

### สรุปผลการวิจัยและข้อเสนอแนะ

ในการวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบการประมาณค่าของตัวแปรตาม ในสมการถดถอยเชิงเส้นอย่างง่าย เมื่อศึกษาวิเคราะห์กับข้อมูลที่ค่าสังเกตของตัวแปรตามเป็นค่าที่ถูกตัดทิ้งทางขวา โดยการประมาณค่าพารามิเตอร์ ด้วยวิธีการ 3 วิธี คือ วิธีกำลังสองต่ำสุด วิธีประมาณด้วยภาวะน่าจะเป็นสูงสุด และวิธีการของบัคเลย์และเจมส์

วิธีดำเนินการวิจัยครั้งนี้ ใช้วิธีการจำลองแบบมอนติคาร์โล ทำงานด้วยเครื่องคอมพิวเตอร์ AMDAHL 5860 โดยใช้ภาษาฟอร์แทรน (FORTRAN) จำลองข้อมูลให้มีลักษณะตามแผนการทดลองที่กำหนด และกำหนดให้เครื่องคอมพิวเตอร์ ทำการทดลองซ้ำ ๆ กัน 1,000 ครั้ง ในแต่ละสถานการณ์ที่สนใจศึกษา

#### 5.1 สรุปผลการวิจัย

การวิจัย จะให้ผลสรุปของความสามารถการประมาณค่าพารามิเตอร์ ในสมการถดถอยเชิงเส้นอย่างง่าย เมื่อค่าสังเกตของตัวแปรตามมีค่าที่ถูกตัดทิ้งทางขวา การเปรียบเทียบความสามารถของแต่ละวิธีการ โดยการศึกษาหาค่าความคลาดเคลื่อนของการประมาณค่าตัวแปรตามกับค่าจริง ทั้ง 3 วิธี เพื่อหาว่าวิธีการใดจะดีที่สุด ในแต่ละสถานการณ์ที่ได้ทำการทดลองสรุปได้ดังนี้

สำหรับขนาดตัวอย่างทุกขนาด ( $NM = 10, 15, 30, 50, 70$ ) สัดส่วนของข้อมูลที่ถูกตัดทิ้งทุกระดับ ( $P = 10\%, 20\%, 30\%, 40\%$ ) และการแจกแจงของค่าที่ถูกตัดทิ้งทุกแบบ คือ แบบสมมาตรแบบปกติตัดปลายทางซ้าย แบบไวบูลล์ตัดปลายทางซ้าย แบบเชิงเส้น ( $U, TN, TW, LI$ )

1. วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด จะให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าตัวแปรตามต่ำกว่าวิธีการของบัคเลย์และเจมส์ในทุกสถานการณ์ที่ทำการศึกษาดังนั้น วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด จึงเป็นวิธีการที่มีประสิทธิภาพในการประมาณค่า

พารามิเตอร์ที่ดีที่สุดใในสามวิธีการที่ศึกษา

2. วิธีการของบัคเลย์และเจมส์ จะให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าตัวแปรตาม กรณีที่เป็นตัวอย่างขนาดเล็ก  $NM = 10, 15, 30$  เมื่อสัดส่วนของข้อมูลที่ถูกต้อง  $P=10\%$  และการแจกแจงของค่าที่ถูกตัดทิ้งทุกแบบ วิธีการของบัคเลย์และเจมส์จะให้ค่า RMSE ใกล้เคียงกับวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด กรณีที่เป็นตัวอย่างขนาดใหญ่  $NM = 50, 70$  เมื่อสัดส่วนของข้อมูลที่ถูกต้อง  $P=10\%, 20\%, 30\%, 40\%$  และการแจกแจงของค่าที่ถูกตัดทิ้งทุกแบบ วิธีการของบัคเลย์และเจมส์ จะให้ค่า RMSE ใกล้เคียงกับวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด ทั้งนี้ วิธีการของบัคเลย์และเจมส์ จะให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าตัวแปรตามสูงกว่าวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด ในทุกสถานการณ์ที่ศึกษา แต่วิธีการของบัคเลย์และเจมส์ จะให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าตัวแปรตามต่ำกว่าวิธีกำลังสองต่ำสุด ในทุกสถานการณ์ที่ทำการศึกษา

3. วิธีกำลังสองต่ำสุดจะให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าตัวแปรตามสูงกว่าวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด และวิธีการของบัคเลย์และเจมส์ ในทุก ๆ สถานการณ์ที่ทำการศึกษา

4. ปัจจัยที่มีผลกระทบต่อค่า RMSE มีดังนี้ กรณีที่ขนาดตัวอย่างคงที่ ค่า RMSE ของวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดจะไม่แปรผันตามสัดส่วนของข้อมูลที่ถูกต้อง แต่วิธีกำลังสองต่ำสุด และวิธีการของบัคเลย์และเจมส์จะแปรผันโดยตรงตามสัดส่วนของข้อมูลที่ถูกต้อง และกรณีขนาดตัวอย่างเพิ่มขึ้นขณะที่สัดส่วนของข้อมูลที่ถูกต้องคงที่ ค่า RMSE ของแต่ละวิธีการจะแปรผัน โดยตรงตามขนาดตัวอย่าง ส่วนการแจกแจงของค่าที่ถูกตัดทิ้ง ไม่มีผลกระทบต่อประสิทธิภาพของแต่ละวิธีการ

## 5.2 ข้อเสนอแนะ

ผลการวิจัยครั้งนี้ จะเสนอแนะเป็น 2 ด้าน คือ

### 5.3.1 ด้านการนำไปใช้ประโยชน์

กรณีที่ต้องการประมาณค่าพารามิเตอร์ ในสมการถดถอยเชิงเส้นอย่างง่าย เมื่อค่าสังเกตของตัวแปรตามเป็นค่าที่ถูกตัดทิ้งทางขวา โดยประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองต่ำสุด วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด และวิธีการของบักเคอร์และเจมสัน มีแนวทางดังนี้

สำหรับข้อมูลที่น่ามาวิเคราะห์ทุกขนาดตัวอย่าง ( $NM = 10, 15, 30, 50, 70$ ) สัดส่วนของค่าที่ถูกตัดทิ้งทุกระดับ ( $P = 10\%, 20\%, 30\%, 40\%$ ) และการแจกแจงของค่าที่ถูกตัดทิ้งทุกแบบ คือ แบบสม่าเสมอ แบบปกติตัดปลายทางซ้าย แบบไวบูลล์ตัดปลายทางซ้าย แบบเชิงเส้น ( $U, TN, TW, LI$ ) ควรเลือกใช้วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด เพราะเป็นวิธีการที่ให้ค่าความคลาดเคลื่อนของการประมาณค่าตัวแปรตามต่ำที่สุด

### 5.3.2 ด้านการศึกษาวิจัย

5.3.2.1 กรณีที่ตัวแปรตามมีค่าถูกตัดปลายทางซ้าย หรือกรณีที่ตัวแปรตามมีค่าถูกตัดปลายทั้งสองด้าน คือ ตัดปลายทางด้านซ้ายและทางด้านขวา วิธีการประมาณค่าพารามิเตอร์ ในสมการถดถอยเชิงเส้นอย่างง่ายเหล่านี้ อาจจะทำให้ผลของการประมาณแตกต่างจากกรณีที่ตัวแปรตามมีค่าถูกตัดปลายทางขวา จึงน่าจะมีการศึกษาวิจัยในปัญหาต่อไป

5.3.2.2 การวิจัยครั้งนี้ เป็นการศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีการ 3 วิธี โดยการจำลองข้อมูลให้มีการแจกแจงของค่าความคลาดเคลื่อน  $\varepsilon_1$  เป็นแบบปกติ เนื่องจากวิธีการของบักเคอร์และเจมสันใช้ศึกษากับข้อมูลที่การแจกแจงของ  $\varepsilon_1$  ไม่มีรูปแบบเฉพาะ หรือสำหรับการแจกแจงของ  $\varepsilon_1$  ที่ไม่เป็นไปตามข้อสมมติของการวิเคราะห์การถดถอย ดังนั้นในการศึกษาครั้งต่อไป ควรจะศึกษากรณีของความคลาดเคลื่อน  $\varepsilon_1$  มีการแจกแจงในรูปแบบอื่น ๆ ที่ไม่เป็นการแจกแจงแบบปกติ

5.3.2.3 ในการวิจัยครั้งนี้ได้ทำการศึกษา เฉพาะกรณีของความถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression) เท่านั้น ดังนั้นในการวิจัยครั้งต่อไปอาจจะทำการศึกษาในกรณีของความถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)

5.3.2.4 ในการวิจัยครั้งนี้ ได้ทำการศึกษาเมื่อค่าที่ถูกตัดทิ้งของตัวแปรตาม ที่มีการแจกแจงแบบสม่ำเสมอ แบบปกติตัดปลายทางซ้าย แบบไวบูลล์ตัดปลายทางซ้าย และกรณีค่าที่ถูกตัดทิ้งเป็นฟังก์ชันเชิงเส้น รูปแบบเดียวกับกับ  $T_1$  คือ  $C_1 = \beta_0 + \beta_1 X_1 + \epsilon_1$  เท่านั้น ดังนั้น ในการศึกษาครั้งต่อไปอาจจะทำการศึกษาเมื่อกรณีค่าที่ถูกตัดทิ้งที่มีการแจกแจงในรูปแบบอื่น ๆ

5.3.2.5 ในการวิจัยครั้งนี้ ได้ทำการศึกษาเฉพาะค่าที่ถูกตัดทิ้งของตัวแปรตามเป็นตัวแปรเดียว (Univariate) เท่านั้น ดังนั้น ในการวิจัยครั้งต่อไปอาจจะทำการศึกษาในกรณีของค่าที่ถูกตัดทิ้งของตัวแปรตามเป็นสองตัว (Bivariate) เช่น  $T_1 = (T_{11}, T_{12})$  เป็นคู่ของอายุการใช้งานของไต่ด้านซ้ายและด้านขวา เส้นอัตราประมาณโดยเค็มแลนและไมเออร์