การทดสอบภาวะสารูปดีแบบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงสำหรับการทดสอบการแจกแจงปรกติ

Goodness of Fit Test Based on Modified Log-Likelihood Ratio for Normal Distribution

ธณัชช์ศักยภ์ ทรงธรรมบวร และ จุฑาพร เนียมวงษ์ *

Tanutsak Songthamboworn and Jutaporn Neamvonk*

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

Department of Mathematics, Faculty of Science, Burapha University

Received : 27 December 2017

Accepted : 6 February 2018

Published online : 16 February 2018

บทคัดย่อ

ในงานวิจัยนี้ได้นำเสนอตัวสถิติทดสอบภาวะสารูปดีโดยใช้ฟังก์ชันการแจกแจงจากการทดลองซึ่งพัฒนามาจากตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง สำหรับการทดสอบการแจกแจงปรกติ และทำการเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบนี้กับตัวสถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุง สถิติทดสอบคาแมร์-ฟอนมิสที่ปรับปรุง ซึ่งเป็นสถิติทดสอบในกลุ่มที่พัฒนามาจากหลักการพื้นฐานของอัตราส่วนล็อกภาวะน่าจะเป็น โดย Zhang (2002) และสถิติทดสอบแอนเดอร์สัน-ดาร์ลิงแบบดั้งเดิม ค่าวิกฤติของสถิติทดสอบได้ถูกสร้างขึ้นจากข้อมูลจำลอง และพิจารณาประสิทธิภาพในด้านการควบคุมค่าความน่าจะเป็นของผิดพลาดแบบที่ 1 และค่าประมาณกำลังการทดสอบ การแจกแจงของข้อมูลที่ศึกษามี 4 ลักษณะการแจกแจง คือ การแจกแจงใกล้เคียงปรกติ การแจกแจงสมมาตร การแจกแจงไม่สมมาตรที่มีค่าความโด่งมากกว่า 3 และการแจกแจงไม่สมมาตรที่มีค่าความโด่งน้อยกว่า 3 ผลการศึกษาพบว่า สถิติทดสอบทั้ง 4 วิธีสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ เมื่อข้อมูลมีการแจกแจงใกล้เคียงปรกติและการแจกแจงสมมาตรสถิติทดสอบภาวะสารูปดีแบบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงจะมีค่าประมาณกำลังการทดสอบสูงที่สุดในทุกขนาดตัวอย่าง สำหรับข้อมูลที่มีการแจกแจงไม่สมมาตรที่มีค่าความโด่งมากกว่า 3 และน้อยกว่า 3 สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุงจะมีค่าประมาณกำลังการทดสอบสูงที่สุดในทุกขนาดตัวอย่าง

คำสำคัญ : ตัวสถิติทดสอบภาวะสารูปดี อัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง กำลังการทดสอบ

***Corresponding author. E-mail** : jutaporn@buu.ac.th

Abstract

This research presents a goodness of fit test based on modified log-likelihood ratio for normal distribution and compares the efficiency of the proposed test with modified Anderson-Darling test and modified Cramer-von-Mises test which were developed by Zhang (2002), and the original Anderson-Darling tests. The critical values of the tests are obtained through simulation and the efficiency of test is considered as probability of type I error and power of the tests. There are four types of distribution; near normal distribution, symmetric distribution, asymmetric distribution with more than three of kurtosis, and asymmetric distribution with less than three of kurtosis. The results show that the four tests can control type I error probability and the proposed test is the most powerful test for near normal and symmetric distributions with all sizes of sample. The modified Anderson-Darling test is the most powerful test for asymmetric distribution both more than three and less than three of kurtosis with all sizes of sample.

Keywords : goodness of fit test statistic, modified likelihood ratio , power of the test

บทนำ

ปัจจุบันข้อมูลเข้ามามีบทบาทสำคัญในการตัดสินใจดำเนินกิจกรรมต่าง ๆ ด้วยยุคของวิทยาศาสตร์ข้อมูล ดังนั้น การเก็บรวบรวมข้อมูลจริงไม่ว่าจะเป็นจากประชากรหรือจากตัวอย่าง จึงเป็นสิ่งที่ดำเนินการอยู่อย่างแพร่หลายในปัจจุบัน ทั้งภาครัฐและเอกชน การศึกษาข้อมูลตัวอย่างที่เป็นข้อมูลเชิงปริมาณในเบื้องต้นผู้ศึกษามักต้องการทราบถึงลักษณะการกระจายของข้อมูล ซึ่งสามารถดำเนินการได้โดยการสร้างกราฟฮิสโทแกรม หรือกราฟความหนาแน่นความน่าจะเป็น หรือต้องการศึกษาถึงการแจกแจงที่เหมาะสมกับข้อมูลนั้น เช่น การศึกษาการแจกแจงข้อมูลความเร็วลม ปริมาณน้ำฝน ปริมาณการใช้น้ำประปา หรือการศึกษาว่าข้อมูลมีการแจกแจงปรกติหรือไม่ เพื่อนำไปสู่การเลือกใช้สถิติอิงพารามิเตอร์และสถิติไม่อิงพารามิเตอร์ที่ถูกต้องสำหรับการวิเคราะห์ข้อมูลขั้นต่อไป

ในการตรวจสอบว่าข้อมูลมีการแจกแจงตามที่คาดหมายไว้หรือไม่ เราเรียกการทดสอบนั้นว่าเป็นวิธีการทดสอบภาวะสารูปดี (Goodness of fit test) ในการทดสอบว่าข้อมูลมีการแจกแจงแบบปรกติเป็นสิ่งที่สำคัญในการศึกษาเชิงสถิติวิเคราะห์ ตัวสถิติทดสอบที่ใช้ในการทดสอบภาวะสารูปดีแบบดั้งเดิมมีหลายตัวสถิติทดสอบ เช่น ตัวสถิติทดสอบเชพพิโร – วิลค์ (Shapiro Wilk test statistic: SW) ตัวสถิติทดสอบคอลโมโกรอฟ- สมีร์นอฟ (Kolmogorov-Smirnov test statistic: KS) ตัวสถิติทดสอบแอนเดอร์สัน – ดาร์ลิง (Anderson-Darling test statistic: AD) และตัวสถิติทดสอบคาแมร์-ฟอนมิส (Cramer-von Misses test statistic: CVM) เป็นต้น และจากผลงานวิจัยของนักวิจัยหลายท่านพบว่า ตัวสถิติทดสอบเชพพิโร – วิลค์ และ ตัวสถิติทดสอบแอนเดอร์สัน – ดาร์ลิง มีค่ากำลังการทดสอบที่สูงเหมาะสมในการทดสอบภาวะสารูปดีสำหรับการทดสอบการแจกแจงปรกติ (Wasinra, 2006; Wichitchan, 2010; Yap & Sim, 2011; Razali & Wah, 2011; Pinthongpan & Thongteeraparp, 2015)

ตัวสถิติทดสอบภาวะสารูปดีที่ใช้ในการทดสอบการแจกแจงปรกติได้ถูกพัฒนาขึ้นมาใหม่อีกหลายตัวสถิติ ได้แก่ สถิติทดสอบคอลโมโกรอฟ-สมีร์นอฟที่ปรับปรุง (modified Kolmogorov-Smirnov test statistic) สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุง (modified Anderson-Darling test statistic) และสถิติคาแมร์-ฟอนมิสที่ปรับปรุง (modified Cramer-von-Mises test statistic) (Zhang, 2002) และยังมีตัวสถิติทดสอบภาวะสารูปดีที่ถูกพัฒนาขึ้นมาใหม่เพื่อใช้ทดสอบการ

แจกแจงอื่น ๆ ด้วย เช่น ตัวสถิติทดสอบ Liao – Shimokawa ที่ปรับปรุงเพื่อใช้ในการทดสอบการแจกแจงปรกติแบบเบ้ (Potas *et al.,* 2014) ตัวสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ปรับปรุงใหม่ (new modified Anderson-Darling test) เพื่อทดสอบการแจกแจงแบบเบ้ ได้แก่ การแจกแจงแกมมาและการแจกแจงล็อกนอร์มอล (Saethow & Neamvonk, 2014) เป็นต้น ในการทดสอบการแจกแจงปรกติตัวสถิติทดสอบภาวะสารูปดีทั้งแบบดั้งเดิมและตัวสถิติทดสอบที่ถูกพัฒนาขึ้นมาใหม่ได้ถูกนำมาเปรียบเทียบประสิทธิภาพ ซึ่งพบว่าตัวสถิติทดสอบที่ถูกพัฒนาขึ้นมาใหม่ได้แก่ สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุงและสถิติคาแมร์-ฟอนมิสที่ปรับปรุง จะมีกำลังการทดสอบสูงกว่าสถิติแบบดั้งเดิม เมื่อข้อมูลที่ทดสอบมีการแจกแจงแตกต่างจากการแจกแจงปรกติ แต่หากข้อมูลที่ทดสอบมีการแจกแจงใกล้เคียงการแจกแจงปรกติ สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงแบบดั้งเดิมจะมีประสิทธิภาพสูงกว่าสถิติทดสอบที่พัฒนาขึ้นมาใหม่ (Zhang, 2002)

เมื่อพิจารณาถึงที่มาของตัวสถิติทดสอบที่ถูกพัฒนาขึ้นมาใหม่ พบว่าตัวสถิติทั้งสองนี้ถูกพัฒนาขึ้นมาบนพื้นฐานของตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็น (log-likelihood ratio statistic) ซึ่งแตกต่างจากตัวสถิติดั้งเดิมที่ถูกพัฒนาขึ้นมาบนพื้นฐานของตัวสถิติเพียร์สันไคกำลังสอง (Pearson Chi-square statistic) (Cressie & Read, 1984) และยังพบว่ามีตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง (modified log-likelihood ratio statistic) ที่สามารถนำมาพัฒนาเป็นตัวสถิติทดสอบภาวะสารูปดีใหม่ได้ ดังนั้น ผู้วิจัยจึงสนใจศึกษาตัวสถิติทดสอบภาวะสารูปดีแบบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง และเปรียบเทียบกำลังการทดสอบที่นำเสนอใหม่กับสถิติทดสอบแบบดั้งเดิมคือ ตัวสถิติทดสอบแอนเดอร์สัน-ดาร์ลิง และสถิติทดสอบที่ถูกพัฒนาขึ้นมาโดย Zhang (2002) คือตัวสถิติทดสอบคาแมร์-ฟอนมิสที่ปรับปรุง ตัวสถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุง สำหรับการทดสอบการแจกแจงปรกติ

วิธีดำเนินการวิจัย

1. ตัวสถิติทดสอบ

กำหนดให้ X เป็นตัวแปรสุ่มที่มีค่าสถิติอันดับเป็น $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ และมีฟังก์ชันการแจกแจงความน่าจะเป็นคือ $F_0(x)$ ตัวสถิติทดสอบภาวะสารูปสนิทดีถูกสร้างขึ้นจาก

$$Z = \int_{-\infty}^{\infty} Z_x \, dw(x) \tag{1}$$

โดยที่ Z_x เป็นสถิติทดสอบพัฒนามาจากงานวิจัยของ Cressie และ Read (1984) ในกรณีที่มีการเปรียบเทียบฟังก์ชันการแจกแจง 1 กลุ่มคือ ฟังก์ชันการแจกแจงที่ได้จากการทดลอง ($F_n(x)$) กับฟังก์ชันการแจกแจงที่คาดหวัง ($F_0(x)$) วงศ์ของสถิติทดสอบแสดงดังสมการ

$$2nI_x^{\lambda} = \frac{2n}{\lambda(\lambda+1)}\left[F_n(x)\left\{\frac{F_n(x)}{F_0(x)}\right\}^{\lambda} + \{1-F_n(x)\}\left\{\frac{1-F_n(x)}{1-F_0(x)}\right\}^{\lambda} - 1 \right] \tag{2}$$

เมื่อกำหนดให้ค่า $\lambda = 1$ ในสมการ (2) จะได้ตัวสถิติเพียร์สันไคกำลังสอง (Pearson Chi-square statistic)

$$2nI_x^1 = \frac{n\left[F_n(x) - F_0(x)\right]^2}{F_0(x)\left[1 - F_0(x)\right]} \quad (3)$$

และเมื่อกำหนดให้ $\lambda = 0$ จะได้ตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็น (log likelihood ratio statistic)

$$2nI_x^0 = 2n\left[F_n(x)\log\left\{\frac{F_n(x)}{F_0(x)}\right\} + \left\{1 - F_n(x)\right\}\log\left\{\frac{1 - F_n(x)}{1 - F_0(x)}\right\}\right] \quad (4)$$

ซึ่งได้ถูกนำมาพัฒนาตัวสถิติทดสอบภาวะสารูปดีต่าง ๆ ดังนี้

Anderson และ Darling (1954) ได้กำหนดให้ Z_x เป็นตัวสถิติเพียร์สันไคกำลังสองในสมการ (3) และฟังก์ชันถ่วงน้ำหนักเป็น $w(x) = F_0(x)$ จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิง (Anderson-Darling test statistic : AD) ดังนี้

$$AD = -\frac{2}{n}\sum_{i=1}^{n}\left[\left(i - \frac{1}{2}\right)\log\left\{F_0\left(X_{(i)}\right)\right\} + \left(n - i + \frac{1}{2}\right)\log\left\{1 - F_0\left(X_{(i)}\right)\right\}\right] - n \quad (5)$$

Zhang (2002) ได้กำหนด Z_x เป็นตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็นในสมการ (4) และฟังก์ชันถ่วงน้ำหนักเป็น $w(x) = \frac{1}{F_n(x)\{1 - F_n(t)\}}dF_n(x)$ จะได้ตัวสถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุง (Modified Anderson-Darling test statistic : ZAD) ดังนี้

$$ZAD = -\sum_{i=1}^{n}\left[\frac{\log\left\{F_0\left(X_{(i)}\right)\right\}}{n - i + \frac{1}{2}} + \frac{\log\left\{1 - F_0\left(X_{(i)}\right)\right\}}{i - \frac{1}{2}}\right] \quad (6)$$

และ Zhang (2002) ได้กำหนดให้ Z_x ที่เป็นตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็นในสมการ (4) และกำหนดฟังก์ชันถ่วงน้ำหนักเป็น $w(x) = \frac{1}{F_0(x)\{1 - F_0(t)\}}dF_0(x)$ จะได้ตัวสถิติทดสอบคาแมร์-ฟอนมิสที่ปรับปรุง (Modified Cramer-von-Mises test statistic: $ZCVM$) ดังนี้

$$ZCVM \approx \sum_{i=1}^{n}\left[\log\left\{\frac{\frac{1}{F_0\left(X_{(i)}\right)} - 1}{\left(n - \frac{1}{2}\right)\Big/\left(i - \frac{3}{4}\right) - 1}\right\}_i\right]^2 \quad (7)$$

ตัวสถิติทดสอบที่นำเสนอ ถูกพัฒนามาจากแนวคิดข้างต้น โดยตัวสถิติจากสมการ (2) กำหนดให้ $\lambda = -1$ ซึ่งเป็นสถิติอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง (modified log likelihood ratio statistics) (Cressie และ Read, 1984)

$$Z_x = -2n\left[F_0(x)\log\left[\frac{F_n(x)}{F_0(x)}\right] + \{1-F_0(x)\}\log\left[\frac{1-F_n(x)}{1-F_0(x)}\right]\right]$$

กำหนดให้ ฟังก์ชันถ่วงน้ำหนัก $dw(x) = \frac{1}{F_0(x)\{1-F_0(x)\}}dF_0(x)$ เมื่อแทนค่าฟังก์ชันของตัวสถิติและฟังก์ชันถ่วงน้ำหนักลงในสมการ (1) และปรับการปริพันธ์เป็นการหาผลรวมในรูปแบบไม่ต่อเนื่องสำหรับข้อมูลตัวอย่างจะได้สถิติทดสอบภาวะสารูปดีแบบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุง (goodness of fit test based on modified log-likelihood Ratio : $MLAD$) ดังนี้

$$MLAD = -2\sum_{i=1}^{n}\left(\frac{1}{1-F_0(X_{(i)})}\log\left[\frac{i-\frac{1}{2}}{nF_0(X_{(i)})}\right] + \frac{1}{F_0(X_{(i)})}\log\left[\frac{n-i+\frac{1}{2}}{n\left(1-F_0(X_{(i)})\right)}\right]\right) \qquad (6)$$

2. ขอบเขตของการวิจัย

1. สร้างค่าวิกฤติของตัวสถิติทดสอบชนิดต่าง ๆ เพื่อการทดสอบการแจกแจงปรกติ โดยทำการจำลองข้อมูลที่มีการแจกแจงปรกติ ที่มีค่าพารามิเตอร์บอกตำแหน่ง ซึ่งแทนค่าเฉลี่ยของประชากรให้ มีค่าเป็น -10, -5, 0, 5 และ 10 และมีค่าพารามิเตอร์บอกขนาด ซึ่งแทนค่าความแปรปรวนของประชากรให้ มีค่าเป็น 0.25, 0.5, 1, 2 และ 4 จำนวนทั้งสิ้น 25 ชุดพารามิเตอร์

2. ศึกษาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยทำการจำลองข้อมูลที่มีการแจกแจงปรกติ ที่มีค่าพารามิเตอร์ เป็น (4, 1), (0, 9), (-7, 0.81), (1, 0.25), (-4, 0.16), (15, 0.04) และ (-20, 3) ซึ่งพารามิเตอร์ที่กำหนดนี้ จะมีค่าต่างไปจากชุดที่ใช้ในการหาค่าวิกฤติ เพื่อเป็นการทดสอบว่า ค่าวิกฤติที่ได้นั้นมีประสิทธิภาพหรือไม่

3. ศึกษากำลังการทดสอบ โดยทำการจำลองข้อมูลที่มีการแจกแจงต่าง ๆ ตามขอบเขตของสัมประสิทธิ์ค่าความเบ้ (γ_1) และค่าสัมประสิทธิ์ความโด่ง (γ_2) ตามแนวคิดการกำหนดค่าของ Shapiro (1968) (Shapiro, 1968 อ้างถึงใน Rinnakorn & Kamon, 2007) จึงทำการกำหนดการแจกแจงพร้อมทั้งกำหนดค่าพารามิเตอร์ดังตารางที่ 1

ตารางที่ 1 แสดงค่าพารามิเตอร์ สัมประสิทธิ์ความเบ้และสัมประสิทธิ์ความโด่งของการแจกแจงที่สอดคล้องกับประเภทของการแจกแจง

ลักษณะการแจกแจง	**เงื่อนไข**	**การแจกแจงที่เข้ากับเงื่อนไข**
การแจกแจงใกล้เคียงการแจกแจงปรกติ	γ_1 =0 2.5 < γ_2 < 4.5	การแจกแจงที่ t(8) t(10) t(20)
การแจกแจงสมมาตร	γ_1 =0 γ_2 > 4.5	การแจกแจงที่ t(5) t(6) t(7)
การแจกแจงไม่สมมาตร	$\lvert\gamma_1\rvert$ =0.3 γ_2 > 3	กาแจกแจงแกมมา Gamma(1,1) การแจกแจงบีตา Beta(7,2) การแจกแจงล๊อกนอร์มอล LogN(0,0.433342)
	$\lvert\gamma_1\rvert$ =0.3 γ_2 < 3	การแจกแจงแกมมา Gamma(4,1) การแจกแจงบีตา Beta(2,5) การแจกแจงล็อกนอร์มอล LogN(0,0.326613)

4. ในการสร้างข้อมูลแต่ละการแจกแจง จะทำภายใต้ขนาดตัวอย่างเป็น 10, 20, 30, 50,70, 100 และ 200 และกำหนดระดับนัยสำคัญในการทดสอบที่ระดับ 0.05

3. วิธีการศึกษา

ขั้นตอนในการวิจัยแบ่งออกเป็น 3 ขั้นตอน

ขั้นตอนที่ 1 สร้างค่าวิกฤติของตัวสถิติทดสอบ AD , ZAD , $ZCVM$ และ $MLAD$ สำหรับทดสอบการแจกแจงปรกติ

1. สร้างข้อมูลที่มีลักษณะการแจกแจงปรกติที่มีพารามิเตอร์ และขนาดตัวอย่างตามที่กำหนดไว้ในขอบเขตการศึกษา

2. ประมาณค่าพารามิเตอร์$\left(\mu,\sigma^2\right)$ ของการแจกแจงปรกติด้วยวิธีภาวะน่าจะเป็นสูงสุด (maximum likelihood method) จากนั้นนำไปแทนค่าในฟังก์ชันการแจกแจงปรกติ

3. คำนวณค่าสถิติของสถิติทดสอบ AD , ZAD , $ZCVM$ และ $MLAD$

4. ทำซ้ำ ข้อ 1 ถึง ข้อ 3 จำนวน 100,000 รอบ แล้วนำค่าสถิติมาเรียงลำดับจากน้อยไปมาก และหาค่าเปอร์เซ็นไทล์ที่ 95 ซึ่งเป็นค่าวิกฤติที่ระดับนัยสำคัญที่ระดับ 0.05 ในแต่ละตัวสถิติ

5. นำค่าวิกฤติที่ได้ในทุกพารามิเตอร์ที่มีขนาดตัวอย่างเดียวกันมาหาค่าเฉลี่ย เพื่อเป็นตัวแทนค่าวิกฤติในแต่ละขนาดตัวอย่าง และทำซ้ำในทุกขนาดตัวอย่าง

ขั้นตอนที่ 2 การคำนวณค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1

1. สร้างตัวอย่างสุ่มที่มีการแจกแจงปรกติที่มีพารามิเตอร์และขนาดตัวอย่าง ตามที่ได้กำหนดไว้ในขอบเขตการศึกษา ประมาณค่าพารามิเตอร์ μ และσ^2 **ด้วยวิธีภาวะน่าจะเป็นสูงสุด** เพื่อนำไปใช้ในการหาค่าฟังก์ชันสะสมของการแจกแจงปรกติ

2. คำนวณค่าสถิติของตัวสถิติทดสอบ AD , ZAD , $ZCVM$ และ $MLAD$

3. เปรียบเทียบค่าสถิติของตัวทดสอบในข้อ 2 กับค่าวิกฤติที่ได้จากขั้นตอนที่ 1 ที่ขนาดตัวอย่างเดียวกัน แล้วพิจารณาผลการทดสอบสมมุติฐานหลัก

4. ทำซ้ำข้อ 1 ถึงข้อ 3 เป็นจำนวน 10,000 รอบ

5. คำนวณหาค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ($\hat{\alpha}$) ได้โดยการนับ จำนวนครั้งในการปฏิเสธสมมุติฐานหลักหารด้วยจำนวนรอบที่ทำซ้ำ นั่นคือ 10,000 รอบ

6. นำค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ($\hat{\alpha}$) จากข้อ 5 เทียบกับเกณฑ์การพิจารณาความสามารถในการควบคุมค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของ Cochran (1952) นั่นคือ

ถ้าค่า $\hat{\alpha}$ อยู่ในช่วง [0.040 , 0.060] จะถือว่าค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 มีค่าใกล้เคียงกับระดับนัยสำคัญที่กำหนด แสดงว่าตัวสถิติทดสอบสามารถควบคุมค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

ถ้าค่า $\hat{\alpha}$ อยู่นอกช่วง [0.040 , 0.060] จะถือว่าค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 มีค่าไม่ใกล้เคียงกับระดับนัยสำคัญที่กำหนด แสดงว่าตัวสถิติทดสอบไม่สามารถควบคุมค่าความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

ขั้นตอนที่ 3 การคำนวณค่าประมาณกำลังการทดสอบ

1. สร้างตัวอย่างสุ่มที่ไม่ใช่การแจกแจงปรกติ ในที่นี้จะสร้างตัวอย่างสุ่มที่มีการแจกแจงที การแจกแจงบีตา การแจกแจงแกมมาและการแจกแจงล็อกนอร์มอลที่มีพารามิเตอร์และขนาดตัวอย่าง ตามที่ได้กำหนดไว้ในขอบเขตการศึกษา ดังตารางที่ 1 ประมาณค่าพารามิเตอร์ μ และ σ^2 เพื่อนำไปใช้ในการหาค่าฟังก์ชันสะสมของการแจกแจงปรกติ

2. คำนวณค่าสถิติของตัวสถิติทดสอบ AD , ZAD , $ZCVM$ และ $MLAD$

3. เปรียบเทียบค่าสถิติของตัวทดสอบในข้อ 2 กับค่าวิกฤติที่ได้จากขั้นตอนที่ 1 ที่ขนาดตัวอย่างเดียวกัน แล้วพิจารณาผลการทดสอบสมมุติฐานหลัก

4. ทำซ้ำข้อ 1 ถึงข้อ 3 เป็นจำนวน 10,000 รอบ

5. คำนวณหาค่าประมาณกำลังการทดสอบ โดยการนับจำนวนครั้งในการปฏิเสธสมมุติฐานหลักหารด้วยจำนวนรอบที่ทำซ้ำ นั่นคือ 10,000 รอบ

6. นำค่าประมาณกำลังการทดสอบที่ได้จากข้อ 5 ของแต่ละตัวสถิติทดสอบมาเปรียบเทียบกัน โดยตัวสถิติทดสอบที่มีค่ากำลังการทดสอบมากที่สุดจะมีประสิทธิภาพมากที่สุด

ผลการวิจัยและวิจารณ์ผล

1. ค่าวิกฤติของตัวสถิติทดสอบสำหรับทดสอบการแจกแจงปรกติ

จากการศึกษาพบว่า สถิติทดสอบ AD , ZAD , $ZCVM$ และ $MLAD$ ในแต่ละตัวสถิติ เมื่อขนาดตัวอย่างเท่ากัน และค่าพารามิเตอร์ต่างกัน ค่าวิกฤติของสถิติทดสอบจะมีค่าใกล้เคียงกัน แต่เมื่อขนาดตัวอย่างต่างกันค่าวิกฤติของสถิติทดสอบจะมีค่าแตกต่างกัน ซึ่งจะเห็นได้ว่าค่าวิกฤติของสถิติทดสอบไม่ได้ขึ้นอยู่กับพารามิเตอร์ของการแจกแจง แต่ขึ้นอยู่กับขนาดตัวอย่าง ในแต่ละตัวสถิตินำค่าวิกฤติที่ขนาดตัวอย่างเดียวกันแต่พารามิเตอร์ต่างกันมาหาค่าเฉลี่ย เพื่อเป็นตัวแทนค่าวิกฤติที่ขนาดตัวอย่างที่กำหนดจะได้ค่าแสดงดังตารางที่ 2

ตารางที่ 2 ค่าวิกฤติเฉลี่ยของแต่ละตัวสถิติทดสอบ สำหรับทดสอบการแจกแจงปรกติ ที่ระดับนัยสำคัญ 0.05

n	AD	ZAD	$ZCVM$	$MLAD$
10	0.687	3.497	6.646	6.718
20	0.722	3.452	9.163	12.467
30	0.732	3.422	10.665	16.595
50	0.740	3.387	12.577	22.176
70	0.744	3.368	13.833	25.947
100	0.746	3.351	15.145	29.836
200	0.749	3.327	17.683	36.530

2. การควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

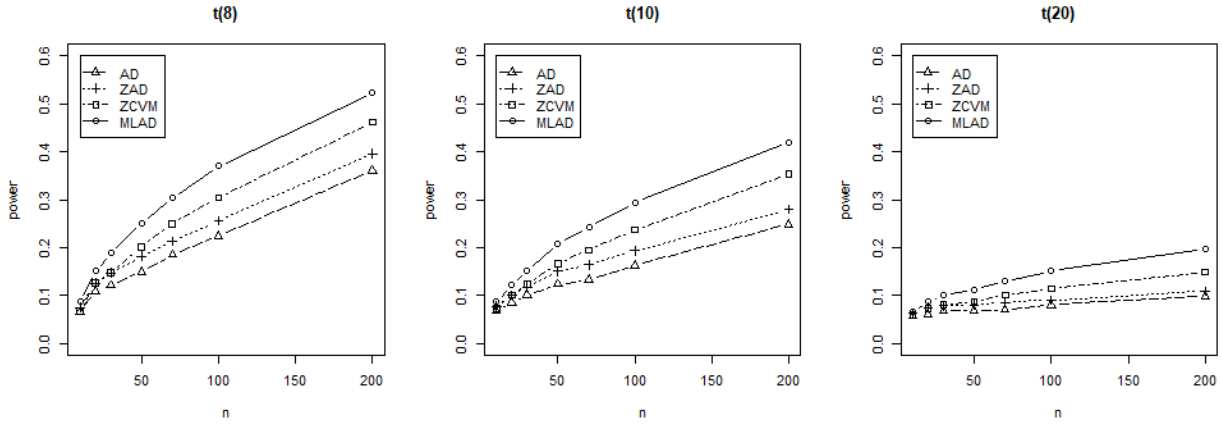
จากการจำลองข้อมูลเพื่อคำนวณหาค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยใช้ค่าวิกฤติที่ได้จากตารางที่ 2 พบว่า ในทุกชุดค่าพารามิเตอร์ที่ใช้ในการทดสอบ ค่าประมาณความน่าจะเป็นที่ได้ทุกตัวมีค่าอยู่ในช่วง [0.040, 0.060] ซึ่งสอดคล้องกับเกณฑ์การพิจารณาความสามารถในการควบคุมความผิดพลาดแบบที่ 1 ของ Cochran (1952) นั่นคือ ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 มีค่าใกล้เคียงกับระดับนัยสำคัญที่กำหนด ดังแสดงในตารางที่ 3

ตารางที่ 3 ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 สำหรับสถิติทดสอบทั้ง 4 แบบ ที่ระดับนัยสำคัญ 0.05

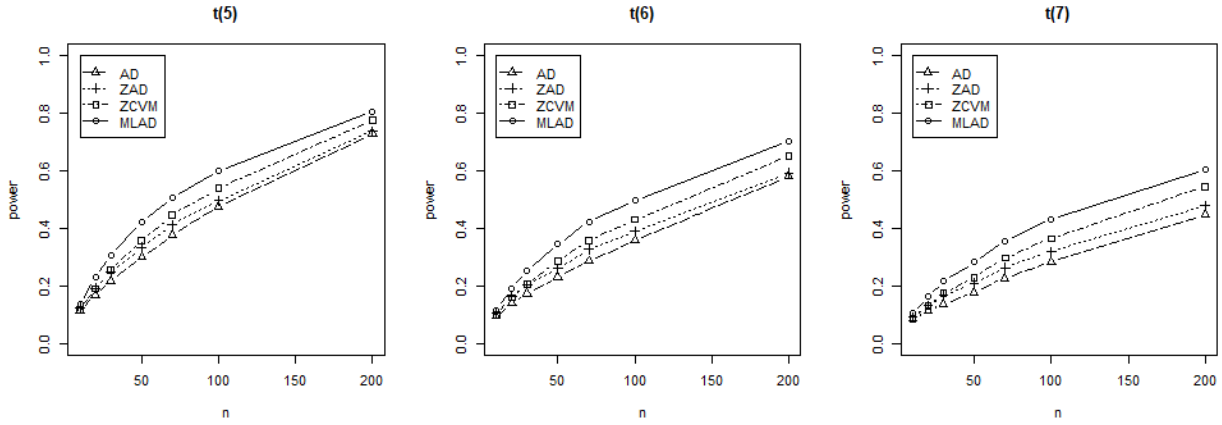
n	AD	ZAD	$ZCVM$	$MLAD$
10	0.050	0.049	0.050	0.049
20	0.050	0.050	0.050	0.050
30	0.050	0.052	0.051	0.051
50	0.050	0.050	0.050	0.050
70	0.051	0.052	0.052	0.051
100	0.050	0.051	0.051	0.050
200	0.049	0.050	0.050	0.049

3. ค่าประมาณกำลังการทดสอบ

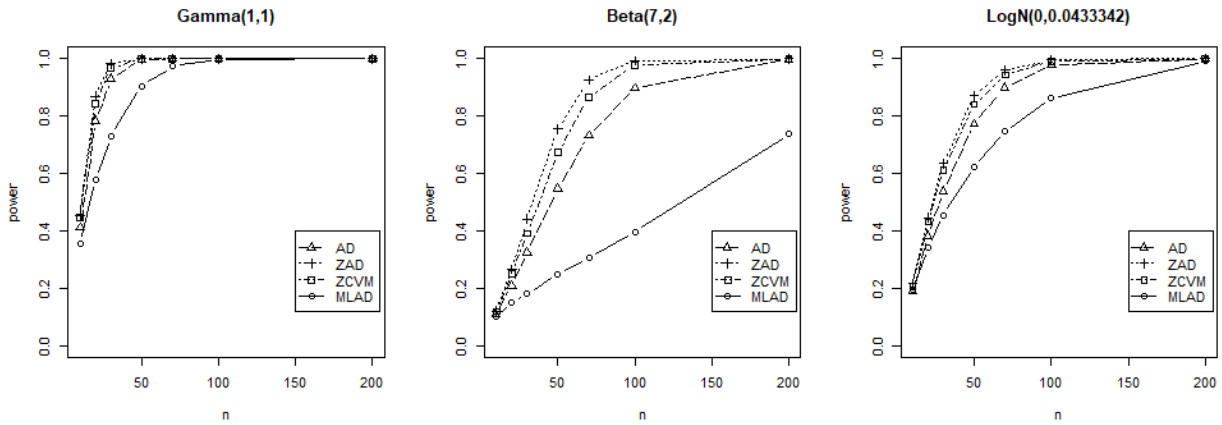
ในการศึกษากำลังการทดสอบ แบ่งการศึกษาออกเป็น 4 กรณี คือ ข้อมูลมีการแจกแจงใกล้เคียงปรกติ ข้อมูลมีการแจกแจงสมมาตร ข้อมูลมีการแจกแจงไม่สมมาตร เมื่อมีค่าความโด่งมากกว่า 3 และข้อมูลมีการแจกแจงไม่สมมาตร เมื่อมีค่าความโด่งน้อยกว่า 3 ผลการศึกษาค่าประมาณกำลังการทดสอบแสดงในภาพที่ 1



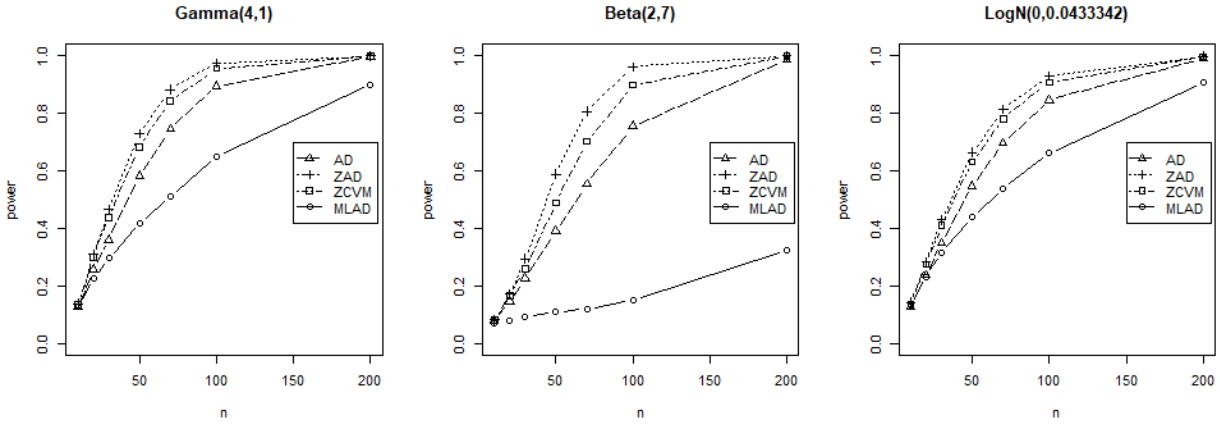
ภาพที่ 1 ค่าประมาณกำลังการทดสอบที่ได้จากการจำลอง เมื่อข้อมูลมีการแจกแจงใกล้เคียงปรกติ



ภาพที่ 2 ค่าประมาณกำลังการทดสอบที่ได้จากการจำลอง เมื่อข้อมูลมีการแจกแจงสมมาตร



ภาพที่ 3 ค่าประมาณกำลังการทดสอบที่ได้จากการจำลอง เมื่อข้อมูลมีการแจกแจงไม่สมมาตรและมีค่าความโด่งมากกว่า 3



ภาพที่ 4 ค่าประมาณกำลังการทดสอบที่ได้จากการจำลอง เมื่อข้อมูลมีการแจกแจงไม่สมมาตรและมีค่าความโด่งน้อยกว่า 3

จากภาพที่ 1 และ ภาพที่ 2 กรณีที่จำลองข้อมูลมีการแจกแจงใกล้เคียงการแจกแจงปรกติ ที่การแจกแจง t(8), t(10) และ t(20) และ กรณีที่ข้อมูลมีการแจกแจงสมมาตร ที่การแจกแจง t(5), t(6) และ t(7) จะพบว่าตัวสถิติทดสอบ $MLAD$ จะมีค่าประมาณกำลังการทดสอบมากที่สุดในทุกขนาดตัวอย่าง รองลงมาคือ ตัวสถิติทดสอบ $ZCVM$ ตัวสถิติทดสอบ ZAD และตัวสถิติทดสอบ AD ตามลำดับ

ในขณะที่ข้อมูลมีการแจกแจงที่ไม่สมมาตร ทั้งที่มีค่าความโด่งมากกว่า 3 และมีค่าความโด่งน้อยกว่า 3 ที่แสดงผลในภาพที่ 3 และภาพที่ 4 พบว่า ตัวสถิติทดสอบ ZAD จะมีค่าประมาณกำลังการทดสอบมากที่สุดในทุกขนาดตัวอย่าง รองลงมาคือ ตัวสถิติทดสอบ $ZCVM$ ตัวสถิติทดสอบ AD และตัวสถิติทดสอบ $MLAD$ ตามลำดับ

เมื่อพิจารณาในทางปฏิบัติ การทดสอบว่าข้อมูลมีการแจกแจงปรกติหรือไม่ ผู้วิจัยมักจะสร้างกราฟฮิสโทแกรมเพื่อตรวจสอบการแจกแจงของข้อมูลเบื้องต้นหากกราฟฮิสโทแกรมแสดงลักษณะการแจกแจงที่ไม่สมมาตรอย่างชัดเจนแล้ว ผู้วิจัยอาจจะไม่จำเป็นต้องทดสอบอีกก็ได้ หากลักษณะของข้อมูลมีการแจกแจงที่สมมาตร หรือใกล้เคียงปรกติ ผู้วิจัยอาจจะไม่สามารถสรุปได้อย่างชัดเจนว่าข้อมูลที่ได้นี้มีการแจกแจงปรกติหรือไม่ ดังนั้นสถิติทดสอบภาวะสารูปดีแบบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงจึงเป็นสถิติที่มีประสิทธิภาพอีกวิธีหนึ่ง

สรุปผลการวิจัย

ในการศึกษาเปรียบเทียบประสิทธิภาพของสถิติทดสอบภาวะสารูปดี 4 แบบ ในการทดสอบการแจกแจงปรกติ พบว่าตัวสถิติทดสอบภาวะสารูปดีแบบอัตราส่วนล็อกภาวะน่าจะเป็นที่ปรับปรุงจะมีค่าประมาณกำลังการทดสอบสูงที่สุด เมื่อข้อมูลมีการแจกแจงใกล้เคียงการแจกแจงปรกติ และข้อมูลมีการแจกแจงแบบสมมาตร สำหรับตัวสถิติทดสอบแอนเดอร์สัน-ดาร์ลิงที่ปรับปรุงจะมีค่ากำลังการทดสอบสูงที่สุดเมื่อข้อมูลมีการแจกแจงไม่สมมาตร

เอกสารอ้างอิง

Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, *49*(268), 765-769.

Cochran, W. G. (1952). The X^2 test of goodness of fit. *Annals of Mathematical Statistics, 23*, 315-345.

Cressie, N., & Read, R.C.T. (1984). Multinomial goodness-of-fit tests. *The Journal of the Royal Statistical Society. Series B (Methodological), 46*(3), 440-464.

Pinthongpan, A. & Thongteeraparp, A. (2015). Comparison of the Goodness of Fit Tests for Normality. *Academic Journal Pranakorn Rajabhat University, 6*(1), 240-249. (in Thai)

Potas, N., Sarısoy, E.E. & Kara, M. (2014). A modified goodness-of-fit test based on likelihood ratiofor the skew-normal distribution. *Applied Mathematical Sciences, 8*, 3869 – 3887.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson–Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

Rinnakorn, C., & Kamon, B. (2007). A Power Comparison of Goodness-of-fit Tests for Normality Based on the Likelihood Ratio and the Non-likelihood Ratio. *Thailand Statistician, 5*, 57-68.

Saethow, P. & Neamvonk, J. (2014). Modified Anderson-Darling test based on the likelihood ratio for skewed distributions. In *Proceedings of the International Conference on Applied Statistics 2014*, (pp. 27-34), KhonKaen, Thailand.

Wasinra, S. (2006). Comparison of the test for Normality using goodness of fit test based on the likelihood ratio. Master thesis in Statistics, Kasetsart University. (in Thai)

Wichitchan, S. (2010). Effects of statistical analysis under normal distribution assumption on data from logistic distribution. Master thesis in Statistics, Chulalongkorn University. (in Thai)

Yap, B. W. & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation, 81*(12), 2141–2155.

Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society. Series B (Methodological), 64*(2), 281-294.