第三章 研究方法

第一節 簡介 Fuzzy Model

Fuzzy 理論發展至今有三十年左右，相關的研究成果迅速成長，應用觸角也伸向科技領域的許多方面，發展速度超過許多應用數學的分支。目前，這門學科在人工智慧、自動控制、圖像識別、醫療診斷、決策支援、管理科學、氣象預報、環境評估等各種領域應用有豐碩的成果。但較廣為人知的應用還是在自動控制的領域。在財務金融方面，近年來的研究普遍認為金融市場序列呈現非線性，Fuzzy 模型也是一種非線性模型，以 Fuzzy 為基礎的預測方法有些成功的案例，本論文嘗試用 Takagi 和 Sugeno 在 1985 年提出的 Fuzzy 模型來預測匯率，研究模糊理論的特性是否能準確的預測匯率波動，並和隨機漫步模型進行預測能力上的比較。接下來先簡單介紹 Fuzzy 模型。

〔一〕模糊理論起源

在我們的生活中，到處存在著模糊的現象，比如說：”這杯水很燙”、”那個女生很瘦”這樣的敘述，到底水溫幾度算燙呢？女生幾公斤算瘦呢？這種敘述隱含了人類思維概念、語意表達、感覺判斷等都具有模

糊的特質，所謂模糊，就是不精確、模稜兩可、不確定、曖昧性的意思。雖然每個人對於”燙”或”瘦”的定義都不一樣，而且受到個人成長經驗、主觀判斷、當時心理狀態…等的影響，我們並不能精確的估測對方所說的燙是幾度，瘦是幾公斤，但是人與人之間的對於此種經驗的差距並不大，而能達到溝通的效果。

模糊理論是由美國加州大學柏克萊分校 Zadeh〔1965〕教授提出的一種定量表達工具，以數學的方法來解決生活中普遍存在的模糊現象，尤其在表現人類語言特有的模糊現象方面有不錯的成果。在傳統的控制中，研究者必須建立精確嚴謹的數學模型來達成控制，但實際上，有很多情況是無法明確的決定模型的數值，也很難建立精確的數學模型。Zadeh 教授嘗試以人類的思維方式去簡化問題的複雜度，利用歸屬度函數以 0 到 1 之間的數值來描述一個元素屬於某一個集合的程度，這 0 到 1 之間的值為歸屬度中，若歸屬度為 1，表示這個元素完全屬於這個集合，歸屬度為 0，是完全不屬於這個集合，介於兩者之間，則是灰色地帶，也就具有模糊的概念。藉由模糊的思考來簡化問題，並不要求百分之百的精確，達到與傳統控制方法相同的目的。

模糊集合是模糊理論的基礎，因此下面將會探討模糊集合的特質、集

合運算、歸屬函數(Membership function)等。為了能更了解模糊集合的精神，在說明模糊集合的概念及演算時，先介紹傳統集合的概念及演算。

〔二〕傳統集合

集合是由一些具有某種共同屬性的元素匯集起來的組織，集合中的每一個成員稱為此集合的元素，這裡的集合是以表達明確事物為主，習慣上稱做「傳統集合」、「明確集合」或「Crisp 集合」。所謂明確就是可以用“是”與“否”做出清楚的判斷。

考慮具體的問題時，通常會把考慮的對象限制在一定範圍內，這個範圍就稱為論域或全集合。例如討論“奇數”與“偶數”的概念，可以拿所有的自然數為論域，論域可以寫成，U={自然數論域}={1、2、3、4、…}，其中奇數的集合寫成 A={奇數}={1、3、5…}，偶數集合寫成 B={偶數}={2、4、6…}。其中也可以討論 A 集合與 B 集合的運算問題，包括交集〔∩〕、聯集〔∪〕、差集、補集…等。

接著是集合裡面很重要的概念，即特徵函數(Characteristic function)。對 U 的子集合 A 而言，我們定義 $S_A(x)$ 為 A 集合的特徵函數

$$S_A(x) = \begin{cases} 1 & if \quad x \in A \\ 0 & if \quad x \notin A \end{cases}$$

如上述公式所示，若 χ 為 A 的元素，則其特徵函數 $S_A(x)$ 為 1；反之若 χ 不是 A 的元素，則其特徵函數 $S_A(x)$ 為 0。利用特徵函數可以表現集合論中"非此即彼"的二值邏輯特性，其值域可寫成 $S_A(x)$={1，0}。

〔三〕模糊集合

雖然在日常生活中很自然的使用這些敘述，像“這杯水很燙”…等，但我們很難具體的定義所謂的“燙”是幾度？若攝氏 60 度到 80 度的水稱做“燙”，攝氏 59 度的水就稱做“不燙”，這種一分為二的方式在多數的情況下，是不合理而且不適當的。若區分“燙”與“不燙”的界限能夠緩和一些，則這種不自然的情況就會消失。或許可以用“燙的程度”來考量，若攝氏 80 度的水，燙的程度是 0.9，而 60 度的水，燙的程度是 0.6，隨著不同的溫度，燙的程度也會跟著變化，這個問題也能獲得根本上的解決。這種重新擴張定義的集合，由 Zadeh 教授提出，稱之為模糊集合。

Zadeh 教授把傳統集合論特徵函數從非 0 即 1 的二值選擇，推廣為可從 0 到 1 之間的任何值作選擇，這種新的特徵函數，稱為歸屬函數。歸屬函數可以描述模糊集合的性質，是模糊理論的最基本概念，透過歸屬函

數，才能對模糊集合進行量化，也才能利用精確的數學方法分析和處理模糊性資訊。一般常用的歸屬度函數有三角形、梯形、高斯函數…等。

若歸屬函數 $\mu_A(x)$ 為離散型，此時論域 U 裡的元素有“有限個”，則模糊集合 A 可表達如下，其中“ / ”為分隔的意思：

$$A = \mu_A(x_1)/x_1 + \mu_B(x_2)/x_2 + ... + \mu_A(x_n)/x_n$$

若歸屬函數 $\mu_A(x)$ 為連續型，此時論域 U 裡的元素有“無限多個”，模糊集合 A 可表達如下：

$$A = \int_U \mu_A(x)/x$$

〔四〕模糊集合的基本運算

模糊集合的運算主要以聯集、交集、補集最為常用，概念也和傳統集合一樣，任何演算的定義基本上都可以用這三種運算組合來表達。先假設 C 和 D 為兩個模糊集合，μ_C 和 μ_D 為集歸屬函數，以下簡單介紹常用的運算方法：

$\boxed{\text{模糊聯集}(C \cup D)}$ “$\vee$”為模糊運算子，這裡取 max 運算。

$$\mu_{C\cup D}(x) = \mu_C(x) \vee \mu_D(x) = \max\{\mu_C(x), \mu_D(x)\}$$

$\boxed{\text{模糊交集}(C \cap D)}$，“$\wedge$”為模糊運算子，這裡取 min 運算。

$$\mu_{C \cap D}(x) = \mu_C(x) \wedge \mu_D(x) = \min\{\mu_C(x), \mu_D(x)\}$$

$\boxed{\text{模糊補集}(\overline{A})}$

$$\mu_{\overline{A}}(x) = 1 - \mu_A(x)$$

此外，在明確集合中所運用的定律，也同樣適用於模糊集合，包括：恆等律、交換律、結合律、分配律、狄摩根定律…等。

〔五〕模糊推論〔推論引擎〕

模糊邏輯的推論過程與二元邏輯的推論過程類似，其藉由模糊邏輯的運算，模擬人類思考判斷的方式。若我們只考慮兩個輸入、一個輸出的系統，並假設只有兩個模糊規則，則模糊規則可以表示成：

If x_1 is A_1 and x_2 is B_1 Then y is C_1

If x_1 is A_2 and x_2 is B_2 Then y is C_2

為了方便討論，我們定義：在上式中，If後面的 χ is A 為該模糊規則的前件部，在 Then 後面的 y is B稱為後件部。

以下介紹一種模糊推論方式：Min-Min-Max，由這種方式，能讓我們更了解模糊推論的過程。

Min-Min-Max推論方式，可以分成三個步驟，並把概念繪製成〔圖2〕：

〔1〕計算這兩個規則前件部的歸屬度w_i

$$w_i = \mu_{A_i}(x_1) \wedge \mu_{B_i}(x_2) = \min\{\mu_{A_i}(x_1), \mu_{B_i}(x_2)\} \qquad \text{i=1,2}$$

〔2〕計算這兩個規則後見部的歸屬度$C(y)$

$$C_i(y) = \min\{w_i, C_i(y)\} \qquad \text{i=1,2}$$

〔3〕整合這兩個規則，推論出最後的結果$C^*(y)$

$$C^*(y) = \max\{C_1(y), C_2(y)\}$$

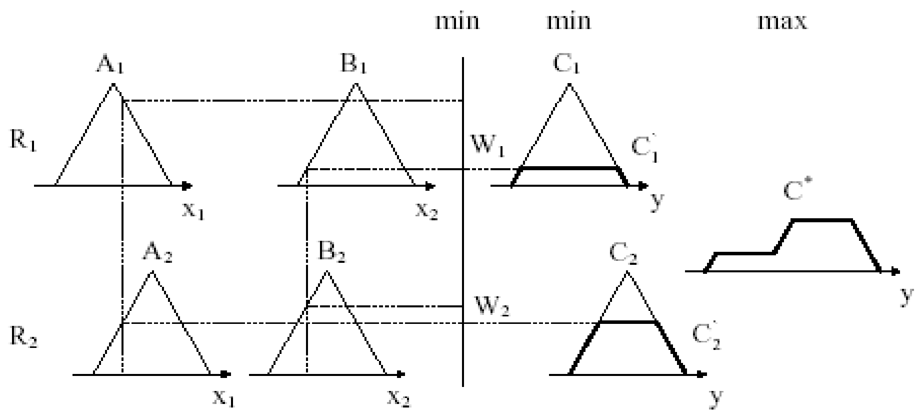

圖 2：Min-Min-Max 的模糊推論

〔六〕模糊系統架構

模糊系統被用於各種領域，故有時候又被稱為模糊專家系統、模糊邏輯控制或是模糊模型…等。模糊系統是以言語化的模糊規則為主體，所以為了將輸入值與模糊規則相結合，必須先將輸入值模糊化，經過推論引擎的作用，產生的結果也需去模糊化，才能讓結果轉換成外界可接受的明確數值。整個模糊系統架構基本上包含了五個主要的部分：

1. 模糊化機構

2. 資料庫

3. 規則庫

4. 推論引擎

5. 去模糊化機構

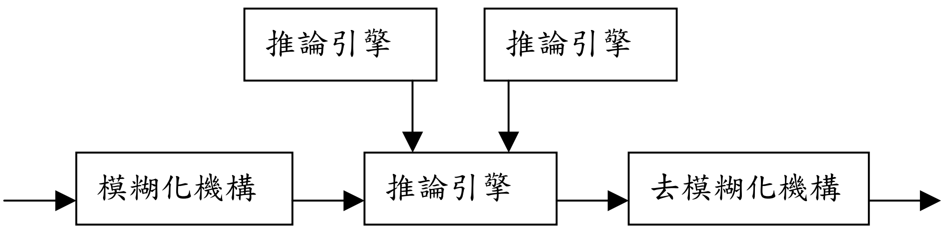

圖 3：去模糊化機構

第二節 Takagi and Sugeno's Fuzzy Model

〔一〕簡介 Takagi and Sugeno's Fuzzy Model

本研究所採用的 Fuzzy 模型是 Takagi 和 Sugeno 兩位學者在 1985 年所提出的。Takagi-Sugeno Fuzzy 模型是以片段的線性模式為基礎，將非線性系統分解成許多更簡單的線性子問題，如此能巧妙地解決匯率的複雜性與非線性，這種局部解構法接近人類處理問題方式，觀念也是簡潔和直覺，融合了線性與非線性的優點。表達如下：

$$R^i : If \quad x_1 \quad is \quad A_1^i \quad and \quad x_2 \quad is \quad A_2^i, \ldots, and \quad x_m \quad is \quad A_m^i$$

$$then \quad y^i = a_0^i + a_1^i x_1 + \ldots + a_m^i x_m$$

$$\hat{y} = \frac{\sum_{i=1}^{c} w^i y^i}{\sum_{i=1}^{c} w^i}, \quad w^i = MIN_{j=1}^{m} A_j^i(x_j)$$

其中 $R^i(i = 1,2,\ldots,c)$ 表示第 i 條規則，$x_i(j - 1,2,\ldots,m)$ 表示第 j 個輸入，y^i 表示第 i 條模糊規則的輸出，$A_1^i, A_2^i, \ldots, A_m^i$ 是以不同的歸屬度函數代表的模糊變數，一般常見的有鐘型、梯形和三角形。

由以上的介紹我們可以知道，Takagi-Sugeno Fuzzy 模型是數個線性系統的組合來近似一個非線性的系統，利用模糊化來解構所有的輸入，經過

模糊推論，去模糊化後產生數條線性的方程式，來代表每一組輸入與輸出的關係。Takagi-Sugeno Fuzzy 模型中，各個規則的後件部可以用線性系統理論的標準工具加以分析。由以下範例我們可以更清楚的了解這個觀念：

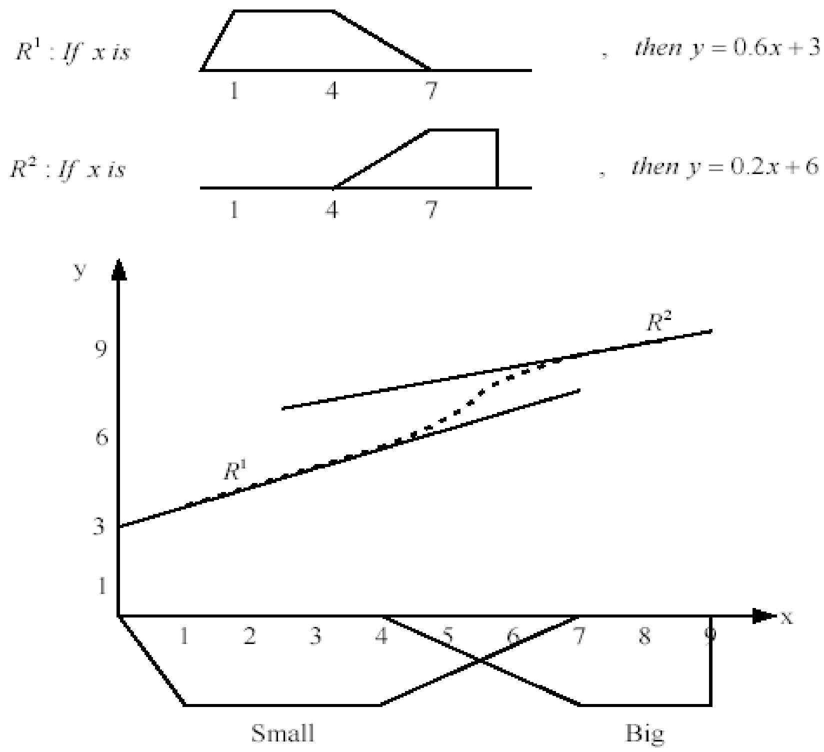

圖 4：Takagi-Sugeno Fuzzy 模型的範例

本論文用來預測匯率的 Takagi-Sugeno Fuzzy 模型，係採用朱修明先生於 2001 年 6 月所發表的模型為主體，由郭子文學姊修改而成。模型以 MATLAB 程式語言寫成。

〔二〕Takagi-Sugeno Fuzzy 模型的建模流程

本模型建模流程〔朱修明，1995〕是先將未知系統的輸出資料做模糊分類，使用 FCM 演算法計算出每一類的中心向量，再利用 SC 準則來決定出合適的分類群數。這個分類群數便用來當作模糊規則庫中的規則數目。

此規則數目將未知系統分成數個線性系統，使用 FCRM 演算法找出輸入/輸出資料對與這些線性系統的模糊關係，以歸屬值矩陣表示。利用前件部參數建構法和後件部參數建構法，初步求出模糊規則中前件部和後件部的係數值。

之後再用最陡坡降法同時對前件部和後件部係數做最佳化調整，直到找出一組係數值能使得該模糊系統達到設定條件的要求。這整個流程圖可由〔圖五〕看出。Takagi-Sugeno Fuzzy 模型的詳細研究方法可以參考朱修明〔1995〕的論文。

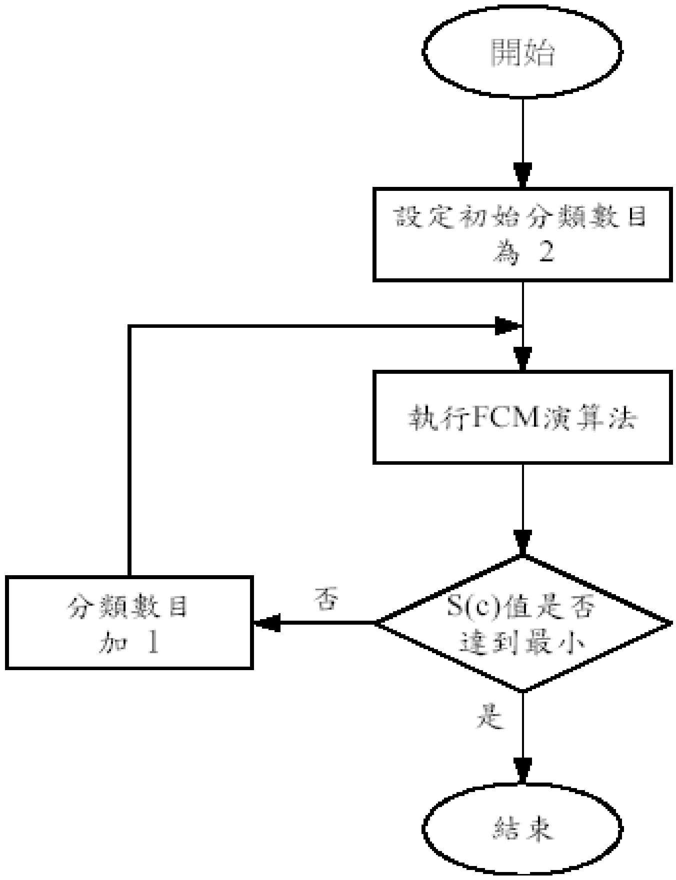

圖 5：建立模糊模型之流程

第三節 決策樹模型

決策樹學習法可說是歸納是學習法裡最簡單的一支，且實際操作上並不困難，所以應用的層面非常廣泛。決策樹主要的功能在於，藉由分類以知的事例來建立一個樹狀結構，並從中歸納出事例裡的某些規律，而產生出來的決策樹也能用來做樣本外預測。

在決策樹演算結果的樹狀圖裡，每個內部節點〔Internal Node〕代表對某個屬性的測試〔Test〕，其下的每個分支〔Branch〕代表此屬性的一個可能值，或多個可能值的集合，最後每個樹葉節點〔Leaf Node〕對應的是一個目標類別〔Target Class〕。決策樹一般都是自上而下來生成的，選擇分割的方法有好幾種，但目的都是一致的，就是對目標類嘗試進行最佳的分割，從根到葉子節點都有一條路徑，這樣一條路徑就是一條規則。決策樹可以是二叉的，也可以是多叉的，

當資料不完整、過於稀疏或是含有雜訊時，所建構的決策樹通常會有過度配適的情形，所產生的決策樹會過於複雜。產生過度配適的原因有兩種：一是樣本本身的屬性太多，決策樹學習法容易選到和種類不相關的屬性，也就是說，就是自由度太大。二是偏誤〔Bias〕，不同演算法在尋找

測試屬性的時候，都有自己的偏好，所以很有可能會找到演算法所偏好但不是真正與種類相關的屬性，因此決策樹完成後還會做適當的修剪〔Pruning〕。

進行決策樹修剪時有兩個常見的方法：一、預先修剪〔Pre-Pruning〕，二、事後修剪〔Post-Pruning〕。

預先修剪是以提早停止決策樹的生長來達到修剪的目的，當樹停止生長時，末端節點即成為樹的樹葉。樹葉的標籤〔Label〕，為該節點訓練集合中佔有比例最大的類別，停止決策樹生長是在決策樹建立前，事先設定好一個門檻〔Threshold〕，當分支節點滿足門檻值的設定，就停止該分支繼續成長。反之，事後修剪是先建立一個完整的樹，再將其分支移除的做法，移除分支的依據是計算該分支的錯誤率〔Error Rate〕，最末端未被移除的分支節點就變成樹葉。目前以事後修剪較受到歡迎，因為許多學者認為，預先修剪法所需設定的門檻會過於主觀。

決策樹經過生長和修剪後，就可以從中讀取事例中隱藏的規律，每個由樹根〔Root〕開始，到某個樹葉節點〔Leaf Node〕的路徑〔Path〕，都能代表一條分類規則。值得注意的是，每條決策樹歸納出來的規則，只能

參考到一個事例，所以決策樹的表達能力與零階邏輯相同，因此決策樹的缺點之一，就是無法描述在一個測試裡，同時參考兩個事例的情況。

本篇論文用 See5/Cubist 的套裝軟體來對匯率進行策樹的分析。

第四節 隨機漫步模型

Meese 和 Rogoff〈1983〉指出隨機漫步的模型比其他匯率決定模型更為準確，這個研究引起很大的衝擊。許多經濟學者開始探討匯率的波動是否真的隨機不可預測，許多模型的預測結果也拿來跟隨機漫步模型相比。

為了衡量 Takagi-Sugeno Fuzzy 模型和 Cubist 決策樹模型的表現，本論文採用隨機漫步模型〔Random Walk〕來和這兩個模型做預測能力的比較，並採用均方差〔MSE〕為比較的標準。在這裡所用的隨機漫步模型是：對第 n 期匯率的預測值為第 n-1 期的匯率值，以數學式來表示：

$$E(x_n(\Omega_{n-1})) = x_{n-1}, \quad \forall n$$

第五節 過度配適問題

所謂過度配適，就是用模型解釋過去的資料時，準確度很高，但當要用此模型預測未來時，則預測的效果很差。本論文中，在探討過度配適的問題時，乃沿用一般常用的“樣本內”及“樣本外”的配適比較。也就是當我們要說明，Takagi-Sugeno Fuzzy 模型預測的精確度，並非是過度配適下所產生的精確度，樣本內的誤差若很小，預測未來時，樣本外的誤差也應該很小才合理，這樣就沒有過度配適的問題。

而 Cubist 決策樹模型的過度配是問題在本章第三節有提到，即看產生出來的規則是否過於複雜，與誤差觀察的方式相比，其實是一體的兩面。

第六節 Diebold-Mariano 檢定

本論文對於 Takagi-Sugeno Fuzzy 模型和 Cubist 決策樹模型預測出來的結果與隨機漫步的預測結果做 Diebold-Mariamo 檢定，觀察此兩種預測方法的準確度是否優於、劣於或等於隨機漫步。以下簡單介紹 Diebold-Mariamo 檢定的統計量：

Diebold-Mariamo 檢定統計量不需要假設損失函數〔loss function〕的形態，也不要求預測誤差是常態分配，允許誤差具有序列相關、異質變異，所以 Dieblod-Mariamo 檢定法適合做為不同預測能力經濟上顯著性的比較。大部分經濟決策考量下的模型判斷，模型的損失函數通常不是二次式，有些甚至不能用二次式表示，所以 Diebold-Mariamo 檢定法不要求損失函數的形式及誤差的限制，使這個檢定方法具有相當強的檢定力，因此受到廣泛的應用。

假設有一組時間序列資料$\{y_{it}\}_{t=1}^{T}$，比較兩種預測方法的優劣。兩種預測值分別是$\{\hat{y}_{it}\}_{t=1}^{T}$和$\{\hat{y}_{jt}\}_{t=1}^{T}$，令$\{e_{it}\}_{t=1}^{T}$和$\{e_{jt}\}_{t=1}^{T}$為這兩組預測值的預測誤差，$g(e_{it})$和$g(e_{jt})$是損失函數。若虛無假設是$\{\hat{y}_{jt}\}_{t=1}^{T}$預測的準確度和$\{\hat{y}_{it}\}_{t=1}^{T}$一樣好，則假設檢定寫成：

$$H_0 : E[g(e_{it})] = E[g(e_{jt})]$$

現在定義 $d_t \equiv [g(e_{it}) - g(e_{jt})]$，則假設檢定可以改寫為：

$$H_0 : E[d_t] = 0$$

$$H_1 : E[d_t] > 0$$

$$H_2 : E[d_t] < 0$$

若接受 H_0，表示 $\{\hat{y}_{jt}\}_{t=1}^{T}$ 和 $\{\hat{y}_{it}\}_{t=1}^{T}$ 一樣準確；接受 H_1，表示 $\{\hat{y}_{jt}\}_{t=1}^{T}$ 優於 $\{\hat{y}_{it}\}_{t=1}^{T}$；若接受 H_2，則表示 $\{\hat{y}_{it}\}_{t=1}^{T}$ 優於 $\{\hat{y}_{jt}\}_{t=1}^{T}$。

如果 $\{d_t\}_{t=1}^{T}$ 是 covariance stationary 和 short memory，則可用下列分配：

$$\sqrt{T}(\bar{d} - u) \xrightarrow{d} N(0, 2\pi f_d(0))$$

其中，

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T} d_t$$

$$f_d(0) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} \gamma_d(\tau)$$

$$\gamma_d(\tau) = E[(d_t - u)(d_{t-\tau} - u)]$$

$$u = E[d_t]$$

檢定統計量為：

$$S_1 = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{T}}}$$

在 97.5%的信心水準下，$\alpha = 0.025$，若 $S_1 > Z_{\alpha/2} = 2.24$，則拒絕 H_0，接受 H_1，即 $\{\hat{y}_{jt}\}_{t=1}^{T}$ 優於 $\{\hat{y}_{it}\}_{t=1}^{T}$；若 $S_1 < Z_{\alpha/2} = -2.24$，則拒絕 H_0，接受 H_2，即 $\{\hat{y}_{it}\}_{t=1}^{T}$ 優於 $\{\hat{y}_{jt}\}_{t=1}^{T}$。

用這個方法可以檢定出 Takagi-Sugeno Fuzzy 模型及 Cubist 決策樹模型的預測準確度和隨機漫步的準確度孰優孰劣。

圖 6：檢定鐘型圖