หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

รูปแบบทางสถิติที่ใช้ในการวิเคราะห์

จากการได้ศึกษาถึงปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์จากตัวแปรต่าง ๆ ที่คาดว่าจะมีอิทธิพลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ในเชิงบรรยายแล้ว บทนี้จะทำการวิเคราะห์ตัวแปรเหล่านั้นในทางสถิติ เพื่อดูว่าตัวแปรใดมีความสัมพันธ์อย่างไรต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ โดยการใช้สมการถดถอยและค่าสหสัมพันธ์ในการวิเคราะห์

ทฤษฎีการถดถอยและสหสัมพันธ์

การวิเคราะห์การถดถอย (Regression Analysis) เป็นการใช้เทคนิคทางสถิติเข้ามาช่วยหาความสัมพันธ์ระหว่างตัวแปรตัวหนึ่งหรือหลายตัว ซึ่งเรียกว่า ตัวแปรอิสระ (Independent Variable) ว่ามีอิทธิพลหรือมีความสัมพันธ์ต่อตัวแปรอีกตัวหนึ่ง ซึ่งเรียกว่า ตัวแปรตาม (Dependent Variable) อย่างไร โดยแสดงในรูปสมการถดถอย (Regression Equation) ส่วนการวิเคราะห์ค่าสหสัมพันธ์ (Correlation Analysis) เป็นเรื่องเกี่ยวกับการวัดระดับความสัมพันธ์ระหว่างตัวแปรต่าง ๆ โดยวัดออกมาเป็นสัมประสิทธิ์ของการตัดสินใจ (Coefficient of Determination ; r^2) และสัมประสิทธิ์สหสัมพันธ์ (Coefficient of Correlation, r) ซึ่งรวมไปถึงการทดสอบความสัมพันธ์นี้ โดยใช้ F-test หรือ T-test

การวิเคราะห์ที่เป็นการนำเอาตัวแปรตามตัวหนึ่งกับตัวแปรอิสระเพียงตัวเดียวมาทำการวิเคราะห์ เรียกว่า การถดถอยหรือสหสัมพันธ์อย่างง่าย (Simple Regression or Correlation) แต่ถ้าตัวแปรอิสระมีมากกว่า 1 ตัว เรียกว่า การถดถอยหรือสหสัมพันธ์เชิงซ้อน (Multiple Regression or Correlation)

รูปแบบการถดถอยเชิงซ้อน

สมการถดถอยเชิงซ้อน เป็นการแสดงให้เห็นอิทธิพลของตัวแปรอิสระหลาย ๆ ตัวที่มีต่อ

ตัวแปรตามร่วมกัน โดยปกติรูปแบบที่ใช้ในการวิเคราะห์มี 2 แบบ คือ

1. รูปแบบเส้นตรง (Linear Regression Model)

2. รูปแบบที่ไม่ใช่เส้นตรง (Non-Linear Regression Model)

เพื่อเป็นการง่ายและสะดวกในการวิเคราะห์การถดถอย จะถือว่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระเป็นไปในลักษณะที่เป็นเส้นตรง ดังนั้น ถ้ากำหนดให้ Y เป็นตัวแปรตาม และ X เป็นตัวแปรอิสระ สมการเส้นตรงของการถดถอยเชิงซ้อน จะมีรูปแบบดังนี้

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \text{---} + \beta_n X_n + E$$

Y = ตัวแปรตาม

α = ตัวคงที่และเป็นค่าของ Y เมื่อค่าของ X แต่ละตัวมีค่าเป็นศูนย์

$\beta_1, \text{---}, \beta_n$ = สัมประสิทธิ์ของตัวแปรอิสระแต่ละตัว ซึ่งเป็นค่าที่แสดงว่า เมื่อ X แต่ละตัวเปลี่ยนแปลงไป 1 หน่วย Y จะเปลี่ยนไปเท่าใด ซึ่งเรียกว่า พารามิเตอร์ถดถอย หรือความชันของเส้นถดถอย

$X_1, \text{---}, X_n$ = ตัวแปรอิสระ

E = ความคลาดเคลื่อนเชิงสุ่ม ซึ่งแสดงว่าค่า Y ที่หาได้แตกต่างไปจากเส้นถดถอยที่แท้จริงเท่าไร

สมการของการถดถอยข้างต้นนี้เป็นสมการถดถอยของประชากร ซึ่งมีจำนวนที่ใหญ่มากและยากที่จะทราบจำนวนที่แน่นอน ในการวิเคราะห์จึงต้องทำการสุ่มตัวอย่าง (Sampling) จากตัวอย่างชุดใดชุดหนึ่งไปสู่ประชากรอันเป็นที่มาของตัวอย่างชุดนั้น เส้นถดถอยของตัวอย่างชุดหนึ่งเป็นเพียงเส้นถดถอยเส้นหนึ่งในหลาย ๆ เส้นที่ได้จากตัวอย่างชุดต่าง ๆ ซึ่งทำการสุ่มตัวอย่างมาจากประชากรกลุ่มเดียวกัน

จากการสุ่มตัวอย่างประชากร จะได้สมการถดถอยโดยประมาณดังนี้

$$\hat{Y} = a + b_1 x_1 + b_2 x_2 + \text{---} + b_n x_n$$

$\hat{Y}$ = ค่ากะประมาณของ Y

a = ค่ากะประมาณของ α

$b_1, \text{---}, b_n$ = ค่ากะประมาณของ β

$x_1, ---, x_n$ = ตัวแปรอิสระ

เพื่อให้การอ้างอิงเกี่ยวกับความสัมพันธ์ของประชากรจากข้อมูลตัวอย่าง เป็นการอ้างอิงที่ใช้ได้ จึงจำเป็นจะต้องกำหนดข้อสมมติบางประการ คือ

1. จำนวนค่าสังเกต (Observation) ทั้งหมดต้องมากกว่าจำนวนพารามิเตอร์ที่ต้องการประมาณ

2. ตัวแปรอิสระแต่ละตัว เป็นอิสระจากกัน

3. ความสัมพันธ์ของประชากรจะต้องเป็นไปในลักษณะที่เป็นเส้นตรง

4. ค่า E แต่ละค่าเป็นอิสระจากกันและมีการแจกแจงเป็นปกติ

<u>การประมาณค่าของพารามิเตอร์</u> (The Estimation of Parameters)

การหาค่าของ a และ b ในสมการถดถอยเชิงซ้อนเส้นตรง วิธีที่สะดวกที่สุดคือ ใช้วิธีกำลังสองน้อยที่สุด (Least Square Method) ตามวิธีนี้เป็นการแก้สมการเส้นตรงเพื่อหาตัวไม่ทราบค่าคือ ค่า a 1 ตัว และค่า b ตามจำนวนแปรตาม ซึ่งในการวิจัยนี้มีตัวแปรอิสระทั้งหมด 5 ตัว สมการจะสามารถเขียนได้ดังนี้

$$\hat{Y} = a + b_1x_1 + b_2x_2 + --- + b_5x_5$$

สมการที่จะต้องแก้ คือ

$$\Sigma Y = na + b_2\Sigma x_3 + b_2\Sigma x_2 + --- + b_5\Sigma x_5 \text{ ----------(1)}$$

$$\Sigma x_1y = a\Sigma x_1 + b_1\Sigma x_1^2 + b_2\Sigma x_1x_2 + --- + b_5\Sigma x_1x_5 \text{ -------(2)}$$

$$\Sigma x_2y = a\Sigma x_2 + b_1\Sigma x_1x_2 + b_2\Sigma x_2^2 + --- + b_5\Sigma x_2x_5 \text{ -------(3)}$$

$$\Sigma x_5y = a\Sigma x_5 + b_1\Sigma x_1x_5 + b_2\Sigma x_2x_5 + --- + b_5\Sigma x_5^2 \text{ -------(6)}$$

จากการแก้สมการดังกล่าวข้างต้น จะทำให้ทราบค่า a และค่า $b_1, b_2 ---, b_5$ เพื่อนำไปแทนค่าในสมการ $\hat{Y}$ ซึ่งเป็นสมการถดถอยที่ต้องการ

<u>ความคลาดเคลื่อนมาตรฐานของการกะประมาณ</u> (Standard Error of Estimates)

ความคลาดเคลื่อนมาตรฐานของการกะประมาณ เป็นมาตรการที่ใช้วัดการกระจัดกระจายของข้อมูลที่รวบรวมได้ ว่ามีการกระจัดกระจายไปจากเส้นถดถอยที่คำนวณได้มากน้อยเพียงใด ถ้าค่าของ Y ที่สังเกตุได้แตกต่างไปจากเส้นถดถอยมาก ค่า Y ที่กะประมาณจากเส้นนี้ก็มีโอกาสที่จะแตกต่างไปจากค่าที่เกิดขึ้นจริงได้ง่าย และไม่อาจนำไปใช้ประโยชน์ในการคาดคะเนล่วงหน้าได้ดีเท่าที่ควร ในทางตรงกันข้ามถ้าค่าของ Y ที่สังเกตุได้อยู่ใกล้กับเส้นถดถอยแล้ว การกะประมาณค่าของ Y จากเส้นถดถอยที่คำนวณได้ก็จะมีความเชื่อถือได้มากขึ้น

<u>สัมประสิทธิ์ของการตัดสินใจ</u> (Coefficient of determination : r^2)

สัมประสิทธิ์ของการตัดสินใจ เป็นมาตรการวัดการปรับที่ดีที่สุด (Goodness of Fit) โดยชี้แจงให้เห็นว่า ตัวแปรอิสระมีส่วนทำให้เกิดการเปลี่ยนแปลงขึ้นในตัวแปรตามมากน้อยเพียงใด สัญญลักษณ์ที่ใช้แทนสัมประสิทธิ์ของการตัดสินใจคือ r^2 ค่าของ r^2 จะเป็นบวกเสมอและมีค่าไม่เกิน 1 ($0 \leq r^2 \leq 1$) กล่าวคือ ถ้า r^2 มีค่าเป็นศูนย์ จะชี้ให้เห็นว่าตัวแปรทั้งสองไม่มีความสัมพันธ์กันเลย แต่ถ้าค่าของ r^2 เข้าใกล้ 1 มากเพียงใด แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์กันมากเท่านั้น

<u>สัมประสิทธิ์สหสัมพันธ์</u> (Coefficient of Correlation ; r)

สัมประสิทธิ์สหสัมพันธ์ เป็นมาตรการที่ใช้วัดความสัมพันธ์ระหว่างตัวแปรอิสระที่มีต่อตัวแปรตาม สัญญลักษณ์ที่ใช้คือ r ค่าของ r จะอยู่ระหว่าง +1 กับ -1 กล่าวคือ ถ้าค่า r เท่ากับ ศูนย์ แสดงว่าตัวแปรทั้งสองไม่มีความสัมพันธ์กันเลย แต่ถ้าค่า r มีค่าใกล้ 1 มากเพียงใด แสดงว่าความสัมพันธ์ระหว่างตัวแปรทั้งสองมีมากเท่านั้น และเป็นไปในทิศทางเดียวกัน แต่ถ้าค่า r มีค่าใกล้ -1 มากเพียงใด ก็แสดงว่ามีความสัมพันธ์กันในทิศทางตรงกันข้าม ค่าของ r จะหาได้จากการถอดรากกำลังที่ 2 ของ r^2

<u>สัมประสิทธิ์ของการตัดสินใจบางส่วน</u> (Coefficient of Partial Detemination)

สัมประสิทธิ์ของการตัดสินใจตามที่กล่าวมาแล้ว เป็นการวัดอิทธิพลของตัวแปรอิสระหลาย ๆ

ตัวที่มีต่อตัวแปรตามร่วมกัน แต่สัมประสิทธิ์ของการตัดสินใจบางส่วนนี้จะเป็นการวัดอิทธิพลของตัวแปรอิสระทีละตัวกับตัวแปรตามว่า ตัวแปรอิสระตัวใดจะมีอิทธิพลต่อตัวแปรตามมากน้อยกว่ากัน โดยในการวิเคราะห์เพื่อหาสัมประสิทธิ์ของการตัดสินใจบางส่วนของตัวแปรอิสระตัวหนึ่งกับตัวแปรตามนั้น จะถือว่าตัวแปรอิสระตัวอื่นอยู่คงที่

<u>การทดสอบความมีนัยสำคัญทางด้านความสัมพันธ์</u>

บางครั้งตัวอย่างที่สุ่มมานั้น อาจทำให้เห็นว่า ตัวแปรตามกับตัวแปรอิสระทุกตัวมีความสัมพันธ์กัน ซึ่งจริง ๆ แล้วอาจเป็นเรื่องที่เกิดขึ้นโดยบังเอิญก็ได้ ดังนั้น จึงต้องทำการทดสอบอีกครั้ง จากการศึกษาข้างต้นเราทราบว่า ถ้าตัวแปรตามไม่มีความสัมพันธ์กันกับตัวแปรอิสระทุกตัวเลย ค่าสัมประสิทธิ์การถดถอย β_1, β_2 ---, β_5 จะเท่ากับศูนย์ เพราะฉะนั้น จึงตั้งสมมติฐานเพื่อใช้ในการทดสอบได้ดังนี้

$$H_o : \beta_i = 0 \quad (i = 1, 2, 3, 4, 5)$$

$$H_a : \beta_i \neq 0$$

ตัวสถิติที่ใช้ในการทดสอบสมมติฐาน คือ

$$F = \frac{\text{Mean Square (Due to Regression)}}{\text{Error Mean Square}}$$

ถ้าค่าของ F ที่คำนวณได้นี้มากกว่าค่า $F_{1-\frac{\alpha}{2}, (k-1)(n-k)}$ ซึ่งได้จากการเปิดตารางโดย n คือ จำนวนตัวอย่าง และ k คือ จำนวนพารามิเตอร์ทั้งหมด ในการวิเคราะห์แล้วก็จะปฏิเสธสมมติฐาน H_o และยอมรับสมมติฐาน H_a นั่นคือ ยอมรับว่าตัวแปรตามมีความสัมพันธ์กับตัวแปรอิสระทุกตัว

สำหรับการทดสอบว่า ตัวแปรอิสระแต่ละตัวมีความสัมพันธ์กับตัวแปรตามหรือไม่ก็อาจใช้ F-test ทดสอบได้เช่นกัน โดยสมมติฐานที่ใช้ทดสอบจะเป็นดังนี้

$$H_o : \beta_i = 0 \quad (i = 1, 2, ___, 5)$$

$$H_a : \beta_i \neq 0$$

ตารางที่ 4.1 ข้อมูลที่ใช้ในการวิเคราะห์สมการถดถอย

พ.ศ.	ปริมาณเงินฝากของธนาคารพาณิชย์ (Y)	มูลค่าผลิตภัณฑ์ในประเทศ (X_1)	ปริมาณเงินหมุนเวียน (X_2)	อัตราดอกเบี้ยเงินฝาก (X_3)	จำนวนสาขาของธนาคารพาณิชย์ (X_4)	ปริมาณเงินให้กู้ยืม (X_5)
2507	11,901.95	75,473.52	11,985.27	7.05	446	8,517.20
2508	14,141.60	84,303.0	12,916.8	7.0	455	8,930.70
2509	17,240.99	97,021.36	14,038.79	6.7	475	10,131.51
2510	19,399.54	99,443.53	14,422.68	6.43	516	11,561.71
2511	22,247.16	105,296.66	15,586.47	6.31	549	13,192.16
2512	24,845.33	113,174.29	15,835.21	6.16	584	15,648.42
2513	28,646.96	119,778.85	17,134.36	6.17	628	19,019.47
2514	33,946.40	126,807.02	18,812.02	6.14	663	27,815.61
2515	40,472.83	135.510.87	20,761.62	5.85	712	29,971.32
2516	42,537.80	156,801.59	21,677.41	5.07	760	37,140.62
2517	43,449.62	157,073.38	19,340.65	4.66	827	40,079.03
2518	48,505.70	163,881.64	19,348.73	4.42	876	45,851.11
2519	56,885.19	176.314.76	21,564.44	4.25	1,046	51,155.68
2520	64,519.34	182,737.27	21,850.98	3.55	1,165	60,581.10
2521	71,997.35	214,854.84	24,187.54	3.29	1,279	73,551.18
2522	72,639.10	231,381.03	26,434.53	3.37	1,363	82,513.81
2523	74,731.17	238,079.18	24,846.33	3.75	1,458	77,971.64
2524	79,477.64	242,464.22	22,616.83	3.61	1,536	79,606.89
2525	94,480.26	247,995.54	22,958.56	3.33	1,623	89,916.99
2526	114,699.32	262,301.69	23,111.64	3.21	1,689	116,373.67
2527	139,439.78	277,786.12	26,144.59	3.19	1,749	134,964.96

ตารางที่ 4.2 ผลการวิเคราะห์ตัวแปรอิสระแต่ละตัวต่อตัวแปรตาม

ตัวแปรอิสระ	ค่าสัมประสิทธิ์ของการตัดสินใจ (r^2)	ค่าสัมประสิทธิ์สหสัมพันธ์ (r)	F test	Standard Error	Standardized Coefficient
1. มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น	0.9530	0.9762	364.763	6450.787	0.97620
2. ปริมาณเงินหมุนเวียน	0.7615	0.8726	57.468	14527.671	0.87263
3. อัตราดอกเบี้ยเงินฝาก	0.8755	-0.9357	126.598	10495.315	-0.93569
4. จำนวนสาขาของธนาคารพาณิชย์	0.9646	0.9822	491.123	5593.270	0.98216
5. ปริมาณเงินให้กู้ยืม	0.9857	0.9928	1244.43	3552.009	0.99284

หมายเหตุ รายละเอียดแสดงในภาคผนวก ข.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ตัวสถิติที่ใช้ในการทดสอบสมมติฐาน คือ

$$F = \frac{\text{Mean Square (Due to } X_i)}{\text{Error Mean Square}}$$

ถ้าค่าของ F ที่คำนวณได้มากกว่าค่า F จากการเปิดตารางก็แสดงว่า ปฏิเสธสมมติฐาน H_o และยอมรับสมมติฐาน H_a นั่นคือ ยอมรับว่าตัวแปรอิสระที่ทดสอบตัวนั้นมีความสัมพันธ์กับตัวแปรตาม

<u>รูปแบบที่ใช้ในการวิเคราะห์ข้อมูล</u>

ปัจจัยสำคัญที่มีอิทธิพลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ตามที่ได้กล่าวไว้ในบทที่ 3 จะนำมาเป็นตัวแปรในการวิเคราะห์สมการถดถอยดังนี้.-

y = ปริมาณเงินฝากของธนาคารพาณิชย์ (หน่วย : ล้านบาท)

x_1 = มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น (หน่วย : ล้านบาท)

x_2 = ปริมาณเงินหมุนเวียน (หน่วย : ล้านบาท)

x_3 = อัตราดอกเบี้ยเงินฝาก (หน่วย : ร้อยละ)

x_4 = จำนวนสาขาของธนาคารพาณิชย์ในประเทศไทย (หน่วย : จำนวนแห่ง)

x_5 = ปริมาณเงินให้กู้ยืม (หน่วย : ล้านบาท)

สำหรับข้อมูลที่ใช้ในการวิเคราะห์นี้ได้แสดงไว้ในตารางที่ 4.1 จากข้อมูลดังกล่าวเมื่อนำมาพิจารณาถึงความสัมพันธ์ของตัวแปรอิสระแต่ละตัวที่มีต่อตัวแปรตามด้วยวิธีการถดถอยอย่างง่าย (Simple Regression) ตามตารางที่ 4.2 ผลของการวิเคราะห์เป็นดังนี้.-

1. มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น ผลการวิเคราะห์ระหว่างมูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้นต่อปริมาณเงินฝากของธนาคารพาณิชย์ ปรากฏว่าค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) มีค่าเท่ากับ 0.9530 แสดงว่าการเปลี่ยนแปลงไปของปริมาณเงินฝากของธนาคารพาณิชย์ เป็นผลสืบเนื่องมาจากการเปลี่ยนแปลงของมูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้นถึง 95.30% และเป็นไปในทิศทางเดียวกัน กล่าวคือ จะเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงไปตามกัน และมีความสัมพันธ์กัน 0.9762 ซึ่งถ้านำสมการถดถอยที่ได้นี้ไปใช้ในการกะประมาณปริมาณเงินฝากของธนาคารพาณิชย์ ความผิด-

พลาดของการกะประมาณจะอยู่ระหว่าง ± 6,450.787 ล้านบาท โดยมีระดับนัยสำคัญ 95%

2. ปริมาณเงินหมุนเวียน ผลการวิเคราะห์ระหว่างปริมาณเงินหมุนเวียนต่อปริมาณเงินฝากของธนาคารพาณิชย์ ปรากฏว่า ค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) มีค่าเท่ากับ 0.7615 แสดงว่า การเปลี่ยนแปลงไปของปริมาณเงินฝากของธนาคารพาณิชย์เป็นผลสืบเนื่องมาจากการเปลี่ยนแปลงของปริมาณเงินหมุนเวียน 76.15% และจะเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงไปตามกัน โดยมีความสัมพันธ์กัน 0.8726 ซึ่งถ้านำสมการถดถอยนี้ไปใช้ในการกะประมาณปริมาณเงินฝากของธนาคารพาณิชย์ ความผิดพลาดของการกะประมาณจะอยู่ระหว่าง ± 14,527.671 ล้านบาท โดยมีระดับนัยสำคัญ 95%

3. อัตราดอกเบี้ยเงินฝาก ผลการวิเคราะห์ระหว่างอัตราดอกเบี้ยเงินฝากต่อปริมาณเงินฝากของธนาคารพาณิชย์ ปรากฏว่า ค่าความสัมพันธ์ของการตัดสินใจ (r^2) มีค่าเท่ากับ 0.8755 แสดงว่าการเปลี่ยนแปลงไปของปริมาณเงินฝากของธนาคารพาณิชย์เป็นผลสืบเนื่องมาจากการเปลี่ยนแปลงของอัตราดอกเบี้ยเงินฝาก 87.55% และเป็นไปในทิศทางตรงกันข้าม กล่าวคือ ถ้าอัตราดอกเบี้ยเงินฝากเพิ่มขึ้น ปริมาณเงินฝากของธนาคารพาณิชย์จะลดลง แต่ถ้าอัตราดอกเบี้ยเงินฝากลดลง ปริมาณเงินฝากของธนาคารพาณิชย์ก็จะเพิ่มขึ้น โดยมีความสัมพันธ์กัน -0.9357 และถ้านำสมการถดถอยนี้ไปใช้ในการกะประมาณปริมาณเงินฝากของธนาคารพาณิชย์ ความผิดพลาดของการกะประมาณจะอยู่ระหว่าง ± 10,495.315 ล้านบาท โดยมีระดับนัยสำคัญ 95%

4. จำนวนสาขาของธนาคารพาณิชย์ในประเทศไทย ผลการวิเคราะห์ระหว่างจำนวนสาขาของธนาคารพาณิชย์กับปริมาณเงินฝากของธนาคารพาณิชย์ ปรากฏว่า ค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) มีค่าเท่ากับ 0.9646 แสดงว่า การเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ เป็นผลสืบเนื่องมาจากการเปลี่ยนแปลงของจำนวนสาขาของธนาคารพาณิชย์ถึง 96.46% และเป็นไปในทิศทางเดียวกัน โดยมีความสัมพันธ์กัน 0.9822 ซึ่งถ้านำสมการถดถอยที่ได้นี้ไปใช้ในการกะประมาณปริมาณเงินฝากของธนาคารพาณิชย์ ความผิดพลาดของการกะประมาณจะอยู่ระหว่าง ± 5,593.270 ล้านบาท โดยมีระดับนัยสำคัญ 95%

5. ปริมาณเงินให้กู้ยืม ผลการวิเคราะห์ระหว่างปริมาณเงินให้กู้ยืมกับปริมาณเงินฝากของธนาคารพาณิชย์ ปรากฏว่า ค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) มีค่าเท่ากับ 0.9857 แสดงว่าการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ เป็นผลสืบเนื่องมาจากการเปลี่ยน

แปลงของปริมาณเงินให้กู้ยืม 98.57% และเป็นไปในทิศทางเดียวกัน โดยจะมีความสัมพันธ์กัน 0.9928 ซึ่งถ้านำสมการถดถอยที่ได้นี้ไปใช้ในกะประมาณปริมาณเงินฝากของธนาคารพาณิชย์ ความผิดพลาดของการกะประมาณจะอยู่ระหว่าง ± 3,552.009 ล้านบาท โดยมีระดับนัยสำคัญ 95%

จากผลการวิเคราะห์ข้างต้น แสดงให้เห็นว่า ตัวแปรอิสระแต่ละตัวต่างก็มีความสัมพันธ์ต่อตัวแปรตาม และความสัมพันธ์จะมากหรือน้อยอย่างไรดูได้จากค่าของสัมปริสทธิ์ของการตัดสินใจ (r^2) และสัมประสิทธิ์สหสัมพันธ์ (r) ส่วนการวัดว่าตัวแปรอิสระตัวใดที่มีค่าความสัมพันธ์มากกว่าก็อาจดูได้จากค่า Standardized Coefficient ซึ่งค่านี้เป็นค่าที่ปรับข้อมูลให้อยู่ในหน่วยเดียวกัน เพื่อให้สามารถนำมาเปรียบเทียบกันได้ ผลการวิเคราะห์ปรากฏว่า ปริมาณเงินให้กู้ยืมเป็นตัวแปรอิสระที่มีค่าความสัมพันธ์กับตัวแปรตามสูงสุด รองลงมาคือจำนวนสาขาของธนาคารพาณิชย์, มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น, อัตราดอกเบี้ยเงินฝาก และปริมาณเงินหมุนเวียน ซึ่งถ้านำไปใช้ในการกะประมาณปริมาณเงินฝากของธนาคารพาณิชย์ โดยอาศัยตัวแปรอิสระตัวใดตัวหนึ่งแล้ว ตัวแปรอิสระที่มีค่าความสัมพันธ์สูงกว่าจะสามารถให้ผลการกะประมาณได้ดีกว่า

ในการวิเคราะห์ตัวแปรอิสระต่าง ๆ ที่คาดว่าจะมีอิทธิพลหรือมีความสัมพันธ์ต่อตัวแปรตาม การนำเอาตัวแปรอิสระเข้ามามากขึ้นจะทำให้การกะประมาณค่าของตัวแปรตามได้ดีขึ้น ดังนั้นจึงนำเอาตัวแปรอิสระทั้ง 5 มาวิเคราะห์ร่วมกัน แต่ในบางครั้งตัวแปรอิสระบางตัวไม่ควรนำมาพิจารณา เนื่องจากตัวแปรอิสระนั้น ๆ มีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ ที่อยู่ในสมการถดถอย ซึ่งถ้ารวมตัวแปรอิสระนั้นเข้าไปในสมการแล้ว จะทำให้สมการถดถอยนั้นใช้ในการกะประมาณค่าของตัวแปรตามได้ไม่ดีเท่าที่ควร ฉะนั้น จึงต้องมีขั้นตอนในการเลือกว่าตัวแปรอิสระใดควรอยู่ในสมการ เพื่อจะทำให้สมการถดถอยนั้นมีค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) มีค่าสูงสุด และให้ค่าความคลาดเคลื่อนมาตรฐานของการกะประมาณ (Standard Error of Estimates) ต่ำสุด นั่นคือ การเลือกสมการถดถอยที่ดีที่สุด

<u>การเลือกรูปแบบสมการถดถอย</u>

วิธีการเลือกสมการถดถอยในโปรแกรม SPSS^x นั้นมีอยู่ทั้งหมด 3 วิธีคือ[1] การเลือก

---

[1]Marya J. Norusis, <u>Introductory Statistics Guide SPSS^x</u> ; Mcgraw-Hill Book Company 1982, Page 160

ตัวแปรแบบไปข้างหน้า (Forward Selection) การกำจัดตัวแปรแบบถอยหลัง (Backward Elimination) และการถดถอยแบบขั้นบันได (Stepwise Regression) ซึ่งวิธีการเลือกทั้ง 3 นี้ไม่มีการกำหนดว่าวิธีไหนจะเป็นวิธีที่ดีที่สุด แต่ขึ้นอยู่กับว่าสมการการถดถอยที่ได้จากแต่ละวิธีนั้น สมการใดจะเป็นสมการการถดถอยที่ดีที่สุดก็ให้เลือกสมการของวิธีนั้น จากคำสั่ง REGRESSION ในโปรแกรม SPSS^{x} ยังสามารถควบคุมตัวแปรอิสระให้เข้าไปในสมการได้ตามความต้องการ และยังสามารถแก้ปัญหาการเกิดความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) ที่เข้าไปในสมการ เพื่อไม่ให้ความสัมพันธ์สูงกว่าเกณฑ์ที่กำหนด

วิธีคัดเลือกตัวแปรอิสระที่กำหนดโดยโปรแกรม SPSS^{x} มีดังต่อไปนี้ คือ

1. การเลือกตัวแปรแบบไปข้างหน้า (Forward Selection) การคัดเลือกตัวแปรอิสระตามวิธีนี้จะพิจารณาจากค่าความสัมพันธ์ (Correlation) ระหว่างตัวแปรอิสระแต่ละตัวที่มีต่อตัวแปรตาม จะเริ่มจากตัวแปรอิสระจะถูกคัดเลือกเข้าไปในสมการทีละตัวจนกระทั่งไม่มีตัวแปรใดถูกคัดเลือกเข้าไปในสมการได้อีก โดยใช้เกณฑ์ของระดับนัยสำคัญ (Significant T; Sig.T) ที่กำหนด ใช้สัญญลักษณ์ PIN(Probability of T-to-Enter) ซึ่งในที่นี้กำหนดไว้ 0.05 โปรแกรม SPSS^{x} จะคำนวณว่าความน่าจะเป็นที่จะยอมรับสมมติฐานหลักของค่าสถิติ $H_o : \beta_i = 0$ ค่า Sig.T ที่คำนวณได้จะถูกนำไปเปรียบเทียบกับค่า PIN ที่กำหนด ถ้า Sig.T $\leq$ PIN จะปฏิเสธสมมติฐาน H_o แสดงว่าตัวแปรอิสระตัวนั้นจะถูกคัดเลือกเข้าไปในสมการ โดยจะเลือกตัวแปรอิสระที่มีค่า Sig.T ต่ำสุด แล้วจึงคำนวณค่า Sig.T แต่ละตัวใหม่อีก โดยคำนึงถึงตัวแปรอิสระที่มีค่า Sig.T ต่ำสุดอีก ทำเช่นนี้จนกว่าจะไม่มีตัวแปรเข้าไปได้อีก

ในบางครั้งอาจไม่มีตัวแปรอิสระใดที่มีค่า Sig.T ตามที่กำหนด ทำให้ไม่มีตัวแปรอิสระใดที่จะถูกนำเข้าสมการ จึงอาจพิจารณาในอีกขั้นหนึ่ง โดยใช้คำสั่ง Enter ซึ่งเป็นการบังคับให้ตัวแปรอิสระทุกตัวเข้าไปในสมการ โดยพิจารณาค่าความสัมพันธ์บางส่วน (Partial Correlation) ระหว่างตัวแปรอิสระแต่ละตัวที่มีต่อตัวแปรตาม ตัวแปรอิสระส่วนที่เหลือนี้ ตัวใดที่มีค่า (Partial Correlation) สูงสุด และมีค่า Tolerance อยู่ในขอบเขตที่กำหนด ตัวแปรอิสระนั้นก็จะถูกนำเข้ามาในสมการ การกำหนดค่า Tolerance ขึ้นเพื่อกำจัดตัวแปรที่มีค่าความสัมพันธ์กับตัวแปรอิสระอื่นสูง (Multicollinearity) ให้ออกจากสมการ

Tolerance[1] คือ ค่าที่ใช้กำหนดความน่าจะเป็นที่จะตรวจสอบความสัมพันธ์ของตัวแปรอิสระก่อนที่จะเข้าไปในสมการว่ามีความสัมพันธ์กับตัวแปรที่อยู่ในสมการก่อนแล้วหรือไม่ เพื่อขจัดปัญหาการเกิดการชับซ้อน (Multicollionearity) ซึ่งจะทำให้ค่าสัมประสิทธิ์การตัดสินใจ (r^2) สูงเกินความเป็นจริง Tolerance ของตัวแปรอิสระหนึ่ง ๆ คำนวณได้จากสัดส่วนของความแปรปรวนจากตัวแปรอิสระนั้น ๆ กับความแปรปรวนของตัวแปรอิสระที่อยู่ในสมการ ตัวแปรอิสระที่จะผ่านเกณฑ์เข้าไปในสมการก่อนจะต้องมีค่า Tolerance น้อยที่สุด ซึ่งแสดงว่ามีความสัมพันธ์กับตัวแปรอิสระที่อยู่ในสมการน้อยที่สุดนั่นเอง ในที่นี้จะกำหนดค่า Tolerance เท่ากับ 0.01

2. <u>การกำจัดตัวแปรแบบถอยหลัง</u> (Backward Elimination) การเลือกตามวิธีนี้จะนำตัวแปรอิสระทั้งหมดใส่เข้าไปในสมการ แล้วจะคัดตัวแปรอิสระออกจากสมการทีละตัวจนไม่สามารถคัดตัวแปรอิสระออกจากสมการได้อีก โดยใช้หลักเกณฑ์ของระดับนัยสำคัญ (Significant T; Sig.T) ที่กำหนด ใช้สัญญลักษณ์ POUT ซึ่งกำหนดไว้ = 0.10 จะคำนวณค่า Sig.T ของตัวแปรอิสระทีละตัว แล้วเปรียบเทียบ Sig.T ของแต่ละตัวแปรอิสระ Sig ของตัวแปรอิสระใดที่มีค่ามากที่สุด แสดงว่าโอกาสที่จะยอมรับสมมติฐาน $H_o : \beta_i = 0,\ i=1,2,\ldots,5$ มีมากที่สุดคือตัวแปรอิสระนั้นควรจะถูกตัดออกจากสมการ และจะคำนวณค่า Sig.T แต่ละตัวใหม่ โดยไม่นำค่าของตัวแปรอิสระที่ถูกตัดออกไปมาคำนวณ แล้วจึงเริ่มตัดตัวแปรที่มี Sig.T มากที่สุดออกไปอีก ทำเช่นนี้จนกว่าจะไม่มี Sig.T ของตัวแปรอิสระใดมีค่ามากกว่าหรือเท่ากับ POUT ที่กำหนดคือ 0.10 จึงหยุดการคัดเลือกตัวแปรอิสระออกจากสมการ

3. <u>การถดถอยแบบขั้นบันได</u> (Stepwise Regression) วิธีนี้เป็นวิธีที่นิยมใช้กันมากที่สุด เพราะเป็นวิธีที่รวมวิธีการถดถอยแบบไปข้างหน้า (Forward Selection) และการกำจัดตัว

[1] รศ.ชูศักดิ์ อุดมศรี, อ.ศิริชัย พงษ์วิชัย "SPSS^x" คณะพาณิชย์ศาสตร์ จุฬาฯ พ.ศ. 2528, หน้า 216

แปรแบบถอยหลัง (Backward Elimination) เข้าด้วยกัน โดยจะคัดเลือกตัวแปรเข้ามาในสมการตามวิธีของการถดถอยแบบไปข้างหน้า ตัวแปรอิสระที่จะเข้ามาในสมการจะต้องมีค่า Sig.T ≤ PIN ที่กำหนด และมีค่าความสัมพันธ์บางส่วนในระดับสูงด้วย (Highest Partial Correlation) ซึ่งตัวแปรอิสระใดที่มีค่าความสัมพันธ์บางส่วนสูงสุดก็จะได้รับการนำมาในสมการก่อน เมื่อคัดตัวแปรอิสระเข้ามาในสมการจนหมดแล้วก็จะพิจารณาดูว่าตัวแปรอิสระในสมการตัวใดที่จะต้องถูกกำจัดออกจากสมการตามวิธีการกำจัดตัวแปรแบบถอยหลัง ซึ่งวัดโดยค่า POUT ที่กำหนด ถ้าค่า Sig.T ที่อยู่ในสมการมีค่ามากกว่า POUT ที่กำหนดไว้ ตัวแปรอิสระนั้นก็จะต้องกำจัดออกจากสมการไปทำเช่นนี้จนกระทั่งไม่มีตัวแปรอิสระใดที่จะเข้าหรือออกจากสมการอีก ในการกำหนดค่า PIN และ POUT นั้นจะต้องกำหนดค่า PIN ให้ต่ำกว่าค่า POUT เพื่อป้องกันการคัดตัวแปรซ้ำ

สำหรับข้อมูลที่ใช้ในการวิเคราะห์นี้ ได้แสดงไว้ในตารางที่ 4.1 จากข้อมูลดังกล่าวได้นำเข้าโปรแกรม SPSS^{x} เพื่อเลือกรูปแบบสมการการถดถอย ซึ่งผลที่ได้จากการคำนวณด้วยเครื่องคอมพิวเตอร์ (ภาคผนวก) สมการในวิธี Stepwise Regression เป็นสมการที่ดีที่สุดสำหรับข้อมูลชุดนี้ เนื่องจากค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) สูงกว่าวิธีอื่น และค่า Standard Error มีค่าต่ำที่สุดเมื่อเทียบกับวิธีอื่น ๆ ดังนั้น ผลของการวิเคราะห์จึงได้ดังนี้

ก. ตัวแปรอิสระที่ถูกเลือกเข้าในสมการถดถอยมี 1 ตัวคือ ปริมาณเงินให้กู้ยืม (X_5)

ข. ตัวแปรอิสระที่ถูกคัดออกจากสมการถดถอยมี 4 ตัว

1. มูลค่าผลิตภัณฑ์ในประเทศทั้งสิ้น (X_1)
2. ปริมาณเงินทุนหมุนเวียน (X_2)
3. อัตราดอกเบี้ยเงินฝาก (X_3)
4. จำนวนสาขาของธนาคารพาณิชย์ในประเทศไทย (X_4)

ค. สมการถดถอยที่หาได้ มีรูปแบบดังนี้

$$Y = 9009.698646 + .88554\ (X_5)$$

Y = ปริมาณเงินฝากของธนาคารพาณิชย์ (ล้านบาท)

X_5 = ปริมาณเงินให้กู้ (ล้านบาท)

ง. ค่าสัมประสิทธิ์ของการตัดสินใจ (r^2) มีค่าเท่ากับ 0.9857 แสดงว่า การเปลี่ยนแปลงไปของปริมาณเงินฝากของธนาคารพาณิชย์เป็นผลสืบเนื่องมาจากการเปลี่ยนแปลงของปริมาณเงินให้กู้ยืมถึง 98.57%

จ. ทิศทางความสัมพันธ์ สังเกตุได้จากเครื่องหมายของค่าสัมประสิทธิ์ซึ่งจากการวิเคราะห์พบว่า มีค่าเท่ากับ 0.99285 และเครื่องหมายเป็น + แสดงว่าการเปลี่ยนแปลงไปของปริมาณเงินฝากของธนาคารพาณิชย์จะมีทิศทางไปในทิศทางเดียวกันกับการเปลี่ยนแปลงไปร่วมกันของปริมาณของเงินให้กู้ยืม กล่าวคือจะเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงไปตามกัน

ฉ. ความผิดพลาดจากการนำสมการถดถอยไปใช้ในการกะประมาณ ถ้านำสมการนี้ไปกะประมาณเงินฝากของธนาคารพาณิชย์ความผิดพลาดจะอยู่ระหว่าง ±3,551.99080 ล้านบาท

ช. การทดสอบความมีนัยสำคัญทางด้านความสัมพันธ์ ตัวสถิติที่ใช้ในการทดสอบคือ F โดยค่า F ที่คำนวณได้จากสมการถดถอยเท่ากับ 1244.44 เมื่อเทียบกับค่า F ที่ได้จากการเปิดตาราง ณ ระดับนัยสำคัญ 0.05 โดยมี Degree of Freedom เป็น (1,18) มีค่าเท่ากับ 4.41 ค่า F ที่คำนวณได้จากสมการถดถอยจึงมีค่าสูงกว่าค่า F ที่ได้จากการเปิดตาราง แสดงว่าปฏิเสธสมมติฐาน H_o และยอมรับว่าเงินฝากของธนาคารพาณิชย์มีความสัมพันธ์กับปริมาณเงินให้กู้ยืม

<u>การตีความหมายจากสมการถดถอย</u>

ผลการวิเคราะห์ข้อมูลระหว่างปี พ.ศ. 2507-2526 ทำให้ทราบว่า ปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงของปริมาณเงินฝากของธนาคารพาณิชย์ ได้แก่ ปริมาณเงินให้กู้ ซึ่งมีรูปแบบสมการถดถอยดังนี้

$$Y = 9009.69864 + .88554\ (X_5)$$

Y คือ ปริมาณเงินฝากของธนาคารพาณิชย์

X_5 คือ ปริมาณเงินให้กู้ยืม

จากสมการข้างต้น ทำให้ทราบความสัมพันธ์ระหว่างปริมาณเงินฝากของธนาคารพาณิชย์กับตัวแปรอิสระ ซึ่งในข้อมูลนี้มีเพียงตัวเดียวคือ ปริมาณเงินให้กู้ยืม (X_5) โดยมีค่าสัมประสิทธิ์เป็น 0.88554 หมายความว่า ปริมาณเงินให้กู้ยืมมีความสัมพันธ์ไปในทางเดียวกับปริมาณเงินฝากของธนาคารพาณิชย์ กล่าวคือ ถ้าปริมาณเงินกู้เพิ่มขึ้น 1 ล้านบาท ปริมาณเงินฝากของธนาคารพาณิชย์จะเพิ่มขึ้นประมาณ 885,540.- บาท เมื่อปัจจัยอื่น ๆ คงที่ ซึ่งเป็นไปตามทฤษฎีการสร้างเงินฝากของธนาคารพาณิชย์ ปริมาณเงินที่ธนาคารให้กู้ยืมไปนั้นเป็นการสร้างเงินฝากขั้นสอง เมื่อธนาคารให้กู้เงินไป 1 ล้านบาท เงินนั้นจะถูกนำไปใช้จ่ายต่าง ๆ และส่วนหนึ่งที่รั่วไหลออกไปจากระบบธนาคาร

ส่วนที่เหลือก็จะกลับเข้ามาเป็นเงินฝากของธนาคาร มีผลทำให้ปริมาณเงินฝากของธนาคารเพิ่มขึ้น

<u>การทดลองนำสมการถดถอยไปใช้ในการพยากรณ์</u>

จากสมการถดถอยที่หามาได้นั้น จะนำมาพยากรณ์ปริมาณเงินฝากสำหรับปี พ.ศ. 2527 โดยใช้ค่าจริงของปริมาณเงินให้กู้ยืมปี 2527 ผลของการพยากรณ์เป็นดังนี้

Y_c = 9009.698646 + .88554 (134,964.96)

Y_c = 128,526.57

Y = 139,439.78

$Y-Y_c$ = 10,913.21

$\frac{Y-Y_c}{Y} \times 100$ = 7.83%

เมื่อ Y_c คือ ปริมาณเงินฝากของธนาคารพาณิชย์ที่คำนวณได้จากสมการถดถอย

Y คือ ปริมาณเงินฝากของธนาคารพาณิชย์ที่เกิดขึ้นจริง

$Y-Y_c$ คือ ความแตกต่างระหว่างปริมาณเงินฝากของธนาคารพาณิชย์ที่เกิดขึ้นจริงกับที่คำนวณได้จากสมการถดถอย

$\frac{Y-Y_c}{Y} \times 100$ คือ จำนวนร้อยละของความแตกต่างระหว่างปริมาณเงินฝากของธนาคารพาณิชย์ที่เกิดขึ้นจริงกับที่คำนวณได้จากสมการถดถอย

จะเห็นได้ว่า การพยากรณ์ปริมาณเงินฝากของธนาคารพาณิชย์จากสมการถดถอยที่หามาได้นั้นมีความผิดพลาดร้อยละ 7.43 จากที่เกิดขึ้นจริง อย่างไรก็ตาม ในการนำสมการถดถอยที่หาได้ไปใช้ในการพยากรณ์ จะทำได้ก็ต่อเมื่อต้องทราบค่าของตัวแปรอิสระที่เข้าอยู่ในสมการเสียก่อน ดังนั้นจึงได้นำวิธีการพยากรณ์อีกแบบหนึ่งมาใช้

<u>การวิเคราะห์อนุกรมเวลาแบบบ๊อกซ์และเจนกินส์</u>

การพยากรณ์โดยวิธีบ๊อกซ์และเจนกินส์ (Box & Jenkins) นี้ เป็นการพยากรณ์โดยอาศัย

หลักสถิติทางด้านอนุกรมเวลาเข้ามาช่วยในการประมาณการ เป็นการวิเคราะห์ข้อมูลอนุกรมเวลาแผนใหม่ (Modern Time Serries Analysis) ซึ่งในการวิเคราะห์นี้ไม่จำเป็นต้องกำหนดรูปแบบของสมการขึ้นก่อน แต่รูปแบบต่าง ๆ จะถูกกำหนดขึ้นมาเองในระหว่างการวิเคราะห์ โดยอาศัยฟังก์ชั่นสหสัมพันธ์ในตัวเอง (Autocorrelation Function) และค่าฟังก์ชั่นสหสัมพันธ์ในตัวเองบางส่วน (Partial Autocorrelation Function)

<u>รูปแบบต่าง ๆ ของอนุกรมเวลา</u>

รูปแบบอนุกรมเวลาจำแนกออกได้เป็น 3 ประเภท คือ

1. รูปแบบ Autoregressive (AR) มีลักษณะสำคัญดังนี้ ค่าของตัวแปร ณ เวลาใดก็ตามจะมีความสัมพันธ์โดยตรงกับค่าของตัวแปรนั้นในช่วงเวลาที่ผ่านมาแล้วซึ่งรูปแบบโดยทั่วไปของ Autoregressive ถ้าเขียนในรูปของ Autoregressive of Order p (AR(p)) จะเป็นดังนี้

$$(\tilde{y}_t-\mu) = \phi_1(\tilde{y}_{t-1}-\mu) + \phi_2(\tilde{y}_{t-2}-\mu) + \text{---} \phi_p(\tilde{y}_{t-p}-\mu) + e_t$$

หรือ $$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \text{---} \phi_p y_{t-p} + e_t$$

p คือ อันดับของรูปแบบ Autoregressive ซึ่งบอกให้ทราบว่าค่าของตัวแปร ณ เวลาหนึ่งมีความสัมพันธ์กับค่าของตัวแปรในอดีตที่ผ่านมากี่หน่วยเวลา

$\phi_1, \phi_2 \text{---} \phi_n$ คือ ค่าพารามิเตอร์ของรูปแบบ

ในทางทฤษฎีฟังก์ชั่นสหความสัมพันธ์ในตัวเอง (Autocorrelation Function) ของโมเดล AR จะมีค่าค่อย ๆ ลดลงเมื่อระยะเวลาผ่านไป

2. รูปแบบ Moving Average (MA) มีลักษณะที่สำคัญคือ ค่าของตัวแปร ณ เวลาใดก็ตามจะมีความสัมพันธ์โดยตรงกับค่าความคลาดเคลื่อนในการกะประมาณเทอมแปรปรวนสุ่ม (Disturbance Terms) ในช่วงเวลาที่ผ่านมาแล้ว ซึ่งรูปแบบโดยทั่วไปของ Moving Average ถ้าเขียนในรูปของ Moving Average of Order q (MA(q))

$$(\tilde{y}_t-\mu) = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \text{---} - \theta_q e_{t-q}$$

หรือ $$y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

q คือ อันดับของรูปแบบ Moving Average ซึ่งบอกให้ทราบว่าค่าของตัวแปร ณ เวลาหนึ่งมีความสัมพันธ์กับเทอมแปรปรวนสุ่มในอดีตที่ผ่านมากี่หน่วยเวลา

$\theta_1, \theta_2, \cdots \theta_q$ คือ ค่าพารามิเตอร์ของรูปแบบ

ในทางทฤษฎีฟังก์ชั่นสหสัมพันธ์ในตัวเอง (Autocorrelation Function) ของโมเดล MA จะมีค่าเป็นศูนย์เมื่อพ้นหน่วยเวลาที่เท่ากับอันดับของโมเดล MA

จะเห็นได้ว่ารูปแบบ AR และ MA มีความสัมพันธ์กัน โดยที่รูปแบบ AR ทุก ๆ อันดับสามารถเขียนในรูปของ MA (∞) ได้ และรูปแบบของ MA ก็สามารถเขียนในรูปของ AR (∞) ได้เช่นกัน แต่ในทางปฏิบัติจะพยายามเลือกรูปแบบที่ง่ายที่สุดโดยคำนึงถึงรูปแบบที่มีจำนวนพารามิเตอร์น้อยที่สุดในกรณีที่โมเดลนั้นแสดงถึงสิ่งเดียวกัน เมื่อพิจารณาถึงสหสหสัมพันธ์ในตัวเองบางส่วน (Partial Autocorrelation) จะพบว่าสหสัมพันธ์ในตัวเองบางส่วนของรูปแบบ AR มีค่าเป็นศูนย์เมื่อพ้นหน่วยเวลาที่มีค่าเท่ากับอันดับของรูปแบบ AR นั้น แต่สำหรับสหสัมพันธ์ในตัวเองบางส่วนของรูปแบบ MA จะมีค่าค่อย ๆ ลดลงไปตามเวลาที่ผ่านไป ตามรูปที่ 4.1

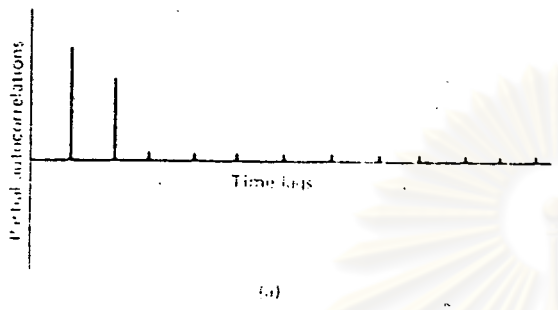
3. รูปแบบ Mixed Autoregressive - Moving Average (ARMA) จะมีลักษณะร่วมกันของ 2 รูปแบบที่กล่าวมาข้างต้น คือ ค่าของตัวแปร ณ เวลาใดก็ตามจะมีความสัมพันธ์โดยตรงกับค่าของตัวแปรนั้นและเทอมแปรปรวนสุ่มในช่วงเวลาที่ผ่านมาแล้ว รูปแบบของ ARMA ถ้าเขียนในรูปของ ARMA (p,q) โดยที่ค่า p ในวงเล็บแทนค่า Order ของ AR และค่า q ใช้แทน Order ของ MA จะเป็นดังนี้

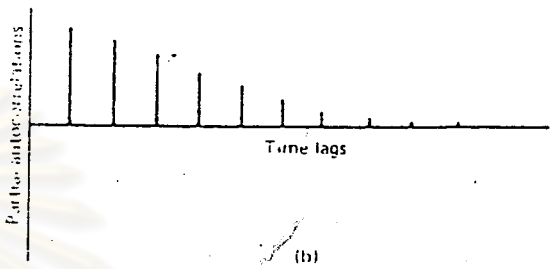
$$(\tilde{y}_t - \mu) - \phi_1(\tilde{y}_{t-1} - \mu) - \cdots - \phi_p(\tilde{y}_{t-p} - \mu) = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

หรือ $$y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} = e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

ในกรณีที่ $\theta_1 = \theta_2 = \cdots = \theta_q = 0$ รูปแบบ ARMR (p,q) ก็คือ

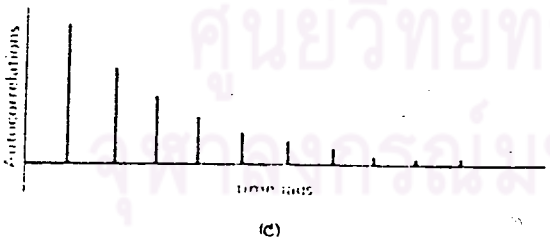
รูปที่ 4.1 แสดงความสัมพันธ์ของรูปแบบ AR และ MA

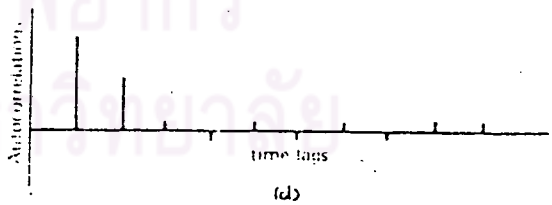




(a) แสดงสหสัมพันธ์ในตัวเองบางส่วนของรูปแบบ AR มีค่าเป็นศูนย์เมื่อพ้นหน่วยเวลาที่มีค่าเท่ากับอันดับของรูปแบบ AR นั้น

(b) แสดงสหสัมพันธ์ในตัวเองบางส่วนของรูปแบบ MA จะมีค่าค่อย ๆ ลดลงไป ตามเวลาที่ผ่านไป





(c) แสดงสหสัมพันธ์ในตัวเองของรูปแบบ AR ซึ่งจะมีค่าค่อย ๆ ลดลงไปเมื่อเวลาผ่านไป

(d) แสดงสหสัมพันธ์ในตัวเองของรูปแบบ MA จะมีค่าเป็นศูนย์เมื่อพ้นหน่วยเวลาที่มีค่าเท่ากับอันดับของรูปแบบ MA นั้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปแบบ AR(p) และถ้า $\phi_1 = \phi_2 \text{ --- } \phi_p = 0$ รูปแบบ ARMA (p,q) ก็คือรูปแบบ MA (q)

รูปแบบอนุกรมเวลาของ AR, MA และ ARMA ทั้ง 3 รูปแบบที่ได้อธิบายข้างต้นนั้น เป็นรูปแบบที่อยู่ภายใต้ข้อสมมติฐานว่ามาจาก Stationary Process แต่บางครั้งจะพบว่าข้อมูลบางชุดเป็นแบบ Non-Stationary Process ดังนั้นจะต้องแปลงข้อมูลนั้นให้เป็น Stationary เสียก่อน ซึ่งวิธีการแปลงข้อมูลก็คือ การหาผลต่าง (In-tegrated) ในทางปฏิบัติพบว่า การหาผลต่างเพียง 1 หรือ 2 ครั้ง ก็จะมีผลทำให้ข้อมูลนั้นเป็น Stationary ได้ ดังนั้น จะได้รูปแบบอีกแบบหนึ่งที่เรียกว่า Autoregressive In-tegrated Moving Average เขียนเป็นตัวย่อว่า ARIMA ซึ่งถ้าเขียนในรูปของ ARIMA of Order (p, d, q) แล้วค่า p จะแทนค่าของ AR d จะเป็นค่า In-tegrated และ q จะแทนค่าของ MA

ฉะนั้น ในวิธีของ Box & Jenkins Time-Series จึงมักใช้ ARIMA Model ในการหารูปแบบต่าง ๆ ของอนุกรมเวลา Y_t โดยกำหนดค่าของ p, d และ q ซึ่งถ้าหากข้อมูลเบื้องต้นมีลักษณะเป็น Stationary อยู่แล้ว ค่า d ก็จะเท่ากับ 0 แต่ถ้าเป็นลักษณะของ Non-stationary ค่า d ก็จะมีค่าเป็น 1 หรือ 2 ตามลำดับ เช่น

ถ้า p = 1, d = 0, q = 0 ก็คือ ARIMA (1,0,0)

ถ้า p = 0, d = 0, q = 1 ก็คือ ARIMA (0,0,1)

ถ้า p = 0, d = 1, q = 0 ก็คือ ARIMA (0,1,1)

ถ้า p = 1, d = 2, q = 0 ก็คือ ARIMA (1,2,0)

ถ้า p = 1, d = 1, q = 1 ก็คือ ARIMA (1,1,1,)

รูปแบบต่าง ๆ ของอนุกรมเวลาที่กล่าวมานี้ จะเป็นวิธีที่จะนำไปใช้ในการกำหนดรูปแบบขั้นต้น (Identification) ตามขั้นตอนที่ 1 ของการพยากรณ์ด้วยวิธี Box & Jenkins ซึ่งวิธีการของแต่ละขั้นตอนจะเป็นดังนี้

1. กำหนดรูปแบบขั้นต้น (Identification) ในขั้นนี้จะพิจารณาถึงค่าของ Auto-correlation function (ρ_k) ซึ่งค่าของ ρ_k จะไม่มีหน่วย แต่มีค่าอยู่ระหว่าง $-1 \leq \rho_k \leq 1$ เมื่อ k มีค่าเท่ากับ 0, 1, 2, --- ในทางปฏิบัติแล้วเราไม่สามารถหาค่าที่แท้จริงของ ρ_k ได้ แต่จะหาค่าประมาณของ ρ_k ได้จากค่าสังเกตุตัวอย่าง (r_k) ดังนั้น จึงต้องหา

Autocorrelation coefficient (r_k)

จากข้อมูลตัวอย่างที่ต้องการศึกษา เมื่อพิจารณาความสัมพันธ์ของค่า r_k กับ k ของตัวอย่างแล้วต้องพิจารณาดูว่าเป็นแบบ Stationary หรือได้รับอิทธิพลจากฤดูกาลหรือไม่ ถ้าอนุกรมเวลานั้นมาจาก Non-stationary หรือได้รับอิทธิพลจากฤดูกาล ก็จะต้องทำการแปลงข้อมูลนั้นให้มีลักษณะเป็น Stationary เสียก่อน

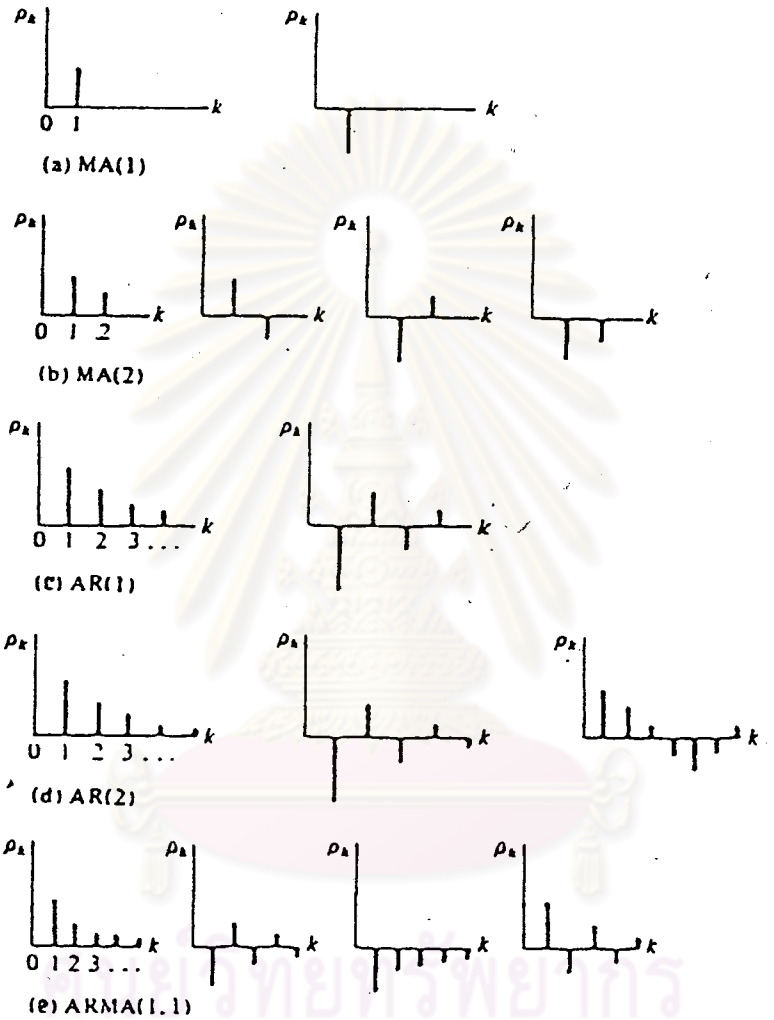
เมื่อข้อมูลอนุกรมเวลามีลักษณะเป็น Stationary แล้ว ก็จะนำไปเปรียบเทียบกับรูปแบบมาตรฐานของรูปแบบอนุกรมเวลาแบบต่าง ๆ (ρ_k) แล้วเลือกรูปแบบที่มีลักษณะคล้ายกับของอนุกรมเวลาชุดนั้น และในกรณีที่มีรูปแบบให้เลือกหลายแบบก็ให้เลือกรูปแบบที่มีค่าพารามิเตอร์น้อยที่สุดก่อน

รูปที่ 4.2 แสดงลักษณะของ ρ_k สำหรับรูปแบบลักษณะต่าง ๆ

2. ประมาณค่าพารามิเตอร์ (Parameter Estimation) เมื่อกำหนดรูปแบบได้แล้ว จะต้องคำนวณหาค่าพารามิเตอร์ทั้งหมดของรูปแบบนั้น ด้วยวิธีกำลังสองน้อยที่สุด ซึ่งในการคำนวณนี้จำเป็นต้องอาศัยเครื่องคอมพิวเตอร์ในการคำนวณ

3. ตรวจสอบความเหมาะสมของรูปแบบ (Diagnostic Checking) จะพิจารณาจากเทอมแปรปรวนสุ่มที่คำนวณได้จากโมเดลหลังจากแทนค่าอนุกรมเวลาและค่าพารามิเตอร์ในโมเดลนั้นแล้ว เทอมแปรปรวนสุ่มที่คำนวณได้จากโมเดลนี้จะขอเรียกว่า $\hat{e}_t$; t = 1, 2, ..., N การตรวจสอบว่ามีความเป็นอิสระต่อกันหรือไม่นั้น ทำได้โดยการเปรียบเทียบค่าสหสัมพันธ์ในตัวเองของ $\hat{e}_t$ ณ เวลาต่าง ๆ กับพิสัยแห่งความเชื่อถือ (Confidence Interval) ที่แต่ละหน่วยเวลานั้น ถ้าค่าสหสัมพันธ์ในตัวเองของ $\hat{e}_t$ ณ หลายหน่วยเวลาอยู่นอกพิสัยแห่งความเชื่อถือแสดงว่าค่าประมาณของเทอมแปรปรวนสุ่มยังไม่เป็นอิสระต่อกัน การตรวจสอบค่าสหสัมพันธ์ในตัวเองของ $\hat{e}_t$; t = 1, 2, ... , N กับพิสัยแห่งความเชื่อถือที่แต่ละหน่วยเวลานี้ เป็นการตรวจสอบความเป็นอิสระของ $\hat{e}_t$ แบบรวมกระทำโดยการเปรียบเทียบผลบวกของค่าสหสัมพันธ์ในตัวเองยกกำลังสองของ $\hat{e}_t$ กับค่า Chi Square χ^2 (υ) ซึ่งมีค่าองศาอิสระ (Degree of Freedom) เท่ากับ υ ซึ่ง υ เท่ากับ k-r เมื่อ k คือจำนวนของ residual autocorrelation coefficient ที่เก็บมาศึกษา และ r คือจำนวนพารามิเตอร์ทั้งหมดของรูปแบบนั้น ๆ ถ้า χ^2 ที่คำนวณได้มีค่าน้อยกว่า $\chi^2_{1-\alpha}$ จากตารางสถิติด้วยองศาของความอิสระ k-r แล้ว แสดงว่ารูปแบบขั้นต้นที่เลือกมานั้นเหมาะสมแล้ว แต่ถ้า χ^2 ที่คำนวณมีค่ามากกว่าก็ต้องเลือกพิจารณาเลือกโมเดลอนุกรมเวลาใหม่ แล้วกระทำตามขั้นตอนที่ 2 และ 3 อีกจนกว่าจะได้โมเดลที่

รูปที่ 4.2 แสดงลักษณะของ ρ_k สำหรับรูปแบบลักษณะต่าง ๆ



รูปแบบ MA(1) ค่า $P_k = 0$ เมื่อ $k \geq 2$

รูปแบบ MA(2) ค่า $P_k = 0$ เมื่อ $k \geq 3$

สำหรับรูปแบบ AR(1) , AR(2) และ ARMA(1,1) ค่า ρ_k จะค่อยๆจางหายไป

ตารางที่ 4.3 ปริมาณเงินฝากของธนาคารพาณิชย์

หน่วย : ล้านบาท

พ.ศ.	ไตรมาสที่ 1	ไตรมาสที่ 2	ไตรมาสที่ 3	ไตรมาสที่ 4
2507	10,308.4	10,647.1	11,350.1	11,798.4
2508	12,088.9	12,418.5	13,363.3	14,141.1
2509	14,678.9	15,475.9	16,886.2	17,528.0
2510	18,219.9	18,795.8	19,687.1	20,545.4
2511	21,047.6	21,661.7	23,091.5	24,140.7
2512	24,714.2	25,522.6	26,629.9	27,666.6
2513	28,572.9	29,003.0	30,547.6	31,884.8
2514	33,209.4	34,485.9	36,013.1	37,758.6
2515	40,914.8	42,824.4	45,253.5	48,405.5
2516	51,209.0	54,413.0	56,321.3	58,744.7
2517	61,835.5	66,170.8	69,366.8	74,603.0
2518	78,272.0	81,258.8	84,246.3	87,698.3
2519	91,112.4	95,443.2	100,69.08	106,418.1
2520	113,522.2	118,772.6	124,364.3	131,079.9
2521	138,644.7	144,810.4	149,199.3	158,523.7
2522	162,673.5	166,141.1	165,555.5	174,624.4
2523	184,842.5	193,629.3	202,491.7	214,994.1
2524	225,834.9	235,217.4	239,373.8	257,698.3
2525	271,848.2	286,086.2	297,091.8	322,357.2
2526	346,400.6	364,826.7	375,711.4	406,035.6
2527	423,442.8	499,583.0	463,089.1	497,827.9

ที่มา : ธนาคารแห่งประเทศไทย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

เหมาะสม เมื่อได้ตรวจสอบรูปแบบที่เหมาะสมแล้ว การพยากรณ์ซึ่งเป็นสิ่งที่เราต้องการก็จะทำได้ง่าย เนื่องจากสามารถคำนวณหาได้ทันทีจากรูปแบบที่ได้มานั้น โดยใช้เครื่องคอมพิวเตอร์ช่วย

<u>การพยากรณ์ปริมาณเงินฝากของธนาคารพาณิชย์ด้วยวิธี Box & Jenkins</u>

ข้อมูลที่นำมาใช้ในการพยากรณ์ในวิธีนี้ เป็นปริมาณเงินฝากของธนาคารพาณิชย์ตั้งแต่ปี พ.ศ. 2507-2526 โดยกำหนดให้เป็นรายไตรมาส มีทั้งหมด 80 ค่าสังเกตุ ดังตารางที่ 4.3 เมื่อนำข้อมูลเข้าโปรแกรม SPSS^{X} เพื่อประมวลผลตามวิธีอนุกรมเวลา Box & Jenkins ในขั้นแรกจะคำนวณหาค่า Autocorrelation Function ซึ่งจะได้ผลดังรูปที่ 4.3 ค่า r_k ค่อย ๆ ลดลงทีละน้อยไม่หมดไปอย่างรวดเร็ว ดังนั้น จึงต้องหาผลต่างครั้งที่ 1 และ 2 ตามรูปที่ 4.4 และ 4.5 รูปแบบที่ได้จากการหาผลต่างครั้งที่ 2 เมื่อนำไปเปรียบเทียบกับรูปแบบมาตรฐานในรูปที่ 4.2 จะเห็นได้ว่าอนุกรมของข้อมูลชุดนี้มีลักษณะเหมือน MA(2) คือมีเด่นอยู่แท่งที่ 1 และ 2 ดังนั้น จึงกำหนดให้อนุกรมของข้อมูลชุดนี้มีรูปแบบเป็น MA(2) เมื่อได้รูปแบบแล้วขั้นต่อไปจะเป็นการประมาณค่าพารามิเตอร์ ซึ่งได้ผลตามรูปที่ 4.6 ดังนี้

1. ค่าพารามิเตอร์ MA Lag ที่ 1 = -.45625
   ค่าพารามิเตอร์ MA Lag ที่ 2 = -.70625

2. การตรวจสอบรูปแบบโดยเปรียบเทียบค่า Chi-Square ที่คำนวณได้กับค่าที่ได้จากการเปิดตารางสถิติ มีผลดังนี้

Degree of Freedom	Chi-SQ ที่คำนวณได้	Chi-SQ ที่ได้จากการเปิดตาราง
4	6.81	9.49
10	18.68	18.30
16	25.98	26.30
22	31.67	33.90
23	31.77	35.20

จะเห็นได้ว่า ค่า Chi-Square ที่คำนวณได้จากค่าความคลาดเคลื่อนกะประมาณของข้อมูลอนุกรมเวลาชุดนี้มีค่าน้อยกว่าค่าที่ได้จากการเปิดตารางสถิติ แสดงว่ารูปที่เลือกนั้นสามารถนำไปใช้ในการพยากรณ์ได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3. สมการอนุกรมเวลาที่ได้จากการแทนค่าพารามิเตอร์ เป็นดังนี้.-

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

$$Y_t = e_t - (-.45625)e_{t-1} - (-.70625)e_{t-2}$$

เมื่อ Y_t คือ ค่าพยากรณ์ของปริมาณเงินฝากของธนาคารพาณิชย์ ณ เวลาที่ t

θ_1 คือ ค่าพารามิเตอร์

e_t คือ ค่าความคลาดเคลื่อนในการกะประมาณ ณ เวลาที่ t

e_{t-1} คือ ค่าความคลาดเคลื่อนในการกะประมาณ ณ เวลาที่ t-1

e_{t-2} คือ ค่าความคลาดเคลื่อนในการกะประมาณ ณ เวลาที่ t-2

จากสมการข้างต้น เมื่อนำไปใช้ในการพยากรณ์ จะได้ผลการพยากรณ์ดังที่แสดงไว้ในรูปที่ 4.7

<u>เปรียบเทียบค่าที่คำนวณได้กับค่าที่เกิดขึ้นจริง โดยวิธีการพยากรณ์แบบ Box & Jenkins</u>

พ.ศ. 2527	ค่าที่เกิดขึ้นจริง	ค่าพยากรณ์	ผลต่าง	ผลต่าง %
ไตรมาสที่ 1	423,442.8	431,617	-8,174.2	1.93
ไตรมาสที่ 2	449,583.0	454,971	-5,388.0	1.20
ไตรมาสที่ 3	463,089.1	465,856	-2,766.9	0.60
ไตรมาสที่ 4	497,827.9	496,180	-1,647.9	0.33

จะเห็นได้ว่า ค่าพยากรณ์ที่ได้กับตัวเลขที่เกิดขึ้นจริงแตกต่างกันโดยเฉลี่ยประมาณ 1% ซึ่งถือว่าค่าพยากรณ์ที่ได้ใกล้เคียงกับตัวเลขที่เกิดขึ้นจริง และได้ผลดีกว่าค่าที่พยากรณ์ได้จากสมการถดถอย

<u>การวิเคราะห์อนุกรมเวลาแบบคลาสสิค</u> (Classical Time Serries Analysis)

อนุกรมเวลาแบบคลาสสิค เป็นอนุกรมเวลาอีกแบบหนึ่งที่นิยมใช้กันมากในการวิเคราะห์

รูปที่ 4.3 รูปแบบอนุกรมเวลาของปริมาณเงินฝาก

VARIABLE -- VAR SERIES LENGTH - 30
DEGREE OF NONSEASONAL DIFFERENCING 0 DEGREE OF SEASONAL DIFFERENCING 0

MEAN VALUE OF THE PROCESS
0.103270+06

STANDARD DEVIATION OF THE PROCESS
0.101500+06

AUTOCORRELATION FUNCTION FOR VARIABLE VAR
ALTOCORRELATIONS *
TWO STANDARD ERROR LIMITS .

LAG	AUTO. COEF.	STAND. ERR.	-1 .75 .5 .25 0 .25 .5 .75 1
1	0.936	0.108	
2	0.879	0.108	
3	0.821	0.107	
4	0.766	0.106	
5	0.715	0.105	
6	0.671	0.105	
7	0.627	0.104	
8	0.585	0.103	
9	0.545	0.103	
10	0.509	0.102	
11	0.472	0.101	
12	0.436	0.100	
13	0.402	0.100	
14	0.370	0.099	
15	0.339	0.098	
16	0.310	0.097	
17	0.282	0.096	
18	0.255	0.096	
19	0.227	0.095	
20	0.199	0.094	
21	0.170	0.093	
22	0.142	0.092	
23	0.115	0.092	
24	0.088	0.091	
25	0.063	0.090	

***** BOX-JENKINS ANALYSIS REQUIRES 5224 BYTES OF WORKSPACE FOR IDENTIFICATI

รูปที่ 4.4 รูปแบบที่ได้จากการหาผลต่างครั้งที่ 1

```
* * * * * * * * * * * * * * * * * * * * * * * A R I M A     A N A L Y S I S * * * * *

VARIABLE VAR    CONTAINS THE TIME SERIES

 DEGREE OF NONSEASONAL DIFFERENCING =    1

 DEGREE OF SEASONAL DIFFERENCING =    0

 SEASONAL SPAN =    4

MEAN VALUE OF THE PROCESS
  0.50092D+04

STANDARD DEVIATION OF THE PROCESS
  0.60227D+04

AUTOCORRELATION FUNCTION FOR VARIABLE VAR
AUTOCORRELATIONS *
TWO STANDARD ERROR LIMITS .

      AUTO.  STAND.
LAG   CORR.    ERR.  -1   -.75   -.5   -.25    0    .25   .5    .75    1
                     :----:-----:-----:-----:----:-----:-----:-----:
  1   0.684  0.110                          .    :    .                *
  2   0.665  0.109                          .    :    .              *
  3   0.638  0.108                          .    :    .              *
  4   0.568  0.108                          .    :    .              *
  5   0.450  0.107                          .    :    .         *
  6   0.420  0.106                          .    :    .       *
  7   0.406  0.105                          .    :    .       *
  8   0.468  0.105                          .    :    .         *
  9   0.295  0.104                          .    :    .  *
 10   0.305  0.103                          .    :    .  *
 11   0.292  0.102                          .    :    .  *
 12   0.321  0.102                          .    :    .  *
 13   0.215  0.101                          .    :    *
 14   0.177  0.100                          .    :    *
 15   0.174  0.099                          .    :   *.
 16   0.193  0.098                          .    :    *
 17   0.104  0.098                          .    :  * .
 18   0.127  0.097                          .    :  *.
 19   0.131  0.096                          .    :  *.
 20   0.173  0.095                          .    :   *.
 21   0.112  0.094                          .    :  * .
 22   0.110  0.094                          .    :  * .
 23   0.110  0.093                          .    :  * .
 24   0.091  0.092                          .    :  * .
 25   0.055  0.091                          .    : *  .
```

รูปที่ 4.5 รูปแบบที่ได้จากการหาผลต่างครั้งที่ 2

```
* * * * * * * * * * * * * * * * * * * * * * * A R I M A   A N A L Y S I S * *

VARIABLE VAR      CONTAINS THE TIME SERIES

 DEGREE OF NONSEASONAL DIFFERENCING =   1

 DEGREE OF SEASONAL DIFFERENCING =   1

 SEASONAL SPAN =   4

MEAN VALUE OF THE PROCESS
 -0.139200+04

STANDARD DEVIATION OF THE PROCESS
 0.261360+04

AUTOCORRELATION FUNCTION FOR VARIABLE VAR
AUTOCORRELATIONS *
TWO STANDARD ERROR LIMITS .

     AUTO.  STAND.
LAG  CORR.   ERR.  -1   -.75   -.5   -.25    0    .25   .5   .75   1
                   :----:----:----:----:----:----:----:----:
  1  0.523  0.112                          .    :    .    *    .
  2  0.368  0.112                          .    :    .  *      .
  3  0.108  0.111                          .    :  *      .
  4 -0.150  0.110                          .  *   :    .
  5  0.004  0.109                          .    *   .
  6 -0.069  0.109                          .   *:   .
  7 -0.171  0.108                          .  :  *  .
  8  0.289  0.107                          .    :    .  *
  9  0.192  0.106                          .    :    *
 10  0.216  0.105                          .    :    *
 11  0.046  0.104                          .    :*   .
 12  0.044  0.104                          .    :*   .
 13 -0.021  0.103                          .    *    .
 14 -0.042  0.102                          .   *:    .
 15  0.019  0.101                          .    *    .
 16 -0.012  0.100                          .    *    .
 17 -0.047  0.099                          .   *:    .
 18 -0.121  0.098                          .  *  :   .
 19 -0.157  0.096                          . *   :   .
 20 -0.020  0.097                          .    *    .
 21  0.000  0.096                          .    *    .
 22  0.115  0.095                          .    :  *      .
 23  0.147  0.094                          .    :    *    .
 24  0.121  0.093                          .    :  *      .
 25  0.053  0.092                          .    :*   .

*** BOX-JENKINS REQUIRES        5224 BYTES OF WORKSPACE FOR IDENTIFICATION ***
```

รูปที่ 4.6 การประมาณค่าพารามิเตอร์

```
NONLINEAR ESTIMATION RESULTS

   PAR  LAG    ESTIMATE        STD ERROR      T RATIO

    MA    1    -.45625          .88899D-01    -5.1322
    MA    2    -.70625          .88981D-01    -7.9371

COVARIANCE MATRIX OF THE ESTIMATES

   PAR  LAG
    MA    1    .79031D-02    .21691D-02
    MA    2    .21691D-02    .79176D-02

CORRELATION MATRIX OF THE ESTIMATES

   PAR  LAG
    MA    1   1.00000  0.27421
    MA    2   0.27421  1.00000

MEAN VALUE OF RESIDUAL SERIES
    0.24338D+02

STANDARD DEVIATION OF RESIDUAL SERIES
    0.19135D+04

VARIANCE OF RESIDUAL SERIES
    0.36614D+07

DIAGNOSTIC CHI-SQUARE STATISTICS FOR RESIDUAL SERIES OF VARIABLE VAR

  LAG   CHI-SQ.   D.F.   PROB.
    6     6.81      4   0.1462
   12    18.68     10   0.0445
   18    25.98     16   0.0543
   24    31.67     22   0.0832
   25    31.77     23   0.1051


16 FEB 86    DEPOSITS QUATERLY DATA
16:32:31     **C H U L A L O N G K O R N**  I B M - 3 0 3 1   O S / V S 1
```

รูปที่ 4.7 ผลการพยากรณ์ปริมาณเงินฝาก

FORECASTS FOR VARIABLE VAR WITH ORIGIN AT 80 AND 95.00% CONFIDENCE LIMITS

OBS	LOW CONF LIM	FORECAST	UPP CONF LIM	STAND ERROR
81	.427420+06	.431620+06	.435820+06	2141.8
82	.447550+06	.454970+06	.462390+06	3783.6
83	.454130+06	.465860+06	.477580+06	5980.7
84	.481350+06	.496180+06	.511010+06	7564.5
85	.501860+06	.521760+06	.541670+06	10154.
86	.520070+06	.545120+06	.570160+06	12774.
87	.525070+06	.556000+06	.586930+06	15779.
88	.550450+06	.586520+06	.622200+06	18298.
89	.569040+06	.611910+06	.654170+06	21561.
90	.586520+06	.635260+06	.684000+06	24864.
91	.590310+06	.646150+06	.701980+06	28483.
92	.614340+06	.676470+06	.738600+06	31692.

GRAPHIC DISPLAY OF FORECASTS FOR VARIABLE VAR

DEFINITIONS OF SYMBOLS
DATA = *
FORECASTS AT LEAD 1 = 1
ESTIMATED 95% CONFIDENCE LIMITS = .
FORECAST FUNCTION = 0
OVERLAP = X

OBS.	DATA		1.00+05 3.00+05 5.00+05 7.00+05 9.00+05
41	51835.5	:	X
42	66170.8	:	X
43	69366.8	:	X.
44	74603.0	:	.X
45	78272.0	:	X
46	81258.8	:	X
47	84246.3	:	X
48	87698.3	:	X.
49	91112.4	:	.X
50	95443.2	:	.X
51	100691.	:	X
52	106418.	:	X.
53	113522.	:	.X
54	118773.	:	X
55	124364.	:	X
56	131080.	:	XX
57	138645.	:	X
58	144810.	:	X.
59	149119.	:	XX
60	158524.	:	X
61	162674.	:	X.
62	166141.	:	X.
63	165556.	:	X.
64	174624.	:	.X
65	184843.	:	X
66	193689.	:	.X
67	202492.	:	X
68	214994.	:	X
69	225835.	:	X.
70	235217.	:	.X
71	239374.	:	X
72	257698.	:	.X
73	271848.	:	XX
74	286086.	:	X.
75	297092.	:	.X
76	322357.	:	X
77	346401.	:	X
78	364827.	:	X.
79	375711.	:	X
80	406036.	:	X
81	431617.	F :	.X
82	454971.	F :	.X
83	465856.	F :	X.
84	496180.	F :	.0.
85	521761.	F :	.0.
86	545116.	F :	.0 .
87	556001.	F :	. 0.
88	586325.	F :	.0 .
89	611906.	F :	. 0 .
90	635260.	F :	. 0 .
91	646145.	F :	. 0 .
92	676469.	F :	. 0 .

*** BOX-JENKINS REQUIRES 10440 BYTES OF WORKSPACE FOR FORECASTING ***

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

อนุกรมเวลาแบบนี้จะพิจารณาถึงส่วนประกอบต่าง ๆ ของข้อมูลในอนุกรมเวลาชุดนั้น โดยหาว่าส่วนประกอบอะไรบ้างที่ได้ถูกผสมผสานกันขึ้นมาเป็นอนุกรมเวลา และวัดผลที่สืบเนื่องมาจากปรากฏการณ์ของส่วนประกอบเหล่านี้ จากนั้นจึงนำเอาผลที่ได้ไปใช้เป็นส่วนประกอบในการวางแผนงานในอนาคต

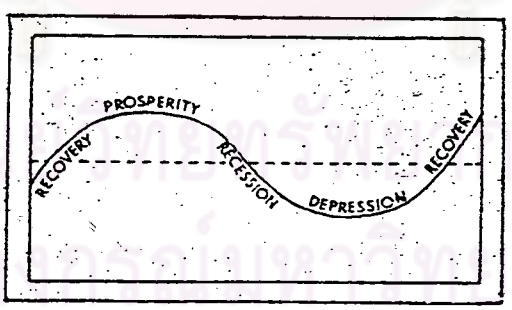
ส่วนประกอบของอนุกรมเวลา ได้แก่

1. ค่าแนวโน้ม (Trend หรือ Secular Trend ; T) เป็นการเคลื่อนไหวในระยะเวลาที่ค่อนข้างยาวนาน ซึ่งแสดงถึงทิศทางของอนุกรมเวลาชุดนั้น

2. การเปลี่ยนแปลงตามฤดูกาล (Seasonal Variation ; S) การผันผวนตามฤดูกาล โดยปกติมักจะเกิดขึ้นซ้ำ ๆ กันทุก ๆ ระยะเวลาเดียวกัน ซึ่งในการพิจารณาความเคลื่อนไหวตามฤดูกาลนี้ หน่วยของระยะเวลาอาจเป็นงวด 3 เดือน รายเดือน รายสัปดาห์ รายวัน หรือแม้กระทั่งรายชั่วโมงก็ได้

3. การเปลี่ยนแปลงตามวัฏจักร (Cyclical Variation ; C) วัฏจักรทางธุรกิจโดยทั่วไปประกอบด้วย 4 ระยะคือ ระยะเวลารุ่งเรือง (Prosperity) ระยะเวลาฝืดเคือง (Recession) ระยะเวลาตกต่ำ (Depression) และฟื้นตัว (Recovery)

รูปที่ 4.8 วัฏจักรทางธุรกิจ



4. การเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ (Irregular Variation ; I) เป็นเหตุการณ์ที่เกิดขึ้นโดยไม่อาจคาดการณ์ได้ล่วงหน้า และมีผลต่อข้อมูลอนุกรมเวลาชุดนั้น

ข้อมูลอนุกรมเวลาชุดหนึ่ง ๆ อาจกล่าวได้ว่าเป็นผลหรืออิทธิพลจากปัจจัยทั้ง 4 ที่กล่าวมาข้างต้น จึงอาจเขียนเป็นสมการได้ ดังนี้

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

เมื่อนำข้อมูลปริมาณเงินฝากจากตารางที่ 4.3 ซึ่งมีทั้งหมด 80 ค่าสังเกตุมาประมวลผลตามวิธีอนุกรมเวลาแบบ Classical โดยใช้โปรแกรมสำเร็จรูป TSER 3 จะได้ผลการพยากรณ์ตามรูปที่ 4.9 (รายละเอียดแสดงในภาคผนวก ข.)

<u>เปรียบเทียบค่าที่คำนวณได้กับค่าที่เกิดขึ้นจริง</u>

<u>พ.ศ. 2527</u>	<u>ค่าที่เกิดขึ้นจริง</u>	<u>ค่าพยากรณ์</u>	<u>ผลต่าง</u>	<u>ผลต่าง %</u>
ไตรมาสที่ 1	423,442.8	264,925.72	158,517.08	37.44
ไตรมาสที่ 2	449,583.0	267,542.50	182,040.50	40.49
ไตรมาสที่ 3	463,087.1	271,138.07	191,951.03	41.45
ไตรมาสที่ 4	497,827.9	276,929.14	220,898.76	44.37

จะเห็นได้ว่า ค่าพยากรณ์ที่ได้กับตัวเลขที่เกิดขึ้นจริงแตกต่างกันโดยเฉลี่ยประมาณ 41%

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 4.9 ผลการพยากรณ์โดยวิธีอนุกรมเวลาแบบคลาสสิค

FORECAST FOR NEXT 12 OBSERVATIONS

OBSERVATION	TREND VALUE	SEASONAL INDEX	FORECAST
81	264174.739	1.0023437	264925.734
82	268147.751	.997742841	267542.499
83	272100.764	.996389767	271138.073
84	276093.777	1.00302563	276929.135
85	280066.79	1.00284276	280862.954
86	284039.804	.997742841	283399.681
87	288012.817	.996389767	286972.735
88	291985.83	1.00302563	293869.271
89	295958.843	1.00284276	296800.184
90	299931.856	.997742841	299254.862
91	303904.869	.996389767	302807.398
92	307877.882	1.00302563	308809.407

CHOOSE BY NUMBER FROM THE FOLLOWING :

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย