3 研究方法

圖 3 爲本文主要研究流程圖。本文以 Glosten, Jagannathan and Runkle (1993) 提出的 GJR-GARCH 模型爲基礎, 白噪音分配加入厚尾與不對稱的特性, 針對不同分配假設進行配適, 再以模型衡量指標比較不同分配配適程度。

此外以一步預測法 (one-step-ahead forecasting) 做樣本外預測, 進一步結合動態極值理論與動態歷史模擬法, 透過回溯測試 (back-testing) 檢定在不同分配假設下風險值預測的可靠度。最後使用損失函數 (loss function) 選擇最適的模型。

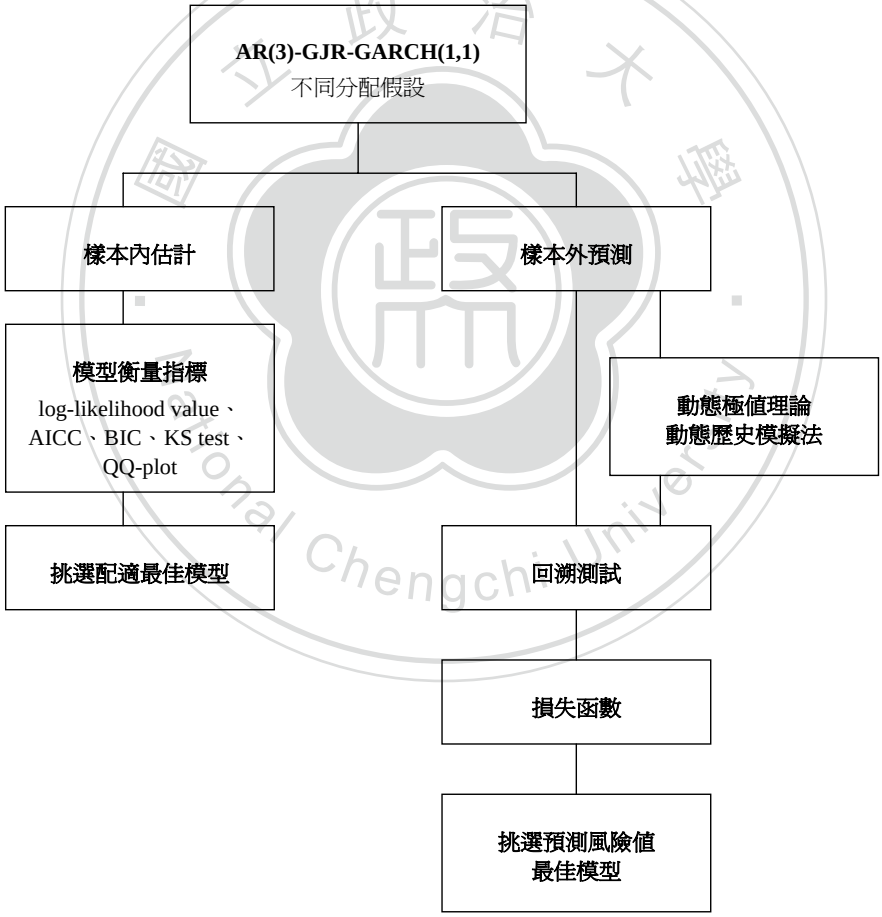

圖 3: 研究流程

3.1 AR-GJR-GARCH

重新檢視第 (11) 式, 以 GARCH 模型估計風險值, 其風險值估計爲

$$VaR_\alpha = \hat{\mu} + \hat{\sigma}_t WN^{-1}(\alpha)$$

風險值估計可分成三個部份, $\hat{\mu}$ 爲條件一階動差的估計, $\hat{\sigma}_t$ 爲條件二階動差的估計, $WN^{-1}(\alpha)$ 爲白噪音分配第 α 分量。因此若希望準確的估計風險值, 必須掌握其一階動差、二階動差及白噪音的分配。

爲了進一步捕捉條件變異數在報酬率正負之間可能存在不對稱的情形, 在此, 我們使用 Glosten, Jagannathan and Runkle (1993) 提出的 GJR-GARCH 模型。GJR-GARCH 模型以 0 爲門檻, 刻劃前期報酬率正負之間條件變異數不對稱的情形。而爲了捕捉條件一階動差, 進一步加入自我相關。一般化的 AR(p)-GJR-GARCH(m, s) 模型可表示爲

$$\begin{aligned}
r_t &= \mu_t + \varepsilon_t \\
\mu_t &= \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} \\
h_t &= \alpha_0 + \sum_{i=1}^{m} (\alpha_i + I_i \gamma_i)\, \varepsilon_{t-1}^2 + \sum_{i=1}^{s} \beta_i h_{t-i} \\
\varepsilon_t &\sim WN(0, h_t)
\end{aligned}$$

其中 I_i 爲指標函數, 當 $\varepsilon_{t-1} \leq 0$ 時, $I_i = 1$; 當 $\varepsilon_{t-1} > 0$ 時, $I_i = 0$ 。爲確保變異數爲正且模型穩定, 還須滿足 $0 \leq \alpha_i, \beta_i, \gamma_i \leq 1$, $\sum \alpha_i + \sum \beta_i + 0.5 \sum \gamma_i < 1$ 的條件。

3.2 白噪音設定

實證資料顯示, 財務資料通常存在厚尾及不對稱的特性, 白噪音爲常態分配的假設下, 模型配適與風險值估計結果通常都不盡理想。因此愈來愈多的研究改採用厚尾及不對稱的分配來配適, 文獻上最常使用的厚尾分配可分爲 Student t 分配一族與 EPD 一族。

圖 4 顯示 Student t 分配一族與 EPD 一族之間的關係, 本文白噪音的機率分配假設爲 normal 、skew-normal 、Student t 、skew-t 、EPD 、SEPD 、AEPD 等七種分配, 以最大概似法參數估計, 並以概似比檢定 (likelihood ratio test, LR test) 虛無假設

$$M : \hat{\theta} = \theta$$

雖然 EPD 分配在參數 $p = 2$ 與 Student t 分配自由度 $\nu \to \infty$ 時, 皆收斂至常態分配, 但在兩者皆爲厚尾分配時 (EPD 分配參數 $p < 2$), 機率分配是不同的型態。

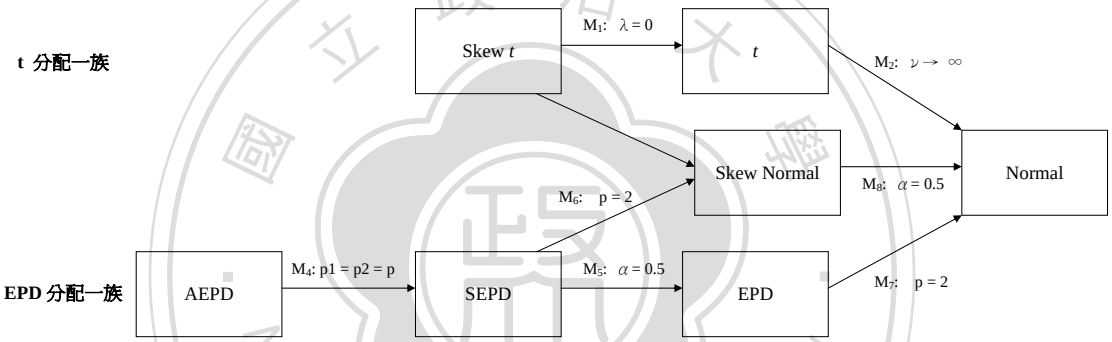

圖 4: Student t 分配一族與 EPD 一族關係圖

Student t 分配一族

Student t 分配爲最廣泛使用的厚尾分配, 其機率密度函數表示如下

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \tag{14}$$

爲了更符合白噪音期望值爲 0、變異數爲 1 的要求, 在此我們採用標準化 t 分配, 其機率密度函數表示如下

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi(\nu-2)}\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu-2}\right)^{-\frac{\nu+1}{2}} \tag{15}$$

Hansen (1994) 提出一般化的 Student t 分配, 增加一不對稱參數 λ , 一般稱之爲 skew-t 分配。機率密度函數表示如下

$$f(x) = \begin{cases} bc\left(1+\frac{1}{\nu-2}\left(\frac{bx+a}{1-\lambda}\right)^2\right)^{-\frac{\nu+1}{2}} & \text{if } x < -\frac{a}{b} \\ bc\left(1+\frac{1}{\nu-2}\left(\frac{bx+a}{1+\lambda}\right)^2\right)^{-\frac{\nu+1}{2}} & \text{if } x \geq -\frac{a}{b} \end{cases} \tag{16}$$

其中

$$a = 4\lambda c\left(\frac{\nu-2}{\nu-1}\right),\ b^2 = 1+3\lambda^2-a^2\ ,\ c = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)\Gamma\left(\frac{\nu}{2}\right)}}$$

skew-t 分配期望值爲 0 , 變異數爲 1 , 參數 ν 爲自由度。若 $\lambda > 0$ 時, skew-t 分配爲右偏分配; 若 $\lambda < 0$ 時, skew-t 分配爲左偏分配; 若 $\lambda = 0$, skew-t 分配退化爲標準化的 t 分配; 若 $\lambda = 0$ 且 $\nu \to \infty$, skew-t 分配退化爲標準常態分配。

以最大概似法參數估計, 對數概似函數爲

$$l(\theta; r) = \sum_{t=1}^{T}\left\{-\ln h_t + \log f_y\left(\frac{r_t-\mu_t}{h_t}\middle|\beta\right)\right\}$$

其中 $\beta = (\nu, \lambda)'$, $f_y(\cdot|\beta)$ 爲 skew-t 分配機率密度函數。

因此白噪音假設爲 t 分配一族, 風險值估計爲

$$VaR = \hat{\mu}_t + \hat{\sigma}_t F^{-1}(\alpha) \tag{17}$$

其中 F^{-1} 爲 skew-t 分配的分量函數

$$F^{-1}(u;\nu,\lambda) = \begin{cases} \frac{1}{b}\left((1-\lambda)\sqrt{\frac{\nu-2}{\nu}}G^{-1}\left(\frac{u}{1-\lambda};\nu\right) - a\right) & \text{if } \ 0 < u < \frac{1-\lambda}{2} \\ \frac{1}{b}\left((1+\lambda)\sqrt{\frac{\nu-2}{\nu}}G^{-1}\left(\frac{u+\lambda}{1+\lambda};\nu\right) - a\right) & \text{if } \ \frac{1-\lambda}{2} \leq u < 1 \end{cases}$$

G^{-1} 爲 Student t 分配自由度爲 ν 的分量函數。

EPD 分配一族

EPD 分配又稱爲 generalized error distribution (GED) 或 generalized laplace distribution , [4] 機率密度函數表示如下

$$f(x) = \frac{1}{\sigma}K_{EP}(p)\exp\left(-\frac{1}{p}\left|\frac{x-\mu}{\sigma}\right|^p\right) \tag{18}$$

其中 $K_{EP}(p) = \dfrac{1}{2p^{1/p}\Gamma(1+1/p)}$

EPD 分配參數 $\mu = E(x) = Med(x)$, $\sigma = (E|x-\mu|^p)^{1/p}$, 在此, 我們使用標準 EPD 分配 (令 $\mu = 0$, $\sigma = 1$) 。標準 EPD 分配當 $p > 2$ 時爲細尾分

[4]一般常見的 GED 分配函數型式表示如下

$$g(z) = \frac{p}{2^{1+1/p}\lambda\Gamma(1/p)}\exp\left(-0.5\left|\frac{z}{\lambda}\right|^p\right)$$

其中

$$\lambda = \left(\frac{2^{-2/p}\Gamma(1/p)}{\Gamma(3/p)}\right)$$

GED 分配可視爲將標準 EPD 分配標準化後, 使其期望值爲 0 , 變異數爲 1 。

配; 當 $p < 2$ 時, 爲厚尾分配; 當 $p = 2$ 時, 標準 EPD 分配退化爲標準常態分配; 當 $p = 1$ 時, 標準 EPD 分配退化爲拉普拉斯分配 (Laplace distribution)。

Zhu and Zinde - Walsh (2009) 提出更爲一般化的 EPD 分配, 稱之爲 asymmetric exponential power distribution (AEPD), 其機率密度函數表示如下

$$f(x) = \begin{cases} \left(\dfrac{\alpha}{\alpha^*}\right)\dfrac{1}{\sigma}K_{EP}(p_1)\exp\left(-\dfrac{1}{p_1}\left|\dfrac{x-\mu}{2\alpha^*\sigma}\right|^{p_1}\right) & \text{if } x < \mu \\ \left(\dfrac{1-\alpha}{1-\alpha^*}\right)\dfrac{1}{\sigma}K_{EP}(p_2)\exp\left(-\dfrac{1}{p_2}\left|\dfrac{x-\mu}{2(1-\alpha^*)\sigma}\right|^{p_2}\right) & \text{if } x < \mu \end{cases} \tag{19}$$

其中 $K_{EP}(p)$ 與 EPD 分配定義相同, $\alpha^* = \dfrac{\alpha K_{EP}(p_1)}{\alpha K_{EP}(p_1) + (1-\alpha)K_{EP}(p_2)}$

若 AEPD 參數 $p_1 = p_2 = p$ 時, AEPD 退化爲 skewed exponential power distribution (SEPD), [5] 機率密度函數表示如下

$$f(x) = \begin{cases} \dfrac{1}{\sigma}K_{EP}(p)\exp\left(-\dfrac{1}{p}\left|\dfrac{x-\mu}{2\alpha\sigma}\right|^{p}\right) & \text{if } x < \mu \\ \dfrac{1}{\sigma}K_{EP}(p)\exp\left(-\dfrac{1}{p}\left|\dfrac{x-\mu}{2(1-\alpha)\sigma}\right|^{p}\right) & \text{if } x \geq \mu \end{cases} \tag{20}$$

若 SEPD 參數 $\alpha < \frac{1}{2}$ 時, 爲右偏分配; 若 $\alpha > \frac{1}{2}$ 時, 爲左偏分配。

標準 AEPD 分配 ($\mu = 0$, $\sigma = 1$) 期望值與變異數爲

$$E(x) = \frac{1}{B}\left[(1-\alpha)^2\frac{p_2\Gamma(2/p_2)}{\Gamma^2(1/p_2)} - \alpha^2\frac{p_1\Gamma(2/p_1)}{\Gamma^2(1/p_1)}\right]$$

$$\begin{aligned} Var(x) =& \frac{1}{B^2}\left\{(1-\alpha)^3\frac{p_2^2\Gamma(3/p_2)}{\Gamma^3(1/p_2)} - \alpha^3\frac{p_1^2\Gamma(3/p_1)}{\Gamma^3(1/p_1)}\right. \\ & \left. - \left[(1-\alpha)^2\frac{p_2\Gamma(2/p_2)}{\Gamma^2(1/p_2)} - \alpha^2\frac{p_1\Gamma(2/p_1)}{\Gamma^2(1/p_1)}\right]^2\right\} \end{aligned}$$

其中 $B = \alpha K_{EP}(p_1) + (1-\alpha)K_{EP}(p_2)$

[5]文獻上 SEPD 有許多種函數型態, 但透過函數轉換仍可得到一致的結果, 本文採用 Zhu and Zinde - Walsh (2009) 的型式。

雖然標準 AEPD 分配不保證期望值恆等於 0 , 變異數恆等於 1 , 若直接以標準 AEPD 分配配適, 並不滿足白噪音的要求, 但其期望值與變異數爲參數 α, p_1, p_2 的函數。因此參數估計時可透過標準化的動作, 使白噪音仍滿足期望值爲 0 , 變異數爲 1 。

以最大概似法參數估計, 對數概似函數爲

$$l(\theta; r) = \sum_{t=1}^{T} \left\{ \ln \delta - \ln h_t + \log f_y \left(\omega + \delta \frac{r_t - \mu_t}{h_t} \middle| \beta \right) \right\}$$

其中 $\beta = (\alpha, p_1, p_2)'$, $f_y(\cdot|\beta)$ 爲 APED 分配機率密度函數, $\omega = \omega(\beta)$ 爲 AEPD 分配的期望值, $\delta = \delta(\beta)$ 爲 AEPD 分配的變異數。

因此白噪音假設爲 EPD 分配一族, 風險值估計爲

$$VaR = \hat{\mu}_t + \hat{\sigma}_t H^{-1}(\alpha) \tag{21}$$

其中 H^{-1} 爲標準 AEPD 分配的分量函數

$$H^{-1}(u; \alpha, p_1, p_2) = \begin{cases} \frac{1}{b}\left((1-\lambda)\sqrt{\frac{\nu-2}{\nu}} G^{-1}\left(\frac{u}{1-\lambda}\right) - a \right) & \text{if } 0 < u < \frac{1-\lambda}{2} \\ \frac{1}{b}\left((1+\lambda)\sqrt{\frac{\nu-2}{\nu}} G^{-1}\left(\frac{u+\lambda}{1+\lambda}\right) - a \right) & \text{if } \frac{1-\lambda}{2} \leq u < 1 \end{cases}$$

G^{-1} 爲 Gamma 分配的分量函數

3.3 模型配適

本節以對數概似值 (log-likelihood value , L) 、$AICC$ 及 BIC , 比較白噪音在不同分配假設下模型整體配適程度。

$$AICC = -2L + \frac{2T(k+1)}{T-k-2}$$

$$BIC = -2L + \frac{k\ln(T)}{T}$$

另外針對白噪音分配進行檢定, 理論上若白噪音爲某一分配, 估計出的標準化殘差應與該分配一致。在此, 我們採用 Kolmogorov-Smirnov (KS) test 與 QQ-plot 檢驗標準化殘差與理論分配是否一致。

Kolmogorov-Smirnov test

KS test 爲無母數方法, 透過比較樣本資料與理論分配機率密度函數的差距, 來檢定樣本資料是否服從該分配。在此, 我們利用 KS test 檢定標準化殘差與白噪音設定的分配是否一致。實證分配 $F_n(x)$ 定義爲

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{X_i \le x}$$

在給定特定分配的累積機率密度函數 $F(x)$ 下, KS test 檢定統計量爲

$$D_n = \sup_x |F_n(x) - F(x)|$$

在虛無假設爲眞情況下

$$\sqrt{n}D_n \to \sup_t |B(F(t))| \tag{22}$$

其中 $B(t)$ 爲 Brownian bridge

由於 KS test 檢定統計量不是標準的分配, 雖然絕大多數的統計軟體皆內建常用的機率分配, 但特殊的機率分配 (如本文使用的 skew-t 分配、AEPD 分配) 仍無法直接得到其 p-value , 因此本文採用兩樣本的 KS test 。

所謂兩樣本的 KS test 則是透過比較兩樣本資料機率密度函數的差距, 來檢定此兩樣本資料是否取自相同母體。其檢定統計量爲

$$D_{n_1,n_2} = \sup_x |F_{n_1}(x) - F_{n_2}(x)|$$

在虛無假設爲眞情況下

$$\sqrt{\frac{n_1 n_2}{n_1 + n_2}} D_n \to \sup_t |B(F(t))| \tag{23}$$

以下列步驟實行兩樣本 KS 檢定

步驟 1、產生 T 個 0 到 1 之間的等差數列, $U = \left\{ \frac{i}{T} \right\}_{i=1}^{T}$, $U \in [0,1]$

步驟 2、將 u_i 代入欲檢定分配的分量函數 $x_i = F^{-1}(u_i \,|\, \hat{\theta})$

步驟 3、使用兩樣本的 KS test , 在虛無假設爲眞情況下, 當 $T \to \infty$ 時, 第 (23) 式會收斂到第 (22) 式

QQ-plot

KS test 雖然可以幫助我們檢定樣本資料是否服從某一特定分配, 但風險值著重在分配的尾端, KS test 卻無法提供這方面的資訊。因此我們再透過 QQ-plot , 觀察標準化殘差與設定的分配在尾端的差異。

首先將標準化殘差排序, 得到排序後的樣本資料, 再與欲比較分配對應的分量繪製成 QQ-plot 。QQ-plot 中直線代表的該特定分配的理論分量, 若樣本分量與理論分量重合, 表示樣本與認定的分配一致。而在 QQ-plot 的左半部, 若樣本分量低於理論分量, 表示樣本分配左尾較該特定分配厚尾; 若樣本分量高於理論分量, 表示樣本分配左尾較該特定分配細尾。QQ-plot 的右半部則呈現相反的狀況, 若樣本分量高於理論分量, 則樣本分量右尾較理論分量厚尾; 若樣本分量低於理論分量, 表示樣本分配右尾較該特定分配細尾。

3.4 回溯測試 (Back-testing)

回溯測試爲利用眞實樣本資料與預測的風險值做比較, 檢定模型預測風險值的可靠性。本文透過一步預測法, 以樣本外預測的結果來檢定白噪音在不同分配假設下, 眞實資料穿透出估計風險值的比例與預期是否相等。若穿透率與預期穿透比例相近, 則認爲估計的風險值具有可信的預測能力。

Unconditional Converage Test

Kupiec (1995) 提出以概似比檢定估計出的風險值是否正確, 假設過去 T 期內樣本資料穿透風險值的次數爲 $n = \sum_{t=1}^{T} I_t$, I_t 爲指標函數, 若 $r_t < VaR_t$ 時, $I_t = 1$; 若 $r_t \geq VaR_t$ 時,$I_t = 0$ 。因此樣本穿透次數可視爲一柏努力試驗, 服從二項分配

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

虛無假設

$$H_0 : \hat{p} = p \quad \text{v.s.} \quad H_1 : \hat{p} \neq p$$

檢定統計量爲

$$LR_{UC} = 2\left[\ln\left(\left(\frac{N}{T}\right)^N \left(1-\frac{N}{T}\right)^{T-N}\right) - \ln\left(p^N (1-p)^{T-N}\right)\right] \tag{24}$$

在虛無假設爲眞條件下, $LR_{UC} \longrightarrow \chi^2(1)$

Conditional Converage Test

由於 Unconditional Converage Test 檢定僅針對穿透率是否與預期相等, 未考慮到兩相臨時間點穿透率是否獨立, Chridtofersen (1998) 提出聯合檢定

1、穿透率與預期相等 $\hat{p} = p$

2、穿透率各期之間爲獨立 $\pi = \pi_{01} = \pi_{11}$

檢定統計量爲

$$LR_{CC} = -2\left[\ln\left(p^N (1-p)^{T-N}\right) + \ln((1-\pi_{01}^{n_{00}})\,\pi_{01}^{n_{01}}\,(1-\pi_{11}^{n_{10}})\,\pi_{11}^{n_{11}})\right] \quad (25)$$

其中 n_{ij} 爲前期狀態爲 i 時, 當期狀態爲 j 的樣本個數;

π_{ij} 爲前期狀態爲 i 時, 當期狀態爲 j 的比例

在虛無假設爲眞條件下, $LR_{CC} \to \chi^2(2)$

圖 5 爲分別以靜態預測 (Variance-Covariance 模型) 與動態預測 (GARCH 模型) 估計在 DGP 爲 GARCH (1,1) 左尾 0.05 的風險值。[6] 靜態預測所估計的風險值爲平坦的直線, 對波動的適應性較差; 動態預測所估計的風險值則隨時間改變, 雖然兩者 Unconditional Converage Test 皆不拒絕虛無假設, 但靜態預測在 Conditional Converage Test 拒絕虛無假設, 表示各期之間穿透率不是獨立, 而且穿透的點大部分皆集中在高波動的時期。

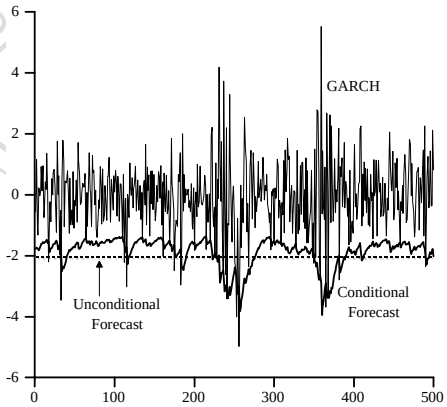

圖 5: Unconditional Forecast 與 Conditional Forecast

[6]圖 5 摘錄至 Lopez (1999) , DGP 爲 GARCH (1,1) , 參數 $h_{t+1} = 0.075 + 0.10\varepsilon_t^2 + 0.85h_t$

3.5 損失函數 (Loss Function)

回溯測試雖然可以檢定樣本資料穿透以不同模型估計風險值的機率是否與預期相符, 但在風險管理的角度, 我們能希望可以找到在通過回溯測試的模型中, 使損失最小的模型。目前仍沒有公認最佳的方法可以判斷各種模型下估計出的風險值何者爲最佳。因此我們採用 Lopez (1999) 提出的損失函數, 一般化的型式表示爲

$$C_{t+1} = \begin{cases} f(r_{t+1}, VaR_{t+1}) & \text{if } r_{t+1} < VaR_{t+1} \\ g(r_{t+1}, VaR_{t+1}) & \text{if } r_{t+1} \geq VaR_{t+1} \end{cases}$$

其中 $f(x,y)$ 與 $g(x,y)$ 可以爲任意函數形式, 只要滿足 $f(x,y) > g(x,y)$ 即可。Lopez (1999) 提出二次方形式的損失函數

$$C_{t+1} = \begin{cases} 1 + (r_{t+1} - VaR_{t+1})^2 & \text{if } r_{t+1} < VaR_{t+1} \\ 0 & \text{if } r_{t+1} \geq VaR_{t+1} \end{cases} \tag{26}$$

Blanco and Ihle (1999) 則建議採用比例形式的損失函數

$$C_{t+1} = \begin{cases} \dfrac{r_{t+1} - VaR_{t+1}}{VaR_{t+1}} & \text{if } r_{t+1} < VaR_{t+1} \\ 0 & \text{if } r_{t+1} \geq VaR_{t+1} \end{cases} \tag{27}$$

損失函數的概念是當樣本資料穿透風險值, 則給予一個正的分數, 損失函數的和愈大, 表示導致的損失愈大。觀察損失函數設定的形式, 損失函數的缺點 (同時也可視爲優點) 在於以不同方法估計出的風險值, 會傾向給予較爲保守的方法較低的分數, 但這也較符合風險管理的角度。

我們先透過回溯測試挑選出可接受的模型, 再透過損失函數來決定最合適的模型。