บทที่ 3

การดำเนินการวิจัย

3.1 ความนำ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อนำเสนอวิธีการประมาณค่าพารามิเตอร์ p แบบช่วงจากข้อมูลตัวอย่างแบบผกผันโดยวิธีบูทแสตรป และทำการตรวจสอบความผิดพลาดและความเชื่อถือได้ของวิธีการประมาณดังกล่าว กับวิธีการเดิม ได้แก่ วิธีแบบวาล์วโดยอาศัยวิธีภาวะน่าจะเป็นสูงสุด วิธีแบบวาล์วโดยอาศัย วิธีความแปรปรวนต่ำสุดอย่างเอกรูป วิธีอาศัยการแจกแจงแบบไคสแควร์ และวิธีสกอร์ โดยใช้การจำลองข้อมูล

เกณฑ์ที่ใช้ในการตรวจสอบความผิดพลาด และความเชื่อถือได้ของวิธีการดังกล่าวจะพิจารณาจากค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นซึ่งมีรายละเอียดเกี่ยวกับแผนการวิจัย ขั้นตอนการดำเนินการวิจัย และขั้นตอนการจำลองข้อมูลที่ใช้ในการวิจัยนำเสนอตามลำดับดังนี้

3.2 แผนการวิจัย

ตามวัตถุประสงค์ที่ได้กล่าวมาแล้วในบทที่ 1 ผู้วิจัยได้กำหนดสถานการณ์ต่าง ๆ สำหรับการตรวจสอบความผิดพลาด และความเชื่อถือได้ของวิธีการประมาณค่าแบบช่วงทั้ง 5 วิธี ดังต่อไปนี้

1. สร้างตัวแปรสุ่ม Y ที่มีการแจกแจงทวินามลบ ซึ่ง Y แทนจำนวนครั้งของการเกิดคุณลักษณะที่ไม่สนใจก่อนที่จะเกิดคุณลักษณะที่สนใจครั้งที่ r มีพารามิเตอร์ r และ p โดยกำหนดจำนวนครั้งของการเกิดคุณลักษณะที่สนใจ (r) เท่ากับ 1, 3, 5, 7, 10, 20, 30 และ 50

2. กำหนดค่าความน่าจะเป็นในการเกิดคุณลักษณะที่สนใจ (p) เท่ากับ 0.001, 0.01, 0.1 และ 0.5

จะได้ชุดข้อมูลดังแสดงในตารางที่ 3.1

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ตารางที่ 3.1 ชุดข้อมูลที่ใช้ในการวิจัย

กรณีที่	r	p	กรณีที่	r	p	กรณีที่	r	p	กรณีที่	r	p
1	1	0.001	9	5	0.001	17	10	0.001	25	30	0.001
2	1	0.01	10	5	0.01	18	10	0.01	26	30	0.01
3	1	0.1	11	5	0.1	19	10	0.1	27	30	0.1
4	1	0.5	12	5	0.5	20	10	0.5	28	30	0.5
5	3	0.001	13	7	0.001	21	20	0.001	29	50	0.001
6	3	0.01	14	7	0.01	22	20	0.01	30	50	0.01
7	3	0.1	15	7	0.1	23	20	0.1	31	50	0.1
8	3	0.5	16	7	0.5	24	20	0.5	32	50	0.5

3. กำหนดระดับความเชื่อมั่น (1-α)100% ที่ทำการวิจัยเท่ากับ 95% และ 99%

4. ทำการทดลองซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ที่กำหนด

3.3 ขั้นตอนการดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัยมีดังต่อไปนี้

1. จำลองข้อมูลตามสถานการณ์ต่าง ๆ ที่กำหนดไว้ในแผนการวิจัย

2. คำนวณช่วงความเชื่อมั่นด้วยวิธีการประมาณค่าแบบช่วงทั้ง 5 วิธี

โดยนำข้อมูลที่ได้จากข้อ 1 มาทำการคำนวณช่วงความเชื่อมั่นโดยใช้ตัวประมาณค่าแบบช่วงทั้ง 5 ดังต่อไปนี้

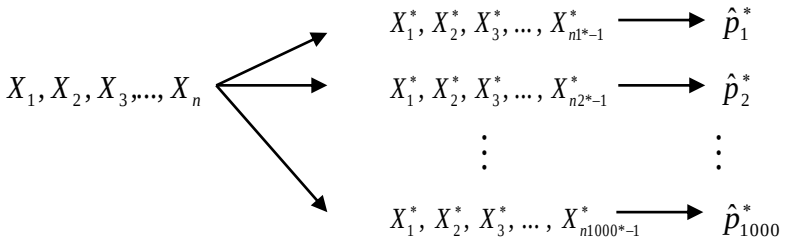
2.1 วิธีบูทแสตรป

เริ่มจากการทำการทดลองสุ่มแบบแบร์นูลลี ซ้ำๆ กันไปจนกว่าจะเกิดผลสำเร็จจนครบจำนวน r ครั้งตามที่กำหนด จากตัวแปรสุ่ม Y คือ จำนวนครั้งของการเกิดคุณลักษณะที่ไม่สนใจจากการทำการทดลองสุ่มแบบแบร์นูลลีที่ซ้ำเดิมอย่างเป็นอิสระต่อกันจนกว่าจะเกิดคุณลักษณะที่สนใจครั้งที่ r ตัวแปรสุ่ม Y จึงมีรูปแบบการแจกแจงแบบทวินามลบ

จะได้ชุดตัวอย่าง $X_1, X_2, X_3, \ldots, X_n$ เมื่อ $n = r + Y$ แล้วนำข้อมูลตัวอย่างชุดดังกล่าวนี้ซึ่งจะเรียกว่าชุดตัวอย่างตั้งต้น มาทำการสุ่มซ้ำทีละหน่วยแบบใส่คืนจนได้ข้อมูลตัวอย่างจากการสุ่มซ้ำดังกล่าวนี้ ซึ่งมีคุณลักษณะที่สนใจจนครบจำนวน r ครั้งตามที่กำหนด ซึ่งเมื่อทำการสุ่มซ้ำไปเช่นนี้เป็นจำนวนทั้งสิ้น 1000 รอบ ก็จะทำให้ได้รับข้อมูลตัวอย่างจำนวนทั้งสิ้น 1000 ชุด ที่ได้รับจากการสุ่มซ้ำจากชุดตัวอย่างตั้งต้น นั่นคือกำหนดให้ $X_1^*, X_2^*, X_3^*, \ldots, X_{n1^*-1}^*$, $X_1^*, X_2^*, X_3^*, \ldots, X_{n2^*-1}^*$,

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

$\ldots, X_1^*, X_2^*, X_3^*, \ldots, X_{n1000^*-1}^*$ ซึ่งจะเรียกชุดดังกล่าว 1000 ชุด นี้ว่า ตัวอย่างบูทแสตรป (Bootstrap Sample) ที่ได้จากการสุ่มซ้ำ 1000 รอบสำหรับนำมาทำการประมาณค่าพารามิเตอร์ p



รูป 3.1 แสดงขั้นตอนการสุ่มตัวอย่างด้วยวิธีบูทแสตรปจากข้อมูลตัวอย่างแบบผกผัน

ซึ่งจะได้ค่าประมาณพารามิเตอร์ p เป็นจำนวนทั้งสิ้น 1000 ค่า จากนั้นนำค่าประมาณพารามิเตอร์ p ทั้งหมด 1000 ค่า มาทำการจัดเรียงลำดับจากค่าน้อยที่สุดไปหาค่ามากที่สุด แล้วจึงทำการคำนวณหาค่าตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(\alpha/2)$ สำหรับนำมาใช้เป็นขีดจำกัดล่างของช่วงความเชื่อมั่น $\hat{p}_{BL}$ และทำการคำนวณหาค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(1-\alpha/2)$ สำหรับนำมาใช้เป็นขีดจำกัดบน $\hat{p}_{BU}$ ก็จะทำให้ได้รับช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับพารามิเตอร์ p โดยวิธีบูทแสตรป คือ $[\hat{p}_{BL}, \hat{p}_{BU}]$

2.2 วิธีแบบวาล์วโดยอาศัยวิธีภาวะน่าจะเป็นสูงสุด วิธีแบบวาล์วโดยอาศัยวิธีความแปรปรวนต่ำสุดอย่างเอกรูป วิธีอาศัยการแจกแจงแบบไคสแควร์ และวิธีสกอร์

เมื่อได้ตัวแปรสุ่ม Y แล้วนำมาหาค่าประมาณแบบจุดโดยใช้ตัวประมาณแบบภาวะน่าจะเป็นสูงสุด แล้วคำนวณหาช่วงความเชื่อมั่นด้วยวิธีการประมาณค่าแบบช่วงทั้ง 4 วิธีดังต่อไปนี้

2.2.1 วิธีแบบวาล์วโดยอาศัยวิธีภาวะน่าจะเป็นสูงสุด (Wald-Type Confidence Interval Based on MLE: WM)

$$\left[\max\left\{\hat{P}-Z_{\alpha/2}\sqrt{\frac{\hat{P}^2(1-\hat{P})}{r}},0\right\}, \min\left\{\hat{P}+Z_{\alpha/2}\sqrt{\frac{\hat{P}^2(1-\hat{P})}{r}},1\right\}\right]$$

2.2.2 วิธีแบบวาล์วโดยอาศัยวิธีความแปรปรวนต่ำสุดอย่างเอกรูป (Wald-Type Confidence Interval Based on UMVUE: WU)

$$\left[\max\left\{\tilde{P}-Z_{\alpha/2}\sqrt{\frac{\tilde{P}(1-\tilde{P})}{N-2}},0\right\}, \min\left\{\hat{P}+Z_{\alpha/2}\sqrt{\frac{\tilde{P}(1-\tilde{P})}{N-2}},1\right\}\right]$$

2.2.3 วิธีอาศัยการแจกแจงแบบไคสแควร์ (Confidence Interval Based on Chi-square Distribution: CS)

$$\frac{\chi^2_{2r,\alpha/2}}{2(r+y)}, \frac{\chi^2_{2r,1-\alpha/2}}{2(r+y)}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

2.2.4 วิธีสกอร์ (Score Confidence Interval: SC)

$$\left[\max\left\{\frac{\left(2Nr - rZ_{\alpha/2}^2\right) - \sqrt{r^2Z_{\alpha/2}^4 - 4Nr^2Z_{\alpha/2}^2 + 4N^2rZ_{\alpha/2}^2}}{2N^2}, 0\right\}, \min\left\{\frac{\left(2Nr - rZ_{\alpha/2}^2\right) + \sqrt{r^2Z_{\alpha/2}^4 - 4Nr^2Z_{\alpha/2}^2 + 4N^2rZ_{\alpha/2}^2}}{2N^2}, 1\right\}\right]$$

3. ดำเนินการตามข้อ 1-2 ซ้ำ จำนวน 1,000 ครั้งในแต่ละสถานการณ์

4. พิจารณาว่าช่วงความเชื่อมั่นที่ได้ในข้อ 2 คลุมค่าพารามิเตอร์ p หรือไม่

การพิจารณาว่าช่วงความเชื่อมั่นที่ได้นั้นคลุมค่าพารามิเตอร์ p หรือไม่นั้น พิจารณาจากช่วงความเชื่อมั่นใดที่ครอบคลุมค่าพารามิเตอร์ $L(p) \le p \le U(p)$ โดยกำหนดให้ช่วงความเชื่อมั่นที่คลุมค่าพารามิเตอร์ p เท่ากับ 1 และให้ช่วงความเชื่อมั่นที่ไม่คลุมค่าพารามิเตอร์ p เท่ากับ 0

5. คำนวณความกว้างของช่วงความเชื่อมั่น

นำช่วงความเชื่อมั่นที่คลุมค่าพารามิเตอร์ p มาคำนวณความกว้างของช่วงความเชื่อมั่น โดยจะคำนวณจากผลต่างระหว่างขีดจำกัดบน และขีดจำกัดล่าง สามารถคำนวณได้ดังนี้

ความกว้างของช่วงความเชื่อมั่น $= U - L$

เมื่อ U แทน ขีดจำกัดบน และ L แทน ขีดจำกัดล่าง จากนั้นทำการบวกสะสมไว้

6. คำนวณค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากวิธีการประมาณค่าแบบช่วงทั้ง 5 วิธี

6.1 การคำนวณค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ทำได้โดยการนับจำนวนครั้งของช่วงความเชื่อมั่นที่ค่าพารามิเตอร์ p ตกอยู่ในช่วงความเชื่อมั่นที่สร้างขึ้น แล้วนำมาหารด้วยจำนวนรอบของการทดลองซ้ำ นั่นคือ

ค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p = จำนวนครั้งของช่วงความเชื่อมั่นที่ค่าพารามิเตอร์ p ตกอยู่ในช่วงความเชื่อมั่นที่สร้างขึ้น / 1,000

6.2 การคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น ทำได้โดยการนำผลรวมของผลต่างระหว่างขีดจำกัดบน และขีดจำกัดล่างมาหารด้วยจำนวนรอบของการทดลองซ้ำ

7. เปรียบเทียบค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ที่ได้จากวิธีการประมาณค่าแบบช่วงทั้ง 5 วิธี กับระดับความเชื่อมั่นที่กำหนด ซึ่งเกณฑ์ที่ใช้ในการพิจารณาในลักษณะของการตรวจสอบว่าวิธีการประมาณค่าแบบช่วงวิธีใดให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ไม่น้อยกว่าระดับความเชื่อมั่นที่กำหนดนั้น จะดำเนินการด้วยการทดสอบสมมติฐาน

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

การทดสอบสมมติฐานในกรณีนี้เป็นการทดสอบสมมติฐานเกี่ยวกับความน่าจะเป็นของการเกิดผลสำเร็จสำหรับประชากรอนันต์ที่เกิดขึ้นภายใต้ฟังก์ชันความน่าจะเป็นแบบแบร์นูลลี หรือค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ ภายใต้สมมติฐานที่ว่า [Doganaksoy, N. and Schmee, J., 1993] $H_0 : 1-\alpha = 1-\alpha_0$ เทียบกับ $H_1 : 1-\alpha \neq 1-\alpha_0$ ที่ระดับนัยสำคัญ0.05 และสถิติที่ใช้ในการทดสอบ คือ สถิติทดสอบ Z

เมื่อพิจารณาความน่าจะเป็นของการยอมรับ H_0 เมื่อ H_0 เป็นจริง ที่ระดับนัยสำคัญ 0.05

$$\Pr\left(\left|\frac{(1\text{-}\hat{\alpha})-(1-\alpha_0)}{\sqrt{\frac{(1\text{-}\alpha_0)(\alpha_0)}{n}}}\right| \leq Z_{1-\alpha/2}\Big|_{1-\alpha=1-\alpha_0}\right) \approx 1-0.05$$

และด้วยคุณสมบัติของความสมมาตร จะได้ว่า

$$\Pr\left((1-\alpha_0)-Z_{0.975}\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}} \leq 1-\hat{\alpha} \leq (1-\alpha_0)+Z_{0.975}\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}}\right) \approx 0.95$$

ดังนั้น จะได้ว่าวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ p ทั้ง 5 วิธีดังกล่าว ให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ไม่น้อยกว่าระดับความเชื่อมั่นที่กำหนด ถ้า $1-\hat{\alpha}$ มีค่าดังนี้

$$(1-\alpha_0)-(1.96)\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}} \leq 1-\hat{\alpha} \leq (1-\alpha_0)+(1.96)\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}}$$

เมื่อ $1-\alpha_0$ แทน ระดับความเชื่อมั่นที่กำหนด ได้แก่ 0.95 และ 0.99

$1-\hat{\alpha}$ แทน ค่าความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ซึ่งได้จากการจำลองข้อมูล

n แทน จำนวนครั้งของการทดลอง 1, 000 ครั้ง

เมื่อกำหนดให้สร้างช่วงความเชื่อมั่น ที่ระดับความเชื่อมั่น 95%

จะเป็นการทดสอบว่าวิธีการประมาณค่าแบบช่วงทั้ง 5 วิธีดังกล่าว ให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p เท่ากับ 0.95 หรือไม่ ที่ระดับนัยสำคัญ 0.05 ภายใต้สมมติฐานที่ว่า $H_0 : 1\text{-}\alpha = 0.95$

$H_1 : 1\text{-}\alpha \neq 0.95$

สถิติที่ใช้ในการทดสอบ คือ สถิติทดสอบ Z

$$Z = \frac{(1\text{-}\hat{\alpha})-(1-\alpha_0)}{\sqrt{\frac{(1\text{-}\alpha_0)(\alpha_0)}{n}}}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ที่ระดับนัยสำคัญ 0.05 ดังนั้นหากต้องการสร้างช่วงความเชื่อมั่นที่นำมาใช้เป็นเกณฑ์ในการตัดสิน ให้พิจารณาจากความน่าจะเป็นในการยอมรับ H_0 เมื่อ H_0 เป็นจริง ที่ระดับนัยสำคัญ 0.05 คือ

$$\Pr\left((1-\alpha_0)-Z_{0.975}\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}} \le 1-\hat{\alpha} \le (1-\alpha_0)+Z_{0.975}\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}}\right) \approx 0.95$$

$$0.95-(1.96)\sqrt{[0.95(1-0.95)]/1{,}000} \le 1-\hat{\alpha} \le 0.95+(1.96)\sqrt{[0.95(1-0.95)]/1{,}000}$$

$$0.936 \le 1-\hat{\alpha} \le 0.964$$

เมื่อกำหนดให้สร้างช่วงความเชื่อมั่น ที่ระดับความเชื่อมั่น 99%

จะเป็นการทดสอบว่าวิธีการประมาณค่าแบบช่วงทั้ง 5 วิธีดังกล่าว ให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p เท่ากับ 0.99 หรือไม่ ที่ระดับนัยสำคัญ 0.05 ภายใต้สมมติฐานที่ว่า $H_0: 1\text{-}\alpha = 0.99$

$H_1: 1\text{-}\alpha \neq 0.99$

สถิติที่ใช้ในการทดสอบ คือ สถิติทดสอบ Z

$$Z = \frac{(1\text{-}\hat{\alpha})-(1-\alpha_0)}{\sqrt{\frac{(1\text{-}\alpha_0)(\alpha_0)}{n}}}$$

ที่ระดับนัยสำคัญ 0.05 ดังนั้นหากต้องการสร้างช่วงความเชื่อมั่นที่นำมาใช้เป็นเกณฑ์ในการตัดสิน ให้พิจารณาจากความน่าจะเป็นในการยอมรับ H_0 เมื่อ H_0 เป็นจริง ที่ระดับนัยสำคัญ 0.05 คือ

$$\Pr\left((1-\alpha_0)-Z_{0.975}\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}} \le 1-\hat{\alpha} \le (1-\alpha_0)+Z_{0.975}\sqrt{\frac{(1-\alpha_0)(\alpha_0)}{n}}\right) \approx 0.99$$

$$0.99-(1.96)\sqrt{[0.99(1-0.99)]/1{,}000} \le 1-\hat{\alpha} \le 0.99+(1.96)\sqrt{[0.99(1-0.99)]/1{,}000}$$

$$0.984 \le 1-\hat{\alpha} \le 0.996$$

ทั้งนี้ เนื่องจากช่วงความเชื่อมั่นที่ดีควรจะให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ไม่น้อยเกินไปหรือมากเกินไป การที่ช่วงความเชื่อมั่นให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์น้อยกว่าหรือมากกว่าระดับความเชื่อมั่นที่กำหนดแสดงว่าช่วงความเชื่อมั่นนั้นแคบหรือกว้างเกินไปจนทำให้การครอบคลุมค่าพารามิเตอร์เกิดความผิดพลาดขึ้น

8. เปรียบเทียบค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นเฉพาะวิธีการประมาณค่าแบบช่วงที่ให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ไม่น้อยกว่าระดับความเชื่อมั่นที่กำหนด โดยถ้าวิธีการประมาณค่าแบบช่วงวิธีใดให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุด จะถือว่าวิธีการประมาณค่าแบบช่วงวิธีนั้นเหมาะสมที่สุดสำหรับสถานการณ์นั้นๆ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

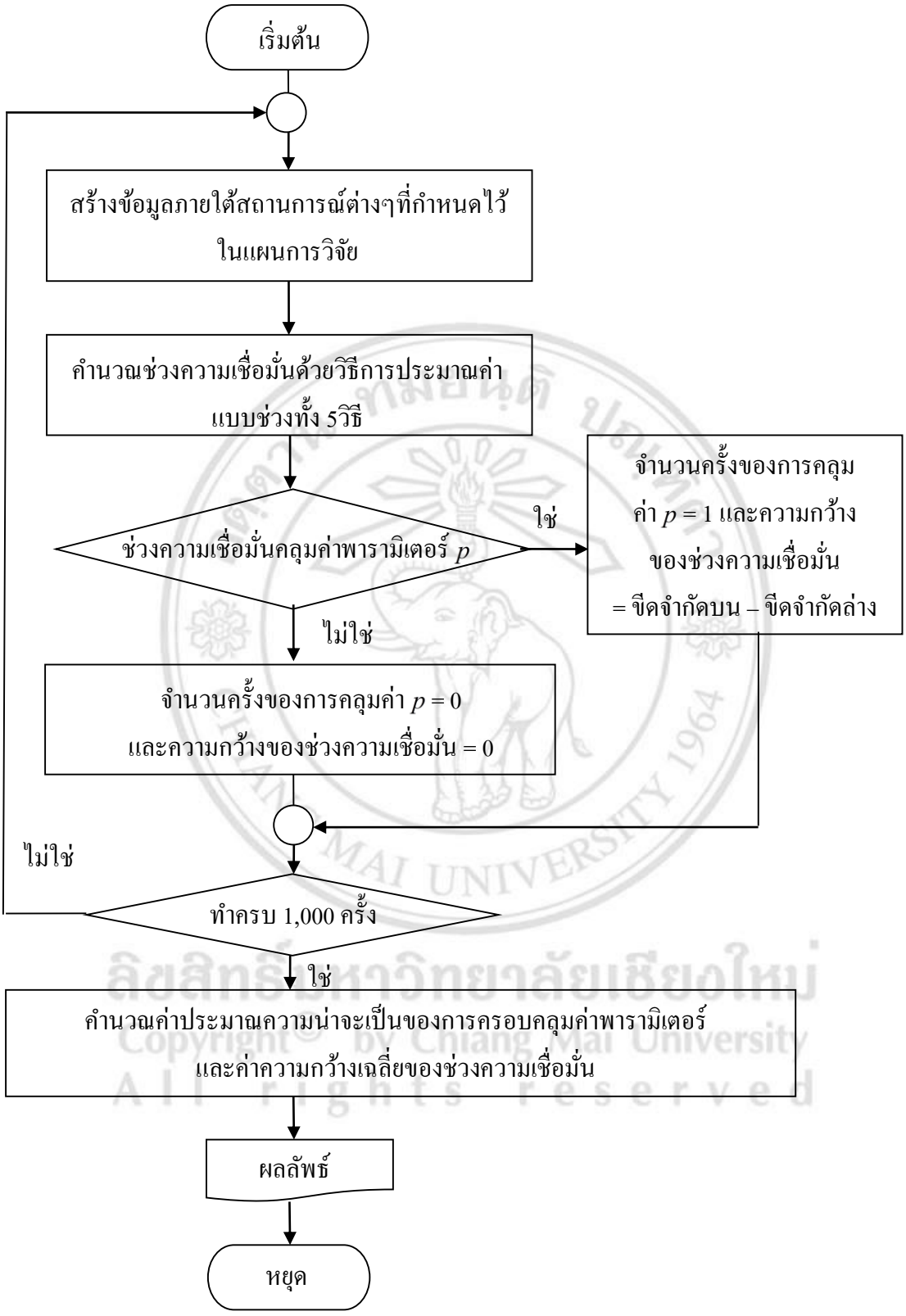
9. สรุปผลการวิจัย

การสรุปผลการวิจัยเราจะทำการสรุปผลการวิจัยว่าวิธีการประมาณค่าแบบช่วงวิธีใดที่ให้ค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p ไม่น้อยกว่าระดับความเชื่อมั่นที่กำหนด และให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดหรือแคบที่สุด วิธีการประมาณค่าแบบช่วงวิธีนั้นเป็นวิธีการประมาณค่าที่เหมาะสมที่สุดสำหรับแต่ละสถานการณ์ที่ได้กำหนดไว้ในแผนการวิจัย

3.4 ขั้นตอนการทำงานของโปรแกรม

ในการวิจัยครั้งนี้ ผู้วิจัยได้ใช้โปรแกรม MINITAB 14.0 ซึ่งเป็นโปรแกรมสำเร็จรูปทางสถิติมาใช้ในการสร้างข้อมูล คำนวณช่วงความเชื่อมั่น คำนวณค่าประมาณความน่าจะเป็นของการครอบคลุมค่าพารามิเตอร์ p และคำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น โดยขั้นตอนในการทำงานของโปรแกรมแสดงดังรูปภาพ 3.1

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved



รูป 3.2 ขั้นตอนการทำงานของโปรแกรม

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved