2、模糊統計介紹：

2.1 隸屬度函數與模糊數

過去數學理論中的集合論和人類思維的最大不同，在於數學的精準性及排它性。在傳統集合理論中，若x爲A集合中的任一元素，則$x \in A$，表示x具有某種性質；若$x \notin A$，則表示x不爲A集合中的元素，且不具有此一性質。至於其它情形則不可能同時存在。但集合論忽略了一項重要的事實，那就是x可能同時存在於A或A的餘集合中，而非全然屬於一子集合上，它是一種程度上的隸屬性關係。

傳統的統計方法透過一般的抽樣調查，往往得到的結果是單一數值的資料，或是固定尺度的選擇。面對人類想法的多元，這樣的抽樣資料，並不能完整表達人類的想法。模糊邏輯的應用在於，其主張個人的喜好程度不需要非常清晰或是有條理的，這種概念與布林邏輯理念相互對立。依照布林邏輯中A和B的比較，結果可能有三種：(1) $A>B$ (2) $A<B$ (3) $A=B$，但是人類的主觀思維遠比布林邏輯來得複雜許多，尤其是人類思維中具有許多不明確的偏好。因此，若能讓受訪者根據自己的想法，利用隸屬度函數或區間值來表達對於問題真正屬意的程度，則更可以完整地呈現受訪者真實的想法。然而，要建立一個足以表達模糊概念的隸屬度函數，並不是一件容易的事。理由是在於隸屬度函數會因個人主觀意識而有所不同，所以隸屬度函數的建立往往是最具有爭議性的。因此，目前的解決方法是根據經驗法則，或是利用已有的統計資料來輔助，利用已有的資訊，再進一步進行隸屬度函數的訂定。

隸屬度函數是模糊理論的基礎，是以 0 到 1 之間的實數來表達元素對集合的隸屬程度。假設A是X論域上的一個模糊子集，A的隸屬度函數以μ_A表示，且滿足

$$\mu_A : X \to [0,1]$$

的對應關係。Zadeh(1965)在模糊集合論中提到，若 $\mu_A(x)$ 的值接近於 1，則表示 x 隸屬於 A 的程度很高；若 $\mu_A(x)$ 的值接近於 0，則表示 x 隸屬於 A 的程度很低。因此，利用 0 ~ 1 之間的數值，可以清楚地表達元素對集合隸屬的程度。

隸屬度函數可分爲離散型與連續型兩種。離散型的隸屬度函數是直接給予有限的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（鐘形）函數）用來描述模糊集合。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。

一般來說，當資料具有不確定性與模糊性時，此時模糊數的定義就格外重要，我們定義模糊數如下：

定義 2.1 離散型模糊數

設 U 為一個論域，令 $\{L_1, L_2, ..., L_n\}$ 為論域 U 的因子集，u 為一個對應到 $[0,1]$ 間的實數函數，即 $u : U \rightarrow [0,1]$。假若佈於論域 U 之一個述句 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，則在離散的情形下，述句 X 的模糊數可表示為：

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

其中＋代表或的意思，$\frac{\mu_i(x)}{L_i}$ 表示述句 X 隸屬於因子集 L_i 的程度。

定義 2.2 連續型模糊數

設 U 為一個論域，令 $\{L_1, L_2, ..., L_n\}$ 為論域 U 的因子集，u 為一個對應到 $[0,1]$ 間的實數函數，即 $u : U \rightarrow [0,1]$。假若佈於論域 U 之一個述句 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，則在連續的情形下，述句 X 的模糊數可表示為：

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

定義 2.3 離散型模糊數的反模糊化值

設 x 是一模糊數，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 $x_f=\sum_{i=1}^{n}m_iL_i$ 為模糊數 x 的反模糊化值。

定義 2.4 連續型模糊數的反模糊化值

設 x 為一連續型模糊數，其隸屬度函數為

$$\mu_U(X)=\begin{cases} f(x) & if \quad x\in[a,b] \\ 0 & if \quad x\notin[a,b] \end{cases},$$

則 $x_f=\dfrac{\int_a^b xf(x)dx}{\int_a^b f(x)dx}$ 為模糊數 x 的反模糊化值。

2.2 模糊樣本排序

在傳統的統計分析方法中，排序是一個很重要的步驟，利用排序可以從中了解資料的分布型態、集中趨勢等。但如果樣本資料是樣糊數或是模糊區間值時，找到有效且簡易的排序方法是很重要的課題。

有關排序相關的文獻很多，Cheng(1988)利用計算距離的方法來比較模糊數大小；Liou 與 Wang (1992)則是針對具有隸屬度函數之模糊數，利用積分的方法來比較模糊數的期望值大小，以用來做模糊數之排序；Kaufmann 與 Gupta (1988)

針對具有三角形隸屬度函數的模糊數，利用其三頂點的值，提出一個排序的三法則；Chen 與 Hwang (1992)將模糊排序的方法依喜好關係、模糊平均數及分散度、模糊評分、語意表達分爲四種。

對於不同的隸屬度函數類型，適用的模糊排序方法也會有所不同。在特殊情況下我們會依欲求之目標來做調整，因此有些時候的計算是較煩瑣不易的。再者，連續型的模糊區間樣本，在排序問題上遠比離散型的模糊樣本來得複雜，而有關連續型的模糊區間樣本排序方法，也較少文獻提及。文章前面有提到，一般都將模糊區間數視爲一個具有凸隸屬度函數的模糊數，再以積分的方法找出各模糊數的模糊中心，再以模糊中心之大小決定各模糊數的排序。然而，這樣的排序會有實質上的問題存在，尤其是隸屬度函數的認定，往往是較主觀，而有的排序方法只適用於特殊型態之隸屬度函數模型，有的則是計算較爲煩瑣。

基於以上的考量，我們針對日常生活較常見的均勻分配模糊資料型態，嘗試以軟計算及配合模糊理論，針對多值邏輯模糊數及模糊區間數，提出有效且簡易的排序方法，並將此構思應用於檢定方法上。

定義 2.5 離散型模糊樣本排序

設 U 為一個論域，x_1, x_2 為離散型模糊樣本，i.e. $x_i = \sum_{j=1}^{k_i} \frac{m_j}{L_j}$。

令 $x_{1f} = \sum_{i=1}^{k_1} m_i * L_i$，$x_{2f} = \sum_{j=1}^{k_2} m_j * L_j$ 為模糊樣本 x_1, x_2 的反模糊化值。則

(1)若 $x_{1f} < x_{2f}$，則我們稱 x_1 此模糊數小於 x_2，記為 $x_1 \prec x_2$。

(2)若 $x_{1f} > x_{2f}$，則我們稱 x_1 此模糊數大於 x_2，記為 $x_1 \succ x_2$。

(3)若 $x_{1f} = x_{2f}$ 時，則我們稱 x_1 此模糊數等於 x_2，記為 $x_1 \approx x_2$。

定義 2.6 連續型模糊樣本排序

設 U 為一個論域，x_1, x_2 為連續型模糊樣本，設 $x_1 = [x_{1l}, x_{1u}]$，$x_2 = [x_{2l}, x_{2u}]$，

令 $x_{1f}=\frac{x_{1l}+x_{1u}}{2}$ ， $x_{2f}=\frac{x_{2l}+x_{2u}}{2}$ 且 $d_{1f}=\left|x_{1u}-x_{1f}\right|$ ， $d_{2f}=\left|x_{2u}-x_{2f}\right|$ ，則

(1)若 $x_{1f}<x_{2f}$ ，則我們稱 x_1 此模糊數小於 x_2 ，記為 $x_1 \prec x_2$ 。

(2)若 $x_{1f}>x_{2f}$ ，則我們稱 x_1 此模糊數大於 x_2 ，記為 $x_1 \succ x_2$ 。

(3)當 $x_{1f}=x_{2f}$ 時，則再比較 d_{1f} 及 d_{2f} ，若 $d_{1f}>d_{2f}$ ，則稱 x_1 此模糊數大於 x_2 ，記為 $x_1 \prec x_2$ ，反之，則 $x_1 \succ x_2$ 。若 $d_{1f}=d_{2f}$ ，則稱 x_1 此模糊數等於 x_2 ，記為 $x_1 \approx x_2$ 。

2.3 模糊樣本中位數

在無母數統計方法中，常以中位數代表資料的集中趨勢量數，因爲中位數比起平均數而言，較不會被極端値影響，故中位數比較穩健。

定義 2.7 離散型模糊中位數

設 U 為一論域，令 $L=\{L_1, L_2, ..., L_n\}$ 為佈於論域 U 上的 k 個語言變數，$\{x_i=\frac{m_{i1}}{L_1}+\frac{m_{i2}}{L_2}+...+\frac{m_{ik}}{L_k}, i=1,2,...,n\}$ 為一組模糊樣本($\sum_{j=1}^{k} m_{ij}=1$)。令 x_{if} 為對應於模糊樣本 x_i 之反模糊化值。

令 $x_{(i)}$ 為 x_i 排序後而得的有序樣本值，則定義離散型模糊樣本中位數為：

$$Fmedian(X)=\begin{cases} x_{(\frac{n}{2})+1} & \text{若}\mathbf{n}\text{為奇數} \\ (x_{(\frac{n}{2})}+x_{(\frac{n}{2})+1})/2 & \text{若}\mathbf{n}\text{為偶數} \end{cases}$$

定義 2.8 連續型模糊中位數

設U為一論域，令$\{x_i=[a_i,b_i], i=1,2,...,n\}$為自恥中抽出的一組糊區間樣本。令$m_i$為$x_i$之中點，$l_i$為$x_i$的長度，則模糊樣本中位數為以$median\{m_j\}$為中心，以$l_i$的中位數為直徑的區間。

$$Fmedian=(m,r),\ m=median\{m_j\},\ r=\frac{median\{l_i\}}{2}$$ 。