บทที่ 3

วิธีดำเนินการวิจัย

สิ่งที่ต้องการศึกษาสำหรับการวิจัยครั้งนี้ คือการทดสอบความเท่ากันของสัมประสิทธิ์ความถดถอยเชิงเส้นเมื่อความแปรปรวนของความคลาดเคลื่อนสุ่มไม่เท่ากัน ซึ่งประกอบด้วยสมการถดถอยเชิงเส้น 2 สมการ การประมาณค่าสัมประสิทธิ์ความถดถอยเชิงเส้นด้วยวิธีกำลังสองน้อยที่สุด วิธีทดสอบเข้า วิธีทดสอบโทโยดา วิธีทดสอบเซลเนอร์-ธีล-กุปตา ศึกษาความแกร่งของวิธีทดสอบและความไวของวิธีทดสอบทั้ง 3 วิธี เพื่อเป็นเกณฑ์พิจารณาความเหมาะสมทั้ง 3 วิธี การศึกษาความแกร่งของวิธีทดสอบจะพิจารณาจากค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และความไวของวิธีทดสอบพิจารณาจากอำนาจการทดสอบ ทั้งความแกร่งและอำนาจการทดสอบ ทำการศึกษาเมื่อความคลาดเคลื่อนมีการแจกแจงแบบปกติ เนื่องจากการแจกแจงแบบปกติ เป็นการแจกแจงตามข้อสมมติเบื้องต้นของแบบจำลองความถดถอย ในกรณีที่ขนาดตัวอย่างทั้งสองชุดมีขนาดเท่ากัน ขนาดตัวอย่างที่สนใจศึกษาของแต่ละชุด คือ (20 20) (30 30) (50 50) ในกรณีที่ขนาดตัวอย่างไม่เท่ากัน ขนาดตัวอย่างในแต่ละชุดจะเป็น (20 30) (20 50) และ (30 50) จำนวนตัวแปรอิสระที่ศึกษาคือ 2 3 4 และ 5 อัตราส่วนความแปรปรวนของความคลาดเคลื่อนประชากรที่ 1 ต่อประชากรที่ 2 ($\sigma_A^2 : \sigma_B^2$) เป็น 1:1 1:2 1:3 1:5 1:10 1:20 2:1 3:1 5:1 10:1 และ 20:1

ในการวิจัยครั้งนี้จะใช้วิธีมอนติคาร์โลในการสร้างข้อมูลให้มีลักษณะตามที่ต้องการศึกษา ขั้นตอนต่าง ๆ ของการวิจัย ตลอดจนโปรแกรมที่ใช้ในการวิจัยมีรายละเอียดดังต่อไปนี้

3.1 ใช้วิธีมอนติคาร์โลเพื่อสร้างข้อมูลที่จะนำมาใช้ในการวิจัย

วิธีมอนติคาร์โล เป็นวิธีที่ใช้ในการจำลองตัวแบบทางคณิตศาสตร์ที่นิยมกันมากในปัจจุบัน วิธีมอนติคาร์โลใช้หลักการจำลองตัวเลขสุ่ม มาช่วยในการหาคำตอบของปัญหาต่าง ๆ ที่ทำการศึกษา

ในการวิจัยครั้งนี้ได้ใช้วิธีมอนติคาร์โลเพื่อสร้างข้อมูลให้มีสภาพการณ์ตามที่ต้องการศึกษา โดยแบ่งขั้นตอนต่าง ๆ ดังนี้

3.1.1 สร้างตัวเลขสุ่ม ตัวเลขสุ่ม (random number) เป็นหลักสำคัญของวิธีมอนติคาร์โล เนื่องจากวิธีมอนติคาร์โลใช้ตัวเลขสุ่มช่วยในการหาคำตอบของปัญหา ซึ่งลักษณะของตัวเลขสุ่มจะมีการแจกแจงแบบยูนิฟอร์ม (Uniform Distribution) ในช่วง (0 , 1) การวิจัยครั้งนี้ได้ใช้วิธีของไวท์และสมิดท์ (White & Schmidt 1975 , 421) ซึ่งเสนอไว้ว่าลักษณะของตัวเลขสุ่มที่เกิดขึ้นมีการแจกแจงแบบยูนิฟอร์มในช่วง (0 , 1) และเป็นอิสระกัน รายละเอียดแสดงไว้ในภาคผนวก ก.

3.1.2 นำปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่มที่ได้ จากนั้นก็ทำการทดลองโดยใช้ขั้นตอนของการสุ่ม โดยกระทำในลักษณะซ้ำ ๆ กันเพื่อหาคำตอบของปัญหา

3.2 แผนการทดลอง

กำหนดสถานการณ์ต่าง ๆ ที่จะศึกษาโดยสุ่มตัวอย่างจากประชากรแบบเดียวกัน ตามลักษณะที่สนใจศึกษา คือ ประชากรที่มีการแจกแจงแบบปกติ อัตราส่วนของความคลาดเคลื่อนประชากรที่หนึ่งต่อประชากรที่สอง เป็น 1:1 1:2 1:3 1:5 1:10 1:20 2:1 3:1 5:1 10:1 และ 20:1 จำนวนตัวแปรอิสระเป็น 2 3 4 และ 5 ขนาดตัวอย่างแต่ละกลุ่มเป็น 20 30 และ 50 โดยแบ่งกรณีกลุ่มตัวอย่างเท่ากันเป็น (20 20) (30 30) (50 50) และกรณีตัวอย่างไม่เท่ากันเป็น (20 30) (20 50) (30 50) ซึ่งมีรายละเอียดดังแสดงไว้ในตารางที่ 3.2.1

ตารางที่ 3.2.1 แสดงขนาดตัวอย่างของแต่ละกลุ่มและจำนวนตัวแปรอิสระ

จำนวนตัวแปรอิสระ	ขนาดตัวอย่างแต่ละกลุ่ม (n_1 , n_2)
2	(20 , 20) (30 , 30) (50 , 50) (20 , 30) (20 , 50) (30 , 50)
3	(20 , 20) (30 , 30) (50 , 50) (20 , 30) (20 , 50) (30 , 50)
4	(20 , 20) (30 , 30) (50 , 50) (20 , 30) (20 , 50) (30 , 50)
5	(20 , 20) (30 , 30) (50 , 50) (20 , 30) (20 , 50) (30 , 50)

3.3 ขั้นตอนการวิจัย

ขั้นตอนต่าง ๆ ของการวิจัยมีดังนี้

1. สร้างโปรแกรมย่อย สำหรับสร้างการแจกแจงของความคลาดเคลื่อน (ε) ซึ่งมีการแจกแจงแบบปกติ

2. การสร้างข้อมูล (X , Y) ที่มีความสัมพันธ์เชิงเส้นตรง

3. การแปลงตัวแปรอิสระให้เป็นมาตรฐานเดียวกัน

4. การหาค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้นด้วยวิธีกำลังสองน้อยที่สุด

5. การหาค่าความน่าจะเป็นของการเกิดความคลาดเคลื่อนประเภทที่ 1

6. การหาค่าอำนาจการทดสอบ

ซึ่งมีรายละเอียดของแต่ละขั้นตอนดังนี้

3.3.1 การสร้างโปรแกรมย่อยสำหรับการแจกแจงแบบปกติของความคลาดเคลื่อน (ε)

โปรแกรมที่ใช้ในการวิจัย เขียนด้วยภาษาฟอร์แทรนโฟร์ (Fortran IV) โดยใช้เครื่องคอมพิวเตอร์ IBM 370/3031 การสร้างความคลาดเคลื่อนที่มีลักษณะการแจกแจงแบบปกติจะใช้ตัวเลขสุ่ม ซึ่งมีการแจกแจงแบบยูนิฟอร์มในช่วง (0 , 1) เป็นพื้นฐานในการสร้าง รายละเอียดการสร้างลักษณะการแจกแจงของความคลาดเคลื่อนเป็นดังนี้

เนื่องจากความคลาดเคลื่อนมีการแจกแจงแบบปกติ จึงใช้โปรแกรมย่อยสำหรับสร้างการแจกแจงแบบปกติโดยวิธีของเกาส์ (Gauss) ซึ่งเป็นวิธีสร้างการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น μ และมีความแปรปรวนเป็น σ^2 โดยใช้โปรแกรมย่อยนอร์ม (Subroutine Norm) ด้วยคำสั่ง CALL NORM (AMEAN, SIGMA, A) โดย AMEAN และ $(SIGMA)^2$ เป็นค่าเฉลี่ยและความแปรปรวนที่ถูกส่งมาจากโปรแกรมหลัก (Main Program) A เป็นค่าผลลัพธ์ ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น AMEAN และความแปรปรวนเป็น $(SIGMA)^2$ โดยศึกษาเมื่อความคลาดเคลื่อนมีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนเป็น 10 20 30 50 100 และ 200 ตามลำดับ

3.3.2 การสร้างข้อมูล (X , Y) ที่มีความสัมพันธ์เชิงเส้นตรง

การสร้างข้อมูล (X , Y) ที่มีความสัมพันธ์เชิงเส้นตรงนั้น จะทำการสร้างเมตริกซ์ X ซึ่งเป็นค่าคงที่ก่อน แล้วจึงสร้างค่า $\underset{\sim}{Y}$ ที่มีความสัมพันธ์เชิงเส้นกับเมตริกซ์ X ตามรูปแบบดังนี้ คือ $\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ เมื่อ $\underset{\sim}{\beta}$ เป็นค่าพารามิเตอร์ที่ถูกกำหนด และ $\underset{\sim}{\varepsilon}$ เป็นความคลาดเคลื่อนที่มีการแจกแจงแบบปกติดังที่กล่าวมาแล้วใน 3.3.1 สำหรับค่าเมตริกซ์ X นั้นสร้างจากการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น 10 และความแปรปรวนเป็น 4 ค่าคงที่ที่ได้จะมีสัมประสิทธิ์ความแปรผัน 20% ในการสร้างข้อมูลนั้นจะเริ่มจากการกำหนดขนาดตัวอย่างที่ต้องการศึกษา จำนวนตัวแปรอิสระค่าพารามิเตอร์ $\underset{\sim}{\beta}$ ค่าเฉลี่ยและความแปรปรวน แล้วสร้างค่าเมตริกซ์ X จากนั้นจะใช้คำสั่งสุ่มตัวอย่างเพื่อสร้าง $\underset{\sim}{\varepsilon}$ ที่มีลักษณะการแจกแจงแบบปกติ แล้วจึงสร้างค่า $\underset{\sim}{Y}$ ตามรูปแบบความสัมพันธ์ดังกล่าว

หมายเหตุ ในการสร้างค่าเมตริกซ์ X และจากการทดลองกระทำที่ขนาดตัวอย่าง (20 20) (20 30) (20 50) จำนวนตัวแปรอิสระเป็น 2 แล้วพบว่า ค่าเฉลี่ยและความแปรปรวนไม่มีผลต่อค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ

3.3.3 การแปลงค่าตัวแปรอิสระให้เป็นค่ามาตรฐานเดียวกัน

เนื่องจากตัวแปรอิสระมีมากกว่า 1 ตัว ค่าของตัวแปรอิสระแต่ละตัวอาจจะมีหน่วยที่ต่างกัน เพื่อทำให้ตัวแปรอิสระมีหน่วยเดียวกัน จึงต้องแปลงค่าตัวแปรอิสระใช้เป็นค่ามาตรฐาน โดยใช้วิธีการแปลงค่าตัวแปรอิสระแบบ Unit Length Scaling

3.3.4 การหาค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้นโดยใช้วิธีกำลังสองน้อยที่สุด

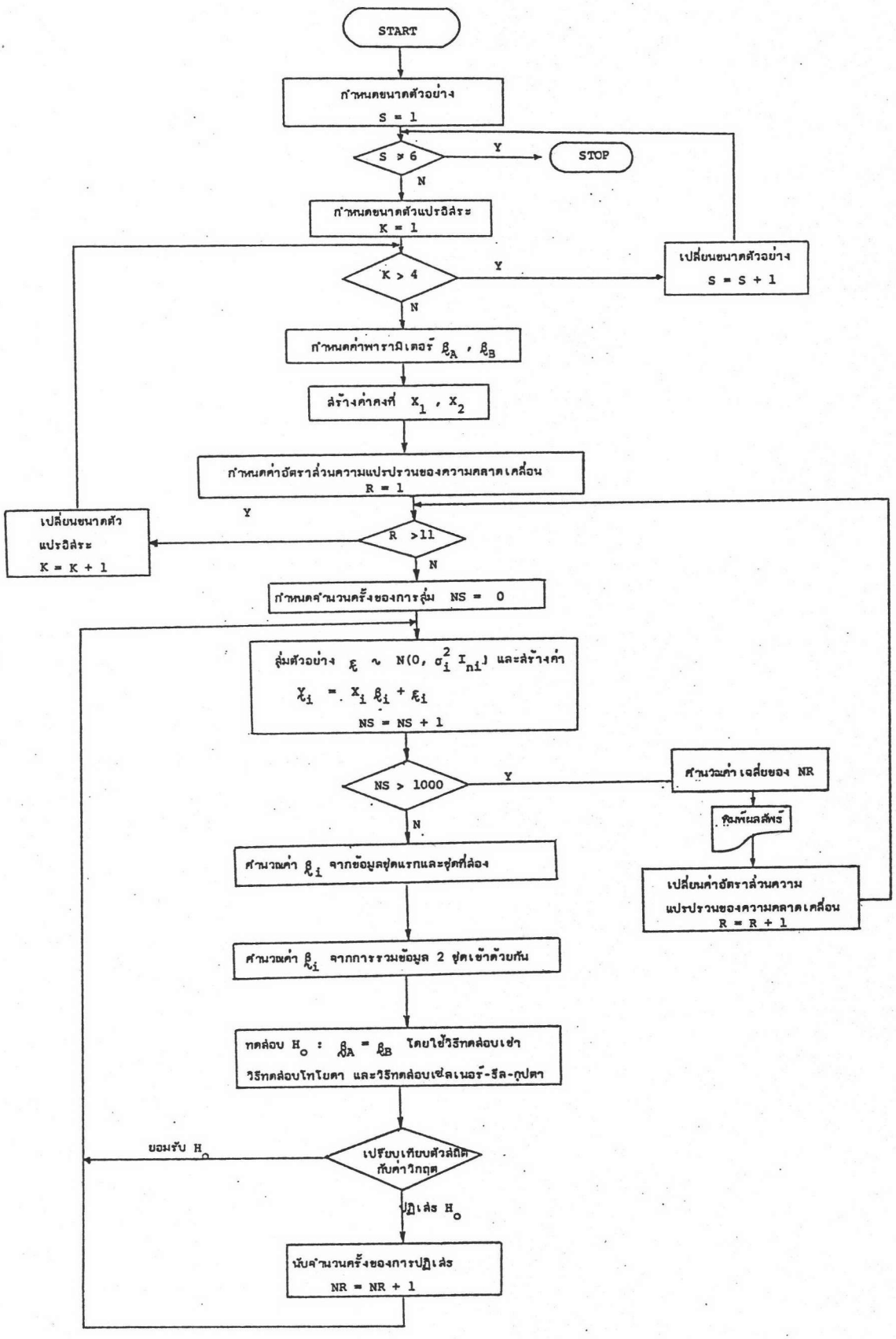
ในการหาค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้นโดยใช้วิธีกำลังสองน้อยที่สุดนั้นได้ทำการประมาณค่าสัมประสิทธิ์ความถดถอยเชิงเส้นของข้อมูลทั้ง 3 ชุด คือ ข้อมูลชุดที่หนึ่ง ข้อมูลชุดที่สอง และข้อมูลซึ่งได้จากการรวมข้อมูลชุดที่หนึ่งและข้อมูลชุดที่สองเข้าด้วยกัน โดยใช้คำสั่ง CALL BETA (N , K, X, Y, B, Al) เมื่อ N เป็น จำนวนตัวอย่าง K เป็นจำนวนสัมประสิทธิ์ความถดถอยเชิงเส้น X เป็นเมตริกซ์ค่าคงที่ Y เป็นเมตริกซ์ของตัวแปรตามซึ่ง N K X Y ถูกส่งค่ามาจากโปรแกรมหลัก B เป็นค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้น Al เป็นค่าส่วนกลับของผลคูณเมตริกซ์ $X'X$ นั่นคือ $Al = (X'X)^{-1}$ โดยที่ B และ

A1 เป็นค่าผลลัพธ์ที่ได้จากโปรแกรมย่อยนี้

3.3.5 การหาค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 สำหรับการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ β

เมื่อสร้างข้อมูล (X , Y) ที่มีความสัมพันธ์เชิงเส้นตามรูปแบบที่ต้องการได้แล้ว จากนั้นหาค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้นโดยวิธีกำลังสองน้อยที่สุด ขั้นต่อไปคือการทดลองเพื่อหาค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 โดยจะต้องคำนวณค่าสถิติของสถิติทดสอบแต่ละวิธีก่อน คือ วิธีทดสอบเซา วิธีทดสอบโทโยดา วิธีทดสอบเซลเนอร์-ธีล-กุปตา เมื่อคำนวณค่าสถิติครบทุกวิธีแล้ว ก็จะนำค่าสถิติที่ได้มาเปรียบเทียบกับค่าวิกฤตเพื่อที่จะตัดสินใจว่าจะปฏิเสธ หรือยอมรับสมมติฐาน H_o ในกรณีที่ปฏิเสธ H_o ให้นับจำนวนครั้งที่ปฏิเสธ จากนั้นก็ย้อนกลับไปสุ่มตัวอย่างชุดใหม่จนกระทั่งครบ 1,000 ครั้ง แล้วคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เมื่อพารามิเตอร์ $\beta_A = \beta_B$ ขั้นต่อไปจะเปลี่ยนความแปรปรวนของความคลาดเคลื่อนของประชากร เพื่อให้อัตราส่วนความแปรปรวนของความคลาดเคลื่อนเป็นไปตามที่ต้องการศึกษา โดยแต่ละค่าอัตราส่วนความแปรปรวนของความคลาดเคลื่อนจะสุ่มตัวอย่างซ้ำ ๆ กัน จนครบ 1,000 ครั้ง และคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 เมื่อครบทุกค่าอัตราส่วนของความแปรปรวนของความคลาดเคลื่อนแล้ว ขั้นต่อไปก็จะเปลี่ยนจำนวนตัวแปรอิสระจนครบตามที่ศึกษาในแต่ละค่าจำนวนตัวแปรอิสระก็จะคำนวณค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 เมื่อครบทุกค่าจำนวนตัวแปรอิสระแล้วก็เปลี่ยนขนาดตัวอย่าง แต่ละขนาดตัวอย่างก็จะคำนวณค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ทำเช่นนี้จนครบขนาดตัวอย่างที่ศึกษา ซึ่งสรุปเป็นผังงานได้ดังแสดงไว้ในรูปที่ 3.1

รูปที่ 3.1 แสดงผังงานของการหาค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 สำหรับการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ β



3.3.6 การหาค่าอำนาจการทดสอบสำหรับการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ β

การหาค่าอำนาจการทดสอบ สำหรับการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ β กระทำเช่นเดียวกับการหาค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 เพียงแต่กำหนดค่าพารามิเตอร์ให้ผิดไปจากสมมติฐานที่ต้องการทดสอบ นั่นคือกำหนดค่าพารามิเตอร์ $\beta_A \neq \beta_B$ ในที่นี้จะพิจารณาค่าพารามิเตอร์ β_B ที่แตกต่างจากค่าพารามิเตอร์ β_A 10 ค่า กล่าวคือ ค่าพารามิเตอร์ β มีค่าเป็น .6 .7 .8 .85 .9 1.1 1.2 1.25 1.3 และ 1.4 และค่าอัตราส่วนความแปรปรวนของความคลาดเคลื่อนเป็น 1:1 1:2 1:3 1:5 2:1 3:1 5:1 กรณีที่ค่าอัตราส่วนความแปรปรวนของความคลาดเคลื่อนเป็น 1:10 1:20 10:1 20:1 จะพิจารณาค่าพารามิเตอร์ β_B 10 ค่าเช่นกันคือ .5 .6 .7 .8 .9 1.1 1.2 1.3 1.4 และ 1.5

ขั้นตอนเริ่มจากการสร้างข้อมูล (X , Y) ที่มีความสัมพันธ์เชิงเส้นตรงตามแบบจำลองที่ต้องการ จากนั้นหาค่าประมาณสัมประสิทธิ์ความถดถอยเชิงเส้นโดยใช้วิธีกำลังสองน้อยที่สุด การทดลองเพื่อหาค่าอำนาจของการทดสอบ โดยการหาค่าสถิติของสถิติทดสอบแต่ละวิธีคือ วิธีการทดสอบเซ่า วิธีการทดสอบโทโยดา และวิธีการทดสอบเซลเนอร์-ธีล-ทุปตา เพื่อที่จะนำค่าสถิติที่ได้มาเปรียบเทียบกับค่าวิกฤตว่าจะปฏิเสธหรือยอมรับสมมติฐาน H_o เมื่อค่าพารามิเตอร์มีค่าต่าง ๆ ดังกล่าวข้างต้น ในกรณีที่ปฏิเสธสมมติฐาน H_o ให้นับจำนวนครั้งของการปฏิเสธแล้วย้อนกลับไปสุ่มตัวอย่างใหม่จนครบ 500 ครั้ง คำนวณค่าอำนาจการทดสอบทำเช่นนี้จนครบทุกค่าของพารามิเตอร์ ขั้นต่อไปเปลี่ยนอัตราส่วนความแปรปรวนของความคลาดเคลื่อนของประชากรตามที่ต้องการศึกษา แต่ละค่าอัตราส่วนความแปรปรวนของความคลาดเคลื่อนจะสุ่มตัวอย่างซ้ำ ๆ กัน 500 ครั้ง แล้วคำนวณค่าอำนาจการทดสอบ จนครบทุกค่าอัตราส่วนความแปรปรวนของความคลาดเคลื่อน จากนั้นก็จะเปลี่ยนค่าจำนวนตัวแปรอิสระจนครบตามที่ต้องการ โดยแต่ละค่าของจำนวนตัวแปรอิสระก็จะคำนวณค่าอำนาจการทดสอบ เมื่อครบค่าของจำนวนตัวแปรอิสระ ก็จะเปลี่ยนขนาดของตัวอย่าง ซึ่งแต่ละขนาดตัวอย่างก็จะคำนวณค่าอำนาจการทดสอบ ทำเช่นนี้จนครบขนาดตัวอย่างที่ศึกษา ซึ่งสรุปเป็นผังงานได้ดังแสดงไว้ในรูป 3.2

รูปที่ 3.2 แสดงผังงานการหาค่าอำนาจการทดสอบสำหรับการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ β