บทที่ 11
การวิเคราะห์อนุกรมเวลา
(Time Series Analysis)

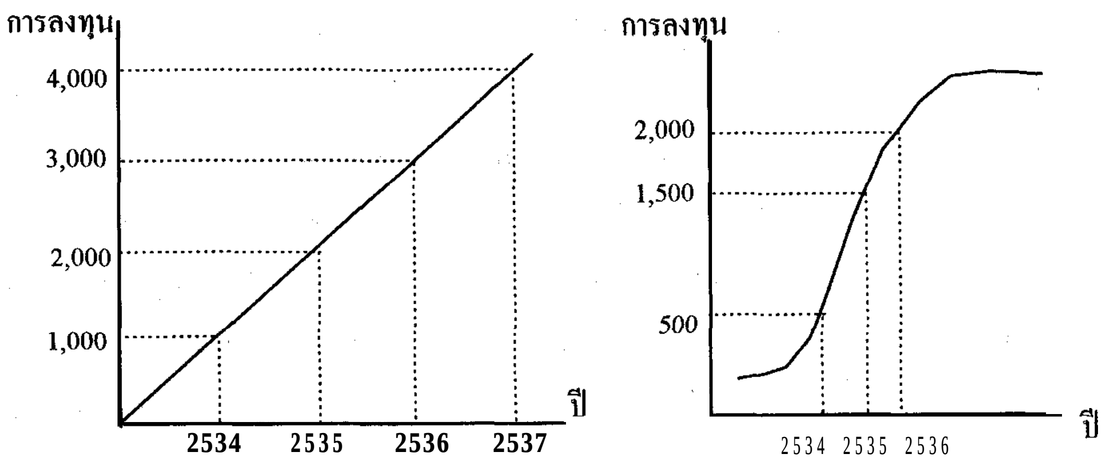
ข้อมูลอนุกรมเวลา (time series data) หมายถึง ข้อมูลที่เกิดขึ้นตามลำดับเวลาโดยมีช่วงห่างของเวลาต่างๆ กัน หรือเป็นข้อมูลที่แสดงการเปลี่ยนแปลงไปตามลำดับเวลาที่ต่อเนื่องกัน โดยที่หน่วยของเวลาอาจเป็น ชั่วโมง วัน สัปดาห์ เดือน ปี เช่น รายได้ประชาชาติรายปี การลงทุนรายปี การใช้จ่ายเพื่อการบริโภครายเดือน การวิเคราะห์อนุกรมเวลาเป็นการศึกษาหารูปแบบการเปลี่ยนแปลงของตัวแปร ซึ่งมีการเปลี่ยนแปลงไปตามกาลเวลาตั้งแต่อดีตจนถึงปัจจุบัน เพื่อนำรูปแบบมาใช้วิเคราะห์หาค่าพยากรณ์ ของตัวแปรนั้นๆ การพยากรณ์สามารถหาได้โดยการวิเคราะห์หาค่าพยากรณ์ของตัวแปรนั้นๆ การพยากรณ์สามารถหาได้โดย การวิเคราะห์ความถดถอย (regression analysis) เป็นการพยากรณ์ที่ใช้แสดงความสัมพันธ์ระหว่างตัวแปรที่ต้องการพยากรณ์กับตัวแปรตัวอื่น เช่น $C = a+bY+dW$ ดังรายละเอียดที่ได้อธิบายในบทที่ 10 และการวิเคราะห์อนุกรมเวลา (time - series analysis) เป็นการพยากรณ์ค่าในอนาคตของตัวแปรตัวใดตัวหนึ่ง โดยอาศัยรูปแบบการเปลี่ยนแปลงตามลำดับเวลาในอดีตของตัวแปรตัวนั้น โดยไม่สนใจความสัมพันธ์ระหว่างตัวแปรนั้นกับตัวแปรตัวอื่นๆ เลย เทคนิคการวิเคราะห์อนุกรมเวลามีหลายวิธี ผู้วิเคราะห์จะต้องเลือกใช้เทคนิคที่เหมาะสมต่อสิ่งที่ต้องการพยากรณ์ ในบทนี้จะได้อธิบายถึง

1. องค์ประกอบของอนุกรมเวลา (Components of a time - series)
2. การพยากรณ์ด้วยวิธี (Exponential Smoothing)
3. การพยากรณ์ด้วย (Trend projection)
4. การพยากรณ์ด้วยวิธีแยกส่วนประกอบ (Decomposition method)

11.1 องค์ประกอบของอนุกรมเวลา

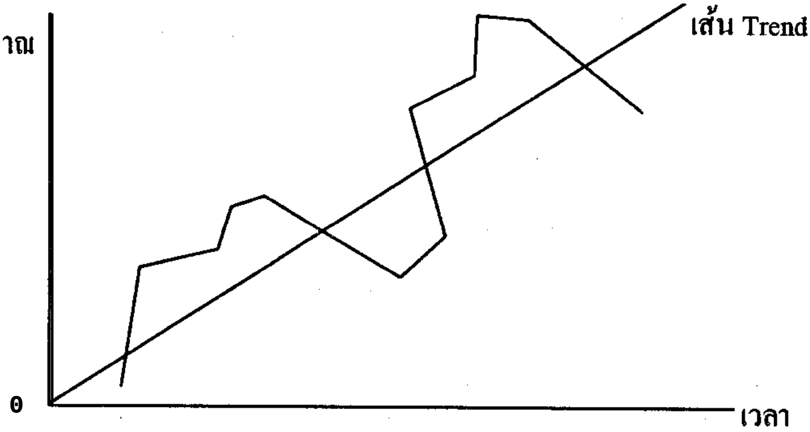
ข้อมูลอนุกรมเวลามีองค์ประกอบที่แบ่งได้เป็น 4 ส่วนคือ

11.1.1 ค่าแนวโน้ม (trend) ใช้แทนด้วยตัว "T" หมายยถึง ข้อมูลซึ่งมีการเปลี่ยนแปลงเป็นเวลายาวนานพอที่จะเห็นแนวโน้มของการเปลี่ยนแปลงของข้อมูลในอนาคต ปกติระยะเวลาที่จะสามารถเห็นแนวโน้มการเปลี่ยนแปลงมากกว่า 1 ปี ลักษณะของเส้นแนวโน้มอาจจะเป็นเส้นตรง หรือเส้นโค้งก็ได้ ดังแสดงในรูป 11.1 (ก) และ (ข)



รูปที่ 11.1 แสดงเส้นแนวโน้ม

11.1.2 องค์ประกอบที่เป็นวัฏจักร (Cyclical component) ใช้แทนด้วย "C" หมายถึงข้อมูลที่มีการเปลี่ยนแปลงขึ้น - ลงตามลำดับเวลารอบๆ เส้นแนวโน้ม ในช่วงเวลาที่ยาวนานกว่า 1 ปี ช่วงเวลาของคลื่นการเปลี่ยนแปลงแต่ละอันอาจใช้เวลา 1 ปี 2 ปี 3 ปี 4 ปี ฯลฯ ก็ได้ โดยไม่มีกำหนดเวลา เช่น เศรษฐกิจเจริญรุ่งเรือง 3 ปี เศรษฐกิจตกต่ำ 4 ปี ระยะฟื้นฟู 2 ปี และระยะถดถอย 5 ปี



รูปที่ 11.2 แสดงองค์ประกอบวัฏจักร

11.1.3 องค์ประกอบที่เป็นการผันแปรตามฤดูกาล (Seasonal variation component) ใช้แทนด้วย "S" หมายถึงข้อมูลที่เปลี่ยนแปลงขึ้นลงตามฤดูกาล ซึ่งจะเปลี่ยนแปลงคล้ายๆ กันในช่วงเวลาเดียวกันของแต่ละปี เวลาของแต่ละฤดูกาล อาจเป็น 1 เดือน 2 เดือน 3 เดือน หรือ 6 เดือน หรืออาจเป็นชั่วโมงในแต่ละวัน หรือเป็นวันในแต่ละเดือน เช่น ข้อมูลการจราจรติดขัด ในแต่ละวัน ข้อมูลการขายของเทศกาลปีใหม่ ข้อมูลปริมาณฝนตกในเดือนพฤษภาคม ความแตกต่างระหว่างการผันแปรตามฤดูกาลกับการผันแปรวัฏจักร คือ

ก. ระยะเวลาของการเกิดความผันแปรตามวัฏจักร 1 รอบ จะยาวนานกว่าความผันแปรตามฤดูกาล

ข. ระยะเวลาการเกิดความผันแปรตามฤดูกาลค่อนข้างแน่นอนกว่า ช่วงเวลาของการเกิดความผันแปรเชิงวัฏจักร

9.1.4 องค์ประกอบที่เป็นความผันแปรที่ไม่สม่ำเสมอ (Irregular variation) ใช้แทนด้วย "I" หมายถึง การผันแปรของข้อมูลอย่างไม่มีรูปแบบที่แน่นอน ดังนั้น การพยากรณ์ที่ใช้ข้อมูลที่มีความผันแปรที่ไม่สม่ำเสมอ จะให้ผลที่ไม่น่าเชื่อถือ ตัวอย่างปัจจัยที่ทำให้เกิดการผันแปรที่ไม่สม่ำเสมอหรือไม่แน่นอน เช่น สงคราม น้ำท่วม แผ่นดินไหว การปฏิวัติรัฐประหาร

ข้อมูลแต่ละตัวจะมีองค์ประกอบทั้ง 4 สามารถเขียนรูปองค์ประกอบได้ 2 ลักษณะ คือ

$$Y_t = T_t + C_t + S_t + I_t \qquad \text{.......} \qquad (11.1)$$

หรือ

$$Y_t = T_t . C_t . S_t . I_t \qquad \text{.......} \qquad (11.2)$$

โดยองค์ประกอบทั้ง 4 เป็นเสมือนสภาพแวดล้อมที่ทำให้เกิดข้อมูลดังกล่าว

11.2 เทคนิคการพยากรณ์ด้วยวิธีหาค่าเฉลี่ยเคลื่อนที่ (The moving average)

เป็นเทคนิคการพยากรณ์ที่ใช้อนุกรมเวลา เพื่อพยากรณ์ในระยะสั้นและระยะปานกลาง เป็นการขจัดอิทธิพลของความผันแปรที่ไม่สม่ำเสมอหรือไม่แน่นอนออกไป การหาค่าเฉลี่ยเคลื่อนที่เป็นวิธีการหาค่าเฉลี่ยโดยการทิ้งข้อมูลเก่าที่สุดออกไปทีละตัว และรวมข้อมูลใหม่สุดเข้ามาแทนที่ทีละตัว ค่าเฉลี่ยที่ได้จะใช้พยากรณ์สำหรับระยะเวลาถัดไป

$$Y_{T+1} = M_T = \frac{(X_T + X_{T-1} + X_{T-2} + \ldots + X_{T-n+1})}{n}$$

M_T	=	moving average ณ เวลา T
Y_{T+1}	=	มูลค่าที่ได้พยากรณ์สำหรับเวลาถัดไป
X_T	=	ค่าที่แท้จริง ณ เวลา T
n	=	จำนวนข้อมูลในการหาค่าเฉลี่ยเคลื่อนที่
X_{T-i}	=	ค่าที่แท้ จริง ณ เวลาก่อยปีที่ T ไป i ปี i = 1, 2, ...

ตารางที่ 11.1 แสดงการหาค่าเฉลี่ยเคลื่อนที่

Year	Quarter	T	Sales X_T	M Moving Total	Moving - Average Forcast F_T+1	e_T
2531	1	1	500			
	2	2	450			
	3	3	300			
	4	4	650	1900		
2532	1	5	700	2100	475	-125
	2	6	550	2200	525	-25
	3	7	450	2250	550	+100
	4	8	600	2300	562.5	-37.5
2533	1	9	800	2400	575	-22.5
	2	10	650	2500	600	-50
	3	11	400	2550	625	+225
	4	12	750	2600	635.5	-114.5
2534	1	13			650	
	2	14				
	3	15				
	4	16				

การหาค่าเฉลี่ยเคลื่อนที่ของไตรมาสแรกของปี 1988 คือ

$$M = \frac{X_4 + X_3 + X_2 + X_1}{4} = 1900$$

$$F_{4+1} = \frac{M_4}{4} = \frac{x_, + x_, + x_, + x_,}{4} = \frac{1900}{4} = 475$$

$$F_6 = \frac{M_5}{4} = \frac{X_4 + X_3 + X_2 + X_1}{4} = \frac{2100}{4} = 525$$

ค่าความคลาดเคลื่อนของการพยากรณ์ คือ $e_5 = F_5 - X_5$

$= 475 - 700$

$= -225$

11.3 การพยากรณ์ด้วย Exponential Smoothing Method

การพยากรณ์ด้วยวิธีนี้ใช้เพื่อการพยากรณ์ระยะสั้น และระยะปานกลาง โดยมีวัตถุประสงค์เหมือนกรณีค่าเฉลี่ยเคลื่อนที่ คือขจัดอิทธิพลของความไม่สม่ำเสมอหรือความไม่แน่นอนออกไป วิธีนี้ไม่ได้ให้น้ำหนักของข้อมูลทุกตัวไม่เท่ากัน

แบบจำลองของ Exponentail smoothing คือ

$F_{t+1} = \propto Y_t + (1 - \propto) F_t$ .... (11.4)

F_{t+1} = ค่าพยากรณ์ในปีที่ $t + 1$

Y_t = ค่าสังเกตของข้อมูลอนุกรมเวลาในปีที่ t

F_t = ค่าพยากรณ์ของปีที่ t

$\propto$ = ค่าถ่วงน้ำหนัก มีค่าคงที่ $0 \leq \propto \leq 1$
จะมีค่าเท่าไรต้องสมมติขึ้นเอง

ตัวอย่าง $F_2 = \propto Y_1 + (1 - \propto) F_1$ .... (11.5)

$F_3 = \propto Y_2 + (1 - \propto) F_2$ .... (11.6)

ปรับรูปสมการใหม่ได้

$F_{t+1} = \propto Y_t + F_t - \propto F_t$

$= F_t + \propto(Y_t - F_t)$ .... (11.7)

$Y_t - F_t$ = ค่าคลาดเคลื่อนของการพยากรณ์ในปีที่ t

ตารางที่ 11.2 แสดงการขายน้ำมันต่อสัปดาห์ ใน 12 สัปดาห์

สัปดาห์	การขาย (1000 แกลลอน)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22

สมมติ $\alpha = 0.2$

$$F_3 = 0.2(Y_2) + 0.8(F_2) = 0.2(21) + 0.8(17) = 17.8$$

$$F_4 = 0.2(Y_3) + 0.8(F_3) = 0.2(19) + 0.8(17.8) = 18.04$$

$$F_{13} = 0.2(Y_{12}) + 0.8(F_{12}) = 0.2(22) + 0.8(18.48) = 19.18$$

ในการหาค่าพยากรณ์ ผู้พยากรณ์จะต้องสมมติค่า α หลายค่า แล้วหาค่าพยากรณ์เมื่อ α มีค่าต่างๆ ตามที่สมมติ หลังจากนั้นให้หาค่า mean square error (MSE) ซึ่งหาได้โดยการหารค่า sum square error ด้วย จำนวนค่าสังเกตลบหนึ่ง แล้วเปรียบเทียบค่า MSE ที่หาได้ ค่า MSE ที่มีค่าต่ำสุดเราจะใช้ผลของการพยากรณ์ชุดนั้นเพื่อการพยากรณ์

ตารางที่ 11.3 แสดงการหาค่า MSE เมื่อ $\alpha = 0.2$

สัปดาห์ (t)	การขายทั่วไป (Y_t)	ค่าพยากรณ์ (E_t)	ค่าพยากรณ์คลาดเคลื่อน (Y_t-F_t)	(ค่าคลาดเคลื่อน)^2 $(Y_t - F_t)^2$
1	17			
2	21	17.00	4.00	16.00
3	19	17.80	1.20	1.44
4	23	18.04	4.96	24.60
5	18	19.03	- 1.03	1.06
6	16	18.83	- 2.83	8.01
7	20	18.26	1.74	3.03
8	18	18.61	- 0.61	0.37
9	22	18.49	3.51	12.32
10	20	19.19	0.81	0.66
11	15	19.35	- 4.35	18.92
12	22	18.48	3.52	12.39
				SSE = 98.80

$$MSE = SSE / (n - 1) = 98.80 / 11 = 8.98$$

ตารางที่ 11.4 แสดงการหาค่า MSE เมื่อ $\propto = 0.3$

สัปดาห์ (t)	การขายทั่วไป (Y_t)	ค่าพยากรณ์ (E_t)	ค่าพยากรณ์คลาดเคลื่อน (Y_t-F_t)	(ค่าคลาดเคลื่อน)^2 (Y_t - F_t)^2
1	17			16.00
2	21	17.00	4.00	0.64
3	19	18.20	0.80	20.79
4	23	18.44	4.56	3.28
5	18	19.81	-1.81	10.69
6	16	19.27	-3.27	2.92
7	20	18.29	1.71	0.64
8	18	18.80	-0.80	11.83
9	22	18.56	3.44	0.17
10	20	19.59	0.41	22.18
11	15	19.71	-4.71	13.69
12	22	18.30	3.70	
				SSE = 102.83

$$MSE = SSE/(n-1) = 102.83/11 = 9.35 \quad \text{.....................} (11.8)$$

ค่า MSE เมื่อ $\propto = 0.2$ จะมีค่าต่ำสุด ดังนี้ จะให้ค่าพยากรณ์ที่แน่นยำมากกว่าเมื่อใช้ $\propto = 0.3$ ดังนั้น เราจะใช้ค่าพยากรณ์ เมื่อ $\propto = 0.2$

11..4 การพยากรณ์ ด้วยวิธี Trend projection

รูปสมการ

$$T_t = b_0 + b_1 t \quad \text{..} (11.9)$$

T_t = ค่าพยากรณ์

b_0 = จุดตัดแกนตั้งของเส้นแนวโน้ม

b_1 = สโลปของเส้นแนวโน้ม

t = เวลา

สูตรการหาค่า b_1

$$b_1 = \frac{\sum tY_t - \sum t \sum Y_t / n}{\sum t^2 - (\sum t)^2 / n} \qquad \text{(11.10)}$$

$$b_0 = \overline{Y} - b_1 \overline{t} \qquad \text{(11.11)}$$

$\overline{t}$ = ค่าเฉลี่ยของ t $= \frac{\sum t}{n}$

$\overline{Y}$ = ค่าเฉลี่ยของ Y $= \frac{\sum Y}{n}$

11.5 วิธีแยกส่วนประกอบ (Decomposition method)

เป็นการพยากรณ์โดยใช้ข้อมูลอนุกรมเวลาเมื่อพยากรณ์ระยะปานกลาง โดยการแยกส่วนประกอบของข้อมูลออกเป็น 4 ส่วน คือ แนวโน้มความผันแปรตามวัฏจักร ความผันแปรตามฤดูกาล และความผันแปรที่ไม่แน่นอนหรือไม่สม่ำเสมอ ทั้งนี้เพื่อพิจารณาว่าส่วนประกอบใดมีอิทธิพลต่อการเปลี่ยนแปลงของข้อมูลบ้าง แล้วคำนวณหาอิทธิพลของส่วนประกอบแต่ละอัน นำค่าองค์ประกอบแต่ละอันไปพยากรณ์ ค่าในอนาคตของข้อมูลนั้นๆ โดยสมมติว่า ปัจจัยหรือสภาพแวดล้อมที่เกี่ยวข้องกับข้อมูลอนุกรมเวลาในอนาคตเหมือนกับในอดีต สมมติว่าตัวแปรหรือข้อมูลอนุกรมเวลามีองค์ประกอบเป็นแบบคูณ

$$Y_t = T_t \cdot C_t \cdot S_t \cdot I_t \quad ; \quad t = 1, 2, \ldots \qquad \text{(11.12)}$$

ขั้นตอนของการแยกส่วนประกอบ มีดังนี้

1. หาค่าความผันแปรตามฤดูกาล (S) โดย

- หาค่าเฉลี่ยเคลื่อนที่ (moving average = MA) ค่าเฉลี่ยเคลื่อนที่เป็นค่าที่ขจัดอิทธิพลของฤดูกาลและความผันแปรความไม่แน่นอนหรือไม่สม่ำเสมอ (I) องค์ประกอบที่เหลืออยู่ในค่าเฉลี่ยเคลื่อนที่ คือ ค่าแนวโน้มและการผันแปรตามวัฏจักร

$$MA = T \times C \quad \text{.......................} \quad (11.13)$$

- หา S x I โดยนำค่า MA ไปหารข้อมูลดิบ (Y)

$$S.I = \frac{Y}{MA} = \frac{T x C x S x I}{T x C} = S x I \quad \text{.......................} \quad (11.14)$$

- หาค่า S โดยขจัดค่า I ออกไป จาก S x I โดยการหาค่าเฉลี่ยของ S x I การหาค่าเฉลี่ยของ S x I โดยตัดค่าสูงสุด และค่าต่ำสุดออกไป ค่าเฉลี่ยของ S x I หรือค่า S

2. หาค่าแนวโน้ม โดยใช้สมการ

$$T = a + bt \quad \text{..........................} \quad (11.15)$$

3. หาค่าความผันแปรตามวัฏจักร (C) โดยหารค่าเฉลี่ยเคลื่อนที่ด้วย แนวโน้ม

$$C = \frac{MA}{T} = \frac{T x C}{T} = c \quad . \quad . \quad . \quad . \quad . \quad (11.16)$$

4. หาค่าความผันนแปรที่ไม่สม่ำเสมอ (I) โดยการหารค่า Y ด้วยผลคูณของ แนวโน้ม ความผันแปรตามวัฏจักรและความผันแปรตามฤดูกาล

$$I = \frac{Y}{TxCxS} = \frac{TxCxSxI}{TxCxS} = I \quad \text{........................} \quad (11.17)$$

5. หาค่าประมาณของค่า Y ในอนาคตได้ โดยใช้สมการ

$$Y = T x C x S \quad \text{..........................} \quad (11.18)$$

ประโยชน์ของการแยกส่วนประกอบของข้อมูลอนุกรมเวลา คือ

1. ทำให้ทราบว่าปัจจัยใดมีอิทธิพลต่อการเปลี่ยนแปลงมากที่สุด
2. นำองค์ประกอบแต่ละตัวไปพยากรณ์ค่าของตัวแปรระยะต่างๆ เช่น

"T" ใช้พยากรณ์ทิศทางการเปลี่ยนแปลงของตัวแปรว่า จะเพิ่มหรือลดลงในอัตราเท่าใด และใช้เพื่อการพยากรณ์ระยะยาว เช่นการวางแผนการส่งออก

"S" ใช้พยากรณ์ระยะสั้น สามารถใช้พิจารณาจากฤดูกาลได้ว่า ฤดูกาลใดมีอิทธิพลต่อยอดขายมาก

ข้อจำกัดของวิธีนี้คือ

1. ต้องใช้ข้อมูลจำนวนมาก

2. สมมติฐานที่การเปลี่ยนแปลงข้อมูลในอนาคตเหมือนกับปัจจุบัน เป็นการสมมติที่ผิดจากความเป็นจริง

3. วิธีทำค่อนข้างยุ่งยาก

ตารางที่ 11.5 แสดงการหาค่าเฉลี่ยคลาดที่ครั้งละ 4 ไตรมาส

ปีที่	ไตรมาสที่	ปริมาณการส่งออก	ค่าเฉลี่ยเคลื่อนที่ MA	ค่าเฉลี่ยของค่าเฉลี่ยเคลื่อนที่
1	1	4.8		
	2	4.1		
	3	6.0	5.350	5.475
	4	6.5	5.600	5.738
2	1	5.8	5.875	5.975
	2	5.2	6.075	6.188
	3	6.8	6.300	6.325
	4	7.4	6.350	6.400
3	1	6.0	6.450	6.538
	2	5.6	6.625	6.675
	3	7.5	6.725	6.763
	4	7.8	6.800	6.838
4	1	6.3	6.875	6.938
	2	5.9	7.000	7.075
	3	8.0	7.150	
	4	8.4		

ที่มา : Anderson, David R., Sweeney, Dennis J., and Williams, Thomas, A., "Statistic for Business and Economics". 1994, p.692.

ตารางที่ 11.6 แสดงการหาค่าองค์ประกอบตามฤดูกาลและองค์ประกอบที่ไม่สม่ำเสมอ

(Seasonal and Irregular Components)

ปีที่ (1)	ไตรมาสที่ (2)	ปริมาณการส่งออก (3)	ค่าเฉลี่ยเคลื่อนที่ MA (4)	องค์ประกอบ S x T (5) = (3) / (4)
1	1	4.8		
	2	4.1		
	3	6.0	5.475	1.096
	4	6.5	5.738	1.133
2	1	5.8	5.975	0.971
	2	5.2	6.188	0.840
	3	6.8	6.325	1.075
	4	7.4	6.400	1.156
3	1	6.0	6.538	0.918
	2	5.6	6.675	0.839
	3	7.5	6.763	1.109
	4	7.8	6.838	1.141
4	1	6.3	6.938	0.908
	2	5.9	7.075	0.834
	3	8.0		
	4	8.4		

การหา seasonal index ก็เพื่อใช้ขจัดผลกระทบที่เกิดจากฤดูกาลออกไป

ตารางที่ 11.7 คำนวณหาค่าดัชนีตามฤดูกาล (seasonal index)

ไตรมาส	S_t I_t	ดัชนีตามฤดูกาล
1	.971, .918, .908	.93
2	.840, .839, .834	.84
3	1.096, 1.075, 1.109	1.09
4	1.133, 1.156, 1.141	1.14

หาค่าดัชนีตามฤดูกาลที่ปรับปรุงแล้ว โดย การหารผลบวกของค่าดัชนีราคาที่ยังไม่ได้ปรับปรุงด้วยจำนวนฤดูกาล หรือ

$$\text{ดัชนีตามฤดูกาลที่ปรับปรุง} = \frac{\text{ผลบวกของดัชนีตามฤดูกาลที่ยังไม่ได้ปรับปรุง}}{\text{จำนวนฤดูกาล}} \quad \text{.. (11.19)}$$

$$= \frac{.93 + .89 + 1.09 + 1.14}{4} = \frac{4}{4} = 1$$

หมายเหตุ ในกรณีที่ผลบวกดัชนีตามฤดูกาลมีค่าเท่ากับจำนวนฤดูกาลพอดี ไม่ต้องหาค่าดัชนีตามฤดูกาลที่ปรับปรุงแล้ว

ตารางที่ 11.8 แสดงการหาข้อมูลอนุกรมเวลาที่ไม่รวมองค์ประกอบฤดูกาล

(Deseasonalizeing the time series)

ปี (1)	ไตรมาส (2)	ปริมาณการส่งออก (3)	ดัชนีตามฤดูกาล (4)	ดัชนีตามฤดูกาล (5) = (3) / (4)
1	1	4.8	.93	5.16
	2	4.1	.84	4.88
	3	6.0	1.09	5.50
	4	6.5	1.14	5.70
2	1	5.8	.93	6.24
	2	5.2	.84	6.19
	3	6.8	1.09	6.24
	4	7.4	1.14	6.49
3	1	6.0	.93	6.45
	2	5.6	.84	6.67
	3	7.5	1.09	6.88
	4	7.8	1.14	6.84
4	1	6.3	.93	6.77
	2	5.9	.84	7.02
	3	8.0	1.09	7.34
	4	8.4	1.14	7.37

การหาร Deseasonalizing the time series **คือการแยกดัชนีฤดูกาบ** (seasonal index) **ออกไปจากอนุกรมเวลา**

ตารางที่ 11.9 แสดงการหาค่าเพื่อใช้คำนวณค่าแนวโน้ม

t	อนุกรมเวลาที่ขจัดค่าแนวโน้มตามฤดูกาลแล้ว (Y_t)	tY_t	t^2
1	5.16	5.16	1
2	4.88	9.76	4
3	5.50	16.50	9
4	5.70	22.80	16
5	6.24	31.20	25
6	6.19	37.14	36
7	6.24	43.68	49
8	6.49	51.92	64
9	6.45	58.05	81
10	6.67	66.70	100
11	6.88	75.68	121
12	6.84	82.08	144
13	6.77	88.01	169
14	7.02	98.28	196
15	7.34	110.10	225
16	7.37	117.92	256
136	101.74	914.98	1,496

$$\bar{t} = 136 / 16 = 8.5$$

$$\bar{Y} = 101.74 / 16 = 6.359$$

$$b_1 = \frac{914.98 - (136)(101.74)/16}{1496 - (136)^2/16} = .148$$

$b_0 \quad = \quad 6.359 - .148\ (8.5) \quad = \quad 5.101$

$\therefore \quad T \quad = \quad 6.359 - .148(.85) \quad = \quad 5.101$

ใช้ค่าอนุกรมเวลาที่ขจัดเรื่องฤดูกาลออกไปแล้วไปคำนาณหาค่าแนวโน้ว ตามสมการ

$$T_t \quad = \quad b_0 + b_1 t \qquad \text{.......................................} \quad (11.20)$$

T_t = แนวโน้มของการส่งออก ณ เวลา t

b_0 = จุดตัดแกนตั้งของเส้นแนวโน้ม

b_1 = สโลปของเส้นแนวโน้ม

$$b_1 \quad = \quad \frac{\sum tY_t - (\sum t \sum Y_t)/n}{\sum t^2 - (\sum t)^2/n} \qquad \text{.................................} \quad (11.21)$$

$$b_1 \quad = \quad \overline{Y} - b_1 \overline{t}$$

ดังนั้น $T_t \quad = \quad 5.101 + .148\,t$ ใช้สมการนี้เพื่อการพยากรณ์ค่าแนวโน้ม หาค่าแนวโน้มในไตรมาสที่ 1, 2, 3, 4 ของปีที่ 5 ได้ดังนี้

$T_{17} \quad = \quad 5.101 + .148\ (17) \quad = \quad 7617$

$T_{18} \quad = \quad 5.101 + .148\ (18) \quad = \quad 7765$

$T_{19} \quad = \quad 5.101 + .148\ (19) \quad = \quad 7913$

$T_{20} \quad = \quad 5.101 + .148\ (20) \quad = \quad 8061$

ตารางที่ 11.10 แสดงการแยกองค์ประกอบเพื่อหาค่า Irregular Component

t	Y_t	T	MA	C = MA/T	S	I
1	4.8	5.249			.93	
2	4.1	5.397			.84	
3	6.0	5.545	5.475	.99	1.09	1.002
4	6.5	6.285	5.738	.91	1.14	,996
5	5.8	5.841	5.975	1.022	.93	
6	5.2	5.989	6.188	1.033	.84	
7	6.8	6.137	6.325	1.031	1.09	
8	7.4	6.285	6.400	1.018	1.14	
9	6.0	6.433	6.538	1.016	.93	
10	5.9	6.581	6.675	1.014	.84	
11	7.5	6.729	6.763	1.001	1.09	
12 ,	7.8	6.877	6.838	,994	1.14	
13	6.3	7.025	6.938	,987	.93	
14	5.9	7.173	7.075	,986	.84	
15	8.0	7.321			1.09	
16	8.4	7.469			1.14	

การหาค่าความแปรปรวนที่ไม่สม่ำเสมอ (Irregular Variation)

$T_3\ C_3\ S_3 \quad = (5.545)\ (.99)\ (1.09) = \quad 5.983$

$T_4\ C_4\ S_4 \quad = (6.285)\ (.91)\ (1.14) = \quad 6.520$

$Y_3 = \quad T_3 \times C_3 \times S_3 \times I_3 = \quad 6$

$$\therefore I_3 = \frac{T_3\ C_3\ S_3\ I_3}{T_3\ C_3\ S_3} = \frac{6}{5.983} = 1.002$$

ในทำนองเดียวกัน

$$I_4 = \frac{6}{(6.520)} = .996$$

11.6 การพยากรณ์ด้วย Box - Jenkins

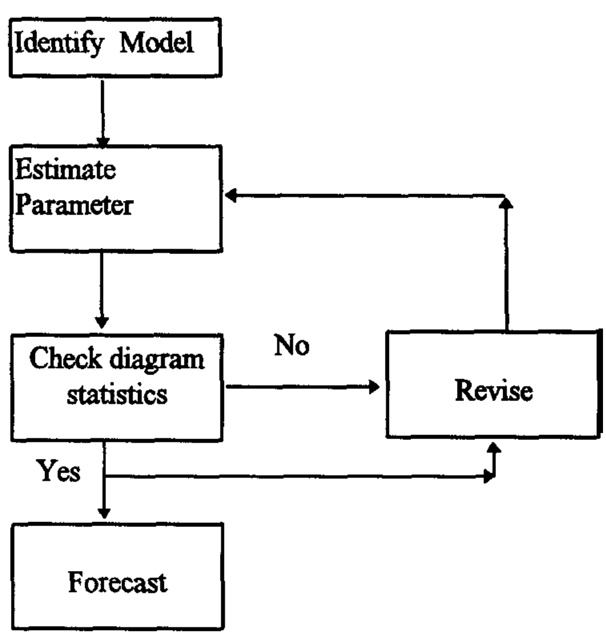
วิเคราะห์ด้วยวิธี Box - Jenkins หรือ univariate Box - Jenkins ซึ่งมีขบวนการ 3 ขั้นตอนดังนี้

1. ระบบแบบจำลอง (Model identification) โดยการใช้ค่าสถิติต่างๆ ที่คำนวณจากการวิเคราะห์ข้อมูลในอดีต

2. การประมาณและการอธิบายแบบจำลอง (Model estimation and verification) แบบจำลองที่ดี คือ ค่าที่ได้จะต้องใกล้เคียงข้อมูลตามความเป็นจริงมากที่สุดเท่าที่จะมากได้

3. การพยากรณ์ (Forecasting) ใช้แบบจำลองเพื่อการพยากรณ์ และเพื่อพัฒนาช่วงความน่าเชื่อถือที่สามารถวัดความไม่แน่นอนที่สัมพันธ์จากการพยากรณ์

ภาพที่ 11.3 แสดงขบวนการสร้างแบบจำลองของ the Box - Jenkins



การวิเคราะห์ Box - Jenkins มีข้อสมมติว่าข้อมูลจะผันผวนรอบๆ ค่าคงที่ ซึ่งเป็นค่าเฉลี่ยของข้อมูลอนุกรมเวลา นั่นคือ ข้อมูลจะมีลักษณะเป็น stationary

แบบจำลองของ Box-Jenkins

1. สมมติให้ y คือค่าสังเกต

y_t คือ ค่าสังเกต ณ ช่วงเวลาที่มีความสม่ำเสมอ

$$y_t = \{ y_1 , y_2 , y_3 , \ldots\ldots, y_n \} \qquad \ldots\ldots\ldots (11.22)$$

วัตถุประสงค์ของ Box - Jenkins คือจะศึกษาว่าค่าสังเกตเกิดจากความสัมพันธ์อย่างเป็นระบบตลอดเวลาหรือไม่ โดยค่า y แต่ละค่าจะแยกเป็น 2 ส่วนคือ explainable variable หรือ predictable variable (p_t) และ ค่าคลาดเคลื่อน (ε_t)

$$y_t = p_t + \varepsilon_t \qquad \ldots\ldots\ldots (11.23)$$

$$\varepsilon_t = y_t - p_t \qquad \ldots\ldots\ldots (11.24)$$

สมการที่ 11.24 แสดงว่า ผลต่างระหว่างค่าสังเกตตามความเป็นจริง กับค่าพยากรณ์ คือค่าตกค้าง (residual) ซึ่งแทนด้วยสัญญลักษณ์ ε_t

ข้อสมมติเกี่ยวกับค่าตกค้า คือ $E(\varepsilon_t) = 0$ และ $E(\varepsilon_t . \varepsilon_{t-1}) = 0$

การที่ $E(\varepsilon_t . \varepsilon_{t-1}) = 0$ แสดงว่า ε_t และ ε_{t-1} ไม่มีความสัมพันธ์กัน ค่าสังเกต ณ เวลา t (y_t) เป็นสัดส่วนโดยตรงกับค่า y ณ ช่วงเวลาก่อน (y_{t-1}) ค่าความสัมพันธ์ระหว่าง y_t และ y_{t-1} คือ ϕ_1 หรืออธิบายอีกนับหนึ่งได้ว่า การประมาณการที่ดีที่สุดของค่าปัจจุบันของ y คือ ค่าในช่วงเวลาก่อน บวกกับค่าคาดเคลื่อน อนึ่ง ค่าพยากรณ์ (y_t) อาจจะสัมพันธ์กับค่าในเวลาก่อนมากกว่า 1 ค่า เช่น

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \qquad \ldots\ldots\ldots (11.25)$$

สมการนี้เรียกว่า a second order autoregression model แสดงว่า y_t มีความสัมพันธ์กับค่าในอดีต 2 ค่า (y_{t-1}, y_{t-2}) บวกกับค่าคลาดเคลื่อน (ε_t) ϕ_1 และ ϕ_2 เป็นค่าแสดงความสัมพันธ์ระหว่าง y_t กับ y_{t-1} และ y_{t-2} ตามลำดับ แบบจำลองเป็นลำดับที่เท่าไรขึ้นอยู่กับค่า p ซึ่งเป็นค่าสูงสุด เช่น p = 5 เรียกสมการนี้ว่า a autoregression model นอกจากนี้ เราอาจจะตัดค่าของระยะต้นๆ หรือค่าของตัวแปรที่อยู่ในลำดับต่ำๆ ออกได้ เช่น

$$y_t = \phi_3 y_{t-3} + \phi_2 y_{t-} + \varepsilon_t \qquad \text{......................... (11.26)}$$

2. รูปแบบทั่วไปของแบบจำลอง Box - Jenkins อันที่ 2 เป็นกรอบของค่าเฉลี่ยเคลื่อนที่ (MA) แบบจำลองค่าเฉลี่ยเคลื่อนที่เชื่อมค่าปัจจุบันของอนุกรมเวลากับค่าคลาดเคลื่อน ที่เกิดขึ้นในช่วงเวลาก่อน แทนที่จะสัมพันธ์กับค่าอนุกรมเวลาของมันเองตามความเป็นจริง

ดังนั้น เราสามารถเขียนแบบจำลองค่าเฉลี่ยเคลื่อนที่ ได้ ดังนี้

$$y_t = \delta - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \text{.........} \theta_q \varepsilon_{t-q} + \varepsilon_t \qquad \text{....................... (11.27)}$$

δ = ค่าเฉลี่ยซึ่งข้อมูลอนุกรมเวลาผันผวนไปรอบๆ

θ_1 = คือค่าพารามิเตอร์ของค่าเฉลี่ยเคลื่อนที่ $i = 1, 2, \ldots. q$ เป็นค่าที่เราต้องประเมิน

ε_{t-q} = ค่าคลาดเคลื่อน

t = ระยะเวลา

q = จำนวนช่วงเวลาที่ย้อนหลังในแบบจำลอง

พิจารณาแบบจำลองที่มีเพียงค่าเดียว

$$y_t = -\theta_1 \varepsilon_{t-1} + \varepsilon \qquad \text{......................... (11.28)}$$

หมายความว่า ค่าสังเกต ของอนุกรมเวลา ณ เวลา t (y_t) เป็นสัดส่วนโดยตรงกับค่าคลาดเคลื่อน ค่าความสัมพันธ์นี้วัดโดย θ_1 เช่นเดียวกับกรณีของแบบจำลอง AR นั่นคือ

แบบจำลอง MA สามารถขยายให้รวมจำนวนตัวแปรทั้งหมด โดยไม่จำเป็นต้องตัวแปรตรงกลางไว้ด้วย เช่น

$$y_t = -\theta_3 t_{t-3} + \varepsilon_t \quad \text{.........................} (11.29)$$

ข้อสังเกต แบบจำลอง MA สมมติว่า ค่าปัจจุบันของอนุกรม เป็นผลโดยตรงของค่าคลาดเคลื่อนในอดีต

3. แบบจำลองสุดท้ายของ Box - Jenkins คือ ARMR ซึ่งสมการจะรวมทั้ง AR และ MR

ภายใต้กรอบของ Box - Jenkins จะมีแบบจำลอง 3 รูปแบบคือ

1. Autoregression (AR) model
2. Moving average (MA) model
3. ARMA model เป็นแบบจำลองที่ผสมระหว่าง AR และ MA

1. เมื่อแบบจำลองเป็น AR ค่าปัจจุบันของตัวแรกขึ้นอยู่กับค่าของตัวแปรตัวเดียวกับระยะเวลาก่อน (previous values) บวกค่าคลาดเคลื่อน (error term)

$$y_t = \delta + \phi_1 y_{t-1} - \phi_2 y_{t-2} - \text{.........} \phi_q y_{t-q} + \varepsilon_t \quad \text{.....................} (11.30)$$

ค่า ϕ_i คือ ค่าพารามิเตอร์ที่ต้องประมาณค่า

δ คือ ค่าคงที่ ซึ่งเกี่ยวพันธ์กับค่าแนวโน้มในอนุกรม

y_t คือ ค่าสังเกตตามความเป็นจริง

p หมายถึง จำนวนช่วงเวลาย้อนหลัง

แบบจำลองนี้เรียกว่า autoregressive เพราะ ค่าสังเกตตามความเป็นจริงขึ้นอยู่กับตัวของมันเองในช่วงเวลาย้อนหลัง เพื่อทำให้ง่ายกับความเข้าใจ สมมติว่า แบบจำลองมีพารามิเตอร์เพียงตัวเดียว และค่าเฉลี่ยหรือค่าคงที่ (δ) เป็นศูนย์ แบบจำลองคือ

$$y_t = \phi_1 y_{t-1} - \varepsilon_t \quad \text{..................................} (11.31)$$

แบบจำลองนี้มีเพียง one order และ ϕ_1 คือ พารามิเตอร์ของ AR ค่า $\phi_1 y_{t-1}$ คือค่าพยากรณ์ ε_t คือ ค่าคลาดเคลื่อนซึ่งเป็นค่าสุ่ม อธิบายความสัมพันธ์ของสมการนี้ได้ดังนี้

$$y_t = \delta + \phi_1 y_{t-1} - \phi_2 y_{t-2} - \dots\dots\dots \phi_q y_{t-p}$$

$$- \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots\dots\dots \theta_q \varepsilon_{t-q} + \varepsilon_t \quad \dots\dots\dots\dots\dots\dots (11.32)$$

ลำดับของแบบจำลองอ่านจากค่า p และ q ปกตอจะเขียนด้วย ARMA (p, q) สมการได้รวมค่า δ ไว้ด้วย ข้อสมมติที่ $\delta = 0$ เป็นกรณีพิเศษเท่านั้น

ถ้าข้อมูลอนุกรมเวลามีลักษณะเป็น nonstationary การจะใช้ the Box - jenkins methodology เพื่อการพยากรณ์ จะต้องปรับข้อมูลให้เป็น stationary โดย ขบวนการของการแตกต่าง (the process of differencing)

เทคนิคของการแตกต่าง (The technique of differencing)

1. HI the first - order difference

$$\Delta y_t = y_t - y_{t-1} \quad \dots\dots\dots\dots\dots\dots (11.33)$$

2. หา the second -order difference

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} \quad \dots\dots\dots\dots\dots\dots (11.34)$$

ตารางที่ 11.11 แสดงการหาค่าความแตกต่างรอบที่ 1 และรอบที่ 2

Y	ความแตกต่างรอบที่ 1	ความแตกต่างรอบที่ 2
15		
28	13	
45	17	4
66	21	4
91	25	4
120	29	4

ถ้าข้อมูลตามความเป็นจริงไม่ได้มีลักษณะเป็นแนวโน้ม การที่จะทำให้ข้อมูลเป็น stationary ในกรณีที่เป็น parabolic trend เราต้องหา the second difference ในกรณีที่เป็น s - curve เราต้องแปลงข้อมูลอนุกรมเวลาให้อยู่ในรูป logarithm แล้วจึงหา first difference ผลที่ได้คือ ข้อมูลที่เป็น stationary ในกรณีที่ข้อมูลอนุกรมมีลักษณะเป็น linear trend ถ้าหาความแตกต่างรอบที่ 1 แล้วยังไม่มีคุณสมบัติเป็น stationary เราต้องหาความแตกต่างรอบที่ 2 และถ้ายังไม่มีคุณสมบัติเป็น stationary อีก ขั้นต่อไปต้องแปลงข้อมูลอนุกรมเวลาให้เป็นค่า logarithm แล้วจึงหาค่า ความแตกต่างของข้อมูล การหาค่าผลต่าง เพราะเราต้องการขจัดค่าแนวโน้มนั่นเอง