บทที่ 1

บทนำ

ที่มาและความสำคัญของปัญหา

การประกันภัยเป็นเครื่องมือที่ใช้ในการเตรียมการต่อการสูญเสียหรือความเสียหายในอนาคตซึ่งเกิดขึ้นเนื่องจากเหตุการณ์ที่ไม่อาจคาดเดาได้ เช่น การเกิดอัคคีภัย และอุบัติเหตุทางรถยนต์ เป็นต้น ในธุรกิจประกันภัยผู้ที่มีบทบาทสำคัญมากผู้หนึ่งในการดำเนินงาน ก็คือ นักคณิตศาสตร์ประกันภัย หน้าที่ที่สำคัญของนักคณิตศาสตร์ประกันภัยคือการคำนวณอัตราเบี้ยประกันภัย การคำนวณเงินสำรอง และอื่นๆ ในการวิเคราะห์ คำนวณอัตราหรือค่าต่างๆดังกล่าว โดยทั่วไปจะพบว่าข้อมูลความสูญเสียที่นำมาใช้นั้นเป็นข้อมูลที่ไม่สมบูรณ์ (Incomplete Data) เช่น ข้อมูลถูกตัดปลาย (Truncated Data) ทำให้อัตราเบี้ยประกันภัยที่กำหนดขึ้นอาจมีค่าสูงหรือต่ำเกินไป และการคำนวณเงินสำรองและค่าอื่นๆ อาจเกิดความผิดพลาดได้มาก ดังนั้นจึงควรได้มีการศึกษารูปแบบการแจกแจงและการประมาณค่าพารามิเตอร์ของความสูญเสียให้ใกล้เคียงกับความเป็นจริงมากที่สุด เพื่อให้การคำนวณอัตราเบี้ยประกันภัย และการคำนวณค่าอื่นๆ มีความถูกต้องยิ่งขึ้น

ข้อมูลที่นำมาวิเคราะห์ในที่นี้คือข้อมูลที่บริษัทมีการบันทึกไว้จัดเป็นข้อมูลที่ไม่สมบูรณ์ ตัวอย่างเช่น ในการประกันภัยรถยนต์ได้กำหนดเงื่อนไขข้อหนึ่งไว้ว่า ผู้เอาประกันภัยจะต้องรับผิดชอบเองต่อความสูญเสียที่มีค่าต่ำกว่าค่ารับผิดส่วนแรก (Deductible) หรืออาจกล่าวได้ว่าค่ารับผิดส่วนแรกเป็นความสูญเสียที่ผู้เอาประกันภัยจะต้องรับผิดชอบในเบื้องต้น ดังนั้นความสูญเสียที่มีค่าต่ำกว่าค่ารับผิดส่วนแรกจะไม่มีการบันทึกไว้ ลักษณะของข้อมูลประเภทนี้เรียกว่า ข้อมูลถูกตัดปลายทางซ้าย (Left-Truncated Data) ส่วนความสูญเสียที่มีค่าสูงกว่าค่ารับผิดส่วนแรกจะมีการบันทึกไว้โดยบันทึกเป็นความสูญเสียที่เกิดขึ้นจริง(ตามความตกลง)แม้ว่าบริษัทจะชดใช้เฉพาะค่าความสูญเสียส่วนที่เกินจากค่ารับผิดส่วนแรก เพราะในทางปฏิบัติบริษัทจะชดใช้ค่าความสูญเสียที่เกิดขึ้นให้กับคู่กรณีและจะเรียกเก็บค่ารับผิดส่วนแรกกับผู้เอาประกันภัยในภายหลัง

ในการวิจัยครั้งนี้ นอกจากสนใจที่จะศึกษากรณีข้อมูลความสูญเสียมีค่าถูกตัดปลายทางซ้ายแล้วยังต้องการศึกษาในกรณีที่ข้อมูลเป็นแบบกลุ่ม (Grouped Data) ซึ่งเป็นอีกแบบหนึ่งของ

การบันทึกข้อมูลนอกเหนือจากการเก็บเป็นรายกรมธรรม์ (Individual Data) โดยจะกำหนดช่วงของความสูญเสียแล้วเก็บในรูปของความถี่ซึ่งก็คือจำนวนกรมธรรม์

จากปัญหาดังกล่าว ผู้วิจัยจึงสนใจที่จะทำการศึกษาการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสียภายใต้ข้อมูลแบบกลุ่มซึ่งถูกตัดปลายทางซ้าย (Left-Truncated Grouped Data) ในที่นี้จะศึกษาลักษณะของค่าความสูญเสียภายใต้การแจกแจง 3 แบบ คือ การแจกแจงแบบลอกนอร์มอล (Lognormal Distribution) การแจกแจงแบบไวบูลล์ (Weibull Distribution) และการแจกแจงแบบพาเรโต (Pareto Distribution) การแจกแจงทั้ง 3 แบบนี้เป็นแบบเบ้ทางขวาซึ่งเป็นการแจกแจงความสูญเสียที่พบบ่อย และจะศึกษาวิธีการประมาณพารามิเตอร์ 3 วิธี ดังต่อไปนี้

1. วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation Method)
2. วิธีการประมาณแบบไค-สแควร์ต่ำสุด (Minimum Chi-Square Estimation Method)
3. วิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ (Cramer-Von Mises Minimum Distance Estimation Method)

วัตถุประสงค์ของการวิจัย

เพื่อศึกษาวิธีการประมาณค่าพารามิเตอร์ทั้ง 3 วิธีที่กล่าวไว้ข้างต้น สำหรับการแจกแจงความสูญเสีย เมื่อข้อมูลเป็นแบบกลุ่มและข้อมูลถูกตัดปลายทางซ้าย

สมมติฐานทางการวิจัย

การประมาณพารามิเตอร์ด้วยวิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ จะให้ค่าประมาณใกล้เคียงกับค่าจริงมากกว่าวิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด และวิธีการประมาณแบบไค-สแควร์ต่ำสุด เมื่อข้อมูลเป็นแบบกลุ่มและข้อมูลถูกตัดปลายทางซ้าย

ข้อตกลงเบื้องต้น

กำหนดให้

$$Y = X \quad , X > d$$

$$F_Y(x) = P(X \le x / X > d)$$

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

เมื่อ

Y คือ ตัวแปรสุ่มของความสูญเสียที่มีการบันทึก

X คือ ตัวแปรสุ่มของความสูญเสียที่เกิดขึ้นจริง

d คือ ค่ารับผิดส่วนแรก

$F_Y(x)$ คือ ฟังก์ชันการแจกแจงของ Y

$F_X(x)$ คือ ฟังก์ชันการแจกแจงของ X

นอกจากนี้ กำหนดให้

$(C_{i-1}, C_i]$ คือ ช่วงของค่าความสูญเสีย

k คือ จำนวนกลุ่ม

n คือ จำนวนกรมธรรม์

ขอบเขตของการวิจัย

1. ศึกษาภายใต้ลักษณะการแจกแจงของความสูญเสีย 3 แบบ ดังนี้

1.1 การแจกแจงแบบลอกนอร์มอล ที่ μ=1,2 ,σ=0.5 (CV=53%) , μ=1 ,σ=1 (CV=131%) และ μ=1 ,σ=2 (CV=732%) [1]

1.2 การแจกแจงแบบไวบูลล์ ที่ c =0.1 ,τ=2 (CV=52%) , c =0.1 ,τ=1.5 (CV=67%) , c =0.1 ,τ=1 (CV=100%) และ c =1 ,τ=0.3 (CV=540%)

1.3 การแจกแจงแบบพาเรโต ที่ λ=10 ,α=2.5 (CV=223%) , λ=10 ,α=4 (CV=141%) และ λ=10 ,α=7 (CV=118%) [2]

---

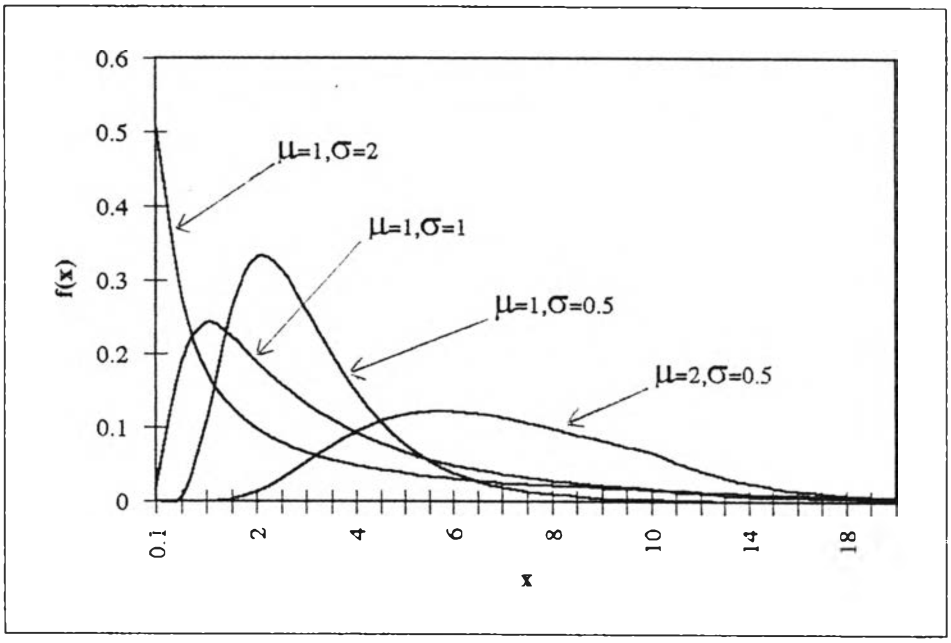
[1] ผู้วิจัยกำหนดพารามิเตอร์เพื่อให้ได้ลักษณะเส้นโค้งที่แตกต่างกัน และเห็นลักษณะการตัดข้อมูลอย่างเด่นชัด

[2] ผู้วิจัยได้ทดลองด้วยค่าพารามิเตอร์หลายๆค่า พบว่าลักษณะเส้นโค้งที่ได้แตกต่างกันน้อยมาก

2. ศึกษาเมื่อข้อมูลมีค่ารับผิดส่วนแรก (d) 4 ระดับ คือ 0.5 , 1.0 , 2.0 และ 3.0 ตามลำดับ

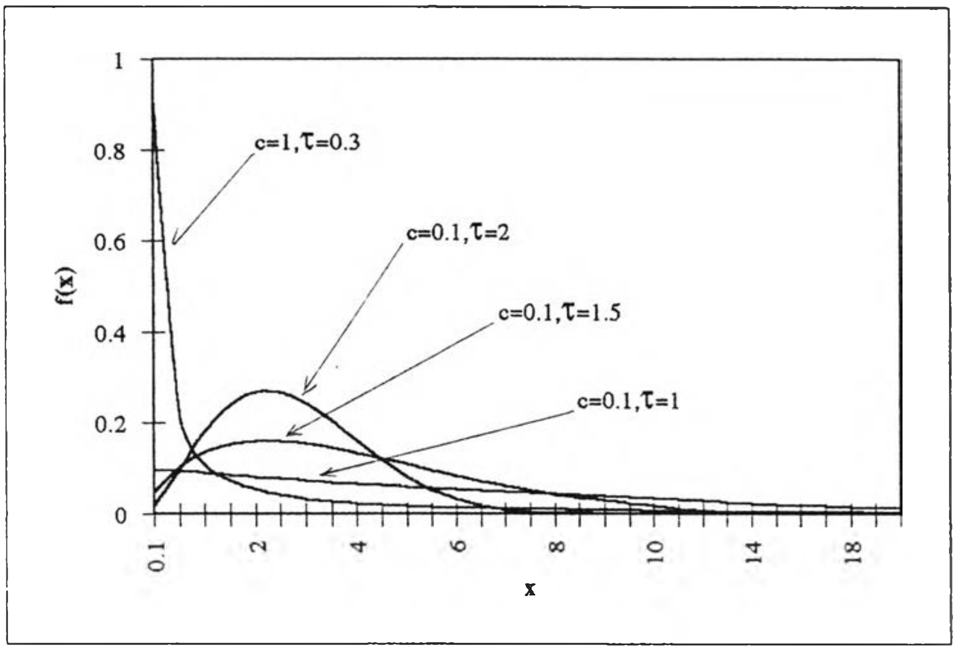
3. ขนาดตัวอย่าง (n) ที่นำมาศึกษามี 4 ระดับ คือ 100, 200, 300 และ 500 ตามลำดับ

4. จำนวนกลุ่มที่ศึกษาเท่ากับรากที่สองของขนาดตัวอย่าง $\left(\sqrt{n}\right)$ [3]

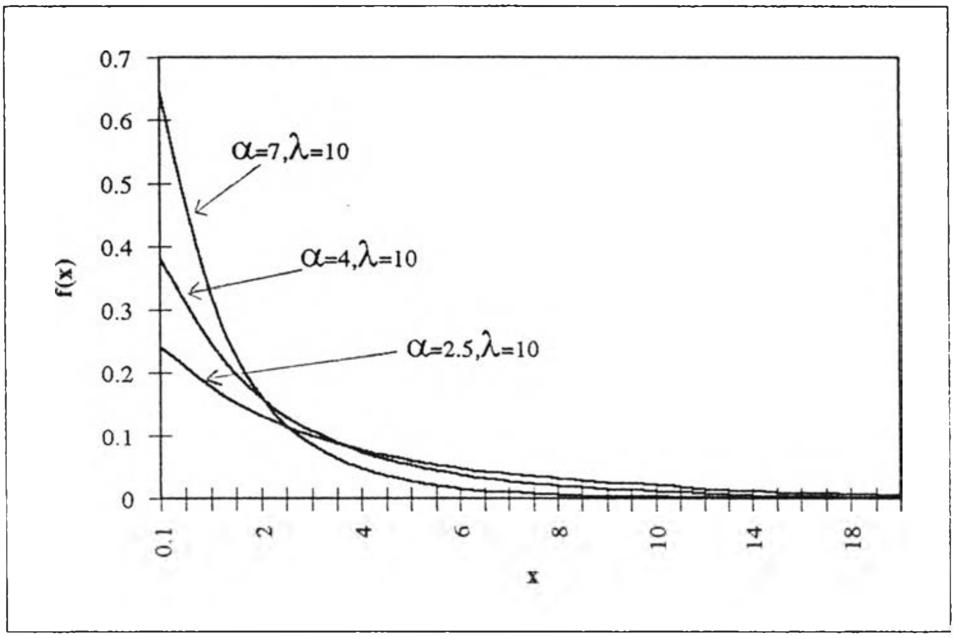


รูปที่ 1.1 แสดงการแจกแจงแบบลอกนอร์มอลที่ $\mu = 1$, $\sigma = 0.5,1,2$ และ $\mu = 2$, $\sigma = 0.5$

[3] Banks, J., and Carson, J.S. Discrete-Event System Simulation. (New jersey : Prentice-Hall , 1984), p.351.



รูปที่ 1.2 แสดงการแจกแจงแบบไวบูลล์ที่ $c = 0.1$, $\tau = 1,1.5,2$ และ $c = 1$, $\tau = 0.3$



รูปที่ 1.3 แสดงการแจกแจงแบบพาเรโตที่ $\lambda = 10$, $\alpha = 2.5,4$ และ 7

เกณฑ์การตัดสินใจ

เกณฑ์การตัดสินใจว่าการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสียด้วยวิธีใดใช้ได้ดีกว่า จะพิจารณาโดยการเปรียบเทียบค่าความคลาดเคลื่อนระหว่างค่าประมาณพารามิเตอร์กับค่าจริงในรูปของค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (Root Mean Squares Error : RMSE) ซึ่งวิธีการใดให้ค่า RMSE ต่ำกว่า จะเป็นวิธีการประมาณที่ดีกว่า

คำจำกัดความ

ค่ารับผิดส่วนแรก (Deductible) คือ ค่าความสูญเสียที่ผู้เอาประกันภัยจะต้องรับผิดชอบในเบื้องต้น บริษัทผู้รับประกันภัยจะรับผิดชอบเฉพาะกรณีที่ความสูญเสียมีค่ามากกว่าค่ารับผิดส่วนแรก มีหน่วยเป็นพันบาท

ข้อมูลถูกตัดปลายทางซ้าย (Left-Truncated Data) คือ ข้อมูลที่มีการบันทึกเฉพาะความเสียหายที่มากกว่าค่า ๆ หนึ่งหรือคือค่าตัดปลายทางซ้าย ซึ่งในที่นี้คือค่ารับผิดส่วนแรก

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเป็นแนวทางให้นักวิจัยมีผลสรุปในการเลือกวิธีประมาณค่าพารามิเตอร์สำหรับการแจกแจงความสูญเสียได้อย่างมีประสิทธิภาพและเหมาะสม เมื่อข้อมูลถูกตัดปลายทางซ้ายและเป็นแบบกลุ่ม
2. เพื่อช่วยในการคำนวณเบี้ยประกันภัย และค่าอื่นๆที่เกี่ยวข้องในธุรกิจประกันภัย ซึ่งจะทำให้การวางแผนการจัดการทางด้านการเงินของบริษัทผู้รับประกันมีประสิทธิภาพมากยิ่งขึ้น