

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในการวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองเพื่อเปรียบเทียบตัวประมาณวิถีกำลังสองน้อยที่สุด ตัวประมาณริจรีเกรสชัน และ ตัวประมาณลาเท็นรูทรีเกรสชัน เมื่อเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระในระดับต่าง ๆ ทั้งนี้ตัวประมาณที่ใช้ในการวิจัย คือ ตัวประมาณวิถีกำลังสองน้อยที่สุด (OLS) ตัวประมาณริจรีเกรสชัน โดยวิธีของ Hoerl , Kennard and Baldwin (HKB) วิธีของ Lawless and Wang (LW) และ ตัวประมาณลาเท็นรูทรีเกรสชัน โดยทำการทดลองภายใต้สถานการณ์ต่าง ๆ ดังนี้

1. เมื่อการแจกแจงของความคลาดเคลื่อนของข้อมูลมีการแจกแจงแบบปกติ แบบปกติปลอมปน และ แบบลอกนอร์มอล
2. เมื่อขนาดตัวอย่างเท่ากับ 10 , 30 และ 50
3. เมื่อจำนวนตัวแปรอิสระเท่ากับ 3 และ 5
4. เมื่อระดับความสัมพันธ์เท่ากับ $[0.11-0.20]$, $[0.21-0.30]$, ... , $[0.91-1.00]$

สำหรับการวิจัยครั้งนี้ใช้การจำลองข้อมูลด้วยเทคนิคของมัลติโนมอลทั้งนี้อาศัยโปรแกรมฟอร์แทน 77 ในการสร้างข้อมูลและกำหนดสถานการณ์ต่าง ๆ ตามที่กำหนดไว้ข้างต้น โดยทำการทดลองซ้ำ 1,000 รอบ ในแต่ละสถานการณ์

5.1 สรุปผลการวิจัย

ในการพิจารณาว่า ตัวประมาณตัวใดมีความเหมาะสมในการประมาณค่าสัมประสิทธิ์ความถดถอยพหุ จะพิจารณาจากค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (AMSE) ซึ่งหากตัวประมาณตัวใดมีค่า AMSE ต่ำ แสดงว่าตัวประมาณตัวนั้นสามารถประมาณค่าสัมประสิทธิ์ความถดถอยพหุได้ดี ใน

การสรุปผลการวิจัย จะทำการสรุปโดยแยกตามลักษณะการแจกแจงของข้อมูลที่จำลองขึ้น ซึ่งสามารถสรุปได้ดังต่อไปนี้

5.1.1 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ (โดยที่ค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ 1.00)

กรณีจำนวนตัวแปรอิสระเท่ากับ 3 พบว่าตัวประมาณที่ดีที่สุด คือ ตัวประมาณ LR โดยให้ค่า AMSE ต่ำที่สุดในทุกกรณีของระดับความสัมพันธ์และขนาดตัวอย่างที่ทำการทดลอง รองลงมา คือ ตัวประมาณ HKB , LW และ OLS ตามลำดับ ยกเว้น ในกรณีที่ขนาดตัวอย่างมีขนาดเล็ก (เท่ากับ 10) และระดับความสัมพันธ์มีค่าสูง (มากกว่า 0.9) ตัวประมาณ HKB จะเป็นตัวประมาณที่ดีที่สุด

กรณีจำนวนตัวแปรอิสระเท่ากับ 5 พบว่า โดยทั่วไป ตัวประมาณ LR ให้ค่า AMSE ต่ำที่สุด รองลงมา คือ ตัวประมาณ HKB , LW และ OLS ตามลำดับ อย่างไรก็ตาม เมื่อขนาดตัวอย่างเล็ก (เท่ากับ 10) และเมื่อระดับความสัมพันธ์มีแนวโน้มเพิ่มสูงขึ้น ตัวประมาณ HKB จะมีแนวโน้มที่ดีขึ้น และจะให้ค่า AMSE ต่ำกว่าตัวประมาณ LR

สำหรับกรณีที่ค่าความแปรปรวนเปลี่ยนไป พบว่า เมื่อข้อมูลมีการกระจายน้อย (ค่าความแปรปรวนเท่ากับ 0.30) โดยส่วนใหญ่ผลลัพธ์ที่ได้ยังคงเช่นเดิมทั้งในกรณีที่จำนวนตัวแปรอิสระเท่ากับ 3 และ 5 ส่วนกรณีที่ข้อมูลมีการกระจายมาก (ค่าความแปรปรวนเท่ากับ 3.00) พบว่า ตัวประมาณ HKB จะมีแนวโน้มดีขึ้นและเป็นตัวประมาณที่ดีที่สุด รองลงมา คือ ตัวประมาณ LR , LW และ OLS ตามลำดับ

นอกจากนี้เมื่อระดับความสัมพันธ์มีค่าสูงขึ้น ค่า AMSE ของตัวประมาณ LR , HKB , LW จะมีค่าต่ำกว่าค่า AMSE ของตัวประมาณ OLS มาก และค่า AMSE ของตัวประมาณทั้ง 3 ตัว จะเข้าใกล้ตัวประมาณ OLS มากขึ้น เมื่อตัวแปรอิสระเกิดพหุสัมพันธ์กันน้อยลง

5.1.2 เมื่อความคลาดเคลื่อนข้อมูลมีการแจกแจงแบบปกติปลอมปน (โดยค่า C เท่ากับ 3 , 5 และค่า P เท่ากับ 5% และ 10%)

เมื่อกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 และกำหนดค่าสเกลแฟคเตอร์น้อย

(C เท่ากับ 3) พบว่า โดยส่วนใหญ่ ตัวประมาณ LR เป็นตัวประมาณที่ดีที่สุด รองลงมา คือ ตัวประมาณ HKB , LW และ OLS ตามลำดับ โดยที่ตัวประมาณ HKB และ LW ให้ค่า AMSE ใกล้เคียงกัน แต่ในกรณีที่ขนาดตัวอย่างเล็ก (เท่ากับ 10) และระดับความสัมพันธ์มีแนวโน้มสูงขึ้น ตัวประมาณ HKB มีแนวโน้มดีขึ้น รองลงมา คือ ตัวประมาณ LR , LW และ OLS ตามลำดับ สำหรับกรณีที่ค่า C สูงขึ้น (C เท่ากับ 10) พบว่าเมื่อเปอร์เซ็นต์ของการปลอมปนมีค่าน้อย (P เท่ากับ 5%) ตัวประมาณ LR ยังคงเป็นตัวประมาณที่ดีที่สุดในทุกกรณี ยกเว้น เมื่อระดับความสัมพันธ์มีค่าสูงขึ้นและขนาดตัวอย่างมีขนาดเล็กหรือปานกลาง (เท่ากับ 10 หรือ 30) ตามลำดับ ตัวประมาณ HKB จะให้ค่า AMSE ที่ต่ำกว่าตัวประมาณ LR ส่วนกรณีที่เปอร์เซ็นต์การปลอมปนมีค่าสูง (P เท่ากับ 10%) โดยส่วนใหญ่ตัวประมาณ HKB จะเป็นตัวประมาณที่ดีที่สุดในทุกกรณี รองลงมาคือ ตัวประมาณ LW , OLS และ LR ตามลำดับ นอกจากนี้เมื่อพิจารณาค่า AMSE ของตัวประมาณ พบว่ามีแนวโน้มเพิ่มสูงขึ้น เมื่อค่า C และ P สูงขึ้น

กรณีจำนวนตัวแปรอิสระเท่ากับ 5 พบว่า โดยส่วนใหญ่ ผลสรุปที่ได้ใกล้เคียงกับจำนวนตัวแปรอิสระเท่ากับ 3 นั่นคือ ตัวประมาณ LR เป็นตัวประมาณที่ดีที่สุด รองลงมา คือ ตัวประมาณ HKB , LW และ OLS ตามลำดับ ยกเว้นเมื่อ C และ P มีค่าสูง (C เท่ากับ 10 และ P เท่ากับ 10% ตามลำดับ) ตัวประมาณ HKB เป็นตัวประมาณที่ดีที่สุด รองลงมา คือ ตัวประมาณ LW , OLS และ LR ตามลำดับ

5.1.3 เมื่อความคลาดเคลื่อนมีการแจกแจงแบบลอกนอร์มอล (โดยที่ค่าเฉลี่ย μ เท่ากับ 0 และค่าความแปรปรวนเท่ากับ 1.00)

พบว่า โดยส่วนใหญ่เมื่อกำหนดจำนวนตัวแปรอิสระเท่ากับ 3 และ 5 จะให้ผลสรุปที่ใกล้เคียงกัน กล่าวคือ ตัวประมาณ LR เป็นตัวประมาณที่ดีที่สุดในทุกกรณีของระดับความสัมพันธ์และขนาดตัวอย่างที่ทำการวิจัย รองลงมา คือ ตัวประมาณ HKB , LW และ OLS ตามลำดับ ทั้งนี้เมื่อพิจารณาค่า AMSE ที่ได้ จะเห็นว่า ในกรณีจำนวนตัวแปรอิสระเท่ากับ 5 และเมื่อระดับความสัมพันธ์มีแนวโน้มสูงขึ้น โดยที่ขนาดตัวอย่างมีขนาดเล็ก ค่า AMSE ของตัวประมาณ OLS จะมีค่าสูงมากขึ้น และการเพิ่มจำนวนตัวแปรอิสระมีผลทำให้ค่า AMSE สูงขึ้นด้วย นอกจากนี้ผู้วิจัยได้ทำการวิจัยเพิ่มเติม โดยการเปลี่ยนค่าความแปรปรวนของข้อมูล พบว่า ค่าความแปรปรวนไม่มีผลต่อตัวประมาณ แต่มีผลทำให้ค่า AMSE สูงขึ้น เมื่อกำหนดค่าความแปรปรวนเพิ่มขึ้น

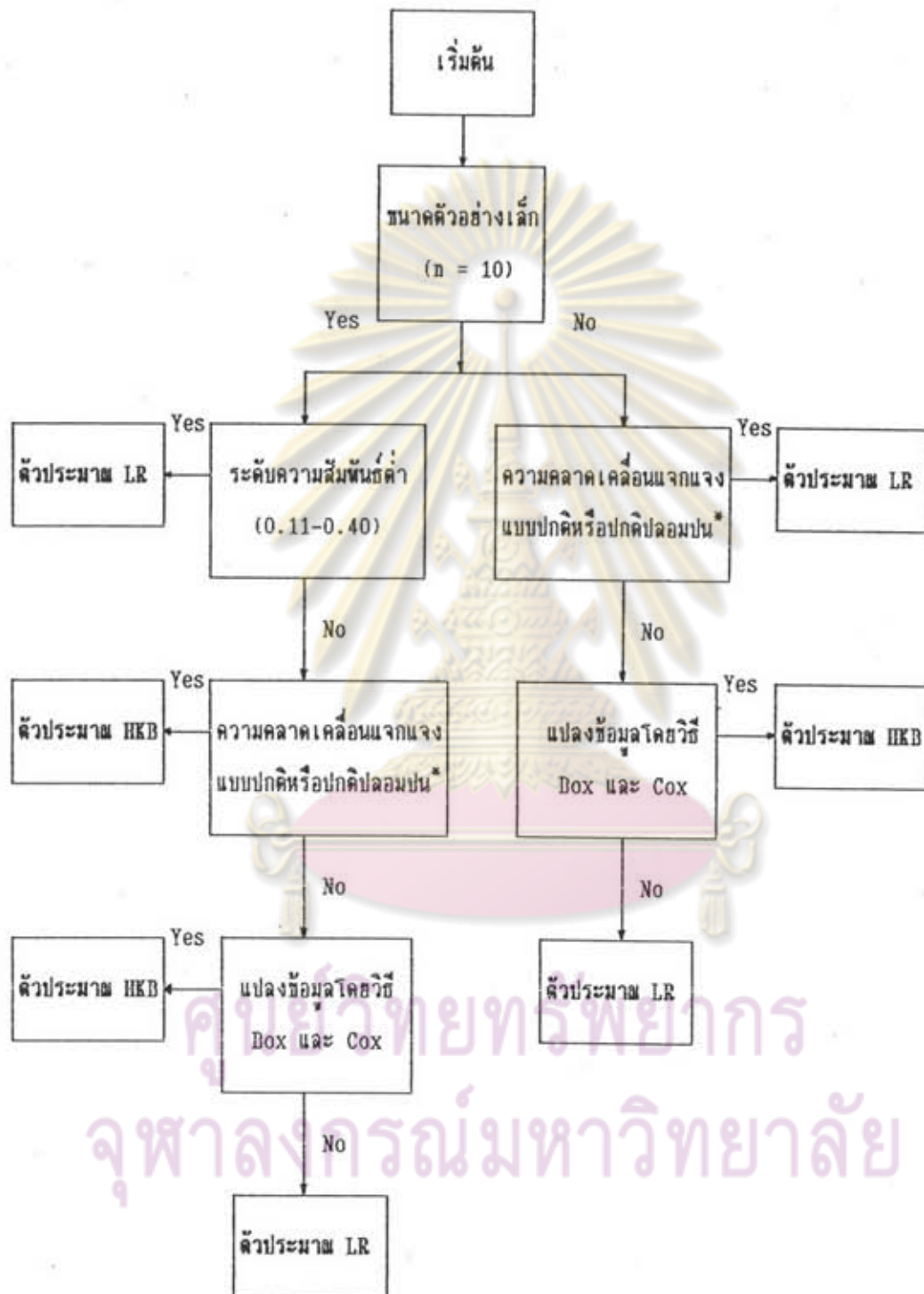
นอกเหนือจากนี้ ผู้วิจัยได้ทำการแปลงข้อมูลให้มีการแจกแจงเป็นแบบปกติ โดยอาศัยวิธีการของ Box และ Cox (ดูภาคผนวก ก.) พบว่า โดยส่วนใหญ่ ตัวประมาณ HKB ให้ค่า AMSE ต่ำที่สุดในทุกกรณีของขนาดตัวอย่าง ระดับความสัมพันธ์ และ จำนวนตัวแปรอิสระ ที่ทำการวิจัย รองลงมา คือ ตัวประมาณ LR , LW และ OLS ตามลำดับ โดยค่า AMSE ของตัวประมาณทั้ง 4 วิธี จะมีค่าต่ำกว่ากรณีที่ไม่ได้ทำการแปลงข้อมูล (ดูตารางที่ 1.21 - 1.26 ในภาคผนวก ข.)

5.2 ข้อเสนอแนะ

ในด้านการวิจัยเพิ่มเติม สามารถทำการศึกษาเพิ่มเติมได้ดังนี้

1. ทำการศึกษาเพิ่มเติมในกรณีที่ขนาดตัวอย่างมากกว่า 50
2. ทำการศึกษาเพิ่มเติมเมื่อจำนวนตัวแปรอิสระมากกว่า 5
3. ทำการศึกษาเพิ่มเติมเมื่อความคลาดเคลื่อนของข้อมูลมีการแจกแจงในรูปแบบอื่น ๆ เช่น การแจกแจงแบบไวบูลล์ การแจกแจงแบบเอกซ์โปเนนเชียล
4. ทำการศึกษาเพิ่มเติมเมื่อข้อมูลมีค่าความแปรปรวนอื่น ๆ นอกเหนือจากที่ได้ทำการวิจัย
5. ทำการศึกษาเปรียบเทียบตัวประมาณความถดถอยวิธีอื่น ๆ เช่น Regression Principle Component หรือ วิธีรีดจ์รีเกรสชั่น โดยกำหนดวิธีหาค่า K โดยวิธีอื่น ๆ เช่น วิธี Dempsfer , Scbatroff and Wermuth's (D.S.W.) , วิธี Kasarda and Sbibi's (K.S.) (Panos Panopoulos , 1989)
6. ก่อนทำการศึกษากการวิเคราะห์ความถดถอยหาค่าการตรวจสอบก่อนว่าเกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระหรือไม่ โดยดูจากค่าพหุสัมพันธ์ ซึ่งสามารถหาได้โดยวิธีของ Marquatt วิธีของ Chaterjee and Price (ดวงใจ ชูรักษ์ , 2529) เพื่อให้สามารถเลือกใช้ตัวประมาณที่เหมาะสมได้ถูกต้อง
7. ในกาวิจัยครั้งนี้ ผู้วิจัยทำการสรุปผลการวิจัย โดยพิจารณาจากการแจกแจงของความคลาดเคลื่อนเป็นหลัก ซึ่งในทางปฏิบัติอาจพิจารณาจากปัจจัยอื่น ๆ เช่น ขนาดตัวอย่าง ระดับความสัมพันธ์ หรือ จำนวนตัวแปรอิสระ เป็นต้น

แผนผังแสดงการเลือกให้ตัวประมาณสัมประสิทธิ์ความถดถอย



หมายเหตุ * กรณีความคลาดเคลื่อนแจกแจงแบบปกติปลอมปน โดยที่สเกลแฟคเตอร์ เท่ากับ 10 และ เปอร์เซ็นต์การปลอมปนเท่ากับ 10% ตัวประมาณ HKB เป็นตัวประมาณที่เหมาะสมที่สุดในทุกกรณี