



## บทที่ 5

### สรุปผลการวิจัย และข้อเสนอแนะ

การวิจัยในครั้งนี้ มีวัตถุประสงค์เพื่อต้องการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณค่าพารามิเตอร์ในสมการถดถอยเชิงเส้นพหุคูณ เมื่อตัวแปรอิสระมีพหุสัมพันธ์และ/หรือตัวแปรตามมีค่าผิดปกติ โดยการศึกษาเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณการถดถอย 5 ตัว ดังนี้ ตัวประมาณกำลังสองน้อยสุด ตัวประมาณค่าสัมบูรณ์น้อยสุด ตัวประมาณริตจ ตัวประมาณริตจที่มีค่าสัมบูรณ์น้อยสุด และตัวประมาณริตจที่มีการถ่วงน้ำหนัก

วิธีการดำเนินการวิจัยครั้งนี้ ใช้วิธีการจำลองด้วยเทคนิคการจำลองมอนติคาร์โล และทำการเขียนโปรแกรมโดยใช้โปรแกรมภาษาฟอร์แทรน เพื่อสร้างข้อมูลตามสถานการณ์ต่าง ๆ ที่กำหนดขึ้นมา โดยกำหนดให้ในแต่ละสถานการณ์ทำการจำลองจำนวน 1,000 ครั้ง

#### 5.1 สรุปผลการวิจัย

จากการทดลองเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณการถดถอยทั้ง 5 ตัว ได้ข้อสรุปดังนี้

##### 1. กรณีที่ตัวแปรอิสระมีพหุสัมพันธ์

###### 1.1 เมื่อสหสัมพันธ์ในตัวแปรอิสระมีค่าน้อยกว่าหรือเท่ากับ 0.5

ในทุกระดับขนาดตัวอย่าง (20 ถึง 60) พบว่า ตัวประมาณ LS จะให้ค่า RMSE ต่ำที่สุด และตัวประมาณ WRID จะให้ค่า RMSE สูงที่สุด ยกเว้นในบางกรณี ได้แก่ กรณีระดับสหสัมพันธ์เท่ากับ 0.5 และขนาดตัวอย่างเท่ากับ 20 และ 30 พบว่า ตัวประมาณ RID จะให้ค่า RMSE ต่ำที่สุด

###### 1.2 เมื่อสหสัมพันธ์ในตัวแปรอิสระมีค่ามากกว่า 0.5

ในทุกระดับขนาดตัวอย่าง (20 ถึง 60) พบว่า ตัวประมาณ RID จะให้ค่า RMSE ต่ำที่สุด และตัวประมาณ LAV จะให้ค่า RMSE สูงที่สุด ยกเว้นในบาง

กรณี ได้แก่ กรณีขนาดตัวอย่างเท่ากับ 20 , 30 และ 35 และระดับสหสัมพันธ์เท่ากับ 0.9 และ 0.95 พบว่า ตัวประมาณ WRID จะให้ค่า RMSE สูงที่สุด

## 2. กรณีที่ตัวแปรตามมีค่าผิดปกติ

ในทุกกรณีขนาดตัวอย่าง( 20 ถึง 60 ) ทุกระดับค่าผิดปกติของตัวแปรตาม ( เล็กน้อย ปานกลาง และรุนแรง ) และทุกระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน ( 0.05 , 0.08 , 0.10 และ 0.15 ) พบว่า ตัวประมาณ LAV จะให้ค่า RMSE ต่ำที่สุด ในขณะที่ตัวประมาณ RLAV , RID , LS และ WRID จะให้ค่า RMSE สูงขึ้น ตามลำดับ

## 3. กรณีที่ตัวแปรอิสระมีพหุสัมพันธ์และตัวแปรตามมีค่าผิดปกติ

### 3.1 เมื่อสหสัมพันธ์ในตัวแปรอิสระมีค่าน้อยกว่าหรือเท่ากับ 0.5

ในทุกกรณีขนาดตัวอย่าง( 20 ถึง 60 ) ทุกระดับค่าผิดปกติของตัวแปรตาม ( เล็กน้อย ปานกลาง และรุนแรง ) และทุกระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน ( 0.05 , 0.08 , 0.10 และ 0.15 ) พบว่า ตัวประมาณ LAV จะให้ค่า RMSE ต่ำที่สุด ในขณะที่ตัวประมาณ RLAV , RID , LS และ WRID จะให้ค่า RMSE สูงขึ้น ตามลำดับ ยกเว้นในบางกรณี ได้แก่ กรณีระดับสหสัมพันธ์เท่ากับ 0.5 และขนาดตัวอย่างเท่ากับ 20 , 30 และ 35 พบว่า ตัวประมาณ RLAV จะให้ค่า RMSE ต่ำที่สุด

### 3.2 เมื่อสหสัมพันธ์ในตัวแปรอิสระมีค่ามากกว่า 0.5

ในทุกกรณีขนาดตัวอย่าง( 20 ถึง 60 ) ทุกระดับค่าผิดปกติของตัวแปรตาม ( เล็กน้อย ปานกลาง และรุนแรง ) และทุกระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน ( 0.05 , 0.08 , 0.10 และ 0.15 ) พบว่า ตัวประมาณ RLAV จะให้ค่า RMSE ต่ำที่สุด และตัวประมาณ LS จะให้ค่า RMSE สูงที่สุด ยกเว้นในบางกรณี ได้แก่ กรณีระดับสัดส่วนการปลอมปนของความคลาดเคลื่อนเท่ากับ 0.05 และระดับสหสัมพันธ์เท่ากับ 0.7 พบว่า ตัวประมาณ WRID จะให้ค่า RMSE สูงที่สุด

4. ผลกระทบจากปัจจัยต่าง ๆ ที่มีต่อค่า RMSE ที่ได้จากการประมาณค่าสัมประสิทธิ์การถดถอยด้วยตัวประมาณพารามิเตอร์ทั้ง 5 ตัว คือ LS , LAV , RID , RLAV และ WRID พบว่า

4.1 ค่า RMSE จะแปรผันตามกับระดับความรุนแรงของค่าผิดปกติในตัวแปรตาม เมื่อกำหนดให้ปัจจัยอื่น ๆ คงที่ ( ระดับขนาดตัวอย่าง ระดับสหสัมพันธ์ในตัวแปรอิสระ และระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน )

4.2 ค่า RMSE จะแปรผันตามกับระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน เมื่อกำหนดให้ปัจจัยอื่น ๆ คงที่ ( ระดับขนาดตัวอย่าง ระดับสหสัมพันธ์ในตัวแปรอิสระ และระดับค่าผิดปกติของตัวแปรตาม )

4.3 ค่า RMSE จะแปรผันตามกับระดับสหสัมพันธ์ในตัวแปรอิสระ เมื่อกำหนดให้ปัจจัยอื่น ๆ คงที่ ( ระดับขนาดตัวอย่าง ระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน และระดับค่าผิดปกติของตัวแปรตาม )

4.4 ค่า RMSE จะแปรผกผันกับระดับขนาดตัวอย่าง เมื่อกำหนดให้ปัจจัยอื่น ๆ คงที่ ( ระดับสหสัมพันธ์ในตัวแปรอิสระ ระดับสัดส่วนการปลอมปนของความคลาดเคลื่อน และระดับค่าผิดปกติของตัวแปรตาม )

## 5.2 ข้อเสนอแนะ

ผลการวิจัยในครั้งนี้ มีข้อเสนอแนะแบ่งเป็น 2 ด้าน คือ

1. ด้านการนำไปใช้ประโยชน์ ( เมื่อตัวแปรอิสระมีพหุสัมพันธ์และ/หรือตัวแปรตามมีค่าผิดปกติ )

### 1.1 การตรวจสอบข้อมูล

เริ่มด้วยการตรวจสอบว่า ลักษณะของข้อมูลนั้น ตัวแปรอิสระมีพหุสัมพันธ์และ/หรือตัวแปรตามมีค่าผิดปกติหรือไม่ ต้องแยกตรวจสอบเป็น 2 กรณี คือ การตรวจสอบการมีค่าผิดปกติในตัวแปรตาม และการตรวจสอบการมีพหุสัมพันธ์ในตัวแปรอิสระ โดยมีวิธีดำเนินการตรวจสอบ ดังต่อไปนี้

#### 1.1.1 การตรวจสอบค่าผิดปกติในตัวแปรตาม

การตรวจสอบค่าผิดปกติในตัวแปรตาม เราอาจพิจารณาจากแผนภาพการกระจาย ( Scatter Diagram ) ของข้อมูลของตัวแปรตาม ว่ามีค่าสังเกตที่มีค่าสูงหรือต่ำกว่าค่า

สังเกตส่วนใหญ่ในข้อมูลชุดเดียวกันหรือไม่ แต่วิธีที่นิยมใช้กัน คือ ใช้เกณฑ์การหาค่าผิดปกติ ด้วย Box Plot ซึ่งได้แสดงรายละเอียดไว้ในภาคผนวก ก. (หน้า 306)

### 1.1.2 การตรวจสอบพหุสัมพันธ์ในตัวแปรอิสระ

การตรวจสอบว่าตัวแปรอิสระมีความสัมพันธ์กันหรือไม่ เราอาจพิจารณาจาก แผนภาพการกระจายของตัวแปรอิสระแต่ละคู่ โดยพิจารณาควบคู่ไปกับค่าสหสัมพันธ์ (Correlation) ของตัวแปรอิสระแต่ละคู่ และนอกจากนี้ เราอาจพิจารณาจากค่า VIF (Variance Inflation Factor) ซึ่งได้แสดงรายละเอียดไว้ในภาคผนวก ข. (หน้า 374)

### 1.2 การเลือกใช้ตัวประมาณ

ถ้าพบว่าตัวแปรอิสระมีพหุสัมพันธ์และตัวแปรตามมีค่าผิดปกติเกิดขึ้นพร้อมกัน โดยที่

- มีสหสัมพันธ์ในระดับต่ำถึงปานกลาง ( 0.1 ถึง 0.5 ) ในทุกระดับขนาดตัวอย่าง ( 20 ถึง 60 ) ควรใช้ตัวประมาณค่าสัมบูรณ์น้อยสุดในการประมาณค่าพารามิเตอร์
- มีสหสัมพันธ์ในระดับสูง ( มากกว่า 0.5 ถึง 0.99 ) ในทุกระดับขนาดตัวอย่าง ( 20 ถึง 60 ) ควรใช้ตัวประมาณการถดถอยวิธีที่มีค่าสัมบูรณ์ต่ำสุดในการประมาณค่าพารามิเตอร์

ถ้าพบว่าตัวแปรอิสระมีพหุสัมพันธ์เพียงอย่างเดียว โดยที่

- มีสหสัมพันธ์ในระดับต่ำ ( 0.1 ถึง 0.5 ) ในทุกระดับขนาดตัวอย่าง ( 20 ถึง 60 ) ควรใช้ตัวประมาณกำลังสองน้อยสุดในการประมาณค่าพารามิเตอร์
- มีสหสัมพันธ์ในระดับสูง ( มากกว่า 0.5 ถึง 0.99 ) ในทุกระดับขนาดตัวอย่าง ( 20 ถึง 60 ) ควรใช้ตัวประมาณการถดถอยวิธีในการประมาณค่าพารามิเตอร์

ถ้าพบว่าตัวแปรตามมีค่าผิดปกติเพียงอย่างเดียว ในทุกระดับขนาดตัวอย่าง ( 20 ถึง 60 ) ควรใช้ตัวประมาณค่าสัมบูรณ์น้อยสุดในการประมาณค่าพารามิเตอร์ ในทุกสถานการณ์

## 2. ด้านการวิจัย

2.1 ควรทำการศึกษาเพื่อปรับปรุงตัวประมาณจุดที่มีการถ่วงน้ำหนัก ซึ่งควร จะทำการทดลองโดยใช้ค่า  $W$  หรือค่า  $k$  ในรูปแบบอื่น ๆ ที่น่าสนใจต่อไป ที่จะทำให้ตัว ประมาณมีประสิทธิภาพดีขึ้นได้ เช่น ทดลองใช้ค่า  $k^*$  โดยอิงพื้นฐานจากตัวประมาณ LAV เป็นต้น

2.2 ควรทำการศึกษาโดยใช้ตัวประมาณที่น่าสนใจอื่น ๆ เช่น ตัวประมาณที่ สร้างจากวิธีการหาค่าประกอบหลัก (Principal Components) หรือ ตัวประมาณภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator) เป็นต้น

2.3 ควรทำการศึกษาในกรณีตัวแปรอิสระมีความสัมพันธ์กันในรูปแบบอื่น ๆ เช่น ตัวแปรอิสระตัวหนึ่งเป็นผลรวมเชิงเส้น (Linear Combination) ของตัวแปรอิสระตัวอื่น ๆ

2.4 ควรทำการศึกษาโดยการหาค่า VIF เป็นค่าวัดระดับความสัมพันธ์ในตัวแปร อิสระ แทนการหาค่าสัมประสิทธิ์สหสัมพันธ์