

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

งานวิจัยด้านต่างๆ โดยทั่วไปจำเป็นต้องนำวิธีการทางสถิติมาใช้ในการวิเคราะห์ข้อมูล โดยการหาข้อสรุปเกี่ยวกับลักษณะที่สำคัญของประชากรโดยใช้ข้อมูลจากตัวอย่าง ซึ่งเมื่อพิจารณาในแง่ของการวิเคราะห์และสรุปผล เราอาจแบ่งวิธีการทางสถิติออกเป็น 2 ประเภทตามลักษณะของข้อมูล คือ สถิติที่ใช้พารามิเตอร์ (*Parametric Statistics*) หรือสถิติที่ทราบรูปแบบการแจกแจงของข้อมูล และสถิติที่ไม่ใช้พารามิเตอร์ (*Nonparametric Statistics*) หรือสถิติที่ไม่ทราบรูปแบบการแจกแจงของข้อมูล

การนำสถิติที่ใช้พารามิเตอร์มาใช้ในการทดสอบสมมติฐานนั้น เราต้องทราบข้อสมมติเบื้องต้น (*Assumptions*) ของการทดสอบ เช่น มีข้อสมมติว่าประชากรที่ศึกษามีการแจกแจงแบบปกติ (*Normal Distribution*) ในทางปฏิบัติจะพบเสมอว่านักวิจัยมักจะละเลยหรือไม่ทดสอบว่าข้อมูลที่จะนำมาวิเคราะห์มีลักษณะการแจกแจงอย่างไร ซึ่งถือว่าเป็นผลเสีย เพราะถ้าข้อมูลนั้นมาจากประชากรที่มีการแจกแจงที่ไม่เป็นไปตามข้อสมมติ การนำวิธีการทางสถิติที่ใช้พารามิเตอร์มาวิเคราะห์ก็มักจะทำให้ผลที่ได้จากการวิเคราะห์มีความคลาดเคลื่อนสูง ดังนั้นก่อนการนำข้อมูลไปประมาณค่าหรือทดสอบสมมติฐาน เราควรตรวจสอบก่อนว่าข้อมูลที่จะนำมาวิเคราะห์นั้นมาจากประชากรที่มีการแจกแจงตามข้อสมมติหรือไม่

ข้อสมมติของการแจกแจงในสถิติที่ใช้พารามิเตอร์ที่พบบ่อยโดยทั่วไป คือ ข้อสมมติที่ว่าประชากรมีการแจกแจงแบบปกติ ฉะนั้นจึงได้มีผู้คิดค้นวิธีการทดสอบการแจกแจงแบบปกติขึ้นมาหลายวิธีด้วยกัน แม้ในปัจจุบันก็ยังมีวิธีการคิดค้นแบบทดสอบใหม่ๆ กันขึ้นมาอีก โดยมุ่งหวังให้มีอำนาจการทดสอบ (*power of test*) ที่ดีกว่า

การทดสอบว่าประชากรมีการแจกแจงแบบปกติหรือไม่ และการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกตินั้น มีนักสถิติหลายท่านได้เสนอตัวสถิติที่จะใช้ในการทดสอบการแจกแจงแบบปกติและทำการศึกษาเปรียบเทียบอำนาจการทดสอบ โดยจะขอกล่าวเฉพาะบางผลงานวิจัยที่เกี่ยวข้องดังนี้

Shapiro และ *Wilk* (1965) ได้เสนอตัวสถิติทดสอบ *Shapiro-Wilk (W)* ซึ่งเป็นตัวสถิติที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติ เมื่อขนาดตัวอย่างน้อยกว่าหรือเท่ากับ 50 โดยมีการนำค่าสัมประสิทธิ์เชิงเส้นที่เหมาะสมกับข้อมูลที่มีการเรียงลำดับ (*order sample*) มาใช้ในการคำนวณ

Shapiro และคณะ (1968) เป็นผู้ริเริ่มในการศึกษาอำนาจการทดสอบของตัวสถิติต่างๆ ในการทดสอบการแจกแจงแบบปกติ โดยทำการศึกษาตัวสถิติทดสอบ 9 ตัว คือ ตัวสถิติทดสอบ Shapiro-Wilk (W) ตัวสถิติทดสอบ Standard Third Moment ($\sqrt{b_1}$) ตัวสถิติทดสอบ Standard Fourth Moment (b_2) ตัวสถิติทดสอบ Kolmogorov-Smirnov (K) ตัวสถิติทดสอบ Cramer-von Mises (W^2) ตัวสถิติทดสอบ Anderson-Darling (A^2) ตัวสถิติทดสอบ Durbin ตัวสถิติทดสอบ Chi-square (χ^2) และตัวสถิติทดสอบ Studentized Range (u) ภายใต้การแจกแจง 12 การแจกแจง ซึ่งมีค่าพารามิเตอร์ที่แตกต่างกันรวมเป็น 45 การแจกแจง ได้ผลสรุปโดยย่อดังนี้

1. ตัวสถิติทดสอบ W ใช้ได้ดีในการทดสอบทั่วไป
2. ตัวสถิติทดสอบ u มีอำนาจการทดสอบสูงเมื่อประชากรมีการแจกแจงแบบสมมาตรทางสั้น (Symmetric Short-Tailed) และมีอำนาจการทดสอบต่ำเมื่อประชากรมีการแจกแจงแบบไม่สมมาตรทางสั้น (Asymmetric Short-Tailed) และไม่สมมาตรหางยาว (Asymmetric Long-Tailed)
3. ตัวสถิติทดสอบ $\sqrt{b_1}$ และตัวสถิติทดสอบ b_2 ใช้ในการทดสอบได้ดี แต่มีอำนาจการทดสอบต่ำกว่าตัวสถิติทดสอบ W

Filliben (1975) ได้เสนอตัวสถิติทดสอบ Probability Plot Correlation Coefficient หรือตัวสถิติทดสอบ Filliben (r) ซึ่งเป็นตัวสถิติที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติ โดยดัดแปลงมาจากตัวสถิติทดสอบ Shapiro-Wilk (W) และได้ทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกติ 8 ตัว คือ ตัวสถิติทดสอบ Standard Third Moment ($\sqrt{b_1}$) ตัวสถิติทดสอบ Standard Fourth Moment (b_2) ตัวสถิติทดสอบ a ตัวสถิติทดสอบ Studentized Range (u) ตัวสถิติทดสอบ Shapiro-Wilk (W) ตัวสถิติทดสอบ Shapiro-Francia (W') ตัวสถิติทดสอบ D'Agostino (D) และตัวสถิติทดสอบ Filliben (r) ภายใต้การแจกแจง 51 การแจกแจง เมื่อขนาดตัวอย่างเท่ากับ 20 และภายใต้การแจกแจง 25 การแจกแจง เมื่อขนาดตัวอย่างเท่ากับ 50 พบว่า ตัวสถิติทดสอบ r มีอำนาจการทดสอบสูง เมื่อประชากรมีการแจกแจงแบบใกล้เคียงปกติ (Near Normal) สมมาตรหางยาว (Symmetric Long-Tailed) ไม่สมมาตรทางสั้น (Asymmetric Short-Tailed) และไม่สมมาตรหางยาว (Asymmetric Long-Tailed) และเมื่อเปรียบเทียบอำนาจการทดสอบกับตัวสถิติอื่นๆ พบว่า ตัวสถิติทดสอบ r มีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบ W เมื่อประชากรมีการแจกแจงแบบใกล้เคียงปกติ (Near Normal) และสมมาตรหางยาว (Symmetric Long-Tailed)

สมพิศ โชติวิฑูระธารากร (2530) ทำวิทยานิพนธ์เรื่องการเปรียบเทียบอำนาจการทดสอบของตัวสถิติบางตัวที่ใช้ในการทดสอบการแจกแจงปกติ ได้ทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกติ 5 ตัว คือ ตัวสถิติทดสอบไคกำลังสอง (χ^2) ตัวสถิติทดสอบ Studentized Range (u) ตัวสถิติทดสอบ Shapiro-Wilk (W) ตัวสถิติทดสอบ

Filliben (r) และตัวสถิติทดสอบ *Hannu oja* (T_1 และ T_2) ภายใต้การแจกแจง 2 ลักษณะที่สำคัญ คือ การแจกแจงแบบปกติ และการแจกแจงแบบไม่ปกติ ซึ่งพบว่าในด้านการนำไปใช้ประโยชน์ควรเลือกใช้ตัวสถิติทดสอบ W เมื่อขนาดตัวอย่างน้อยกว่าหรือเท่ากับ 50 เนื่องจากให้อำนาจการทดสอบสูงเป็นส่วนใหญ่ และสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี แต่เมื่อขนาดตัวอย่างมากกว่า 50 และประชากรมีการแจกแจงแบบสมมาตรหางยาว (*Symmetric Long-Tailed*) และไม่สมมาตรหางยาว (*Asymmetric Long-Tailed*) ควรเลือกใช้ตัวสถิติทดสอบ r

Zhang (2002) ได้เสนอตัวสถิติที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบต่างๆ ที่ต้องการทดสอบ โดยใช้วิธีการใหม่ของการ *parameterize* ซึ่งไม่เพียงแต่ทำให้ได้ตัวสถิติทดสอบเก่าๆ ได้แก่ ตัวสถิติทดสอบ *Kolmogorov-Smirnov* (K) ตัวสถิติทดสอบ *Anderson-Darling* (A^2) และตัวสถิติทดสอบ *Cramer-von Mises* (W^2) ยังทำให้ได้ตัวสถิติทดสอบใหม่ด้วย ได้แก่ ตัวสถิติทดสอบ Z_K ตัวสถิติทดสอบ Z_A และ ตัวสถิติทดสอบ Z_C และได้ทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบต่างๆ 7 ตัว คือ ตัวสถิติทดสอบ *Kolmogorov-Smirnov* (K) ตัวสถิติทดสอบ *Anderson-Darling* (A^2) ตัวสถิติทดสอบ *Cramer-von Mises* (W^2) ตัวสถิติทดสอบไคกำลังสอง (χ^2) ตัวสถิติทดสอบ Z_K ตัวสถิติทดสอบ Z_A และ ตัวสถิติทดสอบ Z_C พบว่า ตัวสถิติทดสอบใหม่ทั้ง 3 ตัวให้อำนาจการทดสอบได้ดีกว่าตัวสถิติทดสอบเก่า

Zhang และ Wu (2004) ได้ทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกติ 6 ตัว คือ ตัวสถิติทดสอบ Z_K ตัวสถิติทดสอบ Z_A ตัวสถิติทดสอบ Z_C ตัวสถิติทดสอบ *Shapiro-Wilk* (W) ตัวสถิติทดสอบ *Anderson-Darling* (A^2) และตัวสถิติทดสอบ *D'Agostino* (D) ภายใต้การแจกแจงแบบต่างๆ ได้แก่ การแจกแจงแบบเบตา การแจกแจงแบบที การแจกแจงแบบแกมมา การแจกแจงแบบไวบูลล์ และการแจกแจงแบบลอกนอร์มอล ได้ผลสรุปโดยย่อดังนี้

1. ตัวสถิติทดสอบใหม่ทั้ง 3 ตัว คือ ตัวสถิติทดสอบ Z_K ตัวสถิติทดสอบ Z_A และตัวสถิติทดสอบ Z_C มีอำนาจการทดสอบสูง เมื่อประชากรมีการแจกแจงที่ต่างไปจากปกติ
2. โดยทั่วไป ตัวสถิติทดสอบ Z_A และตัวสถิติทดสอบ Z_C มีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบ *Shapiro-Wilk* (W) ตัวสถิติทดสอบ *Anderson-Darling* (A^2) และตัวสถิติทดสอบ Z_K
3. ตัวสถิติทดสอบ Z_K มีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบ A^2
4. ตัวสถิติทดสอบ *D'Agostino* (D) มีอำนาจการทดสอบต่ำที่สุด

ตัวสถิติทดสอบแต่ละตัวมีอำนาจการทดสอบและความไวที่แตกต่างกันในสถานการณ์ที่แตกต่างกัน ดังนั้นในการวิจัยครั้งนี้ ผู้วิจัยจึงสนใจที่จะทำการศึกษาเพื่อเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติ 3 ตัว ซึ่งมีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบอื่นๆ ในสถานการณ์ต่างๆ กันดังกล่าวข้างต้น ได้แก่ ตัวสถิติทดสอบ *Filliben* (r) ตัวสถิติทดสอบ Z_A และตัวสถิติทดสอบ Z_C ซึ่งเป็นตัวสถิติที่ไม่ใช่พารามิเตอร์ทั้งหมด โดยจากการศึกษาพบว่ายังไม่มีผู้ใดเปรียบเทียบตัวสถิติทดสอบทั้ง 3 ตัวนี้ร่วมกัน เพื่อให้ผู้วิเคราะห์สามารถเลือกตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกติที่เหมาะสมกับแต่ละสถานการณ์

1.2 วัตถุประสงค์ของการวิจัย

ในการวิจัยครั้งนี้มีวัตถุประสงค์ของการวิจัยดังนี้

1. เพื่อศึกษาเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (*Type I Error*) และอำนาจการทดสอบ (*Power of the Test*) ภายใต้สถานการณ์ต่างๆ ของตัวสถิติ 3 ตัวที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติ คือ

1.1 ตัวสถิติทดสอบ *Filliben* (r)

1.2 ตัวสถิติทดสอบ Z_A

1.3 ตัวสถิติทดสอบ Z_C

2. เพื่อหาข้อสรุปว่าตัวสถิติที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติแต่ละตัว เหมาะที่จะใช้ในกรณีใด

1.3 สมมติฐานของการวิจัย

ในการวิจัยครั้งนี้มีสมมติฐานของการวิจัย คือ

ตัวสถิติทดสอบใหม่ Z_A และ Z_C ให้อำนาจการทดสอบใกล้เคียงกัน และมีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบ *Filliben* (r)

1.4 ขอบเขตของการวิจัย

ในการวิจัยครั้งนี้กระทำภายใต้ขอบเขตดังนี้

1. ศึกษาความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติ 3 ตัว ได้แก่ ตัวสถิติทดสอบ r ตัวสถิติทดสอบ Z_A และตัวสถิติทดสอบ Z_C

2. ในการกำหนดสมมติฐานของการทดสอบ เนื่องจากข้อมูลแต่ละชุดให้ค่าเฉลี่ยและความแปรปรวนแตกต่างกัน ดังนั้นเพื่อความสะดวกและเป็นบรรทัดฐานเดียวกัน ในการวิจัยจึงจะทำการ แปลงข้อมูลแต่ละชุดให้เป็นมาตรฐาน โดยให้ค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ดังนั้นจึงกำหนดสมมติฐานของการทดสอบ คือ H_0 : ประชากรมีการแจกแจงแบบ $N(0,1)$ และ H_1 : ประชากรไม่ได้มีการแจกแจงแบบ $N(0,1)$

สำหรับการทดสอบสมมติฐานภายใต้ค่าของพารามิเตอร์อื่นๆ ของการแจกแจงแบบปกติ นั้น สามารถทดสอบได้ทำนองเดียวกัน

3. สำหรับการเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 กำหนดให้ประชากรมีการแจกแจงแบบปกติ โดยกำหนดค่าของพารามิเตอร์ ได้แก่ ค่าเฉลี่ยของประชากร (μ) เท่ากับ 0 และความแปรปรวนของประชากร (σ^2) เท่ากับ 1 เพื่อทดสอบภายใต้สมมติฐานว่าง H_0 : ประชากรมีการแจกแจงแบบ $N(0,1)$

4. สำหรับการเปรียบเทียบอำนาจการทดสอบ กำหนดให้ประชากรมีการแจกแจงแบบต่างๆ ได้แก่ การแจกแจงแบบลอการิธึม (Lognormal Distribution) การแจกแจงแบบเบตา (Beta Distribution) การแจกแจงแบบแกมมา (Gamma Distribution) และการแจกแจงแบบที (Student's T Distribution) และการแจกแจงแบบจอห์นสัน (Johnson Distribution)

การกำหนดค่าพารามิเตอร์ของแต่ละการแจกแจงนั้น จะใช้ค่าสัมประสิทธิ์ความเบ้ (γ_1) และสัมประสิทธิ์ความโด่ง (γ_2) เป็นตัวกำหนดค่าพารามิเตอร์

สำหรับการแบ่งลักษณะความเบ้และความโด่งนั้น จะกำหนดตามหลักเกณฑ์ของ Shapiro และคณะ (1968) โดยแบ่งเป็น 5 กรณี ดังนี้

- 4.1 ลักษณะใกล้เคียงการแจกแจงแบบปกติ (Near Normal)
- 4.2 ลักษณะที่สมมาตรและหางสั้น (Symmetric Short -Tailed)
- 4.3 ลักษณะที่สมมาตรและหางยาว (Symmetric Long -Tailed)
- 4.4 ลักษณะที่ไม่สมมาตรและหางสั้น (Asymmetric Short -Tailed)
- 4.5 ลักษณะที่ไม่สมมาตรและหางยาว (Asymmetric Long -Tailed)

ในงานวิจัยครั้งนี้สนใจศึกษาในกรณีไม่ทราบค่าพารามิเตอร์ และทำการประมาณค่าพารามิเตอร์ โดยใช้คุณสมบัติของตัวประมาณไม่เอนเอียง (*Unbiased Estimator*) (รายละเอียดอยู่ในหัวข้อ 3.1.3 บทที่ 3)

5. การหาค่าอำนาจการทดสอบ ทำได้โดยแปลงข้อมูลที่ได้จากการแจกแจงแบบต่างๆ ให้มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ซึ่งการแปลงข้อมูลนั้นเพื่อเป็นการกำหนดสมมติฐานว่าง (H_0) ให้อยู่ในรูปแบบเดียวกันสำหรับทุกการแจกแจง โดยการแปลงข้อมูลนั้นไม่ทำให้ค่าสถิติทดสอบเปลี่ยนแปลงไป (รายละเอียดอยู่ในหัวข้อ 3.1.4 บทที่ 3)

6. ขนาดตัวอย่างที่ทำการศึกษา คือ 10 15 20 30 40 50 60 70 และ 80

7. กำหนดระดับนัยสำคัญ 3 ระดับ คือ 0.01 0.05 และ 0.10

8. ในการศึกษาวิจัยครั้งนี้ใช้โปรแกรมคอมพิวเตอร์ภาษาฟอร์แทรน (*Fortran Power Station*) หาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบด้วยวิธีการจำลองด้วยเทคนิคมอนติคาร์โล กำหนดจำนวนรอบทำซ้ำ 1,000 รอบในแต่ละสถานการณ์

1.5 เกณฑ์การประเมินตัวสถิติทดสอบ

ในการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกตินั้น ในการวิจัยครั้งนี้จะพิจารณาจาก

1. ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ซึ่งวัดจากสัดส่วนความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างนั้นเป็นจริง

ในการวิจัยครั้งนี้จะใช้เกณฑ์การเปรียบเทียบวิธีประมาณค่าแบบช่วงโดยใช้การทดสอบทวินาม (*binomial test*) (รายละเอียดอยู่ในหัวข้อ 3.2.3 บทที่ 3)

2. อำนาจการทดสอบ ซึ่งวัดจากสัดส่วนความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างนั้นเป็นเท็จ โดยจะเปรียบเทียบเฉพาะตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้เท่านั้น

1.6 คำจำกัดความ

ในการวิจัยครั้งนี้มีคำจำกัดความที่ใช้ดังนี้

1. ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (*Type I Error*) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง (*Null Hypothesis*) เมื่อสมมติฐานว่างนั้นเป็นจริง

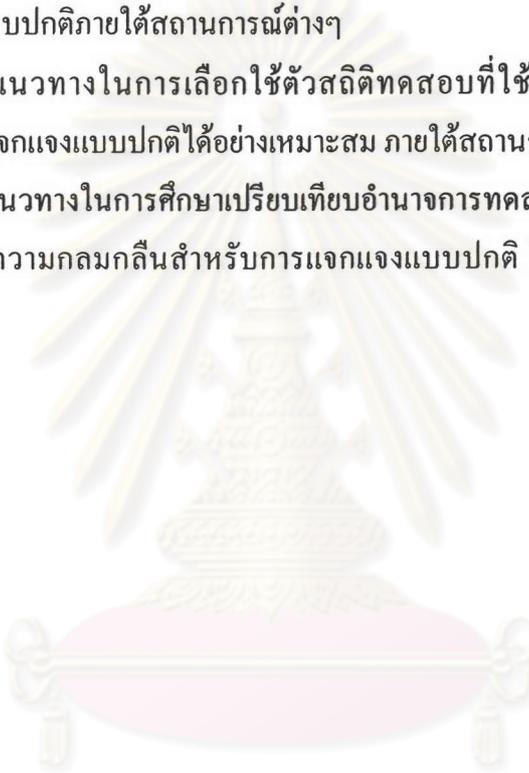
2. ความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 2 (*Type II Error*) หมายถึง ความน่าจะเป็นที่จะยอมรับสมมติฐานว่าง (*Null Hypothesis*) เมื่อสมมติฐานว่างนั้นเป็นเท็จ

3. อำนาจการทดสอบ (*Power of the Test*) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างนั้นเป็นเท็จ ซึ่งมีค่าเท่ากับ $1 - \beta$ เมื่อ β คือ ค่าความน่าจะเป็นที่เกิดความคลาดเคลื่อนประเภทที่ 2

1.7 ประโยชน์ที่คาดว่าจะได้รับ

ในการวิจัยครั้งนี้คาดว่าจะได้รับประโยชน์ดังนี้

1. ทราบอำนาจการทดสอบของตัวสถิติทดสอบที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติภายใต้สถานการณ์ต่างๆ
2. เพื่อเป็นแนวทางในการเลือกใช้ตัวสถิติทดสอบที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติได้อย่างเหมาะสม ภายใต้สถานการณ์ต่างๆ
3. เพื่อเป็นแนวทางในการศึกษาเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบที่ใช้ในการทดสอบเทียบความกลมกลืนสำหรับการแจกแจงแบบปกติ โดยใช้ตัวสถิติทดสอบอื่นๆต่อไป



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย