

การเปรียบเทียบประสิทธิภาพของสถิติทดสอบในการทดสอบ ความสมมาตรของการแจกแจงด้วยวิธีการสถิติแบบไม่อิงพารามิเตอร์ Comparison of Testing Statistics for Symmetry by Nonparametric Methods

ธานี จิตรูว* และสำรวม จงเจริญ

สาขาวิชาสถิติ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

แขวงคลองจั่น เขตบางกะปิ กรุงเทพมหานคร 10240

Tanee Jitrua* and Samruam Chongcharoen

Department of Statistics, Faculty of Science, National Institute of Development Administration,

Khlong Chan, Bang Kapi, Bangkok, 10240

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบในการทดสอบความสมมาตรของการแจกแจงด้วยวิธีการแบบไม่อิงพารามิเตอร์ ซึ่งประกอบด้วยสถิติทดสอบด้วยวิธีโมเมนต์ (M) สถิติทดสอบของกูปตา (G) สถิติทดสอบด้วยวิธีแบบง่ายปรับปรุงใหม่ (S) และสถิติทดสอบทริเบิล (T) ในการวัดประสิทธิภาพของสถิติทดสอบนั้นพิจารณาจากความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบของสถิติทดสอบภายใต้สถานการณ์ที่กำหนดคือประชากรมีการแจกแจงแบบเบต้า แบบจอห์นสันไม่มีขอบเขต แบบเอฟ แบบไคสแควร์ แบบโลจิสติกส์ กำหนดระดับนัยสำคัญ 0.01 และ 0.05 ที่ขนาดตัวอย่าง 25, 50, 100 และ 300 จำลองข้อมูลภายใต้สถานการณ์ต่าง ๆ ทำซ้ำในแต่ละสถานการณ์จำนวน 1,000 รอบ และใช้ข้อมูลจริงประกอบการศึกษาในครั้งนี้ด้วย ผลการศึกษาพบว่าสถิติทดสอบ T และ S สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ทั้งหมดในทุกกรณี ส่วนสถิติทดสอบ M สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้เฉพาะที่ขนาดตัวอย่างตั้งแต่ 100 ขึ้นไปและสถิติทดสอบ G ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้เมื่อประชากรมาจากการแจกแจงแบบเบต้า และแบบจอห์นสันไม่มีขอบเขต เมื่อเปรียบเทียบอำนาจการทดสอบพบว่าสถิติทดสอบ T มีอำนาจการทดสอบที่สูงที่สุดรองลงมาเป็นสถิติทดสอบ S และสถิติทดสอบ M ตามลำดับ

คำสำคัญ : ความน่าจะเป็นของความผิดพลาดประเภทที่ 1; อำนาจการทดสอบ; ความสมมาตรของการแจกแจง

Abstract

The objective of this research was to compare the efficiency of four symmetry test statistics. The tests are moment test (M), gupta test (G), modified simple test (S) and triple

test (T). The efficiency from these test statistics were considered from probability of type I error and the power of the test measured of these following distribution beta, unbounded Johnson, F, Chi-square, and logistics distributions under study varying from sample size 25, 50, 100, and 300 at 0.01 and 0.05 level of significance. The 1,000 Monte Carlo simulations technique were run for these situation and real data are used in this research as well. The result of this research show that T and S test statistics could control the probability of type I error for all situations, M test statistics could control the probability of type I error at the sample size greater than 100, G test statistics cannot control the probability of type I error for beta and unbounded Johnson distribution. When we considered the power of the test, in overall, T test statistics performed best, follow by S and M test statistics respectively.

Keywords: type I error; power of the test; symmetry of distribution

1. บทนำ

สมบัติความสมมาตรของการแจกแจงของข้อมูล เป็นสมบัติหนึ่งที่มีความสำคัญในระเบียบวิธีการทางสถิติทั้งทางสถิติพรรณนาและสถิติอนุมาน ในด้านสถิติพรรณนา ความสมมาตรหรือความเบ้เป็นหนึ่งในสี่มิติสำคัญนอกเหนือจากการวัดแนวโน้มเข้าสู่ศูนย์กลาง การวัดการกระจายและการวัดความโด่ง ถ้าเพียงแต่ทราบค่าตัวเลขสี่มิติสำคัญนี้ แม้ไม่มีโอกาสได้รับทราบข้อมูลดิบมาก่อน ก็จะสามารถทราบลักษณะโดยรวมและสาระสำคัญอื่น ๆ ของข้อมูลนั้นคร่าว ๆ ในด้านสถิติอนุมาน เราจำเป็นต้องทราบการแจกแจงของข้อมูลว่าเป็นแบบใด มีความสมมาตรหรือความเบ้หรือไม่ ก่อนที่จะเลือกใช้ตัวสถิติในการประมาณค่าและทดสอบสมมุติฐาน รวมถึงการแปลงข้อมูลให้มีการแจกแจงแบบปกติ ทั้งนี้มีรายงานการวิจัยหลายฉบับที่กล่าวถึงความสำคัญของรูปแบบการแจกแจงของประชากรมีผลกระทบของความสมมาตร และความเบ้ที่มีต่อการทดสอบสมมุติฐานและการเลือกใช้ตัวสถิติให้เหมาะสม ได้แก่ งานวิจัยของบุญกอง [1] ได้ศึกษาเปรียบเทียบการจำแนกกลุ่มสองกลุ่มระหว่างการวิเคราะห์การถดถอยพหุและการวิเคราะห์จำแนกประเภทเมื่อตัวแปร

อิสระมีการแจกแจงแบบเบ้ พบว่าลักษณะการแจกแจงของตัวแปรอิสระมีอิทธิพลต่อประสิทธิภาพการจำแนกกลุ่ม โดยที่การแจกแจงแบบปกติจะให้ประสิทธิภาพสูงกว่าการแจกแจงแบบเบ้ หรืองานวิจัยของพรรณภัทร [2] ได้ศึกษาเปรียบเทียบสถิติทดสอบสถานะการสุภาพดี (Goodness of fit test) ในวิทยานิพนธ์ เรื่อง สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับตามอัตราส่วนความควรจะเป็นสำหรับการแจกแจงแบบเบ้ พบว่าสถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่นำเสนอจะมีประสิทธิภาพเมื่อข้อมูลมาจากการแจกแจงของประชากรที่มีลักษณะเบ้ขวา เป็นต้น

การทดสอบข้อมูลว่าเป็นข้อมูลที่มาจากการประชากรที่มีการแจกแจงแบบสมมาตรหรือไม่นั้นจะสามารถทดสอบได้หลายวิธี ประกอบด้วย

1.1 การตรวจสอบด้วยวิธีการภาพ

เป็นวิธีการตรวจสอบที่ง่ายและสะดวกที่สุด การตรวจสอบโดยพิจารณาจากกราฟต่าง ๆ ได้แก่ histogram, probability plot, box plot เป็นต้น การตรวจสอบโดยพิจารณาจากกราฟต่าง ๆ ต้องใช้ด้วยความระมัดระวังเพราะข้อสรุปจะขึ้นอยู่กับการพิจารณาของตัวบุคคลซึ่งมีความเห็นแตกต่างกัน

1.2 การทดสอบทางสถิติ

วิธีการนี้ใช้สถิติทดสอบเพื่อตัดสินว่าข้อมูลมีการแจกแจงเป็นแบบสมมาตรหรือไม่ นั่นคือ การทดสอบสมมติฐาน H_0 : ประชากรมีการแจกแจงแบบสมมาตร เทียบกับ H_1 : ประชากรมีการแจกแจงแบบไม่สมมาตร ซึ่งมีหลายวิธีการทดสอบให้เลือกใช้ส่วนใหญ่เป็นวิธีการทดสอบทางด้านสถิติไม่อิงพารามิเตอร์ (nonparametric statistics) ซึ่งเป็นกระบวนการอนุมานเชิงสถิติโดยไม่จำเป็นต้องทราบการแจกแจงของประชากรที่ศึกษา (distribution free) เพียงแต่ทราบว่าค่าที่ได้มาซึ่งตัวอย่างเป็นอิสระซึ่งกันและกัน โดยที่มีการแจกแจงที่เหมือนกัน สามารถใช้ได้กับข้อมูลที่มีมาตรวัดทุกระดับค่าวนได้ง่ายและมีขั้นตอนการทดสอบไม่ยุ่งยากซับซ้อน ได้แก่ การทดสอบแบบง่ายดั้งเดิม (simple test) การทดสอบแบบง่ายปรับปรุงใหม่ (modified simple test) การทดสอบด้วยวิธีโมเมนต์ (moment test) การทดสอบด้วยวิธีของกูปตา (Gupta test) การทดสอบด้วยวิธีทริเบิล (triple test) การทดสอบด้วยวิธีของเดวิดและจอห์นสัน (David and Johnson test) เป็นต้น ซึ่งพบว่าผู้วิจัยหลายท่านได้ศึกษาเกี่ยวกับสถิติทดสอบความสมมาตรของการแจกแจง เช่น งานวิจัยของ Bhattacharya และคณะ [3] ได้ศึกษาเปรียบเทียบประสิทธิภาพของสถิติทดสอบความสมมาตรระหว่างการทดสอบด้วยวิธีโมเมนต์และการทดสอบด้วยวิธีของเดวิดและจอห์นสัน ผลการศึกษาพบว่าสถิติทดสอบด้วยวิธีโมเมนต์มีประสิทธิภาพการทดสอบที่ดีก็ต่อเมื่อตัวอย่างมาจากประชากรที่มีการแจกแจงใกล้เคียงการแจกแจงแบบปกติ และตัวอย่างมีขนาดตั้งแต่ 100 ขึ้นไป ส่วนสถิติทดสอบด้วยวิธีของเดวิดและจอห์นสันไม่ค่อยมีประสิทธิภาพสำหรับการทดสอบการแจกแจงแบบสมมาตร หรือในงานวิจัยของ Zheng และ Gaswirth [4] ได้ศึกษาเปรียบเทียบประสิทธิภาพของ

สถิติทดสอบความสมมาตรระหว่าง การทดสอบแบบง่ายดั้งเดิมของ Carbilio และ Masaro [4] กับการทดสอบแบบง่ายปรับปรุงใหม่ของ Miao และ Gastwirth [4] โดยประชากรมาจากการแจกแจงแบบเบต้า แบบที่ แบบยูนิฟอร์ม และแบบปกติ ผลการศึกษาพบว่าสถิติทดสอบความสมมาตรของการแจกแจงแบบง่ายปรับปรุงใหม่ มีประสิทธิภาพในการทดสอบที่ดีกว่าสถิติทดสอบแบบง่ายดั้งเดิม

จากที่กล่าวมาข้างต้น ผู้วิจัยสนใจที่จะทำการเปรียบเทียบประสิทธิภาพของสถิติทดสอบความสมมาตรของการแจกแจง 4 วิธี คือ การทดสอบแบบง่ายปรับปรุงใหม่ การทดสอบด้วยวิธีโมเมนต์ การทดสอบด้วยวิธีของกูปตา และการทดสอบด้วยวิธีทริเบิล เนื่องจากยังไม่มีข้อสรุปแน่ชัดว่าสถิติทดสอบใดที่กล่าวมาจะมีประสิทธิภาพที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ดีและมีอำนาจการทดสอบที่ดีภายใต้สถานการณ์เดียวกัน

2.วัตถุประสงค์ในการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบความสมมาตรของการแจกแจง 4 วิธี คือ การทดสอบแบบง่ายปรับปรุงใหม่ (modified simple test) การทดสอบด้วยวิธีโมเมนต์ (moment test) การทดสอบด้วยวิธีของกูปตา (Gupta test) และการทดสอบด้วยวิธีทริเบิล (triple test) รวมทั้งเพื่อหาข้อสรุปสำหรับตัวสถิติที่เหมาะสมในการทดสอบความสมมาตรของการแจกแจงของประชากรในแต่ละสถานการณ์

3. สถิติทดสอบที่ใช้ในการศึกษา

กำหนดตัวอย่างสุ่ม $X_1, X_2, \dots, X_n$ มีความเป็นอิสระกันและมาจากประชากรที่มีการแจกแจงแบบต่อเนื่องและไม่ทราบค่ามัธยฐาน โดยมีสมมติฐานในการทดสอบ คือ H_0 : ประชากรมีการแจกแจงแบบ

สมมาตร H_1 : ประชากรมีการแจกแจงแบบไม่สมมาตร ซึ่งสถิติทดสอบที่ใช้ประกอบด้วยสถิติต่อไปนี้

3.1 Triple test statistics (T -test)

เป็นสถิติไม่อิงพารามิเตอร์ ใช้ทดสอบความสมมาตรของการแจกแจง Randles [5,6] สร้างตัวสถิติทดสอบนี้เรียกว่าสถิติทดสอบ T โดยมีสูตรคำนวณ

$$\text{คือ } T = \frac{V}{\sigma} \text{ เมื่อ } V = \sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n f^*(X_i, X_j, X_k)$$

คือ ผลรวมของ $f^*(X_i, X_j, X_k)$ ทั้งหมด โดย

$$f^*(X_i, X_j, X_k) = \left[\text{sign}(X_i + X_j - 2X_k) \right] + \left[\text{sign}(X_i + X_k - 2X_j) \right] + \left[\text{sign}(X_j + X_k - 2X_i) \right]$$

$$\text{ซึ่ง } 1 \leq i < j < k \leq n \text{ และ } \text{sign}(\Delta) = \begin{cases} -1; \Delta < 0 \\ 0; \Delta = 0 \\ 1; \Delta > 0 \end{cases}$$

คือ ฟังก์ชันสำหรับการทดสอบเครื่องหมายเป็นได้ 3 ค่า คือ -1, 0, 1 เมื่อ Δ คือ ค่าที่อยู่ภายในวงเล็บของ

$$\text{ฟังก์ชัน } \text{sign} \text{ คือ } \sigma^2 = \left[\frac{(n-3)(n-4)}{(n-1)(n-2)} \sum_{t=1}^4 B_t^2 + \frac{(n-3)}{(n-4)} \sum_{s=1}^{n-1} \sum_{t=s+1}^n B_{s,t}^2 + \frac{n(n-1)(n-2)}{6} - \left\{ 1 - \frac{(n-3)(n-4)(n-5)}{n(n-1)(n-2)} \right\} T^2 \right] \text{ คือ ค่าความแปร}$$

$$\text{ปรวนของตัวอย่าง } B_t = \left[\sum_{j=t+1}^{n-1} \sum_{k=j+1}^n f^*(X_t, X_j, X_k) \right. \\ \left. + \sum_{j=1}^{t-1} \sum_{k=t+1}^n f^*(X_j, X_t, X_k) + \sum_{j=1}^{t-1} \sum_{k=j+1}^{t-1} f^*(X_j, X_k, X_t) \right]$$

$$\text{สำหรับ } t = 1, \dots, n \text{ คือ ผลรวมของ } f^*(X_t, X_j, X_k) \text{ ที่เป็นไปได้ทั้งหมด ซึ่งมี } X_t \text{ อยู่ } B_{s,t} = \sum_{j=1}^{s-1} f^*(X_j, X_s, X_t) + \sum_{j=s+1}^{t-1} f^*(X_s, X_j, X_t) + \sum_{j=t+1}^n f^*(X_s, X_t, X_j) \text{ สำหรับ } 1 \leq s < t \leq n \text{ คือ ผลรวมของ } f^*(X_t, X_j, X_k) \text{ ที่เป็นไปได้ทั้งหมด ซึ่งมี } X_s \text{ และ } X_t \text{ อยู่ และ } n \text{ คือ}$$

$$\text{ขนาดตัวอย่าง จะปฏิเสธสมมติฐาน } H_0 \text{ (ประชากรมีการแจกแจงแบบสมมาตร) เมื่อ } |T| \geq Z_{\alpha/2} \text{ เมื่อ } Z_{\alpha/2} \text{ คือ ค่าวิกฤตจากการเปิดตารางแจกแจงแบบปกติมาตรฐาน}$$

ขนาดตัวอย่าง จะปฏิเสธสมมติฐาน H_0 (ประชากรมีการแจกแจงแบบสมมาตร) เมื่อ $|T| \geq Z_{\alpha/2}$ เมื่อ $Z_{\alpha/2}$ คือ ค่าวิกฤตจากการเปิดตารางแจกแจงแบบปกติมาตรฐาน

3.2 Moment test statistics (M -test)

กูปตา [7] เป็นผู้สร้างตัวสถิตินี้ โดยประชากรต้องมีการแจกแจงแบบต่อเนื่องและโมเมนต์อันดับที่หนึ่งถึงอันดับที่หกต้องหาค่าได้ เรียกว่าสถิติ

$$\text{ทดสอบ } M \text{ มีสูตรในการคำนวณ คือ } M = \frac{\left(\frac{m_3}{m_2^{3/2}} \right)}{\left(\frac{\sigma}{\sqrt{n}} \right)}$$

$$\text{เมื่อ } m_r = \frac{\sum_{i=1}^n (X_i - \bar{X})^r}{n} \text{ คือ โมเมนต์ศูนย์กลาง}$$

$$\text{อันดับที่ } r; \sigma^2 = \frac{(m_6 - 6m_2m_4 + 9m_3^2)}{nm_2^3} \text{ คือ}$$

ความแปรปรวนตัวอย่าง และ n คือ ขนาดตัวอย่าง จะปฏิเสธสมมติฐาน H_0 (ประชากรมีการแจกแจงแบบสมมาตร) เมื่อ $|M| \geq z_{\alpha/2}$ เมื่อ $Z_{\alpha/2}$ คือ ค่าวิกฤตจากการเปิดตารางแจกแจงแบบปกติมาตรฐาน

3.3 Modified simple test statistics (S -test)

เป็นการทดสอบความสมมาตรของการแจกแจงแบบง่ายโดยการเปรียบเทียบระหว่างค่าเฉลี่ยและค่ามัธยฐาน Miao และคณะ [4] เป็นผู้พัฒนาตัวสถิติทดสอบนี้เรียกว่าสถิติทดสอบ S โดยมีสูตรในการ

$$\text{คำนวณ คือ } S = \frac{(\bar{X} - \theta)}{\left([0.5708/n]^{1/2} \right) J} \text{ เมื่อ}$$

$$J = \frac{\left(\sum_{i=1}^n |X_i - \theta| \right)}{n(n\pi/2)^{1/2}} \text{ คือ ส่วนเบี่ยงเบนมาตรฐานของ}$$

ตัวอย่าง; $\bar{X}$ คือค่าเฉลี่ยตัวอย่าง; θ คือ มัธยฐาน และ n คือ ขนาดตัวอย่าง จะปฏิเสธสมมติฐาน H_0

(ประชากรมีการแจกแจงแบบสมมาตร) เมื่อ $|S| \geq Z_{\alpha/2}$ เมื่อ $Z_{\alpha/2}$ คือ ค่าวิกฤตจากการเปิดตารางแจกแจงแบบปกติมาตรฐาน

3.4 Gupta test statistics (G -test)

เป็นสถิติไม่อิงพารามิเตอร์สำหรับทดสอบความสมมาตรของการแจกแจง โดยกฤษดา [7,8] สร้างตัวสถิติทดสอบนี้เรียกว่าตัวสถิติ G โดยมีสูตรในการคำนวณ คือ $G = \frac{C - (1/4)}{\left(\frac{\sigma}{n}\right)^{1/2}}$ เมื่อ $C = \frac{2(A_1 - A_2)}{n(n-1)}$,

$$A_1 = \sum_{j=2}^n \sum_{i=1}^{j-1} \delta_{ij} \text{ คือ ผลรวมของ } \delta_{ij} \text{ ทั้งหมด ซึ่ง}$$

$$\delta_{ij} = \begin{cases} 1; & (X_i + X_j) \geq 2\theta \\ 0; & (X_i + X_j) \leq 2\theta \end{cases}; \quad j = 2, \dots, n, \quad i = 1, \dots, j-1$$

คือ ฟังก์ชันสำหรับทดสอบเครื่องหมายเป็นได้สองค่า คือ 1 และ 0; θ คือ ค่ามัธยฐาน โดย $A_2 = \frac{[(n-2)/2][n/2]}{2}$ เมื่อ $[\Delta]$ คือ จำนวนที่เต็มที่สุดที่มีค่าน้อยกว่าหรือเท่ากับ Δ , Δ คือ ค่าที่อยู่ภายในเครื่องหมาย $[\]$, n คือ ขนาดตัวอย่าง;

$$\sigma^2 = (1/12) + \frac{\{1 - (2t/A_3)\}^3}{4} \text{ คือ ความแปรปรวน}$$

ตัวอย่าง $A_3 = \frac{D}{2n^{4/5}}; \quad D = \max \left\{ 1, \sum_{i=1}^n a_i \right\}$

คือ ฟังก์ชันสำหรับหาค่ามากที่สุดเปรียบเทียบระหว่าง 1

$$\text{และ } \sum_{i=1}^n a_i; \quad a_i = \begin{cases} 1; & \left\{ \theta - (n)^{-1/5} \right\} \leq X_i \leq \left\{ \theta + (n)^{-1/5} \right\} \\ 0; & \text{otherwise} \end{cases},$$

$i = 1, \dots, n$ คือ ฟังก์ชันสำหรับทดสอบเครื่องหมาย

$$\text{เป็นได้สองค่า คือ 1 และ 0; } t = \frac{Z(\alpha^*/2)}{(\theta_U - \theta_L)(3n)^{1/2}};$$

θ_L และ θ_U คือ ค่าในลำดับที่ L และ U ตามลำดับ

$$\text{ของ } \{(X_i + X_j)/2, \quad i \leq j = 1, \dots, n\} \text{ โดย } L = C_{\alpha^*};$$

$$U = [n(n+1)/2] + 1 - C_{\alpha^*};$$

C_{α^*} คือ จำนวนเต็มที่มากที่สุดที่น้อยกว่าหรือเท่ากับ

$$\frac{n(n+1)}{4} - Z_{\alpha^*/2} \left\{ \frac{n(n+1)(2n+1)}{24} \right\}^{1/2}; \quad Z_{(\alpha^*/2)}$$

คือ ค่าวิกฤตจากการเปิดตารางแจกแจงแบบปกติสำหรับระดับนัยสำคัญ $\alpha^*/2$ ใด ๆ จะปฏิเสธสมมติฐาน H_0 (ประชากรมีการแจกแจงแบบสมมาตร) เมื่อ $|G| \geq Z_{\alpha/2}$ เมื่อ $Z_{\alpha/2}$ คือ ค่าวิกฤตจากการเปิดตารางแจกแจงแบบปกติมาตรฐาน

4. วิธีการวิจัย

4.1 การพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1

4.1.1 จำลองข้อมูลให้มีการแจกแจงแบบสมมาตรรูปแบบต่าง ๆ ที่ขนาดตัวอย่าง 25, 50, 100 และ 300 ซึ่งประกอบด้วย การแจกแจงแบบเบต้า การแจกแจงแบบโลจิสติกส์ การแจกแจงแบบจอห์นสันแบบไม่มีขอบเขต

4.1.2 คำนวณค่าสถิติทดสอบทั้ง 4 การทดสอบ

4.1.3 เปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ดำเนินการดังนี้

(1) นำค่าสถิติที่คำนวณได้เทียบกับค่าวิกฤติ เพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมติฐานหลัก

(2) ทำซ้ำ 1,000 ครั้ง

(3) บันทึกจำนวนครั้งที่ปฏิเสธสมมติฐานหลักกรณีสมมติหลักเป็นจริง (H_0 : ประชากรมีการแจกแจงแบบสมมาตร) ที่ระดับนัยสำคัญ 0.01 และ 0.05 ประมาณความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ด้วยการนับจำนวนครั้งที่ปฏิเสธสมมติฐานหลักหารด้วย 1,000 จะได้ค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1

(4) นำค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ที่ได้ไปเทียบกับเกณฑ์ของ Cochran [9] ว่าสามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้หรือไม่ เมื่อเกณฑ์ของ Cochran คือ ที่ระดับนัยสำคัญ 0.01 ถ้าค่าที่ได้อยู่ในช่วง $[0.007, 0.015]$ ที่ระดับนัยสำคัญ 0.05 ถ้าค่าที่ได้อยู่ในช่วง $[0.04, 0.06]$

4.2 การเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ

4.2.1 จำลองข้อมูลให้มีการแจกแจงแบบไม่สมมาตรรูปแบบต่าง ๆ ที่ขนาดตัวอย่าง 25, 50, 100 และ 300 ซึ่งประกอบด้วยการแจกแจงแบบไคสแควร์ การแจกแจงแบบเบต้า การแจกแจงแบบเอฟ

4.2.2 คำนวณค่าสถิติทดสอบเฉพาะสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้

4.2.3 การหาอำนาจการทดสอบของสถิติทดสอบดำเนินการดังนี้

- (1) นำค่าสถิติที่คำนวณได้เทียบกับค่าวิกฤติเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมติฐานหลัก
- (2) ทำซ้ำ 1,000 ครั้ง
- (3) บันทึกจำนวนครั้งที่ปฏิเสธสมมติฐานหลักกรณีสมมติฐานหลักไม่จริง (H_0 : ประชากรมีการแจกแจงแบบไม่สมมาตร) ที่ระดับนัยสำคัญ 0.01 และ 0.05 หาอำนาจการทดสอบด้วยการนับจำนวนครั้งที่ปฏิเสธสมมติฐานหลัก หาดด้วย 1,000

4.3 วิเคราะห์และสรุปผลเกี่ยวกับประสิทธิภาพของสถิติทดสอบที่ใช้ในแต่ละการทดสอบ

นำค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบสำหรับการแจกแจงแบบสมมาตรในแต่ละกรณีมาวิเคราะห์ เพื่อสรุปหาสถิติทดสอบที่เหมาะสม

5. ผลการวิจัย

5.1 ความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1

ตารางที่ 1 แสดงค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ที่คำนวณได้จากโปรแกรม SAS สามารถสรุปได้ดังนี้

5.1.1 สถิติทดสอบ T และ S สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ทั้งหมดในทุกการแจกแจง ขนาดตัวอย่าง และระดับนัยสำคัญเนื่องจากค่าความน่าจะเป็นที่ได้อยู่ในขอบเขตที่สามารถควบคุมได้ตามเกณฑ์ของ Cochran และค่าที่ได้ส่วนใหญ่จะมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนด คือ 0.01 และ 0.05

5.1.2 สถิติทดสอบ M สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ในทุกการแจกแจงและระดับนัยสำคัญ ที่ขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป เนื่องจากค่าความน่าจะเป็นที่ได้อยู่ในขอบเขตที่สามารถควบคุมได้ตามเกณฑ์ของ Cochran และค่าที่ได้ส่วนใหญ่จะมีค่าน้อยกว่าระดับนัยสำคัญที่กำหนด คือ 0.01 และ 0.05

5.1.3 สถิติทดสอบ G ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้เมื่อประชากรมาจากการแจกแจงแบบเบต้า และการแจกแจงแบบจอห์นสันไม่มีขอบเขต เนื่องจากค่าความน่าจะเป็นที่ได้ส่วนใหญ่อยู่นอกขอบเขตที่สามารถควบคุมได้ตามเกณฑ์ของ Cochran แม้ว่า会增加ขนาดตัวอย่างจาก 25 เป็น 300 หรือเพิ่มระดับนัยสำคัญจาก 0.01 เป็น 0.05 แล้วก็ตาม

เมื่อพิจารณาโดยรวมแล้วสถิติทดสอบ G ไม่สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ในหลายกรณีดังนั้นในการวิจัยครั้งนี้จะไม่นำเอาสถิติทดสอบ G มาพิจารณาเปรียบเทียบอำนาจการทดสอบ ส่วนสถิติทดสอบ M จะพิจารณา

เปรียบเทียบอำนาจการทดสอบที่ขนาดตัวอย่างตั้งแต่ 100 ขึ้นไป

ตารางที่ 1 ค่าความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ของสถิติทดสอบ M สถิติทดสอบ S สถิติทดสอบ T และสถิติทดสอบ G ที่ระดับนัยสำคัญ 0.01 และ 0.05 สำหรับขนาดตัวอย่าง 25, 50, 100 และ 300

การแจกแจง	ขนาดตัวอย่าง	ระดับนัยสำคัญ							
		0.01				0.05			
		M	S	T	G	M	S	T	G
Logistics (5,2)	25	0.001	0.012*	0.014*	0.025	0.037	0.052*	0.056*	0.078
	50	0.006	0.008*	0.013*	0.007*	0.045*	0.053*	0.055*	0.042*
	100	0.007*	0.009*	0.013*	0.010*	0.043*	0.050*	0.057*	0.054*
	300	0.010*	0.009*	0.007*	0.011*	0.050*	0.042*	0.059*	0.050*
Logistic (9,3)	25	0.003	0.012*	0.011*	0.013*	0.034	0.051*	0.046*	0.069
	50	0.003	0.008*	0.011*	0.010*	0.036	0.051*	0.046*	0.050*
	100	0.007*	0.008*	0.010*	0.013*	0.041*	0.047*	0.050*	0.043*
	300	0.009*	0.011*	0.011*	0.010*	0.051*	0.051*	0.050*	0.049*
Beta (5,5)	25	0.002	0.012*	0.011*	0.000	0.032	0.057*	0.047*	0.000
	50	0.005	0.014*	0.011*	0.000	0.040*	0.049*	0.051*	0.000
	100	0.008*	0.008*	0.009*	0.000	0.042*	0.046*	0.055*	0.000
	300	0.009*	0.011*	0.012*	0.000	0.048*	0.048*	0.051*	0.000
Beta (20,20)	25	0.001	0.012*	0.010*	0.000	0.028	0.049*	0.053*	0.000
	50	0.004	0.009*	0.015*	0.000	0.029	0.044*	0.056*	0.000
	100	0.008*	0.008*	0.009*	0.000	0.042*	0.046*	0.055*	0.000
	300	0.009*	0.011*	0.012*	0.000	0.047*	0.048*	0.051*	0.000
Unbounded Johns (0,1.5,0,2)	25	0.001	0.010*	0.013*	0.020	0.018	0.049*	0.052*	0.078
	50	0.003	0.008*	0.016*	0.010*	0.034	0.056*	0.056*	0.050*
	100	0.007*	0.010*	0.013*	0.008*	0.044*	0.047*	0.048*	0.045*
	300	0.007*	0.011*	0.009*	0.011*	0.051*	0.051*	0.049*	0.051*
Unbounded Johns (0,5,0,2)	25	0.001	0.010*	0.011*	0.000	0.032	0.049*	0.053*	0.020
	50	0.004	0.012*	0.014*	0.000	0.037	0.041*	0.050*	0.023
	100	0.007*	0.009*	0.012*	0.002	0.044*	0.048*	0.052*	0.020
	300	0.009*	0.010*	0.011*	0.001	0.048*	0.051*	0.051*	0.027

*สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้

5.2 อำนาจการทดสอบ

ตารางที่ 2 แสดงค่าอำนาจการทดสอบที่คำนวณได้จากโปรแกรม SAS สามารถสรุปได้รายละเอียดต่าง ๆ ดังนี้

5.2.1 ทุกกรณีที่ศึกษาเมื่อขนาดตัวอย่างใหญ่ขึ้น คือ เพิ่มจาก 25 เป็น 300 ค่าอำนาจทดสอบของสถิติทดสอบ M , S และ T จะมีค่าสูงขึ้น โดยค่าอำนาจทดสอบมากที่สุดเมื่อขนาดตัวอย่างเป็น 300

5.2.2 ในทุกกรณีที่ศึกษาเมื่อเพิ่มระดับนัยสำคัญให้สูงขึ้นจาก 0.01 เป็น 0.05 ค่าอำนาจทดสอบของสถิติทดสอบ M , S และ T จะมีค่าสูงขึ้น โดยค่าอำนาจทดสอบมากที่สุดเมื่อระดับนัยสำคัญเป็น 0.05

ตารางที่ 2 ค่าอำนาจทดสอบของสถิติทดสอบ M สถิติทดสอบ S และสถิติทดสอบ T ที่ระดับนัยสำคัญ 0.01 และ 0.05 สำหรับขนาดตัวอย่าง 25, 50, 100 และ 300

การแจกแจง	ขนาดตัวอย่าง	ระดับนัยสำคัญ					
		0.01			0.05		
		M	S	T	M	S	T
χ^2_5	25	-	0.089	0.268	-	0.279	0.480
	50	-	0.254	0.656	-	0.512	0.817
	100	0.491	0.599	0.954	0.800	0.829	0.980
	300	0.893	0.988	1.000	0.984	0.999	1.000
χ^2_{30}	25	-	0.013	0.038	-	0.070	0.097
	50	-	0.037	0.085	-	0.089	0.207
	100	0.087	0.080	0.208	0.369	0.209	0.428
	300	0.612	0.280	0.505	0.540	0.539	0.928
B(10,1)	25	-	0.267	0.589	-	0.541	0.731
	50	-	0.612	0.924	-	0.822	0.970
	100	0.642	0.945	0.998	0.893	0.980	1.000
	300	0.790	1.000	1.000	0.997	1.000	1.000
B(10,8)	25	-	0.004	0.012	-	0.027	0.045
	50	-	0.009	0.016	-	0.044	0.055
	100	0.010	0.012	0.024	0.054	0.059	0.123
	300	0.021	0.019	0.241	0.105	0.061	0.321
F(20,20)	25	-	0.103	0.249	-	0.312	0.451
	50	-	0.309	0.619	-	0.565	0.791
	100	0.273	0.623	0.952	0.761	0.827	0.977
	300	0.315	0.996	1.000	0.889	0.991	1.000
F(100,100)	25	-	0.027	0.051	-	0.091	0.118
	50	-	0.047	0.106	-	0.132	0.288
	100	0.120	0.089	0.260	0.447	0.264	0.404
	300	0.728	0.395	0.787	0.928	0.612	0.945

5.2.3 ในทุกกรณีที่ศึกษาอำนาจการทดสอบของสถิติทดสอบทั้ง 3 จะค่อย ๆ ลดต่ำลงเมื่อค่าความเบ้ลดลง เนื่องจากเมื่อองศาแห่งความอิสระของประชากรมีค่าสูงขึ้น มีผลทำให้การแจกแจงเปลี่ยนรูปร่าง จนเข้าใกล้การแจกแจงแบบสมมาตรมากขึ้น

5.2.4 การเปรียบเทียบอำนาจการทดสอบพบว่าสถิติทดสอบ T จะมีอำนาจการทดสอบที่สูงกว่าสถิติทดสอบ M และ S ในทุกกรณี ส่วนสถิติทดสอบ M และ S พบว่าที่ระดับความเบ้ต่ำสถิติทดสอบ M จะมีอำนาจการทดสอบที่สูงกว่าสถิติทดสอบ S แต่ถ้าความเบ้มากขึ้นสถิติทดสอบ S จะมีอำนาจการทดสอบที่สูงกว่า

ตารางที่ 3 ค่าสถิติทดสอบ M , S และ T และสัมประสิทธิ์ความเบ้ต่าง ๆ ซึ่งคำนวณมาจากข้อมูลเงินรับฝากและเงินให้สินเชื่อของธนาคารพาณิชย์จำแนกตามจังหวัด ณ เดือนธันวาคม พ.ศ. 2555 จากเว็บไซต์ธนาคารแห่งประเทศไทย

Symmetry test ($\alpha = 0.01$)			
Test	Statistic	p-value	Shape of distribution
Pearson's moment coefficient of skewness	2.2167	-	Right skewness
Bowley's measure of skewness	0.8810	-	Right skewness
Kelly's measure of skewness	0.8956	-	Right skewness
Pearson's measure of skewness	1.4992	-	Right skewness
T -test statistics	29.2196	0.0000	Not symmetry
M -test statistics	4.2832	0.0000	Not symmetry
S -test statistics	10.9617	0.0000	Not symmetry

7. สรุป

สถิติทดสอบ T และสถิติทดสอบ S เป็นตัวสถิติที่มีประสิทธิภาพสูงสำหรับการทดสอบความสมมาตรของการแจกแจงของประชากร เนื่องจากสามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ดี และมีอำนาจการทดสอบสูง ส่วนสถิติทดสอบ M จะมีประสิทธิภาพสูง เมื่อตัวอย่างมี

6. การศึกษาสถิติทดสอบโดยใช้ข้อมูลจริง

ตารางที่ 3 แสดงค่าสถิติทดสอบ M , S และ T และสัมประสิทธิ์ความเบ้ต่าง ๆ สำหรับทดสอบความสมมาตรของการแจกแจงซึ่งคำนวณมาจากข้อมูลจริง (ข้อมูลเงินรับฝากและเงินให้สินเชื่อของธนาคารพาณิชย์จำแนกตามจังหวัด ณ เดือนธันวาคม พ.ศ. 2555 จากเว็บไซต์ธนาคารแห่งประเทศไทย) และฮิสโตแกรม (รูปที่ 1) พบว่าสถิติทดสอบ M , S และ T ให้ผลทดสอบตรงกับฮิสโตแกรม และค่าสัมประสิทธิ์ความเบ้ต่าง ๆ ทั้งหมดซึ่งเป็นไปตามข้อสรุปที่ได้จากการศึกษาโดยการจำลองข้อมูล

ขนาดตั้งแต่ 100 ขึ้นไป

ผลการศึกษานี้ได้ผลสอดคล้องกับงานวิจัยของ Bhattacharya และคณะ [4] ที่ว่าสถิติทดสอบด้วยวิธีโมเมนต์ (M) มีประสิทธิภาพการทดสอบที่ดีก็ต่อเมื่อตัวอย่างมีขนาดตั้งแต่ 100 ขึ้นไป สอดคล้องกับงานวิจัยของ Miao และคณะ [3] ที่ว่าสถิติทดสอบแบบง่ายปรับปรุงใหม่ (S) มีประสิทธิภาพสูงสำหรับ

ทดสอบการแจกแจงแบบสมมาตรสามารถใช้ได้กับทั้งตัวอย่างขนาดเล็กและใหญ่ และสอดคล้องกับตำราทางสถิติของ Hollander และคณะ [6] ที่ว่าสถิติทดสอบโดยวิธีทริบเปิล (T) เป็นตัวสถิติทดสอบที่มีประสิทธิภาพสูง คือ สามารถควบคุมความน่าจะเป็นของความผิดพลาดประเภทที่ 1 ได้ดีและมีอำนาจการทดสอบสูงที่ขนาดตัวอย่างที่เหมาะสมคือตั้งแต่ 25 ขึ้นไป

8. รายการอ้างอิง

- [1] บุญกอง ทะกลโยธิน, 2535, การเปรียบเทียบการจำแนกกลุ่ม 2 กลุ่ม ระหว่างการวิเคราะห์การถดถอยทวิและการวิเคราะห์การจำแนกประเภท เมื่อตัวแปรอิสระมีการแจกแจงแบบเบ้, วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.
- [2] พรรณภัทร แซ่โ้ว, 2557, สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับตามอัตราส่วนความควรจะเป็นสำหรับการแจกแจงแบบเบ้, วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยบูรพา, ชลบุรี.
- [3] Bhattacharya, P.K., Gastwirth, J.L. and Wright, A., 1982, Two modified Wilcoxon test for symmetry about and unknown location parameter, *Biometrika J.* 69: 377-382.
- [4] Zheng, T. and Gastwirth, J.L., 2010, On bootstrap test of symmetry about an unknown median, *J. Data Sci.* 8: 413-427.
- [5] Randles, R.H., Fligner, M.A., Policello, G.E. and Wolfe, D.A., 1980, An asymptotically distribution-free tests for symmetry versus asymmetry, *J. Am. Stat. Assoc.* 75: 168-172.
- [6] Hollander, M., Wolfe, D.A. and Chicken, E., 2014, *Nonparametric Statistical Methods*, 3rd Ed., John Wiley & Sons, New Jersey, 819 p.
- [7] Gupta, M.K., 1967, An asymptotically nonparametric test of symmetry, *Ann. Math. Stat.* 38: 849-866.
- [8] Hollander, M. and Wolfe, D.A., 1973, *Nonparametric Statistical Methods*, John Wiley & Sons, New Jersey, 503 p.
- [9] Cochran, W.G., 1954, Some methods for strengthening the common χ^2 tests, *Biometrics* 10: 417-451.