หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

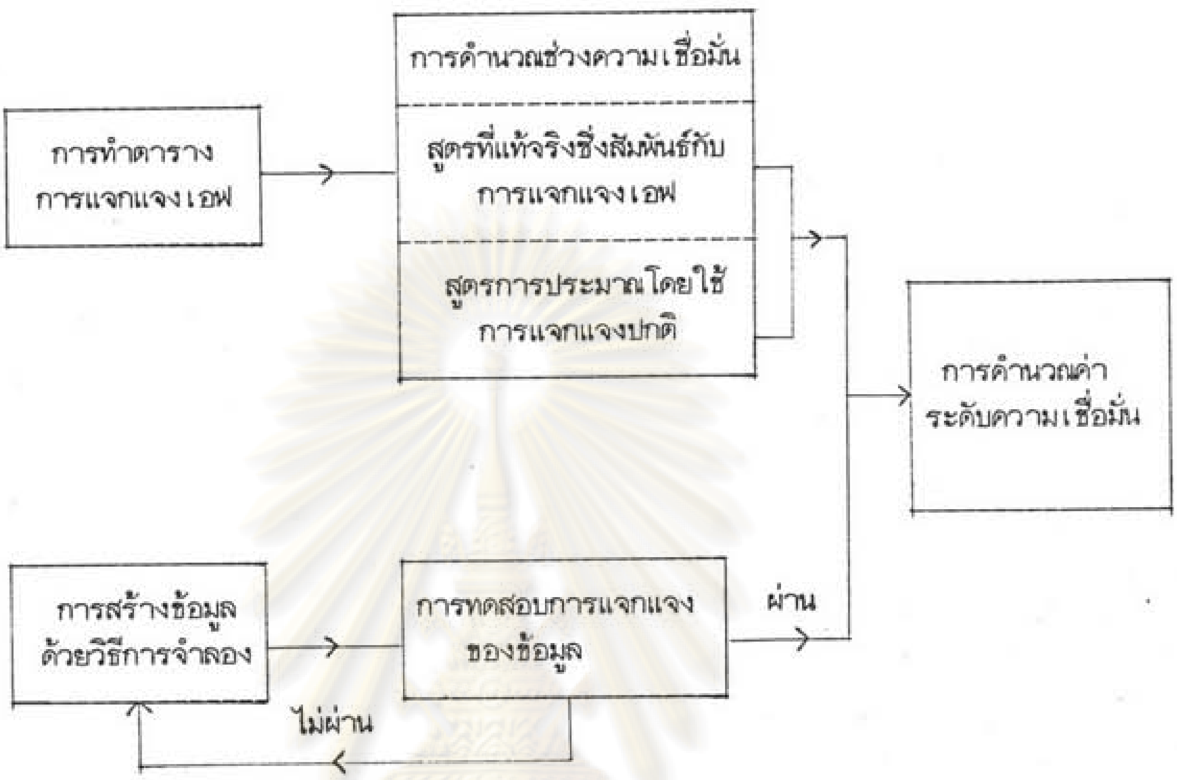
บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ ต้องการทำตารางสำเร็จรูป แสดงช่วงความเชื่อมั่นสำหรับค่าพารามิเตอร์ p ของประชากรที่มีการแจกแจงทวินาม โดยคำนวณจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ ในกรณีตัวอย่างขนาดเล็ก และมีเป้าหมายที่จะศึกษาความเหมาะสมและความเป็นไปได้ของการใช้สูตร การประมาณโดยใช้การแจกแจงปกติในการประมาณค่าพารามิเตอร์ p ดังกล่าวภายใต้เงื่อนไขตัวอย่างขนาดเล็ก

ขั้นตอนการดำเนินงานวิจัยสามารถสรุปออกมาดังรูป 3.1

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.1 แสดงแผนผังของการดำเนินงานวิจัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากแผนผังสามารถแบ่งการดำเนินงานออกเป็น 5 ส่วนดังนี้

1. การทำตารางการแจกแจงแบบเอฟ
2. การสร้างช่วงความเชื่อมั่น
3. การสร้างข้อมูล
4. การทดสอบข้อมูล
5. การคำนวณค่าระดับความเชื่อมั่น

รายละเอียดของงานทั้ง 5 ส่วน ตลอดจนเครื่องมือที่ต้องใช้ในแต่ละงาน อาทิ เช่น โปรแกรมคอมพิวเตอร์ และอื่น ๆ แสดงตามลำดับดังต่อไปนี้

3.1 <u>การทำตารางการแจกแจงความน่าจะเป็นแบบเอฟ</u>

เนื่องจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ ที่ใช้เป็นสูตรในการคำนวณช่วงความเชื่อมั่นสำหรับพารามิเตอร์ p ของประชากรที่มีการแจกแจงทวินาม ตามวัตถุประสงค์ของการวิจัยในครั้งนี้ มีความจำเป็นที่ต้องใช้ค่าตัวแปรสุ่ม ที่มีการแจกแจงความน่าจะเป็นแบบเอฟ ที่มีค่าองศาอิสระ และ ที่ค่าเปอร์เซนต์ไตล์ต่าง ๆ ซึ่งในการวิจัยได้กำหนดขอบเขตไว้ดังนี้

1. กำหนดขนาดตัวอย่าง $n = 1$ ถึง 30

สำหรับตัวแปรสุ่มเอฟที่มีค่าองศาอิสระเป็น (v1,v2) ภายใต้ค่าขนาดตัวอย่าง n ค่าหนึ่ง ๆ สูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ ต้องการค่าเอฟมาแทนค่าในสูตร สำหรับทุกค่า (v1,v2) ที่สอดคล้องกับสมการ

$$v1 + v2 = 2n + 2 \quad \text{เมื่อ } v1 \text{ และ } v2 \text{ เป็นเลขคู่บวก} \qquad (3.1.1)$$

ดังนั้นสำหรับค่า $n = 1$ ถึง 30 จะต้องใช้ค่าตัวแปรสุ่มเอฟที่มีค่าองศาอิสระ

(v1,v2) อยู่ภายในเซต

{ (v1,v2) | v1 + v2 $\leqslant$ 62 , v1 และ v2 เป็นเลขคู่บวก }

2. กำหนดค่าสัมประสิทธิ์ความเชื่อมั่น 8 ค่า คือ 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.98 และ 0.99

จากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟพบว่า จะต้องใช้ค่าตัวแปรสุ่มเอฟที่เปอร์เซนต์ไตล์เท่ากับ 75, 80, 85, 90, 95, 97.5, 99 และ 99.5 ในการแทนค่าลงในสูตร เพื่อให้ได้ค่าสัมประสิทธิ์ความเชื่อมั่นเป็น 0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.98 และ 0.99 ตามลำดับ ดังนั้นจึงต้องการตารางการแจกแจงความน่าจะเป็นแบบเอฟที่ระดับนัยสำคัญ 8 ระดับ คือ 0.25, 0.20, 0.15, 0.10, 0.05, 0.025, 0.01 และ 0.005

จากการรวบรวมตารางการแจกแจงความน่าจะเป็นแบบเอฟ ที่ระดับนัยสำคัญต่าง ๆ พบว่า ที่ระดับนัยสำคัญ 6 ระดับ คือ 0.25, 0.10, 0.05, 0.025, 0.01 และ 0.005 มีค่าเอฟไม่ครบตามที่ต้องการ จึงมีความจำเป็นที่จะต้องสร้างค่าเอฟเพิ่ม

ส่วนที่ระดับนัยสำคัญ 0.20 และ 0.15 พบว่า ยังไม่มีผู้ใดสร้างตารางที่ระดับนัยสำคัญนี้ จึงต้องทำตารางการแจกแจงความน่าจะเป็นแบบเอฟ ที่ระดับนัยสำคัญ 0.20 และ 0.15 ทั้งตาราง ทั้งนี้ค่าองศาอิสระจะอยู่ภายในขอบเขตที่กำหนดตามข้อ 1

ขั้นตอนในการทำตารางการแจกแจงความน่าจะเป็นแแบบเอฟ มีดังนี้

3.1.1 <u>การวิเคราะห์ปัญหา</u>

สำหรับตัวแปรสุ่ม X ที่มีการแจกแจงแบบเอฟ ให้

F แทนฟังก์ชันการแจกแจงแบบเอฟ

f แทนฟังก์ชันความน่าจะเป็นแแบบเอฟ

ฟังก์ชันการแจกแจงของตัวแปรสุ่ม X สามารถเขียนได้ในรูปอินทิกรัลของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ดังนี้

$$F(t) = Pr(X < t) = \int_{0}^{t} f(x)dx$$

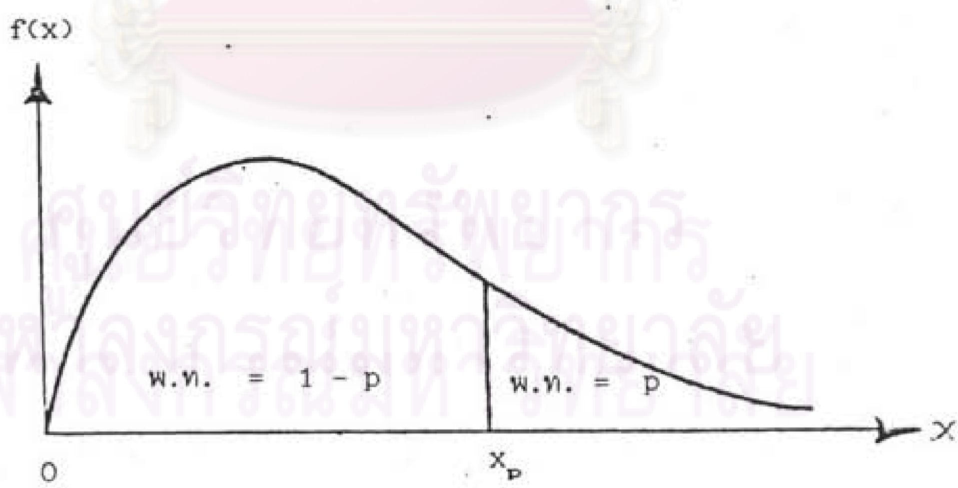
กำหนดให้ p เป็นค่าความน่าจะเป็นมีค่าอยู่ในช่วง [0,1]

x_p เป็นจุดที่ทำให้ $F(x_p) = p$

เราจะเรียก x_p ว่าเป็นค่าเปอร์เซนต์ไตล์ที่ 100p ของฟังก์ชันการแจกแจง F เป้าหมายคือการหาค่า x_p เมื่อกำหนดค่า p ให้ โดยหาจากสมการ

$$F(x_p) = p \qquad (3.1.2)$$

ความหมายเชิงเรขาคณิตของสมการ 3.1.2 แสดงดังรูป 3.2



รูปที่ 3.2 แสดงจุดเปอร์เซนต์ไตล์ที่ 100p ของการแจกแจงเอฟ

การแก้สมการ 3.1.2 คือการหาคำตอบของอินทิกรัล

$$\int_{0}^{x_p} f(x)dx = p \qquad (3.1.3)$$

เมื่อ f คือ ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ที่มีการแจกแจงเอฟ การแก้สมการ 3.1.3 ไม่อาจใช้วิธีการอินทิเกรตได้โดยตรง เนื่องจากฟังก์ชันเอฟมีความซับซ้อนมาก จำเป็นที่จะต้องอาศัยวิธีการอินทิเกรตเชิงตัวเลขมาช่วยในการหาคำตอบ

3.1.2 <u>วิธีการประมาณค่าอินทิกรัล</u>

วิธีที่ใช้ประมาณค่าอินทิกรัลในสมการ 3.1.3 มีด้วยกันหลายวิธี ในการวิจัยครั้งนี้ ได้เลือกใช้วิธีการประมาณโดยใช้กฎสี่เหลี่ยมคางหมู* ซึ่งเป็นวิธีที่ง่ายที่สุด

ขั้นตอนการประมาณค่าอินทิกรัลโดยวิธีกฎสี่เหลี่ยมคางหมู แสดงโดยโปรแกรมย่อย TRAP ดังรูป 3.3

```
      SUBROUTINE TRAP

      H = (XMAX-XMIN)/M
      SUM = 0.0
      X = XMIN + H

      DO 10 I  =  2,M,1
      SUM = SUM + F(X)
   10 X = X + H

      AREA = (H/2)*(F(XMIN) + 2*SUM + F(XMAX))
      RETURN
      END
```

รูปที่ 3.3 แสดงโปรแกรมย่อยที่ใช้คำนวณค่าอินทิกรัลโดยใช้กฎสี่เหลี่ยมคางหมู

---

*รายละเอียดของวิธีนี้กล่าวเอาไว้ในหัวข้อ 2.6 ของบทที่ 2

คำอธิบาย

1. [XMIN, XMAX] คือช่วงของการอินทิเกรต

2. M คือจำนวนค่า x ที่ใช้ในการประมาณค่าอินทิกรัล

3. H คือระยะห่างระหว่างค่า x คู่หนึ่ง ๆ ซึ่งจะต้องเท่ากันตลอดช่วงของการอินทิเกรต

4. F(x) คือค่าของฟังก์ชัน F ที่จุด x

5. AREA คือค่าประมาณของอินทิกรัล

3.1.3 <u>การคำนวณค่าเอฟ</u>

การหาค่า x_p ซึ่งเป็นคำตอบของสมการ 3.1.3 โดยใช้วิธีกฎสี่เหลี่ยมคางหมูในการประมาณค่าอินทิกรัล ไม่สามารถทำได้โดยตรง เนื่องจากวิธีกฎสี่เหลี่ยมคางหมูจำเป็นที่จะต้องระบุช่วงของการอินทิเกรตที่แน่ชัด แต่ปรากฏว่าจุด x_p ที่ต้องการหาค่านั้นคือค่าสุดท้ายของการอินทิเกรต ทำให้ระบุช่วงของการอินทิเกรตไม่ได้

วิธีแก้ไขที่ทำได้คือ การกำหนดระยะห่างระหว่างค่า x คู่หนึ่ง ๆ เสียก่อน แล้วจึงทำการคำนวณค่าอินทิกรัล โดยเพิ่มค่า x ไปทีละค่าจนกว่าจะได้ค่า x ที่ทำให้สมการ 3.1.3 เป็นจริง*

การวิจัยครั้งนี้ได้เลือกค่า H = 0.001 และได้ทำการดัดแปลงโปรแกรมย่อย TRAP เพื่อใช้ในการคำนวณหาค่า x_p ใหม่ เป็นโปรแกรมย่อย MOTRAP ดังรูป 3.4

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

---

*ในทางปฏิบัติ จะไม่สามารถหาค่า x_p ที่ทำให้สมการ 3.1.3 เป็นจริงได้ เนื่องจากเงื่อนไข "เท่ากับ" ในการเปรียบเทียบนิพจน์ทางคณิตศาสตร์ โดยการคำนวณด้วยคอมพิวเตอร์ มีโอกาสเกิดขึ้นน้อยมาก จนถือกันว่าเป็นเรื่องที่แทบจะเป็นไปไม่ได้ในหมู่นักคอมพิวเตอร์ วิธีแก้ไขที่นิยมทำกันคือ การเปรียบเทียบด้วยเงื่อนไข "มากกว่า-หรือเท่ากับ" แทน

```
      SUBROUTINE MOTRAP

      SUM = 0.0
      M = 1000.0*UPPERX
      X = X + 0.001

      DO 20 I  =  1,M,1
      SUM = SUM + 0.0005*F(Y) + 0.0005*F(X)
      IF(SUM.LT.P) GO TO 10
      GO TO 30
   10 X = X + 0.001

   20 CONTINUE
   30 XP = X
      RETURN
      END
```

รูปที่ 3.4 แสดงโปรแกรมย่อยที่ใช้คำนวณค่าเอฟ

คำอธิบาย

1. P คือค่าทางขวามือของสมการ 3.1.3

2. ต้องมีการกำหนดค่า UPPERX ซึ่งจะต้องมากกว่า x_p (ในโปรแกรมใช้ XP แทน x_p)

3. ในทางคอมพิวเตอร์ ไม่อาจเปรียบเทียบค่า SUM เท่ากับ P เนื่องจากเงื่อนไขดังกล่าวมีโอกาสเกิดน้อยมาก จึงต้องใช้การเปรียบเทียบค่า SUM น้อยกว่า P แทน และเมื่อเงื่อนไข SUM น้อยกว่า P ไม่เป็นจริง ก็จะได้ว่า ณ จุดนั้น ค่า SUM มากกว่าหรือเท่ากับ P ค่า x ที่จุดนี้ จะถูกกำหนดให้เป็นค่า x_p

4. ค่า x_p ที่ได้ คือ ค่าเอฟซึ่งเป็นเปอร์เซนต์ไตล์ที่ 100p และจะถูกบรรจุลงในตารางการแจกแจงความน่าจะเป็นแบบเอฟ ที่ระดับนัยสำคัญ 1 - p

3.1.4 การตรวจสอบความถูกต้อง

ในการตรวจสอบความถูกต้องเชื่อถือได้ ของค่าเอฟที่คำนวณได้ มีหลักในการตรวจสอบอยู่ 3 ประการที่สำคัญ ดังนี้

1. สำหรับกรณีที่ค่าองศาอิสระ (v1,v2) มีในตารางที่ปรากฏอยู่แล้ว

ถ้าค่าเอฟคำนวณที่เปอร์เซนต์ไตล์ $100p_1$ และ เปอร์เซนไตล์ $100p_2$ คำนวณได้เท่ากับค่าในตาราง จะถือว่าถูกต้อง และค่าเอฟที่คำนวณได้ที่เปอร์เซนไตล์ 100p ใด ๆ ซึ่งค่า p อยู่ภายในช่วง (p_1, p_2) ก็จะถือว่าถูกต้องด้วย

2. สำหรับกรณีที่ v1 มีในตาราง แต่ v2 ไม่มี

ให้พิจารณาความถูกต้องของค่าเอฟ ที่มีค่าองศาอิสระเป็น (v1,v2*) และ (v1,v2**) โดยที่ v2* น้อยกว่า v2** และต่างมีค่าอยู่ในตารางที่ปรากฏอยู่แล้ว ถ้าค่าทั้งสองถูกต้อง และ v2* < v2 < v2** จะถือว่าค่าเอฟดังกล่าวถูกต้อง

3. สำหรับกรณีที่ v1 ไม่มีในตาราง แต่ v2 มีในตาราง

ให้พิจารณาความถูกต้องของค่าเอฟ ที่มีค่าองศาอิสระเป็น (v1*,v2) และ (v1**,v2) โดยที่ v1* น้อยกว่า v1** และต่างมีค่าอยู่ในตารางที่ปรากฏอยู่แล้ว ถ้าค่าทั้งสองถูกต้อง และ v1* < v1 < v1** จะถือว่าค่าเอฟดังกล่าวถูกต้อง*

---

*ในกรณีที่ค่าองศาอิสระ (v1,v2) ไม่มีอยู่ในตารางที่ปรากฏอยู่แล้วทั้งคู่ การพิจารณาความถูกต้องของค่าเอฟจะต้องรอเอาผลของข้อ 1, 2 และ 3 มาใช้เป็นข้อฐานในการอ้างอิง จากนั้นอาศัยหลักของข้อ 2 และ 3 มาประยุกต์เข้าด้วยกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.2 การสร้างช่วงความเชื่อมั่น

สูตรที่ใช้ในการคำนวณช่วงความเชื่อมั่นสำหรับพารามิเตอร์ p ของประชากรที่มีการแจกแจงทวินาม ที่ใช้ในการวิจัยครั้งนี้ มี 2 สูตร คือ

1. สูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ
2. สูตรการประมาณโดยใช้การแจกแจงปกติ*

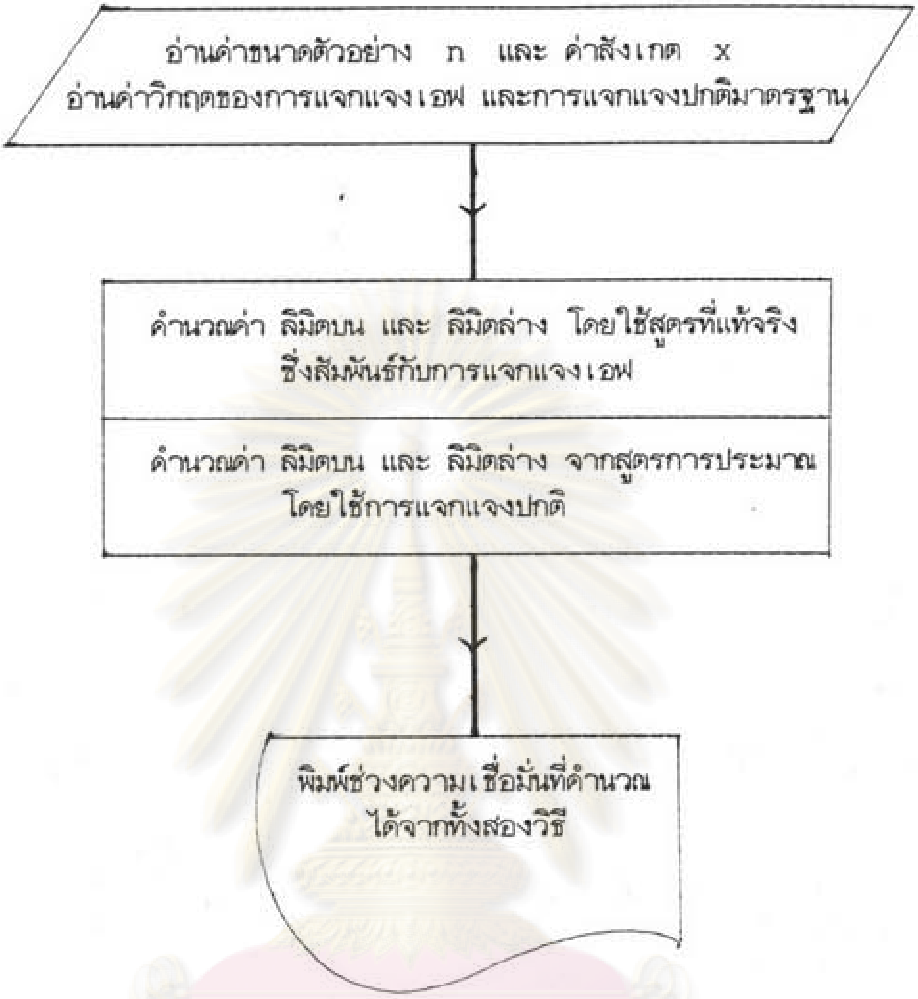
การคำนวณค่าลิมิตบน และลิมิตล่าง ของช่วงความเชื่อมั่นที่คำนวณจากสูตรทั้งสอง ได้เขียนเป็นโปรแกรมภาษา FORTRAN 77 โดยใช้กับเครื่องคอมพิวเตอร์ Micro VAX3600

ขั้นตอนการทำงานของโปรแกรมได้แสดงในรูปของผังงานดังรูป 3.5

---

*รายละเอียดของสูตรทั้งสอง กล่าวเอาไว้ในหัวข้อ 2.7 ของบทที่ 2

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร

รูปที่ 3.5 แสดงผังงานของโปรแกรมคำนวณค่า ลิมิตบน และ ลิมิตล่าง ของช่วงความเชื่อมั่น

จุฬาลงกรณ์มหาวิทยาลัย

สำหรับการคำนวณค่าลิมิตบน และลิมิตล่าง โดยสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ และสูตรการประมาณโดยใช้การแจกแจงปกติ แสดงดังโปรแกรมย่อย COMP1 และ COMP2 ดังรูป 3.6 และ 3.7 ตามลำดับ

```
   SUBROUTINE COMP1
   INTEGER K(31)
   REAL FL(31),FU(31),PL(31),PU(31)

   L   =  N + 1
   PL(1) =  0.0
   PU(1) =  (2*FU(1))/(2*N + 2*FU(1))
   PL(L) =  (2*N)/(2*FL(L) + 2*N)
   PU(L) =  1.0

   DO 10 I = 2,N,1

   PL(I) =  (2*K(I))/(2*(N-K(I)+1)*FL(I) + 2*K(I))
   PU(I) =  (2*(K(I)+1)*FU(I))/(2*(N-K(I)) + 2*(K(I)+1)*FU(I))

10 CONTINUE
   RETURN
   END
```

รูปที่ 3.6 แสดงโปรแกรมย่อยคำนวณค่าลิมิตบน และลิมิตล่างของช่วงความเชื่อมั่นที่คำนวณจาก สูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ

```
      SUBROUTINE COMP2
      INTEGER K(31)
      REAL Z(31),P(31),Q(31),CL(31),CU(31),ER(I)

      L  =  N + 1

      DO 10 I = 1,L,1

      P(I)  =  1.0*K(I)/N
      Q(I)  =  1.0 - P(I)
      ER(I) =  Z(I)*SQRT(P(I)*Q(I)/N)
      CL(I) =  P(I) - ER(I)
      CU(I) =  P(I) + ER(I)

   10 CONTINUE
      RETURN
      END
```

รูปที่ 3.7 แสดงโปรแกรมย่อยคำนวณค่าลิมิตบน และ ลิมิตล่างของช่วงความเชื่อมั่น ที่คำนวณจากสูตรการประมาณโดยใช้การแจกแจงปกติ

คำอธิบาย

1. N คือค่าขนาดตัวอย่าง n

2. K(I) คือค่าของตัวแปรชุดมิติ L = N + 1 ใช้แทนค่า X = 0,1,2, ... ,N โดยกำหนดให้ K(1)=0, K(2)=1, ... , K(L)=N

3. PL(I), PU(I) คือค่าลิมิตล่างและลิมิตบน ของช่วงที่คำนวณจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ

4. CL(I), CU(I) คือค่าลิมิตล่างและลิมิตบน ของช่วงที่คำนวณจากสูตรการประมาณโดยใช้การแจกแจงปกติ

5. I เป็นตัวดัชนีของตัวแปรชุด มีค่าตั้งแต่ 1 ถึง N + 1

ดังนั้นสำหรับค่าขนาดตัวอย่าง n ค่าใด ๆ สามารถสร้างช่วงความเชื่อมั่นที่แตกต่างกันได้ทั้งหมด n + 1 ช่วง ตามค่าที่เป็นไปได้ของค่าสังเกต x

6. ค่า FL(I) คือค่าของตัวแปรสุ่มเอฟ ที่มีค่าองศาอิสระเท่ากับ (2(n-x+1), 2x) : x = I - 1

FU(I) คือค่าของตัวแปรสุ่มเอฟ ที่มีค่าองศาอิสระเท่ากับ (2(x+1), 2(n-x)) : x = I - 1

ค่า FL(I) และ FU(I) หาได้จากตารางการแจกแจงความน่าจะเป็นแบบเอฟที่ระดับนัยสำคัญ $\alpha/2$

7. Z คือค่าของตัวแปรสุ่มปกติมาตรฐาน ที่ให้ค่าของฟังก์ชั่นความน่าจะเป็นสะสม เท่ากับ $1 - \alpha/2$

3.3 <u>การสร้างข้อมูล</u>

ข้อมูลที่ใช้ในการวิจัยครั้งนี้ ได้มาจากการจำลองขึ้นในเครื่องมินิคอมพิวเตอร์รุ่น Micro VAX3600 โดยเขียนโปรแกรมคอมพิวเตอร์ด้วยภาษา FORTRAN 77

ประชากรที่ทำการศึกษาในการวิจัยครั้งนี้ คือ ประชากรที่มีการแจกแจงทวินาม การสร้างลักษณะการแจกแจงของข้อมูลให้มีการแจกแจงแบบทวินาม จำเป็นจะต้องใช้ตัวเลขสุ่มที่มีการแจกแจงยูนิฟอร์มในช่วง (0,1) เป็นพื้นฐานในการสร้าง สำหรับโปรแกรมคอมพิวเตอร์ที่ใช้สร้างตัวเลขสุ่ม ในการวิจัยครั้งนี้ ได้เลือกใช้ตัวผลิตเลขสุ่มของ IMSL (International Mathematics and Statistical Libraries) ที่ชื่อ LLRANDOM* ซึ่งสามารถเขียนเป็นโปรแกรมย่อยฟังก์ชั่น RAND ดังรูป 3.8

---

*ตัวผลิตเลขสุ่ม LLRANDOM คิดค้นโดย P.A.W. Lewis, A.S. Goodman และ J.M. Miller เมื่อปี ค.ศ. 1969. ต่อมาได้ถูกพัฒนาโดย G.P. Learmonth และ P.A.W. Lewis ในปี ค.ศ. 1973.

```
FUNCTION RAND(IX)
INTEGER IX,FLT

IX   = IX*16807
IF(IX.LT.0) IX = 1 + (IX + 2147483647)
FLT  = IX
RAND = FLT*0.4656613E-9
RETURN
END
```

รูปที่ 3.8 แสดงโปรแกรมย่อยฟังก์ชัน ที่ใช้ผลิตค่าเลขสุ่ม

คำอธิบาย

IX คือ ค่าเริ่มต้นที่ต้องป้อนเข้าโปรแกรม ค่าของ IX จะเป็นเลขจำนวนเต็มบวกใด ๆ ก็ได้ แต่ต้องไม่เกินค่า 2147483647

RAND คือ ค่าของตัวแปรสุ่มที่มีการแจกแจงยูนิฟอร์มในช่วง (0,1)

วิธีในการสร้างค่าของตัวแปรสุ่มให้มีการแจกแจงเป็นรูปแบบต่าง ๆ มีด้วยกันหลายวิธี[1] ความเหมาะสมของแต่ละวิธี จะขึ้นอยู่กับชนิดของการแจกแจง และเงื่อนไขประกอบบางประการ สำหรับการสร้างค่าของตัวแปรสุ่มให้มีการแจกแจงทวินาม ซึ่งเป็นลักษณะการแจกแจงของประชากรที่ศึกษาในการวิจัยครั้งนี้ ได้เลือกใช้วิธี คอนโวลูชัน[2] ซึ่งมีหลักการโดยย่อเป็นดังนี้

---

[1]ศึกษาได้จาก Averill M. Law and W. David Kelton , Simulation Modeling and Analysis (New York: McGraw-Hill Book Co., 1982), pp. 242-252.

[2]เรื่องเดียวกัน, หน้า 249-250 และหน้า 266.

1. ในการสร้างค่าตัวแปรสุ่ม X ให้มีการแจกแจงทวินาม ที่มีพารามิเตอร์เท่ากับ (n.p) จะทำการสร้างค่าตัวแปรสุ่ม $Y_1, Y_2, \ldots, Y_n$ ซึ่งต่างเป็นอิสระจากกัน และมีการแจกแจงแบบเบอร์นูลี ที่มีพารามิเตอร์ p เหมือนกัน

2. คำนวณค่าของตัวแปรสุ่ม X โดยเอาค่าของตัวแปรสุ่ม Y ทั้ง n ตัว มาบวกเข้าด้วยกัน ผลบวกที่ได้จะกำหนดให้เป็นค่าของตัวแปรสุ่ม X

การสร้างค่าตัวแปรสุ่ม Y ให้มีการแจกแจงแบบเบอร์นูลี ที่มีพารามิเตอร์เท่ากับ p สามารถทำได้โดยวิธีการแปลงย้อนกลับ[3] ซึ่งเขียนเป็นโปรแกรมย่อยฟังก์ชั่น JBER ดังรูป 3.9

```
      FUNCTION JBER(P,IX)
      U = RAND(IX)
      IF(U.LE.P) GO TO 10
      JBER = 0
      GO TO 20
   10 JBER = 1
   20 RETURN
      END
```

รูปที่ 3.9 แสดงโปรแกรมย่อยฟังก์ชัน ที่ใช้ผลิตค่าตัวแปรสุ่มเบอร์นูลี

คำอธิบาย

P คือ ค่าความน่าจะเป็นที่จะเกิดความสำเร็จ เป็นค่าพารามิเตอร์ที่ต้องป้อนเข้าโปรแกรม พร้อมกับค่าเริ่มต้น IX

JBER คือ ค่าของตัวแปรสุ่มเบอร์นูลี ที่มีพารามิเตอร์เท่ากับ p ที่ผลิตได้

[3] เรื่องเดียวกัน, หน้า 242-247 และหน้า 263.

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

โปรแกรมย่อยฟังก์ชัน NBINIOM ในรูป 3.10 แสดงการผลิตค่าตัวแปรสุ่ม X ตามวิธีคอนโวลูชัน

```
      FUNCTION NBINOM(N,P,IX)

      M = 0
      DO 10 I = 1,N,1
      L = JBER(P,IX)
      M = M + L
   10 CONTINUE
      NBINOM = M
      RETURN
      END
```

รูปที่ 3.10 แสดงโปรแกรมย่อยฟังก์ชั่น ที่ใช้ผลิตค่าตัวแปรสุ่มทวินาม

คำอธิบาย

N และ P เป็นค่าพารามิเตอร์ที่ต้องป้อนเข้าโปรแกรมพร้อมกับค่าเริ่มต้น IX

NBINOM คือค่าของตัวแปรสุ่มทวินามที่มีพารามิเตอร์เท่ากับ (n,p) ที่ผลิตได้

3.4 การทดสอบข้อมูล

ปกติแล้ว ในการจำลองค่าตัวแปรสุ่มให้มีการแจกแจงตามรูปแบบที่กำหนด โดยอาศัยเลขสุ่มเป็นพื้นฐานในการจำลอง จะต้องมีการตรวจสอบว่า ค่าของตัวแปรสุ่มที่จำลองขึ้น มีรูปแบบของการแจกแจงตรงตามที่กำหนดจริงหรือไม่ ปัจจัยสำคัญที่มีผลกระทบต่อการแจกแจงของตัวแปรสุ่ม คือคุณสมบัติของตัวเลขสุ่มที่ใช้ในการจำลอง และวิธีการที่ใช้ในการสร้างค่าตัวแปรสุ่ม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตัวเลขสุ่มที่ดีจะต้องมีคุณสมบัติ 2 ประการคือ

1. ตัวเลขสุ่มต้องมีการแจกแจงแบบยูนิฟอร์ม ในช่วง (0,1)
2. ตัวเลขสุ่มแต่ละตัวจะต้องเป็นอิสระจากกัน

ในการวิจัยครั้งนี้ได้เลือกใช้ตัวผลิตเลขสุ่ม LLRANDOM ของ IMSL ซึ่งเพย์น[4] กล่าวว่า: " ตัวผลิตเลขสุ่ม LLRANDOM ได้ถูกทำการทดสอบคุณสมบัติความเป็นอิสระและความเป็นยูนิฟอร์ม โดยการทดสอบหลายชนิดอย่างกว้างขวาง และในการทดลองที่ทำกันมายังพบว่า ตัวผลิตเลขสุ่มนี้ ให้ค่าเลขสุ่มที่ใกล้เคียงกับเลขสุ่มในอุดมคติ* มากกว่าตัวผลิตเลขสุ่มใด ๆ ที่เคยมีมา ".

การเลือกใช้ตัวผลิตเลขสุ่ม LLRANDOM จึงสามารถค้ำประกันผลกระทบจากคุณสมบัติของตัวเลขสุ่มที่ใช้ในการจำลองได้เป็นอย่างมาก จึงไม่จำเป็นต้องเสียเวลาในการตรวจสอบคุณสมบัติของตัวเลขสุ่มอีก

สำหรับวิธีในการสร้างค่าตัวแปรสุ่มให้มีการแจกแจงแบบทวินาม การวิจัยครั้งนี้ได้เลือกใช้วิธีคอนโวลูชัน ซึ่งมีทฤษฎีทางสถิติสนับสนุนอย่างหนักแน่น สามารถใช้ค้ำประกันความเชื่อถือได้ของวิธีการสร้างค่าตัวแปรสุ่มทวินามได้เป็นอย่างดี

---

[4]Jame A. Payne, <u>Introduction to Simulations Programing Techniques and Method of Analysis</u> (New York: McGraw-Hill Book Co., 1982), pp. 311.

*ตัวเลขสุ่มในอุดมคติ คือตัวเลขสุ่มที่มีการแจกแจงยูนิฟอร์มในช่วง (0,1) และเป็นอิสระจากกันอย่างแท้จริง

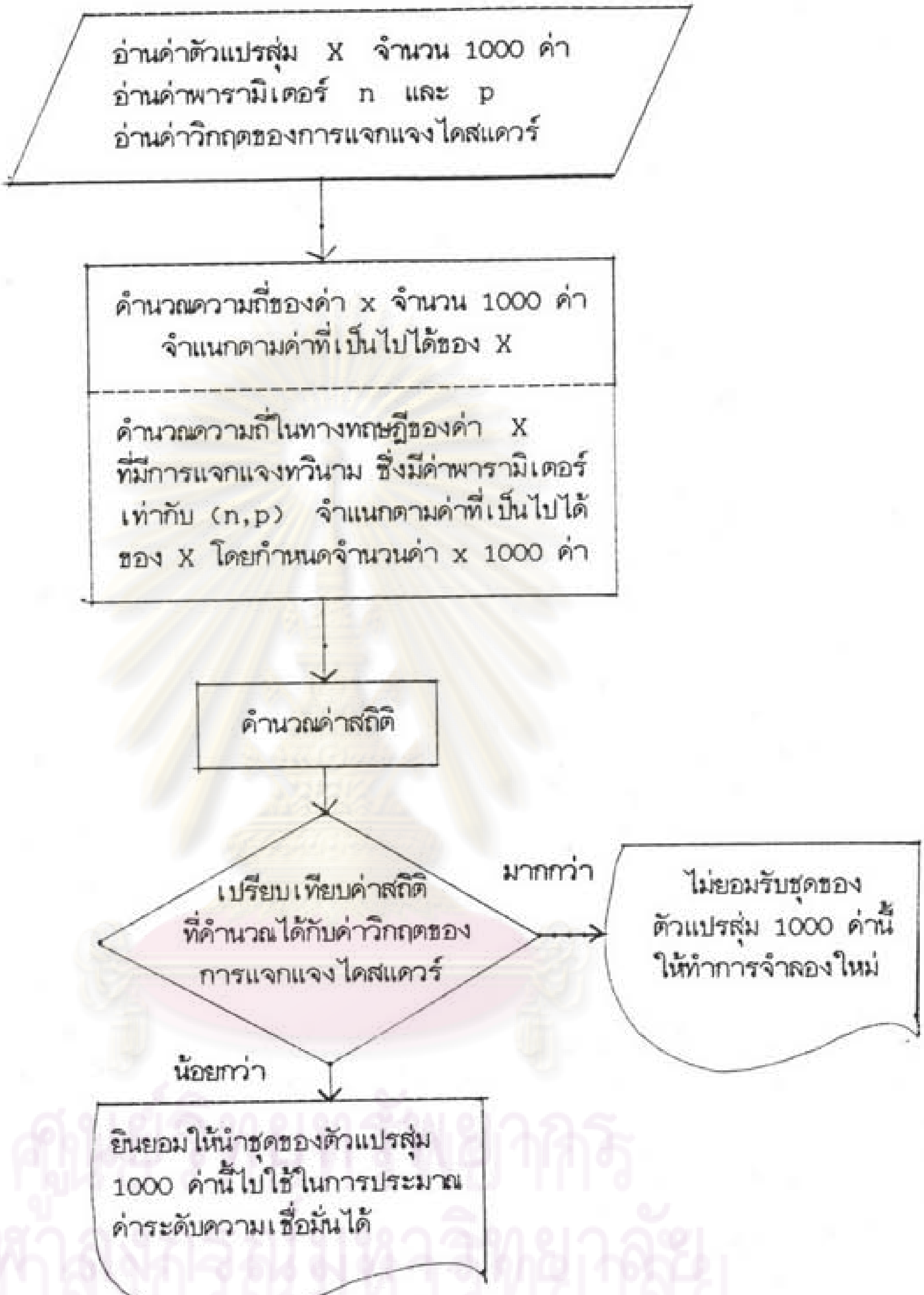
อย่างไรก็ตาม ยังคงมีความจำเป็นที่จะต้องทำการตรวจสอบข้อมูลที่สร้างขึ้น ซึ่งเป็นค่าของตัวแปรสุ่มที่กำหนดให้มีการแจกแจงแบบทวินาม ว่ามีการแจกแจงแบบทวินามหรือไม่ เนื่องจากข้อมูลดังกล่าว ต้องถูกนำไปใช้ในการประมาณค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่นที่สร้างขึ้น การแจกแจงของข้อมูล จะมีผลกระทบต่อระดับความเชื่อมั่นที่ประมาณได้ ถ้าข้อมูลไม่ได้มีการแจกแจงแบบทวินาม การสรุปผลเกี่ยวกับค่าระดับความเชื่อมั่นย่อมเชื่อถือไม่ได้ และอาจนำไปสู่ข้อสรุปที่ผิดพลาดได้ง่าย

การวิจัยครั้งนี้ได้เลือกใช้ วิธีการทดสอบไคสแควร์* ในการตรวจสอบการแจกแจงของข้อมูลที่สร้างขึ้น ขั้นตอนของการทดสอบแสดงในรูปผังงาน ดังรูป 3.11

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

---

*รายละเอียดของวิธีการทดสอบไคสแควร์ ได้แสดงไว้ในหัวข้อ 2.5



รูปที่ 3.11 แสดงขั้นตอนการทดสอบค่าของตัวแปรสุ่มที่ได้จากการจำลอง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.5 การคำนวณค่าระดับความเชื่อมั่น

การคำนวณค่าระดับความเชื่อมั่น ของช่วงความเชื่อมั่นที่คำนวณจากสูตรการประมาณด้วยการแจกแจงปกติ และสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ จัดเป็นงานชิ้นสุดท้ายของวิธีดำเนินงานวิจัย ผลลัพธ์ที่ได้จากงานวิจัยนี้ คือผลการวิจัย ซึ่งจะถูกนำไปวิเคราะห์ตามหลักวิชาทางสถิติ ก่อนที่จะทำการสรุปผลการวิจัย

เนื่องจากค่าระดับความเชื่อมั่น ของช่วงความเชื่อมั่นที่สร้างขึ้นสำหรับการประมาณค่าแบบช่วงของพารามิเตอร์ p ในประชากรที่มีการแจกแจงทวินาม จะขึ้นกับค่าพารามิเตอร์ (n,p) ของตัวแปรสุ่ม X ที่ใช้กำหนดประชากร ดังนั้นสำหรับค่าของพารามิเตอร์ (n,p) คู่ใด ๆ จะสามารถคำนวณค่าระดับความเชื่อมั่นได้ 1 ค่า เมื่อเทียบกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด 1 ค่า

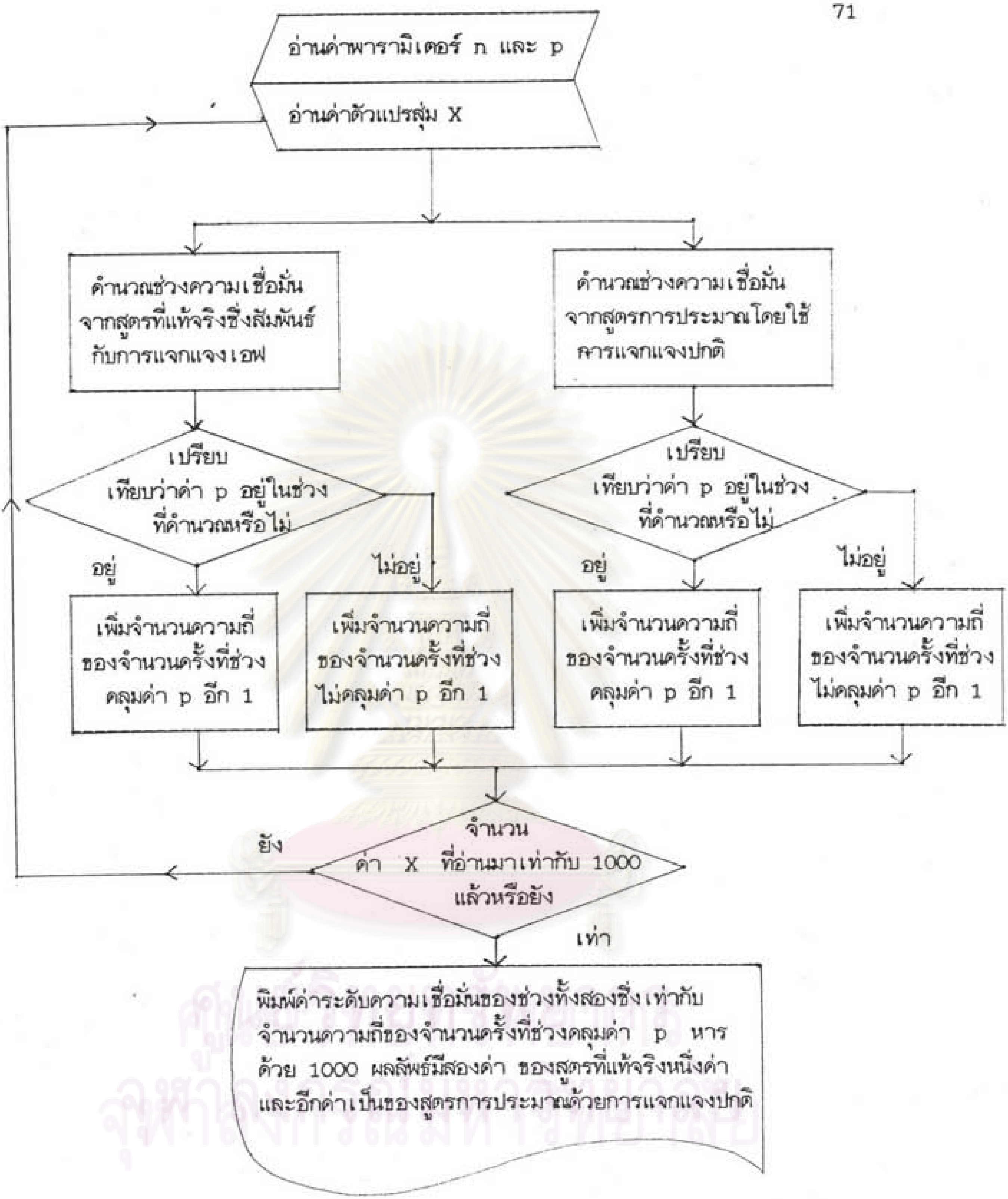
ในการวิจัยครั้งนี้ ได้กำหนดขอบเขตของค่าพารามิเตอร์ไว้ดังนี้

1. ขนาดตัวอย่าง n จำนวน 30 ค่า มีค่าตั้งแต่ 1 ถึง 30
2. ค่าสัดส่วนของความสำเร็จในประชากร p จำนวน 19 ค่า มีค่าตั้งแต่ 0.05 เพิ่มค่าครั้งละ 0.05 จนถึง 0.95

และได้กำหนดขอบเขตของค่าสัมประสิทธิ์ความเชื่อมั่น จำนวน 8 ค่า ปริมาณงานในการคำนวณค่าระดับความเชื่อมั่นจึงเท่ากับ $30 \times 19 \times 8 = 4560$ ครั้ง

ขั้นตอนในการคำนวณค่าระดับความเชื่อมั่น 1 ครั้ง แสดงโดยผังงาน ดังรูป 3.12

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.12 แสดงขั้นตอนของการคำนวณค่าระดับความเชื่อมั่น

จากผังงาน จะพบว่า ในการคำนวณค่าระดับความเชื่อมั่น 1 ครั้ง จะได้ผลลัพธ์เป็นค่าระดับความเชื่อมั่น 2 ค่า โดยเป็นของช่วงความเชื่อมั่นที่คำนวณจากสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ 1 ค่า และอีกค่าเป็นของช่วงความเชื่อมั่นที่คำนวณจากสูตรการประมาณด้วยการแจกแจงปกติ

สำหรับการวิจัยครั้งนี้ ได้เลือกใช้วิธีการจำลองค่าตัวแปรสุ่ม X ชุดใหม่ทุกครั้งที่ใช้ในการคำนวณค่าระดับความเชื่อมั่นแต่ละครั้ง เพื่อให้ค่าระดับความเชื่อมั่นที่คำนวณได้สำหรับแต่ละค่าสัมประสิทธิ์ความเชื่อมั่นเป็นอิสระจากกัน ภายใต้ค่าพารามิเตอร์ (n,p) คู่หนึ่ง ๆ โดยมีเหตุผลดังต่อไปนี้

1. เนื่องจากผลของการศึกษาที่ได้จากวิธีการจำลอง จะถูกกำหนดโดยเงื่อนไขเริ่มต้นของการจำลอง สำหรับในการวิจัยครั้งนี้ เงื่อนไขดังกล่าว คือ ค่าเริ่มต้นที่ป้อนเข้าโปรแกรมที่ใช้สร้างเลขสุ่ม ซึ่งกำหนดโดยผู้วิจัย ความเชื่อถือได้ของผลการวิจัย จะขึ้นอยู่กับค่าเริ่มต้นที่ผู้วิจัยเป็นผู้กำหนดในลักษณะของการส่งผลกระทบต่อกันเป็นทอด ๆ ดังนี้

ก. ค่าของเลขสุ่มจะขึ้นกับค่าเริ่มต้นที่ป้อนเข้าโปรแกรม

ข. ค่าของตัวแปรสุ่ม X จะขึ้นกับค่าของเลขสุ่ม

ค. ค่าระดับความเชื่อมั่นจะขึ้นอยู่กับชุดของตัวแปรสุ่ม X จำนวน 1000 ค่า

ถ้าชุดของตัวแปรสุ่มจำนวน 1000 ค่าดังกล่าว ไม่ได้เป็นตัวแทนที่ดีของประชากร ค่าระดับความเชื่อมั่นที่คำนวณได้ อาจแตกต่างจากความเป็นจริงมาก ทำให้ผลการวิจัยที่ได้บิดเบือนไปจากความจริง

2. วิธีการจำลองค่าตัวแปรสุ่ม X เพียง 1 ชุด เพื่อใช้ในการคำนวณค่าระดับความเชื่อมั่น สำหรับค่าสัมประสิทธิ์ความเชื่อมั่นทั้ง 8 ค่า จะมีข้อเสียที่เห็นได้ชัดเจนจากเหตุผลข้อ 1 กล่าวคือ ถ้าข้อมูลหรือค่าของชุดตัวแปรสุ่มไม่ได้เป็นตัวแทนที่ดีของประชากร การสรุปผลจะผิดพลาดได้ง่าย เนื่องจากผลของการคำนวณค่าระดับความเชื่อมั่นของช่วงความเชื่อมั่นทั้ง 8 ช่วง จะขึ้นอยู่กับชุดของตัวแปรสุ่ม X เพียงชุดเดียว ทำให้ผลการคำนวณของช่วงความเชื่อมั่นทั้ง 8 ช่วง ออกมาในทิศทางเดียวกัน

3. วิธีการจำลองค่าตัวแปรสุ่ม X ชุดใหม่ทุกครั้งที่ทำการคำนวณค่าระดับความเชื่อมั่น สำหรับค่าสัมประสิทธิ์ความเชื่อมั่นแต่ละค่า จากทั้งหมด 8 ค่า จะช่วยแก้ปัญหาที่เกิดในข้อ 2 ได้ ในกรณีที่มีชุดของตัวแปรสุ่มบางชุด ไม่ได้เป็นตัวแทนที่ดีของประชากร เนื่องจากผลการคำนวณของช่วงความเชื่อมั่นทั้ง 8 ช่วง จะมีความขัดแย้งในด้านทิศทางปรากฏให้เห็น โดยผลที่ได้จะไม่ออกมาในทิศทางเดียวกัน ทำให้โอกาสที่จะทำการสรุปผลผิดพลาดมีน้อยลง

หัวข้อ 3.1 ถึง 3.5 ได้บรรยายขั้นตอนการดำเนินการวิจัยไว้อย่างละเอียด เทคนิคทางคณิตศาสตร์ที่ใช้ในการวิจัยครั้งนี้ มีชื่อเรียกว่า เทคนิคมอนติคาร์โล

วิธีมอนติคาร์โล เป็นสาขาหนึ่งของคณิตศาสตร์เชิงทดลอง ที่ใช้แก้ปัญหาทางคณิตศาสตร์ โดยใช้ตัวเลขสุ่มที่มีการแจกแจงยูนิฟอร์มในช่วง (0,1) เป็นเครื่องมือในการคำนวณ เพื่อหาคำตอบของปัญหา วิธีการคำนวณใด ๆ ก็ตาม ที่ใช้ตัวเลขสุ่มเป็นเครื่องมือในการหาคำตอบ จะเรียกว่าวิธีมอนติคาร์โลทั้งสิ้น

โดยปกติการใช้เทคนิคมอนติคาร์โล จะประกอบไปด้วยขั้นตอน 3 ขั้นตอนดังนี้

1. การสร้างตัวเลขสุ่ม

การใช้ตัวเลขสุ่ม เป็นสิ่งที่สำคัญมากในวิธีมอนติคาร์โล ลักษณะของตัวเลขสุ่มจะมีการแจกแจงแบบยูนิฟอร์มในช่วง (0,1) สำหรับวิธีการสร้างตัวเลขสุ่ม ได้มีพัฒนาการมาตั้งแต่ในปี ค.ศ. 1940 เป็นต้นมา โดยพยายามคิดค้นวิธีการสร้างตัวเลขสุ่มที่ให้ชุดของตัวแปรสุ่มจำนวนมากตัวที่สุด ก่อนที่จะเกิดการซ้ำของวงจรการผลิต ทั้งนี้ตัวเลขสุ่มแต่ละตัวจะเป็นอิสระจากกัน และมีการแจกแจงยูนิฟอร์มในช่วง (0,1)

2. การนำตัวเลขสุ่มไปประยุกต์ใช้กับปัญหาที่ต้องการศึกษา

ขั้นตอนนี้จะขึ้นอยู่กับลักษณะของปัญหาที่ต้องการศึกษา บางปัญหาอาจจะไม่ใช้ตัวเลขสุ่มโดยตรง แต่จะมีขั้นตอนบางขั้นตอนในปัญหานั้นที่จะต้องใช้ตัวเลขสุ่ม

3. การทดลองกระทำซ้ำโดยใช้กระบวนการของการสุ่ม

เมื่อประยุกต์ปัญหาให้ใช้กับตัวเลขสุ่มได้แล้ว ขั้นตอนต่อไปคือ การทำการทดลองโดยอาศัยกระบวนการของการสุ่ม มากระทำในลักษณะที่ซ้ำ ๆ กัน เพื่อหาคำตอบของปัญหาที่ต้องการศึกษา

จากขั้นตอน 3 ขั้นตอนของวิธีมอนติคาร์โล ซึ่งเป็นเทคนิคที่ใช้แก้ปัญหาในการวิจัยครั้งนี้ จะพบว่า

ก. ขั้นตอนที่ 1 จะอยู่ในหัวข้อ 3.3 การสร้างข้อมูล

ข. ขั้นตอนที่ 2 คือ หัวข้อ 3.2 การสร้างช่วงความเชื่อมั่น หัวข้อ 3.3 (ในส่วนที่เหลือจากขั้นตอนที่ 1) และหัวข้อ 3.4 การทดสอบข้อมูล

ค. ขั้นตอนที่ 3 คือ หัวข้อ 3.5 การคำนวณค่าระดับความเชื่อมั่น

สำหรับหัวข้อ 3.1 การทำตารางการแจกแจงแบบเอฟ อาจจัดให้อยู่ในขั้นตอนที่ 2 แต่ผู้วิจัยมองว่าเป็นการเตรียมข้อมูล ที่ต้องใช้สำหรับการคำนวณช่วงความเชื่อมั่นตามสูตรที่แท้จริงซึ่งสัมพันธ์กับการแจกแจงเอฟ ในหัวข้อ 3.2 และจะเห็นว่าหัวข้อ 3.2 เป็นขั้นตอนที่ไม่มีการใช้ตัวเลขสุ่ม แต่เนื่องจากผลลัพธ์ของหัวข้อนี้ จะต้องถูกนำไปใช้ในหัวข้อ 3.5 จึงได้รวมเอาหัวข้อ 3.2 ไว้ในขั้นตอนที่ 2 ของเทคนิคมอนติคาร์โล

อนึ่งในการใช้เทคนิคมอนติคาร์โลครั้งนี้ ได้ทำเป็นโปรแกรมคอมพิวเตอร์ที่เขียนขึ้นด้วยภาษา FORTRAN 77 โดยใช้กับเครื่องมินิคอมพิวเตอร์ Micro VAX3600 รายละเอียดของโปรแกรมทุกโปรแกรมได้แสดงไว้ในภาคผนวก ค.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย