東海大學資訊管理研究所

碩士學位論文

結合修正後的多尺度熵與多尺度潘凱圖應用於比特幣市場價格動態分析之研究

Bitcoin price dynamics analysis using modified multiscale entropy and multiscale poincaré plots

指導教授： 余心淳 博士

研 究 生： 林耕希 撰

中華民國 108 年 01 月

東海大學資訊管理學系碩士學位

考試委員審定書

資訊管理學系研究所 林耕希 君所提之論文

結合修正後的多尺度熵與多尺度潘凱圖
應用於比特幣市場價格動態分析之研究

經本考試委員會審查，符合碩士資格標準。

學位考試委員會 召集人： 謝明光 （簽章）

委　員： 余心淳

黃宗瑞

鄭辰仰

張伊婷

中　華　民　國　108　年　1　月　11　日

誌謝

這篇畢業論文能夠順利的完成，首先我要大力感謝我的指導老師 余心淳老師，從一開始的選題、資料文獻的蒐集、到最終論文的完成，在各方各面老師都花費了極大的心血，不論是研究的進行到論文的修改，老師都給我許多的建議以及指導。並且除了在論文的研究以外，余老師也給予了我許多幫助，在我低潮的時期給予相當多的鼓勵，也常常詢問我狀態如何。也感謝張伊婷老師，在各方面都不吝指導並且給予我許多的意見以及幫助。老師們的指導讓我順利的完成了這篇論文。

其次我要感謝所有來參與口試的口試委員們，口委給予了我許多意見是一開始在研究中沒有思考到，甚至是從未想到過的東西，感謝口委們的意見，讓我能夠修改並且更加完善我的論文。

再者，要感謝的是所有系上指導過我的老師們，在系上的課程中，我學到了許多的知識，每每在走廊遇到老師時，老師們也總會關心我的近況以及鼓勵我，讓我感受到滿滿的關心。也要感謝我的朋友以及同學們，有你們的陪伴以及幫助，我才有動力能夠完成論文。

最後我要感謝我的家人，是你們把我扶養長大，也教給我了許多人生的道理以及知識。你們是我最堅強的後盾，讓我能夠毫無後顧之憂的認真進行研究，也在我攻讀碩士的路上，給了我最大的支持以及鼓勵，謝謝你們!

論文名稱：結合修正後的多尺度熵與多尺度潘凱圖應用於比特幣市場價格動態分析之研究。

校所名稱：東海大學資訊管理學系研究所

畢業時間：2019年01月

研究生：林耕希　　　　　　　　　　　　　指導教授：余心淳

論文摘要：

比特幣是一種使用區塊鏈技術為基礎的加密虛擬貨幣，因其蘊含與傳統貨幣不同的新穎概念，受到了全球金融市場和媒體的極大關注，比特幣不僅成為金融市場上支付、投資或避險的工具，更成為數位金融科技熱潮下資產數位化的重要實例。本研究的主要目的就是分析比特幣交易市場的複雜度，以便了解比特幣市場交易價格的動態變化中是否存在著富有意義的結構，並研究在世界經濟及比特幣歷史上重大事件的發生，是否會影響到比特幣市場的交易價格。本研究使用了 2010 年 07 月 31 日至 2018 年 07 月 31 日的比特幣價格資料做為樣本，並視比特幣市場交易價格的變動為一個非線性的時間序列。但因比特幣交易價格的資料為一個較短的時間序列，所以本研究中使用修正後的多尺度熵來分析比特幣交易市場中的複雜度，並將其與白雜訊及粉紅雜訊做比較。本研究另外也使用多尺度的潘凱圖來對比特幣交易價格之時間序列進行圖形的分析，以了解其中的自我相似性。而由在修正後的多尺度熵資料分析的結果中顯示，比特幣市場的複雜度非常的低，表示比特幣市場與白雜訊相似，是一個較為隨機的時間序列。再對照多尺度潘凱圖分析的結果，也顯示出比特幣市場價格的自我相似性相當的低，是類似於白雜訊的隨機時間序列，證明我們使用這兩個方法的結果是相互吻合的。在另一項研究結果中我們也發現，在歷史上發生的重大事件，會影響到比特幣市場的活絡程度。整體而言，本論文的研究顯示比特幣交易價格的變動走勢是一個複雜度低、隨機性較高、自我相似度低、未來可預測性也較低的時間序列，同時比特幣市場與交易價格也容易受到事件的影響而波動。

關鍵詞：複雜度、比特幣價格、修正後的多尺度熵、多尺度潘凱圖

Title of Thesis：Bitcoin price dynamics analysis using modified multiscale entropy and multiscale poincaré plots

Name of Institute: Tunghai University, Graduate Institute of Information Management

Graduation Time：01/2019

Student Name：Keng-Hsi Lin　　　　　　　　　　Advisor Name：Hsin-Chun Yu

Abstract：

Bitcoin, a cryptocurrency based on the application of blockchain technology, has attracted great attention from the global financial market and media for its novel concept different from traditional currencies. Bitcoin has become not only an instrument for payment, investment or risk aversion in the financial market, but also an important case of asset digitization in the context of the boom of digital finance technology. The purpose of the study was mainly to analyze the complexity of Bitcoin trading market, to facilitate understanding whether there is meaningful structure in the dynamic changes of Bitcoin market trading price and studying whether the occurrence of significant events in the world economy and the history of Bitcoin will affect Bitcoin market trading price. In the study, Bitcoin price data during the period from 31 July 2010 to 31 July 2018 were used as samples, and changes in Bitcoin market trading price were regarded as a non-linear time series. Bitcoin trading price data were a short-term time series, therefore, modified multiscale entropy was used in the study to analyze the complexity of Bitcoin trading market and compare with white noise and pink noise. In addition, multiscale poincaré plots were used in the study to implement graphic analysis to the time series of Bitcoin trading price, so as to understand the self-similarity herein. However, the results of data analysis using modified multiscale entropy show that the complexity of Bitcoin market is very low, which indicates that Bitcoin market is similar to white noise and is a relatively random time series. Moreover, the results of analysis using multiscale poincaré plots also show that the self-similarity of Bitcoin market price is very low, and is a random time series similar to white noise. Therefore, it has been proven that the results obtained by us applying such two methods are consistent. It is also found in another study that, the significant events occurred in the history will have an impact on the liquidity of Bitcoin market. Overall, the study shows that changes and movements of Bitcoin trading price is a time series featuring a low complexity, a high randomness, a low self-similarity and also a low future predictability. In addition, both Bitcoin market and trading price are prone to be affected by events.

Keywords: Complexity, Bitcoin price, Modified multiscale entropy, Multiscale poincaré plots

TUNGHAI UNIVERSITY 1955

目次

頁次

表次

頁次

TUNGHAI UNIVERSITY 東海大學 1955

圖次

頁次

東海大學 TUNGHAI UNIVERSITY 1955

第一章、 緒論

第一節、研究背景及動機

加密貨幣(Cryptocurrency)的興起，獲得全球金融市場、投資大眾、媒體以及社會關注的焦點。直至今日，全球加密貨幣的種類已經超過了 3000 種，其中第一個也是現在最受歡迎的加密貨幣，便是源自於日裔美籍工程師中本聰(Satoshi Nakamoto)於 2008 年發表了一篇題目為「比特幣：點對點的網路電子現金系統」(Bitcoin: A Peer-to-Peer Electronic Cash System)的論文[1]中所推出的比特幣(Bitcoin)電子現金交易系統。比特幣有著去中心化的特性並以點對點的形式進行電子現金交易，因此比特幣有著交易速度快速、成本低的特色。另外比特幣為了避免通貨膨脹的問題，從一上市開始就已經固定了貨幣的供應量，其發行總量不會超過 2100 萬枚[2]，不像各國的法定貨幣可以依照經濟現況來調節。現今比特幣不僅已成為金融市場上支付、投資或避險的工具，更成為數位金融科技熱潮下資產數位化的重要實例。

比特幣是一種使用區塊鏈(Blockchain)為基礎的加密虛擬貨幣，大眾經由「挖礦」(Mining)的行為來獲取比特幣。中本聰於 2009 年推出了第一個比特幣的客戶端軟體，並且經過挖礦之後獲得了世界上第一批的 50 個比特幣[2]。而在 2010 年產生了第一筆以比特幣進行的交易，是由一位在美國佛羅里達州的用戶 Andrey Nikolaevich Kolmogorov 以 10000 枚比特幣購買了一個價值 25 美元的披薩，方才首次開啟了比特幣的交易，並成為虛擬通貨交易的一環[2]。這些加密貨幣與傳統貨幣蘊含著不同的新穎概念，比特幣的價值並不是源於各國政府的中央銀行或企業對其提供價值保障，而是以民眾及商家對其價值的看法決定。傳統貨幣的好處是，有政府的中央銀行提供保障價值，而央行也可以視情況發行或者回收貨幣，以控制匯率，但缺點就是若發生類似通貨膨脹或金融風暴的情形，傳統貨幣也會受到打擊。

由於比特幣的交易是無法撤回的，所以接受交易的零售商也不會因為被欺騙而導致收不到款項的情事發生，因此有許多企業與網路平台願意接受使用比特幣來進行交易與付款，像是微軟、Paypal、及 DELL 等大型公司，當然也有企業及商家是不承認與接受比特幣作為交易貨幣。透過這些接受比特幣企業的推波，讓比特幣逐漸融入了主流的金融市場，並在貨幣市場中取得一定的地位，使其能夠讓全球

各地的金融交易更加便利快捷，也使得從未謀面的客戶充分信任彼此，並安心的進行交易。比特幣具有著長遠的發展前景，它是一個類似傳統貨幣卻又不同於傳統貨幣的一種新型虛擬貨幣，能夠給人們帶來一個更快、更安全也更有效率的電子支付系統。

比特幣剛誕生之時，目的是創造一個線上版本的貨幣，讓政府或銀行無法對交易進行惡意的干預，有些人認為當年市場以紙幣取代黃金時，人們也都抱持著懷疑的態度，而現今的比特幣就類似當年的紙幣，將來能夠作為在數字貨幣 2.0(或稱為加密貨幣 2.0)時代進入經濟的市場，成為未來貨幣的合法交易媒介。但也有些人認為，比特幣無法作為貨幣踏入經濟市場，因為其波動性太大，讓人覺得這是一種投機性的投資，而預期比特幣終究會走向泡沫化。

對於比特幣，各國政府態度不一，這樣新興的一樣東西，讓許多國家在探索如何對其進行約束，又有些國家支持比特幣在虛擬通貨交易市場的發展，並立法承認。例如在美國，政府將比特幣當作合法，並且受到聯邦證券法管理。美國監管機構美國商品期貨交易委員會（Commodity Futures Trading Commission, CFTC）甚至核准美國芝加哥選擇權交易所（Chicago Board Options Exchange, CBOE）於 2017 年 12 月 10 日首度開放比特幣期貨交易，成為全球第一個推出比特幣商品的大型交易所；在歐洲各國，如德國、英國、法國等國家，皆承認比特幣合法，並將虛擬貨幣納入法規保護，也針對比特幣進行稅收；在日本也修法承認比特幣在法律上的定位。另一方面有些國家反對比特幣的擴張，並下達禁令，像是在中國市場就對虛擬貨幣不是這麼友善了，中國央行認為比特幣的使用存在風險，並認為其是一種虛擬商品，不具有與貨幣相同之法律地位，不能也不應當在市場上作為貨幣流通。中國的金融監管機構在 2017 年更是全面封殺比特幣，禁止各大交易所進行比特幣的交易。在印尼央行也發布了法規，宣布禁止使用包含比特幣在內的加密貨幣作為支付工具。

比特幣市值從誕生開始一路的平緩，幾乎沒有什麼太大的價格波動，而在 2017 年開始才大幅上漲，最高的市值曾在 2017 年 12 月時到達了 3200 億美元的驚人數字，24 小時內的交易額也到達了幾十億，其帶動了其他的加密貨幣在區塊鏈產業交易市場的蓬勃發展，如以太幣(Ethereum)、比特幣現金（Bitcoin Cash）、瑞波幣（Ripple）與萊特幣（Litecoin）等。比特幣發展至今，在短短幾年的時間內飆漲，表面看似風光，但是其中也包含了許許多多的風險，因為交易是匿名的，也具有不受政府控管難以追蹤的特性，所以很容易流為犯罪的工具；另外在安全性上也有隱

憂，曾經有被駭客入侵，盜取用戶大筆比特幣的案例，導致許多的比特幣用戶損失嚴重。

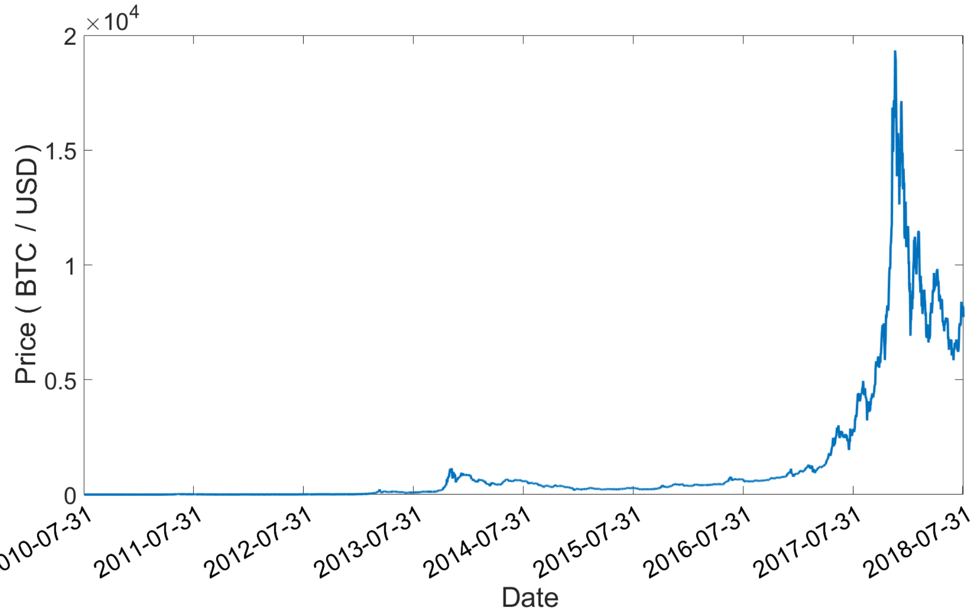


圖 1- 1 比特幣交易價格變動的時間序列(2010/07/31~2018/07/31)

在圖 1-1 中顯示自 2010 年 7 月 31 日至 2018 年 7 月 31 日這段期間內比特幣市場交易價格(以一比特幣價值多少美金為單位)在時間序列上的變動情形。從圖中我們不難看出比特幣的發展，特別是在交易價格波動之處的附近，均有發生過幾個對比特幣意義重大的事件，像是 2013 年 4 月發生的歐洲塞浦路斯事件導致價格波動；2013 年 11 月美國參議院的聽證會上，首次稱比特幣為一樣合法的金融工具，讓比特幣價格上升；2014 年 2 月 Mt. Gox 被盜事件使得原本向上的價格止住上升並下降了；2017 年 5 月的 WannaCrypt 勒索病毒事件，使得許多人知道了比特幣的存在，讓其價格在短短幾個月內飆升了好幾倍；之後在 2018 年 1 月又有許多企業與網路平台宣布禁止使用加密貨幣，讓原本狂升不止的比特幣價格又因此急遽下降，這些事件都在一定程度內影響著比特幣的價格升降，是值得我們探討的因素。

比特幣是一種具革命性的科技，具有著徹底改變銀行業務與商業經營模式的潛力，而且可能使得全世界的人們進入現代數位化、全球化的整合經濟。如果這場貨幣革命真的能夠成功，將會顛覆許多原本令大家視為理所當然的事情，也會對現有的系統造成嚴重的威脅。第一，虛擬貨幣能夠取代目前金融機構進行的大部分業務，必然能夠對金融機構造成極大的影響。第二，虛擬貨幣的交易具有著匿名性，

政府可能就無法掌握所有的交易，若虛擬貨幣的交易量擴大，會有著阻礙稅務徵收的風險存在。第三，虛擬貨幣的出現會造成資金外逃(Capital Flight)的問題，當國家的人民不再信任自己國家貨幣的未來發展，將導致購買比特幣這種較為穩定的貨幣來避險的行為發生。綜上所述，本論文認為虛擬貨幣的議題相當值得關注以及對其進行探討研究。

第二節、研究目的及目標

本研究的方法會使用修正後的多尺度熵(Modified Multiscale Entropy, 以下簡稱 *ModMSE*)演算方法，來對比特幣市場時間序列的交易資料中進行複雜度的計算與研究分析。藉以了解在比特幣市場中是否存在著富有意義的變動模式與結構，以看出其市場是否具有著規律性，並且探討於比特幣歷史上發生的重大事件對於比特幣市場交易價格是否在一定程度上會造成影響，再透過 *MSP* 進行分析，以印證我們使用 *ModMSE* 方式分析出的結果。我們希望能夠在此論文的研究中，找到比特幣市場中的結構，即其是否存在著規律性，以及在經濟上或比特幣歷史上發生的重大事件，是否對於比特幣的市場有著影響。

第三節、相關文獻探討

關於比特幣的交易市場，過去曾經有過許多學者進行過各種不同類型的研究，在 Böhme 等人的研究[3]中，清楚的介紹了對比特幣網路的調查，包含了比特幣貨幣的創建、治理、及存在其中的風險等的詳細背景，以及比特幣未來的發展與應用方向。Lahmiri 及 Bekiros 的研究[4]中，使用混沌理論(Chaos Theory)來分析比特幣市場的價格，他們以 2013 年 2 月 26 日為界，將比特幣價格分為平穩以及較為動盪的兩個時期，他們發現在比特幣市場的不確定性非常高，認為比特幣市場有著高度投機的特性。Kim 的研究[5]以 16 種不同的貨幣與比特幣的交易，探討比特幣在市場上的交易成本，他們發現因為比特幣僅在網路空間中推出並交易，無須其他複雜的交易系統，導致比特幣在外匯市場中成本較低，能夠作為將一種貨幣轉換為另一種其他的貨幣的中介交易貨幣，而銀行或信用卡公司為了尋求利用這種低成本的交易，會助長這種新型貨幣在市場上的發展。Baur 等人的研究[6]中，比較了比特幣與傳統貨幣之中美金以及黃金之間的差異性，他們發現到比特幣中的收益率、波動性等特徵與黃金及普通的法定貨幣(如美金)有很大的不同，比起美金跟黃金，

在比特幣這種加密貨幣上具有獨特的風險收益特徵，波動性高於其他兩者許多。

Yonghong 等人在他們的研究中[7]使用赫斯特指數(Hurst Exponents)及標準差的計算，探討比特幣虛擬貨幣市場是否會隨著時間而帶有長期記憶(Long-Term Memory)的特性。而他們發現到比特幣市場中存在長期記憶的現象，表示可以由比特幣過去的市價，來推斷未來的價格，但是其市場的效率低落，是一個風險較高的市場。在 Demir 等人的研究[8]中發現了經濟政策的不確定性指數(Economic Policy Uncertainty, EPU)能夠影響比特幣市場的發展，結論是 EPU 的變化與比特幣市場呈現負相關(Negative Correlation)的關係，所以 EPU 同時也具有預測比特幣市場的能力。Urquhart 的研究中[9]則是探索為什麼比特幣會受到經濟市場如此大的關注，他使用 Google 的趨勢搜尋資料來作為代表投資者關注的變數，以進一步研究 Google 趨勢與比特幣交易量之間的關係。研究結果發現，比特幣的波動性及交易量皆會影響到在第二天 Google 關鍵字的搜尋次數，據此推論為投資者對比特幣市場活動關注行為之間的關聯性。

在 Bouri, Shahzad 及 Roubaud 等三人的研究中[10]，進行了包含比特幣(Bitcoin)、瑞波幣(Ripple)、以太幣(Ethereum)、萊特幣(Litecoin)、新經幣(Nem)、達世幣(Dash)及恆星幣(Stellar)等 7 種加密貨幣交易市場間波動的研究，他們認為各種加密貨幣市場間存在著相關性，而結果發現各種加密貨幣的市場間會互相影響，當其中一種加密貨幣的市價呈現爆炸性成長的態勢，則其它加密貨幣的交易價格也會受到影響而成長。而 Giudici 與 Abu-Hashish 的研究[11]中討論了影響比特幣市場的要素，發現比特幣與美金及人民幣的兌換匯率之間有正相關(Positive Correlation)的關係，但是比特幣的市價與黃金及石油價格並無關係。在文獻[12][13][14][15]中，作者們則是分別使用了各種不同的方式去評估比特幣市場的效率。[12]中提出使用多重分形(Multifractality)研究比特幣的時間序列，接著[13]中提出使用長期記憶估算器(Long-Range Dependence Estimators)來分析，[14]中使用冪律指數(Power-Law Exponent)進行分析，而[15]中則使用效率指數(Efficiency Index)來評估比特幣市場的效率。

從 2009 至今 2018 年，比特幣出現已經經過了 10 個年頭，在這麼長的時間中，其相關交易市場也是經歷了許許多多高度的變化，並蘊含許多豐富的資訊在內。本論文的研究目的之一是想要分析比特幣的交易市場與歷史中蘊含著多少的資訊或訊息量。資訊是代表著一個東西或現象能傳達出的訊息有多少，以二進制編碼的資

料為例，當一串二進制編碼的資料內容全部都是 0 的時候，因為其最整齊與一致，因此我們所得到的資訊與訊息量是最少的；相反的，一串完全隨機的二進制編碼序列，則是資訊量最大的模式。因此我們可以知道，所謂一個東西資訊與訊息量的多寡，也就是它在特徵表現上的隨機性有多高。在一個時間序列中資料內容越不隨機，代表它越有跡可循，也代表其具有可預測性。熵是度量的工具，是用來計算與度量一個系統是穩定、有序、單調且具結構的狀態，還是相反地呈現隨機、無序、混亂且缺乏結構的狀態。因此本論文將使用熵的計算方式，來對比特幣的交易市場的變動情形進行分析，探索其在時間上的非線性變化，是否具有結構性，亦或是完全隨機的一個結果。

在自然界中的生態、系統等事物狀態的演進，都是漸漸地從有序變為無序的過程，而熵就是用來量化並描述這類型的問題而產生的，它是一種對於混亂(Disorder)與複雜度(Complexity)的度量。比方說在打撞球時，一開始所有的球都是整齊地擺放，此時熵的值就很低，但一開始打球，球就會四散在球檯上的各處，球檯上的混亂程度就是增加了，代表的就是整個球檯上的熵值增加了，我們很難一桿將球打回最初開始之前在球檯上整齊的樣子，這樣的機率是微乎其微的。這證明了在一個系統中，隨著時間經過，其熵的值是只增不減的。

熵是對不確定性的測量，熵越高，代表能夠傳達越多的資訊，反之熵則越低，則表示代表的訊息與資訊量較少。訊息熵的概念源於 1948 年提出的 Shannon 熵[16]，後來的發展演變出許多不同的熵量測方式，如 1991 年 Pincus 提出的近似熵(Approximate Entropy, *ApEn*)[17]，用以量化一個時間序列資料的規律性以及預測性；Richman 在 2000 年將近似熵改良後提出的樣本熵(Sample Entropy, *SampEn*)[18]，其改進了近似熵中自我配對的問題，讓計算能夠更加精確；Costa 於 2002 年提出的在由多個尺度的樣本熵來研究一個時間序列資料特性的多尺度熵(Multiscale Entropy, *MSE*)[19]，它可以用來量化系統的複雜度；2013 年 Wu 等人將多尺度熵改良後能夠分析更小時間序列，並且更加精確的修正多尺度熵(Modified Multiscale Entropy, *ModMSE*)[20]，其改變了原本 *MSE* 的粗粒化(Coarse Graining)呈現方式以去除雜訊，進而達到更精確的計算。雖然上述這些熵計算的方式均有差異，但都可以用來分析與量化一個時間序列數據上的可預測性與複雜度。

在 Gulko 1999 年的研究[21]中，首次提出了將熵應用於分析金融市場的時間序列為開端，其使用了最大熵(Maximum Entropy)的方式對金融市場進行分析，後來

更衍生出了許多使用不同熵的計算方法來分析金融市場的研究。而 Darbellay 與 Wuertz 的研究[22]，使用夏儂熵(Shannon Entropy)的演算法來分析了多個不同的金融時間序列，並且證實了熵在金融市場分析的可行性。2018 年 Li 與 Shang 的研究[23]中，使用了複雜性分析來分析在金融、股票市場中的複雜度，以及價格在時間序列上的波動性，他們發現了歷史價格與未來波動之間的相關性，其呈現負相關的關係，並且找到了在金融市場的時間序列變動資料中所存在的規則。2014 年 Xia 等人的研究[24]中，使用了多尺度熵 *MSE* 來分析各區域的金融時間序列，他們發現歐洲金融市場的熵值低於亞洲卻高於美洲，了解到使用 *MSE* 的演算方式能夠有效的探索與分類出世界上不同區域金融市場交易的時間序列資料在特徵上的差異，並進而提升後續研究分析的正確性。

而熵在其它的方面也有不同的應用，Martina 等人在 2011 年的研究[25]中，使用多尺度熵 *MSE* 來監測原油(Crude Oil)價格走向的趨勢變化，藉由 *MSE* 能夠測量市場動態中的複雜度及多樣性的優勢，並找出幾個時間軸上重大的社會事件與經濟事件，進而將其與原油市場交易的熵值相互比對與分析，在結果中發現使用 *MSE* 的熵計算方法能夠找到原油市場的非線性變動結構與社會事件和經濟事件之間的相關性，並且在短期市場中的效率高於長期市場。Ortiz-Cruz 等人在 2012 年的研究[26]中，也同樣使用 *MSE* 研究了原油市場的複雜度以及效率，他們發現原油市場的複雜度取決於時間尺度的規模，在小時間的尺度上熵值較高，而較大的時間尺度則較低，這代表由長期來看，原油價格的可預測性較高。

在另一個應用研究方面，Fan 等人於 2016 年的研究[27]中，使用多尺度熵 *MSE* 來分析歐盟第三階段碳市場(Carbon Market)的複雜度、以及整個市場的內在演化，並且研究碳市場價格的明顯波動，是否與社會政治事件相吻合。而由研究結果顯示，碳市場的價格在長期看來有較高的可預測性，並且碳市場的複雜度與各類社會事件(如氣候會議、金融危機等)的發生相對應。而 Yin 等人在 2018 年的研究[28]中，使用改良後的多尺度熵 *ModMSE* 研究中國幾大城市(北京、重慶、天津、上海、深圳、湖北、廣東)的碳市場交易特性，並以複雜度來分析市場的效率，發現中國的碳市場整體複雜度遠低於歐洲碳市場。

在時間序列資料的分析上，使用熵計算的方式還可搭配其它演算方式來提升資料探索分析的深度與準確度。例如，在 2016 年 Henriques 等人的研究[29]中，配合時間序列資料在多尺度間隔的使用，研究發展出了多尺度的潘凱圖(Multiscale

Poincaré Plot, 以下簡稱 *MSP*)視覺化分析方式，來對一個人的心跳間距數據進行分析，以分析對方是否患有慢性心臟衰竭(Chronic Heart Failure)與心房顫動(Atrial fibrillation)的疾病，並且發現此圖形能夠在時間序列的研究中起到作用。在 2017 年 Humeau-Heurtier 等人的研究[30]中，也使用了 *MSP* 對白雜訊(White Noise)及粉紅雜訊(Pink Noise, 或稱 1/f Noise)來進行分析，並且了解到當潘凱圖(Poincaré Plot)上的圖形呈現圓形形狀的散布時，代表其為不相關時間序列資料的常態分布(Normal Distribution)；而呈現橢圓形形狀時，則顯示相鄰尺度的樣本之間的正相關(Positive Correlation)關係。而這樣的 *MSP* 就是由常用於生物醫學領域的潘凱圖所演化改良出來的，其用以分析心跳間隔的時間序列，並以此畫出圖像來判斷此患者的身體健康是否出現了異常，在 Stein 與 Reddy 於 2005 年發表的研究[31]中，使用了潘凱圖來分析心臟病患者的心跳變異率(Heart Rate Variability)，潘凱圖擁有簡單運算和適合短時間資料分析的優點，結果顯示病患的生理參數經運算後，透過潘凱圖中所呈現的散布模式可以正確的指出心臟是否發生病變。

本論文的第一章為介紹論文的大綱及介紹本研究之研究目的與研究方法，也敘述了我們所預期能夠達成的研究結果，並且對過去所做之文獻研究探討，了解過去學者們進行過什麼相關方面的研究，以幫助了解本研究之研究背景；第二章將對在本研究中所需使用的分析方法進行詳細的介紹，包含了各種熵的計算與應用，以及潘凱圖的理論、特性、使用的參數、計算方式和應用的方式等等；第三章介紹研究使用之數據來源，以及如何處理原始的數據，並且詳述研究流程與步驟，最後說明資料分析的方式；第四章討論研究結果，對研究結果加以分析以得出結論，並對我們的研究結果與過去的研究進行比對；最後，第五章對本論文研究做一個總結，並說明本研究成果的貢獻。

第二章、 複雜度與熵的計算

本章節主要介紹本論文所使用之研究方法。本研究專注於從系統複雜度的角度來研究比特幣交易市場價格變動的結構，而要研究複雜度的問題，我們使用熵的計算來量測與檢視一個時間序列資料的複雜度高低與計算系統中無序和不確定性的方式，並且其可以依照計算的尺度分為單尺度熵(Single-Scale Entropy)以及多尺度熵(Multi-Scale Entropy)二種不同的模式來處理。在此我們首先將介紹複雜度之定義，還有各種不同類型之熵計算方法的意義以及其計算方式，最後還有潘凱圖(Poincaré Plot)的計算以及應用方式。

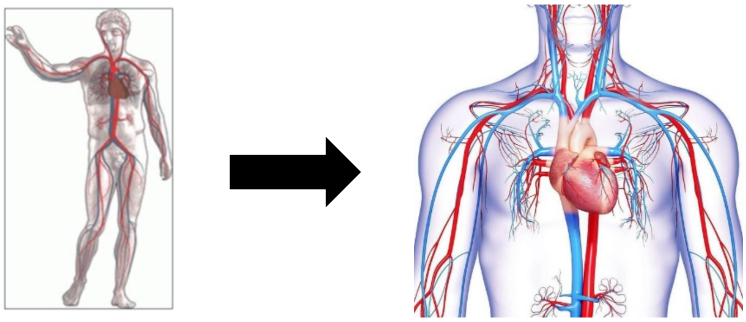
第一節、複雜度(Complexity)

如果一個系統的動態變化可以在時間軸上依序的取樣和紀錄下來，那複雜度便是一種可以用來估測時間序列資料的混亂程度或規律性的方式，數學家柯爾莫哥洛夫(Andrey Nikolaevich Kolmogorov)定義一個系統的複雜程度，便是衡量與描述這個系統所需要信息量的一個尺度。所謂的尺度(Scale)，指的是觀看一個系統在時間或空間上的規模，或是看時間序列資料在時間軸上的解析度，尺度小代表我們將系統中的資料一一檢視；而尺度大則代表著綜觀系統的整體。若一個時間序列的複雜度越高，表示生成此數據序列背後的系統，有著較高的自我相似性。

在生理醫學相關的研究中，經常使用熵來量測與描述生物體的系統複雜度與疾病特徵。例如，大部分的病理研究上顯示較為健康的人其系統複雜度是較高的，而身體不健康甚至有疾病的人其系統的複雜度較低；另外在年輕人與老年人的生理訊號比較中，年輕人的複雜度也較高，表示身體活動性、協調性與機能較好，而老年人則反之。總結在生物系統中，複雜度反映了其物種在環境中的適應力，複雜度傾向愈高的物種，可以在不斷變化的環境中有更快的調適性、控制性、穩定性與生存性，而隨時間保持不變，不易被環境所影響。

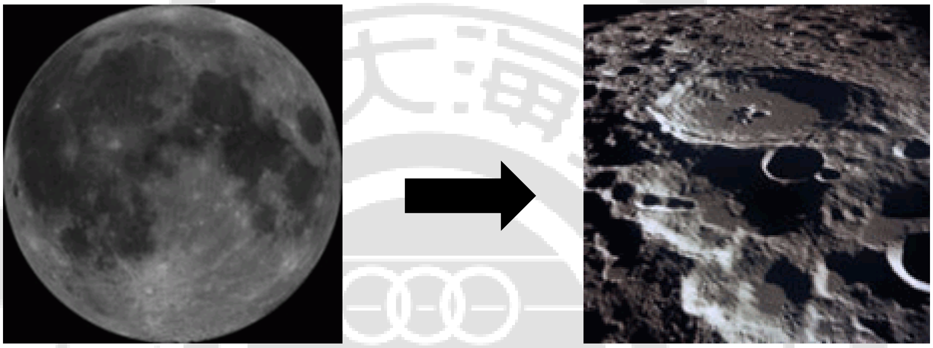
複雜度跟所謂的自我相似程度(Self-Similarity)有著直接的關係，以生物體的血管分布為例，在圖 2-1 中我們可以看見，在大尺度下觀察整個生物體的血管分布時，會發現血管都是由幾條大血管分成許多小血管的方式將血液傳至全身；而當我們把尺度縮小，只看單一器官時，也可以發現血管是從大分至小的方式供給血液。不論在大小尺度下，血管的分布都相當的雷同，這代表著自我相似程度高，這樣的

複雜度也較高。



大尺度　　　　　　　　　　　　小尺度

圖 2- 1 由尺度大小介紹複雜度



大尺度　　　　　　　　　　　　小尺度

圖 2- 2 以月球表面說明混亂

(出處: 複雜度計算: 以 Visual Signal 分析人體站立時平衡度之案例)

另一方面，複雜與混亂(Confusion)是不同的意涵。舉例來說，我們由圖 2-2 來看可以了解，以月球表面來看，在大尺度下觀看整個月球，因為距離遙遠的關係我們看見的月球表面看似相當的光滑，但當在小尺度下觀看時，會發現月球表面充滿坑洞，明顯的凹凸不平。在不同的尺度下觀看，並不具任何的自我相似性，這便不是複雜，而就只是混亂而已。因此，一個系統越混亂就表示此系統在時間或空間上具有自我相似性的表徵就越低，即低複雜度的情形就會越明顯。

圖 2-3 表示著傳統的熵觀念，若熵值越高，則表示這一個系統是一個越規則、有規律的情形，而若是熵值越低，則表示此系統越混亂且不規則。

而圖 2-4 中則顯示的是新的複雜度的觀念，複雜度越高，表示整個系統越健康，而系統太過於規則或是太過混亂都屬於複雜度低，也就是這個系統屬於一個不健康的情形。

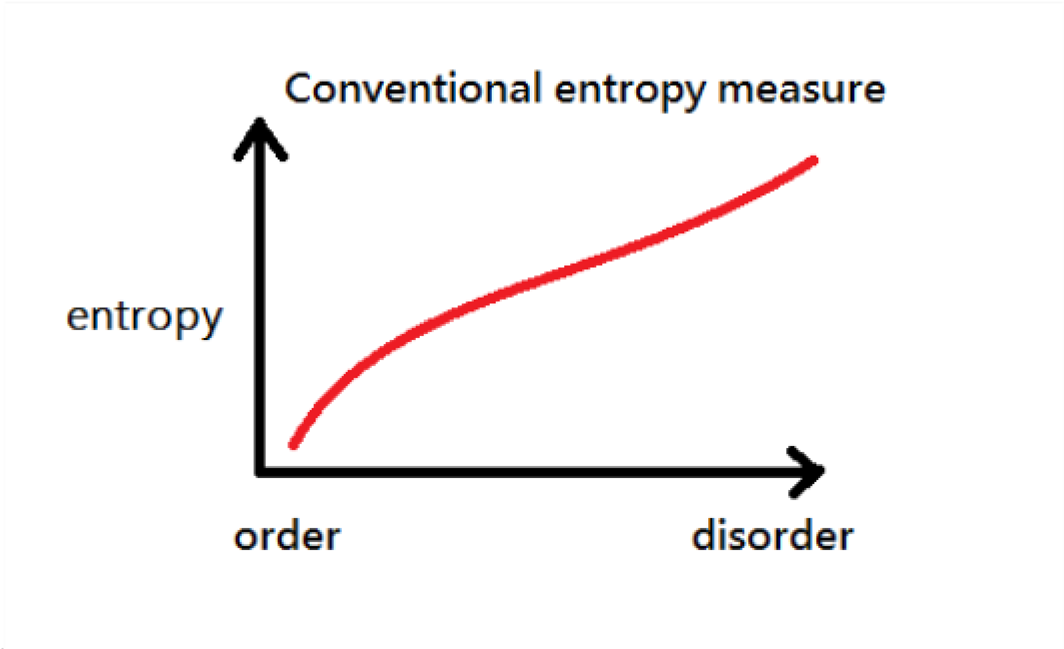


圖 2- 3 傳統的熵觀念

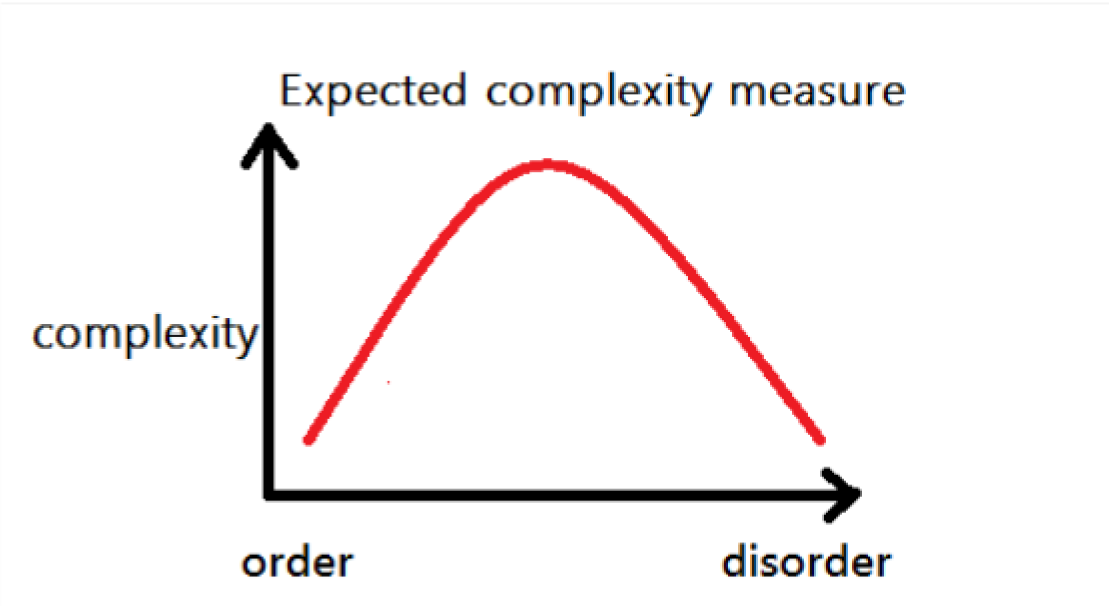


圖 2- 4 複雜度的觀念

(出處: 逸奇科技: 時頻分析與生醫訊號)

第二節 、單尺度熵(Single-Scale Entropy, SSE)

單尺度熵可以用來測量一個時間序列資料的規律性，並且瞭解此序列是否混亂，並且隱含著可預測性。而近似熵(Approximate Entropy)及樣本熵(Sample Entropy)

便是屬於單尺度熵的兩種計算方式，常應用於生物醫學訊號的分析處理。分述如下：

1. 近似熵 (Approximate Entropy, ApEn)

近似熵是由 Pincus 在 1991 年時提出[17]，用來衡量不同維度下時間序列的隨機程度，可用來描述時間序列中的不規則性和複雜程度。如果時間序列的近似熵 *ApEn* 值越大，代表序列的複雜度越大，不規則性也越強。近似熵以 *ApEn*(*m, r, N*)表示，維度為 *m* 及 *m+1*，而相似容限(Similarity Criterion)為 *r* (*r*>0)，相似容限代表的是在時間序列的兩資料比較，為判斷此兩資料點是否為相似的依據，數據長度為 *N*。

我們以時間序列 $X_i = \{x_1,\ x_2, \ldots,\ x_N\}$為例，其近似熵 *ApEn*(*m, r, N*)的計算方式如下，將總長度為 *N* 的時間序列X_i，組成維度為 *m* 的向量X_i^m，其中 $m \leq n$，則

$$X_i^m = \{x_i,\ x_{i+1}, \ldots,\ x_{i-m+1}\},\ i = 1,\ 2, \ldots,\ N - m + 1$$

首先計算各個 *i* 值之間的最大差距值，為了比較兩個數據間的差距，我們以 X_i^m代表序列中以 *m* 為維度第 *i* 筆資料，$X_i^m(i = 1, 2, \ldots,\ N - m + 1)$，並且以 X_j^m代表序列中以 *m* 為維度的第 *j* 筆資料 $X_j^m(j = 1, 2, \ldots,\ N - m + 1)$。另外，我們用 *d* [*X, Y*] 來代表 *X* 與 *Y* 兩者之間的距離，其定義為

$$d[X_i^m,\ X_j^m] = \max|x_{i+k} - x_{j+k}|,\ k = 0,\ 1, \ldots,\ m - 1$$

若其距離 $d[X_i^m,\ X_j^m]$ 小於我們所設定的可容許誤差範圍 *r*，則代表 X_i^m 與 X_j^m 二者之間是有關係的，於是可計算出所有 $d[X_i^m, X_j^m] < r$ 的個數的比率 $P_i^m(r)$，

$$P_i^m(r) = \frac{1}{N+m-1} num\{d[X(i),\ X(j) < r]\}, i = 1,\ 2, \ldots,\ N - m + 1$$

接下來將$P_i^m(r)$取對數，並且將所有的$P_i^m(r)$加總計算其平均$P^m(r)$

$$P^m(r) = \frac{1}{N+m-1} \sum_{i=1}^{N-m+1} \ln P_i^m(r)$$

將維度改為 *m*+1，並重複上述步驟，得到$P^{m+1}(r)$，最後便能得到近似熵值為

$$ApEn(m,\ r) = \lim_{n \to \infty} [P^m(r) - P^{m+1}(r)]$$

而在實際估算時間序列中，將 *N* 取有限值後，可以計算出近似熵的值為

$$ApEn(m,\ r,\ N) = P^m(r) - P^{m+1}(r)$$

在近似熵的參數設定中，Pincus 建議以 *N* > 500，*m* = 2，而且 *r* 的值介於 0.1σ 與 0.25σ 之間(σ為整個時間序列之標準差)會得到較佳的計算結果[17]。

2. 樣本熵(Sample Entropy, SampEn)

樣本熵由 Richman 在 2000 年時，將 Pincus 的近似熵改進後提出[18]，主要是

改進了近似熵中會自我比對的特性，以減少計算熵值出現的誤差。其主要的目的也是一種衡量時間序列的複雜度的方式，當樣本熵值越小，表示時間序列的自我相似度越高；相反的，當樣本熵的值越大，表示時間序列的自我相似度越低，時間序列的複雜度就越高。樣本熵以 *SampEn*(*m, r, N*)表示，維度為 *m* 及 *m+1*，相似容限為 *r* (*r*>0)，而相似容限一樣代表的是在兩個時間序列資料之間的比較，判斷此兩序列資料是否為為相似的依據，而比較的數據長度為 *N*。

我們一樣以時間序列 $X_i = \{x_1,\ x_2, \dots,\ x_N\}$ 為例，其樣本熵的計算方式如下，將總長度為 N 的時間序列 X_i，組成維度為 m 的向量X_i^m，其中 $m \leq n$

$$X_i^m = \{x_i,\ x_{i+1}, \dots,\ x_{i-m+1}\},\ i =\ 1,\ 2, \dots,\ N-m+1$$

首先計算各個 *i* 值之間的最大差距值，為了比較兩個數據間的差距，我們以 X_i^m代表序列中以 m 為維度第 i 筆資料，$X_i^m(i = 1,\ 2, \dots,\ N-m+1)$，並且以 X_j^m代表序列中以 m 為維度的第 j 筆資料 $X_j^m(j = 1,\ 2, \dots,\ N-m+1)$，而我們用 $d\,[X, Y]$來代表 X 與 Y 兩者之間的距離

$$d\left[X_i^m,\ X_j^m\right] = \max\left|x_{i+k} - x_{j+k}\right|,\ k = 0,\ 1, \dots,\ m-1$$

並計算所有 $d\left[X_i^m,\ X_j^m\right] < r$ 的個數，但因為其不對自身進行比對，所以進行限制，i 不能夠與 j 相同

$$P_i^m(r) = \frac{1}{N+m-1} num\{d[X(i),\ X(j) < r]\}, i =\ 1,\ 2, \dots,\ N-m+1,\ i \neq j$$

接下來對所有的$P_i^m(r)$加總並進行取平均值的動作得到$P^m(r)$

$$P^m(r) = \frac{1}{N+m-1} \sum_{i=1}^{N-m+1} P_i^m(r)$$

計算維度 m+1 下$P^{m+1}(r)$的比率值，並重複上述步驟，得到$P^m(r)$及$P^{m+1}(r)$，便能計算樣本熵

$$SampEn(m,r) = \lim_{n\to\infty}\left\{-\ln\left[\frac{P^{m+1}(r)}{P^m(r)}\right]\right\}$$

在實際的時間序列估算中，N 取有限的數值後，能得到時間序列之樣本熵為

$$SampEn(m,r,N) = -\ln\left[\frac{P^{m+1}(r)}{P^m(r)}\right]$$

第三節、多尺度熵 (Multi-Scale Entropy, MSE)

多尺度熵由 Costa 於 2005 年提出[19]，改進了近似熵以及樣本熵僅在同一個尺度上計算複雜度，考慮了一個非線性時間序列在不同尺度上的複雜度，其被應用在生理訊號的分析上。Costa 提出，在兩個時間序列的比較中，若一時間序列的熵

值在大部分的尺度上皆大於另一個時間序列的熵值，代表前者的複雜度較高。單尺度熵用以測量一個時間序列的規則性，並且提供可預測性，而多尺度熵則是計算一個時間序列在不同尺度上的自我相似性。

我們以時間序列 $X_i = \{x_1,\ x_2, \dots,\ x_N\}$為例，多尺度熵的計算方式如下，先將原始的時間序列按照不同的尺度，進行粗粒化的動作，獲得處理後的時間序列y_j^{τ}，粗粒化的過程如圖 2-5 所示

$$y_j^{\tau} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} X_i,\ 1 \le \mathrm{j} \le \mathrm{N}/\tau$$

並且計算各個 y_j^{τ} 時間序列的 *SampEn*，多尺度熵便是在各個尺度下的樣本熵值，在多尺度熵 *MSE* 的參數中，最低維度為 m，相似容限為 r (r>0)，而y_j^{τ}是經過粗粒化處理後的時間序列

$$MSE = \{\tau \mid SampEn(m,\ r,\ y_j^{\tau})\}$$

使用多尺度熵的優勢之一是這個熵的計算方法，不像近似熵與樣本熵需要大量數據才能夠進行分析，多尺度熵只需要時間序列中少量的數據(約 500 筆)，就能夠精確的估算出其複雜度，也不會因為出現偏差值而產生極大的誤差，準確性非常之高。

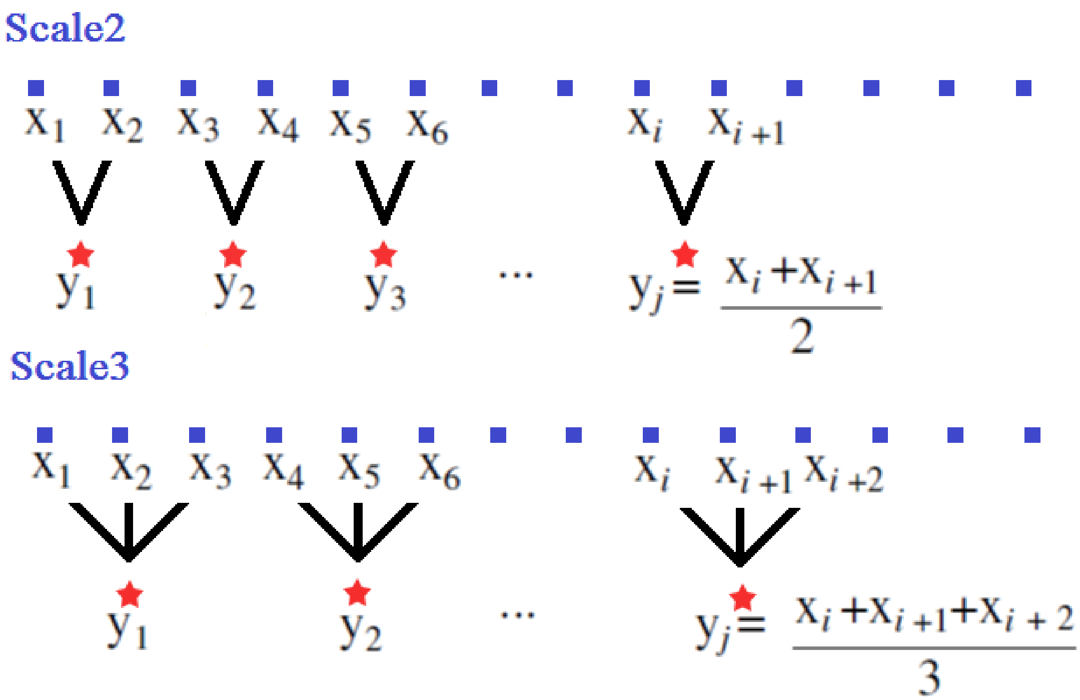


圖 2- 5 粗粒化的方式及計算[19]

粗粒化(Coarse-Graining)方法與多尺度技術就是將時間序列依尺度不同進行處

理，如圖 2-5 所示，在尺度為 2 時，將原始的時間序列兩兩為一組進行算術平均，如將 x_1、x_2 平均變成新的時間序列的第一筆數據 y_1，再將 x_3、x_4 進行平均得到 y_2，以此類推，將整個時間序列都做完粗粒化的動作後，便能得到尺度 2 的新時間序列 y_j^2；當尺度變為 3 時也是依照同樣的道理，將原始時間序列中的資料每三個為一組進行一次平均的計算，將整個時間序列處理過後，便能得到尺度為 3 時的新時間序列 y_j^3。

第四節、白雜訊(White Noise)與粉紅雜訊(Pink Noise, 或稱為 1/f Noise)

雜訊(Noise)在電子學中是指訊號在傳輸過程中受到外在能量的干擾，而這些能量即為雜訊。而在這些雜訊中，白雜訊與粉紅雜訊帶有著明顯的特徵，能夠使我們加以分析。白雜訊在時間序列上是一個完全隨機的訊號，而粉紅雜訊則是有長期相關性(Long-Range Correlation)的一種訊號。

圖 2-6 為使用 *MSE* 對白雜訊以及粉紅雜訊進行分析所得出之結果，我們可以看見，白雜訊的 *MSE* 值隨著尺度的上升而呈現單調下降的情形，而粉紅雜訊的 *MSE* 值則不太因為尺度的變化而有所變動。因為白雜訊時間序列隨機的特性，在尺度較小，尺度為 1 和 2 時，白雜訊的熵值是大於粉紅雜訊的，但隨著尺度的上升，白雜

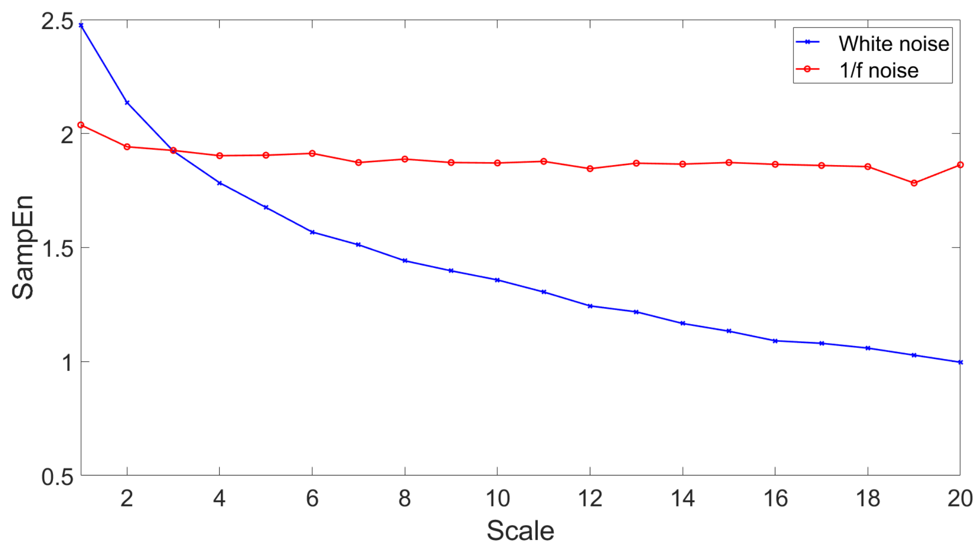


圖 2- 6 白雜訊及粉紅雜訊的 *MSE* 比較

訊的熵值呈現下降的趨勢，而熵值便都低於了粉紅雜訊。而粉紅雜訊則因其是一個

具有長期相關性的時間序列，所以不論尺度為多少熵值都穩定幾乎沒有什麼改變，代表著其時間序列的自我相似性很高。

第五節、修正後的多尺度熵(Modified Multiscale Entropy, ModMSE)

ModMSE 由 Wu 等人提出[20]，他們將原本 *MSE* 的粗粒化過程進行更精確的平均，原本的 *MSE* 因為粗粒化的過程，減少了時間序列的長度，所以當將傳統的 *MSE* 使用在短的時間序列上時，會因為時間序列的資料筆數較短或不足，造成其在熵值的計算上不夠精確。在這方面 *ModMSE* 便改善了這個問題，他們使用了移動平均(Moving Average)來進行粗粒化的動作，雖然這樣子的運算使得計算所需要的時間大幅增加，但會使其粗粒化後的數據保有原先的時間序列的連續性，以及去除其中的雜訊，因此 *ModMSE* 較為適合使用在較短的時間序列中。

以時間序列 $X_i = \{x_1, x_2, \ldots, x_N\}$為例，修正後的多尺度熵計算方式如下，將原始的時間序列按照不同的尺度，進行粗粒化的動作，獲得處理後的時間序列z_j^τ

$$z_j^\tau = \frac{1}{\tau}\sum_{i=j}^{j+\tau-1} X_i\,,\ 1 \le j \le N-\tau+1$$

接著計算各個 z_j^τ 時間序列的 *SampEn*，便為修正後多尺度熵的值

$$ModMSE = \{\tau \mid SampEn(m,\ r,\ z_j^\tau)\}$$

在圖 2-7 中，介紹了 *ModMSE* 中粗粒化的方式，當尺度為 2 時，將原始的時間序列資料兩兩相鄰做算術平均，如將 x_1、x_2 的值平均後變成新的時間序列的第一筆數據 y_1，同樣的再將 x_2、x_3 進行平均後得到 y_2，以此類推，將整個時間序列做完粗粒化的動作後，便能得到在尺度為 2 時的新時間序列 y_j^2，這樣子的方式比 *MSE* 好的地方在於能夠確保時間序列上的連續性。

ModMSE 因粗粒化的處理方式與 *MSE* 不同，在粗粒化後能夠兼顧到時間序列的連續性與去雜訊性，所以在計算上又會較 *MSE* 來的精準。假設一短時間序列只有 12 筆數據，當尺度為 2 時，傳統 *MSE* 粗粒化後會變為只有 6 筆數據再進行熵計算的動作，而尺度為 3 時，粗粒化過後會只有 4 筆數據進行熵計算。而 *ModMSE* 在經過粗粒化處理後，不論尺度為多少都會有 9 筆的數據可以進行熵計算。但缺點就是為了要計算大量的數據，計算時間會拉長許多，因此 *ModMSE* 較適合針對短的時間序列來進行分析。

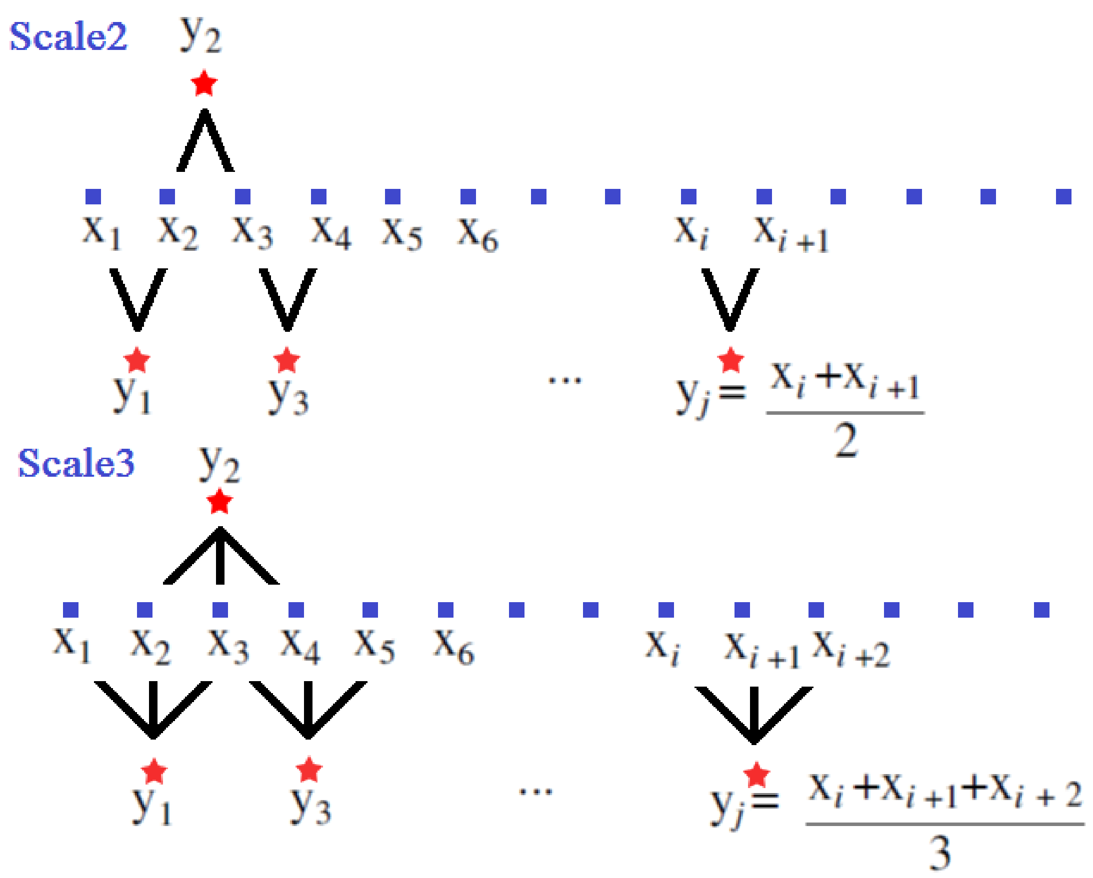


圖 2- 7 *ModMSE* 之粗粒化的方式及計算(本研究自行繪製)

第六節、潘凱圖(Poincaré Plot)

在生理醫學中，心跳是極重要的一個生理訊號參數，為了分析心跳的變異率並將其視覺化的表示，便有了潘凱圖的誕生。潘凱圖擁有簡單運算和適合短時間資料分析的優點。在[32]中，首次使用了潘凱圖進行分析，其是一種使用幾何的方式將病患的心跳間距或間期(R-R Interval)繪製在 2D 的圖形上，可用來有效地應用在病患的心律變動性(Heart Rate Variability)相關的分析上。

圖 2-8 為典型潘凱圖之範例，在潘凱圖中，*X* 軸代表的是心跳的間距 *RR*(n)，而 *Y* 軸代表下一次的心跳間距 *RR*(n+1)，圖形每一個座標點代表的意義為一次心跳間距與下一次的關係。一般在潘凱圖中會察看幾個重要的要素，*SD1* 表示垂直於 *X*=*Y* 這條線上，橢圓形短軸的長度，代表的是短期的心跳變異率，而 *SD2* 表示橢圓形在平行於 *X*=*Y* 這條線上，橢圓形長軸的長度，代表長期的心跳變異率，而 *SD1*/*SD2* 則表示 *SD1* 與 *SD2* 之間的比率，其代表的意義為副交感神經以及交感神經之間活動的比例。

關於潘凱圖中 *SD1* 以及 *SD2* 的計算方式如下，先將心跳間距的時間序列 *RR* 組成兩集合 *X* 與 *Y*

$RR = \{RR_1, RR_2, \ldots, RR_{n-1}, RR_n\}$

$X = \{ RR_1, RR_2, \ldots, RR_{n-2}, RR_{n-1}\}$

$Y = \{ RR_2, RR_3, \ldots, RR_{n-1}, RR_n\}$

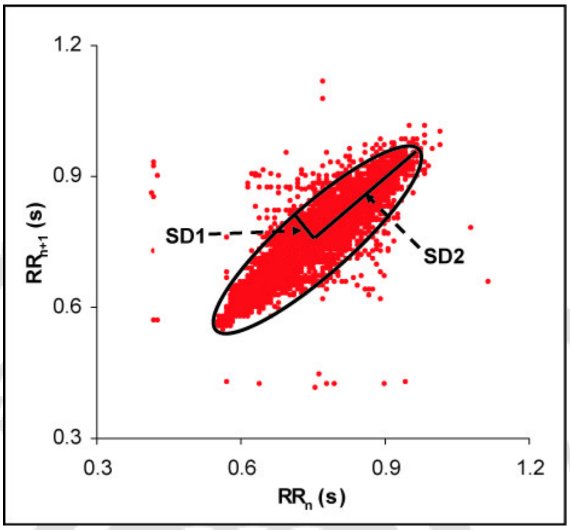


圖 2- 8 典型 Poincaré Plot

接著求出以 *X* 與 *Y* 為兩軸所畫出之圖形的質心點(X_c, Y_c)或稱資料分布中心

$$X_c = \frac{1}{n}\sum_{i=1}^{n} X_i\,,\ Y_c = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

接下來進一步算出各個資料的數據點橢圓形兩軸之間的距離 d_i^1 與 d_i^2

$$d_i^1 = \frac{1}{\sqrt{2}}|(X_i - X_c) - (Y_i - Y_c)|,\ d_i^2 = \frac{1}{\sqrt{2}}|(X_i - X_c) + (Y_i - Y_c)|$$

最後就能夠得出此圖形之 *SD1* 及 *SD2* 分別表示如下為

$$SD1^2 = \frac{1}{n}\sum_{i=1}^{n}(d_i^1)^2\,,\ SD2^2 = \frac{1}{n}\sum_{i=1}^{n}(d_i^2)^2$$

第七節、多尺度潘凱圖(Multiscale Poincaré Plot, MSP)

多尺度潘凱圖 *MSP* 為 2016 年在 Henriques 等人的研究[29]中提出，原本的潘凱圖已廣泛應用於心率變異分析中，而他們引入了於多個不同尺度上觀看潘凱圖的方式，來探測並分析一個生理信號時間序列，以觀察出患者是否患有慢性心臟衰竭(Chronic Heart Failure)與心房顫動(Atrial Fibrillation)的疾病。

在 *MSP* 的繪製中，首先要先將原始的時間序列做單一尺度粗粒化的處理

$$y_j^s = \frac{1}{s}\sum_{i=(j-1)s+1}^{j\tau} x_i\,, 1 \le j \le N/s$$

接著將各尺度的y_j^s序列，繪出其潘凱圖，便是能夠在多個尺度上比較的 MSP 了。圖 2-9 為以 *MSP* 圖形呈現之範例，若分析一時間序列 $X_i = \{x_1, x_2, \ldots, x_N\}$，在這麼多圖中每個圖的 X 軸代表的是分析的時間序列中一個數據點(例如 x_i)，Y 軸則代表的是此數據點的下一個數據(例如 x_{i+1})，每個點代表的就是時間序列中任一數據與下一個數據點的差異。接著由左而右，由上而下代表使用不同尺度對時間序列進行粗粒化後，再將此時間序列繪製成潘凱圖，便可以對不同尺度的圖形進行比較。而圖中顏色代表的是該區域的數據點密度，其數值介於 0 與 1 之間，顏色越偏藍色表示該區域的數據點密度越小，而越偏紅色的區域則表示其數據點的密度越大。

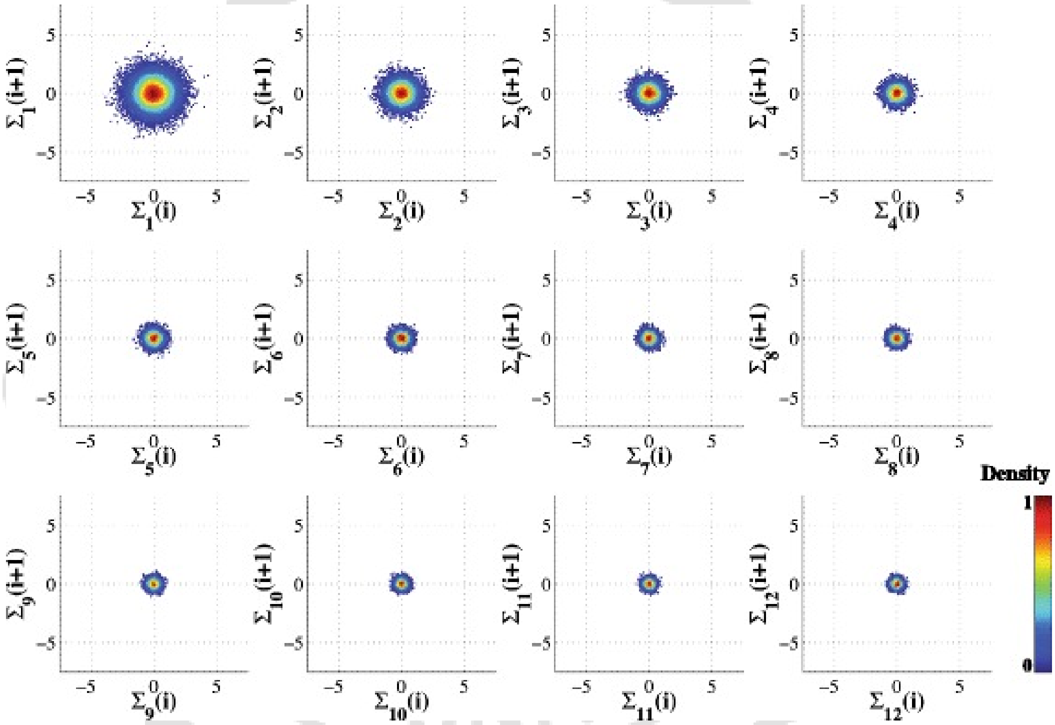


圖 2- 9 多尺度潘凱圖的範例

在[30]的研究中發現當 *MSP* 圖形呈現正圓型形狀時，表示這個時間序列是一個由不相關隨機的數據點所形成的常態分布，而當圖形呈現橢圓形形狀時，表示連續數據點之間為正相關，代表在時間序列中有著相關性的結構存在。

在本章節中，我們介紹了包含複雜度的定義，以及各種不同熵計算方式的定義與區別，最後介紹了潘凱圖與多尺度潘凱圖的計算。近似熵 *ApEn* 在 1991 年由 Pincus 提出，用來衡量時間序列中的複雜度；而 2000 年時 Richman 將近似熵改良，提出了樣本熵 *SampEn* 這個計算方式，其改進了近似熵中需要進行自我比對的特

性，將計算熵值出現的誤差減少；Costa 在 2005 年時提出了多尺度熵 *MSE*，改進了原本近似熵與樣本熵只在單一尺度上計算複雜度的方式，改為在多個尺度上進行熵值的計算，其使分析所需要的數據大大減少，只需要少量數據便能夠計算出精確的結果；而 Wu 等人於 2013 年提出修正後的多尺度熵 *ModMSE*，改進了多尺度熵在粗粒化上不連續的特性，以消除資料中雜訊，使其可以應用於短的時間序列進行運算。而接下來在第三、四章中，本論文將應用本章節中所介紹之方法，對比特幣市場交易價格之時間序列進行複雜度與其它的特性分析。

第三章、研究方法與資料處理

本章節介紹本論文的研究方法與資料處理的方式，首先介紹研究的流程包含研究目的的確認、相關研究文獻的探討、資料分析方法的選擇、數據的蒐集、比特幣的資料分析、研究分析結果與討論等步驟。

第一節、研究流程

本研究由一開始訂定題目，訂定我們的研究目的與目標，接著探討相關文獻，了解過去學者們進行相關研究以進行參考，接著找尋數據以進行分析，將數據使用各種方法進行分析，最後得出結果，而後進行結果的分析與討論，探討是否有達到我們所預期之結果。圖 3-1 為本研究之流程圖。

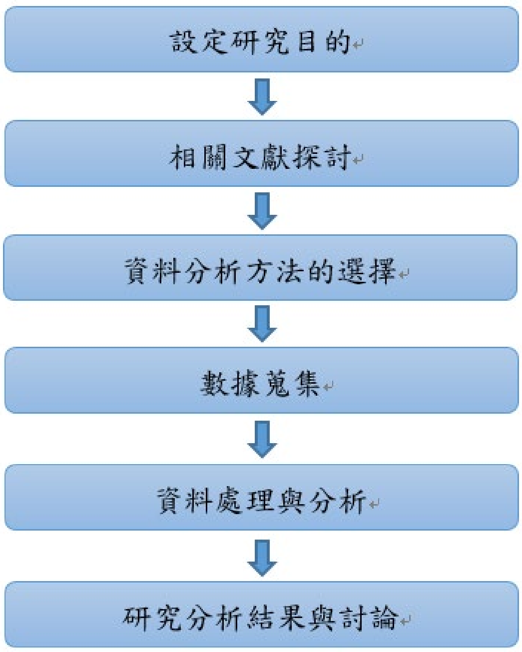


圖 3- 1　研究流程圖

其中有關研究目的的確認、相關研究文獻的探討、資料分析方法的選擇等三項已分別於本論文的第一章與第二章做了完整的說明。我們的研究目的是為了探索比特幣市場的複雜度，以分析比特幣市場中是否存在有規律性，並且探討經濟及彼特幣市場發生的重大事件是否會影響到市場的變化。在第一章的四到八頁我們進行了比特幣市場及各種資料分析方式的相關文獻探討，而我們選擇的資料分析方式則是使用多尺度熵 *MSE* 以及多尺度潘凱圖 *MSP* 來進行資料分析。

第二節、比特幣交易市場資料數據之蒐集與整理

本研究使用之數據，是使用 CoinmarketMap(網址：https://coinmarketcap.com/)上取得的比特幣價格數據，圖 3-2 便是 CoinmarketMap 的網站，此網站匯集了各大交易所的每日單價，我們可以在這個網站上下載到比特幣市場過去的所有數據，數據範圍由 2010 年 7 月 31 日開始，一直到 2018 年的 7 月 31 日為止，總計 8 年時間內收集 2922 筆交易日的收盤價格數據。

Charts Markets Social Tools Historical Data

Historical data for Bitcoin

Currency in USD

Jul 31, 2010 - Jul 31, 2018

Date	Open*	High	Low	Close**	Volume	Market Cap
Jul 31, 2018	8,181.20	8,181.53	7,696.93	7,780.44	5,287,530,000	133,688,476,220
Jul 30, 2018	8,221.58	8,235.50	7,917.50	8,180.48	5,551,400,000	140,547,190,784
Jul 29, 2018	8,205.82	8,272.26	8,141.18	8,218.46	4,107,190,016	141,185,844,808
Jul 28, 2018	8,169.06	8,222.85	8,110.77	8,192.15	3,988,750,080	140,716,560,550
Jul 27, 2018	7,950.40	8,262.66	7,839.76	8,165.01	5,195,879,936	140,235,573,607
Jul 26, 2018	8,176.85	8,290.33	7,878.71	7,951.58	4,899,089,920	136,553,079,708
Jul 25, 2018	8,379.66	8,416.87	8,086.36	8,181.39	5,845,400,064	140,482,950,401
Jul 24, 2018	7,716.51	8,424.27	7,705.50	8,424.27	7,277,689,856	144,634,918,474
Jul 23, 2018	7,414.71	7,771.50	7,409.10	7,711.11	5,132,480,000	132,375,368,459
Jul 22, 2018	7,417.80	7,537.95	7,383.82	7,418.49	3,695,460,096	127,336,615,249
Jul 21, 2018	7,352.72	7,437.64	7,262.41	7,419.29	3,726,609,920	127,333,468,174
Jul 20, 2018	7,467.40	7,594.67	7,323.26	7,354.13	4,936,869,888	126,202,570,251
Jul 19, 2018	7,378.20	7,494.46	7,295.46	7,466.86	5,111,629,824	128,122,730,711
Jul 18, 2018	7,315.32	7,534.99	7,280.47	7,370.78	6,103,410,176	126,458,165,406
Jul 17, 2018	6,739.65	7,387.24	6,684.17	7,321.04	5,961,950,208	125,588,963,706
Jul 16, 2018	6,357.01	6,741.75	6,357.01	6,741.75	4,725,799,936	115,638,203,962
Jul 15, 2018	6,272.70	6,403.46	6,256.51	6,359.64	3,285,459,968	109,074,579,938
Jul 14, 2018	6,247.50	6,298.19	6,212.22	6,276.12	2,923,670,016	107,631,453,835
Jul 13, 2018	6,235.03	6,310.55	6,192.24	6,238.05	3,805,400,064	106,968,440,793
Jul 12, 2018	6,396.78	6,397.10	6,136.42	6,228.81	3,770,170,112	106,798,864,820

圖 3- 2　CoinmarketMap 數據的部分網頁

我們將這些蒐集到的比特幣每日交易價格的時間序列繪製成圖 3-3，圖中顯示在 2017 年以前比特幣市場交易價格的變動較小，而 2017 年以後價格的變動則相對較大。如同其它文獻[4][6][13][14][24][25][26][27][28]中的研究方法一樣，因比特幣價格原始時間序列的波動性較大，本研究使用每日價格對數報酬(Daily Price Return)的計算方法，其公式為 $X_t = \log(P_t) - \log(P_{t-1})$，來對比特幣市場交易價格序列做初步的處理以計算其熵值，以便於進行後續的資料分析，並將其結果繪製為圖 3-4。

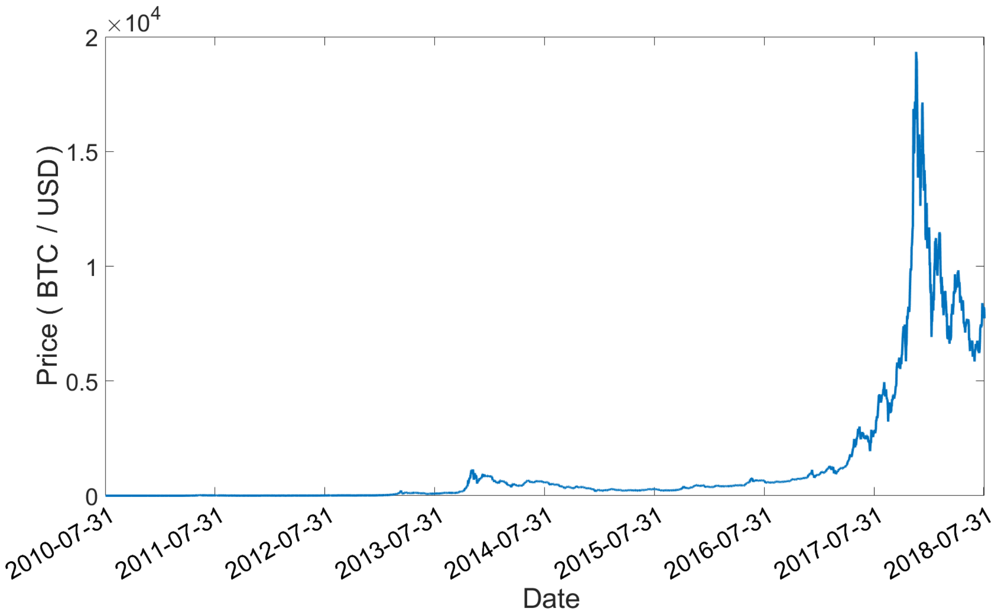


圖 3- 3　比特幣美金價格的時間序列(2010/07/31~2018/07/31)

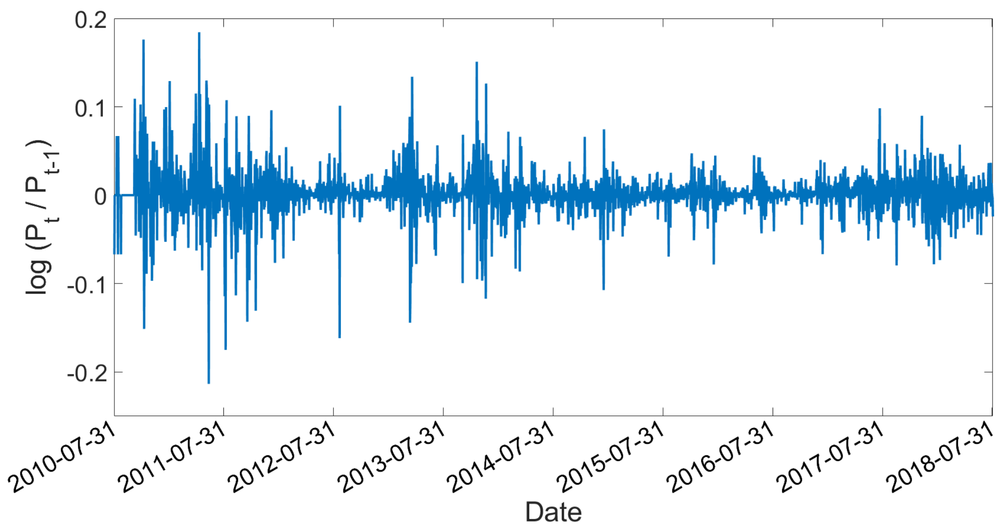


圖 3- 4　比特幣每日美金價格對數報酬的時間序列(2010/07/31~2018/07/31)

第三節、資料處理與分析

接著我們整理出了在比特幣歷史上的重大事件，如圖 3-5 中所示，以研究這些事件是否會對比特幣的市場造成一定程度上的影響。圖中所標示的事件序列資訊是我們從 Google 搜尋和媒體報導[33]中，自行彙集整理並標出了各個事件所發生的時間點，並在圖中以黑色的數字標號勾勒及標示出來，在表 3-1 中能夠找到各編號對應之事件，而紅色的英文標號則是表示著在時間序列上比特幣價格大幅變動

的事件與時間點。

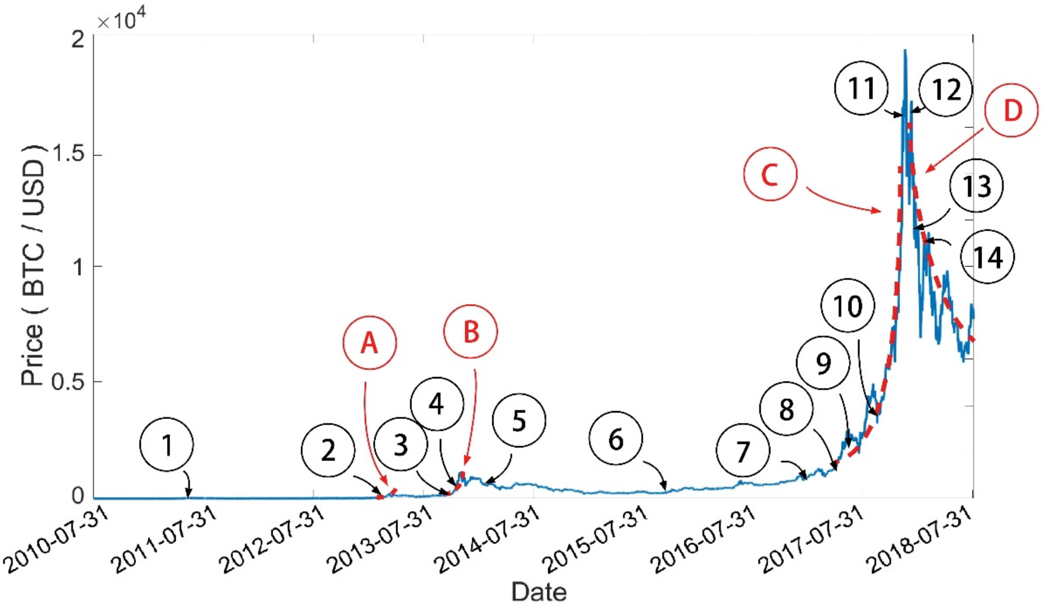


圖 3- 5　比特幣歷史上發生之重大事件(2010/07/31~2018/07/31)

並且因比特幣價格波動較大的關係，在圖 3-5 中我們再將整個時間序列切分為三個區塊並放大來檢視，以方便看出比特幣市場在較平穩之時期中價格的波動現象。圖 3-6 為 2010/07/31 起至 2013/02/26 止，這是一段比特幣才剛問世，那時是一段比特幣價格最平穩而幾乎沒有什麼大波動的時期；圖 3-7 為 2013/02/26 起至 2016/07/31 止，這段時期比特幣慢慢被世界所看見，發生了些許事件並使得其價格開始稍微有些波動，但整體而言波動的幅度也並不大；而圖 3-8 則為 2016/07/31 起至 2018/07/31 止，在這段時間中，比特幣因一些相關的金融事件而受到世界關注，大量的人群開始投資比特幣，而這便成為了比特幣價格波動最劇烈的一段時期。

2011 年的 6 月，比特幣與各國貨幣的兌換交易平台相繼上線，而時代雜誌及富士比等美國主流媒體都發表了比特幣的相關文章，吸引了群眾的目光，造成一波比特幣的價格上升，上升到了 32 美元的價格。接著投資者們漸漸意識到了比特幣這樣去中心化的貨幣，是一種劃時代的發展，使得比特幣的價格開始有著些許的波動。

2013 年 4 月歐洲發生塞浦路斯事件，大眾開始拋棄銀行監管之金融產業，轉而關注像是比特幣這樣的虛擬貨幣作為避險投資，大量的入資下使得比特幣價格

表 3- 1 比特幣歷史上發生之重大事件

編號	日期	事件內容
1	2011.06	比特幣受到時代雜誌及富比士關注，吸引群眾目光 (Bitcoin was attentioned by TIME & Forbes, attracting people)
2	2013.04	歐洲塞浦路斯事件 (Cyprus Event)
3	2013.10	世界第一台比特幣提款機 (The world's first bitcoin ATM)
4	2013.11	美國參議院聽證會上，首次稱比特幣為合法金融工具 (First time announce Bitcoin is Legal in US Senate)
5	2014.02	交易平台 Mt. Gox 宣稱被盜 85 萬枚比特幣 (Mt. Gox was stolen 850,000 bitcoins)
6	2015.10	歐盟承認比特幣並交易免稅 (EU exempting currency transactions from value added tax)
7	2017.04	比特幣在日本成為合法支付方式 (Japan officially recognized bitcoin as a method of payment)
8	2017.05	WannaCrypt 勒索病毒盛行 (WannaCrypt ransomware attack)
9	2017.06	中國央行禁止比特幣 (Bank of China banned Bitcoin trading)
10	2017.09	中國關閉所有比特幣交易所 (China closed all of the virtual currency exchange)
11	2017.12	芝加哥推出比特幣期貨合約 (CBOE launched Bitcoin futures contract)
12	2018.01	Facebook 宣布禁止虛擬貨幣廣告 (Facebook banned the advertisement of virtual currency)
13	2018.02	Coincheck 遭駭客入侵，盜走市值 580 億日圓的虛擬貨幣 (Coincheck was stolen Virtual currency value of 58 billions)
14	2018.03	Google. Twitter 禁止虛擬貨幣廣告 (Google, Twitter banned the advertisement of virtual currency)

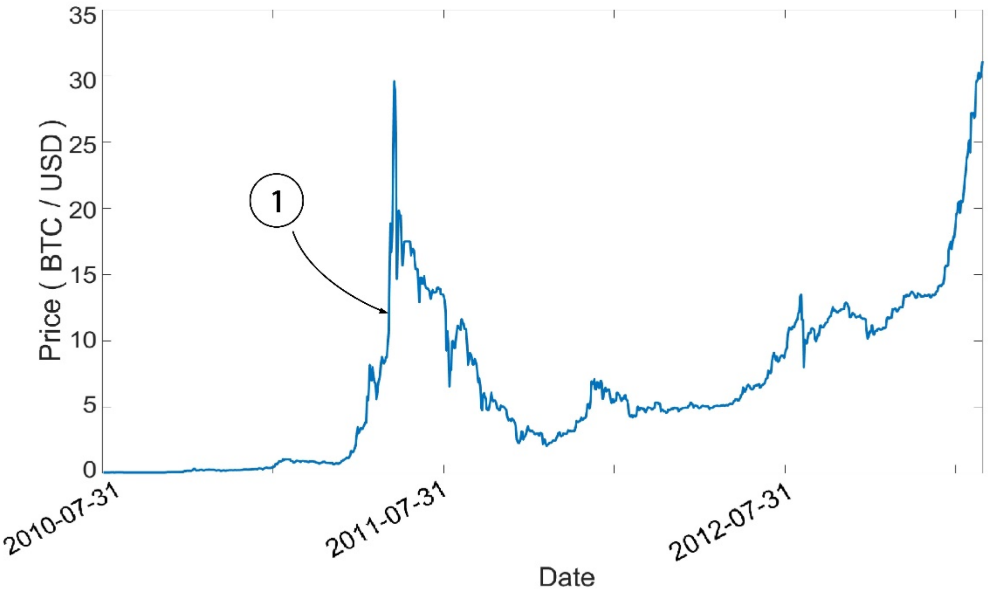


圖 3- 6　比特幣歷史上發生之重大事件(2010/07/31~2013/02/26)

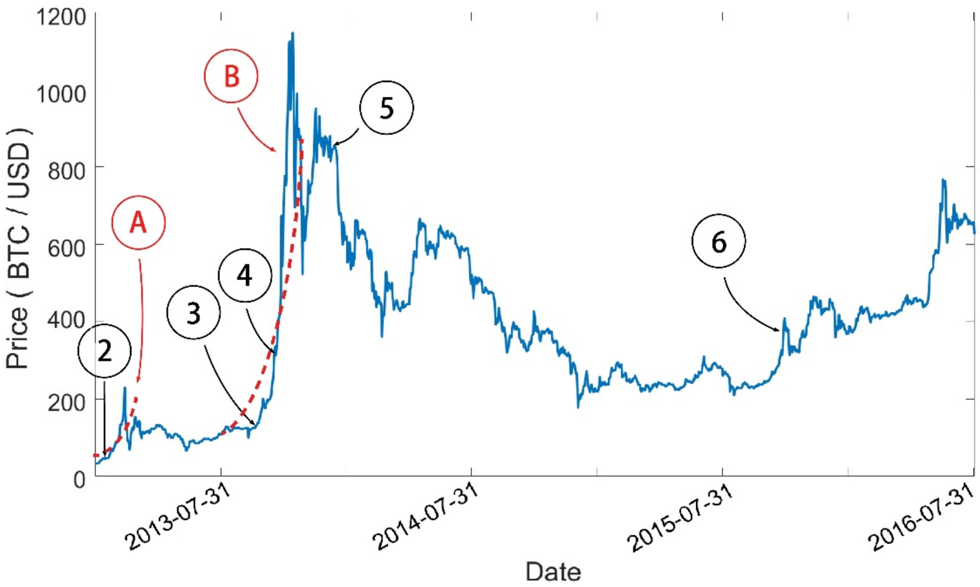


圖 3- 7 比特幣歷史上發生之重大事件(2013/02/26~2016/07/31)

開始大幅上漲，圖 3-7 中 A 曲線便代表著比特幣備受關注的情況下，比特幣價格因此由 30 美元暴漲至 200 多美元。2013 年 10 月 29 日，世界上第一台比特幣提款機在加拿大溫哥華問世，允許使用者將比特幣兌換為加幣提款，也可以透過存入現金購入比特幣。2013 年 11 月 19 日，美國司法部與美國證交會的代表在出席參議院的一個聽證會時，首次稱比特幣為一種合法的金融工具，並對比特幣的未來表示看好。這也造成了圖 3-7 中的 B 曲線，比特幣的快速上漲。

2014 年的 2 月，比特幣龍頭交易平台 Mt. Gox 宣布其平台內 85 萬枚比特幣遭到駭客透過系統漏洞竊取，而公司破產，此事件使比特幣安全性遭到質疑，也造成了價格下降的現象。2015 年 10 月 22 日，歐盟法院裁定比特幣交易不需要徵增值稅。

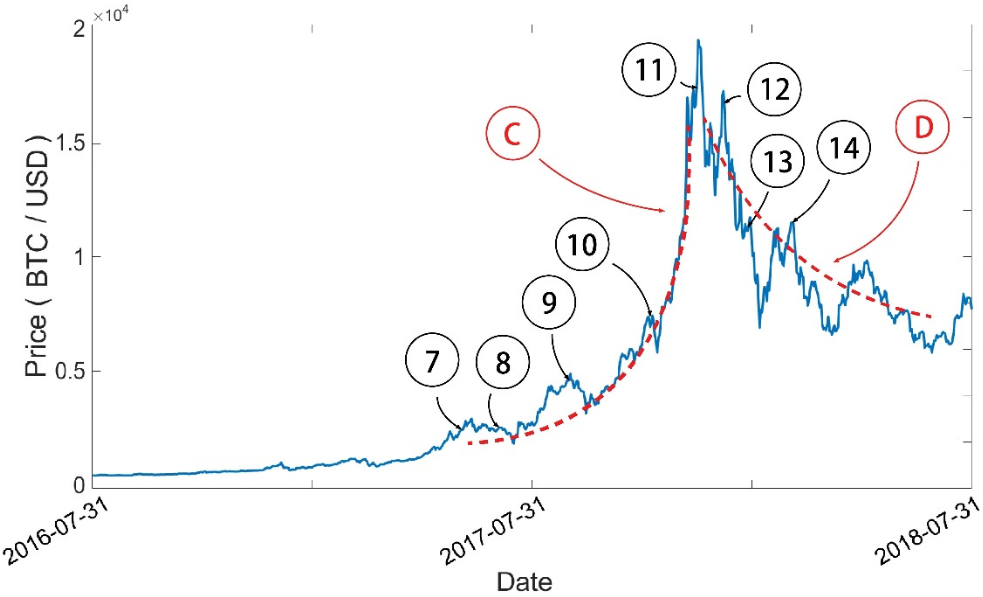


圖 3- 8 比特幣歷史上發生之重大事件(2016/07/31~2018/07/31)

2017 年的 4 月，比特幣在日本成為了合法的支付方式。同年 5 月 12 日 WannaCrypt 勒索病毒出現，全球許多地方受到了大規模攻擊，總計超過 23 萬台電腦受到病毒攻擊，此病毒將使用者電腦內部檔案加密，並要求使用者支付比特幣來贖回自己的檔案。此次事件使得幾乎全世界所有人都知道比特幣的存在，導致有許多人開始想要投資比特幣，又或者是開始挖礦獲得比特幣，圖 3-8 中的 C 曲線，就是表示著因此次事件造成的比特幣價格瘋狂上升的情形，最誇張的時期來到了將近 20000 美元的價格。6 月 15 日中國央行宣布國內禁止比特幣，並且在 9 月 8 日關閉境內所有的比特幣交易所，這兩個消息宣布後使得價格有小幅度的下降。而在 12 月 10 日美國監管機構美國商品期貨交易委員會核准美國芝加哥選擇權交易所開放比特幣期貨交易合約，成為全球第一個推出比特幣商品的大型交易所。在 2018 年的 1 月，Facebook 這個大型平台因比特幣手續費暴漲，而宣布了禁止所有虛擬貨幣的廣告。在 2 月時 Coincheck 平台遭到駭客攻擊，被駭客盜走市值 580 萬日圓的虛擬貨幣，雖然被盜走的並非比特幣，但也著實影響了虛擬貨幣的價格。3 月 Google 與大型社群平台 Twitter 也宣布了禁止所有虛擬貨幣廣告出現。這三件事件

造成了圖 3-8 中的 D 曲線，比特幣價格從最高點開始快速下降的情形。

第三節、資料分析方式

本研究使用多種方式分析比特幣市場的結構。

1. 使用 *MSE* 與 *ModMSE* 分別對比特幣市場做處理，並且比較兩種不同方法之間的差異性，以此來檢視 *ModMSE* 較 *MSE* 好的部分，並且決定後續分析將使用何種方式。
2. 使用 *ModMSE* 對比特幣市場與白雜訊、粉紅雜訊進行處理，並分析相關性以及差異性，以了解其中所存在的意義。
3. 使用 *ModMSE* 對比特幣歷史上重大事件後六個月的數據進行處理，以分析這些事件是否有影響到市場，因為我們的資料數據並不算多，而比特幣的時間序列波動較大，所以我們選用六個月進行分析，各個事件的時間點中便不會有覆蓋到相同的時間段。
4. 使用熱區圖對比特幣市場進行處理，以達到能夠讓視覺化分析的效果，。
5. 使用 *MSP* 對比特幣市場與白雜訊、粉紅雜訊進行分析，分析相關性以及差異性，並對使用熵方法找出之結果進行印證。

在本章中，我們介紹了本論文的研究流程，並介紹流程圖，接著介紹本研究所使用的研究數據以及如何將數據轉化為我們所需要的資料之處理方式。也呈現了本研究自行彙整的，在經濟及比特幣市場中發生的重大事件，將這些事件繪製成各張圖片以及表格，並且詳細的介紹了事件以及價格大幅變動的時間區段。最後介紹本論文將進行之分析，使用到第二章所介紹的各種資料分析方式，以及本章節介紹的數據，我們將會在接下來的第四章介紹本研究的分析結果與討論。

第四章、研究結果與討論

本章將根據第三章所提到的研究方法與流程進行資料的處理與分析，並在 Matlab 軟體工具上撰寫程式，以進行相關的熵計算分析與成果的視覺化繪圖，並且進行結果之討論。

第一節、研究結果

1. 使用 *MSE* 與 *ModMSE* 對比特幣市場進行分析

在第二章中曾提及修正後的多尺度熵 *ModMSE* 計算方法，是由 Wu 等人[20]提出對之前 Costa[19]所提出原始的多尺度熵 *MSE* 計算進行了改良。*ModMSE* 適合處理較短的時間序列，並可同時達到去雜訊與提升熵計算正確性的效果。在 Xia 等人的研究[28]中便使用了 *ModMSE* 來進行碳市場的交易價格分析與研究。相較於其它類似使用 *MSE* 熵計算的研究，本研究所使用的比特幣交易價格時間序列的數據長度僅有 2922 筆較為短少，並且在分析各事件時僅會使用到半年 182 筆的資料，為一相當短的時間序列，因此採用了 *ModMSE* 方法來進行多尺度熵的計算分析。

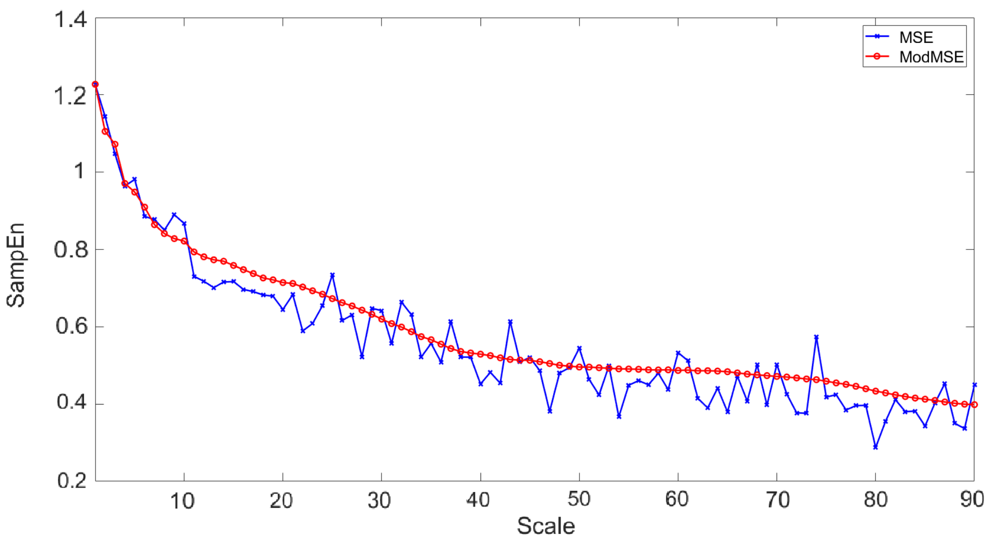


圖 4- 1　分別使用 *MSE* 與 *ModMSE* 多尺度熵方法來計算

比特幣歷年交易價格時間序列之比較

由圖 4-1 中可以觀察到，針對相同的一個時間序列數據，分別使用 *MSE* 和 *ModMSE* 兩個方法來計算不同尺度的熵值，從結果的曲線來看兩者的趨勢是差不

多的，都是由高往低的方向遞減。而相較於 *MSE* 的熵值曲線會有起伏不定的現象，*ModMSE* 的熵值曲線是較為平滑的。這表示使用 *ModMSE* 比 *MSE* 能夠更適合處理數據量較少的時間序列資料，也說明 *ModMSE* 的熵值計算方法能夠保留原時間序列的連續性與趨勢，並在去除雜訊後能夠更精確的做複雜度的分析研究。

2. 使用 *ModMSE* 對比特幣市場、白雜訊及粉紅雜訊進行分析

白雜訊是一個全部隨機的時間序列訊號，而粉紅雜訊則是一個具有自相關性的時間序列訊號。在 Wu 等人的研究[20]中，他們使用了 *ModMSE* 對白雜訊及粉紅雜訊進行分析比較；而在 Fan 等人的研究[27]及 Yin 等人的研究[28]中，也對白雜訊進行分析，並與他們研究的時間序列來進行比較。因此本論文以此方式來對比特幣市場交易價格動態進行分析，並與白雜訊及粉紅雜訊的多尺度熵值曲線走式進行相互之間的比較。

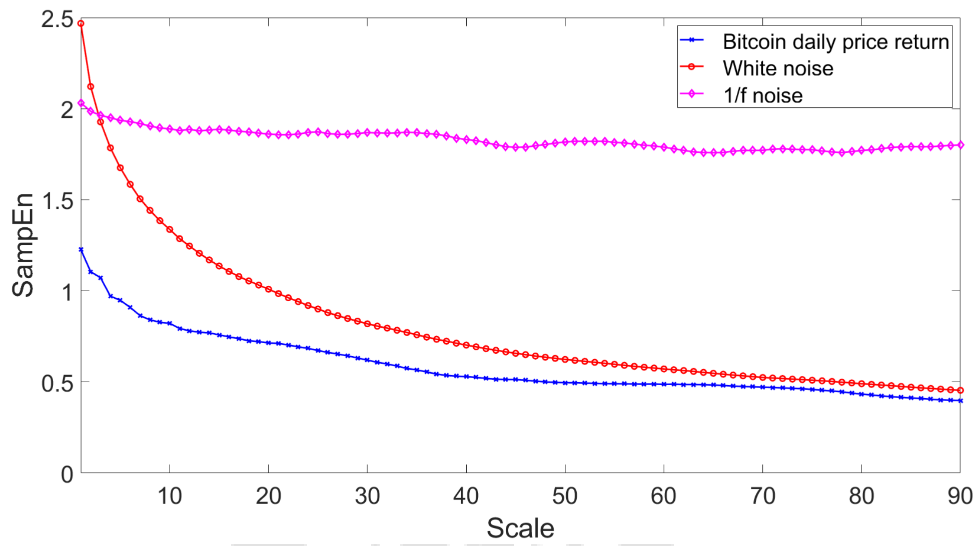


圖 4- 2　比特幣市場價格與白雜訊、粉紅雜訊時間序列多尺度熵值之比較

圖 4-2 展示了我們使用 *ModMSE* 對比特幣市場以及白雜訊、粉紅雜訊之分析結果。我們能夠看見，比特幣市場的多尺度熵值趨勢與白雜訊較為相似，在尺度越高的情況下，兩者的熵值皆為單調遞減下降的情形，而粉紅雜訊則是不管尺度多高，熵值皆為維持在一個差不多的數值上。在這項結果裡面我們可以推論得知，三者之中，比特幣交易市場時間序列的複雜度最低，表示其隨機度最高；粉紅雜訊則是複雜度最高與隨機度最低，是屬於一個有自相關性特徵的時間序列；而白雜訊介於二者之間，但是在低尺度時白雜訊的熵值與複雜度最高。這也與白雜訊及粉紅雜訊的

定義吻合，白雜訊是一完全隨機的時間序列，而粉紅雜訊就是一個自相關的時間序列，因此我們可以得知比特幣市場的價格動態是類似於白雜訊的情形，甚至比白雜訊更為隨機且自相關性更低。

3. 對比特幣重大事件後六個月的數據進行分析

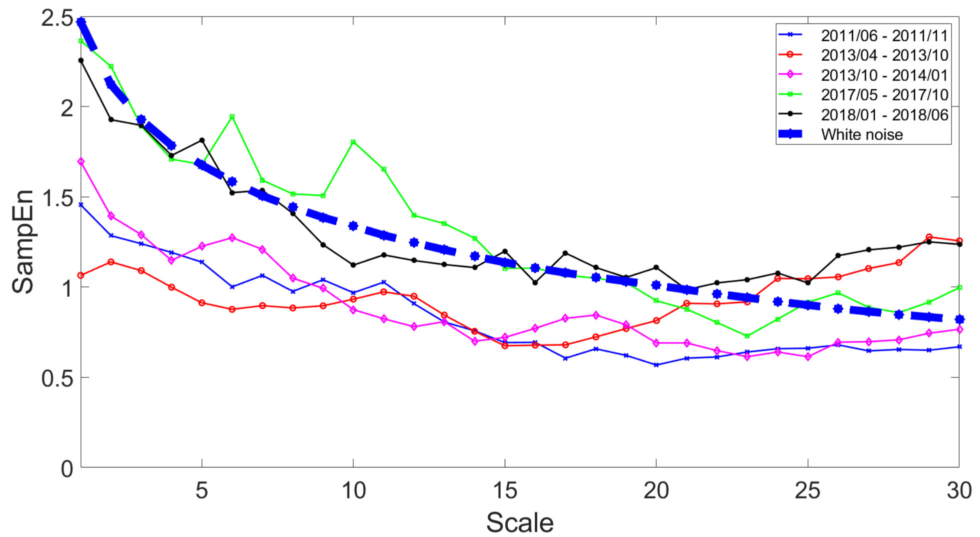


圖 4- 3　比特幣重大事件後數據與白雜訊之比較

圖 4-3 呈現的是將五個影響比特幣的重大事件，在其發生點之後六個月內的所有交易價格數據進行 *ModMSE* 的多尺度熵計算與分析。本研究在此取六個月的參考基準，作為觀看每一個金融事件發生後比特幣價格波動被影響的時間長度範圍，其原因仍為原始資料量不多的關係，若時間取的太長則事件之間的時間範圍將會相互重疊造成干擾。經 *ModMSE* 對這五個事件進行分析，再將這些分析數據對應白雜訊進行比較。此五個重大事件分別為 2011 年 6 月比特幣受富士比及時代雜誌報導，使群眾關注後的比特幣上漲、2013 年 4 月歐洲塞浦路斯事件後比特幣的上漲、2013 年 10 月美國稱比特幣為合法金融工具並使比特幣價格上漲、2017 年 5 月 WannaCrypt 勒索病毒讓全世界看見比特幣而價格暴漲以及 2018 年 1 月起各大平台宣布禁止虛擬貨幣廣告使得比特幣價格大幅下降等五個事件。

由圖 4-3 我們能夠看見，在經過重大的事件過後，熵值較原先的比特幣市場幾乎都是有所上升的，甚至在 2017 年 5 月及 2018 年 1 月這兩個事件後的數據更是使得熵值幾乎與白雜訊相同甚至是高於了白雜訊的情形。這代表了發生這些重大事件是有影響到比特幣市場的，它們使得市場中的交易變得較為活絡。以及在 2013

年 4 月的事件影響後，比特幣的熵值隨著尺度上升時熵值幾乎都趨於平穩而不是單調遞減的狀態，這代表著事件影響到了市場，使得市場開始有了一些結構的出現。

4. 使用熱區圖對比特幣市場進行分析

在 Fan 等人[27]與 Yin 等人對炭交易市場的研究中[28]，皆用到了熱區圖來對資料進行分析，熱區圖的好處便是能夠將資料視覺化，讓人一眼就能夠明瞭資料散布的區域與顏色所代表數值高低的意義。在圖中，X 軸表示的是年代，而 Y 軸表示的則是多尺度熵的尺度，右方顏色條代表的是熵值的高低，色彩越偏向藍色則表示著熵值越低，而色彩越偏向紅色則代表著熵值越高。

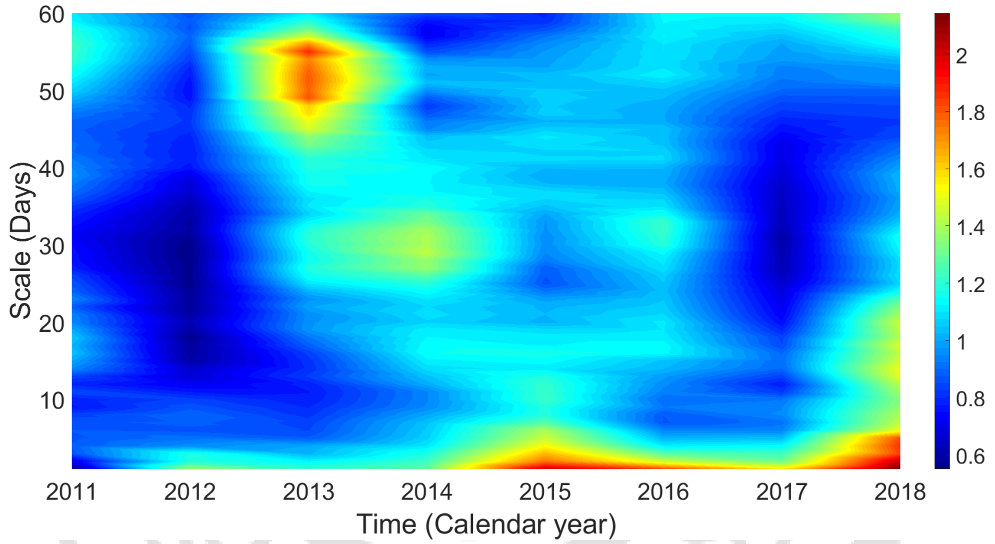


圖 4- 4　比特幣市場之熱區圖(2011 年~2018 年)

由圖 4-4 的熱區圖我們可以看見許多信息，在尺度小的時候熵值幾乎都較大，而尺度大時幾乎都有較小的熵值，這說明了比特幣市場交易價格的波動短期內較長期來的複雜。而在 2011~2013 年的期間，熵值都是非常非常低的，說明此時的比特幣市場極為平穩，非常少出現波動的現象。在 2013~2014 年中，熵值沒有隨尺度上升而單調下降，反而維持在一定的值中，在 2017~2018 年中，我們看見熵值相當高的現象，這說明了在這兩段時間點，市場是較為活絡的。而這張圖也印證了在圖 4-3 中的事件，在 2013 年及 2018 年兩個事件中都對彼特幣市場造成了影響。

5. 使用多尺度潘凱圖來對比特幣市場、白雜訊及粉紅雜訊進行分析

在 Henriques 等人的研究[29]中，使用了多尺度潘凱圖 *MSP* 對白雜訊與粉紅雜

訊進行分析，並解釋了其在 *MSP* 中所代表的意義。因此本研究也使用 *MSP* 對比特幣每日價格對數報酬(Daily Price Return)的時間序列進行分析，並將其與白雜訊、粉紅雜訊的時間序列經 *MSP* 計算之結果進行比較。

使用 *MSP* 的方面，我們採用了四種尺度大小，分別為 1、7、30 與 60 來代表一天、一周、一個月以及兩個月的時間長度來進行檢視與分析，以便可以觀察分析比特幣市場在短期(一天、一周)，或者在長期的時間(一個月、兩個月)範圍內的變異性。

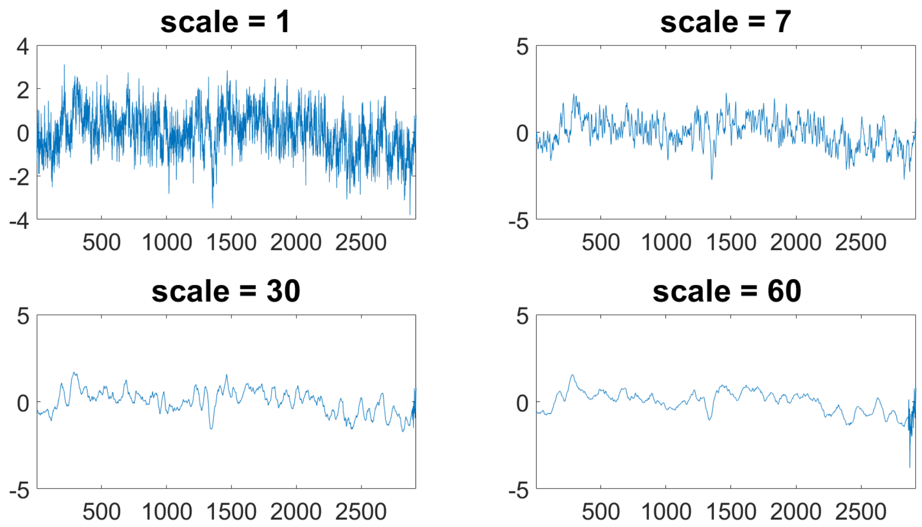


圖 4- 5　在不同尺度下粉紅雜訊的時間序列

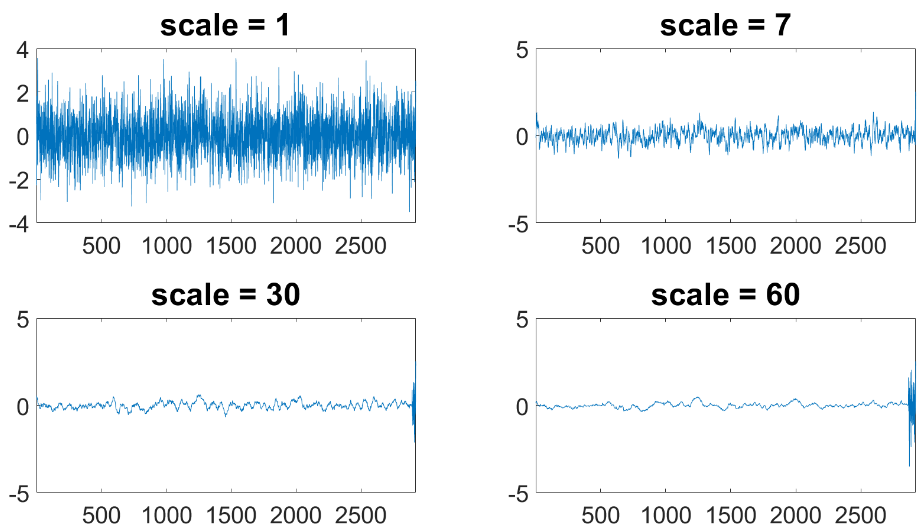


圖 4- 6　在不同尺度下白雜訊的時間序列

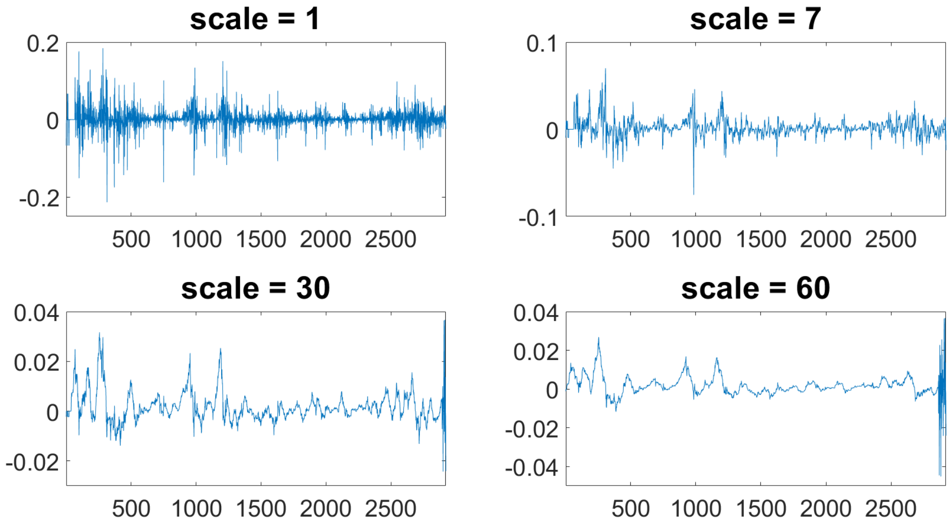


圖 4- 7　在不同尺度下比特幣每日價格對數報酬數據的時間序列

圖 4-5、4-6、4-7 中所呈現的分別是粉紅雜訊、白雜訊、以及比特幣市場的時間序列在上述四種尺度下的時間序列情形，而 *MSP* 就是用不同尺度的序列進行潘凱圖的運算來達到多個尺度的分析結果。

圖 4-8、4-9 與 4-10 分別呈現粉紅雜訊、白雜訊以及比特幣市場的每日價格對

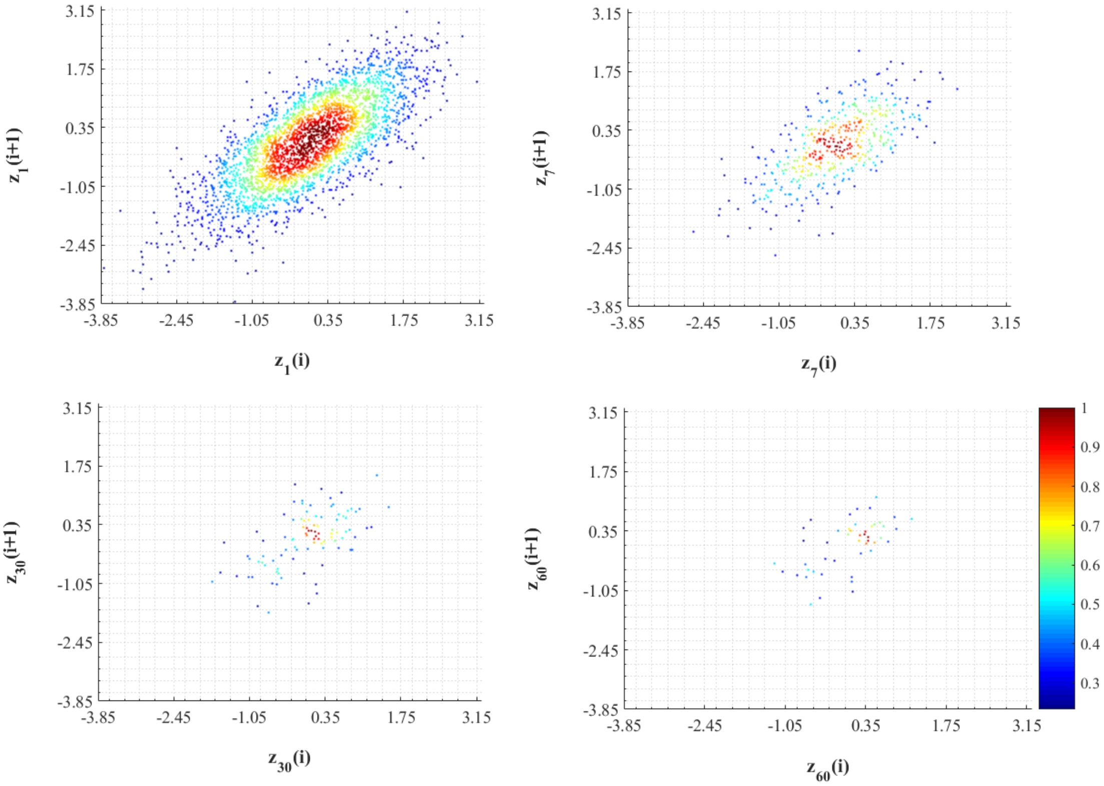


圖 4- 8　粉紅雜訊時間序列的 *MSP* 分析

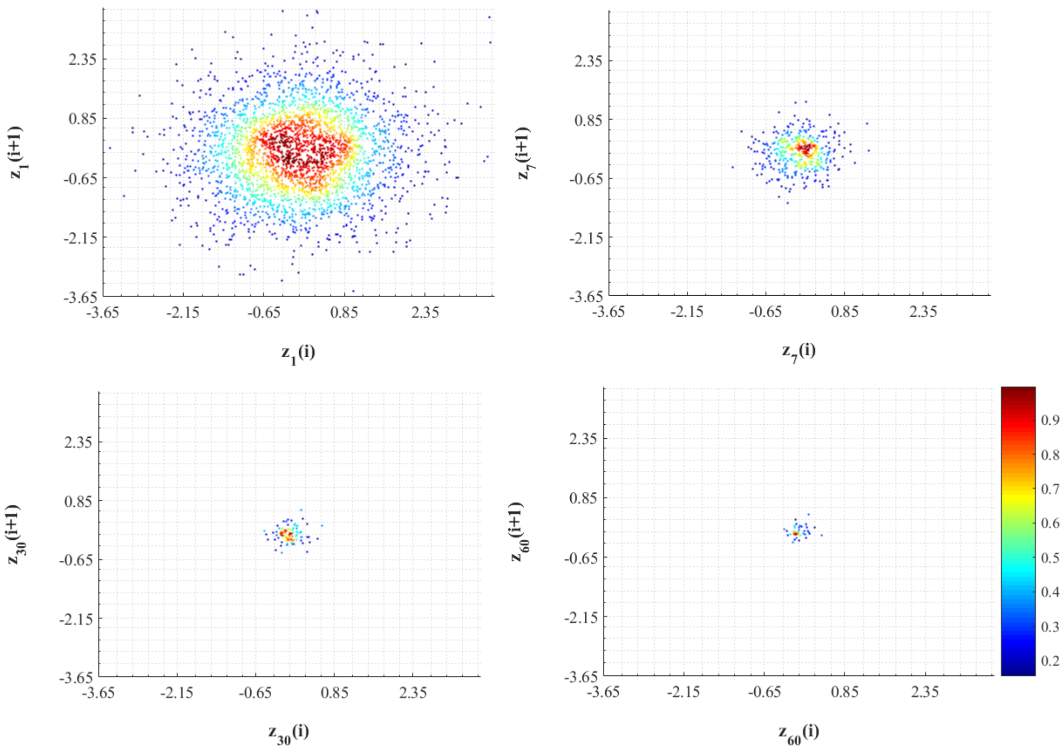


圖 4- 9　白雜訊時間序列的 *MSP* 分析

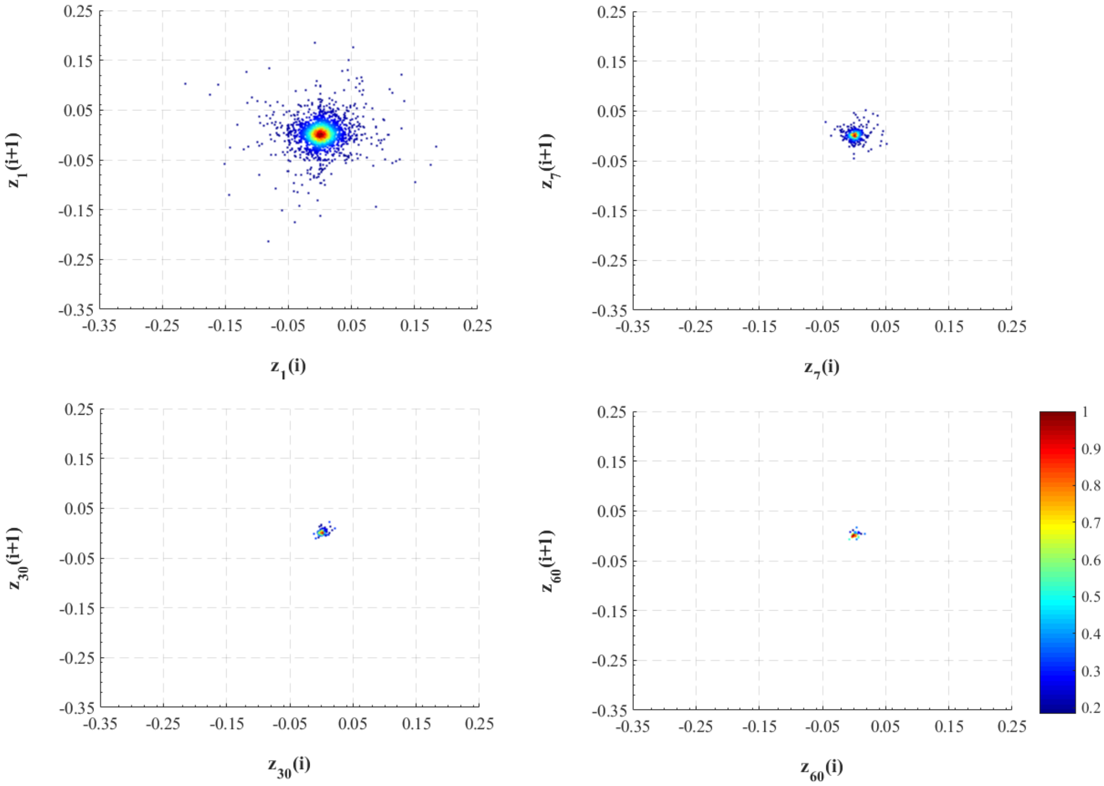


圖 4- 10　比特幣每日價格對數報酬時間序列的 *MSP* 分析

數報酬等三個時間序列的 *MSP* 分析結果，在白雜訊及粉紅雜訊的分析中我們皆使用了與比特幣市場每日價格對數報酬序列相同的資料數量(2922 筆)。在 *MSP* 的圖中，*X* 軸代表時間序列中一個數據的數值，*Y* 軸則代表 *X* 軸上相對應鄰近的下一個數據，因此 *MSP* 圖中的每一個座標點代表的是時間序列上每一筆資料(*X* 座標)與下一筆資料(*Y* 座標)之間的差距。而在圖中顏色代表的是資料聚集的密度，其數值範圍介於 0(藍色)與 1(紅色)之間，顏色越接近紅色代表的是這一區域的資料密度相當高，而顏色越接近藍色則表示這一區域的資料密度相對較低。

表 4- 1 白雜訊、粉紅雜訊與比特幣市場的 *SD1*、*SD2* 與 *SD1*/*SD2*

	1/f Noise		
	SD1	**SD2**	**SD1/SD2**
1	0.5088	1.3198	0.39
7	0.4734	1.0510	0.45
30	0.4376	0.8296	0.53
60	0.3801	0.7200	0.53
	White Noise		
	SD1	**SD2**	**SD1/SD2**
1	0.9817	1.0177	0.96
7	0.3871	0.4008	0.97
30	0.1831	0.1936	0.95
60	0.1215	0.1453	0.84
	Bitcoin		
	SD1	**SD2**	**SD1/SD2**
1	0.0240	0.0254	0.94
7	0.0101	0.0105	0.96
30	0.0043	0.0061	0.70
60	0.0040	0.0047	0.85

在圖 4-8 中我們可以看見，粉紅雜訊的 *MSP* 圖形呈現橢圓形的分布，代表著時間序列數據有著自我相似性，而圖 4-9 與 4-10 中，白雜訊與比特幣市場的圖形呈現圓形的分布，代表兩者在時間序列為隨機性高的時間序列，並沒有自我相似

性的存在。而在的 *MSP* 圖形中比特幣市場的圖形較白雜訊密集的原因，便是因為原始的時間序列資料中，比特幣市場的波動幅度較小(見圖 4-7)，而白雜訊的波動幅度較大(見圖 4-6)，這才導致了在圖形上看起來雖然同樣都是圓形，但比特幣的圖形較白雜訊的圖形聚集。

而由表 4-1 結果中我們也可以看見，白雜訊的圖形不管在什麼樣的尺度下，*SD1*/*SD2* (表示橢圓形的橢圓度) 都相近於 1，這表示圖形呈現圓形的分布，這說明了此時間序列是一個相關隨機的數據點所形成，與白雜訊完全隨機的本質相同。而粉紅雜訊在不由不同尺度上皆為橢圓形的分布情形，其表示此時間序列數據點之間的正相關關係，代表著其中有一定的結構存在，與粉紅雜訊這樣自相關的序列本質相互符合。

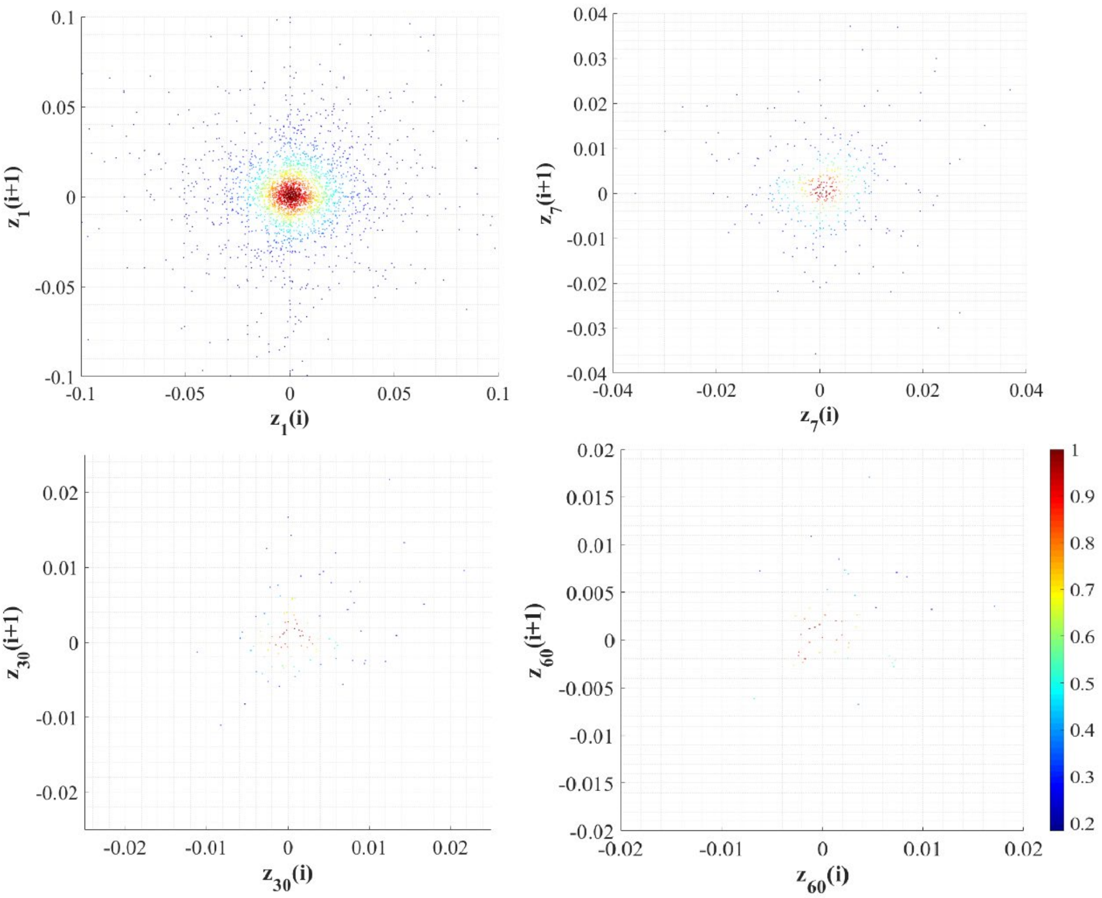


圖 4- 11　放大檢視比特幣市場的 *MSP* 分析

因為比特幣市場在 *MSP* 上呈現的資料較為密集，所以我們將圖片放大檢視，於圖 4-11 中，我可以發現到比特幣市場的圖形也是與白雜訊的圖形非常相近，並

且經計算後其 *SD1/SD2* 也相當接近於 1，表示其呈現一個正圓形分布的現象，代表比特幣市場的隨機分布，此結果也印證了我們以 *ModMSE* 分析出的結果，比特幣市場是較為隨機且自相關性較低的一個時間序列。

本章節介紹了本論文分析的所有結果，在這些結果中我們能夠發現到，Wu 等人在研究中對 *MSE* 進行改良出的 *ModMSE*，確實較前者更加精確。而用 *ModMSE* 進行的分析，對隨機時間序列白雜訊進行分析，尺度越大時，熵值呈現單調下降的情形，而對自相似的時間序列粉紅雜訊進行分析，在不同的尺度下，熵值皆在一個不怎麼變動的值。而我們最終發現比特幣市場的時間序列與白雜訊相似，尺度越大熵值下降，且有著較低的熵值，代表著比特幣的時間序列是較為隨機，比較無法預測其發展。而當世界發生了什麼重大事件時，則會很容易的就影響到比特幣市場的波動。而在 *MSP* 的分析中我們可以看見，比特幣市場也確實是與白雜訊相似，是個複雜度較低、較為隨機的一個時間序列，這與我們使用 *ModMSE* 的研究結果是相互吻合的。綜上所述，本研究歸納出了比特幣市場是一個波動性較為隨機，很難與策其未來發展，並且容易受到外在事件所影響的時間序列。

第五章、結論

本論文共有五個章節，在第一章說明了研究動機與目的，以及牽涉到與本論文相關研究的文獻探討與應用領域。在第二章敘述了系統複雜度的解釋與介紹了各式熵的計算方式與特性，以及潘凱圖的計算方式。而第三章介紹了研究數據的取得、資料處理以及研究方法的部分。第四章進行分析並且展示本論文的研究結果。而本章將介紹本論文的總結以及研究的貢獻。

第一節、研究目的回顧

在第一章中提到了本論文的研究目標以及期許，希望能夠藉由 *ModMSE* 多尺度熵的分析方式，來了解比特幣市場交易價格的時間序列中是否存在著富有意義的變動模式及結構，並分析其中是否存在著規律性。並且探討自比特幣發行交易至今，分別在經濟市場以及比特幣市場中所發生的各大事件是否會影響到比特幣市場的交易價格。最後再使用多尺度潘凱圖 *MSP* 的圖形分析來印證在 *ModMSE* 中分析的結果。

而在我們的研究中，我們發現第一點：找到比特幣市場交易價格的時間序列中中是否存在著富有意義的變動模式及結構的方面，答案是否定的。在結果中我們認為比特幣市場的交易價格是一個隨機的時間序列，其中並沒有一個結構存在，即沒有規律性可循，所以可由此推論比特幣市場未來的可預測性較低。而在各大重大事件對於比特幣市場的影響方面，我們確認了在經濟市場及比特幣市場中發生的各個事件，都會在一定程度上影響到了比特幣市場的價格波動，這項分析便符合了我們原先所設定的研究期許。而在 *MSP* 的方面，則是發現了在比特幣市場的價格波動中，不存在著自我相似性，雖然研究分析中沒有找到比特幣市場中的結構存在，但也印證了同時使用 *ModMSE* 與 *MSP* 兩種方法來相互進行交叉的分析驗證，可分別得到了相似的結果，提升了研究結果的準確性。

第二節、研究方法回顧

在本論文的第二章及第三章中分別介紹了在本研究中使用的研究方法與資料分析方式。本研究採用了適合處理資料量較少的修正後多尺度熵 *ModMSE* 來計算比特幣市場價格的複雜度，並且使用隨機性高與自我相似性低的白雜訊之時間序

列，以及隨機性低且自我相似性高的粉紅雜訊之時間序列，以進行三者相互之間的比較。

我們在第二章中明確介紹複雜度的定義還有白雜訊與粉紅雜訊的特性，以及各式熵的計算方式，包含單一尺度的近似熵 *ApEn* 和樣本熵 *SampEn*，以及多尺度熵 *MSE* 與修正後的多尺度熵 *ModMSE* 等，亦對本論文另一種分析方式多尺度潘凱圖做了詳細的介紹。

在第三章中我們介紹了本論文的研究流程，並且介紹本研究分析之比特幣市場交易價格數據從何取得，以及將取得後的資料以每日價格對數報酬的方式處理並呈現，而後再繼續進行我們下一步的分析。在此也分別整理出在經濟市場以及比特幣市場中發生的重大事件，並繪製為圖表來檢視，最後介紹與說明本研究分析資料的方式與步驟。

第三節、研究結論

本論文使用了修正後的多尺度熵 *ModMSE*、多尺度的潘凱圖 *MSP*、以及熱區圖等，進行了對比特幣市場每日價格對數報酬之時間序列的研究。並且也對白雜訊及粉紅雜訊的時間序列進行模擬分析，來計算與比較比特幣交易市場價格、白雜訊以及粉紅雜訊時間序列在不同尺度上的熵值與複雜度，以檢視三者在這些特徵上的差異處。

在第四章所呈現的研究結果中顯示，透過由 *ModMSE* 計算分析可發現比特幣市場交易價格時間序列的熵值趨勢走向與白雜訊的特徵相近，都是隨著尺度上升而單調遞減，表示兩者皆是一個複雜度低的時間序列，而比特幣市場的複雜度更低於白雜訊；而在多尺度 *MSP* 的分析結果中，可以發現到比特幣交易市場價格序列依舊與白雜訊有著相似的結果。兩者在 *MSP* 圖形中皆呈現著正圓形的相關分布圖形，表示其自我相似度不高。由 *ModMSE* 與 *MSP* 這兩個分析方法中可以發現，其結論是相互吻合的，比特幣市場交易價格的時間序列，其變動特徵與白雜訊類似，都呈現較為隨機並且沒有自我相似性的狀態。而在對事件是否影響到比特幣市場的方面，我們也發現各個事件皆對比特幣市場的熵值與自我相似性方面有所影響，因此我們可以得到比特幣市場是一個較隨機、沒有自我相似性、其中也沒有明顯結構存在，並且容易受到外來事件影響的一個市場。

第四節、研究貢獻

總結本論文的研究貢獻，可歸納成以下四點：

1. 經過我們盡力的搜尋已公開的學術研究與文獻，就我們所認知，本論文為第一篇使用修正後的多尺度熵 *ModMSE* 對比特幣市場的交易價格動態做分析，分析其時間序列中複雜度的研究。
2. 而在我們搜尋發現，過去所有使用多尺度潘凱圖 *MSP* 的相關研究，幾乎都是用於生理醫學訊號的領域，以解決心臟疾病診斷的問題。而本論文是第一篇將多尺度 *MSP* 應用在分析生理醫學訊號以外領域的研究，我們利用 *MSP* 簡易的計算與資訊視覺化的優勢，來對比特幣市場的時間序列數據進行處理，分析在不同尺度下時間序列中鄰近資料之間的自我相似性。
3. 我們可以發現比特幣市場，透過 *ModMSE* 以及 *MSP* 兩種資料科學的實證分析方法中，都能夠獲得相同的結果。而在過去不論是多尺度熵以及多尺度潘凱圖中在生理上的研究中，比特幣市場的時間序列與身體不健康或者是有特殊疾病的人有著相同的特徵。而從現在來看，回顧比特幣交易市場的歷史發展與價格走勢，確實符合本論文所做出之結論，換言之，比特幣是一種風險較大，並不是一個適合、健康與穩定的投資。
4. 在現在的許多研究中，有人使用機器學習來對比特幣進行分析，並且想要預測比特幣的未來價格，機器學習是一種讓機器藉由過去的資料進行學習，並且用來推敲未來可能發生的未知情形的方式，而其適合使用在一個很難用人為的方式來寫出一個規則描述的東西，而比特幣的市場正是這樣子的一個例子。所以本研究能夠讓想使用機器學習分析比特幣市場，並且預測比特幣未來變化的未來研究有所幫助。

第五節、未來研究發展

本論文的研究成果與結論可提供給相關問題研究的一個解決模式與參考經驗。因此在本研究未來的發展方向，我們希望也能透過多尺度熵的計算以及潘凱圖的呈現，對其它亦屬於財務金融方面的時間序列(如股票市場、期貨等)複雜度分析相關的學術研究議題與應用，期望可以將這些分析的方式用在更多不同的領域上。

參考文獻

1. Nakamoto S. (2008), Bitcoin: a peer-to-peer electronic cash system, Available: https://bitcoin.org/bitcoin.pdf
2. BENJAMIN WALLACE (2011), The Rise and Fall of Bitcoin, Available: https://www.wired.com/2011/11/mf-bitcoin/
3. Bohme R., Christin N., Edelman B., Moore M. (2015), Bitcoin: economic, technology and governance, *Journal of Economic Perspectives*, *29*(2), 213-238
4. S. Lahmiri, S. Bekiros (2018), Chaos, randomness and multi-fractality in bitcoin market, *Chaos Solitons Fractals*, *106*, 28-34
5. T. Kim (2017), On the transaction cost of bitcoin, *Finance Research Letters*, *23*, 300-305
6. Baur D.G., Dimpfl T., Kuck K. (2018), Bitcoin, gold and the US dollar – A replication and extension, *Finance Research Letters*, *25*, 103-110
7. Y. Jiang, H. Nie, W. Ruan (2018), Time-varying long-term memory in Bitcoin market, *Finance Research Letters*, *25*, 280-284
8. E. Demir, G. Gozgor, C.K.M. Lau, S.A. Vigne (2018), Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation, *Finance Research Letters*, *26*, 145-149
9. A. Urquhart (2018), What causes the attention of Bitcoin?, *Economics Letters*, *166*, 40-44
10. Bouri E., Shahzad S.J.H., Roubaud D. (2018), Co-explosivity in the cryptocurrency market, *Finance Research Letters*, doi: 10.1016/j.frl.2018.07.005
11. Giudici P., Abu-Hashish I. (2018), What determines bitcoin exchange prices? A network VAR approach, *Finance Research Letters*, doi: 10.1016/j.frl.2018.05.013
12. T. Takaishi (2018), Statistical properties and multifractality of bitcoin, *Physica A*, *506* , 507-519
13. Tawari A.K., Jana R.K., Das D., Roubaud D. (2018), Informational efficiency of Bitcoin—An extension, *Economics Letters*, *163*, 106-109
14. S. Begušić, Z. Kostanjčar, H.E. Stanley, B. Podobnik (2018), Scaling properties of extreme price fluctuations in bitcoin markets, *Physica A*, *510*, 400-406
15. Ladislav Kristoufek (2018), On Bitcoin markets (in)efficiency and its evolution *Physica A*, *503*, 257-262

16. Shannon C.E. (1948), A mathematical theory of communication., *The Bell System Technical Journal*, *27*, 379-423.
17. Pincus S.M. (1991), Approximate entropy as a measure of system complexity., *Proc. Natl. Acad. Sci. USA*, *88*, 2297-2301.
18. J. Richman, J.S. Moorman (2000), Physiological time-series analysis using approximate entropy and sample entropy, *Am J Physiol Heart Circ Physiol*, *278*(6), 2039-2049
19. Costa. M., Goldberger. A.L., Peng. C. (2002), Multiscale entropy analysis of complex physiologic time series., Physical Review Letters, *89*(6), 068102
20. Wu S.D., Wu C.W., Lee K.Y., Lin S.G., Modified multiscale entropy for short-term time series analysis, Physica A, 392 (2013), pp. 5865-5873
21. L. Gulko (1999), The entropic market hypothesis, *International Journal of Theoretical and Applied Finance*, *2*, 293-329
22. G.A. Darbellay, D. Wuertz (2000), The entropy as a tool for analysing statistical dependences in financial time series, *Physica A*, *287*, 429-439
23. Chao Li, Pengjian Shang (2018), Complexity analysis based on generalized deviation for financial markets, *Physica A*, 494, 118-128
24. Jianan Xia, Pengjian Shang, Jing Wang, Wenbin Shi (2014), Classifying of financial time series based on multiscale entropy and multiscale time irreversibility, *Physica A*, *400*, 151-158
25. Martina E., Rodriguez E., Escarela-Perez R., Alvarez-Ramirez J. (2011), Multiscale entropy analysis of crude oil price dynamics , *Energy Economics*, *33*, 936-947
26. Ortiz-Cruz A., Rodriguez E., Ibarra-Valdez C., Alvarez-Ramirez J. (2012), Efficiency of crude oil markets: Evidences from informational entropy analysis, *Energy Policy*, *41*, 365-373
27. Fan. X., Li. S., Tian. L. (2016), Complexity of carbon market from multi-scale entropy analysis., *Physica A*, *452*, 79-85.
28. Jiuli Yin, Cui Su, Yongfen Zhang, Xinghua Fan (2018), Complexity Analysis of Carbon Market Using the Modified Multi-Scale Entropy, *Entropy*, *20*(6), doi: 10.3390/e20060434
29. Teresa S. Henriques, Sara Mariani, Anton Burykin, Filipa Rodrigues, Tiago F. Silva, Ary L. Goldberger (2016), Multiscale poincaré plots for visualizing the structure of heartbeat time series, *BMC Medical Informatics and Decision Making*, doi: 10.1186/s12911-016-0252-0

30. Anne Humeau-Heurtier, Guillaume Mahé, Gilles Hunault, Lydie Gascoin, Pierre Abraham (2017), Multiscale Poincaré plot analysis of time series from laser speckle contrast imaging data, *Biomedical Signal Processing and Control*, *38*, 361-369
31. Phyllis K Stein, Anand Reddy (2005), Non-Linear Heart Rate Variability and Risk Stratification in Cardiovascular Disease, *BMC Medical Informatics and Decision Making*, *5*(3), 210-220
32. P.W. Kamen, A.M. Tonkin (1995), Application of the Poincaré plot to heart rate variability: a new measure of functional status in heart rate failure, *Australian and New Zealand journal of medicine*, *25*, 18-26
33. 链向财经 (2018), 2013 年-2017 年比特币大事件, Available: https://mifengcha.com/news/5b74162ace79d2cf9bccf1af