บทที่ 4

ทฤษฎีและแนวคิดที่ใช้ในการศึกษา

ในการศึกษาถึงความสัมพันธ์ระหว่างราคาปิดของหลักทรัพย์กับช่วงเวลาต่างๆในอดีต ความผันผวนของข้อมูลและการนำข้อมูลไปพยากรณ์ราคาที่คาดว่าจะเกิดขึ้นในอนาคต โดยใช้แนวคิดและทฤษฎีเกี่ยวกับการวิเคราะห์อนุกรมเวลาการทดสอบความเป็น Stationary ของข้อมูลและการทดสอบ Unit Root, แบบจำลอง Autoregressive Conditional Heteroskedasticity (ARCH), แบบจำลอง Generalized Autoregressive Conditional Heteroskedasticity (GARCH), แบบจำลอง ARCH-in-mean (ARCH-M), แบบจำลอง GARCH-in-mean (GARCH-M) และการตรวจสอบรูปแบบ ดังต่อไปนี้

4.1 การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

การวิเคราะห์อนุกรมเวลาเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านไป ลักษณะของการเปลี่ยนแปลงอาจมีหรือไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใดและสามารถพยากรณ์การเปลี่ยนแปลงข้อมูลในอนาคตได้ การวิเคราะห์ข้อมูลอนุกรมเวลาจะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน (ศิริลักษณ์ เล็กสมบูรณ์, 2531)

4.2 การทดสอบความนิ่งของข้อมูล (Stationary) และการทดสอบ Unit root

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง (Stationary) คือข้อมูลที่ค่าเฉลี่ยและความแปรปรวนของกระบวนการเชิงสุ่มนั้นมีค่าคงที่เมื่อเวลาได้เปลี่ยนไปและค่าความแปรปรวนระหว่างสองคาบเวลา

ขึ้นอยู่กับความล้าหลัง (Lag) ระหว่างคาบเวลาทั้งสอง (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542) โดยสามารถเขียนเป็นสมการทางคณิตศาสตร์ตามสมการที่ (4.1)-(4.3)ได้ดังนี้

ค่าเฉลี่ย (Mean) : $E(x_t) = \text{constant} = \mu$ (4.1)

ความแปรปรวน (Variance) : $V(x_t) = \text{constant} = \sigma^2$ (4.2)

ความแปรปรวนร่วม (Covariance) : $\text{cov}(x_t, x_{t+k}) = E(x_t - \mu)(x_{t+k} - \mu) = \sigma_k - \mu$ (4.3)

โดยที่ x_t แทนข้อมูลอนุกรมเวลาซึ่งเป็นกระบวนการเชิงสุ่ม

ในการวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลานั้น ข้อมูลจะต้องมีลักษณะนิ่งเนื่องด้วยข้อมูลอนุกรมเวลานั้นมาจากกระบวนการเชิงสุ่ม (Random Process) การนำข้อมูลอนุกรมเวลาไปใช้โดยไม่ได้ทำการตรวจสอบว่าข้อมูลนั้นมีลักษณะนิ่งนั้น ถ้าข้อมูลอนุกรมเวลานั้นไม่นิ่งจะทำให้ค่าสถิติที่เกิดขึ้นมีการแจกแจงไม่มาตรฐาน (Nonstandard Distributions) ซึ่งทำให้การนำไปใช้เปรียบเทียบกับค่าในตารางมาตรฐาน (Standard Tables) ไม่ถูกต้องเนื่องจากค่าต่างๆนั้นมีสมมุติฐานว่าข้อมูลนั้นมีการแจกแจงมาตรฐาน (Standard Distributions) ทำให้นำไปสู่การลงความเห็นที่ผิดพลาดและความสัมพันธ์ที่ไม่แท้จริง (Spurious Regression) กล่าวคือค่า R^2 มีค่าสูงมากและได้ค่าสถิติ t มีนัยสำคัญหรือสูงเกินกว่าความเป็นจริง (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542)

ในการใช้ข้อมูลอนุกรมเวลาจึงต้องทำการทดสอบว่าข้อมูลที่นำมาใช้มีลักษณะนิ่งหรือไม่ ซึ่งจะใช้การทดสอบ Unit Root โดยในการศึกษานี้จะพิจารณาเฉพาะวิธีของ Dickey-Fuller โดยวิธี DF (Dickey–Fuller Test) และ ADF (Augmented Dickey-Fuller Test) ซึ่งกำหนดโดยสมการที่ (4.4)

$$x_t = \rho x_{t-1} + \varepsilon_t \tag{4.4}$$

โดยกำหนดสมมติฐานหลัก $H_0 : \rho = 1$

และ $H_1 : |\rho| < 1$

ถ้ายอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่ง แต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง และการทดสอบนี้ยังสามารถแปลงสมการได้ดังนี้ คือ

กรณีไม่มีค่าคงที่และแนวโน้มเวลา $\Delta x_t = \theta x_{t-1} + \varepsilon_t$ (4.5)

กรณีมีค่าคงที่ $\Delta x_t = \alpha + \theta x_{t-1} + \varepsilon_t$ (4.6)

กรณีมีทั้งค่าคงที่และแนวโน้มเวลา $\Delta x_t = \alpha + \beta t + \theta x_{t-1} + \varepsilon_t$ (4.7)

โดยกำหนดสมมติฐานหลัก $H_0 : \theta = 0$

และสมมติฐานรอง $H_1 : \theta < 0$

การยอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่ง แต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง นอกจากนี้ถ้าสมการที่ (4.5) (4.6) และ (4.7) นำไปเข้ากระบวนการอัตถดถอย (Autoregressive Processes) จะได้ดังนี้

กรณีไม่มีทั้งค่าคงที่และแนวโน้มเวลา $$\Delta x_t = \theta x_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta x_{t-i} + \varepsilon_t \qquad (4.8)$$

กรณีมีเฉพาะค่าคงที่ $$\Delta x_t = \alpha + \theta x_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta x_{t-i} + \varepsilon_t \qquad (4.9)$$

กรณีมีทั้งค่าคงที่และแนวโน้มเวลา $$\Delta x_t = \alpha + \beta t + \theta x_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta x_{t-i} + \varepsilon_t \qquad (4.10)$$

ซึ่งสมการที่ (4.8) (4.9) และ (4.10) เป็นการทดสอบ Augmented Dickey-Fuller Test (ADF) นั่นเอง ซึ่งพัฒนามาจากวิธี Dickey-Fuller Test เพื่อแก้ปัญหา Serial Correlation ในการตรวจสอบว่าข้อมูลมีลักษณะนิ่งหรือไม่โดยการเปรียบเทียบค่าสถิติ t ที่คำนวณได้กับค่าวิกฤติ (Critical Value) ในตาราง ADF

4.3 แบบจำลอง Autoregressive Conditional Heteroskedasticity (ARCH)

ในการวิเคราะห์อนุกรมเวลาส่วนใหญ่แล้วจะมีการกำหนด Stochastic Variable ให้มีความแปรปรวนคงที่ (Homoskedastic) ซึ่งในการประยุกต์ใช้กับบางข้อมูลนั้นค่าความแปรปรวนของค่าความคลาดเคลื่อน (Error Term) จะไม่ใช่ฟังก์ชั่นของตัวแปรอิสระแต่มีค่าเปลี่ยนแปลงไปตามช่วงเวลาขึ้นอยู่กับขนาดของความคลาดเคลื่อนที่เกิดขึ้นในอดีต และในบางงานศึกษา เช่น แบบจำลองของเงินเฟ้อ อัตราดอกเบี้ยหรือผลตอบแทนจากตลาดหลัก ในบางคาบเวลาจะมีค่าความผันผวน (Volatility) สูง (และความคลาดเคลื่อนขนาดใหญ่) ตามด้วยคาบเวลาที่มีค่าความผันผวน (Volatility) ต่ำ (และมีค่าความคลาดเคลื่อนที่มีขนาดเล็ก) สรุปได้ว่าค่าความแปรปรวนของค่าความคลาดเคลื่อนจากการถดถอยจะขึ้นอยู่กับค่าความผันผวน (Volatility) ของความคลาดเคลื่อนในอดีตที่ผ่านมา (Enders, 1995)

ความเป็นไปได้ในการหาค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาไปพร้อมกันนั้น ในขั้นต้นการพยากรณ์อย่างมีเงื่อนไขจะมีความแม่นยำเหนือกว่าการพยากรณ์อย่างไม่มีเงื่อนไขมาก ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) ซึ่งแสดงได้ดังสมการ (4.11)

$$x_t = a_0 + a_1 x_{t-1} + \epsilon_t \tag{4.11}$$

และต้องการพยากรณ์ x_{t+1} การพยากรณ์อย่างมีเงื่อนไขของ x_{t+1} ดังนี้ คือ

$$E_t x_{t+1} = a_0 + a_1 x_t \tag{4.12}$$

ถ้าเราใช้ค่าเฉลี่ยอย่างมีเงื่อนไขในการพยากรณ์ x_{t+1} ค่าความคลาดเคลื่อนของความแปรปรวนอย่างมีเงื่อนไขที่พยากรณ์ได้ดังสมการ (4.13)

$$E_t[(x_{t+1} - a_0 - a_1 x_t)^2] = E_t \epsilon_{t+1}^2 = \sigma^2 \tag{4.13}$$

ถ้าเปลี่ยนไปใช้การพยากรณ์อย่างไม่มีเงื่อนไขแล้ว ผลที่จะใช้เป็นค่าเฉลี่ยในช่วง Long-Run ของลำดับ $\{x_t\}$ ซึ่งเท่ากับ $\frac{a_0}{(1-a_1)}$ จะได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไขตามสมการ (4.14) คือ

$$\begin{aligned} E\{[x_{t+1} - \frac{a_0}{(1-a_1)}]^2\} &= E[(\epsilon_{t+1} + a_1 \epsilon_t + a_1^2 \epsilon_{t-1} + a_1^3 \epsilon_{t-2} + \ldots)^2] \\ &= \frac{\sigma^2}{(1-a_1^2)} \end{aligned} \tag{4.14}$$

เมื่อ $\frac{1}{(1-a_1^2)} > 1$ ค่าความแปรปรวนที่ได้จากการพยากรณ์อย่างไม่มีเงื่อนไขจะสูงกว่าแบบมีเงื่อนไข ดังนั้นในการพยากรณ์อย่างมีเงื่อนไขจึงมีความเหมาะสมกว่า ในลักษณะเดียวกันถ้าความแปรปรวนของ $\{\epsilon_t\}$ ไม่เป็นค่าคงที่ จะสามารถประมาณค่าแนวโน้มของการเปลี่ยนแปลงความแปรปรวนโดยใช้ ARMA model อธิบายได้โดยให้ $\{\hat{\epsilon}_t\}$ แทนส่วนที่เหลือ (Residuals) ที่ได้จากการประมาณจากสมการ (4.11) ดังนั้นค่าความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) ของ x_{t+1} จะได้ดังสมการ (4.15)

$$\begin{aligned} Var(x_{t+1} \mid x_t) &= E_t[(x_{t+1} - a_0 - a_1 x_t)^2] \\ &= E_t \epsilon_{t+1}^2 \end{aligned} \tag{4.15}$$

และจากที่ให้ $E_t \in_{t+1}^2$ เท่ากับ σ_{t+1}^2 จึงแสดงว่าค่าความแปรปรวนอย่างมีเงื่อนไขไม่ใช่ค่าคงที่และจะได้แบบจำลองในการประมาณค่าส่วนที่เหลือ (Residuals) ออกมาดังสมการ (4.16)

$$\hat{\in}_t^2 = \alpha_0 + \alpha_1 \hat{\in}_{t-1}^2 + \ldots + \alpha_q \hat{\in}_{t-q}^2 + v_t \tag{4.16}$$

เมื่อ v_t = white noise process

ถ้าค่าของ $\alpha_1, \alpha_2, \ldots, \alpha_q$ เท่ากับศูนย์ ค่าความแปรปรวนจากการประมาณจะเท่ากับค่าคงที่ α_0 อีกนัยหนึ่ง คือ ค่าความแปรปรวนอย่างมีเงื่อนไขของ x_t จะมีการเปลี่ยนแปลงสอดคล้องกับ Autoregression ในสมการ (4.16) ดังนั้นจะสามารถใช้สมการ (4.16) ในการพยากรณ์ค่าความแปรปรวนอย่างมีเงื่อนไขที่เวลา $t+1$ ดังสมการ (4.17)

$$E_t \hat{\in}_{t+1}^2 = \alpha_0 + \alpha_1 \hat{\in}_t^2 + \alpha_2 \hat{\in}_{t-1}^2 + \ldots + \alpha_q \hat{\in}_{t+1-q}^2 \tag{4.17}$$

จากเหตุผลที่กล่าวมา สมการที่ (4.16) เรียกว่า Autoregressive Conditional Heteroskedastic (ARCH) model และสมการ (4.17) เป็น ARCH(q) สมการ (4.17) ค่า $E_t \hat{\in}_{t+1}^2$ หรือ σ_{t+1}^2 จะประกอบด้วย 2 องค์ประกอบ คือ ค่าคงที่และความผันผวน (Volatility) ในคาบเวลาที่ผ่านมา ซึ่งเขียนได้เป็นส่วนเหลือกำลังสองของคาบในอดีต (ARCH term) ส่วนค่าสัมประสิทธิ์ $(\alpha_0, \alpha_1, \ldots, \alpha_q)$ สามารถหาค่าได้โดยใช้วิธี Maximum Likelihood

4.4 แบบจำลอง Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

Bollerslev (1986) ได้ขยายมาจาก ARCH model โดยมีขั้นตอน คือ ให้ค่าความคลาดเคลื่อนจากกระบวนการเป็นดังสมการ (4.18)

$$\in_t = v_t \sqrt{h_t} \tag{4.18}$$

เมื่อ $\sigma_v^2 = 1$

และ

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \in_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{4.19}$$

เมื่อ $\{v_t\}$ คือ white noise process ที่เป็นค่าอิสระจากเหตุการณ์ในอดีต ($\in_{t-i}$) ค่าเฉลี่ยอย่างมีเงื่อนไขและไม่มีเงื่อนไขของ $\in_t$ จะมาจาก h_t ในสมการ (4.19)

GARCH(p,q) นั้นใช้กระบวนการ Autoregressive และ Moving Average ในการหา Heteroskedastic Variance ได้ดังสมการ (4.20)

$$E_{t-1} \in_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \in_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{4.20}$$

ถ้ากำหนดให้ค่า $p = 0$ และ $q = 1$ จะได้เป็น ARCH(1) หรือถ้าค่า β_i ทั้งหมดมีค่าเป็น 0 แบบจำลอง GARCH(p,q) จะเทียบเท่ากับแบบจำลอง ARCH (q) คุณสมบัติที่สำคัญของแบบจำลอง GARCH คือค่าความแปรปรวนอย่างมีเงื่อนไขของ disturbances ของค่า x_t สร้างขึ้นมาจากกระบวนการ ARMA จึงสามารถคาดได้ว่าส่วนเหลือจากการทำ ARMA จะแสดงถึงรูปแบบคุณลักษณะเดียวกัน เช่น ถ้าการประมาณค่า $\left\{x_t\right\}$ ด้วยกระบวนการ ARMA ค่าสหสัมพันธ์ในตัวเอง (Autocorrelation Function หรือ ACF) ซึ่งเป็นสหสัมพันธ์ระหว่างตัวแปรสุ่มที่หน่วยเวลาห่างกันของกระบวนการเดียวกันและสหสัมพันธ์ในตัวเองส่วนย่อย (Partial Autocorrelation Function หรือ PACF) (รายละเอียดดังภาคผนวก ง) ของส่วนเหลือควรจะบ่งถึงกระบวนการ white noise และ ACF ของกำลังสองของส่วนเหลือนำมาช่วยในการระบุถึงลำดับของกระบวนการ GARCH

4.5 แนวคิดเกี่ยวกับแบบจำลอง ARCH – in – mean (ARCH-M)

พื้นฐานจากแนวคิด ARCH ได้มีการขยายเพื่อให้ค่าเฉลี่ยของลำดับอนุกรมขึ้นอยู่กับความแปรปรวนอย่างมีเงื่อนไขของตัวเอง ซึ่งเรียกว่า ARCH-M เพื่อที่จะมีความเหมาะสมในการศึกษา โดยมีหลักการพื้นฐาน คือ ผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยง (Risk Averter) จะต้องการชดเชยสำหรับการถือสินทรัพย์ที่มีความเสี่ยง (Risk Premium) เมื่อให้ความเสี่ยงในทรัพย์สินสามารถวัดค่าได้จากความแปรปรวนของผลตอบแทน ค่าชดเชยความเสี่ยงจะเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของผลตอบแทน Engle et al. (1987) ที่ได้แสดงแนวคิดนี้โดยนำส่วนที่เหลือของผลตอบแทนจากการถือครองทรัพย์สินที่มีความเสี่ยง ดังสมการ (4.21)

$$x_t = \mu_t + \in_t \qquad (4.21)$$

เมื่อ $x_t =$ ผลตอบแทนส่วนเกิน จากการถือครองทรัพย์สินในระยะยาวเปรียบเทียบกับหนึ่งคาบเวลาของผลตอบแทนอ้างอิง

$\mu_t =$ ค่าชดเชยความเสี่ยงที่จำเป็นในการโน้มน้าวผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยงให้ถือครองทรัพย์สินในระยะยาวแทนที่จะเป็นแค่คาบเวลาเดียว

$\in_t =$ Unforecastable Shock ของส่วนที่เหลือของผลตอบแทนในการถือทรัพย์สินในระยะยาว

ในการอธิบายสมการ (4.21) จะให้ค่าคาดหวังของส่วนที่เหลือของผลตอบแทนจากการถือครองทรัพย์สินในระยะยาวมีค่าเท่ากับค่าชดเชยความเสี่ยงดังสมการ (4.22)

$$E_{t-1} x_t = \mu_t \qquad (4.22)$$

สมมติว่า ค่าชดเชยความเสี่ยงเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของ $\in_t$ อีกนัยหนึ่ง คือ ยิ่งค่าความแปรปรวนอย่างมีเงื่อนไขของผลตอบแทนมีขนาดใหญ่ ผลตอบแทนของการโน้มน้าวให้คนหันมาถือครองทรัพย์สินในระยะยาวก็จะยิ่งมีค่ามากขึ้นตาม ถ้า h_t เป็นความแปรปรวนอย่างมีเงื่อนไขของ $\in_t$ ค่าชดเชยความเสี่ยงสามารถแสดงได้เป็น

$$\mu_t = \beta + \delta h_t \, , \; \delta > 0 \qquad (4.23)$$

เมื่อ h_t คือ ARCH(q) process

$$h_t = \alpha_0 + \sum_{i=0}^{q} \alpha_i \in_{t-i}^2 \qquad (4.24)$$

จากสมการ (4.21) (4.23) และ (4.24) ประกอบขึ้นเป็นพื้นฐานของแบบจำลอง ARCH-M จากสมการ (4.21) และ (4.23) ค่าเฉลี่ยอย่างมีเงื่อนไขของ x_t จะขึ้นอยู่กับค่าความแปรปรวนอย่างมีเงื่อนไข h_t จากสมการ (4.24) ความแปรปรวนอย่างมีเงื่อนไขได้จากกระบวนการ ARCH(q) ซึ่งเป็นการอธิบายว่า ถ้าความแปรปรวนอย่างมีเงื่อนไขเป็นค่าคงที่ (เช่น ถ้า $\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$) แบบจำลอง ARCH-M จะกลับไปสู่กรณีเดิมของค่าชดเชยความเสี่ยงคงที่

4.6 แบบจำลอง GARCH – in – mean (GARCH-M)

จากแบบจำลอง GARCH(p,q) ในสมการ (4.19) ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนเหลือกำลังสองของกระบวนการนี้ (Bollerslev, 1986) Engel(1987) ขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักกันในชื่อ GARCH-in-mean หรือ GARCH-M ผลตอบแทนจากหลักทรัพย์สามารถแสดงในรูปของแบบจำลอง GARCH(p,q)-M ดังสมการ (4.25) ถึง (4.27)

$$x_t = \mu_t + \delta_1 h_t^{1/2} + \varepsilon_t \tag{4.25}$$

$$\varepsilon_t / \psi_{t-1} \sim N(0, h_t) \tag{4.26}$$

$$h_t = \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i})^2 \tag{4.27}$$

เมื่อ x_t คือ ผลตอบแทนจากหลักทรัพย์

μ_t คือ ค่าเฉลี่ย x_t อย่างมีเงื่อนไขต่อข้อมูลในอดีต(ψ_{t-1})และตามสมการข้อจำกัด $\omega > 0, \alpha_i > 0$ และ $\beta_i \geq 0$ เพื่อให้แน่ใจว่าค่าความแปรปรวนอย่างมีเงื่อนไข (h_t) นั้นเป็นบวก

$h_t^{1/2}$ ในสมการที่ (4.25) เพื่อใช้แสดงความสัมพันธ์โดยตรงถึง trade off ระหว่างความเสี่ยงและผลตอบแทนที่คาดหวัง

อิทธิพลอย่างมีนัยสำคัญของความผันผวนในผลตอบแทนจากหลักทรัพย์ถูกตรวจจับด้วยสัมประสิทธิ์ $h_t^{1/2}$ (δ_1) ในสมการ (4.25) ซึ่งอธิบายแทนดัชนีของความสัมพันธ์ของการหลีกเลี่ยงความเสี่ยง ค่าสัมประสิทธิ์ δ_1 ที่เป็นบวกอย่างมีนัยสำคัญบอกถึงผู้ลงทุนในหลักทรัพย์เมื่อเผชิญกับระดับความเสี่ยงที่สูงขึ้นก็ต้องการค่าชดเชยความเสี่ยงที่มากขึ้นตามไปด้วย

4.7 การตรวจสอบรูปแบบ (Diagnostic Checking)

การสร้างสมการพร้อมทั้งประมาณค่าพารามิเตอร์แล้วนั้น จะต้องทำการตรวจสอบรูปแบบว่าสมการพยากรณ์ที่ได้มามีความเหมาะสมหรือไม่และรูปแบบใดของสมการที่ดีที่สุด โดยใช้การทดสอบต่างๆ ดังนี้

4.7.1 การทดสอบ Ljung-Box Q-Statistics

การทดสอบ Ljung-Box Q-Statistics เป็นการทดสอบว่าสหสัมพันธ์ในตัวเองในส่วนเหลือทุกช่วงเวลาที่ห่างกัน k มีความเป็นอิสระกันหรือไม่
โดยมีสมมติฐานดังนี้

$$H_0 : \rho_1(\hat{a}_t) = \rho_2(\hat{a}_t) = ... = \rho_k(\hat{a}_t) = 0$$

และ

$$H_1 : \rho_1(\hat{a}_t) \neq \rho_2(\hat{a}_t) \neq ... \neq \rho_k(\hat{a}_t) \neq 0$$

คำนวณตามสมการที่ (4.28) คือ

$$Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{r_j^2}{T-j} \tag{4.28}$$

เมื่อ r_j คือ สหสัมพันธ์ในตัวเองลำดับที่ j โดยที่ $j = 1, ..., k$

T คือ จำนวนของค่าสังเกต

ภายใต้ส่วนเหลือจากการประมาณด้วยแบบจำลอง ARIMA ค่า Q_{LB} มีการแจกแจงแบบไคสแควร์ (χ^2) ด้วยระดับความเป็นอิสระ (Degree of Freedom) เท่ากับจำนวนของสหสัมพันธ์ในตัวเองลบด้วยจำนวนของพารามิเตอร์ Autoregressive (AR) และ Moving Average (MA) ที่ได้มาจากการประมาณหรือ *k-m*

จะยอมรับสมมติฐานหลักเมื่อ $Q_{LB} \leq \chi^2_{\alpha,k-m}$ คือส่วนเหลือเป็นอิสระต่อกันที่ความล่า k และถ้าปฏิเสธสมมติฐานหลักเมื่อ $Q_{LB} > \chi^2_{\alpha,k-m}$ คือ เกิดสหสัมพันธ์ในตัวเองอย่างน้อยหนึ่งค่าในส่วนเหลือที่ไม่เท่ากับศูนย์

4.7.2 เกณฑ์การเลือกรูปแบบของแบบจำลองที่ดีที่สุด (Information Criteria)

ในการหารูปแบบของแบบจำลอง เมื่อได้รูปแบบของแบบจำลองที่เหมาะสมหลายรูปแบบจึงต้องมีแนวทางในการเลือกรูปแบบของแบบจำลองที่ดีที่สุด โดยการพิจารณาค่า Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) รูปแบบที่ให้ค่า AIC และ SC น้อยที่สุดจะเป็นรูปแบบที่ดีที่สุด โดย Akaike Information Criterion (AIC) คำนวณตามสมการที่ (4.29) และ Schwartz Criterion (SC) คำนวณตามสมการที่ (4.30)

Akaike Informaiton Criterion (AIC)	$-2\ell/\eta+2k/\eta$	(4.29)
Schwartz Criterion (SC)	$-2\ell/\eta+k\log\eta/\eta$	(4.30)

โดยที่ k เป็นจำนวนของพารามิเตอร์ที่ทำการประมาณค่า

η เป็นจำนวนของค่าสังเกต

ℓ เป็นค่าของ log likelihood function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว

มหาวิทยาลัยเชียงใหม่ Chiang Mai University