บทที่ 2

กรอบแนวคิดทางทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

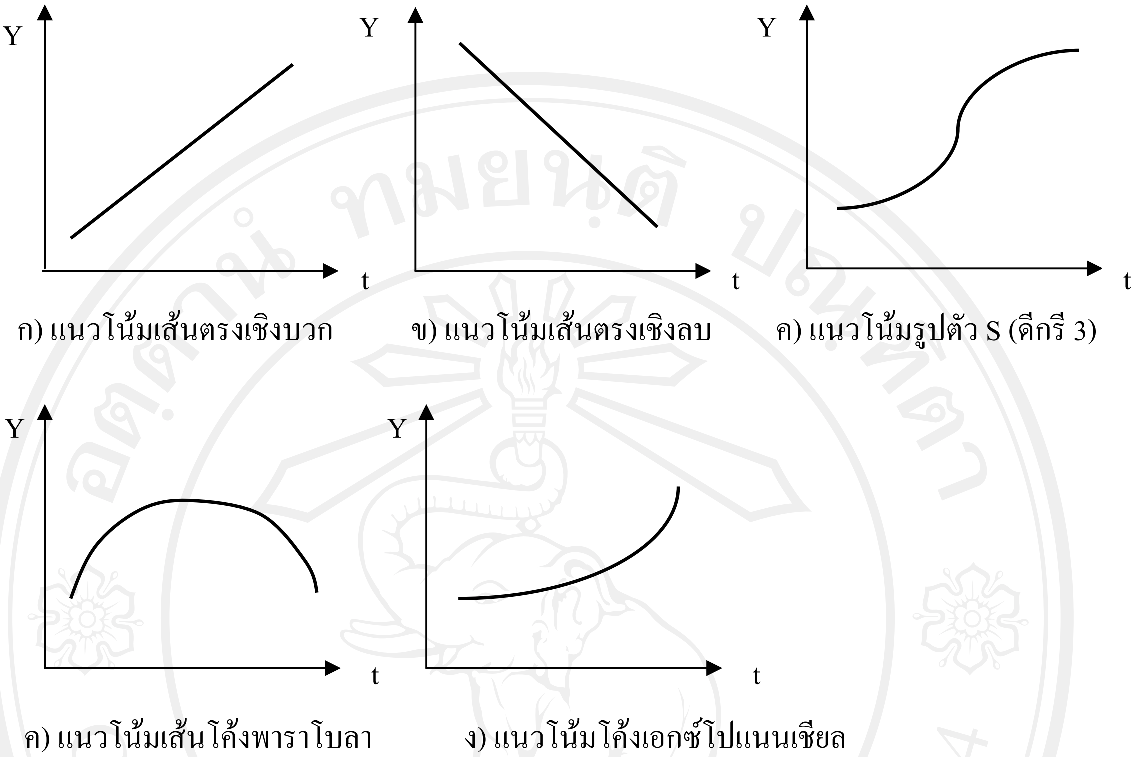
2.1.1 การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

2.1.1.1 องค์ประกอบของอนุกรมเวลา

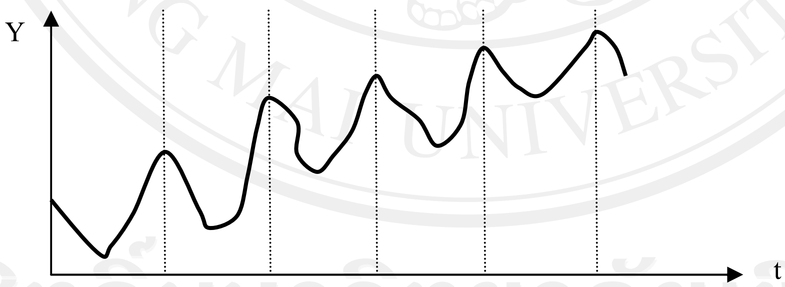
วิธีการพยากรณ์อนุกรมเวลาที่ถือว่า การเปลี่ยนแปลงของอนุกรมเวลาใดๆ อาจจะขึ้นอยู่กับองค์ประกอบ 4 องค์ประกอบ ได้แก่ การเปลี่ยนแปลงอันเนื่องมาจากอิทธิพลของแนวโน้ม (Trend) ฤดูกาล (Seasonal) วัฏจักร (Cyclical) หรือ ความไม่สม่ำเสมอ (Irregular) บางอนุกรมเวลาอาจจะมีองค์ประกอบเพียงอย่างเดียว บางอนุกรมเวลาอาจจะมี 2 หรือ 3 องค์ประกอบ หรือบางอนุกรมเวลาอาจจะมีครบทั้ง 4 องค์ประกอบ เป็นต้น ทั้งนี้ขึ้นอยู่กับลักษณะของอนุกรมเวลานั้นๆ และระยะเวลาที่เก็บรวบรวมข้อมูล ดังรายละเอียดที่จะเสนอต่อไปในการพิจารณาองค์ประกอบของอนุกรมเวลาต่อไป

ก) การเปลี่ยนแปลงแนวโน้ม (Trend Variation) แทนด้วยสัญลักษณ์ **T** ลักษณะของอนุกรมเวลาที่สามารถเห็นการเปลี่ยนแปลงอันเนื่องมาจากแนวโน้มได้นั้น จะต้องเก็บรวบรวมในระยะยาวพอสมควร ซึ่งหน่วยที่เก็บจะมีหน่วยเป็นวัน สัปดาห์ เดือน ไตรมาส หรือปีก็ได้ เมื่อนำอนุกรมเวลาดังกล่าวมาพล็อตกราฟ อาจจะมีการเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงเป็นเส้นตรง (Linear) หรือไม่เป็นเส้นตรง (Non-linear) ก็ได้ ดังรูป 2.1

ข) การเปลี่ยนแปลงฤดูกาล (Seasonal Variation) แทนด้วยสัญญาลักษณ์ **S** อนุกรมเวลาที่สามารถเห็นลักษณะการเปลี่ยนแปลงฤดูกาลได้นั้น ระยะเวลาการเก็บรวบรวมข้อมูลจะเก็บเป็นรายสัปดาห์ รายเดือน หรือรายไตรมาส เมื่อนำอนุกรมเวลาดังกล่าวมาพล็อตกราฟ ก็จะเห็นการเปลี่ยนแปลงอนุกรมเวลาคล้ายกันในช่วงเวลาเดียวกัน ซึ่งการเปลี่ยนแปลงนั้นจะเปลี่ยนแปลงในระยะเวลาสั้นๆ และสามารถคำนวณค่าการเปลี่ยนแปลงที่เปรียบเทียบ แต่ละช่วงเวลาออกมาเป็นตัวเลขได้ เรียกว่า ดัชนีฤดูกาล (Seasonal Index) ดังรูป 2.2



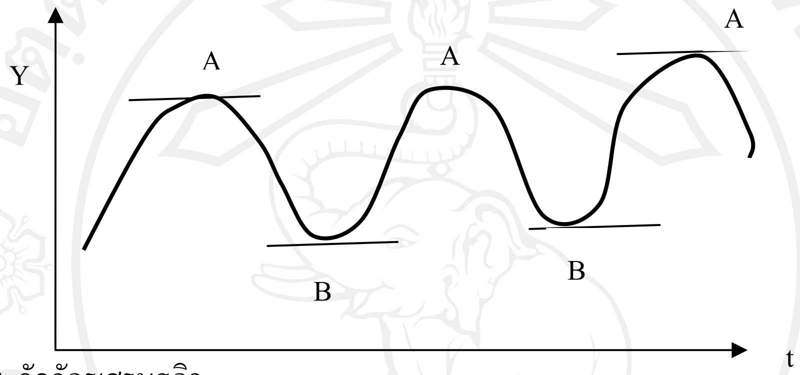
รูป 2.1 ลักษณะแนวโน้มของอนุกรมเวลา



รูป 2.2 ลักษณะอนุกรมเวลาที่มีฤดูกาล

จากรูปที่ 2.2 จะเห็นการเคลื่อนไหวแต่ละช่วงของปีที่ 1 มีการเคลื่อนไหวซ้ำกันหรือคล้ายกันในช่วงเดียวกันของ ปีที่ 2 ส่วนปัจจัยที่ก่อให้เกิดการเปลี่ยนแปลงฤดูกาล ได้แก่ อุณหภูมิ สภาพภูมิอากาศ วัฒนธรรม เทศกาลต่างๆ อนุกรมเวลาที่มีการเปลี่ยนแปลงของฤดูกาล เช่น ยอดจำหน่ายเครื่องปรับอากาศของบริษัทเครื่องปรับอากาศแห่งหนึ่ง ยอดจำหน่ายเสื้อผ้ากันหนาว ราคาสินค้าเกษตรในแต่ละชนิดของแต่ละเดือน ปริมาณการใช้ไฟฟ้าของประชาชนในประเทศในแต่ละเดือน ยอดจำหน่ายเครื่องดื่มและแอลกอฮอล์ในแต่ละเดือนของบริษัท เป็นต้น

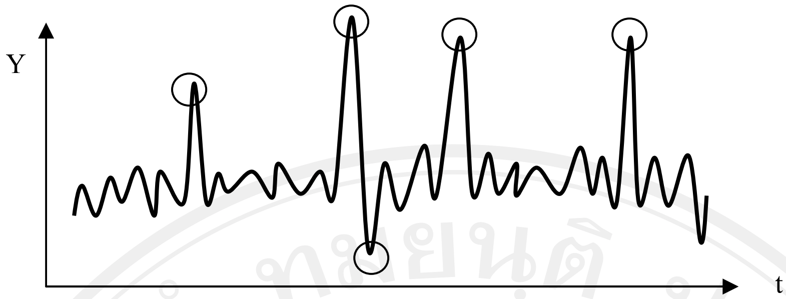
ค) การเปลี่ยนแปลงวัฏจักร (Cyclical Variation) แทนด้วยสัญลักษณ์ **C** ลักษณะของที่สามารถเห็นการเปลี่ยนแปลงของวัฏจักรได้ จะต้องมีระยะเวลาเก็บในระยะยาวเป็นเวลาหลายสิบปี ลักษณะการเคลื่อนไหวของอนุกรมเวลา จะคล้ายกันกับการเปลี่ยนแปลงฤดูกาลต่างกันที่หน่วยของอนุกรมเวลามักจะเก็บเป็นรายปี เพื่อเห็นวัฏจักรที่ชัดเจนและการเปลี่ยนแปลงอาจจะมีช่วงเวลาและขนาดของอนุกรมไม่เท่ากัน เช่น การดำเนินกิจการบางอย่างมีช่วงเวลาการเฟื่องฟู (Prosperity) ระยะตกต่ำ (Depression) การถดถอย (Recession) หรือ การฟื้นตัว (Recovery) ที่แตกต่างกัน กิจการบางกิจการใช้เวลาฟื้นตัวมาก บางกิจการก็ใช้เวลาฟื้นตัวน้อย เป็นต้น



รูป 2.3 วัฏจักรเศรษฐกิจ

จากรูปที่ 2.3 จะพบว่าวัฏจักรเศรษฐกิจของกิจการใดกิจการหนึ่งจะเกิดระยะเฟื่องฟู (Prosperity) ที่จุด A แล้วก็จะลดต่ำลงที่เรียกว่าระยะถดถอย (Recession) ตั้งแต่จุด A มาถึงจุด B เมื่อถึงจุด B จะเรียกระยะต่ำสุด (Depression) แล้วจะเริ่มต้นใหม่มีลักษณะคล้ายๆเดิม ซึ่งวงจรแต่ละรอบอาจจะไม่เท่ากันก็ได้ เช่น ในรอบแรกอาจจะใช้เวลาในการฟื้นตัว 5 ปี ส่วนรอบสองอาจจะใช้เวลา 3 ปี เป็นต้น

ง) การเปลี่ยนแปลงที่ไม่สม่ำเสมอ (Irregular Variation) แทนด้วยสัญลักษณ์ **I** เป็นการเปลี่ยนแปลงของอนุกรมเวลาที่ไม่มีแบบแผนที่แน่นอน ทำให้เป็นอุปสรรคต่อการพยากรณ์ สาเหตุของการเปลี่ยนแปลงเกิดจากเหตุการณ์ที่ผิดปรกติ เช่น วาตภัย อัคคีภัย อุทกภัย ภาวะสงคราม การนัดหยุดงาน การปฏิวัติ รวมถึงปัจจัยอื่นๆ ที่นองเหนือจากแนวโน้ม ฤดูกาล และวัฏจักร ลักษณะการเปลี่ยนแปลงที่ผิดปรกติอาจจะเกิดได้กับอนุกรมทุกๆ อนุกรม โดยลักษณะการเกิดจะไม่มีรูปแบบที่แน่นอน ดังรูป 2.4



รูป 2.4 การเปลี่ยนแปลงของอนุกรมที่ไม่สม่ำเสมอ (Irregular Variation)

จากรูปที่ 2.4 จะพบว่าข้อมูลบางช่วงมีค่าสูงผิดปกตหรือต่ำผิดปกติ ซึ่งไม่สามารถทำนายรูปแบบได้ว่าจะเกิดขึ้นเมื่อไรหรือเกิดขึ้นอย่างไร องค์ประกอบของอนุกรมเวลานี้จะเป็นอุปสรรคของการพยากรณ์ หรือเป็นส่วยที่ทำให้การพยากรณ์มีความคลาดเคลื่อนสูงนั่นเอง

2.1.1.2 รูปแบบของอนุกรมเวลา

ตามทฤษฎีการพยากรณ์อนุกรมเวลาแบบคลาสสิก เมื่อพิจารณาว่าอนุกรมเวลามีองค์ประกอบครบทั้ง 4 องค์ประกอบแล้ว องค์ประกอบทั้ง 4 องค์ประกอบ อาจจะรวมตัวกันได้หลายรูปแบบ เช่น

รูปแบบเชิงบวก (Additive Model) $Y_t = T + S + C + I$ (1)

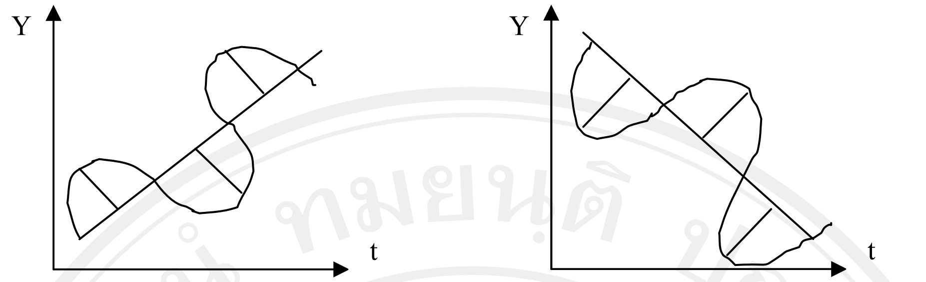
รูปแบบเชิงคูณ (Multiplicative Model $Y_t = T \times S \times C \times I$ (2)

รูปแบบผสม (Mixed Model) $Y_t = (T \times S) + (C \times I)$ (3)

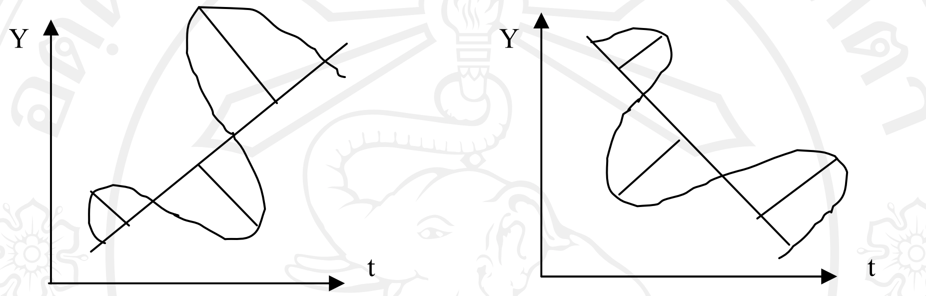
$Y_t = (T + S) \times (C + I)$ (4)

หรือ รูปแบบอื่น ๆ

การวิเคราะห์อนุกรมเวลาเพื่อพิจารณาว่าองค์ประกอบของอนุกรมอยู่ในรูปแบบใดนั้นวิธีการแบบคลาสสิกนี้ยังไม่สามารถสรุปได้แน่นอน เพียงใช้ข้อสังเกตจากรูปกราฟว่า ถ้าเส้นอนุกรมเวลาแกว่งออกจากเส้นแนวโน้มในเดือนเดียวกันแต่แตกต่างกันคงที่หรือไม่แตกต่างกันอนุกรมเวลาจะมีรูปแบบบวก ดังรูปที่ 2.5 ก ถึงรูปที่ 2.5 ข และถ้าเส้นอนุกรมเวลาแกว่งออกจากเส้นแนวโน้มในเดือนเดียวกันแต่ต่างปีกันไม่คงที่หรือแตกต่างกันใยลักษณะเพิ่มขึ้นหรือลดลงอนุกรมเวลานั้นจะมีรูปแบบคูณ ดังรูปที่ 2.5 ค – ง



ก) รูปแนวโน้มเพิ่มขึ้นฤดูกาลคงที่ ข) รูปแนวโน้มลดลงฤดูกาลคงที



ค) รูปแนวโน้มเพิ่มขึ้นฤดูกาลเพิ่มขึ้น ง) รูปแนวโน้มลดลงฤดูกาลเพิ่มขึ้น

รูป 2.5 การเคลื่อนไหวของอนุกรมเวลาที่มีรูปแบบเชิงบวกและเชิงคูณ

2.1.1.3 การวิเคราะห์แนวโน้มของอนุกรมเวลา

การวิเคราะห์อนุกรมเวลาที่รวบรวมในช่วงระยะยาวพอสมควร โดยหน่วยที่รวบรวมได้อาจจะมีหน่วยเป็น รายวัน สัปดาห์ เดือน ไตรมาส หรือปี ซึ่งอนุกรมเวลาดังกล่าวอาจจะมีการเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงเป็นเส้นตรง (Linear) หรือไม่เป็นเส้นตรง (Non-linear) เรียกว่า อนุกรมเวลาที่มีแนวโน้มเข้ามาเกี่ยวข้อง สำหรับการวิเคราะห์แนวโน้มของอนุกรมนั้น มีขั้นตอนดังนี้

ขั้นที่ 1: นำอนุกรมเวลามาทดสอบว่ามีแนวโน้มหรือไม่ ซึ่งมีสองวิธี คือ ทดสอบด้วยสถิติทดสอบ และพิจารณาจากกราฟว่ามีแนวโน้มลักษณะใด

ขั้นที่ 2: ถ้าพิจารณาแนวโน้มโดยอาศัยกราฟ ให้เปรียบเทียบกราฟกับฟังก์ชันทางคณิตศาสตร์ ซึ่งอนุกรมเวลาอาจจะมีแนวโน้มเป็นเส้นตรง (Linear Trend) หรือแนวโน้มไม่เป็นเส้นตรง (Non-linear Trend)

ขั้นที่ 3: สร้างสมการแนวโน้มที่เหมาะสมกับลักษณะของอนุกรมเวลานั้นๆ ซึ่งโดยปกติเมื่อสร้างรูปแบบสมการพยากรณ์ที่เหมาะสมได้แล้วจะมีการพยากรณ์ค่าแนวโน้มใน 2 ลักษณะ คือ

1) การพยากรณ์ย้อนหลังในอดีตเพื่อตรวจสอบความแม่นยำของสมการแนวโน้มที่ใช้พยากรณ์อนุกรมเวลา ซึ่งบางครั้งอนุกรมเวลาหนึ่งชุดอาจมีรูปแบบการพยากรณ์มากกว่าหนึ่ง รูปแบบขึ้นไปก็ได้ กล่าวคือ ในทางปฏิบัติบางอนุกรมเวลาอาจจะมีองค์ประกอบแนวโน้มที่ไม่ชัดเจนทำให้ผู้พยากรณ์จำเป็นต้องสร้างสมการพยากรณ์แนวโน้มในหลายรูปแบบ เช่น อาจจะสร้างสมการแนวโน้มเส้นตรงหรือเส้นโค้งอื่นๆ ทั้งนี้ขึ้นอยู่กับความสามารถเชิงคณิตศาสตร์ของผู้พยากรณ์ซึ่งค่าความคลาดเคลื่อนของการพยากรณ์และคุณสมบัติของค่าคลาดเคลื่อนจะช่วยให้ผู้พยากรณ์ตัดสินใจใช้รูปแบบของการพยากรณ์ที่ดีได้ ดังนั้นเพื่อเปรียบเทียบรูปแบการพยากรณ์ที่ดีได้ ดังนั้นเพื่อเปรียบเทียบรูปแบบการพยากรณ์ต่างๆดังกล่าว จึงต้องอาศัยค่าความคลาดเคลื่อนของแต่ละรูปแบบและนำค่าความคลาดเคลื่อนนั้นมาเปรียบเทียบกัน ถ้ารูปแบบใดที่มีความคลาดเคลื่อนมนการพยากรณ์ต่ำก็ควรจะใช้รูปแบบนั้น

2) พยากรณ์ค่าที่จะเกิดขึ้นในอนาคตหรือการพยากรณ์ช่วงเวลาล่วงหน้า (lead time) ในระยะสั้นหรือระยะยาวก็ได้ ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ของการพยากรณ์ และเทคนิคการพยากรณ์ที่เลือกใช้

2.1.1.4 การหาแนวโน้มเชิงเส้นตรงของอนุกรมเวลา

สมการแนวโน้มเชิงเส้นตรง มีวิธีสร้างได้หลายวิธี ได้แก่ 1) วิธีการเลือกจุด (Selected Points Method) 2) วิธีเฉลี่ยทีละครึ่ง(Semi Average Method) 3) วิธีกำลังสองน้อยที่สุด (Least Square Method) วิธีกำลังสองน้อยที่สุดนับเป็นที่นิยมใช้มากว่าวิธีอื่นๆ เพราะมีเหตุผลทางคณิตศาสตร์ที่น่าเชื่อถือมากกว่าวิธีอื่น มีวิธี ดังนี้

ก) การหาสมการพยากรณ์เชิงเส้นตรงด้วยวิธีกำลังสองน้อยที่สุด

การหาฟังก์ชันเชิงเส้นตรงอย่างง่าย ด้วยวิธีกำลังสองน้อยที่สุดอย่างง่าย เป็นวิธีการใช้ตัวแปรตัวหนึ่ง พยากรณ์ตัวแปรอีกตัวแปรหนึ่ง โดยกำหนดตัวแปรตามเป็น Y และตัวแปรอิสระเป็น t ความสัมพันธ์ของตัวแปรทั้งสองนี้สามารถเขียนในรูปตัวแบบของสมการถดถอยเส้นตรง ได้ดังนี้

ให้ Y_t แทน ค่าจริงของอนุกรมเวลา ณ เวลา t ที่คาดว่ามีแนวโน้มเป็นองค์ประกอบ

$\hat{T}_t$ แทน ค่าพยากรณ์ของอนุกรมเวลา ณ เวลา t ในที่จะแทน $\hat{T}_t = \hat{Y}_t$

$$Y_t = \alpha + \beta t + \varepsilon_t \quad (5)$$

โดยมี α เป็นพารามิเตอร์ที่แสดงค่าจุดตัดแกน Y และ β เป็นพารามิเตอร์ที่ค่าความชันของเส้นตรง ซึ่งหมายความว่า ถ้า t เปลี่ยนไป 1 หน่วยค่า Y_t จะเปลี่ยนไป β หน่วย ความคลาดเคลื่อนควรมีคุณสมบัติ ดังนี้

1) ผลรวมของค่าคลาดเคลื่อนเท่ากับศูนย์เสมอ $\left(\sum_{t=1}^{n} \varepsilon_t\right)$

2) ความแปรปรวนของความคลาดเคลื่อน σ_e^2 ตลอดช่วงเวลา t ต้องคงที่เสมอ

3) ค่าความคลาดเคลื่อนที่เวลา i และj ต้องเป็นอิสระกัน

4) ความคลาดเคลื่อนควรมีการแจกแจงแบบปกติ $\left(\varepsilon_t \approx Niid(0, \sigma_\varepsilon^2)\right)$

แต่ในทางปฏิบัติไม่สามารถนำข้อมูลจริงทั้งหมดมาคำนวณได้ ต้องอาศัยข้อมูลที่เป็นตัวแทนบางส่วนมาคำนวณและเรียกค่าสถิติ และค่าที่คำนวณได้จะเป็นค่าที่นำไปประมาณค่าจากประชากรทั้งหมดหรือเรียกค่าพารามิเตอร์ กำหนดให้

a เป็น ค่าประมาณของ α และ b เป็น ค่าประมาณของ β

$\hat{Y}_t$ เป็น ค่าประมาณของ Y_t และ e_t เป็น ค่าประมาณของ ε_t

ดังนั้นสมการตัวอย่างที่ใช้แทนประชากรดังนี้

$$\hat{Y}_t = a + bt \tag{6}$$

ซึ่งจะพยากรณ์แนวโน้มเชิงเส้นตรง คือ $\hat{Y}_t = a + b \times t$ ในการหาค่า a และ b ที่ทำให้ค่าคลาดเคลื่อนของการพยากรณ์ต่ำสุดจากการพิจารณาค่าผลรวมของค่าคลาดเคลื่อน $\left(\sum e_t\right)$ อาจจะมีค่าเป็น 0 ทั้งๆ ที่การพยากรณ์มีค่าคลาดเคลื่อน

ดังนั้นการพิจารณาความคลาดเคลื่อน จึงพิจารณาผลรวมกำลังสองของค่าคลาดเคลื่อน $\left(\sum e_t^2\right)$ เมื่อ $\sum_{t=1}^{n} e_t^2 = \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2$ ถ้าค่าผลรวมกำลังสองของค่าคลาดเคลื่อนมีค่าสูงการพยากรณ์ก็จะมีการคลาดเคลื่อนสูงตาม หรือถ้าค่าผลรวมกำลังสองของค่าคลาดเคลื่อนมีค่าต่ำการพยากรณ์ก็จะมีความคลาดเคลื่อนต่ำด้วยเช่นกัน หลักการผลรวมกำลังสองจองค่าคลาดเคลื่อนน้อยที่สุด (Least Square method) มีวิธีการหาค่า a และ b ดังนี้

ให้

$$Q = \sum_{t=1}^{n} e_t^2 \tag{7}$$

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

$$= \sum_{t=1}^{n} \left(Y_t - a - bt\right)^2$$

$$\frac{\partial Q}{\partial a} = \frac{\partial}{\partial a}\left(\sum_{t=1}^{n}\left(Y_t - a - bt\right)^2\right) = 0 \qquad (8)$$

$$\frac{\partial Q}{\partial b} = \frac{\partial}{\partial b}\left(\sum_{t=1}^{n}\left(Y_t - a - bt\right)^2\right) = 0 \qquad (9)$$

เรียกสมการ (8) และ (9) ว่าสมการปกติ (Normal Equation) และ สามารถแก้สมการหาค่า a และ b ได้ดังนี้

$$b = \frac{\sum_{t=1}^{n}(t \times Y) - (n \times \bar{t} \times \bar{Y})}{\sum_{t=1}^{n} t^2 - (n \times \bar{t}^2)} \qquad (10)$$

$$a = \bar{Y} - b\bar{t} \qquad (11)$$

จากสมการที่ (10) และ (11) เพื่อความสะดวกในการคำนวณจะปรับ $\sum t = 0$

จะได้
$$b = \frac{\sum_{t=1}^{n}\left(t \times Y\right)}{\sum_{t=1}^{n} t^2} \qquad (12)$$

$$a = \bar{Y} \qquad (13)$$

2.1.1.5 แนวโน้มแบบไม่เป็นเส้นตรง (Non-linear Trend)

เมื่อนำอนุกรมเวลามาพล็อตกราฟแล้ว พบว่าแนวโน้มไม่เป็นเส้นตรง ซึ่งอาจจะอยู่ในรูปโพลิโนเมียล เอ็กซ์โปเนนเซียล ลอกการิทึม หรือรูปแบบฟังก์ชันอื่นๆทางคณิตศาสตร์ ได้ดังนี้

ก) รูปแบบแนวโน้มโพลิโนเมียลดีกรี 2 (Parabola Trend)

สำหรับอนุกรมเวลาที่มีแนวโน้มแบบพาราโบลา จะมีรูปแบบพยากรณ์ ดังนี้

$$\hat{Y} = a + bt + ct^2 \qquad (14)$$

การหาค่า a ,b และ c เพื่อเป็นตัวแทนที่ดีของพารามิเตอร์นั้น ในที่นี้จะหาโดยวิธีกำลังสองน้อยที่สุด

ให้ $$Q = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$$

$$Q = \sum_{i=1}^{n} \left(Y_i - a - bt_i - ct_i^2\right)^2$$

ให้ $$\frac{\partial Q}{\partial a} = 0$$

จะได้ $$na + b\sum_{i=1}^{n} t_i + c\sum_{i=1}^{n} t_i^2 = \sum_{i=1}^{n} Y_i \qquad (15)$$

ให้ $$\frac{\partial Q}{\partial b} = 0$$

$$na + b\sum_{i=1}^{n} t_i^2 + c\sum_{i=1}^{n} t_i^3 = \sum_{i=1}^{n} \left(t_i \times Y_i\right) \qquad (16)$$

ให้ $$\frac{\partial Q}{\partial c} = 0$$

จะได้ $$\sum_{i=1}^{n} t_i^2 + b\sum_{i=1}^{n} t_i^3 + \sum_{i=1}^{n} t_i^4 = \sum_{i=1}^{n} \left(t_i^2 \times Y_i\right) \qquad (17)$$

เรียกสมการ (15) ,(16) และ (17) ว่าสมการปกติ (Normal Equation) การหาค่า a ,b และ c จากสมการ (15) ,(16) และ (17) นำมาเขียนในรูปเมตริกซ์ได้ดังนี้

$$\begin{bmatrix} N & \sum t & \sum t^2 \\ \sum t & \sum t^2 & \sum t^3 \\ \sum t^2 & \sum t^3 & \sum t^4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum (t \times Y) \\ \sum (t^2 \times Y) \end{bmatrix}$$

ให้ $$A = \begin{bmatrix} N & \sum t & \sum t^2 \\ \sum t & \sum t^2 & \sum t^3 \\ \sum t^2 & \sum t^3 & \sum t^4 \end{bmatrix} \quad \tilde{X} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$ และ $$c = \begin{bmatrix} \sum Y \\ \sum (t \times Y) \\ \sum (t^2 \times Y) \end{bmatrix}$$

จะได้สมการในรูป $$A\tilde{X} = B \qquad (18)$$

จากสมการ (13) นำ A^{-1} คูณเข้าทั้งสองข้าง

จะได้สมการดังนี้ $$AA^{-1}\tilde{X} = A^{-1}B \qquad (19)$$

นำสมการที่ (19) ไปหาเมตริกซ์ $\tilde{X}$ ตามหลักการแก้สมการระบบเมตริกซ์ และกำหนดให้ $\sum t = 0$ และ $\sum t^3 = 0$ จะได้

$$a = \frac{\left(\sum t^2\right)\sum\left(Y \times t^2\right) - \left(\sum Y\right)\left(\sum t^4\right)}{\left(\sum t^2\right)^2 - n\sum t^2} \qquad (20)$$

$$b = \frac{\sum(t \times Y)}{\sum t^2} \tag{21}$$

$$c = \frac{\sum Y - na}{\sum t^2} \tag{22}$$

ข) รูปแบบแนวโน้มเอ็กซ์โปเนนเชียล (Exponential Trend)

สำหรับอนุกรมเวลาที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล จะมีรูปแบบพยากรณ์ดังนี้

$$\hat{Y} = a \times b^t \tag{23}$$

จากสมการ (14) ทำให้ในรูป log จะได้

$$\log \hat{Y} = \log a + (t \times \log b) \tag{24}$$

การหาค่า log a และ log b นั้นในที่นี้จะหาโดยวิธีกำลังสองน้อยที่สุด ซึ่งมีวิธีหาเช่นเดียวกับ การหาค่า a และ b ในสมการพยากรณ์แนวโน้มเชิงเส้นตรง ซึ่งจะได้

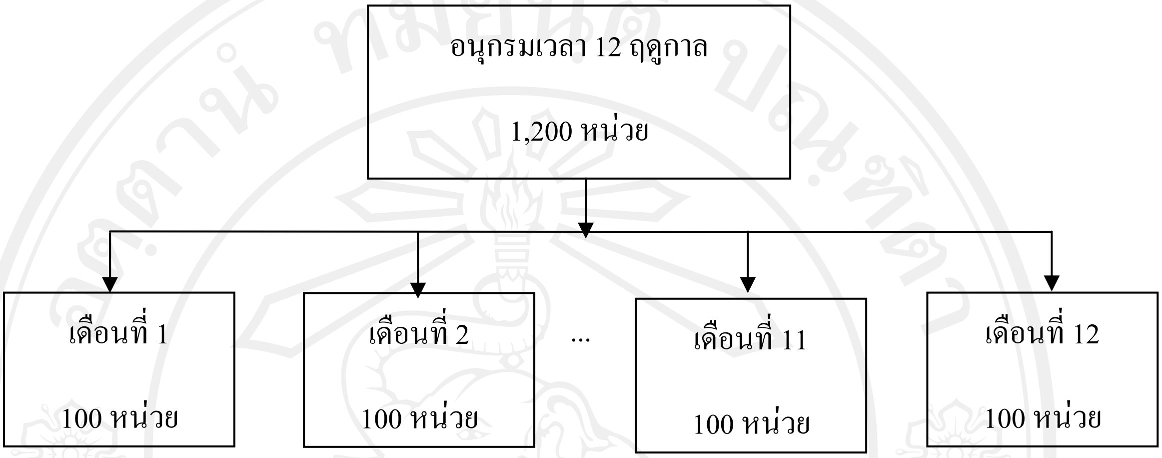
$$\log a = \frac{\sum \log Y}{n} \tag{25}$$

$$\log b = \frac{\sum(t \times \log Y)}{\sum t^2} \tag{26}$$

2.1.1.6 ดัชนีฤดูกาล (Seasonal Index)

การวิเคราะห์หาค่าดัชนีฤดูกาล ทำได้เช่นเดียวกับการหาแนวโน้มของอนุกรมเวลา โดยนำอนุกรมเวลามาพล็อตกราฟ เมื่อพิจารณาลักษณะการเปลี่ยนแปลงของฤดูกาล โดยทั่วไปถ้ารวบรวมอนุกรมเวลาเป็นรายเดือน จะคำนวณหาค่าดัชนีฤดูกาลได้ 12 ค่า หรืออนุกรมเวลานี้มี ฤดูกาล 12 ฤดูกาลนั่นเอง ซึ่งถือว่าแต่ละเดือนคือแต่ละฤดูกาล

ในการหาค่าดัชนีฤดูกาลนั้น จะอยู่ภายใต้ข้อกำหนดที่ว่า โดยเฉลี่ย ในเดือนค่าดัชนีจะเท่ากับ 100 ถ้าไม่มีอิทธิพลจากฤดูกาลเข้ามาเกี่ยวข้อง ดังนั้น 1 ปี ซึ่งมี 12 เดือน ก็จะมีผลรวมทั้งหมดเท่ากับ 1,200 พอดี ดังรูป 2.6



รูป 2.6 ค่าของอนุกรมใดๆที่มีอิทธิพลของฤดูกาล กรณีของข้อมูลรายเดือน

จากรูป 2.6 จะพบว่า ถ้าอนุกรมใดๆ (Y) มี 12 ฤดูกาลและผลรวมของค่า Y ทั้งหมดเท่ากับ 1,200 หน่วย นั่นคือ

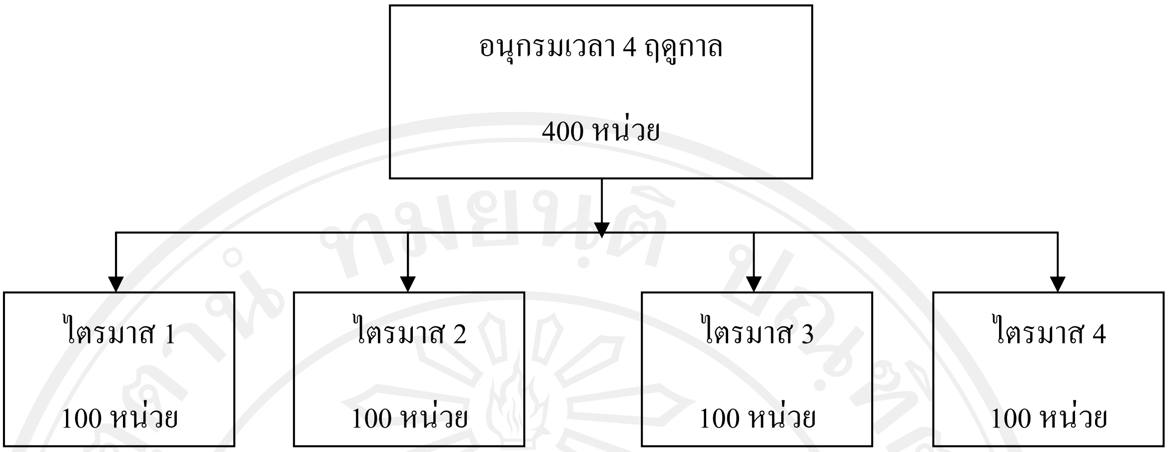
ค่าของ Y ในเดือนที่ 1 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

ค่าของ Y ในเดือนที่ 2 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

: :

ค่าของ Y ในเดือนที่ 12 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

แต่ถ้าดัชนีฤดูกาลในแต่ละเดือนแตกต่างจาก 100 แสดงว่าภายในเดือนนั้นมีอิทธิพลจากฤดูกาลนั่นเอง สำหรับอนุกรมเวลาที่มี 4 ฤดูกาล หรือฤดูกาลหนึ่งมี 3 เดือน ซึ่งเรียกว่ารายไตรมาสก็จะคำนวณดัชนีฤดูกาลได้ 4 ค่า ซึ่งแต่ละฤดูกาลก็จะมีค่าดัชนีเฉลี่ยฤดูกาลละ 100 เช่นกัน และผลรวมทั้งหมดจะเท่ากับ 400 พอดี ดังรูป 2.7



รูป 2.7 ค่าของอนุกรมใดๆที่มีอิทธิพลของฤดูกาล กรณีของข้อมูลรายไตรมาส

จากรูป 2.7 จะพบว่าอนุกรมเวลาใดๆ (Y) มี 4 ฤดูกาล และผลรวมของค่า Y ทั้งหมดเท่ากับ 400 หน่วย นั่นคือ

ค่าของ Y ในไตรมาสที่ 1 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

ค่าของ Y ในไตรมาสที่ 2 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

ค่าของ Y ในไตรมาสที่ 3 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

ค่าของ Y ในไตรมาสที่ 4 จะมีค่าเท่ากับ 100 หน่วย หรือ 100 %

ในการอธิบายความหมายก็เหมือนกันกับอนุกรมเวลาที่มี 12 ฤดูกาลดังที่กล่าวไว้ข้องต้น สำหรับวิธีการคำนวณหาค่าดัชนีฤดูกาลมี 3 วิธีดังนี้

ก) วิธีอัตราส่วนเฉลี่ยอย่างง่าย (Ratio to Simple Average)

การหาค่าดัชนีฤดูกาล โดยวิธีอัตราส่วยเฉลี่ยอย่างง่าย มีขั้นตอนดังนี้

1) คำนวณหาค่าเฉลี่ยเลขคณิตของแต่ละปี

2) คำนวณอัตราส่วยร้อยละของข้อมูลในแต่ละเดือนหรือแต่ละไตรมาส โดยใช้ค่าเฉลี่ยเลขคณิตแต่ละปีที่คำนวณได้หารข้อมูลแต่ละตัวในปีนั้นๆ

3) คำนวณหาค่าเฉลี่ยเลขเลขคณิตของค่าฤดูกาลเฉพาะ โดยเฉลี่ยในแต่ละฤดูกาลของทุกๆปี

4) ถ้าหากผลรวมของค่าดัชนีฤดูกาลในข้อ 3) ไม่เท่ากับ 1,200 เมื่อนุกรมเวลาเป็น รายเดือน หรือไม่เท่ากับ 400 เมื่อนุกรมเวลาเป็นรายไตรมาสจะต้องปรับค่าดัชนีฤดูกาลเฉพาะให้มีผลรวมเท่ากับ 1,200 หรือ 400 ดังกล่าว และดัชนีฤดูกาลที่จะได้ จะเป็นดัชนีฤดูกาลจริง

ข) วิธีอัตราส่วนเฉลี่ยเคลื่อนที่ (Ratio to Moving Average)

การหาค่าดัชนีฤดูกาล โดยวิธีอัตราส่วนเฉลี่ยเคลื่อนที่ มีขั้นตอนดังต่อไปนี้

1) คำนวณค่าเฉลี่ยเคลื่อนที่ 4 ไตรมาส หรือค่าเฉลี่ยเคลื่อนที่ 12 เดือนขึ้นอยู่กับอนุกรมเวลา ถ้าอนุกรมเวลาเป็นรายไตรมาสจะทำให้ปลายทั้งสองหายไปข้างละ 2 ตัว หรือถ้าอนุกรมเวลา เป็นรายเดือนจะทำให้ปลายทั้งสองหายไปข้างละ 6 ตัว

2) จากข้อ 1) จะได้ค่าเฉลี่ยเคลื่อนที่ครั้งที่ 1ซึ่งค่าที่ได้ไม่สามารถนำมาเป็นตัวแทนอนุกรมเวลาในแต่ละไตรมาสหรือรายเดือนได้ จึงต้องเฉลี่ยให้อยู่กึ่งกลางรายไตรมาส หรือรายเดือนจึงถือเป็นค่าแทนรายไตรมาสหรือรายเดือนนั้นๆ แล้วคูณด้วย 100

3) หาค่าเฉลี่ยในแต่ละรายไตรมาสหรือแต่ละเดือน เฉลี่ยตามจำนวนปีของอนุกรมเวลาดัชนีฤดูกาลที่ได้ในข้อ 3) เป็นดัชนีฤดูกาลเฉพาะ ซึ่งต้องปรับให้เป็นผลรวมเป็น 400 เมื่ออนุกรมเวลารายไตรมาส หรือปรับให้ผลรวมเป็น 1,200 เมื่ออนุกรมเวลาเป็นรายเดือน

ค) วิธีอัตราส่วนต่อแนวโน้ม (Ratio to Trend Method)

การหาค่าดัชนีฤดูกาล โดยวิธีอัตราส่วนต่อแนวโน้มนี้ จะเหมาะกับอนุกรมเวลาที่มีองค์ประกอบของแนวโน้มรวมอยู่ด้วย มีขั้นตอนดังต่อไปนี้

1) คำนวณค่าแนวโน้มจากสมการแนวโน้มที่คำนวณได้

2) คำนวณค่าอัตราส่วนร้อยละต่อแนวโน้ม $\left(\frac{Y}{\hat{Y}} \times 100\right)$ เมื่อ $\hat{Y}$ แทนแนวโน้มของอนุกรมเวลา กล่าวคือ ถ้าจากการวิเคราะห์อนุกรมเวลาทราบว่าอนุกรมเวลามีองค์ประกอบของ T, S, C และ I ในเชิงคูณ ในขั้นนี้จะเป็นการขจัดแนวโน้มออกจากอนุกรมเวลา

3) หาค่าเฉลี่ยของอัตราส่วนร้อยละต่อแนวโน้มแต่ละไตรมาสด้วยวิธีเฉลี่ยเลขคณิตหรือมัธยฐานก็ได้ เพื่อกำจัดอิทธิพลเนื่องจากการเปลี่ยนแปลงตามวัฏจักร และการเปลี่ยนแปลงที่ไม่สม่ำเสมอ

4) ค่าเฉลี่ยที่ได้ในข้อ 3) จะถือเป็นค่าดัชนีฤดูกาลที่ยังไม่ปรับผลรวมเป็น 400 เมื่ออนุกรมเวลาเป็นรายไตรมาสหรือผลรวมเป็น 1,200 เมื่อนุกรมเวลาเป็นรายเดือน จึงปรับให้ได้ผลรวมเป็นดังที่กล่าวจึงจะได้ดัชนีฤดูกาลที่แท้จริง

2.1.1.7 การเปลี่ยนแปลงวัฏจักร (Cyclical Variation)

การวิเคราะห์อนุกรมเวลาเพื่อประมาณวัฏจักร หาโดยวิธีขจัดอิทธิพลขององค์ประกอบอื่นๆ ออกไปจากอนุกรมเวลาได้แก่ แนวโน้ม ฤดูกาล และการเปลี่ยนแปลงที่ไม่สม่ำเสมอ ซึ่งการขจัดอาจจะกระทำพร้อมๆกันหรือขจัดทีละองค์ประกอบก็ได้ มีขั้นตอนดังนี้

1) กำจัดแนวโน้มและฤดูกาล ถ้าอนุกรมเวลามีองค์ประกอบดังกล่าว

$$\frac{Y}{\hat{T}} = \frac{T \times C \times I}{\hat{T}} = C \times I \tag{27}$$

หรือ
$$\frac{Y}{\hat{T} \times \hat{S}} = \frac{T \times C \times S \times I}{\hat{T} \times \hat{S}} = C \times I \tag{28}$$

2) ขจัดการเปลี่ยนแปลงที่ไม่สม่ำเสมอออกจากอนุกรมเวลาโดยเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก จะถ่วงแบบไหนขึ้นอยู่กับลักษณะการขึ้นลงของอนุกรมเวลา วิธีที่นิยมใช้ คือ จะใช้น้ำหนักถ่วงทวิ ซึ่งได้จากสัมประสิทธิ์ทวินาม เช่น ต้องการเฉลี่ย 3 คาบเวลา จะใช้ถ่วงน้ำหนักเป็น 1:2:1 และค่าที่ได้จะเป็นค่าการเปลี่ยนแปลงตามวัฏจักร แล้วคูณด้วย 100

2.1.1.8 การเปลี่ยนแปลงที่ไม่สม่ำเสมอ (Irregular Variation)

การวิเคราะห์หาค่าการเปลี่ยนแปลงที่ไม่สม่ำเสมอ สามารถคำนวณได้จากการขจัดองค์ประกอบของอนุกรมเวลาอื่นๆ ออกไป เมื่อขจัดออกไปแล้ว ส่วนที่เหลือ คือ การเปลี่ยนแปลงที่ไม่สม่ำเสมอ ซึ่งในทางปฏิบัติไม่นิยมคำนวณหาค่านี้เพราะไม่มีรูปแบบที่แน่นอน ถ้าหากจะคำนวณค่านี้ก็สามารถทำได้โดย

กรณีที่ 1: ข้อมูลรายปีอนุกรมเวลาประกอบด้วย T, C และ I เมื่อ $Y = T \times C \times I$ มีขั้นตอนการหา I ดังนี้

1) ขจัดค่าแนวโน้ม โดยการหาค่าแนวโน้มแล้วนำไปหารข้อมูลเดิม

$$\frac{Y}{\hat{T}} = \frac{T \times C \times I}{\hat{T}} = C \times I \tag{29}$$

2) ขจัดค่าเปลี่ยนแปลงตามวัฏจักรโดยการหาค่าการเปลี่ยนแปลงตามวัฏจักรแล้วนำไปหารข้อมูลจากข้อ 1)

$$\frac{T \times C \times I}{\hat{T} \times \hat{C}} = I \tag{30}$$

จะได้ข้อมูลที่ปราศจากค่าแนวโน้มและการเปลี่ยนแปลงตามวัฏจักร คงเหลือแต่การเปลี่ยนแปลงเนื่องจากความไม่สม่ำเสมอ

กรณีที่ 2: ข้อมูลรายเดือน หรือรายไตรมาสอนุกรมเวลาประกอบด้วย T,S,C และ I อยู่ในรูป $Y = T \times S \times C \times I$ มีขั้นตอนการหาค่า I ดังนี้

1) ขจัดค่าแนวโน้ม (T)

2) ขจัดค่าเปลี่ยนแปลงตามฤดูกาล (S)

3) ขจัดค่าเปลี่ยนแปลงตามวัฏจักร (C)

จะได้ค่าเปลี่ยนแปลงจากความไม่สม่ำเสมอ $\dfrac{T \times S \times C \times I}{\hat{T} \times \hat{S} \times \hat{C}} = I$

(ลำปาง แสงจันทร์ ,2549)

2.1.2 การทดสอบความนิ่งของข้อมูล (Unit Root Test)

การวิเคราะห์ข้อมูลทางเศรษฐศาสตร์ ซึ่งข้อมูลที่ใช้เป็นข้อมูลแบบอนุกรมเวลา(Time series data) มีความจำเป็นที่จะต้องมีการทดสอบว่า ข้อมูลที่เราใช้นั้นมีความนิ่ง (Stationary) หรือไม่ เพราะหากเราละเลยที่จะทดสอบข้อมูล แล้วข้อมูลเกิดมีความไม่นิ่ง (Non-stationary) จะทำให้สมการถดถอยระหว่างตัวแปรอนุกรมเวลาสองตัวแปร จะได้ค่า R^2 ที่มีค่าสูงมากและค่าสถิติ t จะมีนัยสำคัญ ทั้งที่ตัวแปรทั้งสองนั้นไม่มีความสัมพันธ์กันเลยในทางเศรษฐศาสตร์ (Ender, 1995:216; Gujarati, 1995: 709)

เหตุที่ทำให้ได้ค่า R^2 สูงเช่นนี้เป็นเพราะอนุกรมเวลามีแนวโน้ม ไม่ใช่เนื่องจากความสัมพันธ์ที่แท้จริงระหว่างตัวแปรอนุกรมเวลาทั้งสองตัวแปร ส่วนกรณีที่ค่าสถิติ t มีนัยสำคัญเป็นเพราะอนุกรมเวลาทั้งสองมีแนวโน้มที่แข็งแรงมาก (Strong Trend) โดยสรุปแล้วการละเลยการทดสอบความนิ่งของข้อมูลอาจนำไปสู่การลงความเห็นที่ผิดพลาดได้ในที่สุดข้อมูลที่มีลักษณะนิ่ง (Stationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean)และความแปรปรวน (Variance) เท่ากันตลอดระยะเวลาที่ศึกษาส่วนข้อมูลที่มีลักษณะไม่นิ่ง (Non-stationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย(Mean) และความแปรปรวน (Variance) ไม่เท่ากันตลอดระยะเวลาที่ศึกษาทั้งนี้การวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลาส่วนมากจะพบปัญหาความไม่นิ่งของข้อมูล ซึ่งสามารถแก้ไขได้ด้วยการทำให้ข้อมูลมีความนิ่งเสียก่อน โดยอาจใช้วิธีการหาผลต่าง (Difference) ของข้อมูล การแปลงให้อยู่ในรูป Logarithm หรือการทดสอบหาความสัมพันธ์ของตัวแปรในระยะยาว (Cointegration) เป็นต้น

ในการการทดสอบ Unit root หรือ อันดับความสัมพันธ์ของข้อมูล(Order of Integration) เป็นการทดสอบตัวแปรทางเศรษฐกิจต่างๆ ที่ใช้ในสมการเพื่อดูความเป็น Station (I(0); integrated of order 0) หรือ Non-stationary โดยส่วนมากแล้วจะนิยมการทดสอบโดยวิธี Dickey-Fuller test ซึ่งสามารถแบ่งออกได้เป็น 2 วิธี คือ

วิธีที่ 1 Dickey-Fuller Test (DF)

วิธีนี้จะทำการทดสอบตัวแปรที่เคลื่อนไหวไปตามช่วงเวลาที่มีลักษณะเป็น Autoregressive model โดยพิจารณาสมการ 3 รูปแบบที่แตกต่างกัน ดังนี้

$$\Delta X_t = \theta x_{t-1} + \varepsilon_t \quad \text{(Random walk process)} \tag{31}$$

$$\Delta X_t = \alpha + \theta x_{t-1} + \varepsilon_t \quad \text{(Random walk with drift)} \tag{32}$$

$$\Delta X_t = \alpha + \beta t + \theta x_{t-1} + \varepsilon_t \quad \text{(Random walk with drift and linear time trend)} \tag{33}$$

โดยที่ ΔX_t คือ ค่าความแตกต่างครั้งที่ 1 ของตัวแปรที่ทำการศึกษา

α, β, θ คือ ค่าคงที่

t คือ ค่าคงที่

ε_t คือ ตัวแปรสุ่มที่มีค่าเฉลี่ยเท่ากับศูนย์ และค่าความแปรปรวนที่คงที่ หรือ $\varepsilon_t - iid\left(0, \delta_\varepsilon^2\right)$

การทดสอบ จะพิจารณาค่า θ โดยเปรียบเทียบค่าสถิติ t (t-statistic) ที่คำนวณได้กับค่าที่เหมาะสมจากตาราง Dickey-Fuller ซึ่งมีสมมติฐานการทดสอบ ดังนี้

$H_0 : \theta = 0$: non-stationary

$H_1 : \theta < 0$: stationary

ถ้ายอมรับ $H_0 : \theta = 0$ จะได้ว่า ตัวแปรที่สนใจ (X_t) มี Unit root หรือ X_t มีลักษณะเป็น Non-stationary

แต่ถ้ายอมรับ $H_0 : \theta \neq 0$ จะได้ว่าตัวแปรที่สนใจ (X_t) ไม่มี Unit root หรือ X_t มีลักษณะเป็น Stationary

วิธีที่ 2 Augmented Dickey-Fuller Test (ADF)

เป็นการทดสอบ Unit root อีกวิธีหนึ่งที่พัฒนามาจาก DF Test เนื่องจากวิธี DF ไม่สามารถทำการทดสอบตัวแปรในกรณีที่เป็น Serial correlation ในค่าความคลาดเคลื่อน Error term ε_t ที่มีลักษณะความสัมพันธ์กันเองในระดับสูง โดยมีสมการดังนี้

$$\Delta X_t = \theta x_t + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \varepsilon_t$$ กรณีไม่มีค่าคงที่และแนวโน้ม (34)

$$\Delta X_t = \alpha + \theta x_t + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \varepsilon_t$$ กรณีมีเฉพาะค่าคงที่ (35)

$$\Delta X_t = \alpha + \beta t + \theta x_t + \sum_{j=1}^{p} \phi_j \Delta X_{t-j} + \varepsilon_t$$ กรณีมีทั้งค่าคงที่และแนวโน้มเวลา (36)

การใส่ค่า Lagged Term (ρ) ว่ามีจำนวนเท่าใดจึงจะเหมาะสมสำหรับแต่ละข้อมูลอนุกรมนั้นมีหลักในการเลือก Lag length ที่เสนอโดย Ender (1995) ว่า ควรจะเริ่ม Lag length ที่มีค่าที่มากพอ แล้วพิจารณาความมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญต่างๆ (σ = 0.01 ,0.05 และ 0.1) เมื่อพบว่าที่ Lag length ที่เลือกมีค่า t – statistic ที่ไม่มีนัยสำคัญทางสถิติ ณ ระดับนัยสำคัญร้อยละ 10 แล้วจึงทำการลด Lag length ลงที่ละ 1 ช่วงจนกระทั่งสามารถปฏิเสธสมมติฐานว่าง (Null Hypothesis)

2.1.3 Long Memory Test

หลักการของ Long Memory จะมีความสัมพันธ์ที่ใกล้เคียงในการประมาณค่า mean ของ stationary การใช้ Long Memory Test จะใช้กับข้อมูลที่มีปัญหาสหสัมพันธ์ตัวคลาดเคลื่อน (Autocorrelation) หมายความว่าค่าสังเกตไม่เป็นอิสระต่อกันสูง แต่จะใช้ทดสอบในระยะเวลาที่ไกลกว่ากระบวนการ stationary สมมติให้ว่า y_t เป็นข้อมูลอนุกรมเวลาที่มีความนิ่ง จะมีระบบความจำระยะยาว (long memory) หรือการขึ้นอยู่ต่อกันในระยะยาว (long range dependence) เมื่อมี autocorrelation function ดังนี้

$$\rho(k) \to C_p k^{-\alpha} \quad \text{เมื่อ } k \to \infty \tag{37}$$

โดยที่ C_p คือ ค่าคงที่ที่เป็นค่าบวก

α คือ จำนวนจริงที่มีค่าระหว่าง 0 ถึง 1

ดังนั้น autocorrelation function ของกระบวนการระบบความจำระยะยาว (long memory) จะเสื่อมลง (decay) อย่างช้า ๆ ใน อัตราไฮเพอร์โบลิก (hyperbolic) โดยข้อเท็จจริงแล้วมันเสื่อมลงอย่างช้า ๆ เนื่องจาก autocorrelation นั้นไม่ใช่ผลรวมทั้งหมด ดังที่

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty \tag{38}$$

สำหรับกระบวนการของข้อมูลอนุกรมเวลาที่มีความนิ่งจะมี autocorrelation function ที่ประกอบไปด้วยข้อมูลที่เหมือนกับความหนาแน่นแบบสเปกตรัม (spectral density) ของมัน ซึ่งสมการความหนาแน่นแบบสเปกตรัมสามารถแสดงดังนี้

$$f(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{ikw} \tag{39}$$

โดยที่ w คือ ความถี่ฟูเรียร์ (Fourier frequency) (Hamilton, 1994) จากสมการที่ (37) สามารถแสดงได้ดังนี้

$$f(w) \to C_f w^{\alpha-1} \quad \text{เมื่อ } w \to 0 \tag{40}$$

เมื่อ C_f คือ ค่าคงที่ที่เป็นค่าบวก ดังนั้นกระบวนการของระบบความจำระยะยาว (long memory) มีความหนาแน่นแบบสเปกตรัม (spectral density) ที่มีความโน้มเอียงเข้าสู่ค่าอนันต์ (infinity) ณ ความถี่เท่ากับศูนย์ แทน α โดยใช้

$$H = 1 - \frac{\alpha}{2} \in (0.5, 1) \tag{41}$$

ซึ่งทราบกันว่านี่คือ Hurst coefficient (Hurst, 1951) ที่จะวัดความมีระบบความจำระยะยาว (long memory) ใน y_t ซึ่งค่า H ที่มากแสดงถึงระบบความจำระยะยาว (long memory) ที่ยาวกว่า

จากหลักการของคุณสมบัติของขนาด (scaling property) จากสมการที่ (37) และคุณสมบัติของขอบเขตความถี่ (frequency domain property) จากสมการที่ (40) Granger and Joyeux (1980) สามารถจำลองเชิงพารามิเตอร์โดยรวมกระบวนการปริพันธ์ (integrated process) เป็นกระบวนการหาผลต่างที่เป็นจำนวนเศษส่วน (fractionally integrated process) ซึ่งให้การหาผลต่างที่เป็นจำนวนเศษส่วน (fractionally integration) ในข้อมูลอนุกรมเวลา (time series data) y_t มีลักษณะดังนี้

$$(1-L)^d (y_t - \mu) = u_t \tag{42}$$

โดยที่ L คือ lag operator

d คือ fractional difference parameter

μ คือ ค่าคาดหมายของ y_t

u_t คือ พจน์รบกวนในระบบความจำระยะสั้นที่นิ่งและมีค่าเฉลี่ยเท่ากับศูนย์

โดยปกติแล้วเมื่อข้อมูลอนุกรมเวลานั้นถูกพบว่าไม่นิ่ง (non-stationary) เมื่อทำการหาผลต่าง (difference) หนึ่งครั้ง (d=1) ก็จะนิ่ง (stationary) อย่างไรก็ตามข้อมูลอนุกรมเวลาทางการเงินและเศรษฐกิจบางข้อมูลก็มีคุณสมบัติการคงอยู่ (persistent) ของ autocorrelation function อย่างสูง ซึ่งหมายความว่ากระบวนการหาผลต่าง (difference) ที่เป็นจำนวนเต็มนั้นอาจจะมากเกินไป ซึ่งแสดงให้เห็นจากความหนาแน่นแบบสเปกตรัม (spectral density) หายไป ณ ความถี่เท่ากับศูนย์ของผลต่าง (difference) ของข้อมูลอนุกรมเวลา การจำลองกระบวนการของระบบความจำระยะยาว (long memory) จึงหลีกเลี่ยงที่จะหาผลต่างของ y_t เป็นจำนวนเต็ม (integer) โดยให้ค่า d สามารถเป็นจำนวนเศษส่วนได้ ซึ่ง fractional difference filter สามารถแสดงได้ดังนี้

$$(1-L)^d = \sum_{k=0}^{\infty}\binom{d}{k}(-1)^k L^k \tag{43}$$

โดยสัมประสิทธิ์ทวินาม (binomial coefficient) มีค่าดังนี้

$$\binom{d}{k} = \frac{d!}{k!(d!-k!)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)} \tag{44}$$

เมื่อพิจารณาค่า d สามารถแสดงให้ทราบว่า

เมื่อค่า $|d| > 1/2$ แล้ว y_t นั้นมีคุณสมบัติ non-stationary

$0 < d < 1/2$ แล้ว y_t นั่น stationary และมี long memory

$-1/2 < d < 0$ แล้ว y_t นั้น stationary และเป็น short memory

เมื่อ fractionally integrated ข้อมูลอนุกรม y_t มีคุณสมบัติของระบบความจำระยะยาว (long memory) สามารถแสดงได้ว่า

$$d = H - 1/2 \tag{45}$$

ดังนั้น d และ H สามารถใช้สลับกันได้ Hosking (1981) ได้แสดงว่าคุณสมบัติด้านขนาด (scaling property) จากสมการที่ (37) และคุณสมบัติด้านขอบเขตความถี่ (frequency domain property) จากสมการที่ (40) นั้นสอดคล้องกันเมื่อ 0 < d < ½

2.1.3.1 Test for Long Memory: การทดสอบ R/S

การทดสอบ R/S test ถูกพัฒนาโดย Harold Edwin Hurst ในช่วง 1960 และ Mandelbrot & Wallis(1969) ใช้ในการคำนวณค่า พารามิเตอร์ H ที่ใช้วัด ความหนาแน่นของ long range dependenceในอนุกรมเวลา

อนุกรมเวลาของช่วง T แบ่งออกเป็น n sub-series ของช่วง m และสำหรับทุกๆ sub-series โดยแต่ละ sub-series จะมีค่า m = 1,…,n, to เพื่อหาค่ามัธยฐาน (E_m) และส่วนเบี่ยงเบนมาตรฐาน (S_m) และขจัด ค่ามัธยฐานตัวอย่าง ที่ $Z_{i,m} = X_{i,m} – E_m$ สำหรับ i = 1,…..,m

หลังจากนั้นจึง สร้างอนุกรมเวลาโดยใช้รูปแบบของ $W_{i,m} = \sum_{j=1}^{i} Z_{j,m}$ โดยที่ i = 1,..,m และ เพื่อหา ระยะของ $R_m = \max\{W_{1,m,\ldots,}W_{n,m}\} - \min\{W_{1,m,\ldots,}W_{n,m}\}$

การกำหนด rescaled range ของ R_m โดยใช้ $\frac{Rm}{Sm}$ แบบเดียวกับในกรณีของอนุกรมเวลา สามารถหาค่า R, S และ H ตามข้อกำหนดต่อไปนี้

$$R = k \times T^{0.5} \tag{46}$$

กำหนดให้ R คือ ระยะ ที่อยู่ในตัวแปร , k คือ ค่าคงที่ และ T คือ ช่วงความยาวของเวลา

กำหนดให้ R/S คือ rescaled range, m จำนวนครั้งของกาสำรวจ, k คือ ค่าคงที่ และ H คือ Hurst exponent จะสามารถนำมาใช้ในอนุกรมเวลาขนาดใหญ่ได้

$$\frac{R}{S} = k \times m^{H} \tag{47}$$

ค่า Hurst exponent หาได้จาก :

$$\log(R/S)\text{m} = \log k + \text{H} \log m \tag{48}$$

และมีข้อกำหนดว่า :

ถ้า H value = 0.5 อนุกรมเวลาจะเป็นเคลื่อนไปอย่างสุ่ม และ เป็นอิสระ

ถ้า H value =(0, 0.5) อนุกรมเวลาจะเป็นแบบไม่คงตัว กระบวนการจะคลอบคลุมเพียงแค่วงแคบ เมื่อเทียบกับกรณีการเคลื่อนไปอย่างสุ่ม

ถ้า H value =(0.5, 1) อนุกรมเวลา จะเป็นชุดข้อมูลที่คงตัว กระบวนการจะคลอบคลุมเป็นวงกว้าง เมื่อเทียบกับกรณีการเคลื่อนไปอย่างสุ่ม

2.1.3.2 Test for Long Memory: การทดสอบ Modified R/S

Classical R/S test เสนอโดย Hurst (1951) ใช้ในการทดสอบข้อมูลอนุกรมเวลาว่ามีลักษณะของระบบความจำระยะยาว (long memory) โดยสมมติให้ข้อมูลอนุกรมเวลา $\{x_t\}, t = 1,2,...,N$ โดยที่ $y_j = \sum_{i=1}^{j} x_i,\ j = 1,2,...,N$ และ ความแปรปรวนของตัวอย่าง $S_j^2 = j^{-1}\sum_{i=1}^{j}(x_i - j^{-1}y_j)^2,\ j = 1,2,...,N$ ซึ่งจะได้ R/S-statistic ดังนี้ (Wang, et al. 2549)

$$R/S(j) = 1/S_j\left[\max_{0\le t\le j}(y_t - \frac{t}{j}y_j) - \min_{0\le t\le j}(y_t - \frac{t}{j}y_j)\right] \quad (49)$$

Lo (1991) ได้เสนอ Modified R/S ซึ่งพัฒนามาจาก การทดสอบ เนื่องจาก R/S test มีความอ่อนไหวอย่างมากต่อ short-range dependence โดยแทนที่ S_j ในสมการที่ (23) เป็น S_q (modified standard deviation) ที่คำนึงถึง autocovariances ใน q lags แรก ซึ่งจะช่วยละอิทธิพลของ short-range dependence ที่อยู่ในข้อมูล แทนที่จะพิจารณาหลาย ๆ lag อย่างในสมการที่ (23) ก็พิจารณาเพียง lag j = N โดย S_q มีค่าดังนี้

$$S_q = \left(\frac{1}{N}\sum_{j=1}^{N}(x_j - \bar{x}_N)^2 + \frac{2}{N}\sum_{j=1}^{q}\omega_j(q)\left[\sum_{i=j=1}^{N}(x_i - \bar{x}_N)(x_{i-j} - \bar{x}_N)\right]\right)^{1/2} \quad (50)$$

โดยที่ $\bar{x}_N$ เป็นค่าเฉลี่ยของอนุกรมเวลา

$$\omega_j(q) = 1 - \frac{j}{q+1}, q < N \quad (51)$$

ดังนั้นจะได้ Lo's Modified R/S statistic ดังสมการ

$$Q_{N,q} = \frac{1}{S_q}\left\{\max_{0\le i\le N}\sum_{j=1}^{i}(x_j - \bar{x}_N) - \min_{0\le i\le N}\sum_{j=1}^{i}(x_j - \bar{x}_N)\right\} \quad (52)$$

Lo การมีอยู่ของระบบความจำระยะยาว (long memory) $N^{-1/2}Q_{N,q}$ จะเบนเข้าหาช่วง Brownian Bridge

$$W = \max_{0 \le r \le 1} V(r) - \min_{0 \le r \le 1} V(r) \tag{53}$$

โดย V คือ Standard Brownian bridge

$$V(r) = B(r) - rB(1) \tag{54}$$

โดยที่ B คือ Standard Brownian motion

การกระจายของตัวแปรสุ่ม W รู้จักกันใน

$$P(W \le x) = 1 + 2\sum_{j=1}^{\infty}(1 - 4x^2 j^2)e^{-2x^2 j^2} \tag{55}$$

Lo ใช้ช่วง [0.809,1.862] ที่ 95% (asymptotic) ยอมรับ null hypothesis

H_0 = { ไม่มี long memory, เช่น H = 0.5}

H_1 = { มี long memory, เช่น 0.5< H <1}

ค่า critical อยู่ในช่วงตามตาราง 2.1

ตาราง 2.1 ค่า Critical Value ของ Modified R/S test

ระดับความน่าจะเป็น	Critical Value
0.5%	0.721
2.5%	0.809
5%	0.861
10%	0.927
90%	1.620
95%	1.747
97.5%	1.862
99.5%	2.098

ที่มา : International Journal of Human and Social Sciences (Long Memory Analysis of USD/TRL Exchange Rate)

2.1.3.3 Test for Long Memory: GPH Test

กระบวนการ GPH Test ถูกพัฒนาโดย J. Geweke และ S. Porter-Hudak (1983) เพื่อแสดงถึง การประมาณค่า OLS estimator ของ d จากสมการถดถอย

$$\ln\left[I(\xi)\right] = a - \hat{d}\ln\left[\sin^2\left(\frac{\xi_\lambda}{2}\right)\right] + e_\lambda \qquad , \lambda = 1,2,...,V \tag{56}$$

โดยที่

$$I(\xi) = \frac{1}{2\pi T}\left|\sum_{t=1}^{T} e^{it\xi}(x_t - \bar{x})\right|^2 \tag{57}$$

และสมการ คือ Periodogram (การประมาณค่าความหนาแน่น ของ spectral) ของค่า x ที่ frequency เหมือนกันกับ ค่า bandwidth v ถูกเลือกไว้สำหรับ (ξ) , $T \to \infty$ $v \to \infty$ แต่ $\frac{v}{T} \to 0$ แนวคิดของ Geweke and Porter-Hudak (1983) พิจารณาว่า อิทธิพลของ T จะอยู่ระหว่าง (0.5,0.6) และสมมติฐานหลักของกระบวนการความทรงจำระยะยาว ความชันของสมการถดถอย d เท่ากับศูนย์ และ ค่า t-statistics สามารถใช้ในการแสดงผลการทดสอบได้

2.1.4 แบบจำลอง ARIMA (Autoregressive Integrated Moving Average)

แบบจำลอง Autoregressive Integrated Moving Average (ARIMA) ศึกษาโดย Box และ Jenkins (1976) แต่ Wold (1938) ได้เป็นผู้ให้พื้นฐานทางทฤษฎีของกระบวนการหรือระบบ ARIMA บนพื้นฐานของ Wold แบบจำลอง ARIMA ได้ถูกพัฒนาขึ้นในสามทิศทาง ซึ่งได้แก่ ขั้นตอบสนองการประมาณค่าและการบ่งชี้ที่มีประสิทธิภาพและกระบวนการประมาณ (efficient identification and estimation and estimation procedures) (สำหรับกระบวนการหรือระบบ AR,MA และ ARMA) การคลอบคลุมไปถึงผลลัพธ์ที่ได้รวบรวมเอาอนุกรมเวลาเชิงฤดูกาล (seasonal time series) และการขยายของเขตไปเพื่อรวมเอากระบวนการหรือระบบไม่นิ่ง (non-stationary process (ARIMA)) เข้าไว้ด้วย (ทรงศักดิ์ ศรีบุญจิตต์, 2547)

Autoregressive Process (AR(p)) แสดงให้เห็นว่าข้อมูลอนุกรมเวลาขึ้นอยู่กับค่าของตัวมันเองในอดีต โดย p คือ จำนวนระยะห่าง (lag) ของข้อมูลในอดีตจากปัจจุบัน

Moving Average Process (MA(q)) แสดงให้เห็นว่าข้อมูลอนุกรมเวลาขึ้นอยู่กับความคาดเคลื่อนในปัจจุบันและความคลาดเคลื่อนอดีต โดย q คือ จำนวนของระยะห่าง (lag) ของค่าคลาดเคลื่อนในอดีต

Autoregressive and Moving Average Process (ARMA(p,q)) เป็นการรวมกันระหว่าง AR กับ MA นั่นคือ ข้อมูลอนุกรมเวลาขึ้นอยู่กับค่าของข้อมูลอนุกรมเวลาในอดีต และค่าความคลาดเคลื่อนทั้งในปัจจุบันและในอดีต

การวิเคราะห์ข้อมูลโดยวิธีนี้เป็นวิธีวิเคราะห์ข้อมูลอนุกรมเวลาที่อาศัยขบวนการสโตคาสติก (Stochastic Process) โดยถือว่าข้อมูลที่เกิดขึ้นตามเวลาที่เปลี่ยนแปลงไปมีลักษณะการเกิดที่เป็นไปตามกฎความน่าจะเป็น ข้อมูลที่ใช้จะต้องมีลักษณะที่นิ่ง (stationary) โดยเขียนในรูปสมการดังนี้

$$y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-p} \quad (58)$$

โดยที่ y_t คือ ค่าสังเกตของอนุกรมเวลา ณ เวลาที่ t

δ คือ ค่าคงที่

$\phi_1, \phi_2, \ldots, \phi_p$ คือ พารามิเตอร์ของ Autoregressive

$\theta_1, \theta, \ldots, \theta_q$ คือ พารามิเตอร์ของ Moving Average

p คือ อันดับของ Autoregressive

q คือ อันดับของ Moving Average

ε_t คือ ค่าความคาดเคลื่อน ณ เวลาที่ t

2.1.5 แบบจำลอง ARFIMA (The Autoregressive Fractionally Integrated Moving Average)

แบบจำลอง ARFIMA ได้เสนอโดย Granger และ Joyeux (1980) ต่อมา Hosking (1981) ได้ใช้แบบจำลองนี้กับข้อมูลชุดความจำระยะยาว (Long memory) ซึ่งแบบจำลอง ARFIMA(p,d,q) สามารถเขียนสมการความสัมพันธ์ ได้ดังนี้

$$\phi(\beta \Delta^d y_t) = \delta + \theta(\beta)\varepsilon_t \quad (59)$$

กับ

$$\phi(\beta) = 1 - \phi_1(\beta) - \phi_2(\beta)^2 - \ldots. - \phi_p(\beta)^p \quad (60)$$

และ

$$\theta(\beta) = 1 - \theta_1(\beta) - \theta_2(\beta)^2 - \ldots. - \theta_q(\beta)^q \quad (61)$$

โดย

$$\delta = \text{ ค่าคงที่}$$

$$\theta(\beta) = \text{ การทำงานของ moving-average ที่ order q}$$

$$\varepsilon_t = \text{ error term ของสมการ}$$

$$\phi(\beta) = \text{ การทำงานของ autoregressive ที่ order p}$$

$$\Delta^d y_t = \text{ผลต่างที่ order d ของข้อมูลอนุกรมเวลา } y_t$$

ถ้า d อยู่ที่ (0 ,0.5) ในแบบจำลอง ARFIMA อธิบายว่า Long memory หรือ Long range มี ความสัมพันธ์กันในทิศทางบวก

ถ้า d อยู่ที่ (-0.5 ,0) ในแบบจำลอง ARFIMA อธิบายว่า Intermediate memory หรือ Long range มี ความสัมพันธ์กันในทิศทางลบ

ถ้า d อยู่ที่ (0.5 ,1) ในแบบจำลอง ARFIMA อธิบายว่า เป็นค่าเฉลี่ยย้อยหลัง และไม่มีผลกระทบระยะยาวกับมูลค่าในอนาคต

ถ้าเป็นใน Short memory d=0 จะสอดคล้องกับค่ามาตรฐานตามวิธี ARMA (ชูเกียรติ ชัยบุญศรี และคณะ, 2553)

2.1.6 เกณฑ์การเลือกรูปแบบของแบบจำลองที่ดีที่สุด

ในการหารูปแบบของแบบจำลอง เมื่อได้รูปแบบของแบบจำลองที่เหมาะสมหลายรูปแบบต้องมีแนวทางในการเลือกรูปแบบของแบบจำลองที่ดีที่สุด โดยพิจารณาจากค่า Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) รูปแบบจำลองที่ให้ค่า AIC และ SC น้อยที่สุดจะเป็นรูปแบบที่ดีที่สุด โดย Akaike Information Criterion (AIC) และ Schwartz Criterion (SC) สามารถคำนวณได้ดังนี้

Akaike Information Criterion (AIC) $-2l/\eta + 2k/\eta$

Schwartz Criterion (SC) $-2l/\eta + k\log\eta/\eta$

โดยที่ k เป็นจำนวนพารามิเตอร์ที่ทำการประมาณค่า

η เป็นจำนวนของค่าสังเกต

l เป็นค่าของ Log Likelihood function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว

2.1.7 ค่าความคลาดเคลื่อนร้อยละเฉลี่ย (The Mean Absolute Percentage Error : MAPE)

ในทางสถิติแล้ว MAPE คือ เครื่องมือที่มีความแม่นยำ ในการหาค่า Time series ที่มีความเหมาะสมทางสถิติ โดยเฉพาะอย่างยิ่งในการหาแนวโน้มโดยปกติแล้ว MAPE ถูกใช้ในการ แสดงค่าที่เหมาะสม โดยการแสดงเป็นเปอร์เซ็นต์ ซึ่งสูตรของ MAPE คือ

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - \hat{Y}_t}{Y_t}\right|$$

เมื่อให้ Y_t คือค่าที่แท้จริง และ F_t คือค่าที่คาดการไว้

ค่าความแตกต่างระหว่าง Y_t และ $\hat{Y}_t$ ได้แสดงให้เห็นโดย ค่าที่แท้จริง Y_t อีกครั้ง ค่าสัมบูรณ์ที่ได้จากการคำนวณนี้ คือผลบวกของทุกๆจุดที่เหมาะสมหรือทุกๆจุดที่คาดการณ์ ไว้ในล้วงเวลาและได้แสดงให้เห็นอีกครั้งด้วยจำนวนที่เหมาะสม n ทำให้เกิดข้อผิดพลาดด้านเปอร์เซ็นต์ ดังนั้นจึงสามารถเปรียบเทียบข้อผิดพลาดของ Time series ที่เหมาะสมซึ่งมีความต่างกัน

ประโยชน์ของ ค่า MAPE คือความสามารถในการเปรียบเทียบระหว่างความแตกต่างของแบบจำลองการพยากรณ์ และมีความชัดเจนในการแปรความหมาย ตัวชี้นำของการแปลความหมาย MAPE เป็นดังนี้

- ถ้าหากค่า MAPE น้อยกว่า 10% การทำนายจะมี“ความแม่นยำสูงมาก”
- ถ้าหากค่า MAPE อยู่ระหว่าง 10%-20% การทำนายจะมี“ความแม่นยำสูง”
- ถ้าหากค่า MAPE อยู่ระหว่าง20-50% การทำนายจะมี“ความแม่นยำปานกลาง”
- ถ้าหากค่า MAPE สูงกว่า 50% การทำนายจะ “ไม่มีความแม่นยำ” (Lewis, 1982)

2.1.8 รากที่สองของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสอง (Root Mean Square Error: RMSE)

ค่ารากที่สองของค่าเฉลี่ยค่าคลาดเคลื่อนกำลังสอง เป็นค่าที่แสดงความคลาดเคลื่อนระหว่างค่าที่ประมาณได้จากแบบจำลองกับค่าข้อมูลจริง ซึ่งถ้าค่า RMSE มีค่าเข้าใกล้ศูนย์แสดงว่าแบบจำลองนี้มีความคลาดเคลื่อนน้อยสามารถใช้เป็นตัวแทนข้อมูลจริงได้อย่างเหมาะสม ซึ่งมีสมการดังนี้

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{Y}_t - Y_t)^2}$$

เมื่อให้ $\hat{Y}_t$ คือ ค่าประมาณจากแบบจำลอง

Y_t คือ ค่าที่แท้จริง / ข้อมูลจริง

2.1.9 ค่าสัมประสิทธิ์ Theil (Theil's Ineqvality Coefficient: U)

ค่า Theil's Ineqvality Coefficient (U) หลักในการพิจารณาค่าคลาดเคลื่อนนั้นเหมือนกับค่า RMSE แต่ต่างกันตรงที่ค่า Theil's Ineqvality Coefficient จะมีค่าระหว่าง 0 ถึง 1 ซึ่งถ้าเข้าใกล้ 0 แสดงว่าแบบจำลองนี้สามารถใช้เป็นตัวแทนข้อมูลจริงได้อย่างเหมาะสม ซึ่งมีสมการดังนี้

$$U = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{Y}_t - Y_t)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{Y}_t)} + \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t)^2}}$$

เมื่อให้ $\hat{Y}_t$ คือ ค่าประมาณจากแบบจำลอง

Y_t คือ ค่าที่แท้จริง / ข้อมูลจริง

2.2 เอกสารและงานวิจัยที่เกี่ยวข้อง

แบบจำลองอาร์รีมาและอาร์ฟีม่าได้รับความนิยมเป็นอย่างมากในการใช้พยากรณ์ข้อมูลอนุกรมเวลา แสดงได้ดังนี้

นริสา สมุทรสาคร (2547) ได้ทำการศึกษาการพยากรณ์ทองคำโดยใช้วิธีอารีมา โดยมีวัตถุประสงค์เพื่อพยากรณ์ราคาทองคำ 2 ประเภทคือ ราคาขายทองแท่งและทองรูปพรรณ โดยการพยากรณ์ใช้ข้อมูลเป็นรายเดือนจำนวน 120 เดือน ตั้งแต่ปี 2537 ถึง 2546 ซึ่งเก็บรวบรวมข้อมูลจากธนาคารแห่งประเทศไทย โดยใช้แบบจำลองอารีมาด้วยวิธีบอกส์และเจนส์กินส์ (Box - Jennkins) จากการศึกษาพบว่า ข้อมูลราคาทองคำแท่งแลทองรูปพรรณมีลักษณะไม่นิ่ง จากการทดสอบความนิ่งของข้อมูล พบว่าข้อมูลนิ่งที่ระดับ (1) ทั้งนี้เนื่องจากการพิจารณาคอเรลโลแกรม ผลปรากฏว่าแบบจำลอง AR (2) MA (2) MA(5) มีความเหมาะสมมากที่สุดสำหรับข้อมูลทองแท่งและทองรูปพรรณ เมื่อทำการตรวจสอบความถูกต้องของแบบจำลองพบว่าแบบจำลองมีลักษณะ White Noise ที่ระดับสำคัญทางสถิติที่ 1% แบบจำลองทั้งสองให้ค่า Root Mean Squared Error (RMSE)และค่า Theil's Inequality Coefficient (U) ที่ต่ำที่สุด ดังนั้นแบบจำลองดังกล่าวจึงมีความเหมาะสมที่สุดที่จะเป็นตัวแทนของราคาขายทองแท่งและทองรูปพรรณเพื่อการพยากรณ์ในอนาคต โดยผลการพยากรณ์ราคาขายทองแท่งระหว่างเดือนมกราคมถึงเดือนเมษายน 2547 ราคาพยากรณ์ที่ได้เท่ากับ 7,817.89 7,715.80 7,755.11 และ 7,761.17 บาทต่อบาททองคำ ตามลำดับ ส่วนราคา

ขายทองรูปพรรณระหว่างเดือนมกราคมถึงเดือนเมษายน 2547 ราคาพยากรณ์ที่ได้เท่ากับ 7817.89 7, 893.76 7, 915.98 7,917.87 ตามลำดับ

จิตราภรณ์ ผั้นศิริ (2547) ได้ทำการศึกษารูปแบบและพยากรณ์ราคาส่งออกข้าวของไทยโดยวิธีอารีมาในรูปแบบของ Box-Jenkins พบว่าข้อมูลราคาส่งออกข้าวมีลักษณะไม่นิ่ง จึงได้ทำการหาผลต่างลำดับที่ 1 และจากการพิจารณา Correlogram เพื่อหาแบบจำลองที่เหมาะสมคือ AR(1)และ AR(19) โดยมีค่าสัมประสิทธ์ 0.360 และ 0.228 ตามลำดับ อย่างมีนัยสำคัญทางสถิติที่ระดับ 1% หลังจากนั้นได้ตรวจสอบความถูกต้องของแบบจำลองพบว่า ค่าประมาณการของความคลาดเคลื่อนมีลักษณะเชิงสุ่ม (White noise) ที่ระดับนัยสำคัญทางสถิติ 10% จากค่ารากที่สองของค่าเฉลี่ยค่าความคลาดเคลื่อนกำลังสองและค่าสัมประสิทธิ์ Theil ที่มีค่าต่ำสุด เมื่อได้แบบจำลองที่เหมาะสมคือ AR(1) และ AR(19) จึงได้นำไปพยากรณ์ราคาส่งออกข้าวในช่วงเดือนมกราคม ถึงเดือนเมษายน พ.ศ.2547 ต่อไปได้ค่าเท่ากับ 205 204 202 และ 201 เหรียญสหรัฐต่อตันตามลำดับ

ราชพล สุนทรศรี (2548) ได้ทำการศึกษาการพยากรณ์ราคาน้ำมันดิบโดยวิธีอารีมาโดยมีวัตถุประสงค์เพื่อที่จะวิเคราะห์และพยากรณ์การเคลื่อนไหวของราคาน้ำมันดิบที่นำเข้าจากต่างประเทศ คือ ประเทศโอมาน ประเทศดูไบ ประเทศไนจีเรียและประเทศอังกฤษ การพยากรณ์ใช้ข้อมูลเป็นรายเดือนจำนวน 271 เดือน ตั้งแต่เดือนมกราคม พ.ศ. 2527 ถึงเดือนตุลาคมพ.ศ.2547 และข้อมูลรายไตรมาส ตั้งแต่ไตรมาสที่ 1 พ.ศ. 2527 ถึงไตรมาสที่ 3 พ.ศ. 2547 โดยใช้แบบจำลองอารีมา ช่วยในการวิเคราะห์โดยวิธีบ็อกซ์และเจนกินส์ จากการศึกษาโดยการทดสอบ unit root ของราคาน้ำมันดิบที่นำเข้าจากทั้ง 4 ประเทศ พบว่าข้อมูลราคาน้ำมันมี unit root จึงทำการหาผลลำดับที่ 1 จากการพิจารณาคอเรลโลแกรมผลปรากฏว่าแบบจำลอง AR(1) MA(1), AR(1) AR(2) MA(1), AR(2) MA(1) MA(2) และ AR(2) MA(1) MA(2) และ AR(1) MA(1), AR(1) MA(2), AR(1)AR(2)MA(1)MA(2) และ AR(1) AR(2) MA(1) MA(2) ของข้อมูลแบบรายไตรมาส เปนีแบบจำลองที่มีความเหมาะสมมากที่สุดของการพยากรณ์ราคาน้ำมันดิบที่นำเข้าจากประเทศโอมานประเทศดูไบประเทศไนจีเรียและประเทศอังกฤษตามลำดับ เมื่อทำการทดสอบความถูกต้องของแบบจำลองพบว่าค่าส่วนเหลือของทุกแบบจำลองมีลักษณะเป็น white noise ที่ระดับนัยสำคัญทางสถิติที่ 0.01 และทุกแบบจำลองข้างต้นจึงมีความเหมาะสมที่จะเป็นตัวแทนของราคาน้ำมันดิบในอนาคตที่นำเข้าจากแต่ละประเทศ โดยผลการพยากรณ์ราคาน้ำมันดิบที่นำเข้าจากประเทศโอมานระหว่างเดือนพฤศจิกายน พ.ศ. 2547 ถึงเดือนมกราคม พ.ศ. 2548 คือ 38.63 38.25 และ 38.13 ดอลลาร์/บาร์เรลส่วนผลการพยากรณ์ราคาน้ำมันดิบที่นำเข้าจากตลาดดูไบระหว่างเดือน

พฤศจิกายน พ.ศ. 2547 ถึงเดือนมกราคม พ.ศ. 2548 คือ 36.71 36.31 และ 35.89 ดอลลาร์/บาร์เรล และผลการพยากรณ์ราคาน้ำมันดิบที่นำเข้าจากตลาดไนจีเรียระหว่างเดือนพฤศจิกายน พ.ศ. 2547 ถึงเดือนมกราคม พ.ศ. 2548 คือ 49.44 48.94 และ 49.43 ดอลลาร์/บาร์เรล ผลการพยากรณ์ราคาน้ำมันดิบที่นำเข้าจากประเทศโอมาน ดูไบ ไนจีเรียและอังกฤษ ตั้งแต่ไตรมาสที่ 4 พ.ศ. 2547 ถึงไตรมาสที่ 2 พ.ศ. 2548 คือ 33.55 34.50 33.65 ; 34.67 34.73 35.95 ; 40.57 39.46 41.16 และ 40.74 40.44 42.74 ดอลลาร์/บาร์เรล ตามลำดับดังนั้นจึงสรุปได้ว่าการพยากรณ์ราคาน้ำมันที่นำเข้าจากต่างประเทศจะเป็นประโยชน์ต่อการกำหนดมาตรการเพื่อรองรับปัญหาที่อาจเกิดจากความผันผวนของราคาน้ำมันในอนาคต

ณันทิวร เทียนแป้น (2552) วัตถุประสงค์ในการศึกษาครั้งนี้ประกอบด้วย หนึ่ง เพื่อทดสอบการมีอยู่ของ long memoryในข้อมูลราคาน้ำมันดิบไลท์สวีทและน้ำมันสำเร็จรูปเบนซินรายวันในตลาดฟิวเจอร์ไนเม็กซ์ สอง เพื่อหาแบบจำลองอาร์ฟีมา ที่เหมาะสมและพยากรณ์การเปลี่ยนแปลงของราคาน้ำมันดิบไลท์สวีท และน้ำมันสำเร็จรูปเบนซินรายวันในอนาคตจากผลการศึกษาและทดสอบการมี long memory ด้วยวิธี R/S Statistic และ GPH Test ในราคาน้ำมันดิบไลท์สวีทและน้ำมันสำเร็จรูปเบนซินจากตลาดล่วงหน้าไนเม็กซ์และสร้างแบบจำลองเพื่อการพยากรณ์ระหว่างวันที่ 2 กุมภาพันธ์ 2552 ถึง วันที่ 27 กุมภาพันธ์ 2552 ซึ่งได้ผลการศึกษาต่างกัน คือ ประการที่หนึ่ง ราคาน้ำมันดิบไลท์สวีทมี long memory และแบบจำลองที่เหมาะสมต่อการพยากรณ์ราคาคือ ARFIMA(10, 0.1142, 0) เมื่อ $d \in (0.5)$ โดยร้อยละของค่าความเบี่ยงเบนมีค่าอยู่ในช่วง 1.19 ถึง 23.09 ประการที่สอง ราคาน้ำมันสำเร็จรูปเบนซินมี long memoryแต่ไม่สามารถเลือกแบบจำลองอาร์ฟีมาที่เหมาะสมได้ เพราะ $d \notin (0,0.5)$ จึงได้ใช้แบบจำลองARIMA ในการวิเคราะห์ราคาน้ำมันเบนซิน คือ ARIMA(4, 1, 4) โดยร้อยละของค่าความเบี่ยงเบนมีค่าอยู่ในช่วง 0.73 ถึง 17.86 ดังนั้นการทดสอบข้อมูลอนุกรมเวลาเพื่อศึกษาการมี long memory และทดสอบความแม่นยำของการพยากรณ์ด้วยแบบจำลองอาร์ฟีมาควรใช้กับข้อมูลอนุกรมเวลาที่มีความถี่สูง หรือถ้าข้อมูลที่มีจำนวนไม่เกิน 1,000 ค่าสังเกต ควรจะใช้แบบจำลองอารีมาจึงจะเหมาะสมสำหรับการพยากรณ์ข้อมูลอนุกรมเวลา