

ช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยประชากรของการแจกแจงเรย์ลีย์: กรณีศึกษา ความเร็วคลื่นเฉือนของดินในประเทศไทย

Confidence Interval for the Population Mean of Rayleigh Distribution with Application to Shear wave Velocity of Soils in Thailand

ธนกฤต เตียวพงษ์พันธุ์ และ ภทรวรรณ แสงนวกิจ*

Thanakit Teawpongpan and Patarawan Sangnawakij*

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ประเทศไทย

Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Thailand

บทคัดย่อ

การแจกแจงเรย์ลีย์เป็นการแจกแจงความน่าจะเป็นแบบต่อเนื่องชนิดหนึ่งที่ถูกนำไปใช้ในหลากหลายด้าน โดยเฉพาะทางด้านวิศวกรรม และทางการแพทย์ การแจกแจงนี้ถูกพบบ่อยครั้งในปรากฏการณ์ทางธรรมชาติ เนื่องจากการแจกแจงเรย์ลีย์เกี่ยวข้องกับข้อมูลเชิงปริมาณในการศึกษาครั้งนี้ผู้วิจัยจึงได้นำเสนอวิธีประมาณค่าแบบช่วงสำหรับค่าเฉลี่ยประชากร การสร้างช่วงความเชื่อมั่นในครั้งนี้อาศัยวิธีของวาลด์ วิธีสร้างช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ และวิธีช่วงความเชื่อมั่นน้อยทั่วไป ผู้วิจัยได้ทำการศึกษาประสิทธิภาพของช่วงความเชื่อมั่นโดยอาศัยการจำลองและใช้เกณฑ์การวัดประสิทธิภาพจากค่าความน่าจะเป็นคัมรวมและความยาวเฉลี่ยของช่วงความเชื่อมั่น ผลที่ได้จากการจำลองพบว่าช่วงความเชื่อมั่นน้อยทั่วไปมีประสิทธิภาพที่ดีในทุกสถานการณ์ และดีกว่าช่วงความเชื่อมั่นที่นำมาเปรียบเทียบกับการศึกษาครั้งนี้ผู้วิจัยได้นำวิธีการทางทฤษฎีมาประยุกต์ใช้กับข้อมูลจริงที่เกี่ยวข้องกับความเร็วคลื่นเฉือนของดินในจังหวัดลำพูน ประเทศไทย ทั้งนี้เพื่อเป็นการยืนยันผลที่ได้จากการจำลอง

คำสำคัญ: การประมาณค่าแบบช่วง, การแจกแจงเรย์ลีย์, ความน่าจะเป็นคัมรวม, ช่วงความเชื่อมั่นน้อยทั่วไป

ABSTRACT

Rayleigh distribution is one of the continuous probability distributions applied in several areas, especially in engineering and medicine. This probability model is also widely occurred in natural phenomena. Since Rayleigh distribution is related to quantitative data, we are interested in studying interval estimation for the population mean. The methods based on Wald-type confidence interval, large-sample confidence interval, and generalized confidence intervals are used in construction. Then, the performance of the proposed confidence intervals is conducted using simulations in terms of coverage probability and expected length. The results illustrate that the novel generalized confidence interval performs better than the compared confidence intervals in all cases in the study. Finally, a real data set on the shear wave velocity of soils in Lamphun province of Thailand is used in computation, to confirm the results from simulation.

KEYWORDS: Interval Estimation, Rayleigh Distribution, Coverage Probability, Generalized Confidence Interval

*Corresponding Author: patarawan@mathstat.sci.tu.ac.th

Received: 01/02/2023; Revised: 19/06/2023; Accepted: 14/07/2023

1. บทนำ

การประมาณค่าพารามิเตอร์ (Parameter estimation) เป็นวิธีการทางสถิติที่ใช้ในการศึกษาค่าจริงของประชากร โดยอาศัยข้อมูลจากตัวอย่างซึ่งเป็นส่วนหนึ่งของประชากร ที่สนใจศึกษา การประมาณค่าพารามิเตอร์แบ่งออกเป็น 2 แบบ ได้แก่ การประมาณค่าแบบจุด และการประมาณ

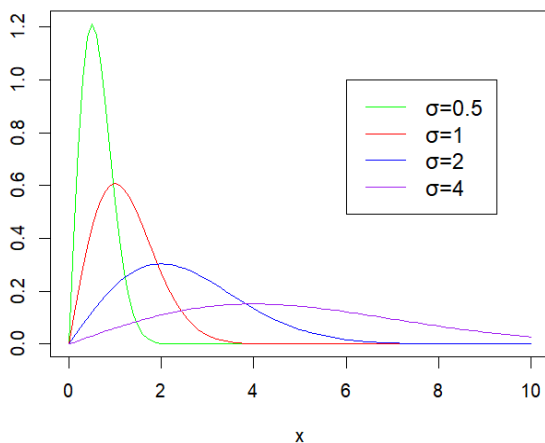
ค่าแบบช่วง โดยการประมาณค่าแบบจุดจะเป็นการประมาณค่าพารามิเตอร์จากค่าสถิติเพียงค่าเดียว ซึ่งต่างจากการประมาณค่าแบบช่วงที่ให้ช่วงค่าสถิติสองค่า อย่างไรก็ตามเนื่องจากการประมาณค่าแบบช่วงสามารถอธิบายระดับความเชื่อมั่น (Confidence level) ที่ค่าสถิติจะประมาณค่าพารามิเตอร์ได้ถูกต้อง (Casella & Berger,

2002) ดังนั้นจึงเป็นวิธีที่สามารถให้ผลสรุปได้มากกว่าการประมาณค่าแบบจุด

การแจกแจงเรย์ลีย์ (Rayleigh distribution) เป็นการแจกแจงชนิดต่อเนื่องที่เกี่ยวข้องกับตัวแปรสุ่มที่มีค่ามากกว่าศูนย์ การแจกแจงนี้ถูกพบบ่อยครั้งในข้อมูลที่เกิดขึ้นจากปรากฏการณ์ทางธรรมชาติ เช่น ความเร็วลม ปริมาณฝุ่นละอองขนาดเล็กในอากาศ และความเร็วคลื่นเฉือนของดินซึ่งเป็นสัญญาณของการเกิดแผ่นดินไหว (Pishgar-Komleh et al., 2015) นอกจากนี้ยังพบได้ในข้อมูลที่เกี่ยวข้องกับสัญญาณภาพถ่ายทางดาวเทียม และสัญญาณภาพถ่ายอัลตราซาวด์ (Ultrasound) (Sarti et al., 2005; Ahn et al., 2012; Nar et al., 2018) สำหรับทฤษฎีความน่าจะเป็นเกี่ยวกับการแจกแจงเรย์ลีย์มีดังนี้ กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงเรย์ลีย์ด้วยพารามิเตอร์กำหนดรูปร่าง (Shape parameter) σ ในที่นี้แทนด้วย $Ra(\sigma)$ แล้ว ฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability density function) ของ X คือ

$$f_X(x; \sigma) = \frac{x}{\sigma^2} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} \quad (1)$$

โดยที่ $x > 0$ และ $\sigma > 0$ ค่าคาดหวังและความแปรปรวนของตัวแปรสุ่ม X คือ $E(X) = \psi = \sigma\sqrt{\pi}/2$ และ $Var(X) = (4-\pi)\sigma^2/2$ ตามลำดับ โดยที่ $\pi \approx 3.142$ (Ateeq et al., 2019) การแจกแจงเรย์ลีย์มีลักษณะเบ้ขวา โดยความเบ้ของข้อมูลจะขึ้นอยู่กับค่าพารามิเตอร์ σ ดังแสดงในรูปที่ 1 จากรูปพบว่าเมื่อ σ มีค่าน้อยข้อมูลจะมีลักษณะเบ้มาก แต่เมื่อ σ มีค่ามากขึ้นความเบ้ของข้อมูลจะลดลง



รูปที่ 1 กราฟความน่าจะเป็นของการแจกแจงเรย์ลีย์ เมื่อกำหนดค่า σ ต่างกัน

จากที่ได้กล่าวไปข้างต้นว่าการแจกแจงเรย์ลีย์ถูกนำไปประยุกต์ใช้ในหลายด้าน ตัวอย่างงานวิจัยที่เกี่ยวข้องมีดังนี้ การสร้างแบบจำลองความเร็วคลื่นเฉือนของดินแบบบูรณาการสำหรับแหล่งก๊าซไกรนิงเกน (Groningen) ในประเทศเนเธอร์แลนด์ เพื่อใช้จำลองความเสี่ยงในการเกิดแผ่นดินไหว (Kruiver et al., 2017) การประมาณค่าพลังงานและความเร็วลม ณ เกาะเซนต์มาร์ติน ประเทศบังกลาเทศ (Islam et al., 2017) และการพยากรณ์ความเร็วลมเพื่อทำนายสภาพอากาศในมณฑลเจียงซู ประเทศจีน (Xu et al., 2020) ในทางทฤษฎีมีนักวิจัยได้นำเสนอวิธีประมาณค่าพารามิเตอร์ในการแจกแจงเรย์ลีย์ดังนี้ Elfattah et al. (2006) ได้ศึกษาผลกระทบของรูปแบบการสุ่มตัวอย่างแบบเซ็นเซอร์ (Censored sampling) ที่มีผลต่อการประมาณค่าพารามิเตอร์ σ โดยอาศัยวิธีภาวะน่าจะเป็นสูงสุด Dey & Dey (2009) ได้ศึกษาวิธีประมาณค่าแบบจุดสำหรับ σ โดยอาศัยวิธีเบย์ส (Bayesian's method) ที่มีการแจกแจงก่อนเป็นฟังก์ชันสูญเสีย (Loss function) และฟังก์ชันความน่าเชื่อถือได้ (Reliability function) Ahmed et al. (2013) ได้ทำการเปรียบเทียบประสิทธิภาพของการประมาณค่าแบบจุดสำหรับ σ ด้วยวิธีประมาณค่าแบบเบย์ที่มีการแจกแจงก่อนเป็นฟังก์ชันสูญเสีย และวิธีภาวะน่าจะเป็นสูงสุด Seo et al. (2016) ได้สร้างช่วงความเชื่อมั่นที่แน่นอน (Exact confidence interval) ในการแจกแจงเรย์ลีย์สำหรับพารามิเตอร์ σ โดยอาศัยปริมาณหมุน (Pivotal quantity) ที่มีการแจกแจงไคกำลังสอง (Chi-square distribution) และ Liebenberg & Allison (2023) ได้ศึกษาวิธีการสร้างสถิติทดสอบภาวะสารูปดี (Goodness of fit test) สำหรับการทดสอบการแจกแจงเรย์ลีย์

จากการทบทวนวรรณกรรมที่เกี่ยวข้องจะพบว่ามีผู้เสนอตัวประมาณค่าพารามิเตอร์แบบจุดและแบบช่วงสำหรับ σ ในการแจกแจงเรย์ลีย์อย่างต่อเนื่อง แต่การประมาณค่าแบบช่วงสำหรับค่าเฉลี่ยประชากร (ψ) ในการแจกแจงเรย์ลีย์นั้น ยังไม่พบว่ามีผู้ที่ทำการศึกษามากนัก โดยเฉพาะอย่างยิ่งการสร้างช่วงความเชื่อมั่นที่อาศัยวิธีของวาลด์จากการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีสร้างช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ที่อาศัยตัวประมาณโมเมนต์ รวมทั้งวิธีการสร้างช่วงความเชื่อมั่นนัยทั่วไปนั้น ยังไม่มีผู้ทำการศึกษา ดังนั้นการนำเสนอช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยประชากรในการแจกแจงเรย์ลีย์ และการศึกษาประสิทธิภาพของช่วงความเชื่อมั่นที่ได้จากการ

จำลองจึงเป็นวัตถุประสงค์หลักของงานวิจัยฉบับนี้ นอกจากการศึกษาทางด้านทฤษฎีแล้วผู้วิจัยได้นำข้อมูลจริงซึ่งเป็นค่าความเร็วคลื่นเฉือนของดินในจังหวัดลำพูนมาใช้ในการคำนวณซึ่งรายละเอียดจะกล่าวถึงในบทประยุกต์ต่อไป

2. วิธีการทางสถิติ

กำหนดให้ $X = (X_1, X_2, \dots, X_n)$ เป็นตัวอย่างสุ่ม (Random sample) ขนาด n จากประชากรที่มีการแจกแจงเรย์ลีย์ด้วยพารามิเตอร์ σ ซึ่งไม่ทราบค่า และมีฟังก์ชันความหนาแน่นความน่าจะเป็นตามสมการ (1) ค่าคาดหวังของ X คือ $\psi = \sigma\sqrt{\pi/2}$ โดยที่ $\sigma > 0$ และ $\pi \approx 3.142$ ในการศึกษาครั้งนี้ ψ คือ พารามิเตอร์ที่สนใจศึกษาและต้องการหาค่าประมาณ

การประมาณค่าแบบจุดของ σ และ ψ โดยอาศัยวิธีโมเมนต์จะพิจารณาโมเมนต์ที่ k รอบจุดกำเนิดของประชากร ($E(X^k)$) และโมเมนต์ที่ k รอบจุดกำเนิดของตัวอย่าง ($\sum_{i=1}^n X_i^k / n$) โดยที่ $k = 1, 2, \dots$ และเนื่องจาก $\sum_{i=1}^n X_i^k / n$ เป็นตัวประมาณไม่เอนเอียงของ $E(X^k)$ ดังนั้น

$$E(X^k) = \frac{1}{n} \sum_{i=1}^n X_i^k$$

เมื่อกำหนดให้ $k = 1$ จะได้ว่า

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{และ} \quad \sigma\sqrt{\pi/2} = \bar{X}$$

กล่าวคือ $\hat{\sigma}_{mm} = \bar{X}\sqrt{2/\pi}$ เป็นตัวประมาณโมเมนต์สำหรับ σ และจะได้ว่าตัวประมาณโมเมนต์ของ ψ คือ $\hat{\psi}_{mm} = \hat{\sigma}_{mm}\sqrt{\pi/2}$ สำหรับวิธีภาวะน่าจะเป็นสูงสุดเป็นวิธีที่ใช้ในการประมาณค่าพารามิเตอร์ที่ทำให้ล็อกของฟังก์ชันภาวะน่าจะเป็น (Log-likelihood function) ของพารามิเตอร์ เมื่อทราบ $(X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n)$ มีค่าสูงสุด และเนื่องจาก X_i มีการแจกแจง $Ra(\sigma)$ โดยที่ $i = 1, 2, \dots, n$ ดังนั้นล็อกของฟังก์ชันภาวะน่าจะเป็นของ σ คือ

$$\ln L(\sigma; x_i) = \ln \prod_{i=1}^n x_i - 2n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n x_i \quad (2)$$

เมื่อหาอนุพันธ์อันดับหนึ่งของ (2) เทียบกับ σ และกำหนดให้เท่ากับศูนย์แล้ว จะได้ตัวประมาณภาวะน่าจะเป็นสูงสุดของ σ คือ $\hat{\sigma}_{MLE} = \sqrt{\sum_{i=1}^n X_i^2 / 2n}$ อาศัยสมบัติไม่แปรเปลี่ยน (Invariance property) ดังนั้นตัว

ประมาณภาวะน่าจะเป็นสูงสุดของ ψ คือ $\hat{\psi}_{MLE} = \hat{\sigma}_{MLE}\sqrt{\pi/2}$

ตัวประมาณค่าเฉลี่ยประชากรที่ได้จากวิธีโมเมนต์ และวิธีภาวะน่าจะเป็นสูงสุดข้างต้นจะถูกนำไปใช้ในการสร้างช่วงความเชื่อมั่นในหัวข้อต่อไป

1. ช่วงความเชื่อมั่นจากวิธีของวาลด์

วิธีของวาลด์เป็นวิธีที่ใช้ในการสร้างช่วงความเชื่อมั่นที่อาศัยสมบัติของตัวประมาณภาวะน่าจะเป็นสูงสุด ถ้าสมมติให้ $\hat{\psi}_{MLE}$ เป็นตัวประมาณภาวะน่าจะเป็นสูงสุดของ $\psi = \sigma\sqrt{\pi/2}$ เมื่อตัวอย่างมีขนาดใหญ่ ($n \rightarrow \infty$) แล้ว $\hat{\psi}_{MLE}$ จะมีการแจกแจงปกติโดยประมาณ (Approximate normal distribution) ด้วยค่าเฉลี่ย ψ และความแปรปรวน $\text{Var}(\hat{\psi})$ ดังนั้นในกรณีของตัวประมาณ $\hat{\psi}_{MLE}$ ที่ได้นำเสนอไปข้างต้นจะหาความแปรปรวนได้จาก

$$\text{Var}(\hat{\psi}_{MLE}) = \text{Var}\left(\sqrt{\frac{\pi}{2}}\hat{\sigma}_{MLE}\right) = \frac{\pi}{2} \text{Var}(\hat{\sigma}_{MLE})$$

โดยที่ $\text{Var}(\hat{\sigma}_{MLE}) = \text{Var}\left(\sqrt{\sum_{i=1}^n X_i^2} / 2n\right)$ เมื่ออาศัยวิธีเดลตา (Delta method) จากนิยาม

$$\text{Var}(g(x)) \cong \text{Var}(X)[g'(x)]^2_{X=E(X)}$$

(Casella & Berger, 2002) ในการประมาณความแปรปรวนของ $g(x) = \sqrt{\sum_{i=1}^n X_i^2}$ จะได้ว่า

$$\text{Var}\left(\sqrt{\sum_{i=1}^n X_i^2}\right) \cong \frac{(4-\pi)n\pi\sigma^2}{8} \quad \text{และ} \quad \text{Var}(\hat{\sigma}_{MLE}) \cong \frac{(4-\pi)\pi\sigma^2}{16}$$

อาศัยการพิสูจน์ข้างต้นทำให้ได้ความแปรปรวนโดยประมาณของ $\hat{\psi}_{MLE}$ ดังนี้

$$\text{Var}(\hat{\psi}_{MLE}) \cong \left(\frac{(4-\pi)\pi^2}{32}\right)\sigma^2 \quad (3)$$

ดังนั้นช่วงความเชื่อมั่น $(1-\alpha) 100\%$ สำหรับ ψ จากวิธีของวาลด์ คือ

$$CI_{NMLE} = (l_{NMLE}, u_{NMLE}) \\ = (\hat{\psi}_{MLE} - z_{1-\alpha/2}\sqrt{\text{var}(\hat{\psi}_{MLE})}, \hat{\psi}_{MLE} + z_{1-\alpha/2}\sqrt{\text{var}(\hat{\psi}_{MLE})})$$

เมื่อ $\text{var}(\hat{\psi}_{MLE})$ คือ ตัวประมาณค่าความแปรปรวนของ $\hat{\psi}_{MLE}$ ซึ่ง $\text{var}(\hat{\psi}_{MLE})$ หาได้จากการแทนค่า σ^2 ในสมการ (3) ด้วย $\hat{\sigma}_{MLE}^2 = \sum_{i=1}^n X_i^2 / 2n$ และ $z_{1-\alpha/2}$ คือเปอร์เซ็นต์ไทล์ที่ $(1-\alpha/2) \times 100$ ของการแจกแจงปกติมาตรฐาน (Standard normal distribution)

2. ช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ที่อาศัย $\hat{\psi}_{mm}$

การสร้างช่วงความเชื่อมั่นในหัวข้อนี้อาศัยทฤษฎีขีดจำกัดส่วนกลาง (Central limit theorem) เนื่องจาก $\hat{\psi}_{mm} = \hat{\sigma}_{mm} \sqrt{\frac{\pi}{2}} = \bar{X}$ ดังนั้น ปริมาณหมุน (Pivotal quantity) ที่ได้จาก $\frac{\hat{\psi}_{mm} - \psi}{\sqrt{\text{Var}(\hat{\psi}_{mm})}}$ จึงมีการแจกแจงปกติมาตรฐาน เมื่อผู้วิจัยทำการหาความแปรปรวนของ $\hat{\psi}_{mm}$ จึงได้ว่า

$$\text{Var}(\hat{\psi}_{mm}) = \text{Var}\left(\sqrt{\frac{\pi}{2}} \hat{\sigma}_{mm}\right) = \frac{\pi}{2} \text{Var}(\hat{\sigma}_{mm})$$

โดยที่ $\text{Var}(\hat{\sigma}_{mm}) = (4-\pi)\sigma^2 / n\pi$ เมื่อแทนค่าลงในสมการข้างต้นจะได้

$$\text{Var}(\hat{\psi}_{mm}) = \left(\frac{4-\pi}{2n}\right)\sigma^2 \quad (4)$$

ดังนั้นช่วงความเชื่อมั่น $(1-\alpha) 100\%$ สำหรับ ψ จากวิธีที่อาศัยตัวประมาณโมเมนต์ $\hat{\psi}_{mm}$ คือ

$$CI_{Nmm} = (l_{Nmm}, u_{Nmm}) \\ = (\hat{\psi}_{mm} - z_{1-\alpha/2} \sqrt{\text{var}(\hat{\psi}_{mm})}, \hat{\psi}_{mm} + z_{1-\alpha/2} \sqrt{\text{var}(\hat{\psi}_{mm})})$$

เมื่อ $\text{var}(\hat{\psi}_{mm})$ คือ ตัวประมาณค่าความแปรปรวนของ $\hat{\psi}_{mm}$ ซึ่ง $\text{var}(\hat{\psi}_{mm})$ หาได้จากการแทนค่า σ^2 ในสมการ (4) ด้วย $\hat{\sigma}_{MLE}^2 = 2\bar{X}^2 / \pi$

3. ช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ที่อาศัย $\hat{\sigma}_{MLE}^2$

การสร้างช่วงความเชื่อมั่นในหัวข้อนี้อาศัยสมบัติของตัวประมาณภาวะนั้นจะเป็นสูงสุดสำหรับ σ^2 นั่นคือ $\hat{\sigma}_{MLE}^2 = \sum_{i=1}^n X_i^2 / 2n$ เนื่องจากทราบว่า $E(\hat{\sigma}_{MLE}^2) = \sigma^2$ และตัวแปรสุ่ม $Y = \sum_{i=1}^n X_i^2$ มีการแจกแจงแกมมา (Gamma distribution) ด้วยพารามิเตอร์ n และ $2\sigma^2$

หรือ $\text{Gamma}(n, 2\sigma^2)$ เมื่อ $\sigma^2 > 0$ และฟังก์ชันความหนาแน่นความน่าจะเป็นของ Y คือ

$$f_Y(y; \sigma) = \frac{1}{\Gamma(n)(2\sigma^2)^n} y^{n-1} e^{-y/(2\sigma^2)}$$

โดยที่ $y > 0$ และ n คือ ขนาดตัวอย่าง สำหรับ $E(Y) = 2n\sigma^2$ และ $\text{Var}(Y) = 4n\sigma^4$ ดังนั้น ความแปรปรวนของ $\hat{\sigma}_{MLE}^2$ จึงหาได้จาก

$$\text{Var}(\hat{\sigma}_{MLE}^2) = \text{Var}\left(\frac{1}{2n} \sum_{i=1}^n X_i^2\right) = \frac{\sigma^4}{n} \quad (5)$$

ในกรณีที่ตัวอย่างมีขนาดใหญ่ $\hat{\sigma}_{MLE}^2$ จะมีการแจกแจงปกติโดยประมาณด้วยค่าเฉลี่ย σ^2 และความแปรปรวน σ^4/n ดังนั้นช่วงความเชื่อมั่น $(1-\alpha) 100\%$ จากวิธีตัวอย่างขนาดใหญ่สำหรับ ψ ที่อาศัยตัวประมาณ $\hat{\sigma}_{MLE}^2$ คือ

$$CI_{MLE} = (l_{MLE}, u_{MLE}) = \left(\sqrt{\frac{\pi}{2}} \sqrt{L_{ML}}, \sqrt{\frac{\pi}{2}} \sqrt{U_{ML}}\right)$$

$$\text{เมื่อ } L_{ML} = \hat{\sigma}_{MLE}^2 - z_{1-\alpha/2} \sqrt{\text{var}(\hat{\sigma}_{MLE}^2)},$$

$U_{ML} = \hat{\sigma}_{MLE}^2 + z_{1-\alpha/2} \sqrt{\text{var}(\hat{\sigma}_{MLE}^2)}$ และ $\text{var}(\hat{\sigma}_{MLE}^2)$ คือ ตัวประมาณค่าความแปรปรวนของ $\hat{\sigma}_{MLE}^2$ ซึ่งหาได้จากการแทนค่า σ ในสมการ (5) ด้วย σ_{MLE}

4. ช่วงความเชื่อมั่นนัยทั่วไป

ช่วงความเชื่อมั่นนัยทั่วไปเป็นการสร้างช่วงความเชื่อมั่นสำหรับพารามิเตอร์ด้วยการหาเปอร์เซ็นต์ไทล์ของปริมาณหมุนนัยทั่วไป (Generalized pivotal quantity) ตามแนวคิดของ Weerahandi (1993) โดยปริมาณหมุนดังกล่าวมีสมบัติดังต่อไปนี้

1) ถ้าให้ $X=x$ แล้ว ฟังก์ชัน $R(X;x,\theta)$ ปรากฏจากพารามิเตอร์ที่ไม่ทราบค่า

2) ค่าสังเกตของ $R(X;x,\theta)$ หรือ $R(x;x,\theta)$ มีค่าเท่ากับ θ และไม่ขึ้นกับพารามิเตอร์รบกวน (Nuisance parameter) เมื่อ θ คือ พารามิเตอร์ทั่วไปในการแจกแจงความน่าจะเป็น

สำหรับการศึกษาค้นคว้าครั้งนี้ผู้วิจัยได้นำเสนอปริมาณหมุนนัยทั่วไปใหม่เพื่อใช้ในการสร้างช่วงความเชื่อมั่นสำหรับ ψ โดยมีข้อตกลงเบื้องต้นดังนี้ กำหนดให้ $X=(X_1, X_2, \dots, X_n)$ เป็นตัวอย่างสุ่มขนาด n จากการแจกแจง $Ra(\sigma)$ เมื่อ $\sigma > 0$ และเนื่องจาก $Y = \sum_{i=1}^n X_i^2$ มีการแจกแจง $\text{Gamma}(n, 2\sigma^2)$ ดังนั้นฟังก์ชันก่อกำเนิด

โมเมนต์ (Moment generating function) ของ $W = Y / \sigma^2$ จึงหาได้จาก

$$M_W(t) = E(e^{tW}) = M_Y\left(\frac{t}{\sigma^2}\right) = \frac{1}{(1-2t)^n}$$

ในที่นี้ W มีการแจกแจง $\text{Gamma}(n, 1/2)$ ซึ่งพบว่าการแจกแจงไม่ขึ้นอยู่กับการพารามิเตอร์และไม่มีพารามิเตอร์รบกวนเข้ามาเกี่ยวข้อง ดังนั้นจึงสามารถหาปริมาณหมูนัยทั่วไปสำหรับ σ^2 ได้จาก

$$R(X; x, \sigma^2) = \frac{Y}{W} \quad (6)$$

จาก (6) พบว่า 1) $R(X; x, \sigma^2)$ ไม่ขึ้นอยู่กับการพารามิเตอร์ และ 2) เมื่อแทนค่า $X = x$ แล้ว $R(x; x, \sigma^2) = \sigma^2$ ซึ่งเป็นไปตามเงื่อนไขของ Weerahandi (1993) ดังนั้นช่วงความเชื่อมั่นนัยทั่วไป $(1-\alpha) 100\%$ ของ σ^2 คือ

$$(R_{\alpha/2}(X; x, \sigma^2), R_{1-\alpha/2}(X; x, \sigma^2))$$

และจะสามารถหาช่วงความเชื่อมั่นนัยทั่วไป $(1-\alpha) 100\%$ สำหรับ ψ ได้จาก

$$CI_{GCI} = (l_{GCI}, u_{GCI}) \\ = \left(\sqrt{\pi/2} \sqrt{R_{\alpha/2}(X; x, \sigma^2)}, \sqrt{\pi/2} \sqrt{R_{1-\alpha/2}(X; x, \sigma^2)} \right)$$

โดยที่ $R_{\alpha/2}(X; x, \sigma^2)$ และ $R_{1-\alpha/2}(X; x, \sigma^2)$ คือ เปอร์เซ็นไทล์ที่ $(\alpha/2) \times 100$ และ $(1-\alpha/2) \times 100$ ของ $R(X; x, \sigma^2)$ ตามลำดับ

4. การจำลองและผลลัพธ์

การศึกษาประสิทธิภาพของช่วงความเชื่อมั่น CI_{NMLE} , CI_{Nmm} , CI_{MLE} และ CI_{GCI} ในครั้งนี้อาศัยการจำลองข้อมูลภายใต้สถานการณ์ต่าง ๆ โดยใช้โปรแกรม R (R Core Team, 2023) ผู้วิจัยกำหนดขนาดตัวอย่าง (n) เท่ากับ 10, 20, 30, 50, 100 และ 200 เพื่อแทนตัวอย่างขนาดเล็กไปจนถึงขนาดใหญ่ จำลองค่าของตัวแปรสุ่ม X จากการแจกแจงเรย์ลีย์ที่มีพารามิเตอร์ σ เท่ากับ 0.5, 1, 2 และ 3 จากรูปที่ 1 จะเห็นว่าลักษณะของข้อมูลจะแตกต่างกัน ทั้งนี้ในการสร้างเลขสุ่มผู้วิจัยได้อาศัยฟังก์ชัน *rrayleigh* จากโปรแกรม R (Wolodzko, 2022) และเมื่อคำนวณค่าเฉลี่ยประชากร (ψ) จึงมีค่าเท่ากับ 0.63, 1.25, 2.51 และ 3.76 จากนั้นจึงหาช่วงความเชื่อมั่นของ ψ จาก 4 วิธีโดยกำหนดระดับความเชื่อมั่น $(1-\alpha) 100\%$ เท่ากับ 0.95

การจำลองในแต่ละสถานการณ์ข้างต้น ผู้วิจัยกำหนดจำนวนการทำซ้ำ 10,000 รอบ แล้วคำนวณค่าความน่าจะเป็นคัมรวม (CP) และค่าความยาวช่วงโดยเฉลี่ย (EL) จาก

$$CP = \frac{1}{10,000} \sum_{i=1}^{10,000} (L_i \leq \psi \leq U_i) \quad \text{และ} \\ EL = \frac{1}{10,000} \sum_{i=1}^{10,000} (U_i - L_i)$$

โดยที่ $\sum_{i=1}^{10,000} (L_i \leq \psi \leq U_i)$ คือ ผลรวมของจำนวนครั้งที่ค่าขอบเขตล่าง L_i และขอบเขตบน U_i ในการทำซ้ำรอบที่ i ครอบคลุมค่าพารามิเตอร์ ψ และ $U_i - L_i$ คือความยาวของช่วงความเชื่อมั่นในการทำซ้ำรอบที่ i สำหรับช่วงความเชื่อมั่นที่มีประสิทธิภาพดีนั้นจะต้องมีค่าความน่าจะเป็นคัมรวมใกล้เคียงกับระดับความเชื่อมั่นที่กำหนด คือ $1-\alpha = 0.95$ และมีค่าความยาวช่วงโดยเฉลี่ยสั้นในแต่ละสถานการณ์ ทั้งนี้ผู้วิจัยได้พิจารณาขอบเขตที่ค่าความน่าจะเป็นคัมรวมจากการจำลองไม่แตกต่างจาก 0.95 อย่างมีนัยสำคัญทางสถิติจาก

$$(1-\alpha) \pm z_{1-\alpha} \sqrt{\frac{\alpha(1-\alpha)}{10,000}} = 0.95 \pm 1.645 \sqrt{\frac{0.05(0.95)}{10,000}} = \\ (0.9464, 0.9536)$$

ช่วงความเชื่อมั่นที่สอดคล้องกับเกณฑ์ดังกล่าวจะแสดงให้เห็นว่าเป็นวิธีที่สามารถประมาณค่าพารามิเตอร์ได้ถูกต้องและแม่นยำ

ตารางที่ 1 แสดงค่าความน่าจะเป็นคัมรวมและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น 95% สำหรับ ψ จากการจำลอง ซึ่งผลสรุปที่ได้มีดังนี้

1) จากการศึกษาพบว่าค่าความน่าจะเป็นคัมรวมของ CI_{NMLE} ไม่เป็นไปตามระดับความเชื่อมั่นที่กำหนด กล่าวคือมีค่าสูงกว่า 0.95 อย่างมากในทุกกรณีที่ทำการศึกษา ส่วน CI_{MLE} และ CI_{Nmm} มีค่าความน่าจะเป็นคัมรวมใกล้เคียง 0.95 เฉพาะในกรณีที่ $n \geq 50$ และ $n \geq 30$ ตามลำดับ เท่านั้น

2) CI_{GCI} ให้ค่าความน่าจะเป็นคัมรวมใกล้เคียง 0.95 ในทุกกรณีที่ทำการศึกษา โดยค่าความน่าจะเป็นคัมรวมที่ได้นั้นมีความคงเส้นคงวา (Consistency) กล่าวคือเมื่อ n , σ หรือ ψ มีค่าเปลี่ยนไป ความน่าจะเป็นคัมรวมของ CI_{GCI} ยังคงมีค่าใกล้เคียง 0.95 เสมอ

ตารางที่ 1 ค่าความน่าจะเป็นคัมรวมและค่าความยาวช่วงโดยเฉลี่ยของช่วงความเชื่อมั่น 95% สำหรับค่าเฉลี่ยประชากรในการแจกแจงเรย์ลีย์ ที่ได้จากการจำลอง

n	พารามิเตอร์		ความน่าจะเป็นคัมรวม			ความยาวช่วงโดยเฉลี่ย				
	σ	ψ	NMLE	MLE	Nmm	GCI	NMLE	MLE	Nmm	GCI
10	0.5	0.63	0.9985	0.9024	0.9347	0.9511	0.9964	0.4062	0.4064	0.4200
	1	1.25	0.9989	0.9018	0.9325	0.9463	1.9895	0.8111	0.8116	0.8385
	2	2.51	0.9987	0.9060	0.9356	0.9486	3.9952	1.6288	1.6289	1.6837
	3	3.76	0.9979	0.9040	0.9342	0.9515	5.9662	2.4324	2.4324	2.5142
20	0.5	0.63	1.0000	0.9287	0.9439	0.9513	1.0015	0.2799	0.2869	0.2849
	1	1.25	1.0000	0.9217	0.9388	0.9497	2.0036	0.5600	0.5738	0.5702
	2	2.51	1.0000	0.9253	0.9387	0.9467	4.0133	1.1217	1.1501	1.1426
	3	3.76	1.0000	0.9258	0.9384	0.9470	6.0158	1.6813	1.7232	1.7123
30	0.5	0.63	1.0000	0.9350	0.9440	0.9525	1.0060	0.2275	0.2347	0.2303
	1	1.25	1.0000	0.9309	0.9439	0.9464	2.0066	0.4537	0.4684	0.4593
	2	2.51	1.0000	0.9279	0.9414	0.9463	4.0146	0.9078	0.9369	0.9189
	3	3.76	1.0000	0.9371	0.9447	0.9491	6.0267	1.3628	1.4068	1.3793
50	0.5	0.63	1.0000	0.9418	0.9450	0.9489	1.0068	0.1751	0.1817	0.1764
	1	1.25	1.0000	0.9437	0.9509	0.9544	2.0127	0.3501	0.3633	0.3527
	2	2.51	1.0000	0.9406	0.9465	0.9495	4.0263	0.7004	0.7268	0.7054
	3	3.76	1.0000	0.9392	0.9445	0.9483	6.0337	1.0495	1.0891	1.0572
100	0.5	0.63	1.0000	0.9472	0.9470	0.9510	1.0078	0.1233	0.1285	0.1237
	1	1.25	1.0000	0.9450	0.9486	0.9492	2.0146	0.2466	0.2568	0.2473
	2	2.51	1.0000	0.9476	0.9529	0.9515	4.0291	0.4931	0.5136	0.4946
	3	3.76	1.0000	0.9420	0.9469	0.9487	6.0393	0.7391	0.7697	0.7413
200	0.5	0.63	1.0000	0.9486	0.9501	0.9496	1.0077	0.0870	0.0908	0.0871
	1	1.25	1.0000	0.9480	0.9507	0.9494	2.0165	0.1741	0.1817	0.1743
	2	2.51	1.0000	0.9480	0.9494	0.9478	4.0329	0.3481	0.3633	0.3485
	3	3.76	1.0000	0.9503	0.9489	0.9523	6.0432	0.5217	0.5444	0.5223

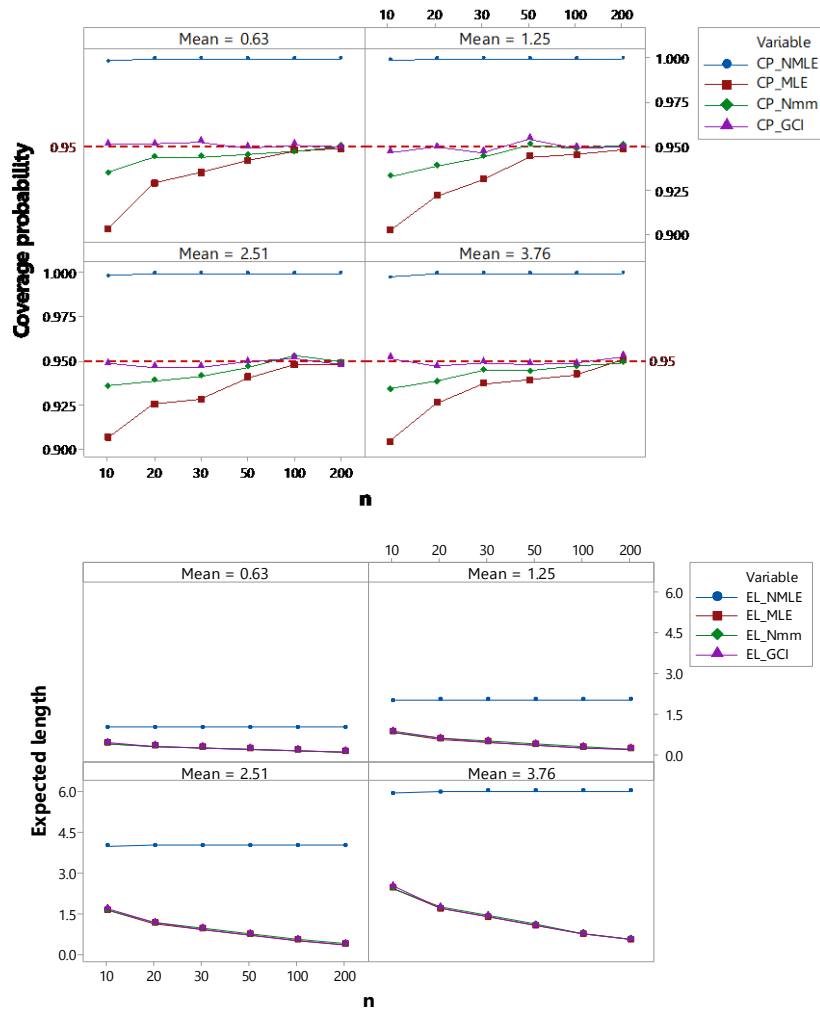
3) เมื่อพิจารณาประสิทธิภาพของตัวประมาณในเรื่องความยาวเฉลี่ยของช่วงความเชื่อมั่นพบว่า CI_{GCI} , CI_{Nmm} และ CI_{MLE} มีค่าความยาวช่วงโดยเฉลี่ยไม่แตกต่างกันมากนัก ในขณะที่ CI_{NMLE} ที่มีค่าความยาวช่วงโดยเฉลี่ยมากที่สุดเมื่อเทียบกับช่วงความเชื่อมั่นจากวิธีการอื่น ๆ อย่างไรก็ตามจะเห็นได้ว่าความยาวช่วงโดยเฉลี่ยที่ได้จากแต่ละวิธีจะมีค่าน้อยลง เมื่อ n มีค่าเพิ่มขึ้น และจะแปรผันตามค่าของ ψ

จากผลการจำลองจึงสรุปได้ว่า CI_{GCI} เป็นช่วงความเชื่อมั่นที่มีประสิทธิภาพสูงที่สุดในการศึกษาครั้งนี้ เนื่องจากสามารถประมาณค่าเฉลี่ยประชากรในการแจกแจงเรย์ลีย์ได้ด้วยค่าความน่าจะเป็นที่กำหนด และมีความยาวช่วงที่เหมาะสมในทุกสถานการณ์ดังแสดงในรูปที่ 2 ดังนั้นผู้วิจัยจึงเสนอให้ใช้ CI_{GCI} ในทางปฏิบัติสำหรับกรณีที่ตัวอย่างมีขนาดมากกว่า 30 หน่วย CI_{MLE} สามารถใช้เป็นวิธีทางเลือกในการประมาณค่าเฉลี่ยประชากรของ

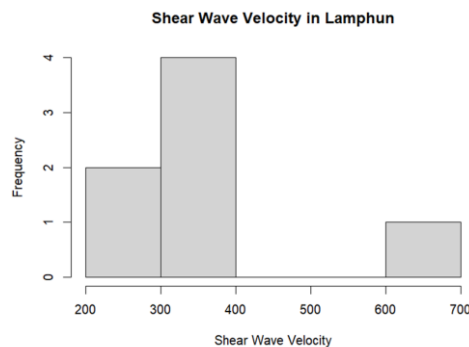
การแจกแจงเรย์ลีย์ได้ เนื่องจากเป็นวิธีที่ให้ค่าความน่าจะเป็นคัมรวมที่ยอมรับได้และมีความยาวช่วงสั้น นอกจากนั้น CI_{MLE} เป็นวิธีที่ไม่ต้องอาศัยโปรแกรมคอมพิวเตอร์ช่วยในการคำนวณ เนื่องจากมีสูตรปิดในการหาช่วงความเชื่อมั่น เชื่อมั่น 95% สำหรับสัดส่วนของผู้โดยสารที่ติดเชื้อไวรัส

การประยุกต์ใช้ข้อมูลจริง

ความเร็วคลื่นเฉือนของดิน (Shear-wave velocity of soils) คือ สมบัติของดินที่สามารถขยายแรงแผ่นดินไหวให้มีความรุนแรงมากขึ้นได้ ในงานวิศวกรรมธรณีเทคนิคจึงนิยมสำรวจและวัดค่าดังกล่าวมาทำการศึกษาแหล่งกำเนิดคลื่นแผ่นดินไหว ในปัจจุบันความเร็วคลื่นเฉือนของดินสามารถหาได้จากการลงภาคสนาม



รูปที่ 2 กราฟแสดงค่าความน่าจะเป็นคุ่มรวม (บน) และค่าความยาวโดยเฉลี่ย (ล่าง) ของช่วงความเชื่อมั่น 95% สำหรับค่าเฉลี่ยประชากรในการแจกแจงเรย์ลีย์ ที่ได้จากการจำลอง



รูปที่ 3 ฮิสโตแกรมของข้อมูลความเร็วคลื่นเฉือนของดิน (หน่วย: เมตร/วินาที) ในจังหวัดลำพูน

(วัดในหลุมเจาะสำรวจ) และสำรวจคลื่นไหวสะเทือน พื้นผิวแบบหลายช่องรับสัญญาณ (Multi-channel Analysis of Surface Wave: MASW) วิธีหลังนี้เป็นวิธีใหม่ที่มีราคาไม่สูงมากนัก และสามารถทำได้รวดเร็วกว่าวิธีแรก (ฐานบ ธิติมากร และคณะ, 2551) อย่างไรก็ตามถึงแม้ประเทศไทยจะไม่ได้รับผลกระทบโดยตรงจากแผ่นดินไหวในระดับรุนแรงมากนัก แต่ก็มีก่เกิดแผ่นดินไหวที่ประชาชนสามารถรับรู้ถึงแรงสั่นสะเทือนได้

จากข้อมูลสถิติพบว่าตั้งแต่ พ.ศ. 2550 ถึง พ.ศ. 2559 อัตราการเกิดแผ่นดินไหวในประเทศไทยมีมากขึ้น (โชติ เนื่องนันท์ และคณะ, 2561) การเกิดแผ่นดินไหวเป็นปรากฏการณ์ทางธรรมชาติที่ไม่สามารถคาดเดาการเกิดล่วงหน้าได้ แต่ด้วยวิธีการที่ทันสมัยจึงทำให้นักวิจัยทราบแนวโน้มในการเกิดเหตุการณ์ดังกล่าวได้เช่นกัน และหนึ่งในนั้นคือการวิเคราะห์คลื่นพื้นผิวจากค่าความเร็วคลื่นเฉือน

ตารางที่ 2 การตรวจสอบความเหมาะสมของตัวแบบการแจกแจงความน่าจะเป็นของข้อมูลตัวอย่าง

เกณฑ์/การแจกแจง	การแจกแจงเรย์ลีย์	การแจกแจงแบบเลขชี้กำลัง	การแจกแจงโคกำลังสอง
AIC	90.6010	98.5239	154.0379
BIC	90.5469	98.4699	153.9838

ตารางที่ 3 ช่วงความเชื่อมั่น 95% สำหรับค่าเฉลี่ยของความเร็วคลื่นเฉือนของดิน (เมตร/วินาที) ในจังหวัดลำพูน

วิธี	ขีดจำกัดล่าง	ขีดจำกัดบน	ค่าความยาวช่วง
CI_{NMLE}	65.7395	607.3215	541.5819
CI_{MLE}	171.3344	444.0158	272.6814
CI_{Nmm}	222.4714	503.6477	281.1763
CI_{GCI}	244.0206	525.4918	281.4712

การศึกษาในครั้งนี้ผู้วิจัยได้นำค่าความเร็วคลื่นเฉือนของดินในจังหวัดลำพูน ซึ่งเป็นจังหวัดที่มีพื้นที่ใกล้เคียงมีรอยเลื่อนที่มีพลังหลายรอยเลื่อน และมีประวัติของการเกิดแผ่นดินไหวในพื้นที่มาทำการศึกษา ข้อมูลที่ใช้ในครั้งนี้ถูกเก็บรวบรวมโดย ธนบดี แร่นาค (2557) จากการตรวจวัดด้วยเครื่อง MASW แบบ 24 ช่องสัญญาณ ค่าความเร็วคลื่นเฉือนที่วัดได้จาก 7 สถานีที่ระดับความลึก 30 เมตร (เมตร/วินาที) แสดงในรูปที่ 3 เมื่อคำนวณค่าสถิติเบื้องต้นแล้วพบว่า ค่าเฉลี่ยตัวอย่าง $\bar{X}$ เท่ากับ 363.0596 ค่ามัธยฐานเท่ากับ 344.9510 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 120.2079 เมตร/วินาที อย่างไรก็ตามเพื่อให้การประมาณค่าสอดคล้องกับการแจกแจงที่แท้จริงของข้อมูล ผู้วิจัยจะทำการตรวจสอบความเหมาะสมของตัวแบบความน่าจะเป็น (Model selection) สำหรับข้อมูลชุดนี้โดยอาศัยเกณฑ์ข้อสนเทศของอะกะอิเกะ (Akaike information criterion: AIC) และเกณฑ์ข้อสนเทศของเบส์ (Bayesian information criterion: BIC) ก่อน จากการทบทวนวรรณกรรมพบว่าความเร็วคลื่นเฉือนของดินอาจมีการแจกแจงเรย์ลีย์ การแจกแจงแบบเลขชี้กำลัง และการแจกแจงโคกำลังสองได้ (Pasari & Dikshit, 2015) ดังนั้นเมื่อพิจารณาค่า AIC และ BIC ของการแจกแจงต่าง ๆ ในตารางที่ 2 แล้วพบว่า การแจกแจงเรย์ลีย์มีค่า AIC และ BIC ต่ำที่สุด ดังนั้นจึงเป็นการแจกแจงที่เหมาะสมที่สุดสำหรับข้อมูลชุดนี้ จากนั้นผู้วิจัยได้ทำการทดสอบความมีนัยสำคัญของการแจกแจงโดยใช้สถิติทดสอบคอลโมโกรอฟ-สมิรโนฟ (Kolmogorov-Smirnov test) ภายใต้สมมติฐาน คือ

H_0 : ความเร็วคลื่นเฉือนของดินมีการแจกแจงเรย์ลีย์

H_1 : ความเร็วคลื่นเฉือนของดินไม่มีการแจกแจงเรย์ลีย์

ที่ระดับนัยสำคัญ 0.05 พบว่าความเร็วคลื่นเฉือนของดินที่ทำการศึกษาในครั้งนี้มีการแจกแจงเรย์ลีย์ (ค่าสถิติเท่ากับ

0.3983 และค่า p-value เท่ากับ 0.1630) ข้อมูลชุดนี้จึงสามารถนำไปใช้ในการประมาณค่าเฉลี่ยประชากรจาก CI_{NMLE} , CI_{Nmm} , CI_{MLE} และ CI_{GCI} ได้ผลที่ได้จากตารางที่ 3 พบว่า CI_{MLE} , CI_{Nmm} และ CI_{GCI} ให้ค่าความยาวช่วงความเชื่อมั่นไม่แตกต่างกันมากนัก ในขณะที่ CI_{NMLE} มีความยาวช่วงมากที่สุด ซึ่งสอดคล้องกับผลที่ได้จากการจำลอง แต่อย่างไรก็ตามผลจากการจำลองแสดงให้เห็นชัดเจนว่าเมื่อตัวอย่างมีขนาดเล็ก CI_{GCI} ให้ประสิทธิภาพดีที่สุด ดังนั้นจึงสรุปได้ว่าค่าเฉลี่ยของความเร็วคลื่นเฉือนของดินในจังหวัดลำพูนมีค่าอยู่ระหว่าง 244.0206 ถึง 525.4918 เมตร/วินาที และเมื่อพิจารณาตามเกณฑ์ National Earthquake Hazards Reduction Program (NEHRP) แล้วพบว่าชั้นดินในบริเวณและระยะเวลาที่นำข้อมูลมาทำการศึกษาจัดอยู่ในประเภทดินแข็ง (Stiff soil) ค่อนไปทางดินแน่นมากและหินผุ (Dense soil and soft rock) ซึ่งอาจเกิดแผ่นดินไหวได้มาก

สรุปและอภิปรายผลการศึกษา

ในการศึกษาครั้งนี้ผู้วิจัยได้นำเสนอช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยประชากรในการแจกแจงเรย์ลีย์ 4 วิธี คือ วิธีของวาลด์ วิธีช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ที่อาศัยตัวประมาณโมเมนต์ วิธีช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่ที่อาศัย $\hat{\sigma}_{MLE}^2$ และวิธีช่วงความเชื่อมั่นนัยทั่วไป จากการจำลองภายใต้สถานการณ์ต่าง ๆ พบว่าช่วงความเชื่อมั่นนัยทั่วไปมีประสิทธิภาพดีที่สุด กล่าวคือเป็นวิธีที่ให้ค่าความน่าจะเป็นคุ้มครองใกล้เคียงกับระดับความเชื่อมั่นที่กำหนด ทั้งนี้เป็นผลเนื่องมาจากวิธีนี้ได้ถูกพิจารณาภายใต้การแจกแจงที่แท้จริงของปริมาณหมุน (ไม่ได้อาศัยการแจกแจงปกติโดยประมาณอย่างวิธีการอื่น) ในทางปฏิบัติจึงสามารถนำวิธีช่วงความเชื่อมั่นนัยทั่วไป (CI_{GCI}) ไปใช้ได้กับทุกขนาดตัวอย่าง ในขณะที่ช่วงความเชื่อมั่นสำหรับตัวอย่างขนาดใหญ่นั้นได้

ถูกสร้างขึ้นมาโดยอาศัยทฤษฎีที่จำกัดส่วนกลางจึงทำให้มีประสิทธิภาพดีเมื่อตัวอย่างมีขนาดใหญ่พอ ซึ่งการศึกษานี้แนะนำให้ใช้เมื่อตัวอย่างมีขนาดมากกว่า 50 หน่วย ช่วงความเชื่อมั่นนี้ทั่วไปนอกจากจะมีประสิทธิภาพที่ดีกว่าช่วงความเชื่อมั่นที่นำมาเปรียบเทียบกับความน่าจะเป็นคุ่มรวมแล้ว ยังมีความยาวของช่วงสั้นซึ่งแสดงให้เห็นว่าสามารถประมาณค่าเฉลี่ยได้ใกล้เคียงกับค่าจริง และวิธีนี้สามารถใช้โปรแกรมคอมพิวเตอร์ช่วยในการหาขีดจำกัดบนและขีดจำกัดล่างได้สะดวก การศึกษาในครั้งนี้ได้นำวิธีที่นำเสนอในทางทฤษฎีมาประยุกต์ใช้กับข้อมูลจริง ซึ่งผลที่ได้พบว่ามีความสอดคล้องกับผลที่ได้จากการจำลอง

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณบรรณาธิการและกรรมการที่ได้ให้ข้อเสนอแนะสำหรับบทความวิจัยฉบับนี้

เอกสารอ้างอิง

- Ahmed, A. A., Ahmad, S. P., & Reshi, J. A. (2013). Bayesian analysis of Rayleigh distribution. *International Journal of Scientific and Research Publications*, 3(10), 1-9.
- Ahn, C. Y., Jung, Y. M., Kwon, O. I., & Seo, J. K. (2012). Fast segmentation of ultrasound images using robust Rayleigh distribution decomposition. *Pattern Recognition*, 45(9), 3490-3500.
- Ateeq, K. Qasim, T. B., & Alvi, A. R. (2019). An extension of Rayleigh distribution and applications. *Cogent Mathematics & Statistics*, 6(1622191), 1-16.
- Casella, G., & Berger, R. L. (2002). *Statistical Inference*. California, USA: Duxbury Press.
- Dey, S. (2009). Comparison of Bayes estimators of the parameter and reliability function for Rayleigh distribution under different loss functions. *Malaysian Journal of Mathematical Sciences*, 3(2), 249-266.
- Dey, S., & Dey, T. (2014). Statistical inference for the Rayleigh distribution under progressively type-II censoring with binomial removal. *Applied Mathematical Modelling*, 38(3), 974-982.
- Abd Elfattah, A. M., Hassan, A. S., & Ziedan, D. M. (2006). Efficiency of maximum likelihood estimators under different censored sampling schemes for Rayleigh distribution. *Interstat*, 1, 1-16.
- Islam, D. K., Chaichana, T., Dussadee, N., & Intaniwet, A. (2017). Statistical distribution and energy estimation of the wind speed at Saint Martin's Island, Bangladesh. *International Journal of Renewable Energy*, 15(1), 77-88.
- Kruiver, P. P., Dedem, E. V., Romijn, R., Lange, G. D., Korff, M., Stafleu, J., Gunnink, J. L., Marek, A. R., Bommer, J. J., Elk, J. A., & Doornhof, D. (2017). An integrated shear-wave velocity model for the Groningen gas field, The Netherlands. *Bulletin of Earthquake Engineering*, 15(9), 3555-3580.
- Liebenberg, S. C., & Allison, S. J. (2023). A review of goodness-of-fit tests for the Rayleigh distribution. *Austrian Journal of Statistics*, 1-22.
- Nar, F., Okman, O. E., Özgür, A., & Çetin, M. (2018). Fast target detection in radar images using Rayleigh mixtures and summed area tables. *Digital Signal Processing*, 77(2018), 86-101.
- Pasari, S., & Dikshit, O. (2015). Earthquake interevent time distribution in Kachchh, Northwestern India. *Earth, Planets and Space*, 67(1), 1-17.
- Pishgar-Komleh, S. H., Keyhani, A., & Sefeedpari, P. (2015). Wind speed and power density analysis based on Weibull and Rayleigh distributions (a case study: Firouzkooch county of Iran). *Renewable and Sustainable Energy Reviews*, 42, 313-322.
- R Core Team. (2023). R: A Language and Environment for Statistical Computing. Vienna, Austria.
- Sarti, A., Corsi, C., Mazzini, E., & Lamberti, C. (2005). Maximum likelihood segmentation of ultrasound images with Rayleigh distribution. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 52(6), 947-960.
- Seo, J. I., Jeon, J. W., & Kang, S. B. (2016). Exact interval inference for the two-parameter Rayleigh distribution based on the upper record values. *Journal of Probability and Statistics*, 2016, 1-5.
- Weerahandi, S. (1993). Generalized confidence intervals. *Journal of the American Statistical Association*, 88(423), 899-905.
- Wolodko, T. (2020). *Additional Univariate and Multivariate Distributions*. <https://cran.r-project.org/web/packages/extraDistr/extraDistr.pdf>
- Xu, W., Ning, L., & Luo, Y. (2020). Wind speed forecast based on post-processing of numerical weather predictions using a gradient boosting decision tree algorithm. *Atmosphere*, 11(7), 1-38.
- Nungun, C., Phungkham, N., Srikongruk, S., Mameechai, D., Thongperm, P., & Noolek, W. (2561). Trends in earthquakes in Thailand and Burma. *Journal of Thaksin University*, 21(3), 168-173. (in Thai)
- Thitimakorn, T., Luang-Sunthornkul, N., Junthamas, P., & Hansiri, A. (2551). Determining the shear wave velocity of the Bangkok soil layer From surface wave analysis. *Substance Engineering, Research and Development Edition*, 19(4), 30-35. (in Thai)
- Raenak, T. (2557). Determination of soil shear wave velocity using an MASW surface wave receiver belt. Mueang District area Lamphun Province [Unpublished Bachelor thesis]. Chulalongkorn University. (in Thai)