บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบของตัวสถิติทดสอบ 4 ตัวคือ

1) ตัวสถิติทดสอบ Gini
2) ตัวสถิติทดสอบ Lorenz
3) ตัวสถิติทดสอบ Kolmogorov-Smirnov
4) ตัวสถิติทดสอบ Anderson-Darling

โดยใช้การจำลองข้อมูลด้วยวิธีมอนติคาร์โล (Monte Carlo Method) ซึ่งเขียนโปรแกรมคอมพิวเตอร์ด้วยภาษาฟอร์แทรน 77 (Fortran Programming) บนเครื่อง PC และขั้นตอนแผนการทดลอง วิธีการวิจัยจะนำเสนอเป็นลำดับดังนี้

3.1 ขั้นตอนการวิจัย

กำหนดสถานการณ์สำหรับหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 3 ตัวดังนี้

3.1.1 เลือกตัวอย่างสุ่มจากประชากร โดยกำหนดประชากรให้มีการแจกแจงดังต่อไปนี้

ก) ประชากรมีการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ ซึ่งในที่นี้กำหนดค่าพารามิเตอร์ที่สนใจศึกษาดังต่อไปนี้

1. พารามิเตอร์สเกล(θ) มีค่าเท่ากับ 1 เมื่อ

- พารามิเตอร์ตำแหน่ง (β) มีค่าเท่ากับ 0.5 ,1,2 และ 5 ตามลำดับ

2. พารามิเตอร์ตำแหน่ง (β) มีค่าเท่ากับ 1

- พารามิเตอร์สเกล (θ) มีค่าเท่ากับ 0.5,1,2 และ 5 ตามลำดับ

ข) การแจกแจงแบบลอกนอร์มอล โดยสนใจศึกษาเมื่อพารามิเตอร์

$\mu = 0$ และ $\sigma^2 = 0.7 , 0.9$ ตามลำดับ

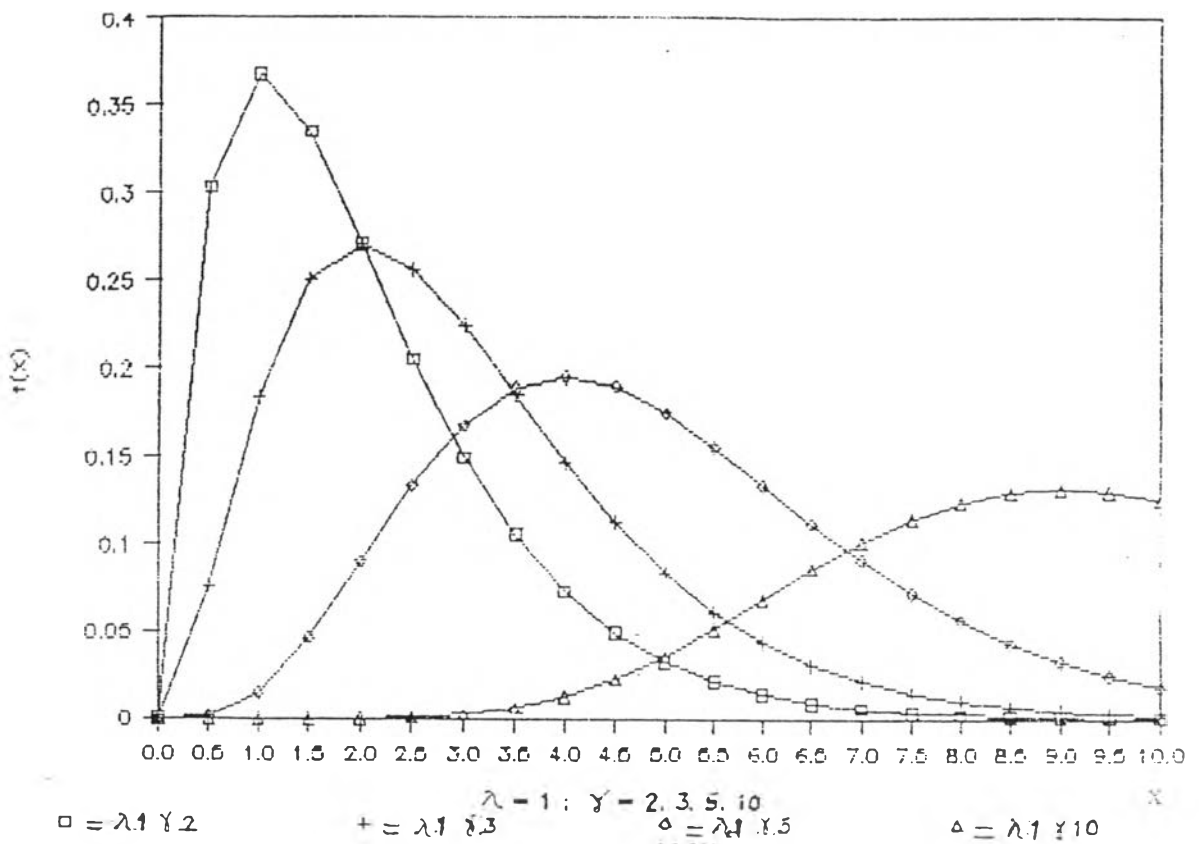
ค) การแจกแจงแบบไวบูลล์ โดยสนใจศึกษาเมื่อพารามิเตอร์

$\beta = 1$ และ $\alpha = 0.5 ,2.0$ ตามลำดับ

ง) การแจกแจงแบบแกมมา โดยสนใจศึกษาเมื่อพารามิเตอร์

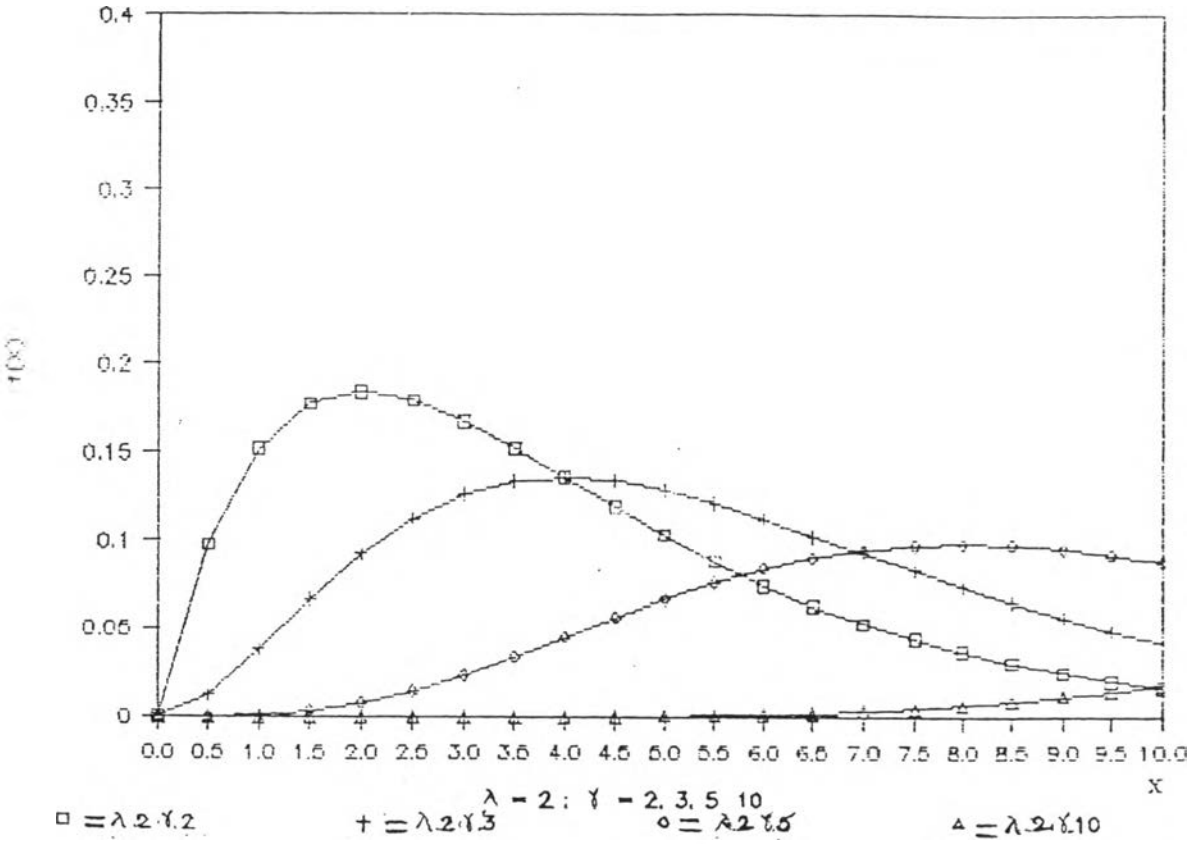
$\lambda = 1$ และ $\gamma = 1, 2, 3$ ตามลำดับ

สำหรับการแจกแจงแบบแกมมา , การแจกแจงแบบไวบูลล์ และการแจกแจงแบบลอกนอร์มอล มีเกณฑ์ที่พิจารณาค่าแปรเปลี่ยนพารามิเตอร์ กล่าวคือจะพิจารณาให้พารามิเตอร์ที่เปลี่ยนนั้นมีผลทำให้รูปกราฟมีค่าใกล้เคียงกับการแจกแจงเลขชี้กำลัง 2 พารามิเตอร์ และค่อย ๆ มีรูปแบบที่แตกต่างกันมากขึ้น ซึ่งในแต่ละการแจกแจงจะใช้สัมประสิทธิ์ความแปรผัน(Coefficient of Variable:C.V) เป็นตัวกำหนดค่าพารามิเตอร์

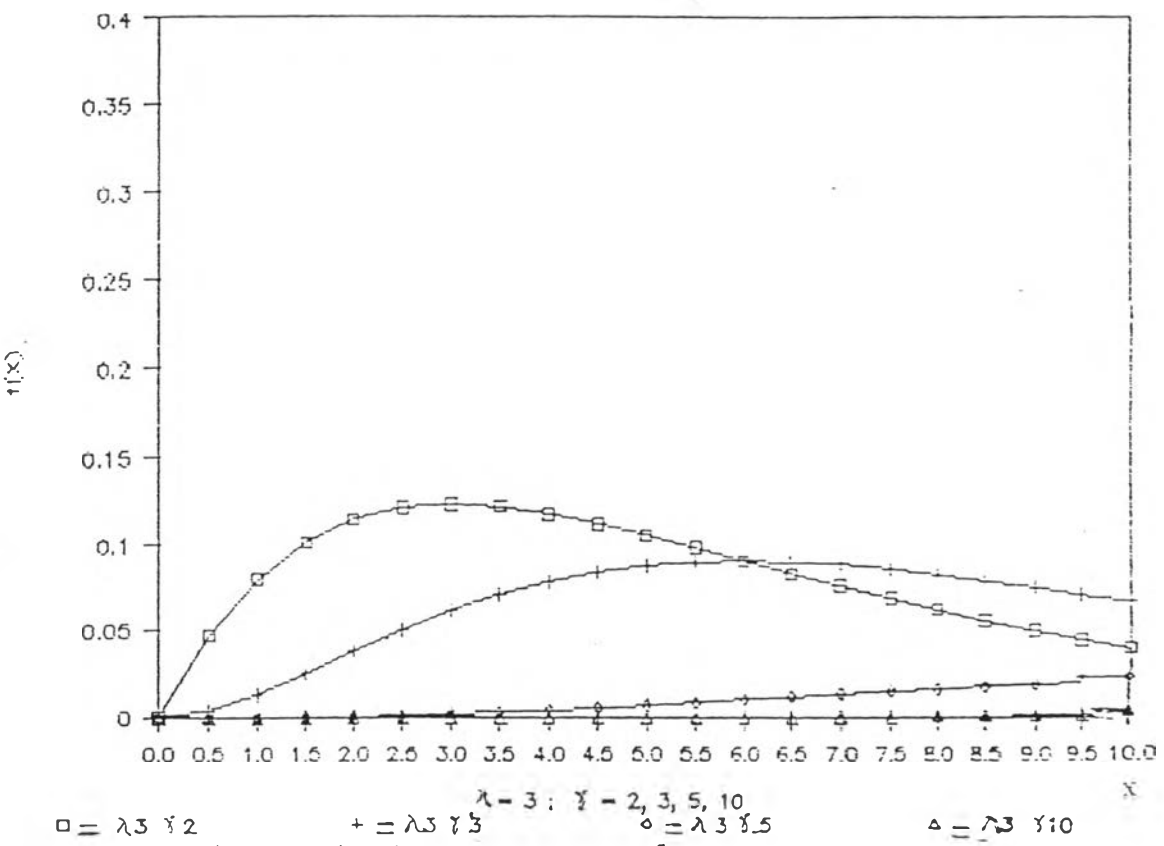


รูปที่ 3.1.1.1 รูปกราฟแสดงการแจกแจงแบบแกมมา เมื่อ $\lambda = 1$ และ $\gamma = 2, 3, 5$ และ 10

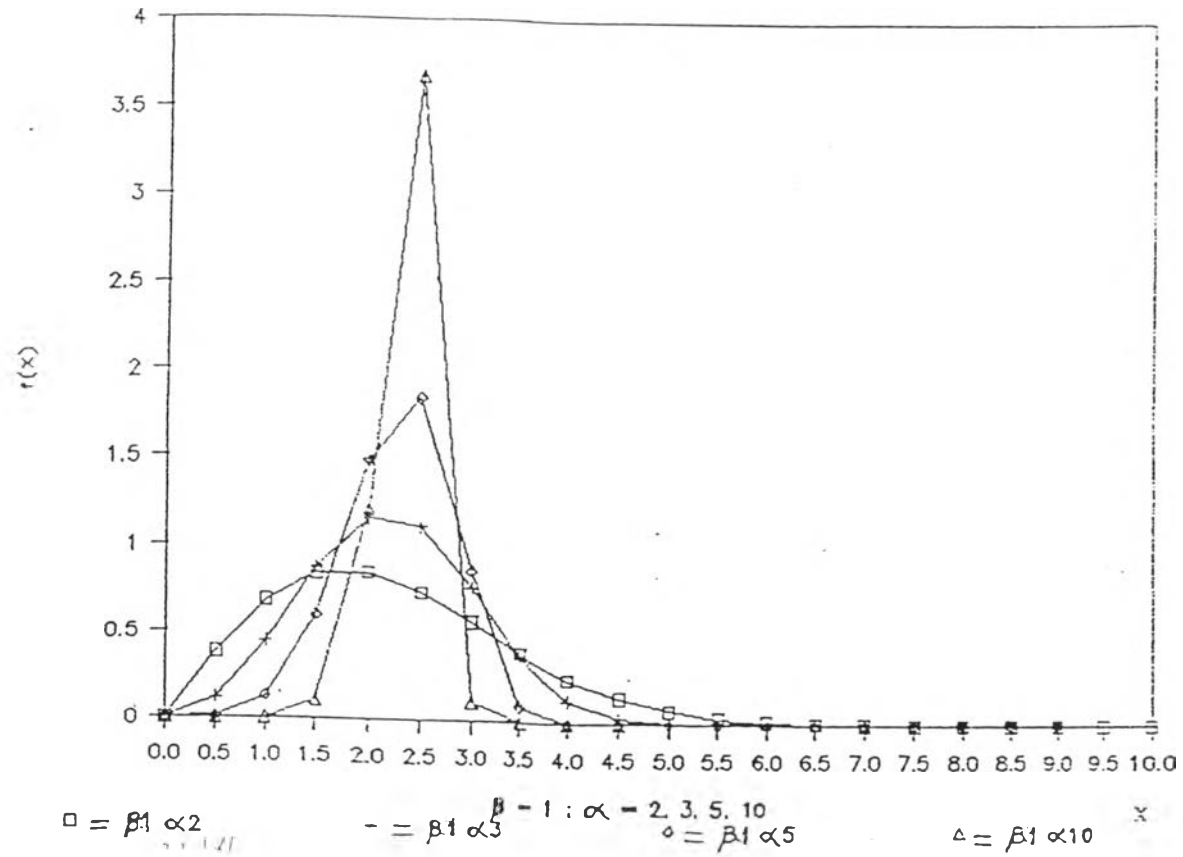
หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



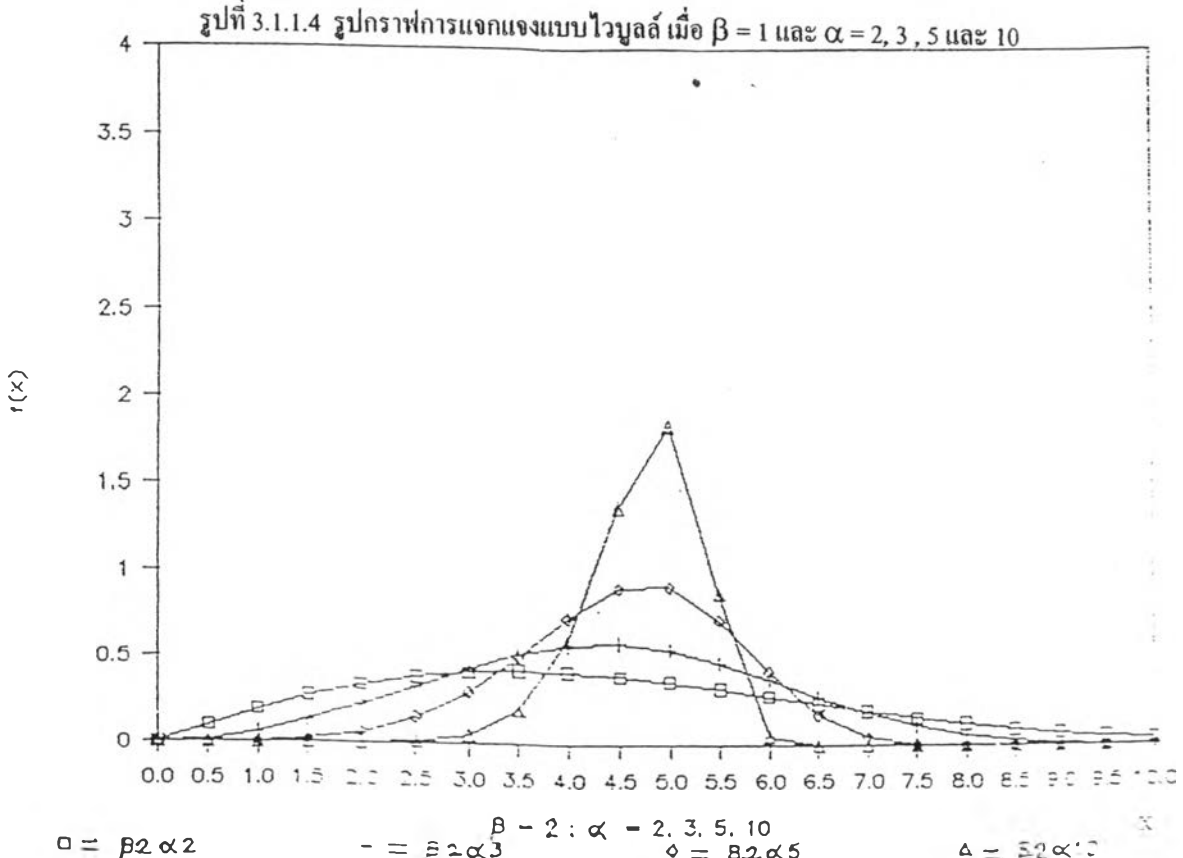
รูปที่ 3.1.1.2 รูปกราฟการแจกแจงแบบแกมมา เมื่อ $\lambda = 2$ และ $\gamma = 2, 3, 5$ และ 10



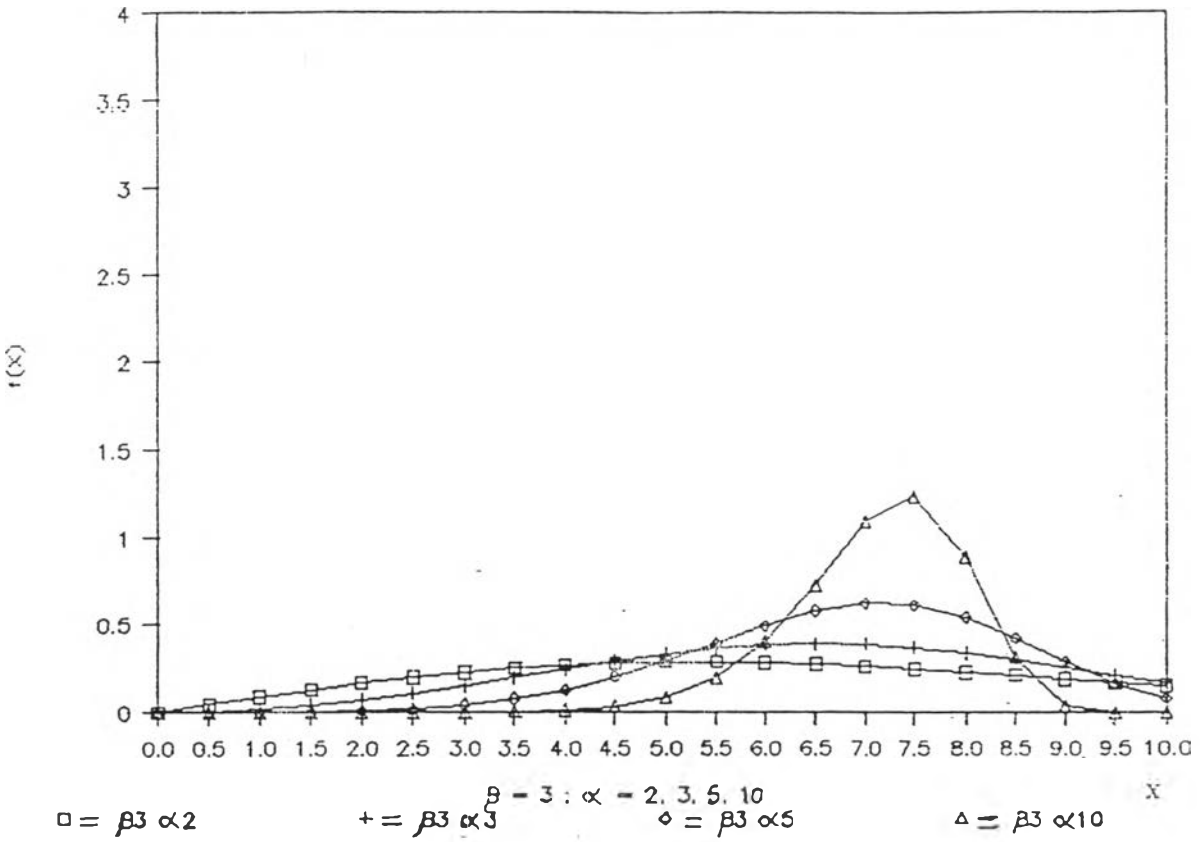
รูปที่ 3.1.1.3 รูปกราฟการแจกแจงแบบแกมมา เมื่อ $\lambda = 3$ และ $\gamma = 2, 3, 5$ และ 10



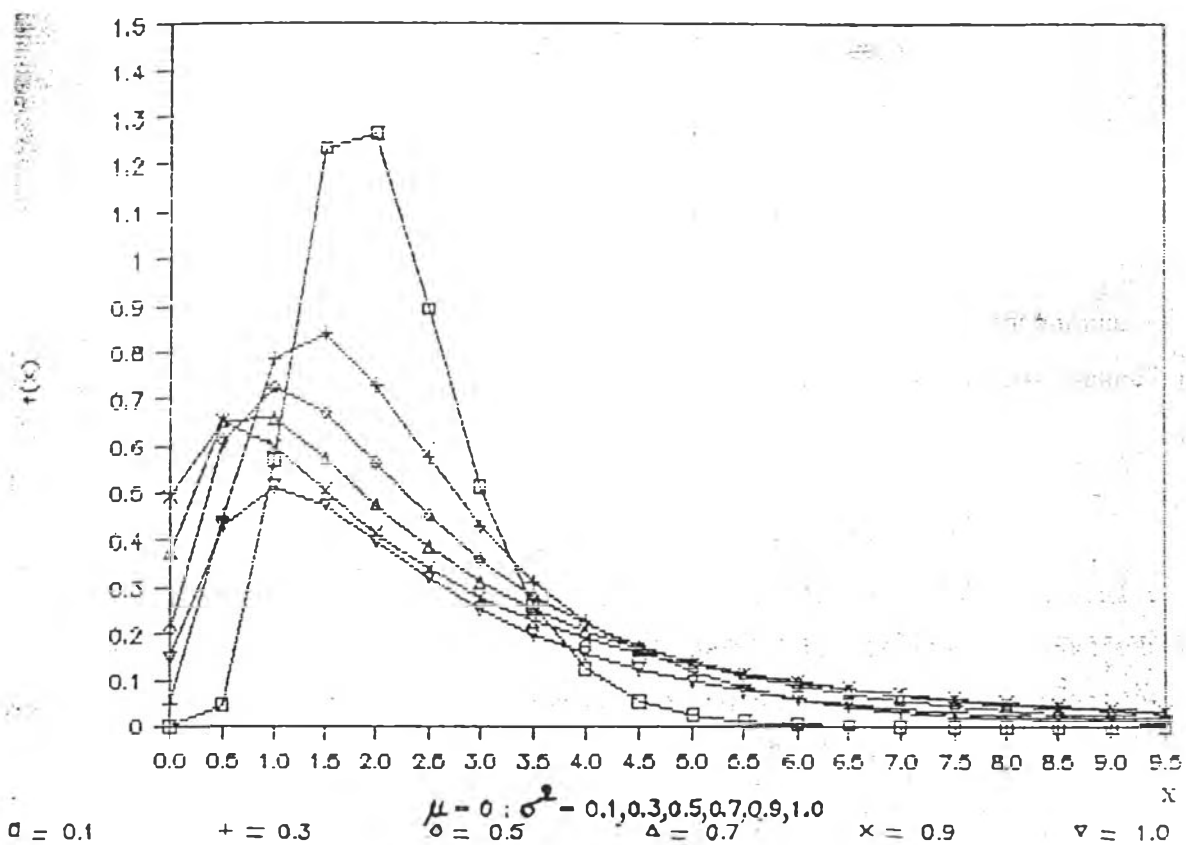
รูปที่ 3.1.1.4 รูปกราฟการแจกแจงแบบไวบูลล์ เมื่อ $\beta = 1$ และ $\alpha = 2, 3, 5$ และ 10



รูปที่ 3.1.1.5 รูปกราฟการแจกแจงแบบไวบูลล์ เมื่อ $\beta = 2$ และ $\alpha = 2, 3, 5$ และ 10



รูปที่ 3.1.1.6 รูปกราฟการแจกแจงแบบไวบูลล์ เมื่อ $\beta = 3$ และ $\alpha = 2, 3, 5$ และ 10



รูปที่ 3.1.1.7 รูปกราฟการแจกแจงแบบลอกนอร์มอล $\mu = 0$ และ $\sigma^2 = 0.1, 0.3, 0.5, 0.7, 0.9$ และ 1.0

3.1.2 กำหนดจำนวนขนาดตัวอย่างที่ศึกษาคือ 10 ,20,30,50 และ 70 ตามลำดับ

3.1.3 กำหนดระดับนัยสำคัญ 3 ระดับ คือ 0.01 , 0.05 และ 0.10 ตามลำดับ

3.1.4 กรณีที่ข้อมูลที่ถูกตัดปลายจะศึกษาในกรณีที่ข้อมูลมีค่าถูกตัดปลายทางขวาจำนวน 10% , 20% และ 30% ตามลำดับ โดยพิจารณาได้ดังนี้

ให้ R เป็นเปอร์เซ็นต์ข้อมูลที่มีค่าถูกตัดปลายทางขวา

จำนวนข้อมูลที่มีค่าถูกตัดปลายในแต่ละกรณีจะแสดงในตารางที่ 3.1 ข้างล่างนี้

ตารางที่ 3.1 แสดงจำนวนข้อมูลที่ถูกตัดปลายทางขวา

R	n=10	n=20	n=30	n=50	n=70
10%	1	2	3	5	7
20%	2	4	6	10	14
30%	3	6	9	15	21

สำหรับข้อมูลตัดปลายนั้นมีวิธีการจำลองข้อมูลดังนี้คือ

1) สร้างเลขสุ่มที่มีการแจกแจงแบบเอกรูป อยู่ในช่วง [0,1] และเป็นอิสระซึ่งกันและกัน โดยตัวเลขสุ่มจากการแจกแจงแบบเอกรูปจะนำไปใช้ในการจำลองตัวแปรสุ่มเลขชี้กำลัง 2 พารามิเตอร์ ตัวแปรสุ่มไวบูลล์ ตัวแปรสุ่มแกมมา และตัวแปรสุ่มลอกนอร์มอล

2) กำหนดเปอร์เซ็นต์ข้อมูลที่ถูกตัดปลาย คือ 10%, 20% และ 30%

3) กำหนดจุดที่พิจารณาให้เป็นจุดตัดปลายทางขวา ซึ่งในการวิจัยจะพิจารณาทำการเปรียบเทียบ 3 จุด คือ $\mu+0.5\sigma$, $\mu+1.0\sigma$ และ $\mu+2.0\sigma$ ตามลำดับ

4) ทำการสุ่มข้อมูลให้เป็นไปตามการแจกแจงที่ต้องการศึกษา โดยหากตัวเลขที่สุ่มได้นั้นเกินจุดตัดปลายที่กำหนด จะให้ข้อมูลนั้นเท่ากับค่าที่อยู่ตำแหน่งจุดตัดพอดี และสุ่มจนครบตามจำนวนเปอร์เซ็นต์ที่ถูกตัดปลาย และครบตามจำนวนข้อมูลที่ต้องการทดสอบ

3.2 การดำเนินการวิจัย

การดำเนินการวิจัยแบ่งออกเป็น 3 ขั้นตอนที่สำคัญ ดังนี้

3.2.1 การสร้างการแจกแจงของประชากรตามลักษณะที่กำหนดในแผนการทดลอง

3.2.2 การคำนวณค่าสถิติทดสอบทั้ง 4 ตัว

3.2.3 การหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ

รายละเอียดของแต่ละขั้นตอนมีดังนี้

3.2.1 การสร้างการแจกแจงของประชากรตามลักษณะที่กำหนดในแผนการทดลอง ใช้โปรแกรมฟอร์แทรนกับเครื่อง PC โดยใช้เลขสุ่มซึ่งมีการแจกแจงแบบเอกรูปในช่วง [0,1](U(0,1)) เป็นพื้นฐานในการสร้าง โดยเลขสุ่ม U(0.1) ที่ได้ควรมีคุณสมบัติดังต่อไปนี้

ก) ตัวเลขที่ได้มีการกระจายของความน่าจะเป็นแบบเอกรูปและเป็นอิสระซึ่งกันและกัน

ข) อนุกรมของตัวเลขที่ได้สามารถสร้างซ้ำเดิมได้ (reproducible)

ค) อนุกรมของตัวเลขไม่ซ้ำเดิมในช่วงที่ต้องการใช้ตัวเลขแบบสุ่ม หมายความว่าขนาดความยาวของอนุกรมตัวเลขต้องยาวพอสำหรับใช้งาน

ง) ใช้เวลาสั้นๆ ในการสร้างตัวเลขแบบสุ่ม

จ) ใช้หน่วยความจำของคอมพิวเตอร์น้อย

สำหรับโปรแกรมย่อยที่ใช้ในการสร้างตัวเลขแบบสุ่ม คือ

SUBROUTINE RAND(IX, IY, YFL) ดังที่ได้แสดงไว้ในภาคผนวก ส่วนรายละเอียดในการสร้างการแจกแจงต่างๆ เป็นดังนี้

3.2.1.1 การผลิตเลขสุ่มที่มีการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์

ขั้นตอนนี้จะสร้างประชากรให้มีการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ ให้มีลักษณะตามที่กำหนดไว้ในแผนการทดลองโดยใช้เลขสุ่มที่มีการแจกแจงเอกรูป (U(0,1)) เป็นพื้นฐานในการสร้าง และเขียนโปรแกรมด้วยภาษาฟอร์แทรนกับเครื่อง PC โปรแกรมที่ใช้ในการสร้างเลขสุ่มที่มีการแจกแจงแบบเอกรูปแสดงไว้ในภาคผนวก เมื่อสร้างเลขสุ่มที่มีการแจกแจงแบบเอกรูปแล้วจะนำเลขสุ่มแบบเอกรูปที่ได้มาสร้างการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ ซึ่งฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x)=\begin{cases}\dfrac{1}{\theta}\exp\left(\dfrac{-(x-\beta)}{\theta}\right) & ,x\geq\beta \quad ,\theta>0 \quad ,-\infty<\beta<\infty \\ 0 & \text{อื่นๆ}\end{cases}$$

เมื่อ β เป็นพารามิเตอร์ตำแหน่ง

และ θ เป็นพารามิเตอร์สเกล

การสร้างเลขสุ่มให้มีการแจกแจงเลขชี้กำลัง 2 พารามิเตอร์ จะใช้เทคนิคการแปลงผกผัน (inverse transformation) ซึ่งเป็นเทคนิคที่ใช้ในการแปลงตัวแปรสุ่มที่มีลักษณะการแจกแจง

แบบเอกรูปให้เป็นรูปแบบของตัวแปรสุ่มที่มีลักษณะการแจกแจงตามแบบที่เราต้องการ สำหรับการสร้างตัวแปรสุ่มให้มีการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ มีขั้นตอนดังนี้

1) หาฟังก์ชันการแจกแจงสะสมของการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ได้

$$F(x) = 1 - \exp(\frac{-(x-\beta.)}{\theta})$$

2) ฟังก์ชันการแจกแจงสะสมที่ได้จากข้อ 1) มีการแจกแจงแบบเอกรูปในช่วง [0,1] ดังนั้นให้

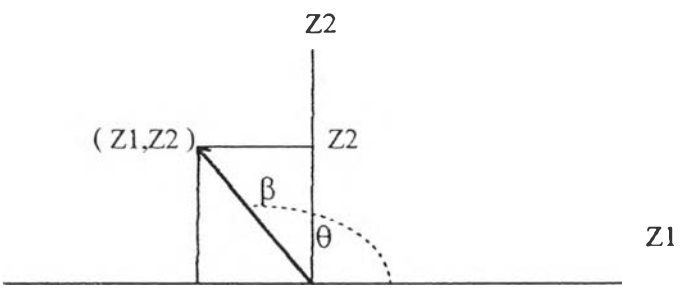
$$F(x) = 1 - \exp(\frac{-(x-\beta.)}{\theta}) = U$$ โดยที่ U คือเลขสุ่มแบบเอกรูป U(0,1)

3) หาค่า x ในเทอมของ U ได้ $x = \beta - \theta \ln(1-U)$

สำหรับรายละเอียดในแต่ละขั้นตอนรวมทั้งคำสั่งที่ใช้ในการสร้างตัวแปรให้มีการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์แสดงไว้ในภาคผนวก

<u>3.2.1.2 การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติ</u>

การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติใช้วิธีการของ Box และ Muller (ค.ศ. 1958) โดยผลิตเลขสุ่มที่มีการแจกแจงปกติมาตรฐาน ที่มีค่าเฉลี่ย 0 และค่าความแปรปรวนเป็น 1 พร้อมๆ กัน 2 ค่า โดยพิจารณาจากรูปต่อไปนี้



Z1

พิจารณาจากรูปจะได้

$$Z1 = \beta\cos(\theta) \quad (1)$$

$$Z2 = \beta\sin(\theta) \quad (2)$$

จาก (1) และ (2) พิสูจน์ได้ว่า β และ θ เป็นอิสระกัน และ $\beta^2 = z_1^2 + z_2^2$ มีการแจกแจงแบบไคกำลังสอง (Chi-square) ด้วยระดับความเสรีเท่ากับ 2 ซึ่งเทียบเท่ากับการแจกแจงแบบเลขชี้

กำลัง ค่าเฉลี่ยเท่ากับ 2 และ θ มีการแจกแจงเอกรูปอยู่ในช่วง 0 ถึง 2π เรเดียน(ridian) ดังนั้นจึงสามารถใช้วิธีการแปลงผกผัน (Inverse tranformation) สร้างเลขสุ่ม ของ β และ θ ได้ดังนี้

$$\beta = (-2 \ln U)^{1/2} \qquad (3)$$

เมื่อ U เป็นเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง [0,1] เพราะฉะนั้นแทนค่า β และ θ ใน (1) และ (2) จะได้เลขสุ่มการแจกแจงปกติมาตรฐาน Z1 , Z2 อิสระกัน

$$Z1 = (-2 \ln R1)^{1/2} \cos (2\pi R1)$$

$$Z2 = (-2 \ln R2)^{1/2} \sin (2\pi R2)$$

จากนี้เมื่อต้องการเลขสุ่มที่มีการแจกแจงปกติที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 ทำได้โดยการแปลงเลขสุ่ม Z1 , Z2 โดยอาศัยฟังก์ชัน :

$$X1 = \mu + \sigma Z1$$

$$X2 = \mu + \sigma Z2$$

ซึ่งจะได้ X1 และ X2 เป็นอิสระกันและต่างก็มีการแจกแจงปกติที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2

โปรแกรมย่อยที่ใช้ในการสร้างเลขสุ่มให้มีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวน เท่ากับ σ^2 คือ

FUNCTION NORMAL (DMEAN , SIGMA) ดังแสดงในภาคผนวก

<u>3.2.1.3 การผลิตเลขสุ่มที่มีการแจกแจงแบบลอกนอร์มอล</u>

การแจกแจงแบบลอกนอร์มอลมีฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} \exp[-(\ln x - \mu)^2 / 2\sigma^2 & ,x > 0 \quad ,\sigma > 0 \\ 0 & ,\text{อื่นๆ} \end{cases}$$

เมื่อ μ และ σ^2 เป็นค่าเฉลี่ย และความแปรปรวนของ Y โดยที่ $Y = \ln X$ ซึ่ง X คือตัวแปรสุ่มที่มีการแจกแจงแบบลอกนอร์มอล และ Y มี σ^2 เป็นพารามิเตอร์สเกล และ μ เป็นพารามิเตอร์รูปทรง ค่าเฉลี่ยและความแปรปรวนของการแจกแจงแบบลอกนอร์มอลคือ $\exp(\mu+\sigma^2/2)$ และ $\exp(2\mu+\sigma^2) *[\exp(\sigma^2)-1]$ ตามลำดับ

สำหรับการสร้างเลขสุ่มที่มีการแจกแจงแบบลอกนอร์มอลหาได้จากการหา exponential ของฟังก์ชัน NORMAL (DMEAN , SIGMA) เมื่อ DMEAN และ SIGMA เป็นค่าเฉลี่ยและความแปรปรวนของการแจกแจงปกติ ดังแสดงในภาคผนวก สำหรับการวิจัยครั้งนี้กำหนดให้ค่า $\mu = 0$ และ $\sigma^2 = 0.7$ และ 0.9 ตามลำดับ

3.3.1.4 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบไวบูลล์</u>

การแจกแจงแบบไวบูลล์มีฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x) = \begin{cases} \dfrac{\alpha . x^{\alpha-1} \exp[-(x/\beta)^{\alpha}}{\beta^{\alpha}} & ,0 < x < \infty \quad ,\beta > 0 \quad ,\alpha > 0 \\ 0 & ,\text{อื่นๆ} \end{cases}$$

เมื่อ β เป็น พารามิเตอร์สเกล

α เป็น พารามิเตอร์รูปทรง

ค่าคาดหวังและความแปรปรวนของการแจกแจงแบบไวบูลล์ คือ $\beta\Gamma(1+1/\alpha)$ และ $\beta^2[\Gamma(1+2/\alpha) - \Gamma^2(1+1/\alpha)]$

การสร้างตัวแปรสุ่มให้มีการแจกแจงแบบไวบูลล์อาศัยเทคนิคการแปลงผกผัน มีขั้นตอนดังนี้

1) หาฟังก์ชันการแจกแจงสะสมของการแจกแจงแบบไวบูลล์ได้

$$F(x) = 1 - \exp[-(x/\beta)^{\alpha}] \quad , x > 0$$

2) ฟังก์ชันการแจกแจงสะสมที่ได้ในข้อ 1) มีการแจกแจงแบบเอกรูปในช่วง [0,1] ดังนั้นให้

$$F(x) = 1 - \exp[-(x/\beta)^{\alpha}] = U \text{ โดยที่ U คือตัวเลขสุ่มแบบเอกรูป}$$

3) หาค่าของ x ในเทอมของ U ได้ $x = \beta[-\ln(U)]^{1/\alpha}$

สำหรับคำสั่งในการสร้างตัวแปรให้มีการแจกแจงแบบไวบูลล์ คือ

FUNCTION WEIBUL (ALPHA , BETA) ดังแสดงในภาคผนวก

ในการวิจัยครั้งนี้กำหนดพารามิเตอร์ $\beta = 1$ และ $\alpha = 0.5$ และ 2.0

3.3.1.5 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบแกมมา</u>

การแจกแจงแบบแกมมามีฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x) = \begin{cases} \dfrac{x^{\gamma-1} \exp(-x/\lambda)}{\lambda^{\gamma}\Gamma(\gamma)} & ,x > 0 \quad ,\gamma > 0 \quad ,\lambda > 0 \\ 0 & ,\text{อื่นๆ} \end{cases}$$

เมื่อ λ เป็น พารามิเตอร์สเกล

γ เป็น พารามิเตอร์รูปทรง

การสร้างเลขสุ่มให้มีการแจกแจงแบบแกมมา ใช้คุณสมบัติ reproductive property เมื่อ x_i เป็นตัวแปรสุ่มจากการแจกแจงแบบ Gamma(G) แล้ว $x = \sum_{i=1}^{n} nx_i$ มีรูปแบบเป็น Gamma(γ,λ) ซึ่ง

$\gamma = \sum_{i=1}^{n} \gamma_i$ ดังนั้นเมื่อ γ เป็นตัวเลขจำนวนเต็ม หรือ $\gamma = m$ สามารถผลิตได้โดยการรวมตัวแปรสุ่มแบบเลขชี้กำลังที่เป็นอิสระ m ตัวดังนี้

$$x = \sum_{i=1}^{n} (-\lambda . \ln U_i)$$

$$= -\lambda \sum_{i=1}^{n} \ln U_i$$

$$= -\lambda . \ln(\prod_{i=1}^{n} U_i)$$

เมื่อ U_i เป็นเลขสุ่มแบบสม่ำเสมอในช่วง (0,1)

สำหรับกรณีที่ค่า γ ไม่เป็นจำนวนเต็ม มีวิธีการผลิตตัวเลขสุ่มดังนี้

การผลิตเลขสุ่มให้มีการแจกแจงแบบแกมมา ที่มีพารามิเตอร์ γ และ λ มีขั้นตอนดังนี้

1. คำนวณ $a = \sqrt{2\gamma - 1}$

$$b = 2.\gamma - 2\ln 2 + 1/a$$

2. ผลิตเลขสุ่ม RD_1 และ RD_2 จากโปรแกรมย่อย SUBROUTINE RAND (IX,RD)

3. คำนวณ $x = \gamma [RD_1 / (1 - RD_2)]^a$

4. ถ้า $x > b - \ln(RD_1^2 . RD_2)$ ไม่ยอมรับค่า x ให้กลับไปทำขั้นที่ 2 แต่ถ้า $x \le b - \ln(RD_1^2 . RD_2)$ ได้ค่า X= x ซึ่งจะมีค่าเฉลี่ย และความแปรปรวนเท่ากับ γ เพราะฉะนั้นเมื่อต้องการหาค่า X ที่มีค่าเฉลี่ย $\gamma\lambda$ และค่าความแปรปรวน $\gamma\lambda^2$ ให้ทำในขั้นที่ 5 ต่อไป

5. ค่า $X = x\lambda$

สำหรับคำสั่งในการสร้างเลขสุ่มแบบแกมมา คือ

FUNCTION GAMMA1 (GAM1 ,LAM1) ดังแสดงในภาคผนวก

ในการวิจัยครั้งนี้กำหนดพารามิเตอร์ $\lambda = 1$ และ $\gamma = 1, 2, 3$ ตามลำดับ

3.2.2 การคำนวณค่าตัวสถิติทดสอบ 4 วิธี

เราสุ่มตัวอย่างจากประชากรโดยใช้โปรแกรมย่อยที่เขียนในภาคผนวกตามขนาดตัวอย่างและเปอร์เซ็นต์ของข้อมูลที่ถูกตัดปลายที่กำหนดในแผนการทดลองแล้วนำข้อมูลที่ได้ไปคำนวณค่าต่าง ๆ ตามสูตรของตัวสถิติทดสอบแต่ละวิธีที่เสนอในบทที่ 2 เมื่อได้ค่าของตัวสถิติทดสอบแต่ละตัวให้เปรียบเทียบกับค่าวิกฤติ โดยตัวสถิติทดสอบ G เปรียบเทียบกับค่าวิกฤติที่เปิดจาก percentile of Gini เมื่อขนาดตัวอย่างน้อยกว่าหรือเท่ากับ 20 และเปรียบเทียบกับค่าวิกฤติที่เปิดจากตารางการแจกแจงปกติมาตรฐาน เมื่อขนาดตัวอย่างมากกว่า 20 ตัวสถิติทดสอบ L เปรียบเทียบกับค่าวิกฤติที่เปิดจาก percentile of Lorenz เมื่อขนาดตัวอย่างน้อยกว่าหรือเท่ากับ 20 และเปรียบเทียบกับค่าวิกฤติที่เปิดจากตารางการแจกแจงปกติมาตรฐาน เมื่อขนาดตัวอย่างมากกว่า 20 ตัวสถิติ

ทดสอบ K-S เปรียบเทียบกับค่าวิกฤติที่เปิดจากตาราง Kolmogorov - Smirnov ตัวสถิติทดสอบ Anderson- Darling เปรียบเทียบกับค่าวิกฤติที่เปิดจากตารางที่ปรับปรุงแล้วของM.A.Stephens ซึ่งการยอมรับหรือการปฏิเสธสมมติฐานว่างนั้นให้ถือเกณฑ์ในบทที่ 2

3.2.3 การหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ

1) วิธีการดำเนินการในขั้นแรกจะทำการทดสอบความคลาดเคลื่อนประเภทที่ 1 โดยทำการสุ่มตัวอย่างจากการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ ตามสถานการณ์ที่กำหนดไว้ในขั้นตอนการวิจัย เพื่อนำมาทดสอบกับตัวสถิติทดสอบ 4 ตัว คือตัวสถิติทดสอบ G, ตัวสถิติทดสอบ L ,ตัวสถิติทดสอบ K-S และตัวสถิติทดสอบ A-D จากนั้นคำนวณค่าตัวสถิติทดสอบแต่ละตัวพร้อมทั้งเปรียบเทียบกับค่าวิกฤติของแต่ละตัวสถิติทดสอบที่กล่าวถึงในบทที่ 2 ทำซ้ำ ๆ กันเป็นจำนวน 1,000 ครั้ง เมื่อครบแล้วให้นับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง H_0: การแจกแจงเป็นแบบเลขชี้กำลัง 2 พารามิเตอร์ โดยสัดส่วนจำนวนครั้งที่จะปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นจริง หาได้จากการหารจำนวนครั้งของการปฏิเสธสมมติฐานว่างด้วย 1,000 ซึ่งเป็นจำนวนครั้งในการทดลอง ทำเช่นนี้จนครบทุกสถานการณ์ของการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์ จากนั้นพิจารณาเปรียบเทียบกับเกณฑ์การประมาณแบบช่วงคือ ถ้าสัดส่วนจำนวนครั้งปฏิเสธสมมติฐาน H_0 มีค่าอยู่ในช่วง [0,0.017] , [0,0.061] และ [0,0.012] ณ ระดับนัยสำคัญ 0.01 , 0.05 และ 0.10 ตามลำดับ จะถือว่าตัวสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 จากนั้นพิจารณาดูว่าตัวสถิติทดสอบใดสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้จึงนำมาทดสอบหาอำนาจการทดสอบต่อไป

2)ในกรณีหาค่าอำนาจการทดสอบมีวิธีดำเนินการคือทำการสุ่มตัวอย่างจากการแจกแจงแบบแบบแกมมา , การแจกแจงแบบไวบูลล์ และการแจกแจงแบบลอกนอร์มอล ตามสถานการณ์ที่กำหนดไว้ในขั้นตอนการวิจัย เพื่อนำมาทดสอบกับตัวสถิติทดสอบ 4 ตัว คือตัวสถิติทดสอบ G, ตัวสถิติทดสอบ L ,ตัวสถิติทดสอบ K-S และตัวสถิติทดสอบ A-D จากนั้นคำนวณค่าตัวสถิติทดสอบแต่ละตัวพร้อมทั้งเปรียบเทียบกับค่าวิกฤติของแต่ละตัวสถิติทดสอบที่กล่าวถึงในบทที่ 2 ทำซ้ำ ๆ กันเป็นจำนวน 1,000 ครั้ง เมื่อครบแล้วให้นับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง H_0: การแจกแจงเป็นแบบเลขชี้กำลัง 2 พารามิเตอร์ โดยความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 2 หาได้จากการหารจำนวนครั้งของการปฏิเสธสมมติฐานว่างด้วย 1,000 ซึ่งเป็นจำนวนครั้งในการทดลอง ทำเช่นนี้จนครบทุกสถานการณ์ของการแจกแจงแบบแกมมา , การแจกแจงแบบไวบูลล์ และการแจกแจงแบบลอกนอร์มอล จากนั้นพิจารณาเปรียบเทียบกับเกณฑ์การประมาณแบบช่วงคือ ถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 2 มีค่าอยู่ในช่วง [0,0.017] , [0,0.061] และ [0,0.012] ณ ระดับ

นัยสำคัญ 0.01 , 0.05 และ 0.10 ตามลำดับจะถือว่าตัวสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 2 จากนั้นพิจารณาเปรียบเทียบตัวสถิติทดสอบแต่ละตัวว่ามีอำนาจการทดสอบสูงสุด ในกรณีใดบ้าง และหากโดยส่วนใหญ่ตัวสถิติทดสอบตัวใดมีอำนาจสูงสุดในทุกกรณีจะถือว่าตัวสถิติทดสอบตัวนั้นมีอำนาจการทดสอบสูงสุด

เกณฑ์ในการทดสอบค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จะใช้เกณฑ์การเปรียบเทียบโดยการคำนวณช่วงความเชื่อมั่นของอัตราความคลาดเคลื่อนที่ระบุ ซึ่งมีวิธีการคำนวณดังนี้

วิธีคำนวณเกณฑ์ในการตัดสินอัตราความคลาดเคลื่อนที่ระบุ (nominated) ซึ่งสามารถคำนวณจากช่วงความเชื่อมั่นของ p เมื่อ p หมายถึงโอกาสที่เกิดจากความคลาดเคลื่อนประเภทที่ 1 ดังนี้

$$0 \le \hat{p} \le p + Z_\alpha \sqrt{\frac{pq}{n}}$$

โดยที่ α คือระดับนัยสำคัญหรือ Type I Error ที่กำหนดในการทดสอบ ในการวิจัยครั้งนี้กำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.01, 0.05 และ 0.10

1.ระดับนัยสำคัญ $\alpha = 0.01$

จะได้ว่า $p = 0.01 \quad ,q = 1 - p = 0.99 \quad ,n = 1{,}000 \quad ,Z_\alpha = 2.326$

ดังนั้นจะได้ $0.01+2.326\{0.01(0.99)/1000)^{1/2}\} = 0.017$

2. ระดับนัยสำคัญ $\alpha = 0.05$

จะได้ว่า $p = 0.05 \quad ,q = 1 - p = 0.95 \quad ,n = 1{,}000 \quad ,Z_\alpha = 1.645$

ดังนั้นจะได้ $0.05+1.645\{0.05(0.95)/1000)^{1/2}\} = 0.061$

3. ระดับนัยสำคัญ $\alpha = 0.10$

จะได้ว่า $p = 0.10 \quad ,q = 1 - p = 0.90 \quad ,n = 1{,}000 \quad ,Z_\alpha = 1.282$

ดังนั้นจะได้ $0.10+1.282\{0.10(0.90)/1000)^{1/2}\} = 0.112$

สรุปช่วงความเชื่อมั่นสำหรับ p = 0.01 คือ $0 \le \hat{p} \le 0.017$

p = 0.05 คือ $0 \le \hat{p} \le 0.016$

p = 0.10 คือ $0 \le \hat{p} \le 0.112$

สถานการณ์สำหรับการวิจัยแบ่งออกได้ดังนี้

ก. ระดับนัยสำคัญ 3 ระดับ คือ 0.01 , 0.05 และ 0.10

ข. ขนาดตัวอย่างของประชากร 5 ระดับ คือ 10 , 20 , 30 , 50 และ 70

ค. ตัวสถิติทดสอบ 4 วิธี คือ G , L ,K-S และ A-D

ง. ประชากร 15 ประชากร

จ. เปอร์เซ็นต์ของข้อมูลคัดปลาย 3 สถานการณ์ กล่าวคือ

- ข้อมูลคัดปลายทางขวา 10% , 20% และ 30% ของจำนวนข้อมูลตามลำดับ

ซึ่งในกรณีนี้ใช้ตัวสถิติทดสอบ 4 ตัวคือ ตัวสถิติทดสอบ G , ตัวสถิติทดสอบ L ,ตัวสถิติทดสอบ K-S และตัวสถิติทดสอบ A-D

ดังนั้น จำนวนสถานการณ์ที่ใช้ทดลองวิจัย = สถานการณ์ที่วิเคราะห์ข้อมูลสมบูรณ์ +
สถานการณ์ที่วิเคราะห์โดยเกิดข้อมูลคัดปลายทางขวา

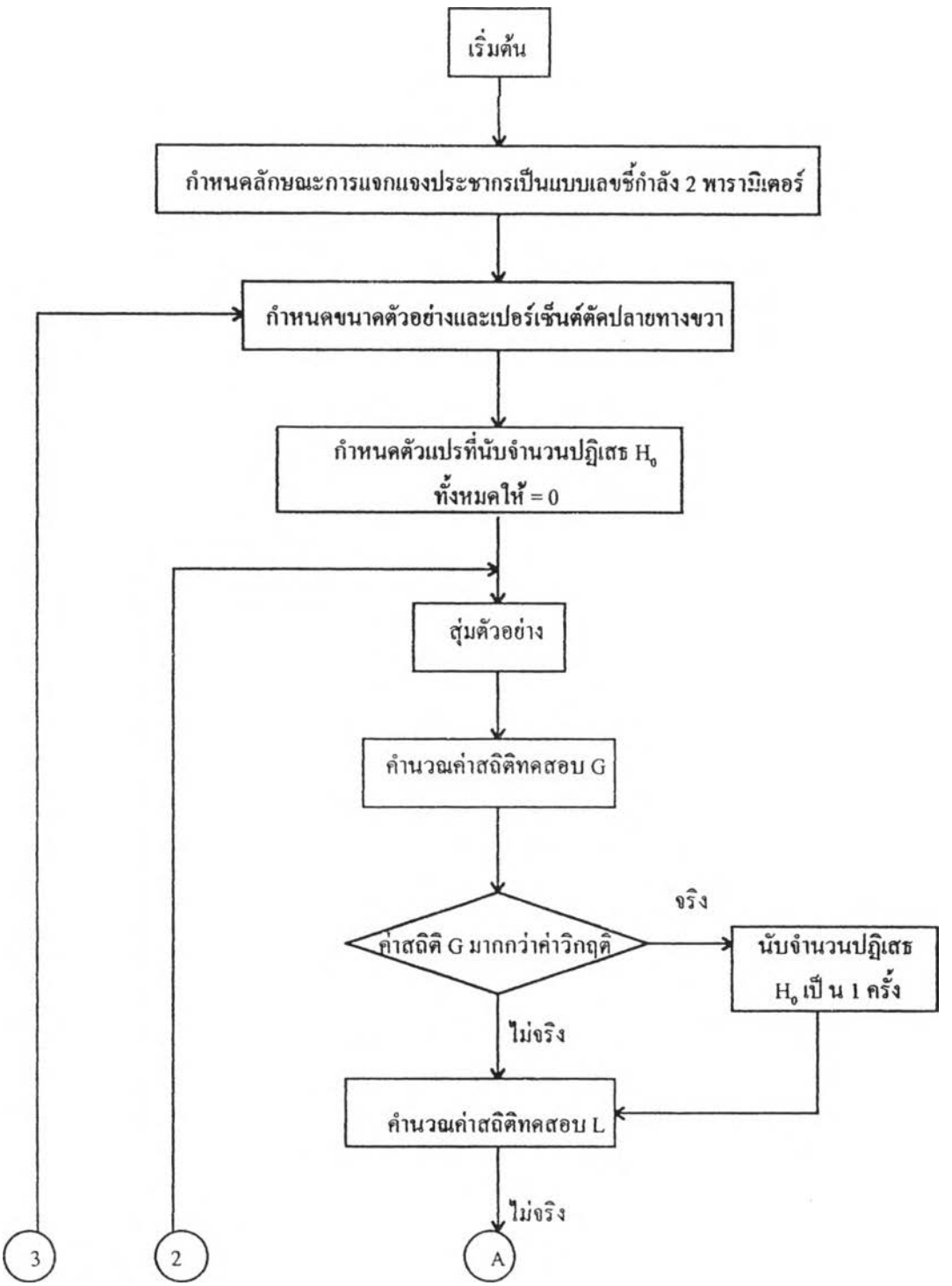
= (3 x 5 x 4 x 15) + (3 x 5 x 4 x 15 x 3)

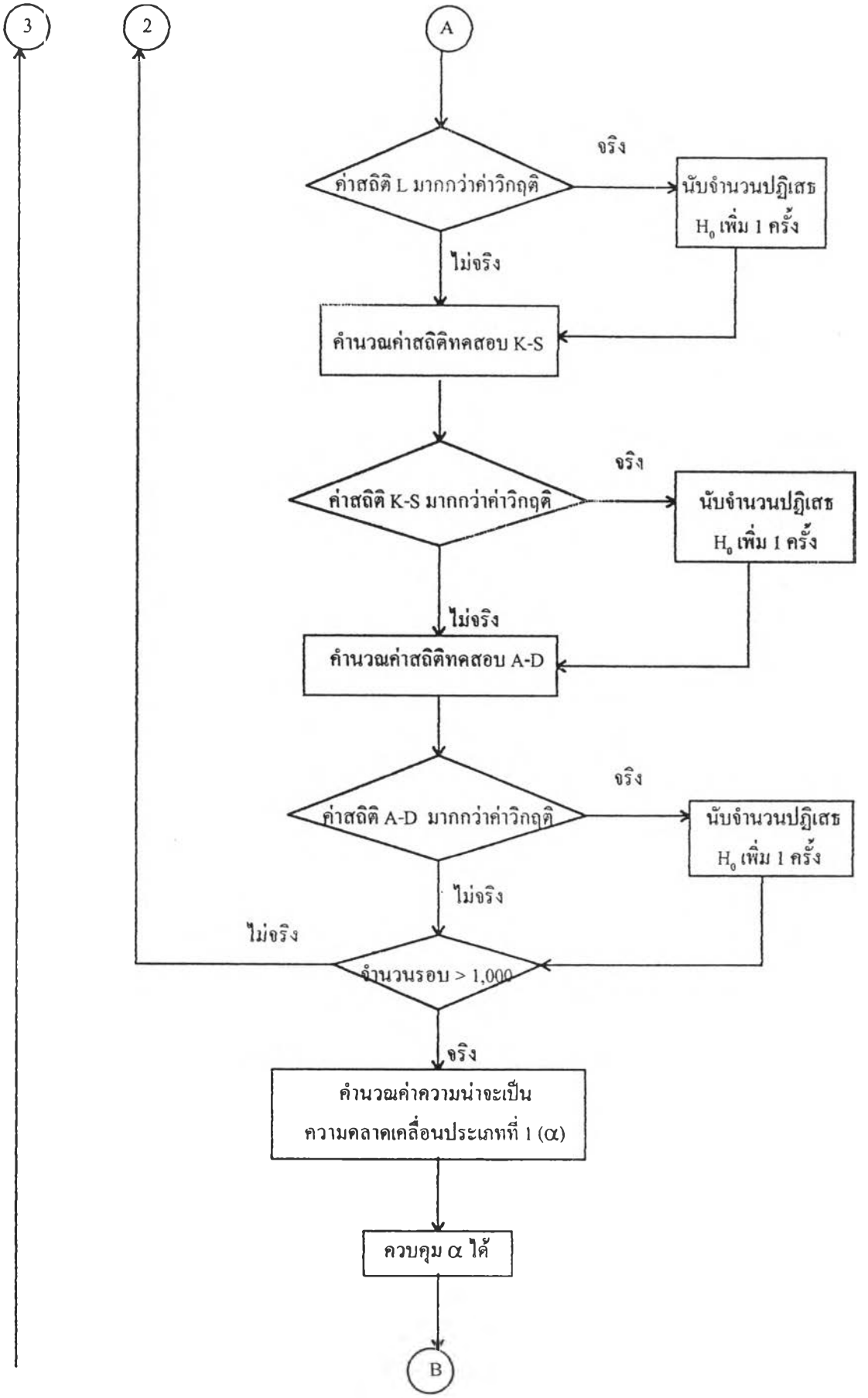
= 900 + 2700

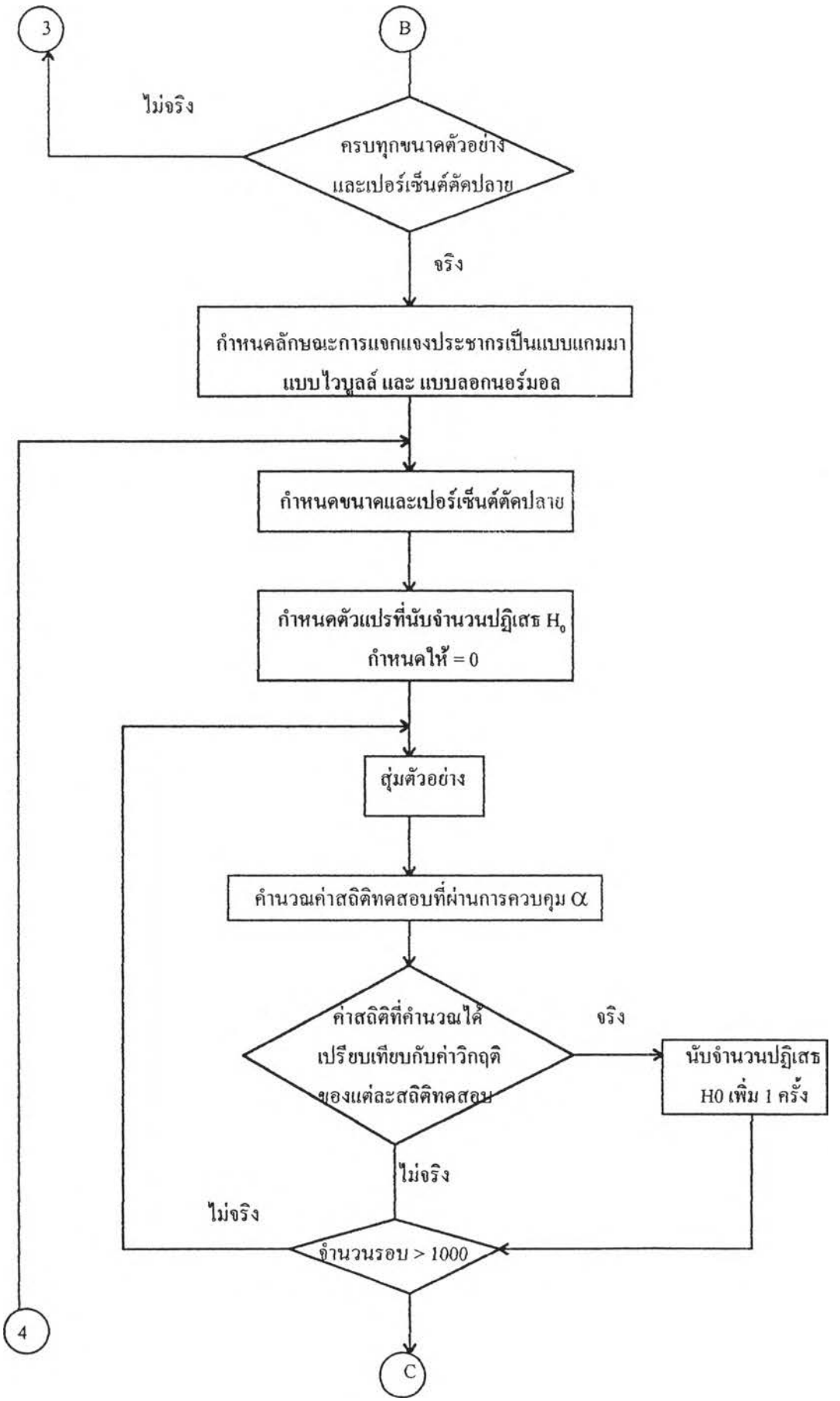
= 3600 สถานการณ์

3.3 ขั้นตอนในการทำงานของโปรแกรม

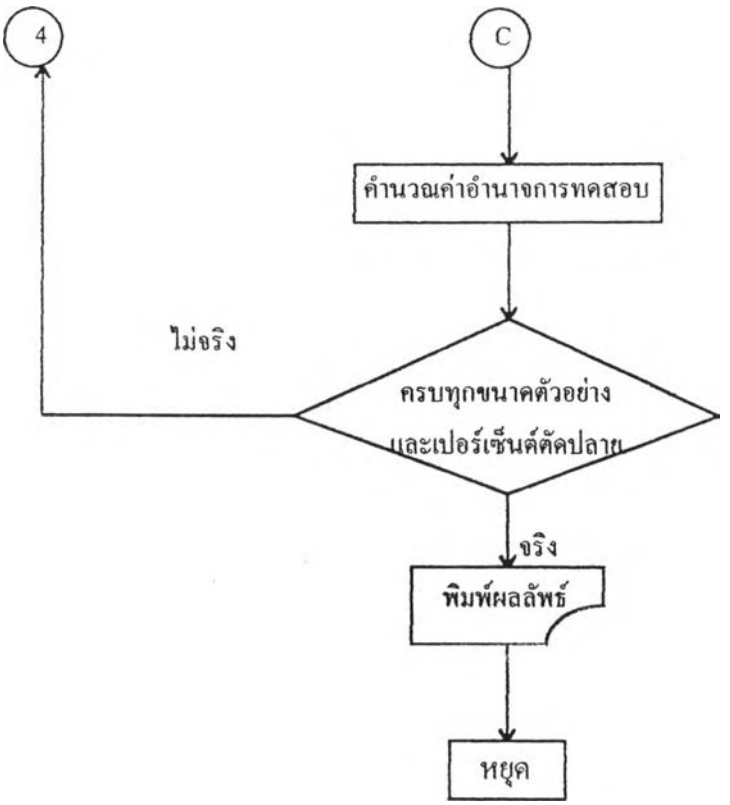
ขั้นตอนในการทำงานของโปรแกรมการหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบ ในส่วนของโปรแกรมอธิบายเป็นลำดับขั้นตอนได้ดังนี้




3
2
A
ค่าสถิติ L มากกว่าค่าวิกฤติ
จริง
นับจำนวนปฏิเสธ H_0 เพิ่ม 1 ครั้ง
ไม่จริง
คำนวณค่าสถิติทดสอบ K-S
ค่าสถิติ K-S มากกว่าค่าวิกฤติ
จริง
นับจำนวนปฏิเสธ H_0 เพิ่ม 1 ครั้ง
ไม่จริง
คำนวณค่าสถิติทดสอบ A-D
ค่าสถิติ A-D มากกว่าค่าวิกฤติ
จริง
นับจำนวนปฏิเสธ H_0 เพิ่ม 1 ครั้ง
ไม่จริง
ไม่จริง
จำนวนรอบ > 1,000
จริง
คำนวณค่าความน่าจะเป็น
ความคลาดเคลื่อนประเภทที่ 1 (α)
ควบคุม α ได้
B


3
B
ไม่จริง
ครบทุกขนาดตัวอย่าง
และเปอร์เซ็นต์ตัดปลาย
จริง
กำหนดลักษณะการแจกแจงประชากรเป็นแบบแกมมา
แบบไวบูลล์ และ แบบลอกนอร์มอล
กำหนดขนาดและเปอร์เซ็นต์ตัดปลาย
กำหนดตัวแปรที่นับจำนวนปฏิเสธ H_0
กำหนดให้ = 0
สุ่มตัวอย่าง
คำนวณค่าสถิติทดสอบที่ผ่านการควบคุม α
ค่าสถิติที่คำนวณได้
เปรียบเทียบกับค่าวิกฤติ
ของแต่ละสถิติทดสอบ
จริง
นับจำนวนปฏิเสธ
H0 เพิ่ม 1 ครั้ง
ไม่จริง
ไม่จริง
จำนวนรอบ > 1000
4
C

I18996298



รายละเอียดของแผนผังโปรแกรมในการคำนวณความน่าจะเป็นของค่าความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติที่ใช้ทดสอบการแจกแจงแบบเลขชี้กำลัง 2 พารามิเตอร์จะเสนอเป็นขั้นตอนต่อไปนี้

<u>ขั้นที่ 1</u>

สร้างประชากรให้มีการแจกแจงลักษณะต่าง ๆ ดังนี้

การแจกแจงแบบเลขชี้กำลังที่มี 2 พารามิเตอร์ โดยใช้คำสั่ง FUNCTION EXPO (BETA1,TETA1)

การแจกแจงแบบลอกนอร์มอล โดยใช้คำสั่ง EXP (NORMAL(DMEAN,SIGMA)

การแจกแจงแบบไวบูลล์ โดยใช้คำสั่ง FUNCTION WEIBUL (ALPHA,BETA)

การแจกแจงแบบแกมมา โดยใช้คำสั่ง FUNCTION GAMMA1 (GAM1 , LAM1)

การกำหนดขนาดตัวอย่างและการคำนวณค่าตัวสถิติทดสอบ 4 ตัวคือ G , L , K-S และ A-Dทำได้โดยการเลือกประชากรครั้งละ 1 ประชากร และจะหยุดการทดลองเมื่อทำการทดลองครบทุกกรณีที่กำหนดของแต่ละประชากร

ขั้นที่ 2

กำหนดขนาดตัวอย่างและเปอร์เซ็นต์ข้อมูลตัดปลายทางขวา

โดยที่ n คือขนาดตัวอย่าง

R คือ เปอร์เซ็นต์ข้อมูลที่ถูกตัดปลายทางขวา

ขั้นที่ 3

กำหนดตัวแปรที่นับจำนวนการปฏิเสธ H_0 ให้มีค่าเท่ากับ 0

โดยที่ G01 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ G ณ ระดับนัยสำคัญ 0.01

G05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ G ณ ระดับนัยสำคัญ 0.05

G10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ G ณ ระดับนัยสำคัญ 0.10

L01 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ L ณ ระดับนัยสำคัญ 0.01

L05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ L ณ ระดับนัยสำคัญ 0.05

L10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ L ณ ระดับนัยสำคัญ 0.10

KS01 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ K-S ณ ระดับนัยสำคัญ 0.01

KS05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ K-S ณ ระดับนัยสำคัญ 0.05

KS10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ K-S ณ ระดับนัยสำคัญ 0.10

AD01 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ A-D ณ ระดับนัยสำคัญ 0.01

AD05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ A-D ณ ระดับนัยสำคัญ 0.05

AD10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ A-D ณ ระดับนัยสำคัญ 0.10

ขั้นที่ 4

จัดเรียงข้อมูลซึ่งได้จากขั้นที่ 1 โดยเรียงข้อมูลจากน้อยไปมาก โดยใช้ SUBROUTINE RANK (X)

ขั้นที่ 5

คำนวณตัวสถิติทดสอบ G , L , K-S และ A-D

ขั้นที่ 6

ทำการเปรียบเทียบตัวสถิติทดสอบที่คำนวณได้ในขั้นที่ 5 กับค่าวิกฤติที่กำหนดขึ้น ถ้าการเปรียบเทียบปรากฏผลว่ายอมรับสมมติฐาน H_0 ให้พิมพ์ผลลัพธ์แต่ถ้าการเปรียบเทียบปรากฏผลว่าปฏิเสธสมมติฐาน H_0 ให้นับจำนวนครั้งที่ปฏิเสธสมมติฐานเพิ่มขึ้น 1 ครั้งของตัวสถิติทดสอบแต่ละตัว แล้วให้พิมพ์ผลลัพธ์ตัวสถิติทดสอบแต่ละตัวและ ทำการทดลองซ้ำตั้งแต่ ขั้นตอนที่ 4 ถึงขั้นตอนที่ 6 1,000 ครั้ง ในแต่ละสถานการณ์เพื่อหาผลลัพธ์ของค่าความน่าจะเป็นของค่าคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบ G , L , K-S และ A-D ตามลำดับ