

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงบรรยาย เพื่อประสงค์ใช้วิธีของบ็อกซ์และเงินกินส์ ในการพยากรณ์ข้อมูลอนุกรมเวลาที่มีและไม่มี การเปลี่ยนแปลงเนื่องจากฤดูกาล พร้อมทั้งตรวจสอบผลการพยากรณ์ด้วยวิธีบ็อกซ์และเงินกินส์ กับวิธีการพยากรณ์ 3 แบบ คือ การวิเคราะห์ การถดถอย การเฉลี่ยเคลื่อนที่ และการปรับให้เรียบแบบเอ็กซ์โปเนนเชียล โดยให้เกณฑ์ในการเปรียบเทียบจากการวัดความคลาดเคลื่อน 6 แบบ โดยมีวิธีดำเนินการวิจัยดังนี้

ลักษณะและแหล่งที่มาของข้อมูล

ฐานข้อมูล (data base) ที่ใช้ในการวิจัยครั้งนี้มี 2 ฐาน ฐานข้อมูลแรก คือ ฐานข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ได้แก่ ปริมาณการดื่มสิ่งพิมพ์ ของศูนย์บรรณสารสถานศึกษาทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ตั้งแต่เดือนมิถุนายน ปีการศึกษา 2535 ถึงเดือนตุลาคมปีการศึกษา 2540 รวมเป็นอนุกรมเวลารายเดือน 65 เดือน ฐานข้อมูลชุดนี้ประกอบด้วยข้อมูลอนุกรมเวลา 3 ชุด จำนวนคาบสังเกตสิ่งพิมพ์ คือ หนังสือทั่วไป หนังสือตำรา และวิทยานิพนธ์ ฐานข้อมูลฐานที่สอง คือ ฐานข้อมูลอนุกรมเวลาที่ไม่มีการเปลี่ยนแปลงเนื่องจากฤดูกาล ประกอบด้วย ข้อมูลอนุกรมเวลาจำนวนนักเรียน 2 ชุด จำนวนตามระดับการศึกษา ได้แก่ จำนวนนักเรียนระดับประถมศึกษา และระดับมัธยมศึกษา ปีการศึกษา 2480-2539 เป็นอนุกรมเวลารายปี 60 ปี

แหล่งที่มาของข้อมูล ข้อมูลสำหรับการวิจัยครั้งนี้เป็นข้อมูลอนุกรมเวลาจากแหล่งสถิติภูมิ ซึ่งเป็นรายงานสถิติที่หน่วยงานจัดทำขึ้น ประกอบด้วยแหล่งที่มา 2 แหล่ง คือ ข้อมูลจำนวนนักเรียนระดับประถมศึกษาและระดับมัธยมศึกษา มาจากหนังสือสถิติรายปีประเทศไทย ส่วนข้อมูลปริมาณการดื่มสิ่งพิมพ์ อันได้แก่ หนังสือทั่วไป หนังสือตำรา และวิทยานิพนธ์ มาจากบันทึกสถิติรายเดือนของศูนย์บรรณสารสถานศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัย เป็นแบบบันทึกข้อมูลที่ผู้วิจัยสร้างขึ้นเอง ดังตาราง 15-16

ตาราง 15 แบบบันทึกข้อมูลปริมาณการเข้าชมสิ่งพิมพ์ประเภท..... (เล่ม)

ปีการศึกษา เดือน	2535	2536	2537	2538	2539	2540
มิถุนายน						
.						
.						
.						
พฤษภาคม						

ตาราง 16 แบบบันทึกข้อมูลจำนวนนักเรียน (คน)

ปีการศึกษา	จำนวนนักเรียนระดับประถมศึกษา	จำนวนนักเรียนระดับมัธยมศึกษา
2480		
.		
.		
.		
2539		

การเก็บรวบรวมข้อมูล

การวิจัยครั้งนี้ ผู้วิจัยบันทึกข้อมูลด้วยตนเอง โดยมิขึ้นตอนการรวบรวมข้อมูล ดังต่อไปนี้

1. ขั้นตอนการติดต่อขอความร่วมมือ ผู้วิจัยติดต่อขอความร่วมมือไปยังเจ้าหน้าที่บรรณารักษ์ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เพื่อขอบันทึกข้อมูลปริมาณการเข้าชมสิ่งพิมพ์ เจ้าหน้าที่บรรณารักษ์จึงดำเนินการประสานงานไปยัง

เจ้าหน้าที่ฝ่ายบริการของศูนย์บรรณสารสนเทศทางการศึกษา เพื่อให้ผู้วิจัยไปสำรวจและบันทึกข้อมูลในขั้นตอนต่อไป นอกจากนี้ข้อมูลจำนวนนักเรียนบางท่าน ผู้วิจัยติดต่อขอความร่วมมือ ไปยังเจ้าหน้าที่บรรณารักษ์ ของห้องสมุดสำนักงานสถิติแห่งชาติ เพื่อขอบันทึกข้อมูลจากหนังสือสถิติรายปีประเทศไทย

2. ขั้นตอนการสำรวจและบันทึกข้อมูล ผู้วิจัยสำรวจและบันทึกข้อมูลปริมาณการพิมพ์พิมพ์ 3 ประเภท ได้แก่ ปริมาณการพิมพ์หนังสือทั่วไป หนังสือตำรา และวิทยานิพนธ์ของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย จากฝ่ายบริการของศูนย์บรรณสารสนเทศทางการศึกษา ลงในแบบบันทึกข้อมูลตามตาราง 15 ส่วน ข้อมูลจำนวนนักเรียนระดับประถมศึกษาและระดับมัธยมศึกษา ผู้วิจัยสำรวจข้อมูลจากหนังสือสถิติรายปีประเทศไทย ที่มีอยู่ในศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปรากฏว่ามีข้อมูลไม่ครบถ้วนสมบูรณ์ ผู้วิจัยจึงไปศึกษาค้นคว้าข้อมูลเพิ่มเติมจากหนังสือสถิติรายปีประเทศไทย ที่มีอยู่ในห้องสมุดสำนักงานสถิติแห่งชาติ จนกระทั่งได้ข้อมูลครบถ้วนสมบูรณ์ ผู้วิจัยจึงบันทึกข้อมูลลงในแบบบันทึกข้อมูลตามตาราง 16

3. ขั้นตอนการบรรณาธิกรณ (edit) ข้อมูล ผู้วิจัยตรวจสอบความถูกต้องและความสมบูรณ์ของข้อมูลที่ได้จากการรวบรวมข้อมูลทั้งหมด

การรวบรวมข้อมูลทั้ง 3 ขั้นตอน ใช้เวลาในการดำเนินการประมาณ 2 สัปดาห์ คือ ในช่วงวันที่ 17 - 30 พฤศจิกายน 2540

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล ผู้วิจัยใช้เครื่องคอมพิวเตอร์ PC โปรแกรม SAS และโปรแกรม SPSS for Windows ในส่วนของโปรแกรม SAS ผู้วิจัยติดต่อขอความช่วยเหลือขอใช้โปรแกรมที่ศูนย์คอมพิวเตอร์ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ในการวิเคราะห์ข้อมูล ผู้วิจัยนำข้อมูลอนุกรมเวลารายปีที่มีค่าสังเกต 60 ค่า จำนวน 2 ชุด และอนุกรมเวลารายเดือนที่มีค่าสังเกต 65 ค่า จำนวน 3 ชุด มาทำการวิเคราะห์โดยแบ่งขั้นตอนการวิเคราะห์ออกเป็น 4 ตอน ในตอนที่ 1, 2 และ 4 ใช้โปรแกรม

SPSS for Windows ส่วนตอนที่ 3 ใช้โปรแกรม SAS และโปรแกรม SPSS for Windows การวิเคราะห์ข้อมูลในแต่ละตอน มีรายละเอียดดังต่อไปนี้

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐาน ค่าความแปร ค่าความโค้ง และค่าการกระจายของข้อมูล

ตอนที่ 2 การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และการเปลี่ยนแปลงเนื่องจากฤดูกาล ผู้วิจัยทำการตรวจ 2 ขั้นตอน คือ ขั้นตอนแรกตรวจสอบโดยการพล็อตกราฟเป็นการนำอนุกรมเวลาตามเชิงกราฟ เพื่อดูลักษณะการเคลื่อนไหวของอนุกรมเวลา ส่วนขั้นตอนที่สองตรวจสอบโดยใช้การวิเคราะห์การถดถอย ในขั้นตอนนี้เป็นการนำผลจากการพล็อตกราฟมาตรวจสอบว่า ข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงที่แสดงแนวโน้มและ/หรือมีการเปลี่ยนแปลงเนื่องจากฤดูกาลจริงหรือไม่ โดยตรวจสอบทั้งรูปแบบบวกและรูปแบบลบ

การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้ม และการเปลี่ยนแปลงเนื่องจากฤดูกาล โดยใช้การวิเคราะห์การถดถอย ต้องตรวจสอบ 2 ขั้นตอน ขั้นแรกเป็นการตรวจสอบว่าข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงที่แสดงแนวโน้มหรือไม่ ถ้าผลการทดสอบ พบว่า ข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้ม ในขั้นที่สอง นักวิจัยต้องทำการตรวจสอบการเปลี่ยนแปลงเนื่องจากฤดูกาลต่อไป แต่ถ้าผลการทดสอบ พบว่า ข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงที่แสดงแนวโน้ม ในขั้นที่สอง นักวิจัยต้องทำการตรวจสอบทั้งการเปลี่ยนแปลงที่แสดงแนวโน้ม และการเปลี่ยนแปลงเนื่องจากฤดูกาล

ตอนที่ 3 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีบ็อกซ์และเจนกินส์ ขั้นตอนการวิเคราะห์ในตอนนี้ มี 5 ขั้นตอน คือ

ขั้นที่ 1 การกำหนดโมเดล โดยพิจารณาจากลักษณะของฟังก์ชันอัตสหสัมพันธ์และฟังก์ชันอัตสหสัมพันธ์พาร์เรียล เพื่อตรวจสอบว่าข้อมูลมีลักษณะคงที่หรือไม่ ถ้าค่าฟังก์ชันอัตสหสัมพันธ์มีค่าลดลงอย่างรวดเร็ว แสดงว่าอนุกรมเวลานั้นเป็นอนุกรมคงที่ ผู้วิจัยสามารถดำเนินการในขั้นที่ 3 ได้เลย แต่ถ้าฟังก์ชันอัตสหสัมพันธ์มีค่าลดลงอย่างช้า ๆ แสดงว่า อนุกรมเวลานั้นเป็นอนุกรมไม่คงที่และ/หรือมีการเปลี่ยนแปลงเนื่องจากฤดูกาลซ้ำๆ เกือบข้อหวัช ต้องดำเนินการในขั้นที่ 2 ต่อไป

ขั้นที่ 2 การหาผลต่าง (regular differencing) และ / หรือผลต่างฤดูกาล (seasonal differencing) เพื่อให้ได้ออนุกรมเวลาดคงที่

ขั้นที่ 3 การประมาณค่าพารามิเตอร์ในโมเดล

ขั้นที่ 4 การตรวจสอบความเหมาะสมของโมเดล โดยใช้สถิติทดสอบ Box-Ljung Chi-Square Statistic

$$Q'_m = N(N-d+2) \sum_{k=1}^m (r_k^2(a_k)) / (N-d-k)$$

ขั้นที่ 5 การนำโมเดลไปใช้ในพยากรณ์ 5 ช่วงเวลาดังหน้า

ในตอนต่อไป ผู้วิจัยต้องตรวจสอบผลการพยากรณ์ด้วยวิธีบ็อกซ์และเจ็นกินส์ กับผลการพยากรณ์ด้วยวิธีวิเคราะห์การถดถอย วิธีเรถี่ยเคลื่อนที่ และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล ในตอนนี้ ผู้วิจัยจึงนำข้อมูลอนุกรมเวลาที่ 5 จุด มาพยากรณ์ด้วยวิธีวิเคราะห์การถดถอย และวิธีปรับให้เรียบแบบเอ็กซ์โปเนนเชียล โดยใช้โปรแกรม SPSS for Windows ส่วนวิธีการเรถี่ยเคลื่อนที่ ใช้โปรแกรม SAS แล้วนำเสนอผลการพยากรณ์ 5 ช่วงเวลาดังหน้า โดยกราฟต่อกราฟ

ตอนที่ 4 การตรวจสอบผลการพยากรณ์ การวิเคราะห์ในขั้นตอนนี้เป็นการนำผลจากการพยากรณ์ด้วยวิธีบ็อกซ์และเจ็นกินส์ มาเปรียบเทียบกับผลการพยากรณ์ด้วยวิธีวิเคราะห์การถดถอย วิธีเรถี่ยเคลื่อนที่ และวิธีทำให้เรียบแบบเอ็กซ์โปเนนเชียล โดยมีขั้นตอนดังนี้

ขั้นที่ 1 พล็อตกราฟค่าข้อมูลอนุกรมเวลาที่เกิดขึ้นจริง กับค่าผลการพยากรณ์ด้วยวิธีบ็อกซ์และเจ็นกินส์ วิเคราะห์การถดถอย วิธีเรถี่ยเคลื่อนที่ และวิธีทำให้เรียบแบบเอ็กซ์โปเนนเชียล

ขั้นที่ 2 คำนวณความคลาดเคลื่อนในการพยากรณ์แต่ละวิธี ด้วยค่าวัดความคลาดเคลื่อน 6 แบบ

- ให้
- m - วิธีการพยากรณ์
 - F_m - ค่าพยากรณ์จากวิธี m
 - F_{rw} - ค่าพยากรณ์จากวิธีเดินอย่างสุ่ม (random walk)
 - Z - ค่าตั้งภาคที่เกิดขึ้นจริง
 - S - จำนวนอนุกรมเวลา

สูตรที่ 1 รากที่สองของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error หรือ RMSE)

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

สูตรที่ 2 ค่ามัธยฐานของค่าสัมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Median Absolute Percentage Error หรือ MdAPE)

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

เมื่อนำค่าสังเกต APE_m มาเรียงลำดับ (rank-ordered) จะได้

$$MdAPE_m = (S+1) / 2 \quad ; S \text{ เป็นจำนวนคู่}$$

$$MdAPE_m = (S/2) + 1 \quad ; S \text{ เป็นจำนวนคี่}$$

สูตรที่ 3 ค่าเฉลี่ยของค่าสัมบูรณ์ของความคลาดเคลื่อนวัดในรูปร้อยละ (Mean Absolute Percentage Error หรือ MAPE)

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

$$MAPE_m = \frac{\sum_{s=1}^S APE_m}{S}$$

สูตรที่ 4 ค่าเฉลี่ยเรขาคณิตของความคลาดเคลื่อนสัมบูรณ์สัมพัทธ์ (Geometric Mean of the Relative Absolute Error หรือ GMRAE)

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

$$GMRAE_m = \left[\prod_{s=1}^S RAE_m \right]^{1/S}$$

สูตรที่ 5 ค่ามัธยฐานของค่าสัมบูรณ์สัมพัทธ์ของความคลาดเคลื่อน (Median Relative Absolute Error หรือ MdRAE)

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

เมื่อนำค่าสังเกต RAE_m มนวิเรียงลำดับ (rank-ordered) จะได้

$$MdRAE_m = (S+1) / 2 \quad ; S \text{ เป็นจำนวนคู่}$$

$$MdRAE_m = (S/2) + 1 \quad ; S \text{ เป็นจำนวนคู่}$$

สูตรที่ 6 ร้อยละที่ดีกว่า (Percent Better)

$$\text{Percent Better}_m = \frac{\sum_{i=1}^S j_i}{S} \times 100$$

เมื่อ $j_i = 1$ ถ้า $|F_m - Z| < |F_{rw} - Z|$

$j_i = 0$ ในกรณีอื่น ๆ

ขั้นที่ 3 เปรียบเทียบผลการพยากรณ์ทุกช่วงเวลา และขนาดความคลาดเคลื่อนในการพยากรณ์ ด้วยวิธีบ็อกซ์และแจนกินส์ กับวิธีวิเคราะห์การถดถอย วิธีการเฉลี่ยเคลื่อนที่ และวิธีทำให้เวียบแบบเอ็กซ์โปเนนเชียล นอกจากนี้ผู้วิจัยได้หาความสัมพัทธ์ของค่าความคลาดเคลื่อน เพื่อดูความสัมพัทธ์ระหว่างค่าความคลาดเคลื่อนของผลการพยากรณ์ในแต่ละวิธี

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย