บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการวิจัยเรื่องการเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่ไม่คงที่ ระหว่างวิธีบ๊อกซ์และเจนกิ้นส์ที่ใช้โมเดลสมการเชิงโครงสร้างและตัวบ่งชี้นำนั้น มีแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ซึ่งผู้วิจัยได้ศึกษาจากตำรา เอกสาร และงานวิจัยต่าง ๆ แยกนำเสนอเป็น 5 ตอน ประกอบด้วย ตอนที่ 1 มโนทัศน์เกี่ยวกับเทคนิคการพยากรณ์ ตอนที่ 2 ประเภทข้อมูลอนุกรมเวลาและการพยากรณ์วิธีบ๊อกซ์และเจนกิ้นส์ ตอนที่ 3 มโนทัศน์เกี่ยวกับโมเดลสมการเชิงโครงสร้าง (SEM) และตัวบ่งชี้นำ (leading indicator) ตอนที่ 4 การตรวจสอบความคลาดเคลื่อนของการพยากรณ์ และตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับการพยากรณ์ ในแต่ละตอนมีรายละเอียดดังต่อไปนี้

ตอนที่ 1 มโนทัศน์เกี่ยวกับเทคนิคการพยากรณ์

สาระในตอนนี้ ผู้วิจัยนำเสนอเกี่ยวกับเทคนิคการพยากรณ์เชิงปริมาณ เนื่องจากการวิจัยครั้งนี้เป็นลักษณะการพยากรณ์ในเชิงปริมาณ ผู้วิจัยจึงนำเสนอเทคนิคการพยากรณ์เชิงปริมาณเพียงหลักการโดยภาพรวมของแต่ละเทคนิคเท่านั้น และนำเสนอรายละเอียดของเทคนิคที่นำไปใช้ในการวิจัยครั้งนี้ในตอนต่อไป

การพยากรณ์เป็นกระบวนการศึกษาข้อมูลในอดีตแล้วใช้วิธีการทางสถิติวิเคราะห์ข้อมูลในอดีต เพื่อจะคาดการณ์ล่วงหน้าได้อย่างถูกต้องหรือใกล้เคียงความจริงมากที่สุด (วิชิต หล่อจีระชุณห์กุล และคณะ, 2524 ; ทรงศิริ แต้สมบัติ, 2539) เทคนิคการพยากรณ์ที่ใช้กันอยู่ในปัจจุบันนี้สามารถแบ่งออกได้เป็น 2 ประเภท คือ เทคนิคการพยากรณ์เชิงคุณภาพ (qualitative forecasting techniques) และเทคนิคการพยากรณ์เชิงปริมาณ (quantitative forecasting techniques) แต่ที่นิยมใช้กันมากคือ เทคนิคการพยากรณ์เชิงปริมาณ เทคนิคการพยากรณ์เชิงปริมาณที่นิยมใช้กันนี้สามารถจัดแบ่งออกเป็น 2 ประเภท คือ การวิเคราะห์อนุกรมเวลา (time series analysis) และโมเดลเชิงสาเหตุ (causal models)

จากการศึกษาเอกสาร และงานวิจัยที่เกี่ยวข้อง (สุรางค์ ลิขิตอรุณรัตน์, 2536 ; วัลลภา อุนวิจิตร, 2539 ; นงลักษณ์ วิรัชชัย, 2542 ; O'Donovan, 1983 ; Brockwell & Davis, 1996 ; Brinkman & McIntype, 1997 ; Pindyck & Rubinfeld, 1998) พบว่าโมเดลเชิงสาเหตุ

(causal models) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ที่เป็นเหตุและผลเนื่องจากการเกิดเหตุการณ์ต่าง ๆ ต่อกัน ความสัมพันธ์จะแสดงในรูปแบบเหตุและผล ตัวแปรที่นำมาศึกษาจะต้องมีมากกว่า 1 ตัวแปร มีวิธีการที่ใช้ในการวิเคราะห์ 2 วิธีคือ การวิเคราะห์การถดถอย (regression analysis) และโมเดลทางเศรษฐศาสตร์ (economics models) แต่ละวิธีมีหลักการในการวิเคราะห์ดังนี้

1. การวิเคราะห์การถดถอย (regression analysis) เป็นการศึกษาความผันแปรของตัวแปรตาม (dependent variable) ที่สามารถอธิบายได้ด้วยความผันแปรของตัวแปรอิสระ (independent variable) อาจจะมี 1 ตัวแปรหรือมากกว่า 1 ตัวแปรก็ได้ หรืออาจกล่าวได้ว่าเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ โดยจะวิเคราะห์ความแปรปรวนในตัวแปรตามแยกตามข้อมูลของตัวแปรอิสระเพื่อประมาณค่าและพยากรณ์ค่าของตัวแปรตามเมื่อทราบค่าของตัวแปรอิสระ และประมาณค่าความแปรปรวนในตัวแปรตามที่อธิบายได้ด้วยตัวแปรอิสระ โมเดลที่ได้อาจจะเป็นโมเดลการถดถอยที่เป็นเส้นตรง (linear) หรือไม่เป็นเส้นตรง (nonlinear) ค่าสัมประสิทธิ์การถดถอย (regression coefficient) ที่ได้บอกให้ทราบถึงอัตราการเปลี่ยนแปลงของค่าตัวแปรตามเมื่อค่าตัวแปรอิสระเปลี่ยนแปลงไปหนึ่งหน่วย โมเดลที่ได้สามารถที่จะตรวจสอบได้ว่ามีความเหมาะสมที่จะนำไปใช้หรือไม่ โดยการพิจารณาค่าสถิติบางค่าที่อธิบายประสิทธิภาพของโมเดล เช่น R^2 (coefficient of determination) นอกจากนี้ค่าสถิติดังกล่าวบอกให้ทราบว่า โมเดลถดถอยที่สร้างขึ้นนั้นสามารถใช้อธิบายความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระได้ดีหรือไม่อีกด้วย

การวิเคราะห์การถดถอยนี้นอกจากจะวิเคราะห์ในโมเดลเชิงสาเหตุแล้ว สามารถที่จะนำไปวิเคราะห์ข้อมูลอนุกรมเวลาได้อีกด้วย โดยที่ตัวแปรอิสระในการวิเคราะห์อนุกรมเวลาได้แก่ ตัวแปรเวลา ตัวแปรฤดูกาลซึ่งมีลักษณะเป็นตัวแปร dummy และ/หรือตัวแปรตรีโกณมิติ โดยตัวแปรเวลาเป็นตัวแปรในการวัดแนวโน้ม ส่วนตัวแปร dummy และ/หรือตัวแปรตรีโกณมิติเป็นตัวแปรในการวัดอิทธิพลของฤดูกาล

2. โมเดลทางเศรษฐศาสตร์ (economics models) เป็นลักษณะกลุ่มของโมเดลที่สร้างขึ้นเพื่อแสดงความสัมพันธ์ในเชิงเศรษฐศาสตร์ระหว่างตัวแปรที่เกี่ยวโยงกัน

การวิเคราะห์อนุกรมเวลามีวิธีการที่ใช้ในการวิเคราะห์ด้วยกัน 5 วิธีคือ วิธีง่าย (naive method) วิธีเฉลี่ยเคลื่อนที่ (moving average method) วิธีแยกส่วนประกอบ

(decomposition method) และวิธีบ๊อกซ์และเจนกิ้นส์ (Box-Jenkins method) โดยวิธีการวิเคราะห์อนุกรมเวลาทั้ง 5 วิธีดังกล่าวสามารถนำมาประยุกต์ใช้ร่วมกันได้ (ทรงศิริ แต้สมบัติ, 2539 ; Makridakis, Wheelwright, and McGee, 1978 ; Hanke & Reitsch, 1992 ; Bowerman & O'Connell, 1993 ; Newbold & Bos, 1994 ; Mason, Lind, and Marchal, 1999) วิธีการวิเคราะห์อนุกรมเวลาแต่ละวิธีมีหลักการดังนี้

1. วิธีง่าย (naive method) ค่าพยากรณ์ในอนาคตจะมีค่าเป็นสัดส่วนของค่าสังเกตล่าสุดซึ่งสัดส่วนเป็นอย่างไรนั้นผู้พยากรณ์เป็นผู้กำหนดขึ้น

2. วิธีเฉลี่ยเคลื่อนที่ (moving average method) เป็นวิธีการวิเคราะห์อนุกรมเวลาที่ปรับค่าสังเกตในอนุกรมเวลาให้เรียบเพื่อให้ทราบถึงลักษณะของอนุกรมเวลาอย่างคร่าว ๆ เช่น อนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากแนวโน้มและเหตุการณ์ที่ผิดปกติ ค่าเฉลี่ยเคลื่อนที่จะแสดงแนวโน้ม อนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากฤดูกาลและเหตุการณ์ที่ผิดปกติ ค่าเฉลี่ยเคลื่อนที่จะแสดงฤดูกาล หรืออนุกรมเวลาที่มีการเคลื่อนไหวเนื่องจากแนวโน้ม ฤดูกาลและเหตุการณ์ที่ผิดปกติ ค่าเฉลี่ยเคลื่อนที่จะแสดงแนวโน้มและฤดูกาล เป็นต้น โดยมีหลักการของการปรับให้เรียบ (smoothing method) เป็นการเฉลี่ยน้ำหนักของข้อมูลหรือค่าสังเกตล่าสุดชุดหนึ่ง โดยการเฉลี่ยน้ำหนักให้กับค่าสังเกตแต่ละค่าที่เลือกมานั้นมีผู้นำเสนอไว้หลายลักษณะ ในที่นี้ขอแยกนำเสนอตามลักษณะของอนุกรมเวลาดังนี้

2.1 อนุกรมเวลาที่ไม่มีแนวโน้มและฤดูกาล วิธีการเฉลี่ยเคลื่อนที่ที่นำเสนอมี 2 วิธี ได้แก่ วิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average method : SMA) และวิธีเฉลี่ยเคลื่อนที่แบบถ่วงน้ำหนัก (weighted moving average method : WMA) โดยทั้งสองวิธีนี้มีหลักการที่ต่างกันก็คือ SMA จะให้น้ำหนักกับค่าสังเกตแต่ละค่าเท่ากัน นั่นคือค่าพยากรณ์ที่ได้มาจากการเฉลี่ยค่าสังเกตจำนวนหนึ่ง แต่ WMA จะให้น้ำหนักกับค่าที่เกิดขึ้นล่าสุดมากกว่าค่าที่เกิดขึ้นก่อนหน้า นั่นคือค่าพยากรณ์ที่ได้มาจากการเฉลี่ยค่าสังเกตจำนวนหนึ่งเช่นเดียวกันแต่น้ำหนักที่ให้ในแต่ละค่าสังเกตไม่เท่ากัน

2.2 อนุกรมเวลาที่มีแนวโน้มแต่ไม่มีฤดูกาล อนุกรมเวลาที่มีแนวโน้มนั้นอาจจะมีลักษณะของแนวโน้มได้ในหลายลักษณะแตกต่างกันไป ซึ่งมีทั้งที่เป็นเส้นตรง (linear) หรือไม่เป็นเส้นตรง (nonlinear) ดังนั้นการที่จะปรับข้อมูลอนุกรมเวลาให้เรียบจึงมีลักษณะที่ต่างกันตามรูปแบบของแนวโน้ม ถ้าอนุกรมเวลาชุดนั้นมีแนวโน้มเป็นเส้นตรง วิธีเฉลี่ยเคลื่อนที่ที่

ควรจะนำมาใช้ในการปรับให้เรียบคือ วิธีเฉลี่ยเคลื่อนที่สองครั้ง (double moving average method : DMA) โดยที่ค่าจุดตัดแกน Y และค่าความชันของสมการแนวโน้มเส้นตรงจะได้จากการทำเฉลี่ยเคลื่อนที่สองครั้งของอนุกรมเวลา แต่ถ้าอนุกรมเวลาชุดนั้นมีแนวโน้มแบบเอกซ์โปเนนเชียล วิธีเฉลี่ยเคลื่อนที่ที่ควรจะนำมาใช้คือ วิธีเฉลี่ยเคลื่อนที่ของเปอร์เซ็นต์การเปลี่ยนแปลง (moving average of percentage change method : MAPC) โดยที่การประมาณอัตราการเพิ่มหรืออัตราการลดของค่าสังเกตต่อหน่วยเวลาจะทำโดยใช้หลักการของการเฉลี่ยเคลื่อนที่

3. วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียล (exponential smoothing method) เป็นวิธีการวิเคราะห์อนุกรมเวลาที่มีหลักการเดียวกันกับวิธีเฉลี่ยเคลื่อนที่ จะแตกต่างกันเฉพาะจำนวนค่าสังเกตในอดีตที่นำมาพิจารณา โดยจะใช้ค่าในอดีตทั้งหมดและน้ำหนักที่ให้กับค่าสังเกตแต่ละค่าไม่เท่ากันหรือมีลักษณะลดหลั่นแบบเอกซ์โปเนนเชียล ค่าปรับน้ำหนัก (smoothing constant) ที่ใช้ในการปรับน้ำหนักในวิธีนี้จะมีค่าอยู่ระหว่าง 0 ถึง 1 วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียลนี้มีผู้นำเสนอไว้หลายวิธีขึ้นอยู่กับลักษณะของอนุกรมเวลาดังนี้

3.1 อนุกรมเวลาไม่มีแนวโน้มและฤดูกาล วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียลที่นำเสนอคือ วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียบแบบง่าย (simple exponential smoothing method : SES) การปรับให้เรียบทำโดยการให้น้ำหนักกับค่าสังเกตที่เกิดขึ้นล่าสุดมากที่สุดและจะลดหลั่นกันไปสำหรับค่าสังเกตที่อยู่ห่างออกไป

3.2 อนุกรมเวลามีแนวโน้มแต่ไม่มีฤดูกาล การปรับให้เรียบจะมีลักษณะต่าง ๆ กันตามลักษณะของแนวโน้มอนุกรมเวลา มีผู้นำเสนอไว้ 3 วิธีคือ วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียลแบบสองครั้ง (double exponential smoothing method : DES) วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียลแบบเส้นตรง (linear exponential smoothing method : LES) และวิธีปรับให้เรียบแบบเอกซ์โปเนนเชียลแบบสามครั้ง (triple exponential smoothing method : TES) สองวิธีแรกเหมาะสมกับอนุกรมเวลาที่มีแนวโน้มแบบเส้นตรง โดยที่วิธี DES ค่าจุดตัดแกน Y และค่าความชันในสมการแนวโน้มสุดท้ายที่จะนำไปใช้ในการพยากรณ์ได้มาจากค่าจุดตัดแกน Y และค่าความชันในอดีตเช่นเดียวกันกับวิธี SES แต่การให้น้ำหนักกับค่าจุดตัดแกน Y และค่าความชันในอดีตแตกต่างกัน วิธี LES มีค่าปรับน้ำหนักสองค่า โดยที่ค่าปรับน้ำหนักค่าหนึ่งจะเป็นของค่าจุดตัดแกน Y ส่วนค่าปรับน้ำหนักอีกค่าหนึ่งจะเป็นค่าของความชัน ส่วนวิธี TES จะใช้เมื่ออนุกรมเวลามีแนวโน้มกำลังสอง (quadratic trend) ในรูปแบบแนวโน้มลักษณะนี้จะมีสัมประสิทธิ์การถดถอยจำนวน 3 พารามิเตอร์ การประมาณค่าพารามิเตอร์จะทำในลักษณะเดียวกันกับการประมาณค่า

คงที่โดยวิธี SES และการประมาณค่าจุดตัดบนแกน Y และค่าความชันโดยวิธี DES และวิธี LES เนื่องจากพารามิเตอร์ 3 ค่าจะมีค่าปรับน้ำหนัก 3 ค่า

3.3 อนุกรมเวลาไม่มีแนวโน้มแต่มีฤดูกาล จะใช้วิธีการปรับให้เรียบแบบเอกซ์โปเนนเชียลแบบฤดูกาล (seasonal exponential smoothing method : SSES) สมการที่ใช้ในการพยากรณ์สร้างขึ้นโดยใช้หลักการปรับให้เรียบดังกล่าวมาแล้ว โดยมีค่าปรับน้ำหนัก 2 ค่า สำหรับค่าคงที่และค่าวัดฤดูกาล

3.4 อนุกรมเวลาที่มีแนวโน้มแบบเส้นตรงและมีฤดูกาล ควรใช้วิธีปรับให้เรียบแบบเอกซ์โปเนนเชียลของฮอลท์และวินเทอร์ (Holt-Winters exponential smoothing method : HWS) สมการพยากรณ์ประกอบด้วยส่วนของแนวโน้มและส่วนของฤดูกาลที่สร้างขึ้นโดยใช้หลักการของการปรับให้เรียบดังกล่าวมาแล้ว โดยมีค่าปรับน้ำหนัก 3 ค่า สำหรับค่าจุดตัดแกน Y ค่าความชัน และค่าวัดอิทธิพลฤดูกาล

4. วิธีแยกส่วนประกอบ (decomposition method) เป็นวิธีการวิเคราะห์อนุกรมเวลาโดยการแยกอนุกรมเวลาออกเป็นส่วน ๆ ซึ่งอนุกรมเวลาแต่ละชุดจะประกอบด้วย 4 ส่วน ที่มีอิทธิพลหรือสามารถอธิบายในแต่ละช่วงเวลาของอนุกรมเวลาได้ และสามารถวัดได้อย่างชัดเจน คือ การวัดส่วนประกอบแนวโน้มในรูปสมการแนวโน้ม วัดส่วนประกอบฤดูกาลในรูปแผนแบบฤดูกาล (seasonal pattern) ที่อธิบายด้วย seasonal index หรือ seasonal factor วัดส่วนประกอบวัฏจักรในรูปแผนแบบวัฏจักร (cyclical pattern) ที่อธิบายด้วย cyclical index และการวัดเหตุการณ์ที่ผิดปกติ โดยเริ่มจากการนำอนุกรมเวลาชุดนั้นมาสร้างกราฟเพื่อพิจารณาการเคลื่อนไหวของอนุกรมเวลา ทำการสร้างเส้นแนวโน้มระยะยาวและหาค่าแนวโน้มด้วยวิธีใดวิธีหนึ่ง ได้แก่ วิธีมือ วิธีเฉลี่ยครึ่ง วิธีผลรวมบางส่วน วิธีเฉลี่ยเคลื่อนที่ วิธีปรับให้เรียบ และวิธีกำลังสองน้อยที่สุด ถ้าอนุกรมเวลามีการเคลื่อนไหวเนื่องจากอิทธิพลของฤดูกาล จะหาค่าวัดอิทธิพลของฤดูกาลเมื่อรูปแบบเป็นแบบบวก และดัชนีฤดูกาลเมื่อรูปแบบเป็นแบบคูณได้ กรณีที่รูปแบบเป็นคูณจะหาดัชนีฤดูกาลด้วยวิธีเฉลี่ยแบบง่าย (average percentage หรือ simple average method) วิธีสัดส่วนกับแนวโน้ม (percentage trend หรือ ratio to trend method) หรือวิธีสัดส่วนกับค่าเฉลี่ยเคลื่อนที่ (percentage moving average หรือ ratio to moving average method) สำหรับกรณีรูปแบบบวกจะทำได้ทำนองเดียวกัน โดยการแปลงจากสัดส่วนให้เป็นผลต่าง เมื่อได้ค่าวัดอิทธิพลของฤดูกาล หรือดัชนีฤดูกาลแล้วนำไปปรับอนุกรมเวลาเดิมจะได้อนุกรมเวลาใหม่เรียกว่า อนุกรมเวลาปรับฤดูกาล (deseasonalized series) เมื่อปรับแนวโน้มออกจากอนุกรม

เวลาปรับฤดูกาล จะได้อนุกรมเวลาปรับแนวโน้มและฤดูกาล อนุกรมเวลาใหม่ที่ได้จะประกอบด้วยอิทธิพลของวัฏจักรและเหตุการณ์ที่ผิดปกติ เมื่อทำการเฉลี่ยเคลื่อนที่แบบ 3, 5 หรือ 7 ช่วงเวลากับอนุกรมเวลาชุดใหม่จะได้อนุกรมเวลาที่ปรับแนวโน้ม ฤดูกาล และอิทธิพลของเหตุการณ์ที่ผิดปกติ ถ้ามีอิทธิพลของวัฏจักรเข้ามาเกี่ยวข้องกับการเคลื่อนไหวของอนุกรมเวลา จะใช้อนุกรมเวลาใหม่ในการพิจารณาอิทธิพลของวัฏจักร ถ้ามีการเคลื่อนไหวของวัฏจักรเข้ามาเกี่ยวข้องจะต้องพิจารณาต่อไปว่าแผนแบบวัฏจักรที่เกิดขึ้นกับอนุกรมเวลาชุดนั้นเป็นอย่างไร คงที่หรือไม่ และแต่ละวัฏจักรประกอบด้วยกี่ช่วงเวลา การจะนำไปพยากรณ์ก็ได้จากการรวบรวมผลการวิเคราะห์ที่ได้ตั้งแต่การเริ่มต้นในขั้นแรกจนขั้นสุดท้าย รวมทั้งต้องใช้ข้อมูลอื่น ๆ ประกอบการพยากรณ์ด้วย

5. วิธีบ๊อกซ์และเจนกิ้นส์ (Box-Jenkins method) เป็นวิธีการวิเคราะห์อนุกรมเวลาโดยการหาโมเดลที่เหมาะสมให้กับอนุกรมเวลา โดยมีโมเดล $ARMA(p,q)xSARMA(P,Q)_L$ แต่มีข้อตกลงเบื้องต้นของโมเดลว่าอนุกรมเวลาที่นำมาใช้ในการสร้างโมเดลต้องมีคุณสมบัติที่คงที่ (stationary) ถ้าอนุกรมเวลาที่นำมาใช้ในการวิเคราะห์มีลักษณะหรือคุณสมบัติที่ไม่คงที่ (nonstationary) ต้องแปลงอนุกรมเวลาชุดนั้นให้มีลักษณะที่คงที่ก่อนที่จะนำมาใช้ในการวิเคราะห์ในขั้นตอนต่อไป การแปลงอนุกรมเวลาให้อยู่ในลักษณะที่คงที่ทำได้โดยการหาผลต่าง (regular differencing) และ/หรือผลต่างฤดูกาล (seasonal differencing) ส่วนการแปลงข้อมูลด้วยฟังก์ชันต่าง ๆ เช่น ลอการิทึม รากกำลังสอง เป็นต้น จะทำเมื่ออนุกรมเวลามีค่าความแปรปรวนไม่คงที่หรือการรวมส่วนประกอบของอนุกรมเวลาเป็นแบบคูณ ในการวิเคราะห์จะใช้ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ (autocorrelation coefficients) และค่าสัมประสิทธิ์อัตตะสหสัมพันธ์บางส่วน (partial autocorrelation coefficients) เป็นหลักในการพิจารณาคัดเลือกโมเดล ARMA(p,q) แล้วตรวจสอบความเหมาะสมของโมเดลที่ได้จากการคัดเลือก เมื่อได้โมเดลที่เหมาะสมแล้ว นำโมเดลที่เหมาะสมไปใช้ในการพยากรณ์ต่อไป การพยากรณ์ด้วยวิธีนี้จะอธิบายในรายละเอียดมากขึ้นในตอนต่อไป

จากวิธีการหาค่าพยากรณ์ของเทคนิคการพยากรณ์เชิงปริมาณในส่วนของการวิเคราะห์อนุกรมเวลา สามารถสรุปสิ่งที่ต้องศึกษาในการพยากรณ์แต่ละวิธีได้ดังตาราง 1

ตอนที่ 2 ประเภทข้อมูลอนุกรมเวลาและการพยากรณ์วิธีบ๊อกซ์และเจนกิ้นส์

สาระในตอนนี้ ผู้วิจัยนำเสนอประเภทของข้อมูลอนุกรมเวลาที่กำหนดโมเดลโดยวิธีของบ๊อกซ์และเจนกิ้นส์ ในส่วนของการพยากรณ์ด้วยเทคนิคบ๊อกซ์และเจนกิ้นส์ผู้วิจัยนำเสนอ

เป็น 3 ตอน ประกอบด้วย ตอนที่ 1 ฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน ตอนที่ 2 โมเดลที่ใช้ในการวิเคราะห์อนุกรมเวลาด้วยวิธีบ๊อกซ์และเจนกิ้นส์ และตอนที่ 3 ขั้นตอนการพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีบ๊อกซ์และเจนกิ้นส์ มีรายละเอียดดังนี้

ตาราง 1 สิ่งที่ศึกษาในวิธีการหาค่าพยากรณ์ของเทคนิคการพยากรณ์อนุกรมเวลา

สิ่งที่ศึกษาในการพยากรณ์	Naive Method	Moving Average	Exponential Smoothing	Classical Decom-position	Box-Jenkins Method
ประเภทของอนุกรมเวลา					✓
อัตตะสหสัมพันธ์และอัตตะสหสัมพันธ์บางส่วน					✓
รูปแบบของเทคนิค	✓	✓	✓	✓	✓
ขั้นตอนการหาค่าพยากรณ์	✓	✓	✓	✓	✓
ตัวอย่างการคำนวณค่าพยากรณ์	✓	✓	✓	✓	✓

2.1 ประเภทข้อมูลอนุกรมเวลา

อนุกรมเวลา หมายถึงชุดของข้อมูลที่บันทึกในช่วงระยะเวลาหนึ่งอย่างต่อเนื่อง อาจจะเป็นข้อมูลรายวัน รายสัปดาห์ รายเดือน รายไตรมาส หรือรายปี เป็นต้น ช่วงเวลาห่างที่เก็บข้อมูลมาจะเท่ากันหรือไม่เท่ากันก็ได้ จะใช้สัญลักษณ์ $\{y_t\}$ แทนชุดของอนุกรมเวลา โดยที่ t = 1, 2, ... ซึ่ง t แทนช่วงเวลาของการบันทึกข้อมูล ตัวอย่างเช่น ยอดขายรายไตรมาสของห้างสรรพสินค้าแห่งหนึ่งตั้งแต่เริ่มเปิดดำเนินการในปี 1962 ผลผลิตกรดซัลฟูริกในแต่ละปีตั้งแต่ปี 1970 จำนวนคนไข้รายวันในโรงพยาบาลของรัฐแห่งหนึ่ง เป็นต้น (ทรงศิริ แต้สมบัติ, 2539 : Makridakis, Wheelwright, and McGee, 1978 ; Hanke & Reitsch, 1992 ; Bowerman & O'Connell, 1993 ; Mason, Lind, and Marchal, 1999)

อนุกรมเวลาแต่ละชุดนั้นมีลักษณะหรือแบบแผนการเคลื่อนไหวที่แตกต่างกัน โดยการเคลื่อนไหวของอนุกรมเวลาเป็นแบบใดนั้น การพิจารณาขั้นต้นสามารถพิจารณาได้จากการพล็อตกราฟระหว่างช่วงเวลากับค่าสังเกต การเคลื่อนไหวของอนุกรมเวลาขึ้นอยู่กับส่วนประกอบต่าง ๆ ซึ่งส่วนประกอบหลักของอนุกรมเวลาได้แก่ แนวโน้ม (trend) อิทธิพลของฤดูกาล (seasonal effect) อิทธิพลของวัฏจักร (cyclical effect) และเหตุการณ์ที่ผิดปกติ (irregular effect) แต่ละอนุกรมเวลาจะมีส่วนประกอบมากกว่าหนึ่งส่วนประกอบได้ รายละเอียดของแต่ละส่วนประกอบมีดังนี้

1. แนวโน้ม (trend) หมายถึงการเคลื่อนไหวของอนุกรมเวลาในระยะยาวอาจจะมีลักษณะแนวโน้มที่เพิ่มขึ้น (upward trend) หรือแนวโน้มลดลง (downward trend) นอกจากนี้ลักษณะแนวโน้มอนุกรมเวลามีลักษณะที่เป็นไปได้อีกหลายลักษณะ เช่น แนวโน้มเส้นตรง (linear trend) แนวโน้มกำลังสอง (quadratic trend) แนวโน้มแบบเอกซ์โปเนนเชียล (exponential trend) เป็นต้น ดังแผนภาพ 3

2. อิทธิพลของฤดูกาล (seasonal effect) หมายถึงการเคลื่อนไหวของอนุกรมเวลามีผลสืบเนื่องมาจากฤดูกาล นั่นคือลักษณะการเคลื่อนไหวนั้นจะเกิดขึ้นซ้ำแล้วซ้ำอีกในช่วงเวลาหนึ่ง ปัจจัยที่อาจก่อให้เกิดอิทธิพลของฤดูกาลต่ออนุกรมเวลามีได้หลายปัจจัย เช่น สภาพอากาศ อุณหภูมิ วัฒนธรรม เป็นต้น ดังแผนภาพ 3

3. อิทธิพลของวัฏจักร (cyclical effect) อนุกรมเวลาที่เก็บรวบรวมในระยะยาวหลายปี การเคลื่อนไหวอาจจะแสดงอิทธิพลของวัฏจักรที่มีลักษณะทำนองเดียวกันกับอิทธิพลของฤดูกาล โดยวัฏจักรหนึ่งจะครอบคลุมระยะเวลาหลายปี แต่ละช่วงจะมีการเคลื่อนไหวไม่แตกต่างกันมากนัก ดังแสดงในแผนภาพ 3

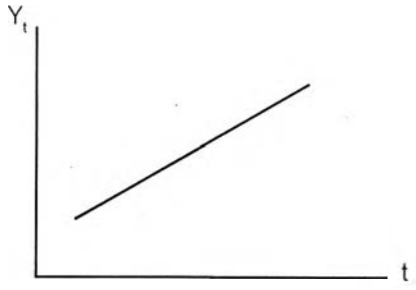
4. เหตุการณ์ที่ผิดปกติ (irregular effect) เป็นการเคลื่อนไหวของอนุกรมเวลาเฉพาะส่วนที่ไม่มีแผนแบบที่แน่นอน เหตุการณ์ผิดปกตินี้ส่วนใหญ่จะเป็นเหตุการณ์ที่ไม่ได้คาดคิดมาก่อนหรือไม่เกิดบ่อยครั้ง เช่น อุบัติภัยทางธรรมชาติ อุบัติเหตุ การปฏิวัติ การนัดหยุดงาน เป็นต้น รวมถึงปัจจัยอื่น ๆ ที่ไม่ใช่เนื่องจากแนวโน้ม อิทธิพลของฤดูกาล และอิทธิพลของวัฏจักร

จากแต่ละส่วนประกอบของอนุกรมเวลาดังกล่าวข้างต้นประกอบกันเป็นโมเดลของอนุกรมเวลามีด้วยกัน 2 ลักษณะคือ โมเดลมีการรวมกันแบบบวก (additive component model) และโมเดลมีการรวมกันแบบคูณ (multiplicative component model) การรวมกันของโมเดลทั้งสองลักษณะสามารถเขียนเป็นสมการได้ดังนี้

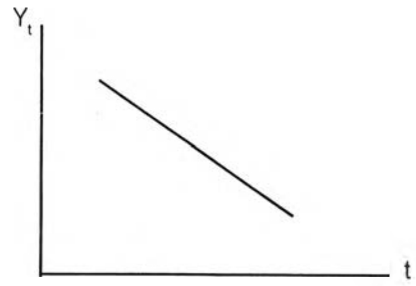
$$Z = T + S + C + I \qquad \text{แบบบวก}$$

$$Z = T \times S \times C \times I \qquad \text{แบบคูณ}$$

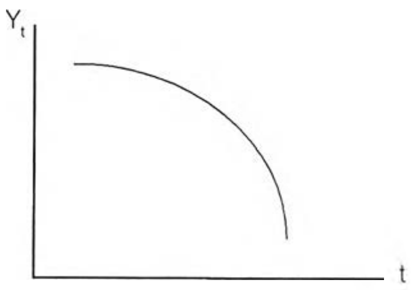
ถ้าพิจารณาช่วงเวลาของการเก็บรวบรวมข้อมูลเข้ามาร่วมด้วย พบว่าอนุกรมเวลารายปีจะไม่มีอิทธิพลเนื่องจากฤดูกาล แต่อนุกรมเวลารายเดือนหรือรายไตรมาสจะมีอิทธิพลเนื่องจากฤดูกาลเข้ามาเกี่ยวข้องด้วย



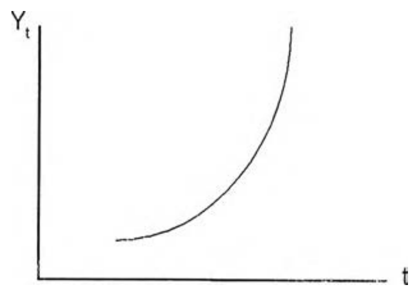
(ก) แนวโน้มเส้นตรงในอัตราที่เพิ่มขึ้น



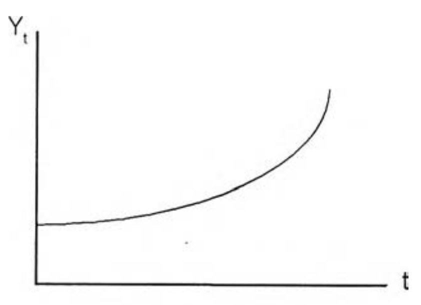
(ข) แนวโน้มเส้นตรงในอัตราที่ลดลง



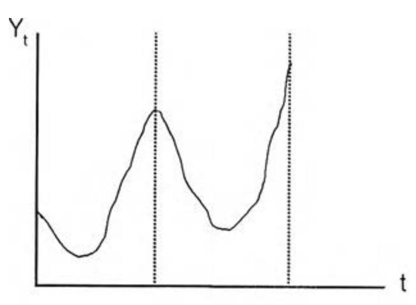
(ค) แนวโน้มแบบกำลังสองลดลงในอัตราที่เพิ่มขึ้น



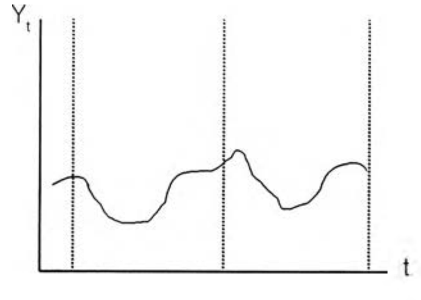
(ง) แนวโน้มแบบกำลังสองเพิ่มขึ้นในอัตราที่เพิ่มขึ้น



(จ) แนวโน้มแบบเอกซ์โปเนนเชียล



(ฉ) อิทธิพลของฤดูกาลในช่วงเวลา 2 ปี



(ช) อิทธิพลของวัฏจักรในช่วงเวลา 20 ปี

แผนภาพ 3 การเคลื่อนไหวของอนุกรมเวลาที่ขึ้นอยู่กับส่วนประกอบ

จากส่วนประกอบของอนุกรมเวลาที่กล่าวมาแล้วข้างต้น สามารถจำแนกออกได้เป็น 2 ประเภทคือ ส่วนประกอบที่วัดค่าได้ (deterministic component) และส่วนประกอบที่วัดค่าไม่ได้ (stochastic component) ส่วนประกอบที่วัดค่าได้ประกอบด้วย แนวโน้ม วัดได้จากสมการแนวโน้มที่สร้างขึ้น ซึ่งวัดได้ทั้งขนาดและทิศทาง อิทธิพลของฤดูกาลสามารถวัดได้ด้วยดัชนีฤดูกาล (seasonal index) หรือค่าวัดอิทธิพลของฤดูกาล (seasonal factor) และอิทธิพลของวัฏจักรสามารถวัดได้ด้วยดัชนีวัฏจักร (cyclical index) หรือค่าวัดอิทธิพลของวัฏจักร (cyclical factor) ส่วนประกอบของอนุกรมเวลาที่ไม่สามารถวัดค่าได้ได้แก่ เหตุการณ์ที่ผิดปกติ

การวิเคราะห์เบื้องต้นว่าการเคลื่อนไหวของอนุกรมเวลาประกอบด้วยส่วนประกอบอะไรบ้าง และมีการรวมตัวของส่วนประกอบเป็นแบบบวกหรือแบบคูณมีความจำเป็นต่อการวิเคราะห์อนุกรมเวลาในขั้นตอนต่อไปเป็นอย่างมาก การตรวจสอบว่าการเคลื่อนไหวของอนุกรมเวลามีส่วนประกอบอะไรบ้างนั้น อาจทำได้โดยพิจารณาด้วยสายตาจากกราฟหรือใช้การทดสอบสมมติฐานต่าง ๆ ทั้งแบบใช้พารามิเตอร์และไม่ใช้พารามิเตอร์ ส่วนการตรวจสอบว่าอนุกรมเวลามีรูปแบบการรวมกันของโมเดลเป็นแบบบวก (additive model) หรือแบบคูณ (multiplicative model) ยังไม่มีผู้พัฒนาหลักการในการตรวจสอบ ดังนั้นผู้พยากรณ์อาจจะพิจารณาจากกราฟระหว่างช่วงเวลากับค่าสังเกต ถ้าเส้นอนุกรมเวลาแกว่งออกจากเส้นแนวโน้มในเดือนเดียวกันแต่ต่างปีกันไม่ต่างกัน รูปแบบการรวมกันของโมเดลเป็นแบบบวก แต่ถ้าเส้นอนุกรมเวลาแกว่งออกจากเส้นแนวโน้มในเดือนเดียวกันแต่ต่างปีกันอาจจะต่างกันในทิศทางที่เพิ่มขึ้นหรือลดลง รูปแบบการรวมกันของโมเดลเป็นแบบคูณ การกำหนดรูปแบบของอนุกรมเวลาว่าเป็นแบบบวกหรือแบบคูณนั้น จำเป็นต้องทำก่อนการวิเคราะห์อนุกรมเวลาเพราะวิธีการพยากรณ์จะแตกต่างกันตามรูปแบบที่กำหนด การกำหนดรูปแบบของอนุกรมเวลาผิดพลาดทำให้ค่าพยากรณ์ต่างจากค่าจริงมากหรือให้ค่าความคลาดเคลื่อนของการพยากรณ์สูง

อนุกรมเวลาที่จะนำมาศึกษาเพื่อประโยชน์ในการพยากรณ์มีลักษณะต่าง ๆ กันหลายลักษณะดังได้กล่าวแล้วเกี่ยวกับส่วนประกอบของอนุกรมเวลา ประเภทของอนุกรมเวลาที่กำหนดโมเดลโดยวิธีบ๊อกซ์และเจนกิ้นส์ (Box-Jenkins method) แบ่งออกเป็น 3 ประเภทคือ

1. อนุกรมเวลาคงที่ (stationary time series) เป็นอนุกรมเวลา $\{y_t\}$ ที่มีค่าเฉลี่ยและค่าความแปรปรวนของ Y_t คงที่ นั่นคือค่าเฉลี่ย $E(Y_t)$ และค่าความแปรปรวน $V(Y_t)$ มีค่าคงที่สำหรับแต่ละเวลา t ซึ่งอนุกรมเวลาที่มีแนวโน้มและ/หรือฤดูกาลจะมี $E(Y_t)$ ไม่คงที่ อนุกรมเวลาที่

คงที่จะต้องมีค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ที่เวลาล้าหลัง k ขึ้นอยู่กับค่า k อย่างเดียว อนุกรมเวลาที่จะกำหนดโมเดล ARMA(p,q) ให้จะต้องเป็นอนุกรมเวลาที่คงที่เท่านั้น ดังแผนภาพ 4

2. อนุกรมเวลาไม่คงที่ (nonstationary time series) เป็นข้อมูลที่เคลื่อนไหวไม่แน่นอน หรือมีความผันแปรของ Y_t สูงจะเป็นลักษณะของอนุกรมเวลาที่ $V(Y_t)$ ไม่คงที่ จะหาโมเดล ARMA(p,q) ให้กับอนุกรมเวลาดังกล่าวไม่ได้จะต้องแปลงอนุกรมเวลานั้นให้เป็นอนุกรมเวลาใหม่ที่มีคุณสมบัติที่คงที่เสียก่อน จึงจะหารูปแบบ ARMA(p,q) ให้กับอนุกรมเวลาใหม่ได้ และจะไม่สามารถหาค่ากลางที่เป็นศูนย์กลางได้ ดังแผนภาพ 4 การแปลงอนุกรมเวลาเดิมให้เป็นอนุกรมเวลาใหม่ที่คงที่จะทำได้ด้วยวิธีการดังนี้

2.1 หาผลต่าง (regular differencing) ของอนุกรมเวลา นั่นคือถ้าอนุกรมเวลา $\{y_t\}$ มีแนวโน้ม จะแปลงให้เป็นอนุกรมเวลาใหม่ที่ไม่มีแนวโน้ม $\{z_t\}$ โดย $Z_t = \nabla^d Y_t$ โดย d เป็นลำดับของการหาผลต่าง เช่น เมื่อ d = 1 ; $Z_t = \nabla Y_t = Y_t - Y_{t-1}$ เมื่อ d = 2 ; $Z_t = \nabla^2 Y_t = \nabla(Y_t - Y_{t-1}) = \nabla Y_t - \nabla Y_{t-1} = Y_t - Y_{t-1} - Y_{t-1} + Y_{t-2} = Y_t - 2Y_{t-1} + Y_{t-2}$ เป็นต้น จำนวนครั้งที่หาผลต่างจะขึ้นอยู่กับว่าเมื่อหาผลต่างแล้วอนุกรมเวลาใหม่คงที่หรือไม่ ถ้ายังไม่คงที่ต้องหาผลต่างต่อไป โดยทั่วไปถ้าอนุกรมเวลามีแนวโน้มเป็นแบบเส้นตรง จะใช้ d = 1 ส่วนเมื่ออนุกรมเวลามีแนวโน้มกำลังสอง (quadratic trend) จะใช้ d = 2

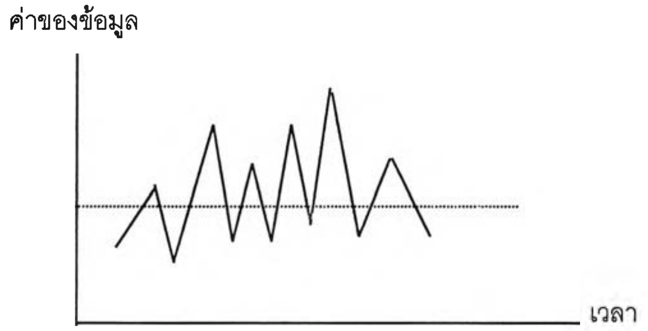
2.2 หาผลต่างฤดูกาล (seasonal differencing) ของอนุกรมเวลา ถ้าอนุกรมเวลามีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง จะแปลงอนุกรมเวลาเดิม $\{y_t\}$ ให้เป็นอนุกรมเวลาใหม่ที่ไม่มีฤดูกาล $\{z_t\}$ โดย $Z_t = \nabla^D{}_L Y_t$ โดย D เป็นลำดับของการหาผลต่าง และ L เป็นจำนวนฤดูกาลต่อปี เช่น อนุกรมเวลารายเดือน (L = 12) เมื่อ D = 1 $Z_t = \nabla_{12} Y_t = Y_t - Y_{t-12}$ เมื่อ D = 2 $Z_t = \nabla^2{}_{12} Y_t = \nabla_{12}(Y_t - Y_{t-12}) = Y_t - Y_{t-12} - Y_{t-12} + Y_{t-24} = Y_t - 2Y_{t-12} + Y_{t-24}$ เป็นต้น จำนวนครั้งที่หาผลต่างฤดูกาลจะขึ้นอยู่กับว่าเมื่อหาผลต่างฤดูกาลแล้วอนุกรมเวลาใหม่คงที่หรือไม่ ถ้ายังไม่คงที่ต้องหาผลต่างต่อไป

2.3 หาผลต่างและผลต่างฤดูกาล (regular and seasonal differencing) กรณีที่อนุกรมเวลามีทั้งแนวโน้มและฤดูกาล การปรับให้อนุกรมเวลาคงที่นั้นจะทำได้โดยการหาผลต่างและผลต่างฤดูกาลควบคู่กัน เช่น อนุกรมเวลารายเดือน (L = 12) เมื่อ d = 1 และ D = 1 จะแปลงอนุกรมเวลาเดิม $\{y_t\}$ ให้เป็นอนุกรมเวลาใหม่ที่ไม่มีแนวโน้มและฤดูกาล $\{z_t\}$ โดย $Z_t = \nabla\nabla_{12} Y_t = \nabla(Y_t - Y_{t-12}) = \nabla Y_t - \nabla Y_{t-12} = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$ เป็นต้น จำนวน

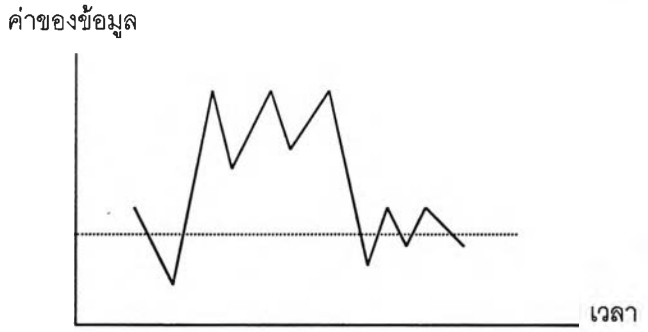
ครั้งที่หาผลต่างและผลต่างฤดูกาลจะขึ้นอยู่กับว่าเมื่อหาผลต่างทั้งสองแล้วอนุกรมเวลาใหม่คงที่หรือไม่ ถ้ายังไม่คงที่ต้องหาผลต่างต่อไป

2.4 หาลอการิทึมของค่าสังเกตในอนุกรมเวลา เป็นการแปลงอนุกรมเวลาเดิม $\{y_t\}$ ให้เป็นอนุกรมเวลาใหม่ $\{z_t\}$ โดย $Z_t = \log(Y_t)$ การแปลงนี้จะทำเมื่อความผันแปรของอนุกรมเวลาไม่คงที่ นั่นคือ $V(Y_t)$ ไม่คงที่สำหรับค่า t ใด ๆ

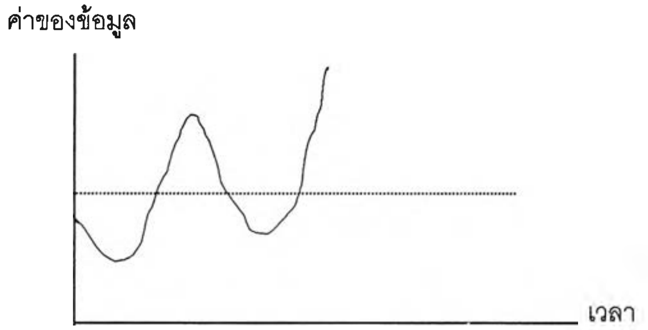
3. อนุกรมเวลาฤดูกาล (seasonal time series) เป็นข้อมูลที่มีการเคลื่อนไหวขึ้นลงตามระยะเวลาเป็นช่วงที่แน่นอน และลักษณะการเคลื่อนไหวในระยะเวลาหนึ่งเหมือนกับช่วงเวลาอื่น ๆ ซ้ำ ๆ กัน ดังแผนภาพ 4



(ก) ลักษณะของอนุกรมเวลาคงที่



(ข) ลักษณะของอนุกรมเวลาไม่คงที่



(ค) ลักษณะของอนุกรมเวลาฤดูกาล

แผนภาพ 4 ลักษณะของอนุกรมเวลาที่กำหนดโดยโมเดลของวิธีบ๊อกซ์และเจนกิ้นส์

2.2 การพยากรณ์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์

สาระในตอนนี้ ผู้วิจัยนำเสนอเกี่ยวกับฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function) ฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation function) และโมเดลในการวิเคราะห์อนุกรมเวลาด้วยวิธีบ๊อกซ์และเจนกิ้นส์ เนื่องจากสาระในตอนต้นนี้เป็นสิ่งจำเป็นที่ผู้พยากรณ์ต้องทำความเข้าใจเพื่อจะได้เป็นประโยชน์สำหรับการเลือกโมเดลที่เหมาะสมให้กับการวิเคราะห์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์ได้อย่างถูกต้อง ในส่วนท้ายของตอนนี้ผู้วิจัยได้นำเสนอขั้นตอนการวิเคราะห์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์อย่างละเอียดเพื่อให้มีความเข้าใจในแต่ละขั้นตอนของการวิเคราะห์ได้อย่างละเอียด โดยมีรายละเอียดดังนี้

2.2.1 ฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation function) เป็นฟังก์ชันที่จะนำไปใช้ในการกำหนดโมเดล ARMA(p,q) ให้กับอนุกรมเวลา โดยการพิจารณาจากค่าสัมประสิทธิ์ที่ได้จากแต่ละฟังก์ชันประกอบกัน มีรายละเอียดดังนี้

(1) ฟังก์ชันอัตตะสหสัมพันธ์ (autocorrelation function) เป็นความสัมพันธ์ที่เกิดระหว่างข้อมูลอนุกรมเวลาชุดเดียวกันแต่ต่างเวลากัน ตัวอย่างเช่น ข้อมูล Y_t มีอัตตะสหสัมพันธ์กับข้อมูล Y_{t-1} หมายความว่า ข้อมูลทั้งสองชุดมีความสัมพันธ์กัน โดยที่ข้อมูล Y_t เป็นข้อมุลชุดเดียวกับ Y_{t-1} แต่ระยะเวลาต่างกัน 1 ช่วงเวลาให้หลัง ความสัมพันธ์ที่เกิดขึ้นเรียกว่าอัตตะสหสัมพันธ์อันดับที่ 1 (autocorrelation at lag 1) สำหรับการวัดความสัมพันธ์จะใช้ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ (coefficient of autocorrelation : r_k) เมื่อ k เป็นระยะเวลาที่แตกต่างกัน (lag time period) เมื่อ n คือ จำนวนข้อมูลอนุกรมเวลา มีสูตรการคำนวณสำหรับตัวอย่างดังนี้

$$r_k = \frac{\sum_{t=1}^{n-k}\left(Y_t - \bar{Y}\right)\left(Y_{t+k} - \bar{Y}\right)}{\sum_{t=1}^{n}\left(Y_t - \bar{Y}\right)^2} \qquad \text{สำหรับ } k = 1, 2, \ldots$$

และ

$$V(r_k) = \frac{1}{n}\left(1 + 2\sum_{j=1}^{k-1} r_j^2\right) \qquad \text{สำหรับ } k = 1, 2, \ldots$$

เมื่อ r_k แทน สัมประสิทธิ์อัตตะสหสัมพันธ์ที่เวลาล้าหลัง k

t แทน ช่วงเวลา

n แทน ช่วงเวลาสุดท้าย

k แทน เวลาล้าหลัง

Y_t แทน ตัวแปรเริ่มต้น

$\bar{Y}$ แทน ค่าเฉลี่ยของค่าสังเกต มีค่าเท่ากับ $\frac{\sum_{t=1}^{n} Y_t}{n}$

ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ที่ได้มีค่าอยู่ระหว่าง –1 ถึง 1 เมื่อขนาดของ r_k ที่วัดด้วย $|r_k|$ มีค่าเข้าใกล้ 1 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลามีสหสัมพันธ์กันสูง เมื่อ $|r_k|$ มีค่าเข้าใกล้ 0 แสดงว่ามีสหสัมพันธ์กันต่ำ ถ้า r_k มีค่าน้อยกว่า 0 แสดงว่ามีความสัมพันธ์กันในทิศทางที่ตรงกันข้ามหรือทางลบ และถ้า r_k มีค่ามากกว่า 0 แสดงว่ามีความสัมพันธ์กันในทิศทางเดียวกันหรือทางบวก ในการวัดค่าสัมประสิทธิ์อัตตะสหสัมพันธ์ พบว่า $r_k = r_{-k}$ กล่าวคือการวัดสหสัมพันธ์ระหว่าง Y_t และ Y_{t-k} หรือระหว่าง Y_t และ Y_{t+k} ซึ่งต่างก็เป็นค่าสังเกตที่อยู่ห่างกัน k ค่า มีค่าเท่ากัน

ค่า r_k เป็นค่าวัดลักษณะของตัวอย่างที่เป็นค่าประมาณค่าของวัดลักษณะของประชากร ρ_k เมื่อ ρ_k มีค่าเป็น 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 $S^2_{r_k}$ ซึ่ง $S^2_{r_k}$ มีค่าใกล้เคียงกับ 1/n การทราบลักษณะการแจกแจงของ r_k จะทำให้สามารถทดสอบสมมติฐานเกี่ยวกับ ρ_k ได้ โดยมีสมมติฐานดังนี้

$$H_0 : \rho_k = 0$$

$$H_1 : \rho_k \neq 0$$

จะใช้ทดสอบสถิติ $Z = \frac{r_k}{S_{rk}} = \frac{r_k}{\sqrt{\frac{1}{n}}} = \sqrt{n}\, r_k$ ที่มีช่วงวิกฤติ $|Z| \geq Z_{\frac{\alpha}{2}}$ ที่ระดับนัยสำคัญ α อาจจะใช้ทดสอบสถิติ r_k ที่มีช่วงวิกฤติเป็น $|r_k| \geq Z_{\frac{\alpha}{2}}\sqrt{\frac{1}{n}}$ โดยทั่วไปการทดสอบจะทำที่ระดับนัยสำคัญ .05 จึงมีช่วงวิกฤติเป็น $|r_k| \geq \frac{1.96}{\sqrt{n}} \approx \frac{2}{\sqrt{n}}$

(2) ฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (partial autocorrelation) เป็นความสัมพันธ์ที่เกิดระหว่างข้อมูลอนุกรมเวลาชุดเดียวกัน Y_t กับ Y_{t-k} สำหรับ k = 1, 2, ..., n โดยกำหนดให้ความสัมพันธ์ที่เกิดระหว่าง Y_t กับข้อมูลที่ช่วงเวลาอื่น ๆ (ยกเว้นข้อมูล Y_{t-k}) ถูกขจัดออกทั้งหมดก่อน สำหรับการวัดความสัมพันธ์จะใช้ค่าสัมประสิทธิ์อัตตะสหสัมพันธ์บางส่วน (coefficient of partial autocorrelation : r_{kk}) และมีค่าความแปรปรวน คือ VAR (r_{kk}) = 1/n เมื่อ n คือ จำนวนข้อมูลอนุกรมเวลา มีสูตรการคำนวณสำหรับตัวอย่างดังนี้ คือ

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{(k-1)j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{(k-1)j} r_j}$$

สำหรับ k = 1 ; $r_{kk} = r_1$

สำหรับ k = 1, 2, ..., n

ค่า r_{kk} เป็นค่าวัดลักษณะของตัวอย่างที่เป็นค่าประมาณของค่าวัดลักษณะของประชากร ρ_{kk} โดยที่ r_{kk} มีการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวนเป็น $\sigma^2_{r_{kk}}$ หรือ r_{kk} ~ N(0, $\sigma^2_{r_{kk}}$) จะประมาณ $\sigma^2_{r_{kk}}$ ด้วย $S^2_{r_{kk}}$ ซึ่ง $S^2_{r_{kk}}$ = 1/n ดังนั้น การทดสอบสมมติฐานว่า r_{kk} มาจาก ρ_{kk} ที่เป็น 0 หรือไม่ สามารถเขียนเป็นสมมติฐานได้ดังนี้

$H_0 : \rho_{kk} = 0$

$H_1 : \rho_{kk} \neq 0$

จะใช้ทดสอบสถิติ r_{kk} ที่มีช่วงวิกฤติเป็น $|r_{kk}| \geq \frac{2}{\sqrt{n}}$ โดยทั่วไปการทดสอบจะทำที่ระดับนัยสำคัญ .05

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

ลักษณะที่ 1 ค่า r_k และ r_{kk} ลดลงอย่างรวดเร็วเป็น 0 (cut off) เรียกว่า มีค่าต่ำสุด ถ้า r_k และ r_{kk} ของอัตตะสหสัมพันธ์และอัตตะสหสัมพันธ์บางส่วนของตัวอย่างมีค่าโด่งสุด (spike) ที่เวลาล้าหลัง k มีขนาดใหญ่ จะปฏิเสธสมมติฐานหลัก นั่นคืออัตตะสหสัมพันธ์และอัตตะสหสัมพันธ์บางส่วนทางทฤษฎี $\rho_k = 0$, $\rho_{kk} = 0$ และ ρ_k , ρ_{kk} มีการวัดความสัมพันธ์เชิงเส้นระหว่างค่าอนุกรมเวลาที่เป็นไปได้ทั้งหมด แยกโดยหน่วยของเวลาล้าหลัง k ผู้พยากรณ์สามารถ

ตัดสินค่าโด่งสุดที่เวลาล้าหลัง k ของอัตตะสหสัมพันธ์และอัตตะสหสัมพันธ์บางส่วนของตัวอย่าง โดยพิจารณาค่าสถิติ t ที่สัมพันธ์กับ r_k และ r_{kk}

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 ใน 3 แบบดังนี้

แบบที่ 1 ลดลงอย่างช้า ๆ แบบเอกซ์โปเนนเชียล (damped exponential)

แบบที่ 2 ลดลงอย่างช้า ๆ แบบฟังก์ชันตรีโกณมิติแบบไซน์ (damped sine-wave)

แบบที่ 3 แบบที่ 1 รวมกับแบบที่ 2

ลักษณะฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของกลุ่มตัวอย่างอนุกรมเวลาทั้งสองลักษณะดังกล่าวข้างต้น แสดงดังแผนภาพ 5

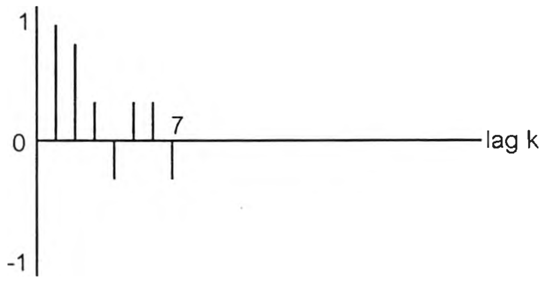
ในการกำหนดโมเดล ARMA(p,q) ให้กับอนุกรมเวลาจะทำโดยการเปรียบเทียบระหว่างค่า r_k และ r_{kk} จากอนุกรมเวลากับ ρ_k และ ρ_{kk} ของโมเดลต่าง ๆ หรือเปรียบเทียบจากคอเรลโลแกรม ค่า r_k และ r_{kk} จะหาได้แต่อาจจะมีค่าเล็กจนกระทั่งสรุปได้ว่ามาจาก ρ_k และ ρ_{kk} ที่มีค่าเป็น 0 ดังนั้นการสรุปว่า r_k และ r_{kk} มีค่าเล็กจนกล่าวได้ว่ามีค่าเป็น 0 นั้น จะทำให้การเปรียบเทียบคอเรลโลแกรมทำได้สะดวกขึ้น จึงควรได้มีการพิจารณาก่อนการเปรียบเทียบทุกครั้งว่าค่า r_k และ r_{kk} มีค่าน้อยจนถือว่าเป็น 0 ได้หรือไม่

2.2.2 โมเดลในการวิเคราะห์อนุกรมเวลาด้วยวิธีบ๊อกซ์และเจนกิ้นส์

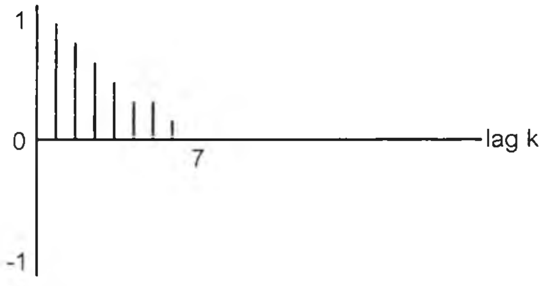
โมเดลที่ใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา มี 3 โมเดลดังนี้ (ทรงศิริ แต้สมบัติ, 2539 ; Makridakis, Wheelwright, and McGee, 1982 ; Pindyck & Rubinfeld, 1991 ; Hanke & Reitsch, 1992 ; Bowerman & O'Connell, 1993)

โมเดล 1 โมเดลกระบวนการเฉลี่ยเคลื่อนที่ (moving average models)

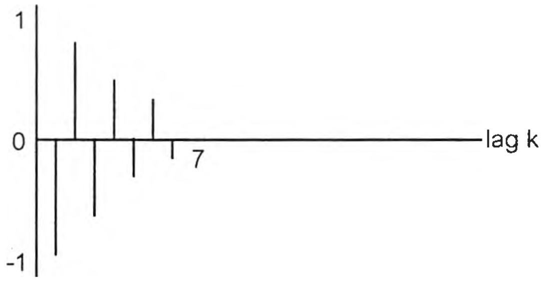
กระบวนการเฉลี่ยเคลื่อนที่ลำดับที่ q (moving average process of order q) เป็นโมเดลที่ค่าสังเกต Y_t เกิดขึ้นจากการให้น้ำหนักเฉลี่ยของความคลาดเคลื่อนอย่างสุ่ม (random disturbances) ที่เกิดขึ้นก่อนหน้า q ช่วงเวลา เขียนแทนด้วย MA(q) และเขียนเป็นสมการได้ดังนี้



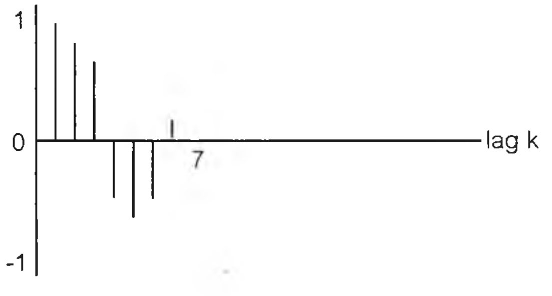
(ก) ลดลงอย่างรวดเร็วหลังเวลาล้าหลัง k



(ข) ลดลงอย่างช้า ๆ แบบเอกซ์โปเนนเชียล



(ค) ลดลงอย่างช้า ๆ แบบเอกซ์โปเนนเชียลที่ไม่กระเพื่อม



(ง) ลดลงอย่างช้า ๆ แบบฟังก์ชันตรีโกณมิติแบบไซน์

แผนภาพ 5 ฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของกลุ่มตัวอย่างที่ลดลงอย่างรวดเร็วและลดลงอย่างช้า ๆ

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

เมื่อ θ_0 แทน ค่าคงที่

$\theta_1, \dots, \theta_q$ แทน พารามิเตอร์กระบวนการเฉลี่ยเคลื่อนที่ (MA) ตัวที่ q

ε_t แทน ค่าความคลาดเคลื่อนช่วงเวลา t

ในโมเดลกระบวนการเฉลี่ยเคลื่อนที่ $\theta_1, \dots, \theta_q$ เป็นพารามิเตอร์ที่อาจจะเป็นบวกหรือลบได้ ดังนั้นเครื่องหมายหน้า θ_i อาจจะเป็นบวกหรือลบได้ นอกจากนี้มีข้อตกลงว่าให้ความคลาดเคลื่อนอย่างสุ่มเป็นอิสระต่อกันข้ามช่วงเวลา แต่ละเทอมของ ε_t สมมติให้เป็นตัวแปรสุ่มปกติ (normal random variable) ที่มีค่าเฉลี่ยเป็น 0 ความแปรปรวนเป็น σ^2_ε และความแปรปรวนร่วม $\gamma_k = 0$ สำหรับ $k \neq 0$ ในทางปฏิบัติลำดับของกระบวนการเฉลี่ยเคลื่อนที่ (MA) มักมีค่าไม่เกิน 2 นั่นคือ $q \leq 2$

ฟังก์ชันอัตตะสหสัมพันธ์ (ρ_k) มีลักษณะ cut off คือ

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

และ

$$\rho_k = 0 \qquad ; k > q$$

ส่วนฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (ρ_{kk}) มีลักษณะ dies down แบบ damped exponential หรือ damped sine-wave หรือทั้งสองแบบรวมกัน

โมเดล 2 โมเดลกระบวนการถดถอยในตัวเอง (autoregressive models)

กระบวนการถดถอยในตัวเองลำดับที่ p (autoregressive process of order p) เป็นโมเดลที่ค่าสังเกต Y_t เกิดขึ้นจากการให้น้ำหนักเฉลี่ยของค่าสังเกตที่เกิดขึ้นก่อนหน้า p ช่วงเวลา ซึ่งความคลาดเคลื่อนอย่างสุ่มจะเกิดขึ้นกับช่วงเวลาปัจจุบัน เขียนแทนด้วย AR(p) และเขียนเป็นสมการได้ดังนี้

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

เมื่อ Y_t แทน ค่าสังเกต

θ_0 แทน ค่าคงที่ซึ่งสัมพัทธ์กับค่าเฉลี่ยของกระบวนการเฟ้นสุ่ม (stochastic process)

ϕ_i แทน พารามิเตอร์กระบวนการถดถอยในตัวเอง (AR) ตัวที่ i

ε_t แทน ค่าความคลาดเคลื่อนช่วงเวลา t (random shock)

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

$$\rho_{kk} \neq 0 \qquad ; k = 1, \ldots, q$$

และ

$$\rho_{kk} = 0 \qquad ; k > q$$

ในทางปฏิบัติลำดับของกระบวนการถดถอยในตัวเอง (AR) มักมีค่าไม่เกิน 2 นั่นคือ $p \leq 2$

โมเดล 3 โมเดลรวมกันของกระบวนการถดถอยในตัวเองและกระบวนการเฉลี่ยเคลื่อนที่ (mixed autoregressive-moving average models)

ในกระบวนการสุ่มคงที่ (stationary random processes) บางครั้งไม่สามารถที่จะใช้โมเดลการเฉลี่ยเคลื่อนที่หรือโมเดลกระบวนการถดถอยในตัวเองในการสร้างโมเดลได้เพียงโมเดลใดโมเดลหนึ่ง เนื่องจากในแต่ละโมเดลต่างก็มีคุณภาพของกระบวนการทั้งสองโมเดล จึงได้นำโมเดลทั้งสองมารวมกันในการสร้างโมเดล เรียกว่า กระบวนการรวมกันของกระบวนถดถอยในตัวเองและกระบวนการเฉลี่ยเคลื่อนที่ลำดับ (p,q) (mixed autoregressive-moving average process of order (p,q)) เขียนแทนด้วย ARMA(p,q) และเขียนเป็นสมการได้ดังนี้

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q}$$

เมื่อ $\theta_0 = \mu(1 - \phi_1 - \ldots - \phi_q)$

ฟังก์ชันอัตตะสหสัมพันธ์ (ρ_k) และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน (ρ_{kk}) มีลักษณะ dies down แบบ damped exponential หรือ damped sine-wave หรือทั้งสองแบบรวมกัน

โมเดลที่ใช้ในการวิเคราะห์อนุกรมเวลาด้วยวิธีบ็อกซ์และเจนกิ้นส์ทั้ง 3 โมเดลข้างต้น สามารถนำมาสรุปเป็นลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนสำหรับการวิเคราะห์อนุกรมเวลาคงที่ (Bowerman & O'Connell, 1993) ดังตาราง 2

ตาราง 2 ลักษณะทางทฤษฎีของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของโมเดลการวิเคราะห์อนุกรมเวลาคงที่

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

จากการแบ่งประเภทของอนุกรมเวลาที่กำหนดโมเดลโดยวิธีบ๊อกซ์และเจนกิ้นส์ 3 ประเภท ได้แก่ อนุกรมเวลาคงที่ อนุกรมเวลาไม่คงที่ และอนุกรมเวลาฤดูกาล มีลักษณะของโมเดล ARMA(p,d,q) แตกต่างกันไปตามลักษณะของประเภทอนุกรมเวลา มีรายละเอียดดังนี้

สำหรับอนุกรมเวลาคงที่ โมเดล ARMA(p,q) มี p เป็นอันดับของ AR และ q เป็นอันดับของ MA จำนวนพารามิเตอร์ในโมเดลจะเท่ากับ p+q+1 โมเดลที่กำหนดให้กับอนุกรมเวลามักจะเป็นโมเดลที่มีจำนวนพารามิเตอร์น้อย ในทางปฏิบัติมักจะไม่เกิน 3 พารามิเตอร์ ดังตาราง 3 (ทรงศิริ แต้สมบัติ, 2539) แสดงโมเดล ARMA ต่าง ๆ และเงื่อนไขของพารามิเตอร์ที่ปรากฏในโมเดล สำหรับเงื่อนไขของพารามิเตอร์ในส่วนของ AR อันดันที่ 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 การพิจารณาว่าค่าพารามิเตอร์ $\phi_1, \ldots, \phi_p$ ใด ที่จะทำให้โมเดล AR คงที่จะทำได้โดย

1. จากโมเดล AR(p)

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

หรือ

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \ldots - \phi_p Y_{t-p} = \theta_0 + \varepsilon_t$$

จะเขียนโมเดลในเทอมของ backward operator ได้เป็น

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

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

ตาราง 3 โมเดล ARMA(p,q) และเงื่อนไขความคงที่และอินเวอร์ติเบิ้ล

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)	$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$	$\phi_2 + \phi_1 < 1$
		$\phi_2 - \phi_1 < 1$
		$-1 < \phi_2 < 1$
MA(1)	$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1}$	$-1 < \theta_1 < 1$
MA(2)	$Y_t = \theta_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$	$\theta_2 + \theta_1 < 1$
		$\theta_2 - \theta_1 < 1$
		$-1 < \theta_2 < 1$
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 ซึ่งเป็นคุณสมบัติของโมเดล MA(q) ที่ทำให้ค่าคลาดเคลื่อนของการพยากรณ์ ε_t ในเทอมของ Y_t, Y_{t-1}, ... ได้ การพิจารณาว่าค่าพารามิเตอร์ $\theta_1, \ldots, \theta_q$ ใดที่ทำให้โมเดล MA(q) เป็น invertible ทำได้โดย

1. จากโมเดล MA(q)

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

หรือ

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

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

สำหรับอนุกรมเวลาที่มีแนวโน้มเพียงอย่างเดียว นั่นคือมีค่าเฉลี่ย $E(Y_t)$ ไม่คงที่สำหรับแต่ละค่าของ t การปรับอนุกรมเวลาดังกล่าวให้เป็นอนุกรมเวลาใหม่ที่มีลักษณะคงที่จะทำได้โดยการหาผลต่าง นั่นคือจากอนุกรมเวลา $\{y_t\}$ แปลงเป็นอนุกรมเวลาใหม่ $\{z_t\}$ เมื่ออนุกรมเวลา $\{y_t\}$ มีลักษณะคงที่แล้วจะหาโมเดล ARMA(p,q) ให้กับอนุกรมเวลา $\{z_t\}$ เป็น Z_t~ARMA (p,q) และสำหรับโมเดลของอนุกรมเวลา $\{y_t\}$ เป็น Y_t~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 4

ตาราง 4 โมเดล ARIMA(p,1,q) ของอนุกรมเวลา $\{z_t\}$ และ $\{y_t\}$

โมเดลARIMA(p,1,q)	อนุกรมเวลา $\{z_t\}$	อนุกรมเวลา $\{y_t\}$
ARIMA(0,1,0)	$Z_t = \theta_0 + \varepsilon_t$	$Y_t = \theta_0 + Y_{t-1} + \varepsilon_t$
ARI(1,1)	$Z_t = \theta_0 + \phi_1 Z_{t-1} + \varepsilon_t$	$Y_t = \theta_0 + (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t$
ARI(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)	$Z_t = \theta_0 + \phi_1 Z_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$	$Y_t = \theta_0 + (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t - \theta_1\varepsilon_{t-1}$

ลักษณะของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของโมเดล ARIMA (p,1,q) แสดงในตาราง 5 (ทรงศิริ แต้สมบัติ, 2539)

สำหรับอนุกรมเวลาที่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โมเดลที่จะใช้ได้แก่ $SARIMA(P,D,Q)_L$ (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 $\{y_t\}$ ที่ไม่คง

ที่เนื่องจากฤดูกาลเป็นอนุกรมเวลาชุดใหม่ $\{z_t\}$ ที่คงที่ ตัวอย่างของโมเดล SARMA(P,Q)$_L$ ของอนุกรมเวลา $\{y_t\}$ ดังตาราง 6

ตาราง 5 ลักษณะของ $\rho_k(Y_t)$, $\rho_k(Z_t)$ และ $\rho_{kk}(Z_t)$ สำหรับโมเดล ARIMA(p,d,q)

โมเดล ARIMA(p,d,q) ของ Y_t	ฟังก์ชันอัตตะ-สหสัมพันธ์ของ Y_t	ฟังก์ชันอัตตะ-สหสัมพันธ์ของ Z_t	ฟังก์ชันอัตตะสหสัมพันธ์บางส่วนของ Z_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

ตาราง 6 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)$_L$	$Y_t = \theta_0 + \varepsilon_t - \theta_L \varepsilon_{t-L} - \theta_{2L} \varepsilon_{t-2L}$
SARMA(1,1)$_L$	$Y_t = \theta_0 + \phi_L Y_{t-L} + \varepsilon_t - \theta_L \varepsilon_{t-L}$

คุณสมบัติที่สำคัญของโมเดล SARMA(P,Q)$_L$ เมื่อ L = 12 ของอนุกรมเวลาที่มีลักษณะคงที่ได้แก่ ρ_k และ ρ_{kk} ลักษณะของ ρ_k และ ρ_{kk} แสดงดังตาราง 7

ตาราง 7 ลักษณะของ $\rho_k(Z_t)$ และ $\rho_{kk}(Z_t)$ สำหรับโมเดล SARMA(P,Q)$_{12}$

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

ตัวอย่างของโมเดล SARIMA(P,1,Q)$_{12}$ ของอนุกรมเวลา $\{y_t\}$ หรือโมเดล SARMA(P,Q)$_{12}$ ของอนุกรมเวลา $\{z_t\}$ เป็นดังนี้

(1) $Z_t \sim SAR(1)_{12}$ หรือ $Y_t \sim SARI(1,1)_{12}$ 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

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

ในกรณีที่อนุกรมเวลามีทั้งแนวโน้มและอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง โมเดลที่จะใช้ได้แก่ ARIMA(p,d,q)xSARIMA(P,D,Q)$_L$ ผู้พยากรณ์ต้องแปลงอนุกรมเวลาให้อยู่ในลักษณะที่คงที่ โดยการหาผลต่างและผลต่างฤดูกาลควบคู่กัน โดย d เป็นจำนวนครั้งที่หาผลต่าง และ D เป็นจำนวนครั้งที่หาผลต่างฤดูกาลเพื่อทำให้อนุกรมเวลา $\{y_t\}$ ที่ไม่คงที่เนื่องจากมีแนวโน้มและอิทธิพลของฤดูกาลเป็นอนุกรมเวลาชุดใหม่ $\{z_t\}$ ที่คงที่ โดย $Z_t = \nabla^d \nabla_L^D Y_t$ เช่น สำหรับอนุกรมเวลารายปี (L = 12) และมีโมเดล $Y_t \sim ARI(1,1)xSARI(1,1)_{12}$ หรือ $Z_t \sim AR(1)xSAR(1)_{12}$ สามารถเขียนสมการได้ดังนี้

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

2.2.3 ขั้นตอนการพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีบ๊อกซ์และเจนกิ้นส์

การใช้วิธีการพยากรณ์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์ ถ้าอนุกรมเวลาที่นำมาใช้ในการวิเคราะห์มีลักษณะที่คงที่ (stationary) สามารถที่จะนำอนุกรมเวลานั้นมาวิเคราะห์ตามขั้นตอนของวิธีบ๊อกซ์และเจนกิ้นส์ได้ทันที แม้ว่าอนุกรมเวลาที่นำมาวิเคราะห์มีลักษณะที่ไม่คงที่ก็ตาม แต่ขั้นตอนในการวิเคราะห์เหมือนกันจะต่างกันในส่วนของการทำให้อนุกรมเวลาที่ไม่คงที่เป็นอนุกรมเวลาที่คงที่ กล่าวคือถ้าอนุกรมเวลามีแนวโน้มให้หาผลต่างของอนุกรมเวลาจนได้อนุกรมเวลาชุดใหม่ที่มีลักษณะคงที่ ถ้าอนุกรมเวลามีอิทธิพลของฤดูกาลให้หาผลต่างฤดูกาลของอนุกรมเวลาจนได้อนุกรมเวลาชุดใหม่ที่มีลักษณะคงที่ ถ้าอนุกรมเวลามีแนวโน้มและอิทธิพลของฤดูกาลให้หาทั้งผลต่างและผลต่างฤดูกาลจนได้อนุกรมเวลาที่มีลักษณะคงที่ และถ้าอนุกรมเวลามีค่าความแปรปรวนไม่คงที่ให้แปลงอนุกรมเวลาเดิมเป็นอนุกรมเวลาใหม่โดยการหาลอการิทึม ($Z_t = \ln Y_t$) เป็นต้น จนกว่าอนุกรมเวลาใหม่จะมีค่าความแปรปรวนคงที่

เมื่ออนุกรมเวลาที่นำมาวิเคราะห์มีลักษณะที่คงที่แล้ว นำอนุกรมเวลาที่ได้นั้นไปวิเคราะห์ตามขั้นตอนการวิเคราะห์ของวิธีบ๊อกซ์และเจนกิ้นส์ ประกอบด้วย 4 ขั้นตอน ได้แก่ การกำหนดโมเดล การประมาณค่าพารามิเตอร์ในโมเดล การตรวจสอบความเหมาะสมของโมเดล และการพยากรณ์ค่าในอนาคตจากสมการพยากรณ์ รายละเอียดในแต่ละขั้นตอนมีดังนี้

1. การกำหนดโมเดล (identification) เป็นการกำหนดโมเดล ARMA(p,q) สำหรับอนุกรมเวลาที่อนุกรมเวลาเดิมมีลักษณะที่คงที่ และโมเดล ARMA(p,q)xSARMA(P,Q)$_L$ สำหรับอนุกรมเวลาที่อนุกรมเวลาเดิมมีลักษณะที่ไม่คงที่ การกำหนดโมเดลนี้อาศัยการเปรียบเทียบคอเรลโลแกรมของ r_k และ r_{kk} ของตัวอย่างกับคอเรลโลแกรมของ ρ_k และ ρ_{kk} ของโมเดลต่าง ๆ ซึ่ง k = 1, 2, ... และ k = L, 2L, ... สำหรับโมเดล ARMA(p,q) และ SARMA(P,Q)$_L$ ตามลำดับ หากโมเดลของอนุกรมเวลามีแผนแบบเหมือนกันกับคอเรลโลแกรมของโมเดลใดมากที่สุดจะกำหนดโมเดลนั้นให้กับอนุกรมเวลา โมเดลที่กำหนดในขั้นตอนนี้อาจจะมีมากกว่า 1 โมเดลได้ ในกรณีที่ลักษณะของ r_k และ r_{kk} ไม่ชัดเจนว่าตรงกับโมเดลใด เนื่องจากค่า r_k และ r_{kk} เป็นเพียงค่าประมาณของ ρ_k และ ρ_{kk} และอาจจะมีค่าน้อยจนสรุปได้ว่ามาจาก ρ_k และ ρ_{kk} ที่มีค่าเท่ากับ 0 หรืออาจจะมีค่ามากจนสรุปได้ว่ามาจาก ρ_k และ ρ_{kk} ที่ไม่เท่ากับ 0 ดังนั้นก่อนการพิจารณากำหนดโมเดลควรจะพิจารณา r_k ที่มีค่าใหญ่ และ r_k ที่มีค่าเล็ก โดยพิจารณาว่าอยู่ในช่วงหรือนอกช่วงวิกฤติก่อน กล่าวคือพิจารณาว่า $|r_k|$ มีค่ามากกว่า $\frac{2}{\sqrt{n}}$ หรือไม่ ในส่วนของ r_{kk} ก็ทำในทำนองเดียวกันกับ r_k

กรณีที่อนุกรมเวลา $\{y_t\}$ มีค่าเฉลี่ยสูงเมื่อเทียบกับค่าส่วนเบี่ยงเบนมาตรฐาน จะทำให้ค่าของสถิติทดสอบสำหรับการทดสอบสมมติฐาน

$H_0 : \mu = 0$

$H_1 : \mu \neq 0$

มีค่าสูงและทำให้ปฏิเสธสมมติฐานหลัก เนื่องจากพารามิเตอร์ θ_0 ในทุกโมเดลสามารถเขียนได้ในรูปฟังก์ชันของ μ ดังนั้นถ้าค่า μ มีค่าไม่เป็น 0 ค่า θ_0 ก็จะไม่เป็น 0 ด้วย ดังนั้นโมเดล ARMA (p,q) และ SARMA$(P,Q)_L$ ที่กำหนดจะมีพารามิเตอร์ θ_0 ในโมเดลด้วยหรือไม่นั้น ขึ้นอยู่กับค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของอนุกรมเวลา $\{y_t\}$

2. การประมาณค่าพารามิเตอร์ในโมเดล (estimation) เป็นขั้นตอนที่ผู้พยากรณ์ทำต่อเนื่องจากการกำหนดโมเดลสำหรับอนุกรมเวลา ค่าประมาณที่ได้เป็นค่าเริ่มต้นที่จะนำไปคำนวณหาค่าประมาณสุดท้ายสำหรับพารามิเตอร์นั้น ๆ ค่าประมาณสุดท้ายที่ได้ถือว่าเป็นค่าที่ดีที่สุดสำหรับโมเดลนั้น และเมื่อนำค่าประมาณสุดท้ายที่ได้ไปสร้างเป็นสมการพยากรณ์จะทำให้การพยากรณ์ค่ามีความคลาดเคลื่อนน้อยที่สุด (Makridakis, Wheelwright and McGee, 1978 ; Newbold & Bos, 1994) การประมาณค่าพารามิเตอร์สามารถกระทำได้ 2 แบบ ได้แก่ การประมาณค่าแบบง่ายและการประมาณค่าด้วยวิธีกำลังสองน้อยที่สุด (ทรงศิริ แต้สมบัติ, 2539) การประมาณแบบจุดทำได้ด้วยวิธีการต่าง ๆ ดังนี้

2.1 การประมาณค่าแบบง่าย จะใช้สมการที่แสดงความสัมพันธ์ระหว่าง ρ_k และพารามิเตอร์ในโมเดล จำนวนสมการที่จะใช้เท่ากับอันดับที่ปรากฏในโมเดล เมื่อประมาณ ρ_k ด้วย r_k ในสมการความสัมพันธ์และแก้สมการจะได้ค่าประมาณของพารามิเตอร์ ส่วนค่าประมาณของ θ_0 จะได้จากสมการที่แสดงความสัมพันธ์ระหว่าง $E(Y_t)$ กับพารามิเตอร์ในโมเดล เช่น

(1) สำหรับกระบวนการถดถอยในตัวเองลำดับที่ 1 หรือ AR(1)

กระบวนการถดถอยในตัวเองลำดับที่ p หรือ AR(p) ใช้สมการ Yule-Walker ดังนี้

$$\rho_1 = \phi_1 + \phi_2\rho_1 + \ldots + \phi_p\rho_{k-1}$$

$$\rho_2 = \phi_1\rho_1 + \phi_2 + \ldots + \phi_p\rho_{k-2}$$

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \ldots + \phi_p$$

เมื่อ $\rho_1, \ldots, \rho_k$ แทน สัมประสิทธิ์อัตตะสหสัมพันธ์ทางทฤษฎี สำหรับ 1, ..., k

$\phi_1, \ldots, \phi_p$ แทน ค่าสัมประสิทธิ์ของกระบวนการถดถอยในตัวเองลำดับที่ p

จากสมการข้างต้นสามารถหา $\hat{\phi}_1$ ได้จากสมการความสัมพันธ์ดังต่อไปนี้

$$\rho_1 = \phi_1$$

และ $\hat{\theta}_0$ ได้จากสมการความสัมพันธ์ดังต่อไปนี้

$$\mu = \frac{\theta_0}{1-\phi_1}$$

โดยที่ ρ_1 แทนค่าด้วย r_1 และ μ แทนค่าด้วย $\bar{Y}$ จะได้ $\hat{\phi}_1 = r_1$ และ $\hat{\theta}_0 = \bar{y}(1-r_1)$

(2) สำหรับกระบวนการถดถอยในตัวเองลำดับที่ 2 หรือ AR(2)

กระบวนการถดถอยในตัวเองลำดับที่ 2 ใช้สมการ Yule-Walker เช่นเดียวกันกับข้างต้น โดยการแก้สมการหา $\hat{\phi}_1$ และ $\hat{\phi}_2$ จากสมการความสัมพันธ์ดังต่อไปนี้

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

$$\rho_2 = \phi_1\rho_1 + \phi_2$$

และหา $\hat{\theta}_0$ ได้จากสมการความสัมพันธ์ดังต่อไปนี้

$$\mu = \frac{\theta_0}{1-\phi_1-\phi_2}$$

โดยที่ ρ_1 และ ρ_2 แทนค่าด้วย r_1 และ r_2 ส่วน μ แทนค่าด้วย $\bar{Y}$

(3) สำหรับกระบวนการเฉลี่ยเคลื่อนที่ลำดับที่ 1 หรือ MA(1)

กระบวนการเฉลี่ยเคลื่อนที่ลำดับที่ q หรือ MA(q) สามารถเขียนเป็นสมการในรูปของอัตตะสหสัมพันธ์ทางทฤษฎีได้ดังนี้

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

$$\rho_k = 0 \qquad \text{สำหรับ } k > q$$

จากสมการข้างต้นสามารถหา $\hat{\theta}_1$ ได้จากสมการความสัมพันธ์ดังต่อไปนี้

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

และหา $\hat{\theta}_0$ ได้จากสมการความสัมพันธ์

$$\mu = \theta_0$$

โดยที่ ρ_1 แทนค่าด้วย r_1 และ μ แทนค่าด้วย $\bar{Y}$ ซึ่งค่า $\hat{\theta}_1$ ที่ได้จะมี 2 ค่า ค่าที่เลือกใช้จะเป็นค่าที่ทำให้โมเดลเป็น invertible

2.2 การประมาณค่าด้วยวิธีกำลังสองน้อยที่สุด ค่าประมาณของพารามิเตอร์ในโมเดลจะเป็นค่าประมาณที่ทำให้ผลรวมกำลังสองของการประมาณน้อยที่สุด การหาค่าประมาณนี้จะทำไม่ได้โดยการแก้สมการ จะต้องใช้การวิเคราะห์ตัวเลขและคอมพิวเตอร์ช่วยในการคำนวณ เพราะมีการคำนวณที่ยุ่งยากและซ้ำกันหลายครั้งทำให้ใช้เวลาในการคำนวณมาก การกำหนดค่าเริ่มต้นใช้ค่าประมาณแบบง่ายดังที่ได้กล่าวมาแล้ว เพื่อจะนำค่าเริ่มต้นที่ได้ไปประมาณค่าสุดท้ายต่อไป

3. การตรวจสอบโมเดล (diagnostic checking) จากการกำหนดโมเดลให้กับอนุกรมเวลานั้นโมเดลที่ได้อาจจะไม่ใช่โมเดลที่เหมาะสมสำหรับอนุกรมเวลาก็เป็นได้ ควรมีการตรวจสอบความเหมาะสมของโมเดลที่กำหนดขึ้นมา การตรวจสอบความเหมาะสมของโมเดลมีผู้เสนอไว้ด้วยกันหลายวิธี ควรจะใช้หลายวิธีรวมกัน วิธีตรวจสอบส่วนใหญ่จะใช้ค่าความคลาดเคลื่อนของการพยากรณ์ (e_t) ที่เป็นผลต่างระหว่างค่าจริง (Y_t) และค่าพยากรณ์ $(\hat{Y}_t)$ เป็นหลักในการพิจารณาการตรวจสอบความเหมาะสมของโมเดล (ทรงศิริ แต้สมบัติ, 2539 ; Makridakis, Wheelwright and McGee, 1978 : Hanke & Reitsch, 1992 ; Bowerman & O'Connell, 1993 ; Newbold & Bos, 1994) มีรายละเอียดในแต่ละวิธีดังนี้

3.1 พิจารณาว่าอนุกรมเวลา $\{e_t\}$ มีลักษณะการเคลื่อนไหวที่เป็นอิสระกันหรือไม่ นั่นคือเป็นการตรวจสอบข้อจำกัดของโมเดลที่กำหนดว่า ε_t จะต้องมีการแจกแจงที่เป็นอิสระกัน หากพบว่า $\{e_t\}$ มีลักษณะการเคลื่อนไหวที่เป็นอิสระกันจะสรุปว่าโมเดลที่กำหนดนั้นเหมาะสมกับอนุกรมเวลาแล้ว ซึ่งจะทำโดยการทดสอบสมมติฐานดังนี้

$H_0 : \rho_k(e_t) = 0$

$H_1 : \rho_k(e_t) \neq 0$

สำหรับ k = 1, ...

การทดสอบนี้จะปฏิเสธสมมติฐานหลักที่ระดับนัยสำคัญ .05 เมื่อ $|r_k(e_t)| \geq \frac{2}{\sqrt{n}}$ เมื่อ n เป็นจำนวนช่วงเวลาของอนุกรมเวลา $\{e_t\}$

3.2 พิจารณาค่าความแปรปรวนของค่าคลาดเคลื่อน $S^2 = \sum \frac{e_t^2}{n}$ ซึ่ง n เป็นจำนวนค่าคลาดเคลื่อน โมเดลที่เหมาะสมจะให้ S^2 ต่ำ

3.3 พิจารณาค่าพารามิเตอร์ในโมเดลมีค่าเป็น 0 หรือไม่ นั่นคือเมื่อ $\theta, \hat{\theta}$ และ $s_{\hat{\theta}}$ เป็นพารามิเตอร์ ค่าประมาณ และค่าคลาดเคลื่อนมาตรฐานของค่าประมาณ $\hat{\theta}$ ตามลำดับ โดยทำการทดสอบสมมติฐานดังนี้

$$H_0 : \theta = 0$$
$$H_1 : \theta \neq 0$$

ใช้สถิติทดสอบ $Z = \frac{\hat{\theta}}{s_{\hat{\theta}}}$ การปฏิเสธสมมติฐานหลักจะทำเมื่อ $|Z| \geq Z_{\frac{\alpha}{2}}$ ที่ระดับนัยสำคัญ α

3.4 พิจารณา $\rho_k(e_t) = 0$ สำหรับ k = 1, ..., m หรือไม่ นั่นคือพิจารณาว่าค่าคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน m ช่วงเวลาเป็นอิสระกันหรือไม่ กำหนดสมมติฐานในการทดสอบดังนี้

$$H_0 : \rho_1(e_t) = \ldots = \rho_m(e_t) = 0$$
$H_1 : \rho_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เท่ากับ 0

สำหรับ k = 1, ..., m

การทดสอบสมมติฐานข้างต้นนี้จะใช้การทดสอบของ Box-Pierce หรือการทดสอบของ Box-Ljung การกำหนดสมมติฐานหลักและสมมติฐานเลือกเหมือนกัน ช่วงวิกฤติเดียวกัน ซึ่งต่างก็เป็นการทดสอบสมมติฐานโดยใช้สถิติไค-สแควร์ที่มีสูตรในการคำนวณต่างกันดังนี้

$$Box-Pierce \quad Q = n\sum_{k=1}^{m} r_k^2(e_t)$$

$$Box-Ljung \quad Q_m = n(n+2)\sum_{k=1}^{m} \frac{r_k^2(e_t)}{n-k}$$

เมื่อ n แทน ขนาดของอนุกรมเวลา $\{e_t\}$

m แทน ช่วงเวลาห่างสูงสุดของ e_t ในอนุกรมเวลา $\{e_t\}$ ที่นำมาพิจารณา

a แทน จำนวนพารามิเตอร์ทั้งหมดในโมเดล ซึ่งรวม θ_0 ด้วย

สถิติทดสอบดังกล่าวมีช่วงวิกฤติ Q หรือ $Q_m \geq \chi^2_{\alpha,m-a}$ ถ้าปฏิเสธสมมติฐานหลักสรุปได้ว่ามี $\rho_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เท่ากับ 0 สำหรับ k = 1, ..., m นั่นคือมีอัตตะสหสัมพันธ์ระหว่างค่า

ความคลาดเคลื่อนของการพยากรณ์ที่อยู่ห่างกัน k ค่า หรือสรุปได้ว่าโมเดลที่กำหนดให้กับอนุกรมเวลายังไม่เหมาะสม

4. การพยากรณ์ (forecast) เมื่อได้โมเดลและโมเดลที่ได้ผ่านการตรวจสอบความเหมาะสมของโมเดลแล้ว ผู้พยากรณ์สามารถนำโมเดลอนุกรมเวลาที่ได้ไปใช้ในการพยากรณ์ในอนาคต โดยที่การพยากรณ์ค่าในอนาคตสามารถทำได้ 2 แบบคือ การพยากรณ์แบบจุด (point forecast) และการพยากรณ์แบบช่วง (interval forecast) ในลักษณะของการพยากรณ์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์ถ้าพยากรณ์ค่าในช่วงเวลาที่ห่างไปจากค่าสังเกตมาก ๆ ค่าพยากรณ์ที่ได้จะอาศัยสารสนเทศจากข้อมูลจริงหรือค่าสังเกตที่มีอยู่น้อยลงมาเท่านั้น โดยนำค่าจากการพยากรณ์ในช่วงเวลาที่ผ่านมาใช้ในการพยากรณ์ในเวลาข้างหน้าต่อไปเรื่อย ๆ ซึ่งค่าพยากรณ์ที่ได้เหล่านี้มีข้อตกลงว่าความคลาดเคลื่อนในการพยากรณ์มีค่าเท่ากับ 0 ทำให้ความแม่นยำลดลงมากถ้ายิ่งพยากรณ์ในช่วงเวลาที่ห่างออกไป อนุกรมเวลาที่พยากรณ์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์จึงมีความเหมาะสมกับการพยากรณ์ในระยะสั้นและควรพยากรณ์ล่วงหน้าเพียง 1 ช่วงเวลา จึงเกิดความแม่นยำที่สุด (ทรงศิริ แต้สมบัติ, 2539 ; Makridakis, Wheelwright and McGee, 1978 ; Hanke & Reitsch, 1992 ; Bowerman & O'Connell, 1993 ; Newbold & Bos, 1994) หลักการของการพยากรณ์ทั้งสองแบบได้แก่ การพยากรณ์แบบจุด และการพยากรณ์แบบช่วง มีรายละเอียดดังนี้

4.1 การพยากรณ์แบบจุด ค่าพยากรณ์ในอนาคตจะได้จากการแทนค่าสังเกตในอดีตหรือค่าความคลาดเคลื่อนในอดีตในสมการพยากรณ์ ในสมการพยากรณ์จะแทนค่าประมาณของพารามิเตอร์ที่ได้ประมาณขึ้นในขั้นตอนที่สอง สำหรับโมเดล ARMA(p,q) ซึ่ง $p+q \leq 2$ มีสมการพยากรณ์ $\hat{Y}_t(l)$ ดังแสดงในตาราง 8 (ทรงศิริ แต้สมบัติ, 2539) สมการพยากรณ์จะเขียนได้ทั้งในเทอมของพารามิเตอร์ μ และพารามิเตอร์อื่น หรือในเทอมของพารามิเตอร์ θ_0 และพารามิเตอร์อื่น การใช้สมการพยากรณ์จะทำโดยการแทนค่าประมาณของพารามิเตอร์ในโมเดลลักษณะของโมเดล ARMA(p,q) ซึ่ง $p+q \leq 2$ เมื่อกำหนด $\hat{Y}_t(l)$ เป็นค่าพยากรณ์ ณ เวลา t+1 ใช้ค่าสังเกต t ค่า มี

$$\hat{Y}_t(l) = \begin{cases} Y_{t+1} & ; l \leq 0 \\ \hat{Y}_t(l) & ; l \geq 0 \end{cases}$$

และ $\hat{e}_t(l)$ เป็นค่าความคลาดเคลื่อนของการพยากรณ์ ณ เวลา t+1 ใช้ค่าสังเกต t ค่า มี

$$\hat{e}_t(l) = \begin{cases} e_{t+1} & ; l \leq 0 \\ 0 & ; l \geq 0 \end{cases}$$

ตาราง 8 สมการพยากรณ์ ARMA(p,q) ในเทอมของ $\hat{\mu}$ และ $\hat{\theta}_0$

โมเดล	สมการพยากรณ์ $\hat{Y}_t(l)$			
	ในเทอมของ $\hat{\mu}$		ในเทอมของ $\hat{\theta}_0$	
white noise	$\hat{\mu}$	,l≥1	$\hat{\theta}_0$	,l≥1
AR(1)	$\hat{\mu}+\hat{\phi}_1^{l}(Y_t-\hat{\mu})$	,l≥1	$\hat{\theta}_0+\hat{\phi}_1 Y_t$	,l=1
			$\hat{\theta}_0+\hat{\phi}_1\hat{Y}_t(l-1)$	,l≥2
AR(2)	$\hat{\mu}+\hat{\phi}_1(Y_t-\hat{\mu})+\hat{\phi}_2(Y_{t-1}-\hat{\mu})$	,l=1	$\hat{\theta}_0+\hat{\phi}_1 Y_t+\hat{\phi}_2 Y_{t-1}$	,l=1
	$\hat{\mu}+\hat{\phi}_1(\hat{Y}_t(l)-\hat{\mu})+\hat{\phi}_2(Y_t-\hat{\mu})$	,l=2	$\hat{\theta}_0+\hat{\phi}_1\hat{Y}_t(l)+\hat{\phi}_2 Y_t$	,l=2
	$\hat{\mu}+\hat{\phi}_1(\hat{Y}_t(l-1)-\hat{\mu})+\hat{\phi}_2(\hat{Y}_t(l-2)-\hat{\mu})$	,l≥3	$\hat{\theta}_0+\hat{\phi}_1\hat{Y}_t(l-1)+\hat{\phi}_2\hat{Y}_t(l-2)$	,l≥3
MA(1)	$\hat{\mu}-\hat{\theta}_1 e_t$	,l=1	$\hat{\theta}_0-\hat{\theta}_1 e_t$	,l=1
	$\hat{\mu}$	,l≥2	$\hat{\theta}_0$	,l≥2
MA(2)	$\hat{\mu}-\hat{\phi}_1 e_t-\hat{\theta}_2 e_{t-1}$	,l=1	$\hat{\theta}_0-\hat{\theta}_1 e_t-\hat{\theta}_2 e_{t-1}$	,l=1
	$\hat{\mu}-\hat{\theta}_2 e_t$	,l=2	$\hat{\theta}_0-\hat{\theta}_2 e_t$	,l=2
	$\hat{\mu}$	,l≥3	$\hat{\theta}_0$	,l≥3
ARMA(1,1)	$\hat{\mu}+\hat{\phi}_1(Y_t-\hat{\mu})-\hat{\theta}_1 e_t$	,l=1	$\hat{\theta}_0+\hat{\phi}_1 Y_t-\hat{\theta}_1 e_t$	,l=1
	$\hat{\mu}+\hat{\phi}_1^{l}(Y_t-\hat{\mu})$	,l≥2	$\hat{\theta}_0+\hat{\phi}_1\hat{Y}_t(l-1)$	,l≥2

4.2 การพยากรณ์แบบช่วง ณ เวลา t+1 ค่าจริงเป็น Y_{t+1} แต่ ณ เวลา t จะหาค่าจริงที่เวลา t+1 ไม่ได้ จะได้เพียงค่าพยากรณ์ $\hat{Y}_t(l)$ ซึ่งค่าพยากรณ์ $\hat{Y}_t(l)$ อาจจะมีค่าต่างจากค่าจริง Y_{t+1} มาก หรือค่าคลาดเคลื่อนของการพยากรณ์ ณ เวลา t+1 $e_t(l)$ มีค่าสูง การพยากรณ์อีกแบบหนึ่งที่ให้ความเชื่อมั่นแก่ผู้พยากรณ์ในการพยากรณ์ค่า ณ เวลา t+1 ได้แก่ การพยากรณ์แบบช่วง การพยากรณ์แบบช่วงนั้น ช่วงของการพยากรณ์ค่า Y_{t+1} จะขึ้นกับค่าพยากรณ์แบบจุด $\hat{Y}_t(l)$ การสร้างช่วงของการพยากรณ์จะใช้ผลที่ค่าคลาดเคลื่อนของการพยากรณ์ ณ เวลา t+1 $e_t(l)$ มีการแจกแจงประมาณด้วยการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวน $\sigma^2_{e_t(l)}$ ดังนั้น $e_t(l)$ จะอยู่ในช่วงของการประมาณที่มีขีดจำกัด $\pm z_{\frac{\alpha}{2}}\sigma_{e_t(l)}$ ที่ระดับความเชื่อมั่น (1-α)100 เปอร์เซ็นต์ สามารถเขียนในรูปของความน่าจะเป็นได้ดังนี้

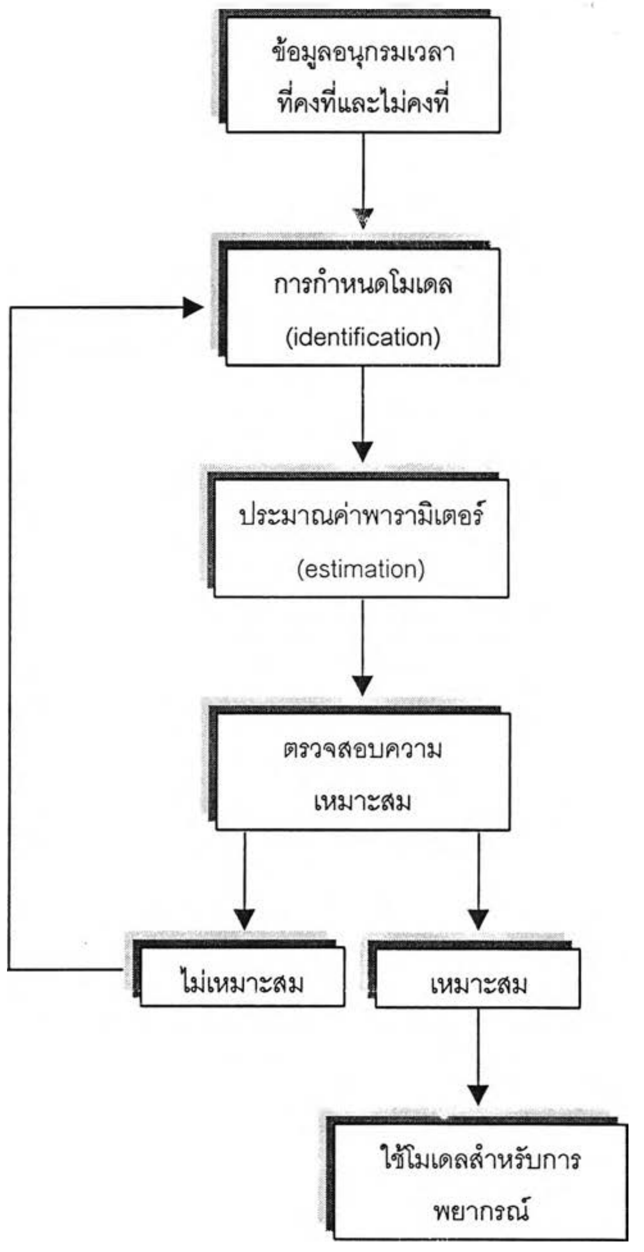
$$P\left(-z_{\alpha/2}\sigma_{e_t(l)} < e_t(l) < z_{\alpha/2}\sigma_{e_t(l)}\right)=1-\alpha$$

เมื่อแทน $e_t(l)$ ด้วย $Y_{t+1}-\hat{Y}_t(l)$ จะได้

$$P\left(\hat{Y}_t(l)-z_{\alpha/2}\sigma_{e_t(l)} < Y_{t+1} < \hat{Y}_t(l)+z_{\alpha/2}\sigma_{e_t(l)}\right)=1-\alpha$$

ดังนั้นค่าจริง Y_{t+1} จะอยู่ในช่วงของการพยากรณ์ที่มีขีดจำกัด $\hat{Y}_t(l) \pm z_{\alpha/2}\sigma_{e_t(l)}$ ด้วยความเชื่อมั่น $(1-\alpha)100$ เปอร์เซ็นต์ กรณีที่ไม่ทราบค่า $\sigma_{e_t(l)}$ จะใช้ค่าประมาณ $s_{e_t(l)}$ แทน

ขั้นตอนการพยากรณ์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์ สามารถสรุปเป็นแผนภาพได้ดังแผนภาพ 6



แผนภาพ 6 ขั้นตอนการพยากรณ์ด้วยเทคนิคบ๊อกซ์และเจนกิ้นส์

ตอนที่ 3 มโนทัศน์เกี่ยวกับโมเดลสมการเชิงโครงสร้างและตัวบ่งชี้นำ

ผู้วิจัยนำเสนอแยกเป็น 2 ตอน ประกอบด้วย ตอนที่ 1 มโนทัศน์เกี่ยวกับโมเดลสมการเชิงโครงสร้าง และตอนที่ 2 มโนทัศน์เกี่ยวกับตัวบ่งชี้นำ ดังนี้

3.1 มโนทัศน์เกี่ยวกับโมเดลสมการเชิงโครงสร้าง

ผู้วิจัยได้รวบรวมเกี่ยวกับหลักการของโมเดลสมการเชิงโครงสร้าง จากการศึกษางานของ ศิริชัย กาญจนวาสี (2532) นงลักษณ์ วิรัชชัย (2538) และ Joreskog & Sorbom (1989) พบว่า โมเดลสมการเชิงโครงสร้าง เป็นโมเดลที่พัฒนามาจากโมเดลทางคณิตศาสตร์ ใช้สำหรับการทดสอบเชิงสาเหตุ สำหรับทฤษฎีที่มีความซับซ้อนมีความเป็นไปได้สำหรับข้อมูลที่ไม่ใช่การทดลอง เนื่องจากการวิจัยทางสังคมศาสตร์ส่วนใหญ่เป็นการศึกษาความจริงจากสภาพธรรมชาติ และปรากฏการณ์ต่าง ๆ ที่เกิดขึ้นในสังคมไม่มีการศึกษาตัวแปรใดตัวแปรหนึ่งเหมือนการวิจัยเชิงทดลอง

โมเดลสมการเชิงโครงสร้าง (structural equation models) หมายถึง โมเดลแสดงความสัมพันธ์โครงสร้างเชิงเส้นระหว่างตัวแปรซึ่งประกอบด้วยโมเดลการวัด (measurement model) และโมเดลสมการโครงสร้าง (structural equation model) ในโมเดลการวัดสามารถผ่อนคลายข้อตกลงเบื้องต้นในการวัดที่ไม่มีความคลาดเคลื่อนของโมเดลแบบดั้งเดิม โมเดลสมการโครงสร้างครอบคลุมความสัมพันธ์เชิงโครงสร้างแบบเส้นทุกรูปแบบ จึงสามารถวิเคราะห์ข้อมูลได้ไม่ว่าจะเป็นการวิเคราะห์ความสัมพันธ์เชิงสาเหตุโมเดลอิทธิพลทางเดียวหรืออิทธิพลย้อนกลับ

โมเดลการวัด (measurement model) เป็นโมเดลซึ่งแสดงความสัมพันธ์ระหว่างตัวแปรสังเกตได้กับตัวแปรแฝง โมเดลการวัดจะประกอบด้วยชุดของตัวแปรที่สังเกตได้ คือ ตัวแปรอิสระสังเกตได้และชุดของตัวแปรตามสังเกตได้ โมเดลนี้สามารถเขียนเป็นสมการในรูปเวคเตอร์จำแนกตามลักษณะของตัวแปรสังเกตได้ดังนี้

$X = (\Delta X)(\xi) + \delta$ สำหรับโมเดลการวัดตัวแปรอิสระสังเกตได้

$Y = (\Delta Y)(\eta) + \varepsilon$ สำหรับโมเดลการวัดตัวแปรตามสังเกตได้

โมเดลสมการโครงสร้าง (structural equation model) เป็นโมเดลที่ระบุความสัมพันธ์เชิงโครงสร้างระหว่างตัวแปรแฝง โมเดลนี้สามารถเขียนเป็นสมการในรูปเวคเตอร์ได้ดังนี้

$$\eta = (\beta)(\eta) + (\Gamma)(\xi) + \zeta$$

เวคเตอร์แต่ละเวคเตอร์ในสองสมการข้างต้นนี้ เป็นเวคเตอร์ของตัวแปรในโมเดลแทนด้วยอักษรกรีก ในทางปฏิบัติมักจะใช้อักษรภาษาอังกฤษแทนอักษรกรีกเหล่านี้เพื่อความสะดวกในการเขียนโมเดลมากขึ้น สามารถเขียนอักษรภาษาอังกฤษแทนอักษรกรีกเหล่านั้น และอธิบายความหมายของแต่ละเวคเตอร์ดังนี้ (นงลักษณ์ วิรัชชัย, 2542)

X แทน เวคเตอร์ของตัวแปรภายนอกสังเกตได้ X ขนาด (NX x 1)

Y แทน เวคเตอร์ของตัวแปรภายในสังเกตได้ Y ขนาด (NY x 1)

ξ แทน เวคเตอร์ตัวแปรภายนอกแฝง K ขนาด (NK x 1)

η แทน เวคเตอร์ตัวแปรภายในแฝง E ขนาด (NE x 1)

δ แทน เวคเตอร์ความคลาดเคลื่อน d ในการวัดตัวแปร X ขนาด (NX x 1)

ε แทน เวคเตอร์ความคลาดเคลื่อน e ในการวัดตัวแปร Y ขนาด (NY x 1)

ζ แทน เวคเตอร์ความคลาดเคลื่อน z ของตัวแปร E ขนาด (NE x 1)

นอกจากนี้ในสมการดังกล่าวมีเมทริกซ์พารามิเตอร์อิทธิพลเชิงสาเหตุ หรือสัมประสิทธิ์การถดถอยรวม 4 เมทริกซ์ที่ควรทราบ และเมทริกซ์พารามิเตอร์ความแปรปรวน-ความแปรปรวนร่วมอีก 4 เมทริกซ์ที่ไม่ได้นำเสนอไว้ในสมการข้างต้น แต่จะขอกล่าวถึงเพื่อความเข้าใจในการศึกษาเกี่ยวกับโมเดลลิสเรล มีรายละเอียดดังนี้

ΔX = LX แทน เมทริกซ์สัมประสิทธิ์การถดถอยของ X บน K ขนาด (NX x NK)

ΔY = LY แทน เมทริกซ์สัมประสิทธิ์การถดถอยของ Y บน E ขนาด (NY x NE)

Γ = GA แทน เมทริกซ์อิทธิพลเชิงสาเหตุจาก K ไป E ขนาด (NE x NK)

β = BE แทน เมทริกซ์อิทธิพลเชิงสาเหตุระหว่าง E ขนาด (NE x NE)

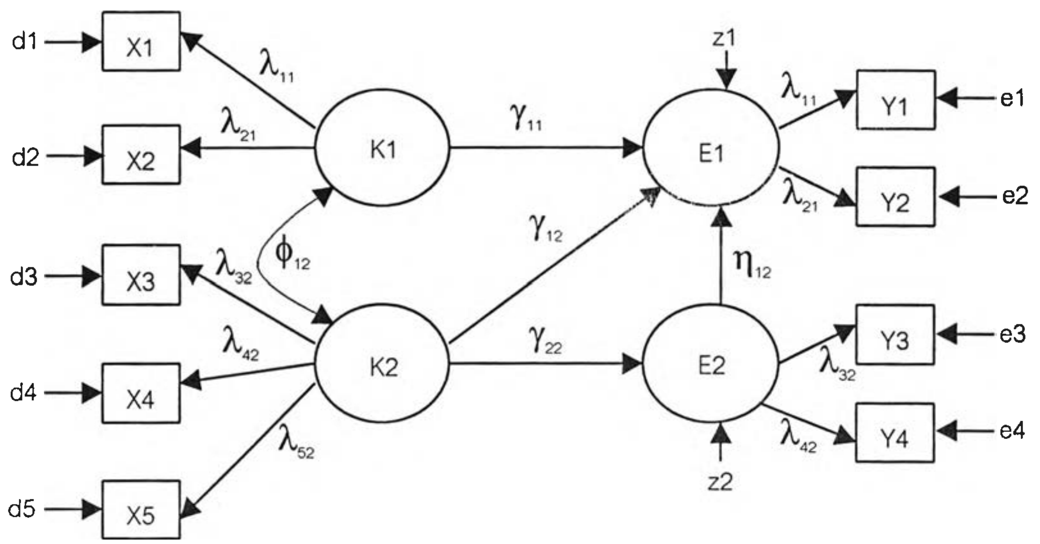
Φ = PH แทน เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมระหว่างตัวแปรภายในแฝง K ขนาด (NK x NK)

ψ = PS แทน เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน z ขนาด (NE x NE)

Θ_δ = TD แทน เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมระหว่างความคลาดเคลื่อน d ขนาด (NX x NX)

Θ_{ε} = TE แทน เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมระหว่าง ความคลาดเคลื่อน e ขนาด (NY x NY)

ในการนำข้อมูลในธรรมชาติมาเพื่อศึกษาความสัมพันธ์ในเชิงสาเหตุต้องทำด้วยความรอบคอบ โดยนักวิจัยต้องดำเนินการตามขั้นตอนที่สำคัญคือ ขั้นแรก คัดเลือกตัวแปรมาศึกษาให้ครบถ้วน จากการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง และจากการวิเคราะห์เชิงตรรกะของผู้วิจัย ขั้นที่สอง สร้างโมเดลแสดงความสัมพันธ์ระหว่างตัวแปร และขั้นสุดท้าย นำโครงสร้างความสัมพันธ์มาตรวจสอบความสอดคล้องกับข้อมูลเชิงประจักษ์ เพื่อความเข้าใจเกี่ยวกับโมเดลขอนำเสนอตัวอย่างโมเดลสมการเชิงโครงสร้างแบบเต็มรูป ดังแผนภาพ 7



แผนภาพ 7 ตัวอย่างโมเดลสมการเชิงโครงสร้างแบบเต็มรูป

ในการวิเคราะห์โมเดลสมการเชิงโครงสร้างหรือโมเดลลิสเรลนี้ ควรทราบถึงข้อตกลงเบื้องต้นสำหรับโมเดลดังกล่าวนี้ ซึ่งนงลักษณ์ วิรัชชัย (2542) ได้กล่าวสรุปไว้ดังนี้

1.ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งหมดในโมเดลเป็นความสัมพันธ์เชิงเส้นแบบบวก และเป็นความสัมพันธ์เชิงสาเหตุ

2. ลักษณะการแจกแจงของตัวแปรทั้งตัวแปรภายนอกและตัวแปรภายในและความคลาดเคลื่อนต้องเป็นการแจกแจงแบบปกติ ความคลาดเคลื่อน e, d และ z ต้องมีค่าเฉลี่ยเป็น 0 กรณีเป็นตัวแปรทวิวิภาค (dichotomous variables) ที่มีค่าเฉลี่ยใกล้ 0.5 ให้ค่าประมาณพารามิเตอร์ที่มีความแกร่ง และสามารถนำมาวิเคราะห์โมเดลนี้ได้

3. ความคลาดเคลื่อน e และตัวแปรแฝง E เป็นอิสระต่อกัน

4. ความคลาดเคลื่อน d และตัวแปรแฝง K เป็นอิสระต่อกัน

5. ความคลาดเคลื่อน z และตัวแปรแฝง K เป็นอิสระต่อกัน

6. ความคลาดเคลื่อน e, d และ z เป็นอิสระต่อกัน

7.สำหรับกรณีการวิเคราะห์ข้อมูลอนุกรมเวลาที่มีการวัดข้อมูลมากกว่า 2 ครั้ง การวัดตัวแปรต้องไม่ได้รับอิทธิพลจากช่วงเวลาเหลื่อม (time lag) ระหว่างการวัด

ในการดำเนินงานเพื่อวิเคราะห์โมเดลสมการเชิงโครงสร้าง (SEM) แบ่งได้เป็น 6 ขั้นตอน (นงลักษณ์ วิรัชชัย, 2542) ดังนี้

ขั้นตอนที่ *1* การกำหนดข้อมูลจำเพาะของโมเดล เป็นการกำหนดค่าเมทริกซ์ทั้ง 8 เมทริกซ์ ให้สอดคล้องกับโมเดลที่ใช้ในการวิเคราะห์ เพื่อจะนำไปใช้ในการเขียนโปรแกรมสำหรับการประมาณค่าพารามิเตอร์ด้วยโปรแกรมลิสเรลต่อไป การกำหนดค่าให้กับเมทริกซ์ทั้ง 8 เมทริกซ์นั้นสามารถทำได้ 3 แบบด้วยกัน คือ 1) พารามิเตอร์กำหนด (fixed parameter) เป็นพารามิเตอร์ที่ผู้วิเคราะห์ต้องการให้มีขนาดอิทธิพลเป็น 0 2) พารามิเตอร์บังคับ (constrained parameter) เป็นพารามิเตอร์ที่ต้องประมาณค่า แต่มีเงื่อนไขกำหนดให้พารามิเตอร์มีค่าเฉพาะคงที่ เช่น มีค่าเป็น 1 หรือมีค่าเท่ากับพารามิเตอร์ตัวอื่น ๆ และ 3) พารามิเตอร์อิสระ (free parameter) เป็นพารามิเตอร์ที่ต้องการประมาณค่าและไม่ได้บังคับให้มีค่าเป็นค่าใดค่าหนึ่ง

ขั้นตอนที่ *2* การระบุความเป็นได้ค่าเดียวของโมเดล เป็นขั้นตอนที่สืบเนื่องมาจากขั้นตอนแรก เหตุเพราะว่าโมเดลสมการเชิงโครงสร้างเมื่อนำมาประมาณค่าพารามิเตอร์จะต้องมีการระบุความเป็นได้ค่าเดียวของพารามิเตอร์ก่อนที่จะประมาณค่า ฉะนั้นการระบุความเป็นได้ค่าเดียวของโมเดลนั้นคือการระบุว่าโมเดลนั้นสามารถนำมาประมาณค่าพารามิเตอร์ได้เป็นค่าเดียวหรือไม่ ถ้าจำนวนสมการเท่ากับจำนวนพารามิเตอร์ที่ไม่ทราบค่าแต่ละตัว เรียกโมเดลนั้นว่าโมเดลระบุความเป็นได้ค่าเดียวได้พอดี ถ้าจำนวนสมการมากกว่าจำนวนพารามิเตอร์ที่ไม่ทราบค่าในโมเดลเรียกโมเดลนั้นว่า โมเดลระบุความเป็นได้ค่าเดียวเกินพอดี และถ้าจำนวนสมการน้อยกว่าจำนวนพารามิเตอร์ที่ไมทราบค่าเรียกโมเดลนั้นว่า โมเดลระบุความเป็นได้ค่าเดียวไม่พอดี และโมเดลประเภทนี้จะไม่สามารถประมาณค่าพารามิเตอร์ได้ (Pedhazur, 1982 อ้างถึงในนงลักษณ์ วิรัชชัย, 2542)

ขั้นตอนที่ *3* การประมาณค่าพารามิเตอร์จากโมเดล ใช้หลักการนำเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมจากกลุ่มตัวอย่าง (ข้อมูลเชิงประจักษ์) มาเปรียบเทียบกับเมทริกซ์ความแปรปรวน-ความแปรปรวนร่วมที่สร้างขึ้นจากพารามิเตอร์ที่ประมาณได้ ถ้ามีค่าใกล้เคียงกัน หมายความว่าโมเดลที่เป็นสมมติฐานมีความกลมกลืนกับข้อมูลเชิงประจักษ์

ขั้นตอนที่ 4 การทดสอบเปรียบเทียบความสอดคล้องระหว่างข้อมูลเชิงประจักษ์กับโมเดลที่เป็นสมมติฐาน สถิติวัดระดับความกลมกลืนที่สำคัญมี 4 ค่า (นงลักษณ์ วิรัชชัย, 2538 ; สุปรียา ไข่มุกข์, 2540 ; สายรุ้ง แสงแจ้ง, 2540 ; Joreskog & Sorbom, 1989) ได้แก่

(1) ค่าสถิติไคสแควร์ (chi-square statistic) การคำนวณค่าไค-สแควร์ คำนวณจากผลคูณขององศาอิสระ (degree of freedom) กับค่าฟังก์ชันความกลมกลืน ถ้าค่าสถิติไค-สแควร์มีค่าสูงมาก แสดงว่าฟังก์ชันความกลมกลืนนั้นมีความแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติ นั่นคือโมเดลไม่มีความสอดคล้องกับข้อมูลเชิงประจักษ์ ถ้าค่าสถิติไค-สแควร์มีค่าต่ำมากและใกล้ศูนย์ แสดงว่าโมเดลสอดคล้องกับข้อมูลเชิงประจักษ์

(2) ดัชนีวัดระดับความกลมกลืน (goodness of fit index = GFI) เป็นอัตราส่วนของผลต่างระหว่างฟังก์ชันความกลมกลืนก่อนปรับโมเดล ค่าดัชนีที่เข้าใกล้ 1.00 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

(3) ดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (adjusted goodness of fit index = AGFI) เป็นการนำค่าดัชนี GFI มาปรับแก้โดยคำนึงถึงขนาดขององศาอิสระ (degree of freedom) จำนวนตัวแปรและขนาดของกลุ่มตัวอย่าง ค่าดัชนีที่เข้าใกล้ 1.00 แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

(4) ดัชนีรากกำลังสองเฉลี่ยของเศษ (root mean squared residual = RMR) เป็นดัชนีที่บอกขนาดของเศษที่เหลือโดยเฉลี่ยจากการเปรียบเทียบระดับความกลมกลืนของโมเดลกับข้อมูลเชิงประจักษ์ ค่าดัชนียิ่งเข้าใกล้ศูนย์ แสดงว่าโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์

ขั้นตอนที่ 5 การปรับโมเดล ในการปรับโมเดลพิจารณาจากค่าดัชนีดัดแปรโมเดล (model modification indicies : MI) และค่าที่คาดว่าจะเปลี่ยนแปลง (expected change : EPC) ค่าทั้งสองนี้เป็นค่าสถิติสำหรับการกำหนดค่าพารามิเตอร์ในโมเดลให้เป็นพารามิเตอร์อิสระเพื่อทำให้ค่าไค-สแควร์ลดลงและโมเดลมีความสอดคล้องกับข้อมูลเชิงประจักษ์มากขึ้น โดยพิจารณาพารามิเตอร์ที่มีค่า MI และ EPC สูงกว่าพารามิเตอร์อื่น ๆ ที่ยังไม่ได้กำหนดไว้ในโมเดล ทั้งนี้ในการปรับนั้นต้องพิจารณาทฤษฎีและเอกสารงานวิจัยที่เกี่ยวข้องร่วมด้วย

ขั้นตอนที่ 6 การแปลความหมายผลการวิเคราะห์

ในการวิจัยครั้งนี้ผู้วิจัยประยุกต์ใช้การพยากรณ์ด้วยวิธีบ๊อกซ์และเจนกิ้นส์กับโมเดลสมการเชิงโครงสร้าง โดยการนำโมเดล ARMA(p,q) ที่ได้จากการเลือกด้วยวิธีบ๊อกซ์และเจนกิ้นส์ มาตรวจสอบความเหมาะสมของโมเดลที่เลือกมาด้วยโมเดลสมการเชิงโครงสร้าง โมเดล ARMA(p,q) ที่ได้นั้นจะต้องนำมาเขียนในอยู่ในรูปของโมเดลสมการเชิงโครงสร้าง โมเดลดังกล่าวจะต้องมีลักษณะที่คงที่เท่านั้น ตัวอย่างลักษณะการเขียนโมเดล ARMA(p,q) ให้อยู่ในลักษณะของโมเดลสมการเชิงโครงสร้าง ผู้วิจัยได้นำเสนอไว้ในบทที่ 1 ดังแผนภาพ 2 เมื่อได้โมเดลสมการเชิงโครงสร้างแล้ว ก็ดำเนินการวิเคราะห์ตามขั้นตอนของโมเดลสมการเชิงโครงสร้างเพื่อตรวจสอบความเหมาะสมของโมเดลที่สร้างขึ้น

3.2 มโนทัศน์เกี่ยวกับตัวบ่งชี้นำ

ตัวบ่งชี้นำ (leading indicator) เป็นอนุกรมเวลาที่มีความเกี่ยวข้องสัมพันธ์กับอนุกรมเวลาที่สนใจศึกษาหรือทำการพยากรณ์ ใช้ช่วยในการพยากรณ์ค่าในอนาคตของอนุกรมเวลาที่สนใจทำการพยากรณ์ (ทรงศิริ แต้สมบัติ, 2539 ; SPSS Inc., 1994) โดยศึกษาความสัมพันธ์ของอนุกรมเวลาชุดที่สนใจ $\{Y_t\}$ กับอนุกรมเวลาอีกชุดหนึ่ง $\{X_t\}$ ซึ่งต่างก็มีแบบแผนของวัฏจักรไม่แตกต่างกัน แต่ช่วงเวลาการเริ่มและจบของวัฏจักรอาจต่างกัน ดังนั้นเมื่อทราบความสัมพันธ์ระหว่างอนุกรมเวลาทั้งสอง นั่นคือจะทำการประมาณค่าหรือการพยากรณ์จากการทราบว่า X_{t-k} มีส่วนในการอธิบายความผันแปรของ Y_t

การหาค่า k ที่เหมาะสมที่ X_{t-k} มีส่วนในการอธิบายความผันแปรของ Y_t จะพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ข้ามกลุ่มของ X_t และ Y_t ที่เวลาล้าหลัง k (sample cross-correlation coefficient of lag k : $r_{xy}(k)$) จะเรียกค่าสัมประสิทธิ์สหสัมพันธ์ข้ามกลุ่มของ X_t และ Y_t ที่เวลาล้าหลัง k ว่าค่า ccf ที่เวลาล้าหลัง k มีสูตรในการคำนวณดังนี้

$$r_{xy}(k) = \frac{\sum\left(X_t - \bar{X}\right)\left(Y_{t+k} - \bar{Y}\right)}{\sqrt{\sum\left(X_t - \bar{X}\right)^2 \sum\left(Y_t - \bar{Y}\right)^2}}$$

การอธิบายความหมายของ $r_{xy}(k)$ จะเป็นในทำนองเดียวกับ r_{xy} หรือค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปร X และ Y นั่นคือ $r_{xy}(k)$ จะบอกทิศทางและขนาดของสหสัมพันธ์ระหว่าง X_{t-k} และ Y_t การพิจารณาว่า $r_{xy}(k)$ มาจาก $\rho_{xy}(k)$ ที่มีค่าต่างจาก 0 หรือไม่ สำหรับการทดสอบสมมติฐาน $H_0 : \rho_{xy}(k) = 0$ กับ $H_1 : \rho_{xy}(k) \neq 0$ จะใช้ตัวทดสอบสถิติ $r_{xy}(k)$ การพิจารณาว่าจะรับหรือปฏิเสธ H_0 จะใช้ช่วงวิกฤติ $|r_{xy}(k)| \geq \frac{2}{\sqrt{n-k}}$ เมื่อ n-k เป็นจำนวนคู่ที่ลดลงของ X และ Y จากทั้งหมด n คู่ ซึ่งเป็นการทดสอบที่ระดับนัยสำคัญประมาณ .05 เมื่อปฏิเสธ H_0 จะสรุปได้ว่า X_{t-k} มีส่วนในการอธิบายการเคลื่อนไหวของ Y_t

กรณีที่ทราบว่าการเคลื่อนไหวของอนุกรมเวลา $\{Y_t\}$ และ $\{X_t\}$ มีอิทธิพลของวัฏจักรเข้ามาเกี่ยวข้องด้วย และ X_{t-k} มีส่วนในการอธิบายความผันแปรของ Y_t ดังนั้นหากทราบลักษณะการเคลื่อนไหวของ $\{X_t\}$ ในอดีตจะทำให้ประมาณค่าหรือพยากรณ์ค่าของ Y_t ได้ สมการพยากรณ์ที่จะสร้างเมื่อทราบว่า X_{t-k} มีส่วนในการอธิบายความผันแปรของ Y_t ทำได้ 2 ลักษณะดังนี้

1. กำหนดรูปแบบการถดถอยเส้นตรง $Y_t = \beta_0 + \beta_1 X_{t-k} + \varepsilon_t$ และหาสมการพยากรณ์ $Y_t = b_0 + b_1 X_{t-k}$ โดยวิธีกำลังสองน้อยที่สุด

2. สร้างสมการพยากรณ์แบบเอกซ์โปเนนเชียล จากการทราบลักษณะการเปลี่ยนแปลงของค่า Y_t และการเปลี่ยนแปลงของค่า X_t ในแต่ละช่วงเวลาและทราบว่า X_{t-k} มีส่วนในการอธิบายความผันแปรของ Y_t

$$\hat{Y}_{t+1} = Y_t \left[1 + \frac{(\bar{Y}|\%\Delta Y_t|)}{(\bar{X}|\%\Delta X_t|)} \Delta X_{t+1-k} \right]$$

โดย

$\% \Delta Y_t$ = % การเปลี่ยนแปลงจาก Y_{t-1} เป็น Y_t

$\% \Delta X_t$ = % การเปลี่ยนแปลงจาก X_{t-1} เป็น X_t

ΔX_{t+1-k} = สัดส่วนการเปลี่ยนแปลงจาก X_{t-k} เป็น X_{t+1-k}

ในการวิจัยครั้งนี้ ผู้วิจัยได้นำโมเดล transfer function มาใช้ในการวิเคราะห์ตัวบ่งชี้นำ โมเดล transfer function เป็นรูปแบบที่ใช้พยากรณ์ค่าในอนาคตของอนุกรมเวลา $\{y_t\}$

ด้วยค่าในอดีตของอนุกรมเวลา $\{y_t\}$ และค่าของอนุกรมเวลาอื่นที่เกี่ยวข้องเรียกว่าอนุกรมเวลา $\{x_t\}$ ซึ่งอาจจะมีมากกว่าหนึ่งอนุกรมเวลา กรณีที่ $\{x_t\}$ และ $\{y_t\}$ เป็นอนุกรมเวลาที่ไม่คงที่ จะหา $\{z_t^x\}$ และ $\{z_t^y\}$ ซึ่งเป็นอนุกรมเวลาใหม่ที่มีลักษณะที่คงที่ ที่ได้จากการแปลง $\{x_t\}$ และ $\{y_t\}$ โดยการหาผลต่าง

$$Z_t^y = \theta_0 + \omega_0 \frac{\left(1-\omega_1 B - \ldots - \omega_s B^s\right)}{\left(1-\delta_1 B - \ldots - \delta_r B^r\right)} B^b Z_t^x + N_t$$

หรือ

$$Z_t^y = \theta_0 + \omega_0 \frac{\omega(B)}{\delta(B)} B^b Z_t^x + N_t$$

เมื่อ $\omega(B) = 1-\omega_1 B - \omega_2 B^2 - \ldots - \omega_s B^s$, $\delta(B) = 1-\delta_1 B - \delta_2 B^2 - \ldots - \delta_r B^r$, θ_0 และ ω_0 เป็นค่าคงที่ และ N_t เป็นความคลาดเคลื่อน

สมการ transfer function สร้างจากรูปแบบ transfer function ที่กำหนดขึ้นโดยเริ่มจากการทำ prewhitening อนุกรมเวลา $\{x_t\}$ และ $\{y_t\}$ เป็นอนุกรมเวลาใหม่ $\{\alpha_t\}$ และ $\{\beta_t\}$ จากสัมประสิทธิ์สหสัมพันธ์ข้ามกลุ่มของอนุกรมเวลา $\{\alpha_t\}$ และ $\{\beta_t\}$ นำไปกำหนดรูปแบบ transfer function ขั้นตอนการสร้างสมการ transfer function มีรายละเอียดดังนี้

1. การกำหนดโมเดล มีรายละเอียดดังนี้

1.1 กรณีอนุกรมเวลา $\{x_t\}$ มีลักษณะที่ไม่คงที่ สร้างอนุกรมเวลาใหม่ $\{z_t^x\}$ ที่มีลักษณะคงที่ โดยการหาผลต่าง และ/หรือหาค่าลอการิทึม

1.2 สร้างอนุกรมเวลา $\{z_t^y\}$ ที่มีลักษณะคงที่โดยการหาผลต่าง และ/หรือหาค่าลอการิทึม ในทำนองเดียวกันกับการสร้างอนุกรมเวลาใหม่ $\{z_t^x\}$

1.3 กำหนดโมเดล ARMA(p,q) ให้กับ $\{z_t^x\}$ และประมาณค่าพารามิเตอร์ในโมเดล ARMA ของอนุกรมเวลา $\{z_t^x\}$

1.4 จากโมเดล ARMA(p,q) ของ z_t^x ที่ได้ กำหนดให้เป็นโมเดลของ z_t^y ด้วย หาความคลาดเคลื่อนจากโมเดลทั้งสองได้เป็น α_t และ β_t จะเรียกการหาอนุกรมเวลาของความคลาดเคลื่อน $\{\alpha_t\}$ และ $\{\beta_t\}$ จาก $\{z_t^x\}$ และ $\{z_t^y\}$ ตามลำดับ ว่าการทำ prewhitening กับอนุกรมเวลา $\{z_t^x\}$ และ $\{z_t^y\}$

1.5 หาค่าสัมประสิทธิ์สหสัมพันธ์ข้ามกลุ่มระหว่าง α_t และ β_t ที่เวลาล้าหลัง k นั่นคือหาค่า $r_k(\alpha_t, \beta_t)$

1.6 จาก $r_k(\alpha_t, \beta_t)$ พิจารณาหาค่า r, s และ b ซึ่งค่า r, s และ b ที่ได้จะให้รูปแบบ preliminary transfer function เป็น

$$Z_t^y = \theta_0 + \omega_0 \frac{(1-\omega_1 B - \ldots - \omega_s B^s)}{(1-\delta_1 B - \ldots - \delta_r B^r)} B^b Z_t^x + N_t$$

โดย b = k ค่าแรกของ $r_k(\alpha_t, \beta_t)$ ที่ไม่เป็น 0 หรือจำนวนช่วงเวลาก่อนที่ X_t จะเริ่มมีผลต่อ Y_t

s = จำนวน lag ที่เริ่มนับจาก lag ที่ b ที่เกิดก่อนการ dies down ของ $r_k(\alpha_t, \beta_t)$ หรือจำนวนค่า Z_t^x ที่มีผลต่อค่า Z_t^y

r = ค่าที่บอกลักษณะของ $r_k(\alpha_t, \beta_t)$ หลังจาก lag ที่ b+s ซึ่งอาจจะเป็นแบบ damped exponential หรือแบบ sine-wave โดย r = 1 เมื่อ dies down แบบ damped exponential และ r = 2 เมื่อ dies down แบบ sine-wave

2. การประมาณค่าพารามิเตอร์ มีรายละเอียดดังนี้

2.1 จากโมเดล preliminary transfer function หาค่าประมาณของพารามิเตอร์และหาอนุกรมเวลาของความคลาดเคลื่อน $\{N_t\}$

2.2 หาค่า $r_k(N_t)$ และ $r_{kk}(N_t)$ หรือสร้างคอเรลโลแกรมของ $\{N_t\}$ เพื่อพิจารณาว่า N_t มีความเป็นอิสระต่อกันหรือไม่ กรณีที่ไม่เป็นอิสระต่อกันจะกำหนดโมเดล ARMA (p,q) ให้กับ $\{N_t\}$

2.3 สร้างโมเดล final transfer function จากรูปแบบ preliminary transfer function ที่ได้ในขั้นแรก และโมเดล ARMA ของ $\{N_t\}$ ที่ได้จากข้อข้างต้นนี้

3. การตรวจสอบโมเดล เป็นสิ่งที่ต้องกระทำเพื่อทำการตรวจสอบว่าโมเดล final transfer function ที่กำหนดเหมาะสมหรือไม่ โดยการทดสอบพารามิเตอร์ในโมเดลและพิจารณาคอเรลโลแกรมของความคลาดเคลื่อน

4. การพยากรณ์ ทำในทำนองเดียวกันกับการพยากรณ์ด้วยวิธีบ็อกซ์และเจนกิ้นส์ นั่นคือสามารถพยากรณ์ได้ทั้งแบบจุด และแบบช่วง

ตอนที่ 4 การตรวจสอบความคลาดเคลื่อนในการพยากรณ์

การพยากรณ์ข้อมูลอนุกรมเวลาและการประมาณค่าพารามิเตอร์ จำเป็นที่จะต้องมีการตรวจสอบว่ามีความถูกต้องและความคลาดเคลื่อนมากน้อยเพียงใด โดยทั่วไปมี 2 ขั้นตอน คือ

1. การตรวจสอบฟังก์ชันอัตตะสหสัมพันธ์ของโมเดลอนุกรมเวลา ที่สามารถเปรียบเทียบกับฟังก์ชันอัตตะสหสัมพันธ์ของกลุ่มตัวอย่างในอนุกรมแรก

2. ค่าวัดความถูกต้องของการพยากรณ์ (Pindyck and Rubinfeld, 1985 อ้างถึงใน บำเพ็ญ ปิดชิด, 2540 และ ทรงศิริ แต้สมบัติ, 2539) ค่าความคลาดเคลื่อนของการพยากรณ์ (forecast error : e_t) เป็นผลต่างของค่าจริงและค่าพยากรณ์ $\left(e_t = Y_t - \hat{Y}_t\right)$ ความคลาดเคลื่อนจะมากถ้าค่าจริงห่างจากค่าพยากรณ์มาก และจะน้อยถ้าค่าพยากรณ์ใกล้เคียงกับค่าจริง ค่าความคลาดเคลื่อนของการพยากรณ์จะมีค่าเท่ากับ 0 เมื่อพยากรณ์ในช่วงเวลาล่วงหน้า ค่าวัดความถูกต้องของการพยากรณ์สามารถจำแนกได้เป็น 2 กลุ่ม คือ ค่าวัดความถูกต้องของการพยากรณ์ในรูปความแปรปรวน หรือค่าวัดในรูปค่าเฉลี่ยของผลรวมกำลังสองของความคลาดเคลื่อน และค่าวัดความถูกต้องของการพยากรณ์ในรูปความคลาดเคลื่อนเฉลี่ย หรือค่าเฉลี่ยของค่าสมบูรณ์ของความคลาดเคลื่อนในรูปร้อยละ แต่พบว่าไม่มีค่าวัดความถูกต้องของการพยากรณ์ค่าใดดีที่สุด (Armstrong, 1982) ฉะนั้นการเลือกค่าวัดความถูกต้องของการพยากรณ์ควรเลือกใช้ตามสถานการณ์ แต่ในการวิจัยครั้งนี้ผู้วิจัยเลือกใช้ค่าวัดความถูกต้องของการพยากรณ์ 6 ค่า ได้แก่ RMSE, MAPE, GMRAE, MdAPE, MdRAE และ Percent Better โดยค่าวัดความถูกต้องของการพยากรณ์ 5 ค่า ซึ่งได้แก่ RMSE, GMRAE, MdAPE, MdRAE และ Percent Better ไม่มีเกณฑ์ที่ใช้ในการพิจารณาความเหมาะสม ในขณะที่ค่าวัดความถูกต้องของการพยากรณ์ด้วย MAPE มีเกณฑ์ในการพิจารณาความเหมาะสมดังนี้ (Lewis, 1982)

ค่าที่ได้	*ความหมาย*
น้อยกว่า 10%	มีความถูกต้องสูง
10 – 20%	มีความถูกต้อง
20 – 50%	ยังยอมรับได้
มากกว่า 50%	ไม่มีความถูกต้อง

ค่าวัดความถูกต้องของการพยากรณ์แต่ละค่ามีสูตรในการคำนวณ ดังนี้

กำหนดให้ m = วิธีการพยากรณ์

F_m = ค่าพยากรณ์จากวิธี m

F_{rw} = ค่าพยากรณ์จากวิธี random walk

Z = ค่าสังเกตที่เกิดขึ้นจริง

S = จำนวนอนุกรมเวลา

แบบที่ 1 รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE)

$$RMSE_m = \left(\frac{\sum_{s=1}^{s} (F_m - Z)^2}{S} \right)^{\frac{1}{2}}$$

แบบที่ 2 ค่าเฉลี่ยของค่าสมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (MAPE)

$$APE_m = \left| \frac{F_m - Z}{Z} \right| x\,100$$

$$MAPE_{m_i} = \frac{\sum_{i=1}^{s} APE_{m_i}}{S}$$

แบบที่ 3 ค่าเฉลี่ยเรขาคณิตของความคลาดเคลื่อนสมบูรณ์สัมพัทธ์ (GMRAE)

$$RAE_m = \left| \frac{F_m - Z}{F_{rw} - Z} \right|$$

$$GMRAE_m = \left(\prod_{i=1}^{s} RAE_{m_i} \right)^{\frac{1}{s}}$$

แบบที่ 4 ค่ามัธยฐานของค่าสมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (MdAPE)

$$APE_m = \left| \frac{F_m - Z}{Z} \right| x\,100$$

เมื่อนำค่าสังเกต APE_m มาเรียงลำดับ จะได้

$$MdAPE_m = \frac{S+1}{2}$$

เมื่อ S เป็นจำนวนคี่

$$MdAPE_m = \left(\frac{S}{2}\right) + 1$$

เมื่อ S เป็นจำนวนคู่

แบบที่ 5 ค่ามัธยฐานของค่าสมบูรณ์สัมพัทธ์ของความคลาดเคลื่อน (MdRAE)

$$RAE_m = \left|\frac{F_m - Z}{F_{rw} - Z}\right|$$

เมื่อนำค่าสังเกต RAE_m มาเรียงลำดับ จะได้

$$MdRAE_m = \frac{S+1}{2}$$

เมื่อ S เป็นจำนวนคี่

$$MdRAE_m = \left(\frac{S}{2}\right) + 1$$

เมื่อ S เป็นจำนวนคู่

แบบที่ 6 ร้อยละที่ดีกว่า (Percent Better)

$$Percent\ Better = \frac{\sum_{i=1}^{s} j_i}{S} x\,100$$

เมื่อ $j_s \begin{cases} 1 & \text{ถ้า } |F_m - Z| < |F_{rw} - Z| \\ 0 & \text{ในกรณีอื่น ๆ} \end{cases}$

ตอนที่ 5 งานวิจัยที่เกี่ยวข้องกับการพยากรณ์

สาระในตอนนี้ผู้วิจัยนำเสนองานวิจัยที่เกี่ยวข้องกับการพยากรณ์ที่ใช้เทคนิคการพยากรณ์เชิงปริมาณ แบ่งออกได้เป็น 2 ประเภท คือ การพยากรณ์โดยใช้โมเดลความสัมพันธ์เชิงเหตุและผล และข้อมูลอนุกรมเวลา มีดังนี้

งานวิจัยที่ใช้การพยากรณ์ความสัมพันธ์เชิงเหตุและผล ได้แก่ งานวิจัยของนงลักษณ์ วิรัชชัย (2513) พรรณมาศ คันฉาย (2513) บุญธรรม กิจปรีดาบริสุทธิ์ (2513) นิตยา ภัสสรศิริ (2513) และสมหวัง พิธิยานุวัฒน์ (2514) เป็นลักษณะการพยากรณ์ความต้องการอัตรา

กำลังคนในสาขาต่าง ๆ ของประเทศไทยโดยใช้สูตรเศรษฐมิติทางการศึกษา นอกจากนี้มีงานวิจัยที่นำโมเดลเศรษฐมิติมาประยุกต์ใช้กับการพยากรณ์ทางการศึกษา ได้แก่ อังคณา พัฒนาผลไพบูลย์ (2531) ได้ทำการสร้างโมเดลเศรษฐมิติเพื่อพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ปีการศึกษา 2531 – 2540 โดยมีขั้นตอนการดำเนินการวิจัยดังนี้ ในขั้นแรก ก็กำหนดโมเดลในการพยากรณ์จำนวนครูซึ่งได้คัดเลือกตัวแปรที่มีความสัมพันธ์กับจำนวนครู จำนวน 11 ตัวแปร และเก็บรวบรวมข้อมูลจำนวน 16 ปี จากแหล่งทุติยภูมิมาทำแผนการกระจัดกระจาย เพื่อดูลักษณะความสัมพันธ์ของตัวแปรอธิบายกับตัวแปรตามแล้วคำนวณหาสัมประสิทธิ์สหสัมพันธ์ หลังจากนั้นก็คัดตัวแปรที่มีความสัมพันธ์อย่างชัดเจนออกไป แล้วเลือกตัวแปรที่สำคัญกว่าไว้ 5 ตัว ขั้นที่ 2 ประมาณค่าพารามิเตอร์ในโมเดล ขั้นที่ 3 ประเมินค่าที่ประมาณได้ของพารามิเตอร์โดยใช้ F-test เมื่อพบว่ามีนัยสำคัญจึงทดสอบค่าพารามิเตอร์โดยใช้ t-test และขั้นสุดท้าย ทำการประเมินประสิทธิภาพของโมเดลการพยากรณ์โดยตรวจสอบกับข้อมูลในอดีต แล้วนำโมเดลไปพยากรณ์จำนวนครู ผลการพยากรณ์เมื่อใช้โมเดลที่ 1 และโมเดลที่ 2 ได้จำนวนครู 96,650-115,977 คน และ 95,654-115,101 คน ตามลำดับ

นงนุช อินทรวงษ์โชติ (2538) ได้ทำการพยากรณ์จำนวนครูโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ปีการศึกษา 2538 – 2550 แต่มีตัวแปรเพิ่มขึ้นอีก 6 ตัวแปร และมีการตรวจสอบประสิทธิภาพของโมเดลพยากรณ์กับข้อมูลในอนาคตด้วย ผลการพยากรณ์เมื่อใช้โมเดลที่ 1 และโมเดลที่ 2 ได้จำนวนครู 131,513-175,132 คน และ 131,600-175,422 คน ตามลำดับ

งานวิจัยที่ใช้วิธีการวิเคราะห์อนุกรมเวลา (time series methods) ในการพยากรณ์ เช่นงานวิจัยของ วันพร เหลืองอาภาพงศ์ (2519) ใช้รูปแบบอนุกรมเวลาของบ๊อกซ์และเจนกิ้นส์ในการคาดคะเนปริมาณการส่งออกของสินค้าเกษตรกรรมที่สำคัญของไทย ได้แก่ ข้าว ยาง และข้าวโพด ตั้งแต่เดือนมกราคมถึงเดือนธันวาคมในปี 2519 โดยอาศัยข้อมูลอนุกรมเวลาเป็นรายเดือนตั้งแต่ปี พ.ศ. 2513 – 2518 พบว่า อนุกรมเวลาของปริมาณส่งออกข้าวและยางเป็นอนุกรมเวลาคงที่ รูปแบบที่มีความเหมาะสมกับข้อมูล คือ ARIMA(2,0,0) ส่วนข้าวโพดเป็นอนุกรมเวลาฤดูกาล รูปแบบที่มีความเหมาะสมกับข้อมูล คือ ARIMA(0,1,1)xSARIMA$(0,1,1)_{12}$ ซึ่งเมื่อนำเอามาคาดคะเนพบว่าอยู่ในเกณฑ์เฉลี่ยไม่แตกต่างจากปีที่ผ่านมามากนัก ในปี พ.ศ. 2522 บุษบา พิกุลผล ได้เปรียบเทียบรูปแบบที่ใช้คาดคะเนจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยโดยใช้การวิเคราะห์อนุกรมเวลากับเทคนิคของบ๊อกซ์และเจนกิ้นส์ โดยใช้ข้อมูลที่เป็นรายเดือนของจำนวนนักท่องเที่ยวที่เข้ามาในประเทศไทยตั้งแต่ปี พ.ศ. 2506 – 2520 พบว่ารูปแบบที่เหมาะสม คือ รูปแบบที่อธิบายการเปลี่ยนแปลงเนื่องจากวัฏจักร และเหตุการณ์ผิดปกติร่วมกัน

โดยใช้เทคนิคของบ๊อกซ์และเจนกิ้นส์ และปรับด้วยค่าแนวโน้มของข้อมูลและการเปลี่ยนแปลงเนื่องจากฤดูกาล ซึ่งให้ผลการพยากรณ์ที่มีความแม่นยำสูง

งานวิจัยในต่างประเทศ เช่น Pflaumer (1992) ใช้วิธีบ๊อกซ์และเจนกิ้นส์พยากรณ์ประชากรของสหรัฐอเมริกาปี 1990 – 2080 โดยใช้ข้อมูลประชากรของสหรัฐอเมริกาซึ่งเป็นอนุกรมเวลารายปีตั้งแต่ปี 1990 – 1988 แล้วนำข้อมูลประชากรรายปีมาหาผลต่างลำดับที่ 1 และ 2 เมื่อได้อนุกรมเวลาคงที่ แล้วสร้างโมเดลพยากรณ์ พบว่า ในอนาคตประชากรของสหรัฐอเมริกาจะเพิ่มขึ้นเป็น 251 – 489 ล้านคน

งานวิจัยทางการศึกษาที่ได้ประยุกต์ใช้วิธีวิเคราะห์อนุกรมเวลาของบ๊อกซ์และเจนกิ้นส์ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษา ได้แก่ บำเพ็ญ ปิดชิด (2540) โดยพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีและไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล โดยพยากรณ์ 5 ช่วงเวลาล่วงหน้า และตรวจสอบผลการพยากรณ์กับผลที่ได้จากวิธีการวิเคราะห์การถดถอย วิธีการเฉลี่ยเคลื่อนที่และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล โดยใช้การวัดความคลาดเคลื่อน 6 แบบ ได้แก่ RMSE, MAPE, GMRAE, MdAPE, MdRAE และ Percent Better เป็นเกณฑ์ โดยใช้ฐานข้อมูลอนุกรมเวลารายเดือน 3 ชุดที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ ปริมาณการยืมหนังสือทั่วไป หนังสือสำรอง และวิทยานิพนธ์ 65 ช่วงเวลา ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และชุดของข้อมูลอนุกรมเวลารายปี 2 ชุดที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ จำนวนนักเรียนระดับประถมศึกษาและระดับมัธยมศึกษา 60 ช่วงเวลา พบว่าผลการพยากรณ์ข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลด้วยวิธีการวิเคราะห์การถดถอย และผลการพยากรณ์ข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลด้วยวิธีบ๊อกซ์และเจนกิ้นส์มีขนาดความคลาดเคลื่อนน้อยที่สุด

งานวิจัยของ van Buuren (1997) ได้ทำการทดสอบความกลมกลืนของโมเดล ARMA ในอนุกรมเวลา โดยการใช้ SEM (structural equation models) พบว่า การประมาณค่าพารามิเตอร์ในโมเดล MA และ ARMA สามารถประมาณค่าโดยไม่มีความลำเอียง แต่สำหรับโมเดล AR พบว่า มีความลำเอียงในการประมาณค่าพารามิเตอร์ 5-10% หลังจากนั้นได้นำวิธีนี้ไปประยุกต์ใช้ในการแก้ปัญหาที่เกี่ยวกับการประเมินถึงการเปลี่ยนแปลงของร่างกายในระหว่างตั้งครรภ์ ศึกษาผลกระทบของการแทรกแซงทางนโยบายเพื่อป้องกันสิ่งที่ไม่ถูกต้อง และศึกษาอิทธิพลของภูมิอากาศที่เป็นสาเหตุให้เกิดการขาดงาน ต่อมาได้มีงานวิจัยของ Molenaar (1999) ได้ให้ข้อเสนอแนะเกี่ยวกับงานวิจัยของ van Buuren (1997) ในส่วนของการเกิดความลำเอียงในการ

ประมาณค่าพารามิเตอร์ในโมเดล MA ว่าเนื่องมาจากการทดลองจำลองข้อมูล เพราะในการวิจัยของ van Buuren ไม่ได้ใช้ข้อมูลจริงแต่จะใช้การจำลองข้อมูลในการวิจัย

งานวิจัยของ McGinnis (1994) ใช้โมเดลอนุกรมเวลาวิเคราะห์ด้วยโมเดล transfer function โดยวิเคราะห์ความสัมพันธ์ระหว่างองค์ประกอบทางเศรษฐศาสตร์กับนโยบายพัฒนาของรัฐบาลท้องถิ่น นโยบายพัฒนาของรัฐบาลท้องถิ่นวัดได้จากกิจกรรมในการสร้างอาคาร ข้อมูลอนุกรมเวลาทุกชุดได้มาจาก Cobb County มลรัฐ Georgia พบว่าจำนวนของการอนุญาตให้สร้างที่อยู่อาศัย จำนวนรวมการก่อสร้างทั้งหมดที่ได้รับอนุญาต และจำนวนการก่อสร้างทางพาณิชยกรรมและอุตสาหกรรมที่ได้รับอนุญาต ได้รับอิทธิพลเนื่องมาจากการตัดสินใจทางด้านเศรษฐศาสตร์ของรัฐบาล ส่วนงานวิจัยของ Chen และ Fomby (1999) ใช้โมเดล transfer function ในการวิเคราะห์โมเดลอนุกรมเวลารายเดือนเช่นเดียวกัน แต่เพิ่มการพยากรณ์โมเดล ARIMA เพื่อใช้ในการเปรียบเทียบผลการพยากรณ์ที่ได้ โมเดลอนุกรมเวลาที่ใช้มีรูปแบบการเปลี่ยนแปลงเนื่องจากฤดูกาลอย่างคงที่ การพยากรณ์ในการวิจัยนี้เป็นการพยากรณ์ระยะยาว ผลการวิจัยพบว่า ไม่มีโมเดล transfer function ใดที่มีผลการพยากรณ์เช่นเดียวกับโมเดล ARIMA อาจเนื่องมาจากการใช้ตัวแปรอื่นเข้ามาร่วมวิเคราะห์ซึ่งเป็นหลักการของโมเดล transfer function มีคุณภาพต่ำในการพยากรณ์ระยะยาว