บทที่ 4
ระเบียบวิธีการศึกษา

ในการศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาถึงมีอยู่ของ long memory ในข้อมูลราคาน้ำมันดิบไลท์สวีทและน้ำมันเบนซินรายวัน และเพื่อหาแบบจำลอง ARFIMA ที่เหมาะสมเพื่อพยากรณ์การเปลี่ยนแปลงของราคาน้ำมันดิบไลท์สวีทและน้ำมันเบนซินรายวันในอนาคต

4.1 ข้อมูลที่ใช้ในการศึกษา

การศึกษาในครั้งนี้จะใช้ข้อมูลอนุกรมเวลาของราคาน้ำมัน 2 ประเภทคือ น้ำมันดิบไลท์สวีท (Light Sweet) และ น้ำมันเบนซิน(Gasoline) ที่ได้มีกระบวนการซื้อขายในตลาดล่วงหน้า ด้วยสัญญาส่งมอบเดือนที่ 3 , 6 , 9 , และ 12 ในตลาด NYMEX (New York Mercantile Exchange) ประเทศสหรัฐอเมริกา ซึ่งเป็นข้อมูลประเภททุติยภูมิแบบรายวัน โดยราคาน้ำมันดิบไลท์สวีท เริ่มตั้งแต่วันที่ 30 มีนาคม 2526 ถึง 31 มกราคม 2552 รวมจำนวนค่าสังเกตุ 6,488 ค่า และราคาน้ำมันเบนซิน เริ่มตั้งแต่วันที่ 10 มีนาคม 2548 ถึง 31 มกราคม 2552 รวมจำนวนค่าสังเกตุ 838 ค่า ซึ่งข้อมูลจะอ้างอิงจากฐานข้อมูล Reuters(2007) ในศูนย์การเงินการลงทุน มหาวิทยาลัยเชียงใหม่

4.2 แบบจำลองที่ใช้ในการศึกษา

ในการประมาณค่าตัวแปรและการพยากรณ์จะใช้แบบจำลอง ARFIMA จากสมการ (2.11) ประยุกต์ให้เป็น 2 สมการ คือ

$$\phi(L)(1-L)^{\delta}[(1-L)^{m}y_{it}-\mu]=\theta(L)\in_{it} \qquad (2.11)$$

เมื่อ $-0.5 < d < 0.5$ นั่นคือพารามิเตอร์ผลต่างคือ $d = \delta + m$ โดยที่

1.) ถ้า $-0.5 < d < 0$ แล้วข้อมูลจะมีลักษณะเป็น short memory และ stationary

2.) ถ้า $0 < d < 0.5$ แล้วข้อมูลจะมีลักษณะเป็น long memory และ stationary

โดยเมื่อ y_i คือ ค่าสังเกตหรือราคาน้ำมัน

i คือ 1 และ 2

y_{1t} คือ ราคาน้ำมันดิบไลท์สวีท(Light Sweet)รายวัน ลำดับที่ t

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

y_{2t} คือ ราคาน้ำมันเบนซิน(Gasoline)รายวัน ลำดับที่ t

$\in_{1t}$ คือ ค่าความคลาดเคลื่อนราคาน้ำมันดิบไลท์สวีท(Light Sweet)รายวัน ลำดับที่ t

$\in_{2t}$ คือ ค่าความคลาดเคลื่อนราคาน้ำมันเบนซินรายวัน(Gasoline) ลำดับที่ t

t คือ ลำดับเวลาของค่าสังเกตที่ 1 , 2 , 3 , …,T โดยที่ T คือ ลำดับค่าสังเกตสุดท้ายของข้อมูล

μ คือ ค่าคงที่

d คือ degree of differencing

δ คือ fractional

m คือ จำนวนครั้งของการทำผลต่าง y_t

ϕ(L) และ θ (L) คือ โพลีโนเมียลของค่าล่าหรือค่าล้าหลัง (lag polynomial)

4.3 วิธีการวิเคราะห์ข้อมูล

การทดสอบการมี long memory และ ทดสอบความสามารถในการพยากรณ์ราคาน้ำมันโดยใช้เทคนิคของแบบจำลอง ARFIMA ซึ่งขั้นตอนของการศึกษามีดังนี้

1) ทำการทดสอบทางสถิติของ Long Memory เพื่อเป็นการทดสอบแบบคร่าวๆ โดยย่อถึงการมี long memory ในข้อมูลอนุกรมเวลาราคาน้ำมันที่ได้นำมาทดสอบ ประกอบด้วยการทดสอบ 2 วิธี ดังนี้

1.1) R/S Statistic จากสมการ (2.16) คือ

$$Q_T = \frac{1}{s_T}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(y_j - \bar{y}) - \min_{1\le k\le T}\sum_{j=1}^{k}\left(y_j - \bar{y}\right)\right] \qquad (2.16)$$

และ Modified R/S Statistic จากสมการ (2.17) คือ

$$\widetilde{Q}_T = \frac{1}{\hat{\sigma}_T(q)}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(y_j - \bar{y}) - \min_{1\le k\le T}\sum_{j=1}^{k}(y_j - \bar{y})\right] \qquad (2.17)$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ซึ่ง modified R/S Statistic จะถูกคำนวณ และ banwidth(q) จะได้ความแปรปรวนระยะยาวมาแล้วถูกเลือกให้เป็น $4(T/100)^{1/4}$ ถ้าไม่มี long memory แล้วค่า R/S statistic จะลู่เข้าสู่ตัวแปรสุ่มด้วยอัตรา $T^{1/2}$ แต่ถ้ามี long memory แล้วค่า Modified R/S statistic จะลู่เข้าสู่ตัวแปรสุ่มด้วยอัตรา T^{H} เมื่อ H คือ Hurst coefficient

โดยอยู่ภายใต้สมมุติฐานคือ H_0: ไม่มี long memory

H_1 : มี long memory

1.2) การทดสอบโดยใช้ GPH Test จากสมการ (3.19) เพื่อการทดสอบ long memory โดยการประมาณค่า d (fractional difference parameter) ซึ่งจะทำการถดถอยสมการด้วยวิธีการ periodogram ของค่า $f(\omega_j)$

$$\ln f(\omega_j) = \beta - d \ln[4\sin^2(\frac{\omega_j}{2})] + e_j \qquad (2.19)$$

เมื่อ $j = 1, 2, \ldots, n_f(T)$ หรือ $n_f(T)$ คือ จำนวนความถี่ และทำการคำนวณค่า t-statistic จากสมการ(3.20) ซึ่งมีการกระจายแบบปกติมาตรฐาน คือ

$$t_{d=0} = \hat{d}.\left(\frac{\pi^2}{6\sum_{j=1}^{n_f}(U_j - \bar{U})^2}\right)^{-1/2} \qquad (2.20)$$

เมื่อ $\bar{U}$ คือ ค่าเฉลี่ยตัวอย่างของ U_j โดย $j = 1, 2, \ldots, n_f$

โดยอยู่ภายใต้สมมุติฐานคือ H_0: ไม่มี long memory

H_1 : มี long memory

2) ทำการประมาณค่า Long Memory Parameter เพื่อประมาณค่าพารามิเตอร์ H และ d ดังนี้

2.1) วิเคราะห์ด้วยวิธีคำนวณค่าจาก Modified R/S statistic โดยการใช้ k_i หรือ ค่าสังเกตต่างๆ ตามลำดับในตัวอย่าง เมื่อ k_1 ควรใหญ่เพียงพอ จากนั้นเพิ่มจำนวนของค่าสังเกตโดยตัวแปร f ซึ่งเป็นการคำนวณ R/S statistic โดยการใช้ $k_i = fk_{i-1}$ ค่าสังเกตต่างๆ ตามลำดับ เมื่อ $i = 2,..., s$ เส้นที่เหมาะสมของ Modified R/S statistic เหล่านั้นเทียบกับ k_i , $i = 1,..., s$ และได้ผลการประมาณของสัมประสิทธิ์ Hurst (H) เป็นเส้นที่เหมาะสม(fitted line)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.2) วิเคราะห์ด้วยวิธี Periodogram คือ การหาค่า d จากการถดถอยสมการ (3.19) และคำนวณค่า t-statistic จากสมการ (3.20)

3) การสร้างแบบจำลอง ARFIMA(p,d,q) ในอนุกรมเวลาที่มีความนิ่ง(stationary) โดยโดยประมาณค่าด้วยวิธีการ Maximum Likelihood ถ้าหากข้อมูลมีความนิ่ง(stationary) จากการทดสอบแบบจำลองที่ p และ q เท่ากับศูนย์แล้ว จะกำหนดค่า m เท่ากับศูนย์ ส่วนการประมาณค่า p และ q สามารถหาได้จากค่า $p \leq p_{\max}$ และ $q \leq q_{\max}$ ซึ่งได้จากการเลือกค่า Baysian Information Criterior (BIC) ต่ำสุด

4) ทำการพยากรณ์แบบจำลองด้วยข้อมูลอนุกรมเวลาราคาน้ำมันและเปรียบเทียบกับข้อมูลจริงเพื่อทดสอบความแม่นยำของแบบจำลอง ARFIMA ที่ได้สร้างขึ้นมา

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved