



金融資料分析 —R語言的運用

文／吳牧恩 (東吳大學數學系助理教授)



在大數據蔚為風潮的時代，金融科技(FinTech)成為最近相當火熱的話題。以金融交易來說，市面上早有許多軟體，可針對各種商品進行策略回測，如MultiChart、TradeStation、MT4等。可想而知，想擁有這些軟體都須付出一定的費用，那麼假如今天有一個免費使用的工具，一樣可以回測各種交易策略，會不會引起使用者極大興趣？這工具就是「R語言」。

什麼是R語言？R語言是一種適合統計研究、資料分析、資料視覺化的程式語言。由於其簡單易學、免費開源(Open Source)的特性，在這個資料科學與資料大爆發的年代，R語言身價也跟著水漲船高。可以想見。如果將R語言運用於金融交易，那即使是一般投資人，也有能力開發屬於自己的看盤軟體，進行自動下單。

上面提到R語言「開源」的特色，是指全球R語言「玩家」，針對特別功能或是個別領域所開發的R語言套件 (Package)，無償分享給大家使用。以金融交易來說，大概就屬quantmod^[1]與quantstrat^[2]兩個套件最受歡迎。quantmod套件用來抓取金融資料、繪製行情走勢，而quantstrat套件則用來與quantmod做搭配，進行策略回測與績效分析。本文以quantmod套件作示範。

抓取資料與繪製線圖的好工具： quantmod

R語言可到其官方網站進行下載(<https://www.r-project.org/>)。安裝好R語言後，在R Console輸入install.packages("quantmod")指令進行quantmod套件安裝，隨後執行library(quantmod)便可開始使用。quantmod為Jeffrey A. Ryan^[3]所開發，其最方便的兩大功能分別為下載(getSymbols)歷史資料與繪製(chartSeries)行情走勢。例如Yahoo Finance上提供的全球股價資料、Oanda提供的外匯資料、Google Finance、Fred提供的各類金融商品資料等。

以臺灣加權指數為例，其在Yahoo Finance的代號為^TWII，只要輸入getSymbols("^TWII")指令，便可下載臺灣加權指數的日K資料。而下載後的資料也可用想要的視覺化方式呈現，例如只看2015年第四季的大盤日K線，輸入chartSeries(TWII, subset="2015-10-01::2015-12-31")即可，如圖1所示。

此外，quantmod亦可畫出各種熟悉的技術指標：均線、布林通道、KD、RSI指標等，且參數可自由設定。例如圖2是以23ma為均線，上下1.5標準差的布林通道(傳統布林通道參數為20ma，上下2標準差)，並搭配RSI指標。

以上這些都是傳統的技术指標，坊間看盤軟體也做得到。那什麼是一般看盤軟體做不到的？下

面以週選擇權研究為例，用quantmod繪製「特殊需求」的K線。

週三K線說明

期交所推出的週選擇權於2012年11月正式上市，三年半的時間發展迅速，實務上不僅成為一般投資人與法人的避險投資工具，學術界對週選擇權的研究也日益增多。而週選擇權是每週三上午 8:45開始交



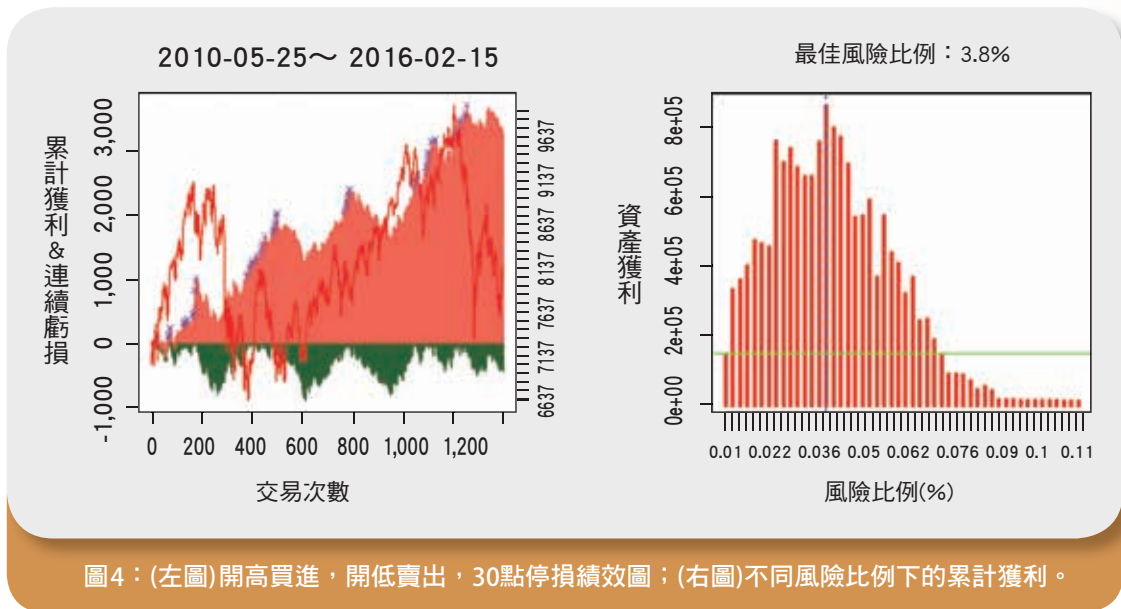
圖1：使用quantmod畫出2015年第四季臺灣加權股價指數。



圖2：使用quantmod畫出布林通道與RSI指標。



圖3：使用quantmod畫出2015年「週三K線」與季線。



易，次週三下午 1:30結束交易，其週期涵蓋6個交易日，與傳統週K線5個交易日不同。為了研究方便，以 quantmod 繪製每週三開盤，下個週三收盤的「週三K」。換句話說，在週三K線裡，每根K棒與下根K棒之間，都有週三整天的時間重疊(Overlap)，與傳統週K的開盤收盤結構有些許不同，而這樣的特殊功能，一般看盤軟體並不容易提供。圖3為用R語言畫出的週三K線(周三開盤，下周三收盤)，藍色曲線為週三K線的12ma，當季線使用。

交易理論與實務：使用R語言

就資料分析與策略回測而言，quantmod與quantstrat確實提供極佳的方便性，但R語言的功能當然不僅於此。R語言尤其適合統計分析與資料視覺化，以下舉兩個在臺股期貨歷史資料的實作案例：

案例1：最佳風險比例

當研發好一個交易策略，首先遇到的便是資金管理問題，尤其是計算最佳風險比例。茲以下述策略作說明：

◆**策略**：臺股期貨當沖交易，開高買進，開低賣出；停損30點，否則收盤平倉。

此策略在2010年5月25日~2016年2月15日期間累計的績效(不考慮滑價與手續費)如圖4.左所示。紫色x為獲利創新高的點，紅色曲線為相對應的臺股期貨走勢，以收盤價呈現。

左圖呈現的是採單口買進、賣出的累計損益點數，無任何資金管理機制；圖4.右則是使用15萬元資金，在不同風險比例下執行此策略的損益結果。從這實驗可看出，最佳風險比例為3.8% (藍色虛線)，資金成長可達到80多萬元(以小型臺指期貨計算)，也就是淨賺約65萬元。

但若是沒用資金管理機制，則只獲利約3,000點(圖4.左)，淨賺15萬元左右。此外，當風險比例達到6.5%以上，此策略便開始虧損。換句話說，3.8%獲利最多，6.5%卻開始虧損，這中間的差距僅有2.7%。由此可見，若沒有經過精準的分析，一個原本可獲利的策略，很有可能陷入高風險的窘境。

案例2：停損與停利的影響

研發交易策略時，必定會考慮如何停損與停利，而好的「停損停利機制」對績效絕對有重大影

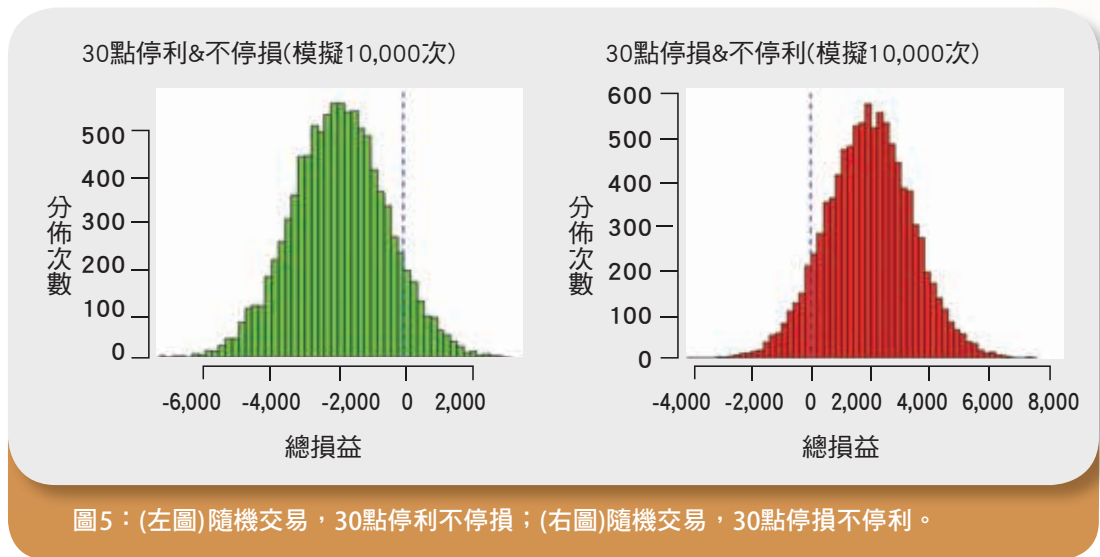


圖5：(左圖)隨機交易，30點停利不停損；(右圖)隨機交易，30點停損不停利。

響。下面是一個有趣的實驗，藉此說明其重要性：

- ◆策略I：開盤隨機交易，買賣機率各50%，30點停損不停利，收盤平倉。
- ◆策略II：開盤隨機交易，買賣機率各50%，30點停利不停損，收盤平倉。

由於是隨機交易，每次回測結果都不一樣，於是將策略I與策略II重複模擬10,000次，再觀察其損益分佈。圖5.左(綠色)與圖5.右(紅色)分別為為策略I(停利不停損)與策略II(停損不停利)各模擬回測10,000次後的損益分佈。一個有趣的現象是，圖5.右的紅色曲線幾乎像是圖5.左的綠色曲線直接往右平移(藍色虛線為損益兩平)。換句話說，不停損改成不停利，正報酬的機會也相對提高，這與一般人的投資行為似乎不同，此實驗除了說明台股期貨的動量效應外，也驗證了「停損停利機制」在策略研發上的重要性。

結語

使用R語言可做金融資料分析、策略回測、資金管理等研究案例。在資料科學的年代，最重要的還是要有Data，資料就像食材，而R語言是烹飪工具，等待善於調理的廚師(交易者)，烹飪出美麗的佳餚(策略績效)！🔗

〈以上內容不代表本刊及期交所立場〉



【註1】：Ryan, Jeffrey A. "quantmod: Quantitative financial modelling framework." *R package version 0.3-5*. URL <http://www.quantmod.com> URL <http://r-forge.r-project.org/projects/quantmod> (2008).
 【註2】：Georgakopoulos, Harry. "Backtesting with Quantstrat." *Quantitative Trading with R*. Palgrave Macmillan US, 2015. 147-175.
 【註3】：<http://www.quantmod.com/>