บทที่ 3
วิธีดำเนินการวิจัย

3.1 แผนการทดลอง

ในการวิจัยครั้งนี้ได้กำหนดสถานการณ์ต่างๆที่ต้องการจะศึกษา ดังนี้

5. จำนวนตัวแปรอิสระที่ใช้ในการศึกษาเท่ากับ 3 และ 6
6. ขนาดตัวอย่างที่ใช้ในการศึกษาเท่ากับ 10 , 15 , 20 , 25 และ 30 ตามลำดับ
7. ลักษณะการแจกแจงของค่าความคลาดเคลื่อนที่จะศึกษามีดังนี้
   7.1 การแจกแจงปกติ จะศึกษาในกรณีที่มีค่าเฉลี่ย (μ) เท่ากับ 0 และ 1 ค่าความแปรปรวน (σ^2) เท่ากับ 1, 4 และ 9 ตามลำดับ
   7.2 การแจกแจงปกติปลอมปน จะศึกษาในกรณีที่ค่าเฉลี่ย (μ) เท่ากับ 0 ค่าความแปรปรวน (σ^2) เท่ากับ 1 ค่าสเกลแฟคเตอร์ (C) เท่ากับ 10 และเปอร์เซ็นต์การปลอมปน (P) เป็น 25%, 50% และ 75% ตามลำดับ
   7.3 การแจกแจงแกมมา จะศึกษาในกรณีที่มีค่า β = 1 และ α = 0.5, 1, 2 และ 4 ตามลำดับ

3.2 ขั้นตอนการวิจัย

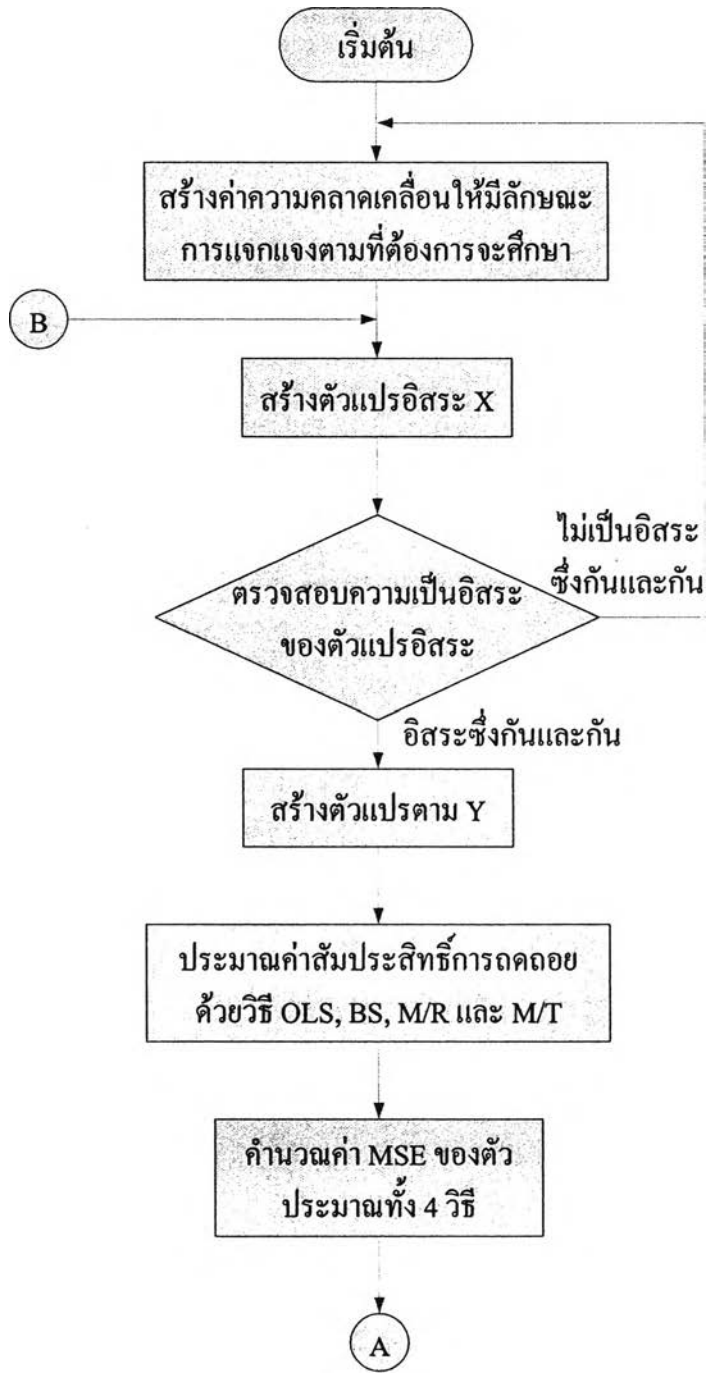
ขั้นตอนในการวิจัยมีดังนี้

1. สร้างข้อมูลของความคลาดเคลื่อนให้มีลักษณะตามที่ต้องการศึกษา
2. สร้างข้อมูลของตัวแปรอิสระ (X) โดยที่ตัวแปรอิสระแต่ละตัวมีการแจกแจงปกติด้วยค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 และตัวแปรอิสระแต่ละตัวจะต้องเป็นอิสระซึ่งกันและกัน โดยทำการทดสอบความเป็นอิสระด้วยค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's correlation coefficient) ที่ระดับนัยสำคัญ α=0.05
3. สร้างข้อมูลของตัวแปรตาม (y) จากรูปแบบความสัมพันธ์ $y = X\beta + \varepsilon$ โดยกำหนดให้ β เป็นค่าคงที่ใดๆกล่าวคือสำหรับในกรณีที่ตัวแปรอิสระเท่ากับ 3จะกำหนดให้ $\beta = (1,1,1,1)'$ และในกรณีที่ตัวแปรอิสระเท่ากับ 6 จะกำหนดให้ $\beta = (1,1,1,1,1,1,1)'$ ซึ่งค่าความคลาดเคลื่อน ε มีการแจกแจงตามที่กำหนดไว้ในขอบเขตของการวิจัย

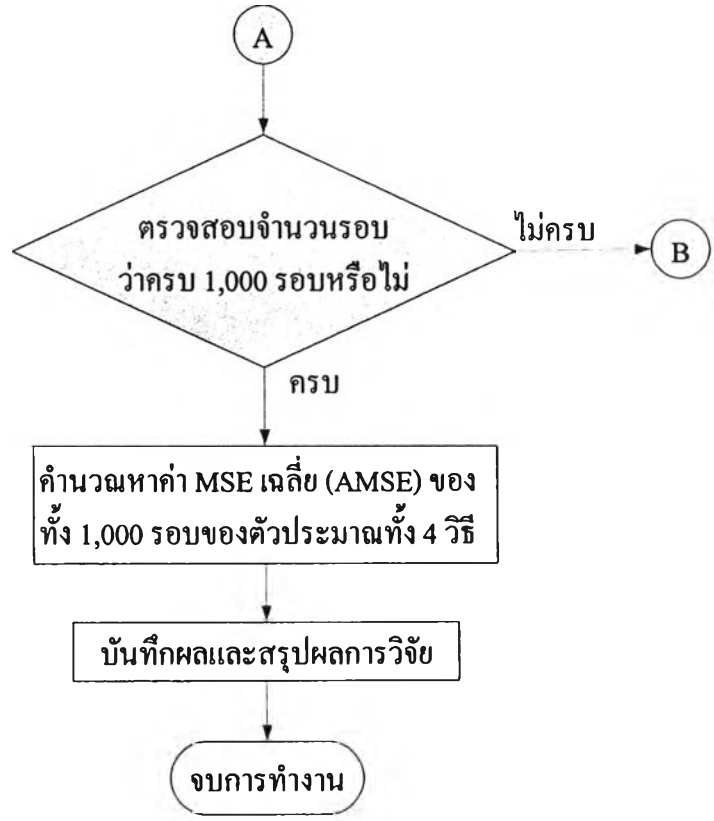
4. ประมาณค่าสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยสุด (Ordinary Least Squares method (OLS)) วิธีบูตสแตรป (Bootstrap method (BS)) และวิธีตัวประมาณ M โดยใช้เกณฑ์ความแกร่งของแรมเซย์ (M-estimator method by Ramsey (M/R)) และตูกี (M-estimator method by Tukey (M/T))
5. กรณีที่ข้อจำกัดไม่เป็นจริง ให้กลับไปทำขั้นตอนที่ 1 และ 2 อีกครั้ง โดยสร้างข้อมูลชุดใหม่ขึ้นมาแทนแทนข้อมูลชุดเดิม จากนั้นทำตามขั้นตอนที่ 3 และ 4 ใหม่ต่อไป
6. หาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error (MSE)) และทำการจำลองซ้ำใหม่จำนวน 1,000 รอบ
7. หาค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (average of mean square error (AMSE)) ค่าส่วนเบี่ยงเบนมาตรฐานของของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (standard deviation of average of mean square error (SD)) และค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (ratio of different average mean square error (RDAMSE)) ของสัมประสิทธิ์การถดถอยเชิงเส้นพหุคูณด้วยวิธีกำลังสองน้อยสุด วิธีบูตสแตรป และวิธีตัวประมาณ M โดยใช้เกณฑ์ความแกร่งของตูกีและแรมเซย์ ในกรณีที่ข้อจำกัดเป็นจริง
8. บันทึกผล เปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองเฉลี่ย และค่าอัตราส่วนผลต่างของค่าเฉลี่ยของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของแต่ละวิธีพร้อมทั้งสรุปผลที่ได้

ขั้นตอนการวิจัยดังกล่าว สามารถแสดงเป็นผังงาน ดังรูปต่อไปนี้

รูปที่ 3.1 ผังงานแสดงขั้นตอนการวิจัยในการเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 4 วิธี



รูปที่ 3.1 ผังงานแสดงขั้นตอนการวิจัยในการเปรียบเทียบค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของการประมาณค่าสัมประสิทธิ์การถดถอยทั้ง 4 วิธี (ต่อ)



3.3 ขั้นตอนในการหาตัวประมาณ

7.3.1 ขั้นตอนในการหาตัวประมาณบูตสแตรป

ในการหาตัวประมาณค่าสัมประสิทธิ์ความถดถอยด้วยวิธีบูตสแตรปมีขั้นตอนดังต่อไปนี้

(1) คำนวณหาค่าตัวประมาณกำลังสองน้อยสุดจากสูตร $\hat{\beta} = (X'X)^{-1}X'y$

(2) หาค่าประมาณของค่าความคลาดเคลื่อนจาก $\hat{\varepsilon} = y - \hat{y}$

(3) สุ่มค่า $\hat{\varepsilon}_i$ แบบคืนที่ขนาด n กล่าวคือเมื่อได้ค่า $\hat{\varepsilon}$ จากขั้นตอนที่ 2 มาแล้วจะทำการสุ่มค่า $\hat{\varepsilon}_i$ จากสมาชิกในเวกเตอร์ $\hat{\varepsilon}$ จะได้ค่า ε_1^* เป็นสมาชิกตัวแรกเก็บไว้ในเวกเตอร์ ε^* จากนั้นทำการสุ่มค่า $\hat{\varepsilon}_i$ ใหม่จากสมาชิกในเวกเตอร์ $\hat{\varepsilon}$ โดยที่เวกเตอร์ยังมีสมาชิกเท่าเดิม (เนื่องจากเป็นการสุ่มแบบคืนที่) จนได้สมาชิกครบในเวกเตอร์ ε^* จะได้ $\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_n^*$ เป็นสมาชิกในเวกเตอร์ ε^*

(4) คำนวณหาค่า $\underset{\sim}{y}^*$ จาก $\underset{\sim}{y}^* = X\underset{\sim}{\hat{\beta}} + \underset{\sim}{\varepsilon}^* = \underset{\sim}{\hat{y}} + \underset{\sim}{\varepsilon}^*$

(5) คำนวณหาค่าตัวประมาณกำลังสองน้อยสุด เพื่อให้ได้ตัวประมาณ $\underset{\sim}{\hat{\beta}}^*$ จากสูตร

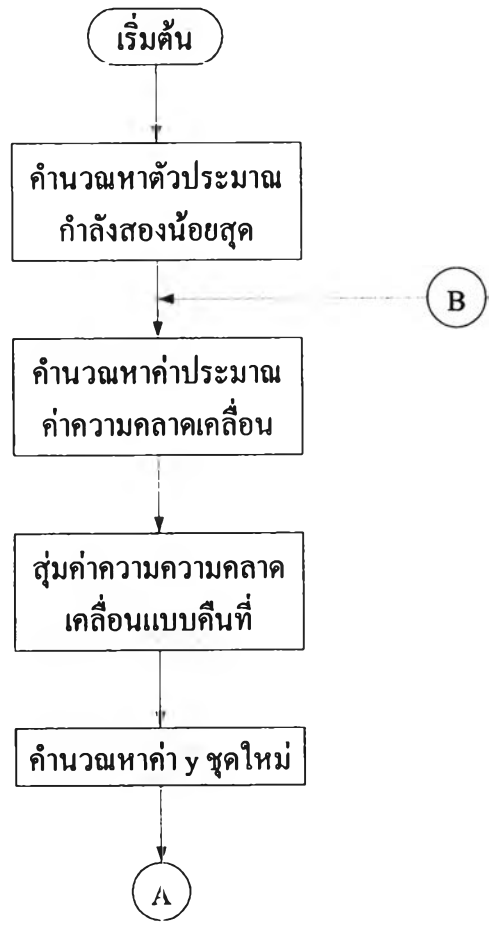
$$\underset{\sim}{\hat{\beta}}^* = (X'X)^{-1}X'\underset{\sim}{y}^*$$

(6) เก็บค่า $\underset{\sim}{\hat{\beta}}^*$ ไว้ใน $\underset{\sim}{\hat{\beta}}^{*i}$ แล้วกลับไปทำขั้นตอนที่ 2 ใหม่ จนครบทั้งหมด 100 รอบ (จำนวนรอบของการประมาณด้วยวิธีบูตสแตรป) จะได้ค่า $\underset{\sim}{\hat{\beta}}^{*i}$ ทั้งหมด 100 ค่า

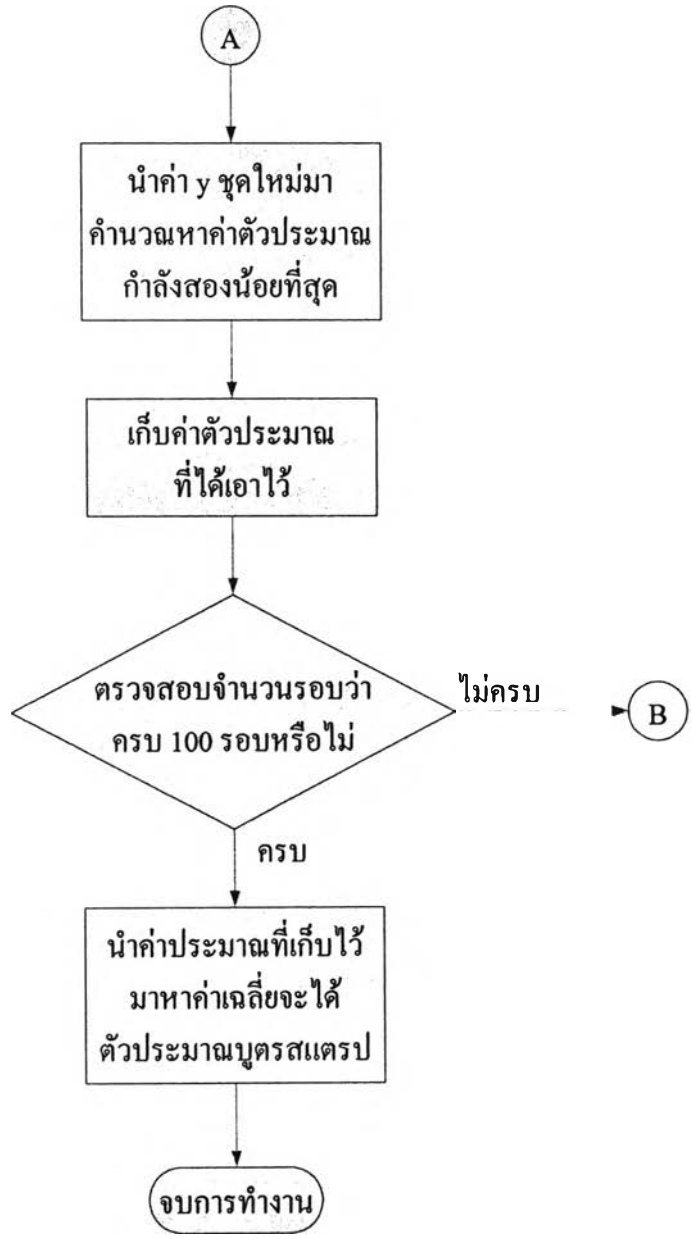
(7) คำนวณหาค่าตัวประมาณบูตสแตรป $\underset{\sim}{\bar{\hat{\beta}}}^*$ จาก $\underset{\sim}{\bar{\hat{\beta}}}^* = \dfrac{\sum_{i=1}^{B}\underset{\sim}{\hat{\beta}}^{*i}}{B}$ เมื่อ B = 100

จากขั้นตอนดังกล่าว สามารถสรุปอยู่ในรูปของแผนงานได้ดังนี้

รูปที่ 3.2 ผังงานแสดงขั้นตอนในการหาตัวประมาณบูตสแตรป



รูปที่ 3.2 ผังงานแสดงขั้นตอนในการหาตัวประมาณบูตสแตรป (ต่อ)



7.3.2 ขั้นตอนในการหาตัวประมาณ M

ในการหาตัวประมาณค่าสัมประสิทธิ์ความถดถอยด้วยวิธีตัวประมาณ M มีขั้นตอนดังต่อไปนี้

(1) คำนวณหาตัวประมาณที่แกร่งของพารามิเตอร์สเกล (s)

(2) สร้างเกณฑ์ความแกร่ง ψ ซึ่งในที่นี้ใช้เกณฑ์ความแกร่งของแรมเซย์และตูกี

(3) คำนวณหาค่า w_{i0}

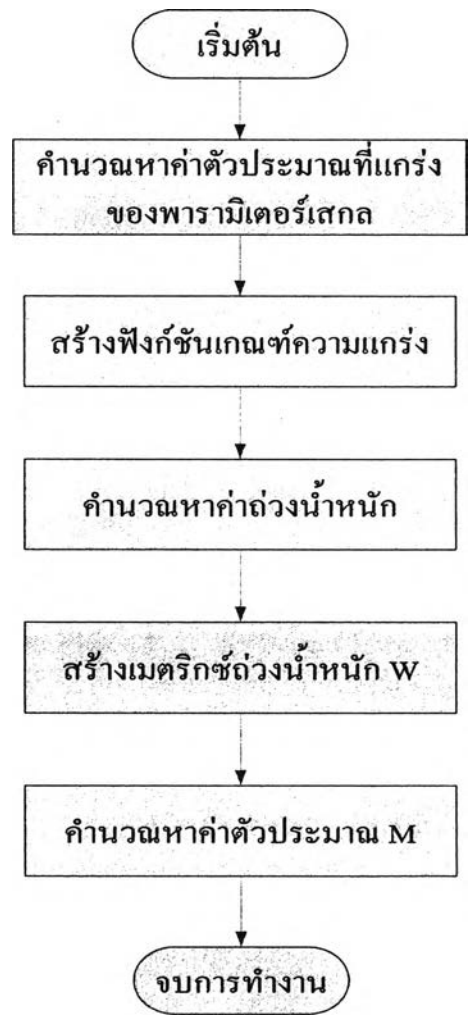
เมื่อ $w_{i0} = \begin{cases} \psi \dfrac{|(y_i - x_i'\beta_0)/s|}{(y_i - x_i'\beta_0)/s} & ; y_i \neq x_i\hat{\beta}_0 \\ 1 & ; y_i = x_i\hat{\beta}_0 \end{cases}$

(4) สร้างเมทริกซ์ถ่วงน้ำหนัก W ซึ่งเป็นเมทริกซ์ทแยงมุมขนาด n × n เมื่อสมาชิกตามเส้นทแยงมุมหลักเป็น $w_{10}, w_{20}, \ldots, w_{n0}$

(5) คำนวณหาค่าตัวประมาณ M $\hat{\beta}$ จากสูตร $\hat{\beta} = (X'W_0X)^{-1}X'W_0\,y$

จากขั้นตอนดังกล่าวข้างต้นสามารถสรุปอยู่ในรูปของแผนงานได้ดังนี้

รูปที่ 3.3 ผังงานแสดงขั้นตอนในการหาตัวประมาณ M



I 22612129

7.4 โปรแกรมที่ใช้ในการวิจัย

ในงานวิจัยครั้งนี้ โปรแกรมที่ใช้ในการวิจัยทั้งหมดเขียนด้วยโปรแกรม S-PLUS2000 ซึ่งในแต่ละสถานการณ์ของการทดลองจะมีลักษณะการทำงานของโปรแกรมเหมือนกันโดยรายละเอียดของโปรแกรม คำอธิบายแต่ละคำสั่งและขั้นตอนการทำงานของโปรแกรมแต่ละวิธีจะแสดงไว้ในภาคผนวก ค