บทที่ 3

ระเบียบวิธีการศึกษา

3.1 กรอบแนวคิดและแบบจำลองที่ใช้ในการศึกษา

การศึกษาเพื่อประมาณค่าความผันผวนของราคาหลักทรัพย์ในหมวดอาหารและเครื่องดื่มในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้แบบจำลอง ARIMA, ARFIMA-GARCH, EGRACH, FIGARCH, FIEGARCH ของราคาปิดของหลักทรัพย์ในอดีตเพื่อประมาณค่าความผันผวนในอนาคต

3.2 ข้อมูลที่ใช้ในการศึกษา

การศึกษาครั้งนี้ใช้ข้อมูลอนุกรมเวลาทุติยภูมิ (Secondary data) โดยใช้ข้อมูลราคาปิดรายวันของหลักทรัพย์หมวดอาหารและเครื่องดื่ม ขอบเขตข้อมูลที่ใช้เริ่มตั้งแต่วันที่ 2 มกราคม พ.ศ. 2549 ถึงวันที่ 30 มิถุนายน พ.ศ. 2554 รวมทั้งสิ้น 1434 วัน นำมาแปลงเป็นรูปผลตอบแทน

โดยการรวบรวมข้อมูลผ่านศูนย์การเงินและการลงทุน (Finance and Investment Center : FIC) มหาวิทยาลัยเชียงใหม่หลักทรัพย์ที่จะนำมาใช้ในการศึกษาคัดเลือกจากหลักทรัพย์หมวดอาหารและเครื่องดื่ม 4 หลักทรัพย์ ได้แก่

บริษัท ไทยยูเนียน โฟรเซ่น โปรดักส์ จำกัด (มหาชน) : TUF

บริษัท เจริญโภคภัณฑ์อาหาร จำกัด (มหาชน) : CPF

บริษัท น้ำตาลขอนแก่น จำกัด (มหาชน) : KSL

บริษัท น้ำมันพืชไทย จำกัด (มหาชน) : TVO

3.3 วิธีการศึกษา

3.3.1 ปรับข้อมูลให้อยู่ในรูปผลตอบแทน

สามารถปรับข้อมูลราคาหลักทรัพย์ให้อยู่ในรูปผลตอบแทนได้ดังสูตรต่อไปนี้

$$y_t = \log\left(\frac{P_t}{P_{t-1}}\right) \qquad (3.1)$$

y_t คือ อัตราผลตอบแทนของหลักทรัพย์ ณ เวลาที่ t

P_t คือ ราคาปิดของหลักทรัพย์ ณ เวลาที่ t

P_{t-1} คือ ราคาปิดของหลักทรัพย์ ณ เวลาที่ t-1

3.3.2 การทดสอบความนิ่ง หรือ Unit Root Test

นำข้อมูลที่จะทำการศึกษามาทดสอบว่ามีลักษณะที่นิ่งหรือไม่ โดยใช้การทดสอบ Unit Root โดยวิธี Dickey – Fuller test (DF) และ Augmented Dickey-Fuller test (ADF) ดังสมการต่อไปนี้

การทดสอบ DF (Dickey – Fuller test)

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \tag{3.2}$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t \tag{3.3}$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t \tag{3.4}$$

โดยกำหนดสมมติฐานหลัก $H_0: \theta = 0$

และสมมติฐานรอง $H_1: \theta < 0$

การทดสอบ DF (Dickey – Fuller test)

$$\Delta X_t = \theta X_{t-1} + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \varepsilon_t \tag{3.5}$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \varepsilon_t \tag{3.6}$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \varepsilon_t \tag{3.7}$$

โดยกำหนดสมมติฐานหลัก $H_0: \theta = 0$

และสมมติฐานรอง $H_1: \theta < 0$

โดย X_t คือ ราคาหลักทรัพย์ TUF, CPF, KSL, TVO ณ เวลาที่ t

X_{t-1} คือ ราคาหลักทรัพย์ TUF, CPF, KSL, TVO ณ เวลาที่ t-1

$\alpha, \beta, \theta, \phi$ คือ ค่าพารามิเตอร์

ε_t คือ ค่าความคลาดเคลื่อนเชิงสุ่มของราคาหลักทรัพย์

หากยอมรับสมมติฐานหลักแสดงว่าข้อมูลมีลักษณะที่ไม่นิ่ง (Non-Stationary) ให้ทำการทดสอบข้อมูลระดับผลต่างลำดับที่ 1 (1st difference) หากปฏิเสธสมมติฐานหลักแสดงว่าข้อมูลนั้นมีลักษณะนิ่ง

3.3.3 การทดสอบ Long memory

ขั้นตอนทดสอบ Long memory เป็นกระบวนการ Fractionally integrated เป็นการหาค่า mean ของ stationary หรือการหาค่ากลางระหว่าง I(0) และ I(1) โดยจะใช้วิธีการ Modified R/S test ซึ่งเป็นวิธีการทดสอบของ Lo ที่ปรับปรุงมาจาก Classical R/S test โดย Modified R/S statistic จะมีลักษณะดังสมการ

$$Q_{N,q} = \frac{1}{S_q}\left\{\max_{0\le i\le N}\sum_{j=1}^{i}\left(x_j-\bar{x}_N\right) - \min_{0\le i\le N}\sum_{j=1}^{i}\left(x_j-\bar{x}_N\right)\right\} \quad (3.8)$$

การมีอยู่ของระบบความจำระยะยาว (long memory) ค่าสถิติ $N^{-1/2}Q_{N,q}$ จะเบนเข้าหาช่วง Brownian Bridge

$$W = \max_{0\le r\le 1} V(r) - \min_{0\le r\le 1} V(r) \quad (3.9)$$

โดย V คือ Standard Brownian bridge

$$V(r) = B(r) - rB(1) \quad (3.10)$$

โดยที่ B คือ Standard Brownian motion

การกระจายของตัวแปรสุ่ม W รู้จักกันใน

$$P(W \le x) = 1 + 2\sum_{j=1}^{\infty}(1-4x^2j^2)e^{-2x^2j^2} \quad (3.11)$$

ค่า critical อยู่ในช่วงตามตารางที่ 3.1 Lo ใช้ช่วง [0.809,1.862] ที่ %95 (asymptotic) ยอมรับ null hypothesis

H_0 = { ไม่มี long memory, เช่น H = 0.5}

H_1 = { มี long memory, เช่น 0.5< H <1}

ตาราง 3.1 ค่า Critical Value ของ Modified R/S test

ระดับความน่าจะเป็น	Critical Value
0.5%	0.721
2.5%	0.809
5%	0.861
10%	0.927
90%	1.620
95%	1.747
97.5%	1.862
99.5%	2.098

ที่มา : Alptekin, N. 2006. Long Memory Analysis of USD/TRL Exchange Rat. International Journal of Human and Social Sciences.

3.3.4 การประมาณค่าพารามิเตอร์ด้วยแบบจำลอง ARIMA, ARFIMA

โดยแบบจำลอง ARIMA เป็นแบบจำลองที่พัฒนาขึ้นเพื่อใช้ในการพยากรณ์ค่าอนุกรมเวลา แต่ไม่สามารถใช้กับข้อมูลที่มีลักษณะของระบบความจำระยะยาว (long memory) โดยเขียนในรูปสมการดังนี้

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (3.12)$$

โดยที่ y_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลาที่ t

δ คือ ค่าคงที่

$\phi_1, \phi_2, \ldots, \phi_p$ คือ พารามิเตอร์ของ Autoregressive

$\theta_1, \theta_2, \ldots, \theta_q$ คือ พารามิเตอร์ของ Moving Average

P คือ อันดับของ Autoregressive

q คือ อันดับของ Moving Average

ε_t คือ ค่าความคาดเคลื่อน ณ เวลาที่ t

ส่วนแบบจำลอง ARFIMA ได้ถูกพัฒนามาจากแบบจำลอง ARIMA ใช้กับข้อมูลรายวัน ซึ่งสามารถนำไปใช้กับข้อมูลที่มีลักษณะของระบบความจำระยะยาว (long memory) ได้ แบบจำลอง ARFIMA จะมีประสิทธิภาพในการอธิบายได้ทั้ง short memory และ long memory โดยรูปแบบสมการจะเขียนเป็น ARFIMA (p, d, q)

$$\phi(L)(1-L)^d y_t = \delta + \theta(L)\varepsilon_t \quad (3.13)$$

และ

$$\phi(L) = 1 - \phi(L) - \phi_2(L)^2 - \ldots - \phi_p(L)^p \quad (3.14)$$

และ

$$\theta(L) = 1 - \theta(L) - \theta_2(L)^2 - \ldots - \theta_q(L)^q \quad (3.15)$$

โดย

δ = ค่าคงที่

$\theta(L)$ = moving average operator at order q

ε_t = error term

$\phi(L)$ = autoregressive operator at order p

L = lag operator

ถ้า d อยู่ที่ (0 ,0.5) ในแบบจำลอง ARFIMA อธิบายว่า มีลักษณะเป็น Long memory

ถ้า d อยู่ที่ (-0.5 ,0) ในแบบจำลอง ARFIMA อธิบายว่า มีลักษณะเป็น Short memory (Hosking, 1981)

3.3.5 การประมาณค่าความผันผวนด้วยแบบจำลอง GARCH, EGARCH, FIGARCH, FIEGARCH

แบบจำลอง GARCH นำเสนอโดย Bollerslev (1986) เป็นแบบจำลองความผันผวนโดยมีลักษณะสมการดังนี้

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j} \quad (3.16)$$

โดยที่ค่าสัมประสิทธิ์ $\alpha_i(i = 0, \ldots, p)$ และ $\beta_j(j = 0, \ldots, q)$ ทั้งหมดถูกสมมติว่าเป็นค่าบวก เพื่อที่ค่าความแปรปรวนอย่างมีเงื่อนไข (conditional variance) h_t จะได้มีค่าเป็นบวกเสมอ โดยแบบจำลองนี่คือ Generalized Autoregressive Conditional Heteroskedasticity หรือ GARCH(q,p)

แบบจำลอง EGARCH ถูกนำเสนอโดย Nelson (1991) ต่อมา Bollerslev และ Mikkelsen (1996) ได้นำมานำเสนอใหม่ โดยรูปแบบสมการ EGARCH เขียนได้ดังนี้

$$\ln(h_t) = \alpha_0 + [1 - \varphi(L)]^{-1}[1 + \psi(L)]g(z_{t-1}) \qquad (3.17)$$

โดยที่ $g(z_t) = \theta z_t + \gamma[|z_t| - E(|z_t|)]$

$z_t = \varepsilon_t/\sigma_t$

$\varphi(L)$ คือ พหุนามค่าล่าลำดับที่ p

$\psi(L)$ คือ พหุนามค่าล่าลำดับที่ q

θ คือ ผลกระทบทางเครื่องหมาย

γ คือ ผลกระทบทางขนาด

θ ทำให้มีความไม่สมมาตรของข้อมูลข่าวสารหรือผลกระทบที่มีต่อความผันผวน ถ้า $\theta < 0$ ข่าวร้ายหรือการเปลี่ยนแปลงที่ไม่คาดการณ์ทางลบ (negative shock) จะส่งผลกระทบความผันผวนมากกว่าข่าวสารด้านบวกหรือข่าวดี และ γ แสดงถึงอัตราที่ข้อมูลข่าวสารที่มากระทบต่อความผันผวนเบี่ยงเบนไปจากค่าเฉลี่ย

แบบจำลอง FIGARCH ถูกสร้างโดย Baillie (1996) โดยตั้งอยู่บนรูปแบบของ GARCH และสมมติให้เกิด heteroscedasticity ในความแปรปรวนในเวลาที่แตกต่างกันของข้อมูล แบบจำลอง FIGARCH มาจากแบบจำลอง GARCH และ IGARCH เพื่อใช้กับ long memory กระบวนการ GARCH (p, q) สามารถจำลองเป็นกระบวนการ ARMA โดยใช้ lag operator

$$h_t = \alpha_0[1 - \beta(L)]^{-1} + [1 - [1 - \beta(L)]^{-1}\phi(L)(1 - L)^d]\varepsilon_t^2 \qquad (3.18)$$

$(1\text{-}L)^d$: fractional difference operator

แบบจำลอง FIGARCH (p, d, q) ถูกแปลงมาจาก GARCH มาตรฐาน d = 0 และ IGARCH model เมื่อ d = 1.

แบบจำลอง FIEGARCH สร้างขึ้นเนื่องจากความไม่สมมาตรระหว่างความผิดปกติทางบวกและลบ (leverage effect) Bollerslev และ Mikkelsen (1991) ได้ขยายแบบจำลอง FIGARCH เป็น FIEGARCH โดยลักษณะทั่วไปของ FIEGARCH(p,d,q) คือ

$$\ln(h_t) = \omega + \phi(L)^{-1}(1-L)^{-d}[1+\psi(L)]g(z_{t-1}) \tag{3.19}$$

โดย $\omega = \alpha_0[1-\beta(L)]^{-1}$

เมื่อ d = 0 กระบวนการ FIEGARCH จะลดรูปเป็น EGARCH และเมื่อ d = 1 กระบวนการ FIEGARCH จะกลายเป็น integrated EGARCH (IEGARCH)

3.3.6 การหาค่า Akaike Information criteria (AIC), Schwartz Information criteria (SC)

เพื่อเลือกรูปแบบของแบบจำลองที่ดีที่สุด โดยพิจารณาจากค่า Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) รูปแบบจำลองที่ให้ค่า AIC และ SC น้อยที่สุดจะเป็นรูปแบบที่ดีที่สุด โดย Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) สามารถคำนวณได้ดังนี้

Akaike Information Criterion (AIC) $-2l/n + 2k/n$

Schwartz Criterion (SC) $-2l/n + k\log n/n$

โดยที่ k เป็นจำนวนพารามิเตอร์ที่ทำการประมาณค่า

n เป็นจำนวนของค่าสังเกต

l เป็นค่าของ Log Likelihood function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว

3.3.7 ทดสอบค่าความคลาดเคลื่อนร้อยละเฉลี่ย (The Mean Absolute Percentage Error : MAPE)

เพื่อวัดความแม่นยำของแบบจำลองที่ได้ ซึ่งสูตรของ MAPE คือ

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

เมื่อให้ A_t คือค่าที่แท้จริง และ F_t คือค่าที่คาดการไว้

ค่าสัมบูรณ์ที่ได้จากการคำนวณนี้ คือผลบวกของทุกๆจุดที่เหมาะสมหรือทุกๆจุดที่คาดการณ์ไว้ในล้วงเวลาและได้แสดงให้เห็นอีกครั้งด้วยจำนวนที่เหมาะสม n ทำให้เกิดข้อผิดพลาดด้านเปอร์เซ็นต์ ดังนั้นจึงสามารถเปรียบเทียบข้อผิดพลาดของ Time series ที่เหมาะสมซึ่งมีความต่างกัน