บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

ในการวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบการประมาณค่าพารามิเตอร์ที่มีประสิทธิภาพที่สุด สำหรับการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย โดยวิธีการประมาณค่าพารามิเตอร์ 3 วิธีด้วยกัน คือ วิธีโมเมนต์คัดแปลง วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด และวิธีกำลังสองต่ำสุดเทียม

วิธีดำเนินการวิจัยครั้งนี้ ใช้วิธีการจำลองแบบมอนติคาร์โล ทำงานโดยใช้กับเครื่องคอมพิวเตอร์ AMSAHL 5860 โดยใช้ภาษาฟอร์แทรน (FORTRAN 77) จำลองข้อมูลให้มีลักษณะตามแผนการทดลองที่กำหนด และกำหนดให้เครื่องคอมพิวเตอร์ ทำการทดลองซ้ำ ๆ กัน 1,000 ครั้ง ในแต่ละสถานการณ์ที่สนใจศึกษา

สรุปผลการวิจัย

การวิจัย จะให้ผลสรุปของความสามารถการประมาณค่าพารามิเตอร์ สำหรับการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย การเปรียบเทียบแต่ละวิธีการ โดยการศึกษาหาค่าความคลาดเคลื่อนของการประมาณค่าตัวแปรตามกับค่าจริง ทั้ง 3 วิธี เพื่อหาวิธีการใดจะดีที่สุด ในแต่ละสถานการณ์ที่ได้ทำการทำลองสรุปได้ดังนี้

สำหรับขนาดตัวอย่างทุกขนาด (N=10,30,50,70,100) จุดตัดปลายทุกระดับ (D=500,1000, 1500, 2000,3000) และการแจกแจงที่ถูกตัดปลายทางซ้ายทุกแบบคือ แบบลอกนอร์มอลตัดปลายทางซ้าย แบบเอกซโพเนนเชียลตัดปลายทางซ้าย แบบพาเรโต้ตัดปลายทางซ้าย (TLN,TEP,TPR)

1. การแจกแจงแบบลอกนอร์มอลตัดปลายทางซ้าย การประมาณพารามิเตอร์ด้วยวิธีโมเมนต์คัดแปลงจะให้ค่าความคลาดเคลื่อน RMSE ต่ำที่สุด วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดจะให้ค่า RMSE ใกล้เคียงกับวิธีโมเมนต์คัดแปลง ส่วนวิธีกำลังสองต่ำสุดเทียมให้ค่า RMSE สูงกว่าทุกวิธี ในทุกสถานการณ์ที่ทำการศึกษา

ปัจจัยที่มีผลกระทบต่อค่า RMSE ของการประมาณค่าพารามิเตอร์สำหรับการแจกแจงแบบลอกนอร์มอลตัดปลายทางซ้ายมีดังนี้คือ จุดตัดปลายและขนาดของตัวอย่างซึ่งจะเห็นว่า เมื่อจุดตัดปลายคงที่ค่า RMSE จะมีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น และเมื่อขนาดตัวอย่างคงที่ค่า RMSE จะมีค่าเพิ่มขึ้นเมื่อจุดตัดปลายเพิ่มขึ้น

2. การแจกแจงแบบเอกซโพเนนเชียลตัดปลายทางซ้าย การประมาณพารามิเตอร์ด้วยวิธีโมเมนต์คัดแปลงจะให้ค่าความคลาดเคลื่อน RMSE ต่ำที่สุด วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดจะให้ค่า RMSE ใกล้เคียงกับวิธีโมเมนต์คัดแปลง ส่วนวิธีกำลังสองต่ำสุดเทียมให้ค่า RMSE สูงกว่าทุกวิธี ในทุกสถานการณ์ที่ทำการศึกษา

ปัจจัยที่มีผลกระทบต่อค่า RMSE ของการประมาณค่าพารามิเตอร์สำหรับการแจกแจงแบบเอกซโพเนนเชียลตัดปลายทางซ้ายมีดังนี้คือ จุดตัดปลายและขนาดของตัวอย่างซึ่งจะเห็นว่าเมื่อจุดตัดปลายคงที่ค่า RMSE จะมีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น และเมื่อขนาดตัวอย่างคงที่ค่า RMSE จะมีค่าเพิ่มขึ้นเพียงเล็กน้อยเมื่อจุดตัดปลายเพิ่มขึ้นสำหรับทุกวิธีการประมาณค่าพารามิเตอร์ที่ทำการศึกษา

3. การแจกแจงแบบพาเรโต้ตัดปลายทางซ้าย การประมาณพารามิเตอร์ด้วยวิธีโมเมนต์คัดแปลงจะให้ค่าความคลาดเคลื่อน RMSE ต่ำที่สุด วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดจะให้ค่า RMSE ใกล้เคียงกับวิธีโมเมนต์คัดแปลง ส่วนวิธีกำลังสองต่ำสุดเทียมให้ค่า RMSE สูงกว่าทุกวิธีในทุกสถานการณ์ที่ทำการศึกษา

ปัจจัยที่มีผลกระทบต่อค่า RMSE ของการประมาณค่าพารามิเตอร์สำหรับการแจกแจงแบบพาเรโต้ตัดปลายทางซ้ายมีดังนี้คือ จุดตัดปลายและขนาดของตัวอย่างซึ่งจะเห็นว่า เมื่อจุดตัดปลายคงที่ค่า RMSE จะมีค่าลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นสำหรับวิธีโมเมนต์คัดแปลงและวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด ยกเว้นวิธีกำลังสองต่ำสุดเทียมค่า RMSE เพิ่มขึ้นตามขนาดตัวอย่าง และเมื่อขนาดตัวอย่างคงที่ค่า RMSE จะมีแนวโน้มที่เพิ่มขึ้นเมื่อจุดตัดปลายเพิ่มขึ้น

4. สำหรับทุกการแจกแจงที่ทำการวิจัยในงานวิจัยครั้งนี้สามารถสรุปได้ว่าวิธีโมเมนต์คัดแปลง จะให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าพารามิเตอร์ต่ำกว่าวิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุดและวิธีกำลังสองต่ำสุดเทียม ในทุกสถานการณ์ที่ทำการศึกษา และปัจจัยที่มีผลต่อค่า RMSE คือ จุดตัดปลายและขนาดตัวอย่าง โดยที่ เมื่อจุดตัดปลายคงที่ค่า RMSE จะมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น และเมื่อขนาดตัวอย่างคงที่ค่า RMSE จะมีแนวโน้มเพิ่มขึ้นเมื่อจุดตัดปลายเพิ่มขึ้น สำหรับทุกการแจกแจงที่ทำการศึกษา

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้ จะเสนอแนะ 2 ด้านคือ

1. ด้านการนำไปใช้

กรณีที่ต้องการประมาณค่าพารามิเตอร์ สำหรับการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย 3 แบบ คือ ลอกนอร์มอลตัดปลายทางซ้าย เอกซโพเนนเชียลตัดปลายทางซ้าย พาเรโต้ตัดปลายทางซ้าย โดยประมาณค่าพารามิเตอร์ 3 วิธีด้วยกันคือ วิธีโมเมนต์คัดแปลง วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด และวิธีกำลังสองต่ำสุดเทียม มีแนวทางดังนี้

กรณีที่ 1 เมื่อทราบการแจกแจงของข้อมูล

สำหรับขนาดตัวอย่างทุกขนาด (N=10,30,50,70,100) จุดตัดปลายทุกระดับ (D=500, 1000, 1500, 2000,3000) และการแจกแจงที่ถูกตัดปลายทางซ้ายทุกแบบคือ แบบลอกนอร์มอลตัดปลายทางซ้าย แบบเอกซโพเนนเชียลตัดปลายทางซ้าย แบบพาเรโต้ตัดปลายทางซ้าย (TLN, TEP, TPR) เมื่อไม่ทราบจุดตัดปลายควรเลือกใช้วิธีโมเมนต์คัดแปลง เพราะจะเป็นวิธีที่ให้ค่าคลาดเคลื่อนของการประมาณค่าพารามิเตอร์ต่ำสุด

กรณีที่ 2 เมื่อไม่ทราบการแจกแจงของข้อมูล

นำข้อมูลที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลายมาคัดเลือกตัวแบบ ซึ่งสามารถทำได้หลายวิธี คือ วิธีการใช้เส้นโค้งโอกีฟ (Ogive) หรือ วิธี Mean Residual Life Function ($e_n(x)$) สำหรับวิธีแรกนั้นทำการหากราฟแสดงความสัมพันธ์ระหว่างสมการโอกีฟ (H(x)) กับข้อมูล (x) ซึ่งประมาณค่าสมการโอกีฟหาได้จาก

$$H'(x) = h(x) = \frac{F_n(c_i) - F_n(c_{i-1})}{c_i - c_{i-1}} = \frac{f_n(c_i)}{n(c_i - c_{i-1})} \quad , \quad c_{i-1} < x \le c_i$$

โดยที่ $f_n(c_i)$ หมายถึงฟังก์ชันแสดงความถี่ของข้อมูล และ $F_n(c_i)$ หมายถึงฟังก์ชันแสดงความถี่สะสมของข้อมูล ทำการเปรียบเพื่อหาดูลักษณะการแจกแจงที่ใกล้เคียงกับข้อมูลมากที่สุด[1]

ส่วนวิธี Mean Residual Life Function (e(x)) นั้นเป็นการหากราฟแสดงความสัมพันธ์ระหว่าง e(x) กับ x ซึ่งสามารถหา e(x) ได้จาก

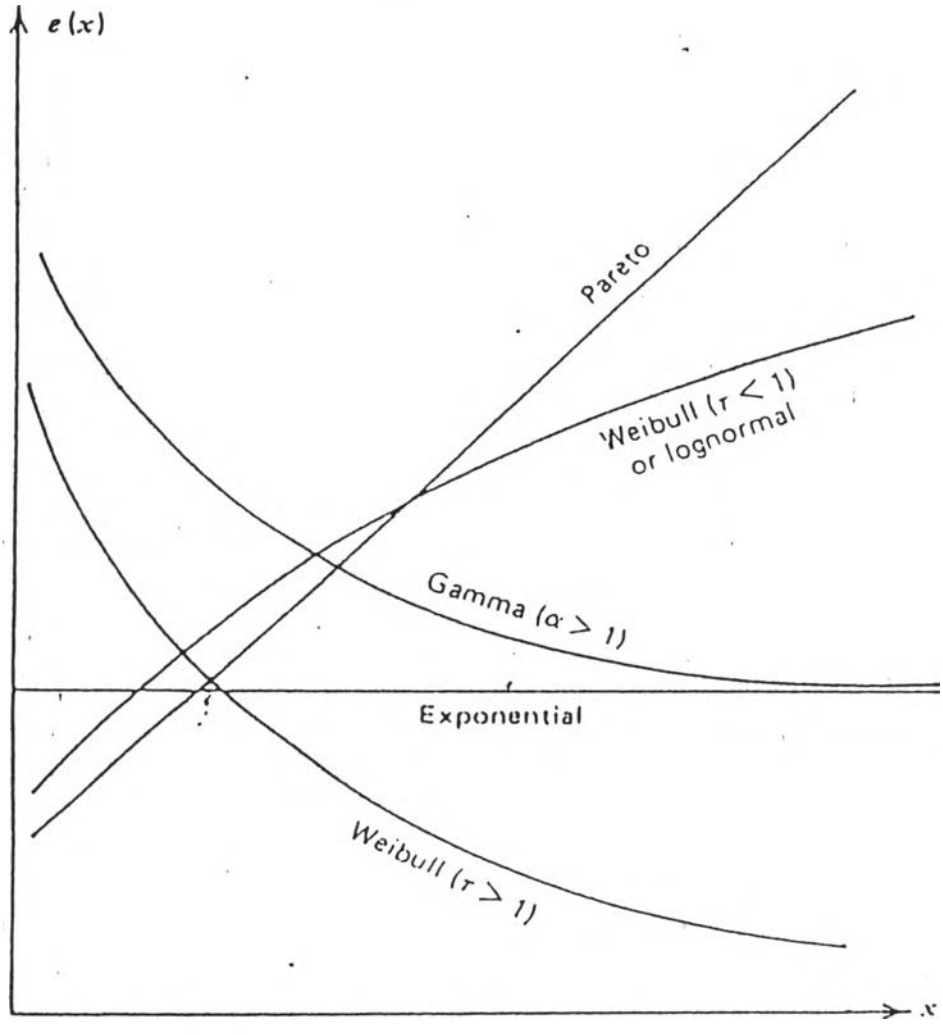
$$e(x) = \int_x^{\infty} (w - x) \frac{f(w)}{\Pr(X \ge x)} d(w) = \frac{\int_x^{\infty} [1 - F(w)] d(w)}{1 - F(w)} \quad , \quad w > x > 0$$

---

[1] Robert V. Hogg & Stuart A. Klugman. Loss Distribution. (New York: John Wiley & Sons,1984), pp. 63-65.

ทำการเปรียบเพื่อหาดูลักษณะการแจกแจงที่ใกล้เคียงกับข้อมูลมากที่สุด โดยเปรียบเทียบได้จากรูปกราฟต่อไปนี้[2]

รูปที่ 5.1 แสดง Mean residual life function e(x) ของการแจกแจงแบบต่าง ๆ



เมื่อได้รูปแบบของการแจกแจงแล้ว สำหรับขนาดตัวอย่างทุกขนาด (N=10, 30, 50, 70, 100) จุดตัดปลายทุกระดับ (D=500, 1000, 1500, 2000,3000) และการแจกแจงที่ถูกตัดปลายทางซ้ายทุกแบบคือ แบบลอกนอร์มอลตัดปลายทางซ้าย แบบเอกซโพเนนเชียลตัดปลายทางซ้าย แบบพาเรโต้ตัดปลายทางซ้าย (TLN,TEP,TPR) เมื่อไม่ทราบจุดตัดปลาย ควรเลือกใช้วิธีโมเมนต์ดัดแปลง เพราะจะเป็นวิธีที่ให้ค่าคลาดเคลื่อนของการประมาณค่าพารามิเตอร์ต่ำสุด

---

[2] Robert V. Hogg & Stuart A. Klugman. Loss Distribution. (New York: John Wiley & Sons,1984), pp. 107-109.

2. ด้านการศึกษาวิจัย

2.1 กรณีของการแจกแจงที่ถูกตัดปลายทางขวาหรือการแจกแจงที่ถูกตัดปลายทางทั้ง 2 ด้านคือ ทั้งด้านซ้ายและด้านขวา เมื่อไม่ทราบจุดตัดปลาย วิธีการประมาณค่าพารามิเตอร์สำหรับการแจกแจงเหล่านี้อาจจะให้ผลของการประมาณแตกต่างหรือให้ผลที่ไม่แตกต่างจากกรณีของการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย จึงน่าจะมีการศึกษาวิจัยในปัญหาต่ำไป

2.2 ในการวิจัยครั้งนี้ เป็นการศึกษาวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีการ 3 วิธี คือ วิธีโมเมนต์ดัดแปลง วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด วิธีกำลังสองต่ำสุดเทียม ในการศึกษาครั้งต่อไปควรจะศึกษาวิธีอื่นที่แตกต่างจาก 3 วิธีนี้

2.3 ในการวิจัยครั้งนี้ได้ทำการศึกษาการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย 3 แบบ คือ แบบลอกนอร์มอลตัดปลายทางซ้าย แบบเอกซโพเนนเชียลตัดปลายทางซ้าย แบบพาเรโต้ตัดปลายทางซ้าย ดังนั้นในการศึกษาครั้งต่อไปควรนำการแจกแจงแบบอื่นๆ มาทำการศึกษา

2.4 ในการวิจัยครั้งนี้ได้นำเอาวิธีนิวตัน-ราฟสัน (Newton-Raphson method) เข้ามาช่วยในการหาค่าประมาณพารามิเตอร์ ซึ่งวิธีนิวตัน-ราฟสันนั้นการลู่เข้าให้ได้คำตอบนั้นขึ้นอยู่กับค่าเริ่มต้นและสมการที่เราใช้ในการประมาณ ค่าเริ่มต้นที่ใกล้เคียงกับค่าพารามิเตอร์จริงจะทำให้การลู่เข้าเร็วขึ้น ส่วนรูปแบบสมการนั้นบางรูปแบบการลู่เข้าจะเร็วบางรูปแบบจะไม่ลู่เข้า ดังนั้นหากต้องนำเอาวิธีนิวตัน-ราฟสันมาใช้ในการศึกษาครั้งต่อไปควรทดลองหาค่าเริ่มต้นที่ใกล้เคียงและหารูปแบบสมการที่ทำให้ลู่เข้าได้เร็วขึ้น

รูปที่ 5.2 แสดงขั้นตอนของการนำไปใช้งาน