บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ใช้การจำลองข้อมูลด้วยวิธีมอนติคาร์โล (Monte Carlo simulation technique) เพื่อศึกษาเปรียบเทียบประสิทธิภาพของวิธีการทดสอบความเหมาะสมของตัวแบบกรณีข้อมูลไม่มีการทำซ้ำ 3 วิธี โดยพิจารณาจากค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่ากำลังการทดสอบ มีรายละเอียดของการดำเนินการวิจัยทั้งหมด 5 ขั้นตอนดังนี้

1. การจำลองข้อมูลให้มีการแจกแจงและความแปรปรวนตามที่ต้องการศึกษา
2. แบ่งกลุ่มของข้อมูลให้มีจำนวนกลุ่มตามต้องการ
3. คำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่ากำลังการทดสอบของสถิติทดสอบทั้ง 3 วิธี คือ วิธีของเดรปเปอร์และสมิทธิ์ วิธีของซูและยาง วิธีของมิลเลอร์และนีลล์
4. วิเคราะห์เปรียบเทียบประสิทธิภาพของตัวสถิติทั้ง 3 ชนิด

การจำลองข้อมูลให้มีการแจกแจงและความแปรปรวนตามที่ต้องการศึกษา

ในการวิจัยนี้เป็นการศึกษาโดยอาศัยวิธีการจำลองค่า (Simulation study) ด้วยโปรแกรม R เพื่อนำมาสร้างเป็นข้อมูล 2000 ชุด ซึ่งเป็นตัวแบบที่ประกอบไปด้วยตัวแปรอิสระ X ตัวแปรตาม Y และความคลาดเคลื่อนสุ่ม ตัวแบบที่ศึกษามีดังนี้

1. ตัวแบบเส้นตรง
2. ตัวแบบพหุนามลำดับที่ 2
3. ตัวแบบพหุนามลำดับที่ 3
4. ตัวแบบตรีโกณมิติแบบที่ 1
5. ตัวแบบตรีโกณมิติแบบที่ 2
6. ตัวแบบเลขชี้กำลัง

ตัวแบบที่ 1 เป็นตัวแบบภายใต้สมมติฐานว่าง ตัวแบบที่ 2 -6 เป็นตัวแบบภายใต้สมมติฐานแย้ง

ขั้นตอนการสร้างข้อมูล 1 ชุด

1. จำลองค่า (Simulate) ตัวแปรอิสระ X โดยสุ่มจากการแจกแจงยูนิฟอร์ม (1, 10) ให้มีขนาดตัวอย่างตามที่กำหนด

2. กำหนดค่าพารามิเตอร์ ภายใต้สถานการณ์ต่าง ๆ ที่ต้องการศึกษา

ตารางที่ 3.1

แสดงตัวแบบและค่าพารามิเตอร์ของแต่ละตัวแบบ

ตัวแบบที่	ตัวแบบที่ศึกษา	ค่าพารามิเตอร์
1.เส้นตรง	$Y = \beta_0 + \beta_1 X + \varepsilon$	$\beta_0 =$ 2 $\beta_1 =$ 0.0001 0.001 0.01 0.05 0.1 0.5 1 1.5 2
2.พหุนามลำดับที่ 2	$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon$	$\beta_0 = \beta_1 =$ 2 $\beta_2 =$ 0.0001 0.001 0.01 0.05 0.1 0.5 1
3.พหุนามลำดับที่ 3	$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \varepsilon$	$\beta_0 = \beta_1 =$ 2 $\beta_2 =$ 0.2 $\beta_3 =$ 0.0001 0.001 0.01 0.05 0.1 0.5 1
4.ตัวแบบตรีโกณมิติ แบบที่ 1	$Y = \beta_0 + \beta_1 \sin X_1 + \varepsilon$	$\beta_0 =$ 2 $\beta_1 =$ 0.01 0.1 0.5 1.5 2
5.ตัวแบบตรีโกณมิติ แบบที่ 2	$Y = \beta_0 + \{\beta_1 \cos(X_1) + \beta_1 \sin(X_1)\} + \varepsilon$	$\beta_0 =$ 2 $\beta_1 =$ 0.01 0.1 0.5 1.5 2
6.ตัวแบบเลขชี้กำลัง	$Y = \exp(\beta_0 + \beta_1 X).\varepsilon$	$\beta_0 =$ 2 $\beta_1 =$ 0.01 0.1 0.5 1 1.5

หมายเหตุ ตัวแบบที่ 1 เป็นตัวแบบภายใต้สมมติฐานว่าง ตัวแบบที่ 2 -6 เป็นตัวแบบภายใต้สมมติฐานแย้ง

3. จำลองค่าความคลาดเคลื่อนสุ่มตามขนาดตัวอย่างให้มีคุณสมบัติตามการแจกแจงที่ต้องการ

4. สร้างค่าตัวแปรตาม Y ให้มีขนาดตัวอย่างตามต้องการ โดยมีความสัมพันธ์กับตัวแปร X ตามที่กำหนดไว้ในตารางที่ 3.1

แบ่งกลุ่มของข้อมูลให้มีจำนวนกลุ่มตามต้องการ

เมทริกซ์ที่ใช้เพื่อแสดงกลุ่มของข้อมูลจะแทนด้วยเมทริกซ์ **Z** เมทริกซ์ดังกล่าวมีขนาด $n \times c$ โดย n คือขนาดตัวอย่าง c คือจำนวนกลุ่ม เมทริกซ์ **Z** จะประกอบไปด้วยตัวเลข 0 และ1 เท่านั้น ค่าในคอลัมน์ที่ไม่เป็น 0 จะมีความสัมพันธ์กับขนาดตัวอย่างในกลุ่มที่ j $j = 1, 2, ..., c$ ยกตัวอย่างเช่น ข้อมูลมีขนาดตัวอย่างเท่ากับ 11 ต้องการแบ่งกลุ่มออกเป็น 3 กลุ่ม ดังนี้

หน่วยตัวอย่างค่าที่ 1, 2, 3, 4 อยู่กลุ่ม 1

หน่วยตัวอย่างค่าที่ 5, 6, 7 อยู่กลุ่ม 2

หน่วยตัวอย่างค่าที่ 8, 9, 10, 11 อยู่กลุ่ม 3

รูปแบบของ **Z** จะแสดงได้ดังนี้

$$\mathbf{Z}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

วิธีของมิลเลอร์และนีลล์จะกำหนดค่าสังเกตในแต่ละกลุ่มเท่ากับ 2, 3, 4, 5 จำนวนกลุ่มจึงมีดังต่อไปนี้

ตารางที่ 3.2

จำนวนกลุ่มสำหรับวิธีของมิลเลอร์และนีลล์

ขนาดตัวอย่าง	วิธีของมิลเลอร์และนีลล์
15	c = (7, 5, 4, 3)
50	c = (25, 17, 13, 10)
100	c = (50, 33, 25, 20)

หมายเหตุ วิธีของมิลเลอร์และนีลล์มีค่าสังเกตในแต่ละกลุ่มเท่ากับ 2, 3, 4, 5

วิธีของเดรปเปอร์และสมิทธิ์ กับ วิธีของซูและยาง มีการแบ่งกลุ่มของข้อมูล 2 วิธี คือ

1. แบ่งกลุ่มของข้อมูลโดยใช้ฟังก์ชัน cutree วิธีนี้โปรแกรมจะแบ่งข้อมูลเป็นกลุ่มเอง โดยแต่ละกลุ่มอาจมีขนาดตัวอย่างในแต่ละกลุ่มไม่เท่ากัน หลักการของฟังก์ชัน cutree คือข้อมูลที่มีค่าใกล้เคียงกันอยู่ในกลุ่มเดียวกันในขณะที่ต่างกลุ่มกันมีค่าแตกต่างกัน

2. แบ่งกลุ่มของข้อมูลให้ในแต่ละกลุ่มให้มีขนาดตัวอย่างเท่าๆ กัน วิธีนี้ผู้วิจัยจะกำหนดให้ข้อมูลในแต่ละกลุ่มให้มีขนาดตัวอย่างเท่าๆ กัน ยกเว้นกรณีที่ไม่สามารถจัดให้เท่ากันทุกกลุ่มจะให้กลุ่มสุดท้ายที่จัดมีขนาดแตกต่างจากกลุ่มอื่น เช่น n = 15 ต้องการจัดให้แต่ละกลุ่มมีขนาดตัวอย่างเท่ากับ 4 สามารถจัดได้ 4 กลุ่ม โดยมีขนาดตัวอย่างในแต่ละกลุ่ม ดังนี้

กลุ่ม 1 ขนาดตัวอย่างเท่ากับ 4

กลุ่ม 2 ขนาดตัวอย่างเท่ากับ 4

กลุ่ม 3 ขนาดตัวอย่างเท่ากับ 4

กลุ่ม 4 ขนาดตัวอย่างเท่ากับ 3

จำนวนกลุ่มในแต่ละขนาดตัวอย่าง (n =15, 50, 100) มีดังนี้

ตารางที่ 3.3

จำนวนกลุ่มสำหรับวิธีของเดรปเปอร์และสมิทธิ์ กับ วิธีของซูและยางกรณีแบ่งกลุ่มของข้อมูลโดยใช้ฟังก์ชัน cutree และการแบ่งกลุ่มของข้อมูลให้ในแต่ละกลุ่มให้มีขนาดตัวอย่างเท่าๆ กัน

ขนาดตัวอย่าง	จำนวนกลุ่ม (c)	
	วิธีของเดรปเปอร์และสมิทธิ์	วิธีของซูและยาง
15	c = 3,4,5,7	c = 3,4,5,7
50	c = 3,4,5,7,10,13,17,25	c = 3,4,5,7,10,13,17,25
100	c = 3,4,5,7,10,13,17,20,25,33,50	c = 3,4,5,7,10,13,17,20,25,33,50

ตารางที่ 3.4

ขนาดตัวอย่างในแต่ละกลุ่มสำหรับวิธีของเดรปเปอร์และสมิทธิ์ กับ วิธีของซูและยาง กรณีแบ่งกลุ่มของข้อมูลให้ในแต่ละกลุ่มให้มีขนาดตัวอย่างเท่าๆ กัน

ขนาดตัวอย่าง	จำนวนกลุ่ม	กลุ่มที่ (G_m)	ขนาดตัวอย่างในแต่ละกลุ่ม
15	3	G_1, G_2, G_3	5, 5, 5
	4	G_1, G_2, G_3, G_4	4, 4, 4,3
	5	$G_1, G_2, G_3, ..., G_5$	3, 3, 3, 3, 3
	7	$G_1, G_2, G_3, ..., G_{13}$	2, 2, 2, 2, 2, 2, 3

ตารางที่ 3.4 (ต่อ)

ขนาดตัวอย่างในแต่ละกลุ่มสำหรับวิธีของเดรปเปอร์และสมิทธิ์ กับ วิธีของซูและยาง

กรณีการแบ่งข้อมูลในแต่ละกลุ่มให้มีขนาดตัวอย่างเท่าๆ กัน

ขนาดตัวอย่าง	จำนวนกลุ่ม	กลุ่มที่ (G_m)	ขนาดตัวอย่างในแต่ละกลุ่ม
50	3	G_1, G_2, G_3	17, 17, 16
	4	G_1, G_2, G_3, G_4	12, 12, 12, 14
	5	$G_1, G_2, G_3, ..., G_5$	10, 10, 10, 10, 10
	7	$G_1, G_2, G_3, ..., G_7$	7, 7, 7, 7, 7, 7, 8
	10	$G_1, G_2, G_3, ..., G_{10}$	5, 5, 5, 5, 5, 5, 5, 5, 5, 5
	13	$G_1, G_2, G_3, ..., G_{13}$	4, 4, 4,…,2
	17	$G_1, G_2, G_3, ..., G_{17}$	3, 3, 3,…,2
	25	$G_1, G_2, G_3, ..., G_{25}$	2, 2, 2,…,2
100	3	G_1, G_2, G_3	33, 33, 34
	4	G_1, G_2, G_3, G_4	25, 25, 25, 25
	5	$G_1, G_2, G_3, ..., G_5$	20, 20, 20, 20, 20
	7	$G_1, G_2, G_3, ..., G_7$	14, 14, 14, 14, 14, 16
	10	$G_1, G_2, G_3, ..., G_{10}$	10, 10, 10,…,10
	13	$G_1, G_2, G_3, ..., G_{13}$	8, 8, 8,…,4
	17	$G_1, G_2, G_3, ..., G_{17}$	6, 6, 6,…, 4
	20	$G_1, G_2, G_3, ..., G_{20}$	5, 5, 5,…,5
	25	$G_1, G_2, G_3, ..., G_{25}$	4, 4, 4,…,4
	33	$G_1, G_2, G_3, ..., G_{33}$	3, 3, 3,…,4
	50	$G_1, G_2, G_3, ..., G_{50}$	2, 2, 2,…,2

ฟังก์ชัน cutree

ฟังก์ชัน cutree ในโปรแกรม R เป็นฟังก์ชันที่ใช้ในการจัดกลุ่มข้อมูลซึ่งเกิดจากการใช้คำสั่ง hclust เพื่อให้ได้จำนวนกลุ่มตามต้องการ คำสั่งมีดังนี้

cutree(tree , k)

เมื่อ tree คือ ข้อมูลที่จะใช้ในจัดกลุ่มต้องอยู่ในรูปเมทริกซ์

k คือ จำนวนกลุ่มที่ต้องการ

สำหรับฟังก์ชัน cutree นี้จะมีการจัดข้อมูลที่มีค่าใกล้เคียงกันอยู่ในกลุ่มเดียวกัน

คำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่ากำลังการทดสอบของสถิติทดสอบทั้ง 3 วิธี คือ วีธีของเดรปเปอร์และสมิทธิ์ วิธีของซูและยาง วิธีของมิลเลอร์และนีลล์

1. **สถิติทดสอบที่เสนอโดยเดรปเปอร์และสมิทธิ์**

1.1 สร้างแบบจำลองข้อมูลตามสถานการณ์ที่ต้องการศึกษา

1.2 จัดกลุ่มข้อมูล

1.3 สร้างเมทริกซ์ $\mathbf{X}_{\mathbf{mean}}$ โดยแทนที่ค่าของตัวแปรอิสระ $\mathbf{X}$ ด้วยค่าเฉลี่ยของตัวแปรอิสระ $\mathbf{X}$ ภายในกลุ่มเดียวกัน

1.4 นำเมทริกซ์ $\mathbf{X}_{\mathbf{mean}}$ จากข้อ 1.3 และ $\mathbf{Y}$ ที่ได้จากการจำลองในขั้นตอนการสร้างข้อมูล 1 ชุด มาคำนวณค่าผลบวกกำลังสอบของความคลาดเคลื่อนแทนด้วย $\mathrm{SSE}_{\mathrm{mean}}$

$$\mathrm{SSE}_{\mathrm{mean}} = \mathbf{Y}'\left\{\mathbf{I}_{\mathbf{n}} - \mathbf{X}_{\mathbf{mean}}\left(\mathbf{X}'_{\mathbf{mean}}\mathbf{X}_{\mathbf{mean}}\right)^{-1}\mathbf{X}'_{\mathbf{mean}}\right\}\mathbf{Y}$$

1.5 คำนวณค่าผลบวกกำลังสองของความคลาดเคลื่อนที่แท้จริงจากตัวแบบ แทนด้วย SSPE

$$\mathrm{SSPE} = \sum_{i=1}^{c}\sum_{j=1}^{n_i}(\mathrm{Y}_{ij} - \overline{\mathrm{Y}}_i)^2$$

1.6 คำนวณค่าผลบวกกำลังสองของความคลาดเคลื่อนเนื่องจากตัวแบบไม่เหมาะสมจาก $\mathrm{SSE}_{\mathrm{mean}} - \mathrm{SSPE}$ แทนด้วย SSLF

1.7 คำนวณสถิติทดสอบที่ใช้ทดสอบสมมติฐานว่างของวิธีที่เสนอโดยเดรปเปอร์และสมิทธิ์

$$F_{DS} = \frac{SSLF/(n-k-1)}{SSPE/(n-c)}$$

เมื่อ n คือ ขนาดตัวอย่าง

c คือ จำนวนกลุ่ม

k คือ จำนวนตัวแปรอิสระ

จะปฏิเสธสมมติฐานว่าง $\left(H_0\right)$ เมื่อ F_{DS} ที่คำนวณมีค่ามากกว่า $F_{(c-k-1,n-c)}$

1.8 เปรียบเทียบค่าสถิติทดสอบของเดรปเปอร์และสมิทธิ์กับบริเวณวิกฤติ

1.9 กระทำการทดลองซ้ำ ๆ กัน 2000 ครั้ง ต่อ 1 สถานการณ์ที่กำหนดขึ้นเองของตัวแบบพารามิเตอร์ ค่าของตัวแปรอิสระ X แล้วทำการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง

1.10 คำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1ของวิธีเดรปเปอร์และสมิทธิ์โดยสร้างข้อมูลให้มีรูปแบบเชิงเส้นตรง จะได้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการนำจำนวนที่ปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นจริงหารด้วย 2000

1.11 คำนวณหาค่ากำลังการทดสอบ (Power of the test) ของวิธีเดรปเปอร์และสมิทธิ์ โดยสร้างข้อมูลให้มีรูปแบบไม่เชิงเส้น จะได้ค่ากำลังการทดสอบจากการนำจำนวนที่ปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นเท็จหารด้วย

2. สถิติทดสอบที่เสนอโดยซูและยาง

2.1 สร้างแบบจำลองข้อมูลตามสถานการณ์ที่ต้องการศึกษา

2.2 สร้างเมทริกซ์ **Z**

2.3 นำค่าตัวแปรอิสระ X ที่ได้จากการจำลองในขั้นตอนการสร้างข้อมูล 1 ชุด มาแทนในตำแหน่งที่กำหนดในเมทริกซ์ **Z** จะได้เมทริกซ์ **X**

2.4 นำเมทริกซ์ **X** มาคำนวณค่าผลบวกกำลังสองของความคลาดเคลื่อน จะแทนด้วย SSE_x

$$SSE_x = \mathbf{Y}'\left\{\mathbf{I_n} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\right\}\mathbf{Y}$$

2.5 คำนวณค่าผลบวกกำลังสองของความคลาดเคลื่อนคำนวณจากตัวแบบ $\mathbf{E}(\mathbf{Y}) = \mathbf{X\beta} + \mathbf{W\alpha}$ แทนด้วย SSE_{XW} โดยมีการเพิ่มเทอมกำลังสองของตัวแปรอิสระจะได้เมทริกซ์ $\mathbf{X_w}$ เพื่อใช้ในการคำนวณ SSE_{XW}

$$SSE_{XW} = \mathbf{Y}'\left\{\mathbf{I_n} - \mathbf{X_w}\left(\mathbf{X}'_\mathbf{w}\mathbf{X_w}\right)^{-1}\mathbf{X}'_\mathbf{w}\right\}\mathbf{Y}$$

2.6 คำนวณค่าสถิติทดสอบ F_{SY} จาก

$$F_{SY} = \frac{\left. SSE_x - SSE_{xw} \middle/ (r_x - r_{xw}) \right.}{\left. SSE_{xw} \middle/ (n - r_{xw}) \right.}$$

เมื่อ n คือ ขนาดตัวอย่าง

r_X คือ rank ของเมทริกซ์ $\mathbf{X}$

r_{XW} คือ rank ของเมทริกซ์ $\mathbf{X_w}$

จะปฏิเสธสมมติฐานว่าง (H_0) เมื่อ F_{SY} ที่คำนวณมีค่ามากกว่า $F_{r_X - r_{XW}, n-r_{XW}}$

2.7 เปรียบเทียบค่าสถิติทดสอบของซูและยางกับบริเวณวิกฤติ

2.8 กระทำการทดลองซ้ำ ๆ กัน 2000 ครั้ง ต่อ 1 สถานการณ์ที่กำหนดขึ้นเองของตัวแบบ พารามิเตอร์ ค่าของตัวแปรอิสระ X แล้วทำการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง

2.9 คำนวณค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ของ วิธีซูและยาง โดยสร้างข้อมูลให้มีรูปแบบเชิงเส้นตรง จะได้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการนำจำนวนที่ปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นจริง หารด้วย 2000

2.10 คำนวณหาค่ากำลังการทดสอบ (Power of the test) ของวิธีซูและยาง โดยสร้างข้อมูลให้มีรูปแบบไม่เชิงเส้น จะได้ค่ากำลังการทดสอบจากการนำจำนวนที่ปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นเท็จหารด้วย 2000

3. สถิติทดสอบมิลเลอร์และนีลล์

3.2 สร้างแบบจำลองข้อมูลตามสถานการณ์ที่ต้องการศึกษา

3.2 นำตัวแปรอิสระ X ที่ได้จากการจำลองค่ามาเรียงจากน้อยไปมาก เพื่อใช้ในการสร้างเมทริกซ์ **X** โดยค่าตัวแปรอิสระ X ค่าน้อยจะอยู่กลุ่มแรก

3.3 สร้างเมทริกซ์ **X** ซึ่งมีรูปแบบ $\begin{bmatrix} 1 & X_i \end{bmatrix}$; $i = 1, 2, ..., n$

3.4 สร้างเมทริกซ์ **Z** สำหรับการจัดกลุ่มให้มีขนาดตัวอย่างกลุ่มละ 2, 3, 4 และ5

3.5 คำนวณค่าสถิติทดสอบ F_B จาก

$$F_B = \frac{\| P_B\ Y \|^2 \big/ \dim B}{\| P_{W_E}\ Y \|^2 \big/ \dim W_E}$$

เมื่อ $\mathbf{A} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, $\mathbf{P} = \mathbf{AX}$ โดยที่ $\mathbf{X} = = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ 1 & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix}$

$\mathbf{MB} = \mathbf{A} - \mathbf{P}(\mathbf{P}'\mathbf{P})\mathbf{P}'$

$\| P_B\ Y \|^2$ คือ $\mathbf{Y}'(\mathbf{MB}'\mathbf{MB})\mathbf{Y}$

$\dim B$ คือ rank ของเมทริกซ์ **MB**

$\| P_{W_E} Y \|^2$ คือ $\left\{ \mathbf{I_n} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}' \right\} - \| P_B\ Y \|^2$

$\dim W_E$ คือ ผลต่างระหว่าง rank ของเทอม $\left\{ \mathbf{I_n} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}' \right\}$ และ $\dim B$ และจะได้ค่า F_{W_E} จาก $1/F_B$

3.6 คำนวณค่า $\mathbf{a_n}$ โดยที่จำลองค่า e ให้มีการแจกแจง $N(0, \sigma^2 I_n)$ จำนวน 10000 ค่า หาค่า F_{BR_m} จากการคำนวณค่าสถิติทดสอบเช่นเดียวกับข้อ 3.5 แต่แทนที่ **Y** ด้วย e แล้วนำ F_{BR_m} มาหาระดับนัยสำคัญในแต่ละขนาดของค่าสังเกตในแต่ละกลุ่ม แทนด้วย Fcompvals เลือกค่า Fcompvals ที่น้อยที่สุดจาก Fcompvals ที่ได้จากในแต่ละขนาดของค่าสังเกตในแต่ละกลุ่มในที่นี้กำหนดให้เท่ากับ 2, 3, 4 และ 5 แทนด้วย inf จะได้ $\mathbf{a_n}$ จากควอไทล์ของ inf

3.7 นำ $\mathbf{a_n}$ มาหาพื้นที่ใต้เส้นโค้งโดยใช้ฟังก์ชัน $\mathbf{qf}(1-\mathbf{a_n}, \dim \mathbf{B}, \dim \mathbf{W_E})$ แทนด้วย Bcpt

3.8 นำ $\mathbf{a_n}$ มาหาพื้นที่ใต้เส้นโค้งโดยใช้ฟังก์ชัน $\mathbf{qf}(1-\mathbf{a_n}, \dim \mathbf{W_E}, \dim \mathbf{B})$ แทนด้วย WScpt

3.9 หาผลต่างระหว่าง F_B กับ Bcpt แทนด้วย diffB และผลต่างระหว่าง F_{W_E} กับ WScpt แทนด้วย diffWS เปรียบเทียบค่า diffB diffWS แล้วเลือกค่าที่สูงที่สุด แทนด้วย Tval

3.10 ทดสอบสมมติฐานโดยจะปฏิเสธสมมติฐานว่างเมื่อ Tval > 0

3.11 กระทำการทดลองซ้ำ ๆ กัน 2000 ครั้ง ต่อ 1 สถานการณ์ที่กำหนดขึ้นเองของตัวแบบ พารามิเตอร์ ค่าของตัวแปรอิสระ X แล้วทำการนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง

3.12 คำนวณค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทที่ 1 ของวิธีมิลเลอร์และนีลล์ โดยสร้างข้อมูลให้มีรูปแบบเชิงเส้นตรง จะได้ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากการนำจำนวนที่ปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นจริงหารด้วย 2000

3.13 คำนวณหาค่ากำลังการทดสอบ (Power of the test) ของวิธีมิลเลอร์และนีลล์ โดยสร้างข้อมูลให้มีรูปแบบไม่เชิงเส้น จะได้ค่ากำลังการทดสอบจากการนำจำนวนที่ปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นเท็จหารด้วย 2000

วิเคราะห์เปรียบเทียบประสิทธิภาพของตัวสถิติทั้ง 3 ชนิด

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการตรวจสอบความเหมาะสมของตัวแบบ กรณีข้อมูลไม่ซ้ำด้วยการทดสอบ 3 วิธี โดยพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ในแต่ละสถานการณ์ โดยทำการเปรียบเทียบค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง (τ) ว่าอยู่ในขอบเขตที่ยอมรับได้หรือไม่ โดยกำหนดระดับนัยสำคัญของการทดสอบ (α) เท่ากับ0.05 ซึ่งเกณฑ์ที่ใช้ในการพิจารณาความสามารถในการควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่1 จะใช้เกณฑ์ของ Bradley (Rheinheimer,D.C. และ Penpield,D.A.2001: 373 - 391) ดังรายละเอียดต่อไปนี้

ถ้าค่า τ เป็นค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่เกิดจากการทดลอง เมื่อ τ อยู่ในช่วง $(0.5\alpha, 1.5\alpha)$ จะถือว่าการทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ หมายความว่า ที่ระดับนัยสำคัญ (α) เท่ากับ 0.05 ค่า τ จะต้องอยู่ในช่วง

(0.025, 0.075) จึงจะถือว่าการทดสอบนั้นสามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้เท่ากับค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ระบุ

จากผลการทดลองถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของการทดสอบอยู่นอกขอบเขตที่ระบุจะถือว่าการทดสอบนั้นไม่สามารถควบคุมอัตราความคลาดเคลื่อนประเภทที่ 1 ได้ ซึ่งแยกเป็น 2 กรณี คือ

1. กรณีที่ค่าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่ามากกว่าขอบเขตบนของเกณฑ์ที่ใช้พิจารณา จะถือว่าอัตราความคลาดเคลื่อนประเภทที่ 1 มีค่ามากกว่าระดับนัยสำคัญที่ระบุ $(\tau > \alpha)$

2. กรณีที่ค่าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าน้อยกว่าขอบเขตล่างของเกณฑ์ที่ใช้พิจารณา จะถือว่าอัตราความคลาดเคลื่อนประเภทที่ 1 มีค่าน้อยกว่าระดับนัยสำคัญที่ระบุ $(\tau < \alpha)$

ส่วนกรณีที่สามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จะถือว่าอัตราความคลาดเคลื่อนประเภทที่ 1 มีค่าเท่ากับค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ระบุ $(\tau = \alpha)$

พิจารณาค่ากำลังการทดสอบในแต่ละสถานการณ์ ของสถิติทดสอบทั้ง 3 วิธี หากค่ากำลังการทดสอบมีค่าใกล้เคียงหรือเท่ากับ 1 แสดงว่าการทดสอบดังกล่าวมีประสิทธิภาพค่อนข้างสูงในสถานการณ์นั้น