

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ผู้วิจัยทำการเปรียบเทียบค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลาง โดยทำการศึกษาแผนแบบการทดลองแบบแฟกทอเรียลในกรณีที่ปัจจัย ในแผนแบบการทดลองเป็นตัวแปรเชิงปริมาณ โดยทำการศึกษาจากแผนแบบการทดลองแบบแฟกทอเรียลที่มีการทำซ้ำที่จุดศูนย์กลาง ซึ่งมีขนาด 2^2 โดยศึกษากรณีที่ความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็นรูปแบบกำลังหนึ่ง หรือในกรณีที่ไม่มีความสัมพันธ์ระหว่างปัจจัย และกรณีที่ความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็นรูปแบบกำลังสอง โดยหาข้อสรุปของการศึกษาโดยใช้การจำลองข้อมูลด้วยวิธีมอนติคาร์โล (Monte Carlo Method) ผู้วิจัยขอแนะนำเสนอขั้นตอนเป็นลำดับดังนี้

การจำลองข้อมูลด้วยวิธีมอนติคาร์โล (Simulation by Monte Carlo Method)

วิธีมอนติคาร์โลเป็นเทคนิคในการจำลองตัวแบบทางคณิตศาสตร์ ซึ่งหลักการเบื้องต้นนั้น ต้องจำลองตัวเลขสุ่ม (Random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา โดยในการวิจัยนี้ใช้เทคนิคมอนติคาร์โลในการสร้างข้อมูลที่มีสภาพการแจกแจงตามที่ต้องการศึกษา ขั้นตอนสำคัญของการจำลองข้อมูลด้วยวิธีมอนติคาร์โลมี 3 ขั้นตอนคือ

ขั้นตอนที่ 1 การสร้างเลขสุ่ม

การสร้างเลขสุ่มเป็นสิ่งที่สำคัญมากในวิธีมอนติคาร์โล ทั้งนี้ก็เพราะว่าหลักการของมอนติคาร์โลนั้นจะใช้เลขสุ่มมาช่วยในการหาคำตอบของปัญหา ลักษณะของเลขสุ่มที่ดีจะมีการแจกแจงสม่ำเสมอ (Uniform Distribution) ในช่วง $[0,1]$ และเป็นอิสระซึ่งกันและกัน

ขั้นตอนที่ 2 การประยุกต์ของปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม

ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจใช้เลขสุ่มโดยตรง แต่บางปัญหาอาจใช้ตัวเลขสุ่มเพียงบางขั้นตอนของปัญหาเท่านั้น

ขั้นตอนที่ 3 การทดลองกระทำ

เมื่อประยุกต์ปัญหาที่ใช้กับเลขสุ่มได้แล้วขั้นตอนต่อไปก็คือการทดลอง โดยใช้กระบวนการสุ่ม (random process) มากระทำในลักษณะซ้ำ ๆ กัน (replication) เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

แผนการทดลอง

ผู้วิจัยกำหนดสถานการณ์ในการหาค่าเฉลี่ยของพารามิเตอร์ไว้ศูนย์กลางของแผนแบบการทดลองทั้ง 2 วิธีดังนี้

1. เลือกตัวอย่างอย่างสุ่มจากประชากรที่มีการแจกแจงเดียวกัน การแจกแจงที่สนใจศึกษา คือ การแจกแจงแบบปกติที่มีค่าเฉลี่ย 0, 5, 10 และ 20 และส่วนเบี่ยงเบนมาตรฐาน 5, 10, 15, 20, 25 และ 30

2. กำหนดขนาดตัวอย่างที่ศึกษา คือ 7, 8, 9 และ 10 ซึ่งเป็นแผนแบบการทดลองแบบแฟกทอเรียลในกรณีที่มีการทำซ้ำที่จุดศูนย์กลาง โดยในที่นี้การทำซ้ำที่จุดศูนย์กลางเท่ากับ 3, 4, 5 และ 6

3. ปัจจัยหลัก (main effect) ที่ใช้ในการศึกษาครั้งนี้แบ่งออกดังนี้

3.1 แผนแบบการทดลองแบบแฟกทอเรียลขนาด 2^2 กรณีความสัมพันธ์ระหว่าง ตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็นรูปแบบกำลังหนึ่ง (กรณีไม่มีผลกระทบร่วมระหว่างปัจจัย) ทำให้มีปัจจัยหลัก 2 ปัจจัย และแต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

3.2 แผนแบบการทดลองแบบแฟกทอเรียลขนาด 2^2 กรณีความสัมพันธ์ระหว่าง ตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็นรูปแบบกำลังสอง ซึ่งแบ่งออกเป็น 7 รูปแบบ ดังนี้

1. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็นรูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 3 ปัจจัย และซึ่งแต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

2. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็นรูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 3 ปัจจัย และแต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

3. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็น
รูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{22} x_2^2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 3 ปัจจัย และแต่ละ
ปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

4. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็น
รูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 4 ปัจจัย และ
แต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

5. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็น
รูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{22} x_2^2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 4 ปัจจัย และ
แต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

6. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็น
รูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 4 ปัจจัย และ
แต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

7. รูปแบบของความสัมพันธ์ของตัวแปรตามกับตัวแปรในแผนแบบการทดลองเป็น
รูปแบบ $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon_i$ ซึ่งทำให้มีปัจจัยหลัก 5
ปัจจัย และแต่ละปัจจัยสร้างจากการแจกแจงปกติค่าเฉลี่ย 10 ความแปรปรวน 10

ขั้นตอนในการศึกษาวิจัย

ขั้นตอนในการศึกษาวิจัยมีดังนี้คือ

1. กำหนดลักษณะของการแจกแจงความผิดพลาด ขนาดตัวอย่าง (จำนวนการทำซ้ำที่
จุดศูนย์กลาง)

2. การคำนวณบริเวณที่ใช้ในแผนแบบการทดลองขั้นที่สองของแต่ละวิธี

2.1 แผนแบบการทดลองแบบแฟกทอเรียล (Factorial experiments)

2.2 แผนแบบการทดลองแบบการหมุน (Rotated Design)

3. การสร้างข้อมูลของตัวแปรตาม (y) จากตัวแปรในแผนแบบการทดลอง (X) และความ
คลาดเคลื่อนที่มีลักษณะการแจกแจงตามที่ต้องการศึกษา (ε) โดยให้ตัวแปรตามมีความ
สัมพันธ์เชิงเส้นในพารามิเตอร์กับตัวแปรในแผนแบบการทดลอง โดยพารามิเตอร์คำนวณจาก
ปัจจัยหลัก (main effect) ของแต่ละปัจจัยหารด้วยสอง

4. การหาสัมประสิทธิ์การถดถอยของแต่ละวิธี

5. การหาความคลาดเคลื่อนกำลังสองเฉลี่ยของแต่ละวิธี

6. การหาค่าเฉลี่ยของพารามิเตอร์ไว้ศูนย์กลางแต่ละวิธี
7. สรุปผลในรูปตาราง

รายละเอียดของแต่ละขั้นตอนเป็นดังนี้

ขั้นตอนที่ 1

การกำหนดลักษณะของการแจกแจงของความคลาดเคลื่อน ขนาดตัวอย่าง กำหนดตามแผนการทดลองที่ได้กำหนดไว้แล้วข้างต้น

ขั้นตอนที่ 2

การกำหนดบริเวณที่สนใจของแผนแบบการทดลองทั้งสองวิธีดังนี้

1. แผนแบบการทดลองแบบแฟคทอเรียล กำหนดเมทริกซ์ของแผนแบบการทดลอง

ดังนี้

$$D = \begin{bmatrix} \pm g & \pm g & \dots & \pm g \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

โดยที่ $g^2 = n / [(d_1 + d_2) * (n - n_0)]$

หรือ $g^2 = n / 3 * (n - n_0)$

ซึ่ง $g \leq 1$

เมื่อ n คือจำนวนตัวอย่างทั้งหมด

n_0 คือจำนวนตัวอย่าง (การทำซ้ำ) ที่จุดศูนย์กลาง (center point)

d_1 คือจำนวนของกำลังที่ใช้ในการประมาณ ในที่นี้เท่ากับ 1

และ d_2 เท่ากับ $d_1 + 1$ ซึ่งในที่นี้เท่ากับ 2

2. แผนแบบการทดลองแบบการหมุน กำหนดเมทริกซ์ของแผนแบบการทดลอง

ดังนี้

$$D' = \begin{bmatrix} -d & d & -1 & 1 & 0 \\ -1 & 1 & d & -d & 0 \end{bmatrix}$$

โดยที่ $d^2 = \sqrt{(2g^2) - 1}$

ซึ่ง $d \leq 1$

ขั้นตอนที่ 3

การสร้างข้อมูลที่มีลักษณะการแจกแจงที่ต้องการศึกษาจะใช้โปรแกรมภาษาฟอร์แทรน 77 (Fortran 77) บนเครื่อง AMDAHL 5860 ซึ่งการสร้างข้อมูลที่มีลักษณะการแจกแจงแบบต่าง ๆ นั้นจะต้องใช้ตัวเลขสุ่มเป็นพื้นฐานในการสร้าง ฟังก์ชันที่ใช้ในการสร้างเลขสุ่มคือ RANDOM(IX) เมื่อสร้างเลขสุ่มแล้วจะนำตัวเลขสุ่มที่ได้มาสร้างข้อมูลที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ ความแปรปรวน σ^2 โดยจะได้จากโปรแกรมย่อย NORM (MEANS,SIGMA,X) ซึ่งรายละเอียดโปรแกรมย่อยแสดงดังนี้

```

SUBROUTINE NORM (MEANS,SIGMA,X)
REAL MEANS
COMMOM/SEED/ISEED, KK
PI=3.1415926
IF (KK.EQ.1) GOTO 10
R1 = RANDOM(ISEED)
R2 = RANDOM(ISEED)
Z1 = SQRT(-2*ALOG(R1))*COS(2*PI*R2)
Z2 = SQRT(-2*ALOG(R1))*SIN(2*PI*R2)
X = Z1*SIGMA+MEANS
KK = 1
RETURN
10 X = Z2*SIGMA+MEANS
KK = 0
RETURN
END

```

ข้อมูลที่สร้างได้แก่ ตัวแปรในแผนแบบการทดลองกำหนดให้มีค่า -1 0 และ 1 ปัจจัยหลักแต่ละปัจจัยสร้างจากการแจกแจงแบบปกติที่มีค่าเฉลี่ย 10 ความแปรปรวน 100 ความคลาดเคลื่อน (ϵ) สร้างตามขอบเขตที่กำหนด จากนั้นจะสร้างตัวแปรตาม (y) ให้มีความสัมพันธ์เชิงเส้นตรงในพารามิเตอร์กับตัวแปรในแผนแบบการทดลอง โดยมีลักษณะการแจกแจงความคลาดเคลื่อนที่กำหนด ซึ่งมีรูปแบบดังนี้

$$y = Z*(X) \beta + \epsilon$$

เมื่อ $Z^*(X)$ เป็นเมทริกซ์ตัวแปรในแผนแบบการทดลองที่อยู่ในรูปของ X
 β เป็นค่าพารามิเตอร์ที่กำหนด ซึ่งได้จากปัจจัยหลักหารด้วยสอง
 และ ε เป็นความคลาดเคลื่อนซึ่ง $\varepsilon \sim N(\mu, \sigma^2)$

ขั้นตอนที่ 4

การหาสัมประสิทธิ์การถดถอย ซึ่งจะคำนวณจากโปรแกรมย่อย BETA(N,KE1, XE,G,Y,B) โดยทั้งสองวิธีจะต้องใช้ตัวแปรทุกตัวให้อยู่ในสมการ เนื่องจากเมทริกซ์ของ แผนแบบการทดลองทั้งสองวิธีมีลำดับชั้นไม่เต็ม จึงต้องนำตัวผกผันทั่วไป (G) มาใช้ในการ หาสัมประสิทธิ์การถดถอย โดยตัวผกผันทั่วไปที่ใช้ในการวิจัยครั้งนี้ ได้มาจากการคำนวณ จากโปรแกรมสำเร็จรูป SAS

โดยที่

- N คือจำนวนตัวอย่าง
- KE1 คือจำนวนตัวแปรในแผนแบบการทดลอง
- XE คือตัวแปรในแผนแบบการทดลอง
- G คือตัวผกผันทั่วไปของ XE
- Y คือตัวแปรตาม
- B คือสัมประสิทธิ์การถดถอย

โดยขั้นตอนการทำงานของโปรแกรมย่อย BETA(N,KE1, XE,G,Y,B) คือเป็น โปรแกรมที่ใช้สำหรับหาสัมประสิทธิ์การถดถอย โดยหาได้จากสูตรดังนี้

$$\beta^0 = GX' y$$

ขั้นตอนที่ 5

การหาความคลาดเคลื่อนกำลังสองเฉลี่ยของแต่ละวิธี ซึ่งคำนวณจากโปรแกรมย่อย MSQE(N,KE1, XE,Y,B,SSE,MSE, YP,RES)

โดยที่

- N คือจำนวนตัวอย่าง
- KE1 คือจำนวนตัวแปรในแผนแบบการทดลอง
- XE คือตัวแปรในแผนแบบการทดลอง
- Y คือตัวแปรตาม
- B คือสัมประสิทธิ์การถดถอย

SSE คือผลรวมของความคลาดเคลื่อนกำลังสอง

MSE คือความคลาดเคลื่อนกำลังสองเฉลี่ย

YP คือค่าพยากรณ์ของ Y

RES คือความคลาดเคลื่อนกำลังสอง

โดยขั้นตอนการทำงานของโปรแกรมย่อย MSQE(N,KE1,XE,Y,B,SSE, MSE,YP,RES) คือเป็นโปรแกรมที่ใช้หาค่าพยากรณ์ของ Y คือ YP แต่ละค่า แล้วนำมาหาความคลาดเคลื่อนกำลังสองโดยการนำค่าจริง (Y) ลบด้วยค่าพยากรณ์แล้วยกกำลังสอง และหาผลรวมของความคลาดเคลื่อนกำลังสอง จากนั้นนำมาหาความคลาดเคลื่อนกำลังสองเฉลี่ย

ขั้นตอนที่ 6

การหาค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลาง ของแต่ละวิธี ซึ่งคำนวณจากโปรแกรมย่อย NON(B,T,V,N,TRAC,MSE,PE,BIAS,AMD,ZET)

โดยที่

B คือสัมประสิทธิ์การถดถอย

T คือเมทริกซ์ T

V คือเมทริกซ์ V

N คือจำนวนตัวอย่าง

TRAC คือผลรวมแนวทแยงมุมของ $T^{-1}L$ หาค่าด้วย p_2 ในที่นี้ p_2 เท่ากับ 3

MSE คือความคลาดเคลื่อนกำลังสองเฉลี่ย

PE คือค่า g ในแต่ละแผนแบบการทดลอง

BIAS คือค่าความเอนเอียงของการประมาณ

AMD คือค่าพารามิเตอร์ไร้ศูนย์กลาง

ZET คือค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลาง

โดยขั้นตอนการทำงานของโปรแกรมย่อย NON(B,T,L,N,TRAC,MSE,PE,BIAS, AMD,ZET) คือเป็นโปรแกรมย่อยสำหรับหาค่าความเอนเอียงของการประมาณ ค่าพารามิเตอร์ไร้ศูนย์กลาง และค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลาง ซึ่งการหาค่าความเอนเอียงของการประมาณสามารถหาได้จากสูตร

$$\begin{aligned} B_1 &= (n/\sigma^2) \int_R \beta_2' (x_1 A - x_1')' (x_1 A - x_2') \beta_2 dx \\ &= (n/\sigma^2) \beta_2' T \beta_2 / \sigma^2 \end{aligned}$$

และการหาค่าพารามิเตอร์ไร้ศูนย์กลางสามารถหาได้จากสูตร

$$\lambda = n\beta_1 L\beta_2 / \sigma^2$$

และการหาค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลางสามารถหาได้จากสูตร

$$\begin{aligned} \lambda_{ave}^* &= \int_{\phi^*} \lambda d\beta_2 / \int_{\phi^*} d\beta_2 \\ &= \delta \text{tr}(\mathbf{T}^{-1}\mathbf{L}) / p_2 \end{aligned}$$

ขั้นตอนที่ 7

เมื่อหาค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลางในแต่ละวิธีได้แล้ว นำมาสรุปผลลงในตาราง เพื่อแสดงการเปรียบเทียบและศึกษาแนวโน้มของค่าเฉลี่ยของพารามิเตอร์ไร้ศูนย์กลางของแต่ละวิธี

ศูนย์วิทยพัชกร
จุฬาลงกรณ์มหาวิทยาลัย