

การเปรียบเทียบวิธีการประมาณค่าของช่วงความเชื่อมั่นเชิงเส้นกำกับ และช่วงความเชื่อมั่นที่แท้จริงสำหรับสัดส่วนของประชากร ที่มีการแจกแจงทวินามและการประยุกต์

Comparison of Methods on Asymptotic and Exact Confidence Intervals for a Binomial Proportion with Application

ชนกานต์ สังข์บุญชู*, ปิยพล ไพจิตร

Chanakan Sungboonchoo*, Piyaphon Paichit

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร ประเทศไทย

Department of Statistics, Faculty of Science, Silpakorn University, Thailand

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 2 ประเภท คือ ช่วงความเชื่อมั่นที่แท้จริง และช่วงความเชื่อมั่นเชิงเส้นกำกับ โดยศึกษาเปรียบเทียบวิธีการประมาณค่าช่วงความเชื่อมั่น 5 วิธี ได้แก่ วิธีวัลด์ (Wald) วิธี Wilson score วิธี Agresti-Coull วิธี New และวิธี generalized Agresti-Coull จำลองข้อมูลโดยใช้วิธีมอนติคาร์โลให้มีการแจกแจงทวินามด้วยพารามิเตอร์ ขนาดตัวอย่าง (n) มีค่าเท่ากับ 10, 30, 50 และ 100 กำหนดสัดส่วน (p) มีค่าเท่ากับ 0.1, 0.3, 0.5, 0.7 และ 0.9 และกำหนดระดับความเชื่อมั่นเท่ากับ 95% โดยทำการทดลองซ้ำ 10,000 ครั้งในแต่ละสถานการณ์ ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่น คือ ค่าความน่าจะเป็นคัมรวม และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น ผลการวิจัยสรุปได้ดังนี้ ช่วงความเชื่อมั่นวิธี Wilson score มีประสิทธิภาพดีที่สุดหลายสถานการณ์ที่ศึกษา ซึ่งช่วงความเชื่อมั่นนี้เหมาะกับสถานการณ์ที่ $n=10$ เมื่อ $p=0.5$ และ $n=30$ เมื่อ p มีค่าเข้าใกล้ 0 หรือ 1 และ $n=50$ สำหรับทุกค่าของ p ยกเว้น $p=0.5$ ช่วงความเชื่อมั่นวิธี Agresti-Coull มีประสิทธิภาพเป็นอันดับที่สอง ซึ่งเหมาะกับ $n=10$ เมื่อ $p=0.3$ และ 0.7 และ $n=30$ เมื่อ $p=0.3-0.7$ ส่วนช่วงความเชื่อมั่นวิธี New มีประสิทธิภาพเป็นอันดับที่สาม ซึ่งเหมาะกับ $n=10$ และ 100 เมื่อ p มีค่าเข้าใกล้ 0 หรือ 1 ช่วงความเชื่อมั่นวิธี generalized Agresti-Coull มีประสิทธิภาพเป็นอันดับที่สี่ ซึ่งเหมาะกับ $n=100$ เมื่อ $p=0.3$ และ 0.7 และช่วงความเชื่อมั่นวิธีวัลด์ (Wald) มีประสิทธิภาพที่ไม่ค่อยดี นอกจากนั้นช่วงความเชื่อมั่นทั้ง 5 วิธีสามารถนำไปประยุกต์ใช้กับข้อมูลจริงในสถานการณ์ต่าง ๆ

คำสำคัญ: ช่วงความเชื่อมั่น, พารามิเตอร์ค่าสัดส่วนทวินาม, ความน่าจะเป็นคัมรวม, การจำลองมอนติคาร์โล

ABSTRACT

The objective of this research was to compare efficiency of the methods for estimating two types of confidence interval for a binomial proportion, which are exact confidence interval and asymptotic confidence interval by studying comparison upon 5 methods for estimating confidence interval. These 5 methods are Wald method, Wilson Score method, Agresti-Coull method, New method, and generalized Agresti-Coull method. The data was simulated by using Monte Carlo method, utilizing binomial distribution with parameters n and p that determine sizes of the sample (n) equal to 10, 30, 50, and 100. Meanwhile, the values of proportion (p) were determined as 0.1, 0.3, 0.5, 0.7, and 0.9 with confidence level 95% by repeating experiment 10,000 times in each certain circumstance. Moreover, the criteria used to compare efficiency of confidence intervals was coverage probability as well as average length. According to research findings, it can be summarized as follows. Confidence interval of Wilson Score method showed the best efficiency in diverse circumstances in case study. Besides, this confidence interval was appropriate for the circumstance at $n=10$ when $p=0.5$ and $n=30$ when p was close to 0 or 1, and $n=50$ for every value of p , except $p=0.5$. Furthermore, confidence interval of Agresti-Coull method had the most second efficiency, which was appropriate for $n=10$ when $p=0.3$ and 0.7, and $n=30$ when $p=0.3-0.7$, whereas confidence interval of New method had the third most efficiency, which was appropriate for $n=10$ and 100 when value of p was close to 0 or 1. At the same time, confidence interval of generalized Agresti-Coull method had the fourth most efficiency, which was appropriate for $n=100$ when $p=0.3$ and 0.7.

However, confidence interval of Wald method possess quite low precision. In addition, confidence interval of these 5 methods can be applicable to the real data in any circumstances.

KEYWORDS: Confidence interval, Binomial proportion parameter, Coverage probability, Monte Carlo simulation

*Corresponding Author: sungboonchoo_c@silpakorn.edu

Received: 12/01/2023; Revised: 19/06/2023; Accepted: 14/07/2023

1. บทนำ

การอนุมานเชิงสถิติ (statistical inference) คือ กระบวนการหาข้อสรุปเกี่ยวกับลักษณะของประชากร โดยใช้ข้อมูลจากตัวอย่างที่ได้มาโดยวิธีการเชิงสุ่มจากประชากรนั้น ในการอนุมานเชิงสถิติแบ่งออกเป็น 2 ลักษณะ คือ การทดสอบสมมติฐานและการประมาณค่า ซึ่งการประมาณค่ามี 2 แบบ คือ การประมาณค่าแบบจุด (point estimation) เป็นการประมาณค่าพารามิเตอร์ด้วยค่าสถิติเพียงค่าเดียว ซึ่งตัวประมาณแบบจุดที่ได้มีโอกาสคลาดเคลื่อนไปจากค่าพารามิเตอร์ได้มาก และการประมาณค่าแบบช่วง (interval estimation) เป็นการประมาณค่าพารามิเตอร์ด้วยช่วง ๆ หนึ่งหรือค่า 2 ค่าที่เป็นขีดจำกัดล่าง (lower limit) และขีดจำกัดบน (upper limit) ของช่วงความเชื่อมั่น ค่า 2 ค่าดังกล่าวสร้างจากตัวประมาณแบบจุดและการแจกแจงภาวะน่าจะเป็นของตัวประมาณนั้น โดยคาดว่าพารามิเตอร์มีแนวโน้มจะอยู่ในช่วงนี้หรืออยู่ระหว่างค่า 2 ค่านี้อย่างมีความเชื่อมั่นระดับหนึ่ง ซึ่งจะเห็นได้ว่าการประมาณค่าแบบช่วงสามารถบอกขอบเขตของค่าพารามิเตอร์ที่สนใจได้แม่นยำกว่าการประมาณค่าแบบจุด

การประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินามนั้นมีการศึกษากันอย่างแพร่หลาย ซึ่งพารามิเตอร์สำหรับสัดส่วน (p) ของการแจกแจงทวินามนี้แทนความน่าจะเป็นของการเกิดสิ่งที่สนใจ โดยให้ตัวแปรสุ่ม X แทน จำนวนครั้งของการเกิดสิ่งที่สนใจจากการทดลองสุ่มที่กระทำซ้ำ ๆ กัน n ครั้ง โดยการทดลองแต่ละครั้งเป็นอิสระต่อกัน ซึ่งในการกระทำแต่ละครั้งจะมีผลลัพธ์ที่เกิดขึ้นเป็นไปได้ 2 ลักษณะ คือ ผลสำเร็จ (success) และผลไม่สำเร็จหรือผลล้มเหลว (failure) นั่นคือ $X \sim B(n, p)$ และมีฟังก์ชันมวลความน่าจะเป็น $f(x) = \binom{n}{x} p^x (1-p)^{n-x}, x=0,1,2,\dots,n$ การประมาณค่าสัดส่วนนั้นมีความสำคัญทั้งในด้านการแพทย์ เช่น Komoltri (2012) ได้ศึกษาการประมาณค่าสัดส่วนโดยผู้วิจัย

ต้องการทราบค่าความไว (sensitivity) และความจำเพาะ (specificity) ของ ultrasound, mammogram เทียบกับผลการตรวจทางพยาธิวิทยาของชิ้นเนื้อ (gold standard) ในการวินิจฉัยมะเร็งเต้านม ค่าต่าง ๆ เหล่านี้จะรายงานเป็นสัดส่วนของความไวหรือช่วงความเชื่อมั่นสำหรับสัดส่วนของค่าความไวในประชากร และในด้านสิ่งแวดล้อม เช่น Muenmee & Bootdee (2021) ได้ศึกษาความเข้มข้นของฝุ่น $PM_{2.5}$ ที่ปลดปล่อยมาจากพื้นที่เขตอุตสาหกรรมในอำเภอปลวกแดง จังหวัดระยอง และนำมาทำการประเมินความเสี่ยงต่อสุขภาพของมนุษย์ด้วยการคำนวณจากค่าสัดส่วนความเสี่ยง (Hazard Quotient; HQ) เป็นต้น

ข้อมูลทวิภาค (binary data) จะพบได้บ่อยในการวิเคราะห์ทางสถิติ ซึ่งจะเป็นข้อมูลของตัวแปรที่มีค่าที่เป็นไปได้เพียง 2 ค่า เช่น สำเร็จ-ไม่สำเร็จ ใช่-ไม่ใช่ เป็นต้น และช่วงความเชื่อมั่นที่นำมาใช้สำหรับการประมาณค่าสัดส่วนของข้อมูลทวิภาคหรือสัดส่วนของประชากรที่มีการแจกแจงทวินาม คือ ช่วงความเชื่อมั่นของวิธี Wald (Wald, 1943) ซึ่งเป็นวิธีดั้งเดิมและมีวิธีการคำนวณที่ง่ายจึงนิยมใช้กันอย่างแพร่หลาย จากการศึกษางานวิจัยส่วนใหญ่ที่เกี่ยวข้องกับการประมาณค่าสัดส่วนโดยใช้ช่วงความเชื่อมั่นของวิธี Wald จะพบปัญหาที่สำคัญ คือ Brown et al. (2001) พบว่าช่วงความเชื่อมั่นของวิธี Wald มีความกว้างของค่าความน่าจะเป็นคลุมรวม (coverage probability) สูงในกรณีที่ค่าพารามิเตอร์สัดส่วนมีค่าเข้าใกล้ 0 หรือ 1 Agresti & Coull, 1998; Brown et al. (2002) ได้กล่าวถึงช่วงความเชื่อมั่นของวิธี Wald มีประสิทธิภาพที่ไม่ดีนักเนื่องจากให้ค่าความน่าจะเป็นคลุมรวมต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด จึงมีงานวิจัยที่เสนอวิธีการปรับปรุงช่วงความเชื่อมั่นสำหรับการประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินามหลากหลายวิธี ซึ่งจะมีช่วงความเชื่อมั่นของวิธี Wilson score (Wilson, 1927) เป็นอีกทางเลือกหนึ่ง เพราะเป็นวิธีที่มีสูตรการคำนวณช่วงความเชื่อมั่นที่ง่ายและมีประสิทธิภาพที่ดีกว่าช่วงความเชื่อมั่นของวิธี Wald แต่จะมี

ประสิทธิภาพที่ไม่ดีนักสำหรับในกรณีที่ค่าพารามิเตอร์สัดส่วนมีค่าเข้าใกล้ 0 หรือ 1 โดยทั่วไปช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามจะแบ่งออกเป็น 2 ประเภท ได้แก่ 1) ช่วงความเชื่อมั่นที่แท้จริง (exact confidence interval) ซึ่งโดยทั่วไปช่วงความเชื่อมั่นประเภทนี้จะไม่มีรูปแบบของสูตรช่วงความเชื่อมั่นที่ชัดเจน สำหรับการสร้างช่วงความเชื่อมั่นที่แท้จริงจะนิยมใช้วิธีแก้ปัญหาคณิตศาสตร์เชิงตัวเลข (numerical solution) เช่น ช่วงความเชื่อมั่นของวิธี Clopper-Pearson (Clopper & Pearson, 1934) และ 2) ช่วงความเชื่อมั่นเชิงเส้นกำกับ (asymptotic confidence interval) จะสร้างขึ้นโดยใช้ทฤษฎีเชิงเส้นกำกับ (asymptotic theory) ซึ่งลักษณะทั่วไปของช่วงความเชื่อมั่นประเภทนี้จะไม่มีรูปแบบของสูตรการคำนวณช่วงความเชื่อมั่นที่ชัดเจน เช่น ช่วงความเชื่อมั่นของวิธี Wald (Wald) วิธี Wilson score และวิธี Agresti-Coull เป็นต้น Cao et al. (2022) ได้ทำการปรับปรุงช่วงความเชื่อมั่นจากวิธี Wilson score เพื่อแก้ไขปัญหาในกรณีที่ค่าพารามิเตอร์สัดส่วนมีค่าเข้าใกล้ 0 หรือ 1 จึงนำเสนอช่วงความเชื่อมั่นสำหรับการประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินามแบบใหม่ คือ วิธี New จะมีลักษณะเป็นช่วงความเชื่อมั่นที่แท้จริง นั่นคือวิธี New จะเป็นวิธีการคำนวณที่ซับซ้อนไม่มีสูตรการคำนวณช่วงความเชื่อมั่นที่ชัดเจน จึงต้องเขียนโปรแกรมคอมพิวเตอร์ในการคำนวณหาช่วงความเชื่อมั่น ซึ่งปรับปรุงโดยใช้วิธี Cornish-Fisher expansion (Fisher & Cornish, 1960) ในการประมาณค่าควอนไทล์ของการแจกแจงทวินามและสร้างช่วงความเชื่อมั่นโดยใช้วิธีสกอร์ (score method) และทำการศึกษเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับการประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 7 วิธี ได้แก่ 1) วิธี New 2) วิธี Wilson score 3) วิธี Agresti-Coull 4) วิธี Jeffreys 5) วิธี Mid-P 6) วิธี modified midpoint Wilson score และ 7) วิธี randomized plausible จากผลการศึกษาพบว่า เมื่อค่าพารามิเตอร์สัดส่วนมีค่าเข้าใกล้ 0 หรือ 1 (<0.17 หรือ >0.83) ช่วงความเชื่อมั่นของวิธี New จะมีประสิทธิภาพดีที่สุด และเมื่อค่าพารามิเตอร์สัดส่วนมีค่าอยู่ในช่วง 0.17-0.83 ช่วงความเชื่อมั่นของวิธี Wilson score และวิธี modified midpoint Wilson score จะมีประสิทธิภาพดีที่สุด อย่างไรก็ตามสำหรับในกรณีค่าพารามิเตอร์สัดส่วนมีค่าอยู่ในช่วง

0.17-0.83 วิธี New จะมีประสิทธิภาพดี แต่แนะนำให้ใช้วิธี Wilson score เพราะมีสูตรการคำนวณที่ง่ายกว่าสูตรของวิธี modified midpoint Wilson score และ วิธี New Kim & Jang (2022) ได้ศึกษาประสิทธิภาพของช่วงความเชื่อมั่นของวิธี Wilson score และวิธี Agresti-Coull พบว่าช่วงความเชื่อมั่นทั้ง 2 วิธีนี้มีความเอนเอียงแบบเป็นระบบ (systematic bias) และมีความกวัดแกว่งของค่าความน่าจะเป็นกลุ่มรวม จึงได้ปรับปรุงช่วงความเชื่อมั่นจากวิธี Agresti-Coull โดยนำเสนอช่วงความเชื่อมั่นสำหรับการประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินาม คือ วิธี generalized Agresti-Coull โดยทำการปรับลดความเอนเอียงแบบเป็นระบบ ด้วยวิธี saddlepoint (saddlepoint approximation) ซึ่งช่วงความเชื่อมั่นที่นำเสนอนี้มีลักษณะเป็นช่วงความเชื่อมั่นเชิงเส้นกำกับ นั่นคือมีสูตรการคำนวณช่วงความเชื่อมั่นที่ชัดเจน และทำการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับการประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 4 วิธี ได้แก่ 1) วิธี generalized Agresti-Coull 2) วิธี Wald 3) วิธี Wilson score และ 4) วิธี Agresti-Coull จากการศึกษาพบว่า ช่วงความเชื่อมั่นของวิธี Wilson score มีประสิทธิภาพดีที่สุดสำหรับค่าความน่าจะเป็นกลุ่มรวม ถึงแม้ว่าจะมีค่าความกว้างของช่วงความเชื่อมั่นที่กว้างมากในทางกลับกันช่วงความเชื่อมั่นของวิธี Agresti-Coull จะให้ค่าความน่าจะเป็นกลุ่มรวมที่ค่อนข้างคงที่กว่าของวิธี Wilson score แต่จะมีความกว้างของช่วงความเชื่อมั่นที่มากกว่าวิธี Wilson score ส่วนวิธี generalized Agresti-Coull จะให้ค่าความน่าจะเป็นกลุ่มรวมที่ค่อนข้างคงที่กว่าของวิธี Wilson score และมีความกว้างของช่วงความเชื่อมั่นที่น้อยกว่าวิธี Agresti-Coull สำหรับทุกขนาดตัวอย่างที่ทำการศึกษา

จากการศึกษาวิธีการปรับปรุงช่วงความเชื่อมั่นแบบต่างๆ ในงานวิจัยนี้จึงสนใจศึกษาเปรียบเทียบวิธีการประมาณค่าของช่วงความเชื่อมั่นทั้ง 2 ประเภท คือ ช่วงความเชื่อมั่นเชิงเส้นกำกับและช่วงความเชื่อมั่นที่แท้จริงสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินาม โดยทำการศึกษเปรียบเทียบ 5 วิธี ได้แก่ 1) วิธี Wald 2) วิธี Wilson score 3) วิธี Agresti-Coull 4) วิธี New และ 5) วิธี generalized Agresti-Coull โดยเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่นวิธีต่างๆ

มี 2 เกณฑ์ คือ ค่าความน่าจะเป็นค้ำรวม (coverage probability) และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (average length) ซึ่งวิธีที่มีประสิทธิภาพดีจะต้องให้ค่าความน่าจะเป็นค้ำรวมเป็นไปตามค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดและมีค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุด และการประยุกต์ใช้โดยการแนะนำช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามให้เหมาะสมกับสถานการณ์ต่าง ๆ ที่ศึกษา

2. วิธีการทดลอง

การเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินาม 5 วิธี ได้แก่ 1) วิธีไวล์ด (Wald) 2) วิธี Wilson score 3) วิธี Agresti-Coull 4) วิธี New และ 5) วิธี generalized Agresti-Coull มีขั้นตอนการวิจัยดังต่อไปนี้

1. กำหนดขนาดตัวอย่าง (n) เท่ากับ 10, 30, 50 และ 100 และกำหนดค่าพารามิเตอร์สำหรับสัดส่วนเท่ากับ 0.1, 0.3, 0.5, 0.7 และ 0.9
2. กำหนดระดับความเชื่อมั่น $(1-\alpha)100\%$ เท่ากับ 95%
3. ใช้โปรแกรม R เวอร์ชัน 4.2.2 ในการจำลองข้อมูลที่มีการแจกแจงทวินามที่มีค่าพารามิเตอร์ (n, p) โดยใช้การจำลองมอนติคาร์โล (Monte Carlo simulation) ซึ่งทำการทดลองซ้ำ 10,000 ครั้งในแต่ละสถานการณ์
4. คำนวณช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามที่ศึกษาในงานวิจัยนี้ทั้งหมด 5 วิธี ดังนี้

4.1 วิธีไวล์ด (Wald)

Wald (1943) ได้เสนอวิธีการสร้างช่วงความเชื่อมั่นโดยใช้ทฤษฎีบทขีดจำกัดส่วนกลาง (central limit theorem) ภายใต้การแจกแจงปกติ โดยใช้ตัวประมาณค่าพารามิเตอร์สำหรับสัดส่วน คือ $\hat{p} = \frac{X}{n}$ ซึ่งเป็นตัวประมาณภาวน่าจะเป็นสูงสุด (maximum likelihood estimator $\hat{p}$) โดยที่ จะมีค่าเฉลี่ยเท่ากับ p และมีความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ และจากทฤษฎีบทขีดจำกัดส่วนกลาง ถ้าขนาดตัวอย่าง (n) มีขนาดใหญ่ ($n \rightarrow \infty$) จะได้ว่า $Z = \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \sim N(0,1)$ และสามารถประมาณค่าช่วงความเชื่อมั่นได้โดย

$$P\left[-Z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < Z_{\alpha/2}\right] = 1 - \alpha \text{ กำหนดให้}$$

ขีดจำกัดล่างและขีดจำกัดบนของช่วงความเชื่อมั่นไวล์ดเขียนแทนด้วยสัญลักษณ์ L_W และ U_W ตามลำดับ ดังนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับสัดส่วน (p) ของประชากรที่มีการแจกแจงทวินามของวิธีไวล์ด เป็นดังนี้

$$L_W \text{ คือ } \hat{p} - Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$U_W \text{ คือ } \hat{p} + Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

เมื่อ $\hat{p}$ คือ สัดส่วนของการเกิดสิ่งที่สนใจในตัวอย่าง ซึ่ง $\hat{p} = \frac{X}{n}$; X แทน จำนวนครั้งของการเกิดสิ่งที่สนใจในตัวอย่างสุ่มขนาด n และ $Z_{\alpha/2}$ คือ ค่าควอนไทล์ที่

$1-\alpha/2$ ของการแจกแจงปกติมาตรฐาน

4.2 วิธี Wilson score

Wilson (1927) ได้พัฒนาช่วงความเชื่อมั่นโดยใช้แนวคิดของวิธีสกอร์ (score method) โดยสร้างช่วงความเชื่อมั่นภายใต้การทดสอบสมมติฐาน $H_0: p = p_0$ เทียบกับ $H_1: p \neq p_0$ ที่ระดับนัยสำคัญ α และจะยอมรับสมมติฐานหลัก H_0 เมื่อ

$$\frac{|\hat{p} - p_0|}{\sqrt{\frac{p_0(1-p_0)}{n}}} < Z_{\alpha/2} \quad (1)$$

จะสามารถหาค่าพารามิเตอร์ p_0 โดยกำหนดให้ $p = p_0$ ได้จากการจัดสมการที่ (1) ให้อยู่ในรูปของสมการกำลังสอง (quadratic equation) แล้วแก้สมการหาค่าพารามิเตอร์ p_0 กำหนดให้ขีดจำกัดล่างและขีดจำกัดบนของช่วงความเชื่อมั่นวิธี Wilson score เขียนแทนด้วยสัญลักษณ์ L_{WS} และ U_{WS} ตามลำดับ ดังนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับสัดส่วน (p) ของประชากรที่มีการแจกแจงทวินามของวิธี Wilson score เป็นดังนี้

$$L_{WS} \text{ คือ } \frac{\left(\hat{p} + \frac{Z_{\alpha/2}^2}{2n}\right) - Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{Z_{\alpha/2}^2}{4n^2}}}{1 + \frac{Z_{\alpha/2}^2}{n}}$$

$$U_{WS} \text{ คือ } \frac{\left(\hat{p} + \frac{Z_{\alpha/2}^2}{2n}\right) + Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{Z_{\alpha/2}^2}{4n^2}}}{1 + \frac{Z_{\alpha/2}^2}{n}}$$

เมื่อ $\hat{p}$ คือ สัดส่วนของการเกิดสิ่งที่สนใจในตัวอย่าง
ซึ่ง $\hat{p} = \frac{X}{n}$; X แทน จำนวนครั้งของการเกิดสิ่งที่สนใจ
ในตัวอย่างสุ่มขนาด n และ $Z_{\alpha/2}$ คือ ค่าควอนไทล์ที่
 $1-\alpha/2$ ของการแจกแจงปกติมาตรฐาน

4.3 วิธี Agresti-Coull

Agresti & Coull (1998) ได้ทำการปรับปรุงช่วงความเชื่อมั่นจากวิธีวิลด์โดยใช้หลักการเพิ่มค่า สังเกต 4 ค่า นั่นคือ การเพิ่มผลลัพธ์ของค่าสังเกตที่สนใจและไม่สนใจอย่างละ 2 ค่า (adding two successes and two failures) ในจำนวนครั้งของการทดลอง (n) ทำให้จำนวนครั้งของผลลัพธ์ที่เป็นไปได้ทั้งหมดเพิ่มขึ้นอีก 4 ครั้ง และกำหนดให้ $Z_{\alpha/2} \approx 2$ ซึ่งช่วงความเชื่อมั่นของวิธี Agresti-Coull กับช่วงความเชื่อมั่นของวิธีวิลด์จะแตกต่างกันที่ค่าประมาณของพารามิเตอร์สำหรับสัดส่วน จะเห็นได้ว่า
 $\tilde{p} = \frac{X+2}{n+4}$ กำหนดให้ขีดจำกัดล่างและขีดจำกัดบนของช่วงความเชื่อมั่นวิธี Agresti-Coull เขียนแทนด้วยสัญลักษณ์ L_{AC} และ U_{AC} ตามลำดับ ดังนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับสัดส่วน (p) ของประชากรที่มีการแจกแจงทวินามของวิธี Agresti-Coull เป็นดังนี้

$$L_{AC} \text{ คือ } \tilde{p} - Z_{\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$$

$$U_{AC} \text{ คือ } \tilde{p} + Z_{\alpha/2} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}$$

เมื่อ $\tilde{p}$ คือ สัดส่วนของการเกิดสิ่งที่สนใจในตัวอย่าง ซึ่ง
 $\tilde{p} = \frac{X+2}{n+4}$; X แทน จำนวนครั้งของการเกิดสิ่งที่สนใจ
ในตัวอย่างสุ่มขนาด n , $\tilde{n} = n+4$ และ $Z_{\alpha/2}$ คือ
ค่าควอนไทล์ที่ $1-\alpha/2$ ของการแจกแจงปกติมาตรฐาน

4.4 วิธี New

Cao et al. (2022) ได้พัฒนาช่วงความเชื่อมั่นจากวิธี Wilson score จึงนำเสนอช่วงความเชื่อมั่นสำหรับการประมาณค่าสัดส่วนของประชากรที่มีการแจกแจงทวินามคือ วิธี New ซึ่งปรับปรุงโดยใช้วิธี Cornish-Fisher expansion (Fisher & Cornish, 1960) ในการประมาณค่าควอนไทล์ของการแจกแจงทวินาม โดยจะใช้คิวมุลแลนต์ (cumulant) ที่ 3 (κ_3) และ 4 (κ_4) ของการแจกแจงทวินามในการประมาณค่าควอนไทล์ที่ q (q -quantile (λ_q)) นั่นคือ

$$\lambda_q \approx z_q + \frac{z_q^2 - 1}{6} \kappa_3 + \frac{z_q^3 - 3z_q}{24} \kappa_4 - \frac{2z_q^3 - 5z_q}{36} \kappa_3^2$$

$$\kappa_3 = (1-2\pi) / \sqrt{n\pi(1-\pi)}$$

$$\kappa_4 = (1-6\pi(1-\pi)) / \sqrt{n\pi(1-\pi)}$$

และสร้างช่วงความเชื่อมั่นโดยใช้วิธีสโคร์ (score method) กำหนดให้ขีดจำกัดล่างและขีดจำกัดบนของช่วงความเชื่อมั่นวิธี New เขียนแทนด้วยสัญลักษณ์ L_N และ U_N ตามลำดับ ดังนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับสัดส่วน (π) ของประชากรที่มีการแจกแจงทวินามของวิธี New เป็นดังนี้

$$L_N \text{ คือ } \frac{p-\pi}{\sqrt{\pi(1-\pi)/n}} = \lambda_{1-\alpha/2}$$

$$U_N \text{ คือ } \frac{p-\pi}{\sqrt{\pi(1-\pi)/n}} = \lambda_{\alpha/2}$$

ช่วงความเชื่อมั่นของวิธี New จะมีวิธีการคำนวณที่ซับซ้อน ไม่มีสูตรการคำนวณช่วงความเชื่อมั่นที่ชัดเจน จึงสามารถใช้โค้ดโปรแกรม R จากงานวิจัยของ Cao et al. (2022) ในการคำนวณหาช่วงความเชื่อมั่น

4.5 วิธี generalized Agresti-Coull

Kim & Jang (2022) ได้นำเสนอช่วงความเชื่อมั่นวิธี generalized Agresti-Coull ซึ่งสร้างโดยใช้วิธีการประมาณ saddlepoint (saddlepoint approximation) และมีพารามิเตอร์ 3 ค่าซึ่งเป็นองค์ประกอบในการสร้างช่วงความเชื่อมั่นเพื่อปรับลดค่าความเอนเอียง (bias) ระหว่างค่าความน่าจะเป็นคัมรวม (coverage probability) และระดับความเชื่อมั่น (confidence level) ซึ่งพารามิเตอร์ทั้ง 3 ค่านี้ ได้แก่ a, c และ d การปรับค่าความเอนเอียงของพารามิเตอร์จะขึ้นอยู่กับขนาดตัวอย่าง (n) โดยที่พารามิเตอร์ a จะควบคุมตำแหน่งของค่าประมาณแบบจุดของช่วงความเชื่อมั่น ส่วนพารามิเตอร์ c และ d จะควบคุมค่าขีดจำกัดบนและขีดจำกัดล่างของช่วงความเชื่อมั่นของค่าประมาณแบบจุดแต่ละค่า กำหนดให้ขีดจำกัดล่างและขีดจำกัดบนของช่วงความเชื่อมั่นวิธี generalized Agresti-Coull เขียนแทนด้วยสัญลักษณ์ L_{gAC} และ U_{gAC} ตามลำดับ ดังนั้นช่วงความเชื่อมั่น $(1-\alpha)100\%$ สำหรับสัดส่วน (p) ของประชากรที่มีการแจกแจงทวินามของวิธี generalized Agresti-Coull เป็นดังนี้

$$L_{gAC} \text{ คือ } \tilde{p}_a - Z_{\alpha/2} \sqrt{\frac{\tilde{p}_c(1-\tilde{p}_c)}{n+d}}$$

$$U_{gAC} \text{ คือ } \tilde{p}_a + Z_{\alpha/2} \sqrt{\frac{\tilde{p}_c(1-\tilde{p}_c)}{n+d}}$$

เมื่อ $\tilde{p}_a$ และ $\tilde{p}_c$ คือ สัดส่วนของการเกิดสิ่งที่สนใจในตัวอย่าง ซึ่ง $\tilde{p}_a = \frac{X+a}{n+2a}$ และ $\tilde{p}_c = \frac{X+c}{n+2c}$; X แทนจำนวนครั้งของการเกิดสิ่งที่สนใจในตัวอย่างสุ่มขนาด n และ $Z_{\alpha/2}$ คือ ค่าควอนไทล์ที่ $1-\alpha/2$ ของการแจกแจงปกติมาตรฐาน

สำหรับช่วงความเชื่อมั่น 95% สำหรับสัดส่วน (p) ของประชากรที่มีการแจกแจงทวินาม ค่าพารามิเตอร์

$$a = 2.1 - \frac{1}{\sqrt{n}}, c = 1.7 - \frac{2}{\sqrt{n}}, d = 4 - \frac{1}{\sqrt{n}}$$

5. คำนวณค่าประมาณความน่าจะเป็นคัมรวม ($\hat{c}$) สำหรับช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี สามารถคำนวณได้ดังนี้

$$\text{ค่าประมาณความน่าจะเป็นคัมรวม} = \frac{\text{จำนวนครั้งทั้งหมดที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ } p}{M}$$

เมื่อ M คือ จำนวนรอบในการทำซ้ำ ในที่นี้ $M=10,000$ รอบ

6. เปรียบเทียบค่าประมาณความน่าจะเป็นคัมรวม ($\hat{c}$) ที่ได้สำหรับช่วงความเชื่อมั่นทั้ง 5 วิธี ว่ามีค่าเท่ากับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (c_0) อย่างมีนัยสำคัญหรือไม่ ทำได้โดยการทดสอบสมมติฐานทางสถิติดังนี้

$H_0 : c \geq c_0$ (ค่าความน่าจะเป็นคัมรวมมีค่าไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด)

$H_1 : c < c_0$ (ค่าความน่าจะเป็นคัมรวมมีค่าต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด)

$$\text{ตัวสถิติทดสอบ คือ } Z = \frac{\hat{c} - c_0}{\sqrt{\frac{c_0(1-c_0)}{M}}}$$

เมื่อ M คือ จำนวนรอบในการทำซ้ำเท่ากับ 10,000 รอบ

$$\text{จะปฏิเสธสมมติฐาน } H_0 \text{ เมื่อ } \frac{\hat{c} - c_0}{\sqrt{\frac{c_0(1-c_0)}{M}}} < -Z_\alpha$$

$$\text{หรือ } \hat{c} < c_0 - Z_\alpha \sqrt{\frac{c_0(1-c_0)}{M}}$$

เมื่อกำหนดระดับนัยสำคัญของการทดสอบ (α) เท่ากับ 0.05 และค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด (c_0) คือ 0.95 จะได้ว่า

$$\hat{c} < 0.95 - 1.645 \sqrt{\frac{0.95(1-0.95)}{10,000}} \text{ หรือ } \hat{c} < 0.946$$

ดังนั้น วิธีการประมาณค่าช่วงความเชื่อมั่นที่ให้ค่าประมาณความน่าจะเป็นคัมรวม ($\hat{c}$) ไม่ต่ำกว่า

ค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ($c_0 = 0.95$) เมื่อค่าประมาณความน่าจะเป็นคัมรวมมีค่าไม่ต่ำกว่า 0.946 ($\hat{c} \geq 0.946$)

ถ้าช่วงความเชื่อมั่นใดให้ค่าประมาณความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด แสดงว่าช่วงความเชื่อมั่นนั้นมีประสิทธิภาพ และจะพิจารณาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น แต่ถ้าช่วงความเชื่อมั่นใดให้ค่าประมาณความน่าจะเป็นคัมรวมต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด แสดงว่าช่วงความเชื่อมั่นนั้นไม่มีประสิทธิภาพ และจะไม่นำมาพิจารณาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น

7. คำนวณค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (average length; AL) โดยคำนวณค่าขีดจำกัดล่างและขีดจำกัดบนของช่วงความเชื่อมั่นทั้ง 5 วิธี แล้วนำไปคำนวณโดยใช้สูตร

$$AL = \frac{\sum_{i=1}^{10,000} (U_i - L_i)}{10,000}$$

เมื่อ U_i และ L_i คือ ขีดจำกัดบนและขีดจำกัดล่างของช่วงความเชื่อมั่น ในการทำซ้ำรอบที่ i

โดยที่ $i = 1, 2, \dots, 10,000$

ถ้าช่วงความเชื่อมั่นใดให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุด แสดงว่าช่วงความเชื่อมั่นนั้นมีประสิทธิภาพมากที่สุด

8. เปรียบเทียบค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นทั้ง 5 วิธี

9. สรุปผลการศึกษาของช่วงความเชื่อมั่นทั้ง 5 วิธีในแต่ละสถานการณ์

3. ผลและวิจารณ์ผลการวิจัย

ผลการศึกษางานวิจัยนี้ได้ทำการเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินาม 5 วิธี ได้แก่ 1) วิธีวัลด์ (Wald) 2) วิธี Wilson score 3) วิธี Agresti-Coull 4) วิธี New และ 5) วิธี generalized Agresti-Coull ซึ่งเกณฑ์ที่ใช้ในการเปรียบเทียบประสิทธิภาพมี 2 เกณฑ์ คือ 1) ค่าความน่าจะเป็นคัมรวม และ 2) ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น ซึ่งวิธีที่มีประสิทธิภาพดีจะต้องให้ค่าความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด นั่นคือวิธีการประมาณค่าช่วงความเชื่อมั่นที่ให้ค่าประมาณความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

0.95 เมื่อค่าประมาณความน่าจะเป็นคัมรวมมีค่าไม่ต่ำกว่า 0.946 และมีค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุด โดยการใช้การจำลองมอนติคาร์โลในสถานการณ์ต่าง ๆ ที่ศึกษา ดังแสดงในตารางที่ 1 และ 2 ตามลำดับ

ตารางที่ 1 แสดงค่าประมาณความน่าจะเป็นคัมรวมของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี ผลการศึกษาพบว่าวิธี Wald (Wald) ให้ค่าประมาณความน่าจะเป็นคัมรวมต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในหลายสถานการณ์ ยกเว้นในกรณีที่ $n=30$ เมื่อ $p=0.3-0.7$ และ $n=100$ เมื่อ $p=0.3$ และ 0.7 และมีแนวโน้มที่จะให้ค่าประมาณความน่าจะเป็นคัมรวมที่มากขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น วิธี Wilson score จะให้ค่าประมาณความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด ในกรณีที่ $n=10$ เมื่อ $p=0.5$ $n=30$ เมื่อ $p=0.1, 0.5$ และ 0.9 และ $n=50$ สำหรับทุกค่าของ p ยกเว้น $p=0.5$ ช่วงความเชื่อมั่นวิธี Wald (Wald) และวิธี Wilson score จะให้ค่าประมาณความน่าจะเป็นคัมรวมที่มีความกว้างสูงในกรณีที่ค่าพารามิเตอร์สัดส่วนมีค่าเข้าใกล้ 0 หรือ 1 ส่วนวิธี Agresti-Coull ให้ค่าประมาณความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในหลายสถานการณ์ ยกเว้นในกรณีที่ $n=10$ เมื่อ $p=0.1$ และ 0.9 และ $n=50$ และ 100 เมื่อ $p=0.5$ สำหรับวิธี New จะให้ค่าประมาณความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในหลายสถานการณ์ที่ศึกษา ยกเว้นในกรณีที่ $n=10$ เมื่อ $p=0.3$ และ 0.7 $n=30$ เมื่อ $p=0.1$ และ 0.9 $n=50$ เมื่อ $p=0.1, 0.5$ และ 0.9 และ $n=100$ เมื่อ $p=0.5$ และวิธี generalized Agresti-Coull จะให้ค่าประมาณความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดในหลายสถานการณ์ ยกเว้นในกรณีที่ $n=10$ สำหรับทุกค่าของ p ยกเว้น $p=0.5$ $n=30$ เมื่อ $p=0.3$ และ 0.7 และ $n=50$ และ 100 เมื่อ $p=0.5$ ซึ่งผลการศึกษาสอดคล้องกับกราฟในรูปที่ 1 แสดงค่าประมาณความน่าจะเป็นคัมรวมของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี โดยที่ W แทนวิธี Wald, WS แทนวิธี Wilson score, AC แทนวิธี Agresti-Coull, N แทนวิธี New และ gAC แทนวิธี generalized Agresti-Coull จะพบว่าเมื่อขนาดตัวอย่างใหญ่ขึ้นจะทำให้ช่วงความเชื่อมั่นทั้ง 5 วิธี มีแนวโน้มให้ค่าประมาณความน่าจะเป็นคัมรวมเพิ่มขึ้น

ตารางที่ 2 แสดงค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี ผลการศึกษาพบว่าช่วงความเชื่อมั่นทั้ง 5 วิธีให้ค่าประมาณความกว้างเฉลี่ยของช่วงความเชื่อมั่นลดลงเมื่อตัวอย่างมีขนาดใหญ่ขึ้นในทุกระดับของค่า p วิธี Wilson score ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดในกรณีที่ $n=10$ เมื่อ $p=0.5$ $n=30$ เมื่อ $p=0.1$ และ 0.9 และ $n=50$ สำหรับทุกค่าของ p ยกเว้น $p=0.5$ วิธี Agresti-Coull จะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดในกรณีที่ $n=10$ เมื่อ $p=0.3$ และ 0.7 และ $n=30$ เมื่อ $p=0.3-0.7$ ส่วนวิธี New ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดในกรณีที่ $n=10$ และ 100 เมื่อ $p=0.1$ และ 0.9 และวิธี generalized Agresti-Coull จะให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดในกรณีที่ $n=100$ เมื่อ $p=0.3$ และ 0.7 ซึ่งผลการศึกษาสอดคล้องกับกราฟในรูปที่ 2 แสดงค่าประมาณความกว้างเฉลี่ยของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี โดยที่ W แทนวิธี Wald, WS แทนวิธี Wilson score, AC แทนวิธี Agresti-Coull, N แทนวิธี New และ gAC แทนวิธี generalized Agresti-Coull

4. การประยุกต์ใช้กับข้อมูลจริง

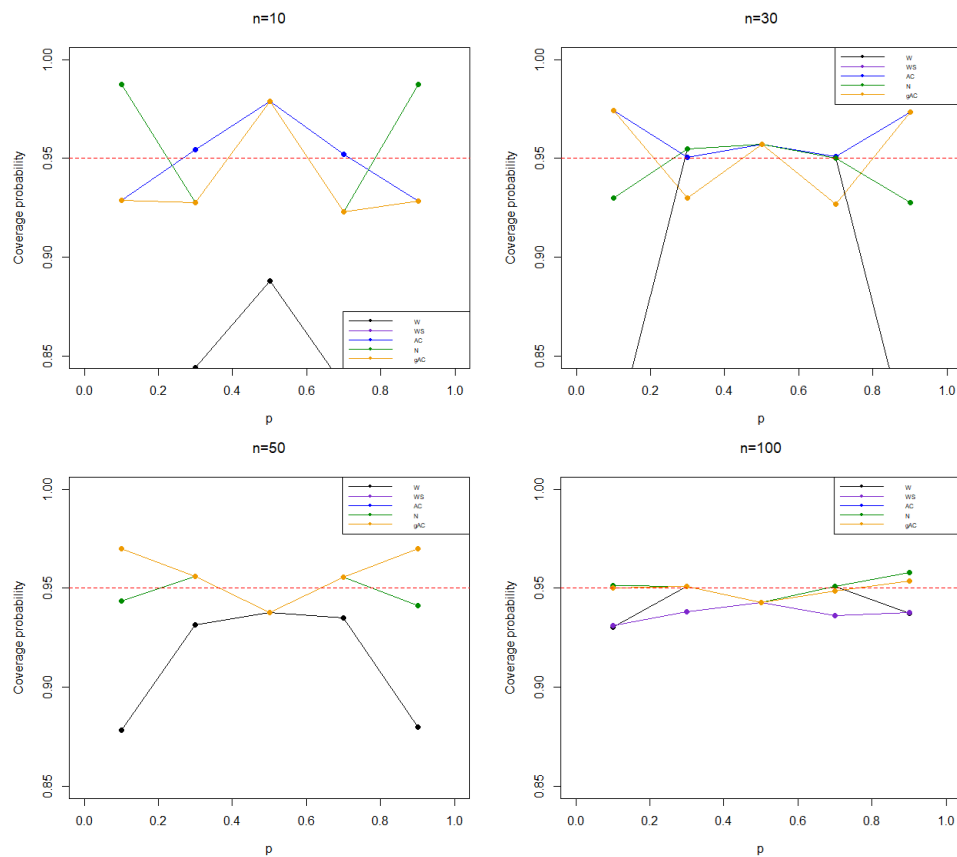
การศึกษาเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี สามารถนำมาประยุกต์ใช้กับข้อมูลในสถานการณ์จริง ซึ่งข้อมูลที่น่านำมาประยุกต์ใช้ในงานวิจัยนี้มาจากงานวิจัยของ Mizumoto et al. (2020) ได้ศึกษาข้อมูลเกี่ยวกับผู้โดยสารในเรือสำราญ Diamond Princess ที่ล่องเรืออยู่ในเมืองโยโกฮามะ ประเทศญี่ปุ่น ซึ่งมีผู้โดยสารในเรือได้ตรวจพบผู้ติดเชื้อไวรัสโคโรนา (COVID-19) จำนวน 634 คน ซึ่งพบว่าเป็นชาวญี่ปุ่นจำนวน 270 คน เป็นชาวอเมริกันจำนวน 88 คน และที่เหลือเป็นบุคคลที่มีสัญชาติต่าง ๆ จาก 28 ประเทศ

ในงานวิจัยนี้ได้นำข้อมูลมาทำการศึกษา 2 กรณี ได้แก่ กรณี 1) ศึกษาช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของผู้โดยสารที่ติดเชื้อไวรัสโคโรนาในเรือสำราญที่เป็นชาวอเมริกัน และกรณี 2) ศึกษาช่วงความเชื่อมั่น

ตารางที่ 1 ค่าประมาณความน่าจะเป็นคัมรวมของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี

ขนาดตัวอย่าง (n)	พารามิเตอร์ p	วิธี				
		Wald	Wilson score	Agresti-Coull	New	Generalized Agresti-Coull
10	0.1	0.6540	0.9288	0.9288	0.9876*	0.9288
	0.3	0.8444	0.9278	0.9544*	0.9278	0.9278
	0.5	0.8881	0.9788*	0.9788*	0.9788*	0.9788*
	0.7	0.8350	0.9230	0.9521*	0.9230	0.9230
	0.9	0.6469	0.9284	0.9284	0.9874*	0.9284
30	0.1	0.8066	0.9744*	0.9744*	0.9302	0.9744*
	0.3	0.9548*	0.9300	0.9507*	0.9548*	0.9300
	0.5	0.9573*	0.9573*	0.9573*	0.9573*	0.9573*
	0.7	0.9502*	0.9270	0.9508*	0.9502*	0.9270
	0.9	0.8023	0.9733*	0.9733*	0.9275	0.9733*
50	0.1	0.8783	0.9701*	0.9701*	0.9436	0.9701*
	0.3	0.9315	0.9560*	0.9560*	0.9560*	0.9560*
	0.5	0.9377	0.9377	0.9377	0.9377	0.9377
	0.7	0.9349	0.9555*	0.9555*	0.9555*	0.9555*
	0.9	0.8799	0.9701*	0.9701*	0.9411	0.9701*
100	0.1	0.9305	0.9313	0.9500*	0.9513*	0.9500*
	0.3	0.9509*	0.9383	0.9509*	0.9509*	0.9509*
	0.5	0.9428	0.9428	0.9428	0.9428	0.9428
	0.7	0.9510*	0.9361	0.9488*	0.9510*	0.9488*
	0.9	0.9373	0.9378	0.9535*	0.9581*	0.9535*

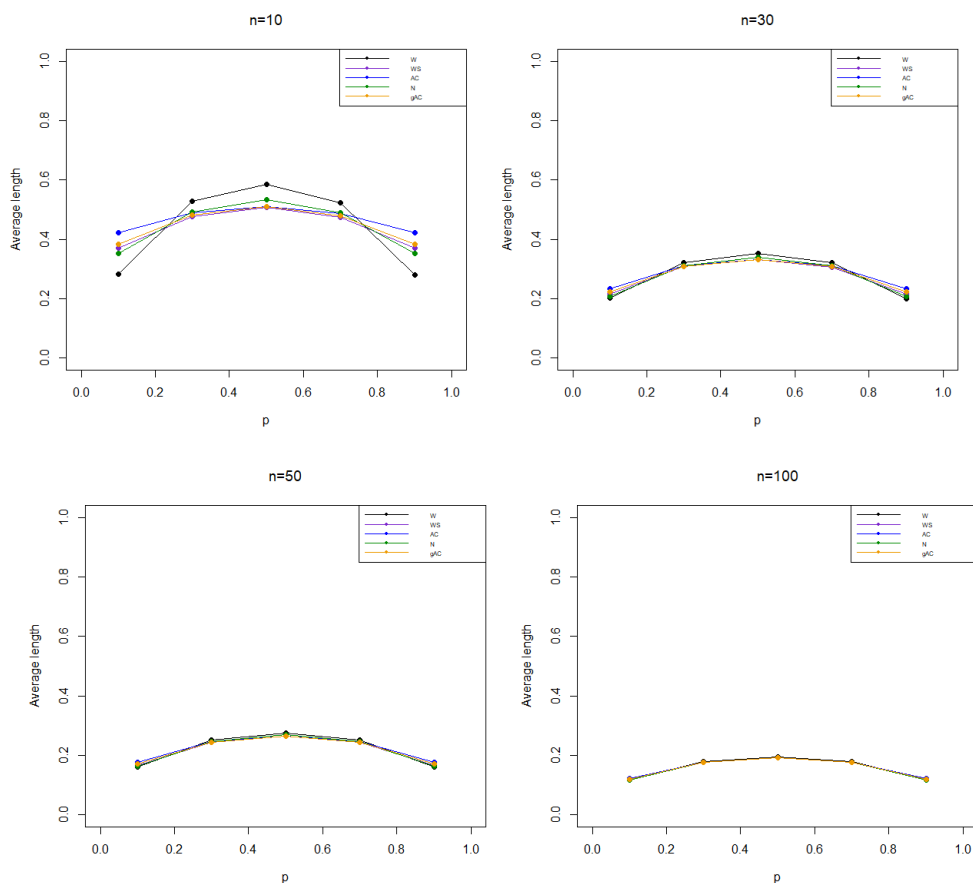
หมายเหตุ: * หมายถึง วิธีการประมาณช่วงความเชื่อมั่นให้ค่าความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

รูปที่ 1 ค่าประมาณความน่าจะเป็นคัมรวมของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี กรณีขนาดตัวอย่าง (n) แตกต่างกัน

ตารางที่ 2 ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี

ขนาดตัวอย่าง (n)	พารามิเตอร์ p	วิธี				
		Wald	Wilson score	Agresti-Coull	New	Generalized Agresti-Coull
10	0.1	0.2818	0.3696	0.4217	0.3534**	0.3839
	0.3	0.5283	0.4762	0.4896**	0.4932	0.4824
	0.5	0.5856	0.5066**	0.5099	0.5330	0.5108
	0.7	0.5235	0.4737	0.4880**	0.4900	0.4801
	0.9	0.2797	0.3691	0.4214	0.3527**	0.3835
30	0.1	0.2011	0.2148**	0.2335	0.2080	0.2229
	0.3	0.3216	0.3070	0.3102**	0.3121	0.3087
	0.5	0.3517	0.3318	0.3317**	0.3393	0.3322
	0.7	0.3212	0.3067	0.3100**	0.3118	0.3084
	0.9	0.2000	0.2141**	0.2330	0.2071	0.2223
50	0.1	0.1607	0.1662**	0.1761	0.1630	0.1710
	0.3	0.2509	0.2437**	0.2453	0.2463	0.2446
	0.5	0.2744	0.2646	0.2644	0.2682	0.2647
	0.7	0.2509	0.2437**	0.2453	0.2462	0.2446
	0.9	0.1612	0.1666**	0.1765	0.1634	0.1714
100	0.1	0.1157	0.1175	0.1214	0.1166**	0.1196
	0.3	0.1786	0.1759	0.1765	0.1768	0.1763**
	0.5	0.1950	0.1914	0.1913	0.1925	0.1914
	0.7	0.1785	0.1759	0.1764	0.1767	0.1762**
	0.9	0.1160	0.1178	0.1217	0.1169**	0.1199

หมายเหตุ: ** หมายถึง วิธีการประมาณช่วงความเชื่อมั่นให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุดโดยพิจารณาจากช่วงความเชื่อมั่นที่ให้ค่าความน่าจะเป็นคัมรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดซึ่งแสดงในตารางที่ 1



รูปที่ 2 ค่าประมาณความกว้างเฉลี่ยของช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามทั้ง 5 วิธี กรณีขนาดตัวอย่างแตกต่างกัน

ตารางที่ 3 ช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินาม และค่าความกว้างของช่วงความเชื่อมั่นทั้ง 5 วิธี เมื่อนำมาประยุกต์ใช้กับข้อมูลจริง

วิธีการประมาณค่าช่วงความเชื่อมั่น	กรณีศึกษา			
	กรณี 1 ($p = 88/634 = 0.14$)		กรณี 2 ($p = 270/634 = 0.43$)	
	ช่วงความเชื่อมั่น 95%	ค่าความกว้างของช่วงความเชื่อมั่น	ช่วงความเชื่อมั่น 95%	ค่าความกว้างของช่วงความเชื่อมั่น
วิธี Wald (Wald)	(0.1119, 0.1657)	0.0538	(0.3874, 0.4644)	0.0770
วิธี Wilson score	(0.1141, 0.1679)	0.0538	(0.3879, 0.4647)	0.0768
วิธี Agresti-Coull	(0.1141, 0.1681)	0.0540	(0.3880, 0.4647)	0.0767
วิธี New	(0.1135, 0.1674)	0.0539	(0.3878, 0.4646)	0.0767
วิธี generalized Agresti-Coull	(0.1142, 0.1681)	0.0540	(0.3880, 0.4647)	0.0767

ตารางที่ 4 วิธีการประมาณค่าช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามที่เหมาะสมกับสถานการณ์ต่าง ๆ ที่ศึกษา

ขนาดตัวอย่าง (n)	พารามิเตอร์ p				
	0.1	0.3	0.5	0.7	0.9
10	N	AC	WS	AC	N
30	WS	AC	AC	AC	WS
50	WS	WS	-	WS	WS
100	N	gAC	-	gAC	N

หมายเหตุ: WS แทน วิธี Wilson score; AC แทน วิธี Agresti-Coull; N แทน วิธี New และ gAC แทน วิธี generalized Agresti-Coull

- แทน ไม่มีวิธีการประมาณค่าช่วงความเชื่อมั่นใดที่เหมาะสม

ความเชื่อมั่น 95% และค่าความกว้างของช่วงความเชื่อมั่นทั้ง 5 วิธี สำหรับสัดส่วนของผู้โดยสารที่ติดเชื้อไวรัสโคโรนาในเรือสำราญที่เป็นชาวอเมริกันและชาวญี่ปุ่น ดังแสดงในตารางที่ 3

ผลจากตารางที่ 3 จะพบว่าวิธีการประมาณค่าช่วงความเชื่อมั่นทั้ง 5 วิธี ในกรณีที่ศึกษาทั้ง 2 กรณี คือกรณีที่ตัวอย่างมีขนาดใหญ่ ($n = 634$) และค่า p มีค่าเข้าใกล้ 0 ($p = 0.14$) และค่า $p = 0.43$ จะให้ค่าประมาณช่วงความเชื่อมั่น 95% และค่าความกว้างของช่วงความเชื่อมั่นสำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินามที่ใกล้เคียงกันในทุกกรณี เนื่องจากตัวอย่างที่นำมาศึกษามีขนาดใหญ่ ซึ่งจะพบว่าจะมีความสอดคล้องกับผลการศึกษาเชิงจำลองในกรณีที่ตัวอย่างมีขนาดใหญ่ ช่วงความเชื่อมั่นทั้ง 5 วิธีจะมีประสิทธิภาพมากขึ้นในการคำนวณหาค่าประมาณช่วงความเชื่อมั่น 95% และค่าความกว้างของช่วงความเชื่อมั่นจึงมีค่าที่ใกล้เคียงกันในทุกกรณี สำหรับการศึกษกรณี 1 ช่วงความเชื่อมั่นวิธี Wald และวิธี Wilson score จะมีค่าประมาณความกว้างของช่วงความเชื่อมั่นน้อยที่สุดจึงมีประสิทธิภาพมากที่สุด รองลงมาคือช่วงความเชื่อมั่นวิธี New วิธี Agresti-Coull และวิธี generalized Agresti-Coull ส่วนกรณีที่ศึกษาที่ 2 ช่วงความ

เชื่อมั่นวิธี Agresti-Coull วิธี New และวิธี generalized Agresti-Coull จะให้ค่าประมาณความกว้างของช่วงความเชื่อมั่นน้อยที่สุดจึงมีประสิทธิภาพมากที่สุด รองลงมาคือช่วงความเชื่อมั่น วิธี Wilson score และวิธี Wald

สรุปผลการวิจัย

ผลจากการศึกษาเปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินาม 5 วิธี ได้แก่ 1) วิธี Wald 2) วิธี Wilson score 3) วิธี Agresti-Coull 4) วิธี New และ 5) วิธี generalized Agresti-Coull โดยพิจารณาประสิทธิภาพของช่วงความเชื่อมั่นที่ให้ค่าความน่าจะเป็นคุ่มรวมไม่ต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดไว้คือ 0.95 และให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุด จะพบว่าช่วงความเชื่อมั่นที่เหมาะสมกับสถานการณ์ต่าง ๆ ที่ศึกษาโดยจำแนกตามขนาดตัวอย่าง (n) และ ค่าพารามิเตอร์ p แสดงดังตารางที่ 4

ตารางที่ 4 แสดงวิธีการประมาณค่าช่วงความเชื่อมั่น 95% สำหรับสัดส่วนของประชากรที่มีการแจกแจงทวินาม

ที่เหมาะสมกับสถานการณ์ต่าง ๆ ที่ศึกษาโดยจำแนกตามขนาดตัวอย่าง (n) และค่าพารามิเตอร์ p จะพบว่าช่วงความเชื่อมั่นวิธี Wilson score มีประสิทธิภาพดีที่สุดใหลายสถานการณ์ที่ศึกษา ซึ่งช่วงความเชื่อมั่นนี้เหมาะกับสถานการณ์ที่ตัวอย่างมีขนาดเล็ก $n=10$ เมื่อ $p=0.5$ และตัวอย่างขนาดกลาง $n=30$ เมื่อ p มีค่าเข้าใกล้ 0 หรือ 1 ($p=0.1, 0.9$) และ $n=50$ สำหรับทุกค่าของ p ยกเว้น $p=0.5$ ช่วงความเชื่อมั่นวิธี Agresti-Coull มีประสิทธิภาพเป็นอันดับที่สอง ซึ่งเหมาะกับตัวอย่างขนาดเล็ก $n=10$ เมื่อ $p=0.3$ และ 0.7 และตัวอย่างขนาดกลาง $n=30$ เมื่อ $p=0.3-0.7$ ส่วนช่วงความเชื่อมั่นวิธี New มีประสิทธิภาพเป็นอันดับที่สาม ซึ่งเหมาะกับตัวอย่างขนาดเล็ก $n=10$ และตัวอย่างขนาดใหญ่ $n=100$ เมื่อ p มีค่าเข้าใกล้ 0 หรือ 1 ($p=0.1, 0.9$) ช่วงความเชื่อมั่นวิธี generalized Agresti-Coull มีประสิทธิภาพเป็นอันดับที่สี่ ซึ่งเหมาะกับตัวอย่างขนาดใหญ่ $n=100$ เมื่อ $p=0.3$ และ 0.7 และช่วงความเชื่อมั่นวิธี Wald มีประสิทธิภาพที่ไม่ดีนักเนื่องจากให้ค่าความน่าจะเป็นคุ่มรวมต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนดไว้ในหลายสถานการณ์ที่ศึกษาแต่มีแนวโน้มให้ประสิทธิภาพดีเมื่อตัวอย่างมีขนาดใหญ่ นอกจากนี้การประยุกต์ช่วงความเชื่อมั่นทั้ง 5 วิธีนำไปใช้กับข้อมูลจริงในสถานการณ์ต่าง ๆ จะให้ผลการศึกษาประสิทธิภาพของช่วงความเชื่อมั่นแต่ละวิธีมีความสอดคล้องกับผลการศึกษาเชิงจำลอง

เอกสารอ้างอิง

- Agresti, A., & Coull, B. A. (1998). Approximate is better than "Exact" for interval estimation of binomial proportions. *The American Statistician*, 52(2), 118-126. <https://doi.org/10.1080/00031305.1998.10480550>
- Brown, L. D., Cai, T., & Dasgupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, 16(2), 101-133. <https://doi.org/10.1214/ss/1009213286>
- Brown, L. D., Cai, T., & Dasgupta, A. (2002). Confidence intervals for a binomial proportion and asymptotic expansions. *The Annals of Statistics*, 30(1), 160-201. <https://doi.org/10.1214/aos/1015362189>
- Cao, Y., Xu, Y., Tan, M. T., Chen, P., & Duan, C. (2022). A simple and improved score confidence interval for a single proportion. *Communications in Statistics-Theory and Methods*, 51(8), 2659-2675. <https://doi.org/10.1080/03610926.2020.1779747>
- Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4), 404-413.
- Fisher, A. R., & Cornish, E. A. (1960). The percentile points of distributions having known cumulants. *American Society for Quality*, 2(2), 209-225.
- Kim, J., & Jang, W. (2022). A generalized Agresti-Coull type confidence interval for a binomial proportion. *Journal of the Korean Statistical Society*, 51, 356-377. <https://doi.org/10.1007/s42952-021-00143-3>
- Komoltri, J. (2012). Sample size calculation. *Journal of Mental Health of Thailand*, 20(3), 192-198. (in Thai)
- Mizumoto, K., Kagaya, K., Zarebski, A., & Chowell, G. (2020). Estimating the Asymptomatic Proportion of Coronavirus Disease 2019 (COVID-19) Case on Board the Diamond Princess Cruise Ship, Yokohama, Japan, 2020. *Euro Surveill*, 25(10), 1-5. <https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180>
- Muenmee, S., & Bootdee, S. (2021). Health risk assessment of exposure PM2.5 from industrial area in Pluak Daeng District, Rayong province. *Naresuan Phayao Journal*, 14(3), 95-110. (in Thai)
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is Large. *American Mathematical Society*, 54(3), 426-482.
- Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158), 201-212. <https://doi.org/10.2307/2276774>