บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการศึกษาวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงความสูญเสีย เมื่อข้อมูลเป็นแบบกลุ่มและถูกตัดปลายทางซ้าย โดยประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบไค-สแควร์ต่ำสุด และวิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์ ทั้ง 3 วิธีจะใช้การกระทำวนซ้ำ (Iterative) จนได้ค่าประมาณพารามิเตอร์ ในแต่ละวิธีการประมาณพารามิเตอร์ภายใต้สถานการณ์เมื่อขนาดตัวอย่างมี 4 ขนาด คือ 100, 200, 300 และ 500 จุดตัดปลายทางซ้ายที่ค่ารับผิดส่วนแรก มี 4 ระดับคือ 0.5, 1.0, 2.0 และ 3.0 และการแจกแจงของความสูญเสีย 3 แบบ คือ การแจกแจงแบบลอกนอร์มอล การแจกแจงแบบไวบูลล์ และการแจกแจงแบบพาเรโต โดยที่ในแต่ละสถานการณ์มีการทำซ้ำ 1,000 ครั้ง

การวิจัยครั้งนี้ใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) ในการสร้างสถานการณ์ต่างๆ ดังนั้นจะขอกล่าวถึงวิธีการจำลองโดยใช้เทคนิคมอนติ-คาร์โล แล้วจึงแสดงรายละเอียดของขั้นตอนในการวิจัยและโปรแกรมที่ใช้ในการวิจัยตามลำดับ

วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคที่ใช้แก้ปัญหาในการคำนวณทางคณิตศาสตร์นั้นมีอยู่หลายวิธี วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โลเป็นวิธีหนึ่งที่นิยมนำมาใช้แก้ปัญหากันอย่างแพร่หลายในปัจจุบัน หลักการสำคัญของการจำลองโดยใช้เทคนิคดังกล่าวคือการใช้เลขสุ่ม (Random Numbers) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

ขั้นตอนของวิธีการจำลองด้วยเทคนิคมอนติคาร์โลแบ่งออกเป็น 3 ขั้นตอนดังนี้

1. การสร้างเลขสุ่ม ลักษณะของเลขสุ่มที่ถูกสร้างขึ้นจะต้องมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) ตัวเลขสุ่มแต่ละตัวเป็นอิสระต่อกัน และมีช่วงยาวก่อนจะเกิดเลขสุ่มซ้ำ

2. การนำตัวเลขสุ่มมาประยุกต์ใช้กับปัญหาที่ต้องการศึกษา ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่นำไปผลิตเลขสุ่มที่มีการแจกแจงแบบอื่นต่อไป

3. การทดลองกระทำ เมื่อนำตัวเลขมาประยุกต์ใช้กับปัญหาที่ต้องการศึกษาได้แล้ว ขั้นต่อไปคือการทดลองโดยใช้กระบวนการของเลขสุ่ม (Random Process) มากระทำในลักษณะซ้ำๆ กัน (Replication) เพื่อหาคำตอบที่ต้องการ

แผนการทดลอง

การวิจัยครั้งนี้ กำหนดสถานการณ์ต่างๆเพื่อศึกษาวิธีการประมาณค่าพารามิเตอร์ 3 วิธี สำหรับการแจกแจงความสูญเสียเมื่อข้อมูลเป็นแบบกลุ่มและถูกตัดปลายทางซ้าย ดังนี้

1. กำหนดลักษณะการแจกแจงความสูญเสีย 3 แบบ คือ

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%)

1.2 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. กำหนดให้ข้อมูลมีค่ารับผิดส่วนแรก (d) 4 ระดับ คือ 0.5 , 1.0 , 2.0 และ 3.0

3. กำหนดขนาดตัวอย่าง (n) 4 ขนาด คือ 100, 200, 300 และ 500

4. กำหนดจำนวนกลุ่มข้อมูล (k) เท่ากับ รากที่สองของขนาดตัวอย่าง ($\sqrt{n}$)

ขั้นตอนในการวิจัย

แบ่งเป็น 4 ขั้นตอนหลัก คือ

1. จำลองข้อมูลความสูญเสีย x แบบกลุ่มและถูกตัดปลายทางซ้ายจากการแจกแจงที่กำหนด

2. การหาค่าพารามิเตอร์เริ่มต้น

3. การประมาณค่าพารามิเตอร์ ด้วยวิธีการทั้ง 3 วิธี จะได้ค่าประมาณพารามิเตอร์ของแต่ละวิธีการ

4. การหาค่าความคลาดเคลื่อนจากการประมาณค่าพารามิเตอร์ และทำการเปรียบเทียบ

รายละเอียดแต่ละขั้นตอนเป็นดังนี้

1. จำลองข้อมูลความสูญเสีย x_i แบบกลุ่มและถูกตัดปลายทางซ้าย จากการแจกแจงที่กำหนด

ขั้นตอนที่ 1 จำลองข้อมูลความสูญเสีย x_i , i = 1 , . . . , n โดยที่ $x_i > d$

เมื่อ X มีการแจกแจงแบบ

1. แบบลอกนอร์มอล (รายละเอียดในภาคผนวก ก)
2. แบบไวบูลล์ (รายละเอียดในภาคผนวก ก)
3. แบบพาเรโต (รายละเอียดในภาคผนวก ก)

ขั้นตอนที่ 2 นำข้อมูลมาเรียงลำดับและจัดกลุ่ม

โดยให้ $(C_0, C_1]$, $(C_1, C_2]$, . . . , $(C_{k-1}, C_k]$ เป็นช่วงกว้างของข้อมูล k กลุ่ม ซึ่งมีความกว้างเท่ากันทุกกลุ่ม และมีความถี่ของข้อมูลในแต่ละกลุ่มเป็น f_1 , f_2 , . . . , f_k ตามลำดับ

2. การกำหนดค่าพารามิเตอร์เริ่มต้น ในแต่ละการแจกแจงจะใช้ค่าพารามิเตอร์เริ่มต้นค่าเดียวกันทั้ง 3 วิธี

2.1 การแจกแจงแบบลอกนอร์มอล มี μ_0 และ σ_0 เป็นพารามิเตอร์เริ่มต้น ซึ่งจะหาค่าพารามิเตอร์ได้ดังนี้

$$\sigma_0 = \sqrt{\ln m_2 - 2\ln m_1}$$

$$\mu_0 = \ln m_1 - \frac{\sigma_0^2}{2}$$

โดยที่ m_1 , m_2 เป็นโมเมนต์ที่หนึ่งและสองของข้อมูล ตามลำดับ

2.2 การแจกแจงแบบไวบูลล์ มี c_0 และ τ_0 เป็นพารามิเตอร์เริ่มต้น ซึ่งจะหาค่าพารามิเตอร์ได้ดังนี้

$$\tau_0 = \frac{\ln\left[\dfrac{\ln\left(1 - F_n(C_i)\right)}{\ln\left(1 - F_n(C_j)\right)}\right]}{\ln\left(\dfrac{C_i}{C_j}\right)}$$

$$c_0 = -\frac{\ln\left(1 - F_n(C_i)\right)}{C_i^{\tau_0}}$$

โดยที่ C_i , C_j เป็นค่าขอบเขตบนของข้อมูลความสูญเสียกลุ่มที่ i , j

2.3 การแจกแจงแบบพาเรโต มี α_0 และ λ_0 เป็นพารามิเตอร์เริ่มต้น ซึ่งจะหาค่าพารามิเตอร์เริ่มต้น

$$\alpha_0 = \frac{2\left(m_1^2 - m_2\right)}{2m_1^2 - m_2}$$

$$\lambda_0 = \left(\alpha_0 - 1\right)m_1$$

โดยที่ m_1 , m_2 เป็นโมเมนต์ที่หนึ่งและสองของข้อมูล ตามลำดับ

3. การประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณ 3 วิธี มีรายละเอียดดังต่อไปนี้

3.1 วิธีการประมาณด้วยภาวะน่าจะเป็นสูงสุด

ฟังก์ชันภาวะน่าจะเป็นของความสูญเสียเมื่อข้อมูลเป็นแบบกลุ่มและถูกตัดปลายทางซ้าย แสดงได้ดังนี้

$$L = \prod_{i=1}^{k}\left[F_Y(C_i;\theta) - F_Y(C_{i-1};\theta)\right]^{f_i}$$

$$= \prod_{i=1}^{k} P_i^{f_i}$$

เมื่อ P_i เป็นความน่าจะเป็นที่จะเกิดความสูญเสียในกลุ่มที่ i สำหรับแต่ละการแจกแจง (แสดงไว้ในบทที่ 2)

θ เป็นพารามิเตอร์ใดๆ

ประมาณค่าพารามิเตอร์ θ โดยหาอนุพันธ์บางส่วนของลอกของภาวะน่าจะเป็นเทียบกับ θ และให้สมการอนุพันธ์บางส่วนเท่ากับ 0

$$\ln L = \sum_{i=1}^{k} f_i \ln P_i$$

$$\frac{\partial \ln L}{\partial \theta} = \sum_{i=1}^{k} \frac{f_i}{P_i} \frac{\partial P_i}{\partial \theta}$$

เนื่องจากสมการการหาอนุพันธ์บางส่วนของลอกของฟังก์ชันภาวะน่าจะเป็นของการแจกแจงที่ใช้ในการศึกษาครั้งนี้ไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจึงใช้วิธีนิวตัน-ราฟสันมาช่วยในการแก้สมการหาค่าประมาณพารามิเตอร์ โดยมีขั้นตอนการคำนวณดังนี้

ขั้นตอนที่ 1 การคำนวณค่า g_1 , g_2 , g_{11} , g_{12} , g_{21} , g_{22} สำหรับแต่ละการแจกแจง โดยสมการที่จะกล่าวต่อไปนี้สามารถใช้กับทุกการแจกแจง

$$g_1(\theta_1,\theta_2) = \frac{\partial \ln L}{\partial \theta_1} = \sum_{i=1}^{k} \frac{f_i}{P_i} \frac{\partial P_i}{\partial \theta_1}$$

$$g_2(\theta_1,\theta_2) = \frac{\partial \ln L}{\partial \theta_2} = \sum_{i=1}^{k} \frac{f_i}{P_i} \frac{\partial P_i}{\partial \theta_2}$$

$$g_{11}(\theta_1,\theta_2) = \frac{\partial^2 \ln L}{\partial \theta_1^2} = \sum_{i=1}^{k} f_i \left[\frac{\frac{\partial^2 P_i}{\partial \theta_1^2}}{P_i} - \left(\frac{\frac{\partial P_i}{\partial \theta_1}}{P_i} \right)^2 \right]$$

$$g_{12}(\theta_1,\theta_2) = \frac{\partial^2 \ln L}{\partial \theta_2 \partial \theta_1} = \sum_{i=1}^{k} f_i \left[\frac{\frac{\partial^2 P_i}{\partial \theta_2 \partial \theta_1}}{P_i} - \frac{\left(\frac{\partial P_i}{\partial \theta_1} \right)\left(\frac{\partial P_i}{\partial \theta_2} \right)}{P_i^2} \right]$$

$$= \frac{\partial^2 \ln L}{\partial \theta_1 \partial \theta_2} = g_{21}(\theta_1,\theta_2)$$

$$g_{22}(\theta_1,\theta_2) = \frac{\partial^2 \ln L}{\partial \theta_2^2} = \sum_{i=1}^{k} f_i \left[\frac{\frac{\partial^2 P_i}{\partial \theta_2^2}}{P_i} - \left(\frac{\frac{\partial P_i}{\partial \theta_2}}{P_i} \right)^2 \right]$$

เมื่อ θ_1, θ_2 เป็นพารามิเตอร์ของแต่ละการแจกแจง

สมการการหาอนุพันธ์อันดับหนึ่งและสองของ P_i เมื่อเทียบกับพารามิเตอร์แต่ละตัวสำหรับแต่ละการแจกแจงจะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวณค่าประมาณพารามิเตอร์ θ_1, θ_2 สำหรับแต่ละการแจกแจง โดยคำนวณจากสมการดังนี้

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \theta_{10} \\ \theta_{20} \end{bmatrix} + \begin{bmatrix} g_{11}(\theta_{10}, \theta_{20}) & g_{12}(\theta_{10}, \theta_{20}) \\ g_{21}(\theta_{10}, \theta_{20}) & g_{22}(\theta_{10}, \theta_{20}) \end{bmatrix}^{-1} \begin{bmatrix} -g_1(\theta_{10}, \theta_{20}) \\ -g_2(\theta_{10}, \theta_{20}) \end{bmatrix}$$

เมื่อ θ_{10}, θ_{20} เป็นค่าพารามิเตอร์เริ่มต้นของ θ_1, θ_2 สำหรับแต่ละการแจกแจง

จะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่ 1 และ 2 ตามลำดับ จนกระทั่งได้ค่าประมาณพารามิเตอร์ที่ต้องการ

3.2 วิธีการประมาณแบบไค-สแควร์ต่ำสุด

ตัวสถิติไค-สแควร์ของความสูญเสียเมื่อข้อมูลเป็นแบบกลุ่มและถูกตัดปลายทางซ้าย แสดงได้ดังนี้

$$\begin{aligned} Q &= \sum_{i=1}^{k} \frac{\left[f_i - n\left(F_Y(C_i;\theta) - F_Y(C_{i-1};\theta)\right)\right]^2}{n\left(F_Y(C_i;\theta) - F_Y(C_{i-1};\theta)\right)} \\ &= \sum_{i=1}^{k} \frac{\left[f_i - nP_i\right]^2}{nP_i} \end{aligned}$$

เมื่อ P_i เป็นความน่าจะเป็นที่จะเกิดความสูญเสียในกลุ่มที่ i สำหรับแต่ละการแจกแจง (แสดงไว้ในบทที่ 2)

θ เป็นพารามิเตอร์ใดๆ

ประมาณค่าพารามิเตอร์ θ โดยหาอนุพันธ์บางส่วนของตัวสถิติไค-สแควร์เทียบกับ θ และให้สมการอนุพันธ์บางส่วนเท่ากับ 0

$$\frac{\partial Q}{\partial \theta} = \sum_{i=1}^{k} \left[n - \frac{f_i^2}{nP_i^2} \right] \frac{\partial P_i}{\partial \theta}$$

เนื่องจากสมการการหาอนุพันธ์บางส่วนของตัวสถิติไค-สแควร์ของการแจกแจงที่ใช้ในการศึกษาครั้งนี้ไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจึงใช้วิธีนิวตัน-ราฟสันมาช่วยในการแก้สมการหาค่าประมาณพารามิเตอร์ โดยมีขั้นตอนการคำนวณดังนี้

ขั้นตอนที่ 1 การคำนวณค่า g_1 , g_2 , g_{11} , g_{12} , g_{21} , g_{22} สำหรับแต่ละการแจกแจง โดยสมการที่จะกล่าวต่อไปนี้สามารถใช้กับทุกการแจกแจง

$$g_1(\theta_1,\theta_2) = \frac{\partial Q}{\partial \theta_1} = \sum_{i=1}^{k}\left(n-\frac{f_i^2}{nP_i^2}\right)\frac{\partial P_i}{\partial \theta_1}$$

$$g_2(\theta_1,\theta_2) = \frac{\partial Q}{\partial \theta_2} = \sum_{i=1}^{k}\left(n-\frac{f_i^2}{nP_i^2}\right)\frac{\partial P_i}{\partial \theta_2}$$

$$g_{11}(\theta_1,\theta_2) = \frac{\partial^2 Q}{\partial \theta_1^2} = \sum_{i=1}^{k}\left[\left(n-\frac{f_i^2}{nP_i^2}\right)\frac{\partial^2 P_i}{\partial \theta_1^2}+\frac{2f_i^2}{nP_i^3}\left(\frac{\partial P_i}{\partial \theta_1}\right)^2\right]$$

$$g_{12}(\theta_1,\theta_2) = \frac{\partial^2 Q}{\partial \theta_2 \partial \theta_1} = \sum_{i=1}^{k}\left[\left(n-\frac{f_i^2}{nP_i^2}\right)\frac{\partial^2 P_i}{\partial \theta_2 \partial \theta_1}+\frac{2f_i^2}{nP_i^3}\left(\frac{\partial P_i}{\partial \theta_1}\right)\left(\frac{\partial P_i}{\partial \theta_2}\right)\right]$$

$$= \frac{\partial^2 Q}{\partial \theta_1 \partial \theta_2} = g_{21}(\theta_1,\theta_2)$$

$$g_{22}(\theta_1,\theta_2) = \frac{\partial^2 Q}{\partial \theta_2^2} = \sum_{i=1}^{k}\left[\left(n-\frac{f_i^2}{nP_i^2}\right)\frac{\partial^2 P_i}{\partial \theta_2^2}+\frac{2f_i^2}{nP_i^3}\left(\frac{\partial P_i}{\partial \theta_2}\right)^2\right]$$

เมื่อ θ_1,θ_2 เป็นพารามิเตอร์ของแต่ละการแจกแจง

สมการการหาอนุพันธ์อันดับหนึ่งและสองของ P_i เมื่อเทียบกับพารามิเตอร์แต่ละตัวสำหรับแต่ละการแจกแจงจะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวณค่าประมาณพารามิเตอร์ θ_1,θ_2 สำหรับแต่ละการแจกแจง โดยคำนวณจากสมการดังนี้

$$\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}=\begin{bmatrix}\theta_{10}\\ \theta_{20}\end{bmatrix}+\begin{bmatrix}g_{11}(\theta_{10},\theta_{20}) & g_{12}(\theta_{10},\theta_{20})\\ g_{21}(\theta_{10},\theta_{20}) & g_{22}(\theta_{10},\theta_{20})\end{bmatrix}^{-1}\begin{bmatrix}-g_1(\theta_{10},\theta_{20})\\ -g_2(\theta_{10},\theta_{20})\end{bmatrix}$$

เมื่อ θ_{10},θ_{20} เป็นค่าพารามิเตอร์เริ่มต้นของ θ_1,θ_2 สำหรับแต่ละการแจกแจง

จะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่ 1 และ 2 ตามลำดับ จนกระทั่งได้ค่าประมาณพารามิเตอร์ที่ต้องการ

3.3 วิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์

ตัวสถิติของคราเมอร์-วอน ไมส์ของความสูญเสียเมื่อข้อมูลเป็นแบบกลุ่มและถูกตัดปลายทางซ้าย แสดงได้ดังนี้

$$K = \sum_{i=1}^{k}\left[F_n(C_i)-F_Y(C_i;\theta)\right]^2$$

เมื่อ θ เป็นพารามิเตอร์ใดๆ

ประมาณค่าพารามิเตอร์ θ โดยหาอนุพันธ์บางส่วนของตัวสถิติของคราเมอร์-วอน ไมส์เทียบกับ θ และให้สมการอนุพันธ์บางส่วนเท่ากับ 0

$$\frac{\partial K}{\partial\theta} = -2\sum_{i=1}^{k}\left[F_n(C_i)-F_Y(C_i;\theta)\right]\frac{\partial F_Y(C_i)}{\partial\theta}$$

เนื่องจากสมการการหาอนุพันธ์บางส่วนของตัวสถิติของคราเมอร์-วอน ไมส์ของการแจกแจงที่ใช้ในการศึกษาครั้งนี้ไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจึงใช้วิธีนิวตัน-ราฟสันมาช่วยในการแก้สมการหาค่าประมาณพารามิเตอร์ โดยมีขั้นตอนการคำนวณดังนี้

ขั้นตอนที่ 1 การคำนวณค่า g_1 , g_2 , g_{11} , g_{12} , g_{21} , g_{22} สำหรับแต่ละการแจกแจง โดยสมการที่จะกล่าวต่อไปนี้สามารถใช้กับทุกการแจกแจง

$$g_1(\theta_1,\theta_2) = \frac{\partial K}{\partial \theta_1} = -2\sum_{i=1}^{k}\left[F_n(C_i)-F_Y(C_i)\right]\frac{\partial F_Y(C_i)}{\partial \theta_1}$$

$$g_2(\theta_1,\theta_2) = \frac{\partial K}{\partial \theta_2} = -2\sum_{i=1}^{k}\left[F_n(C_i)-F_Y(C_i)\right]\frac{\partial F_Y(C_i)}{\partial \theta_2}$$

$$g_{11}(\theta_1,\theta_2) = \frac{\partial^2 K}{\partial \theta_1^2} = 2\sum_{i=1}^{k}\left[\left(\frac{\partial F_Y(C_i)}{\partial \theta_1}\right)^2 - \left(F_n(C_i)-F_Y(C_i)\right)\frac{\partial^2 F_Y(C_i)}{\partial \theta_1^2}\right]$$

$$g_{12}(\theta_1,\theta_2) = \frac{\partial^2 K}{\partial \theta_2 \partial \theta_1} = 2\sum_{i=1}^{k}\left[\left(\frac{\partial F_Y(C_i)}{\partial \theta_1}\right)\left(\frac{\partial F_Y(C_i)}{\partial \theta_2}\right) - \left(F_n(C_i)-F_Y(C_i)\right)\frac{\partial^2 F_Y(C_i)}{\partial \theta_2 \partial \theta_1}\right]$$

$$= \frac{\partial^2 K}{\partial \theta_1 \partial \theta_2} = g_{21}(\theta_1,\theta_2)$$

$$g_{22}(\theta_1,\theta_2) = \frac{\partial^2 K}{\partial \theta_2^2} = 2\sum_{i=1}^{k}\left[\left(\frac{\partial F_Y(C_i)}{\partial \theta_2}\right)^2 - \left(F_n(C_i)-F_Y(C_i)\right)\frac{\partial^2 F_Y(C_i)}{\partial \theta_2^2}\right]$$

เมื่อ θ_1,θ_2 เป็นพารามิเตอร์ของแต่ละการแจกแจง

สมการการหาค่าอนุพันธ์อันดับหนึ่งและสองเมื่อเทียบกับพารามิเตอร์แต่ละตัวสำหรับแต่ละการแจกแจงจะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวณค่าประมาณพารามิเตอร์ θ_1,θ_2 สำหรับแต่ละการแจกแจง โดยคำนวณจากสมการดังนี้

$$\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} = \begin{bmatrix}\theta_{10}\\ \theta_{20}\end{bmatrix} + \begin{bmatrix} g_{11}(\theta_{10},\theta_{20}) & g_{12}(\theta_{10},\theta_{20}) \\ g_{21}(\theta_{10},\theta_{20}) & g_{22}(\theta_{10},\theta_{20})\end{bmatrix}^{-1}\begin{bmatrix}-g_1(\theta_{10},\theta_{20})\\ -g_2(\theta_{10},\theta_{20})\end{bmatrix}$$

เมื่อ θ_{10},θ_{20} เป็นค่าพารามิเตอร์เริ่มต้นของ θ_1,θ_2 สำหรับแต่ละการแจกแจง

จะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่ 1 และ 2 ตามลำดับ จนกระทั่งได้ค่าประมาณพารามิเตอร์ที่ต้องการ

4. การหาค่าความคลาดเคลื่อนจากการประมาณ โดยนำค่าประมาณพารามิเตอร์มาเปรียบเทียบกับพารามิเตอร์จริง เพื่อคำนวณหาค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (RMSE) เนื่องจากในการทดลองได้จำลองข้อมูลซ้ำๆกันจำนวน 1,000 ครั้ง ในแต่ละสถานการณ์ ให้ i แทนรอบที่ทำซ้ำ $i = 1,2,...,1000$ เนื่องจากพารามิเตอร์ที่ต้องการประมาณมี 2 พารามิเตอร์ ดังนั้นจะต้องหาค่าเฉลี่ยของค่า RMSE ของการประมาณทั้ง 2 พารามิเตอร์ การหาค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองหาได้จากสูตรดังนี้

$$RMSE\left(\theta_p\right) = \left[\sum_{i=1}^{1000} \frac{\left(\hat{\theta}_{pi} - \theta_p\right)^2}{1000}\right]^{\frac{1}{2}} \quad ; \qquad p = 1,2$$

$$RMSE = \frac{RMSE\left(\theta_1\right) + RMSE\left(\theta_2\right)}{2}$$

จากนั้นจึงนำไปเปรียบเทียบเพื่อหาว่าวิธีการใดให้ค่า RMSE ของการประมาณค่าพารามิเตอร์ต่ำที่สุดจะเป็นวิธีที่ประมาณค่าพารามิเตอร์ของแต่ละสถานการณ์ได้ดีที่สุด

โปรแกรมที่ใช้ในการวิจัย

โปรแกรมที่ใช้ในการวิจัยทั้งหมด เขียนด้วยภาษาฟอร์แทรน (FORTRAN) โดยใช้กับเครื่อง AMDAHL 5860 ซึ่งในแต่ละสถานการณ์ของการทดลอง ลักษณะการทำงานของโปรแกรมจะเหมือนกัน สำหรับรายละเอียดของโปรแกรมจะแสดงไว้ในภาคผนวก ข ซึ่งจะเป็นโปรแกรมการทำงานของแต่ละวิธีการ คือ วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด วิธีการประมาณแบบไค-สแควร์ต่ำสุด และวิธีการประมาณแบบระยะต่ำสุดของคราเมอร์-วอน ไมส์

รูปที่ ธ.1 แสดงผังงานสำหรับหาค่าความคลาดเคลื่อนจากการประมาณด้วยวิธีการทั้ง 3 วิธี