หอสมุดกลาง ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลองซึ่งต้องการศึกษาและเปรียบเทียบการแก้ไขปัญหาข้อมูลผิดปกติในการวิเคราะห์ข้อมูลที่ได้จากแผนการทดลองแบบสุ่มตลอดเมื่อปัจจัยทดลองเป็นปัจจัยคงที่ ด้วยวิธีการแปลงข้อมูลโดยการใช้เลขยกกำลัง(The Power transformation)โดยสร้างความคลาดเคลื่อนให้มีการแจกแจงแบบปกติ ซึ่งการจำลองข้อมูลในแต่ละสถานการณ์จะใช้เทคนิคมอนติคาร์โลโดยใช้โปรแกรม S-PLUS 2000 กับเครื่อง PC ดังนั้นรายละเอียดของแผนการดำเนินการวิจัย จะกล่าวในรายละเอียดต่างๆดังต่อไปนี้

3.1 การจำลองด้วยเทคนิคมอนติคาร์โล

เนื่องจากเทคนิคมอนติคาร์โลเป็นเทคนิคที่ถูกนำมาใช้ในการแก้ปัญหาต่างๆเป็นเวลานานแล้วและก็ยังเป็นวิธีที่นิยมใช้กันอยู่ในปัจจุบันและได้มีการพัฒนาเพื่อการนำไปใช้ในสาขาวิชาต่างๆมากขึ้นเช่น สาขาคณิตศาสตร์ สาขาการวิจัยดำเนินงาน เป็นต้น

เทคนิคมอนติคาร์โลเป็นเทคนิคที่ใช้ตัวเลขสุ่มมาช่วยในการแก้ปัญหาในการคำนวณทางคณิตศาสตร์ได้ และช่วยในการหาคำตอบของปัญหาที่ยังไม่แน่ใจในผลที่จะเกิดขึ้น

ตัวเลขสุ่มมีประโยชน์ ดังต่อไปนี้

3.1.1 ทำให้การเลือกตัวอย่างไม่มีความเอนเอียงในการทดลองหรือการสำรวจเรื่องต่างๆ ทั้งนี้เพราะเลขสุ่มมาจากแนวคิดเกี่ยวกับการคำนวณความน่าจะเป็น

3.1.2 เลขสุ่มจะทำให้ได้มาซึ่งรูปแบบต่างๆหรือวิธีการที่สลับซับซ้อนโดยการสร้างสถานการณ์จำลอง (Simulation)

3.1.3 การใช้เลขสุ่มอาจทำเพื่อศึกษาคุณสมบัติทางทฤษฎีของกระบวนการทางสถิติ ที่มีความสำคัญสำหรับการประมาณค่า ตลอดจนนำไปสู่คำอธิบายเกี่ยวกับอำนาจการทดสอบทางสถิติ

3.1.4 เพื่อหาคำตอบในปัญหาทางคณิตศาสตร์ โดยการพิจารณาจากการแจกแจงความน่าจะเป็นของปัญหานั้นๆ

3.2 แผนการดำเนินการวิจัย

สำหรับการวิจัยครั้งนี้ผู้วิจัยได้กำหนดสถานการณ์ต่างๆเพื่อทำการศึกษาการแก้ไขปัญหาข้อมูลผิดปกติในการวิเคราะห์ข้อมูลที่ได้จากแผนการทดลองแบบสุ่มตลอดเมื่อปัจจัยทดลองเป็นปัจจัยคงที่และทำการเปรียบ เทียบวิธีการแปลงข้อมูลรูปแบบต่างๆเพื่อแก้ไขได้ดังนี้

3.2.1 อิทธิพลของวิธีทดลองในแผนแบบการทดลองสุ่มตลอดเป็นแบบคงที่

3.2.2 จำนวนวิธีทดลองในแผนการทดลองคือ 3, 4 และ 5

3.2.3 จำนวนซ้ำในแต่ละวิธีทดลองคือ 5, 6, 7 และ 8

3.2.4 ค่าเฉลี่ยของข้อมูลตอบสนองในแต่ละวิธีทดลองเท่ากับ 50

3.2.5 กำหนดจำนวนค่าผิดปกติที่ศึกษาเป็น 5% และ 10% ของข้อมูล

3.2.6 กำหนดค่าสัมประสิทธิ์ความแปรผัน(Coeffcient of Variation) ของแต่ละวิธีการทดลองเท่ากับ 30%,40% และ 50%

3.2.7 ความคลาดเคลื่อนมีการแจกแจงแบบปกติและมีความแปรปรวนแต่ละวิธีทดลองเท่ากัน

3.2.8 กำหนดส่วนเบี่ยงเบนมาตรฐานของแต่ละวิธีทดลองเป็นฟังก์ชันกับค่าสัมประสิทธิ์ความแปรผันและค่าเฉลี่ยของวิธีทดลอง กล่าวคือ

$$C.V. = \frac{SD}{\overline{X}} \times 100$$

$$SD = \frac{C.V. \times \overline{X}}{100}$$

จะได้ส่วนเบี่ยงเบนมาตรฐาน คือ 15 20 และ 25 ตามลำดับ

3.2.9 กำหนดความแตกต่างระหว่างอิทธิพลของวิธีทดลองแบ่งออกเป็น 3 ระดับคือ

3.2.9.1 ความแตกต่างระหว่างอิทธิพลของวิธีทดลองมีความแตกต่างกันน้อย ค่า ϕ อยู่ระหว่าง (0,1.5)

3.2.9.2 ความแตกต่างระหว่างอิทธิพลของวิธีทดลองมีความแตกต่างกันปานกลาง ค่า ϕ อยู่ระหว่าง [1.5,3.0)

3.2.9.3 ความแตกต่างระหว่างอิทธิพลของวิธีทดลองมีความแตกต่างกันมาก ค่า ϕ มากกว่า 3.0

3.2.10 ระดับนัยสำคัญของการทดสอบเอฟ (α) ที่ศึกษาคือ 0.01 และ 0.05

3.2.11 กำหนดการกระทำซ้ำในแต่ละสถานการณ์เป็น 500 รอบ

3.3 ขั้นตอนในการดำเนินการวิจัย

ขั้นตอนในการดำเนินการวิจัย แบ่งเป็น 7 ขั้นตอน ดังนี้

3.3.1 สร้างความคลาดเคลื่อนให้มีการแจกแจงตามที่กำหนดในแผนดำเนินการวิจัย

3.3.2 สร้างอิทธิพลของวิธีทดลอง (τ_i) ให้มีความแตกต่างกัน

3.3.3 การสร้างข้อมูลตามตัวแบบจากแผนการทดลองแบบสุ่มตลอด

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

3.3.4 การสร้างข้อมูลตอบสนองให้เป็นค่าผิดปกติ

3.3.5 การแปลงข้อมูลจากข้อมูลที่สร้างขึ้น

3.3.6 การหาค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงและค่าสัดส่วนการปฏิเสธสมมติฐานว่าง และอำนาจการทดสอบของการทดสอบเอฟ

3.3.7 เปรียบเทียบค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลง ค่าสัดส่วนของการปฏิเสธสมมติฐานว่างและอำนาจการทดสอบของการทดสอบเอฟ

ซึ่งรายละเอียดแต่ละขั้นตอนเป็นดังนี้

3.3.1 สร้างความคลาดเคลื่อนให้มีการแจกแจงตามที่กำหนดในแผนดำเนินการวิจัย

ในการวิจัยครั้งนี้ได้ได้ทำการสร้างการแจกแจงของความคลาดเคลื่อนเป็นแบบปกติโดยในโปรแกรม S-PLUS 2000 จะใช้ฟังก์ชัน rnorm(n, μ, sd) ในการสร้างการแจกแจงแบบปกติ โดย n แทนจำนวนซ้ำ μ แทนค่าเฉลี่ย และ sd แทนค่าเบี่ยงเบนมาตรฐาน ซึ่งผู้วิจัยได้เสนอไว้ในภาคผนวก

3.3.2 สร้างอิทธิพลของวิธีทดลอง (τ_i) ให้มีความแตกต่างกัน

โดยการพิจารณา $\sum_{i=1}^{k} \tau_i = 0$ ซึ่งจะกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลองโดยใช้ ϕ เป็นตัวกำหนด โดยกำหนดจาก

$$\phi = \sqrt{\frac{n\sum_{i=1}^{k} \tau_i^2}{k\sigma^2}}$$

ในกรณีที่จำนวนวิธีทดลองเท่ากับ 3 สามารถกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลองได้สะดวกขึ้น โดยกำหนดให้

$$D = \tau_{max} - \tau_{min}$$

$$\tau_i = \frac{(\tau_{max} + \tau_{min})}{2} \quad ; i = 1, 2, \ldots, k$$

โดยที่ $\tau_{max} = \frac{D}{2}, \tau_{min} = -\frac{D}{2}$ และ $\tau_i = 0$ เมื่อ i ไม่ใช่ค่า max และ min

ในที่นี้ τ_{max} หมายถึง ค่าที่มากที่สุดของอิทธิพลวิธีทดลอง

τ_{min} หมายถึง ค่าที่น้อยที่สุดของวิทธิพลวิธีทดลอง

D หมายถึง ค่าความแตกต่างระหว่างค่าที่มากที่สุดและค่าที่น้อยที่สุดของอิทธิพลวิธีทดลอง

ดังนั้นในการกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลอง โดยใช้ ϕ เป็นตัวกำหนด จะทำได้ดังนี้

$$\phi = D\sqrt{\frac{n}{2k\sigma^2}}$$

ในกรณีที่จำนวนวิธีทดลองเท่ากับ 4 และ 5 สามารถกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลองได้สะดวกขึ้น โดยกำหนดให้

$$D = 2(\tau_{max} - \tau_{min})$$

$$\tau_i = \frac{(\tau_{max} + \tau_{min})}{2} \quad ; i = 1, 2, \ldots, k$$

โดยที่ $\tau_{max} = \frac{D}{4}, \tau_{min} = -\frac{D}{4}$ และ $\tau_i = 0$ เมื่อ i ไม่ใช่ค่า max และ min

ดังนั้นในการกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลอง โดยใช้ ϕ เป็นตัวกำหนด จะทำได้ดังนี้

$$\phi = \frac{D}{2}\sqrt{\frac{n}{k\sigma^2}}$$

ในกรณีที่จำนวนวิธีทดลองเท่ากับ 6 สามารถกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลองได้สะดวกขึ้น โดยกำหนดให้

$$D = 3(\tau_{max} - \tau_{min})$$

$$\tau_i = \frac{(\tau_{max} + \tau_{min})}{2} \quad ; i = 1, 2, \ldots, k$$

โดยที่ $\tau_{max} = \frac{D}{6}, \tau_{min} = -\frac{D}{6}$ และ $\tau_i = 0$ เมื่อ i ไม่ใช่ค่า max และ min

ดังนั้นในการกำหนดกลุ่มความแตกต่างระหว่างอิทธิพลของวิธีทดลอง โดยใช้ ϕ เป็นตัวกำหนด จะทำได้ดังนี้

$$\phi = D\sqrt{\frac{n}{6k\sigma^2}}$$

โดยจะแสดงตัวอย่างของการสร้างอิทธิพลของวิธีทดลองไว้ในภาคผนวก ก

3.3.3 การสร้างข้อมูลตามตัวแบบจากแผนการทดลองแบบสุ่มตลอด

สร้างตัวแปรสุ่มของความคลาดเคลื่อน ε_{ij} ที่มีการแจกแจงแบบปกติ โดยมีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น σ^2 ขึ้นมาก่อนแล้วจึงสร้างค่า y_{ij} ตามตัวแบบดังนี้คือ

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

โดย τ_i เป็นอิทธิพลของวิธีทดลองที่กำหนดขึ้น

3.3.4 การสร้างข้อมูลตอบสนองให้เป็นค่าผิดปกติ

สร้างข้อมูลให้เป็นค่าผิดปกติโดยทำให้ค่าข้อมูลเหล่านั้นเป็นค่าสูงสุดหรือค่าต่ำสุด โดยจะใช้เกณฑ์ของ Box plot กล่าวคือ

ค่าของข้อมูลตอบสนองที่มีค่าสูงกว่า $Q_3 + 1.5IQR$ และ ค่าของข้อมูลตอบสนองที่มีค่าต่ำกว่า $Q_1 - 1.5IQR$ จะถือว่าข้อมูลตอบสนองตัวนั้นเป็นค่าผิดปกติโดยที่ IQR คือ ค่าพิสัยควอไทล์ (Inter Quartile Range) มีค่าเท่ากับ $Q_3 - Q_1$

3.3.5 การแปลงข้อมูลจากข้อมูลที่สร้างขึ้น

กำหนดจำนวนวิธีทดลอง จำนวนซ้ำในแต่ละวิธีทดลอง ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน แล้วทำการสร้างชุดข้อมูลสุ่มโดยโปรแกรมที่แสดงในภาคผนวกตามลักษณะการแจกแจงของความคลาดเคลื่อนแบบปกติและให้มีข้อมูลที่เป็นค่าผิดปกติ และนำข้อมูลที่ได้มาทำการแปลงข้อมูลโดยใช้พารามิเตอร์ยกกำลังตามหลักการของ Box และ Cox ซึ่งรายละเอียดเกี่ยวกับการแปลงข้อมูลได้กล่าวไว้แล้วในบทที่2

3.3.6 การหาค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงและค่าสัดส่วนการปฏิเสธสมมติฐานว่าง และอำนาจการทดสอบของการทดสอบเอฟ

เมื่อสร้างข้อมูลตามตัวแบบที่ต้องการศึกษาและทำการแปลงข้อมูลด้วยเลขยกกำลังค่าต่างๆแล้วจะทำการคำนวณค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงและคำนวณค่า p-value ของการทดสอบเอฟ และเปรียบเทียบค่า p-value กับระดับนัยสำคัญที่กำหนด ขั้นตอนต่อไปคือ การหาค่าสัดส่วนของการปฏิเสธสมมติฐานว่างและอำนาจการทดสอบของการทดสอบเอฟ ซึ่งสรุปขั้นตอนดังนี้

3.3.6.1 สร้างอิทธิพลของวิธีทดลองโดยการกำหนดค่า τ_i ให้มีค่าเป็น 0 ทุกค่าในแต่ละวิธีทดลอง เมื่อพิจารณาหาค่าสัดส่วนของการปฏิเสธสมมติฐานว่างของการทดสอบเอฟ และให้ τ_i มีค่าไม่เท่ากับ 0 ในบางค่า โดยให้ $\sum_{i=1}^{k}\tau_i = 0$ เมื่อพิจารณาหาค่าอำนาจการทดสอบ

3.3.6.2 คำนวณค่าสัดส่วนของการปฏิเสธสมมติฐานว่างของการทดสอบเอฟเมื่อ $\tau_i = 0$ และคำนวณค่าอำนาจการทดสอบเมื่อ τ_i มีค่าไม่เท่ากับ 0 ในบางค่า

3.3.6.3 เปลี่ยนค่าส่วนเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน จนกระทั่งครบทุกสถานการณ์ โดยในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

3.3.7 เปรียบเทียบค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลง ค่าสัดส่วนของการปฏิเสธสมมติฐานว่างและอำนาจการทดสอบของการทดสอบเอฟ

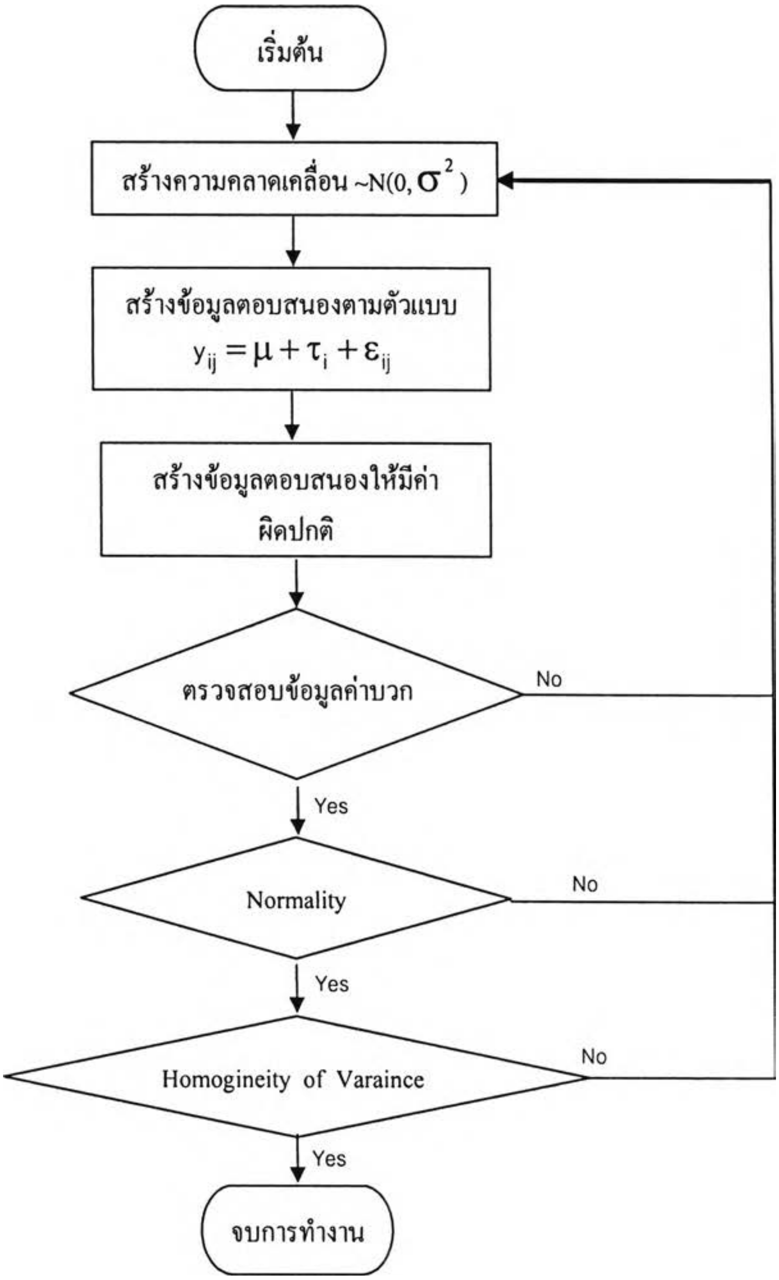
เปรียบเทียบค่าสัดส่วนของจำนวนค่าผิดปกติที่ลดลงเป็นหลักว่าวิธีการแปลงข้อมูลด้วยเลขยกกำลังค่าใดสามารถแก้ไขปัญหาได้มากที่สุด โดยจะพิจารณาร่วมกับสัดส่วนของการปฏิเสธสมมติฐานว่างที่น้อยและอำนาจการทดสอบของการทดสอบเอฟที่สูง ค่านั้นจะเป็นค่าที่เหมาะสมที่สุดในการแก้ไขปัญหาข้อมูลที่มีค่าผิดปกติ

3.4 แผนผังแสดงการทำงานของโปรแกรม

ลักษณะการทำงานของโปรแกรมในการวิจัยครั้งนี้ ใช้ S-PLUS 2000 ในการประมวลผลข้อมูลโดยมีขั้นตอนการทำงานดังรูปที่ 3.4.1-3.4.2 ส่วนโปรแกรมการทำงานตามลำดับขั้นดังแสดงในรูปที่ 3.4.1-3.4.2 นั้นได้เสนอไว้ในภาคผนวก ค

หอสมุดกลาง ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 3.1 แสดงผังงานในการสร้างข้อมูลตามสถานการณ์ที่ศึกษา

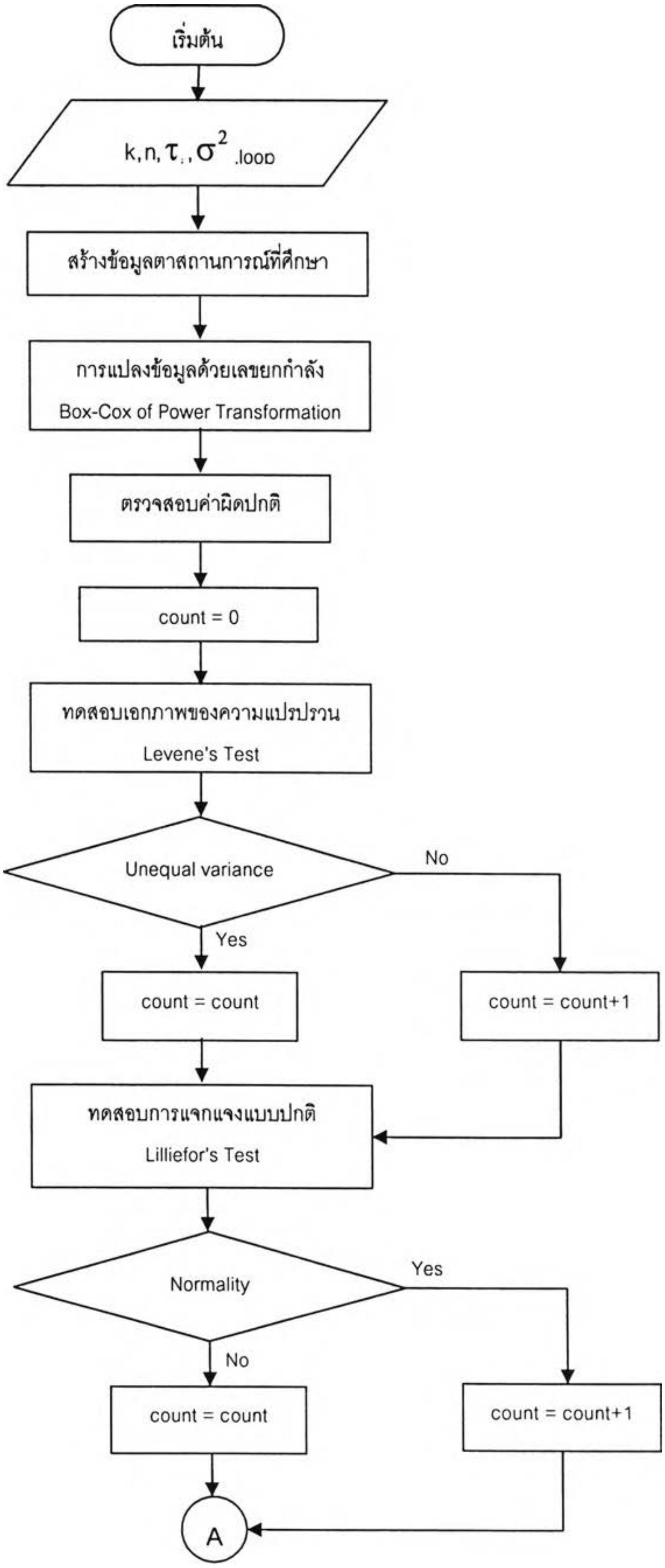


<u>หมายเหตุ</u>

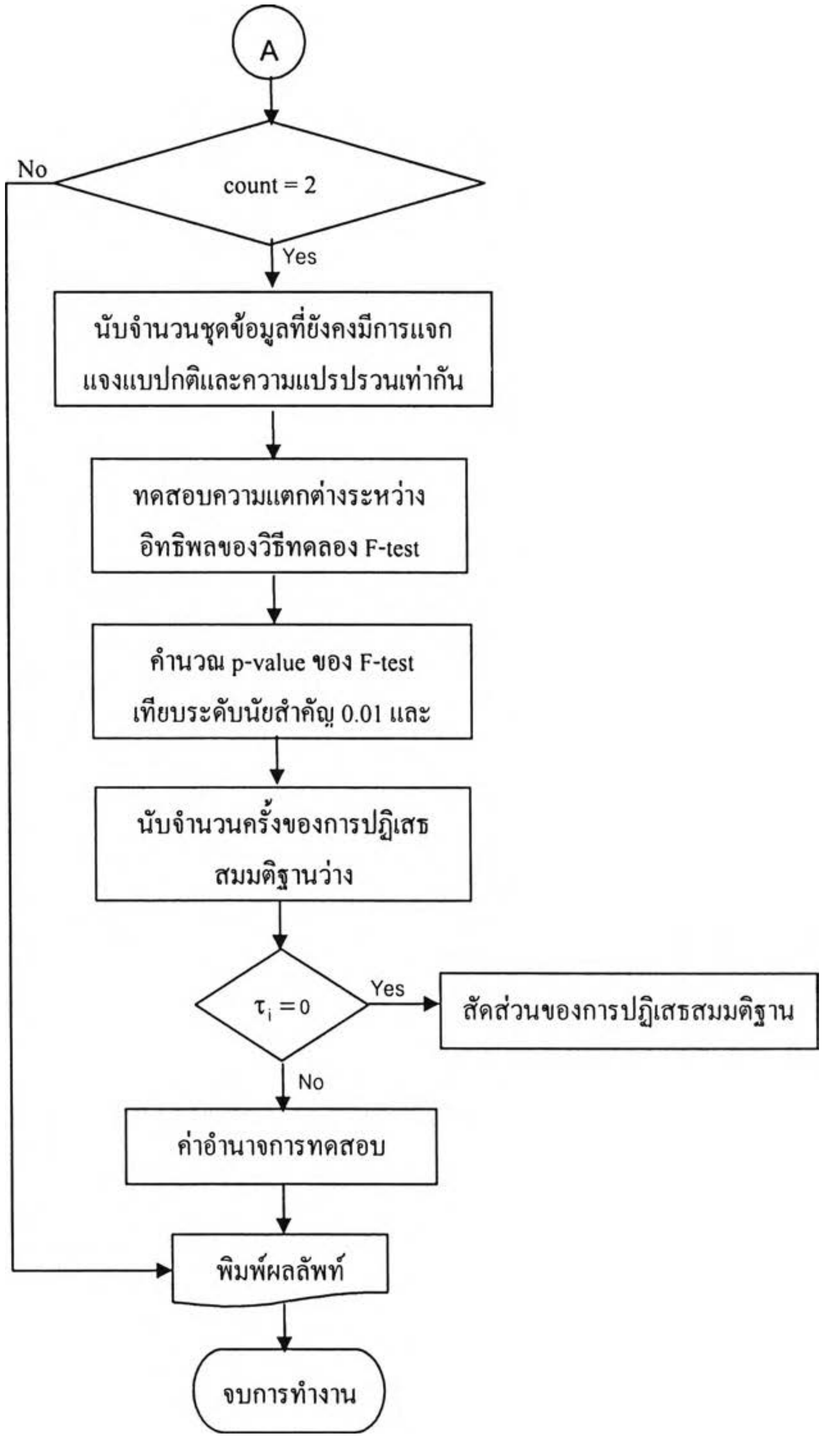
Normality แทน "ข้อมูลตอบสนองมีการแจกแจงแบบปกติหรือไม่"

Homogineity of variance แทน "ข้อมูลตอบสนองมีความแปรปรวนเท่ากันหรือไม่"

รูปที่ 3.2 แสดงผังงานในการทำงานของโปรแกรมการแปลงข้อมูล



รูปที่ 3.2 แสดงผังงานในการทำงานของโปรแกรมการแปลงข้อมูล (ต่อ)



<u>หมายเหตุ</u>

count แทน ตัวแปรการนับจำนวนชุดข้อมูลที่ยังคงมีการแจกแจงแบบปกติและความแปรปรวนเท่ากันภายหลังการแปลงข้อมูล