

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ในการวิจัยเรื่องการเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่คงที่และไม่คงที่จากโมเดลอริมา โมเดลอริมาอินเตอร์เวนชัน และโมเดลการถดถอยนั้น ผู้วิจัยได้ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้องและนำเสนอวรรณคดีที่เกี่ยวข้องในบทนี้ โดยแยกนำเสนอเป็น 5 ตอน ดังนี้

- ตอนที่ 1 มโนทัศน์เกี่ยวกับเทคนิคการพยากรณ์และอนุกรมเวลา
- ตอนที่ 2 วิธีของบ็อกซ์และเจนกินส์และโมเดลอริมาอินเตอร์เวนชัน
- ตอนที่ 3 มโนทัศน์เกี่ยวกับการวิเคราะห์การถดถอย
- ตอนที่ 4 ความคลาดเคลื่อนในการพยากรณ์
- ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 มโนทัศน์เกี่ยวกับเทคนิคการพยากรณ์และอนุกรมเวลา

เทคนิคการพยากรณ์

เทคนิคการพยากรณ์มี 2 ประเภท คือ เทคนิคการพยากรณ์เชิงคุณภาพ (Qualitative Forecasting) และเทคนิคการพยากรณ์เชิงปริมาณ (Quantitative Forecasting) แต่ที่นิยมใช้คือเทคนิคการพยากรณ์เชิงปริมาณ ซึ่งแบ่งออกเป็น 2 ประเภทดังต่อไปนี้

1. เทคนิคอนุกรมเวลา (Time Series Analysis)

1.1 วิธีเฉลี่ยเคลื่อนที่ (Moving Average Method)

เป็นวิธีที่ใช้ในการพยากรณ์เหตุการณ์ล่วงหน้าระยะสั้น โดยสามารถทำการเฉลี่ยเคลื่อนที่ 3 เดือน 6 เดือน หรือ 1 ปี ฯลฯ แล้วแต่ค่าสังเกต ซึ่งวิธีนี้แบ่งออกเป็นหลายประเภท ดังนี้

1.1.1 วิธีเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average Method : SMA) เป็นวิธีการหาค่าพยากรณ์จากค่าสังเกตล่าสุดจำนวนหนึ่ง นิยมใช้วิธีนี้พยากรณ์เพียง 3 ค่า หรือ

5 ค่าลวงหน้าเท่านั้น การคำนวณทำได้จากนำค่าสังเกตมาหาค่าเฉลี่ยซึ่งการเฉลี่ยนี้เป็นการเฉลี่ยที่ให้น้ำหนักกับค่าสังเกตแต่ละค่าเท่ากัน กรณีที่ใช้ค่าสังเกต k ค่า จะให้ค่าพยากรณ์ ณ เวลา $t+1$ ดังนี้

$$\hat{Y}_{t+1} = \frac{Y_t + \dots + Y_{t-k+1}}{k}$$

จำนวนค่าสังเกตที่จะนำมาหาค่าเฉลี่ยนั้นจะมีจำนวนไม่แน่นอน จำนวนที่เหมาะสมคือจำนวนที่ทำให้ค่าพยากรณ์มีค่าใกล้เคียงกับค่าจริงมากที่สุด

1.1.2 วิธีเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก (Weighted Moving Average Method : WMA) โดยวิธี WMA จะเหมือนกับวิธี SMA จะต่างกันเพียงน้ำหนักที่ให้กับแต่ละค่าสังเกตจะต่างกัน กรณีที่ใช้ค่าสังเกต k ค่า จะให้ค่าพยากรณ์ ณ เวลา $t+1$ ดังนี้

$$\hat{Y}_{t+1} = w_1 Y_t + w_2 Y_{t-1} + \dots + w_k Y_{t-k+1}$$

เมื่อ w_i เป็นค่าถ่วงน้ำหนัก และ $\sum w_i = 1$

1.1.3 วิธีเฉลี่ยเคลื่อนที่สองครั้ง (Double Moving Average Method : DMA) โดยที่การเฉลี่ยเคลื่อนที่สองครั้งจะทำให้การเคลื่อนไหวของอนุกรมเวลาในการปรับครั้งที่สองเรียกว่าการเฉลี่ยเคลื่อนที่ครั้งแรก สมมติว่าอนุกรมเวลาเดิมเป็น $\{y_t\}$ อนุกรมเวลาใหม่ที่ได้จากการเฉลี่ยเคลื่อนที่ครั้งแรกเป็น $\{MA_t\}$ และอนุกรมเวลาที่ได้จากการเฉลี่ยเคลื่อนที่ครั้งที่สองเป็น $\{MA'_t\}$ ดังนั้นในการเฉลี่ยเคลื่อนที่ k ช่วงเวลา ค่า MA_t จะเป็นค่าประมาณแนวโน้ม ณ เวลา $t - (k - 1)/2$ หรือ $T_{t-(k-1)/2}$ ส่วนค่า MA'_t จะเป็นค่าประมาณแนวโน้ม ณ เวลา $t - (k - 1)$ หรือเป็น $T_{t-(k-1)}$ ซึ่ง

$$T_{t-(k-1)/2} = T_t - ((k-1)/2)\beta_1 = \beta_0 + \beta_1 (t - ((k-1)/2))$$

$$T_{t-(k-1)} = T_t - (k-1)\beta_1 = \beta_0 + \beta_1 (t - (k-1))$$

จะใช้ $MA_t - MA'_t$ ในการประมาณค่าดังนี้

$$T_{t-(k-1)/2} - T_{t-(k-1)} = ((k-1)/2)\beta_1$$

จะได้ $\hat{\beta}_1(t) = (2/(k-1))(MA_t - MA'_t)$

และใช้	$2MA_t - MA'_t$ ในการประมาณค่า $\beta_0 + \beta_1 t$
จะได้	$\hat{T}_t(t) = 2MA_t - MA'_t$
เมื่อ	$\hat{T}_t(t) =$ ค่าประมาณแนวโน้ม ณ เวลา t $=$ ค่าประมาณของ $\beta_0 + \beta_1 t$ ณ เวลา t $=$ ค่าประมาณของ T_t $\hat{\beta}_1(t) =$ ค่าประมาณของ β_1 ณ เวลา t $=$ อัตราการเพิ่มหรือลดลงของ Y เมื่อเวลาเพิ่มขึ้น 1 ช่วงเวลา ณ เวลา t

ดังนั้นสมการพยากรณ์ ณ เวลา $t + p$ เมื่อใช้ค่าสังเกต t ค่าคือ

$$\hat{Y}_{t+p}(t) = T_t(t) + p\beta_1(t)$$

1.1.4 วิธีเฉลี่ยเคลื่อนที่ของเปอร์เซ็นต์การเปลี่ยนแปลง (Moving Average of Percentage Change Method : MAPC) โดยข้อมูลอนุกรมเวลาจะมีแนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential) ซึ่งเป็นลักษณะของแนวโน้มที่มีการเพิ่มขึ้นหรือลดลงเป็นเปอร์เซ็นต์ที่เท่ากันในแต่ละช่วงเวลา นั่นคือเมื่อ

$$\begin{aligned}
 PC_t &= \text{อัตราการเปลี่ยนแปลง ณ เวลา } t \\
 &= (Y_t - Y_{t-1})/Y_{t-1} \times 100 \\
 &= (Y_t - Y_{t-1}) \times 100
 \end{aligned}$$

$$\begin{aligned}
 MPC_t &= \text{ค่าเฉลี่ยเคลื่อนที่ } k \text{ ช่วงเวลา ณ เวลา } t \\
 &= \sum(PC_{t+k}/k)
 \end{aligned}$$

สมการพยากรณ์ ณ เวลา t สำหรับ p ช่วงเวลาล่วงหน้าคือ

$$\hat{Y}_{t+p}(t) = Y_t \left(1 + \left(\frac{MPC_t}{100} \right) \right)^p \quad \text{สำหรับ } p = 1, 2, \dots$$

1.2 วิธีปรับให้เรียบ (Smoothing Method)

เป็นวิธีที่ใช้ในการพยากรณ์เหตุการณ์ล่วงหน้าระยะสั้น และเหมาะสำหรับพยากรณ์ค่าของตัวแปรที่มักจะมีการเปลี่ยนแปลงไม่มากนักในหนึ่งหน่วยเวลา ซึ่งวิธีนี้แบ่งออกเป็นหลายประเภท ดังนี้

1.2.1 วิธีปรับให้เรียบเอ็กซ์โปเนนเชียลแบบง่าย (Simple Exponential Smoothing : SES) ค่าพยากรณ์จะใช้ค่าสังเกตที่ผ่านมาทั้งหมด โดยน้ำหนักที่ให้กับค่าสังเกตแต่ละค่าไม่เท่ากัน น้ำหนักที่ให้กับค่าที่เกิดขึ้นล่าสุดจะมากและลดหลั่นไปสำหรับค่าสังเกตที่อยู่ห่างออกไป ซึ่งการลดหลั่นของน้ำหนักนี้จะเป็นการลดแบบเอ็กซ์โปเนนเชียล (exponential) น้ำหนักจะเป็นเท่าไรนั้นขึ้นอยู่กับค่าปรับน้ำหนัก (smoothing constant : α) ที่มีค่าอยู่ระหว่าง 0 ถึง 1 สมการพยากรณ์สำหรับการพยากรณ์ ณ เวลา t เมื่อใช้ค่าสังเกต Y_t, Y_{t-1}, \dots มีรูปแบบดังนี้

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1 - \alpha)Y_{t-1} + \alpha(1 - \alpha)^2 Y_{t-2} + \dots$$

โดยที่ค่าถ่วงน้ำหนักทั้งหมดจะรวมกันเป็น 1 อาจเขียนสมการพยากรณ์ใหม่ได้อีก 2 แบบ ได้แก่ แบบปรับให้เรียบ (smoothing form) และแบบปรับค่าเดิมด้วยความคลาดเคลื่อน (error correction form) ดังนี้

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)Y_t$$

และ

$$\hat{Y}_{t+1} = Y_t + \alpha e_t$$

สมการพยากรณ์แบบปรับให้เรียบจะได้จากการเฉลี่ยแบบถ่วงน้ำหนักของค่าจริงและค่าพยากรณ์ ณ เวลา t ส่วนสมการพยากรณ์แบบปรับค่าเดิมด้วยความคลาดเคลื่อนจะได้จากการปรับค่าพยากรณ์ ณ เวลา t ด้วย α เท่าของความคลาดเคลื่อน ณ เวลา t

1.2.2 วิธีปรับให้เรียบเอ็กซ์โปเนนเชียลแบบดับเบิล (Double Exponential Smoothing : DES) หลักการของวิธี DES จะคล้ายกับวิธี DMA เพียงแต่ใช้ A_t แทน MA_t และ A'_t แทน MA'_t โดย

$$A_t = \text{ค่าปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ณ เวลา } t$$

$$A'_t = \text{ค่าปรับให้เรียบครั้งที่สองแบบเอ็กซ์โปเนนเชียล ณ เวลา } t$$

ซึ่ง A_t หาได้จากการเฉลี่ยถ่วงน้ำหนักระหว่าง A_{t-1} และ Y_t และ A'_t หาได้จากการเฉลี่ยถ่วงน้ำหนักระหว่าง A'_{t-1} และ A_t นั่นคือ

$$A_t = (1 - \alpha)A_{t-1} + \alpha Y_t$$

$$A'_t = (1 - \alpha)A'_{t-1} + \alpha A_t$$

จะได้

$$\hat{T}_t(t) = \text{ค่าแนวโน้ม ณ เวลา } t = 2A_t - A'_t$$

$$\hat{\beta}_1(t) = \text{ค่าประมาณของ } \beta_1 \text{ ณ เวลา } t = (\alpha/(1 - \alpha))(A_t - A'_t)$$

1.2.3 วิธีปรับให้เรียบเอ็กซ์โปเนนเชียลแบบเส้นตรง (Linear Exponential Smoothing : LES) มีวิธีการคล้ายกับวิธี DES เมื่อ

$$\hat{T}_t(t) = \text{ค่าแนวโน้ม ณ เวลา } t$$

$$\hat{\beta}_1(t) = \text{ค่าประมาณของ } \beta_1 \text{ ณ เวลา } t$$

และสมการพยากรณ์ ณ เวลา $t+1$ เป็น

$$\hat{Y}_{t+1}(t) = \hat{T}_t(t) + \hat{\beta}_1(t)$$

ในการปรับค่าแนวโน้มและค่าประมาณของ β_1 จะใช้สมการปรับที่ขึ้นอยู่กับค่าปรับน้ำหนัก 2 ค่า ได้แก่ α และ γ ซึ่งต่างก็มีค่าอยู่ระหว่าง 0 ถึง 1

$$\hat{T}_t(t) = (1 - \alpha)\hat{Y}_t(t) + \alpha Y_t$$

$$\hat{\beta}_1(t) = (1 - \gamma)\hat{\beta}_1(t-1) + \gamma(\hat{T}_t(t) - \hat{T}_{t-1}(t-1))$$

จะได้สมการพยากรณ์ ณ เวลา t พยากรณ์ p ช่วงเวลาล่วงหน้าเป็น

$$\hat{Y}_{t+p}(t) = \hat{T}_t(t) + p\hat{\beta}_1(t) \quad \text{สำหรับ } p = 1, 2, \dots$$

1.2.4 วิธีปรับให้เรียบเอ็กซ์โปเนนเชียลแบบฤดูกาล (Simple Seasonal Exponential Smoothing : SSES) สมการพยากรณ์จะสร้างขึ้นโดยใช้หลักการปรับให้เรียบ มีค่าปรับน้ำหนัก 2 ค่า สำหรับค่าคงที่และค่าวัดฤดูกาล ค่าปรับน้ำหนักจะมีค่าอยู่ระหว่าง 0 ถึง 1 เมื่ออนุกรม

เวลามีรูปแบบต่างกัน นั่นคือมีรูปแบบบวกหรือรูปแบบคูณ การปรับให้เรียบจะใช้หลักการที่ไม่ต่างกันนัก โดยที่สมการพยากรณ์มีรูปแบบดังนี้

$$\begin{aligned} \text{รูปแบบบวก} \quad \hat{Y}_{t+p}(t) &= \hat{T}_t(t) + \hat{S}_{t+p}(t) \\ \text{รูปแบบคูณ} \quad \hat{Y}_{t+p}(t) &= \hat{T}_t(t) \hat{S}_{t+p}(t) \end{aligned}$$

1.2.5 วิธีปรับให้เรียบเอ็กซ์โปเนนเชียลของ Holt และ Winters (Holt – Winters Exponential Smoothing : HWS) สมการพยากรณ์จะประกอบด้วยส่วนของแนวโน้มและส่วนของฤดูกาลที่สร้างขึ้นโดยใช้หลักการของการปรับให้เรียบ การปรับน้ำหนัก 3 ค่า สำหรับค่าจุดตัดแกน Y ค่าวัดความลาดชันและค่าวัดอิทธิพลของฤดูกาล ค่าปรับน้ำหนักจะมีค่าอยู่ระหว่าง 0 ถึง 1 เมื่ออนุกรมเวลามีรูปแบบต่างกัน นั่นคือมีรูปแบบบวกหรือรูปแบบคูณ การปรับให้เรียบจะใช้หลักการที่ไม่ต่างกันมากนัก โดยที่สมการพยากรณ์มีรูปแบบดังนี้

$$\begin{aligned} \text{รูปแบบบวก} \quad \hat{Y}_{t+p}(t) &= \hat{T}_{t+p}(t) + \hat{S}_{t+p}(t) \\ \text{รูปแบบคูณ} \quad \hat{Y}_{t+p}(t) &= \hat{T}_{t+p}(t) \hat{S}_{t+p}(t) \end{aligned}$$

1.3 วิธีแยกส่วนประกอบ (Decomposition Method)

เป็นวิธีที่ใช้ในการพยากรณ์เหตุการณ์ล่วงหน้าระยะสั้น ซึ่งเป็นการวิเคราะห์อนุกรมเวลาที่ระบุงค์ประกอบ 4 องค์ประกอบที่มีอิทธิพลในแต่ละช่วงเวลาของอนุกรม คือ การเปลี่ยนแปลงที่แสดงแนวโน้ม (trend) การเปลี่ยนแปลงตามวัฏจักร (cyclical) การเปลี่ยนแปลงเนื่องจากฤดูกาล (seasonal) และการเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ (irregular) แล้วนำองค์ประกอบที่ได้ไปใช้พยากรณ์ ค่าพยากรณ์ในอนาคตจะได้จากการรวมค่าการวัดส่วนประกอบของอนุกรมเวลา โดยรายละเอียดเกี่ยวกับส่วนประกอบของอนุกรมเวลาและการรวมตัวของอนุกรมเวลา มีดังนี้

ส่วนประกอบของอนุกรมเวลา

การเคลื่อนไหวของอนุกรมเวลาขึ้นอยู่กับส่วนประกอบต่างๆ ซึ่งส่วนประกอบหลักของอนุกรมเวลา มีดังนี้ (เอกชัย ชัยประเสริฐสิทธิ, 2525 ; ทรงศิริ แต่สมบัติ, 2539)

1. แนวโน้ม (Secular Trend หรือ T) หมายถึง การเคลื่อนไหวของอนุกรมเวลาในระยะยาวซึ่งอาจจะเป็นแนวโน้มขึ้นหรือลง (upward หรือ downward trend) ค่าแนวโน้มเป็นการ

เคลื่อนไหวในระยะเวลาที่ค่อนข้างจะยาวนาน ค่าแนวโน้มอาจมีลักษณะเป็นเส้นตรง เส้นโค้ง หรือลักษณะใดก็ได้

2. อิทธิพลของฤดูกาล (Seasonal Variation หรือ S) หมายถึง การเคลื่อนไหวของอนุกรมเวลาที่เกิดขึ้นซ้ำแล้วซ้ำอีกในช่วงเวลาหนึ่ง หน่วยของระยะเวลาอาจเป็นรายสัปดาห์ รายเดือน แต่ส่วนใหญ่จะเป็น 1 ปี ปัจจัยที่ก่อให้เกิดอิทธิพลของฤดูกาลมีหลายปัจจัย เช่น สภาพอากาศ อุณหภูมิ วัฒนธรรม

3. อิทธิพลของวัฏจักร (Cyclical Variation หรือ C) วัฏจักรหนึ่งจะครอบคลุมระยะเวลาหลายปี วัฏจักรที่พบเสมอได้แก่ วัฏจักรธุรกิจ (business cycle) แบบแผนวัฏจักรธุรกิจจะแบ่งออกเป็น 3 ช่วง ได้แก่ ช่วงรุ่งเรือง (growth) ช่วงคงที่ (maturity) และช่วงตกต่ำ (decline) ซึ่งลักษณะพิเศษของการเคลื่อนไหวแบบนี้คือการพุ่งสูงขึ้นช่วงระยะเวลาหนึ่งและตามด้วยการตกต่ำช่วงระยะเวลาหนึ่ง

4. เหตุการณ์ที่ผิดปกติ (Irregular Variation หรือ I) เป็นการเคลื่อนไหวของอนุกรมเวลาเฉพาะส่วนที่ไม่มีแผนแบบที่แน่นอน เหตุการณ์ผิดปกตินี้ส่วนใหญ่จะเป็นเหตุการณ์ที่ไม่ได้คาดคิดมาก่อนหรือไม่เกิดบ่อยครั้ง เช่น การนัดหยุดงานของคนงาน ไปจนถึงเรื่องใหญ่ๆ เช่น สงครามโลก เหตุการณ์เหล่านี้มิได้อยู่ภายใต้กฎที่แน่นอนอย่างใดอย่างหนึ่ง ซึ่งมักจะเกิดขึ้นตามโอกาสหรือโดยบังเอิญ จึงอาจเรียกอีกอย่างหนึ่งว่าการเปลี่ยนแปลงขึ้นลงที่เป็นเชิงสุ่ม (random fluctuation)

การรวมตัวของปัจจัยทั้ง 4 ประเภท ที่ประกอบขึ้นเป็นอนุกรมเวลามีลักษณะของการรวมตัวได้ 2 แบบ ดังนี้

แบบที่ 1 การรวมตัวในเชิงบวก หรือตัวแบบในเชิงบวก (Additive Model)

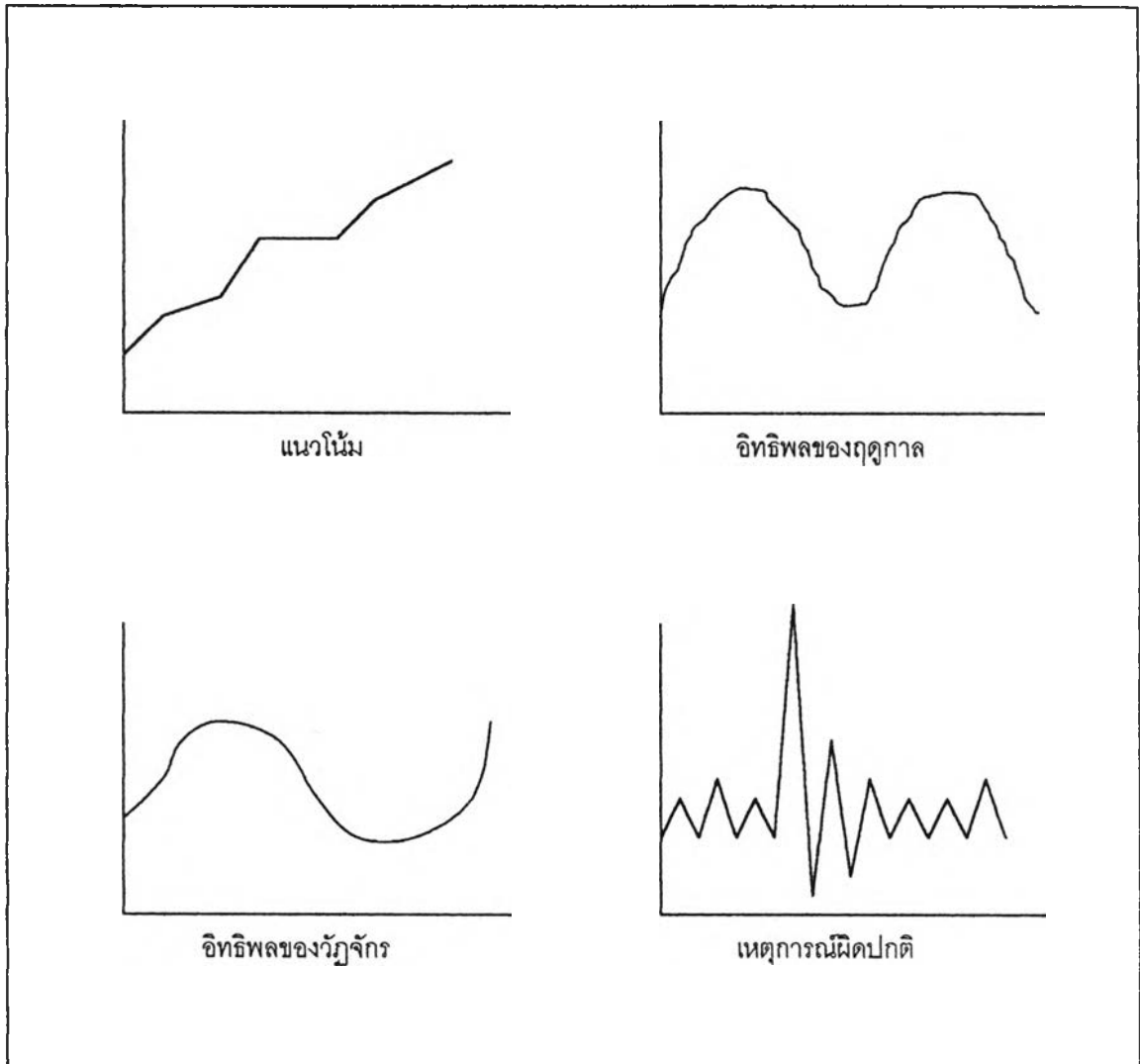
$$Y = T + S + C + I$$

แบบที่ 2 การรวมตัวในเชิงคูณ หรือตัวแบบในเชิงคูณ (Multiplicative Model)

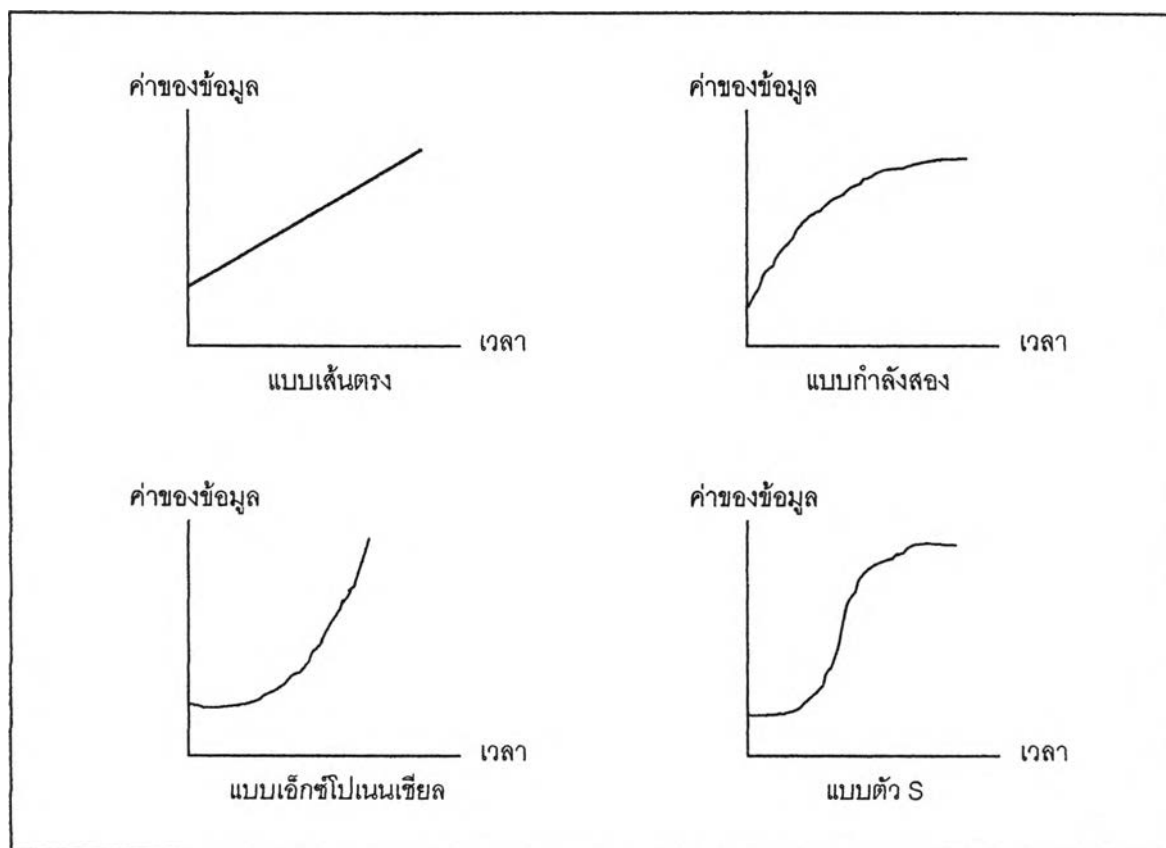
$$Y = T \times S \times C \times I$$

1.4 การวิเคราะห์แนวโน้ม (Trend Analysis)

เป็นการพยากรณ์แนวโน้มของเหตุการณ์ในอดีตต่อเนื่องไปถึงเหตุการณ์ในอนาคต แนวโน้มอาจจะมีลักษณะต่างๆ กัน เช่น แนวโน้มเส้นตรง (linear trend) แนวโน้มกำลังสอง (quadratic trend) แนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential trend) และแนวโน้มแบบตัว S (S-shaped trend) ดังแผนภาพ 3



แผนภาพ 2 ส่วนประกอบของอนุกรมเวลา



แผนภาพ 3 รูปแบบของแนวโน้มประเภทต่างๆ

1.5 วิธีบ็อกซ์และเจ็นกินส์ (Box-Jenkins Method)

เป็นเทคนิคที่ต้องการข้อมูลอย่างน้อย 50 ค่าขึ้นไป โดยอาศัยฟังก์ชันอัตโนมัติสหสัมพันธ์ (autocorrelation function) และฟังก์ชันอัตโนมัติสหสัมพันธ์บางส่วน (partial autocorrelation function) ในการวิเคราะห์ (Sullivan and Claycombe, 1977 ; Montgomery, Johnson and Gardiner, 1990) ซึ่งมีวิธีการทั้งสิ้น 4 ขั้นตอน ดังนี้ (Makridakis, Wheelwright and Mcgee, 1978 ; ทรงศิริ แต่สมบัติ, 2539)

- การกำหนดโมเดล (Identification)
- การประมาณค่าพารามิเตอร์ (Estimation the Parameters)
- การตรวจสอบความเหมาะสมของโมเดล (Diagnostic Checking)
- การพยากรณ์ (Forecasting)

ซึ่งรายละเอียดของทั้ง 4 ขั้นตอนจะกล่าวถึงในตอนต่อไป

2. เทคนิคที่ใช้ความสัมพันธ์ระหว่างตัวแปร (Causal Model)

2.1 การวิเคราะห์การถดถอยแบบง่าย (Simple Regression Model)

ค่าของตัวแปรอิสระ (independent variable) หรือ X เพียง 1 ตัวสามารถพยากรณ์ตัวแปรตาม (dependent variable) หรือ Y ได้ 1 ตัว ดังสมการ

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

2.2 การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Model)

ค่าของตัวแปรอิสระ (independent variable) หรือ X หลายตัวสามารถพยากรณ์ตัวแปรตาม (dependent variable) หรือ Y ได้ 1 ตัว ดังสมการ

$$Y = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

ซึ่งรายละเอียดของการวิเคราะห์การถดถอยแบบง่ายและการวิเคราะห์การถดถอยแบบพหุคูณ จะกล่าวถึงในตอนต่อไป

2.3 การวิเคราะห์ทางเศรษฐมิติ (Econometric Model)

หมายถึง วิธีการวิเคราะห์ปรากฏการณ์ทางเศรษฐศาสตร์ในเชิงปริมาณ โดยใช้ข้อมูลที่เป็นตัวเลขหรือข้อมูลซึ่งสามารถตีความหมายเป็นตัวเลขได้ นำมาเป็นวัตถุดิบสำคัญในการวิเคราะห์ โดยอาศัยทฤษฎีทางเศรษฐศาสตร์ คณิตศาสตร์ และสถิติ เป็นเครื่องมือในการวิเคราะห์

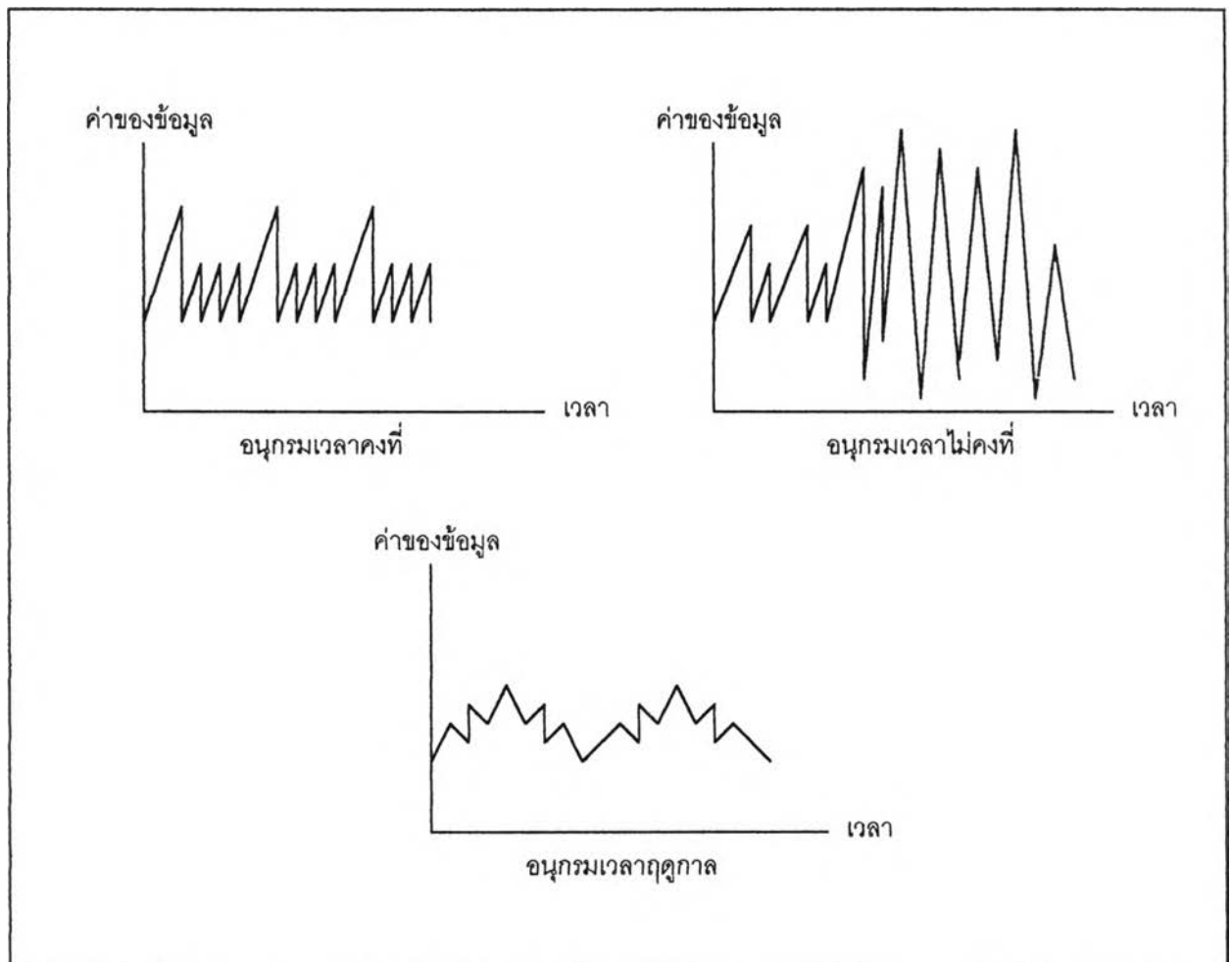
ประเภทของอนุกรมเวลา

อนุกรมเวลา (Time Series) หมายถึง ค่าของข้อมูลที่เกิดขึ้นในระยะเวลาต่างๆ กัน ซึ่งได้รับการจัดเรียงลำดับโดยคำนึงถึงเวลาแห่งการเกิดขึ้นหรือปรากฏเป็นเกณฑ์ โดยจุดบอเวลาจะอยู่ห่างจากกันแต่ละคู่เป็นระยะเท่าๆ กัน เช่น ยอดขายสินค้าเป็นรายเดือน ข้อมูลนี้เรียกว่าข้อมูลอนุกรมเวลา (Time Series) ซึ่งแบ่งออกได้เป็น 3 ประเภท ดังนี้ (วันพร เหลืองอาภาพงศ์, 2519 ; บุชบา พิกุลผล, 2522)

1. อนุกรมเวลาคงที่ (Stationary Time Series) ข้อมูลจะเคลื่อนไหวไปรอบๆ ค่าตัวกลาง (Mean) และการเคลื่อนไหวจะเป็นไปในลักษณะคงที่ พหาค่าหนึ่งเพิ่มขึ้นค่าที่ตามมาจะลดลง

2. อนุกรมเวลาไม่คงที่ (Nonstationary Time Series) ข้อมูลจะเคลื่อนไหวไม่แน่นอน มีการเปลี่ยนแปลงไปตามระยะเวลาและจะไม่สามารถหาตัวกลางที่เป็นศูนย์กลางได้

3. อนุกรมเวลาฤดูกาล (Seasonal Time Series) ข้อมูลจะมีการเคลื่อนไหวขึ้นลงตามระยะเวลาเป็นช่วงที่แน่นอน เช่น 3 เดือน 6 เดือน 1 ปี เป็นต้น และลักษณะการเคลื่อนไหวระยะเวลาหนึ่งจะคล้ายหรือเหมือนกันกับช่วงเวลาอื่นๆ ซ้ำๆ กัน



แผนภาพ 4 ประเภทของอนุกรมเวลา

ตาราง 1 การเปรียบเทียบเทคนิคการพยากรณ์เชิงปริมาณตามปัจจัยต่างๆ

ปัจจัย	การพยากรณ์เชิงปริมาณ							
	Time Series Analysis					Causal Model		
	Moving Average	Smoothing Method	Decomposition	Trend Analysis	Box Jenkins	Simple Regression	Multiple Regression	Econometric Model
1.ระยะเวลา								
-สั้นมาก	✓	✓	✓		✓			
-สั้น	✓				✓		✓	✓
-ปานกลาง					✓		✓	✓
-ยาว				✓		✓	✓	✓
2.ลักษณะข้อมูล								
-แนวนอน	✓	✓	✓		✓			
-แนวโน้ม			✓		✓	✓	✓	
-ฤดูกาล			✓		✓			
-วัฏจักร			✓		✓			
3.ชนิดของเทคนิค								
-อนุกรมเวลา	✓	✓	✓	✓	✓	✓	✓	
-ความสัมพันธ์ของข้อมูล					✓	✓	✓	✓
-ใช้เทคนิคทางสถิติ					✓	✓	✓	✓
-ไม่ใช้เทคนิคทางสถิติ	✓	✓	✓	✓				

ที่มา : เศกสรร บัวทรัพย์ (2531)

ตอนที่ 2 วิธีของบ็อกซ์และเจ็นกินส์และโมเดลอริมาอินเตอร์เวนชัน

วิธีของบ็อกซ์และเจ็นกินส์

Autoregressive Integrated Moving Average (ARIMA) Model ได้รับการศึกษาอย่างกว้างขวางโดย George Box และ Gwilym Jenkins (1976) และชื่อของทั้งสองคนถูกนำมาใช้กับกระบวนการ ARIMA ในการประยุกต์ใช้กับการวิเคราะห์อนุกรมเวลา การพยากรณ์ และการควบคุม โดยที่ Autoregressive (AR) Model ถูกนำมาใช้ครั้งแรกโดย Yule (1926) และนำมาใช้ต่อมาโดย Walker (1931) ในขณะที่ Moving Average (MA) Model ถูกนำมาใช้ครั้งแรกโดย Slutsky (1937) และนำมาใช้ต่อมาโดย Wold (1938) วิธีการดังกล่าวแสดงไว้ดังแผนภาพ 5

วิธีการของบ็อกซ์และเจ็นกินส์มี 4 ขั้นตอน ดังนี้

1. การกำหนดโมเดล (Identification)

Nick T. Thomopoulos(1980) กล่าวว่า วัตถุประสงค์ของขั้นตอนนี้เป็นการเลือกโมเดลสำหรับใช้ในการพยากรณ์ซึ่งจะต้องมีความเหมาะสมกับอนุกรมเวลาที่ต้องการศึกษา โดยจะแบ่งออกเป็น 2 ตอน ดังนี้

1.1 อัตตะสหสัมพันธ์ (autocorrelation) และอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation)

1.1.1 อัตตะสหสัมพันธ์ (autocorrelation)

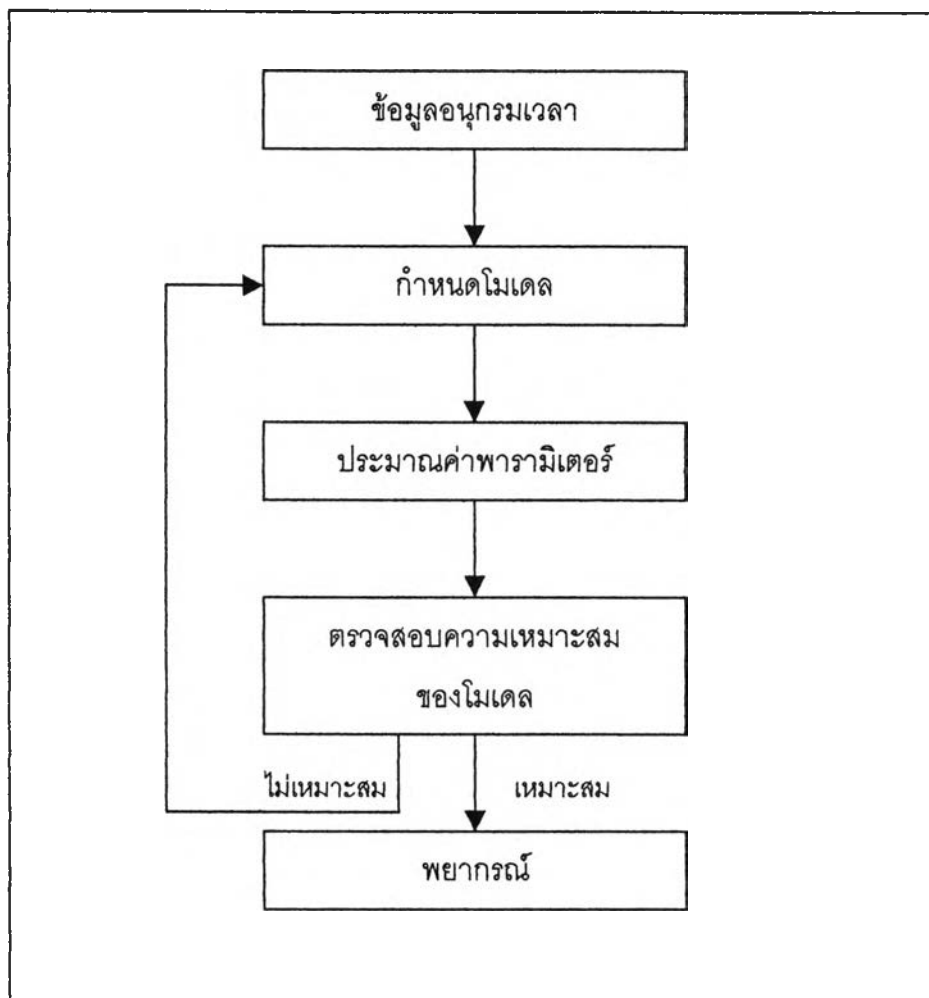
สถิติที่สำคัญที่ใช้ในการวิเคราะห์อนุกรมเวลาคือสัมประสิทธิ์อัตตะสหสัมพันธ์ (autocorrelation coefficient) (Makridakis , Wheelwright and McGee, 1978) ซึ่งเป็นการวัดความสัมพันธ์ระหว่างค่าที่ต่อเนื่องของตัวแปรที่เหมือนกัน โดยที่สัมประสิทธิ์อัตตะสหสัมพันธ์จะคล้ายคลึงกับสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ซึ่งจะอธิบายความสัมพันธ์ซึ่งกันและกัน (mutual relationship) ระหว่างค่าของตัวแปร แต่จะแตกต่างกันในระหว่างคาบเวลา (time period)

จากตาราง 2 พบว่า สามารถสร้างตัวแปรตัวหนึ่งจากตัวแปรอีกตัวหนึ่งได้ ซึ่งตัวแปรที่สร้างขึ้นใหม่นี้ เรียกว่า ตัวแปรเวลาล่าหลัง (time lag variables) โดยสามารถสร้างตัวแปรเวลาล่าหลัง Y_{t-1} , Y_{t-2} โดยการย้ายข้อมูลจากช่วงเวลา t ลงมาเป็นช่วงเวลา t_2 และจากช่วงเวลา t_2 ลงมาเป็นช่วงเวลา t_3 ตามลำดับ ซึ่งผลจากการย้ายนี้ทำให้ค่าใน Y_{t-1} หายไป 1 ค่า และ Y_{t-2} หายไป 2 ค่า

จากข้อมูลในตาราง 2 สามารถเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \quad \dots\dots\dots(7)$$

จากสมการ (7) พบว่า Y_t เป็น Linear Combination ตัวแปร Y_{t-1} และ Y_{t-2} ถูกสร้างขึ้นโดยการเคลื่อนที่ของช่วงเวลาที 1 และ 2 ตามลำดับ ผลที่ได้คือในคอลัมน์ Y_{t-1} นั้นค่าที่ 1 จะสูญหาย ส่วนในคอลัมน์ Y_{t-2} นั้นค่าที่ 1 และ 2 จะสูญหาย



แผนภาพ 5 ขั้นตอนการวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีบ็อกซ์และเจ็นกินส์

ตาราง 2 การสร้างตัวแปรเวลาล้าหลัง (time lag variables)

ช่วงเวลา	ตัวแปรเริ่มต้น	ตัวแปรเวลาล้าหลังที่ 1	ตัวแปรเวลาล้าหลังที่ 2
t	Y_t	Y_{t-1}	Y_{t-2}
1	13	-	-
2	8	13	-
3	15	8	13
4	4	15	8
5	4	4	15
6	12	4	4
7	11	12	4
8	7	11	12
9	14	7	11
10	12	14	7

ดังนั้น Simple Correlation Coefficient ระหว่าง Y_t และ Y_{t-1} สามารถคำนวณได้โดยใช้สมการดังต่อไปนี้

$$r_{Y_t Y_{t-1}} = \frac{\text{Covariance ระหว่าง } Y_t \text{ และ } Y_{t-1}}{(\text{standard deviation ของ } Y_t) (\text{standard deviation ของ } Y_{t-1})}$$

$$= \frac{\sum_{t=2}^n (Y_t - \bar{Y}_t)(Y_{t-1} - \bar{Y}_{t-1})}{\sqrt{\sum_{t=1}^n (Y_t - \bar{Y}_t)^2} \sqrt{\sum_{t=2}^n (Y_{t-1} - \bar{Y}_{t-1})^2}} \quad \dots\dots\dots(8)$$

อนุกรม Y_t ถูกสมมติให้เป็น stationary ทั้งค่าเฉลี่ยและความแปรปรวน ดังนั้นค่าเฉลี่ยของทั้ง Y_t และ Y_{t-1} จึงเท่ากันและจะตัดตัวห้อยทิ้งโดยให้ $\bar{Y} = \bar{Y}_t = \bar{Y}_{t-1}$ จึงทำให้ได้สมการดังต่อไปนี้

$$r_{Y_t Y_{t-1}} = \frac{\sum_{t=2}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sqrt{\sum_{t=1}^n (Y_t - \bar{Y})^2}} \quad \dots\dots\dots(9)$$

เพื่อความสะดวกจะเขียน r_1 เป็น autocorrelation coefficient สำหรับเวลาดำหลังที่ 1 และในทำนองเดียวกัน autocorrelation coefficient สำหรับเวลาดำหลังที่ 1, 2, 3, ..., k สามารถหาได้โดย r_k ดังนี้

$$r_k = \frac{\sum_{t=2}^n (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sqrt{\sum_{t=1}^n (Y_t - \bar{Y})^2}} \quad \dots\dots\dots(10)$$

ความหมายของ r_k จะเป็นทำนองเดียวกับค่าสัมประสิทธิ์สหสัมพันธ์ นั่นคือเป็นค่าวัดสหสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน k ช่วงเวลา คือค่าวัดสหสัมพันธ์ระหว่าง Y_t และ Y_{t+k} (ทรงศิริ แต่สมบัตติ, 2539)

การวัดความสัมพันธ์เชิงเส้นระหว่างค่าสังเกตของอนุกรมเวลาจะแยกโดยใช้หน่วยของเวลาดำหลังที่ k ซึ่ง r_k มีค่าอยู่ระหว่าง -1 และ 1 ถ้าค่าสังเกตแยกโดยหน่วยของเวลาดำหลังที่ k มีแนวโน้มเคลื่อนไหวเชิงเส้นด้วยความชันเป็นบวก ค่า r_k จะเข้าใกล้ 1 แต่ถ้าความชันเป็นลบ ค่า r_k จะเข้าใกล้ -1 (บำเพ็ญ ปิตชิต, 2540)

1.1.2 อัตตะสหสัมพันธ์บางส่วน (partial autocorrelation)

Makridakis, Wheelwright and McGee (1978) กล่าวว่า อัตตะสหสัมพันธ์บางส่วน (Partial autocorrelation) เป็นการวัดทางสถิติที่คล้ายคลึงกับอัตตะสหสัมพันธ์ (Autocorrelation) และเป็นการวัดระดับของความสัมพันธ์ระหว่าง Y_t และ Y_{t+k} เมื่อมีผลกระทบของเวลาดำหลังที่ 1, 2, 3, ..., k มีวัตถุประสงค์คือใช้ในการวิเคราะห์อนุกรมเวลาที่จะช่วยระบุโมเดล ARIMA ที่เหมาะสมสำหรับการพยากรณ์

สัมประสิทธิ์อัตตะสหสัมพันธ์บางส่วน (Partial autocorrelation coefficient) ของลำดับที่ m ถูกจำกัดความเป็น สัมประสิทธิ์การถดถอยในตัวเอง (Autoregressive coefficient)

ของโมเดล AR(m) ตัวอย่างเช่นในสมการ (11) – (15) เป็นการอธิบายกระบวนการ AR(1) , AR(2) , ..., AR(m-1) , AR(m) ตามลำดับ และพบว่า $\phi_1, \phi_2, \phi_3, \dots, \phi_{m-1}, \phi_m$ เป็น first m partial autocorrelation coefficient สำหรับอนุกรมเวลา

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t \quad \dots\dots\dots(11)$$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \quad \dots\dots\dots(12)$$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \varepsilon_t \quad \dots\dots\dots(13)$$

⋮
⋮
⋮

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_{m-1} Y_{t-m+1} + \varepsilon_t \quad \dots\dots\dots(14)$$

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_{m-1} Y_{t-m+1} + \phi_m Y_{t-m} + \varepsilon_t \quad \dots\dots\dots(15)$$

ดังนั้นเพื่อความสะดวกในการประมาณค่า $\phi_1, \phi_2, \phi_3, \dots, \phi_{m-1}, \phi_m$ ซึ่งเป็นพื้นฐานของสัมประสิทธิ์อัตโนมัติ การประมาณค่าเหล่านี้สามารถทำได้ตามวิธีดังต่อไปนี้

คูณทั้งสองข้างของสมการ (11) ด้วย Y_{t-1} จะได้ว่า

$$Y_{t-1} Y_t = \phi_1 Y_{t-1} Y_{t-1} + Y_{t-1} \varepsilon_t \quad \dots\dots\dots(16)$$

ค่าคาดหวังของสมการ (16) คือ

$$E(Y_{t-1} Y_t) = \phi_1 E(Y_{t-1} Y_{t-1}) + E(Y_{t-1} \varepsilon_t)$$

ซึ่งสามารถเขียนสมการใหม่ได้เป็น

$$\gamma_1 = \phi_1 \gamma_0 \quad \dots\dots\dots(17)$$

เมื่อ $E(Y_{t-1} Y_t) = \gamma_1$, $E(Y_{t-1} Y_{t-1}) = \gamma_0$ และ $E(Y_{t-1} \varepsilon_t) = 0$

นำ γ_0 มาหารทั้งสองข้างของสมการ (17) จะได้ว่า

$$\rho_1 = \phi_1$$

เมื่อ $\rho_1 = (\gamma_1/\gamma_0)$ เป็นการอธิบาย first autocorrelation ดังนั้น $\phi_1 = \rho_1$ จึงทำให้ first partial autocorrelation เหมือนกับ first autocorrelation เมื่อคูณทั้งสองข้างของสมการ (11) ด้วย Y_{t-k} จะได้ค่าคาดหวัง และเมื่อหารทั้งสองข้างของสมการ (17) โดย γ_0 จึงเรียกสมการนี้ว่า

Yule-Walker Equations ซึ่งสามารถแก้สมการหาค่า $\phi_1, \phi_2, \phi_3, \dots, \phi_{m-1}$ และ ϕ_m ได้ ซึ่งค่าเหล่านี้สามารถใช้ประมาณค่า partial autocorrelation ของเวลาล่าหลังที่ m ขึ้นไป

1.2 โมเดลการวิเคราะห์อนุกรมเวลาด้วยวิธีบ็อกซ์และเจ็นกินส์

แบ่งออกเป็น 3 ประเภท ดังนี้

1.2.1 โมเดลกระบวนการถดถอยในตัวเอง (Autoregressive Model) อันดับ p ใช้สัญลักษณ์ AR(p)

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad \dots\dots\dots(18)$$

เมื่อ Y_t = ค่าสังเกต
 θ_0 = ค่าคงที่
 ϕ_p = พารามิเตอร์กระบวนการถดถอยในตัวเอง (AR) ตัวที่ p
 ε_t = ค่าความคลาดเคลื่อน ณ ช่วงเวลา t (random shock)

สมการ (18) เรียกว่า autoregressive เพราะคล้ายกับสมการถดถอย คือ $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_p X_p + e$ ซึ่งแตกต่างกันตรงที่ $X_1 = Y_{t-1}, X_2 = Y_{t-2}, X_p = Y_{t-p}$ และตัวแปรอิสระเป็นค่า lag ของตัวแปรตามทีเวลาล่าหลัง 1, 2, 3, ..., p

จากสมการ (18) สามารถเขียนในเทอมของ backward – shift operator ได้ดังนี้

$$Y_t = \theta_0 + (\phi_1 B + \phi_2 B^2 + \dots + \phi_p B^p) Y_t + \varepsilon_t \quad \dots\dots\dots(19)$$

หรือ

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Y_t = \theta_0 + \varepsilon_t \quad \dots\dots\dots(20)$$

โมเดล AR ไม่มีข้อจำกัดของพารามิเตอร์ที่ทำให้โมเดลเป็นอินเวอร์ทิเบิล แต่มีข้อจำกัดของพารามิเตอร์ที่ทำให้โมเดลเป็นสเตชันนารี พังก์ชันอัตตะสหสัมพันธ์ r_k มีค่ามากในค่าแรกๆ และลดลงอย่างช้าๆ เมื่อ k มีค่ามากขึ้นเกือบถึงศูนย์ หรือ tails off มีลักษณะ damped exponential หรือ damped sine wave หรือทั้ง 2 แบบรวมกัน พังก์ชันอัตตะสหสัมพันธ์บางส่วนจะมี

ค่าลดลงอย่างรวดเร็ว ค่า ϕ_{kk} มีค่าเป็นศูนย์หลัง lag p หรือ cuts off โดยที่โมเดล AR จะมี order เป็น p

$$\phi_{kk} \neq 0 \quad ; \quad k \leq p$$

$$\phi_{kk} = 0 \quad ; \quad k \geq p$$

โมเดลกระบวนการถดถอยในตัวเองมีกระบวนการที่น่าสนใจ 2 กระบวนการดังนี้

1.2.1.1 กระบวนการถดถอยในตัวเองลำดับที่หนึ่ง

(The first-order autoregressive process) หรือ AR(1) process

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t \quad \dots\dots\dots(21)$$

AR(1) process เรียกอีกอย่างหนึ่งว่า The Markov Process เพราะค่าสังเกต ณ เวลา t ขึ้นกับค่าสังเกต ณ เวลา t - 1

ข้อจำกัดที่จะต้องมีสำหรับพารามิเตอร์ ϕ_1 ที่ทำให้โมเดลเป็นสเตชันนารี คือ $|\phi_1| < 1$ ฟังก์ชันอัตโนมัติบางส่วนจะ cuts off หลัง lag 1 และฟังก์ชันอัตโนมัติลดลงแบบ exponential

1.2.1.2 กระบวนการถดถอยในตัวเองลำดับที่สอง (The second-order autoregressive process) หรือ AR(2) process

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \quad \dots\dots\dots(22)$$

ข้อจำกัดของพารามิเตอร์ที่ทำให้โมเดลเป็นสเตชันนารีคือ

$$\phi_1 + \phi_2 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$|\phi_2| < 1$$

ฟังก์ชันอัตโนมัติบางส่วนบางส่วนจะ cuts off หลัง lag 2 และฟังก์ชันอัตโนมัติจะลดลงในลักษณะของ damped exponentials และ/หรือ damped sine waves

1.2.2 โมเดลกระบวนการเฉลี่ยเคลื่อนที่ (Moving Average Model) อันดับ ที่ q ใช้สัญลักษณ์ $MA(q)$

$$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad \dots\dots\dots(23)$$

เมื่อ $Y_t =$ ค่าสังเกต
 $\theta_0 =$ ค่าคงที่
 $\theta_q =$ พารามิเตอร์กระบวนการเฉลี่ยเคลื่อนที่ (MA) ตัวที่ q
 $\varepsilon_t =$ ค่าความคลาดเคลื่อน ณ ช่วงเวลา t (random shock)

จากสมการ (12) เขียนในเทอมของ backward – shift operator ได้ดังนี้

$$Y_t = \theta_0 + (1 - \theta_1 B^{-1} - \theta_2 B^{-2} - \dots - \theta_q B^{-q}) \varepsilon_t \quad \dots\dots\dots(24)$$

โมเดล MA ไม่มีข้อจำกัดของพารามิเตอร์ที่ทำให้โมเดลเป็นสเตชันนารี แต่มีข้อจำกัดของพารามิเตอร์ที่ทำให้โมเดลเป็นอินเวอร์ทิเบิล ฟังก์ชันอัตตะสหัมพันธ์บางส่วนมีลักษณะ damped exponential หรือ damped sine waves หรือทั้ง 2 แบบรวมกันเป็น tails off ฟังก์ชันอัตตะสหัมพันธ์มีลักษณะ cuts off ค่า r_k ลดลงอย่างรวดเร็ว จะมีลำดับ q เมื่อหลัง lag q มีค่าเป็นศูนย์

โมเดลกระบวนการเฉลี่ยเคลื่อนที่มีกระบวนการที่น่าสนใจ 2 กระบวนการ
 ดังนี้คือ

1.2.2.1 กระบวนการเฉลี่ยเคลื่อนที่ลำดับที่หนึ่ง (The first-order moving average process) หรือ $MA(1)$ process

$$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad \dots\dots\dots(25)$$

ข้อจำกัดที่ทำให้โมเดลเป็นอินเวอร์ทิเบิล คือ $|\theta_1| < 1$ ฟังก์ชัน
 อัตตะสหัมพันธ์ cuts off หลัง lag 1 และฟังก์ชันอัตตะสหัมพันธ์บางส่วนมีลักษณะ damped
 exponential

1.2.2.2 กระบวนการเฉลี่ยเคลื่อนที่ลำดับที่สอง (The second-order moving average process) หรือ MA(2) process

$$Y_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} \quad \dots\dots\dots(26)$$

ข้อจำกัดที่ทำให้โมเดลเป็นอินเวอร์ทิเบิลคือ

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$|\theta_2| < 1$$

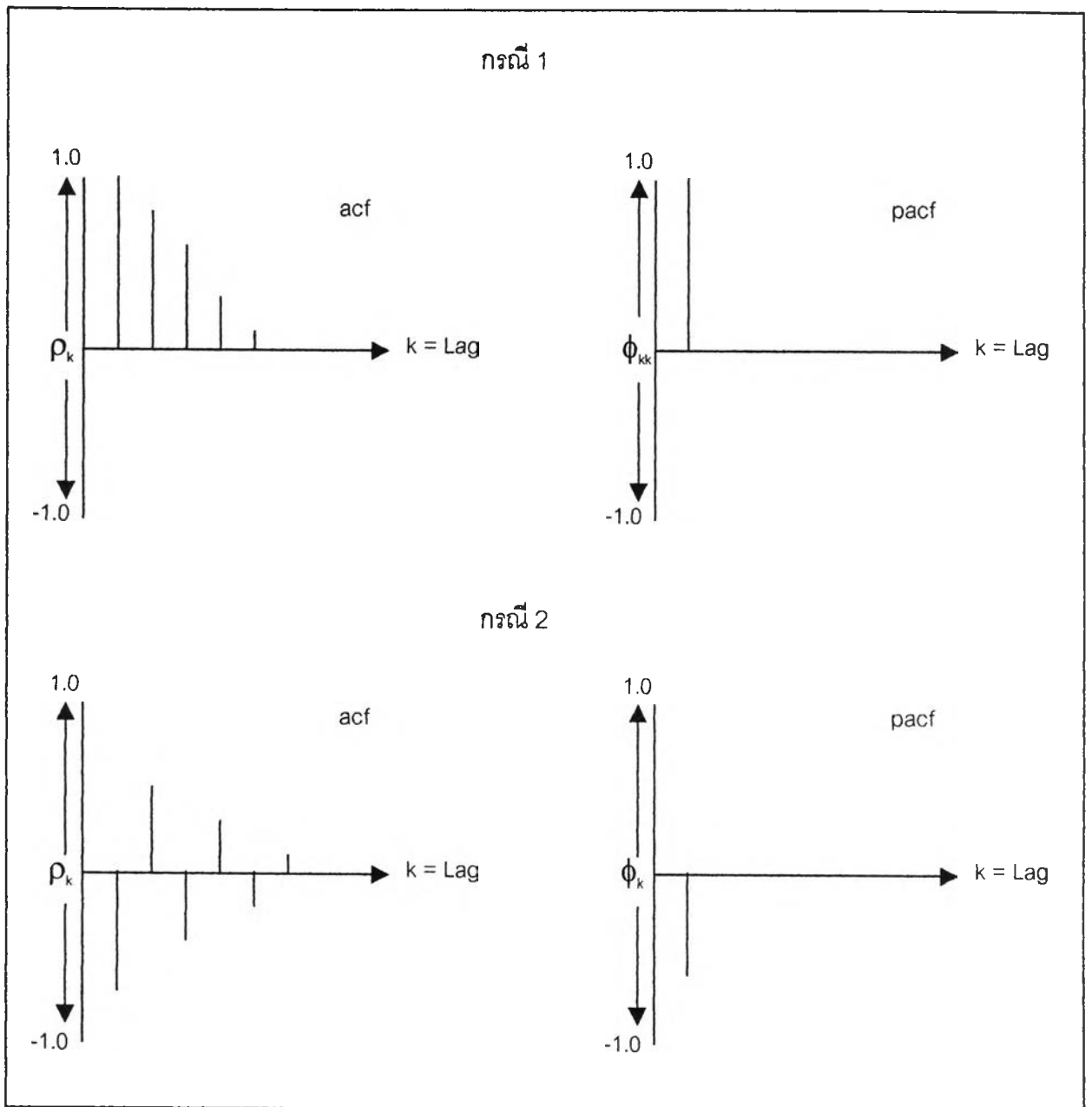
ฟังก์ชันอัตตาสหสัมพันธ์บางส่วนบางส่วนจะลดลงมีลักษณะผสมของ damped exponential และ/หรือ damped sine waves และฟังก์ชันอัตตาสหสัมพันธ์จะ cuts off หลัง lag2

1.2.3 โมเดลการรวมกันของกระบวนการถดถอยในตัวเองและกระบวนการเฉลี่ยเคลื่อนที่ (Mixed autoregressive- moving Average Model) อันดับที p และ q ใช้สัญลักษณ์ ARMA (p,q)

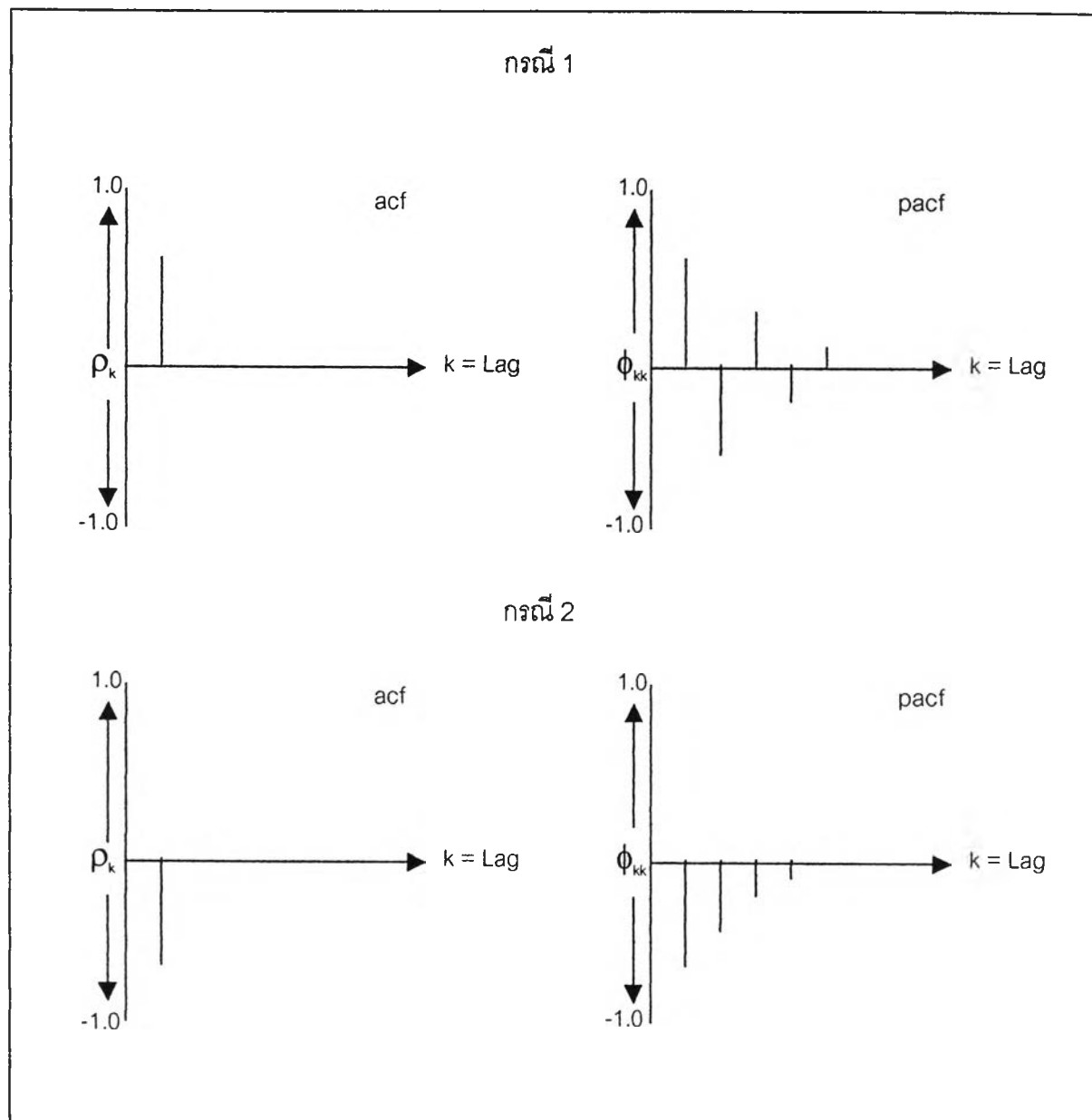
$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q} + \varepsilon_t \quad \dots\dots\dots(27)$$

ตาราง 3 ลักษณะของฟังก์ชันอัตตาสหสัมพันธ์ (Autocorrelation Function) และฟังก์ชันอัตตาสหสัมพันธ์บางส่วน (Partial Autocorrelation Function) ของอนุกรมเวลาคงที่

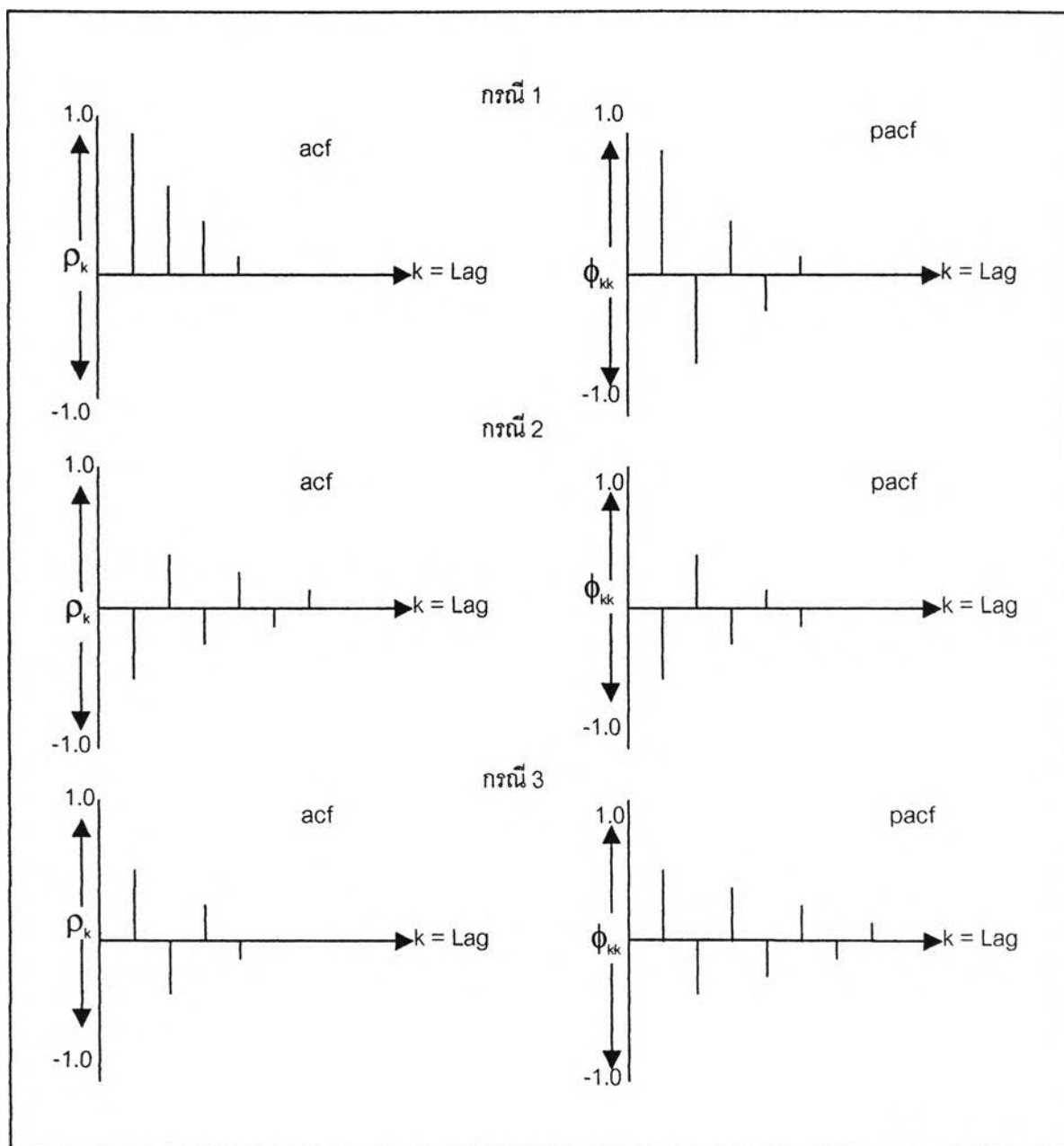
กระบวนการ	ACF	PACF
AR(1)	ค่า ρ_k ลดลงอย่างรวดเร็วขณะที่ $k > 1$	ค่า ϕ_{kk} จะมีค่าสูงที่ $k=1$ และเท่ากับ 0 เมื่อ $k>1$
AR(2)	ค่า ρ_k ลดลงอย่างรวดเร็วขณะที่ $k > 1$	ค่า ϕ_{kk} จะมีค่าสูงที่ $k=1,2$ และเท่ากับ 0 เมื่อ $k>2$
MA(1)	ค่า ρ_k จะมีค่าสูงที่ $k=1$ และเท่ากับ 0 เมื่อ $k>1$	ค่า ϕ_{kk} ลดลงอย่างรวดเร็ว ขณะที่ $k>1$
MA(2)	ค่า ρ_k จะมีค่าสูงที่ $k=1,2$ และเท่ากับ 0 เมื่อ $k>2$	ค่า ϕ_{kk} ลดลงอย่างรวดเร็ว ขณะที่ $k>1$
ARMA(1,1)	ค่า ρ_k ลดลงอย่างรวดเร็วหลังจาก Lag $k=1$	ค่า ϕ_{kk} ลดลงอย่างรวดเร็วหลัง Lag $k=1$



แผนภาพ 6 กระบวนการ AR(1)



แผนภาพ 7 กระบวนการ MA(1)



แผนภาพ 8 กระบวนการ ARMA(1,1)

จากที่ผู้วิจัยได้กล่าวถึงประเภทของข้อมูลอนุกรมเวลามี 3 ประเภท ได้แก่ อนุกรมเวลาคงที่ อนุกรมเวลาไม่คงที่ และอนุกรมเวลาฤดูกาล อาจกล่าวได้ว่ามีลักษณะของโมเดล ARIMA (p,d,q) แตกต่างกันไปตามลักษณะของประเภทอนุกรมเวลา ดังรายละเอียดต่อไปนี้

สำหรับอนุกรมเวลาคงที่ โมเดล ARMA(p,q) มี p เป็นอันดับของ AR และ q เป็นอันดับของ MA จำนวนพารามิเตอร์ในโมเดลจะเท่ากับ $p + q + 1$ โมเดลที่กำหนดให้กับอนุกรมเวลามักจะเป็นโมเดลที่มีจำนวนพารามิเตอร์น้อย ในทางปฏิบัติมักจะไม่เกิน 3 พารามิเตอร์ (ทรงศิริ แต่สมบัติ, 2539) ดังตาราง 4 แสดงโมเดล ARMA ต่างๆ และเงื่อนไขของพารามิเตอร์ที่ปรากฏในโมเดล โดย Y_t เป็นค่าสังเกต ณ เวลา t ในอนุกรมเวลา $\{y_t\}$ และ ε_t เป็นค่าคลาดเคลื่อนของการพยากรณ์ ณ เวลา t จะกำหนดให้ ε_t มีการแจกแจงที่เป็นอิสระกัน มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ^2 นั่นคือ $\varepsilon_t \sim IN(0, \sigma^2)$ สำหรับเงื่อนไขของพารามิเตอร์ในส่วนของ AR อันดับที่ p เป็นเงื่อนไขที่กำหนดให้โมเดล ARMA(p,q) เป็นสเตชันนารี ส่วนเงื่อนไขของพารามิเตอร์ในส่วนของ MA ที่อันดับ q จะเป็นเงื่อนไขที่กำหนดให้โมเดล MA(q) เป็นอินเวอร์ทิเบิล

ตาราง 4 โมเดล ARMA(p,q) และเงื่อนไขสเตชันนารี (stationary) และอินเวอร์ทิเบิล (invertible)

รูปแบบ ARMA(p,q)	รูปแบบ	เงื่อนไข
White noise	$Y_t = \theta_0 + \varepsilon_t$	
AR(1)	$Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t$	$-1 < \phi_1 < 1$
AR(2)	$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$	$\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$ $-1 < \phi_2 < 1$
MA(1)	$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1}$	$-1 < \theta_1 < 1$
MA(2)	$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$	$\theta_2 + \theta_1 < 1$ $\theta_2 - \theta_1 < 1$ $-1 < \theta_2 < 1$
ARMA(1,1)	$Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$	$-1 < \phi_1 < 1$ $-1 < \theta_1 < 1$

สเตชันนารีจะเป็นคุณสมบัติของโมเดล AR(p) ซึ่งเป็นคุณสมบัติที่ทำให้ $E(Y_t)$ และ $V(Y_t)$ คงที่ และ $cov(Y_t, Y_{t-k})$ จะขึ้นกับ lag k เพียงอย่างเดียว การพิจารณาว่าค่าพารามิเตอร์ ϕ_1, \dots, ϕ_p ไตที่จะทำให้โมเดล AR เป็นสเตชันนารี จะทำได้ตามขั้นตอนต่อไปนี้

1. จากโมเดล AR(p)

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

หรือ

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \theta_0 + \varepsilon_t$$

จะเขียนโมเดลในเทอมของ backward operator ได้เป็น

$$(1 - \phi_1 B - \dots - \phi_p B^p) Y_t = \theta_0 + \varepsilon_t$$

2. หาคำตอบของสมการ $1 - \phi_1 B - \dots - \phi_p B^p = 0$ จะได้ค่าของ B จำนวน p ค่า จะเลือก B เพียง 1 ค่าที่อยู่นอก unit circle นั่นคือ $|B|$ ต้องมีค่ามากกว่า 1 เงื่อนไขดังกล่าวของ B จะเป็นเงื่อนไขของสเตรชันนารี

อินเวอร์ติเบิลจะเป็นคุณสมบัติของโมเดล MA(q) ซึ่งเป็นคุณสมบัติที่ทำให้หาค่าคลาดเคลื่อนของการพยากรณ์ ε_t ในเทอมของ Y_t, Y_{t-1}, \dots ได้ การพิจารณาว่าค่าพารามิเตอร์ $\theta_1, \dots, \theta_q$ ใดที่จะทำให้โมเดล MA(q) เป็นอินเวอร์ติเบิล จะทำได้ตามขั้นตอนต่อไปนี้

1. จากโมเดล MA(q)

$$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

หรือ

$$Y_t = \theta_0 + (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t$$

2. หาคำตอบของสมการ $1 - \theta_1 B - \dots - \theta_q B^q = 0$ จะได้ค่าของ B จำนวน q ค่า จะเลือก B เพียง 1 ค่าที่อยู่นอก unit circle นั่นคือ $|B|$ มีค่ามากกว่า 1 ซึ่ง B ที่มีค่าดังกล่าวจะเป็นเงื่อนไขของอินเวอร์ติเบิล

สำหรับอนุกรมเวลาที่มีแนวโน้มเพียงอย่างเดียว นั่นคือ มีค่าเฉลี่ย $E(Y_t)$ ไม่คงที่ สำหรับแต่ละค่าของ t การปรับอนุกรมเวลาดังกล่าวให้เป็นอนุกรมเวลาใหม่ที่เป็นอนุกรมเวลาที่คงที่ จะทำได้โดยการหาผลต่าง (ทรงศิริ แต่สมบัติ, 2539) นั่นคือ จากอนุกรมเวลา $\{y_t\}$ แปลงเป็นอนุกรมเวลาใหม่ $\{z_t\}$ ซึ่ง $Z_t = \nabla^d Y_t$ เมื่ออนุกรมเวลา $\{y_t\}$ เป็นอนุกรมเวลาที่คงที่แล้วจะหาโมเดล ARMA(p,q) ให้

กับอนุกรมเวลา $\{Z_t\}$ ได้เป็น $Z_t \sim \text{ARMA}(p,q)$ และสำหรับโมเดลของอนุกรมเวลา $\{y_t\}$ ได้เป็น $Y_t \sim \text{ARIMA}(p,d,q)$ ซึ่งผู้วิจัยขอยกตัวอย่างโมเดล $\text{ARIMA}(p,1,q)$ ของอนุกรมเวลา $\{y_t\}$ ซึ่งมีลักษณะของสมการดังต่อไปนี้

$$\begin{aligned} \text{ARIMA}(0,1,0) \quad Z_t &= \theta_0 + \varepsilon_t \\ \text{หรือ} \quad Y_t &= \theta_0 + Y_{t-1} + \varepsilon_t \\ \\ \text{ARI}(1,1) \quad Z_t &= \theta_0 + \phi_1 Z_{t-1} + \varepsilon_t \\ \text{หรือ} \quad Y_t &= \theta_0 + (1 + \phi_1)Y_{t-1} - \phi_2 + \varepsilon_t \\ \\ \text{ARI}(2,1) \quad Z_t &= \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \varepsilon_t \\ \text{หรือ} \quad Y_t &= \theta_0 + (1 + \phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} - \phi_2 Y_{t-3} + \varepsilon_t \\ \\ \text{IMA}(1,1) \quad Z_t &= \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ \text{หรือ} \quad Y_t &= \theta_0 + Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ \\ \text{IMA}(1,2) \quad Z_t &= \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \\ \text{หรือ} \quad Y_t &= \theta_0 + Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \\ \\ \text{ARIMA}(1,1,1) \quad Z_t &= \theta_0 + \phi_1 Z_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \\ \text{หรือ} \quad Y_t &= \theta_0 + (1 + \phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t - \phi_1 \varepsilon_{t-1} \end{aligned}$$

ตาราง 5 ลักษณะของ $\rho_k(Y_t)$, $\rho_k(Z_t)$ และ $\rho_{kk}(Z_t)$ สำหรับโมเดล $\text{ARIMA}(p,d,q)$

รูปแบบของอนุกรมเวลา $\{y_t\}$	ลักษณะของ $\rho_k(Y_t)$	ลักษณะของ $\rho_k(Z_t)$	ลักษณะของ $\rho_{kk}(Z_t)$
Random Walk	ลดลงอย่างช้าๆ	ทุก ρ_k เป็น 0	ทุก ρ_{kk} เป็น 0
ARI(1,1)	ลดลงอย่างช้าๆ	ค่าลดลงเร็วใกล้ 0	ρ_{kk} เป็น 0 สำหรับ $k = 2, \dots$
ARI(2,1)	ลดลงอย่างช้าๆ	ค่าลดลงเร็วใกล้ 0	ρ_{kk} เป็น 0 สำหรับ $k = 3, \dots$
IMA(1,1)	ลดลงอย่างช้าๆ	ρ_k เป็น 0 สำหรับ $k = 2, \dots$	ค่า ρ_{kk} ลดลงเร็วใกล้ 0
IMA(1,2)	ลดลงอย่างช้าๆ	ρ_k เป็น 0 สำหรับ $k = 3, \dots$	ค่า ρ_{kk} ลดลงเร็วใกล้ 0
ARIMA(1,1,1)	ลดลงอย่างช้าๆ	ค่าลดลงเร็วใกล้ 0	ค่าลดลงเร็วใกล้ 0

สำหรับอนุกรมเวลาที่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โมเดลที่จะใช้ได้แก่ SARIMA(P,D,Q)_L (seasonal integrated autoregressive and moving average ที่ลำดับ P,D,Q) โดยที่ P เป็นอันดับของ SAR (seasonal autoregressive) Q เป็นอันดับของ SMA (seasonal moving average) และ D เป็นจำนวนครั้งที่หาผลต่างฤดูกาลเพื่อทำให้อนุกรมเวลา {y_t} ที่ไม่คงที่เนื่องจากฤดูกาลเป็นอนุกรมเวลาชุดใหม่ {Z_t} ที่เป็นอนุกรมเวลาที่คงที่ โดย $Z_t = \nabla_L^D Y_t$ ตัวอย่างของโมเดล SARMA(P,Q)_L ของอนุกรมเวลา {y_t} ดังตาราง 6 นอกจากนี้คุณสมบัติที่สำคัญของโมเดล SARMA(P,Q)_L เมื่อ L = 12 ของอนุกรมเวลาที่คงที่ ได้แก่ ρ_k และ ρ_{kk} ลักษณะของ ρ_k และ ρ_{kk} ดังตาราง 7

ตาราง 6 โมเดล SARMA(P,Q)_L ของอนุกรมเวลา {y_t}

โมเดล SARMA(P,Q) _L	อนุกรมเวลา {y _t }
SAR(1) _L	$(1 - \phi_L B^L)Y_t = \theta_0 + \varepsilon_t$ หรือ $Y_t = \theta_0 + \phi_L Y_{t-L} + \varepsilon_t$
SAR(2) _L	$(1 - \phi_L B^L - \phi_{2L} B^{2L})Y_t = \theta_0 + \varepsilon_t$ หรือ $Y_t = \theta_0 + \phi_L Y_{t-L} + \phi_{2L} Y_{t-2L} + \varepsilon_t$
SMA(1) _L	$Y_t = \theta_0 + (1 - \phi_L B^L)\varepsilon_t$ หรือ $Y_t = \theta_0 + \varepsilon_t - \theta_L \varepsilon_{t-L}$
SMA(2) _L	$Y_t = \theta_0 + (1 - \theta_L B^L - \theta_{2L} B^{2L})\varepsilon_t$ หรือ $Y_t = \theta_0 + \varepsilon_t - \theta_L \varepsilon_{t-L} - \theta_{2L} \varepsilon_{t-2L}$
SARMA(1,1) _L	$(1 - \phi_L B^L)Y_t = \theta_0 + (1 - \theta_L B^L)\varepsilon_t$ หรือ $\theta_0 + \phi_L Y_{t-L} + \varepsilon_t - \theta_L \varepsilon_{t-L}$

ตาราง 7 ลักษณะของ $\rho_k(Z_t)$ และ $\rho_{kk}(Z_t)$ สำหรับโมเดล SARMA(P,Q)₁₂

โมเดล	ลักษณะของ $\rho_k(Z_t)$	ลักษณะของ $\rho_{kk}(Z_t)$
SAR(1) ₁₂	$\rho_{12}, \rho_{24}, \dots$ มีค่าลดลงเร็วใกล้ 0	$\rho_{kk} = 0$ สำหรับ $k = 24, 36, \dots$
SAR(2) ₁₂	$\rho_{12}, \rho_{24}, \dots$ มีค่าลดลงเร็วใกล้ 0	$\rho_{kk} = 0$ สำหรับ $k = 36, 48, \dots$
SMA(1) ₁₂	$\rho_k = 0$ สำหรับ $k = 24, 36, \dots$	$\rho_{12,12}, \rho_{24,24}, \dots$ มีค่าลดลงเร็วใกล้ 0
SMA(2) ₁₂	$\rho_k = 0$ สำหรับ $k = 36, 48, \dots$	$\rho_{12,12}, \rho_{24,24}, \dots$ มีค่าลดลงเร็วใกล้ 0

ตัวอย่างของโมเดล SARIMA(P,1,Q)₁₂ ของอนุกรมเวลา {y_t} หรือโมเดล SARMA(P,Q)₁₂ ของอนุกรมเวลา {Z_t} เป็นดังนี้

$$1. Z_t \sim \text{SAR}(1)_{12} \text{ หรือ } Y_t \sim \text{SAR}(1,1)_{12} \text{ เมื่อ } Z_t = \nabla_{12} Y_t$$

$$Z_t = \theta_0 + \phi_{12} Z_{t-12} + \varepsilon_t$$

$$\text{หรือ } Y_t = \theta_0 + (1 + \phi_{12}) Y_{t-12} - \phi_{12} Y_{t-24} + \varepsilon_t$$

$$2. Z_t \sim \text{SMA}(1)_{12} \text{ หรือ } Y_t \sim \text{SIMA}(1,1)_{12} \text{ เมื่อ } Z_t = \nabla_{12} Y_t$$

$$Z_t = \theta_0 + (1 - \phi_{12} B^{12}) \varepsilon_t$$

$$\text{หรือ } Y_t = \theta_0 + Y_{t-12} + \varepsilon_t - \theta_{12} \varepsilon_{t-12}$$

$$3. Z_t \sim \text{SARMA}(1,1)_{12} \text{ หรือ } Y_t \sim \text{SARIMA}(1,1,1)_{12} \text{ เมื่อ } Z_t = \nabla_{12} Y_t$$

$$Z_t = \theta_0 + \phi_{12} Z_{t-12} + \varepsilon_t - \theta_{12} \varepsilon_{t-12}$$

$$\text{หรือ } Y_t = \theta_0 + (1 + \phi_{12}) Y_{t-12} - \phi_{12} Y_{t-24} + \varepsilon_t - \theta_{12} \varepsilon_{t-12}$$

สำหรับอนุกรมเวลาที่มีทั้งแนวโน้มและอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โมเดลที่จะใช้ได้แก่ ARIMA(p,d,q)×SARIMA(P,D,Q)_L ซึ่งอนุกรมเวลาจะมีลักษณะไม่คงที่ การกำจัดอิทธิพลของแนวโน้มและฤดูกาลจะทำได้โดยการหาผลต่างและผลต่างฤดูกาล เพื่อปรับอนุกรมเวลาเดิม {y_t} ที่ไม่คงที่ให้เป็นอนุกรมเวลาใหม่ {Z_t} ที่คงที่ เมื่อ d เป็นจำนวนครั้งที่หาผลต่าง และ D เป็นจำนวนครั้งที่หาผลต่างฤดูกาล โดยที่ $Z_t = \nabla^d \nabla_L^D Y_t$ เช่น อนุกรมเวลารายปี (L = 12) และมีโมเดล $Y_t \sim \text{ARI}(1,1) \times \text{SARI}(1,1)_{12}$ หรือ $Z_t \sim \text{AR}(1) \times \text{SAR}(1)_{12}$ สามารถเขียนเป็นสมการได้ดังนี้

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_{12} Z_{t-12} - \phi_1 \phi_{12} Z_{t-13} + \varepsilon_t$$

2. การประมาณค่าพารามิเตอร์ (Estimation the Parameters)

การประมาณค่าพารามิเตอร์จะทำได้ 2 วิธี ดังนี้ (ทรงศิริ แต่สมปิติ, 2539)

2.1 การประมาณแบบง่าย จะใช้สมการที่แสดงความสัมพันธ์ระหว่าง ρ_k และพารามิเตอร์ในโมเดล จำนวนสมการที่ใช้จะเท่ากับอันดับที่ปรากฏในโมเดล เมื่อประมาณ ρ_k ด้วย r_k ในสมการความสัมพันธ์และแก้สมการจะได้ค่าประมาณของพารามิเตอร์ ส่วนค่าประมาณของ θ_0 จะได้จากสมการที่แสดงความสัมพันธ์ระหว่าง $E(Y_t)$ กับพารามิเตอร์ในโมเดล เช่น

2.1.1 สำหรับ AR(1) จะหา $\hat{\phi}_1$ ได้จากการแก้สมการ $\rho_1 = \phi_1$ และ $\hat{\theta}_0$ ได้จากการแก้สมการ $\mu = \theta_0 / (1 - \phi_1)$ เมื่อแทน ρ_1 และ μ ด้วย r_1 และ \bar{Y} ในสมการความสัมพันธ์ทั้งสองตามลำดับ จะได้ค่าประมาณ $\hat{\theta}_0$ และ $\hat{\phi}_1$

2.1.2 สำหรับ AR(2) จะหา $\hat{\phi}_1$ และ $\hat{\phi}_2$ ได้จากการแก้สมการ $\rho_1 = \phi_1 / (1 - \phi_2)$ และ $\rho_2 = \phi_1 + \rho_1 + \phi_2$ และ $\hat{\theta}_0$ จะได้จากการแก้สมการ $\mu = \theta_0 / (1 - \phi_1 - \phi_2)$ เมื่อแทน ρ_1, ρ_2 และ μ ด้วย r_1, r_2 และ \bar{Y} ในสมการแสดงความสัมพันธ์ทั้งสาม จะได้ค่าประมาณ $\hat{\theta}_0, \hat{\phi}_1, \hat{\phi}_2$

2.1.3 สำหรับ MA(1) จะหา $\hat{\theta}_1$ ได้จากการแก้สมการ $\rho_1 = -\theta_1 / (1 + \theta_1^2)$ และ $\hat{\theta}_0$ ได้จาก $\mu = \theta_0$ เมื่อแทน ρ_1 และ μ ด้วย r_1 และ \bar{Y} ในสมการแสดงความสัมพันธ์ทั้งสองตามลำดับ จะได้ค่าประมาณ $\hat{\theta}_0$ และ $\hat{\theta}_1$ ซึ่ง $\hat{\theta}_1$ ที่ได้จะมีค่าสองค่า ค่าที่เลือกใช้จะเป็นค่าที่ทำให้โมเดลเป็นอินเวอร์ทิเบิล

2.2 การประมาณด้วยวิธีกำลังสองน้อยที่สุด จะต้องหาค่าประมาณที่ทำให้ Sum Square of Error มีค่าต่ำที่สุด

$$\text{เช่น AR(1)} \quad Y_t = \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

$$\text{เพราะฉะนั้น Sum Square of Error} = \sum (Y_t - \hat{\theta}_0 - \hat{\phi}_1 Y_{t-1})^2$$

ตาราง 8 สูตรที่ใช้ประมาณค่าพารามิเตอร์สำหรับรูปแบบ ARIMA(p,d,q)

ลำดับ (p,d,q)	พารามิเตอร์ที่จะประมาณค่า	สูตร	ขอบเขตของพารามิเตอร์
(0,d,1)	θ_1	$\rho_1 = -\theta_1 / (1 + \theta_1^2)$	$-1 < \theta_1 < 1$
(0,d,2)	θ_1 และ θ_2	$\rho_1 = [-\theta_1(1 - \theta_2)] / [1 + \theta_1^2 + \theta_2^2]$ $\rho_2 = -\theta_2 / [1 + \theta_1^2 + \theta_2^2]$	$-1 < \theta_2 < 1$ $\theta_2 - \theta_1 < 1$ $\theta_2 - \theta_1 < 1$
(1,d,0)	ϕ_1	$\phi_1 = \rho_1$	$-1 < \phi_1 < 1$
(2,d,0)	ϕ_1 และ ϕ_2	$\phi_1 = [1 - \rho_2] / [1 - \rho_1^2]$ $\phi_2 = [\rho_2 - \rho_1^2] / [1 - \rho_1]$	$-1 < \phi_2 < 1$ $\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$
(1,d,1)	ϕ_1 และ θ_1	$\rho_1 = [(1 - \theta_1 \phi_1)(\phi_1 - \theta_1)] / [1 + \theta_1^2 - 2\phi_1 \theta_1]$ $\rho_2 = \rho_1 \phi_1$	$-1 < \phi_1 < 1$ $-1 < \theta_1 < 1$

3. การตรวจสอบความเหมาะสมของโมเดล (Diagnostic Checking)

หลักในการพิจารณาการตรวจสอบความเหมาะสมของโมเดลมี 3 วิธี ได้แก่ (ทรงศิริ แต่สมบัติ, 2539)

3.1 พิจารณาว่าอนุกรมเวลา $\{e_t\}$ มีลักษณะการเคลื่อนไหวที่เป็นอิสระกันหรือไม่ นั่นคือเป็นการตรวจสอบข้อจำกัดของโมเดลที่กำหนดว่า ε_t จะต้องมีการแจกแจงที่เป็นอิสระกัน หากพบว่า $\{e_t\}$ มีลักษณะการเคลื่อนไหวที่เป็นอิสระกันจริง จะสรุปว่าโมเดลที่กำหนดนั้นเหมาะสมกับอนุกรมเวลาแล้วซึ่งทำได้โดยการทดสอบสมมติฐาน $H_0: \rho_k(e_t) = 0$ กับ $H_1: \rho_k(e_t) \neq 0$ สำหรับ $k = 1, 2, \dots$ การทดสอบนี้จะเป็นการทดสอบทำนองเดียวกันกับการทดสอบ $H_0: \rho_k(Y_t) = 0$ กับ $H_1: \rho_k(Y_t) \neq 0$ สำหรับ $k = 1, 2, \dots$ ที่ผ่านมา การปฏิเสธ H_0 ที่ระดับนัยสำคัญ .05 จะทำเมื่อ $|r_k(e_t)| \geq 2/\sqrt{k}$ เป็นขนาดของอนุกรมเวลา $\{e_t\}$

3.2 พิจารณาว่าค่าพารามิเตอร์ในโมเดลมีค่าเป็น 0 หรือไม่ นั่นคือเมื่อ $\theta, \hat{\theta}$ และ $S\hat{\theta}$ เป็นพารามิเตอร์ ค่าประมาณ และค่าคลาดเคลื่อนมาตรฐานของค่าประมาณ $\hat{\theta}$ ตามลำดับ การทดสอบสมมติฐาน $H_0: \theta = 0$ กับ $H_1: \theta \neq 0$ จะใช้ตัวทดสอบสถิติ $Z = \hat{\theta}/S\hat{\theta}$ การปฏิเสธ H_0 จะทำเมื่อ $|Z| \geq Z_{\alpha/2}$ ที่ระดับนัยสำคัญ α เช่น จากโมเดล AR(2) มีพารามิเตอร์ ϕ_1 และ ϕ_2 สำหรับทดสอบว่า ϕ_2 เป็น 0 หรือไม่ จะเป็นการทดสอบว่าแทนที่โมเดลจะเป็น AR(2) ควรจะเป็น AR(1) ใชหรือไม่ นั่นคือ

$$\begin{aligned} H_0: \phi_2 = 0 \text{ หรือ } Y_t &= \theta_0 + \phi_1 Y_{t-1} + \varepsilon_t && \text{หรือ } Y_t \sim \text{AR}(1) \\ \text{กับ } H_1: \phi_2 \neq 0 \text{ หรือ } Y_t &= \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t && \text{หรือ } Y_t \sim \text{AR}(2) \end{aligned}$$

3.3 พิจารณาว่า $\rho_k(e_t) = 0$ สำหรับ $k = 1, \dots, m$ หรือไม่ นั่นคือพิจารณาว่าค่าคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน $1, \dots, m$ ช่วงเวลาเป็นอิสระกันหรือไม่ จะกำหนดสมมติฐานเป็น $H_0: \rho_1(e_t) = \dots = \rho_m(e_t) = 0$ กับ $H_1: \rho_k(e_t)$ อย่างน้อย 1 ค่าไม่เป็น 0 สำหรับ $k = 1, \dots, m$ การทดสอบสมมติฐานจะทำโดยการทดสอบของ Box และ Pierce หรือการทดสอบของ Box และ Ljung ซึ่งต่างก็เป็นการทดสอบสมมติฐานที่กำหนดสมมติฐานหลักและสมมติฐานทางเลือกเหมือนกัน มีช่วงวิกฤตเดียวกัน แต่มีตัวทดสอบสถิติต่างกันนั่นคือ

$$3.3.1 \quad H_0: \rho_1(e_t) = \dots = \rho_m(e_t) = 0$$

$$H_1: \rho_k(e_t) \text{ อย่างน้อย 1 ค่าไม่เท่ากับ 0 สำหรับ } k = 1, \dots, m$$

3.3.2 ตัวทดสอบสถิติ

$$\text{Box - Pierce} \quad Q = n \sum_{k=1}^m r_k^2(e_t)$$

$$\text{Box - Ljung} \quad Q_m = n(n+2) \sum_{k=1}^m (r_k^2(e_t)) / (n-k)$$

- เมื่อ n = ขนาดของอนุกรมเวลา $\{e_t\}$
 m = ช่วงเวลาห่างสูงสุดของ e_t ในอนุกรมเวลา $\{e_t\}$ ที่นำมาพิจารณา
 a = จำนวนพารามิเตอร์ทั้งหมดในรูปแบบ ซึ่งรวมทั้ง θ_0 ด้วย

3.3.3 ช่วงวิกฤต Q หรือ $Q_m \geq \chi^2_{\alpha, n-a}$

3.3.4 สรุปผลการทดสอบ เมื่อปฏิเสธ H_0 จะสรุปได้ว่ามี $\rho_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เท่ากับ 0 สำหรับ $k = 1, \dots, m$ นั่นคือมีสหสัมพันธ์บางส่วนระหว่างค่าความคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน k ค่า หรือสรุปได้ว่ารูปแบบที่กำหนดให้กับอนุกรมเวลายังไม่เหมาะสม

4. การพยากรณ์ (Forecasting)

การพยากรณ์ค่าในอนาคตทำได้ 2 แบบ คือ การพยากรณ์แบบจุด (point forecasting) และการพยากรณ์แบบช่วง (interval forecasting) ซึ่งมีรายละเอียดดังต่อไปนี้ (ทรงศิริ แต่สมบัติ, 2539)

4.1 การพยากรณ์แบบจุด (point forecasting) ค่าพยากรณ์ในอนาคตจะได้จากการแทนค่าสังเกตในอดีตหรือค่าความคลาดเคลื่อนในอดีตในสมการพยากรณ์ ในสมการพยากรณ์จะแทนค่าประมาณของพารามิเตอร์ที่ได้ประมาณขึ้นในขั้นตอนที่สอง สำหรับโมเดล ARMA(p,q) ซึ่ง $p+q \leq 2$ มีสมการพยากรณ์ $\hat{Y}_t(l)$ ดังแสดงในตาราง 9 สมการพยากรณ์จะเขียนได้ทั้งในเทอมของพารามิเตอร์ μ และพารามิเตอร์อื่น หรือในเทอมของพารามิเตอร์ θ_0 และพารามิเตอร์อื่น การใช้สมการพยากรณ์จะทำโดยการแทนค่าประมาณของพารามิเตอร์ในรูปแบบ ลักษณะของโมเดล ARMA(p,q) ซึ่ง $p+q \leq 2$ เมื่อกำหนด $\hat{Y}_t(l)$ เป็นค่าพยากรณ์ ณ เวลา $t+1$ ใช้ค่าสังเกต t ค่า มี

$$\hat{Y}_t(l) = \begin{cases} Y_{t+1} & \text{เมื่อ } l \leq 0 \\ \hat{Y}_t(l) & \text{เมื่อ } l \geq 1 \end{cases}$$

และ $e_t(1)$ เป็นค่าคลาดเคลื่อนของการพยากรณ์ ณ เวลา $t+1$ ใช้ค่าสังเกต t ค่า มี

$$\hat{e}_t(1) = \begin{cases} e_{t+1} & \text{เมื่อ } l \leq 0 \\ 0 & \text{เมื่อ } l \geq 1 \end{cases}$$

4.2 การพยากรณ์แบบช่วง (interval forecasting) ณ เวลา $t + 1$ ค่าจริงเป็น Y_{t+1} จะหาค่าจริงที่เวลา $t + 1$ ไม่ได้ จะได้เพียงค่าพยากรณ์ $\hat{Y}_t(l)$ ซึ่งค่าพยากรณ์ $\hat{Y}_t(l)$ อาจจะมีค่าต่างจากค่าจริง Y_{t+1} มาก หรือค่าคลาดเคลื่อนของการพยากรณ์ ณ เวลา $t + 1$ $e_t(l)$ มีค่าสูง การพยากรณ์อีกแบบหนึ่งที่ทำให้ความเชื่อมั่นแก่ผู้พยากรณ์ในการพยากรณ์ค่า ณ เวลา $t + 1$ ได้แก่ การพยากรณ์แบบช่วง การพยากรณ์แบบช่วงนั้น ช่วงของการพยากรณ์ค่า Y_{t+1} จะขึ้นกับค่าพยากรณ์แบบจุด $\hat{Y}_t(l)$ การสร้างช่วงของการพยากรณ์จะใช้ผลที่ค่าคลาดเคลื่อนของการพยากรณ์ ณ เวลา $t + 1$ $e_t(l)$ มีการแจกแจงประมาณด้วยการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวน $\sigma_{e_t}^2(l)$ ดังนั้น $e_t(l)$ จะอยู่ในช่วงของการประมาณที่มีขีดจำกัด $\pm Z_{\alpha/2} \sigma_{e_t}(l)$ ที่ระดับความเชื่อมั่น $(1-\alpha)100$ เปอร์เซนต์ ซึ่งจะเขียนในรูปของความน่าจะเป็นได้เป็น

$$P(-Z_{\alpha/2} \sigma_{e_t}(l) < e_t(l) < Z_{\alpha/2} \sigma_{e_t}(l)) = 1 - \alpha$$

เมื่อแทน $e_t(l)$ ด้วย $Y_{t+1} - \hat{Y}_t(l)$ จะได้

$$P(\hat{Y}_t(l) - Z_{\alpha/2} \sigma_{e_t}(l) < Y_{t+1} < \hat{Y}_t(l) + Z_{\alpha/2} \sigma_{e_t}(l)) = 1 - \alpha$$

ดังนั้นค่าจริง Y_{t+1} จะอยู่ในช่วงของการพยากรณ์ที่มีขีดจำกัด $\hat{Y}_t(l) \pm Z_{\alpha/2} \sigma_{e_t}(l)$ ด้วยความเชื่อมั่น $(1 - \alpha)100$ เปอร์เซนต์ กรณีที่ไม่ทราบค่า $\sigma_{e_t}(l)$ จะใช้ค่าประมาณ $S_{e_t}(l)$ แทน

ตาราง 9 สมการพยากรณ์และค่าความแปรปรวนของค่าคลาดเคลื่อน

รูปแบบ	สมการพยากรณ์ $\hat{Y}_t(l)$		$\sigma_{e(l)}^2$
	ในเทอมของ $\hat{\mu}$	ในเทอมของ $\hat{\theta}_0$	
White noise	$\hat{\mu}$, $l \geq 1$	$\hat{\theta}_0$, $l \geq 1$	σ^2
AR(1)	$\hat{\mu} + \hat{\phi}_1(Y_t - \hat{\mu})$, $l \geq 1$	$\hat{\theta}_0 + \hat{\phi}_1 Y_t$, $l = 1$ $\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1)$, $l \geq 2$	$((1 - \phi_1^{2l}) / (1 - \phi_1^2)) \sigma^2$
AR(2)	$\hat{\mu} + \hat{\phi}_1(Y_t - \hat{\mu}) + \hat{\phi}_2(Y_{t-1} - \hat{\mu})$, $l = 1$ $\hat{\mu} + \hat{\phi}_1(\hat{Y}_t(l) - \hat{\mu}) + \hat{\phi}_2(Y_t - \hat{\mu})$, $l = 2$ $\hat{\mu} + \hat{\phi}_1(\hat{Y}_t(l-1) - \hat{\mu}) + \hat{\phi}_2(\hat{Y}_t(l-2) - \hat{\mu})$, $l \geq 3$	$\hat{\theta}_0 + \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1}$, $l = 1$ $\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l) + \hat{\phi}_2 Y_t$, $l = 2$ $\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1) + \hat{\phi}_2 \hat{Y}_t(l-2)$, $l \geq 3$	$(1 + \sum \psi_j^2) \sigma^2$ $\psi_0 = 1, \psi_1 = \phi_1$ $\psi_j = \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2}$, $j > 1$
MA(1)	$\hat{\mu} - \hat{\theta}_1 e_t$, $l = 1$ $\hat{\mu}$, $l \geq 2$	$\hat{\theta}_0 - \hat{\theta}_1 e_t$, $l = 1$ $\hat{\theta}_0$, $l \geq 2$	σ^2 , $l = 1$ $(1 + \theta_1^2) \sigma^2$, $l \geq 2$
MA(2)	$\hat{\mu} - \hat{\phi}_1 e_t - \hat{\theta}_2 e_{t-1}$, $l = 1$ $\hat{\mu} - \hat{\theta}_2 e_t$, $l = 2$ $\hat{\mu}$, $l \geq 3$	$\hat{\theta}_0 - \hat{\theta}_1 e_t - \hat{\theta}_2 e_{t-1}$, $l = 1$ $\hat{\theta}_0 - \hat{\theta}_2 e_t$, $l = 2$ $\hat{\theta}_0$, $l \geq 3$	σ^2 , $l = 1$ $(1 + \theta_1^2) \sigma^2$, $l = 2$ $(1 + \theta_1^2 + \theta_2^2) \sigma^2$, $l \geq 3$
ARMA(1,1)	$\hat{\mu} + \hat{\phi}_1(Y_t - \hat{\mu}) - \hat{\theta}_1 e_t$, $l = 1$ $\hat{\mu} + \hat{\phi}_1(Y_t - \hat{\mu})$, $l \geq 2$	$\hat{\theta}_0 + \hat{\phi}_1 Y_t - \hat{\theta}_1 e_t$, $l = 1$ $\hat{\theta}_0 + \hat{\phi}_1 \hat{Y}_t(l-1)$, $l \geq 2$	$[1 + (\phi_1 - \theta_1)^2 (1 - \phi_1^{2l-2}) / (1 - \phi_1^2)] \sigma^2$

ที่มา : ทงศิริ แต่สมบัติ (2539)

โมเดลอริมาอินเตอร์เวนชัน (ARIMA Intervention Model)

โมเดลอริมาอินเตอร์เวนชันได้รับการพัฒนาโดย Box และ Tiao (1975), Campbell (1963) และ Campbell กับ Stanley (1966) ซึ่งเป็นโมเดลที่ใช้กันอย่างแพร่หลายในการวิจัยทางสังคมศาสตร์ นอกจากนี้โมเดลอริมาอินเตอร์เวนชันยังเป็นกรณีเฉพาะ (special case) ของ Transfer Function อีกด้วย เราสามารถจำแนกเหตุการณ์ที่ทำให้เกิดอินเตอร์เวนชันออกเป็น 2 ประเภท ดังนี้

1. Common intervention คือ อินเตอร์เวนชันที่มักเกิดขึ้นและมีการเปลี่ยนแปลงอยู่บ่อยๆ ได้แก่ การออกกฎหมายใหม่ (new laws) , เหตุการณ์ทางเศรษฐศาสตร์ (economic events) , การเปลี่ยนแปลงนโยบาย , แผ่นดินไหว (earthquakes) , สงคราม (wars) , ความเสียหายทางธรรมชาติ (natural disasters)

2. Company-specific intervention คือ อินเตอร์เวนชันที่เกิดขึ้นเพื่อผลประโยชน์ทางธุรกิจ ได้แก่ การส่งเสริมผลผลิต (product promotions) , การแนะนำผลิตภัณฑ์ใหม่ (new product introductions) , การรณรงค์ให้มีการโฆษณา (major advertising campaign)

วัตถุประสงค์หลักของอริมาอินเตอร์เวนชัน คือ เพื่อประมาณผลกระทบ รวมไปถึงการเกิดอินเตอร์เวนชันและลดผลกระทบของอินเตอร์เวนชันที่มีต่ออนุกรมเวลาในการพยากรณ์ในส่วนที่ไม่เกิดอินเตอร์เวนชันและยังช่วยในการปรับปรุงความถูกต้องของการพยากรณ์หลังการเกิดอินเตอร์เวนชันอีกด้วย (Newbold and Bos, 1994 ; Makridakis , Wheelwright and Hyndman, 1998 ; ทรงศิริ แต่สมบัติ, 2539)

วิธีการของอริมาอินเตอร์เวนชันมีขั้นตอนดังต่อไปนี้

1. การกำหนดโมเดลก่อนการเกิดอินเตอร์เวนชัน

มีวิธีการคำนวณเช่นเดียวกับขั้นตอนการกำหนดโมเดลในวิธีของบ็อกซ์และเจนกินส์ ดังรายละเอียดที่กล่าวไว้แล้วในตอนต้น

2. การตรวจสอบความเหมาะสมของโมเดลก่อนการเกิดอินเตอร์เวนชัน

มีวิธีการคำนวณเช่นเดียวกับขั้นตอนการตรวจสอบความเหมาะสมของโมเดลในวิธีของบ็อกซ์และเจนกินส์ ดังรายละเอียดที่กล่าวไว้แล้วในตอนต้น

3. การกำหนดโมเดลอินเตอร์เวนชัน

องค์ประกอบของโมเดล ARIMA Intervention

Intervention Model = Intervention Function + ARIMA noise Model

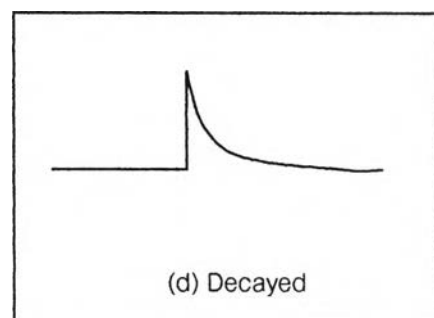
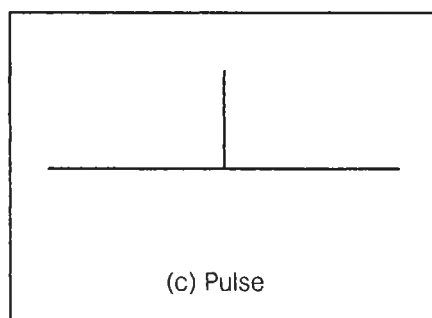
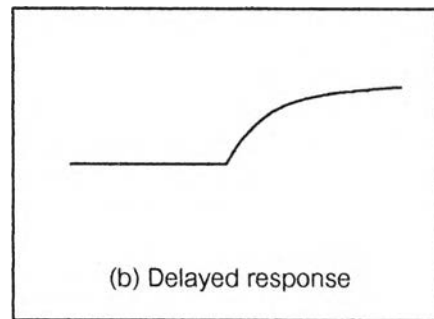
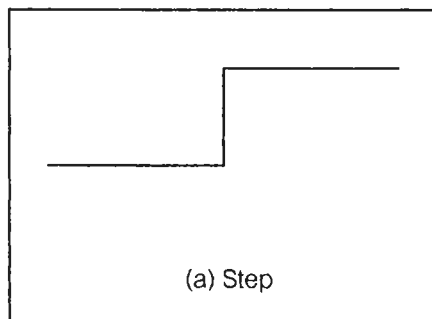
$$Y_t = f(I_t) + N_t$$

เมื่อ intervention function ถูกระบุโดย $f(I_t)$

$N_t =$ โมเดล ARIMA(p,d,q) \times SARIMA(P,D,Q)_L

$$I_t = \begin{cases} 1 & \text{เมื่อเกิดอินเตอร์เวนชัน} \\ 0 & \text{เมื่อไม่เกิดอินเตอร์เวนชัน} \end{cases}$$

โดยตัวแปรดัมมี่ I_t ที่มีค่าเป็น 0 หรือ 1 อาจจะมีลักษณะเป็นแบบ Step - based interventions หรือ Pulse - Based interventions ดังแผนภาพ 9



แผนภาพ 9 รูปแบบของตัวแปรอินเตอร์เวนชัน

Step – based interventions

รูปแบบทั่วไปของอินเตอร์เวนชันจะเป็น step functions และสมมติให้อินเตอร์เวนชันเกิดขึ้น ณ เวลา u เราสามารถให้คำจำกัดความตัวแปรดัมมี ได้ดังนี้

$$X_t = \begin{cases} 0, & t < u \\ 1, & t \geq u \end{cases}$$

เป็น 0 เมื่อเป็นเหตุการณ์ก่อนการเกิดอินเตอร์เวนชัน (preintervention)
 1 เมื่อเป็นเหตุการณ์หลังการเกิดอินเตอร์เวนชัน (postintervention)

ซึ่งเรียกว่า step intervention เพราะกราฟของ X_t เมื่อเทียบกับ t จะคล้ายคลึงกับขั้นบันไดดังแผนภาพ 9a

ในบางกรณีพบว่าอินเตอร์เวนชันจะให้ผลที่เพิ่มขึ้นหรือลดลงในตัวแปรพยากรณ์แต่จะไม่ให้ผลในทันทีทันใด การกำหนดรูปแบบในตัวแปรพยากรณ์จะใช้ Transfer Function ดังสมการต่อไปนี้

$$v(B) = \frac{\omega}{1 - \delta B} X_t$$

เมื่อ X_t เป็น step variable ซึ่งผลของการตอบสนองจะแสดงดังแผนภาพ 9b

Pulse – based interventions

อินเตอร์เวนชันบางประเภทจะมีผลกระทบชั่วคราวต่ออนุกรมเวลา ในกรณีนี้เรานิยมใช้ pulse function มากกว่า step function ซึ่งเราสามารถให้คำจำกัดความ pulse variable ได้ดังนี้

$$X_t = \begin{cases} 0, & t \neq u \\ 1, & t = u \end{cases}$$

โดยที่ $X_t = 0$ เมื่อไม่เกิดอินเตอร์เวนชัน
 $X_t = 1$ เมื่อเกิดอินเตอร์เวนชัน

ผลกระทบของอินเตอร์เวนชันจะตอบสนอง ณ จุดที่เกิดอินเตอร์เวนชันแต่จะไม่ตอบสนองในบริเวณอื่น ซึ่งรูปร่างของอินเตอร์เวนชันที่ตอบสนองจะแสดงดังแผนภาพ 9c สิ่งนี้อาจเป็นประโยชน์ต่อการกำหนดรูปแบบเหตุการณ์ที่ผิดปกติ เช่น การสไตรค์ของคนงานในโรงงานอุตสาหกรรม สำหรับอินเตอร์เวนชันซึ่งมีผลกระทบอย่างรวดเร็ว แต่ค่อยๆ ลดลงอย่างช้าๆ เราสามารถนิยามโมเดลได้ดังนี้

$$Y_t = a + \frac{\omega}{1 - \delta B} X_t + N_t$$

ลักษณะของอินเตอร์เวนชันจะแสดงดังแผนภาพ 9d โดยที่ค่าของ ω จะเป็นขนาดของการเพิ่มขึ้นอย่างรวดเร็ว และ Y_t และ σ จะเป็นอัตราของการลดลง

รูปแบบของอริมาอินเตอร์เวนชัน

แบ่งออกได้เป็น 2 แบบใหญ่ๆ ดังนี้

แบบที่ 1 Zero-Order Intervention Functions

แผนภาพ 10 แสดง common intervention function ที่จำแนกออกเป็น 4 ประเภท ซึ่งเรียกว่า "Zero-order intervention functions" รายละเอียดมีดังต่อไปนี้

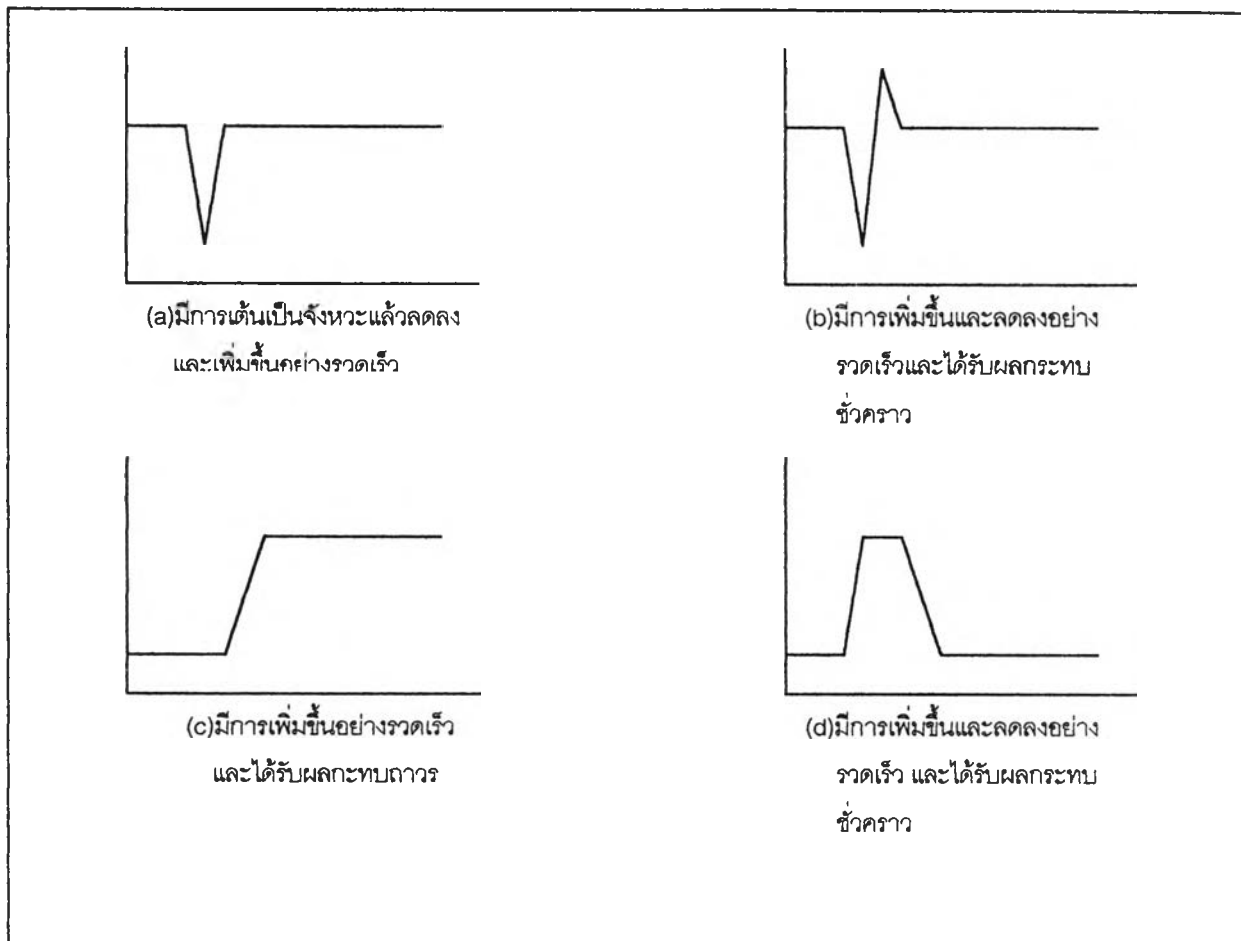
ประเภทที่ 1 (แผนภาพ 10a) เป็น simple intervention model ซึ่งลดลงแล้วเพิ่มขึ้นอย่างรวดเร็วและเกิดผลกระทบชั่วคราว ตัวอย่างเช่น เมื่อผลกระทบของอินเตอร์เวนชันมีจำนวนมากและไม่ทราบสาเหตุของการเกิดอินเตอร์เวนชัน เราจะเรียกว่าเกิดเหตุการณ์ที่มีลักษณะเป็น "outlier" ในขณะที่ outlier บางตัวจะมีลักษณะเป็น eyeballed ที่มีค่าปกติ (normal value) นอกจากนี้โมเดลอินเตอร์เวนชัน (Intervention Model) ยังสามารถนำมาใช้ในการประมาณค่าผลกระทบของ outlier ดังสมการต่อไปนี้

$$Y_t = \theta_0 + \omega_0 I_t$$

เมื่อ $I_t = 1$ สำหรับ $t = 9$ นอกเหนือจากนั้นมีค่าเป็น 0

$\theta_0 =$ ค่าศูนย์กลางปกติ (Normal central value) ของอนุกรม

$\omega_0 =$ ผลกระทบของอินเตอร์เวนชัน



แผนภาพ 10 รูปแบบ Zero-order intervention function

ประเภทที่ 2 (แผนภาพ 10b) เป็น complex intervention model เขียนเป็นสมการได้

ดังนี้

$$\begin{aligned}
 Y_t &= \theta_0 + (\omega_0 - \omega_1 B) I_t \\
 &= \theta_0 + \omega_0 I_t - \omega_1 I_{t-1}
 \end{aligned}$$

เมื่อ $I_t = 1$ สำหรับ $t = 9$ นอกเหนือจากนั้นมีค่าเป็น 0

θ_0 = ค่าศูนย์กลางปกติ (Normal central value)

ω_0 = ผลกระทบขั้นต้นของอินเตอร์เวนชัน (Initial intervention effect)

$-\omega_1$ = ผลกระทบของอินเตอร์เวนชันที่ลดลง (Delayed intervention effect)

ประเภทที่ 3 (แผนภาพ 10c) เป็น simple intervention model ที่มีการเพิ่มขึ้นอย่างรวดเร็วในลักษณะถาวร ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \omega_0 I_t$$

เมื่อ $I_t = 1$ สำหรับ $t > 8$ นอกเหนือจากนั้นเป็น 0

ประเภทที่ 4 (แผนภาพ 10d) เป็น different intervention model ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + (1 - B^5)\omega_0 I_t$$

เมื่อ $I_t = 1$ สำหรับ $t > 4$ นอกเหนือจากนั้นเป็น 0

แบบที่ 2 First-Order Intervention Functions

แผนภาพ 11 แสดง intervention model ซึ่งแบ่งออกเป็น 4 ประเภท ซึ่งมีตัวหารของฟังก์ชันอินเตอร์เวนชันเพิ่มเข้ามา ได้แก่ $\delta_1, 1-\delta_1 B$ และ $\delta_1 B$ ซึ่งตัวหารเหล่านี้จะทำให้เกิดอินเตอร์เวนชันที่มีลักษณะค่อยๆ เพิ่มขึ้นหรือลดลง ดังแผนภาพ 11 โดยที่ $\delta_1 B$ จะระบุสถานภาพที่เป็น first - order function เมื่อ first order มีความสัมพันธ์กับอำนาจของ B (power of B) ซึ่งมีรายละเอียดที่น่าสนใจดังต่อไปนี้

ประเภทที่ 1 (แผนภาพ 11a) แสดงการค่อยๆ ลดลงและได้รับผลกระทบถาวรของ first-order transfer function ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \frac{\omega_0}{(1 - \delta_1 B)} I_t \quad \dots\dots\dots(28)$$

คูณทั้ง 2 ข้างของสมการ (28) ด้วย $(1 - \delta_1 B)$

$$(1 - \delta_1 B)Y_t = (1 - \delta_1 B)\theta_0 + \omega_0 I_t \quad \dots\dots\dots(29)$$

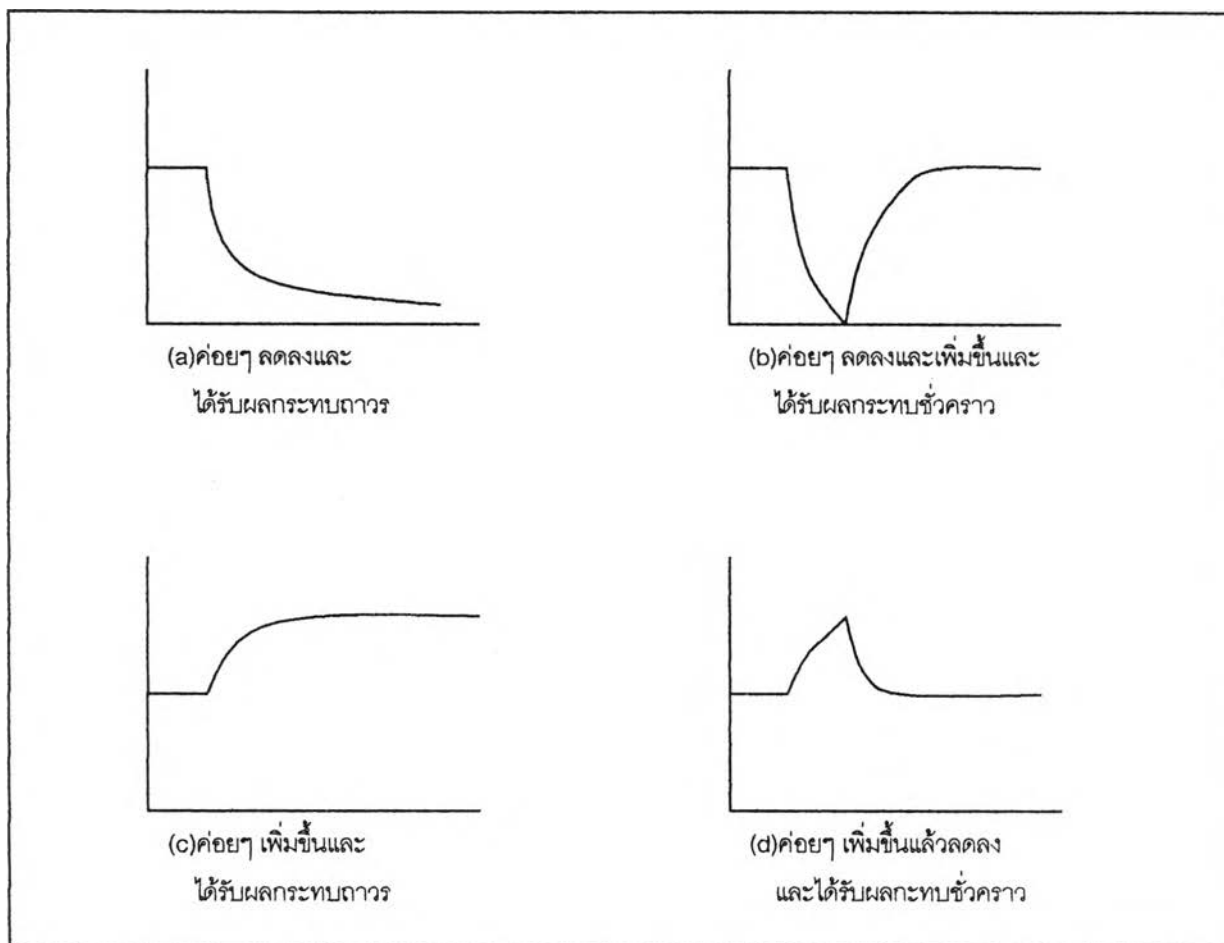
เพราะ backshift ของ θ_0 เป็นค่าคงที่ ดังนั้นจึงตัด B ออกจากสมการ (29) จะได้ว่า

$$Y_t - \delta_1 Y_{t-1} = (1 - \delta_1)\theta_0 + \omega_0 I_t$$

$$Y_t = \delta_1 Y_{t-1} + (1 - \delta_1)\theta_0 + \omega_0 I_t$$

จากแผนภาพ 11a เป็นรูปแบบของโมเดลที่ลดลงแบบเอ็กซ์โปเนนเชียล เมื่อผลกระทบมีค่าลดลง Y_t ซึ่งการลดลงของค่าสังเกตนี้เราเรียกว่า Taylor Series เมื่อ n มีค่าเข้าใกล้ ∞ ดังนั้นผลของฟังก์ชันอินเตอร์เวนชันที่ได้คือ asymptotic impact value

$$\text{asymptotic value} = \lim_{t \rightarrow \infty} \frac{\omega_0}{(1 - \delta_1)}$$



แผนภาพ 11 รูปแบบ First-order intervention functions

ประเภทที่ 2 (แผนภาพ 11b) แสดงการค่อยๆ ลดลงแล้วเพิ่มขึ้นและได้รับผลกระทบชั่วคราวของ first-order transfer function ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \frac{\omega_0(1-B^5)}{(1-\delta_1 B)} I_t \quad \dots\dots\dots(30)$$

คูณทั้ง 2 ข้างของสมการด้วย $1-\delta_1 B$ จะได้ว่า

$$(1-\delta_1 B)Y_t = (1-\delta_1 B)\theta_0 + \omega_0(1-B^5)I_t$$

$$Y_t - \delta_1 Y_{t-1} = (1-\delta_1)\theta_0 + \omega_0(1-B^5)I_t$$

$$Y_t = \delta_1 Y_{t-1} + (1-\delta_1)\theta_0 + \omega_0(1-B^5)I_t$$

ประเภทที่ 3 (แผนภาพ 11c) แสดงการค่อยๆ เพิ่มขึ้นและได้รับผลกระทบแบบถาวรของ first – order transfer function ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \frac{\omega_0}{(1-\delta_1 B)} I_t \quad \dots\dots\dots(31)$$

จะเห็นได้ว่า สมการ (28) เหมือนกับสมการ (31) แต่ต่างกันตรงที่ค่าของ ω_0 ในสมการ (28) มีค่าเป็นลบ ส่วนค่าของ ω_0 ในสมการ (31) มีค่าเป็นบวก

ประเภทที่ 4 (แผนภาพ 11d) แสดงการค่อยๆ เพิ่มขึ้นแล้วลดลง และได้รับผลกระทบแบบชั่วคราวของ first – order transfer function ซึ่งเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \frac{\omega_0(1-B^5)}{(1-\delta_1 B)} I_t \quad \dots\dots\dots(32)$$

จะเห็นได้ว่า สมการ (30) เหมือนกับสมการ (32) แต่ต่างกันตรงที่ค่าของ ω_0 ในสมการ (30) มีค่าเป็นลบ ส่วนค่าของ ω_0 ในสมการ (32) มีค่าเป็นบวก

รูปแบบทั่วไปของอริมาอินเตอร์เวนชัน

รูปแบบของฟังก์ชันอินเตอร์เวนชันดังแผนภาพ 10 และ 11 เป็นกรณีเฉพาะของ Transfer Function ที่นำเสนอโดย Box, Jenkins และ Reinsel (1994) แต่เมื่อพิจารณาารูปแบบทั่วไปแล้วจะมีลักษณะดังนี้

$$Y_t = \frac{\omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s}{(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)} X_{t-b} + N_t$$

เมื่อ y_t และ x_t = ค่าเบี่ยงเบนจากค่าเฉลี่ยของอนุกรมเวลาที่คงที่หรืออนุกรมเวลาที่ไมคงที่แต่แปลงเป็นอนุกรมเวลาที่คงที่แล้ว

$$N_t = \frac{(1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)}{(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)} \times e_t$$

- r = ระดับของกระบวนการถดถอยในตัวเองหรือ อำนาจของ B ในตัวหาร
 $s+1$ = จำนวนของเทอม (ω) ในตัวเศษ
 b = การลดผลกระทบของอินเตอร์เวนชันโดยคุณด้วย X_{t-b}

4. การประมาณค่าพารามิเตอร์จากโมเดลอริมาอินเตอร์เวนชัน

การประมาณค่าพารามิเตอร์จะทำได้ 2 วิธี ดังนี้

4.1 การประมาณแบบง่าย

4.2 การประมาณด้วยวิธีกำลังสองน้อยที่สุด

รายละเอียดในการคำนวณได้กล่าวไว้แล้วในขั้นตอนการประมาณค่าพารามิเตอร์ด้วยวิธีของบ็อกซ์และเจนกินส์

5. การตรวจสอบความเหมาะสมของโมเดลอริมาอินเตอร์เวนชัน

นำโมเดลอริมาอินเตอร์เวนชันที่ได้จากขั้นตอนที่ 3 มาตรวจสอบความเหมาะสมโดยการทำ autocorrelation check ของค่าความคลาดเคลื่อน โดยมีเงื่อนไขดังนี้

ยอมรับ H_0 → โมเดลอริมาอินเตอร์เวนชันที่กำหนดมีความเหมาะสม

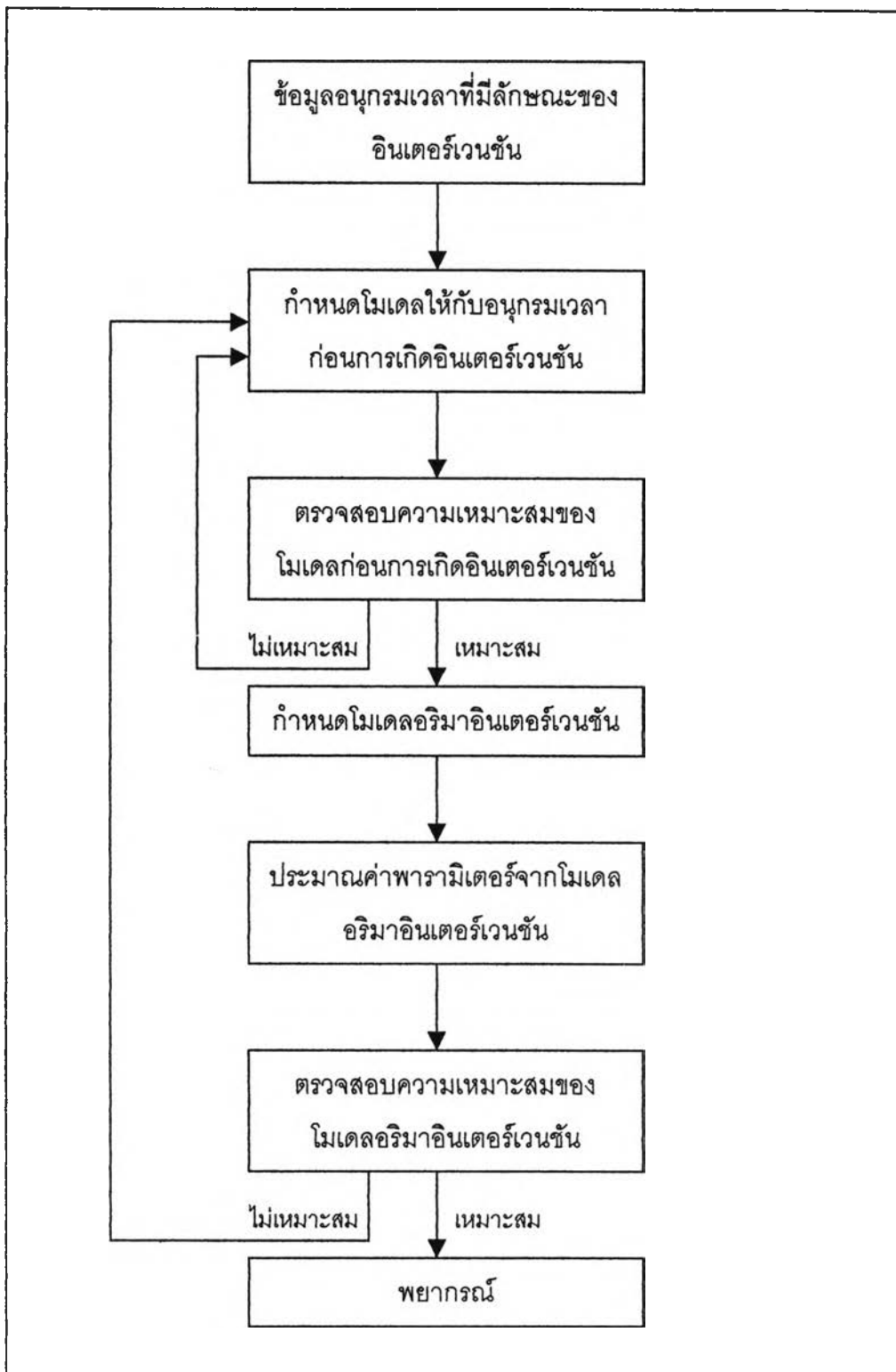
ปฏิเสธ H_0 → โมเดลอริมาอินเตอร์เวนชันที่กำหนดไม่มีความเหมาะสม

6. การพยากรณ์

การพยากรณ์ค่าในอนาคตทำได้ 2 วิธี ดังนี้

1. การพยากรณ์แบบจุด (point forecasting)
2. การพยากรณ์แบบช่วง (interval forecasting)

รายละเอียดในการคำนวณได้กล่าวไว้แล้วในขั้นตอนการพยากรณ์ด้วยวิธีของบ็อกซ์และเจนกินส์



แผนภาพ 12 ขั้นตอนการวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีอริมาอินเตอร์เวนชัน (ARIMA Intervention Analysis)

ตอนที่ 3 มโนทัศน์เกี่ยวกับการวิเคราะห์การถดถอย

การวิเคราะห์การถดถอย (Regression Analysis) เป็นวิธีการที่ได้ชื่อมาจาก Sir Francis Galton เมื่อปี ค.ศ. 1899 ซึ่งเป็นการศึกษาในเชิงเปรียบเทียบเกี่ยวกับแนวโน้มของลักษณะความสูงของบุตรหลานที่บิดามารดาสูงและเตี้ย ปรากฏว่าความสูงของบุตรหลานจะถอยเข้าหาส่วนสูงเฉลี่ยของประชากร ดังนั้นการวิเคราะห์ถดถอยจึงเป็นวิธีการที่ใช้ในการพิจารณาถึงรูปแบบที่เป็นไปได้ของความสัมพันธ์ระหว่างตัวแปร มีวัตถุประสงค์เพื่อใช้เป็นประโยชน์ในการพยากรณ์ (Prediction) หรือการประมาณ (Estimation) ค่าของตัวแปรตัวหนึ่งซึ่งเรียกว่าตัวแปรตามหรือตัวแปรเกณฑ์ (Dependent Variable) โดยใช้ค่าของตัวแปรอีกตัวหนึ่งหรืออีกชุดหนึ่งที่กำหนดให้ ซึ่งเรียกว่า ตัวแปรอิสระหรือตัวแปรทำนาย (Independent Variable) เป็นตัวประมาณ

3.1 การวิเคราะห์การถดถอยอย่างง่าย

การวิเคราะห์การถดถอยอย่างง่ายเป็นวิธีการทางสถิติที่ใช้ในการประมาณค่าและการพยากรณ์ค่าของตัวแปรตัวหนึ่งโดยใช้ค่าของข้อมูลอีกตัวหนึ่งเป็นตัวพยากรณ์ ตัวแปรที่ใช้ในการพยากรณ์เรียกว่าตัวแปรอิสระ (Independent Variable) หรือตัวพยากรณ์ (Predictor) ส่วนผลที่ได้เรียกว่าตัวแปรตาม (Dependent Variable) หรือผลที่วัดได้ (Outcome)

สมการการถดถอย (Regression Equation)

$$Y_i = \alpha + \beta_1 X_i + \epsilon_i$$

Y_i = ค่าของตัวแปรตามที่ได้จากการพยากรณ์

α = ค่า Y-intercept (จุดที่เส้นกราฟตัดแกน Y) ของกลุ่มประชากร

β_1 = สัมประสิทธิ์การถดถอยของกลุ่มประชากร

X_i = ค่าของตัวแปรอิสระหรือตัวพยากรณ์

ϵ_i = ค่าความคลาดเคลื่อนที่ i

การคำนวณค่าสัมประสิทธิ์การถดถอยและค่า Y-intercept

สัมประสิทธิ์การถดถอย

$$\beta = \frac{N(\sum XY) - (\sum X)(\sum Y)}{N(\sum X^2) - (\sum X)^2}$$

เมื่อ N = จำนวนสมาชิกทั้งหมด

X = ค่าของตัวแปรอิสระ

Y = ค่าของตัวแปรตาม

ค่า Y-intercept

$$\alpha = Y - \beta X$$

เมื่อ X = ค่าเฉลี่ยของตัวแปรอิสระ

Y = ค่าเฉลี่ยของตัวแปรตาม

การทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอย

หลังจากที่คำนวณค่าสัมประสิทธิ์การถดถอยของกลุ่มตัวอย่างแล้วต้องทดสอบความมีนัยสำคัญของสัมประสิทธิ์การถดถอยเพื่ออ้างอิงกลับไปหากกลุ่มประชากรว่าตัวแปรอิสระนั้นสามารถใช้ในการพยากรณ์ค่าของตัวแปรตามได้

$$r = b \left(\frac{S_x}{S_y} \right)$$

ค่าสถิติ

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$$

เมื่อ r = สัมประสิทธิ์สหสัมพันธ์

n = จำนวนคู่หรือจำนวนสมาชิก

3.2 การวิเคราะห์การถดถอยพหุคูณ

ถ้ามีตัวแปรอิสระ k ตัว (X_1, X_2, \dots, X_k) ที่มีความสัมพันธ์กับตัวแปรตาม Y โดยที่ความสัมพันธ์อยู่ในรูปเชิงเส้น จะได้สมการการถดถอยพหุคูณดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

โดยที่

k = จำนวนตัวแปรอิสระที่ใช้ในการถดถอย

Y = ค่าของตัวแปรตาม (Dependent Variable)

X_k = ค่าของตัวแปรอิสระ (Independent Variable) ตัวที่ k

β_0 = ส่วนตัดแกน Y เมื่อกำหนดให้ $X_1 = X_2 = \dots = X_k = 0$

$\beta_1, \beta_2, \dots, \beta_k$ = สัมประสิทธิ์ความถดถอยเชิงส่วน (Partial Regression Coefficient)

ε = ค่าความแตกต่างหรือความคลาดเคลื่อน (Error or Residual) ของการประมาณค่า Y

การวิเคราะห์การถดถอยพหุคูณในกรณีที่มีตัวแปรอิสระมากกว่าสองตัว

การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยใช้วิธีการทางพีชคณิตในการแก้สมการ เพื่อหาค่าสัมประสิทธิ์ $\beta_1, \beta_2, \dots, \beta_k$ มีความยุ่งยากในการคำนวณมาก จึงนำวิธีการทางเมตริกซ์มาใช้ในการวิเคราะห์การถดถอยพหุคูณ ซึ่งจะช่วยให้การวิเคราะห์ทำได้สะดวกและลดความยุ่งยากซับซ้อนลง การวิเคราะห์การถดถอยพหุคูณที่มีตัวแปรอิสระหลายๆ สูตรที่ผ่านมามักกล่าวจะยุ่งยาก วิธีการที่ใช้ทุกๆ ไปคือใช้หลักของเมตริกซ์ซึ่งเขียนเป็นสูตรได้ดังนี้

$$R\beta = R_y$$

เมื่อ R = เมตริกซ์ของค่าสหสัมพันธ์ระหว่างตัวแปรอิสระด้วยกัน

β = เวกเตอร์ของค่าสัมประสิทธิ์การถดถอยหรือ Beta weight

R_y = เวกเตอร์ของค่าสหสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระแต่ละตัว

จากสูตรสามารถเขียนในรูปเมตริกซ์ ได้ดังนี้

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1k} \\ r_{21} & r_{22} & r_{23} & \cdots & r_{2k} \\ r_{31} & r_{32} & r_{33} & \cdots & r_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & r_{k3} & \cdots & r_{kk} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} r_{y1} \\ r_{y2} \\ r_{y3} \\ \vdots \\ r_{yk} \end{bmatrix}$$

ทั้งหมด จากสูตรข้างต้นเมื่อใช้วิธีการคูณเมตริกซ์มาใช้ ก็จะแก้สมการหาค่า $\beta_1, \beta_2, \dots, \beta_k$ ได้

ขั้นตอนในการคำนวณด้วยเมตริกซ์ มีดังนี้

1. เขียนสมการให้อยู่ในรูป $R\beta = R_y$
2. เมื่อต้องการหาค่าของ β ต้องนำ inverse ของ R คูณทั้งสองข้าง
จะได้ $\hat{\beta} = R^{-1}R_y$

ความคลาดเคลื่อนมาตรฐานของการประมาณค่า (Standard error of estimate)

การหาความคลาดเคลื่อนในการพยากรณ์

$$SE_{est} = \sqrt{\frac{SS_{res}}{n-k-1}}$$

$$= \sqrt{\frac{\sum (y - \hat{y})^2}{n-k-1}}$$

เมื่อ SE_{est} = ค่าความคลาดเคลื่อนมาตรฐานในการพยากรณ์

SS_{res} = ค่าของ $\sum (y - \hat{y})^2$

n = จำนวนข้อมูล

k = จำนวนตัวแปรอิสระ

df = $n-k-1$

การทดสอบความมีนัยสำคัญของค่าสัมประสิทธิ์การถดถอยพหุคูณ

การทดสอบความมีนัยสำคัญของสมการการถดถอยพหุคูณ โดยใช้ R^2 เป็นการทดสอบว่าค่าสหสัมพันธ์แบบพหุในประชากรเท่ากับ 0 หรือไม่ สถิติที่ใช้ทดสอบคือ

$$F = \frac{R^2 / k}{(1 - R^2)(n - k - 1)}$$

เมื่อ k = จำนวนตัวแปรอิสระ

n = จำนวนกลุ่มตัวอย่าง

df_1 = k

df_2 = $n - k - 1$

การทดสอบความมีนัยสำคัญของค่าสัมประสิทธิ์การถดถอยพหุคูณ คือ การทดสอบว่ามีตัวแปรใดบ้างที่มีอิทธิพลต่อตัวแปรตามอย่างมีนัยสำคัญ สถิติที่ใช้ทดสอบความมีนัยสำคัญของค่าน้ำหนักตัวพยากรณ์ คือ

$$t_{bi} = \frac{b_i}{S_{bi}}$$

เมื่อ t_{bi} = ค่า t-ratio

b_i = สัมประสิทธิ์สหสัมพันธ์การถดถอยพหุคูณของตัวแปรอิสระตัวที่ i

S_{bi} = ค่าความคลาดเคลื่อนมาตรฐานของค่าสัมประสิทธิ์การถดถอยพหุคูณของตัวแปรอิสระตัวที่ i

df = $n - k - 1$

3.3 การวิเคราะห์การถดถอยกับอนุกรมเวลา

การวิเคราะห์อนุกรมเวลาที่ใช้หลักของการวิเคราะห์การถดถอย โดยอนุกรมเวลาเป็นตัวแปรตามและมีตัวแปรอิสระที่มีความหมายต่างๆ กันดังนี้ (ทรงศิริ แต่สมบัติ, 2539)

1. ตัวแปรเวลา (time variable) เป็นตัวแปรที่ใช้เมื่อพิจารณาแล้วว่าอนุกรมเวลามีแนวโน้ม ค่าของตัวแปรเวลาที่ใช้ในการวิเคราะห์จะเป็นค่ารหัส (coded value) ที่ทำให้การคำนวณทำได้ง่าย รวดเร็ว และสามารถอธิบายสมการแนวโน้มได้ง่าย ซึ่งการกำหนดให้ตัวแปรอิสระเป็นตัวแปร

เวลา (t) นั้นไม่ว่ารูปแบบการถดถอยจะเป็นแบบใด จะต้องกำหนดค่าให้กับตัวแปรเวลาก่อนซึ่งมีขั้นตอนดังนี้

1) t จะมีค่าเริ่มต้นเป็นเท่าใดก็ได้ แต่ค่า t จะต้องเพิ่มขึ้นเท่าๆ กันในแต่ละช่วงเวลา เช่น อนุกรมเวลารายปีที่เก็บรวบรวมตั้งแต่ปี 2530 ถึง 2534 มีจำนวนปีที่เก็บรวบรวมเท่ากับ 5 ปี จะกำหนดค่าของ t ที่มีหน่วยเป็นปี และเพิ่มขึ้นทีละ 1 ได้หลายลักษณะ เช่น ค่าของตัวแปรเวลา t_1 , t_2 , และ t_3 ดังตารางต่อไปนี้

ตาราง 10 ค่าสังเกตและค่าของตัวแปรเวลา t_1 , t_2 , และ t_3

ปี	ค่าสังเกต	t_1	t_2	t_3
2530	Y_1	-2	0	1
2531	Y_2	-1	1	2
2532	Y_3	0	2	3
2533	Y_4	1	3	4
2534	Y_5	2	4	5

จากตาราง 10 จะเห็นได้ว่า t_1 มีค่าเป็น 0 ที่ปี 2532, t_2 มีค่าเป็น 0 ที่ปี 2530 และ t_3 มีค่าเป็น 0 ที่ปี 2539 ซึ่งการกำหนดค่าให้กับตัวแปร t_1 จะมีประโยชน์ต่อการคำนวณค่าประมาณของพารามิเตอร์ในโมเดล เพราะผลรวมของค่า t มีค่าเท่ากับ 0 ($\sum t = \sum t^3 = 0$)

2) ค่าสังเกตที่เก็บรวบรวมในหนึ่งช่วงเวลาจะถือเป็นตัวแทนของช่วงเวลา ณ กลางช่วงเวลานั้น ได้แก่

2.1) อนุกรมเวลารายปี แต่ละค่าสังเกตเป็นค่ากลางปี หรือต้นเดือน

กรกฎาคม (7/1/XX)

2.2) อนุกรมเวลารายไตรมาส ในแต่ละปีจะมีค่าสังเกต 4 ค่า แต่ละค่าสังเกตเป็นค่ากลางไตรมาสหรือกลางเดือนกุมภาพันธ์ กลางเดือนพฤษภาคม กลางเดือนสิงหาคม และกลางเดือนพฤศจิกายน (2/15/XX, 5/15/XX, 8/15/XX, 11/15/XX)

2.3) อนุกรมเวลารายเดือน ในแต่ละปีจะมีค่าสังเกต 12 ค่า แต่ละค่าจะเป็นค่ากลางเดือนในแต่ละเดือน (1/15/XX, ..., 12/15/XX)

รูปแบบแนวโน้มของอนุกรมเวลา $\{y_t\}$ เมื่อตัวแปร t แทนเวลา มีลักษณะต่างๆ กันดังต่อไปนี้

1. แนวโน้มแบบเส้นตรง มีรูปแบบเป็น $Y_t = \beta_0 + \beta_1 t + \varepsilon_t$

2. แนวโน้มแบบ quadratic มีรูปแบบเป็น $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$ โดยที่ค่าสูงสุดหรือค่าต่ำสุดของแนวโน้มเกิดขึ้นเมื่อ t มีค่าเท่ากับ $-\beta_1/2\beta_2$ การพยากรณ์โดยใช้รูปแบบนี้ในช่วงเวลาล่วงหน้ามากอาจคลาดเคลื่อนได้ เนื่องจากแนวโน้มทางขึ้นอาจเปลี่ยนเป็นทางลง หรือแนวโน้มทางลงอาจเปลี่ยนเป็นทางขึ้น ผลต่างของค่าสังเกตครั้งที่ 1 มีแนวโน้มเส้นตรงที่มีจุดตัดแกนตั้งที่ $\beta_1 + \beta_2$ และมีความลาดชันของเส้นแนวโน้มเส้นตรงเท่ากับ $2\beta_2$ ผลต่างของค่าสังเกตของอนุกรมครั้งที่ 2 จะไม่มีแนวโน้ม แต่จะมีค่าอยู่รอบค่าคงที่ค่าหนึ่ง การใช้รูปแบบแนวโน้มที่มีลำดับสูงกว่า quadratic จะทำได้ แต่จะมีปัญหาของการพยากรณ์ค่าในอนาคตมากขึ้นกว่ากรณีการใช้รูปแบบแนวโน้มแบบ quadratic ดังนั้นจึงมักจะไม่มีการใช้รูปแบบแนวโน้มที่สูงกว่า quadratic

3. แนวโน้มแบบ exponential มีรูปแบบเป็น $Y_t = \beta_0 + \beta_1^t \varepsilon_t$ มีลักษณะแนวโน้มเป็นเส้นโค้ง โดยมี β_0 เป็นค่าคงที่ และ $\beta_1 = 1 + \beta_2$ โดย β_2 เป็นสัดส่วนที่เพิ่มขึ้นหรือลดลงต่อหน่วยของ t ถ้า β_2 มีค่าเป็นบวกแสดงว่าเมื่อ t เพิ่มขึ้น 1 หน่วย ค่าสังเกตจะมีค่าเพิ่มขึ้นร้อยละ $\beta_2 \times 100$ หน่วย ถ้า β_2 มีค่าเป็นลบแสดงว่าเมื่อ t มีค่าเพิ่มขึ้น 1 หน่วย ค่าสังเกตจะมีค่าลดลงร้อยละ $\beta_2 \times 100$ หน่วย

4. แนวโน้มแบบ modified exponential มีรูปแบบเป็น $Y_t = \beta_0 + \beta_1 \beta_2^t \varepsilon_t$ โดย β_2 จะเป็นค่าสูงสุดหรือต่ำสุดที่ Y_t มีค่าได้ และจุดตัดของเส้นโค้งบนแกน Y คือ $\beta_0 + \beta_2$ โดยที่เส้นแนวโน้มแบบ modified exponential จะแสดงลักษณะของแนวโน้มที่เพิ่มขึ้นหรือลดลงของค่าสังเกตตามเวลา

5. แนวโน้มแบบ Gompertz มีรูปแบบเป็น $Y_t = \beta_0 \beta_1^{\beta_2^t} \varepsilon_t$ มีแนวโน้มแบบเส้นโค้งลักษณะแบบตัว S ซึ่งรูปแบบเส้นโค้งแบบนี้เป็นลักษณะของโค้งการเจริญเติบโต (growth curve) ที่เพิ่มขึ้นด้วยอัตราที่เพิ่มขึ้นจนถึงจุดหนึ่งแล้วจะเพิ่มขึ้นด้วยอัตราที่ลดลง เส้นโค้งจะตัดแกนตั้งที่ $\beta_0 \beta_1$ และมีค่าไม่เกิน β_0

6. แนวโน้มแบบ logistic มีรูปแบบเป็น $Y_t = 10^{\beta_0 + \beta_1 t + \beta_2 \varepsilon_t}$ มีแนวโน้มแบบเส้นโค้งลักษณะแบบตัว S แทนโค้งการเจริญเติบโตที่มีลักษณะทำนองเดียวกับแนวโน้มแบบ Gompertz เส้นโค้งจะตัดแกนตั้งที่ $1/(\beta_0 + \beta_1)$ และมีค่าไม่เกิน $1/\beta_0$

2. ตัวแปรดัมมี่ (dummy variable) คือตัวแปรที่สร้างขึ้นเพื่อวัดอิทธิพลของแนวโน้มและ/หรือฤดูกาล จำนวนตัวแปรดัมมี่ที่สร้างขึ้นจะมีจำนวนเท่ากับ $L - 1$ โดยที่ L เป็นจำนวนฤดูกาลต่อปี ตัวอย่างเช่น กรณีอนุกรมเวลารายเดือน ใน 1 ปีจะมี 12 ฤดูกาลหรือ 12 เดือน ดังนั้น L จะเท่ากับ 12 และในกรณีของอนุกรมเวลารายไตรมาส ใน 1ปีจะมี 4 ฤดูกาลหรือ 4 ไตรมาส ดังนั้น L จะเท่ากับ 4 เราจะกำหนดค่าของตัวแปรดัมมี่ X_i สำหรับ $i = 1, \dots, L-1$ ดังนี้

$$X_{it} = \begin{cases} 1 & \text{เมื่อค่าสังเกต } Y_t \text{ อยู่ในฤดูกาลที่ } i \\ 0 & \text{เมื่อค่าสังเกต } Y_t \text{ ไม่อยู่ในฤดูกาลที่ } i \end{cases}$$

จะกำหนดรูปแบบเส้นตรงที่อธิบาย L ฤดูกาลได้ดังนี้

$$Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_{L-1} X_{L-1,t} + \varepsilon_t$$

หรือ
$$Y_t = \alpha + \beta_1 X_{1t} + \dots + \beta_{L-1} X_{L-1,t} + \beta_L X_{Lt} + \varepsilon_t$$

เมื่อ α = ค่า Y - intercept

β_i = ค่าวัดอิทธิพลของฤดูกาลที่ i เมื่อเทียบกับฤดูกาลที่ L สำหรับ

$i = 1, \dots, L-1$ เมื่อ β_i เป็นบวกแสดงว่าฤดูกาลที่ i มีอิทธิพลเหนือ

ฤดูกาลที่ L และเมื่อ β_i เป็นลบแสดงว่าฤดูกาลที่ i มีอิทธิพลน้อยกว่า

ฤดูกาลที่ L

รูปแบบที่แสดงแนวโน้มและอิทธิพลของฤดูกาล L ฤดูกาล มีรูปแบบต่างๆ ดังต่อไปนี้

1.กรณีที่แนวโน้มเป็นแบบเส้นตรง และมีการรวมตัวกับฤดูกาลเป็นแบบบวก

$$Y_t = \alpha + \beta t + \beta_1 X_{1t} + \dots + \beta_{L-1} X_{L-1,t} + \varepsilon_t$$

เมื่อ α = ค่า Y - intercept

β = อัตราการเพิ่มขึ้นหรือลดลงของ Y_t เมื่อ t เพิ่มขึ้น 1 หน่วย

β_l = ค่าวัดอิทธิพลของฤดูกาลที่ l เมื่อเทียบกับฤดูกาลที่ L ที่กำหนดสำหรับ $l = 1, \dots, L-1$

2. กรณีที่มีแนวโน้มเป็นแบบ quadratic และมีการรวมตัวกับฤดูกาลเป็นแบบบวกจะมีรูปแบบเป็น

$$Y_t = \alpha + \beta t + \gamma t^2 + \beta_1 X_{1,t} + \dots + \beta_{L-1} X_{L-1,t} + \varepsilon_t$$

โดยที่รูปแบบทางขวาจะประกอบไปด้วยส่วนประกอบ 3 ส่วน คือ ส่วนที่ 1 จะอธิบายแนวโน้มแบบ quadratic ได้แก่ $\alpha + \beta t + \gamma t^2$ ส่วนที่ 2 จะอธิบายอิทธิพลของฤดูกาล L ฤดูกาล โดย β_l เป็นค่าวัดอิทธิพลของฤดูกาลที่ l และส่วนที่ 3 จะอธิบายความคลาดเคลื่อนที่เกิดขึ้นหรือปัจจัยที่มีอิทธิพลต่อ Y ที่นอกเหนือจากแนวโน้มและฤดูกาล โดยที่ทั้ง 3 ส่วนมีการรวมตัวกันแบบบวก

3. กรณีที่มีแนวโน้มแบบ exponential และมีการรวมตัวกับฤดูกาลเป็นแบบคูณ

$$Y_t = \alpha \beta^t \beta_1^{X_{1,t}} \dots \beta_{L-1}^{X_{L-1,t}} \varepsilon_t$$

หรือ

$$Y_t = \alpha \beta^t \beta_1^{X_{1,t}} \dots \beta_{L-1}^{X_{L-1,t}} \beta_L^{X_{L,t}} \varepsilon_t$$

สำหรับรูปแบบทางด้านขวาจะประกอบไปด้วยส่วนประกอบ 3 ส่วน คือ ส่วนประกอบที่ 1 จะอธิบายแนวโน้มแบบ exponential ได้แก่ $\alpha \beta^t$ ส่วนที่ 2 จะอธิบายอิทธิพลของฤดูกาล โดยที่ β_l เป็นดัชนีฤดูกาลที่ l เมื่อมีฤดูกาลที่ L เป็นฐาน กำหนดให้ β_L เป็นดัชนีฤดูกาลของฤดูกาลที่ L มีค่าเป็น 1 กรณีที่ β_l มีค่าน้อยกว่า 1 แสดงว่าฤดูกาลที่ l มีอิทธิพลน้อยกว่าฤดูกาลที่ L และส่วนที่ 3 จะอธิบายความคลาดเคลื่อนที่เกิดขึ้นหรือปัจจัยที่มีอิทธิพลต่อ Y ที่นอกเหนือจากแนวโน้มและฤดูกาล โดยที่ทั้ง 3 ส่วนมีการรวมตัวกันแบบคูณ

จากรูปแบบคูณนี้จะแปลงให้อยู่ในรูปแบบเชิงเส้นตรงได้โดยการหาลอการิทึมของ Y_t ($\ln Y_t$) จะได้รูปแบบของ $\ln Y_t$ เป็นรูปแบบเชิงเส้นตรง

$$\ln Y_t = \ln \alpha + (\ln \beta) t + (\ln \beta_1) X_{1,t} + \dots + (\ln \beta_{L-1}) X_{L-1,t} + \ln \varepsilon_t$$

หรือ

$$Y_t' = \alpha' + \beta' t + \beta_1' X_{1,t} + \dots + \beta_{L-1}' X_{L-1,t} + \varepsilon_t'$$

โดยที่ $Y_t' = \ln Y_t$, $\alpha' = \ln \alpha$, $\beta' = \ln \beta$, $\beta_1' = \ln \beta_1$, ... และ $\beta_{L-1}' = \ln \beta_{L-1}$ จะสร้างสมการพยากรณ์ได้จากรูปแบบใหม่โดยการประมาณค่าพารามิเตอร์ในรูปแบบใหม่ α' , β' , β_1'

และ β_{L-1}' โดยวิธีกำลังสองน้อยที่สุดได้ค่าประมาณ a' , b' , b_1' , ... และ b_{L-1}' ได้สมการพยากรณ์จากรูปแบบเชิงเส้นตรงดังนี้

$$\hat{Y}_t = a' + b't + b_1' X_{1t} + \dots + b_{L-1}' X_{L-1,t}$$

จะแปลงรูปแบบเชิงเส้นตรงนี้ให้เป็นสมการพยากรณ์ที่มาจากรูปแบบที่มีแนวโน้มแบบ exponential และมีการรวมตัวกับฤดูกาลเป็นแบบคูณ ดังนี้

$$\hat{Y}_t = ab^t b_1^{X_{1t}} \dots b_{L-1}^{X_{L-1,t}}$$

ซึ่งจะหา a , b , b_1 , ... , b_{L-1} ได้จาก $a = e^a$, $b = e^b$, $b_1 = e^{b_1}$, ... , และ $b_{L-1} = e^{b_{L-1}}$ โดยที่ตัวแปรต้นมีที่สร้างขึ้นจำนวน $L - 1$ ตัวแปรนั้น จะแทนฤดูกาลได้ก็จำนวน $L - 1$ ฤดูกาล ไม่จำเป็นจะต้องแทนเฉพาะฤดูกาลที่ 1 ถึง $L - 1$ เสมอไป

จากรูปแบบที่กำหนดจะหาสมการแนวโน้มและ/หรือฤดูกาลได้โดยวิธีกำลังสองน้อยที่สุด แต่ก่อนที่จะสร้างสมการควรได้มีการตรวจสอบอนุกรมเวลาก่อนว่ามีแนวโน้มและ/หรือฤดูกาลจริงหรือไม่ โดยการวิเคราะห์การถดถอยซึ่งทำตามขั้นตอนดังต่อไปนี้ (ทรงศิริ แต่สมบัติ , 2541)

1) ทดสอบว่ามีแนวโน้มหรือไม่เมื่อมีฤดูกาลอยู่แล้ว โดยการทดสอบแบบ t บางส่วน หรือแบบ F บางส่วน

$$\text{โดยที่ } t = b_1 / S_{b_1}$$

$$\text{และช่วงวิกฤตคือ } |t| \geq t_{\alpha, 2, n-5}$$

2) หากพบว่ามีแนวโน้มจึงทดสอบว่ามีฤดูกาลด้วยหรือไม่ โดยการทดสอบแบบ F บางส่วนจากรูปแบบแนวโน้มและฤดูกาล

$$\text{โดยที่ } F_1 = [\text{SSR}(S_1, S_2, S_3/t)/3] / [\text{SSE}(t, S_1, S_2, S_3)/(n-5)]$$

$$\text{และช่วงวิกฤตคือ } F_1 \geq F_{\alpha, (3, n-5)}$$

3) หากพบว่าไม่มีแนวโน้มจากผลการทดสอบสมมติฐานในข้อ 1 จะทดสอบว่ามีฤดูกาลเพียงอย่างเดียวหรือไม่ โดยการทดสอบแบบ F

$$\text{โดยที่ } F_2 = [\text{SSR}(S_1, S_2, S_3)/3] / [\text{SSE}(S_1, S_2, S_3)/(n-4)]$$

$$\text{และช่วงวิกฤตคือ } F_2 \geq F_{\alpha, (3, n-4)}$$

จากการทดสอบสมมติฐานตามขั้นตอนที่ได้กล่าวถึงจะได้ผลสรุปว่าการเคลื่อนไหวของอนุกรมเวลาเนื่องมาจากแนวโน้มและ/หรือฤดูกาลหรือไม่ ซึ่งทำให้กำหนดรูปแบบให้กับอนุกรมเวลาที่จะนำไปวิเคราะห์ต่อไปได้อย่างถูกต้อง หากอนุกรมเวลามีการเคลื่อนไหวเนื่องจากอิทธิพลของฤดูกาลเพียงอย่างเดียว จะใช้รูปแบบฤดูกาลและสร้างสมการฤดูกาลด้วยวิธีกำลังสองน้อยที่สุดเป็น

$$\hat{Y}_t = b_0 + b_1X_{1t} + b_2X_{2t} + b_3X_{3t}$$

เมื่อ $b_0 =$ ค่าคงที่

$b_l =$ ค่าวัดอิทธิพลของฤดูกาลที่ l สำหรับ $l = 1, 2, 3$

การอธิบายความหมายของ b_l จะอธิบายเทียบกับฤดูกาลที่ 4 โดยสมมติว่ามีเทอม b_4X_{4t} ในสมการฤดูกาล b_4 มีค่าเป็น 0 สำหรับ b_l ที่มีค่าเป็นบวกแสดงว่าด้วยอิทธิพลของฤดูกาลที่ l ค่าสังเกตในฤดูกาลที่ l มีค่าสังเกตสูงกว่าในฤดูกาลที่ 4 เท่ากับ $|b_l|$ หน่วย และสำหรับ b_l ที่มีค่าเป็นลบแสดงว่าด้วยอิทธิพลของฤดูกาลที่ l ค่าสังเกตในฤดูกาลที่ l มีค่าต่ำกว่าค่าสังเกตในฤดูกาลที่ 4 อยู่ $|b_l|$ หน่วย สมการฤดูกาลข้างต้นจะเขียนได้อีกแบบหนึ่งเมื่อ $\bar{b} = \sum_{i=1}^4 b_i/4$ เป็น

$$\hat{Y}_t = (b_0 + \bar{b}) + (b_1 - \bar{b})X_{1t} + (b_2 - \bar{b})X_{2t} + (b_3 - \bar{b})X_{3t} - \bar{b}X_{4t}$$

หรือจะเขียนเป็นสมการฤดูกาลใหม่ได้เป็น

$$\hat{Y}_t = b'_0 + b'_1X_{1t} + b'_2X_{2t} + b'_3X_{3t} + b'_4X_{4t}$$

เมื่อ $b'_0 = b_0 + \bar{b}$, $b'_l = b_l - \bar{b}$ สำหรับ $l = 1, 2, 3$ และ $b'_4 = -\bar{b}$ โดยที่ b'_0 เป็นค่าคงที่และ b'_l เป็นค่าวัดอิทธิพลของฤดูกาลที่ l สำหรับ $l = 1, 2, 3, 4$ ดังนั้นเมื่อ b_l มีค่าเป็นบวกแสดงว่าอิทธิพลของฤดูกาลที่ l ทำให้ค่าสังเกตมีค่าสูงกว่าปกติคือ \bar{b} หรือ 0 เท่ากับ $|b'_l|$ หน่วยและเมื่อ b_l มีค่าน้อยกว่า 0 แสดงว่าอิทธิพลของฤดูกาลที่ l ทำให้ค่าสังเกตมีค่าต่ำกว่าปกติคือ \bar{b} หรือ 0 เท่ากับ $|b'_l|$ หน่วย

3. ตัวแปรอิสระเป็น lag ของตัวแปรตาม จะเรียกรูปแบบที่มีลักษณะดังกล่าวว่ารูปแบบการถดถอยแบบออโต (autoregressive model) ตัวอย่างของรูปแบบการถดถอยแบบออโตที่ตัวแปรเป็น lag ที่ 1 ของตัวแปรตาม หรือรูปแบบ AR(1) ตัวแปรอิสระเป็น lag ที่ 1 และ lag ที่ 2 ของตัวแปรตาม หรือรูปแบบ AR(2) รูปแบบเป็นดังนี้

$$\text{รูปแบบ AR(1)} \quad Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

$$\text{รูปแบบ AR(2)} \quad Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \varepsilon_t$$

4.ตัวแปรตรีโกณมิติ จะใช้เพื่ออธิบายการเคลื่อนไหวของฤดูกาล เนื่องจากการเคลื่อนไหวของอนุกรมเวลาเนื่องจากฤดูกาลเป็นการเคลื่อนไหวแบบคลื่นที่มีแบบแผนเดียวกันในแต่ละปี การเคลื่อนไหวแบบคลื่นนี้จะอธิบายได้ด้วยฟังก์ชันตรีโกณมิติของเวลา ได้แก่ $\sin(2\pi t/L)$, $\sin(4\pi t/L)$, $\cos(2\pi t/L)$, $\cos(4\pi t/L)$ โดยที่ L เป็นจำนวนฤดูกาลต่อปี จะกำหนดรูปแบบแนวโน้มและฤดูกาลในเทอมของตัวแปรเวลาและตัวแปรตรีโกณมิติ แยกเป็นกรณีรูปแบบบวกและรูปแบบคูณดังนี้

4.1 กรณีรูปแบบบวก จะกำหนดรูปแบบแนวโน้มและฤดูกาลได้ดังนี้

$$Y_t = \beta_0 + \beta_1 t + \beta_2 \sin(2\pi t/L) + \beta_3 \cos(2\pi t/L) + \varepsilon_t$$

$$\text{หรือ} \quad Y_t = \beta_0 + \beta_1 t + \beta_2 \sin(2\pi t/L) + \beta_3 \cos(2\pi t/L) + \beta_4 \sin(4\pi t/L) + \beta_5 \cos(4\pi t/L) + \varepsilon_t$$

4.2 กรณีรูปแบบคูณ จะกำหนดรูปแบบแนวโน้มและฤดูกาลได้ดังนี้

$$Y_t = \beta_0 + \beta_1 t + \beta_2 \sin(2\pi t/L) + \beta_3 \cos(2\pi t/L) + \beta_4 t \sin(2\pi t/L) + \beta_5 t \cos(2\pi t/L) + \varepsilon_t$$

หรือ

$$Y_t = \beta_0 + \beta_1 t + \beta_2 \sin(2\pi t/L) + \beta_3 \cos(2\pi t/L) + \beta_4 \sin(4\pi t/L) + \beta_5 \cos(4\pi t/L) \\ + \beta_6 t \sin(2\pi t/L) + \beta_7 t \cos(2\pi t/L) + \beta_8 t \sin(4\pi t/L) + \beta_9 t \cos(4\pi t/L) + \varepsilon_t$$

ตอนที่ 4 ความคลาดเคลื่อนในการพยากรณ์

หลังจากที่มีการพยากรณ์ข้อมูลอนุกรมเวลาและประมาณค่าพารามิเตอร์แล้ว นักวิจัยต้องตรวจสอบผลการพยากรณ์ โดยกระบวนการตรวจสอบมี 2 ขั้นตอน คือ

1. การตรวจสอบฟังก์ชันอัตโนมัติสหสัมพันธ์ของโมเดลอนุกรมเวลา ที่สามารถเปรียบเทียบเทียบกับฟังก์ชันอัตโนมัติสหสัมพันธ์ของกลุ่มตัวอย่างในอนุกรมแรก

2. การวัดความคลาดเคลื่อน (Pindyck and Rubinfeld , 1985 อ้างถึงใน บำเพ็ญ ปิตชิต , 2540) วิธีการวัดความคลาดเคลื่อนแบ่งออกเป็น 2 กลุ่ม ได้แก่ กลุ่มแรกคือการวัดความคลาดเคลื่อนในรูปความแปรปรวน กลุ่มที่สองคือการวัดความคลาดเคลื่อนในรูปความคลาดเคลื่อนเฉลี่ย ซึ่งการวัดความคลาดเคลื่อนที่ใช้ในการวิจัยครั้งนี้มี 6 แบบ โดยค่าวัดความคลาดเคลื่อน 5 แบบ ได้แก่

RMSE, GMRAE, MdAPE, MdRAE และ Percent Better ไม่มีเกณฑ์ที่ใช้ในการพิจารณาความเหมาะสม ในขณะที่ค่าวัดความคลาดเคลื่อนในการพยากรณ์ด้วยวิธี MAPE มีเกณฑ์ในการพิจารณาความเหมาะสมดังนี้ (Lewis, 1982 อ้างถึงใน เอกภพ ยานะวิมุตติ, 2543)

ค่าที่ได้	ความหมาย
น้อยกว่า 10%	มีความถูกต้องสูง
10 – 20%	มีความถูกต้อง
20 – 50%	ยังยอมรับได้
มากกว่า 50%	ไม่มีความถูกต้อง

หลังจากการคำนวณผลจากการพยากรณ์ทั้ง 3 วิธีแล้วถ้าพบว่าวิธีใดมีค่า RMSE, MAPE, GMRAE, MdRAE และ MdRAE น้อยที่สุด และมีค่า Percent Better มากที่สุด แสดงว่าวิธีนั้นเป็นวิธีที่มีความคลาดเคลื่อนในการพยากรณ์น้อยที่สุด โดยรายละเอียดเกี่ยวกับการวัดความคลาดเคลื่อนทั้ง 6 แบบ มีดังนี้

กำหนดให้	m	=	วิธีการพยากรณ์
	\hat{Y}_t	=	ค่าพยากรณ์ที่เวลา t
	F_{rw}	=	ค่าพยากรณ์จากวิธีเดินอย่างสุ่ม (random walk)
	Y_t	=	ค่าสังเกตที่เวลา t
	S	=	ขนาดของอนุกรมเวลา

แบบที่ 1 รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย

(Root Mean Squared Error : RMSE)

$$RMSE_m = \left(\frac{\sum_{t=1}^S (\hat{Y}_t - Y_t)^2}{S} \right)^{\frac{1}{2}}$$

แบบที่ 2 ค่ามัธยฐานของค่าสัมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ
(Median Absolute Percentage Error : MdAPE)

$$APE_m = \left| \frac{\hat{Y}_t - Y_t}{Y_t} \right| \times 100$$

เมื่อนำค่าสังเกต APE_m มาเรียงลำดับ จะได้

$$MdAPE_m = (S+1)/2 \quad \text{เมื่อ } S \text{ เป็นจำนวนคี่}$$

$$MdAPE_m = (S/2)+1 \quad \text{เมื่อ } S \text{ เป็นจำนวนคู่}$$

แบบที่ 3 ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ
(Mean Absolute Percentage Error : MAPE)

$$APE_m = \left| \frac{\hat{Y}_t - Y_t}{Y_t} \right| \times 100$$

$$MAPE_m = \frac{\sum_{t=1}^S APE_m}{S}$$

แบบที่ 4 ค่าเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์สัมพัทธ์

(Geometric Mean of the Relative Absolute Error : GMRAE)

$$RAE_m = \frac{|\hat{Y}_t - Y_t|}{|F_{rw} - Y_t|}$$

$$GMRAE_m = \left[\prod_{t=1}^S RAE_m \right]^{\frac{1}{S}}$$

แบบที่ 5 ค่ามัธยฐานของค่าสัมบูรณ์สัมพัทธ์ของความคลาดเคลื่อน
(Median Relative Absolute Error : MdRAE)

$$RAE_m = \frac{|\hat{Y}_t - Y_t|}{|F_{rw} - Y_t|}$$

เมื่อนำค่าสังเกต RAE_m มาเรียงลำดับ จะได้

$$MdRAE_m = (S+1)/2 \quad \text{เมื่อ } S \text{ เป็นจำนวนคี่}$$

$$MdRAE_m = (S/2)+1 \quad \text{เมื่อ } S \text{ เป็นจำนวนคู่}$$

แบบที่ 6 ร้อยละที่ดีกว่า (Percent Better)

$$Percent\ Better_m = \frac{\sum_{s=1}^S j_s}{S} \times 100$$

$$\text{เมื่อ } j_s = 1 \text{ ถ้า } |\hat{Y}_t - Y_t| < |F_{rw} - Y_t|$$

$$j_s = 0 \text{ ในกรณีอื่นๆ}$$

ตาราง 11 การเปรียบเทียบประสิทธิภาพการวัดความคลาดเคลื่อนในการพยากรณ์

เกณฑ์ในการเปรียบเทียบ การวัดความคลาดเคลื่อน	ความเที่ยง	ความตรงเชิง โครงสร้าง	การป้องกัน ค่าสุดโต่ง	ความไว	ความสัมพันธ์ต่อ การตัดสินใจ
RMSE	P	F	P	G	G
Percent Better	G	F	G	P	P
MAPE	F	G	P	G	F
MdAPE	F	G	G	P	F
GMRAE	F	G	F	G	P
MdRAE	F	G	G	P	P

ที่มา : J. Scott Armstrong และ Fred Collopy (1992)

เมื่อ P = Poor (ต่ำ)

F = Fair (ปานกลาง)

G = Good (สูง)

จากตาราง 11 จะเห็นได้ว่าการวัดความคลาดเคลื่อนทั้ง 6 แบบมีทั้งข้อดีและข้อด้อยแตกต่างกันไป ตัวอย่างเช่น RMSE มีความไว ความสัมพันธ์ต่อการตัดสินใจและความตรงเชิงโครงสร้างอยู่ในเกณฑ์ที่ดีแต่มีความเที่ยงและการป้องกันค่าสุดโต่งอยู่ในเกณฑ์ที่ต่ำ แต่ในทางกลับกัน MdRAE มีความไว ความสัมพันธ์ต่อการตัดสินใจอยู่ในเกณฑ์ที่ต่ำ แต่มีความเที่ยง การป้องกันค่าสุดโต่งและความตรงเชิงโครงสร้างอยู่ในเกณฑ์ที่ดี และการวัดความคลาดเคลื่อน 4 แบบที่เหลือจะมีข้อดีและข้อด้อยแตกต่างกันไปซึ่งมีลักษณะดังเช่นการวัดความคลาดเคลื่อนทั้ง 2 แบบที่กล่าวมา ดังนั้นผู้วิจัยจึงเลือกใช้การวัดความคลาดเคลื่อนทั้ง 6 แบบ คือ RMSE, Percent Better, MAPE, MdAPE, GMRAE และ MdRAE เพื่อเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์สำหรับการวิจัยในครั้งนี้

ตอนที่ 5 งานวิจัยที่เกี่ยวข้อง

ชะไมพร ธรรมวัฒน์ไพศาล (2522) ทำการวิจัยเรื่องวิธีการประมาณค่าที่ขาดหายไปในการวิเคราะห์การถดถอย โดยวิธีที่ศึกษามี 5 วิธี ได้แก่ วิธีกำลังสองน้อยที่สุด วิธีอันดับศูนย์ วิธีอันดับศูนย์ดัดแปลง วิธีถดถอยอันดับหนึ่ง และวิธีถดถอยสองชั้น และเพื่อให้เห็นขั้นตอนในการดำเนินการโดยวิธีต่างๆ อย่างชัดเจน จึงกำหนดให้ข้อมูลของราคาข้าวสารเจ้าชนิด 100% เป็นตัวแปรตาม โดยคิดว่าราคาข้าวสารเจ้าชนิด 100% นี้จะเปลี่ยนแปลงขึ้นอยู่กับราคาข้าวเปลือกเจ้าชนิด 100% ราคาข้าวสารเจ้าชนิด 5% และปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ โดยราคาดังกล่าวจะเป็นราคาขาย ณ ตลาดกรุงเทพฯ แล้วนำข้อมูลมาทำให้ขาดหายไปโดยวิธีสุ่ม หลังจากนั้นทำการประมาณค่าโดยวิธีต่างๆ ทั้ง 5 วิธีดังที่ได้กล่าวมาแล้ว พบว่า วิธีถดถอยสองชั้นให้ค่าสัมประสิทธิ์แห่งการตัดสินใจ (R^2) สูงกว่าวิธีอื่นๆ และราคาข้าวสารเจ้าชนิด 100% จะเปลี่ยนแปลงขึ้นอยู่กับราคาข้าวเปลือกเจ้าชนิด 100% ราคาข้าวสารเจ้าชนิด 5% และปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ

บุษบา พิภูลผล (2522) ทำการวิจัยเรื่องการเปรียบเทียบรูปแบบที่ใช้คาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทย โดยใช้การวิเคราะห์อนุกรมเวลากับเทคนิคของบ็อกซ์และเจ็นกินส์ โดยศึกษารูปแบบที่เหมาะสมกับข้อมูลรายเดือนของจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยตั้งแต่ปี พ.ศ. 2506 ถึง พ.ศ. 2520 ซึ่งข้อมูลนี้จะใช้ในการคาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยต่อไป และจากการศึกษาพบว่ารูปแบบที่เหมาะสมที่สุดคือรูปแบบที่อธิบายการเปลี่ยนแปลงเนื่องจากวัฏจักรและเหตุการณ์ผิดปกติร่วมกัน โดยใช้เทคนิคของบ็อกซ์และเจ็นกินส์ และปรับด้วยค่าแนวโน้มของข้อมูล และการเปลี่ยนแปลงเนื่องจากฤดูกาล

วัลลภ โรจนศิริวณิชย์ (2528) ทำการวิจัยเรื่องการพยากรณ์และการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงของราคาหลักทรัพย์ โดยการวิจัยในครั้งนี้ได้เลือกศึกษาหลักทรัพย์ 3 ประเภท คือ ธนาคาร สถาบันการเงินอื่นๆ และกลุ่มพาณิชย์อุตสาหกรรม ที่มีมูลค่าการซื้อขายมากที่สุดตั้งแต่ปี พ.ศ. 2520 ถึง พ.ศ. 2526 และมีการเปลี่ยนแปลงของราคาอย่างต่อเนื่อง โดยศึกษาอนุกรมเวลาราคารายวันและรายสัปดาห์ และจากการศึกษาพบว่าอนุกรมเวลาบ็อกซ์-เจนกินส์ใช้ในการพยากรณ์ได้ดีกว่าเทคนิคการเคลื่อนที่ซ้ำสองครั้งและเทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้ง ยกเว้นในกรณีของข้อมูลราคารายสัปดาห์ของธนาคารกรุงเทพจะใช้เทคนิคการเคลื่อนที่ซ้ำสองครั้งได้ดีที่สุด และในกรณีของบริษัทมหาชนและไซโลจะใช้เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้งใช้ได้ดีที่สุด

พรศิริ หมั่นไชยศรี (2529) ทำการวิจัยเรื่องการเปรียบเทียบวิธีการประมาณค่าสูญหายในการวิเคราะห์ตัวแปรพหุ ซึ่งวิธีการประมาณค่าสูญหายทั้ง 4 วิธี ได้แก่ วิธีที่ใช้ค่าเฉลี่ย วิธีวิเคราะห์ความถดถอยพหุเชิงเส้น วิธีวิเคราะห์ความถดถอยพหุเชิงเส้นดัดแปลง และวิธีวิเคราะห์ส่วนประกอบหลัก โดยใช้ความคลาดเคลื่อนเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบสถานการณ์ต่างๆ ซึ่งจำลองการทดลองขึ้นโดยใช้เทคนิคมอนติคาร์โล แต่ละสถานการณ์จะแตกต่างกันขึ้นอยู่กับขนาดของกลุ่มตัวอย่าง จำนวนตัวแปร และขนาดความสัมพันธ์ระหว่างตัวแปร โดยมีสถานการณ์แตกต่างกันทั้งหมด 106 สถานการณ์ และจากการศึกษาพบว่าถ้ามีข้อมูลสูญหายเกิดขึ้น เราสามารถเลือกวิธีการประมาณค่าสูญหายวิธีใดก็ได้ใน 4 วิธีนี้ แต่วิธีการประมาณค่าสูญหายที่ง่ายที่สุดและใช้เวลาในการประมวลผลน้อยที่สุดคือวิธีค่าเฉลี่ย

สมทรง จิรพัฒน์กุล (2529) ทำการวิจัยเรื่องการศึกษาปัจจัยที่มีผลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ในประเทศไทย โดยกำหนดปัจจัยที่คาดว่าจะมีอิทธิพลต่อการเปลี่ยนแปลงปริมาณเงินฝากของธนาคารพาณิชย์ขึ้นทั้งหมด 5 ปัจจัย ได้แก่ มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น ปริมาณเงินหมุนเวียน อัตราดอกเบี้ยเงินฝาก จำนวนสาขาของธนาคารพาณิชย์ในประเทศไทย และปริมาณเงินให้กู้ยืม โดยกำหนดให้เป็นตัวแปรอิสระ และให้เงินฝากของธนาคารพาณิชย์เป็นตัวแปรตาม และเมื่อนำไปวิเคราะห์ตามทฤษฎีการถดถอยและสหสัมพันธ์เชิงซ้อน พบว่ามีตัวแปรอิสระที่ถูกเลือกเข้ามาในสมการถดถอยเพียงตัวเดียว คือ ปริมาณเงินให้กู้ยืม แสดงว่าปริมาณเงินให้กู้ยืมเป็นปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ โดยมีความสัมพันธ์ไปในทิศทางเดียวกัน

เกศินี กมลรัตน์ (2530) ทำการวิจัยเรื่องการศึกษาเปรียบเทียบเทคนิคการพยากรณ์ที่เหมาะสมกับลักษณะข้อมูล โดยข้อมูลที่น่ามาใช้ในการวิจัยครั้งนี้คือข้อมูลเงินออกรักษาเข้า ซึ่งเทคนิคที่น่ามาใช้มีทั้งสิ้น 5 วิธี ได้แก่ เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลครั้งเดียว เทคนิคการทำให้เรียบแบบเอกซ์โปเนนเชียลซ้ำสองครั้ง เทคนิคการพยากรณ์แบบการกรองแบบปรับได้ เทคนิคการวิเคราะห์อนุกรมเวลาแบบคลาสสิก และเทคนิคการวิเคราะห์อนุกรมเวลาบ็อกซ์และเจ็นกินส์ จากการศึกษาพบว่า ในกรณีที่ข้อมูลมีการเปลี่ยนแปลงเนื่องจากฤดูกาลมาก มีค่าเปลี่ยนแปลงน้อยหรือมาก มีค่าผิดปกติหรือไม่ก็ตาม และขนาดตัวอย่างเพิ่มขึ้นเป็น 50 , 60 , 70 , 80 , 90 , 100 , 110 และ 120 วิธีการพยากรณ์ที่ดีที่สุดคือการใช้เทคนิคการวิเคราะห์อนุกรมเวลาบ็อกซ์และเจ็นกินส์ ยกเว้นในกรณีที่ข้อมูลมีการเปลี่ยนแปลงเนื่องจากฤดูกาลน้อย

ปราณี รัตน์ (2531) ทำการวิจัยเรื่องการประมาณค่าสัมประสิทธิ์การถดถอยพหุ เมื่อความผิดพลาดมีการแจกแจงแบบเบ้ และมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ โดยในการวิจัยได้ทำการจำลองข้อมูลขนาดตัวอย่างด้วยโปรแกรมคอมพิวเตอร์โดยใช้เทคนิคมอนติคาร์โล ซึ่งการแจกแจงของค่าผิดปกติที่ใช้ในการศึกษาครั้งนี้มี 2 แบบ โดยแบบที่ 1 คือการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติซึ่งจะทำการแจกแจงปกติปลอมปนที่มีสเกลแฟคเตอร์เท่ากับ 3 และ 10 และการแจกแจงที่ซึ่งมีระดับความเป็นอิสระเท่ากับ 4 และ 8 และแบบที่ 2 คือการแจกแจงแบบเบ้ซึ่งจะทำการแจกแจงลอกนอร์มอลที่มี $\mu = 0$ และ $\sigma^2 = 1$, การแจกแจงแกมมา และการแจกแจงไวบูลล์ โดยเปรียบเทียบวิธีกำลังสองน้อยที่สุดกับวิธี M – estimator ซึ่งใช้เกณฑ์ความแกร่งของ Ramsay โดยเกณฑ์ที่ใช้ในการเปรียบเทียบคือค่าเฉลี่ยของค่าสัมพัทธ์ของค่าเฉลี่ยความผิดพลาดกำลังสอง และค่าเฉลี่ยของค่าสัมบูรณ์ของค่าแตกต่างของอัตราส่วนค่าเฉลี่ยความผิดพลาดกำลังสอง ผลการวิจัยพบว่าวิธี M – estimator ที่ใช้เกณฑ์ของ Ramsay ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุได้ดีในกรณีที่ค่าผิดพลาดมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ และวิธีกำลังสองน้อยที่สุดกับวิธี M – estimator ซึ่งใช้เกณฑ์ความแกร่งของ Ramsay ใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุได้ใกล้เคียงกันในกรณีที่ค่าผิดพลาดมีการแจกแจงแบบเบ้

ทรงพันธ์ ชุณหสวัสดิกุล (2532) ทำการวิจัยเรื่องการประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยที่ค่าประมาณสเกลเปลี่ยนไป โดยในการวิจัยได้จำลองเหตุการณ์ต่างๆ ด้วยเครื่องคอมพิวเตอร์โดยใช้เทคนิคมอนติคาร์โล ซึ่งการแจกแจงของค่าผิดพลาดที่ใช้ในการศึกษาครั้งนี้มี 2 แบบ โดยแบบที่ 1 คือการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ ซึ่งจะทำการแจกแจงแบบปกติปลอมปนและการแจกแจงแบบที่ และแบบที่ 2 คือการแจกแจงแบบเบ้ ซึ่งจะทำการแจกแจงแบบ

ลอกนอร์มอล การแจกแจงแบบแกมมา และการแจกแจงแบบไวบูลล์ โดยเกณฑ์ที่ใช้ในการเปรียบเทียบคืออัตราส่วนผลต่างของค่าเฉลี่ยความผิดพลาดกำลังสอง และอัตราส่วนผลต่างความผิดพลาดกำลังสอง ผลการวิจัยพบว่า วิธี M – estimator ซึ่งใช้ the median absolute deviation โดยใช้เกณฑ์ความแกร่งของ Huber เป็นวิธีการประมาณค่าที่ดีที่สุด ในขณะที่ค่าผิดพลาดมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ และวิธีกำลังสองน้อยที่สุดกับวิธี M – estimator ซึ่งใช้ the standard deviation of location, the median absolute deviation และ the modified biweight A – estimator โดยใช้เกณฑ์ความแกร่งของ Huber เป็นวิธีการประมาณค่าที่ให้ผลและมีประสิทธิภาพแตกต่างกัน

ชุติมา ชัยมุสิก (2533) ทำการวิจัยเรื่องการวิเคราะห์การถดถอยเชิงซ้อนเมื่อข้อมูลของตัวแปรอิสระสูญหาย โดยในการวิจัยได้จำลองสถานการณ์ที่ใช้ในการศึกษาจำนวน 432 สถานการณ์โดยใช้เทคนิคมอนติคาร์โล ซึ่งวิธีที่ใช้ในการประมาณข้อมูลของตัวแปรอิสระที่สูญหายมี 4 วิธีคือ วิธีวิเคราะห์ความถดถอย วิธี Maximum Likelihood วิธีค่าเฉลี่ย และวิธีค่ามัธยฐาน โดยใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของสมการถดถอยของวิธีที่ไม่มีข้อมูลสูญหายเป็นเกณฑ์ในการเปรียบเทียบ ผลการวิจัยพบว่า โดยส่วนใหญ่วิธีค่าเฉลี่ยให้ผลดีที่สุด ยกเว้นในกรณีที่ขนาดตัวอย่างน้อยและจำนวนตัวแปรอิสระมากพบว่าวิธีวิเคราะห์ความถดถอยจะให้ผลดีที่สุด แต่ถ้าตัวอย่างมีขนาดใหญ่และจำนวนตัวแปรอิสระมีน้อย การตัดขาดของข้อมูลสูญหายทั้งจะไม่แสดงผลกระทบต่อผลการวิเคราะห์การถดถอยด้วยวิธีกำลังสองน้อยที่สุด

สมนิตย์ เจียมธีระนาถ (2534) ทำการวิจัยเรื่องการเปรียบเทียบผลการวิเคราะห์พหุคูณจากการเลือกตัวทำนายเข้าสู่สมการแบบไปข้างหน้า ถอยหลัง และขั้นบันได เมื่อความสัมพันธ์ระหว่างตัวแปรทำนายแตกต่างกัน โดยได้จำลองข้อมูลด้วยเทคนิคมอนติคาร์โล โดยกำหนดให้ข้อมูลมีการแจกแจงเป็นแบบปกติเมื่อมีตัวแปรทำนาย 5 ตัว โดยที่ตัวแปรทำนายแต่ละคู่มีความสัมพันธ์กันใน 3 ระดับคือ 0.00 – 0.30, 0.30 – 0.70 และ 0.70 – 1.00 และตัวแปรทำนายแต่ละตัวมีความสัมพันธ์กับตัวแปรเกณฑ์อยู่ระหว่าง 0.30 – 1.00 จากการศึกษาพบว่า ในกรณีที่ความสัมพันธ์ระหว่างตัวแปรทำนายอยู่ในระดับเดียวกัน ค่าเฉลี่ยของ R^2 ที่ได้จาก 3 วิธีจะมีค่าเกือบเท่ากัน แตกต่างกันอย่างไม่มีนัยสำคัญที่ $\alpha = 0.05$ ความแปรปรวนของ R^2 มีค่าต่ำ และใกล้เคียงกัน และในกรณีที่ความสัมพันธ์ระหว่างตัวแปรทำนายอยู่ในระดับต่างกัน ค่าเฉลี่ยของ R^2 ที่ได้จากวิธีคัดตัวแปรวิธีเดียวกัน มีค่าแตกต่างกันอย่างมีนัยสำคัญที่ $\alpha = 0.05$ ความแปรปรวนของ R^2 มีค่าต่ำ และใกล้เคียงกัน นอกจากนี้ตัวแปรทำนายที่ถูกคัดเข้าสู่สมการในอันดับต่างๆ จากวิธีคัดตัวแปรแบบไปข้างหน้า (Forward Selection)

และวิธีถดถอยแบบขั้นบันได (Stepwise Regression) จะคล้ายคลึงกันและมีร้อยละในการเข้าสมการในแต่ละอันดับใกล้เคียงกันเมื่อตัวแปรทำนายมีความสัมพันธ์ในระดับเดียวกัน

ธิดารัตน์ จันทร์วี (2539) ทำการวิจัยเรื่องการพยากรณ์ความต้องการใช้ไฟฟ้าเพื่อการผลิต ไฟฟ้าระยะสั้น โดยใช้ข้อมูลทุติยภูมิจากแผนกวางแผนการผลิตไฟฟ้าระยะสั้น ซึ่งมีลักษณะข้อมูลเป็นอนุกรมเวลารายเดือน สำหรับเทคนิคการพยากรณ์ที่ใช้ในการวิจัยครั้งนี้มีทั้งสิ้น 3 วิธี ได้แก่ วิธีการพยากรณ์ของบ็อกซ์และเจนกินส์ เทคนิคการปรับให้เรียบ และวิธีพยากรณ์ของแผนกวางแผนการผลิตไฟฟ้าระยะสั้น ซึ่งพิจารณาประกอบกับค่าเฉลี่ยต่ำสุดของเปอร์เซ็นต์ค่าความคลาดเคลื่อนสมบูรณ์ จากการศึกษพบว่าวิธีการพยากรณ์โดยวิธีของบ็อกซ์และเจนกินส์เป็นวิธีที่เหมาะสมที่สุดในการพยากรณ์ความต้องการไฟฟ้าเพื่อการผลิตไฟฟ้าระยะสั้น

บังอร กุมพล (2539) ทำการวิจัยเรื่องการวิเคราะห์การถดถอยเชิงเส้นพหุ เมื่อตัวแปรตามมีค่าถูกตัดทิ้งทางขวากรณีค่าตัดทิ้งประเภทที่ 1 โดยวิธีการประมาณค่าพารามิเตอร์ที่ใช้ในการวิจัยครั้งนี้ ได้แก่ วิธีการของสมิท (Smith Method) วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation Method) และวิธีโมดิไฟด์แอกชูวเรียล (Modified Actuarial Method) ข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองด้วยเทคนิคมอนติคาร์โลซึ่งขนาดตัวอย่างเป็น 10, 20, 30, 50 และ 100 และสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.1, 0.2, 0.3, 0.4 และ 0.5 ในการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของแต่ละวิธีนั้นจะพิจารณาหาค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ของการประมาณค่าตัวแปรตามเป็นเกณฑ์ในการตัดสินใจจากการศึกษาพบว่าในกรณีที่มีสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.1, 0.2 และ 0.3 ณ ทุกระดับของขนาดตัวอย่างและจำนวนตัวแปรอิสระ วิธีการของสมิทจะให้ค่า RMSE ต่ำกว่าวิธีอื่น ในกรณีที่สัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.1 โดยที่ขนาดตัวอย่างเท่ากับ 10 และจำนวนตัวแปรอิสระเท่ากับ 5 วิธีการโมดิไฟด์แอกชูวเรียลจะให้ค่า RMSE ต่ำกว่าวิธีอื่น ในกรณีที่สัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.2 และ 0.3 โดยที่จำนวนตัวแปรอิสระเท่ากับ 7 วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดจะให้ค่า RMSE ต่ำกว่าวิธีอื่น และในกรณีที่สัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.4 และ 0.5 ณ ทุกระดับของขนาดตัวอย่างและจำนวนตัวแปรอิสระ วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดจะให้ค่า RMSE ต่ำกว่าวิธีอื่น

บำเพ็ญ ปิณฑิต (2540) ทำการวิจัยเรื่องการประยุกต์ใช้วิธีวิเคราะห์อนุกรมเวลาของบ็อกซ์และเจนกินส์ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีและไม่มี การเปลี่ยนแปลงเนื่อง จากฤดูกาล โดยทำการพยากรณ์ 5 ช่วงเวลาล่วงหน้าและตรวจสอบผลการพยากรณ์กับผลที่ได้จากวิธี

การวิเคราะห์การถดถอย วิธีการเฉลี่ยเคลื่อนที่ และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล โดยใช้การวัดความคลาดเคลื่อน 6 แบบ ได้แก่ RMSE , MAPE , GMRAE , MdAPE , MdRAE และ Percent Better เป็นเกณฑ์ ซึ่งฐานข้อมูลที่ใช้ในการวิจัยครั้งนี้มีทั้งสิ้น 2 ฐาน ได้แก่ ปริมาณการยืมสิ่งพิมพ์จากศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์จุฬาลงกรณ์มหาวิทยาลัย และจำนวนนักเรียนระดับประถมศึกษา จากการศึกษาพบว่าผลการพยากรณ์ข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ด้วยวิธีการวิเคราะห์การถดถอย และผลการพยากรณ์ข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ด้วยวิธีบ็อกซ์และเจนกินส์มีขนาดความคลาดเคลื่อนน้อยที่สุด

เอกภพ ยานะวิมุติ (2543) ทำการวิจัยเรื่องการเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่ไม่คงที่ระหว่างเทคนิคบ็อกซ์และเจนกินส์ที่ใช้โมเดลสมการเชิงโครงสร้างและตัวบ่งชี้ นำ โดยทำการพยากรณ์ 5 ช่วงเวลาล่วงหน้าและตรวจสอบผลการพยากรณ์ด้วยค่าวัดความคลาดเคลื่อน 6 แบบ ได้แก่ RMSE , MAPE , GMRAE , MdAPE , MdRAE และ Percent Better เป็นเกณฑ์ ซึ่งฐานข้อมูลที่ใช้ในการศึกษาค้างนี้ ได้แก่ ฐานข้อมูลของบำเพ็ญพิตชิต (2540) และเก็บรวบรวมเพิ่มเติมจากบันทึกสถิติ เป็นชุดของข้อมูลอนุกรมเวลารายเดือน 4 ชุดที่มีลักษณะไม่คงที่ ได้แก่ ปริมาณการยืมหนังสือทั่วไป หนังสือสำรอง วิทยานิพนธ์ และผู้เข้าใช้บริการ จำนวน 90, 48, 90 และ 90 ช่วงเวลา ตามลำดับ จากการศึกษาพบว่าผลการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการยืมหนังสือทั่วไป และหนังสือสำรองด้วยวิธีบ็อกซ์และเจนกินส์ และผลการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการยืมวิทยานิพนธ์ด้วยวิธีบ็อกซ์และเจนกินส์ที่ใช้โมเดลสมการเชิงโครงสร้าง มีขนาดความคลาดเคลื่อนน้อยที่สุด

Lakshman K. (1989) ทำการวิจัยในรูปแบบของกรณีศึกษาเรื่องการวิเคราะห์อินเตอร์เวนชันโดยใช้อนุกรมควบคุมและตัวแปรภายนอกในโมเดลทรานสเฟอร์ฟังก์ชัน (Intervention analysis using control series and exogenous variables in a transfer function model : A case study) โดยในการวิจัยนี้เป็นกรณีศึกษาที่ประยุกต์ใช้กับการวิจัยเชิงทดลอง โดยมีการนำระบบ split cable - TV มาใช้กับครอบครัวที่เป็นกลุ่มควบคุมและกลุ่มทดลอง ซึ่งความถี่ของการใช้ระบบนี้กับครอบครัวทั้ง 2 กลุ่มมีความแตกต่างกัน จากการศึกษาพบว่าเกิดอินเตอร์เวนชัน (intervention) กับครอบครัวที่เป็นกลุ่มทดลอง ดังนั้นการวิจัยนี้จึงนำวิธีอริมาอินเตอร์เวนชัน (ARIMA Intervention Analysis) และวิธีทรานสเฟอร์ฟังก์ชัน (Transfer Function Analysis) มาใช้ ผลการวิจัยพบว่ากลุ่มควบคุมและตัวแปรภายนอกสามารถปรับเข้าหากันได้

Connie J. Callahan (1998) ทำการวิจัยเรื่องวิกฤตการณ์ของโมเดลอินเตอร์เวนชันสำหรับครู (Crisis intervention model for teachers) จากการศึกษาพบว่าในแต่ละปีของประเทศสหรัฐอเมริกาในบริเวณโรงเรียนจะเกิดอาชญากรรมประมาณ 3 ล้านครั้ง และในแต่ละวันมีจำนวนนักเรียน 160,000 คนที่โดดเรียน ซึ่งเป็นสิ่งที่น่าวิตกกังวลอย่างมาก ดังนั้นผู้ที่แก้ปัญหาลำดับนี้คือครู ซึ่งครูเป็นบุคคลที่มีความสำคัญที่จะให้ความรู้ ให้ความช่วยเหลือในด้านทักษะ และช่วยแก้ปัญหาให้กับนักเรียน

Judy A. Temple และคณะ (1998) ทำการวิจัยเรื่องการป้องกันการออกจากโรงเรียนกลางคันจากการจัดตั้งศูนย์พ่อแม่และเด็กในมลรัฐชิคาโก (Can Early Intervention Prevent High School Dropout? Evidence from the Chicago Child-Parent Centers) โดยในมลรัฐชิคาโกมีการเสนอกองทุนรัฐบาลให้กับเด็กอายุ 3 – 9 ปี จึงทำให้เกิดข้อมูลที่มีลักษณะเป็นอินเตอร์เวนชันขึ้น และพบว่าเมื่อเด็กเหล่านี้อายุ 17 ปี จะมีการถูกไล่ออกกลางคัน ดังนั้นจากผลการวิจัยพบว่าแนวทางในการแก้ปัญหาคือควรมีการจัดตั้งศูนย์พ่อแม่และเด็ก (Child – Parent Centers) ซึ่งพบว่าช่วยลดปัญหาการออกจากโรงเรียนกลางคันได้ประมาณ 7 – 8%

Shi-Ruei S. Fang และคณะ (1998) ทำการวิจัยเรื่องผลกระทบเกี่ยวกับโปรแกรมทางการศึกษาของผู้ปกครองที่มีต่อความเปลี่ยนแปลงในตัวพ่อและแม่ (The Impact of a Parent Education Program during Mothers' and Fathers' Transition to Parenthood) ซึ่งเป็นการประเมินผลกระทบของอินเตอร์เวนชันในด้านการศึกษาของพ่อแม่ โดยใช้โปรแกรมการเรียนรู้ของครอบครัวในเด็กวัยทารกและทอเดเดอร์ (Family Infant and Toddler Learning Program หรือ FITL) กับชนชั้นกลาง 76 ครอบครัวที่พ่อแม่มีการศึกษาดีและมีบุตรที่มีอายุเฉลี่ยประมาณ 16.9 เดือน ซึ่งโปรแกรมนี้จะเน้นในด้านการสาริตและปฏิบัติ จากการศึกษาพบว่าการเปลี่ยนแปลงในตัวพ่อมีการเพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติ แต่การเปลี่ยนแปลงในตัวแม่มีการลดลงอย่างมีนัยสำคัญทางสถิติ

Janice M. Swanson และคณะ (1999) ทำการวิจัยเรื่องอิทธิพลของจิตวิทยาทางการศึกษาที่ทำให้เกิดอินเตอร์เวนชันต่อความเสี่ยงทางเพศและการปรับเปลี่ยนของจิตวิทยาทางสังคมในวัยรุ่นเกี่ยวกับโรคทางเพศสัมพันธ์ (Effect of psycho-educational interventions on sexual health risks and psycho-social adaptation in young adults with genital herpes) โดยศึกษาจากกลุ่มตัวอย่างจำนวน 252 คนที่สุ่มจากวิธี Randomized Controlled Trial (RCT) พบว่ากลุ่มทดลองซึ่งเป็น

กลุ่ม Psycho-Educational Intervention (PEI) มีระดับนัยสำคัญในเรื่องของความรู้เกี่ยวกับการใช้ถุงยางอนามัยเพื่อป้องกันโรคทางเพศสัมพันธ์มากกว่ากลุ่มควบคุม

Cynthia F. (2000) ทำการวิจัยเรื่องประสิทธิภาพของโมเดลอินเตอร์เวนชันในการเสี่ยงต่อผู้อ่านในโรงเรียนมัธยมศึกษา (An effective (and affordable) intervention model for at-risk high school readers) จากการศึกษาพบว่าการเกิดอินเตอร์เวนชันกับนักเรียนในโรงเรียนในเรื่องของการเรียนรู้ จึงมีการนำโมเดลที่ประกอบด้วย tailor – made มาใช้ในงานวิจัยครั้งนี้ ซึ่งวิธีการแก้ไขมี 3 วิธี ได้แก่ มีการส่งเสริมและประเมินศักยภาพของนักเรียน มีการพัฒนาโปรแกรมที่ใช้สอน และจัดให้มีความเสมอภาคระหว่างครูผู้สอน

John D. และคณะ (2000) ทำการวิจัยเรื่องผลกระทบของโรคที่ติดต่อทางเพศสัมพันธ์ของผู้ชายผิวดำกับพฤติกรรมการใช้ถุงยางอนามัย (Impact of a Culturally Appropriate STD/AIDS Education Intervention on Black Male Adolescents' Sexual and Condom Use Behavior) โดยการวิจัยในครั้งนี้ใช้วิดีโอเทปเป็นสื่อในการพัฒนาส่งเสริมการใช้ถุงยางอนามัยในผู้ชายอเมริกันเชื้อสายแอฟริกันที่มีอายุ 15 – 19 ปี จำนวน 562 คน และมีนักการศึกษาทางด้านสุขภาพมาสาธิตการใช้ถุงยางอนามัย จากการศึกษาพบว่าเกิดอินเตอร์เวนชันกับคนกลุ่มนี้ในด้านพฤติกรรมการใช้ถุงยางอนามัยที่เพิ่มมากขึ้นและมีการเสี่ยงต่อโรคที่ติดต่อทางเพศสัมพันธ์น้อยลง

Jon P. Nelson (2000) ทำการวิจัยเรื่องการวิเคราะห์อนุกรมเวลาอินเตอร์เวนชันที่นำมาประยุกต์ใช้กับการวิเคราะห์การล้มละลายของผู้บริโภค (Consumer Bankruptcies and the Bankruptcy Reform Act : A Time Series Intervention Analysis) โดยการวิจัยในครั้งนี้ใช้ ARIMA intervention analysis ในการพยากรณ์การล้มละลายของผู้บริโภคในปี 1978 โดยใช้ข้อมูลรายไตรมาสที่ปรับแล้วในปี 1960 ถึงปี 1995 ซึ่ง 36% ในช่วงเวลาหลังการกระทำ (post-act period) สัมพันธ์กับช่วงเวลาก่อนการกระทำ (pre-act period) จากการศึกษาพบว่าผู้บริโภคมีการล้มละลายเพิ่มขึ้นในปีหนึ่งๆ ประมาณ 72,400 ราย

Scott R. McConnell (2000) ทำการวิจัยเรื่องการประเมินการเกิดอินเตอร์เวนชันและการศึกษาสำหรับเด็กพิเศษ : การวางรากฐานสำหรับโครงการในอนาคต (Assessment in Early Intervention and Early Childhood Special Education : Building on the Past to Project into Our Future) โดยการวิจัยนี้มีการบูรณาการในส่วนของการเกิดอินเตอร์เวนชันกับการศึกษาสำหรับเด็กพิเศษ จากการศึกษาพบว่ามีแนวทางปฏิบัติ 3 แนวทางที่เป็นเป้าหมายของการประเมินการศึกษา

สำหรับเด็กพิเศษ ได้แก่ จัดให้มีการประเมินความก้าวหน้า มีการปรับและขัดเกลารูปแบบการวิจัย และแนวทางการใช้วิธีอินเทอร์เน็ตเวบนั้นจะต้องมีความสัมพันธ์ที่เป็นอันหนึ่งอันเดียวกัน ซึ่งผลจากการนำแนวทางทั้ง 3 แนวทางมาปฏิบัติพบว่าเป็นการเตรียมการวางรากฐานในอนาคตที่ดีสำหรับเด็กแต่ละคน

ตาราง 12 สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์หอนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอริมาอินเทอร์เน็ตเวบนั้น และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
ชะไมพร ธรรมวัฒน์ไพศาล (2522)*** เรื่อง "วิธีการประมาณค่าที่ขาดหายไปในการวิเคราะห์การถดถอย"	ในการวิจัยพบว่าตัวแปรต้นคือ ราคาข้าวเปลือกเจ้าชนิด 100%, ราคาข้าวสารเจ้าชนิด 5% และปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ โดยที่ตัวแปรตามคือ ราคาข้าวสารเจ้าชนิด 100% แล้วนำข้อมูลเหล่านี้มาทำให้ขาดหายไปโดยวิธีสุ่มแล้วประมาณค่าโดยวิธีต่างๆ 5 วิธี ได้แก่ วิธีกำลังสองน้อยที่สุด, วิธีอันดับศูนย์, วิธีอันดับศูนย์ดัดแปลง, วิธีถดถอยอันดับหนึ่ง และวิธีถดถอยสองชั้น ผลการวิจัยพบว่าวิธีถดถอยสองชั้นให้ค่าสัมประสิทธิ์แห่งการตัดสินใจ (R^2) สูงที่สุด ดังนั้นจึงเป็นวิธีที่ใช้ในการประมาณค่าที่ขาดหายไปที่ดีที่สุด
บุษบา พิภพผล (2522)* เรื่อง "การเปรียบเทียบรูปแบบที่ใช้คาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยโดยใช้การวิเคราะห์หอนุกรมเวลากับเทคนิคของบ็อกซ์และเจนกินส์"	ในการวิจัยได้ศึกษารูปแบบที่เหมาะสมกับข้อมูลรายเดือนของจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทย ตั้งแต่ปี พ.ศ. 2506 – 2520 ผลการวิจัยพบว่ารูปแบบที่เหมาะสมที่สุดคือรูปแบบที่อธิบายการเปลี่ยนแปลงเนื่องจากวัฏจักรและเหตุการณ์ผิดปกติร่วมกัน โดยใช้เทคนิคของบ็อกซ์และเจนกินส์และปรับด้วยค่าแนวโน้มของข้อมูลและการเปลี่ยนแปลงเนื่องจากฤดูกาล
วัลลภ โรจนศิริวิณัย (2528)* เรื่อง "การพยากรณ์และการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงราคาของหลักทรัพย์"	ในการวิจัยได้ศึกษานักทรัพย์ 3 ประเภท คือ ธนาคาร สถาบันการเงิน และกลุ่มพาณิชย์อุตสาหกรรม หลักทรัพย์ที่เลือกมาศึกษา ได้แก่ ธนาคารกรุงเทพ จำกัด, ธนาคารกสิกรไทย จำกัด, ธนาคารกรุงศรีอยุธยา จำกัด, บริษัทเงินทุนอุตสาหกรรมแห่งประเทศไทย, บริษัทชลประทานซีเมนต์, บริษัทปูนซีเมนต์ไทย และบริษัทมาบุญครองออปทิซและไซโล ผลการวิจัยพบว่าหอนุกรมเวลาบ็อกซ์และเจนกินส์ใช้ในการพยากรณ์ได้ดีกว่าการเคลื่อนที่ซ้ำสองครั้งและเทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง ยกเว้นข้อมูลราคาขายสดปันของธนาคารกรุงเทพ การเคลื่อนที่ซ้ำสองครั้งใช้ได้ดีที่สุด และบริษัทมาบุญครองออปทิซและไซโล เทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้งใช้ได้ดีที่สุด

ตาราง 12 (ต่อ) สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์หอนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของไมเคิลอริมา อินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
พรศิริ นมื่นไชยศรี (2529)*** เรื่อง "การเปรียบเทียบวิธีการประมาณค่าสูญหายในการวิเคราะห์ตัวแปรพหุ"	ในการวิจัยได้จำลองสถานการณ์ทั้งหมด 106 สถานการณ์ โดยใช้เทคนิคมอนติคาร์โล โดยสมมติว่าในแต่ละสถานการณ์มีข้อมูลสูญหายเกิดขึ้น และเลือกใช้วิธีการประมาณค่าสูญหายจากทั้งหมด 4 วิธี คือ วิธีที่ใช้ค่าเฉลี่ย, วิธีวิเคราะห์ความถดถอยพหุเชิงเส้น, วิธีวิเคราะห์ความถดถอยพหุเชิงเส้นดัดแปลง และวิธีวิเคราะห์ส่วนประกอบหลัก โดยใช้ค่าความคลาดเคลื่อนเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบ ผลการวิจัยพบว่าวิธีการประมาณค่าสูญหายที่ง่ายที่สุดและใช้เวลาในการประมวลผลน้อยที่สุดคือวิธีที่ใช้ค่าเฉลี่ย
สมทรง จิรพัฒน์กุล (2529)**** เรื่อง "การศึกษาปัจจัยที่มีผลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ในประเทศไทย"	ในการวิจัยครั้งนี้มีตัวแปรต้นทั้งหมด 5 ตัว ได้แก่ มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น, ปริมาณเงินหมุนเวียน, อัตราดอกเบี้ยเงินฝาก, จำนวนสาขาของธนาคารพาณิชย์ในประเทศไทย และปริมาณเงินให้กู้ยืม ส่วนตัวแปรอิสระคือเงินฝากของธนาคารพาณิชย์ โดยใช้ข้อมูลที่เกิดขึ้นจริงระหว่างปี พ.ศ. 2507 - 2526 ซึ่งเทคนิคที่นำมาใช้ในการพยากรณ์มี 3 วิธี คือ วิธีการวิเคราะห์การถดถอย, วิธีบ็อกซ์และเจนกินส์ และอนุกรมเวลาแบบคลาสสิก ผลการวิจัยพบว่า วิธีการพยากรณ์ของบ็อกซ์และเจนกินส์เป็นวิธีที่ดีที่สุด เพราะให้ผลที่คลาดเคลื่อนจากความเป็นจริงต่ำสุดคือประมาณ 1%
เกศินี กมลรัตน์ (2530)* เรื่อง "การศึกษาเปรียบเทียบเทคนิคการพยากรณ์ที่เหมาะสมกับลักษณะข้อมูล"	ในการวิจัยได้ทำการศึกษาเพื่อหาว่าข้อมูลลักษณะใดมีความเหมาะสมที่จะใช้กับเทคนิคการพยากรณ์ในลักษณะใด โดยศึกษาจากข้อมูลเงินออกรักษาเข้า ผลการวิจัยพบว่าเมื่อข้อมูลมีการเปลี่ยนแปลงเนื่องจากฤดูกาลมาก มีค่าเปลี่ยนแปลงน้อยหรือมาก และมีค่าผิดปกติหรือไม่ก็ตาม ในช่วงขนาดตัวอย่าง 50, 60, 70, 80, 90, 100, 110 และ 120 การวิเคราะห์หอนุกรมเวลาบ็อกซ์และเจนกินส์จะใช้ในการพยากรณ์ได้ดีที่สุด

ตาราง 12 (ต่อ) สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์หอนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอริมา อินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
<p>ปราณี รัตน์ (2531)*** เรื่อง "การประมาณค่าสัมประสิทธิ์การถดถอยพหุเมื่อความผิดพลาดมีการแจกแจงแบบเบ้และมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ"</p>	<p>ในการวิจัยได้ทำการจำลองข้อมูลขนาดตัวอย่างด้วยเทคนิคมอนติคาร์โล โดยทำการเปรียบเทียบวิธีกำลังสองน้อยที่สุดกับวิธี M - estimator ซึ่งใช้เกณฑ์ความแกร่งของ Ramsay ซึ่งเกณฑ์การเปรียบเทียบที่ใช้คือค่าเฉลี่ยของค่าสัมพัทธ์ของค่าเฉลี่ยความผิดพลาดกำลังสอง และค่าเฉลี่ยของค่าสัมบูรณ์ของค่าแตกต่างของอัตราส่วนค่าเฉลี่ยความผิดพลาดกำลังสอง จากการศึกษาพบว่าวิธีกำลังสองน้อยที่สุดและวิธี M - estimator สามารถใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุได้ใกล้เคียงกันในกรณีที่ค่าผิดพลาดมีการแจกแจงแบบเบ้ และวิธี M - estimator สามารถใช้ในการประมาณค่าสัมประสิทธิ์การถดถอยพหุได้ดีกว่าวิธีกำลังสองน้อยที่สุดในกรณีที่ค่าผิดพลาดมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ</p>
<p>ทรงพันธ์ ชุนหวัดศิริกุล (2532)*** เรื่อง "การประมาณค่าสัมประสิทธิ์การถดถอยพหุโดยที่ค่าประมาณสเกลเปลี่ยนไป"</p>	<p>ในการวิจัยได้จำลองเหตุการณ์ต่างๆ ด้วยเทคนิคมอนติคาร์โล โดยจะเปรียบเทียบวิธีกำลังสองน้อยที่สุดกับวิธี M-estimator ซึ่งใช้ the standard of location , the median absolute deviation และ the modified biweight A-estimator โดยใช้เกณฑ์ความแกร่งของ Huber และใช้เกณฑ์การเปรียบเทียบของอัตราส่วนผลต่างของค่าเฉลี่ยความผิดพลาดกำลังสอง และอัตราส่วนผลต่างความผิดพลาดกำลังสอง จากการศึกษาพบว่าในกรณีที่ค่าผิดพลาดมีการแจกแจงแบบหางยาวกว่าการแจกแจงแบบปกติ จะใช้ the median absolute deviation เป็นตัวประมาณสเกลซึ่งใช้เกณฑ์ความแกร่งของ Huber ซึ่งเป็นวิธีการประมาณค่าที่ดีที่สุด และในกรณีที่ค่าผิดพลาดมีการแจกแจงแบบเบ้ พบว่าแต่ละวิธีจะใช้ประมาณค่าการแจกแจงในแต่ละแบบได้มีประสิทธิภาพแตกต่างกัน</p>
<p>ชุติมา ชัยมุสิก (2533)*** เรื่อง "การวิเคราะห์การถดถอยเชิงซ้อนเมื่อข้อมูลของตัวแปรอิสระสูญหาย"</p>	<p>จากการศึกษาพบว่าวิธีการประมาณข้อมูลสูญหายในการวิเคราะห์การถดถอยเชิงซ้อนทั้ง 4 วิธี คือ วิธีการวิเคราะห์ความถดถอย วิธี Maximum - Likelihood วิธีค่าเฉลี่ย และวิธีมัธยฐาน ให้ผลต่างกันตามสถานการณ์ต่างๆ ซึ่งโดยส่วนใหญ่วิธีค่าเฉลี่ยให้ผลดีที่สุด ยกเว้นเมื่อมีขนาดตัวอย่างน้อยและจำนวนตัวแปรอิสระมากพบว่าวิธีวิเคราะห์การถดถอยจะให้ผลดีที่สุด แต่ถ้าตัวอย่างมีขนาดใหญ่และจำนวนตัวแปรอิสระมีน้อย การตัดขาดของข้อมูลสูญหายทั้งจะไม่มีผลกระทบต่อผลการวิเคราะห์การถดถอยด้วยวิธีกำลังสองน้อยที่สุด</p>

ตาราง 12 (ต่อ) สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์หอนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอริมา อินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
<p>สมนิตย์ เจียมธีระนาถ (2534)*** เรื่อง "การเปรียบเทียบผลการวิเคราะห์หุนุกรมจากการเลือกตัวทำนายเข้าสู่สมการแบบไปข้างหน้า ถอยหลัง และขั้นบันได เมื่อความสัมพันธ์ระหว่างตัวแปรทำนายแตกต่างกัน"</p>	<p>ในการวิจัยได้ทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล โดยกำหนดให้ข้อมูลมีการแจกแจงเป็นแบบปกติ เมื่อมีตัวแปรทำนาย 5 ตัว และตัวแปรทำนายแต่ละตัวมีความสัมพันธ์กับตัวแปรเกณฑ์อยู่ระหว่าง 0.30 – 1.00 ผลการวิจัยพบว่าเมื่อความสัมพันธ์ระหว่างตัวแปรทำนายอยู่ในระดับเดียวกัน ค่าเฉลี่ยของ R^2 ที่ได้จากวิธีการคัดเลือกตัวแปรทำนายเข้าสู่สมการถดถอย 3 วิธี คือ การเลือกตัวแปรแบบไปข้างหน้า การกำจัดตัวแปรแบบถอยหลัง และการถดถอยแบบขั้นบันได มีค่าเกือบเท่ากัน แต่ในกรณีที่เมื่อความสัมพันธ์ระหว่างตัวแปรทำนายอยู่ในระดับต่างกัน ค่าเฉลี่ยของ R^2 ที่ได้จาก 3 วิธีมีค่าแตกต่างกัน</p>
<p>ธิดารัตน์ จันทวี (2539)* เรื่อง "การพยากรณ์ความต้องการใช้ไฟฟ้าเพื่อการผลิตไฟฟ้าระยะสั้น"</p>	<p>ในการวิจัยครั้งนี้เป็นการศึกษาหาวิธีการพยากรณ์ที่เหมาะสมสำหรับการพยากรณ์ข้อมูลปริมาณความต้องการใช้ไฟฟ้า เพื่อวางแผนการผลิตไฟฟ้าระยะสั้น โดยพิจารณาจากค่าเฉลี่ยต่ำสุดของเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์ ซึ่งเทคนิคที่นำมาใช้ประกอบด้วยวิธีการพยากรณ์ของบ็อกซ์และเจนกินส์ เทคนิคการปรับให้เรียบ และวิธีการพยากรณ์ของแผนกวางแผนการผลิตไฟฟ้าระยะสั้น ผลการวิจัยพบว่าวิธีการพยากรณ์ของบ็อกซ์และเจนกินส์มีความเหมาะสมที่สุด</p>
<p>บังอร กุมพล (2539)*** เรื่อง "การวิเคราะห์การถดถอยเชิงเส้นพหุ เมื่อตัวแปรตามมีค่าถูกตัดทิ้งทางขวา กรณีค่าตัดทิ้งประเภทที่ 1"</p>	<p>ในการวิจัยได้ทำการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล โดยวิธีการประมาณค่าพารามิเตอร์ที่ใช้ในการศึกษานี้มี 3 วิธี คือ วิธีการของสมิท วิธีการประมาณด้วยภาวะนำจะเป็นสูงสุด และวิธีการโมดิไฟด์แอกซ์เรียล ซึ่งในการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในแต่ละวิธีนั้นจะพิจารณาค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) ของการประมาณค่าตัวแปรตามเป็นเกณฑ์ในการตัดสินใจ ผลการวิจัยพบว่าค่า RMSE จะแปรผกผันกับขนาดตัวอย่างและจะแปรผันโดยตรงกับสัดส่วนของข้อมูลที่ถูกต้อง</p>
<p>บำเพ็ญ ปัดชิด (2540)* เรื่อง "การประยุกต์ใช้วิธีวิเคราะห์หอนุกรมเวลาบ็อกซ์และเจนกินส์ในการพยากรณ์ข้อมูลหอนุกรมเวลาทางการศึกษาที่มีและไม่มีเปลี่ยนแปลงเนื่องจากฤดูกาล"</p>	<p>ในการวิจัยเป็นการประยุกต์ใช้วิธีของบ็อกซ์และเจนกินส์ในการพยากรณ์ข้อมูลหอนุกรมเวลาที่มีและไม่มีเปลี่ยนแปลงเนื่องจากฤดูกาล และตรวจสอบผลการพยากรณ์กับผลที่ได้จากวิธีการวิเคราะห์การถดถอย วิธีการเฉลี่ยเคลื่อนที่ และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ผลการวิจัยพบว่าผลการพยากรณ์ข้อมูลหอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลด้วยวิธีการวิเคราะห์การถดถอย และผลการพยากรณ์ข้อมูลหอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลด้วยวิธีของบ็อกซ์และเจนกินส์มีขนาดความคลาดเคลื่อนน้อยที่สุด</p>

ตาราง 12 (ต่อ) สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์อนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอธิบายอินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
<p>เอกภพ ยานะวิมุติ (2543)* เรื่อง “การเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่ไม่คงที่ระหว่างเทคนิคบ็อกซ์และเจนกินส์ที่ใช้โมเดลสมการเชิงโครงสร้างและตัวบ่งชี้หน้า”</p>	<p>ในการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อต้องการเปรียบเทียบระหว่างวิธีบ็อกซ์และเจนกินส์ที่ใช้โมเดลสมการเชิงโครงสร้างกับวิธีบ็อกซ์และเจนกินส์ที่ใช้ตัวบ่งชี้หน้า ผลการวิจัยพบว่า ผลการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการยืมหนังสือทั่วไป และหนังสือสำรองด้วยวิธีบ็อกซ์และเจนกินส์ และผลการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการยืมวิทยานิพนธ์ด้วยวิธีบ็อกซ์และเจนกินส์ที่ใช้โมเดลสมการเชิงโครงสร้าง มีขนาดความคลาดเคลื่อนน้อยที่สุด</p>
<p>Lakshman K. และคณะ (1989)** เรื่อง “กรณีศึกษา : การวิเคราะห์อินเตอร์เวนชันโดยใช้อนุกรมควบคุมและตัวแปรภายนอกในโมเดลทรานสเฟอร์ฟังก์ชัน”</p>	<p>ในการวิจัยนี้เป็นกรณีศึกษาที่ประยุกต์ใช้กับการวิจัยเชิงทดลอง โดยมีการนำระบบ split cable – TV มาใช้กับครอบครัวที่เป็นกลุ่มควบคุมและกลุ่มทดลอง ซึ่งความถี่ของการใช้ระบบนี้กับครอบครัวทั้ง 2 กลุ่มมีความแตกต่างกัน จากการศึกษพบว่าเกิดอินเตอร์เวนชัน (intervention) กับครอบครัวที่เป็นกลุ่มทดลอง ดังนั้นการวิจัยนี้จึงนำวิธีอธิบายอินเตอร์เวนชัน (ARIMA Intervention Analysis) และวิธีทรานสเฟอร์ฟังก์ชัน (Transfer Function Analysis) มาใช้ ผลการวิจัยพบว่ากลุ่มควบคุมและตัวแปรภายนอกสามารถปรับเข้าหากันได้</p>
<p>Connie J. Callahan (1998)** เรื่อง “วิกฤตการณ์ของโมเดลอินเตอร์เวนชันสำหรับครู”</p>	<p>จากการสำรวจพบว่าในแต่ละปีของประเทศสหรัฐอเมริกา บริเวณโรงเรียนจะเกิดอาชญากรรมประมาณ 3 ล้านครั้ง ซึ่งวัยรุ่นเป็นเหยื่อของการเกิดอาชญากรรมมากกว่าวัยอื่นๆ และในแต่ละวันมีจำนวนนักเรียนโดดเรียนประมาณ 160,000 คน ผลการวิจัยพบว่าผู้ที่สามารถแก้ปัญหาเหล่านี้ได้ดีที่สุดคือครู เพราะครูสามารถให้ความปลอดภัยและช่วยเหลือนักเรียนในด้านทักษะได้</p>
<p>Judy A. Temple และคณะ (1998)** เรื่อง “การป้องกันการออกจากโรงเรียนกลางคันจากการตั้งศูนย์พ่อแม่และเด็กในมลรัฐชิคาโก”</p>	<p>จากการสำรวจพบว่าในมลรัฐชิคาโกมีเด็กที่ออกจากโรงเรียนกลางคันอายุประมาณ 17 ปี ผลการศึกษพบว่าการจัดตั้งศูนย์พ่อแม่และเด็ก (Child – Parent Centers) ช่วยลดการออกจากโรงเรียนกลางคันของนักเรียนในระดับมัธยมศึกษาได้ประมาณ 7 – 8 %</p>

ตาราง 12 (ต่อ) สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์หอนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอริมา อินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
<p>Shi – Ruei S. Fang และคณะ (1998)**</p> <p>เรื่อง “ผลกระทบเกี่ยวกับโปรแกรมทางการศึกษาของผู้ปกครองที่มีต่อความเปลี่ยนแปลงในตัวเอง”</p>	<p>มีการเกิดอินเตอร์เวนชันกับระดับการศึกษาของพ่อและแม่ ผลการวิจัยพบว่าแนวทางในการแก้ปัญหาคือการนำโปรแกรมการเรียนรู้ในครอบครัววัยทารกและทอเดเตอร์ (Family Infant and Toddler Learning program : FITL) มาใช้กับพ่อและแม่ โดยโปรแกรมนี้นำมาใช้กับครอบครัวที่เป็นชนชั้นกลาง 76 ครอบครัว และมีอายุเฉลี่ยของบุตรประมาณ 16.9 เดือน จากการศึกษาพบว่า การเปลี่ยนแปลงในตัวพ่อมีการเพิ่มขึ้นอย่างมีนัยสำคัญทางสถิติ แต่การเปลี่ยนแปลงในตัวแม่มีการลดลงอย่างมีนัยสำคัญทางสถิติ</p>
<p>Janice M. Swanson และคณะ (1999)**</p> <p>เรื่อง “อิทธิพลของจิตวิทยาทางการศึกษาที่ทำให้เกิดอินเตอร์เวนชันต่อความเสี่ยงทางเพศและการปรับเปลี่ยนของจิตวิทยาทางสังคมในวัยรุ่นเกี่ยวกับโรคทางเพศสัมพันธ์</p>	<p>การวิจัยนี้เป็นการทดสอบผลลัพธ์ของกลุ่ม Psycho – Educational Intervention (PEI) ที่มีความเสี่ยงต่อสุขภาพทางเพศ ได้แก่ ความรู้ พฤติกรรม และภาระของโรค (disease burden) โดยทำการศึกษากับคนจำนวน 252 คน ที่ถูกสุ่มโดยวิธี Randomized Controlled Trial (RCT) ซึ่งจะแบ่งออกเป็นกลุ่มควบคุมและกลุ่มทดลอง ผลการวิจัยพบว่า กลุ่มทดลองมีระดับนัยสำคัญในเรื่องของความรู้เกี่ยวกับการใช้ถุงยางอนามัยมากกว่ากลุ่มควบคุม</p>
<p>Cynthia F. (2000)**</p> <p>เรื่อง “ประสิทธิภาพของโมเดลอินเตอร์เวนชันในการเสี่ยงต่อผู้อ่านในโรงเรียนมัธยมศึกษา”</p>	<p>มีการเกิดอินเตอร์เวนชันกับนักเรียนในโรงเรียนมัธยมศึกษาในเรื่องของการรู้หนังสือ จากการศึกษาพบว่า แนวทางในการแก้ปัญหาคือจะต้องแก้ไขในตัวครูผู้สอน โดยจะต้องให้ครูผู้สอนมีความเสมอภาคและเท่าเทียมกัน</p>
<p>John D. และคณะ (2000)**</p> <p>เรื่อง “ผลกระทบของโรคที่ติดต่อกับเพศสัมพันธ์ของผู้ชายผิวดำกับพฤติกรรมการใช้ถุงยางอนามัย”</p>	<p>การวิจัยนี้ศึกษาจากผู้ชายอเมริกันเชื้อสายแอฟริกันจำนวน 562 คน เกี่ยวกับพฤติกรรมการเสี่ยงของโรคทางเพศสัมพันธ์ วิธีการศึกษาคือให้บุคคลเหล่านี้ดูวิดีโอเทปและมีนักการศึกษาทางด้านสุขภาพมาสาธิตการใช้ถุงยางอนามัย ซึ่งผลการวิจัยพบว่าเกิดอินเตอร์เวนชันกับบุคคลที่มีกลุ่มอายุ 15 – 19 ปี ในด้านพฤติกรรมการใช้ถุงยางอนามัยที่เพิ่มมากขึ้น ซึ่งส่งผลต่อการเสี่ยงต่อโรคทางเพศสัมพันธ์ที่น้อยลง</p>

ตาราง 12 (ต่อ) สรุปรายชื่อผู้วิจัย หัวข้อวิจัย และผลการวิจัยที่ใช้วิธีการวิเคราะห์หอนุกรมเวลาด้วยวิธีของบ็อกซ์และเจนกินส์ วิธีของบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอริมา อินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย

ผู้แต่งและหัวข้อวิจัย	ผลการวิจัย
Jon P. Nelson (2000)** เรื่อง "การวิเคราะห์หอนุกรมอินเตอร์เวนชันที่นำมาประยุกต์ใช้กับการวิเคราะห์การล้มละลายของผู้บริโภค"	ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลรายไตรมาสตั้งแต่ปี ค.ศ. 1960 – 1995 เกี่ยวกับการล้มละลายของผู้บริโภค โดยจะนำข้อมูลเหล่านี้ไปพยากรณ์การล้มละลายของผู้บริโภคในปี ค.ศ. 1978 ซึ่งวิธีที่ใช้ในการพยากรณ์คือ ARIMA Intervention Analysis ผลการวิจัยพบว่าในปี ค.ศ. 1978 มีจำนวนผู้ที่ล้มละลายเพิ่มขึ้นประมาณ 72,400 ราย
Scott R. McConnell (2000)** เรื่อง "การประเมินการเกิดอินเตอร์เวนชันและการศึกษาสำหรับเด็กพิเศษ : การวางรากฐานสำหรับโครงการในอนาคต"	การวิจัยนี้มีการบูรณาการในส่วนของการเกิดอินเตอร์เวนชันกับการศึกษาสำหรับเด็กพิเศษ ซึ่งมี 3 แนวทาง ได้แก่ จัดให้มีการประเมินความก้าวหน้า มีการปรับและขัดเกลารูปแบบการวิจัย และการใช้การวิเคราะห์ด้วยวิธีอินเตอร์เวนชัน ซึ่งผลการวิจัยพบว่าทั้ง 3 แนวทางนี้เป็นการเตรียมการวางรากฐานในอนาคตสำหรับในเด็กแต่ละคน

หมายเหตุ	*	หมายถึง	การวิจัยที่ใช้วิธีของบ็อกซ์และเจนกินส์
	**	หมายถึง	การวิจัยที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน
	***	หมายถึง	การวิจัยที่ใช้วิธีการวิเคราะห์การถดถอย
	****	หมายถึง	การวิจัยที่ใช้วิธีของบ็อกซ์และเจนกินส์และวิธีการวิเคราะห์การถดถอย