บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงทดลองเพื่อหาข้อสรุปในการเปรียบเทียบวิธีการหาค่าประมาณแบบช่วงของ β_0 และ β_1 ในสมการถดถอยเชิงเส้นเชิงเดียว โดยจะทำการเปรียบเทียบ 3 วิธีคือ วิธีแบบฉบับ วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักและปรับให้เหมาะสม และวิธีบูตสแตรป โดยจะจำลองข้อมูลด้วยเทคนิคมอนติคาร์โลเพื่อกำหนดรูปแบบและปัจจัยที่คาดว่าจะมีผลต่อการศึกษาได้ตามต้องการ ตามรายละเอียดดังต่อไปนี้

3.1 แผนการทดลอง

งานวิจัยนี้ได้กำหนดสถานการณ์ที่จะใช้ทำการศึกษาดังนี้

1. ขนาดตัวอย่างที่จะใช้ทำการศึกษาเป็น 10 , 20 , 30 , 40 , 50 และ 60
2. ระดับความเชื่อมั่นที่ใช้คำนวณช่วงประมาณคือ 90% , 95% และ 99%
3. ค่าพารามิเตอร์ β_0 กับ β_1 ที่ใช้จำลองข้อมูลแบ่งเป็นสองกรณีดังนี้
   - กรณีแรก β_0 = 1 และ β_1 = 2
   - กรณีที่สอง β_0 = -1 และ β_1 = -3
4. ค่าพารามิเตอร์และสัมประสิทธิ์ความโด่งของคลาดเคลื่อนของสมการถดถอยจะกำหนดจากสัมประสิทธิ์ความเบ้ดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สัมประสิทธิ์ความเบ้	การแจกแจง	สัมประสิทธิ์ความโด่ง	พารามิเตอร์
0	ปกติ	3	$\mu = 0$, $\sigma^2 = 1$
0.25	แลมดาของตูกีร์	2	λ_1 = -1.465 , λ_2 = 0.2748 λ_3 = 0.0105 , λ_4 = 0.7034
		3.2	λ_1 = -0.237 , λ_2 = 0.1619 λ_3 = 0.0831 , λ_4 = 0.1300
		6.2	λ_1 = -0.064 , λ_2 = 0.1753 λ_3 = -0.0781 , λ_4 = -0.0876
	แกมมา	3.09	$\lambda = 1$, $\gamma = 64$
	ลอกนอร์มอล	3.11	$\mu = 0$, $\sigma = 0.0830$
0.5	แกมมา	3.38	$\lambda = 1$, $\gamma = 16$
	ลอกนอร์มอล	3.45	$\mu = 0$, $\sigma = 0.1641$
1	แลมดาของตูกีร์	4.2	λ_1 = -0.787 , λ_2 = 0.1142 λ_3 = 0.0212 , λ_4 = 0.1244
		5.4	λ_1 = -0.445 , λ_2 = -0.0317 λ_3 = -0.0101 , λ_4 = -0.0242
		8.4	λ_1 = -0.233 , λ_2 = -0.2079 λ_3 = -0.0744 , λ_4 = -0.1141
	แกมมา	4.5	$\lambda = 1$, $\gamma = 4$
	ลอกนอร์มอล	4.83	$\mu = 0$, $\sigma = 0.3143$
1.5	แกมมา	6.37	$\lambda = 1$, $\gamma = 1.778$
	ลอกนอร์มอล	7.25	$\mu = 0$, $\sigma = 0.4435$
2	แลมดาของตูกีร์	11.4	λ_1 = -0.629 , λ_2 = -0.1181 λ_3 = -0.0202 , λ_4 = -0.0867
		12.6	λ_1 = -0.539 , λ_2 = -0.1644 λ_3 = -0.0342 , λ_4 = -0.1104
		15.6	λ_1 = -0.409 , λ_2 = -0.2532 λ_3 = -0.646 , λ_4 = -0.1472
	แกมมา	9	$\lambda = 1$, $\gamma = 1$
	ลอกนอร์มอล	10.86	$\mu = 0$, $\sigma = 0.5514$
2.5	แกมมา	12.38	$\lambda = 1$, $\gamma = 0.64$
	ลอกนอร์มอล	15.86	$\mu = 0$, $\sigma = 0.6409$

ตารางที่ 3.1 แสดงการแจกแจงที่ใช้ในการวิจัย ค่าพารามิเตอร์ ค่าสปส.ความเบ้ และค่าสปส.ความโด่ง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.2 การสร้างรูปแบบการแจกแจงของค่าคลาดเคลื่อนตามที่กำหนด

การผลิตตัวแปรสุ่มให้มีการแจกแจงตามที่กำหนด ได้ทำการจำลองด้วยการใช้เทคนิคมอนติคาร์โล โดยการเขียนโปรแกรมคอมพิวเตอร์ภาษาวิชวลเบสิก (Visual Basic)

1. <u>การผลิตเลขสุ่มที่มีการแจกแจงเอกรูปในช่วง (0 , 1)</u>

ภาษาวิชวลเบสิกมีฟังก์ชันที่ใช้ในการผลิตเลขสุ่มอยู่แล้วคือ Rnd และมีคำสั่ง Randomize ซึ่งใช้สำหรับสร้างตารางเลขสุ่มใหม่

2. <u>การผลิตเลขสุ่มให้มีการแจกแจงแบบแลมดาของตูกีร์</u>

เริ่มต้นโดยกำหนดค่าความเบ้และความโด่งตามที่ต้องการ แล้วจึงหาค่า λ_1 , λ_2 , λ_3 และ λ_4 จากตาราง Ramberg มาใส่ในสมการต่อไปนี้

$$X = f^{-1}(X) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2} \quad , 0 \le p \le 1$$

ซึ่งสามารถเขียนโปรแกรมย่อย Gen_Lam ได้ดังนี้

```
Private Sub Gen_Lam()
        rd1 = Rnd
        err(i) = lam1 + ( rd1^lam3 - ( 1 - rd1 )^lam4 ) / lam2
End Sub
```

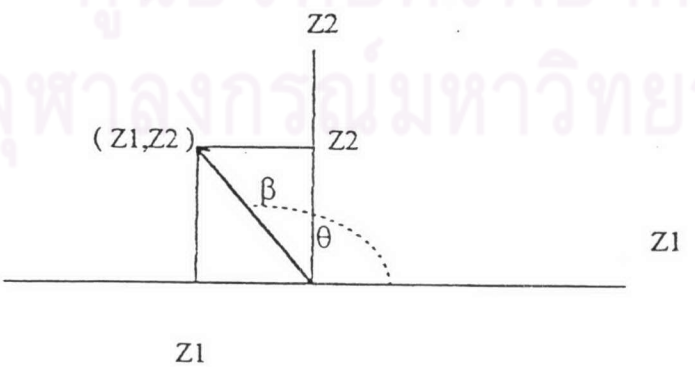
โดย rd1 เป็นตัวแปรที่ใช้เก็บค่าของตัวเลขสุ่มที่สร้างขึ้น

lam1 , lam2 , lam3 , lam4 คือค่าพารามิเตอร์ที่กำหนดตามตารางของ Ramberg

err(i) คืออาเรย์ (Array) ของค่าคลาดเคลื่อน (สำหรับสถานการณ์ที่ค่าคลาดเคลื่อนแจกแจงแบบแลมดาของตูกีร์)

3. <u>การผลิตเลขสุ่มให้มีการแจกแจงแบบปกติ</u>

สามารถทำได้โดยใช้วิธีของ Box และ Muller (1958) ซึ่งผลิตเลขสุ่มที่มีการแจกแจงปกติมาตรฐานที่มีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 1 พร้อมกัน 2 ค่าคือ Z_1 และ Z_2 ดังรูป



ภาพประกอบที่ 3.1 แนวคิดการสร้างเลขสุ่มให้มีการแจกแจงแบบปกติ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากรูปจะได้ว่า $Z_1 = \beta \cos(\theta)$

$Z_2 = \beta \sin(\theta)$

จากสมการ สามารถพิสูจน์ได้ว่า β และ θ เป็นอิสระกัน และ β^2 ซึ่งเท่ากับ $Z_1^2 + Z_2^2$ มีการแจกแจงแบบเลขชี้กำลัง (Exponential Distribution) ที่มีค่าเฉลี่ยเท่ากับ 2 ส่วน θ มีการแจกแจงเอกรูปในช่วง 0 ถึง 2π เรเดียน เมื่อใช้วิธีการแปลงผกผัน (Inverse Transformation) จะได้สมการต่อไปนี้

$$\beta = [-2\ln(RD_1)]^{1/2}$$

$$\theta = 2\pi (RD_2)$$

เมื่อ RD_1 และ RD_2 เป็นเลขสุ่มอิสระจากการแจกแจงเอกรูปในช่วง (0 , 1)

จากนี้เราสามารถสร้างเลขสุ่มที่มีการแจกแจงปกติมาตรฐานได้ 2 ชุดคือ

$$Z_1 = [-2\ln(RD_1)]^{1/2}\cos(2\pi RD_2)$$

$$Z_2 = [-2\ln(RD_1)]^{1/2}\sin(2\pi RD_2)$$

เมื่อได้ตัวแปรสุ่มที่มีการแจกแจงปกติมาตรฐานแล้ว เราสามารถแปลงให้มีค่าเฉลี่ยเป็น μ และความแปรปรวนเป็น σ^2 โดยอาศัยสมการ

$$X = \mu + \sigma Z$$

โปรแกรมย่อยฟังก์ชัน Norm สำหรับผลิตตัวแปรสุ่มที่มีการแจกแจงปกติ [$N(\mu , \sigma)$] มีดังนี้

```
Function Norm(a As Single , b As Single) As Single
        rd1 = Rnd
        Z = Sqr( -2 * Log( rd1 ) ) * cos( 2 * pi * rd1 )
        Norm = Z * b + a
End Function
```

โดย rd1 คือตัวแปรที่ใช้เก็บค่าเลขสุ่มที่สร้างขึ้น

Z คือตัวแปรสุ่มแบบปกติมาตรฐาน

Norm คือตัวแปรสุ่มแบบปกติ (μ , σ^2) ที่ผลิตได้

a คือค่าเฉลี่ย (ซึ่งรับค่ามาจากโปรแกรมที่เรียกใช้)

b คือค่าเบี่ยงเบนมาตรฐาน (ซึ่งรับค่ามาจากโปรแกรมที่เรียกใช้)

pi คือค่าพาย (π)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4. <u>การผลิตเลขสุ่มที่มีการแจกแจงแกมมา</u>

การผลิตเลขสุ่มที่มีการแจกแจงแกมมา ที่มีพารามิเตอร์ γ และ λ มีขั้นตอนดังนี้

1) คำนวณ $a = \sqrt{2\gamma - 1}$

$b = 2\gamma - 2\ln 2 + \frac{1}{a}$

2) ผลิตเลขสุ่ม RD_1 และ RD_2

3) คำนวณ $c = \gamma\left(\frac{RD_1}{1 - RD_1}\right)^a$

4) คำนวณ $d = b - \ln(RD_1^2 RD_2)$

5) ถ้า $c > d$ เราจะไม่ยอมรับค่า c และกลับไปทำขั้นที่ 2 แต่ถ้า $c \leq d$ จะยอมรับค่า c ซึ่งมีค่าเฉลี่ย และความแปรปรวนเท่ากับ γ

6) ถ้าต้องการค่า c ที่มีค่าเฉลี่ยเท่ากับ $\frac{\gamma}{\lambda}$ และความแปรปรวนเท่ากับ $\frac{\gamma}{\lambda^2}$ ให้แปลงค่าโดยนำค่า c ไปหารด้วย λ

โปรแกรมย่อย Gen_Gam สำหรับผลิตตัวแปรสุ่มที่มีการแจกแจงแบบแกมมามีดังนี้

```
Private Sub Gen_Gam()
        a = Sqr( 2 * gam_c – 1 )
        b = 2 * gam_c – 2 * Log( 2 ) + 1 / a
        Do
                rd1 = Rnd
                rd2 = Rnd
                c = gam_c * ( rd1 / ( 1 - rd1 ) )^a
                d = b - Log( ( rd1^2 ) * rd2 )
        Loop Until c <= d
        err(i) = c
End Sub
```

โดย gam_c คือพารามิเตอร์ γ ที่ป้อนเข้าไป

err(i) คืออาเรย์ (Array) ของค่าคลาดเคลื่อน (สำหรับสถานการณ์ที่ค่าคลาดเคลื่อนแจกแจงแบบแกมมา)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

5.<u>การผลิตเลขสุ่มที่มีการแจกแจงลอกนอร์มอล</u>

เนื่องจากการแจกแจงลอกนอร์มอลมีความสัมพันธ์กับการแจกแจงปกติ ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงปกติ เราจะได้ตัวแปรสุ่มที่มีการแจกแจงลอกนอร์มอลคือ y = exp(X) ดังนั้นการผลิตตัวแปรสุ่มที่มีการแจกแจงลอกนอร์มอลจึงมีขั้นตอนดังนี้

1) ผลิตตัวแปรสุ่มแบบปกติที่มีค่าเฉลี่ย μ และความแปรปรวนเท่ากับ σ^2
2) สร้างตัวแปรสุ่มที่มีการแจกแจงลอกนอร์มอลได้จากค่าชี้กำลังของตัวแปรสุ่มแบบปกติที่ได้จากข้อ 1

โปรแกรมย่อย Gen_Lnorm สำหรับผลิตตัวแปรสุ่มที่มีการแจกแจงแบบลอกนอร์มอลมีดังนี้

```
Private Sub Gen_Lnorm( )
        a = Norm( Logn_m , Logn_std )
        err(i) = Exp(a)
End Sub
```

โดย Logn_m และ Logn_std คือค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงปกติตามที่กำหนด

a คือตัวแปรที่มีการแจกแจงปกติโดยมีค่าเฉลี่ยคือ Logn_m และความแปรปรวนคือ $(Logn_std)^2$ ซึ่งได้จากการเรียกใช้โปรแกรมย่อยฟังก์ชัน Norm( )

err(i) คืออาเรย์ (Array) ของค่าคลาดเคลื่อน (สำหรับสถานการณ์ที่ค่าคลาดเคลื่อนแจกแจงแบบลอกนอร์มอล)

3.3 การสร้างตัวแปร (X , Y) ของสมการถดถอย

เมื่อจำลองค่าคลาดเคลื่อนได้แล้ว จากนั้นก็จะนำข้อมูลที่ได้ไปสร้างตัวแปร (X , Y) ของสมการถดถอย ตามโปรแกรมย่อยต่อไปนี้

```
For i = 1 To n
        x(i) = Norm( m , std )
        If  t_lam.value Then Call Gen_Lam
                ElseIf gam.value Then Call Gen_Gam
                Else Call Gen_Lnorm
        End if
        y(i) = b0 + b1 * x(i) + err(i)
Next i
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โดย m และ std คือค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของการแจกแจงปกติ

b0 และ b1 คือค่าพารามิเตอร์ β_0 และ β_1 ที่ใช้ในการจำลองข้อมูล ซึ่งได้กำหนดไว้แล้วในขอบเขตของการวิจัย

err(i) คืออาเรย์ (Array) ของค่าคลาดเคลื่อนที่มีการแจกแจงในรูปแบบต่างๆซึ่งได้จากการเรียกใช้โปรแกรมย่อยและฟังก์ชัน

x(i) คืออาเรย์ (Array) ของตัวแปรอิสระที่มีการแจกแจงปกติโดยมีค่าเฉลี่ยคือ m และความแปรปรวนคือ std^2 ซึ่งได้จากการเรียกใช้โปรแกรมย่อยฟังก์ชัน Norm( )

y(i) คืออาเรย์ (Array) ของตัวแปรตามที่สร้างจากสมการถดถอยตามเงื่อนไขที่กำหนด

ในแต่ละสถานการณ์จะทำการคำนวณซ้ำ 1,000 รอบ โดยใช้ $x_1, x_2, \ldots, x_n$ ชุดเดียวกันแต่จะเปลี่ยน $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ ไปในแต่ละรอบ และจะจำลอง $x_1, x_2, \ldots, x_n$ ใหม่เมื่อเปลี่ยนสถานการณ์

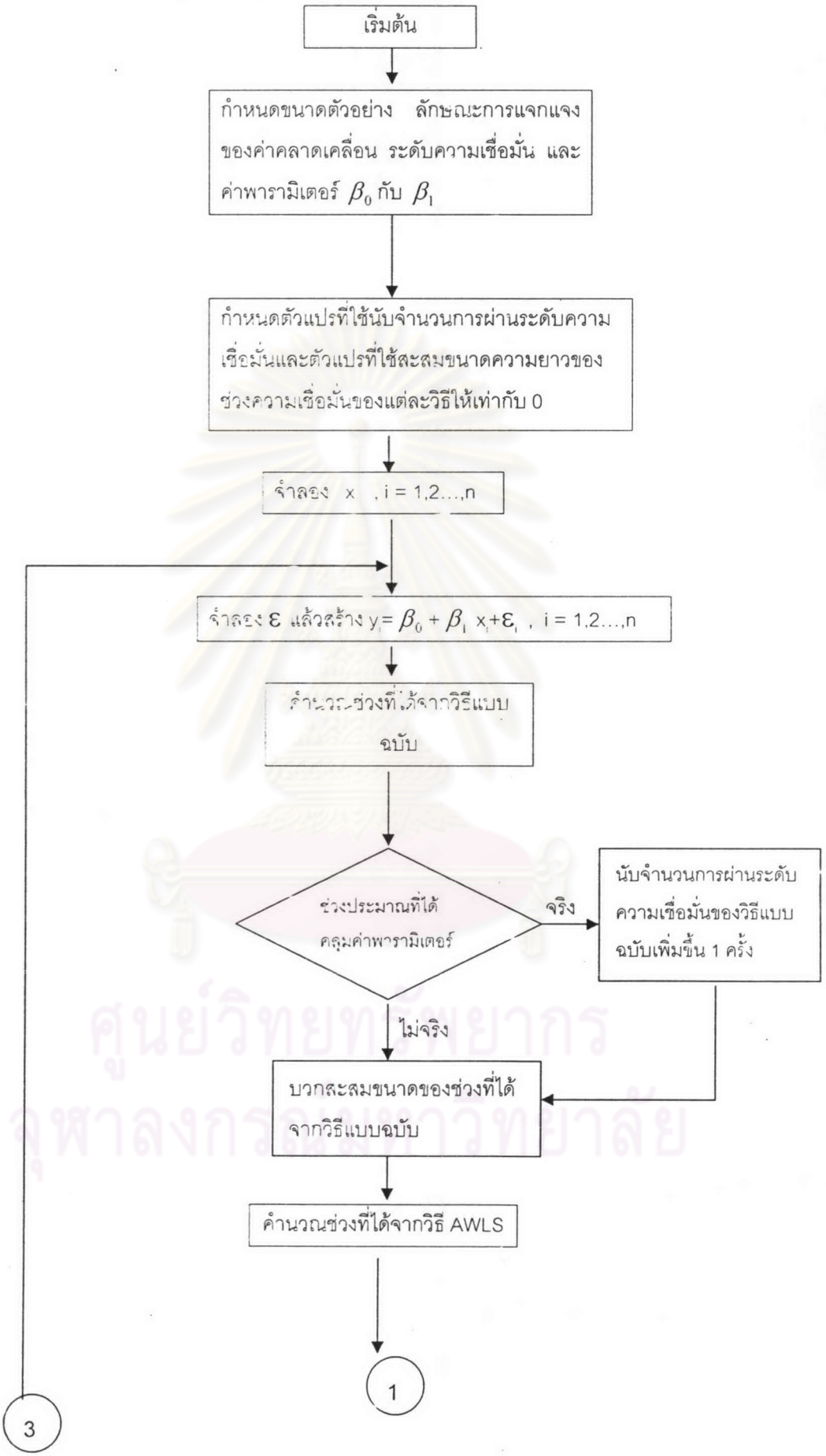
3.4 การคำนวณขนาดของช่วงความเชื่อมั่น

เมื่อได้ทำการจำลองข้อมูลตามสถานการณ์ที่กำหนดไว้ในแผนการทดลองแล้ว ก็จะนำข้อมูลที่ได้ไปคำนวณหาขนาดของช่วงความเชื่อมั่นของ β_0 และ β_1 ตามรายละเอียดที่กล่าวไว้ในบทที่ 2 และจะทำการคำนวณโดยโปรแกรมที่แสดงไว้ในภาคผนวก

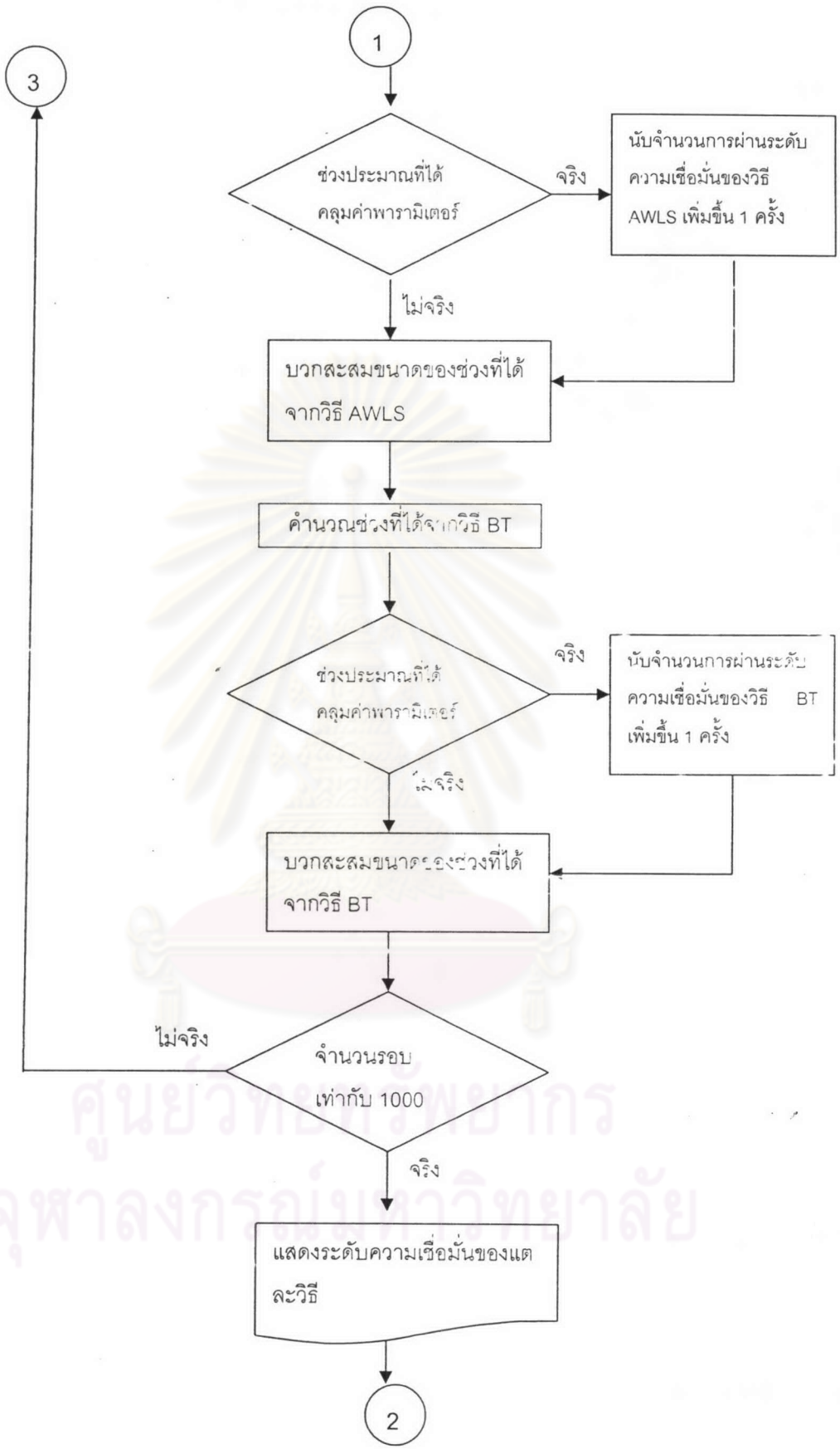
3.5 ขั้นตอนการทำงานของโปรแกรม

ในการวิจัยนี้ใช้โปรแกรมภาษาวิชวลเบสิกในการคำนวณ และจะทำการคำนวณทีละสถานการณ์ ขั้นตอนในการทำงานของโปรแกรมมีดังนี้

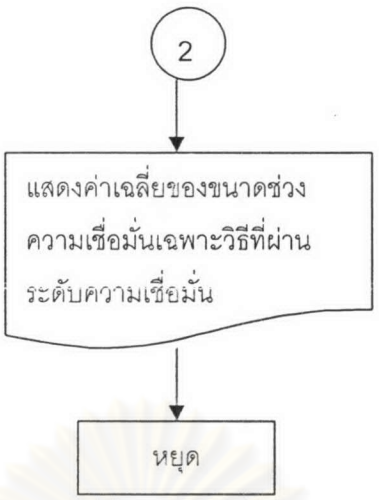
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาพประกอบที่ 3.2 แสดงขั้นตอนทั่วไปในการคำนวณขนาดของช่วงความเชื่อมั่น และทดสอบการผ่านระดับความเชื่อมั่น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย