บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ ได้นำวิธีการบูทสแตรปมาประยุกต์ใช้กับตัวสถิติ 3 ชนิด คือ White's test , Breusch-Pagan's Test และ Szroeter's test เพื่อใช้ในการตรวจสอบความไม่คงที่ของความแปรปรวนของความคลาดเคลื่อน โดยจะทำการศึกษาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทั้ง 3 ชนิด ภายใต้สถานการณ์ที่ว่า ความคลาดเคลื่อนมีความแปรปรวนไม่คงที่และมีความแปรปรวน 2 รูปแบบ คือ รูปแบบการคูณและรูปแบบการบวก ณ ความรุนแรงของปัญหาที่ระดับต่างๆ ทั้งนี้เทคนิคที่ใช้ในการหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบ คือ เทคนิคมอนติคาร์โล

<u>เทคนิคมอนติคาร์โล</u>

เทคนิคมอนติคาร์โลเป็นหนึ่งในหลาย ๆ วิธี ที่ใช้ในการแก้ปัญหาในการคำนวณทางคณิตศาสตร์ และเป็นวิธีที่นิยมกันอย่างแพร่หลายในปัจจุบัน Hammerslay และ Handscomb (1964) กล่าวว่า วิธีมอนติคาร์โลเป็นสาขาหนึ่งของคณิตศาสตร์เชิงทดลอง ซึ่งหลักการของวิธีมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่ม (Random Number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา เพราะตัวเลขสุ่มมีประโยชน์หลายประการ คือ

1. ทำให้การเลือกตัวอย่างไม่มีการเอนเอียง ในการสำรวจหรือทดลองในเรื่องนั้นๆ ทั้งนี้เพราะตัวเลขสุ่มมาจากแนวคิดเกี่ยวกับการคำนวณความน่าจะเป็น
2. ตัวเลขสุ่มทำให้ได้มาซึ่งรูปแบบต่างๆ หรือวิธีการที่สลับซับซ้อน โดยการสร้างสถานการณ์จำลอง (Simulation)
3. การใช้ตัวเลขสุ่ม อาจทำเพื่อศึกษาคุณสมบัติทางทฤษฎีของกระบวนการทางสถิติที่มีความสำคัญสำหรับการประมาณค่า ตลอดจนนำไปสู่คำอธิบายเกี่ยวกับอำนาจการทดสอบทางสถิติ (Power of the test)
4. เพื่อหาคำตอบในปัญหาทางคณิตศาสตร์ โดยจะพิจารณาจากการแจกแจงความน่าจะเป็นของปัญหานั้นๆ

ในการวิจัยครั้งนี้ จะใช้วิธีมอนติคาร์โลในการสร้างข้อมูลที่มีสภาพการแจกแจงตามที่ต้องการ ซึ่งขั้นตอนของวิธีมอนติคาร์โลนี้แบ่งออกได้เป็น 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การสร้างตัวเลขสุ่ม การใช้ตัวเลขสุ่มเป็นสิ่งที่สำคัญมากในเทคนิคมอนติคาร์โล ทั้งนี้เพราะว่าหลักการของวิธีมอนติคาร์โลนั้นจะใช้ตัวเลขสุ่มมาช่วยในการหาคำตอบของปัญหา แต่วิธีการสร้างตัวเลขสุ่มนั้นมีหลายวิธี แต่วิธีที่ดีนั้นลักษณะของตัวเลขสุ่มที่เกิดขึ้นจะต้องมีการแจกแจงแบบยูนิฟอร์มในช่วง [0,1] และเป็นอิสระต่อกัน

ขั้นตอนที่ 2 การประยุกต์ปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม ซึ่งขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่อาจจะมีขั้นตอนอื่นหลายๆ ขั้นตอน ซึ่งขั้นตอนเหล่านี้มีบางขั้นตอนที่ต้องใช้ตัวเลขสุ่ม

ขั้นตอนที่ 3 การทดลองกระทำ เมื่อประยุกต์ปัญหาให้ใช้กับตัวเลขสุ่มได้แล้ว ขั้นต่อไปคือการทดลอง โดยใช้กระบวนการของการสุ่ม(Random Process) มากระทำในลักษณะที่ซ้ำๆ กัน (Replication) เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

3.1 การวางแผนการทดลอง

การวิจัยครั้งนี้ เป็นการศึกษาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบ 3 ชนิดข้างต้น ในการตรวจสอบปัญหาความแปรปรวนของความคลาดเคลื่อนไม่คงที่ ณ ความรุนแรงของปัญหาระดับต่างๆ กัน โดยใช้เกณฑ์ของ Prakash (1979) ในการวัดระดับความรุนแรงที่ว่า เมื่อ I.L. ของความแปรปรวนมีค่าน้อยกว่า 0.5 ถือว่า ความคลาดเคลื่อนมีความแปรปรวนไม่คงที่อยู่ในระดับต่ำหรือไม่รุนแรงเท่าใดนัก และถ้าค่า I.L. ของความแปรปรวนมีค่ามากกว่า 0.5 ขึ้นไปถือว่าความรุนแรงของปัญหาความคลาดเคลื่อนมีความแปรปรวนไม่คงที่อยู่ในระดับสูง และจะเพิ่มความรุนแรงขึ้นเรื่อยๆ ตามการเพิ่มขึ้นของค่า I.L. ของความแปรปรวน โดยค่า I.L. ของความแปรปรวนมีสูตรในการคำนวณดังนี้

$$I.L.(\sigma_i^2) = \left\{ \frac{1}{n-1}\left[\sum_{i=1}^{n}\sigma_i^4 - \frac{\left(\sum_{i=1}^{n}\sigma_i^2\right)^2}{n}\right] \Bigg/ \left[\frac{1}{n}\sum_{i=1}^{n}\sigma_i^2\right]^2 \right\}^{\frac{1}{2}}$$

โดยที่ความแปรปรวนมีรูปแบบการคูณและรูปแบบการบวก ซึ่งมีการแปรค่าตามพารามิเตอร์ r และ λ ตามลำดับ รวมทั้งมีการกำหนดสถานการณ์ต่างๆ ดังนี้

3.1.1 กำหนดค่าพารามิเตอร์ $\beta_0 = 1$ และ $\beta_1 = 1$

3.1.2 กำหนดให้ความคลาดเคลื่อนมีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเท่ากับศูนย์และมีรูปแบบความแปรปรวนดังนี้

(1) <u>รูปแบบการคูณ</u> (Multiplicative model)

$$\sigma_i^2 = V(\varepsilon_i) = E(\varepsilon_i^2) = KX_i^r$$

(2) <u>รูปแบบการบวก</u> (Additive model)

$$\sigma_i^2 = V(\varepsilon_i) = E(\varepsilon_i^2) = K^2(1+\lambda X_i)^2$$

เมื่อ K เป็นค่าคงที่ใดๆ ซึ่งกำหนดให้ในที่นี้ $K = 1$

3.1.3 กำหนดค่าพารามิเตอร์ r หรือ λ โดยที่

(1) สำหรับรูปแบบการคูณ

จะทำการพิจารณาเลือกค่า r จำนวน 60 ค่าในช่วง 0 ถึง 20 ซึ่งประกอบด้วย 0.2 , 0.4 , 0.8 , ... , 10 , 11, ... , 20 ทั้งนี้เพราะเป็นช่วงที่ทำให้ความคลาดเคลื่อนมีความแปรปรวนอยู่ในระดับต่ำถึงระดับสูง

(2) สำหรับรูปแบบการบวก

จะทำการพิจารณาค่า λ ทั้งหมด 33 ค่า ซึ่งประกอบด้วยค่าที่อยู่ในช่วง 0 ถึง 1 ซึ่งประกอบด้วย 0.05 , 0.10 , 0.15 , ... , 0.95 เนื่องจากอัตราการเพิ่มขึ้นของอำนาจการทดสอบของสถิติแต่ละชนิดค่อนข้างสูง หลังจากนั้นได้ทำการทดลองเพิ่มค่า λ ในช่วง 1 ถึง 10 ซึ่งได้แก่ 1 , 2 , 3 , ... , 10 และพิจารณาค่า λ เพิ่มอีกจำนวน 6 ค่า ซึ่งได้แก่ λ = 10 , 50 , 100 , 500 , 1,000

3.1.4 การสร้างค่าคงที่ X

ในงานวิจัยครั้งนี้ จะทำการสร้างค่าคงที่ X แต่ละตัวให้มีการแจกแจงแบบปกติ โดยกำหนดให้มีค่าเฉลี่ยเท่ากับ 5 และความแปรปรวนเท่ากับ 1

3.2 ขั้นตอนในการวิจัย

เขียนโปรแกรมคอมพิวเตอร์ด้วยภาษา S-Plus 2000 เพื่อสร้างข้อมูลให้เป็นไปตามแผนการทดลอง แล้วคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทั้ง 3 ชนิด ซึ่งแบ่งขั้นตอนต่างๆ ในการวิจัยเป็นดังนี้

3.2.1 การสร้างข้อมูล (X_i, Y_i) ที่มีความสัมพันธ์กันเชิงเส้นตรง เมื่อจำนวนตัวแปรอิสระเท่ากับ 1 โดยมีขั้นตอนย่อยๆ ดังนี้

3.2.1.1 กำหนดค่าต่างๆ ไม่ว่าจะเป็นขนาดตัวอย่าง (n) ค่าพารามิเตอร์เริ่มต้น (β_0, β_1) รวมทั้งค่าเฉลี่ย ความแปรปรวนและการแจกแจงของความคลาดเคลื่อนดังที่ได้กล่าวมาแล้วข้างต้น

3.2.1.2 สร้างค่าคงที่ X_i

3.2.1.3 สร้างการแจกแจงของความคลาดเคลื่อนตามที่กำหนดไว้ในข้างต้น โดยใช้หลักการเช่นเดียวกับการสร้างค่าคงที่ X_i

3.2.1.4 สร้างค่า Y_i จากความสัมพันธ์ดังนี้

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

3.2.1.5 กระทำตามข้อ 3.2.1.1 – 3.2.1.4 ซ้ำๆ กันตามขนาดตัวอย่างที่กำหนด นั่นคือ ทำซ้ำๆ กัน n ครั้ง ซึ่งทำให้ได้ X_i และ Y_i ที่มีความสัมพันธ์กันในเชิงเส้นตรง

3.2.2 การหาตัวสถิติทดสอบ จะทำการคำนวณค่าของตัวสถิติทดสอบ 3 ชนิด ได้แก่

3.2.1.1 ตัวสถิติทดสอบของไวท์ (White's Statistic) ประกอบด้วยขั้นตอนย่อยๆ ดังต่อไปนี้

1) คำนวณหาค่าสัมประสิทธิ์การถดถอย (β_0, β_1) และความคลาดเคลื่อน $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ โดยวิธี OLS

2) จากสมการ $\hat{\tilde{\varepsilon}}^2 = X_{white}\tilde{\alpha} + \tilde{v}$

หาค่า $\hat{\tilde{\alpha}} = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix} = (X'_{white}X_{white})^{-1}X'_{white}\hat{\tilde{\varepsilon}}^2$

และค่า $R^2 = \dfrac{SSR}{SST}$

เมื่อ $X_{white} = \begin{pmatrix} 1 & X_i & X_i^2 \end{pmatrix}$

$$SSR = \sum_{i=1}^{n}(\hat{\alpha}_0 + \hat{\alpha}_1 X_{1i} + \hat{\alpha}_2 X_{1i}^2)\hat{\varepsilon}_i^2 - \frac{\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2\right)^2}{n}$$

และ $SST = \sum_{i=1}^{n}(\hat{\varepsilon}_i^2)^2 - \dfrac{\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2\right)^2}{n}$

3) คำนวณหาสถิติทดสอบจาก $W = nR^2$

3.2.1.2 ตัวสถิติทดสอบของบรูชและพาแกน (Breusch-Pagan's statistic) ซึ่งประกอบด้วยขั้นตอนย่อยๆ ดังต่อไปนี้

1) คำนวณหาค่าประมาณความคลาดเคลื่อนโดยวิธี OLS $(\hat{\varepsilon}_i)$ เช่นเดียวกับการหาสถิติทดสอบของไวท์

2) คำนวณหาค่าประมาณความแปรปรวนของความคลาดเคลื่อนจาก

สูตร $\hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$

3) คำนวณค่าเวกเตอร์ $\tilde{g}$ โดยสมาชิกในเวกเตอร์ คำนวณได้จาก

$$g_i = \frac{\hat{\varepsilon}_i^2}{\hat{\sigma}^2} - 1 \quad , \quad i = 1,2,\ldots,n$$

4) สถิติทดสอบของบรูชและพาแกน คือ

$$BP = \frac{1}{2}\tilde{g}'z(z'z)^{-1}z'\tilde{g}$$

เมื่อ z เป็นเมตริกซ์ขนาด $n \times 2$ ซึ่งมีสมาชิกในแถวที่ i คือ

$\tilde{z}_i = (1, \ln X_i)$ สำหรับความแปรปรวนที่มีรูปแบบการคูณ

และ $\tilde{z}_i = (1, X_i)$ สำหรับความแปรปรวนที่มีรูปแบบการบวก

3.2.1.3 ตัวสถิติทดสอบของสโรเตอร์ (Szroeter's statistic) ซึ่งประกอบด้วยขั้นตอนย่อยๆ ดังต่อไปนี้

1) คำนวณหาค่าประมาณความคลาดเคลื่อนโดยวิธี OLS $(\hat{\varepsilon}_i)$ เช่นเดียวกับทั้ง 2 วิธีข้างต้น

2) สถิติทดสอบของสโรเตอร์ คือ

$$SZ = n(\tilde{h} - \bar{h}) \Big/ \left[2\sum_{i=1}^{n} (h_i - \bar{h})^2 \right]^{\frac{1}{2}}$$

โดยที่ $h_i = i$, $\tilde{h} = \sum_{i=1}^{n} w_i h_i$, $w_i = \hat{\varepsilon}_i^2 \Big/ \sum_{i=1}^{n} \hat{\varepsilon}_i^2$,

$$\bar{h} = \frac{1}{n}\sum_{i=1}^{n} h_i$$

3.2.3 การหาการแจกแจงเชิงประจักษ์ (Empirical Distribution) และค่าวิกฤตของสถิติทดสอบ ซึ่งประกอบด้วยขั้นตอนย่อยๆ ดังนี้

3.2.3.1 หาค่าความคลาดเคลื่อนจากวิธีบูทสแตรป (bootstrapped error : ε_i^*) โดยการสุ่มค่าประมาณความคลาดเคลื่อน $(\hat{\varepsilon}_i)$ ที่คำนวณได้ในข้อ 3.2.2 ทั้ง n ค่าแบบใส่คืน

3.2.3.2 หาค่า bootstrapped dependent variable $\left(y_i^*\right)$ จากสมการดังนี้

$$Y_i^* = \hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon_i^*$$

3.2.3.3 หาค่าประมาณสัมประสิทธิ์การถดถอย $\left(\hat{\beta}_0^*, \hat{\beta}_1^*, \hat{\beta}_2^*\right)$ และค่าประมาณความคลาดเคลื่อน $\left(\hat{\varepsilon}_i^*\right)$ ด้วยวิธีบูทสแตรป จาก OLS

3.2.3.4 หาค่าบูทสแตรปของไวท์ (W^*) ค่าบูทสแตรปของบรูชและพาแกน (BP^*) และค่าบูทสแตรปของสโรเตอร์ (SZ^*) โดยอาศัยหลักการเช่นเดียวกับการหาค่าสถิติทดสอบของไวท์สถิติทดสอบของบรูชและพาแกน และสถิติทดสอบของสโรเตอร์ แต่เปลี่ยนค่าประมาณความคลาดเคลื่อนเดิมเป็นค่าประมาณความคลาดเคลื่อนจากวิธีบูทสแตรป นั่นคือ จาก $\hat{\varepsilon}_i$ เป็น $\hat{\varepsilon}_i^*$

3.2.3.5 กระทำตามข้อ 3.2.3.1 – 3.2.3.4 ซ้ำๆ กันตามจำนวนรอบของการทำบูทสแตรปที่กำหนด นั่นคือ 200 รอบ

3.2.3.6 นำค่าบูทสแตรปไวท์ทั้ง 200 ค่ามาสร้างการแจกแจงเชิงประจักษ์ และหาค่าวิกฤตของการทดสอบที่ระดับนัยสำคัญที่กำหนด ส่วนในกรณีค่าบูทสแตรปของบรูชและพาแกนและบูทสแตรปของสโรเตอร์ก็กระทำในลักษณะเดียวกัน

3.2.4 การตัดสินใจ

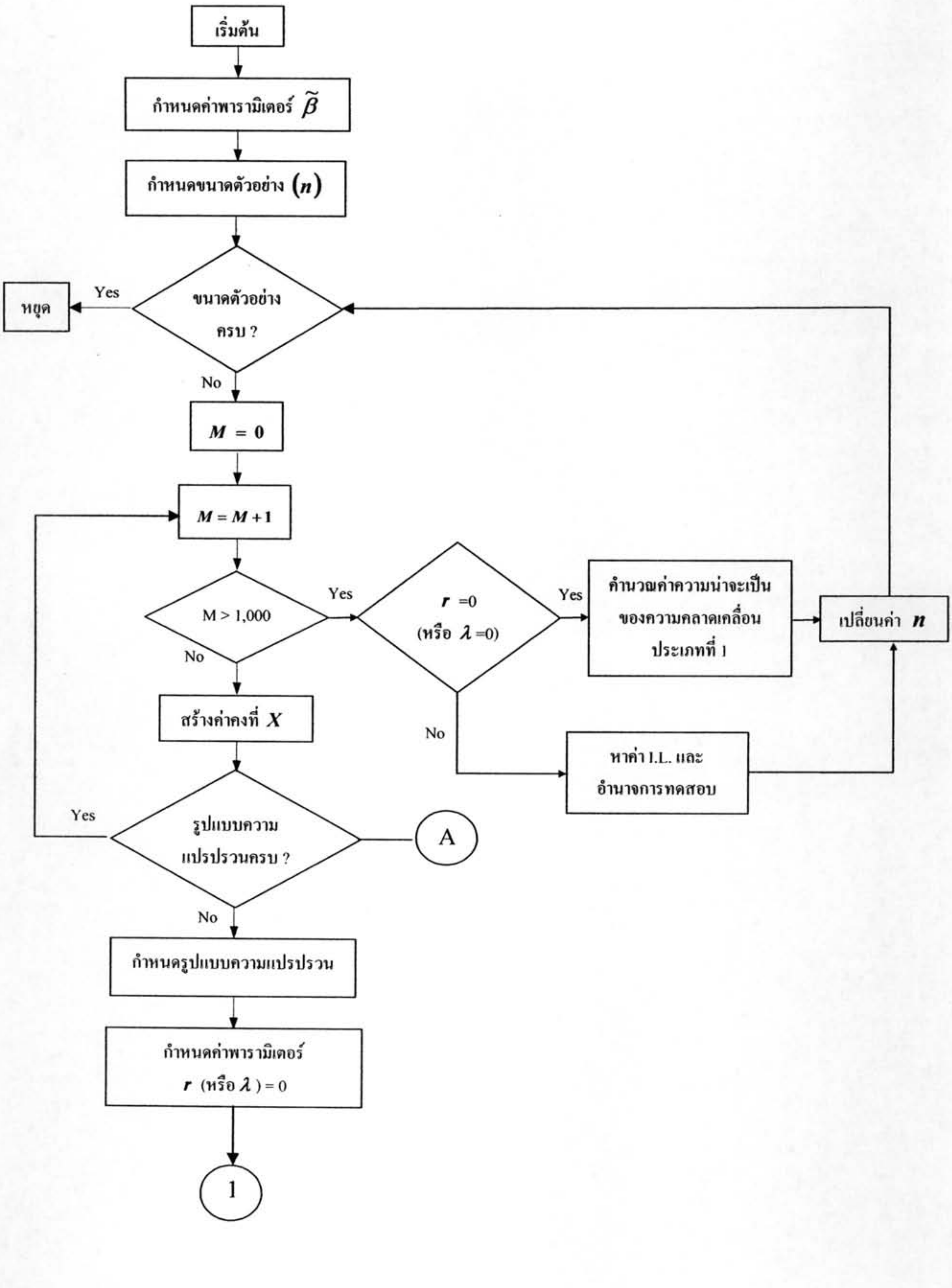
เปรียบเทียบค่าสถิติทดสอบที่คำนวณได้ในข้อ 3.2.2 กับค่าวิกฤตที่คำนวณได้ในข้อ 3.2.3 เพื่อตัดสินใจว่าจะปฏิเสธหรือยอมรับสมมติฐานว่าง ในกรณีที่ปฏิเสธสมมติฐานว่างให้นับจำนวนครั้งที่ปฏิเสธด้วย

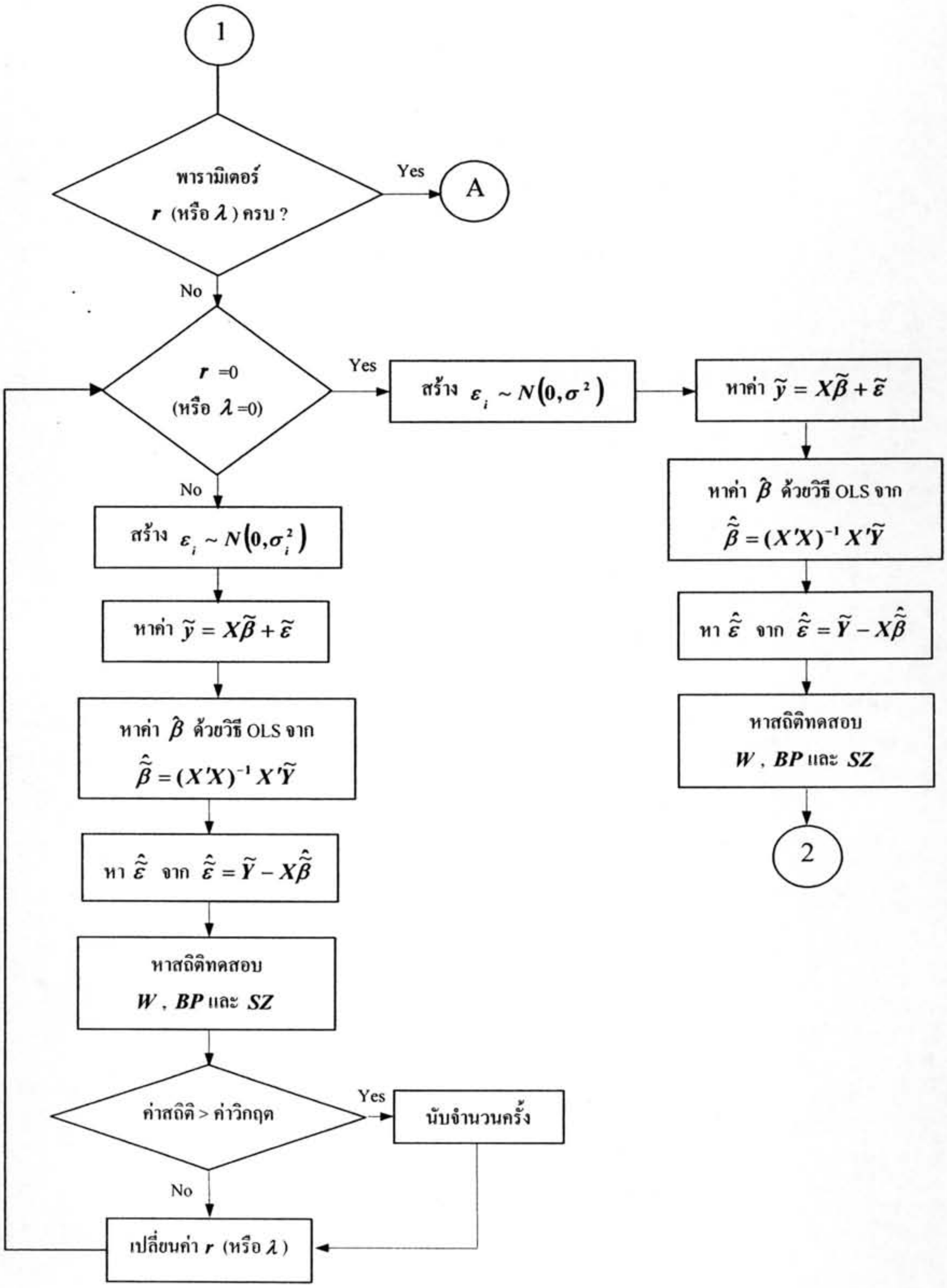
3.2.5 การหาค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 และอำนาจในการทดสอบ

3.2.5.1 กรณีที่ r หรือ λ มีค่าเท่ากับศูนย์ จะทำการคำนวณหาค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1

3.2.5.2 กรณีที่ r หรือ λ มีค่าไม่เท่ากับศูนย์ จะทำการคำนวณหาอำนาจในการทดสอบ สำหรับพารามิเตอร์ r หรือ λ ในแต่ละค่า จะทำการหาค่าระดับความไม่คงที่ (I.L.) ของความแปรปรวน หลังจากนั้นจะทำการเปลี่ยนค่าพารามิเตอร์ r หรือ λ จนกระทั่งครบทุกค่าตามที่ต้องการ ในทั้งสองรูปแบบความแปรปรวน โดยในแต่ละค่าของพารามิเตอร์จะสุ่มตัวอย่างซ้ำๆ กัน 1000 ครั้งตามกระบวนการของเทคนิคมอนติคาร์โล ขั้นตอนต่อไปจึงเปลี่ยนขนาดตัวอย่าง และทำการศึกษาในลักษณะเช่นเดิมในทุกขนาดตัวอย่าง แล้วจึงเปลี่ยนจำนวนตัวแปรอิสระเป็น 2

รูปที่ 3.1 แสดงขั้นตอนในการวิจัยเพื่อหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจในการทดสอบ




1
พารามิเตอร์
r (หรือ λ) ครบ ?
Yes
A
No
r =0
(หรือ λ =0)
Yes
สร้าง ε_i ~ N(0,σ²)
หาค่า ỹ = Xβ̃ + ε̃
หาค่า β̂ ด้วยวิธี OLS จาก
β̃̂ = (X'X)⁻¹ X'Ỹ
หา ε̃̂ จาก ε̃̂ = Ỹ − Xβ̃̂
หาสถิติทดสอบ
W , BP และ SZ
2
No
สร้าง ε_i ~ N(0,σ_i²)
หาค่า ỹ = Xβ̃ + ε̃
หาค่า β̂ ด้วยวิธี OLS จาก
β̃̂ = (X'X)⁻¹ X'Ỹ
หา ε̃̂ จาก ε̃̂ = Ỹ − Xβ̃̂
หาสถิติทดสอบ
W , BP และ SZ
ค่าสถิติ > ค่าวิกฤต
Yes
นับจำนวนครั้ง
No
เปลี่ยนค่า r (หรือ λ)

รูปที่ 3.2 แสดงขั้นตอนที่ใช้ในการหาค่าวิกฤตด้วยวิธีการบูทสแตรป