# บทที่ 3
# ระเบียบวิธีวิจัย

ในการศึกษาการวิเคราะห์ความผันผวนและพยากรณ์มูลค่าหน่วยลงทุนกองทุนหุ้นระยะยาวโดยใช้แบบจำลอง ARIMA-GARCH ใช้แนวคิดและทฤษฎีเกี่ยวกับการวิเคราะห์อนุกรมเวลา การทดสอบความนิ่งของข้อมูล (stationary) โดยการทดสอบ Unit Root และสร้างแบบจำลองที่ดีที่สุดเพื่อประมาณค่าความผันผวนและพยากรณ์มูลค่าหน่วยลงทุนในอนาคต แบบจำลองที่ใช้ คือ แบบจำลอง ARIMA (Autoregressive Intergrated Movingaverage) แบบจำลอง Autoregressive Conditional Heteroscadasticity (ARCH) แบบจำลอง Gerneralized Conditional Heteroscadasticity (GARCH) และวิธีการพยากรณ์

## 3.1 การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

อนุกรมเวลา คือ เซตของข้อมูลเชิงปริมาณที่มีการจัดเก็บในช่วงเวลาที่ติดต่อกัน ส่วนข้อมูลอนุกรมเวลา (Time Series Data) คือ ชุดของข้อมูลที่มีการเก็บรวบรวมตามระยะเวลาที่ติดต่อกันอย่างเป็นระบบโดยทั่วไปข้อมูลอนุกรเวลาประกอบด้วยองค์ประกอบ 4 ส่วน คือ แนวโน้ม (Trend: T), ฤดูกาล (Seasonal: S), วัฏจักร (Cycle: C) และเหตุการณ์ที่ผิดปกติหรือเหตุการณ์ความไม่แน่นอน (Irregular: I) สำหรับรูปแบบของอนุกรมเวลาโดยทั่วไปนั้นมีอยู่ 2 รูปแบบคือ

ก) รูปแบบบวก $$Y = T + S + C + I \tag{1}$$

ข) รูปแบบคูณ $$Y = T \times S \times C \times I \tag{2}$$

โดยที่ Y คือ อนุกรมเวลา
T คือ อิทธิพลของแนวโน้ม
S คือ อิทธิพลของฤดูกาล
C คือ อิทธิพลของวัฏจักร
I คือ อิทธิพลของความไม่แน่นอน

การวิเคราะห์อนุกรมเวลาเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้น หรือการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านไป ลักษณะของการเปลี่ยนแปลงอาจมีหรือไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีตก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะ

การเปลี่ยนแปลงควรอยู่ในรูปแบบใด และสามารถพยากรณ์การเปลี่ยนแปลงข้อมูลในอนาคตได้ การวิเคราะห์ข้อมูลอนุกรมเวลานี้ จะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน (ศิริลักษณ์ เล็กสมบูรณ์ 2531)

นำข้อมูลมูลค่าทรัพย์สินสุทธิ (NAV) ซึ่งเป็นข้อมูลทุติยภูมิ (Secondary data) มาวิเคราะห์ข้อมูลอนุกรมเวลา (time series data) ได้สมการในการทดสอบดังนี้

$$\mathrm{Nav}_{1} = \rho\,\mathrm{Nav}_{t-1} + \varepsilon_{t} \qquad (3)$$

กำหนดให้

$\mathrm{Nav}_1$ คือ ตัวแปรที่เราทำการศึกษาได้แก่ มูลค่าทรัพย์สินสุทธิหรือมูลค่าหน่วยลงทุน

t คือ แนวโน้มเวลา

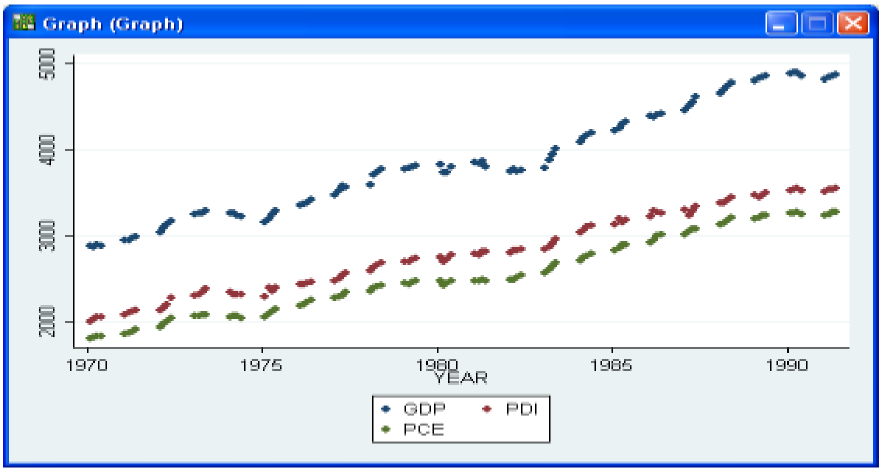
$\varepsilon_t$ คือ ตัวแปรสุ่มโดยมีการแจกแจงแบบปกติที่เป็นอิสระต่อกันและเหมือนกัน

นำข้อมูลอนุกรมเวลามาทดสอบ Stationary แบ่งได้เป็น 3 วิธีหลัก ๆ คือ

1) การทดสอบโดยใช้กราฟ
2) การทดสอบโดยใช้ Correlogram test
3) การทดสอบโดยใช้ Unit Root test

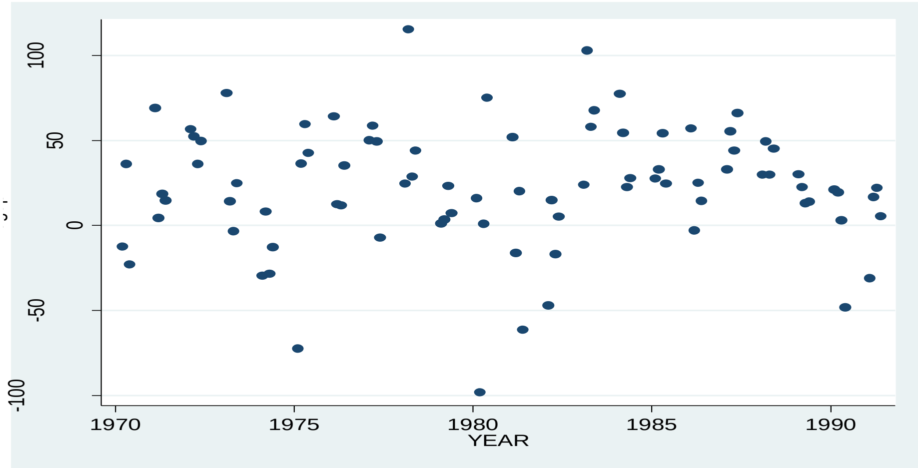
### 3.1.1 การทดสอบโดยใช้กราฟ

การทดสอบโดยใช้กราฟ ข้อมูลที่ไม่เสถียรจะมีรูปแบบ ดังนี้




**ภาพที่ 3.1** ตัวอย่างข้อมูลที่ไม่มีเสถียรภาพ

การทดสอบโดยใช้กราฟข้อมูลที่มีเสถียรภาพจะกระจายอยู่รอบ ๆ ค่าเฉลี่ย




**ภาพที่ 3.2** ตัวอย่างข้อมูลที่มีเสถียรภาพ

**3.1.2 การทดสอบโดยใช้ Correlogram test (Correlogram and Q-statistics)**

การพิจารณา Correlogram and Q-statistics มีเงื่อนไขในการพิจารณาว่ากราฟ Correlogram ของAutocorrelation ของตัวแปรที่กำลังพิจารณาจะต้องไม่มีลักษณะการลดลงแบบ Exponential และค่า Q-statistics ที่คำนวณได้มีค่ามากกว่าค่าวิกฤติของ Chi-square ณ ระดับนัยสำคัญ

## 3.2 การทดสอบความนิ่งของข้อมูล (Stationary) และการทดสอบ Unit Root

การทดสอบอนุกรมเวลา เป็นสิ่งที่ควรกระทำก่อนที่จะนำข้อมูลข้อมูลอนุกรมเวลามาใช้ในการวิเคราะห์ โดยเฉพาะเงื่อนไขความนิ่งของอนุกรมเวลา (Stationary) ซึ่งเป็นเงื่อนไขที่สำคัญในการนำข้อมูลอนุกรมเวลามาใช้ ดังนั้นถ้าหากอนุกรมเวลาที่นำมาใช้ไม่คงที่จะต้องทำให้อนุกรมเวลาดังกล่าวคงที่ก่อน โดยการหาผลต่างของอนุกรมเวลา การคงที่ของอนุกรมเวลา หมายถึง อนุกรมเวลาที่อยู่ในสภาวะสมดุลเชิงสถิติ (Statistical equilibrium) ซึ่งก็คือ การที่คุณสมบัติทางสถิติของอนุกรมเวลาไม่เปลี่ยนแปลงตามกาลเวลา

การทดสอบ Unit Root เป็นการทดสอบว่าข้อมูลที่นำมาศึกษามีนิ่งหรือไม่ สามารถทำได้โดยการทดสอบ DF (Dickey – Fuller Test) ซึ่งเสนอโดย Dickey และ Fuller ในปี 1981 และวิธีการทดสอบ ADF (Augmented Dickey – Fuller Test) ซึ่งเสนอโดย Said และ Dickey ในปี 1984

ข้อมูลที่มีลักษณะนิ่ง (Stationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) เท่ากันตลอดระยะเวลาที่ศึกษา

เมื่อสมมติให้ตัวแปร $X_t$ เป็นอนุกรมเวลาที่คงที่ (Stationary) ดังนั้น ตัวแปร $X_t$ จะมีคุณสมบัติดังนี้

Mean: $E(X_t) = \mu$ (4)

Variance: $Var(X_t) = E(X_t - \mu)^2 = \sigma^2$ (5)

Covariance: $E[(X_t - \mu)(X_{t+k} - \mu)] = \gamma_k$ (6)

ส่วนข้อมูลที่มีลักษณะไม่นิ่ง (Non-Stationary) หมายถึงข้อมูลอนุกรมที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) ไม่เท่ากันตลอดระยะเวลาที่ศึกษา

เมื่อสมมติให้ตัวแปร $X_t$ เป็นอนุกรมเวลาที่ไม่คงที่ (Nonstationary) ดังนั้น ตัวแปร $X_t$ จะมีคุณสมบัติดังนี้

Mean: $E(X_t) = t\mu$ (7)

Variance: $Var(X_t) = E(X_t - \mu)^2 = t\sigma^2$ (8)

Covariance: $E[(X_t - \mu)(X_{t+k} - \mu)] = t\gamma_k$ (9)

**3.2.1 วิธีการทดสอบ DF (Dickey – Fuller Test)**

มีสมการที่ต้องทดสอบอยู่ 3 สมการ (At Level) คือ

$\Delta X_t = \gamma X_{t-1} + \varepsilon_1$ (random walk process) (10)

$\Delta X_t = \alpha + \gamma X_{t-1} + \varepsilon_1$ (random walk with drift) (11)

$\Delta X_t = \alpha + \beta_t + \gamma X_{t-1} + \varepsilon_1$ (random walk with drift และมี linear Time and trend) (12)

สมมติฐานที่ทดสอบ

$H_0 : \gamma = 0$

$H_1 : \gamma \neq 0$

ถ้ายอมรับ $H_0$ แสดงว่า $X_t$ มีลักษณะไม่นิ่ง(Nonstationary) มีการเปลี่ยนแปลงเมื่อเวลาเปลี่ยนแปลงและถ้าปฏิเสธ $H_0$ แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง (Statioary)

สมการที่ใช้ทดสอบตามวี DF

$$\Delta Nav_1 = \theta \, \Delta Nav_{t-1} + \varepsilon_t \quad (13)$$

$$\Delta Nav_1 = \alpha + \theta \, \Delta Nav_{t-1} + \varepsilon_t \quad (14)$$

$$\Delta Nav_1 = \alpha + \beta_t + \theta \, \Delta Nav_{t-1} + \varepsilon_t \quad (15)$$

### 3.2.2 วิธีการทดสอบ ADF (Augmented Dickey–Fuller Test)

การทดสอบ ADF (Augmented Dickey–Fuller Test) พัฒนามาจากวิธี Dickey – Fuller Test เพื่อแก้ปัญหา Serial Correlation หรือแก้ปัญหา Autocorrelation ซึ่งค่าสถิติที่ได้จะขาดความถูกต้องดังนั้น จึงได้มีการนำเสนอให้ปรับสมการของวิธีการ Dickey-Fuller ใหม่ โดยใส่ตัวแปรล่า (Lag) ของ X ในลำดับที่สูงขึ้น แล้วเรียกวิธีการนี้ว่า Augmented Dickey–Fuller Test ในการตรวจสอบว่าข้อมูลนั้นมีลักษณะนิ่งหรือไม่ โดยการเปรียบเทียบค่าสถิติ t ที่คำนวณได้กับค่าวิกฤตในตาราง ADF (อัครพงศ์ อั้นทอง, 2550)

มีสมการที่ต้องทดสอบอยู่ 3 สมการ (At Level) คือ

$$\Delta X_t = \gamma X_{t-1} + \sum_{i=1}^{\rho} \phi i \Delta X_{t-1} + \varepsilon_1 \quad \text{(random walk process)} \quad (16)$$

$$\Delta X_t = \alpha + \gamma X_{t-1} + \sum_{i=1}^{\rho} \phi i \Delta X_{t-1} + \varepsilon_1 \quad \text{(random walk with drift)} \quad (17)$$

$$\Delta X_t = \alpha + \beta_t + \gamma X_{t-1} + \sum_{i=1}^{\rho} \phi i \Delta X_{t-1} + \varepsilon_1 \quad \text{(random walk with linear and trend)} \quad (18)$$

สมมติฐานที่ทดสอบ

$H_0 : \gamma = 0$

$H_1 : \gamma \neq 0$

ถ้ายอมรับ $H_0$ แสดงว่า $X_t$ ข้อมูลมีลักษณะไม่นิ่ง (Nonstationary) และถ้าปฏิเสธ $H_0$ แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง (Stationary) ข้อมูลที่มีลักษณะนิ่ง (Stationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) เท่ากันตลอดระยะเวลาที่ศึกษา ส่วนข้อมูลที่มีลักษณะไม่นิ่ง (Non-Stationary) หมายถึงข้อมูลอนุกรมที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) ไม่เท่ากันตลอดระยะเวลาที่ศึกษา

สมการที่ใช้ในการทดสอบตามวิธี ADF

$$\Delta Nav_1 = \theta\ Nav_{t-1} + \sum_{j=1}^{p} \lambda_j\ \Delta Nav_{t-1} + \varepsilon_t \quad (19)$$

$$\Delta Nav_1 = \alpha + \theta\ Nav_{t-1} + \sum_{j=1}^{p} \lambda_j\ \Delta Nav_{t-1} + \varepsilon_t \quad (20)$$

$$\Delta Nav_1 = \alpha + \beta_t + \theta\ Nav_{t-1} + \sum_{j=1}^{p} \lambda_j\ \Delta Nav_{t-1} + \varepsilon_t \quad (21)$$

**3.3 การเลือกแบบจำลองจากการทดสอบ Unit Root โดยการทดสอบสัมประสิทธิ์ของการถดถอย (Deterministic Regressors)**

เป็นการทดสอบว่า แบบจำลองใดเป็นแบบจำลองที่เหมาะสมที่สุดระหว่างกรณีของแบบจำลองที่ไม่มีค่าคงที่และแนวโน้มเวลา (None) แบบจำลองที่มีค่าคงที่ (Intercept) และแบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา (Trend and Intercept) โดยการทดสอบการมีนัยสำคัญทางสถิติของสัมประสิทธิ์ของตัวถดถอย (ค่าคงที่หรือค่าแนวโน้มเวลา) โดยขั้นตอนการทดสอบดังนี้

**ขั้นตอนที่ 1** เริ่มการทดสอบจากแบบจำลองกรณีที่มีทั้งค่าคงที่และแนวโน้มเวลาตามสมการ

$$\Delta X_t = a_0 + \beta + \gamma X_{t-1} + a_{2t} + \sum \beta i \Delta X_{t-1} + \varepsilon_1 \quad (22)$$

ทำการทดสอบสมมติฐานว่าง $H_0 : \gamma = 0$ โดยใช้ $\tau$ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่าง แสดงว่า ข้อมูล $X_t$ มีลักษณะนิ่งแล้ว และเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา

**ขั้นตอนที่ 2** ถ้าเกิดการยอมรับสมมติฐานว่างในขั้นตอนที่ 1 แสดงว่าในแบบจำลองดังกล่าวมีตัวถดถอยที่ไม่จำเป็นอยู่ในสมการ ซึ่งอาจทำให้อำนาจการทดสอบของสมการลดลง ดังนั้น จึงต้องมีการทดสอบนัยสำคัญทางสถิติของค่าแนวโน้ม ($a_{2t}$) ที่อยู่ในสมการ โดยการทดสอบสมมติฐานว่า $H_0 : a_2 = \gamma = 0$ โดยใช้ $\phi$ statistic ถ้าค่าสัมประสิทธิ์ของค่าแนวโน้มไม่มีนัยสำคัญทางสถิติให้ข้ามไปขั้นตอนที่ 3 อย่างไรก็ตาม ถ้าค่าสัมประสิทธิ์ของค่าแนวโน้มมีนัยสำคัญทางสถิติให้ทำการทดสอบความไม่นิ่งของข้อมูลอีกครั้ง โดยใช้ การแจกแจงแบบปกติมาตรฐาน (standardized normal distribution) ถ้าเกิดการปฏิเสธสมมติฐานว่าแสดงว่า ข้อมูล $X_t$ มีลักษณะนิ่งแล้วจะเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา แต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล $X_t$ มีลักษณะไม่นิ่ง

**ขั้นตอนที่ 3** ทำการประเมินค่าแบบจำลอง ที่ปราศจากค่าแนวโน้มเวลา และทดสอบ unit root โดยใช้ $\tau_\mu$ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่าข้อมูล $X_t$ มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่ปราศจากค่าแนวโน้มเวลา แต่ถ้าเกิดการยอมรับสมมติฐานว่างให้ทำการทดสอบความไม่นิ่งของข้อมูลอีกครั้งโดยใช้ การแจกแจงแบบปกติมาตรฐาน (standardized normal distribution) ถ้าเกิดการปฏิเสธสมมติฐานว่างแสดงว่า ข้อมูล $X_t$ มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่มีทั้งค่าคงที่และแนวโน้มเวลา แต่ถ้าเกิดการยอมรับสมมติฐานว่างแสดงว่า ข้อมูล $X_t$ มีลักษณะไม่นิ่ง

**ขั้นตอนที่ 4** ทำการประมาณค่าแบบจำลองที่ปราศจากค่าแนวโน้มเวลา และค่าคงที่ และทดสอบ Unit root โดยใช้ $\tau$ statistic ถ้าเกิดการปฏิเสธสมมติฐานว่าง แสดงว่าข้อมูล $X_t$ มีลักษณะนิ่งแล้วและเลือกใช้แบบจำลองที่ปราศจากค่าแนวโน้มเวลาและค่าคงที่ แต่ถ้าเกิดการยอมรับสมมติฐานว่าง แสดงว่าข้อมูล $X_t$ มีลักษณะไม่นิ่ง

ทั้งนี้ การวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลา ส่วนมากจะพบปัญหาความไม่นิ่งของข้อมูล เนื่องจากข้อมูลนั้นมาจากกระบวนการเชิงสุ่ม (random process) เมื่อนำข้อมูลอนุกรมเวลาไปใช้โดยไม่ได้ทำการตรวจสอบว่าข้อมูลมีลักษณะนิ่งนั้น ค่าสถิติที่เกิดขึ้นจะมีการแจกแจงไม่มาตรฐาน (standard distributions) ทำให้นำไปสู่การลงความเห็นที่ผิดพลาดและความสัมพันธ์ที่ไม่แท้จริง (spurious regression) กล่าวคือ R มีค่าสูงมากและได้ค่าสถิติ t-test มีนัยสำคัญหรือสูงเกินกว่าความเป็นจริง ซึ่งสามารถแก้ไขได้ด้วยการทำให้ข้อมูลมีความนิ่งเสียก่อน โดยอาจใช้วิธีการหาผลต่าง (Difference) ของข้อมูล การแปลงให้อยู่ในรูป Logarithm หรือการทดสอบหาความสัมพันธ์ของตัวแปรในระยะยาว (Cointegration) เป็นต้น

## 3.4 การเลือกแบบจำลองที่ดีที่สุด โดยการตรวจสอบรูปแบบ (Diagnostic Checking)

หลังจากการสร้างสมการพร้อมทั้งประมาณค่าพารามิเตอร์แล้ว สิ่งสำคัญคือจะต้องทำคือการตรวจสอบรูปแบบว่าสมการที่ได้มามีความเหมาะสมหรือไม่ และรูปแบบใดของสมการที่ดีที่สุด

### 3.4.1 เกณฑ์ในการเลือกแบบจำลองที่ดีที่สุด (Information Criteria)

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

Akaike Information Criterion (AIC) $= -2\iota/\eta + 2\kappa/\eta$ (23)

Schwartz Criterion (SC) $= -2\iota/\eta + \kappa \log \eta/\eta$ (24)

โดยที่ $\kappa$ เป็นจำนวนของพารามิเตอร์ที่ทำการประมาณค่า

$\eta$ เป็นจำนวนของค่าสังเกต

$\iota$ เป็นค่าของ log likelihood function ที่ใช้พารามิเตอร์ที่ถูกประมาณค่า k ตัว

**3.4.2 แบบจำลอง ARIMA (Autoregressive Intergrated Movingaverage)**

แบบจำลองตัวแบบ ARIMA (p, d, q) มีส่วนประกอบที่สำคัญ 3 ส่วน ได้แก่ Auto Regressive AR : (p), Intergrated (I) และ Moving Average MA : (q) สำหรับ AR (p) เป็นรูปแบบที่แสดงว่าค่าสังเกต $y_t$ ขึ้นอยู่กับค่าของ $y_{t-1},\ldots, y_{t-p}$ หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ MA (q) เป็นรูปแบบที่แสดงค่าสังเกต $y_t$ ขึ้นอยู่กับค่าความคลาดเคลื่อน $\varepsilon_{t-1},\ldots, \varepsilon_{t-q}$ หรือความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า ส่วน Integrated (I) เป็นการหาผลต่าง (Difference) ของอนุกรมเวลา เหตุผลสำคัญที่ต้องหาผลต่างของอนุกรมเวลา เนื่องจากแบบจำลอง ARIMA จะต้องใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีคุณสมบัติคงที่ (Stationary) เท่านั้น ในกรณีที่ข้อมูลอนุกรมเวลาที่ใช้ในการวิเคราะห์สมบัติไม่คงที่ (Nonstationary) จะต้องทำการแปลงข้อมูลอนุกรมเวลาดังกล่าวให้มีคุณสมบัติคงที่ก่อน โดยการหาผลต่างของข้อมูลอนุกรมเวลา หรือการหาค่า Natural logarithm ของอนุกรมเวลาก่อนที่จะนำข้อมูลไปใช้สร้างแบบจำลอง ARIMA

$$\phi(B)\nabla^{d}y_t = \delta + \theta(B)\varepsilon_t \tag{25}$$

โดยที่

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \tag{26}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_p B^p \tag{27}$$

$y_t$ = ค่าสังเกตในอนุกรมเวลา ณ เวลา t

B = Backward shift operation โดยที่ $B_m = \nabla y_{t-m}$

d = จำนวนครั้งของการหาผลต่างเพื่อให้อนุกรมเวลามีคุณสมบัติคงที่ (stationary)

p = อันดับของออโตรีเกรสซีฟ (Autoregressive Order)

q = อันดับของค่าเฉลี่ยเคลื่อนที่ (Moving Average)

$\delta$ = ค่าคงที่ (Constant Term)

$\phi_1,\ldots, \phi p$ = พารามิเตอร์ของ ออโตรีเกรสซีฟ (Autoregressive parameter)

$\theta_1,\ldots, \theta p$ = พารามิเตอร์ของ ค่าเฉลี่ยเคลื่อนที่ (Moving-Average parameter)

$\varepsilon_t$ = กระบวนการ white noise ซึ่งก็คือ ค่าความคลาดเคลื่อน ณ เวลา t ภายใต้ข้อสมมติว่าความคลาดที่คนละเวลาเป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนคงที่ [$\varepsilon_t \sim N(0, \sigma^2)$]

จากสมการข้างต้นอาจเขียนได้ใหม่เป็น

$$\nabla^{d} y_t = \delta_{+} \phi \nabla^{d} y_{t-1} + \phi \nabla^{d} y_{t-2} + \ldots + \phi \nabla^{d} y_{t-p} + \varepsilon_t - \theta_t \varepsilon_{t-1} - \theta_q \varepsilon_{t-q} \quad (28)$$

จากรูปแบบทั่วไปตามสมการข้างต้นนำไปใช้ในการกำหนดรูปแบบที่เหมาะสม และประมาณค่าต่อไป ซึ่งอนุกรมเวลาที่จะนำมาวิเคราะห์ ด้วยวิธีของ Box and Jenkins นี้ ต้องมีเงื่อนไขบางประการเกี่ยวกับค่าพารามิเตอร์ในตัวแบบเพื่อให้อนุกรมเวลามีคุณสมบัติคงที่ (Stationary) และคุณสมบัติผกผัน (Invertibility) สำหรับคุณสมบัติคงที่ (Stationary) เป็นคุณสมบัติของรูปแบบ AR (p) ซึ่งเป็นคุณสมบัติที่ทำให้ E $(y_t)$ และ V $(y_t)$ คงที่ และ Cov $(y_t,\ldots,y_{t-k})$ มีค่าคงที่ ขึ้นกับว่า Lag k อย่างเดียว สวนคุณสมบัติผกผัน (Invertible) เป็นคุณสมบัติของรูปแบบ MA (q) ซึ่งเป็นคุณสมบัติที่ทำให้ค่าคลาดเคลื่อนของการพยากรณ์ $\varepsilon_t$ ในเทอมของ $y_t, y_{t-1}$ มีค่าคงที่ (ทรงศิริ แต้สมบัติ, 2539) มีขั้นตอนของวิธีการ Box and Jenkins

**ขั้นตอนที่ 1** กำหนดรูปแบบ (Identification) เพื่อหารูปแบบที่คาดว่าเหมาะสมให้กับอนุกรมเวลา โดยใช้วิธีพิจารณาเปรียบเทียบจาก Correlogram ของค่า $r_k$ และ $r_{kk}$ ของอนุกรมเวลา

**ขั้นตอนที่ 2** การประมาณค่าสัมประสิทธิ์ (Estimation) ในรูปแบบ โดยทั่วไปใช้วิธีการประมาณค่าสัมประสิทธิ์ด้วยวิธีกำลังสองน้อยที่สุดแบบธรรมดา (Ordinary Least Square Method: OLS)

**ขั้นตอนที่ 3** การตรวจสอบรูปแบบ (Diagnostic checking) เมื่อกำหนดรูปแบบและประมาณค่าพารามิเตอร์ในรูปแบบแล้ว ต้องตรวจสอบอีกครั้งว่ารูปแบบที่กำหนดมีความเหมาะสมจริงหรือไม่ โดยการพิจารณาค่าสหสัมพันธ์ในตัวเองของค่าความคลาดเคลื่อน (ดูจากกราฟ Correlogram) การทดสอบค่าพารามิเตอร์ในรูปแบบ โดยการพิจาณาจากค่าสถิติ t (t-statistic) และการทดสอบความเหมาะสมของรูปแบบโดยการทดสอบของ Box and Pierce หรือการทดสอบของ Box and Ljung [Q-statistic]

**ขั้นตอนที่ 4** การพยากรณ์ (Forecasting) นำสมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนดและผ่านการตรวจสอบรูปแบบ มาพยากรณ์ค่าในอนาคต โดยสามารถทำได้ทั้งการพยากรณ์แบบจุด (Point forecast) และการพยากรณ์แบบช่วง (Interval Forecast) การพยากรณ์โดยวิธี Box and Jenkins จะให้ค่าพยากรณ์ไปข้างหน้าที่ดีในช่วงสั้นๆ และต้องมีอนุกรมเวลาที่ยาวพอสมควร (อัครพงศ์ อั้นทอง, 2550)

**3.4.3 แบบจำลอง Autoregressive Conditional Heteroscedasticity (ARCH)**

ในการวิเคราะห์อนุกรมเวลาส่วนใหญ่แล้วจะมีการกำหนด stochastic variable ให้มีความแปรปรวนคงที่ (homoscedastic) ซึ่งในการประยุกต์ใช้กับบางข้อมูลนั้นค่าความแปรปรวนของค่าความคลาดเคลื่อน(error term) จะไม่ใช่ฟังก์ชั่นของตัวแปรอิสระแต่มีค่าเปลี่ยนแปลงไปตามช่วงเวลาขึ้นอยู่กับขนาดของความคลาดเคลื่อนที่เกิดขึ้นในอดีต และในบางงานศึกษา เช่น แบบจำลองของเงินเฟ้อ อัตราดอกเบี้ยหรือผลตอบแทนจากตลาดหลักทรัพย์ ในบางคาบเวลาจะมีค่าความผันผวน(volatility) สูง (และความคลาดเคลื่อนที่มีขนาดใหญ่) ตามด้วยคาบเวลาที่มีค่าความผันผวน(volatility) ต่ำ (และมีค่าความคลาดเคลื่อนที่มีขนาดเล็ก) สรุปได้ว่าค่าความแปรปรวนของค่าความคลาดเคลื่อนจากการถดถอยจะขึ้นอยู่กับค่าความผัน(volatility) ของความคลาดเคลื่อนในอดีตที่ผ่านมา (สธนพล วิเชียรรัตนพันธ์, 2547; อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2547)

ความเป็นไปได้ในการหาค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาไปพร้อมกันนั้น ในขั้นต้นการพยากรณ์อย่างมีเงื่อนไขจะมีความแม่นยำเหนือกว่าการพยากรณ์อย่างไม่มีเงื่อนไข ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) ซึ่งแสดงได้ดังนี้

$$X_t = a_0 + a_1 X_{t-1} + \varepsilon_1 \tag{29}$$

และการพยากรณ์อย่างมีเงื่อนไขของ $X_{t-1}$ ดังนี้คือ

$$E_t X_{t-1} = a_0 + a_1 X_t \tag{30}$$

และค่าเฉลี่ยแบบมีเงื่อนไขในการพยากรณ์ $X_{t-1}$ ค่าความคลาดเคลื่อนของความแปรปรวนอย่างมีเงื่อนไขที่พยากรณ์ได้ดังนี้คือ

$$E_t\left[(X_{t-1} - a_0 - a_1 X_t)^2\right] = E_t\, \varepsilon^2_{t-1} = \sigma^2 \tag{31}$$

ถ้าเปลี่ยนไปใช้การพยากรณ์อย่างมีเงื่อนไขแล้ว ผลที่จะใช้เป็นค่าเฉลี่ยในช่วง Long-Run ของลำดับ $\{X_t\}$ ซึ่งเท่ากับ $\dfrac{a_0}{(1-a_1)}$ จะได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไขดังนี้

$$E_t\left\{\left(X_{t-1} - \frac{a_0}{(1-a_1)}\right)^2\right\} = E\left[\left(\varepsilon_{t-1} + a_1 \varepsilon_1 + a_1^2 \varepsilon_{t-1} + a_1^3 \varepsilon_{t-2} + \ldots\right)^2\right] \tag{32}$$

เนื่องจาก $\frac{1}{(1-a_1)^2} > 1$ เพราะฉะนั้นความแปรปรวน(variance) จากการพยากรณ์แบบไม่มีเงื่อนไข (unconditional variance) จึงมีค่าสูงกว่าความแปรปรวนของ{ $\varepsilon_t$} ไม่เป็นค่าคงที่ จะสามารถประมาณค่าแนวโน้มของการเปลี่ยนแปลงความแปรปรวนโดยใช้ ARMA Model โดยให้ $\varepsilon_t$ แทนส่วนที่เหลือที่ได้จากการประมาณจากสมการ ดังนั้น ค่าความแปรปรวนแบบมีเงื่อนไขของ $X_{t-1}$ จะได้ดังนี้

$$\mathrm{Var}\,(X_{t+1} \mid X_t) \;=\; E\left[(X_{t+1} - a_0 - a_1 X_t)^2\right] \;=\; E_t\,\varepsilon^2_{t+1} \tag{33}$$

และจากที่ให้ $E_t\,\varepsilon^2_{t+1} = \sigma^2_{t+1}$ จึงแสดงว่าความแปรปรวนอย่างมีเงื่อนไขไม่ใช่ค่าคงที่และจะได้แบบจำลองในการประมาณค่าส่วนที่เหลือ (residuals) ออกมาดังนี้

$$\hat{\varepsilon}^2_t = \alpha_0 + \alpha_1 \hat{\varepsilon}^2_{t-1} + \ldots + \alpha_q \hat{\varepsilon}^2_{t-q} + V_t \tag{34}$$

เมื่อ $V_t$ = white noise process

ถ้าค่าของ $\alpha_1$, $\alpha_2$,..., $\alpha_q$ มีค่าเท่ากับศูนย์ ความแปรปรวนที่ประมาณค่ามาได้ (estimated variance) จะมีค่าคงที่หรือคงตัว (constant variance) $\alpha_0$ อีกนัยหนึ่ง คือ ค่าความแปรปรวนอย่างีเงื่อนไขของ $X_t$ จะมีการเปลี่ยนแปลงสอดคล้องกับในสมการ (34) และค่าพยากรณ์สามารถเขียนได้ดังนี้

$$E_t\,\hat{\varepsilon}^2_t = \alpha_0 + \alpha_1 \hat{\varepsilon}^2_{t-1} + \ldots + \alpha_q \hat{\varepsilon}^2_{t-1-q} \tag{35}$$

จากเหตุผลที่กล่าวมา สมการ(34) เรียกว่า Autoregressive Conditional Heteroscedastic (ARCH) Model และสมการ(35) เป็น ARCH(q) โดยค่า $E_t\,\hat{\varepsilon}^2_t$ หรือ $\sigma^2_{t-1}$ จะประกอบด้วย 2 องค์ประกอบ คือ ค่าคงที่และความผันผวน(volatility) ในคาบเวลาที่ผ่านมา ซึ่งเขียนได้เป็นส่วนเหลือกำลังสองของคาบเวลาในอดีต (ARCH term) ส่วนค่าสัมประสิทธิ์ ($\alpha_1$, $\alpha_2$,..., $\alpha_q$) สามารถหาค่าได้โดยวิธี Maximum likelihood

### 3.4.4 แบบจำลอง Generalized Autoregressive Conditional Heterosedasticity (GARCH)

แบบจำลอง ARCH ของ Engle ได้มีการพัฒนาต่อโดย Bollerslev ในปี 1986 ด้วยการให้ความแปรปรวนแบบมีเงื่อนไข (conditional variance) มีลักษณะเป็น ARMA process โดยที่ให้ error process มีลักษณะดังนี้

$$\varepsilon_1 = V_t \int \sigma_1^2 \tag{36}$$

โดยที่ความแปรปรวนของ $V_t = \sigma_1^2 = 1$ และ

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_1 \varepsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-1}^2 \tag{37}$$

เมื่อ { $V_t$ } คือ White noise process ที่เป็นค่าอิสระจากเหตุการณ์ในอดีต{ $\varepsilon_{t-1}$ } ค่าเฉลี่ยแบบมีเงื่อนไขของ $\varepsilon_t$ จะเท่ากับศูนย์ ประเด็นสำคัญในการหาความแปรปรวนแบบมีเงื่อนไขของ $\varepsilon_t$ ถูกกำหนดโดย

$$E_{t-1}\ \varepsilon_t^2 = \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_1 \varepsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-1}^2 \tag{38}$$

ดังนั้น ความแปรปรวนแบบมีเงื่อนไขของ $\varepsilon_t$ ถูกกำหนดโดย $\sigma_t^2$ ในสมการ (16) แบบจำลองนี้จึงเรียกว่า Generalized Autoregressive Conditional Heterosedasticity (GARCH) หรือ GARCH (p, q) นั้นใช้กระบวนการ Autoregressive และ Moving Averge ในการหาความแปรปรวนที่มีลักษณะ Heteroscedastic Variance จะเห็นว่าถ้า p=0 และ q=1 เป็น GARCH (0,1) หรือคือ ARCH(1) นั่นเอง โดยสรุปว่าถ้า $\beta_1$ ทั้งหมดมีค่าเป็นศูนย์แบบจำลอง GARCH (p, q) จะเทียบเท่ากับแบบจำลอง ARCH (q) คุณสมบัติที่สำคัญของแบบจำลอง GARCH คือค่าความแปรปรวนอย่างมีเงื่อนไขของ disturbances ของค่า $X_t$ สร้างขึ้นมาจากกระบวนการ ARMA จึงสามารถคาดได้ว่าส่วนเหลือจากการทำ ARMA จะแสดงถึงรูปแบบคุณลักษณะเดียวกันเช่น ถ้าการประมาณค่า { $X_t$ } ด้วยกระบวนการทำ ARMA ค่า Autocorrelation Function (ACF) ซึ่งเป็นสหสัมพันธ์ระหว่างตัวแปรสุ่มที่หน่วยเวลาห่างกันของกระบวนการเดียวกันและ Patial Autocorrelation Function (PACF) ของส่วนที่เหลือ (residuals) ควรจะบ่งถึงกระบวนการ White noise และ ACF ของกำลังสองของส่วนที่เหลือ (aquared residuals) นำมาช่วยในการระบุถึงลำดับ (order) ของกระบวนการ GARCH (สธนพล วิเชียรรัตนพันธ์, 2547; อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2547)

### 3.4.5 แบบจำลอง Exponencial GARCH (EGARCH)

แบบจำลอง GARCH ต่าง ๆ นอกจากใช้ได้อย่างประสบผลสำเร็จ แต่ก็มีข้อเสียอยู่สองประการในการประยุกต์ใช้กับการตั้งหรือคำนวณค่าทรัพย์สินประเภททุน ประการแรก คือ ในกระบวนการ GARCH แบบสมมาตรนั้น ถ้ามีความผิดปกติ หรือ shock เกิดขึ้นๆไม่ว่าในทางบวกหรือทางลบ แต่อยู่ในระดับหรือขนาดเดียวกัน ซึ่งให้ระดับความไม่แน่นอนที่เท่ากันแล้ว ค่าความแปรปรวนอย่างมีเงื่อนไขก็จะเพิ่มขึ้นในทางบวกหรือทางลบ แต่อยู่ในระดับหรือขนาดเดียวกัน ซึ่งให้ระดับความไม่แน่นอนที่เท่ากันแล้ว ค่าความแปรปรวนอย่างมีเงื่อนไขก็จะเพิ่มขึ้นในทางบวกหรือทางลบอย่างมากจนน่าตกใจ (Engle and Bollerslev, 1986) อย่างไรก็ตาม Black (1976) ได้พบความสัมพันธ์ที่เป็นลบหรือตรงกันข้ามกันระหว่างผลตอบแทนในปัจจุบันกับความไม่แน่นอนที่เกิดจากความอ่อนไหว (Volatility) ในอนาคต เช่น ความไม่แน่นอนมักจะสูงเมื่อมีข่าวร้าย และจะลดลงเมื่อมีข่าวดี (Nelson, Daniel B, 1991) ลักษณะความไม่สมมาตรของความแปรปรวนแบบมีเงื่อนไขนี้ ผู้เขียนงานวิชาการหลายคนเรียกว่า leverage effect คือ อิทธิพลจากค่ายกกำลัง ซึ่งแบบจำลอง GARCH แบบเส้นตรงไม่สามารถกำหนดความไม่แน่นอนที่อ่อนไหวในอนาคต กล่าวอีกนัยหนึ่งคือ เฉพาะขนาดของค่าความคลาดเคลื่อนจากประมาณการถดถอยโดยมีการทอดระยะเวลา (Lagged Residuals) เท่านั้นที่มีส่วนกำหนดค่าความแปรปรวนอย่างมีเงื่อนไข แต่ความเป็นบวกหรือลบของค่าความคลาดเคลื่อนไม่มีส่วนเกี่ยวข้อง (Nelson, Daniel B, 1991) ซึ่งข้อจำกัดนี่เป็นจุดสำคัญประการแรกที่ก่อให้มีการพัฒนาแบบจำลอง EGARCH

Nelson (1991) นำเสนอแบบจำลอง EGARCH หรือ Exponencial ARCH Model ที่มีค่ายกกำลังสูง(Exponential) เพื่อแก้ไขข้อจำกัดที่ปรากฏในแบบจำลอง GARCH (1,1) ให้ความแปรปรวนอย่างมีเงื่อนไข (นนทพร จำปาวัน, 2550)

สมการความแปรปรวน

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \alpha \left[ \frac{|\varepsilon_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - \sqrt{\frac{2}{\pi}} \right] \qquad (40)$$

### 3.4.6 แบบจำลอง GARCH-in-Mean (GARCH-M)

จากแบบจำลอง GARCH (p,q) ในสมการที่ (38) ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนที่เหลือกำลังสองของกระบวนการนี้ (Bollerslev, 1986) Engle (1987) ขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักกันในชื่อ GARCH-in-Mean หรือ GARCH-M (ภัทร์ ตั้งตระกูล, 2546)

สมการความแปรปรวน

$$\sigma_{t}^{2} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^{2} + \beta\sigma_{t-1}^{2} \qquad (41)$$

อิทธิพลอย่างมีนัยสำคัญของความผันผวนในผลตอบแทนของหลักทรัพย์ถูกตรวจจับด้วยค่าสัมประสิทธิ์ $\sigma(\beta)$ ในสมการที่ (41) ซึ่งอธิบายแทนดัชนีของความสัมพันธ์ของการหลีกเลี่ยง ความเสี่ยง ค่าสัมประสิทธิ์ $\beta$ ที่เป็นบวกอย่างมีนัยสำคัญบอกถึงผู้ลงทุนในหลักทรัพย์เมื่อเผชิญกับระดับความเสี่ยงที่สูงขึ้นก็ต้องการค่าชดเชยความเสี่ยงที่มากขึ้นตามไปด้วย

## 3.5 วิธีการพยากรณ์

ประโยชน์ของข้อมูลอนุกรมเวลาที่สำคัญ ก็คือการนำข้อมูลดังกล่าวมาใช้วิเคราะห์เพื่อพยากรณ์ค่าอนุกรมเวลาในอนาคตสามารถทำได้หลายวิธีเช่น การวิเคราะห์แนวโน้ม (Time trend) วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing), วิธีถดถอยเชิงพหุ (Multiple regression) และวิธีบอกซ์-เจนกินส์ (Box-jenkins)

### 3.5.1 วิธีการวิเคราะห์แนวโน้ม (Time trend)

แบบจำลองที่ใช้การวิเคราะห์แนวโน้มโดยทั่วไปมักจะใช้สมการแนวโน้มที่ประมาณค่าสัมประสิทธิ์ด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Square: OLS) ซึ่งเป็นวิธีการที่ทำให้ผลรวมกำลังสองของผลต่างระหว่างค่าแนวโน้มกับข้อมูลที่ใช้มีค่าน้อยที่สุด (Least-square Error) วิธีการนี้เป็นวิธีการหนึ่งที่นิยมนำมาใช้ในการพยากรณ์แนวโน้ม เนื่องจากเป็นวิธีการที่ง่ายในการคำนวณ และสามารถสร้างสมการพยากรณ์ได้หลายรูปแบบทั้งที่เป็นเส้นตรงหรือไม่ใช่เส้นตรงก็ได้ หรือสามารถประยุกต์ใช้วัดอิทธิพลของฤดูกาลได้ด้วย สำหรับวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีการกำลังสองน้อยที่สุด (OLS) สามารถทำได้ดังนี้

$$\hat{y} = \alpha + \beta t \qquad (42)$$

โดยที่ $\hat{y}$ คือ ค่าพยากรณ์ y

$\alpha, \beta$ คือ ค่าพารามิเตอร์

t คือ เวลา โดยที่ t = 1, 2,…, n

สมการข้างต้นเป็นสมการแนวโน้มเส้นตรง แต่ในบางกรณีเส้นสมการแนวโน้มอาจเป็นเส้นโค้งก็ได้ ดังนั้นการใช้สมการแนวโน้มเส้นตรงในการพยากรณ์อาจทำให้เกิดความคลาดเคลื่อน รูปแบบสมการแนวโน้มที่ไม่เป็นเส้นตรงที่มักจะพบอยู่เสมอ ได้แก่

1) แนวโน้มโพลิโนเมียล (Polynomial trend)

$$\hat{y} = \alpha_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \ \ldots + \beta_n t^n \tag{43}$$

2) แนวโน้มโพลิโนเมียลที่มักพบอยู่เสมอ ก็คือแนวโน้มพาราโบลา (Parabola trend) ดังนี้

$$\hat{y} = \alpha_0 + \beta_1 t + \beta_2 t^2 \tag{44}$$

3) แนวโน้มเอ็กซ์โปเนนเชียล (Exponential trend)

$$\hat{y} = \alpha_0 + \beta^{t} \tag{45}$$

4) แนวโน้มเอ็กซ์โปเนนเชียลลำดับที่สอง (Second exponential trend)

$$\hat{y} = \alpha_0 + \beta_1^{t} \beta_2^{t2} \tag{46}$$

การเลือกรูปแบบของสมการแนวโน้มแบบใดนั้น ขึ้นอยู่กับลักษณะการกระจายของข้อมูลอนุกรมเวลา ดังนั้นก่อนการสร้างสมการพยากรณ์จะต้องทำการทดสอบหรือตรวจสอบว่าตัวแปรหรือข้อมูลที่ต้องการพยากรณ์นั้นมีลักษณะของการกระจาย และมีเส้นกราฟแนวโน้มประเภทใด แล้วเลือกรูปแบบของสมการพยากรณ์ให้สอดคล้องกับรูปแบบการกระจายและเส้นแนวโน้ม เพื่อให้สมการและค่าพยากรณ์ที่ได้จากการคำนวณมีความแม่นยำและคลาดเคลื่อนน้อยที่สุด

### 3.5.2 วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing)

**วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Exponential smoothing)** เป็นวิธีพยากรณ์ที่เหมาะสมกับการพยากรณ์ในระยะสั้นและปานกลาง วิธีการนี้ได้ให้ความสำคัญกับข้อมูลล่าสุด และความสำคัญของข้อมูลที่ห่างออกไปจะลดลง วิธีการ Exponential smoothing ที่นิยมในปัจจุบันมีอยู่ 3 วิธี ดังนี้

1) Single Exponential Smoothing (SES)

เป็นวิธีการทำให้เรียบอย่างง่าย โดยใช้การหาค่าเฉลี่ยแบบถ่วงน้ำหนัก และสมมติให้ค่าน้ำหนักหรือค่าความสำคัญของข้อมูล คือ $\alpha$ (Alpha) วิธีการนี้มีเงื่อนไขว่า ข้อมูล

อนุกรมเวลาที่ใช้ในการวิเคราะห์จะต้องไม่มีแนวโน้ม (Trend) และอิทธิพลของฤดูกาล (Seasonality) หมายความว่า ข้อมูลอนุกรมเวลาที่นำมาใช้การวิเคราะห์ต้องมีลักษณะคงที่ สำหรับสมการที่ใช้ในการพยากรณ์สามารถแสดงได้ดังนี้

$$\hat{y}_t = \alpha y_1 + (1+\alpha)\hat{y}_{t-1} \qquad (47)$$

หรือ

$$\hat{y}_t = \alpha \sum_{n=0}^{t-1} (1+\alpha)^n y_{t-n} \qquad (48)$$

โดยที่ $y_t$ = ข้อมูล ณ เวลาที่ t; t = 1, 2,…, n

$\alpha$ = ค่าถ่วงน้ำหนักความสำคัญที่ให้แก่ข้อมูล ณ เวลาที่ t ($0 \leq \alpha \leq 1$)

$\hat{y}_t$ = ค่าประมาณหรือพยากรณ์ของข้อมูล ณ เวลาที่ t

$\hat{y}_{t-1}$ = ค่าประมาณหรือค่าพยากรณ์ของข้อมูล ณ เวลา t-1

2) Holt's Two-Parameter Method

วิธี Holt's Two-Parameter Method เป็นวิธีที่ Holt (1957) ได้ปรับปรุงวิธี Single Exponential Smoothing (SES) ขึ้นใหม่เพื่อให้สามารถใช้ข้อมูลอนุกรมเวลาที่มีแนวโน้ม (trend) ของเวลา เรียกวิธีการนี้ว่า "Holt's Two-Parameter Method" วิธีการนี้ให้ความสำคัญกับข้อมูลล่าสุดและแนวโน้มเวลา ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 2 ค่า คือ $\alpha$ (Alpha) และ $\beta$ (Beta) โดยมีสมการที่ใช้ในการพยากรณ์ดังนี้

$$\hat{y}_{t+k} = \alpha + bK \quad ; \hat{y}_{t+k} = \text{ค่าพยากรณ์ ณ เวลาที่ } t+k \qquad (49)$$

ค่า a และ b คำนวณจาก

$$a_t = \alpha_{yt} + (1+\alpha)(a_{t-1} + b_{t-1}) \qquad (50)$$

$$b_t = \beta(a_t - a_{t-1}) + 1 - Bb_{t-1} \qquad (51)$$

โดยที่

$\alpha$ = ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับพยากรณ์ ($0 \leq \alpha \leq 1$)

$\beta$ = ค่าคงที่ที่ทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม ($0 \leq \beta \leq 1$)

3) Holt-Winters-Trend and Seasonal (Three-Parameter)

วิธีนี้ถูกพัฒนาเพิ่มขึ้นจากวิธีการของ Holt โดย Winters (1960) ได้พัฒนาให้วิธีการนี้สามารถวิเคราะห์ข้อมูลอนุกรมเวลาที่มีทั้งแนวโน้ม (Trend) และฤดูกาล (Seasonality) วิธีการนี้ให้ความสำคัญกับข้อมูลล่าสุดแนวโน้มเวลาและฤดูกาล ดังนั้นจึงมีค่าคงที่ในการทำให้เรียบ 3 ค่า คือ $\alpha$ (Alpha), $\beta$ (Beta) และ $\gamma$ (Gamma) โดยมีสมการที่ใช้ในการพยากรณ์ดังนี้

Holt-Wintes-Mulitplicative

$$\hat{y}_{t+k} = (a+bk)c_{t+k} \quad ; \hat{y}_{t+k} = \text{ค่าพยากรณ์ ณ เวลาที่ } t+k \tag{52}$$

ค่า a,b และ c คำนวณจาก

$$a_{t} = \alpha \frac{y_t}{c_{t-s}} + (1+\alpha)(a_{t-1} + b_{t-1}) \tag{53}$$

$$b_{t} = \beta(a_t - a_{t-1}) + (1-B)b_{t-1} \tag{54}$$

$$c_{t} = \gamma \frac{y_t}{a_t} + (1-\gamma)c_{t-1} \tag{55}$$

โดยที่ $\alpha$ = ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับพยากรณ์ ($0 \leq \alpha \leq 1$)

$\beta$ = ค่าคงที่ที่ทำให้เรียบระหว่างแนวโน้มจริงกับค่าประมาณของแนวโน้ม ($0 \leq \beta \leq 1$)

$\gamma$ = ค่าคงที่ที่ทำให้เรียบระหว่างฤดูกาลจริงกับค่าประมาณของฤดูกาล ($0 \leq \gamma \leq 1$)

Holt-Wintes-Additive

$$\hat{y}_{t+k} = (a+bk)c_{t+k} \quad ; \hat{y}_{t+k} = \text{ค่าพยากรณ์ ณ เวลาที่ } t+k \tag{56}$$

ค่า a,b และ c คำนวณจาก

$$a_{t} = \alpha(y_{t-1} - c_{t-1}) + (1-\alpha)(a_{t-1} + b_{t-1}) \tag{57}$$

$$b_{t} = \beta(a_t - a_{t-1}) + 1 - Bb_{t-1} \tag{58}$$

$$c_{t} = \gamma(y_t - a_{t+1}) - \gamma c_{t-1} \tag{59}$$

โดยที่ $\alpha$ = ค่าคงที่ที่ทำให้เรียบระหว่างข้อมูลกับพยากรณ์ ($0 \leq \alpha \leq 1$)

$\beta$ = ค่าคงที่ที่ทำให้เรียบระหว่างแน้วโน้มจริงกับค่าประมาณของแนวโน้ม ($0 \leq \beta \leq 1$)

$\gamma$ = ค่าคงที่ที่ทำให้เรียบระหว่างฤดูกาลจริงกับค่าประมาณของฤดูกาล ($0 \leq \gamma \leq 1$)

## 3.6 การประเมินค่าการพยากรณ์

การศึกษาครั้งนี้ใช้ค่าสถิติ RMSE (Root Mean Squared Error) และ MAPE (Mean Absolute Percentage Error) ในการประเมินค่าการพยากรณ์ เพื่อทดสอบความแม่นยำของผลการพยากรณ์

**Root Mean Squared Error (RMSE)**

$$\text{RMSE} = \sqrt{\sum_{t=T+1}^{T+h} \left(\hat{Y}_t - Y_t\right)^2 / h} \tag{60}$$

**Mean Absolute Percentage Error (MAPE)**

$$\text{MAPE} = 100 \sum_{t=T+1}^{T+h} \left| \frac{\hat{Y}_t - Y_t}{Y_t} \right| / h \tag{61}$$