บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเพื่อเปรียบเทียบความคลาดเคลื่อนของการพยากรณ์ด้วยวิธีของเบย์ (Bayesian method) วิธีบ็อกซ์-เจนกิ้นส์ (Box-Jenkins: B-J) และวิธีการพยากรณ์ร่วม (combined forecasting) โดยพิจารณาแยกวิธีการพยากรณ์ร่วม (combined forecasting) ออกเป็น 2 วิธี ได้แก่ วิธีที่ 1 ได้จากการนำค่าพยากรณ์ด้วยวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลมารวมกันโดยให้น้ำหนักตามแบบของ Newbold และ Granger แบบที่ 1 และวิธีที่ 2 ได้จากการนำค่าพยากรณ์ด้วยวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลมารวมกันโดยให้น้ำหนักตามแบบของ Newbold และ Granger แบบที่ 2 โดยการเลือกวิธีพยากรณ์เดี่ยวที่เหมาะสมกับข้อมูลพิจารณาจาก 3 กรณีตามลักษณะของข้อมูลอนุกรมเวลา คือ กรณีแรก เมื่อข้อมูลไม่มีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) กรณีที่สอง เมื่อข้อมูลมีอิทธิพลของแนวโน้มเพียงอย่างเดียวจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสำหรับข้อมูลที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential smoothing for exponential trend: ESE) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) และกรณีที่สาม เมื่อข้อมูลมีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบโฮลท์-วินเทอร์ (Holt-Winters smoothing: HWS) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) โดยใช้ค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าความคลาดเคลื่อนเฉลี่ย (ME) เป็นเกณฑ์ในการเปรียบเทียบ มีวิธีดำเนินการวิจัย ดังนี้

ลักษณะและแหล่งที่มาของข้อมูล

ฐานข้อมูลสำหรับการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิ 2 ฐาน ฐานแรก คือ ปริมาณการยืมสิ่งพิมพ์ซึ่งแยกเป็นหนังสือภาษาไทย หนังสือภาษาอังกฤษ และวิทยานิพนธ์ ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ตั้งแต่เดือนมิถุนายนปีการศึกษา 25๖, ถึง เดือนตุลาคม ปีการศึกษา 2542 ในการวิจัยครั้งนี้ผู้วิจัยได้เก็บรวบรวมข้อมูลเพิ่มเติม โดยเริ่มตั้งแต่เดือนพฤศจิกายน ปีการศึกษา 2542 ถึงเดือนกันยายน ปีการศึกษา 2544 รวมเป็นอนุกรมเวลา 88 เดือน และข้อมูลฐานที่สอง คือ จำนวนนักเรียนในระดับประถมศึกษา และระดับมัธยมศึกษา ระหว่าง ปี 2480-2539 ในการวิจัยครั้งนี้ผู้วิจัยได้เก็บรวบรวมข้อมูลเพิ่มเติม โดยเริ่มตั้งแต่ปี 2540-2543 รวมเป็นอนุกรมเวลา 64 ปี

แหล่งที่มาของข้อมูลสำหรับการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิจาก 2 แหล่ง คือ ในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยใช้ข้อมูลจากฐานข้อมูลของทัศนีย์ อินทนู (2543) คือ ปริมาณการยืมสิ่งพิมพ์ ซึ่งแยกเป็นหนังสือภาษาไทย หนังสือภาษาอังกฤษ และวิทยานิพนธ์ ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งมีลักษณะเป็นข้อมูลรายเดือน และผู้วิจัยได้เก็บรวบรวมข้อมูลอนุกรมเวลาชุดนี้เพิ่มเติมเพื่อให้มีความเป็นปัจจุบันมากขึ้น ส่วนข้อมูลทางการศึกษาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยใช้ข้อมูลจากฐานข้อมูลของบำเพ็ญ ปิดชิด (2540) ได้แก่ จำนวนนักเรียนในระดับประถมศึกษา และระดับมัธยมศึกษา มีลักษณะข้อมูลเป็นรายปี และผู้วิจัยได้เก็บรวบรวมข้อมูลอนุกรมเวลาชุดนี้เพิ่มเติมเพื่อให้มีความเป็นปัจจุบันมากขึ้น

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นแบบบันทึกข้อมูลของทัศนีย์ อินทนู (2543) และบำเพ็ญ ปิดชิด (2540) สร้างขึ้น ดังตาราง 13-14

ตาราง 13 แบบบันทึกข้อมูลปริมาณการสิ่งพิมพ์ประเภท.....(เล่ม)

ปีการศึกษา / เดือน	2537	2538	2539	2540	2541	2542	2543
มิถุนายน							
.							
.							
.							
พฤษภาคม							

ตาราง 14 แบบบันทึกข้อมูลจำนวนนักเรียน (คน)

ปีการศึกษา	จำนวนนักเรียนระดับประถมศึกษา	จำนวนนักเรียนระดับมัธยมศึกษา
2480		
.		
.		
.		
2543		

การเก็บรวบรวมข้อมูล

การวิจัยครั้งนี้ ผู้วิจัยบันทึกข้อมูลด้วยตนเองโดยมีขั้นตอนการเก็บรวบรวมข้อมูลดังนี้

1. ขั้นการติดต่อขอความร่วมมือ ผู้วิจัยติดต่อขอความร่วมมือไปยังเจ้าหน้าที่บรรณารักษ์ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัยเพื่อบันทึกข้อมูลการยืมปริมาณการยืมสิ่งพิมพ์ และติดต่อไปยังเจ้าหน้าที่บรรณารักษ์ของห้องสมุดสำนักงานสถิติแห่งชาติ เพื่อขอข้อมูลจำนวนนักเรียนระดับประถมศึกษาและระดับมัธยมศึกษา

2. สำรวจและบันทึกข้อมูล

3. ขั้นบรรณาธิกรณ์ข้อมูลเพื่อตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูลที่ได้เก็บรวบรวมมาทั้งหมด

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล ผู้วิจัยวิเคราะห์ข้อมูลโดยใช้โปรแกรม SPSS for Windows version 10.05, SAS version 6.12 และ Microsoft office 97 (Excel) วิเคราะห์ข้อมูลแบ่งออกเป็น 6 ตอน ดังนี้

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐาน ค่าความเบ้ ค่าความโด่ง และค่าสัมประสิทธิ์การกระจายของข้อมูล

ตอนที่ 2 การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และ/หรือการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยแบ่งการตรวจสอบออกเป็น 2 ตอน คือ ตอนแรก เป็นการตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และ/หรือการเปลี่ยนแปลงเนื่องจากฤดูกาล สำหรับข้อมูลรายเดือน และตอนที่สอง เป็นการตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้มเพียงอย่างเดียว สำหรับข้อมูลรายปี มีรายละเอียดดังต่อไปนี้

ตอนที่ 2.1 การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และ/หรือการเปลี่ยนแปลงเนื่องจากฤดูกาล สำหรับข้อมูลอนุกรมเวลารายเดือน โดยการนำข้อมูลอนุกรมเวลามาพล็อตกราฟ เพื่อดูลักษณะการเคลื่อนไหวของอนุกรมเวลา และตรวจสอบโดยใช้การวิเคราะห์การถดถอย เพื่อตรวจสอบว่าข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้ม และ/หรือการเปลี่ยนแปลงเนื่องจากฤดูกาลจริงหรือไม่ โดยตรวจสอบทั้งรูปแบบบวกและรูปแบบคูณ

การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และ/หรือการเปลี่ยนแปลงเนื่องจากฤดูกาลโดยใช้การวิเคราะห์การถดถอย ต้องตรวจสอบตามขั้นตอนดังนี้ คือ

1. ตรวจสอบว่าข้อมูลอนุกรมเวลามีแนวโน้มเส้นตรงหรือไม่ จากโมเดลแบบบวก (Additive model) $y_t = \beta_0 + \beta_1 t + \beta_2 x_{1t} + \beta_3 x_{2t} + \ldots + \beta_{12} x_{11t} + \varepsilon_t$ ใช้สถิติทดสอบ partial t

$$t = \frac{b_1}{s_{b_1}}$$

เพื่อทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นตรง

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นตรง

หรือทดสอบ $H_0 : \beta_1 = 0$

$H_a : \beta_1 \neq 0$

จากโมเดลแบบคูณ (Mutiplicative model)

$$y_t = \beta_0\beta_1^t\beta_2^{x1t}\beta_3^{x2t}\ldots\beta_{12}^{x11t}\varepsilon_t$$

ซึ่งแปลงให้เป็นโมเดลแบบบวก โดยการหาลอการิทึม ดังนี้

$$\ln y_t = (\ln\beta_0) + (\ln\beta_1)t + (\ln\beta_2)x_{1t} + \ldots + (\ln\beta_{12})x_{11t} + (\ln\varepsilon_t)$$

$$y'_t = \beta'_0 + \beta'_1 t + \beta'_2 x_{1t} + \beta'_3 x_{2t} + \ldots + \beta'_{12} x_{11t} + \varepsilon'_t$$

ทดสอบ $H_0 : \beta'_1 = 0$

$H_a : \beta'_1 \neq 0$

หรือตรวจสอบว่าข้อมูลอนุกรมเวลามีแนวโน้มเส้นโค้งหรือไม่ จากโมเดลแบบบวก (Additive model) $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 x_{1t} + \beta_4 x_{2t} + \ldots + \beta_{13} x_{11t} + \varepsilon_t$ ใช้สถิติทดสอบ partial t

$$t = \frac{b_2}{s_{b2}}$$

เพื่อทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นโค้ง

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นโค้ง

ทดสอบ $H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

2. หากพบว่ามีแนวโน้มเส้นตรงจากผลการตรวจสอบสมมติฐานในข้อที่ 1 จึงตรวจสอบว่ามีฤดูกาลด้วยหรือไม่ โดยการทดสอบแบบ partial F จากรูปแบบแนวโน้มและฤดูกาล มีรายละเอียดดังนี้

2.1 เมื่อพบว่าข้อมูลมีการเปลี่ยนแปลงเนื่องจากแนวโน้มเส้นตรง จึงตรวจสอบต่อไปว่าข้อมูลอนุกรมเวลามีการมีการเปลี่ยนแปลงเนื่องจากฤดูกาลหรือไม่ จากโมเดลแบบบวก (Additive model) $y_t = \beta_0 + \beta_1 t + \beta_2 x_{1t} + \beta_3 x_{2t} + \ldots + \beta_{12} x_{11t} + \varepsilon_t$ ใช้สถิติทดสอบ partial F

$$F = \frac{\left|SSR(t, x_1,\ldots,x_{11}) - SSR(t)\right|/(11)}{SSE(t, x_1,\ldots,x_{11})/(n-13)}$$

เพื่อทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงเนื่องจากฤดูกาล

หรือทดสอบ $H_0 : \beta_2 = \beta_3 = \ldots = \beta_{12} = 0$

$H_a : \beta_2, \beta_3, \ldots, \beta_{12}$ มีค่าไม่เป็น 0 อย่างน้อยหนึ่งค่า

จากโมเดลแบบคูณ (Mutiplicative model)

$$y_t = \beta_0\beta_1^t\beta_2^{x1t}\beta_3^{x2t}\ldots\beta_{12}^{x11t} + \varepsilon_t$$

ซึ่งแปลงให้เป็นโมเดลแบบบวก โดยการหาลอการิทึม ดังนี้

$$\ln y_t = (\ln\beta_0) + (\ln\beta_1)t + (\ln\beta_2)x_{1t} + \ldots + (\ln\beta_{12})x_{11t} + (\ln\varepsilon_t)$$

$$y'_t = \beta'_0 + \beta'_1 t + \beta'_2 x_{1t} + \beta'_3 x_{2t} + \ldots + \beta'_{12} x_{11t} + \varepsilon'_t$$

ทดสอบ $H_0 : \beta'_2 = \beta'_3 = \ldots = \beta'_{12} = 0$

$H_a : \beta'_2, \beta'_3, \ldots, \beta'_{12}$ มีค่าไม่เป็น 0 อย่างน้อย หนึ่งค่า

2.2 เมื่อพบว่าข้อมูลมีการเปลี่ยนแปลงเนื่องจากแนวโน้มเส้นโค้ง จึงตรวจสอบต่อไปว่าข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงเนื่องจากฤดูกาลหรือไม่ จากโมเดลแบบบวก (Additive model) $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 x_{1t} + \beta_4 x_{2t} + \ldots + \beta_{13} x_{11t} + \varepsilon_t$ ใช้สถิติทดสอบ partial F

$$F = \frac{\left|SSR(t, t^2, x_1, \ldots, x_{11}) - SSR(t, t^2)\right|/(11)}{SSE(t, t^2, x_1, \ldots, x_{11})/(n-14)}$$

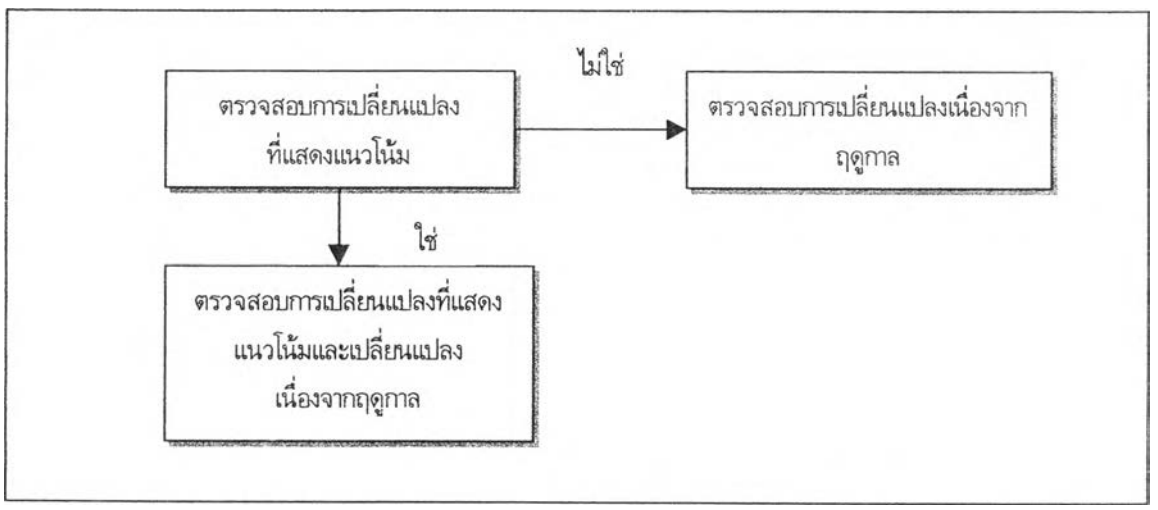
เพื่อทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงเนื่องจากฤดูกาล

ทดสอบ $H_0 : \beta_3 = \beta_4 = \ldots = \beta_{13} = 0$

$H_a : \beta_3, \beta_4, \ldots, \beta_{13}$ มีค่าไม่เป็น 0 อย่างน้อยหนึ่งค่า

3. หากพบว่าไม่มีแนวโน้มจากผลการตรวจสอบสมมติฐานในข้อที่ 1 จะตรวจสอบว่ามีฤดูกาลอย่างเดียวหรือไม่ โดยการทดสอบแบบ overall F จากรูปแบบฤดูกาล ดังภาพ 8



ภาพ 8 การตรวจสอบการเปลี่ยนแปลงแนวโน้มและ/หรือฤดูกาล

ตอนที่ 2.2 การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม สำหรับข้อมูลอนุกรมเวลารายปี โดยการนำข้อมูลอนุกรมเวลามาพล็อตกราฟ เพื่อดูลักษณะการเคลื่อนไหวของอนุกรมเวลา และตรวจสอบโดยใช้การวิเคราะห์การถดถอย เพื่อตรวจสอบว่าข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มจริงหรือไม่ โดยการทดสอบแบบ partial t

1. ตรวจสอบว่าข้อมูลมีแนวโน้มแบบเส้นโค้งหรือไม่ จากโมเดลที่แสดงแนวโน้มแบบเส้นโค้ง คือ $y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$ ใช้สถิติทดสอบ partial t

$$t = \frac{b_2}{s_{b2}}$$

เพื่อทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นโค้ง

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มเส้นโค้ง

หรือทดสอบ $H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

2. ตรวจสอบว่าข้อมูลมีแนวโน้มแบบเอ็กซ์โปเนนเชียลหรือไม่ จากโมเดลที่แสดงแนวโน้มแบบเอ็กซ์โปเนนเชียล คือ $y_t = \beta_0 \beta_1^t \varepsilon_t$

ซึ่งแปลงให้เป็นโมเดลแบบบวก โดยการหาลอการิทึม ดังนี้

$$\ln y_t = (\ln \beta_0) + (\ln \beta_1)t + (\ln \varepsilon_t)$$

$$y'_t = \beta'_0 + \beta'_1 t + \varepsilon'_t$$

โดยใช้สถิติทดสอบ partial t

$$t = \frac{b'_1}{s'_{b_1}}$$

เพื่อทดสอบสมมติฐานว่า H_0 : ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้มแบบเอ็กซ์โปเนนเชียล

H_a : ข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงที่แสดงแนวโน้มแบบเอ็กซ์โปเนนเชียล

หรือทดสอบ $H_0 : \beta'_1 = 0$

$H_a : \beta'_1 \neq 0$

ตอนที่ 3 พยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีของเบย์ ในแต่ละชุดข้อมูลโดยใช้โปรแกรม Microsoft office 97 (Excel) ซึ่งมี ขั้นตอน ดังนี้

ขั้นที่ 1 กำหนดโมเดลแนวโน้มที่เหมาะสมให้กับข้อมูลอนุกรมเวลา 24 ช่วงเวลาสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล และ 20 ช่วงเวลาสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล โมเดลแนวโน้มที่นิยมใช้ ได้แก่

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

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

แนวโน้มแบบ exponential $y_t = b_1 b_2^t \varepsilon_t$

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

ขั้นที่ 2 กำหนดค่าเริ่มต้น (prior information) สำหรับ $\bar{\mathbf{b}}'$, $\mathbf{V}'$ และ σ_ε^2 โดยใช้วิธีวิเคราะห์การถดถอย (regression analysis) หาค่าเริ่มต้นดังกล่าว

ขั้นที่ 3 คำนวณการแจกแจงภายหลัง (posterior distribution) ของ $\bar{\mathbf{b}}'$ ตามหลักการของเบย์

ขั้นที่ 4 จากสมการพยากรณ์ที่ได้จากขั้นที่ 3 เสนอผลการพยากรณ์ 12 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่มีการแปลงเนื่องจากฤดูกาล และ 5 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล

ตอนที่ 4 พยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีบ็อกซ์-เจนกิ้นส์ ในแต่ละชุดข้อมูลโดยใช้โปรแกรม SPSS for Windows version 10.05 ซึ่งมี 4 ขั้นตอน คือ

ขั้นที่ 1 กำหนดรูปแบบ โดยพิจารณาจากผลการตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และ/หรือฤดูกาลในตอนที่ 2 แล้วจึงนำมากำหนดค่าผลต่าง และ/หรือผลต่างฤดูกาลของค่าสังเกตในอนุกรมเวลา พิจารณาคอแรลโลแกรมเพื่อกำหนดโมเดลจากลักษณะของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน

ขั้นที่ 2 การประมาณค่าพารามิเตอร์ในโมเดลที่กำหนดในขั้นที่ 1

ขั้นที่ 3 การตรวจสอบความเหมาะสมของโมเดลโดยพิจารณาจากคลอแรลโลแกรมของความคลาดเคลื่อน การทดสอบค่าพารามิเตอร์ในรูปแบบด้วยการทดสอบแบบ partial t และทดสอบโดยใช้สถิติทดสอบ Box-Ljung Chi-Square Statistic

$$Q'_m = n(n+2)\sum_{k=1}^{m}\frac{r_k^2(e_t)}{n-k}$$

ขั้นที่ 4 นำสมการพยากรณ์ที่เหมาะสม หากมีมากกว่าหนึ่งสมการที่เหมาะสมในขั้นที่ 3 จะเลือกเพียงหนึ่งสมการเพื่อการพยากรณ์โดยการเปรียบเทียบค่าความแปรปรวนของค่าความคลาดเคลื่อนในแต่ละสมการที่เป็นไปได้ ไปใช้ในการพยากรณ์ 12 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่มีการแปลงเนื่องจากฤดูกาล และ 5 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล

ตอนที่ 5 พยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีพยากรณ์ร่วม ในแต่ละชุดข้อมูล

ขั้นที่ 1 แบ่งออกเป็น 2 วิธี ได้แก่ วิธีที่ 1 ได้จากการนำค่าพยากรณ์จากวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลมารวมกันโดยให้น้ำหนักตามแบบของ Newbold และ Granger แบบที่ 1 และวิธีที่ 2 ได้จากการนำค่าพยากรณ์จากวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลมารวมกันโดยให้น้ำหนักตามแบบของ Newbold และ Granger แบบที่ 2 โดยพิจารณาเลือกวิธีพยากรณ์เดี่ยวที่เหมาะสมกับลักษณะของข้อมูลจาก 3 กรณี คือ กรณีแรก เมื่อข้อมูลไม่มีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีเฉลี่ยเคลื่อนที่แบบง่าย (simple moving average: SMA) และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลครั้งเดียว (single exponential smoothing: SES) กรณีที่สอง เมื่อข้อมูลมีอิทธิพลของแนวโน้มเพียงอย่างเดียวจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของบราวน์ (Brown's double exponential smoothing: DES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 2 ครั้งตามแบบของโฮลท์ (Holt's linear exponential smoothing: LES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล 3 ครั้งตามแบบของบราวน์ (Brown's triple exponential smoothing: TES) วิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียลสำหรับข้อมูลที่มีแนวโน้มแบบเอ็กซ์โปเนนเชียล (exponential smoothing for exponential trend: ESE) และวิธีการวิเคราะห์การถดถอย (Regression analysis: REG) และกรณีที่สาม เมื่อข้อมูลมีอิทธิพลของแนวโน้มและฤดูกาลจะเลือกวิธีพยากรณ์เดี่ยวจากวิธีปรับให้เรียบแบบโฮลท์-

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

ขั้นที่ 2 คำนวณหาค่าน้ำหนักของแต่ละวิธี จากสูตรดังนี้

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

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

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

ขั้นที่ 3 คำนวณหาค่าพยากรณ์จากสูตร

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

เมื่อ n = จำนวนข้อมูล

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

$\hat{y}_t$ = ค่าพยากรณ์ ณ เวลาที่ t

w_i = น้ำหนักของวิธีการพยากรณ์ที่ i

p = จำนวนวิธีการพยากรณ์

โดยที่น้ำหนักแบบที่ 1 จากงานวิจัยของ Winkler และ Makridakis (1983) กำหนดให้ υ มีค่า 3, 6, 9 และ 12 ในงานวิจัยนี้เพื่อให้มีความละเอียดยิ่งขึ้นผู้วิจัยจึงกำหนดให้ υ มีค่าตั้งแต่ 1, 2, ..., 20 และน้ำหนักแบบที่ 2 จากงานวิจัยของ Winkler และ Makridakis (1983) กำหนดให้ υ มีค่า 3, 6, 9 และ 12 กับ β มีค่าเท่ากับ .5, .7 และ .9 ในงานวิจัยนี้เพื่อให้มีความละเอียดยิ่งขึ้นผู้วิจัยจึงกำหนดให้ υ มีค่า 3, 6, 9 และ 12 กับ β มีค่าเท่ากับ .1, .2, ..., .9 เพื่อหาค่าน้ำหนักที่ให้ค่าพยากรณ์ที่มีความคลาดเคลื่อน SSE ต่ำที่สุด

ขั้นที่ 4 นำเสนอผลการพยากรณ์ 12 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล และ 5 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล

ตอนที่ 6 ตรวจสอบผลการพยากรณ์ โดยนำผลจากการพยากรณ์ในตอนที่ 3, ตอนที่ 4 และตอนที่ 5 มาเปรียบเทียบกัน ซึ่งในการศึกษาครั้งนี้ผู้วิจัยได้ตัดข้อมูลให้เหลือ 24 ค่าสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล และ 20 ค่าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจาก

ฤดูกาล เพื่อทำการเปรียบเทียบผลการพยากรณ์ของวิธีบ็อกซ์-เจนกิ้นส์ และวิธีการพยากรณ์ร่วม กับวิธีของเบย์ในกรณีที่ข้อมูลมีขนาดเล็ก เหตุผลที่ตัดข้อมูลให้เหลือเพียง 24 ค่าสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล เนื่องมาจากการศึกษาของ de Alba และ Mendoza (2001) ได้ใช้ข้อมูลรายเดือน 2 ปี (24 เดือน) ทำการพยากรณ์ด้วยวิธีของเบย์เปรียบเทียบกับวิธีบ็อกซ์-เจนกิ้นส์ทั้งกรณีที่ใช้ข้อมูลทั้งหมด และข้อมูลมีขนาดเท่ากันกับวิธีของเบย์ และเหตุผลที่ตัดข้อมูลให้เหลือเพียง 20 ค่าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล เพราะในการศึกษาครั้งนี้ผู้วิจัยต้องการทำการศึกษาการพยากรณ์ด้วยวิธีของเบย์สำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล จะให้ผลสอดคล้องกับวิธีของเบย์สำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาลหรือไม่ และ 20 ค่าที่ตัดให้เหลือใช้แนวคิดคล้ายกันกับข้อมูลรายเดือน คือใช้ข้อมูลรายปีในช่วง 2 ทศวรรษ (20 ปี) โดยมีขั้นตอนตรวจสอบผลการพยากรณ์ ดังนี้

ขั้นที่ 1 พล็อตกราฟค่าข้อมูลอนุกรมเวลาที่เกิดขึ้นจริง กับค่าผลการพยากรณ์ด้วยวิธีของเบย์ วิธีบ็อกซ์-เจนกิ้นส์ และวิธีพยากรณ์ร่วม 2 วิธี ซึ่งวิธีบ็อกซ์-เจนกิ้นส์ และวิธีพยากรณ์ร่วมได้แยกเป็น 2 กรณี คือกรณีแรกใช้ข้อมูล 76 ช่วงเวลา (เก็บไว้ 12 ช่วงเวลาเพื่อเปรียบเทียบค่าพยากรณ์กับค่าจริง) ในการพยากรณ์ 12 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่มีการแปลงเนื่องจากฤดูกาล และ 59 ช่วงเวลา (เก็บไว้ 5 ช่วงเวลาเพื่อเปรียบเทียบค่าพยากรณ์กับค่าจริง) ในการพยากรณ์ 5 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ซึ่งต่อไปผู้วิจัยจะเรียกว่า การพยากรณ์สำหรับข้อมูลขนาดใหญ่ และกรณีที่สองใช้ข้อมูล 24 ช่วงเวลา (เก็บไว้ 12 ช่วงเวลาเพื่อเปรียบเทียบค่าพยากรณ์กับค่าจริง) ในการพยากรณ์ 12 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล และ 20 ช่วงเวลา (เก็บไว้ 5 ช่วงเวลาเพื่อเปรียบเทียบค่าพยากรณ์กับค่าจริง) ในการพยากรณ์ 5 ช่วงเวลาล่วงหน้าสำหรับข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ซึ่งต่อไปผู้วิจัยจะเรียกว่า การพยากรณ์สำหรับข้อมูลขนาดเล็ก ดังนั้นมีวิธีการพยากรณ์ที่นำมาพิจารณารวม 7 วิธี ได้แก่

1. วิธีของเบย์สำหรับข้อมูลขนาดเล็ก (BAYES)
2. วิธีบ็อกซ์-เจนกิ้นส์สำหรับข้อมูลขนาดใหญ่ (B-J)
3. วิธีบ็อกซ์-เจนกิ้นส์สำหรับข้อมูลขนาดเล็ก (B-J/2T)
4. วิธีการพยากรณ์ร่วมโดยให้น้ำหนักแบบของ Newbold และ Granger แบบที่ 1 สำหรับข้อมูลขนาดใหญ่ (COMB1)
5. วิธีการพยากรณ์ร่วมโดยให้น้ำหนักแบบของ Newbold และ Granger แบบที่ 1 สำหรับข้อมูลขนาดเล็ก (COMB1/2T)
6. วิธีการพยากรณ์ร่วมโดยให้น้ำหนักแบบของ Newbold และ Granger แบบที่ 2 สำหรับข้อมูลขนาดใหญ่ (COMB2)
7. วิธีการพยากรณ์ร่วมโดยให้น้ำหนักแบบของ Newbold และ Granger แบบที่ 2 สำหรับข้อมูลขนาดเล็ก (COMB2/2T)

ขั้นที่ 2 คำนวณค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) และค่าความคลาดเคลื่อนเฉลี่ย (ME) ในการพยากรณ์แต่ละวิธีจากสูตร

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

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

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

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

y_t = ข้อมูลจริง ณ เวลาที่ t

n = จำนวนข้อมูล

ขั้นที่ 3 เปรียบเทียบผลความคลาดเคลื่อนของการพยากรณ์ทั้ง 7 วิธี

ตอนที่ 7 เลือกวิธีการพยากรณ์ที่มีค่าเปอร์เซ็นต์ความคลาดเคลื่อนสมบูรณ์เฉลี่ย (MAPE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) น้อยที่สุดสำหรับข้อมูลแต่ละชุดเพื่อนำมาพยากรณ์ข้อมูลอนุกรมเวลาล่วงหน้า และพิจารณาวิธีการพยากรณ์จากค่าความคลาดเคลื่อนเฉลี่ย (ME) ดังนี้

ถ้า ME มีค่าสูงเป็นบวก แสดงว่า ค่าพยากรณ์ส่วนใหญ่น้อยกว่าค่าจริง

ถ้า ME มีค่าสูงเป็นลบ แสดงว่า ค่าพยากรณ์ส่วนใหญ่มากกว่าค่าจริง

ถ้า ME มีค่าเข้าใกล้ 0 แสดงว่า ค่าพยากรณ์ใกล้เคียงกับค่าจริง