บทที่3
วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการศึกษาและเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของตัวแบบถดถอยโลจิสติคด้วยวิธีความควรจะเป็นสูงสุด วิธีการถ่วงน้ำหนัก และวิธีปรับแก้เบื้องต้น ซึ่งในการเปรียบเทียบจะพิจารณาจากค่า AMH ข้อมูลที่ใช้ในการวิจัยนี้เป็นการจำลองสุ่มด้วยเทคนิคการจำลองแบบมอนติคาร์โล (Monte carlo simulation technique) สร้างสถานการณ์การจำลองโดยใช้โปรแกรม S – plus กับเครื่อง PC สำหรับแผนการทดลองและขั้นตอนในการวิจัยรวมทั้งรายละเอียดเกี่ยวกับการทำงานของโปรแกรมที่ใช้ในการวิจัยจะนำเสนอในลำดับต่อไป

3.1 แผนการดำเนินการวิจัย

ในการดำเนินการวิจัยครั้งนี้กำหนดสถานการณ์ต่าง ๆ สำหรับศึกษา และเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของตัวแบบถดถอยโลจิสติคด้วย วิธีความควรจะเป็นสูงสุด วิธีการถ่วงน้ำหนัก และวิธีปรับแก้เบื้องต้น ดังนี้

1. ตัวแปรอิสระ(X) ที่ศึกษามี 3 ระดับ คือ 3 , 5 และ 7 ตัว
2. ตัวแปรอิสระแต่ละตัวเป็นค่าคงที่
3. ตัวแปรอิสระแต่ละตัวไม่มีความสัมพันธ์กัน
4. ตัวแปรอิสระเป็นข้อมูลเชิงปริมาณที่มีการแจกแจงดังนี้

   - การแจกแจงปกติ (Normal Distribution)
     ฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}(x_i-\mu)^2}$$

เมื่อ $E(X)=\mu, V(X)=\sigma^2$

5. ตัวแปรตาม Y_i เป็นอิสระซึ่งกันและกันและมีการแจกแจงแบบทวินามด้วยพารามิเตอร์ n_i และ $\pi(x_i)$ ซึ่งกำหนดดังนี้

   5.1 กำหนด n_i=n เท่ากับ 10 20 และ 30

   5.2 $\pi(x_i)=\frac{e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}{1+e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}$; i=1,2,...,m

6. จำนวนกลุ่มที่ทำการศึกษา(m) คือ 30 , 90 , 150 และ 210
7. กำหนดค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจของประชากร $\bar{\pi}(x_i)$ เท่ากับ 0.1 , 0.3 , 0.5 และ 0.8
8. กำหนด $\beta=1$
9. กำหนดจำนวนการกระทำซ้ำในแต่ละสถานการณ์เป็น 500 ครั้ง
10. การวิจัยครั้งนี้ได้ใช้วิธีการจำลองสุ่ม (Simulation) ให้มีสถานการณ์ตามที่กำหนดด้วยโปรแกรม S – plus

3.2 ขั้นตอนในการดำเนินการวิจัย

ขั้นตอนในการดำเนินการวิจัยมีดังนี้ คือ

3.2.1 สร้างข้อมูลที่ใช้ในการวิจัย

3.2.2 คำนวณหาค่าประมาณพารามิเตอร์ด้วยวิธี

3.2.2.1 ความควรจะเป็นสูงสุด

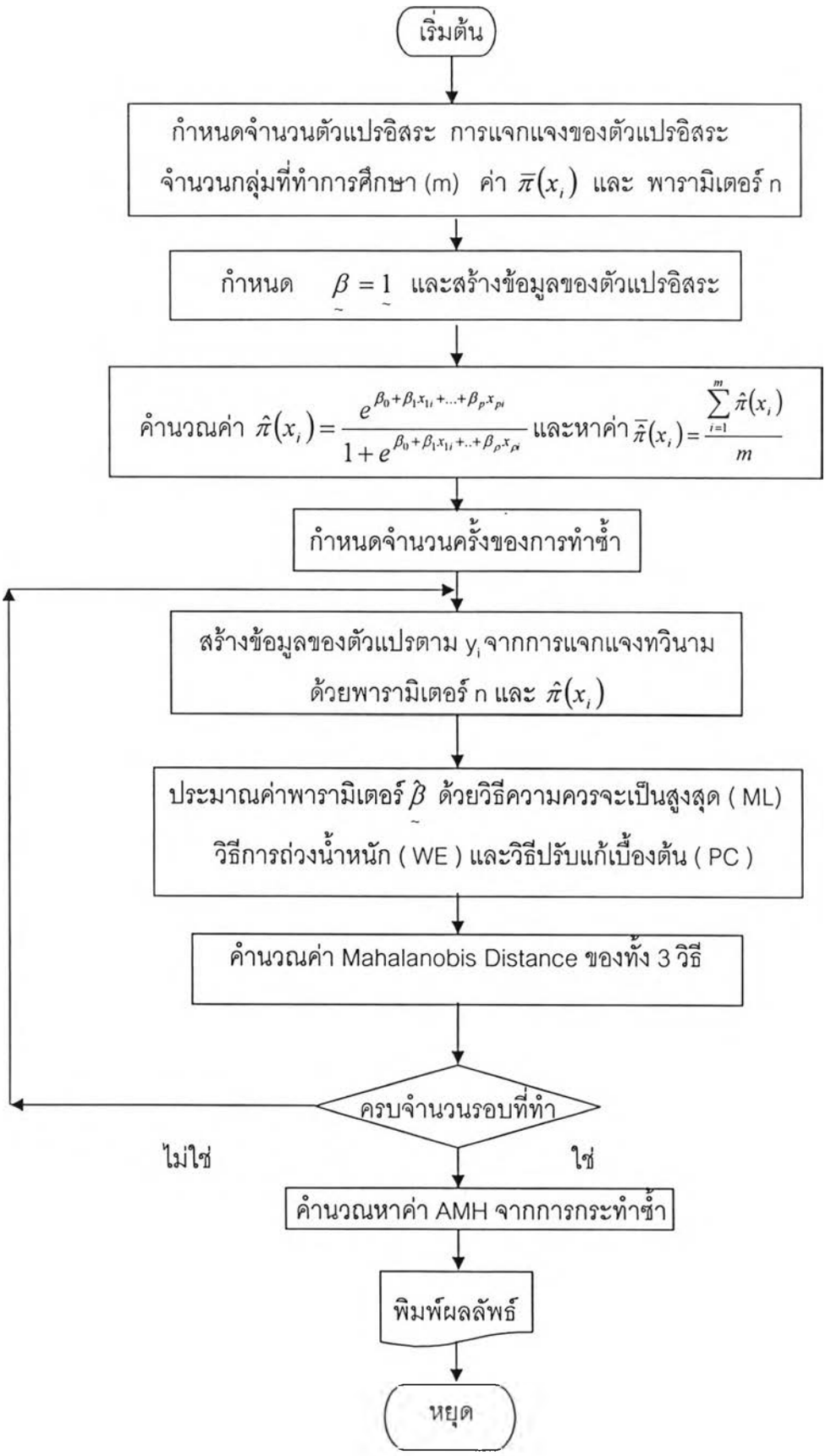
3.2.2.2 การถ่วงน้ำหนัก

3.2.2.3 ปรับแก้เบื้องต้น

3.2.3 คำนวณค่า AMH จากการกระทำซ้ำ

3.2.4 ทำการเปรียบเทียบและสรุปผลในรูปตาราง

ผังงานแสดงขั้นตอนในการวิจัย



รายละเอียดของแต่ละขั้นตอน

3.2.1 สร้างข้อมูลที่ใช้ในการวิจัย

ในการสร้างข้อมูลที่ใช้ในการวิจัยมีรายละเอียดและขั้นตอนดังนี้

3.2.1.1 สร้างตัวแปรอิสระ X โดยกำหนดให้เป็นค่าคงที่ตามจำนวนตัวแปรอิสระ การแจกแจงเป็นแบบปกติ จำนวนกลุ่มที่ทำการศึกษา ค่า $\bar{\pi}(x_i)$ และ พารามิเตอร์ n ที่กำหนดไว้ในแผนการดำเนินการวิจัยข้างต้น โดยใช้โปรแกรม S-plus

3.2.1.2 สร้างข้อมูลตัวแปรตาม y_i จากการแจกแจงทวินามด้วยพารามิเตอร์ n และ $\hat{\pi}(x_i)$ เมื่อ $\hat{\pi}(x_i)$ คำนวณจาก $\hat{\pi}(x_i)=\frac{e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}{1+e^{\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi}}}$ โดยที่กำหนดให้ $\beta=1$ และหาค่า $\bar{\hat{\pi}}(x_i)$ จาก $\bar{\hat{\pi}}(x_i)=\frac{\sum_{i=1}^{m}\hat{\pi}(x_i)}{m}$

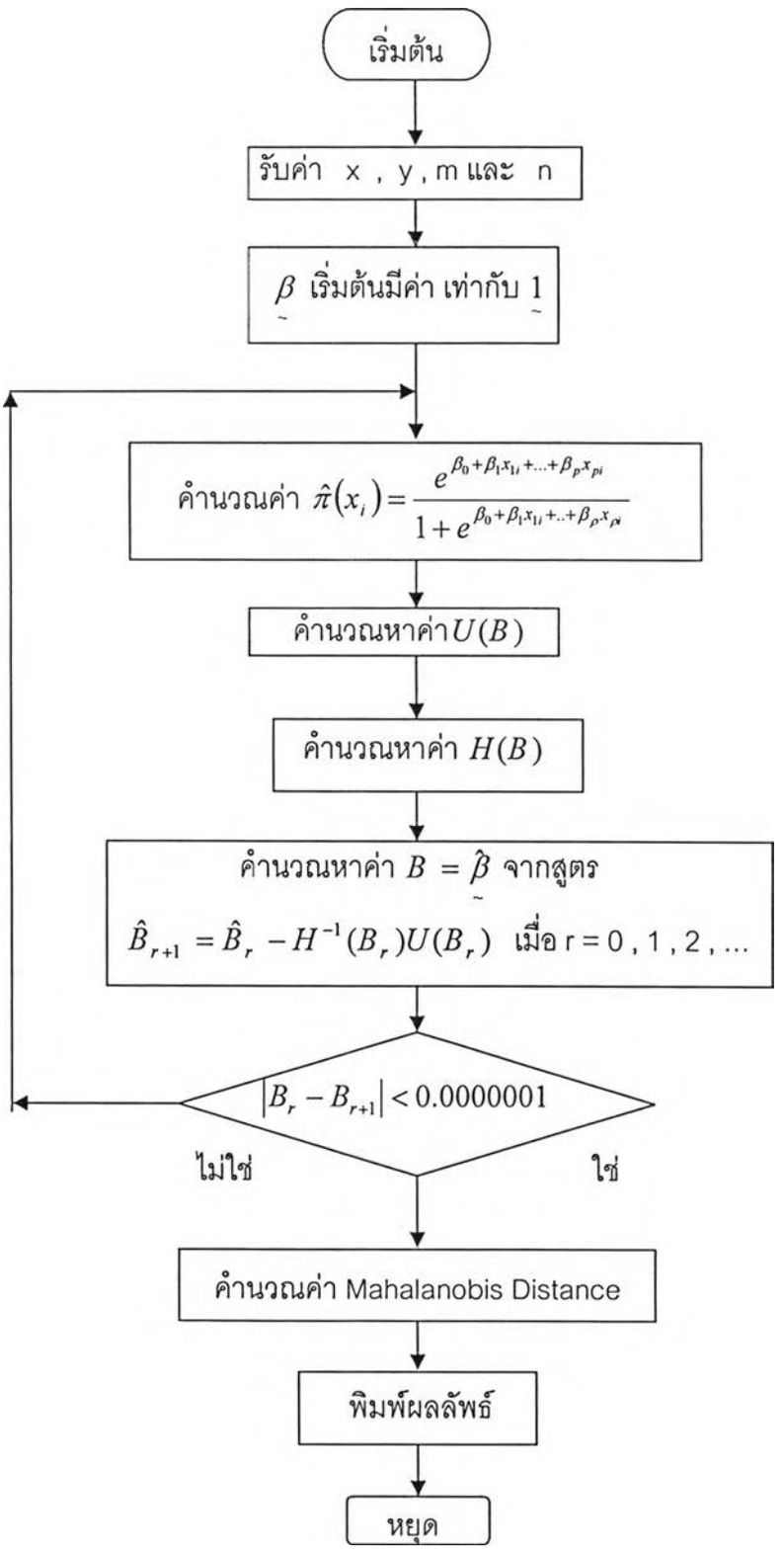
3.2.2 หาค่าประมาณพารามิเตอร์

มีขั้นตอนดังนี้

3.2.2.1 วิธีความควรจะเป็นสูงสุด

พิจารณาตัวประมาณค่าพารามิเตอร์ที่มีฟังก์ชันความควรจะเป็นสูงสุด ซึ่งจะใช้ Newton–Raphson โดยทำการหาอนุพันธ์อันดับหนึ่งของลอการิทึมของฟังก์ชันความควรจะเป็น $l(\beta)$ เทียบกับ $\beta_0, \beta_1, \ldots, \beta_p$ ซึ่งคือ $U(B)$ และหาอนุพันธ์อันดับที่สองของ $l(\beta)$ ซึ่งคือ $H(B)$ จากนั้นนำค่า $U(B)$ และ $H(B)$ มาแทนลงในสมการ $\hat{B}_{r+1}=\hat{B}_r-H^{-1}(B_r)U(B_r)$ โดยที่ค่า β เริ่มต้นเท่ากับ 1 จะมีการทำซ้ำจนกว่าจะได้ค่า $\hat{\beta}$ ที่มีค่าไม่ต่างกันมากเป็นที่ยอมรับได้ สำหรับโปรแกรม S-plus ใช้คำสั่ง ml.logit(x,y,m,n) เมื่อ x แทนเมทริกซ์ของตัวแปรอิสระ y แทนเวกเตอร์ของตัวแปรตาม m แทน จำนวนกลุ่ม และ n แทนค่าของ ni ซึ่งขั้นตอนเป็นดังนี้

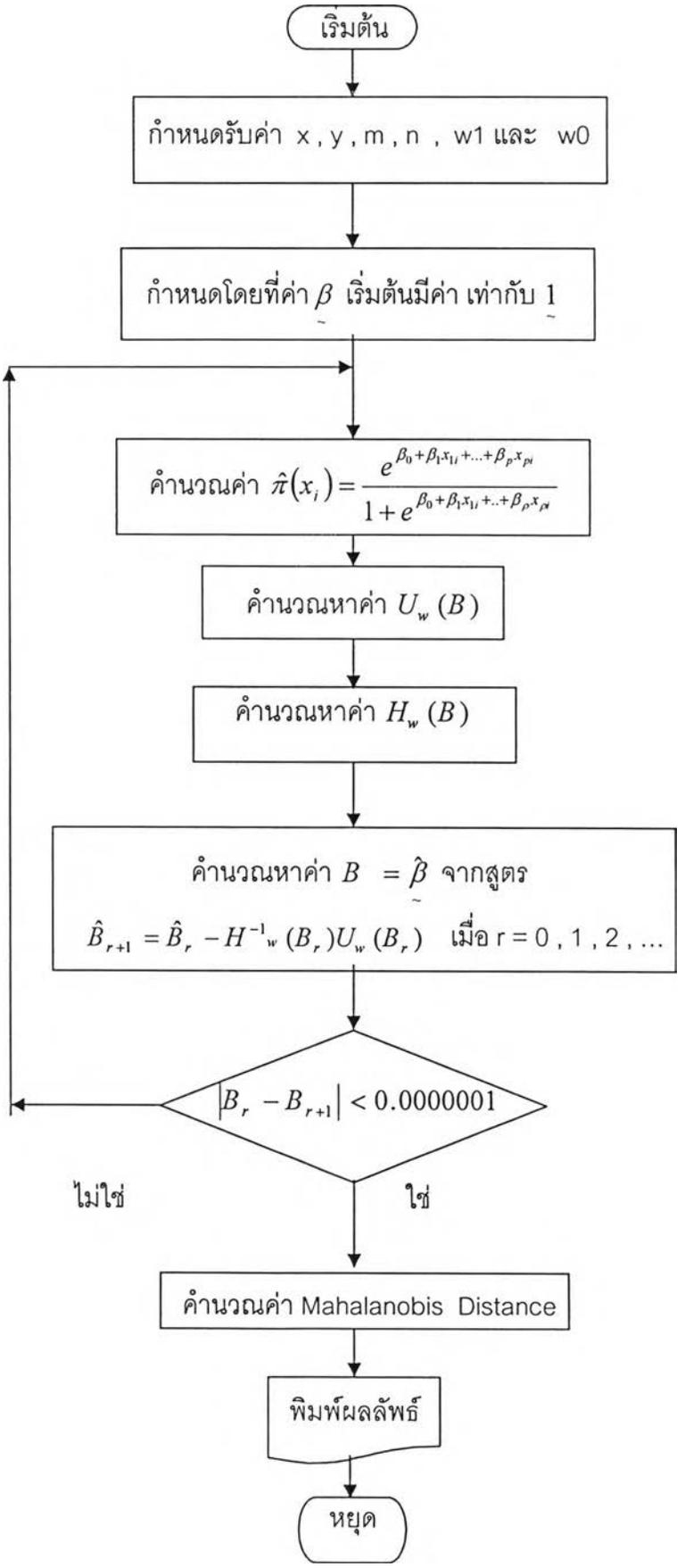
วิธีความควรจะเป็นสูงสุด



3.2.2.2 **วิธีการถ่วงน้ำหนัก**

วิธีการถ่วงน้ำหนักนี้จะใช้ตัวถ่วงน้ำหนัก คือ $w_1 = \frac{\bar{\pi}(x_i)}{\hat{\bar{\pi}}(x_i)}$ และ $w_0 = \frac{(1-\bar{\pi}(x_i))}{(1-\hat{\bar{\pi}}(x_i))}$ ซึ่ง กำหนดค่า $\bar{\pi}(x_i)$ มีค่าเท่ากับ 0.1 , 0.3 , 0.5 และ 0.8 หาค่า $\hat{\bar{\pi}}(x_i) = \frac{\sum_{i=1}^{m} \hat{\pi}(x_i)}{m}$ ซึ่งในการหาตัวประมาณค่าพารามิเตอร์โดยวิธีการถ่วงน้ำหนักนี้ จะใช้วิธี Newton – Raphson โดยทำการหาอนุพันธ์อันดับหนึ่งของลอการิทึมของฟังก์ชันความควรจะเป็น $l_w(\beta)$ เทียบกับ $\beta_0, \beta_1, ..., \beta_p$ ซึ่งคือ $U_w(B)$ และหาอนุพันธ์อันดับที่สองของ $l_w(\beta)$ ซึ่งคือ $H_w(B)$ จากนั้นนำค่า $U_w(B)$ และ $H_w(B)$ มาแทนลงในสมการ $\hat{B}_{r+1} = \hat{B}_r - H^{-1}{}_w(B_r)U_w(B_r)$ โดยที่ค่า β เริ่มต้นมีค่า เท่ากับ 1 จะมีการทำซ้ำจนกว่าจะได้ค่า $\hat{\beta}$ ที่มีค่าไม่ต่างกันมากเป็นที่ยอมรับได้ สำหรับโปรแกรม S-plus ใช้คำสั่ง we.logit (x,y,m,n,w1,w0) เมื่อ x แทนเมทริกซ์ของตัวแปรอิสระ y แทนเวกเตอร์ของตัวแปรตาม m แทน จำนวนกลุ่ม n แทนค่าของ ni w1 และ w0 แทน สัดส่วนการถ่วงน้ำหนัก ซึ่งขั้นตอนเป็นดังนี้

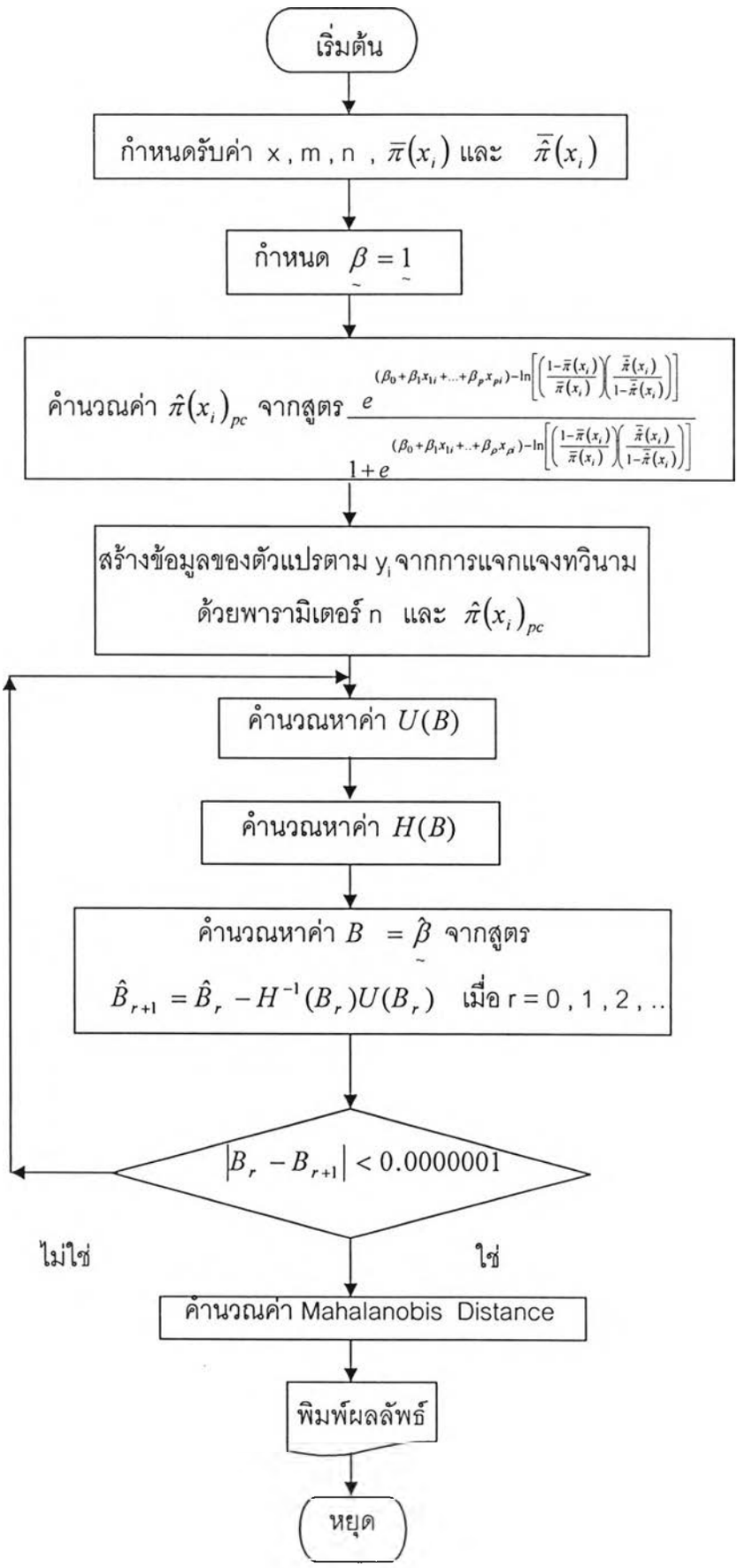
วิธีการถ่วงน้ำหนัก



3.2.2.3 **วิธีปรับแก้เบื้องต้น**

วิธีการปรับแก้เบื้องต้นนี้จะเป็นการใช้ค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจของประชากรคือ $\bar{\pi}(x_i)$ซึ่งกำหนดให้มีค่าเท่ากับ 0.1 , 0.3 , 0.5 และ 0.8 และหาค่า $\bar{\hat{\pi}}(x_i)=\frac{\sum_{i=1}^{m}\hat{\pi}(x_i)}{m}$ ซึ่งวิธีจะปรับค่า $\hat{\pi}(x_i)$ ให้เป็นค่า $\hat{\pi}(x_i)_{pc}$ โดยมีค่าเท่ากับ $\frac{e^{(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi})-\ln\left[\left(\frac{1-\bar{\pi}(x_i)}{\bar{\pi}(x_i)}\right)\left(\frac{\bar{\hat{\pi}}(x_i)}{1-\bar{\hat{\pi}}(x_i)}\right)\right]}}{1+e^{(\beta_0+\beta_1 x_{1i}+\ldots+\beta_p x_{pi})-\ln\left[\left(\frac{1-\bar{\pi}(x_i)}{\bar{\pi}(x_i)}\right)\left(\frac{\bar{\hat{\pi}}(x_i)}{1-\bar{\hat{\pi}}(x_i)}\right)\right]}}$ ซึ่งจะใช้ Newton – Raphson โดยทำการหาอนุพันธ์อันดับหนึ่งของลอการิทึมของฟังก์ชันความควรจะเป็น $l(\beta)$ เทียบกับ $\beta_0, \beta_1, \ldots, \beta_p$ ซึ่งคือ $U(B)$ และหาอนุพันธ์อันดับที่สองของ $l(\beta)$ ซึ่งคือ $H(B)$ จากนั้นนำค่า $U(B)$ และ $H(B)$ มาแทนลงในสมการ $\hat{B}_{r+1}=\hat{B}_r-H^{-1}(B_r)U(B_r)$ โดยที่ค่า β เริ่มต้นเท่ากับ 1 จะมีการทำซ้ำจนกว่าจะได้ค่า $\hat{\beta}$ ที่มีค่าไม่ต่างกันมากเป็นที่ยอมรับได้ สำหรับโปรแกรม S-plus ใช้คำสั่ง pc.logit(x,m,n,inde, pi.M,pi.Mhat) เมื่อ x แทนเมทริกซ์ของตัวแปรอิสระ m แทน จำนวนกลุ่ม n แทนค่าของ ni pi.M และ pi.Mhat แทน ค่าเฉลี่ยความน่าจะเป็นของเหตุการณ์ที่สนใจของประชากรและตัวอย่าง ตามลำดับซึ่งขั้นตอนเป็นดังนี้

วิธีปรับแก้เบื้องต้น



3.2.3 คำนวณหาค่าAMH จากการกระทำซ้ำ

หลังจากได้ค่าประมาณพารามิเตอร์ของตัวประมาณพารามิเตอร์ของตัวแบบถดถอยโลจิสติคแต่ละวิธีแล้ว คำนวณหาค่า AMH จากการกระทำซ้ำ

3.2.4 ทำการเปรียบเทียบและสรุปผลในรูปตาราง

เมื่อได้ค่า AMH ของแต่ละวิธีแล้ว นำผลที่ได้มาสรุปลงในตารางเพื่อแสดงการเปรียบเทียบและศึกษาแนวโน้มของแต่ละวิธีการ