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

模糊資料相關係數的估算及應用

Correlation Evaluation with Fuzzy Data and its Applications

碩專班學生：黃玟綾 撰

指導教授：吳柏林 博士

中華民國 101 年 8 月 27 日

摘要

在統計學上，研究分析兩變數間是否存在某程度的相關性，常使用皮爾森相關係數(Pearson’s Correlation Coefficient)，來表達兩實數變數間線性相關的強度及方向，但是當資料並非明確實數，而是反映人類思維不確定性的模糊資料時，傳統的相關分析方法並不適用於貼切傳達模糊資料的訊息。本研究探討有序性離散型模糊資料的相關性，提出四種類型的相關係數定義，提供研究者對此類型資料間相關程度合理的分析方法。本研究實證分析以教育心理問卷，來探究影響未成年抽菸學生菸癮程度的相關性因素。

關鍵字：模糊相關係數、模糊資料、模糊統計

Abstract

In statistical studies, we want to analyze whether there is some degree of correlation between two variables. The Pearson correlation coefficient is often used to convey the strength and direction of the linear correlation between the two real number variables. When the information is not clear real number, but fuzzy information to reflect the uncertainty of the human mind, the traditional analysis method does not apply to appropriate to convey the message of the fuzzy data. This study proposed four types of definition of the correlation coefficients between the ordered discrete fuzzy data. The empirical research used analysis of educational psychology questionnaire to explore the relevance factors of the degree of correlation of underage smoking student smokers.

Keywords： Fuzzy correlation coefficient, Fuzzy data, Fuzzy statistics

目次

表目次

1. 前言

人類的思維主要是來自於對自然現象和社會現象的認知意識，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。模糊理論的產生即是參考人類思維方式對環境所用的模糊測度與分類原理，給予較穩健的描述方式，以處理多元複雜的曖昧和不確定現象(吳柏林，2005)。

自 Lotfi A. Zadeh(1965)提出模糊集合理論(fuzzy set theory)以來，提供了解決人類思維模糊問題的方法，以模糊邏輯(fuzzy logic)為基礎理論，利用邏輯測度並運用模糊統計來代替傳統統計方法。隸屬度函數是模糊理論的基本概念，用隸屬度函數來描述模糊集合，對模糊集合進行量化，用數學方法來處理模糊資訊進而探討分析。

統計分析廣泛應用於自然科學、社會科學領域，欲探討兩變數間線性相關的強度，傳統上我們使用皮爾森相關係數(Pearson’s correlation coefficient)來表達出兩者間關係程度與方向。然而傳統皮爾森相關係數處理的資料皆數明確的實數，若欲處理的資料為模糊資料時，則須探究其它方式來求出模糊相關係數(fuzzy correlation coefficient)，以達到分析研究之目的。模糊資料的相關係數之相關研究，有其值得探討的意義。近年來模糊理論應用的相關研究不少。Lowen(1990)、Tseng & Klein(1992)、Ruspini(1991)、Wu & Sun(1996)、Kosko(1993)、Guarisom Rizzoli, &Werthner(1992)則提出許多模糊理論在社會科學計量的重要觀念與方法。Liu 與 Kao(2002)定義了模糊資料相關係數公式。Hung 與 Wu(2001,2002)應用期望區間及 α-截集($\alpha - cuts$)方法來分析模糊數間的相關性。江彥聖(2008)、林立夫(2011)提出模糊相關係數的定義，然而這兩篇文獻中對於離散型有序性模糊資料的相關係數求法，皆是先轉換成區間模糊資料，再給予區間模糊資料的相關係數定義。雖文獻上提供許多不同的公式，但模糊相關係數的估算，仍有待發展研究。

隸屬度函數通常可分為離散型(discrete)與連續型(continuous)兩類(吳柏林，2005)，探討模糊資料的相關性便可分為:離散型模糊數與離散型模糊數、連續型模糊數與連續型模糊數、離散型模糊數與連續型模糊數，然而離散型模糊數間的相關係數算法相關的文獻裡，提供的方法較為繁複，本研究探討離散型有序性模糊數間的相關性估算方法，提出四個不同種類的相關係數定義。應用傳統皮爾森相關係數來探究模糊相關係數，以及應用母數統計的史比爾曼等級相關係數(Spearman Rank Correlation)來探討模糊相關係數，還應用謝名娟、吳柏林（2012）所提出計算連續型區間模糊數之相關係數模糊區間的定義，來探究離散型有序性模糊資料相關係數。本研究在心理統計研究的問卷方法上，提供了資料相關性分析探討的進一步發展。

2. 研究方法

2.1 模糊理論

模糊數及軟計算

語言的意旨在於和世界溝通。藉研究語意，我們將 更了解這個世界共識。然而即使人們用相同的字眼，文法或語法。欲成功的交 流卻仍嫌不夠，我們希望對相同的東西亦要有同樣程度意指。也就是說，我們必須有一個有關語意的協議。因此如何分析人類的意指與推論，如何測量它真正感覺的程度，越來越重要。我們借語意計量方法來幫助我們理解語言，進而來檢視我們的理論。若能建立一語言計算模型，便擁有和世界溝通的有利工具(吳柏林,2005)。

隸屬度函數的概念

隸屬度函數通常可分為離散型(discrete)與連續型(continuous)兩類。離散型隸屬度函數是直接給定有限模糊集合內每個元素的隸屬度，並以向量的形式表現出來，而連續型隸屬度函數則有幾種常用的函數形式（S-函數、Z-函數、π-函數、三角形函數、梯形函數、高斯(指數)函數）來描述模糊集合。函數定義的表現，可以是無限模糊集合的元素及其隸屬度之間的關係，也可以是有限模糊集合的元素及其隸屬度之間的關係(吳柏林,2005)。

定義 2.1：隸屬度函數(吳柏林，2005)

設在論域U上給定映射μ，及μ ： $U \to [0\ 1]$確定了U上一個模糊數(Fuzzy)集合$\tilde{A}$， $\mu_{\tilde{A}}$稱為$\tilde{A}$的隸屬度函數， $\mu_{\tilde{A}}(u)$稱為u對$\tilde{A}$的隸屬度，其表示u屬於$\tilde{A}$的程度。

定義 2.2：模糊數（*Fuzzy Number*）(吳柏林，2005)

設U為一論域，令$\{L_1, L_2, ..., L_n\}$為論域U的因子集。 μ為一對應到[0, 1]間的實數函數，即 μ ： $U \rightarrow [0\ 1]$ 。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$\{\mu_1(X), \mu_2(X), ..., \mu_n(X)\}$表示，

(1)當U為離散時，則述句X之模糊數可表示成：

$$\mu_U(X) = \frac{\mu_1(X)}{L_1} + \frac{\mu_2(X)}{L_2} + ... + \frac{\mu_n(X)}{L_n}$$

其中「+」是或的意思，$\frac{\mu_i(X)}{L_i}$表示述句X隸屬於因子集L_i的程度，且

$\sum_{i=1}^{n} \mu_i(X) = 1$。

(2)當U為連續時，則述句X之模糊數可表示成：

$$\mu_U(X) = \int_{x \in X} \frac{\mu_i(X)}{L_i}$$ 。

例 2.2:抽菸的學生對評斷「自我管理能力」認知程度模糊數表示

問卷調查五位有抽菸的高二學生，X為有抽菸青少年對自我管理能力的自我認知評斷程度，以模糊數表示為$\mu_U(X)$，其論域為$U = \{$ L_1=非常差，L_2 = 稍微差，L_3 = 普通，L_4 = 稍微良好，L_5 = 非常良好$\}$。

表 2.1 五位抽菸的高二學生自我管理能力的自我認知評斷程度的隸屬度

學生	L_1=非常差	L_2＝稍微差	L_3＝普通	L_4＝稍微良好	L_5＝非常良好
A				0.8	0.2
B		0.6	0.2	0.2	
C			0.6	0.4	
D		0.7	0.3		
E	0.2	0.3	0.4	0.1	

其中 E 同學對「自我管理能力的自我認知評斷程度」的隸屬度函數為

$\{\mu_1(X)=0.2,\mu_2(X)=0.3,\mu_3(X)=0.4,\mu_4(X)=0.1,\mu_5(X)=0\}$

亦可以模糊數表示為

$$\mu_U(X)=\frac{0.2}{L_1}+\frac{0.3}{L_2}+\frac{0.4}{L_3}+\frac{0.1}{L_4}+\frac{0}{L_5}$$

研究者在進行離散型模糊問卷時，必須針對所要研究的目標主題，設計各種語言變數，並賦予每個語言變數所對應的權重數值，以便於將模糊統計資料進行分析研究。

定義 2.3：離散型模糊數的語言變數之偏好權重(吳柏林，2005)

設U為一論域，令$L=\{L_1,L_2,...,L_n\}$為佈於論域U上的n個語言變數，對應各語言變數，分別賦予語言變數之偏好權重為$r=\{r_1,r_2,...,r_n\}$，若偏好權重$r=\{r_1,r_2,...,r_n\}$為一有序序列，且$o\le r_1<r_2<...<r_n\le 1$，則定義$r=\{r_1,r_2,...,r_n\}$為偏好遞增序列；反之，若$o\ge r_1>r_2>...>r_n\ge 1$，則定義$r=\{r_1,r_2,...,r_n\}$為偏好遞減序列。

例 2.3：抽菸的學生對「校園團體生活」喜好程度

調查五位會抽菸的高二學生對校園團體生活喜好程度隸屬度如表 2.2

表 2.2　五位抽菸的高二學生對校園團體生活喜好程度之隸屬度

學生	L_1=非常不喜歡	L_2 = 稍微不喜歡	L_3 = 普通	L_4 = 稍微喜歡	L_5 = 非常喜歡
A		0.2	0.2	0.6	
B			0.2	0.7	0.1
C			0.5	0.5	
D				0.3	0.7
E			0.1	0.2	0.7

我們分別賦予五個語言變數的喜愛程度權重為:

非常不喜歡=1，稍微不喜歡=2，普通=3，稍微喜歡=4，非常喜歡=5

此為一偏好遞增序列。

定義 2.4：離散型模糊數的重心

設 X 為一模糊數，語言變數 $\{L_1, L_2, ..., L_n\}$ 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，且 $\sum_{i=1}^{n} \mu_{L_i}(X) = 1$，則稱 $c_x = \sum_{i=1}^{n} m_i L_i$

為模糊數 X 的重心。

例 2.4：抽菸的學生對「校園團體生活」喜好程度的重心

根據表 2.2 ，會抽菸的五位高二學生對校園團體生活喜好程度的調查資料，我們分別賦予五個語言變數的喜愛程度權重為:

非常不喜歡=1，稍微不喜歡=2，普通=3，稍微喜歡=4，非常喜歡=5

則這五位學生資料的重心為:

學生 A: $0.2\times2+0.2\times3+0.6\times4=2.2$

學生 B: $0.2\times3+0.7\times4+0.1\times5=3.9$

學生 C: $0.5\times3+0.5\times4=3.5$

學生 D: $0.3\times4+0.7\times5=4.7$

學生 E: $0.1\times3+0.2\times4+0.7\times5=4.6$

定義 2.5：離散型模糊數的標準差

設 X 為一模糊數，語言變數 $\{L_1, L_2, ..., L_n\}$ 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，且 $\sum_{i=1}^{n}\mu_{L_i}(X)=1$，則稱 $c_x=\sum_{i=1}^{n}m_iL_i$ 為模糊數 X 的重心，稱 $d_x=\dfrac{\sum_{i=1}^{n}m_i\left|L_i-c_x\right|}{n-1}$ 模糊數 X 的標準差。

例 2.5：抽菸的學生對「校園團體生活」喜好程度的標準差

根據表 2.2 ，會抽菸的五位高二學生對校園團體生活喜好程度的調查資料，我們分別賦予五個語言變數的喜愛程度權重為:

非常不喜歡=1，稍微不喜歡=2，普通=3，稍微喜歡=4，非常喜歡=5

則這五位學生資料的標準差為:

學生 A: $\dfrac{0.2\left|2-2.2\right|+0.2\left|3-2.2\right|+0.6\left|4-2.2\right|}{5-1}=0.32$

學生 B: $\dfrac{0.2\left|3-3.9\right|+0.7\left|4-3.9\right|+0.1\left|5-3.9\right|}{5-1}=0.09$

學生 C: $\dfrac{0.5|3-3.5|+0.5|4-3.5|}{5-1}=0.125$

學生 D: $\dfrac{0.3|4-4.7|+0.7|5-4.7|}{5-1}=0.105$

學生 E: $\dfrac{0.1|3-4.6|+0.2|4-4.6|+0.7|5-4.6|}{5-1}=0.125$

定義 2.6: 離散型模糊數的反模糊化值

設 X 為一模糊數，語言變數$\{L_1, L_2, ..., L_n\}$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，且$\sum_{i=1}^{n}\mu_{L_i}(X)=1$，則稱 $c_x=\sum_{i=1}^{n}m_iL_i$ 為模糊數 X 的重心，稱 $d_x=\dfrac{\sum_{i=1}^{n}m_i(L_i-c_x)}{n-1}$ 模糊數 X 的標準差。稱 $X_f=c_x+d_x$ 離散型模糊樣本 X 的反模糊化值。

例 2.6：抽菸的學生對「校園團體生活」喜好程度的反模糊化值

根據表 2.2，會抽菸的五位高二學生對校園團體生活喜好程度的調查資料，我們分別賦予五個語言變數的喜愛程度權重為:

非常不喜歡=1，稍微不喜歡=2，普通=3，稍微喜歡=4，非常喜歡=5

則這五位學生資料的反模糊化值為:

學生 A: $2.2+0.32=2.52$

學生 B: $3.9+0.09=3.99$

學生 C: $3.5+0.125=3.625$

學生 D: $4.7+0.105=4.805$

學生 E: $4.6+0.125=4.725$

2.2 模糊相關係數

傳統實數樣本之線性相關係數

如果我們想要了解 X 、Y 兩個變數間的相關程度，可直接將 (X,Y) 的資料散佈圖畫出來，由散佈圖的呈現狀況，可約略觀察出兩組資料之間的相關性。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。其定義為：

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

(1) $\rho>0$ 時，稱 X 與 Y 兩變數為正相關；

(2) $\rho<0$ 時，稱 X 與 Y 兩變數為負相關；

(3) $\rho=0$ 時，稱 X 與 Y 兩變數沒有關係存在，或稱統計無關。

在實務應用上，母體的變異數 σ_X^2 、σ_Y^2 以及他們之間的共變異數 $Cov(X,Y)$ 並不容易得到，所以，我們用樣本相關係數 γ 來估算 ρ ，其定義為；

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

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。

傳統實數樣本相關係數的特性如下：

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) $-1\le\gamma\le1$。

(2) γ 值的絕對值大小表示線性關係的強度，正、負號表示相關的方向。即：

1. $\gamma>0$ ，(x,y) 散佈的圖形為一帶狀且從左下方至右上方。
2. $\gamma<0$ ，(x,y) 散佈的圖形為一帶狀且從左上方至右下方。
3. $\gamma=1$，所有 (x,y) 均落於一正斜率直線上(完全線性正相關)。
4. $\gamma=-1$，所有 (x,y) 均落於一負斜率直線上(完全線性負相關)。

5. $|\gamma|$ 值越大，則表示線性關係強度越大。

(3) γ 的值愈接近 0，則線性相關愈弱。

史比爾曼等級相關檢定(Spearman Rank Correlation Test)

我們要分析兩個隨機變數之間的關係，在實數樣本相關分析中，我們使用皮爾森相關係數(Pearson’s Correlation Coefficient) ρ 來表達兩變數肩線性相關的強度。但是在判定兩變數之皮爾森相關係數 ρ 是否為零時，檢定方法必須基於樣本的母群體來自常態分配。若樣本的母群體不滿足常態分配的假設，那麼我們必須使用史比爾曼等級相關係數，對資料作等級判定。史比爾曼等級相關係數記作 γ_s，

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

其中 $d_i = R(X_i) - R(Y_i)$ 表示成對樣本之等級差；$R(X_i)$ 表示 X_i 在 X 中之等級，$R(Y_i)$ 表示 Y_i 在 Y 中之等級，n 為樣本數。

史比爾曼等級相關係數 r_s 的解釋與皮爾森相關係數 ρ 類似：

(1) 當兩個變數 X 、Y 的等級順序完全一致時，$r_s = 1$；

(2) 當兩個變數 X 、Y 的等級順序完全相反時，$r_s = -1$；

(3) $r_s = 0$ 則表示兩個變數 X 、Y 無關。

但是史比爾曼等級相關係數的計算過程比皮爾森相關係數 ρ 簡單，而且適用於等級資料的相關係數檢定。它的優點是，不受極端值(extreme value)的影響。這在處理屬質資料，或觀察值的數值大小較不客觀時，例如滿意度指數調查，藝術音樂體育的評分，藥物的有效性等，提供一個很好的統計檢定方法。

無母數統計使用的時機：

(1) 母體並不服從常態分配或母體來自某一未知分配時。

(2) 抽樣樣本資料以類別尺度表示時。

(3) 抽樣樣本資料以等級順序尺度表現時。

定義 2.7:離散型模糊數之相關係數第一類型

設 $X=\{x_1,x_2,...,x_n\}$、$Y=\{y_1,y_2,...,y_n\}$ 為兩組在語言變數 $\{L_1,L_2,...,L_n\}$ 為論域 U 中的有序數列，$c_{xi}=\sum_{k=1}^{n} m_k L_k$ 為模糊數 x_i 的重心，$c_{yi}=\sum_{k=1}^{n} m_k L_k$ 為模糊數 y_i 的重心，$d_{xi}=\dfrac{\sum_{k=1}^{n} m_k (L_k-c_x)}{n-1}$ 為模糊數 x_i 的標準差，$d_{yi}=\dfrac{\sum_{k=1}^{n} m_k (L_k-c_x)}{n-1}$ 為模糊數 y_i 的標準差。

X 、Y 兩組離散型模糊數重心的相關係數為：

$$r_c=\frac{\sum_{i=1}^{n}(c_{ci}-\overline{c_x})(c_{yi}-\overline{c_y})}{\sqrt{\sum_{i=1}^{n}(c_{ci}-\overline{c_x})^2}\sqrt{\sum_{i=1}^{n}(c_{yi}-\overline{c_y})^2}}$$

X 、Y 兩組離散型模糊數標準差的相關係數為：

$$r_d=\frac{\sum_{i=1}^{n}(d_{xi}-\overline{d_x})(d_{yi}-\overline{d_y})}{\sqrt{\sum_{i=1}^{n}(d_{xi}-\overline{d_x})^2}\sqrt{\sum_{i=1}^{n}(d_{yi}-\overline{d_y})^2}}$$

X 、Y 兩組離散型模糊數的相關係數為：

$$r_{FC}=\beta_1 r_c+\beta_2 r_d \text{ ，其中 } \beta_1+\beta_2=1 \text{ 。}$$

β_1 、β_2 的選擇依實際狀況需求做調整，例如若認為重心的相關係數比離散程度的相關係數重要，可建議選擇 $\beta_1=0.7$ 、$\beta_2=0.3$ 。

離散型模糊數在統計上，其重心一開始事資料研究者想分析的，但人類的思維對語言變數有著模糊性，由定義 2.7 所得到的相關係數是一個實數，並非模糊數，在模糊統計的觀點，希望得到的模糊相關係數為一個區間模糊數。在考慮離

散程度相關係數對模糊相關係數的影響之下，將兩組離散型模糊數離散程度(標準差)的相關係數做調整。

定義 2.8:離散型模糊數之相關係數第二類型

設 $X=\{x_1,x_2,...,x_n\}$ 、 $Y=\{y_1,y_2,...,y_n\}$ 為兩組模糊樣本，在語言變數 $\{L_1,L_2,...,L_m\}$ 為論域 U 中的有序數列。

$$\Delta\gamma_d = 1-\frac{\ln(1+|\gamma_d|)}{|\gamma_d|} \text{；其中 } r_d = \frac{\sum_{i=1}^{n}(d_{xi}-\overline{d_x})(d_{yi}-\overline{d_y})}{\sqrt{\sum_{i=1}^{n}(d_{xi}-\overline{d_x})^2}\sqrt{\sum_{i=1}^{n}(d_{yi}-\overline{d_y})^2}}$$

X 、 Y 兩組離散型模糊數的模糊相關係數為：

(i) 若 $r_c \ge 0, r_d \ge 0$ ，模糊相關係數區間 $=(r_c, \min(1, r_c+\Delta r_d))$

(i) 若 $r_c \ge 0, r_d < 0$ ，模糊相關係數區間 $=(r_c-\Delta r_d, r_c)$

(i) 若 $r_c < 0, r_d \ge 0$ ，模糊相關係數區間 $=(r_c, r_c+\Delta r_d)$

(i) 若 $r_c < 0, r_d < 0$ ，模糊相關係數區間 $=(\max(-1, r_c-\Delta r_d), r_c)$

定義 2.9:離散型模糊數之相關係數第三類型

設 $X=\{x_1,x_2,...,x_n\}$ 、 $Y=\{y_1,y_2,...,y_n\}$ 為兩組在語言變數 $\{L_1,L_2,...,L_m\}$ 為論域 U 中的有序數列，兩組模糊樣本 X 、 Y 的反模糊化值分別為 X_{fi} 、 Y_{fi} ， $i=1,2,...,n$ ，則 X 、 Y 的相關係數：

$$\gamma_s = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n(n^2-1)}$$

其中 $d_i = R(X_{fi}) - R(Y_{fi})$ 表示成對樣本的等級差，$R(X_{fi})$ 表示 X_{fi} 在 X 中之等級，$R(Y_{fi})$ 表示 Y_{fi} 在 Y 中之等級，n 為樣本數。

定義 2.10:離散型模糊數之相關係數第四類型

設 $X = \{x_1, x_2, ..., x_n\}$、$Y = \{y_1, y_2, ..., y_n\}$ 為兩組在語言變數 $\{L_1, L_2, ..., L_m\}$ 為論域 U 中的有序數列，兩組模糊樣本 X 、Y 的反模糊化值分別為 X_{fi} 、Y_{fi} ，$i = 1, 2, ..., n$ ，則 X 、Y 的相關係數：

$$\gamma_p = \frac{\sum_{i=1}^{n}(x_{fi} - \overline{x_f})(y_{fi} - \overline{y_f})}{\sqrt{\sum_{i=1}^{n}(x_{fi} - \overline{x_f})^2}\sqrt{\sum_{i=1}^{n}(y_{fi} - \overline{y_f})^2}}$$

2.3 模糊相關係數的性質

性質 2.1：

因為標準差相關係數範圍為 $-1 \le \gamma_d \le 1$，所以修正標準差相關係數

$\Delta\gamma_d = 1 - \dfrac{\ln(1+|\gamma_d|)}{|\gamma_d|}$，其範圍為 $0 \le \Delta\gamma_d \le 0.3069$。

性質 2.2：

第一類型相關係數 $r_{FC} = \beta_1 r_c + \beta_2 r_d$，其中 $\beta_1 + \beta_2 = 1$，因 $-1 \le \gamma_c \le 1$、$-1 \le \gamma_d \le 1$，所以 $-1 \le \gamma_{FC} \le 1$。

1. 若 $\gamma_{FC} = 1$ 時，稱兩變數 X 與 Y 完全正相關。
2. 若 $\gamma_{FC} > 0$ 時，稱兩變數 X 與 Y 正相關。
3. 若 $\gamma_{FC} = -1$ 時，稱兩變數 X 與 Y 完全負相關。
4. 若 $\gamma_{FC} < 0$ 時，稱兩變數 X 與 Y 負相關。
5. 若 $\gamma_{FC} = 0$ 時，稱兩變數 X 與 Y 沒有關係存在或統計無關。
6. $|\gamma_{FC}|$ 值越接近 1 時，兩變數 X 與 Y 線性相關愈強。
7. $|\gamma_{FC}|$ 值越接近 0 時，兩變數 X 與 Y 線性相關愈弱。

性質 2.3：

第三類型相關係數 γ_s，是將模糊資料反模糊化值進行排序後，每個資料有了等級後而求得 γ_s，其範圍為 $-1 \le \gamma_s \le 1$。

1. 若 $\gamma_s = 1$，稱兩變數 X 與 Y 的等級順序完全一致，稱兩變數完全正相關。
2. 若 $\gamma_s > 0$，稱兩變數 X 與 Y 正相關。
3. 若 $\gamma_s = -1$，稱兩變數 X 與 Y 的等級順序完全相反，稱兩變數完全負相關。

4. 若$\gamma_s < 0$，稱兩變數 X 與Y 正相關。

5. 若$\gamma_s = 0$，稱兩變數 X 與Y 沒有關係存在或統計無關。

6. $|\gamma_s|$ 值越接近1時，兩變數 X 與Y 線性相關愈強。

7. $|\gamma_s|$ 值越接近 0 時，兩變數 X 與Y 線性相關愈弱。

性質 2.4：

第四類型相關係數γ_p，其範圍為$-1 \le \gamma_p \le 1$。

1. 若$\gamma_p = 1$時，稱兩變數 X 與Y 完全正相關。
2. 若$\gamma_p > 0$時，稱兩變數 X 與Y 正相關。
3. 若$\gamma_p = -1$時，稱兩變數 X 與Y 完全負相關。
4. 若$\gamma_p < 0$時，稱兩變數 X 與Y 負相關。
5. 若$\gamma_p = 0$時，稱兩變數 X 與Y 沒有關係存在或統計無關。
6. $|\gamma_p|$ 值越接近1時，兩變數 X 與Y 線性相關愈強。
7. $|\gamma_p|$ 值越接近 0 時，兩變數 X 與Y 線性相關愈弱。

3. 實證分析

政府近年對菸害管理很注重，然而其實未成年抽菸問題卻一直衝擊著青少年健康及校園環境菸害問題，這是很棘手且該被正視與關懷探討的課題，為了能更有有效達到教育輔導及實際協助這些青少年，由教育心理的立場出發去思考這問題，其根本應該先了解未成年抽菸者其心理層面及其周遭環境的影響因素，也就是我們要能更客觀、更多面性地去了解學生，除了訪談外，作適當性的心理問卷作，有其價值在。然而孩子們的思維及實際面向具模糊性的觀感，所以在此將心理教育研究問卷融入模糊理論，用有序性語言變數來設計問題，並對此離散型資料作分析。

本問卷研究對象是針對研究者所任教的高二有抽菸記錄的學生中隨機抽 14 位來進行調查。

欲研究探討分析，有抽菸記錄的學生，其「菸癮輕重程度」與以下幾個面向的相關程度:「與家人相處」、「與師長應對」、「與同儕相處」、「自我管理能力」、「自我健康意識」。

問卷的每個問題給予五個語言變數，並分別給予其所對應之權重。

「菸癮輕重程度」:

問題：你菸癮的程度？

權重：非常不重=1、稍微不重=2、普通=3、稍微重=4、非常重=5

「與家人相處」

問題：你喜歡與家人相處的模式嗎？

權重：非常不喜歡=1、稍微不喜歡=2、普通=3、稍微喜歡=4、非常不喜歡=5

「與師長應對」：

問題：你與師長的應對態度？

權重：非常差=1、稍微差=2、普通=3、稍微良好=4、非常良好=5

「與同儕相處」：

問題：你與同儕相處關係？

權重：非常差=1、稍微差=2、普通=3、稍微良好=4、非常良好=5

「自我管理能力」：

問題：你的自我管理能力如何？

權重：非常差=1、稍微差=2、普通=3、稍微良好=4、非常良好=5

「自我健康意識」：

問題：你擔心抽菸影響你的健康嗎？

權重：非常不擔心=1、稍微不擔心=2、普通=3、稍微擔心=4、非常不擔心=5

表 3.1 「問題：你菸瘾的程度？」的模糊資料與重心與標準差與反模糊化值與大到小排序等級

	非常不重	稍微不重	普通	稍為重	非常重	重心cx	標準差 dx	反模糊化值Xf	排序R
1號	1					1	0	1	14
2號		0.7	0.3			2.3	0.105	2.405	10
3號			1			3	0	3	7
4號			0.3	0.7		3.7	0.105	3.805	3
5號		0.2	0.8			2.8	0.08	2.88	8
6號				1		4	0	4	2
7號			0.8	0.2		3.2	0.08	3.28	5
8號		0.4	0.5	0.1		2.7	0.14	2.84	9
9號		1				2	0	2	13
10號			0.5	0.5		3.5	0.125	3.625	4
11號			0.2	0.4	0.4	4.2	0.16	4.36	1
12號		0.8	0.2			2.2	0.08	2.28	11
13號		0.5		0.5		3	0.25	3.25	6
14號		0.8	0.2			2.2	0.08	2.28	11

表 3.2 「問題：你喜歡與家人相處的模式嗎？」的模糊資料與重心與標準差與反模糊化值與大到小排序等級

	非常不喜歡	稍微不喜歡	普通	稍微喜歡	非常喜歡	重心cx	標準差 dx	反模糊化值Xf	排序R
1號			0.4	0.6		3.6	0.12	3.72	3
2號		0.8	0.2			2.2	0.08	2.28	11
3號		0.2	0.5	0.3		3.1	0.135	3.235	6
4號			0.5	0.5		3.5	0.125	3.625	4
5號	0.4	0.3	0.3			1.9	0.18	2.08	12
6號	0.2	0.2	0.6			2.4	0.18	2.58	10
7號				0.3	0.7	4.7	0.105	4.805	1
8號	0.1	0.3	0.5	0.1		2.6	0.17	2.77	9
9號		0.3	0.4	0.3		3	0.15	3.15	8
10號		0.1	0.5	0.4		3.3	0.14	3.44	5
11號	0.8	0.2				1.2	0.08	1.28	14
12號	0.8		0.2			1.4	0.16	1.56	13
13號		0.1	0.2	0.7		3.6	0.14	3.74	2
14號		0.2	0.5	0.3		3.1	0.135	3.235	6

表 3.3　「問題：你與師長的應對態度？」的模糊資料與重心與標準差與反模糊化值與大到小排序等級

	非常差	稍微差	普通	稍微良好	非常良好	重心cx	標準差 dx	反模糊化值Xf	排序R
1號			0.2	0.6	0.2	4	0.1	4.1	4
2號		0.3	0.7			2.7	0.105	2.805	13
3號		0.8	0.2			2.2	0.08	2.28	14
4號			0.5	0.5		3.5	0.125	3.625	7
5號		0.1	0.2	0.7		3.6	0.14	3.74	6
6號		0.2			0.8	4.4	0.24	4.64	2
7號				0.2	0.8	4.8	0.08	4.88	1
8號		0.3	0.4	0.2	0.1	3.1	0.185	3.285	10
9號		0.2	0.5	0.3		3.1	0.135	3.235	11
10號		0.1	0.5	0.4		3.3	0.14	3.44	9
11號			0.2	0.2	0.6	4.4	0.18	4.58	3
12號		0.3	0.5	0.2		2.9	0.135	3.035	12
13號		0.3		0.7		3.4	0.21	3.61	8
14號		0.2		0.8		3.6	0.16	3.76	5

表 3.4　「問題：你與同儕相處關係？」的模糊資料與重心與標準差與反模糊化值與大到小排序等級

	非常差	稍微差	普通	稍微良好	非常良好	重心cx	標準差 dx	反模糊化值Xf	排序R
1號			0.4	0.4	0.2	3.8	0.16	3.96	12
2號			0.5	0.5		3.5	0.125	3.625	14
3號			0.5	0.4	0.1	3.6	0.15	3.75	13
4號				0.1	0.9	4.9	0.045	4.945	2
5號			0.1	0.3	0.6	4.5	0.15	4.65	4
6號		0.2	0.2		0.6	4	0.3	4.3	8
7號				0.2	0.8	4.8	0.08	4.88	3
8號		0.1	0.3	0.2	0.4	3.9	0.23	4.13	10
9號			0.1	0.7	0.2	4.1	0.09	4.19	9
10號		0.1	0.2	0.5	0.2	3.8	0.17	3.97	11
11號					1	5	0	5	1
12號		0.1		0.3	0.6	4.4	0.18	4.58	7
13號		0.1	0.1	0.1	0.7	4.4	0.21	4.61	5
14號			0.3		0.7	4.4	0.21	4.61	5

表 3.5　「問題：你的自我管理能力如何？」的模糊資料與重心與標準差與反模糊化值與大到小排序等級

	非常差	稍微差	普通	稍微良好	非常良好	重心cx	標準差 dx	反模糊化值Xf	排序R
1號			0.4	0.4	0.2	3.8	0.16	3.96	12
2號			0.5	0.5		3.5	0.125	3.625	14
3號			0.5	0.4	0.1	3.6	0.15	3.75	13
4號				0.1	0.9	4.9	0.045	4.945	2
5號			0.1	0.3	0.6	4.5	0.15	4.65	4
6號		0.2	0.2		0.6	4	0.3	4.3	8
7號				0.2	0.8	4.8	0.08	4.88	3
8號		0.1	0.3	0.2	0.4	3.9	0.23	4.13	10
9號			0.1	0.7	0.2	4.1	0.09	4.19	9
10號		0.1	0.2	0.5	0.2	3.8	0.17	3.97	11
11號					1	5	0	5	1
12號		0.1		0.3	0.6	4.4	0.18	4.58	7
13號		0.1	0.1	0.1	0.7	4.4	0.21	4.61	5
14號			0.3		0.7	4.4	0.21	4.61	5

表 3.6　「問題：你擔心抽菸影響你的健康嗎？」的模糊資料與重心與標準差與反模糊化值與大到小排序等級

	非常擔心	稍微不擔心	普通	稍微擔心	非常擔心	重心cx	標準差 dx	反模糊化值Xf	排序R
1號	1					1	0	1	14
2號	0.4	0.6				1.6	0.12	1.72	13
3號			1			3	0	3	7
4號				0.5	0.5	4.5	0.125	4.625	2
5號	0.3	0.6	0.1			1.8	0.12	1.92	12
6號		0.8	0.2			2.2	0.08	2.28	9
7號				1		4	0	4	4
8號			0.5	0.4	0.1	3.6	0.15	3.75	6
9號		1				2	0	2	11
10號		0.1	0.9			2.9	0.045	2.945	8
11號					1	5	0	5	1
12號	0.3	0.5	0.2			1.9	0.135	2.035	10
13號	0.1	0.1			0.8	4.3	0.28	4.58	3
14號			0.2	0.8		3.8	0.08	3.88	5

第一類相關係數

將「菸癮輕重程度」與五項模糊資料求重心的相關係數γ_c，標準差的相關係數γ_d。

表 3.7　菸癮程度與五項因素之「重心的相關係數γ_c」與「標準差的相關係數γ_d」

	與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
γ_c	-0.1305	0.3227	0.4277	-0.4784	0.6419
γ_d	-0.2165	0.3762	-0.1001	0.4380	0.6752

第一類模糊相關係數為：$r_{FC} = \beta_1 r_c + \beta_2 r_d$，其中$\beta_1 + \beta_2 = 1$。

我們分別給予不同的幾組(β_1, β_2)來列出r_{FC}。

表 3.8　菸癮程度與五項因素之第一類型相關係數$r_{FC} = \beta_1 r_c + \beta_2 r_d$

	與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
(0.2,0.8)	-0.1993	0.3655	0.0054	0.2547	0.6685
(0.3,0.7)	-0.1907	0.3601	0.0582	0.1630	0.6652
(0.4,0.6)	-0.1821	0.3548	0.1110	0.0714	0.6619
(0.5,0.5)	-0.1735	0.3495	0.1638	-0.0202	0.6586
(0.6,0.4)	-0.1649	0.3441	0.2165	-0.1119	0.6552
(0.7,0.3)	-0.1563	0.3388	0.2693	-0.2035	0.6519
(0.8,0.2)	-0.1477	0.3334	0.3221	-0.2952	0.6486

第二類相關係數

將離散型模糊資料標準差的相關係數γ_d，修正成$\Delta\gamma_d = 1 - \dfrac{\ln(1+|\gamma_d|)}{|\gamma_d|}$ 。

表 3.9　　菸癮程度與五項因素之「重心的相關係數γ_c」與「標準差的相關係數γ_d」與「標準差的修正相關係數$\Delta\gamma_d$」

	與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
γ_c	-0.1305	0.3227	0.4277	-0.4784	0.6419
γ_d	-0.2165	0.3762	-0.1001	0.4380	0.6752
$\Delta\gamma_d$	0.0948	0.1512	0.0469	0.1706	0.2359

接下來觀察γ_c、γ_d的正負，來判定出模糊相關係數區間。
菸癮程度與此五項之相關係數區間。

表 3.10　　菸癮程度與五項因素之第二類型相關係數區間

與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
(-0.2253 , -0.1305)	(0.3227 , 0.4739)	(0.3877 , 0.4277)	(-0.4784 , -0.3078)	(0.6419 , 0.8778)

第三類相關係數

將離散型資料轉換成反模糊化值 X_f，再將這些反模糊化值進行排序(樣本次序等級)，接著求史比爾等級相關係數 γ_s。

表 3.11　　菸癮程度與五項因素之第三類型史比爾曼等級相關係數 γ_s

	與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
γ_s	-0.0110	0.4088	0.4901	-0.3868	0.6681

第四類相關係數

將離散型模糊資料求反模糊化值，反模糊化值為實數，用以求傳統皮爾森相關 γ_p。

表 3.12　　菸癮程度與五項因素之第四類型皮爾森相關係數 γ_p

	與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
γ_p	-0.1339	0.3379	0.4490	-0.4981	0.6707

將四個種類的相關係數整理如下，而第一類型相關係數，因研究者探討選擇視重心影響為重，故取$\beta_1 = 0.7$及$\beta_1 = 0.8$的權重所求得的相關係數區間，列於此總表內。

表 3.13　菸癮程度與五項因素之四種類型相關係數彙整表

	與家人相處	與師長應對	與同儕相處	自我管理能力	自我健康意識
第一類 $\beta_1 = 0.7$ $\beta_2 = 0.3$	-0.1563	0.3388	0.2693	-0.2035	0.6519
第一類 $\beta_1 = 0.8$ $\beta_2 = 0.2$	-0.1477	0.3334	0.3221	-0.2952	0.6486
第二類	(-0.2253 , -0.1305)	(0.3227 , 0.4739)	(0.3877 , 0.4277)	(-0.4784 , -0.3078)	(0.6419 , 0.8778)
第三類 γ_s	-0.0110	0.4088	0.4901	-0.3868	0.6681
第四類 γ_P	-0.1339	0.3379	0.4490	-0.4981	0.6707

由此資料可得知以下幾點：

1. 菸癮越重的學生，對自我健康的憂慮也重，兩者呈現正向相關，為中度相關。對提供教育輔導研究者訊息，需要專業來協助這些學生們改善。
2. 菸穩程度與家人相處模式，統計不顯著，是低度負相關。
3. 菸癮程度與對師長應對，呈現低度正相關，該分析得知，不見得菸癮重的孩子對師長應對就是如刻板印象中偏差，別因為孩子們抽菸問題，而漠視對他們良善的本質，該多給予關懷與協助。
4. 菸癮程度與同儕相處關係呈現中度正相關，分析得會抽菸的學生同儕人際關係不見得不好。
5. 菸癮程度與自我管理能力呈現中度負相關，表示有某種程度相關性，由數據分析傳達出，菸癮越重者，其自我管理能力有較差的可能性。

4. 結論

因為人類思維與主觀意識的模糊與複雜，傳統的二元邏輯並不能反映出實際情形，欲更貼切傳達出人類真實想法，許多問題的回答遍呈現模糊的回答方式，如對於薪資的要求、購買房屋可接受之價錢等，以及回答含有語言變數的問題，如心理研究問卷常有的有序性語言變數，面對這些模糊資料，此時傳統的估算方式便顯不足，研究問卷量化統計分析具重要意義與價值。

面對含有序性語言變數的離散型模糊資料，欲探討其間的相關性，本論文提供四個種類的模糊相關係數估算：第一類是由重心相關係數γ_c及標準差相關係數γ_d依適當比重相加；第二類是由重心的相關係數γ_c及標準差的修正相關係數$\Delta\gamma_d$形成的模糊相關係數區間；第三類是將資料反模糊化值X_f用來進行無母數統計中的史比爾曼等級相關係數γ_s；第四類是將資料反模糊化值X_f用來進行傳統皮爾森相關係數γ_p。第二類相關係數是模糊區間，其他三類的相關係數均為實數，將這四種相關係數對資料一起做分析，會更具合理性與客觀性。

研究者對未來的研究方向提出幾點建議：

1. 本論文的修正標準差相關係數γ_d是採用對數轉換運算成$\Delta\gamma_d$，為縮小標準差相關係數的影響力，未來研究者考慮加入調整係數，或可考慮探究標準差相關係數影響力的意義，進一步探究如何擴大及縮小其影響，以達到實際狀況的考量。
2. 後續研究上，可對非有序性的離散型模糊數問卷資料探討其相關係數的估算方法，以補足本研究未探討之遺憾。

5. 参考文獻

[1] 吳柏林(2005)。模糊統計導論－方法與應用。臺北：五南圖書公司。

[2] 吳柏林(2002)。現代統計學。臺北：前程。

[3] 吳柏林、謝名娟(2010)。現代教育與心理統計學。新北市： Airiti Press。

[4] 謝名娟、吳柏林(2012)。高中學生時間運用與學習表現關聯之研究：模糊相關的應用。教育政策論壇，15卷，第1期，157-176。

[5] 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。

[6] 林立夫 (2011)。模糊資料相關係數及在數學教育之應用。碩士論文，國立政治大學，台北市。

[7] Lowen, R. (1990) A fuzzy language interpolation theorem. *Fuzzy Sets and Systems, 34*, 33-38.

[8] Hung,T.N., Berlin, W.(2006).*Fundamentals of Statistics with Fuzzy Data*.Springer Verlag.

[9] Tseng, T. and Klein, C. (1992) A new Algorithm for fuzzy multicriteria decision making. *International Journal of Approximate Reasoning. 6*, 45-66.

[10] Ruspini, E. (1991) Approximate Reasoning: past, present, future. *Information Sciences. 57*, 297-317.

[11] Kosko, B. (1993). *Fuzzy thinking : the new science of fuzzy logic*. Hyperion, New York.

[12] Guariso, G. , Rizzoli, A., & Werthner, H. (1992) Identification of model structure via qualitive simulation. *IEEE Trans. on Systems, Man, and cybernetics.,22(5)*, 1075-1086.

[13] 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

[14] 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.

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

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

[17] Zimmermann, H. J. (1991) *Fuzzy Set Theory and Its Applications.* Boston: Kluwer Academic.

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