บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในการวิจัยครั้งนี้ต้องการศึกษาวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสีย เมื่อข้อมูลความสูญเสียเป็นแบบกลุ่มและถูกตัดปลายทางซ้าย โดยประมาณค่าพารามิเตอร์ด้วยวิธีการ 3 วิธี คือ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบไค-สแควร์ต่ำสุด และวิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์

วิธีดำเนินการวิจัยครั้งนี้ ใช้วิธีการจำลองด้วยเทคนิคมอนติคาร์โล เขียนโปรแกรมด้วยภาษาฟอร์แทรน (FORTRAN) โดยใช้กับเครื่องคอมพิวเตอร์ AMDAHL 5860 จำลองข้อมูลเพื่อให้มีลักษณะตามแผนการทดลองที่กำหนด ทำการทดลองซ้ำๆกัน 1,000 ครั้ง ในแต่ละสถานการณ์ที่สนใจศึกษา

สรุปผลการวิจัย

การวิจัย จะให้ผลสรุปของประสิทธิภาพของการประมาณค่าพารามิเตอร์สำหรับการแจกแจงความสูญเสีย เมื่อข้อมูลเป็นแบบกลุ่มและข้อมูลถูกตัดปลายทางซ้าย การเปรียบเทียบประสิทธิภาพของแต่ละวิธีการ โดยการศึกษาหาค่าความคลาดเคลื่อนของการประมาณค่าพารามิเตอร์ทั้ง 3 วิธี เพื่อหาว่าวิธีการใดจะดีที่สุด ในแต่ละสถานการณ์ที่ได้ทำการทดลองสรุปได้ดังนี้

1. เมื่อข้อมูลมีการแจกแจงแบบลอกนอร์มอล ($\mu=2,\sigma=0.5$ และ $\mu=1,\sigma=0.5,1,2$) ค่าสัมประสิทธิ์ความแปรผันจะแปรผันตามค่าพารามิเตอร์แสดงรูปร่าง (σ)

เมื่อตัวอย่างมีขนาดเล็ก (n) = 100 , 200 สำหรับ $\mu=1,\sigma=0.5,1$ และ $\mu=2,\sigma=0.5$ ที่ระดับค่ารับผิดส่วนแรก (d) ทุกระดับ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุดเป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำที่สุด ส่วน $\mu=1,\sigma=2$ วิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ เป็น วิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำสุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ ที่ระดับค่ารับผิดส่วนแรกทุกระดับ

เมื่อตัวอย่างมีขนาดใหญ่ (n) = 300 , 500 สำหรับ $\mu=1,\sigma=0.5,1$ และ $\mu=2,\sigma=0.5$ ที่ระดับค่ารับผิดส่วนแรก (d) ทุกระดับ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุดเป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำที่สุด ส่วน $\mu=1,\sigma=2$ ที่ระดับค่ารับผิดส่วนแรกต่ำ (d) = 0.5 , 1.0 วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุดเป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณ

ต่ำที่สุด แต่เมื่อระดับค่ารับผิดส่วนแรกสูง (d) = 2.0 และ 3.0 วิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ เป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำสุดเมื่อเปรียบเทียบกับวิธีการอื่น

2. เมื่อข้อมูลมีการแจกแจงแบบไวบูลล์ (c=1, τ=0.3 และ c=0.1, τ=1,1.5,2)

ค่าสัมประสิทธิ์ความแปรผันจะแปรผันตามค่าพารามิเตอร์แสดงรูปร่าง (τ)

เมื่อตัวอย่างมีขนาดเล็ก (n) = 100 , 200 สำหรับ c=0.1, τ =1,1.5 และ 2 ที่ระดับค่ารับผิดส่วนแรก (d) ทุกระดับ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุดเป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำที่สุด ส่วน c=1, τ =0.3 วิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ เป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำสุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ ที่ระดับค่ารับผิดส่วนแรกทุกระดับ

เมื่อตัวอย่างมีขนาดใหญ่ (n) = 300 , 500 สำหรับ c=0.1, τ =1,1.5,2 และ c=1, τ =0.3 ที่ระดับค่ารับผิดส่วนแรก (d) ทุกระดับ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุดเป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำที่สุด

3. เมื่อข้อมูลมีการแจกแจงแบบพาเรโต (λ =10, α =2.5,4,7)

ค่าสัมประสิทธิ์ความแปรผันจะแปรผันตามค่าพารามิเตอร์แสดงรูปร่าง (α)

ที่ทุกขนาดตัวอย่าง สำหรับ λ =10, α =2.5,4 และ 7 วิธีการประมาณแบบไค-สแควร์ต่ำสุด เป็นวิธีที่ให้ความคลาดเคลื่อนของการประมาณต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ ที่ทุกระดับของค่ารับผิดส่วนแรก

ปัจจัยที่มีอิทธิพลต่อค่าความคลาดเคลื่อน RMSE ของการประมาณ ทั้ง 3 วิธี คือ

1. รูปร่างของข้อมูลที่ถูกตัด

1.1 กรณีข้อมูลมีการแจกแจงแบบลอกนอร์มอล เมื่อ σ เพิ่มขึ้น (CV เพิ่มขึ้น) ค่าความคลาดเคลื่อน RMSE จะมีแนวโน้มเพิ่มขึ้น

1.2 กรณีข้อมูลมีการแจกแจงแบบไวบูลล์ เมื่อ τ เพิ่มขึ้น (CV ลดลง) ค่าความคลาดเคลื่อน RMSE จะมีแนวโน้มเพิ่มขึ้น

1.3 กรณีข้อมูลมีการแจกแจงแบบพาเรโต ทุกระดับของ α จะให้รูปร่างของข้อมูลที่มีลักษณะคล้ายการแจกแจงแบบเอกซโพเนนเชียล และเมื่อ α มีค่าเพิ่มขึ้น (CV ลดลง) กราฟของการแจกแจงจะมีความชันเพิ่มขึ้น ค่าความคลาดเคลื่อน RMSE จะมีแนวโน้มเพิ่มขึ้น

2. สำหรับทุกการแจกแจง กรณีพารามิเตอร์แต่ละตัวคงที่ และขนาดตัวอย่างคงที่ เมื่อค่ารับผิดส่วนแรกเพิ่มขึ้น ค่าความคลาดเคลื่อน RMSE มีแนวโน้มเพิ่มมากขึ้น

3. สำหรับทุกการแจกแจง กรณีพารามิเตอร์แต่ละตัวคงที่ และค่ารับผิดส่วนแรกคงที่ เมื่อขนาดตัวอย่างเพิ่มขึ้น ค่าความคลาดเคลื่อน RMSE มีแนวโน้มลดลง

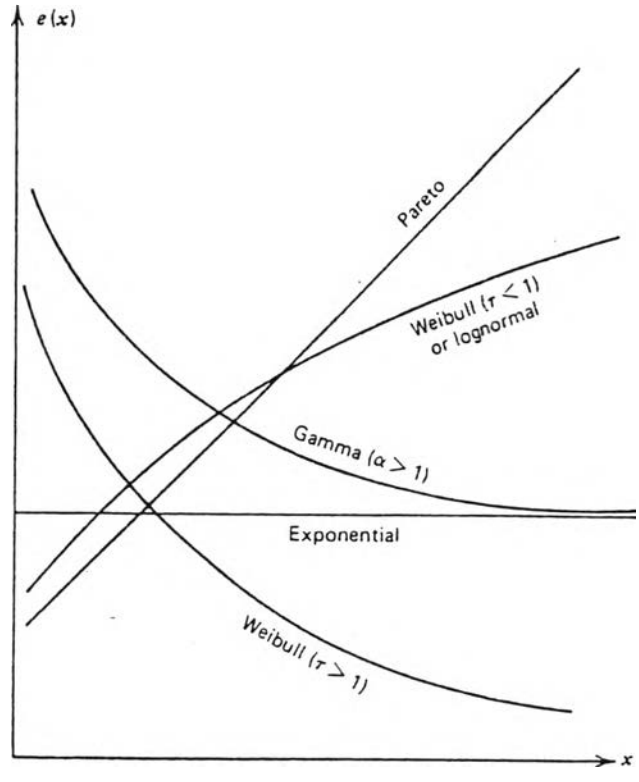
ข้อเสนอแนะ

เพื่อเป็นแนวทางในการเลือกวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสีย และเพื่อเป็นแนวทางสำหรับการศึกษาวิจัยนอกเหนือจากการวิจัยครั้งนี้ ดังนั้นผู้วิจัยขอเสนอแนะเป็น 2 ด้าน คือ

1. ด้านการนำไปใช้ประโยชน์

ในการวิจัยครั้งนี้ จะสนใจเฉพาะขั้นตอนการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสียซึ่งเป็นขั้นตอนหนึ่งในการทำงานเท่านั้น ดังนั้นผู้วิจัยจะขอกล่าวถึงขั้นตอนการทำงานทั้งหมดหลังจากได้ข้อมูลมาแล้ว โดยจะกล่าวรวมถึงการนำผลการวิจัยในครั้งนี้ไปใช้ประโยชน์ในขั้นตอนการประมาณค่าพารามิเตอร์ ขั้นตอนการทำงานมีดังนี้คือ

1) การคัดเลือกตัวแบบเพื่อมาตรวจสอบ วิธีการคัดเลือกเบื้องต้น อาจใช้เส้นโค้งโอกีฟ (Ogive) และฮิสโตแกรมแสดงความถี่สัมพัทธ์ (Relative Frequency Histogram) ว่ามีลักษณะคล้ายการแจกแจงแบบใด แต่ถ้าข้อมูลมีลักษณะถูกตัดปลายทางซ้าย จะใช้ Mean Residual Life Function $e(x)$ ในเชิงทฤษฎี ซึ่งได้แสดงกราฟไว้ในรูปที่ 5.1 เพื่อเปรียบเทียบกับกราฟของ Empirical Mean Residual Life $e_n(x)$ (ดูรายละเอียดได้จาก Loss Distributions .,R.V. Hogg , S.A. Klugman.,1984) จากนั้นจะเลือกตัวแบบที่กราฟแสดง $e(x)$ คล้ายกับ $e_n(x)$ ซึ่งอาจเลือกได้หลายการแจกแจง



รูปที่ 5.1 แสดงกราฟของ Mean Residual Life Function $e(x)$ ของการแจกแจงต่างๆ

2) การประมาณค่าพารามิเตอร์ในตัวแบบ ซึ่งเป็นขั้นตอนที่สนใจในการวิจัยครั้งนี้ โดยจะเลือกใช้วิธีการที่ให้ค่าความคลาดเคลื่อน RMSE ของการประมาณค่าพารามิเตอร์ซึ่งโดยส่วนใหญ่แล้วต่ำกว่าวิธีการอื่น ๆ ซึ่งแนวทางการเลือกใช้มีดังนี้

กรณีข้อมูลมีการแจกแจงแบบลอกนอร์มอล เมื่อตัวอย่างมีขนาดเล็ก (n = 100 , 200) สำหรับค่ารับผิดส่วนแรกมีค่าต่ำ (d = 0.5 , 1.0) ควรเลือกใช้วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด แต่ถ้าค่ารับผิดส่วนแรกมีค่าสูง (d = 2.0 , 3.0) หรือเมื่อตัวอย่างมีขนาดใหญ่ (n = 300 , 500) ควรเลือกใช้วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด หรือวิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ [1]

กรณีข้อมูลมีการแจกแจงแบบไวบูลล์ ควรเลือกใช้วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด

---

[1] สำหรับการแจกแจงแบบลอกนอร์มอลและการแจกแจงแบบไวบูลล์ เมื่อข้อมูลมีการกระจายมาก (CV สูงมาก) จะทำให้วิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์เริ่มดีขึ้น

กรณีข้อมูลมีการแจกแจงแบบพาเรโต ควรเลือกใช้วิธีการประมาณแบบไค-สแควร์ต่ำสุด

2. ด้านการศึกษาวิจัย

2.1 กรณีข้อมูลความสูญเสียมีค่าถูกตัดปลายทั้งสองด้าน คือ ทางซ้ายและทางขวา วิธีการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสีย อาจจะให้ผลของการประมาณแตกต่างจากกรณีข้อมูลถูกตัดปลายทางซ้าย จึงน่าจะมีการศึกษาวิจัยในปัญหาต่อไป

2.2 ในการวิจัยครั้งนี้ ได้ทำการศึกษาเมื่อความสูญเสียมีการแจกแจงเป็นแบบใดแบบหนึ่ง ซึ่งในที่นี้เป็นแบบลอกนอร์มอล ไวบูลล์ และพาเรโต ดังนั้น ในการศึกษาครั้งต่อไป อาจทำการศึกษาเมื่อการแจกแจงของความสูญเสียเป็นแบบผสม (Mixed Distributions) หรือแบบอื่น ๆ

รูปที่ 5.2 แผนผังแสดงลำดับขั้นตอนการนำไปใช้ประโยชน์