บทที่ 3

วิธีดำเนินงานวิจัย

การวิจัยนี้เป็นการวิจัยเชิงทดลอง เพื่อหาตัวสถิติทดสอบที่มีประสิทธิภาพสูงที่สุดที่ใช้ในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติ กรณีที่ไม่ทราบพารามิเตอร์ประชากร โดยพิจารณาจากตัวสถิติทดสอบที่มีกำลังการทดสอบสูงสุด เมื่อตัวสถิติทดสอบนั้นสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ได้ ซึ่งในการวิจัยมีขั้นตอนดังนี้

1. การหาความน่าจะเป็นของความผิดพลาดประเภทที่ 1
2. การทดสอบความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1
3. การหากำลังการทดสอบของตัวสถิติทดสอบ
4. การเปรียบเทียบกำลังการทดสอบ

โดยกำหนดสมมติฐานในการทดสอบคือ

H_0 : ประชากรมีการแจกแจงแบบปกติ

H_1 : ประชากรไม่มีการแจกแจงแบบปกติ

แต่เพื่อให้เป็นบรรทัดฐานเดียวกันในการวิจัยครั้งนี้จึงทำการแปลงข้อมูลให้มีค่ามาตรฐาน คือมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ซึ่งการแปลงข้อมูลให้มีมีค่ามาตรฐานนั้นไม่ทำให้ค่าที่ได้จากตัวสถิติทดสอบเปลี่ยนแปลงไป ดังแสดงในภาคผนวก ข ดังนั้นสมมติฐานในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติคือ

H_0 : ประชากรมีการแจกแจงแบบปกติมาตรฐาน

H_1 : ประชากรไม่มีการแจกแจงแบบปกติมาตรฐาน

3.1 การหาความน่าจะเป็นของความผิดพลาดประเภทที่ 1

ตัวสถิติทดสอบที่นำมาศึกษาในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติเพื่อหาความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ประกอบด้วย

- ตัวสถิติทดสอบ Z_A (Zhang and Wu, 2005)
- ตัวสถิติทดสอบ Z_C (Zhang and Wu, 2005)
- ตัวสถิติทดสอบ Z_K (Zhang and Wu, 2005)
- ตัวสถิติทดสอบ Anderson-Darling (A^2) (Anderson and Darling, 1954)
- ตัวสถิติทดสอบ Shapiro-Wilk (W) (Shapiro and Wilk, 1965)
- ตัวสถิติทดสอบ Shapiro-Francia (W') (Shapiro and Francia, 1972)

ขั้นตอนในการหาความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 มีดังนี้

3.1.1 กำหนดสมมติฐาน

H_0 : ประชากรมีการแจกแจงแบบปกติมาตรฐาน

H_1 : ประชากรไม่มีการแจกแจงแบบปกติมาตรฐาน

3.1.2 กำหนดการแจกแจงเป็นแบบปกติมาตรฐาน (N(0,1)) ซึ่งมี ค่าเฉลี่ย เท่ากับ 0 ($\mu = 0$) และความแปรปรวน เท่ากับ 1 ($\sigma^2 = 1$)

3.1.3 กำหนดขนาดตัวอย่างในการศึกษาเท่ากับ 10, 50 และ100

3.1.4 กำหนดระดับนัยสำคัญในการทดสอบ คือ 0.05

3.1.5 ทำการจำลองข้อมูลให้มีการแจกแจงเป็นแบบปกติมาตรฐาน ที่มีขนาดตัวอย่างตามที่กำหนด โดยในการวิจัยนี้จะใช้โปรแกรมสำเร็จรูป Minitab 14 for Windows

3.1.6 ทำการคำนวณค่าสถิติของตัวสถิติทดสอบแต่ละตัว แล้วเปรียบเทียบกับค่าวิกฤติของตัวสถิติทดสอบนั้นตามเกณฑ์การตัดสินใจของตัวสถิติทดสอบเพื่อพิจารณาการยอมรับหรือปฏิเสธสมมติฐาน H_0 ตามที่กำหนดในข้อ 3.1.1

3.1.7 ทำซ้ำในแต่ละลักษณะการแจกแจง ขนาดตัวอย่าง และระดับนัยสำคัญ จำนวน 1000 ครั้ง แล้วนับจำนวนการปฏิเสธสมมติฐาน (H_0)

3.1.8 คำนวณหาความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จากอัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง (H_0) เมื่อสมมติว่าง (H_0) จริง เทียบกับจำนวนครั้งที่ทำการทดลอง (ในที่นี้คือ N=1000)

3.2 การทดสอบความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่1

ในการทดสอบว่าตัวสถิติทดสอบแต่ละตัวสามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้หรือไม่นั้นจะใช้การทดสอบทวินาม โดยกำหนดระดับนัยสำคัญในการทดสอบนี้คือ 0.05 ซึ่งมีสมมติฐานทดสอบคือ

H_{01} : $\alpha \leq \alpha_0$

H_{11} : $\alpha > \alpha_0$

สถิติทดสอบคือ

$$Z = \frac{\alpha_{\cdot} - \alpha_0}{\sqrt{\frac{\alpha_0(1-\alpha_0)}{(N=1000)}}}$$

กำหนด α แทน ความน่าจะเป็นของความผิดพลาดประเภทที่ 1

$\alpha_{\cdot}$ แทน ความน่าจะเป็นของความผิดพลาดประเภทที่ 1 จากการทดลอง (คืออัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่างเมื่อสมมติว่างจริง เทียบกับจำนวนครั้งที่ทำการทดลอง)

α_0 แทน ระดับนัยสำคัญที่กำหนดในการศึกษา คือ 0.05

N แทน จำนวนครั้งที่ทำการทดลองเท่ากับ 1,000 ครั้ง

ดังนั้นเกณฑ์การพิจารณาความสามารถในการควบคุมความผิดพลาดประเภทที่ 1 จากที่กำหนดระดับนัยสำคัญในการทดสอบทวินามเท่ากับ 0.05 คือตัวสถิติทดสอบจะสามารถควบคุมความผิดพลาดประเภทที่ 1 ได้ ถ้า $Z \leq 1.645$ ดังนั้นในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (α_0) เท่ากับ 0.05 จะได้

$$Z \leq Z_{0.95}$$

$$\frac{\alpha_{\cdot} - 0.05}{\sqrt{\frac{0.05(1-0.05)}{1000}}} \leq 1.645$$

$$\alpha_{\cdot} \leq 0.061$$

ดังนั้นตัวสถิติทดสอบสามารถควบคุมความผิดพลาดประเภทที่ 1 ที่ $\alpha_0 = 0.05$ ได้ ถ้า $\alpha_{\cdot}$ มีค่าไม่เกิน 0.061

3.3 การหากำลังการทดสอบของตัวสถิติทดสอบ

การหากำลังการทดสอบของตัวสถิติทดสอบที่นำมาศึกษาในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติทั้ง 6 ตัว มีขั้นตอนในการหาดังนี้

3.3.1 กำหนดสมมติฐาน

H_0 : ประชากรมีการแจกแจงแบบปกติมาตรฐาน

H_1 : ประชากรไม่มีการแจกแจงแบบปกติมาตรฐาน

3.3.2 กำหนดการแจกแจงของประชากรที่ไม่เป็นการแจกแจงแบบปกติ คือ จะกำหนดให้ประชากรมีการแจกแจงแบบที, การแจกแจงแบบล็อกนอร์มัล, การแจกแจงแบบบีตา และการแจกแจงแบบไวบูลล์ โดยจะใช้สัมประสิทธิ์ความเบ้ (Coefficient of Skewness : γ_1) และสัมประสิทธิ์ความโด่ง (Coefficient of Kurtosis : γ_2) เป็นตัวกำหนดพารามิเตอร์ ดังนี้

ประชากรมีการแจกแจงแบบที

เมื่อประชากรมีการแจกแจงแบบที จะกำหนดค่าพารามิเตอร์ดังตาราง 3.1

ตารางที่ 3.1

แสดงค่าพารามิเตอร์, สัมประสิทธิ์ความเบ้ และสัมประสิทธิ์ความโด่ง

เมื่อประชากรมีการแจกแจงแบบที

พารามิเตอร์รูปร่าง (ν)	ความเบ้ (γ_1)	ความโด่ง (γ_2)
5	0	9.00
6	0	6.00
7	0	5.00
8	0	4.50
10	0	4.00
14	0	3.60
34	0	3.20

ประชากรมีการแจกแจงแบบล็อกนอร์มัล

เมื่อประชากรมีการแจกแจงแบบล็อกนอร์มัล จะกำหนดค่าพารามิเตอร์ดังตาราง 3.2

ตารางที่ 3.2

แสดงค่าพารามิเตอร์,สัมประสิทธิ์ความเบ้ และสัมประสิทธิ์ความโด่ง

เมื่อประชากรมีการแจกแจงแบบล็อกนอร์มัล

พารามิเตอร์		ความเบ้	ความโด่ง
μ	σ^2	(γ_1)	(γ_2)
0	0.0269	0.5	3.448
0	0.0988	1.0	4.830
0	0.1967	1.5	7.251
0	0.3040	2.0	10.864
0	0.4108	2.5	15.850

ประชากรมีการแจกแจงแบบบีตา

เมื่อประชากรมีการแจกแจงแบบบีตา จะกำหนดค่าพารามิเตอร์ดังตาราง 3.3

ตารางที่ 3.3

แสดงค่าพารามิเตอร์, สัมประสิทธิ์ความเบ้ และสัมประสิทธิ์ความโด่ง

เมื่อประชากรมีการแจกแจงแบบบีตา

พารามิเตอร์		ความเบ้ (γ_1)	ความโด่ง (γ_2)
รูปร่าง (α)	รูปร่าง (β)		
1.00	1.00	0	1.800
1.50	1.50	0	2.000
2.25	2.25	0	2.200
3.50	3.50	0	2.400
13.50	13.50	0	2.800
7.00	2.00	-0.768	3.312
5.00	1.00	-1.183	4.200
2.00	1.08	-0.498	2.324
2.28	5.00	0.500	2.754
2.00	1.00	-0.566	2.400
2.00	5.00	0.596	2.880
0.50	1.00	0.639	2.143
2.00	4.00	0.468	2.625

ประชากรมีการแจกแจงแบบไวบูลล์

เมื่อประชากรมีการแจกแจงแบบไวบูลล์ จะกำหนดค่าพารามิเตอร์ดังตาราง 3.4

ตารางที่ 3.4

แสดงค่าพารามิเตอร์, สัมประสิทธิ์ความเบ้ และสัมประสิทธิ์ความโด่ง

เมื่อประชากรมีการแจกแจงแบบไวบูลล์

พารามิเตอร์		ความเบ้	ความโด่ง
รูปร่าง (α)	รูปร่าง (β)	(γ_1)	(γ_2)
2.211	1.00	0.50	3.03
1.563	1.00	1.00	4.16
1.211	1.00	1.50	6.13
1.000	1.00	2.00	9.00
0.896	1.00	2.50	12.83

ในการวิจัยนี้จะอาศัยหลักเกณฑ์ของ ชาพิโร และคณะ (Shapiro, Wilk and Chen, 1968) ในการแบ่งลักษณะความเบ้และความโด่งเป็น 5 กรณี โดยกำหนด

- กรณีที่ 1 : $\gamma_1 = 0$, $2.5 \leq \gamma_2 \leq 4.5$
  จะเรียกว่าการแจกแจงแบบใกล้เคียงปกติ (Near Normal)
- กรณีที่ 2 : $\gamma_1 = 0$, $\gamma_2 > 4.5$
  จะเรียกว่าการแจกแจงสมมาตรและหางยาว (Symmetric Long-tailed)
- กรณีที่ 3 : $\gamma_1 = 0$, $\gamma_2 < 2.5$
  จะเรียกว่าการแจกแจงสมมาตรและหางสั้น (Symmetric Short-tailed)
- กรณีที่ 4 : $|\gamma_1| > 0.3$, $\gamma_2 > 3.0$
  จะเรียกว่าการแจกแจงไม่สมมาตรและหางยาว (Asymmetric Long-tailed)
- กรณีที่ 5 : $|\gamma_1| > 0.3$, $\gamma_2 < 3.0$
  จะเรียกว่าการแจกแจงไม่สมมาตรและหางสั้น (Asymmetric Short-tailed)

ตารางที่ 3.5

แสดงลักษณะการแจกแจงที่ไม่เป็นแบบปกติ

ลักษณะที่	ลักษณะการแจกแจง	การแจกแจง	ความเบ้	ความโด่ง
1	การแจกแจงใกล้เคียงการแจกแจงแบบปกติ	Be(13.5,13.5)	0	2.8
2	การแจกแจงใกล้เคียงการแจกแจงแบบปกติ	t(34)	0	3.2
3	การแจกแจงใกล้เคียงการแจกแจงแบบปกติ	t(14)	0	3.6
4	การแจกแจงใกล้เคียงการแจกแจงแบบปกติ	t(10)	0	4.0
5	การแจกแจงใกล้เคียงการแจกแจงแบบปกติ	t(8)	0	4.5
6	การแจกแจงสมมาตรและหางยาว	t(7)	0	5.0
7	การแจกแจงสมมาตรและหางยาว	t(6)	0	6.0
8	การแจกแจงสมมาตรและหางยาว	t(5)	0	9.0
9	การแจกแจงสมมาตรและหางสั้น	Be(1,1)	0	1.8
10	การแจกแจงสมมาตรและหางสั้น	Be(1.5,1.5)	0	2.0
11	การแจกแจงสมมาตรและหางสั้น	Be(2.25,2.25)	0	2.2
12	การแจกแจงสมมาตรและหางสั้น	Be(3.5,3.5)	0	2.4
13	การแจกแจงไม่สมมาตรและหางยาว	W(2.211,1)	0.50	3.030
14	การแจกแจงไม่สมมาตรและหางยาว	LN(0,0.1641)	0.50	3.448
15	การแจกแจงไม่สมมาตรและหางยาว	W(1.563,1)	1.00	4.160
16	การแจกแจงไม่สมมาตรและหางยาว	LN(0,0.3143)	1.00	4.830
17	การแจกแจงไม่สมมาตรและหางยาว	W(1.211,1)	1.50	6.130
18	การแจกแจงไม่สมมาตรและหางยาว	LN(0,0.4435)	1.50	7.251
19	การแจกแจงไม่สมมาตรและหางยาว	W(1,1)	2.00	9.000
20	การแจกแจงไม่สมมาตรและหางยาว	LN(0,0.5514)	2.00	10.864
21	การแจกแจงไม่สมมาตรและหางยาว	W(0.896,1)	2.50	12.830
22	การแจกแจงไม่สมมาตรและหางยาว	LN(0,0.6409)	2.50	15.850
23	การแจกแจงไม่สมมาตรและหางยาว	Be(7,2)	-0.768	3.312
24	การแจกแจงไม่สมมาตรและหางยาว	Be(5,1)	-1.183	4.200

ตารางที่ 3.5 (ต่อ)

แสดงลักษณะการแจกแจงที่ไม่เป็นแบบปกติ

ลักษณะที่	ลักษณะการแจกแจง	การแจกแจง	ความเบ้	ความโด่ง
25	การแจกแจงไม่สมมาตรและหางสั้น	Be(2,1.08)	-0.498	2.324
26	การแจกแจงไม่สมมาตรและหางสั้น	Be(2.28,5)	0.500	2.754
27	การแจกแจงไม่สมมาตรและหางสั้น	Be(2,1)	-0.566	2.400
28	การแจกแจงไม่สมมาตรและหางสั้น	Be(2,5)	0.596	2.880
29	การแจกแจงไม่สมมาตรและหางสั้น	Be(0.5,1)	0.639	2.143
30	การแจกแจงไม่สมมาตรและหางสั้น	Be(2,4)	0.468	2.625

ดังนั้นสามารถสรุปลักษณะการแจกแจงได้ดังตาราง 3.6

ตารางที่ 3.6

แสดงจำนวนลักษณะการแจกแจง

กรณี	ลักษณะการแจกแจง	จำนวนลักษณะการแจกแจง
1	การแจกแจงใกล้เคียงการแจกแจงแบบปกติ	5
2	การแจกแจงสมมาตรและหางยาว	3
3	การแจกแจงสมมาตรและหางสั้น	4
4	การแจกแจงไม่สมมาตรและหางยาว	12
5	การแจกแจงไม่สมมาตรและหางสั้น	6

3.3.3 กำหนดขนาดตัวอย่างที่ศึกษาเท่ากับ 10, 50 และ 100

3.3.4 กำหนดระดับนัยสำคัญในการทดสอบ คือ 0.05

3.3.5 ทำการจำลองข้อมูลตามการแจกแจง และขนาดตัวอย่างต่างๆตามที่กำหนด โดยในการวิจัยนี้จะใช้โปรแกรมสำเร็จรูป Minitab 14 for Windows

3.3.6 แปลงข้อมูลที่ได้จาก 3.3.5 ให้มีค่ามาตรฐาน คือค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 โดยแปลงข้อมูล X_i ให้อยู่ในรูป X_i' คือ

$$X_i' = \frac{X_i - \mu}{\sigma}$$

เพื่อให้ X_i' มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ซึ่งโดยปกตินั้นในการดำเนินงานวิจัยมักจะไม่ทราบค่าพารามิเตอร์ จึงต้องทำการประมาณค่าพารามิเตอร์และพารามิเตอร์ในที่นี้คือค่าเฉลี่ยประชากร (μ) และความแปรปรวนประชากร (σ^2) โดยการใช้คุณสมบัติของตัวประมาณไม่เอนเอียงและความแปรปรวนต่ำสุด (Minimum Variance Unbiased Estimator) นั้น เป็นที่ทราบดีว่า $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ และ $S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$ เป็นตัวประมาณพารามิเตอร์ที่ดีที่สุด ของค่าเฉลี่ยประชากรและความแปรปรวนประชากรตามลำดับ โดยจะเรียก $\bar{X}$ ว่าค่าเฉลี่ยตัวอย่าง และเรียก S^2 ว่าความแปรปรวนตัวอย่าง

ดังนั้นจะได้

$$X_i' = \frac{X_i - \bar{X}}{S}$$

3.3.7 ทำการคำนวณค่าสถิติของตัวสถิติทดสอบแต่ละตัวคือ

- ตัวสถิติทดสอบ Z_A (Zhang and Wu, 2005)
- ตัวสถิติทดสอบ Z_C (Zhang and Wu, 2005)
- ตัวสถิติทดสอบ Z_K (Zhang and Wu, 2005)
- ตัวสถิติทดสอบ Anderson-Darling (A^2) (Anderson and Darling, 1954)
- ตัวสถิติทดสอบ Shapiro-Wilk (W) (Shapiro and Wilk, 1965)
- ตัวสถิติทดสอบ Shapiro-Francia (W') (Shapiro and Francia, 1972)

แล้วทำการเปรียบเทียบค่าสถิติที่ได้กับค่าวิกฤติของตัวสถิติทดสอบแต่ละตัว เพื่อพิจารณาการยอมรับหรือปฏิเสธสมมติฐาน H_0 ตามเกณฑ์การตัดสินใจของตัวสถิติทดสอบนั้นๆ กล่าวคือ

เกณฑ์การตัดสินใจของตัวสถิติทดสอบ Z_A จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่า $10Z_A$ - 32 ที่คำนวณได้มีค่ามากกว่าค่าวิกฤตที่ได้จากตาราง ก6 Percentage points of $10Z_A$ – 32 for testing normality (ภาคผนวก ก) ณ ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนด

เกณฑ์การตัดสินใจของตัวสถิติทดสอบ Z_C จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่า Z_C ที่คำนวณได้มีค่ามากกว่าค่าวิกฤตที่ได้จากตาราง ก7 Percentage points of Z_C for testing normality (ภาคผนวก ก) ณ ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนด

เกณฑ์การตัดสินใจของตัวสถิติทดสอบ Z_K คือจะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่า Z_K ที่คำนวณได้มีค่ามากกว่าค่าวิกฤตที่ได้จากตาราง ก8 Percentage points of Z_K for testing normality (ภาคผนวก ก) ณ ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนด

เกณฑ์การตัดสินใจของตัวสถิติทดสอบ A^2 จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่า A^2 ที่คำนวณได้มีค่ามากกว่าค่าวิกฤตที่ได้จากตาราง ก1 Percentage points of Anderson-Darling test (ภาคผนวก ก) ณ ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนด

เกณฑ์การตัดสินใจของตัวสถิติทดสอบ W จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่า W ที่คำนวณได้มีค่าน้อยกว่าค่าวิกฤตที่ได้จากตาราง ก3 Percentage points of Shapiro-Wilk test (ภาคผนวก ก) ณ ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนด

เกณฑ์การตัดสินใจของตัวสถิติทดสอบ W' จะปฏิเสธสมมติฐานว่าง H_0 เมื่อค่า W' ที่คำนวณได้มีค่าน้อยกว่าค่าวิกฤตที่ได้จากตาราง ก5 Percentage points of Shapiro-Francia test (ภาคผนวก ก) ณ ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนด

3.3.8 ทำซ้ำตามการแจกแจง ขนาดตัวอย่าง และระดับนัยสำคัญที่กำหนดแต่ละสถานการณ์จำนวน 1000 ครั้ง แล้วนับจำนวนการปฏิเสธสมมติฐาน (H_0)

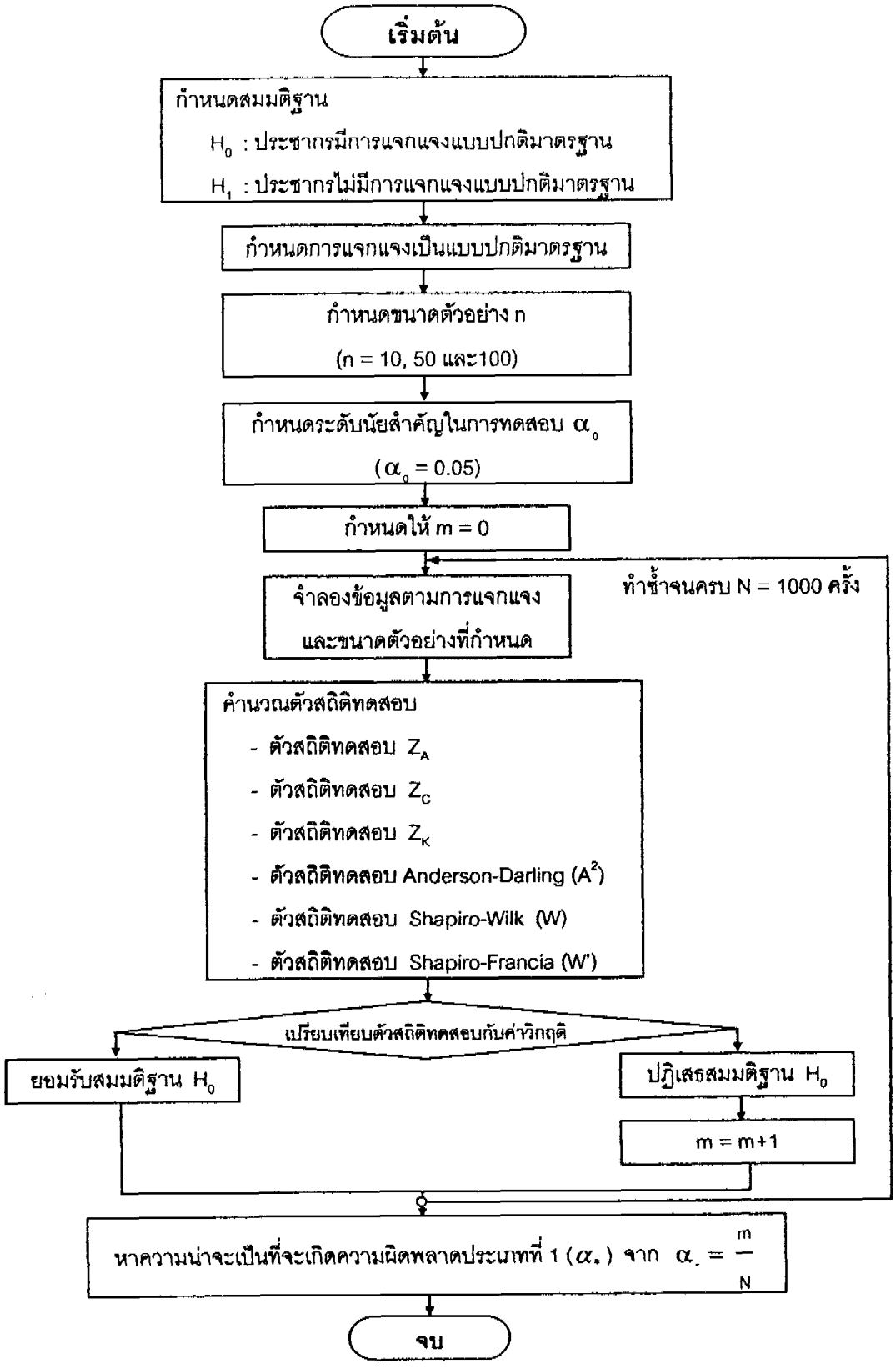
3.3.9 คำนวณหากำลังการทดสอบของตัวสถิติทดสอบแต่ละตัวจาก อัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง (H_0) เมื่อสมมติว่าง (H_0) จริง เทียบกับจำนวนครั้งที่ทำการทดลอง (ในที่นี้คือ N=1000)

3.4 การเปรียบเทียบกำลังการทดสอบ

ในการเปรียบเทียบกำลังการทดสอบเพื่อเลือกตัวสถิติทดสอบในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติที่ดีที่สุด จะพิจารณาเลือกจากตัวสถิติทดสอบที่ให้กำลังการทดสอบที่สูงที่สุดในแต่ละสถานการณ์

แผนภาพ 3.1

แสดงการหาความน่าจะเป็นของความผิดพลาดประเภทที่ 1



แผนภาพ 3.2

แสดงการหากำลังการทดสอบ