บทที่ 2

อุตสาหกรรมสิ่งทอและเทคนิคการพยากรณ์เชิงสถิติ

2.1 อุตสาหกรรมสิ่งทอ

อุตสาหกรรมสิ่งทอเป็นอุตสาหกรรมที่มีบทบาทสำคัญต่อเศรษฐกิจของประเทศไทยเป็นอย่างมาก เพราะเป็นอุตสาหกรรมที่สำคัญในการรองรับแรงงานจำนวนมาก โดยมีการจ้างงานสูงสุดในภาคอุตสาหกรรม นอกจากนี้ประเทศไทยยังคงเป็นผู้ผลิตและผู้ส่งออกชั้นนำประเทศหนึ่งของโลก โดยมีมูลค่าการส่งออกสิ่งทอเป็นอันดับที่ 20 ของโลก

ประเภทสิ่งทอส่งออกที่สำคัญของไทย ได้แก่ เสื้อผ้าสำเร็จรูป รองลงมาคือผ้าผืน ด้ายเส้นใยประดิษฐ์ ตามลำดับ

2.1.1 การผลิตในอุตสาหกรรมสิ่งทอไทย[1]

อุตสาหกรรมสิ่งทอ เป็นอุตสาหกรรมหมวดใหญ่ของไทยที่ประกอบขึ้นด้วยอุตสาหกรรมการผลิตผลิตภัณฑ์สิ่งทอสำเร็จรูปประเภทต่าง ๆ ได้แก่ เสื้อผ้า เครื่องนุ่งห่มรวมตลอดถึงผลิตภัณฑ์ที่เป็นผลผลิตจากการถักทอเส้นใยต่าง ๆ เช่น ด้าย ผ้า เชือก และพรม เป็นต้น ลักษณะการผลิตจะมีความสัมพันธ์ต่อเนื่องกันทั้งระบบ โดยอุตสาหกรรมสิ่งทอในประเทศไทย[2] แบ่งได้เป็น 3 กลุ่มใหญ่ตามขั้นตอนการผลิต คือ

---

[1] กองวิจัยสินค้าและการตลาด. **สิ่งทอปี 2532 และแนวโน้มปี 2533** (กรุงเทพมหานคร: ฝ่ายวิจัยสินค้าอุตสาหกรรม, 2532), หน้า 7.

[2] หน่วยการอุตสาหกรรม. **สรุปภาวะธุรกิจ-อุตสาหกรรม 2535 และแนวโน้ม 2536** (กรุงเทพมหานคร: ฝ่ายวิชาการ, 2535), หน้า 52.

- กลุ่มอุตสาหกรรมสิ่งทอขั้นต้น ได้แก่ การผลิตเส้นใยธรรมชาติ เส้นใยสังเคราะห์และเส้นใยกึ่งสังเคราะห์ ลักษณะโดยทั่วไปเป็นอุตสาหกรรมที่ต้องใช้เครื่องจักร เทคโนโลยี และเงินทุนสูง
- กลุ่มอุตสาหกรรมสิ่งทอขั้นกลาง ได้แก่ การปั่นด้าย การทอผ้า ถักผ้า การฟอกย้อม พิมพ์ และแต่งสำเร็จ อุตสาหกรรมสิ่งทอขั้นนี้ลักษณะการผลิตจะใช้ทั้งทุนและแรงงานมาก แต่แนวโน้มจะมีการใช้เครื่องจักรในการผลิตแทนแรงงานคนมากขึ้น
- กลุ่มอุตสาหกรรมสิ่งทอขั้นปลาย ได้แก่ การตัดเย็บเสื้อผ้า ถุงเท้า ถุงมือผ้า เครื่องยกทรง ส่วนประกอบเครื่องแต่งกาย และเคหะสิ่งทอ

จำนวนผู้ประกอบการในภาคอุตสาหกรรมสิ่งทอ ในปี 2540 มีจำนวนไม่น้อยกว่า 4,849 โรง จำแนกเป็นโรงงานผลิตเส้นใยประดิษฐ์ 18 โรง โรงงานปั่นด้าย 158 โรง โรงงานฟอกย้อม แต่งพิมพ์สำเร็จ 426 โรง โรงงานทอผ้า ถักผ้า 1,391 โรง และโรงงานเสื้อผ้าสำเร็จรูป 2,856 โรง โรงงานที่ผลิตมีทั้งขนาดใหญ่ ขนาดกลาง และขนาดเล็ก ส่วนใหญ่ตั้งอยู่ในบริเวณรอบกรุงเทพมหานคร และจังหวัดใกล้เคียง โดยมีปริมาณการผลิตจำแนกในแต่ละขั้นตอนการผลิต ดังนี้

2.1.1.1 อุตสาหกรรมผลิตเส้นใย

อุตสาหกรรมผลิตเส้นใย แบ่งเป็น 2 ประเภทใหญ่ ๆ คือ เส้นใยธรรมชาติ ส่วนใหญ่จะเป็นเส้นใยฝ้าย และเส้นใยประดิษฐ์ หรือเส้นใยสังเคราะห์ ซึ่งสังเคราะห์จากผลิตภัณฑ์ปิโตรเลียม ถ่านหิน หรือจากเยื่อไม้ เช่น เส้นใยโพลีเอสเตอร์ ไนล่อน และเรยอง

2.1.1.2 อุตสาหกรรมปั่นด้าย

ผลิตภัณฑ์เส้นด้าย แบ่งเป็น 2 ประเภท คือ Yarn ซึ่งใช้เป็นวัตถุดิบในอุตสาหกรรมทอผ้า และเครื่องนุ่งห่ม และ Thead เป็นเส้นด้ายที่ใช้สำหรับเย็บผ้า ขนาดของอุตสาหกรรมปั่นด้ายเพื่อใช้เย็บผ้าจะมีขนาดเล็กเมื่อเทียบกับอุตสาหกรรมปั่นด้ายเพื่อใช้ในอุตสาหกรรมทอผ้าและเครื่องนุ่งห่ม

2.1.1.3 อุตสาหกรรมฟอกย้อมและแต่งพิมพ์สำเร็จ

ระบบการผลิตในอุตสาหกรรมนี้แบ่งออกเป็น 2 ประเภท คือ การฟอกย้อมเส้นด้าย(Yarn Dyeing)อันเป็นการผลิตในโรงงานปั่นด้ายที่มีหม้อฟอกย้อมของตนเอง และการฟอกย้อมผ้าผืน(Piece Dyeing)ซึ่งเป็นโรงงานขนาดใหญ่ ใช้เทคนิคการผลิตที่ทันสมัยและลงทุนสูงทำการฟอกย้อมผ้าผืนให้แก่โรงงานทอผ้าในเครือและรับจ้างฟอกย้อมทั่วไป

2.1.1.4 อุตสาหกรรมทอผ้า-ถักผ้า

ผลิตภัณฑ์ผ้าที่ประเทศไทยผลิตได้ ประกอบด้วย ผ้าทอ และผ้าถัก ซึ่งสามารถยกตามชนิดของวัตถุดิบเป็น ผ้าฝ้ายทอ ผ้าฝ้ายถัก ผ้าทอจากใยสังเคราะห์ และผ้าถักจากใยสังเคราะห์

2.1.1.5 อุตสาหกรรมเสื้อผ้าสำเร็จรูป

การผลิตเสื้อผ้าสำเร็จรูปเป็นกระบวนการผลิตขั้นสุดท้ายของอุตสาหกรรมสิ่งทอ ปัจจุบันการผลิตขยายตัวเพิ่มขึ้นอย่างรวดเร็ว เนื่องจากรัฐบาลได้ให้การสนับสนุนอุตสาหกรรมประเภทนี้อย่างมาก เฉพาะอย่างยิ่งในด้านการส่งเสริมการส่งออก จนกลายเป็นภาคการผลิตที่ทำรายได้สูงสุดของประเทศ และก่อให้เกิดการจ้างงานมากที่สุด

ผลิตภัณฑ์สิ่งทอที่ประเทศผลิตได้ในแต่ละขั้นตอนการผลิต ส่วนใหญ่จะนำไปใช้เป็นวัตถุดิบในการผลิตผลิตภัณฑ์สิ่งทอขั้นต่อเนื่องอื่น ๆ ตามลำดับ จนถึงขั้นตอนการผลิตสุดท้าย คือเสื้อผ้าสำเร็จรูป และจากความเจริญก้าวหน้าอย่างมากของอุตสาหกรรมสิ่งทอในรอบทศวรรษที่ผ่านมาส่งผลให้ความต้องการผลิตภัณฑ์สิ่งทอประเภทต่าง ๆ ในประเทศขยายตัวเพิ่มในอัตราสูงตลอดมา โดยเฉพาะในผลิตภัณฑ์ขั้นต้นและขั้นกลาง คือ เส้นใย เส้นด้ายและผ้าผืน เพื่อใช้เป็นวัตถุดิบหลักในการผลิตผลิตภัณฑ์เสื้อผ้าสำเร็จรูปเพื่อส่งออกซึ่งขยายตัวอย่างมาก

2.1.2 <u>ตลาดส่งออกสิ่งทอสำคัญของไทย</u>

ตลาดส่งออกเครื่องนุ่งห่มของไทยที่สำคัญ ได้แก่ สหรัฐอเมริกา ตามด้วยญี่ปุ่น เยอรมนี และสหราชอาณาจักร โดยในปี 2540 ส่วนแบ่งของทั้ง 4 ตลาดแรกมีสัดส่วน

ร้อยละ 62.48 ซึ่งสหรัฐอเมริกามีส่วนแบ่งสูงถึงร้อยละ 42.93 คิดเป็นมูลค่าการส่งออก 42,859.6 ล้านบาท(ดังตารางที่ 2.1)

ตารางที่ 2.1 ตลาดส่งออกที่สำคัญของเครื่องนุ่งห่มของไทยในปี 2540

ประเทศ	มูลค่า(ล้านบาท)	มูลค่า(ล้านดอลลาร์สหรัฐ)	ส่วนแบ่ง(%)
1. สหรัฐอเมริกา	42,859.6	1396.08	42.93
2. ญี่ปุ่น	10,195.1	332.09	10.21
3. เยอรมนี	4,724.2	153.88	4.73
4. สหราชอาณาจักร	4,602.1	149.91	4.61
รวม 4 ประเทศ			62.48

ที่มา : สถิติสิ่งทอไทย 2540/2541 กองอุตสาหกรรมสิ่งทอ กรมส่งเสริมอุตสาหกรรม

สำหรับการส่งออกผ้าผืนนั้น ตลาดส่งออกที่สำคัญ ได้แก่ สหรัฐอเมริกา ตามด้วยสหรัฐอาหรับเอมิเรตส์ ฮ่องกงและสหราชอาณาจักร โดยมีส่วนแบ่งร้อยละ 11.95 9.35 8.83 และ 5.76 ตามลำดับ(ดังตารางที่ 2.2)

ตารางที่ 2.2 ตลาดส่งออกที่สำคัญของผ้าผืนของไทยในปี 2540

ประเทศ	มูลค่า(ล้านบาท)	มูลค่า(ล้านดอลลาร์สหรัฐ)	ส่วนแบ่ง(%)
1. สหรัฐอเมริกา	3,554.1	115.77	11.95
2. สหรัฐอาหรับเอมิเรตส์	2,781.2	90.59	9.35
3. ฮ่องกง	2,625.6	85.52	8.83
4. สหราชอาณาจักร	1,712.9	55.79	5.76
รวม 4 ประเทศ			35.89

ที่มา : สถิติสิ่งทอไทย 2540/2541 กองอุตสาหกรรมสิ่งทอ กรมส่งเสริมอุตสาหกรรม

ส่วนตลาดส่งออกเส้นด้ายของไทยนั้น ประเทศที่นำเข้าเส้นด้ายที่สำคัญได้แก่ ฮ่องกง ตามด้วยสหรัฐอเมริกา ญี่ปุ่น และเกาหลีใต้ โดยมีส่วนแบ่งใกล้เคียงกันเท่ากับ ร้อยละ 13.31 8.58 8.35 และ 7.12 ตามลำดับ(ดังตารางที่ 2.3)

ตารางที่ 2.3 ตลาดส่งออกที่สำคัญของเส้นด้ายของไทยในปี 2540

ประเทศ	มูลค่า(ล้านบาท)	มูลค่า(ล้านดอลลาร์สหรัฐ)	ส่วนแบ่ง(%)
1. ฮ่องกง	2,392.4	77.93	13.31
2. สหรัฐอเมริกา	1,541.9	50.22	8.58
3. ญี่ปุ่น	1,501.1	48.90	8.35
4. เกาหลีใต้	1,280.6	41.71	7.12
รวม 4 ประเทศ			37.36

ที่มา : สถิติสิ่งทอไทย 2540/2541 กองอุตสาหกรรมสิ่งทอ กรมส่งเสริมอุตสาหกรรม

ถึงแม้ว่าในปัจจุบันประเทศไทยจะได้รับผลกระทบจากภาวะเศรษฐกิจที่ตกต่ำ อันส่งผลต่อเนื่องมายังอุตสาหกรรมการส่งออกสินค้าของไทยนั้น หากแต่ในอนาคตด้วยความร่วมมือของรัฐบาลในการเร่งรัดที่จะแก้ไขปัญหาทางด้านเศรษฐกิจให้ดีขึ้นและยังคงมีนโยบายให้การส่งเสริม สนับสนุนในอุตสาหกรรมการส่งออกสินค้าของไทยแล้ว ประเทศไทยก็จะยังคงรักษาความเป็นผู้นำในด้านการส่งออกอุตสาหกรรมสิ่งทอไว้ได้ต่อไป

2.2 <u>เทคนิคการพยากรณ์เชิงสถิติ</u>

เทคนิคการพยากรณ์เชิงปริมาณที่นำมาใช้ในการศึกษาวิจัยในครั้งนี้ ประกอบไปด้วยวิธีการวิเคราะห์อนุกรมเวลา(Time Series Analysis Method) และวิธีการวิเคราะห์การถดถอย(Regression Analysis Method)

การวิเคราะห์อนุกรมเวลาเพื่อการพยากรณ์ เป็นการวิเคราะห์การเคลื่อนไหวของข้อมูลอนุกรมเวลาของตัวแปร เพื่อหาตัวแบบสำหรับใช้พยากรณ์ค่าของตัวแปรนั้นในอนาคต โดยพิจารณาสหสัมพันธ์ของข้อมูล ณ คาบเวลาต่าง ๆ วิธีการวิเคราะห์อนุกรมเวลาที่ใช้ในการศึกษาวิจัยนี้ ประกอบด้วย 3 วิธีคือ วิธีพยากรณ์โดยใช้เทคนิคการปรับให้เรียบ วิธีพยากรณ์บอกซ์-เจนกินส์และวิธีแยกองค์ประกอบ

การวิเคราะห์การถดถอย เป็นวิธีการทางสถิติอย่างหนึ่งที่ใช้ศึกษาถึงรูปแบบความสัมพันธ์ระหว่างตัวแปร 2 ตัวหรือมากกว่า โดยที่ตัวแปรหนึ่งเรียกว่า"ตัวแปรตาม" (dependent variable) หรือ "ตัวแปรผล"(response variable) และตัวแปรอื่น ๆ ที่เหลือเรียกว่า "ตัวแปรอิสระ"(independent variables) หรือ "ตัวแปรให้ค่าพยากรณ์หรือ

ค่าทำนาย"(predictor variables) ซึ่งมีวัตถุประสงค์เพื่อที่จะอธิบายตัวแปรตามในรูปของฟังก์ชันของตัวแปรอิสระ โดยในการวิจัยครั้งนี้เป็นการศึกษาถึงปัจจัยต่าง ๆ ที่เป็นตัวกำหนดค่าของอุปสงค์ต่อสิ่งทอแต่ละประเภทที่ส่งออกไปยังตลาดต่างประเทศ

2.2.1 <u>เทคนิคการปรับให้เรียบ</u>

เทคนิคการปรับให้เรียบสำหรับการพยากรณ์ข้อมูลอนุกรมเวลามีหลายวิธีด้วยกัน ซึ่งในการจะการพยากรณ์เลือกใช้วิธีใดขึ้นอยู่กับลักษณะการเคลื่อนไหวของข้อมูลว่ามีการเคลื่อนไหวอย่างไร โดยในการวิจัยครั้งนี้ใช้วิธีการปรับให้เรียบครั้งเดียวแบบเอกซโพเนนเชียล วิธีการปรับให้เรียบสองครั้งแบบเอกซโพเนนเชียล วิธีพารามิเตอร์สองตัวของโฮลท์ และวิธีของวินเตอร์ แสดงรายละเอียดดังต่อไปนี้

2.2.1.1 วิธีการปรับให้เรียบครั้งเดียวแบบเอกซโพเนนเชียล(Single Exponential Smoothing)

วิธีนี้เป็นวิธีพยากรณ์สำหรับอนุกรมเวลาที่มีค่าเฉลี่ยไม่คงที่ โดยเป็นการหาค่าเฉลี่ยแบบถ่วงน้ำหนักซึ่งจะให้น้ำหนักกับข้อมูลปัจจุบันมากที่สุด และให้น้ำหนักลดลงเรื่อยๆ สำหรับข้อมูลอดีตตามลำดับ เมื่อเขียนกราฟแสดงการลดลงของน้ำหนักจะมีรูปแบบเอกซโพเนนเชียล

ตัวแบบพยากรณ์ค่าในอนาคต $Y_{t+\ell}$ (ℓ =1,2,...) จากเวลาปัจจุบัน t: :

$$\hat{Y}_t(\ell) = S_t = \alpha Y_t + (1-\alpha)S_{t-1},\ t=1,2,...$$

ซึ่ง α คือ ค่าคงที่หรือสัมประสิทธิ์ปรับให้เรียบ(smoothing constant) มีค่าอยู่ระหว่าง 0 และ 1 นักพยากรณ์จะต้องเลือกกำหนดค่า α ซึ่งโดยทั่วไปจะเลือกค่า α อยู่ระหว่าง 0.05 และ 0.3[3] ซึ่งถ้าระดับค่าเฉลี่ยเปลี่ยนแปลงช้า ๆ ควรเลือก α ค่าเล็ก สำหรับ S_t มี

[3] Bovas Abraham and Johannes Ledolter, **Statistical Method For Forecasting** (New York: John Wiley&Sons, 1983), p.86.

ชื่อเรียกว่า "ตัวสถิติปรับให้เรียบ"(smoothed statistic) ซึ่งการคำนวณหาค่า S_t จะต้องใช้ค่า S_{t-1} , S_{t-2} , ..., S_1 , S_0 นั่นคือ เมื่อทราบค่าเริ่มต้น S_0 จะสามารถหาค่า S ตัวต่อ ๆ ไปได้ ฉะนั้น นักพยากรณ์จะต้องกำหนดค่าเริ่มต้น S_0 และตัวอย่างการเลือกกำหนดค่า S_0 เช่น ใช้ค่าเฉลี่ยของข้อมูลอนุกรมเวลาที่มีอยู่ $[S_0 = \bar{Y} = (1/T)(Y_1 + Y_2 + \ldots + Y_t)]$ โดยเฉพาะกรณีที่ระดับค่าเฉลี่ยของอนุกรมเวลาเปลี่ยนแปลงช้า ๆ

อีกหนทางหนึ่งในการเลือกค่า α คือทดลองแปรเปลี่ยนค่า α เช่น เริ่มจาก α = 0.01 ต่อไปเป็น 0.02 , 0.03 ,...... และแต่ละค่า α คำนวณค่า S_t และหาค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสอง(MSE) จากนั้น เปรียบเทียบค่า MSE ทั้งหมด และเลือกค่า α ที่ให้ MSE ต่ำสุด

2.2.1.2 วิธีการปรับให้เรียบสองครั้งแบบเอกซโพเนนเชียล(Double Exponential Smoothing Method)

วิธีนี้เหมาะสำหรับข้อมูลที่มีการเคลื่อนไหวแบบแนวโน้มเชิงเส้นตรง(Linear Trend Data) และไม่มีการเคลื่อนไหวแบบฤดูกาล(Seasonal Data) ซึ่งมาจากแนวคิดการปรับให้เรียบครั้งเดียวแบบเอกซโพเนนเชียล นำมาขยายผลใช้กับข้อมูลอนุกรมเวลาที่มีแนวโน้มเชิงเส้นไม่คงที่ตลอดช่วงเวลา T มีสูตรพยากรณ์ค่าจริง $Y_{t+\ell}$ ที่เวลา $T+\ell$ จากเวลาปัจจุบัน T ดังนี้

$$\hat{Y}_T(\ell) = \left(2 + \frac{\alpha \ell}{1-\alpha}\right) S_T^{[1]} - \left(1 + \frac{\alpha \ell}{1-\alpha}\right) S_T^{[2]}$$

ซึ่ง

$$S_T^{1} = \alpha Y_T + (1-\alpha) S_{T-1}^{[1]}$$

$$S_T^{[2]} = \alpha S_T^{[1]} + (1-\alpha) S_{T-1}^{[2]}$$

การคำนวณ $S_T^{[1]}$ และ $S_T^{[2]}$ ต้องการทราบค่า $S_{T-2}^{[1]}$, $S_{T-2}^{[2]}$, $S_{T-2}^{[3]}$,.....

$S_0^{[1]}$,$S_0^{[2]}$ และดังนั้นต้องเริ่มด้วยค่า $S_0^{[1]}$ และ $S_0^{[2]}$ เราประมาณค่า $S_0^{[1]}$ และ $S_0^{[2]}$ ได้ดังนี้

$$S_0^{[1]} = \hat{\beta}_0 - \frac{(1-\alpha)}{\alpha}\hat{\beta}_1$$

$$S_0^{[1]} = \hat{\beta}_0 - \frac{1-\alpha}{\alpha}\hat{\beta}_1$$

ซึ่ง
$$\hat{\beta}_0 = \bar{Y} - \frac{T+1}{2}\hat{\beta}_1$$

2.2.1.3 วิธีพารามิเตอร์สองตัวของโฮลท์(Holt's Two-Parameter Method)

วิธีการของโฮลท์มีลักษณะคล้ายกับวิธีการปรับให้เรียบสองครั้งแบบเอกซโพเนนเชียล

$$\hat{\beta}_1 = \frac{12\sum_{t=1}^{T}(t-(T+1)/2)Y_t}{T^3 - T}$$

แต่มีลักษณะทั่วไปมากกว่า มีสูตรพยากรณ์ดังนี้

ซึ่ง
$$\hat{Y}_t(\ell) = S_T + \ell\hat{\beta}_t$$

ตัวสถิติปรับระดับ $S_T = \alpha Y_T + (1-\alpha)(S_{T-1} + \hat{\beta}_{T-1})$

ตัวสถิติปรับแนวโน้ม $\hat{\beta}_T = \gamma(S_T - S_{T-1}) + (1-\gamma)\hat{\beta}_{T-1}$

วิธีการของโฮลท์ใช้พารามิเตอร์ปรับให้เรียบสองตัวคือ α $(0<\alpha<1)$ และ γ $(0<\gamma<1)$ ซึ่งนักพยากรณ์จะต้องกำหนดค่าทั้งสองนี้ และกำหนดค่าเริ่มต้น S_1 และ β_1

2.2.1.4 วิธีการพยากรณ์ของวินเตอร์(Winter's Forecast Method)

วิธีการพยากรณ์ของวินเตอร์เป็นวิธีการพยากรณ์ที่พิจารณาองค์ประกอบฤดูกาลร่วมด้วย ดังนั้นจึงเหมาะกับอนุกรมเวลาที่มีองค์ประกอบฤดูกาล วิธีการของวินเตอร์เป็นการขยายผลของวิธีการพารามิเตอร์สองตัวของโฮลท์ โดยเพิ่มพารามิเตอร์หรือค่าคงที่ปรับให้เรียบอีกหนึ่งตัวรวมเป็นสามตัวคือ ค่าคงที่ปรับให้เรียบสำหรับระดับ (α_1) ค่าคงที่ปรับให้เรียบ

สำหรับแนวโน้มหรือความชัน (α_2) และค่าคงที่ปรับให้เรียบสำหรับฤดูกาล (α_3) ค่าทั้งสามมีค่าอยู่ระหว่าง 0 และ 1

ตัวแบบอนุกรมเวลาตัวแบบหนึ่งของวินเตอร์ ซึ่งประกอบด้วยองค์ประกอบแนวโน้มและฤดูกาลมีสมการดังนี้ และมีชื่อเรียกว่าตัวแบบผลคูณของวินเตอร์

$$Y_t = (\mu_t + \beta_t t)I_t + \varepsilon_t$$

ซึ่ง μ_t , β_t , I_t เป็นพารามิเตอร์แสดงระดับ ความชัน และฤดูกาล ของอนุกรมเวลาตามลำดับ และ ε_t คือความคลาดเคลื่อนสุ่ม โดยมีข้อสมมติพื้นฐานคือ มีค่าเฉลี่ยศูนย์ความแปรปรวนคงที่ และไม่มีสหสัมพันธ์กัน

ตัวแบบข้างต้นเหมาะสมกับอนุกรมเวลามีการแกว่งหรือการผันแปรของฤดูกาลเป็นสัดส่วนโดยตรงกับระดับของอนุกรม กล่าวคือ การแกว่งจะมากขึ้นขณะที่ระดับของอนุกรมเพิ่มขึ้น ส่วนการผันแปรของ ε_t ไม่ขึ้นอยู่กับระดับของอนุกรม

จากตัวแบบข้างต้น วิเคราะห์ได้สูตรพยากรณ์ดังนี้

$$\hat{Y}_t(\ell) = \left(\hat{\mu}_t + \ell\hat{\beta}_t\right)\hat{I}_{t+l-m}, t = m, m+1, \ldots$$

ซึ่ง

$$\hat{\mu}_t = \alpha_1\left(Y_t / \hat{I}_{t-m}\right) + \left(1-\alpha_1\right)\left(\hat{\mu}_{t-1} + \hat{\beta}_{t-1}\right)$$

$$\hat{\beta}_t = \alpha_2\left(\hat{\mu}_t - \hat{\mu}_{t-1}\right) + \left(1-\alpha_2\right)\hat{\beta}_{t-1}$$

$$\hat{I}_t = \alpha_3\left(Y_t/\hat{\mu}_t\right) + \left(1-\alpha_3\right)\hat{I}_{t-m}$$

m = ความยาวของคาบฤดูกาล เช่น m = 12 (สำหรับอนุกรมเวลารายเดือน) หรือ m = 4 (สำหรับอนุกรมเวลารายสามเดือน)

การคำนวณค่าพยากรณ์ $\hat{Y}_t(\ell)$ ต้องกำหนดค่าเริ่มต้นของ $\hat{\mu}_t$, $\hat{\beta}_t$ และ $\hat{I}_t$ นอกเหนือจากการกำหนดค่าคงที่ปรับให้เรียบ α_1 , α_2 , α_3 และหนทางหนึ่งในการกำหนดค่าเริ่มต้นคือให้

$$\hat{\mu}_m = (Y_1 + Y_2 + \ldots + Y_m)/m$$

$$\hat{I}_t = Y_t / \hat{\mu}_m \quad \text{สำหรับ } t = 1,2,\ldots,m$$

$$\hat{\beta}_m = 0$$

2.2.2 วิธีการตัวแบบบอกซ์-เจนกินส์ หรือตัวแบบ ARIMA

วิธีบอกซ์-เจนกินส์เป็นวิธีการสร้างตัวแบบพยากรณ์สำหรับอนุกรมเวลาที่นำเอาสหสัมพันธ์ของอนุกรมเวลา $\{\ldots Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, \ldots\}$ ที่ปรากฏไปวิเคราะห์ใช้ประโยชน์ โดยพิจารณาสหสัมพันธ์ระหว่าง Y ที่ตำแหน่งเวลาหรือคาบเวลา t (Y_t) และ Y ที่ตำแหน่งเวลา หรือคาบเวลาต่าง ๆ ที่ผ่านมา($Y_{t-1}, Y_{t-2},\ldots$) โดยเมื่อได้ตัวแบบแล้ว ตัวแบบนี้จะแสดงความสัมพันธ์ระหว่าง Y_t กับ $Y_{t-1}, Y_{t-2},\ldots$ และจะใช้ตัวแบบนี้พยากรณ์ค่า $Y_{t+1}, Y_{t+2},\ldots$ในอนาคต ซึ่งต่างจากวิธีการพยากรณ์อนุกรมเวลาโดยทั่วไป เช่นวิธีการปรับให้เรียบที่กล่าวไปแล้ว รวมทั้งการวิเคราะห์การถดถอยที่ใช้วิธี OLS จะมีข้อสมมติพื้นฐานข้อหนึ่งคือ อนุกรมเวลา $\{\ldots Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, \ldots\}$ ไม่มีสหสัมพันธ์ ซึ่งบ่อยครั้งพบว่าข้อสมมตินี้ไม่เป็นจริง นั่นคือ มีหลายกรณีที่อนุกรมเวลา $\{\ldots Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, \ldots\}$ มีสหสัมพันธ์ ถ้าเป็นเช่นนี้ การใช้วิธีที่มีข้อสมมติตัวแปรอนุกรมเวลาไม่มีสหสัมพันธ์อาจจะไม่เหมาะสม ทั้งนี้เพราะวิธีการเหล่านี้ไม่ได้นำสหสัมพันธ์ที่ปรากฏไปใช้ประโยชน์ในการสร้างตัวแบบพยากรณ์ ฉะนั้นจึงได้มีผู้คิดค้นหาวิธีการสร้างตัวแบบพยากรณ์สำหรับอนุกรมเวลาที่ได้นำเอาสหสัมพันธ์ที่ปรากฏไปวิเคราะห์ใช้ประโยชน์ และโดยทั่วไปวิธีเหล่านี้จะให้ผลพยากรณ์ที่ดีกว่า ในหลายวิธีเหล่านี้ วิธีหนึ่งที่เป็นที่รู้จักและใช้กันมากคือวิธีบอกซ์-เจนกินส์ ซึ่งพัฒนาขึ้นโดย George E.P,Box และ Gwilym M. Jenkins

ตัวแบบบอกซ์-เจนกินส์โดยทั่วไป จะใช้พยากรณ์ค่าในช่วงเวลาข้างหน้าที่เป็นระยะสั้น หรือปานกลาง ทั้งนี้เพราะตัวแบบโดยทั่วไปจะให้ความสำคัญหรือน้ำหนักมากกับข้อมูล

ในอดีตระยะใกล้มากกว่าข้อมูลอดีตที่ห่างไกลออกไปมาก ๆ เราอาจจะไม่เคยพบตัวแบบบอกซ์-เจนกินส์ ที่แสดงความสัมพันธ์ระหว่าง Y_t กับ Y ที่ห่างไกลกันมาก ๆ

แนวคิดของการพัฒนาตัวแบบบอกซ์-เจนกินส์ มาจากการศึกษาวิเคราะห์กระบวนการเชิงเส้น หรือตัวกรองเชิงเส้น (linear filter)

$$Y_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots \qquad (2.2.1)$$

นั่นคือ พิจารณาอนุกรมเวลาหรือค่าสังเกต Y_t เกิดจากผลบวกเชิงเส้นของตัวแปรสุ่ม a_t, a_{t-1},...ที่ไม่มีสหสัมพันธ์กัน เราเรียกตัวแปรสุ่ม a_t, a_{t-1},... ว่าค่าผิดพลาดสุ่ม หรือเรียกว่ากระตุกสุ่ม(random shocks) และสมมติว่าแต่ละตัวมีค่าเฉลี่ยเท่ากับศูนย์และมีความแปรปรวนคงที่ และโดยทั่วไปสมมติด้วยว่ามีการแจกแจงปกติ

ในตัวกรองเชิงเส้น หรือตัวแบบเชิงเส้น (2.2.1) พารามิเตอร์ μ คือ ค่าระดับค่าเฉลี่ยของ Y_t เมื่ออนุกรมเวลาอยู่ในสภาพคงที่ และพารามิเตอร์ ψ_1 , ψ_2 ,.... เป็นน้ำหนักที่ให้กับตัวแปรสุ่ม a_{t-1} , a_{t-2},...

กระบวนการหรือตัวแบบเชิงเส้น (2.2.1) จะไม่ให้ประโยชน์ถ้ามีพารามิเตอร์จำนวนอนันต์(จำนวนไม่รู้จบ) เพราะฉะนั้น จะสร้างตัวแบบที่ประกอบด้วยพารามิเตอร์จำนวนจำกัด และเพียงพอที่จะอธิบายอนุกรมเวลาที่พิจารณา

2.2.2.1 ตัวแบบภายใต้ภาวะคงที่ (Stationary Models)

2.2.2.1.1ตัวแบบอัตถดถอย [Autoregressive (AR) Models]

จากตัวแบบเชิงเส้น (2.2.1) พัฒนาตัวแบบเฉพาะกลุ่มหนึ่งเรียกว่า "ตัวแบบอัตถดถอยอันดับ p" หรือ "กระบวนการอัตถดถอยอันดับ p" [autoregressive process of order p] แทนด้วยอักษรย่อ AR(p) และมีรูปแบบทั่วไปดังนี้

$$Y_t - \mu = \phi_1 (Y_{t-1} - \mu) + \phi_2 (Y_{t-2} - \mu) + \ldots + \phi_p (Y_{t-p} - \mu) + a_t$$

หรือ

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + a_t \qquad (2.2.2)$$

หรือ

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + ... + \phi_p Z_{t-p} + a_t$$

โดยให้ $Z_t = Y_t - \mu$, $Z_{t-1} = Y_{t-1} - \mu$,... และ $c = \mu(1 - \phi_1 - \phi_2 - \phi_p)$

และ ϕ_1 , ϕ_2 ......, μ เป็นพารามิเตอร์ที่โดยทั่วไปไม่ทราบค่า จะประมาณค่าจากข้อมูล

จากตัวแบบ AR(p) ข้างต้น อาจเขียนในรูปแบบสั้น ๆ :

$$\phi_p(B)Z_t = a_t \tag{2.2.3}$$

โดยที่ $\phi_p(B) = \left(1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p\right)$

และ $BZ_t = Z_{t-1}$, $B^2 Z_t = Z_{t-2}$,...., $B^p Z_t = Z_{t-p}$

เราเรียก B ว่า "ตัวถอยหลังเวลา" (backward-shift operator)

โดยทั่วไปในทางปฏิบัติ อันดับ p จะไม่สูงมาก เช่น 1, 2 , หรือ 3 ถ้า p = 1, 2 เขียนกระบวนการคงที่หรือตัวแบบคงที่ AR(1) และ AR(2) ได้ดังนี้

ตัวแบบ AR(1) :

$$Y_t = c + \phi_1 Y_{t-1} + a_t \tag{2.2.4}$$

โดยมีเงื่อนไขของการเป็นตัวแบบคงที่ $-1 < \phi_1 < 1$

ตัวแบบ AR(2) :

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t \tag{2.2.5}$$

โดยมีเงื่อนไขของการเป็นตัวแบบคงที่ดังนี้

$$\phi_1 + \phi_2 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$-1 < \phi_2 < 1$$

2.2.2.1.2 ตัวแบบค่าเฉลี่ยเคลื่อนที่ [Moving Average (MA) Models]

จากตัวแบบเชิงเส้น (2.2.1) พัฒนาตัวแบบเฉพาะอีกกลุ่มหนึ่งเรียกว่า "ตัวแบบค่าเฉลี่ยเคลื่อนที่อันดับ q" หรือ "กระบวนการค่าเฉลี่ยเคลื่อนที่อันดับ q " [moving average process of order q] แทนด้วยอักษรย่อ MA(q) และมีตัวแบบทั่วไปดังนี้

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \qquad (2.2.6)$$

หรือ

$$Z_t = \theta_q(B) a_t \qquad (2.2.7)$$

ซึ่ง $Z_t = Y_t - \mu$

$$\theta_q(B) = \left(1 - \theta_1 B - \theta_2 B - \ldots - \theta_q B^q\right)$$

ตัวแบบบอกซ์-เจนกินส์ ยังมีเงื่อนไขที่ต้องสอดคล้องอีกเงื่อนไขหนึ่งนอกเหนือจากเงื่อนไขคงที่ (stationarity) คือเงื่อนไข "ผกผันได้" (invertibility) ซึ่งพบว่าตัวแบบ AR(p) , $p<\infty$, มีคุณสมบัติผกผันได้เสมอ แต่อาจไม่คงที่ ขณะที่ตัวแบบ MA(q) , $q<<\infty$, มีคุณสมบัติคงที่เสมอ แต่อาจจะผกผันไม่ได้ เพราะฉะนั้น ต้องตรวจสอบคุณสมบัติคงที่ในตัวแบบ AR และตัวสอบคุณสมบัติผกผันได้ในตัวแบบ MA

ทำนองเดียวกันกับตัวแบบ AR ในทางปฏิบัติจะมีอันดับ q ไม่สูงมาก เช่น 1, 2, หรือ 3 ถ้า $q = 1$ และ 2 จะได้ตัวแบบ MA(1) และ MA(2) ดังนี้

ตัวแบบ MA(1) :

$$Y_t = \mu + a_t - \theta_1 a_{t-1} \qquad (2.2.8)$$

โดยมีเงื่อนไขของการเป็นตัวแบบคงที่ $-1 < \theta_1 < 1$

ตัวแบบ MA(2)

$$Y_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \qquad (2.2.9)$$

โดยมีเงื่อนไขของการเป็นตัวแบบคงที่ดังนี้

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$-1 < \theta_2 < 1$$

2.2.2.1.3 ตัวแบบอัตถดถอย-ค่าเฉลี่ยเคลื่อนที่

[Autoregressive-Moving Average (ARMA) Models]

ในบางกรณี การใช้ตัวแบบหรือกระบวนการผสมระหว่างตัวแบบ AR และ MA จะเป็นตัวแบบที่ประหยัด แทนการใช้ตัวแบบ AR อันดับสูง ๆ ตัวแบบเดียว หรือแทนการใช้ตัวแบบ MA อันดับสูง ๆ ตัวแบบเดียว เราจะใช้สัญลักษณ์ ARMA(p,q) หมายถึงตัวแบบผสมอัตถดถอยอันดับ p และค่าเฉลี่ยเคลื่อนที่อันดับ q และมีรูปแบบดังนี้

$$\begin{aligned} Y_t = {} & \mu + \phi_1 (Y_{t-1} - \mu) + \phi_2 (Y_{t-2} - \mu) + \ldots + \\ & \phi_p (Y_{t-p} - \mu) + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \\ & \theta_q a_{t-q} \end{aligned} \qquad (2.2.10)$$

หรือ $\phi_p(B) Z_t = \theta_q(B) a_t$ (2.2.11)

อันดับ p และ q สำหรับตัวแบบ ARMA จะไม่สูงมากในทางปฏิบัติ ถ้า $p = 1$ และ $q = 1$ จะได้ตัวแบบ ARMA(1,1) :

$$Y_t = \mu + \phi_1 (Y_{t-1} - \mu) + a_t - \theta_1 a_{t-1}$$

หรือ

$$Y_t = c + \phi_1 Y_{t-1} + a_t - \theta_1 a_{t-1} \qquad (2.2.12)$$

โดยมีเงื่อนไขของความคงที่และผกผันได้ ดังนี้

$$-1< \phi_1 <1 \quad \text{และ} \quad -1< \theta_1 <1$$

2.2.2.2 ตัวแบบภายใต้ภาวะไม่คงที่(Nonstationary Models) และตัวแบบ ARIMA

ถ้าข้อมูลอนุกรมเวลาหรือกระบวนการ Y_t ไม่อยู่ในสภาพคงที่ในค่าเฉลี่ย และ/หรือ ความแปรปรวน นักพยากรณ์จะต้องแปลงข้อมูลนั้นให้อยู่ในสภาพคงที่ก่อนพิจารณากำหนดตัวแบบ

กรณีที่อนุกรมเวลาไม่คงที่ในค่าเฉลี่ย จะต้องทำการแปลงข้อมูลซึ่งใช้วิธีการทำผลต่าง โดยนำข้อมูลมาลบกันได้เป็นข้อมูลชุดใหม่ นอกจากนี้การทำผลต่างดังกล่าว อาจจะต้องทำมากกว่าหนึ่งครั้งจึงจะทำให้อนุกรมมีค่าเฉลี่ยคงที่ ซึ่งโดยทั่วไป ถ้าอนุกรมมีแนวโน้ม มักจะทำผลต่างสองครั้งจึงจะคงที่ การทำผลต่างไม่ควรทำหลายครั้งมากเกินความจำเป็น เพราะจะมีผลทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูง

ในกรณีที่อนุกรมเวลาไม่คงที่ในความแปรปรวน หรือมีการเคลื่อนไหวเป็นเส้นโค้ง วิธีการแปลงข้อมูลที่ใช้กันมากคือ ใส่ ln ในอนุกรม Y_t ได้เป็นข้อมูลอนุกรมใหม่ $X_t = \ln Y_t$ ซึ่ง $Y_t > 0$, $t = 1, 2, \ldots, n$ วิธีนี้มักจะใช้เมื่อความแปรปรวนแปรผันตามระดับค่าเฉลี่ย เช่น ความแปรปรวนมากขึ้นขณะที่อนุกรมมีแนวโน้มเพิ่มขึ้น หรือมีระดับค่าเฉลี่ยสูงขึ้น บางกรณีการใช้ ln อาจจะไม่ได้ผลนักพยากรณ์ควรทดลองใช้วิธีอื่น เช่น ใช้วิธีหารากที่สองได้ $X_t = \sqrt{Y_t}$, $Y_t > 0$, $t = 1, 2, \ldots, n$

ถ้าอนุกรมเวลาไม่คงที่ทั้งในค่าเฉลี่ยและความแปรปรวน ในกรณีนี้จะต้องแปลงข้อมูลให้คงที่ทั้งในค่าเฉลี่ย และความแปรปรวน และโดยทั่วไปจะแปลงข้อมูลในเรื่องความแปรปรวนก่อน แล้วจึงทำผลต่างในข้อมูลที่ถูกแปลงแล้ว ทั้งนี้เพราะ ถ้าทำผลต่างก่อนอาจแปลงแก้ความแปรปรวนไม่ได้ เช่นถ้าผลต่างมีค่าเป็นลบ และถ้าใช้วิธีหาค่า ln หรือรากที่สอง จะหาค่า ln หรือรากที่สองไม่ได้

เมื่ออนุกรมเวลามีสภาพไม่คงที่ หรือไม่เคลื่อนไหวรอบค่าเฉลี่ยคงที่ค่าหนึ่งค่าเดียว จะต้องแปลงข้อมูลดังที่กล่าวไปแล้ว ฉะนั้น ถ้ามีการทำผลต่าง d ครั้ง จะเขียนตัวแบบผสมเป็น ARIMA (Autoregressive Integrated Moving Average Models) ด้วยอันดับ (p,d,q) มีรูปแบบทั่วไปดังนี้

ตัวแบบ ARIMA(p,d,q) :

$$\phi_p(B)(1-B)^d Y_t = \delta + \phi_q(B)a_t \qquad (2.2.13)$$

หรือ

$$\phi_p(B)\, W_t = \delta + \phi_q(B)a_t \qquad (2.2.14)$$

ซึ่งให้ $W_t = (1-B)^d Y_t$ และ δ (อาจมีค่าเท่ากับศูนย์) เป็นพารามิเตอร์แสดงระดับค่าเฉลี่ยคงที่ของอนุกรม W_t และ Y_t เป็นอนุกรมที่ถูกแปลงให้มีความแปรปรวนคงที่แล้ว ถ้าอนุกรมแรกเริ่มไม่คงที่ในความแปรปรวน

2.2.2.3 ตัวแบบ ARIMA เมื่อมีองค์ประกอบฤดูกาล

องค์ประกอบที่เป็นฤดูกาลในที่นี้ จะไม่จำกัดเฉพาะการแปรผันของข้อมูลคล้ายคลึงกันในระยะเวลาห่างกัน 1 ปี (คล้ายคลึงกันจากปีหนึ่งไปยังอีกปีหนึ่ง) แต่จะรวมถึงความแปรผันคล้ายคลึงกันในช่วงเวลาห่างอื่น ๆ ด้วย เช่น คล้ายคลึงกันทุก ๆ 7 วัน ในกรณีนี้จะมีคาบฤดูกาล $s = 7$

สำหรับอนุกรมเวลาที่มีองค์ประกอบฤดูกาลโดยมีคาบเวลาของฤดูกาล $s > 1$ ตัวแบบอนุกรมเวลาในส่วนที่เป็นฤดูกาลจะมีโครงสร้างเป็นไปได้เหมือนกับองค์ประกอบที่ไม่ใช่ฤดูกาล นั่นคือจะมีตัวแบบ ARIMA ด้วยอันดับ $(P,D,Q)_S$ ซึ่ง P คืออันดับในส่วนของกระบวนการ AR, Q คืออันดับในส่วนของกระบวนการ MA , และ D คือจำนวนครั้งทำผลต่างอนุกรมเวลาห่างกัน s คาบเวลา

เมื่อนำองค์ประกอบในส่วนที่ไม่ใช่ฤดูกาล และส่วนที่เป็นฤดูกาลมาผนวกเข้าด้วยกัน จะได้ตัวแบบ ARIMA ที่แสดงส่วนประกอบทั้งสอง และตัวแบบทั่วไปตัวแบบหนึ่งคือตัวแบบในรูปผลคูณ ARIMA (p,d,q) $(P,D,Q)_S$ มีรูปแบบดังนี้

ตัวแบบ ARIMA(p,d,q) $(P,D,Q)_S$:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Y_t = \delta + \theta_q(B)\Theta_Q(B^s)a_t \qquad (2.2.15)$$

โดยที่ $\Phi_p\left(B^s\right) = \left(1-\Phi_s B^s - \Phi_{2s} B^{2s} - \ldots - \Phi_{Ps} B^{Ps}\right)$

$$\Theta_Q\left(B^s\right) = \left(1-\Theta_s B^s - \Theta_{2s} B^{2s} - \ldots - \Theta_{Qs} B^{Qs}\right)$$

2.2.2.4 ขั้นตอนวิธีการสร้างตัวแบบบอกซ์-เจนกินส์ หรือ ตัวแบบ ARIMA

กรรมวิธีสร้างตัวแบบบอกซ์-เจนกินส์ หรือ ตัวแบบ ARIMA จะมีขั้นตอนใหญ่ ๆ ทำนองเดียวกันกับกรรมวิธีสร้างตัวแบบการถดถอย กล่าวคือ ประกอบด้วยขั้นตอน

1) กำหนดตัวแบบทดลอง(Identification)
   คือ กำหนดตัวแบบที่พิจารณาว่าจะเป็นตัวแบบที่เหมาะสมหรือมีความเพียงพอในเชิงสถิติ
2) จากตัวแบบที่เลือก ประมาณค่าองค์ประกอบหรือค่าพารามิเตอร์ของตัวแบบ (Estimation)
3) วินิจฉัยตัวแบบ (Diagnostic checking) ตัวแบบที่เลือกและประมาณค่าพารามิเตอร์แล้ว อาจยังมีความไม่เหมาะสมเพียงพอในเชิงสถิติ จึงควรตรวจสอบวินิจฉัยตัวแบบ ถ้าพบว่าตัวแบบที่กำหนดยังไม่สอดคล้องข้อสมมติ หรือขาดคุณสมบัติในเชิงสถิติ หรือยังมีรูปแบบไม่เหมาะสม จะทำการปรับแก้ตัวแบบใหม่และประมาณค่าพารามิเตอร์ของตัวแบบใหม่ และตรวจสอบความเพียงพอและรูปแบบของตัวแบบใหม่อีก จนกว่าจะพบว่าตัวแบบพยากรณ์ผ่านการทดสอบ มีความเหมาะสมเพียงพอเชิงสถิติ เมื่อผ่านขั้นนี้ก็เข้าสู่ขั้นพยากรณ์

2.2.2.4.1การกำหนดตัวแบบ ARIMA(p,d,q) $(P,D,Q)_s$

การกำหนดตัวแบบ ARIMA นักพยากรณ์จะต้องพิจารณากำหนดอันดับ p,d, และ q และต้องกำหนดอันดับ P,D,Q, และ s ด้วย ถ้าตรวจพบว่าข้อมูลอนุกรมเวลามีองค์ประกอบฤดูกาลด้วยคาบฤดูกาล s

อันดับ p และ q คืออันดับของกระบวนการ AR และ MA ในส่วนที่ไม่ใช่ฤดูกาล และ P และ Q คืออันดับของกระบวนการ AR และ MA ในส่วนที่เป็นฤดูกาล สำหรับ d

คือจำนวนครั้งทำผลต่างอนุกรมเวลาเมื่ออนุกรมเวลาในส่วนที่ไม่ใช่ฤดูกาลมีสภาพไม่คงที่ในค่าเฉลี่ย

การพิจารณากำหนดอันดับ (p,d,q) และ $(P,D,Q)_S$ จะพิจารณาแยกจากกันแต่ใช้หลักการพิจารณาเหมือนกัน

กระบวนการ AR และ MA ต่างมีรูปแบบโครงสร้างเฉพาะสำหรับอันดับ p และ q ของฟังก์ชันอัตสหสัมพันธ์ ACF (Autocorrelation Function) แทนด้วย ρ_k และโครงสร้างของฟังก์ชันอัตสหสัมพันธ์ย่อย PACF (Patial Autocorrelation Function) แทนด้วย ϕ_{kk} ซึ่ง k หมายถึงคาบเวลาห่างระหว่างอนุกรม และเรียกคาบเวลานี้ว่า "แล็ก k" (lag k) ฉะนั้น ρ_1 หมายถึงอัตสหสัมพันธ์ที่แล็ก 1 หรือ อัตสหสัมพันธ์ของอนุกรมเวลาที่ห่างกัน 1 หน่วย หรือ 1 คาบเวลา (Y_t, Y_{t+1}), t = 1,2,...., ซึ่งวัดความสัมพันธ์เชิงเส้นระหว่างอนุกรมเวลาที่ห่างกัน 1 คาบเวลาและ ρ_2 คืออัตสหสัมพันธ์ของอนุกรมเวลาที่ห่างกัน 2 คาบเวลา (Y_t, Y_{t+2}), t = 1,2,...., สำหรับ ϕ_{kk} เป็นอัตสหสัมพันธ์ที่วัดความสัมพันธ์ระหว่างอนุกรมเวลาที่ห่างกัน k คาบเวลา (Y_t, Y_{t+k}) โดยพิจารณาจากผลกระทบจากอนุกรมเวลา $Y_{t+1}, Y_{t+2}, \ldots, Y_{t+k-1}$ เข้ามาด้วย ค่าของ ρ_k และ ϕ_{kk} ต่างมีค่าอยู่ระหว่าง -1 และ 1 ตัวอย่างสูตรของฟังก์ชันเหล่านี้เช่น

กระบวนการ AR(1) : $Y_t = \delta + \phi Y_{t-1} + a_t$

$$\rho_k = \phi^k \quad \text{สำหรับ } k = 0,1,2,\ldots\ldots$$

$$\phi_{11} = \rho_1 = \phi \,,\ \phi_{kk} = 0 \quad \text{สำหรับ } k = 2,3,\ldots.$$

กระบวนการ MA(1) : $Y_t = \mu + a_t - \theta a_{t-1}$

$$\rho_1 = \frac{-\theta}{1+\theta^2} \,,\ \rho_k = 0 \quad \text{สำหรับ } k = 2,3,\ldots\ldots$$

$$\phi_{kk} = \frac{-\theta^k\left(1-\theta^2\right)}{1-\theta^{2(k+1)}} \quad \text{สำหรับ } k = 1,2,\ldots.$$

เพราะฉะนั้นการกำหนดอันดับ จะประมาณค่า ρ_k และ ϕ_{kk} โดยใช้ข้อมูลอนุกรมเวลาที่นำมาวิเคราะห์ แทนค่าประมาณด้วย $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ และเรียกว่า "ฟังก์ชันอัตสหสัมพันธ์ตัวอย่าง" SACF(Sample Autocorrelation Function) และ "ฟังก์ชันอัตสหสัมพันธ์ย่อยตัวอย่าง" SPACF(Sample Partial Autocorrelation Function)

ค่าประมาณ $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ซึ่งจะคำนวณค่าโดยใช้ข้อมูลอนุกรมเวลา มีสูตรทั่วไปดังนี้

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2} \quad \text{สำหรับ } k = 0, 1, 2, \ldots$$

$$\hat{\phi}_{11} = \hat{\rho}_1$$

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_{kk} - \sum_{j=1}^{k-1}\hat{\phi}_{k-1,j}\left(\hat{\rho}_{k-j}\right)}{1 - \sum_{j=1}^{k-1}\hat{\phi}_{k-1,j}\left(\hat{\rho}_j\right)} \quad \text{สำหรับ } k = 2, 3, \ldots$$

ซึ่ง $\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j}$ $(k = 3, 4, \ldots; j = 1, 2, \ldots, k-1)$

y_t = ข้อมูลที่เวลา t

การคำนวณค่า $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ที่แล็ก k = 1,2,....,m เพื่อให้มีจำนวนค่ามากพอที่จะพิจารณาโครงสร้างการแปรผันของอัตสหสัมพันธ์ได้ง่าย จะคำนวณค่า $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ แต่ละประเภท จำนวน m = n/4 โดยประมาณ จากค่าประมาณเหล่านี้ จะเปรียบเทียบลักษณะแปรผันกับโครงสร้างแปรผันของ $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ทางทฤษฎีและเลือกตัวแบบ ARIMA ที่โครงสร้างของ $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ เข้ากับโครงสร้าง $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ของ ARIMA นั้นมากที่สุดด้วยวิธีการเลือกตัวแบบ ARIMA โดยการเปรียบเทียบโครงสร้างแปรผันของค่าตัวอย่าง $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ กับโครงสร้างแปรผันของค่าทางทฤษฎี $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ของ

กระบวนการ AR , MA, และ ARMA ที่อันดับต่าง ๆ เพื่อจะได้เลือกกระบวนการ และอันดับเป็นตัวแบบ ARIMA ทดลอง สำหรับข้อมูลอนุกรมเวลาที่พิจารณา

สรุปลักษณะแปรผันของ ACF และ PACF ของกระบวนการอนุกรมเวลาคงที่สำหรับกระบวนการพื้นฐานดังนี้

กระบวนการ	ACF	PACF
AR(1)	ค่า ρ_k ลดลงอย่างรวดเร็ว ขณะที่ $k > 1$	ค่า ϕ_{kk} จะมีค่าสูงที่ $k = 1$ และเท่ากับ 0 เมื่อ $k > 1$
AR(2)	ค่า ρ_k ลดลงอย่างรวดเร็ว ขณะที่ $k > 2$	ค่า ϕ_{kk} จะมีค่าสูงที่ $k = 1,2$ และเท่ากับ 0 เมื่อ $k > 2$
MA(1)	ค่า ρ_k จะมีค่าสูงที่ $k = 1$ และเท่ากับ 0 เมื่อ $k >1$	ค่า ϕ_{kk} ลดลงอย่างรวดเร็ว ขณะที่ $k >1$
MA(2)	ค่า ρ_k จะมีค่าสูงที่ $k = 1,2$ และเท่ากับ 0 เมื่อ $k >2$	ค่า ϕ_{kk} ลดลงอย่างรวดเร็ว ขณะที่ $k > 2$
ARMA(1,1)	ค่า ρ_k ลดลงอย่างรวดเร็ว หลังจากแล็ก $k = 1$	ค่า ϕ_{kk} ลดลงอย่างรวดเร็ว หลังจากแล็ก $k = 1$

สำหรับกระบวนการ AR , MA, และ ARMA ที่อันดับอื่น ๆ พิจารณาได้ในทำนองเดียวกัน

ในองค์ประกอบที่เป็นฤดูกาล s การกำหนดอันดับ P และ Q พิจารณาทำนองเดียวกันกับองค์ประกอบที่ไม่เป็นฤดูกาล โดยพิจารณาโครงสร้างแปรผันของอัตสหสัมพันธ์ $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ที่แล็กฤดูกาล s, 2s, 3s,... เปรียบเทียบกับโครงสร้างของ $\hat{\rho}_k$ และ $\hat{\phi}_{kk}$ ทางทฤษฎี ซึ่งมีลักษณะตามที่กล่าวไปแล้วข้างต้น

2.2.2.4.2 การประมาณค่าพารามิเตอร์ในตัวแบบ ARIMA

เมื่อนักพยากรณ์เลือกตัวแบบ ARIMA ทดลองได้แล้ว ขั้นต่อไปคือประมาณค่าของ

พารามิเตอร์ที่ปรากฏในตัวแบบ วิธีการประมาณค่าพารามิเตอร์มีด้วยกันหลายวิธี วิธีหนึ่งที่ใช้กันมากคือ วิธีกำลังสองน้อยที่สุดแบบไม่เป็นเชิงเส้น (nonlinear least-squares method)

เมื่อใช้โปรแกรมคอมพิวเตอร์ในเรื่อง ARIMA ในขั้นประมาณค่าพารามิเตอร์ นอกจากจะได้ค่าประมาณของพารามิเตอร์แล้ว จะมีค่าของตัวสถิติต่าง ๆ ปรากฏออกมาด้วยซึ่งให้ประโยชน์ในการทดสอบเชิงสถิติว่า องค์ประกอบหรือพารามิเตอร์นั้นควรมีอยู่ในตัวแบบหรือไม่ ซึ่งนั่นคือเป็นหนทางหนึ่งในการพิจารณาว่า ตัวแบบที่พิจารณานั้นเหมาะสมเพียงพอหรือไม่ในเชิงสถิติ

2.2.2.4.3การวินิจฉัยตัวแบบ ARIMA

ภายหลังที่ประมาณค่าพารามิเตอร์ของตัวแบบ ARIMA แล้ว นักพยากรณ์ควรจะตรวจสอบตัวแบบก่อนที่จะตัดสินใจนำตัวแบบนั้นไปใช้พยากรณ์ เนื่องจากตัวแบบที่พิจารณาคัดเลือกในขั้นแรกนั้นอาจยังเลือกไม่ถูกต้องเหมาะสม จึงควรวินิจฉัย และถ้าพบว่ายังไม่เหมาะสมควรกลับไปขั้นที่ 1 พิจารณาปรับปรุงแก้ไขตัวแบบใหม่ และดำเนินการขั้นที่ 2 ประมาณค่าและวินิจฉัยในขั้นที่ 3 กรรมวิธีจะกระทำซ้ำ ๆ เช่นนี้จนกว่าจะได้ตัวแบบที่เหมาะสมเพียงพอในเชิงสถิติ

การวินิจฉัยตัวแบบ จะทำการตรวจสอบคุณสมบัติเชิงสถิติของค่าผิดพลาดสุ่ม $\{a_t\}$ และการทดสอบว่า ค่าผิดพลาดสุ่มมีอัตสหสัมพันธ์หรือไม่ จะเป็นการตรวจสอบที่สำคัญมากที่สุดในการวินิจฉัยความเพียงพอในเชิงสถิติของตัวแบบ ARIMA ฉะนั้นการทดสอบ จะคำนวณค่า SACF และ SPACF ของค่าเศษเหลือตกค้าง $e_t = y_t - \hat{y}_t$ ซึ่งเป็นค่าประมาณของ a_t ที่แล็ก k ต่าง ๆ และทดสอบด้วยค่าของตัวสถิติที , t , สำหรับทดสอบว่าค่าผิดพลาดสุ่มมีอัตสหสัมพันธ์หรือไม่ที่แต่ละแล็ก $k = 1,2,3,....,m$ และทดสอบอัตสหสัมพันธ์รวมหรือพร้อมกัน k แล็ก ด้วยตัวสถิติไคกำลังสอง(chi-squared test) ว่าค่าผิดพลาดไม่มีอัตสหสัมพันธ์ k แล็กแรก นอกจากการวินิจฉัยตัวแบบด้วยการทดสอบเชิงสถิติแล้ว นักพยากรณ์อาจตรวจสอบด้วยวิธีการอื่น ๆ ด้วยเช่น การเขียนกราฟของเศษเหลือตกค้างกับแกนเวลา ถ้าพบว่าค่าของเศษเหลือตกค้างกระจายเป็นแนวในลักษณะขนานรอบค่าเฉลี่ยศูนย์ แสดงเหตุผลได้ว่า ค่าผิดพลาดมีค่าเฉลี่ยศูนย์ และมีความแปรปรวนคงที่ แต่ถ้าการกระจายของค่าเศษเหลือตกค้างมีรูปแบบต่างไปจากแนวขนาน ควรพิจารณาปรับปรุงแก้ไขตัวแบบ ซึ่งอาจจะพบว่าความแปรปรวนยังไม่คงที่(ถ้าความแปรปรวน

แปรเปลี่ยนตามเวลา) ต้องปรับให้คงที่ด้วยวิธีการแปลงข้อมูล เป็นต้น

ผลจากการวินิจฉัยตัวแบบนอกจากจะช่วยตรวจสอบว่าตัวแบบที่กำลังพิจารณาเหมาะสมเพียงพอในเชิงสถิติหรือไม่แล้ว ยังให้แนวทางในการปรับปรุงแก้ไขตัวแบบด้วย ถ้าพบว่าตัวแบบยังไม่เหมาะสม กล่าวคือจากลักษณะของ SACF และ SPACF ของเศษเหลือตกค้าง อาจพบว่า ควรเพิ่มองค์ประกอบ MA เข้าในตัวแบบ ถ้ายังไม่มีองค์ประกอบ MA หรือ เพิ่มอันดับของ MA ให้มากขึ้น หรืออาจพบว่าควรเพิ่มองค์ประกอบ AR หรืออันดับของ AR ในตัวแบบ เป็นต้น

2.2.3 <u>การวิเคราะห์การถดถอย (Regression Analysis)</u>

การวิเคราะห์การถดถอย เป็นเทคนิคเชิงสถิติเทคนิคหนึ่งสำหรับการศึกษาวิเคราะห์และจำลองรูปแบบความสัมพันธ์พึ่งพิงเชิงคณิตศาสตร์ระหว่างตัวแปรสองกลุ่ม ตัวแปรกลุ่มหนึ่งเรียกว่า "ตัวแปรตาม" (dependent variable) หรือ "ตัวแปรผล" (response variable) มีหนึ่งตัวแปร เป็นตัวแปรที่นักสถิติหรือนักพยากรณ์ สนใจที่จะศึกษาลักษณะการเปลี่ยนแปลงหรือพยากรณ์ค่าหรือควบคุม โดยศึกษาวิเคราะห์หารูปแบบความสัมพันธ์กับตัวแปรอีกกลุ่มหนึ่ง เรียกตัวแปรในกลุ่มนี้ว่า "ตัวแปรอิสระ" (independent variables) หรือ "ตัวแปรให้ค่าพยากรณ์หรือค่าทำนาย" (predictor variables) ตัวแปรในกลุ่มนี้อาจมีหนึ่งหรือมากกว่าหนึ่งตัวแปร และรูปแบบความสัมพันธ์เชิงคณิตศาสตร์หรือเชิงสถิติที่ได้เรียกว่า "ตัวแบบการถดถอย" หรือ "สมการถดถอย" จากสมการการถดถอยสามารถอธิบายลักษณะการเปลี่ยนแปลงของตัวแปรตาม หรือพยากรณ์ค่าของตัวแปรตามหรือใช้ในการควบคุมตัวแปรตาม โดยใช้รูปแบบสมการและค่าของตัวแปรอิสระ

การหาสมการถดถอย จะมีตัวแปรตาม ให้เป็น Y และมีตัวแปรอิสระกลุ่มหนึ่งให้เป็น $x_1, x_2, \ldots, x_k$ $(k \geq 1)$ โดยทั่วไปเมื่อ Y มีการแจกแจง เราจะใช้วิธีคาดคะเนหรือพยากรณ์ค่าของ Y ด้วยค่าเฉลี่ย $E(Y|x_1, x_2, \ldots, x_k)$ หรือเขียนสั้น ๆ $E(Y)$ ของ Y ซึ่งโดยปกติ $E(Y)$ จะเป็นฟังก์ชัน $g(\underline{x})$ ของ $x_1, x_2, \ldots, x_k$ $[\underline{x} = (x_1, x_2, \ldots, x_k)]$

สมมติ $E(Y)$ หรือ $E(Y/x_1, x_2, \ldots, x_k) = g(\underline{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$ ฉะนั้นสำหรับ $x_1 = x_{i1}, x_2 = x_{i2}, \ldots, x_k = x_{ik}$ ได้ $E(Y) = E(Y_i) = g(\underline{x}_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}$ ซึ่งอาจจะต่างจากค่าจริงหรือค่าสังเกต Y_i แทนผลต่างนี้ด้วย ε_i นั่นคือ

$$\varepsilon_i = Y_i - E(Y_i)$$

จากนี้ได้ว่า

$$\begin{aligned} Y_i &= E(Y_i) + \varepsilon_i \\ &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \qquad (2.3.1) \\ & \qquad i = 1,2,\ldots,n \end{aligned}$$

ผลต่าง ε_i ($i = 1,2,\ldots,n$) เป็นค่าคลาดเคลื่อนหรือค่าผิดพลาดซึ่งรวมค่าผิดพลาดต่าง ๆที่เกิดขึ้น คือสมการ (2.3.1) ยังขาดตัวแปรอิสระที่มีผลต่อค่าของ Y เพราะฉะนั้น ε_i จะรวมผลกระทบต่อค่า Y จากตัวแปรอิสระอื่น ๆ ที่ไม่ใช่ $x_1, x_2, \ldots, x_k$ และอาจมีค่าผิดพลาดซึ่งเกิดขึ้นได้อย่างสุ่ม (โดยไม่ได้ตั้งใจ) ในการวัดค่าหรือบันทึกค่าของ Y_i ดังนั้น พิจารณา ε_i เป็นตัวแปรสุ่ม มีชื่อเรียกว่า "ค่าผิดพลาดสุ่ม" (random error)

พารามิเตอร์ β_j ($j \neq 0$) คือ อัตราเปลี่ยนแปลงโดยเฉลี่ยใน Y เมื่อค่าของตัวแปร x_j เปลี่ยนแปลงไปหนึ่งหน่วยขณะที่ค่าของตัวแปร x_i ($i \neq j$) ที่เหลือทั้งหมดไม่เปลี่ยนแปลง สำหรับ β_0 คือ ค่าเฉลี่ยของ Y เมื่อตัวแปรอิสระทุกตัวมีค่าเป็นศูนย์ แต่ทั้งนี้ β_0 จะไม่ให้ความหมายถ้าหากตัวแบบ (2.3.1) ไม่ครอบคลุมกรณีตัวแปรอิสระทุกตัวเท่ากับศูนย์

คุณสมบัติพื้นฐานของ ε_i คือมีค่าเฉลี่ยเท่ากับศูนย์ ความแปรปรวนเท่ากับ σ^2 สำหรับทุกค่า i และความแปรปรวนร่วมของ $(\varepsilon_i, \varepsilon_j)$, $i \neq j$, ทุกคู่เท่ากับศูนย์หรือไม่มีสหสัมพันธ์กัน และ ε_i มีการแจกแจงความน่าจะเป็นแบบปกติ นั่นคือ ε_i มีข้อสมมติดังนี้

$$E(\varepsilon_i) = 0, \; i = 1,2,\ldots,n \qquad (2.3.2)$$

$$V(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2 \;, \; i=1,2,\ldots,n \qquad (2.3.3)$$

$$Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0 \;, \; i \neq j; \; i,j = 1,2,\ldots,n \qquad (2.3.4)$$

$$\varepsilon_i \sim N(0, \sigma^2) \;, \quad i = 1, 2, \ldots, n \qquad (2.3.5)$$

เราเรียกตัวแบบหรือสมการ (2.3.1) ว่า "ตัวแบบการถดถอยเชิงเส้น" (linear regression model) ซึ่งเป็นการอธิบายความสัมพันธ์ Y บนกลุ่มตัวแปรอิสระ $x_1, x_2, \ldots, x_k$ พารามิเตอร์หรือค่าคงที่ไม่ทราบค่า $\beta_0, \beta_1, \ldots, \beta_k$ มีชื่อเรียกว่า "สัมประสิทธิ์การถดถอย" (regression coefficients)

ตัวแบบการถดถอยที่มีตัวแปรอิสระเพียงหนึ่งตัวมีชื่อเรียกว่า "ตัวแบบการถดถอยเชิงเดียว" (simple regression model) ในกรณีที่มีตัวแปรอิสระมากกว่าหนึ่งตัวมีชื่อเรียกว่า "ตัวแบบการถดถอยเชิงพหุคูณ" (multiple regression model)

ดังนั้น ตัวแบบหรือสมการถดถอยเชิงเส้นเชิงเดียว หรือตัวแบบการถดถอยเชิงเส้นอย่างง่าย (simple linear regression model) จะแสดงความสัมพันธ์เป็นเชิงเส้นระหว่างตัวแปรตาม Y และตัวแปรอิสระ x หนึ่งตัว เขียนเป็นตัวแบบทั่วไปได้ดังนี้

$$\begin{aligned} Y &= E(Y|x) + \varepsilon \\ &= \alpha + \beta x + \varepsilon \quad [E(Y|x) = \alpha + \beta x] \end{aligned}$$

หรือ

$$Y_i = \alpha + \beta x_i + \varepsilon_i \qquad i = 1, 2, \ldots, n \tag{2.3.6}$$

โดยที่ ε_i มีข้อสมมติ (2.3.2) ถึง (2.3.5) ซึ่งข้อสมมติของ ε_i และ Y_i เสมอเหมือนกัน

2.2.3.1 ตัวแบบการถดถอยเชิงเส้นพหุคูณ

. (Multiple Linear Regression Models)

ในกรณีวิเคราะห์รูปแบบความสัมพันธ์พึ่งพิงเชิงเส้นของตัวแปรตาม Y บนตัวอิสระมากกว่าหนึ่งตัวแปร $x_1, x_2, \ldots, x_k$ เราจะเรียกรูปแบบความสัมพันธ์นั้นว่า "ตัวแบบการถดถอยเชิงเส้นพหุคูณ" หรือ "สมการการถดถอยเชิงเส้นพหุคูณ" และมีรูปแบบทั่วไปดังนี้

$$\begin{aligned} Y &= E(Y) + \varepsilon \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \end{aligned} \tag{2.3.7}$$

โดยที่ค่าเฉลี่ย E(Y) หรือ $E(Y|x_1, x_2, \ldots, x_k)$ ของ Y เท่ากับ $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$ เมื่อกำหนดตัวแปรอิสระ $x_1, x_2, \ldots, x_k$ สำหรับตัวอย่างสุ่ม $(x_{i1}, x_{i2}, \ldots, x_{ik}, Y_i)$,

$i = 1, 2, \ldots, n$, นั่นคือ $x_1 = x_{i1}$, $x_2 = x_{i2}, \ldots, x_k = x_{ik}, Y = Y_i$, และให้ $\varepsilon = \varepsilon_i$ สำหรับ $i = 1, 2, \ldots, n$ เขียนตัวแบบ (2.3.16) ได้ใหม่ดังนี้

$$Y_i = E(Y_i) + \varepsilon_i$$

$$= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \quad , i = 1,2,\ldots, n \qquad (2.3.8)$$

และ ε_i , $i=1,2,\ldots,n$, มีข้อสมมติ (2.3.2) ถึง (2.3.5)

วิธีประมาณค่าของพารามิเตอร์มีหลายวิธีด้วยกัน วิธีหนึ่งที่ใช้กันมากคือ "วิธีกำลังสองน้อยสุดแบบสามัญ" หรือวิธี OLS (Ordinary Least Squares Method) ด้วยวิธี OLS ตัวประมาณแบบกำลังสองน้อยที่สุด $B_0, B_1, \ldots, B_k$ ของ $\beta_0, \beta_1, \ldots, \beta_k$ ตามลำดับ คำนวณหาได้โดยการแก้สมการ

$$\left.\frac{\partial Q}{\partial \beta_i}\right|_{\beta_0 = B_0, \ldots, \beta_k = B_k} = 0, \quad i = 0, 1, 2, \ldots, k$$

โดยที่

$$Q = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 x_{i1} \ldots - \beta_k x_{ik}\right)^2$$

หรือนั่นคือ โดยการแก้ระบบสมการ (2.3.9):

$$nB_0 + B_1 \sum X_{i1} + B_2 \sum_{i=1}^{n} X_{i2} + \ldots + B_k \sum_{i=1}^{n} X_{ik} = \sum_{i=1}^{n} Y_i$$

$$B_0 \sum_{i=1}^{n} X_{i1} + B_1 \sum X^2{}_{i1} + B_2 \sum_{i=1}^{n} X_{i1} X_{i2} + \ldots + B_k \sum_{i=1}^{n} X_{i1} X_{ik} = \sum_{i=1}^{n} X_{i1} Y_i$$

$$B_0 \sum_{i=1}^{n} X_{i2} + B_1 \sum X_{i2} X_{i1} + B_2 \sum_{i=1}^{n} X^2{}_{i2} + \ldots + B_k \sum_{i=1}^{n} X_{i2} X_{ik} = \sum_{i=1}^{n} X_{i2} Y_i$$

. . .

$$B_0 \sum_{i=1}^{n} X_{ik} + B_1 \sum X_{ik} X_{i1} + B_2 \sum_{i=1}^{n} X_{ik} X_{i2} + \ldots + B_k \sum_{i=1}^{n} X^2{}_{ik} = \sum_{i=1}^{n} X_{ik} Y_i$$

(2.3.9)

โดยทั่วไป การวิเคราะห์การถดถอยเชิงพหุคูณ จะใช้เมตริกซ์เป็นเครื่องมือซึ่งจะทำให้การวิเคราะห์สะดวกมากขึ้น ฉะนั้นให้

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ Y_n \end{bmatrix} \qquad \underline{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1k} \\ 1 & X_{21} & X_{22} & \dots & X_{2K} \\ \dots & & & & \\ 1 & X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix}$$

$$\underline{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_k \end{bmatrix} \qquad \underline{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_n \end{bmatrix} \qquad \underline{B} = \begin{bmatrix} B_0 \\ B_1 \\ . \\ . \\ B_k \end{bmatrix}$$

ดังนั้น $\underline{Y}$ เป็นเมตริกซ์ขนาด $n \times 1$ ของตัวแปรสุ่ม หรือเวกเตอร์ขนาด n ของตัวแปรสุ่ม ,$\underline{\beta}$ เป็นเวกเตอร์ขนาด k+1 ของพารามิเตอร์, $\underline{\varepsilon}$ เป็นเวกเตอร์ขนาด n ของตัวแปรสุ่มค่าผิดพลาด , $\underline{B}$ เป็นเวกเตอร์ขนาด k+1 ของตัวประมาณของพารามิเตอร์, และ $\underline{X}$ เป็นเมตริกซ์ขนาด $n \times (k+1)$ ของค่าคงที่ 1 และค่าของตัวแปรอิสระ

โดยการใช้สัญลักษณ์เมตริกซ์ข้างต้น เขียนระบบสมการ (2.3.8) ได้ใหม่ดังนี้

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon} \qquad (2.3.10)$$

และเขียนข้อสมมติ (2.3.2) ถึง (2.3.5) ได้สั้นๆดังนี้

$$\underline{\varepsilon} \sim N_n(\underline{0}, \underline{I}\sigma^2)$$

ซึ่งหมายความว่า $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ เป็นอิสระกัน และต่างมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน σ^2 $(N(0, \sigma^2))$

ในทำนองเดียวกัน เขียนข้อสมมติของ Y_i , $i = 1,2,\ldots,n$ ได้ดังนี้

$$\underline{Y} \sim N_n\left(\underline{X}\underline{\beta}\,,\underline{I}\sigma^2\right) \qquad (2.3.11)$$

จากระบบสมการปกติ (2.3.9) เขียนในเทอมของเมตริกซ์ได้สั้น ๆ เป็น

$$\left(\underline{X}'\underline{X}\right)\underline{B} = \underline{X}'\underline{Y} \qquad (2.3.12)$$

ซึ่ง $\underline{X}'$ หมายถึง เมตริกซ์สับเปลี่ยน (transposed matrix) ของเมตริกซ์ $\underline{X}$

$$\underline{X} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_{11} & x_{21} & \ldots & x_{n1} \\ x_{12} & x_{22} & \ldots & x_{n2} \\ . & . & \ldots & . \\ x_{1k} & x_{2k} & \ldots & x_{nk} \end{bmatrix}$$

การแก้สมการ (2.3.12) หา $\underline{B}$ จะสมมติว่าหาเมตริกซ์ผกผัน $\left(\underline{X}'\underline{X}\right)^{-1}$ ของเมตริกซ์ $\underline{X}'\underline{X}$ ได้ ซึ่งเป็นจริงโดยทั่วไปในทางปฏิบัติ เพราะฉะนั้น ตัวประมาณแบบกำลังสองน้อยที่สุดสามัญคือ

$$\underline{B} = \left(\underline{X}'\underline{X}\right)^{-1}\underline{X}'\underline{Y} \qquad (2.3.13)$$

และตัวแบบพยากรณ์ค่า Y หรือตัวประมาณของค่าเฉลี่ย E(Y) ของ Y เมื่อกำหนด $x_1 = x_{01}$, $x_2 = x_{02}, \ldots, x_k = x_{0k}$ คือ

$$\hat{Y} = \underline{x}_0{}'\,\underline{B}$$

$$= B_0 + B_1x_{01} + B_2x_{02} + \ldots + B_kx_{0k}$$

ซึ่ง $\underline{x}_0{}' = \left(x_{01}, x_{02}, \ldots, x_{0k}\right)$

2.2.3.2 ขั้นตอนการสร้างตัวแบบการถดถอยเชิงเส้น

กรรมวิธีการสร้างตัวแบบการถดถอยเชิงเส้น มีขั้นตอนที่ควรดำเนินการดังนี้

1 กำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ

กำหนดรูปแบบความสัมพันธ์ อาจใช้ทฤษฎีที่เกี่ยวข้อง เช่น ทฤษฎีเศรษฐศาสตร์ มีการกำหนดรูปแบบความสัมพันธ์ที่เรียกว่าตัวแบบเศรษฐศาสตร์ ซึ่งอาจจะนำมาประยุกต์ได้กับเรื่องที่ศึกษา ในกรณีที่ไม่สามารถหาทฤษฎีใดมาประยุกต์ได้ นักพยากรณ์จะพิจารณารูปแบบความสัมพันธ์ โดยอาศัยข้อมูลที่มีอยู่ของตัวแปรตาม และของตัวแปรอิสระ ซึ่งสามารถพิจารณาโดยใช้กราฟดังนี้

1.1 โดยทั่วไปจะมีข้อสมมติว่าตัวแปรตาม Y มีการแจกแจงความน่าจะเป็นแบบปกติ ฉะนั้น ควรเขียนกราฟแผนภาพแบบจุด หรือแผนภาพฮิสโทแกรม (histogram diagram) เพื่อดูลักษณะการกระจาย หรือการแจกแจงของ Y ว่าเข้ารูปลักษณะแบบสมมาตรหรือไม่ ถ้าพบว่ามีลักษณะไม่สมมาตร โดยเบ้ไปทางซ้ายหรือขวามาก ควรที่จะแปลงข้อมูลของ Y เพื่อให้เข้าลักษณะการแจกแจงแบบสมมาตร วิธีการแปลงค่าของ Y อาจจะทดลองด้วยแบบต่างๆ เช่น $\sqrt{Y}$, $1/\sqrt{Y}$, $1/Y$, $\ln Y$, หรือ $\log_{10} Y$

1.2 เขียนกราฟระหว่างตัวแปรตาม กับตัวแปรอิสระทีละตัว เพื่อพิจารณากำหนดรูปแบบความสัมพันธ์เป็นคู่ ๆ ระหว่างตัวแปรตามและตัวแปรอิสระ

1.3 การแปลงตัวแบบการถดถอยไม่เป็นเชิงเส้นให้เป็นตัวแบบการถดถอยเชิงเส้น ตัวแบบเริ่มแรกที่นักพยากรณ์เลือก อาจไม่อยู่ในลักษณะเหมือนตัวแบบทั่วไป (2.3.1) ซึ่งอยู่ในรูปแบบเชิงเส้นทั้งในเทอมของพารามิเตอร์ และในเทอมของตัวแปร ในกรณีของตัวแปร เราสามารถแปลงให้อยู่ในแบบเชิงเส้นได้ง่ายดังตัวอย่างต่อไปนี้

ตัวอย่างการแปลงตัวแปร

$\ln Y = \alpha + \beta x + \varepsilon$ ให้ $Y' = \ln Y$ ได้ตัวแบบ

$$Y' = \alpha + \beta x + \varepsilon$$

การแปลงตัวแปรตาม Y ในตัวอย่าง ถ้า Y [หรือ] มีคุณสมบัติ (2.3.2) ถึง (2.3.5) เราสามารถใช้วิธี OLS และวิธีการอนุมานเชิงสถิติทั่วไปในการวิเคราะห์ และ

เมื่อได้ค่าพยากรณ์ของ Y' จะแปลงกลับได้ค่าพยากรณ์ของ Y เช่นจาก $Y'=\ln Y$ ได้ $Y=e^{Y'}$

2. การประมาณค่าพารามิเตอร์ของตัวแบบการถดถอย

เมื่อนักพยากรณ์กำหนดตัวแบบการถดถอยเป็นตัวแบบทดลองได้แล้ว ซึ่งอาจจะมีมากกว่าหนึ่งตัวแบบ ขั้นตอนต่อไปก็คือ ประมาณค่าพารามิเตอร์ในตัวแบบ ในกรณีของตัวแบบการถดถอยเชิงเส้น และภายใต้ข้อสมมติ (2.3.2) ถึง (2.3.5) วิธีมาตรฐานที่ใช้กันทั่วไปคือวิธี OLS ในการประมาณค่าพารามิเตอร์ และใช้วิธีวิเคราะห์สถิติมาตรฐานทั่วไปในการอนุมานเกี่ยวกับพารามิเตอร์ ตลอดจนค่าพยากรณ์แบบช่วง

3. การวินิจฉัยความเพียงพอของตัวแบบการถดถอย

งานขั้นกำหนดรูปแบบของตัวแบบ และงานประมาณค่าพารามิเตอร์ของตัวแบบเป็นเพียงงานครึ่งหนึ่งของกรรมวิธีสร้างตัวแบบพยากรณ์ ตัวแบบพยากรณ์ที่ได้จากงานขั้นที่ 1 และ 2 อาจยังไม่เหมาะสมหรือยังไม่เพียงพอที่จะใช้พยากรณ์ นักพยากรณ์จึงควรตรวจสอบและทำการเปรียบเทียบคัดเลือกตัวแบบพยากรณ์ ถ้าตรวจสอบพบว่า ตัวแบบที่กำลังพิจารณายังขาดความเหมาะสม จะกลับไปทำงานในขั้นที่ 1 ถึง 3 ซ้ำ ๆ จนกว่าจะได้ตัวแบบที่เหมาะสมเพียงพอในเชิงสถิติที่จะใช้พยากรณ์ค่าต่อไป

เนื่องจากการอนุมานเชิงสถิติเกี่ยวกับพารามิเตอร์ ซึ่งเกี่ยวกับการทดสอบข้อสมมติฐานต่าง ๆ และการประมาณค่าแบบช่วงของค่าพารามิเตอร์ ตลอดจนการทดสอบและการประมาณค่าแบบช่วงของค่าเฉลี่ย $E(Y)$ ของ Y และของค่าพยากรณ์ของค่าจริง Y โดยทั่วไปการทดสอบและการประมาณค่าดังกล่าว จะกระทำภายใต้ข้อสมมติ (2.3.2) ถึง (2.3.5) ดังนั้น จึงจำเป็นต้องตรวจสอบความเหมาะสม หรือความเพียงพอของตัวแบบด้วยการตรวจสอบคุณสมบัติของค่าผิดพลาด ε_i แต่เนื่องจากไม่ทราบค่าจริง ε_i ฉะนั้นจะตรวจสอบคุณสมบัติของค่าเศษเหลือตกค้าง e_i ซึ่ง $e_i = y_i - \hat{y}_i$ เป็นค่าประมาณของ ε_i เมื่อตัวแบบถูกต้องเพียงพอ และตรวจสอบว่าค่า e_i $(i=1,2,\dots,n)$ มีคุณสมบัติสอดคล้อง (2.3.2) ถึง (2.3.5) หรือไม่ นั่นคือ มีค่าเฉลี่ยเท่ากับศูนย์, มีความแปรปรวนคงที่ , ไม่มีอัตสหสัมพันธ์ , และมีการแจกแจงแบบปกติ หรือไม่ นอกจากนี้อาจพบว่า ตัวแบบไม่เหมาะสมเนื่องมาจากตัวแบบมีรูปแบบยังไม่ถูกต้องเหมาะสม

วิธีการตรวจสอบนักพยากรณ์อาจเลือกใช้วิธีกราฟ หรือวิธีการทดสอบเชิงสถิติ ซึ่งวิธีเชิงสถิติเป็นวิธีที่มีระเบียบ(formal) หรือมีทฤษฎี อย่างไรก็ตาม ถึงแม้ว่าวิธีกราฟอาจจะไม่เป็นวิธีเชิงระเบียบทฤษฎี แต่โดยทั่วไปก็เป็นวิธีการที่เพียงพอที่จะใช้วินิจฉัยตัวแบบ และเป็นวิธีการง่ายที่ใช้โดยทั่ว ฉะนั้นในที่นี้จะกล่าวถึงเฉพาะวิธีกราฟ

โดยวิธีกราฟ เราจะเขียนกราฟของค่าเศษเหลือกตกค้าง e_i หรือค่าเศษเหลือตกค้างมาตรฐาน(standardized residuals) $e_i / \sqrt{MSE}$ กับตัวแปรต่าง ๆ

1. ตัวแปร $\hat{y}_i$
2. ตัวแปรอิสระ x_i แต่ละตัว
3. ตัวแปรเวลา ถ้าข้อมูลเป็นอนุกรมเวลา
4. เขียนกราฟความน่าจะเป็นแบบปกติ (normal probability plot) หรือ แผนภาพฮิสโทแกรมของ e_i หรือ $e_i / \sqrt{MSE}$

ถ้ากราฟระหว่าง e_i หรือ $e_i / \sqrt{MSE}$ และ $\hat{y}_i$, และตัวแปรอิสระแต่ละตัว, และกับเวลา(ถ้าเป็นอนุกรมเวลา) ทั้งหมดมีรูปแบบการกระจายของจุดเป็นแนวขนานแสดงว่าตัวแบบการถดถอยเชิงเส้นมีรูปแบบเหมาะสม ค่าผิดพลาดมีค่าเฉลี่ยศูนย์และความแปรปรวนคงที่ นอกจากนี้ในกรณีที่เป็นกราฟ e_i กับเวลา แสดงด้วยว่าความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระไม่แปรเปลี่ยนตามเวลา หรือไม่มีอิทธิพลของเวลา และค่าคลาดเคลื่อนสุ่มไม่มีอัตสหสัมพันธ์กัน แต่ถ้ากราฟมีรูปแบบไม่เป็นแนวขนาน แสดงว่าค่าผิดพลาดไม่สอดคล้องกับคุณสมบัติบางข้อหรือทุกข้อในข้อ (2.3.2) ถึง (2.3.4) หรือตัวแบบยังมีรูปแบบไม่ถูกต้อง

การตรวจสอบค่าผิดพลาดมีสหสัมพันธ์กันหรือไม่ โดยปกติจะตรวจสอบเมื่อข้อมูลเป็นอนุกรมเวลาและวิธีการตรวจสอบมีหลายวิธี เช่นใช้วิธีพิจารณาค่าของฟังก์ชันอัตสหสัมพันธ์ตัวอย่าง[Sample Autocorrelation Function (SACF)] $\left| r_k \right|$ ของ $e_1, e_2, \ldots, e_n$ ที่คาบเวลาห่างกัน k (k = 1 , 2 , . . .) เปรียบเทียบกับค่าตัดสินเชิงสถิติหรือค่าวิกฤต (Critical value) $2 / \sqrt{n}$ โดยประมาณ ที่ระดับนัยสำคัญ (significance level) 0.05 หรือใช้ค่าของตัวสถิติ Durbin-Watson สำหรับตรวจสอบอัตสหสัมพันธ์ที่คาบเวลาห่างกัน 1 (k=1) เมื่อพบว่าค่าผิดพลาดมีสหสัมพันธ์กัน ในการวิเคราะห์

การถดถอยเชิงเส้น จะมีวิธีการแก้ปัญหานี้โดยเฉพาะ เช่นวิธีการแปลงตัวแปร และวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก เป็นต้น

การตรวจสอบคุณสมบัติว่าค่าผิดพลาดมีการแจกแจงปกติหรือไม่ อาจใช้วิธีเขียนแผนภาพฮิสโทแกรม และ/หรือกราฟความน่าจะเป็นแบบปกติ ถ้าพบว่าแผนภาพฮิสโทแกรมมีลักษณะโค้งสมมาตร หรือเส้นกราฟความน่าจะเป็นเป็นแนวเส้นตรง ในกรณีนี้ก็มีเหตุผลว่าค่าผิดพลาดมีการแจกแจงเข้ารูปแบบปกติ แต่ถ้าฮิสโทแกรมเบ้มาก หรือเส้นกราฟความน่าจะเป็นไม่เป็นแนวเส้นตรง ในกรณีนี้ ค่าผิดพลาดจะไม่สอดคล้องคุณสมบัติการแจกแจงปกติ ควรปรับปรุงแก้ไขตัวแบบ เช่นใช้วิธีการแปลงค่าตัวแปรตาม Y ตามที่กล่าวไปแล้ว อย่างไรก็ตาม การแจกแจงที่ไม่เข้าสู่การแจกแจงแบบปกติอาจไม่เป็นปัญหามากนัก และอาจพบว่าเมื่อแก้ปัญหาในข้ออื่น ๆ ที่กล่าวก่อนหน้านี้ ปัญหาการแจกแจงไม่เป็นแบบปกติอาจจะถูกแก้ไขไปด้วยโดยไม่ต้องแก้ไขโดยตรง

นอกจากการตรวจสอบคุณสมบัติต่าง ๆ ของค่าผิดพลาดสุ่ม และความถูกต้องของตัวแบบการถดถอยเชิงเส้นแล้ว ในการวิเคราะห์การถดถอย จะมีการคำนวณค่าของตัวสถิติต่าง ๆ สำหรับทดสอบผลกระทบของแต่ละตัวแปรอิสระต่อค่าตัวแปรตาม Y ว่ามีมากพอที่จะยอมรับได้หรือไม่ในเชิงสถิติ รวมทั้งการตรวจสอบตัวแปรอิสระพร้อมกันทุกตัว (ถ้ามีมากกว่าหนึ่งตัว) ว่าตัวแปรอิสระทั้งกลุ่มมีผลต่อค่า Y มากพอหรือไม่ หรือช่วยให้ค่าพยากรณ์ถูกต้องมากขึ้นหรือไม่ (จากเดิมที่ใช้ค่าเฉลี่ย Y เท่านั้นโดยไม่อาศัยตัวแปรอิสระใด ๆ) ถ้าพบว่าตัวแปรอิสระตัวใดมีผลกระทบน้อยจนไม่อาจยอมรับได้ในเชิงสถิติ ก็ควรตัดตัวแปรนั้นออกจากตัวแบบ การตรวจสอบหรือการทดสอบทางสถิติเกี่ยวกับผลกระทบของตัวแปรอิสระทีละตัว เราใช้ตัวสถิติทดสอบที (t statistic) และการทดสอบผลกระทบรวมกันเป็นกลุ่ม เราใช้ตัวสถิติทดสอบเอฟ (F statistic)

4. การเปรียบเทียบตัวแบบพยากรณ์

ถ้านักพยากรณ์ได้สร้างตัวแบบทดลองขึ้นมาหลายตัวแบบ และแต่ละตัวแบบได้ผ่านการวินิจฉัยแล้ว เมื่อต้องการคัดเลือกตัวแบบหนึ่งเป็นตัวแบบที่จะใช้พยากรณ์ นักพยากรณ์อาจจะคัดเลือกโดยใช้วิธีการเปรียบเทียบค่าความถูกต้องของค่าพยากรณ์

2.2.4 วิธีแยกองค์ประกอบ(Decomposition Method)

อนุกรมเวลา $\{Y_t\}$ จะมีการเคลื่อนไหวตามกาลเวลาในรูปแบบต่าง ๆ เช่นมีแนวโน้ม และมีฤดูกาลฉะนั้น หนทางหนึ่งในการวิเคราะห์อนุกรมเวลาเพื่อหาตัวแบบพยากรณ์ คือ วิเคราะห์หาองค์ประกอบในข้อมูลอนุกรมเวลาที่ศึกษามาสร้างเป็นตัวแบบพยากรณ์

องค์ประกอบของอนุกรมเวลา จำแนกได้เป็น 3 องค์ประกอบหลักคือ องค์ประกอบแนวโน้ม-วัฏจักร(trend-cycle component) องค์ประกอบฤดูกาล(seasonal component) และองค์ประกอบไม่ปรกติหรือองค์ประกอบส่วนเหลือ(irregular or remainder component) และดังนั้นโดยวิธีแยกองค์ประกอบจะได้ตัวแบบสำหรับการพยากรณ์ค่า Y_t และเรียกตัวแบบที่ได้ว่าตัวแบบแยกองค์ประกอบ ซึ่งจำแนกได้เป็นสองตัวแบบคือ

1. ตัวแบบเชิงบวก(Additive Model)

$$Y_t = T_t + S_t + \varepsilon_t$$

2. ตัวแบบเชิงคูณ(Multiplicative Model)

$$Y_t = T_t \times S_t \times \varepsilon_t$$

โดยที่

Y_t คือค่าของอนุกรมเวลา ณ คาบเวลา t

ε_t คือค่าคลาดเคลื่อนสุ่ม

T_t คือองค์ประกอบแนวโน้ม-วัฏจักร ณ คาบเวลา t ซึ่งโดยทั่วไปจะอยู่ในรูปแบบพหุนามอันดับต่ำ(low-order polynomial) เช่น

$T_t = \beta_0 + \beta_1 t$ และ $T_t = \beta_0 + \beta_1 t + \beta_2 t^2$

S_t คือองค์ประกอบฤดูกาล ณ คาบเวลา t ซึ่งอาจจะอยู่ในรูปแบบของตัวแปรบ่งชี้ฤดูกาล(seasonal indicators) เช่น

$S_t = \alpha_1 X_{1t} + \alpha_2 X_{2t} + \ldots + \alpha_{11} X_{11t}$ (X_{it} เป็นตัวแปรบ่งชี้ฤดูกาล)

หรืออยู่ในรูปฟังก์ชันตรีโกณ เช่น

$$S_t = \phi_1 \sin\left(\frac{2\pi t}{12}\right) + \phi_2 \cos\left(\frac{2\pi t}{12}\right) + \phi_3 \sin\left(\frac{4\pi t}{12}\right) + \phi_4 \cos\left(\frac{4\pi t}{12}\right)$$

ตัวแบบเชิงบวก เหมาะสำหรับข้อมูลอนุกรมเวลาที่มีความแปรปรวนคงที่ ไม่แปรผันตามเวลาหรือการแกว่งของฤดูกาลไม่แปรผันตามระดับของอนุกรมเวลา มิฉะนั้น ตัวแบบเชิงคูณจะเหมาะสมกว่า อย่างไรก็ตาม เราสามารถใช้วิธีการแปลงข้อมูลอนุกรมเวลาที่ไม่คงที่ในความแปรปรวนหรือแปลงตัวแบบเชิงคูณเป็นตัวแบบเชิงบวกได้ โดยการใส่ ln ซึ่งจะได้ตัวแบบ

$$\ln Y_t = \ln T_t + \ln S_t + \ln \varepsilon_t$$

นั่นคือ ใช้ตัวแบบเชิงบวกกับข้อมูลที่ใส่ ln

ตัวอย่างตัวแบบพยากรณ์แบบแยกองค์ประกอบ :

$$Y_t = b_0 + b_1 t_t + b_2 X_{1t} + b_3 X_{2t} + \ldots + b_{12} X_{11t}$$

โดย t แทนคาบเวลา

X_{it} (i = 1,2, . . .,11) แทนตัวบ่งชี้ สำหรับเดือนที่ i ในคาบเวลา t (ให้ $X_{it} = 1$ สำหรับเดือนที่ i ในคาบเวลา t มิฉะนั้นให้ $X_{it} = 0$)