บทที่ 3

วิธีการศึกษา

ในการศึกษาครั้งนี้มีวัตถุประสงค์หลักเพื่อพัฒนาเทคนิคการคัดเลือกตัวแบบโดยใช้ค่าการลงจุดตำแหน่งและประยุกต์ใช้ในการหาการแจกแจงที่เหมาะสมสำหรับปริมาณฝนตามฤดูกาลในลุ่มน้ำปิงตอนบน ดังนั้นในบทนี้จะกล่าวถึงขอบเขตของการจำลองขอมูล สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนาด้วยค่าการลงจุดตำแหน่ง การประมาณค่าพารามิเตอร์ รวมถึงขั้นตอนดำเนินการศึกษา

3.1 ขอบเขตของการจำลองข้อมูล

กำหนดสถานการณ์ต่างๆ ในการจำลองข้อมูล ดังนี้

3.1.1 ลักษณะการแจกแจงของข้อมูล

ในการศึกษาครั้งนี้ ทำการจำลองข้อมูลที่มีการแจกแจงเบ้ขวา โดย

1) การแจกแจงแกมมา

สามารถคำนวณหาค่าสัมประสิทธิ์ความเบ้ได้โดย

$$\gamma_X = \frac{2}{\sqrt{\alpha}}$$

2) การแจกแจงกัมเบล

มีค่าสัมประสิทธิ์ความเบ้เป็นค่าคงที่

$$\gamma_x = 1.1396$$

3) การแจกแจงค่าสุดขีดวางนัยทั่วไป

สามารถคำนวณหาค่าสัมประสิทธิ์ความเบ้ได้โดย

$$\gamma_x = Sign(\alpha)\frac{3\Gamma(1+\alpha)\Gamma(1+2\alpha)-\Gamma(1+3\alpha)-2\Gamma^3(1+\alpha)}{\left[\Gamma(1+2\alpha)-\Gamma^2(1+\alpha)\right]^{3/2}}$$

4) การแจกแจงล็อกนอร์มอล

สามารถคำนวณหาค่าสัมประสิทธิ์ความเบ้ได้โดย

$$\gamma_x = \sqrt{\exp(\alpha^2)-1}\ \exp(\alpha^2)+2$$

5) การแจกแจงเพียร์สันประเภทที่ 3

สามารถคำนวณหาค่าสัมประสิทธิ์ความเบ้ได้โดย

$$\gamma_X = \frac{2\beta}{|\beta|\sqrt{\alpha}}$$

6) การแจกแจงรีเวิร์สไวบูล

สามารถคำนวณหาค่าสัมประสิทธิ์ความเบ้ได้โดย

$$\gamma_x = \frac{\Gamma\left(1+\frac{3}{\alpha}\right)-3\Gamma\left(1+\frac{1}{\alpha}\right)\Gamma\left(1+\frac{2}{\alpha}\right)+2\Gamma^3\left(1+\frac{1}{\alpha}\right)}{\left[\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma^2\left(1+\frac{1}{\alpha}\right)\right]^{3/2}}$$

3.1.2 ขนาดตัวอย่าง

ในแต่ละการแจกแจงกำหนดขนาดตัวอย่างเท่ากับ 20, 30, 50 และ 100 ซึ่งในที่นี้ได้กำหนดขนาดตัวอย่างให้ใกล้เคียงกับจำนวนข้อมูลจริง นั่นคือข้อมูลปริมาณฝนตามฤดูกาลในลุ่มน้ำปิงตอนบน

3.1.3 การกำหนดระดับนัยสำคัญ

ในการทดสมสมมติฐานได้กำหนดระดับนัยสำคัญ (α) เท่ากับ 0.05

3.2 สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนาโดยการใช้ค่าการลงจุดตำแหน่ง

สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งมีหลักการคือการวัดระยะห่างระหว่างฟังก์ชันการแจกแจงเชิงประจักษ์ $(F_n(x))$ กับการแจกแจงที่คาดหมาย $(F(x))$ ดังนี้

$$AD = n\int_{-\infty}^{\infty}\frac{[F_n(x)-F(x)]^2}{[F(x)\{1-F(x)\}]}dF(x)$$

$$\frac{1}{n}AD = \int_{-\infty}^{\infty} \frac{[F_n(x) - F(x)]^2}{F(x)\{1 - F(x)\}} dF(x)$$

$$= \int_{-\infty}^{x_1} \frac{F^2(x)}{F(x)\{1 - F(x)\}} dF(x) + \int_{x_1}^{x_2} \frac{[F_n(x) - F(x)]^2}{F_n(x)\{1 - F(x)\}} dF(x) + \ldots + \int_{x_n}^{\infty} \frac{\{1 - F(x)\}^2}{F(x)\{1 - F(x)\}} dF(x)$$

จาก Reimann-Stieltjes Integral จะได้

$$= \sum_{i=1}^{n} \left(-\ln(1 - \min(F(x))) - \ln(\max(F(x))) - 1\right)$$

$$AD = \left[\sum_{i=1}^{n} \left(\frac{2i-1}{n}\right)\left(\ln(F(x)) + \ln(1 - F(x_{n-i+1}))\right)\right] - n$$

$$= -n - \frac{1}{n}\sum_{i=1}^{n} (2i-1)\left[\log F(x_i) + \log[1 - F(x_{n+1-i})]\right]$$

เมื่อ $(2i-1)$ คือ ค่าคงที่ที่ได้จาก Riemann – Stieltjes Integral

i คือ ข้อมูลลำดับที่ i ; $i = 1, 2, 3, \ldots$

ในส่วนนี้ผู้วิจัยได้นำเสนอสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งใหม่ซึ่งพัฒนามาจากสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง โดยให้ความสนใจในการเปลี่ยนค่าคงที่ที่ได้จาก Riemann – Stieltjes Integral โดยใช้แนวคิดของค่าการลงจุดตำแหน่ง ดังนี้

3.2.1 ค่าการลงจุดตำแหน่งของ Weibull (Weibull, 1939)
เป็นค่าการลงจุดตำแหน่งที่มีการใช้งานอย่างแพร่หลาย โดยกำหนดให้ $\alpha = 0$ เป็นค่าการลงจุดตำแหน่งที่เหมาะสำหรับทุกการแจกแจง มีสูตรทั่วไปดังนี้

$$P_i = \frac{i}{n+1}$$

จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนา ดังนี้

$$AD_1 = -n - \sum_{i=1}^{n} \left(\frac{i}{n+1}\right)\left[\log F(x_i) + \log[1 - F(x_{n-i+1})]\right]$$

$$AU_1 = \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n} \left\{2 - \frac{i}{n+1}\right\} \log\{1 - F(x_i)\}$$

$$AL_1 = -\frac{3n}{2} + 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n} \frac{i}{n+1} \log\{F(x_i)\}$$

3.2.2 ค่าการลงจุดตำแหน่งของ Beard (Beard, 1943)

กำหนดให้ $\alpha = 0.3175$ เป็นค่าการลงจุดตำแหน่งที่เหมาะสำหรับการแจกแจงปกติ มีสูตรทั่วไปดังนี้

$$P_i = \frac{i - 0.3175}{n + 0.365}$$

จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนา ดังนี้

$$AD_2 = -n - \sum_{i=1}^{n} \left(\frac{i - 0.3175}{n + 0.365} \right) \left[\log F(x_i) + \log[1 - F(x_{n-i+1})] \right]$$

$$AU_2 = \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n} \left\{ 2 - \frac{i - 0.3175}{n + 0.365} \right\} \log\{1 - F(x_i)\}$$

$$AL_2 = -\frac{3n}{2} + 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n} \frac{i - 0.3175}{n + 0.365} \log\{F(x_i)\}$$

3.2.3 ค่าการลงจุดตำแหน่งของ Blom (Blom, 1958)

กำหนดให้ $\alpha = 0.375$ เป็นค่าการลงจุดตำแหน่งที่เหมาะสำหรับการแจกแจงปกติ มีสูตรทั่วไปดังนี้

$$P_i = \frac{i - 0.375}{n + 0.25}$$

จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนา ดังนี้

$$AD_3 = -n - \sum_{i=1}^{n} \left(\frac{i - 0.375}{n + 0.25} \right) \left[\log F(x_i) + \log[1 - F(x_{n-i+1})] \right]$$

$$AU_3 = \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n} \left\{ 2 - \frac{i - 0.375}{n + 0.25} \right\} \log\{1 - F(x_i)\}$$

$$AL_3 = -\frac{3n}{2} + 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n} \frac{i - 0.375}{n + 0.25} \log\{F(x_i)\}$$

3.2.4 ค่าการลงจุดตำแหน่งของ Cunnane (Cunnane, 1977)

กำหนดให้ $\alpha = 0.40$ เป็นค่าการลงจุดตำแหน่งที่เหมาะสำหรับการแจกแจงค่าสุดขีดวางนัยทั่วไป และการแจกแจงเพียร์สันประเภทที่ 3 มีสูตรทั่วไปดังนี้

$$P_i = \frac{i-0.40}{n+0.20}$$

จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนา ดังนี้

$$AD_4 = -n - \sum_{i=1}^{n}\left(\frac{i-0.40}{n+0.20}\right)\left[\log F(x_i) + \log\left[1 - F(x_{n-i+1})\right]\right]$$

$$AU_4 = \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n}\left\{2 - \frac{i-0.40}{n+0.20}\right\}\log\{1 - F(x_i)\}$$

$$AL_4 = -\frac{3n}{2} + 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n}\frac{i-0.40}{n+0.20}\log\{F(x_i)\}$$

3.2.5 ค่าการลงจุดตำแหน่งของ Gringorten (Gringorten, 1963)

กำหนดให้ $\alpha = 0.44$ เป็นค่าการลงจุดตำแหน่งที่เหมาะสำหรับการแจกแจงเลขชี้กำลัง การแจกแจงกัมเบลและการแจกแจงค่าสุดขีดวางนัยทั่วไป มีสูตรทั่วไปดังนี้

$$P_i = \frac{i-0.44}{n+0.12}$$

จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนา ดังนี้

$$AD_5 = -n - \sum_{i=1}^{n}\left(\frac{i-0.44}{n+0.12}\right)\left[\log F(x_i) + \log\left[1 - F(x_{n-i+1})\right]\right]$$

$$AU_5 = \frac{n}{2} - 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n}\left\{2 - \frac{i-0.44}{n+0.12}\right\}\log\{1 - F(x_i)\}$$

$$AL_5 = -\frac{3n}{2} + 2\sum_{i=1}^{n} F(x_i) - \sum_{i=1}^{n}\frac{i-0.44}{n+0.12}\log\{F(x_i)\}$$

3.3 การประมาณค่าพารามิเตอร์

ประมาณค่าพารามิเตอร์ด้วยวิธี $L-moment$ ซึ่งคำนวณได้จากสูตรดังต่อไปนี้

3.3.1 การแจกแจงแกมมา

$$\hat{\alpha} = \left(1 - \frac{\lambda_2}{\lambda_1}\right)\left(\frac{0.7213 + \left(1 - \frac{\lambda_2}{\lambda_1}\right)(-0.5947)}{1 + \left(1 - \frac{\lambda_2}{\lambda_1}\right)\left(-2.1817 + \left(1 - \frac{\lambda_2}{\lambda_1}\right)(1.2113)\right)}\right)$$

$$\hat{\beta} = \lambda_2 / \hat{\alpha}$$

3.3.2 การแจกแจงกัมเบล

$$\hat{\beta} = \frac{\lambda_2}{\log(2)}$$

$$\hat{\xi} = \lambda_1 - (\beta\gamma)$$

เมื่อ γ คือ Euler's Constant $(\gamma = 0.5772)$

3.3.3 การแจกแจงล็อกนอร์มอล

$$\hat{\alpha} = \left(-\frac{\lambda_3}{\lambda_2}\right)\left(\frac{2.0466543 + (-3.654437)\left(\frac{\lambda_3}{\lambda_2}\right)^2 + (1.8396733)\left(\frac{\lambda_3}{\lambda_2}\right)^4 + (-0.20360244)\left(\frac{\lambda_3}{\lambda_2}\right)^6}{1 + (-0.20182173)\left(\frac{\lambda_3}{\lambda_2}\right)^2 + (1.2420401)\left(\frac{\lambda_3}{\lambda_2}\right)^4 + (-0.21741801)\left(\frac{\lambda_3}{\lambda_2}\right)^6}\right)$$

$$\hat{\beta} = \frac{\lambda_2 \hat{\alpha} \exp[-\hat{\alpha}/2]}{1 - 2\Phi(-\hat{\alpha}/\sqrt{2})}$$

$$\hat{\xi} = \lambda_1 - \left(\frac{\hat{\beta}}{\hat{\alpha}}\right)\left(1 - \exp[\hat{\alpha}^2/2]\right)$$

3.3.4 การแจกแจงเพียร์สันประเภทที่ 3

$$\hat{\xi} = \lambda_1 - \hat{\alpha}\hat{\beta}$$

$$\hat{\beta} = \frac{\lambda_2 \sqrt{\pi}\,\Gamma(\hat{\alpha})}{\Gamma(\hat{\alpha} + 0.5)}$$

$$\hat{\alpha} = \begin{bmatrix} \dfrac{1 + 0.2906(3\pi\tau_3^2)}{(3\pi\tau_3^2) + 0.1882(3\pi\tau_3^2)^2 + 0.0442(3\pi\tau_3^2)} & \textit{for } 0 < |\tau_3| < 1/3 \\ \dfrac{0.36067(1-|\tau_3|) - 0.59567(1-|\tau_3|)^2 + 0.25361(1-|\tau_3|)^3}{1 - 2.78861(1-|\tau_3|) + 2.56096(1-|\tau_3|)^2 - 0.77045(1-|\tau_3|)^3} & \textit{for } 1/3 \le |\tau_3| < 1 \end{bmatrix}$$

3.3.5 การแจกแจงค่าสุดขีดวางนัยทั่วไป

$$\hat{\xi} = \lambda_1 - \frac{\hat{\beta}}{\hat{\alpha}}\left(1 - \Gamma(1 + \hat{\alpha})\right)$$

$$\hat{\beta} = \frac{\lambda_2 \hat{\alpha}}{\left(1 - 2^{-\hat{\alpha}}\right)\Gamma\left(1 + \hat{\alpha}\right)}$$

$$\hat{\alpha} = 7.859c + 2.9554c^2$$

เมื่อ $c = \frac{2}{\left(3 + \tau_3\right)} - \frac{\log(2)}{\log(3)}$

3.3.6 การแจกแจงรีเวิร์สไวบูล

$$\hat{\xi} = \lambda_1 - \hat{\beta}\Gamma\left(\frac{1}{\hat{\alpha}} + 1\right)$$

$$\hat{\beta} = \frac{\lambda_2}{\Gamma\left(\frac{1}{\hat{\alpha}} + 1\right)\left(1 - \frac{1}{2^{1/\hat{\alpha}}}\right)}$$

$$\hat{\alpha} = 7.859c + 2.9554c^2$$

เมื่อ $c = \frac{2}{\left(3 + \tau_3\right)} - \frac{\log(2)}{\log(3)}$

3.4 ขั้นตอนการดำเนินการศึกษา

3.4.1 การจำลองข้อมูล

ขั้นตอนการดำเนินการศึกษาสำหรับการจำลองข้อมูลในสถานการณ์ต่างๆแบ่งออกเป็น 3 ขั้นตอนได้แก่

ขั้นตอนที่ 1 การสร้างค่าวิกฤตสำหรับทดสอบการแจกแจง ดังนี้

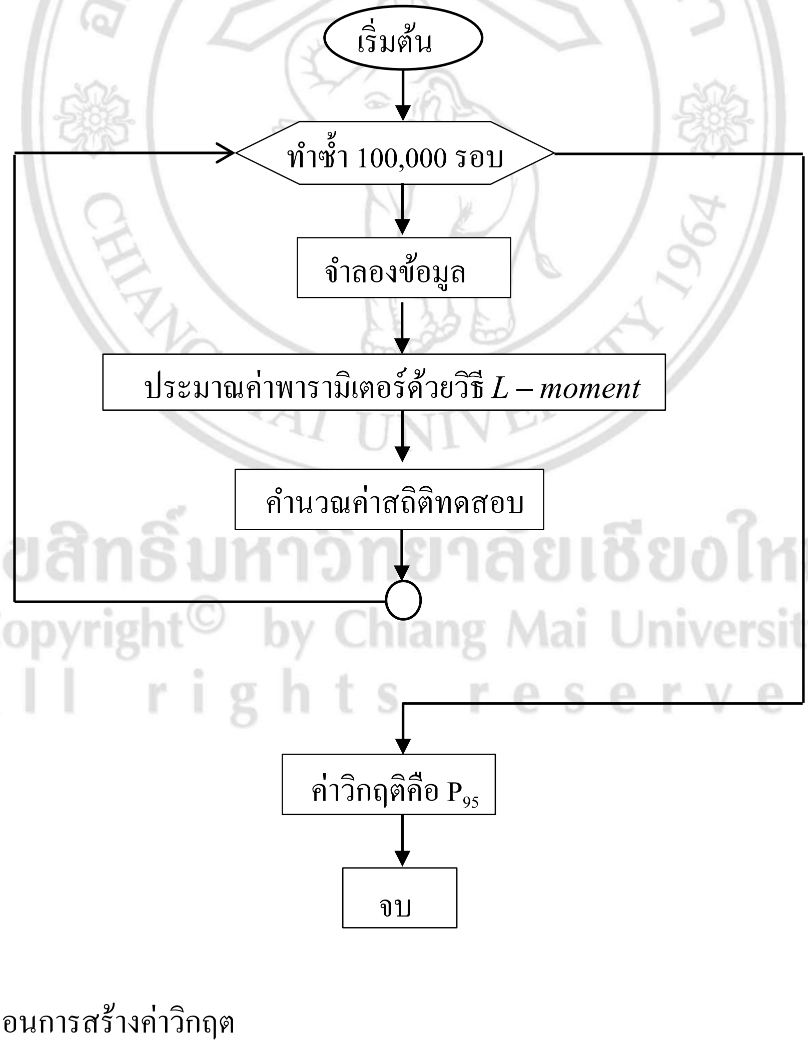
1) จำลองที่มีการแจกแจงแกมมาที่ค่า $\beta = 50$ และ $\alpha = 1, 2, 5, 10, 500$
2) ประมาณค่าพารามิเตอร์ของแต่ละการแจกแจงด้วยวิธี $L - moment$ จากนั้นนำไปแทนค่าในฟังก์ชันการแจกแจง
3) คำนวณค่าสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ปรับปรุงของ Ahmad สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนาโดยการใช้ค่าการลงจุดตำแหน่ง และค่าเกณฑ์การทดสอบ AIC และ BIC
4) ทำซ้ำข้อ 1) ถึง 3) จำนวน 100,000 รอบ
5) นำสถิติทดสอบแต่ละวิธีที่คำนวณได้ในข้อ 4) มาเรียงลำดับเพื่อหาตำแหน่งเปอร์เซ็นไทล์ที่ 95 ซึ่งเป็นค่าวิกฤตที่ระดับนัยสำคัญ 0.05 ในแต่ละสถานการณ์

6) ทำซ้ำข้อ 1) ถึง 5) จนกระทั่งครบทุกพารามิเตอร์และทุกขนาดตัวอย่าง

7) นำค่าวิกฤตในข้อ 6) มาหาค่าเฉลี่ยในทุกค่าพารามิเตอร์ของแต่ละขนาดตัวอย่าง เพื่อให้เป็นตัวแทนของค่าวิกฤตในขนาดตัวอย่างนั้นๆ

8) สำหรับการแจกแจงอื่นก็ทำในลักษณะเดียวกัน โดยการแจกแจงกัมเบลจำลองที่ค่า $\xi = 200$ และ $\beta = 1,10,30,50,85$ การแจกแจงล็อกนอร์มอลจำลองที่ค่า $\beta = 5.5$ $\xi = -20$ $\alpha = 0.01, 0.2, 0.4, 0.6, 1$ การแจกแจงค่าสุดขีดวางนัยทั่วไปจำลองที่ค่า $\beta = 62, \xi = 200$ และ $\alpha = -0.2, -0.1, 0.1, 0.2, 0.3$ การแจกแจงรีเวิร์สไวบูลจำลองที่ค่า $\beta = 450, \xi = 600$ และ $\alpha = 4, 6, 20, 50, 100$ การแจกแจงเพียร์สันประเภทที่ 3 จำลองที่ค่า $\beta = 55, \xi = 30$ และ $\alpha = 1, 2, 5, 10, 500$

รายละเอียดขั้นตอนการจำลองข้อมูลแสดงดังภาพต่อไปนี้



ภาพที่ 3.1 ขั้นตอนการสร้างค่าวิกฤต

ขั้นตอนที่ 2 การคำนวณค่าความผิดพลาดเชิงประจักษ์แบบที่ 1

โดยกำหนดสมมติฐานของการทดสอบ ดังนี้

H_0 : ข้อมูลที่จำลองมาจากประชากรที่มีการแจกแจงที่สนใจศึกษา

H_1 : ข้อมูลที่จำลองไม่ได้มาจากประชากรที่มีการแจกแจงที่สนใจศึกษา

1) จำลองที่มีการแจกแจงแกมมาที่ค่า $\beta = 50$ และ $\alpha = 2, 5, 10$
2) ประมาณค่าพารามิเตอร์ของแต่ละการแจกแจงด้วยวิธี $L-moment$ จากนั้นนำไปแทนค่าในฟังก์ชันการแจกแจง
3) คำนวณค่าสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกปรับปรุงของ Ahmad สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนาโดยการใช้ค่าการลงจุดตำแหน่ง และค่าเกณฑ์การทดสอบ AIC และ BIC
4) เปรียบเทียบค่าสถิติทดสอบที่ได้ในข้อ 3) กับค่าวิกฤตที่ขนาดตัวอย่างและการแจกแจงที่สนใจแล้วพิจารณาผลการทดสอบสมมติฐาน ดังนี้

$h = 0$ เมื่อปฏิเสธสมมติฐานหลัก

$h = 1$ เมื่อไม่มีเหตุผลเพียงพอที่จะปฏิเสธสมมติฐานหลัก

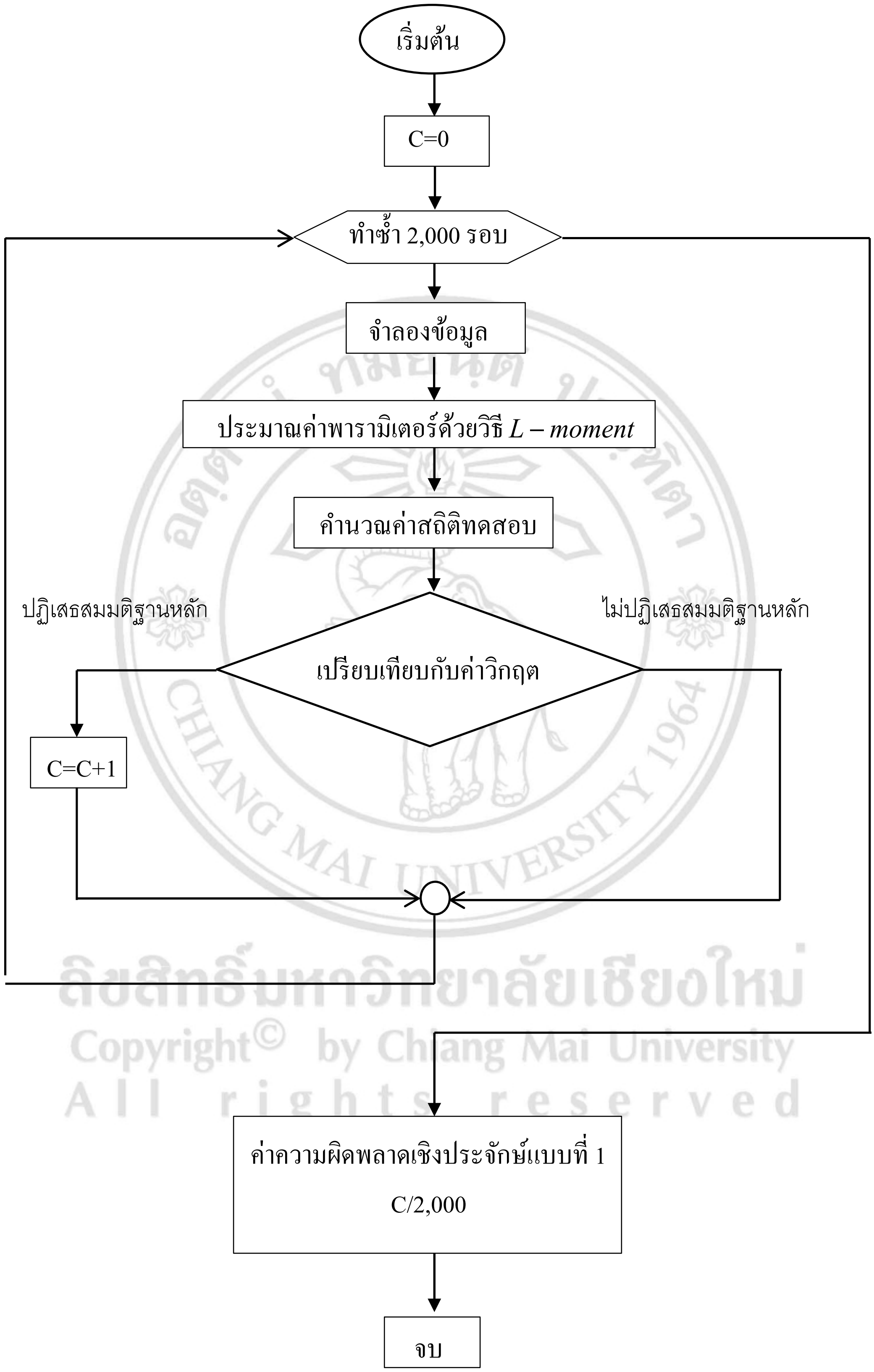
5) ทำซ้ำ 1) ถึง 4) เป็นจำนวน 2,000 รอบ
6) คำนวณหาค่าความผิดพลาดเชิงประจักษ์แบบที่ 1 ได้โดยการนำจำนวนครั้งในการปฏิเสธสมมติฐานหลักหารด้วยจำนวนรอบที่ทำซ้ำ 2,000 รอบ
7) นำค่าความผิดพลาดเชิงประจักษ์แบบที่ 1 $(\hat{\alpha})$ เทียบกับเกณฑ์การพิจารณาความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ของ Cochran (Cochran, 1954) และ Bradley (Bradley, 1978) โดยที่ที่ระดับนัยสำคัญ 0.05

เกณฑ์ของ Cochran ค่า $\hat{\alpha}$ จะอยู่ในช่วง [0.040 , 0.060]

เกณฑ์ของ Bradley ค่า $\hat{\alpha}$ จะอยู่ในช่วง [0.025 , 0.075]

8) สำหรับการแจกแจงอื่นก็ทำในลักษณะเดียวกัน โดยการแจกแจงกัมเบลจำลองที่ค่า $\xi = 200$ และ $\beta = 1, 10, 50$ การแจกแจงล็อกนอร์มอลจำลองที่ค่า $\beta = 5.5$ $\xi = -20$ $\alpha =, 0.2, 0.4, 0.6$ การแจกแจงค่าสุดขีดวางนัยทั่วไปจำลองที่ค่า $\beta = 62, \xi = 200$ และ $\alpha = -0.1, 0.1, 0.2$ การแจกแจงรีเวิร์สไวบูลจำลองที่ค่า $\beta = 450, \xi = 600$ และ $\alpha = 6, 20, 50$ การแจกแจงเพียร์สันประเภทที่ 3 จำลองที่ค่า $\beta = 55, \xi = 30$ และ $\alpha = 2, 5, 10$

รายละเอียดขั้นตอนการจำลองข้อมูลแสดงดังภาพต่อไปนี้



ภาพที่ 3.2 ขั้นตอนการคำนวณค่าผิดพลาดเชิงประจักษ์แบบที่ 1

ขั้นตอนที่ 3 การคำนวณอำนาจการทดสอบเชิงประจักษ์

การคำนวณอำนาจการทดสอบจะพิจารณาเฉพาะสถิติทดสอบที่มีค่าความผิดพลาดเชิงประจักษ์แบบที่ 1 อยู่ในช่วงเกณฑ์พิจารณาการพิจารณาความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ของ Cochran และ Bradley โดยกำหนดพารามิเตอร์ของแต่ละการแจกแจงดังนี้

การแจกแจงแกมมา จำลองที่ค่า $\alpha = 5$ และ $\beta = 50$

การแจกแจงกัมเบล จำลองที่ค่า $\beta = 10$ และ $\xi = 200$

การแจกแจงล็อกนอร์มอล จำลองที่ค่า $\alpha = 0.4, \beta = 1$ และ $\xi = 20$

การแจกแจงค่าสุดขีดวางนัยทั่วไป จำลองที่ค่า $\alpha = 0.2, \beta = 62$ และ $\xi = 200$

การแจกแจงรีเวิร์สไวบูล จำลองที่ค่า $\alpha = 6$, $\beta = 450$ และ $\xi = 600$

การแจกแจงเพียร์สันประเภทที่ 3 จำลองที่ค่า $\alpha = 10, \beta = 55$ และ $\xi = 30$

โดยกำหนดสมมติฐานของการทดสอบ ดังนี้

H_0 : ข้อมูลที่จำลองมาจากประชากรที่มีการแจกแจงที่สนใจศึกษา

H_1 : ข้อมูลที่จำลองไม่ได้มาจากประชากรที่มีการแจกแจงที่สนใจศึกษา

1) จำลองข้อมูลที่มีการแจกแจงดังนี้

1.1) เมื่อทดสอบการแจกแจงแกมมา ให้จำลองข้อมูลที่มีการแจกแจงกัมเบล การแจกแจงล็อกนอร์มอล การแจกแจงค่าสุดขีดวางนัยทั่วไป การแจกแจงรีเวิร์สไวบูลและการแจกแจงเพียร์สันประเภทที่ 3

1.2) เมื่อทดสอบการแจกแจงกัมเบล ให้จำลองข้อมูลที่มีการแจกแจงแกมมา การแจกแจงล็อกนอร์มอล การแจกแจงค่าสุดขีดวางนัยทั่วไป การแจกแจงรีเวิร์สไวบูลและการแจกแจงเพียร์สันประเภทที่ 3

1.3) เมื่อทดสอบการแจกแจงล็อกนอร์มอล ให้จำลองข้อมูลที่มีการแจกแจงกัมเบล การแจกแจงแกมมา การแจกแจงค่าสุดขีดวางนัยทั่วไป การแจกแจงรีเวิร์สไวบูลและการแจกแจงเพียร์สันประเภทที่ 3

1.4) เมื่อทดสอบการแจกแจงค่าสุดขีดวางนัยทั่วไป ให้จำลองข้อมูลที่มีการแจกแจงกัมเบล การแจกแจงล็อกนอร์มอล การแจกแจงแกมมา การแจกแจงรีเวิร์สไวบูลและการแจกแจงเพียร์สันประเภทที่ 3

1.5) เมื่อทดสอบการแจกแจงรีเวิร์สไวบูล ให้จำลองข้อมูลที่มีการแจกแจงกัมเบล การแจกแจงล็อกนอร์มอล การแจกแจงค่าสุดขีดวางนัยทั่วไป การแจกแจงแกมมาและการแจกแจงเพียร์สันประเภทที่ 3

1.6) เมื่อทดสอบการแจกแจงเพียร์สันประเภทที่ 3 ให้จำลองข้อมูลที่มีการแจกแจงกัมเบล การแจกแจงล็อกนอร์มอล การแจกแจงค่าสุดขีดวางนัยทั่วไป การแจกแจงรีเวิร์สไวบูลและการแจกแจงแกมมา

2) ประมาณค่าพารามิเตอร์ของแต่ละการแจกแจงด้วยวิธี $L-moment$ จากนั้นนำไปแทนค่าในฟังก์ชันการแจกแจง

3) คำนวณค่าสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนาโดยการใช้ค่าการลงจุดตำแหน่ง และค่าเกณฑ์การทดสอบ AIC และ BIC

4) เปรียบเทียบค่าสถิติทดสอบที่ได้ในข้อ 3) กับค่าวิกฤตที่ขนาดตัวอย่างและการแจกแจงที่สนใจแล้วพิจารณาผลการทดสอบสมมติฐาน ดังนี้

$h=0$ เมื่อปฏิเสธสมมติฐานหลัก

$h=1$ เมื่อไม่มีเหตุผลเพียงพอที่จะปฏิเสธสมมติฐานหลัก

5) ทำซ้ำ 1) ถึง 4) เป็นจำนวน 2,000 รอบ

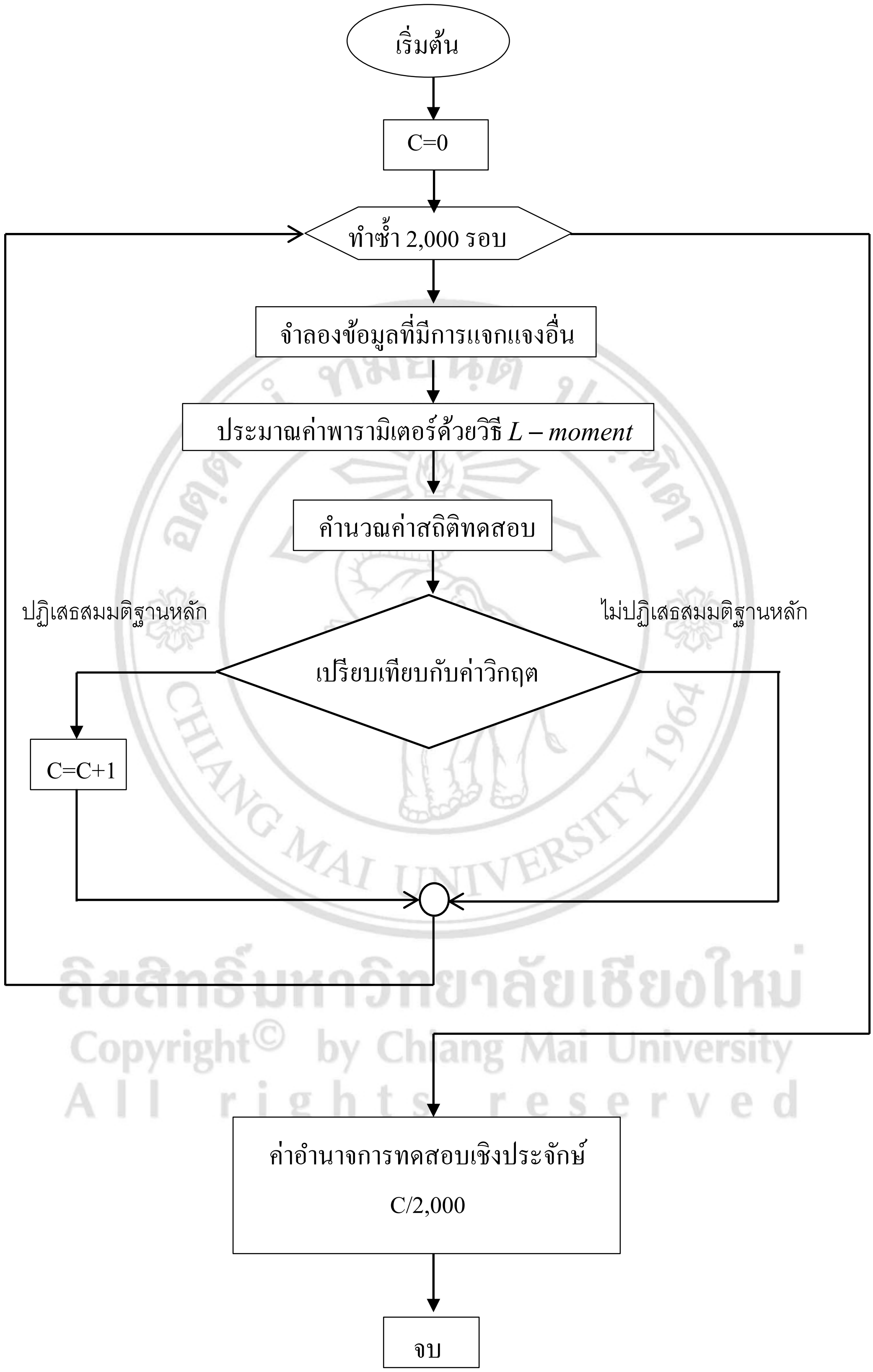
6) คำนวณหาค่าอำนาจการทดสอบ ได้โดยการนำจำนวนครั้งในการปฏิเสธสมมติฐานหลักหารด้วยจำนวนรอบที่ทำซ้ำ 2,000 รอบ

7) สำหรับเกณฑ์การทดสอบ AIC และ BIC เปรียบเทียบโดยการหาร้อยละของการตัดสินใจที่ถูกต้อง

8) นำค่าอำนาจการทดสอบที่ได้ใน 6) ของแต่ละสถิติทดสอบในขนาดตัวอย่างเดียวกันมาเทียบกัน โดยสถิติทดสอบใดมีค่าอำนาจการทดสอบสูงที่สุด นั่นคือ สถิติทดสอบนั้นมีประสิทธิภาพมากที่สุด

9) สรุปผล

รายละเอียดขั้นตอนการจำลองข้อมูลแสดงดังภาพต่อไปนี้



ภาพที่ 3.3 ขั้นตอนการคำนวณค่าอำนาจการทดสอบเชิงประจักษ์

3.4.2 การหาการแจกแจงที่เหมาะสมสำหรับปริมาณฝนตามฤดูกาล

สำหรับการหาการแจกแจงที่แท้จริงสำหรับข้อมูลปริมาณฝนตามฤดูกาลในเขตภาคเหนือตอนบนของไทยนั้น มีขั้นตอนการวิเคราะห์ใกล้เคียงกับการจำลองสถานการณ์ที่กล่าวไปข้างต้น ดังนี้

1) นำข้อมูลปริมาณฝนรายเดือนจากสถานีสำรวจปริมาณฝนของศูนย์อุทกวิทยาและบริหารน้ำภาคเหนือตอนบน กรมชลประทาน ทั้ง 9 สถานี มาทำการสะสมเป็นปริมาณฝนราย 4 เดือน (ปริมาณฝนตามฤดูกาลของประเทศไทย) โดยที่ปริมาณฝนตามฤดูกาลแบ่งตามฤดูกาลของประเทศไทยซึ่งแบ่งได้ 3 ฤดูกาล ได้แก่ ฤดูร้อนประมาณเดือนกุมภาพันธ์ถึงเดือนพฤษภาคม ฤดูฝนประมาณเดือนมิถุนายนถึงเดือนกันยายนและฤดูหนาวประมาณเดือนตุลาคมถึงเดือนมกราคมของปีถัดไป

2) ประมาณค่าพารามิเตอร์ของแต่ละการแจกแจงด้วยวิธี $L-moment$ จากนั้นนำไปแทนค่าในฟังก์ชันการแจกแจง

3) คำนวณค่าสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ปรับปรุงของ Ahmad สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ถูกพัฒนาโดยการใช้ค่าการลงจุดตำแหน่ง และค่าเกณฑ์การทดสอบ AIC และ BIC

4) เปรียบเทียบค่าสถิติทดสอบที่ได้ในข้อ 3) กับค่าวิกฤตที่การแจกแจงที่สนใจแล้วพิจารณาผลการทดสอบสมมติฐาน ดังนี้

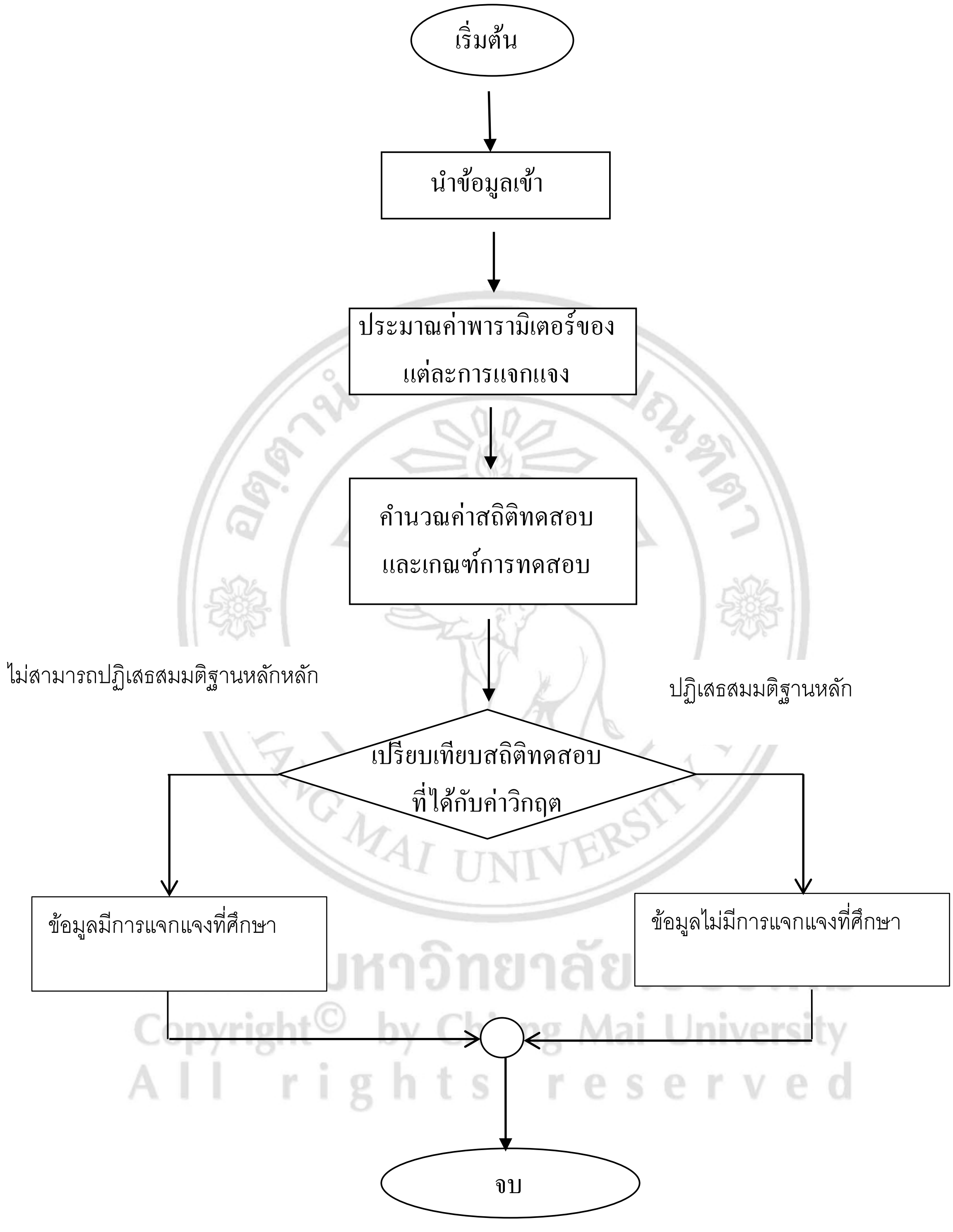
$h = 0$ เมื่อปฏิเสธสมมติฐานหลัก

$h = 1$ เมื่อไม่มีเหตุผลเพียงพอที่จะปฏิเสธสมมติฐานหลัก

5) สำหรับเกณฑ์การทดสอบ AIC และ BIC เปรียบเทียบโดยการหาร้อยละของการตัดสินใจที่ถูกต้อง

6) สรุปผล

รายละเอียดขั้นตอนการทำงานในการหาการแจกแจงที่เหมาะสมสำหรับข้อมูลปริมาณฝนตามฤดูกาลในลุ่มน้ำปิงตอนบน แสดงดังภาพดังต่อไปนี้



ภาพที่ 3.4 ขั้นตอนการทำงานในการหาการแจกแจงที่เหมาะสมสำหรับข้อมูลปริมาณฝนตามฤดูกาลในลุ่มน้ำปิงตอนบน