

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงทดลอง โดยต้องการศึกษาเปรียบเทียบวิธีการประมาณค่าสัดส่วนประชากรแบบช่วง ระหว่างวิธีการประมาณด้วยการแจกแจงแบบปกติเมื่อไม่ใช้ค่าปรับแก้ไขเพื่อความต่อเนื่อง วิธีการประมาณด้วยการแจกแจงแบบปกติเมื่อใช้ค่าปรับแก้ไขเพื่อความต่อเนื่อง และวิธีการประมาณด้วยการแจกแจงแบบเอฟ โดยวิธีการประมาณด้วยการแจกแจงแบบปกติทั้งกรณีที่ใช้และไม่ใช้ค่าปรับแก้ไขเพื่อความต่อเนื่องนั้น จะพิจารณาเปรียบเทียบในลักษณะที่ผสมผสานกับค่าความคลาดเคลื่อนมาตรฐานอีก 2 ค่า รวมเป็นการศึกษาเปรียบเทียบทั้งหมด 5 วิธี เพื่อศึกษาว่าวิธีการประมาณค่าสัดส่วนประชากรแบบช่วงวิธีใดดีที่สุดในกลุ่มของวิธีทั้งหมดที่ผู้วิจัยนำมาศึกษาเปรียบเทียบ โดยในขั้นตอนแรกจะศึกษาถึงระดับความเชื่อมั่นที่ได้จากแต่ละวิธีการประมาณก่อน แล้วจึงคัดเลือกวิธีการประมาณที่ให้ค่าระดับความเชื่อมั่นสูงกว่าหรือเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด มาเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นอีกทีหนึ่ง โดยทำการศึกษา ณ ระดับค่าสัมประสิทธิ์ความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99% ขนาดตัวอย่าง 49 ขนาดคือ n มีค่าตั้งแต่ 2 ถึง 50 และในแต่ละระดับขนาดตัวอย่างจะแปรค่า p ทั้งสิ้น 19 ค่า คือ p มีค่าตั้งแต่ 0.05 ถึง 0.95 โดยค่า p เพิ่มขึ้นทีละ 0.05

ในการวิจัยครั้งนี้ ใช้เทคนิควิธีการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Method) สร้างสถานการณ์การทดลอง โดยใช้โปรแกรมภาษาฟอร์แทรน 77 (FORTRAN 77) กับเครื่อง AMDAHL 5860 สำหรับแผนการทดลอง ขั้นตอนในการทดลอง และโปรแกรมที่ใช้ในการวิจัย จะนำเสนอตามลำดับดังต่อไปนี้

3.1 แผนการทดลอง

ได้กำหนดสถานการณ์ต่างๆ สำหรับศึกษาเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่นสำหรับค่าสัดส่วนประชากร จากวิธีการประมาณ 5 วิธี ดังนี้

3.1.1 กำหนดค่าสัมประสิทธิ์ความเชื่อมั่น 3 ระดับ คือ 90%, 95% และ 99%

3.1.2 ในแต่ละระดับค่าสัมประสิทธิ์ความเชื่อมั่น กำหนดขนาดตัวอย่าง n มีค่าตั้งแต่ 2 ถึง 50

3.1.3 ในแต่ละระดับค่าสัมประสิทธิ์ความเชื่อมั่น และแต่ละระดับขนาดตัวอย่าง จะแปรค่าสัดส่วนประชากร p เป็น 19 ค่า คือ p มีค่าตั้งแต่ 0.05 ถึง 0.95 โดยค่า p เพิ่มขึ้นทีละ 0.05

รวมเป็นการแปรสถานการณ์ทั้งสิ้น 2,793 สถานการณ์ โดยการเปรียบเทียบจะพิจารณาจากค่าระดับความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากวิธีการประมาณทั้ง 5 วิธี เพื่อหาวิธีการประมาณที่ดีที่สุดในแต่ละสถานการณ์ต่อไป

3.2 ขั้นตอนในการวิจัย

แบ่งออกเป็น 3 ขั้นตอน ดังนี้

3.2.1 การสร้างตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม

3.2.2 การคำนวณช่วงความเชื่อมั่นด้วยวิธีการประมาณ 5 วิธี

3.2.3 การหาค่าระดับความเชื่อมั่น และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น ซึ่งรายละเอียดสำหรับแต่ละขั้นตอนเป็นดังนี้

3.2.1 การสร้างตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม

การสร้างตัวแปรสุ่มที่มีลักษณะการแจกแจงแบบทวินามนั้น จะต้องใช้ตัวเลขสุ่มซึ่งมีการแจกแจงแบบสม่ำเสมอในช่วง $(0,1)$ เป็นพื้นฐานในการสร้าง โดยเลขสุ่มที่ได้ควรมีคุณสมบัติดังต่อไปนี้ (Banks 1984: 258-259)

- ก) ตัวเลขสุ่มที่ได้มีการกระจายของความน่าจะเป็นแบบสม่ำเสมอ และเป็นอิสระซึ่งกันและกัน
- ข) อนุกรมของตัวเลขที่ได้สามารถสร้างซ้ำได้ (reproducible)
- ค) อนุกรมของตัวเลขไม่ซ้ำเดิมในช่วงที่ต้องการใช้ตัวเลขแบบสุ่ม หมายความว่าขนาดความยาวของอนุกรมตัวเลขต้องยาวพอสำหรับใช้งาน
- ง) ใช้เวลาสั้น ๆ ในการสร้างตัวเลขแบบสุ่ม
- จ) ใช้หน่วยความจำของคอมพิวเตอร์น้อย

สำหรับโปรแกรมที่ใช้สร้างตัวเลขสุ่มซึ่งมีการแจกแจงแบบสม่ำเสมอในการวิจัยครั้งนี้ ได้สร้างเป็นฟังก์ชันย่อย คือ

FUNCTION RAND (IX) คังรายละเอียดที่ได้แสดงไว้ในภาคผนวก ซึ่งค่าของฟังก์ชัน RAND คือค่าตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอนั่นเอง

การแจกแจงแบบทวินาม มีฟังก์ชันความน่าจะเป็นในรูปของ

$$f(x; n, p) = \binom{n}{x} p^x q^{n-x} \quad ; x = 0, 1, 2, \dots, n$$

เมื่อ n และ p เป็นค่าพารามิเตอร์ของการแจกแจงแบบทวินาม โดยที่ n เป็นจำนวนครั้งของการทดลอง แต่ครั้งความน่าจะเป็นของการเกิดความสำเร็จเป็น p และความน่าจะเป็นของการเกิดความสำเร็จไม่สำเร็จเป็น $q = 1-p$ ค่าเฉลี่ยและค่าความแปรปรวนของการแจกแจงแบบทวินาม คือ np และ npq ตามลำดับ

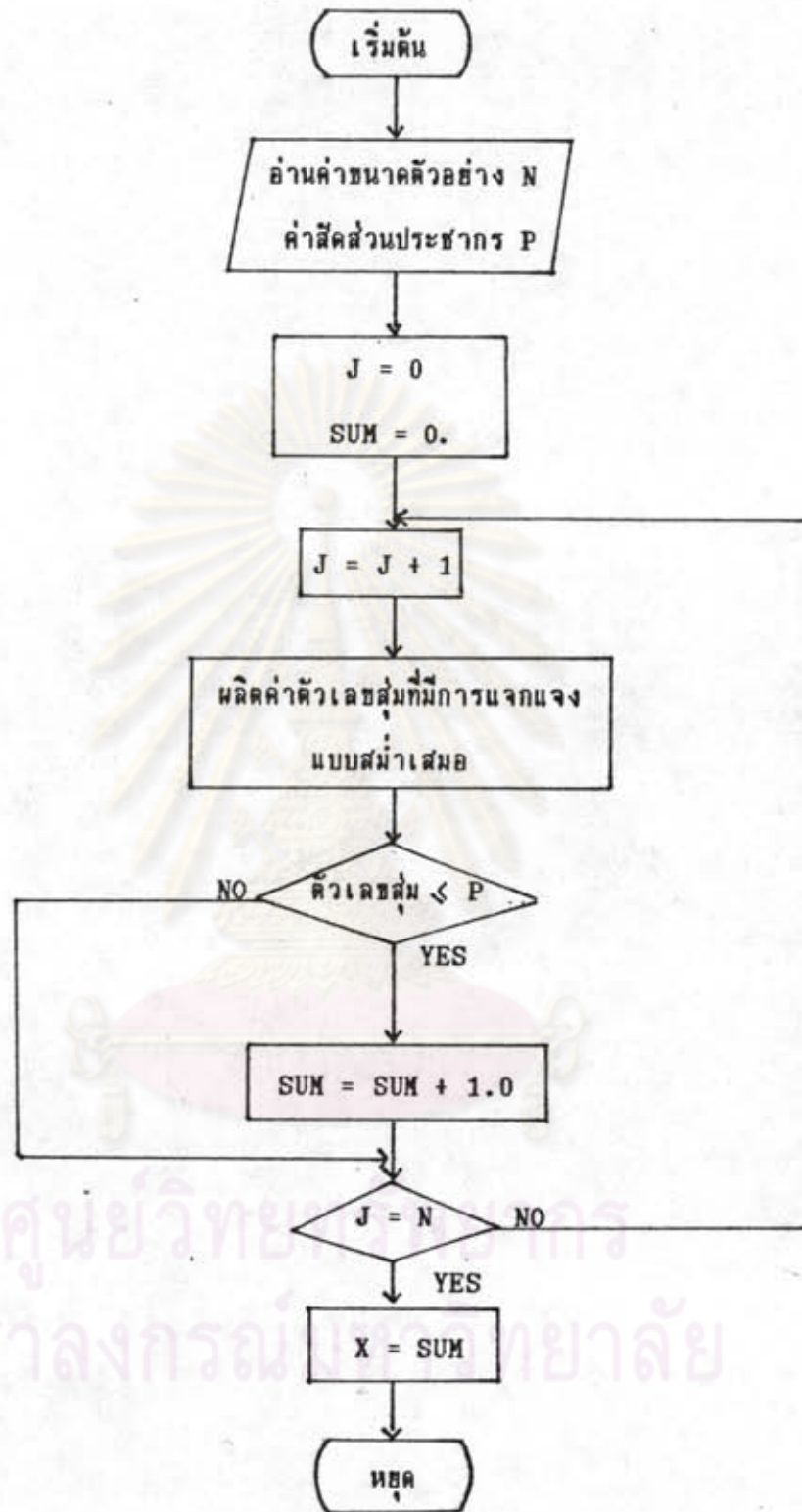
โปรแกรมที่ใช้ในการผลิตตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม ได้สร้างในลักษณะของโปรแกรมย่อย คือ

SUBROUTINE BIN (IX, N, P, X) ซึ่งรายละเอียดแสดงไว้ในภาคผนวก

การใช้โปรแกรมย่อยนี้ใช้คำสั่ง CALL BIN (IX,N,P,X) โดยค่า N เป็นจำนวนครั้งของการทดลองแบบทวินาม และค่า P เป็นค่าความน่าจะเป็นของการเกิดความสำเร็จในการทดลองแต่ละครั้ง ซึ่งก็คือค่าพารามิเตอร์ของการแจกแจงแบบทวินามที่ถูกกำหนดขึ้นและถูกส่งมาจากโปรแกรมหลัก การสร้างตัวแปรสุ่มที่มีการแจกแจงแบบทวินามนี้จะต้องเรียกใช้ตัวเลขสุ่มจาก FUNCTION RAND (IX) มาทำการสร้างเงื่อนไขว่าจะทำการนับสะสมเป็นผลสำเร็จถ้าตัวเลขสุ่มที่ถูกส่งมามีค่าน้อยกว่าหรือเท่ากับค่า P ที่กำหนด โดยจะทำการเรียกฟังก์ชันย่อย RAND ซ้ำ หรือผลิตค่าเลขสุ่มซ้ำเท่ากับค่าพารามิเตอร์ N ส่วนผลบวกสะสมที่ได้ให้เป็น X ซึ่งก็คือจำนวนครั้งของผลสำเร็จในการทดลอง N ครั้ง ฉะนั้น X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบทวินามนั่นเอง และในแต่ละขนาดตัวอย่าง N จะทำการผลิตตัวแปรสุ่มทวินามซ้ำ 2,000 ครั้ง

สรุปการสร้างค่าของตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม 1 ค่า ได้ด้วยผังลำดับงานของโปรแกรม รูป 3.1 ต่อไปนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูป 3.1

3.2.2 การคำนวณช่วงความเชื่อมั่นด้วยวิธีประมาณ 5 วิธี

3.2.2.1 การประมาณโดยใช้การแจกแจงแบบเอเฟ

การประมาณค่าสัดส่วนประชากรแบบช่วงโดยใช้การแจกแจงแบบเอเฟนี้ จะต้องอาศัยค่าตัวแปรสุ่มแบบเอเฟเป็นพื้นฐานในการสร้าง ซึ่งการวิจัยครั้งนี้ผู้วิจัยได้สร้างตารางการแจกแจงแบบเอเฟขึ้นใช้เอง โดยค่าตัวแปรสุ่มแบบเอเฟที่สร้างขึ้นได้จากความสัมพันธ์ระหว่างการแจกแจงแบบเอเฟ กับการแจกแจงแบบเบต้า

ในการประมาณค่าสัดส่วนประชากรแบบช่วงโดยใช้การแจกแจงแบบเอเฟ จากสูตรการประมาณ

$$\text{ขีดความเชื่อมั่นล่าง (PL)} = \frac{x}{x + (n - x + 1)F1} \dots\dots\dots (3.1)$$

$$\text{ขีดความเชื่อมั่นบน (PU)} = \frac{(x + 1)F2}{(x + 1)F2 + (n - x)} \dots\dots\dots (3.2)$$

เมื่อ x คือ จำนวนครั้งของผลสำเร็จในตัวอย่าง

n คือ ขนาดตัวอย่าง

$$F1 = F[1-\alpha/2; 2(n - x + 1), 2x]$$

$$F2 = F[1-\alpha/2; 2(x + 1), 2(n - x)]$$

ดังนั้น จะทำการคำนวณช่วงความเชื่อมั่นจากวิธีนี้ได้ จะต้องทราบค่า $F1$ และ $F2$ สำหรับโปรแกรมที่ใช้ผลิตค่าตัวแปรสุ่มแบบเอเฟ ได้สร้างในลักษณะของโปรแกรมย่อย คือ

SUBROUTINE MDBETI (PP,A,B,XX) ซึ่งรายละเอียดแสดงไว้ในภาคผนวก
ค่า PP คือค่า $1-\alpha/2$ ที่จะต้องกำหนดขึ้น เพื่อส่งไปยังโปรแกรมย่อย MDBETI ค่า A

และ B คือค่าพารามิเตอร์ค่าแรกและค่าที่สองของฟังก์ชันความหนาแน่นแบบเบตาไม่สมบูรณ์
ส่วนค่า XX คือค่าที่สอดคล้องกับสมการ

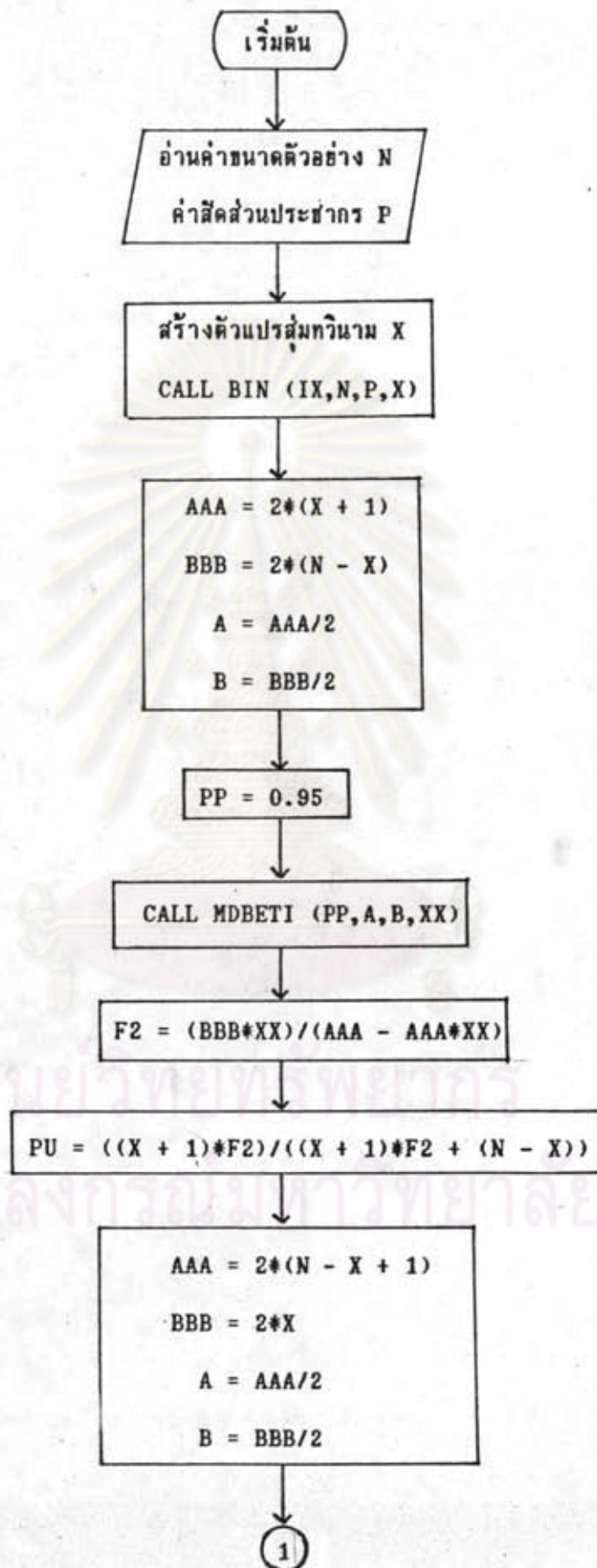
$$P(F \leq f) = I_{XX}^{(m/2, n/2)} = \frac{1}{B(m/2, n/2)} \int_0^{XX} t^{m/2-1} (1-t)^{n/2-1} dt$$

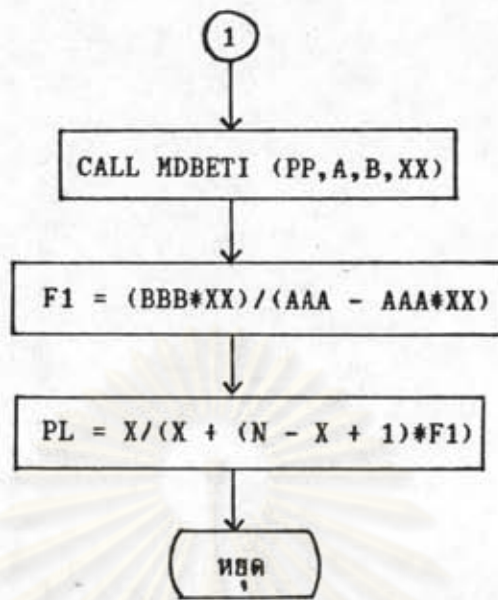
จากค่า XX ที่ได้ จะนำมาใช้คำนวณหาค่าของ $F1$ และ $F2$ ตามความสัมพันธ์
ระหว่างการแจกแจงแบบเอฟกับการแจกแจงแบบเบตา (ที่กล่าวไว้ในหัวข้อที่ 2.3)

สำหรับค่าของ x ในสมการ (3.1) และ (3.2) ได้จากการเรียกใช้โปรแกรม
ชื่อ BIN(IX, N, P, X)

หังลำดับงานรูป 3.2 แสดงขั้นตอนการทำงานของโปรแกรม เมื่อใช้การแจกแจง
แบบเอฟประมาณช่วงความเชื่อมั่น 95% หนึ่งช่วง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย





รูป 3.2 (ต่อ)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.2.2.2 การประมาณโดยใช้การแจกแจงแบบปกติ

เมื่อสร้างตัวแปรสุ่มที่มีลักษณะการแจกแจงแบบทวินามได้แล้ว การคำนวณช่วงความเชื่อมั่นสำหรับค่าสัดส่วนประชากร p ได้สร้างในลักษณะของโปรแกรมย่อย คือ

SUBROUTINE METHOD (IX,N,A,P,X,P190,P195,P199,P290,P295,P299,
*P390,P395,P399,P490,P495,P499)

ซึ่งรายละเอียดแสดงไว้ในภาคผนวก เมื่อเรียกใช้โปรแกรมย่อย METHOD จากโปรแกรมหลัก ค่า X ซึ่งเป็นจำนวนครั้งของผลสำเร็จในตัวอย่างไม่ขนาด N จะถูกส่งมายังโปรแกรมย่อย METHOD เพื่อทำการคำนวณช่วงความเชื่อมั่นด้วยวิธีการประมาณทั้ง 4 วิธีดังนี้

วิธีที่ 1 การประมาณค่าสัดส่วนประชากรแบบช่วงด้วยวิธีการแจกแจงแบบปกติ เมื่อไม่ใช้ค่าปรับแก้ไขเพื่อความต่อเนื่อง และใช้ความคลาดเคลื่อนมาตรฐานสำหรับค่า \hat{p} คือ $(\hat{p}(1-\hat{p})/n)^{1/2}$ มีขั้นตอนการคำนวณคือ

ขั้นที่ 1 นำจำนวนครั้งของผลสำเร็จหารด้วยขนาดตัวอย่าง ค่าที่ได้คือค่าสัดส่วนตัวอย่าง \hat{p}

ขั้นที่ 2 คำนวณขีดความเชื่อมั่นบนและขีดความเชื่อมั่นล่างที่แต่ละระดับความเชื่อมั่น

ระดับความเชื่อมั่น 90%

$$\text{จาก } (PL, PU) = (\hat{p} - Z_{1-\alpha/2} SE(\hat{p}), \hat{p} + Z_{1-\alpha/2} SE(\hat{p}))$$

$$PL = \hat{p} - 1.645(\hat{p}(1-\hat{p})/n)^{1/2}$$

$$PU = \hat{p} + 1.645(\hat{p}(1-\hat{p})/n)^{1/2}$$

ระดับความเชื่อมั่น 95%

$$PL = \hat{p} - 1.960(\hat{p}(1-\hat{p})/n)^{1/2}$$

$$PU = \hat{p} + 1.960(\hat{p}(1-\hat{p})/n)^{1/2}$$

ระดับความเชื่อมั่น 99%

$$PL = \hat{p} - 2.576(\hat{p}(1-\hat{p})/n)^{1/2}$$

$$PU = \hat{p} + 2.576(\hat{p}(1-\hat{p})/n)^{1/2}$$

วิธีที่ 2 การประมาณค่าสัดส่วนประชากรแบบช่วงด้วยการแจกแจงแบบปกติเมื่อไม่
ใช้ค่าปรับแก้ไขเพื่อความต่อเนื่อง และใช้ความคลาดเคลื่อนมาตรฐานสำหรับค่า \hat{p} คือ
 $(\hat{p}(1-\hat{p})/(n-1))^{1/2}$

ขั้นตอนการคำนวณวิธีนี้จะเหมือนกับวิธีที่ 1 เพียงแต่แทนค่า $(\hat{p}(1-\hat{p})/n)^{1/2}$
ด้วย $(\hat{p}(1-\hat{p})/(n-1))^{1/2}$

วิธีที่ 3 การประมาณค่าสัดส่วนประชากรแบบช่วงด้วยการแจกแจงแบบปกติเมื่อใช้ค่า
ปรับแก้ไขเพื่อความต่อเนื่อง และใช้ความคลาดเคลื่อนมาตรฐานสำหรับค่า \hat{p} คือ
 $(\hat{p}(1-\hat{p})/n)^{1/2}$

ค่าปรับแก้ไขเพื่อความต่อเนื่องที่ใช้ในการวิจัยครั้งนี้คือ ค่าปรับแก้ไขเพื่อความ
ต่อเนื่องของเฮจส์ (Yates's Correction for Continuity) ซึ่งมีค่าเท่ากับ $(2n)^{-1}$
มีการคำนวณดังนี้

ระดับความเชื่อมั่น 90%

$$\text{จาก } (PL, PU) = (\hat{p} - Z_{1-\alpha/2} SE(\hat{p}) - (2n)^{-1}, \hat{p} + Z_{1-\alpha/2} SE(\hat{p}) + (2n)^{-1})$$

$$PL = \hat{p} - \{ 1.645(\hat{p}(1-\hat{p})/n)^{1/2} + (2n)^{-1} \}$$

$$PU = \hat{p} + \{ 1.645(\hat{p}(1-\hat{p})/n)^{1/2} + (2n)^{-1} \}$$

ระดับความเชื่อมั่น 95%

$$PL = \hat{p} - \{ 1.960(\hat{p}(1-\hat{p})/n)^{1/2} + (2n)^{-1} \}$$

$$PU = \hat{p} + \{ 1.960(\hat{p}(1-\hat{p})/n)^{1/2} + (2n)^{-1} \}$$

ระดับความเชื่อมั่น 99%

$$PL = \hat{p} - \{ 2.576(\hat{p}(1-\hat{p})/n)^{1/2} + (2n)^{-1} \}$$

$$PU = \hat{p} + \{ 2.576(\hat{p}(1-\hat{p})/n)^{1/2} + (2n)^{-1} \}$$

วิธีที่ 4 การประมาณค่าสัดส่วนประชากรแบบช่วงด้วยการแจกแจงแบบปกติเมื่อใช้ค่าปรับแก้ไขเพื่อความต่อเนื่อง และใช้ความคลาดเคลื่อนมาตรฐานสำหรับค่า \hat{p} คือ $(\hat{p}(1-\hat{p})/(n-1))^{1/2}$

ขั้นตอนการคำนวณวิธีนี้ จะเหมือนกับวิธีที่ 3 เพียงแต่แทนค่า $(\hat{p}(1-\hat{p})/n)^{1/2}$ ด้วย $(\hat{p}(1-\hat{p})/(n-1))^{1/2}$

3.2.3 การหาค่าระดับความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น

ตามที่กล่าวมาแล้วว่า ในการพิจารณาคัดเลือกวิธีการประมาณจะพิจารณาเป็น 2 ขั้นตอน โดยในขั้นตอนแรกจะศึกษาถึงระดับความเชื่อมั่นของช่วงความเชื่อมั่นที่ได้จากแต่ละวิธีการประมาณ แล้วจึงคัดเลือกวิธีการประมาณที่ให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด มาทำการเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น ดังนั้นในขั้นตอนแรก ที่แต่ละระดับค่าพารามิเตอร์ (n, p) จากวิธีการประมาณทั้ง 5 วิธี เมื่อทำการคำนวณช่วงความเชื่อมั่นสำหรับค่า p ที่แต่ละระดับความเชื่อมั่นเสร็จเรียบร้อยแล้ว จะทำการตรวจสอบว่าช่วงความเชื่อมั่นที่คำนวณได้คลุมค่าพารามิเตอร์ p หรือไม่ หากช่วงความเชื่อมั่นของวิธีการประมาณใดคลุมค่าพารามิเตอร์ จะทำการนับจำนวนครั้งและบวกสะสมไว้ โดยในแต่ละขนาดตัวอย่าง n และแต่ละค่าสัดส่วนประชากร p จะคำนวณช่วงความเชื่อมั่นซ้ำ 2,000 ครั้ง ค่าบวกสะสมที่ได้คือจำนวนครั้งที่ทั้งหมดที่ช่วงความเชื่อมั่นคลุมค่า p เมื่อนำค่านี้หารด้วย 2,000 ค่าที่ได้คือค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่นที่คำนวณได้จากแต่ละวิธีการประมาณ ส่วนการคำนวณค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นทำได้โดยเมื่อคำนวณช่วงความเชื่อมั่นจากแต่ละวิธีการประมาณได้แล้ว จะหาผลต่างระหว่างขีดความเชื่อมั่นบนและขีดความเชื่อมั่นล่าง ผลต่างที่ได้จะบวกสะสมเก็บเอาไว้ แล้วหาค่าเฉลี่ยเมื่อทำการคำนวณช่วงความเชื่อมั่นครบ 2,000 ครั้ง เพื่อทำการเปรียบเทียบกันระหว่างวิธีต่าง ๆ การทดลองข้างต้นจะกระทำในแต่ละค่าพารามิเตอร์ (n, p) ที่กำหนดในงานวิจัยครั้งนี้

3.3 การเปรียบเทียบค่าระดับความเชื่อมั่นและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น

ในการพิจารณาค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่น กรณีในการพิจารณาว่าค่าระดับความเชื่อมั่นที่ได้จากการทดลองมีค่าไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด จะอาศัยจากการทดสอบสมมติฐานโดยใช้ตัวสถิติ Z ดังนั้นที่ระดับค่าสัมประสิทธิ์ความเชื่อมั่น 90%, 95% และ 99% หากวิธีการประมาณใดให้ค่าระดับความเชื่อมั่นจากการทดลองไม่ต่ำกว่า 0.8890, 0.9405 และ 0.9843 ตามลำดับ จะถือว่าวิธีการประมาณนั้นให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

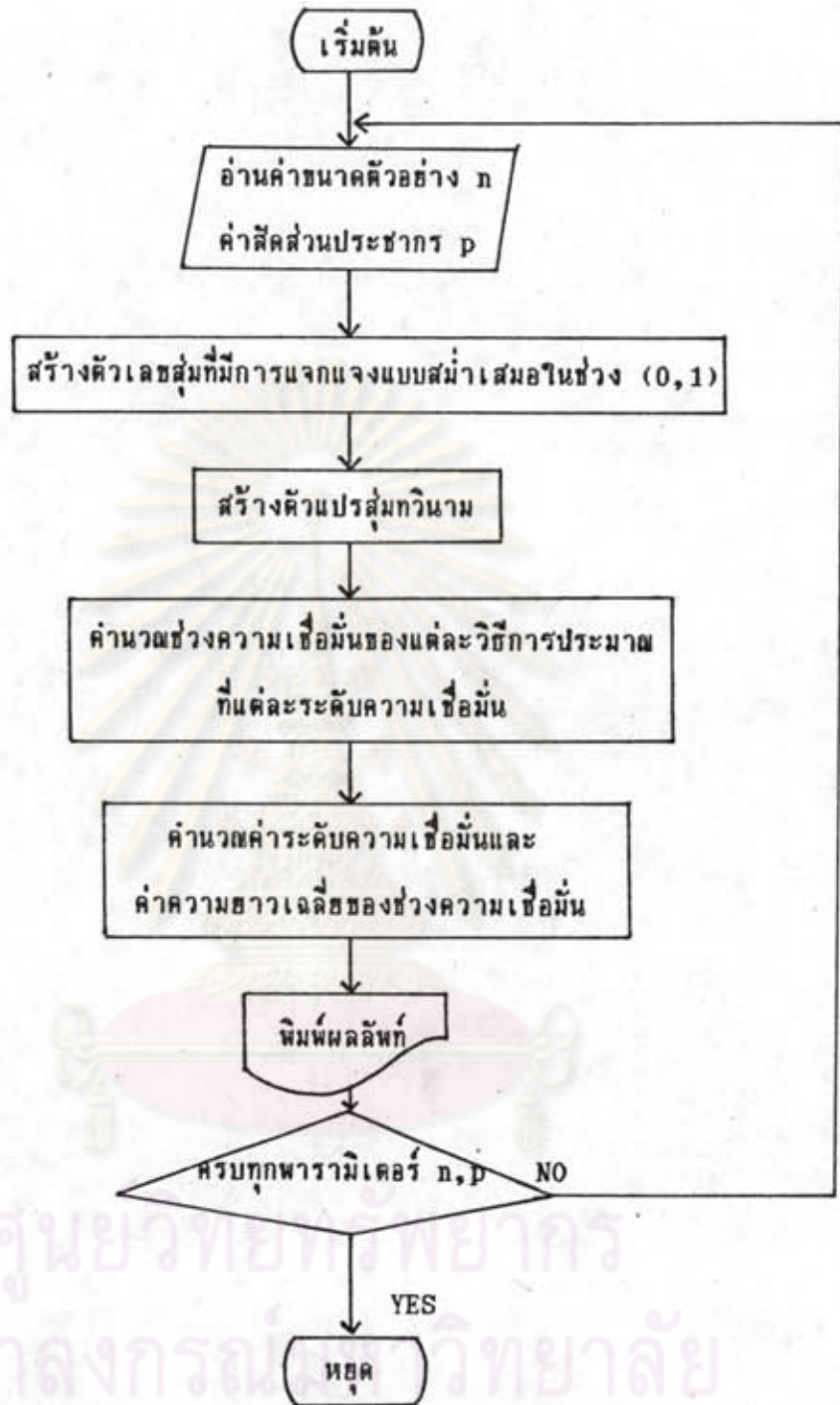
สำหรับการเปรียบเทียบค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากการทดลองนั้น จะเปรียบเทียบเฉพาะในกรณีที่ วิธีการประมาณนั้นให้ค่าระดับความเชื่อมั่นไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดเท่านั้น

3.4 ขั้นตอนในการทำงานของโปรแกรม

ขั้นตอนในการทำงานของโปรแกรมการหาค่าระดับความเชื่อมั่น และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น สรุปเป็นผังงานระบบได้ดังรูป 3.3



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูป 3.3