หอสมุดกลาง จุฬาลงกรณมหาวิทยาลัย

บทที่ ๒
วิธีการประมาณเชิงสถิติ

๒.๑ วิธีการประมาณเชิงสถิติโดยใช้วิธี Time Series Decomposition

อนุกรมเวลา (Time Series) ซึ่งโดยทั่ว ๆ ไปประกอบด้วย

๑) ทางโน้มตามลำดับเวลา (Secular trend) เป็นการเคลื่อนไหวขึ้นลงของข้อมูลในระยะยาว แสดงการเคลื่อนไหวขึ้นลงของข้อมูลในช่วงเวลายาว ซึ่งอาจเป็นเส้นตรงหรือเส้นโค้งก็ได้

๒) การแปรผันเนื่องจากฤดูกาล (Seasonal movement) เป็นการเคลื่อนไหวซึ่งค่อนข้างสม่ำเสมอภายในระยะเวลา ๑ ปี ความเคลื่อนไหวชนิดนี้เกิดขึ้นเนื่องจากการเปลี่ยนแปลงของฤดูกาล

๓) การเคลื่อนไหวเป็นวัฏจักร (Cyclical movement) เป็นการเคลื่อนไหวคล้ายการแปรผันอันเนื่องจากฤดูกาล แต่ช่วงเวลานานกว่า ๑ ปี ดังนั้นจึงคล้ายกับ secular trend ด้วย แต่จะต่างกันที่ว่าการเคลื่อนไหวเป็นวัฏจักรนั้นมีทั้งระยะเวลารุ่งเรืองถึงสุดยอดจนกระทั่งตกต่ำสุด

๔) ความเคลื่อนไหวโดยบังเอิญ (Irregular movement) เป็นการเคลื่อนไหวที่อยู่นอกเหนือธรรมดา เป็นการเปลี่ยนแปลงซึ่งเป็นผลกระทบกระเทือนจากภาวะสงคราม ความหายนะ การหยุดงาน หรือเหตุอื่น ๆ ซึ่งล้วนเป็นเหตุการณ์ที่เกิดขึ้นเฉพาะคราว ซึ่งเราไม่สามารถจะคาดคะเนได้ล่วงหน้า

<u>โมเดลอนุกรมเวลา</u>

ถ้ากำหนดให้ Y = ข้อมูลที่สนใจศึกษา

T = ทางโน้มตามลำดับเวลา

S = การแปรผันเนื่องจากฤดูกาล

C = การเคลื่อนไหวเป็นวัฏจักร

I = การแปรผันโดยบังเอิญ

ส่วนประกอบของอนุกรมเวลาอาจจำแนกได้เป็น ๒ ลักษณะคือ

๑) ในลักษณะผลบวก $Y = T + S + C + I$

หอสมุดกลาง จุฬาลงกรณมหาวิทยาลัย

๒) ในลักษณะผลคูณ

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

แต่โดยทั่ว ๆ ไปมักนิยมใช้แบบผลคูณ เพราะส่วนมากของข้อมูลที่เกี่ยวกับอนุกรมเวลานั้น เป็นข้อมูลในทางเศรษฐกิจและโมเดลต่าง ๆ ที่ใช้วิเคราะห์ข้อมูลในทางเศรษฐกิจมักนิยมในรูปของผลคูณ

การประมาณโดยวิธี Time Series Decomposition อาศัยการวิเคราะห์ส่วนประกอบของอนุกรมเวลาที่อยู่ในลักษณะของผลคูณคือ

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

เนื่องจากส่วนที่เป็น I คือการเคลื่อนไหวโดยบังเอิญ เป็นการเคลื่อนไหวที่เราไม่สามารถคาดคะเนได้ล่วงหน้า เพราะเป็นการเคลื่อนไหวที่ขึ้นอยู่กับธรรมชาติหรือเหตุผิดปกติต่าง ๆ ซึ่งนาน ๆ จะเกิดขึ้นสักครั้งจึงจะไม่นำมารวมในการวิเคราะห์ และส่วนที่เป็น C คือการเคลื่อนไหวเป็นวัฏจักร เป็นการเคลื่อนไหวในระยะยาวเป็นสิบ ๆ ปีขึ้นไป แต่ในการทำวิจัยครั้งนี้เราทำกับข้อมูลเพียง ๑๒๐ เดือนจึงคาดว่าส่วนที่เป็น C จะไม่มีผลต่อข้อมูล ดังนั้น ส่วนประกอบของอนุกรมเวลาสำหรับการวิเคราะห์ในที่นี้จะมีเพียง ๒ อย่างคือ

$$Y = T \times S$$

ข้อมูลที่เก็บรวบรวมมาทำการวิเคราะห์ในที่นี้เป็นปริมาณสินค้าเป็นรายเดือน เป็นเวลา ๑๐ ปี คือตั้งแต่ พ.ศ.๒๕๐๖ – พ.ศ.๒๕๑๕ ดังนั้นจึงสมมติให้

Y_{ij} = ปริมาณสินค้าเดือนที่ j ของปีที่ i

T_i = ทางโน้มตามลำดับเวลาของปีที่ i

S_j = ดรรชนีการแปรผันเนื่องจากฤดูกาลของเดือนที่ j

ε_{ij} = error term

โดยตั้งอยู่บนข้อกำหนด (assumption) ที่ว่า

๑) $\sum_{j=1}^{12} S_j = 12$

๒) ε_{ij} มีการกระจายแบบปกติ มีค่าเฉลี่ยเท่ากับ ๑ และความแปรปรวนเท่ากับ σ^2

จะได้โมเดลที่ใช้ในการศึกษาปริมาณสินค้าเข้าและสินค้าออก คือ

$$Y_{ij} = T_i \times S_j \times \varepsilon_{ij}$$

ในการทำวิทยานิพนธ์เรื่องนี้ มุ่งที่จะพยากรณ์ปริมาณสินค้าเข้าและสินค้าออกในระยะ ๑๐ ปีข้างหน้าให้ได้ใกล้เคียงความจริงที่สุดเท่าที่จะทำได้ ดังนั้นในการประมาณจึงจำเป็นต้องขจัดส่วนที่เป็นความแปรผันเนื่องจากฤดูกาลออกจากข้อมูลเสียก่อน ให้เหลือแต่ส่วนที่เป็น trend อย่างเดียว เพื่อใช้ส่วนที่เป็น trend นี้มาประมาณปริมาณสินค้าเข้าและสินค้าออกใน ๑๐ ปีข้างหน้า ซึ่งจะทำให้ได้ปริมาณที่ขึ้นกับเวลาล้วน ๆ แล้วจึงจะนำส่วนที่เป็นการแปรผันเนื่องจากฤดูกาลเข้ามาปรับอีกครั้งด้วยวิธีการนี้ จะทำให้ได้ปริมาณสินค้าเข้าและสินค้าออกที่เป็นไปโดยธรรมชาติของอนุกรมเวลา ซึ่งคาดว่าน่าจะให้ผลใกล้เคียงกับความเป็นจริง

๒.๑.๑ <u>วิธีเฉลี่ยเคลื่อนที่</u> (Moving Average Method)

วิธีกำจัด S เพื่อหาส่วนที่เป็นทางโน้มตามลำดับเวลา (T_i) ทำได้หลายวิธี วิธีที่นิยมและใช้กันทั่วไปคือ วิธีเฉลี่ยเคลื่อนที่ เพราะวิธีนี้สามารถจะขจัดการกระเพื่อมขึ้นลงของข้อมูลให้ลดลงได้ และเนื่องจากข้อมูลที่เก็บมาในการทำวิทยานิพนธ์เรื่องนี้ มีการกระเพื่อมขึ้นลงมาก เพื่อที่จะขจัดการกระเพื่อมขึ้นลงของปริมาณสินค้าแต่ละอย่าง จึงเลือกวิธีเฉลี่ยเคลื่อนที่มาใช้ในการคำนวณหาทางโน้มตามลำดับเวลา (T_i) วิธีนี้ค่อย ๆ เฉลี่ย (หามัชฌิมเลขคณิต) ข้อมูลทีละหมู่ในการเฉลี่ยค่อย ๆ ทิ้งข้อมูลตัวแรกทีละตัวแล้วใช้ข้อมูลใหม่ถัดไปแทนที่ทีละตัว ข้อมูลที่เก็บมาในการทำวิทยานิพนธ์เรื่องนี้เป็นรายเดือนใน ๑ ปี (period) มี ๑๒ เดือน จึงใช้วิธีเฉลี่ยเคลื่อนที่ซึ่งมีชื่อเฉพาะว่า Centered 12 - month moving average method เพราะวิธีนี้เป็นวิธีพิเศษที่ใช้เฉพาะกรณีที่ใช้ข้อมูลหมู่ละ ๑๒ ตัว หาค่าเฉลี่ยเคลื่อนที่ซึ่งมีวิธีการดังนี้

๒.๑.๑.๑ หาผลรวมของ ๑๒ เดือน ผลรวมของ ๑๒ เดือนแรกคือผลรวมของเดือนมกราคม ปี ๒๕๐๖ – เดือนธันวาคม ปี ๒๕๐๖ ต่อไปหาผลรวมของ ๑๒ เดือนที่สอง โดยทิ้งข้อมูลเดือนมกราคม ปี ๒๕๐๖ แต่ใช้ข้อมูลเดือนมกราคม ปี ๒๕๐๗ แทน ผลรวมของ ๑๒ เดือนที่สองจึงเริ่มแต่เดือนกุมภาพันธ์ ปี ๒๕๐๖ ถึงเดือนมกราคมปี ๒๕๐๗ แล้วหาผลรวมของ ๑๒ เดือนที่สาม โดยทิ้งข้อมูลเดือนกุมภาพันธ์ ปี ๒๕๐๖ แต่ใช้ข้อมูลเดือนกุมภาพันธ์

ปี ๒๕๐๗ แทน ผลรวมของ ๑๒ เดือนที่สามจึงเริ่มแต่เดือนมีนาคม ปี ๒๕๐๖ จนถึงเดือนกุมภาพันธ์ ปี ๒๕๐๗ ทำเช่นนี้ต่อไปเรื่อย ๆ จนกระทั่งหมดข้อมูลในปี ๒๕๑๕ ซึ่งผลรวมของ ๑๒ เดือนสุดท้าย จะเป็นผลรวมของข้อมูลตั้งแต่เดือนมกราคมปี ๒๕๑๕ ถึงเดือนธันวาคมปี ๒๕๑๕

๒.๑.๑.๒ หาค่าเฉลี่ย ๑๒ เดือนโดยเอา ๑๒ ไปหารผลรวม ๑๒ เดือนแต่ละตัวที่ได้จาก ๒.๑.๑.๑

๒.๑.๑.๓ หา Centered 12 - month moving โดยเอาผลที่ได้จากข้อ ๒.๑.๑.๒ มาเฉลี่ยทีละ ๒ ตัวอีกทีหนึ่ง

๒.๑.๒ <u>การคำนวณหาดรรชนีการแปรผันเนื่องจากฤดูกาล</u>

เนื่องจากอนุกรมเวลาของข้อมูลแต่ละตัวที่เก็บมาได้ตามโมเดลในหัวข้อ ๒.๑ จะประกอบด้วยทางโน้มตามลำดับเวลา และการแปรผันเนื่องจากฤดูกาล และค่า Centered 12 month moving average ที่คำนวณได้ ตามหัวข้อ ๒.๑.๑.๓ คือ ค่าทางโน้มตามลำดับเวลา ดังนั้นเมื่อนำค่าทางโน้มตามลำดับเวลาไปหารข้อมูลที่เก็บมาได้ก็จะเหลือแต่การแปรผันเนื่องจากฤดูกาล ดูตามโมเดลจะได้

$$\frac{Y_{ij}}{T_i} = S_j \times \varepsilon_{ij}$$

$$i = 1, 2, 3, \ldots\ldots, 10$$

$$j = 1, 2, 3, \ldots\ldots, 12$$

อัตราส่วนที่ได้นี้เรียกว่า อัตราส่วนต่อทางโน้มตามลำดับเวลา (Ratio to trend) มักนิยมทำเป็นร้อยละ โดยเอา ๑๐๐ ไปคูณ

$$\frac{Y_{ij}}{T_i} \times 100 = 100 \times S_j \times \varepsilon_{ij} = S_{ij}$$

S_{ij} ที่ได้นี้จะเป็นค่าการแปรผันเนื่องจากฤดูกาลของเดือนที่ j ในปีที่ i ต้องนำมาเฉลี่ยอีกครั้ง เพื่อที่จะได้ค่าการแปรผันเนื่องจากฤดูกาลประจำเดือนที่ j (S_j) (Seasonal Index) เพียงค่าเดียว ในการเฉลี่ยก็เอาค่า S_{ij} ของทุก ๆ ปี มาหามัชฌิมเลขคณิต ถ้าดูตามโมเดลจะได้ดังนี้

$$\frac{\sum_{i=1}^{N} S_{ij}}{N} = 100 S_j \times \frac{\sum_{i=i}^{N} \varepsilon_{ij}}{N}$$

โดยที่ N = จำนวนปีที่มีข้อมูล

แต่เรากำหนดให้ ε_{ij} มีการกระจายปกติมีค่าเฉลี่ย ๑ มีความแปรปรวน σ^2

นั่นคือ $E(\varepsilon_i) = 1$ และ $E(\varepsilon_i - \bar{\varepsilon})^2 = \sigma^2$

ดังนั้น $E(\varepsilon_i) = \frac{\sum_{i=1}^{N} \varepsilon_{ij}}{N} = 1$

$$\therefore \frac{\sum_{i=1}^{N} S_{ij}}{N} = 100 S_j$$

และเนื่องจากเรากำหนดให้ $\sum_{j=1}^{12} S_j = 12$

ดังนั้น เราจึงจำเป็นต้องปรับปรุงค่าดรรชนีการแปรผันเนื่องจากฤดูกาล เพื่อที่จะให้ผลรวมเท่ากับ ๑๒ โดยเอา $\frac{๑๒๐๐}{\sum S_j}$ ไปคูณ S_j แต่ละตัว ที่ต้องทำเช่นนี้เพื่อให้แน่ใจได้ว่า ในฤดูกาลหนึ่งของผลการเปลี่ยนแปลงเนื่องจากฤดูกาลจริง ๆ เท่านั้น มิได้เนื่องมาจากการลดหรือเพิ่มของค่าระดับเฉลี่ยของปริมาณสินค้าแต่ละชนิด ดรรชนีการแปรผันเนื่องจากฤดูกาลของสินค้าทั้ง ๖ ชนิด แสดงไว้ในตาราง ก. ๑, ก.๒, ก.๓, ก.๔, ก.๕ ก.๖ ในภาคผนวก

^ จะได้มาจากการเปลี่ยนแปลงเนื่องจากฤดูกาล

๒.๑.๓ การปรับข้อมูลด้วยผลการแปรผันเนื่องจากฤดูกาล

เพื่อให้การพยากรณ์ปริมาณสินค้าเข้าและสินค้าออกในระยะเวลา ๑๐ ปีข้างหน้าใกล้เคียงกับความเป็นจริง และเนื่องจากเรากำหนดให้อนุกรมเวลาประกอบด้วย T x S ดังนั้น จึงจำเป็นต้องกำจัดการแปรผันเนื่องจากฤดูกาล หรือที่เรียกว่า deseasonalizing โดยการเอาค่าดรรชนีการแปรผันเนื่องจากฤดูกาลซึ่งหาได้จากหัวข้อ ๒.๑.๒ ของแต่ละเดือน (S_j) ไปหารข้อมูลในเดือนเดียวกันในแต่ละปีแล้วทำเป็นร้อยละ ข้อมูลที่ได้จะเป็นข้อมูลที่ได้

ขจัดการแปรผันเนื่องจากฤดูกาลจะเหลือแต่ข้อมูลที่มีความแปรปรวนเนื่องจากทางโน้มตามลำดับเวลา (trend) อย่างเดียว

๒.๑.๔ <u>การลากเส้นทางโน้ม (trend) โดยใช้ข้อมูลที่ขจัดการแปรผันเนื่องจากฤดูกาลแล้ว</u>

ข้อมูลที่บันทึกตามลำดับเวลา (อนุกรมเวลา) ถ้านำมาเขียนกราฟโดยให้แกน X เป็นเวลาแกน Y เป็นปริมาณจะได้จุดกระจัดกระจายทั่วไปบนกราฟ เรียกว่าแผนภาพกระจาย (scatter diagram)ก็จะมองเห็นทางโน้มจากแผนภาพกระจายนี้ได้ ฉะนั้นถ้าเราลากเส้นขึ้นเส้นหนึ่งแสดงทางโน้มประมาณเราก็อาจทำได้ แต่เส้นที่ลากขึ้นนี้อาจลากขึ้นได้หลายทาง ข้อสำคัญอยู่ที่ว่าเส้นไหนจะเป็นเส้นที่ดีที่สุด หลักในการลากเส้นทางโน้มก็คือ พยายามทำให้ส่วนเบี่ยงเบนเหนือเส้นทางโน้มและใต้เส้นทางโน้มจากข้อมูลเดิมเมื่อรวมแล้วเท่ากันหรือต่างกันน้อยที่สุด กล่าวคือ

$$\sum (Y - Y_c) = 0$$

ดังได้กล่าวแล้วว่า ทางโน้มตามลำดับเวลาเป็นข้อมูลที่บันทึกตามลำดับเวลาที่เกิดแสดงการเคลื่อนไหวขึ้นลงของข้อมูลและการวิเคราะห์ทางโน้มตามลำดับเวลาก็คือ การพยายามศึกษาและวัดความเปลี่ยนแปลงต่าง ๆ ของข้อมูลในช่วงระยะเวลาหนึ่ง โดยการลากเส้นทางโน้ม เพราะเส้นที่เขียนขึ้นนี้จะช่วยประมาณตัวเลขในอนาคตได้ การลากเส้นทางโน้มโดยทั่วไปทำได้ ๒ แบบคือ

๑. โดยการกะหรือประมาณด้วยตา

๒. โดยการคำนวณ

สำหรับการคำนวณนั้นมีหลายวิธี แต่วิธีที่นิยมกันมากที่สุด เพราะให้ความถูกต้องมากกว่าวิธีใด ๆ คือ วิธีกำลังสองน้อยที่สุด (Leest square) วิธีนี้ประมาณทางโน้มโดยยึดหลักความคลาดเคลื่อนน้อยที่สุด กล่าวคือ พยายามทำให้ผลรวมของผลต่างระหว่างค่าของข้อมูลที่ได้จากเส้นทางโน้มกับข้อมูลเดิมยกกำลังสองแล้วได้ค่าน้อยที่สุด ดังนั้นการหาทางโน้มโดยวิธีนี้จึงต้องตั้งต้นด้วยการสมมติสมการทางโน้มขึ้นในเทอมของตัวที่ยังไม่ทราบค่าแล้วดำเนินการหาค่าพวกนี้ โดยใช้หลักความคลาดเคลื่อนจะต้องน้อยที่สุดและโดยอาศัยวิชาแคลคูลัสอนุพันธ์ของความคลาดเคลื่อนจะต้องมีค่าเป็นศูนย์

ให้ Y = ค่าจริงของข้อมูล

Y_c = ค่าที่อ่านได้จากเส้นที่เขียนขึ้น

E = ความคลาดเคลื่อนทั้งหมด

$\therefore E = \sum (Y - Y_c)^2$

Y_c ซึ่งเป็นเส้นทางโน้มที่เขียนขึ้นนั้น จะเป็นเส้นตรงหรือเส้นโค้งขึ้นอยู่กับข้อมูลเดิม ถ้าเขียนกราฟของข้อมูลแล้ว แผนภาพกระจาย (scatter diagram) ที่ได้อยู่ในแนวเส้นตรง เราก็จะใช้สมการเส้นตรงในการลากเส้นทางโน้ม โดยที่เส้นตรงที่ได้จะมีสมการเป็น $Y_c = a + bX$ แต่ถ้าแผนภาพกระจายที่ได้ไม่อยู่ในแนวเส้นตรง แสดงว่า จะลากเส้นทางโน้มของข้อมูลชุดนี้เป็นเส้นตรงไม่ได้ เพราะผลจะผิดพลาดมากต้องใช้เส้นโค้งแทน โดยพิจารณาถึงการใช้เส้นโค้งตามกำลัง (degree) ต่าง ๆ แล้วแต่ลักษณะของข้อมูล คือ ถ้ามี curve อันเดียวใช้กำลังสอง (second degree curve) ถ้ามี curve ๒ อันใช้สมการกำลังสาม (third degree curve) เป็นต้น ถ้าเป็นสมการกำลังสองจะได้สมการเป็น $Y_c = a + bX + cX^2$ โดยที่ทั้งสองสมการนี้เราให้

X เป็นตัวแปรอิสระ (independent variable)

Y_c เป็นตัวแปรไม่อิสระ (dependent variable)

a, b, c เป็นตัวคงที่ไม่ทราบค่า (unknown constants) ซึ่งเราจะต้องคำนวนหา

จากการทดสอบโดยคำนวณหาผลต่างครั้งที่ ๑ และที่ ๒ ของค่า trend ณ X ตำแหน่งต่าง ๆ เพื่อดูว่าเส้นที่ใช้ปรับนั้นควรเป็นชนิดใดแน่ปรากฏว่า มีข้าวโพดเพียงอย่างเดียวที่มีแนวโน้มเป็นเส้นตรงอีก ๕ ชนิดคือ ข้าว ยางพารา ผลิตภัณฑ์เหล็กและเหล็กกล้า น้ำมันเชื้อเพลิงและกระดาษมีแนวโน้มเป็นเส้นโค้ง คือเป็น second degree curve ดังนั้นจึงเขียนกราฟข้อมูลของข้าวโพดด้วยเส้นตรงส่วนกราฟข้อมูลของสินค้าอีก ๕ ชนิด เขียนด้วยเส้นโค้งพาราโบลาและได้ค่า unknown constants และสมการของเส้นที่ใช้เขียนข้อมูลของสินค้าแต่ละชนิดดังนี้

ข้าว $a = ๑๗๖,๕๘๐.๖๒$

$b = -๒,๕๖๑.๗๗$

c = ๒๐.๘๒

Y_c = ๑๗๖,๕๘๐.๖๒ − ๒,๕๖๑.๗๗ X + ๒๐.๘๒ $X^{๒}$

ยางพารา a = ๑,๕๖๕,๖๒๓.๕๑

b = − ๖,๗๘๔.๒๓

c = ๑๓๗.๒๑

Y_c = ๑,๕๖๕,๖๒๓.๕๑ − ๖,๗๘๔.๒๓ X + ๑๓๗.๒๑ $X^{๒}$

ข้าวโพด a = ๕๖,๒๖๘.๔๕

b = ๔๐๗.๒๒

Y_c = ๕๖,๒๖๘.๔๕ + ๔๐๗.๒๒ X

ผลิตภัณฑ์เหล็กและเหล็กกล้า

a = ๒๓,๓๑๒,๔๒๐.๔๑

b = ๓๘๔,๐๑๕.๐๗

c = ๖๐๓.๗๔

Y_c = ๒๓,๓๑๒,๔๒๐.๔๑ + ๓๘๔.๐๑๕.๐๗ X + ๖๐๓.๗๔ $X^{๒}$

น้ำมันเชื้อเพลิง

a = ๑,๔๐๓,๒๗๓.๒๔

b = ๔,๓๘๖.๗๑

c = ๒๓๔.๗๓

Y_c = ๑,๔๐๓,๒๗๓.๒๔ + ๔,๓๘๖.๗๑ X + ๒๓๔.๗๓ $X^{๒}$

กระดาษ a = ๖,๐๒๕,๖๓๗.๐๗

b = ๕๓,๗๔๘.๕๕

c = ๒๔๘.๘๑

Y_c = ๖,๐๒๕.๐๗ + ๕๓,๗๔๘.๕๕X + ๒๔๘.๘๑ $X^{๒}$

๒.๑.๕ <u>การประมาณเชิงสถิติโดยใช้ทางโน้มตามลำดับเวลา (trend)</u>

ในการประมาณเชิงสถิติโดยใช้ทางโน้มตามลำดับเวลา ทำโดยการแทนค่า X ด้วย ๑๒๑, ๑๒๒, ๑๒๓, ... ๒๔๐ เข้าไปในสมการทางโน้มตามลำดับเวลาซึ่งคำนวณ

ได้จากหัวข้อ ๒.๑.๔ ก็จะได้ค่า Y_c ของเดือนต่าง ๆ ตั้งแต่มกราคม ๒๕๑๖ – ธันวาคม ๒๕๒๕ แต่ค่าที่คำนวณได้นี้เป็นค่าประมาณของปริมาณสินค้าที่ยังมิได้รวมผลกระทบกระเทือนจากการแปรผันเนื่องจากฤดูกาลด้วย จึงต้องนำเอาดรรชนีการแปรผันเนื่องจากฤดูกาลประจำเดือนที่ j (S_j) คูณเข้าไปในปริมาณของเดือนที่ j เราจะได้ปริมาณสินค้าเข้าและสินค้าออกแต่ละชนิดที่คาดว่าจะส่งออกซึ่งได้รับผลกระทบกระเทือนจากทางโน้มตามลำดับเวลา และการแปรผันเนื่องจากฤดูกาลรวมอยู่ด้วย (แสดงไว้ในตาราง ข.๑, ข.๒, ข.๓, ข.๔, ข.๕, ข.๖) ในภาคผนวก)

๒.๒ วิธีการประมาณเชิงสถิติโดยใช้วิธี Exponential Weighted Moving Average

การประมาณด้วยวิธีนี้เป็นการประมาณที่ค่าที่ประมาณได้แต่ละคาบเวลาขึ้นอยู่กับข้อมูลเดิมและการประมาณเมื่อคาบเวลาที่แล้วซึ่งมีวิธีการดังนี้

ถ้าให้ $S_{t,T}$ = ปริมาณที่ประมาณได้ล่วงหน้า T เดือนในอนาคต

S_t = ปริมาณที่แท้จริงที่เดือนที่ t

$\tilde{S}_t$ = ปริมาณที่คาดว่าจะเป็นที่ได้ปรับด้วยความแปรปรวนเนื่องจากฤดูกาลแล้วที่เดือนที่ t

F_t = ค่าประมาณของการเปลี่ยนแปลงเนื่องจากฤดูกาลที่เดือนที่ t

R_t = ค่าประมาณของปริมาณทางโน้มตามลำดับเวลาที่เดือนที่ t

A, B, C = ค่าคงที่ซึ่งมีค่าอยู่ระหว่าง ๐ และ ๑ ซึ่งเราจะต้องเลือกให้เหมาะสมเพื่อใช้เป็นตัวถ่วงน้ำหนัก

การพยากรณ์ล่วงหน้า T เดือน จะได้จากสมการ

$$S_{t,T} = \left[\tilde{S}_t + TR_t\right] F_{t-L+T} \quad \text{.....} \quad (1)$$

$$T = 1, 2, \ldots, L$$

$$\tilde{S}_t = A\frac{S_t}{F_{t-L}} + (1-A)(\tilde{S}_{t-1} + R_{t-1}) \quad \text{.....} \quad (2)$$

$$0 \leq A \leq 1$$

$$F_t = B\frac{S_t}{\tilde{S}_t} + (1-B)F_{t-L} \quad \text{.....} \quad (3)$$

$$0 \leq B \leq 1$$

$$R_t = C(\tilde{S}_t - \tilde{S}_{t-1}) + (1 - C)R_{t-1} \quad \ldots\ldots\ldots\ldots (4)$$

$$0 \leq C \leq 1$$

t = ๑, ๒, ..., ๖๐

T = ๑, ๒, ..., ๑๒

การพยากรณ์ที่ $T > ๑๒$ ทำได้โดยการเลือกค่า F ที่เหมาะสม

แต่การพยากรณ์โดยวิธี Exponential Weighted Moving Average จำเป็นต้องใช้ค่าเริ่มต้น กล่าวคือ เมื่อจะเริ่มการพยากรณ์จะต้องมีค่า $\tilde{S}$ แรกเริ่ม R แรกเริ่ม F แรกเริ่มของเดือนต่าง ๆ ทั้ง ๑๒ เดือน และค่า A, B, C ที่เหมาะสม (ค่า A, B, C เป็นค่าที่อยู่ระหว่าง ๐ และ ๑ ซึ่งเราไม่ทราบว่าค่าใดจะเหมาะสมที่สุดสำหรับข้อมูลแต่ละอย่างจำเป็นต้องทดลองหาเพื่อเลือกค่าที่เหมาะสม) ในการหาค่าเริ่มต้นและเลือกค่า A, B, C ที่เหมาะสมทำดังนี้คือ

๒.๒.๑ เฉลี่ยปริมาณสินค้าต่อเดือน (แต่ละชนิด) ในแต่ละปี สมมติให้เป็นค่า V_i

๒.๒.๒ ค่าประมาณแรกเริ่ม ($\tilde{S}$ แรกเริ่ม) = V_1 ปริมาณสินค้าเฉลี่ยต่อเดือนในปีที่ ๑ คือ ปี ๒๕๑๑

๒.๒.๓ ค่าประมาณแรกเริ่มของ R (R แรกเริ่ม) = $(V_{(H \div L)} - V_1)/H-L$

จากการทดลองใช้ $H = ๖๐$ และ $H = ๔๘$ พบว่าใช้ $H = ๔๘$ ให้ค่าผิดพลาดน้อยกว่า ดังนั้น ในที่นี้จึงใช้ $H = ๔๘$ และ $L = ๑๒$

∴ ค่าประมาณแรกของ R ที่จะทำวิจัยในเรื่องนี้ = $(V_4 - V_1)/๓๖$ (คือ (ค่าเฉลี่ยของปี ๒๕๑๔ – ค่าเฉลี่ยของปี ๒๕๑๑) หารด้วย ๓๖)

๒.๒.๔ ค่าแปรผันเนื่องจากฤดูกาลสำหรับเดือนที่ t

$$F_t = \frac{S_t}{V_i - \left(\frac{L+1}{2} - j\right) \times R_{\text{แรกเริ่ม}}}$$

โดยที่ t = ๑, ๒, ..., H

S_t = ปริมาณที่แท้จริงในเดือนที่ t

j = ตำแหน่งของเดือนในปีนั้น เช่นเดือนมกราคม $j = 1$, ...

ธันวาคม $j = ๑๒$

๒.๒.๕ การแปรผันเนื่องจากฤดูกาลที่คำนวณได้ในแต่ละเดือนของแต่ละปี ต้องนำมาเฉลี่ยเพื่อที่จะได้ดรรชนีการแปรผันเนื่องจากฤดูกาลในแต่ละเดือน เช่นการแปรผันเนื่องจากฤดูกาลในเดือนมกราคมของแต่ละปีนำมาเฉลี่ยเพื่อที่จะได้ดรรชนีการแปรผันเนื่องจากฤดูกาลของเดือนมกราคม

๒.๒.๖ เพื่อให้มั่นใจว่าในวัฏระหนึ่งของการแปรผันเนื่องจากฤดูกาลจะได้มาจากการแปรผันเนื่องจากฤดูกาลจริง ๆ เท่านั้น ไม่ใช่เนื่องมาจากการลดหรือเพิ่มขึ้นของระดับค่าเฉลี่ยของปริมาณสินค้า กล่าวคือ $\sum_{j=1}^{12} F_j = ๑๒$ จึงต้องปรับปรุงค่าการแปรผันเนื่องจากฤดูกาล

$$F_j = \text{ค่าเฉลี่ย } F_j \left(\frac{L}{\sum_{j=1}^{L} \text{ค่าเฉลี่ย} F_j} \right) \qquad ,j=1...,12$$

F_j แต่ละตัวที่ได้คือค่าแรกเริ่ม

จากวิธีการตั้งแต่ ๒.๒.๑ – ๒.๒.๖ เราจะได้ค่า $\tilde{S}_{\text{แรกเริ่ม}}$, $R_{\text{แรกเริ่ม}}$ และ $F_{\text{แรกเริ่ม}}$ ที่เดือนต่าง ๆ ทั้ง ๑๒ เดือน ที่จะใช้เริ่มต้นในการประมาณที่คาบเวลาที่ ๑ แล้วก็ใช้สมการ (๑) – (๔) ในการประมาณโดยใส่ค่า A, B, C ที่อยู่ระหว่าง ๐ ถึง ๑ เข้าไปในการประมาณตั้งแต่เดือนที่ ๑ – เดือนที่ ๖๐ เพื่อเลือกค่า A, B, C ที่เหมาะสมโดยในทางปฏิบัติจะทำตามลำดับขั้นดังนี้

ขั้นที่ ๑ ใช้สมการ (๒) คำนวณค่า S_t ด้วยการใช้ $\tilde{S}_{\text{แรกเริ่ม}}$ $R_{\text{แรกเริ่ม}}$ และค่า F_{t-L} ซึ่งได้ทำการคำนวณไว้แล้ว

ขั้นที่ ๒ ใช้สมการ (๓) คำนวณค่า F_t ซึ่งจะใช้เป็น F_{t-L} ในการคำนวณเดือนต่อไป

ขั้นที่ ๓ ใช้สมการ (๔) คำนวณค่า R_t ซึ่งจะใช้เป็น R_{t-1} ในการคำนวณเดือนต่อไป

ขั้นที่ ๔ ใช้สมการ (๑) ประมาณปริมาณล่วงหน้า

ขั้นที่ ๕ ใช้ $\tilde{S}_t$ ที่ประมาณได้เป็น $\tilde{S}_{t-1}$ เพื่อประมาณปริมาณในเดือนต่อไป

เมื่อได้ค่าประมาณของเดือนที่ ๔๙ ถึงเดือนที่ ๖๐ แล้วก็นำมาเปรียบเทียบกับ

ปริมาณจริงที่มีอยู่คือ A, B, C เท่าใดที่ให้ค่าประมาณใกล้เคียงค่าจริงมากที่สุด โดยพิจารณาจากความคลาดเคลื่อนมาตรฐานของการประมาณ

$$\frac{\sum_{i=1}^{N} (d_i - \bar{d})^2}{N - 1}$$

โดยที่ d_i = ผลต่างระหว่างค่าจริงกับค่าที่ประมาณได้แต่ละกาบเวลา

จากผลของการทดลองหาค่า A, B, C ต่าง ๆ ของปริมาณสินค้าต่าง ๆ ทั้ง ๖ ชนิด ปรากฏผลดังนี้

ข้าว	A = ๐.๘๔,	B = ๐.๔๒,	C = ๐.๐๐
ข้าวโพด	A = ๐.๖๓,	B = ๐.๐๐,	C = ๐.๐๐
ยางพารา	A = ๐.๐๒,	B = ๐.๐๐	C = ๐.๒๗
ผลิตภัณฑ์น้ำมันเชื้อเพลิง	A = ๐.๐๐,	B = ๐.๐๔	C = ๐.๐๕
ผลิตภัณฑ์เหล็กและเหล็กกล้า	A = ๐.๐๒,	B = ๐.๒๔	C = ๐.๔๕
ผลิตภัณฑ์กระดาษ	A = ๐.๐๒,	B = ๐.๒๓	C = ๐.๔๓

เมื่อได้ค่า A, B, C ที่เหมาะสมแล้ว ก็ใช้ค่านั้นในการประมาณเดือนที่ ๖๑ – ๑๘๐ โดยใช้สมการ (๑) – (๔) ดูรายละเอียดการคำนวณจากโปรแกรม (ในการศึกษาครั้งนี้ผู้เขียนทำการคำนวณโดยใช้เครื่องจักรคำนวณ CDC ๓๖๐๐ ของสถาบัน AIT ภาษาที่ใช้คือ FORTAN IV)

```
      PROGRAM FORECAST
      DIMENSION S(200), V(200), FAVER(12), STDV(20,20,20),
     $ F(200) , ITITLE(10), WAA(20), WBB(20), WCC(20)
      DIMENSION IP(100)
    1 READ 04,  ITITLE
    4 FORMAT (10A8)
      READ 05, ND, NH, NL
    5 FORMAT (3I10)
      IF (ND. EQ. 0)                                    STOP
      READ 05, NTABLE, NFORCAST
      READ 06, WA1, WA2, WA3
      READ 06, WB1, WB2, WB3
      READ 06, WC1, WC2, WC3
    6 FORMAT (3F10. 2)
      READ 07, (S(K), K = 1, ND)
    7 FORMAT (5F15. 2)

C     FIND AVERAGE SALES PER PERIOD FOR EACH YEAR, V(I)
      NT = 0
      I  = 1
   10 V(I) = 0.
   11 NT = NT + 1
      V(I) = V(I) + S(NT)
      IF (NT, NE, I * NL) GO TO                         11
      V(I) = V(I)/NL
      IF (NT, EQ, NH) GO TO                             12
      I = I + 1
      GO TO                                             10
```

```
C       FIND THE PREVIOUS ESTIMATE OF R (TREND) AND S (SALES)
   12 RLAST =(V(I) - V(1))/(NH - NL)
      SLAST = V(1)

C       FIND SEASONAL FACTORS
      CONST = (NL + 1,)/2,
      NT = 0
      I  = 1
   16 J = 1
   17 NT = NT + 1
      F(NT) = S(NT)/(V(I) - (CONST - J)  * RLAST)
      IF (NT. EQ. I * NL)   GO TO                    18
      J = J + 1
      GO TO                                          17
   18 IF (NT. EQ. NH) GO TO                          19
      I = I + 1
      GO TO                                          16

C **    FIND AVERAGE SEASONAL FACTORS
   19 M = NH/NL
      ESUM = 0;
      J = 1
   20 FAVER(J) = F(J)
```

```
      K = 1
   21 JKL = J + K * NL
      FAVER (J) = FAVER(J) + F(JKL)
      IF (K. EQ. M - 1) GO TO                                    22
      K = K + 1
      GO TO                                                      21
   22 FAVER(J) = FAVER(J)/M
      FSUM = FSUM + FAVER(J)
      IF (J.EQ.NL)  GO TO                                        23
      J = J + 1
      GO TO                                                      20
C     FIND NORMALIZED AVERAGE FACTORS
   23 U = NL/FSUM
      DO 24 J=1,NL
   24 FAVER(J) = U*FAVER(J)
      SSAVE = SLAST
      RSAVE = RLAST
C     GENERATE WEIGHTS
      NWA = (WA2 - WA1)/WA3 + 1
      NWB = (WB2 - WB1)/WB3 + 1
      NWC = (WC2 - WC1)/WC3 + 1
      DO 40 IA=1, NWA
      WA = (IA-1) * WA3 + WA1
      WAA(IA) = WA
      A = 1. - WA
      DO 40 IB=1, NWB
      WB = (IB-1) * WB3 + WB1
      WBB(IB) = WB
      B = 1  - WB
      DO 40 IC=1, NWC
      WC = (IC-1) * WC3 + WC1
      WCC(IC) = WC
```

```
      C = 1. - WC
      DO 30 J=1, NL
   30 F(J) = FAVER(J)
      SLAST = SSAVE
      RLAST = RSAVE
      STD = 0.
      NT = 0
   31 J = 1
   32 NT = NT + 1
      SCUR = WA * S(NT)/F(J) + A * (SLAST + RLAST)
      RCUR = WC * (SCUR - SLAST) + C*RLAST
      F(J) = WB * S(NT)/SCUR + B*F(J)
      JL = J + NL
      F(JL) = F(J)
      SLAST = SCUR
      RLAST = RCUR
      IF (NT. GT. NH)  GO TO                    35
      IF (J. EQ. NL)  GO TO                     34
   33 J = J + 1
      GO TO                                     32
   34 IF (NT. NE. NH)  GO TO                    31
C     FIND FORCASTS FOR FUTURE PERIODS
   35 ST1 = (SCUR+RCUR) * F(J+1)
C     FIND STANDARD DIVIATION
      ERROR = S(NT+1) - ST1
      STD = STD + ERROR*ERROR
      IF (NT.EQ.ND - 1)  GO TO                  36
      IF (J.EQ.NL)  GO TO                       31
      GO TO                                     33
   36 STD = SQRT(STD/(ND-NH-1,))
      STDV(IA,IB,IC) = STD
   40 CONTINUE
      PRINT 50,  NTABLE
   50 FORMAT (//1H1,5X,*TABLE*, I2,*:    STANDARD DIVIATIONS OF*,
     $       * FORECAST ERRORS*)
```

หอสมุดกลาง จุฬาลงกรณ์มหาวิทยาลัย

```
      PRINT 51, ITITLE
   51 FORMAT (20X,10A8)
      DO 56 IA=1,NWA
   53 PRINT 54,  WAA(IA), (WCC(IC), IC=1,NWC)
   54 FORMAT (//6X,*WA =*F5.2,34X,*WC*/14X,*WB*,2X,10F10,2/)
      PRINT 55
      DO 56 IB=1,NWB
      PRINT 55,  WBB(IB), (STDV(IA,IB,IC), IC=1,NWC)
   55 FORMAT (8X,F10.2,10F10)
   56 CONTINUE
      SMEAN = S(1)
      SCALE = S(1)
      DO 57 I=2,ND
      SMEAN = SMEAN + S(I)
   57 IF (SCALE. LT, S(I))  SCALE = S(I)
      SMEAN = SMEAN/ND
      INT = SCALE
      DO157 I=1,20
      INT = INT/10
  157 IF (INT.LT.10)   GO TO                    158
  158 SCALE = INT + 1
      SCALE = SCALE * 10,** I/100.
      DO 59 I=1,100
   59 IP(I) = 1H
      STD = 999999999.
      DO 60 IA = 1, NWA
      DO 60 IB = 1, NWB
      DO 60 IC = 1, NWC
      IF (STD.LE.STDV(IA,IB,IC))  GO TO          60
      STD = STDV(IA,IB,IC)
      WA = WAA(IA)
      WB = WBB(IB)
      WC = WCC(IC)
   60 CONTINUE
      PRINT 58,  WA,  WB,  WC,  SCALE
```

```
58 FORMAT (1H1,5X, *WA =  *F5.2,5X,  *WB =  *F5.2,5X*WC = *F5.2/
  $       *   T    ACTUAL   FORCAST  %  ERROR * , 5X, *SCALE =1

  $   *, F8  / 9X,1H+,7X,1H *  )
    A = 1. - WA
    B = 1. - WB
    C = 1. - WC
    DO 61 J=1, NL
61  F(J) = FAVER(J)
    SLAST = SSAVE
    RLAST = RSAVE
    NT = 0
62  J = 1
63  NT = NT + 1
    SCUR = WA * S(NT)/F(J) + A  *(SLAST + RLAST)
    RCUR = WC * (SCUR - SLAST) + C *RLAST
    F(J) = WB  *S(NT)/SCUR + B *F(J)
    JL = J + NL
    F(JL) = F(J)
    SLAST = SCUR
    RLAST = RCUR
    ST1 = (SCUR + RCUR)  *F(J+1)
    ERROR = S(NT + 1) - ST1
    IA = ST1/SCALE + 1.5
    IB = S(NT+1)/SCALE + 1,5
    IF (IA.GT.100)  IA = 100
    IF (IA.GT.0)  IP(IA) = 1H *
    IP(IB) = 1H+
    II = NT + 1
    ERROR = ERROR/SMEAN  *100
    PRINT 64,   II,   S(II),   ST1,-ERROR, IP
64  FORMAT (I3,F9,F10,F8.2,3X,1HI,100A1)
    IF (IA.GT.0)   IP(IA)  =  1H
    IP(IB) = 1H
    IF (NT.EQ.ND - 1)  GO TO                    70
    IF (J.EQ.NL)  GO TO                         62
```

```
      J = J + 1
      GO TO                                              63
   70 IF (NFORCAST.LE.ND)  GO TO                         1
      J = J + 1
      NT = ND
      SCUR = WA  *S(NT)/F(J) + A  * (SLAST + RLAST)
      F(J) = WB  * S(NT)/SCUR + B  * F(J)
      RCUR = WC  * (SCUR - SLAST) + C  * RLAST
   71 JJ = NT - ND + 1
      J = JJ - (JJ-1)/NL * NL
      S(NT+1) = (SCUR+JJ  * RCUR)  * F(J)
      IF (NT.EQ.NFORCAST-1)  GO TO                       72
      NT = NT + 1
      GO TO                                              71
   72 ND1 = ND + 1
      PRINT 172
  172 FORMAT (1H1)
      PRINT 73,  (NT, S(NT), NT=ND1,NFORCAST)
   73 FORMAT (I8, F13.2,I8,F13.2,I8,F13.2,I8,F13.2,I8,F13.2)
      GO TO                                              1
      END
```