

## บทที่ 5

### สรุปผลการวิจัยและข้อเสนอแนะ

ในการศึกษาเปรียบเทียบการประมาณค่าของตัวแปรตามในสมการถดถอยเชิงเส้นพหุ โดยทำการประมาณค่าพารามิเตอร์ด้วยวิธีการ 3 วิธีการคือ วิธีการสมิท วิธีการประมาณด้วยภาชนะ น่าจะเป็นสูงสุด และวิธีการโมดิไฟด์แอกซ์เรียล ซึ่งข้อมูลที่นำมาวิเคราะห์นั้นค่าสังเกตของตัวแปรตามเป็นข้อมูลที่ถูกตัดทิ้งทางขวาและเป็นการตัดทิ้งประเภทที่ 1 โดยใช้วิธีจำลองแบบมอนติคาร์โล ด้วยภาษาฟอร์แทรน (FORTRAN) จำลองข้อมูลให้มีลักษณะตามแผนการทดลองที่กำหนดและทำการทดลองซ้ำ ๆ กัน 800 ครั้งในแต่ละสถานการณ์ที่ทำการศึกษา สามารถสรุปผลการวิจัยได้ดังนี้

#### 5.1 สรุปผลการวิจัย

1. วิธีการประมาณด้วยภาชนะน่าจะเป็นสูงสุด เป็นวิธีการที่ให้ค่า RMSE ของการประมาณค่าตัวแปรตามสูงกว่าวิธีการอื่น ๆ เมื่อสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.1 , 0.2 , 0.3 ทุกขนาดตัวอย่างและจำนวนตัวแปรที่ทำการศึกษา ยกเว้นกรณีที่ตัวแปรอิสระเท่ากับ 7 ในสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.2 และ 0.3 ขนาดตัวอย่างเท่ากับ 10 พบว่าวิธีการนี้จะให้ค่า RMSE ต่ำกว่าวิธีการอื่น ๆ

เมื่อสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.4 , 0.5 พบว่าวิธีการประมาณด้วยภาชนะน่าจะเป็นสูงสุด เป็นวิธีการที่ให้ค่า RMSE ของการประมาณค่าตัวแปรตามต่ำกว่าวิธีการอื่น ๆ ทุกขนาดตัวอย่างและจำนวนตัวแปรที่ทำการศึกษา

2. วิธีการสมิท เป็นวิธีการที่ให้ค่า RMSE ของการประมาณค่าตัวแปรตามต่ำกว่าวิธีการอื่น ๆ เมื่อสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.1 , 0.2 , 0.3 ทุกขนาดตัวอย่างที่ทำการศึกษา ยกเว้นกรณีที่ขนาดตัวอย่างเท่ากับ 10 ตัวแปรอิสระเท่ากับ 5 และ 7

เมื่อสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.4 , 0.5 พบว่าวิธีการสมิท เป็นวิธีการที่ให้ค่า RMSE ของการประมาณค่าตัวแปรตามต่ำกว่าวิธีการโมดิไฟด์แอกซ์เรียล และสูงกว่าวิธีการประมาณด้วยภาชนะน่าจะเป็นสูงสุด

3. วิธีการโมดิไฟด์แอกซ์เรียล เป็นวิธีการที่ให้ค่า RMSE ใกล้เคียงกับวิธีการสมิททุกสถานการณ์ที่ทำการศึกษา

#### 4. ปัจจัยที่มีผลกระทบต่อค่า RMSE มีดังนี้

4.1 กรณีที่ขนาดตัวอย่างคงที่ ค่า RMSE ของทั้ง 3 วิธีการจะแปรผันตามสัดส่วนของข้อมูลที่ถูกตัดทิ้งนั่นคือเมื่อสัดส่วนของข้อมูลที่ถูกตัดทิ้งเพิ่มขึ้น ค่า RMSE ก็จะเพิ่มขึ้นด้วย

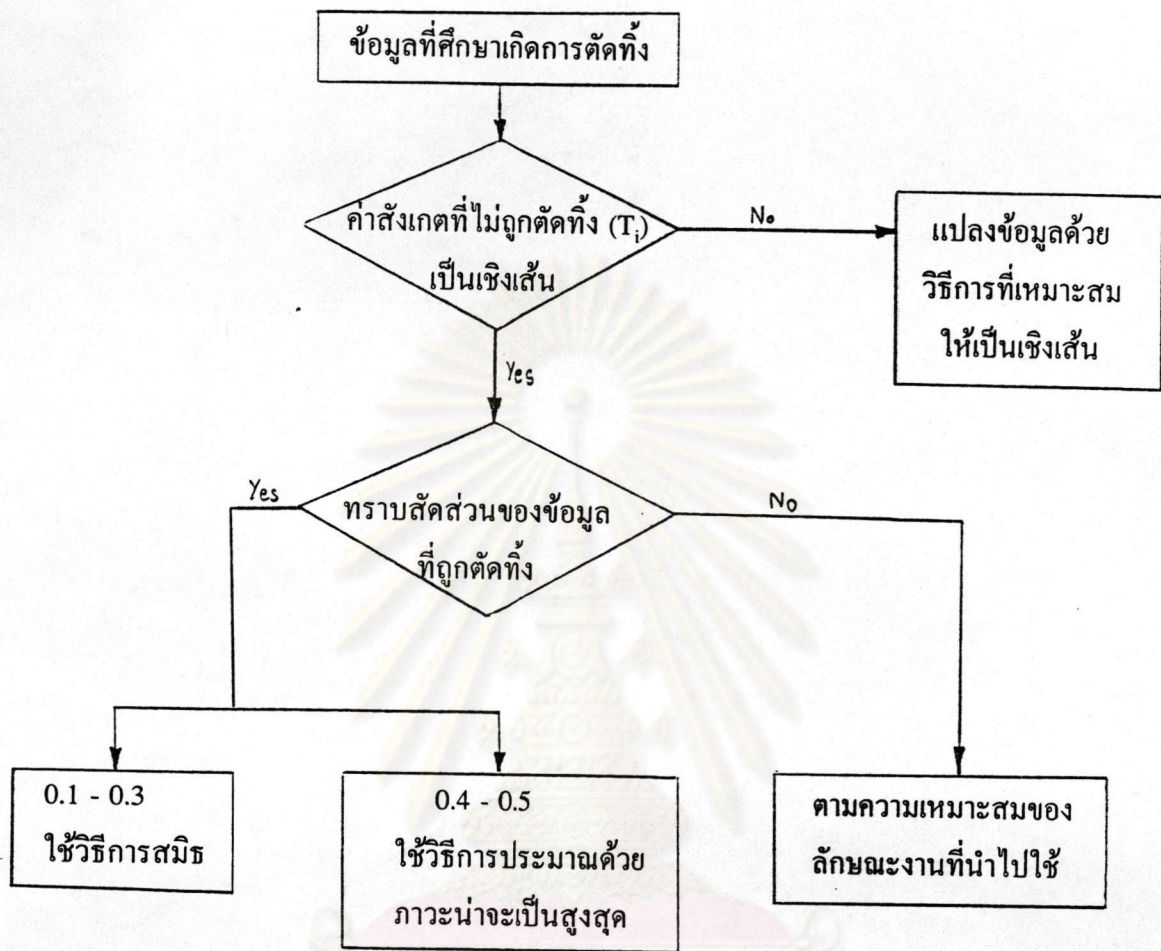
4.2 กรณีที่สัดส่วนของข้อมูลที่ถูกตัดทิ้งคงที่ ค่า RMSE ของทั้ง 3 วิธีการจะแปรผกผันตามขนาดตัวอย่าง นั่นคือ เมื่อขนาดตัวอย่างเพิ่มมากขึ้นค่า RMSE ก็จะมีค่าลดลง

กรณีที่ต้องการประมาณค่าพารามิเตอร์ในสมการถดถอยเชิงเส้นพหุ เมื่อตัวแปรตามบางค่าเป็นค่าที่ถูกตัดทิ้งทางขวาและเป็นการตัดทิ้งประเภทที่ 1 เมื่อสัดส่วนของข้อมูลที่ถูกตัดทิ้งเป็น 0.1 , 0.2 , 0.3 ควรเลือกใช้วิธีการสมิธ กรณีที่สัดส่วนของข้อมูลที่ถูกตัดทิ้งมีค่ามากขึ้น วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดเป็นวิธีการที่ควรนำมาใช้ ซึ่งสามารถสรุปผลได้ดังตารางที่ 5.1 ดังนี้

ขนาดตัวอย่าง	จำนวนตัวแปรอิสระ	สัดส่วนของข้อมูลที่ถูกตัดทิ้ง				
		0.1	0.2	0.3	0.4	0.5
10	x=3	SM	SM	SM	ML	ML
	x=5	AC	SM	SM	ML	ML
	x=7	SM	ML	ML	ML	ML
20	x=3	SM	SM	SM	ML	ML
	x=5	SM	SM	SM	ML	ML
	x=7	SM	SM	SM	ML	ML
30	x=3	SM	SM	SM	ML	ML
	x=5	SM	SM	SM	ML	ML
	x=7	SM	SM	SM	ML	ML
50	x=3	SM	SM	SM	ML	ML
	x=5	SM	SM	SM	ML	ML
	x=7	SM	SM	SM	ML	ML
100	x=3	SM	SM	SM	ML	ML
	x=5	SM	SM	SM	ML	ML
	x=7	SM	SM	SM	ML	ML

ตารางที่ 5.1 สรุปผลการศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ทั้ง 3 วิธีการ

หรือสามารถสรุปเป็นผังงานการนำไปใช้งานได้ดังนี้



รูปที่ 5.2 แสดงผังงานการนำไปใช้งาน

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

## 5.2 ข้อเสนอแนะ

1. ในการศึกษาครั้งนี้เป็นการศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีการ 3 วิธี โดยในการจำลองข้อมูลนั้นให้ค่าความคลาดเคลื่อน ( $\epsilon_i$ ) มีการแจกแจงแบบปกติ เนื่องจากวิธีการของสมิธนั้นสามารถใช้ศึกษาข้อมูลที่มีความคลาดเคลื่อนมีการแจกแจงเป็นรูปแบบใดก็ได้ ไม่มีรูปแบบเฉพาะ ไม่เป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอย ดังนั้นในการศึกษาครั้งต่อไปควรจะศึกษาในกรณีของค่าความคลาดเคลื่อนมีการแจกแจงในรูปแบบอื่น

2. เมื่อข้อมูลเกิดเป็นข้อมูลที่ถูกต้องทั้ง นอกจากจะส่งผลในการวิเคราะห์ข้อมูลแล้ว ยังเกิดปัญหาในด้านอื่น ๆ อีก เช่น การทดสอบสมมติฐานเกี่ยวกับ 2 ประชากร ซึ่งมีวิธีที่ใช้ในการแก้ปัญหาคือวิธีการของ Gehan Test, Mantel-Hacnszel test, Tarone-Ware Class of tests, Efron test น่าจะเป็นแนวทางในการศึกษาค้นคว้าต่อไป

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย