บทที่ 3
การดำเนินการวิจัย

การวิจัยครั้งนี้มีลักษณะเป็นการวิจัยเชิงทดลอง โดยต้องการศึกษาและเปรียบเทียบตัวประมาณค่าพารามิเตอร์แสดงตำแหน่ง และพารามิเตอร์แสดงสเกล ของข้อมูลที่มีการแจกแจงแบบปกติที่มีค่าผิดปกติ โดยตัวประมาณค่าที่นำมาศึกษาได้แก่

1. ตัวประมาณค่าพารามิเตอร์แสดงตำแหน่ง ได้แก่
   1.1 ตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงตำแหน่ง
   1.2 ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปสำหรับพารามิเตอร์แสดงตำแหน่ง
   1.3 ตัวประมาณค่าที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงตำแหน่ง
2. ตัวประมาณค่าพารามิเตอร์แสดงสเกล ได้แก่
   2.1 ตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงสเกล
   2.2 ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปสำหรับพารามิเตอร์แสดงสเกล
   2.3 ตัวประมาณค่าที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงสเกล

ผู้วิจัยจะทำการเปรียบเทียบความเที่ยงตรงของตัวประมาณค่าพารามิเตอร์แสดงตำแหน่งและพารามิเตอร์แสดงสเกล โดยจะทำการเปรียบเทียบความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณที่ได้จากทุกวิธี ซึ่งแต่ละวิธีจะประมาณค่าภายใต้ข้อมูลที่มีการแจกแจงแบบปกติที่มีค่าผิดปกติตามที่กำหนด โดยให้มีขนาดตัวอย่าง คือ 20, 30, 50 และ 70 ซึ่งจำแนกลักษณะข้อมูลที่มีค่าผิดปกติดังนี้

(1) ข้อมูลที่มีค่าผิดปกติอยู่ในระดับไม่รุนแรง

(2) ข้อมูลที่มีค่าผิดปกติอยู่ในระดับรุนแรง

ในการวิจัยครั้งนี้ ผู้วิจัยได้อาศัยเทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Method) สร้างสถานการณ์การทดลอง โดยใช้โปรแกรมภาษาฟอร์แทรน (Fortran) สำหรับแผนการทดลอง ดังนั้นส่วนแรกผู้วิจัยจะกล่าวถึงวิธีจำลองโดยใช้เทคนิคมอนติคาร์โล และรายละเอียดของขั้นตอนการวิจัยตามลำดับ ส่วนรายละเอียดของโปรแกรมที่ใช้ในการวิจัยจะแสดงไว้ในภาคผนวก

3.1 วิธีจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคที่ใช้แก้ปัญหาในการคำนวณทางคณิตศาสตร์นั้นมีอยู่หลายวิธี วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โลเป็นวิธีหนึ่งที่นิยมมาใช้แก้ปัญหากันอย่างแพร่หลายในปัจจุบัน ซึ่งหลักการของการจำลองโดยใช้เทคนิคดังกล่าวจะใช้เลขสุ่ม (Random Numbers) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

ขั้นตอนของวิธีการจำลองด้วยเทคนิคมอนติคาร์โลที่ใช้ในปัจจุบันแบ่งได้เป็น 3 ขั้นตอนดังนี้คือ

1. การสร้างตัวเลขสุ่ม การใช้ตัวเลขสุ่มเป็นสิ่งที่สำคัญมากในเทคนิคนี้ ทั้งนี้เพราะว่าหลักการจำลองด้วยเทคนิคมอนติคาร์โลนั้น จะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหา โดยลักษณะของตัวเลขสุ่มที่นำมาใช้จะมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) สำหรับวิธีการสร้างตัวเลขสุ่มมีผู้เสนอไว้หลายวิธี แต่วิธีที่ดีนั้นลักษณะของเลขสุ่มที่ถูกสร้างขึ้นจะต้องมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) ตัวเลขสุ่มแต่ละตัวเป็นอิสระต่อกัน และมีช่วงยาวก่อนจะเกิดเลขสุ่มซ้ำ (มีวัฏจักรยาว)

2. การนำตัวเลขสุ่มมาประยุกต์ใช้กับปัญหาที่ต้องการศึกษา ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่จะนำไปผลิตตัวเลขสุ่มที่มีการแจกแจงแบบอื่นต่อไป

3. การทดลองกระทำ เมื่อนำตัวเลขสุ่มมาประยุกต์ให้เข้ากับปัญหาที่ต้องการศึกษาได้แล้ว ขั้นตอนต่อไปคือ การทดลองโดยใช้กระบวนการของการสุ่ม (Random Process) มากระทำในลักษณะซ้ำๆกันหลายๆครั้ง เพื่อหาคำตอบที่ต้องการ

3.2 แผนการทดลอง

การวิจัยครั้งนี้ทำการประมาณค่าพารามิเตอร์แสดงตำแหน่ง (μ) และพารามิเตอร์แสดงสเกล (σ) เมื่อข้อมูลมีการแจกแจงแบบปกติที่มีค่าผิดปกติ โดยเปรียบเทียบตัวประมาณที่มีความแกร่ง ตัวประมาณค่าที่ได้จากวิธีบูตสแตรป และตัวประมาณค่าที่ไม่เอนเอียงสำหรับประมาณค่าพารามิเตอร์แสดงตำแหน่งและพารามิเตอร์แสดงสเกลภายใต้ลักษณะข้อมูลที่มีค่าผิดปกติมีดังนี้

(1) เลือกกลุ่มตัวอย่างจากประชากร โดยกำหนดให้ประชากรมีฟังก์ชันความหนาแน่นดังนี้

(1.1) การแจกแจงแบบปกติ โดยที่กำหนดพารามิเตอร์ μ เท่ากับ 2 และ σ เท่ากับ 2, 4, 6, 8, 10

(1.2) การแจกแจงแบบปกติปลอมปนในตำแหน่งซึ่งมีฟังก์ชันความหนาแน่นอยู่ในรูป

(1.2.1) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(1, 2^2) + p.N(C, 2^2) \quad ; \quad -\infty<x<\infty$$

(1.2.2) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(1, 2^2) + p.L(\theta, \beta) \quad ; -\infty<x<\infty, \beta>0$$

โดยกำหนดอัตราส่วนการปลอมปน (p) เท่ากับ 0.05, 0.10, 0.15 และ 0.20 กำหนดสเกลแฟกเตอร์ (C) เท่ากับ 3, 10 กำหนด $\theta = C$, $\beta = 2$

(1.3) การแจกแจงแบบปกติปลอมปนในสเกล ซึ่งมีฟังก์ชันความหนาแน่นอยู่ในรูป

(1.3.1) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(2, 2^2) + p.N(2, C^2 2^2) \quad ; \quad -\infty<x<\infty$$

(1.3.2) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(2, 2^2) + p.(\theta, C\beta) \quad ; \quad -\infty<x<\infty, -\infty<\theta<\infty, \beta>0$$

(1.3.3) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(2, 2^2) + p.\,Expo(\lambda) \quad ; \quad -\infty<x<\infty, \lambda>0$$

โดยกำหนดอัตราส่วนการปลอมปน (p) เท่ากับ 0.05, 0.1, 0.15 และ 0.20 กำหนดสเกลแฟกเตอร์ (C) เท่ากับ 3, 10 กำหนด $\theta = 2$, $\lambda = \frac{1}{2C}$, $\beta = 2$

(2) ในทุกการแจกแจงของประชากร จะศึกษาในกรณีที่มีขนาดตัวอย่างเท่ากับ 20, 30, 50 และ 70

(3) ทำการเปรียบเทียบความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าทุกตัวเพื่อหาตัวประมาณที่เหมาะสมสำหรับการประมาณค่าพารามิเตอร์แสดงตำแหน่ง และพารามิเตอร์แสดงสเกล

3.3 ขั้นตอนการดำเนินการวิจัย

ขั้นตอนการวิจัยมีดังนี้ คือ

(1) สร้างข้อมูลให้มีการแจกแจงตามลักษณะที่ต้องการศึกษา และมีขนาดตัวอย่างตามที่ต้องการศึกษา

(2) คำนวณค่าตัวประมาณพารามิเตอร์แสดงตำแหน่ง และค่าตัวประมาณพารามิเตอร์แสดงสเกล จากตัวประมาณค่าที่มีความแกร่ง ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปและตัวประมาณที่ไม่เอนเอียง

(3) คำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าพารามิเตอร์แสดงตำแหน่ง และความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าพารามิเตอร์แสดงสเกล

(4) เปรียบเทียบความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณเพื่อพิจารณาเลือกตัวประมาณค่าที่เหมาะสมในแต่ละสถานการณ์

(5) สรุปผลการวิจัยในแต่ละสถานการณ์

สำหรับรายละเอียดของแต่ละขั้นตอนมีดังต่อไปนี้

(1) สร้างข้อมูลให้มีการแจกแจงตามลักษณะที่ต้องการศึกษา และมีขนาดตัวอย่างตามที่ต้องการศึกษา

การสร้างลักษณะการแจกแจงของประชากรให้มีลักษณะตามที่ต้องการศึกษานั้นใช้โปรแกรมฟอร์แทรน โดยที่สร้างลักษณะการแจกแจงต่างๆ จะต้องใช้เลขสุ่ม (Random Number) ซึ่งมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) เป็นพื้นฐานในการสร้างการแจกแจงต่างๆ มีดังนี้

(1.1) การแจกแจงแบบปกติ

การแจกแจงแบบปกติมีฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad ; \quad -\infty < x < \infty, \sigma > 0, -\infty < x < \infty$$

การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติใช้วิธีของ Box และ Muller โดยการสร้างเลขสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน ที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 ใน function dmerm(dmean, sigma) สำหรับรายละเอียดแสดงไว้ในภาคผนวก

(1.2) การแจกแจงแบบปกติปลอมปนในตำแหน่งซึ่งมีฟังก์ชันความหนาแน่นอยู่ในรูป

(1.2.1) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(1, 2^2) + p.N(C,2^2) \quad ; \quad -\infty<x<\infty$$

การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติปลอมปนในตำแหน่งฟังก์ชันที่ใช้ คือ function ncale (c, p, dmean, sigma, w, n) สำหรับรายละเอียดแสดงไว้ในภาคผนวก

(1.2.2) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(1, 2^2) + p.L(C,\beta) \quad ; \quad -\infty<x<\infty, \beta>0$$

การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติปลอมปนในตำแหน่งฟังก์ชันที่ใช้ คือ function LLscale (c, p, dmean, sigma, w, n) สำหรับรายละเอียดแสดงไว้ในภาคผนวก

(1.3) การแจกแจงแบบปกติปลอมปนในสเกล ซึ่งมีฟังก์ชันความหนาแน่นอยู่ในรูป

(1.3.1) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(2, 2^2) + p.N(2,C^2 2^2) \quad ; \quad -\infty<x<\infty$$

การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติปลอมปนในสเกล ฟังก์ชันที่ใช้ คือ function scale (c, p, dmean, sigma, w, n) สำหรับรายละเอียดแสดงไว้ในภาคผนวก

(1.3.2) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(2, 2^2) + p.L(\theta,C.\beta) \quad ; \quad -\infty<x<\infty, -\infty<\theta<\infty, \beta>0$$

การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติปลอมปนในสเกล ฟังก์ชันที่ใช้ คือ function Lscale (c, p, dmean, sigma, w, n) สำหรับรายละเอียดแสดงไว้ในภาคผนวก

(1.3.3) ฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = (1-p).N(2, 2^2) + p.Expo(\lambda) \quad ; \quad -\infty<x<\infty, \lambda>0$$

การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติปลอมปนในสเกล ฟังก์ชันที่ใช้ คือ function escale (c, p, dmean, sigma, w, n) สำหรับรายละเอียดแสดงไว้ในภาคผนวก

(1.4) เมื่อสร้างข้อมูลให้มีลักษณะการแจกแจงตามที่ต้องการศึกษาแล้ว จะนำค่าข้อมูลมาตรวจสอบค่าผิดปกติโดยใช้กราฟแบบ Box-Whisker เพื่อหาค่าสเกลแฟกเตอร์ ให้กำหนดระดับค่าผิดปกติ ซึ่งเงื่อนไขการตรวจสอบได้แสดงไว้ในบทที่ 1

(1.5) เมื่อตรวจสอบค่าผิดปกติโดยใช้กราฟแบบ Box-Whisker พบว่าค่าสเกลแฟกเตอร์เท่ากับ 3 สอดคล้องกับกรณีที่ค่าผิดปกติอยู่ในระดับไม่รุนแรง และค่าสเกลแฟกเตอร์เท่ากับ 10 สอดคล้องกับกรณีที่ค่าผิดปกติอยู่ในระดับรุนแรง

(2) การคำนวณหาค่าตัวประมาณพารามิเตอร์แสดงตำแหน่ง และค่าตัวประมาณพารามิเตอร์แสดงสเกล จากตัวประมาณค่าที่มีความแกร่ง ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปและตัวประมาณที่ไม่เอนเอียง

(2.1) ตัวประมาณค่าพารามิเตอร์แสดงตำแหน่งเป็น

(2.1.1) ตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงตำแหน่ง

การคำนวณตัวประมาณค่าที่มีความแกร่งคำนวณได้จากสูตร

กำหนดให้ $d_i = |X_i - m|$

$$MAD = med\{d_1, d_2, \ldots, d_n\}$$

เมื่อ X_i แทน ข้อมูลตัวที่ i

m แทน ค่ามัธยฐานของข้อมูล X_i

MAD แทน ค่ามัธยฐานของข้อมูล d_i

โดยตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงตำแหน่งคือ

$$M_{bi}(X) = \frac{\sum_{i=1}^{n} X_i W_{bi}(U_i)}{\sum_{i=1}^{n} W_{bi}(U_i)}$$

เมื่อฟังก์ชันถ่วงน้ำหนัก คือ

$$W_{bi}(U_i) = (1-U_i^2)^2 \quad ; \quad |U_i| < 1$$

$$= 0 \quad ; \quad U_i \text{ เป็นค่าอื่น ๆ}$$

โดยที่ $U_i = \dfrac{(X_i - m)}{c\,MAD}$

c แทน ค่าคงที่ที่เป็นบวก

(2.1.2) ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปสำหรับพารามิเตอร์แสดงตำแหน่ง

ตัวประมาณที่ได้จากวิธีบูตสแตรปมีขั้นตอนดังนี้

ขั้นตอนที่ 1 ทำการสร้างข้อมูลให้มีการแจกแจงตามที่ต้องการศึกษา $X_i \sim F$; $i = 1, 2, \ldots, n$ เมื่อ F เป็นการแจกแจงที่ต้องการศึกษา

ขั้นตอนที่ 2 จากค่า $X_1, \ldots, X_n$ ทำการสุ่มตัวอย่างขนาด n แบบใส่คืนได้ $X_1^*, \ldots, X_n^*$ เมื่อ X_j^* คือตัวอย่างที่สุ่มได้ตัวที่ j จากข้อมูล $X_1, \ldots, X_n$

ขั้นตอนที่ 3 นำค่าข้อมูล $X_1^*, \ldots, X_n^*$ ไปคำนวณหาค่าเฉลี่ย และค่ามัธยฐาน

1. ค่าเฉลี่ยของข้อมูล $X_1^*, \ldots, X_n^*$ คือ

$$\bar{X}^* = \frac{\sum_{i=1}^{n} X_i^*}{n}$$

2. ค่ามัธยฐานของข้อมูล $X_1^*, \ldots, X_n^*$ คือ

2.1 ทำการเรียงข้อมูล $X_1^*, \ldots, X_n^*$ จากน้อยไปมาก

2.2 ค่ามัธยฐานของข้อมูล $X_1^*, \ldots, X_n^*$ คือ

$$\widetilde{X}^* = \frac{X^*_{(\frac{n}{2})} + X^*_{(\frac{n}{2}+1)}}{2}$$

เมื่อ $X^*_{(\frac{n}{2})}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}$

$X^*_{(\frac{n}{2}+1)}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}+1$

ขั้นตอนที่ 4 ทำตามขั้นตอนที่ 2 - 3 เท่ากับจำนวนครั้งที่ต้องการทำบูตสแตรป

ขั้นตอนที่ 5 หาค่าเฉลี่ยของค่าประมาณที่ได้จากขั้นตอนที่ 3 ดังนั้นตัวประมาณที่ได้จากวิธีบูตสแตรป คือ

$$BMEAN = \frac{\sum_{j=1}^{B} \overline{X}^*_j}{B}$$

$$BMEDIAN = \frac{\sum_{j=1}^{B} \widetilde{X}^*_j}{B}$$

เมื่อ B แทน จำนวนครั้งของการทำบูตสแตรป

(2.1.3) ตัวประมาณค่าที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงตำแหน่ง

ตัวประมาณที่ไม่เอนเอียงในการวิจัยครั้งนี้ได้พิจารณาตัวประมาณค่า 2 ตัว คือ

1. ค่าเฉลี่ยของข้อมูล คือ

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

2. ค่ามัธยฐานของข้อมูล

2.1 ทำการเรียงข้อมูลจากน้อยไปมาก

2.2 ค่ามัธยฐานของข้อมูล คือ

$$\widetilde{X} = \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2}$$

เมื่อ

$X_{(\frac{n}{2})}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}$

$X_{(\frac{n}{2}+1)}$ แทน ค่าข้อมูลตัวที่ $\frac{n}{2}+1$

(2.2)ตัวประมาณค่าพารามิเตอร์แสดงสเกล แบ่งเป็น

(2.2.1) ตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงสเกล

กำหนดให้ $d_i = |X_i - m|$

$MAD = med\{d_1, d_2, \ldots, d_n\}$

เมื่อ X_i แทน ข้อมูลตัวที่ i

m แทน ค่ามัธยฐานของข้อมูล X_i

MAD แทน ค่ามัธยฐานของข้อมูล d_i

โดยตัวประมาณค่าที่มีความแกร่งสำหรับพารามิเตอร์แสดงสเกลคือ

$$S_{mb,c} = \frac{n}{(n-1)^{\frac{1}{2}}} \frac{\left[\sum_{|U_i|<1}((X_i - m)^2 W_{bi}(U_i))^2\right]^{\frac{1}{2}}}{\left|\sum_{|U_i|<1} W_{bi}(U_i)\right|}$$

เมื่อฟังก์ชันถ่วงน้ำหนัก คือ

$$W_{bi}(U_i) = (1-U_i^2)^2 \quad ; \quad |U_i| < 1$$

$$= 0 \quad ; \quad U_i \text{ เป็นค่าอื่น ๆ}$$

โดยที่ $U_i = \dfrac{(X_i - m)}{c\,MAD}$

c แทน ค่าคงที่ที่เป็นบวก

$$S_{msi} = \frac{n * 2.1 * MAD}{(n-1)^{\frac{1}{2}}} \tan gent^{-1}\left(\frac{\left[\sum_{|U_i|<\pi}\sin^2(U_i)\right]^{\frac{1}{2}}}{\left|\sum_{|U_i|\le\pi}\cos(U_i)\right|}\right)$$

เมื่อฟังก์ชันถ่วงน้ำหนักอยู่ในรูป

$$\Psi(U_i) = \sin(U_i) \quad ; \quad |U_i| \le \pi$$

$$= 0 \quad ; \quad U_i \text{ เป็นค่าอื่น ๆ}$$

โดยที่ $U_i = \dfrac{(X_i - m)}{c\,MAD}$

c แทน ค่าคงที่ที่เป็นบวก

(2.2.2)ตัวประมาณค่าที่ได้จากวิธีบูตสแตรปสำหรับพารามิเตอร์แสดงสเกล

ขั้นตอนที่ 1 ทำการสร้างข้อมูลให้มีการแจกแจงตามที่ต้องการศึกษา $X_i \sim F$; $i = 1, 2, \ldots, n$ เมื่อ F เป็นการแจกแจงที่ต้องการศึกษา

ขั้นตอนที่ 2 จากค่า $X_1, \ldots, X_n$ ทำการสุ่มตัวอย่างขนาด n แบบใส่คืนได้ $X_1^*, \ldots, X_n^*$ เมื่อ X_j^* คือตัวอย่างที่สุ่มได้ตัวที่ j จากข้อมูล $X_1, \ldots, X_n$

ขั้นตอนที่ 3 นำค่าข้อมูล $X_1^*, \ldots, X_n^*$ ไปคำนวณหาค่าตัวประมาณที่เอนเอียง และไม่เอนเอียง

1. ตัวประมาณค่าที่เอนเอียง คือ

$$SD^* = \sqrt{\frac{\sum_{i=1}^{n}(X_i^* - \bar{X}^*)^2}{n-1}}$$

2. ตัวประมาณค่าที่ไม่เอนเอียง คือ

$$UNSD^* = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{\sum_{i=1}^{n}(X_i^* - \bar{X}^*)^2}{2}}$$

เมื่อ X_i แทน ค่าข้อมูลตัวที่ i

$$\bar{X}^* = \frac{\sum_{i=1}^{n} X_i^*}{n}$$

$$\Gamma(n) = (n-1)!$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

ขั้นตอนที่ 4 ทำตามขั้นตอนที่ 2 - 3 เท่ากับจำนวนครั้งที่ต้องการทำบูตสแตรป

ขั้นตอนที่ 5 หาค่าเฉลี่ยของค่าประมาณที่ได้จากขั้นตอนที่ 3 ดังนั้นตัวประมาณที่ได้จากวิธีบูตสแตรป คือ

$$BSD = \frac{\sum_{i=1}^{B} SD^*}{B}$$

$$BUNSD = \frac{\sum_{i=1}^{B} UNSD^*}{B}$$

เมื่อ B แทน จำนวนครั้งของการทำบูตสแตรป

(2.2.3) ตัวประมาณที่เอนเอียงสำหรับพารามิเตอร์แสดงสเกล

ตัวประมาณค่าที่เอนเอียงสำหรับพารามิเตอร์แสดงสเกล คือ

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

ตัวประมาณค่าที่ไม่เอนเอียงสำหรับพารามิเตอร์แสดงสเกล คือ

$$UNSD = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{2}}$$

เมื่อ Xi แทน ค่าข้อมูลตัวที่ i

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

(3)คำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าพารามิเตอร์แสดงตำแหน่งและพารามิเตอร์แสดงสเกล จากตัวประมาณที่มีความแกร่ง ตัวประมาณที่ได้จากวิธีบูตสแตรป และตัวประมาณที่ไม่เอนเอียงภายใต้ข้อมูลที่มีการแจกแจงแบบปกติที่มีค่าผิดปกติ โดยมีการทำซ้ำ 500 รอบ ดังนี้

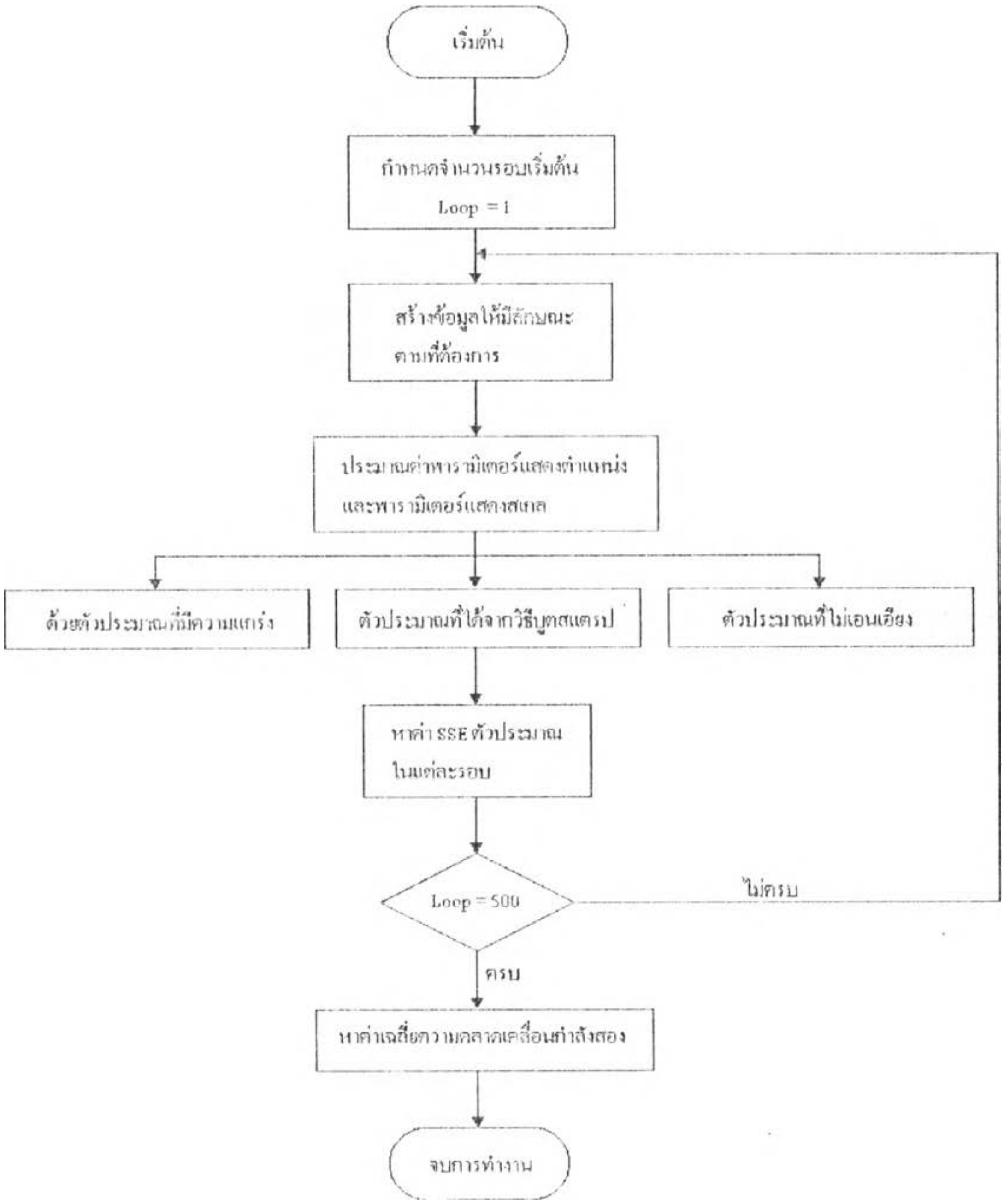
$$MSE = \frac{\sum_{i=1}^{500}(\hat{\theta}_i - \theta)^2}{500}$$

เมื่อ θ แทน ค่าพารามิเตอร์ที่ต้องการประมาณ

$\hat{\theta}_i$ แทนค่าประมาณของพารามิเตอร์รอบที่ i

ในการคำนวณความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าทุกตัวภายใต้สถานการณ์ที่กำหนดโดยทุกสถานการณ์ผู้วิจัยจะทำการทดลองซ้ำ 500 ครั้ง จนครบทุกสถานการณ์ ซึ่งมีขั้นตอนสรุปเป็นผังงานดังรูปที่ 3.1

รูปที่ 3.1 รูปแสดงผังงานสำหรับการหาความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ



(4)ทำการเปรียบเทียบความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณค่าโดยการเปรียบเทียบนั้น ถ้าตัวประมาณค่าตัวใดให้ค่าเฉลี่ยความคลาดเคลื่อนต่ำสุดในแต่ละสถานการณ์ แสดงว่าตัวประมาณค่านั้นเหมาะสมที่ใช้ในการประมาณค่าพารามิเตอร์แสดงตำแหน่ง และพารามิเตอร์แสดงสเกลในสถานการณ์นั้น

(5) สรุปผลการวิจัยในแต่ละสถานการณ์