บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีลักษณะเป็นการวิจัยเชิงทดลองซึ่งจำลองขึ้นด้วยการทำงานของเครื่องคอมพิวเตอร์โดยใช้เทคนิควิธีการจำลองแบบมอนติคาโล (Monte Carlo simulation method) ด้วยโปรแกรมภาษาปาสคาล เพื่อหาข้อสรุปในการเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับค่าเฉลี่ยของการแจกแบบล็อกนอร์มอล 3 วิธี ซึ่งได้แก่

1. วิธีการประมาณของคอกซ์ (Cox's method)
2. วิธีการประมาณแบบคอนเซอเวทีฟ (Conservative method)
3. วิธีการประมาณแบบพาราเมตริกบูทสแทร็พ (Parametric bootstrap method)

ในการเปรียบเทียบวิธีการประมาณแบบช่วงแต่ละวิธี ขั้นตอนแรกจะทำการศึกษาค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองที่ได้จากแต่ละวิธีการประมาณก่อน แล้วจึงคัดเลือกวิธีการประมาณที่ให้ค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลองไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด จากนั้นจะหาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นของแต่ละวิธีการประมาณเพื่อเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นต่อไป

3.1 แผนการดำเนินการวิจัย

ในการทดลองครั้งนี้กำหนดสถานการณ์ต่างๆดังนี้

3.1.1 ขนาดตัวอย่าง มีค่าตั้งแต่ 5 ถึง 50

3.1.2 กำหนดค่าสัมประสิทธิ์ความเชื่อมั่น เท่ากับ 0.90, 0.95 และ 0.99

3.1.3 กำหนดพารามิเตอร์ ของการแจกแจงแบบล็อกนอร์มอลจาก C.V. ดังนี้

C.V.(%)	μ	σ^2
10	1,3	0.0100
50	1,3	0.2232
100	1,3	0.6932
150	1,3	1.1787
200	1,3	1.6095
250	1,3	1.9811
300	1,3	2.3026

3.2 ขั้นตอนในการดำเนินการวิจัย

สำหรับการดำเนินการวิจัยมีขั้นตอนดังนี้

3.2.1 สร้างข้อมูลที่ใช้ในการวิจัย

3.2.2 คำนวณช่วงความเชื่อมั่นด้วยวิธีการประมาณทั้ง 3 วิธี

3.2.3 คำนวณค่าสัมประสิทธิ์ความเชื่อมั่น และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น

3.2.4 เปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น

3.2.5 สรุปผลการวิจัยในแต่ละสถานการณ์

โดยมีรายละเอียดในแต่ละขั้นตอนดังนี้

3.2.1. <u>การสร้างข้อมูลที่ใช้ในการวิจัย</u>

การสร้างข้อมูลให้มีการแจกแจงตามลักษณะที่ต้องการศึกษานั้น จะต้องใช้เลขสุ่ม (Random number) ซึ่งมีการแจกแจงแบบสม่ำเสมอในช่วง (0, 1) เป็นพื้นฐานในการสร้าง โดยในการวิจัยครั้งนี้ต้องการสร้างตัวแปรสุ่มที่มีการแจกแจงแบบแบร์นูลลี และมีขั้นตอนดังนี้

1) การสร้างเลขสุ่มที่เป็นอิสระกันและมีการแจกแจงเอกรูปในช่วง [0,1][1]

วิธีการคณิตศาสตร์ในการจำลองเลขสุ่ม(เทียม) มีหลายวิธีการ สำหรับวิธีการที่ได้รับความนิยมใช้กันมากในปัจจุบัน คือ วิธีสมภาค (Congruential Method) ซึ่งมีสูตรหรือตัวแบบหนึ่งที่ใช้กันมาก คือ

$$X_i = (c + aX_{i-1}) \bmod m , \quad i = 1, 2, \ldots$$

โดยที่ค่า c , a และ m เป็นค่าคงที่จำนวนเต็มค่าไม่เป็นลบ และความหมายของตัวแบบ คือ X_i เป็นเศษเหลือที่เป็นจำนวนเต็มที่ได้จากการหาร $(c + aX_{i-1})$ ด้วย m นั่นคือ $X_i = c + aX_{i-1} - mk_i$ ซึ่ง $k_i = \lfloor (c + aX_{i-1})/m \rfloor$ (หมายถึง จำนวนเต็มใหญ่ที่สุดที่น้อยกว่าหรือเท่ากับผลหาร $(c + aX_{i-1})/m$) ดังนั้นค่าเป็นไปได้ของ X_i คือ 0 , 1 , … , m-1 และก่อนที่จะได้ค่าของ $X_1, X_2, \ldots$ ต้องกำหนดค่าของ c , a , m และ X_0 เราเรียก X_0 ว่า ซีด (seed) หรือ ค่าเริ่มต้น (starting value) จาก X_i ที่ได้จากการคำนวณนำมาหาค่า R_i ซึ่ง

$$R_i = \frac{X_i}{m} \quad , i = 1 , 2 , \ldots \qquad \ldots\ldots\ldots(1.1)$$

จะได้ R_i มีค่าอยู่ในช่วง [0,1) เรียก $R_1, R_2, R_3 \ldots$ ว่า เลขสุ่มเทียม หรือ เลขสุ่มคล้าย

ตัวแบบจำลองสมภาคแบบผลคูณที่ใช้กันมากตัวแบบหนึ่ง ซึ่งได้ผ่านการตรวจสอบคุณสมบัติแล้วหลายประการ คือ กำหนด c = 0, $m = 2^{31}-1$ = 2147483647, $a = 7^5$ = 16807 และ X_0 เป็นจำนวนเต็มบวกที่ไม่เกิน m (ควรเป็นเลขคี่) ฟังก์ชันการจำลองเลขสุ่มที่เป็นอิสระกันและมีการแจกแจงเอกรูปในช่วง [0,1] คือ subroutine random

2) การสร้างเลขสุ่มที่มีการแจกแจงแบบปกติ[2]

ในการสร้างเลขสุ่มที่มีการแจกแจงแบบปกติจะใช้วิธีโพลาร์ ซึ่งวิธีโพลาร์ได้จากการดัดแปลงวิธีของ Box และ Muller

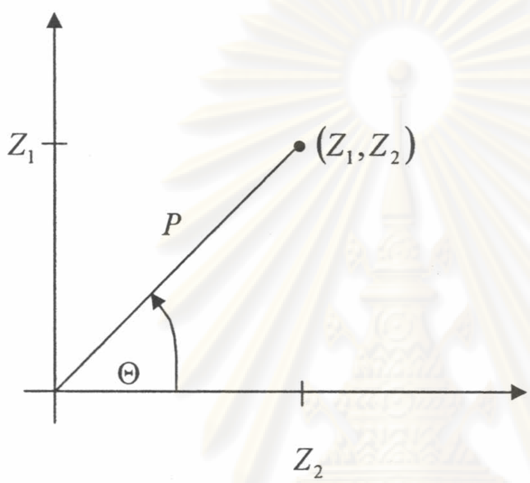
[1] ที่มา : มานพ วราภักดิ์, การจำลองเบื้องต้น (กรุงเทพฯ: ศูนย์ผลิตตำราเรียนสถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2547), หน้า 43

[2] เรื่องเดียวกัน, หน้า 145

George E. P. Box และ Mervin E. Muller (1958) ได้คิดค้นวิธีการจำลองตัวแปรสุ่มปกติมาตรฐาน *N*(0,1) โดยใช้การแปลงตัวแปรสุ่ม คือ จากตัวแปรสุ่มมาตรฐานอิสระกัน Z_1 และ Z_2 ได้จุดบนระนาบในระบบพิกัดคาร์ทีเซียน (Cartesian coordinates) ดูรูป 2.1 จากตัวแปรสุ่มในระบบพิกัดคาร์ทีเซียนแปลงเป็นตัวแปรสุ่มในระบบพิกัดเชิงขั้ว (polar coordinates) เป็นจุด
(P, Θ) โดยที่

$$Z_1 = P\cos\Theta$$

$$Z_2 = P\sin\Theta$$



รูป 2.1

จากนี้เมื่อทราบการแจกแจงของ P และของ Θ เราจะจำลอง P และ Θ จากการแจกแจงนั้นและเมื่อแทนค่าจะได้ Z_1 และ Z_2 ซึ่งวิธีการนี้เรียกว่า วิธีการบอกซ์ มูลเลอร์

การแปลง $z_1 = \rho\cos\theta$ และ $z_2 = \rho\sin\theta$ เป็นการแปลงแบบหนึ่งต่อหนึ่ง (one-to-one transformation) จากปริภูมิ $R_{Z_1,Z_2} = \{(z_1, z_2) : -\infty < z_1 < \infty, -\infty < z_2 < \infty\}$ ของ (Z_1, Z_2) ไปยังปริภูมิ $R_{P,\Theta} = \{(\rho,\theta) : 0 \le \rho < \infty, 0 \le \theta \le 2\pi\}$ โดยมีจาโคเบียน (Jacobian) ของการแปลง J ดังนี้

$$J = \begin{vmatrix} \dfrac{\partial z_1}{\partial \rho} & \dfrac{\partial z_1}{\partial \theta} \\ \dfrac{\partial z_2}{\partial \rho} & \dfrac{\partial z_2}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{vmatrix}$$

$$= \rho\left(\cos^2\theta + \sin^2\theta\right) = \rho$$

เพราะฉะนั้น โดยเทคนิคของการแปลงในทฤษฎีความน่าจะเป็น ได้ว่า P และ Θ มีฟังก์ชันความหนาแน่นร่วม (joint density function)

$$f_{P,\Theta}(\rho,\theta) = f_{Z_1,Z_2}(\rho\cos\theta, \rho\sin\theta)|J|$$

เนื่องจาก Z_1 และ Z_2 มีฟังก์ชันความหนาแน่นร่วม

$$\begin{aligned} f_{Z_1,Z_2}(z_1,z_2) &= f_{Z_1}(z_1) f_{Z_2}(z_2) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_1^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_2^2} \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}\left(z_1^2+z_2^2\right)} \end{aligned}$$

ดังนั้น โดยการแทนค่า ได้ผลลัพธ์

$$\begin{aligned} f_{P,\Theta}(\rho,\theta) &= \frac{1}{2\pi} e^{-\frac{1}{2}\rho^2} \cdot \rho \\ &= \frac{1}{\sqrt{2\pi}} \rho e^{-\frac{1}{2}\rho^2} \quad , \; 0 \le \theta \le 2\pi, 0 \le \rho < \infty \\ &= f_\Theta(\theta) f_P(\rho) \end{aligned}$$

โดยที่

$f_\Theta(\theta) = \dfrac{1}{2\pi}, \quad 0 \le \theta \le 2\pi$ เป็นฟังก์ชันของ θ เท่านั้น ไม่ขึ้นอยู่กับ ρ และ $f_P(\rho) = \rho e^{-\frac{1}{2}\rho^2}, \rho \ge 0$ เป็นฟังก์ชัน ρ เท่านั้น ไม่ขึ้นอยู่กับ θ เพราะฉะนั้น โดยคุณสมบัติของตัวแปรสุ่มอิสระ ได้ว่า P และ Θ เป็นอิสระกัน (เชิงสถิติ) นอกจากนี้ พบว่า f_Θ เป็นฟังก์ชันความหนาแน่นของการแจกแจงเอกรูป $U(0,2\pi)$ และ f_P เป็นฟังก์ชันความหนาแน่นเช่นกัน ซึ่งเป็นฟังก์ชันความหนาแน่นของการแจกแจงเรย์ลี (Rayleigh distribution) เพราะฉะนั้น ในการจำลอง Z_1 และ Z_2 เราจะจำลอง Θ และ P อย่างอิสระกัน โดยจำลอง P จาก $f_P(\rho) = \rho\exp\left(-\frac{1}{2}\rho^2\right)$ ซึ่งด้วยวิธีการแปลงผกผัน ได้ตัวแบบจำลอง $P = \sqrt{-2\ell n R_1}$, $R_1 \sim U(0,1)$ และ จำลอง Θ จากการแจกแจง $U(0,2\pi)$ ได้ $\Theta = 2\pi R_2$, $R_2 \sim U(0,1)$ ดังนั้น ตัวแบบจำลองตัวแปรสุ่ม $Z_1 \sim N(0,1)$ และ $Z_2 \sim N(0,1)$ อิสระกันคือ

$$\begin{aligned} Z_1 &= \sqrt{-2\ell n R_1}\cos(2\pi R_2) \\ Z_2 &= \sqrt{-2\ell n R_1}\sin(2\pi R_2) \end{aligned} \tag{2.1}$$

โดยที่ R_1, $R_2 \sim U(0,1)$ และเป็นอิสระกัน (พิสูจน์กลับได้ว่า Z_1 และ Z_2 อิสระกัน และต่างมีการแจกแจง $N(0,1)$)

การใช้ตัวแบบ (2.1) เราต้องจำลองเลขสุ่มสองตัว แต่เราก็ได้ค่าของตัวแปรสุ่ม $N(0,1)$ สองค่าเป็นอิสระกัน ในทางปฏิบัติ จะใช้เฉพาะสูตรใดสูตรหนึ่งของ (2.1) ก็ได้

วิธีโพลาร์ของ Marsaglia, MacLaren และ Bray (1964) ดัดแปลงวิธีของ Box และ Muller โดยหลีกเลี่ยงการคำนวณ cosine และ sine ด้วยการใช้วิธีรับ-ปฏิเสธ ดังนี้

สุ่มจุด (V_1,V_2) ในระนาบสี่เหลี่ยมจัตุรัสขนาด 2x2 โดยมีจุดกลางที่จุด (0,0) นั่นคือ จำลอง V_1 จาก $U(-1,1)$ และจำลอง V_2 จาก $U(-1,1)$ อย่างอิสระกัน จะได้จุด (V_1,V_2) อยู่ในสี่เหลี่ยมจัตุรัส (ดูรูป 2.2 ประกอบ) จะสุ่มจุดดังกล่าวจนกว่าจะได้จุดอยู่ในวงกลมรัศมี 1 ซึ่งมีจุดศูนย์กลางที่จุด (0,0) บรรจุอยู่ในสี่เหลี่ยมจัตุรัส จุดที่ไม่ตกอยู่ในวงกลมจะตัดออก ไม่พิจารณา นั่นคือ จนกว่าจะได้จุด (V_1,V_2) ซึ่ง $V_1^2+V_2^2\le 1$ ดังนั้น (V_1,V_2) มีการแจกแจงร่วมแบบเอกรูปบนระนาบวงกลมรัศมี 1 โดยมีฟังก์ชันความหนาแน่นร่วม (แบบมีเงื่อนไข) :

$$f_{V_1,V_2|C}(v_1,v_2|C)=\frac{f_{V_1,V_2}(v_1,v_2)}{P((V_1,V_2)\in C)}$$

โดยที่

$$C=\{(v_1,v_2):v_1^2+v_2^2\le 1\}$$

$$f_{V_1,V_2}(v_1,v_2)=f_{V_1}(v_1)f_{V_2}(v_2)$$

$$=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\quad -1\le v_1\le 1,\quad -1\le v_2\le 1$$

$$P((V_1,V_2)\in C)=P(V_1^2+V_2^2\le 1)$$

$$=\frac{\text{พื้นที่วงกลมที่มีรัศมียาว 1}}{\text{พื้นที่สี่เหลี่ยมจัตุรัสขนาด 2X2}}$$

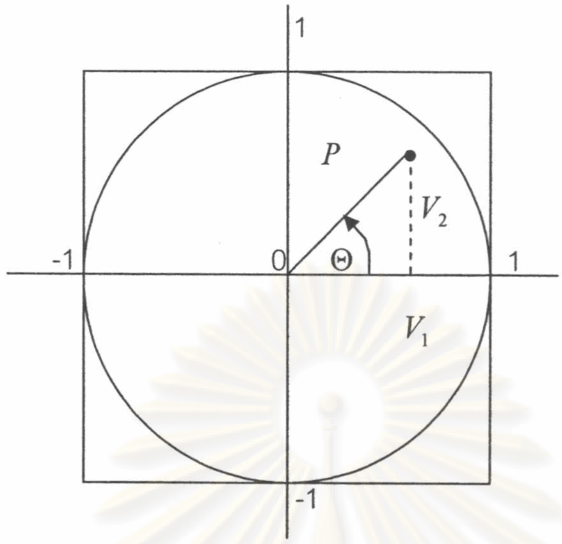
$$=\frac{\pi}{4}$$

หรือคำนวณหาค่า $P(V_1^2+V_2^2\le 1)$ โดยใช้อินทิกรัล :

$$P(V_1^2+V_2^2\le 1)=4\int_0^1\int_0^{\sqrt{1-v_2^2}}\frac{1}{4}dv_1dv_2=\frac{\pi}{4}$$

เพราะฉะนั้น

$$f_{V_1,V_2|C}(v_1,v_2|C)=\frac{1/4}{\pi/4}=\frac{1}{\pi},\quad v_1^2+v_2^2\le 1,\quad -1\le v_1,v_2\le 1$$



รูป 2.2

สำหรับจุด (V_1, V_2) อยู่ในวงกลมแปลงเป็นจุด (P, Θ) ในพิกัดเชิงขั้ว ได้การแปลง

$$P = \sqrt{V_1^2 + V_2^2}$$

$$\Theta = \tan^{-1}\left(\frac{V_2}{V_1}\right)$$

จากนี้ พิจารณาการแจกแจงความน่าจะเป็นร่วมของ P และ Θ

ได้ว่า การแปลง $\rho = \sqrt{v_1^2 + v_2^2}$, $\theta = \tan^{-1}\left(\frac{v_2}{v_1}\right)$ เป็นการแปลงแบบหนึ่งต่อหนึ่ง จาก $R_{V_1,V_2} = \{(v_1, v_2) : 0 \le v_1^2 + v_2^2 \le 1, -1 \le v_1, v_2 \le 1\}$ ไปยัง $R_{P,\Theta} = \{(\rho, \theta) : 0 \le \rho \le 1,$ $0 \le \theta \le 2\pi\}$ และได้ว่า

$$\frac{\partial(\rho,\theta)}{\partial(v_1,v_2)} = \begin{vmatrix} \dfrac{v_1}{\sqrt{v_1^2+v_2^2}} & \dfrac{v_2}{\sqrt{v_1^2+v_2^2}} \\ \dfrac{-v_2}{v_1^2+v_2^2} & \dfrac{v_1}{v_1^2+v_2^2} \end{vmatrix} = \frac{1}{\sqrt{v_1^2+v_2^2}}$$

เพราะฉะนั้น ได้จาโคเบียนของการแปลงผกผัน

$$J = \frac{\partial(v_1,v_2)}{\partial(\rho,\theta)} = \sqrt{v_1^2+v_2^2} = \rho$$

และดังนั้น P , Θ มีฟังก์ชันความหนาแน่นร่วม

$$f_{P,\Theta}(\rho,\theta) = f_{V_1,V_2|C}(v_1, v_2|C)|J| = \frac{\rho}{\pi}$$

$$= \frac{1}{2\pi} 2\rho, \quad 0 \le \theta \le 2\pi, \quad 0 \le \rho \le 1$$
$$= f_{\Theta}(\theta) f_P(\rho)$$

โดยที่ $f_{\Theta}(\theta) = \frac{1}{2\pi}, \ 0 \le \theta \le 2\pi$ และ $f_P(\rho) = 2\rho, \ 0 \le \rho \le 1$ เพราะฉะนั้น P และ Θ เป็นอิสระกัน และได้ว่า $\Theta \sim U(0, 2\pi)$ และแสดงได้ง่ายว่า $P^2 = V_1^2 + V_2^2 \sim U(0,1)$ และเป็นอิสระกับมุม Θ ดังนั้น จะจำลอง $\cos\Theta$ และ $\sin\Theta$ ด้วยการจำลอง (V_1, V_2) ในวงกลม และให้

$$\cos\Theta = \frac{V_1}{P} = \frac{V_1}{\sqrt{V_1^2 + V_2^2}}$$
$$\sin\Theta = \frac{V_2}{P} = \frac{V_2}{\sqrt{V_1^2 + V_2^2}}$$

เพราะฉะนั้น จากตัวแบบจำลอง (2.1) เขียนใหม่เป็น

$$Z_1 = \sqrt{-2\ell n P^2} \, \frac{V_1}{\sqrt{V_1^2 + V_2^2}}$$
$$Z_2 = \sqrt{-2\ell n P^2} \, \frac{V_2}{\sqrt{V_1^2 + V_2^2}}$$

(ใช้ P^2 เป็นเลขสุ่มได้ เพราะว่า $P^2 = V_1^2 + V_2^2 \sim U(0,1)$) หรือ

$$Z_1 = V_1 \sqrt{\frac{-2\ell n(V_1^2 + V_2^2)}{V_1^2 + V_2^2}}$$
$$Z_2 = V_2 \sqrt{\frac{-2\ell n(V_1^2 + V_2^2)}{V_1^2 + V_2^2}} \tag{2.2}$$

สำหรับตัวแบบ (2.2) มีขั้นตอนวิธีดังนี้

(1) จำลองเลขสุ่ม R_1 และ R_2

(2) $V_1 = 2R_1 - 1, \ V_2 = 2R_2 - 1$ (จำลอง V_1, V_2 จาก $U(-1,1)$)

(3) $S = V_1^2 + V_2^2$

(4) ถ้า $S > 1$ กลับไปขั้นตอย (1)

(5) $W = \sqrt{\frac{-2\ell n S}{S}}$

(6) $Z_1 = V_1 W, \ Z_2 = V_2 W$

ด้วยวิธีโพลาร์ มีประสิทธิภาพหรือความน่าจะเป็นที่จุด (V_1, V_2) จะอยู่ในวงกลม หรือความน่าจะเป็นที่จะได้ตัวแปรสุ่ม (Z_1, Z_2) เท่ากับ $P(S \le 1) = P(V_1^2 + V_2^2 \le 1) = \frac{\pi}{4} \approx 0.7854$ เพราะฉะนั้น วิธีโพลาร์จะมีจำนวนรอบทำซ้ำขั้นตอน (1) – (3) จนกว่าจะได้ (Z_1, Z_2) หนึ่งคู่เท่ากับ $\frac{\pi}{4} \approx 1.273$ ครั้งโดยเฉลี่ย

3) **การคำนวณค่าเฉลี่ยและความแปรปรวน**

เมื่อจำลองได้ตัวแปรสุ่มที่มีการแจกแจงแบบปกติ ให้นำค่าตัวแปรสุ่มมาคำนวณหาค่าเฉลี่ยและความแปรปรวน ดังนี้

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2$$

โดยที่ $\overline{Y}$ และ S^2 คือค่าเฉลี่ยและความแปรปรวนตามลำดับ

3.2.2 <u>การคำนวณช่วงความเชื่อมั่นของวิธีการประมาณทั้ง 3 วิธี</u>

ในการคำนวณช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยของการแจกแจงแบบล็อกนอร์มอลตามสูตรของแต่ละวิธีการประมาณ ทำได้ดังนี้

1) วิธีการประมาณของคอกซ์ (Cox's method)

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับ ค่าเฉลี่ย (θ) คือ

$$\left[\exp\left(\left\{\overline{Y} + \frac{S^2}{2}\right\} - Z_{\alpha/2}\sqrt{\left\{\frac{S^2}{n} + \frac{S^4}{2(n-1)}\right\}}\right), \exp\left(\left\{\overline{Y} + \frac{S^2}{2}\right\} + Z_{\alpha/2}\sqrt{\left\{\frac{S^2}{n} + \frac{S^4}{2(n-1)}\right\}}\right)\right]$$

2) วิธีการประมาณแบบคอนเซอเวทีฟ (Conservative method)

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับ ค่าเฉลี่ย (θ) คือ

$$\left[\exp\left(\left\{\overline{Y} + \frac{S^2}{2} - \frac{t_{(1-\alpha/2),(n-1)}}{\sqrt{n}}\sqrt{S^2\left(1+\frac{S^2}{2}\right)}\right\}\right), \exp\left(\left\{\overline{Y} + \frac{S^2}{2} + \frac{q_{\alpha/2,(n-1)}}{\sqrt{n}}\sqrt{S^2\left(1+\frac{S^2}{2}\right)}\right\}\right)\right]$$

โดยที่ $t_{(1-\alpha/2),(n-1)}$ เป็นเปอร์เซนไทล์ที่ $1-\alpha/2$ ของการแจกแจงแบบที ที่มีองศาแห่งความเป็นอิสระเท่ากับ $n-1$ และ

$$q_{\alpha/2,(n-1)} = \sqrt{\left(\frac{n}{2}\right)\left(\frac{(n-1)}{\chi^2_{\alpha/2,(n-1)}} - 1\right)}$$

โดยที่ $\chi^2_{\alpha/2,(n-1)}$ เป็นเปอร์เซนไทล์ที่ $\alpha/2$ ของการแจกแจงแบบไคกำลังสอง ที่มีองศาแห่งความเป็นอิสระเท่ากับ $n-1$

3) วิธีการประมาณแบบพาราเมตริกบูทสแทร็พ (Parametric bootstrap method)

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับ ค่าเฉลี่ย (θ) คือ

$$\left[\exp\left(\bar{Y} + \frac{S^2}{2} + t_1^*\sqrt{\frac{S^2\left(1+\frac{S^2}{2}\right)}{n}}\right), \exp\left(\bar{Y} + \frac{S^2}{2} + t_0^*\sqrt{\frac{S^2\left(1+\frac{S^2}{2}\right)}{n}}\right)\right]$$

โดยค่า t_0^*และ t_1^* ได้จากกระบวนการทางบูทสแทร็พ ดังนี้

1. สร้างตัวแปรสุ่ม $N_i^* \sim N(0,1)$ และ $\chi_i^{2*} \sim \chi^2_{n-1}, i = 1,\ldots,B, N_i^*$ และ χ_i^{2*} เป็นอิสระกัน
2. คำนวณ T_i^* จากสมการที่ (1)(ในหัวข้อ 2.3.3) โดยแทนค่า σ, N, χ^2_{n-1} ด้วย S, N_i^*, χ_i^{2*} ตามลำดับ
3. เรียงค่า T_i^* ตามลำดับ นั่นคือ $T^*_{(1)} < T^*_{(2)} < \ldots < T^*_{(B)}$
4. ประมาณค่า t_1 ด้วย $t_1^* = T_{[(1-\alpha/2)B]}$ และ ประมาณค่า t_0 ด้วย $t_0^* = T_{[(\alpha/2)B]}$ โดยที่ $[d]$แสดงถึงจำนวนเต็มที่มากที่สุดซึ่งมีค่าน้อยกว่าหรือเท่ากับ d

โดย $B = 2{,}000$

3.2.3 <u>การคำนวณค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น</u>

การคำนวณหาค่าสัมประสิทธิ์ความเชื่อมั่นของวิธีการประมาณแต่ละวิธี ทำได้โดยการตรวจสอบว่าช่วงความเชื่อมั่นที่คำนวณจากแต่ละวิธีการประมาณครอบคลุมค่าพารามิเตอร์ θ หรือไม่ หากช่วงความเชื่อมั่นที่คำนวณได้ของวิธีการประมาณใดครอบคลุมค่าพารามิเตอร์ θ จะทำการนับจำนวนครั้งและบวกสะสมค่าไว้ โดยในแต่ละสถานการณ์จะคำนวณช่วงความเชื่อมั่นซ้ำกัน 1,000 ครั้ง ผลบวกสะสมที่ได้คือจำนวนครั้งทั้งหมดที่ช่วงความเชื่อมั่นคลอบคลุมค่าพารามิเตอร์ θ ซึ่งจะนำไปคำนวณค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองดังนี้

ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลอง

$$= \frac{\text{จำนวนครั้งทั้งหมดที่ช่วงความเชื่อมั่นคลอบคลุมค่าพารามิเตอร์ } \theta}{1{,}000}$$

ส่วนการคำนวณหาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นนั้น จะทำการคำนวณเฉพาะในสถานการณ์ที่ช่วงความเชื่อมั่นที่คำนวณได้คลอบคลุมค่าพารามิเตอร์ θ เท่านั้น โดยคำนวณหาผลต่างระหว่างขีดจำกัดความเชื่อมั่นบนและขีดจำกัดความเชื่อมั่นล่าง นำผลต่างที่ได้มาบวกสะสมไว้ ทำจนครบ 1,000 ครั้ง นำไปคำนวณค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นดังนี้

ความยาวเฉลี่ยของช่วงความเชื่อมั่น

$$= \frac{\text{ผลรวมของผลต่างระหว่างขีดจำกัดบนและขีดจำกัดล่างของช่วงความเชื่อมั่นทั้งหมด 1,000 ช่วง}}{1{,}000}$$

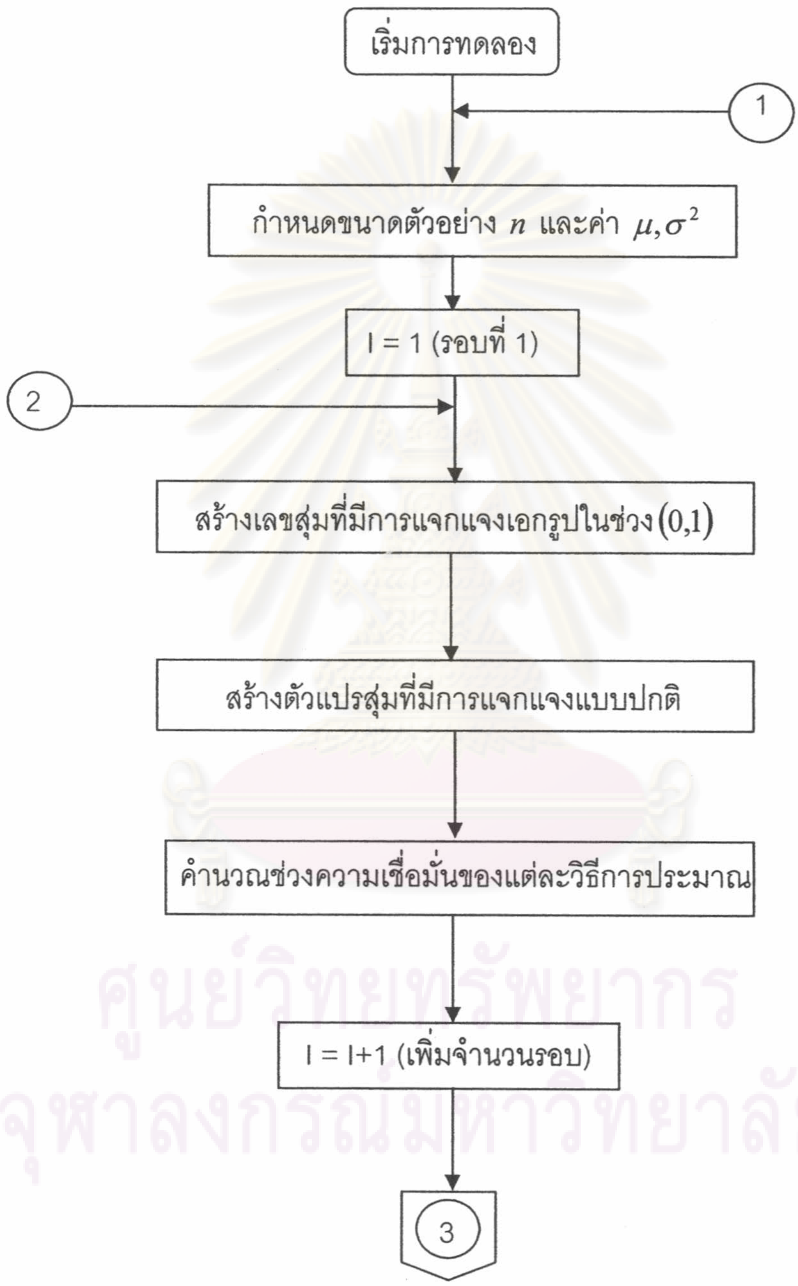
3.2.4 <u>เปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น</u>

ในการตรวจสอบว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการทดลองมีค่าเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดหรือไม่นั้น ผู้วิจัยจะอาศัยการทดสอบสมมติฐานโดยใช้ตัวสถิติ Z ดังนั้น ถ้าวิธีการประมาณใดให้ค่าสัมประสิทธิ์ความเชื่อมั่นของช่วงความเชื่อมั่นจากการทดลองในแต่ละสถานการณ์ ที่สัมประสิทธิ์ความเชื่อมั่น 0.90, 0.95 และ 0.99 มีค่าไม่ต่ำกว่า 0.8878, 0.9387 และ 0.9827 ตามลำดับ (รายละเอียดการคำนวณค่าเหล่านี้ ได้กล่าวไว้ในหัวข้อที่ 2.4) เราจะสรุปได้ว่าวิธีการประมาณนั้นให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในสถานการณ์นั้น

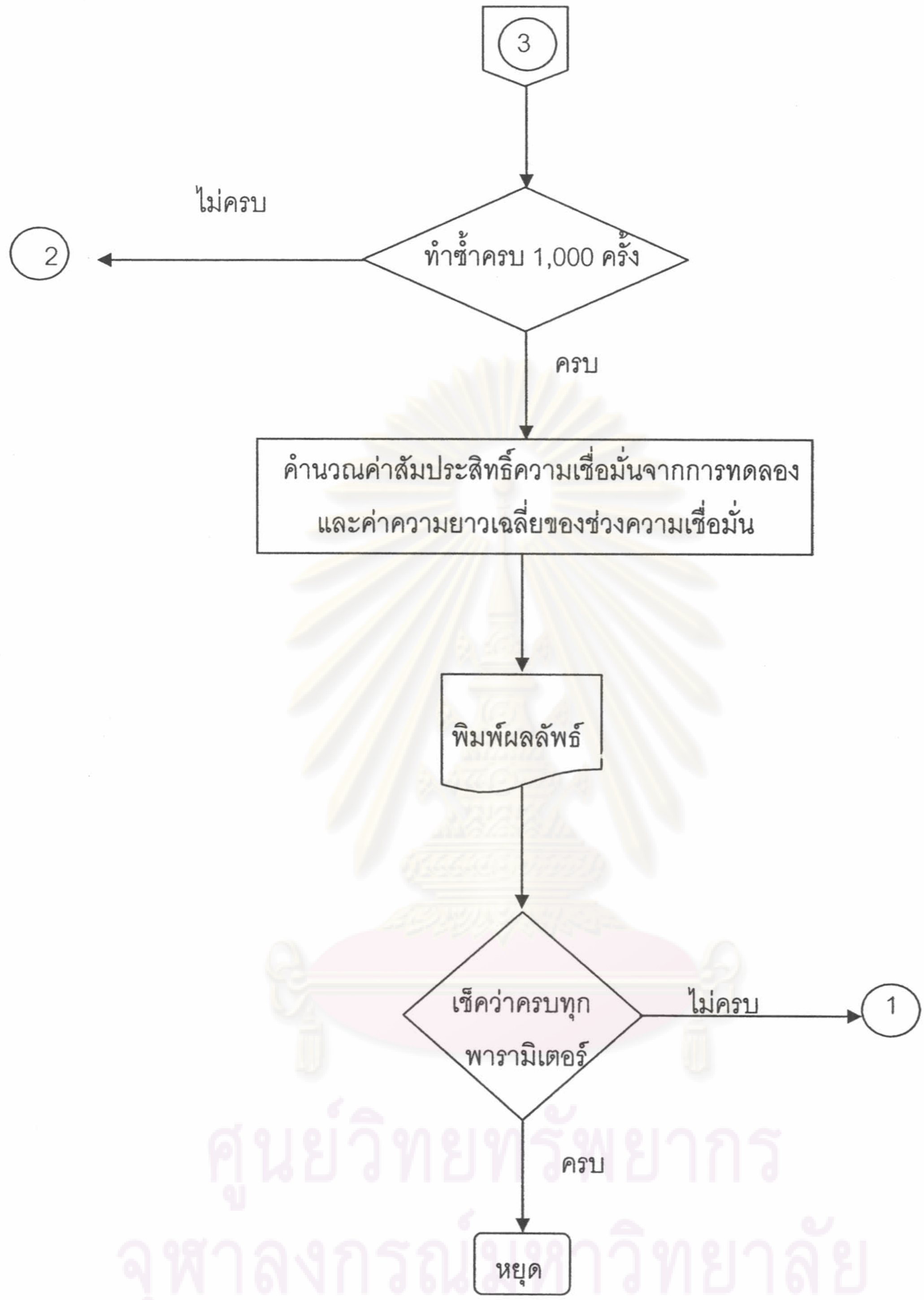
เมื่อทำการทดลองและตรวจสอบแล้วทราบว่า วิธีการประมาณใดให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่ต่ำกว่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในแต่ละสถานการณ์ ให้นำมาเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นว่าวิธีการประมาณวิธีใด ให้ค่าความยาวเฉลี่ยต่ำสุดในแต่ละสถานการณ์

3.3 **ขั้นตอนในการทำงานของโปรแกรม**

สำหรับขั้นตอนในการทำงานของโปรแกรมการคำนวณค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลอง และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น สามารถสรุปเป็นผังงานได้ตามรูปที่ 3.1



รูปที่ 3.1


3
ไม่ครบ
2
ทำซ้ำครบ 1,000 ครั้ง
ครบ
คำนวณค่าสัมประสิทธิ์ความเชื่อมั่นจากการทดลอง
และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น
พิมพ์ผลลัพธ์
เช็คว่าครบทุก
พารามิเตอร์
ไม่ครบ
1
ครบ
หยุด

รูปที่ 3.1 (ต่อ)