

การทดสอบภาวะสารูปดีสำหรับการแจกแจงແປປາ 4 พารามิเตอร์

The Goodness-of-Fit Test for Four - Parameter Kappa Distribution

เบญจวรรณ รัตนวงศ์ คณิงนิจ ประคำมินทร์ และ นิภาดา พากัดดี*

Benjawan Rattanawong, Khanuengnij Prakhamin and Nipada Papukdee *

Received: 12 July 2021, Revised: 5 July 2022, Accepted: 16 August 2022

บทคัดย่อ

ในการศึกษาครั้งนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติการทดสอบภาวะสารูปดีสำหรับการแจกแจงແປປາ 4 พารามิเตอร์ โดยสถิติทดสอบที่ใช้ในการศึกษามีทั้งหมด 6 ตัว ได้แก่ สถิติทดสอบคอลโมโกรอฟ-สมิรโนฟ (Kolmogorov-Smirnov, KS) สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง (Anderson-Darling, AD) สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งปรับปรุง (Modified Anderson-Darling, B^2) สถิติทดสอบครามเมอร์-วอน-มิสส์ (Cramer-von Mises, CVM) สถิติทดสอบซาง-แอนเดอร์สัน-ดาร์ลิ่ง (Zhang Anderson-Darling, ZAD) และสถิติทดสอบซาง-ครามเมอร์-วอน-มิสส์ (Zhang Cramer von-Mises, $ZCVM$) สำหรับการแจกแจงແປປາ 4 พารามิเตอร์ (Four-parameter kappa distribution : K4D) สำหรับการหาค่าประมาณกำลังการทดสอบ การแจกแจงของข้อมูลที่ศึกษา มี 2 ลักษณะการแจกแจง คือ การแจกแจงสมมาตรและการแจกแจงไม่สมมาตร ได้ใช้หลักการของมอนติคาร์โล (Monte-Carlo Method) ในการจำลองสถานการณ์ข้อมูล โดยทำซ้ำ 10,000 รอบ ผลศึกษาพบว่า เมื่อข้อมูลมีการแจกแจงแบบสมมาตรและไม่สมมาตร สถิติทดสอบที่ให้กำลังการทดสอบสูงที่สุดคือ สถิติทดสอบ ZAD

คำสำคัญ: การแจกแจงແປປາ 4 พารามิเตอร์, การทดสอบภาวะสารูปดี, กำลังการทดสอบ

สาขาวิชาสถิติประยุกต์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตขอนแก่น อำเภอเมือง จังหวัดขอนแก่น 40000
Applied Statistics Department, Faculty of Engineering, Rajamangala University of Technology Isan, Khon Kaen Campus, Muang, Khon Kaen 40000, Thailand.

* ผู้รับผิดชอบประสานงาน ไปรษณีย์อิเล็กทรอนิกส์ (Corresponding author, e-mail): nipada.pa@rmu.ac.th Tel: 08 8552 1606

ABSTRACT

The objective of this study was to investigate the comparison of the power of test statistic by a goodness of fit test for Four Parameter Kappa Distribution. The test with 6 types of goodness-of-fit-test include Kolmogorov-Smirnov (KS), Anderson-Darling (AD), Modified Anderson-Darling (B^2), Cramer-von Mises (CVM), Zhang Anderson-Darling (ZAD) and Zhang Cramer von-Mises ($ZCVM$). For estimating function, the power of test statistics with two distributions was symmetry distribution and asymmetric distribution. However, the researchers were applying Monte-Carlo method for simulation with 10,000 replicated. The results show that, the most powerful test statistics are ZAD , $ZCVM$ and AD for symmetric distribution, and the ZAD is the most powerful test for asymmetric distribution.

Key words: four - parameter kappa distribution : K4D, goodness-of-fit-test, power of test

บทนำ

ทฤษฎีค่าสุดขีด เป็นสถิติที่มุ่งเน้นไปที่การวิเคราะห์เหตุการณ์สุดขีด จากหนังสือของ Coles (2001) ได้กล่าวว่า ทฤษฎีค่าสุดขีดอธิบายพฤติกรรมของการแจกแจงที่ระดับสูงสุดหรือต่ำสุด อันที่จริงแล้วทฤษฎีค่าสุดขีดได้ถูกประยุกต์ใช้ในสาขาต่าง ๆ สำหรับการสร้างแบบจำลองเพื่อทำนายเหตุการณ์รุนแรงซึ่งอาจมีผลกระทบที่ไม่พึงประสงค์สูง การแจกแจงค่าสุดขีดน้อยทั่วไป (Generalized Extreme Value) เป็นแบบจำลองหนึ่งในทฤษฎีค่าสุดขีดซึ่งประกอบด้วยพารามิเตอร์ 3 ชนิด คือ พารามิเตอร์ระบุตำแหน่ง พารามิเตอร์ระบุขนาด และพารามิเตอร์ระบุรูปร่าง ในการเลือกใช้แบบจำลองที่ดีและเหมาะสมนั้นก็จะนำไปสู่ความแม่นยำในการพยากรณ์ค่าต่าง ๆ ที่สนใจศึกษา หากจำนวนพารามิเตอร์ของแบบจำลองนั้นยิ่งเพิ่มมากขึ้นก็อาจทำให้แบบจำลองนั้นมีความแม่นยำขึ้นได้ การแจกแจงแบบ 4 พารามิเตอร์ (Four - parameter kappa distribution; K4D) จะประกอบไปด้วยพารามิเตอร์ระบุตำแหน่ง พารามิเตอร์ระบุขนาด และพารามิเตอร์ระบุรูปร่างตัวที่ 1 และตัวที่ 2 และเป็น

การแจกแจงที่มีความสัมพันธ์กับการแจกแจงค่าสุดขีดน้อยทั่วไป เมื่อพารามิเตอร์ระบุรูปร่างตัวที่ 2 มีค่าเข้าใกล้ 0 จะลดรูปเป็นการแจกแจงค่าสุดขีดน้อยทั่วไป นอกเหนือจากนั้นการพิจารณาแบบจำลองที่ดีที่สุดหรือเหมาะสมที่สุดสำหรับข้อมูลที่มีค่าสุดขีดอีกประการหนึ่ง คือ การทดสอบภาวะสารูปดี (Goodness-of-fit Test) ของแบบจำลอง โดยสามารถพิจารณาจากกราฟ และพิจารณาจากสถิติทดสอบภาวะสารูปดีร่วมกัน การทดสอบภาวะสารูปดีเป็นวิธีที่ใช้ในการทดสอบระหว่างแบบจำลองทางสถิติและค่าที่สังเกตได้ ซึ่งการทดสอบภาวะสารูปดีได้ถูกนำไปใช้กันอย่างแพร่หลายเพื่อตรวจสอบแบบจำลองที่ถูกพัฒนาหรือสร้างขึ้นมา

Liao and Shimokawa (1999) ได้พัฒนาตัวสถิติขึ้นมาใหม่เรียกว่า Liao-Shimokawa (L_n) พร้อมทั้งเปรียบเทียบกำลังการทดสอบของสถิติทดสอบคอลโมโกรอฟ-สมิรโนฟ (Kolmogorov-Smirnov, KS) สถิติทดสอบแอนเดอร์สัน-ดาร์ลิง (Anderson-Darling, AD) และสถิติทดสอบครามเมอร์-วอน-มิสส์ (Cramer-vonMises, CVM) สำหรับการแจกแจงแบบไวบูล พบว่า สถิติทดสอบ L_n มีค่าประมาณ

กำลังการทดสอบสูงสุด ต่อจากนั้น Zempléni (2004) ได้เสนอการทดสอบที่ปรับปรุงมาจากสถิติทดสอบ AD เรียกว่า สถิติทดสอบแอนเดอร์สัน-ดาร์ลิงปรับปรุง (Modified Anderson-Darling, B^2) เพื่อทดสอบภาวะสารถูปดีสำหรับค่าสุดขีดวงนัยทั่วไป เมื่อประมาณค่าพารามิเตอร์ด้วยวิธีการประมาณค่าภาวะสูงสุด พบว่า สถิติทดสอบ B^2 ให้ค่าประมาณกำลังการทดสอบสูงกว่าสถิติทดสอบ AD เนื่องจากมีความไวต่อการตรวจจับข้อมูลที่มีการแจกแจงที่เบ้เข้ามาสูงกว่าสถิติทดสอบ AD และในปีเดียวกัน Laio (2004) ได้เปรียบเทียบประสิทธิภาพของสถิติทดสอบ AD และสถิติทดสอบ CVM สำหรับการแจกแจงค่าสุดขีด พบว่า สถิติทดสอบ AD มีประสิทธิภาพดีกว่าสถิติทดสอบ CVM นอกจากนี้ Zhang (2002); Zhang and Wu (2005) ได้พัฒนาสถิติทดสอบ B^2 สถิติทดสอบ CVM และสถิติทดสอบ KS ขึ้นมาใหม่ ซึ่งเป็นที่รู้จักกันในนามของสถิติทดสอบซาง-แอนเดอร์สัน-ดาร์ลิง (Zhang Anderson-Darling, ZAD) สถิติทดสอบซาง-ครามเมอร์-วอน-มิสส์ (Zhang Cramer von-Mises, $ZCVM$) และสถิติทดสอบซาง-คอลโมโกรอฟ-สมิร์นอฟ (Zhang Kolmogorov Smirnov KS) ซึ่งเป็นอีกทางเลือกของการทดสอบภาวะสารถูปดี พบว่า สถิติทดสอบ ZAD และสถิติทดสอบ $ZCVM$ เป็นสถิติทดสอบที่มีค่าประมาณกำลังการทดสอบมากกว่าสถิติทดสอบ AD สถิติทดสอบ CVM และสถิติทดสอบ ZKS สำหรับข้อมูลที่มีการแจกแจงปกติทั้งนี้สถิติทดสอบ ZAD และสถิติทดสอบ $ZCVM$ มีประสิทธิภาพเมื่อข้อมูลมีการแจกแจงปกติ Shabri and Jemain (2008) ได้ศึกษาการทดสอบภาวะสารถูปดีแบบเดิมหลายชนิด ได้แก่ สถิติทดสอบ AD สถิติทดสอบ CVM สถิติทดสอบ KS และสถิติทดสอบ Shapiro-Wilk test (w) พบว่า สถิติทดสอบ AD มีค่าประมาณกำลังการทดสอบสูงสุด สำหรับข้อมูลที่มีการแจกแจงค่าสุดขีด

นัยทั่วไป และการแจกแจงกัมเบล Abidin *et al.* (2012) ได้ศึกษาค่าประมาณกำลังการทดสอบของสถิติทดสอบ AD สถิติทดสอบ B^2 สถิติทดสอบ CVM สถิติทดสอบ ZAD สถิติทดสอบ $ZCVM$ และสถิติทดสอบ L_n สำหรับการแจกแจงแบบกัมเบล พบว่า สถิติทดสอบ AD มีกำลังการทดสอบสูงสุด ดังนั้นจากงานวิจัยที่ผ่านมา ยังไม่มีการวิจัยเพื่อศึกษาและเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบภาวะสารถูปดีสำหรับการแจกแจงแบบกัมเบล พบว่า 4 พารามิเตอร์ ผู้วิจัยจึงสนใจที่จะศึกษาการเปรียบเทียบกำลังการทดสอบของสถิติทดสอบที่ใช้ในการทดสอบภาวะสารถูปดีสำหรับการแจกแจงแบบกัมเบล 4 พารามิเตอร์ ซึ่งเป็นการแจกแจงใหม่ที่ Hosking (1994) ได้พัฒนาขึ้น การแจกแจงนี้สามารถใช้กับข้อมูลทางด้านอุทกวิทยาได้ดีกว่าการแจกแจงอื่น ๆ โดยสถิติทดสอบที่จะใช้ในการทดสอบภาวะสารถูปดีมีทั้งหมด 6 ชนิด ได้แก่ สถิติทดสอบ KS สถิติทดสอบ AD สถิติทดสอบ B^2 สถิติทดสอบ CVM สถิติทดสอบ ZAD และสถิติทดสอบ $ZCVM$ ซึ่งผู้วิจัย พบว่า ยังไม่มีงานวิจัยใดที่เสนอการทดสอบภาวะสารถูปดีสำหรับการแจกแจงแบบกัมเบล 4 พารามิเตอร์ ด้วยสถิติทดสอบทั้ง 6 ตัวดังกล่าว ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาประเด็นดังกล่าว เพื่อให้เกิดประโยชน์สูงสุดในการเลือกแบบจำลองการแจกแจงที่เหมาะสมกับข้อมูลในทางอุทกวิทยาต่อไป

วิธีดำเนินการวิจัย

1. ตัวสถิติทดสอบ

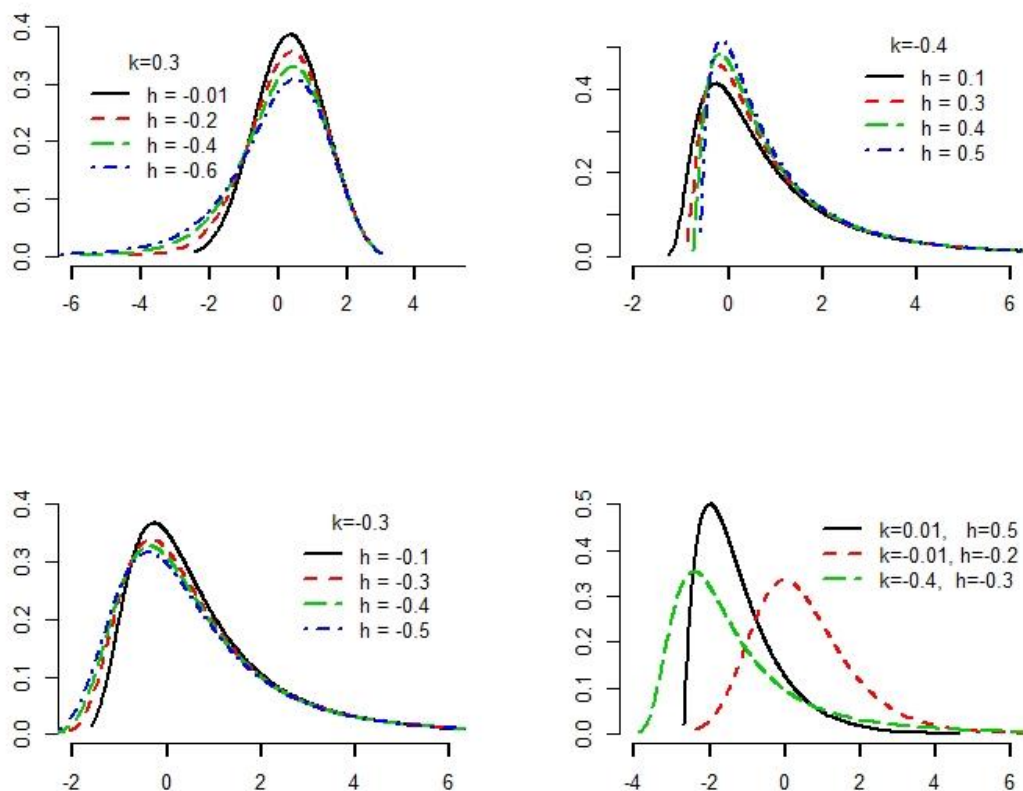
1) ทฤษฎีการแจกแจงแบบกัมเบล 4 พารามิเตอร์
ทฤษฎีการแจกแจงแบบกัมเบล 4 พารามิเตอร์ ($K4D$) ถูกพัฒนาขึ้น โดย Hosking (1994) โดยมีฟังก์ชันการแจกแจงสะสม (Cumulative distribution function; CDF) ดังสมการ (1)

$$F(x) = \left\{ 1 - h \left[1 - \frac{k}{\delta} (x - \omega) \right]^{1/k} \right\}^{1/h}, \quad k \neq 0; h \neq 0 \quad (1)$$

เมื่อ ω แทน พารามิเตอร์ระบุตำแหน่ง (location parameter)

δ แทน พารามิเตอร์ระบุขนาด (scale parameter)

k, h แทน พารามิเตอร์ระบุรูปร่าง (shape parameter)



ภาพที่ 1 ฟังก์ชันความน่าจะเป็นของการแจกแจงแบบลา 4 พารามิเตอร์ เมื่อพารามิเตอร์รูปร่างเปลี่ยนแปลง

2) สถิติการทดสอบภาวะสารูปดีที่ใช้ในการวิจัย

2.1) คอลโมโกรอฟ-สมิรโนฟ (Kolmogorov-Smirnov, KS)

(KS)

สถิติคอลโมโกรอฟ-สมิรโนฟ คือ ค่าที่มากที่สุดของค่าสัมบูรณ์ของผลต่างระหว่างความน่าจะเป็นสะสมที่ได้จากการสังเกต และความน่าจะเป็นสะสมที่ได้ในเชิงทฤษฎีที่กำหนดข้อตกลงเบื้องต้น คือกลุ่มตัวอย่างที่ได้มาโดยการสุ่มและเป็นอิสระกันที่ไม่ทราบฟังก์ชันของการแจกแจง $G(x)$

ถ้ากำหนดให้ $F(x)$ แทนฟังก์ชันของการแจกแจงชนิดหนึ่งแล้วจะสามารถเขียนสมมติฐานของการทดสอบได้ดังนี้

$$H_0: G(X) = F(X) \quad \text{สำหรับค่าของ } X$$

$$\text{คู่กับ} \quad H_1: G(X) \neq F(X) \quad \text{สำหรับค่าของ } X$$

อย่างน้อยที่สุด 1 ค่าการคำนวณสถิติทดสอบ คือ จำนวนการแจกแจงความถี่สะสมของตัวอย่างให้แทนด้วย $s(x)$ แล้วเปรียบเทียบกับค่าการแจกแจงของ

ประชากรที่เป็นสมมติฐานของการทดสอบแทนด้วย $F(x)$ สำหรับแต่ละค่าของ x แล้วหาความแตกต่างสัมบูรณ์ของการแจกแจงทั้ง 2 นี้ คือ $|S(x) - F(x)|$ ค่าสถิติทดสอบ K คือ ค่าที่มากที่สุดของแตกต่างนี้ ถ้าค่า F เข้าใกล้ 0 แสดงว่าการแจกแจงของตัวอย่าง

ใกล้เคียงกันกับการแจกแจงที่กำหนด ทำให้สรุปได้ว่ายอมรับ H_0 ถ้าค่าสถิติทดสอบ K มีค่ามากพอจะสรุปว่าการแจกแจงของตัวอย่างแตกต่างจากการแจกแจงที่กำหนดในสมมติฐาน อย่างน้อยที่สุด 1 จุด ทำให้สรุปได้ว่าปฏิเสธ H_0 ดังสมการ (2)

$$K = \max |S(X) - F(Z_x)| \quad (2)$$

2.2) แอนเดอร์สัน-ดาร์ลิ่ง (Anderson-Darling, AD)

สถิติแอนเดอร์สัน-ดาร์ลิ่ง เป็นสถิติที่ใช้ทดสอบภาวะสภาวะปกติของการแจกแจง เกิดขึ้นในปี

ค.ศ. 1954 แอนเดอร์สัน-ดาร์ลิ่ง ได้เสนอสถิติทดสอบที่ใช้ทดสอบภาวะสภาวะปกติ เมื่อข้อมูลอยู่ในสเกลเรียงอันดับ (Ordinal Scale) และลักษณะการแจกแจงของข้อมูลเป็นแบบต่อเนื่อง ซึ่งคำนวณได้จากสมการ (3)

$$AD = -\frac{2}{n} \sum_{i=1}^n \left(i - \frac{1}{2}\right) \log\{F_0(X_{(i)})\} + \left(n - i + \frac{1}{2}\right) \log\{1 - F_0(X_{(i)})\} - n \quad (3)$$

2.3) แอนเดอร์สัน-ดาร์ลิ่งปรับปรุง (Modified Anderson-Darling, B^2), (Zempléni, 2004) สามารถคำนวณค่าสถิติทดสอบ ได้จากสมการ (4)

$$B^2 = -\sum_{i=1}^n \frac{2i-1}{n} \{\ln[1 - F(x_{n+1-i})]\} + \frac{n}{2} - \sum_{i=1}^n \frac{F(x_i)}{n} \quad (4)$$

โดยที่ n คือ ขนาดตัวอย่าง

$F(x_i)$ คือ ฟังก์ชันการแจกแจงสะสมของข้อมูลตามลักษณะการแจกแจงที่ศึกษา

2.4) คราเมอร์-วอน-มิสส์ (Cramer-von Mises, CVM)

สามารถคำนวณค่าสถิติทดสอบได้จากสูตร (5)

$$CVM = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i) - \frac{2i-1}{2n}\right]^2 \quad (5)$$

โดยที่ n คือ ขนาดตัวอย่าง

$F(x_i)$ คือ ฟังก์ชันการแจกแจงสะสมของข้อมูลตามลักษณะการแจกแจงที่ศึกษา

2.5) ขาง-แอนเดอร์สัน-ดาร์ลิ่ง (Zhang Anderson-Darling, ZAD), (Zhang, 2002)

ได้กำหนด Z_x เป็นตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็น (log likelihood ratio statistic) ดังสมการ (6)

$$2nI_x^0 = 2n \left[\frac{F_n(x) \log \left\{ \frac{F_n(x)}{F_0(x)} \right\} + (1 - F_n(x)) \log \left\{ \frac{1 - F_n(x)}{1 - F_0(x)} \right\}}{2} \right] \quad (6)$$

และฟังก์ชันถ่วงน้ำหนักเป็น $w(x) = \frac{1}{E_n(x)\{1 - F_n(t)\}}$ จะได้สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่งที่ปรับปรุง

(Zhang Anderson-Darling, ZAD) ดังสมการ (7)

$$ZAD = \sum_{i=1}^n \left[\frac{\log \{F_x(X_{(i)})\}}{n - i + \frac{1}{2}} + \frac{\log \{F_x(X_{(i)})\}}{i - \frac{1}{2}} \right] \quad (7)$$

2.6) ขาง-คราเมอร์-วอน-มิสส์ (Zhang Cramer von-Mises, ZCVM), (Zhang, 2002)

ได้กำหนด Z_x เป็นตัวสถิติอัตราส่วนล็อกภาวะน่าจะเป็น (log likelihood ratio statistic) ในสมการ (3) และฟังก์ชันถ่วงน้ำหนักเป็น $w(x) = \frac{1}{F_0(x)\{1 - F_0(t)\}}$ จะได้สถิติทดสอบคราเมอร์-วอนมิส ที่ปรับปรุง (Zhang

Cramer von-Mises, ZCVM) ดังสมการ (8)

$$ZCVM = \sum_{i=1}^n \left[\log \left\{ \frac{\frac{1}{F_0(X_{(i)})} - 1}{\left(n - \frac{1}{2}\right) / \left(i - \frac{3}{4}\right) - 1} \right\} \right]^2 \quad (8)$$

2. ขอบเขตของการวิจัย

1) สร้างค่าวิกฤตของตัวสถิติทดสอบชนิดต่างๆ เพื่อการทดสอบการแจกแจงแคปป์ 4 พารามิเตอร์ โดยการจำลองข้อมูลที่มีการแจกแจงแคปป์ 4 พารามิเตอร์ ที่มีค่าพารามิเตอร์ระบุตำแหน่ง มีค่าเป็น 0 พารามิเตอร์ระบุขนาดมีค่าเป็น 1 และพารามิเตอร์ระบุรูปร่าง มีค่าเป็น -0.4 ถึง 0.4 และ -1.2 ถึง 1.2 ($\mu = 0$, $\sigma = 1$, $-0.4 < k < 0.4$ และ $-1.2 < h < 1.2$)

2) ศึกษาความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 โดยการจำลองข้อมูลที่มีการแจกแจงแคปป์ 4 พารามิเตอร์

3) ศึกษาค่าประมาณกำลังการทดสอบ โดยการจำลองข้อมูลที่มีการแจกแจงต่าง ๆ ซึ่งแบ่ง

ออกเป็น 2 กลุ่มได้แก่ กลุ่มที่ 1 การแจกแจงสมมาตร ประกอบด้วย 1) การแจกแจงปกติ (Normal distribution, $X \sim N(0,20)$) 2) การแจกแจงโคชี (Cauchy distribution, $X \sim \text{Cauchy}(0,1)$) และ 3) การแจกแจงลอจิสติก (Logistic distribution, $X \sim \text{Logis}(9, 3)$) กลุ่มที่ 2 การแจกแจงไม่สมมาตร ประกอบด้วย 1) การแจกแจงไคกำลังสอง (Chi-square distribution, $X \sim \chi^2(10)$) และ 2) การแจกแจงแบบเลขชี้กำลัง (Exponential distribution, $X \sim \text{Exp}(1)$)

4) ในการสร้างข้อมูลแต่ละการแจกแจง จะทำภายใต้ขนาดตัวอย่าง 15 30 50 และ 100 และกำหนดระดับนัยสำคัญ 0.10 0.05 และ 0.01

3. วิธีการศึกษา

วิธีการศึกษามีขั้นตอนในการดำเนินงานดังนี้

1) สร้างค่าวิกฤตของตัวสถิติทดสอบ KS , AD , B^2 , CVM , ZAD และ $ZCVM$ สำหรับการแจกแจงแคลป 4 พารามิเตอร์

1.1 สร้างข้อมูลที่มีลักษณะการแจกแจงแคลป 4 พารามิเตอร์ ที่มีพารามิเตอร์และขนาดตัวอย่างตามที่กำหนดไว้ในขอบเขตการศึกษา

1.2 คำนวณค่าสถิติของสถิติทดสอบ KS , AD , B^2 , CVM , ZAD และ $ZCVM$

1.3 ทำซ้ำข้อ 1.1) ถึง ข้อ 1.2) จำนวน 10,000 รอบ แล้วนำค่าสถิติมาเรียงลำดับจากน้อยไปมากและหาค่าเปอร์เซ็นต์ไทล์ที่ 90 95 และ 99 ซึ่งเป็นค่าวิกฤตที่ระดับนัยสำคัญ 0.10 0.05 และ 0.01 ในแต่ละตัวสถิติ

2) การคำนวณค่าประมาณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1

2.1 สร้างตัวอย่างสุ่มที่มีการแจกแจงแคลป 4 พารามิเตอร์ ที่มีพารามิเตอร์และขนาดตัวอย่างตามที่กำหนดไว้ในขอบเขตการศึกษา

2.2 คำนวณค่าสถิติของตัวสถิติทดสอบ KS , AD , B^2 , CVM , ZAD และ $ZCVM$

2.3 เปรียบเทียบค่าสถิติของสถิติทดสอบในข้อ 2.2 กับค่าสถิติที่ได้จากขั้นตอนที่ 1 ที่ขนาดตัวอย่างเดียวกัน แล้วพิจารณาผลการทดสอบสมมติฐานหลัก

2.4 ทำซ้ำข้อ 2.1) ถึงข้อ 2.3) เป็นจำนวน 10,000 รอบ

2.5 คำนวณค่าประมาณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (τ) ได้โดยการนับจำนวนครั้งในการปฏิเสธสมมติฐานหลักหารด้วยจำนวนรอบที่ทำซ้ำ นั่นคือ 10,000 รอบ

โดยทั่วไปการทดสอบภาวะสารูปดีจะพบผลการทดสอบที่ได้อาจมีความผิดพลาดเกิดขึ้นได้ 2 ลักษณะ คือ ความผิดพลาดแบบที่ 1 ซึ่งความผิดพลาดที่เกิดจากการปฏิเสธสมมติฐานหลัก เมื่อสมมติฐาน

หลักเป็นจริง ซึ่งความน่าจะเป็นที่จะเกิดความผิดพลาดลักษณะนี้แบบนี้คือ ระดับนัยสำคัญ (α) และความผิดพลาดแบบที่ 2 ซึ่งความผิดพลาดที่เกิดจากการยอมรับสมมติฐานหลัก เมื่อสมมติฐานหลักเป็นเท็จ ซึ่งความน่าจะเป็นที่จะไม่เกิดความผิดพลาดในลักษณะแบบนี้คือ $(1-\beta)$ คือ กำลังของการทดสอบ

2.6 นำค่าประมาณความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 จากข้อ 2.6) เทียบกับเกณฑ์การพิจารณาความสามารถในการควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 (Bradley, 1978) นั่นคือ

ค่า τ อยู่ในช่วง $[0.05, 0.15]$ กรณีกำหนดระดับนัยสำคัญเท่ากับ 0.1

ค่า τ อยู่ในช่วง $[0.025, 0.075]$ กรณีกำหนดระดับนัยสำคัญเท่ากับ 0.05

ค่า τ อยู่ในช่วง $[0.0005, 0.015]$ กรณีกำหนดระดับนัยสำคัญเท่ากับ 0.01

3) การคำนวณค่าประมาณกำลังการทดสอบ

3.1 สร้างตัวอย่างสุ่มที่ไม่ใช่การแจกแจงแคลป 4 พารามิเตอร์ ในที่นี้จะสร้างตัวอย่างสุ่มที่มีการแจกแจง 2 กลุ่ม ได้แก่ กลุ่มที่ 1 การแจกแจงสมมาตร ประกอบด้วย การแจกแจงปกติ การแจกแจงโคชี การแจกแจงลอจิสติก กลุ่มที่ 2 การแจกแจงไม่สมมาตร ประกอบด้วย การแจกแจงไคกำลังสอง และการแจกแจงแบบเลขชี้กำลังที่มีพารามิเตอร์และขนาดตัวอย่างตามที่กำหนดไว้ในขอบเขตการศึกษา

3.2 คำนวณค่าสถิติสถิติทดสอบ KS , AD , B^2 , CVM , ZAD และ $ZCVM$

3.3 เปรียบเทียบค่าสถิติทดสอบในข้อ 3.2) กับค่าวิกฤตที่ได้จากขั้นตอนที่ 1 ที่ขนาดตัวอย่างเดียวกันแล้วพิจารณาผลการทดสอบสมมติฐานหลัก

3.4 ทำซ้ำข้อ 3.1) ถึงข้อ 3.3) เป็นจำนวน 10,000 รอบ

3.5 คำนวณหาค่าประมาณกำลังการทดสอบ โดยการนับจำนวนครั้งในการปฏิเสธสมมติฐานหลัก หาดด้วยจำนวนรอบที่ทำซ้ำ นั่นคือ 10,000 รอบ

3.6 นำค่าประมาณกำลังการทดสอบที่ได้จากข้อ 3.5 ของแต่ละตัวสถิติทดสอบมาเปรียบเทียบกัน โดยสถิติทดสอบที่มีค่ากำลังการทดสอบมากที่สุดจะมีประสิทธิภาพมากที่สุด

ZAD เมื่อระดับนัยสำคัญเท่ากันและขนาดตัวอย่างแตกต่างกัน ค่าวิกฤตของสถิติทดสอบจะแตกต่างกัน นั่นคือ เมื่อขนาดตัวอย่างเพิ่มขึ้นค่าวิกฤตของสถิติทดสอบจะลดลง ส่วนสถิติทดสอบ B^2 และสถิติทดสอบ ZCVM เมื่อระดับนัยสำคัญเท่ากันและขนาดตัวอย่างแตกต่างกัน ค่าวิกฤตของสถิติทดสอบจะแตกต่างกัน นั่นคือ ขนาดตัวอย่างเพิ่มขึ้นค่าวิกฤตของสถิติทดสอบจะเพิ่มขึ้น ดังแสดงในตารางที่ 1

ผลการวิจัยและวิจารณ์ผล

1. ค่าวิกฤตของการทดสอบภาวะสารูปดี

จากการศึกษา พบว่า สถิติทดสอบ KS สถิติทดสอบ AD สถิติทดสอบ CVM และสถิติทดสอบ

ตารางที่ 1 ค่าวิกฤตสำหรับสถิติทดสอบ KS สถิติทดสอบ AD สถิติทดสอบ B^2 สถิติทดสอบ CVM สถิติทดสอบ ZAD และสถิติทดสอบ ZCVM เพื่อใช้ทดสอบการแจกแจงแบบป้า 4 พารามิเตอร์ จำแนกตามระดับนัยสำคัญและขนาดตัวอย่าง

ระดับนัยสำคัญ (α)	ขนาด ตัวอย่าง (n)	สถิติ GOF					
		KS	AD	B^2	CVM	ZAD	ZCVM
0.10	15	0.3052	1.9587	17.9957	0.3494	3.7926	17.7244
	30	0.2181	1.9509	34.1859	0.3484	3.5768	19.8378
	50	0.1731	1.9378	54.1131	0.3481	3.4831	21.1777
	100	0.1206	1.9278	107.5746	0.3459	3.3925	22.9989
0.05	15	0.3382	2.5328	19.1186	0.4634	3.9344	22.0642
	30	0.2422	2.4948	35.6465	0.4621	3.6535	24.2208
	50	0.1923	2.5029	55.8594	0.4624	3.5311	25.7847
	100	0.1337	2.4608	109.8714	0.4562	3.417	27.6992
0.01	15	0.4073	3.8947	21.3659	0.7258	4.3047	33.9511
	30	0.2902	3.8778	38.4631	0.7277	3.8473	36.367
	50	0.2300	3.8937	59.3131	0.7370	3.6495	37.549
	100	0.1586	3.7462	114.3337	0.7123	3.4768	40.1447

2. ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1

จากการจำลองข้อมูลเพื่อคำนวณหา ค่าประมาณความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยใช้ค่าวิกฤตที่ได้จากตารางที่ 1 พบว่า เมื่อ จำแนกตามระดับนัยสำคัญ สามารถแสดงค่าประมาณ ความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของการ ทดสอบโดยสถิติทดสอบแต่ละตัว และเปรียบเทียบกับเกณฑ์ของ Bradley ได้ดังนี้ ที่ระดับนัยสำคัญ $\alpha = 0.10$ เมื่อใช้เกณฑ์ของ Bradley พบว่า เมื่อ $n = 15$ และ $n = 30$ สถิติทดสอบทุกตัวสามารถ ควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ เมื่อ $n = 50$ สถิติทดสอบทุกตัวสามารถควบคุมความคลาดเคลื่อน

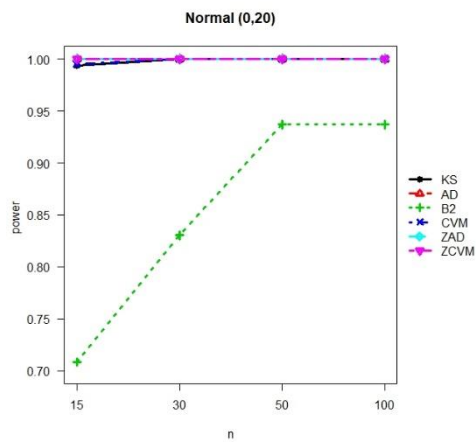
แบบที่ 1 ได้ (ยกเว้นสถิติทดสอบ KS และสถิติ ทดสอบ B^2) และเมื่อ $n = 100$ มีเพียงสถิติทดสอบ KS เท่านั้นที่สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ ที่ระดับนัยสำคัญ $\alpha = 0.05$ เมื่อใช้เกณฑ์ของ Bradley พบว่า สถิติทดสอบทุกตัวไม่สามารถควบคุม ความคลาดเคลื่อนแบบที่ 1 ได้ (ยกเว้น สถิติทดสอบ B^2 เมื่อ $n = 15$ สถิติทดสอบ ZAD สถิติทดสอบ ZCVM เมื่อ $n = 50$ และสถิติทดสอบ AD สถิติ ทดสอบ CVM เมื่อ $n = 100$) ที่ระดับนัยสำคัญ $\alpha = 0.01$ เมื่อใช้เกณฑ์ของ Bradley พบว่า สถิติทดสอบ ทุกตัวไม่สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ ดังแสดงในตารางที่ 2

ตารางที่ 2 ค่าประมาณความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนแบบที่ 1 ของการทดสอบการแจกแจงแบบป 4 พารามิเตอร์ จำแนกตามระดับนัยสำคัญและขนาดตัวอย่าง

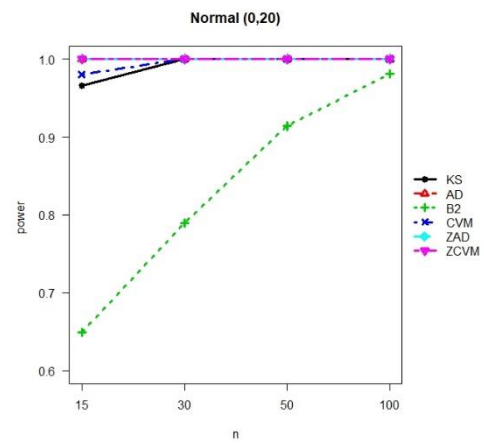
ระดับนัยสำคัญ (α)	ขนาดตัวอย่าง (n)	สถิติ GOF					
		KS	AD	B^2	CVM	ZAD	ZCVM
0.10	15	0.0625*	0.1250*	0.1250*	0.1250*	0.0625*	0.1250*
	30	0.0625*	0.0625*	0.0625*	0.0625*	0.0625*	0.0625*
	50	0.1875	0.1250*	0.1875	0.1250*	0.0625*	0.0625*
	100	0.1250*	0.1875	0.2500	0.1875	0.1875	0.1875
0.05	15	0.0000	0.0000	0.0625*	0.0000	0.0000	0.0000
	30	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	50	0.1250	0.1250	0.1875	0.1250	0.0625*	0.0625*
	100	0.0000	0.0625*	0.1875	0.0625*	0.1875	0.1875
0.01	15	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	30	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	50	0.0625	0.0000	0.0000	0.0000	0.0000	0.0000
	100	0.0000	0.0000	0.0625	0.0000	0.0625	0.0000

หมายเหตุ * คือ สถิติทดสอบที่สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้

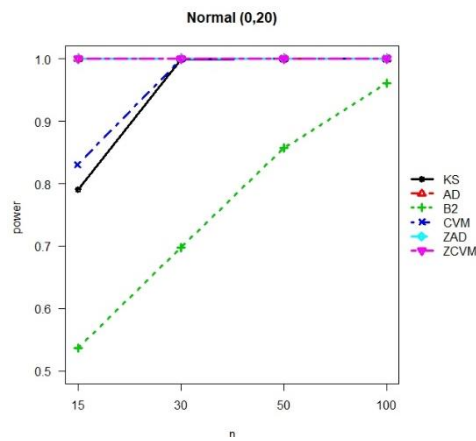
3. ค่าประมาณกำลังการทดสอบ



(ก.) ระดับนัยสำคัญ 0.10

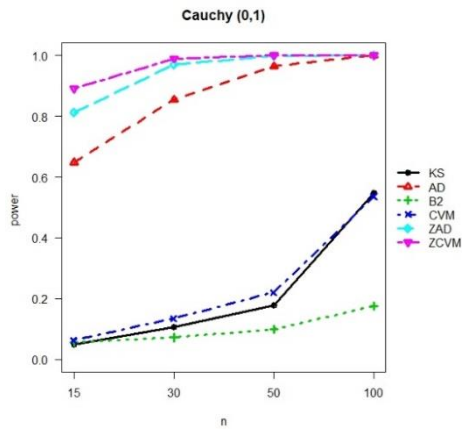


(ข.) ระดับนัยสำคัญ 0.05

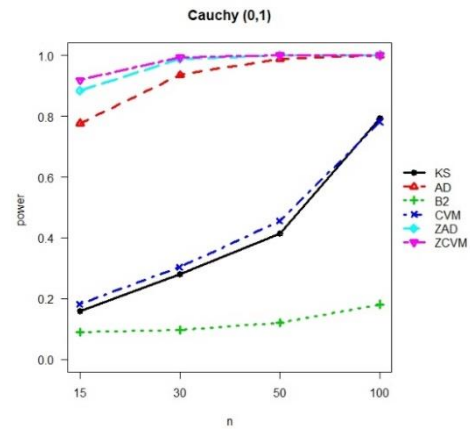


(ค.) ระดับนัยสำคัญ 0.10

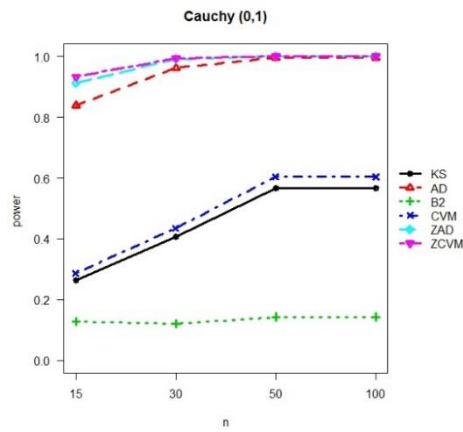
ภาพที่ 2 การเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบแต่ละตัว ที่ระดับนัยสำคัญ α เมื่อกำหนดให้สมมติฐานรอง H_1 เป็นการแจกแจงปกติ



(ก.) ระดับนัยสำคัญ 0.10

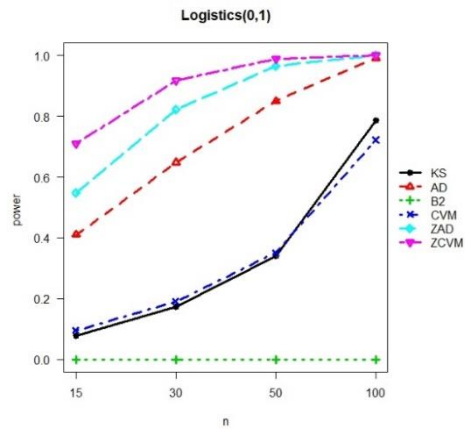


(ข.) ระดับนัยสำคัญ 0.05

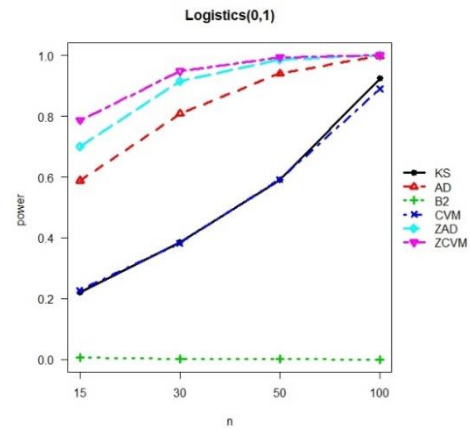


(ค.) ระดับนัยสำคัญ 0.01

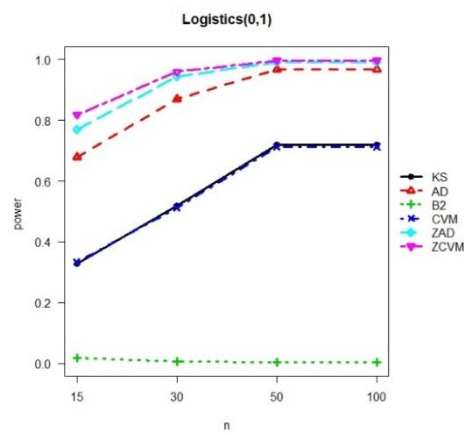
ภาพที่ 3 การเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบแต่ละตัว ที่ระดับนัยสำคัญ α เมื่อกำหนดให้ สมมติฐานรอง H_1 เป็นการแจกแจงโคชี



(ก.) ระดับนัยสำคัญ 0.10

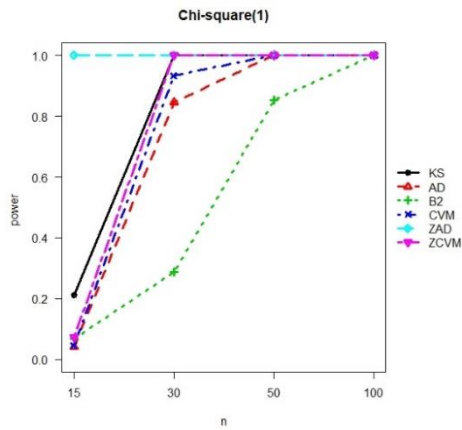


(ข.) ระดับนัยสำคัญ 0.05

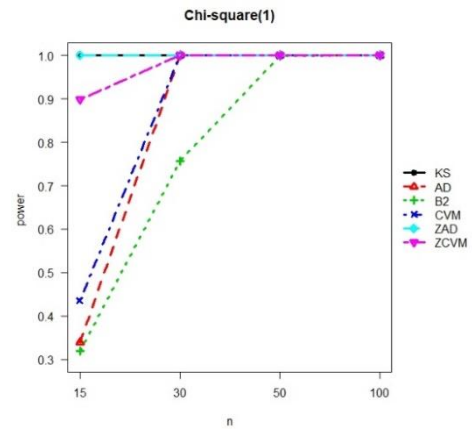


(ค.) ระดับนัยสำคัญ 0.01

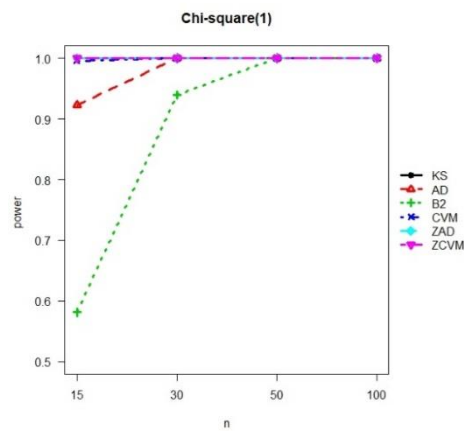
ภาพที่ 4 การเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบแต่ละตัว ที่ระดับนัยสำคัญ α เมื่อ กำหนดให้สมมติฐานรอง H_1 เป็นการแจกแจงลอจิสติก



(ก.) ระดับนัยสำคัญ 0.10

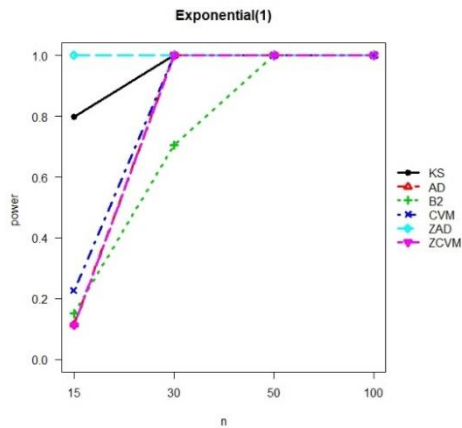


(ข.) ระดับนัยสำคัญ 0.05

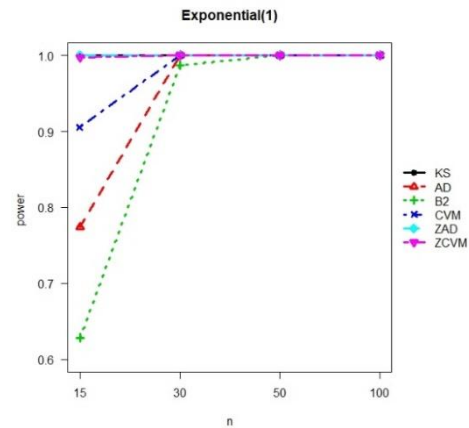


(ค.) ระดับนัยสำคัญ 0.01

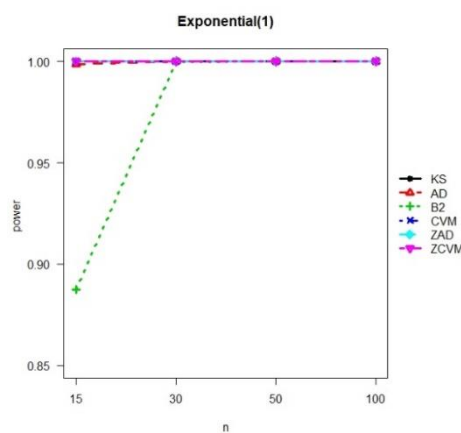
ภาพที่ 5 การเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบแต่ละตัว ที่ระดับนัยสำคัญ α เมื่อกำหนดให้ สมมติฐานรอง H_1 เป็นการแจกแจงไคกำลังสอง



(ก.) ระดับนัยสำคัญ 0.10



(ข.) ระดับนัยสำคัญ 0.05



(ค.) ระดับนัยสำคัญ 0.01

ภาพที่ 6 การเปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบแต่ละตัว ที่ระดับนัยสำคัญ α เมื่อกำหนดให้ สมมติฐานรอง H_1 เป็นการแจกแจงแบบเลขชี้กำลัง

จากภาพที่ 2-6 เปรียบเทียบค่าประมาณกำลังการทดสอบที่ใช้การทดสอบภาวะสารูปดีสำหรับการแจกแจงแบบเลขชี้กำลัง 4 พารามิเตอร์ ทั้งหมด 6 ตัว ได้แก่ สถิติทดสอบ KS สถิติทดสอบ AD สถิติทดสอบ B^2 สถิติทดสอบ CVM สถิติทดสอบ ZAD และสถิติทดสอบ ZCVM สำหรับการแจกแจงแบบเลขชี้กำลัง 4 พารามิเตอร์ เมื่อพิจารณาในภาพรวม พบว่า สถิติทดสอบ ZAD มีค่าประมาณกำลังการทดสอบสูงที่สุดในทุกการแจกแจง รองลงมาคือ สถิติทดสอบ ZCVM สถิติทดสอบ AD สถิติทดสอบ KS และสถิติทดสอบ CVM ซึ่งไม่สอดคล้องกับงานวิจัยของ Shabri and Jemain (2008) ได้ศึกษาการทดสอบภาวะสารูปดีแบบเดิมหลายชนิด ได้แก่ สถิติทดสอบ AD

สถิติทดสอบ CVM สถิติทดสอบ KS และสถิติทดสอบ W พบว่า สถิติทดสอบ AD มีค่าประมาณกำลังการทดสอบสูงที่สุด สำหรับข้อมูลมีการแจกแจงค่าสุดขีดน้อยทั่วไป และการแจกแจงกัมเบล และงานของ Abidin และคณะ ได้ศึกษาค่าประมาณกำลังการทดสอบของสถิติทดสอบ AD สถิติทดสอบ B^2 สถิติทดสอบ CVM สถิติทดสอบ ZAD สถิติทดสอบ ZCVM และสถิติทดสอบ L_n สำหรับการแจกแจงแบบกัมเบล พบว่า สถิติทดสอบ AD มีกำลังการทดสอบสูงที่สุด (Abidin *et al.*, 2012) เนื่องจากการแจกแจงค่าสุดขีดน้อยทั่วไปเป็นการแจกแจงเป็นข้อมูลที่มีการแจกแจงหางหนา (Heavy-tailed Distribution) กว่า การแจกแจงแบบเลขชี้กำลัง 4 พารามิเตอร์ เนื่องจากมี

พารามิเตอร์ระบุตำแหน่ง ขนาด และรูปร่าง ซึ่งมี 3 พารามิเตอร์ ซึ่งเกิดจากการลดรูปร่างของพารามิเตอร์ ระบุรูปร่างตัวที่ 2 ของการแจกแจงแคปป์ 4 พารามิเตอร์ ส่วนการแจกแจงกัมเบลจะประกอบไปด้วยพารามิเตอร์ ระบุขนาดและรูปร่าง ถึงแม้ว่าการแจกแจงทั้งสองเป็น วงศ์ของการแจกแจงแคปป์ 4 พารามิเตอร์ผลการวิจัย อาจจะไม่สามารถคล้อยเพราะการตรวจจับการแจกแจงหาง หนาไม่มาก ควรที่จะเลือกใช้สถิติทดสอบ ZAD เพื่อ เป็นทางเลือกใช้สถิติทดสอบภาวะสารูปดีสำหรับการ แจกแจงแคปป์ 4 พารามิเตอร์ต่อไป และหากพิจารณา ตามลักษณะการแจกแจง พบว่า เมื่อข้อมูลมีการแจก แจกแจงสมมาตร สถิติทดสอบที่ให้ค่าประมาณกำลังการ ทดสอบสูงที่สุดและใกล้เคียงกันคือ สถิติทดสอบ ZAD สถิติทดสอบ ZCVM และสถิติทดสอบ AD ส่วนสถิติ ทดสอบที่ให้กำลังการทดสอบรองลงมาคือ สถิติทดสอบ CVM และสถิติทดสอบ KS ส่วนสถิติ ทดสอบที่ให้กำลังการทดสอบต่ำที่สุด คือ สถิติ ทดสอบ B^2 ซึ่งสอดคล้องกับงานวิจัยของ Zhang (2002); Zhang and Wu (2005) ได้ทำการทดสอบกับ ข้อมูลที่มีการแจกแจงปกติ พบว่า สถิติทดสอบ ZAD และสถิติทดสอบ ZCVM เป็นสถิติทดสอบที่มีกำลัง การทดสอบสูงที่สุด และเมื่อข้อมูลมีการแจกแจงไม่ สมมาตร พบว่า ทุกตัวสถิติทดสอบให้กำลังการทดสอบ สูงใกล้เคียงกัน โดยในกลุ่มนี้สถิติทดสอบ ZAD ให้ กำลังการทดสอบสูงที่สุด ส่วนสถิติที่ให้กำลังการ ทดสอบต่ำที่สุด คือสถิติทดสอบ B^2

สรุป

จากผลการศึกษาการเปรียบเทียบกำลังการ ทดสอบที่ใช้การทดสอบภาวะสารูปดีสำหรับการแจก แจกแจงแคปป์ 4 พารามิเตอร์ ทั้งหมด 6 ตัว ได้แก่ สถิติ ทดสอบ KS สถิติทดสอบ AD สถิติทดสอบ B^2 สถิติ ทดสอบ ZCVM สถิติทดสอบ ZAD และสถิติ ทดสอบ ZCVM สำหรับการแจกแจงแคปป์ 4

พารามิเตอร์ ในการศึกษาค่าประมาณความน่าจะเป็น ความคลาดเคลื่อนแบบที่ 1 พบว่าที่ระดับนัยสำคัญ $\alpha = 0.10$ เมื่อใช้เกณฑ์ของ Bradley พบว่า เมื่อ $n = 15$ และ $n = 30$ ทุกตัวสถิติทดสอบสามารถควบคุม ความคลาดเคลื่อนแบบที่ 1 ได้ เมื่อ $n = 50$ ทุกตัว สถิติทดสอบสามารถควบคุมความคลาดเคลื่อนแบบ ที่ 1 ได้ (ยกเว้นสถิติทดสอบ KS และสถิติทดสอบ B^2) และเมื่อ $n = 100$ มีเพียงสถิติทดสอบ KS เท่านั้น ที่สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ เมื่อ ข้อมูลมีการแจกแจงแบบสมมาตรและไม่สมมาตร พบว่า สถิติทดสอบ ZAD มีค่าประมาณกำลังการ ทดสอบสูงที่สุดในทุกการแจกแจง รองลงมาคือ สถิติ ทดสอบ ZCVM สถิติทดสอบ AD สถิติทดสอบ KS และสถิติทดสอบ CVM เพื่อเป็นแนวทางในการ เลือกใช้สถิติทดสอบภาวะสารูปดีสำหรับการแจก แจกแจงที่มีลักษณะหางหนาบาง ๆ ควรเลือกใช้ สถิติ ทดสอบ AD ถ้าหากการแจกแจงที่มีหางหนาไม่มาก ควรเลือกใช้สถิติทดสอบ ZAD เพื่อใช้ในการ ตรวจสอบแบบจำลองที่ใช้ว่ามีความเหมาะสม หรือ ไม่ ซึ่งเป็นสิ่งที่จำเป็นในการพิจารณา แบบจำลอง

กิตติกรรมประกาศ

งานวิจัยเล่มนี้สำเร็จลุล่วงได้ด้วยความกรุณา จากอาจารย์สาขาวิชาสถิติประยุกต์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขต ขอนแก่น ที่ได้ให้ข้อคิดเห็นและคำแนะนำที่เป็น ประโยชน์ในครั้งนี้ ขอขอบพระคุณคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขต ขอนแก่น ที่ได้ให้ทุนสนับสนุนการวิจัย งบประมาณ รายได้ปีงบประมาณ พ.ศ. 2562 ในครั้งนี้ คุณค่าและ ประโยชน์จากการทำวิจัยในครั้งนี้ ผู้วิจัยขอขอบพระคุณ พระคุณบิดา มารดา บुरพจารย์และผู้มีพระคุณ

ทุกท่านที่มีได้เอ่ยนาม และหากมีข้อบกพร่องด้วยประการใด ผู้วิจัยขอน้อมรับไว้ด้วยความขอบคุณยิ่ง

เอกสารอ้างอิง

- Abidin, N.Z., Adam, M. and Midi, H. 2012. The Goodness-of-fit Test for Gumbel Distribution : A Comparative Study. **MATEMATIKA** 28(1): 35-48.
- Bradley, J.V. 1978. Robustness. **British Journal of Mathematical and Statistical Psychology** 31(2): 144-152.
- Coles, S. 2001. **An Introduction to Statistical Modeling of Extremes Values**. Springer, London.
- Hosking, J.R. 1994. The four-parameter kappa distribution. **IBM Journal of Research and Development** 38(3): 251-258.
- Laio, F. 2004. Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. **Water Resources Research** 40(1): 1-10.
- Liao, M. and Shimokawa, T. 1999. A new goodness-of-fit test for type-1 extreme value and 2-parameter Weibull distributions with estimated parameters. **Journal of Statistical Computation and Simulation** 64(1): 23-48.
- Shabri, A. and Jemain, A.A. 2008. The Anderson-Darling test statistic of the generalized extreme value. **Malaysian Journal of Industrial and Applied Mathematics** 24(1): 85-97.
- Zempléni, A. 2004. Goodness-of-fit Test in Extreme Value Applications. **Collaborative Research Center** 386: 1-13.
- Zhang, J. 2002. Powerful goodness-of-fit tests based on the likelihood ratio. **Journal of the Royal Statistical Society** 64(2): 281-294.
- Zhang, J. and Wu, Y. 2005. Likelihood-ratio tests for normality. **Computational Statistics and Data Analysis** 49(3): 709-721.