บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณค่าสัมประสิทธิ์ในสมการถดถอยเชิงเส้นระหว่างวิธีกำลังสองต่ำสุด (Least Squares Method) และวิธีบูตสแตรป (Bootstrap Method) โดยจะศึกษาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) ของตัวประมาณและฟังก์ชั่นเชิงเส้นของตัวประมาณทั้ง 2 วิธีแล้วทำการเปรียบเทียบในลักษณะประสิทธิภาพสัมพัทธ์ (Relative Efficiency) ของตัวประมาณที่ได้จากวิธีกำลังสองต่ำสุด และวิธีบูตสแตรปโดยสร้าง (generate) ข้อมูลให้ความคลาดเคลื่อน (ε) มีการแจกแจงแบบปกติ (Normal Distribution) แบบยูนิฟอร์ม(Uniform Distribution) แบบโลจิสติค (Logistic Distribution) แบบดับเบิ้ลเอ็กซ์โปเนนเชียล (Double Ex - ponential Distribution) และแบบปกติปลอมปน (Scale Contaminated Normal Distribution) ทั้งนี้เนื่องจากการแจกแจงแบบต่าง ๆ ดังกล่าว ยกเว้นการแจกแจงแบบปกติเป็นการแจกแจงที่มีลักษณะการกระจายไปทางหางมากหรือมีหางยาวกว่าปกติ ซึ่งเป็นลักษณะที่สนใจศึกษา ยกเว้นการแจกแจงแบบยูนิฟอร์ม ที่ศึกษาเนื่องจากต้องการทราบผลสรุปในกรณีที่ ε มีการแจกแจงไม่เป็นทั้งแบบปกติ และไม่มีลักษณะการกระจายไปทางหางมากหรือหางยาวกว่าปกติ และที่สนใจศึกษากรณีที่ ε มีการแจกแจงแบบปกติด้วย เพราะว่าการประมาณค่าพารามิเตอร์โดยวิธีบูตสแตรปไม่จำเป็นต้องมีข้อกำหนดเกี่ยวกับการแจกแจงของ ε คือ ε อาจมีลักษณะการแจกแจงแบบใดก็ได้ที่ไม่ทราบ ซึ่งอาจเป็นได้ทั้งแบบปกติและไม่ปกติ สำหรับรูปแบบของการแจกแจงแบบปกติปลอมปนจะทำการศึกษาเมื่อมีเปอร์เซ็นต์การปลอมปนเป็น 1% 5% 10% และ 25% และมีค่าสเกลแฟคเตอร์ 2 ระดับคือ 3 และ 10 ซึ่งถ้าสเกลแฟคเตอร์มีค่าสูงจะทำให้เกิดค่าผิดปกติได้ และจากการศึกษาที่ผ่านมาสเกลแฟคเตอร์ที่มีค่าน้อยกว่า 3 จะมีโอกาสของการ

เกิดค่าผิดปกติน้อย ส่วนสเกลแฟคเตอร์ที่มีค่ามากกว่า 10 จะมีโอกาสของการเกิดค่าผิดปกติมาก จึงทำการศึกษาเพียง 2 ระดับดังกล่าว ขนาดตัวอย่างและจำนวนตัวแปรอิสระที่สนใจศึกษาในครั้งนี้มี 3 ชุด คือ

- ชุดที่ 1 (n_1) = 50 จำนวนตัวแปรอิสระ (p_1) = 5
- ชุดที่ 2 (n_2) = 10 จำนวนตัวแปรอิสระ (p_2) = 4
- ชุดที่ 3 (n_3) = 5 จำนวนตัวแปรอิสระ (p_3) = 3

ทั้งนี้ เทคนิคที่ใช้ในการหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ วิธีมอนติคาร์โล ซึ่งเป็นเทคนิคอย่างหนึ่งที่ใช้แก้ปัญหาในการคำนวณทางคณิตศาสตร์ โดยเฉพาะในกรณีที่ไม่สามารถหาคำตอบโดยวิธีทางทฤษฎีได้

เนื่องจากวิธีมอนติคาร์โลเป็นเทคนิคที่ใช้ในการวิจัยครั้งนี้ ดังนั้นในตอนแรกของบทนี้จะกล่าวถึงวิธีมอนติคาร์โลก่อน แล้วจึงแสดงรายละเอียดของแผนการทดลองขั้นตอนการวิจัยและโปรแกรมที่ใช้ในการวิจัยตามลำดับ ซึ่งรายละเอียดต่าง ๆ เป็นดังนี้

3.1 วิธีมอนติคาร์โล (Monte Carlo Method)

เทคนิคที่ใช้สำหรับแก้ปัญหาในการคำนวณทางคณิตศาสตร์นั้นมีอยู่หลายวิธี วิธีมอนติ - คาร์โล ก็เป็นเทคนิคอย่างหนึ่งที่จะใช้แก้ปัญหานี้ได้ และเป็นวิธีที่นิยมใช้กันอย่างแพร่หลายในปัจจุบัน Hammersley และ Handscomb (1964:2) กล่าวว่า วิธีมอนติคาร์โลเป็นสาขาหนึ่งของคณิตศาสตร์เชิงทดลอง ซึ่งหลักการของวิธีมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่ม (Random Number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา

ในการวิจัยครั้งนี้ ได้ใช้วิธีมอนติคาร์โลในการสร้างข้อมูลที่มีสภาพการแจกแจงตามที่ต้องการและใช้ในการประมาณค่าสัมประสิทธิ์ความถดถอย ซึ่งขั้นตอนของวิธีมอนติคาร์โล แบ่งเป็น 3 ขั้นตอน ดังนี้

1. การสร้างตัวเลขสุ่ม (Random Number) การใช้ตัวเลขสุ่มเป็นสิ่งที่สำคัญมากในวิธีมอนติคาร์โล ทั้งนี้เพราะว่าหลักการของวิธีมอนติคาร์โลนั้น จะใช้ตัวเลขสุ่มมาช่วยในการ

หาคำตอบของปัญหา ลักษณะของตัวเลขสุ่มจะมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) สำหรับวิธีการสร้างตัวเลขสุ่มมีผู้เสนอไว้หลายวิธี แต่วิธีที่ดีนั้นลักษณะของตัวเลขสุ่มที่เกิดขึ้นจะต้องมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) และเป็นอิสระต่อกัน

2. การประยุกต์ปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่อาจจะมีขั้นตอนอื่นหลาย ๆ ขั้นตอน ซึ่งขั้นตอนเหล่านี้มีบางขั้นตอนที่ต้องใช้ตัวเลขสุ่ม

3. การทดลองกระทำ เมื่อประยุกต์ปัญหาให้ใช้กับตัวเลขสุ่มได้แล้ว ขั้นต่อไปก็คือการทดลอง โดยใช้กระบวนการของการสุ่ม (Random Process) มากระทำในลักษณะที่ซ้ำ ๆ กัน (Replication) เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

3.2 แผนการทดลอง

ในการวิจัยครั้งนี้ เป็นการศึกษาเปรียบเทียบประสิทธิภาพของตัวประมาณ เมื่อสุ่มตัวอย่างจากประชากรที่มีการแจกแจงของความคลาดเคลื่อนแบบต่าง ๆ โดยใช้สัญลักษณ์แทนการแจกแจงเป็นดังนี้คือ

N แทนการแจกแจงแบบปกติ

U แทนการแจกแจงแบบยูนิฟอร์ม

L แทนการแจกแจงแบบโลจิสติค

D แทนการแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียล

C แทนการแจกแจงแบบปกติปลอมปน

สำหรับการแจกแจงแบบปกติปลอมปน กำหนดเปอร์เซ็นต์ของการปลอมปนเป็น 1% 5% 10% และ 25% ที่สเกลแฟคเตอร์ 2 ระดับคือ 3 และ 10 ซึ่งแสดงรายละเอียดดังตาราง 3.1

ตาราง 3.1 แสดงค่าสเกลแฟคเตอร์และเปอร์เซ็นต์การปลอมปนที่ใช้ในการวิจัย

(3,1)	(3,5)	(3,10)	(3,25)
(10,1)	(10,5)	(10,10)	(10,25)

3.3 ขั้นตอนในการวิจัย

ขั้นตอนในการวิจัยแบ่งออกเป็น 4 ขั้นตอน คือ

1. การสร้างโปรแกรมย่อย (Subroutine) สำหรับสร้างการแจกแจงของความคลาดเคลื่อน (ε) ตามที่กำหนด

2. การสร้างข้อมูล (x,y) ที่มีความสัมพันธ์กันเชิงเส้นตรง

3. การหาค่าประมาณของค่าสัมประสิทธิ์ความถดถอย

4. การหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ

ซึ่งรายละเอียดสำหรับแต่ละขั้นตอนเป็นดังนี้

3.3.1 การสร้างโปรแกรมย่อยสำหรับสร้างการแจกแจงของความคลาดเคลื่อน (ε) ตามที่กำหนด

การสร้างลักษณะการแจกแจงของความคลาดเคลื่อนทุกรูปแบบตามที่กำหนดไว้ในแผนการทดลองนั้นใช้โปรแกรมภาษาฟอร์แทรน(Fortran 77) โดยใช้กับเครื่อง VAX-11/750 ซึ่งการสร้างลักษณะการแจกแจงแบบต่าง ๆ นั้น จะต้องใช้ตัวเลขสุ่ม (Random Number) ซึ่งมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) เป็นพื้นฐานในการสร้าง สำหรับโปรแกรมที่ใช้สร้างตัวเลขสุ่มในการวิจัยครั้งนี้ใช้วิธีที่ White และ Schmidt (1975:421) เสนอไว้ ซึ่งรายละเอียดในการสร้างการแจกแจงแบบต่าง ๆ เป็นดังนี้

3.3.1.1 การแจกแจงแบบปกติ

โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบปกติในการวิจัยครั้งนี้ใช้วิธีของ Gauss ซึ่งเป็นวิธีที่สร้างการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 1 สำหรับค่าเฉลี่ยและความแปรปรวนอื่น จะใช้วิธีแปลงข้อมูลในรูป X' = DMEAN + (SIGMA)X โดย DMEAN และ $(SIGMA)^2$ คือ ค่าเฉลี่ยและค่าความแปรปรวนที่ต้องการ สำหรับรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL NORMAL (DMEAN,SIGMA,X) ค่า DMEAN และ $(SIGMA)^2$ เป็นค่าพารามิเตอร์ ซึ่งกำหนดค่าเฉลี่ยและความแปรปรวนซึ่งจะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์คือค่า X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น DMEAN และความแปรปรวนเป็น $(SIGMA)^2$

3.3.1.2 การแจกแจงแบบยูนิฟอร์ม

โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบยูนิฟอร์ม ใช้วิธี Inverse Transformation ซึ่งรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL UNIFRM (A,B,X) โดยค่า A และ B เป็นค่าพารามิเตอร์ที่กำหนดขึ้นเพื่อให้ได้ค่าเฉลี่ยและความแปรปรวนตามที่กำหนด ค่า A และ B จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์ที่ได้คือ X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบยูนิฟอร์มในช่วง (A,B) ที่มีค่าเฉลี่ยเป็น $\frac{A+B}{2}$ และความแปรปรวนเป็น $\frac{(B-A)^2}{12}$

3.3.1.3 การแจกแจงแบบโลจิสติค

โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบโลจิสติค ใช้วิธี Inverse Transformation ซึ่งรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL LOGIST (ALPHA,BETA,X) ค่า ALPHA และ BETA เป็นค่าพารามิเตอร์ที่กำหนดขึ้นเพื่อให้ได้ค่าเฉลี่ยและความแปรปรวนตามที่กำหนด ค่า ALPHA และ BETA จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์ที่ได้คือ X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบโลจิสติคที่มีค่าเฉลี่ยเป็น ALPHA และความแปรปรวนเป็น $\frac{1}{3}\pi^2 (BETA)^2$

I17499327

3.3.1.4 การแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียล

โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียล ใช้วิธี Inverse Transformation ซึ่งรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL DOUBLE (ALPHA,BETA,X) ค่า ALPHA,BETA เป็นค่าพารามิเตอร์ที่กำหนดขึ้นเพื่อให้ได้ค่าเฉลี่ย และความแปรปรวนตามที่กำหนด ค่า ALPHA และ BETA จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์ที่ได้คือ X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบดับเบิ้ลเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเป็น ALPHA และความแปรปรวนเป็น $2(BETA)^2$

3.3.1.5 การแจกแจงแบบปกติปลอมปน

โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบปกติปลอมปนนั้น ใช้วิธีแปลงข้อมูลจากการแจกแจงแบบปกติ โดยพิจารณาการแจกแจงแบบปกติปลอมปนของตัวแปร X ซึ่งมาจาก $F = (1-p)\ N(\mu,\sigma^2) + p\ N(\mu,c^2\sigma^2)$ เมื่อ p เป็นเปอร์เซ็นต์การปลอมปน c เป็นสเกลแฟคเตอร์ที่จะทำให้เกิดค่าผิดปกติ ($c > 0$) หมายความว่าตัวแปร X จะมาจากการแจกแจงแบบ $N(\mu,\sigma^2)$ ด้วยความน่าจะเป็น $(1-p)$ และมาจากการแจกแจงแบบ $N(\mu,c^2\sigma^2)$ ด้วยความน่าจะเป็น p สำหรับรายละเอียดแสดงไว้ในภาคผนวก ก การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL SCNRML (C,P,DMEAN,SIGMA,X) ค่า C และ P เป็นค่าของสเกลแฟคเตอร์และเปอร์เซ็นต์การปลอมปน ส่วนค่า DMEAN และ $(SIGMA)^2$ เป็นค่าเฉลี่ยและความแปรปรวนที่กำหนด ซึ่งค่า C P DMEAN และ SIGMA จะถูกส่งมาจากโปรแกรมหลัก ส่วนผลลัพธ์ที่ได้คือ X ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติปลอมปนที่มีค่าเฉลี่ยเป็น DMEAN และความแปรปรวนเป็น $(SIGMA)^2$

3.3.2 การสร้างข้อมูล (x,y) ที่มีความสัมพันธ์กันเชิงเส้นตรง

การสร้างข้อมูล (x,y) ที่มีความสัมพันธ์กันเชิงเส้นตรงนั้น ในการวิจัยครั้งนี้จะทำการสร้างค่า X ซึ่งเป็นค่าคงที่ก่อน แล้วจึงสร้างค่า Y ที่มีความสัมพันธ์เชิงเส้นตรงกับ

ค่า X ตามรูปแบบดังนี้คือ $Y = X\beta + \varepsilon$

เมื่อ Y เป็นเวคเตอร์ของตัวแปรตามขนาด n x 1

X เป็นเมตริกซ์ของตัวแปรอิสระขนาด n x p ซึ่งมีค่าคงที่

β เป็นเวคเตอร์ของค่าพารามิเตอร์ที่ถูกกำหนดขึ้นมาขนาด p x 1

ε เป็นเวคเตอร์ของความคลาดเคลื่อนขนาด n x 1 ซึ่งมีรูปแบบการแจกแจงเป็นแบบต่างๆ ดังที่กล่าวมาแล้วในหัวข้อ 3.3.1

n เป็นขนาดของข้อมูลตัวอย่าง

p เป็นจำนวนตัวแปรอิสระ

ซึ่งในการสร้างข้อมูลจะเริ่มจากการกำหนดขนาดของตัวอย่างที่ต้องการศึกษา จำนวนตัวแปรอิสระ ค่าของพารามิเตอร์ ค่าเฉลี่ย ค่าความแปรปรวน และลักษณะการแจกแจงของ ε แล้วสร้างค่าคงที่ X จากนั้นก็จะใช้คำสั่งสุ่มตัวอย่างเพื่อสร้าง ε ที่มีลักษณะการแจกแจงเป็นแบบต่าง ๆ ตามที่ต้องการศึกษา แล้วจึงสร้างค่า Y ตามรูปแบบความสัมพันธ์ดังกล่าว

3.3.3 การหาค่าประมาณของค่าสัมประสิทธิ์ความถดถอย

3.3.3.1 โดยวิธีกำลังสองต่ำสุด (Least Square Method)

จากข้อมูล (x,y) ที่ได้จะนำมาหาค่าประมาณของพารามิเตอร์ที่ทำให้ผลบวกของค่าความคลาดเคลื่อนกำลังสอง ($\sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \varepsilon'\varepsilon$) มีค่าต่ำสุดจะได้ว่า $\hat{\beta} = (X'X)X'Y$ เป็นค่าประมาณของค่าสัมประสิทธิ์ความถดถอยที่ได้จากวิธีนี้

3.3.3.2 โดยวิธีบูตสแตรป (Bootstrap Method)

ทำนองเดียวกันจากข้อมูล (x,y) ที่ได้จะนำมาหาค่าประมาณของพารามิเตอร์ ซึ่งมีขั้นตอนการคำนวณดังนี้

1. จากค่า $\hat{\beta}$ ที่ได้โดยวิธีกำลังสองต่ำสุด นำหาคำนวณหาค่า $\hat{\varepsilon} = Y - X\hat{\beta}$

2. จากค่า $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n$ ทำการสุ่มตัวอย่างขนาด n แบบใส่คืน (with replacement) สำหรับรายละเอียดในการสุ่มแสดงไว้ในภาคผนวก ก ได้ $\varepsilon^*_1, \varepsilon^*_2, \ldots, \varepsilon^*_3$ เมื่อ ε^*_i คือตัวอย่างที่สุ่มได้ตัวที่ i จาก $\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n$

3. สร้างรูปแบบ (model) ใหม่ ดังนี้

$$Y^* = X\hat{\beta} + \varepsilon^*$$

4. คำนวณหาค่า $\hat{\beta}^*$ โดยวิธีกำลังสองต่ำสุด ได้ว่า

$$\hat{\beta}^* = (X'X)X'Y^*$$

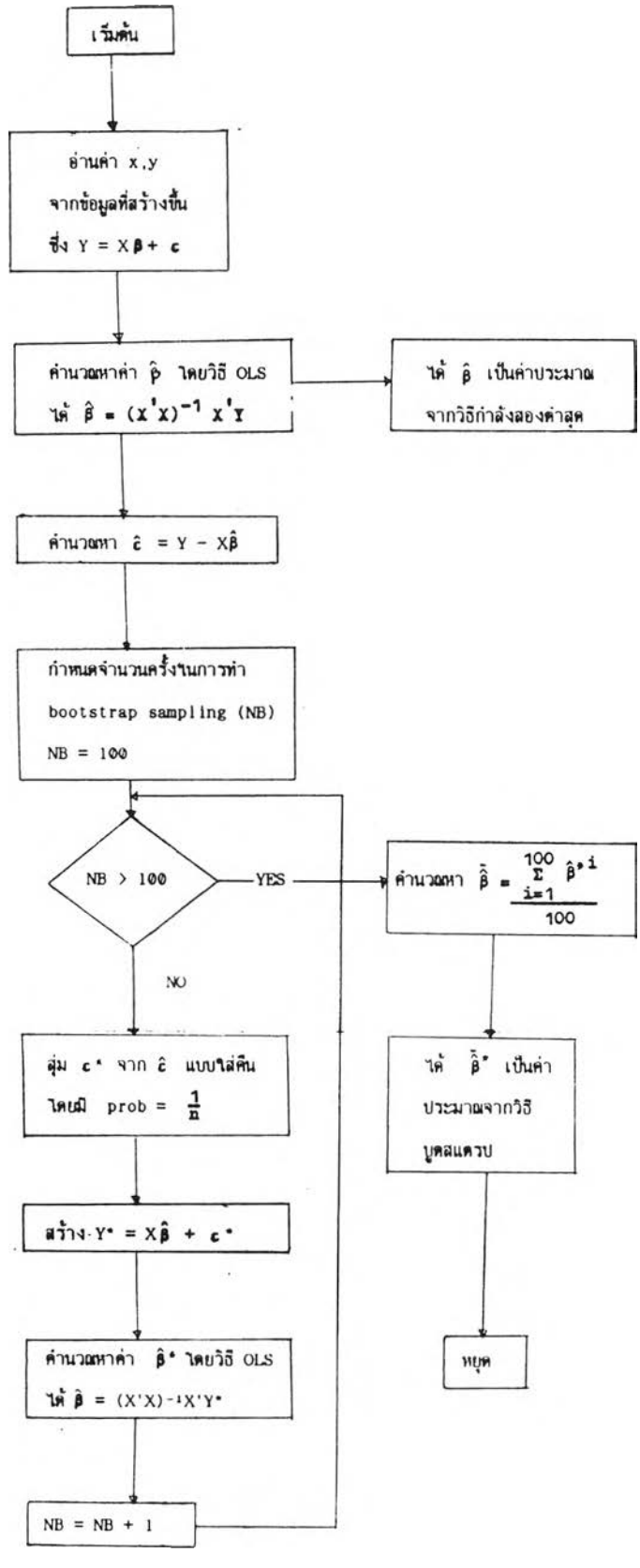
5. กระทำตามขั้นในข้อ 2-4 100 ครั้ง จะได้

$$\hat{\beta}^{*1}, \hat{\beta}^{*2}, \ldots, \hat{\beta}^{*100}$$

6. คำนวณหาค่า $\bar{\hat{\beta}}^* = \sum_{i=1}^{100} \hat{\beta}^{*i} / 100$

และใช้ $\bar{\hat{\beta}}^*$ เป็นค่าประมาณของสัมประสิทธิ์ความถดถอยที่ได้จากวิธีบูตสแตรป ซึ่งพอจะสรุปเป็นแผนผังงานได้ดังรูป 3.1

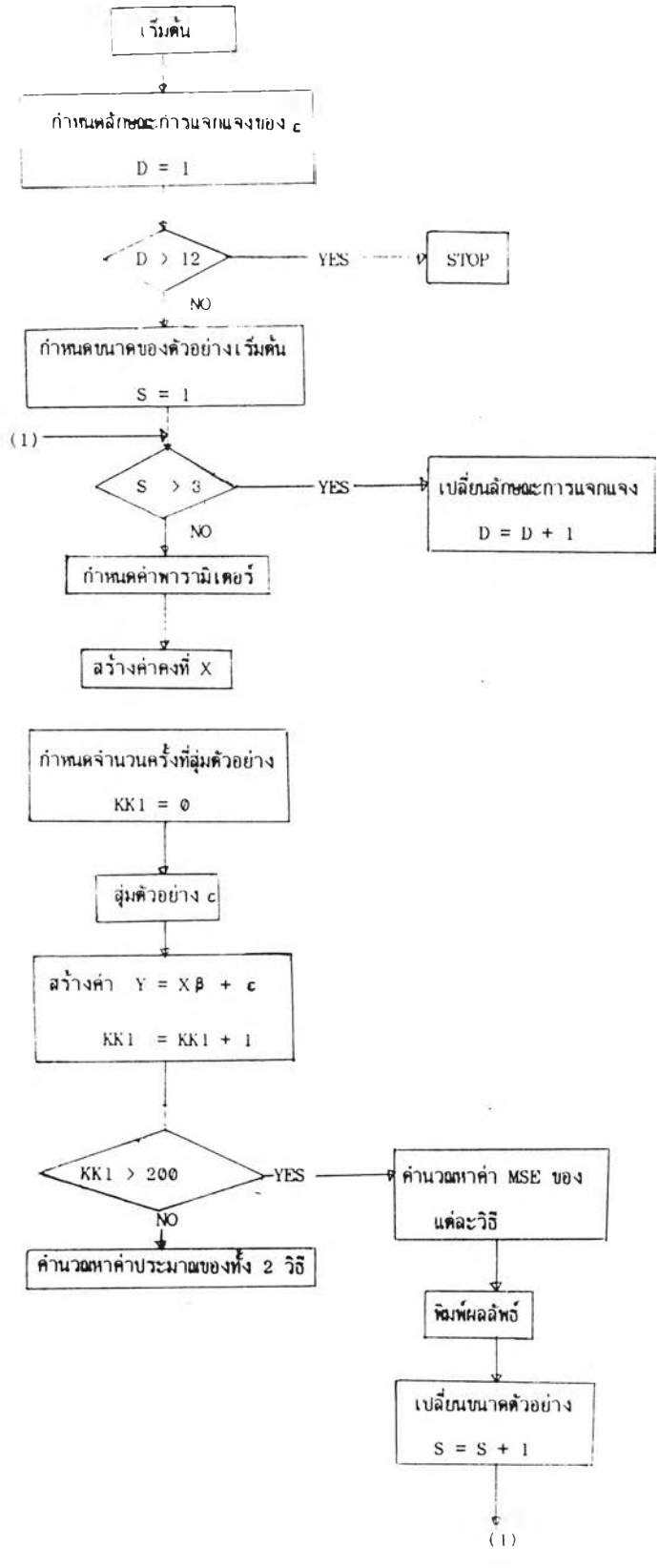
รูป 3.1 แสดงแผนผังงานสำหรับการหาค่าประมาณของค่าสัมประสิทธิ์ความถดถอยโดยวิธีกำลังสองต่ำสุด และโดยวิธีบูตสแตรป



3.3.4 การหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ (MSE)

เมื่อคำนวณหาค่าประมาณของพารามิเตอร์ทั้ง 2 วิธีได้แล้ว ก็นำค่าประมาณที่ได้ลบออกจากค่าพารามิเตอร์แล้วยกกำลังสองบวกสะสมไว้ เพื่อใช้ในการคำนวณหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแต่ละตัว พร้อมทั้งคำนวณหาผลบวกของตัวประมาณแล้วนำมาลบออกจากผลบวกของพารามิเตอร์ เพื่อใช้ในการหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของฟังก์ชั่นเชิงเส้นตัวประมาณที่ได้ ทำการสุ่มตัวอย่างความคลาดเคลื่อนชุดใหม่ที่มีการแจกแจงแบบเดียวกัน แต่ยังคงใช้ค่าคงที่ X ชุดเดิม จนกระทั่งครบ 200 ครั้ง จึงทำการคำนวณหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณแต่ละตัว และฟังก์ชั่นเชิงเส้นของตัวประมาณที่ได้จากแต่ละวิธี ในการแจกแจงเดียวกันนี้จะเปลี่ยนขนาดตัวอย่างจนครบทุกขนาดที่ต้องการศึกษา โดยในแต่ละขนาดตัวอย่างจะสุ่มตัวอย่างซ้ำกัน 200 ครั้ง แล้วคำนวณหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเช่นเดียวกัน ขั้นต่อไปจะเปลี่ยนลักษณะการแจกแจงของความคลาดเคลื่อน (ε) จนครบทุกการแจกแจง โดยแต่ละการแจกแจงของความคลาดเคลื่อนจะใช้ขนาดตัวอย่างครบทุกขนาด และแต่ละขนาดของตัวอย่างจะคำนวณหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ซึ่งพอจะสรุปเป็นแผนผังงานได้ดังรูป 3.2

รูป 3.2 แสดงแผนผังงานสำหรับการหาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย(MSE)ของตัวประมาณแต่ละตัว และฟังก์ชันเชิงเส้นของตัวประมาณที่ได้จากวิธีกำลังสองต่ำสุด และวิธีบูตสแตรป



3.4 โปรแกรมที่ใช้ในการวิจัย

โปรแกรมที่ใช้ในการวิจัยทั้งหมดเขียนด้วยภาษาฟอร์แทรน 77 ซึ่งแสดงรายละเอียดไว้ในภาคผนวก ข โดยขั้นตอนและลักษณะการทำงานของโปรแกรมเป็นดังตาราง 3.2

ตาราง 3.2 แสดงลักษณะการทำงานของโปรแกรที่ใช้ในการวิจัย

ขั้นตอน	ลักษณะการทำงาน	โปรแกรมย่อยที่เรียกใช้
1	คำนวณหาค่าประมาณของค่าสัมประสิทธิ์ความถดถอยโดยวิธีกำลังสองต่ำสุด	-โปรแกรมสร้างตัวเลขสุ่ม -โปรแกรมสร้างค่าคงที่ X -โปรแกรมสร้างลักษณะการแจกแจงของความคลาดเคลื่อน -โปรแกรมการคำนวณหาค่าประมาณของสัมประสิทธิ์ความถดถอย
2	คำนวณหาค่าประมาณของค่าสัมประสิทธิ์ความถดถอยโดยวิธีบูตสแตรป	-โปรแกรมการสุ่มตัวอย่างแบบใส่คืน -โปรแกรมการคำนวณหาค่าประมาณของสัมประสิทธิ์ความถดถอย
3	คำนวณหาค่าคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ และฟังก์ชั่นเชิงเส้นของตัวประมาณที่ได้จากทั้ง 2 วิธี	-จัดพิมพ์ค่าต่างๆ ตามต้องการ

3.5 ข้อมูลที่ใช้ในการวิจัย

3.5.1 ลักษณะการแจกแจงของความคลาดเคลื่อน (ε) มี 5 ลักษณะ แต่ละลักษณะจะมีค่าเฉลี่ยเท่ากับ 0 และมีค่าความแปรปรวนเท่ากับ 100 สำหรับการแจกแจงแบบปกติปลอมปนจะมีเปอร์เซนต์การปลอมปนเป็น 1% 5% 10% 25% และมีค่าสเกลแฟคเตอร์ 2 ระดับคือ 3 และ 10 รายละเอียดสามารถแสดงได้ดังตาราง 3.3

ตาราง 3.3 ลักษณะการแจกแจงของความคลาดเคลื่อน (ε) และค่าพารามิเตอร์ที่ใช้ในการวิจัย

ค่าพารามิเตอร์	ลักษณะการแจกแจง				
	N	U	D	L	C
	0	0	0	0	0
	100	100	100	100	100
A		-17.325			
B		17.325			
ALPHA			0	0	
BETA			7.071	5.511	
c_1					3
c_2					10
p_1					1
p_2					5
p_3					10
p_4					25

เมื่อตัวอักษร N U D L และ C ใช้แทนลักษณะการแจกแจงในหัวข้อ 3.2

μ	แทน	ค่าเฉลี่ย
σ^2	แทน	ความแปรปรวน
ALPHA	แทน	ค่า ALPHA
BETA	แทน	ค่า BETA
c_i	แทน	ค่าสเกลแฟคเตอร์ชุดที่ i ; i = 1,2
p_i	แทน	เปอร์เซ็นต์การปลอมปนชุดที่ i ; i = 1,2,3,4
A	แทน	ค่าขอบเขตล่าง
B	แทน	ค่าขอบเขตบน