

บทที่ 1

บทนำ



1.1 ความเป็นมาและความสำคัญของปัญหา

ในการทดลองต่างมีวัตถุประสงค์ที่สำคัญข้อหนึ่ง คือ ต้องการข้อสรุปที่มีประสิทธิภาพ มีความถูกต้อง และน่าเชื่อถือได้ ซึ่งการจะได้ข้อสรุปของการทดลองที่ได้นั้นผู้ทดลองจะต้องเลือกแผนแบบการทดลองให้เหมาะสม กำหนดระดับปัจจัยการทดลอง ควบคุมตัวแปรรบกวน หน่วยทดลอง และเงื่อนไขต่างๆ ของการทดลองให้เหมาะสมกับสิ่งที่ต้องการจะศึกษา เพื่อจะได้นำมาวิเคราะห์และสามารถสรุปผลเชิงเหตุและผลได้ ซึ่งในการวิเคราะห์ข้อมูลทางสถิติที่ใช้ในการทดลองเพื่อศึกษาความแตกต่างของวิธีการทดลอง (Treatment) ที่ให้กับหน่วยทดลอง (Experimental Unit) ว่ามีอิทธิพลแตกต่างกันหรือไม่นั้น จะต้องคำนึงถึงเงื่อนไขหรือข้อสมมติ (Assumptions) ในการวิเคราะห์ด้วย ในที่นี้จะมีข้อสมมติเกี่ยวกับค่าความคลาดเคลื่อน คือ ค่าความคลาดเคลื่อนจะต้องมีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเท่ากับศูนย์ ค่าความแปรปรวนของแต่ละวิธีทดลองเท่ากัน และค่าความคลาดเคลื่อนแต่ละค่าจะต้องเป็นอิสระต่อกัน

ในทางปฏิบัติ ข้อมูลที่ได้จากการทดลองโดยส่วนใหญ่แล้วจะมีจำนวนน้อยเมื่อเทียบกับข้อมูลที่ไม่ได้เกิดจากการทำการทดลอง เช่น จากการทำสำรวจตัวอย่าง อาจเป็นผลให้ข้อมูลที่ได้จากการทำการทดลองนั้นเป็นข้อมูลที่ตรงตามข้อสมมติดังกล่าวได้ยาก และหากข้อมูลไม่เป็นไปตามข้อสมมติดังกล่าวแล้ว อาจทำให้การวิเคราะห์และสรุปผลเกิดความผิดพลาดและมีความน่าเชื่อถือน้อยลงได้ ซึ่งมีวิธีการแก้ปัญหาเพื่อให้ได้ข้อมูลที่เป็นไปตามข้อสมมติในการวิเคราะห์ที่แตกต่างกันในแต่ละข้อสมมติ เช่น การแก้ปัญหาเมื่อข้อมูลที่ได้ไม่ได้มีการแจกแจงแบบปกติโดยการแปลงข้อมูล การแก้ปัญหาเมื่อข้อมูลมีค่าความแปรปรวนไม่เท่ากันโดยการแปลงข้อมูลหรือการถ่วงน้ำหนัก (Weighted) เป็นต้น ซึ่งการแปลงข้อมูลนั้นมีด้วยกันหลายรูปแบบและจะใช้รูปแบบใดนั้นขึ้นอยู่กับรูปแบบของข้อมูล

เพราะฉะนั้นผู้วิจัยจึงสนใจศึกษาหารูปแบบการแปลงข้อมูลที่ใช้เพื่อแก้ไขปัญหาข้อมูลตอบสนองที่ได้จากการทดลองที่ไม่มีการแจกแจงแบบปกติให้มีการแจกแจงแบบปกติและยังคงรักษาข้อสมมติอื่นๆในการวิเคราะห์ข้อมูลทางการทดลองไว้

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาการแก้ไขปัญหาคือข้อมูลตอบสนองที่ได้จากการทดลองที่ไม่มีการแจกแจงแบบปกติ สำหรับการวิเคราะห์ความแปรปรวน โดยการเปรียบเทียบรูปแบบการแปลงข้อมูลที่ใช้ในการแปลงต่างๆ

1.3 ข้อตกลงเบื้องต้น

1.3.1 ตัวแบบเชิงสถิติ (Statistical Model)

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad ; i = 1, 2, \dots, k ; j = 1, 2, \dots, n$$

เมื่อ Y_{ij} = ข้อมูลตอบสนองที่ได้จากการทดลอง จากวิธีทดลองที่ i หน่วยทดลองที่ j

μ = พารามิเตอร์ค่าเฉลี่ยรวม

τ_i = พารามิเตอร์ผลกระทบจากวิธีทดลองที่ i โดยที่ $\sum_{i=1}^k \tau_i = 0$

ε_{ij} = ความคลาดเคลื่อนสุ่มของการทดลองจากวิธีทดลองที่ i หน่วยทดลองที่ j

k = จำนวนวิธีทดลอง

n = จำนวนหน่วยทดลองในแต่ละวิธีทดลอง

1.3.2 การแปลงข้อมูลมีจุดประสงค์เพื่อต้องการแก้ไขปัญหากับข้อมูลตอบสนองที่ไม่มีการแจกแจงแบบปกติให้มีการแจกแจงแบบปกติ ซึ่งข้อมูลที่ได้หลังจากการแปลงแล้วจะต้องสามารถนำไปใช้ในการวิเคราะห์ต่อได้ และข้อมูลจะต้องเป็นไปตามข้อสมมติของการวิเคราะห์ความแปรปรวน คือ

- 1) ความคลาดเคลื่อนสุ่มต้องมีความเป็นอิสระกัน ($\varepsilon_{ij}, \varepsilon_{i'j'}$ เป็นอิสระจากกัน เมื่อ $i \neq i' = 1, 2, \dots, k \quad j \neq j' = 1, 2, \dots, n$)
- 2) ความคลาดเคลื่อนสุ่มต้องมีการแจกแจงแบบปกติ ($\varepsilon_{ij} \sim N(0, \sigma^2)$) จากข้อกำหนดนี้ทำให้ $y_{ij} ; i = 1, 2, \dots, k ; j = 1, 2, \dots, n$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติด้วย
- 3) ความแปรปรวนของความคลาดเคลื่อนในแต่ละกลุ่มต้องเท่ากัน (Homogeneity of Variances) ($\sigma_i^2 = \sigma^2$)

1.4 ขอบเขตของการวิจัย

- 1.4.1 ตัวแบบที่ใช้ศึกษาในการวิจัยครั้งนี้ เป็นตัวแบบเชิงสถิติของข้อมูลตอบสนองจากแผนแบบการทดลองสุ่มตลอดเมื่อปัจจัยทดลองเป็นปัจจัยคงที่ (Completely Randomized Design with Fixed Treatment Factor)
- 1.4.2 การแก้ไขปัญหาข้อมูลตอบสนองที่ได้จากการทดลองที่ไม่มีการแจกแจงแบบปกติ จะใช้วิธีการแปลงข้อมูล ด้วยวิธี

Box-Cox Transformation

$$y' = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log y, & \lambda = 0 \end{cases}$$

เมื่อ $-2 \leq \lambda \leq 2$

- 1.4.3 ลักษณะข้อมูลที่ศึกษา พิจารณาจากการแจกแจงแบบเบ้โดยใช้เกณฑ์ของความเบ้และความโด่ง ซึ่งลักษณะของข้อมูลมีดังนี้
- มีการแจกแจงแบบเบ้ขวา (Positive-Skewed Distribution) ซึ่งกำหนดค่าพารามิเตอร์ต่างๆ คือ
 - ความเบ้ มีค่าตั้งแต่ 0.7 ถึง 2.0
 - ความโด่ง มีค่าตั้งแต่ 3.0 ถึง 11.0
 - มีการแจกแจงแบบเบ้ซ้าย (Negative-Skewed Distribution) ซึ่งกำหนดค่าพารามิเตอร์ต่างๆ คือ
 - ความเบ้ มีค่าตั้งแต่ -2.0 ถึง -0.7
 - ความโด่ง มีค่าตั้งแต่ 3.0 ถึง 11.0
 - จำนวนวิธีการทดลอง (Treatment) ที่ศึกษา คือ 3 4 และ 5 วิธีการทดลอง
 - กำหนดจำนวนหน่วยทดลองในแต่ละวิธีการทดลองให้มีจำนวนที่เท่ากัน ($n_i = n$) โดยกำหนดให้มีจำนวนหน่วยทดลองเท่ากับ 4 5 และ 6
 - กำหนดให้ข้อมูลมีค่าสัมประสิทธิ์ความแปรผัน (Coefficient of Variation: C.V.(%)) เท่ากับ 20% 40% และ 60%
 - กำหนดให้ข้อมูลมีค่าเฉลี่ย (μ) เท่ากับ 50 ซึ่งจะทำให้ได้ค่าส่วนเบี่ยงเบนมาตรฐาน (σ) เท่ากับ 10 20 และ 30 และกำหนดให้ค่าความแปรปรวนของแต่ละวิธีการทดลองมีค่าเท่ากัน ($\sigma_i^2 = \sigma^2$)

- ในการหาอำนาจการทดสอบกำหนดให้อัทธิพลของวิธีทดลอง (τ_i) แตกต่าง

กัน โดยพิจารณา
$$\sum_{i=1}^k \tau_i = 0$$

และใช้ Φ เป็นตัวกำหนด โดยที่

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

(Φ แทน สัมประสิทธิ์ความเบี่ยงเบนของวิธีทดลอง)

- ระดับนัยสำคัญ (α) ของการทดสอบในครั้งนี้ กำหนดที่ระดับ 0.05
- การศึกษาครั้งนี้ ทำการจำลองข้อมูล (Simulation) ด้วยวิธีมอนติคาร์โล (Monte Carlo Method)
- การทดสอบการแจกแจงของความคลาดเคลื่อนของข้อมูล ใช้วิธีของวิกและแชปปีโร (Shapiro-Wilk)
- การทดสอบความเท่ากันของความแปรปรวนของข้อมูลจากการทดลองทุกวิธีทดลอง (Homogeneity of Variance) ใช้วิธีตรวจสอบของเลอวี้น (Levene's Test for Equality of Variances)

1.4.4 การจำลองกระทำซ้ำๆ กัน 500 ครั้งในแต่ละสถานการณ์ของการทดลอง

1.5 วิธีดำเนินการวิจัย

- 1.5.1 จำลองข้อมูลตอบสนองที่ได้จากการทดลอง ด้วยวิธีมอนติคาร์โล (Monte Carlo Method)
- 1.5.2 ทำการตรวจสอบข้อมูลที่ได้จากการจำลองว่ามีลักษณะข้อมูลตรงตามกรณีศึกษาหรือไม่ โดยใช้โมเมนต์ของค่าเฉลี่ยทำการทดสอบความคลาดเคลื่อนของข้อมูลว่ามีความเบ้, ความโด่งตรงตามที่กำหนด
- 1.5.3 ทำการแก้ไขปัญหาข้อมูลที่ไม่มีการแจกแจงปกติโดยการแปลงข้อมูลด้วยวิธีต่างๆ
- 1.5.4 ทดสอบการแจกแจงของความคลาดเคลื่อนของข้อมูลหลังจากผ่านการแก้ไขแล้ว ด้วยวิธีของวิกและแชปปีโร (Shapiro-Wilk)
- 1.5.5 ทำการวิเคราะห์ความแปรปรวนหาค่าอำนาจของการทดสอบของข้อมูลตอบสนองที่ได้จากการแปลงข้อมูลให้มีการแจกแจงแบบปกติ
- 1.5.6 ทำการทดลองซ้ำ 500 ครั้งในแต่ละสถานการณ์

1.6 เกณฑที่ใชในการตัดสินใจ

พิจารณาจากค่าสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลตอบสนองที่ไม่มีการแจกแจงแบบปกติ ค่าสัดส่วนของความสำเร็จในการแก้ปัญหาข้อมูลตอบสนองที่ไม่มีการแจกแจงแบบปกติและมีความเท่ากันของความแปรปรวนภายหลังการแปลงข้อมูล ค่าสัดส่วนของการปฏิเสธสมมติฐานว่างและอำนาจการทดสอบของการทดสอบเอฟ

1.7 คำจำกัดความต่างๆ ที่ใช้ในการวิจัย

วิธีทดลอง (Treatment) หมายถึง ระดับต่างๆของปัจจัยทดลองที่กำหนดสำหรับการทดลองที่ศึกษา

ตัวแปรตอบสนอง (Response Variable (y)) หมายถึง ตัวแปรของข้อมูลที่วัดได้หลังจากการหน่วยทดลองได้รับผลกระทบจากวิธีทดลองใดๆ

ความผิดพลาดประเภทที่ 1 (Type I Error) คือ ความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานว่าง (Null hypothesis (H_0)) เมื่อสมมติฐานว่างนั้นเป็นจริง

ความผิดพลาดประเภทที่ 2 (Type II Error) คือ ความผิดพลาดที่เกิดจากการยอมรับสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นไม่จริง

อำนาจการทดสอบ (Power of the Test) คือ ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างนั้นไม่จริง

1.8 ประโยชน์ที่คาดว่าจะได้รับ

เพื่อเป็นแนวทางในการตัดสินใจเลือกวิธีการแก้ปัญหาข้อมูลตอบสนองในแผนแบบการทดลองสุ่มตลอดที่ไม่มีการแจกแจงแบบปกติ โดยการแปลงข้อมูลด้วยรูปแบบต่างๆ ที่เหมาะสมกับลักษณะของข้อมูลที่ได้จากการทดลอง เมื่อต้องการทำให้ข้อมูลมีการแจกแจงแบบปกติ และยังเป็นไปตามข้อสมมติของการวิเคราะห์ความแปรปรวน เพื่อที่จะสามารถนำข้อมูลที่ได้นี้ไปใช้ในการวิเคราะห์ต่อได้