บทที่ 3
วิธีดำเนินการวิจัย

วิธีการดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงทดลองซึ่งจำลองขึ้นด้วยโปรแกรมในคอมพิวเตอร์ โดยวิธีจำลองข้อมูลแบบ มอนติ คาร์โล (Monte Carlo Simulation Method) จากโปรแกรม MATLAB เพื่อหาการประมาณช่วงความเชื่อมั่นสำหรับขนาดอิทธิพลของข้อมูลที่มีการแจกแจงแบบเบ้ขวา 4 แบบ ได้แก่ การแจกแจงแบบไคกำลังสองการแจกแจงแบบแกมมา และ การแจก ,การแจกแจงแบบลอกนอร์มอล , แจงแบบไวบูลล์ส์ ที่มีระดับความเบ้แตกต่างกัน ดังนี้ 0.5, 1.0, 1.5, 2.0 และ 2.5 แต่มีระดับความโด่งที่ใกล้เคียงกัน กับกลุ่มตัวอย่างที่เป็นอิสระและไม่เป็นอิสระกัน กำหนดขนาดตัวอย่าง = 4, 8, 16, 32, 64 และ128 และมีค่าสัมประสิทธิ์ความเชื่อมั่น = 0.90, 0.95 และ 0.99

โดยแต่ละกรณีจะทำการทดลองซ้ำ 2,000 ครั้ง จากโปรแกรม MATLAB 7.1

แผนการดำเนินงาน

ในการทดลองครั้งนี้กำหนดสถานการณ์ ดังนี้

1. สัมประสิทธิ์ความเบ้ของการแจกแจงแบบต่างๆ = 0.5, 1.0, 1.5, 2.0 และ 2.5
2. ขนาดตัวอย่าง (กำหนดตาม Veichtbauer, 2007)

    2.1ขนาดตัวอย่างกรณีตัวอย่าง 2 กลุ่มเป็นอิสระต่อกัน

    - ขนาดตัวอย่างของทั้งสองกลุ่มไม่เท่ากันเมื่อกำหนดอัตราส่วนเป็น 1 : 3 คือ (2,6), (4,12), (8,24), (16,48) และ (32,96)
    - ขนาดตัวอย่างของทั้งสองกลุ่มไม่เท่ากันเมื่อกำหนดอัตราส่วนเป็น 1 : 7 คือ (2,14), (4,28), (8,56) และ (16,112)

    2.2 ขนาดตัวอย่างกรณีตัวอย่าง 2 กลุ่มไม่เป็นอิสระต่อกัน คือ 8, 16, 32, 64 และ 128
3. ค่าสัมประสิทธิ์ความเชื่อมั่น = 0.90, 0.95 และ 0.99

ขั้นตอนในการวิจัย

1.สร้างข้อมูลให้เป็นไปตามการแจกแจงที่กำหนดสร้างข้อมูลตามการแจกแจงทั้ง 4 แบบ โดยการใช้เลขสุ่ม (random number)

-การสร้างเลขสุ่มที่มีการแจกแจงไคกำลังสอง (XC) เมื่อ $Z_i \sim N(0,1)$ โดยที่ Z_i เป็นอิสระกัน จะได้ว่า $X = \sum_{i=1}^{n} Z_i^2$ ซึ่งมีระดับขั้นความเสรี = n ในการสร้างเลขสุ่มที่มีการแจกแจงแบบนี้ ทำได้โดยสร้างเลขสุ่มที่มีการแจกแจงแบบปกติ และเป็นอิสระต่อกันขึ้นมาเท่ากับจำนวนลำดับขั้นความเสรี แล้วนำเลขสุ่มแต่ละตัวมายกกำลังสอง แล้วเอามาบวกกัน เราจะได้ค่าเลขสุ่มที่มีการแจกแจงแบบไคกำลังสองและมีระดับขั้นความเสรี = n

-การสร้างเลขสุ่มที่มีการแจกแจงแบบลอกนอร์มอล (XL) การแจกแจงแบบลอกนอร์มอลมีความสัมพันธ์กับการแจกแจงแบบปกติ คือ ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย = μ และมีความแปรปรวน = σ^2 แล้วทำให้ $Y = Exp(X)$ มีการแจกแจงแบบลอกนอร์มอล การสร้างเลขสุ่มด้วยพารามิเตอร์ μ และ σ^2 ทำได้โดยสร้างจากค่าชี้กำลังของเลขสุ่มที่มีการแจกแจงแบบปกติที่ได้

-การสร้างเลขสุ่มที่มีการแจกแจงแบบแกมมา (XG) ด้วยพารามิเตอร์ α และ β มี 3 กรณี ดังนี้

1. กรณี $0 < \alpha < 1$ ทำได้โดย หาค่า b จากสมการ $b = (e - \alpha)/e$

1.1 สร้างเลขสุ่ม (R_1, R_2)ที่มีการแจกแจงสม่ำเสมอในช่วง (0,1) และให้ $(P = bR_1)$ ถ้า P>1 ให้ทำข้อ 3

1.2 ให้ $Y = P^{1/\alpha}$ ถ้า $R_2 \le e^{-Y}$ ให้ XG=Y สำหรับกรณีอื่นๆให้กลับไปทำข้อ 1.2

1.3 ถ้า $Y = \ln[(b - p)/\alpha]$ ถ้า $R_2 \le Y^{\alpha - 1}$ ให้ XG=Y สำหรับกรณีอื่นๆให้กลับไปทำข้อ 1.2

2. กรณี $\alpha = 1$ ซึ่งมีการกำหนด β เป็นพารามิเตอร์ของการแจกแจงนี้ = 1 ดังนั้น ถ้า X_i เป็นตัวแปรสุ่มที่มีการแจกแจงแบบแกมมาด้วยพารามิเตอร์ $\alpha = 1$ และ $\beta = 1$ แล้ว X_i จะมีการแจกแจงชี้กำลังด้วยพารามิเตอร์ $\beta = 1$ จึงสามารถสร้างเลขสุ่มได้จากการสร้างเลขสุ่มที่มีการแจกแจงชี้กำลังด้วยพารามิเตอร์ $\beta = 1$ ซึ่งอาศัยเทคนิคการแปลงผกผัน ดังนี้

$XG = -\beta \ln R$; R_i เป็นเลขสุ่มที่มีการแจกแจงสม่ำเสมอในช่วง (0,1)

3. กรณี $\alpha > 1$ มีขั้นตอนดังนี้

3.1 คำนวณค่า $a_1, a_2 a_3, a_4$ จากสมการ

$$a_1 = 1/\sqrt{2\alpha - 1}$$

$a_2 = \alpha - \ln 4$

$a_3 = \alpha + 1/a$

$a_4 = 1 + \ln 4.5$

3.2 สร้างเลขสุ่ม R_1 และ R_2 ที่มีการแจกแจงสม่ำเสมอในช่วง (0,1)

3.3 กำหนด $V = a_1 \ln[R_1/(1-R_1)]$

$Y = \alpha e^v$

$Z = R_1^2 R_2$

$W = a_2 + a_3 V - Y$

ถ้า $W + a_4 - 4.5Z \geq 0$ ให้ XG=Y สำหรับกรณีอื่นๆให้ทำข้อ 3.4

3.4 ถ้า $W \geq \ln Z$ ให้ XG=Y สำหรับกรณีอื่นๆให้กลับไปทำข้อ 3.2

-<u>การสร้างเลขสุ่มที่มีการแจกแจงแบบไวบูลส์ (XW)</u> ด้วยพารามิเตอร์ α และ β โดยใช้สมการ $XW = \beta[-\ln(1-R)]^{1/\alpha}$ เมื่อ R เป็นเลขสุ่มที่มีการแจกแจงสม่ำเสมอในช่วง (0,1)

2. คำนวณช่วงความเชื่อมั่นสำหรับขนาดอิทธิพลมาตรฐานโดยใช้ข้อมูลที่มีการแจกแจงแบบไม่ปกติ ได้แก่ การแจกแจงแบบไคกำลังสอง, การแจกแจงแบบแกมมา, การแจกแจงแบบลอกนอร์มอล และการแจกแจงแบบไวบูลส์ และมีระดับความเบ้และความโด่งตามที่กำหนด โดยคำนวณช่วงความเชื่อมั่นจากวิธีทั้ง 10 วิธี

3. ตรวจสอบว่าช่วงความเชื่อมั่นใดที่เหมาะสมในสถานการณ์ที่กำหนดขึ้นมา โดยใช้โปรแกรม MATLAB เพื่อทำการทดลองซ้ำ 2,000 ครั้ง โดยการคำนวณค่าสัมประสิทธิ์ความเชื่อมั่นในแต่ละสถานการณ์ที่กำหนดในการทดลอง ทำได้โดยนับจำนวนครั้งทั้งหมดที่ช่วงความเชื่อมั่นที่ประมาณได้จะครอบคลุมค่าพารามิเตอร์ แล้วนำค่าที่ได้นี้หารด้วยจำนวนครั้งที่ทำซ้ำทั้งหมด ซึ่งครั้งนี้จะทำการทดลองซ้ำ 2 000,ครั้ง จะได้ค่าสัมประสิทธิ์ความเชื่อมั่นที่คำนวณได้จากการทดลอง ว่า

สัมประสิทธิ์ความเชื่อมั่น = $\frac{1}{2,000}$ (จำนวนครั้งทั้งหมดที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์)

การคำนวณค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น ทำได้โดยคำนวณช่วงความเชื่อมั่นจากแต่ละสูตรและตรวจสอบช่วงความเชื่อมั่นที่ครอบคลุมค่าพารามิเตอร์ แล้วจึงหาผลต่างระหว่างขีดจำกัด

ความเชื่อมั่นบน (U) และขีดจำกัดความเชื่อมั่นล่าง (L) จากนั้นนำผลต่างที่ได้บวกสะสมเอาไว้แล้วจึงหาค่าเฉลี่ยเมื่อทำการคำนวณซ้ำครบ 2,000 ครั้ง จะได้

$$\text{ความยาวเฉลี่ยของช่วงความเชื่อมั่น} = \frac{\sum_{i}^{2000}(U_i - L_i)}{2,000}$$

4. เปรียบเทียบทั้ง 10 วิธีว่า วิธีใดให้ความยาวช่วงความเชื่อมั่นสั้นที่สุดในการประมาณช่วงความเชื่อมั่นสำหรับขนาดอิทธิพลมาตรฐานของกลุ่มตัวอย่างที่มีการแจกแจงแบบเบ้ขวา ทั้ง 4 การแจกแจง โดยกำหนด 5 ความเบ้ และ 3 สัมประสิทธิ์ความเชื่อมั่น ที่แตกต่างกัน ดังนั้น

ตารางที่ 3.1 วิธีที่ใช้ในการคำนวณหาช่วงความเชื่อมั่นสำหรับขนาดอิทธิพล

วิธี	สูตร
gB	$c(m)\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{[c(m)]^2 m\left[1+\tilde{n}g^2\right]}{(m-2)\tilde{n}}-g^2\right)^{\frac{1}{2}}$
dB	$\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{m\left[1+\tilde{n}g^2\right]}{(m-2)\tilde{n}}-\dfrac{d^2}{[c(m)]^2}\right)^{\frac{1}{2}}$
gU	$c(m)\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{1}{\tilde{n}}+\left(1-\dfrac{m-2}{m[c(m)]^2}\right)-d^2\right)^{\frac{1}{2}}$
dU	$\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{1}{\tilde{n}[c(m)]^2}+\left(1-\dfrac{m-2}{m[c(m)]^2}\right)-d^2\right)^{\frac{1}{2}}$
gL1	$c(m)\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{1}{\tilde{n}}+\dfrac{g^2}{2m}\right)^{\frac{1}{2}}$
dL1	$\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{1}{\tilde{n}}+\dfrac{d^2}{2m}\right)^{\frac{1}{2}}$
gL2	$c(m)\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{1}{\tilde{n}}+\dfrac{g^2}{2N}\right)^{\frac{1}{2}}$
dL2	$\dfrac{\left(\overline{X}_E-\overline{X}_C\right)}{S_p}\pm q\times\left(\dfrac{1}{\tilde{n}}+\dfrac{d^2}{2N}\right)^{\frac{1}{2}}$
gH	$a\left(\dfrac{\exp\left[Zg/\sqrt{2}\right]-\exp\left[Zg/\sqrt{2}\right]}{2}\right)$
dH	$a\left(\dfrac{\exp\left[Zd/\sqrt{2}\right]-\exp\left[Zd/\sqrt{2}\right]}{2}\right)$

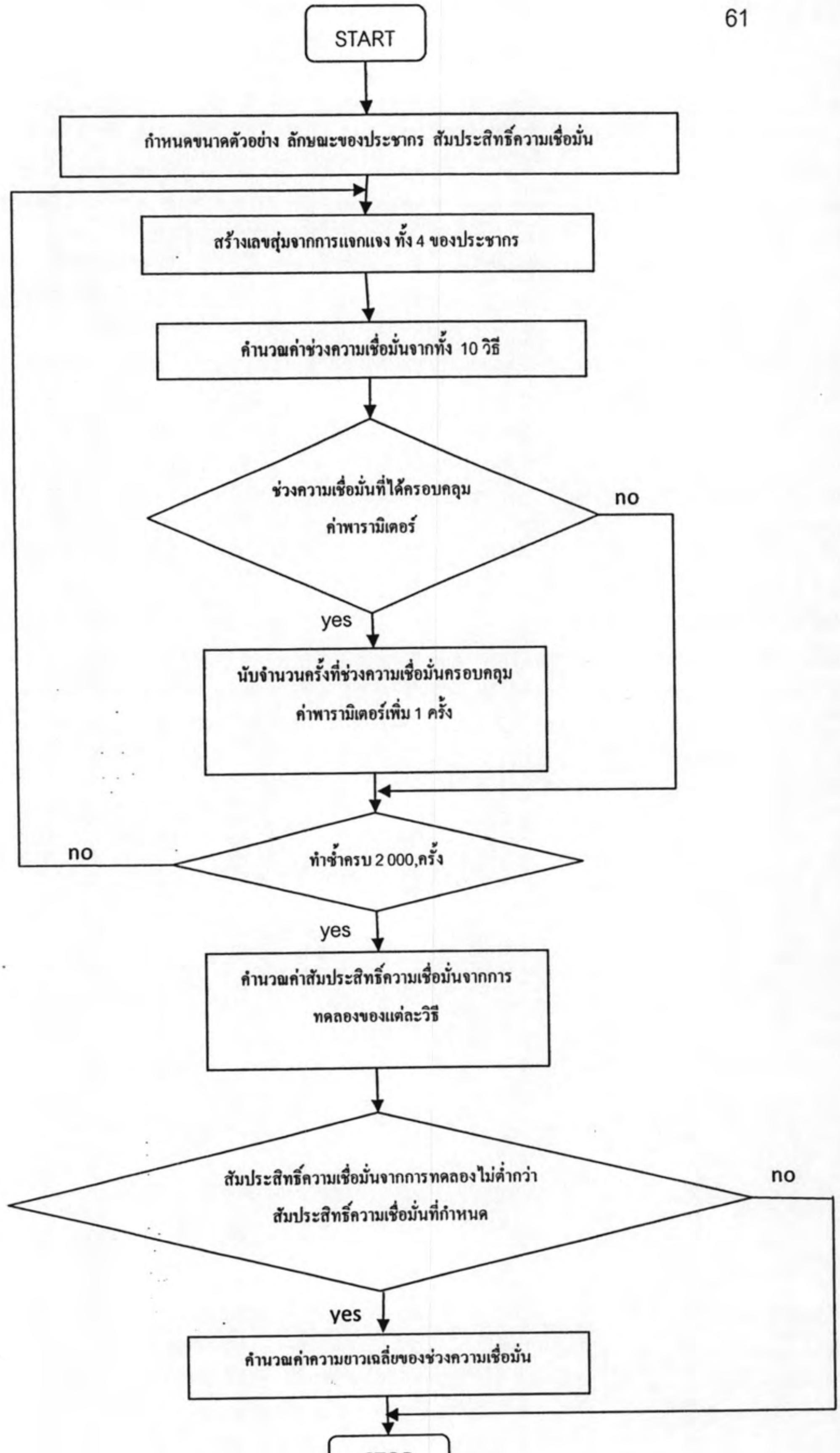
โดยที่

$$g' = c(m)\left(\frac{\overline{X}_E-\overline{X}_C}{S_p}\right)$$

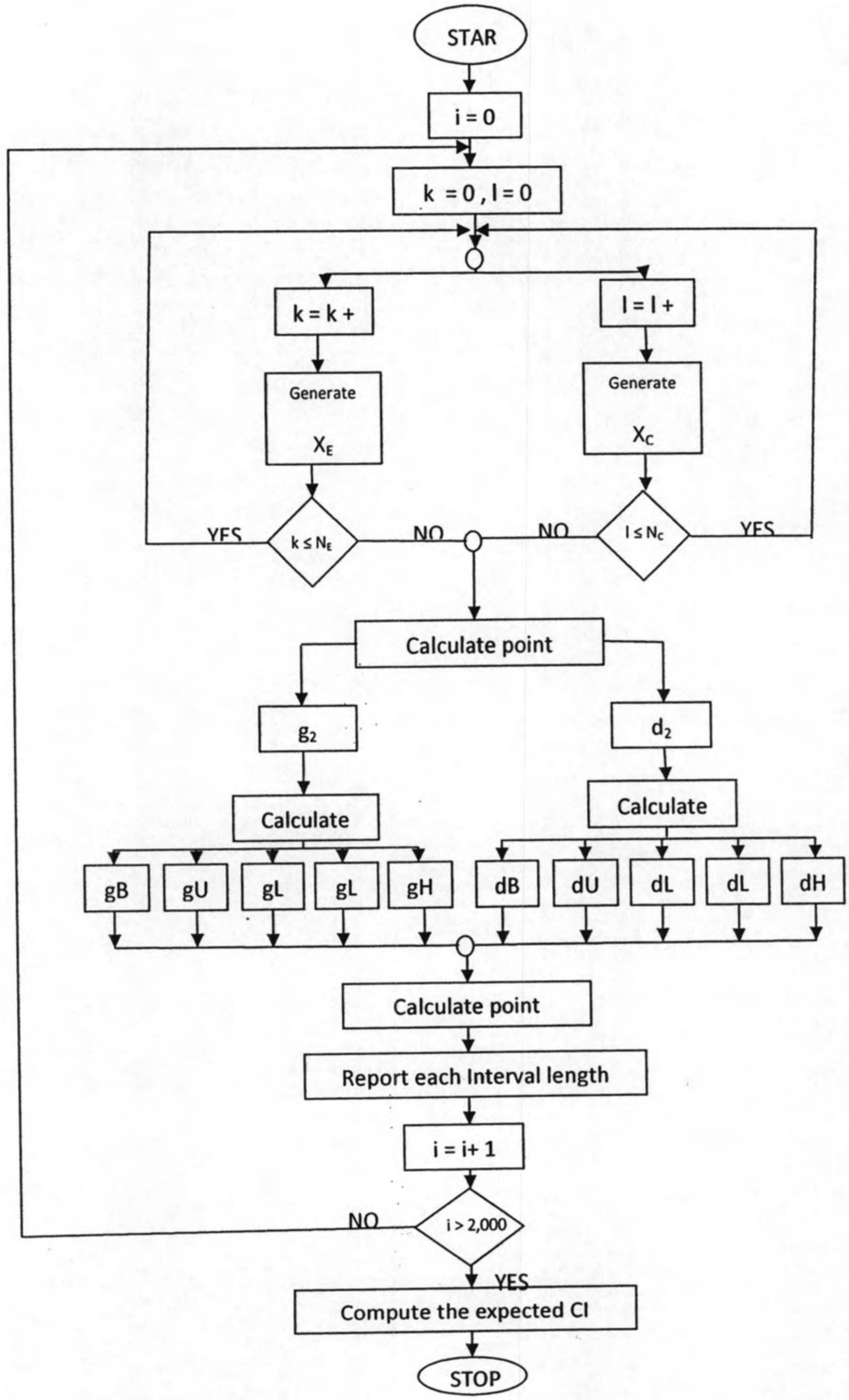
$$q = 100\times\left(1-\frac{\alpha}{2}\right)$$

d	=	$\dfrac{\overline{X}_E - \overline{X}_C}{S_p}$	
c(m)	=	$1 - \dfrac{3}{4m-1}$	
m	=	$n_E + n_C - 2$	two-independent sample case
	=	n	two-dependent sample case with δ_D
$\tilde{n}$	=	$\dfrac{n_E n_C}{(n_E + n_C)}$	two-independent sample case
	=	n	two-dependent sample case with δ_D
N	=	$n_E + n_C$	two-independent sample case
		n	two-dependent sample case with δ_D

5. สรุปผลการวิจัย



ภาพที่ 3.1 แผนผังขั้นตอนการทดลองในการคำนวณช่วงความเชื่อมั่นสำหรับขนาดอิทธิพลมาตรฐาน



ภาพที่ 3.2 แผนผังขั้นตอนการทดลองในการคำนวณช่วงความเชื่อมั่นสำหรับขนาดอิทธิพลมาตรฐาน