บทที่ 2

วรรณคดีที่เกี่ยวข้อง

ในการวิจัยเรื่องการประยุกต์ใช้วิธีวิเคราะห์อนุกรมเวลาบ็อกซ์และเจ็นกิ้นส์ เพื่อพยากรณ์ข้อมูลอนุกรมเวลาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลนั้น ผู้วิจัยได้ศึกษาค้นคว้าเอกสารและงานวิจัยที่เกี่ยวข้อง และนำเสนอวรรณคดีที่เกี่ยวข้องในบทนี้ โดยแยกนำเสนอเป็น 4 ตอน คือตอนที่ 1 เป็นความรู้ทั่วไปของการวิเคราะห์ข้อมูลอนุกรมเวลา ตอนที่ 2 วิธีการของบ็อกซ์และเจ็นกิ้นส์ ตอนที่ 3 วิธีการพยากรณ์ที่ใช้เปรียบเทียบกับวิธีบ็อกซ์และเจ็นกิ้นส์ และตอนที่ 4 เป็นงานวิจัยที่เกี่ยวข้องกับการพยากรณ์ทางการศึกษาและการพยากรณ์โดยใช้วิธีบ็อกซ์และเจ็นกิ้นส์ ในแต่ละตอนมีรายละเอียดดังต่อไปนี้

ตอนที่ 1 การวิเคราะห์ข้อมูลอนุกรมเวลา

สาระในตอนนี้ ผู้วิจัยได้นำเสนอเกี่ยวกับลักษณะและส่วนประกอบของอนุกรมเวลา ขั้นตอนการใช้อนุกรมเวลาในการพยากรณ์ และวิธีการวิเคราะห์อนุกรมเวลาแบบต่าง ๆ ในส่วนของวิธีการวิเคราะห์นั้น ผู้วิจัยนำเสนอในภาพรวม ส่วนรายละเอียดจะนำเสนอในตอนที่ 2 และตอนที่ 3 ต่อไป

ข้อมูลอนุกรมเวลา หมายถึง ชุดของค่าสังเกตที่แทนปรากฏการณ์ที่นักวิจัยต้องการศึกษาในช่วงระยะเวลาหนึ่ง ๆ ตัวอย่างเช่น ข้อมูลของอนุกรมเวลาจำนวนนักเรียนระดับประถมศึกษาในแต่ละปี ตั้งแต่ปีการศึกษา 2531-2536 ประกอบด้วยค่าสังเกตรวม 6 ค่า ค่าสังเกตแต่ละค่ามีความแตกต่างกันตามสภาพการจัดการศึกษาระดับประถมศึกษา ถ้ามีนักเรียนเพิ่มขึ้นในอัตราคงที่ทุกปี ข้อมูลอนุกรมเวลาชุดนี้จะมีแนวโน้มเป็นเส้นตรง โดยทั่วไปนักสถิติแสดงแนวโน้มของข้อมูลอนุกรมเวลาในรูปแผนภูมิเส้น (line graph) และแสดงแนวโน้มของข้อมูลอนุกรมเวลาจำนวนนักเรียนได้ดังภาพ 1ก จะเห็นว่าค่าสังเกตจำนวนนักเรียนในข้อมูลอนุกรมเวลาชุดนี้มีการเปลี่ยนแปลง หรือเคลื่อนไหวไปตามระยะเวลาที่เปลี่ยนแปลง โดยการเปลี่ยนแปลงมีความคงที่ทุกปี

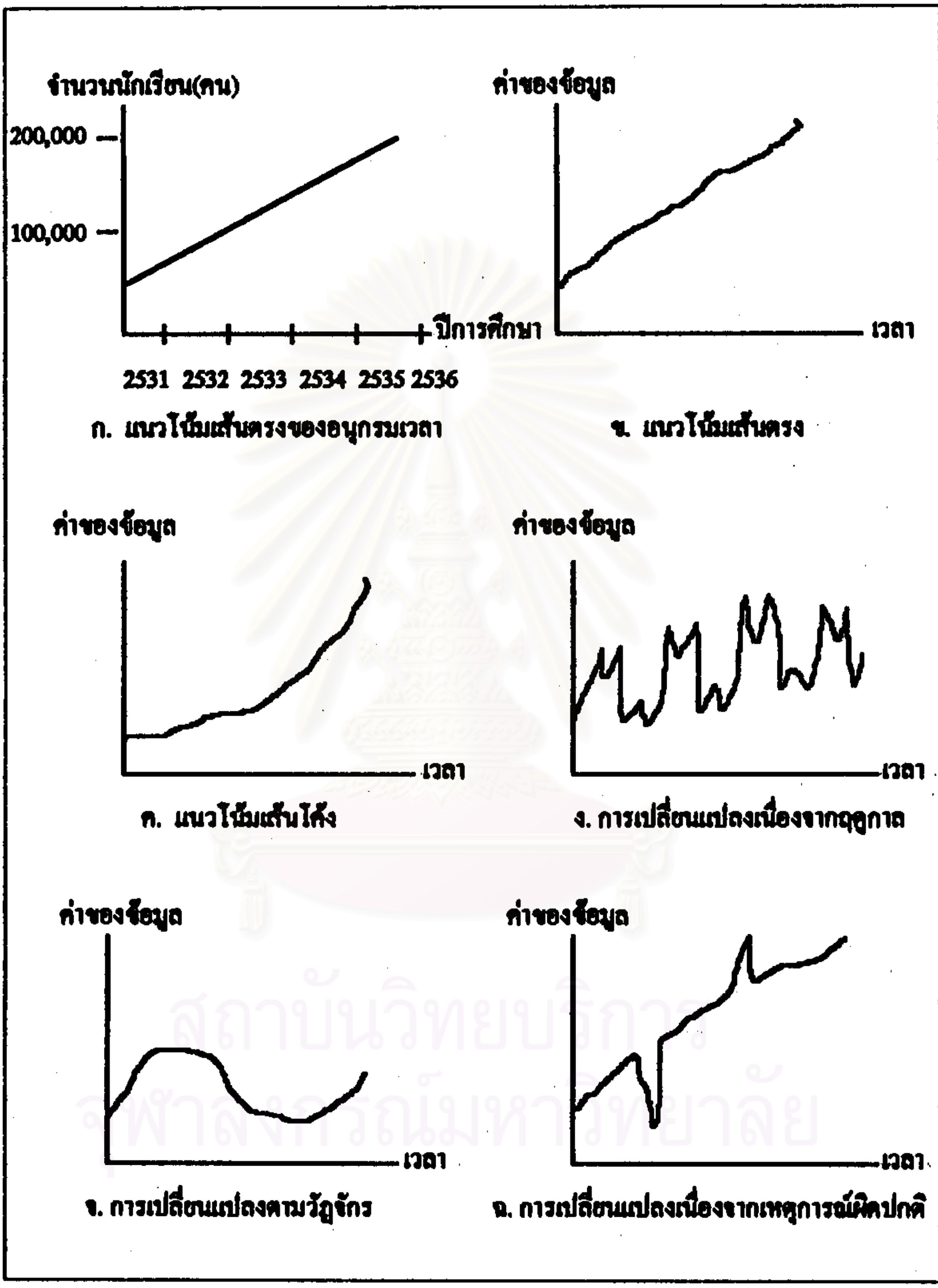
การเปลี่ยนแปลงหรือการเคลื่อนไหวของค่าสังเกตในข้อมูลอนุกรมเวลานั้น เกิดขึ้นได้เนื่องจากปัจจัยหลายประเภท นักสถิติแยกประเภทของลักษณะการเปลี่ยนแปลง ตามลักษณะปัจจัยได้เป็น 4 แบบ คือ การเปลี่ยนแปลงที่แสดงแนวโน้ม การเปลี่ยนแปลงเนื่องจากฤดูกาล การเปลี่ยนแปลงตามวัฏจักร และการเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ (ทรงศิริ แต้สมบัติ, 2539 ; เอกชัย ชัยประเสริฐสิทธิ์, 2527 ; Montgomery, Johnson and Gardiner, 1990) ดังต่อไปนี้

1. การเปลี่ยนแปลงที่แสดงแนวโน้ม (secular trend หรือ T) เป็นการเคลื่อนไหวในระยะเวลาที่ค่อนข้างยาวนาน โดยปกติค่า T แสดงถึงทิศทางที่อนุกรมเวลาชุดนั้น ๆ เปลี่ยนแปลง ค่า T อาจมีลักษณะเป็นเส้นตรง เส้นโค้ง หรือลักษณะอื่นใดก็ได้ ดังตัวอย่างในภาพ 1ข และ 1ค

2. การเปลี่ยนแปลงเนื่องจากฤดูกาล (seasonal variation หรือ S) หรือการเคลื่อนไหวเนื่องจากฤดูกาล (seasonal movement) ข้อมูลมักจะเปลี่ยนแปลงขึ้น ๆ ลง ๆ ซ้ำแล้วซ้ำอีกในช่วงเวลาหนึ่ง ในการพิจารณาศึกษาความเคลื่อนไหวเนื่องจากฤดูกาลนี้ หน่วยของระยะเวลาอาจเป็น รายไตรมาส รายปี รายเดือน รายสัปดาห์ รายวัน หรือรายชั่วโมงก็ได้ดังตัวอย่างในภาพที่ 1ง

3. การเปลี่ยนแปลงตามวัฏจักร (cyclical variation หรือ C) การเคลื่อนไหวเปลี่ยนแปลงของข้อมูลแสดงอิทธิพลของวัฏจักร มีลักษณะทำนองเดียวกับการเคลื่อนไหวเนื่องจากฤดูกาล วัฏจักรหนึ่งจะครอบคลุมระยะเวลาหลายปี แต่ละช่วงมีการเคลื่อนไหวไม่แตกต่างกันมากนัก เช่น วัฏจักรทางธุรกิจประกอบด้วยระยะเวลาหนึ่งที่รุ่งเรือง (prosperity) ติดตามด้วยเศรษฐกิจฝืดเคือง (recession) ตกต่ำ (depression) และฟื้นตัว (recovery) แล้วกลับมาสู่ระยะเวลาที่เศรษฐกิจรุ่งเรืองอีกครั้งหนึ่ง ดังตัวอย่างในภาพที่ 1จ

4. การเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ (irregular variation หรือ I) เป็นปัจจัยที่นักวิจัยไม่อาจคาดการณ์ได้ล่วงหน้า มักจะเกิดขึ้นตามโอกาสหรือโดยบังเอิญ การเปลี่ยนแปลงนี้เกิดขึ้นอย่างสุ่ม (random fluctuation) ดังตัวอย่างในภาพที่ 1ฉ



ก. แนวโน้มเส้นตรงของอนุกรมเวลา ข. แนวโน้มเส้นตรง

ค. แนวโน้มเส้นโค้ง ง. การเปลี่ยนแปลงเนื่องจากฤดูกาล

จ. การเปลี่ยนแปลงตามวัฏจักร ฉ. การเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ

ภาพ 1 ประเภทของลักษณะการเปลี่ยนแปลงแบบต่างๆ ในข้อมูลอนุกรมเวลา

โดยทั่วไปลักษณะของโมเดลอนุกรมเวลาที่พบจะประกอบด้วย การเปลี่ยนแปลงที่แสดงแนวโน้ม การเปลี่ยนแปลงเนื่องจากฤดูกาล การเปลี่ยนแปลงตามวัฏจักร และ การเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ ลักษณะการรวมกันของโมเดลมี 2 แบบ คือ โมเดลที่มีการรวมกันแบบบวก (additive component model) ดังสมการ

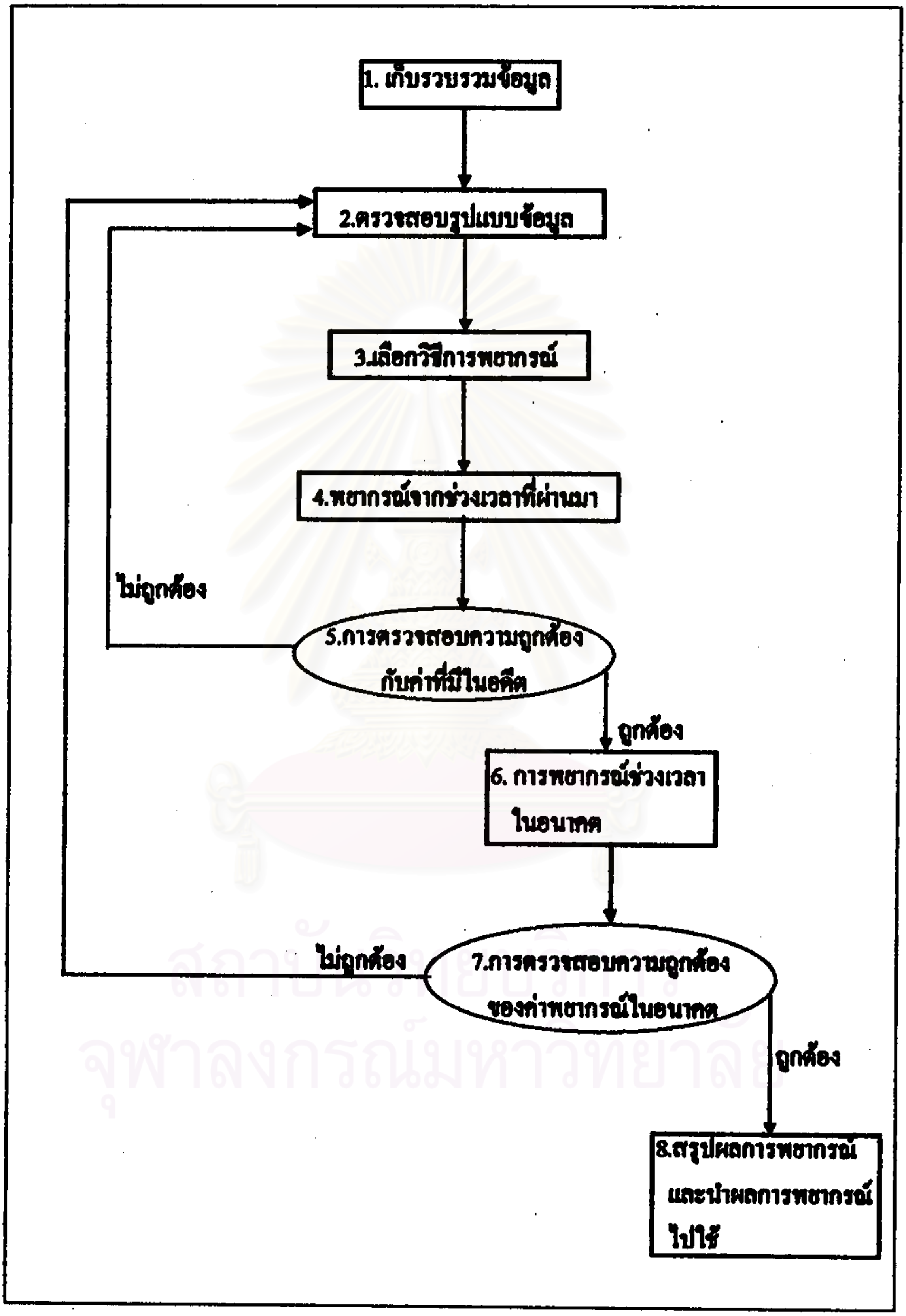
$$Z_t = T_t + S_t + C_t + I_t$$

และโมเดลที่มีการรวมกันแบบคูณ (multiplicative component model) ดังสมการ

$$Z_t = T_t \cdot S_t \cdot C_t \cdot I_t$$

การวิเคราะห์อนุกรมเวลา เป็นการศึกษาถึงการเปลี่ยนแปลงหรือการเคลื่อนไหวของข้อมูลอนุกรมเวลาชุดหนึ่ง ตามช่วงเวลาว่ามีแบบแผนการเคลื่อนไหวสอดคล้องกับโมเดลอนุกรมเวลาชนิดใด เพื่อนำโมเดลอนุกรมเวลาที่ได้ไปใช้ประโยชน์ในการพยากรณ์ข้อมูลในอนาคตต่อไป (เอกชัย ชัยประเสริฐสิทธิ์, 2527 ; Makridakis, Wheelwright and McGee, 1978)

วิธีการวิเคราะห์อนุกรมเวลาเพื่อใช้พยากรณ์ข้อมูลโดยทั่วไป มีขั้นตอนสรุปได้เป็น 8 ขั้นตอน (Hanke and Reitsch, 1992) ขั้นตอนแรก เป็นการเก็บรวบรวมข้อมูล ขั้นตอนที่ 2 เมื่อเก็บรวบรวมข้อมูลหรือค่าสังเกตที่ผ่านมาในอดีตจากช่วงเวลา t_1 ถึงเวลา t_n ได้แล้ว นำข้อมูลที่ได้มาตรวจสอบรูปแบบของข้อมูล ขั้นตอนที่ 3 เป็นการเลือกวิธีพยากรณ์ที่เหมาะสมกับลักษณะของข้อมูล ขั้นตอนที่ 4 เป็นการพยากรณ์ข้อมูลจากช่วงเวลาที่ผ่านมา โดยใช้ข้อมูลจากช่วงเวลา t_1-t_k เป็นฐานพยากรณ์จากช่วงเวลา t_{k+1}..... t_n ขั้นตอนที่ 5 เป็นการตรวจสอบผลการพยากรณ์จากข้อมูลจริงในช่วงเวลา t_{k+1} ..... t_n ถ้านักวิจัยเห็นว่าผลการพยากรณ์ไม่ถูกต้องจะต้องนำข้อมูลไปตรวจสอบโมเดล และเลือกวิธีพยากรณ์ใหม่อีกครั้งหนึ่งตามขั้นตอนที่ 1-5 แต่ถ้านักวิจัยยอมรับผลการพยากรณ์ว่าถูกต้อง ก็ดำเนินการต่อไปตามขั้นตอนที่ 6 คือการพยากรณ์ช่วงเวลาในอนาคต t_{n+1} , t_{n+2}.... ขั้นตอนที่ 7 เป็นการตรวจสอบผลการพยากรณ์ ถ้านักวิจัยเห็นว่าผลการพยากรณ์ไม่ถูกต้อง จะต้องนำข้อมูลไปตรวจสอบโมเดลและเลือกวิธีพยากรณ์ใหม่ตามขั้นตอนที่ 1-6 อีกครั้งหนึ่ง แต่ถ้านักวิจัยยอมรับผลการพยากรณ์ว่าถูกต้อง จึงสรุปผลการพยากรณ์และนำผลการพยากรณ์ไปใช้ในการตัดสินใจ ขั้นตอนทั้ง 8 ขั้นตอนในการพยากรณ์ แสดงไว้ในภาพ 2



ภาพ 2 ขั้นตอนการพยากรณ์ข้อมูลอนุกรมเวลาโดยทั่วไป

วิธีวิเคราะห์อนุกรมเวลาโดยทั่วไปแบ่งได้เป็น 4 วิธีใหญ่ ๆ คือวิธีการแยกส่วนประกอบ (decomposition method) วิธีการเฉลี่ยเคลื่อนที่ (moving averages) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) และ วิธีบ็อกซ์และเจ็นกินส์ (Box-Jenkins) (Newbold and Bos, 1994 ; Sullivan and Claycombe, 1977 ; Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; ทรงศิริ แต้สมบัติ, 2539) วิธีการวิเคราะห์อนุกรมเวลาทั้ง 4 วิธี อาจใช้ผสมผสานกันได้ เช่น ใช้การวิเคราะห์อนุกรมเวลาโดยการแยกส่วนประกอบแล้วนำองค์ประกอบที่ได้ไปปรับให้เรียบแบบเอ็กซ์โปเนนเชียล เป็นต้น วิธีการวิเคราะห์อนุกรมเวลาแต่ละวิธีมีหลักการดังนี้

1. วิธีการแยกส่วนประกอบ (decomposition method) เป็นการวิเคราะห์อนุกรมเวลาโดยการระบุองค์ประกอบ 4 องค์ประกอบ ที่มีอิทธิพลในแต่ละช่วงเวลาของอนุกรม คือ การเปลี่ยนแปลงที่แสดงแนวโน้ม(trend) การเปลี่ยนแปลงตามวัฏจักร (cyclical) การเปลี่ยนแปลงเนื่องจากฤดูกาล (seasonal) และการเปลี่ยนแปลงเนื่องจากเหตุการณ์ผิดปกติ (irregular) แล้วนำองค์ประกอบที่ได้ไปใช้พยากรณ์ ค่าพยากรณ์ในอนาคตจะได้จากการรวมค่าการวัดส่วนประกอบของอนุกรมเวลา วิธีการแยกส่วนประกอบใช้พยากรณ์ได้ทั้งระยะสั้นไม่เกิน 5 ช่วงเวลา และระยะยาวตั้งแต่ 10 ช่วงเวลาขึ้นไป

2. วิธีการเฉลี่ยเคลื่อนที่ (moving averages) เป็นวิธีการวิเคราะห์อนุกรมเวลาโดยการขจัดปัจจัยที่ทำให้เกิดความเคลื่อนไหวของข้อมูลอนุกรมเวลาออกอย่างสุ่ม แล้วนำโมเดลที่ได้ไปใช้พยากรณ์ ค่าพยากรณ์ในอนาคตเป็นค่าที่ได้จากค่าสังเกตในอดีต โดยเฉลี่ยให้น้ำหนักกับค่าสังเกตต่าง ๆ เท่ากันแบบเลขคณิต วิธีการเฉลี่ยเคลื่อนที่ใช้พยากรณ์ระยะสั้น

3. วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (exponential smoothing) เป็นวิธีการวิเคราะห์อนุกรมเวลา โดยการขจัดปัจจัยที่ทำให้เกิดความเคลื่อนไหวของข้อมูลอนุกรมเวลาออกอย่างสุ่ม แล้วนำโมเดลที่ได้ไปใช้พยากรณ์ ค่าพยากรณ์ในอนาคตเป็นค่าที่ได้จากค่าสังเกตในอดีต โดยให้น้ำหนักค่าสังเกตลดหลั่นกันแบบเอ็กซ์โปเนนเชียล คือการเฉลี่ยน้ำหนักของค่าสังเกตจะให้น้ำหนักของค่าสังเกตที่เวลาปัจจุบันมาก และเมื่อช่วงเวลาห่างจากปัจจุบันมากขึ้นจะให้น้ำหนักค่าสังเกตลดลงเรื่อย ๆ แบบเรขาคณิต วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลใช้พยากรณ์ระยะสั้น

วิธีวิเคราะห์อนุกรมเวลาที่กล่าวข้างต้นจะใช้พยากรณ์ข้อมูลในอนาคตได้ ถ้านักวิจัยมีการกำหนดโมเดลความสัมพันธ์ หรือโมเดลการเคลื่อนไหวเปลี่ยนแปลงของข้อมูลในอดีตที่จะใช้ในการพยากรณ์ ซึ่งผู้วิจัยจำเป็นต้องตรวจสอบโมเดลการเคลื่อนไหวเปลี่ยนแปลงของข้อมูลในอดีตก่อนที่จะทำการวิเคราะห์ข้อมูล แต่ในบางกรณี นักวิจัยไม่สามารถกำหนดโมเดลความสัมพันธ์ดังกล่าวได้ จึงมีผู้เสนอวิธีการวิเคราะห์อนุกรมเวลาที่ไม่มีการกำหนดโมเดลความสัมพันธ์ก่อนที่จะทำการวิเคราะห์ แต่โมเดลจะถูกเลือกขึ้นมาในขั้นตอนต่าง ๆ ของการวิเคราะห์ เรียกว่า วิธีบ๊อกซ์และเจ็นกิ้นส์ ดังนั้นวิธีการนี้ผู้วิจัยไม่จำเป็นต้องตรวจสอบโมเดลการเคลื่อนไหวเปลี่ยนแปลงของข้อมูลในอดีตก่อนที่จะทำการวิเคราะห์ข้อมูล

4. วิธีบ๊อกซ์และเจ็นกิ้นส์ (Box-Jenkins) เป็นวิธีการวิเคราะห์อนุกรมเวลาซึ่งใช้พยากรณ์ข้อมูลโดยไม่มีการกำหนดโมเดลความสัมพันธ์ก่อนที่จะทำการวิเคราะห์ แต่มีข้อตกลงเบื้องต้นของโมเดลว่า อนุกรมเวลาที่ใช้ในการวิเคราะห์ต้องเป็นอนุกรมคงที่ (stationary) ถ้าอนุกรมที่ใช้วิเคราะห์เป็นอนุกรมไม่คงที่ (nonstationary) ผู้วิจัยต้องเปลี่ยนให้อยู่ในรูปอนุกรมคงที่โดยการหาผลต่างและ/หรือผลต่างฤดูกาล นอกจากนี้อนุกรมเวลาที่ใช้ในการวิเคราะห์ต้องมีจำนวนค่าสังเกตอย่างน้อย 50 ค่า หลักการพยากรณ์ใช้วิธีการคำนวณทวนซ้ำ (iterative) เพื่อให้ได้ค่าพยากรณ์ที่สอดคล้องกับแบบแผนการเคลื่อนไหวของข้อมูลอนุกรมเวลามากที่สุด วิธีบ๊อกซ์และเจ็นกิ้นส์ใช้พยากรณ์ในระยะเวลาสั้นๆ

วิธีวิเคราะห์อนุกรมเวลาทั้ง 4 วิธีมีประโยชน์ในการใช้งาน ลักษณะข้อมูลที่ใช้ในการพยากรณ์ และความสะดวกในการมีโปรแกรมสำเร็จรูปที่ใช้วิเคราะห์แตกต่างกัน สรุปได้ดังตาราง 1

จากตาราง 1 เมื่อเปรียบเทียบวิธีการวิเคราะห์อนุกรมเวลาทั้ง 4 วิธี สรุปได้ว่าวิธีบ๊อกซ์และเจ็นกิ้นส์ เป็นวิธีที่ใช้ในการพยากรณ์ได้กว้างขวางกว่าวิธีอื่น ๆ และให้ผลการวิเคราะห์ที่มีการตรวจสอบความถูกต้องได้ชัดเจน ไม่ว่าแบบแผนการเคลื่อนไหวของข้อมูลอนุกรมเวลาจะเป็นอย่างไร ผู้วิจัยจึงสนใจใช้วิธีบ๊อกซ์และเจ็นกิ้นส์ในการพยากรณ์ทางการศึกษา และได้เสนอรายละเอียดของวิธีการวิเคราะห์ของบ๊อกซ์และเจ็นกิ้นส์ในตอนต่อไป

ตาราง 1 การเปรียบเทียบวิธีวิเคราะห์อนุกรมเวลา 4 วิธี

ลักษณะ	วิธีวิเคราะห์อนุกรมเวลา			
	การแยกส่วนประกอบ	การเฉลี่ยเคลื่อนที่	การทำให้เรียบแบบเอ็กซ์โปเนนเชียล	บ็อกซ์และเจ็นกิ้นส์
1. ระยะเวลาการพยากรณ์				
- ระยะสั้น	/	/	/	/
- ระยะยาว	/			/
2. ลักษณะของข้อมูล				
- แนวโน้ม	/	/	/	/
- ฤดูกาล	/	/	/	/
- วัฏจักร	/			
3. ขนาดข้อมูลที่ต้องการ	30	10	10	50
4. การตรวจสอบโมเดล	ต้องทำ	ต้องทำ	ต้องทำ	ไม่ต้องทำ
5.ข้อตกลงเบื้องต้นเกี่ยวกับโมเดล	มี	มี	มี	มี
6.การตรวจสอบความสอดคล้องของโมเดลกับข้อมูล	ไม่มี	ไม่มี	ไม่มี	มี
7. มีโปรแกรมสำเร็จรูป	/	/	/	/
8. ค่าใช้จ่ายในการวิเคราะห์				
- ต่ำ		/	/	
- ปานกลาง	/			
- สูง				/

ตอนที่ 2 วิธีการของบ็อกซ์และเจ็นกิ้นส์ (Box-Jenkins)

สาระในตอนนี้จะกล่าวถึงวิธีการวิเคราะห์อนุกรมเวลาบ็อกซ์และเจ็นกิ้นส์ โดยผู้วิจัยแยกนำเสนอเป็น 5 ตอน คือ ตอนที่ 2.1 เป็นการกำหนดโมเดลที่เหมาะสมกับลักษณะของ

ข้อมูล ตอนที่ 2.2 การนำโมเดลที่ได้ไปประมาณค่าพารามิเตอร์ ตอนที่ 2.3 การตรวจสอบความเหมาะสมของโมเดล ตอนที่ 2.4 การนำโมเดลไปใช้พยากรณ์ และตอนที่ 2.5 เป็นการตรวจสอบผลการพยากรณ์

ในการวิเคราะห์ข้อมูลอนุกรมเวลา เพื่อประมาณค่าตัวแปรตามในช่วงเวลาที่ t จากสมการที่เป็นการรวมกันเชิงเส้น (linear combination) ของตัวแปรสุ่มอิสระ เขียนได้ดังสมการ (1)

$$Z_t = \mu + \psi_0 a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots \qquad \text{.................................(1)}$$

เมื่อ ψ_j (j=0, 1, ...) เป็นค่าคงที่ เรียกว่า น้ำหนัก (weight) โดยทั่วไป $\psi_0 = 1$

μ เป็นค่าคงที่ที่กำหนดระดับของกระบวนการ (process)

Z_t เป็นค่าสังเกตที่เวลา t

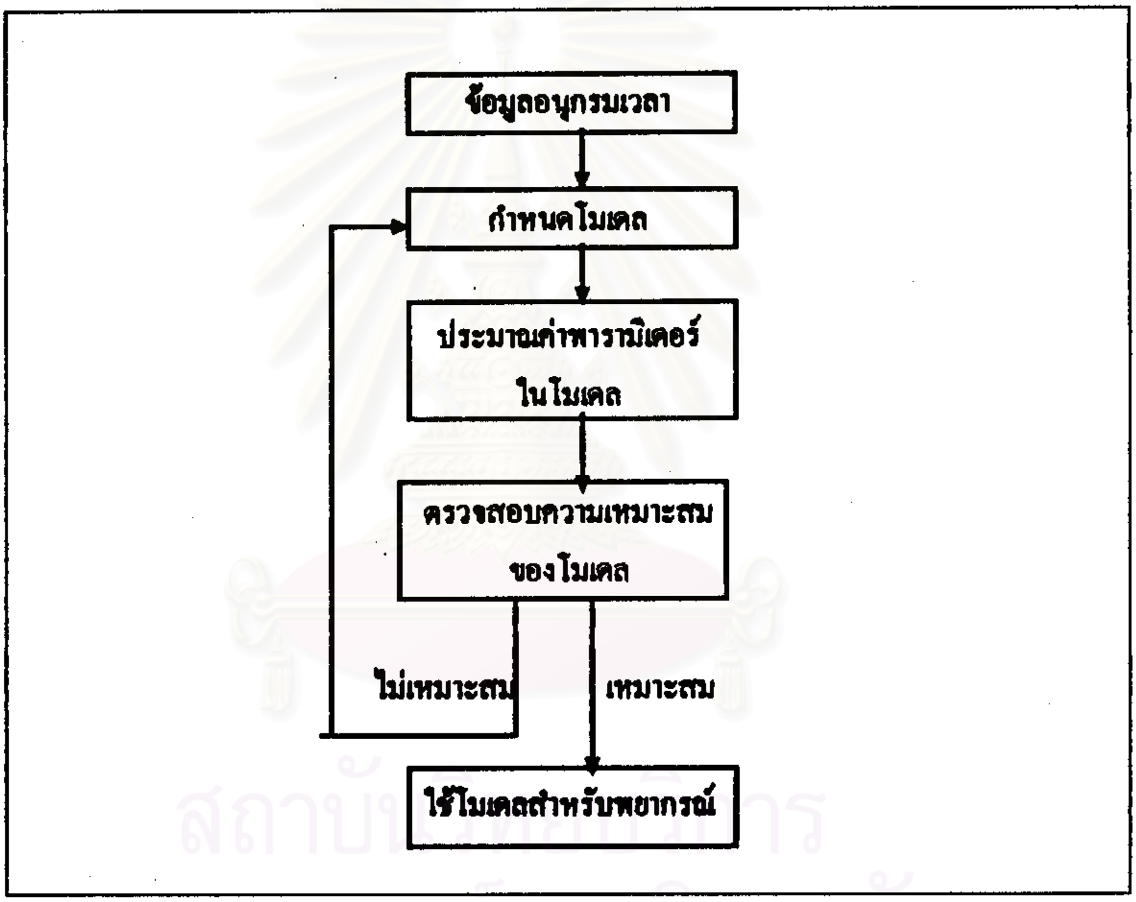
a_t, a_{t-1}, a_{t-2} เป็นตัวแปรสุ่มอย่างอิสระ มีการแจกแจงแบบปกติ ค่าเฉลี่ยเป็น 0 ความแปรปรวนเป็น σ_a^2

ในที่นี้ค่า a_t, a_{t-1}, a_{t-2}.... เป็นเทอมความคลาดเคลื่อนของตัวแปรล้าหลัง (lag variable) และเทอมความคลาดเคลื่อนนี้เป็นตัวแปรสุ่ม ซึ่งในตำราหลายเล่มใช้สัญลักษณ์ e แต่ผู้วิจัยใช้สัญลักษณ์ a ตามต้นฉบับตำราของ Box, Jenkins and Reinsel (1994)

จากสมการ (1) โดยทั่วไปเรียกว่าการกรองเชิงเส้น (linear filter) ซึ่งมี Z_t เป็นตัวแปรตามหรือค่าสังเกตที่กำหนดมาจาก a_t นั่นคือถ้า a_t มีการแจกแจงปกติ Z_t ก็จะแจกแจงปกติด้วย ในโมเดลการกรองเชิงเส้น (linear filter) จะนิยามอนุกรมเวลาตามฟังก์ชันที่เปลี่ยนแปลง ในโมเดลอนุกรมเวลาการกรองเชิงเส้น (linear filter) จะสร้างโมเดลที่แตกต่างกันหลายโมเดล George E.P. Box และGwilym M. Jenkins จึงนำมารวมให้เป็นหน่วยเดียวกัน ดังนั้นโมเดลอนุกรมเวลาที่ได้มาจากการกรองเชิงเส้น (linear filter) โดยทั่วไปจึงเรียกว่า โมเดลบ็อกซ์และเจ็นกิ้นส์ (Box-Jenkins model) (Montgomery, Johnson and Gardiner, 1990)

วิธีบ็อกซ์และเจ็นกิ้นส์เป็นวิธีการพยากรณ์ที่มีความแตกต่างจากวิธีอื่น ๆ เพราะวิธีนี้ไม่มีการกำหนดโมเดลความสัมพันธ์ข้อมูลอนุกรมเวลาก่อนที่จะใช้ในการวิเคราะห์ข้อมูล โดยวิธีนี้ใช้การคำนวณเทวนซ้ำ (iteration) ของการกำหนดโมเดลที่เป็นไปได้จากโมเดลทั่วไป โมเดลจะ

การตรวจสอบข้อมูลจนสามารถอธิบายอนุกรมได้อย่างถูกต้อง ถ้าโมเดลพยากรณ์ข้อมูลในอดีตมีความคลาดเคลื่อนน้อย มีการแจกแจงอย่างสุ่มและเป็นอิสระต่อกัน แสดงว่าโมเดลมีความสอดคล้องดี แต่ถ้าหากโมเดลไม่เหมาะสม กระบวนการนี้ก็จะหาโมเดลใหม่เพื่อปรับปรุงโมเดล ซึ่งกระบวนการนี้จะทำซ้ำ จนกระทั่งได้โมเดลที่มีความเหมาะสม (Hanke and Reitsch, 1992 ; Thomopoulos, 1980; Newbold and Bos, 1994 ; Sullivan and Claycombe, 1977 ; Bowerman and O'Connell, 1993) วิธีการดังกล่าวแสดงได้ดังภาพ 3



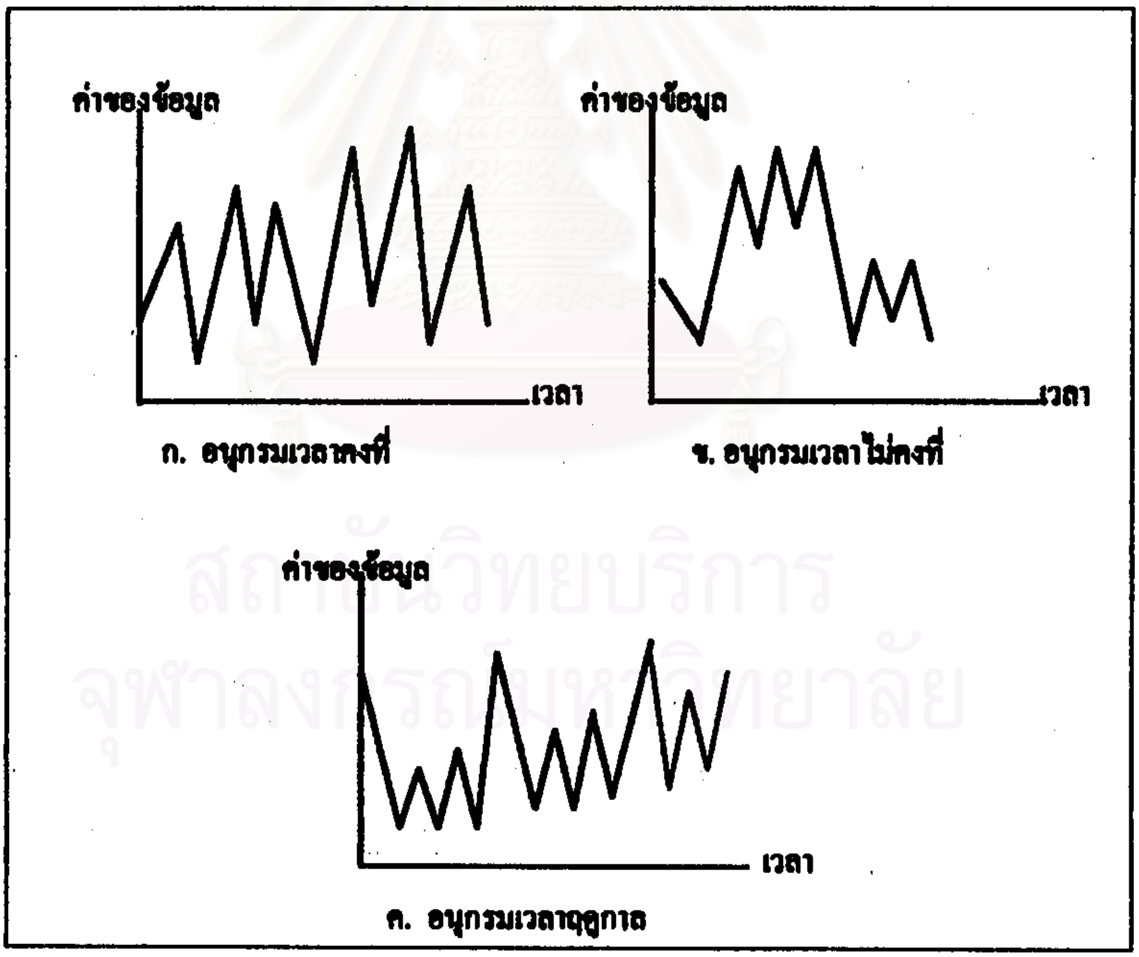
ภาพ 3 วิธีการวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีบ๊อกซ์และเจ็นกิ้นส์

โมเดลจากสมการ (1) เป็นได้ทั้งอนุกรมเวลาที่คงที่และไม่คงที่ ดังนั้นจึงสามารถแบ่งประเภทของอนุกรมเวลา ที่กำหนดโมเดลของบ๊อกซ์และเจ็นกิ้นส์ได้ 3 ประเภท (Newbold and Bos, 1994 ; Sullivan and Claycombe, 1977 ; Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; Makridakis, Wheelwright and McGee, 1978 ; ทรงศิริ แต้สมบัติ, 2539) คือ

1. อนุกรมเวลาคงที่ (stationary time series) หมายถึง ข้อมูลที่เคลื่อนไหวไปรอบ ๆ ค่าเฉลี่ย และการเคลื่อนไหวเป็นไปในลักษณะคงที่ไม่เปลี่ยนแปลงไปตามเวลาที่ผ่านไป เมื่อค่าหนึ่งเพิ่มขึ้นค่าที่ตามมาจะลดลง ดังภาพ 4 ก

2. อนุกรมเวลาไม่คงที่ (nonstationary time series) หมายถึง ข้อมูลที่เคลื่อนไหวไม่แน่นอน เปลี่ยนแปลงไปตามระยะเวลา ดังภาพ 4 ข

3. อนุกรมเวลาฤดูกาล (seasonal time series) หมายถึง ข้อมูลที่มีการเคลื่อนไหวขึ้นลงตามระยะเวลาเป็นช่วงที่แน่นอน และลักษณะการเคลื่อนไหวในระยะเวลาหนึ่งจะคล้าย ๆ กันกับช่วงเวลาอื่น ๆ ซ้ำ ๆ กัน ดังภาพ 4 ค



ภาพ 4 ลักษณะอนุกรมเวลาแบบคงที่ ไม่คงที่ และแบบฤดูกาล

โดยปกติโมเดลอนุกรมเวลาของบ็อกซ์และเจ็นกิ้นส์ ที่ใช้ในการพยากรณ์ต้องเป็นอนุกรมเวลาคงที่ แต่หากอนุกรมเวลาชุดใดเป็นอนุกรมไม่คงที่ ต้องใช้กระบวนการขจัดแนวโน้มเพื่อเปลี่ยน (transformation) อนุกรมเวลานั้นให้อยู่ในรูปของค่าอนุกรมเวลาที่คงที่ก่อน จึงจะนำโมเดลไปใช้ในการพยากรณ์ได้ คือ ถ้ามีข้อมูลอนุกรมเวลาที่เป็นอนุกรมเวลาไม่คงที่ ประกอบด้วยค่าสังเกต n ตัว ได้แก่ อนุกรม $Z_1, Z_2, Z_3, ..., Z_n$ เป็นอนุกรมไม่คงที่ ต้องนำข้อมูลอนุกรมเวลา n ตัวนี้มาเปลี่ยนให้อยู่ในรูปค่าอนุกรมเวลาที่คงที่ โดยการหาผลต่างลำดับที่ 1 ได้ดังนี้

$$W_t = Z_t - Z_{t-1} \quad \text{เมื่อ} \quad t = 2, 3, ..., n$$

นั่นคือ นักวิจัยสามารถเขียนค่าเริ่มต้นและค่าผลต่างลำดับที่ 1 ของอนุกรมเวลาได้ดังนี้

ค่าเริ่มต้น	ค่าผลต่างลำดับที่ 1
Z_1	
Z_2	$W_2 = Z_2 - Z_1$
Z_3	$W_3 = Z_3 - Z_2$
.	.
.	.
.	.
Z_{n-1}	.
Z_n	$W_n = Z_n - Z_{n-1}$

ถ้านักวิจัยหาผลต่างลำดับที่ 1 ได้แล้ว แต่ยังไม่สามารถทำให้ข้อมูลอนุกรมเวลานั้นอยู่ในรูปอนุกรมคงที่ นักวิจัยต้องหาผลต่างลำดับที่ 2, 3, ... จนกว่าจะได้อนุกรมเวลาคงที่ แต่โดยทั่วไป การหาผลต่างมักจะไม่เกินลำดับที่ 2 ก็จะได้อนุกรมคงที่ ผลต่างลำดับที่ 2 หาได้ดังนี้

$$\begin{aligned} W_t &= (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) \\ &= Z_t - 2Z_{t-1} + Z_{t-2} \qquad \text{เมื่อ } t = 3, 4, ..., n \end{aligned}$$

นั่นคือ นักวิจัยสามารถเขียนค่าเริ่มต้นและค่าผลต่างลำดับที่ 2 ของอนุกรมเวลาได้ดังนี้

ค่าเริ่มต้น ค่าผลต่างลำดับที่ 2

Z_1

Z_2

Z_3 $W_3 = Z_3 - 2Z_2 + Z_1$

Z_4 $W_4 = Z_4 - 2Z_3 + Z_2$

. .

. .

. .

Z_n $W_n = Z_n - 2Z_{n-1} + Z_{n-2}$

กำหนดให้สัญลักษณ์ $W_b, W_{b+1}, \ldots W_n$ (เมื่อ $b = 2, 3, \ldots n$) แทน $Z_2, Z_3, \ldots Z_n$ และเรียก $W_b, W_{b+1}, \ldots W_n$ ว่าอนุกรมใหม่ (new series หรือ working series) (Bowerman and O'Connell, 1993)

ขั้นตอนในการวิเคราะห์อนุกรมเวลาของวิธีบ๊อกซ์และเจ็นกิ้นส์มี 4 ขั้นตอน ขั้นตอนแรกเป็น การกำหนดโมเดล (identification) ขั้นตอนที่ 2 การประมาณค่าพารามิเตอร์ในโมเดล (parameter estimation) ขั้นตอนที่ 3 การตรวจสอบความเหมาะสมของโมเดล (diagonostic checking) และขั้นตอนที่ 4 การใช้โมเดลทำการพยากรณ์ (forecasting) ในแต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

ตอนที่ 2.1 การกำหนดโมเดล (Identification)

การพยากรณ์ด้วยโมเดลบ๊อกซ์และเจ็นกิ้นส์ เป็นการพยากรณ์ค่าสังเกตในอนาคตโดยใช้ค่าสังเกตในปัจจุบันและค่าสังเกตในอดีตเป็นฐานในการพยากรณ์ หรือกล่าวอีกนัยหนึ่งว่าเป็นการใช้ชุดของค่าสังเกตในอดีตของตัวแปรตัวเดียวกันมาพยากรณ์ค่าสังเกตในอนาคต เช่น ถ้าให้ t เป็นช่วงเวลาที่จะพยากรณ์ นักวิจัยสามารถพยากรณ์ค่า Z_t ได้จาก Z_{t-1} และพยากรณ์ Z_{t-1} ได้จาก Z_{t-2} หรือสามารถเขียนให้อยู่ในรูปของการรวมกันเชิงเส้นได้ว่า

$$Z_t = a + b_1Z_{t-1} + b_2Z_{t-2} + \ldots + b_kZ_{t-k} + a_t \qquad \text{..........................(2)}$$

จากสมการ (2) จะเห็นได้ว่าตัวแปรตามที่เกิดขึ้นนั้น เกิดจากค่าของตัวแปรตามที่เกิดขึ้นก่อนหรือเรียกว่า ค่าเวลาล้าหลัง (time-lagged) ของตัวแปรตาม ความสัมพันธ์ระหว่าง

ค่าสังเกตในช่วงเวลาหนึ่งกับค่าสังเกตในช่วงเวลาก่อนหน้าเป็นความสัมพันธ์ของค่าสังเกตในตัวแปรตัวเดียวกันและความสัมพันธ์ในลักษณะนี้มีชื่อเรียกว่า อัตตะสหสัมพันธ์ (autocorrelation) ดังนั้นการพยากรณ์ข้อมูลอนุกรมเวลาจึงต้องหาฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function) และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล (partial autocorrelation function) (Makridakis, Wheelwright and McGee, 1978)

ฟังก์ชันอัตตะสหสัมพันธ์ (Autocorrelation Function)

ฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function) ใช้ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ (autocorrelation coefficient) อธิบายความเกี่ยวเนื่องระหว่างค่าของตัวแปรตัวเดียวกันที่อยู่ในช่วงเวลาแตกต่างกัน นักวิจัยจึงสามารถสร้างตัวแปรตัวหนึ่งจากตัวแปรอีกตัวหนึ่งได้ ตัวแปรที่นักวิจัยสร้างขึ้นใหม่นี้ เรียกว่า ตัวแปรเวลาล้าหลัง (time lag variables) เช่น นักวิจัยมีตัวแปร Z_t เป็นตัวแปรเริ่มต้น และสามารถนำข้อมูลในตัวแปร Z_t ไปสร้างเป็นตัวแปรเวลาล้าหลัง Z_{t-1}, Z_{t-2} และ Z_{t-3} ได้ดังตาราง 2

ตาราง 2 การสร้างตัวแปรเวลาล้าหลัง (time lag variables)

ช่วงเวลา	ตัวแปรเริ่มต้น	ตัวแปรเวลาล้าหลังที่ 1	ตัวแปรเวลาล้าหลังที่ 2	ตัวแปรเวลาล้าหลังที่ 3
t	Z_t	Z_{t-1}	Z_{t-2}	Z_{t-3}
1	12	-	-	-
2	15	12	-	-
3	8	15	12	-
4	11	8	15	12
5	4	11	8	15
6	9	4	11	8
7	13	9	4	11
8	12	13	9	4
9	7	12	13	9
10	14	7	12	13

จากตาราง 2 สามารถสร้างตัวแปรเวลาล้าหลัง Z_{t-1}, Z_{t-2} และ Z_{t-3} โดยการย้ายข้อมูลจากช่วงเวลา t_1 ลงมาเป็นช่วงเวลา t_2 จากช่วงเวลา t_2 ลงมาเป็นช่วงเวลา t_3 และจากช่วงเวลา t_3 ลงมาเป็นช่วงเวลา t_4 ตามลำดับ ผลจากการย้ายนี้ทำให้ค่าใน Z_{t-1} หายไป 1 ค่า ค่าใน Z_{t-2} หายไป 2 ค่า และค่าใน Z_{t-3} หายไป 3 ค่า เมื่อนักวิจัยสร้างตัวแปรเวลาล้าหลังได้แล้ว จึงนำมาคำนวนหาค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ระหว่างตัวแปรเริ่มต้นกับตัวแปรเวลาล้าหลัง คือ อัตตะสหสัมพันธ์ระหว่าง Z_t กับ Z_{t-1}, Z_t กับ Z_{t-2} และ Z_t กับ Z_{t-3}

สัมประสิทธิ์อัตตะสหสัมพันธ์ระหว่าง Z_t กับ Z_{t-1} หาได้จาก

$$r_{Z_tZ_{t-1}} = \frac{(\text{Covariance between } Z_t \text{ and } Z_{t-1})}{(\text{S.D.}Z_t)\times(\text{S.D.}Z_{t-1})}$$

$$= \frac{\sum_{t=2}^{n}(Z_t-\overline{Z}_t)(Z_{t-1}-\overline{Z}_{t-1})}{\sqrt{\sum_{t=1}^{n}(Z_t-\overline{Z}_t)^2}\sqrt{\sum_{t=2}^{n}(Z_{t-1}-\overline{Z}_{t-1})^2}} \quad \text{.................(3)}$$

$$= \frac{\sum_{t=2}^{n}(Z_t-\overline{Z})(Z_{t-1}-\overline{Z})}{\sum_{t=1}^{n}(Z_t-\overline{Z})^2} \quad \text{.................(4)}$$

จากสมการ (3) มีตัวห้อยท้าย (subscript) ทั้งตัวเศษและตัวส่วน ทำให้สูตรเกิดความยุ่งยาก จึงมีข้อตกลงเบื้องต้นว่า Z_t เป็นอนุกรมเวลาคงที่ จึงให้ค่าเฉลี่ยของ Z_t และ Z_{t-1} มีค่าเท่ากัน และนำเอาตัวห้อยท้ายออก จะได้ $\overline{Z} = \overline{Z}_t = \overline{Z}_{t-1}$ ส่วนความแปรปรวนของ Z_t และ Z_{t-1} ประมาณค่าได้โดยการใช้ข้อมูล Z_t จากข้อตกลงเบื้องต้นดังกล่าวจึงได้สมการ (4) และโดยทั่วไปอัตตะสหสัมพันธ์ของเวลาล้าหลัง (time lag) ที่ 1, 2, 3, 4, ..., k สามารถเขียนได้ดังสมการ (5) (Newbold and Bos, 1994 ; Sullivan and Clayoombe, 1977 ; Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; Makridakis and Wheelwright, 1989 ; Makridakis, Wheelwright and McGee, 1978 ; ทรงศิริ แต้สมบัติ, 2539)

$$r_k = \frac{\sum_{t=1}^{n-k}(Z_t - \overline{Z})(Z_{t+k} - \overline{Z})}{\sum_{t=1}^{n}(Z_t - \overline{Z})^2} \qquad \text{...............(5)}$$

เมื่อ r_k = สัมประสิทธิ์อัตตะสหสัมพันธ์ที่เวลาล้าหลัง k

t = ช่วงเวลา

n = ช่วงเวลาสุดท้าย

k = เวลาล้าหลัง

Z_t = ตัวแปรเริ่มต้น

$$\overline{Z} = \frac{\sum_{t=1}^{n} Z_t}{n}$$

การวัดความสัมพันธ์เชิงเส้นระหว่างค่าสังเกตของอนุกรมเวลาจะแยกโดยใช้หน่วยของเวลาล้าหลังที่ k ซึ่ง r_k มีค่าอยู่ระหว่าง -1 และ 1 ถ้าค่าสังเกตแยกโดยหน่วยของเวลาล้าหลังที่ k มีแนวโน้มเคลื่อนไหวในรูปเชิงเส้นด้วยความชันเป็นบวก ค่า r_k จะเข้าใกล้ 1 แต่ถ้าความชันเป็นลบ ค่า r_k จะเข้าใกล้ -1

ความคลาดเคลื่อนมาตรฐานของ r_k จะประมาณได้ดังนี้

$$se_{r_k} = \frac{1}{\sqrt{n}} \qquad \text{......................(6)}$$

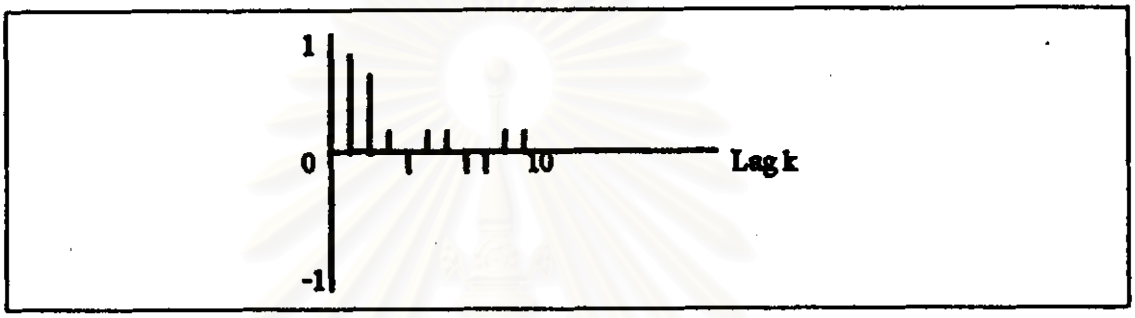
ค่าสถิติ t ในการทดสอบสมมุติฐานว่า พารามิเตอร์ความสัมพันธ์ในตัวเองเป็นศูนย์ หรือ $\rho_k = 0$ คือ

$$t_{r_k} = \frac{r_k}{se_{r_k}} \qquad \text{......................(7)}$$

ฟังก์ชันอัตตะสหสัมพันธ์ของกลุ่มตัวอย่างอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล แสดงได้ในรูปแบบที่แตกต่างกัน 2 ลักษณะ (Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; Makridakis, Wheelwright and McGee, 1978) ดังนี้

ลักษณะที่ 1 ค่า r_k ลดลงอย่างรวดเร็ว เป็น 0 เรียกว่า มีค่าต่ำสุด (cut off) ถ้า r_k ของอัตตะสหสัมพันธ์ของกลุ่มตัวอย่างมีค่าโด่งสุด (spike) ที่เวลาล้าหลัง k มีขนาดใหญ่ จะเทียบเท่ากับการปฏิเสธ H_0 นั่นคืออัตตะสหสัมพันธ์ทางทฤษฎี $\rho_k = 0$ และ ρ_k

มีการวัดความสัมพันธ์เชิงเส้นระหว่างค่าอนุกรมเวลาที่เป็นไปได้ทั้งหมด แยกโดยหน่วยของเวลาล้าหลังที่ k นักวิจัยสามารถตัดสินค่าโด่งสุด (spike) ที่เวลาล้าหลัง k ของอัตตสหสัมพันธ์ของกลุ่มตัวอย่าง โดยการพิจารณาจากค่าสถิติ t ที่สัมพันธ์กับ r_k ดังภาพ 5 แสดงลักษณะของฟังก์ชันอัตตสหสัมพันธ์ของกลุ่มตัวอย่างที่ลดลงอย่างรวดเร็ว (cut off) หลังเวลาล้าหลัง k (k=2)

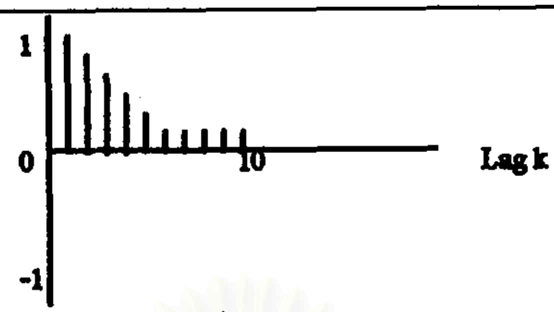


ภาพ 5 ฟังก์ชันอัตตสหสัมพันธ์ของกลุ่มตัวอย่างที่ลดลงอย่างรวดเร็ว (cut off) หลังเวลาล้าหลัง k

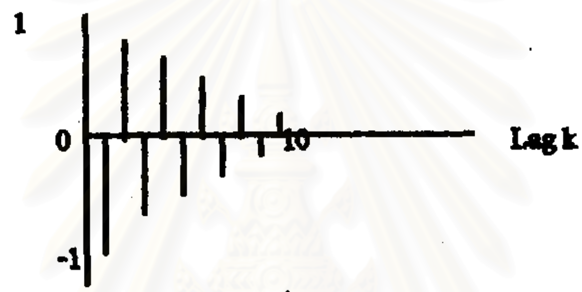
ลักษณะที่ 2 ค่า r_k มีค่ามากในเวลาล้าหลัง k แรก ๆ และลดลงอย่างช้า ๆ เกือบถึง 0 เมื่อ k มีค่ามากขึ้น เรียกว่า dies down หรือ tails off ใน 3 แบบ ดังภาพ 6

แบบที่ 1 ลดลงอย่างช้า ๆ แบบเอ็กซ์โปเนนเชียล (damped exponential)
แบบที่ 2 ลดลงอย่างช้า ๆ แบบฟังก์ชันตรีโกณมิติแบบไซน์ (damped sine - wave)
แบบที่ 3 แบบที่ 1 รวมกันแบบที่ 2

นักวิจัยใช้ลักษณะของฟังก์ชันอัตตสหสัมพันธ์ของกลุ่มตัวอย่างทั้ง 2 ลักษณะดังกล่าวข้างต้น เป็นตัวกำหนดลักษณะของโมเดลได้ 3 โมเดลด้วยกันคือ โมเดล AR(p), MA(q) และ ARMA(p,q) ซึ่งผู้วิจัยจะนำเสนอรายละเอียดในตอนต่อไป



แบบที่ 1.1 ลดลงอย่างช้า ๆ แบบเอ็กโปเนนเชียล (damped exponential) ที่ไม่แกว่งหรือไม่กระเพื่อม (no oscillation)



แบบที่ 1.2 ลดลงอย่างช้า ๆ แบบเอ็กโปเนนเชียล (damped exponential) ที่แกว่งหรือกระเพื่อม (oscillation)

แบบที่ 2 ลดลงอย่างช้า ๆ แบบฟังก์ชันตรีโกณมิติแบบไซน์ (damped sine-wave)

ภาพ 6 ฟังก์ชันอัตตะสหสัมพันธ์ของกลุ่มตัวอย่างที่ลดลงอย่างช้า ๆ (dying - down) แตกต่างกัน

ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล (Partial Autocorrelation Function)

อัตตะสหสัมพันธ์พาร์เชียล เป็นการวัดระดับของความเกี่ยวเนื่องระหว่างตัวแปร Z_t และ Z_{t+k} เมื่ออิทธิพลของเวลาล้าหลัง 1, 2, 3, ..., k-1 ถูกขจัดออกไป (partial out)

สัมประสิทธิ์อัตตะสหสัมพันธ์พาร์เชียลลำดับที่ m มีการให้ความหมายตามสัมประสิทธิ์การถดถอยในตัวเองตัวสุดท้ายของโมเดล AR(m) ดังสมการ (8)-(12) มีการให้ความหมายของกระบวนการ AR(1), AR(2), ..., AR(m-1) และ AR(m) สัมประสิทธิ์ตัวสุดท้ายของ Z_t ในแต่ละสมการคือสัมประสิทธิ์อัตตะสหสัมพันธ์พาร์เชียล นั่นคือ $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, ..., \hat{\phi}_{m-1}$ และ $\hat{\phi}_m$ เป็นสัมประสิทธิ์อัตตะสหสัมพันธ์พาร์เชียลตัวที่ m ของอนุกรมเวลา

$$Z_t = \hat{\phi}_1 Z_{t-1} + a_t \quad \text{..................(8)}$$
$$Z_t = \phi_1 Z_{t-1} + \hat{\phi}_2 Z_{t-2} + a_t \quad \text{..................(9)}$$
$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \hat{\phi}_3 Z_{t-3} + a_t \quad \text{..................(10)}$$

.

.

.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + ... + \hat{\phi}_{m-1} Z_{t-m+1} + a_t \quad \text{..................(11)}$$
$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + ... + \phi_{m-1} Z_{t-m+1} + \hat{\phi}_m Z_{t-m} + a_t \quad \text{..................(12)}$$

จากสมการ (8)-(12) ต้องประมาณค่า $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, ..., \hat{\phi}_{m-1}$ และ $\hat{\phi}_m$ โดยใช้สัมประสิทธิ์อัตตะสหสัมพันธ์ ซึ่งทำการประมาณค่าได้โดยการเอา Z_{t-1} คูณเข้าไปทั้งสองข้างในสมการ (8) จะได้

$$Z_{t-1} Z_t = \phi_1 Z_{t-1} Z_{t-1} + Z_{t-1} a_t \quad \text{..................(13)}$$

จากสมการ (13) เขียนในรูปค่าคาดหวัง (expected value) ได้ว่า

$$E(Z_{t-1} Z_t) = E(\phi_1 Z_{t-1} Z_{t-1}) + E(Z_{t-1} a_t)$$

หรือ

$$\gamma_1 = \phi_1 \gamma_0 \quad \text{..................(14)}$$

เมื่อ

$E(Z_{t-1} Z_t) = \gamma_1$, $E(Z_{t-1} Z_{t-1}) = \gamma_0$ และ $E(Z_{t-1} a_t) = 0$

γ_1 และ γ_0 = ความแปรปรวนร่วมในตัวเองของประชากร (population autocovariances) ในลำดับที่ 1 และ 0

จากสมการ (13) เอา γ_0 หารทั้งสองข้างจะได้

$$\rho_1=\phi_1 \text{ (เมื่อ } \rho_1=\gamma_1/\gamma_0) \qquad ....(15)$$

นั่นคือ สัมประสิทธิ์อัตตะสหสัมพันธ์พาร์เชียลตัวแรกจะเหมือนกันกับสัมประสิทธิ์อัตตะสหสัมพันธ์ตัวแรก และประมาณค่าสัมประสิทธิ์ทั้งสองตัวในกลุ่มตัวอย่างโดยใช้ r_1 โดยปกติขั้นตอนในสมการ (12)-(14) เรียกว่า สมการ Yule-Walker เพื่อประมาณค่า $\phi_1, \phi_2, \phi_3, ..., \phi_{m-1}$ และ ϕ_m และค่าที่ได้จะนำไปใช้ประมาณค่าของอัตตะสหสัมพันธ์พาร์เชียลของเวลาล้าหลังที่ m ขึ้นไป และสามารถเขียนความสัมพันธ์ในตัวเองของข้อมูลจากฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล (partial autocorrelation function) ของกลุ่มตัวอย่าง (sample partial autocorrelation function) ได้ดังสมการ

$$r_{kk} = r_1 \quad ; \quad k=1 \qquad ....(16)$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j}\, r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j}\, r_j} \quad ; \quad k = 2, 3, ... \qquad ....(17)$$

เมื่อ

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

ความคลาดเคลื่อนมาตรฐานของ r_{kk} คือ

$$se_{r_{kk}} = \frac{1}{\sqrt{n}} \qquad ....(18)$$

ค่าสถิติ t ในการทดสอบสมมุติฐาน คือ

$$t_{r_{kk}} = \frac{r_{kk}}{se_{r_{kk}}} \qquad ....(19)$$

ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลของกลุ่มตัวอย่างอนุกรมเวลาที่ไม่มีฤดูกาล แสดงได้ในรูปแบบที่แตกต่างกัน 2 ลักษณะเช่นเดียวกับฟังก์ชันอัตตะสหสัมพันธ์

Box และ Jenkins ได้สร้างโมเดลสำหรับการวิเคราะห์อนุกรมเวลาที่ไม่มีฤดูกาลไว้ 3 โมเดลดังนี้ (Newbold and Bos, 1994 ; Sullivan and Clayoombe, 1977 ; Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; Makridakis, Wheelwright and McGee, 1978 ; Box, Jenkins and Reinsel, 1994 ; Thomopoulos, 1980 ; ทรงศิริ แต้สมบัติ, 2539)

โมเดลที่ 1 กระบวนการถดถอยในตัวเอง (Autoregressive Processes - AR)

กระบวนการถดถอยในตัวเองลำดับที่ p แทนด้วยสัญลักษณ์ AR(p) เป็นโมเดลที่แสดงว่าค่าสังเกต Z_t ขึ้นอยู่กับค่า Z_{t-1}, Z_{t-2}, ..., Z_{t-p} หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + a_t \qquad \ldots\ldots\ldots\ldots(20)$$

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

เมื่อ Z_t = ตัวแปรตาม

δ = ค่าคงที่

ϕ_j = พารามิเตอร์ AR ตัวที่ j

a_t = ค่าความคลาดเคลื่อน ณ ช่วงเวลา t (random shock)

ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down แบบ damped exponential หรือ damped sine-wave หรือทั้งสองแบบรวมกัน ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ cut off คือ

$\rho_{kk} \neq 0$; $k = 1, 2, \ldots, p$

และ $\rho_{kk} = 0$; $k > p$

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

โมเดลที่ 2 กระบวนการการเฉลี่ยเคลื่อนที่ (Moving Average Processes - MA)

กระบวนการการเฉลี่ยเคลื่อนที่ลำดับที่ q แทนด้วยสัญลักษณ์ MA(q) เป็นโมเดลที่แสดงว่าค่าสังเกต Z_t ขึ้นอยู่กับค่าความคลาดเคลื่อน a_{t-1}, a_{t-2}, ..., a_{t-q} หรือความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า

$$Z_t = \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad \text{.........................(21)}$$

เมื่อ δ = ค่าคงที่ = μ

$\theta_1, \theta_2 \dots \theta_q$ = พารามิเตอร์ MA ตัวที่ q

a_t = ค่าความคลาดเคลื่อน ณ ช่วงเวลา t-k

ฟังก์ชันอัตตสหสัมพันธ์ ρ_k มีลักษณะ cut off คือ

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

และ $\rho_k = 0$; $k > q$

ส่วนฟังก์ชันอัตตสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down แบบ damped exponential หรือ damped sine-wave หรือทั้งสองแบบรวมกัน

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

โมเดลที่ 3 กระบวนการรวมกันของการถดถอยในตัวเองและการเฉลี่ยเคลื่อนที่ (Mixed Autoregressive - Moving Average Processes = ARMA)

กระบวนการรวมกันของการถดถอยในตัวเองและการเฉลี่ยเคลื่อนที่แทนด้วยสัญลักษณ์ ARMA(p,q)

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad \text{...........(22)}$$

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

ฟังก์ชันอัตตสหสัมพันธ์ ρ_k และฟังก์ชันอัตตสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down แบบ damped exponential หรือ damped sine-wave หรือทั้งสองแบบรวมกัน

ในโปรแกรม SPSS และ SAS เรียกโมเดลทั้ง 3 โมเดลที่กล่าวมาข้างต้นว่า ARIMA โดยที่

AR(p) = ARIMA(p,0,0)

MA(q) = ARIMA(0,0,q)

ARMA(p,q) = ARIMA(p,0,q)

เมื่อ p เป็นลำดับของกระบวนการถดถอยในตัวเอง

d เป็นระดับของความแตกต่างในการขจัดค่าแนวโน้มที่ไม่คงที่

q เป็นลำดับของกระบวนการการเฉลี่ยเคลื่อนที่

การวิเคราะห์อนุกรมเวลาทั้ง 3 โมเดล สามารถนำมาสรุปเป็นลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์ และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล (Bowerman and O'Connell, 1993) ได้ดังตาราง 3

ตาราง 3 ลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลของโมเดลการวิเคราะห์อนุกรมเวลาคงที่

โมเดล	ฟังก์ชันอัตตะสหสัมพันธ์	ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล
AR(p) = ARIMA(p,0,0)	dies down	cut off หลังเวลาล้าหลัง p
MA(q) = ARIMA(0,0,q)	cut off หลังเวลาล้าหลัง q	dies down
ARMA(p,q) = ARIMA (p,0,q)	dies down	dies down

จากที่ผู้วิจัยได้กล่าวถึงประเภทของข้อมูลอนุกรมเวลาว่ามี 3 ประเภท คือ ข้อมูลอนุกรมเวลาคงที่ อนุกรมเวลาไม่คงที่ และอนุกรมเวลาฤดูกาล เมื่อต้องการพยากรณ์ข้อมูลอนุกรมเวลาแต่ละประเภทจะสร้างโมเดลที่แตกต่างกันเป็น 3 โมเดลใหญ่ ๆ คือ AR(p), MA(q) และ ARMA(p,q) โดยแต่ในละโมเดลยังแบ่งเป็นโมเดลย่อย ๆ ได้อีก ดังรายละเอียดที่จะได้นำเสนอต่อไป

1. อนุกรมเวลาคงที่ (Stationary Time Series)

อนุกรมเวลาคงที่ หมายถึง อนุกรมที่มีโมเดลแบบ ARIMA(p,0,q) โมเดลนี้มีค่า d เป็น 0 เนื่องจากค่า d หมายถึง ระดับของความแตกต่างที่เกิดขึ้นจากการหาผลต่างลำดับต่าง ๆ ของข้อมูล เพื่อขจัดค่าแนวโน้มที่ไม่คงที่ออกจากข้อมูลอนุกรมเวลา ในกรณีที่ข้อมูลมีลักษณะเป็นอนุกรมเวลาคงที่อยู่แล้ว ไม่ต้องมีการขจัดค่าแนวโน้ม และจะไม่มีระดับของความแตกต่าง ดังนั้น ค่า d จึงเป็น 0

เมื่อ d = 0 จะพิจารณารูปแบบของโมเดลได้ 3 กรณี คือ กรณีแรก ARIMA(p,0,0) กรณีที่สอง ARIMA(0,0,q) และกรณีที่สาม ARIMA(p,0,q) ในแต่ละกรณีสามารถแบ่งเป็นโมเดลย่อย ๆ ได้ดังนี้

กรณีที่ 1 ARIMA(p,0,0) หรือ AR(p) แบ่งเป็น 2 โมเดลย่อย คือ

ก. โมเดลการถดถอยในตัวเองลำดับที่ 1 (the first order autoregressive model) หรือ AR(1)

โมเดลนี้ได้จากการแทนค่า p=1, d=0 และ q=0 หรือ ARIMA(1,0,0)

$$Z_t = \delta + \phi_1 Z_{t-1} + a_t \quad \text{.................................(23)}$$

โดยที่ $|\phi_1| < 1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down แบบ damped exponential

$$\rho_k = (\phi_1)^k \quad ; \; k \geq 1$$

ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ cut off หลังเวลาล้าหลังที่ 1

ข. โมเดลการถดถอยในตัวเองลำดับที่ 2 (the second order autoregressive model) หรือ AR(2)

โมเดลนี้ได้จากการแทนค่า p=2, d=0 และ q=0 หรือ ARIMA(2,0,0)

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t \quad \text{.................................(24)}$$

โดยที่ $\phi_2 + \phi_1 < 1$, $\phi_2 - \phi_1 < 1$ และ $|\phi_2| < 1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down แบบ damped exponential และ/หรือ damped sine-wave

$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$

$$\rho_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2$$

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \qquad \text{เมื่อ} ; k \geq 1$$

ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ cut off หลังเวลาล้าหลังที่ 2

กรณีที่ 2 ARIMA(0,0,q) หรือ MA(q) แบ่งเป็น 2 โมเดลย่อย คือ

ก. โมเดลการเฉลี่ยเคลื่อนที่ลำดับที่ 1 (the first order moving average model) หรือ MA(1)

โมเดลนี้ได้จากการแทนค่า p=0, d=0 และ q=1 หรือ ARIMA(0,0,1)

$$Z_t = \delta + a_t - \theta_1 a_{t-1} \quad \text{.................................(25)}$$

โดยที่ $|\theta_1| < 1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ cut off หลังเวลาล้าหลังที่ 1

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$$

$\rho_k = 0 \quad ; k>1$

ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down แบบ damped exponential

ข. โมเดลการเฉลี่ยเคลื่อนที่ลำดับที่ 2 (the second order moving average model) หรือ MA(2)

โมเดลนี้ได้จากการแทนค่า p=0, d=0 และ q=2 หรือ ARIMA(0,0,2)

$$Z_t = \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \quad \text{.................................(26)}$$

โดยที่ $\theta_2 + \theta_1 < 1$, $\theta_2 - \theta_1 < 1$ และ $|\theta_2| < 1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ cut off หลังเวลาล้าหลังที่ 2

$$\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$$

$$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$$

$$\rho_k = 0 \quad ; k>2$$

ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down แบบผสมของ damped exponential และ/หรือ damped sine-wave

กรณีที่ 3 ARIMA(p,0,q) หรือ ARMA(p,q) เมื่อ p=1 และ q=1 จะได้โมเดลดังนี้

ก. โมเดลการรวมกันของการถดถอยในตัวเองกับการเฉลี่ยเคลื่อนที่ลำดับที่ (1,1) (mixed autoregressive - moving average model of order (1,1)) หรือ ARMA(1,1)

โมเดลนี้ได้จากการแทนค่า p=1, d=0 และ q=1 หรือ ARIMA(1,0,1)

$$Z_t = \delta + \phi_1 Z_{t-1} + a_t - \theta_1 a_{t-1} \quad \text{.................................(27)}$$

โดยที่ $|\phi_1| < 1$ และ $|\theta_1| < 1$

ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down แบบ damped exponential

$$\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\theta_1\phi_1}$$

$$\rho_k = \phi_1\rho_{k-1} \quad ; k \geq 2$$

ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down แบบ damped exponential

การวิเคราะห์อนุกรมเวลาคงที่ทั้ง 5 โมเดล สามารถนำมาสรุปเป็นลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล (Bowerman and O'Connell, 1993) ได้ดังตาราง 4

ตาราง 4 ลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลของโมเดลการวิเคราะห์อนุกรมเวลาคงที่

โมเดล	ฟังก์ชันอัตตะสหสัมพันธ์	ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล
ARIMA(p,0,0)		
ก. AR(1)	dies down แบบ damped exponential $\rho_k = (\phi_1)^k$; $k \geq 1$	cut off หลังเวลาล้าหลังที่ 1
ข. AR(2)	dies down แบบผสมของ damped exponential และ/หรือ damped sine wave $\rho_1 = \frac{\phi_1}{1-\phi_2}$ $\rho_2 = \frac{\phi_1^2}{1-\phi_2} + \phi_2$ $\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}$; $k \geq 1$	cut off หลังเวลาล้าหลังที่ 2
ARIMA(0,0,q)		
ก. MA(1)	cut off หลังเวลาล้าหลังที่ 1 $\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$ $\rho_k = 0$; $k>1$	dies down แบบ damped exponential
ข. MA(2)	cut off หลังเวลาล้าหลังที่ 2 $\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$ $\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$ $\rho_k = 0$; $k>2$	dies down แบบผสมของ damped exponential และ/หรือ damped sine wave
ARIMA(p,0,q)		
ก. ARMA(1,1)	dies down แบบ damped exponential $\rho_1 = \frac{(1-\phi_1\theta_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\theta_1\phi_1}$ $\rho_k = \phi_1\rho_{k-1}$; $k \geq 2$	dies down แบบ damped exponential

2 อนุกรมเวลาไม่คงที่ (*Nonstationary Time Series*)

อนุกรมเวลาไม่คงที่ หมายถึง อนุกรมที่มีโมเดลแบบ ARIMA(p,d,q) โมเดลนี้มีค่า d เป็น 1, 2, 3, ... เนื่องจากเมื่อข้อมูลเป็นอนุกรมเวลาไม่คงที่ ทำให้เป็นอนุกรมเวลาคงที่โดยการหาผลต่างลำดับต่าง ๆ ของข้อมูล (regular difference = d)

ถ้า $d=1$, $W_t = Z_t - Z_{t-1}$ ....(28)

$d=2$, $W_t = Z_t - 2Z_{t-1} + Z_{t-2}$ ....(29)

จากผลต่างลำดับที่ 1 ในสมการ (28) ให้ W_t เป็นตัวแบบต่าง ๆ แทน Z_t เรียกกระบวนการ ARMA ในรูปของ W_t ว่า Autoregressive Integrated Moving Average หรือ ARIMA (p,d,q)

เมื่อ $d = 1$ เขียนสมการในรูป W_t ได้ดังนี้

$$W_t = \delta + \phi_1 W_{t-1} + \phi_2 W_{t-2} + ... + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ... - \theta_q a_{t-q} \quad ....(30)$$

หรือเขียนในเทอมของ Z_t ได้เป็น

$$\begin{aligned} Z_t - Z_{t-1} = \; & \delta + \phi_1(Z_{t-1} - Z_{t-2}) + \phi_2(Z_{t-2} - Z_{t-3}) + ... + \phi_p(Z_{t-p} - Z_{t-p-1}) \\ & + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ... - \theta_q a_{t-q} \quad ....(31) \end{aligned}$$

หรือ

$$\begin{aligned} Z_t = \; & \delta + Z_{t-1} + \phi_1(Z_{t-1} - Z_{t-2}) + \phi_2(Z_{t-2} - Z_{t-3}) + ... + \phi_p(Z_{t-p} - Z_{t-p-1}) \\ & + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - ... - \theta_q a_{t-q} \quad ....(32) \end{aligned}$$

ถ้าผลต่างลำดับที่ 1 ของข้อมูลยังไม่ทำให้อนุกรมเวลาเป็นอนุกรมคงที่ ก็หาผลต่างลำดับอื่น ๆ ไปจนกว่าจะเป็นอนุกรมคงที่ ในทางปฏิบัติแล้วมักใช้ผลต่างลำดับที่ 1 และอย่างมากก็ไม่เกินผลต่างลำดับที่ 2 ก็ทำให้โมเดลที่ได้เป็นอนุกรมคงที่ ในที่นี้ผู้วิจัยจึงจะเสนอโมเดลที่มักจะพบบ่อย ๆ ในทางปฏิบัติ คือ โมเดลที่เป็นผลต่างลำดับที่ 1 หรือ $d=1$

เมื่อ $d=1$ จะพิจารณารูปแบบของโมเดลได้ 3 กรณี คือ กรณีแรก ARIMA(p,d,0) กรณีที่สอง ARIMA(0,d,q) และกรณีที่สาม ARIMA(p,d,q) ในแต่ละกรณีสามารถแบ่งเป็นโมเดลย่อย ๆ ได้ดังนี้

กรณีที่ 1 ARIMA(p,d,0) หรือ ARI(p,d) แบ่งเป็น 2 โมเดลย่อย คือ

ก. โมเดลการถดถอยในตัวเองลำดับที่ 1 ที่ไม่คงที่ (the nonstationary first order autoregressive model) หรือ ARIMA(1,1,0) หรือ ARI(1,1)

โมเดลนี้ได้จากการแทนค่า p=1, d=1 และ q=0

$$Z_t = \delta + Z_{t-1} + \phi_1(Z_{t-1} - Z_{t-2}) + a_t \qquad \text{.................................(33)}$$

โดยที่ $|\phi_1|<1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ cut off หลังเวลาล้าหลังที่ 1

ข. โมเดลการถดถอยในตัวเองลำดับที่ 2 ที่ไม่คงที่ (the nonstationary second order autoregressive model) หรือ ARIMA(2,1,0) หรือ ARI(2,1)

โมเดลนี้ได้จากการแทนค่า p=2, d=1 และ q=0 มีสมการดังนี้

$$Z_t = \delta + Z_{t-1} + \phi_1(Z_{t-1} - Z_{t-2}) + \phi_2(Z_{t-2} - Z_{t-3}) + a_t \qquad \text{..............................(34)}$$

โดยที่ $\phi_2+\phi_1<1$, $\phi_2-\phi_1<1$ และ $|\phi_2|<1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ cut off หลังเวลาล้าหลังที่ 2

กรณีที่ 2 ARIMA(0,d,q) หรือ IMA(d,q) แบ่งเป็น 2 โมเดลย่อย คือ

ก. โมเดลการรวมกันของการเฉลี่ยเคลื่อนที่ลำดับที่ (0,1,1) (the integrated moving average model of order (0,1,1)) หรือ IMA(0,1,1)

โมเดลนี้ได้จากการแทนค่า p=0, d=1 และ q=1 ดังนั้น IMA(0,1,1) ก็คือ ARIMA(0,1,1) มีสมการดังนี้

$$Z_t = \delta + Z_{t-1} + a_t - \theta_1 a_{t-1} \qquad \text{..................................(35)}$$

โดยที่ $|\theta_1|<1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ cut off หลังเวลาล้าหลังที่ 1 ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down

ข. โมเดลการรวมกันของการเฉลี่ยเคลื่อนที่ลำดับที่ (0,1,2) (the integrated moving average model of order (0,1,2)) หรือ IMA(0,1,2)

โมเดลนี้ได้จากการแทนค่า p=0, d=1 และ q=2 ดังนั้น IMA(0,1,2) ก็คือ ARIMA(0,1,2) มีสมการดังนี้

$$Z_t = \delta + Z_{t-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \qquad \text{.................................(36)}$$

โดยที่ $\theta_2+\theta_1<1$, $\theta_2-\theta_1<1$ และ $|\theta_1|<1$ ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k

มีลักษณะ cut off หลังเวลาล้าหลังที่ 2 ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down

กรณีที่ 3 ARIMA(p,d,q) แบ่งเป็น 2 โมเดลย่อย คือ

ก. โมเดลการรวมกันของการถดถอยในตัวเองลำดับที่ 1 กับการเฉลี่ยเคลื่อนที่ลำดับที่ 1 (the first order autoregressive integrated first order moving average model) หรือ ARIMA(1,1,1)

โมเดลนี้ได้จากการแทนค่า p=1, d=1 และ q=1 มีสมการดังนี้

$$Z_t = \delta + Z_{t-1} + \phi_1(Z_{t-1} - Z_{t-2}) + a_t - \theta_1 a_{t-1} \quad \text{.................(37)}$$

โดยที่ $|\phi_1|<1$ และ $|\theta_1|<1$

ฟังก์ชันอัตตะสหสัมพันธ์ ρ_k มีลักษณะ dies down ส่วนฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ρ_{kk} มีลักษณะ dies down

ข. โมเดลที่เดินอย่างสุ่ม (the random walk model) หรือ ARIMA(0,1,0)

โมเดลนี้ได้จากการแทนค่า p=0, d=1 และ q=0 มีสมการดังนี้

$$Z_t = Z_{t-1} + a_t \quad \text{...............................(38)}$$

ฟังก์ชันอัตตะสหสัมพันธ์ทุก ρ_k และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลทุก ρ_{kk} เป็น 0

จากโมเดลอนุกรมเวลาไม่คงที่สามารถนำลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลของโมเดลมาสรุปได้ดังตาราง 5 (ทรงศิริ แต้สมบัติ, 2539)

ตาราง 5 ลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลของโมเดลการวิเคราะห์อนุกรมเวลาไม่คงที่ สำหรับโมเดล ARIMA(p,d,q)

โมเดล	ฟังก์ชันอัตตะสหสัมพันธ์ ($\rho_k(Z_t)$)	ฟังก์ชันอัตตะสหสัมพันธ์ ($\rho_k(W_t)$)	ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล ($\rho_k(W_t)$)
random walk	ลดลงอย่างช้า ๆ	ทุก ρ_k เป็น 0	ทุก ρ_{kk} เป็น 0
ARI(1,1)	ลดลงอย่างช้า ๆ	ค่าลดลงเร็วใกล้ 0	ρ_{kk} เป็น 0 สำหรับ k=2,...
ARI(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

1.3 อนุกรมเวลาฤดูกาล (Seasonal Time Series)

ถ้าข้อมูลที่สนใจมีการขึ้นลงตามฤดูกาล โมเดลของอนุกรมเวลาที่มีฤดูกาลยังคงเป็นลักษณะของโมเดล ARMA แต่จะมีการปรับโดยใช้ผลต่างของฤดูกาล (seasonal difference) กับอนุกรมเวลาเดิมดังนี้

$$W_t = \nabla_L Z_t = Z_t - Z_{t-L} \qquad \text{.................................(39)}$$

เมื่อ L คือช่วงเวลาของฤดูกาลที่ทำให้อนุกรม Z_t มีการเปลี่ยนแปลง

ดังนั้นจะหาโมเดลพยากรณ์อนุกรมเวลาได้จาก SARIMA$(P,D,Q)_L$ ซึ่งแสดงผลเกี่ยวข้องกับข้อมูลที่อยู่ในช่วงเวลาที่ติดต่อกัน อาจจะเป็นเดือน เป็นไตรมาส ฯลฯ อนุกรมเวลาฤดูกาลนั้น นอกจากข้อมูลที่เกี่ยวข้องกันระหว่างเดือนแล้วข้อมูลยังเกี่ยวข้องกันระหว่างปีด้วย ซึ่งหาโมเดลได้ดังนี้ ถ้าสมมุติให้ SARIMA(0,1,1) แสดงความสัมพันธ์ของข้อมูลที่อยู่ห่างกัน 12 ช่วง หรือ 12 เดือน จะได้

$$Z_t - Z_{t-12} = a_t - \theta_{12} a_{t-12} \qquad \text{.................................(40)}$$

โดยที่ $Z_t - Z_{t-12}$ คือ ผลต่างของข้อมูลที่ห่างกัน 12 ช่วง

θ_L คือ ค่าพารามิเตอร์ในโมเดลการเฉลี่ยเคลื่อนที่มีฤดูกาล (seasonal moving average model) และถ้าสมมุติให้ SARIMA$(0,1,1)_{12}$ แสดงความสัมพันธ์ของข้อมูลที่อยู่ห่างกัน 1 ช่วง หรือ 1 เดือน มีสมการดังนี้

$$Z_t - Z_{t-1} = a_t - \theta a_{t-1} \qquad \text{.................................(41)}$$

โดยที่ $Z_t - Z_{t-1}$ คือ ผลต่างของข้อมูลที่ห่างกัน 1 ช่วง

a_t คือ ความคลาดเคลื่อนที่เกิดขึ้นอย่างสุ่ม (random shock)

จากโมเดลที่มีฤดูกาล (seasonal model) ข้อมูลที่เกี่ยวข้องกันในแต่ละปี และในปีเดียวกันยังเกี่ยวข้องกันในแต่ละเดือนด้วย ซึ่งมีสมการดังนี้

$$(Z_t - Z_{t-1}) - (Z_{t-12} - Z_{t-13}) = (a_t - \theta a_{t-1}) - \theta_{12}(a_{t-12} - \theta_1 a_{t-13}) \qquad \text{...................(42)}$$

หรือ $Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} + a_t - \theta_1 a_{t-1} - \theta_{12} a_{t-12} - \theta_1 \theta_{12} a_{t-13}$ ....(43)

โมเดลอนุกรมเวลาที่มีฤดูกาล มีลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลดังตาราง 6 (ทรงศิริ แต้สมบัติ, 2539)

ตาราง 6 ลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลของโมเดลการวิเคราะห์อนุกรมเวลาที่มีฤดูกาล สำหรับโมเดล $SARMA(P,Q)_{12}$

โมเดล	ฟังก์ชันอัตตะสหสัมพันธ์	ฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียล
$SAR(1)_{12}$	ρ_{12}, ρ_{24} ค่าลดลงเร็วใกล้ 0	$\rho_{kk} = 0$ สำหรับ $k = 24, 36, \ldots$
$SAR(2)_{12}$	ρ_{12}, ρ_{24} ค่าลดลงเร็วใกล้ 0	$\rho_{kk} = 0$ สำหรับ $k = 36, 48 \ldots$
$SMA(1)_{12}$	$\rho_k = 0$ สำหรับ $k = 24, 36 \ldots$	$\rho_{12,12}$, $\rho_{24,24}$ ค่าลดลงเร็วใกล้ 0
$SMA(2)_{12}$	$\rho_k = 0$ สำหรับ $k = 36, 48 \ldots$	$\rho_{12,12}$, $\rho_{24,24}$ ค่าลดลงเร็วใกล้ 0
$SARMA(1,1)_{12}$	ρ_{12}, ρ_{24} ค่าลดลงเร็วใกล้ 0	$\rho_{12,12}$, $\rho_{24,24}$ ค่าลดลงเร็วใกล้ 0

ในกรณีที่อนุกรมเวลาไม่คงที่เนื่องจากมีการเปลี่ยนแปลงที่แสดงแนวโน้มและการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยต้องเปลี่ยนให้อยู่ในรูปอนุกรมคงที่ โดยการหาผลต่างของแนวโน้มและผลต่างฤดูกาล และเขียนโมเดลได้ว่า

$$ARIMA(p,d,q) \times SARIMA(P,D,Q)_L$$

เช่น $ARIMA(1,1,1) \times SARIMA(1,1,1)_{12}$ สามารถเขียนสมการได้ดังนี้

$$Z_t = (1+\phi_1)Z_{t-1} + (1+\phi_{12})Z_{t-12} - (1+\phi_1+\phi_{12}+\phi_1\phi_{12})Z_{t-13} + (\phi_1+\phi_1\phi_{12})\phi_1)Z_{t-14} - \phi_{12} Z_{t-15}$$
$$+ (\phi_{12}+\phi_1\phi_{12}) Z_{t-16} - \phi_1\phi_{12}Z_{t-17} + a_t - \theta_1 a_{t-1} - \theta_{12} a_{t-12} + \theta_1\theta_{12} a_{t-13}$$

การเขียนโมเดลต้องรวม (intergrate) ทั้งโมเดลอนุกรมเวลาที่มีแนวโน้มและฤดูกาลเข้าด้วยกัน เมื่อกำหนดโมเดลที่เหมาะสมกับอนุกรมเวลา โดยการพิจารณาจากฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์พาร์เชียลได้แล้ว จึงนำโมเดลนั้นไปประมาณค่าพารามิเตอร์ในขั้นตอนต่อไป

ตอนที่ 2.2 การประมาณค่าพารามิเตอร์

เมื่อกำหนดโมเดลได้แล้ว ขั้นต่อไปเป็นการประมาณค่าพารามิเตอร์ด้วยวิธีการโปรแกรมที่เป็นเส้นโค้งหรือไม่เป็นเส้นตรง (non-linear programing) ค่าประมาณที่ได้เป็นค่าประมาณเบื้องต้นซึ่งจะนำไปหาค่าประมาณสุดท้ายของพารามิเตอร์ ค่าประมาณสุดท้ายของพารามิเตอร์ที่ดีที่สุดจะนำไปใช้ในแต่ละโมเดล จะเป็นค่าพารามิเตอร์ที่เมื่อนำไปพยากรณ์ค่าของอนุกรมเวลาแล้วจะมีผลทำให้ ผลบวกกำลังสองของความคลาดเคลื่อนมีค่าน้อยที่สุด (Newbold and Bos, 1994 ; Hanke and Reitsch, 1992 ; Makridakis, Wheelwright and McGee, 1978)

การประมาณค่าพารามิเตอร์สามารถกระทำได้ 2 แนวทาง (Makridakis, Wheelwright andMcGee, 1978) ได้แก่

แนวทางที่ 1 การลองผิดลองถูก (trial and error) เป็นการตรวจสอบค่าความแตกต่างหลาย ๆ ค่าและเลือกค่าที่มีผลรวมความคลาดเคลื่อนกำลังสอง (sum of squared residuals) ต่ำที่สุด

แนวทางที่ 2 การปรับปรุงแบบคำนวณทวนซ้ำ (iterative improvement) ใช้โปรแกรมคอมพิวเตอร์ ประมาณค่าแบบคำนวณทวนซ้ำ

2.2.1 กระบวนการ AR(1) และ AR(2) ที่ไม่มีฤดูกาล

กระบวนการ AR(p) ใช้สมการ Yule-Walker ดังนี้

$$\left.\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{k-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{k-2} \\
&\;\;\vdots \\
\rho_k &= \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \dots + \phi_p
\end{aligned}\right\} \quad \dots\dots\dots\dots\dots\dots(44)$$

เมื่อ $\rho_1, \rho_2, \dots, \rho_k$ เป็นสัมประสิทธิ์อัตตะสหสัมพันธ์ทางทฤษฎี สำหรับช่วงเวลา 1, 2, 3, ..., k

$\phi_1, \phi_2, \dots, \phi_p$ เป็นสัมประสิทธิ์ของกระบวนการ AR(p)

ρ_k เป็นค่าทางทฤษฎีที่ไม่ทราบค่า จึงต้องแทนด้วยค่าเชิงประจักษ์ที่เทียบเท่ากัน คือ r_k (สัมประสิทธิ์สหสัมพันธ์เชิงประจักษ์)

ดังนั้น กระบวนการ AR(1) เขียนสมการใหม่ด้วย $p = 1$ ได้

$$\rho_1 = \phi_1 \qquad \text{.................................(45)}$$

แทนค่า ρ_1 ที่ไม่ทราบค่าด้วย r_1 เพื่อประมาณค่าพารามิเตอร์ ϕ_1 ใน AR(1) จะได้

$$\hat{\phi}_1 = r_1 \qquad \text{.................................(46)}$$

และกระบวนการ AR(2) เขียนสมการ Yule-Walker ของ $p = 2$ ได้ดังนี้

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 , \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \qquad \text{.................................(47)}$$

แทนค่า ρ_1 และ ρ_2 ด้วย r_1 และ r_2 จะได้

$$\left.\begin{aligned} \hat{\phi}_1 &= \frac{r_1(1-r_2)}{1-r_1^2} \\ \hat{\phi}_2 &= \frac{r_2-r_1^2}{1-r_1^2} \end{aligned}\right\} \qquad \text{.................................(48)}$$

2.2.2 กระบวนการ MA(1) และ MA(2) ที่ไม่มีฤดูกาล

อัตตะสหสัมพันธ์ทางทฤษฎีของกระบวนการ MA(q) สามารถเขียนในรูปของสัมประสิทธิ์ MA ได้ดังนี้

$$\rho_k = \frac{-\theta_k + \theta_1\theta_{k-1} + \ldots + \theta_{q-k}\theta_q}{1+\theta_1^2 + \ldots + \theta_q^2}, \quad k = 1, 2, \ldots q$$

$$\rho_k = 0, \qquad k > q \qquad \text{.................(49)}$$

ซึ่ง ρ_k เป็นค่าทางทฤษฎีที่ไม่ทราบค่า การประมาณค่าสัมประสิทธิ์ $\theta_1, \theta_2, ..., \theta_q$ สามารถทำได้โดยการแทนค่า อัตตะสหสัมพันธ์เชิงประจักษ์ คือ r_k ในสมการ แล้วแก้สมการ ดังนั้น กระบวนการ MA(1) เมื่อ $q=1$ จะได้สมการ

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2}, \quad k=2 \quad ....(50)$$

$$\rho_k = 0, \quad k \geq 2$$

แทนค่า r_1 ใน ρ_1 จะได้สมการเส้นโค้ง (quadratio) แก้สมการหาค่า θ_1 ดังนี้

$$\hat{\theta}_1^2 + \left(\frac{1}{r_1}\right)\hat{\theta}_1 + 1 = 0 \quad ....(51)$$

และกระบวนการ MA(2) ที่ $q=2$ จะได้สมการ

$$\rho_1 = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$$

$$\rho_2 = \frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$$

$$\rho_k = 0 \quad , k \geq 3 \quad ....(52)$$

แทนค่า r_1 และ r_2 ใน ρ_1 และ ρ_2 แก้สมการหาค่า $\hat{\theta}_1$ และ $\hat{\theta}_2$

เมื่อนักวิจัยประมาณค่าพารามิเตอร์ได้แล้ว ยังไม่สามารถนำโมเดลไปใช้พยากรณ์ได้ เพราะต้องมีการตรวจสอบโมเดลก่อน วิธีการตรวจสอบโมเดลต้องใช้ค่าประมาณค่าพารามิเตอร์ในการคำนวณความคลาดเคลื่อน และทดสอบว่าค่าความคลาดเคลื่อนระหว่างโมเดลกับข้อมูลเชิงประจักษ์ เป็นไปโดยสุ่มหรือไม่ ดังในตอนที่ 2.3 ที่จะกล่าวในหัวข้อต่อไป

ตอนที่ 2.3 การตรวจสอบความเหมาะสมของโมเดล

การตรวจสอบความเหมาะสมของโมเดล เป็นการตรวจสอบความคลาดเคลื่อนอย่างสุ่ม ซึ่งจะตรวจสอบอัตตะสหสัมพันธ์ของความคลาดเคลื่อนที่เวลาล้าหลังต่าง ๆ โดยการกำหนดสมมติฐานหลักและสมมติฐานเลือกดังนี้ (Newbold and Bos, 1994 ; Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; Makridakis, Wheelwright and McGee, 1978 ; Montgomery, Johnson and Gardine, 1990 ; ทรงศิริ แต้สมบัติ, 2539)

$$H_0 : \rho_1(e_t) = \ldots = \rho_m(e_t) = 0$$

$H_1 : \rho_k(e_t)$ อย่างน้อยหนึ่งค่าที่ไม่เท่ากับ 0 สำหรับ $k = 1, 2, \ldots, m$

การตรวจสอบว่าโมเดลเหมาะสมหรือไม่ ใช้การทดสอบไคสแควร์ ด้วยสถิติ Box-Pierce Chi-Square Statistics ใช้สัญลักษณ์ Q_m ดังสมการ

$$Q_m = (N-d)\sum_{k=1}^{m} r_k^2(e_t) \quad \ldots\ldots\ldots\ldots(53)$$

เมื่อ N = จำนวนข้อมูลที่ใช้ในการวิเคราะห์อนุกรมเวลา

k = เวลาล้าหลังตัวแรกที่ตรวจสอบ

m = เวลาล้าหลังที่มากที่สุดที่ตรวจสอบ

$r_k(e_t)$ = อัตตะสหสัมพันธ์ของความคลาดเคลื่อน ในเวลาล้าหลังที่ k

d = ลำดับผลต่างของข้อมูลที่ทำให้เป็นอนุกรมคงที่

การประมาณค่าการแจกแจงของ Q_m มีองศาความเป็นอิสระ (degree of freedom) เท่ากับ $k-p-q$ ถ้าค่า Q_m ที่คำนวณได้มีค่ามากกว่าไคสแควร์ในตารางที่ $df = k-p-q$ แสดงว่าโมเดลนี้ไม่เหมาะสม จะต้องกลับไปขั้นตอนที่ 1 และ 2 เพื่อเลือกโมเดลและวิเคราะห์ใหม่ จนได้โมเดลที่เหมาะสม และเมื่อผลการทดสอบยอมรับสมมติฐานหลัก (accept null hypothesis) แสดงว่าค่าความคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน 1, 2, ..., m ช่วงเวลาเป็นอิสระต่อกัน จึงสรุปได้ว่าโมเดลที่กำหนดให้กับอนุกรมเวลามีความเหมาะสม ผู้วิจัยสามารถนำไปใช้พยากรณ์ได้

การทดสอบนี้ได้มีการพัฒนาขึ้นต่อมาโดย Box และ Ljung เรียกว่า modified Box-Pierce test สูตรการทดสอบที่พัฒนาใหม่นี้ใช้ตัวทดสอบสถิติ Q_m' แทน Q_m

$$Q_m' = N(N-d+2)\sum_{k=1}^{m} (r_k^2(e_t))/(N-d-k) \quad \ldots\ldots\ldots\ldots(54)$$

การทดสอบของ Box-Pierce หรือการทดสอบของ Box - Ljung ต่างก็เป็นการทดสอบสมมติฐานหลักและสมมติฐานเลือกเหมือนกัน มีช่วงวิกฤตเดียวกัน แต่ต่างกันที่ตัวทดสอบสถิติ

ตอนที่ 2.4 การพยากรณ์ค่า

เมื่อได้โมเดลที่เหมาะสมและผ่านการตรวจสอบความเหมาะสมแล้ว นักวิจัยสามารถนำโมเดลอนุกรมเวลาบ็อกซ์และเจ็นกิ้นส์ ไปใช้พยากรณ์ไปข้างหน้ากี่ช่วงเวลาก็ได้ แต่ยิ่งพยากรณ์ออกไปไกลเท่าไร ค่าพยากรณ์ที่ได้จะอาศัยสารสนเทศจากข้อมูลจริงน้อยลงมากเท่านั้น โดยนำค่าจากการพยากรณ์ในคาบเวลาที่ผ่านมา มาใช้พยากรณ์ในเวลาข้างหน้า ความแม่นยำจึงลดลงมากถ้ายิ่งพยากรณ์ไกลออกไป และถ้าอัตตะสหสัมพันธ์มีเครื่องหมายบวก ค่าพยากรณ์ในอนาคตจะมีแนวโน้มสูงขึ้นเรื่อย ๆ ในทำนองเดียวกันถ้าอัตตะสหสัมพันธ์มีเครื่องหมายลบ ค่าพยากรณ์ในอนาคตจะมีแนวโน้มลดลงเรื่อย ๆ ทั้งที่ในความเป็นจริงแล้วพฤติกรรมในอนาคตมีแนวโน้มที่เป็นไปได้ทั้งเพิ่มขึ้นและลดลง อนุกรมเวลาบ็อกซ์และเจ็นกินส์จึงเหมาะกับการพยากรณ์ในระยะสั้น ๆ และควรพยากรณ์เพียง 1 ช่วงเวลาล่วงหน้าเท่านั้นจึงเกิดความแม่นยำสูงที่สุด เนื่องจากได้อาศัยสารสนเทศของข้อมูลที่เกิดขึ้นจริงมากที่สุด (Newbold and Bos, 1994 ; Hanke and Reitsch, 1992 ; Bowerman and O'Connell, 1993 ; Makridakis, Wheelwright and McGee, 1978 ; Pindyck and Rubinfeld, 1985)

สำหรับสมการที่ใช้พยากรณ์ของโมเดล ARIMA (1,d,0), (2,d,0), (0,d,1), (0,d,2) และ (1,d,1) มีสมการดังตาราง 7 (Thomopoulos, 1980) เมื่อ d คือ ระดับของผลต่างลำดับที่ 0, 1, 2 และ T คือช่วงเวลาที่จะพยากรณ์ล่วงหน้า เช่น นักวิจัยต้องการพยากรณ์ล่วงหน้า 1 ช่วงเวลา แสดงว่า T=1 ซึ่งการพยากรณ์ด้วยโมเดลบ็อกซ์และเจ็นส์กิ้นส์ นักวิจัยจะพยากรณ์ไปข้างหน้ากี่ช่วงเวลาก็ได้ ดังนั้น T จึงมีค่าเป็นได้ตั้งแต่ 1, 2, 3, ... ตามความต้องการของนักวิจัย

ตาราง 7 สมการพยากรณ์ของวิธีบ๊อกซ์และเจ็นกิ้นส์

โมเดล	สมการพยากรณ์
ARIMA (1,0,0)	$\hat{Z}_t(1) = \phi_1 Z_t + (1-\phi_1)\bar{Z}$ $\hat{Z}_t(T) = \phi_1 \hat{Z}_t(T-1) + (1-\phi_1)\bar{Z}$;T = 2,3,...
ARIMA (1,2,0)	$\hat{Z}_t(1) = (2+\phi_1)Z_t - (1+2\phi_1)Z_{t-1} + \phi_1 Z_{t-2} - \theta_1 a_t$ $\hat{Z}_t(2) = (2+\phi_1)\hat{Z}_t(1) - (1+2\phi_1)Z_t + \phi_1 Z_{t-1}$ $\hat{Z}_t(3) = (2+\phi_1)\hat{Z}_t(2) - (1+2\phi_1)\hat{Z}_t(1) + \phi_1 Z_t$ $\hat{Z}_t(T) = (2+\phi_1)\hat{Z}_t(T-1) - (1+2\phi_1)\hat{Z}_t(T-2) + \phi_1 \hat{Z}_t(T-3)$;T = 4,5,...
ARIMA (2,0,0)	$\hat{Z}_t(1) = \phi_1 Z_t + \phi_2 Z_{t-1} + (1-\phi_1-\phi_2)\bar{Z}$ $\hat{Z}_t(2) = \phi_1 \hat{Z}_t(1) + \phi_2 Z_t + (1-\phi_1-\phi_2)\bar{Z}$ $\hat{Z}_t(T) = \phi_1 \hat{Z}_t(T-1) + \phi_2 \hat{Z}_t(T-2) + (1-\phi_1-\phi_2)\bar{Z}$; T = 3,4,...
ARIMA (2,1,0)	$\hat{Z}_t(1) = (1+\phi_1)Z_t + (\phi_2-\phi_1)Z_{t-1} - \phi_2 Z_{t-2}$ $\hat{Z}_t(2) = (1+\phi_1)\hat{Z}_t(1) + (\phi_2-\phi_1)Z_t - \phi_2 Z_{t-1}$ $\hat{Z}_t(3) = (1+\phi_1)\hat{Z}_t(2) + (\phi_2-\phi_1)\hat{Z}_t(1) - \phi_2 Z_t$ $\hat{Z}_t(T) = (1+\phi_1)\hat{Z}_t(T-1) + (\phi_2-\phi_1)\hat{Z}_t(T-2) - \phi_2 \hat{Z}_t(T-3)$; T = 4,5,...
ARIMA (2,2,0)	$\hat{Z}_t(1) = (2+\phi_1)Z_t + (\phi_2-2\phi_1-1)Z_{t-1} + (\phi_1-2\phi_2)Z_{t-2} + \phi_2 Z_{t-3}$ $\hat{Z}_t(2) = (2+\phi_1)\hat{Z}_t(1) + (\phi_2-2\phi_1-1)Z_t + (\phi_1-2\phi_2)Z_{t-1} + \phi_2 Z_{t-2}$ $\hat{Z}_t(3) = (2+\phi_1)\hat{Z}_t(2) + (\phi_2-2\phi_1-1)\hat{Z}_t(1) + (\phi_1-2\phi_2)Z_t + \phi_2 Z_{t-1}$ $\hat{Z}_t(4) = (2+\phi_1)\hat{Z}_t(3) + (\phi_2-2\phi_1-1)\hat{Z}_t(2) + (\phi_1-2\phi_2)\hat{Z}_t(1) + \phi_2 Z_t$ $\hat{Z}_t(T) = (2+\phi_1)\hat{Z}_t(T-1) + (\phi_2-2\phi_1-1)\hat{Z}_t(T-2) + (\phi_1-2\phi_2)\hat{Z}_t(T-3) + \phi_2 \hat{Z}_t(T-4)$; T = 5,6,...
ARIMA (0,0,1)	$\hat{Z}_t(1) = -\theta_1 a_t + \bar{Z}$ $\hat{Z}_t(T) = \bar{Z}$;T = 2,3,...
ARIMA (0,1,1)	$\hat{Z}_t(1) = -\theta_1 a_t + Z_t$ $\hat{Z}_t(T) = Z_t(T-1)$;T = 2,3,...
ARIMA (0,2,1)	$\hat{Z}_t(1) = -\theta_1 a_t + 2Z_t - Z_{t-1}$ $\hat{Z}_t(2) = 2\hat{Z}_t(1) - Z_t$ $\hat{Z}_t(T) = 2\hat{Z}_t(T-1) - \hat{Z}_t(T-2)$;T = 3,4,...

ตาราง 7 (ต่อ)

โมเดล	สมการพยากรณ์
ARIMA (0,0,2)	$\hat{Z}_t(1) = -\theta_1 a_t - \theta_2 a_{t-1} + \overline{Z}$ $\hat{Z}_t(2) = -\theta_2 a_t + \overline{Z}$ $\hat{Z}_t(T) = \overline{Z}$;T= 3,4,...
ARIMA (0,1,2)	$\hat{Z}_t(1) = -\theta_1 a_t - \theta_2 a_{t-1} + Z_t$ $\hat{Z}_t(2) = -\theta_2 a_t + \hat{Z}_t(1)$ $\hat{Z}_t(T) = \hat{Z}_t(T-1)$;T = 3,4,...
ARIMA (0,2,2)	$\hat{Z}_t(1) = -\theta_1 a_t - \theta_2 a_{t-2} + 2Z_t - Z_{t-1}$ $\hat{Z}_t(2) = -\theta_2 a_t + 2\hat{Z}_t(1) - Z_t$ $\hat{Z}_t(T) = 2\hat{Z}_t(T-1) - \hat{Z}_t(T-2)$;T = 3,4,...
ARIMA (1,0,1)	$\hat{Z}_t(1) = \phi_1 Z_t - \theta_1 a_t + (1-\phi_1)\overline{Z}$ $\hat{Z}_t(T) = \phi_1 \hat{Z}_t(T-1) + (1-\phi_1)\overline{Z}$;T = 2,3,...
ARIMA (1,1,1)	$\hat{Z}_t(1) = (1+\phi_1)Z_t - \phi_1 Z_{t-1} - \theta_1 a_t$ $\hat{Z}_t(2) = (1+\phi_1)\hat{Z}_t(1) - \phi_1 Z_t$ $\hat{Z}_t(T) = (1+\phi_1)\hat{Z}_t(T-1) - \phi_1 \hat{Z}_t(T-2)$;T = 3,4,...
ARIMA (1,2,1)	$\hat{Z}_t(1) = (2+\phi_1)Z_t - (1+2\phi_1)Z_{t-1} + \phi_1 Z_{t-2} - \theta_1 a_t$ $\hat{Z}_t(2) = (2+\phi_1)\hat{Z}_t(1) - (1+2\phi_1)Z_t + \phi_1 Z_{t-1}$ $\hat{Z}_t(3) = (2+\phi_1)\hat{Z}_t(2) - (1+2\phi_1)\hat{Z}_t(1) + \phi_1 Z_t$ $\hat{Z}_t(T) = (2+\phi_1)\hat{Z}_t(T-1) - (1+2\phi_1)\hat{Z}_t(T-2) + \phi_1\hat{Z}_t(T-3)$;T = 4,5,...

ตอนที่ 2.5 การตรวจสอบผลการพยากรณ์

หลังจากที่มีการพยากรณ์ข้อมูลอนุกรมเวลาและประมาณค่าพารามิเตอร์แล้ว นักวิจัยต้องตรวจสอบผลการพยากรณ์ว่ามีความถูกต้องเพียงใด กระบวนการตรวจสอบผลการพยากรณ์โดยทั่วไปจะมี 2 ขั้นตอน คือ ขั้นตอนแรกเป็นการตรวจสอบฟังก์ชันอัตตะสหสัมพันธ์ของโมเดลอนุกรมเวลา ที่สามารถเปรียบเทียบกับฟังก์ชันอัตตะสหสัมพันธ์ของกลุ่มตัวอย่างในอนุกรมแรก ขั้นตอนที่สอง คือการวัดความคลาดเคลื่อน (Pindyck and Rubinfeld, 1985)

การวัดความคลาดเคลื่อน จึงมีบทบาทสำคัญในการตรวจสอบความถูกต้องของเครื่องวัด วิธีการวัดความคลาดเคลื่อนในการพยากรณ์มีหลายวิธี เช่น Weigend et al. (1990) ; Weigend (1991); Nowlan and Hinton (1992 อ้างถึงใน Lachtermacher and Fuller, 1994) ใช้ความแปรปรวนสัมพัทธ์เฉลี่ย (Average Relative Variance หรือ ARV) Gorr et al. (1992 อ้างถึงใน Lachtermacher and Fuller, 1994) ใช้ความคลาดเคลื่อนเฉลี่ย (Mean Error หรือ ME) ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (Mean Absolute Error หรือ MAE) รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error หรือ RMSE) Coakley et al. (1992 อ้างถึงใน Lachtermacher and Fuller, 1994) และ Hill et al. (1991) ใช้ความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error หรือ MSE) และค่าเฉลี่ยของค่าสมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Mean Absolute Percentage Error หรือ MAPE) Chung and Lee (1992 อ้างถึงใน Lachtermacher and Fuller, 1994) ใช้รากที่สองของค่าเฉลี่ยของกำลังสองของความคลาดเคลื่อนปกติ (Normalized Root Mean Squared Error หรือ NRMSE) เป็นต้น

วิธีการวัดความคลาดเคลื่อนที่กล่าวมาข้างต้นแยกได้เป็น 2 กลุ่ม กลุ่มแรกคือการวัดความคลาดเคลื่อนในรูปความแปรปรวน หรือความคลาดเคลื่อนในรูปค่าเฉลี่ยของผลรวมกำลังสองของความคลาดเคลื่อน กลุ่มที่สอง คือการวัดความคลาดเคลื่อนความคลาดเคลื่อนในรูปความคลาดเคลื่อนเฉลี่ย หรือค่าเฉลี่ยของค่าสมบูรณ์ของความคลาดเคลื่อนในรูปร้อยละ ดังตาราง 8

ตาราง 8 วิธีการวัดความคลาดเคลื่อนในตัวอย่างงานวิจัย

งานวิจัยของ	วิธีการวัดความคลาดเคลื่อน	
	กลุ่มที่ 1	กลุ่มที่ 2
1.Weigned et al.	ARV	-
2.Nowlan and Hinton	ARV	-
3.Gorr et al.	RMSE	ME, MAE
4.Coakley et al.	MSE	MAPE
5.Hill et al.	MSE	MAPE
6.Chung and Lee	NRMSE	-

จากที่กล่าวมาจะเห็นได้ว่าการวัดความคลาดเคลื่อนในการพยากรณ์มีหลายวิธี นักวิจัยจึงพยายามหาข้อสรุปว่าควรจะเลือกใช้วิธีการใดวัดความคลาดเคลื่อนของการพยากรณ์ Armstrong and Collopy (1992) ได้ศึกษาเปรียบเทียบการวัดความคลาดเคลื่อนของการพยากรณ์ 6 แบบ คือ รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error หรือ RMSE) ค่ามัธยฐานของค่าสมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Median Absolute Percentage Error หรือ MdAPE) ค่าเฉลี่ยของค่าสมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Mean Absolute Percentage Erro r หรือ MAPE) ค่าเฉลี่ยเรขาคณิตของความคลาดเคลื่อนสมบูรณ์สัมพันธ์ (Geometric Mean of the Relative Absolute Error หรือ GMRAE) ค่ามัธยฐานของค่าสมบูรณ์สัมพัทธ์ของความคลาดเคลื่อน (Median Relative Absolute Error หรือ MdRAE) และ ร้อยละที่ดีกว่า (Percent Better) โดยเปรียบเทียบในด้าน ความเที่ยง ความตรงเชิงโครงสร้าง ความไว การป้องกันค่าสุดโต่ง ความสัมพันธ์ต่อการตัดสินใจ ได้ผลดังตาราง 9

ตาราง 9 การเปรียบเทียบการวัดความคลาดเคลื่อนในวิธีการพยากรณ์โดยทั่วไป

การวัดความคลาดเคลื่อน เกณฑ์ในการเปรียบเทียบ	RMSE	Percent Better	MAPE	MdAPE	GMRAE	MdRAE
ความเที่ยง	p	g	f	f	f	f
ความตรงเชิงโครงสร้าง	f	f	g	g	g	g
การป้องกันค่าสุดโต่ง	p	g	p	g	f	g
ความไว	g	p	g	p	g	p
ความสัมพันธ์ต่อการตัดสินใจ	g	p	f	f	p	p

เมื่อ p = poor (ต่ำ)
f = fair (ปานกลาง)
g = good (สูง)

จากตาราง 9 จะเห็นได้ว่า ไม่มีการวัดความคลาดเคลื่อนวิธีใดที่ดีที่สุดในทุกเกณฑ์ ดังนั้น การเลือกวิธีวัดความคลาดเคลื่อนควรเลือกใช้ตามสถานการณ์ เช่น การตรวจสอบความถูกต้องของเครื่องวัดควรใช้ GMRAE ถ้าข้อมูลอนุกรมเวลามีจำนวนน้อยควรใช้ MdRAE ถ้าข้อมูลอนุกรมเวลามีจำนวนมากควรใช้ MdAPE และในการเปรียบเทียบวิธีการพยากรณ์ส่วนใหญ่มักจะใช้ RMSE ซึ่งจากการศึกษาครั้งนี้ Armstrong and Collopy (1994) เสนอแนะ

ว่าไม่ควรใช้ เพราะมีความเที่ยงต่ำ ส่วน MAPE ไม่ควรใช้ในกรณีที่คาดว่าจะมีความคลาดเคลื่อนมาก เพราะ MAPE มีความลำเอียงในการประมาณค่าต่ำ

อย่างไรก็ตาม Makridakis (1993) ได้เสนอแนะว่าการวัดความถูกต้องในทางทฤษฎีและทางปฏิบัติ ควรใช้ MAPE เพราะเมื่อปรับจุดอ่อนของ MAPE ตามวิธีของ Makridakis แล้ว จะทำให้ MAPE ง่ายต่อการแปลความหมาย แต่ในการวิจัยครั้งนี้ผู้วิจัยใช้วิธีตรวจสอบผลการพยากรณ์โดยการวัดความคลาดเคลื่อน 6 แบบ คือ RMSE, Percent Better, MAPE, MdAPE, GMRAE และ MdRAE

เนื่องจากในงานวิจัยครั้งนี้ ผู้วิจัยสนใจเปรียบเทียบประสิทธิภาพในการพยากรณ์ด้วยวิธีบ็อกซ์และเจนกิ้นส์ กับการพยากรณ์ 3 วิธี คือ วิธีการวิเคราะห์การถดถอย วิธีการเฉลี่ยเคลื่อนที่ และวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ในตอนต่อไปผู้วิจัยจึงได้นำเสนอวิธีการดังกล่าวตามลำดับ

ตอนที่ 3 วิธีการพยากรณ์ที่ใช้เปรียบเทียบกับวิธีบ็อกซ์และเจนกิ้นส์

สาระในตอนนี้ผู้วิจัยแยกนำเสนอเป็น 3 ตอน ตอนแรกกล่าวถึงการวิเคราะห์การถดถอยกับข้อมูลอนุกรมเวลา ตอนต่อไปเป็นวิธีการเฉลี่ยเคลื่อนที่ และตอนสุดท้ายเป็นวิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ดังนี้

ตอนที่ 3.1 การวิเคราะห์การถดถอยกับข้อมูลอนุกรมเวลา

การวิเคราะห์ข้อมูลอนุกรมเวลาที่ใช้หลักการวิเคราะห์การถดถอย มีตัวแปรอนุกรมเวลาเป็นตัวแปรตาม และมีตัวแปรอิสระที่มีความหมายต่าง ๆ กันไป (ทรงศิริ แต้สมบัติ, 2539 ; Thomopoulos, 1980) ในที่นี้ผู้วิจัยจะกล่าวถึงตัวแปรอิสระใน 2 ความหมาย

<u>*กรณีที่ 1*</u> ตัวแปรอิสระเป็นตัวแปรเวลา (time variable) ตัวแปรนี้ใช้เมื่ออนุกรมเวลามีแนวโน้ม รูปแบบแนวโน้มมีหลายแบบทั้งนี้ขึ้นอยู่กับลักษณะของอนุกรมเวลา แต่รูปแบบแนวโน้มที่พบมากได้แก่

แนวโน้มเส้นตรง (linear trend) $$Z_t = \beta_0 + \beta_1 t + \varepsilon_t$$

แนวโน้มเส้นโค้ง (quadratic) $$Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$

แนวโน้มเอ็กซ์โปเนนเชียล (exponential) $$Z_t = \beta_0 \beta_1^t \varepsilon_t$$

การใช้ตัวแปรเวลาเป็นตัวแปรอิสระนั้น ไม่ว่ารูปแบบการถดถอยจะเป็นแบบใดต้องกำหนดค่าให้กับตัวแปรเวลา (t) ค่า t จะมีค่าเริ่มต้นเป็นเท่าใดก็ได้ แต่จะต้องมีการเพิ่มขึ้นเท่า ๆ กัน ในแต่ละช่วงเวลา

__กรณีที่ 2__ ตัวแปรอิสระเป็นตัวแปรดัมมี่ (dummy variable) เป็นตัวแปรที่กำหนดขึ้นเพื่อบอกว่าค่าสังเกตในอนุกรมเวลาเกิดขึ้นในฤดูกาลใด ค่าตัวแปรดัมมี่จะเป็น 1 เมื่อค่าสังเกตอยู่ในฤดูกาลที่กำหนด แต่ถ้าค่าสังเกตไม่อยู่ในฤดูกาลที่กำหนดค่าตัวแปรดัมมี่จะเป็น 0 จำนวนตัวแปรดัมมี่ที่สร้างขึ้นมีจำนวนเท่ากับ L-1 ตัว โดย L เป็นจำนวนฤดูกาลต่อปี นั่นคือ

$$D_{it} = \begin{cases} 1 & \text{เมื่อค่าสังเกต } Z_t \text{ อยู่ในฤดูกาลที่ } i \\ 0 & \text{เมื่อค่าสังเกต } Z_t \text{ ไม่อยู่ในฤดูกาลที่ } i \end{cases}$$

จึงเขียนโมเดลเชิงเส้นที่อธิบาย L ฤดูกาลได้ดังนี้

$$Z_t = \alpha + \beta_1 D_{1t} + \ldots + \beta_{L-1} D_{L-1,t} + \beta_L D_{Lt} + \varepsilon_t \qquad \text{.....................(55)}$$

เมื่อ α คือ ค่าจุดตัด (Z-intercept)

β_i คือ ค่าขนาดอิทธิพลฤดูกาลที่ i เมื่อเทียบกับฤดูกาลที่ L ; i = 1, 2, ...,L-1

β_L คือ ค่าขนาดอิทธิพลของฤดูกาลที่ L จะกำหนดให้มีค่าเป็น 0

ในกรณีที่อนุกรมเวลามีแนวโน้มและมีการรวมตัวกับฤดูกาลเป็นแบบบวก เขียนโมเดลได้ว่า

$$Z_t = \alpha + \beta t + \beta_1 D_{1t} + \ldots + \beta_{L-1} D_{L-1,t} + \varepsilon_t \qquad \text{...........................(56)}$$

ในกรณีที่อนุกรมเวลามีแนวโน้มและมีการรวมตัวกับฤดูกาลเป็นแบบคูณ ต้องแปลงให้อยู่ในรูปแบบบวก โดยการหาลอการิทึม (logarithm) ของ Z_t ค่าของ $\ln Z_t$ เขียนสมการได้ดังนี้

$$Z_t = \alpha \beta^t \beta_1^{D_{1t}} \ldots \beta_{L-1}^{D_{L-1t}} \varepsilon_t$$
$$\ln Z_t = \ln \alpha + (\ln \beta) t + (\ln \beta_1) D_{1t} + \ldots + (\ln \beta_{L-1}) D_{L-1t} + \ln \varepsilon_t$$
$$Z_t' = \alpha' + \beta' t + \beta_1' D_{1t} + \ldots + \beta_{L-1}' D_{L-1t} + \varepsilon_t' \quad \ldots\ldots(57)$$

ตอนที่ 3.2 วิธีการเฉลี่ยเคลื่อนที่

วิธีการเฉลี่ยเคลื่อนที่ เป็นวิธีการหาค่าพยากรณ์ที่มีการเฉลี่ยน้ำหนักของค่าสังเกตที่ผ่านมาแบบเลขคณิต โดยให้น้ำหนักแต่ละค่าสังเกตเท่ากัน เช่น กรณีที่ใช้ค่าสังเกต N ค่า จะได้ค่าพยากรณ์ ณ ช่วงเวลา t+1 ดังนี้ (Sullivan and Claycombe, 1977 ; Bowerman and O'Connell, 1993 ; Makridakis and Wheelwright , 1989 ; Pecar, 1994 ; ทรงศิริ แต้สมบัติ, 2539)

$$F_{t+1} = (Z_t + Z_{t-1} + \ldots + Z_{t-N+1}) / N \quad \ldots\ldots(58)$$

ค่าพยากรณ์ที่ได้จากสมการเป็นค่าพยากรณ์ที่เกิดจากการเฉลี่ยเคลื่อนที่ และเรียกการเฉลี่ยเคลื่อนที่ในลักษณะนี้ว่า การเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average)

การเฉลี่ยเคลื่อนที่แบบง่าย เป็นการขจัดอิทธิพลอย่างสุ่มออกจากอนุกรมเวลา วิธีการเฉลี่ยเคลื่อนที่ทำได้โดยการนำชุดของค่าสังเกตมาหาค่าเฉลี่ย แล้วใช้ค่าเฉลี่ยนั้นพยากรณ์ในช่วงเวลาถัดไป จำนวนค่าสังเกตที่จะนำมาหาค่าเฉลี่ย มีจำนวนไม่แน่นอน จำนวนที่เหมาะสม คือ จำนวนที่ทำให้ค่าพยากรณ์มีค่าใกล้เคียงกับความเป็นจริงมากที่สุด นั่นคือทำให้ค่า SSE, MSE หรือ RMSE มีค่าน้อยที่สุด เช่น ข้อมูลอนุกรมเวลาที่มีค่าสังเกตรวม 10 ค่า นำมาหาค่าพยากรณ์ ณ ช่วงเวลา t=11 โดยการเฉลี่ยเคลื่อนที่ 3 ช่วงเวลา และ 5 ช่วงเวลา ได้ผลดังตาราง 10

จากตาราง 10 จะเห็นได้ว่า ค่าพยากรณ์ที่ได้จากวิธีเฉลี่ยเคลื่อนที่ 5 ช่วงเวลา ให้ค่าพยากรณ์ที่ใกล้เคียงความเป็นจริงมากกว่าการเฉลี่ยเคลื่อนที่ 3 ช่วงเวลา เพราะการเฉลี่ยเคลื่อนที่ 5 ช่วงเวลามีค่า MSE น้อยกว่าวิธีเฉลี่ยเคลื่อนที่ 3 ช่วงเวลา ดังนั้นในการพยากรณ์จึงควรเลือกใช้การเฉลี่ยเคลื่อนที่ 5 ช่วงเวลา

ตาราง 10 การพยากรณ์ด้วยการเฉลี่ยเคลื่อนที่ 3 ช่วงเวลา และ 5 ช่วงเวลา

ช่วงเวลา t	ค่าสังเกต Z_t	การเฉลี่ยเคลื่อนที่ 3 ช่วงเวลา	การเฉลี่ยเคลื่อนที่ 5 ช่วงเวลา
1	18	-	-
2	20	-	-
3	17	-	-
4	9	18.3	-
5	31	15.3	-
6	16	19	19
7	22	18.7	18.6
8	13	23	19
9	29	17	18.2
10	17	21.3	22.2
11	-	19.7	19.4
MSE		87.9	40.1

เทคนิคการพยากรณ์ด้วยการเฉลี่ยเคลื่อนที่แบบง่าย สามารถเขียนสมการและสมการพยากรณ์ได้ดังนี้

$$M_t^* = (Z_t + Z_{t-1} + ... + Z_{t-N+1})/N \qquad \text{...(59)}$$

$$F_{t+1} = S_t = (Z_t + Z_{t-1} + ... + Z_{t-N+1})/N \qquad \text{...(60)}$$

เมื่อ M_t^* คือ การเฉลี่ยเคลื่อนที่แบบง่าย ณ ช่วงเวลา t

F_{t+1} คือ ค่าพยากรณ์ ณ ช่วงเวลา $t+1$

S_t คือ ค่าปรับเรียบ ณ ช่วงเวลา t

Z_t คือ ค่าสังเกต ณ ช่วงเวลา t

N คือ จำนวนช่วงเวลาที่ใช้ในการเฉลี่ย

การเฉลี่ยเคลื่อนที่แบบง่ายที่กล่าวมา จะใช้ได้ดีเมื่ออนุกรมเวลามีลักษณะเป็นอนุกรมคงที่ แต่ถ้าอนุกรมเวลาเป็นอนุกรมไม่คงที่ แสดงว่าอนุกรมนั้นต้องมีแนวโน้มเพิ่มขึ้นหรือลดลง วิธีการเฉลี่ยเคลื่อนที่แบบง่ายก็จะมีข้อจำกัด จึงมีการพัฒนาวิธีการเฉลี่ยเคลื่อนที่แบบใหม่ขึ้นมา เพื่อลดข้อจำกัดดังกล่าว และเรียกวิธีนี้ว่า การเฉลี่ยเคลื่อนที่สองครั้ง (double moving average) (Pecar, 1994)

การเฉลี่ยเคลื่อนที่สองครั้ง ทำให้การเคลื่อนไหวของอนุกรมเวลาในการปรับครั้งที่สองเรียบกว่าการเฉลี่ยเคลื่อนที่ครั้งแรก การเฉลี่ยเคลื่อนที่สองครั้งสามารถคำนวณได้จาก

$$M_t^{**} = (M_t^{*} + M^{*}_{t-1} + \dots + M^{*}_{t-N+1}) / N \qquad \text{.................................(61)}$$

เมื่อ M_t^{**} คือ การเฉลี่ยเคลื่อนที่ครั้งที่สอง ณ ช่วงเวลา t

M_t^{*} คือ การเฉลี่ยเคลื่อนที่ครั้งแรก ณ ช่วงเวลา t

เช่น ข้อมูลอนุกรมเวลาที่มีค่าสังเกตรวม 10 ค่า มีแนวโน้มเป็นเส้นตรง สามารถหาค่าพยากรณ์ด้วยวิธีการเฉลี่ยเคลื่อนที่สองครั้งขนาด 3×3 ได้ดังตาราง 11

ตาราง 11 การพยากรณ์ด้วยวิธีการเฉลี่ยเคลื่อนที่สองครั้ง ขนาด 3×3

ช่วงเวลา t	ค่าสังเกต Z_t	การเฉลี่ยเคลื่อนที่ครั้งแรก M_t^{*}	การเฉลี่ยเคลื่อนที่ครั้งที่สอง M_t^{**}	ค่าประมาณ a_t	ค่าประมาณ b_t	ค่าพยากรณ์ F_{t+1}
1	18	-	-	-	-	-
2	20	-	-	-	-	-
3	17	18.3	-	-	-	-
4	9	15.3	-	-	-	-
5	31	19	17.6	20.5	1.5	22.0
6	17	18.7	17.7	19.7	1.0	20.7
7	22	23	20.2	25.8	2.8	28.6
8	13	17	19.6	14.5	-2.6	11.9
9	29	21.3	20.4	22.2	0.9	23.1
10	17	19.7	19.3	20	0.3	20.3

จากตารางสามารถสร้างสมการพยากรณ์ในรูปสมการเชิงเส้นได้ว่า

$$F_{t+T} = a_t + b_t T \qquad \text{.................................(62)}$$

โดยหาค่าพารามิเตอร์ a และ b ได้จาก

$$a_t = 2M_t^{*} - M_t^{**}$$

$$b_t = (2/N-1)(M_t^{*} - M_t^{**})$$

เมื่อ T คือ ช่วงเวลาล่วงหน้าที่จะพยากรณ์

a_t คือ ค่าประมาณแนวโน้ม (β_0) ณ ช่วงเวลา t

b_t คือ ค่าประมาณของ β_1 ณ ช่วงเวลา t

ในกรณีที่อนุกรมเวลามีแนวโน้มเป็นแบบเอ็กซ์โปเนนเชียล ลักษณะของแนวโน้มมีการเปลี่ยนแปลงเพิ่มขึ้นหรือลดลงเป็นเปอร์เซนต์ที่เท่ากันในแต่ละช่วงเวลา ต้องใช้วิธีเฉลี่ยเคลื่อนที่ของเปอร์เซนต์การเปลี่ยนแปลง (ทรงศิริ แต้สมบัติ, 2539) และเขียนสมการพยากรณ์ ณ ช่วงเวลา t ได้ว่า

$$F_{t+T} = Z_t\,(1+(MPC_t/100))^T \qquad \text{..............................(63)}$$

เมื่อ PC_t = อัตราการเปลี่ยนแปลง ณ ช่วงเวลา t

= $(Z_t - Z_{t-1})/Z_{t-1} \times 100$

= $(Z_t/Z_{t-1})\,100$

MPC = การเฉลี่ยเคลื่อนที่ N ช่วงเวลา ณ ช่วงเวลา t

= $\sum(PC_{t-i+1}/N)$

เช่น ข้อมูลอนุกรมเวลาที่มีค่าสังเกตรวม 8 ค่า มีลักษณะการเคลื่อนไหวเป็นแนวโน้มแบบเอ็กซ์โปเนนเชียล เมื่อทำการเฉลี่ยเคลื่อนที่ของเปอร์เซนต์การเพิ่มขึ้นหรือลดลงของค่าสังเกตใน 4 ช่วงเวลา (N=4) ได้ผลดังตาราง 12

ตาราง 12 การพยากรณ์ด้วยวิธีการเฉลี่ยเคลื่อนที่ของเปอร์เซ็นต์การเปลี่ยนแปลง

ช่วงเวลา t	ค่าสังเกต Z_t	ผลต่าง ∇Z_t	อัตราการเปลี่ยนแปลง PC_t	การเฉลี่ยเคลื่อนที่ MPC_t	ค่าพยากรณ์ F_{t+1}
1	21	-	-	-	-
2	18	-3	-14.3	-	-
3	15	-3	-16.7	-	-
4	17	2	3.3	-	-
5	11	-6	-35.3	-15.7	9.3
6	13	2	8.2	-10.1	11.7
7	10	-3	-23.1	-11.7	8.8
8	9	-1	-10	-15.1	7.7

จากตารางจะเห็นว่า ณ ช่วงเวลา $t=8$ หาการเฉลี่ยเคลื่อนที่ของเปอร์เซ็นต์การเปลี่ยนแปลงได้เป็น -15.1 แสดงว่าค่าสังเกตมีค่าลดลงช่วงเวลาละ 15.1 เปอร์เซ็นต์ โดยใช้ค่าสังเกต ณ ช่วงเวลา 8 มี Z_8 เป็นฐาน ได้สมการพยากรณ์ คือ

$$F_{8+T} = Z_8 (1 - 0.151)^T \qquad \text{.............................(64)}$$

เป็นที่น่าสังเกตว่า ถ้าการเฉลี่ยเคลื่อนที่ของเปอร์เซ็นต์การเปลี่ยนแปลงมีค่าต่ำกว่า 0 แนวโน้มแบบเอ็กซ์โปเนนเชียล จะมีแนวโน้มในทางลง แต่ถ้าการเฉลี่ยเคลื่อนที่ของเปอร์เซ็นต์การเปลี่ยนแปลงมีค่ามากกว่า 0 แนวโน้มแบบเอ็กซ์โปเนนเชียล จะมีแนวโน้มในทางขึ้น

จากที่กล่าวมาว่าการเฉลี่ยเคลื่อนที่ เป็นการให้น้ำหนักของค่าสังเกตในการเฉลี่ยเท่ากันหมด แต่ค่าสังเกตของอนุกรมเวลาส่วนใหญ่จะมีสารสนเทศ (information) ที่ใกล้ช่วงเวลา t มากกว่าค่าสังเกตที่ห่างออกไปจากช่วงเวลา t คือ $< Z_{t-M+1} < ... < Z_{t-1} < Z_t$ จึงเกิดแนวคิดเกี่ยวกับการเฉลี่ยเคลื่อนที่ชนิดพิเศษขึ้น เรียกว่า การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล (Pocar, 1994) ซึ่งจะกล่าวถึงรายละเอียดในตอนที่ 3.3

ตอนที่ 3.3 การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล

การปรับให้เรียบแบบเอ็กซ์โปเนนเชียล เป็นการเฉลี่ยน้ำหนักของค่าสังเกต โดยให้น้ำหนักของค่าสังเกตที่ช่วงเวลา t มาก และให้น้ำหนักของค่าสังเกตก่อนหน้าช่วงเวลา t ลดลงเรื่อย ๆ แบบเรขาคณิตหรือให้น้ำหนักลดหลั่นกันไปแบบเอ็กซ์โปเนนเชียล (Newbold and Bos, 1994 ; Sullivan and Claycombe, 1977 ; Bowerman and O'Connell, 1993 ; Makridakis, Wheelwright and McGee, 1978; Pecar, 1994 ; ทรงศิริ แต้สมบัติ, 2539)

โดยทั่วไปเราพยากรณ์ค่าสังเกตตัวใหม่ (new forecast) ได้จากความถูกต้องของการพยากรณ์ตัวที่ผ่านมา และการวัดความแตกต่างระหว่างค่าที่เกิดขึ้นจริงกับค่าพยากรณ์ (Pecar, 1994) นั่นคือ

ค่าพยากรณ์ตัวใหม่ = ค่าพยากรณ์ตัวเก่า + ความคลาดเคลื่อน ....(65)

(new forecast) (old forecast) (error fraction)

ค่าความคลาดเคลื่อน (error fraction) คือส่วนหนึ่งของความคลาดเคลื่อนที่เกิดขึ้นจริงเท่านั้น เมื่อคูณค่าความคลาดเคลื่อนที่เกิดขึ้นจริงด้วยค่าคงที่ค่าหนึ่ง ที่มีค่าระหว่าง 0 ถึง 1 และเรียกค่าคงที่ตัวนี้ว่า ค่าคงที่ของการปรับน้ำหนัก (smoothing constant) หรือ α ดังนั้นจากสมการ (65) สามารถอธิบายในรูปสมการคณิตศาสตร์ได้ดังนี้

ให้ F_{t+1} คือ ค่าพยากรณ์ในช่วงเวลาถัดไป

F_t คือ ค่าพยากรณ์ในช่วงเวลาปัจจุบัน (ค่าเริ่มต้น)

Z_t คือ ค่าสังเกตที่เกิดขึ้นจริงในช่วงเวลาปัจจุบัน

α คือ ค่าคงที่ของการปรับน้ำหนัก

$$F_{t+1} = F_t + \alpha(Z_{t-1} - F_t)$$

$$F_{t+1} = \alpha Z_t + (1-\alpha)F_t$$

$$F_t = \alpha Z_{t-1} + (1-\alpha)F_{t-1} \qquad \text{.............................(66)}$$

การเขียนสมการพยากรณ์แบบเอ็กซ์โปเนนเชียล สามารถเขียนได้ 2 แบบ คือ

1. แบบปรับน้ำหนัก (smoothing form)

$$F_{t+1} = \alpha Z_t + (1-\alpha)F_t$$

2. แบบปรับค่าเดิมด้วยความคลาดเคลื่อน (error correction form)

$$F_{t+1} = F_t + \alpha e_t$$

แต่ในงานวิจัยครั้งนี้ผู้วิจัยนำเสนอเฉพาะสมการแบบปรับน้ำหนักเท่านั้น เพราะการคำนวณไม่ว่าจะใช้สมการปรับค่าแบบใดจะให้ผลไม่ต่างกัน

จากสมการ (66) อิทธิพลของค่าคงที่ในการปรับน้ำหนัก (α) จะลดลงแบบเอ็กซ์โปเนนเชียล และเมื่อรวมน้ำหนักทั้งหมดแล้วจะต้องไม่เกิน 1 การกำหนด α ที่เหมาะสมกับอนุกรมเวลา ต้องดูที่ลักษณะของอนุกรมเวลา ถ้าอนุกรมเวลาเป็นอนุกรมคงที่ ควรกำหนด α เท่ากับ 0.1 - 0.3 (Pecar, 1994) แต่อย่างไรก็ตาม นักวิจัยไม่จำเป็นต้องกำหนด α เอง เพราะโปรแกรมสำเร็จรูปสามารถกำหนด α ที่เหมาะสมได้ ซึ่งค่า α ที่เหมาะสมจะทำให้ SSE, MSE หรือ RMSE มีค่าน้อยที่สุด การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่ใช้สำหรับอนุกรมเวลาคงที่ที่กล่าวมาในเบื้องต้นเรียกว่า การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing)

ในกรณีที่อนุกรมเวลาเป็นอนุกรมไม่คงที่ การสร้างสมการพยากรณ์จะมีลักษณะคล้ายคลึงกับวิธีการเฉลี่ยเคลื่อนที่สองครั้ง คือ ต้องใช้วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสองครั้ง (double exponential smoothing) และการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสามครั้ง (triple exponential smoothing)

จากสมการ (66) ถ้าให้ $F_{t+1} = S_t^*$ และ $F_t = S_{t-1}^*$ สามารถเขียนสมการเอ็กซ์โปเนนเชียลได้ดังนี้

1. การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว เป็นวิธีการที่ใช้กับอนุกรมเวลาที่ไม่มีแนวโน้มและฤดูกาล

$$S_t^* = \alpha Z_t + (1-\alpha) S_{t-1}^* \qquad \text{.......................(67)}$$

2. การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสองครั้ง เป็นวิธีการที่ใช้กับอนุกรมเวลาที่มีแนวโน้มแบบเส้นตรง

$$S_t^{**} = \alpha S_t^* + (1-\alpha) S_{t-1}^{**} \qquad \text{.......................(68)}$$

สมการนี้มีลักษณะเช่นเดียวกับสมการการเคลื่อนที่สองครั้ง คือ มีพารามิเตอร์ a และ b

$$a_t = 2S_t^* - S_t^{**}$$

$$b_t = (\alpha/(1-\alpha))(S_t^* - S_t^{**})$$

จะได้สมการพยากรณ์ คือ

$$F_{t+T} = a_t + b_tT \qquad \text{.........................(69)}$$

3. การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสามครั้ง เป็นวิธีการที่ใช้กับอนุกรมเวลาที่มีแนวโน้มแบบเส้นโค้ง (quadratic)

$$S_t^{***} = \alpha S_t^{**} + (1-\alpha) S_{t-1}^{***} \qquad \text{............................(70)}$$

สมการนี้จะคล้ายกับสมการการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสองครั้ง แต่ต่างกันที่จำนวนพารามิเตอร์ คือ สมการการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสามครั้งมีพารามิเตอร์ 3 ตัว คือ a, b และ c

$$a_t = 3S_t^* - 3S_t^{**} + S_t^{***}$$

$$b_t = \{\alpha/(2(1-\alpha)^2)\}\{(6-5\alpha)S_t^* - 2(5-4\alpha)S_t^{**} + (4-3\alpha)S_t^{***}\}$$

$$c_t = \{\alpha^2/(2(1-\alpha)^2)\}\{S_t^* - 2S_t^{**} + S_t^{***}\}$$

จะได้สมการพยากรณ์ คือ

$$F_{t+T} = a_t + b_tT + c_tT^2 \qquad \text{..............................(71)}$$

การสร้างสมการพยากรณ์ จะเกี่ยวข้องกับค่าพารามิเตอร์เริ่มต้น a_0, b_0, c_0 จากอนุกรมเวลาที่มีอยู่ โดยใช้ค่าสังเกตเพียงบางค่า แล้วใช้วิธีกำลังสองน้อยที่สุด จากนั้นหา S_0^*, S_0^{**} และ S_0^{***}

$$S_0^* = a_0 - \{(1-\alpha)/\alpha\}b_0 + \{((1-\alpha)(2-\alpha))/a^2\}c_0$$

$$S_0^{**} = a_0 - \{2(1-\alpha)/\alpha\}b_0 + \{(2(1-\alpha)(3-2\alpha))/a^2\}c_0$$

$$S_0^{***} = a_0 - \{3(1-\alpha)/\alpha\}b_0 + \{(3(1-\alpha)(4-3\alpha))/a^2\}c_0$$

เมื่อได้ S_0^*, S_0^{**}, S_0^{***} แล้วใช้หา S_t^*, S_t^{**} และ S_t^{***} สำหรับค่า t ต่อ ๆ ไปได้

4. การปรับให้เรียบเอ็กซ์โปเนนเชียลแบบโฮลท์-วินเทอร์ (Holt - Winters) เป็นวิธีวิเคราะห์อนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากแนวโน้มและอิทธิพลของฤดูกาล รูปแบบมีทั้งแบบบวกและแบบคูณ หลักการของวิธีการปรับให้เรียบเอ็กซ์โปเนนเชียลแบบโฮลท์-วินเทอร์

มีลักษณะเช่นเดียวกับวิธีการปรับให้เรียบเอ็กซ์โปเนนเชียลแบบอื่น ๆ แต่วิธีการปรับให้เรียบเอ็กซ์โปเนนเชียลแบบโฮลท์-วินเทอร์ มีค่าคงที่ในการปรับน้ำหนัก 3 ค่า คือ

α เป็นค่าคงที่ของการปรับน้ำหนักสำหรับแนวโน้ม (trend)

γ เป็นค่าคงที่ของการปรับน้ำหนักสำหรับความชัน (slope)

δ เป็นค่าคงที่ของการปรับน้ำหนักสำหรับฤดูกาล (season)

ค่า α, γ, δ มีค่าอยู่ระหว่าง 0 ถึง 1

ถ้าให้ Z_t เป็นค่าสังเกตของอนุกรมเวลา

S_t เป็นค่าประมาณแนวโน้ม

b_t เป็น slope ณ ช่วงเวลา t

I_t เป็นค่าประมาณองค์ประกอบของฤดูกาล ณ ช่วงเวลา t

s เป็นจำนวนฤดูกาล

จะได้สมการพยากรณ์ในรูปแบบบวก คือ

$$F_{t+T} = S_t + b_tT + F_{t+T-s} \qquad \text{.............................(72)}$$

เมื่อ $S_t = \alpha(Z_t - I_{t-s}) + (1-\alpha)(S_{t-1} + b_{t-1})$

$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\, b_{t-1}$

$I_t = \delta(Z_t - S_t) + (1-\delta)\, I_{t-s}$

และได้สมการพยากรณ์ในรูปแบบคูณ คือ

$$F_{t+T} = (S_t + b_tT)\, I_{t+T-s} \qquad \text{.................................(73)}$$

เมื่อ $S_t = \alpha(Z_t / I_{t-s}) + (1-\alpha)(S_{t-1} + b_{t-1})$

$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\, b_{t-1}$

$I_t = \delta(Z_t / S_t) + (1-\delta)\, I_{t-s}$

5. การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่แนวโน้มค่อย ๆ ลดลง (damped trend exponential smoothing) ใช้ในกรณีที่อนุกรมเวลามีแนวโน้มค่อย ๆ ลดลง จะได้สมการพยากรณ์ดังนี้

$$F_{t+T} = S_t + \sum \phi b_t T \qquad \text{................................(74)}$$

เมื่อ $S_t = \alpha Z_t + (1-\alpha)(S_{t-1} + \phi b_{t-1})$

$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\,\phi\, b_{t-1}$

$0 < \phi \leq 1$; ϕ คือ องค์ประกอบที่ลดลง (damping factor)

6. การปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่แนวโน้มเอ็กซ์โปเนนเชียล (exponential trend exponential smoothing) ใช้ในกรณีที่อนุกรมเวลามีแนวโน้มแบบเอ็กซ์โปเนนเชียล จะได้สมการพยากรณ์ดังนี้

$$F_{t+T} = S_t (b_t)^T \qquad \text{.................................(75)}$$

เมื่อ $S_t = \alpha Z_t + (1-\alpha) S_{t-1} b_{t-1}$

$b_t = \gamma (S_t / S_{t-1}) + (1-\gamma) b_{t-1}$

วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลมีอยู่หลายวิธี นักวิจัยควรเลือกใช้ให้เหมาะสมกับลักษณะการเคลื่อนไหวของข้อมูลอนุกรมเวลาชุดนั้น ๆ ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 วิธี โดยคำนึงถึงลักษณะการเคลื่อนไหวของข้อมูล กล่าวคือ ข้อมูลปริมาณการยืมหนังสือทั่วไป หนังสือสำรอง และวิทยานิพนธ์ เป็นข้อมูลอนุกรมเวลาที่มีแนวโน้มแบบเส้นตรงและมีการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยจึงเลือกใช้วิธีการปรับให้เรียบเอ็กซ์โปเนนเชียลแบบโฮลท์-วินเทอร์ ข้อมูลจำนวนนักเรียนระดับประถมศึกษา มีการเปลี่ยนแปลงที่แสดงแนวโน้มแบบเส้นโค้ง ผู้วิจัยจึงเลือกใช้วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสามครั้ง ส่วนข้อมูลจำนวนนักเรียนระดับมัธยมศึกษามีการเปลี่ยนแปลงที่แสดงแนวโน้มแบบเอ็กซ์โปเนนเชียล ผู้วิจัยจึงเลือกใช้วิธีการปรับให้เรียบแบบเอ็กซ์โปเนนเชียลที่มีแนวโน้มเอ็กซ์โปเนนเชียล

ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

สาระในตอนนี้ ผู้วิจัยนำเสนองานวิจัยที่เกี่ยวข้องกับการพยากรณ์ทางการศึกษาและการพยากรณ์ที่ใช้วิธีบ็อกซ์และเจ็นกิ้นส์ ดังนี้

ในปี พ.ศ. 2513 ณัฐฑิณา วัฒนายากูล ได้พยากรณ์ความต้องการครูระดับประถมศึกษาของประเทศไทยระหว่างปีการศึกษา 2515-2519 จากการเพิ่มของจำนวนนักเรียน โดยใช้ข้อมูลจำนวนนักเรียนระดับประถมศึกษาของประเทศไทยระหว่างปีการศึกษา 2508-2513 และจำนวนประชากรอายุ 7 ปีบริบูรณ์ ตั้งแต่ พ.ศ. 2509-2519 ซึ่งได้จากการคำนวณมาพยากรณ์จำนวนนักเรียนในอนาคต โดยใช้อัตราส่วนแนวโน้มจำนวนนักเรียนในอดีต แล้วนำมาพยากรณ์จำนวนครูที่ต้องการจากจำนวนนักเรียนที่พยากรณ์ได้โดยใช้อัตราส่วนจำนวนนักเรียนต่อครูของกระทรวงศึกษาธิการ คือ 30 : 1 ในชั้นประถมศึกษาตอนต้น และ 25 : 1 ในชั้น

ประถมศึกษาตอนปลาย อัตราส่วนต่อครูของยูเนสโก คือ 38 : 1 และอัตราส่วนจำนวนนักเรียนต่อครูที่เป็นจริงในปัจจุบันจากผลของการวิจัยเป็น 32 : 7 : 1 ทั้งในระดับประถมศึกษาตอนต้นและตอนปลาย และผลการวิจัยพบว่า การพยากรณ์ความต้องการครูโดยใช้อัตราส่วนดังกล่าว ได้ตัวเลขจำนวนครูสูงสุดไปหาต่ำสุดเรียงตามลำดับคือ อัตราส่วนของกระทรวงศึกษาธิการ อัตราส่วนที่เป็นจริงในปัจจุบันจากผลการวิจัย และ อัตราส่วนของยูเนสโก (ณัดดิยา วัฒนาชยากูล, 2513)

ต่อมาในปี พ.ศ. 2518 นภาพร สิงหทัต ได้ทำการพยากรณ์ความต้องการครูในระดับประถมศึกษาเช่นเดียวกับ ณัดดิยา วัฒนาชยากูล แต่พยากรณ์จำนวนนักเรียนจากประชากรอายุ 5-13 ปี ประมาณจำนวนนักเรียนประถมศึกษาปีที่ 1 โดยใช้ร้อยละของประชากรที่เข้าเรียน และพยากรณ์จำนวนนักเรียนในชั้นอื่น ๆ และพยากรณ์ความต้องการครูในอนาคตใช้วิธีเดียวกับงานวิจัยของ ณัดดิยา วัฒนาชยากูล ผลการวิจัยพบว่า การพยากรณ์ความต้องการครูโดยใช้อัตราส่วนจำนวนนักเรียนต่อครู ได้ตัวเลขจำนวนครูสูงสุดไปหาต่ำสุดเรียงตามลำดับ คือ อัตราส่วนของกระทรวงศึกษาธิการ (30 : 1) อัตราส่วนที่เป็นจริงในปัจจุบันจากผลการวิจัย (30 : 16 : 1) และ อัตราส่วนของยูเนสโก (40 : 1) แต่อย่างไรก็ตาม ความต้องการครูในอนาคตมีจำนวนลดลง ซึ่งเป็นเครื่องชี้ว่า มาตรฐานอัตราส่วนจำนวนนักเรียนต่อครูควรลดลง จะทำให้ครูสามารถดูแลนักเรียนได้อย่างทั่วถึง (นภาพร สิงหทัต, 2518)

ในปี พ.ศ. 2521 วรรณพร วิเชียรวงศ์ ทำการพยากรณ์ความต้องการครูระดับประถมศึกษาและมัธยมศึกษาของอำเภอธัญญบุรี จังหวัดปทุมธานี ในปีการศึกษา 2521-2526 โดยใช้ประชากรที่เกิดในปี พ.ศ.2509-2520 มาปรับด้วยอัตราความครบถ้วนของการจดทะเบียนและอัตราการตายตามหมวดอายุ เพื่อให้ได้ประชากรที่แท้จริง แล้วพยากรณ์จำนวนนักเรียนในอนาคต โดยใช้อัตราการเข้าเรียนและอัตราส่วนแนวโน้มของจำนวนนักเรียน แล้วจึงนำมาพยากรณ์จำนวนครูที่ต้องการ โดยใช้อัตราส่วนจำนวนนักเรียนต่อครู ผลการวิจัยพบว่า ในปีการศึกษา 2521-2525 จำนวนครูประถมศึกษาและมัธยมศึกษาที่ต้องการมีจำนวนลดลง ส่วนในปีการศึกษา 2526 จำนวนครูที่ต้องการเพิ่มขึ้น (วรรณพร วิเชียรวงศ์, 2521)

ต่อมาในปี พ.ศ. 2525 กานต์ กุณาศล พยากรณ์ความต้องการครูโรงเรียนประถมศึกษาของจังหวัดกาญจนบุรี ในปีการศึกษา 2524-2529 โดยพยากรณ์ความต้องการครูจาก 3 โมเดล โมเดลแรกจะหาจำนวนนักเรียนจากประชากรในโรงเรียนทั้งหมด ตั้งแต่ พ.ศ.2524-

2529 โมเดลที่ 2 จะหาจำนวนนักเรียนจากประชากรวัยเรียนทั้งหมดที่จะเข้าเรียนชั้นประถมศึกษาปีที่ 1 ตั้งแต่ พ.ศ. 2524-2529 เป็นรายอำเภอ แล้วนำมาพยากรณ์ความต้องการครูจากเกณฑ์ ครู : นักเรียน (1 : 25) ส่วนโมเดลที่ 3 พยากรณ์ความต้องการครูเป็นรายโรงเรียน โดยจะพยากรณ์จากขนาดของโรงเรียน ผลการวิจัยพบว่า โมเดลที่ 3 ต้องการครูจำนวนมากที่สุด ส่วนโมเดลที่ 2 ต้องการครูน้อยที่สุด (กานต์ กุณาศล, 2525)

จะเห็นได้ว่า งานวิจัยที่กล่าวมาทั้ง 4 เรื่อง เป็นการพยากรณ์ความต้องการครูโดยใช้เกณฑ์อัตราส่วนนักเรียนต่อครู การพยากรณ์จำนวนนักเรียนใช้อัตราส่วนแนวโน้มจำนวนนักเรียน อัตราการเข้าเรียน จำนวนประชากรอายุ 7 ปีและ อายุ 5-13 ปี นอกจากนี้ยังมีการพยากรณ์ความต้องการอัตรากำลังคนในสาขาต่าง ๆ ของประเทศไทยโดยใช้ สูตรเศรษฐมิติทางการศึกษา เช่น ในปี พ.ศ. 2513 นงลักษณ์ วิรัชชัย (2513) นิตยา ภัสสรศิริ (2513) บุญธรรม กิจปรีดาบริสุทธิ์ (2513) พรรณมาศ คันฉาย (2513) และ สมหวัง พิธิยานุวัฒน์ (2513) ศึกษาสูตรเศรษฐมิติทางการศึกษาของทินเบอร์เกน แสดงความสัมพันธ์ระหว่างพื้นฐานการพัฒนาทางเศรษฐกิจและระบบการศึกษาของประเทศ ในรูปความสัมพันธ์มหภาค และผู้วิจัยศึกษาเปรียบเทียบงานวิจัยของนักวิจัยที่กล่าวมา ดังรายละเอียดในตาราง 13

ตาราง 13 การเปรียบเทียบงานวิจัยที่ใช้สูตรเศรษฐมิติทางการศึกษาของทินเบอร์เกน

นักวิจัย	วัตถุประสงค์	ข้อมูลที่ใช้	จำนวนค่าสัมประสิทธิ์ที่หา
นงลักษณ์ วิรัชชัย	หาสูตรเศรษฐมิติทางการศึกษามาใช้เพื่อการเร่งรัดพัฒนาประเทศไทยโดยไม่ต้องรับความช่วยเหลือจากต่างประเทศ	ใช้ข้อมูลทางการศึกษา กำลังคน และผลิตภัณฑ์รวมภายในประเทศระหว่าง พ.ศ. 2502-2511	12 ค่า

ตาราง 13 (ต่อ)

นักวิจัย	วัตถุประสงค์	ข้อมูลที่ใช้	จำนวนค่าสัมประสิทธิ์ที่หา
พรรณมาศ กันฉาย	หาสูตรเศรษฐมิติทางการศึกษามาใช้เพื่อการเร่งรัดพัฒนาประเทศไทยโดยต้องรับความช่วยเหลือจากต่างประเทศ	ใช้ข้อมูลทางการศึกษา กำลังคน และผลิตภัณฑ์รวมภายในประเทศระหว่าง พ.ศ. 2503-2511	6 ค่า
บุญธรรม กิจปรีดาบริสุทธิ์	หาสูตรเศรษฐมิติทางการศึกษาและกระสวนความเจริญที่สมดุลย์สำหรับประเทศไทย	ใช้ข้อมูลทางการศึกษา กำลังคน และผลิตภัณฑ์รวมภายในประเทศระหว่าง พ.ศ. 2503-2511	9 ค่า
นิตยา ภัสสรศิริ	หาสูตรเศรษฐมิติทางการศึกษาเพื่อประเมินกำลังคนในสาขาวิทยาศาสตร์และสาขาอื่นของประเทศไทย	ใช้ข้อมูลทางการศึกษา กำลังคน และผลิตภัณฑ์รวมภายในประเทศระหว่าง พ.ศ. 2503-2512	12 ค่า
สมหวัง พิธิยานุวัฒน์	หาสูตรเศรษฐมิติทางการศึกษาเพื่อประเมินกำลังคนในสาขาเกษตรกรรม อุตสาหกรรม และบริการ ในประเทศไทย	ใช้ข้อมูลทางการศึกษา กำลังคน และ ผลิตภัณฑ์รวมภายในประเทศระหว่าง พ.ศ. 2503-2511	6 ค่า

จากผลการวิจัย นักวิจัยสามารถหาสูตรเศรษฐมิติทางการศึกษาได้ดังตาราง 14

ตาราง 14 สูตรเศรษฐมิติทางการศึกษา

นักวิจัย	สูตรเศรษฐมิติที่หาได้
นงลักษณ์ วิรัชชัย	สูตรเศรษฐมิติพื้นฐานทางการศึกษา $N_t^2 = 1.887\ V_t$ $N_t^2 = (1-0.074)\,N_{t-1}^2 + m_t^2$ $m_t^2 = n_{t-1}^2 - n^3$ $m_t^3 = n_{t-1}^3$ $N_t^3 = (1-0.103)\ N_{t-1}^3 + m_t^3$ $N_t^3 = 0.785\,V_t + 0.059\,n_t^2 + 0.094\,n_t^3$ สูตรเศรษฐมิติทางการศึกษาภาคขยาย $N_t^2 = 1.887\ V_t$ $N_t^2 = (1-0.074)\,N_{t-1}^2 + m_t^2$ $m_t^2 = 0.997\ (0.36\,n_{t-1}^2 + 0.124\,n_{t-1}^3 - 1.239\,n_t^3)$ $m_t^3 = 0.998\ (1.136\,)n_{t-1}^3$ $N_t^3 = (1-0.103)\ N_{t-1}^3 + m_t^3$ $N_t^3 = 0.785\,V_t + 0.059\,n_t^2 + 0.094\,n_t^3$
พรรณมาศ คันฉาย	สูตรเศรษฐมิติพื้นฐานทางการศึกษา $N_t^2 = 0.038\ V_t$ $N_t^2 = 0.9256\,N_{t-1}^2 + m_t^2$ $m_t^2 = n_{t-1}^2 - n^3$ $m_t^3 = n_{t-1}^3$ $N_t^3 = 0.8982\ N_{t-1}^3 + m_t^3$ $N_t^3 = 0.0171\,V_t + 0.0589\,n_t^2 + 0.0916\,n_t^3$
บุญธรรม กิจปรีดาบริสุทธิ์	สูตรเศรษฐมิติทางการศึกษาภาคขยาย $N_t^2 = 1.4130\ V_t$ $N_t^2 = 0.9359\,N_{t-1}^2 + m_t^2$ $m_t^2 = 0.3561n_{t-1}^2 + 0.1465\,n_{t-1}^3 - n_t^3$ $m_t^3 = 0.8535\ n_{t-1}^3$ $N_t^3 = 0.9308\ N_{t-1}^3 + m_t^3$ $N_t^3 = 0.5726\,V_t + 0.0578\,n_t^2 + 0.0880\,n_t^3$

ตาราง 14 (ต่อ)

นักวิจัย	สูตรเศรษฐมิติที่หาได้
นิตยา ภัสสรศิริ	สูตรเศรษฐมิติทางการศึกษาภาคขยาย $N_t^{2} = 1.942\ V_t$ $N_t^{2} = 0.926\, N_{t-1}^{2} + m_t^{2}$ $m_t^{2} = 0.360 n_{t-1}^{2} + 0.135\, n_{t-1}^{3} + 0.082\, n_{t-1}^{3s2} - n_t^{2}$ $m_t^{3s1} = 0.865\, n_{t-1}^{3s1}$ $m_t^{3s2} = 0.918\, n_{t-1}^{3s2}$ $N_t^{3s1} = 0.909\ N_{t-1}^{3s1} + m_t^{3s1}$ $N_t^{3s2} = 0.909\ N_{t-1}^{3s2} + m_t^{3s2}$ $N_t^{3s1} = 0.508\ V_t^{s1}$ $N_t^{3s2} = 0.937\, V_t^{s2} + 0.059\, n_t^{2} + 0.094\, n_t^{3}$ $n_t^{3} = n_t^{3s1} + n_t^{3s2}$
สมหวัง พิธิยานุวัฒน์	สูตรเศรษฐมิติทางการศึกษาภาคขยาย $N_t^{2a} = 2.13087\ V_t^{a}$ $N_t^{2i} = 1.23530\, V_t^{i}$ $N_t^{2s} = 2.45006\, V_t^{s}$ $N_t^{2} = N_t^{2a} + N_t^{2i} + N_t^{2s}$ $N_t^{2} = 0.9304\, N_{t-1}^{2} + m_t^{2}$ $m_t^{2} = 0.3561\, n_{t-1}^{2} + 0.1465\, n_{t-1}^{3} - n_t^{3}$ $m_t^{3} = 0.8535\, n_{t-1}^{3}$ $N_t^{3} = 0.9087\ N_{t-1}^{3} + m_t^{3}$ $N_t^{3a} = 0.54463\ V_t^{a}$ $N_t^{3i} = 0.44715\ V_t^{i}$ $N_t^{3s} = 2.32394\, V_t^{s} + 0.0586\, n_t^{2} + 0.0912\, n_t^{3}$ $N_t^{3} = N_t^{3a} + N_t^{3i} + N_t^{3s}$

เมื่อ N เป็นสต็อคกำลังคน

V เป็นผลิตภัณฑ์รวมภายในประเทศ

m เป็นผู้เข้าสู่กำลังแรงงานใหม่

n เป็นจำนวนนักเรียน

2 เป็นดัชนีบอกระดับมัธยมศึกษา

3 เป็นดัชนีบอกระดับอุดมศึกษา

a เป็นดัชนีบอกสาขาเกษตรกรรม

i เป็นดัชนีบอกสาขาอุตสาหกรรม

s เป็นดัชนีบอกสาขาบริการ

s1 เป็นดัชนีบอกสาขาวิทยาศาสตร์

s2 เป็นดัชนีบอกสาขาอื่น ๆ

t เป็นดัชนีบอกคาบเวลา

นอกจากนี้ยังมีการพยากรณ์ความต้องการครู โดยใช้สูตรเศรษฐมิติหรือโมเดลเศรษฐมิติ ศึกษาความสัมพันธ์เชิงสาเหตุของตัวแปร เช่น งานวิจัยของ อังคณา พัฒนผลไพบูลย์ (2531) และ นงนุช อินทรวงษ์โชติ (2538) ซึ่งมีรายละเอียดดังนี้

ในปี พ.ศ. 2531 อังคณา พัฒนผลไพบูลย์ สร้างโมเดลทางเศรษฐมิติเพื่อพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ปีการศึกษา 2531-2540 โดยดำเนินการตามขั้นตอน 4 ขั้นตอน คือขั้นแรก กำหนดโมเดลในการพยากรณ์จำนวนครู โดยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง เพื่อคัดเลือกตัวแปรที่มีความสัมพันธ์กับจำนวนครู จากการศึกษาได้ตัวแปร 11 ตัว คือ จำนวนนักเรียน ผลิตภัณฑ์ประชาชาติภายในประเทศ หลักสูตรหรืออัตราส่วนวิชาชีพต่อวิชาสามัญ เทคโนโลยีทางการศึกษา จำนวนห้องเรียน ขนาดของโรงเรียน อัตราการสอนของครู จำนวนคาบที่เรียนต่อสัปดาห์ อัตราส่วนจำนวนนักเรียนต่อครู อัตราส่วนจำนวนนักเรียนต่อห้องเรียน และอัตราการปลดเกษียณ เก็บรวบรวมข้อมูลจำนวน 16 ปี (2515-2530) จากแหล่งทุติยภูมิ มาทำแผนภาพกระจัดกระจาย เพื่อหาลักษณะความสัมพันธ์ของตัวแปรอธิบายกับตัวแปรตาม คำนวณหาสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอธิบายและตัวแปรตาม พิจารณาปัญหาความสัมพันธ์ภาวะร่วมเส้นตรงพหุ (multicollinearity) ของตัวแปรอธิบาย โดยตัดตัวแปรที่มีความสัมพันธ์อย่างเด่นชัดออก เลือกตัวแปรที่สำคัญกว่าไว้ 5 ตัว ซึ่งเป็นตัวแปรอธิบายที่เหมาะสมในสมการ ขั้นที่ 2 ประมาณค่าพารามิเตอร์ในโมเดลหรือค่าสัมประสิทธิ์ของโมเดล ขั้นที่ 3 ประเมินค่าที่ประมาณได้ของพารามิเตอร์โดยใช้ F-test เมื่อมีนัยสำคัญแล้วทดสอบค่าพารามิเตอร์ของตัวแปรอธิบายแต่ละตัว โดยใช้ t-test จะได้ค่าพารามิเตอร์หรือค่าสัมประสิทธิ์ของตัวแปรอธิบายที่มีความเหมาะสม ขั้นสุดท้ายประเมิน

ประสิทธิภาพของโมเดลการพยากรณ์ โดยตรวจสอบกับข้อมูลในอดีต แล้วใช้โมเดลไปพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ปีการศึกษา 2531-2540 ผลการวิจัยได้โมเดล 2 โมเดล คือ

1. $Q = -151.6686 + 15.2960X_1 + 75.7035 \log G_{t-1}$
   (2.0243) (3.4710)
   $R^2 = 0.9953$ SE. est = 2.1263
2. $Q = 87.2475 + 0.0567\, S + 5.0969X_2 - 76.5132 \log L$
   (0.0028) (2.3080) (30.0134)
   $R^2 = 0.9926$ SE. est = 2.7737

โดยที่ Q คือ จำนวนครู
X_1 คือ ตัวแปรหุ่นแก้การเปลี่ยนระบบการศึกษา
$\log G_{t-1}$ คือ log ของGDP ย้อนหลัง 1 ปี
S คือ จำนวนนักเรียนลด
X_2 คือ ตัวแปรหุ่นแก้จำนวนนักเรียน
$\log L$ คือ log ของอัตราการสอน

ผลการพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ปีการศึกษา 2531-2540 เมื่อใช้โมเดลที่ 1 ได้จำนวนครู 96,650 -115,977 คน และโมเดลที่ 2 ได้จำนวนครู 95,654 - 115,101 คน

ต่อมาใน พ.ศ. 2538 นงนุช อินทรวงษ์โชติ สร้างโมเดลทางเศรษฐมิติเพื่อพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ปีการศึกษา 2538-2550 เช่นเดียวกับอังคณา พัฒนผลไพบูลย์ แต่มีตัวแปรเพิ่มขึ้นอีก 6 ตัว คือ งบประมาณทางการศึกษา รายได้ต่อหัวต่อปีของประชากร จำนวนโรงเรียน จำนวนโรงเรียนที่เปิดใหม่ อัตราการเรียนต่อ ม. 1 และตัวแปรอื่น ๆ ที่เนื่องมาจากตัวแปรจำนวนขนาดของโรงเรียน งานวิจัยของนงนุช อินทรวงษ์โชติ นั้นนอกจากจะตรวจสอบประสิทธิภาพของโมเดลการพยากรณ์กับข้อมูลในอดีตแล้ว ยังตรวจสอบกับข้อมูลในอนาคตด้วย โดยใช้ข้อมูลจากผลการพยากรณ์ของอังคณา พัฒนผลไพบูลย์ ผลการวิจัยได้โมเดล 2 โมเดล คือ

1. $Q = 96.0391 + 0.0295\,S - 2.7862\,F + 0.0019\,So$

   $(0.0044) \quad (0.6021) \quad (0.0006)$

   $R^2 = 0.9979 \quad SE.\ est = 0.6063$

2. $Q = 221.3984 + 0.0089\,S + 0.0051\,GDP - 0.0014\,Nso - 3.5012\,E$

   $(0.0012) \quad (0.0002) \quad (0.0003) \quad (0.2440)$

   $R^2 = 0.9999 \quad SE.\ est = 0.1205$

โดยที่ Q คือ จำนวนครู

S คือ จำนวนนักเรียน

GDP คือ ผลิตภัณฑ์ประชาชาติภายในประเทศ

So คือ จำนวนโรงเรียน

NSo คือ จำนวนโรงเรียนที่เปิดใหม่

F คือ จำนวนนักเรียนต่อครู

E คือ จำนวนนักเรียนต่อห้อง

ผลการพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษาปีการศึกษา 2538-2550 เมื่อใช้โมเดลที่ 1 ได้จำนวนครู 131,513-175,132 คน และโมเดลที่ 2 ได้จำนวนครู 131,600-175,422 คน (นงนุช อินทรวงษ์โชติ, 2538)

เท่าที่ผู้วิจัยศึกษาพบว่า การพยากรณ์ทางการศึกษา เช่นการพยากรณ์จำนวนครูและจำนวนนักเรียนที่ผ่านมานั้น ยังไม่มีการนำวิธีบ๊อกซ์และเจ็นกิ้นส์มาใช้ในการพยากรณ์ ทั้งนี้อาจจะเป็นเพราะว่าข้อมูลอนุกรมเวลาทางการศึกษายังไม่ค่อยมีการเก็บรวบรวมไว้มากนัก ปริมาณข้อมูลอาจไม่เพียงพอที่จะใช้วิธีของบ๊อกซ์และเจ็นกิ้นส์ได้ แต่การพยากรณ์ในสาขาอื่น ๆ เช่น การพยากรณ์ทางธุรกิจและทางประชากรศาสตร์ วิธีการของบ๊อกซ์และเจ็นกิ้นส์ ได้รับความนิยมอย่างแพร่หลาย มีงานวิจัยในประเทศและต่างประเทศหลายเรื่อง ดังเช่น งานวิจัยของ วันพร เหลืองอาภาพงศ์ (2519) และPflaumer (1992) ดังนี้

เมื่อปี พ.ศ. 2519 วันพร เหลืองอาภาพงศ์ ใช้วิธีการกำหนดโมเดลอนุกรมเวลาของบ๊อกซ์และเจ็นกินส์มาวิเคราะห์หาโมเดลที่เหมาะสมกับข้อมูลอนุกรมเวลาปริมาณส่งออกข้าว ยาง และข้าวโพดของประเทศไทยเป็นรายเดือน ตั้งแต่ปี พ.ศ. 2513-2518 พร้อมทั้งพยากรณ์ปริมาณ

ที่จะส่งออกของผลิตผลทั้งสามชนิดตั้งแต่เดือนมกราคมถึงเดือนธันวาคม 2519 ผลการวิจัยพบว่า อนุกรมเวลาปริมาณส่งออกข้าว และยาง เป็นอนุกรมเวลาคงที่ โมเดลที่มีความเหมาะสมกับข้อมูลคือ ARIMA(2,0,0) ส่วนอนุกรมเวลาปริมาณส่งออกข้าวโพด เป็นอนุกรมเวลาฤดูกาล โมเดลที่มีความเหมาะสมกับข้อมูลคือ ARIMA(0,1,1) x SARIMA$(0,1,1)_{12}$ และผลการพยากรณ์ปริมาณที่จะส่งออกของผลิตผลทั้งสามชนิด ตั้งแต่เดือนมกราคมถึงเดือนธันวาคม 2519 ได้ค่าพยากรณ์อยู่ในเกณฑ์ไม่แตกต่างจากปีที่ผ่านมาเท่าใดนัก (วันพร เหลืองอาภาพงศ์, 2519)

และในปี ค.ศ. 1992 Pflaumer ใช้วิธีบ็อกซ์และเจ็นกินส์ พยากรณ์ประชากรในระยะยาวของสหรัฐอเมริกาปี 1990-2080 โดยใช้ข้อมูลประชากรของสหรัฐอเมริกา ซึ่งเป็นอนุกรมเวลารายปีตั้งแต่ปี 1900-1988 เป็นข้อมูลปีฐาน แล้วนำข้อมูลประชากรรายปีมาหา ผลต่างลำดับที่ 1 และ 2 จนกระทั่งได้อนุกรมคงที่ และสร้างโมเดลพยากรณ์ ผลการวิจัยพบว่า ในอนาคตประชากรของสหรัฐอเมริกา จะเพิ่มขึ้นเป็น 251-489 ล้านคน (Pflaumer, 1992)

ในเมืองไทย ยังไม่ค่อยมีการใช้ข้อมูลอนุกรมเวลามาพยากรณ์เพื่อวางแผนในระดับประเทศ ผู้วิจัยจึงสนใจที่จะใช้วิธีบ็อกซ์และเจ็นกิ้นส์มาพยากรณ์ข้อมูลอนุกรมเวลาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล โดยผู้วิจัยเลือกฐานข้อมูลเกี่ยวกับปริมาณการยืมสิ่งพิมพ์ ได้แก่ ปริมาณการยืมหนังสือทั่วไป หนังสือสำรอง และวิทยานิพนธ์ ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เป็นตัวแทนอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล และใช้ฐานข้อมูลเกี่ยวกับจำนวนนักเรียนระดับประถมศึกษาและระดับมัธยมศึกษา เป็นตัวแทนอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล