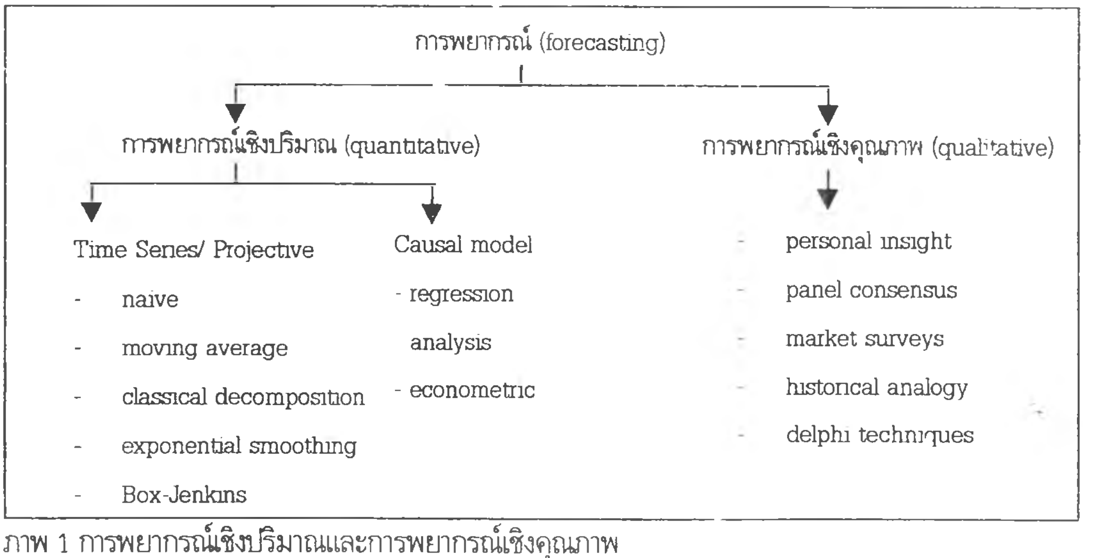
บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การเปรียบเทียบความคลาดเคลื่อนของวิธีพยากรณ์ด้วยวิธีของเบย์ (Bayesian method) วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) และวิธีการพยากรณ์ร่วม (combined forecasting) ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีการเปลี่ยนแปลงและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยได้ทำการศึกษาค้นคว้าจาก หนังสือ เอกสาร บทความ และงานวิจัยต่าง ๆ ที่เกี่ยวข้องนำเสนอโดยแบ่งออกเป็น 7 ตอน ประกอบด้วย ตอนที่ 1 ความรู้ทั่วไปเกี่ยวกับการพยากรณ์ ตอนที่ 2 วิธีการพยากรณ์ข้อมูลอนุกรมเวลา ตอนที่ 3 การพยากรณ์ด้วยวิธีของเบย์ ตอนที่ 4 การพยากรณ์ด้วยวิธีบ็อกซ์-เจนกิ้นส์ ตอนที่ 5 วิธีการพยากรณ์ร่วม ตอนที่ 6 เกณฑ์ในการเปรียบเทียบความสามารถของวิธีการพยากรณ์ ตอนที่ 7 งานวิจัยที่เกี่ยวข้อง รายละเอียดแต่ละตอนมีดังต่อไปนี้

ตอนที่ 1 ความรู้ทั่วไปเกี่ยวกับการพยากรณ์

การพยากรณ์ แบ่งได้ 2 กลุ่ม คือ การพยากรณ์เชิงคุณภาพ (qualitative forecasting) และการพยากรณ์เชิงปริมาณ (quantitative forecasting) ซึ่งการพยากรณ์เชิงปริมาณจำแนกตามแนวคิดได้เป็น 2 ประเภท คือ ประเภทที่ 1 มีแนวคิดว่าพฤติกรรมในอดีตเพียงพอที่จะพยากรณ์พฤติกรรมในอนาคตได้ หรือเรียกว่าการวิเคราะห์อนุกรมเวลา (time series) ซึ่งมีวิธีการพยากรณ์ ได้แก่ วิธีง่าย (naive) วิธีการเฉลี่ยเคลื่อนที่ (moving average) วิธีการแยกส่วนประกอบ (classical decomposition) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) และวิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins) ประเภทที่ 2 มีแนวคิดว่าพฤติกรรมของสิ่งที่จะพยากรณ์มีสิ่งอื่น ๆ ซึ่งมีความสัมพันธ์บางลักษณะกับสิ่งที่จะพยากรณ์เป็นตัวกำหนด หรือเรียกว่าโมเดลเชิงสาเหตุ (causal model) ได้แก่ วิธีการวิเคราะห์การถดถอย (regression annalysis) และวิธีการพยากรณ์เชิงเศรษฐมิติ (econometric) ดังแสดงในแผนภาพ 1



ภาพ 1 การพยากรณ์เชิงปริมาณและการพยากรณ์เชิงคุณภาพ

วิธีการพยากรณ์มีมากมายหลายวิธี แต่ละวิธีมีคุณลักษณะ ข้อดีและข้อเสียแตกต่างกัน ไม่มีวิธีใดดีและสมบูรณ์ที่สุด (Makridakis, Wheelwright and Hyndman, 1998) ดังนั้นจึงควรคำนึงถึงปัจจัยในการเลือกวิธีการเลือกวิธีพยากรณ์ ดังต่อไปนี้

1. ระยะเวลาในการพยากรณ์ (time horizon or lead time) หมายถึง ระยะเวลาที่ผู้พยากรณ์จะพยากรณ์เหตุการณ์ที่น่าสนใจแต่ละเหตุการณ์ล่วงหน้านับจากปัจจุบัน จะยาวนานเท่าใดนั้นขึ้นอยู่กับการนำค่าพยากรณ์นั้นไปใช้ ซึ่งแบ่งเป็น 4 ระยะ (O'Donovan, 1983) ดังนี้

1.1 ระยะใกล้ (immediate term) เป็นช่วงเวลาน้อยกว่า 1 เดือน

1.2 ระยะสั้น (short term) เป็นช่วงเวลาระหว่าง 1-3 เดือน

1.3 ระยะปานกลาง (medium term) เป็นเวลาระหว่าง 3 เดือนถึง 2 ปี

1.4 ระยะยาว (long term) เป็นช่วงเวลามากกว่า 2 ปี

2. ลักษณะโครงสร้างของข้อมูล (pattern of data) หมายถึง ลักษณะการเปลี่ยนแปลงของข้อมูลตามช่วงระยะเวลา เมื่อนำเอาข้อมูลที่ต้องการพยากรณ์มาพล๊อตกราฟ ข้อมูลจะแสดงรูปร่างลักษณะปรากฏออกมา ข้อมูลที่มีลักษณะต่าง ๆ ก็จะมีวิธีที่เหมาะสมสำหรับข้อมูลแต่ละชนิดนั้น ๆ สามารถแบ่งลักษณะข้อมูลได้ 4 ประเภท (O'Donovan, 1983; Makridakis, Wheelwright, and Hyndman, 1998 อ้างถึงใน ทัศนีย์ อินทนุ, 2543) ดังนี้

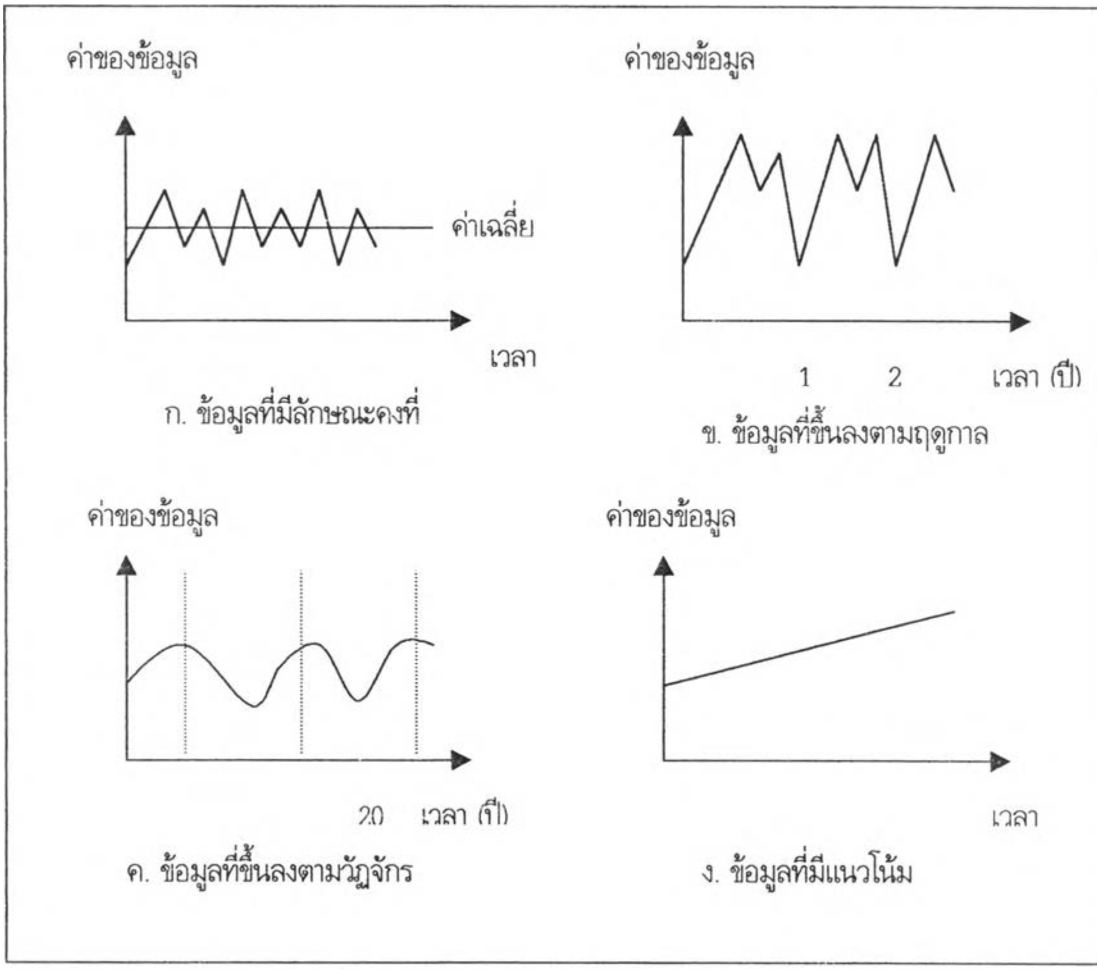
2.1 ข้อมูลที่สม่ำเสมอตามแนวนอนหรือมีลักษณะคงที่ (horizontal data pattern: H) เป็นข้อมูลที่มีลักษณะของรูปแบบไม่เพิ่มขึ้นหรือลดลงไปในทางใดทางหนึ่ง คือ ข้อมูลที่เกิดขึ้นจะอยู่ในภาวะสมดุลรอบ ๆ ค่าเฉลี่ย ความน่าจะเป็นที่จะเกิดข้อมูลมากกว่าค่านี้มีเท่ากับความน่าจะเป็นที่จะเกิดข้อมูลน้อยกว่าค่านี้ วิธีที่เหมาะสมกับข้อมูลประเภทนี้ได้แก่ วิธีการเฉลี่ยเคลื่อนที่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล เป็นต้น ดังภาพ 2ก

2.2 ข้อมูลที่ขึ้นลงตามฤดูกาล (seasonal data pattern: S) เป็นข้อมูลที่มีลักษณะเกิดขึ้นซ้ำ ๆ กันในช่วงเวลาสั้น ๆ เช่น 1 วัน 1 สัปดาห์ 1 เดือน ฯลฯ โดยเกิดการเคลื่อนไหวขึ้นลงซ้ำกันในช่วงเวลาเดียวกันซึ่งจะแสดงการเปลี่ยนแปลงอันเกิดจากฤดูกาลในลักษณะของแบบ (pattern) ที่เกิดขึ้นซ้ำกันในแต่ละช่วง วิธีที่เหมาะสมกับข้อมูลประเภทนี้ ได้แก่ วิธีบ็อกซ์-เจนกิ้นส์ ดังภาพ 2ข

2.3 ข้อมูลที่ขึ้นลงตามวัฏจักร (cyclical data pattern: C) มีลักษณะคล้ายฤดูกาลแต่ช่วงของวัฏจักรจะยาวกว่าฤดูกาลมาก วัฏจักรหนึ่งจะครอบคลุมระยะเวลาหลายปี ส่วนมากมักจะขึ้นลงตามวัฏจักรทางเศรษฐกิจ วิธีที่เหมาะสมกับข้อมูลประเภทนี้ ได้แก่ วิธีบ็อกซ์-เจนกิ้นส์ ดังภาพ 2ค

2.4 ข้อมูลที่มีแนวโน้ม (trend data pattern: T) เป็นข้อมูลที่มีการเพิ่มขึ้นหรือลดลงอย่างสม่ำเสมอตามระยะเวลา ค่อนข้างมีแบบแผน อาจมีลักษณะเป็นเส้นตรง เส้นโค้ง หรือลักษณะอื่น ๆ วิธีที่เหมาะสมกับข้อมูลประเภทนี้ ได้แก่ วิธีการวิเคราะห์การถดถอย และวิธีการพยากรณ์เชิงเศรษฐมิติ เป็นต้น ดังภาพ 2ง

ลักษณะข้อมูลทั้ง 4 ประเภทนี้ บ่อยครั้งที่พบว่ามีหลายลักษณะปนกันอยู่ หรือบางครั้งก็มีเพียงลักษณะใดลักษณะหนึ่งเท่านั้น



ภาพ 2 ลักษณะการเปลี่ยนแปลงข้อมูล

3. ความแม่นยำหรือความถูกต้อง (accuracy) ซึ่งเป็นวิธีการวัดความถูกต้องมี 2 แบบ คือ

3.1 วิธีทางสถิติที่มีการทดสอบนัยสำคัญ ได้แก่ การตรวจสอบความถูกต้องของพารามิเตอร์หรือสมการที่ใช้เป็นตัวแบบพยากรณ์ เช่น t-test, F-test ซึ่งใช้ในวิธีการวิเคราะห์การถดถอย เช่น ทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นตรง

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นตรง

ถ้าการทดสอบมีนัยสำคัญทางสถิติ แสดงว่า ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นตรง

3.2 วิธีที่ไม่ได้มีการทดสอบนัยสำคัญ ได้แก่ ค่าเบี่ยงเบนสมบูรณ์เฉลี่ย (mean absolute deviation: MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error: MSE) รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (root mean squarre error: RMSE) และเปอร์เซ็นความคลาดเคลื่อนสมบูรณ์เฉลี่ย (mean absolute percentage error: MAPE) เป็นต้น โดยวัดจากความแตกต่างของข้อมูลจริงกับค่าพยากรณ์ ถ้าค่าดังกล่าวมีค่าน้อย แสดงว่า วิธีการพยากรณ์มีความถูกต้อง

4. ค่าใช้จ่าย (cost) ซึ่งแบ่งได้เป็น 3 ประเภท คือ ค่าใช้จ่ายในการพัฒนาหรือนำเอาวิธีมาใช้ ค่าใช้จ่ายในการเก็บข้อมูล และค่าใช้จ่ายในการพยากรณ์ วิธีที่มีค่าใช้จ่ายสูง ได้แก่ วิธีบ็อกซ์-เจนกิ้นส์ และวิธีพยากรณ์เชิงเศรษฐมิติ ส่วนวิธีที่เสียค่าใช้จ่ายน้อย ได้แก่ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล

5. การนำไปประยุกต์ใช้ (application) การเลือกวิธีพยากรณ์ให้เหมาะสมกับสถานการณ์มีสิ่งที่ควรพิจารณา คือ ความง่ายต่อการเข้าใจ และความง่ายต่อการวิเคราะห์ข้อมูล วิธีการพยากรณ์ที่มีความซับซ้อนมักมีความถูกต้องสูง แต่ผู้บริหารที่นำค่าพยากรณ์ไปใช้ในการตัดสินใจอาจจำเป็นต้องใช้วิธีที่มีความซับซ้อนน้อยกว่า และมีความถูกต้องน้อยกว่า ปัจจัยที่สำคัญอีกประการหนึ่ง คือ ความเข้าใจต่อการพยากรณ์และความสามารถในการวิเคราะห์ ผู้ที่จะนำค่าพยากรณ์ไปใช้ควรเข้าใจวิธีพยากรณ์นั้น ๆ เป็นอย่างดี เข้าใจถึงความสามารถของวิธีนั้น และนำไปประยุกต์ใช้ได้ วิธีที่ง่ายต่อความเข้าใจและตีความหมาย ได้แก่ วิธีหาค่าเฉลี่ยเคลื่อนที่ ส่วนวิธีที่ยากต่อความเข้าใจและตีความ ได้แก่ วิธีบ็อกซ์-เจนกิ้นส์ และวิธีการพยากรณ์เชิงเศรษฐมิติ เป็นต้น

จากปัจจัยในการเลือกวิธีพยากรณ์ดังกล่าว สามารถเปรียบเทียบการพยากรณ์ได้ดังตาราง 1 (O'Donovan, 1983 และทรงศิริ แต้สมบัติ, 2539)

ตาราง 1 การเปรียบเทียบวิธีพยากรณ์

ปัจจัย	วิธีการพยากรณ์								
	N	MA	SES	HWS	D	BJ	REG	ECO	MBJ
ระยะเวลาในการพยากรณ์									
-ระยะใกล้	/	-	/	/	/	/	-	-	-
-ระยะสั้น	-	/	/	/	/	/	/	/	/
-ระยะปานกลาง	-	/	-	-	-	-	/	/	/
-ระยะยาว	-	-	-	-	-	-	/	/	/
ลักษณะของข้อมูล									
-ลักษณะคงที่	/	/	/	-	/	/	-	-	/
-ฤดูกาล	-	-	-	/	/	/	/	/	/
-วัฏจักร	-	-	-	-	/	-	/	/	/
-แนวโน้ม	-	-	-	/	/	/	/	/	-
ขนาดของข้อมูลที่ต้องการ	น้อย	10	10	15	30	30	30	น้อย	60
แนวโน้ม (S-ช่วงฤดูกาล)				2(S)	6(S)	6(S)	6(S)		8(S)
ค่าใช้จ่าย									
-ต่ำ	/	/	/	/	-	-	-	-	-
-ปานกลาง	-	-	-	-	/	-	/	-	-
-สูง	-	-	-	-	-	/	-	/	/

หมายเหตุ : N - วิธีง่าย
MA - การเฉลี่ยเคลื่อนที่
SES - วิธีปรับให้เรียบเอ็กซ์โปเนนเชียลแบบง่าย
HWS - วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์
D - การแยกส่วนประกอบ
BJ - วิธีบ็อกซ์-เจนกิ้นส์
REG - การวิเคราะห์การถดถอย
ECO - การพยากรณ์เชิงเศรษฐมิติ
MBJ - วิธีบ็อกซ์-เจนกิ้นส์แบบพหุ

ตอนที่ 2 วิธีการพยากรณ์ข้อมูลอนุกรมเวลา

ข้อมูลอนุกรมเวลา หมายถึง ข้อมูลที่ได้จากค่าสังเกตชุดหนึ่ง ซึ่งถูกจัดเรียงกันตามลำดับการเกิดขึ้นก่อนหลัง โดยที่ข้อมูลชุดนี้จะถูกเก็บมา ณ ช่วงเวลาที่ห่างเท่า ๆ กัน

ข้อมูลอนุกรมเวลามีส่วนประกอบแบ่งออกเป็น 4 ส่วน (O'Donovan, 1983; ทรงศิริ แต้สมบัติ, 2539; กัลยา วานิชย์บัญชา, 2542) ได้แก่

1. แนวโน้ม (trend: T) เป็นข้อมูลที่มีการเปลี่ยนแปลงในระยะเวลาที่นานพอที่จะเห็นแนวโน้มการเคลื่อนไหวของอนุกรมเวลา ซึ่งอาจเป็นแนวโน้มขึ้นหรือลง (upward หรือ downward trend) แนวโน้มมีลักษณะเป็นไปได้ต่าง ๆ กัน เช่น แนวโน้มเส้นตรง (linear trend) แนวโน้มเส้นโค้ง (quadratic trend) แนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential trend) และแนวโน้มแบบตัว S (S-shape trend) เป็นต้น

2. อิทธิพลของฤดูกาล (seasonal: S) การเคลื่อนไหวมีผลเนื่องจากฤดูกาล การเคลื่อนไหวจะเกิดขึ้นซ้ำแล้วซ้ำอีกในช่วงเวลาหนึ่งส่วนใหญ่จะเป็นหนึ่งปี ปัจจัยที่ก่อให้เกิดอิทธิพลของฤดูกาลมีได้หลายปัจจัย เช่น สภาพอากาศ อุณหภูมิ วัฒนธรรม กำหนดการตามปฏิทินที่หน่วยงานกำหนดขึ้น อนุกรมเวลาที่ใช้ในการพิจารณาอิทธิพลของฤดูกาลมักจะเป็นอนุกรมเวลารายเดือน หรือรายไตรมาส ที่มีการเก็บรวบรวมไว้อย่างน้อย 2 ปีขึ้นไป ซึ่งค่าวัดอิทธิพลของฤดูกาลขึ้นอยู่กับรูปแบบในการรวมตัวของแนวโน้มและฤดูกาล ถ้ารูปแบบเป็นแบบบวกจะเรียกว่า ค่าวัดอิทธิพลของฤดูกาล (seasonal factor) หรือถ้ารูปแบบเป็นแบบคูณจะเรียกว่า ค่าดัชนีฤดูกาล (seasonal index)

3. อิทธิพลของวัฏจักร (cyclical: C) อนุกรมเวลาที่เก็บรวบรวมในระยะยาวหลายปี การเคลื่อนไหวอาจจะแสดงอิทธิพลของวัฏจักรที่มีลักษณะทำนองเดียวกันกับอิทธิพลของฤดูกาล โดยวัฏจักรหนึ่งจะครอบคลุมระยะเวลาหลายปี แต่ละช่วงจะมีการเคลื่อนไหวไม่แตกต่างกันมากนัก

4. เหตุการณ์ที่ผิดปกติหรือความผันแปรที่ไม่แน่นอน (irregular variation: I) เป็นการเคลื่อนไหวของอนุกรมเวลาเฉพาะส่วนที่ไม่มีแบบแผนที่แน่นอน เหตุการณ์ผิดปกตินี้ส่วนใหญ่จะเป็นเหตุการณ์ที่ไม่ได้คาดคิดมาก่อนหรือไม่เกิดบ่อยครั้ง เช่น น้ำท่วม พายุ อุบัติเหตุ ปฏิวัติ สงคราม การนัดหยุดงาน ข่าวลือ เป็นต้น รวมถึงปัจจัยอื่น ๆ ที่ไม่ใช่เนื่องจากแนวโน้ม อิทธิพลของฤดูกาล และอิทธิพลของวัฏจักร

ข้อมูลอนุกรมเวลาแต่ละชุด ไม่จำเป็นต้องประกอบไปด้วยส่วนประกอบทั้ง 4 ส่วน และในกรณีที่เป็นข้อมูลรายปี จะไม่มีอิทธิพลของฤดูกาล (ทรงศิริ แต้สมบัติ, 2539)

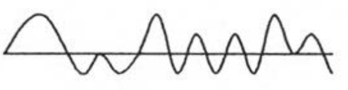
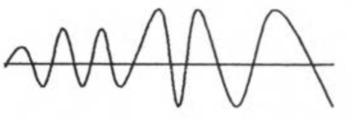
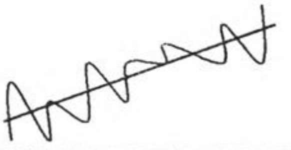
การรวมส่วนประกอบของอนุกรมเวลาเป็นได้ทั้งการรวมตัวแบบบวก (additive model) และการรวมตัวแบบคูณ (multiplicative model) ซึ่งรูปแบบการรวมตัวแบบบวก คือ

$$y = T + S + C + I$$

และรูปแบบการรวมตัวแบบคูณ คือ

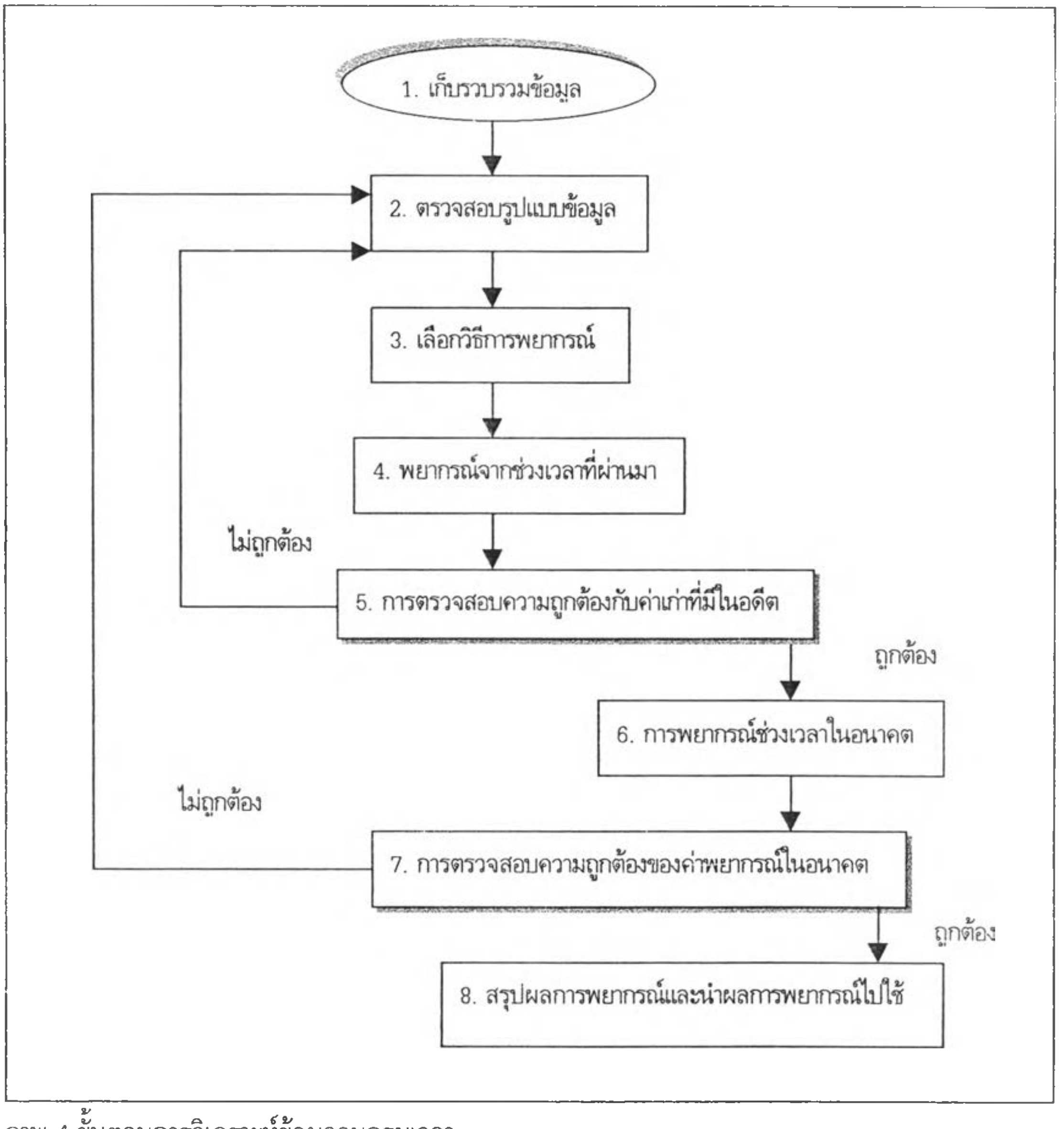
$$y = T \times S \times C \times I$$

การรวมตัวแบบบวกลักษณะการเคลื่อนไหวของอนุกรมเวลาเฉพาะในส่วนของอิทธิพลของฤดูกาลในแต่ละปีจะไม่มีการเปลี่ยนแปลง แต่ถ้าเป็นการรวมตัวแบบคูณ อิทธิพลของฤดูกาลในแต่ละปีจะมีการแกว่งมากขึ้นหรือน้อยลง ดังภาพ 3

การรวมตัว / ลักษณะแนวโน้ม	แบบบวก	แบบคูณ
ไม่มีแนวโน้ม		
เส้นตรง		

ภาพ 3 การรวมตัวของแนวโน้มและฤดูกาลแบบบวกและแบบคูณ

วิธีการวิเคราะห์อนุกรมเวลาเพื่อใช้พยากรณ์ข้อมูลโดยทั่วไป มีขั้นตอนสรุปได้ 8 ขั้นตอน (Hanke and Reitsc,1992 อ้างถึงในบำเพ็ญ ปิดชิด, 2540) ขั้นตอนแรกเป็นการเก็บรวบรวมข้อมูล ขั้นตอนที่ 2 เมื่อเก็บรวบรวมข้อมูลหรือค่าสังเกตที่ผ่านมาในอดีตจากช่วงเวลา t_1 ถึงเวลา t_n ได้แล้ว นำข้อมูลที่ได้มาตรวจสอบรูปแบบของข้อมูล ขั้นตอนที่ 3 เป็นการเลือกวิธีพยากรณ์ที่เหมาะสมกับลักษณะของข้อมูล ขั้นตอนที่ 4 เป็นการพยากรณ์ข้อมูลจากช่วงเวลาที่ผ่านมา โดยใช้ข้อมูลจากช่วงเวลา t_1- t_n เป็นฐานพยากรณ์จากช่วงเวลา $t_{n+1...}$ t_n ขั้นตอนที่ 5 เป็นการตรวจสอบผลการพยากรณ์จากข้อมูลจริงในช่วงเวลา $t_{n+1...}$ t_n ถ้านักวิจัยเห็นว่าผลการพยากรณ์ไม่ถูกต้องจะต้องนำข้อมูลไปตรวจสอบโมเดล และเลือกวิธีพยากรณ์ใหม่อีกครั้งหนึ่ง ตามขั้นตอนที่ 1-5 แต่ถ้านักวิจัยยอมรับผลการพยากรณ์ว่าถูกต้อง ก็ดำเนินต่อไปตามขั้นตอนที่ 6 คือการพยากรณ์ช่วงเวลาในอนาคต t_{n+1}, t_{n+2}... ขั้นตอนที่ 7 เป็นการตรวจสอบผลการพยากรณ์ ถ้านักวิจัยเห็นว่าผลการพยากรณ์ไม่ถูกต้อง จะต้องนำข้อมูลไปตรวจสอบโมเดลและเลือกวิธีพยากรณ์ใหม่ตามขั้นตอนที่ 1-6 อีกครั้งหนึ่ง แต่ถ้านักวิจัยยอมรับผลการพยากรณ์ว่าถูกต้อง จึงสรุปผลการพยากรณ์และนำผลการพยากรณ์ไปใช้ในการตัดสินใจ ขั้นตอนทั้ง 8 ขั้นตอน ดังภาพ 4



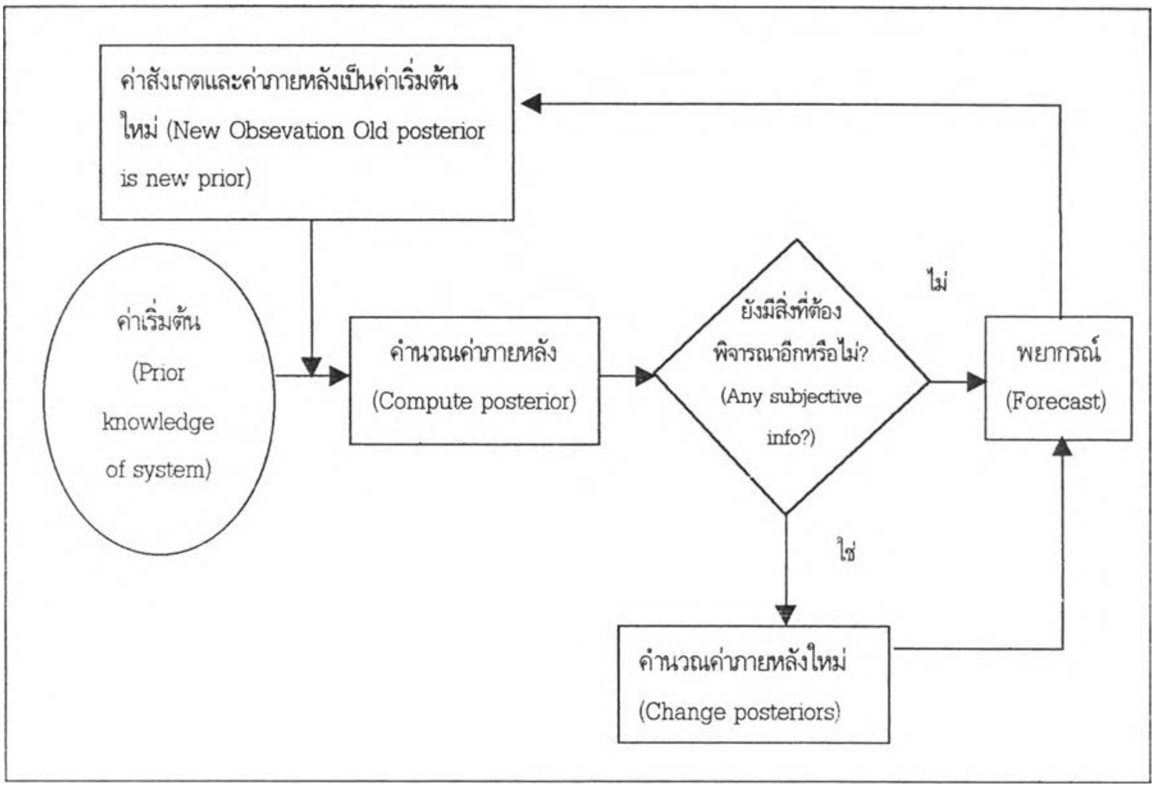
ภาพ 4 ขั้นตอนการวิเคราะห์ข้อมูลอนุกรมเวลา

ตอนที่ 3 การพยากรณ์ด้วยวิธีของเบย์

นอกจากจะมีการนำวิธีของเบย์ (Bayesian method) มาใช้ค่อนข้างมากในการประยุกต์ใช้วิธีการทางสถิติ เช่น การประมาณค่าพารามิเตอร์ การทดสอบสมมติฐาน เป็นต้น แล้วยังมีการนำเอาวิธีของเบย์มาใช้ในการพยากรณ์ข้อมูลอนุกรมเวลาด้วย ซึ่งการนำข้อมูลอนุกรมเวลามาใช้ในการพยากรณ์ด้วยวิธีของเบย์ ทำได้ดังแผนภาพ 5 (Fildes, 1984)

การพยากรณ์ในยุคแรก ๆ มักไม่มีการนำข้อมูลในอดีตมาใช้จะมีเพียงการใช้ความเป็นอัตนัยหรือวิจารณญาณ (subjective) เป็นพื้นฐานในการตัดสินใจเลือกโมเดลการพยากรณ์ อย่างไรก็ตามเมื่อมีข้อมูลอนุกรมเวลาซึ่งข้อมูลอนุกรมเวลาดังกล่าวจะแทนลักษณะการเกิดของเหตุการณ์ต่าง ๆ จึงได้มีการปรับปรุงการพยากรณ์เพื่อลดความเสี่ยงในการตัดสินใจให้น้อยลง ตัวอย่างของการพยากรณ์ที่ใช้ข้อมูลอนุกรมเวลาเป็นพื้นฐานในการตัดสินใจเลือกโมเดลการพยากรณ์ เช่น การพยากรณ์ยอดขายสินค้าตามฤดูกาล

เป็นต้น วิธีการนำข้อมูลอนุกรมเวลามาใช้เพื่อลดความเสี่ยงในการตัดสินใจ จะเรียกว่าวิธีการของเบย์ (Bayesian methods) โดยสมมติว่าการพยากรณ์แบบดั้งเดิมเป็นเพียงการประมาณค่าพารามิเตอร์ในโมเดลการพยากรณ์ จะมีการปรับค่าประมาณของพารามิเตอร์ให้เหมาะสมได้ด้วยหลักการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ (Bayesian Estimation) ผู้วิจัยจะกล่าวถึงวิธีของเบย์ 3 ส่วน ได้แก่ ส่วนแรก ความรู้เบื้องต้นเกี่ยวกับการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ ส่วนที่ 2 การประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ และส่วนที่สาม การพยากรณ์ ดังรายละเอียดต่อไปนี้



ภาพ 5 The Bayesian forecasting model

ส่วนที่ 1 ความรู้เบื้องต้นเกี่ยวกับการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์

กำหนดให้ x เป็นตัวแปรสุ่มที่มีลักษณะฟังก์ชันความหนาแน่นน่าจะเป็น (probability density function) f และไม่ทราบค่าพารามิเตอร์ θ จะเขียนฟังก์ชันความหนาแน่นด้วย $f(x|\theta)$ ซึ่งแสดงถึงการแจกแจงความน่าจะเป็นของ x สมมติให้พารามิเตอร์ θ เป็นตัวแปรสุ่มที่มีฟังก์ชันความหนาแน่นน่าจะเป็น $h_0(\theta)$ ที่เรียกว่า ฟังก์ชันความหนาแน่นเบื้องต้น (prior density) ของ θ การแจกแจงความน่าจะเป็นนี้ใช้วัดความเป็นอัตนัยหรือวิจารณญาณของความเชื่อเกี่ยวกับ θ กล่าวคือ ถ้าค่อนข้างเชื่อมั่นเกี่ยวกับค่าของ θ จะเลือกการแจกแจงเบื้องต้น (prior distribution) ที่มีค่าความแปรปรวนต่ำ แต่ถ้าค่อนข้างไม่เชื่อมั่นเกี่ยวกับค่าของ θ จะเลือกการแจกแจงเบื้องต้นด้วยความแปรปรวนสูง

การประมาณค่าพารามิเตอร์ θ โดยวิธีของเบย์ทำได้โดยกำหนดฟังก์ชันความหนาแน่นเบื้องต้น $h_0(\theta)$ และให้ y เป็นข้อมูลที่เกิดขึ้นที่มีฟังก์ชันความหนาแน่นน่าจะเป็น $f(y|\theta)$ ขึ้นอยู่กับค่าพารามิเตอร์ θ

โดยทั่วไป y เป็นค่าสังเกตของกลุ่มตัวอย่าง $x_1, x_2, ..., x_t$ ซึ่งเป็นฟังก์ชันของ $x_1, x_2, ..., x_t$ ค่าประมาณพารามิเตอร์ของ θ จะขึ้นกับรูปแบบของการแจกแจงที่เปลี่ยนไป $h_1(\theta|y)$ ที่เรียกว่า การแจกแจงภายหลัง (posterior distribution) จากทฤษฎีของเบย์

$$h_1(\theta|y) = \frac{h_0(\theta)f(y|\theta)}{\int_\theta h_0(\theta)f(y|\theta)d\theta} = \frac{h_0(\theta)f(y|\theta)}{g(y)} \tag{1}$$

จะเห็นว่า $h_1(\theta|y)$ ได้จากการรวมความเป็นอัตนัยเกี่ยวกับ θ และข้อมูล y จะประมาณค่าพารามิเตอร์ θ ด้วย θ^* ซึ่งเป็น

$$\theta^* = \int_\theta \theta h_1(\theta|y)d\theta \tag{2}$$

จะกล่าวถึงตัวอย่างของวิธีของเบย์ในกรณีต่าง ๆ ดังนี้

กรณีที่ 1 เมื่อกำหนดตัวแปร y มีการแจกแจงปกติ ที่มีค่าเฉลี่ย θ และความแปรปรวน σ_y^2 ซึ่งทราบค่า σ_y^2 $[y \sim N(\theta, \sigma_y^2)]$ มีฟังก์ชันความหนาแน่นเป็น

$$f(y|\theta) = (2\pi\sigma_y^1)^{-1/2} \exp\left[-\frac{1}{2}\left(\frac{y-\theta}{\sigma_y}\right)^2\right] \quad ; -\infty < y < \infty \tag{3}$$

และกำหนดลักษณะการแจกแจงเบื้องต้นของ θ เป็นแบบปกติที่มีค่าเฉลี่ย $\bar{\theta}'$ และความแปรปรวน υ'_θ $[\theta \sim N(\bar{\theta}', \upsilon'_\theta)]$ นั่นคือ

$$h_0(\theta) = (2\pi\upsilon'_\theta)^{-1/2} \exp\left[-\frac{(\theta-\bar{\theta}')^2}{2\upsilon'_\theta}\right] \tag{4}$$

ฟังก์ชันความหนาแน่นน่าจะเป็นภายหลังของ θ เมื่อกำหนดว่า y เกิดขึ้น คือ

$$h_1(\theta|y) = \frac{2\pi(\upsilon'_\theta\sigma_y^2)^{-1/2}\exp\left\{\frac{1}{2}\left[(\theta-\bar{\theta}')^2/\upsilon'_\theta + (y-\theta)^2/\sigma_y^2\right]\right\}}{\int 2\pi(\upsilon'_\theta\sigma_y^2)^{-1/2}\exp\left\{-\frac{1}{2}\left[(\theta-\bar{\theta}')^2/\upsilon'_\theta + (y-\theta)^2/\sigma_y^2\right]\right\}d\theta}$$

$$= \left(2\pi\frac{\upsilon'_\theta\sigma_y^2}{\upsilon'_\theta+\sigma_y^2}\right)^{-1/2} \exp\left\{-\frac{1}{2}\frac{\left[\theta-(y\upsilon'_\theta+\bar{\theta}'\sigma_y^2)/(\upsilon'_\theta+\sigma_y^2)\right]^2}{\upsilon'_\theta\sigma_y^2/(\upsilon'_\theta+\sigma_y^2)}\right\}$$

นั่นคือการแจกแจงภายหลังของ θ เป็นแบบปกติที่มีค่าเฉลี่ย

$$\bar{\theta}'' \equiv E(\theta|y) = \frac{y\upsilon'_\theta + \bar{\theta}'\sigma_y^2}{\upsilon'_\theta + \sigma_y^2} \tag{5}$$

และความแปรปรวน

$$\upsilon''_\theta \equiv V(\theta|y) = \frac{\upsilon'_\theta\sigma_y^2}{\upsilon'_\theta + \sigma_y^2} \tag{6}$$

พบว่า การแจกแจงภายหลังเป็นแบบปกติ เมื่อการแจกแจงเบื้องต้นเป็นแบบปกติ และทราบค่าความแปรปรวนในกระบวนการสุ่มปกติ

กรณีที่ 2 สมมติว่าอนุกรมเวลาอธิบายได้โดยโมเดลคงที่ (the constant model)

$$y_t = b + \varepsilon_t \tag{7}$$

ซึ่ง b คือ ค่าเฉลี่ยที่ไม่ทราบค่า และ ε_t คือ องค์ประกอบสุ่ม ที่สมมติให้มีการแจกแจงปกติด้วยค่าเฉลี่ย 0 และทราบค่าความแปรปรวน σ_ε^2 $[\varepsilon_t \sim N(0,\sigma_\varepsilon^2)]$

ดังนั้น y_t ณ ช่วงเวลา t มีการแจกแจงความน่าจะเป็น (probability distribution) $N(b,\sigma_\varepsilon^2)$ เมื่อทราบค่าความแปรปรวนนั่นคือจะหาฟังก์ชัน $f(y_t|b)$ ได้ ต้องการประมาณค่าพารามิเตอร์ b ด้วยวิธีของเบย์

สมมติจุดเริ่มต้นในกระบวนการพยากรณ์ หรือที่เวลา 0 กำหนดการแจกแจงเบื้องต้นของ b เป็น $N(\bar{b}',\upsilon'_b)$ หรือฟังก์ชันความหนาแน่นเบื้องต้นเป็น $h_0(b)$

เมื่อทราบค่าสังเกต y_1 จะนำค่า y_1 ไปปรับปรุงค่าเฉลี่ย $\bar{b}'$ และความแปรปรวน υ'_b โดยการแจกแจงภายหลังทราบค่า y_1 เป็น $N[\bar{b}''(1),\upsilon''_b(1)]$

ซึ่ง $$\bar{b}''(1) \equiv E(b|y_1) = \frac{x_1\upsilon'_b + \bar{b}'\sigma_\varepsilon^2}{\upsilon'_b + \sigma_\varepsilon^2}$$

และ $$\upsilon''_b(1) \equiv V(b|y_1) = \frac{\upsilon'_b\sigma_\varepsilon^2}{\upsilon'_b + \sigma_\varepsilon^2}$$ มีฟังก์ชันความหนาแน่นภายหลัง เป็น $h_1(b|y_1)$

เมื่อทราบค่า x_2 จะนำค่า x_2 ไปปรับปรุงค่าเฉลี่ย $\bar{b}''(1)$ และความแปรปรวน $\upsilon''_b(1)$ เปลี่ยนจาก $h_1(b|y_1)$ ไปเป็น $h_2(b|y_1,y_2)$ ดังนี้

$$h_2(b|y_1,y_2) = \frac{h_1(b|y_1)f(y_2|b)}{\int_b h_1(b|y_1)f(y_2|b)db}$$

โดยการแจกแจงภายหลังทราบค่า y_1 และ y_2 เป็น $N[\bar{b}''(2),\upsilon''_b(2)]$

ซึ่ง $$\bar{b}''(2) \equiv E(b|y_1,y_2) = \frac{\bar{x}\upsilon'_b + \bar{b}'(\sigma_\varepsilon^2/2)}{\upsilon'_b + (\sigma_\varepsilon^2/2)}$$

และ $$\upsilon''_b(2) \equiv V(b|y_1, y_2) = \frac{\upsilon'_b\sigma_\varepsilon^2}{2\upsilon'_b + \sigma_\varepsilon^2}$$

ซึ่ง $\bar{y} = (y_1 + y_2)/2$ มีฟังก์ชันความหนาแน่นภายหลังทราบค่า y_1 และ y_2 เป็น $h_2(b|y_1, y_2)$ จะพิสูจน์ได้ว่า $h_2(b|y_1, y_2) = h_2(b|\bar{y})$ เพราะ $\bar{x}$ เป็นค่าสถิติที่เพียงพอ (sufficient statistics) สำหรับการประมาณค่า b

ในทำนองเดียวกันเมื่อทราบค่า y_T อาจจะแสดงหลังจากสังเกต y_T ทำการคำนวณค่าภายหลังได้ ที่มีการแจกแจงเป็น $N[\bar{b}''(T), \upsilon''_b(T)]$

ซึ่ง $$\bar{b}''(T) = \frac{\bar{x}\upsilon'_b + \bar{b}'(\sigma_\varepsilon^2/T)}{\upsilon'_b + (\sigma_\varepsilon^2/T)} \tag{8}$$

และ $$\upsilon''_b(T) = \frac{\upsilon'_b\sigma_\varepsilon^2}{T\upsilon'_b + \sigma_\varepsilon^2} \tag{9}$$

$$\bar{x} = \frac{\sum_{t=1}^{T} x_t}{T}$$

ดังนั้น ตัวประมาณค่า b ด้วยวิธีของเบย์ หลังจากช่วงเวลาที่ T คือ $b^*(T) = \bar{b}''(T)$ โดย

$$b^*(T) = \frac{T}{n' + T}\bar{y} + \frac{n'}{n' + T}\bar{b}' \tag{10}$$

ซึ่ง $n' = \sigma_\varepsilon^2/\upsilon'_b$ จะเห็นว่าตัวประมาณค่า b ด้วยวิธีของเบย์ คือ ค่าเฉลี่ยน้ำหนักของค่าเฉลี่ยกลุ่มตัวอย่าง และ $\bar{y}$ ค่าเริ่มต้น $\bar{b}'$

นอกจากนี้ $b^*(T)$ จากสมการ (10) จะเขียนได้ใหม่เป็น

$$b^*(T) = \alpha x_T + (1-\alpha)b^*(T-1) \tag{11}$$

ซึ่ง $$\alpha = \alpha(T) = \frac{1}{n' + T} = \frac{\upsilon'_b}{T\upsilon'_b + \sigma_\varepsilon^2} \tag{12}$$

จากสมการ (11) แสดงให้เห็นว่าค่าของ y_T ได้จากการทำให้เป็นปัจจุบันที่แต่ละช่วงเวลาซึ่งมีรูปแบบเช่นเดียวกับสมการที่ได้จากวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) จากสมการ (12) จะเห็นว่า α เป็นฟังก์ชันของเวลา T ที่มีค่าต่ำลงเมื่อเวลา T เพิ่มขึ้น จาก $\upsilon''_b(T) = \alpha(T)\sigma_\varepsilon^2$ ทำให้สรุปได้ว่าความไม่แน่นอนในการประมาณค่า b จะลดลงถึง 0 เมื่อ T เข้าใกล้อนันต์ (infinitely large) ในทำนองเดียวกัน น้ำหนักของ $\bar{b}'$ ที่เป็นค่าเริ่มต้นจะลดลงเมื่อเวลา T เพิ่มขึ้น อย่างไรก็ตามน้ำหนักที่สอดคล้องจะได้จากข้อมูลอนุกรมเวลาที่เกิดขึ้นจริง ดังนั้นที่ช่วงเวลาจะพบว่ามีข้อมูลเพียงพอในการลดความแปรปรวนของค่าภายหลังในระดับต่ำตามที่ต้องการ

ในกรณีที่ไม่ทราบค่าความแปรปรวน σ_ε^2 จะกำหนด σ_ε^2 ให้เป็นตัวแปรสุ่มเหมือนกับค่า b แล้วนิยามการแจกแจงความน่าจะเป็นร่วมกันเบื้องต้น (a prior joint probabilit; distribution) ของ b และ σ_ε^2 การคำนวณค่าภายหลังเพื่อให้เป็นปัจจุบันจะใช้ประโยชน์จากทั้งค่าเฉลี่ยของกลุ่มตัวอย่างและความแปรปรวนของกลุ่มตัวอย่าง จากค่าสังเกต T ในกรณีที่ทราบค่าความแปรปรวน มีข้อแนะนำเกี่ยวกับค่าเริ่มต้นและรายละเอีย:ที่ยวกับการคำนวณค่าภายหลังไว้ Raiffa และ Schlaifer (1961 อ้างถึงใน Montgomery, Johnson, and Gardiner, 1990)

สมการพยากรณ์ที่ได้จากการแจกแจงภายหลังที่เวลา $T+\tau$ จะเขียนเป็น

$$\hat{x}_{T+\tau}(T) = b^*(T) \tag{13}$$

เนื่องจากความไม่แน่นอนในการประมาณค่าของ b วัดโดยความแปรปรวนภายหลัง $\upsilon_b^*(T)$ ความแปรปรวนค่าพยากรณ์ $\hat{x}_{T+\tau}(T)$ คือ

$$V[\hat{x}_{T+\tau}(T)] = \upsilon_b^* \tag{14}$$

จะมีความแปรปรวนของความคลาดเคลื่อนในการพยากรณ์ คือ

$$V[e_\tau(T+\tau)] = V[x_{T+\tau} - \hat{x}_{T+\tau}(T)] = \sigma_\varepsilon^2 + \upsilon_b^* \tag{15}$$

ความแปรปรวนของความคลาดเคลื่อนในการพยากรณ์ดังกล่าวเป็นอิสระกันสำหรับการพยากรณ์ล่วงหน้า τ ช่วงเวลาจะคำนวณขีดจำกัดของ $100(1-\alpha)$ ช่วงความเชื่อมั่นสำหรับสำหรับ $x_{T+\tau}$ คือ

$$b^*(T) \pm Z_{\alpha/2}\sqrt{\sigma_\varepsilon^2 + \upsilon_b^*} \tag{16}$$

ซึ่ง $Z_{\alpha/2}$ คือ $100(\alpha/2)$ เปอร์เซ็นต์ไทล์ของการแจกแจงปกติมาตรฐาน

ส่วนที่ 2 การประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์

ในตอนนี้ผู้วิจัยจะเสนอการประมาณค่าพารามิเตอร์ด้วยวิธีของเบย์ และโมเดลการพยากรณ์อนุกรมเวลาเชิงเส้น พารามิเตอร์ที่ไม่ทราบค่า คือ $\beta_1, \beta_2, \ldots, \beta_k$ การวิเคราะห์สำหรับแต่ละโมเดลใช้วิธีการวิเคราะห์ถดถอย (regression methods) โมเดลแนวโน้มที่นิยมใช้ ได้แก่

แนวโน้มเส้นตรง $y_t = b_1 + b_2 t + \varepsilon_t$

แนวโน้มแบบเส้นโค้ง $y_t = b_1 + b_2 t + b_3 t^2 + \varepsilon_t$

แนวโน้มแบบ exponential $y_t = b_1 b_2^{\,t}\,\varepsilon_t$

แนวโน้มและฤดูกาล $y_t = b_1 + b_2 t + b_3 \sin\frac{2\pi t}{12} + b_4 \cos\frac{2\pi t}{12} + \varepsilon_t$

กำหนดโมเดลอนุกรมเวลา คือ $y_t = b_1 z_1(t) + b_2 z_2(t) + \ldots + b_k z_k(t) + \varepsilon_t$

$$= \sum_{i=1}^{k} b_i z_i(t) + \varepsilon_t \tag{17}$$

ซึ่ง $\{b_i\}$ เป็นเซตของค่าคงที่

$\{z_i(t)\}$ เป็นฟังก์ชันเชิงคณิตศาสตร์ของ t และเป็นตัวแปรอิสระในโมเดล

$\varepsilon_t = N(0, \sigma_\varepsilon^2)$ โดยสมมติให้ ε_t เป็นอิสระจาก ε_{t+j} สำหรับทุก ๆ j บ่อยครั้งที่ $z_1(t) = 1$ ดังนั้น b_1 เป็นค่าคงที่ในโมเดล

ในทางปฏิบัติจะเขียนโมเดลในรูปแบบเมทริกซ์ โดยนิยามให้

$$\mathbf{z}(t) = [z_1(t),\ z_2(t), \ldots, z_k(t)]^t$$

$$\mathbf{b} = [b_1, b_2, \ldots, b_k]^t$$

สัญลักษณ์ " t " คือ ทรานสโพสของเมทริกซ์ จะเขียนสมการ (17) เป็น

$$y_t = \mathbf{b}^t \mathbf{z}(t) + \varepsilon_t \tag{18}$$

การแจกแจงความน่าจะเป็นของ x_t คือ $N(\mathbf{b}^t\mathbf{z}(t), \sigma_\varepsilon^2)$ มีฟังก์ชันความหนาแน่นน่าจะเป็น $f\left(x_t \middle| \mathbf{b}, \sigma_\varepsilon^2\right)$

ขั้นตอนการพยากรณ์ประกอบด้วย การประมาณค่า $\mathbf{b}$ โดยสมมติว่าทราบค่า σ_ε^2 สำหรับในกรณีที่ไม่ทราบค่า ศึกษาเพิ่มเติมใน Raiffa and Schlaifer (1961 อ้างถึงใน Montgomery, Johnson, and Gardiner, 1990)

กำหนดการแจกแจงเบื้องต้น (Prior distribution) ของ $\mathbf{b}$ เป็น $N(\bar{\mathbf{b}}', \mathbf{V}')$ ซึ่ง $E(b_i) = \bar{b}_i'$, $Var(b_i) = \upsilon_{ii}'$ และ $Cov(b_i, b_j) = \upsilon_{ij}'$ ซึ่ง $\bar{\mathbf{b}}' = E(\mathbf{b})$ และ $\mathbf{V}'$ เป็น variance - covariance เมทริกซ์ของการแจกแจงเบื้องต้น นั่นคือ $\mathbf{b}$ มีการแจกแจงแบบพหุปกติ (the multivariate normal distribution) และมีฟังก์ชันเบื้องต้นเป็น

$$h_0(\mathbf{b}) = (2\pi)^{-(1/2)k} \left|\mathbf{V}'^{-1}\right|^{1/2} \exp\{-\frac{1}{2}[\mathbf{b} - \bar{\mathbf{b}}']^t \mathbf{V}'^{-1}[\mathbf{b} - \bar{\mathbf{b}}']\} \tag{19}$$

เพื่อความสะดวกนิยามเมทริกซ์ $\mathbf{G}'$ ดังนี้

$$\mathbf{G}' \equiv \sigma_\varepsilon^2 \mathbf{V}'^{-1} \tag{20}$$

ดังนั้น

$$\mathbf{G}'^{-1} = \frac{\mathbf{V}'}{\sigma_\varepsilon^2} \tag{21}$$

หลังจากช่วงเวลา T จะมีค่าสังเกต $y_1, y_2, \ldots, y_T$ และต้องการปรับปรุงการแจกแจงเบื้องต้นจาก

$$\mathbf{Y} = [y_1, y_2, \ldots, y_T]'$$

และ $$\mathbf{Z} = \begin{bmatrix} \mathbf{z}'(1) \\ \mathbf{z}'(2) \\ \vdots \\ \mathbf{z}'(T) \end{bmatrix} = \begin{bmatrix} z_1(1) & z_2(1) & \cdots & z_k(1) \\ z_1(2) & z_2(2) & \cdots & z_k(2) \\ \cdots & \cdots\vdots\cdots & \cdots & \cdots \\ z_1(T) & z_2(T) & \cdots & z_k(T) \end{bmatrix}$$

จะประมาณค่า **b** จากอนุกรมเวลาที่เกิดขึ้นจริง ด้วยวิธีกำลังสองน้อยที่สุด ได้สมการปกติ (normal equation)

$$\mathbf{Z}^t\mathbf{Z}\hat{\mathbf{b}} = \mathbf{Z}^t\mathbf{Y} \tag{22}$$

กำหนด $\mathbf{G} = \mathbf{Z}^t\mathbf{Z}$ และ $\mathbf{g} = \mathbf{Z}^t\mathbf{Y}$ จะเขียนสมการปกติได้ใหม่เป็น

$$\mathbf{G}\hat{\mathbf{b}} = \mathbf{g} \tag{23}$$

ซึ่ง $$\mathbf{E}(\hat{\mathbf{b}}|\mathbf{b}) = \mathbf{b} \tag{24}$$

และมี variance – covariance เมทริกซ์ คือ

$$Cov(\hat{\mathbf{b}}) \equiv \mathbf{V} = \mathbf{G}^{-1}\sigma_\varepsilon^2 \tag{25}$$

สมาชิกที่ i, j ของ $\mathbf{V}$ คือ $\upsilon_{ij} \equiv Cov(\hat{b}_i, \hat{b}_j)$ นั่นคือ เมื่อทราบค่า **b** และ σ_ε^2 แล้ว $\hat{\mathbf{b}}$ มีการแจกแจงปกติแบบพหุ $N(\mathbf{b}, \mathbf{V})$ ที่มีฟังก์ชันความหนาแน่นน่าจะเป็น

$$f(\hat{\mathbf{b}}|\mathbf{b}; \mathbf{Z}, \sigma_\varepsilon^2) = (2\pi)^{-(1/2)k}\left|\mathbf{V}^{-1}\right| \exp\{\frac{1}{2}[\hat{\mathbf{b}} - \mathbf{b}]^t\mathbf{V}^{-1}[\hat{\mathbf{b}} - \mathbf{b}]\} \tag{26}$$

สำหรับการเจาะจงค่า **z** และ σ_ε^2 และกำหนด **b** การแจกแจงของ $\hat{\mathbf{b}}$ เป็น $N(\bar{\mathbf{b}}', \mathbf{V}_m)$ มีฟังก์ชันความหนาแน่นน่าจะเป็น

$$g(\hat{\mathbf{b}}) = \int_{-\alpha}^{\alpha} f(\hat{\mathbf{b}}|\mathbf{b})h_0(\mathbf{b})d\mathbf{b} \tag{27}$$

ซึ่ง $$\mathbf{V}_m = \mathbf{V}' + \mathbf{V} = (\mathbf{G}'^{-1} + \mathbf{G}^{-1})\sigma_\varepsilon^2 \tag{28}$$

ดังนั้นการแจกแจงภายหลังของ **b** เมื่อกำหนดว่า $\hat{\mathbf{b}}$ ต้องเกิดขึ้น ณ เวลา T เป็น $N(\bar{\mathbf{b}}'', \mathbf{v}'')$ ที่มีฟังก์ชันความหนาแน่นน่าจะเป็น

$$h_T(\mathbf{b}|\hat{\mathbf{b}}) = \frac{h_0(\mathbf{b})f(\hat{\mathbf{b}}|\mathbf{b})}{g(\hat{\mathbf{b}})} \tag{29}$$

โดยค่าเฉลี่ย $\bar{\mathbf{b}}''$ และ $\mathbf{V}''$ พิจารณาจาก

$$\mathbf{V}''^{-1} = \mathbf{V}'^{-1} + \mathbf{V}^{-1} \tag{30}$$

$$\bar{\mathbf{b}}'' = \mathbf{V}''(\mathbf{V}'^{-1}\bar{\mathbf{b}}' + \mathbf{V}^{-1}\hat{\mathbf{b}}) \tag{31}$$

จะเขียนสมการ (30) เป็น

$$\mathbf{G}'' = \mathbf{G}' + \mathbf{G} = \sigma_{\varepsilon}^{2}\mathbf{V}^{-1} + \mathbf{Z}^{t}\mathbf{Z} \tag{31}$$

และเขียนสมการ (31) ได้อีกแบบหนึ่งเป็น

$$\mathbf{b}'' = \mathbf{G}''^{-1}(\mathbf{G}'\bar{\mathbf{b}}' + \mathbf{G}\hat{\mathbf{b}}) = \mathbf{G}''^{-1}(\mathbf{G}'\bar{\mathbf{b}}' + \mathbf{g}) \tag{32}$$

ผลจากข้างบนได้มาจาก Raiffa and Schlaifer (1961 อ้างถึงใน Montgomery, Johnson, and Gardiner, 1990)

จะสังเกตได้ว่าเมื่อการแจกแจงเบื้องต้นเป็นการแจกแจงปกติแบบพหุ การแจกแจงภายหลังจะมีการแจกแจงปกติแบบพหุด้วย

ส่วนที่ 3 การพยากรณ์ (Forecasting)

จากการแจกแจงภายหลังของ $\mathbf{b}$ ณ เวลา T การพยากรณ์ค่าที่ช่วงเวลา $T+\tau$ จะได้จาก $E(x_{T+\tau}) = \sum_{i=1}^{k} b_i z_i(T+\tau)$ ที่ได้จากวิธีของเบย์ นั่นคือ $b / \hat{b}$ เป็น $N(\bar{\mathbf{b}}'', \mathbf{V}'')$ ที่

$$\sum_{i=1}^{k} \bar{b}_i'' z_i(T+\tau) = \mathbf{z}^{t}(T+\tau)\bar{\mathbf{b}}'' \tag{33}$$

และความแปรปรวน

$$\sum_{i=1}^{k}\sum_{j=1}^{k} z_i(T+\tau) z_j(T+\tau) \upsilon_{ij}'' = \mathbf{z}^{t}(T+\tau)\mathbf{V}''\mathbf{z}(T+\tau) \tag{34}$$

ดังนั้น $y_{T+\tau}$ เป็นตัวแปรที่มีการแจกแจงปกติภายหลังที่มีค่าเฉลี่ยจากสมการ (33) และความแปรปรวน

$$\mathbf{z}^{t}(T+\tau)\mathbf{V}''\mathbf{z}(T+\tau) + \sigma_{\varepsilon}^{2} \tag{35}$$

จากการแจกแจงภายหลังของ $x_{T+\tau}$ ดังกล่าว ความคลาดเคลื่อนในการพยากรณ์สำหรับช่วงเวลา $T+\tau$ จะได้จาก

$$e_{\tau}(T+\tau) = x_{T+\tau} - \hat{x}_{T+\tau}(T)$$

ซึ่ง $e_{\tau}(T+\tau)$ มีการแจกแจงปกติด้วยค่าเฉลี่ย 0 และความแปรปรวนจากสมการ (35)

เพื่อให้สัญลักษณ์สั้นลง จึงนิยามให้ $M(T+\tau)$ เป็นค่าเฉลี่ยที่ได้จากสมการ (33) และ $S(T+\tau)$ เป็นรากที่สองของความแปรปรวน ที่ได้จากสมการ (35) จะหาช่วงความเชื่อมั่น $(1-\alpha)$ ร้อยละของการทำนายสำหรับ $x_{T+\tau}$ เป็น

$$M(T+\tau) \pm Z_{\alpha/2} S(T+\tau) \tag{36}$$

มีนักวิจัยใช้ลักษณะการพยากรณ์ด้วยวิธีของเบย์ เช่น Morrison และ Pike (1977) อ้างถึงใน Fildes (1984) ในช่วงปี 1970 ได้มีการนำวิธีของเบย์มาใช้จนเป็นที่ยอมรับอย่างแพร่หลายโดย Fildes (1979) และยังมีผู้ที่สนใจศึกษาการพยากรณ์ด้วยวิธีนี้อีก คือ Oliver (1987); de Alba (1988 อ้างถึงใน de Alba and Mendoza, 2001); Lenk, (1992); และ Guerrero Elizondo (1997) ได้ศึกษาการพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีของเบย์ ซึ่งเป็นวิธีการที่ค่อนข้างยากเพราะมีเรื่องของความน่าจะเป็นเข้ามาเกี่ยวข้องและใช้ข้อมูลที่มีขนาดใหญ่ แต่ในงานวิจัยนี้ผู้วิจัยจะศึกษาในกรณีที่ขนาดของข้อมูลอนุกรมเวลามีขนาดเล็ก

ตอนที่ 4 การพยากรณ์ด้วยวิธีบ็อกซ์-เจนกิ้นส์

วิธีบ็อกซ์–เจนกิ้นส์ (Box-Jenkins: B-J) เป็นการวิเคราะห์อนุกรมเวลาซึ่งใช้พยากรณ์ข้อมูลโดยไม่มีการกำหนดโมเดลความสัมพันธ์ก่อนที่จะทำการวิเคราะห์ แต่มีข้อตกลงเบื้องต้นของโมเดลว่าอนุกรมเวลาที่ใช้ในการวิเคราะห์จะต้องเป็นอนุกรมคงที่ (stationary) ถ้าอนุกรมที่ใช้วิเคราะห์เป็นอนุกรมไม่คงที่ (nonstationary) ผู้วิจัยต้องเปลี่ยนให้อยู่ในรูปอนุกรมคงที่โดยการหาผลต่างและ/หรือผลต่างฤดูกาล หลักการพยากรณ์ใช้วิธีการคำนวณทวนซ้ำ (iterative) เพื่อให้ได้ค่าพยากรณ์ที่สอดคล้องกับแบบแผนการเคลื่อนไหวของข้อมูลอนุกรมเวลามากที่สุด ซึ่งในการวิเคราะห์อนุกรมเวลาโดยการหารูปแบบที่เหมาะสมให้กับอนุกรมเวลาจะใช้ฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation function) และโมเดลที่เลือกจะอยู่ในกลุ่มของ ARMA(p,q) ดังนั้นผู้ใช้จะต้องมีความรู้เกี่ยวกับโมเดล ARMA(p,q) นอกจากนี้อนุกรมเวลาที่ใช้ในการวิเคราะห์ต้องมีจำนวนค่าสังเกตอย่างน้อย 30 ค่าเพื่อนำมาหาโมเดลในการพยากรณ์ วิธีนี้จึงเป็นวิธีที่ค่อนข้างยุ่งยากสลับซับซ้อน และอาศัยประสบการณ์และความชำนาญในการวิเคราะห์ แต่วิธีนี้ก็ให้ค่าพยากรณ์ที่มีความถูกต้องสูงกว่าวิธีอื่นในการพยากรณ์ในระยะเวลาสั้น ๆ

เนื่องจากรูปแบบความสัมพันธ์ข้อมูลอนุกรมเวลาจะถูกกำหนดขึ้นจากขั้นตอนการวิเคราะห์ ดังนั้นผู้ใช้จำเป็นต้องมีความรู้พื้นฐานและเพื่อที่จะอธิบายกระบวนการของวิธีนี้ง่ายขึ้น ผู้วิจัยจึงแบ่งการพยากรณ์ด้วยวิธีบ็อกซ์-เจนกิ้นส์ ออกเป็น 4 ส่วน คือ ส่วนที่ 1 ประเภทของอนุกรมเวลา ส่วนที่ 2 ฟังก์ชันอัตตะสหสัมพันธ์ และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน ส่วนที่ 3 โมเดลสำหรับการวิเคราะห์ และส่วนที่ 4 ขั้นตอนการวิเคราะห์ โดยมีรายละเอียดดังนี้

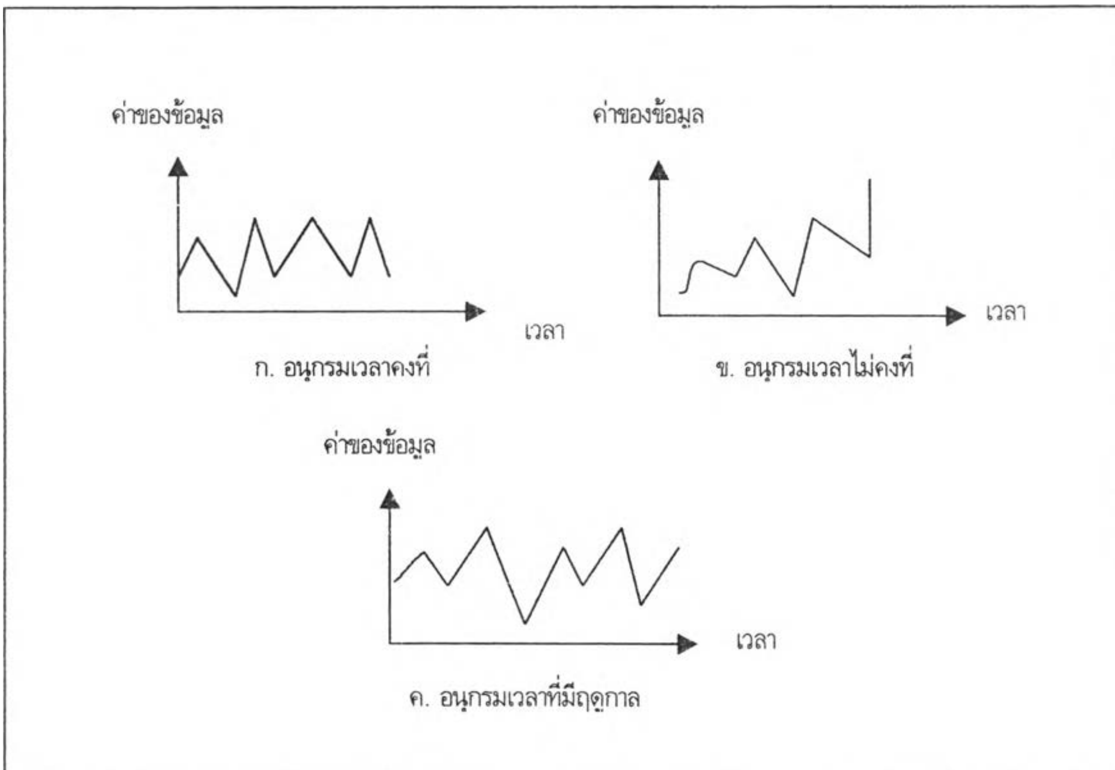
ส่วนที่ 1 ประเภทของอนุกรมเวลา

อนุกรมเวลาในวิธีบ็อกซ์-เจนกิ้นส์ แบ่งได้เป็น 3 ประเภท คือ

1.1 อนุกรมเวลาคงที่ (stationary time series หรือ horizontal) หมายถึง ข้อมูลที่เคลื่อนไหวไปรอบ ๆ ค่าเฉลี่ย และการเคลื่อนไหวเป็นไปในลักษณะคงที่ไม่เปลี่ยนแปลงไปตามเวลาที่ผ่านไป เมื่อค่าหนึ่งเพิ่มขึ้นค่าที่ตามมาจะลดลง ดังภาพ 6ก

1.2 อนุกรมเวลาไม่คงที่ (nonstationary time series) หมายถึง ข้อมูลที่เคลื่อนไหวไม่แน่นอน เปลี่ยนแปลงไปตามระยะเวลา ดังภาพ 6ข

1.3 อนุกรมเวลาที่มีฤดูกาล (seasonal time series) หมายถึง ข้อมูลที่มีการเคลื่อนไหวขึ้นลงตามระยะเวลาเป็นช่วงที่แน่นอน และลักษณะการเคลื่อนไหวในระยะเวลาหนึ่งจะคล้าย ๆ กันกับช่วงเวลาอื่น ๆ ซ้ำ ๆ กัน ดังภาพ 6ค



ภาพ 6 ลักษณะอนุกรมเวลาแบบคงที่ ไม่คงที่ และมีฤดูกาล

โดยปกติรูปแบบอนุกรมเวลาของบ็อกซ์-เจนกิ้นส์ ที่ใช้ในการพยากรณ์จะต้องเป็นอนุกรมเวลาที่คงที่ (stationary time series) ถ้าหากว่าอนุกรมเวลาชุดใดไม่คงที่หรือมีแนวโน้ม จะต้องทำการแปลงข้อมูล (transformation) อนุกรมเวลานั้นให้คงที่ ซึ่งอาจทำได้โดยหาผลต่าง (regular differencing) ลำดับต่าง ๆ ของข้อมูลจนกว่าอนุกรมเวลานั้นจะคงที่ หรือถ้าข้อมูลอนุกรมเวลานั้นมีฤดูกาลก็จะต้องกำจัดอิทธิพลของฤดูกาลออกไปโดยการหาผลต่างฤดูกาล (seasonal differencing) หรือบางครั้งถ้าความแปรปรวนไม่คงที่อาจต้องเปลี่ยนให้อยู่ในรูปของลอการิทึมฐานธรรมชาติ (natural logarithm) เพื่อทำให้อนุกรมเวลานั้นคงที่เสียก่อน แล้วจึงกำหนดรูปแบบของอนุกรมเวลา

ส่วนที่ 2 ฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function: ACF) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation function: PACF)

2.1 อัตตะสหสัมพันธ์ (autocorrelation) เป็นความสัมพันธ์ที่เกิดระหว่างข้อมูลอนุกรมเวลาชุดเดียวกันที่อยู่ในช่วงเวลาแตกต่างกัน ซึ่งในข้อมูลอนุกรมเวลานักวิจัยสามารถสร้างตัวแปรใหม่ตัวหนึ่ง

จากตัวแปรอีกตัวหนึ่งได้ ตัวแปรที่สร้างขึ้นใหม่นี้เรียกว่า **ตัวแปรเวลาล้าหลัง** (lag time variables: y_{t-k}) เมื่อ k เป็นเวลาที่แตกต่าง ดังตาราง 2

ตาราง 2 การสร้างตัวแปรเวลาล้าหลัง (lag time variables)

ช่วงเวลา t	ตัวแปรเริ่มต้น y_t	ตัวแปรล้าหลังเวลาที่ 1 y_{t-1}	ตัวแปรล้าหลังเวลาที่ 2 y_{t-2}	ตัวแปรล้าหลังเวลาที่ 3 Y_{t-3}
1	8			
2	6	8		
3	13	6	8	
4	12	13	6	8
5	14	12	13	6
6	15	14	12	13
7	7	15	14	12
8	9	7	15	14

จากตาราง สามารถสร้างตัวแปรเวลาล้าหลังที่1 (lag1: y_{t-1}) ตัวแปรเวลาล้าหลังที่ 2 (lag2: y_{t-2}) ตัวแปรเวลาล้าหลังที่ 3 (lag3: y_{t-3}) โดยการย้ายข้อมูลจากช่วงเวลา t_1 ลงมาเป็นช่วงเวลา t_2 จากช่วงเวลา t_2 ลงมาเป็นช่วงเวลา t_3 และช่วงเวลา t_3 ลงมาเป็นช่วงเวลา t_4 ตามลำดับ ผลจากการย้ายนี้ทำให้ค่าใน lag1: y_{t-1} หายไป 1 ค่า ค่าใน lag2: y_{t-2} หายไป 2 ค่า และค่าใน lag3: y_{t-3} หายไป 3 ค่า เมื่อนักวิจัยสร้างตัวแปรล้าหลังได้แล้ว จึงนำมาคำนวณค่าความสัมพันธ์ระหว่าง y_t และ y_{t-k} เมื่อ k = 1, 2, 3, ... ซึ่งก็คือ ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ (coefficient of autocorrelation) สามารถหาได้จากสูตร

$$r_k = \frac{\sum_{t=k+1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

เมื่อ	r_k	=	สัมประสิทธิ์อัตตะสหสัมพันธ์ ณ เวลาล้าหลัง k; $-1 < r_k < 1$
	t	=	ช่วงเวลา
	n	=	ช่วงเวลาสุดท้าย
	y_t	=	ค่าของข้อมูล ณ เวลา t
	$\bar{y}$	=	ค่าเฉลี่ยของข้อมูล
	k	=	เวลาล้าหลัง

โดยที่ $\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}$

เมื่อ $\rho_k = 0$, r_k มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวน $\sigma^2_{r_k}$ หรือ $r_k \sim N(0, \sigma^2_{r_k})$ ซึ่งจะประมาณ $\sigma^2_{r_k}$ ด้วย $S^2_{r_k}$ โดยค่าความคลาดเคลื่อนมาตรฐานของ r_k คือ

$$S_{r_k} = \sqrt{\frac{1 + 2\sum_{j=1}^{k-1} r_j^2}{n}}, k = 1,2,\ldots$$

$$\approx \frac{1}{\sqrt{n}}$$

การทราบลักษณะการแจกแจงของ r_k จะทำให้ทดสอบสมมติฐานเกี่ยวกับ ρ_k ได้ นั่นคือการทดสอบ $H_0 : \rho_k = 0$ กับ $H_1 : \rho_k \neq 0$ จะใช้สถิติทดสอบ

$$Z = \frac{r_k}{S_{r_k}}$$

$$= \sqrt{n} r_k$$

ที่มีช่วงวิกฤติ $|Z| \geq Z_{\alpha/2}$ ที่ระดับนัยสำคัญ α ในทำนองเดียวกันอาจจะใช้ตัวทดสอบสถิติ r_k ที่มีช่วงวิกฤติเป็น $|r_k| \geq Z_{\alpha/2}\sqrt{\frac{1}{n}}$ โดยทั่วไปการทดสอบจำทำที่ระดับนัยสำคัญ .05 จึงมีช่วงวิกฤติเป็น $|r_k| \geq \frac{1.96}{\sqrt{n}} \approx \frac{2}{\sqrt{n}}$

ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ (coefficient of autocorrelation: r_k) มีความหมายในทำนองเดียวกันกับสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient: r) คือ ค่าที่มีข้ออธิบายถึงความสัมพันธ์ระหว่างตัวแปรเหมือนกัน แต่ต่างกันที่ r_k ใช้อธิบายความสัมพันธ์ของตัวแปรเดียวกันในช่วงเวลาที่ห่างกัน k ช่วงเวลา หรือก็คือ เป็นค่าวัดสหสัมพันธ์ระหว่าง y_t และ y_{t-1} ซึ่งมีค่าอยู่ระหว่าง –1 ถึง 1 ถ้า $|r_k|$ มีค่าเข้าใกล้ 1 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลามีสหสัมพันธ์กันสูง แต่ถ้า $|r_k|$ มีค่าเข้าใกล้ 0 แสดงว่ามีสหสัมพันธ์กันต่ำ

ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ มีความสำคัญมากในการช่วยกำหนดลักษณะของข้อมูล ถ้าได้ข้อมูลมาจากการสุ่มอย่างสมบูรณ์ (completely random) ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์สำหรับช่วงเวลาที่ตามมา จะมีค่าเข้าใกล้หรือเท่ากับ 0 แต่ถ้าข้อมูลประกอบด้วยอิทธิพลของฤดูกาลหรือวัฏจักร จะพบว่าค่าสัมประสิทธิ์อัตตะสหสัม.พันธ์ที่ช่วงเวลาสูง เช่น lag 12, 24, ... จะมีค่าสูงมาก จะเห็นได้ชัดว่าก่อนที่จะใช้วิธีบ็อกซ์-เจนกิ้นส์ไม่จำเป็นต้องทราบลักษณะของข้อมูลเลย เพราะค่าสัมประสิทธิ์อัตตะสหสัมพันธ์จะช่วยกำหนดรูปแบบของข้อมูลอนุกรมเวลาให้

2.2 อัตตะสหสัมพันธ์บางส่วน (partial autocorrelation) เป็นความสัมพันธ์ที่เกิดระหว่างข้อมูลอนุกรมเวลาชุดเดียวกัน ระหว่าง y_t และ y_{t-k} เมื่อขจัด (partial out) อิทธิพลของเวลาล้าหลัง

1,2,3,...,k-1 ออกไป **ดังนั้น**ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์บางส่วนตัวแรกจะเหมือนกับสัมประสิทธิ์อัตตะสหสัมพันธ์ตัวแรก ค่าความสัมพันธ์ของอัตตะสหสัมพันธ์บางส่วนระหว่าง y_t และ y_{t-k} คือ ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์บางส่วน (coefficient of partial autocorrelation: r_{kk}) ซึ่งสามารถหาได้จากสูตร

$$r_{kk} = \begin{cases} r_1 & , k=1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} r_{k-1,j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j}\, r_j} & , k=2, 3, 4, \ldots \end{cases}$$

เมื่อ $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$, $j = 1,2,\ldots,k-1$

r_{kk} มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวน $\sigma^2_{r_{kk}}$ ประมาณด้วย $S^2_{r_{kk}}$ โดยค่าความคลาดเคลื่อนมาตรฐาน r_{kk} คือ

$$S_{r_{kk}} = \frac{1}{\sqrt{n}} \quad ,k=1,2,\ldots$$

การทราบลักษณะการแจกแจงของ r_{kk} จะทำให้ทดสอบสมมติฐานเกี่ยวกับ ρ_{kk} ได้ นั่นคือ การทดสอบ $H_0 : \rho_k = 0$ กับ $H_1 : \rho_k \neq 0$ จะใช้สถิติทดสอบ r_{kk} และจะปฏิเสธ H_0 ที่ระดับนัยสำคัญ .05 เมื่อ $|r_{kk}| \geq \frac{2}{\sqrt{n}}$

ในการกำหนดโมเดลจะพิจารณาจากชุดรวมของอัตตะสหสัมพันธ์และอัตตะสหสัมพันธ์บางส่วน ซึ่งก็คือ ฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function: ACF) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation function: PACF) แต่ในทางปฏิบัติไม่สามารถหาค่าฟังก์ชันอัตตะสหสัมพันธ์ (ρ_k) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของประชากร (ρ_{kk}) ได้ ดังนั้นจึงพิจารณาจากฟังก์ชันอัตตะสหสัมพันธ์ และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของกลุ่มตัวอย่าง ซึ่งแสดงในรูปแบบที่แตกต่างกัน 2 ลักษณะ ดังนี้

ลักษณะที่ 1 ค่า r_k และ r_{kk} ลดลงอย่างรวดเร็วเป็น 0 เรียกว่า มีค่าต่ำสุด (cuts off)

ลักษณะที่ 2 ค่า r_k และ r_{kk} มีค่ามากในเวลาล้าหลัง k แรก ๆ และลดลงอย่างช้า ๆ เกือบถึง 0 เมื่อ k มีค่ามากขึ้น เรียกว่า dies down หรือ tails off

ส่วนที่ 3 โมเดลสำหรับการวิเคราะห์

วิธีบ็อกซ์-เจนกิ้นส์ จะพิจารณาค่าฟังก์ชันอัตตะสหสัมพันธ์ และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน เพื่อสร้างโมเดลขึ้นมา โดยเรียกชื่อเฉพาะว่า ARMA (autoregressive moving average) โดยโมเดลสำหรับการวิเคราะห์อนุกรมเวลาที่คงที่และไม่มีฤดูกาล Box และ Jenkins กำหนดไว้ 3 โมเดลดังนี้

3.1 โมเดลกระบวนการถดถอยในตัวเอง (autoregressive process) เขียนแทนด้วยสัญลักษณ์ AR(p) หรือ ARMA(p,0) หรือ ARIMA(p,0,0) หมายถึง โมเดลที่แสดงว่าค่าสังเกต y_t จะขึ้นอยู่กับค่าของ $y_{t-1},\dots,y_{t-p}$ หรือค่าสังเกตที่เกิดขึ้นก่อน p ค่า สมการของ AR (p) คือ

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

เมื่อ y_t = ตัวแปรตาม

θ_0 = ค่าคงที่

ϕ_i = พารามิเตอร์ ตัวที่ i, i = 1,2,3,...,p

ε_t = ค่าความคลาดเคลื่อนที่เวลา t

โดยที่ $\theta_0 = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$

พารามิเตอร์ ϕ_i จะต้องมีคุณสมบัติ stationary คือ มีค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาคงที่ และค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ ที่ lag k ขึ้นอยู่กับค่า k อย่างเดียว โดยพิจารณาว่าค่าพารามิเตอร์ $\phi_1,\dots,\phi_p$ ใด ที่ทำให้คำตอบของสมการ $1 - \phi_1 B - \dots - \phi_p B^p = 0$ มี $|B| > 1$ ซึ่งสมการนี้มาจากโมเดลของ AR(p) นั่นเอง คือ

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

หรือ $$y_t - \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} = \theta_0 + \varepsilon_t$$

และเขียนในเทอมของ backward operator B ได้เป็น

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

ดังนั้น AR(p) อาจเขียนใหม่ได้เป็น $\phi_p(B^p) y_t = \theta_0 + \varepsilon_t$

โดยที่ $\phi_p(B) y_t = 1 - \phi_1 B - \dots - \phi_p B^p$

และ $By_t = y_{t-1},\dots,B^p y_t = y_{t-p}$

เรียก B ว่า ตัวถอยหลังเวลา

ฟังก์ชันอัตตะสหสัมพันธ์ (ρ_k) จะมีค่ามากในช่วงแรก ๆ และลดลงอย่างช้า ๆ เกือบถึง 0 เมื่อ k มีค่ามากขึ้น หรือมีลักษณะ dies down แบบ damped exponential หรือ damped sine wave ส่วนฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของประชากร (ρ_{kk}) มีค่าเป็น 0 หลังเวลาล้าหลัง p (lag p) หรือมีลักษณะ cuts off นั่นคือ $\rho_k \neq 0$; k=1,2,3,...,p และ $\rho_k = 0$; k >p

โดยทั่วไปในทางปฏิบัติ ลำดับของกระบวนการ AR มักมีค่าไม่เกิน 2 นั่นคือ $p \leq 2$

3.2 โมเดลกระบวนการการเคลื่อนที่เฉลี่ย (moving average process) เขียนแทนด้วยสัญลักษณ์ MA(q) หรือ ARMA(0,q) หรือ ARIMA(0,0,q) หมายถึง โมเดลที่แสดงว่าค่าสังเกต y_t จะขึ้นอยู่กับค่าของความคลาดเคลื่อน $\varepsilon_{t-1},\ldots,\varepsilon_{t-q}$ หรือความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า สมการของ MA(q) คือ

$$y_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q}$$

เมื่อ y_t = ตัวแปรตาม

θ_0 = ค่าคงที่

θ_i = พารามิเตอร์ ตัวที่ i, i = 1,2,3,...,q

ε_t = ค่าความคลาดเคลื่อนที่เวลา t

โดยที่ $\theta_0 = \mu$

พารามิเตอร์ θ_i จะต้องมีคุณสมบัติ คือ invertible คุณสมบัติที่ทำให้ค่าความคลาดเคลื่อนของการพยากรณ์ ε_t ในเทอมของ $y_t, y_{t-1},\ldots$ ได้ โดยพิจารณาว่าค่าพารามิเตอร์ $\theta,\ldots,\theta_q$ ใด ที่ทำให้คำตอบของ สมการ $1-\theta_1 B-\ldots-\theta_q B^q = 0$ มี $|B| > 1$ ซึ่งสมการนั้นมาจากโมเดลของ MA(q) นั่นเอง คือ

$$y_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q}$$

$$y_t = \theta_0 + (1 - \theta_1 B - \ldots - \theta_q B^q)\varepsilon_t$$

ดังนั้น MA(q) อาจเขียนใหม่ได้เป็น $y_t = \theta_0 + \theta_q(B)\varepsilon_t$

โดยที่ $\theta_q(B) = 1-\theta_1 B - \ldots - \theta_q B^q$

และ $B\varepsilon_t = \varepsilon_{t-1},\ldots,B^q\varepsilon_t = \varepsilon_{t-q}$

ฟังก์ชันอัตตะสหสัมพันธ์ (ρ_k) มีค่าเป็น 0 หลัง lag q หรือมีลักษณะ cuts off นั่นคือ

$\rho_k \neq 0$; k=1,2,3,...,q

และ $\rho_k = 0$; k >q

ส่วนฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (ρ_{kk}) มีค่ามากในช่วงแรก ๆ และลดลงอย่างช้า ๆ เกือบถึง 0 เมื่อ k มีค่ามากขึ้น หรือมีลักษณะ dies down แบบ damped exponential หรือ damped sine wave หรือทั้งสองแบบรวมกัน

โดยทั่วไปในทางปฏิบัติ ลำดับของกระบวนการ MA มักมีค่าไม่เกิน 2 นั่นคือ $q \leq 2$

3.3 โมเดลรวมกันของกระบวนการถดถอยในตัวเองและกระบวนการเฉลี่ยเคลื่อนที่ (mixed autoregressive-moving average models) เขียนแทนด้วยสัญลักษณ์ ARMA(p,q) หรือ ARIMA(p,0,q) มี p เป็นอันดับของ AR และ q เป็นอันดับของ MA โดยโมเดลนี้เป็นการรวมส่วนของโมเดล AR(p) และโมเดล MA(q) เข้าด้วยกัน แสดงว่าค่าสังเกต y_t ขึ้นอยู่กับค่าสังเกตที่เกิดขึ้นก่อน p ค่า และค่าของความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า สมการของ ARMA(p,q) คือ

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q}$$

เมื่อ y_t = ตัวแปรตาม

θ_0 = ค่าคงที่

θ_i = พารามิเตอร์ ตัวที่ i, i = 1,2,3,...,q

ϕ_i = พารามิเตอร์ ตัวที่ i, i = 1,2,3,...,p

ε_t = ค่าความคลาดเคลื่อนที่เวลา t

โดยที่ $\theta_0 = \mu(1-\phi_1-\phi_2-\ldots-\phi_p)$

พารามิเตอร์ ϕ_i จะต้องมีคุณสมบัติ stationary และพารามิเตอร์ θ_i จะต้องมีคุณสมบัติคือ invertible

ฟังก์ชันอัตตะสหสัมพันธ์น (ρ_k) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (ρ_{kk}) มีค่ามากในช่วงแรก ๆ และลดลงอย่างช้า ๆ เกือบถึง 0 เมื่อ k มีค่ามากขึ้น หรือมีลักษณะ dies down แบบ damped exponential หรือ damped sine wave หรือทั้งสองแบบรวมกัน

โมเดลที่ใช้ในการวิเคราะห์อนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกิ้นส์ทั้ง 3 โมเดลข้างต้น สามารถนำมาสรุปเป็นลักษณะทางทฤษฎีของอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนสำหรับการวิเคราะห์อนุกรมเวลาคงที่ (Bowerman และ O'Connell, 1993) ดังตาราง 3

ตาราง 3 ลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของโมเดลการวิเคราะห์อนุกรมเวลาคงที่

โมเดล	ฟังก์ชันอัตตะสหสัมพันธ์	ฟังก์ชันอัตตะสหสัมพันธ์บางส่วน
AR(p) = ARIMA(p,0,0)	ลดลงอย่างช้า ๆ เข้าใกล้ 0	ลดลงอย่างเร็วเข้าใกล้ 0 หลังเวลาล้าหลัง p
MA(q) = ARIMA(0,0,q)	ลดลงอย่างเร็วเข้าใกล้ 0 หลังเวลาล้าหลัง q	ลดลงอย่างช้า ๆ เข้าใกล้ 0
ARMA(p,q) = ARIMA(p,0,q)	ลดลงอย่างช้า ๆ เข้าใกล้ 0	ลดลงอย่างช้า ๆ เข้าใกล้ 0

จากการแบ่งประเภทของอนุกรมเวลาที่กำหนดโมเดลโดยวิธีบ็อกซ์-เจนกิ้นส์ 3 ประเภท ได้แก่ อนุกรมเวลาคงที่ อนุกรมเวลาไม่คงที่ และอนุกรมเวลาฤดูกาล มีลักษณะของโมเดล ARIMA(p,d,q) แตกต่างกันไปตามลักษณะของประเภทของอนุกรมเวลา มีรายละเอียดดังนี้

สำหรับอนุกรมเวลาคงที่ โมเดล ARMA(p,q) มี p เป็นอันดับของ AR และ q เป็นอันดับของ MA จำนวนพารามิเตอร์ในโมเดลจะเท่ากับ p+q+1 โมเดลที่กำหนดให้กับอนุกรมเวลามักจะเป็นอนุกรมเวลาที่มีจำนวนพารามิเตอร์น้อย ในทางปฏิบัติมักจะไม่เกิน 3 พารามิเตอร์ ดังตาราง 4 (ทรงศิริ แต้สมบัติ, 2539) แสดงโมเดล ARMA ต่าง ๆ และเงื่อนไขของพารามิเตอร์ที่ปรากฏในโมเดล สำหรับเงื่อนไขของพารามิเตอร์ในส่วนของ AR อันดับที่ p เป็นเงื่อนไขที่กำหนดให้โมเดล ARMA(p,q) คงที่ (stationary) ซึ่งเป็นคุณสมบัติที่ทำให้ $E(Y_t)$ และ $V(Y_t)$ คงที่ และ $cov(Y_t,Y_{t-k})$ จะขึ้นกับเวลาล้าหลัง k อย่างเดียว การพิจารณาค่าพารามิเตอร์ $\phi_1, \ldots, \phi_p$ ใด ที่จะทำให้โมเดล AR คงที่จะทำได้โดย

1. จากโมเดล AR(p) ในเทอมของ backward operator B คือ

$$(1-\phi_1 B - \ldots - \phi_p B^p) y_t = \theta_0 + \varepsilon_t$$

2. หาค่าตอบของสมการ $(1-\phi_1 B - \ldots - \phi_p B^p) y_t = 0$ จะได้ค่า B จำนวน p ค่า จะเลือก B เพียงหนึ่งค่าที่อยู่นอก unit circle นั่นคือ $|B|$ ต้องมีค่ามากกว่า 1 เงื่อนไขดังกล่าวของ B จะเป็นเงื่อนไขของความคงที่

ตาราง 4 โมเดล ARMA(p,q) และเงื่อนไขความคงที่และ invertible

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

ส่วนเงื่อนไขของพารามิเตอร์ในส่วนของ MA ที่อันดับ q จะเป็นเงื่อนไขที่กำหนดให้โมเดล MA(q) เป็น invertible ซึ่งเป็นคุณสมบัติของ MA(q) ที่ทำให้ค่าความคลาดเคลื่อนของการพยากรณ์ ε_t ในเทอมของ Y_t, Y_{t-1}, ... ได้ การพิจารณาค่าพารามิเตอร์ θ_1, ..., θ_q ใด ที่ทำให้โมเดล MA(q) เป็น invertible ทำได้โดย

1. จากโมเดล MA(q) ในเทอมของ backward operator B คือ

$$y_t = \theta_0 + (1 - \theta_1 B - \ldots - \theta_q B^q)\varepsilon_t$$

2. หาค่าตอบของสมการ $1 - \theta_1 B - \ldots - \theta_q B^q = 0$ จะได้ค่าของ B จำนวน q ค่า จะเลือก B เพียงหนึ่งค่าที่อยู่นอก unit circle นั่นคือ $|B|$ ต้องมีค่ามากกว่า 1 เงื่อนไขดังกล่าวของ B จะเป็นเงื่อนไขของ invertible

สำหรับอนุกรมเวลาที่มีแนวโน้มเพียงอย่างเดียว นั่นคือมีค่าเฉลี่ย $E(Y_t)$ ไม่คงที่ สำหรับแต่ละค่าของ t การปรับอนุกรมเวลาดังกล่าวให้เป็นอนุกรมเวลาใหม่ที่มีลักษณะคงที่จะทำได้โดยการหาผลต่าง นั่นคือจากอนุกรมเวลา $\{Y_t\}$ แปลงเป็นอนุกรมเวลาใหม่ $\{Z_t\}$ เมื่ออนุกรมเวลา $\{Y_t\}$ มีลักษณะคงที่แล้วจะหาโมเดล ARMA(p,q) ให้กับอนุกรมเวลา $\{Z_t\}$ เป็น $Z_t \sim$ ARMA(p,q) และสำหรับโมเดลของอนุกรมเวลา $\{Y_t\}$ $Y_t \sim$ ARIMA(p,d,q) ซึ่งโมเดล ARMA มี p เป็นอันดับของ AR และ q เป็นอันดับของ MA โดย d เป็นจำนวนครั้งที่หาผลต่างเพื่อให้อนุกรมเวลา $\{Z_t\}$ มีลักษณะคงที่ ตัวอย่างของโมเดล ARIMA(p,1,q) ของอนุกรมเวลา $\{Y_t\}$ เป็นดังตาราง 5 และลักษณะฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของโมเดล ARIMA(p,1,q) แสดงในตาราง 6 (ทรงศิริ แต้สมบัติ, 2539)

ตาราง 5 โมเดล ARIMA(p,1,q) ของอนุกรมเวลา $\{Z_t\}$ และ $\{Y_t\}$

โมเดล ARIMA(p,1,q)	อนุกรมเวลา $\{Z_t\}$	อนุกรมเวลา $\{Y_t\}$
ARIMA(0,1,0)	$Z_t = \theta_0 + \varepsilon_t$	$Y_t = \theta_0 + Y_{t-1} + \varepsilon_t$
ARIMA(1,1)	$Z_t = \theta_0 + \phi_1 Z_{t-1} + \varepsilon_t$	$Y_t = \theta_0 + (1+\phi_1) Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t$
ARIMA(2,1)	$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \varepsilon_t$	$Y_t = \theta_0 + (1+\phi_1) Y_{t-1} + (\phi_2 - \phi_1) Y_{t-2}$ $-\phi_2 Y_{t-3} + \varepsilon_t$
IMA(1,1)	$Z_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1}$	$Y_t = \theta_0 + Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$
IMA(1,2)	$Z_t = \theta_0 + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}$	$Y_t = \theta_0 + Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}$
ARIMA(1,1,1)	$Z_t = \theta_0 + \phi_1 Z_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$	$Y_t = \theta_0 + (1+\phi_1) Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t$ $- \theta_1\varepsilon_{t-1}$

สำหรับอนุกรมเวลาที่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โมเดลที่จะใช้ ได้แก่ SARIMA $(P,D,Q)_L$ (seasonal integrated autoregressive and moving average order P,D,Q) โดย P เป็นอันดับของ SAR (seasonal autoregressive) Q เป็นอันดับของ SMA (seasonal moving average) และ D เป็นจำนวนครั้งที่หาผลต่างฤดูกาลเพื่อทำให้อนุกรมเวลา $\{Y_t\}$ ที่ไม่คงที่เนื่องจากฤดูกาลเป็นอนุกรมเวลาชุดใหม่ $\{Z_t\}$ ที่คงที่ ตัวอย่างของโมเดล SARMA$(P,Q)_L$ ของ $\{Y_t\}$ ดังตาราง 7

ตาราง 6 ลักษณะของ $\rho_k(Y_t)$, $\rho_k(Z_t)$ $\rho_{kk}(Z_t)$ สำหรับโมเดล ARIMA(p,d,q)

โมเดล ARIMA(p,d,q) ของ Y_t	ฟังก์ชันอัตตะสหสัมพันธ์ ของ Y_t	ฟังก์ชันอัตตะสหสัมพันธ์ ของ Z_t	ฟังก์ชันอัตตะสหสัมพันธ์ บางส่วนของ Z_t
Random walk	ลดลงอย่างช้า ๆ	ทุก ρ_k เป็น 0	ทุก ρ_{kk} เป็น 0
AR(1,1)	ลดลงอย่างช้า ๆ	ค่าลดลงเร็วใกล้ 0	ρ_{kk} เป็น 0 สำหรับ k =2, ...
AR(2,1)	ลดลงอย่างช้า ๆ	ค่าลดลงเร็วใกล้ 0	ρ_{kk} เป็น 0 สำหรับ k =3, ...
IMA(1,1)	ลดลงอย่างช้า ๆ	ρ_k เป็น 0 สำหรับ k = 2, ...	ค่าลดลงเร็วใกล้ 0
IMA(1,2)	ลดลงอย่างช้า ๆ	ρ_k เป็น 0 สำหรับ k = 3, ...	ค่าลดลงเร็วใกล้ 0
ARIMA(1,1,1)	ลดลงอย่างช้า ๆ	ค่าลดลงเร็วใกล้ 0	ค่าลดลงเร็วใกล้ 0

ตาราง 7 โมเดล $SARMA(P,Q)_L$ ของอนุกรมเวลา $\{Y_t\}$

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

คุณสมบัติที่สำคัญของโมเดล $SARMA(P,Q)_L$ เมื่อ L = 12 ของอนุกรมเวลาที่มีลักษณะคงที่ได้แก่ ρ_k และ ρ_{kk} ลักษณะของ ρ_k และ ρ_{kk} แสดงดังตาราง 8

ตาราง 8 ลักษณะของ $\rho_k(Z_t)$ $\rho_{kk}(Z_t)$ สำหรับโมเดล $SARMA(P,Q)_{12}$

โมเดล $SARMA(P,Q)_{12}$	ฟังก์ชันอัตตะสหสัมพันธ์ของ Z_t	ฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของ Z_t
$SAR(1)_{12}$	ρ_{12}, ρ_{24}, ... มีค่าลดลงเร็วใกล้ 0	$\rho_{kk} = 0$ สำหรับ k = 24, 36, ..
$SAR(2)_{12}$	ρ_{12}, ρ_{24}, ... มีค่าลดลงเร็วใกล้ 0	$\rho_{kk} = 0$ สำหรับ k = 36, 48, ..
$SMA(1)_{12}$	$\rho_k = 0$ สำหรับ k = 24, 36, ..	$\rho_{12,12}$, $\rho_{24,24}$, ... มีค่าลดลงเร็วใกล้ 0
$SMA(2)_{12}$	$\rho_k = 0$ สำหรับ k = 36, 48, ..	$\rho_{12,12}$, $\rho_{24,24}$, ... มีค่าลดลงเร็วใกล้ 0

ตัวอย่างของโมเดล SARMA(P,1,Q)$_{12}$ ของอนุกรมเวลา $\{Y_t\}$ หรือโมเดล SARMA(P,,Q)$_{12}$ ของอนุกรมเวลา $\{Z_t\}$ เป็นดังนี้

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

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

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

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

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

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

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

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

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

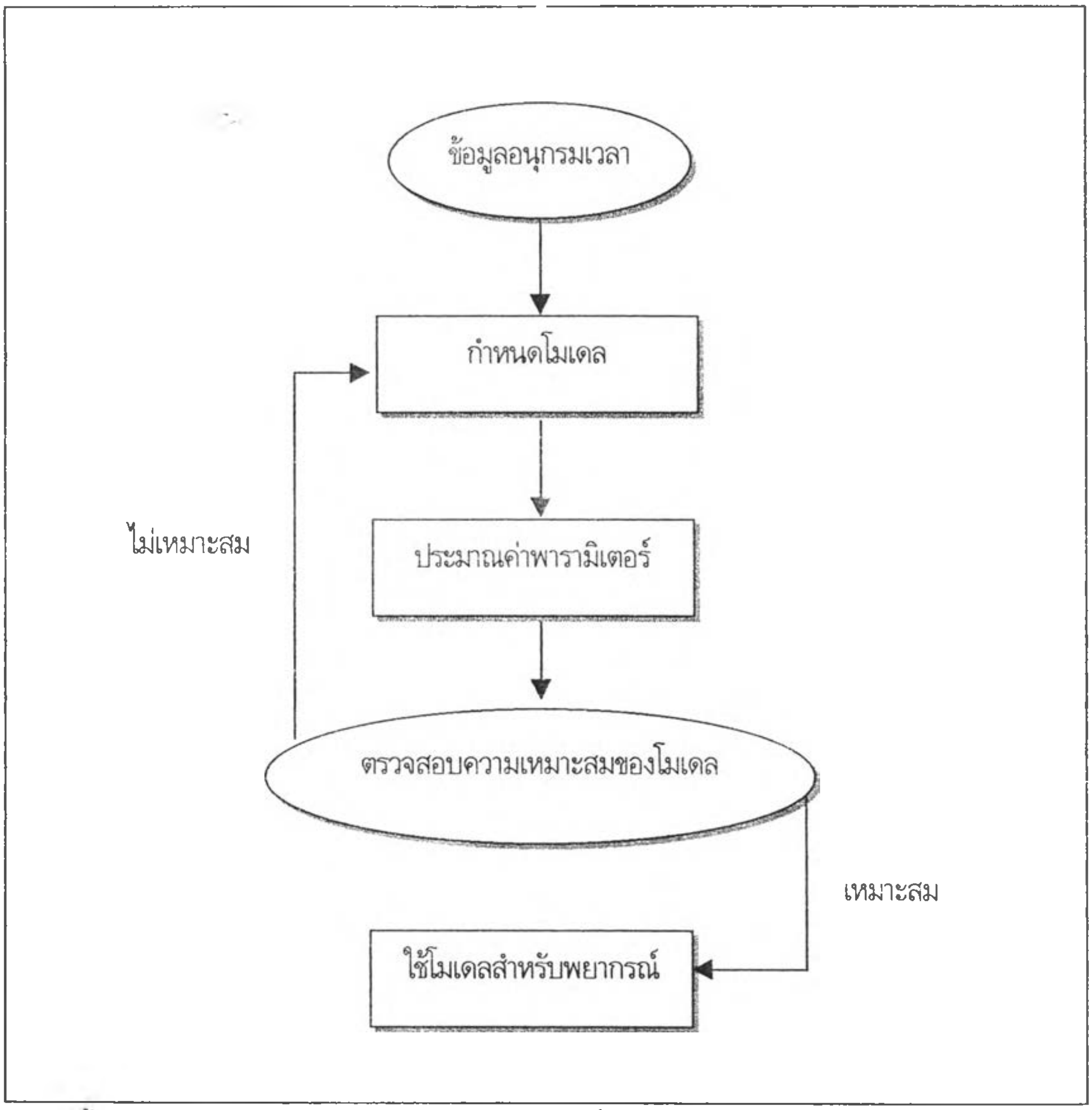
ในกรณีของอนุกรมเวลามีทั้งแนวโน้มและอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โมเดลที่จะใช้ได้แก่ ARIMA(p,d,q)×SARIMA(P,D,Q)$_L$ ผู้พยากรณ์ต้องแปลงอนุกรมเวลาให้อยู่ในลักษณะที่คงที่ โดยการหาผลต่างและผลต่างฤดูกาลควบคู่กับ โดย d เป็นจำนวนครั้งที่หาผลต่าง และ D เป็นจำนวนครั้งที่หาผลต่างฤดูกาล เพื่อทำให้อนุกรมเวลา $\{Y_t\}$ ที่ไม่คงที่เนื่องจากมีแนวโน้มและอิทธิพลของฤดูกาลเป็นอนุกรมเวลาชุดใหม่ $\{Z_t\}$ ที่คงที่ โดย $Z_t = \nabla^d \nabla_L^D Y_t$ เช่นสำหรับอนุกรมเวลารายปี (L = 12) และมีโมเดล $Y_t \sim$ ARI(1,1) ×SARI(1,1)$_{12}$ หรือ $Z_t \sim$ AR(1) ×SAR(1)$_{12}$ สามารถเขียนสมการได้ ดังนี้

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

ส่วนที่ 4 ขั้นตอนการวิเคราะห์

วิธีบ็อกซ์-เจนกิ้นส์ แบ่งขั้นตอนในการวิเคราะห์อนุกรมเวลาไว้ 4 ขั้นตอน คือ การกำหนดรูปแบบ (identification) การประมาณค่าพารามิเตอร์ (estimate parameters) การตรวจสอบความเหมาะสมของรูปแบบ (diagnotic checking) และใช้โมเดลสำหรับพยากรณ์ (model for forecasting) ดังภาพ 7

ถ้าในขั้นตอนการตรวจสอบความเหมาะสมของโมเดล พบว่า โมเดลไม่เหมาะสมก็จะไปกำหนดโมเดลในขั้นตอนที่ 1 ใหม่ จนกว่าจะได้โมเดลที่เหมาะสม แล้วจึงใช้โมเดลนั้นพยากรณ์ข้อมูลในอนาคตต่อไป



ภาพ 7 ขั้นตอนการวิเคราะห์อนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกิ้นส์

ตอนที่ 5 วิธีการพยากรณ์ร่วม

แนวความคิดในการรวมการพยากรณ์ เริ่มต้นประมาณต้นศตวรรษที่ 19 จากประสิทธิภาพของการรวมการพยากรณ์ทำให้เกิดมีการศึกษาเพิ่มเติมขึ้นในเวลาต่อมา และใน ค.ศ. 1963 Barnard (อ้างถึงใน Granger and Newbold, 1986) ได้ทำการรวมวิธีการพยากรณ์ 2 วิธีเข้าด้วยกัน โดยใช้ simple everage ต่อมาใน ค.ศ. 1969 Bates และ Granger ใช้แนวความคิดนี้ของ Barnard ประยุกต์ขึ้นเป็น weighted everage และในปี ค.ศ. 1974 Newbold และ Granger ได้ปรับปรุง weighted everage ให้ดีขึ้น โดยในระยะเริ่มแรกนี้ได้นำเอาวิธีการพยากรณ์ 2 วิธี มารวมกัน เช่น วิธีบ็อกซ์-เจนกิ้นส์รวมกับวิธีการปรับให้เรียบแบบโฮลท์-วินเทอร์ วิธีบ็อกซ์-เจนกิ้นส์รวมกับวิธี stepwise autoregression เป็นต้น และพบว่าวิธีการรวมการพยากรณ์นี้มีความถูกต้องในการพยากรณ์มากกว่าวิธีการพยากรณ์เดี่ยว ดังนั้นต่อมาในปี ค.ศ.

1983 Winkler และ Makridakis จึงได้มีการพัฒนาขึ้นโดยการนำเอาวิธีการพยากรณ์หลายๆ วิธีมารวมกัน โดยใช้ weighted everage ของ Newbold และ Granger

จากผลการค้นคว้าของทัศนีย์ อินทนู (2543) พบว่า Makridakis and Winkler; Cleman and Winkler; Makridakis; Batchelor and Dua ในปี 1983, 1986, 1989 และ 1995 ในการหาค่าพยากรณ์ที่ได้จากการรวมการพยากรณ์ (combining forecast) สรุปได้ว่าการรวมการพยากรณ์มีประโยชน์และข้อดีหลายประการ ได้แก่

1.การรวมการพยากรณ์ (combining forecast) มีค่าความถูกต้องมากกว่า นั่นคือ มีความคลาดเคลื่อนจากการพยากรณ์น้อยกว่าวิธีการพยากรณ์เพียงวิธีเดียว

2.วิธีการพยากรณ์ที่จะนำมารวมกัน (combine) ยิ่งใช้วิธีมากขึ้นเท่าใด ค่าความถูกต้องก็จะยิ่งเพิ่มมากตามไปด้วย และค่าความแปรปรวนของความคลาดเคลื่อนก็จะยิ่งลดลง

3.การนำข้อมูลจากแหล่งข้อมูลหลายแหล่งหรือวิธีการพยากรณ์ที่แตกต่างกันมารวมกัน จะทำให้ค่าการพยากรณ์นั้นมีความถูกต้องมากขึ้น

4.การนำวิธีที่มีความแตกต่างกันมากมารวมกัน จะยังให้ผลการพยากรณ์ได้ดีกว่าการพยากรณ์วิธีเดียว รวมทั้งให้ผลที่ดีกว่าการรวมวิธีที่เหมือนกัน

เมื่อมีการรวมวิธีการพยากรณ์หลายๆ วิธีเข้าด้วยกัน ก็จะต้องมีการให้น้ำหนักของวิธีพยากรณ์เดี่ยวแต่ละวิธีที่นำมารวมกัน การให้น้ำหนักของวิธีที่นำมารวมกัน แบ่งเป็น 2 ประเภท คือ simple everage และ weighted average ซึ่งการให้น้ำหนักแบบ simple everage นั้นจะเป็นการให้น้ำหนักเฉลี่ยที่เท่ากันทุกวิธี นั่นคือให้ความสำคัญในแต่ละวิธีเท่ากัน ส่วนการให้น้ำหนักแบบ weighted average ในแต่ละวิธีที่นำมารวมกันนั้นจะไม่เท่ากัน นั่นคือ ใน weighted average จะให้น้ำหนักตามความสำคัญของแต่ละวิธีที่นำมารวมกัน

ในการวิจัยครั้งนี้ วิธีการพยากรณ์ร่วม (combined forecasting) เลือกวิธีพยากรณ์เดี่ยวที่เหมาะสมกับข้อมูลพิจารณาจาก 3 กรณีตามลักษณะของข้อมูลอนุกรมเวลา คือ กรณีแรก เมื่อข้อมูลไม่มีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) กรณีที่สอง เมื่อข้อมูลมีอิทธิพลของแนวโน้มเพียงอย่างเดียวจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสำหรับข้อมูลที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential smoothing for exponential trend: ESE) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) และกรณีที่สาม เมื่อข้อมูลมีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจาก

วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) แต่ละวิธีมีรายละเอียดดังต่อไปนี้

1. วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) เป็นการปรับข้อมูลอนุกรมเวลาให้เรียบโดยการกำจัดอิทธิพลของเหตุการณ์ที่ผิดปกติ (I) ออกไป ซึ่งจะทำให้เห็นลักษณะที่แท้จริงของข้อมูล โดยการนำชุดของข้อมูลมาหาค่าเฉลี่ย แล้วใช้ค่าเฉลี่ยนั้นพยากรณ์ในช่วงเวลาถัดไป จำนวนข้อมูลที่เหมาะสมในการนำมาหาค่าเฉลี่ย คือ จำนวนที่ทำให้ค่าพยากรณ์มีค่าใกล้เคียงกับความเป็นจริงมากที่สุดหรือค่าความคลาดเคลื่อนน้อยที่สุด วิธีนี้เหมาะกับข้อมูลที่มีลักษณะคงที่หรือสม่ำเสมอตามแนวนอน (horizontal) ใช้พยากรณ์ข้อมูลอนุกรมเวลาในระยะสั้นและระยะปานกลาง ข้อจำกัดของวิธีนี้ คือ ต้องใช้ข้อมูลล่าสุด การคำนวณค่าเฉลี่ยเคลื่อนที่จะต้องมีข้อมูลอย่างน้อย k ค่า และจะมีข้อมูลบางส่วนหายไป โดยมีขั้นตอนดังนี้

1. กำหนดค่าคงที่ของจำนวนเทอมที่จะหาค่าเฉลี่ยหรือจำนวนข้อมูลที่ต้องการหาค่าเฉลี่ย (k)
2. คำนวณค่าพยากรณ์จากสูตร

$$\hat{y}_{t+1} = \frac{y_t + y_{t-1} + \cdots + y_{t-k+1}}{k}$$

และ $\hat{y}_{t+m} = \hat{y}_{t+1}$, $m \geq 2$

เมื่อ y_t = ค่าของของข้อมูล ณ เวลา t

$\hat{y}_{t+1}$ = ค่าพยากรณ์ ณ เวลาที่ t+1

$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, $m \geq 2$

k = จำนวนช่วงเวลาที่ใช้ในการเฉลี่ย

ข้อบกพร่องของวิธีนี้ คือ การให้น้ำหนักแก่ข้อมูลเท่ากันหมด ซึ่งในความเป็นจริงแล้วข้อมูลล่าสุดควรมีความสำคัญมากกว่า เพราะมีสถานการณ์สิ่งแวดล้อมคล้ายในอนาคตมากกว่า ดังนั้นจึงควรให้น้ำหนักแก่ข้อมูลล่าสุดมากกว่า ซึ่งวิธีแก้ไขข้อบกพร่องนี้ คือ วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing)

วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล เป็นการหาค่าเฉลี่ยแบบถ่วงน้ำหนัก โดยให้ความสำคัญแก่ข้อมูลล่าสุดมากที่สุด และความสำคัญจะลดลงเรื่อย ๆ แบบเรขาคณิต หรือแบบเอ็กซ์โปเนนเชียลสำหรับข้อมูลที่มีระยะเวลาห่างไกลออกไป (ทรงศิริ แต้สมบัติ, 2539) ...กำหนดน้ำหนัก หรือค่าคงที่สำหรับปรับให้เรียบ (smoothing constant) ซึ่งมีตั้งแต่ 1 ตัวขึ้นไปขึ้นอยู่กับแต่ละวิธี จะมีค่าอยู่ระหว่าง 0 ถึง 1 เท่านั้น แต่วิธีนี้มีปัญหาในการกำหนดค่าเริ่มต้นว่าควรจะกำหนดโดยใช้ข้อมูลตัวแรกเป็นค่าเริ่มต้น หรือใช้วิธีกำลังสองน้อยที่สุดคำนวณหาค่าเริ่มต้น ซึ่งไม่ปรากฏทางทฤษฎีว่ากำหนดแบบใดดีที่สุด โดยทั่วไปจึงต้องทำเปรียบเทียบกันว่ากำหนดแบบใดที่ให้ผลพยากรณ์ได้ถูกต้องที่สุด แต่ถ้าใช้โปรแกรมสำเร็จรูปปัญหานี้ก็จะหมดไป เนื่องจากโปรแกรมสำเร็จรูปจะกำหนดค่าเริ่มต้นที่ดีที่สุดให้

2. วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีนี้คล้ายกับวิธีการเฉลี่ยเคลื่อนที่แบบง่าย (SMA) คือ เพื่อกำจัดอิทธิพลของความผันแปรไม่แน่นอนไปจากข้อมูล แต่การกำหนดน้ำหนักของข้อมูลแต่ละเวลาต่างกัน คือข้อมูลล่าสุดจะได้รับการถ่วงน้ำหนักมากกว่าข้อมูลในอดีตตัวอื่น ๆ และจะลดหลั่นกันไปสำหรับข้อมูลหรือค่าสังเกตที่อยู่ห่างออกไป น้ำหนักที่ให้ขึ้นอยู่กับค่าปรับน้ำหนัก (α)

วิธีนี้เหมาะสำหรับข้อมูลที่มีลักษณะคงที่ (horizontal) หรือเป็นข้อมูลที่ไม่มีแนวโน้ม และไม่มีอิทธิพลของฤดูกาล ใช้พยากรณ์ข้อมูลอนุกรมเวลาในระยะสั้น ช่วงเวลาของการพยากรณ์ ทำได้ 1 ช่วงเวลาล่วงหน้า ข้อจำกัดของวิธีนี้ คือ ต้องมีค่าสังเกตค่าล่าสุด โดยมีขั้นตอน ดังนี้

1. กำหนดค่าเริ่มต้นของการพยากรณ์
2. กำหนดค่าปรับน้ำหนัก (α) โดยที่ $0<\alpha<1$
3. คำนวณค่าพยากรณ์ จากสูตร

$$\begin{aligned}\hat{y}_{t+1} &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \ldots \\ &= \alpha y_t + (1-\alpha)\hat{y}_t\end{aligned}$$

และ $\hat{y}_{t+m} = \hat{y}_{t+1}$, $m \geq 2$

เมื่อ y_t = ค่าของของข้อมูล ณ เวลา t

$\hat{y}_{t+1}$ = ค่าพยากรณ์ ณ เวลาที่ t+1

$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, $m \geq 2$

α = ค่าปรับน้ำหนัก ; $0<\alpha<1$

การกำหนดค่าปรับน้ำหนัก (α) ควรมีค่าอยู่ระหว่าง .01 ถึง .30 จะใช้ได้ผลดี (Sullivan and Claycombe, 1977; Montogomery; Johnson and Gardiner, 1990; Makridakis; Wheelwright and Hyndman, 1998 อ้างถึงใน ทัศนีย์ อินทนู, 2543) แต่ในทางปฏิบัตินักวิจัยไม่จำเป็นต้องกำหนดน้ำหนักหรือค่าคงที่สำหรับปรับให้เรียบนี้เอง เนื่องจากโปรแกรมสำเร็จรูปจะกำหนดน้ำหนักที่เหมาะสมนี้ให้ โดยค่าน้ำหนักที่เหมาะสม คือ ค่าที่ทำให้ความคลาดเคลื่อน (SSE) มีค่าต่ำสุด จึงจะทำให้ค่าพยากรณ์ใกล้เคียงกับค่าจริง ถ้า α มีค่าใกล้ 1 แสดงว่า ให้ความสำคัญแก่ข้อมูลล่าสุดมากที่สุด แต่ถ้า α มีค่าใกล้ 0 แสดงว่า ค่าพยากรณ์จะไม่คำนึงถึงความผิดพลาดในการพยากรณ์ ณ เวลาที่เพิ่งผ่านมานัก

3. วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) เป็นวิธีที่ใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้มแบบเส้นตรง มีปัจจัยแนวโน้มเพิ่มขึ้นหรือลดลงอย่างสม่ำเสมอตามเวลา และไม่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โดยมีขั้นตอนดังนี้

1. กำหนดค่าเริ่มต้นของการพยากรณ์

2. กำหนดค่าปรับน้ำหนัก (α) โดยที่ $0<\alpha<1$
3. คำนวณค่าพยากรณ์แบบเอ็กซ์โปเนนเชียลอย่างง่ายจากข้อมูล y_t หรือค่าปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ณ เวลา t โดย

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

4. คำนวณค่าพยากรณ์แบบเอ็กซ์โปเนนเชียลอย่างง่ายจาก $\hat{y}_t$ หรือค่าปรับให้เรียบครั้งที่สองแบบเอ็กซ์โปเนนเชียล ณ เวลา t โดย

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

5. คำนวณค่าประมาณพารามิเตอร์

$$a_t = 2A_t - A'_t$$
$$= A_t + \left(1-(1-\alpha)^2\right)e_t$$
$$b_t = \frac{\alpha}{1-\alpha}(A_t - A'_t)$$
$$= b_{t-1} + \alpha^2 e_t$$

6. คำนวณค่าพยากรณ์จากสูตร

$$\hat{y}_{t+m} = a_t + b_t m \quad , m \geq 1$$

เมื่อ y_t = ค่าสังเกต/ข้อมูล ณ เวลา t
$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, $m \geq 1$
α = ค่าปรับน้ำหนัก ; $0<\alpha<1$
a_t = ค่าแนวโน้ม ณ เวลา t
b_t = ค่าประมาณพารามิเตอร์ ณ เวลา t
A_t = ค่าปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ณ เวลา t
A'_t = ค่าปรับให้เรียบครั้งที่สองแบบเอ็กซ์โปเนนเชียล ณ เวลา t

4. วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) ลักษณะคล้ายกับวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลตามแบบของบราวน์ (DES) ต่างกันแต่เพียงวิธีการของโฮลท์ มีค่าคงที่ปรับให้เรียบสำหรับแนวโน้มโดยเฉพาะ ซึ่งจะมีผลให้ติดตามแนวโน้มได้ดีขึ้น โดนใช้พารามิเตอร์คนละตัว จึงทำให้วิธีของโฮลท์มีพารามิเตอร์ 2 ตัว เหมาะกับข้อมูลที่มีแนวโน้มในรูปเส้นตรงที่เพิ่มขึ้นหรือลดลงอย่างสม่ำเสมอ ต้องมีข้อมูลอย่างน้อย 4 ตัว ใช้พยากรณ์ข้อมูลอนุกรมเวลาในระยะใกล้ ระยะสั้น และอาจใช้ในระยะปานกลางได้ โดยมีขั้นตอนดังนี้

1. กำหนดค่าเริ่มต้นของการพยากรณ์
2. กำหนดค่าปรับน้ำหนัก 2 ค่า (α และ β) โดยที่ $0<\alpha<1$, $0<\beta<1$ และทำให้ SSE มีค่าต่ำสุด
3. คำนวณค่าประมาณค่าพารามิเตอร์

$$a_t = \alpha y_t + (1-\alpha)A_t$$
$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

4. คำนวณค่าพยากรณ์จากสูตร

$$\hat{y}_{t+m} = a_t + b_t m \qquad , m \geq 1$$

เมื่อ y_t = ค่าสังเกต/ข้อมูล ณ เวลา t

$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, m≥1

α = ค่าคงที่ปรับให้เรียบระหว่างข้อมูลจริงกับค่าพยากรณ์ ; $0<\alpha<1$

β = ค่าคงที่ปรับให้เรียบระหว่างค่าแนวโน้มจริงกับค่าพยากรณ์ของแนวโน้ม; $0<\beta<1$

a_t, b_t = ค่าประมาณพารามิเตอร์ ณ เวลา t

5. วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) เป็นวิธีที่ใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้มเป็นเส้นโค้งแบบครอดราติก (quadratic) ซึ่งวิธีการจะคล้ายกับวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (DES) แต่เพิ่มพารามิเตอร์เป็น 3 ตัว โดยมีขั้นตอนดังนี้

1. กำหนดค่าเริ่มต้นของการพยากรณ์
2. กำหนดค่าปรับน้ำหนัก(α) โดยที่ $0<\alpha<1$ และทำให้ SES มีค่าต่ำสุด
3. คำนวณค่าปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ณ เวลา t โดย

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

4. คำนวณค่าปรับให้เรียบครั้งที่สองแบบเอ็กซ์โปเนนเชียล ณ เวลา t โดย

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

5. คำนวณค่าปรับให้เรียบครั้งที่สามแบบเอ็กซ์โปเนนเชียล ณ เวลา t โดย

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

6. คำนวณค่าประมาณพารามิเตอร์

$$a_t = 3A_t - 3A'_t + A''_t$$
$$b_t = \frac{\alpha}{2(1-\alpha)^2}\left((6-5\alpha)A_t - 2(5-4\alpha)A'_t + (4-3\alpha)A''_t\right)$$
$$c_t = \frac{\alpha^2}{2(1-\alpha)^2}\left(A_t - 2A'_t + A''_t\right)$$

7. คำนวณค่าพยากรณ์จากสูตร

$$\hat{y}_t = a_t + b_t m + c_t m^2 \qquad , m \geq 1$$

เมื่อ y_t = ค่าสังเกต/ข้อมูล ณ เวลา t

$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, $m \geq 1$

α = ค่าน้ำหนักของการเฉลี่ย; $0<\alpha<1$

a_t, b_t, c_t = ค่าประมาณพารามิเตอร์ ณ เวลา t

A_t = ค่าปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ณ เวลา t

A'_t = ค่าปรับให้เรียบครั้งที่สองแบบเอ็กซ์โปเนนเชียล ณ เวลา t

A''_t = ค่าปรับให้เรียบครั้งที่สามแบบเอ็กซ์โปเนนเชียล ณ เวลา t

6. วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) วิธีนี้ปรับปรุงมาจากวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (LES) เป็นวิธีที่เหมาะสมสำหรับข้อมูลอนุกรมเวลาที่มีแนวโน้มและอิทธิพลของฤดูกาล ใช้พยากรณ์ในระยะสั้นจนถึงระยะปานกลาง ข้อมูลไม่ควรเป็นรายปี เพราะจะทำให้ไม่สามารถแยกอิทธิพลของฤดูกาลได้ รูปแบบอาจจะเป็นทั้งแบบบวกและแบบคูณ มีการประมาณค่าพารามิเตอร์ 3 ตัว โดยมีขั้นตอน ดังนี้

1. กำหนดค่าเริ่มต้นสำหรับการพยากรณ์ ซึ่งมี 2+s ค่า คือค่าเริ่มต้นของการพยากรณ์ 1 ค่า ค่าเริ่มต้นของแนวโน้ม 1 ค่า และค่าเริ่มต้นของฤดูกาล s ค่า เมื่อ s คือ คาบเวลาของฤดูกาลในแต่ละปี เช่น ข้อมูลรายเดือน ค่า s เท่ากับ 12 ข้อมูลรายไตรมาส ค่า s เท่ากับ 4 เป็นต้น
2. กำหนดค่าปรับน้ำหนัก 3 ค่า (α, β และ γ) โดยที่ $0<\alpha<1$, $0<\beta<1$, $0<\gamma<1$ และทำให้ SSE มีค่าต่ำสุด
3. คำนวณค่าประมาณพารามิเตอร์

   3.1 รูปแบบบวก

   Level; $L_t = \alpha(y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + b_{t-1})$

   Trend; $b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$

   Seasonal; $S_t = \gamma(y_t - L_t) + (1-\gamma)S_{t-s}$

   3.2 รูปแบบคูณ

   Level; $L_t = \alpha \dfrac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1})$

   Trend; $b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$

   Seasonal; $S_t = \gamma \dfrac{y_t}{L_t} + (1-\gamma)S_{t-s}$

4. คำนวณค่าพยากรณ์จากสูตร

   4.1 รูปแบบบวก

   $$\hat{y}_{t+m} = L_t + b_t m + S_{t-s+m} \quad , m \geq 1$$

4.2 รูปแบบคูณ

$$\hat{y}_{t+m} = (L_t + b_t m)S_{t-s+m} \quad , m \geq 1$$

เมื่อ y_t = ค่าสังเกต/ข้อมูล ณ เวลา t

$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, $m \geq 1$

L_t, b_t, S_t = ค่าประมาณพารามิเตอร์ ณ เวลา t

α = ค่าปรับให้เรียบระหว่างข้อมูลจริงกับค่าพยากรณ์ ; $0 < \alpha < 1$

β = ค่าคงที่ปรับให้เรียบระหว่างค่าแนวโน้มจริงกับค่าพยากรณ์ของแนวโน้ม; $0 < \beta < 1$

γ = ค่าคงที่ปรับให้เรียบระหว่างฤดูกาลจริงกับค่าพยากรณ์ของฤดูกาล; $0 < \beta < 1$

s = คาบเวลาของฤดูกาลในแต่ละปี

7. วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสำหรับข้อมูลที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential smoothing for exponential trend: ESE) เป็นวิธีที่ใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล และไม่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โดยมีขั้นตอนดังนี้

1. กำหนดค่าเริ่มต้นของการพยากรณ์
2. กำหนดค่าปรับน้ำหนัก 2 ค่า (α และ β) โดยที่ $0 < \alpha < 1$, $0 < \beta < 1$ แล้ทำให้ SSE มีค่าต่ำสุด
3. คำนวณค่าพยากรณ์จากสูตร

$\hat{Y}_{t+1} = S_{t+1}T_{t+1}$ หรือ $\hat{Y}_{t+m} = S_{t+m}T_{t+m}$

จาก $S_t = S_{t-1}T_{t-1} + \alpha e_t$ ดังนั้น $S_{t+1} = S_tT_t + \alpha e_{t+1}$

$T_t = T_{t-1} + \alpha\beta e_t/S_{t-1}$ ดังนั้น $T_{t+1} = T_t + \alpha\beta e_{t+1}/S_t$

เมื่อ y_t = ค่าสังเกต/ข้อมูล ณ เวลา t

$\hat{y}_{t+m}$ = ค่าพยากรณ์ ณ เวลา t+m, $m \geq 1$

α = ค่าคงที่ปรับให้เรียบระหว่างข้อมูลจริงกับค่าพยากรณ์ ; $0 < \alpha < 1$

β = ค่าคงที่ปรับให้เรียบระหว่างค่าแนวโน้มจริงกับค่าพยากรณ์ของแนวโน้ม; $0 < \beta < 1$

S_t = ค่าเริ่มต้นอนุกรมเวลา ณ เวลา t

T_t = ค่าเริ่มต้นแนวโน้ม ณ เวลา t

8. วิธีการวิเคราะห์การถดถอย (Regression analysis: REG) การวิเคราะห์ข้อมูลอนุกรมเวลาที่ใช้หลักการวิเคราะห์การถดถอย มีตัวแปรอนุกรมเวลาเป็นตัวแปรตาม และมีตัวแปรอิสระที่ความหมายต่าง ๆ กันไป (ทรงศิริ แต้สมบัติ, 2539) ในที่นี้ผู้วิจัยจะกล่าวถึงใน 2 ความหมาย ดังนี้

8.1 ตัวแปรเวลา (time variable) เป็นตัวแปรที่ใช้เมื่ออนุกรมเวลามีแนวโน้ม โดยจะกำหนดตัวแปรเวลาเป็นตัวแปรอิสระที่มีค่าเป็นรหัส (coded value) แทนวัน เดือน ไตรมาส หรือปีที่เก็บรวบรวมข้อมูล รูปแบบแนวโน้มมีหลายแบบขึ้นอยู่กับลักษณะของอนุกรมเวลาแนวโน้มที่นิยมใช้ ได้แก่

แนวโน้มเส้นตรง $y_t = \beta_0 + \beta_1 t + \varepsilon_t$

แนวโน้มแบบเส้นโค้ง $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$

แนวโน้มแบบ exponential $y_t = \beta_0 \beta_1^{\,t} \varepsilon_t$

การใช้ตัวแปรเวลาเป็นตัวแปรอิสระนั้น ไม่ว่ารูปแบบการถดถอยจะเป็นแบบใด ต้องกำหนดค่าให้กับตัวแปรเวลา (t) โดยจะมีค่าเริ่มต้นเท่าใดก็ได้ แต่จะต้องเพิ่มขึ้นเท่า ๆ กัน ในแต่ละช่วงเวลา

8.2 ตัวแปรดัมมี่ (dummy variable) เป็นตัวแปรที่กำหนดขึ้นเพื่อบอกว่าค่าสังเกตในอนุกรมเวลาเกิดขึ้นในฤดูกาลใด ค่าตัวแปรดัมมี่จะเป็น 1 เมื่อค่าสังเกตอยู่ในฤดูกาลที่กำหนด ค่าตัวแปรดัมมี่จะเป็น 0 เมื่อค่าสังเกตไม่อยู่ในฤดูกาลที่กำหนด และตัวแปรดัมมี่จะมีน้อยกว่าจำนวน ฤดูกาลอยู่ 1 ดังนี้

แบบบวก

รูปแบบฤดูกาล $y_t = \beta_0 + \beta_2 x_{1t} + \beta_3 x_{2t} + \ldots + \beta_L x_{(L-1)t} + \varepsilon_t$

รูปแบบแนวโน้มเส้นตรงและฤดูกาล

$$y_t = \beta_0 + \beta_1 t + \beta_2 x_{1t} + \beta_3 x_{2t} + \ldots + \beta_L x_{(L-1)t} + \varepsilon_t$$

เมื่อ β_0 = ค่าจุดตัด (y-intercept)

β_1 = ค่าอัตราการเพิ่มขึ้นหรือลดลงของ y_t เมื่อ t เพิ่มขึ้น 1 หน่วย

β_i = ค่าวัดอิทธิพลของฤดูกาลที่ i, i = 2,3,...,L

$$x_{it} = \begin{cases} 1 & \text{เมื่อค่าสังเกต } t \text{ เป็นค่าในฤดูกาลที่ } i \\ 0 & \text{เมื่อค่าสังเกต } t \text{ ไม่เป็นค่าในฤดูกาลที่ } I \end{cases}$$

รูปแบบแนวโน้มเส้นโค้งและฤดูกาล

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 x_{1t} + \beta_4 x_{2t} + \ldots + \beta_{L+1} x_{(L-1)t} + \varepsilon_t$$

แบบคูณ

รูปแบบฤดูกาล $y_t = \beta_0 \beta_2^{x_{1t}} \beta_3^{x_{2t}} \ldots \beta_L^{x_{(L-1)t}} \varepsilon_t$

รูปแบบแนวโน้มเส้นตรงและฤดูกาล $y_t = \beta_0 \beta_1^{\,t} \beta_2^{x_{1t}} \beta_3^{x_{2t}} \ldots \beta_L^{x_{(L-1)t}} \varepsilon_t$

สำหรับรูปแบบคูณ ต้องแปลงให้อยู่ในรูปแบบบวก โดยการหา logarithm คือ แปลง y_t ให้เป็น $\ln y_t$ หรือ y'_t ดังนี้

รูปแบบฤดูกาล

$$\ln y_t = (\ln\beta_0) + (\ln\beta_2)x_{1t} + \ldots + (\ln\beta_L)x_{(L-1)t} + (\ln\varepsilon_t)$$

$$y'_t = \beta'_0 + \beta'_2 x_{1t} + \ldots + \beta'_L x_{(L-1)t} + \varepsilon'_t$$

รูปแบบแนวโน้มเส้นตรงและฤดูกาล

$$\ln y_t = (\ln\beta_0) + (\ln\beta_1)t + (\beta_2)x_{1t} + \ldots + (\ln\beta_L)x_{(L-1)t} + (\ln\varepsilon_t)$$

$$y'_t = \beta'_0 + \beta'_1 t + \beta'_2 x_{1t} + \beta'_3 x_{2t} + \ldots + \beta'_L x_{(L-1)t} + \varepsilon'_t$$

เมื่อ β_0 = ค่าจุดตัด (y-intercept)

β_i = ค่าวัดอิทธิพลของฤดูกาลที่ i, i = 2,3,...,L

$$x_{it} = \begin{cases} 1 & \text{เมื่อค่าสังเกต t เป็นค่าในฤดูกาลที่ i} \\ 0 & \text{เมื่อค่าสังเกต t ไม่เป็นค่าในฤดูกาลที่ i} \end{cases}$$

ซึ่งในการพยากรณ์ร่วม (combined forecasts) นี้จะมีการให้น้ำหนักเฉลี่ย (weighted average) ตามแบบของ Newbold และ Granger (1974) ที่ได้ผลดีที่สุด มี 2 แบบ คือ

1. $$w_i = \frac{\left(\sum_{t=n-\upsilon}^{n-1} e_t^{(i)2}\right)^{-1}}{\sum_{j=1}^{p}\left(\sum_{t=n-\upsilon}^{n-1} e_t^{(j)2}\right)^{-1}}$$

2. $$w_i = \beta\left[\frac{\left(\sum_{t=n-\upsilon-1}^{n-2} e_t^{(i)2}\right)^{-1}}{\sum_{j=1}^{p}\left(\sum_{t=n-\upsilon-1}^{n-2} e_t^{(j)2}\right)^{-1}}\right] + (1-\beta)\left[\frac{\left(\sum_{t=n-\upsilon}^{n-1} e_t^{(i)2}\right)^{-1}}{\sum_{j=1}^{p}\left(\sum_{t=n-\upsilon}^{n-1} e_t^{(j)2}\right)^{-1}}\right], \quad 0<\beta<1$$

ค่าการพยากรณ์ คือ $\hat{y}_t = \sum_{i=1}^{p} w_i \hat{y}_t^{(i)}$

โดยที่ $e_t^{(i)} = \dfrac{\left(y_t - \hat{y}_t^{(i)}\right)}{y_t}$, i = 1,2,3,...,p

เมื่อ n = จำนวนข้อมูล

$\hat{y}_t^{(i)}$ = ค่าพยากรณ์ ณ เวลาที่ t วิธีพยากรณ์ที่ i

$\hat{y}_t$ = ค่าพยากรณ์ ณ เวลาที่ t

w_i = น้ำหนักของวิธีการพยากรณ์ที่ i

p = จำนวนวิธีการพยากรณ์

การให้น้ำหนักเฉลี่ยในการพยากรณ์ร่วมตามแบบของ Newbold และ Granger (1974) ทั้ง 2 แบบ จะเห็นได้ว่ามีการพิจารณาจากค่าความคลาดเคลื่อนของวิธีที่นำมารวมกัน โดยแบบแรกจะให้น้ำหนักเท่ากันสำหรับวิธีการพยากรณ์เดี่ยวที่นำมารวมกัน ส่วนแบบที่สองจะมีการปรับการให้น้ำหนักของวิธีการพยากรณ์เดี่ยวที่นำมารวมกัน โดยมีการคำนึงถึงความสำคัญของน้ำหนักในช่วงเวลาที่ผ่านมาก่อน 1 ช่วงเวลาด้วย โดยให้ความสำคัญของน้ำหนักใน 2 ช่วงเวลาที่ต่างกันลดลงแบบเรขาคณิต

ตอนที่ 6 เกณฑ์ในการเปรียบเทียบความสามารถของวิธีการพยากรณ์

การตัดสินว่าวิธีการพยากรณ์ใดเหมาะสมและมีความถูกต้องในการพยากรณ์มากที่สุด สามารถพิจารณาได้จากค่าความคลาดเคลื่อนระหว่างค่าพยากรณ์ที่ได้กับค่าจริง ซึ่งค่าความคลาดเคลื่อนที่นิยมใช้กันมาก ได้แก่ ค่าความคลาดเคลื่อนเฉลี่ยสมบูรณ์ (mean absolute error: MAE) ค่าเบี่ยงเบนสมบูรณ์เฉลี่ย (mean absolute deviation: MAD) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error: MSE) รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (root mean square error: RMSE) และค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (mean absolute percentage error: MAPE) เป็นต้น

ในการวิจัยครั้งนี้ผู้วิจัยใช้ค่าความคลาดเคลื่อน 3 แบบ คือ ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (mean absolute percentage error: MAPE) เนื่องจากเป็นค่าวัดความถูกต้องของการพยากรณ์ที่วัดจากความคลาดเคลื่อนของการพยากรณ์ที่เทียบกับค่าจริง และไม่มีหน่วยจึงเหมาะที่จะใช้กับการเปรียบเทียบวิธีการพยากรณ์หลายวิธีเมื่อใช้อนุกรมเวลาชุดเดียวกัน (Shearer, 1994; Hanke and Reitsch, 1995; ทรงศิริ แต้สมบัติ, 2539 อ้างถึงใน ทัศนีย์ อินทนู, 2543) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error: MSE) ซึ่งเหมาะสำหรับใช้เปรียบเทียบความถูกต้องของวิธีการพยากรณ์ของข้อมูลอนุกรมเวลาในชุดเดียวกัน (Thompson, 1990 อ้างถึงในทัศนีย์ อินทนู, 2543) และค่าความคลาดเคลื่อนเฉลี่ย (mean error: ME) เป็นค่าที่บอกได้ว่าค่าพยากรณ์นั้นสูงหรือต่ำกว่าค่าจริงทั้งหมดหรือบางส่วน มีสูตร ดังนี้

$$\text{MAPE} = \frac{\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right|}{n} \times 100$$

$$\text{MSE} = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}$$

$$\text{ME} = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)}{n}$$

เมื่อ $\hat{y}_t$ = ค่าพยากรณ์ ณ เวลาที่ t

y_t = ข้อมูลจริง ณ เวลาที่ t

n = จำนวนข้อมูล

สำหรับค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) มีเกณฑ์ดังตาราง 9 (Lewis, 1982)

ตาราง 9 การแปลผลค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย

ค่า MAPE (%)	การแปลผล
<10	ค่าพยากรณ์มีความถูกต้องสูงมาก
10-20	ค่าพยากรณ์มีความถูกต้องดี
20-50	ค่าพยากรณ์เชื่อถือได้
>50	ค่าพยากรณ์ไม่ถูกต้อง

และพิจารณาวิธีการพยากรณ์จากค่าความคลาดเคลื่อนเฉลี่ย (ME) ดังนี้

ถ้า ME มีค่าสูงเป็นบวก แสดงว่า ค่าพยากรณ์ส่วนใหญ่น้อยกว่าค่าจริง

ถ้า ME มีค่าสูงเป็นลบ แสดงว่า ค่าพยากรณ์ส่วนใหญ่มากกว่าค่าจริง

ถ้า ME มีค่าเข้าใกล้ 0 แสดงว่า ค่าพยากรณ์ใกล้เคียงกับค่าจริง

ตอนที่ 7 งานวิจัยที่เกี่ยวข้อง

สาระในตอนนี้ผู้วิจัยนำเสนองานวิจัยที่เกี่ยวข้องกับการพยากรณ์ที่ใช้เทคนิคการพยากรณ์ด้วยวิธีของเบย์ วิธีบ็อกซ์-เจนกิ้นส์ และวิธีพยากรณ์ร่วม ดังนี้

วันพร เหลืองอาภาพงษ์ (2520) ได้ทำพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการส่งออกข้าว ยาง และข้าวโพดด้วยวิธีบ็อกซ์-เจนกิ้นส์ โดยใช้ข้อมูลอนุกรมเวลารายเดือนตั้งแต่ พ.ศ. 2513-2518 พบว่า อนุกรมเวลาปริมาณข้าว และยาง เป็นอนุกรมเวลาที่คงที่ รูปแบบที่เหมาะสมกับข้อมูล คือ ARIMA (2,0,0) ส่วนอนุกรมเวลาปริมาณส่งออกข้าวโพดเป็นอนุกรมเวลาที่มีฤดูกาล รูปแบบที่เหมาะสมกับข้อมูล คือ ARIMA (0,1,0) $\times$ SARIMA $(0,1,1)_{12}$ และ ผลการพยากรณ์ปริมาณส่งออกทั้งสามชนิด ได้ค่าพยากรณ์อยู่ในเกณฑ์ไม่แตกต่างจากปีที่ผ่านมาเท่าใดนัก

บุษบา พิกุลผล (2522) ได้ทำการเปรียบเทียบรูปแบบที่ใช้คาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยด้วยการวิเคราะห์อนุกรมเวลา และวิธีบ็อกซ์-เจนกิ้นส์ โดยการศึกษารูปแบบที่เหมาะสมกับข้อมูลรายเดือนของจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทย ตั้งแต่ พ.ศ. 2506-2520 ซึ่งข้อมูลนี้จะใช้ในการคาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยต่อไป และจากการศึกษาพบว่า รูปแบบที่เหมาะสมที่สุด คือรูปแบบที่อธิบายการเปลี่ยนแปลงเนื่องจากวัฏจักรและเหตุการณ์ผิดปกติร่วมกัน โดยใช้วิธีบ็อกซ์-เจนกิ้นส์ และปรับด้วยค่าแนวโน้มของข้อมูลและการเปลี่ยนแปลงเนื่องจากฤดูกาล

วัลลภ โรจนศิริวณิชย์ (2528) ได้ทำการวิเคราะห์ข้อมูลราคาในอดีต โดยใช้การวิเคราะห์อนุกรมเวลาเพื่อหาว่าหลักทรัพย์ใดควรใช้เทคนิคการพยากรณ์ใด และใช้ข้อมูลย้อนหลังเท่าไรจึงจะเหมาะสม โดยเปรียบเทียบระหว่างอนุกรมบ็อกซ์-เจนกิ้นส์ การเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง และเทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง พบว่า วิธีบ็อกซ์-เจนกิ้นส์ ใช้ในการพยากรณ์ได้ดีกว่าการเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง และเทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง ยกเว้นข้อมูลราคารายสัปดาห์ของธนาคารกรุงเทพ การเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง ใช้ได้ดีที่สุด และบริษัทมาบุญครองอบพืชและไซโล เทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้งใช้ได้ดีที่สุด

เกศิณี กมลรัตน์ (2529) ได้ทำการพยากรณ์ข้อมูลเงินอากรขาเข้าด้วยวิธี SES, วิธี DES, วิธีการกรองแบบปรับได้, การวิเคราะห์อนุกรมเวลาแบบคลาสสิค และ วิธีบ็อกซ์-เจนกิ้นส์ พบว่าเมื่อขนาดตัวอย่างเป็น 50, 60, 70, 80, 90, 100, 110 และ 120 วิธีบ็อกซ์-เจนกิ้นส์ใช้พยากรณ์ได้ดีที่สุด

ธิดารัตน์ จันทวี (2539) ได้ทำการพยากรณ์ความต้องการใช้ไฟฟ้าเพื่อการผลิตไฟฟ้าระยะสั้น ด้วยวิธีการปรับให้เรียบ และวิธีบ็อกซ์-เจนกิ้นส์ พบว่า วิธีการพยากรณ์ที่มีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) ต่ำที่สุดในทุกกรณี คือ วิธีบ็อกซ์-เจนกิ้นส์

อมรรัตน์ ปรารมย์ (2539) ได้ทำการเปรียบเทียบวิธีการพยากรณ์สำหรับข้อมูลอนุกรมเวลา ระหว่างการเลือกใช้วิธีการพยากรณ์แบบใดแบบหนึ่ง กับการใช้ค่าพยากรณ์ร่วม วิธี SES, DES, LES และการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลแบบอัตราส่วน พบว่า ข้อมูลอนุกรมเวลาที่มีขนาดกลาง (n=30) และข้อมูลขนาดใหญ่ (n=50, 70) มีค่าความคลาดเคลื่อนในการพยากรณ์ต่ำที่สุด

บำเพ็ญ ปิดชิด (2540) ได้ทำการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ด้วยวิธี REG, วิธี SMA, วิธี SES และวิธีบ็อกซ์-เจนกิ้นส์ พบว่า การพยากรณ์ข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล วิธีบ็อกซ์-เจนกิ้นส์ มีขนาดความคลาดเคลื่อนน้อยที่สุด

ทัศนีย์ อินหนู (2543) ได้ทำการเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีพยากรณ์เดี่ยว และวิธีพยากรณ์ร่วม โดยใช้ข้อมูลอนุกรมเวลาทางการศึกษาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล วิธี SMA, SES, DES, LES, TES, HWS และ REG พบว่า วิธีการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ ปริมาณการยืมหนังสือภาษาไทย และวิทยานิพนธ์ คือ วิธีการพยากรณ์ร่วม ที่ได้จากการรวมของวิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) กับวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) ซึ่งเป็นวิธีที่เหมาะสมกับข้อมูลทั้ง 2 ชุด โดยให้น้ำหนักแบบที่ 2 แบบที่ 1 ตามลำดับ สำหรับข้อมูลที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ จำนวนครูวิทยาศาสตร์และจำนวนครูคณิตศาสตร์ คือ วิธีพยากรณ์ร่วม ที่ได้จากการนำวิธีการพยากรณ์ทุกวิธีมารวมกัน

เอกภพ ยานะวิมุติ (2543) ได้ทำการเปรียบเทียบความคลาดเคลื่อนระหว่างวิธีบ็อกซ์-เจนกิ้นส์ที่ใช้โมเดลสมการโครงสร้าง กับ วิธีบ็อกซ์-เจนกิ้นส์ที่ใช้ตัวบ่งชี้นำ ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่ไม่คงที่ ด้วยวิธีบ็อกซ์-เจนกิ้นส์, วิธีบ็อกซ์-เจนกิ้นส์ที่ใช้โมเดลเชิงโครงสร้าง และวิธีบ็อกซ์-เจนกิ้นส์ที่ใช้ตัวบ่งชี้นำ พบว่า ผลการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการยืมหนังสือทั่วไป และหนังสือสำรองด้วยวิธีบ็อกซ์-เจนกิ้นส์ มีขนาดความคลาดเคลื่อนน้อยที่สุด

อรุณี หงษ์ศิริวัฒน์ (2543) ได้ทำการเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์จากวิธีบ็อกซ์และเจนกิ้นส์ วิธีของบ็อกซ์และเจนกิ้นส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย ในการพยากรณ์อนุกรมเวลาทางการศึกษาที่คงที่และไม่คงที่ พบว่า วิธีการพยากรณ์ที่เหมาะสมสำหรับข้อมูลอนุกรมเวลาจำนวนผู้เข้าใช้บริการห้องสมุด คือวิธีบ็อกซ์และเจนกิ้นส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และสำหรับข้อมูลอนุกรมเวลาปริมาณการยืมหนังสือระหว่างห้องสมุด คือวิธีบ็อกซ์และเจนกิ้นส์ และวิธีบ็อกซ์และเจนกิ้นส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน

Winkler และ Makridakis (1983) ได้ทำการศึกษาวิธีพยากรณ์ร่วม โดยให้น้ำหนักของ Newbold และ Granger (1974) 5 แบบ จากวิธีการพยากรณ์ 10 วิธี คือ วิธีง่าย (naive) วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของ บราวน์ (Brown's triple exponential smoothing: TES) วิธีการวิเคราะห์การถดถอย (Regression analysis: REG) วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) วิธีการปรับให้เรียบแบบปรับอัตราส่วน (adaptive response rate exponential smoothing: ARRES) และวิธี Automatic AEP โดยใช้ข้อมูล 1,001 ชุด (series) จากแหล่งข้อมูลต่าง ๆ หลายแหล่งและหลายประเภท มีทั้งข้อมูลที่เป็นรายปี รายไตรมาส และรายเดือน พบว่าวิธีการพยากรณ์ร่วมโดยการให้น้ำหนักเฉลี่ยของ Newbold และ Granger แบบที่ 1 และแบบที่ 3 เป็นวิธีที่ดีที่สุดในการพยากรณ์ โดยมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) น้อยที่สุด

Makridakis และ Winkler (1983) ได้ทำการศึกษาเพื่อตรวจสอบความถูกต้องของการพยากรณ์ร่วม โดยใช้ข้อมูลอนุกรมเวลา 111 ชุด และ 1,001 ชุด ทั้งประเภทรายเดือน รายไตรมาส และรายปี ใช้วิธีการพยากรณ์ร่วมโดยให้น้ำหนักแบบ simple everage จากวิธีการพยากรณ์เดี่ยว 14 วิธี ได้แก่ วิธีง่าย (naive) วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) วิธีการปรับให้เรียบแบบปรับอัตราส่วน (adaptive response rate exponential smoothing: ARRES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2

ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) วิธีการวิเคราะห์การถดถอย (Regression analysis: REG) วิธี Automatic AEP วิธีบ๊อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) วิธี Lewandowski's FORSYS system วิธี Parzen's ARIMA methodology และวิธี Bayesian Forecasting โดยใช้ MAPE เป็นเกณฑ์ในการตัดสิน พบว่า ความถูกต้องของการพยากรณ์ร่วมขึ้นอยู่กับจำนวนวิธีที่นำมารวมกัน ถ้าใช้จำนวนวิธีการพยากรณ์ในการนำมารวมกันเพิ่มมากขึ้น ค่าความถูกต้องก็จะเพิ่มตาม

Attwell และ Smith (1991) ได้ทำการประยุกต์ใช้วิธีของเบย์ทำการการพยากรณ์ราคาประมูลที่แต่ละบริษัทเสนอเพื่อประมูลงาน ช่วยในการพิจารณาปัญหาเฉพาะหน้าของบริษัทในการเลือกราคาประมูลงานที่บริษัทอื่นเสนอให้ โดยประยุกต์ใช้ Gate model ในลักษณะของ Dirichlet process ที่มีลักษณะชี้วัดราคาประมูลที่เป็นไปได้ในการเปิดประมูล โมเดลจะให้วิธีการที่ชัดเจนสำหรับการคำนวณค่าความน่าจะเป็น และจะเลือกราคาประมูลต่ำสุดจากราคาที่ถูกเสนอ ความคล่องตัวหรือการเปลี่ยนแปลงได้ของโมเดลทำให้การพยากรณ์มีประสิทธิภาพมากขึ้นสำหรับการใช้งานกับบริษัทต่าง ๆ ในการทำสัญญาประมูล ซึ่งประโยชน์ที่ใช้ Dirichlet model มี 2 ประการ คือ ประการแรก ผลที่ได้เกิดจากการเฟ้นสุ่ม และประการที่สอง ไม่ต้องมีการระบุความน่าจะเป็นต่อเนื่องเหมือนในกรณีของ Gate model โดยทั่วไป

Batchelor และ Dua (1995) ได้ทำการศึกษาโดยใช้ชุดข้อมูลอนุกรมเวลารายเดือน 4 ประเภท คือ real GNP growth (RGNP), inflation in the implicit GNP deflator (PGNP), the growth of corporate profits (PROF) และ the umemployment rate (UNMP) ของนักพยากรณ์ 22 คน โดยมีเป้าหมายในการพยากรณ์ครอบคลุม 10 ปี คือ ค.ศ. 1978-1987 พบว่า ในข้อมูลทุกประเภทยิ่งทำการรวม (combine) ด้วยทฤษฎีและเทคนิคการพยากรณ์ที่หลากหลายก็จะยิ่งให้ผลที่ดีขึ้นทำให้ค่าของความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ลดลง และมีค่าความน่าจะเป็นในการลดลงของความคลาดเคลื่อนได้มากกว่าการที่รวมด้วยทฤษฎีและเทคนิคการพยากรณ์ที่เหมือนหรือคล้ายกัน

Pammer, Fong และ Arnold (2000) ได้ประยุกต์ใช้วิธีของเบย์ในการพยากรณ์การแพร่กระจายสินค้าใหม่ออกสู่ท้องตลาด โดยรวบรวมข้อมูลเบื้องต้นจากสินค้าที่มีอยู่จริง และ/หรือการวินิจฉัยจากการจัดการ ในการวิเคราะห์ข้อมูลนั้นสมมติให้โค้งการแพร่กระจายเป็นฟังก์ชันที่ไม่ลดลงตามเวลาและรูปร่างคงที่ไม่เปลี่ยนแปลง แล้วใช้เทคนิค Markov-chain Monte Carlo คำนวณค่าพยากรณ์ด้วยวิธีของเบย์ นอกจากนี้ยังทดลองใช้กับข้อมูลจริง คือ การพยากรณ์การแพร่กระจายของโทรทัศน์สี โดยใช้ข้อมูลจากโทรทัศน์ขาว-ดำ พบว่า วิธีของเบย์สามารถใช้ได้ดี และยังสามารถนำไปใช้วิเคราะห์กับสถานการณ์อื่นได้หลากหลาย และเชื่อมั่นได้

Settimit และ Smith (2000) ทำการเปรียบเทียบประสิทธิภาพของการใช้ Linear Bayes method, Algebraic approximate method และ Gibbs sampler method สำหรับ non-Gaussian dynamic processes โดยใช้เทคนิค Markov-chain Monte Carlo เพื่อวิเคราะห์อนุกรมเวลาที่มีลักษณะไม่เป็นเส้นตรง และถึงแม้ว่าเทคนิค Markov-chain Monte Carlo จะไม่รวดเร็วแต่ก็สามารถทำให้ศึกษาถึงประสิทธิภาพของการทำนายค่าในอนาคตของวิธีการพยากรณ์ทางพีชคณิตที่มีประโยชน์และรวดเร็ว จากการนี้ทำการวิเคราะห์ด้วย Poisson model พบว่า Linear Bayes method ให้ผลไม่ดีเท่ากับ Algebraic approximate method และ Gibbs sampler method เมื่อโมเดลมีลักษณะเหมาะสมกับข้อมูล

de Alba และ Mendoza (2001) ได้ทำการพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีของเบย์เปรียบเทียบกับวิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) โดยใช้ข้อมูล 2 ฐาน ฐานแรก คือ ข้อมูลปริมาณการใช้ไฟฟ้าของเมืองไอโอวา เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 1971 ถึง เดือนตุลาคม 1978 เป็นข้อมูลที่มีฤดูกาล (seasonal) และฐานที่สอง คือ ค่าใช้จ่ายในการจัดการธนาคาร Maxican เป็นข้อมูลรายเดือนตั้งแต่เดือนมกราคม 1987 ถึง เดือนตุลาคม 1994 เป็นข้อมูลที่มีแนวโน้ม (trend) พบว่า ในกรณีที่ข้อมูลอนุกรมเวลามีขนาดเล็ก วิธีของเบย์มีค่า MAPE ต่ำกว่าวิธีบ็อกซ์-เจนกิ้นส์ สำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลและข้อมูลอนุกรมเวลาที่มีแนวโน้ม

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการใช้วิธีของเบย์ (Bayesian method) วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) และวิธีการพยากรณ์ร่วม (combined forecasting) สามารถสรุปได้ดังตาราง 10, 11 และ 12 ตามลำดับ

ตาราง 10 สรุปงานวิจัยที่เกี่ยวข้องกับการใช้วิธีของเบย์ (Bayesian method)

นักวิจัย	วัตถุประสงค์	ผลการวิจัย
Attwell และ Smith (1991)	ได้ทำการประยุกต์ใช้วิธีของเบย์ทำการการพยากรณ์ราคาประมูลที่แต่ละบริษัทเสนอเพื่อประมูลงาน ช่วยในการพิจารณาปัญหาเฉพาะหน้าของบริษัทในการเลือกราคาประมูลงานที่บริษัทอื่นเสนอให้ โดยประยุกต์ใช้ Gate model ในลักษณะของ Dirichlet process	พบว่า โมเดลจะให้วิธีการที่ชัดเจนสำหรับการคำนวณค่าความน่าจะเป็น และจะเลือกราคาประมูลต่ำสุดจากราคาที่ถูกเสนอ ความคล่องตัวหรือการเปลี่ยนแปลงได้ของโมเดลทำให้การพยากรณ์มีประสิทธิภาพมากขึ้นสำหรับการใช้งานกับบริษัทต่าง ๆ ในการทำสัญญาประมูล ซึ่งประโยชน์ที่ใช้ Dirichlet model มี 2 ประการ คือ ประการแรก ผลที่ได้เกิดจากการเฟ้นสุ่ม และประการที่สอง ไม่ต้องมีการระบุความน่าจะเป็นต่อเนื่องเหมือนในกรณีของ Gate model โดยทั่วไป
Pammer, Fong และ Arnold (2000)	ใช้วิธีของเบย์ในการพยากรณ์การเผยแพร่สินค้าใหม่ออกสู่ตลาด	พบว่า วิธีของเบย์สามารถใช้ได้ดี และยังสามารถนำไปใช้วิเคราะห์กับสถานการณ์อื่นได้หลากหลาย และเชื่อมั่นได้

ตาราง 10 (ต่อ) สรุปงานวิจัยที่เกี่ยวข้องกับการใช้วิธีของเบย์ (Bayesian method)

นักวิจัย	วัตถุประสงค์	ผลการวิจัย
Settimit และ Smith (2000)	เพื่อเปรียบเทียบประสิทธิภาพของการใช้ Linear Bayes method, Algebraic approximate method และ Gibbs sampler method สำหรับ non-Gaussian dynamic processes โดยใช้เทคนิค Markov-chain Monte Carlo โดยที่อนุกรมเวลาที่มีลักษณะไม่เป็นเส้นตรง	พบว่า Linear Bayes method ให้ผลไม่ดีเท่ากับ Algebraic approximate method และ Gibbs sampler method เมื่อโมเดลมีลักษณะเหมาะสมกับข้อมูล
de Alba และ Mendoza (2001)	เพื่อเปรียบเทียบวิธีของเบย์กับวิธีบ็อกซ์-เจนกิ้นส์	พบว่า ในกรณีที่ข้อมูลอนุกรมเวลามีขนาดเล็ก วิธีของเบย์มีค่า MAPE ต่ำกว่าวิธีบ็อกซ์-เจนกิ้นส์ สำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลและข้อมูลอนุกรมเวลาที่มีแนวโน้ม

ตาราง 11 สรุปงานวิจัยที่เกี่ยวข้องกับการใช้วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J)

นักวิจัย	วัตถุประสงค์	วิธีพยากรณ์ที่ใช้	ผลการวิจัย
วันพร เหลืองอาภาพงษ์ (2520)	เพื่อพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการส่งออกข้าว ยาง และข้าวโพด	วิธี B-J	พบว่า อนุกรมเวลาปริมาณข้าว และยางเป็นอนุกรมเวลาที่คงที่รูปแบบที่เหมาะสมกับข้อมูล คือ ARIMA (2,0,0) ส่วนอนุกรมเวลาปริมาณส่งออกข้าวโพดเป็นอนุกรมเวลาที่มีฤดูกาลรูปแบบที่เหมาะสมกับข้อมูลคือ ARIMA (0,1,0) × SARIMA $(0,1,1)_{12}$ และผลการพยากรณ์ปริมาณส่งออกทั้งสามชนิด ได้ค่าพยากรณ์อยู่ในเกณฑ์ไม่แตกต่างจากปีที่ผ่านมาเท่าใดนัก
บุษบา พิกุลผล (2522)	เพื่อเปรียบเทียบรูปแบบที่ใช้คาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทย	เทคนิคการวิเคราะห์อนุกรมเวลา วิธี B-J	พบว่า รูปแบบที่เหมาะสมที่สุด คือรูปแบบที่อธิบายการเปลี่ยนแปลงเนื่องจากวัฏจักรและเหตุการณ์ผิดปกติร่วมกันโดยใช้วิธีบ็อกซ์-เจนกิ้นส์

ตาราง 11 (ต่อ) สรุปงานวิจัยที่เกี่ยวข้องกับการใช้วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J)

นักวิจัย	วัตถุประสงค์	วิธีพยากรณ์ที่ใช้	ผลการวิจัย
วัลลภ โรจนศิริวณิชย์ (2528)	วิเคราะห์ข้อมูลราคาในอดีต โดยใช้การวิเคราะห์อนุกรมเวลาเพื่อหาว่าหลักทรัพย์ใดควรใช้เทคนิคการพยากรณ์ใด และใช้ข้อมูลย้อนหลังเท่าไรจึงจะเหมาะสม โดยเปรียบเทียบระหว่างอนุกรมบ็อกซ์-เจนกิ้นส์ การเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง และเทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง	วิธี B-J, การเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง และเทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง	พบว่า วิธีบ็อกซ์-เจนกิ้นส์ ใช้ในการพยากรณ์ได้ดีกว่าการเฉลี่ยเคลื่อนที่ซ้ำสองครั้ง และเทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้ง ยกเว้นข้อมูลราคารายสัปดาห์ของธนาคารกรุงเทพ การเฉลี่ยเคลื่อนที่ซ้ำสองครั้งใช้ได้ดีที่สุด และบริษัทมาบุญครองอบพืชและไซโล เทคนิคการทำให้เรียบแบบเอ็กซ์โปเนนเชียลซ้ำสองครั้งใช้ได้ดีที่สุด
เกศิณี กมลรัตน์ (2529)	เพื่อเปรียบเทียบวิธีพยากรณ์ที่เหมาะสมกับลักษณะของข้อมูลเงินอากรขาเข้า	วิธี SES, วิธี DES, วิธีการกรองแบบปรับได้, การวิเคราะห์อนุกรมเวลาแบบคลาสสิก และ วิธี B-J	พบว่า ข้อมูลที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลมาก มีค่าเปลี่ยนแปลงน้อยหรือมาก และมีค่าผิดปกติหรือไม่ก็ตาม เมื่อขนาดตัวอย่างเป็น 50, 60, 70, 80, 90, 100, 110 และ 120 วิธีบ็อกซ์-เจนกิ้นส์ใช้พยากรณ์ได้ดีที่สุด
ธิดารัตน์ จันทวี (2539)	เพื่อพยากรณ์ความต้องการใช้ไฟฟ้าเพื่อการผลิตไฟฟ้าระยะสั้น	วิธีการปรับให้เรียบ และ วิธี B-J	พบว่า วิธีการพยากรณ์ที่มีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) ต่ำที่สุดในทุกกรณี คือ วิธีบ็อกซ์-เจนกิ้นส์
บำเพ็ญ ปิดชิด (2540)	เพื่อพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล	วิธี REG, วิธี SMA, วิธี SES และ วิธี B-J	พบว่า การพยากรณ์ข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล วิธีบ็อกซ์-เจนกิ้นส์ มีขนาดความคลาดเคลื่อนน้อยที่สุด
เอกภพ ยานะวิมุติ (2543)	เพื่อเปรียบเทียบความคลาดเคลื่อนระหว่างวิธีบ็อกซ์-เจนกิ้นส์ที่ใช้โมเดลสมการโครงสร้าง กับวิธีบ็อกซ์-เจนกิ้นส์ที่ใช้ตัวบ่งชี้นำในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่ไม่คงที่	วิธี B-J, วิธี B-J ที่ใช้โมเดลเชิงโครงสร้าง และวิธี B-J ที่ใช้ตัวบ่งชี้นำ	พบว่า ผลการพยากรณ์ข้อมูลอนุกรมเวลาปริมาณการยืมหนังสือทั่วไป และหนังสือสำรองด้วยวิธีบ็อกซ์-เจนกิ้นส์ มีขนาดความคลาดเคลื่อนน้อยที่สุด

ตาราง 11 (ต่อ) สรุปงานวิจัยที่เกี่ยวข้องกับการใช้วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J)

นักวิจัย	วัตถุประสงค์	วิธีพยากรณ์ที่ใช้	ผลการวิจัย
อรุณี หงษ์ศิริวัฒน์ (2543)	เพื่อเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์จากวิธีบ็อกซ์และเจนกิ้นส์ วิธีของบ็อกซ์และเจนกิ้นส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย ในการพยากรณ์อนุกรมเวลาทางการศึกษาที่คงที่และไม่คงที่	วิธี B-J, วิธี B-J ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และวิธีการวิเคราะห์การถดถอย	พบว่า วิธีการพยากรณ์ที่เหมาะสมสำหรับข้อมูลอนุกรมเวลาจำนวนผู้เข้าใช้บริการห้องสมุดคือวิธีบ็อกซ์และเจนกิ้นส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และสำหรับข้อมูลอนุกรมเวลาปริมาณการยืมหนังสือระหว่างห้องสมุด คือวิธีบ็อกซ์และเจนกิ้นส์ และวิธีบ็อกซ์และเจนกิ้นส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน

ตาราง 12 สรุปงานวิจัยที่เกี่ยวข้องกับวิธีการพยากรณ์ร่วม (combined forecasting)

นักวิจัย	วัตถุประสงค์	วิธีพยากรณ์ที่เดี่ยวที่นำมารวม	ผลการวิจัย
อมรรัตน์ ปรารมย์ (2539)	เพื่อเปรียบเทียบวิธีการพยากรณ์สำหรับข้อมูลอนุกรมเวลา ระหว่างการเลือกใช้วิธีการพยากรณ์แบบใดแบบหนึ่ง กับการใช้ค่าพยากรณ์ร่วม	วิธี SES, DES, LES และการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลแบบอัตราส่วน	พบว่า ข้อมูลอนุกรมเวลาที่มีขนาดกลาง (n=30) และข้อมูลขนาดใหญ่ (n=50, 70) มีค่าความคลาดเคลื่อนในการพยากรณ์ต่ำที่สุด
ทัศนีย์ อินทนู (2543)	เพื่อเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีพยากรณ์เดี่ยว และวิธีพยากรณ์ร่วม โดยใช้ข้อมูลอนุกรมเวลาทางการศึกษาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล	วิธี SMA, SES, DES, LES, TES, HWS และ REG	พบว่า วิธีการพยากรณ์ที่เหมาะสมกับข้อมูลที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ ปริมาณการยืมหนังสือภาษาไทย และ วิทยานิพนธ์ คือ วิธีการพยากรณ์ร่วม ที่ได้จากการรวมของวิธี HWS กับวิธี REG ซึ่งเป็นวิธีที่เหมาะสมกับข้อมูลทั้ง 2 ชุด โดยให้น้ำหนักแบบที่ 2 แบบที่ 1 ตามลำดับ สำหรับข้อมูลที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ จำนวนครูวิทยาศาสตร์ และจำนวนครูคณิตศาสตร์ คือ วิธีพยากรณ์ร่วม ที่ได้จากการนำวิธีการพยากรณ์ทุกวิธีมารวมกัน โดยให้น้ำหนักแบบที่ 1 และแบบที่ 2 ตามลำดับ

ตาราง 12 (ต่อ) สรุปงานวิจัยที่เกี่ยวข้องกับวิธีการพยากรณ์ร่วม (combined forecasting)

นักวิจัย	วัตถุประสงค์	วิธีพยากรณ์ที่เดี่ยวที่นำมารวม	ผลการวิจัย
Winkler และ Makridakis (1983)	เพื่อศึกษาวิธีพยากรณ์ร่วมโดยใช้น้ำหนักเฉลี่ยของ Newbold และ Granger (1974) 5 แบบ	วิธี naive, SMA, SES, LES, TES, REG, HWS, ARRES และAutomatic AEP	พบว่า วิธีการพยากรณ์ร่วมโดยการให้น้ำหนักเฉลี่ยของ Newbold และ Granger แบบที่ 1 และ 3 เป็นวิธีที่ดีที่สุดในการพยากรณ์ โดยมีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) น้อยที่สุด
Makridakis และ Winkler (1983)	เพื่อตรวจสอบความถูกต้องของการพยากรณ์ร่วม	วิธี naive, SMA, SES, ARRES, DES, LES, TES, REG, HWS, Automatic AEP, B-J, ewandowski FORSYS system, Parzer's ARIMA methodology และ Bayesian forecasting	พบว่า ความถูกต้องของการพยากรณ์ร่วมขึ้นอยู่กับจำนวนวิธีที่นำมารวมกัน โดยยิ่งใช้จำนวนวิธีการพยากรณ์ในการนำมารวมกันเพิ่มมากขึ้นเท่าใด ค่าความถูกต้องก็ยิ่งเพิ่มตาม