บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการศึกษาวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามลบแบบทั่วไป โดยประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณแบบความควรจะเป็นสูงสุด วิธีการประมาณแบบสองโมเมนต์แรกและสัดส่วนศูนย์ วิธีการประมาณแบบไคกำลังสองต่ำสุด และวิธีการประมาณแบบระยะห่างต่ำสุด ทั้ง 4 วิธีจะใช้การกระทำวนซ้ำ (Iterative) จนได้ค่าประมาณพารามิเตอร์ ในแต่ละวิธีการประมาณค่าพารามิเตอร์จะกระทำภายใต้สถานการณ์เมื่อขนาดตัวอย่างมี 5 ขนาด คือ 30 , 50 , 70 , 100 และ 200 โดยในแต่ละสถานการณ์มีการทำซ้ำ 1,000 ครั้ง

การวิจัยครั้งนี้ใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) ในการสร้างสถานการณ์ต่าง ๆ ดังนั้นจะขอกล่าวถึงวิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล แล้วจึงแสดงรายละเอียดของแต่ละขั้นตอนในการวิจัยและโปรแกรมที่ใช้ในการวิจัยตามลำดับ

3.1 วิธีการจำลองโดยใช้เทคนิคมอนติคาร์โล

เทคนิคมอนติคาร์โลเป็นเทคนิคการสร้างข้อมูลโดยการจำลองตัวเลขสุ่ม (Random Number) ด้วยโปรแกรมคอมพิวเตอร์ภาษาฟอร์แทรน (FORTRAN) ซึ่งตัวเลขสุ่มที่ได้จะมีการแจกแจงแบบเอกรูปในช่วง (0,1) โดยตัวเลขสุ่มที่ได้ควรมีคุณสมบัติดังต่อไปนี้

1. ตัวเลขที่ได้มีการกระจายของความน่าจะเป็นแบบสม่ำเสมอและเป็นอิสระซึ่งกันและกัน
2. อนุกรมของตัวเลขที่ได้สามารถสร้างซ้ำเดิมได้ (Reproducible)
3. อนุกรมของตัวเลขไม่ซ้ำเดิมในช่วงที่ต้องการใช้ตัวเลขแบบสุ่ม หมายความว่าขนาดความยาวของอนุกรมตัวเลขต้องยาวพอสำหรับใช้งาน
4. ใช้เวลาสั้นๆ ในการสร้างตัวเลขสุ่ม
5. ใช้หน่วยความจำของคอมพิวเตอร์น้อย

3.2 แผนการทดลอง

การวิจัยครั้งนี้ กำหนดสถานการณ์ต่างๆ เพื่อศึกษาวิธีการประมาณค่าพารามิเตอร์ 4 วิธี สำหรับการแจกแจงทวินามลบแบบทั่วไป ดังนี้

1. กำหนดค่าพารามิเตอร์ให้สอดคล้องกับค่าสัมประสิทธิ์ความแปรผันระดับต่าง ๆ ดังต่อไปนี้

1.1 ค่าสัมประสิทธิ์ความแปรผันเท่ากับ 50 % ศึกษาที่ $\theta = 0.3$, $\beta = 1.5$, $m = 16.9697$ และ $\theta = 0.4$, $\beta = 1.1$, $m = 10.7143$

1.2 ค่าสัมประสิทธิ์ความแปรผันเท่ากับ 75 % ศึกษาที่ $\theta = 0.2$, $\beta = 1.5$, $m = 10.1587$ และ $\theta = 0.3$, $\beta = 1.5$, $m = 7.5421$

1.3 ค่าสัมประสิทธิ์ความแปรผันเท่ากับ 100 % ศึกษาที่ $\theta = 0.3$, $\beta = 1.5$, $m = 4.2424$ และ $\theta = 0.4$, $\beta = 1.1$, $m = 2.6786$

1.4 ค่าสัมประสิทธิ์ความแปรผันเท่ากับ 125 % ศึกษาที่ $\theta = 0.1$, $\beta = 2.0$, $m = 7.1999$ และ $\theta = 0.3$, $\beta = 1.5$, $m = 2.7152$

1.5 ค่าสัมประสิทธิ์ความแปรผันเท่ากับ 150 % ศึกษาที่ $\theta = 0.2$, $\beta = 1.5$, $m = 2.5397$ และ $\theta = 0.3$, $\beta = 1.5$, $m = 1.8855$

1.6 ค่าสัมประสิทธิ์ความแปรผันเท่ากับ 175 % ศึกษาที่ $\theta = 0.1$, $\beta = 2.0$, $m = 3.6735$ และ $\theta = 0.3$, $\beta = 1.5$, $m = 1.3853$

2. กำหนดขนาดตัวอย่าง (n) 5 ขนาด คือ 30 , 50 , 70 , 100 และ 200

3.3 ขั้นตอนในการวิจัย

แบ่งเป็น 4 ขั้นตอนหลัก คือ

1. การผลิตตัวแปรสุ่มให้มีการแจกแจงทวินามลบแบบทั่วไป
2. การหาค่าพารามิเตอร์เริ่มต้น
3. การประมาณค่าพารามิเตอร์ ด้วยวิธีการทั้ง 4 วิธี จะได้ค่าประมาณพารามิเตอร์ของแต่ละวิธีการ

4. การหาค่าความคลาดเคลื่อนจากการประมาณค่าพารามิเตอร์ และทำการเปรียบเทียบ

รายละเอียดแต่ละขั้นตอนเป็นดังนี้

1. การผลิตตัวแปรสุ่มให้มีการแจกแจงทวินามลบแบบทั่วไป

1.1 สร้างค่าตัวเลขสุ่ม (Y) ดังโปรแกรมย่อยสับรูทีน

1.2 การผลิตตัวแปรสุ่มให้มีการแจกแจงทวินามลบแบบทั่วไป โดยการนำตัวเลขสุ่มที่ได้จากการสร้างในข้อ 1.1 มาเปรียบเทียบกับค่าฟังก์ชันการแจกแจง (CDF)

ซึ่งฟังก์ชันการแจกแจง แสดงได้ดังนี้

$$F(a) = \sum_{x=0}^{a} \frac{m}{m+\beta x}\binom{m+\beta x}{x}\theta^{x}(1-\theta)^{m+\beta x-x}$$

โดยที่ $0 < \theta < 1$, $1 \le \beta < \theta^{-1}$, $m > 0$

ถ้าค่าของตัวเลขสุ่มที่ได้มากกว่า $CDF(x-1)$ แต่น้อยกว่าหรือเท่ากับ $CDF(x)$ ก็จะได้ตัวแปรสุ่มที่มีการแจกแจงทวินามลบเท่ากับ x

1.3 นำตัวแปรสุ่มที่ได้มาเรียงลำดับและจัดกลุ่ม

2. การกำหนดค่าพารามิเตอร์เริ่มต้น ในแต่ละวิธีการประมาณค่าจะใช้ค่าพารามิเตอร์เริ่มต้นค่าเดียวกันทั้ง 4 วิธี

ซึ่งการแจกแจงทวินามลบแบบทั่วไป มี θ_0 , β_0 และ m_0 เป็นพารามิเตอร์เริ่มต้น ซึ่งจะหาค่าพารามิเตอร์เริ่มต้นได้ดังนี้

$$\theta_0 = 1-\left(\frac{A}{2}\right)+\sqrt{\left(\frac{A^2}{4}\right)-1}$$

$$\beta_0 = \frac{1}{\theta_0}\left[1-\left\{\frac{m_1(1-\theta_0)}{\mu_2}\right\}^{1/2}\right]$$

$$m_0 = \frac{m_1(1-\beta_0\theta_0)}{\theta_0}$$

เมื่อ $$A = -2 + \left[\frac{\left(m_1 \mu_3 - 3\mu_2^2\right)^2}{m_1 \mu_2^3} \right]$$

โดยที่ m_1, μ_2, μ_3 เป็นโมเมนต์ที่หนึ่ง,โมเมนต์ที่สอง และโมเมนต์ที่สามรอบค่าเฉลี่ยของข้อมูลตามลำดับ

$$m_1 = \frac{\sum_{i=0}^{k} n_i x_i}{n}$$

$$\mu_2 = \frac{n\left(\sum_{i=1}^{k} n_i x_i^2\right) - \left(\sum_{i=1}^{k} n_i x_i\right)^2}{n(n-1)}$$

$$\mu_3 = \frac{\sum_{i=1}^{k} n_i x_i^3 - 3m_1\left(\sum_{i=1}^{k} n_i x_i^2\right) + 2nm_1^3}{n}$$

3. การประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณ 4 วิธี มีรายละเอียดดังต่อไปนี้

3.1 วิธีการประมาณแบบความควรจะเป็นสูงสุด

ฟังก์ชันความควรจะเป็นสำหรับข้อมูลที่มีการแจกแจงทวินามลบแบบทั่วไป แสดงได้ดังนี้

$$L = \prod_{i=1}^{n} \left[P\left(X_i = x_i\right)\right]$$

$$l(\theta, \beta, m) = \ln L$$

$$= \ln\left\{\prod_{i=1}^{n} \left[P\left(X_i = x_i\right)\right]\right\}$$

การประมาณค่าพารามิเตอร์ θ, β และ m ทำได้โดยการหาอนุพันธ์บางส่วนของลอกความควรจะเป็น เทียบกับ θ, β และ m ตามลำดับ และให้สมการอนุพันธ์บางส่วนเท่ากับ 0

เนื่องจากสมการการหาอนุพันธ์บางส่วนของลอกของฟังก์ชันความควรจะเป็นของการแจกแจงทวินามลบแบบทั่วไปไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจำเป็นต้องใช้วิธีการประมาณ ในการวิจัยครั้งนี้เลือกใช้วิธีนิวตัน-ราฟสัน โดยมีขั้นตอนการคำนวนดังนี้

ขั้นตอนที่1 การคำนวนค่า $g_1, g_2, g_3, g_{11}, g_{12}, g_{13}, g_{21}, g_{22}, g_{23}, g_{31}, g_{32}, g_{33}$ ซึ่งสามารถแสดงได้ดังนี้

$$g_1(\theta,\beta,m) = \frac{\partial \ln L}{\partial \theta}$$

$$g_2(\theta,\beta,m) = \frac{\partial \ln L}{\partial \beta}$$

$$g_3(\theta,\beta,m) = \frac{\partial \ln L}{\partial m}$$

$$g_{11}(\theta,\beta,m) = \frac{\partial^2 \ln L}{\partial \theta^2}$$

$$g_{12}(\theta,\beta,m) = \frac{\partial^2 \ln L}{\partial \beta \partial \theta}$$

$$= \frac{\partial^2 \ln L}{\partial \theta \partial \beta}$$

$$= g_{21}(\theta,\beta,m)$$

$$g_{13}(\theta,\beta,m) = \frac{\partial^2 \ln L}{\partial m \partial \theta}$$

$$= \frac{\partial^2 \ln L}{\partial \theta \partial m}$$

$$= g_{31}(\theta,\beta,m)$$

$$g_{22}(\theta,\beta,m) = \frac{\partial^2 \ln L}{\partial \beta^2}$$

$$g_{23}(\theta,\beta,m) = \frac{\partial^2 \ln L}{\partial m \partial \beta}$$

$$= \frac{\partial^2 \ln L}{\partial \beta \partial m}$$

$$= \quad g_{32}(\theta,\beta,m)$$

$$g_{33}(\theta,\beta,m) \quad = \quad \frac{\partial^2 \ln L}{\partial m^2}$$

เมื่อ θ , β , m เป็นพารามิเตอร์ของการแจกแจงทวินามลบแบบทั่วไป

สมการการหาอนุพันธ์อันดับหนึ่งและสองของลอกของฟังก์ชันความควรจะเป็นของการแจกแจงเมื่อเทียบกับพารามิเตอร์แต่ละตัวจะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวนค่าประมาณพารามิเตอร์ θ , β , m โดยคำนวนจากสมการดังนี้

$$\begin{bmatrix}\theta\\ \beta\\ m\end{bmatrix}=\begin{bmatrix}\theta_0\\ \beta_0\\ m_0\end{bmatrix}-\begin{bmatrix}g_{11}(\theta_0,\beta_0,m_0) & g_{12}(\theta_0,\beta_0,m_0) & g_{13}(\theta_0,\beta_0,m_0)\\ g_{21}(\theta_0,\beta_0,m_0) & g_{22}(\theta_0,\beta_0,m_0) & g_{23}(\theta_0,\beta_0,m_0)\\ g_{31}(\theta_0,\beta_0,m_0) & g_{32}(\theta_0,\beta_0,m_0) & g_{33}(\theta_0,\beta_0,m_0)\end{bmatrix}^{-1}\begin{bmatrix}g_1(\theta_0,\beta_0,m_0)\\ g_2(\theta_0,\beta_0,m_0)\\ g_3(\theta_0,\beta_0,m_0)\end{bmatrix}$$

เมื่อ θ_0 , β_0 , m_0 เป็นค่าพารามิเตอร์เริ่มต้นของ θ , β , m

จากนั้นจะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่1 และ 2 ตามลำดับ จนกระทั่งได้ค่าพารามิเตอร์ที่ต้องการ

3.2 วิธีการประมาณแบบสองโมเมนต์แรกและสัดส่วนศูนย์

หลักการของวิธีการนี้คือ จะใช้ความสัมพันธ์ของโมเมนต์ตัวอย่างสองอันดับแรกของการแจกแจงทวินามลบแบบทั่วไป

จัดให้โมเมนต์ตัวอย่างสองอันดับแรกสอดคล้องกับโมเมนต์ของประชากรจะได้สมการ

$$f_1(\theta) \quad = \quad \left(m_2 - m_1^2\right)\left(\ln f_0\right)^2 m_1^{-3}\theta^2 - (1-\theta)[\ln(1-\theta)]^2 \quad = \quad 0$$

เนื่องจากสมการข้างต้นไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจำเป็นต้องใช้วิธีการประมาณ ในการวิจัยครั้งนี้เลือกใช้วิธีนิวตัน-ราฟสัน โดยมีขั้นตอนการคำนวนดังนี้

ขั้นตอนที่ 1 การคำนวนค่า $f_1(\theta)$, , $f_1'(\theta)$

สมการการหาอนุพันธ์ของฟังก์ชันการแจกแจงเมื่อเทียบกับพารามิเตอร์ θ จะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวนค่าประมาณพารามิเตอร์ θ โดยคำนวนจากสมการดังนี้

$$\bar{\theta} = \theta_0 - \frac{f_1(\theta_0)}{f_1'(\theta_0)}$$

เมื่อ $\bar{\theta}$ เป็นพารามิเตอร์ที่ต้องการประมาณ

θ_0 เป็นค่าพารามิเตอร์เริ่มต้นของ $\bar{\theta}$

$f_1'(\theta_0)$ เป็นการหาอนุพันธ์บางส่วนของตัวสถิติ $f_1(\theta_0)$ เทียบกับ θ

จะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่ 1 และ 2 ตามลำดับ จนกระทั่งได้ค่าประมาณพารามิเตอร์ที่ต้องการ

เมื่อได้ค่า $\bar{\theta}$ ซึ่งเป็นตัวประมาณ MOZE ของ θ แล้ว ก็จะสามารถหาตัวประมาณ MOZE อื่น ๆ ได้ดังนี้

$$\bar{m} = \bar{\theta}^{-1}\left[(1-\bar{\theta})m_1^3\left(m_2 - m_1^2\right)^{-1}\right]^{1/2}$$

และ

$$\bar{\beta} = \bar{\theta}^{-1} - \frac{\bar{m}}{m_1}$$

3.3 วิธีการประมาณแบบไคกำลังสองต่ำสุด

ตัวประมาณไคกำลังสองต่ำสุด แสดงได้ดังนี้

$$Q = \sum_{x=0}^{k} \frac{[n_x - nP_x]^2}{nP_x}$$

เมื่อ P_x เป็นความน่าจะเป็นในกลุ่มที่ x สำหรับการแจกแจงทวินามลบแบบทั่วไป (แสดงไว้ในบทที่ 2)

การประมาณค่าพารามิเตอร์ทำได้โดยหาอนุพันธ์บางส่วนของตัวสถิติไคกำลังสองเทียบกับ φ และให้อนุพันธ์บางส่วนเท่ากับ 0

$$\frac{\partial Q}{\partial \varphi} = \sum_{x=0}^{k}\left[n - \frac{n_x^{\,2}}{nP_x^{\,2}}\right]\frac{\partial P_x}{\partial \varphi}$$

เมื่อ φ เป็นพารามิเตอร์ใด ๆ

เนื่องจากสมการการหาอนุพันธ์บางส่วนของตัวสถิติไคกำลังสองของการแจกแจงทวินามลบแบบทั่วไปไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจำเป็นต้องใช้วิธีการประมาณ ในการวิจัยครั้งนี้เลือกใช้วิธีนิวตัน-ราฟสัน โดยมีขั้นตอนการคำนวนดังนี้

ขั้นตอนที่ 1 การคำนวนค่า $g_1, g_2, g_3, g_{11}, g_{12}, g_{13}, g_{21}, g_{22}, g_{23}, g_{31}, g_{32}, g_{33}$ ซึ่งสามารถแสดงได้ดังนี้

$$g_1(\theta,\beta,m) = \frac{\partial Q}{\partial \theta} = \sum_{x=0}^{k}\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial P_x}{\partial \theta}$$

$$g_2(\theta,\beta,m) = \frac{\partial Q}{\partial \beta} = \sum_{x=0}^{k}\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial P_x}{\partial \beta}$$

$$g_3(\theta,\beta,m) = \frac{\partial Q}{\partial m} = \sum_{x=0}^{k}\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial P_x}{\partial m}$$

$$g_{11}(\theta,\beta,m) = \frac{\partial^2 Q}{\partial \theta^2} = \sum_{x=0}^{k}\left[\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial^2 P_x}{\partial \theta^2} + \frac{2n_x^2}{nP_x^3}\left(\frac{\partial P_x}{\partial \theta}\right)^2\right]$$

$$g_{12}(\theta,\beta,m) = \frac{\partial^2 Q}{\partial \beta \partial \theta} = \sum_{x=0}^{k}\left[\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial^2 P_x}{\partial \beta \partial \theta} + \frac{2n_x^2}{nP_x^3}\left(\frac{\partial P_x}{\partial \theta}\right)\left(\frac{\partial P_x}{\partial \beta}\right)\right]$$

$$= \frac{\partial^2 Q}{\partial \theta \partial \beta} = g_{21}(\theta,\beta,m)$$

$$g_{13}(\theta,\beta,m) = \frac{\partial^2 Q}{\partial m \partial \theta} = \sum_{x=0}^{k}\left[\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial^2 P_x}{\partial m \partial \theta} + \frac{2n_x^2}{nP_x^3}\left(\frac{\partial P_x}{\partial \theta}\right)\left(\frac{\partial P_x}{\partial m}\right)\right]$$

$$= \frac{\partial^2 Q}{\partial \theta \partial m} = g_{31}(\theta,\beta,m)$$

$$g_{22}(\theta,\beta,m) = \frac{\partial^2 Q}{\partial \beta^2} = \sum_{x=0}^{k}\left[\left(n - \frac{n_x^2}{nP_x^2}\right)\frac{\partial^2 P_x}{\partial \beta^2} + \frac{2n_x^2}{nP_x^3}\left(\frac{\partial P_x}{\partial \beta}\right)^2\right]$$

$$g_{23}(\theta,\beta,m) = \frac{\partial^2 Q}{\partial m \partial \beta} = \sum_{x=0}^{k}\left[\left(n-\frac{{n_x}^2}{n{P_x}^2}\right)\frac{\partial^2 P_x}{\partial m \partial \beta}+\frac{2{n_x}^2}{n{P_x}^3}\left(\frac{\partial P_x}{\partial \beta}\right)\left(\frac{\partial P_x}{\partial m}\right)\right]$$

$$= \frac{\partial^2 Q}{\partial \beta \partial m} = g_{32}(\theta,\beta,m)$$

$$g_{33}(\theta,\beta,m) = \frac{\partial^2 Q}{\partial m^2} = \sum_{x=0}^{k}\left[\left(n-\frac{{n_x}^2}{n{P_x}^2}\right)\frac{\partial^2 P_x}{\partial m^2}+\frac{2{n_x}^2}{n{P_x}^3}\left(\frac{\partial P_x}{\partial m}\right)^2\right]$$

เมื่อ θ , β , m เป็นพารามิเตอร์ของการแจกแจงทวินามลบแบบทั่วไป

สมการการหาอนุพันธ์อันดับหนึ่งและสองของ P_x เมื่อเทียบกับพารามิเตอร์แต่ละตัวจะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวนค่าประมาณพารามิเตอร์ θ , β , m โดยคำนวนจากสมการดังนี้

$$\begin{bmatrix}\theta\\ \beta\\ m\end{bmatrix}=\begin{bmatrix}\theta_0\\ \beta_0\\ m_0\end{bmatrix}-\begin{bmatrix}g_{11}(\theta_0,\beta_0,m_0) & g_{12}(\theta_0,\beta_0,m_0) & g_{13}(\theta_0,\beta_0,m_0)\\ g_{21}(\theta_0,\beta_0,m_0) & g_{22}(\theta_0,\beta_0,m_0) & g_{23}(\theta_0,\beta_0,m_0)\\ g_{31}(\theta_0,\beta_0,m_0) & g_{32}(\theta_0,\beta_0,m_0) & g_{33}(\theta_0,\beta_0,m_0)\end{bmatrix}^{-1}\begin{bmatrix}g_1(\theta_0,\beta_0,m_0)\\ g_2(\theta_0,\beta_0,m_0)\\ g_3(\theta_0,\beta_0,m_0)\end{bmatrix}$$

เมื่อ θ_0 , β_0 , m_0 เป็นค่าพารามิเตอร์เริ่มต้นของ θ , β , m

จากนั้นจะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่1 และ 2 ตามลำดับ จนกระทั่งได้ค่าพารามิเตอร์ที่ต้องการ

3.4 วิธีการประมาณแบบระยะห่างต่ำสุด

ตัวสถิติของคราเมอร์-วอนไมส์ของการแจกแจงทวินามลบแบบทั่วไป แสดงได้ดังนี้

$$\min K = \sum_{x=0}^{k}\left[F_n(x)-F(x:\varphi)\right]^2$$

เมื่อ φ เป็นพารามิเตอร์ใด ๆ

การประมาณค่าพารามิเตอร์ φ ทำได้โดยหาอนุพันธ์บางส่วนของตัวสถิติของคราเมอร์-วอนไมส์เทียบกับ φ และให้สมการอนุพันธ์บางส่วนเท่ากับ 0

$$\frac{\partial K}{\partial \varphi_i} = -2\sum_{x=0}^{k}[F_n(x)-F(x:\varphi)]\frac{\partial F(x)}{\partial \varphi_i} \quad ; \quad i=1,2,3$$

เนื่องจากสมการการหาอนุพันธ์บางส่วนของตัวสถิติของคราเมอร์-วอน ไมส์ของการแจกแจงทวินามลบแบบทั่วไปไม่อยู่ในรูปของสมการเชิงเส้น จึงไม่สามารถแก้สมการได้โดยตรง ดังนั้นจำเป็นต้องใช้วิธีการประมาณ ในการวิจัยครั้งนี้เลือกใช้วิธีนิวตัน-ราฟสัน โดยมีขั้นตอนการคำนวนดังนี้

ขั้นตอนที่ 1 การคำนวนค่า $g_1, g_2, g_3, g_{11}, g_{12}, g_{13}, g_{21}, g_{22}, g_{23}, g_{31}, g_{32}, g_{33}$ ซึ่งสามารถแสดงได้ดังนี้

$$g_1(\theta,\beta,m) = \frac{\partial K}{\partial \theta} = -2\sum_{x=0}^{k}[F_n(x)-F(x)]\frac{\partial F(x)}{\partial \theta}$$

$$g_2(\theta,\beta,m) = \frac{\partial K}{\partial \beta} = -2\sum_{x=0}^{k}[F_n(x)-F(x)]\frac{\partial F(x)}{\partial \beta}$$

$$g_3(\theta,\beta,m) = \frac{\partial K}{\partial m} = -2\sum_{x=0}^{k}[F_n(x)-F(x)]\frac{\partial F(x)}{\partial m}$$

$$g_{11}(\theta,\beta,m) = \frac{\partial^2 K}{\partial \theta^2} = 2\sum_{x=0}^{k}\left[\left(\frac{\partial F(x)}{\partial \theta}\right)^2 - (F_n(x)-F(x))\frac{\partial^2 F(x)}{\partial \theta^2}\right]$$

$$g_{12}(\theta,\beta,m) = \frac{\partial^2 K}{\partial \beta \partial \theta} = 2\sum_{x=0}^{k}\left[\left(\frac{\partial F(x)}{\partial \theta}\right)\left(\frac{\partial F(x)}{\partial \beta}\right) - (F_n(x)-F(x))\frac{\partial^2 F(x)}{\partial \beta \partial \theta}\right]$$

$$= g_{21}(\theta,\beta,m)$$

$$g_{13}(\theta,\beta,m) = \frac{\partial^2 K}{\partial m \partial \theta} = 2\sum_{x=0}^{k}\left[\left(\frac{\partial F(x)}{\partial \theta}\right)\left(\frac{\partial F(x)}{\partial m}\right) - (F_n(x)-F(x))\frac{\partial^2 F(x)}{\partial m \partial \theta}\right]$$

$$= g_{31}(\theta,\beta,m)$$

$$g_{22}(\theta,\beta,m) = \frac{\partial^2 K}{\partial \beta^2} = 2\sum_{x=0}^{k}\left[\left(\frac{\partial F(x)}{\partial \beta}\right)^2 - (F_n(x)-F(x))\frac{\partial^2 F(x)}{\partial \beta^2}\right]$$

$$g_{23}(\theta,\beta,m) = \frac{\partial^2 K}{\partial m \partial \beta} = 2\sum_{x=0}^{k}\left[\left(\frac{\partial F(x)}{\partial \beta}\right)\left(\frac{\partial F(x)}{\partial m}\right) - (F_n(x) - F(x))\frac{\partial^2 F(x)}{\partial m \partial \beta}\right]$$

$$= g_{32}(\theta,\beta,m)$$

$$g_{33}(\theta,\beta,m) = \frac{\partial^2 K}{\partial m^2} = 2\sum_{x=0}^{k}\left[\left(\frac{\partial F(x)}{\partial m}\right)^2 - (F_n(x) - F(x))\frac{\partial^2 F(x)}{\partial m^2}\right]$$

เมื่อ θ , β , *m* เป็นพารามิเตอร์ของการแจกแจงทวินามลบแบบทั่วไป

สมการการหาอนุพันธ์อันดับหนึ่งและสองเมื่อเทียบกับพารามิเตอร์แต่ละตัวจะแสดงไว้ในบทที่ 2

ขั้นตอนที่ 2 คำนวนค่าประมาณพารามิเตอร์ θ , β , *m* โดยคำนวนจากสมการดังนี้

$$\begin{bmatrix}\theta\\ \beta\\ m\end{bmatrix} = \begin{bmatrix}\theta_0\\ \beta_0\\ m_0\end{bmatrix} - \begin{bmatrix} g_{11}(\theta_0,\beta_0,m_0) & g_{12}(\theta_0,\beta_0,m_0) & g_{13}(\theta_0,\beta_0,m_0)\\ g_{21}(\theta_0,\beta_0,m_0) & g_{22}(\theta_0,\beta_0,m_0) & g_{23}(\theta_0,\beta_0,m_0)\\ g_{31}(\theta_0,\beta_0,m_0) & g_{32}(\theta_0,\beta_0,m_0) & g_{33}(\theta_0,\beta_0,m_0)\end{bmatrix}^{-1}\begin{bmatrix} g_1(\theta_0,\beta_0,m_0)\\ g_2(\theta_0,\beta_0,m_0)\\ g_3(\theta_0,\beta_0,m_0)\end{bmatrix}$$

เมื่อ θ_0 , β_0 , m_0 เป็นค่าพารามิเตอร์เริ่มต้นของ θ , β , *m*

จากนั้นจะใช้กระบวนการทำซ้ำโดยให้ค่าประมาณพารามิเตอร์ที่ได้ในรอบนี้เป็นค่าพารามิเตอร์เริ่มต้นในรอบถัดไป และกลับไปขั้นตอนที่1 และ 2 ตามลำดับ จนกระทั่งได้ค่าพารามิเตอร์ที่ต้องการ

4. การหาค่าความคลาดเคลื่อนจากการประมาณ โดยนำค่าประมาณพารามิเตอร์มาเปรียบเทียบกับพารามิเตอร์จริง เพื่อคำนวนหาค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error : MSE) เนื่องจากในการทดลองได้จำลองข้อมูลซ้ำ ๆ กันจำนวน 1,000 ครั้ง ในแต่ละสถานการณ์ ให้ i แทนจำนวนรอบที่ทำซ้ำ i = 1,2,...,1000 เนื่องจากพารามิเตอร์ที่ต้องการประมาณมี 3 พารามิเตอร์ ดังนั้นจะต้องหาค่าเฉลี่ยของค่า MSE ของการประมาณทั้ง 3 พารามิเตอร์ การหาค่าเฉลี่ยความคลาดเคลื่อนกำลังสองคำนวนได้จากสูตรดังนี้

$$MSE(\hat{\theta}_p) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}_{pi} - \theta_p)^2 \quad ; \quad p = 1,2,3$$

$$MSE = \frac{MSE(\hat{\theta}_1) + MSE(\hat{\theta}_2) + MSE(\hat{\theta}_3)}{3}$$

การหาค่าความแปรปรวนของการประมาณคำนวนได้จากสูตรดังนี้

$$Var(\hat{\theta}_p) = \frac{1}{n}\sum_{i=1}^{n}(\hat{\theta}_{pi} - \hat{\bar{\theta}}_p)^2 \quad ; \quad p = 1,2,3$$

$$Varme = \frac{Var(\hat{\theta}_1) + Var(\hat{\theta}_2) + Var(\hat{\theta}_3)}{3}$$

โดยที่ Varme เป็นค่าเฉลี่ยของความแปรปรวนของทั้ง 3 พารามิเตอร์

การหาค่าความเอนเอียงของการประมาณคำนวนได้จากสูตรดังนี้

$$Bias(\hat{\theta}_p) = \hat{\bar{\theta}}_p - \theta_p \quad ; \quad p = 1,2,3$$

$$Biasme = \frac{\left|Bias(\hat{\theta}_1)\right| + \left|Bias(\hat{\theta}_2)\right| + \left|Bias(\hat{\theta}_3)\right|}{3}$$

โดยที่ Biasme เป็นค่าเฉลี่ยของความเอนเอียงสัมบูรณ์ของทั้ง 3 พารามิเตอร์

จากนั้นจึงนำไปเปรียบเทียบเพื่อหาว่าวิธีการใดให้ค่า MSE ของการประมาณค่าพารามิเตอร์ต่ำที่สุดจะเป็นวิธีที่ประมาณค่าพารามิเตอร์ของแต่ละสถานการณ์ได้ดีที่สุด

3.4 **โปรแกรมที่ใช้ในการวิจัย**

โปรแกรมที่ใช้ในการวิจัยทั้งหมด เขียนด้วยภาษาฟอร์แทรน (FORTRAN) โดยใช้กับเครื่อง AMDAHL 5860 ซึ่งในแต่ละสถานการณ์ของการทดลอง ลักษณะการทำงานของโปรแกรมจะเหมือนกัน สำหรับรายละเอียดของโปรแกรมจะแสดงไว้ในภาคผนวก ข ซึ่งจะเป็นโปรแกรมการทำงานของแต่ละวิธีการ คือ วิธีการประมาณแบบความควรจะเป็นสูงสุด วิธีการประมาณแบบสองโมเมนต์แรกและสัดส่วนศูนย์ วิธีการประมาณแบบไคกำลังสองต่ำสุด และวิธีการประมาณแบบระยะห่างต่ำสุด

รูปที่ 3.1 แสดงผังงานสำหรับหาค่าความคลาดเคลื่อนจากการประมาณด้วยวิธีการทั้ง 4 วิธี