บทที่ 3
ระเบียบวิธีวิจัย

3.1 ระเบียบวิธีวิจัย

ในการศึกษาครั้งนี้ จะทำการพยากรณ์มูลค่าการส่งออกกาแฟดิบแบบรายเดือนตั้งแต่เดือนมกราคม พ.ศ. 2538 ถึงเดือนมกราคม พ.ศ. 2550 จำนวนทั้งหมด 145 เดือน ซึ่งจะใช้วิธีการพรรณนาอธิบาย และวิธีการวิเคราะห์เชิงปริมาณ โดยกำหนดรูปแบบจำลองให้กับอนุกรมเวลาในรูปแบบ ARIMA โดยวิธีของ Box-Jenkins เนื่องจากเป็นวิธีที่จะให้ค่าพยากรณ์ที่มีความแม่นยำโดยเฉพาะการพยากรณ์ในระยะสั้น โดยมีขั้นตอนดังต่อไปนี้

1) การแปลงข้อมูลให้อยู่ในรูปลอการิทึมฐานธรรมชาติ (natural logarithm) เนื่องจากข้อมูลอนุกรมเวลามูลค่าการส่งออกกาแฟดิบมีลักษณะแปรปรวนมาก เพื่อให้ค่าความแปรปรวนของอนุกรมเวลาคงที่สำหรับค่า t ต่างๆ ดังนั้นจึงต้องแปลงข้อมูลอนุกรมเวลาเดิมให้เป็นอนุกรมเวลาใหม่ซึ่งอยู่ในรูปลอการิทึมฐานธรรมชาติ ทั้งนี้ในการศึกษาจะใช้ตัวแปร Y แทนมูลค่าการส่งออกกาแฟดิบรูปลอการิทึมฐานธรรมชาติ

2) การทดสอบความนิ่งของข้อมูล เป็นการพิจารณาว่าข้อมูลอนุกรมเวลานั้น มีลักษณะนิ่ง (stationary) หรือไม่ ด้วยการทดสอบ unit root โดยใช้ ADF-test ดังแสดงในสมการต่อไปนี้

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_1 \Delta X_{t-1} + \varepsilon_t \tag{3.1}$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{\rho} \phi \Delta X_{t-1} + \varepsilon_t \tag{3.2}$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \sum_{i=1}^{\rho} \phi \Delta X_{t-1} + \varepsilon_t \tag{3.3}$$

โดยเปรียบเทียบค่าสถิติ ADF-test ที่คำนวณได้กับค่าวิกฤต Mackinnon ซึ่งมีสมมติฐานดังนี้

$H_0 : \theta = 0$ คือข้อมูลมี unit root หรือมีลักษณะไม่นิ่ง (nonstationary) ซึ่งต้องทำการหาผลต่าง (differencing) ในอันดับต่อไป

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$H_0 : \theta < 0$ คือข้อมูลปราศจาก unit root หรือมีลักษณะนิ่ง (stationary) ณ อันดับนั้นที่ทำการทดสอบ

ในการหา lag length นั้นสำหรับในโปรแกรม Eviews 5.1 ที่ใช้ในการศึกษาครั้งนี้จะสามารถทำการเลือกให้โดยอัตโนมัติ ซึ่งพิจารณาจากค่า Schwarz criterion ที่ต่ำที่สุดและเป็นการเลือก lag length ที่ทำให้ไม่เกิดปัญหาอัตตสัมพันธ์ (Autocorrelation)

หลังจากนั้นจึงทำการทดสอบความนิ่งของข้อมูลอนุกรมเวลาแบบฤดูกาล ซึ่งมีรูปแบบสมการดังนี้

$$\begin{aligned} y_{8,t} &= \pi_1 y_{2,t} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \pi_5 y_{4,t-2} \\ &+ \pi_6 y_{4,t-1} + \pi_7 y_{5,t-2} + \pi_8 y_{5,t-1} + \pi_9 y_{6,t-2} + \pi_{10} y_{6,t-1} \\ &+ \pi_{11} y_{7,t-2} + \pi_{12} y_{7,t-1} + \mu_t + \varepsilon_t \end{aligned} \tag{3.4}$$

โดยจะทำการทดสอบ t-test สำหรับค่า π_1 และค่า π_2 หลังจากนั้นจึงทำการทดสอบ F-test ต่อไป ดังแสดงในตารางที่ 2.1 เมื่อพบว่าข้อมูลมีลักษณะนิ่งแล้วก็จะสามารถนำข้อมูลอนุกรมเวลานั้นมาใช้ในการหารูปแบบ ARIMA(p,d,q) ด้วยวิธีของ Box-Jenkins ต่อไป

3) การกำหนดรูปแบบจำลอง ARIMA (p,d,q) โดยการพิจารณาคอเรลโลแกรม Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) เพื่อจะสามารถระบุได้ว่าแบบจำลองควรจะมี Autoregressive (q) เท่าใด และ Moving average (q) เท่าใด โดยเลือกสร้างไว้หลายแบบจำลองเพื่อหาแบบจำลองที่เหมาะสมที่สุดในการพยากรณ์ ซึ่งมีการกำหนดรูปแบบของ ARIMA(p,d,q) ไว้ดังนี้

$$(1-L)^n (1-L)^s \mu_t = (1-L)^n (1-L)^s \varepsilon_t$$

ในการศึกษาครั้งนี้จะพิจารณารูปแบบ ARIMA(p,d,q) โดยใช้ค่าสถิติเพื่อประกอบการตัดสินใจดังนี้ Akaike info criterion : AIC, Schwarz's Bayeian information criterion : SBC, Adjusted R-squared, Durbin-Watson stat, และค่า F-statistic

4) การประมาณค่าพารามิเตอร์ (Estimation) คือการนำรูปแบบ ARIMA (p,d,q) ที่เลือกจากขั้นตอนการเลือกแบบจำลองที่เหมาะสม โดยพิจารณาความมีนัยสำคัญทางสถิติ ด้วย T – statistic เพื่อนำค่าพารามิเตอร์ที่ได้นั้นนำไปทำการพยากรณ์ มูลค่าการส่งออกต่อไปได้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

5) การตรวจสอบความถูกต้อง (Diagnostic checking) เมื่อทำการหาแบบจำลองที่เหมาะสมและประมาณค่าพารามิเตอร์แล้ว จึงทำการทดสอบแบบจำลองโดยพิจารณาจากค่า Q-statistic จากคอเรลโลแกรม ดังแสดงในสมการที่ 2.22

6) การพยากรณ์ (Forecasting) คือ การพยากรณ์ข้อมูลล่วงหน้า โดยในการศึกษาครั้งนี้จะแบ่งช่วงการพยากรณ์ดังนี้

- Historical forecast เป็นการพยากรณ์ตั้งแต่อดีตจนถึงช่วงเวลาที่พิจารณา คือตั้งแต่ช่วงเวลาที่ 1 ถึง 141 โดยการถดถอยข้อมูลใหม่และทำการพยากรณ์ พิจารณาค่าสถิติ Root Mean Square Error (RMSE) และ Theil's inequality coefficient (U) และค่า Akaike Information Criterion (AIC) ที่ต่ำที่สุด

- Expost forecast เพื่อเปรียบเทียบว่าแบบจำลองใดจะมีความสามารถในการพยากรณ์ที่ดีที่สุด โดยการลดจำนวนข้อมูลลง 4 ค่า จาก 145 ค่าสังเกต เหลือ 141 ค่าสังเกต แล้วทำการถดถอยข้อมูลใหม่ และนำข้อมูลที่พยากรณ์ได้เปรียบเทียบกับข้อมูลจริง โดยพิจารณาจากค่าสถิติ Root Mean Square Error (RMSE) และ Theil's inequality coefficient (U) และค่า Akaike Information Criterion (AIC) ที่ต่ำที่สุด

- Ex-ante forecast เมื่อทราบแบบจำลองที่สามารถพยากรณ์ได้ดีที่สุด แล้วจึงนำเอาแบบจำลองนั้นไปพยากรณ์ในช่วงเวลาถัดไปอีก 4 คาบเวลา กล่าวคือการพยากรณ์ ณ ช่วงเวลาที่ 146-149

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved