บทที่ 1

บทนำ

ที่มาและความสำคัญของปัญหา

ในงานทางด้านคณิตศาสตร์ประกันวินาศภัย การแจกแจงของความเสียหายในเทอมของเงินค่าสินไหมทดแทน (Claims) นั้นมีความสำคัญมากในการปรับเบี้ยประกันภัย (Premium) ในปีต่อไป หรือนำไปคำนวณหาเงินสำรองความเสียหาย (Loss Reserve) ที่จะเกิดขึ้นในปีต่อไป เนื่องจากหากประมาณค่าไม่เป็นจริงหรือไม่ครอบคลุมค่าใช้จ่ายที่เกิดขึ้นแล้ว เป็นสาเหตุหนึ่งที่จะทำให้บริษัทเกิดความตกต่ำ (Ruins) คือเหตุการณ์ที่ค่าใช้จ่ายมากกว่าเงินสำรองค่าสินไหมทดแทนรวมกับเบี้ยประกันรับ

ข้อมูลที่เกิดขึ้นในทางประกันภัยนั้นส่วนใหญ่จะเป็นข้อมูลที่ถูกตัดปลายซึ่งอาจจะเป็นข้อมูลที่ตัดปลายทางซ้ายจากความรับผิดส่วนแรก (Deductible) หรือตัดปลายทางขวา (Censored) จากความรับผิดชอบสูงสุด (Limit) หรือทั้งตัดปลายทางซ้ายและทางขวา การตัดปลายของข้อมูลนั้นหากเราทราบจุดตัดปลายเราก็จะสามารถประมาณพารามิเตอร์ง่ายขึ้น แต่บางครั้งการเก็บข้อมูลของบริษัทนั้นจะไม่ทราบจุดตัดปลาย ดังนั้นในงานวิจัยครั้งนี้ ผู้ทำวิจัยต้องการที่จะศึกษาข้อมูลที่ไม่ทราบจุดตัดปลายและเป็นข้อมูลตัดปลายทางซ้าย โดยทำการศึกษาวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงต่างๆ ของข้อมูลความเสียหายที่ถูกตัดปลายทางซ้ายโดยความรับผิดส่วนแรกและไม่ทราบจุดตัดปลาย

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย 3 วิธี คือ

1. วิธีโมเมนต์ดัดแปลง (Method of Modified Moment)
2. วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด (Method of Maximum Likelihood Estimation)
3. วิธีกำลังสองต่ำสุดเทียม (Method of Pseudo Least-Squares)

สมมติฐานในการวิจัย

การประมาณพารามิเตอร์ของการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย ด้วยวิธีโมเมนต์ดัดแปลง (Method of Modified Moment) ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Root Mean Squares Error ; RMSE) จากการประมาณต่ำที่สุด

ขอบเขตของการวิจัย

1. ในการทำวิจัยครั้งนี้ ทำการเปรียบเทียบการประมาณค่าพารามิเตอร์สำหรับการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย ด้วยวิธี

1.1 วิธีโมเมนต์ดัดแปลง

1.2 วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด

1.3 วิธีกำลังสองต่ำสุดเทียม

2. ศึกษาในกรณี X เป็นตัวแปรสุ่ม มีการแจกแจงดังนี้

2.1 เมื่อตัวแปรสุ่ม X มีการแจกแจงแบบลอกนอร์มอล (Lognormal Distribution) โดยมี ฟังก์ชันความหนาแน่นเป็นดังนี้

$$f_x(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\dfrac{1}{2}\left(\dfrac{\ln(x)-\mu}{\sigma} \right)^2 \right) & , x > 0 \\ 0 & ,\text{อื่น ๆ} \end{cases}$$

โดยที่ μ เป็นพารามิเตอร์กำหนดขนาด (Scale Parameter) , $-\infty < \mu < \infty$ และ σ เป็นพารามิเตอร์กำหนดรูปร่าง (Shape Parameter) , $\sigma > 0$ มีฟังก์ชันการแจกแจงสะสมเป็นดังนี้

$$\begin{aligned} F(x) &= \int_0^x \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{\ln(x)-\mu}{\sigma} \right]^2 \right) dx \\ &= \int_{-\infty}^{z(x)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z^2 \right) dz = \Phi(z(x)) \end{aligned}$$

2.2 เมื่อตัวแปรสุ่ม X มีการแจกแจงแบบเอกซโพเนนเชียล (Exponential Distribution) โดยมีฟังก์ชันความหนาแน่นเป็นดังนี้

$$f_x(x) = \begin{cases} \dfrac{1}{\theta} \exp\left(-\dfrac{1}{\theta} x \right) & , x > 0 \\ 0 & ,\text{อื่น ๆ} \end{cases}$$

โดยที่ θ เป็นพารามิเตอร์กำหนดขนาด (Scale Parameter) , θ > 0 และฟังก์ชันการแจกแจงสะสมเป็นดังนี้

$$F_x(x) = 1 - \exp\left(-\frac{1}{\theta}x\right)$$

2.3 เมื่อตัวแปรสุ่ม X มีการแจกแจงแบบพาเรโต้ (Pareto Distribution) โดยมีฟังก์ชันความหนาแน่นเป็นดังนี้

$$f(x) = \begin{cases} \dfrac{\alpha\lambda^{\alpha}}{(\lambda + x)^{(\alpha+1)}} & , x > 0 \\ 0 & , \text{อื่น ๆ} \end{cases}$$

โดยที่ α เป็นพารามิเตอร์กำหนดรูปร่าง (Shape Parameter) , α > 0 และ λ เป็นพารามิเตอร์กำหนดขนาด (Scale Parameter) , λ > 0 และฟังก์ชันการแจกแจงสะสมเป็นดังนี้

$$F_x(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}$$

3. ศึกษาในกรณีข้อมูลที่มีการตัดปลายทางซ้าย (Left-Truncated Data) เมื่อไม่ทราบจุดตัดปลายเป็นดังนี้

$$W = X \qquad , X > d$$

โดยที่ W และ X เป็นตัวแปรสุ่ม d เป็นจุดตัดปลาย ซึ่งมีฟังก์ชันความหนาแน่น ดังนี้

$$f_w(x) = \frac{f_x(x)}{[1 - F_x(d)]} \qquad , x > d$$

ฟังก์ชันการแจกแจง(สะสม) เป็นดังนี้

$$F_w(x) = \frac{F_x(x) - F_x(d)}{1 - F_x(d)}$$

4. ศึกษาเมื่อขนาดตัวอย่าง 5 ระดับ คือ 10, 30, 50, 70 และ 100 ตามลำดับ

5. ศึกษาเมื่อจุดตัดปลาย 5 ระดับ คือ 500,1000,1500,2000 และ 3000 ตามลำดับ

การวิจัยครั้งนี้จำลองข้อมูลให้มีสถานการณ์ตามที่ต้องการศึกษาโดยใช้เทคนิคการจำลองแบบมอนติ คาร์โล (Monte Carlo Simulation Technique) โดยเขียนโปรแกรมด้วยภาษาฟอร์แทรน ทำการจำลองข้อมูลซ้ำๆกัน 1,000 ครั้ง ในแต่ละสถานการณ์

เกณฑ์การตัดสินใจ

เกณฑ์การตัดสินใจว่าการประมาณด้วยวิธีใดมีความเหมาะสมและมีประสิทธิภาพมากกว่าพิจารณาโดยใช้ ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง(Root Mean Squares Error , RMSE) ช่วยในการพิจารณานั่นคือ ใช้ค่าความแตกต่างกำลังสองของค่าจริงของพารามิเตอร์เทียบกับ

ค่าที่ได้จากการประมาณค่าพารามิเตอร์ในแต่ละวิธี วิธีที่ให้ค่า RMSE ต่ำที่สุดจะเป็นวิธีที่ดีที่สุดในการประมาณค่าพารามิเตอร์ การหาค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองหาได้จากสูตรดังนี้

$$MSE = \frac{1}{1000*n}\sum_{j=1}^{1000}\sum_{i=1}^{n}\left(\theta_{ij} - \hat{\theta}_{ij}\right)^2$$

$$RMSE = \sqrt{MSE}$$

โดยที่ n แทน จำนวนพารามิเตอร์ ; i=1,...,n

j แทน จำนวนรอบที่ทำซ้ำ ; j=1,...,1000

MSE แทน ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง

RMSE แทน ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเป็นแนวทางให้นักคณิตศาสตร์และนักวิจัย เลือกใช้วิธีการประมาณค่าพารามิเตอร์ สำหรับการแจกแจงที่ถูกตัดปลายทางซ้ายเมื่อไม่ทราบจุดตัดปลาย ได้อย่างมีประสิทธิภาพและเหมาะสมในแต่ละสถานการณ์ คือจะทำให้ได้ค่าประมาณของพารามิเตอร์ที่ใกล้เคียงกับค่าจริง

2. เพื่อเป็นแนวทางในการศึกษา และเปรียบเทียบการประมาณค่าพารามิเตอร์สำหรับการแจกแจกแจงที่ถูกตัดปลายแบบอื่นๆ เช่น การแจกแจงที่ถูกตัดปลายทางขวา หรือ การแจกแจงที่ถูกตัดปลายทั้งสองข้าง