บทที่ 4

ทฤษฎีและแนวคิดที่ใช้ในการศึกษา

ในการศึกษาถึงความสัมพันธ์ระหว่างราคาปิดของหลักทรัพย์กับช่วงเวลาต่างๆในอดีต ความผันผวนของข้อมูลและการนำข้อมูลไปพยากรณ์ราคาที่คาดว่าจะเกิดขึ้นในอนาคต โดยใช้แนวคิดและทฤษฎีเกี่ยวกับการวิเคราะห์อนุกรมเวลาการทดสอบความเป็น Stationary ของข้อมูล และการทดสอบ 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 แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใด และสามารถพยากรณ์การเปลี่ยนแปลงข้อมูลในอนาคตได้ การวิเคราะห์ข้อมูลอนุกรมเวลาจะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน (Enders , 1995)

4.2 การทดสอบความนิ่งของข้อมูล (Stationary) และการทดสอบ Unit root

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง (Stationary) คือข้อมูลที่ค่าเฉลี่ยและความแปรปรวนของกระบวนการเชิงสุ่มนั้นมีค่าคงที่เมื่อเวลาได้เปลี่ยนไป และค่าความแปรปรวนระหว่างสองคาบเวลาขึ้นอยู่กับความล่า (Lag) ระหว่างคาบเวลาทั้งสอง โดยสามารถเขียนเป็นสมการได้ดังนี้

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 ข้อมูลจะต้องมีลักษณะนิ่ง เนื่องจากข้อมูลอนุกรมเวลานั้นมาจากกระบวนการเชิงสุ่ม (Random Process) การนำข้อมูลอนุกรมเวลาไปใช้โดยไม่ได้ทำการตรวจสอบว่าข้อมูลนั้นมีลักษณะนิ่งนั้น ค่าสถิติที่เกิดขึ้นจะมีการแจกแจงไม่มาตรฐาน (Nonstandard Distribution) ซึ่งทำให้การนำไปใช้เปรียบเทียบกับค่าในตารางมาตรฐานไม่ถูกต้อง เนื่องจากค่าต่าง ๆ นั้น มีสมมติฐานว่าข้อมูลนั้นมีการแจกแจงมาตรฐาน (Standard Distributions) ทำให้นำไปสู่การลงความเห็นที่ผิดพลาดและความสัมพันธ์ที่ไม่แน่จริง (Spurious Regression) กล่าวคือ R^2 มีค่าสูงมากและได้ค่าสถิติ t – test มีนัยสำคัญหรือสูงเกินกว่าความเป็นจริง

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 $$\Delta x_t = \theta x_{t-1} + \varepsilon_t \tag{4.5}$$

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

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

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

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

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

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 (Enders , 1995)

4.3 แบบจำลอง Autoregressive Conditional Heteroscedasticity (ARCH)

ในแบบจำลองเศรษฐมิติแบบดั้งเดิมได้มีการสมมติให้ความแปรปรวน (variance) ของเทอมความคลาดเคลื่อน (error term) มีค่าคงที่หรือคงตัว (constant) ซึ่ง Enders ได้แสดงให้เห็นว่าข้อมูลเศรษฐกิจอนุกรมเวลา (time series) จำนวนมากในคาบเวลาจำนวนไม่น้อยมีความผันผวนสูงมากตามมาด้วยคาบเวลาที่อนุกรมดังกล่าวค่อนข้างจะมีความสงบซึ่งจะเห็นได้ว่าข้อสมมติที่ว่าความแปรปรวน (variance) ของเทอมความคลาดเคลื่อน (error term) มีค่าคงที่หรือค่าคงตัว (constant) นั้น ไม่น่าจะเป็นข้อสมมติที่เหมาะสมหรือถูกต้อง ซึ่ง Enders กล่าวว่า ในหลายสถานการณ์ เราสนใจแต่เพียงความแปรปรวนแบบมีเงื่อนไข (conditional variance) เท่านั้น เช่นนักลงทุนในตลาดหุ้น อาจจะสนใจในการพยากรณ์อัตราผลตอบแทน (rate of return) และความแปรปรวน (variance) ของหุ้นที่เราถือเท่านั้น ในขณะที่ความแปรปรวนแบบไม่มีเงื่อนไข (unconditional variance ซึ่งคือความแปรปรวนระยะยาวนั่นเอง) อาจจะไม่ใช่สิ่งที่สำคัญ ถ้านักลงทุนวางแผนที่จะซื้อขายหุ้นในช่วงไม่ยาวจนเกินไปนัก

ในการพยากรณ์ความแปรปรวน (variance) วิธีหนึ่งที่เรามักจะคุ้นเคยก็คือ แบบจำลองที่แสดงความสัมพันธ์ระหว่าง y_{t+1} กับ ϵ_{t+1} และ x_t ซึ่งเขียนเป็นสมการได้ดังนี้ (Enders , 1995)

$$x_{t+1} = \epsilon_{t+1} x_t \tag{4.11}$$

โดยที่ x_{t+1} = ตัวแปรที่เรากำลังพิจารณา

ϵ_{t+1} = ตัวเทอมรบกวน white noise(white noise disturbance term) ซึ่งมีความแปรปรวนเท่ากับ σ^2 ซึ่งเป็นค่าคงที่หรือคงตัว (constant)

x_t = ตัวแปรอิสระ (independent variable) ณ คาบเวลา t ซึ่งเป็นตัวแปรที่เราสังเกตได้เช่นเดียวกัน

จากสมการ (4.11) จะสังเกตได้ว่าถ้า x_t มีค่าเท่ากับทุกคาบเวลาและเท่ากับค่าคงตัวหรือค่าคงที่ (constant) ซึ่งสมมติว่าเท่ากับ x สมการ(4.11) จะสามารถเขียนได้ใหม่ดังนี้

$$x_{t+1} = \epsilon_{t+1} \mathrm{x} \tag{4.12}$$

เราจะได้ว่า $\{x_{t+1}\}$ sequence ก็จะมีลักษณะเป็น white noise process ด้วยความแปรปรวน (variance) คงที่หรือคงตัว (constant) อย่างไรก็ตาม $\{x_t\}$ sequence มักจะมีค่าไม่เท่ากัน เพราะฉะนั้นความแปรปรวน ภายใต้เงื่อนไขของ x_t สามารถเขียนได้ดังนี้

$$Var(x_{t+1} \mid x_t) = \sigma^2 x_t^2 \tag{4.13}$$

และถ้าค่าสืบเนื่อง (successive values) ของ $\{x_{t+1}\}$ มี positive serial correlation ความแปรปรวนแบบมีเงื่อนไข (conditional variance) ของ $\{x_t\}$ sequence ก็จะมี positive serial correlation ด้วย ในลักษณะเช่นนี้ $\{x_{t+1}\}$ sequence ก็จะทำให้เกิดคาบเวลาของความผันผวนใน $\{x_t\}$ sequence

ในทางปฏิบัติแล้ว เราอาจจะปรับปรุงแบบจำลองที่กล่าวมาแล้วข้างต้นให้อยู่ในรูปแบบดังนี้

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

โดยที่ e_t = เทอมความคลาดเคลื่อนซึ่งคือ $\ln(\in_t)$ นั่นเอง

และสามารถทำการถดถอยโดยใช้ OLS (Ordinary Least Square regression) แต่ละจุดอ่อนของวิธีก็คือเราสมมติไว้แน่นอนว่า $\{x_{t+1}\}$ เป็นสาเหตุของการเปลี่ยนแปลงความแปรปรวน (variance) และโดยเหตุผลทางทฤษฎีแล้วเราอาจจะไม่มีเหตุผลที่ดีเพียงพอในการเลือกตัวแปร $\{x_{t+1}\}$ ที่เป็นสาเหตุของการเปลี่ยนแปลงของความแปรปรวนได้ และสิ่งที่เป็นจุดอ่อนที่สำคัญอีกประการหนึ่งของแบบจำลองสมการ (4.14) ก็คือ เราได้สมมติว่าเทอมความคลาดเคลื่อน (error term) ซึ่งคือ $\{e_t\}$ sequence มีความแปรปรวนคงที่หรือคงตัว (constant variance) ถ้าข้อสมมติดังกล่าวไม่ถูกต้องก็จะต้องมีการแปลงข้อมูล (data transformation) อีก

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

ความเป็นไปได้ในการหาค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาไปพร้อมกันนั้นในขั้นต้นการพยากรณ์อย่างมีเงื่อนไขจะมีความแม่นยำเหนือกว่าการพยากรณ์อย่างไม่มีเงื่อนไขมาก ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) แสดงได้ดังนี้

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

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

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

ถ้าเปลี่ยนไปใช้การพยากรณ์แบบไม่มีเงื่อนไขแล้ว ผลที่ใช้จะเป็นค่าเฉลี่ยในช่วงระยะยาว ของลำดับ $\{x_t\}$ ซึ่งเท่ากับ $\dfrac{a_0}{(1-a_1)}$ จะได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไขตามสมการ (4.18) คือ

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

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

เมื่อ $\dfrac{1}{(1-a_1^2)} > 1$ ค่าความแปรปรวนที่ได้จากการพยากรณ์อย่างไม่มีเงื่อนไขจะสูงกว่าแบบมีเงื่อนไข ดังนั้นในการพยากรณ์อย่างไม่มีเงื่อนไขจึงมีความเหมาะสมกว่า ในลักษณะเดียวกันถ้าความแปรปรวนของ $\{\epsilon_t\}$ ไม่เป็นค่าคงที่ จะสามารถประมาณค่าแนวโน้มของการเปลี่ยนแปลงความแปรปรวนโดยใช้ ARMA Modelอธิบายได้โดยให้ $\{\hat{\epsilon}_t\}$ แทนส่วนที่เหลือ (Residual) ที่ได้จากการประมาณจากสมการ (4.15) ดังนั้นค่าความแปรปรวนอย่างมีเงื่อนไข (conditional Variance) ของ x_{t+1} จะได้ดังสมการ (4.19)

$$Var(x_{t+1} \mid x_t) = E_t[(x_{t+1} - a_0 - a_1 x_t)^2]$$

$$= E_t \in_{t+1}^2 \tag{4.19}$$

และจากที่ให้ $E_t \in_{t+1}^2$ เท่ากับ σ_{t+1}^2 จึงแสดงว่าค่าความแปรปรวนอย่างมีเงื่อนไขไม่ใช่ค่าคงที่ และจะได้จากแบบจำลองในการประมาณค่าส่วนที่เหลือ (Residuals) ออกมาดังสมการ (4.20)

$$\hat{\in}_t^2 \quad = \alpha_0 + \alpha_1 + \hat{\in}_{t-1}^2 + ... + \alpha_q \hat{\in}_{t-q}^2 + v_t \tag{4.20}$$

เมื่อ v_t = white noise process

ถ้าค่าของ $\alpha_1, \alpha_2, ..., \alpha_q$ เท่ากับศูนย์ ค่าความแปรปรวนจากการประมาณจะเท่ากับค่าคงที่ α_0 อีกนัยหนึ่ง คือ ค่าความแปรปรวนอย่างมีเงื่อนไขของ x_t จะมีการเปลี่ยนแปลงสอดคล้องกับ Autoregression ในสมการ(4.20) ดังนั้นจะสามารถใช้สมการ (4.20) ในการพยากรณ์ค่า ความแปรปรวนอย่างมีเงื่อนไขที่เวลา t+1 ดังสมการ (4.21)

$$E_t \hat{\in}_{t+1}^2 \quad = \alpha_0 + \alpha_1 \hat{\in}_t^2 + \alpha_2 + \hat{\in}_{t-1}^2 + ... + \alpha_q \hat{\in}_{t+1-q}^2 \tag{4.21}$$

จากเหตุผลที่กล่าวมา สมการที่(4.20) เรียกว่า Autoregressive Conditional Heteroscadastic (ARCH) Model และสมการ (4.21) เป็น ARCH (q) สมการ (4.21) ค่า $E_t \hat{\in}_{t+1}^2$ หรือ σ_{t+1}^2 จะประกอบด้วย 2 องค์ประกอบคือ ค่าคงที่และความผันผวน (Volatility) ในคาบเวลาที่ผ่านมา ซึ่งเขียนได้เป็นส่วนเหลือกำลังสองของคาบในอดีต(ARCHterm) ส่วนค่าสัมประสิทธิ์ $(\alpha_0, \alpha_1, ..., \alpha_q)$ สามารถหาค่าได้โดยใช้วิธี Maximum Likelihood

4.4 แบบจำลอง ARCH-in-mean (ARCH-M)

พื้นฐานจากแนวคิด ARCH ได้มีการขยายเพื่อให้ค่าเฉลี่ยของลำดับอนุกรมขึ้นอยู่กับความแปรปรวนอย่างมีเงื่อนไขของตัวเอง ซึ่งเรียกว่า ARCH-M เพื่อที่จะมีความเหมาะสมในการศึกษา โดยมีหลักการพื้นฐาน คือ ผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยง (Risk Average) จะต้องการชดเชยสำหรับการถือสินทรัพย์ที่มีความเสี่ยง (Risk Premium) เมื่อให้ความเสี่ยงในสินทรัพย์สามารถวัดค่าได้จากความแปรปรวนของผลตอบแทน ค่าชดเชยความเสี่ยงจะเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของผลตอบ โดยแนวคิดนี้ได้นำส่วนที่เหลือของผลตอบแทนจากการถือครองทรัพย์สินที่มีความเสี่ยง ดังสมการ (4.22) (Engle , 1987)

$$x_t = \mu_t + \epsilon_t \tag{4.22}$$

เมื่อ x_t = ผลตอบแทนส่วนเกินจากการถือครองทรัพย์สินในระยะยาวเปรียบเทียบกับหนึ่ง คาบเวลาของผลตอบแทนอ้างอิง

μ_t = ค่าชดเชยความเสี่ยงที่จำเป็นในการโน้มน้าวผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยงให้ถือครองทรัพย์สินในระยะยาวแทนที่จะเป็นในคาบเวลาเดียว

ϵ_t = Unforecastable Shock ของส่วนที่เหลือของผลตอบแทนในการถือทรัพย์สินในระยะยาว

ในการอธิบายสมการ (4.22) จะให้ค่าคาดหวังของส่วนที่เหลือของผลตอบแทนจากการถือครองทรัพย์สินในระยะยาวมีค่าเท่ากับค่าชดเชยความเสี่ยงดังสมการ (4.23)

$$E_{t-1}x_t = \mu_t \tag{4.23}$$

ถ้าค่าชดเชยความเสี่ยงเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของ ϵ_t อีกนัยหนึ่งคือ ยิ่งค่าความแปรปรวนอย่างมีเงื่อนไขของผลตอบแทนมีขนาดใหญ่ ผลตอบแทนของการโน้มน้าวให้คนหันมาถือครองทรัพย์สินในระยะยาวก็จะยิ่งมีค่ามากขึ้นตาม ถ้า h_t เป็นความแปรปรวนอย่างมีเงื่อนไขของ ϵ_t ค่าชดเชยความเสี่ยงสามารถแสดงได้เป็น

$$\mu_t = \beta + \delta h_t \qquad , \delta > 0 \tag{4.24}$$

เมื่อ h_t = คือ ARCH (q) process

$$h_t = \alpha_0 + \sum_{i=0}^{q} \alpha_i \hat{\epsilon}_{t-i}^2 \tag{4.25}$$

จากสมการ (4.23), (4.24) และ (4.25) ประกอบขึ้นเป็นพื้นฐานของแบบจำลอง ARCH-M จากสมการ (4.23) และ (4.24) ค่าเฉลี่ยอย่างมีเงื่อนไขของ x_t จะขึ้นอยู่กับค่าความแปรปรวนอย่างมีเงื่อนไข h_t จากสมการ (4.25) ความแปรปรวนอย่างมีเงื่อนไขได้จากกระบวนการ ARCH(q) ซึ่งเป็นการอธิบายว่า ถ้าความแปรปรวนอย่างมีเงื่อนไขเป็นค่าคงที่ (เช่น ถ้า $\alpha_1 = \alpha_2 = ... = \alpha_q = 0$) แบบจำลอง ARCH-M จะกลับไปสู่กรณีเดิมของค่าชดเชยความเสี่ยงคงที่

4.5 แบบจำลอง GARCH

ในแบบจำลอง GARCH ให้ความแปรปรวนแบบมีเงื่อนไข (conditional variance) มีลักษณะเป็น ARMA process โดยที่ให้ error process มีลักษณะดังนี้ (Enders , 1995)

$$\epsilon_t \quad = \quad v_t\sqrt{h_t} \tag{4.26}$$

โดยที่ความแปรปรวนของ $v_t = \sigma_v^2 = 1$ และ

$$h_t \quad = \quad \alpha_0 + \sum_{i=1}^{q}\alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i} \tag{4.27}$$

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

$$E\epsilon_t \quad = \quad Ev_t\sqrt{h_t} = 0 \tag{4.28}$$

ประเด็นที่สำคัญก็คือว่า ความแปรปรวนแบบมีเงื่อนไข (conditional variance) ของ ϵ_t ถูกกำหนดโดย

$$E_{t-1}\epsilon_t^2 = \quad h_t \tag{4.29}$$

ดังนั้นความแปรปรวนแบบมีเงื่อนไขของ ϵ_t จึงถูกกำหนดโดย h_t ในสมการ (4.29) แบบจำลองนี้จึงถูกเรียกว่า generalized autoregressive conditional heteroscedasticity ซึ่งใช้ตัวย่อว่า GARCH (p,q) ซึ่งได้เปิดโอกาสให้มีทั้งส่วนประกอบที่เป็น 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 ทุกตัวมีค่าเท่ากับศูนย์แบบจำลอง GARCH ก็คือ ARCH (q) นั่นเอง และเพื่อที่จะทำให้ความแปรปรวนแบบมีเงื่อนไข (conditional variance) เป็นอันตะ (finite) รากลักษณะเฉพาะ (characteristic roots) ของสมการ (4.29) จะต้องอยู่ในวงกลมหน่วย (unit circle)

เนื่องจากแบบจำลอง GARCH มีลักษณะเป็น ARMA process ACF (autocorrelation function) และ PACF (partial autocorrelation function) ของส่วนตกค้างหรือส่วนที่เหลือ (residuals) จะเป็น

เครื่องชี้เกี่ยวกับ white-noise process อย่างไรก็ตาม ACF (autocorrelation function) ของส่วนที่เหลือหรือส่วนตกค้างกำลังสอง (squared residuals) สามารถช่วยระบุถึง order ของ GARCH process ได้

เนื่องจาก $E_{t-1} \in_t \; = \; \sqrt{h_t}$ เราสามารถเขียนสมการ(4.29) ใหม่ ได้ดังนี้

$$E_{t-1} \in_t^2 \; = \; \gamma_0 + \sum_{i=1}^{q} \gamma_i \in_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{4.30}$$

จะเห็นได้ว่าสมการ(4.30) มีลักษณะคล้ายกับ ARMA (q,p) ใน $\{\in_t^2\}$ sequence มาก ถ้า heteroscedasticity แบบมีเงื่อนไขมีอยู่จริง แผนภาพสหสัมพันธ์ (correlogram) จะเป็นตัวบ่งบอกกระบวนการ (process) ดังกล่าว

ในการที่จะดูว่ามี conditional heteroscedasticity หรือไม่นั้น เราอาจดูได้จากแผนภาพสหสัมพันธ์ (correlogram) สำหรับการสร้างแผนภาพสหสัมพันธ์ (correlogram) ของส่วนที่เหลือหรือส่วนตกค้างกำลังสอง (squared residuals) นั้น ได้สรุปไว้แล้วดังนี้

ขั้นตอนที่ 1 : เราจะต้องประมาณค่า $\{x_t\}$ sequence ก่อนโดยใช้แบบจำลอง ARMA หรือแบบจำลองถดถอย ที่ best-fitting และหลังจากนั้นก็คำนวณหาส่วนตกค้างหรือส่วนที่เหลือกำลังสอง (squared residual) $\in_t^2$ ของแบบจำลองดังกล่าวที่ประมาณค่าได้จากนั้นคำนวณค่าความแปรปรวนจากตัวอย่าง (sample variance) ของส่วนที่เหลือหรือส่วนตกค้าง (residual) σ^2 ด้วยสูตร

$$\sigma^2 \; = \; \sum_{t=1}^{n} \in_t^2 / n$$

โดยที่ n = จำนวนของส่วนที่เหลือหรือส่วนตกค้าง

ขั้นตอนที่ 2 : คำนวณและพล็อตอัตสหสัมพันธ์ตัวอย่าง ของส่วนตกค้างหรือส่วนที่เหลือกำลังสอง ดังนี้

$$r_k \quad = \quad \frac{\sum_{t=k+1}^{n} (\hat{\in}_t - \hat{\sigma}^2)(\hat{\in}_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^{n} (\hat{\in}_t^2 - \hat{\sigma}^2)}$$

ขั้นตอนที่ 3 : เป็นขั้นตอนการทดสอบสมมติฐานว่ามี ARCH หรือ GARCH errors หรือไม่ สำหรับการทดสอบนั้นมี 2 แบบดังนี้

แบบที่ 1 : การทดสอบทีละค่าของ r_k ว่าแตกต่างจากศูนย์อย่างมีนัยสำคัญหรือไม่ โดยใช้ t-test โดยที่ค่าเบี่ยงเบนมาตรฐาน (standard deviation) ของ r_k จะมีค่าโดยประมาณเท่ากับ $1/\sqrt{n}$ ในตัวอย่างขนาดใหญ่ ถ้าผลการทดสอบปรากฏว่า r_k มีค่าไม่แตกต่างไปจากศูนย์เราก็จะสรุปว่าไม่มี ARCH หรือ GARCH errors

แบบที่ 2 : การทดสอบเป็นกลุ่มซึ่งเป็นการทดสอบกลุ่มของสัมประสิทธิ์ของ r_k โดยใช้ Box - Price Q –Statistics ซึ่งมีสูตรดังนี้

$$Q = n\sum_{k=1}^{m} r_k$$

ซึ่งมีการแจกแจงแบบไคสแควร์เชิงเส้นกำกับ (asymptotic χ^2 distribution) ด้วย degree of freedom เท่ากับ m ภายใต้สมมติฐานว่างว่า $\hat{\epsilon}_t^2$ ไม่มีสหสัมพันธ์ (uncorrelation) ถ้าปฏิเสธสมมติฐานว่างดังกล่าวก็หมายความว่าเป็นการปฏิเสธสมมติฐานว่าง

H_0 : no ARCH หรือ GARCH errors

ในทางปฏิบัติแล้วเราจะให้ค่าของ m มีค่าถึง n/4

สำหรับการทดสอบ ARCH ที่เสนอโดย Engle นั้นเป็นการทดสอบ Lagrange multiplier test ซึ่งได้สรุปไว้ดังนี้

ขั้นตอนที่ 1 : หาแบบจำลอง AR(p) หรือแบบจำลองถดถอยที่เหมาะสม ดังนี้

$$x_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_p x_{t-p} + \epsilon_t$$

โดยการใช้วิธี OLS

ขั้นตอนที่ 2 : หาค่า ϵ_t^2 และทำการถดถอย ϵ_t^2 กับ $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots, \epsilon_{t-q}^2$ โดยมีค่าคงที่หรือค่าคงตัวรวมอยู่ในตัวถดถอยด้วย นั่นคือการประมาณค่าแบบจำลอง

$$\epsilon_t^2 = \gamma_0 + \gamma_1 \epsilon_{t-1}^2 + \gamma_2 \epsilon_{t-2}^2 + \ldots + \gamma_q \epsilon_{t-q}^2 + \eta_t$$

โดยที่ η_t = เทอมความคลาดเคลื่อน (error term)

ถ้าไม่มี ARCH หรือ GARCH ค่า γ_1 ...... γ_q จะมีค่าไม่แตกต่างไปจากศูนย์ ซึ่งการถดถอยนี้จะมีค่า R^2 ต่ำทีเดียวภายใต้สมมติฐานว่า (null hypothesis)

H0 : no ARCH errors

สถิติทดสอบ nR^2 จะลู่เข้าหาการแจกแจงแบบไคสแควร์ χ^2 (q) และถ้า nR^2 มีค่าใหญ่มากเพียงพอ เราจะปฏิเสธสมมติฐานว่างดังกล่าว ซึ่งหมายความว่ามี ARCH errors แต่ถ้า nR^2 มีค่าต่ำมากจะนำไปสู่ข้อสรุปที่ว่าไม่มี ARCH errors

4.6 แบบจำลอง GARCH-M

จากแบบจำลอง GARCH (p,q)

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

ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนเหลือกำลังสองของกระบวนการนี้ (Engel,1987) ขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักกันในชื่อ GARCH-M หรือ GARCH in Mean ผลตอบแทนจากหลักทรัพย์สามารถแสดงได้ดังสมการ (4.31) ถึง (4.33)

$$x_t = \mu_t + \delta_1 h_t^{1/2} + \in_t \tag{4.31}$$

$$\in_t / \psi_{t-1} \sim N(0, h_t) \tag{4.32}$$

$$h_t = \omega + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \alpha_i (\in_{t-i})^2 \tag{4.33}$$

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 (4.31) นั้น เพื่อใช้แสดงความสัมพันธ์โดยตรง ถึง trade off ระหว่างความเสี่ยงและผลตอบแทนที่คาดหวัง

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

4.7 การตรวจสอบรูปแบบ (Diagnostic Checking)

การสร้างสมการพร้อมทั้งประมาณค่าพารามิเตอร์แล้วนั้น จะต้องทำการตรวจสอบรูปแบบว่าสมการพยากรณ์ที่ได้มามีความเหมาะสมหรือไม่ และรูปแบบใดของสมการที่ดีที่สุด โดยใช้การทดสอบต่าง ๆ ดังนี้ (Pindyck , 1997)

4.7.1 การทดสอบ Box - Price Q –Statistics

การทดสอบ Box - Price Q –Statistics เป็นการทดสอบว่าสหสัมพันธ์ในตัวเองในส่วนเหลือทุกช่วงเวลาที่ห่างกัน K มีความเป็นอิสระต่อกันหรือไม่
โดยมีสมมติฐานดังนี้ (ภัทร์ ตั้งตระกูล 2545)

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

และ

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

คำนวณตามสมการที่ (4.33) คือ

$$Q-stat = T\sum_{j=1}^{k} {r_j}^2 \tag{4.34}$$

เมื่อ r_j คือ สหสัมพันธ์ในตัวเองลำดับที่ j โดยที่ j = 1,…,k

T คือ จำนวนของค่าสังเกต

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

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

4.7.2 เกณฑ์การเลือกรูปแบบของแบบจำลองที่ดีที่สุด (Information Criteria)

ในการหารูปแบบของแบบจำลอง เมื่อได้รูปแบบของแบบจำลองที่เหมาะสมหลายรูปแบบ จึงต้องมีแนวทางในการเลือกรูปแบบของแบบจำลองที่ดีที่สุด โดยการพิจารณาค่า Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) รูปแบบที่ให้ค่า AIC และ SC น้อยที่สุด จะเป็นรูปแบบที่ดีที่สุด โดย (Pindyck , 1997)

Akaike Information Criterion (AIC) = $-2\ell/\eta + 2k/\eta$ (4.35)

Schwartz Criterion (SC) = $-2\ell/\eta + k\log\eta/\eta$ (4.36)

โดยที่ k เป็นจำนวนของพารามิเตอร์ที่ทำการประมาณค่า

η เป็นจำนวนของค่าสังเกต

ℓ เป็นค่าของ log likelihood function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว