บทที่ 2

กรอบแนวคิด ทฤษฎี และ ผลงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

วิธีการวิเคราะห์แบบ Box-Jenkins เป็นวิธีที่นิยมใช้ในการพยากรณ์ข้อมูลที่เป็นอนุกรมเวลาเนื่องจากมีความแม่นยำและเหมาะสมในการพยากรณ์ข้อมูลระยะสั้นในอนาคตซึ่งให้ค่าพยากรณ์ที่ใกล้เคียงกับความเป็นจริง และเป็นที่ยอมรับอย่างกว้างขวาง ดังนั้นจึงเหมาะสมที่จะนำมาใช้ในการพยากรณ์ในงานวิจัยนี้ด้วย

2.1.1 แนวคิดในการพยากรณ์อนุกรมเวลา

อนุกรมเวลาเป็นค่าสังเกตของกลุ่มหรือชุดข้อมูลที่มีขึ้น ณ ช่วงเวลาต่างๆ โดยแบ่งได้เป็น 2 ลักษณะ คือ อนุกรมเวลาต่อเนื่อง เป็นค่าสังเกตที่กระทำในเวลาที่ต่อเนื่องกัน และอนุกรมเวลาไม่ต่อเนื่อง เป็นค่าสังเกตที่กระทำ ณ จุดเวลาใดเวลาหนึ่งที่ไม่ต่อเนื่องกัน การวิเคราะห์อนุกรมเวลาเป็นการวิเคราะห์หาค่าสังเกตที่มีการเปลี่ยนแปลงที่ไม่มีรูปแบบและการเปลี่ยนแปลงที่มีรูปแบบ ซึ่งการเปลี่ยนแปลงอย่างหลังนี้สามารถที่จะนำมาใช้ในการพยากรณ์รูปแบบอนุกรมเวลาในอนาคตได้ (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542)

ในการศึกษาข้อมูลอนุกรมเวลาจะต้องกำหนดรูปแบบให้กับอนุกรมเวลา เนื่องจากมีความจำเป็นที่จะต้องพิจารณาความมีเสถียรภาพและลักษณะรูปแบบพฤติกรรมว่ามีอิทธิพลของฤดูกาลและแนวโน้มเข้ามาเกี่ยวข้องหรือไม่ เมื่อได้รูปแบบของสมการที่จะใช้ในการวิเคราะห์ ก็จะสามารถพยากรณ์ข้อมูลที่ต้องการศึกษาตามวัตถุประสงค์ของการศึกษาต่อไป

2.1.2 การทดสอบความนิ่งของข้อมูลอนุกรมเวลา (Unit Root Test)

การทดสอบ unit root (ทรงศักดิ์ ศรีบุญจิตต์ และอารี วิบูลย์พงศ์, 2542) ได้อธิบายว่าสามารถทดสอบได้โดยใช้การทดสอบ DF (Dicky and Fuller, 1981) และการทดสอบ ADF (Augmented Dicky-Fuller Test) (Said and Dickey, 1984) สมมติฐานว่าง (null hypothesis) ของการทดสอบ DF (Dicky-Fuller Test) คือ $H_0 : \rho = 1$ จากสมการ(2.1)

$$X_t = \rho X_{t-1} + \varepsilon_t \qquad (2.1)$$

โดยที่ถ้าทดสอบพบว่า $|\rho| < 1$ แล้วแสดงว่า X_t จะมีลักษณะนิ่งและถ้า $\rho = 1$ แล้วข้อมูล X_t จะมีลักษณะไม่นิ่ง อย่างไรก็ตามการทดสอบนี้สามารถทำได้อีกทางหนึ่งซึ่งเหมือนกับสมการ (2.1) กล่าวคือ

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \qquad (2.2)$$

ซึ่ง $X_t = (1+\theta)X_{t-1} + \varepsilon_t$ คือสมการที่ (2.1) นั่นเอง โดยที่ $\rho = (1+\theta)$ ถ้า θ ในสมการ (5) มีค่าเป็นลบ จะได้ว่า ρ ในสมการ (2.1) จะมีค่าน้อยกว่า 1 ดังนั้นสามารถจะสรุปได้ว่าการปฏิเสธ $H_0 : \theta = 0$ ซึ่งเป็นการยอมรับ $H_a : \theta < 0$ หมายความว่า $\rho < 1$ และข้อมูล X_t มี integration of order zero (Charemza and Deadman, 1979 : 131) นั่นคือ X_t มีลักษณะนิ่ง และถ้าเราไม่สามารถปฏิเสธ $H_0 : \theta$ ได้ ก็จะหมายความว่า X_t มีลักษณะไม่นิ่งถ้า X_t เป็นแนวเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (random walk with drift) เราสามารถเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t \qquad (2.3)$$

และถ้า X_t เป็นแนวเดินเชิงสุ่มซึ่งมีความโน้มเอียงทั่วไปรวมอยู่ด้วย (random walk with drift) และมีแนวโน้มตามเวลาเชิงเส้น (linear time trand) เราสามารถเขียนแบบจำลองได้ดังนี้

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t \qquad (2.4)$$

โดยที่ t= เวลา ซึ่งก็จะทำการทดสอบ $H_0 : \theta = 0$ โดย $H_a : \theta < 0$ เช่นเดียวกับที่กล่าวมาข้างต้นโดยสรุปแล้ว Dicky and Fuller (1981) ได้พิจารณาสมการถดถอย 3 รูปแบบที่แตกต่างกันในการทดสอบว่ามี unit root หรือไม่ ซึ่ง 3 สมการดังกล่าวได้แก่

Random walk

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

Random walk with drift

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

Random walk with drift and trand

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

โดยตัวพารามิเตอร์ที่อยู่ในความสนใจในทุกสมการ คือ θ นั่นคือ $\theta = 0$ ข้อมูล X_t จะมี unit root โดยการเปรียบเทียบค่าสถิติ t (t- statistic) ที่คำนวณได้กับค่าที่เหมาะสมที่อยู่ในตาราง Dicky–Fuller (Enders, 1995: 221) หรือกับค่าวิกฤติ MacKinnon (Gujarati, 2003:769)

อย่างไรก็ตามค่าวิกฤติ (critical values) จะไม่เปลี่ยนแปลง ถ้าสมการ (2.2), (2.3), (2.4) ถูกแทนที่โดยกระบวนการเชิงอัตถดถอย (autoregressive processes)

$$\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{\rho} \varphi \Delta X_{t-1} + \varepsilon_t \tag{2.5}$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{\rho} \varphi \Delta X_{t-1} + \varepsilon_t \tag{2.6}$$

$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \sum_{i=1}^{\rho} \varphi \Delta X_{t-1} + \varepsilon_t \tag{2.7}$$

จำนวนของ lagged difference terms ที่จะนำเข้ามารวมในสมการนั้นมีมากพอที่จะทำให้พจน์ค่าความคลาดเคลื่อน (error terms) มีลักษณะเป็น serially independent และเมื่อนำเอาการทดสอบ DF (Dicky – Fuller) Test มาใช้กับสมการ (2.5), (2.6) และ (2.7) เราจะเรียกว่าการทดสอบ ADF (Augmented Dickey – Fuller) Test ค่าสถิติทดสอบ ADF (ADF Test statistic) มีการแจกแจงเชิงเส้นกำกับ (asymptotic distribution) เหมือนกับสถิติ DF (DF statistic) ดังนั้นก็สามารถใช้ค่าวิกฤติ (critical values) แบบเดียวกัน (Gujarati, 2003: 720)

ในกรณีของการหา lag length ที่เหมาะสมนั้น (Enders, 1995: 277) ได้เสนอแนะว่าวิธีหนึ่งในการหา lag length ที่เหมาะสมนั้น (Enders, 1995: 227) ได้เสนอแนะว่าวิธีหนึ่งในการหา lag length ก็คือเริ่มต้นด้วยการให้มี lag length ที่ยาวมากพอและก็ลดขนาดของ lag length ลงโดยใช้ค่าสถิติทดสอบ t (test) หรือค่าสถิติทดสอบ F (F-test) สมมติว่าเราใช้ lag length เท่ากับ n* ถ้าสถิติ t (t-statistic) ของ lag n* ไม่มีนัยสำคัญ ณ ค่าวิกฤติที่กำหนดให้ เราก็จะต้องทำการประมาณค่าการถดถอยใหม่โดยใช้ lag length n* - 1 ทำอย่างนี้ เรื่อยไปจนกระทั่ง lag นั้นมีค่าสถิติ t (t-statistic) แตกต่างไปจากศูนย์ อย่างมีนัยสำคัญ

2.1.3 การเลือกแบบจำลองที่เหมาะสมจากการทดสอบ Unit Root โดยการทดสอบสัมประสิทธิ์การถดถอย (Deterministic Regressors)

เป็นการทดสอบว่า แบบจำลองใดเป็นแบบจำลองที่มีความเหมาะสมที่สุดระหว่าง กรณีของแบบจำลองที่ไม่มีค่าคงที่และแนวโน้มเวลา(None) แบบจำลองที่มีค่าคงที่(Intercept) และ

แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา(Trend and Intercept) โดยการทดสอบความมีนัยสำคัญทางสถิติของสัมประสิทธิ์ของตัวถดถอย(ค่าคงที่หรือค่าแนวโน้มเวลา)โดยมีขั้นตอนการทดสอบดังนี้

ขั้นตอนที่ 1 เริ่มการทดสอบจากแบบจำลองกรณีที่มีทั้งค่าคงที่และแนวโน้มเวลาตามสมการ (2.8)

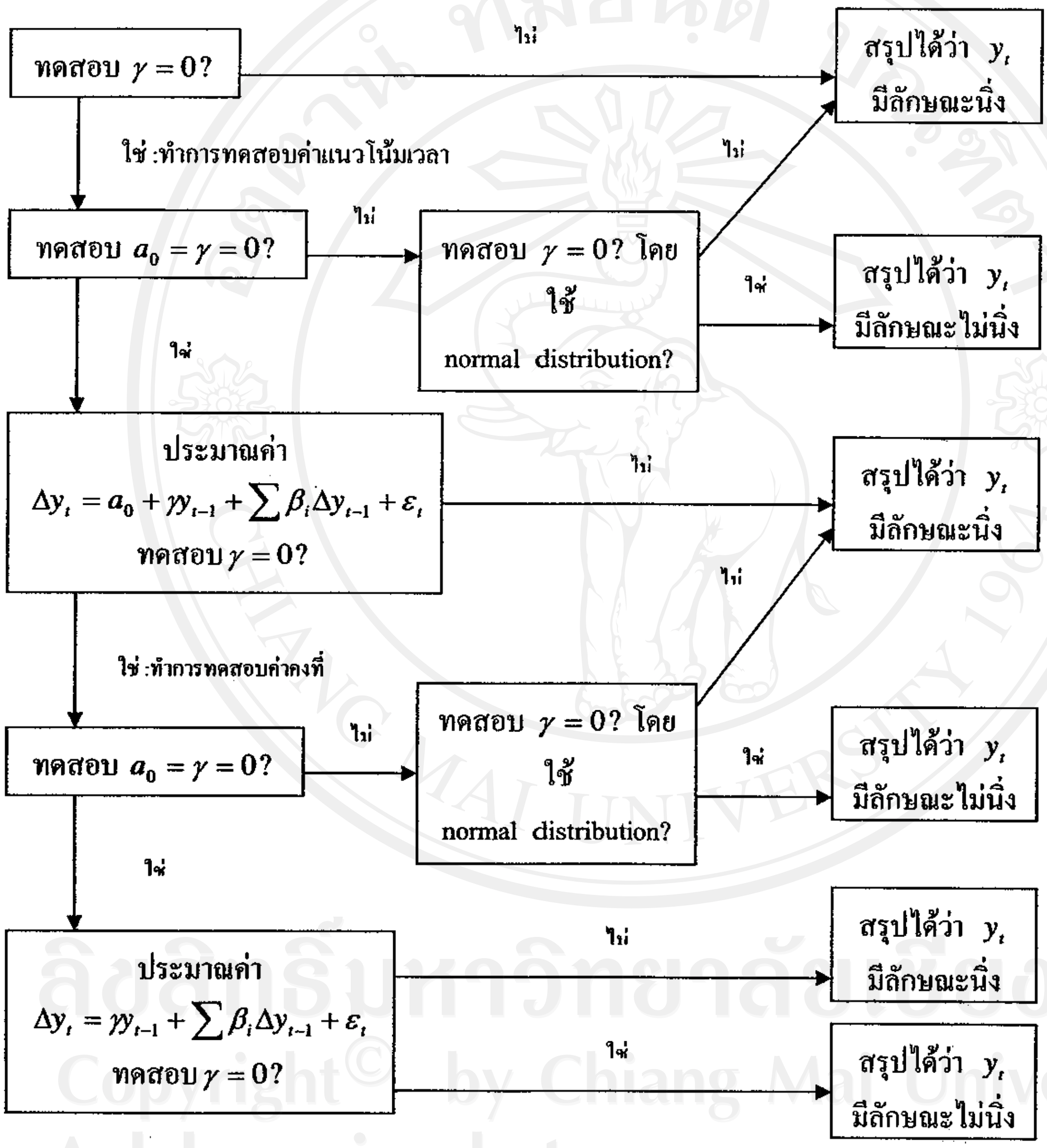
$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum \beta_i \Delta y_{t-i} + \varepsilon_t \tag{2.8}$$

ทำการทดสอบสมมติฐานว่าง H_0: γ=0 โดยใช้ τ_τ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่าง แสดงว่า ข้อมูล y_t มีลักษณะนิ่งแล้ว และเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา

ขั้นตอนที่ 2 ถ้าเกิดการยอมรับสมมติฐานว่างในขั้นตอนที่ 1 แสดงว่าในแบบจำลองดังกล่าวมี ซึ่งการมีตัวถดถอยที่ไม่จำเป็นอยู่ในสมการอาจทำให้อำนาจการทดสอบของสมการลดลง ดังนั้นจึงต้องมีการทดสอบความมีนัยสำคัญทางสถิติของค่าแนวโน้ม $(a_2 t)$ที่อยู่ในสมการ โดยการทดสอบสมมติฐานว่าง H_0: a_2=γ=0โดยใช้ ϕ3 statistic ถ้าค่าสัมประสิทธิ์ของค่าแนวโน้มไม่มีนัยสำคัญทางสถติให้ข้ามไปขั้นตอนที่ 3 อย่างไรก็ตามถ้าค่าสัมประสิทธิ์ของค่าแนวโน้มมีนัยสำคัญทางสถติ ให้ทำการทดสอบความไม่นิ่งของข้อมูลอีกครั้งโดยใช้ การแจกแจงแบบปกติมาตรฐาน (standardized normal distribution)ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่า ข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนว โน้มเวลาแต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล y_t มีลักษณะไม่นิ่ง

ขั้นตอนที่ 3 ทำการประมาณค่าแบบจำลองตามสมการ(3.8) ที่ปราศจากค่าแนวโน้มเวลา และทดสอบ unit root โดยใช้ τ_μ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่าข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่ปราศจากค่าแนวโน้มเวลา แต่ถ้าเกิดการยอมรับสมมติฐานว่างให้ทำการทดสอบความมีนัยสำคัญทางสถิติของค่าคงที่ โดยมีสมมติฐานว่าง H_0: a_0=γ=0 โดยใช้ ϕ1 statistic ถ้าค่าคงที่ไม่มีนัยสำคัญให้ข้ามไปขั้นตอนที่ 4 อย่างไรก็ตามถ้าค่าคงที่มีนัยสำคัญทางสถติ ให้ทำการทดสอบความไม่นิ่งของข้อมูลอีกครั้งโดยใช้ การแจกแจงแบบปกติมาตรฐาน (standardized normal distribution) ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่า ข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลาแต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล y_t มีลักษณะไม่นิ่ง

ขั้นตอนที่ 4 ทำการประมาณค่าแบบจำลองตามสมการ (2.8) ที่ปราศจากค่าแนวโน้มเวลาและค่าคงที่และทดสอบ unit root โดยใช้ τ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่าข้อมูล y_t มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่ปราศจากค่าแนวโน้มเวลาและค่าคงที่แต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล y_t มีลักษณะไม่นิ่ง



รูปที่ 2.1 ขั้นตอนการเลือกแบบจำลองที่เหมาะสม

ที่มา: Enders (1995)

2.1.4 การทดสอบความนิ่งของอนุกรมเวลาแบบฤดูกาล (Seasonal Unit Root Test)

ความสำคัญของการนำข้อมูลอนุกรมเวลามาทำการทดสอบ seasonal unit root เนื่องจากข้อมูลอนุกรมเวลาบางชุดมีความไม่นิ่งของฤดูกาลเข้ามาเกี่ยวข้อง ซึ่งถ้าหากนำข้อมูลที่มีความไม่นิ่งของฤดูกาลมาทำการประมาณค่าแล้วอาจทำให้ผลลัพธ์มีความคลาดเคลื่อนเกิดขึ้นได้ ดังนั้นจึงต้องทำการทดสอบแบบฤดูกาลก่อน โดยการทดสอบมีด้วยกัน 4 แบบคือ ความนิ่งมาตรฐาน (seasonal unit root at zero frequency) ความนิ่งแบบรายครึ่งปี (unit root at the biannual frequency) ความนิ่งแบบรายไตรมาส (unit root with an quarterly frequency) และความนิ่งแบบรายเดือนโดยมีรูปแบบสมการที่ใช้ในการทดสอบดังนี้

$$
\begin{aligned}
X_{8,t} = {} & \pi_1 X_{1,t-1} + \pi_2 X_{2,t-1} + \pi_3 X_{3,t-2} + \pi_4 X_{3,t-1} + \pi_5 X_{4,t-2} \\
& + \pi_6 X_{4,t-1} + \pi_7 X_{5,t-2} + \pi_8 X_{5,t-1} + \pi_9 X_{6,t-2} + \pi_{10} X_{6,t-1} \\
& + \pi_{11} X_{7,t-2} + \pi_{12} X_{7,t-1} + \mu_t + \varepsilon_t
\end{aligned}
\qquad (2.9)
$$

โดยที่

$\pi_1,\ldots,\pi_{12}$ = ค่าสัมประสิทธิ

$$
\begin{aligned}
X_{1,t} &= (1+L)(1+L^2)(1+L^4+L^8)X_t \\
&= X_t + X_{t-1} + X_{t-2} + X_{t-3} + X_{t-4} + X_{t-5} + X_{t-6} + X_{t-7} \\
&\quad + X_{t-8} + X_{t-9} + X_{t-10} + X_{t-11} \\
X_{2,t} &= -(1-L)(1+L^2)(1+L^4+L^8)X_t \\
&= -X_t + X_{t-1} - X_{t-2} + X_{t-3} - X_{t-4} + X_{t-5} - X_{t-6} + X_{t-7} \\
&\quad - X_{t-8} + X_{t-9} - X_{t-10} + X_{t-11} \\
X_{3,t} &= -(1-L^2)(1+L^4+L^8)X_t \\
&= -X_t + X_{t-2} - X_{t-4} + X_{t-6} - X_{t-8} + X_{t-10} \\
X_{4,t} &= -(1-L^4)(1-\sqrt{3}L+L^2)(1+L^2+L^4)X_t \\
&= -X_t + \sqrt{3}X_{t-1} - 2X_{t-2} + \sqrt{3}X_{t-3} - X_{t-4} + X_{t-6} - \sqrt{3}X_{t-7} \\
&\quad + 2X_{t-8} - \sqrt{3}X_{t-9} + X_{t-10}
\end{aligned}
$$

$$X_{5,t} = -(1-L^4)(1+\sqrt{3}L+L^2)(1+L^2+L^4)X_t$$
$$= -X_t - \sqrt{3}X_{t-1} - 2X_{t-2} - \sqrt{3}X_{t-3} - X_{t-4} + X_{t-6} + \sqrt{3}X_{t-7}$$
$$+ 2X_{t-8} + \sqrt{3}X_{t-9} + X_{t-10}$$
$$X_{6,t} = -(1-L^4)(1-L^2+L^4)(1-L+L^2)X_t$$
$$= -X_t + X_{t-1} - X_{t-3} + X_{t-4} - X_{t-6} + X_{t-7} - X_{t-9} + X_{t-10}$$
$$X_{7,t} = -(1-L^4)(1-L^2+L^4)(1+L+L^2)X_t$$
$$= -X_t - X_{t-1} + X_{t-3} + X_{t-4} - X_{t-6} - X_{t-7} + X_{t-9} + X_{t-10}$$
$$X_{8,t} = (1-L^{12})X_t$$
$$= X_t - X_{t-12}$$

μ_t = ค่า deterministic component = $D_1+D_2+\ldots+D_{11}+C$

$D_1,D_2,\ldots,D_{11}$ = Dummy variable

C = ค่าคงที่

ε_t = ค่าความคลาดเคลื่อน

โดยสมมติฐานว่างของการทดสอบความนิ่งแบบมาตรฐาน คือ H_0 : $\pi_1 = 0$ เมื่อทำการทดสอบค่า t-test แล้ว $\pi_1 = 0$ (ยอมรับสมมติฐานว่าง) $X_{8,t}$ มีลักษณะไม่นิ่งแบบมาตรฐาน สำหรับการทดสอบความนิ่งแบบรายครึ่งปี คือ H_0 : $\pi_2 = 0$ เมื่อทำการทดสอบค่า t-test แล้ว $\pi_2 = 0$ (ยอมรับสมมติฐานว่าง) $X_{8,t}$ มีลักษณะไม่นิ่งแบบรายครึ่งปี และการทดสอบความนิ่งแบบรายไตรมาส โดยใช้การทดสอบ F-test สมมติฐานว่าง H_0 : $\pi_3 = \pi_4 = 0$ เมื่อทำการทดสอบแล้วค่า F-test ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติ แสดงว่า $X_{8,t}$ มีลักษณะไม่นิ่งแบบไตรมาส ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ 5 % ใช้ค่าทดสอบที่ได้จาก ตารางแสดงค่าวิกฤติ สำหรับ Seasonal Unit ภาคผนวก ง

2.1.5 แนวคิดการวิเคราะห์ข้อมูลอนุกรมเวลาของ Box-Jenkins

วิธีการวิเคราะห์อนุกรมเวลาของ Box-Jenkins เป็นวิธีการพยากรณ์ข้อมูลอนุกรมเวลาโดยการหารูปแบบที่เหมาะสมให้กับข้อมูลเพื่อใช้ในการพยากรณ์ในช่วงระยะเวลาสั้นๆ การกำหนดรูปแบบที่เหมาะสมให้กับอนุกรมเวลานั้นจะใช้ค่าสัมประสิทธิสหสัมพันธ์แบบออโต (ACF) และค่าสัมประสิทธิสหสัมพันธ์บางส่วนแบบออโต (PACF) ร่วมใช้ในการพิจารณา

รูปแบบ AR(p) หมายถึงรูปแบบที่แสดงว่าค่าสังเกตุ X_t จะขึ้นอยู่กับ $X_{t-1}, X_{t-2} ..., X_{t-p}$ รูปแบบ MA(q) หมายถึงรูปแบบที่แสดงว่าค่าสังเกตุ X_t ขึ้นอยู่กับค่าความคลาดเคลื่อน $\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q}$

โดยรูปแบบของ AR(p) คือ

$$X_t = \theta_0 + \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + \varepsilon_t \tag{2.10}$$

และรูปแบบของ MA(q) คือ

$$X_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t=q} \tag{2.11}$$

ดังนั้นรูปแบบของ ARMA(p,q) คือ

$$X_t = \theta_0 + \phi_1 X_{t-1} + ... + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t=q} \tag{2.12}$$

และรูปแบบของ ARIMA(p,d,q) คือ

$$\Delta^d X_t = \theta_0 + \phi_1 \Delta^d X_{t-1} + ... + \phi_p \Delta^d X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t=q} \tag{2.13}$$

2.1.6 แบบจำลองการพยากรณ์โดยวิธี Box – Jenkins

การพยากรณ์อนุกรมเวลาโดยวิธี Box – Jenkins ในรูปแบบ ARIMA(p,d,q) ต้องพิจารณาว่าอนุกรมเวลาเป็น stationary series หรือไม่ โดยพิจารณาจาก

- ค่าเฉลี่ย E(X_t) คงที่ สำหรับทุกค่าของ t หรือไม่ จะทำได้โดยการแบ่งอนุกรมเวลาออกเป็นส่วน ๆ แล้วหาค่าความแปรปรวนของอนุกรมเวลาแต่ละส่วน ถ้าค่าเฉลี่ยแต่ละส่วนย่อยไม่แตกต่างกันมากจะสรุปได้ว่า E(X_t) คงที่

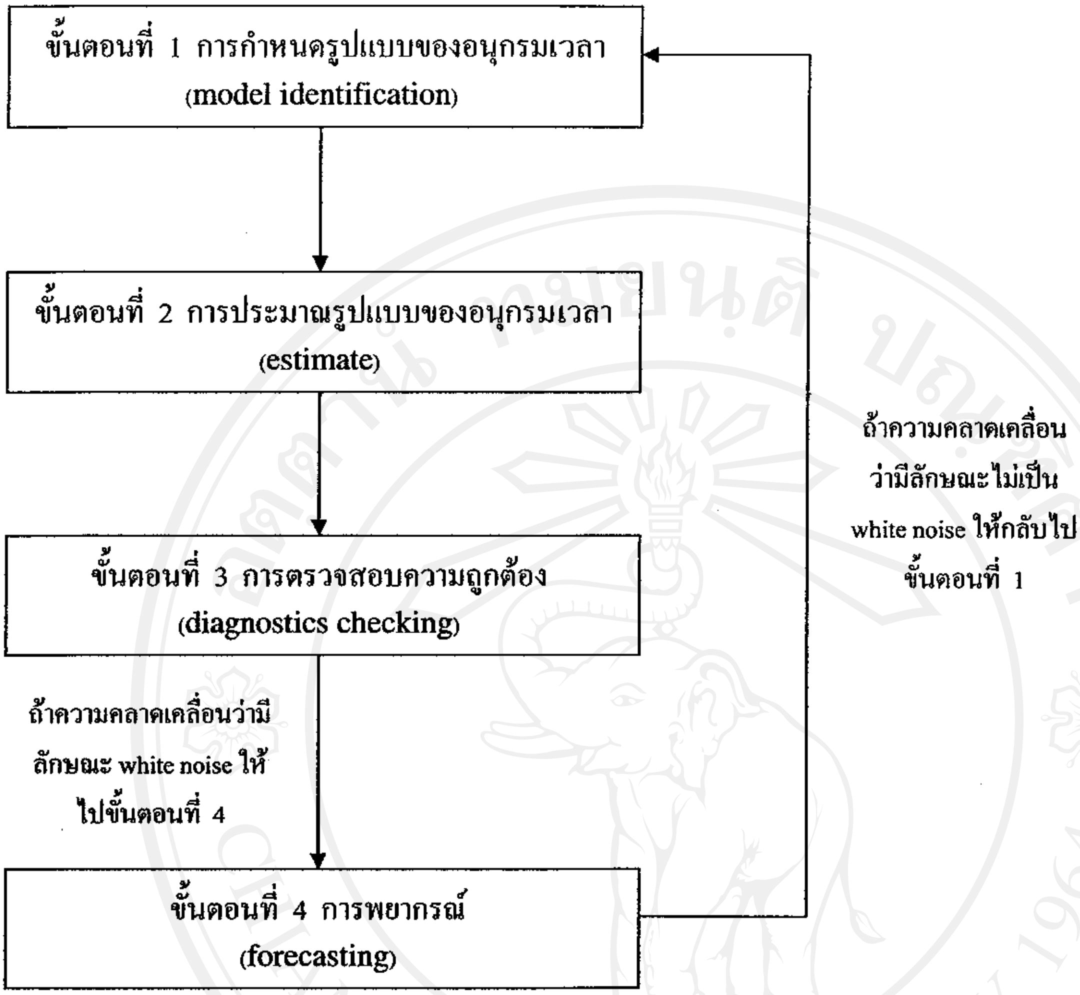
- ค่าความแปรปรวน V(X_t) คงที่ สำหรับทุกค่าของ t หรือไม่ จะทำได้โดยการแบ่งอนุกรมเวลาออกเป็นส่วน ๆ แล้วหาค่าความแปรปรวนของอนุกรมเวลาแต่ละส่วนถ้าค่าความแปรปรวนแต่ละส่วนย่อย ไม่แตกต่างกันมากนักจะสรุปได้ว่า V(X_t) คงที่

- พิจารณาแนวโน้มและปัจจัยฤดูกาล ด้วยการวาดกราฟอนุกรมเวลา ในกรณีที่มีแนวโน้มและปัจจัยฤดูกาลมักจะเห็นชัดเจนได้จากรูปที่เรียกว่า คอเรสโรแกรม (correlogram)

- พิจารณาคอเรลโลแกรม ของสัมประสิทธิ์สหสัมพันธ์ของตัวอย่าง (r_k)

กรณีที่อนุกรมเวลาเป็นแบบ stationary ค่าคอเรลโลแกรม ของสัมประสิทธิ์สหสัมพันธ์ (r_k) จะมีค่าลดลงค่อนข้างเร็ว เมื่อ k มีค่าเพิ่มขึ้นมาก ดังนั้นถ้าค่าสัมประสิทธิ์สหสัมพันธ์ (r_k) มีค่าลดลงค่อนข้างช้าจะเป็นข้อสังเกตว่าอนุกรมเวลาชุดนี้มีแนวโน้ม แต่ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ (r_k) มีค่าลดลงค่อนข้างช้า และมีค่าค่อนข้างสูงที่ k = L, 2L, 3L จะเป็น ข้อสังเกตว่าอนุกรมเวลาชุดนี้มีแนวโน้มและอิทธิพลของฤดูกาลและถ้าการเคลื่อนไหวของค่า คอเรลโลแกรมของสัมประสิทธิ์สหสัมพันธ์ (r_k) มีลักษณะคล้ายลูกคลื่น โดยคลื่นจะครบรอบภายใน 2 ช่วงเวลา แสดงว่าอนุกรมเวลามีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง

เมื่อพิจารณาจากการตรวจสอบแล้วว่าอนุกรมเวลาที่ศึกษาไม่มีลักษณะนิ่ง ก่อนจะทำการกำหนดรูปแบบให้กับอนุกรมเวลาที่ไม่นิ่ง จะต้องแปลงอนุกรมเวลาที่มีลักษณะนิ่ง เสียก่อนโดยการหาผลต่างสำหรับอนุกรมเวลาที่มีแนวโน้ม ถ้าอนุกรมเวลามีทั้งแนวโน้มอิทธิพลฤดูกาลให้หาผลต่างฤดูกาลจนได้อนุกรมเวลาที่มีลักษณะนิ่ง ถ้าอนุกรมเวลามีทั้งแนวโน้มและอิทธิพลฤดูกาลให้หาผลต่างฤดูกาลได้อนุกรมเวลาที่มีลักษณะนิ่ง แต่ถ้าอนุกรมเวลามีความแปรปรวนไม่คงที่ ให้แปลงอนุกรมเวลาเดิมโดยการหา ลอการิทึม $Z = \log(X_t)$ จนกว่าจะได้อนุกรมเวลาใหม่ ที่มีความแปรปรวนคงที่ จากอนุกรมเวลาใหม่เป็น stationary series แล้วจะทำตามขั้นตอนของ Box and Jenkin ดังรูปที่ 2.2



รูปที่ 2.2 ขั้นตอนการพยากรณ์โดยวิธี Box and Jenkins

1) การกำหนดรูปแบบจำลอง (identification)

การกำหนดรูปแบบจำลองให้กับอนุกรมเวลาที่เป็น stationary series เป็นการหารูปแบบ ARMA (p,q) ที่คาดว่าจะเหมาะสมให้กับอนุกรมเวลาโดยที่ autocorrelation : rk คือการวัดความสัมพันธ์ของแต่ละช่วงเวลา โดยมีช่วงเวลาที่ย้อนกลับไป k หน่วยเวลา โดยที่ rk มีค่าเท่ากับ $-1 < rk < 1$ โดยพิจารณาเปรียบเทียบค่า autocorrelation (rk) ของอนุกรมเวลาตัวอย่างกับค่า autocorrelation (rk) ของอนุกรมเวลาของประชากรที่มีช่วงเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตรดังนี้

$$r_k = \frac{\sum_{t=a}^{n-k}(X_{t-q})(X_{t+k-q})}{\sum_{t=a}^{n}(X_{t-q})^2} \qquad (2.14)$$

โดยที่ $X_t = \sum_{t=a}^{n}(X_t)$

q = จำนวนเวลาสุดท้ายที่ย้อนหลัง

Partial Autocorrelation (r_{kk}) คือการวัดสัมพันธ์ของแต่ละช่วงเวลาโดยมีช่วงเวลาย้อนหลังไป k หน่วยเวลา โดยพิจารณาเปรียบเทียบค่า Partial autocorrelation (r_{kk}) ของอนุกรมเวลาตัวอย่างกับ Partial autocorrelation (ρ_{kk}) ของอนุกรมเวลาของประชากรที่มีเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตรดังนี้

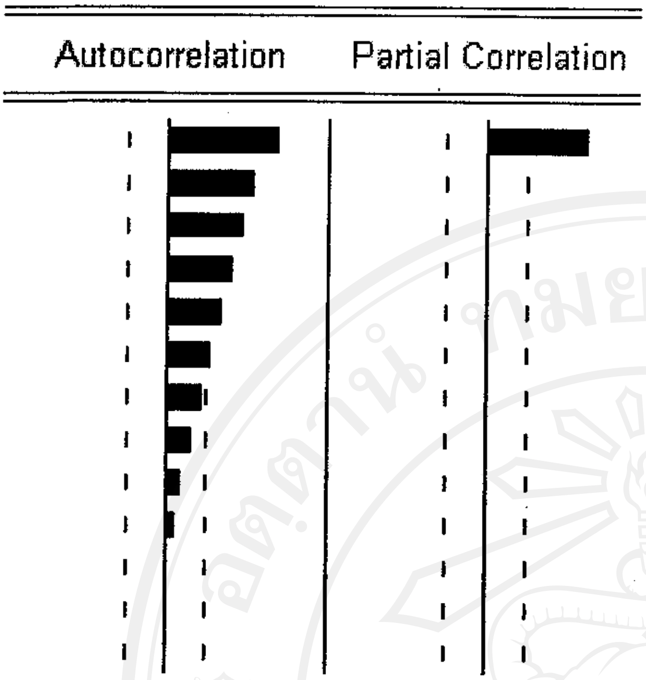
$$r_{kk} = \frac{\sum_{j=1}^{k-1}(r_{k-1})(r_{k-j})}{\sum_{j=1}^{k-1}(r_{k-1,j})(r_j)} \tag{2.15}$$

การกำหนดลำดับขั้น p, q ในแบบจำลอง (identifying the dependence order of model) ขั้นตอนคือการระบุว่าแบบจำลองนี้ควรจะมี autoregressive, p เท่าใด differencing, d ที่ลำดับเท่าใด และ moving average, q เท่าใด โดยพิจารณาจาก ACF และ PACF ซึ่งอาจจะใช้ตารางดังต่อไปนี้พิจารณาร่วม

ตารางที่ 2.1 การพิจารณา ACF และ PACF

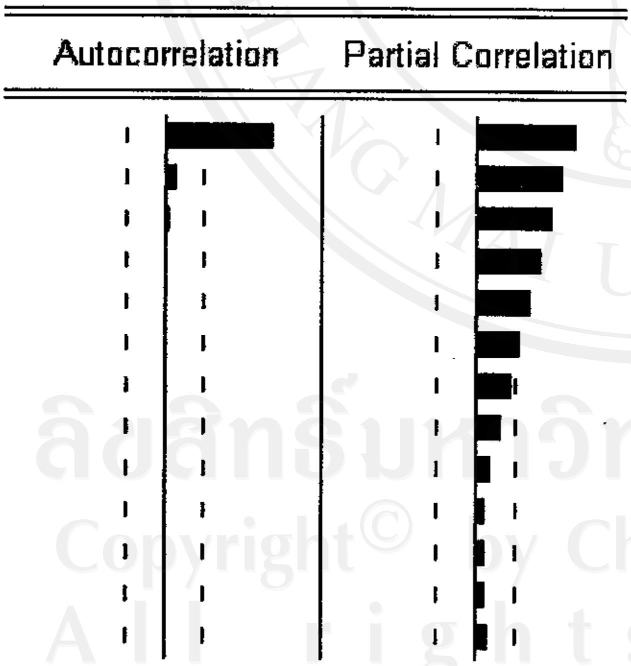
ชนิดของแบบจำลอง	**รูปแบบของ ACF**	**รูปแบบของ PACF**
AR(p)	ลู่โค้งเข้าหาแกน (tails off)	เกิดค่าที่ชัดเจนเพียง p ค่าแล้วหายไป (cut off after lag p)
MA (q)	เกิดค่าที่ชัดเจนเพียง q ค่าแล้วหายไป (cut off after lag q)	ลู่โค้งเข้าหาแกน (tails off)
ARMA (p , q)	ลู่โค้งเข้าหาแกน (tails off)	ลู่โค้งเข้าหาแกน (tails off)

ที่มา: Gujarati (2003)



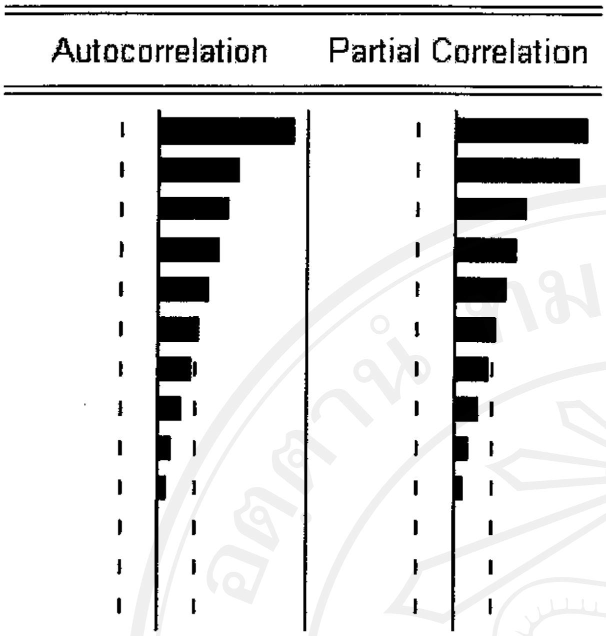
ได้แบบจำลอง AR(1)

รูปที่ 2.3 ตัวอย่างการเกิดแบบจำลอง AR(1)



ได้แบบจำลอง MA(1)

รูปที่ 2.4 ตัวอย่างการเกิดแบบจำลอง MA(1)



ได้แบบจำลอง ARMA(2,1)

รูปที่ 2.5 ตัวอย่างการเกิดแบบจำลอง ARMA(2,1)

จากตารางที่ 2.1 จะสามารถกำหนดรูปแบบของแบบจำลองได้ดังนี้ หากคอเรลโลแกรมของ ACF มีลักษณะโค้งลู่เข้าหาแกนในระนาบ ในขณะที่คอเรลโลแกรม PACF เกิดมีค่าขึ้นมาไม่กี่ค่าแล้วหายไป จำนวนของแท่งของค่าที่เกิดขึ้นมา ให้นับเป็น ค่าที่ p ของ AR(p) ยกตัวอย่างเชื่อเมื่อพิจารณา คอเรลโลแกรมของ ACF ที่โค้งลู่เข้าแกนระนาบ และ PACF ที่มีแท่งคอเรลโลแกรมเกิดขึ้น 1 แท่ง แปลได้ว่าแบบจำลองควรมีลักษณะเป็น AR(1) สำหรับ MA(q) นั้นก็จะมี ACF ที่เกิดขึ้นมาไม่กี่ค่าแล้วหายไป ในขณะที่ PACF จะลู่โค้งเข้าหาแกนระนาบนั้น ยกตัวอย่างเช่น หากค่า ACF เกิดแท่งคอเรลโลแกรมขึ้นเพียง 2 แท่งและหลังจากนั้นก็หายไป ในขณะที่ PACF โค้งลดเข้าหาแกนระนาบ สามารถสรุปได้ว่าแบบจำลองควรมีลักษณะเป็น MA(2) และหาก ACF และ PACF โค้งเข้าหาแกนระนาบทั้งคู่ แบบจำลองควรจะเป็น ARMA (p, q) และเมื่อรวมกันกับการทดสอบความนิ่ง ในขั้นตอนที่ 1 แล้ว จะสามารถหาค่าของ difference จำนวน d ครั้งนั้นก็จะได้แบบจำลอง ARIMA(p, d, q) แต่อย่างไรก็ตามหลักการดังกล่าวก็เป็นเพียงเครื่องช่วยการพิจารณาในระดับหนึ่งเท่านั้นดังนั้นเพื่อประเมินแบบจำลองว่าแบบจำลองใดมีความเหมาะสมที่จะใช้เป็นตัวแทนกลุ่มข้อมูลจริง สามารถพิจารณาได้จากค่าสถิติดังต่อไปนี้เพื่อประกอบในการตัดสินใจ

- ค่ารากที่สองของค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง (root mean square error : RMSE) โดยจะเป็นการวัดค่าความคลาดเคลื่อนระหว่างค่าจริง และค่าที่ประมาณจากแบบจำลองมีความแตกต่างกันมากน้อยเพียงใด ซึ่งหากค่า RMSE มีค่าเท่ากับ 0 จะหมายถึงแบบจำลองที่

ประมาณได้มีค่าเท่ากันกับค่าจริงพอดี ดังนั้นหากว่าค่า RMSE มีค่าน้อยเพียงใดก็แสดงว่าแบบจำลองนั้นสามารถเป็นตัวแทนค่าจริงได้ดีมากเพียงนั้น สามารถพิจารณาสมการค่ารากที่สองของค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสอง (RMSE) ได้ดังนี้

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_t^s - X_t^a)} \tag{2.16}$$

กำหนดให้ X_t^s = ค่าที่ประมาณจากแบบจำลอง

X_t^a = ค่าข้อมูลจริง

T = จำนวนของคาบเวลาที่ใช้ในการประมาณแบบจำลอง

- ค่า Theil's Inequality Coefficient โดยในหลักการเบื้องต้น พบว่าสมการที่ใช้นี้ยังคงมีหลักการที่คล้ายคลึงกันกับ RMSE โดยสิ่งที่ต่างออกไปจาก RMSE คือค่าสถิตินี้จะมีค่าอยู่ระหว่าง 0 และ 1 ทั้งนี้หากค่า U มีค่าเท่ากับศูนย์ นั่นก็หมายความว่าค่าที่ได้จากการประมาณมีค่าเท่ากันพอดีกับค่าที่เป็นข้อมูลจริงแสดงถึงแบบจำลองที่ประมาณได้เป็นแบบจำลองทีแย่ที่สุด ดังนั้นวิธีการพิจารณาค่าสถิตินี้ให้เลือกจากแบบจำลองที่มีค่า U ที่น้อย ๆ

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_t^s - X_t^a)}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_t^s)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(X_t^a)^2}} \tag{2.17}$$

กำหนดให้ X_t^s = ค่าที่ประมาณจากแบบจำลอง

X_t^a = ค่าข้อมูลจริง

T = จำนวนของคาบเวลาที่ใช้ในการประมาณแบบจำลอง

อย่างไรก็ตามยังมีค่าสถิติอีกหลายอย่างที่สามารถนำมาพิจารณาประกอบร่วมกันกับ RMSE และ Theil's Inequality Coefficient เพื่อใช้ในการเลือกแบบจำลองที่เหมาะสมที่สุด อาทิเช่น R^2, Adjusted R^2 และ Akike Information Criterion (AIC) ซึ่งสามารถอธิบายได้ดังต่อไปนี้

- ค่า R^2 คือการวัดค่าตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ดีเพียงใด หากค่านี้เท่ากับ 1 ก็หมายความว่าตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ 100% ในทางกลับกัน หากค่านี้มีค่าเท่ากับ 0 แปลความหมายว่าตัวแปรอิสระเข้าไปในสมการมาก ก็จะทำให้ค่า R^2 มากขึ้นด้วย ซึ่งนับข้อจำกัดของค่าสถิตินี้ โดยสามารถพิจารณารูปแบบสมการได้จากสมการที่ (2.17)

ดังนั้นเพื่อปรับปรุงข้อจำกัดดังกล่าวข้างต้น จึงเกิดค่าสถิติใหม่ คือค่า Adjusted R^2 ($\overline{R}^2$) ซึ่งจะมีการผกผันกันระหว่างตัวแปรที่เพิ่มเข้าไปกับค่า R^2 ที่ได้เพิ่มขึ้นมา ดังแสดงในสมการที่ (2.19)

$$R^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum y_i^2} \qquad (2.18)$$

$$\overline{R}^2 = 1 - \frac{\sum \hat{u}_i^2/(n-k)}{\sum y_i^2(n-1)} \qquad (2.19)$$

- Akaike's Information Criterion (AIC) คือค่าสถิติที่ประยุกต์คล้ายกับ $\overline{R}^2$ แต่ใช้รูปแบบการใส่ค่าลอกาลิทึมฐานธรรมชาติ (natural logarithm) โดยหากค่าสถิตินี้มีค่าน้อยเพียงใดนั้นก็แปลว่าแบบจำลองที่ประมาณได้นั้นสามารถเป็นตัวแทนข้อมูลจริงได้ดีเพียงนั้น ทั้งนี้ค่าสถิตินี้เหมาะที่จะนำไปใช้ในการหาค่าย้อนหลัง (lag length) ที่เหมาะสมอีกด้วย

$$AIC = \left(\frac{2k}{n}\right) + \log\left(\frac{\sum \hat{u}_i^2}{n}\right) \qquad (2.20)$$

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

N = ค่าสังเกตทั้งหมด

- ค่า Schwarz's Bayesian Information Criterion (SBC) คือ วิธีการวัดปรับได้อย่างดี (Goodness of Fit) เป็นวิธีที่ประยุกต์คล้ายกับ Akaike's Information Criterion (AIC) การพิจารณาค่า SBC นั้น ถ้าหากพบว่าค่า SBC มีค่าน้อยเท่าใดแล้วแสดงว่าแบบจำลองนั้นสามารถใช้เป็นตัวแทนของข้อมูลจริงได้อย่างเหมาะสม ซึ่งมีรูปสมการดังนี้

$$SBC = \log\left(\frac{\sum \hat{u}_i^2}{n}\right) + \left(\frac{2k \log n}{n}\right) \qquad (2.21)$$

จากค่าสถิติข้างต้นทั้งหมดจะนำมาใช้ประกอบในการพิจารณาเลือกแบบจำลอง ARIMA(p,d,q) ที่เหมาะสมที่สุด โดยจะคัดเลือกแบบจำลองในขั้นตอนนี้ไว้ 3 – 4 แบบจำลองเพื่อการเลือกอีกครั้งเพื่อที่จะเปรียบเทียบว่าแบบจำลองใดจะมีความสามารถในการพยากรณ์มากที่สุด

2) การประมาณค่าพารามิเตอร์ (Parameter estimation)

การประมาณค่าสัมประสิทธิ์ที่มาจากรูปแบบการถดถอยในตัวเอง (autoregressive; AR : p) และรูปแบบการเคลื่อนที่ของค่าคลาดเคลื่อน (moving average; MA : q) โดยสามารถเลือกใช้วีธีการถดถอยเชิงเส้นอย่างง่าย (simple least square) แต่สามารถที่จะใช้วีธีการถดถอยแบบไม่เป็นเชิงเส้น (nonlinear) ได้เพื่อสร้างความสัมพันธ์ของสมการที่จะสามารถนำไปใช้ในการพยากรณ์หากรูปแบบความสัมพันธ์นั้นเป็นรูปแบบที่มีความเหมาะสมที่สุด

3) การตรวจสอบแบบจำลอง (diagnostics)

เมื่อกำหนดรูปแบบและประมาณค่าพารามิเตอร์ในแบบจำลองจะต้องตรวจสอบทุกครั้งว่ารูปแบบที่กำหนดนั้นมีความเหมาะสมจริงหรือไม่ การตรวจสอบสามารถจะทำได้หลายวิธี ยกตัวอย่างเช่น การพิจารณาคอเรลโลแกรมของอัตสหสัมพันธ์ของกลุ่มตัวอย่าง (r_k) แต่อย่างไรก็ตาม Gujarati ได้เสนอการทดสอบวิเคราะห์ความเหมาะสมของแบบจำลองโดยใช้การทดสอบของ Box – pierce ซึ่งแสดงได้โดยใช้ Q-statistic

$$\text{Q - statistic} \quad = \quad n\sum_{k=1}^{m}\rho_k^2 \qquad (2.22)$$

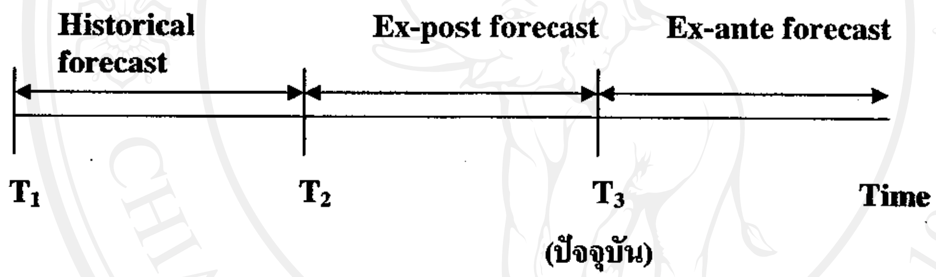
กำหนดให้ n = จำนวนของข้อมูล

m = ค่า lag length

จากสมการ ค่า Q - statistic ของแบบจำลองไม่แตกต่างกัน นั้นจะพบว่ามีการแจกแจงเป็นแบบ chi – square ที่มีดีกรีเท่ากับ m ซึ่งอยู่ภายใต้ข้อสมมุติฐานว่าง สมมุติฐานว่าง คือค่าความคลาดเคลื่อนที่ได้จากการประมาณมีลักษณะเป็น white noise หรือ ε_t มีการกระจายแบบปกติ (normal distribution) มีค่าเฉลี่ยเท่ากับ 0 และมีค่าความแปรปรวนเท่ากับ σ^2I (ε_t ~N(0, σ^2I) แสดงว่า ε_t มีลักษณะปราศจากอัตสหสัมพันธ์ (autocorrelation) ดังนั้นหากตรวจสอบพบว่าแบบจำลองนั้นปราศจากอัตสหสัมพันธ์แล้ว จะใช้แบบจำลองนั้นในการพยากรณ์ต่อไป แต่หากแบบจำลองนั้นไม่เหมาะสมต้องทำตามขั้นตอนที่ 1 เพื่อกำหนดรูปแบบจำลองใหม่

4) การพยากรณ์ (forecasting)

เมื่อได้แบบจำลองที่เหมาะสมภายหลังจากการวิเคราะห์ความถูกต้องแล้ว ก็สามารถนำแบบจำลองใช้ในการพยากรณ์ แต่เนื่องจากการพยากรณ์ข้อมูลไปข้างหน้านั้นจะต้องใช้แบบจำลองที่ให้ค่าประมาณที่แม่นยำที่สุด ดังนั้นการพยากรณ์จึงต้องมีการทดสอบโดยการแบ่งการพยากรณ์ออกเป็น 3 ช่วง คือ Historical forecast (T_1-T_2) และช่วง Ex-post forecast (T_2-T_3) เป็นการพยากรณ์ตั้งแต่อดีตจนถึงช่วงเวลาที่พิจารณา การพยากรณ์แล้วเปรียบเทียบข้อมูลจริงกับข้อมูลที่ได้จากการพยากรณ์ โดยพิจารณาค่า root mean square error (RMSE) และ Theil's inequality coefficient (U) และค่า Akaike Information Criterion (AIC) จะพิจารณาค่าสถิติทั้งสามค่าที่มีค่าน้อยที่สุดซึ่งได้จากการทำการพยากรณ์เมื่อเลือกรูปแบบจำลองที่ดีที่สุดแล้วจึงนำแบบจำลองนั้นมาทำการพยากรณ์แบบ Ex-ante forecast ซึ่งเป็นการพยากรณ์ข้อมูลไปข้างหน้าดังรูป



รูปที่ 2.6 ช่วงเวลาที่ใช้ในการพยากรณ์
ที่มา: Pindyck and Rubinfeld (1998)

2.2 สรุปสาระสำคัญจากเอกสารที่เกี่ยวข้อง

เพียงจันทร์ ลิขิตเอกราช (2535) ศึกษาเรื่องการกระจายรายได้จากการท่องเที่ยว กรณีศึกษางานมหกรรมไม้ดอกไม้ประดับในจังหวัดเชียงใหม่ โดยมีวัตถุประสงค์เพื่อศึกษาแบบแผนค่าใช้จ่ายของนักท่องเที่ยวของเชียงใหม่ การกระจายรายได้จากการท่องเที่ยวไปสู่ธุรกิจขนาดใหญ่ ขนาดกลางหรือ ขนาดเล็ก ที่รองรับการท่องเที่ยว รวมทั้งผู้ใช้แรงงานในธุรกิจเหล่านั้น และต้นทุนระหว่างรัฐและเอกชน และสัดส่วนรายได้ (cost/income sharing) จากงานมหกรรมไม้ดอกไม้ประดับของจังหวัดเชียงใหม่ ที่ตกแก่ภาคเกษตร ภาคอุตสาหกรรมและภาคบริการ โดยทำการศึกษา ศึกษาพฤติกรรมและแบบแผนค่าใช้จ่ายของนักท่องเที่ยวทั้งชาวไทยและต่างประเทศที่มาเที่ยวงานไม้ดอกไม้ประดับจังหวัดเชียงใหม่ 2 ช่วงคือ ช่วงที่จัดงานเทศกาล และช่วงปกติที่ไม่

มีงานเทศกาล โดยใช้แบบสอบถามเก็บข้อมูลนักท่องเที่ยว พบว่าการจัดการงานไม้ดอกไม้ประดับในช่วงวันที่ 7-9 กุมภาพันธ์ 2535 ก่อให้เกิดรายได้แก่จังหวัดเชียงใหม่ 105.28 ล้านบาทและเกิดรายได้เพิ่มจากปกติ 45 ล้านบาท โดยจะกระจายไปสู่ธุรกิจประเภทพาหนะเดินทางมากที่สุดประมาณ 29.3 ล้านบาท ส่วนธุรกิจประเภทขายของที่ระลึกและธุรกิจที่พัก ก็ได้รับรายได้สูงเช่นกัน ตัวแปรที่มีผลต่อการใช้จ่ายของนักท่องเที่ยว คือตัวแปรด้านรายได้ และตัวแปรด้านประชากร

ขวัญฤทัย บุญยะเสนา (2544) ศึกษาเรื่องการวิเคราะห์อุปสงค์การท่องเที่ยวเชิงนิเวศในอำเภออุ้มผางจังหวัดตาก โดยมีวัตถุประสงค์เพื่อศึกษาถึงอุปสงค์ของการท่องเที่ยวเชิงนิเวศในอำเภออุ้มผาง จังหวัดตาก และศึกษาถึงนโยบาย มาตรการและงบประมาณของภาครัฐในการส่งเสริมท่องเที่ยว ในอำเภออุ้มผาง จังหวัดตากและแนวทางแก้ไขปัญหาดังกล่าว โดยทำการศึกษาอุปสงค์การท่องเที่ยวเชิงนิเวศในอำเภออุ้มผาง จังหวัดตาก โดยใช้แบบสอบถามสำรวจความคิดเห็นจากนักท่องเที่ยวที่เดินทางไปท่องเที่ยวที่ อำเภออุ้มผาง จังหวัดตาก ผลการศึกษาพบว่านักท่องเที่ยวที่มามีวัตถุประสงค์เพื่อพักผ่อน พักฟื้น ส่วนใหญ่มากับกลุ่มเพื่อน ใช้รถยนต์เป็นพาหนะ จำนวนวันค้างแรมเฉลี่ย 3 วัน ค่าใช้จ่ายเฉลี่ย 1,484.11 บาทต่อคนต่อวันความสัมพันธ์ของอุปสงค์ของการท่องเที่ยว (จำนวนวันพัก) กับอายุ มีทิศทางตรงกันข้าม แต่ความสัมพันธ์ของอุปสงค์ของการท่องเที่ยว (จำนวนวันพัก) กับจำนวนวันหยุดในรอบปี และจำนวนครั้งที่เคยมาเที่ยวมีทิศทางเดียวกัน

สุพรรณา หัศภาค (2545) ศึกษาเรื่องปัจจัยที่ดึงดูดนักท่องเที่ยวชาวต่างประเทศให้มาท่องเที่ยวจังหวัดเชียงใหม่ โดยมีวัตถุประสงค์เพื่อศึกษาวิเคราะห์ถึงปัจจัยที่ดึงดูดนักท่องเที่ยวชาวต่างประเทศให้เดินทางท่องเที่ยวจังหวัดเชียงใหม่ โดยทำการศึกษาเฉพาะนักท่องเที่ยวชาวต่างประเทศที่มาเยือนและพักค้างคืนในจังหวัดเชียงใหม่ โดยกำหนดปัจจัยทั้งหมด 6 หมวดคือ 1) ปัจจัยด้านศิลปวัฒนธรรม 2) ปัจจัยด้านแหล่งท่องเที่ยวทางธรรมชาติ 3) ปัจจัยด้านการบริการและอัธยาศัยไมตรี 4) ปัจจัยด้านความปลอดภัย 5) ปัจจัยด้านค่าใช้จ่าย และ 6) ปัจจัยอื่นๆ โดยใช้ข้อมูลปฐมภูมิจากการเก็บแบบสอบถาม ส่วนข้อมูลทุติยภูมิจากเอกสารของหน่วยงานราชการ โดยพบว่าปัจจัยด้านศิลปวัฒนธรรมเป็นปัจจัยอันดับแรก รองลงมาคือปัจจัยด้านท่องเที่ยว ปัจจัยด้านการให้บริการและอัธยาศัยไมตรี ปัจจัยด้านความปลอดภัย ปัจจัยด้านค่าใช้จ่าย และปัจจัยอื่นๆ ตามลำดับ

อนัสปรีย์ ไชยวรรณ (2546) ศึกษาเรื่องการประยุกต์ใช้โคอินทิเกรชันและแบบจำลองเอเรอร์คอร์เรคชันกับอัตราแลกเปลี่ยนเงินตราต่างประเทศในภูมิภาคเอเชีย โดยมีวัตถุประสงค์เพื่อประยุกต์ใช้โคอินทิเกรชันและแบบจำลองเอเรอร์คอร์เรคชันกับอัตราแลกเปลี่ยนเงินตราต่างประเทศในภูมิภาคเอเชีย และศึกษาถึงปัจจัยที่มีอิทธิพลต่ออัตราแลกเปลี่ยนเงินตราต่างประเทศของแต่ละประเทศในภูมิภาคเอเชีย โดยทำการศึกษา ศึกษาเฉพาะกลุ่มประเทศเอเชีย 6 ประเทศ คือ ไทย ญี่ปุ่น เกาหลีใต้ มาเลเซีย ฟิลิปปินส์ และสิงคโปร์ โดยใช้เทคนิคโคอินทิเกรชันและแบบจำลองเอเรอร์คอร์เรคชัน โดยใช้ข้อมูลตั้งแต่เดือนกรกฎาคม 2540 ถึง มิถุนายน 2545 พบว่าดัชนีราคาผู้บริโภคเป็นปัจจัยหลัก และอัตราแลกเปลี่ยนมีความสัมพันธ์ในระยะสั้นกับปริมาณเงินโดยเปรียบเทียบ และรายได้ประชาชาติแท้จริงโดยเปรียบเทียบ

ฉัตรสุดา อู้เฉื้อง (2546) ศึกษาเรื่องการวิเคราะห์ปัจจัยที่มีผลต่อการท่องเที่ยวโบราณสถานเวียงกุมกาม โดยมีวัตถุประสงค์เพื่อศึกษาถึงปัจจัยที่ส่งผลกระทบต่อการท่องเที่ยวโบราณสถานเวียงกุมกาม วัดช้างค้ำ วัดกานโถม วัดเจดีย์เหลี่ยมและวัดอีกางอำเภอสารภี จังหวัดเชียงใหม่ โดยทำการศึกษาโดยใช้ข้อมูลจากนักท่องเที่ยวที่มาเที่ยวโบราณสถานเวียงกุมกาม เก็บข้อมูลโดยใช้แบบสอบถามเพื่อวิเคราะห์หาปัจจัยที่มีผลต่อการท่องเที่ยวโบราณสถานเวียงกุมกาม จำนวนแบบสอบถาม 300 ชุด ในช่วงวันหยุดสุดสัปดาห์และวันหยุดนักขัตฤกษ์ในเดือนมีนาคม – เดือนเมษายน 2546 ผลการศึกษาพบว่าปัจจัยที่สำคัญคือ ปัจจัยด้านราคา ปัจจัยด้านร้านอาหาร ปัจจัยด้านของที่ระลึก

จิตราภรณ์ ฝั้นศิริ (2547) ศึกษาเรื่องการพยากรณ์ราคาส่งออกข้าวโดยวิธีอารีมา โดยมีวัตถุประสงค์เพื่อศึกษาความเคลื่อนไหวและพยากรณ์ราคาส่งออกข้าว (ราคา FOB) ของประเทศเทศไทย ชนิดข้าวขาว 100% ชั้น 2 ด้วยวิธีอารีมา โดยทำการศึกษาใช้ข้อมูลราคาส่งออกข้าว (FOB) ของไทย ชนิดข้าวขาว 100% ชั้น 2 เป็นรายเดือนตั้งแต่เดือน มกราคม พ.ศ. 2531 ถึงเดือนธันวาคม พ.ศ. 2546 รวมจำนวน 192 ตัวอย่าง โดยใช้วิธีอารีมา ผลการพยากรณ์ต่อไปข้างหน้า ผลการทดสอบ unit root ที่ระดับ level สัมประสิทธิของ lag length ที่ P-lag เท่ากับ 1 ค่า t-statistic แตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ 1% ทั้งแบบจำลอง without intercept and trend,with intercept and without trend และwith intercept and with trend ค่า ADF test-statistic ที่ระดับ level เปรียบเทียบกับค่าวิกฤตของ Mackinnon ที่ระดับ 1% ทั้ง 3 แบบจำลองมีค่าไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ(ยอมรับ H0) หมายถึงว่ามี unit root ต้องทำ difference สรุปได้ค่าตัวแปรมีลักษณะเป็น I(1) และผลการวิเคราะห์แบบจำลอง ARIMA โดยวิธี Box-Jenkins และการตรวจสอบ

ความถูกต้อง ได้ค่าผล ราคาส่งออกข้าวขาว 100% ชั้น 2 ตั้งแต่เดือนมกราคม ถึงเดือนเมษายน 2547 มีค่าเท่ากับ 205, 204, 201 และ 201 ตามลำดับ

อัญชลี นัสสาสาร (2548) ศึกษาเรื่องการวิเคราะห์รายจ่ายของนักท่องเที่ยวในเทศกาลลอยกระทงจังหวัดเชียงใหม่ โดยมีวัตถุประสงค์เพื่อศึกษาลักษณะต่างๆ ของนักท่องเที่ยวที่มีผลต่อค่าใช้จ่ายทั้งในและนอกช่วงเทศกาลลอยกระทงของนักท่องเที่ยว และศึกษาการกระจายรายได้จากการท่องเที่ยวไปสู่กลุ่มธุรกิจต่างๆที่รองรับการท่องเที่ยวในเทศกาลลอยกระทง โดยทำการศึกษาศึกษาเทศกาลลอยกระทงในจังหวัดเชียงใหม่ในปี พ.ศ. 2546 โดยใช้แบบสอบถามโดยการสัมภาษณ์ตัวอย่าง 2 กลุ่ม 1) นักท่องเที่ยวทั้งในและนอกเทศกาลลอยกระทงที่มาร่วมงานลอยกระทงในจังหวัดเชียงใหม่ ปี พ.ศ. 2546 จำนวน 400 ตัวอย่าง แบ่งเป็นช่วงเทศกาล 200 ตัวอย่าง และนอกช่วงเทศกาล 200 ตัวอย่าง 2) สัมภาษณ์ผู้ประกอบการโรงแรมและที่พัก 30 ราย การวิเคราะห์ข้อมูลโดยหาค่าเฉลี่ยและเปรียบเทียบสัดส่วนร้อยละ และวิเคราะห์แบบจำแนกพหุ (Multiple Classificaton Analysis : MCA) จากการศึกษาพบว่า ค่าใช้จ่ายนักท่องเที่ยวนอกเทศกาล 2,540.67 บาทต่อคนต่อวัน โดยค่าใช้จ่ายส่วนใหญ่เป็นค่าพาหนะรองลงมาคือค่าที่พัก ค่าซื้อสินค้า ค่าของที่ระลึก ตามลำดับ ส่วนนักท่องเที่ยวในเทศกาล ค่าใช้จ่าย 2,149.21 บาทต่อคนต่อวัน โดยค่าใช้จ่ายส่วนใหญ่เป็นค่าพาหนะรองลงมาคือค่าที่พัก ค่าซื้อสินค้า ค่าของที่ระลึก ตามลำดับ

ภารตี ไกรสิทธิ์ (2548) ศึกษาเรื่องการศึกษาโครงสร้าง พฤติกรรม และผลการดำเนินงานของอุตสาหกรรมท่องเที่ยวในจังหวัดเชียงใหม่ โดยมีวัตถุประสงค์เพื่อศึกษาโครงสร้างของอุตสาหกรรมท่องเที่ยวในจังหวัดเชียงใหม่ ในด้านการกระจุกตัวของอุตสาหกรรมบริการท่องเที่ยวโดยพิจารณา อุตสาหกรรมโรงแรม รีสอร์ท และเกสท์เฮาส์ และอุตสาหกรรมบริการนำเที่ยว ผลการดำเนินงาน ความสามารถในการทำกำไรของหน่วยธุรกิจในอุตสาหกรรมท่องเที่ยว 2 หน่วยธุรกิจระหว่างอุตสาหกรรมโรงแรม รีสอร์ท และเกสท์เฮาส์ และอุตสาหกรรมบริการนำเที่ยว ประสิทธิภาพของการดำเนินงานของหน่วยธุรกิจในอุตสาหกรรมท่องเที่ยวของจังหวัดเชียงใหม่ และมูลค่าเพิ่มของอุตสาหกรรมท่องเที่ยวจังหวัดเชียงใหม่ โดยทำการศึกษา ใช้ข้อมูลในปี พ.ศ. 2545 โดยศึกษาโครงสร้างของอุตสาหกรรมท่องเที่ยวทำการศึกษาในหน่วยธุรกิจ 2 หน่วยธุรกิจของจังหวัดเชียงใหม่ระหว่างอุตสาหกรรมโรงแรม รีสอร์ท และเกสท์เฮาส์ และอุตสาหกรรมบริการนำเที่ยว ที่มีการจดทะเบียนนิติบุคคลกับสำนักงานธุรกิจการค้าจังหวัดเชียงใหม่ ซึ่งมีการดำเนินงานในปี พ.ศ. 2545 จำนวน 200 บริษัท โดยใช้การศึกษาโครงสร้างด้วิวิธีวัดการกระจุกตัว ส่วนการศึกษาผลการดำเนินงานใช้อัตราส่วนทางการเงินที่สำคัญๆและดัชนีมูลค่าเพิ่ม ส่วนการศึกษาประสิทธิภาพของอุตสาหกรรม วัดประสิทธิภาพโดยเปรียบเทียบด้วยวิธี Data

Environment Analysis (DEA) ผลการศึกษาพบว่าทั้งอุตสาหกรรมที่พักและบริการนำเที่ยวในจังหวัดเชียงใหม่ มีการกระจุกตัวสูง นั่นคือมีการแข่งขันกันต่ำ การวิเคราะห์อัตราส่วนทางการเงินพบว่าอุตสาหกรรมประสบปัญหาหนี้สินมากโดยเฉพาะโรงแรม รีสอร์ทและเกสท์เฮาท์ขนาดใหญ่จะมีปัญหาด้านสภาพคล่อง การทำกำไรและการชำระหนี้ แต่กิจการที่มีชาวต่างชาติเป็นหุ้นส่วนจะมีปัญหาทางด้านสภาพคล่อง แต่มีประสิทธิภาพในการทำกำไรและชำระหนี้ อุตสากรรมนำเที่ยวทั้งที่เป็นของคนไทยและมีชาวต่างชาติเป็นหุ้นส่วน ต่างก็ประสบปัญหาเช่นเดียวกับอุตสาหกรรมที่พัก