บทที่ 3

แนวความคิดและระเบียบวิธีวิจัย

3.1 แนวคิดการพยากรณ์ข้อมูลอนุกรมเวลา และวิธีการพยากรณ์โดยวิธี Box-Jenkins

การศึกษาอนุกรมเวลาของราคารายเดือนสินค้ากุ้งกุลาดำ ณ ตลาดระดับที่เกษตรกรขายได้ โดยการกำหนดรูปแบบให้กับอนุกรมเวลา เพื่อพิจารณาการเคลื่อนไหวราคาของกุ้งกุลาดำ ว่ามีอิทธิพลของแนวโน้ม และฤดูกาลเข้ามาเกี่ยวข้องกับราคาสินค้ากุ้งกุลาดำที่ทำการศึกษาหรือไม่ พร้อมทั้งนำรูปแบบของสมการที่ได้ทำการวิเคราะห์อนุกรมเวลาแล้ว นำมาพยากรณ์ราคาสินค้าเกษตรที่ทำการศึกษาไปล่วงหน้า เพื่อนำค่าพยากรณ์ดังกล่าวมาใช้คาดคะเนราคาสินค้าเกษตรซึ่งจะเป็นประโยชน์ในการปรับตัว ในการจัดสรรทรัพยากร การผลิต และการตลาดเกษตรกรที่เลี้ยงกุ้งกุลาดำ ให้เป็นไปตามการจัดสรรที่เหมาะสม โดยใช้ราคาที่พยากรณ์เป็นสิ่งชี้นำในการจัดสรรทรัพยากรดังกล่าวแทนการใช้ราคาสินค้าเกษตรในปีก่อนมาจัดสรรทรัพยากร (Dickey and Fuller, 1984)

โดยในการศึกษาครั้งนี้ จะใช้การวิเคราะห์อนุกรมเวลาราคาสินค้ารายเดือน โดยวิธี Box–Jenkins เพราะเป็นวิธีที่จะให้ค่าพยากรณ์ที่มีความถูกต้องสูง โดยเฉพาะในการพยากรณ์ระยะสั้น ซึ่งก่อนทำการพยากรณ์ควรทราบก่อนว่า ระยะเวลาการพยากรณ์ของวิธี Box–Jenkins จะแม่นยำในช่วงระยะเวลาระหว่างหนึ่งสัปดาห์ถึงสามสัปดาห์ หรือหนึ่งเดือนถึงสามเดือน หากต้องการจะใช้พยากรณ์ช่วงเวลาที่ยาวนานกว่านี้ ควรนำข้อมูลที่ทันสมัยมาปรับค่าพยากรณ์ที่ได้ทำไว้แล้ว (Update) เพื่อให้ค่าความคลาดเคลื่อนจากการพยากรณ์ลดลง

ดังนั้น ข้อมูลอนุกรมเวลา (Time-series Data) คือค่าสังเกต (Observation) ชุดหนึ่งซึ่งถูกกำหนดขึ้น ณ เวลาต่างๆ ถ้าค่าสังเกตกระทำในเวลาที่ต่อเนื่องกันจะเรียกว่า ข้อมูลอนุกรมเวลาต่อเนื่อง แต่ถ้าค่าสังเกตกระทำ ณ จุดเวลาที่ไม่ต่อเนื่องกัน เรียกว่า ข้อมูลอนุกรมเวลาไม่ต่อเนื่อง ดังนั้นการวิเคราะห์ข้อมูลอนุกรมเวลาจึงเป็นการวิเคราะห์ค่าสังเกตที่มีการเปลี่ยนแปลงไปตามเวลาที่กระทำ

Enders (1995) กล่าวว่าการพยากรณ์ข้อมูลอนุกรมเวลาต้องอยู่บนข้อสมมติว่า ข้อมูลอนุกรมเวลาต้องมีลักษณะนิ่ง (Stationary) ซึ่งสามารถพิจารณาความนิ่งของข้อมูลได้จาก

ค่าเฉลี่ย:	$E(X_t)$	$= \mu$ = constant	(1)
ค่าความแปรปรวน:	$V(X_t)$	$= \sigma^2$ = constant	(2)
ค่าความแปรปรวนร่วม:	$cov(X_t, X_{t+k})$	$= E(X_t - \mu)(X_{t+k} - \mu)$	(3)

เพราะฉะนั้น ข้อมูลที่มีลักษณะนิ่งจะต้องมีค่าเฉลี่ย ค่าความแปรปรวน ของทุกๆ ค่า ณ เวลา t ใดๆ คงที่ ในขณะที่ความแปรปรวนร่วมระหว่างสองคาบเวลาจะขึ้นอยู่กับช่องว่างระหว่างสองคาบเวลาเท่านั้น ไม่ขึ้นอยู่กับเวลาที่เปลี่ยนไป หากเงื่อนไขใดเงื่อนไขหนึ่งไม่เป็นดังที่กล่าวมา แสดงว่าข้อมูลมีลักษณะไม่นิ่ง (Nonstationary) (Charemza and Deadman, 1992)

ดังนั้น จึงสรุปได้ว่า ข้อมูลจะมีลักษณะนิ่งจะต้องมีค่าเฉลี่ย ค่าความแปรปรวนร่วม คงที่ ทุกๆ เวลา t ใดๆ

3.1.1 การทดสอบความนิ่งของข้อมูลอนุกรมเวลา (Unit Root Test)

การพยากรณ์ข้อมูลอนุกรมเวลาที่ให้ความแม่นยำนั้น จะต้องมีการทดสอบว่าข้อมูลนั้น ค่าเฉลี่ยและค่าความแปรปรวนของตัวแปรมีค่าคงที่หรือไม่ ก็คือการทดสอบว่าข้อมูลมีลักษณะนิ่งหรือไม่ เรียกว่า การทดสอบ Unit Root นั้นสามารถทดสอบได้โดยใช้ Dickey–Fuller test (DF-test) (Dickey and Fuller, 1981) และ Augmented Dickey–Fuller test (ADF-test) (Said and Dickey, 1984) โดยจะตั้งสมมุติฐานว่าง (Null Hypothesis) ของการทดสอบคือ $H_0 : \rho = 1$ การทดสอบเช่นนี้เป็นการทดสอบ DF (Dickey and Fuller, 1981) และการทดสอบ ADF (Said and Dickey, 1984) ดังสมการ

H_0: $\rho = 1$ จากสมการ (4) ด้านล่าง

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{4}$$

ซึ่งเรียกว่า Unit Root Test โดยที่ถ้า $|\rho| < 1$ X_t จะมีลักษณะนิ่ง (Stationary) และถ้า $\rho = 1$ X_t จะมีลักษณะไม่นิ่ง (Nonstationary) อย่างไรก็ตามการทดสอบนี้สามารถทำได้อีกทางหนึ่งซึ่งเหมือนกับสมการ (4)

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \tag{5}$$

ซึ่งก็คือ $X_t = (1+\theta)X_{t-1} + \varepsilon_t$ ซึ่งคือสมการ (4) นั่นเอง โดยที่ $\rho = (1+\theta)$
ถ้า θ ในสมการ (5) มีค่าเป็นลบ จะได้ว่า ρ ในสมการ (4) จะมีค่าน้อยกว่า 1
ดังนั้น สามารถจะสรุปได้ว่า การปฏิเสธ H_0: $\theta = 0$ ซึ่งเป็นการยอมรับ H_a: $\theta < 0$ หมายความว่า $\rho < 1$ และ X_t มี Integration of Order Zero (Charemza and Deadman, 1992) นั่นคือ X_t เป็น Stationary และถ้าเรายอมรับ H_0: $\theta = 0$ ได้ ก็จะหมายความว่า X_t เป็น Non Stationary ถ้า X_t มีแนวคิดเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (Random Walk with Drift) เราสามารถจะเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t \tag{6}$$

และถ้า X_t เป็น Random Walk with Drift และมีแนวโน้มตามเวลาเชิงเส้น (Linear Time Trend) เราสามารถจะเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t \tag{7}$$

โดยที่ t = แนวโน้มของเวลา

ซึ่งการพิจารณาความนิ่งยังคงให้ความสำคัญต่อ θ ดังเช่นการพิจารณา สมการ (5) คือมีสมมติฐานว่าง คือ $H_0 : \theta = 0$ และ $H_a : \theta < 0$ ซึ่งจะเปรียบเทียบค่าสถิติ t (Test-statistic) ที่คำนวณกับค่าวิกฤติที่อยู่ในตาราง Dickey – Fuller (Enders, 1995) หรือกับค่าวิกฤติ MacKinnon (MacKinnon Critical Values) (Gujarati, 2003)

เมื่อมีการแทนที่ของกระบวนการเชิงถดถอย (Autoregressive Process) ลงในสมการ (5) (6) และ (7) จะไม่ทำให้ค่าวิกฤติเปลี่ยนแปลงไป ดังสมการต่อไปนี้

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{p} \phi_i \Delta X_{t-i} + \varepsilon_t \tag{8}$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{p} \phi_i \Delta X_{t-i} + \varepsilon_t \tag{9}$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \sum_{i=1}^{p} \phi_i \Delta X_{t-i} + \varepsilon_t \tag{10}$$

ซึ่งเรียกว่าการทดสอบ ADF โดยที่จำนวนของ Logged Difference Terms ที่จะนำมาใช้ในสมการที่นี้มีมากพอที่จะทำให้พจน์ค่าความคลาดเคลื่อนมีลักษณะเป็น Serially Independent ซึ่งการทดสอบจะใช้ค่าสถิติทดสอบ ADF (ADF Test-statistic) และมีการแจกแจงเชิงเส้นกำกับ ดังนั้นจึงสามารถใช้ค่าวิกฤติแบบเดียวกับการทดสอบ DF (Gujarati, 2003)

สำหรับการเลือก lag length ของWalter Enders (1995) กล่าวว่าควรจะเริ่ม lag length ที่มีค่าที่มากพอ แล้วพิจารณาความมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญต่างๆ (α = 0.01, 0.05 และ0.1) เมื่อพบว่าที่ lag length ที่เลือกมีค่า Test-statistic ที่ไม่มีนัยสำคัญทางสถิติ ที่ระดับ 0.1 แล้ว จึงทำการลด lag length ลงทีละ 1 ช่วง จนกระทั่งทำการปฏิเสธสมมติฐานว่าง คือ ค่า Test-statistic มีนัยสำคัญทางสถิติ

3.1.2 การพยากรณ์โดยวิธี Box-Jenkins

วิธีการวิเคราะห์อนุกรมเวลาของ Box–Jenkins เป็นการวิเคราะห์อนุกรมเวลาโดยการหารูปแบบที่เหมาะสมให้กับอนุกรมเวลา โดยใช้ค่า Autocorrelation Function (ACF) และค่า Partial Autocorrelation Function (PACF) เป็นหลักในการพิจารณา และรูปแบบที่เลือกใช้จะอยู่ในกลุ่มของรูปแบบ ARIMA (p, d, q) หรือเรียก Integrated Autoregressive–Moving Average order p and q ซึ่งเป็นรูปแบบที่กำหนดว่าค่าพยากรณ์ในอนาคตเป็นค่าที่ได้จากการสังเกต หรือค่าพยากรณ์ก่อนหน้า และ ความคลาดเคลื่อนของการพยากรณ์ก่อนหน้า โดยเป็นการรวมส่วนของรูปแบบ AR(p) และรูปแบบ MA(q) เข้าด้วยกัน รูปแบบ AR(p) หมายถึงรูปแบบที่แสดงว่าค่าสังเกต Y_t จะขึ้นอยู่กับค่า $Y_{t-1}, Y_{t-2},\ldots,Y_{t-p}$ หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ MA(q) หมายถึงรูปแบบที่แสดงว่าค่าสังเกต Y_t จะขึ้นอยู่กับค่าของความคลาดเคลื่อน $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots,\varepsilon_{t-q}$

หรือค่าความคลาดเคลื่อนที่เกิดขึ้นก่อนหน้า q ค่าซึ่งรูปแบบ ARIMA (p, d, q) โดยมีการกำหนดรูปแบบดังนี้

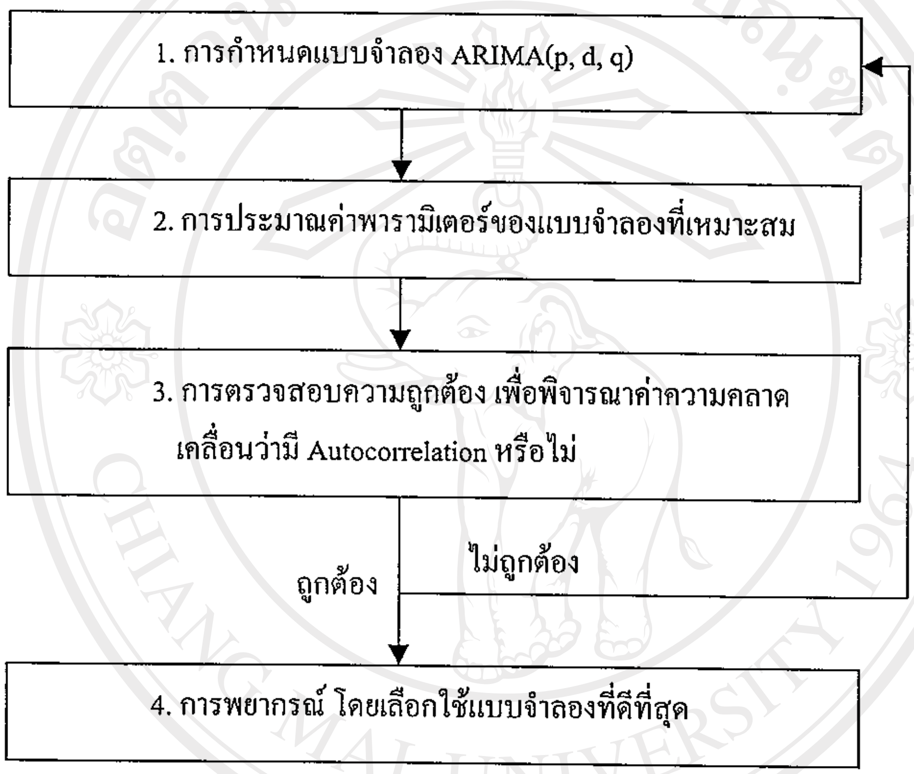
AR(p) คือ $Y_t = \theta_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$

MA(q) คือ $Y_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \ldots - \theta_q\varepsilon_{t-q}$

AR MA(p, q) คือ $Y_t = \theta_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \ldots - \theta_q\varepsilon_{t-q}$

ARIMA(p, d, q) คือ $\Delta^d Y_t = \theta_0 + \phi_1\Delta^d Y_{t-1} + \ldots + \phi_1\Delta^d Y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \ldots - \theta_q\varepsilon_{t-q}$

การพยากรณ์โดยใช้วิธี Box-Jenkins ประกอบด้วย 4 ขั้นตอน คือขั้นตอนที่ 1 การกำหนดแบบจำลอง (Identification) ขั้นตอนที่ 2 การประมาณค่าพารามิเตอร์ในรูปแบบอนุกรมเวลา (Estimation) ขั้นตอนที่ 3 การตรวจสอบความถูกต้อง (Diagnostics Checking) และขั้นตอนที่ 4 การพยากรณ์ (Forecasting) ดังรูป 3.1



รูป 3.1ขั้นตอนการพยากรณ์โดยวิธี Box-Jenkins
ที่มา: Gujarati (2003)

1) การกำหนดแบบจำลอง (Identification)

ให้กับอนุกรมเวลาที่เป็น Stationary Series เป็นการหารูปแบบ ARIMA (p, d, q) ที่คาดว่าจะเหมาะสมให้กับอนุกรมเวลาโดยที่ Autocorrelation: p_k คือการวัดความสัมพันธ์ของแต่ละช่วงเวลา โดยมีช่วงเวลาที่ย้อนหลังไป k หน่วยเวลา โดยที่ p_k มีค่าเท่ากับ $-1 \leq p_k \leq 1$ โดยพิจารณาเปรียบเทียบค่า Autocorrelation (r_k) ของอนุกรมเวลาตัวอย่าง กับค่า Autocorrelation (p_k) ของอนุกรมเวลาของประชากร ที่มีช่วงเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตรดังนี้

$$r_k = \frac{\sum_{t=a}^{n-k}(Y_{t-q})(Y_{t+k-q})}{\sum_{t=a}^{n}(Y_{t-q})^2} \tag{11}$$

โดยที่ $Y_t = \sum_{t=a}^{n}(Y_t)$

q = จำนวนเวลาสุดท้ายที่ย้อนหลัง

Partial Autocorrelation: p_{kk} คือการวัดความสัมพันธ์ของแต่ละช่วงเวลา โดยมีช่วงเวลาที่ย้อนหลังไป k หน่วยเวลา โดยพิจารณาเปรียบเทียบค่า Partial Autocorrelation (r_{kk}) ของอนุกรมเวลาตัวอย่าง กับค่า Partial Autocorrelation (p_{kk}) ของอนุกรมเวลาของประชากร ที่มีช่วงเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตรดังนี้

$$r_{kk} = \frac{\sum_{j=1}^{k-1}(r_{k-1,j})(r_{k-j})}{\sum_{j=1}^{k-1}(r_{k-1,j})(r_j)} \tag{12}$$

การพิจารณาเปรียบเทียบแต่ละรูปแบบ ต้องพิจารณา r_k, r_{kk} กับ p_k และ p_{kk} พร้อมกันหลายๆ ค่า จึงมักจะพิจารณาจากรูปที่เรียกว่าคอเรลโลเกรม (Correlogram) ที่ได้จากการพล๊อต (Plot) r_k, r_{kk}, p_k และ p_{kk} ในช่วงเวลา k ดังนั้นการพิจารณาเปรียบเทียบ จะเป็นการเปรียบเทียบ Correlogram ของค่า Autocorrelation ของอนุกรมเวลาตัวอย่าง (r_k) กับค่า Autocorrelation ของอนุกรมเวลาของประชากร (P_k) และ Correlogram ของค่า Partial Autocorrelation ของอนุกรมเวลาตัวอย่าง (r_{kk}) กับค่า Partial Autocorrelation ของอนุกรมเวลาประชากร (P_{kk}) สำหรับแต่ละรูปแบบจะมี Correlogram ของ P_k และ P_{kk} ต่างกัน อนุกรมเวลาที่จะนำมากำหนดรูปแบบจะต้องเป็นอนุกรมเวลาที่ Stationary เท่านั้น หากไม่เป็น Stationary จะต้องเปลี่ยนแปลงให้เป็น Stationary เสียก่อน

ในขณะที่การพิจารณาความสัมพันธ์ตัวแปรตามกับตัวแปรอิสระจะใช้ ค่า Partial Autocorrelation เป็นตัววัดความสัมพันธ์ของตัวแปรทั้งสอง โดยพิจารณาจากสมการ Yule-Walker (Pindyck and Rubinfeld, 1998) ดังนี้

$$\rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + ... + \phi_p \tag{13}$$

ถ้า k มากกว่า p จะได้

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + ... + \phi_{k-p} \tag{14}$$

เมื่อทราบคอเรลโลแกรมของ ACF และ PACF แล้ว สามารถนำมาหารูปแบบแบบจำลองโดยการพิจารณาลักษณะของ คอเรลโลแกรมของ ACF และ PACF ดังตาราง 3.1

ตาราง 3.1 แสดงการพิจารณา ACF และ PACF

ชนิดของแบบจำลอง	รูปแบบของ ACF	รูปแบบของ PACF
AR(p)	ลู่โค้งเข้าหาแกน (Tails off)	เกิดค่าที่ชัดเจนเพียง p ค่าแล้วหายไป (Cut off after lag p)
MA(q)	เกิดค่าที่ชัดเจนเพียง q ค่าแล้วหายไป (Cut off after lag p)	ลู่โค้งเข้าหาแกน (Tails off)
ARMA(p, q)	ลู่โค้งเข้าหาแกน (Tails off)	ลู่โค้งเข้าหาแกน (Tails off)

ที่มา: Gujarati (2003)

จากตาราง 3.1 จะสามารถกำหนดรูปแบบของแบบจำลองได้ดังต่อไปนี้ หากคอเรลโลแกรมของ ACF มีลักษณะโค้งลู่เข้าหาแกนในระนาบ ในขณะที่คอเรลโลแกรม PACF เกิดมีค่าขึ้นมาไม่กี่ค่าแล้วก็หายไป จำนวนของแท่งของค่าที่เกิดขึ้นมา ให้นับเป็น ค่าที่ p ของ AR(p) ยกตัวอย่าง เช่น เมื่อพิจารณาคอเรลโลแกรมของ ACF ที่โค้งลู่เข้าแกนระนาบ และ PACF ที่มีแท่งคอเรลโลแกรม เกิดขึ้น 1 แท่ง แปลได้ว่าแบบจำลองควรมีลักษณะเป็น AR(1) สำหรับ MA(q) นั้นก็จะมี ACF ที่เกิดขึ้นมาไม่กี่ค่าแล้วหายไป ในขณะที่ PACF จะลู่โค้งเข้าหาแกนระนาบนั้น ยกตัวอย่างเช่น หากค่า ACF เกิดแท่งคอเรลโลแกรมขึ้นเพียง 2 แท่ง และหลังจากนั้นก็หายไป ในขณะที่ PACF โค้งลู่เข้าหาแกนระนาบ สามารถสรุปได้ว่าแบบจำลองควรมีลักษณะเป็น MA(2) และหาก ACF และ PACF โค้งเข้าหาแกนระนาบทั้งคู่ แบบจำลองควรจะเป็น ARMA(p, q) และเมื่อรวมกันกับการทดสอบความนิ่ง (Stationary) ในขั้นตอนที่ 1 แล้ว จะสามารถหาค่าของผลต่าง (Difference) ได้ ซึ่ง

ผลจากการ ค่าของผลต่าง (Difference) จำนวน d ครั้งนั้น ก็จะได้แบบจำลอง ARIMA(p, d, q) แต่หากข้อมูลเมื่อทดสอบแล้วมีความนิ่งนั้น แสดงว่าแบบจำลองควรจะเป็น ARMA(p, q)

อย่างไรก็ตามการพิจารณาดังกล่าวเป็นเพียงการพิจารณาตามหลักการเท่านั้น เพราะการพิจารณารูปแบบ ARIMA(p, d, q) ควรมีการพิจารณาค่าสถิติเพื่อประกอบการตัดสินใจ เช่น ค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE) ค่า Theil Inequality Coefficient (U) ค่า Adjusted R^2 และค่า Akaike Information Criterion (AIC)

1.1) ค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE) คือการวัดค่าความแตกต่างระหว่างค่าความจริงและค่าที่ถูกประมาณจากแบบจำลอง หากค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง มีค่าน้อยแสดงว่าแบบจำลองสามารถประมาณค่าประมาณได้ใกล้เคียงกับค่าจริง (Pindyck and Rubinfeld, 1998) ดังนั้นหากค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง มีค่าเท่ากับศูนย์ จึงหมายถึงไม่เกิดความคลาดเคลื่อนในแบบจำลอง ดังนั้นจึงสามารถพิจารณาค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง ได้ดังนี้

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^s - Y_t^a\right)} \tag{15}$$

กำหนดให้ Y_t^s = ค่าประมาณจากแบบจำลอง

Y_t^a = ค่าที่แท้จริง

T = จำนวนคาบเวลาที่ใช้ในการประมาณแบบจำลอง

1.2) ค่า Theil Inequality Coefficient (U) มีที่มาคล้ายกับค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง แต่ค่า U นั้นจะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้า U มีค่าเท่ากับศูนย์ แสดงว่าค่าที่ได้จากการประมาณเท่ากับค่าจริงของทุกๆ เวลา t และแบบจำลองที่ประมาณได้เป็นแบบจำลองที่ดี แต่ในทางตรงกันข้ามถ้าค่า U มีค่าเท่ากับหนึ่ง แสดงว่าแบบจำลองที่ประมาณเป็นแบบจำลองที่ไม่ดีที่สุด (Pindyck and Rubinfeld, 1998) ดังนั้นสามารถพิจารณาค่า U ได้ดังนี้

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^s - Y_t^a\right)}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^s\right)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t^a\right)^2}} \tag{16}$$

กำหนดให้ Y_t^s = ค่าประมาณจากแบบจำลอง

Y_t^a = ค่าที่แท้จริง

T = จำนวนคาบเวลาที่ใช้ในการประมาณแบบจำลอง

1.3) ค่า Adjusted R-squared (Adjusted R^2) คือการพิจารณาว่าตัวแปรอิสระสามารถอธิบายการเปลี่ยนแปลงตัวแปรตามได้มากน้อยเพียงใด ถ้าหากค่า Adjusted R^2 มีค่าเท่ากับ 1 แสดงว่าตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ 100% แต่ถ้าหากค่า Adjusted R^2 เท่ากับ 0 หมายความว่า ตัวแปรอิสระไม่สามารถอธิบายตัวแปรตามได้เลย ซึ่งค่า Adjusted R^2 นี้เป็นค่าสถิติที่เกิดจากการประยุกต์มาจากค่า R^2 ซึ่งถ้ามีการเพิ่มตัวแปรอิสระเข้าไปในสมการมากขึ้นก็จะทำให้ค่า R^2 สูงขึ้นตาม ดังนั้นจึงมีการเพิ่มระดับความเป็นอิสระเข้าไปในสมการ จึงเรียกว่า Adjusted R^2($\overline{R}^2$) (Gujarati, 2003) โดยสามารถพิจารณาความสัมพันธ์ของ R^2 และ Adjusted R^2 ได้ดังสมการ

$$R^2 = 1 - \frac{\sum \hat{u}_i}{\sum y_i^2} \tag{17}$$

$$\overline{R}^2 = 1 - \frac{\sum \hat{u}_i^2/(n-k)}{\sum y_i^2/(n-1)} \tag{18}$$

1.4) ค่า Akaike Information Criterion (AIC) ค่า AIC เป็นค่าสถิติที่มีการประยุกต์คล้ายกับค่า Adjusted R^2 แต่มีการถ่วงน้ำหนักมากกว่า Adjusted R^2 และยังมีการลอการิทึมฐานธรรมชาติ (Natural Logarithm: ln) ดังนั้นถ้าค่า AIC น้อยจึงหมายถึงแบบจำลองสามารถเป็นตัวแทนข้อมูลจริงได้ดี (Gujarati, 2003) และยังนำค่า AIC ไปใช้ในการหาค่าย้อนหลัง (lag length) ที่เหมาะสมได้อีก สามารถคำนวณค่า AIC ได้ดังนี้

$$\ln AIC = \left(\frac{2k}{n}\right) + \ln\left(\frac{\sum \hat{u}_i^2}{n}\right) \tag{19}$$

กำหนดให้ $\sum \hat{u}_i^2$ = ผลรวมของค่าความคลาดเคลื่อน

n = ค่าสังเกตทั้งหมด

2) การประมาณค่าพารามิเตอร์ในรูปแบบอนุกรมเวลา (Parameter Estimation)

จะทำได้โดยการหาค่าประมาณแบบง่าย หรือค่าประมาณที่ได้จากการวิเคราะห์ตัวเลข (Numerical Analysis) สำหรับค่าประมาณแบบง่าย จะทำโดยการสร้างสมการที่มาจากความสัมพันธ์ระหว่าง P_k และพารามิเตอร์ โดยสมการที่สร้างขึ้นจะมีจำนวนเท่ากับพารามิเตอร์ที่ต้องการประมาณ ส่วนค่าประมาณที่ได้จากการวิเคราะห์ตัวเลขจะได้จากการแก้สมการที่สร้างขึ้นจากวิธีกำลังสองน้อยที่สุด ขั้นตอนของการวิเคราะห์ตัวเลขจะต้องมีการกำหนดค่าประมาณเริ่มต้นซึ่งส่วนใหญ่จะใช้การประมาณแบบง่ายเป็นค่าประมาณเริ่มต้น เมื่อการวิเคราะห์สิ้นสุดจะได้ค่าประมาณสุดท้ายที่จะนำไปใช้ประโยชน์ในการสร้างสมการพยากรณ์

3) การตรวจสอบความถูกต้อง (Diagnostics Checking)

เมื่อกำหนดรูปแบบ และประมาณค่าพารามิเตอร์ในแบบจำลองแล้ว จะต้องตรวจสอบทุกครั้งว่ารูปแบบที่กำหนดนั้นมีความเหมาะสมจริง หรือไม่ การตรวจสอบจะทำได้หลายวิธีได้แก่ การพิจารณาคอเรลโลเกรมของ r_k หรือของค่าคลาดเคลื่อน การทดสอบค่าพารามิเตอร์ในแบบจำลองด้วยการทดสอบแบบ t และการทดสอบความเหมาะสมของแบบจำลองโดยการทดสอบของ Box และ Pierce หรือการทดสอบของ Box และ Ljung หากตรวจสอบ พบว่าแบบจำลองที่กำหนดนั้น เหมาะสมแล้ว จะใช้รูปแบบนั้นในการพยากรณ์ต่อไป แต่หากพบว่าแบบจำลองที่กำหนดนั้นไม่เหมาะสมจะต้องทำตามขั้นตอนที่ 1 เพื่อกำหนดรูปแบบของแบบจำลองใหม่ ซึ่งจะพิจารณาจาก Q-statistic (Gujarati, 2003) ดังสมการ (20) ซึ่งมีการแจกแจงแบบ Chi-square และมีดีกรีเท่ากับ m โดยมีสมมติฐานว่าง คือ พจน์ความคลาดเคลื่อนที่ได้จากการประมาณมีลักษณะเป็น White Noise หมายถึง แบบจำลองไม่มี Autocorrelation ถ้าหากแบบจำลองมี Autocorrelation จะต้องกลับไปทำการกำหนดรูปแบบแบบจำลองใหม่ แต่หากแบบจำลองไม่มี Autocorrelation ก็จะใช้แบบจำลองนั้นทำการพยากรณ์ต่อไป

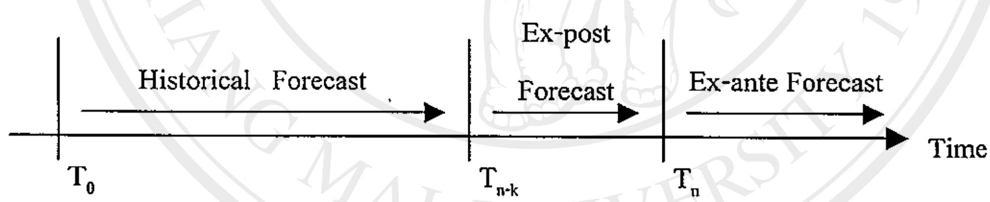
$$Q = n\sum_{k-1}^{m} \hat{\rho}_k^2 \qquad (20)$$

กำหนดให้ n = จำนวนของข้อมูล

m = ค่า lag length

4) การพยากรณ์ (Forecasting)

จะทำได้ทั้งการพยากรณ์แบบจุด (Point Forecast) และการพยากรณ์แบบช่วง (Interval Forecast) โดยการพยากรณ์จะใช้สมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนด และผ่านการตรวจสอบในขั้นตอนที่ผ่านมาแล้ว ก็สามารถนำแบบจำลองใช้ในการพยากรณ์ แต่เนื่องจากการพยากรณ์ข้อมูลไปข้างหน้านั้นจะต้องเป็นแบบจำลองที่ให้ค่าประมาณที่แม่นยำที่สุด ดังนั้นการพยากรณ์จึงต้องมีการทดสอบแบบจำลอง โดยการแบ่งการพยากรณ์ออกเป็น 3 ช่วง คือ ช่วง Historical Forecast อันเป็นการพยากรณ์ตั้งแต่อดีตจนถึงช่วงเวลาที่พิจารณา (T_0-T_{n-k}) การพยากรณ์ช่วง Ex-post forecast คือการพยากรณ์โดยการตัดข้อมูลออกมาส่วนหนึ่งแล้วทำการพยากรณ์แล้วเปรียบเทียบข้อมูลจริงกับข้อมูลที่ได้จากการพยากรณ์ โดยพิจารณาค่า Root Mean Squared Error (RMSE) และ Theil Inequality Coefficient (U) และค่า Akaike Information Criterion (AIC) จะพิจารณาค่าสถิติทั้ง 3 ค่าที่มีค่าต่ำที่สุด ซึ่งได้จากการทำการพยากรณ์ เมื่อเลือกรูปแบบจำลองที่ดีที่สุดได้แล้ว จึงนำแบบจำลองนั้นมาทำการพยากรณ์แบบ Ex-ante Forecast ซึ่งเป็นการพยากรณ์ข้อมูลไปข้างหน้า ดังรูป 3.2



รูป 3.2 ช่วงเวลาที่ใช้ในการพยากรณ์ ($T_0 - T_n$)
ที่มา: Pindyck and Rubinfeld (1998)

3.2 ระเบียบวิธีวิจัย

ในการศึกษาครั้งนี้ จะใช้การวิเคราะห์อนุกรมเวลาราคารายเดือนสินค้ากุ้งกุลาดำ ณ ตลาดระดับที่เกษตรกรขายได้ 2 ขนาด คือขนาดใหญ่ 15-30 ตัว/กิโลกรัม และขนาดกลาง 31-40 ตัว/กิโลกรัมโดยวิธี Box–Jenkins เพราะเป็นวิธีที่จะให้ค่าพยากรณ์ที่มีความถูกต้องสูง โดยเฉพาะในการพยากรณ์ระยะสั้น

Enders (1995) กล่าวว่าการพยากรณ์ข้อมูลอนุกรมเวลาต้องอยู่บนข้อสมมติว่า ข้อมูลอนุกรมเวลาต้องมีลักษณะนิ่ง (Stationary) ดังนั้นการวิเคราะห์ข้อมูลราคากุ้งกุลาดำที่เกษตรกรขาย

ได้ขนาดใหญ่ เพื่อให้ค่าความแปรปรวนของอนุกรมเวลาคงที่สำหรับค่า t ต่างๆ จึงทำการแปลงอนุกรมเวลาเดิม S_t ให้เป็นอนุกรมเวลาใหม่ซึ่งอยู่ในรูปลอการีทึมฐานธรรมชาติ (Natural Logarithm): LLS_t โดยให้ LLS_t เท่ากับ $Ln(LS_t)$ สำหรับข้อมูลราคากุ้งกุลาดำที่เกษตรกรขายได้ขนาดกลาง ทำวิธีเดียวกันได้อนุกรมเวลาใหม่ซึ่งอยู่ในรูปลอการีทึมฐานธรรมชาติ (Natural Logarithm): LMS_t โดยให้ LMS_t เท่ากับ $Ln(MS_t)$ ได้จาก

ค่าเฉลี่ย: $E[Ln(LS_t)] = \mu = \text{constant}$ (21)

ค่าความแปรปรวน: $V[Ln(LS_t)] = \sigma^2 = \text{constant}$ (22)

ค่าความแปรปรวนร่วม: $cov[Ln(LS_t), Ln(LS_{t+k})] = E[Ln(LS_t) - \mu] [Ln(LS_{t+k}) - \mu]$ (23)

หมายเหตุ: 1) ถ้าเป็นข้อมูลราคากุ้งกุลาดำขนาดใหญ่ จะใช้ค่า $Ln(LS_t)$ แทนที่ X_t และ Y_t ในสมการ (4) – (10) และ สมการ (11) (15) – (18) ตามลำดับ สำหรับระเบียบวิธีวิจัย

2) ถ้าเป็นข้อมูลราคากุ้งกุลาดำขนาดกลาง จะใช้ค่า $Ln(MS_t)$ แทนที่ X_t และ Y_t ในสมการ (4) – (10) และ สมการ (11) (15) – (18) ตามลำดับ สำหรับระเบียบวิธีวิจัย

3.2.1 การทดสอบความนิ่งของข้อมูลอนุกรมเวลา (Unit Root Test)

ในการทดสอบ Unit Root ของข้อมูลราคากุ้งกุลาดำที่เกษตรกรขายได้ขนาดใหญ่ เพื่อที่จะพิจารณาว่าข้อมูลมีลักษณะนิ่ง (Stationary): [I(0); Integrated of Order 0] หรือจะมีลักษณะไม่นิ่ง [I(d); d > 0; Integrated of Order d] เพื่อหลีกเลี่ยงความสัมพันธ์ที่ไม่แท้จริง (Spurious Regression) ซึ่งข้อมูลที่มี Unit Root อาจเป็นข้อมูลที่มีค่าเฉลี่ยและความแปรปรวนไม่คงที่เมื่อเวลาเปลี่ยนไป รวมถึงค่าความแปรปรวนร่วมขึ้นอยู่กับเวลาที่เกิดขึ้นจริง

ดังนั้นในการวิเคราะห์ข้อมูลราคากุ้งกุลาดำที่เกษตรกรขายได้ ทั้ง 2 ขนาดที่ทำการแปลงอนุกรมเวลาใหม่ซึ่งอยู่ในรูปลอการีทึมฐานธรรมชาติ จากนั้นทำการทดสอบ Unit Root ทำการทดสอบ ADF (Said and Dickey, 1984) ด้วยวิธีการเลือก lag length ของ Enders (1995) การศึกษาครั้งนี้จะสมมติให้ lag length มีค่าเท่ากับ 3 แล้วพิจารณาความมีนัยสำคัญทางสถิติ (Significant) ณ ระดับนัยสำคัญต่างๆ คือ ระดับนัยสำคัญ 1% 5% และ 10% ($\alpha = 0.01$ 0.05 และ 0.1) หากพบว่า lag length ที่เลือกค่า ADF Test-statistic ไม่มีนัยสำคัญทางสถิติ ณ ระดับความเชื่อมั่นที่ 90% จะทำการทดสอบต่อไป โดยการลดจำนวน lag length ลงทีละ 1 ช่วง จนกระทั่งทำการปฏิเสธสมมติฐานว่าง คือ ค่า Test-statistic มีนัยสำคัญทางสถิติ และพิจารณาค่า ADF Test-statistic

ประกอบด้วย เพื่อพิจารณาการมี Unit Root มีสมมติฐานว่าง คือ H_0: $\theta = 0$ เมื่อปฏิเสธสมมติฐานว่างแสดงว่าข้อมูลไม่มี Unit Root

3.2.2 การกำหนดแบบจำลองอารีมา (ARIMA: p, d, q) โดยวิธี Box - Jenkins

วิธีการวิเคราะห์อนุกรมเวลาของ Box–Jenkins เป็นการวิเคราะห์อนุกรมเวลาโดยการหารูปแบบที่เหมาะสมให้กับอนุกรมเวลา โดยใช้ค่า Autocorrelation Function (ACF) และค่า Partial Autocorrelation Function (PACF) เป็นหลักในการพิจารณา และรูปแบบที่เลือกใช้จะอยู่ในกลุ่มของรูปแบบ ARIMA (p, d, q) หรือเรียก Integrated Autoregressive–Moving Average order p and q ซึ่งเป็นรูปแบบที่กำหนดว่าค่าพยากรณ์ในอนาคตเป็นค่าที่ได้จากการสังเกต หรือค่าพยากรณ์ก่อนหน้า และ ความคลาดเคลื่อนของการพยากรณ์ก่อนหน้า โดยเป็นการรวมส่วนของรูปแบบ AR(p) และรูปแบบ MA(q) เข้าด้วยกัน รูปแบบ AR(p) หมายถึงรูปแบบที่แสดงว่าค่าสังเกต $Ln(LS_t)$ จะขึ้นอยู่กับค่า $Ln(LS_{t-1})$, $Ln(LS_{t-2})$,..., $Ln(LS_{t-p})$ หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ MA(q) หมายถึงรูปแบบที่แสดงว่าค่าสังเกต $Ln(LS_t)$ จะขึ้นอยู่กับค่าของความคลาดเคลื่อน ε_{t-1}, ε_{t-2}, ...,ε_{t-q} หรือค่าความคลาดเคลื่อนที่เกิดขึ้นก่อนหน้า q ค่าซึ่งรูปแบบ ARIMA (p, d, q) โดยมีการกำหนดรูปแบบดังนี้

ARIMA (p, d, q) คือ $$\Delta^d Ln(LS_t) = \theta_0 + \phi_1\Delta^d Ln(LS_{t-1}) + \ldots + \phi_1\Delta^d Ln(LS_{t-p}) + \varepsilon_t - \theta_1\varepsilon_{t-1} - \ldots - \theta_q\varepsilon_{t-q}$$

หมายเหตุ: 1) ถ้าเป็นข้อมูลราคากุ้งกุลาดำขนาดใหญ่ จะใช้ค่า $Ln(LS_t)$

2) ถ้าเป็นข้อมูลราคากุ้งกุลาดำขนาดกลาง จะใช้ค่า $Ln(MS_t)$ แทนที่ ค่า $Ln(LS_t)$

การพยากรณ์โดยใช้วิธี Box-Jenkins ประกอบด้วย 4 ขั้นตอน

1) การกำหนดแบบจำลอง (Identification)

ให้กับอนุกรมเวลาที่เป็น Stationary Series เป็นการหารูปแบบARIMA(p, d, q) ที่คาดว่าจะเหมาะสมให้กับอนุกรมเวลาที่แปลงข้อมูลอนุกรมเวลาให้มีลักษณะนิ่ง โดยการหาผลต่างอันดับที่ 1 แล้วแสดงให้เห็นว่าเป็นข้อมูลอนุกรมเวลามีลักษณะเป็น I(1) จึงสามารถกำหนดแบบจำลองได้เป็น ARIMA โดยการพิจารณาคอเรลโลแกรมของผลต่างอันดับที่ 1 ของข้อมูลอนุกรมเวลา ราคากุ้งกุลาดำขนาดใหญ่ [$Ln(LS_t)$] คือ ($\Delta lnLS_t$) และ พิจารณาคอเรลโลแกรมของผลต่าง

อันดับที่ 1 ของข้อมูลอนุกรมเวลา ราคากุ้งกุลาดำขนาดกลาง [$Ln(MS_t)$] คือ ($\Delta lnMS_t$) ในการกำหนดแบบจำลอง เพื่อหาค่า Autoregressive: AR(p) และ Moving Average: MA(q) โดยพิจารณาจากค่า ACF: Autocorrelation Function และ PACF: Partial Autocorrelation Function จะสามารถกำหนดแบบจำลองที่มีความเหมาะสมได้ อย่างไรก็ตามในการศึกษาครั้งนี้จะพิจารณารูปแบบ ARIMA(p, d, q) โดยใช้ค่าสถิติเพื่อประกอบการตัดสินใจ คือพิจารณาจากค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE) ค่า Theil Inequality Coefficient (U) ค่า Adjusted R^2 และค่า Akaike Information Criterion (AIC)

1.1) ค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error: RMSE) ใช้วัดค่าความแตกต่างระหว่างค่าความจริง และค่าที่ถูกประมาณจากแบบจำลองของผลต่างอันดับที่ 1 ของข้อมูลอนุกรมเวลา ราคากุ้งกุลาดำทั้ง 2 ขนาดในขั้นตอนการกำหนดแบบจำลอง หากได้ค่า RMSE มีค่าต่ำสุด แสดงว่าแบบจำลองสามารถประมาณค่าประมาณได้ใกล้เคียงกับค่าจริง และแสดงว่าแบบจำลองนั้นเป็นแบบจำลองที่เหมาะสมและดี ดังนั้นหากค่า RMSE มีค่าเท่ากับศูนย์ หมายถึงไม่เกิดความคลาดเคลื่อนในแบบจำลอง

1.2) ค่า Theil Inequality Coefficient (U) มีที่มาคล้ายกับค่าเฉลี่ยค่า RMSE แต่ค่า U นั้นจะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้า U มีค่าเท่ากับศูนย์ แสดงว่าค่าที่ได้จากการประมาณเท่ากับค่าจริงของทุกๆ เวลา t และแบบจำลองในขั้นตอนการกำหนดแบบจำลองของราคากุ้งกุลาดำทั้ง 2 ขนาดที่ประมาณได้ เป็นแบบจำลองที่ดี แต่ในทางตรงกันข้ามถ้าค่า U มีค่าเท่ากับหนึ่ง แสดงว่าแบบจำลองในขั้นตอนการกำหนดแบบจำลองของราคากุ้งกุลาดำทั้ง 2 ขนาดที่ประมาณได้ เป็นแบบจำลองที่ไม่ดีที่สุด

1.3) ค่า Adjusted R-squared (Adjusted R^2) การพิจารณาว่าตัวแปรอิสระสามารถอธิบายการเปลี่ยนแปลงตัวแปรตาม ($\Delta lnLS_t$) หรือ ($\Delta lnMS_t$) ได้มากน้อยเพียงใด ถ้าหากค่า Adjusted R^2 มีค่าเท่ากับ 1 แสดงว่าตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ 100% แต่ถ้าหากค่า Adjusted R^2 เท่ากับ 0 หมายความว่า ตัวแปรอิสระไม่สามารถอธิบายตัวแปรตามได้เลย

1.4) ค่า Akaike Information Criterion (AIC) ถ้าค่า AIC น้อย หมายถึงแบบจำลองในขั้นตอนการกำหนดแบบจำลองของราคากุ้งกุลาดำทั้ง 2 ขนาด สามารถเป็นตัวแทนข้อมูลจริงได้ดี และยังนำค่า AIC ไปใช้ในการหาค่าย้อนหลัง (lag length) ที่เหมาะสมได้อีก

2) การประมาณค่าพารามิเตอร์ในรูปแบบอนุกรมเวลา (Parameter Estimation)

หารูปแบบความสัมพันธ์ของแบบจำลองในขั้นตอนการกำหนดแบบจำลองของราคากุ้งกุลาดำทั้ง 2 ขนาด และแทนค่าประมาณการสัมประสิทธิ์พารามิเตอร์จากนั้นพิจารณา Test-statistic เพื่อทดสอบความมีนัยสำคัญทางสถิติ (Significant) ณ ระดับนัยสำคัญต่างๆ คือ ระดับนัยสำคัญ 1%, 5% และ 10% (α = 0.01 0.05 และ0.1)

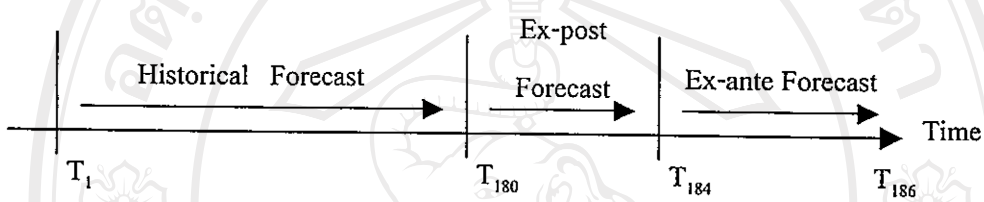
3) การตรวจสอบความถูกต้อง (Diagnostics Checking)

เมื่อกำหนดรูปแบบ และประมาณค่าพารามิเตอร์ในแบบจำลองในขั้นตอนการกำหนดแบบจำลองของราคากุ้งกุลาดำทั้ง 2 ขนาดแล้ว จะต้องตรวจสอบทุกครั้งว่ารูปแบบที่กำหนดนั้นมีความเหมาะสมจริง หรือไม่ การตรวจสอบความถูกต้องของแบบจำลอง ราคากุ้งกุลาดำทั้ง 2 ขนาดได้พิจารณาจากค่า Q-statistic (Gujarati, 2003) โดยวิธี Box and Pierce โดยใช้คุณสมบัติความเป็น White Noise ของค่าความคลาดเคลื่อนที่ประมาณการ (Estimated Residual: ε_t) และพิจารณาความไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ 1%, 5% และ 10% เพื่อพิจารณา ε_t เป็น White Noise มีการกระจายแบบปกติ (Normal Distribution) และไม่มี Autocorrelation และไม่มีความแปรปรวนแตกต่างกัน (Heteroscedasticity) หรือไม่ หากตรวจสอบ พบว่าค่า Q-statistic ของแบบจำลอง ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับต่างๆ ที่กล่าวมา แสดงว่า ε_t เป็น White Noise มีการกระจายแบบปกติ (Normal Distribution) และไม่มีAutocorrelation และไม่มีความแปรปรวนแตกต่างกัน (Heteroscedasticity) แสดงว่าแบบจำลองที่กำหนดนั้น เหมาะสมแล้วจะใช้รูปแบบนั้นในการพยากรณ์ราคากุ้งกุลาดำทั้ง 2 ขนาดต่อไปได้ แต่หากพบว่าแบบจำลองที่กำหนดนั้นไม่เหมาะสม คือมี Autocorrelation จะต้องทำตามขั้นตอน 1) เพื่อกำหนดรูปแบบของแบบจำลองใหม่

4) การพยากรณ์ (Forecasting)

จะทำการพยากรณ์แบบช่วง (Interval Forecast) โดยการพยากรณ์จะใช้สมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนด และผ่านการตรวจสอบในขั้นตอนที่ผ่านมาแล้ว ก็สามารถนำแบบจำลองใช้ในการพยากรณ์ แต่เนื่องจากการพยากรณ์ข้อมูลไปข้างหน้านั้นจะต้องเป็นแบบจำลองที่ให้ค่าประมาณที่แม่นยำที่สุด ดังนั้นการพยากรณ์จึงต้องมีการทดสอบแบบจำลอง โดยการแบ่งการพยากรณ์ออกเป็น 3 ช่วง คือ ช่วง Historical Forecast เป็นการพยากรณ์ตั้งแต่อดีตจนถึงช่วงเวลาที่พิจารณา (T_1-T_{180}) การพยากรณ์ช่วง Ex-post forecast คือการกำหนดการพยากรณ์ ในช่วงเวลาสั้นๆ ซึ่งกำหนดค่าในช่วงของการพยากรณ์ย้อนกลับไป 3 ค่าหรือ 3 ช่วงระยะ

เวลา (T_{181}-T_{183}) เพื่อเปรียบเทียบกับค่าจริงของข้อมูลที่มีอยู่ โดยใช้แบบจำลองในสมการจากแบบ Historical Forecast ซึ่งกำหนดค่าเริ่มต้นจาก T_1-T_{180} โดยพิจารณาค่า Root Mean Squared Error (RMSE) และ Theil Inequality Coefficient (U) ที่มีค่าสถิติต่ำที่สุด ซึ่งได้จากการทำการพยากรณ์ เมื่อเลือกรูปแบบจำลองที่ดีที่สุดได้แล้ว จึงนำแบบจำลองนั้นมาทำการพยากรณ์แบบ Ex-ante Forecast ซึ่งเป็นการพยากรณ์ไปข้างหน้าอีก 1 ไตรมาสหรือ 3 เดือน เนื่องจากการพยากรณ์โดยวิธี ARIMA มีความแม่นยำสำหรับการพยากรณ์ในช่วงเวลาสั้นๆ ในการศึกษาครั้งนี้ จึงได้กำหนดช่วงพยากรณ์ในอนาคตเพียง 3 ช่วงเวลา (T_{184}-T_{186}) ดังรูป 3.3



รูป 3.3 ช่วงเวลาที่ใช้ในการพยากรณ์ (T_1 - T_{186})
ที่มา: Pindyck and Rubinfeld (1998)