บทที่ 3

วิธีการดำเนินการวิจัย

การวิจัยเรื่องนี้ต้องการเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติทดสอบ 3 ตัว ดังที่กล่าวมาแล้วในบทที่ 2 โดยใช้การจำลอง (Simulation) ด้วยเทคนิคมอนติคาร์โล (Monte Carlo Technique) โดยใช้เครื่องคอมพิวเตอร์ AMDAHL 5860 เขียนโปรแกรมด้วยภาษาฟอร์แทรน (Fortran)ทำการจำลองข้อมูลซ้ำ ๆ กันจำนวน 500 ครั้ง ในแต่ละสถานการณ์ของการวิจัย สำหรับขั้นตอนที่สำคัญของการวิจัยเป็นดังนี้

3.1 แผนการทดลอง

เปรียบเทียบความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบของตัวสถิติ 3 ตัว โดยกำหนดสถานการณ์ที่จะทำการเปรียบเทียบดังนี้

3.1.1 สุ่มตัวอย่างจากประชากรโดยกำหนดให้ประชากร มีการแจกแจงดังนี้

ก. การแจกแจงแบบเอกซโปเนนเชียล เมื่อพารามิเตอร์ $\beta=1$

ข. การแจกแจงแบบแกมม่า เมื่อพารามิเตอร์ $\beta=1, \alpha=2,3$ ตามลำดับ

ค. การแจกแจงแบบไวบูลล์ เมื่อพารามิเตอร์ $\beta=1, \alpha=0.5, 2.0$ ตามลำดับ

ง. การแจกแจงแบบลอกนอร์มอล เมื่อพารามิเตอร์ $\mu=0$ และ $\sigma^2= 0.7, 0.9$ ตามลำดับ

จ. การแจกแจงแบบไคสแควร์ ด้วยระดับความเป็นอิสระ 2 และ 3 ตามลำดับ

3.1.2 กำหนดขนาดตัวอย่าง (Sample size) เป็น 5 ระดับ คือ 10,20,30, 50 และ 100 ตามลำดับ

กรณีที่ข้อมูลขาดหาย การวิเคราะห์ข้อมูลพิจารณาได้ดังนี้

ให้ Y_{r1+1} เป็นข้อมูลสุ่มที่เล็กที่สุด

Y_{n-r2} เป็นข้อมูลสุ่มที่ใหญ่ที่สุด

r1 เป็นจำนวนข้อมูลขาดหายทางซ้าย

r2 เป็นข้อมูลขาดหายทางขวา

กรณีต่าง ๆ ที่นำมาวิเคราะห์ สามารถสรุปได้ดังนี้

กรณีที่	n=10		n=20		n=30		n=50		n=100	
	r1	r2	r1	r2	r1	r2	r1	r2	r1	r2
1	0	0	0	0	0	0	0	0	0	0
2	1	0	2	0	3	0	5	0	10	0
3	2	0	4	0	6	0	10	0	20	0
4	3	0	6	0	9	0	15	0	30	0
5	0	1	0	2	0	3	0	5	0	10
6	0	2	0	4	0	6	0	10	0	20
7	0	3	0	6	0	9	0	15	0	30
8	1	1	2	2	3	3	5	5	10	10
9	1	2	2	4	3	6	5	10	10	20
10	2	1	4	2	6	3	10	5	20	10

3.2 ขั้นตอนการทดลอง

แบ่งออกเป็น 3 ขั้นตอน คือ

3.2.1 สร้างการแจกแจงของประชากรตามลักษณะที่กำหนดในแผนการทดลอง

3.2.2 คำนวณค่าของตัวสถิติ 3 ตัว

3.2.3 หาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และหาค่าอำนาจการทดสอบ

รายละเอียดในแต่ละขั้นตอนมีดังนี้

3.2.1 สร้างการแจกแจงของประชากรตามลักษณะที่กำหนดในแผนการทดลอง โดยใช้โปรแกรมภาษาฟอร์แทรน 77 กับเครื่อง AMDAHL 5860 โดยใช้เลขสุ่ม (Random Number) ซึ่งมีการแจกแจงแบบสม่ำเสมอในช่วง (0,1) เป็นพื้นฐานในการสร้าง

สำหรับโปรแกรมย่อยที่ใช้ในการสร้างตัวเลขแบบสุ่ม คือ Subroutine Rand(IX,IY,YFL) ดังแสดงไว้ในภาคผนวก ส่วนรายละเอียดในการสร้างการแจกแจงแบบต่างๆ เป็นดังนี้

3.2.1.1 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบแกมม่า</u>

จากฟังก์ชันความหนาแน่นของการแจกแจงแบบแกมม่าที่เสนอไว้ในบทที่ 1

α เป็น Shape Parameter แสดงถึงรูปร่างของการแจกแจง

β เป็น Scale Parameter แสดงถึงขนาดของการแจกแจง

ซึ่งมีค่าคาดหวังและความแปรปรวนของการแจกแจง คือ $\alpha\beta$ และ $\alpha\beta^2$ ตามลำดับ

กรณีที่ $\alpha = 1$ การแจกแจงแบบแกมม่า คือ การแจกแจงแบบเอกซโปเนนเชียล การสร้างเลขสุ่มที่มีการแจกแจงแบบเอกซโปเนนเชียล นั้นอาศัยเทคนิคการแปลงผกผัน(Inverse Transformation) ซึ่งเป็นเทคนิคที่ใช้ในการแปลงตัวแปรสุ่มที่มีลักษณะการกระจายความน่าจะเป็นแบบสม่ำเสมออยู่ในรูปแบบของตัวแปรสุ่มที่มีลักษณะการกระจายความน่าจะเป็นแบบอื่น ๆ โดยขั้นตอนในการสร้างตัวแปรสุ่มให้มีการแจกแจงแบบเอกซโปเนนเชียลมีดังนี้

ขั้นที่ 1. กำหนดฟังก์ชันการแจกแจงสะสม

$$F(x) = 1-\exp(-(x/\beta)) \quad ; x > 0$$

ขั้นที่ 2. ให้ $F(x) = 1-\exp(-(x/\beta)) = R$ โดยที่ R คือตัวเลขสุ่มสม่ำเสมออยู่ในช่วง (0,1)

ขั้นที่ 3. หาค่าของ x ในเทอมของ R ได้ $x = -\beta \ln R$

สำหรับการสร้างเลขสุ่มที่มีการแจกแจงแบบแกมม่า (α,β) นั้นสร้างได้จากผลรวมของ x_i $(i=1,2,,...,\alpha)$ โดยที่ x_i มีการแจกแจงแบบเอกซโปเนนเชียล ด้วยค่าเฉลี่ย β

กำหนดให้ $x = \Sigma x_i$ จะได้ว่า x Gamma(α,β)

ดังนั้น $x = \Sigma (-\beta \ln R_i)$

$= -\beta \Sigma \ln R_i$

$= -\beta \ln(\prod_{i=1}^{\alpha} R_i)$

คำสั่งในการสร้างตัวแปรให้มีการแจกแจงแบบแกมม่า คือ Function Gamma (Alpha,Beta) ในการวิจัยครั้งนี้กำหนดพารามิเตอร์ Beta(β) = 1 และ Alpha(α) = 1,2 และ 3

3.2.1.2 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบไวบูลล์</u>

จากฟังก์ชันความหนาแน่นของการแจกแจงแบบไวบูลล์ ที่เสนอไว้ในบทที่ 1

α เป็น Shape Parameter แสดงรูปร่างของการแจกแจง

β เป็น Scale Parameter แสดงขนาดของการแจกแจง มีค่าคาดหวังและความแปรปรวนของการแจกแจง คือ $\beta\Gamma(1+1/\alpha)$ และ $\beta^2[\Gamma(1+2/\alpha)- \Gamma^2(1+1/\alpha)]$ การสร้างตัวแปรสุ่มให้มีการแจกแจงแบบไวบูลล์อาศัยการแปลงผกผัน (Inverse Transformat -ion)โดยมีขั้นตอนการสร้างดังนี้

ขั้นที่ 1. กำหนดฟังก์ชันการแจกแจงสะสม $F(x)=1-\exp[-(x/\beta)^{\alpha}]$;$x > 0$

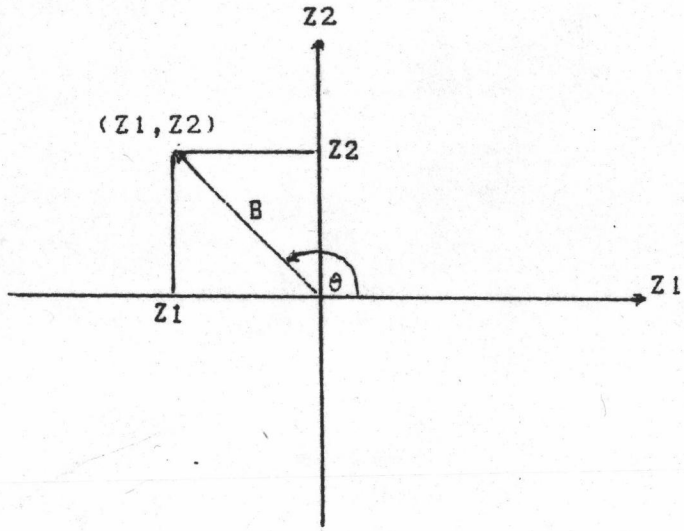
ขั้นที่ 2. ให้ $F(x)=1-\exp[-(x/\beta)^{\alpha}]=R$ โดยที่ R คือตัวเลขสุ่มแบบสม่ำเสมอ

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

คำสั่งในการสร้างตัวแปรให้มีการแจกแจงแบบไวบูลล์ คือ Function Weibull(Alpha, Beta) ดังแสดงในภาคผนวก ในการวิจัยครั้งนี้กำหนดพารามิเตอร์ Beta(β) = 1 และ Alpha(α) = 0.5 และ 2.0

3.2.1.3 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติ</u>

การผลิตเลขสุ่มที่มีการแจกแจงแบบปกติใช้วิธีของ Box และ Muller(ค.ศ.1958) โดยผลิตเลขสุ่มที่มีการแจกแจงปกติมาตรฐาน ที่มีค่าเฉลี่ย 0 และความแปรปรวนเป็น 1 พร้อม ๆ กัน 2 ค่า โดยพิจารณาจากรูปต่อไปนี้



พิจารณาจากรูปจะได้

$$Z1 = B\cos(\theta) \qquad (1)$$

$$Z2 = B\sin(\theta) \qquad (2)$$

เนื่องจาก $B^2 = Z1^2 + Z2^2$ มีการแจกแจงไคสแควร์ด้วยระดับความเป็นอิสระ 2 และเทียบเท่าการแจกแจงเอกซโปเนนเชียล ค่าเฉลี่ยเท่ากับ 2 โดยวิธีการแปลงผกผัน (Inverse Transformation) สามารถสร้างเลขสุ่มที่มีการแจกแจงแบบเอกซโปเนนเชียลได้ดังนี้

$$B = (-2\ln R)^{1/2} \qquad (3)$$

เมื่อ R เป็นเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอในช่วง (0,1)

จากการสมมาตรของการแจกแจงปกติ(Normal Distribution) จะได้ว่ามุม θ มีการแจกแจงแบบสม่ำเสมอระหว่าง 0 ถึง 2 เรเดียน และรัศมี B กับมุม θ เป็นอิสระซึ่งกันและกัน จาก (1),(2),(3) เราสามารถสร้างเลขสุ่มที่มีการแจกแจงปกติมาตรฐาน จากเลขสุ่ม 2 ชุด R1 และ R2 กล่าวคือ

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

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

ในกรณีที่ต้องการเลขสุ่มที่มีการแจกแจงปกติที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 ทำได้โดยการแปลงเลขสุ่ม Z1,Z2 โดยอาศัยฟังก์ชัน

$$\text{Normal} = \mu + \sigma Z1$$

$$\text{Normal} = \mu + \sigma Z2$$

โปรแกรมย่อยที่ใช้ในการสร้างเลขสุ่มให้มีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 คือ Function Normal(Mean,Sigma) แสดงในภาคผนวก

3.2.1.4 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบลอกนอร์มอล</u>

จากฟังก์ชันความหนาแน่นแบบลอกนอร์มอลที่เสนอไว้ในบทที่ 1 เมื่อ μ และ σ^2 เป็นค่าเฉลี่ยและความแปรปรวนของ y โดยที่ y = lnx จะมีการแจกแจงปกติ ซึ่ง $\exp(\sigma^2)$ เป็นScale Parameter และ เป็น Shape Parameter ค่าเฉลี่ยและความแปรปรวนของการแจกแจงแบบลอกนอร์มอล คือ $\exp(2\mu + \sigma^2)$ และ $\exp(2\mu + \sigma^2)[\exp(\sigma^2)-1]$ ตามลำดับ

สำหรับการสร้างเลขสุ่มที่มีการแจกแจงแบบลอกนอร์มอล หาได้โดยการหาค่าของ exponential ของฟังก์ชัน Normal(Mean,Sigma) เมื่อ Mean และ Sigma เป็นค่าเฉลี่ยและความแปรปรวนของการแจกแจงปกติ ดังแสดงในภาคผนวก สำหรับการวิจัยครั้งนี้ กำหนดให้ค่า $\mu = 0$ และ $\sigma^2 = 0.7$ และ 0.9

3.2.1.5 <u>การผลิตเลขสุ่มที่มีการแจกแจงแบบไคสแควร์</u>

กำหนดให้ Z_i มีการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 สามารถพิสูจน์ได้ว่า

$$\chi^2 = \Sigma Z_i^2$$

มีการแจกแจงไคสแควร์ซึ่งมีองศาความเป็นอิสระ n การผลิตเลขสุ่มที่มีการแจกแจงแบบไคสแควร์สามารถสร้างได้ด้วย วิธีผลประสาน(Convolution)โดยการสร้างตัวแปรสุ่ม Z_i ($i = 1,2,\ldots,n$) ให้มีการแจกแจงแบบปกติมาตรฐานแล้วแปลง Z_i ให้เป็น Z_i^2 ซึ่งเป็นตัวแปรสุ่มแบบไคสแควร์

โปรแกรมย่อยที่ใช้ในการสร้างเลขสุ่มที่มี การแจกแจงแบบไคสแควร์ คือ Function CSD(NDF,Mean,Sigma) ดังแสดงในภาคผนวก สำหรับการวิจัยครั้งนี้กำหนดองศาความเป็นอิสระ n=2 และ 3 Mean = 0. และ Sigma = 1

3.2.2 <u>การคำนวณค่าตัวสถิติทดสอบ 3 วิธี</u>

เมื่อสุ่มตัวอย่างจากประชากร โดยใช้โปรแกรมย่อยที่เขียนในภาคผนวก ตามขนาดตัวอย่าง และเปอร์เซนต์ของข้อมูลขาดหายที่กำหนดในแผนการทดลองแล้วนำข้อมูลที่ได้ไปคำนวณค่าต่าง ๆ ตามสูตรของตัวสถิติทดสอบแต่ละวิธีที่เสนอไว้ในบทที่ 2 เมื่อได้ค่าของตัวสถิติทดสอบแต่ละตัวแล้วนำมาเปรียบเทียบกับค่าวิกฤตที่เปิดจากตารางของ Shapiro and Wilk ตัวสถิติทดสอบ Z เปรียบเทียบกับค่าวิกฤต ที่เปิดจากตารางการแจกแจงปกติมาตรฐาน ตัวสถิติ BF เปรียบเทียบกับค่าวิกฤตที่เปิดจากตารางเอฟ ซึ่งการยอมรับหรือการปฏิเสธสมมุติฐานว่างนั้นให้ถือเกณฑ์ในบทที่ 2

3.2.3 <u>การหาความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบ</u>

เมื่อทำการสุ่มตัวอย่างและคำนวณค่าตัวสถิติทดสอบพร้อมทั้งเปรียบเทียบกับค่าวิกฤตซ้ำๆ กันเป็นจำนวน 500 รอบ ในแต่ละสถานการณ์แล้วนับจำนวนครั้งของการปฏิเสธสมมุติฐานว่างในกรณีที่การแจกแจงเป็นแบบเอกซโปเนนเชียล ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 หาได้ โดยเอาจำนวนครั้งของการปฏิเสธสมมุติฐานว่างหารด้วยจำนวนรอบทั้งหมด ส่วนในกรณีการแจกแจงของประชากรเป็นแบบอื่น จะหาค่าอำนาจการทดสอบ โดยการคำนวณเช่นเดียวกันกับการหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

สถานการณ์สำหรับการวิจัยแบ่งออกได้ดังนี้

1.ระดับนัยสำคัญ 2 ระดับ คือ 0.05 และ 0.10

2.ขนาดตัวอย่างของประชากร 5 ระดับ คือ 10,20,30,50 และ 100

3.ตัวสถิติทดสอบ 3 วิธีคือ W,Z และ BF

4.ประชากร 9 ชุด

5.เปอร์เซนต์ของข้อมูลขาดหาย 10 สถานการณ์ คือ

-ข้อมูลขาดหายทางซ้าย 10%,20% และ 30% ของจำนวนข้อมูล

-ข้อมูลขาดหายทางขวา 10%,20% และ 30% ของจำนวนข้อมูล

-จำนวนข้อมูลขาดหายทางซ้ายและทางขวาเท่ากัน 0% และ 10%

-จำนวนข้อมูลขาดหายทางซ้ายและขวาไม่เท่ากัน

ขาดหายทาซ้าย 10% ขาดหายทางขวา 20%

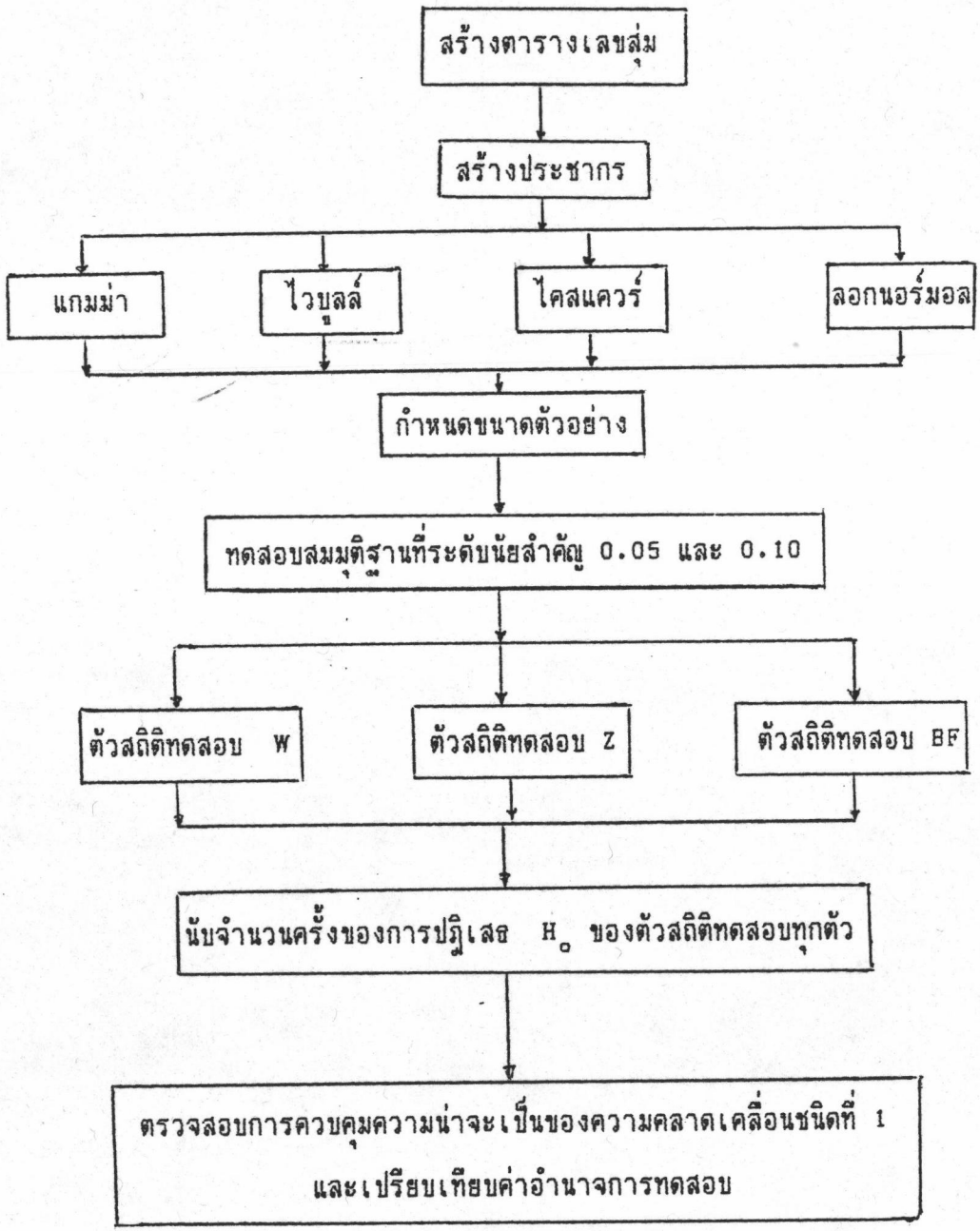
ขาดหายทาซ้าย 20% ขาดหายทางขวา 10%

ดังนั้นจำนวนสถานการณ์ที่ใช้ในการทดลองวิจัย $= 2 \times 5 \times 3 \times 9 \times 10$

$= 2700$ สถานการณ์

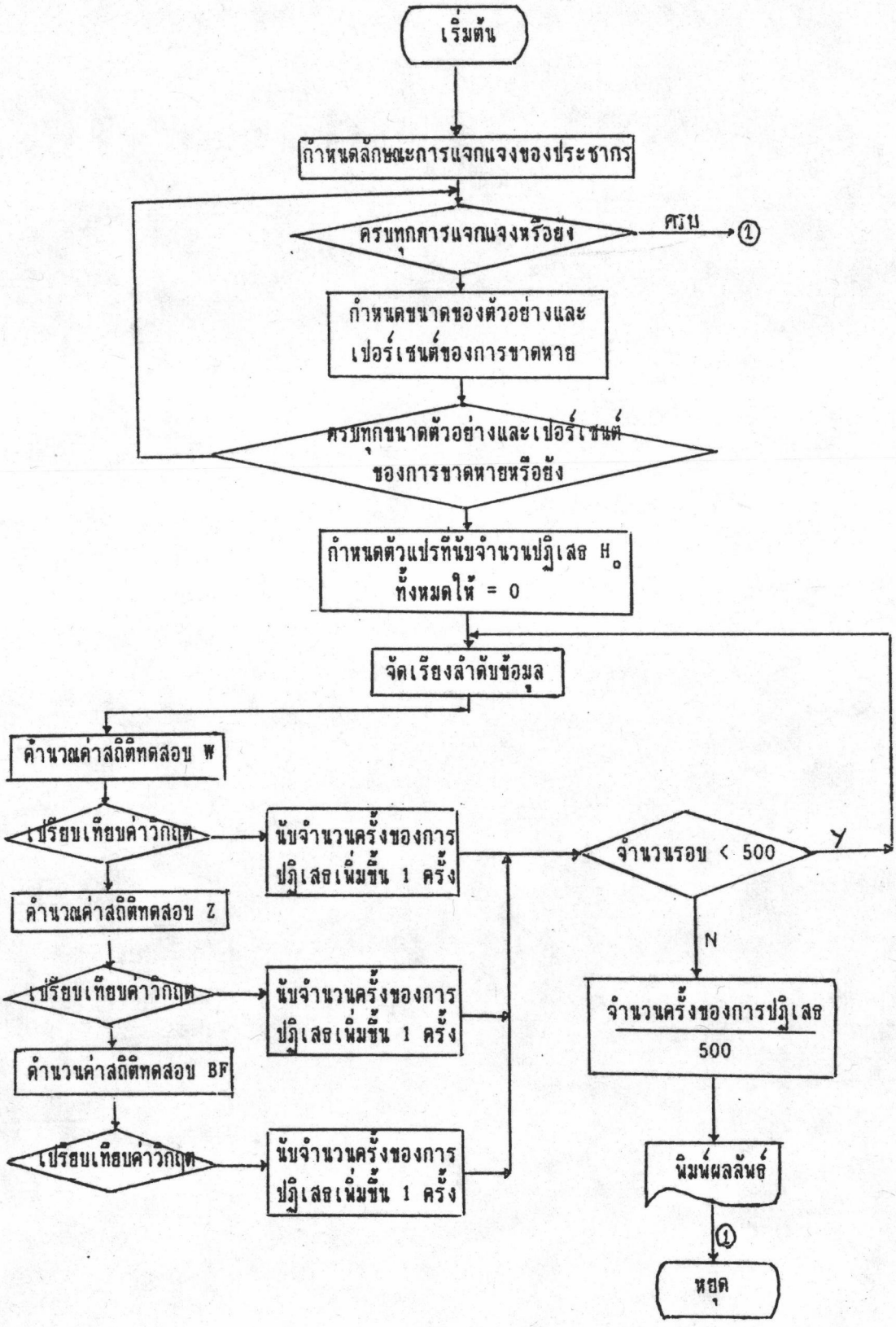
3.3 **ขั้นตอนในการทำงานของโปรแกรม**

โปรแกรมที่เขียนเขียนด้วยภาษาฟอร์แทรน ใช้ในการคำนวณหาค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และค่าอำนาจการทดสอบ ได้นำเสนอไว้ในภาคผนวก ส่วนขั้นตอนการทำงานของโปรแกรม แสดงได้ดังแผนต่อไปนี้



รูปที่ 3.1 แผนผังแสดงขั้นตอนการดำเนินการวิจัย

แผนผังโปรแกรม ที่ใช้ในการคำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และค่าอำนาจการทดสอบ



รายละเอียดของแผนผังโปรแกรม ที่ใช้ในการคำนวณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติที่ใช้ทดสอบ การแจกแจงเอกซโปเนนเชียลจะเสนอเป็นขั้นตอนดังนี้

1. สร้างประชากรให้มีลักษณะต่าง ๆ ดังนี้

การแจกแจงแบบแกมม่า โดยใช้คำสั่ง Function Gamma (Alpha,Beta)

การแจกแจงแบบไวบูลล์ โดยใช้คำสั่ง Function Weibul (Alpha,Beta)

การแจกแจงแบบลอกนอร์มอล โดยใช้คำสั่ง Function Exp(Mean,Sig)

การแจกแจงแบบไคสแควร์ โดยใช้คำสั่ง Function CSD(NDF,Mean,Sig)

การกำหนดขนาดตัวอย่างและการคำนวณค่าตัวสถิติทดสอบทั้ง 3 ตัว (W,Z,BF) สามารถทำได้โดย การเลือกประชากรครั้งละ 1 ประชากร และจะหยุดการทดลอง เมื่อทำการทดลองครบทุกกรณีที่กำหนดของแต่ละประชากร

2. กำหนดขนาดตัวอย่างและเปอร์เซนต์ข้อมูลขาดหาย

โดยที่ n คือ จำนวนขนาดตัวอย่าง

r1 คือ จำนวนข้อมูลขาดหายทางซ้าย

r2 คือ จำนวนข้อมูลขาดหายทางขวา

3.กำหนดตัวแปรที่นับจำนวนการปฏิเสธ H_0 ให้มีค่าเท่ากับ 0 โดยที่

W05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ ณ ระดับนัยสำคัญ 0.05

W10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ ณ ระดับนัยสำคัญ 0.10

Z05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ ณ ระดับนัยสำคัญ 0.05

Z10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ ณ ระดับนัยสำคัญ 0.10

BF05 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ ณ ระดับนัยสำคัญ 0.05

BF10 แทน ตัวแปรที่นับจำนวนการปฏิเสธ H_0 ของตัวสถิติทดสอบ ณ ระดับนัยสำคัญ 0.10

4. จัดเรียงข้อมูลซึ่งได้จากข้อ 1) โดยเรียงข้อมูลจากน้อยไปหามาก โดยใช้ Subroutine Rank(N,Y) พร้อมกับแปรงข้อมูลเป็น $T_i = (n-i+1)(y_i-y_{i-1})$

5. คำนวณตัวสถิติทดสอบ W

6. ทำการเปรียบเทียบค่าสถิติทดสอบที่คำนวณได้ในข้อ 5) กับค่าวิกฤติที่ได้จากตารางถ้าผลการเปรียบเทียบปรากฏว่า ปฏิเสธสมมุติฐาน H_0 ให้นับจำนวนครั้งที่ปฏิเสธสมมุติฐานเพิ่มขึ้น 1 ครั้ง ถ้าการเปรียบเทียบปรากฏว่ายอมรับสมมุติฐาน H_0 ให้ทำข้อ 5) ต่อไป แต่เปลี่ยนตัวสถิติทดสอบเป็น Z และ BF ตามลำดับ

7. ทำการทดลองซ้ำ ตั้งแต่ขั้นตอนที่ 4 จนถึงขั้นตอนที่ 6 จำนวน 500 ครั้ง ในแต่ละสถานาการณ์ เพื่อหาผลลัพธ์ของค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และคำนวณค่าอำนาจการทดสอบของตัวสถิติทดสอบ W,Z และ BF ตามลำดับ