

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

模糊時間序列與區間預測方法探討
-以台灣加權股價指數為例

A study on the Fuzzy time series and
interval forecasting methods
-with case study on the Taiwan
Capitalization Weighted Stock Index

碩專班學生：李栢昌 撰

指導教授：吳柏林 博士

中華民國 一〇一年 八月 二十七日

謝辭

時間一轉眼就過了，終於要從政治大學應用數學所畢業了，當初對於自己教學上的缺陷與不足，有了念在職專班的念頭，在同事的鼓勵之下，我報考了政治大學應用數學所，三年的學習填滿我所有的暑假時光，雖然辛苦也讓我體會到學生的想法，這經驗不但讓我有教學相長的體悟，也讓我對教學工作有更一層的思考。如今在職進修終於完成，從這裡我學習到了不少，從數學教育到測驗評量，從數學繪圖軟體到數學專業科目，都有所精進。感謝這三個暑假以來同窗的同學以及辛勤指導我們的老師群們。特別要感謝我的指導老師，吳教授柏林老師，謝謝您總是關心學生的生活和學習的狀況。

在今年二月多的時候，因為老師一路來的提攜與指導，我才能在模糊統計有些微的認知，也讓我明白如何將模糊區間用在財務管理的領域，讓我的論文更有實用性。最後要感謝我的父母親，讓我無後顧之憂的念書，雖然您們很少問起我進修的事情，一路走來感謝您們給我的自由，希望往後的教學生涯，我能夠給學生帶來新的視野、給家人帶來新的感動，最後，希望我能在未來教學中貢獻所學。

李栢昌 謹誌于
國立政治大學應用數學所
中華民國一百零一年八月

模糊時間序列與區間預測方法探討 -以台灣加權股價指數為例

摘要

台灣加權股價指數(TAIEX)，可以說是台灣最重要的經濟指數之一。在預測的方法中，時間序列分析一直都是熱門的課題，也是最常被使用來研究股價預測的方法。近年來，模糊理論在生醫、財務、社會、電機等各領域都有不錯的應用與發展。本研究欲透過模糊區間的預測，主要是以時間序列預測台灣加權股價指數，來作為模糊區間精確度的探討，並針對區間時間序列進行模式的建構診斷和預測。最後我們將以 2012 年第一季(Q1)，每日交易股價指數的最高價與最低價作為實際研究的例子，同時也比較不同預測方法所得的結果。結果顯示模糊區間預測提供不同於傳統預測方法所得的資訊，希望能提供投資者另一種投資的參考。

關鍵字：台灣加權股價指數(TAIEX)、模糊理論、模糊區間、區間預測

Abstract

Taiwan Weighted Stock Index (TAIEX) is one of Taiwan's most important economic indicators. Among the forecasting methods of time series analysis is always a hot issue on the forecasting methods and is also the most commonly used to make the stock price predictions. In recent years , fuzzy theory makes a great of application and development in various fields , such as , biomedical , financial and social ...etc.. For this study, through the fuzzy interval forecasting is mainly based on time series forecasting TAIEX as fuzzy interval accuracy of the construction of diagnosis and prediction of the mode and interval time series.

Finally, we will take the daily highest / lowest stock index prices data in the first quarter of 2012 (Q1) for actual research example , and will compare different forecasting methods of the results. The results show that the fuzzy interval forecasting differentiated from the traditional one on the basis of these information. We hope to offer investors an alternative investment advice.

Keyword : *Taiwan Capitalization Weighted Stock Index (TAIEX) 、 Fuzzy theory 、 Fuzzy interval 、 Interval forecasting.*

目錄

1. 前言	1
2. 研究方法	3
2.1 時間序列預測方法	3
2.1.1 簡單移動平均法 (Simple Moving Average Method)	3
2.1.2 單一指數平滑法 (Single Exponential Smoothing Method)	3
2.1.3 自回歸移動平均法 (Autoregressive Integrated Moving Average Method)	4
2.2 區間型模糊數及其運算	5
定義 2.1 區間模糊數	5
定義 2.2 區間模糊數加法	6
定義 2.3 區間模糊數減法	6
定義 2.4 反模糊值	7
2.3 模糊區間距離	7
定義 2.5 模糊區間距離 (distance of interval)	7
定義 2.6 區間平均誤差 (Interval Mean Error)	8
2.4 ARIMA 模型配適度檢測	8
定義 2.7 赤池信息量準則	8
3. 實證分析	9
3.1 資料來源	9
3.2 預測方法探討	19
3.3 預測效率分析	20
4. 結論	22
5. 參考文獻	24

1. 前言

近年來由於國際經濟變化加劇，美國的金融風暴、歐洲的債信風險，新興亞洲等國家的經濟崛起，身處台灣的投資大眾在面對瞬息萬變的國際投資環境之外，還要考慮台灣特殊的產業環境與選舉行情等因素，因而各種股市的投資書如雨後春筍般的出現，台灣的投資者依然在投資過程中，對於不確定的股價指數莫衷一是。本研究希望提供不同於傳統對台灣加權股價指數的預測，做為投資台灣股市的一個參考，同時也給予實際的例子作為驗證，期許這些資料能夠給予投資者有用而準確的資訊。

模糊時間序列的演進方面，Song 與 Chissom (1993) 發展了模糊時間數列的定義，以及用模糊關係式發展模糊時間數列的模型。Kunhuang (2001)曾經研究區間長度對模糊時間數列預測結果的影響，Chen (1996)也提出了一個簡單的模糊時間數列預測模式，來預測學生註冊人數，並且得到良好的預測效果。Wu and Hung (2006) 提出模糊認定法則，以作為 ARCH 模式族與 Bilinear 模式族的決策判定標準。Chen (2002)發表高階的模糊時間數列 (high-order fuzzy time series) 來預測阿拉巴馬大學的學生註冊人數。Chen 等人 (2006)對 TAIEX 的預測提出了加權模糊時間數列模型。

在國際上，Chatfield (1993) 是區間預測計算的先驅之一，Huang (2001)研究區間長度對模糊時間數列預測結果的影響，Tseng and Tzeng (2002) 對模糊的季節性 ARIMA 預測模型作估計，而 Chang (2007) 也為區間資料的平均數與變異數設定的假設檢定法則，Vladik, Hung and Wu (2006) 提出網路即時的統計計算方法。更多學者，依照不同的研究課題，發展更適合的區間預測模型，以及對區間效率評估的研究。

綜觀台灣加權股價指數由台灣證券交易所編制以來，世界經濟有越來越緊密的趨勢，台灣股市與國際緊密連結在一起，台灣證券交易所 2012 年第一季的統計資料顯示，2012 年第一季交易日中股價指數最低價為 7005.39 點，最高價為 8170.72 點，國內的政策如復徵證所稅等因素，使得股價受到政策面的影響，因此

決定以第一季的股價作為實證分析的標的，避免預測結果受到政策面的影響。

本文主要分為四個章節，第一章探討相關方法之文獻，以獲取發展理論的依據，並且說明比較區間預測效率差異的目的。第二章說明預測方法分析與模糊區間理論等研究方法，並且建構新的區間距離方式。第三章為實證分析，透過研究2012年第一季台灣加權股價指數（TAIEX）的資料來驗證理論。最後本文的結論將放在第四章。



2. 研究方法

本研究利用模糊區間觀念分別對時間序列進行最低價與最高價區間的探討，並對模糊區間距離與模糊預測的準確度作探討。

2.1 時間序列預測方法

2.1.1 簡單移動平均法 (Simple Moving Average Method)

移動平均 (Moving Average)，是一種分析時間序列數據的常見工具。最常見於股價的日移動平均線或月移動平均線，其優點是不易受到短時間的波動影響預測值，缺點是因平均值需要較多的期數，因此資料的反應較為緩慢，不易把握股價趨勢的高峰與低谷，因實務計算方便而經常被使用。

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \dots + A_{t-n}}{n}$$

A_{t-1} ：代表前 1 期實際資料。

A_{t-n} ：代表前 n 期實際資料。

F_t ：對第 t 期的預測值。

n：期數。

2.1.2 單一指數平滑法 (Single Exponential Smoothing Method)

單一指數平滑法是透過調整係數 α 、前一期的實際資料與前一期的預測資料來做預測，其優點是調整係數重視接近的時點資料，預測結果不易受遠期資料干擾，缺點是，預測資料因為權重的關係，容易受到近時點資料的波動影響。實際操作上，當時間序列資料波動大時，則取較小的 α 值，當時間序列資料波動小時，則取較大的 α 值。

$$F_t = F_{t-1} + \alpha (A_{t-1} - F_{t-1}) = (1 - \alpha)F_{t-1} + \alpha A_{t-1}$$

A_{t-1} ：代表前 1 期實際資料。

F_{t-1} ：代表前 1 期預測資料。

α ：調整係數 ($0 < \alpha < 1$)。

2.1.3 自回歸移動平均法 (Autoregressive Integrated Moving Average Method)

自回歸移動平均法(ARIMA)是由 Box and Jenkins 在 1970 年所提出的預測模式，一個完整的 ARIMA 模型應該包含三個部份來預測時間數列：

- (1) 自我迴歸項 (AR-autogressive terms)；
- (2) 差分處理項 (I-integrated)；
- (3) 移動平均項 (MA-moving average terms)。

經過這三個部份的結合，即可建立一個完整的 ARIMA 預測模式，若一時間數列 $\{X_t\}$ 滿足 ARIMA(p,d,q) 模式，則對任意 t 可寫成

$$\varphi_p(B)(1-B)^d X_t = \theta_q(B)\epsilon_t$$

B : 退後因子 $B X_t = X_{t-1}$ 。

X_t : 是一個時間序列隨機變數。

ϵ_t : 誤差項白噪音(White noise)。

$\varphi_p(B) = (1 - \varphi_1 B^1 - \varphi_2 B^2 - \dots - \varphi_p B^p)$ 稱為 AR(p) 模型。

$\theta_q(B) = (1 - \theta_1 B^1 - \theta_2 B^2 - \dots - \theta_q B^q)$ 稱為 MA(q) 模型。

$(1 - B)^d = \nabla^d$ 稱為 d 階差分。

※ 時間序列預測方法的比較

時間序列的預測方法有很多種，以下我們整理一下由吳柏林老師在 1995 年所提出的預測方法比較，作為預測方法的參考。

表 2.1 各種預測方法成本與準確度比較表

	理論程度	作業時間	作業成本	資料特性	所需資料筆數	預測時間範圍	預測準確度
成長模式	1	1	2	T	20	中期	1
指數平滑法	2	2	1	T	10	短期	2
多變量回歸	3	3	3	T、S	30	中期	3
分解法	4	4	5	A、T、S	30(3S)	短期	4
ARIMA	5	5	4	A、T、S	30(4S)	短期	5
狀態空間	6	6	6	A、T、S	30(4S)	短期	8
計量經濟	7 - ∞	9	9 - ∞	T、S	6	長期	8
非線性	8	8	7	IRREGULAR	40	中期	6
神經網路	9 - ∞	7	8 - ∞	FREE	60(4S)	長期	8

註：

「1」 表最小、最短或最低；

「2」 表最大、最長或最高

A : Stationary

T : Trend

S : Seasonal

資料來源：整理自吳柏林 (1995)

2.2 區間型模糊數及其運算

定義 2.1 區間模糊數 (吳伯林, 2005)

若閉區間 $[a, b]$ 為一區間模糊數， a, b 均為實數， $c = \frac{a+b}{2}$ ， $r = \frac{b-a}{2}$ 分

別表示區間的中心點與半徑，亦可表示為 $[a, b] = (c, r)$ ，其中 $b - a = 2r$ 。

例 2.1 區間模糊數

大學畢業生的起薪通常是一個範圍，無法僅用一個數字來表示，如果大學畢業生起薪 22000 元到 28000 元，表示成 $[22000, 28000]$ ，我們可以視為一個區間模糊數，同時我們可以改寫成 $[22000, 28000]$ 或表示為 $(25000, 3000)$ 。

定義 2.2 區間模糊數加法 (吳伯林, 2005)

$$\begin{aligned} X_1 + X_2 &= [a_1, b_1] + [a_2, b_2] = [a_1 + a_2, b_1 + b_2] \\ &= (c_1, r_1) + (c_2, r_2) = (c_1 + c_2, r_1 + r_2) \end{aligned}$$

例 2.2 $X_1 = [3, 5] = (4, 1)$, $X_2 = [4, 6] = (5, 1)$

$$\begin{aligned} X_1 + X_2 &= [3, 5] + [4, 6] = [3 + 4, 5 + 6] = [7, 11] \\ &= (4, 1) + (5, 1) = (4 + 5, 1 + 1) = (9, 2) \end{aligned}$$

定義 2.3 區間模糊數減法 (吳伯林, 2005)

$$\begin{aligned} X_1 - X_2 &= [a_1, b_1] - [a_2, b_2] = [a_1 - b_2, b_1 - a_2] \\ &= (c_1, r_1) - (c_2, r_2) = (c_1 - c_2, r_1 + r_2) \end{aligned}$$

例 2.3 $X_1 = [3, 5] = (4, 1)$, $X_2 = [4, 6] = (5, 1)$

$$\begin{aligned} X_1 - X_2 &= [3, 5] - [4, 6] = [3 - 6, 5 - 4] = [-3, 1] \\ &= (4, 1) - (5, 1) = (4 - 5, 1 + 1) = (-1, 2) \end{aligned}$$

對於區間模糊數的計算，區間模糊數減法常常被使用於區間預測的誤差定義，爲了讓區間模糊數可以有個反模糊值做比較，同時兼顧中心點與半徑，我們將定義反模糊化函數，定義爲下方的函數式：

$$c_t + \frac{|b_t - a_t|}{2 \ln(e + c_t)}$$

模糊區間誤差的定義對於不同的標的或屬性，往往有不同的定義方式，有時偏重中心點的預測，有時偏重半徑的預測，有時必須兩者兼顧，或者給予某種的權重比例，而對於台灣加權股股價指數而言，由於先天受到漲跌幅的限制（上漲下跌各 7%），使得每日中心點誤差其實遠大於半徑誤差（表 3.1）。根據 2005 年吳柏林教授所提出的區間誤差定義，採取的是中心點誤差與自然對數半徑誤差的方式，有同時兼顧兩者的性質，其運算的公式如下：

$$d(X_t, \hat{X}_t) = |c_t - \hat{c}_t| + \ln(1 + |r_t - \hat{r}_t|)$$

- 根據上面的區間誤差的定義我做少許的修正，這些修正包含三個面向
- 第一、當中心點相同時，其誤差來源主要是半徑的誤差
 - 第二、當半徑相同時，其誤差來源主要是中心點的誤差
 - 第三、當兩者都不同時，我們調整兩者誤差的權重應該相當接近

透過大約 100 次的台灣加權股價指數資料計算發現，我們將半徑誤差給與

修正為 $\left| \frac{|b_t - a_t|}{2 \ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2 \ln(e + \hat{c}_t)} \right|$ 時，可以同時兼顧兩種誤差的權重達各半的

效果，除此之外中心點在分母的用意可以使得指數墊高之後，半徑誤差權重不至於走樣，最後我們將定義表示如下。

$$d(X_t, \hat{X}_t) = |c_t - \hat{c}_t| + \left| \frac{|b_t - a_t|}{2 \ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2 \ln(e + \hat{c}_t)} \right|$$

定義 2.4 反模糊值

令 $\{X_t = [a_t, b_t] = [c_t; r_t], t = 1, 2, \dots, n\}$ 為一個區間時間數列，則 $[a_t, b_t]$ 之反模糊值為

$$c_t + \frac{|b_t - a_t|}{2 \ln(e + c_t)}$$

2.3 模糊區間距離

定義 2.5 模糊區間距離 (distance of interval)

令 $\{X_t = [a_t, b_t] = [c_t; r_t], t = 1, 2, \dots, n\}$ 為一個區間時間數列，預測區間為 $\hat{X}_t = [\hat{a}_t, \hat{b}_t] = [\hat{c}_t; \hat{r}_t]$ ，則區間距離

$$d(X_t, \hat{X}_t) = |c_t - \hat{c}_t| + \left| \frac{|b_t - a_t|}{2 \ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2 \ln(e + \hat{c}_t)} \right|$$

例 2.4 $X_1 = [3,4] = (3.5; 0.5)$, $\hat{X}_1 = [3,6] = (4.5; 1.5)$

$$d(X_1, \hat{X}_1) = |3.5 - 4.5| + \left| \frac{1}{2 \ln(e + 3.5)} - \frac{3}{2 \ln(e + 4.5)} \right| = 1.49$$

$X_2 = [2,6] = (4; 2)$, $\hat{X}_2 = [4,5] = (4.5; 0.5)$

$$d(X_2, \hat{X}_2) = |4 - 4.5| + \left| \frac{4}{2 \ln(e + 4)} - \frac{1}{2 \ln(e + 4.5)} \right| = 0.77$$

$X_3 = [3,5] = (4; 1)$, $\hat{X}_3 = [2,6] = (4; 2)$

$$d(X_3, \hat{X}_3) = |4 - 4| + \left| \frac{2}{2 \ln(e + 4)} - \frac{4}{2 \ln(e + 4)} \right| = 0.52$$

$X_4 = [5,7] = (6; 1)$, $\hat{X}_4 = [5,8] = (6.5; 1.5)$

$$d(X_4, \hat{X}_4) = |6 - 6.5| + \left| \frac{2}{2 \ln(e + 6)} - \frac{3}{2 \ln(e + 6.5)} \right| = 0.98$$

$X_5 = [4,5] = (4.5; 0.5)$, $\hat{X}_5 = [3,8] = (5.5; 2.5)$

$$d(X_5, \hat{X}_5) = |4.5 - 5.5| + \left| \frac{1}{2 \ln(e + 4.5)} - \frac{5}{2 \ln(e + 5.5)} \right| = 1.93$$

定義 2.6 區間平均誤差 (Interval Mean Error)

令 $\{X_t = [a_t, b_t] = [c_t; r_t], t = 1, 2, \dots, n\}$ 為一個區間時間數列，預測區間為 $\hat{X}_t = [\hat{a}_t, \hat{b}_t] = [\hat{c}_t; \hat{r}_t]$ ， $\hat{\varepsilon}_t = d(X_t, \hat{X}_t)$ 為預測區間與實際區間之誤差，則定義區間平均誤差為

$$IME = \frac{1}{k} \sum_{t=n+1}^{n+k} \hat{\varepsilon}_t = \frac{1}{k} \sum_{t=n+1}^{n+k} |c_t - \hat{c}_t| + \left| \frac{|b_t - a_t|}{2 \ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2 \ln(e + \hat{c}_t)} \right|$$

2.4 ARIMA 模型配適度檢測

定義 2.7 赤池信息量準則 (Akaike information criterion、AIC)

赤池信息量準則是一種判斷時間序列模型配適訊息的準則，一般來說，數值愈小，時間序列模型的配適較好。公式如下：

$$AIC(p, q) = -2 * \ln(T) + 2(p + q)$$

其中 T 是最大概似函數值， p 、 q 為 ARIMA(p, d, q) 的參數。

3. 實證分析

本章我們想應用 ARIMA 預測法，對台灣加權股價指數 (TAIEX) 做統計分析，找出每日大盤最低最高價區間價，透過模糊區間的計算給出新的預測方法並且與傳統的預測方法做比較。

3.1 資料來源

資料來源為台灣證券交易所網站 (<http://www.twse.com.tw/ch/>)，下載西元 2012 年第一季所有交易日之最低價與最高價，並將原始資料經過 MINITAB 統計軟體繪出本研究所需之時間序列走勢圖如下。

表 3.1 2012 年台灣加權股價指數 (Q1) 交易日最高價最低價資料

日期	最低價	最高價	價格區間	區間表示
1/2	7071.35	7073.48	[7071.35 , 7073.48]	(7072.415 , 1.065)
1/3	7005.39	7073.79	[7005.39 , 7073.79]	(7039.59 , 34.2)
1/4	7093.07	7116.22	[7093.07 , 7116.22]	(7104.645 , 11.575)
1/5	7099.35	7130.86	[7099.35 , 7130.86]	(7115.105 , 15.755)
1/6	7129.53	7139.04	[7129.53 , 7139.04]	(7134.285 , 4.755)
1/9	7103.09	7103.09	[7103.09 , 7103.09]	(7103.09 , 0)
1/10	7128.86	7181.74	[7128.86 , 7181.74]	(7155.3 , 26.44)
1/11	7203.10	7212.36	[7203.10 , 7212.36]	(7207.73 , 4.63)
1/12	7202.38	7209.74	[7202.38 , 7209.74]	(7206.06 , 3.68)
1/13	7226.24	7249.94	[7226.24 , 7249.94]	(7238.09 , 11.85)
1/16	7241.92	7258.84	[7241.92 , 7258.84]	(7250.38 , 8.46)
1/17	7144.20	7221.08	[7144.20 , 7221.08]	(7182.64 , 38.44)
1/18	7235.54	7245.13	[7235.54 , 7245.13]	(7240.335 , 4.795)
1/30	7382.72	7443.03	[7382.72 , 7443.03]	(7412.875 , 30.155)
1/31	7436.52	7517.08	[7436.52 , 7517.08]	(7476.8 , 40.28)
2/1	7488.97	7566.08	[7488.97 , 7566.08]	(7527.525 , 38.555)
2/2	7578.76	7652.46	[7578.76 , 7652.46]	(7615.61 , 36.85)
2/3	7608.41	7674.99	[7608.41 , 7674.99]	(7641.7 , 33.29)
2/4	7727.29	7787.89	[7727.29 , 7787.89]	(7757.59 , 30.3)
2/6	7677.56	7752.28	[7677.56 , 7752.28]	(7714.92 , 37.36)
2/7	7673.91	7751.76	[7673.91 , 7751.76]	(7712.835 , 38.925)
2/8	7766.38	7869.91	[7766.38 , 7869.91]	(7818.145 , 51.765)
2/9	7840.38	7910.78	[7840.38 , 7910.78]	(7875.58 , 35.2)
2/10	7848.78	7941.63	[7848.78 , 7941.63]	(7895.205 , 46.425)

2/13	7837.58	7917.96	[7837.58 , 7917.96]	(7877.77 , 40.19)
2/14	7829.76	7937.28	[7829.76 , 7937.28]	(7883.52 , 53.76)
2/15	7902.08	8011.87	[7902.08 , 8011.87]	(7956.975 , 54.895)
2/16	7851.62	8010.50	[7851.62 , 8010.50]	(7931.06 , 79.44)
2/17	7862.11	8013.46	[7862.11 , 8013.46]	(7937.785 , 75.675)
2/20	7913.65	7985.50	[7913.65 , 7985.50]	(7949.575 , 35.925)
2/21	7870.26	7973.47	[7870.26 , 7973.47]	(7921.865 , 51.605)
2/22	7897.65	8018.54	[7897.65 , 8018.54]	(7958.095 , 60.445)
2/23	7935.42	7991.92	[7935.42 , 7991.92]	(7963.67 , 28.25)
2/24	7899.62	7974.58	[7899.62 , 7974.58]	(7937.1 , 37.48)
2/29	8046.14	8121.44	[8046.14 , 8121.44]	(8083.79 , 37.65)
3/1	8103.75	8151.70	[8151.70 , 8103.75]	(8127.725 , 23.975)
3/2	8114.14	8170.72	[8114.14 , 8170.72]	(8142.43 , 28.29)
3/3	8087.20	8134.94	[8087.20 , 8134.94]	(8111.07 , 23.87)
3/5	7998.96	8115.49	[7998.96 , 8115.49]	(8057.225 , 58.265)
3/6	7860.87	8028.02	[7860.87 , 8028.02]	(7944.445 , 83.575)
3/7	7824.45	7937.11	[7824.45 , 7937.11]	(7880.78 , 56.33)
3/8	7905.50	8000.93	[7905.50 , 8000.93]	(7953.215 , 47.715)
3/9	7965.68	8027.08	[7965.68 , 8027.08]	(7996.38 , 30.7)
3/12	7927.55	8028.87	[7927.55 , 8028.87]	(7978.21 , 50.66)
3/13	7974.38	8045.61	[7974.38 , 8045.61]	(8009.995 , 35.615)
3/14	8113.76	8170.31	[8113.76 , 8170.31]	(8142.035 , 28.275)
3/15	8088.10	8142.98	[8088.10 , 8142.98]	(8115.54 , 27.44)
3/16	8054.94	8130.52	[8054.94 , 8130.52]	(8092.73 , 37.79)
3/19	7997.05	8123.15	[7997.05 , 8123.15]	(8060.1 , 63.05)
3/20	7965.96	8062.91	[7965.96 , 8062.91]	(8014.435 , 48.475)
3/21	7940.19	8013.67	[7940.19 , 8013.67]	(7976.93 , 36.74)
3/22	7995.57	8075.97	[7995.57 , 8075.97]	(8035.77 , 40.2)
3/23	8030.88	8076.61	[8030.88 , 8076.61]	(8053.745 , 22.865)
3/26	7958.72	8073.93	[7958.72 , 8073.93]	(8016.325 , 57.605)
3/27	7978.74	8033.22	[7978.74 , 8033.22]	(8005.98 , 27.24)
3/28	8013.80	8085.43	[8013.80 , 8085.43]	(8049.615 , 35.815)
3/29	7808.55	7997.87	[7808.55 , 7997.87]	(7903.21 , 94.66)
3/30	7775.75	7933.00	[7775.75 , 7933.00]	(7854.375 , 78.625)
4/2	7859.56	7928.43	[7859.56 , 7928.43]	(7893.995 , 34.435)
4/3	7723.39	7907.62	[7723.39 , 7907.62]	(7815.505 , 92.115)

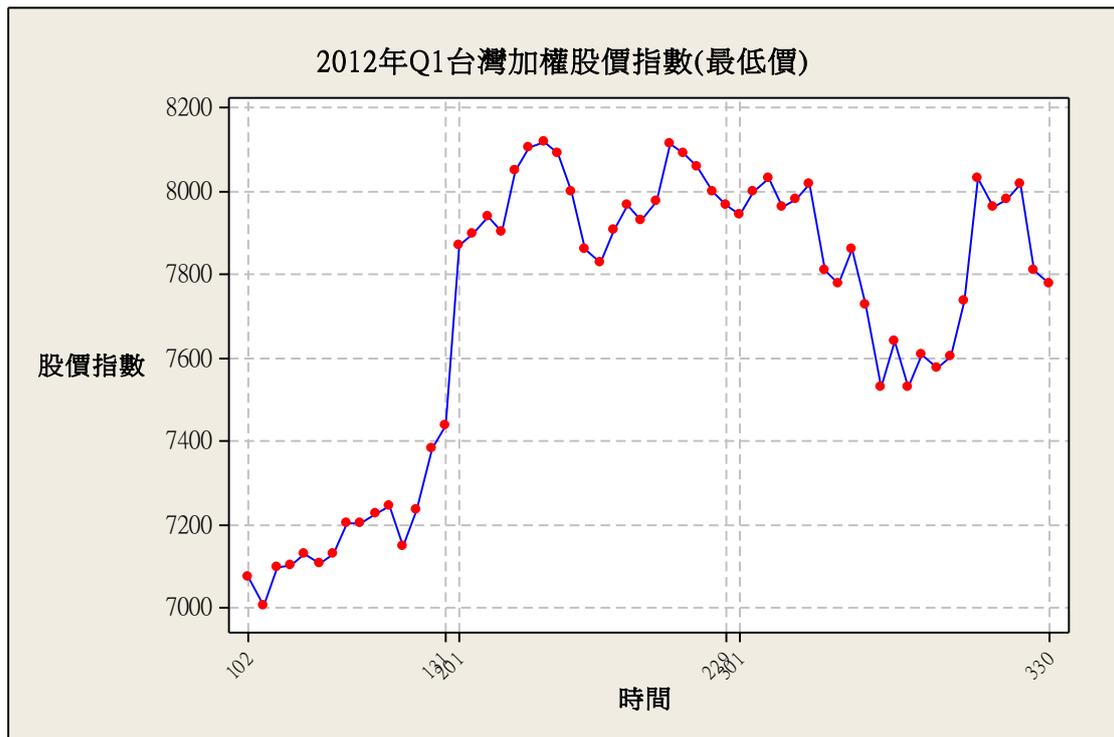


圖 3.1 西元 2012 年 Q1 台灣加權股價指數最低價時間數列走勢圖

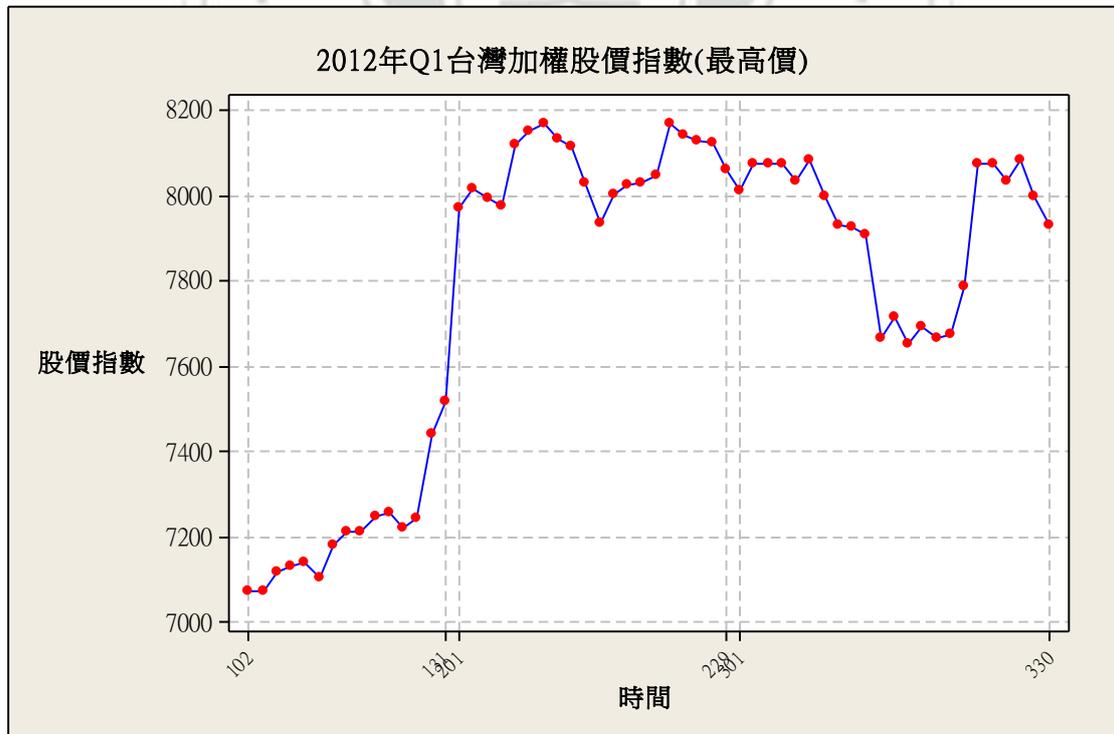


圖 3.2 西元 2012 年 Q1 台灣加權股價指數最高價時間數列走勢圖

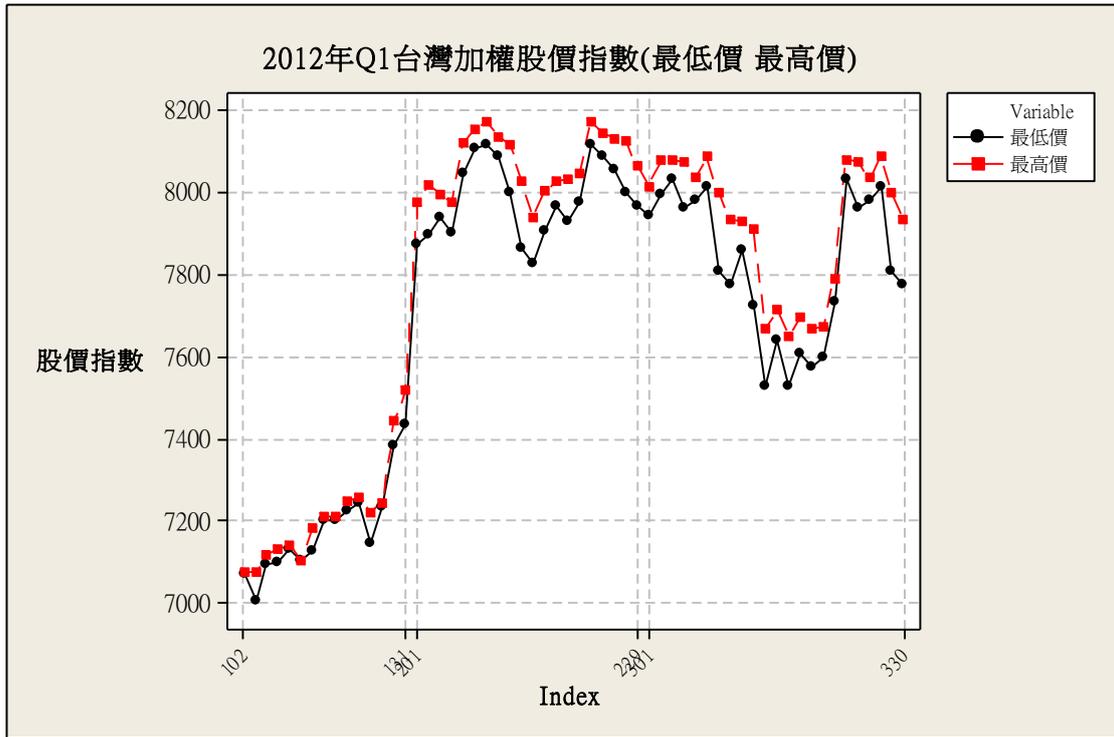


圖 3.3 西元 2012 年 Q1 台灣加權股價指數最低價、最高價時間數列走勢圖

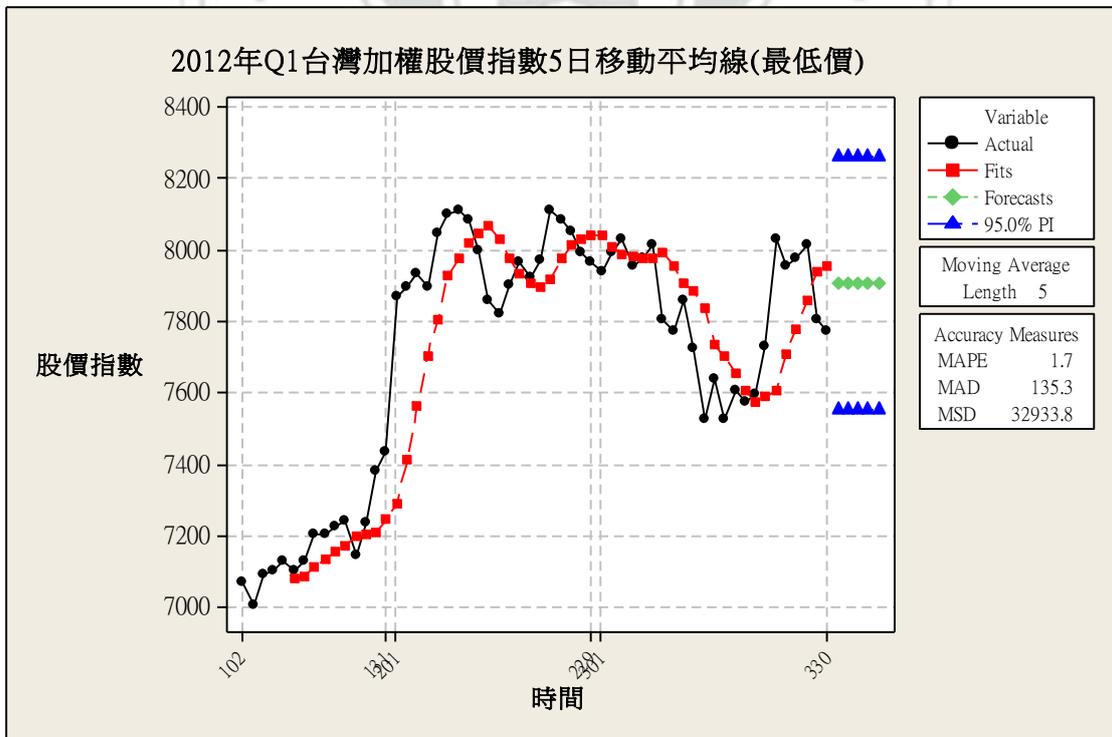


圖 3.4 西元 2012 年 Q1 台灣加權股價指數最低價五日簡單移動平均線走勢圖

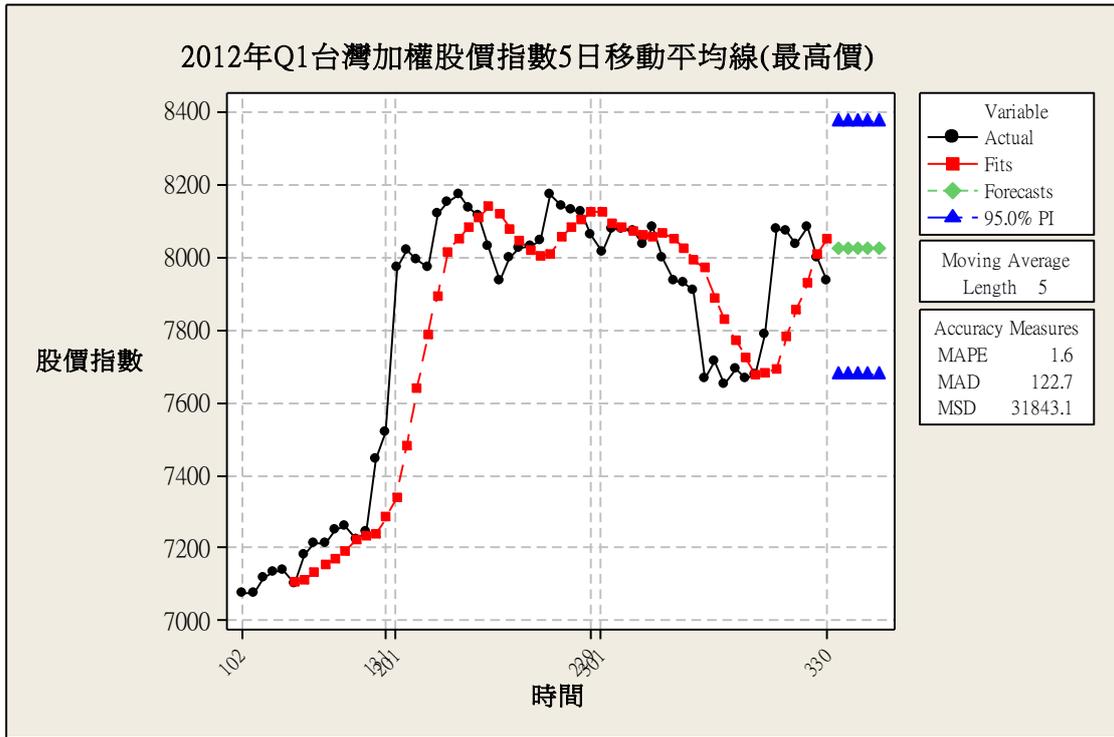


圖 3.5 西元 2012 年 Q1 台灣加權股價指數最高價五日簡單移動平均線走勢圖

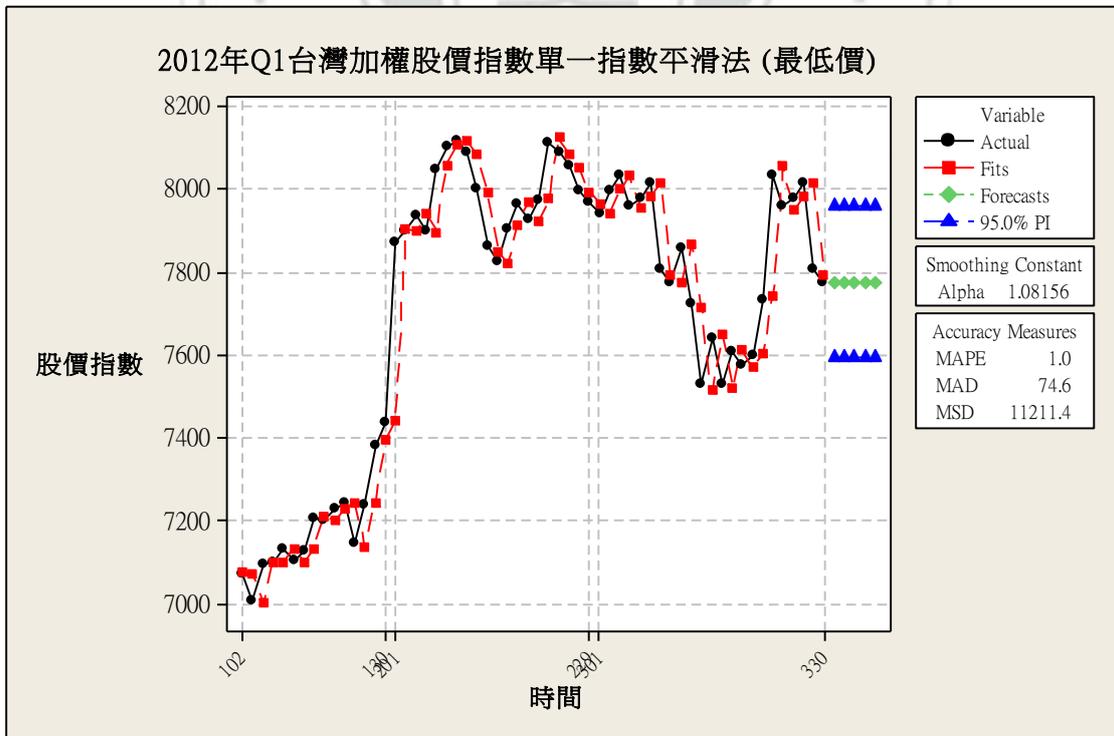


圖 3.6 西元 2012 年 Q1 台灣加權股價指數之最低價指數平滑法

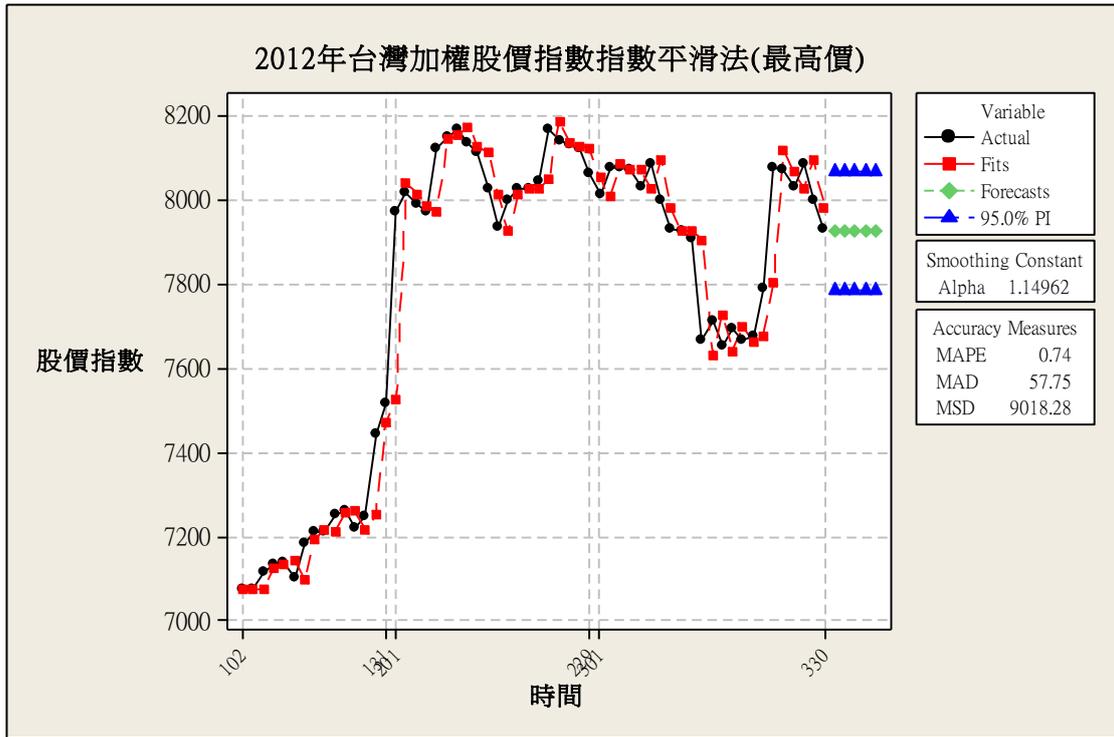


圖 3.7 西元 2012 年 Q1 台灣加權股價指數之最高價指數平滑法

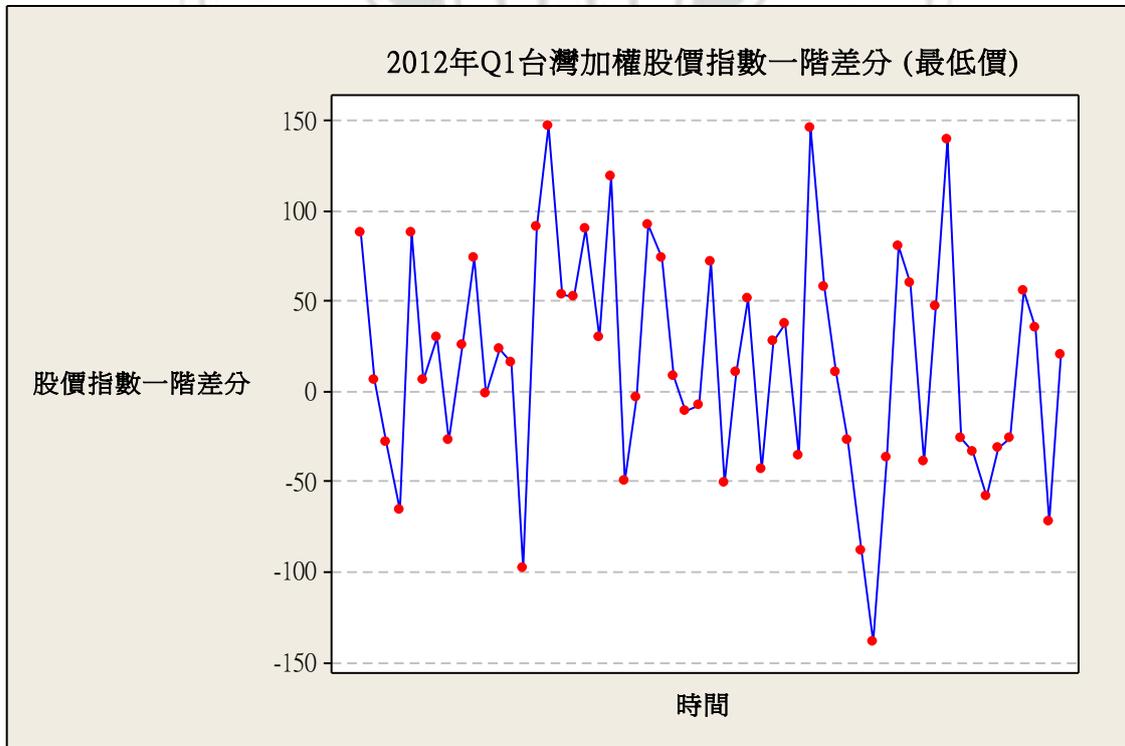


圖 3.8 為 2012 年 Q1 台灣加權股價指數之最低價一階差分時間數列走勢圖

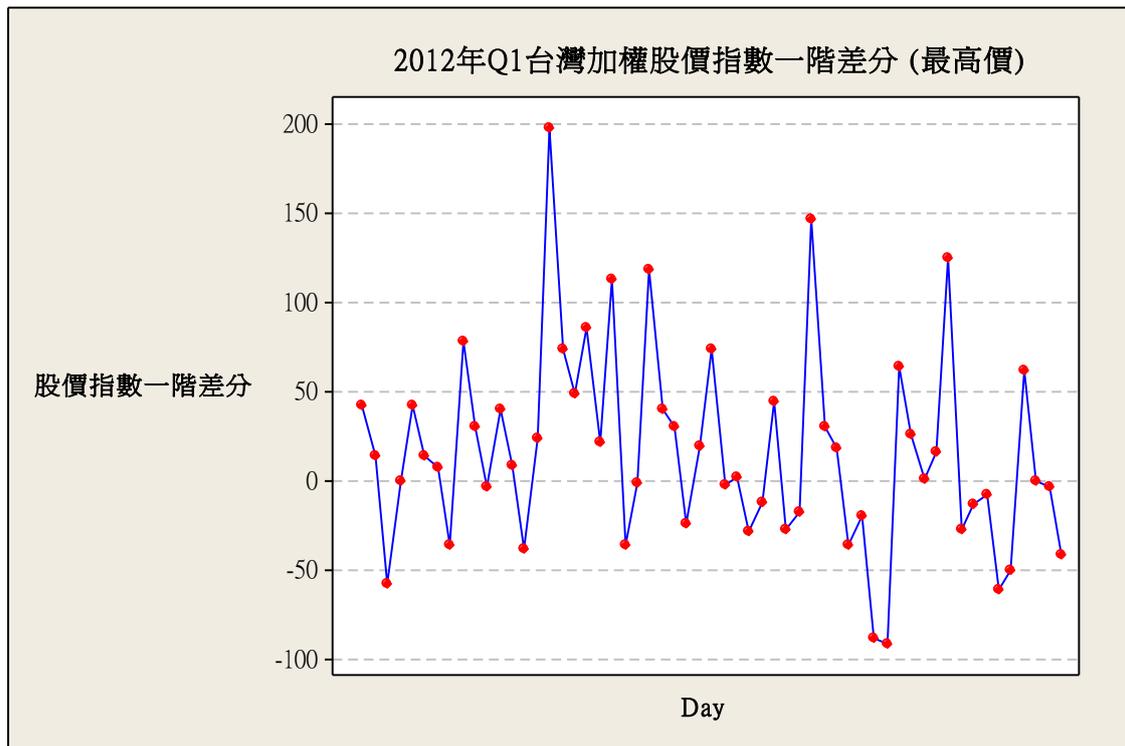


圖 3.9 為 2012 年 Q1 台灣加權股價指數之最高價一階差分時間數列走勢圖。

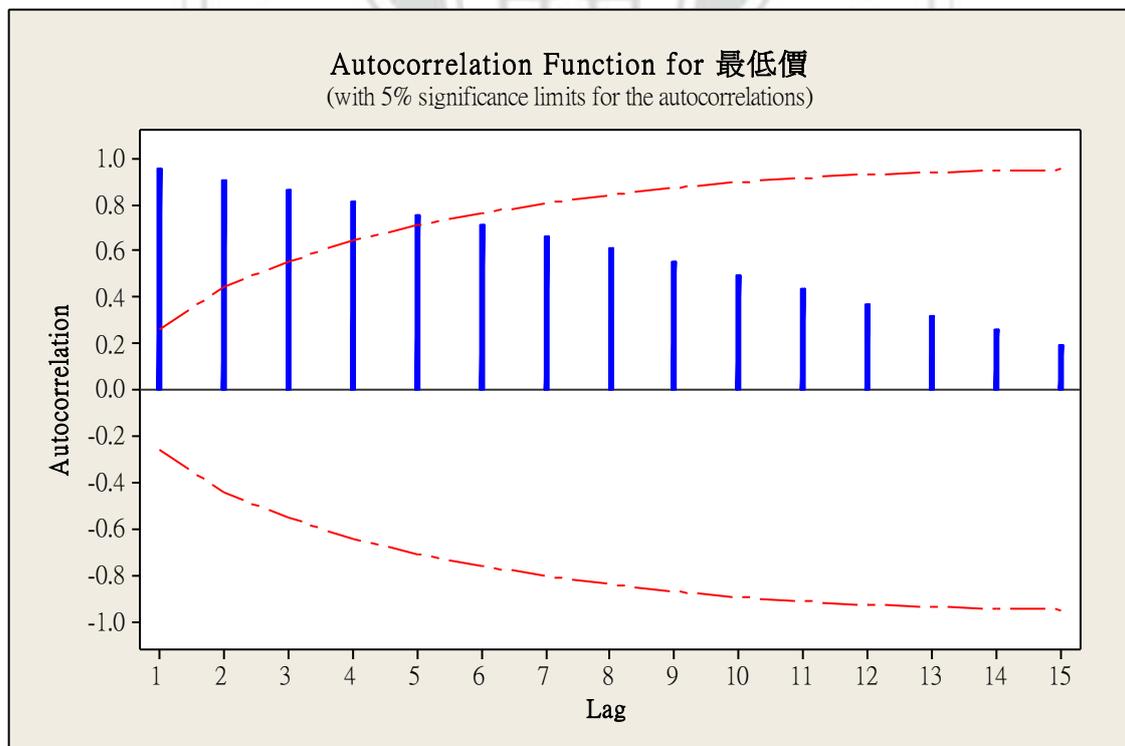


圖 3.10 為 2012 年 Q1 台灣加權股價指數之最低價 ACF 圖。

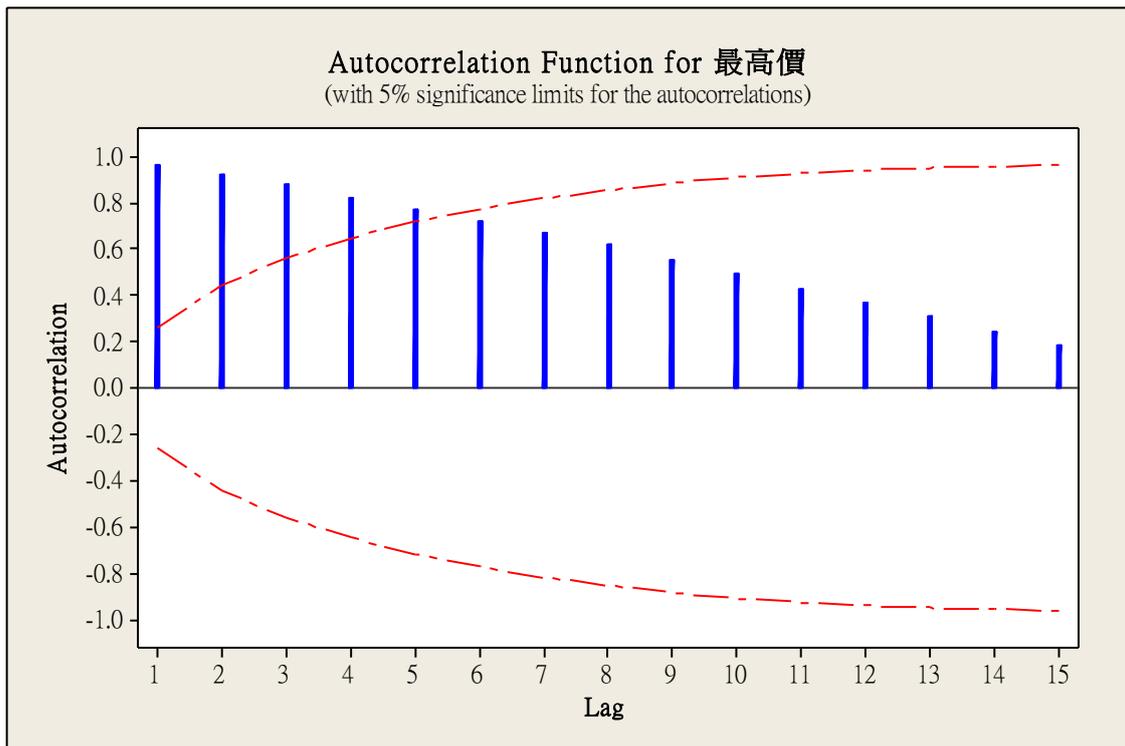


圖 3.11 為 2012 年 Q1 台灣加權股價指數之最高價 ACF 圖。

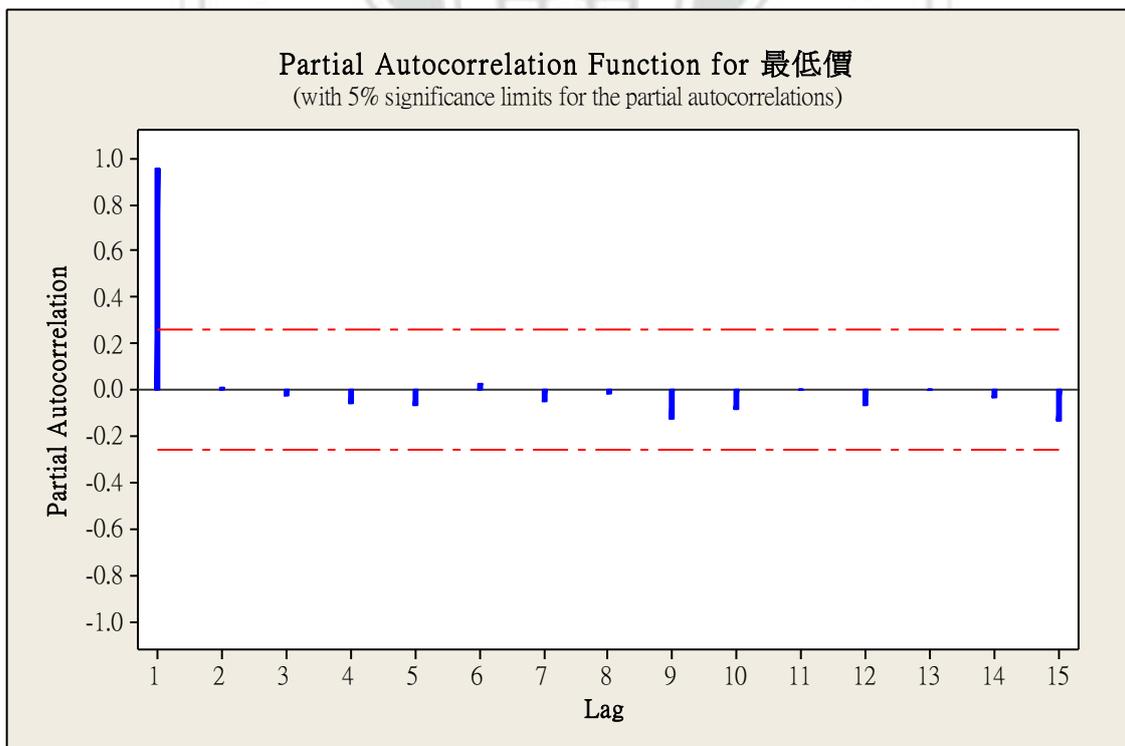


圖 3.12 為 2012 年 Q1 台灣加權股價指數之最低價 PACF 圖。

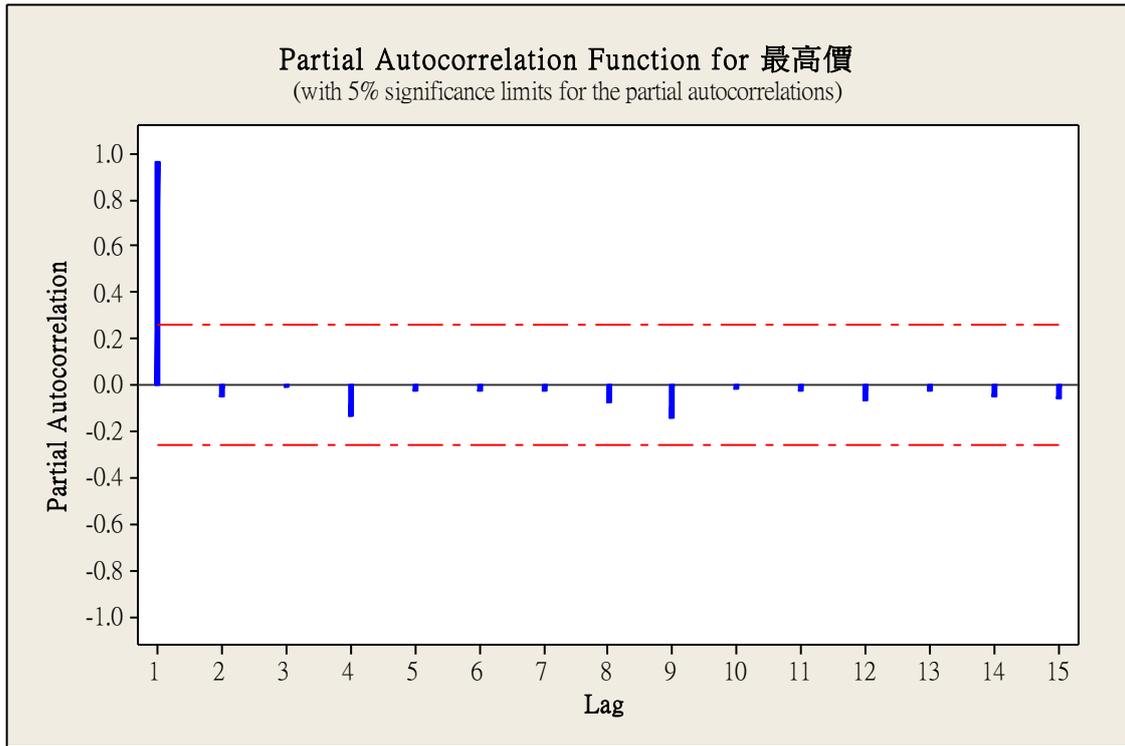


圖 3.13 為 2012 年 Q1 台灣加權股價指數之最高價 PACF 圖。

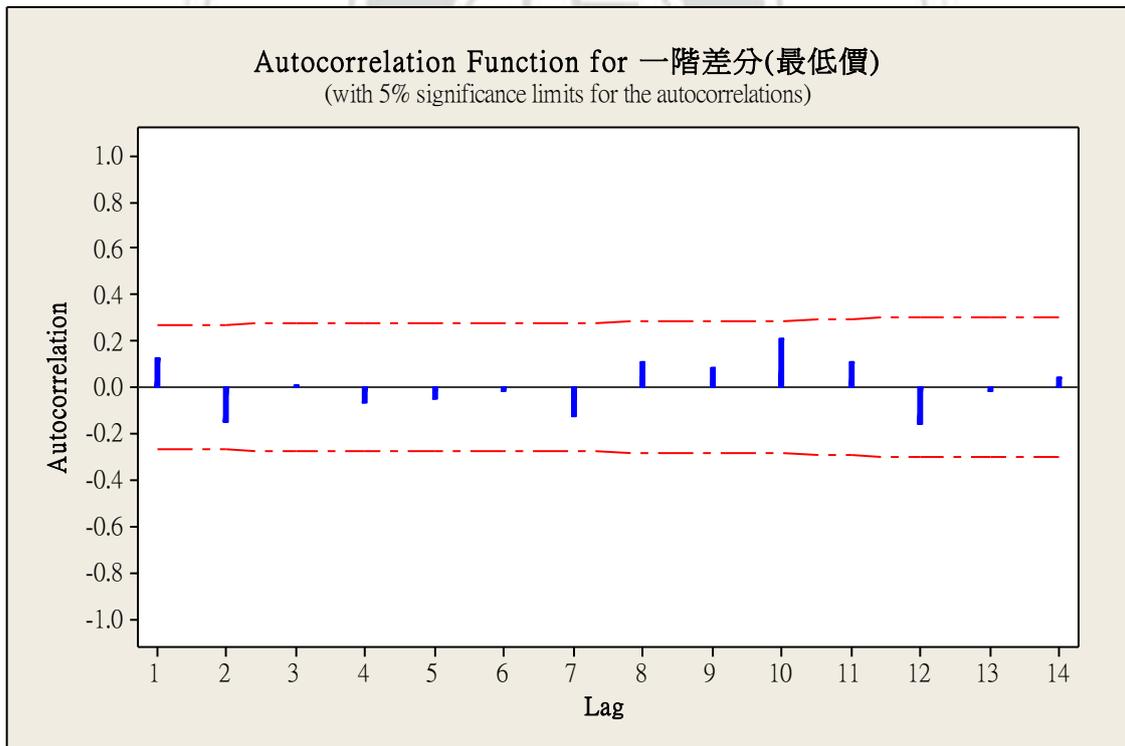


圖 3.14 為 2012 年 Q1 台灣加權股價指數之最低價一階差分 ACF 圖。

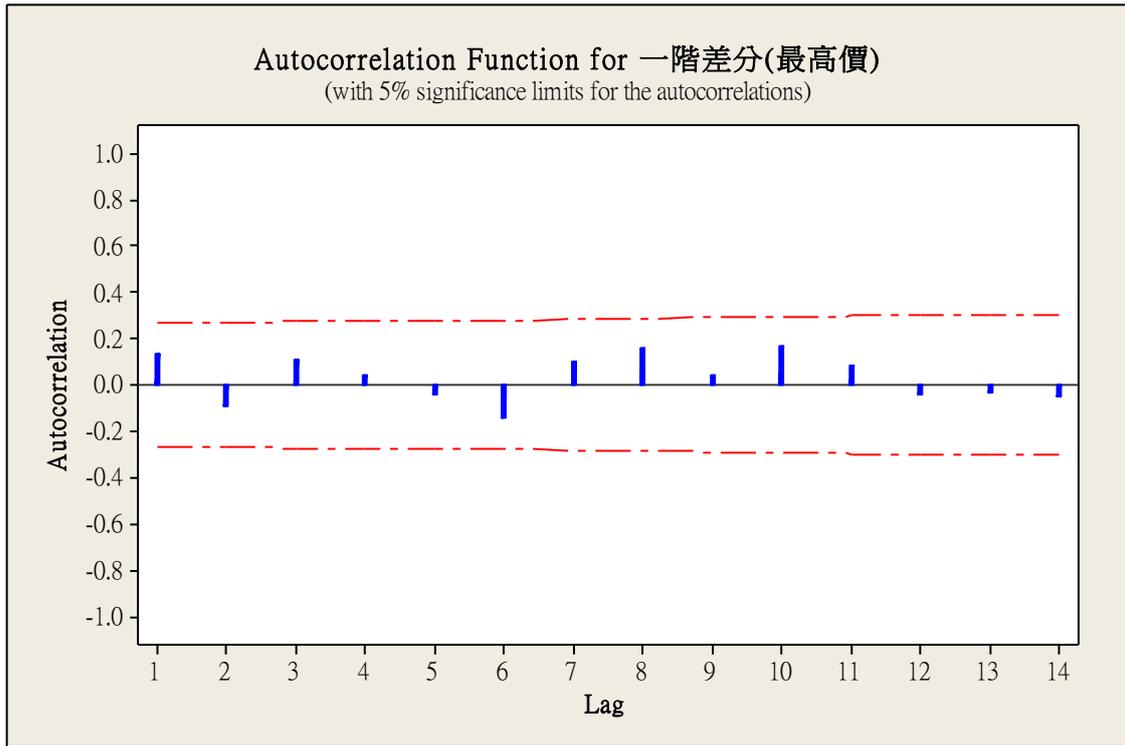


圖 3.15 為 2012 年 Q1 台灣加權股價指數之最高價一階差分 ACF 圖。

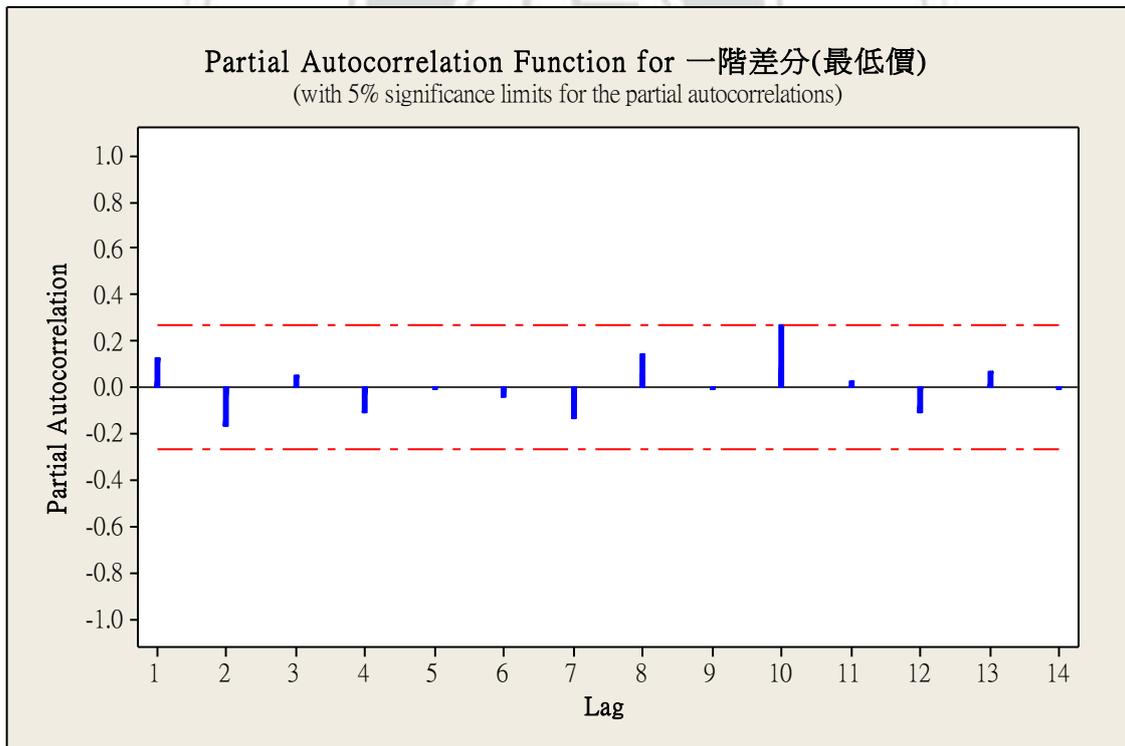


圖 3.16 為 2012 年 Q1 台灣加權股價指數之最低價一階差分 PACF 圖。

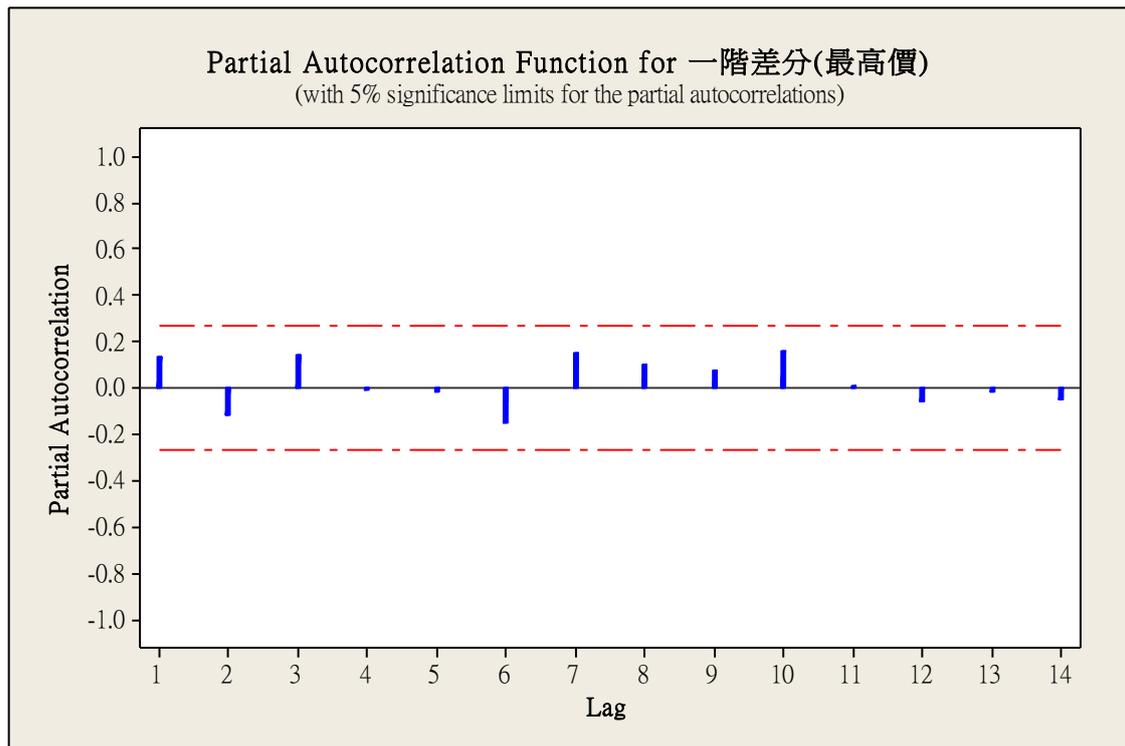


圖 3.17 為 2012 年 Q1 台灣加權股價指數之最高價一階差分 PACF 圖。

3.2 預測方法探討

3.2.1 五日移動平均法 (Simple Moving Average Method)

本研究的標的是大盤股價指數，對於台灣加權股價指數而言，一周只有五天是交易日，為了研究短時間股價變化的方便，我們採用五日移動平均法，同時經過 MINITAB 的演算後，可以得到六期預測值，預測值大部分散落在 7775.75 點到 8012.38 點之間。預測值整理如下方：

	五日移動平均法
預測值 1	[7907.112 , 8024.69]
預測值 2	[7887.280 , 7995.59]
預測值 3	[7836.21 , 7970.47]
預測值 4	[7739.056 , 7886.792]
預測值 5	[7704.956 , 7829.866]
預測值 6	[7655.684 , 7773.444]

3.2.2 單一指數平滑法

在本研究中，單一指數平滑法所採取的調整係數 $\alpha = 0.25$ ，這個參數取得是經過嘗試 0.25、0.5、0.75 後所得，同時也是較常用來預測股價的參數之一，同時經過 MINTAB 的演算後，可以得到六期預測值，預測值大部分散落在 7553.201 點到 7928.43 點之間。預測值整理如下方：

	單一指數平滑法 ($\alpha = 0.25$)
預測值 1	[7843.705 , 7934.812]
預測值 2	[7753.469 , 7914.418]
預測值 3	[7584.39 , 7728.885]
預測值 4	[7624.635 , 7717.151]
預測值 5	[7553.201 , 7667.455]
預測值 6	[7592.553 , 7686.569]

3.2.3 ARIMA 法 (以最低價為例子)

步驟一：判斷是否為定態時間序列

步驟二：判斷適當的參數 (p , d , q)

步驟三：利用 AIC 判別準則找出適合的模型

	ARIMA 法
預測值 1	[7883.811 , 8053.412]
預測值 2	[7934.145 , 8034.913]
預測值 3	[7945.21 , 8010.384]
預測值 4	[7903.233 , 8015.234]
預測值 5	[7752.55 , 7934.494]
預測值 6	[7875.75 , 7937.59]

3.3 預測效率分析

綜觀台灣加權股價指數的波動範圍大約介在 3000 ~ 12000 點，而每天的漲跌幅是上漲或下跌 7%，從這些基本的大盤資訊，我們可以發現要精確預測必須要從預測效率的定義下手，有關預測效率的定義其實已經有很多的論文討論過，如平

均百分誤差(MPE)、平均誤差平方和(MSE)、平均絕對誤差(MAE)、平均絕對百分誤差(MAPE)等等，但大部分都是屬於點誤差的範圍，預測的效果偏重於單個點來做預測，對於區間的資料其實並不適用，特別是本研究為台灣加權股價指數每日指數區間的預測，因此研究的初期採用 2005 年吳柏林教授所提出的區間距離公式

$$d(\mathbf{X}_t, \hat{\mathbf{X}}_t) = |c_t - \hat{c}_t| + \ln(1 + |r_t - \hat{r}_t|) \quad d(X_t, \hat{X}_t) = |c_t - \hat{c}_t| + \ln(1 + |r_t - \hat{r}_t|)$$

為了測試此區間距離公式是否適用於大盤指數，我們拿 100 筆大盤的資料，去實際測試發現，在區間距離公式中，區間中心點誤差項往往因為數字很大而決定了大部分的區間距離取值，對於區間半徑的變動變的不太靈敏，或者我們可以說，因為台灣特殊的漲跌幅限制使得區間半徑誤差項變的不是那麼重要。因此在區間距離方面我們採用新的區間距離定義，如下：

$$d(x_t, \hat{x}_t) = |c_t - \hat{c}_t| + \left| \frac{|b_t - a_t|}{2\ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2\ln(e + \hat{c}_t)} \right|$$

將資料反模糊化(defuzzification)轉換為合適的實數值，使區間距離計算較不容易單單只受到中心點誤差或者是半徑誤差的影響，而且能將所得的數值做運算。

如果我們分析一下新的區間距離定義會發現：

- 第一、當中心點相同時，誤差取決於每天最高最低價的價差(即兩倍半徑)
- 第二、當半徑相同(絕大部分都是受漲跌幅 7% 的限制)，可以看出誤差範圍幾乎取決於每日的中心點誤差
- 第三、當中心點與半徑都不同時，由測試發現 $|c_t - \hat{c}_t|$ 與

$$\left| \frac{|b_t - a_t|}{2\ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2\ln(e + \hat{c}_t)} \right| \text{ 的值幾乎相同，也就是說中心點誤差}$$

與半徑誤差是相當平均的，並不會只受單一因素影響誤差的值。

從上面這三點區間誤差的預測分析大致上符合經驗法則，且不會有任何偏重中心點或半徑的情形，進一步將誤差取平均值，我們提出區間平均誤差(Interval Mean Error: 簡稱 IME) 的公式：

$$IME = \frac{1}{k} \sum_{t=n+1}^{n+k} \hat{\varepsilon}_t = \frac{1}{k} \sum_{t=n+1}^{n+k} |c_t - \hat{c}_t| + \left| \frac{|b_t - a_t|}{2 \ln(e + c_t)} - \frac{|\hat{b}_t - \hat{a}_t|}{2 \ln(e + \hat{c}_t)} \right|$$

並且著手計算各種預測方法的 IME 區間平均誤差值，由於預測的方法屬於較短期的方法，因此只取了六期的預測值，而其計算的結果都記錄結論的表格中。

4. 結論

2012 年 4 月份的台灣加權股價指數，因為政府即將通過證所稅的預期心理，使得股市交易量萎縮，為了不影響本研究的預測效果，本研究決定採用 Q1 (西元 2012 年 1 月到西元 2012 年 3 月) 的資料作為實證分析的材料。

本研究透過三種方法 (1)五日移動平均線法 (2)單一指數平滑法 (3) ARIMA 法做預測分別得到了六組區間預測值，並且透過 IME 計算區間的預測，當作判斷區間預測效力的優劣。

表 4.1 三種不同預測方法預測結果比較

	五日移動平均法	單一指數平滑法	ARIMA 法
預測值 1	[7907.112 , 8024.69]	[7843.705 , 934.812]	[7883.811 , 8053.412]
預測值 2	[7887.28, 7995.59]	[7753.469 ,7914.418]	[7934.145 , 8034.913]
預測值 3	[7836.21 , 7970.47]	[7584.39 , 7728.885]	[7945.21, 8010.384]
預測值 4	[7739.056 , 7886.792]	[7624.635 ,7717.151]	[7903.233 , 8015.234]
預測值 5	[7704.956 , 7829.866]	[7553.201 ,7667.455]	[7752.55 , 7934.494]
預測值 6	[7655.684 , 7773.444]	[7592.553 ,7686.569]	[7875.75 , 7937.59]
實際值	[7859.56 , 7928.43]		
IME 值	60.18862	54.187341	45.35678

由上面三種方法計算而得到的 IME 值計算，可以發現 ARIMA 方法所預測效率最佳，單一指數平滑法次之，五日移動平均法最差。由於台灣加權股價指數在第一季的元月，通常都有所謂的元月行情（紅包行情），從線圖來看也真的是如此，五日移動平均法因為反應比較長時間的資訊，所以預測效率比較差，而 ARIMA 法因為比較貼近資料，因此預測效率最好，單一指數平滑法因為調整這些數據，因此預測效率介於這兩者之間，其實 2012 年 Q1 的股價波動並不大，相對於預測而言，三種方法都算相當的精準，也提供我們投資或研究的數據。

由於預測的方法日新月異，本研究只用了三種常見的預測方法，研究者若要取得更多的資訊，應參考吳柏林教授於 1995 年所提出來的預測比較表 (表 3.2)，台灣加權指數大體而言介在 3000 ~ 12000 點，因此區間的距離公式在預測效率扮演重要的角色，因此建議先收集大量的歷史資料，不斷的測試區間距離公式是否合理，並反覆修正區間距離公式，才能提高預測的效率。同時，因為研究的是經濟數據，難免會受到政府政策的影響，遇到政府有強勢主導政策時，大盤極易受到干擾，應該改採其他的預測方法。本研究只有考慮單純的區間預測效率，對於未來的研究者或投資者，研究的標的甚至可以延伸到房價、外匯市場、黃金市場等。研究的應用可以加入投資金額控管和風險值的概念，讓研究更實用。

5. 參考文獻

中文部分

- [1] 吳柏林 1995 時間序列分析與導論 台北 華泰書局。
- [2] 吳柏林 2005 模糊統計導論方法與應用 台北 五南書局。
- [3] 吳柏林 林玉鈞 2002 模糊時間數列分析與預測—以台灣地區加權股價指數為例 應用數學學報 第 25 卷 第一期 頁 67-76
- [4] 林茂文 1992 時間數列分析與預測 台北 華泰書局。
- [5] 王文俊 認識 Fuzzy 第三版 2005。
- [6] 曾淑惠 2004 多變量模糊時間數列模式之應用以台灣地區高職教師人數之預測為例 教育與心理研究 第 27 卷 第四期 頁 845-861。

英文部分

- [1] H. T. Nguyen and B. Wu (2006) Fundamentals of Statistics with Fuzzy Data. New York : Springer
- [2] S. M. Chen (1996) Forecasting enrollments based on fuzzy time series. Fuzzy sets and systems, 81, 311-319.
- [3] Wu, B. (1995). Model-free forecasting for nonlinear time series: with application in exchange rates. *Computational Statistics and Data Analysis*. **19**, 433-459.
- [4] Wu, B., Chen, C. and Chen (1999) Application of time series analysis in quality control. *Quality Control Journal*. **35(7)**, 68-76.
- [5] Q. Song, B.S. Chissom, "Fuzzy time series and its models," Fuzzy Sets and Systems. Vol. 54, 1993, pp.269-277.
- [6] Q. Song, B.S. Chissom, "Forecasting enrollments with fuzzy time series-part I," Fuzzy Sets and Systems. Vol. 54, 1993, pp.1-9.
- [7] Q. Song, B.S. Chissom, "Forecasting enrollments with fuzzy time series-part II," Fuzzy Sets and Systems Vol. 62, 1994, pp.1-8.