



國立臺灣大學電機資訊學院資訊工程學系

碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

使用極值理論評估具金融報酬特性的合成資料之

風險值與條件風險值

Evaluating the VaR and Conditional VaR of the Synthesized Data

Characteristic of Financial Returns

Using the Extreme Value Theory

蔡永信

Yong-Xin Tsai

指導教授：呂育道 博士

Advisor: Yuh-Daoh Lyuu, Ph.D.

中華民國 103 年 6 月

June, 2014

## 誌謝



因緣際會下，開始對經濟、金融等議題產生濃厚的興趣，因此，在尋找實驗室的時候，希望是針對相關議題作為研究目標，正好呂育道教授的計算理論暨金融計算實驗室有相關的研究，便請呂育道教授指導我的論文，開啟了我對金融計算的研究之窗。

在呂育道教授的指導下，我深深地感到獲益匪淺，不論是他的研究能力，抑或是待人處事的方法，都讓我學習到很多。而能完成這篇論文要感謝我的家人、朋友與指導教授的支持與幫助，在我遇到困難與挫折時，鼓舞我繼續前進與奮鬥，並且努力不懈地想出解決方法，最後完成目標。在研究所求學的期間裡，我增長了許多專業知識與技能，在這日新月異的世界，能與時漸進，跟上新科技的腳步。除此之外，在忙碌的研究之外，我抽空擔任國際志工，接觸不同的人文風情，除了增廣見聞，也體悟到生命不同的風貌與韻味。

最後，感謝所有在我迷惑困惑時，提供建議和幫助的朋友、師友，我將繼續攻讀博士，探索無窮無盡的知識領域。

## 中文摘要



全球金融危機激起人們對黑天鵝事件的關注，黑天鵝事件代表極值發生的事件，預測極值發生的週期與程度是極具有挑戰性。使用極值理論來估算風險值與條件風險值可以計算出在極端事件發生時，在特定信心水準下的資產損失與相對應的期望資產損失。本論文使用具有金融報酬特性的合成資料作為研究資料，可以避開真實市場資料所受的變數繁雜而不利分析的缺點。使用的合成資料分為靜態與動態兩種，靜態資料是由常態分布、Student's  $t$  分布或對數常態分布隨機抽樣產生的資產損失，動態資料是利用自我迴歸移動平均模型、自我迴歸條件異質變異數模型與 Student's  $t$  函數的殘差分布所產生的時間序列。利用極值理論估算風險值與條件風險值的方法區分為靜態方法與動態方法，靜態方法有區塊最大法與穿越門檻值法，動態方法利用動態模型、穿越門檻值法以及拔靴法來作估算。對於靜態資料而言，區塊最大法與穿越門檻值法對於不同的損失分布所估算出的結果，和樣本值或理論值的關係無法維持一定的趨勢。對於動態資料而言，動態方法所算出的結果，不論其變異數模型是否為定數，皆較於樣本值接近理論值但是小於理論值，具有一定的趨勢。

關鍵字：極值理論、風險值、條件風險值、區塊最大法、穿越門檻值模型、自我迴歸移動平均模型、非對稱性冪級數自我迴歸條件異質變異數模型

## 英文摘要



The global financial crisis in 2008 has raised concerns on the events so called “the Black Swans”. Evaluating value at risk and conditional value at risk using the extreme value theory can produce the asset loss and the corresponding expected asset loss at certain confidence level under extreme circumstances. This paper uses the synthetic data characteristic of financial returns as research targets, and in this way, we can avoid the drawbacks of analyzing the real market data which are affected by ambiguous variables. There are two types of synthesized data: static data and dynamic data. Static data are asset losses which are randomly sampled from the normal distribution, the Student’s  $t$  distribution and the log normal distribution, respectively. Dynamic data are time series generated by using the autoregressive moving average model, the autoregressive conditional heteroscedasticity model. There are two methods to evaluate the value at risk and the conditional value at risk using the extreme value theory: the static method and the dynamic method. The static method includes the block maxima method and the peaks over the threshold. The dynamic method integrates the dynamic model, the peaks over the threshold and the bootstrap sampling to evaluate the risk. For static data, there is no definite relation between statistics calculated from samples, theoretical values and the results calculated by the static method. For the dynamic data, there is a definite relation between the results calculated by the dynamic method, statistics calculated from samples, theoretic values and the results calculated by the dynamic method.

Keywords: extreme value theory, value at risk, conditional value at risk, block maxima method, peaks over the threshold, autoregressive moving average model, asymmetric power ARCH model

# 大綱



口試委員會審定書 .....	#
誌謝 .....	i
中文摘要 .....	ii
英文摘要 .....	iii
大綱 .....	iv
表格 .....	v
<b>第一章 研究動機.....</b>	<b>1</b>
<b>第二章 研究方法.....</b>	<b>3</b>
第一節 區塊最大法 (Block Maxima Model; BMM) .....	3
第二節 穿越門檻值法 (Peaks Over the Threshold ; POT) .....	5
第三節 動態方法.....	8
<b>第三章 分析與討論 .....</b>	<b>10</b>
第一節 合成資料說明.....	10
第二節 靜態方法估算風險值與條件風險值.....	12
第三節 動態模型估算風險值與條件風險值.....	15
<b>第四章 結論.....</b>	<b>17</b>
附錄 .....	18
參考資料 .....	19

## 表格



表格 一、靜態合成資料參數與統計摘要 .....	10
表格 二、動態合成資料參數與統計摘要 .....	11
表格 三、區塊最大法的配適結果 .....	12
表格 四、區塊最大法的估算結果 .....	13
表格 五、穿越門檻值法的配適結果 .....	13
表格 六、穿越門檻值法的估算結果 .....	14
表格 七、動態模型標準化殘差尾部 GPD 配適參數 .....	15
表格 八、動態模型估算在信心水準 0.95 與 0.99 下的風險值與條件風險值 .....	16

# 第一章 研究動機



以往，對於系統性風險所導致的極端事件，人們採取較不積極的態度與處置措施，但自從 2008 年全球金融危機（Global Financial Crisis; GFC）[15]之後，所謂的黑天鵝（The Black Swan; TBS）事件[16]，人們開始以積極的預防與管理這些事件，希望能將發生機率低但會造成很大的傷害性的事件做有效的風險控管。

極值理論（Extreme Value Theory; EVT）[5]由數理模型出發，架構良好的理論基礎，針對隨機分布的尾部發展出統計理論以估算極端事件發生的機率與數值；風險值（Value at Risk; VaR）[1][9]是一定的信心水準下風險的量化值，提供資產管理者對於特定機率發生資產損失的數值，使資產管理者可以採取預防性作為，例如調整資產配置來減低風險；條件風險值（Conditional Value at Risk; CVaR）[13]是低於某一機率下的資產損失期望值，可視為風險值的輔助統計量，讓資產管理者更全面地掌握風險。

由於金融時間序列在隨機分布上通常具有厚尾、波動聚集等特徵，使用傳統的金融模型配適（fitting）方法容易忽略隨機分布尾部的特徵而偏重於隨機分布本身的特性，因此，有文獻研究利用極值理論作為風險值與條件風險值的估算方法[3-4,6-8,10-11]，另外，也有論文在實證上發現極值理論估計的極值損失最正確[2]。

由於市場資料變因太繁雜容易影響資料的性質，造成在應用極值理論上的分析難度，因此本論文以合成資料作為分析資料。使用合成資料的好處是可以去除干擾因數，確保資料的純粹度，也可以推算理論值，作為對照數據，讓我們更容易評估極值理論估算出的風險值與條件風險值與理論值接近程度。除此之外，採用合成資料得到的結論，對於日後的市場資料研究上，也能提供參考的效果，提

供分析方向並降低詮釋資料的難度。





## 第二章 研究方法



本論文利用極值理論 (Extreme Value Theory ;EVT) 來捕捉尾部的特徵。由於 EVT 是描述極端事件發生的機率模型，因此，提供了一個適合捕捉極端尾部風險的方法，而一般來說，金融時間序列的隨機分布具有重尾 (heavy tailed) 特徵，例如，報酬分布的尾部[11]。風險值 (Value at Risk; VaR) 和條件風險值 (Conditional Value at Risk; CVaR or Expected Shortfall; ES) 是最常用來量化金融風險的指標，換句話說，極端事件發生時的風險，可透過 EVT 將其模型化並估算出其數值。

EVT 有兩種應用的方法。第一種方法稱為區塊最大法 (Block Maxima Model;BMM)，它利用包含 Gumbel、Fréchet 或 Weibull 的極值分布或以這三者的一般化極值分布 (Generalized Extreme Value Distribution; GEVD) 作為適配模型 (fitting models)，而第二種方法稱為穿越門檻值法 (Peaks Over The Threshold; POT)，它利用一般化帕雷托分布 (Generalized Pareto Distribution; GPD) 作為適配模型。POT 較 BMM 在資料的使用上有效率，這是因為 POT 對極值的認定較 BMM 寬鬆與有彈性，只要超過閾值 (Threshold) 的資料皆視為極值，不似 BMM 是在每一組相同長度的資料組中各只選取一個極值，所以 POT 在資料使用上較不浪費。

### 第一節 區塊最大法 (Block Maxima Model; BMM)

假設有一由相同分布函數  $F$  產生的獨立隨機變數序列  $X_1, \dots, X_n$ ，令  $M_n = \max(X_1, \dots, X_n)$ 。  $M_n$  小於或等於  $x$  的機率  $P$  為：

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{t=1}^n P(X_t \leq x) = F^n(x),$$

由定理 1 可知，  $M_n$  小於或等於  $x$  機率  $P$  的極限分布為極值分布  $G(x)$ 。



定理 1. (Fisher & Tippett, 1928 ; Gnedenko, 1943)

令  $X_1, \dots, X_n$  是一組獨立且相同分布的隨機變數序列，若存在一組正規化 (normalizing) 常數  $a_n > 0$  且  $b_n \in R$ ，使得

$$\lim_{n \rightarrow \infty} P\{a_n^{-1}(\max(X_1, \dots, X_n) - b_n) \leq x\} = G(x),$$

則對於一非退化性分布函數  $G(x)$ ， $G(x)$  屬於下列三種極值分布函數之一：

1. Fréchet：

$$\Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ e^{-x^{-\alpha}}, & x > 0 \end{cases} \quad \alpha > 0$$

2. Weibull：

$$\Psi_\alpha(x) = \begin{cases} e^{-(-x)^\alpha}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0$$

3. Gumbell：

$$\Lambda(x) = e^{-e^{-x}}, x \in R.$$

我們可以藉由一般化，將定理 1 中的三種極值分布函數集成一般化極值分布 (GEVD) 函數 (Von Mises, 1936 ; Jenkinson, 1955) 如下：

$$G_\xi(x) = \begin{cases} e^{-\left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi}}, & \xi \neq 0, \\ e^{-e^{-(x-\mu)/\sigma}}, & \xi = 0 \end{cases}$$

其中， $x$  滿足  $1 + \xi(x - \mu)/\sigma > 0$ 。上式中，極值指數 (extreme value index) 或形狀參數 (shape parameter)  $\xi$  滿足  $-\infty < \xi < \infty$ ，尺度參數 (scale parameter)  $\sigma$  滿足  $\sigma > 0$ ，位置參數 (location parameter)  $\mu$ 。根據形狀參數  $\xi$  的值， $G_\xi$  對應於三種極值分布：



1.  $\xi > 0$  即為 Fréchet 分布，其尾部以冪函數衰減。
2.  $\xi = 0$  即為 Gumbell 分布，其尾部以指數函數衰減。
3.  $\xi < 0$  即為 Weibull 分布，其右端尾部有端點。

因此，利用 GEVD 來對樣本的區塊最大值作配適 (fitting) 可以得到樣本的極值分布。區塊最大法可以用來估算重現水平 (return level)，重現水平為每經過  $k$  個時期，會出現一次超過此重現水平的觀察值並且等於  $G_{\xi}^{-1}(1 - 1/k)$ 。我們可以將重現水平與風險值的關係連接在一起，由於重現水平表示有  $1/k$  的機率出現的觀察值超過此重現水平，換句話說，有  $1 - 1/k$  的機率出現的觀察值小於或等於重現水平，意思等於在信心水準  $c = 1 - 1/k$  下的風險值定義 (參閱附錄)。所以，當信心水準為  $c$  時，其風險值為 [14]：

$$\text{VaR} = \begin{cases} \mu - \left(\frac{\sigma}{\xi}\right) \left(1 - (-\log(1-c))^{-\xi}\right) & , \xi \neq 0 \\ \mu - \sigma \log(-\log(1-c)) & , \xi = 0 \end{cases}$$

## 第二節 穿越門檻值法 (Peaks Over the Threshold ; POT)

若函數  $F(x) \in \text{RV}_{\xi}$  且  $\xi > 0$ ，表示對一緩慢改變 (slowly varying) 的函數  $L$ ， $F(x) = 1 - L(x)x^{-1/\xi}$ 。假設有一由相同分布函數  $F$  產生的獨立隨機變數序列  $X_1, \dots, X_n$ ， $F(x) \in \text{RV}_{\xi}$  且  $\xi > 0$ ，並令  $x_0$  為  $F$  之 (有限或無限) 右端端點，即  $x_0 = \sup\{x \in R : F(x) < 1\} \leq \infty$ 。定義超過門檻值  $u \in R$  之超額損失分布函數為：

$$F_u(x) = P\{X - u \leq x | X > u\} = \frac{F(x+u) - F(u)}{1 - F(u)}, \quad 0 \leq x \leq x_0 - u \dots\dots\dots (1)$$



$F_u(x)$  表示給定損失超過閾值  $u$  時，損失不多過  $x$  的機率。定理 2 證明超額分布的極限分布為一般化帕雷托分布函數 (the generalized Pareto distribution; GPD)。

定理 2. (Balkema & de Haan, 1974; Pickands, 1975)

$G_{\xi, \sigma(u)}$  是一般化帕雷托分布函數，其表示如下：

$$G_{\xi, \sigma(u)}(x) = \begin{cases} 1 - (1 + \xi x / \sigma)^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp(-x/\sigma), & \xi = 0 \end{cases}, \dots\dots\dots (2)$$

其中，規模參數  $\sigma > 0$ ，若形狀參數  $\xi \geq 0$ ，則  $x \geq 0$ ；若形狀參數  $\xi < 0$ ，則  $0 \leq x \leq -\sigma/\xi$ 。對一分布函數  $F$ ， $F \in RV_\xi$  且  $\xi > 0$  若且唯若存在  $\sigma(u) > 0$ ，使得

$$\lim_{u \rightarrow x_0} \sup_{0 \leq x < x_0 - u} |F_u(x) - G_{\xi, \sigma(u)}(x)| = 0。$$

由於  $x$  為超額值 (excess)， $u$  為閾值，可令  $y = u + x$  為超過閾值  $u$  額度  $x$  的變數，且  $u \leq y \leq x_0$ ，則  $G_{\xi, \sigma(u)}(x)$  可表為：

$$G_{\xi, \sigma(u), u}(y) = \begin{cases} 1 - (1 + \xi \frac{y-u}{\sigma})^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp(-\frac{y-u}{\sigma}), & \xi = 0 \end{cases}。 \dots\dots\dots (3)$$

根據形狀參數  $\xi$  的值，有三種極限 GPD：

1.  $\xi > 0$ ， $F$  是重尾分布 (heavy-tailed distributions)，其尾部以冪函數 (power function) 衰減，例如，Pareto、Student's t、Cauchy、Burr、log gamma 和 Fréchet 分布。
2.  $\xi = 0$ ， $F$  其尾部以指數函數衰減，例如，normal、exponential、gamma 和 lognormal 分布。

3.  $\xi < 0$ ， $F$  是具有有限右端點的短尾分布，例如，uniform 和 beta 分布。

將  $y = u + x$  代入 (1)，可得

$$F(y) = (1 - F(u))F_u(x) + F(u) \quad , \quad \dots\dots\dots (4)$$

令  $N_u$  為超過門檻值  $u$  的樣本個數，則

$$F(u) = 1 - \frac{N_u}{n} \quad , \quad \dots\dots\dots (5)$$

將(3)、(5)代入(4)，可得尾部分布估計式為[14]：

$$F(y) = \begin{cases} 1 - \frac{N_u}{n} \left[ 1 + \xi \frac{(y-u)}{\sigma} \right]^{-\frac{1}{\xi}} , & \xi \neq 0 \\ 1 - \frac{N_u}{n} \exp\left(-\frac{y-u}{\sigma}\right) , & \xi = 0 \end{cases} \quad , \quad \dots\dots\dots (6)$$

因此，可利用 GPD 來計算風險值（Value at Risk; VaR）與條件風險值（Conditional

Value at Risk; CVaR 或 Expected Shortfall; ES）。對於  $y > u$ ，尾部機率  $\bar{F}(y)$  為：

$$\begin{aligned} \bar{F}(y) &= P(X > u)P(X > y | X > u) \\ &= \bar{F}(u)P(X - u > y - u | X > u) \\ &= \bar{F}(u)\bar{F}_u(y - u) \\ &= \begin{cases} \bar{F}(u) \left( 1 + \xi \frac{y-u}{\sigma} \right)^{-\frac{1}{\xi}} , & \xi \neq 0 \\ \bar{F}(u) \exp\left(-\frac{y-u}{\sigma}\right) , & \xi = 0 \end{cases} \quad , \end{aligned}$$

根據風險值的定義（參考附錄），在信心水準  $\alpha$  下的風險值即為機率  $\alpha$  下 GPD 的

分位數，可表示為：

$$\text{VaR}_\alpha = q_F(\alpha) = \begin{cases} u + \frac{\sigma}{\xi} \left( \left( \frac{1-\alpha}{\bar{F}(u)} \right)^{-\xi} - 1 \right) , & \xi \neq 0 \\ u - \sigma \log\left(\frac{1-\alpha}{\bar{F}(u)}\right) , & \xi = 0 \end{cases} \quad ,$$

而其相對應的條件風險值為：

$$\text{CVaR}_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_x dx = \begin{cases} \frac{\text{VaR}_\alpha}{1-\xi} + \frac{\sigma - \xi u}{1-\xi}, & \xi \neq 0 \\ \text{VaR}_\alpha + \sigma, & \xi = 0 \end{cases}$$



### 第三節 動態方法

極值理論(EVT)可以應用在動態模型(dynamic model) [10]中， $R_t$  為時間  $t$  之對數報酬， $\{R_t\}$  是穩定的時間序列，其動態模型如下：

$$\begin{cases} R_t = \mu_t + \varepsilon_t, \\ \varepsilon_t = \sigma_t Z_t \end{cases},$$

其中， $\mu_t$  是時間  $t$  時的預期報酬率， $\varepsilon_t$  是時間  $t$  時的誤差， $\sigma_t$  是時間  $t$  時的波動率， $Z_t$  是時間  $t$  時滿足  $F_Z(z)$  分布的隨機變數。

假設  $\mu_t$  滿足級數  $(p, q)$  的自我迴歸移動平均模型 (autoregressive moving average model of order  $(p, q)$ ; ARMA  $(p, q)$ ) 且  $\sigma_t$  滿足級數  $(p, q)$  的非對稱性冪級數自我迴歸條件異質變異數模型 (asymmetric power ARCH model of order  $(p, q)$ )，則動態模型可以進一步表為：

$$\begin{cases} R_t = \mu_t + \varepsilon_t \\ \varepsilon_t = \sigma_t Z_t \\ \mu_t = \sum_{i=1}^p \phi_i \mu_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \\ \sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta \end{cases},$$

其中， $\delta, \omega, \alpha, \beta > 0$  且  $|\gamma| < 1$ 。

我們將利用上述動態模型來配適合成資料，並可計算出隨機變數  $Z_t$ ，接著再

以 POT 分別對  $Z_t$  的分布函數  $F_Z(z)$  之右、左端尾部配適出 GPD 參數值，分別為  $\text{GPD}(\hat{\xi}^{(1)}, \hat{\beta}^{(1)})$  和  $\text{GPD}(\hat{\xi}^{(2)}, \hat{\beta}^{(2)})$ 。

利用拔靴 (bootstrap) 法和  $F_Z(z)$  的尾部配適出的  $\text{GPD}(\hat{\xi}^{(1)}, \hat{\beta}^{(1)})$ 、 $\text{GPD}(\hat{\xi}^{(2)}, \hat{\beta}^{(2)})$  來模擬  $Z_{t+1}, \dots, Z_{t+h}$  (Danielsson 與 de Vries, 1997c)，以計算  $h > 1$  日後之風險值，重置取樣的演算法如下：

1. 假設有  $n$  個殘差樣本，其次序統計值由大到小分別為  $z_{(1)}, \dots, z_{(k)}, \dots, z_{(n-k)}, \dots, z_{(n)}$ 。  
隨機從中選取一殘差  $z$ 。
2. 若  $z > z_{(k+1)}$ ，則利用  $\text{GPD}(\hat{\xi}^{(1)}, \hat{\beta}^{(1)})$  抽樣出一超額值  $y^{(1)}$ ，令  $z = z_{(k+1)} + y^{(1)}$ 。
3. 若  $z < z_{(n-k)}$ ，則利用  $\text{GPD}(\hat{\xi}^{(2)}, \hat{\beta}^{(2)})$  抽樣出一超額值  $y^{(2)}$ ，令  $z = z_{(n-k)} + y^{(2)}$ 。
4. 否則， $z = z$ 。

由上述重置抽樣演算法可知，模擬之取樣分布如下[10]：

$$F_Z(z) = \begin{cases} \frac{k}{n} \left( 1 + \hat{\xi}^{(2)} \frac{|z - z_{(n-k)}|}{\hat{\beta}^{(2)}} \right)^{-1/\hat{\xi}^{(2)}}, & z < z_{(n-k)} \\ \frac{1}{n} \sum_{i=1}^n 1_{\{z_i \leq z\}}, & z_{(n-k)} \leq z \leq z_{(k+1)} \\ 1 - \frac{k}{n} \left( 1 + \hat{\xi}^{(1)} \frac{|z - z_{(k+1)}|}{\hat{\beta}^{(1)}} \right)^{-1/\hat{\xi}^{(1)}}, & z > z_{(k+1)} \end{cases}$$

經過重置抽樣  $h$  次之後，可得一組路徑，即  $h$  日後之報酬。重複上述動作 1000 次，將所得之 1000 組  $h$  日後之報酬以 POT 配適，利用本章第二節所述方法，計算在信心水準  $\alpha$  下的  $h$  日後報酬之 VaR 和 CVaR。

## 第三章 分析與討論

### 第一節 合成資料說明



本節將說明合成資料的性質與統計值。這些合成資料分為靜態與動態兩種類型，靜態合成資料代表資產損失，動態合成資料代表日對數報酬 (daily log return)，靜態合成資料具有獨立與相同隨機分布 (i.i.d.) 的特性，動態合成資料則無。另外，根據台灣經濟新報資料庫的數據顯示，自 2006 年起至 2013 年止，每年台灣證券交易日平均約為 248.6 天，我們選擇資料期間約四年，即每組合成資料的樣本數為 1000 個。表格一為靜態合成資料 Data 1-3 的隨機分布模型設定、參數值以及統計摘要，資料樣本是抽樣自隨機分布模型，因此，除 Data 1 之外，Data 2 和 Data 3 的樣本分布具有重尾特性。

	Data 1	Data 2	Data 3
分布模型	Normal	Student's t	Log Normal
樣本數	1000	1000	1000
參數值	平均值：0；標準差：1	自由度：2.100004	平均值：0；標準差：1
統計摘要			
最小值	-3.8676857	-4.0130858	0.0193746
中位數	-0.0501418	0.0362514	1.031124
最大值	3.2333201	4.3018314	24.61767
平均值	-0.0473285	0.0270547	1.645078
標準差	1.0290541	1.0551507	2.044874
1% 分位數	-2.4281244	-2.5835362	0.0961076
5% 分位數	-1.7596587	-1.6812342	0.1939124
95% 分位數	1.5964619	1.7200769	5.1459189
99% 分位數	2.2157278	2.6731476	8.8826206

表格一、靜態合成資料參數與統計摘要

表格二為動態合成資料 Data 4-6 的時間序列模型設定、參數值以及統計摘要，Data 4-6 都具有自我迴歸的特性且標準殘差機率分布皆是 Student's t 函數，



Data 4 與 Data 5 的條件變異數是常數值，因此不隨時間改變，Data 6 的條件變異數具有非對稱與槓桿作用並隨著時間改變。

	Data 4	Data 5	Data 6
樣本數	1000	1000	1000
平均值模型	ARMA(2,0)	ARMA(1,0)	ARMA(1,0)
參數設定	$\varphi_1 = -0.0126806$ $\varphi_2 = 0.0197513$	$\varphi_1 = 0.0354501$	$\varphi_1 = 0.0631909$
變異數模型	GARCH(0,0)	GARCH(0,0)	APARCH(1,1)
參數設定	$\omega = 0.0001572$	$\omega = 0.0001135$	$\omega = 0.0002010$ $\alpha_1 = 0.9999998$ $\beta_1 = 0.6676622$ $\gamma_1 = -0.0528787$ $\delta = 0.8363420$
分布模型	Student's t	Student's t	Student's t
參數設定	2.1000000	2.1000000	2.1000000
統計摘要			
最小值	-4.988509e-02	-4.248082e-02	-5.251681e-02
中位數	-1.282438e-04	-9.261252e-05	-1.381030e-05
最大值	5.973905e-02	5.086999e-02	5.961762e-02
平均值	-9.206700e-05	-8.051886e-05	4.448447e-05
標準差	6.353634e-03	5.405311e-03	5.456616e-03
1% 分位數	-1.79073e-02	-1.544724e-02	-1.404707e-02
5% 分位數	-7.756171e-03	-6.831265e-03	-4.44739e-03
95% 分位數	7.710881e-03	6.536045e-03	3.344346e-03
99% 分位數	1.727889e-02	1.476519e-02	1.617252e-02

表格 二、動態合成資料參數與統計摘要

## 第二節 靜態方法估算風險值與條件風險值

本節將分析使用區塊最大法 (BMM) 與穿越門檻值法 (POT) 所得之風險值、條件風險值與其理論值、樣本值之差異。表格三是利用區塊最大法配適 Data 1-3 後所得的參數值，Data 1 配適得到的形狀參數為負值，其餘兩組資料配適得到的形狀參數皆為正值。

	Data 1	Data 2	Data 3
分布函數	Normal	Student's t	Log Normal (右尾)
區塊大小	10	10	10
位置參數	1.2578403	1.45919513	3.4367301
尺度參數	0.4764328	0.58583832	1.9259280
形狀參數	-0.1652787	0.03382198	0.4621842

表格 三、區塊最大法的配適結果

由表格 四可知，Data 1 和 Data 2 利用區塊最大法所算出風險值和條件風險值的估算值都較理論值小，而 Data 3 利用區塊最大法所算出風險值和條件風險值的估算值都較理論值大。另一方面，Data 1 和 Data 2 利用區塊最大法所算出的風險值和條件風險值估算值都較樣本值大，而 Data 3 除信心水準 0.95 下的風險值略小於樣本值之外，其餘的估算值也比樣本值大。



	Data 1	Data 2	Data 3
區塊最大法估算值			
$VaR_{0.95}^{BM}$	2.376091	3.289652	15.71399
$VaR_{0.99}^{BM}$	2.792760	4.375096	34.19867
$CVaR_{0.95}^{BM}$	2.629309	3.968846	30.08258
$CVaR_{0.99}^{BM}$	2.984345	5.085328	64.31384
樣本值			
$VaR_{0.95}$	2.185469	3.302413	17.61461
$VaR_{0.99}$	2.683588	3.997241	27.77107
$CVaR_{0.95}$	2.539559	3.766499	23.48106
$CVaR_{0.99}$	2.932827	4.680861	29.54723
理論值			
$VaR_{0.95}$	2.567875	9.064028	13.03809
$VaR_{0.99}$	3.088890	19.82683	21.95270
$CVaR_{0.95}$	2.888189	17.53800	18.91264
$CVaR_{0.99}$	3.366435	37.95408	30.15263

表格 四、區塊最大法的估算結果


表格 五是利用區塊最大法配適 Data 1-3 後所得到的參數值，可以發現 Data 1-2 的形狀參數皆為負值，而 Data 3 配適得到的形狀參數接近零。

	Data 1	Data 2	Data 3
分布函數	Normal	Student's t	Log Normal
位置參數	1.0	2.1	7.5
尺度參數	0.6107301	0.756037	6.833295
形狀參數	-0.2301586	-0.270565	1.075773e-07

表格 五、穿越門檻值法的配適結果

由表格 六可以發現 Data 1 和 Data 3 利用穿越門檻執法所估算出的風險值與條件風險值皆大於理論值，且 Data 3 的估算值遠大於理論值，但 Data 2 利用穿越門檻值法所估算出的風險值與條件風險值，除了信心水準 0.95 下的風險值之外，其

餘皆小於理論值。另一方面，Data 1-3 利用穿越門檻執法所估算出的風險值與條件風險值皆大於樣本值，其中，Data 3 得到的估算值大於樣本值甚多。



	Data 1	Data 2	Data 3
穿越門檻值法估算值			
$\text{VaR}_{0.95}^{\text{POT}}$	2.321903	3.651899	27.97073
$\text{VaR}_{0.99}^{\text{POT}}$	2.734118	4.090502	38.9685
$\text{CVaR}_{0.95}^{\text{POT}}$	2.571043	3.916462	34.80402
$\text{CVaR}_{0.99}^{\text{POT}}$	2.90613	4.26166	45.8018
樣本值			
$\text{VaR}_{0.95}$	1.583641	1.842208	4.781287
$\text{VaR}_{0.99}$	2.087136	2.795539	13.36124
$\text{CVaR}_{0.95}$	1.915886	2.454667	9.532688
$\text{CVaR}_{0.99}$	2.334944	3.406972	19.80156
理論值			
$\text{VaR}_{0.95}$	1.644854	2.827332	5.180252
$\text{VaR}_{0.99}$	2.326348	6.530806	10.24047
$\text{CVaR}_{0.95}$	2.062713	5.723482	8.557227
$\text{CVaR}_{0.99}$	2.665204	12.61582	15.22796

表格 六、穿越門檻值法的估算結果

### 第三節 動態模型估算風險值與條件風險值

本節將分析使用動態方法所得之風險值、條件風險值與其理論值、樣本值之差異。我們利用表格 七所列的平均值模型、變異數模型與殘差分布函數分別對 Data 4-6 作配適，將配適後得到的標準化殘差分左、右尾作 GPD 配適得到如表格 七中的位置、尺度與形狀參數值。

	Data 4	Data 5	Data 6
平均值模型	ARMA(2,0)	ARMA(1,0)	ARMA(1,0)
變異數模型	GARCH(0,0)	GARCH(0,0)	APARCH(1,1)
殘差分布函數	Student's t	Student's t	Student's t
右尾 GPD 配適參數			
位置參數	0.65	0.4	1.25
尺度參數	0.4754582	0.3463844	0.9891233
形狀參數	0.3807341	0.4274878	0.3621177
左尾 GPD 配適參數			
位置參數	0.7	0.45	1.75
尺度參數	0.5183988	0.3516828	1.0668650
形狀參數	0.2404024	0.3560271	0.3595357

表格 七、動態模型標準化殘差尾部 GPD 配適參數

為了比較動態模型所估算出風險值與條件風險值的準確度，我們計算了樣本值與理論值，樣本值是利用第二章第三節所述的抽樣方法，計算十萬組路徑，不用 POT 估算風險值與條件風險值，而是直接以這些樣本計算風險值與條件風險值，而理論值是利用合成資料的參數模型及殘差分布函數，抽樣出十萬組路徑並直接以這些樣本計算風險值與條件風險值。我們估算了五日後的報酬風險值（ $\text{VaR}^{h=5}$ ）與條件風險值（ $\text{CVaR}^{h=5}$ ），其結果如表格 八所示。比較動態模型估算的結果與樣本值發現，Data 4-6 利用動態模型所計算出的風險值與條件風險值均較樣本值大。此外，將動態模型估算的結果與理論值比較後，動態模型估算的結

果確較理論值小，顯示透過動態模型估算無法完全掌握實際的報酬風險分布，但是比缺乏 POT 估算的樣本值更接近理論值。



	Data 4	Data 5	Data 6
動態模型			
$\text{VaR}_{0.95}^{h=5}$	0.0434507	0.0328840	0.0025929
$\text{VaR}_{0.99}^{h=5}$	0.0669701	0.0582264	0.0052407
$\text{CVaR}_{0.95}^{h=5}$	0.0585616	0.0498168	0.0044469
$\text{CVaR}_{0.99}^{h=5}$	0.0856590	0.0829288	0.0083634
樣本值			
$\text{VaR}_{0.95}^{h=5}$	0.02182674	0.0188037	0.001461294
$\text{VaR}_{0.99}^{h=5}$	0.04088665	0.03550512	0.003634361
$\text{CVaR}_{0.95}^{h=5}$	0.03302242	0.02866820	0.003044188
$\text{CVaR}_{0.99}^{h=5}$	0.05216540	0.04607359	0.006661762
理論值			
$\text{VaR}_{0.95}^{h=5}$	0.0953288	0.0831061	0.0541113
$\text{VaR}_{0.99}^{h=5}$	0.1924434	0.1677148	0.2071283
$\text{CVaR}_{0.95}^{h=5}$	0.1689549	0.1468922	0.7504487
$\text{CVaR}_{0.99}^{h=5}$	0.3372686	0.2915243	3.256589

表格 八、動態模型估算在信心水準 0.95 與 0.99 下的風險值與條件風險值

## 第四章 結論



在本論文中，我們使用極值理論（EVT），以靜態與動態的方法，估算具有金融市場特性的合成資料之風險值與條件風險值，並比較估算出的結果與樣本值、理論值間的差異，得到以下的結論。對於損失分布為常態分布（normal distribution）與 Student's t 分布的靜態資料，利用區塊最大法估計出的風險值與條件風險值比起樣本值更接近理論值，但對於損失分布為對數常態分布（lognormal distribution）的靜態資料，卻會超過理論值。對於損失分布為常態與對數常態靜態資料，利用穿越門檻值法估計出的風險值與條件風險值超過理論值，但對於損失分布為 Student's t 分布的靜態資料，穿越門檻值法估計出的風險值與條件風險值小於理論值，顯示出區塊最大法與穿越門檻值法對於不同的損失分布無法維持一定的趨勢。另一方面，動態模型所估算出的風險值與條件風險值對於具有自我回歸的對數報酬時間序列，不論其變異數模型是否為定數，皆較於樣本值接近理論值但小於理論值，具有一定的趨勢。

## 附錄



定義 1. 風險值 (value at risk ; VaR) (Jorion , 1997)

給定信心水準  $\alpha \in (0,1)$  , 在期間  $t$  內且信心水準  $\alpha$  下的資產風險值是指最小的  $k \in R$  , 使得資產在期間  $t$  內的損失  $L$  超越  $k$  的機率為  $1-\alpha$  。風險值的數學表示：

$$\text{VaR}_{\alpha}(L) = \inf \{l \in R : P(L > l) \leq 1 - \alpha\} = \inf \{l \in R : F_L(l) \geq \alpha\} ,$$

其中 ,  $P(L > l)$  代表資產損失  $L$  大於  $l \in R$  的機率函數 ,  $F_L(l)$  代表資產損失  $L$  小於或等於  $l \in R$  的機率分布函數。

定義 2. 條件風險值 (conditional value at risk ; CVaR) (Rockafellar & Uryasev , 2002)

條件風險值又稱為尾部風險值 (tail VaR)、期望差額 (expected shortfall) 或平均超額損失 (mean excess loss)。根據定義 1, 我們可以定義條件風險值的數學表示為：

$$\text{CVaR}_{\alpha} = \frac{1}{1-\alpha} \int_{\alpha}^1 \text{VaR}_x dx = \frac{1}{1-\alpha} \int_{\alpha}^1 F_L^{-1}(x) dx ,$$


其中 ,  $\alpha$  為信心水準 ,  $F_L^{-1}(x)$  為資產損失  $L$  的機率分布函數的反函數 , 即分位數函數 (quantile functions)。



## 參考資料



- [1] Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. “Coherent measures of risk.” *Mathematical Finance* 9, no. 3 (1999): 203–228.
- [2] Bekiros, Stelios D., and Dimitris A. Georgoutsos. “Estimation of Value-at-Risk by extreme value and conventional methods: a comparative evaluation of their predictive performance.” *Journal of International Financial Markets, Institutions and Money* 15, no. 3 (2005): 209–228.
- [3] Danielsson, Jon, and Casper G. De Vries. “Value-at-risk and extreme returns.” *Annales d'Economie et de Statistique* 60 (2000): 239–270.
- [4] Danielsson, Jon, Laurens de Haan, Liang Peng, and Casper G. de Vries. “Using a bootstrap method to choose the sample fraction in tail index estimation.” *Journal of Multivariate Analysis* 76, no. 2 (2001): 226–248.
- [5] De Haan, Laurens, and Ana Ferreira. *Extreme value theory: An introduction*. New York: Springer, 2007.
- [6] Drees, Holger. “Extreme quantile estimation for dependent data, with applications to finance.” *Bernoulli* 9, no. 4 (2003): 617–657.
- [7] Embrechts, Paul, Sidney I. Resnick, and Gennady Samorodnitsky. “Extreme value theory as a risk management tool.” *North American Actuarial Journal* 3, no. 2 (1999): 30–41.
- [8] Gençay, Ramazan, and Faruk Selçuk. “Extreme value theory and value-at-risk: relative performance in emerging markets.” *International Journal of Forecasting* 20, no. 2 (2004): 287–303.

- 
- [9] Jorion, Philippe. *Value at risk: the new benchmark for controlling market risk*. Vol. 2. New York: McGraw–Hill, 1997.
- [10] McNeil, Alexander J., and Rüdiger Frey. “Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach.” *Journal of Empirical Finance* 7, no. 3 (2000): 271–300.
- [11] Poon, Ser-Huang, Michael Rockinger, and Jonathan Tawn. “Extreme value dependence in financial markets: Diagnostics, models, and financial implications.” *Review of Financial Studies* 17, no. 2 (2004): 581–610.
- [12] Resnick, Sidney I. *Heavy-tail phenomena: Probabilistic and statistical modeling*. New York: Springer, 2007.
- [13] Rockafellar, R. Tyrrell, and Stanislav Uryasev. “Conditional value-at-risk for general loss distributions.” *Journal of Banking & Finance* 26, no. 7 (2002): 1443–1471.
- [14] Singh, Abhay K., David E. Allen, and Powell J. Robert. “Extreme market risk and extreme value theory.” *Mathematics and Computers in Simulation* 94 (2013): 310–328.
- [15] Shiller, Robert J. *The subprime solution: How today’s global financial crisis happened, and what to do about it*. Princeton, NJ: Princeton University Press, 2008.
- [16] Taleb, Nassim Nicholas. “Black swans and the domains of statistics.” *The American Statistician* 61, no. 3 (2007): 198–200.