บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยในครั้งนี้เป็นลักษณะการวิจัยแบบเชิงทดลองซึ่งจำลองขึ้นด้วยการทำงานของเครื่องคอมพิวเตอร์ โดยใช้เทคนิคมอนติคาร์โล ซิมูเลชั่น (Monte Carlo Simulation) เพื่อหาผลสรุปในการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ (ค่าเฉลี่ยประชากร ; μ) ด้วยวิธีแฮนเซ็น-เฮอร์วิทซ์ และวิธีเอล-บาดรี้ เมื่อมีการไม่ตอบแบบสอบถามทางไปรษณีย์ มีวิธีดำเนินการวิจัยเป็นไปตามขั้นตอนตามลำดับดังนี้คือ

3.1 <u>การวางแผนการทดลอง</u>

3.1.1 สร้างประชากรให้มีการแจกแจงแบบปกติ การแจกแจงที่ใกล้เคียงปกติ เช่น การแจกแจงแบบโลจิสติค และการแจกแจงแบบไม่ปกติ เช่น การแจกแจงแบบเอกซ์โปเนนเชียล

3.1.2 ในแต่ละการแจกแจงของประชากรนั้น สร้างลักษณะของประชากรออกเป็น 2 ลักษณะคือ ลักษณะของประชากรที่เป็น นอนเซ้นซิทิฟ พอพพิวเลชั่น และลักษณะของประชากรที่เป็น เซ้นซิทิฟ พอพพิวเลชั่น กล่าวคือ

3.1.2.1 สำหรับ นอนเซ้นซิทิฟ พอพพิวเลชั่น สร้างลักษณะประชากรโดยให้เครื่องคอมพิวเตอร์ สุ่มตัวเลขออกมาตามลักษณะการแจกแจงแบบต่าง ๆ ดังกล่าวให้มีขนาดเท่ากับ 2,000 (N = 2,000)

3.1.2.2 สำหรับ เซ้นซิทิฟ พอพพิวเลชั่น สร้างลักษณะประชากรโดยให้เครื่องคอมพิวเตอร์ สุ่มตัวเลขออกมาตามลักษณะการแจกแจงแบบต่าง ๆ ให้มีขนาดเท่ากับ 5,000 แล้วจัดเรียงตัวเลขที่สุ่มได้ (ที่สร้างขึ้นมา) โดยเรียงจากน้อยไปหามาก และแบ่งตัวเลขที่เรียงแล้วออกเป็น 2 กลุ่ม คือ กลุ่มที่มีค่าของข้อมูลน้อยกับกลุ่มที่มีค่าของข้อมูลมาก โดยกำหนดสัดส่วนของกลุ่มข้อมูลที่เรียงแล้วดังต่อไปนี้

ให้กลุ่มที่ 1 เป็นกลุ่มที่มีค่าของข้อมูลน้อย เช่น กลุ่มที่มีรายได้ต่ำโดยกำหนดสัดส่วนเป็น 40% 50% และ 60% ของขนาดประชากรขนาด 5,000 หน่วย ตามลำดับ

ให้กลุ่มที่ 2 เป็นกลุ่มที่มีค่าของข้อมูลมาก เช่น กลุ่มที่มีรายได้สูง โดยกำหนดสัดส่วนเป็น 60% 50% และ 40% ของขนาดประชากรขนาด 5,000 หน่วย ตามลำดับ

3.1.3 ในแต่ละลักษณะของประชากร คำนวณค่าประมาณของค่าเฉลี่ยของประชากร ($\bar{x}_i$) และค่าความแปรปรวนของค่าเฉลี่ยที่ได้จากจำนวนชุดตัวอย่างหลาย ๆ ชุด ($V(\bar{x}_{ij})$) เมื่อ $i = 1$ หมายถึงวิธีแฮนเซ็น-เฮอร์วิทซ์ $i = 2$ หมายถึง วิธีเอล-บาดรี้ $j = 1,2,\ldots,$ IREP = จำนวนชุดตัวอย่าง กล่าวคือ

3.1.3.1 สำหรับลักษณะประชากรที่เป็นนอนเซ็นซิทิฟ พอพิวเลชั่น คำนวณ $\bar{x}_i$, $V(\bar{x}_{ij})$; $i = 1,2$; $j = 1,2 \ldots,$ IREP ทั้ง 2 วิธี คือ วิธีแฮนเซ็น-เฮอร์วิทซ์และวิธีเอล-บาดรี้ ที่ส่งแบบสอบถามทางไปรษณีย์ 3 ครั้ง (L = 3)

3.1.3.2 สำหรับลักษณะประชากรที่เป็น เซ็นซิทิฟ พอพิวเลชั่น คำนวณ $\bar{x}_i$, $V(\bar{x}_{ij})$; $i = 1,2$; $j = 3$ ชุด โดยที่

ให้กลุ่มที่ 1 คำนวณค่าประมาณต่าง ๆ ดังนี้คือ

- โดยวิธีเฮนเซ็น-เฮอร์วิทซ์ คำนวณ $\bar{x}_1$, $V(\bar{x}_{1j})$
- โดยวิธีเอล-บาดรี้ คำนวณ $\bar{x}_2$, $V(\bar{x}_{2j})$

  เมื่อทำการส่งแบบสอบถามทางไปรษณีย์ 3 ครั้ง

ให้กลุ่มที่ 2 คำนวณค่าประมาณต่าง ๆ ดังนี้คือ

- โดยวิธีแฮนเซ็น-เฮอร์วิทซ์ คำนวณ $\bar{x}_1$, $V(\bar{x}_{1j})$
- โดยวิธีเอล-บาดรี้ คำนวณ $\bar{x}_2$, $V(\bar{x}_{2j})$

  เมื่อทำการส่งแบบสอบถามทางไปรษณีย์ 3 ครั้ง

3.1.4 นำค่า $V(\bar{x}_{ij})$ เมื่อ $i = 1$ หมายถึง วิธีแฮนเซ็น-เฮอร์วิทซ์ , $i=2$ หมายถึงวิธีเอล-บาดรี้ และ $j = 1,2, \ldots,$ IREP = จำนวนชุดตัวอย่าง เปรียบเทียบกัน

ในลักษณะของประชากรทั้ง 2 ลักษณะ

3.2 วิธีทดลอง

เขียนโปรแกรมคอมพิวเตอร์ภาษาฟอร์แทรน 4 (FORTRAN IV) ซึ่งใช้กับเครื่องคอมพิวเตอร์ IBM 370/3031 เพื่อสร้างข้อมูลให้เป็นไปตามแผนการทดลอง และคำนวณค่าประมาณต่าง ๆ ของทั้ง 2 วิธีคือ แอนเซิน-เฮอร์วิทซ์ และวิธีเอล-บาดรี้ ด้วยเทคนิคมอนติคาร์โล ตามลำดับดังนี้คือ

3.2.1 สร้างข้อมูลเพื่อให้เป็นไปตามการแจกแจงของประชากร

3.2.1.1 สร้างตัวเลขสุ่มให้มีการแจกแจงแบบสม่ำเสมอ

การสร้างตัวเลขสุ่มโดยใช้วิธี Multiplication Congruential กล่าวคือ

กำหนดให้ $R_i = R_{i-1}\ (2^P+K) \bmod (2^{31})$ ....(1)

เมื่อ P เป็นค่าตัวเลขตั้งแต่ 3 ถึง 30

K เป็นจำนวนเต็มที่

R_o เป็นค่าเริ่มต้นซึ่งเป็นจำนวนเต็มคี่

ตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ ซึ่งมีค่าระหว่าง 0 และ 1 หาได้จาก

$$U_i = \frac{R_i}{(2^{31}-1)} \quad ....(2)$$

โปรแกรมย่อยที่ใช้ในการสร้างตัวเลขสุ่มนี้คือ

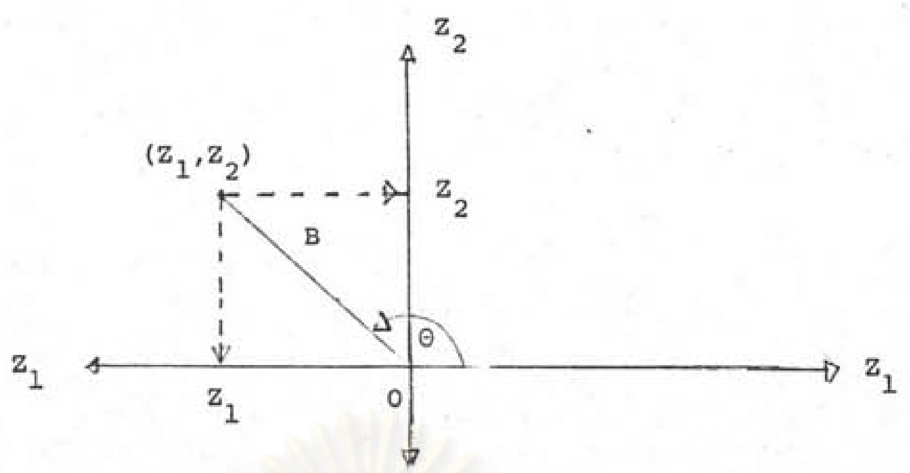
```
      SUBROUTINE RANDU(IX , IY, YFL)
      IY  =  IX * 65539
      IF (IY) 5, 6, 6
    5 IY  =  IY + 2147483647 + 1
    6 YFL = IY
      YFL =  YEL* 0.4656613E-09
      IX  = IY
      RETURN
      END
```

3.2.1.2 แปลงตัวเลขสุ่ม ให้มีการแจกแจงแบบปกติ แบบใกล้เคียงปกติ และการแจกแจงแบบไม่ปกติ โดยใช้วิธีดังต่อไปนี้

3.2.1.2.1 การสร้างตัวเลขสุ่มให้มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน 1

โปรแกรมย่อยที่ใช้ในการสร้างการแจกแจงแบบปกติในการวิจัยครั้งนี้ จะใช้เทคนิค ไดเรค ทรานส์ฟอร์เมชั่น (Direct Transformation) ของ Box-Muller (1958 : 610-611) ซึ่งจะใช้จำนวนคู่ของตัวแปรที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน 1 ซึ่งแต่ละคู่จะเป็นอิสระต่อกัน

พิจารณา ตัวแปร Standard Normal Random Variables 2 ตัวคือ Z_1 และ Z_2 ซึ่งเขียนเป็นรูป Polar Coordinates ได้ดังนี้



$$\text{เมื่อ} \quad \left.\begin{aligned} Z_1 &= B\cos\Theta \\ Z_2 &= B\sin\Theta \end{aligned}\right\} \qquad \cdots\cdots\cdots \; (3.2.1.2.1.1)$$

และ $B^2 = Z_1^2 + Z_2^2$ มีการแจกแจงแบบไค-สแควร์ ที่มีจำนวนองศาความเป็นอิสระเท่ากับ 2 หรือมีการแจกแจงแบบ เอ็กซ์โปเนนเซียลที่มีค่าเฉลี่ยเท่ากับ 2 ดังนั้นรัศมี B หาได้ดังนี้

$$\text{จาก} \quad \mathrm{Gam}\left(\frac{k}{2}, \frac{1}{2}\right) \sim \lambda^2(k)$$

$$\therefore \quad \mathrm{Gam}\left(1, \frac{1}{2}\right) \sim \lambda^2(2)$$

$$\text{แต่} \quad \mathrm{Gam}(1,\lambda) \sim \exp(\lambda) \sim \exp\left(\frac{1}{2}\right)$$

จาก การแจกแจงแบบเอ็กซ์โปเนนเซียลมีฟังก์ชันความหนาแน่นคือ

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \quad \text{เมื่อ} \quad x = B^2 \\ 0 & , x < 0 \end{cases}$$

$$\therefore F(x) = \int_{\infty}^{x} f(t)\,dt = \begin{cases} 1 - e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

$$\text{และ} \quad E(x) = \frac{1}{\lambda} = 2$$

ให้ $F(x) = R_1$ บนช่วงของ x

$$\therefore \quad 1 - e^{-\lambda x} = R_1$$

$$e^{-\frac{1}{2}x} = 1 - R_1$$

$$\therefore \quad x = -2\ln(1-R_1) \qquad \text{.......... (3.2.1.2.1.2)}$$

เรียกสมการ(3.2.1.2.1.2) ว่าเป็น A Random Variate Generator for the Exponential Distribution หรือ

$$x = F^{-1}(R_1)$$

แทนค่า $1 - R_1$ ด้วย R_1 จะได้เป็น

$$x = B^2 = -2\ln R_1 \qquad \text{...........(3.2.1.2.1.3)}$$

เพราะว่า $1 - R_1$ และ R_1 มีการแจกแจงแบบ U (0,1) เหมือนกันและจากสมการ (3.2.1.2.1.3)

$$B = (-1\ln R_1)^{1/2} \qquad \text{...........(3.2.1.2.1.4)}$$

สำหรับมุม Θ จะมีค่าอยู่ระหว่าง 0 ถึง 2Π

ให้ $\Theta \sim U(0,2\Pi)$ และ B และ Θ เป็น Mutually Independent

จากสมการ (3.2.1.2.1.1)และ (3.2.1.2.1.4)

$$\left.\begin{aligned} Z_1 &= (-2\ln R_1)^{1/2}\cos(2\Pi R_2) \\ Z_2 &= (-2\ln R_1)^{1/2}\sin(2\Pi R_2) \end{aligned}\right\} \qquad \text{......(3.2.1.2.1.5)}$$

โดยที่ Z_1 และ Z_2 จะเป็นการ Generating two Independent Standard Normal

Variates Z_1 และ Z_2 จาก Two Independent Random Number R_1 และ R_2

โปรแกรมย่อยที่ใช้คือ

```
      SUBROUTINE NORM(ISEED, SIGMA, MEAN, ZZ)
      REAL MEAN
      DATA PI/3.14159/, L1/0/
      IF (L1.EQ. 1) GOTO 10
      CALL RANDU(ISEED, IYY , YYFL)
      RONE  =  YYFL
      CALL RANDU(ISEED, IYY, YYFL)
      RTWO  =  YYFL
      ZONE  =  SQRT (-2* ALOG(RONE))*COS(2*PI*RTWO)
      ZTWO  =  SQRT (-2* ALOG(RONE))*SIN(2*PI*RTWO)
      ZZ    =  ZONE*SIGMA + MEAN
      L1    =  1
      RETURN
   10 ZZ    =  ZTWO*SIGMA+MEAN
      L1    =  0
      RETURN
      END
```

ผลจากการเขียน SUBROUTINE NORM (ISEED, SIGMA, MEAN, ZZ) เมื่อ SIGMA และ MEAN เป็นพารามิเตอร์ 2 ตัว ซึ่งกำหนดค่าส่วนเบี่ยงเบนมาตรฐานและค่าเฉลี่ยประชากร เป็นค่าที่ส่งมาจากโปรแกรมหลัก (Main Program) ส่วนผลลัพธ์คือค่า ZZ ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ MEAN และความแปรปรวนเท่ากับ $(SIGMA)^2$

3.2.1.2.2 การสร้างตัวเลขสุ่มให้มีการแจกแจงแบบใกล้เคียงปกติ เช่น การแจกแจงแบบโลจิสติค

การแจกแจงแบบโลจิสติคมีฟังก์ชันความหนาแน่น (p.d.f) ดังนี้

$$f(x) = \frac{1}{\beta} \frac{e^{-\frac{(x-\alpha)}{\beta}}}{\left[1 + e^{-\frac{(x-\alpha)}{\beta}}\right]^2} \quad , -\infty < x < \infty$$

การสร้างการแจกแจงแบบโลจิสติค จะใช้วิธี Inserse Transformation ซึ่งแสดงได้ดังนี้

$$F(x) = \int_{-\infty}^{x} f(t)dt \quad , -\infty < x < \infty$$

$$= \int_{-\infty}^{x} \frac{e^{-\frac{(t-\alpha)}{\beta}}}{\beta\left[1+e^{-\frac{(t-\alpha)}{\beta}}\right]^2} . dt$$

$$= \int_{-\infty}^{x} \frac{e^{-\frac{(t-\alpha)}{\beta}}}{\left[1+ e^{-\frac{(t-\alpha)}{\beta}}\right]^2} \, d\left(\frac{t-\alpha}{\beta}\right)$$

$$= -\int_{-\infty}^{x} \frac{d \, . \, e^{-\frac{(t-\alpha)}{\beta}}}{\left[1 + e^{-\frac{(t-\alpha)}{\beta}}\right]^2}$$

$$= -\int_{-\infty}^{x} \frac{d\left[1 + e^{-\frac{(t-\alpha)}{\beta}}\right]}{\left[1+ e^{-\frac{(t-\alpha)}{\beta}}\right]^2}$$

$$= \frac{1}{\left[1 + e^{-\frac{(x-\alpha)}{\beta}}\right]}$$

$$\therefore \quad F(x) \cdot \left[1 + e^{-\frac{(x-\alpha)}{\beta}}\right] = 1$$

$$e^{-\frac{(x-\alpha)}{\beta}} = \left[\frac{1}{F(x)} - 1\right] = \frac{1 - F(x)}{F(x)}$$

$$\therefore \quad x = \alpha + \beta\left[\ell n\,(F(x)) - \ell n(1-F(x))\right]$$

หรือ $x = \alpha + \beta\left[\ell n(YFL) - \ell n(1-YFL)\right]$ เมื่อ YFL มีการแจกแจงแบบสม่ำเสมอ U (0 , 1)

เขียนเป็นโปรแกรมย่อยได้ดังนี้

```
SUBROUTINE LOGIS(ALPHA, BETA,ZZ,SIGMA,MEAN)
REAL MEAN
CALL RANDU (ISEED, IY, YFL)
W     =  ALOG(YFL)-ALOG(1-YFL)
ALPHA = MEAN
BETA  = 1.7320*SIGMA/3.14159
ZZ    =  ALPHA + W*BETA
RETURN
END
```

เมื่อ ALPHA(α) และ BETA(β) เป็นพารามิเตอร์ 2 ตัว ซึ่งกำหนดค่าเฉลี่ยเท่ากับ α และความแปรปรวน เท่ากับ $\frac{1}{3}(\Pi)^2(\beta)^2$ โดยที่ค่าเฉลี่ยและค่าความแปรปรวน($(\mathrm{SIGMA})^2$) ส่งมาจากโปรแกรมหลักและผลลัพธ์คือค่า ZZ $= \alpha + \beta[\ell n(\mathrm{YFL}) - \ell n(1-\mathrm{YFL})]$ ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบโลจิสติคที่มีค่าเฉลี่ย $= \alpha$ และความแปรปรวน $= \frac{1}{3}\ \Pi^2\ \beta^2$

3.2.1.2.3 <u>การสร้างตัวเลขสุ่มให้มีการแจกแจงแบบไม่ปกติ เช่น การแจกแจงแบบเอ็กซ์โปเนนเชียล</u>

การแจกแจงแบบเอ็กซ์โปเนนเชียลมีฟังก์ชันความหนาแน่น (p.d.f) ดังนี้

$$f(x) = \frac{1}{\beta}\ e^{-x/\beta} \quad , x \geq 0$$

การสร้างการแจกแจงแบบเอ็กซ์โปเนนเชียลจะใช้วิธี Inverse-Transformation ซึ่งแสดงได้ดังนี้

$$F(x) = \int_0^x f(t)dt \quad , \quad x \geq 0$$

$$= \int_0^x \frac{1}{\beta}\ e^{-t/\beta}\ dt$$

$$= -\int_0^x e^{-t/\beta}\ d(-t/\beta)$$

$$= -\left[e^{-t/\beta}\right]_0^x$$

$$= 1-e^{-x/\beta} = R$$

Inverse Transformation ในที่นี้คือ

$$1 - R = e^{-x/\beta}$$

โดยที่ (1-R) ยังเป็นเลขสุ่มที่มีการกระจายแบบสม่ำเสมอ ซึ่งมีขอบเขตจำกัดคือ (0 , 1) ดังนั้นจึงอาจแทน 1 - R ด้วย R และสมการข้างต้นจะเปลี่ยนเป็น

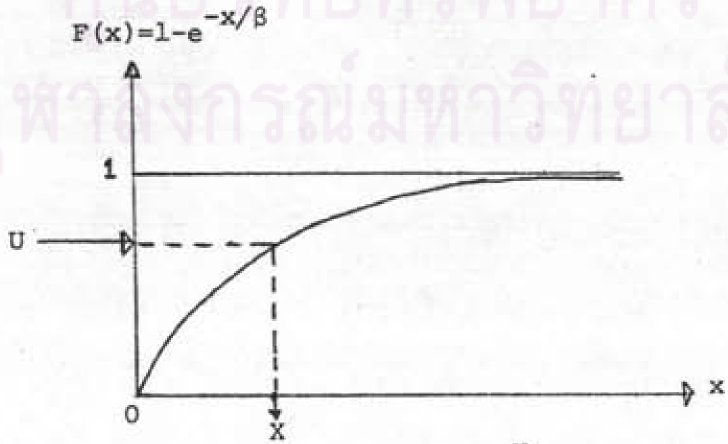
$$R = e^{-x/\beta}$$

$$x = -\beta \ell n R \qquad \text{.............(1)}$$

ดังนั้น $F^{-1}(R) = -\beta \ell n R = x$

ดังนั้น วิธีการหาเลขสุ่มที่มีการกระจายแบบเอ็กซ์โปเนนเชียล จึงต้องหาเลขสุ่มที่มีการกระจายแบบสม่ำเสมอ (R) ขึ้นก่อน แล้วหาเลขสุ่มที่มีการกระจายแบบเอ็กซ์โปเนนเชียล โดยใช้สมการ (1)

และแสดงวิธี Invense-Transformation ซึ่งจะสร้างตัวเลขสุ่มให้มีการแจกแจงแบบเอ็กซ์โปเนนเชียล โดยใช้รูปภาพดังนี้



$$P(X \leqslant x) = F(x) = \int_0^x f(t)dt = 1 - e^{-x/\beta}$$

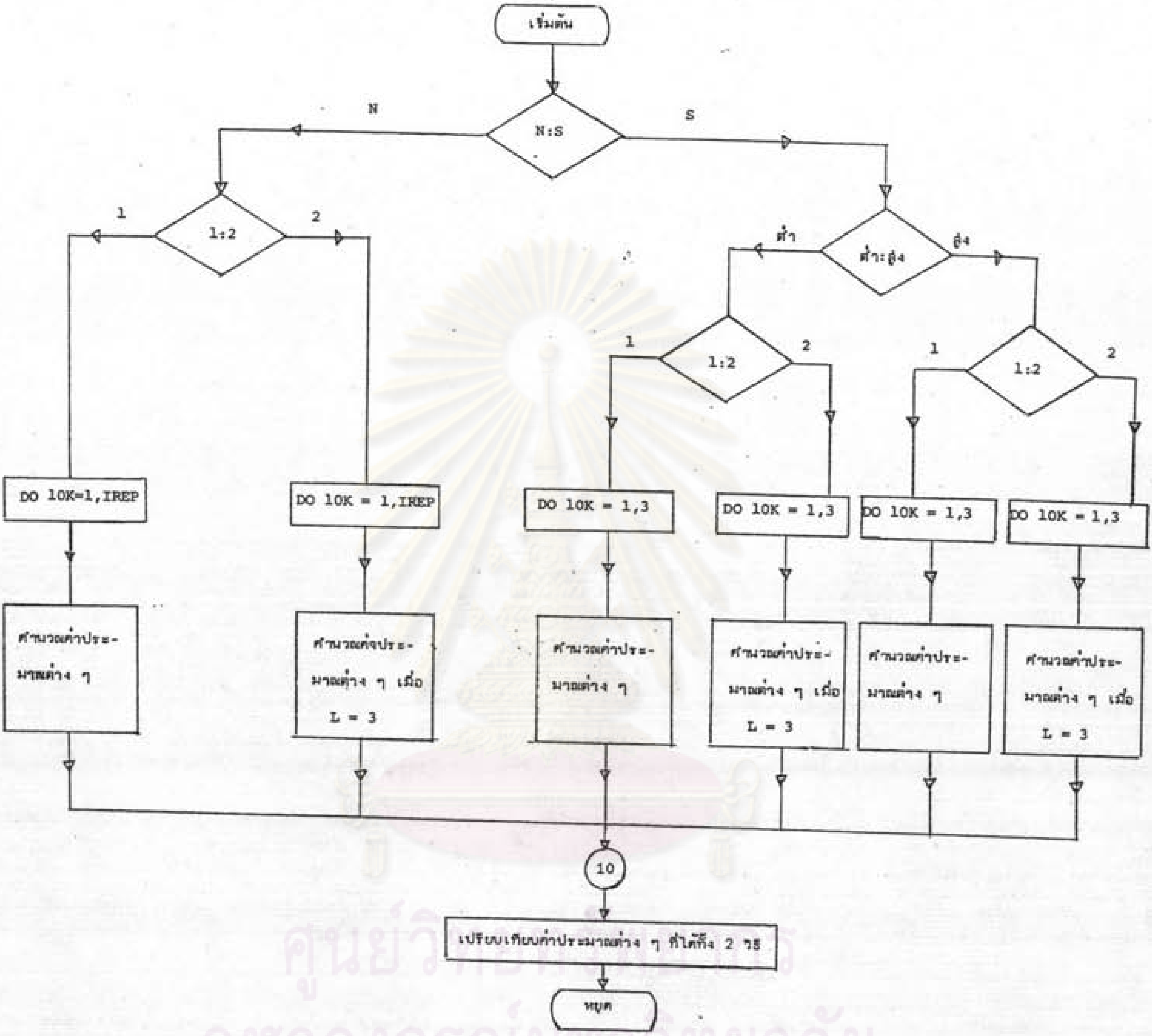
โปรแกรมย่อย ที่ใช้คือ

```
SUBROUTINE EXPO(ISEED, MEAN, XX)
REAL MEAN
CALL RANDU(ISEED, IY, YFL)
XX  =  MEAN*ALOG(YEL)
RETURN
END
```

ผลจากการเขียน SUBROUTINE(ISEED, MEAND, XX) เมื่อ MEAN = β โดยที่ MEAN เป็นค่าที่ส่งมาจากโปรแกรมหลัก ซึ่งเป็นค่าเฉลี่ยของประชากร ผลลัพธ์คือค่า XX ซึ่งเป็นตัวแปรสุ่มที่มีการแจกแจงแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ β และความแปรปรวนเท่ากับ β^2

3.2.2 ในแต่ละการแจกแจงของประชากร สร้างประชากรออกเป็น 2 ลักษณะคือ นอนเซ็นซิทิฟ พอพิวเลชั่น และเซ็นซิทิฟ พอพิวเลชั่น และคำนวณค่าประมาณต่าง ๆ ทั้ง 2 วิธี ซึ่งเขียนเป็นแผนผังได้ดังนี้

แผนผังแสดงขั้นตอนการสร้างลักษณะของประชากรและการคำนวณค่าประมาณทั้ง 2 วิธี



แผนผังที่ 3

<u>หมายเหตุ</u>			
	N	หมายถึง	นอนเซ็นซีทีฟ พอพพิวเลชั่น
	S	"	เซ็นซีทีฟ พอพพิวเลชั่น
	1	"	วิธีของ แฮนเซ็น-เฮอร์วิทซ์
	2	"	วิธีของ เอล-บาดรี้
	ต่ำ	"	กลุ่มของเซ็นซีทีฟ พอพพิวเลชั่น ที่มีค่าของข้อมูลต่ำ (เช่นพวกที่มีรายได้ต่ำ)
	สูง	"	กลุ่มของเซ็นซีทีฟ พอพพิวเลชั่นที่มีค่าของข้อมูลสูง (เช่นพวกที่มีรายได้สูง)
	IREP	"	จำนวนซ้ำ หรือจำนวนชุดตัวอย่างซึ่งจำแนกตาม C.V ของประชากรที่สร้างขึ้น

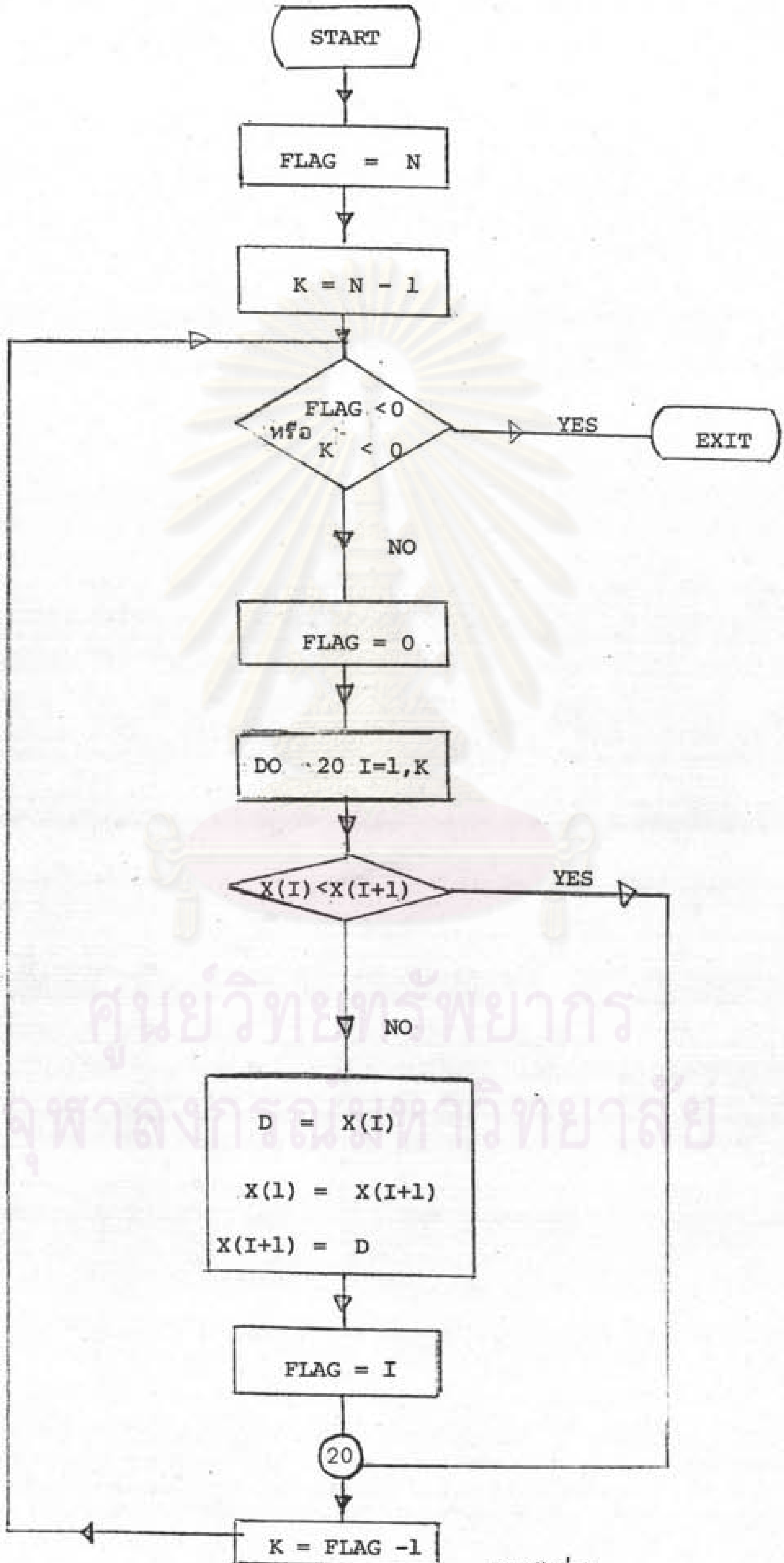
จากแผนผังที่ 3 อธิบายโดยอาศัยโปรแกรมคอมพิวเตอร์ดังนี้

3.2.2.1 การสร้างลักษณะประชากรให้มีลักษณะเป็น นอนเซ้นซีทิพ พอพพิวเลชั่นนั่นคือ การสร้างข้อมูลออกมาตามลักษณะการแจกแจงแบบต่าง ๆ โดยอาศัย SUBROUTINE NORM, SUBROUTINE LOGIS และ SUBROUTINE EXPO ซึ่งเมื่อสร้างข้อมูลออกมาโดยสุ่มตามลักษณะการแจกแจงแบบต่าง ๆ แล้วจะไม่เรียง (SORT) ข้อมูล

3.2.2.2 การสร้างลักษณะประชากรให้มีลักษณะเป็นเซ้นซีทิพ พอพพิวเลชั่นนั่นคือ การสร้างข้อมูลออกมาตามลักษณะการแจกแจงแบบต่าง ๆ เสียก่อน แล้วจึงจัดเรียงข้อมูลที่ได้จากน้อยไปหามาก โดยใช้โปรแกรมการ SORT ซึ่งมีหลักการดังนี้

การเรียง (SORTING) ข้อมูลใช้วิธีบับเบิ้ลซอร์ท (Bubble Sort) โดยเรียงจากค่าน้อยไปหาค่ามาก ซึ่งเขียนแผนผัง แสดงตรรกวิธีของวิธีการจัดเรียงข้อมูลที่เรียกว่า "บับเบิ้ลซอร์ท" ดังนี้

แผนผังแสดงขั้นตอนการเรียงข้อมูลโดยวิธีบับเบิ้ลซอร์ท



แผนผังที่ 4

<u>หมายเหตุ</u>

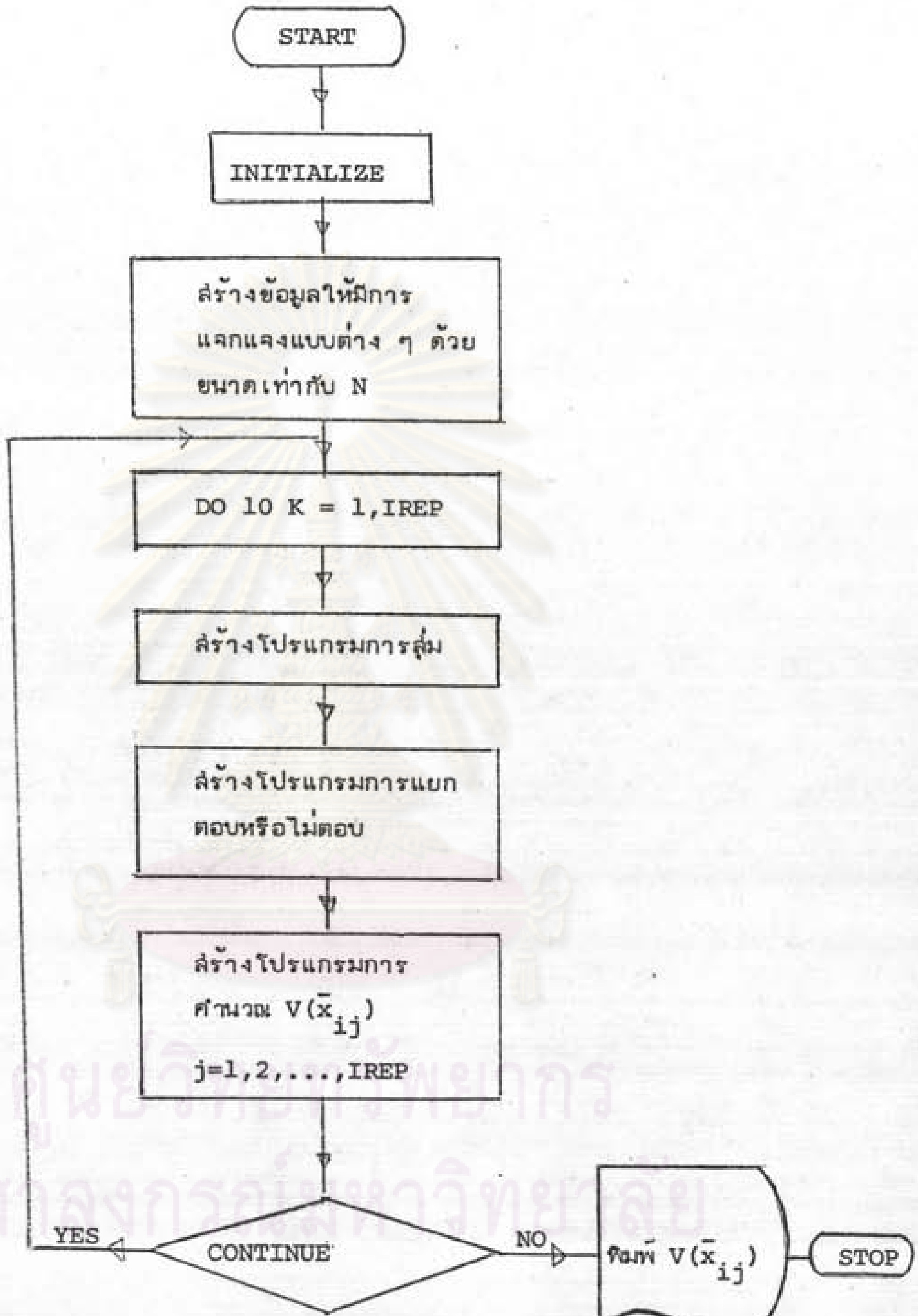
(1) I คือ จำนวนข้อมูลทั้งหมดที่จะนำมาจัดเรียง

(2) X คือ ค่าของข้อมูล

หลังจากที่เรียงข้อมูลแล้ว จึงแยกข้อมูลที่เรียงแล้วออกเป็น 2 กลุ่มคือ กลุ่มของข้อมูลที่มีค่าน้อยกับกลุ่มของข้อมูลที่มีค่ามาก โดยกำหนดตามสัดส่วนที่กำหนดให้คือ 40% 50% และ 60% ของขนาดประชากร 5,000 หน่วย

3.2.2.3 สำหรับการคำนวณค่าประมาณต่าง ๆ ของแต่ละวิธีนั้น เขียนเป็นแผนผังการเขียนโปรแกรมได้ดังนี้

แผนผังการคำนวณค่าประมาณต่าง ๆ ในแต่ละวิธี



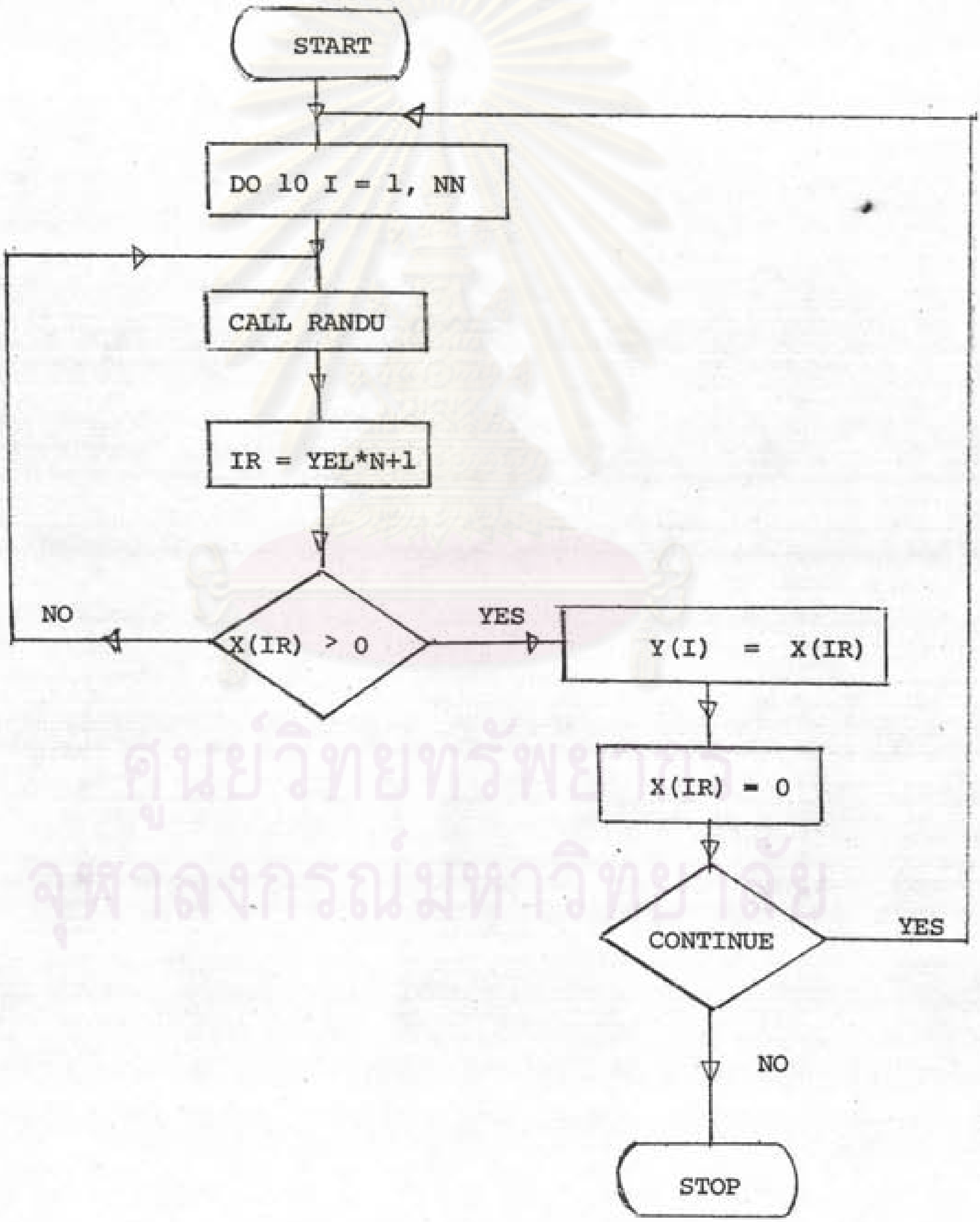
แผนผังที่ 5

จากแผนผังที่ 5 อธิบายรายละเอียดของโปรแกรมต่าง ๆ ดังนี้

3.2.2.3.1 โปรแกรมการสุ่ม (Sampling Program)

อาศัยหลักของการสุ่มตัวอย่างอย่างง่ายแบบไม่ใส่คืน ซึ่งเขียนเป็นแผนผังโปรแกรมดังนี้คือ

แผนผังแสดงขั้นตอนการเขียนโปรแกรมการสุ่ม



แผนผังที่ 6

<u>หมายเหตุ</u>

ให้ NN คือ ขนาดตัวอย่างที่ต้องการ

N คือ ขนาดของประชากรที่ต้องการศึกษา

YEL คือ ค่าของตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ ซึ่งมีค่าอยู่ระหว่าง 0 และ 1 โดยมาจาก SUBROUTINE RANDU

IR คือ ค่าของตัวเลขที่ขึ้นอยู่กับค่า YEL ซึ่งมีค่าเท่ากับ 1 ถึง N เขียนเป็นสูตรได้คือ IR = YFL*N+1

X คือ ค่าของตัวเลขของประชากรที่สร้างขึ้นมาให้มีการแจกแจงแบบต่าง ๆ ดังกล่าวข้างต้น

Y คือ ค่าของตัวเลขของตัวอย่างที่สุ่มได้

จากแผนผังดังกล่าว กรณีที่กำหนดให้ X(IR) = 0 นั้นเพื่อมิให้ตัวเลขที่สุ่มไปแล้ว กลับถูกสุ่มอีกครั้งหนึ่ง นั่นคือต้องการให้สุ่มแบบไม่ใส่คืนนั่นเอง และเขียนโปรแกรมดังนี้

```
   DO 10 I = 1, NN
1  CALL RANDU(ISEED, IY, YFL)
   IR = YFL*N+1
   IF (X(IR). GT.0) GOTO 11
   GOTO 1
11 Y(I) = X(IR)
   X(IR) =  0
10 CONTINUE
```

3.2.2.3.2 <u>โปรแกรมการแยก(ตอบหรือไม่ตอบ)</u>

อาศัยหลักการสุ่มตัวอย่างแบบมีระบบ (Systematic Sampling) ในการแยกตัวอย่างออกเป็น 2 กลุ่ม คือ กลุ่มที่ตอบ (Response Group) กับกลุ่มที่ไม่ตอบ (Nonresponse Group) ซึ่งคุณวิไลพร ธรรมเนียมอินทร์ (2524 หน้า 53) ได้กล่าวถึงวิธีการสุ่มแบบมีระบบ ดังนี้

ถ้าประชากรประกอบด้วยสมาชิก N หน่วยเรียงกันจาก 1 ถึง N แบบสุ่มสมมติว่าต้องการสุ่มตัวอย่างมาเพียง 5 เปอร์เซ็นต์ของจำนวนประชากรทั้งหมดก็เลือกตัวอย่างจากสมาชิกของประชากรที่เรียงกันนั้น จำนวน 1 คน ต่อประชากรทุก ๆ 20 คน โดยเริ่มต้นด้วยการสุ่มหาตัวเลขเริ่มต้น (Random Start) เสียก่อน ซึ่งทำได้ง่าย ๆ โดยการสุ่มเลือกเลข 1 ตัว มาจากเลข 1-20 สมมติได้ 3 ตัวอย่างที่เลือกมาทั้งหมดคือประชากรลำดับที่ 3,23,43.... นั่นคือ ตัวอย่างที่ต้องการทั้งหมดมีขนาด n จะต้องหาตัวเลขช่วงสุ่ม (Sampling Intreval ; I) ให้มีค่าใกล้เคียง $\frac{N}{n}$ แล้วสุ่มหาตัวเลขเริ่มต้นมาจากเลขระหว่าง 1 - I สมมติได้ 10 ประชากรลำดับที่ 10 + I , 10 + 2I , 10 + 3I,... จะถูกบังคับให้เลือกออกมาโดยอัตโนมัติ

ดังนั้นการแยกตัวอย่างดังกล่าว ออกเป็น 2 กลุ่ม (กลุ่มตอบ , กลุ่มไม่ตอบ) โดยอาศัยหลักการสุ่มตัวอย่างแบบมีระบบดังนี้

ให้ Y คือ เลขสุ่มของตัวอย่าง ขนาด NN (ซึ่งเลือกมาจากประชากรขนาด N)

W คือ เลขสุ่มตัวเริ่มต้นที่จะถูกเลือกเป็นตัวอย่างตัวแรกสุดของการเลือกตัวอย่างแต่ละชุด และถูกใช้เป็นตัวกำหนดตำแหน่งของเลขสุ่มตัวที่สองที่จะถูกเลือกเป็นตัวอย่าง เช่น 1, 2, 3, ..., $\frac{n}{n1}$ (ในที่นี้กำหนดให้ W = 1 คือให้ตัวที่ 1 เป็นกลุ่มที่ตอบ)

โดยใช้สูตร $J = W + (I - 1) \times A1$

J คือ ตำแหน่งของเลขสุ่มที่ถูกเลือกเป็นตัวอย่าง (กลุ่มตอบ)

$$A1 = \frac{NN}{M}$$

NN คือ จำนวนตัวอย่างที่สุ่มมาจากประชากรขนาด N

M คือ ตัวอย่างที่ต้องการ หรือจำนวนของกลุ่มที่ตอบ

I คือ ตัวแสดงจำนวนของตัวอย่างมีค่าตั้งแต่ 1 ถึง M

เขียนโปรแกรมได้ดังนี้

```
    W    =  1
    Al   =  NN/M
    NM   =  NN-M
    DO      90 I = 1, M
    J    =  W + (I-1)*Al
    R(J)  = Y(J)
    Y(J)  =   0
90  CONTINUE
    Il   =  0
    DO  100 K = 1 , NM
    IF(Y(K).EQ.0) GØTØ  100
    Il = Il + 1
    GNR(Il) = Y(K)
100 CONTINUE
```

จากโปรแกรมดังกล่าว จะได้กลุ่มตัวอย่างที่ตอบ คือ R(I) ขนาด M และกลุ่มตัวอย่างที่ไม่ตอบคือ GNR(I1) ขนาด NM

3.2.2.3.3 สำหรับโปรแกรมการคำนวณค่าประมาณต่าง ๆ นั้นจะใช้ SUBROUTINE STAT (ดูภาคผนวก)

3.2.2.4 การเปรียบเทียบค่าประมาณต่าง ๆ ที่ได้ทั้ง 2 วิธีนั้น พิจารณาจากค่า $V(\bar{x}_{ij})$ (เมื่อ $i = 1$ หมายถึง วิธีแฮนเซิน-เฮอร์วิทซ์ , $i = 2$ หมายถึง วิธีแอล-บาตรี้ ; $j = 1,2,\ldots,$ IREP เป็นจำนวนชุดตัวอย่างซึ่งแต่ละชุดตัวอย่างจะใช้ตัวอย่างขนาด n_{opt} เท่ากัน)

ถ้าวิธีไหนให้ค่า $V(\bar{x}_{ij})$ น้อยกว่า วิธีประมาณค่าพารามิเตอร์โดยวิธีนั้นจะมีประสิทธิภาพที่ดีกว่า