บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในงานวิจัย ไม่ว่าจะเป็นทางการตลาด การแพทย์ เกษตร หรือแม้แต่สิ่งแวดล้อม จำเป็นต้องอาศัยระเบียบวิธีการทางสถิติในการศึกษา วิเคราะห์ และประมวลผลข้อมูลเพื่อหาข้อสรุปตามสมมติฐานของวิจัย ซึ่งข้อสรุปที่ได้จะต้องมีความน่าเชื่อถือและถูกต้องในระดับที่ยอมรับได้ในทางทฤษฎี ดังนั้นในการเลือกใช้สถิติทดสอบ เพื่อให้เหมาะสมกับลักษณะข้อมูลที่ต้องการวิเคราะห์ สามารถวิเคราะห์ข้อมูลได้อย่างถูกต้อง สอดคล้องกับวิธีการทางสถิติและจุดมุ่งหมายของการวิเคราะห์ จึงควรทราบข้อจำกัด ความเหมาะสม และความแตกต่างของสถิติที่ใช้ทดสอบในแต่ละวิธี รวมถึงลักษณะของข้อมูลที่ศึกษา

ในการวิเคราะห์สถิติเชิงอนุมาน (Inference Statistics) ซึ่งแบ่งวิธีการวิเคราะห์ทางสถิติออกเป็น 2 ประเภท คือ สถิติที่ใช้พารามิเตอร์ (Parametric Statistics) และสถิติที่ไม่ใช้พารามิเตอร์ (Nonparametric Statistics) ซึ่งการวิเคราะห์โดยใช้สถิติที่ใช้พารามิเตอร์จะมีความถูกต้องและน่าเชื่อถือมากกว่าสถิติที่ไม่ใช้พารามิเตอร์ เนื่องจากสถิติที่ใช้พารามิเตอร์เป็นสถิติที่ทราบรูปแบบการแจกแจง มีหลักเกณฑ์และข้อกำหนดเบื้องต้น (Assumption) ที่แน่นอนมากกว่าสถิติที่ไม่ใช้พารามิเตอร์ ซึ่งเรียกว่าสถิติที่เป็นอิสระต่อการแจกแจง (Distribution-free Statistics)

การวิเคราะห์ทางสถิติที่ใช้พารามิเตอร์ต้องเป็นไปตามข้อกำหนดเบื้องต้น ไม่เช่นนั้นจะเกิดความคลาดเคลื่อนในการวิเคราะห์ข้อมูลได้ ซึ่งข้อกำหนดเบื้องต้นหนึ่งที่สำคัญในการวิเคราะห์ คือข้อกำหนดเกี่ยวกับการแจกแจงของประชากร โดยเฉพาะการแจกแจงแบบปกติ ในกรณีที่ข้อมูลมีการแจกแจงปกติและมีลักษณะข้อมูลเป็นไปตามข้อกำหนดเบื้องต้นแล้ว การวิเคราะห์โดยสถิติที่ใช้พารามิเตอร์จะมีประสิทธิภาพและกำลังการทดสอบ (Power of the test) สูงกว่าการวิเคราะห์โดยสถิติที่ไม่ใช้พารามิเตอร์ (Sprent and Smeeton, 2000) สำหรับการทดสอบว่าประชากรมีการแจกแจงปกติหรือไม่ สามารถพิจารณาได้ง่ายโดยการเขียนกราฟการกระจายของข้อมูล แต่เป็นวิธีการที่มีโอกาสผิดพลาดสูง เนื่องจากการพิจารณาจากกราฟนั้นขึ้นอยู่กับดุลพินิจของแต่ละคนจึงทำให้ผลที่แตกต่างกัน ในปัจจุบันจึงมีนักสถิติหลายท่านพยายามเสนอ

ตัวสถิติทดสอบที่ใช้ในการทดสอบรูปแบบการแจกแจงดังกล่าว ซึ่งตัวสถิติทดสอบนั้นอาจแบ่งได้เป็น 2 ประเภท คือตัวสถิติทดสอบที่ใช้อัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test statistic) และไม่ใช้อัตราส่วนภาวะน่าจะเป็น (Non-likelihood ratio test statistic) ซึ่งให้ผลที่ถูกต้องและแน่นอนกว่าวิธีการพิจารณาจากกราฟ อาทิเช่น

ชาง (Zhang, 2002) เสนอตัวสถิติทดสอบ 3 ตัวคือ Z_A , Z_C และ Z_K ในการทดสอบภาวะสารูปสนิทดี (Goodness of fit test) สำหรับการแจกแจงแบบต่างๆ โดยตัวสถิติทดสอบที่เสนอนั้นอาศัยพื้นฐานจากอัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test statistic) และมีการให้ค่าถ่วงน้ำหนักที่เหมาะสมให้กับ อัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test statistic) พบว่าตัวสถิติทดสอบคือ Z_A , Z_C และ Z_K นั้นให้กำลังการทดสอบสูงกว่าตัวสถิติทดสอบ Cramer-von Mises และตัวสถิติทดสอบ Kolmogorov-Smirnov

และต่อมาชางและวู (Zhang and Wu, 2005) ได้ทำการศึกษาและเปรียบเทียบตัวสถิติทดสอบที่ได้อัตราส่วนภาวะน่าจะเป็น (Likelihood ratio test statistic) คือ Z_A, Z_C และ Z_K เพิ่มเติมจากการศึกษาของชาง (Zhang, 2002) เพื่อใช้ทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติ และเสนอว่าในการใช้สถิติทดสอบทั้ง 3 นี้ จะต้องทำการประมาณค่าพารามิเตอร์เสมอถึงแม้ว่าจะทราบค่าพารามิเตอร์ของประชากรหรือไม่ก็ตาม

จากการศึกษาเปรียบเทียบกำลังการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงปกติจากหลายๆ วิธีนี้ พบว่ากำลังการทดสอบของวิธีการทดสอบที่แตกต่างกัน จะขึ้นอยู่กับประเภทของการแจกแจงของประชากรที่ไม่เป็นแบบปกติด้วย (Seier, 2003)

ดังนั้นเพื่อเพิ่มความสะดวกในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติ และหากำลังการทดสอบสำหรับตัวสถิติทดสอบทั้ง 2 ประเภท คือตัวสถิติทดสอบที่ใช้อัตราส่วนภาวะน่าจะเป็นประกอบด้วย ตัวสถิติทดสอบ Z_A, ตัวสถิติทดสอบ Z_C, ตัวสถิติทดสอบ Z_K และตัวสถิติทดสอบ Anderson-Darling (A^2) และตัวสถิติทดสอบที่ไม่ใช้อัตราส่วนภาวะน่าจะเป็น (Non-likelihood ratio test statistic) ประกอบด้วย ตัวสถิติทดสอบ Shapiro-Wilk (W) และ ตัวสถิติทดสอบ Shapiro-Francia (W') ในการพิจารณาหาตัวสถิติทดสอบที่เหมาะสมที่สุดที่ใช้ในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติ สำหรับข้อมูลที่มีลักษณะการแจกแจงที่แตกต่างกัน ผู้วิจัยจึงสนใจศึกษาและเปรียบเทียบกำลังการทดสอบของ ตัวสถิติทดสอบดังกล่าว กรณีที่ไม่ทราบค่าพารามิเตอร์ของประชากร เมื่อข้อมูลมีขนาดตัวอย่างเท่ากับ 10, 50 และ 100 ระดับนัยสำคัญเท่ากับ 0.05

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาความสามารถในการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 และกำลังการทดสอบ ได้แก่ ตัวสถิติทดสอบ Z_A, ตัวสถิติทดสอบ Z_C, ตัวสถิติทดสอบ Z_K, ตัวสถิติทดสอบ Anderson-Darling (A^2), ตัวสถิติทดสอบ Shapiro-Wilk (W) และ ตัวสถิติทดสอบ Shapiro-Francia (W')

2. เพื่อศึกษาความเหมาะสมในการเลือกใช้ตัวสถิติทดสอบ สำหรับข้อมูลที่มีการแจกแจงลักษณะต่างๆ

1.3 สมมติฐานของการวิจัย

ในการศึกษานี้มีสมมติฐานวิจัยคือตัวสถิติทดสอบของชางและวู (Zhang and Wu, 2005) ได้แก่ ตัวสถิติทดสอบ Z_A, ตัวสถิติทดสอบ Z_C และ ตัวสถิติทดสอบ Z_K สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ดี และมีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบ Anderson-Darling (A^2), ตัวสถิติทดสอบ Shapiro-Wilk (W) และตัวสถิติทดสอบ Shapiro-Francia (W') ซึ่งกำลังการทดสอบของตัวสถิติทดสอบจะมีความเหมาะสมกับข้อมูลที่มีลักษณะการแจกแจงของประชากร และขนาดตัวอย่างต่างๆ กัน

1.4 ข้อตกลงเบื้องต้น

1. ในการศึกษากำลังการทดสอบของตัวสถิติทดสอบนั้น ตัวสถิติทดสอบจะต้องสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ได้ จึงจะหากำลังการทดสอบของตัวสถิติทดสอบ และนำกำลังการทดสอบของตัวสถิติทดสอบมาเปรียบเทียบกัน เพื่อเลือกตัวสถิติทดสอบที่เหมาะสมในการทดสอบภาวะสารูปสนิทดีสำหรับการแจกแจงแบบปกติ โดยการเลือกนั้นจะพิจารณาจากกำลังการทดสอบที่มีค่าสูงสุดในแต่ละลักษณะข้อมูล

2. ในการศึกษานี้จะทำการแปลงข้อมูลที่ได้จากการแจกแจงแบบต่างๆ ให้มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 เพื่อให้การทดสอบสมมติฐานในการทดสอบภาวะสารูปสนิทดีสำหรับการทดสอบการแจกแจงแบบปกติเป็นบรรทัดฐานเดียวกัน กล่าวคือ สมมติฐานว่าง (H_0) : ประชากรมีการแจกแจงแบบปกติมาตรฐาน และสมมติฐานแย้ง (H_1) : ประชากรไม่มีการแจกแจงแบบปกติมาตรฐาน

1.5 ขอบเขตของการวิจัย

การดำเนินการวิจัยนี้มีขอบเขต คือ

1. ในการศึกษาการควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ของตัวสถิติทดสอบ Z_A, Z_C, Z_K, A^2, W และ W' จะกำหนดประชากรให้มีการแจกแจงแบบปกติมาตรฐาน ที่มีพารามิเตอร์คือค่าเฉลี่ย (μ) เท่ากับ 0 และความแปรปรวน (σ^2) เท่ากับ 1

2. ในการศึกษาเปรียบเทียบกำลังการทดสอบของตัวสถิติทดสอบ จะกำหนดให้ประชากรมีการแจกแจงแบบที, การแจกแจงแบบล็อกนอร์มัล, การแจกแจงแบบบีตา และการแจกแจงแบบไวบูลล์ ที่มีพารามิเตอร์ซึ่งกำหนดโดยสัมประสิทธิ์ความเบ้ (Coefficient of Skewness : γ_1) และสัมประสิทธิ์ความโด่ง (Coefficient of Kurtosis : γ_2) และจากหลักเกณฑ์ของ ชาพิโร และคณะ (Shapiro, Wilk and Chen, 1968) นั้นจะแบ่งลักษณะความเบ้และความโด่งเป็น 5 กรณี คือ

- ลักษณะใกล้เคียงการแจกแจงแบบปกติ (Near Normal)
- ลักษณะที่สมมาตรและหางยาว (Symmetric Long-tailed)
- ลักษณะที่สมมาตรและหางสั้น (Symmetric Short-tailed)
- ลักษณะที่ไม่สมมาตรและหางยาว (Asymmetric Long-tailed)
- ลักษณะที่ไม่สมมาตรและหางสั้น (Asymmetric Short-tailed)

3. กำหนดขนาดตัวอย่างที่ศึกษาดังนี้
   - กรณีตัวสถิติทดสอบ Z_A, Z_C, Z_K และ Anderson-Darling (A^2) จะกำหนดขนาดตัวอย่างเท่ากับ 10, 50 และ 100
   - กรณีตัวสถิติทดสอบ Shapiro-Wilk (W) จะกำหนดขนาดตัวอย่างเท่ากับ 10 และ 50
   - กรณีตัวสถิติทดสอบ Shapiro-Francia (W') จะกำหนดขนาดตัวอย่างเท่ากับ 50 และ 100
4. กำหนดระดับนัยสำคัญในการทดสอบ คือ 0.05
5. ในการศึกษานี้จะทำซ้ำในแต่ละสถานการณ์จำนวน 1,000 ครั้ง

1.6 เกณฑ์ที่ใช้ในการเปรียบเทียบ

ในการเปรียบเทียบตัวสถิติทดสอบจะใช้กำลังการทดสอบ (Power of the test) ของตัวสถิติทดสอบเป็นเกณฑ์ในการเปรียบเทียบ โดยที่สถิติทดสอบจะต้องสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ได้ ก่อนที่จะนำไปเปรียบเทียบกำลังการทดสอบ ในการพิจารณาความสามารถในการควบคุมความน่าจะเป็นประเภทที่ 1 ของการทดสอบที่ระดับนัยสำคัญต่างๆ จะใช้การทดสอบสมมติฐานว่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้ (หาได้จากอัตราส่วนของจำนวนครั้งในการปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างจริง) นั้นจะเท่ากับความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่กำหนดไว้หรือไม่ ส่วนการหากำลังการทดสอบ หาได้จากอัตราส่วนความน่าจะเป็นในการปฏิเสธสมมติฐานว่างเมื่อสมมติฐานว่างเป็นเท็จ และการสรุปผลว่าตัวสถิติทดสอบใดมีกำลังการทดสอบที่สูงที่สุด หมายความว่าตัวสถิติทดสอบนั้นมีกำลังการทดสอบสูงสุดโดยเปรียบเทียบจากการสถานการณ์เดียวกัน

ในการทดสอบว่าตัวสถิติทดสอบแต่ละตัวว่าสามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ได้หรือไม่นั้นจะใช้การทดสอบทวินาม โดยกำหนดระดับนัยสำคัญคือ 0.05 ซึ่งตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 ที่ระดับ 0.05 ได้จะมีค่าประมาณความน่าจะเป็นที่จะเกิดความผิดพลาดประเภทที่ 1 อยู่ในช่วง [0 , 0.061] (รายละเอียดอยู่ในบทที่ 3)

1.7 คำจำกัดความ

ความผิดพลาดประเภทที่ 1 (Type I error) หมายถึง ความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นจริง และความน่าจะเป็นของความผิดพลาดประเภทนี้คือ ระดับนัยสำคัญ (Significance level) ของการทดสอบ (α)[1]

ความผิดพลาดประเภทที่ 2 (Type II error) หมายถึง ความผิดพลาดที่เกิดจากการยอมรับสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นเท็จ และความน่าจะเป็นของความผิดพลาดประเภทนี้คือ β [1]

กำลังการทดสอบ (Power of the test) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นเท็จ ซึ่งมีค่าเท่ากับ 1- β เมื่อ β คือความน่าจะเป็นของความผิดพลาดประเภทที่ 2 [1]

ความเบ้ (Skewness) ถ้าตัวแปรมีการแจกแจงไม่สมมาตรรอบมัธยฐานหรือค่าเฉลี่ยจะกล่าวว่ามีความเบ้ การแจกแจงมีความเบ้เป็นบวก ถ้าหางทางด้านค่ามากยาวกว่าหางทางด้านค่าน้อย และความเบ้เป็นลบถ้าหางทางด้านค่าน้อยยาวกว่าทางด้านค่ามาก [1]

ความโด่ง (Peakedness or Kurtosis) เป็นมาตรวัดความสูงของการแจกแจงประชากร [1]

1.8 ประโยชน์ที่คาดว่าจะได้รับ

ผลการศึกษาที่ได้สามารถนำไปใช้ประโยชน์ได้ดังนี้

1. เพื่อที่จะสามารถเลือกใช้ตัวสถิติทดสอบในการทดสอบสารูปสนิทดีสำหรับการแจกแจงปกติ ที่มีลักษณะข้อมูลแต่ละแบบได้อย่างเหมาะสม

2. เพื่อเป็นแนวทางในการศึกษาหาวิธีการทางสถิติที่เหมาะสมในการดำเนินงานวิจัยต่อไป

---

[1] Graham Upton and Ian Cook, A Dictionary of Statistics (New York: Oxford University Press, 2004).