

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงบรรยาย เพื่อเปรียบเทียบความคลาดเคลื่อนในการพยากรณ์ข้อมูลอนุกรมเวลาทางการศึกษาที่คงที่และไม่คงที่มีอินเตอร์เวนชัน จากวิธีพยากรณ์ 3 วิธี คือวิธีบ็อกซ์และเจนกินส์ วิธีบ็อกซ์เจนกินส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และการวิเคราะห์การถดถอยที่ใช้ตัวแปรคัมมี โดยใช้เกณฑ์ในการเปรียบเทียบจากการวัดความคลาดเคลื่อน 6 แบบ โดยมีวิธีดำเนินการวิจัยดังนี้

ลักษณะแหล่งที่มาของข้อมูล

ฐานข้อมูลที่ใช้ในการวิจัยครั้งนี้มี 2 ชุด ได้แก่ ชุดแรกเป็นฐานข้อมูลเกี่ยวกับจำนวนผู้เข้าใช้บริการในศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โดยผู้วิจัยใช้ฐานข้อมูลของเอกภพ ยานะวิมุติ (2543) ตั้งแต่เดือนมิถุนายน ปีการศึกษา 2535 ถึงเดือนพฤศจิกายน ปีการศึกษา 2542 และผู้วิจัยทำการเก็บรวบรวมข้อมูลเพิ่มเติมโดยเริ่มตั้งแต่เดือนธันวาคม ปีการศึกษา 2542 ถึงเดือนตุลาคม ปีการศึกษา 2543 และชุดที่สองเป็นฐานข้อมูลเกี่ยวกับปริมาณการยืมหนังสือระหว่างห้องสมุดของอาจารย์ นิสิต และบุคลากร ภายนอกคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย จากศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ซึ่งผู้วิจัยทำการเก็บรวบรวมข้อมูลโดยเริ่มตั้งแต่เดือนมิถุนายน ปีการศึกษา 2535 ถึงเดือนตุลาคม ปีการศึกษา 2543 ดังนั้นฐานข้อมูลที่นำมาใช้ในการวิจัยครั้งนี้จึงประกอบไปด้วยอนุกรมเวลา 2 ชุด เป็นอนุกรมเวลารายเดือน ประกอบไปด้วยจำนวนผู้เข้าใช้บริการในห้องสมุดจำนวน 101 เดือนและปริมาณการยืมหนังสือระหว่างห้องสมุดจำนวน 101 เดือน แหล่งที่มาของการเก็บรวบรวมข้อมูลทั้ง 2 ชุด ได้แก่ บันทึกสถิติรายเดือนของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลครั้งนี้ คือ แบบบันทึกจำนวนผู้เข้าใช้บริการในห้องสมุดและปริมาณการยืมหนังสือระหว่างห้องสมุดของศูนย์บรรณสารสนเทศทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ดังตาราง 13

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล ผู้วิจัยใช้โปรแกรม SPSS (Statistical Package for the Social Science) for Windows version 9.0 โปรแกรม SAS for Windows version 6.12 และโปรแกรม Microsoft Office 97 (EXCEL)

ในการวิเคราะห์ข้อมูล ผู้วิจัยนำข้อมูลอนุกรมเวลารายเดือนที่มีค่าสังเกต 101 ค่า จำนวน 2 ชุด มาทำการวิเคราะห์ โดยแบ่งขั้นตอนการวิเคราะห์เป็น 4 ตอน ในตอนที่ 1 และ 2 ใช้โปรแกรม SPSS ตอนที่ 3 ใช้โปรแกรม SPSS และโปรแกรม SAS ส่วนตอนที่ 4 ใช้โปรแกรม Microsoft Office 97 (EXCEL) มีรายละเอียดดังต่อไปนี้

ตอนที่ 1 การวิเคราะห์ค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐาน ค่าความแปรปรวน ค่าความโค้ง และค่าการกระจายของข้อมูล

ตอนที่ 2 การตรวจสอบความคงที่ การเปลี่ยนแปลงที่แสดงแนวโน้ม และการเปลี่ยนแปลงเนื่องจากฤดูกาล โดยผู้วิจัยทำการตรวจสอบใน 2 ขั้นตอน คือ ขั้นตอนแรก ผู้วิจัยตรวจสอบโดยการพล็อตกราฟเพื่อดูลักษณะการเคลื่อนไหวของข้อมูลอนุกรมเวลา และในขั้นตอนที่สอง ผู้วิจัยตรวจสอบโดยใช้การวิเคราะห์การถดถอย ซึ่งในขั้นตอนนี้เป็นการนำผลจากการพล็อตกราฟมาตรวจสอบว่าข้อมูลอนุกรมเวลา มีการเปลี่ยนแปลงที่แสดงแนวโน้มหรือมีการเปลี่ยนแปลงเนื่องจากฤดูกาล หรือมีการเปลี่ยนแปลงทั้งที่แสดงแนวโน้มและฤดูกาลหรือไม่ โดยตรวจสอบทั้งรูปแบบบวกและรูปแบบคูณ

การตรวจสอบการเปลี่ยนแปลงที่แสดงแนวโน้มและการเปลี่ยนแปลงเนื่องจากฤดูกาล โดยใช้การวิเคราะห์การถดถอยต้องตรวจสอบ 2 ขั้นตอน ขั้นแรกเป็นการตรวจสอบว่าข้อมูลอนุกรมเวลา มีการเปลี่ยนแปลงที่แสดงแนวโน้มหรือไม่ ถ้าผลการตรวจสอบพบว่าข้อมูลอนุกรมเวลาไม่มีการเปลี่ยนแปลงที่แสดงแนวโน้ม ในขั้นที่สองจึงทำการตรวจสอบว่าข้อมูลอนุกรมเวลา มีการเปลี่ยนแปลงเนื่องจากฤดูกาลหรือไม่ แต่ถ้าผลการทดสอบพบว่าข้อมูลอนุกรมเวลา มีการเปลี่ยนแปลงที่แสดงแนวโน้ม ในขั้นที่สองจึงทำการตรวจสอบว่าข้อมูลอนุกรมเวลา มีการเปลี่ยนแปลงที่แสดงแนวโน้มและการเปลี่ยนแปลงเนื่องจากฤดูกาลหรือไม่

ตอนที่ 3 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธีบ็อกซ์และเงินกินส์ วิธีบ็อกซ์และเงินกินส์ที่ใช้เทคนิคของโมเดลอาร์มาอินเตอร์เวนชัน และการวิเคราะห์การถดถอยโดยใช้ตัวแปรดัมมี่

พร้อมทั้งการพยากรณ์อนุกรมเวลาล่วงหน้า 5 ช่วงเวลา ซึ่งผู้วิจัยแยกนำเสนอออกเป็น 3 ตอนตามวิธีการวิเคราะห์ดังนี้

ตอนที่ 3.1 การวิเคราะห์ด้วยวิธีบ็อกซ์และเจนกินส์ การวิเคราะห์นี้ผู้วิจัยนำข้อมูลอนุกรมเวลาจำนวนผู้เข้าใช้บริการห้องสมุดและปริมาณการยืมหนังสือระหว่างห้องสมุดมาวิเคราะห์ และพยากรณ์ค่าในอนาคตล่วงหน้า 5 ช่วงเวลา โดยใช้โปรแกรม SPSS version 9.0 และโปรแกรม SAS ซึ่งมีขั้นตอนในการวิเคราะห์ดังนี้

ขั้นที่ 1 การกำหนดโมเดล โดยพิจารณาจากลักษณะของฟังก์ชันอัตตะสหัสสัมพันธ์และฟังก์ชันอัตตะสหัสสัมพันธ์บางส่วน เพื่อตรวจสอบว่าข้อมูลอนุกรมเวลามีลักษณะคงที่หรือไม่ ถ้าพบว่ามีค่าคงที่ก็สามารถดำเนินการในขั้นที่ 3 ได้เลย แต่ถ้าพบว่าอนุกรมเวลาชุดนั้นไม่คงที่ อาจเนื่องมาจากมีการเปลี่ยนแปลงที่แสดงแนวโน้มและ/หรือมีการเปลี่ยนแปลงเนื่องจากฤดูกาลเข้ามาเกี่ยวข้องด้วยต้องดำเนินการในขั้นที่ 2 ต่อไป

ขั้นที่ 2 การหาผลต่าง (regular differencing) และ/หรือผลต่างฤดูกาล (seasonal differencing) เพื่อให้ได้อนุกรมเวลาที่คงที่ โดยการนำอนุกรมเวลาชุดที่มีเวลาถ่วงน้ำหนักกับ k ไปลบอนุกรมเวลาชุดเดิม ถ้าเป็นการหาผลต่างเพื่อที่จะกำจัดการเปลี่ยนแปลงที่แสดงแนวโน้ม ค่า $k = 1, 2, \dots$ โดยมากจะมีค่าไม่เกิน 2 แต่ถ้าเป็นการกำจัดการเปลี่ยนแปลงเนื่องจากฤดูกาล ค่า k จะเริ่มต้นที่ $k = 12, 24, \dots$

ขั้นที่ 3 การประมาณค่าพารามิเตอร์ เพื่อให้ได้ค่าพารามิเตอร์ที่จะใช้ในการพยากรณ์ และนำค่าพารามิเตอร์ที่ได้ไปสร้างสมการพยากรณ์

ขั้นที่ 4 การตรวจสอบความเหมาะสมของโมเดล ผู้วิจัยใช้สถิติในการตรวจสอบ คือ partial t-test เพื่อตรวจสอบพารามิเตอร์ในโมเดลและ Box-Ljung Chi-square statistic เพื่อตรวจสอบความคลาดเคลื่อนที่ได้จากการกำหนดโมเดล สำหรับการตรวจสอบด้วยสถิติ partial t-test ถ้าโมเดลที่ได้ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบ แสดงว่าโมเดลที่ได้นั้นมีความเหมาะสม แต่ถ้าโมเดลที่ได้ไม่ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบ แสดงว่าโมเดลที่ได้นั้นไม่มีความเหมาะสม จะต้องกลับไปเริ่มต้นใหม่ในขั้นที่ 1 ส่วนการตรวจสอบด้วยสถิติ Box-Ljung Chi-square statistic ถ้าโมเดลที่ได้ไม่ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบ แสดงว่า

โมเดลที่ได้นั้นมีความเหมาะสม แต่ถ้าโมเดลที่ได้ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบ แสดงว่าโมเดลที่ได้นั้นไม่มีความเหมาะสม จะต้องกลับไปเริ่มต้นใหม่ในขั้นที่ 1

ขั้นที่ 5 การนำโมเดลไปใช้ในการพยากรณ์ 5 ช่วงเวลาล่วงหน้า ผู้วิจัยนำสมการพยากรณ์ที่สร้างขึ้นโดยผ่านการตรวจสอบความเหมาะสมแล้วไปพยากรณ์ค่าในอนาคต

ตอนที่ 3.2 การวิเคราะห์ด้วยวิธีบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอินเทอร์เวนชัน การวิเคราะห์ด้วยวิธีนี้ผู้วิจัยนำข้อมูลอนุกรมเวลาจำนวนผู้เข้าใช้บริการห้องสมุดและปริมาณการยืมหนังสือระหว่างห้องสมุดมาวิเคราะห์และพยากรณ์ค่าในอนาคตล่วงหน้า 5 ช่วงเวลา โดยใช้โปรแกรม SPSS version 9.0 และโปรแกรม SAS ซึ่งมีขั้นตอนในการวิเคราะห์ดังนี้

ขั้นที่ 1 การกำหนดโมเดล โดยกำหนดโมเดลจากอนุกรมเวลาก่อนการเกิดอินเทอร์เวนชัน (preintervention) ซึ่งจะพิจารณาจากลักษณะของฟังก์ชันอัตตะสหสัมพันธ์และฟังก์ชันอัตตะสหสัมพันธ์บางส่วน เพื่อตรวจสอบว่าข้อมูลอนุกรมเวลามีลักษณะคงที่หรือไม่ ถ้าพบว่ามี ความคงที่ก็สามารถดำเนินการในขั้นที่ 3 ต่อไป แต่ถ้าพบว่าอนุกรมเวลาชุดนั้นไม่คงที่อาจเนื่องมาจากการเปลี่ยนแปลงที่แสดงแนวโน้มและ/หรือมีการเปลี่ยนแปลงเนื่องจากฤดูกาลเข้ามาเกี่ยวข้องด้วย ต้องดำเนินการในขั้นที่ 2 ต่อไป

ขั้นที่ 2 การหาผลต่าง (regular differencing) และ/หรือผลต่างฤดูกาล (seasonal differencing) เพื่อให้ได้อนุกรมเวลาที่คงที่ โดยการนำอนุกรมเวลาชุดที่มีเวลาล้าหลังเท่ากับ k ไปลบอนุกรมเวลาชุดเดิม ถ้าเป็นการหาผลต่างเพื่อที่จะกำจัดการเปลี่ยนแปลงที่แสดงแนวโน้ม ค่า $k = 1, 2, \dots$ โดยมากจะมีค่าไม่เกิน 2 แต่ถ้าเป็นการกำจัดการเปลี่ยนแปลงเนื่องจากฤดูกาล ค่า k จะเริ่มต้นที่ $k = 12, 24, \dots$

ขั้นที่ 3 การตรวจสอบความเหมาะสมของโมเดลก่อนการเกิดอินเทอร์เวนชัน (preintervention) ผู้วิจัยใช้สถิติในการตรวจสอบ คือ partial t-test เพื่อตรวจสอบพารามิเตอร์ในโมเดลและ Box-Ljung Chi-square statistic เพื่อตรวจสอบความคลาดเคลื่อนที่ได้จากการกำหนดโมเดลก่อนการเกิดอินเทอร์เวนชัน สำหรับการตรวจสอบด้วยสถิติ partial t-test ถ้าโมเดลที่ได้ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบแสดงว่าโมเดลที่ได้นั้นมีความเหมาะสม แต่ถ้าโมเดลที่ได้ไม่ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบแสดงว่าโมเดลที่ได้นั้นไม่มีความเหมาะสม จะต้องกลับไปเริ่มต้นใหม่ในขั้นที่ 1 ส่วนการตรวจสอบด้วยสถิติ Box-Ljung Chi-square statistic ถ้า

โมเดลที่ได้ไม่ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบแสดงว่าโมเดลที่ได้นั้นมีความเหมาะสม แต่ถ้าโมเดลที่ได้ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบแสดงว่าโมเดลที่ได้นั้นไม่มีความเหมาะสม จะต้องกลับไปเริ่มต้นใหม่ในขั้นที่ 1

ขั้นที่ 4 พิจารณาหาฟังก์ชันอินเตอร์เวนชัน โดยการหาฟังก์ชันอินเตอร์เวนชันที่เหมาะสมกับอนุกรมเวลาที่เกิดอินเตอร์เวนชัน

ขั้นที่ 5 กำหนดโมเดลอริมาอินเตอร์เวนชัน โดยนำโมเดลอริมาที่ได้จากขั้นที่ 1 มารวมกับฟังก์ชันอินเตอร์เวนชันที่ได้จากขั้นที่ 4

ขั้นที่ 6 การประมาณค่าพารามิเตอร์ เพื่อให้ได้ค่าพารามิเตอร์ที่จะใช้ในการพยากรณ์และนำค่าพารามิเตอร์ที่ได้ไปสร้างสมการพยากรณ์

ขั้นที่ 7 การตรวจสอบความเหมาะสมของโมเดลอริมาอินเตอร์เวนชัน โดยการนำโมเดลอริมาอินเตอร์เวนชันที่ได้จากขั้นที่ 5 มาตรวจสอบความเหมาะสมด้วยการทำ autocorrelation check ของค่าความคลาดเคลื่อน ถ้าโมเดลที่ได้ยอมรับสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบ ($P > 0.05$) แสดงว่าโมเดลที่ได้นั้นมีความเหมาะสม แต่ถ้าโมเดลที่ได้ปฏิเสธสมมติฐานหลักของสถิติที่ใช้ในการตรวจสอบ แสดงว่าโมเดลที่ได้นั้นไม่มีความเหมาะสม จะต้องกลับไปเริ่มต้นใหม่ในขั้นที่ 1

ขั้นที่ 8 การนำโมเดลไปใช้ในการพยากรณ์ล่วงหน้า 5 ช่วงเวลา ผู้วิจัยนำสมการพยากรณ์ที่สร้างขึ้นโดยผ่านการตรวจสอบความเหมาะสมแล้วไปพยากรณ์ค่าในอนาคต

ตอนที่ 3.3 การวิเคราะห์การถดถอยโดยใช้ตัวแปรดัมมี่

การวิเคราะห์ในขั้นตอนนี้ ผู้วิจัยใช้วิธีการวิเคราะห์การถดถอยโดยใช้ตัวแปรดัมมี่เข้ามาใช้ในการวิเคราะห์ข้อมูลซึ่งมีขั้นตอนนี้

ขั้นที่ 1 พิจารณาว่าข้อมูลอนุกรมเวลาจำนวนผู้เข้าใช้บริการในห้องสมุดและปริมาณการยืมหนังสือระหว่างห้องสมุดมีแนวโน้มและ/หรือฤดูกาลหรือไม่ แล้วทำตามขั้นที่ 2 ต่อไป

ขั้นที่ 2 พิจารณารูปแบบของสมการการถดถอย โดยกำหนดให้ตัวแปรต้นมีเป็นตัวแปรอิสระและกำหนดให้อนุกรมเวลาจำนวนผู้เข้าใช้บริการในห้องสมุดและปริมาณการยืมหนังสือระหว่างห้องสมุดเป็นตัวแปรตาม

ขั้นที่ 3 การพยากรณ์ค่าในอนาคต ผู้วิจัยนำสมการการถดถอยที่เหมาะสมมาพยากรณ์จำนวนผู้เข้าใช้บริการในห้องสมุดและปริมาณการยืมหนังสือระหว่างห้องสมุดในอนาคต 5 ช่วงเวลา

ตอนที่ 4 การตรวจสอบความคลาดเคลื่อนในการพยากรณ์

การวิเคราะห์ในขั้นตอนนี้ ผู้วิจัยเปรียบเทียบผลการพยากรณ์ที่ได้จากวิธีพยากรณ์ทั้ง 3 วิธี โดยอาศัยการตรวจสอบความคลาดเคลื่อนในการพยากรณ์ระหว่างวิธีบ็อกซ์และเจนกินส์ วิธีบ็อกซ์และเจนกินส์ที่ใช้เทคนิคของโมเดลอริมาอินเตอร์เวนชัน และการวิเคราะห์การถดถอยโดยใช้ตัวแปรต้นมี โดยมีขั้นตอนดังนี้

ขั้นที่ 1 นำผลการพยากรณ์ของทั้ง 3 วิธีและค่าที่เป็นจริงมาพล็อตกราฟ

ขั้นที่ 2 คำนวณความคลาดเคลื่อนในการพยากรณ์ ด้วยค่าวัดความคลาดเคลื่อนทั้ง 6 ค่า ได้แก่ RMSE , MAPE , GMRAE , MdAPE , MdRAE และ Percent Better โดยมีสูตรในการคำนวณในบทที่ 2 โดยใช้โปรแกรม Microsoft Office 97 (EXCEL) ถ้าวิธีใดมีค่า RMSE, MAPE, GMRAE, MdAPE และ MdRAE น้อยที่สุด และมีค่า Percent Better มากที่สุด แสดงว่าวิธีนั้นเป็นวิธีที่มีความคลาดเคลื่อนในการพยากรณ์น้อยที่สุด