糊變異數

五位教師對於學生課堂表現的模糊數型式如下：

$$X_{\mathrm{A}} = \frac{0.7}{L_1} + \frac{0.2}{L_2} + \frac{0.1}{L_3} + \frac{0}{L_4} + \frac{0}{L_5}$$

$$X_{\mathrm{B}} = \frac{0}{L_1} + \frac{0}{L_2} + \frac{0.2}{L_3} + \frac{0.5}{L_4} + \frac{0.3}{L_5}$$

$$X_{\mathrm{C}} = \frac{0}{L_1} + \frac{0}{L_2} + \frac{0}{L_3} + \frac{0.3}{L_4} + \frac{0.7}{L_5}$$

$$X_{\mathrm{D}} = \frac{0.1}{L_1} + \frac{0}{L_2} + \frac{0.1}{L_3} + \frac{0.8}{L_4} + \frac{0}{L_5}$$

$$X_{\mathrm{E}} = \frac{0}{L_1} + \frac{0.4}{L_2} + \frac{0.4}{L_3} + \frac{0.2}{L_4} + \frac{0}{L_5}$$

則其模糊樣本變異數為

$$FS^2 = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(m_{i1} - \widehat{\mu}_1)^2}{L_1} + \frac{\frac{1}{n-1}\sum_{i=1}^{n}(m_{i2} - \widehat{\mu}_2)^2}{L_2} + \ldots + \frac{\frac{1}{n-1}\sum_{i=1}^{n}(m_{ik} - \widehat{\mu}_k)^2}{L_k}$$

$$= \frac{0.093}{L_1} + \frac{0.032}{L_2} + \frac{0.023}{L_3} + \frac{0.093}{L_4} + \frac{0.095}{L_5}$$

在例題 2.3 中，教師對學生課堂表現評比的模糊變異數所代表的意義爲：「非常不專心」的隸屬度爲 0.093，「不專心」的隸屬度爲 0.032，「普通」的隸屬度爲 0.023，「專心」的隸屬度爲 0.093，「非常專心」的隸屬度爲 0.095。表示五位教師對於該學生課堂表現中「非常專心」的歧異性最大，但是「非常不專心」與「專心」的隸屬度亦達 0.093，代表著不同教師對於該生的課堂表現，在「非常不專心」、「專心」與「非常專心」這三個語言變數有著較高的變異性。可以推測該生對於不同教師所呈現的課堂表現有所差異。

根據定義 2.3 離散型模糊樣本變異數，我們可以了解到整組離散型模糊資料對於不同語言變數的變異程度。然而，對於單一離散型模糊數，其集中與離散程度如何也是我們所想了解的。爲了了解單一離散型模糊數的集中與離散程度，先定義離散型模糊數的反模糊化值如下，以了解單一離散型模糊數的集中程度。

定義 2.4: 離散型模糊數的反模糊化值 (吳柏林，2005)

設 x 為一模糊數，語言變數 $L=\{L_1,L_2,\cdots,L_k\}$ 為論域 U 中有序的數列，$\mu_{L_i}(X)=m_i$ 為模糊樣本 X 相對於 L_i 的隸屬度函數，其中 $\sum_{i=1}^{n}\mu_{L_i}(X)=1$，則模糊數 x 的反模糊化值為：

$$x_f=\sum_{i=1}^{k}m_iL_i$$

例 2.4: 教師對學生課堂表現的反模糊化值

根據表 2.1，某班級五位任課教師對班上某學生課堂表現所進行的評比。若分別賦予學生課堂表現語言變數的權重為：非常不專心＝0，不專心＝0.25，普通＝0.5，專心＝0.75，非常專心＝1，則五位教師對學生課堂表現評比的反模糊化值為：

教師 A：$0.7\times0+0.2\times0.25+0.1\times0.5=0.1$

教師 B：$0.2\times0.5+0.5\times0.75+0.3\times1=0.775$

教師 C：$0.3\times0.75+0.7\times1=0.925$

教師 D：$0.1\times0+0.1\times0.5+0.8\times0.75=0.65$

教師 E：$0.4\times0.25+0.4\times0.5+0.2\times0.75=0.45$

接著定義離散型模糊數之標準差如下，以了解單一離散型模糊數的離散程度。

定義 2.5: 離散型模糊數之標準差 (江彥聖，2008)

設 x 為一模糊數，語言變數 $L=\{L_1,L_2,\cdots,L_k\}$ 為論域 U 中有序的數列，$\mu_{L_i}(X)=m_i$ 為模糊樣本 X 相對於 L_i 的隸屬度函數，其中 $\sum_{i=1}^{n}\mu_{L_i}(X)=1$。若模糊數 x 的反模糊化值為 $x_f=\sum_{i=1}^{k}m_iL_i$，則此模糊數 x 的標準差為：

$$F_x\sigma=\sqrt{\sum_{i=1}^{k}m_i(L_i-x_f)^2}$$

例 2.5: 考生對 99 年基本學力測驗數學科難易看法的標準差

訪問五位考生對於 99 年基本學力測驗數學科難易度看法整理如表 2.2：

表 2.2 五位考生對 99 年基本學力測驗數學科難易度看法之隸屬度

考生	L_1=非常容易	L_2=容易	L_3=普通	L_4=難	L_5=非常難
A	0.1	0.3	0.5	0.1	
B		0.2	0.4	0.3	0.1
C	0.2		0.5	0.3	
D	0.3	0.4	0.2		0.1
E		0.5	0.4	0.1	

若分別賦予五個語言變數權重為：

非常容易$=0$，容易$=0.25$，普通$=0.5$，難$=0.75$，非常難$=1$。

則考生 A 的難易度看法反模糊化值

$$x_{fA}=0.1\times 0+0.3\times 0.25+0.5\times 0.5+0.1\times 0.75=0.4$$

考生 A 的難易度看法模糊標準差

$$F\sigma_A=\sqrt{0.1\times(0-0.4)^2+0.3\times(0.25-0.4)^2+0.5\times(0.5-0.4)^2+0.1\times(0.75-0.4)^2}=0.2$$

考生 B 的難易度看法反模糊化值

$$x_{fB}=0.2\times 0.25+0.4\times 0.5+0.3\times 0.75+0.1\times 1=0.575$$

考生 B 的難易度看法模糊標準差

$$F\sigma_B=\sqrt{0.2\times(0.25-0.575)^2+0.4\times(0.5-0.575)^2+0.3\times(0.75-0.575)^2+0.1\times(1-0.575)^2}$$
$$=0.225$$

考生 C 的難易度看法反模糊化值

$x_{fC} = 0.2\times 0 + 0.5\times 0.5 + 0.3\times 0.75 = 0.475$

考生 C 的難易度看法模糊標準差

$$F\sigma_C = \sqrt{0.2\times(0-0.475)^2 + 0.5\times(0.5-0.475)^2 + 0.3\times(0.75-0.475)^2} \approx 0.261$$

考生 D 的難易度看法反模糊化值

$x_{fD} = 0.3\times 0 + 0.4\times 0.25 + 0.2\times 0.5 + 0.1\times 1 = 0.3$

考生 D 的難易度看法模糊標準差

$$F\sigma_D = \sqrt{0.3\times(0-0.3)^2 + 0.4\times(0.25-0.3)^2 + 0.2\times(0.5-0.3)^2 + 0.1\times(0.1-0.3)^2}$$
$$\approx 0.292$$

考生 E 的難易度看法反模糊化值

$x_{fE} = 0.5\times 0.25 + 0.4\times 0.5 + 0.1\times 0.75 = 0.4$

考生 E 的難易度看法模糊標準差

$$F\sigma_E = \sqrt{0.5\times(0.25-0.4)^2 + 0.4\times(0.5-0.4)^2 + 0.1\times(0.75-0.4)^2} \approx 0.166$$

在例題 2.5 中，考生 A 及考生 E 對於數學試題難易度看法的反模糊化值均為 0.4，而模糊標準差分別為 0.2 與 0.166，即考生 A 的模糊標準差大於考生 E 的模糊標準差，可以得知考生 A 對於式題難易度看法的分散程度較大。同時由表 2.2 也可以發現，考生 A 的隸屬度函數較為分散。

接著我們將利用定義 2.4 的反模糊化值與定義 2.5 的模糊標準差，將有序性的語言變數加上權重找出離散型模糊數的重心與標準差，再將離散型模糊數轉換成模糊區間。定義如下：

定義 2.6: 離散型模糊樣本之模糊區間 (江彥聖，2008)

設U為一論域，語言變數$L=\{L_1, L_2, \cdots, L_k\}$為論域$U$中有序的數列，其中$\mu_{L_i}(X)=m_i$為模糊樣本$X$相對於$L_i$的隸屬度函數，$\sum_{i=1}^{n}\mu_{L_i}=1$，述句$X$的模糊數為$\mu_U(X)=\frac{\mu_1(X)}{L_1}+\frac{\mu_2(X)}{L_2}+\cdots+\frac{\mu_k(X)}{L_k}$；如果模糊數$x$的反模糊化值為

$x_f=\sum_{i=1}^{k}m_iL_i$；標準差為：$F_x\sigma=\sqrt{\sum_{i=1}^{k}m_i(L_i-x_f)^2}$

則稱此模糊數x的模糊區間為

$$Ix_\alpha=[\max\{0, x_f-\alpha\cdot F_x\sigma\}, \min\{x_f+\alpha\cdot F_x\sigma, 1\}]$$

其中$\alpha\ge 0$為模糊區間半徑的控制係數。

例 2.6: 導師對學校行政協助滿意度之模糊區間

調查五位班級導師對於學校行政協助滿意度，如表 2.3：

表 2.3 五位導師對於學校行政協助滿意度隸屬度

導師	L_1=非常不滿意	L_2=不滿意	L_3=普通	L_4=滿意	L_5=非常滿意
A			0.2	0.8	
B		0.1	0.3	0.5	0.1
C	0.1	0.2	0.3	0.4	
D		0.4	0.3	0.2	0.1
E		0.1	0.4	0.5	

若分別賦予五個語言變數權重為：

非常不滿意$=0$，不滿意$=0.25$，普通$=0.5$，滿意$=0.75$，非常滿意$=1$。

則根據定義 2.4 求得反模糊化值如下

導師 A 對於行政協助滿意度的反模糊化值為:

$$x_{fA}=0.2\times 0.5+0.8\times 0.75=0.7$$

導師 B 對於行政協助滿意度的反模糊化值為:

$$x_{fB} = 0.1 \times 0.25 + 0.3 \times 0.5 + 0.5 \times 0.75 + 0.1 \times 1 = 0.65$$

導師 C 對於行政協助滿意度的反模糊化值為:

$$x_{fC} = 0.1 \times 0 + 0.2 \times 0.25 + 0.3 \times 0.5 + 0.4 \times 0.75 = 0.5$$

導師 D 對於行政協助滿意度的反模糊化值為:

$$x_{fD} = 0.4 \times 0.25 + 0.3 \times 0.5 + 0.2 \times 0.75 + 0.1 \times 1 = 0.5$$

導師 E 對於行政協助滿意度的反模糊化值為:

$$x_{fE} = 0.1 \times 0.25 + 0.4 \times 0.5 + 0.5 \times 0.75 = 0.6$$

根據定義 2.5 求得模糊標準差如下

導師 A 對於行政協助滿意度的模糊標準差為:

$$F\sigma_A = \sqrt{0.2 \times (0.5 - 0.7)^2 + 0.8 \times (0.75 - 0.7)^2} = 0.1$$

導師 B 對於行政協助滿意度的模糊標準差為:

$$F\sigma_B = \sqrt{0.1 \times (0.25 - 0.65)^2 + 0.3 \times (0.5 - 0.65)^2 + 0.5 \times (0.75 - 0.65)^2 + 0.1 \times (1 - 0.65)^2} = 0.2$$

導師 C 對於行政協助滿意度的模糊標準差為:

$$F\sigma_C = \sqrt{0.1 \times (0 - 0.5)^2 + 0.2 \times (0.25 - 0.5)^2 + 0.3 \times (0.5 - 0.5)^2 + 0.4 \times (0.75 - 0.5)^2} = 0.25$$

導師 D 對於行政協助滿意度的模糊標準差為:

$$F\sigma_D = \sqrt{0.4 \times (0.25 - 0.5)^2 + 0.3 \times (0.5 - 0.5)^2 + 0.2 \times (0.75 - 0.5)^2 + 0.1 \times (1 - 0.5)^2} = 0.25$$

導師 E 對於行政協助滿意度的模糊標準差為:

$$F\sigma_E = \sqrt{0.1 \times (0.25 - 0.6)^2 + 0.4 \times (0.5 - 0.6)^2 + 0.5 \times (0.75 - 0.6)^2} \approx 0.166$$

根據定義 2.6，取 $\alpha = 0.5$ 求得模糊區間如下

導師 A 對於行政協助滿意度的模糊區間為:

$$Ix_{A,0.5} = [\max\{0, 0.7 - 0.5 \times 0.1\}, \min\{0.7 + 0.5 \times 0.1, 1\}] = [0.65, 0.75]$$

導師 B 對於行政協助滿意度的模糊區間為:

$$Ix_{B,0.5}=[\max\{0,0.65-0.5\times 0.2\},\min\{0.65+0.5\times 0.2,1\}]\text{=}[0.55,0.75]$$

導師 C 對於行政協助滿意度的模糊區間為:

$$Ix_{C,0.5}=[\max\{0,0.5-0.5\times 0.25\},\min\{0.5+0.5\times 0.25,1\}]=[0.375,0.625]$$

導師 D 對於行政協助滿意度的模糊區間為:

$$Ix_{D,0.5}=[\max\{0,0.5-0.5\times 0.25\},\min\{0.5+0.5\times 0.25,1\}]=[0.375,0.625]$$

導師 E 對於行政協助滿意度的模糊區間為:

$$Ix_{E,0.5}=[\max\{0,0.6-0.5\times 0.166\},\min\{0.6+0.5\times 0.166,1\}]=[0.517,0.683]$$

根據例 2.6，我們將離散型模糊數轉換成區間模糊數的過程整理如圖 2.1

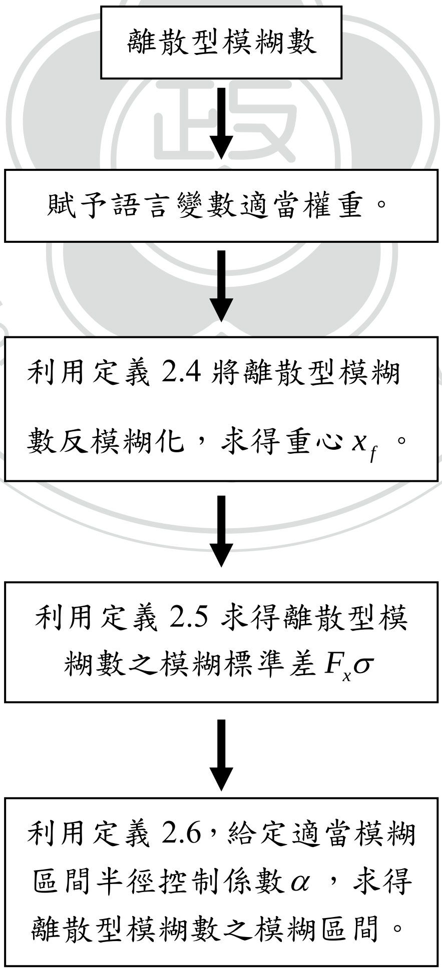

圖 2.1 離散型模糊數轉換成區間模糊數的流程圖

2.3 區間模糊數與軟計算

自 Zadeh (1965)提出模糊理論以來，此思維可以解釋許多符合人類不明確性的實務現象。模糊理論的概念，主要是在於強調個人喜好程度不需非常清晰或數值精確，而區間資料正是反應人類模糊思維的一種形式。Zedah (1999)即建議引用感覺測度(perception measure)和軟計算(soft computing system)做爲模糊函數估計量。

區間資料在統計分析上越來越受到重視，像每日氣溫、薪資收益、股市每日高點與低點、理想的子女數、結婚年齡等。當雇主詢問職員理想的薪資時，人們往往會回答 2~3 萬、3 萬 5 千~4 萬而不是很肯定的回答某個數目。相同的，當別人詢問你每個禮拜花多少時間在運動時，你也不會回答一個切確的時間數，因爲人類的思維有大部分是不確定的、模糊的。本節將以模糊理論的概念定義區間模糊數，並定義區間模糊數的均數與變異數，同時舉例說明

定義 2.7: 區間模糊數（*Interval Fuzzy Number*） (吳柏林，2005)

若 a ,b ,c 和 r 皆為實數，則稱 $X = [x_L \ ,x_U] = [a \ ,b] = (c \ ;r)$ 為區間模糊數。其中 $c = (a+b)/2$ 為區間 $[a \ ,b]$ 之中心，$r = (b-a)/2$ 為區間長度的半徑。而 $l = b-a$ ，代表該區間的長度。

例 2.7: $X_1 = [1 \ ,7] = (4 \ ;3)$ 。

$X_2 = [4 \ ,8] = (6 \ ;2)$ 。

定義 2.8: 連續型模糊樣本均數 *(Fuzzy sample mean for continuous type)*

設 U 為一論域，令 $L = \{L_1, L_2, \cdots, L_k\}$ 為佈於論域 U 上的 k 個語言變數，$\{x_i = [a_i, b_i] = (c_i; r_i) \ , i = 1,2,...,n\}$ 為論域 U 裡的一組模糊樣本。則模糊樣本均數為

$$F\overline{x} = [\frac{1}{n}\sum_{i=1}^{n} a_i \ , \frac{1}{n}\sum_{i=1}^{n} b_i] = (\frac{1}{n}\sum_{i=1}^{n} c_i \ ; \frac{1}{n}\sum_{i=1}^{n} r_i)$$

例 2.8: 五位大學新鮮人對於第一份工作起薪的要求

調查五位大學新鮮人對於第一份工作起薪的要求如表 2.4:

表 2.4 五位大學新鮮人起薪要求(單位:萬元)

大學新鮮人	可接受最低薪資	期望薪資	薪資要求範圍
A	2.2	3.0	2.2 – 3.0
B	2.5	4.5	2.5 – 4.5
C	2.7	3.5	2.7 – 3.5
D	2.8	3.6	2.8 – 3.6
E	1.9	2.8	1.9 – 2.8

五位大學新鮮人起薪要求模糊樣本如下：

$$X_A = [2.2, 3.0] = (2.6; 0.4)$$

$$X_B = [2.5, 4.5] = (3.5; 1.0)$$

$$X_C = [2.7, 3.5] = (3.1; 0.4)$$

$$X_D = [2.8, 3.6] = (3.2; 0.4)$$

$$X_E = [1.9, 2.8] = (2.35; 0.45)$$

求得五位大學新鮮人起薪要求模糊樣本平均數為：

$$F\overline{x} = [\frac{2.2+2.5+2.7+2.8+1.9}{5}, \frac{3.0+4.5+3.5+3.6+2.8}{5}]$$

$$= [2.42, 3.48]$$

$$= (\frac{2.6+3.5+3.1+3.2+2.35}{5}; \frac{0.4+1.0+0.4+0.4+0.45}{5})$$

$$= (2.95; 0.53)$$

定義 2.9: 連續型模糊樣本變異數 (吳柏林，2005)

設U為一論域，令$L=\{L_1,L_2,\cdots,L_k\}$為佈於論域U上的k個語言變數，$\{x_i=[a_i,b_i]=(c_i;r_i), i=1,2,...,n\}$為論域$U$裡的一組模糊樣本，令$c_i=\frac{(a_i+b_i)}{2}$，$r_i=\frac{(b_i-a_i)}{2}$，$\hat{c}=\frac{1}{n}\sum_{i=1}^{n}c_i$，$\hat{r}=\frac{1}{n}\sum_{i=1}^{n}r_i$。則模糊樣本變異數為

$$FS^2=\left(\frac{\sum_{i=1}^{n}(c_i-\hat{c})^2}{n-1};\frac{\sum_{i=1}^{n}(r_i-\hat{r})^2}{n-1}\right)$$

例 2.9: 五位大學新鮮人起薪要求模糊樣本變異數

根據表 2.4，五位大學新鮮人起薪要求模糊樣本如下：

$$X_A=[2.2,3.0]=(2.6;0.4)$$

$$X_B=[2.5,4.5]=(3.5;1.0)$$

$$X_C=[2.7,3.5]=(3.1;0.4)$$

$$X_D=[2.8,3.6]=(3.2;0.4)$$

$$X_E=[1.9,2.8]=(2.35;0.45)$$

求得：

$$\hat{c}=\frac{1}{n}\sum_{i=1}^{n}c_i=\frac{2.6+3.5+3.1+3.2+2.35}{5}=2.95$$

$$\hat{r}=\frac{1}{n}\sum_{i=1}^{n}r_i=\frac{0.4+1.0+0.4+0.4+0.45}{5}=0.53$$

則其模糊樣本變異數為：

$$FS^2=\left(\frac{\sum_{i=1}^{n}(c_i-\hat{c})^2}{n-1};\frac{\sum_{i=1}^{n}(r_i-\hat{r})^2}{n-1}\right)=(0.2175;0.0695)$$

定義 2.10: 區間模糊數之運算 (吳柏林，2005)

令 $X_1 = [a_1, b_1] = (c_1; r_1)$ ， $X_2 = [a_2, b_2] = (c_2; r_2)$ 為兩個區間模糊數

區間相加： $X_1 + X_2 = [a_1, b_1] + [a_2, b_2] = [a_1 + a_2, b_1 + b_2]$

$$= (c_1; r_1) + (c_2; r_2) = (c_1 + c_2; r_1 + r_2)$$

區間相減： $X_1 - X_2 = [a_1, b_1] - [a_2, b_2] = [a_1 - b_2, b_1 - a_2]$

$$= (c_1; r_1) - (c_2; r_2) = (c_1 - c_2; r_1 + r_2)$$

例 2.10: 區間模糊數之相加減

假設 $X_1 = [3,9] = (6;3)$ ， $X_2 = [2,12] = (7;5)$ 則

$$X_1 + X_2 = [3,9] + [2,12] = [5,21] = (13;8)$$

$$X_1 - X_2 = [3,9] - [2,12] = [-9,7] = (-1;8)$$

2.4 模糊線性相關係數

在傳統統計分析中，相關係數在測量兩變數之間的強度關係中扮演著相當重要的角色。然而，人類思維的模糊與不確定性導致許多蒐集到的資料並不能以一個切確的數字來代表，而相對的可能是一段模糊的區間資料。

對於模糊的區間資料，傳統的皮爾森相關系數(Pearson’s Correlation Coefficient)並不足以解釋模糊的概念，考量區間模糊數區間中點與區間半徑兩種對於區間模糊數的影響值，當資料為區間模糊數時，將區間中點與區間半徑兩者的離散程度做加總，得到更具合理性的模糊相關係數。

定義 2.11: 區間模糊數之相關係數(考量區間中點與區間半徑)

若 $X=\{[x_{iL},x_{iU}]\}=\{(c_{xi};r_{xi})\}$ 與 $Y=\{[y_{iL},y_{iU}]\}=\{(c_{yi};r_{yi})\}$ 分別為兩組區間樣本資料。$\overline{c}$ 為 c 的算術平均數，$\overline{r}$ 為 r 的算術平均數。則 X 、Y 的相關係數為:

$$Fr_{x,y}=\frac{\sum_{i=1}^{n}[(c_{xi}-\overline{c}_x)(c_{yi}-\overline{c}_y)+(r_{xi}-\overline{r}_x)(r_{yi}-\overline{r}_y)]}{\sqrt{\sum_{i=1}^{n}[(c_{xi}-\overline{c}_x)^2+(r_{xi}-\overline{r}_x)^2]}\sqrt{\sum_{i=1}^{n}[(c_{yi}-\overline{c}_y)^2+(r_{yi}-\overline{r}_y)^2]}} \quad (2.2)$$

例 2.11: 國中學生數學科每週讀書時間與成績的相關係數

調查五位國中學生數學科每週讀書時間與成績狀況如表 2.5:

表 2.5 五位國中學生數學科每週讀書時間與成績狀況表

學生	讀書時間(小時)	數學成績狀況(分)
A	0-1	30-70
B	1-2	50-90
C	3-5	70-100
D	0-5	20-60
E	4-7	65-95

將五位學生的讀書時間以 X 表示、數學成績狀況以 Y 表示，則五位學生的讀書時間與數學成績調查資料可表示成：

$$X_A = [0,1] = (0.5;0.5)$$

$$X_B = [1,2] = (1.5;0.5)$$

$$X_C = [3,5] = (4;1)$$

$$X_D = [0,5] = (2.5;2.5)$$

$$X_E = [4,7] = (5.5;1.5)$$

$$Y_A = [30,70] = (50;20)$$

$$Y_B = [50,90] = (70;20)$$

$$Y_C = [70,100] = (85;15)$$

$$Y_D = [20,60] = (40;20)$$

$$Y_E = [65,95] = (80;15)$$

五位學生讀書時間與成績狀況分佈如圖 2.2

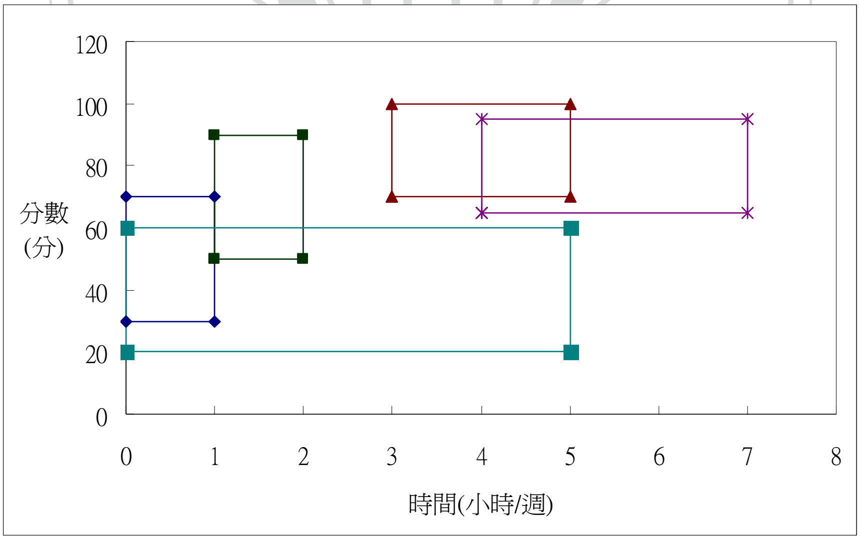

圖 2.2 五位學生讀書時間與成績狀況分佈圖

根據定義 2.11，我們求得五位學生每週數學科讀書時間與數學成績狀況的相關係數為：

$$Fr_{x,y} \sim 0.59$$

區間模糊數的區間半徑在模糊數的意義中，意味著人類思維不確定性強弱的大小，模糊區間數的區間半徑越大，代表著思維的不確定性越高。資料間的區間半徑離散如果越大，代表著彼此思維間的不確定性差異程度越大。考慮區間模糊數區間半徑的影響力，因此本研究將模糊相關係數做另一定義，對皮爾森相關係數的求法做部分調整，將每個樣本區間中心的離散程度與數學常數e做區間半徑離散程度的次方相乘，同時爲了避免相關係數超出我們的認知範圍$(-1\sim1)$，分母根號內也分別乘以對等的數學常數e的半徑離散次方，而得到擴大區間半徑影響力的模糊相關係數。

定義 2.12: 區間模糊數之相關係數(擴大區間半徑影響力)

若 $X=\{[x_{iL},x_{iU}]\}=\{(c_{xi};r_{xi})\}$ 與 $Y=\{[y_{iL},y_{iU}]\}=\{(c_{yi};r_{yi})\}$ 分別為兩組區間樣本資料。e 為數學常數，$\bar{c}$ 為 c 的算術平均數，$\bar{r}$ 為 r 的算術平均數。則 X 、Y 的相關係數為：

$$Fr_{x,y}=\frac{\sum_{i=1}^{n}[(c_{xi}-\bar{c}_x)(c_{yi}-\bar{c}_y)\cdot e^{|r_{xi}-\bar{r}_x|+|r_{yi}-\bar{r}_y|}]}{\sqrt{\sum_{i=1}^{n}[(c_{xi}-\bar{c}_x)^2\cdot e^{2|r_{xi}-\bar{r}_x|}]}\sqrt{\sum_{i=1}^{n}[(c_{yi}-\bar{c}_y)^2\cdot e^{2|r_{yi}-\bar{r}_y|}]}} \quad (2.3)$$

例 2.12: 國中學生每週上網時間與數學週考成績的相關係數

調查五位國中三年級學生，每週上網時間與週考成績狀況如表 2.6:

表 2.6 五位國中學生每週上網時間與數學週考成績狀況表

學生	上網時間(小時)	週考成績狀況(分)
A	1.5-5	50-90
B	5-10	30-70
C	3-5	80-100
D	0-2	40-60
E	5-7	60-95

將五位學生的讀書時間以 X 表示、數學成績狀況以 Y 表示，則五位學生的讀書時間與數學成績調查資料可表示成：

$$X_A = [1.5, 5] = (3.25; 1.75)$$

$$X_B = [5, 10] = (7.5; 2.5)$$

$$X_C = [3, 5] = (4; 1)$$

$$X_D = [0, 2] = (1; 1)$$

$$X_E = [5, 7] = (6; 1)$$

$$Y_A = [50, 90] = (70; 20)$$

$$Y_B = [30, 70] = (50; 20)$$

$$Y_C = [80, 100] = (90; 10)$$

$$Y_D = [20, 60] = (40; 20)$$

$$Y_E = [65, 95] = (80; 15)$$

五位學生每週上網時間與數學週考成績狀況分佈如圖 2.3

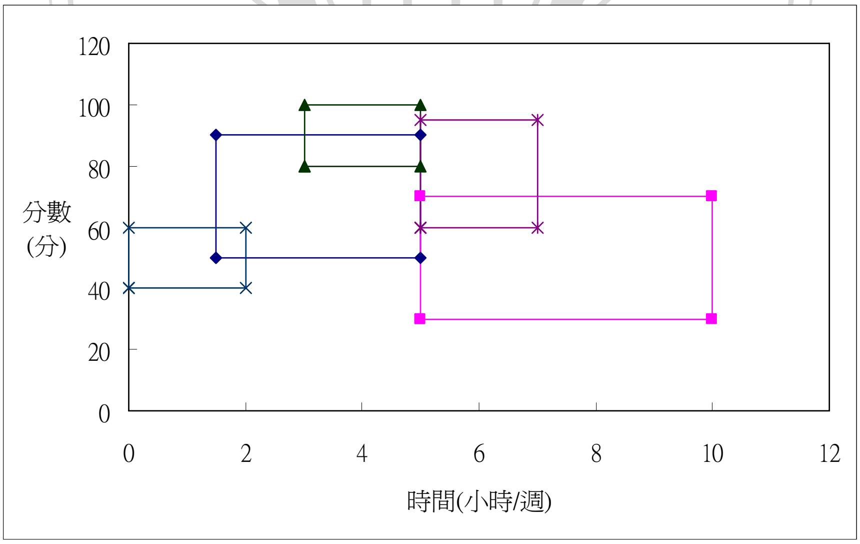

圖 2.3 五位學生每週上網時間與數學週考成績分佈圖

根據定義 2.12，我們求得五位學生每週上網時間與數學週考成績狀況的相關係數為：

$$Fr_{x,y} \sim 0.07$$

對於區間模糊數在統計上的意義，區間模糊數的區間中點才是統計資料研究者一開始想要分析處理的。只是受限於人類知識語言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，不適合過度擴大區間模糊數的區間半徑對於模糊相關係數的影響力。公式 2.3 對於區間模糊數區間半徑過大之資料，亦會造成計算上的困難。同時對於同樣的數據資料，會因爲單位的大小不同而產生些許的差異。

而根據定義 2.11 與定義 2.12 所求得的模糊相關係數，不論樣本是否爲模糊數，所得到的結果均爲一實數，而非模糊數。以模糊統計的觀點來看，更希望得到的模糊相關係數是一個模糊數，因此本研究另外對模糊相關係數作更合適的定義，取得模糊相關區間，以作爲更能有效、適當解釋模糊數之間相關性的模糊相關係數。

定義 2.13: 區間模糊數之相關係數(模糊相關區間)

若 $X=\{[x_{iL},x_{iU}]\}=\{(c_{xi};r_{xi})\}$ 與 $Y=\{[y_{iL},y_{iU}]\}=\{(c_{yi};r_{yi})\}$ 分別為兩組區間樣本資料。

$$r_c=\frac{\sum_{i=1}^{n}(c_{xi}-\overline{c}_x)(c_{yi}-\overline{c}_y)}{\sqrt{\sum_{i=1}^{n}(c_{xi}-\overline{c}_x)^2}\sqrt{\sum_{i=1}^{n}(c_{yi}-\overline{c}_y)^2}}$$ 為兩樣本區間中心點相關係數。

$$r_L=\frac{\sum_{i=1}^{n}(x_{iL}-\overline{x}_L)(y_{iL}-\overline{y}_L)}{\sqrt{\sum_{i=1}^{n}(x_{iL}-\overline{x}_L)^2}\sqrt{\sum_{i=1}^{n}(y_{iL}-\overline{y}_L)^2}}$$ 為兩樣本區間最小值相關係數。

$$r_U=\frac{\sum_{i=1}^{n}(x_{iU}-\overline{x}_U)(y_{iU}-\overline{y}_U)}{\sqrt{\sum_{i=1}^{n}(x_{iU}-\overline{x}_U)^2}\sqrt{\sum_{i=1}^{n}(y_{iU}-\overline{y}_U)^2}}$$ 為兩樣本區間最大值相關係數。

$$r_r=\frac{\sum_{i=1}^{n}(r_{xi}-\overline{r}_x)(r_{yi}-\overline{r}_y)}{\sqrt{\sum_{i=1}^{n}(r_{xi}-\overline{r}_x)^2}\sqrt{\sum_{i=1}^{n}(r_{yi}-\overline{r}_y)^2}}$$ 為兩樣本區間半徑相關係數。

取 $FI=[FI_L,FI_U]=[\min(r_L,r_c,r_U),\max(r_L,r_c,r_U)]=(c_{fr};r_{fr})$ 為 X 、Y 的模糊相關

區間，其中 c_{fr} 為模糊相關區間 FI 的中心點，r_{fr} 為模糊相關區間 FI 的半徑。考慮樣本區間半徑之間的相關程度，對模糊相關係數造成的影響，令以 e 為底的自然對數 ln 函數轉換 $\Delta = 1 - \ln(1+|r_r|)^{\frac{1}{|r_r|}}$ 為修正半徑相關係數。則 X 、Y 的模糊相關係數 $Fr_{x,y}$ 定義為：

(1) 若 $c_{fr} \geq 0$ ，則 $Fr_{x,y} = [FI_L + \Delta \times r_{fr}, FI_U]$

(2) 若 $c_{fr} < 0$ ，則 $Fr_{x,y} = [FI_L, FI_U - \Delta \times r_{fr}]$

例 2.13: 工作年資與其期望薪資相關係數。

調查五位科技業員工，整理其工作年資與期望薪資如表 2.7：

表 2.7 五位科技業員工工作年資與期望薪資表

編號	工作年資	期望薪資(千/月)	
甲	3	35-48	$r_c = 0.96$
乙	1	30-43	$r_L = 0.97$
丙	10	60-90	$r_U = 0.96$
丁	8	55-82	Δ=0
戊	5	50-75	相關係數 = [0.96,0.97]

根據定義 2.13，求得五位科技業員工工作年資與期望薪資兩變數間，$r_c = 0.96$ ，$r_L = 0.97$ ，$r_U = 0.96$ ，而因為工作年資為實數變數，所以兩者間的修正半徑相關係數 Δ=0，相關係數為[0.96, 0.97]，呈現相當高的正相關。繪製五位科技業員工工作年資與期望薪資分佈如圖 2.4：

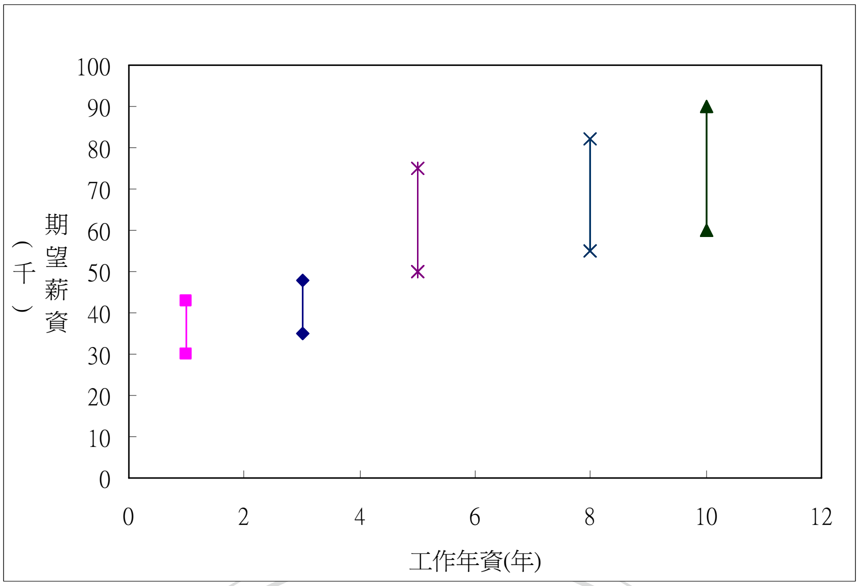

圖 2.4 五位科技業員工工作年資與期望薪資分佈圖

例 2.14: 每週工作時數與其期望薪資相關係數。

調查五位科技業員工，整理其每週工作時數與期望薪資如表 2.8：

表 2.8 五位科技業員工每週工作時數與期望薪資表

編號	每週工作時數(時)	期望薪資(千/月)	
甲	45-60	35-48	$r_c = 0.31$
乙	40-50	30-43	$r_L = 0.52$
丙	50-65	60-90	$r_U = 0.20$
丁	43-52	55-82	Δ=0.077
戊	40-45	50-75	相關係數＝[0.21,0.52]

根據定義 2.13，求得五位科技業員工每週工作時數與期望薪資兩變數間，$r_c = 0.31$，$r_L = 0.52$，$r_U = 0.20$，修正半徑相關係數 Δ=0.077，相關係數為 [0.21, 0.52]。五位科技業員工每週工作時數與期望薪資呈現某種程度的正相關。繪製分布狀況如圖 2.5：

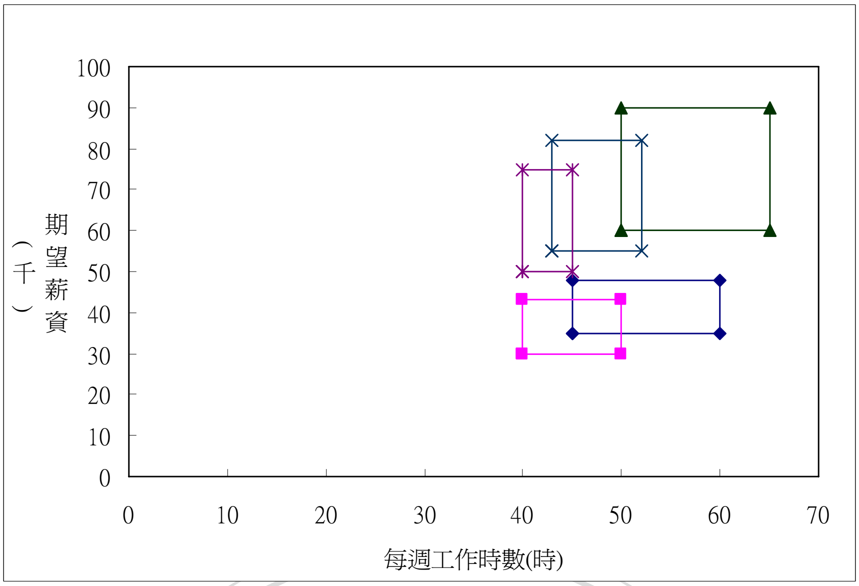

圖 2.5 五位科技業員工工作年資與期望薪資分佈圖

2.5 模糊線性相關係數的性質

性質 2.1:

模糊相關係數的範圍介於-1與1之間。若模糊相關係數$Fr_{x,y}=[a,b]$，其中$a>0$時，我們稱X與Y之間為正相關。若模糊相關係數$Fr_{x,y}=[a,b]$，其中$b<0$時，我們稱X與Y之間為負相關。若模糊相關係數$\left|Fr_{x,y}\right|$的值越接近 1 時，我們稱X與Y之間線性相關越強;當模糊相關係數$\left|Fr_{x,y}\right|$的值越接近0時,則意味著X與Y之間線性相關越弱。

性質 2.2:

離散型模糊數轉換成區間模糊數的過程中，模糊區間半徑的控制係數α，以介於 0 至 2 之間較為恰當。若α值過大時，會造成模糊區間半徑過大而失去其模糊樣本的意義。

性質 2.3:

修正半徑相關係數$\Delta=1-\ln(1+\left|r_r\right|)^{\frac{1}{\left|r_r\right|}}$的範圍為$0\le\Delta\le0.3069$，因為$0\le\left|r_r\right|\le1$。當模糊相關係數區間較大時，修正半徑相關係數對於模糊相關係數區間的影響力相對的比較大。

性質 2.4:

模糊樣本中，若區間模糊數的半徑大小為零(樣本資料為一般實數樣本，非模糊數)時，則模糊相關係數即退化為一般皮爾森相關係數，其中$r_r=\Delta=0$，$Fr_{x,y}=r_c=r_L=r_U$。

性質 2.5:

模糊樣本 X 與 Y 中，若所有區間模糊數的半徑均相同時，所求得之模糊相關係數 $Fr_{x,y}$ 與以區間模糊數區間中點所求得之皮爾森相關係數大小相同。其中 $r_r = \Delta = 0$ ， $Fr_{x,y} = r_c = r_L = r_U$ 。

性質 2.6:

模糊樣本 X 與 Y 中，若樣本 X 中所有區間模糊數的半徑均為某定值 a;樣本 Y 中所有區間模糊數的半徑均為某定值 b ，且 $a \neq b$ ，則所求得之模糊相關係數 $Fr_{x,y}$ 與以區間模糊數的區間中點所求得之皮爾森相關係數大小相同。其中 $r_r = \Delta = 0$ ， $Fr_{x,y} = r_c = r_L = r_U$ 。

3. 實例應用

國中數學教學現場中，教師常要求學生上課要專心聽講、回家要複習當日課程、考前要準備。我們均深知上述的要求會帶給學生課業學習上的穩定學習與表現。學生日常生活中與其數學成績表現值得探討的因數包括：上網的時數、看電視的時間、自我學習數學的時間、閱讀課外讀物的時間等，因為人類思維的不確定性，這些因數所呈現的數值往往會以較具合理反映人類思維的模糊數方式呈現。本章將就國中數學科教學現場所蒐集到的資料，分別探討五人小群體間有序性離散模糊數與區間模糊數之模糊相關係數，以及全班三十位學生所取得之實數資料與區間模糊數資料求模糊相關係數做實例應用。

3.1 有序性離散模糊數與區間模糊數之模糊相關係數

以下就五位國中三年級學生「數學成績表現」、「對數學的喜愛程度」與「每天上網或看電視時數」三組資料，分別求其模糊相關係數。

調查五位國中三年級學生「對數學的喜愛程度」隸屬度、「數學成績表現」區間數與「每天上網或看電視時數」區間數分別如表 3.1、表 3.2、表 3.3：

表 3.1 五位國中三年級學生「對數學的喜愛程度」隸屬度

學生	L_1=非常不喜歡	L_2=不喜歡	L_3=普通	L_4=喜歡	L_5=非常喜歡
A	0.5	0.5			
B			0.3	0.7	
C	0.1	0.3	0.3	0.3	
D			0.2	0.3	0.5
E		0.1	0.9		

表 3.2 五位國中三年級學生「數學成績表現」表

學生	數學成績表現(分)
A	0~70
B	50~90
C	60~85
D	75~95
E	60~90

表 3.3 五位國中三年級學生「每天上網或看電視時數」表

學生	每天上網或看電視時數(分鐘)
A	30~120
B	0~60
C	20~90
D	30~60
E	60~120

依據離散型模糊數轉換成區間模糊數的流程圖(圖 2.1)，將表 3.1 五位國中學生「對數學的喜愛程度」離散型模糊數依步驟轉換成區間型模糊數。

步驟一：賦予語言變數適當權重。

賦予五個語言變數權重：

非常不喜歡=0 ，不喜歡$=0.25$ ，普通=0.5 ，喜歡$=0.75$ ，非常喜歡$=1$ 。

步驟二：利用定義 2.4 將離散型模糊數反模糊化，求得重心 x_f 。

根據定義 2.4，則五位學生反模糊化值如下：

學生 A「對數學喜愛程度」的反模糊化值爲

$$x_{fA} = 0.5 \times 0 + 0.5 \times 0.25 = 0.125$$

學生 B「對數學喜愛程度」的反模糊化值為

$$x_{fB} = 0.3 \times 0.5 + 0.7 \times 0.75 = 0.675$$

學生 C「對數學喜愛程度」的反模糊化值為

$$x_{fC} = 0.1 \times 0 + 0.3 \times 0.25 + 0.3 \times 0.5 + 0.3 \times 0.75 = 0.45$$

學生 D「對數學喜愛程度」的反模糊化值為

$$x_{fD} = 0.2 \times 0.5 + 0.3 \times 0.75 + 0.5 \times 1 = 0.825$$

學生 E「對數學喜愛程度」的反模糊化值為

$$x_{fE} = 0.1 \times 0.25 + 0.9 \times 0.5 = 0.475$$

步驟三：利用定義 2.5 求得離散型模糊數之模糊標準差 $F_x\sigma$ 。

根據定義 2.5，則五位學生之模糊標準差如下：

學生 A「對數學喜愛程度」的模糊標準差為

$$F\sigma_A = \sqrt{0.5 \times (0 - 0.125)^2 + 0.5 \times (0.25 - 0.125)^2} = 0.125$$

學生 B「對數學喜愛程度」的模糊標準差為

$$F\sigma_B = \sqrt{0.3 \times (0.5 - 0.675)^2 + 0.7 \times (0.75 - 0.675)^2} \approx 0.115$$

學生 C「對數學喜愛程度」的模糊標準差為

$$F\sigma_C = \sqrt{0.1 \times (0 - 0.45)^2 + 0.3 \times (0.25 - 0.45)^2 + 0.3 \times (0.5 - 0.45)^2 + 0.3 \times (0.75 - 0.45)^2} \approx 0.245$$

學生 D「對數學喜愛程度」的模糊標準差為

$$F\sigma_D = \sqrt{0.2 \times (0.5 - 0.825)^2 + 0.3 \times (0.75 - 0.825)^2 + 0.5 \times (1 - 0.825)^2} \approx 0.195$$

學生 E「對數學喜愛程度」的模糊標準差為

$$F\sigma_E = \sqrt{0.1 \times (0.25 - 0.475)^2 + 0.9 \times (0.5 - 0.475)^2} = 0.075$$

步驟四：利用定義 2.6，給定適當模糊區間半徑控制係數α，求得離散型模糊數之模糊區間。

給定模糊區間半徑控制系數α=0.5，則五位學生「對數學喜愛程度」的模糊區間如下：

學生 A「對數學喜愛程度」的模糊區間爲

$$Ix_{A,0.5} = [\max\{0, 0.125 - 0.5 \times 0.125\}, \min\{0.125 + 0.5 \times 0.125, 1\}] = [0.0625, 0.1875]$$

學生 B「對數學喜愛程度」的模糊區間爲

$$Ix_{B,0.5} = [\max\{0, 0.675 - 0.5 \times 0.115\}, \min\{0.675 + 0.5 \times 0.115, 1\}] = [0.6175, 0.7325]$$

學生 C「對數學喜愛程度」的模糊區間爲

$$Ix_{C,0.5} = [\max\{0, 0.45 - 0.5 \times 0.245\}, \min\{0.45 + 0.5 \times 0.245, 1\}] = [0.3275, 0.5725]$$

學生 D「對數學喜愛程度」的模糊區間爲

$$Ix_{D,0.5} = [\max\{0, 0.825 - 0.5 \times 0.195\}, \min\{0.825 + 0.5 \times 0.195, 1\}] = [0.7275, 0.9225]$$

學生 E「對數學喜愛程度」的模糊區間爲

$$Ix_{E,0.5} = [\max\{0, 0.475 - 0.5 \times 0.075\}, \min\{0.475 + 0.5 \times 0.075, 1\}] = [0.4375, 0.5125]$$

整理以上步驟所求得之五位學生「對數學喜愛程度」反模糊化值、模糊標準差與模糊區間如表 3.4。

表 3.4 五位國中三年級學生「對數學喜愛程度」模糊區間表

學生	反模糊化值	模糊標準差	區間半徑控制系數	模糊區間
A	0.125	0.125		[0.0625,0.1875]
B	0.675	0.115		[0.6175,0.7325]
C	0.45	0.245	α=0.5	[0.3275,0.5725]
D	0.825	0.195		[0.7275,0.9225]
E	0.475	0.075		[0.4375,0.5125]

五位國中三年級學生「對數學的喜愛程度」與「數學成績表現」模糊區間如表 3.5。

表 3.5 五位學生「對數學喜愛程度」與「數學成績表現」模糊區間表

學生	對數學喜愛程度	數學成績表現(分)
A	[0.0625,0.1875]	0~70
B	[0.6175,0.7325]	50~90
C	[0.3275,0.5725]	60~85
D	[0.7275,0.9225]	75~95
E	[0.4375,0.5125]	60~90

將五位國中三年級學生「對數學喜愛程度」以 X 表示、「數學成績表現」以 Y 表示，五位學生「對數學喜愛程度」與「數學成績表現」資料可表示成：

$$X_A = [0.0625, 0.1875] = (0.125; 0.0625)$$

$$X_B = [0.6175, 0.7325] = (0.675; 0.0575)$$

$$X_C = [0.3275, 0.5725] = (0.45; 0.1225)$$

$$X_D = [0.7275, 0.9225] = (0.825; 0.0975)$$

$$X_E = [0.4375, 0.5125] = (0.475; 0.0375)$$

$$Y_A = [0, 70] = (35; 35)$$

$$Y_B = [50, 90] = (70; 20)$$

$$Y_C = [60, 85] = (72.5; 12.5)$$

$$Y_D = [75, 95] = (85; 10)$$

$$Y_E = [60, 90] = (75; 15)$$

繪製五位學生「對數學喜愛程度」與「數學成績表現」分布狀況如圖 3.1。

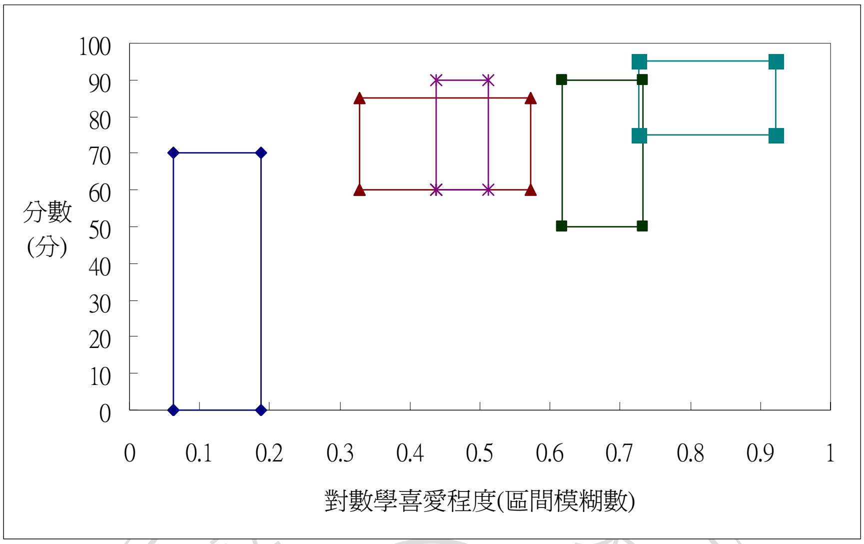

圖 3.1 五位學生「對數學喜愛程度」與「數學成績表現」狀況分布圖

根據定義 2.13，我們求得五位國中三年級學生「對數學喜愛程度」與「數學成績表現」兩組資料間，$r_c = 0.89$，$r_L = 0.83$，$r_U = 0.92$，修正半徑相關係數 Δ=0.17，模糊相關係數爲[0.84, 0.92]。

對於五位國中三年級學生「對數學喜愛程度」與「數學成績表現」這兩組分別爲有序性離散模糊數與區間模糊數，我們利用表 2.1 之步驟將有序性離散模糊數轉換成區間模糊數之後，再以定義 2.13 相關係數的求法求得彼此之間的相關係數。從所求得之結果，相關係數介於 0.84 與 0.92 之間，顯示五位國中三年級學生「對數學喜愛程度」與「數學成績表現」兩組資料之間呈現一定程度的正相關。

根據表 3.2 與表 3.3，整理五位國中三年級學生「每天上網或看電視時數」與「數學成績表現」如表 3.6。

表 3.6 五位學生「每天上網或看電視時數」與「數學成績表現」模糊區間表

學生	每天上網或看電視時數 (分鐘)	數學成績表現(分)
A	30~120	0~70
B	0~60	50~90
C	20~90	60~85
D	30~60	75~95
E	60~120	60~90

將五位國中三年級學生「每天上網或看電視時數」以 X 表示、「數學成績表現」以 Y 表示，五位學生「每天上網或看電視時數」與「數學成績表現」資料可表示成：

$$X_A = [30,120] = (0.125;0.0625)$$

$$X_B = [0,60] = (30;30)$$

$$X_C = [20,90] = (55;35)$$

$$X_D = [30,60] = (45;15)$$

$$X_E = [60,120] = (90;30)$$

$$Y_A = [0,70] = (35;35)$$

$$Y_B = [50,90] = (70;20)$$

$$Y_C = [60,85] = (72.5;12.5)$$

$$Y_D = [75,95] = (85;10)$$

$$Y_E = [60,90] = (75;15)$$

繪製五位學生「每天上網或看電視時數」與「數學成績表現」分布狀況如圖 3.2。

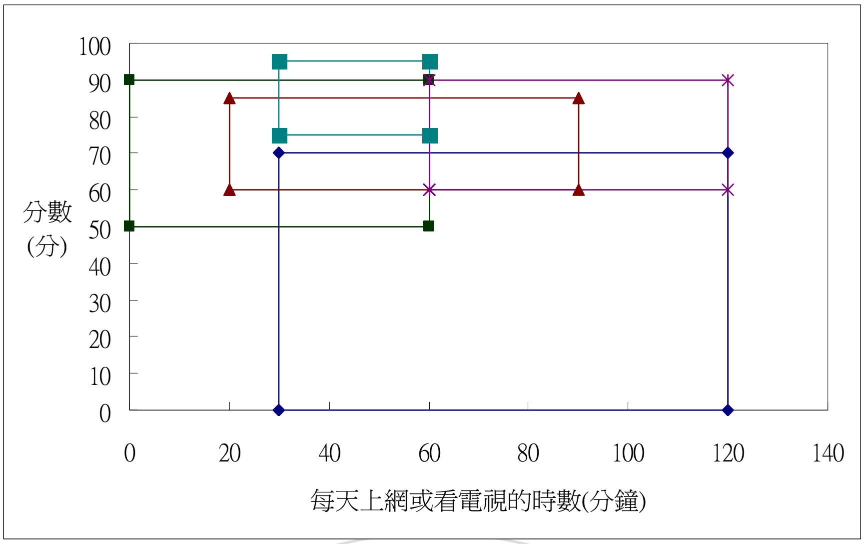

圖 3.2 五位學生「每天上網或看電視時數」與「數學成績表現」狀況分布圖

根據定義 2.13，我們求得五位國中三年級學生「每天上網或看電視時數」與「數學成績表現」兩組資料間，$r_c = -0.34$，$r_L = 0.08$，$r_U = -0.65$，修正半徑相關係數 Δ=0.27 ，相關係數為[-0.65, -0.02]。

對於五位國中三年級學生「每天上網或看電視時數」與「數學成績表現」這兩組均為區間型模糊數的資料，我們直接利用定義 2.13 求得彼此之間的相關係數。從所求得之結果，我們發現所求得之相關係數介於 -0.65 到 -0.02 之間，顯示五位國中三年級學生「每天上網或看電視時數」與「數學成績表現」兩組資料之間呈現某種程度的負相關，而且相關係數的相關區間頗大，我們可以推測學生「每天上網或看電視時數」與「數學成績表現」資料間，受訪者的不確定性亦較大，可能有受訪者數學成績表現良好，但是每天上網或看電視的時數最大值與最小值差距很大；同時也有可能有受訪者每天上網或看電視的時數變動不大，但是數學成績表現起伏很大。

3.2 三十位國中三年級學生影響數學成績表現可能因數之相關係數

影響學生數學成績表現的可能因素很多，這也是教育研究學者想要了解的一大課題。如果將它劃分為個人因素與環境因素，那麼大致上可以指出，影響學生數學成績表現的個人因素有：對於數學的喜好程度、自我要求、智力高低、時間的規劃等。而影響學生數學成績的環境因素有：教師期待、家庭社經地位、同儕影響、居住環境等。這些因素當中，絕大多數是需要較為深入的分析去探討對於學生數學表現的影響力。然而，國中生在時間的規劃上，諸如每週看電視的時間、準備數學科目的時間與閱讀課外讀物的時間等，這些資料是我們可以輕易取得的，我們希望可以藉由相關係數的計算，取得這些因素與數學成績表現的關係，提供教師、家長或學生作為時間規劃的參考。

就我們一般所認為，學生花費越多時間在看電視，自然壓縮到念書的時間，相對的，數學成績表現應該同時也會受到影響；同時，準備數學科目的時間越多，代表該學生較為認真，那麼他的數學成績表現應該也會不錯。而現今各中小學，大力鼓吹閱讀，課外的閱讀對於數學成績是否有影響？本節將以學生「每週看電視時間」、「每週準備數學科目時間」與「每週閱讀課外讀物時間」為主要探討因素，計算上述三個因素與學生數學成績表現之間的相關係數，同時求出個因素彼此之間的相關係數以供參考。

將教學現場中，某國中三年級一個班級為實例，調查班上三十位學生「最近一次數學段考成績」、「自評平時數學成績表現」、「每週看電視時間」、「每週準備數學科目時間」與「每週閱讀課外讀物時間」五項資料數據如表 3.7：

表 3.7 三十位國中三年級學生影響數學成績表現可能因數資料表

編號	段考成績(分)	自評成績(分)	每週看電視時間(時)	準備數學時間(時/週)	課外閱讀時間(時/週)
1	98	85-100	2-8	3-9	0-3
2	77	60-85	5-10	2-7	0.5-4
3	89	50-85	4-8	2.5-8	1-3
4	75	45-80	2-5	3.5-7.5	0.2-1
5	86	60-90	0-4	2.5-8.5	2-4
6	54	35-70	5-15	0.5-5	0-2
7	64	50-70	3.5-6	2.5-5	2-3
8	89	60-85	2.5-4	3-7.5	3-5
9	59	30-60	7-8	1-3	2.5-4
10	97	65-90	5-7	4-8	1-4
11	28	25-55	1-5	0.5-3	4-5
12	31	5-40	0-2	0-2	0-0.5
13	94	70-90	0.5-5	4-7	0.5-2
14	95	80-95	2-3	5-8	1-1.5
15	94	70-90	1-2	8-10	2-2.5
16	97	60-90	2.5-5.5	6-8.5	0-1
17	64	40-70	4-6	1.5-5	0-0.2
18	52	20-55	5-8	3-5	2-2.5
19	42	20-60	1-4	1-4	0-0.1
20	41	30-50	7-9	1-4.5	2-4
21	94	65-90	4-7	5-6	5-6
22	88	60-80	6-8	4-8	1-2
23	34	20-40	10-15	0-1	0-2
24	94	70-95	5.5-8	3-9	1.5-2
25	89	65-90	2-6	5-6	2-6
26	54	40-60	4-7	0-4	0-2
27	86	65-90	5-8	1.5-5	2-8
28	86	65-80	1-3	5-7	3-5
29	53	35-60	7-12	0-2	2-3
30	95	70-90	4-7	2-4	1-3

由表 3.7 資料顯示，除了段考成績爲一實數資料外，其餘均爲區間模糊數。繪製「段考成績」與其他 4 組因素的分布狀況分別如圖 3.3、圖 3.4、圖 3.5、圖 3.6。

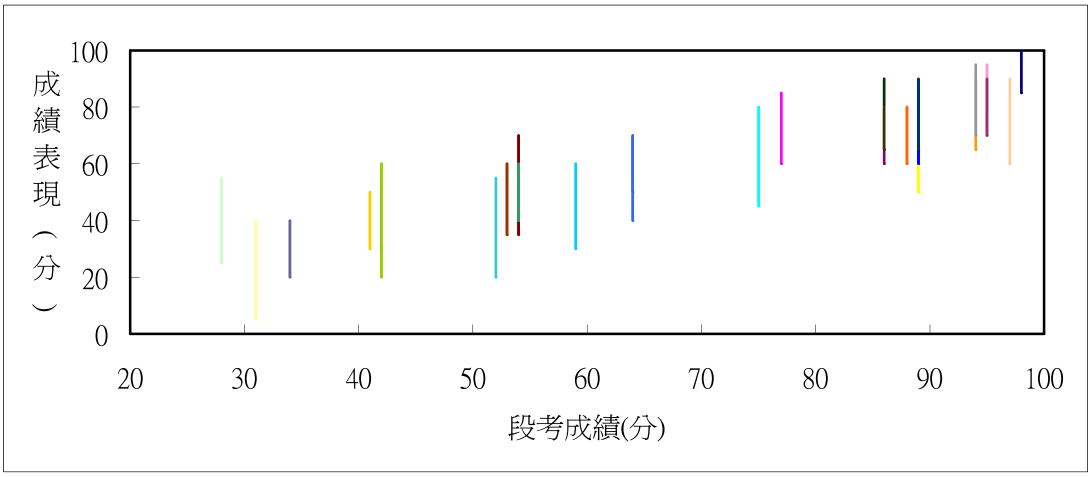

圖 3.3 「段考成績」與「學生自評成績表現」狀況分布圖

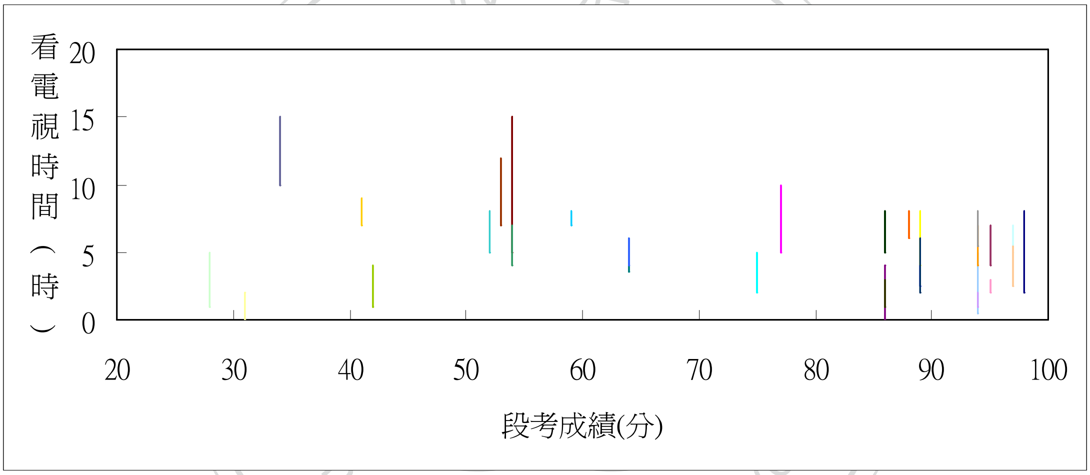

圖 3.4 「段考成績」與「每週看電視時間」狀況分布圖

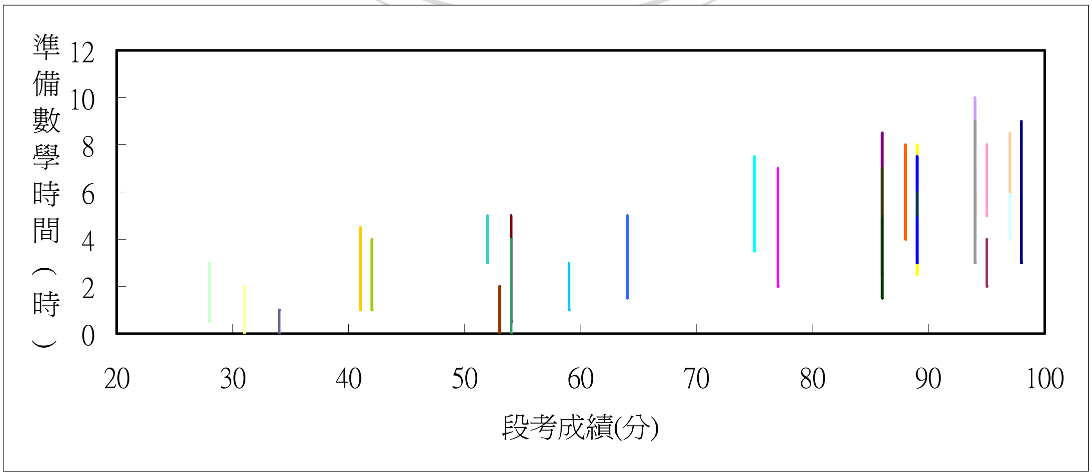

圖 3.5 「段考成績」與「每週準備數學時間」狀況分布圖

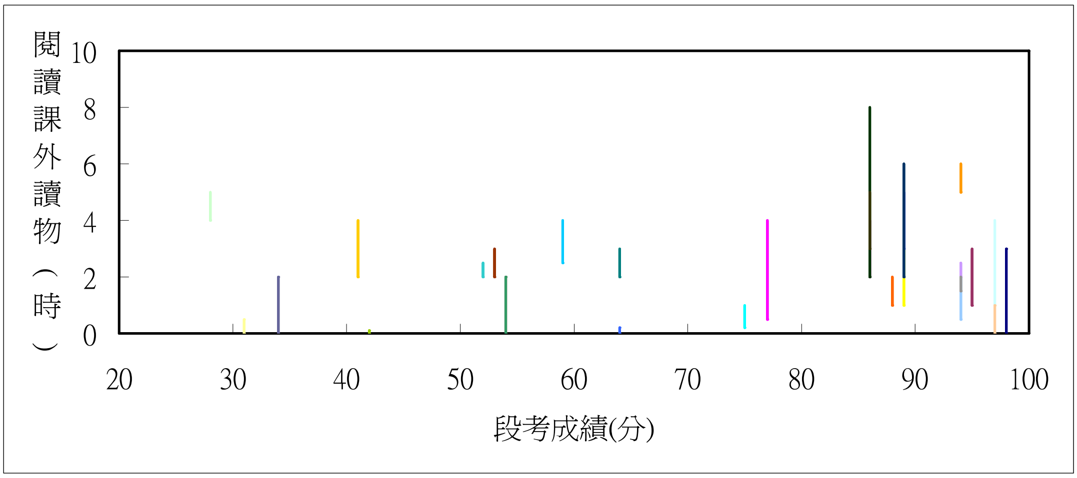

圖 3.6 「段考成績」與「每週閱讀課外讀物時間」狀況分布圖

首先由資料的圖面分布狀況來探討資料間的相關程度。根據圖 3.3，可以判斷三十位學生「段考成績」與「學生自評成績表現」兩組資料之間呈現相當高程度的正相關，學生自評成績表現的中心點越高者，段考成績也相對的比較高。根據圖 3.4，可以判斷學生「段考成績」與「每週看電視時間」約略呈現某種程度的負相關，看電視時間較短者，段考成績表現較好。圖 3.5 與圖 3.3 類似，學生「段考成績」與「每週準備數學時間」也呈現頗高程度的正相關，而圖 3.6 則較難由圖面判斷相關程度。

根據定義 2.13，計算「段考成績」與各影響學生成績表現因素間的相關係數如表 3.8：

表 3.8 「段考成績」與各影響學生成績表現因素間的相關係數

「段考成績」與	r_c	r_L	r_U	Δ	相關係數
學生自評成績表現	0.96	0.94	0.96	0	[0.94, 0.96]
每週看電視時間	-0.27	-0.24	-0.27	0	[-0.27, -0.24]
每週準備數學時間	0.84	0.76	0.82	0	[0.76, 0.84]
課外閱讀時間	0.19	0.08	0.24	0	[0.08, 0.24]

由表 3.8 可以發現，學生自評成績表現與段考成績呈現較高的線性相關，同時相關係數的區間亦較集中，代表多數學生對於自己平時的學習準備有自知之明，而且不確定性亦較低。其次，每週準備數學時間與段考成績也呈現頗高的線性相關，代表學生每週準備數學時間反應了段考成績的表現，但是模糊相關係數的區間較自評成績表現大，顯示不確定性的增加。同時每週看電視時間與數學段考成績呈現微弱的負相關，而課外閱讀時間與數學段考成績則爲微弱的正相關。

根據定義 2.13，列出各因素彼此之間的修正半徑相關係數 Δ 如表 3.9

表 3.9 影響學生數學學業成績表現因素修正半徑相關係數 Δ 表

	段考成績(實數)	自評成績	看電視時間	準備數學時間	課外閱讀時間
段考成績(實數)	Δ =0	Δ =0	Δ =0	Δ =0	Δ =0
自評成績		Δ =0.3069	Δ =0.0817	Δ =0.0365	Δ =0.1041
看電視時間			Δ =0.3069	Δ =0.1010	Δ =0.1149
準備數學時間				Δ =0.3069	Δ =0.0766
課外閱讀時間					Δ =0.3069

由於數學段考成績爲一實數資料，各組數據的區間半徑皆爲 0，所以段考成績與其他資料間的修正半徑相關係數均爲 0。而相同的兩組資料之間，因爲半徑相關係數爲 1，所以修正半徑相關係數 Δ 均爲 0.3069。其餘修正半徑相關係數 Δ 均介於 0 到 0.3069 之間，以「看電視時間」與「課外閱讀時間」的修正半徑相關系數 Δ =0.1149 爲最大，顯示受訪者對於此兩項目之間不確定性的模糊區間半徑較爲一致。

根據定義 2.13，，計算各因素彼此之間的相關係數 $Fr_{x,y}$ 值如表 3.10。

表 3.10 影響學生數學學業成績表現因素相關係數 $Fr_{x,y}$ 表

	段考成績(實數)	自評成績	看電視時間	準備數學時間	課外閱讀時間
段考成績(實數)	1	[0.94, 0.96]	[-0.27, -0.24]	[0.76, 0.84]	[0.08, 0.24]
自評成績		1	[-0.28, -0.24]	[0.69, 0.83]	[0.12, 0.26]
看電視時間			1	[-0.45, -0.38]	[-0.03, 0.06]
準備數學時間				1	[0.03, 0.22]
課外閱讀時間					1

根據表 3.10，可以發現除了「段考成績」與「自評成績」、「準備數學時間」、「課外閱讀時間」呈現正相關之外，「自評成績」與「準備數學時間」、「課外閱讀時間」以及「準備數學時間」與「課外閱讀時間」也呈現正相關。而「看電視時間」與「段考成績」、「自評成績」、「準備數學時間」、「課外閱讀時間」則呈現負相關，至於「看電視時間」與「課外閱讀時間」則較無相關性。與一般我們所認知的狀況相吻合。少看電視，成績會比較好；多閱讀的人，成績也會比較好。

4. 結論與建議

因為人類思維的複雜，對於許多問題回答會以模糊的方式呈現，例如：用於閱讀的時間、對於薪資的要求、購買房屋的可接受價位等。傳統的皮爾森相關係數無法提供模糊統計研究者其模糊樣本彼此的關係為何。本論文有效且便捷的提出區間模糊數相關係數之求法，用以了解區間模糊數之間的相關程度，解決因人類思維的複雜與不明確性的喜好而產生的模糊資訊，提供區間模糊數之間相關程度關係更好的解釋能力。首先將區間模糊數$[a,b]$轉換成$(c;r)$的中心與半徑型態，其中$c=(a+b)/2$為區間$[a\ ,b]$之中心，$r=(b-a)/2$為區間長度的半徑，配合傳統的皮爾森相關係數求法，分別求出區間中心點相關係數、區間最大值相關係數、區間最小值相關係數、區間半徑相關系數，並且利用對數函數轉換取得修正半徑相關系數，即可計算出模糊區間相關系數。而對於有序性離散型模糊數，則可在給予語言變數適當權重後，利用其反模糊化值x_f與標準差$F_x\sigma$，配合給定適當的α值，轉換成區間模糊數以計算其模糊相關係數。

相較於傳統的皮爾森相關係數，本論文所提出的模糊相關係數能處理兩組資料均為區間模糊數的情形，或其中一組為實數資料的情形，若兩組資料均為實數則退化成一般的皮爾森相關係數，適用於變數為不同類型的組合。

以下根據本論文的結果及發現，提出一些建議，以作為應用上及後續研究之參考。

1. 本論文中，用於修正模糊相關係數區間的修正半徑相關係數Δ，採用對數函數轉換，大小介於 0 至 0.3069 之間。未來研究者可以依據區間半徑的影響力考慮加入調整係數，但建議調整後之Δ應介於 0 至 1 之間
2. 在後續研究上，未來可對於非有序性的模糊數提出轉換成區間模糊數的方法或其模糊相關係數的研究與探討，以彌補本研究中未能對非有序性模糊數提出模糊相關係數求法的遺憾。

5. 參考文獻

[1] 江彥聖 (2008)。模糊相關係數及其應用。碩士論文，國立政治大學，台北市。

[2] 吳柏林 (2005)。模糊統計導論：方法與應用。台北：五南。

[3] Buckley, J. J. (2003). *Fuzzy Probabilities: New Approach and Applications*, Physics-Verlag, Heidelberg, Germany.

[4] Buckley, J. J. (2004). *Fuzzy Statistics*, Springer-Verlag, Heidelberg, Germany.

[5] Carlsson, C., Fuller, R. (2001). On possibilistic mean value and variance of fuzzy numbers, *Fuzzy Sets and Systems* 122, 315-326.

[6] Dubois, D. and Prade, H (1987). The mean value of a fuzzy number, *Fuzzy Sets and Systems* 24, 279-300.

[7] Galvo, T., and Mesiar, R. (2001). Generalized Medians, *Fuzzy Sets and Systems* 124, 59-64.

[8] Gebhardt, J., Gil, M. A., and Kruse, R. (1998). *Fuzzy set-theoretic methods in statistics, in: R. Slowinski(Ed.), Handbook on Fuzzy Sets, Fuzzy Sets in Decision Anaysis, Operations Research, and Statistics, vol. 5*, Kluwer Academic Publishers, New York, 311-347.

[9] H. T. Nguyen and B. Wu (2006) *Fundamentals of Statistics with Fuzzy Data.* New York：Springer.

[10] Heilpern, S. (1992). The expected value of a fuzzy number, *Fuzzy Sets and Systems* 47, 81-86.

[11] Hung, W. L. and Wu, J. W., (2001). A note on the correlation of fuzzy numbers by Expected interval, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 9, 517-523.

[12] Hung, W. L. and Wu, J. W., (2002). Correlation of fuzzy numbers by α-cut

method, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 725-735.

[13] Korner, R. (1997). On the Variance of Fuzzy Random Variables, *Fuzzy Sets and Systems* Vol. 92, 83-93.

[14] Liu, S. T. and Kao, C. (2002). Fuzzy measures for correlation coefficient of fuzzy numbers. *Fuzzy Sets and Systems* 128, 267-275

[15] Wu, H. C. (1999). Probability density functions of Fuzzy Random Variables. *Fuzzy Sets and Systems* Vol. 105, 139-158.

[16] Zadeh, L.A., (1965). Fuzzy sets. *Information and Control* 8, 338-353.