國立政治大學理學院應用物理研究所碩士論文

Graduate Institute of Applied Physics, College of Science National Chengchi University Master Thesis

1996-1999 年美國股票群的收益以高頻日移動平均計算之統計與動力性質分析

Statistical and Dynamical Properties of Returns Using High Frequency 1-day Moving Averages For Collections of U.S Stocks Over 1996-1999

王柏淵

Wang, Bo Yuan

指導教授：馬文忠

Advisor：Ma, Wen Jong, Ph.D.

中 華 民 國 一 ○ 二 年 六 月

June, 2013

國立政治大學 National Chengchi University

誌謝

三年的時光匆匆飛逝，轉眼間研究所生涯即將邁入尾聲，準備踏上人生的另一段旅程，心裡總有一股溢於言表的感觸，在這裡誠摯感謝讓我有所蛻變的一切。在碩士求學與撰寫論文過程當中遇到了許多挫折，由於許多人的幫助、支持以及鼓勵，讓我不斷調整腳步與心情，才得以完成此篇論文。

首先感謝我的指導教授馬文忠老師，在論文撰寫期間給予的教導與督促，不厭其煩地解釋所遇到的任何問題，並且從中學習到對於研究的執著與態度；在求學期間，感謝所上的老師傳授各門學科的專業知識，讓我在撰寫論文時助益良多；在論文口試時，感謝本所蕭又新老師、中研院研究員胡進錕老師以及國立東華大學陳企寧老師所給予的寶貴意見與指正，讓此篇碩士論文更趨於完善，在此表示由衷的謝忱。

另外，也感謝三年來陪伴我渡過研究生涯的大家，給我許多的幫助與鼓勵，一起敦促一起學習一起成長。謝謝張凱鈞同學在程式方面的大力幫助，謝謝蕭逸修同學在課業方面的不吝解惑，謝謝陳原孝同學分享的許多事情與人生觀，謝謝王瑞騰同學在忙碌之餘陪我聊聊心事，謝謝董光平同學與謝智勛同學一起討論課業的許多問題，謝謝曾嘉瑤同學撰寫論文期間的陪伴與鼓勵，謝謝張景婷同學課堂筆記的幫助，謝謝佳宏學長解答撰寫論文時遇到的問題，謝謝王世傑學長、黃鈺峰學弟及王帥文學弟在撰寫論文時的幫忙，謝謝秘書敏霞與宜家助理幫忙我處理各種瑣碎的事情，以及陪伴我渡過三年歲月的其他同學、學長姐和學弟妹，感謝有你們的參與，讓我的研究所生涯更加多彩多姿。

最後，感謝政治大學圖書館提供 TAQ 數據以茲研究，讓我可以順利完成這篇論文，以及感謝栽培我的母校 – 政治大學，一個讓我成長與茁壯的地方，在未來人生的大道上不會迷惘。

國立政治大學 National Chengchi University

Abstract

Based on the observations that the mean square log-return obtained from the high-frequency one-day moving averages(HF1MA) of a collection of stocks share similar features with the mean square displacement of a many particle system described by Langevin equation, we carry out a detailed analysis on the time-scale dependence of static as well as dynamic properties for such averages. We analyze the data of a collection of 345 stocks listed in S&P 500 for each month over the years 1996-1999. We examine if the probability distribution meets Lévy distribution in two cases of moving average & non-moving average, and how the selected interval affect the fitted parameters of the probability distribution. Also we calculate the autocorrelation function and analyze the probability density function of log - return at different time scales in detail. Our results show that the central parts of probability density functions are fitted by Lévy with parameter $\alpha \approx 1$ for the averaged data and $\alpha \approx 1.6$ for the non-averaged data. With a newly defined autocorrelation function, we can distinguish dynamic features between the averaged data and the non-averaged data.

國立政治大學
National Chengchi University

Keywords : Langevin equation, Lévy distribution, autocorrelation function, random walk, Brownian motion, diffusion constant.

摘要

本研究著重於隨機漫步的理論與應用，並收集 S&P500 的其中 345 家交易較為頻繁的公司做為實證的數據。根據高頻率交易一天移動平均 (HF1MA)之下的股票觀測其特徵，發現與多粒子系統的均方位移(MSD)的特徵有相似之處，據此，我們進一步對在不同時間尺度靜態和動態屬性進行了詳細的分析。我們在分析 S＆P 500 其中 345 家公司在 1996 – 1999 年各月份的股票數據時，觀察作移動平均的計算對數據的統計分布與動態性質之影響。我們檢驗在有移動平均與沒有移動平均的兩種情況下，市場報酬(log–return)的機率密度函數中心是否符合 Lévy 分布，分析對單月資料進行統計計算之侷限與技巧，同時我們計算自相關函數並對報酬的機率密度函數如何隨時間尺度的變動進行詳細分析。結果顯示在一天的移動平均下，機率密度函數的中心部份符合 Lévy 分布，其 $\alpha \approx 1$；而在沒有一天移動平均下其 $\alpha \approx 1.6$。在新定義的自相關函數中，我們可以分辨在有移動平均與沒有移動平均的情況下其動力性質的特徵。

關鍵字 ： 朗之萬方程，Lévy 穩定分布，自相關函數，隨機行走，布朗運動，擴散係數。

國立政治大學 National Chengchi University

目錄

誌謝.....I

Abstract.....II

中文摘要.....III

圖表目錄.....1

第一章　序論.....3

第二章　理論背景與方法.....7

2.1 布朗運動 & 擴散運動 & 朗之萬方程式.....7

2.1.1 隨機行走(random walk)與布朗運動.....7

2.1.2 布朗運動：簡單的一維隨機行走(random walk)問題.....8

2.1.3 布朗運動：愛因斯坦以機率與擴散的觀點.....10

2.1.4 擴散方程式的解(solution of diffusion equation).....11

2.2 朗之萬方程(Langevin equation)描述布朗運動.....13

2.2.1 朗之萬方程針對速度的描述.....13

2.2.2 朗之萬方程針對 MSD 的描述.....14

2.3 穩定分布(stable distribution).....16

2.4 截尾 Lévy 機率分布(The Truncated Lévy Flight).....18

2.5 自相關函數.....19

第三章　實證分析.....21

3.1 高頻一天移動平均線的定義(definition of High Frequency One-day Moving Average, HF1MA) 21

3.2 (對數)報酬 log–return R(t) 的定義與不同的交易 time scale τ.....23

3.3 股價指數報酬對應於朗之萬理論的實證.....24

3.3.1 隨機行走的模擬.....24

3.3.2 MSLR的意義.....27

3.3.3 HF1MA 與 noMA 情況下的 MSLR 比較.....28

國立政治大學 National Chengchi University

3.3.4 不同時間長度移動平均之下對 MSLR 的影響.....31
3.4 股價指數報酬對應於 Lévy 機率分布的實證.....34
3.4.1 標準普爾指數 (S&P500) 報酬分布的尺度特徵.....36
3.4.2 股價報酬機率密度函數的指數遞減.....39
3.4.3 Lévy 穩定分布中心附近 P(R=0)的 α 值與圖形 fitting.....40
3.4.4 以Lévy穩定分布描述股價報酬P(R=0).....45
3.5 自相關函數.....50
第四章 結論與建議.....54
附錄.....60
參考文獻.....69

國立政治大學
National Chengchi University

圖表目錄

圖 3.1　1996 年 1 月 APPLE 公司的股價變動趨勢圖.....22

圖 3.2 資料選取的方式：Overlapping & non-overlapping.....23

圖 3.3.1A 隨機亂數模擬隨機行走(不同組數之比較).....25

圖 3.3.1B 隨機亂數模擬隨機行走(不同時間之比較).....25

圖 3.3.1C 隨機亂數模擬隨機行走(主曲線之 fitting).....26

圖 3.3.3A　1996 年 1 月 noMA 在所有可能 τ 之下所得到的 MSLR 分布圖.....29

圖 3.3.3B　1996 年 1 月 HF1MA 在所有可能 τ 之下所得到的 MSLR 分布圖.....30

圖 3.3.3C　HF1MA 與 noMA 的圖形比較.....30

圖 3.3.4A　HF1MA 在不同時間尺度下的圖形比較.....32

圖 3.3.4B　不同移動平均的時間長度下的圖形比較.....32

圖 3.3.4C　1996 年連續每 4 週的資料在 HF1MA 下的 MSLR.....33

圖 3.3.4D　1996 年連續每 4 週的資料在 HF1MA 下的所得到的擴散係數與溫度.....34

圖 3.4.1A　1996 年 2 月利用法 1 在不同時間尺度下的股價指數報酬機率密度分布.....37

圖 3.4.1B　1996 年 2 月利用法 2 在不同時間尺度下的股價指數報酬機率密度分布.....38

圖 3.4.2　1996 年 1 月標準化報酬的機率密度分布與厚尾性質.....39

圖 3.4.3A　Lévy 穩定分布 P(R=0)的 α 值利用方法 1 在 HF1MA 計算所得到的圖形.....41

圖 3.4.3B　Lévy 穩定分布 P(R=0)的 α 值利用方法 2 在 1hrMA 計算所得到的圖形.....41

圖 3.4.3C　Lévy 穩定分布 P(R=0)的 α 值利用方法 1 在 noMA 計算所得到的圖形.....42

國立政治大學 National Chengchi University

圖 3.4.3D　Lévy 穩定分布 P(R=0)的 α 值利用方法 2 在 noMA 計算所得到的圖形.....42
圖 3.4.3E　1996 年 2 月在經由 HF1MA(或 1hrMA)與 noMA 的方法後所得到的 α 值計算後的 rescale 圖形.....43
圖 3.4.4A　1996 年 2 月在 HF1MA 的情況下利用方法 1 所得最小 τ 的報酬分布與公式在線性與半對數的圖形.....45
圖 3.4.4B　1996 年 2 月在 HF1hrMA 的情況下利用方法 1 所得最小 τ 的報酬分布與公式在線性與半對數的圖形.....46
圖 3.4.4C　1996 年一整年在 HF1hrMA 的情況下利用方法 2 所得最小 τ 的報酬分布與公式在線性與半對數的圖形.....46
圖 3.4.4D　1996 年 2 月 noMA 的情況下利用方法 2 所得最小 τ 的報酬分布與公式在線性與半對數的圖形.....47
圖 3.4.4E　5 家公司的在 HF1MA 線性及半對數尺度下與公式的分布圖.....48
圖 3.5A　1996 年 2 月在 HF1MA 情況下不同時間尺度下的自相關函數圖形.....51
圖 3.5B　1996 年 2 月在 noMA 情況下不同時間尺度下的自相關函數圖形.....52
圖 3.5C　1996 年 5 月在 noMA 下兩種方法的圖形比較.....53
圖 4A　溫度對擴散係數的關係圖.....55
圖 4B　溫度對 volatility 與溫度對 log-return 的關係圖.....55
表 3.3.2　3 家公司在一週內的股價.....27
表 3.3.3　HF1MA & noMA 的 MSLR 之比較.....31
表 3.4　市場價格平均法.....35
表 3.4.3　Lévy 穩定分布在 MA 與 noMA 情況下的比較.....44
表 3.4.4　Lévy 穩定分布在 HF1MA 與 noMA 情況下的整體比較.....49

國立政治大學 National Chengchi University

第一章　序論

起初科學家利用隨機行走(random walk)的概念，用來描述股價與時間的相依性，股價的起伏不定，猶如一個醉漢的無序行走，或像布朗運動所觀察到的花粉不規則運動。在 1900 年間，Bachelier 提出金融市場的波動遵守隨機行走，其模式可應用到物理中的布朗運動[1]，在他後來的研究中，他給出一些關於布朗運動的函數之分布，而在隨後的科學家研究指出股市波動非完全隨機。

其後越來越多的物理學家積極參與經濟領域的研究，1966 年間，King 發現在一天的時間區間或更長的時間區間之下，不同股市的價格變化通常有高度相關，且在同種類型的公司相關性更高。對於金融市場，越來越多的資料提供市場的微觀細節，基本上它是一個非平衡的問題，其中動力性質扮演一個重要的角色。1977–1979 年間，Epps 研究 4 家公司對數報酬的相關係數，時間區間在 10 分鐘 ~ 3 天之間，發現有 Epps effect。短時間內，公司相關性比較小；長時間來看則相關性比較大。1995 年間，Mantegna 和 Stanley 的研究團隊，分析 S&P500 指數在尺度變換下的機率分布，發現中心分布符合所預期的 Lévy 穩定分布。

物理學家在經濟領域的研究累積了不少成就，由於這個領域是一個多種因素交互影響的學問，因此物理學家們應用物理上的理論，例如布朗運動、氣體分子動力論、描述粒子運動軌跡的朗之萬理論…等，並就原有的理論推廣延伸，對經濟領域提出一套假說，並積極地分析數據，期盼能透過大量的數據研究與歸納，為經濟領域找出並發現新的現象以及了解現象背後的原因，進一步找出合理的解釋並建構一套理論與數學模型，例如眾所週知的 B-S 模型對於期貨與選擇權的定價，就是由布朗運動理論推廣而得。

本論文研究的是標準普爾 500(S&P500)指數其中 345 家公司時間序列的統計性質，研究主要目標為觀察股價指數在經 HF1MA(一天的移動平均, 亦可寫作 1dayMA 或 1MA)計算的數據與未經處理(noMA)的原數據其所含訊息內容之異同，並跟先前研究的結果比較之間的差異，我們簡單分成 3 個部份來探討：

1. 類比股價與粒子的集體行為，從金融數據資料分析系統的巨觀狀態。

首先利用程式模擬隨機行走來驗證理論，並從探討朗之萬方程的原理出發，利用粒子的觀點來類比股價指數，我們已知當時間很短的情況下粒子在流體中的運動，其位移平方的平均值正比於時間的 2 次方；在長時間的情況下正比於時間的 1 次方，在經 HF1MA 計算的數據與原數據(noMA)又有一些不同的特徵出現，我們想藉由此方法來實證股票價格指數是否也有類似的情形發生。結果顯示 HF1MA 的方式是恰當的，因此我們確認股票與粒子間的集體特性是可比較，理由如下：

(1)當價格經過 HF1MA 之後計算所得到的數據(MSLR)在圖形上符合朗之萬方程的描述，在短時間 MSLR 正比於時間的 2 次方，時間夠長情況之下正比於時間的 1 次方。

(2)在 no MA 所得到的結果，時間夠長情況之下符合方程式的描述，在短時間並不符合描述，呈現所謂的粒子受限現象。我們可解讀為在此圖形中，這段時間公司彼此開始相互影響，可視為流體中粒子與粒子之間開始作用，因粒子之間的關聯性交互影響造成運動上有遲滯的現象發生，並與其所在的環境有關，所以在對數圖形中其值並不固定；在長時間時受限因素不存在後復歸為擴散運動的範疇(在對數圖形下斜率等於 1 的部份)，即正比於時間的 1 次方。

(3)另一方面，移動平均(Moving Average, MA)的時間長度，也會影響 MSLR 在短時間之下在對數圖形所對應的斜率，我們發現 1hrMA(1 小時的移動平均)與 1dayMA(1 天的移動平均)的斜率不一樣，前者的斜率較後者斜率小(後者斜率為 2)，且移動平均的時間長度越長，其斜率越接近 2，我們也發現其在轉折之處(斜率從 2 變為 1 時)與所選取的移動平均時間有關，若選取的 MA 時間為 1 天，則在 1 天附近開始出現轉折。

2.在 HF1MA 與 noMA 的條件之下，時間尺度(time scale)如何影響股市價格統計分布。

根據先前研究[2]，S&P500 股價機率分布有時間尺度行為，指數報酬在零附近

的機率密度函數值$P(R=0)$，與時間尺度呈現冪次函數遞減：$P(R=0)\sim\tau^{-1/\alpha}$。並且我們知道在$P(R=0)$附近可以用 Lévy 機率分布恰當地描述。據此，我們想探討在 HF1MA 與 noMA 的條件之下，股價機率分布是否也有雷同的時間尺度特徵，或者呈現截然不同的結果？結果顯示：

(1)時間序列長短會影響統計性，時間越長統計性越好，在做機率分布圖時，X 軸的尺度大小也會影響統計性，所以計算的結果會侷限在一個範圍內。

(2)我們將各家公司的股價指數類比粒子的集體行為，在有無移動平均的統計分析下，計算在作機率密度分布時不同的 X 軸尺度大小下 α 值的變化，結果發現在移動平均之下得到的 α 值接近 1，沒有移動平均之下的 α 值則介於 1.5~1.6 之間，符合穩定分布 (stable distribution) 的描述。

(3) 我們也同樣發現移動平均的時間長短也會影響 α 值的大小，移動平均的時間長度越短則越接近 noMA 的 α 值，因此，我們可藉由這項特性了解 S&P500 股價指數的機率分布。另外，我們已知報酬的機率密度有厚尾分布的特徵，那麼在 HF1MA 作用之下，我們也可看出其斜率近似於 3，符合先前研究[3][4]。

3.股價報酬的自我相關性，移動平均如何影響函數的分布。

股價報酬的自我相關係數隨著時間越大越接近零，意即沒有時間記憶性 (time memoryless)，由先前研究[4]可以知道，股價報酬的自相關函數圖在短時間來看有相關性，特徵時間為 4 分鐘，隨著時間越長其記憶性越差，最後在 20 分鐘左右趨於紊亂。經由HF1MA與noMA的情況下，我們發現有著一些不同的特徵：

(1)我們建立一個新的自相關函數，稱為”時間尺度選取下的自相關函數”(Time Scale Selected Autocorrelation, TSSA)，來表示在不同的時間尺度交易下的自相關函數其分布的異同。

(2) 接著我們再加上移動平均的條件，發現在TSSA的公式下所呈現的 HF1MA 與noMA函數分布圖截然不同，在HF1MA的情況下呈現連續軌跡的特徵，且其轉折與所選取的移動平均的時間尺度 τ 有關。在noMA的情況下呈現離散軌跡的特

徵，其最低點的時間即為所選取的時間尺度 τ，接著回到零附近振盪，隨著時間越大慢慢收斂到零。由於圖形的不同，顯示在TSSA公式下作HF1MA對動力性質有影響。

國立政治大學
National Chengchi University

第二章　理論背景與方法

本章節將介紹一些統計理論，包括穩定分布，TLF，自相關函數，以及簡短介紹物理上的理論包括布朗運動與朗之萬方程。

2.1 布朗運動 & 擴散運動 & 朗之萬方程式

這一小節我們簡單的介紹在 20 世紀初文獻上對於布朗運動的理解與分析方式。random walk 與朗之萬運動方式的概念應用在許多科學的領域，包含：物理、化學、生物、經濟、大氣、材料等學門，其影響十分深遠。

2.1.1 隨機行走 (random walk) 與布朗運動

物理學的理論應用於財務分析，最早可追溯至 1900 年法國物理學家 Louis Bachelier。他提出一個劃時代的主張，股票、債券的價格波動以及它們的基本因素結構都是隨機變化，類似物理學中的布朗運動。因此，Bachelier 將物理學概念發展成經濟模型應用於股市分析，利用布朗運動的數學性質與算式來描述價格波動的隨機變化，同時建構出隨機漫步理論，利用隨機漫步去描述價格走勢。此一想法，在財務學界延用至今，對經濟學有很大的貢獻。

Bachelier 的隨機漫步理論引伸到價格波動的變化可以鐘型曲線作描述[5]。鐘型曲線亦被稱為「高斯曲線」，在「高斯曲線」之下，各情況出現的機率呈現稱之為常態分布(normal distribution)，物件都傾向整體中的典型靠攏，比較異常的情況偶而出現，但一定是佔少數，而且終會走向均衡。採用隨機漫步理論的還有現代投資組合理論 Modern Portfolio Theory、Black-Scholes 選擇權定價模型以及 Long-Term Capital Management (LTCM)。

在 1827 年，英國的植物學家 Robert Brown 在顯微鏡下觀察到懸浮在水中的花粉粒子，會不停地進行不規則的運動，並從花粉與無機膠體粒子的實驗證實了小粒子在水中的隨機行走(random walk) 並非生物體特有的現象，這種類似生命體

的運動特徵引發科學家們研究微小粒子的運動行為[6]。經過許多的實驗與探討，科學家發現這現象應該是微小粒子受到周圍液體分子從各處的連續撞擊，而產生連續且不規則地隨機行走。這種移動我們稱之為布朗運動(Brownian motion) ，是室溫下微小的膠體粒子在溶液中的共同現象。在這之後人們才真正開始從物理的角度來研究布朗運動[7]。

布朗運動具有下列的特性：(1)粒子的運動永不停止；(2)溫度的改變會影響粒子的運動；(3)粒子的運動沒有固定的軌跡，其運動軌跡呈鋸齒狀；(4)粒子的大小影響粒子的運動速度；(5)粒子的成份或密度不會影響粒子的運動[8]。

2.1.2 布朗運動：簡單的一維隨機行走 (random walk) 問題

首先我們考慮一個簡單的問題，在經過一段長時間之後，布朗粒子離開起始位置後的平均位置 (位移期望值) 與均方位移 (Mean Square Displacement)。假設布朗粒子有以下的特徵：

(1) 粒子忽左忽右，猶如一個醉漢走路之前前後後、飄忽不定，此特性正好可以用一維的隨機行走 (random walk) 來描述與討論。

(2) 布朗粒子向左與向右移動的機率 p 與 q 是相等的，即 p = q = 1/2。

(3) 設布朗粒子從 x = 0 處出發，共走了 N 步，其中向右(沿正 x 軸正向)走了 n 步的機率可由二項分配的公式給出機率分布。

在此先簡單介紹二項分配，二項分配的定義如下：

$$RW_N(n) = \begin{cases} \binom{N}{n} p^n \mathrm{q}^{N-n}, & n = 0, 1, 2, \ldots, N \\ 0 \quad , & others \end{cases} \qquad (2.1.2A)$$

$$0 < p < 1 \ , q = 1 - p$$

上式為間斷型隨機變數 n 的機率密度，表示布朗粒子其中一次隨機行走試驗結果的機率，二項分配有如下的性質：

(1) 此試驗由 N 次獨立且完全相同之試驗所組成。

(2) 每次試驗僅有兩種互斥的可能結果稱為”YES”或”NO” 。

(3) 試驗結果為”YES”的機率為 P $(0 < p < 1)$，”NO”的機率為 q，且每次試驗機率 P 為固定不變的。

(4) 隨機變數 n 表示 N 次獨立試驗中”YES”的次數。

假設每一步的長度都是 l，並且向左為 $-l$，向右為 l。當走了 N 步時，其中有 n 步是向右，則某次試驗結果布朗粒子離原點的距離為：$x = [n - (N - n)]l$ 且平均距離為 (n=0~N的所有可能性相加)：

$$\langle x \rangle = \sum_{n=0}^{N} RW_N(n) \cdot [n - (N - n)]l = \sum_{n=0}^{N} RW_N(n) \cdot (2n - N)l \quad (2.1.2B)$$

由二項式定理，

$$(p + q)^N = \sum_{n=0}^{N} RW_N(n) = \sum_{n=0}^{N} \binom{N}{n} p^n \mathrm{q}^{N-n}$$

所以

$$\langle x(N) \rangle = \sum_{n=0}^{N} (2n - N)l \cdot RW_N(n) = \sum_{n=0}^{N} (2n - N)l \cdot \binom{N}{n} p^n \mathrm{q}^{N-n}$$

展開整理後[附錄二]可得結果為

$$\langle x(N) \rangle = 2pNl - Nl = Nl(2p - 1) = Nl(p - q) \qquad (2.1.2\mathrm{C})$$

對於布朗粒子的隨機行走，有 $\mathrm{p} = \mathrm{q} = \frac{1}{2}$，故可得 $\langle x(N) \rangle = 0$。

對於方均位移 $\langle x^2 \rangle$，有如下形式

$$\langle x^2 \rangle = \sum_{n=0}^{N} x^2 \cdot RW_N(n) = \sum_{n=0}^{N} (2n - N)^2 l^2 \frac{N!}{n!\,(N - n)!} p^n \mathrm{q}^{N-n}$$

$$= \sum_{n=0}^{N} (4n^2 - 4nl + N^2) l^2 \frac{N!}{n!\,(N - n)!} p^n \mathrm{q}^{N-n} \qquad (2.1.2D)$$

再利用上述的方式計算後可得結果為

$$\langle x^2 \rangle = Nl^2[(N - 1)(p - q)^2 + 1] \qquad (2.1.2E)$$

可知當 $p \neq q$ 時，布朗粒子會沿著某一個方向漸行漸遠；當在隨機行走的實驗中， $p = q = \frac{1}{2}$ 時，可得 $\langle x^2 \rangle = Nl^2$。又由於步數 N 正比於試驗所用的總時間 t，有 $x^2 \propto t$ 的關係，表示布朗粒子的運動顯然不是均速漂移，且 x^2 與 t 成正比是隨機過程的典型結果。此外，可以進一步求得

$$Var(x) = \langle x^2 \rangle - \langle x \rangle^2 = Nl^2[(p-q)^2 + 1] - [Nl(p-q)]^2 = Nl^2$$

即布朗粒子 N 步隨機行走之後位置的標準差 $\sigma \propto \sqrt{N}$，此為其可到達之最遠可能距離。

2.1.3 布朗運動 ： 愛因斯坦以機率與擴散的觀點

西元1906年，愛因斯坦在發表狹義相對論後，為了要解釋布朗粒子運動的方式，也發表了他以『機率』的觀念探討布朗運動的定量結果，提出了 random walk 的模型去描述此種形式的運動，為其運動形式提出了物理的解釋：布朗粒子的不規則運動是由於環境水分子不間斷地隨機的碰撞所造成。當存在大量的布朗粒子，其密度分布不均勻時，可觀察到布朗粒子的擴散。根據愛因斯坦的研究分析，粒子的運動雖然不規則，但是布朗運動在長時間下的平均移動行為會呈現常態分布，可視為布朗粒子的擴散行為。擴散實際上是粒子作布朗運動而產生位移，現在以擴散的觀點研究粒子的布朗運動。

在一維中以 $\phi(x,t)$ 表布朗顆粒的密度，以 $J(x,t)$ 表布朗顆粒的流量(單位時間內通過單位截面的顆粒數)。菲克定律[附錄五]可導出 diffusion equation

$$J = -D\nabla\phi$$

D 是擴散係數，連續方程式是

$$\frac{\partial \phi}{\partial t} + D \cdot J = 0$$

兩式聯立，得

$$\frac{\partial \phi}{\partial t} = D\nabla^2\phi$$

此為擴散方程式。三維空間下可寫為

$$\frac{\partial n(\vec{r},t)}{\partial t} = D\nabla^2 n(\vec{r},t)$$

其中 $n(\vec{r},t)$ 代表時間 t 於位置 $\vec{r}$ 的布朗粒子數目密度。

2.1.4 擴散方程式的解(solution of diffusion equation)

愛因斯坦是第一位以定量理論詮釋布朗運動，稍後史摩勒丘司基也發表以機率平衡方程式描述布朗運動。依據布朗粒子在時間 t 與位置 x 時的機率 P(x, t)，在每個時間的小間隔 Δt，物體可移動一步 Δl，而其方向則是隨機選取。在無外力作用下，一維空間中的布朗運動可寫成如下形式：

$$\frac{\partial \mathrm{P}}{\partial t} = \mathrm{D}\frac{\partial^2 P}{\partial x^2}$$

若 D 為常數，則此式為線性微分方程(i.e., $\partial(\mathrm{a}\emptyset) = a(\partial\emptyset)$，$\partial(\psi + \emptyset) = \partial\psi + \partial\emptyset$)，其解可為平面波形式：

$$\mathrm{P(x,t)} = e^{ikx}\tilde{P}(k,t)$$

將上式帶入擴散方程式，可得(the Fourier transform method)

$$\frac{\partial \tilde{P}(k,t)}{\partial t} = -\mathrm{D}k^2\tilde{P}(k,t)$$

其解為

$$\tilde{P}(k,t) = \tilde{P}(k,0)e^{-\mathrm{D}k^2 t}$$

利用 inverse Fourier transform，可得一般解:

$$P(x,t) = \frac{1}{2\pi}\int dk\, e^{ikx}\,\tilde{P}(k,t) = \frac{1}{2\pi}\int dk\, e^{ikx}\cdot e^{-\mathrm{D}k^2 t}\cdot\tilde{P}(k,0)$$

利用初始條件：當 t = 0 時，所有粒子均位在 x = 0 處，即

$$\mathrm{n(x,0)} = N\delta(x-0) \quad , \quad P(x,0) = \delta(x-0)$$

則擴散方程在初始條件下的解為

$$\tilde{P}(k,0) = \int dx\, e^{-ikx}\,P(x,0) = \int dx\, e^{-ikx}\,\delta(x) = 1$$

$$\Longrightarrow P(x,t)=\frac{1}{2\pi}\int dk\, e^{ikx}\, e^{-\mathrm{D}k^2 t}=\frac{1}{\sqrt{4\pi Dt}}e^{\frac{-x^2}{4Dt}} \quad (2.1.4\mathrm{A})$$

上式表明，顆粒的機率密度分布是與 t 有關的高斯分布，隨著 t 的增加，顆粒逐漸向兩邊擴散。經由簡單的計算，我們也可以得到長時間三維空間下的位置分布：$P(\vec{r},t)=\frac{1}{(4\pi Dt)^{3/2}}e^{-\frac{r^2}{4Dt}}$，擴散係數 $D={(\Delta l)^2}/{2\Delta t}$。

由 (2.1.4A) 式可求得粒子位移平方的平均值[附錄三]

$$\langle x^2\rangle=\int_{-\infty}^{+\infty}x^2\,P(x,t)dx=2Dt \qquad (2.1.4\mathrm{B})$$

上式亦可稱為 Einstein – Smoluchowski Equation。接著我們可以求出平均位移與標準差：

平均位移(The mean displacement)

$$\langle x\rangle_t=\int dx\cdot x\cdot P(x,t)=0$$

標準差(The standard deviation)

$$\sigma_t=\sqrt{\langle x^2\rangle_t-\langle x\rangle^2{}_t}=\sqrt{2Dt}，D=\frac{(x)^2}{2\Delta t}，其中\ \sigma(t)\propto\sqrt{\frac{t}{\Delta t}}\propto\sqrt{N}$$

t 為總時間，Δt 為每一步的時間間隔，$\frac{t}{\Delta t}$ 即為走 N 步之意。

由此可得均方位移 (The square mean displacement) 的另一表示式：

$$\sqrt{\langle x^2\rangle}\propto\sqrt{N}$$

(2.1.4B)式與朗之萬理論的結果 $\langle x^2\rangle=\frac{2k_BT}{\zeta}t$ ，兩式比較可得，溫度為 T 時顆粒在黏滯阻力係數為 ζ 的介質中的擴散係數為 $\mathrm{D}=\frac{k_BT}{\zeta}$。

由以上討論我們可得到兩個重要的結論：

(1) 粒子的位移平均為零(即$\langle x\rangle=0$)。藉由位移是一個向量的特性，以及布朗運動時粒子移動的等向性，重覆將布朗粒子從原點釋出的實驗，可以得到平均位移為零的結果，圖形呈鐘形分布。

(2) 粒子隨著時間往各個方向運動而遠離原點，將粒子的移動距離先取平方，再取平均，我們將發現位移平方的平均與所經過的時間 t 成正比；在同一時間條

件下，布朗粒子遠離原點的快慢則代表布朗擴散係數 D。

2.2 朗之萬方程 (Langevin equation) 描述布朗運動

在西元 1908 年，郎之萬 (Paul Langevin) 發表了可描述單一布朗粒子運動軌跡的方程，我們現在稱之為『朗之萬』方程式 (Langevin Equation)。

朗之萬 (Langevin) 是依據牛頓第二定律($m\frac{d^2x}{dt^2}=\sum F$)，考慮一個布朗粒子在運動時，同時受到流體的阻力 $m\zeta\left(\frac{dx}{dt}\right)$。與流體分子因熱運動與其碰撞的熱擾動力 $R(t)$。其具體的運動方程式 (Langevin equation) 如下：

$$m\frac{d^2x}{dt^2}=-\zeta v+F(t)+F'(t) \qquad (2.2)$$

在 2.2 式中，m 代表布朗粒子的質量，$-\zeta v=-\zeta\frac{dx}{dt}$ 為粘滯阻力，$\frac{dx}{dt}$ 為 Brownian particle 的速度，根據斯托克斯 (Stokes) 定理的描述，粘滯阻力來自介質分子對顆粒的碰撞，阻力以–ζv 表示，將顆粒看作半徑為 r 的小球，以緩慢的速度 v 在粘滯係數為 η 的流體中運動，此時阻力的大小與顆粒的速度成正比，則有 $\zeta=6\pi\eta r$，ζ 為阻力的比例常數。$F(t)$為隨機作用力(random force)，$F'(t)$為外力。當不存在其它外力時 $F'(t)=0$，朗之萬方程為

$$m\frac{d^2x}{dt^2}=-\zeta\frac{dx}{dt}+F(t) \qquad (2.2')$$

2.2.1 朗之萬方程針對速度的描述

由(2.2)式，經過簡單的計算，我們可得到布朗粒子的運動速度：

$$\vec{v}(t)=\vec{v}(0)e^{-t/\tau}+\int_0^t du\, e^{-(t-u)/\tau}\frac{\vec{F}(u)}{m}$$

其中 τ 代表 relaxation time。右式第一項為布朗粒子的起始速度隨時間呈指數衰減，而右式第二項則表達 random force $\vec{F}$ 的影響。我們可以看出長時間之

後 $\langle v^2(t)\rangle \sim \frac{\tau}{2m^2}$。接著又從氣體分子動力論知道布朗粒子與環境達到熱平衡時所對應的 $\langle v^2(t)\rangle = {3k_BT}/{m}$，其中 k_B 為波茲曼常數，而 T 代表溫度。

2.2.2 朗之萬方程針對 MSD 的描述

由朗之萬過程的推導，對上式等號左右同乘以 x，得

$$m\frac{d^2x}{dt^2}x = -\zeta\frac{dx}{dt}x + xF(t)$$

由 $x\frac{d^2x}{dt^2} = x\ddot{x} = \frac{d}{dt}(x\dot{x}) - \dot{x}^2 = \frac{1}{2}\frac{d^2}{dt^2}x^2 - \dot{x}^2$ 代入上式

左式：$m\frac{d^2x}{dt^2}x = m\left[\frac{1}{2}\frac{d^2}{dt^2}x^2 - \dot{x}^2\right] = \frac{1}{2}m\frac{d^2}{dt^2}x^2 - m\dot{x}^2$

右式：$-\zeta\frac{dx}{dt}x + xF(t) = -\zeta\frac{1}{2}\frac{d}{dt}x^2 + xF(t)$

合併左右兩式得 $\frac{1}{2}m\frac{d^2}{dt^2}x^2 - m\dot{x}^2 = -\zeta\frac{1}{2}\frac{d}{dt}x^2 + xF(t)$

將上式對大量粒子求平均，

$$\langle\frac{1}{2}m\frac{d^2}{dt^2}x^2 - m\dot{x}^2\rangle = \langle -\zeta\frac{1}{2}\frac{d}{dt}x^2 + xF(t)\rangle$$

$$\Rightarrow \frac{1}{2}m\frac{d^2}{dt^2}\langle x^2\rangle - m\langle\dot{x}^2\rangle = -\zeta\frac{1}{2}\frac{d}{dt}\langle x^2\rangle + \langle xF(t)\rangle$$

並利用兩項基本條件：

1. 隨機項 $F(t)$ 與粒子所處位置 x 無關，所以 $\langle F(t)\cdot x\rangle = 0$。
2. 熱力學平衡時，能量均分定理，$\frac{1}{2}m\left\langle\left(\frac{dx}{dt}\right)^2\right\rangle = \frac{1}{2}k_BT$。

將上述兩條件帶入方程式

$$\frac{1}{2}m\frac{d^2}{dt^2}\langle x^2\rangle - m\langle\dot{x}^2\rangle = -\zeta\frac{1}{2}\frac{d}{dt}\langle x^2\rangle + \langle xF(t)\rangle$$

$$\Rightarrow \frac{1}{2}m\frac{d^2}{dt^2}\langle x^2\rangle + \zeta\frac{1}{2}\frac{d}{dt}\langle x^2\rangle - m\langle\dot{x}^2\rangle = 0$$

$$\Rightarrow \frac{1}{2}m\frac{d^2}{dt^2}\langle x^2\rangle + \zeta\frac{1}{2}\frac{d}{dt}\langle x^2\rangle - k_BT = 0$$

國立政治大學 National Chengchi University

$$\Rightarrow \frac{d^2}{dt^2}\langle x^2\rangle + \frac{\zeta}{m}\frac{d}{dt}\langle x^2\rangle - \frac{2k_BT}{m} = 0$$

上式是 $\langle x^2\rangle$ 的二階常係數線性非齊次微分方程，其通解為(設初始條件為：$t = 0$，$\mathrm{x} = 0$)

$$\frac{2k_BT}{m\zeta}t + \frac{2k_BT}{m\zeta^2}e^{-\zeta t} - \frac{2k_BT}{m\zeta^2} \qquad (2.2.2A)$$

由 (2.5.1) 式可再改寫為：

$$\langle x^2\rangle = \frac{2k_BT}{m\zeta}\left[t - \frac{1}{\zeta}\left(1 - e^{-\zeta t}\right)\right] \qquad (2.2.2B)$$

並利用泰勒展開式：

$$e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} \qquad \forall x \in \mathbb{R}$$

(1) 如果 $t \ll \zeta^{-1}$，則 (2.5.2) 式的指數項需要考慮，可改寫為

$$\langle x^2\rangle \approx \frac{2k_BT}{m\zeta}\left\{t - \frac{1}{\zeta}[1 - (1 - \zeta t + (-\zeta t)^2 + \cdots)]\right\}$$

$$\approx \frac{2k_BT}{m\zeta}\{\zeta t^2 + \cdots\} \approx \frac{2k_BT}{m}t^2$$

$$\Rightarrow \langle x^2\rangle \propto t^2$$

(2) 如果 $t \gg \zeta^{-1}$，則 (2.5.2) 式可改寫為

$$\langle x^2\rangle \approx \frac{2k_BT}{m\zeta}\{\,t\,\}$$

$$\Rightarrow \langle x^2\rangle \propto \mathrm{t}$$

(2) 的描述符合 2.1.1 節所提到，在長時間之下，與 t 成正比是隨機過程的典型結果，而前面 2.1.2 節也曾經先提到愛因斯坦的 random walk 模型，經由解出擴散方程得到其機率分布再經由機率密度計算亦可導出 $\langle x^2\rangle = 2Dt$，D為愛因斯坦擴散係數，這樣的關係式比較之後可得到擴散的相對應係數。因此我們得到下述關係：

$$\frac{2k_BT}{m\zeta}t = 2Dt \qquad \Rightarrow \qquad \frac{k_BT}{m\zeta} = D$$

此關係式被稱為 Einstein relation。

國立政治大學 National Chengchi University

2.3 穩定分布 (stable distribution)

在金融市場當中，不同的金融商品諸如股票、期貨、基金…等等，在不同交易時間尺度之下會有不同的價格變化，例如蘋果公司的股票價格，每日的價格變化會不同於每週的價格變化，所以如何有效地探討價格變化背後的意義，我們即採用在不同時間尺度下，討論各種價格變化的機率分布的方式。所以選擇何種機率分布來描述股票價格是非常重要的，以下我們使用穩定分布來討論不同時間尺度之下的股票價格報酬的機率分布。

穩定分布，又稱為 Lévy 穩定分布 (Lévy stable distribution)，是一種連續機率分布。[3]Lévy 和 Khintchine 解決了這個如何決定所有穩定分布應該遵守的函數型式的問題，他們發現最一般的表示式是經由特徵函數 $\varphi(q)$ 來表達，包含四個參數：α、β、γ 與 μ，其中參數 α 是描述機率密度函數的厚尾分布，可稱為特徵指數，範圍是$0 < \alpha \leq 2$；γ 是尺度因子(scale factor)且為正數，並與標準差有關；μ 為實數且為正數，與平均值有關，β 為不對稱參數(asymmetry parameter)，又稱偏態係數，範圍在 $[-1, 1]$ 之間，當分別為正值、0或負值時，分布分別右偏、對稱或左偏。

Lévy 穩定分布只有以下三種已知的解析解形式(analytic form)：

(1) α = ½, β = 1 (Lévy -Smirnov distribution)

(2) α = 1, β = 0 (Lorentz distribution)

(3) α = 2 (Gaussian distribution)

當特徵參數 α 從 2 降到 0 時，分布的尾部會變得越來越”肥”。特別是當參數 $\alpha = 2$ 且 $\beta = 0$ 時，特徵函數會變為常態(高斯)分布的特徵函數之型式 $\varphi(q) = e^{iuq-\gamma q^2}$，此時平均值為 u，標準差為 γ。因此，常態分布是 Lévy 分布的一個特例。

Lévy 分布有兩個基本的重要性質：

(1)**當 $\alpha < 2$ 時，Lévy 分布不存在變異數。**

當 $\alpha < 2$，$|x|$ 的 n 階動差函數 $(\mathrm{n} > \alpha)$ $|x|^n$ 會發散，即所有 $\alpha < 2$ 的 Lévy 穩定過程 (Lévy stable process) 的變異數會發散。

基於長時間的報酬會趨近於零並呈對稱分布的形式，我們考慮對稱分布 ($\beta = 0$) 以及平均值為 0 ($\mu = 0$) 的情況。在這情況之下的特徵函數可寫為如下形式：

$$\varphi(q) = e^{-\gamma|q|^\alpha} \quad (2.3A)$$

而機率分布是它的特徵函數的傅立業轉換，則參數為 α 及尺度因子為 γ 的穩定分布是

$$P_L(x) \equiv \frac{1}{\pi}\int_0^\infty e^{-\gamma|q|^\alpha}\cos(qx)\,dq$$

當 $x = 0$ 時，p_L可改寫為：

$$P_L(x = 0) = \frac{\Gamma(1/\alpha)}{\pi\alpha\gamma^{1/\alpha}}$$

此分布對於在 $|x|$ 很大的情況下可近似為：

$$P_L(|x|) \approx \frac{\gamma\Gamma(1+\alpha)\sin(\pi\alpha/2)}{\pi|x|^{1+\alpha}} \approx |x|^{-(1+\alpha)} \quad (2.3B)$$

事實上對於 x 很大時，P_L 可用冪次定律來描述，且此 stable Lévy processes 的變異數無限大。

(2) 穩定 Lévy 分布的可加性

若隨機變量 ξ、ξ_1 和 ξ_2 為獨立同分布(i.i.d)，對任意正數 b_1 和 b_2，存在正數 b，使得 $\mathrm{b}_1\xi_1 + \mathrm{b}_2\xi_2$ 和 $\mathrm{b}\xi$ 同分布，則稱隨機變量 ξ 具有穩定分布。假設一組隨機變數的分布是 Lévy stable distribution，機率密度函數的特徵函數如 $(2.3A)$ 式 $\varphi(q) = e^{-\gamma|q|^\alpha}$ 所定義，則將 n 組隨機變數相加，得到的另一個隨機變數 S_n，則 S_n 的機率密度函數的特徵函數：

$$\varphi(q) = e^{-n\gamma|q|^\alpha}$$

則將上式經由傅立業轉換，並找出 $S_n = 0$ 的機率密度函數值：

$$P(S_n, n) \equiv \frac{1}{\pi}\int_0^\infty e^{-n\gamma|q|^\alpha}\cos(qS_n)\,dq \xrightarrow{\mathrm{S_n=0}} P(S_n = 0) = \frac{\Gamma(1/\alpha)}{\pi\alpha(n\gamma)^{1/\alpha}} \quad (2.3C)$$

等號兩邊同時取對數可得

$$logP_n\,(S_n = 0) = log\left(\frac{\Gamma(1/\alpha)}{\pi\alpha\gamma^{1/\alpha}}\right) + (-1/\alpha)\,log\,n \qquad (2.3D)$$

由上式可以求得特徵指數 α。

在 1995 年 R.N.Mantegna 和 H.E.Stanley 發表於 Nature[2]的文章中，提到關於股價指數報酬機率密度函數時間尺度行為的研究，資料是 1984 年~1989 年 的 S&P500 股價指數，他們分析了報酬機率密度函數，其所選取的時間範圍是 1 min~1000 min，所得到的 α 值 約為 1.40，說明了

(1) 機率密度函數在靠近零 ($|Z/\sigma| \leq 6$) 的部分，可以用 Lévy stable distribution 來描述；

(2) 機率密度函數在$|Z/\sigma| \geq 6$的部分，則是趨近於 exponential 的遞減。

2.4 截尾Lévy機率分布 (The Truncated Levy Flight)

如果股價指數報酬的時間序列是 Lévy 穩定機率分布，在股價報酬的個數趨向於無限大時，會近似一個冪次函數，如(2.3B)式，而且變異數為無限大。為了避免二階動差無限大的這個問題產生，*Mantegna* et *al* 考慮一個有限變異數的隨機過程來描述這種機率分布，此隨機過程遵守尺度關係。根據 R. N. Mantegna 和 H. E. Stanley 在 1995 年的研究，美國 S&P500 股價指數的報酬機率密度分布，在機率分布中央為 Lévy 穩定機率分布，當股價指數報酬趨向無限大時，機率密度就隨之遞減，顯示變異數為有限值。他們稱這種機率分布叫做 *Truncated Lévy flight* **(TLF)**。

TLF 分布定義如下[3] [9]：

$$\mathrm{P}(x) = \begin{cases} 0 & x > l \\ cP_L(x) & -l < x < l \\ 0 & x < l \end{cases} \qquad (2.4A)$$

這裡的 $P_L(x)$ 是對稱的 Lévy distribution，c 為標準化的常數。

只有 Lévy 分布是穩定分布，**TLF** 分布是不穩定的，但它的變異數是有限值，所以它會收斂至高斯分布，但不像其他的機率分布，它的收斂性很小。

根據第2.1節的描述，如果隨機變數為穩定分布，相加所得到的隨機變數，其機率分布仍為穩定分布。R. N. Mantegna 和 H. E. Stanley 於1994的論文[9]，研究了 TLF 隨機變數相加後的機率分布情況，當 n 很小時，較接近 Lévy flight 的行為；當 n 很大時，較接近高斯行為，他們提出以下近似關係：$n_{\times}(l) \propto l^{\alpha}$，$n_{\times}$ 為 crossover value，經計算後可得：

$$n_{\times} \approx Al^{\alpha} \tag{2.4B}$$

以及

$$A = \left[\frac{\pi\alpha}{2\Gamma(\frac{1}{\alpha})\left[\frac{\Gamma(1+\alpha)\sin(\pi\alpha/2)}{(2-\alpha)}\right]^{1/2}}\right]^{2\alpha/(\alpha-2)} \tag{2.4C}$$

crossover value $n_{\times}$ 隨著截斷長度 l (cutoff length) 的增加而迅速地增加，如果 TLF 隨機變數相加個數在 $n_{\times}$ 個以內，機率密度函數可用 Lévy 穩定分布來描述，大於 $n_{\times}$ 個就可用高斯分布來描述。

2.5 自相關函數 (Autocorrelation)

在機率論和統計學中，相關（Correlation，或稱相關係數），顯示兩個隨機變數之間線性關係的強度和方向。在統計學中，相關的意義是用來衡量兩個變數相對於其相互獨立的距離。

對於不同數據特點，可以使用不同的係數。最常用的是皮爾森相關係數。其定義是兩個變數共變異數除以兩個變數的標準差（變異數的平方根）。

$$\rho_{\mathrm{X,Y}} = \frac{cov(\mathrm{X,Y})}{\sigma_X\sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X\sigma_Y}$$

我們可定義相關係數如下，一階自相關係數 $\boldsymbol{C_1}$ 為

$$C_1 = \frac{\sum_{\mathrm{t=1}}^{\mathrm{n-1}}(\mathrm{x_t}-\bar{x}_t)(\mathrm{x_{t+1}}-\bar{x}_{t+1})}{\sqrt{\sum_{\mathrm{t=1}}^{\mathrm{n-1}}(\mathrm{x_t}-\bar{x}_t)^2\cdot\sum_{\mathrm{t=1}}^{\mathrm{n-1}}(\mathrm{x_{t+1}}-\bar{x}_{t+1})^2}}$$

二階自相關係數 $\boldsymbol{C_2}$ 為

$$C_2 = \frac{\sum_{t=1}^{n-2}(x_t - \bar{x}_t)(x_{t+2} - \bar{x}_{t+2})}{\sqrt{\sum_{t=1}^{n-2}(x_t - \bar{x}_t)^2 \cdot \sum_{t=1}^{n-2}(x_{t+2} - \bar{x}_{t+2})^2}}$$

推廣到 k 階自相關係數 C_k 為

$$C_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x}_t)(x_{t+k} - \bar{x}_{t+k})}{\sqrt{\sum_{t=1}^{n-k}(x_t - \bar{x}_t)^2 \cdot \sum_{t=1}^{n-k}(x_{t+k} - \bar{x}_{t+k})^2}}$$

其中 C 介於 [-1,1]，$0 < C \leq 1$ 表示正相關，$-1 \leq C < 0$ 表示負相關，$C = \pm 1$ 代表完全正(負)相關，$C = 0$ 表示不相關。C 值越靠近 1 表示正相關性越高，越靠近 -1 表示負相關性越高，接近零則表示時間序列上的數值前後越不相關。

H. Eugene Stanley, Yanhui Liu 等人研究報酬的自我相關行為，在 1999 年發表的文章[14]指出，在開盤和收盤時間附近會有較活躍的交易為了消除這些干擾，他們找出所謂當天之內圖形(intraday pattern)，定義為

$$A(t_{day}) \equiv \frac{\sum_{j=1}^{N}\left|G^j(t_{day})\right|}{N}$$

其中，$G^j(t_{day})$ 為每天第 j 家公司在當天某個時刻的股價指數報酬，N 為選取時間序列的全部天數。對原本的資料重新定義後，得到

$$g(t) \equiv \frac{G(t_{day})}{A(t_{day})}$$

$G(t_{day})$ 為當日某時刻的報酬，$g(t)$ 為在 t 時刻標準化過後的價格變化的絕對值，藉此使自相關函數能消除在一天之中所受的外在影響。研究的結果顯示，在半小時內會呈指數式地衰減到零，其行式為

$$C(t) \sim e^{-\frac{t}{\tau}}$$

其中 τ 為特徵時間，約為 4 分鐘，接著遞減大約在 20 分鐘左右至零附近。而我們在討論自相關時所採用的 HF1MA 方式，目的也在於一天之內的移動平均下，可以消除日內所受的干擾。

第三章　實證分析

基於上述理論與統計性質，我們選取的研究對象為 S&P 500 股價指數當中，在 1996 年至 1999 年之間交易較為頻繁的 345 家公司[10]的股票價格，每筆價格為每隔 36 秒記錄一次交易，我們將其對不同的時間尺度下作對數報酬(價格取對數之後再相減)的計算，再以所得到的時間序列做分析與探討。

股市每天交易時間為 AM9:00~PM4:30，共 6.5 個小時(23400s) ，一年交易的時間有 240 天(5156000s)。股票交易視為連續，這裡只考慮股市交易的時候，移除晚上，週末或假日…等股市沒有交易的情形。也就是說，當次收盤與下一次開盤視為連續。

本章前兩節先定義移動平均和對數報酬，以及如何在資料上選取時間尺度，其餘小節則介紹所作的研究及成果，在 3.3 節由股價對數報酬計算 MSLR 來實證朗之萬理論的描述；在 3.4 節詳細探討如何應用 Lévy 機率分布的理論在股價報酬的分析；在 3.5 節則引用新的自相關函數討論有無移動平均的影響。

3.1 高頻一天移動平均線的定義 (definition of High frequency One-day moving average, HF1MA)

所謂高頻是指記錄一次交易的時間為 36 秒，在此我們利用股票技術分析上的「移動平均」(MA)的概念，本研究採用算數移動平均 (Arithmetic Moving Average, AMA)，指的是某序列之前 n 個數值的未作加權算術平均。例如，成交量 (Volume) 的 5 日簡單移動平均是指之前 5 日成交量的算術平均數，性質為圖形上的一點。假設 n 天的成交量為 v_1 至 v_n，則 n 日簡單移動平均方程式為：

$$\mathrm{AMA} = \frac{v_1 + v_2 + \cdots + v_n}{n}$$

所謂線是指由無數個點所構成，然而在實際股票數據上並不可能無限地分割，所以線的稠密性將視所選取資料的時間尺度有關。例如，在一段固定的時間裡，日移動平均比週移動平均有較多的點，也比較符合原指數跳動的趨勢。當我們連

續地計算出 AMA 的數值，並將其連接起來即可得到所謂的「移動平均線」，簡稱**均線**。其計算方法為，從一開始的 AMA 算起，標記為AMA_0，之後每隔一個相同的時間區間 ΔT (time interval) 就有一個新的數值加入，同時最舊的那個數值剔除，由此得到新的移動平均，標記為AMA_1：

$$\mathrm{AMA}_0 = \frac{v_1 + v_2 + \cdots + v_n}{n}$$

$$\mathrm{AMA}_1 = \frac{v_2 + v_3 + \cdots + v_{n+1}}{n}$$

$$= \frac{v_1 + v_2 + \cdots + v_n}{n} - \frac{v_1}{n} + \frac{v_{n+1}}{n}$$

$$= \mathrm{AMA}_0 + \frac{v_{n+1} - v_1}{n}$$

$$\mathrm{AMA}_i = \mathrm{AMA}_0 + \frac{\sum_{k=1}^{i} v_{n+k} - \sum_{k=1}^{i} v_k}{n}$$

再將AMA_1視為初始值，重覆以上式計算後可得各個時間尺度之下的 AMA。

利用上述方法，我們選取 S&P 500 的每個月的價格為總資料長度，由於資料的時間序列屬於高頻的特性，先對資料做 1 天 (即 650 個資料點) 的移動平均 (1MA 或 HF1MA)，經過 1MA 方式調整過後的資料較為平滑，資料點較原始資料點少 649 個，並且曲線可視為連續的路徑。

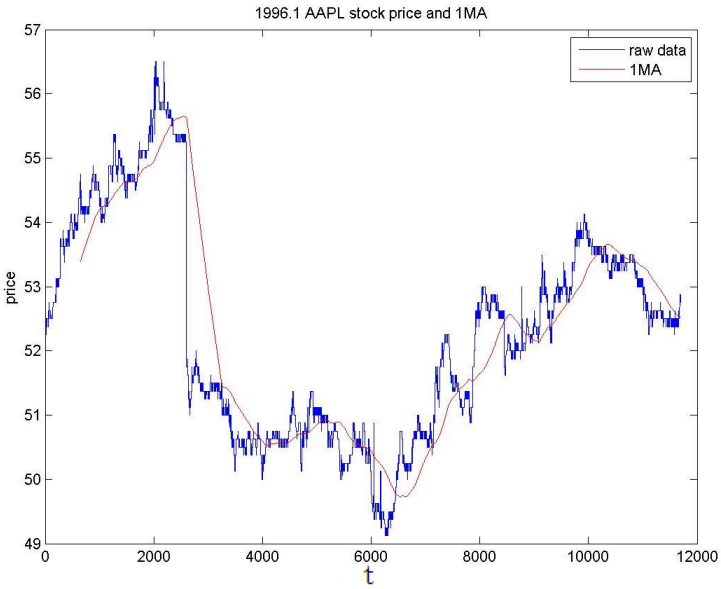


圖 3.1　1996 年 1 月 APPLE 公司的股價變動趨勢圖(藍線)以及經過 1 天的移動平均(紅線)，橫軸為記錄交易的時間間隔 t (36 秒/格)，縱軸為股價(USD)。

3.2 (對數)報酬 log–return R(t)的定義與不同的交易時間尺度 time scale τ

我們定義對數報酬 R(t)（或稱價格變化率），其形式如下：

$$R(\mathrm{t}) \equiv lnp(\mathrm{t}+\tau) - lnp(\mathrm{t}) \cong \frac{p(\mathrm{t}+\tau)-p(\mathrm{t})}{p(\mathrm{t})} \approx \frac{\Delta p}{p} \qquad (3.2A)$$

其中， p 為該時間點的價格指數，τ 為交易的時間尺度， $\frac{\Delta p}{p}$ 為報酬率的百分比。設 $\tau = 36\mathrm{s}$，最初開盤的指數為 p_0，$p_0 = p|_{\mathrm{t=0s}} = p(\mathrm{t}=0) = 1000$，則 36s 後的指數為 p_1，$p_1 = p|_{\mathrm{t=36s}} = p(\mathrm{t}=0+36\mathrm{s}) = 1001$。則取對數 (natural log) 後， $lnp_0 \cong 6.907755$ ； $lnp_1 \cong 6.908754$，$R_1 = R(0) = lnp_1 - lnp_0 = 0.000999$。

接著選取不同的交易時間尺度 τ，相當於隨機漫步中的步數，例如選取 $\tau = 3$ *min* $(60 \times 3 \div 36 = 5$ 個資料點)，代表依序每 3 分鐘 (5 步之後) 把取對數後的收盤價減去 3 分鐘前取對數後的開盤價，而得到一組對數報酬(log–return)。

另外，log – return 的計算皆採取 overlapping 的方式，如下圖所示：

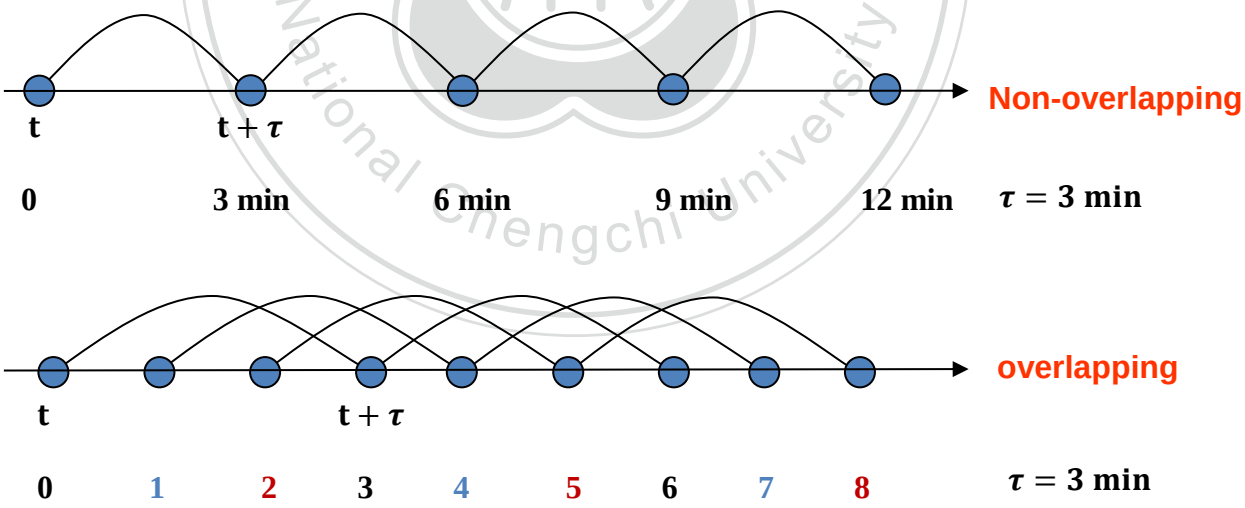


圖 3.2 資料選取的方式：Overlapping & non-overlapping

則考慮在不同交易時間尺度 (time scale τ) 之下的對數報酬，(3.2A)式可改寫為如下形式，其中 p 為該時間點的價格：

$$R_\tau(\mathrm{t}) \equiv lnp(t+\tau) - lnp(\mathrm{t}) \qquad (3.2B)$$

3.3 股價指數報酬對應於朗之萬理論的實證

這一節討論隨機行走與股票價格對應朗之萬方程的結果，在 2.1 節我們曾經提到朗之萬方程式的結果，其中位移平方的平均值(MSD)在長時間之下，正比於時間的二次方，在短時間之下則正比於時間的一次方。在股票市場中，由於價格變動類似隨機行走，所以我們猜測對於長時間之下股票價格也會有同樣的情形出現。利用朗之萬方程式的概念計算後發現，股價報酬的均方位移在長時間與短時間也正比於時間的二次方，符合理論描述之外，並發現無移動平均的報酬在短時間有著不同的特徵。我們也探討移動平均的時間長度如何影響圖形的分布，並經由公式[11]來找出一條主曲線(master curve)，將不同月份的 MSLR 做尺度變換後，發現除了少數月份之外，其餘月份皆可以主曲線來適當地描述。

3.3.1 隨機行走的模擬

本節簡單利用二項式定理的結果來模擬一維隨機行走的理論，我們知道在時間夠長的情況之下，如果選擇向左或向右行走的機率相同，其密度分布可用常態分布描述，形式為 $p(x)\sim N(\mu,\sigma^2)$，這裡分別選取不同組數 (相當於多少次試驗)與不同時間點(相當於不同的步數)的隨機亂數來模擬驗證，亂數值介於 0~1 之間，當亂數值大於 0.5 為向右 1 格(+1)，否則為向左 1 格(-1)。

(1) 100000 組與 10000 組在相同步數(時間)的比較

首先選取不同組數(試驗次數)計算隨機漫步並畫出公式，其公式應符合常態分布的描述，如下形式：

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.3.1A)$$

而由 2.1.4 節得知，隨機行走的機率形式為：

$$p(x,t)=\frac{1}{\sqrt{4\pi Dt}}e^{\frac{-x^2}{4Dt}} \quad (3.3.1B)$$

其中擴散係數 $D={(\Delta l)^2}/{2\Delta t}$，在此設定$\Delta l=1$(每次走一步)，$\Delta t=1$(走一步所

需時間為 1)，可得$D = \frac{1}{2}$。下圖為模擬的結果並選取最大步數 (最長的時間) 做比較，結果顯示在相同的步數(時間)下，組數越多越吻合隨機行走的分布(常態分布)。另一方面，我們也可驗證在相同的組數之下，所行走的步數越多圖形越低擴，因為步數相當於時間，在公式裡時間越長則圖形越低擴，由圖中可看出資料點恰當地落在公式曲線上。

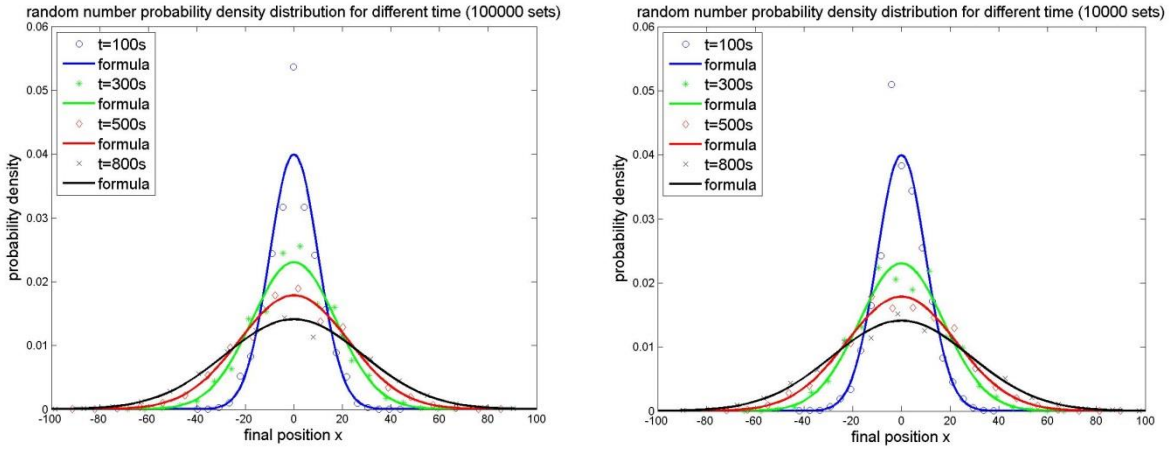


圖 3.3.1A　隨機亂數模擬隨機行走，結果顯示吻合隨機行走的機率分布。

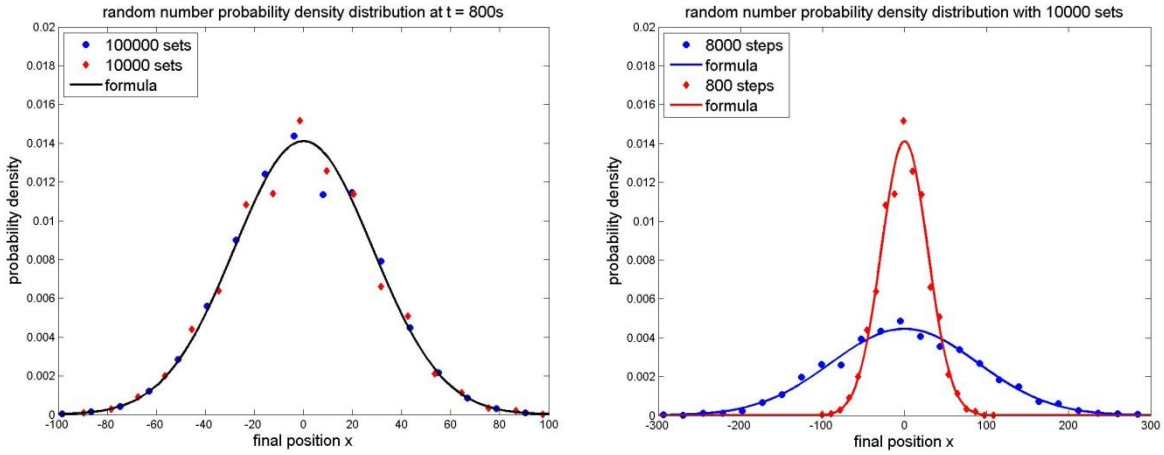


圖 3.3.1B　隨機亂數模擬隨機行走，左圖結果顯示在相同的步數(時間)下，組數越多越吻合隨機行走的分布，右圖結果顯示在相同的組數(試驗次數)下，步數越多越吻合隨機行走的分布

(2) 主曲線(master curve)的 fitting

我們也可以在不同時間的隨機行走找到一條主曲線，使各個時間點曲線消去時間的影響並適當地落在主曲線上，這裡採用方式是變數變換法，對於 3.3.1B 式，由於 $\int_{-\infty}^{\infty} p(x,t)\,dx = 1$，可以找到一個 $\int_{-\infty}^{\infty} \mathrm{A}\cdot P(\mathrm{X})\,d\mathrm{X} = 1$，其中 A 為歸一化常數。令 $\mathrm{X}^2 = \frac{x^2}{4Dt}$ $(D = \frac{1}{2})$，可得 $X = \frac{x}{\sqrt{4Dt}}$ 且 $dX = \frac{dx}{\sqrt{4Dt}}$，代入並計算後可得 $\mathrm{A} = \frac{1}{\sqrt{\pi}}$，則 $A\cdot P(X)$ 為主曲線其值為

$$\mathbf{P} = A\cdot P(X) = \frac{1}{\sqrt{\pi}} e^{-\mathrm{X}^2}$$

所以我們將原本的函數值做變數變換(即 $X = \frac{x}{\sqrt{4Dt}}$)之後再作圖，即可得到所有圖均恰當地分布在主曲線上。

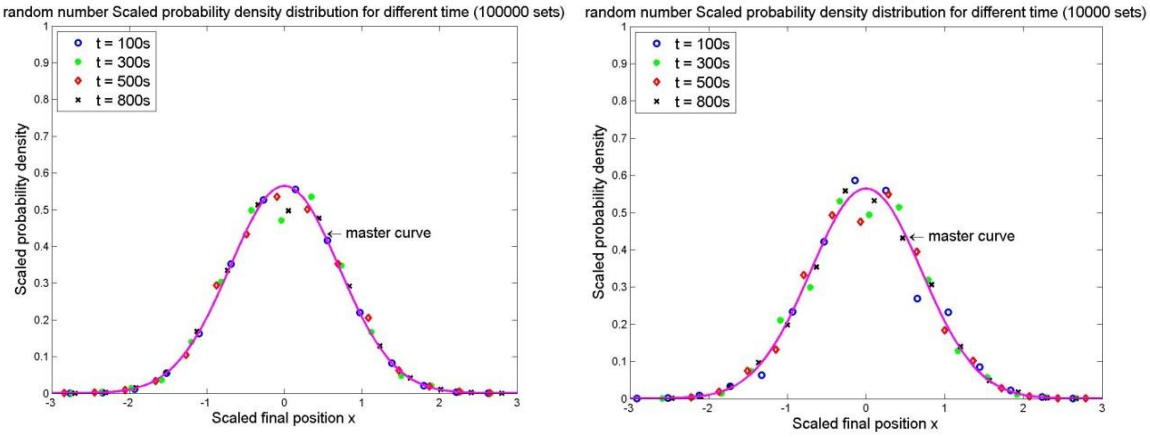


圖 3.3.1C　以隨機亂數模擬隨機行走，左圖為 100000 組隨機試驗次數，右圖為 10000 組隨機試驗次數，結果顯示兩個不同組數的隨機試驗皆落在主曲線上，且組數(試驗次數)越多分布越吻合。

3.3.2 MSLR 的意義

首先我們採用 3.2 節 overlapping 的計算方式以及 (3.2A) 式的定義：

$$R_{f,\tau}(t) \equiv ln\, p_{f,\tau}(t+\tau) - ln\, p_{f,\tau}(t) \cong \frac{\Delta\, p_{f,\tau}(t)}{p_{f,\tau}(t)}$$

式中，f 為正整數且 $1 \le f \le 345$，代表 345 家公司； t 為原始記錄交易的時間點，$t=1$ 代表 36 秒，$t=2$ 代表 72 秒，以此類推； τ 為不同的交易時間尺度。因為 $ln\, p_{f,\tau}(t+\tau) - ln\, p_{f,\tau}(t) \approx \frac{\Delta\, p_{f,\tau}(t)}{p_{f,\tau}(t)}$，則

$$p'_{f,\tau}(t) = \frac{ln\, p_{f,\tau}(t+\tau) - ln\, p_{f,\tau}(t)}{\Delta\, p_{f,\tau}(t)} \approx \frac{\Delta\, p_{f,\tau}(t)}{p_{f,\tau}(t)} \Longrightarrow p'_{f,\tau}(t) \approx \frac{\Delta\, p_{f,\tau}(t)}{p_{f,\tau}(t)} \quad (3.3.2A)$$

所以我們可以將股價的對數報酬視為價格的變化，即 $R(t) \approx p'_f(t)$，類比於朗之萬方程式中粒子位移的變化量。

$R_{f,\tau}(t)$ 可解釋為，第 f 家公司在交易時間尺度為 τ 的情況之下，所得到的各組各個交易時間點的報酬，所以當原始資料有 N 個資料點時，一家公司的報酬有 N-1 組(尾巴減去開頭)。接著對所有公司算出 $R_{f,\tau}(t)$ 再取平方，然後相加取平均，我們稱這一個數為均對數報酬(Mean Square Log Return, MSLR)，記為 $\langle r^2 \rangle$。按此步驟，對所有可能的 τ 都按上述方式計算，得到的整體市場的 MSLR 再畫出圖形。

以例子說明如下，假設市場上有 3 家公司，在一週內的股價如下表所述：

表 3.3.2　3 家公司在一週內的股價

price / firm	Monday	Tuesday	Wednesday	Thursday	Friday
A	10	11	12	13	14
B	10	9	8	7	6
C	4	7	10	7	4

假設 $\tau = 1$ 天，則 $R_{f,\tau}(t) = R_{f,1}(t) \equiv log\, p_{f,1}(t+1) - log\, p_{f,1}(t)$，可表示如下

	t = 1 (Tue-Mon)	t = 2 (Wed-Tue)	t = 3 (Thu-Wed)	t = 4 (Fri-Thu)
$R_{A,1}(t)$	$ln\,11 - ln\,10$	$ln\,12 - ln\,11$	$ln\,13 - ln\,12$	$ln\,14 - ln\,13$
$R_{B,1}(t)$	$ln\,9 - ln\,10$	$ln\,8 - ln\,9$	$ln\,7 - ln\,8$	$ln\,6 - ln\,7$
$R_{C,1}(t)$	$ln\,7 - ln\,4$	$ln\,10 - ln\,7$	$ln\,7 - ln\,10$	$ln\,4 - ln\,7$

取平方後各欄相加

	t = 1 (Tue-Mon)	t = 2 (Wed-Tue)	t = 3 (Thu-Wed)	t = 4 (Fri-Thu)
${R_{A,1}}^2(t)$	0.0414^2	0.0378^2	0.0348^2	0.0322^2
${R_{B,1}}^2(t)$	$(-0.0458)^2$	$(-0.0512)^2$	$(-0.0580)^2$	$(-0.0669)^2$
${R_{C,1}}^2(t)$	$(0.2430)^2$	$(0.1549)^2$	$(-0.1549)^2$	$(-0.2430)^2$
total	甲	乙	丙	丁

最後再把 甲乙丙丁 加起來取平均，可以得到在 $\tau = 1$ day 的 $\langle r^2 \rangle_{\tau=1}$。以此方式對所有可能的 τ ($\tau = 1, 2, 3, 4$ day)，都找出對應的 $\langle r^2 \rangle$，然後畫出圖形。現在我們對 S&P500 中的 345 家公司做上述分析，時間長度為一個月，原始資料的交易時間間隔為每 36 秒記錄一次。

3.3.3 HF1MA 與 noMA 情況下的 MSLR 比較

首先我們不對原始資料作移動平均(noMA)，按照上述步驟得到圖形如圖 3.3.3A 所示，圖上選取前 9000 點，超過 9000 圖形會開始往下掉落，fitting 的直線在短 τ 的部份取值的範圍約在 1~100 點之內，長 τ 的部份大約在 900~5000 點之內。由圖形結果得知，在 $\tau \approx 1000$ 秒 (約 16.7 分鐘) 是一個轉折的時間點，稱為 τ_c，當 $\tau > \tau_c$ 時，斜率大約為 1，即符合隨機行走與朗之萬理論的描述；

在 $\tau < \tau_c$ 時，斜率不固定，並不符合朗之萬理論的描述。

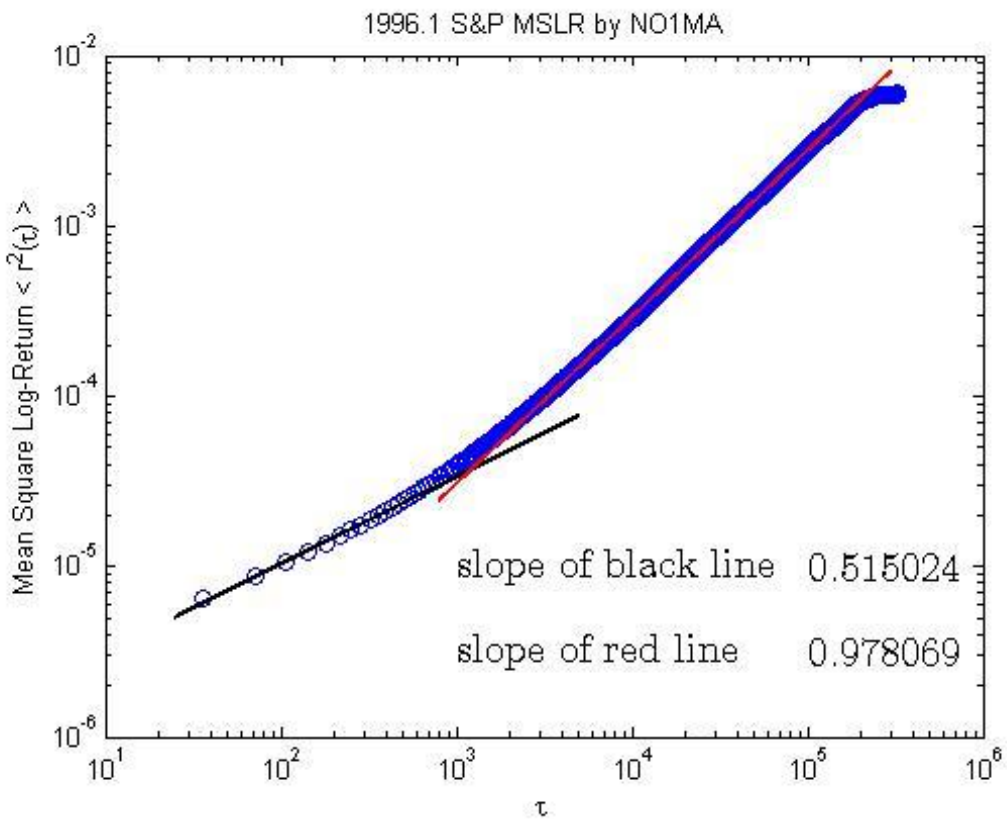


圖 3.3.3A　1996 年 1 月 S&P500 其中 345 家公司，在 noMA (或 NO1MA)下所有可能 τ (交易時間) 之下所得到的整體市場的平均之MSLR 分布圖。長 τ 部份：位移平方正比於時間，圖上截距相當於擴散係數 D 的對數值；短 τ 部份：受粒子(股價)在所處環境(市場)的影響，斜率不固定。

針對這個情形，我們認為可能的原因是，基於朗之萬理論的假設，粒子移動的軌跡是連續的，而原始數據是離散的形式，所以我們以沒有移動平均的報酬來計算，其所得到的圖形會與理論上有所差異，因此我們採用有經過一天的移動平均(HF1MA)的報酬來計算。針對以上假設，我們改先對股票價格做 HF1MA 之後，再按先前的方法算出 MSLR 之後作圖，可得到的結果如圖 3.3.3B 所示。由圖形可得知，經過 HF1MA 後，所得到的圖形較符合我們將報酬對應於朗之萬理論所預期的結果。

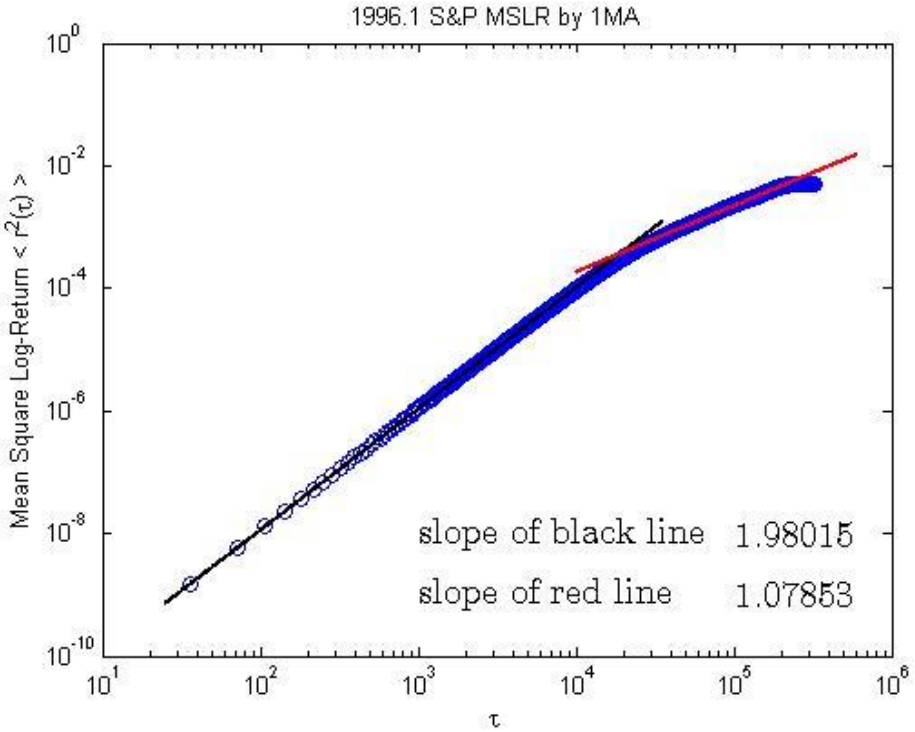


圖 3.3.3B　1996 年 1 月，有 HF1MA (即 1MA) 在所有可能 τ 之下所得到的 MSLR 分布圖。短 τ 部份：microscopic，位移平方正比於時間的平方；長 τ 部份：位移平方正比於時間的一次方。

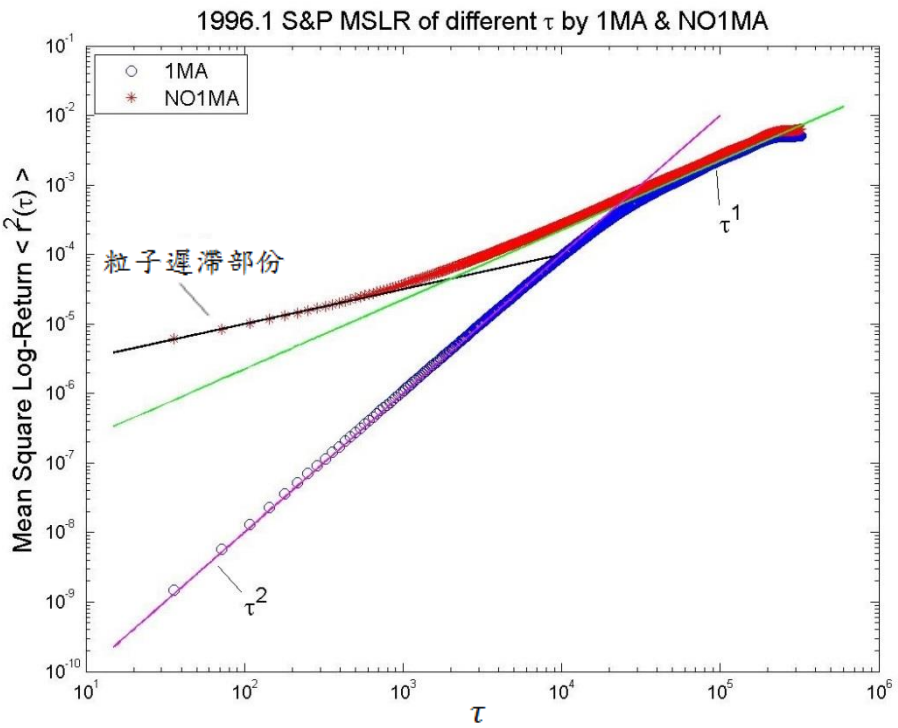


圖 3.3.3C　HF1MA 與 noMA 的圖形比較，由圖形可知在長時間兩者的斜率近乎相同，在短時間則有差異。

根據以上所述簡單整理結論如下表所示：

表 3.3.3　HF1MA & noMA 的 MSLR 之比較

	HF1MA	noMA
轉折時間 τ_c 比較	大(τ ≈ 12400s)	小(τ ≈ 1000s)
短 τ 部份 ($\tau < \tau_c$) microscopic	位移平方正比於時間的二次方，即 $\langle r^2 \rangle \propto \tau^2$，截距相當於溫度係數 T 的對數。	位移平方並不正比於時間的二次方，粒子遲滯現象的部份。
長 τ 部份 ($\tau > \tau_c$) macroscopic	位移平方正比於時間的一次方，即 $\langle r^2 \rangle \propto \tau$，截距相當於擴散係數 D 的對數值。受限資料長度，在靠近資料長度的尾端出現擾動。	同左。

根據先前文獻研究[11]的結果顯示，我們可解讀為在 noMA 的圖形中，在短時間的 $\langle r^2 \rangle$ 分布與粒子的所在的環境有關，粒子與粒子之間作用，可視為公司間的相互影響，其值並不固定，而當長時間之後影響漸漸消失，前面所發生的事所造成的影響漸漸消失(越無沒有記憶性)，復歸為擴散運動的範疇(斜率等於 1 的部份)。

3.3.4 不同時間長度移動平均之下對 MSLR 的影響

我們可進一步觀察在 MA 的情況下，轉折時間 τ_c 的位置是否受移動平均所選擇的時間間隔所影響。在圖 3.3.4A 顯示出在 1hr~5hr 的 MA 後，由小至大的時間間隔的 MA 下， τ_c 逐漸向右移且所對應的 $\langle r^2 \rangle$ 值亦越大，且 τ_c 之前的斜率越趨近於 2。在 0.5hrMA 時，圖形由斜率 1.76 的部份經轉折後較快進入斜率 1 的部份；在 1dayMA 時，圖形由斜率 2 的部份經轉折後進入斜率 1 的部份則較慢(彎曲幅度較大)。

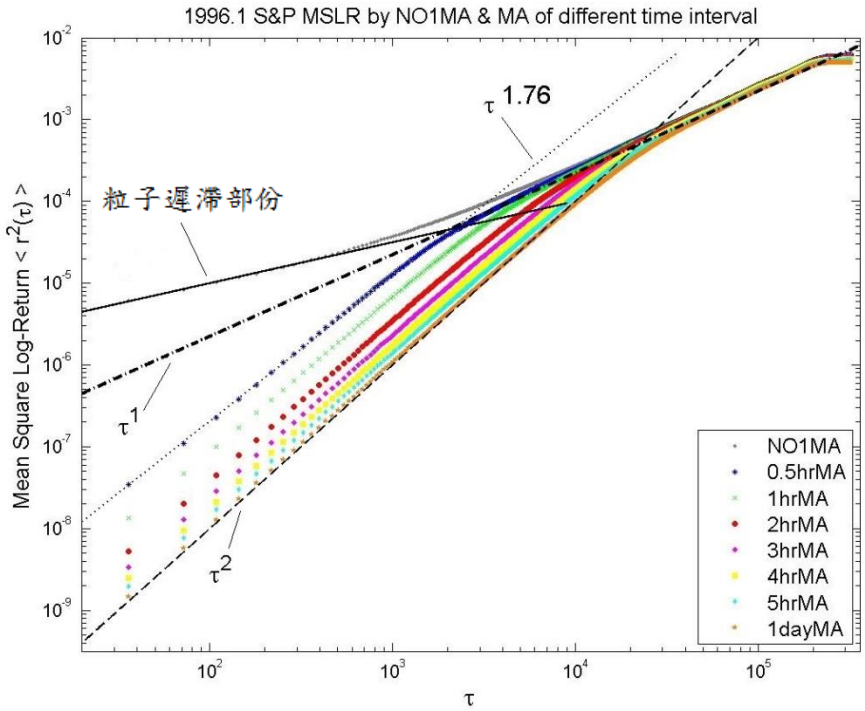


圖 3.3.4A　HF1MA 在不同時間尺度下的圖形比較，由圖形可知在長時間的斜率近乎相同，在短時間的部份斜率則有差異。

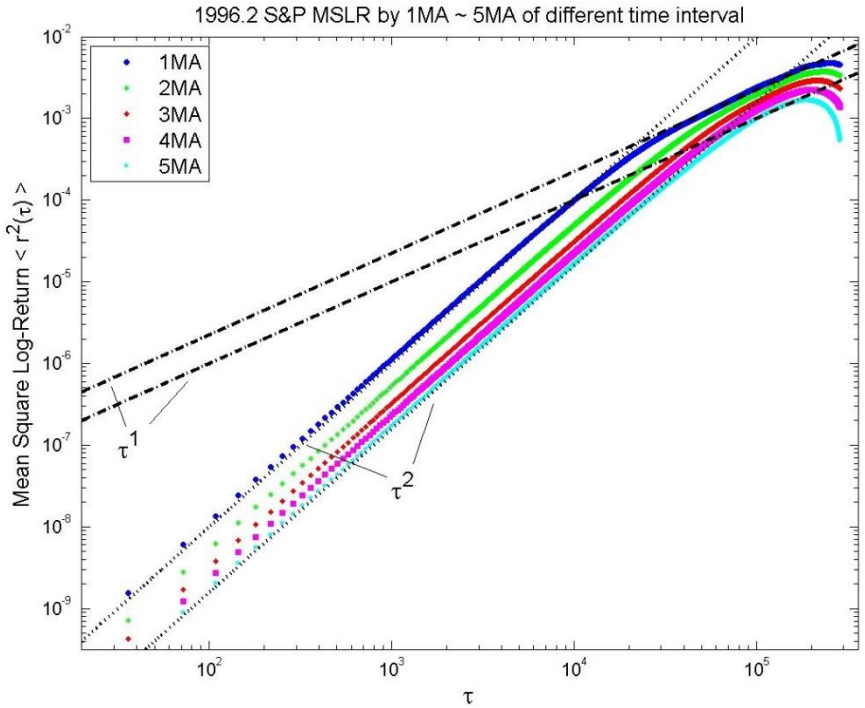


圖 3.3.4B　不同移動平均的時間長度下的圖形比較，移動平均的時間長度從 1 天的移動平均(1MA)到 5 天的移動平均(5MA)，由圖形可知 MA 的長度對轉折時間 τ_c 有影響，末端的掉落受限資料長度所致，統計性較差。

另一方面，由實證可知 MSLR $\langle r^2 \rangle$ 的解有指數形式[11]，

$$\langle r^2 \rangle = \begin{cases} \theta\tau^{\alpha} & where \;\; \alpha \approx 2 \;\; for\, short \;\; \tau \\ D\tau^{\alpha} & where \;\; \alpha \approx 1 \;\; for\, large \;\; \tau \end{cases}$$

其中 $\theta = D/2\tau_0$ 與 D 相當於朗之萬方程的解之中的溫度與擴散係數，由於在短 τ 中的 α 誤差值很小，所以可經由對數尺度下求得 θ，而在長 τ 時的 α 誤差較大，在對數尺度下 D 的誤差也大，所以改在線性尺度下直接找一條線 L：$\mathrm{Y} = D * \tau + b$ 來 fitting 長 τ 的部份以找出 D，經由 θ 與 D 可以找出特徵時間 τ_0 的值。

由朗之萬方程所求得的 MSD 公式形式如下

$$MSD \equiv \langle [R(t+\tau) - R(t)]^2 \rangle = D\left[\tau - \tau_0(1 - e^{-\tau/\tau_0}\right]$$

我們可以找到一條主曲線(master curve)[11]

$$\langle r^{*2} \rangle = \tau^* - 1 + e^{-\tau^*}$$

來 fitting 實證資料經過尺度化之後的 MSLR $\langle r^{*2} \rangle = \langle r^2 \rangle / (D\tau_0)$ 對應於尺度化後的時間 $\tau^* = \tau/\tau_0$，這裡我們對 1996 年的數據採用以每 4 週當作一個月的資料作圖，發現除了幾個較特別的月份之外，每一條 MSLR 幾乎都落在主曲線上。

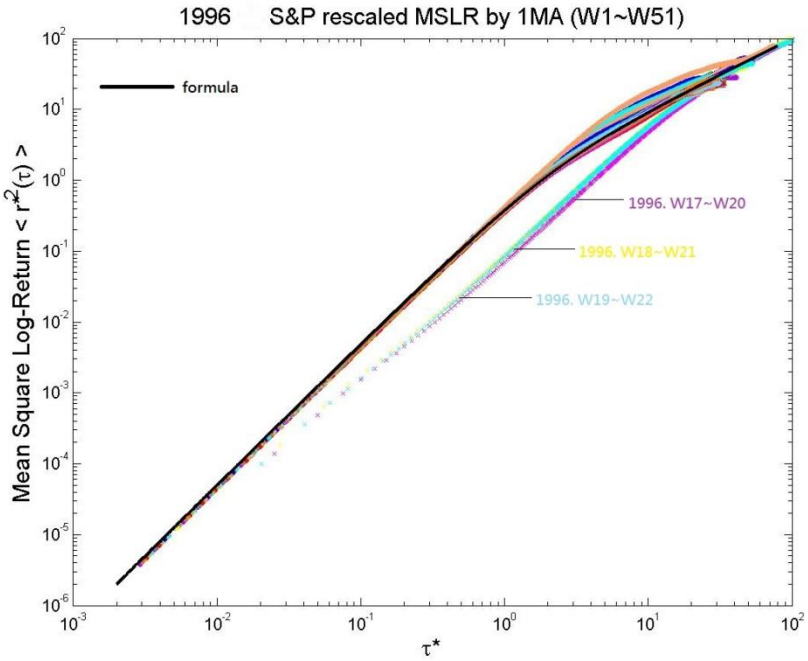


圖 3.3.4C 1996 年連續每 4 週的資料在 HF1MA 下的 MSLR，圖形幾乎都落在主曲線上，其中 W17~W20 剛好為 5 月的那 4 週。

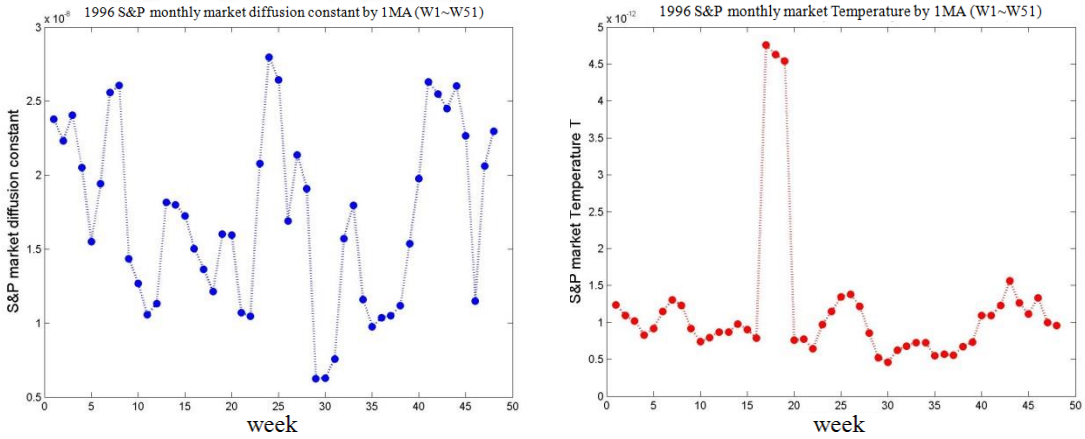


圖 3.3.4D　1996 年在 HF1MA(即 1MA)之下所得到的擴散係數(左)與溫度(右)。圖中每一點代表連續每 4 週的資料，如第 1 點為 1996 年的第 1 週到第 4 週，第 2 點為 1996 年的第 2 週到第 5 週，以此類推。由圖中可約看出兩者的起伏有相似的規律。

3.4 股價指數報酬對應於 Lévy 機率分布的實證

本章節詳細論述股價報酬對 Lévy 機率分布的實證，我們已知股價報酬在機率密度 $P(R=0)$ 附近可以用 Lévy 分布適當地描述，即$P(R=0)\sim\tau^{-1/\alpha}$；而在厚尾部份與時間尺度呈指數函數關係:$P(R>x)\approx\frac{1}{x^k}$。本節討論在 HF1MA 與 noMA 的情況下，以兩種不同的方式計算所得到的結果是否也符合上述所列的性質，結果發現在 HF1MA 的情況下，所得到的 α 值約為 1，比 noMA 的情況下($\alpha\approx1.6$)來得小，顯示移動平均對 α 值的影響，並且移動平均的時間長度與 X 軸尺度皆會影響 α 值，以及尺度轉換之後的報酬在$\mathrm{P(R=0)}$附近的機率密度圖形可以重疊在一起，最後我們也發現，最小交易尺度的報酬$\mathrm{P(R=0)}$附近可以用 Lévy 分布來描述。

這裡所提出的二種計算分析方式，分別介紹如下，我們以法 1 當作主要的驗證方式，而提供法 2 作為佐證。利用這二種計算分析方式，由粒子系統推論到股票市場系統，對股價報酬 (log – return) 做分析與觀察。

法 1：整體股價計算法

此方法的構想為把每一家公司在不同交易尺度 τ 的報酬當作是每個粒子在移動了幾步之後所在的位置，假設一組金融時間序列的資料總長度為一個月(N 筆資料)，交易時間尺度 $\tau = 3\ \text{min}$，當作 $3\ \text{min} \div 36\text{s} = 5$ 個資料點的平移(意即粒子走了 5 步)，並由 3.2B 式求得在 $\tau = 3\ \text{min}$ 交易之下的時間序列 $R_{3min}(\text{t})$，可以得到一組長度為 N – 5+1 = 12347 的 log – return 值。則此時 12347 個資料點當作是 12347 個經過 5 步之後所在位置的隨機漫步試驗。又市場上有 345 家公司，為了能充份利用這組資料，則我們做如下的步驟：12347 x 345 = 4259715，這些資料點代表在 $\tau = 3\ \text{min}$ 時有 4259715 個隨機漫步值。此種計算方法可以使資料個數大幅增加。

法 2：市場價格平均法

以每一個月為資料總長度，先把每一個月各家公司的股價相加取平均後，再作對數報酬。例如市場上有 3 家公司在一週內的交易如下表所示，接著再以平均的價格去計算對數報酬。

表 3.4 市場價格平均法

price / firm	Monday	Tuesday	Wednesday	Thursday	Friday
A	10	11	12	13	14
B	10	9	8	7	6
C	4	7	10	7	4
Average	8	9	10	9	8

國立政治大學 National Chengchi University

3.4.1 標準普爾指數 (S&P500) 報酬分布的尺度特徵

一般而言，在不同時間尺度下的交易，報酬的機率分布的情況會有所不同。當股價交易的時間尺度越長，股價的變化會越大，因此股價指數報酬的機率密度分布圖，也會隨著時間尺度越長越呈低擴的特徵：即圖形向兩側逐漸擴展，中央部分越低的現象，在S&P500的股票指數，根據先前研究成果，例如R. N. Mantegna和H. E. Stanley 在1995年發表於Nature[2]的文章談到相關研究，可以發現上述特徵。其資料選取範圍是1984年~1989年間S&P500的股價指數，時間尺度的範圍大約在1 min~1000 min之間，最大的資料總數是493545筆 ($\tau = 1\text{min}$) ，最小的資料總數是562筆 ($\tau = 1000\text{min}$)。

長期來看，理論上可假設對稱分布 ($\beta = 0$) 以及平均值為0 ($\mu = 0$) 的情況，然而實際數據受限於資料長度，所以不一定平均值為0且對稱，所以會有誤差。一般研究股價指數對數報酬的機率分布，皆取較長的時間長度，這裡我們把時間序列長度限定在1個月，並探討股價報酬在HF1MA及noMA的情況下，在不同的時間尺度其圖形分布的異同，以及在計算時X軸尺度大小(即間隔個數, 簡稱HI)如何影響機率分布圖形，若HI值(間隔個數)越大表示X軸每一格的尺度越小。我們採取上述二種方式來做分析與討論，計算結果與圖形分別如下所示，其中 τ 的範圍選取：1.2 min, 2.4 min, 4.8 min, 9.6 min, 19.2 min, 38.4 min, 76.8 min, 153.6 min, 307.2 min, 這九個時間點來討論。當 τ 超過1天的尺度之後，即 τ 為1228.8 min, 2457.6 min, 4915.2min，圖形幾乎完全無規則可言。

方法 1(mtd1)**：整體股價計算法**

<u>HF1MA</u>：因為資料個數大大地擴展，每月的資料長度變成為 1430000~3555000 筆之間，因為有經過 HF1MA 的步驟，所以圖形會較為平滑。以 4.8 min, 9.6 min, 19.2 min, 38.4 min, 76.8 min 這 5 個時間點作圖，這裡選取的 HI 值(間隔個數)從 100 開始到 300 較佳，超過了 300 之後圖形的效果漸漸變差。

<u>noMA</u>：資料個數如 HF1MA 的個數，在 HI 值(間隔個數)為 500 之前，機率密度的最高點不一定都在報酬等於 0 的位置，此處我們以 4.8 min, 9.6 min, 19.2 min, 38.4 min, 76.8 min 這 5 個時間點作圖。因資料個數大大地擴展，統計性較好，所以圖形如 HF1MA 的平滑，且曲線彼此非常地靠近，但仍可以看出圖形隨時間尺度的增加而遞減的情況。這裡選取的 HI 值(間隔個數)從 800 開始到 1000 以內較佳。

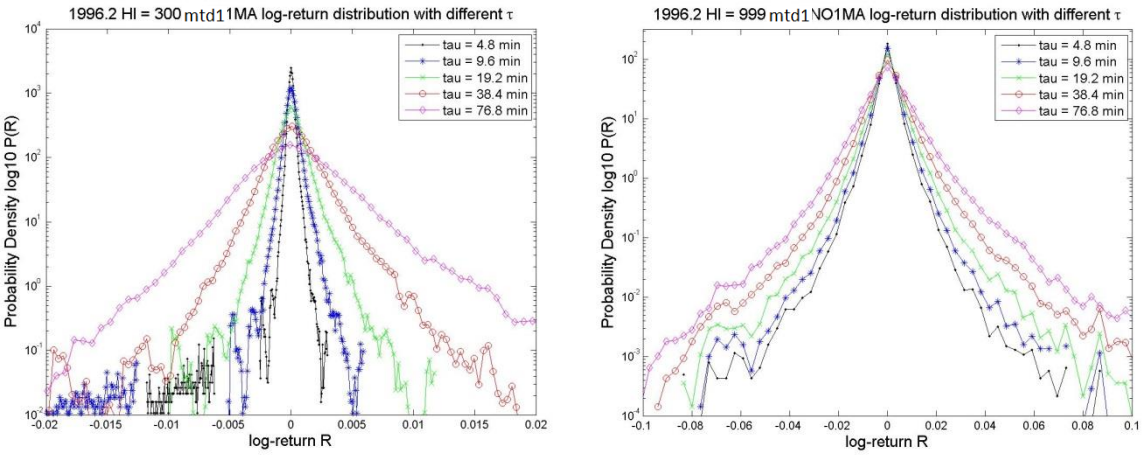


圖 3.4.1A　1996 年 2 月利用方法 1(mtd1)在不同時間尺度下的股價指數報酬機率密度分布，左圖為 1MA 下的分布，右圖為 NO1MA (即 noMA) 下的分布。由上圖可以得到的訊息是，時間尺度 τ 越小越狹長，尺度越大越低擴。

方法 2(mtd2)**：市場價格平均法**

<u>HF1MA</u>：總時間長度為 1 個月(每月大約落在 11700~13000 筆之間)，因為資料的時間長度短，所以 1 天的移動平均(1dayMA)波動較大，只可能大略看出與時間尺度的關係，效果不好。所以此處我們把移動平均的時間尺度縮小為 1 小時(1hrMA)，可以得到較好的效果，HI 值(間隔個數)選擇 10~60 之間的圖形效果較好。為了可以充份在圖形上顯示中心的分布形狀，僅畫出部份時間尺度並選取較好的分布做代表。

<u>noMA</u>：每月的資料長度落在 11700~13000 筆之間，這裡選取的 HI 值(間隔個數)約 80 開始到 550，超過 550 圖形分配效果漸漸變差。為了可以充份在圖形上顯

示中心的分布形狀，僅畫出部分時間尺度並選取較好的分布做代表。

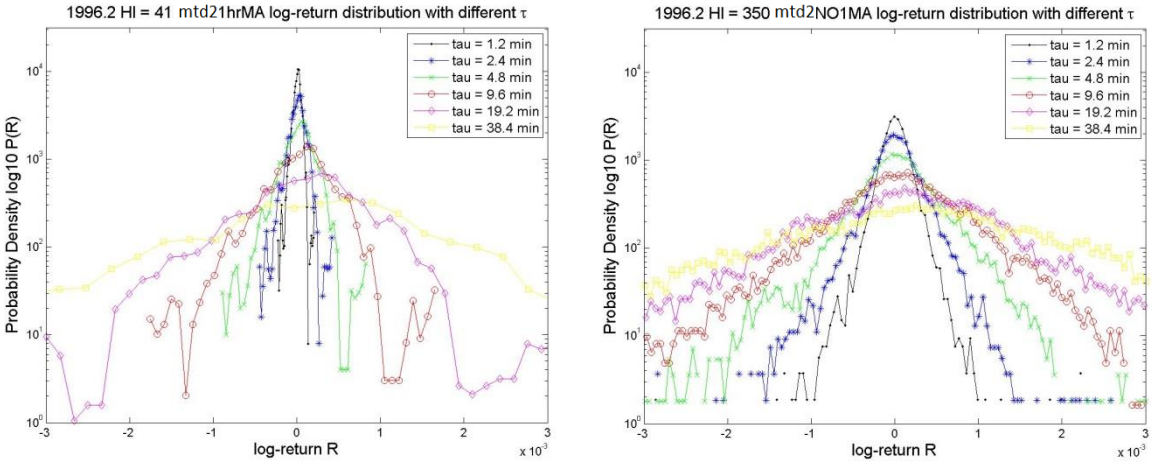


圖 3.4.1B　1996 年 2 月利用方法 2(mtd2)在不同時間尺度下的股價指數報酬機率密度分布，左圖為 1hrMA(一小時的移動平均)下的分布，右圖為 NO1MA(即 noMA)下的分布。由上圖可以得到的訊息是，統計性質的優劣受到時間序列長短所影響。

在這裡我們做一個簡單的小結：

(1) 在短 τ 的部份如 1.2min ~ 9.6min，經過 HF1MA(1 天的移動平均)或 1hrMA(1 小時的移動平均)的圖形較狹長、X 軸寬度較窄且最高點的值比較大；在長 τ 的部份如 36.4min 與 76.8min，經過 1dayMA 或 1hrMA 的圖形、X 軸寬度及最高點的值比較一致。

(2) 對於有無移動平均的情況，兩者在 τ 越短時圖形越狹長，在 τ 越長時圖形越低擴。

(3) 交易的時間尺度越長，其統計性質越差，例如 τ 在超過一天的尺度時，圖形無幾乎沒有規則可言。

以上計算結果符合先前的研究[3]，接下來我們接續驗證圖形中心部份是否可用 Lévy 穩定分布來描述。

3.4.2 股價報酬機率密度函數的指數遞減

根據文獻研究股價報酬機率分布厚尾部份的結果指出，若呈指數函數遞減，形式為 $e^{-\beta x}$；而美國市場的股價資料是冪次函數遞減，形式為 x^{-k}，大於零與小於零的報酬機率分布，其 k 值在對數圖形上大約為 3。

根據過去分析[3]，研究 Argentine (MERVAL), Brazil (BOVESPA), India (BSE 30)與 S&P500 四個指數，得到 k 值的結果大約從 2.3~3.3，認為結果與市場的活絡程度有關，S&P500屬於已開發國家的經濟體而其他指數屬於開發中國家的經濟體。我們依據該文章的方法來檢驗在資料短時間(1個月)且高頻的情況下，機率分布的厚尾部份是否也符合 $\mathrm{p(g > x)} \approx \frac{1}{x^k}$ 的形式，其中 g 為標準化後的報酬。定義標準化的報酬為

$$\mathrm{g} \equiv \frac{R - \langle R \rangle_\tau}{\sigma}$$

其中 $\sigma^2 \equiv \langle R_\tau^2 \rangle - \langle R \rangle_\tau^2$，$\langle \ldots \rangle_\tau$ 表示在這一系列時間尺度 τ 的算術平均。

標準化後的圖形近似截尾 Lévy 分布，其累積機率密度分布後段斜率近似於 3。

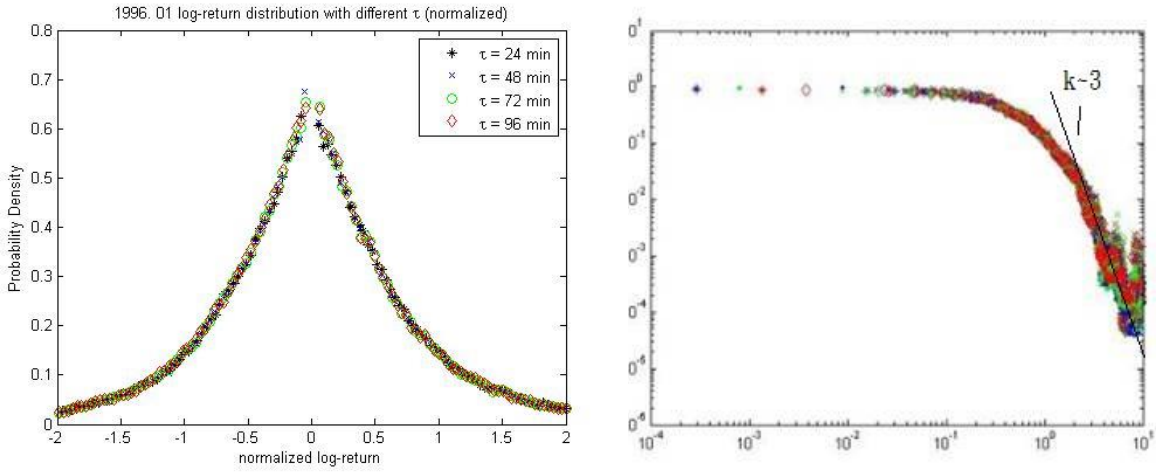


圖 3.4.2　1996 年 1 月在各個不同時間尺度下，報酬標準化後的圖形。左圖為在 HF1MA 之下的圖形，近似截尾 Lévy 分布，右圖為其累積機率密度分布，後段斜率近似於 3。

3.4.3 Lévy 穩定分布中心附近 P(R=0)的 α 值與圖形 fitting

穩定分布常被用作金融數據的分析，例如本華·曼德博(Mandelbrot)發現棉花價格的變化服從穩定分布($\alpha = 1.7$)。為了瞭解 S&P500 股價報酬是否為穩定機率分布，我們利用 Lévy 穩定分布的尺度特徵來觀察 S&P500 的股市資料。

我們分析 S&P 其中345家公司在1996年~1999年當中每個月的股價指數，其指數為每36秒計錄交易一次，對股價根據(3.2A)式的定義，作對數報酬的運算之後再作機率密度函數，接著找出不同時間尺度下，股價報酬在零附近的機率密度函數值，記為P(R=0)，並將其與對應的時間尺度作對數分布圖，若該圖呈現直線，就表示機率密度函數和時間的關係符合(2.3D)式。

一般以整體市場的總指數在長時間之下來找出 α 對時間尺度的對數圖，這裡與一般分析不同的地方是以 345 家公司的計算來類比整體市場行為，並在 HF1MA 與 noMA (即 raw data)的情況下，以上述二種方式觀察異同。

如在2.2節所述，獨立同分布 (i.i.d) 的隨機變量之和以及它們本身具有相同的分布，我們利用(2.3D)式，並令隨機變數是不同時間尺度的股價報酬，(2.3D)式的 n 可以用時間尺度 τ 代換 (因為 $\tau = n \times 36$ 秒，36秒為最小時間單位)，因此我們可以透過(2.1D)式，找出以下的時間尺度關係[8]：

$$\log P(R_\tau = 0) = log\left(\frac{\Gamma(1/\alpha)}{\pi\alpha\gamma^{1/\alpha}}\right) + (-1/\alpha)\, log\, \tau$$

利用不同的 τ 對 $P(R_\tau = 0)$ 和 τ 分別取對數作圖，找出圖形裡的直線斜率即 $-1/\alpha$，而斜率取倒數的負值即是(2.3C)式的 α 值。計算結果與圖形分別如下所示，其中 τ 採 1.2 min, 2.4 min, 4.8 min, 9.6 min, 19.2 min, 38.4 min, 76.8 min, 153.6 min, 307.2 min, 九個時間點。

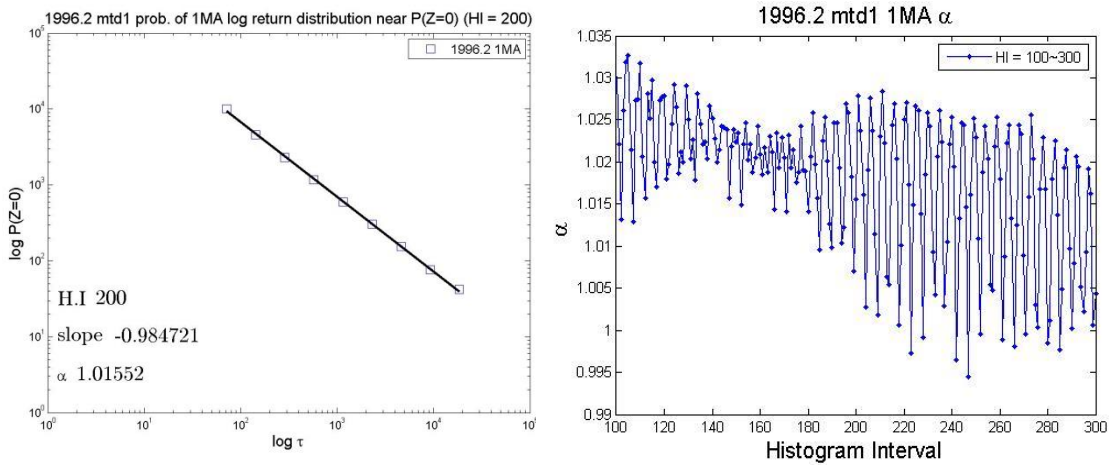


圖 3.4.3A　上圖為利用方法 1(mtd1)計算所得到的圖形，上圖左為 1996 年 2 月經過 HF1MA 後在 P(R=0)附近所 fitting 出來的 α 值，上圖右為 HI 值在 100~300 之間 α 值的分布，圖中可約略看出在 HI = 140~180 時 α 值分布的波動較小。

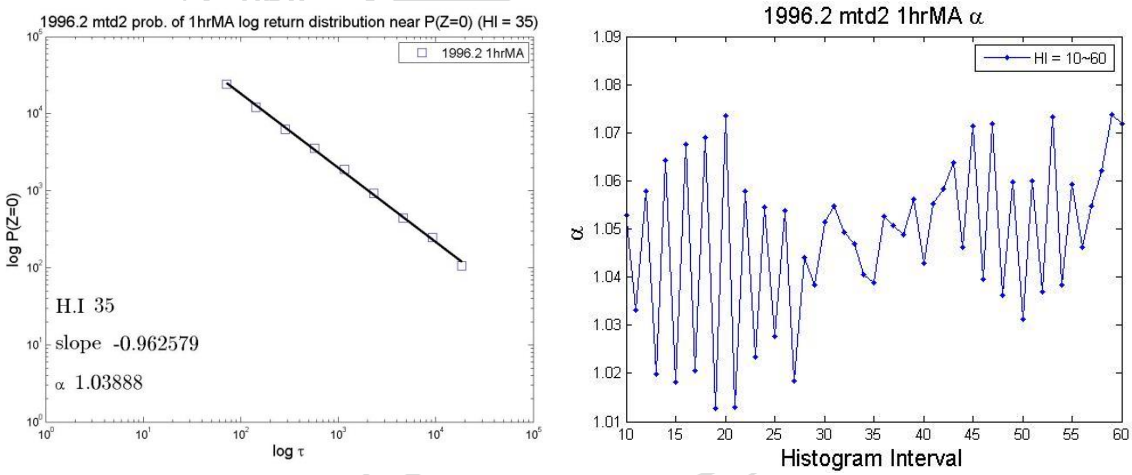


圖 3.4.3B　上圖為利用方法 2(mtd2)計算所得到的圖形，上圖左為 1996 年 2 月經過 1hrMA(一小時的移動平均)後在 P(R=0) 附近所 fitting 出來的 α 值，上圖右為 HI 值在 10~60 之間的分布，圖中可看出約略在 HI = 29~41 時 α 值分布的波動較小。

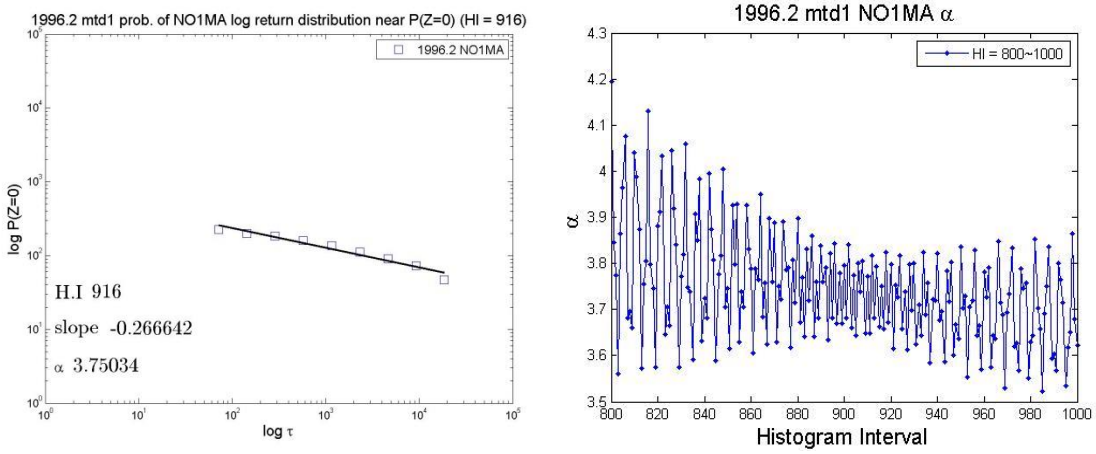


圖 3.4.3C　利用方法 1(mtd1)計算所得到的圖形，左圖為 1996 年 2 月 noMA (即 NO1MA) 在 P(R=0) 附近所 fitting 出來的 α 值，右圖為 HI 值在 800~1000 之間的分布，圖中可約略看出在 HI=890~930 時 α 值分布的波動較小。

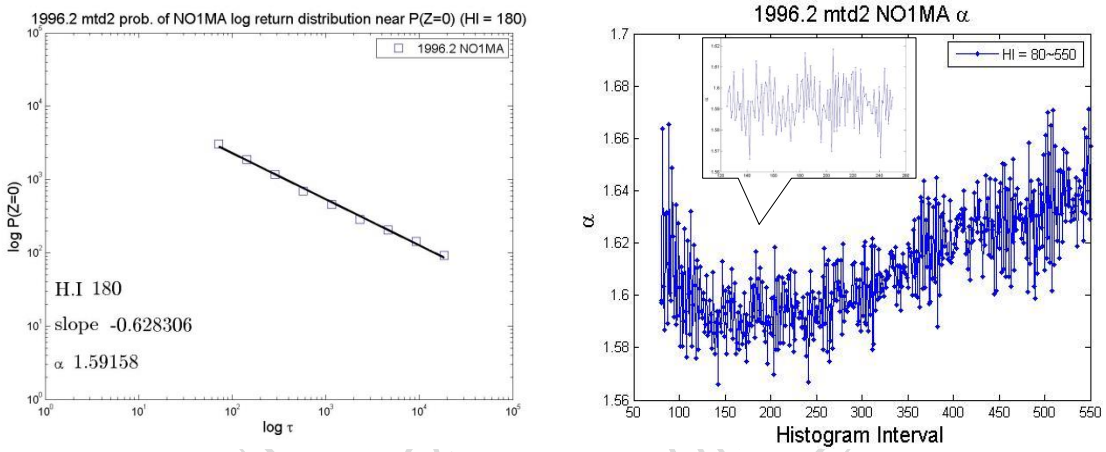


圖 3.4.3D　利用方法 2(mtd2)計算所得到的圖形，左圖為 1996 年 2 月 noMA (即 NO1MA) 在 P(R=0) 附近所 fitting 出來的 α 值，右圖為 HI 值在 80~550 之間的分布，圖中可約略看出在 HI=125~250 時 α 值分布的波動較小。

有了 α 值我們可以對隨機變數做尺度上的調整，基於能 fitting 出 α 值的線上所有點應皆可以重疊在一起。根據 2.3 節所描述之穩定機率分布的概念，由 Lévy 穩定分布的可加性得知，隨機變數加總所形成的新隨機變數，其機率分布形式會和原來隨機變數的機率分布相似。並且經過這樣的調整之後，如果是穩定機率分布，調整後的機率密度函數會重疊在相加時間序列的最小尺度單位上。我

們利用下式，對隨機變數作尺度上的調整(rescale)[2][12][13]：

$$S_\tau = \frac{S}{\tau^{1/\alpha}} \qquad (3.4.3A)$$

$$P(S_\tau, \tau_{min}) = \frac{P(S,\tau)}{\tau^{-1/\alpha}} \qquad (3.4.3B)$$

其中 S_τ 是不同時間尺度的股價指數報酬，經過$(3.4.3A)$式重新尺度調整後得到的新報酬，再以這個新報酬做機率密度函數如$(3.4.3B)$式，以下是利用上述兩式所作的機率密度函數圖形。

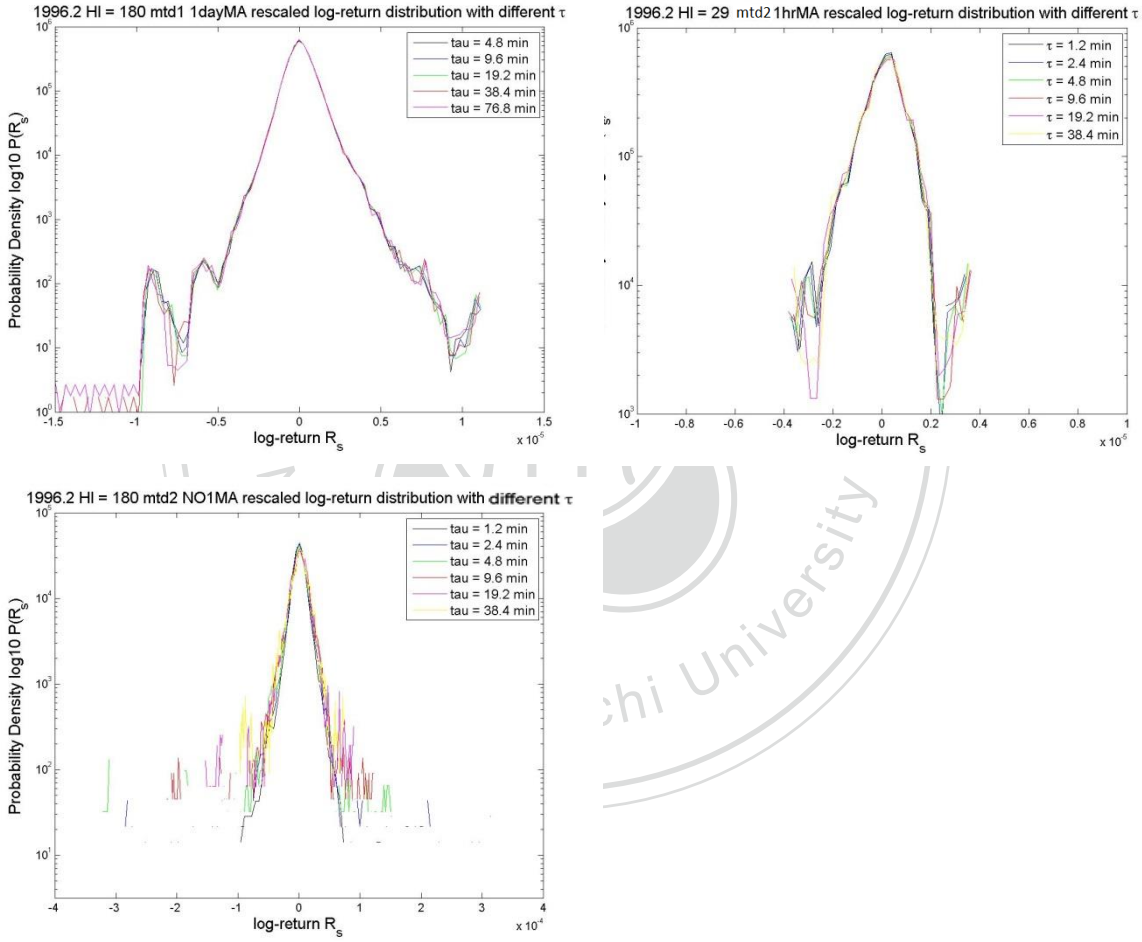


圖 3.4.3E　1996 年 2 月在經由 HF1MA(或 1hrMA)與 noMA (即 NO1MA) 的方法後所得到的 α 值計算後的 rescale 圖形，可知在方法 noMA 的情況下最近文獻的結果。

小結：

相較於過去的研究，這裡我們採用每個月的資料，長度較短且交易時間較短，並分為有移動平均與沒有移動平均的計算，所以得到的結果較一般所預期不盡相同，而我們在 NO1MA 以方法 2 來計算 α 值時，就可以得到較符合預期的結果，顯示可能的訊息是：(1)法 2 在 NO1MA 的概念與 S&P500 指數的編制相近，(2)可能比較可以去除白噪音的干擾，基於時間序列在平均之後，外在干擾的平均值為 0，即 $\langle W_\tau \rangle = 0$，所以股價指數報酬在零附近的機率分布接近 Lévy stable process。

表 3.4.3　Lévy 穩定分布在 HF1MA 與 noMA 情況下的比較

	有移動平均 (HF1MA)	沒有移動平均 (noMA)
方法 1 (mtd1)	1. 移動平均可以消除外在因素的影響，基於白噪音 $\langle W_\tau \rangle = 0$。 2. 在 α 值波動較小的情況之下，$\alpha \approx 1$。	noMA 的整體計算方式，雖然增加資料總數，但也增加外在因素的影響，所以使資料不符合穩定分布，α 值均大於 2。
方法 2 (mtd2)	經過 1 小時的移動平均 (1hrMA)，在 α 值波動較小的情況之下，$\alpha \approx 1.2$。	圖形較為接近先前文獻研究的結果，在 α 值波動較小的情況之下，$\alpha \approx 1.6$。

3.4.4 以 Lévy 穩定分布描述股價報酬在 P(R=0) 的特徵

由圖形可看出報酬的機率分布平均值非常趨近為零且左右近乎對稱，因此我們可以假設決定 Lévy 分布的參數 $x \sim S(\alpha,\beta,\mu,\gamma)$ 的 β 與 μ 均為 0 (不偏與均值為零)。另外，由 (2.3D) 式可以知道

$$P(S_n = 0) = \frac{\Gamma(1/\alpha)}{\pi\alpha(n\gamma)^{1/\alpha}}$$

其中$P(S_n = 0)$ 為實證資料的報酬在 0 附近的機率密度， n 代表有 n 組隨機變數，在先前文獻[2][13]中所研究探討的 n 組股價報酬相加，n 可由時間尺度 Δt 替代，並且其最小時間尺度為 1 min，所以當 $\Delta t = 2min$ 時代表 2 組隨機數相加，此時 $n = \Delta t = 2$。在這裡我們最小時間尺度 τ 為 36 秒，同理可推當 $\tau = 72s$ 時，此時 $n = \tau/36s = 2$，即當 τ 在公式時應視為組數較為恰當而非時間。

由上式求得尺度因子 γ，下面各個圖形是利用 α 與 γ 代入公式所得到的 Lévy 分布和實證資料在線性與半對數的圖形比較。

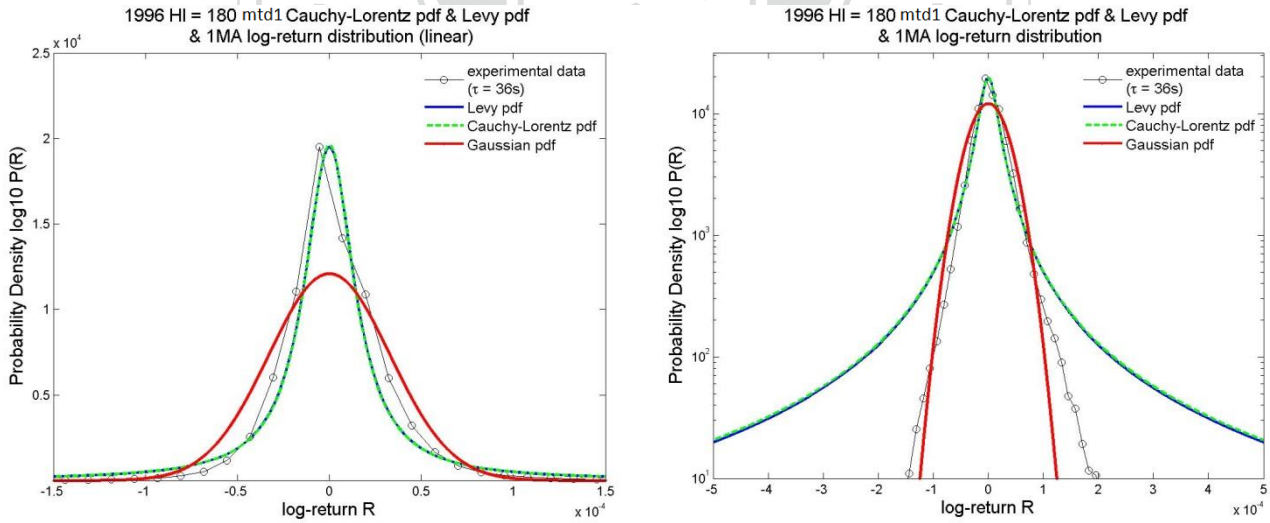


圖 3.4.4A 1996 年 2 月在 HF1MA 的情況下利用方法 1 (mtd1)所得最小 τ 的報酬分布與公式在線性與半對數的圖形，在 P(R=0)的附近符合穩定分布。

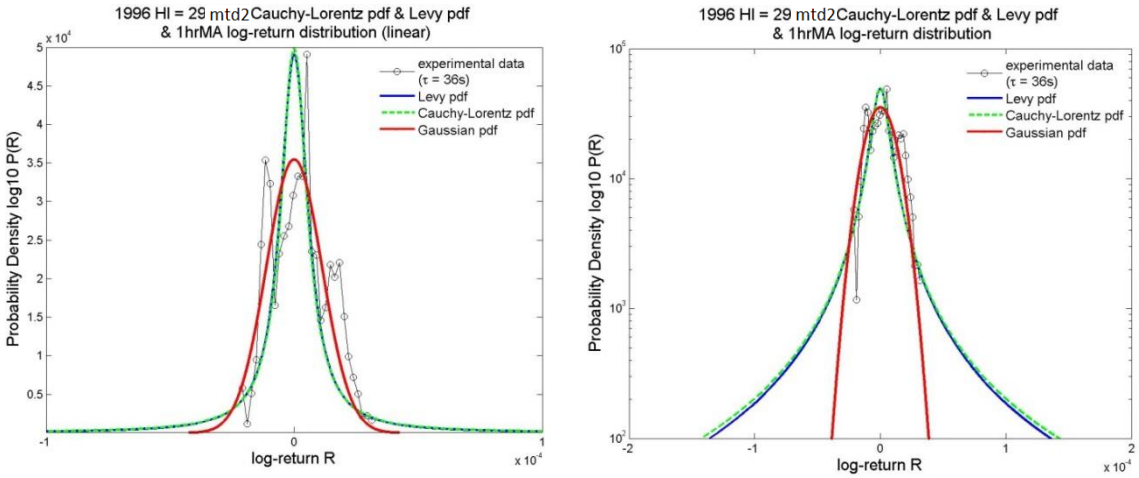


圖 3.4.4B　1996 年 2 月在 1hrMA(一小時的移動平均)的情況下利用方法 2(mtd2)所得最小 τ 的報酬分布與公式在線性與半對數的圖形，圖形零亂的原因可能在於資料總時間長度，所以統計性較差。

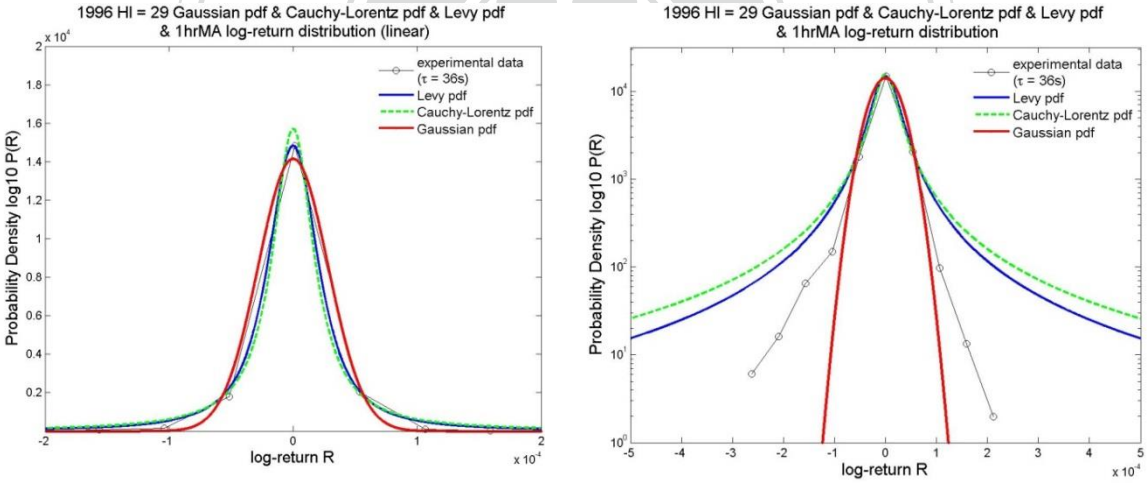


圖 3.4.4C　1996 年一整年在 1hrMA 的情況下利用方法 2(mtd2)所得最小 τ 的報酬分布與公式在線性與半對數的圖形，在提升資料總時間長度之後統計性較佳。

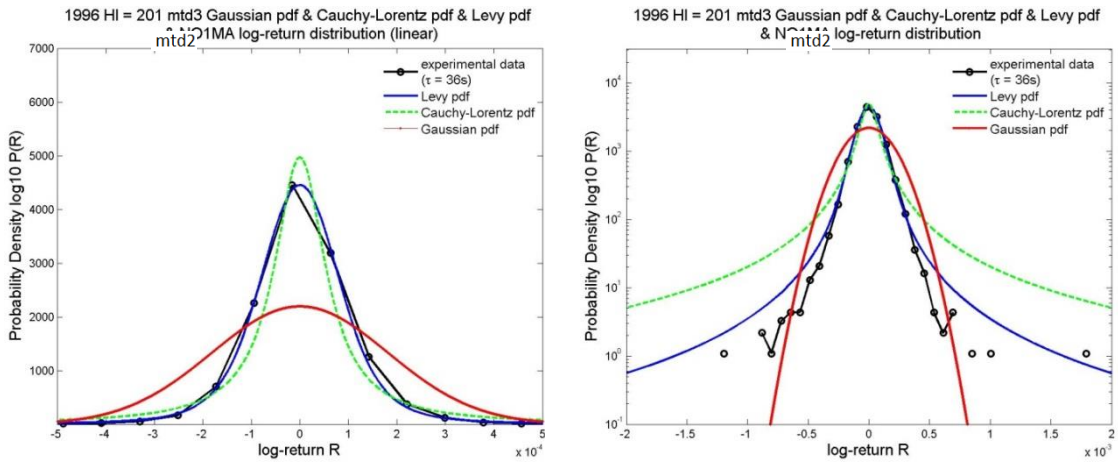
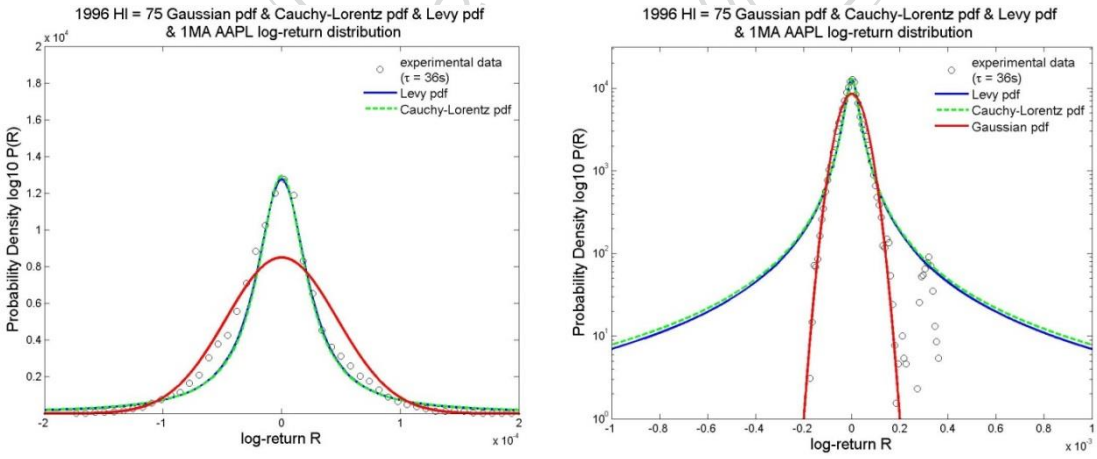


圖 3.4.4D　1996 年 2 月 noMA 的情況下利用方法 2(mtd2)所得最小 τ 的報酬分布與公式在線性與半對數的圖形，看出無論在半對數與線性尺度下的圖形在 P(R=0)的附近皆十分符合穩定分布的描述。

同時我們也分析市場當中 5 家較知名的公司，依序為蘋果(AAPL)、超微(AMD)、奇異(GE)、英特爾(INTC)與微軟(MSFT)。在 HF1MA 的情況下 α 皆為 1 左右，AAPL、AMD 及 MSFT 三家公司的圖形較符合穩定分布，在線性尺度下則 5 家公司仍可看出較符合穩定分布的描述；而在 noMA(即 NO1MA)的情況下則 α 皆大於 2，不符合穩定分布的性質。圖形分別如下所示：



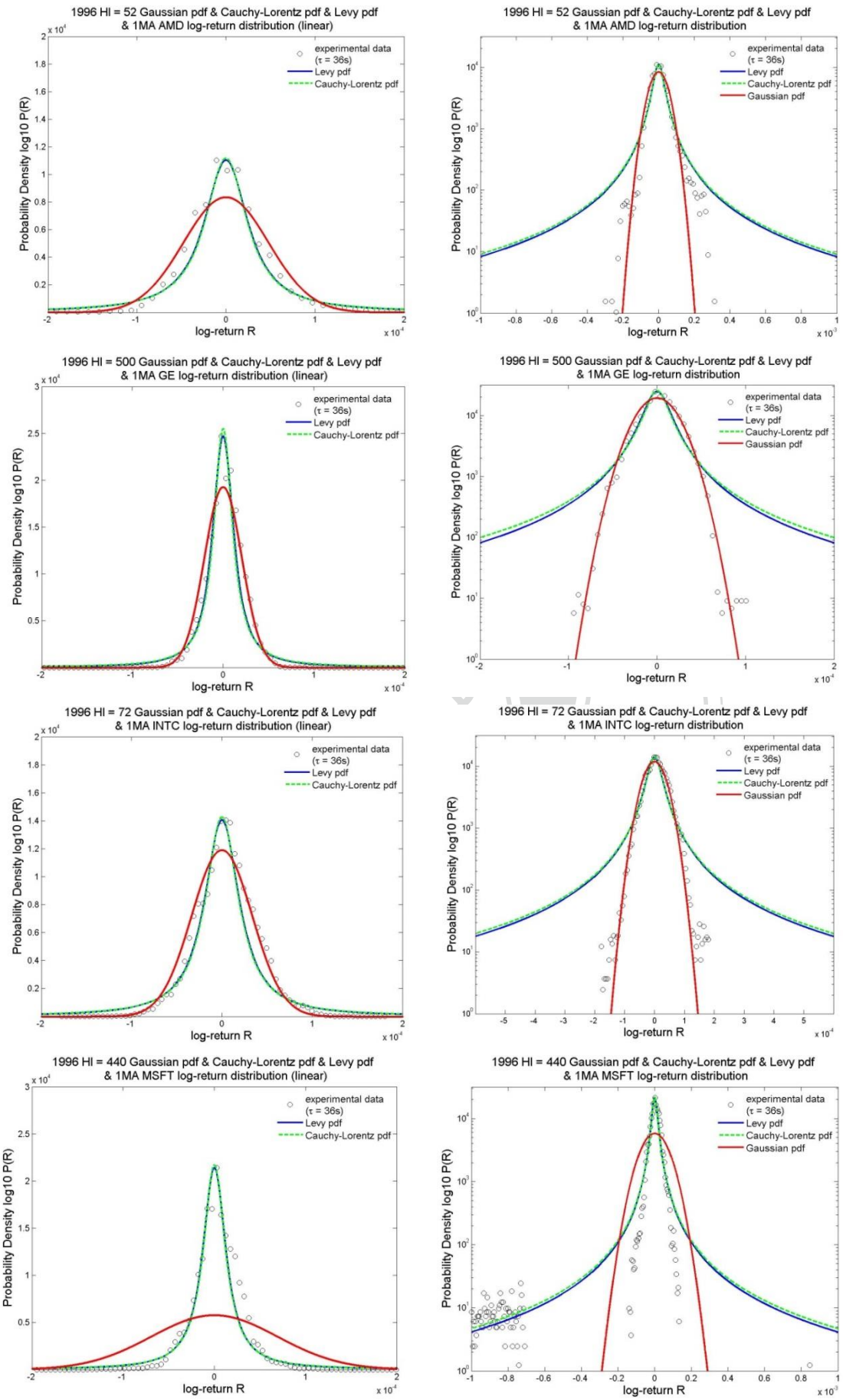


圖 3.4.4E　5 家公司的在 HF1MA 線性及半對數尺度下與公式的分布圖

結論：

在這邊我們討論 S&P 在各個方法下與個別公司的結果，S&P 指數在經由移動平均下的 α 值均接近 1，沒有移動平均的 α 值約為 1.6 左右，而個別公司在有移動平均後的 α 值均接近 1，沒有移動平均的 α 值皆大於 2，可能與單一的公司本身會受到很多外在因素干擾，導致在計算上被 white noise 所影響。我們從以上的實證分析可以得到一些資訊，經由移動平均後得到的 α 值與沒有移動平均的 α 值不同，且資料長度也會影響統計性。

表 3.4.4　Lévy 穩定分布在 HF1MA 與 noMA 情況下的整體比較

	1MA	noMA
5 家各別公司	$\alpha \approx 1$，經由移動平均消除了外在信息影響。	$\alpha > 2$，不符合穩定分布的性質，可能受外在信息干擾。
S&P 方法 1	$\alpha \approx 1$，經由移動平均後消除外在干擾，以整體來計算增加了統計性質，在 P(R=0)附近符合穩定分布。	$\alpha > 2$，雖以整體來計算增加了統計資料總數，但也增加干擾，其結果不符合穩定分布的性質。
S&P 方法 2	$\alpha \approx 1$，資料長度影響圖形，長度越長統計性越好。 在 1hrMA 的情況下 $\alpha \approx 1.2$，推估 MA 的時間越短 α 越大，最後接近 noMA 的 α 值。在作 pdf 可選取的最大間隔數通常比 noMA 的較少。	$\alpha \approx 1.6$，用方法二後可消除外在干擾，在 P(R=0)附近符合穩定分布，接近文獻描述。

3.5 自相關函數

我們知道股價指數具有短期的記憶性，隨著時間越長其記憶性越差，最後趨近於零，表示無相關性，而在 HF1MA 與 noMA 的情況下又會如何影響股價指數？我們引入不同交易時間尺度之下的自相關函數公式，稱為 Time Scale Selected Autocorrelation (TSSA)，公式如下：

$$C_{k,\tau} = \frac{\sum_{t=1}^{n-k}(x_{t,\tau} - \bar{x}_{t,\tau})(x_{t+k,\tau} - \bar{x}_{t+k,\tau})}{\sum_{t=1}^{n-k}(x_{t,\tau} - \bar{x}_{t,\tau})^2} \qquad (3.5A)$$

其中 x_t 為一組時間序列，n 為序列資料點個數，k 表示資料的平移個數，當 $k = 0$，$C_0 = 1$，完全自相關，k = 1 代表平移一格， k = 2 代表平移二格，其目的是在不同的時間尺度 τ 之下，觀察自相關函數分布的變化。

以 1996 年的資料為例，在 HF1MA 以及 noMA 的作法下，計算 S&P 整體市場的 log-return 的自相關函數，原始資料的交易時間間隔為 36 秒 1 次。

所謂不同的 τ 為：設原始交易時間為 36 秒的交易價格為 (假設取對數後)

	t = 0s	t = 36s	t = 72s	t= 108s	t = 144s	t = 180s	……
log p	11	12	13	12	11	12	……

則 τ = 36s 的 log-return 序列為 r_{36s} = [1 , 1 , -1 , -1 , 1 , …]

τ = 72s 的 log-return 序列為 r_{72s} = [2 , 0 , -2 , 0 , …]

τ = 108s 的 log-return 序列為 r_{108s} = [1 , -1 , -1 , …]

τ = 144s 的 log-return 序列為 r_{144s} = [0 , 0 , …]

再以這些資料當作新的時間序列去計算自相關函數。

國立政治大學 National Chengchi University

我們以兩種方式計算自相關函數：

方法 1(mtd1)：自相關平均法

先計算出 345 家公司的自相關函數之後再取平均。

方法 2(mtd2)：整體股價計算法

如 3.4 節所提到的方法 1(mtd1)。

上面兩種方式在有 HF1MA 與 noMA 的情況下分別對各月份作圖，茲舉 1996 年 2 月圖形為例：

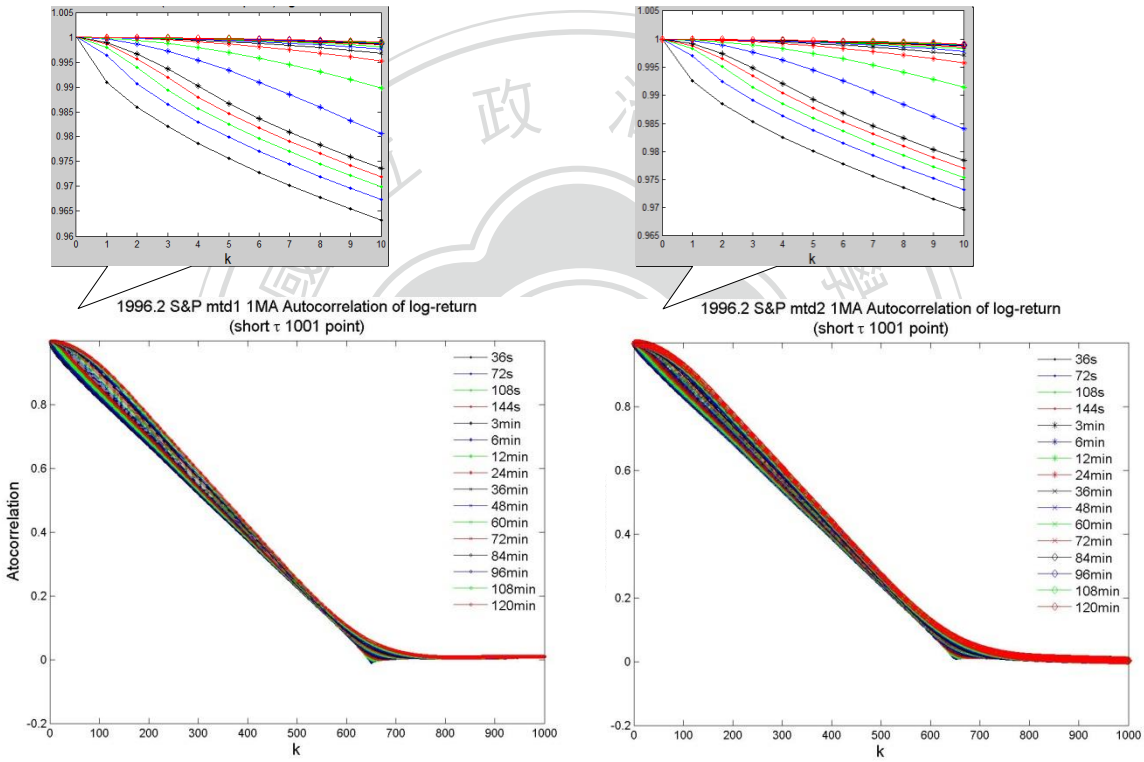


圖 3.5A　1996 年 2 月經由 HF1MA 分別在方法 1(mtd1)與方法 2(mtd2)作不同時間尺度下的圖形比較，小圖為 k=0~10 格的放大，圖形有平滑軌跡的特徵。由圖形可看到時間尺度越長，有圓頭向外擴展，接著變成斜直線，在零附近振盪。

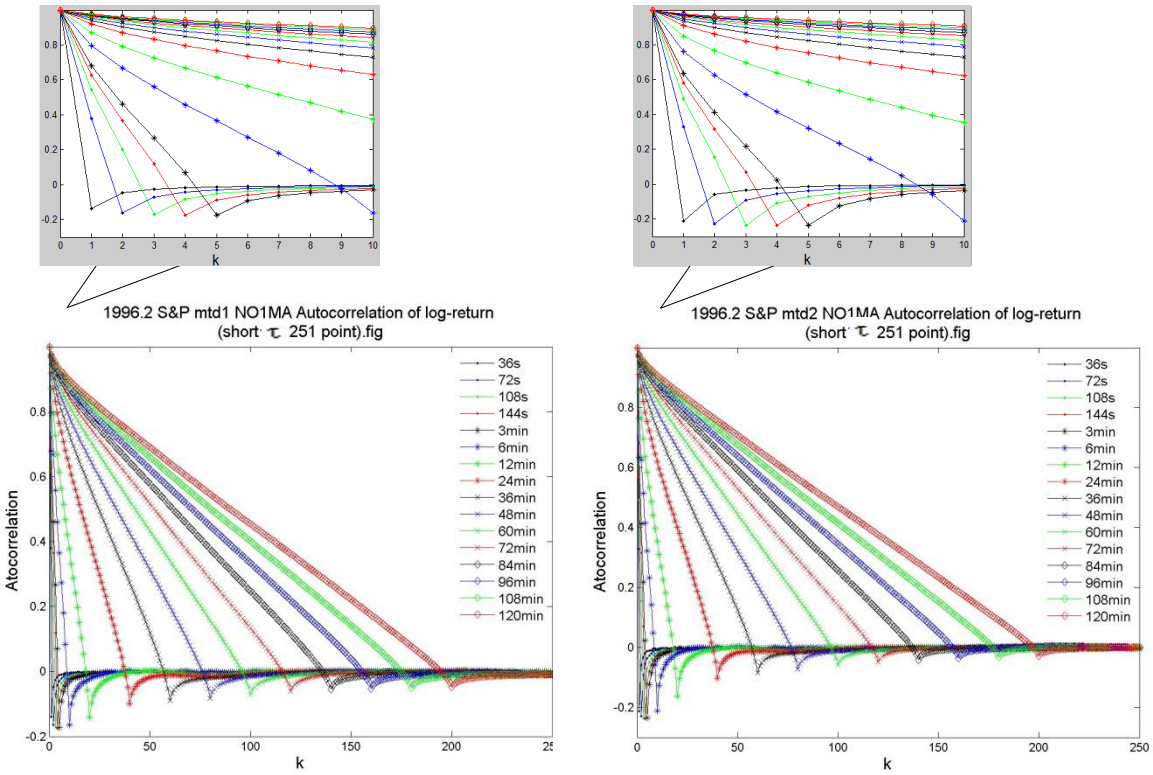


圖 3.5B　1996 年 2 月在 noMA (即 NO1MA) 分別在方法 1(mtd1)與方法 2(mtd2)作不同時間尺度下的圖形比較，小圖為 k=0~10 格的放大，由圖形可看到剛開始皆下落，在時間尺度短的圖形迅速下落到負值再轉折到零附近震盪，時間尺度越長則下落的速度越慢，最後在零附近後振盪。

由以上實證結果我們可以發現，不論方法 1 或方法 2 在 HF1MA 與 noMA 的情況下，圖形都很相似。另外在時間的記憶性方面，歸納如下

(1) HF1MA 的情況下則完全由移動平均的時間長度所掌控，我們可以認為在 HF1MA 的情形下，小尺度的 τ 在剛開始時的交易前後仍較無關聯，基於移動平均消除外在干擾，所以快速下降，而經過較長一段時間則開始有關聯，圖形下降開始變緩。而在大尺度的 τ 在剛開始時前後係數關聯較大，所以較為平滑(圓頭部份)並且下降也沒那麼快，等到時間夠長較無關聯之後，開始下降迅速，最後無論時間尺度長短圖形皆趨近於一致在零附近震盪。

(2) 在 noMA 的情況，最低點與所選的時間尺度有關，並且在短時間尺度時出現

負值的自相關係數，隨著尺度越大最低點越趨近於零。我們發現，在方法 1 時最低相關係數值接不低於 -0.2，在方法 2 最低相關係數值則會超過 -0.2，另一方面也發現平均的統計方式對資料的影響，下圖是 1996 年 5 月在 noMA 的情況下利用兩種方法畫出來的圖形，可明顯看出有無平均的差別，並且方法本身的統計性，方法 1 是單獨公司的 autocorrelation 再取平均，所受干擾較少；方法 2 是以各家公司分別在不同的交易尺度下當作個別的粒子，由於粒子之間會相互作用，相當於公司之間的彼此影響，所以圖形較為不規律。

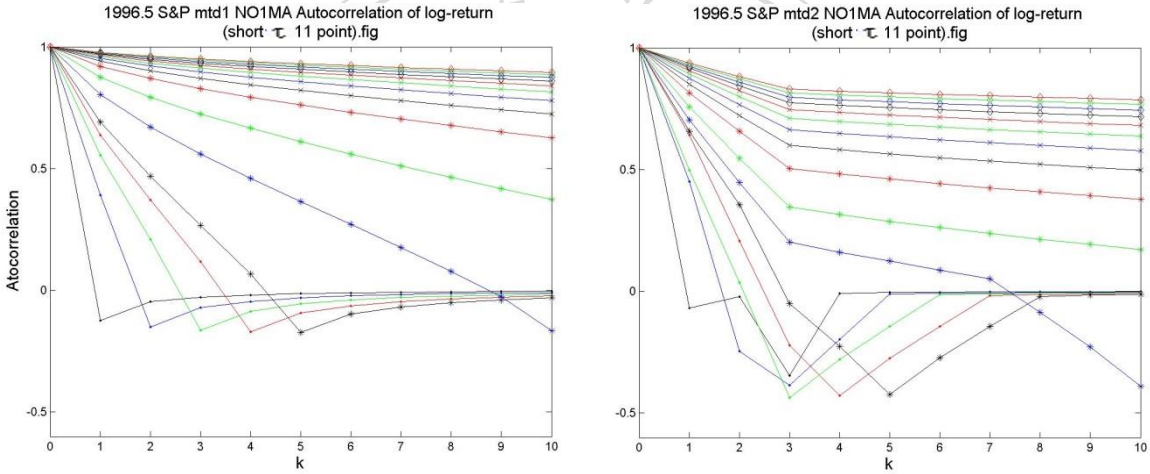


圖 3.5C　1996 年 5 月在 noMA 下方法 1(mtd1)與方法 2(mtd2)的圖形比較，由圖形可看出平均與計算方式不同對圖形的影響。

第四章　結論與建議

我們對以上所討論的幾個主題做簡單的結論與建議，分別討論如下：

一、 股價指數報酬對應於朗之萬理論的實證

我們把對粒子的分析分式用在股票群的系統，而做高頻的移動平均是一個關鍵的手段，在計算上對數據做有系統的分析。利用 HF1MA 的原因是基於 noMA 之下有遲滯運動的發生，接著 HF1MA 後可看出把遲滯項消去的結果。

1. noMA 的數據

在 noMA 在粒子遲滯運動的現象；我們可解讀為在此圖形中，公司在這段時間彼此開始相互影響，可視為流體中粒子與粒子之間開始作用，並與其所在的環境有關，這個環境因素相當於各個公司之間影響程度不一，所以在對數圖形中其所對應的斜率值並不固定；在長時間之後關聯性越來越小，受限因素漸漸不存在後復歸為擴散運動的範疇(在對數圖形下斜率等於 1 部份)，即正比於時間的 1 次方。

2. HF1MA 的數據

粒子移動的軌跡是連續的，而原始數據是離散的形式，所以我們以沒有日移動平均的報酬來計算，其所得到的圖形會與理論上有所差異，當價格經過 HF1MA 之後計算所得到的數據(MSLR)在圖形上符合朗之萬方程的描述，主要原因是經過 HF1MA 之後使資料變平滑，也可以消除遲滯運動的現象，代表交互相關性的記憶被抹掉(把記憶性消除)，所以吻合。另外，由溫度係數與擴散係數可以驗證愛因斯坦關係，並且溫度係數 θ 可以反應 volatility，而與報酬(log-return)較無關係，分別如下圖所示。

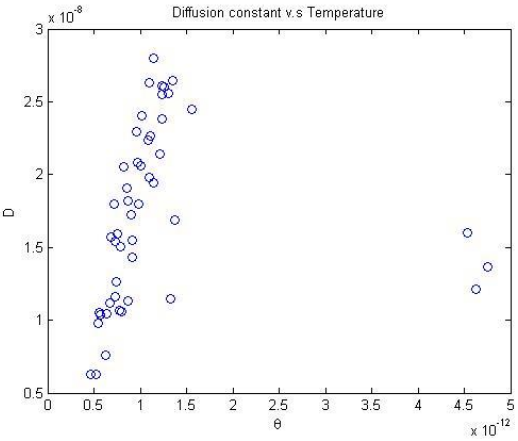


圖 4A　溫度對擴散係數的關係圖。此圖採用圖 3.3.4D 的數據，由圖中可看出溫度對擴散係數存在著線性關係，右邊三個點因為包含 5 月份的數據，波動較大。

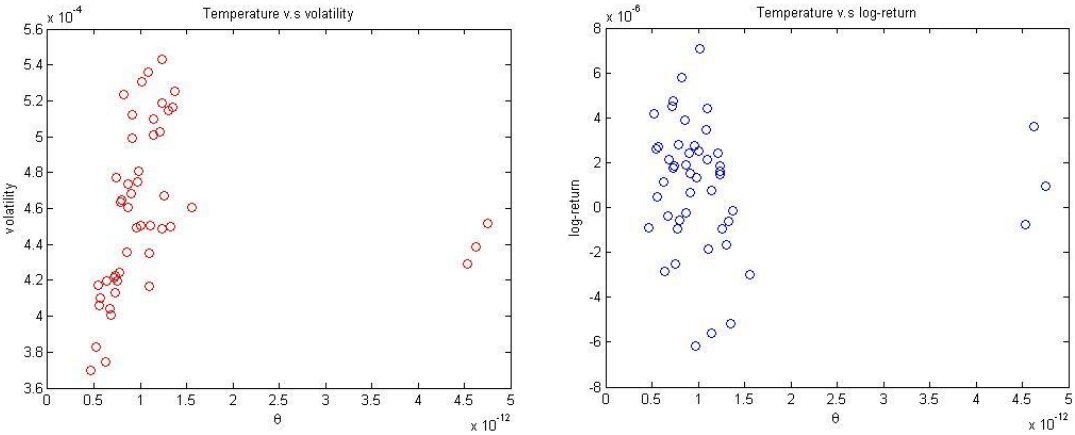


圖 4B　volatility 對溫度的關係圖(左)與 log-return 對溫度的關係圖(右)。由圖可看出溫度係數 θ 可以反應 volatility，而與報酬(log-return)較無關係。右邊三個點因為包含 5 月份的數據，波動較大。

3. 移動平均(MA)的時間長度對影響 MSLR 的影響

另一方面，移動平均(MA)的時間長度，也會影響 MSLR 在短時間之下在對數圖形所對應的斜率，我們發現 1hrMA (1 小時的移動平均)與 1dayMA (一天的移動平均)的斜率不一樣，前者的斜率較後者斜率小(後者斜率為 2)，且移動平均的時間長度越長，其斜率越接近 2，我們也發現其在轉折之處(對數圖下的斜率從 2 轉折到 1 時) 與所選取的移動平均時間有關，若選取的 MA 時間為 1 天，則在 1 天附近開始出現轉折。

4. 建議與未來研究方向

(1) 短時間交互相關性之記憶消除的可能解釋與如何驗證，對於在 HF1MA 的情況下可以符合理論的描述，其另一個推論為：移動平均之後不見得是完全把干擾抹去，有可能是把程度、大小、或影響的範圍都變一樣(標準化或歸一化)，我們可由隨機行走模型上來瞭解交互相關性造成的遲滯運動。

(2) 計算出溫度 T 與擴散係數 D 之後，如何再進一步分析與應用，以瞭解市場上的巨觀性質。

二、 股價指數報酬對應於 Lévy 機率分布的實證

在 HF1MA 與 noMA (即 NO1MA) 的條件之下，股價機率分布的結果顯示：

1. 技術分析方面

(1) 時間序列長短會影響統計性，時間越長統計性越好，在做機率分布圖時，X 軸的尺度大小也會影響統計性，HI 值(間隔個數)越大代表圖形 X 軸刻度越小，容易造成越多的波動，所以計算的結果會侷限在一個範圍內。

(2) 另一方面，計算平均的方式也會影響結果，這裡我們採用兩種計算方式，分別為方法 1 (mtd1)：整體股價計算法與方法 2(mtd2)：市場價格平均法。由於方法 1 的資料數大大增加，所以 HI 值(間隔個數)變大，但在有 HF1MA 的情況下，基於資料已經是”連續”的，所以 X 軸不見得可以分割越細；相反的，在 noMA 的情況下，HI 值(間隔個數)比 HF1MA 的情況下多，其與本身資料的不連續性有關，因為其資料為 0 的個數較多。

(3) 我們觀察 5 家公司(AAPL, AMD, GE, INTC, MSFT)個別的股價指數機率分布性質，在 HF1MA 的情況下，α 值約為 1~1.2 之間，顯示在 P(R=0)的附近有 Lévy 穩定分布的性質；在 noMA 的情況下對數報酬或價格變化的機率分布波動很大，其 α 值皆大於 2，沒有 Lévy 穩定分布的性質，可能的原因是個別公司受到外在的因素干擾，造成序列的不穩定，在 1MA 就可以消除這層影響。

(4) 5 家公司的對數報酬(或價格指數的變化率)長時間來看應該為零，且在

HF1MA 的情況下其 P(R=0)的附近有 Lévy 分布特徵，所以我們假設此分布的特徵是均值 u 為零，圖形不偏 ($\beta = 0$) 的穩定分布，在實際計算上會有些許誤差。然而我們不能保證各個股票之間的對數報酬是獨立同分布 (具有相同的分布形狀和相同的分布參數和機率密度函數)，所以就沒有 Lévy 分布的可加性，我們嘗試的有些方法(文中未列出)所得到的結果不合，可能與此有關。

2. 不同的 X 軸尺度大小與 α 值的變化

我們將各家公司的股價指數在有無移動平均的統計分析下，計算在作機率密度分布時不同的 X 軸尺度大小下 α 值的變化，結果發現在 HF1MA(一天得移動平均)之下以方法 1(mtd1)及方法 2(mtd2)得到的 α 值接近 1，在 noMA(沒有移動平均)之下以方法 2(mtd2)計算得到的 α 值則介於 1.5~1.6 之間，符合穩定分布(stable distribution)的描述。

3. 移動平均的時間長短如何影響 α 值

移動平均的時間長短也會影響 α 值的大小，移動平均的時間長度越短則越接近 noMA 的 α 值，例如：1hrMA(1 小時的移動平均)的 $\alpha \approx 1.2$，1dayMA(1 天的移動平均)的 $\alpha \approx 1$。因此，我們可藉由這項特性了解 S&P500 股價指數的機率分布。

4. 機率密度的厚尾分布特徵

我們已知報酬的機率密度有厚尾分布的特徵，那麼在 HF1MA 作用之下，我們也可看出其斜率近似於 3，符合先前研究。

5. 建議與未來研究方向

(1) 探討在 HF1MA 情況下 $\alpha \approx 1$ 所代表的意義。

(2) 做移動平均的時間長度對 α 值的影響，從 $\alpha \approx 1.6$ 移動到 $\alpha \approx 1$，以及確認超過一天的移動平均之後 α 值是否保持近似為 1，或者超過 1 天的移動平均之後 α 值會變動。例如，作 5 天的移動平均 α 值是否保持為 1。

(3) 在 HF1MA 下不做標準化之後的資料其厚尾分布的特性與資料做標準化之後的比較，以及在 HF1MA 與 noMA 情況下的資料，其厚尾分布的異同。

國立政治大學 National Chengchi University

三、 股價報酬的自我相關性與實證

在股價指數的動力性質方面，因為股票價格沒有速度的概念，於是我們利用價格的變化來類比速度，而速度與時間有關，當我們計算不同交易時間尺度下的報酬自相關，相當於類比不同時間之下的速度自相關，所以我們再加上一個時間項 τ 來表示在不同的時間尺度交易下的自相關函數(TSSA)。接著我們再加上移動平均的條件，發現在 TSSA 的公式下所呈現的 HF1MA 與 noMA 函數分布圖有著一些不同的特徵。由以上實證結果我們可以發現，在 HF1MA 與 noMA 的情況下，不論方法 1(mtd1)的自相關平均法或方法 2(mtd2)的整體股價計算法，圖形都很相似。另外在時間的記憶性方面，歸納如下：

1. 在 noMA 的情況下

(1) 以位移(即對數報酬)的時間相關 $\langle R(t)\,R(t+\Delta t)\rangle$ 對應於粒子系統之速度的時間相關 $\langle v(t)\,v(t+\Delta t)\rangle$。圖形顯示離散軌跡的特徵，其最低點的時間即為所選取的時間尺度 τ，接著回到 0 附近振盪，隨著時間越大慢慢收斂到零。

(2) noMA 的情況下在短 τ 時關聯性不強，與報酬 $R(t)$ 涵蓋的資料範圍有關。τ 越大其 $R(t)$ 所涵蓋的資料範圍越大，可發現相關性越大。

(3) 我們可以發現在函數最大值與最小值之間的點數，基於相關性計算上的結果，其間必有 $\left(\frac{\tau}{36}\right)-1$ 個點。

2. 在 HF1MA 的情況下

(1) 同樣以$\langle R(t)\,R(t+\Delta t)\rangle$ 類比 $\langle v(t)\,v(t+\Delta t)\rangle$，我們看到此函數在移動平均的時間長度(即 1 天)衰減到零，最後回到零值附近振盪，且在 1 天的時間長度內，不同 τ 的曲線彼此不會相交，且由短 τ 到長 τ，函數的衰減由時間 t $(t\sim 0)$ 的一次項的表現轉變為二次項的表現，此顯示平滑軌跡的特徵；相較於離散軌跡的情況，在 1 天之內的曲線隨 τ 的變化相對鈍化。圖形顯示在剛開始是受本身資料性質影響，圖形後端皆慢慢收斂趨於一致，顯示移動平均的作用。

國立政治大學 National Chengchi University

(2) 我們也驗證選取的移動平均時間長度與圖形轉折的位置的確有關，當移動平均的時間長度為 2 天時，轉折的時間也變為在 2 天附近。

由於以上討論我們可以得到一個重要訊息，其顯示在 TSSA 公式下作 HF1MA 對動力性質有影響。而要更瞭解其中細節仍有待日後進一步地分析與探討。

4. 建議與未來研究方向

(1) 在自相關函數方面針對不同的移動時間長度作計算並畫圖，觀察其結果並作更嚴謹的分析。例如以公式來表達在 HF1MA 與 noMA 的情況下，不同時間尺度 τ 的描述，在 noMA 的情況是以離散的觀點表達公式，由 noise 項主導；而 HF1MA 的情況是以連續的觀點來表達公式，可用泰勒展開式來討論。此結果可應用於各種時間序列的分析。

(2) 以此項工作的分析為基礎，進一步對 noMA 的數據建立模型。

國立政治大學 National Chengchi University

附錄

附錄一. Lévy 分布的兩個基本性質

性質一：當 $\alpha < 2$ 時，Lévy 分布不存在變異數。

在(2.3B)式中，當 $\alpha < 2$，$|x|$ 的 n 階動差函數 $(\mathrm{n} > \alpha)$ $|x|^n$ 會發散，即所有 $\alpha < 2$ 的 Lévy 穩定過程 (Lévy stable process) 的變異數會發散。另外，在穩定分布的特例中，所有高斯分布($\alpha = 2$)的動差函數都是有限值。由於 Lévy 分布有厚尾的特徵，所以相應的機率密度函數沒有迅速衰減，所以積分不收斂，即變異數發散或不存在。

性質二：穩定 Lévy 分布的可加性

若隨機變量 ξ、ξ_1 和 ξ_2 為獨立同分布(i.i.d)，對任意正數 b_1 和 b_2，存在正數 b，使得 $\mathrm{b}_1\xi_1 + \mathrm{b}_2\xi_2$ 和 $\mathrm{b}\xi$ 同分布，則稱隨機變量 ξ 具有穩定分布。設 $\varphi(q)$ 是 ξ 的特徵函數，依定義有$\varphi(\mathrm{b}_1 q)\cdot\varphi(\mathrm{b}_2 q) = \varphi(bq)$，且設隨機變數 x 遵守 Lévy 分布，$\varphi(q)$ 是 ξ 的特徵函數，則有

$$\varphi(q) = exp\left[i\mu bq - \gamma b^{\alpha}|q|^{\alpha}\left(1 - i\beta\frac{q}{|q|}\tan\left(\frac{\pi}{2}\alpha\right)\right)\right]$$

$$\varphi(\mathrm{b}_1 q)\cdot\varphi(\mathrm{b}_2 q) = exp\left[i\mu(\mathrm{b}_1 + \mathrm{b}_2)q - \gamma({\mathrm{b}_1}^{\alpha} + {\mathrm{b}_2}^{\alpha})|q|^{\alpha}\left(1 - i\beta\frac{q}{|q|}\tan\left(\frac{\pi}{2}\alpha\right)\right)\right]$$

則條件 $\varphi(\mathrm{b}_1 q)\cdot\varphi(\mathrm{b}_2 q) = \varphi(bq)$ 的成立等價於條件 $\mu b = \mu(\mathrm{b}_1 + \mathrm{b}_2)$ 和 $b^{\alpha} = {\mathrm{b}_1}^{\alpha} + {\mathrm{b}_2}^{\alpha}$ 的成立，即滿足條件 $\mu = 0$ 且 $b^{\alpha} = {\mathrm{b}_1}^{\alpha} + {\mathrm{b}_2}^{\alpha}$。因此，若且唯若當 $\mu = 0$ 時，Lévy 分布是穩定分布，即服從穩定 Lévy 分布的兩個隨機變量之和也服從穩定 Lévy 分布。如果 $X_i \sim S(\alpha, \beta, 0, \gamma_i)$ 相互獨立，那麼 $X \equiv \sum_{i=1}^{N} X_i \sim S(\alpha, \beta, 0, \sum_{i=1}^{N}\gamma_i)$。即相互獨立並服從穩定 Lévy 分布的隨機變量之和仍然服從穩定 Lévy 分布。

附錄二. 二項式定理求平均位移 $\langle x\rangle$ 與均方位移 $\langle x^2\rangle$

$$\langle x\rangle=\sum_{n=0}^{N}RW_N(n)\cdot[n-(N-n)]l=\sum_{n=0}^{N}RW_N(n)\cdot(2n-N)l \quad (2.1.2B)$$

由二項式定理，

$$(p+q)^N=\sum_{n=0}^{N}RW_N(n)=\sum_{n=0}^{N}\binom{N}{n}p^n\mathrm{q}^{N-n}$$

所以

$$\langle x(N)\rangle=\sum_{n=0}^{N}(2n-N)l\cdot RW_N(n)=\sum_{n=0}^{N}(2n-N)l\cdot\binom{N}{n}p^n\mathrm{q}^{N-n}$$

展開可得

$$\sum_{n=0}^{N}2nl\cdot\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n}-\sum_{n=0}^{N}Nl\cdot\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n}$$

對於上述，

$$\text{第一式}=\sum_{n=0}^{N}2nl\cdot\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n}=2l\sum_{n=0}^{N}\frac{N!}{n!\,(N-n)!}np^n\mathrm{q}^{N-n}$$

$$=2l\sum_{n=0}^{N}\frac{N!}{(n-1)!\,(N-n)!}p\cdot p^{n-1}q^{N-n}$$

$$=2pNl\sum_{n=1}^{N-1}\frac{(N-1)!}{(n-1)!\,(N-n)!}p^{n-1}q^{N-n}$$

$$=2pNl\sum_{n=1}^{N-1}\binom{N-1}{n-1}p^{n-1}q^{N-n}$$

令 $\mathrm{y}=n-1$，則 $\mathrm{n}=\mathrm{y}+1$ ，且 $q=1-p$ 代入，可得

$$2pNl\sum_{n=1}^{N-1}\binom{N-1}{y}p^yq^{N-y+1}=2pNl(p+q)^{N-1}=2pNl$$

$$\text{第二式}=\sum_{n=0}^{N}Nl\cdot\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n}$$

$$= Nl\sum_{n=0}^{N}\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n} = Nl$$

合併 $(2.4.1C)$ 與 $(2.4.1D)$ 兩式，則 $(2.4.1B)$ 式可改寫為

$$\langle x(N)\rangle = 2pNl - Nl = Nl(2p-1) = Nl(p-q) \qquad (2.1.2\mathrm{C})$$

對於布朗粒子的隨機行走，有 $\mathrm{p} = \mathrm{q} = \frac{1}{2}$，故可得 $\langle x(N)\rangle = 0$。

對於均方位移 $\langle x^2\rangle$，有如下形式

$$\langle x^2\rangle = \sum_{n=0}^{N} x^2\cdot RW_N(n) = \sum_{n=0}^{N}(2n-N)^2 l^2\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n}$$

$$= \sum_{n=0}^{N}(4n^2-4nl+N^2)l^2\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n} \qquad (2.1.2D)$$

再利用上述的方式計算，其中對於 n^2 項，我們可以先求出$\langle n(n-1)\rangle$

$$\langle n(n-1)\rangle = \sum_{n=0}^{N} n(n-1)\cdot\frac{N!}{n!\,(N-n)!}p^n\mathrm{q}^{N-n}$$

$$= \sum_{n=2}^{N}\frac{N!}{(n-2)!\,(N-n)!}p^n\mathrm{q}^{N-n} = N(N-1)p^2\sum_{n=2}^{N-2}\frac{(N-2)!}{(n-2)!\,(N-n)!}p^{n-2}\mathrm{q}^{N-n}$$

$$= N(N-1)p^2\sum_{n=2}^{N-2}\binom{N-2}{n-2}p^{n-2}q^{N-n}$$

令 $\mathrm{y} = n-2$，則 $\mathrm{n} = \mathrm{y}+2$，且 $q = 1-p$ 代入，可得

$$N(N-1)p^2\sum_{n=2}^{N-2}\binom{N-2}{y}p^y q^{N-y-2}$$

$$= N(N-1)p^2(p+q)^{N-2} = N(N-1)p^2$$

且已知 $\langle n\rangle = Np$，且 $\langle n(n-1)\rangle = \langle n^2\rangle - \langle n\rangle$，則

$$\langle n^2\rangle = \langle n(n-1)\rangle + \langle n\rangle = N(N-1)p^2 - Np$$

將上述結果代入 $(2.1.2D)$ 式，計算後可得結果為

$$\langle x^2\rangle = Nl^2[(N-1)(p-q)^2+1] \qquad (2.1.2E)$$

附錄三. 另一種方式求解均方位移 <Einstein's Derivation of the Diffusion Equation>[15]

τ 為小時間區間，當粒子在 λ 與 $\lambda + d\lambda$ 移動時，其機率密度為 $P(\lambda)$，且 $\lambda = \sqrt{\lambda^2{}_x + \lambda^2{}_y + \lambda^2{}_z}$，$\tau$ 夠小以使粒子的運動與 τ 獨立。如果 $u(\mathbf{r}, t)$ 代表濃度(concentration, i.e. the number of particles per unit volume)，source function $F(\mathbf{r})$ 在時間 $t+\tau$ 所產生的濃度可表示為

$$u(\mathbf{r}, t+\tau) = \int_{-\infty}^{\infty} u(\mathbf{r}+\lambda, t) P(\lambda) d\lambda + F(\mathbf{r}, t) \qquad (A.1)$$

由於 $\tau \ll 1$，由泰勒展開式(Taylor series)可以近似 $u(\mathbf{r}, t+\tau)$ 為

$$u(\mathbf{r}, t+\tau) = u(\mathbf{r}, t) + \tau \frac{\partial}{\partial t} u(\boldsymbol{r}, t) \qquad (A.2)$$

並由泰勒展開式近似 $u(\mathbf{r}+\lambda, t)$ 為

$$\begin{aligned} u(\mathbf{r}+\lambda, t) = u(\mathbf{r}, t) &+ \lambda_x \frac{\partial u}{\partial x} + \lambda_y \frac{\partial u}{\partial y} + \lambda_z \frac{\partial u}{\partial z} \\ &+ \frac{\lambda_x^2}{2!} \frac{\partial^2 u}{\partial x^2} + \frac{\lambda_y^2}{2!} \frac{\partial^2 u}{\partial y^2} + \frac{\lambda_z^2}{2!} \frac{\partial^2 u}{\partial z^2} \\ &+ \lambda_x \lambda_y \frac{\partial^2 u}{\partial x \partial y} + \lambda_x \lambda_z \frac{\partial^2 u}{\partial x \partial z} + \lambda_y \lambda_z \frac{\partial^2 u}{\partial y \partial z} + \cdots \end{aligned} \qquad (A.3)$$

因為 $\tau \ll 1$，所以高階項可以忽略，且粒子移動的距離 λ 也必須要小。(A.1)式可改寫為($u = u(\mathbf{r}, t)$)

$$\begin{aligned} u + \tau \frac{\partial u}{\partial t} &= F(\mathbf{r}, t) + \int_{-\infty}^{\infty} u P(\lambda) d\lambda \\ &+ \int_{-\infty}^{\infty} \left(\lambda_x \frac{\partial u}{\partial x} + \lambda_y \frac{\partial u}{\partial y} + \lambda_z \frac{\partial u}{\partial z} \right) P(\lambda) d\lambda \\ &+ \int_{-\infty}^{\infty} \left(\frac{\lambda_x^2}{2!} \frac{\partial^2 u}{\partial x^2} + \frac{\lambda_y^2}{2!} \frac{\partial^2 u}{\partial y^2} + \frac{\lambda_z^2}{2!} \frac{\partial^2 u}{\partial z^2} \right) P(\lambda) d\lambda \\ &+ \int_{-\infty}^{\infty} \left(\lambda_x \lambda_y \frac{\partial^2 u}{\partial x \partial y} + \lambda_x \lambda_z \frac{\partial^2 u}{\partial x \partial z} + \lambda_y \lambda_z \frac{\partial^2 u}{\partial y \partial z} \right) P(\lambda) d\lambda \end{aligned}$$

假設 $P(\lambda)$ 為標準化(normalized)

$$\int_{-\infty}^{\infty} \mathrm{P}(\lambda)\mathrm{d}\lambda = 1 \Longrightarrow \int_{-\infty}^{\infty} \mathrm{uP}(\lambda)\mathrm{d}\lambda = u \int_{-\infty}^{\infty} \mathrm{P}(\lambda)\mathrm{d}\lambda = u$$

則

$$\tau\frac{\partial u}{\partial t} = \int_{-\infty}^{\infty} \frac{\lambda_x^2}{2}\frac{\partial^2 u}{\partial x^2}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \frac{\lambda_x\lambda_y}{2}\frac{\partial^2 u}{\partial x\partial y}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \frac{\lambda_x\lambda_z}{2}\frac{\partial^2 u}{\partial x\partial z}\mathrm{P}(\lambda)\mathrm{d}\lambda$$

$$+ \int_{-\infty}^{\infty} \frac{\lambda_x\lambda_y}{2}\frac{\partial^2 u}{\partial x\partial y}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \frac{\lambda_y^2}{2}\frac{\partial^2 u}{\partial y^2}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \frac{\lambda_y\lambda_z}{2}\frac{\partial^2 u}{\partial y\partial z}\mathrm{P}(\lambda)\mathrm{d}\lambda$$

$$+ \int_{-\infty}^{\infty} \frac{\lambda_x\lambda_z}{2}\frac{\partial^2 u}{\partial x\partial z}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \frac{\lambda_y\lambda_z}{2}\frac{\partial^2 u}{\partial y\partial z}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \frac{\lambda_z^2}{2}\frac{\partial^2 u}{\partial z^2}\mathrm{P}(\lambda)\mathrm{d}\lambda$$

$$+ \int_{-\infty}^{\infty} \lambda_x\frac{\partial u}{\partial x}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \lambda_y\frac{\partial u}{\partial y}\mathrm{P}(\lambda)\mathrm{d}\lambda + \int_{-\infty}^{\infty} \lambda_z\frac{\partial u}{\partial z}\mathrm{P}(\lambda)\mathrm{d}\lambda + \mathrm{F}(\mathbf{r},\mathrm{t})$$

以矩陣表示之

$$\frac{\partial}{\partial t}\mathrm{u}(\mathbf{r},\mathrm{t}) = \nabla\cdot\mathbf{D}\nabla\mathrm{u}(\mathbf{r},\mathrm{t}) + \mathbf{V}\cdot\nabla\mathrm{u}(\mathbf{r},\mathrm{t}) + F(\boldsymbol{r},t) \qquad \text{(A.4)}$$

D為擴散張量(diffusion tensor)，可表示為

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

其中

$$\mathbf{D}_{ij} = \int_{-\infty}^{\infty} \frac{\lambda_i\lambda_j}{2\tau}P(\lambda)d\lambda = \frac{1}{2\tau}\langle\lambda_i\lambda_j\rangle$$

V 為此群集粒子所可能擁有的漂移速度 (Drift Velocity)

$$\mathbf{V} = \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} \qquad \mathbf{V}_i = \int_{-\infty}^{\infty} \frac{\lambda_i}{\tau}P(\lambda)d\lambda = \frac{1}{\tau}\langle\lambda_i\rangle$$

當 $\lambda_i\lambda_j = \lambda_j\lambda_i$，擴散張量以對角線對稱(即$D_{ij} = D_{ji}$)，對於均向擴散(isotropic diffusion)的情況，$\langle\lambda_i\lambda_j\rangle = 0$，$\mathrm{i} \neq j$ 且 $\langle\lambda_i\lambda_j\rangle = \langle\lambda_i^2\rangle = \langle\lambda_j^2\rangle$，當$i = j$；以及無飄移速度存在，所以 $\mathbf{V} = 0$，所以(A.4)式可改寫為

$$\frac{\partial}{\partial t}\mathrm{u}(\mathbf{r},\mathrm{t}) = \nabla\cdot\begin{pmatrix} D & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{pmatrix}\cdot\nabla\mathrm{u}(\mathbf{r},\mathrm{t}) + F(\boldsymbol{r},t)$$

$$= \mathrm{D}\nabla^2\mathrm{u}(\mathbf{r},\mathrm{t}) + F(\boldsymbol{r},t)$$

其中

$$D = \int_{-\infty}^{\infty}\frac{\lambda^2}{2\tau}P(\lambda)d\lambda = \frac{1}{2\tau}\int_{-\infty}^{\infty}\lambda^2 P(\lambda)d\lambda = \frac{\langle\lambda^2\rangle}{2\tau}$$

可得

$$\langle\lambda^2\rangle = 2D\tau$$

附錄四. Lévy 分布的推導 (Evaluation of the Levy Distribution)[16]

這邊介紹文獻上對於 Lévy 分布的證明，若特徵函數 $P(k) = \exp(-a|k|^{\gamma})$，$0 < \gamma \leq 2$ 等價於機率密度函數 $\mathrm{p(x)}\sim x^{-(1+\gamma)}$，$x \to \infty$。證明：

$$\frac{1}{x^{1+\gamma}} \leftrightarrow e^{(-a|k|^{\gamma})},\ 0 < \gamma \leq 2,\ x \to \infty$$

記號 $\leftrightarrow$ 表示從實部到傅立業空間的轉換

證明：

$0 < \gamma \leq 1$，特徵函數 (characteristic function) 是對稱，則有如下形式

$$\mathrm{p(x)} = \mathrm{Re}[f(x)]$$

其中

$$\mathrm{f(x)} = \frac{1}{\pi}\int_0^{\infty} e^{ikx}e^{-k^{\gamma}}dk$$

$$= \frac{1}{\pi}\left(\left[\frac{1}{ix}e^{ikx}e^{-k^{\gamma}}\right]_{k=0}^{\infty} - \frac{1}{ix}\int_0^{\infty} e^{ikx}\left(-\gamma k^{\gamma-1}e^{-k^{\gamma}}\right)dk\right)$$

$$= \frac{\gamma}{2\pi ix}\int_0^{\infty} dk\, H(k)k^{\gamma-1}e^{-k^{\gamma}}e^{ikx} \quad , \quad x \to \infty$$

其中

$$H(k) = \begin{cases} 1, & k > 0 \\ 0, & k < 0 \end{cases}$$

對於 $0<\gamma<1$，$f(x)$ 在 $k=0$ 是奇異的 (singular)，並且在這區間的最大分布是 $H(k)k^{\gamma-1}$ 的反傅立業 (inverse Fourier form) 形式，可標記為

$$\mathcal{F}^{-1}\left[\frac{1}{(ik)^{\gamma}}\right]\sim\frac{1}{x^{1-\gamma}}$$

$\mathcal{F}^{-1}$ 為反傅立業轉換的標記，並且

$$H(k)\leftrightarrow\delta(x)+\frac{i}{\pi x}\sim\delta(x),\quad x\to\infty$$

接著，使用摺積定理 (convolution theorem)，我們可以得到

$$f(x)\sim\frac{\gamma}{i\pi x}\frac{i^{1-\gamma}}{x^{\gamma}}$$

因此

$$p(x)\sim\frac{1}{x^{1+\gamma}}\ ,\ x\to\infty$$

對於 $1<\gamma<2$，我們可以用兩次分部積分法得到

$$f(x)=\frac{\gamma}{i\pi x}\int_0^{\infty}dk\,k^{\gamma-1}e^{-k^{\gamma}}e^{ikx}$$

$$=\frac{\gamma}{i\pi x}\left[\frac{1}{ix}k^{\gamma-1}e^{-k^{\gamma}}e^{ikx}\right]_{k=0}^{\infty}$$

$$+\frac{\gamma}{\pi x^2}\int_0^{\infty}dk\,e^{ikx}\left[(\gamma-1)k^{\gamma-2}e^{-k^{\gamma}}-\gamma(k^{\gamma-1})^2e^{-k^{\gamma}}\right]$$

$$=\frac{\gamma}{\pi x^2}\int_0^{\infty}dk\,e^{ikx}\left[(\gamma-1)k^{\gamma-2}e^{-k^{\gamma}}-\gamma(k^{\gamma-1})^2e^{-k^{\gamma}}\right],\ x\to\infty$$

結果的第一項是奇異的 (singular)，因此提供最大的貢獻並且我們可以改寫為

$$f(x)\simeq\frac{\gamma(\gamma-1)}{2\pi x^2}\int_{-\infty}^{\infty}H(k)\,e^{ikx}(k^{\gamma-2}e^{-k^{\gamma}})dk$$

以這個例子來說，對於 $1<\gamma<2$，對此積分最大的分貢獻是 $k^{\gamma-2}$ 的反傅立業轉換，因此

$$f(x)\sim\frac{\gamma(\gamma-1)}{\pi x^2}\frac{i^{2-\gamma}}{x^{\gamma-1}}$$

所以

國立政治大學 National Chengchi University

$$p(x) \sim \frac{1}{x^{1+\gamma}}\ ,\ x \to \infty$$

符合先前的對 $\gamma \to 1$ 的近似。

附錄五. 菲克定律

菲克定律由阿道夫•菲克於 1855 年推導出來，用描述擴散作用，可以使用這條定律來求得擴散係數 D。

菲克第一定律

假設從高濃度區域往低濃度流的通量大小與濃度梯度（空間導數）成正比，通過這個假設，菲克第一定律把擴散通量與濃度聯繫起來。在一維空間下的菲克定律如下：

$$\mathrm{J} = -\mathrm{D}\frac{\partial \emptyset}{\partial x}$$

其中

J 為「擴散通量」(於某單位時間內通過某單位面積的物質量)，D 為擴散係數或擴散度，∅ 為濃度（假設為理想混合物），X 為位置其量綱為 長度，例如 m

菲克第二定律

菲克第二定律預測擴散會如何使得濃度隨時間改變：

$$\frac{\partial \emptyset}{\partial t} = \mathrm{D}\frac{\partial^2 \emptyset}{\partial x^2}$$

其中：

∅ 為濃度，t 為時間，D 為擴散係數，X 為位置(長度)

可從菲克第一定律及連續方程式（continuity equation）導出菲克第二定律：(一維情況下)

$$\frac{\partial \emptyset}{\partial t} + \nabla \cdot \mathbf{J} = 0$$

$$\frac{\partial \emptyset}{\partial t} = -\frac{\partial}{\partial x}\mathrm{J} = -\frac{\partial}{\partial x}\left(-\mathrm{D}\frac{\partial \emptyset}{\partial x}\right) = \mathrm{D}\frac{\partial^2 \emptyset}{\partial x^2}$$

上式假設 D 為常數，由此可得上述的菲克方程式。

附錄六. 時間序列資料之平穩性 (stationarity)

若 e_t 為一圍繞在零附近的一個隨機變數，變異數為固定，且具有本期的資料與前期資料無關的特性，也就是 $E(e_t,\ e_s) = 0_{(t \neq s)}$，這樣的數列，我們一般將它稱為白噪音(white noise)。

任一時間序列模型均可由一組獨立同分配 (i.i.d) 的白噪音 $\{e_t\}_{t=1,2,\ldots,\infty}$ 以線性組合而成 $y_t = \mu + \beta_0 e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \cdots$，$e_t \overset{iid}{\sim} N(0, \sigma^2)$

則 y_t 的變異數 $\gamma_0 = Var(y_t) = E(y_t,\ y_t)$

$$E(y_t, y_t) = E[\mu + \beta_0 e_t + \beta_1 e_{t-1} + \beta_2 e_{t-2} + \cdots]$$

$$= {\beta_0}^2 \sigma^2 + {\beta_1}^2 \sigma^2 + {\beta_2}^2 \sigma^2 + \cdots$$

$$\sigma^2 [{\beta_0}^2 + {\beta_1}^2 + {\beta_2}^2 + \cdots]$$

若當 ${\beta_0}^2 + {\beta_1}^2 + {\beta_2}^2 + \cdots = \sum_{i=0}^{\infty} {\beta_i}^2 < \infty$時，則 y_t 即為一平穩的時間序列。

若 $y_t = y_{t-1} + e_t$，例如

$$y_0 = e_0$$

$$y_1 = y_0 + e_1 = e_0 + e_1$$

$$y_2 = y_1 + e_2 = e_0 + e_1 + e_2$$

$$y_3 = y_2 + e_3 = e_0 + e_1 + e_2 + e_3$$

$$\cdots$$

$$y_t = y_{t-1} + e_t = e_0 + e_1 + e_2 + \cdots + e_t$$

則 $Var(y_t) = \sigma^2 + \sigma^2 + \cdots = t \cdot \sigma^2$，此為不平穩的時間序列。

國立政治大學 National Chengchi University

參考文獻

[1]黃文璋(民國 81 年), 布朗運動簡介, 數學傳播, 第 16 卷第 4 期。

[2] R. N. Mantegna, and H. E. Stanley, *Scaling behavior in the dynamics of an economic index*, Nature **376**, 46-49 (1995).

[3] M.P. Beccar Varela - M. Ferraro - S. Jaroszewicz – M.C. Mariani, ”Truncated Levy walks applied to the study of the behavior of Market Indices”

[4] Yanhui Liu, Parameswaran Gopikrishnan, Pierre Cizeau, Martin Meyer, Chung-Kang Peng, and H. Eugene Stanley, *Statistical properties of the volatility of price fluctuations*, Phys. Rev. E. VOL **60** (AUGUST 1999)

[5]陳仁遶(民國 91 年), “布朗運動: 從物理學到財務學”, 數學傳播, 第 26 卷第 1 期, 17-22

[6]陳宣毅(2005年), “布朗運動：從花粉的無規行走到生物與天文”, 物理雙月刊, 廿七卷三期

[7]龐寧寧(2005年), “布朗運動界面成長與擴散現象”, 物理雙月刊, 廿七卷三期

[8]王子瑜、曹恒光(2005 年),”布朗運動、郎之萬方程式、與布朗動力學(Brownian Motion, Langevin Equation, and Brownian Dynamics)”, 物理雙月刊, 廿七卷三期

[9] R. N. Mantegna, and H. E. Stanley, *“Stochastic Process with Ultra-Slow Convergence to a Gaussian The Truncated Levy Flight”*, Phys. Rev. Lett. VOL73, 2946 (NOVENBER 1994).

[10] W.-J. Ma, C.-K. Hu, and R. E. Amritkar. A stochastic dynamic model for stock-stock correlations, Phys. Rev. E 70, 026101 (2004)

[11] Wen-Jong Ma, Shih-Chieh Wang, Chi-Ning Chen, and Chin-Kun Hu, *Crossover behavior of stock returns and mean square displacements of particles governed by the Langevin equation* (2013)

國立政治大學 National Chengchi University

[12] Shih-Chieh Wang, *Cross-correlations in Taiwan stock market – a computational statistical physics approach*, July 2005

[13] 王碩濱，東華大學應用物理研究所

[14] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C. K. Peng, H. E.Stanley, Phys. Rev. E 60, 1390 (1999).

[15] IAENG International Journal of Applied Mathematics, August 2010

[16] IAENG International Journal of Applied Mathematics, August 2010

國立政治大學
National Chengchi University