



### บทที่ 3

#### วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ในสมการถดถอยโลจิสติก ระหว่าง

- ก) วิธีภาวน่าจะเป็นสูงสุด
- ข) ฟังก์ชันจำแนกประเภท
- ค) วิธีกำลังสองน้อยสุดถ่วงน้ำหนัก

โดยศึกษาในกรณีที่ประชากรของตัวแปรอธิบายมีการแจกแจงแบบเบ้ คือ การแจกแจงแบบไวบูลล์ และ แบบชี้กำลัง นอกจากนี้ยังได้นำตัวแปรอธิบายที่มีการแจกแจงแบบปกติมาศึกษาเปรียบเทียบกับ เพราะว่าเป็นการแจกแจงที่จำเป็นสำหรับข้อสมมุติเบื้องต้นสำหรับวิธีฟังก์ชันจำแนกประเภท

ในการเปรียบเทียบวิธีการประมาณจะพิจารณาจาก ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Square Root Mean Squares Error) และ Deviance

สำหรับข้อมูลที่ใช้ในการวิจัย เป็นข้อมูลที่ถูกสร้างขึ้นโดยใช้เทคนิคการจำลองค่าโดยวิธีมอนติคาร์โล (Monte Carlo Method) ผู้เขียนขอแนะนำขั้นตอนของการวิจัยเป็นลำดับดังนี้

#### การจำลองข้อมูลด้วยวิธีมอนติคาร์โล (Simulation by Monte Carlo Method)

วิธีมอนติคาร์โลเป็นสาขาหนึ่งของคณิตศาสตร์เชิงทดลอง เป็นเทคนิคในการจำลองตัวแบบทางคณิตศาสตร์ ซึ่งหลักการเบื้องต้นคือ ต้องจำลองตัวเลขสุ่ม (random number) มาช่วยในการหาคำตอบของปัญหาที่ต้องการศึกษา ขั้นตอนที่สำคัญของการจำลองข้อมูลด้วยวิธีมอนติคาร์โล มี 3 ขั้นตอน คือ

### ขั้นตอนที่ 1 การสร้างตัวเลขสุ่ม

การสร้างตัวเลขสุ่มเป็นสิ่งสำคัญมากในวิธีมอนติคาร์โล ได้มีผู้เสนอวิธีการสร้างตัวเลขสุ่มไว้หลายวิธี แต่วิธีที่ดีนั้นลักษณะของตัวเลขสุ่มที่เกิดขึ้นจะต้องมีการแจกแจงสม่ำเสมอ (Uniform Distribution) ในช่วง  $[0,1]$  และเป็นอิสระซึ่งกันและกัน

### ขั้นตอนที่ 2 การประยุกต์ของปัญหาที่ต้องการศึกษามาใช้กับตัวเลขสุ่ม

ขั้นตอนนี้ขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจใช้ตัวเลขสุ่มโดยตรง แต่บางปัญหาอาจใช้ตัวเลขสุ่มเพียงบางขั้นตอนเท่านั้น

### ขั้นตอนที่ 3 การทดลองกระทำ

เมื่อประยุกต์ปัญหาเพื่อใช้กับตัวเลขสุ่มได้แล้ว ขั้นตอนต่อไปก็คือ การทดลองโดยใช้กระบวนการสุ่ม (random process) มากระทำในลักษณะซ้ำๆ กัน (replication) เพื่อหาคำตอบที่ต้องการของปัญหาที่ต้องการศึกษา

## แผนการทดลอง

แผนการทดลองมีดังนี้

- ก) เลือกตัวอย่างอย่างสุ่มจากประชากรที่มีการแจกแจงเดียวกัน สำหรับประชากรที่มีการแจกแจงแบบเบ้ที่ต้องการศึกษา คือ
- การแจกแจงแบบชี้กำลัง สนใจศึกษาที่  $\theta = 0.5, 1$  และ  $2$
- การแจกแจงแบบไวบูลล์ สนใจศึกษาที่  $\alpha = 2, \beta = 0.5, 1$  และ  $2$
- และการแจกแจงที่นำมาศึกษาเปรียบเทียบ คือ การแจกแจงแบบปกติ สนใจศึกษาที่  $\mu = 2, \sigma^2 = 0.25, 1$  และ  $4$
- ข) กำหนดให้ความผิดพลาด ( $\epsilon$ ) มีการแจกแจงสม่ำเสมอแบบไม่ต่อเนื่อง มีค่าที่เป็นไปได้ 2 ค่า คือ  $-1, 1$
- ค) กำหนดขนาดตัวอย่างที่ต้องการศึกษา คือ  $20, 40, 60$  และ  $80$
- ง) จำนวนตัวแปรอธิบายที่ใช้ในการศึกษา คือ  $1$  ตัว และ  $2$  ตัว
- จ) กำหนดสัดส่วนของตัวแปรตามระหว่าง  $1$  กับ  $0$  จำนวน 10 สัดส่วน ดังนี้  $0.50:0.50, 0.55:0.45, 0.60:0.40, 0.65:0.35, 0.70:0.30, 0.75:0.25, 0.80:0.20, 0.85:0.15, 0.90:0.10$  และ  $0.95:0.05$

## ขั้นตอนในการศึกษาวิจัย

ขั้นตอนในการวิจัยมีดังนี้ คือ

1. การกำหนดลักษณะการแจกแจงของตัวแปรอธิบาย จำนวนตัวแปรอธิบาย ความผิดพลาด ขนาดตัวอย่าง และสัดส่วนของตัวแปรตาม ตามแผนการทดลองข้างต้น
2. การสร้างข้อมูลที่ใช้ในการวิจัย คือ สร้างตัวแปรอธิบาย ความผิดพลาด ตามลักษณะที่กำหนด และสร้างตัวแปรตาม จากตัวแปรอธิบาย และ ความผิดพลาด โดยให้ตัวแปรตามมีความสัมพันธ์เชิงเส้นกับตัวแปรอธิบาย จากนั้นแปลงค่าตัวแปรตาม ให้มี 2 ค่า คือ 1 หรือ 0
3. การคำนวณค่าพารามิเตอร์โดยวิธีทั้ง 3 วิธี
  - ก) วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood)
  - ข) วิธีฟังก์ชันจำแนกประเภท (Discriminant Function)
  - ค) วิธีกำลังสองน้อยสุดถ่วงน้ำหนัก (Weighted Least Squares)
4. การหาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Square Root Mean Squares Error) และ Deviance
5. สรุปผลในรูปของตาราง

## สำหรับรายละเอียดของแต่ละขั้นตอนเป็นดังนี้

### ขั้นตอนที่ 1

การกำหนดลักษณะการแจกแจงของตัวแปรอธิบาย จำนวนตัวแปรอธิบาย ความผิดพลาด ขนาดตัวอย่าง และสัดส่วนของตัวแปรตาม จะกำหนดตามแผนการทดลองที่ได้กล่าวไว้แล้วข้างต้น

### ขั้นตอนที่ 2

การสร้างข้อมูลที่มีลักษณะการแจกแจงที่ต้องการศึกษาจะใช้โปรแกรมภาษาฟอร์แทรน 4 (Fortran IV) บนเครื่อง AMDAHL 5860 ซึ่งการสร้างข้อมูลที่มีลักษณะการแจกแจงแบบต่างๆ นั้น จะต้องใช้ตัวเลขสุ่มเป็นพื้นฐานในการสร้าง โปรแกรมย่อยที่ใช้ในการสร้างตัวเลขสุ่ม คือ RANDOM(IX,RD) เมื่อสร้างตัวเลขสุ่มได้แล้ว นำตัวเลขสุ่มที่ได้มาสร้างข้อมูลที่มีการแจกแจงตามที่สนใจศึกษา ดังนี้

ก) การแจกแจงแบบปกติ

การสร้างข้อมูลที่มีการแจกแจงแบบปกติจะใช้วิธีของ *Box* และ *Muller* (1958) ซึ่งเป็นวิธีที่สร้างการแจกแจงแบบปกติที่มีค่าเฉลี่ย  $\mu$  และ ความแปรปรวน  $\sigma^2$  โปรแกรมย่อย คือ NORMAL(SMEAN,SIGMA,XNOR,IX,IK) รายละเอียดโปรแกรมย่อยแสดงในภาคผนวก ค

ข) การแจกแจงแบบชี้กำลัง

การสร้างข้อมูลที่มีการแจกแจงแบบชี้กำลัง จะใช้เทคนิคการแปลงผกผัน (inverse transformation) ซึ่งเป็นเทคนิคที่ใช้แปลงตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ ในช่วง  $[0,1]$  ให้เป็นตัวเลขสุ่มที่มีการแจกแจงแบบอื่นๆ โปรแกรมย่อย คือ EXPO(THETA,XEXP,IX) รายละเอียดโปรแกรมย่อยแสดงในภาคผนวก ค

ค) การแจกแจงแบบไวบูลล์

การสร้างข้อมูลที่มีการแจกแจงแบบไวบูลล์จะใช้ เทคนิคการแปลงผกผัน (inverse transformation) เช่นเดียวกับการแจกแจงแบบชี้กำลัง โปรแกรมย่อย คือ WEIBUL(ALPHA,BETA,XWEI,IX) รายละเอียดโปรแกรมย่อยแสดงในภาคผนวก ค

ข้อมูลที่สร้างจะได้แก่ ตัวแปรอธิบาย ( $x_1$  สำหรับตัวแปรอธิบาย 1 ตัว และ  $x_1, x_2$  สำหรับตัวแปรอธิบาย 2 ตัว) โดยสร้างจากการแจกแจงตามที่กล่าวแล้วข้างต้น และสร้างความผิดพลาด ( $\varepsilon$ ) จากฟังก์ชัน IDU(I,J) หลังจากที่ได้ข้อมูลตัวแปรอธิบาย และความผิดพลาดแล้ว จึงสร้างตัวแปรตาม ( $y$ ) โดยสร้างให้ตัวแปรตาม มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอธิบาย และความผิดพลาด ซึ่งมีรูปแบบดังนี้

$$y = \mathbf{BX} + \varepsilon$$

เมื่อ  $\mathbf{X}$  เป็นเมทริกซ์ของตัวแปรอธิบาย

$\mathbf{B}$  เป็นเวกเตอร์ของค่าพารามิเตอร์ที่กำหนด

และ  $\varepsilon$  เป็นความผิดพลาด ซึ่ง  $\varepsilon \sim U(-1,1)$

จากนั้นแปลงค่าตัวแปรตาม ( $y$ ) ที่ได้เป็นให้เป็น 1 หรือ 0

### ขั้นตอนที่ 3

หลังจากที่มีข้อมูลพร้อมแล้ว ทำการประมาณค่าพารามิเตอร์จากวิธีทั้ง 3 วิธี จาก MAIN PROGRAM 2 และ โปรแกรมย่อย DISC(B1,B2,X,Y,AA) , WLS(B1,B2,A,NN,PI) ขั้นตอนและโปรแกรมดังกล่าวเป็น ขั้นตอนของวิธีภาวะน่าจะเป็นสูงสุด โปรแกรมย่อยของวิธี ฟังก์ชันจำแนกประเภท และ โปรแกรมย่อยของวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก

### ขั้นตอนที่ 4

หลังจากได้ค่าพารามิเตอร์จากแต่ละวิธีแล้ว จะคำนวณหาค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และ Deviance จากโปรแกรมย่อย SE(B1,B2,X,Y,AA,XRMSE,XDV) สำหรับ วิธีภาวะน่าจะเป็นสูงสุด และฟังก์ชันจำแนกประเภท และจากโปรแกรมย่อย COMP(B1,B2,PI,A,NN,XRMSE,XDV) สำหรับวิธีกำลังสองน้อยสุดถ่วงน้ำหนัก

### ขั้นตอนที่ 5

เมื่อได้ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย และ Deviance ของแต่ละวิธีแล้ว นำผลมาสรุปลงในตารางเพื่อแสดงการเปรียบเทียบและศึกษาแนวโน้มของแต่ละวิธี

## ผังงานแสดงขั้นตอนในการวิจัย

