บทที่ 4
ทฤษฎีและแบบจำลองที่ใช้ในการศึกษา

4.1 ทฤษฎีที่เกี่ยวข้อง

4.1.1 การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

การวิเคราะห์ข้อมูลอนุกรมเวลาจะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน กล่าวคือ การวิเคราะห์อนุกรมเวลาเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาในอดีต การที่อนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาในอดีต ทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใด และสามารถพยากรณ์การเปลี่ยนแปลงข้อมูลในอนาคตได้ (ศิริลักษณ์ เล็กสมบูรณ์, 2531)

4.1.2 การทดสอบความนิ่งของข้อมูล (Unit Root Tests)

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง (Stationary) คือข้อมูลที่ค่าเฉลี่ยและความแปรปรวนของกระบวนการเชิงสุ่ม (Random Process) นั้นมีค่าคงที่เมื่อเวลาได้เปลี่ยนไป และค่าความแปรปรวนระหว่างสองคาบเวลาขึ้นอยู่กับความล่า (Lag) ระหว่างคาบเวลาทั้งสองนั้น (ทรงศักดิ์ ศรีบุญจิตต์ และ อารี วิบูลย์พงศ์, 2542) โดยเขียนเป็นสมการได้ดังนี้

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 คือ ข้อมูลอนุกรมเวลาซึ่งเป็นกระบวนการเชิงสุ่ม

ในการวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลานั้น ข้อมูลจะต้องมีลักษณะนิ่ง เนื่องจากข้อมูลอนุกรมเวลานั้นมาจากกระบวนการเชิงสุ่ม การนำข้อมูลอนุกรมเวลาไปใช้โดยไม่ได้ทำการตรวจสอบว่าข้อมูลนั้นมีลักษณะนิ่งนั้น ค่าสถิติที่เกิดขึ้นจะมีการแจกแจงไม่มาตรฐาน (Nonstandard Distribution) ซึ่งทำให้การนำไปใช้เปรียบเทียบกับค่าในตารางมาตรฐานไม่ถูกต้องเนื่องจากค่าต่าง ๆ นั้น มีสมมติฐานว่าข้อมูล

นั้นมีการแจกแจงมาตรฐาน (Standard Distributions) ทำให้เกิดการลงความเห็นที่ผิดพลาดและความสัมพันธ์ที่ไม่แท้จริง (Spurious Regression) กล่าวคือ R^2 มีค่าสูงมาก และได้ค่าสถิติ t – test มีนัยสำคัญหรือสูงเกินกว่าความเป็นจริง (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542)

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

$$x_t = \rho x_{t-1} + \varepsilon_t \qquad (4.4)$$

โดยกำหนดสมมติฐานหลัก $H_0: \rho = 1$

และสมมติฐานรอง $H_1: |\rho| < 1$

ถ้ายอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่ง แต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง และจากสมการ (4.4) สามารถแปลงเป็นสมการได้ดังนี้ คือ

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

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

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

โดยกำหนดสมมติฐานหลัก $H_0: \theta = 0$

และสมมติฐานรอง $H_1: \theta < 0$

การยอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่งแต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง

นอกจากนี้ถ้าสมการที่ (4.5) (4.6) และ (4.7) เข้าสู่ Autoregressive Processes จะได้สมการดังนี้

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

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

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

ซึ่งสมการที่ (4.8) (4.9) และ (4.10) มีจำนวนของ Lagged Difference Terms ที่เพิ่มเข้ามา การที่ Lagged เพิ่มมากขึ้นจะทำให้เกิดค่าความคลาดเคลื่อน (Error Terms) ที่มีลักษณะเป็น Serial Correlation และเมื่อนำการทดสอบ Augmented Dickey-Fuller Test (ADF) ซึ่งพัฒนามาจากวิธี Dickey-Fuller Test (DF) เพื่อแก้ปัญหา Serial Correlation ในการตรวจสอบว่าข้อมูลนั้นมีลักษณะนิ่งหรือไม่โดยการเปรียบเทียบค่าสถิติ t ที่คำนวณได้กับค่าวิกฤติ MacKinnon (MacKinnon Critical Values) (Enders, 1995: 221 และ Gujarati, 1995: 720)

ในการหาจำนวนของ Lag Length ที่มีความเหมาะสมต่อการนำไปทดสอบนั้น (Enders, 1995, 227) ได้เสนอวิธีที่เหมาะสมหลายวิธี เช่นการกำหนดจำนวนของ Lag Length ที่มีจำนวนมากพอ เช่นที่ P* แล้วดูว่า สัมประสิทธิ์ Lag Length นั้นแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติหรือไม่ โดยการทดสอบด้วยค่าสถิติ t (t-test) ถ้าไม่มีนัยสำคัญทางสถิติให้ทำการลด Lag Length ลงทีละ 1 จนกว่าสัมประสิทธิ์ Lag Length นั้นจะแตกต่างจากศูนย์อย่างมีนัยสำคัญ แต่ในการศึกษาครั้งนี้ได้กำหนดจำนวนของ Lag Length ที่ระดับ 0 และ 1

4.1.3 แบบจำลอง Autoregressive Conditional Heteroscedasticity (ARCH)

ในแบบจำลองเศรษฐมิติแบบดั้งเดิมได้มีการสมมติให้ความแปรปรวนของเทอมความคลาดเคลื่อนมีค่าคงที่หรือคงตัว ซึ่ง Enders (1995) ได้แสดงให้เห็นว่าข้อมูลเศรษฐกิจอนุกรมเวลาจำนวนมากในคาบเวลาจำนวนไม่น้อยมีความผันผวนสูงมาก ตามมาด้วยคาบเวลาที่อนุกรมดังกล่าวค่อนข้างจะมีความสงบซึ่งจะเห็นได้ว่าข้อสมมติที่ว่าความแปรปรวนของเทอมความคลาดเคลื่อน มีค่าคงที่หรือค่าคงตัวนั้น ไม่น่าจะเป็นข้อสมมติที่เหมาะสมหรือถูกต้อง ซึ่ง Enders (1995) กล่าวว่า ในหลายสถานการณ์ เราสนใจแต่เพียงความแปรปรวนแบบมีเงื่อนไขเท่านั้น เช่นนักลงทุนในตลาดหุ้นอาจจะสนใจในการพยากรณ์อัตราผลตอบแทน (Rate of Return) และความแปรปรวนของหุ้นที่เราถือเท่านั้น ในขณะที่ความแปรปรวนแบบไม่มีเงื่อนไข (Unconditional Variance คือความแปรปรวนระยะยาวนั่นเอง) อาจจะไม่ใช่สิ่งที่สำคัญ ถ้านักลงทุนวางแผนที่จะซื้อขายหุ้นในช่วงไม่ยาวจนเกินไปนัก (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542)

วิธีหนึ่งที่มักจะใช้ในการพยากรณ์ความแปรปรวน คือแบบจำลองที่แสดงความสัมพันธ์ระหว่าง x_{t+1} กับ ε_{t+1} และ x_t ซึ่งเขียนเป็นสมการได้ดังนี้

$$x_{t+1} = \varepsilon_{t+1}\, x_t \qquad (4.11)$$

โดยที่ x_{t+1} คือ ตัวแปรที่เรากำลังพิจารณา

ε_{t+1} คือ ตัวเทอมรบกวน white noise (White Noise Disturbance Term) ซึ่งมีความแปรปรวนเท่ากับ σ^2 ซึ่งเป็นค่าคงที่หรือคงตัว (Constant)

x_t คือ ตัวแปรอิสระ (Independent Variable) ณ คาบเวลา t ซึ่งเป็นตัวแปรที่เราสังเกตได้

จากสมการ (4.11) ถ้า x_t มีค่าเท่ากับทุกคาบเวลาและเท่ากับค่าคงตัวหรือค่าคงที่ซึ่งสมมติว่าเท่ากับ x จะสามารถเขียนสมการ (4.11) ใหม่ได้ดังนี้

$$x_{t+1} = \varepsilon_{t+1} x \qquad (4.12)$$

เราจะได้ว่า $\{x_{t+1}\}$ sequence ก็จะมีลักษณะเป็น white noise process ด้วยความแปรปรวนคงที่หรือคงตัว อย่างไรก็ตาม $\{x_t\}$ sequence มักจะมีค่าไม่เท่ากัน ดังนั้นความแปรปรวนภายใต้เงื่อนไขของ x_t สามารถเขียนได้ดังนี้

$$\mathrm{Var}(x_{t+1}|x_t) = \sigma^2 x_t^2 \qquad (4.13)$$

และถ้าค่าสืบเนื่อง (Successive Values) ของ $\{x_{t+1}\}$ มี positive serial correlation ความแปรปรวนแบบมีเงื่อนไข ของ $\{x_t\}$ sequence ก็จะมี positive serial correlation ด้วย ในลักษณะเช่นนี้ $\{x_{t+1}\}$ sequence ก็จะทำให้เกิดคาบเวลาของความผันผวนใน $\{x_t\}$ sequence

ในทางปฏิบัติแล้ว เราอาจจะปรับปรุงแบบจำลองที่กล่าวมาแล้วข้างต้นให้อยู่ในรูปแบบดังนี้

$$\ln x_t = a_0 + a_1 \ln(x_{t+1}) + e_t \qquad (4.14)$$

โดยที่ e_t คือ เทอมความคลาดเคลื่อนซึ่งคือ $\ln(\varepsilon_t)$ นั่นเอง

และสามารถทำการถดถอยโดยใช้ OLS (OLS regression) แต่ละจุดอ่อนของวิธีนี้ก็คือเราสมมติไว้แน่นอนว่า $\{x_{t+1}\}$ เป็นสาเหตุของการเปลี่ยนแปลงความแปรปรวน และโดยเหตุผลทางทฤษฎีแล้วเราอาจจะไม่มีเหตุผลที่ดีเพียงพอในการเลือกตัวแปร $\{x_{t+1}\}$ ที่เป็นสาเหตุของการเปลี่ยนแปลงของความแปรปรวนได้ และสิ่งที่เป็นจุดอ่อนที่สำคัญอีกประการหนึ่งของแบบจำลองสมการ (4.14) ก็คือ เราได้สมมติว่าเทอมความคลาดเคลื่อน ซึ่งคือ $\{e_t\}$ sequence มีความแปรปรวนคงที่หรือไม่คงที่ ถ้าข้อสมมติดังกล่าวไม่ถูกต้องก็จะต้องมีการแปลงข้อมูล (Data Transformation) อีก

ในการวิเคราะห์อนุกรมเวลาส่วนใหญ่แล้วจะมีการกำหนด Stochastic Variable ให้มีความแปรปรวนคงที่ (Homoscadastic) ซึ่งในการประยุกต์ใช้กับบางข้อมูลนั้น ค่าความแปรปรวนของค่าเทอมคลาดเคลื่อน จะไม่ใช่ฟังก์ชั่นของตัวแปรอิสระแต่มีค่าเปลี่ยนแปลงไปตามช่วงเวลาขึ้นอยู่กับขนาดของความคลาดเคลื่อนที่เกิดขึ้นในอดีต หรือกล่าวได้ว่าค่าความแปรปรวนของเทอมคลาดเคลื่อนนั้นขึ้นอยู่กับค่าความผันผวน (Volatility) ของความคลาดเคลื่อนในอดีตที่ผ่านมา

ความเป็นไปได้ในการหาค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาไปพร้อมกันนั้นในขั้นต้น การพยากรณ์อย่างมีเงื่อนไขจะมีความแม่นยำเหนือกว่าการพยากรณ์อย่างไม่มีเงื่อนไขมาก ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) แสดงได้ดังนี้

$$x_t = a_0 + a_1 x_{t-1} + \varepsilon_t \tag{4.15}$$

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

$$E_t x_{t+1} = a_0 + a_1 x_t \tag{4.16}$$

และค่าเฉลี่ยแบบมีเงื่อนไขในการพยากรณ์ x_{t+1} ค่าความคลาดเคลื่อนของความแปรปรวนอย่างมีเงื่อนไขที่พยากรณ์ได้ดังนี้

$$E_t[(x_{t+1} - a_0 - a_1 x_t)^2] = E_t \varepsilon_{t+1}^2 = \sigma^2 \tag{4.17}$$

ถ้าเปลี่ยนไปใช้การพยากรณ์แบบไม่มีเงื่อนไขแล้ว ผลที่ใช้จะเป็นค่าเฉลี่ยในช่วงระยะยาว ของลำดับ $\{x_t\}$ ซึ่งเท่ากับ $\frac{a_0}{(1-a_1)}$ จะได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไข ตามสมการ (4.18) คือ

$$E_t\{[(x_{t+1} - \frac{a_0}{(1-a_1)}]^2 = E[(\varepsilon_{t+1} + a_1\varepsilon_t + a_1^2\varepsilon_{t-1} + a_1^3\varepsilon_{t-2} + \ldots)^2]$$

$$= \frac{\sigma^2}{(1-a_1^2)} \tag{4.18}$$

เมื่อ $\frac{1}{(1-a_1^2)}$ > 1 ค่าความแปรปรวนที่ได้จากการพยากรณ์อย่างไม่มีเงื่อนไขจะสูงกว่าแบบมีเงื่อนไข ดังนั้นในการพยากรณ์อย่างไม่มีเงื่อนไขจึงมีความเหมาะสมกว่า ในลักษณะเดียวกันถ้าความแปรปรวนของ $\{\varepsilon_t\}$ ไม่เป็นค่าคงที่ จะสามารถประมาณค่าแนวโน้มของการเปลี่ยนแปลงความแปรปรวนโดยใช้ ARMA Model อธิบายได้โดยให้ $\{\hat{\varepsilon}_t\}$ แทนส่วนที่เหลือ (Residual) ที่ได้จากการประมาณจากสมการ (4.15) ดังนั้นค่าความแปรปรวนอย่างมีเงื่อนไข ของ x_{t+1} จะได้ดังสมการ (4.19)

$$\begin{aligned} Var(x_{t+1}|x_t) &= E_t[(x_{t+1} - a_0 - a_1x_t)^2] \\ &= E_t\varepsilon_{t+1}^2 \end{aligned} \qquad (4.19)$$

และจากที่ให้ $E_t\varepsilon_{t+1}^2$ เท่ากับ σ_{t+1}^2 จึงแสดงว่าค่าความแปรปรวนอย่างมีเงื่อนไขไม่ใช่ค่าคงที่ และจะได้จากแบบจำลองในการประมาณค่าส่วนที่เหลือออกมาดังสมการ (4.20)

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \ldots + \alpha_q\,\hat{\varepsilon}_{t-1}^2 + v_t \qquad (4.20)$$

โดยที่ v_t คือ white noise process

ถ้าค่าของ $\alpha_1, \alpha_2, \ldots, \alpha_q$ เท่ากับศูนย์ ค่าความแปรปรวนจากการประมาณจะเท่ากับค่าคงที่ α_0 หรือ คือค่าความแปรปรวนอย่างมีเงื่อนไขของ x_t จะมีการเปลี่ยนแปลงสอดคล้องกับ Autoregression ในสมการ (4.20) ดังนั้นจะสามารถใช้สมการ (4.20) ในการพยากรณ์ค่า ความแปรปรวนอย่างมีเงื่อนไขที่เวลา t+1 ดังสมการ (4.21)

$$E_t\hat{\varepsilon}_{t+1}^2 = \alpha_0 + \alpha_1\hat{\varepsilon}_t^2 + \alpha_2\,\hat{\varepsilon}_{t-1}^2 + \ldots + \alpha_q\,\hat{\varepsilon}_{t+1-q}^2 \qquad (4.21)$$

จากเหตุผลที่กล่าวมาสมการ (4.20) เรียกว่า Autoregressive Conditional Heteroscadastic (ARCH) Model และสมการ (4.21) เป็น ARCH (q) ค่า $E_t\hat{\varepsilon}_{t+1}^2$ หรือ σ_{t+1}^2 จะประกอบด้วย 2 ส่วนคือ ค่าคงที่และความผันผวนในคาบเวลาที่ผ่านมา ซึ่งเขียนได้เป็นส่วนเหลือกำลังสองของคาบในอดีต (ARCH term) ส่วนค่าสัมประสิทธิ์ ($\alpha_0, \alpha_1, \ldots, \alpha_q$) สามารถหาค่าได้โดยใช้วิธี Maximum Likelihood

$$P_t = c + \beta_n P_{1-p} + \varepsilon_t + \theta_n \varepsilon_{1-q} + \gamma h_t^{1/2} \tag{4.22}$$

$$h_t = c + \alpha_p \varepsilon_{t-p}^2 + \phi_q h_{t-q} \tag{4.23}$$

โดยที่ P_t คือ ราคาปิดของแต่ละหลักทรัพย์ในเวลาที่ t

ε_t คือ ปัจจัยอื่นที่มีผลต่อการเปลี่ยนแปลงของหลักทรัพย์ในเวลาที่ t

h_t คือ ความแปรปรวนอย่างมีเงื่อนไขของ ε_t

β_n คือ สัมประสิทธิ์ค่า Autoregressive จากการประมาณด้วยสมการ (4.22)

θ_n คือ สัมประสิทธิ์ความคลาดเคลื่อนจากการประมาณด้วยสมการ (4.23)

γ_n คือ สัมประสิทธิ์เทอม GARCH-M จากการประมาณด้วยสมการ (4.22)

α_p คือ สัมประสิทธิ์ ARCH จากการประมาณค่าความล่าที่ p ด้วยสมการ (4.23)

ϕ_q คือ สัมประสิทธิ์ GARCH จากการประมาณค่าความล่าที่ q ด้วยสมการ (4.23)

4.1.4 **แบบจำลอง** Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

Bollerslev (1986) ได้ขยายมาจาก ARCH model โดยมีขั้นตอน คือให้ค่าความคลาดเคลื่อนจากกระบวนการเป็นดังสมการ (4.24)

$$\varepsilon_t = v_t \sqrt{h_t} \tag{4.24}$$

โดยที่ความแปรปรวนของ $v_t = \sigma_v^2 = 1$

และ

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{4.25}$$

เนื่องจาก $\{v_t\}$ เป็น white noise process ซึ่งเป็นอิสระกับ (ε_{t-i}) ค่าเฉลี่ยแบบมีเงื่อนไข (Conditional and Unconditional Means) ของ ε_t จะมีค่าเท่ากับศูนย์ ใส่ค่าคาดหมาย (Expected Value) ของ ε_t จะได้

$$E\varepsilon_t = Ev_t \sqrt{h_t} = 0 \tag{4.26}$$

โดยที่ความแปรปรวนแบบมีเงื่อนไข ของ ε_t ถูกกำหนดโดย

$$E_{t-1}\varepsilon_t^2 = h_t \tag{4.27}$$

ดังนั้นความแปรปรวนแบบมีเงื่อนไขของ ε_t จึงถูกกำหนดโดย h_t ในสมการ (4.27) แบบจำลองนี้จึงถูกเรียกว่า Generalized Autoregressive Conditional Heteroscedasticity หรือ GARCH (p,q) ซึ่งมีทั้งส่วนประกอบที่เป็น Autoregressive Moving Average ในความแปรปรวนที่มีลักษณะ heteroscedastic variance จะเห็นได้ว่า ถ้า p = 0 และ q = 1 เราก็จะได้แบบจำลอง GARCH (0,1) ซึ่งก็คือ ARCH (1) หรือ ARCH (q=1) นั่นเอง โดยสรุปแล้ว ถ้า β_i ทุกตัวมีค่าเท่ากับศูนย์ แบบจำลอง GARCH ก็คือ ARCH (q) นั่นเอง และเพื่อที่จะทำให้ความแปรปรวนแบบมีเงื่อนไข เป็นอันตะ (Finite) รากลักษณะเฉพาะ (Characteristic Roots) ของสมการ (4.27) จะต้องอยู่ในวงกลมหน่วย (Unit Circle)

เนื่องจากแบบจำลอง GARCH มีลักษณะเป็น ARMA process ACF (Autocorrelation Function) และ PACF (Partial Autocorrelation Function) ของส่วนตกค้างหรือส่วนที่เหลือ จะเป็นเครื่องชี้เกี่ยวกับ white-noise process อย่างไรก็ตาม ACF ของส่วนที่เหลือหรือส่วนตกค้างกำลังสอง (Squared Residuals) สามารถช่วยระบุถึง order ของ GARCH process ได้
เนื่องจาก $E_{t-1}\varepsilon_t = \sqrt{h_t}$ เราสามารถเขียนสมการ (4.27) ใหม่ ได้ดังนี้

$$E_{t-1}\varepsilon_t^2 = \gamma_0 + \sum_{i=1}^{q}\gamma_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i} \tag{4.28}$$

จะเห็นได้ว่าสมการ (4.26) มีลักษณะคล้ายกับ ARMA (q,p) ใน $\{\varepsilon_t^2\}$ sequenceมาก ถ้า heteroscedasticity แบบมีเงื่อนไขมีอยู่จริง แผนภาพสหสัมพันธ์ (Correlogram) จะเป็นตัวบ่งบอกกระบวนการ (process) ดังกล่าว

4.1.5 แบบจำลอง ARCH-in-mean (ARCH-M)

พื้นฐานจากแนวคิด ARCH ได้มีการขยายเพื่อให้ค่าเฉลี่ยของลำดับอนุกรมเวลาขึ้นอยู่กับความแปรปรวนอย่างมีเงื่อนไข ซึ่งเรียกว่า ARCH-M เพื่อความเหมาะสมในการศึกษาโดยมีหลักการพื้นฐาน คือ ผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยง (Risk Average) จะต้องการการชดเชยสำหรับการถือสินทรัพย์ที่มีความเสี่ยง (Risk Premium) เมื่อให้ความเสี่ยงในสินทรัพย์สามารถวัดค่าได้จากความแปรปรวนของ

ผลตอบแทน ค่าชดเชยความเสี่ยงจะเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของผลตอบ โดยแนวคิดนี้ได้นำส่วนที่เหลือของผลตอบแทนจากการถือครองทรัพย์สินที่มีความเสี่ยง ดังสมการ (4.29)

$$x_t = \mu_t + \varepsilon_t \tag{4.29}$$

โดยที่ x_t คือ ผลตอบแทนส่วนเกินจากการถือครองทรัพย์สินในระยะยาวเปรียบเทียบกับ หนึ่งคาบเวลาของผลตอบแทนอ้างอิง

μ_t คือ ค่าชดเชยความเสี่ยงที่จำเป็นในการโน้มน้าวผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยงให้ถือครองทรัพย์สินในระยะยาวแทนที่จะเป็นในคาบเวลาเดียว

ε_t คือ Unforecastable Shock ของส่วนที่เหลือของผลตอบแทนในการถือทรัพย์สินในระยะยาว

ในการอธิบายสมการ (4.29) จะให้ค่าคาดหวังของส่วนที่เหลือของผลตอบแทนจากการถือครองทรัพย์สินในระยะยาวมีค่าเท่ากับค่าชดเชยความเสี่ยงดังสมการ (4.30)

$$E_{t-1}x_t = \mu_t \tag{4.30}$$

ถ้าค่าชดเชยความเสี่ยงเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของ ε_t อีกนัยหนึ่งคือ ยิ่งค่าความแปรปรวนอย่างมีเงื่อนไขของผลตอบแทนมีขนาดใหญ่ ผลตอบแทนของการโน้มน้าวให้คนหันมาถือครองทรัพย์สินในระยะยาวก็จะยิ่งมีค่ามากขึ้นตาม ถ้า h_t เป็นความแปรปรวนอย่างมีเงื่อนไขของ ε_t ค่าชดเชยความเสี่ยงสามารถแสดงได้เป็น

$$\mu_t = \beta + \delta h_t \quad , \delta > 0 \tag{4.31}$$

โดยที่ h_t คือ ARCH (q) process

$$h_t = \alpha_0 + \sum_{i=0}^{q} \alpha_i \hat{\varepsilon}_{t-i}^2 \tag{4.32}$$

จากสมการ (4.30), (4.31) และ (4.32) ประกอบขึ้นเป็นพื้นฐานของแบบจำลอง ARCH-M จากสมการ (4.31) และ (4.32) ค่าเฉลี่ยอย่างมีเงื่อนไขของ x_t จะขึ้นอยู่กับค่าความแปรปรวนอย่างมี

เงื่อนไข h_t จากสมการ (4.32) ความแปรปรวนอย่างมีเงื่อนไขได้จากกระบวนการ ARCH(q) ซึ่งเป็นการอธิบายว่า ถ้าความแปรปรวนอย่างมีเงื่อนไขเป็นค่าคงที่ (เช่น ถ้า $\alpha_1 = \alpha_2 = ... = \alpha_q = 0$) แบบจำลอง ARCH-M จะกลับไปสู่กรณีเดิมของค่าชดเชยความเสี่ยงคงที่

4.1.6 แบบจำลอง GARCH-in-mean (GARCH-M)

จากแบบจำลอง GARCH (p,q) ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนเหลือของสมการ (4.31) ดังนี้ (Bollerslev, 1986)

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \qquad (4.33)$$

ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนเหลือกำลังสองของสมการ (4.33)นี้ (Engle; Lilien and Robins, 1987) ขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักกันในชื่อ GARCH-M หรือ GARCH in Mean ผลตอบแทนจากหลักทรัพย์สามารถแสดงได้ดังสมการ (4.34) ถึง (4.36)

$$x_t = \mu_t + \delta_1 h_t^{1/2} + \varepsilon_t \qquad (4.34)$$

$$\varepsilon_t / \psi_{t-1} : N(0, h_t) \qquad (4.35)$$

$$h_t = \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \alpha_i (\varepsilon_{t-i})^2 \qquad (4.36)$$

โดยที่ 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

$h_t^{1/2}$ คือ ความสัมพันธ์โดยตรง ระหว่างความเสี่ยงและผลตอบแทนที่คาดหวัง

ปัจจัยอย่างหนึ่งที่มีอิทธิพลอย่างมีนัยสำคัญของความผันผวนในผลตอบแทนจากหลักทรัพย์สามารถวัดได้ด้วยสัมประสิทธิ์ $h_t^{1/2}$ (δ_1) ในสมการ (4.34) ซึ่งแสดงแทนดัชนีของความสัมพันธ์ของการหลีกเลี่ยงความเสี่ยง ค่าสัมประสิทธ์ (δ_1) ที่เป็นบวกอย่างมีนัยสำคัญบอกถึงผู้ลงทุนในหลักทรัพย์เมื่อเผชิญกับระดับความเสี่ยงที่สูงขึ้นก็ต้องการค่าชดเชยความเสี่ยงที่มากขึ้นตามไปด้วย

4.2 การตรวจสอบรูปแบบ (Diagnostic Checking)

การสร้างสมการพร้อมทั้งประมาณค่าพารามิเตอร์แล้วนั้น จะต้องทำการตรวจสอบรูปแบบว่าสมการพยากรณ์ที่ได้มามีความเหมาะสมหรือไม่ และรูปแบบใดของสมการที่ดีที่สุด โดยใช้การทดสอบต่าง ๆ ดังนี้

4.2.1 การทดสอบ Box-Pierce Q –Statistics

เป็นการทดสอบว่าสหสัมพันธ์ในตัวเองในส่วนเหลือทุกช่วงเวลาที่ห่างกัน K มีความเป็นอิสระต่อกันหรือไม่

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

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

คำนวณตามสมการที่ (4.37) คือ

$$\text{Q-stat} = T\sum_{j=1}^{k}\rho_j^2 \tag{4.37}$$

โดยที่ ρ_j คือ สหสัมพันธ์ในตัวเองลำดับที่ j โดยที่ j = 1,....,k

T คือ จำนวนของค่าสังเกต (Observations)

ภายใต้ส่วนเหลือจากการประมาณด้วยแบบจำลอง ARIMA ค่า Q-stat มีการแจกแจงแบบไคสแควร์ (χ^2) ด้วยระดับความเป็นอิสระ (Degree of Freedom) เท่ากับจำนวนของสหสัมพันธ์ในตัวเอง ลบด้วยจำนวนของพารามิเตอร์ Autoregressive (AR) และ Moving Average (MA) ที่ได้มาจากการประมาณหรือ k-m

และจะยอมรับสมมติฐานหลักเมื่อ Q-stat $\leq \chi^2_{\alpha,k-m}$ คือส่วนเหลือเป็นอิสระต่อกันที่ความล่า k และถ้าปฏิเสธสมมติฐานหลักเมื่อ Q-stat $> \chi^2_{\alpha,k-m}$ คือ เกิดสหสัมพันธ์ในตัวเองอย่างน้อยหนึ่งค่าในส่วนเหลือที่ไม่เท่ากับศูนย์

4.2.2 เกณฑ์การเลือกแบบจำลองที่ดีที่สุด (Information Criteria)

ในการหาแบบจำลอง เมื่อได้แบบจำลองที่เหมาะสมหลายแบบจำลอง จึงต้องมีแนวทางในการเลือกแบบจำลองที่ดีที่สุด โดยการพิจารณาค่า Akaike Information Criterion (AIC) แบบจำลองที่ให้ค่า AIC น้อยที่สุด จะเป็นแบบจำลองที่ดีที่สุด โดย

Akaike Information Criterion (AIC) $= -2\ell/\eta + 2k/\eta$ (4.38)

โดยที่ k คือ จำนวนของพารามิเตอร์ที่ทำการประมาณค่า

η คือ จำนวนของค่าสังเกต

ℓ คือ ค่าของ Log Likelihood Function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว

AIC คือค่าสถิติประยุกต์ที่คล้ายกับ Adjusted R^2 แต่ใช้รูปแบบการใส่ค่าลอกกาลิทึม ฐานธรรม (Natural Logarithm) หากค่า AIC นี้มีค่าน้อยเพียงใด สามารถอธิบายได้ว่าแบบจำลองที่ประมาณได้นั้นสามารถเป็นตัวแทนข้อมูลจริงได้ดีเพียงนั้น อีกทั้งค่า AIC นี้ยังเป็นค่าสถิติที่เหมาะสมที่จะนำไปใช้ในการค่าย้อนหลัง (Lag Length) ที่เหมาะสมได้อีกด้วย (Gujarati: 2003)