

การใช้สารสนเทศจากตัวแปรช่วยเพื่อพัฒนาตัวประมาณอัตราส่วนแบบเลขชี้กำลัง ภายใต้ปัญหาการไม่ตอบสนอง ในการสำรวจตัวอย่าง

Utilizing Auxiliary Information to Develop an Exponential Ratio-type Estimator under the Problem of Non-Response in Sample Surveys

ณภัทรจันทร์ ด่านสวัสดิ์* และ เอกภพ เกตุสมบุญ

Napattchan Dansawad* and Ekkaphop Ketsombun

คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ประเทศไทย

Faculty of Science, King Mongkut's University of Technology, Thailand

Received : 1 September 2025, Received in revised form : 25 October 2025, Accepted : 28 October 2025

Available online : 4 November 2025

บทคัดย่อ

วัตถุประสงค์และที่มา : ในการวิจัยเชิงสำรวจทางสถิติ (Survey Sampling) การประมาณค่าพารามิเตอร์ เช่น ค่าเฉลี่ย (Mean) ค่าสัดส่วน (Proportion) ค่าความแปรปรวน (Variance) หรือผลรวม (Total) ของประชากร โดยอาศัยวิธีการสุ่มตัวอย่างมักเผชิญข้อจำกัดด้านความแม่นยำของค่าประมาณ โดยเฉพาะอย่างยิ่งเมื่อขนาดตัวอย่างมีข้อจำกัดจากปัจจัยในด้านต่าง ๆ เช่น งบประมาณ เวลา หรือทรัพยากรที่ใช้ในการเก็บข้อมูล ซึ่งปัจจัยเหล่านี้ล้วนเป็นปัจจัยที่สำคัญที่ส่งผลโดยตรงที่ทำให้ค่าประมาณที่ได้จากการศึกษา มีความแปรปรวนสูงและยังไม่สามารถสะท้อนค่าที่แท้จริงของประชากรได้อย่างมีประสิทธิภาพ และเพื่อลดข้อจำกัดดังกล่าว นักวิจัยส่วนใหญ่จึงพัฒนาแนวทางการใช้ตัวแปรช่วย (Auxiliary Variable) มาใช้ในการสร้างหรือต่อยอดตัวประมาณ (Estimators) ที่สนใจศึกษา ซึ่งตัวแปรช่วย คือข้อมูลหรือคุณลักษณะของตัวแปรตัวหนึ่ง ๆ ที่นักวิจัยเชื่อมั่นว่า จะมีความสัมพันธ์กับตัวแปรที่สนใจศึกษา (Study Variable) และสามารถเข้าถึงได้ง่ายหรือมีต้นทุนการเข้าถึงตัวแปรดังกล่าวต่ำ ซึ่งการบูรณาการตัวแปรช่วยเข้ากับการประมาณค่าพารามิเตอร์ นอกจากจะช่วยลดค่าความเอนเอียง (Bias) และค่าความแปรปรวน (Variance) ของตัวประมาณแล้ว ยังส่งผลให้ประสิทธิภาพของตัวประมาณต่าง ๆ ที่สนใจศึกษามีค่าที่สูงขึ้นเมื่อเทียบกับวิธีการประมาณแบบดั้งเดิม ด้วยเหตุนี้การศึกษาทบทวนของตัวแปรช่วยในการพัฒนาตัวประมาณจึงมีความสำคัญทั้งในเชิงทฤษฎีและเชิงปฏิบัติซึ่งรวมถึงการประยุกต์ใช้ เนื่องจากผู้วิจัยสามารถนำประโยชน์จากการใช้ตัวแปรช่วยในการสร้างตัวประมาณที่สนใจศึกษาไปใช้ได้หลากหลายสาขา ไม่ว่าจะเป็นการสำรวจประชากรศาสตร์ เศรษฐศาสตร์ เกษตรศาสตร์ อุตสาหกรรม ตลอดจนงานวิจัยทางสังคมศาสตร์ซึ่งต่างก็มีข้อจำกัดด้านทรัพยากรและความสมบูรณ์ของข้อมูล นอกจากนี้การพิจารณาความสัมพันธ์ระหว่างตัวแปรช่วยกับตัวแปรที่สนใจศึกษา ยังช่วยให้เข้าใจคุณลักษณะของประชากรได้อย่างลึกซึ้งยิ่งขึ้น และสามารถนำไปปรับปรุงการวางแผนการสำรวจและการจัดสรรทรัพยากรได้อย่างมีประสิทธิภาพมากขึ้น ซึ่งสอดคล้องกับประเด็นการวิจัยที่มุ่งศึกษาการใช้สารสนเทศจากตัวแปรช่วยเพื่อพัฒนาตัวประมาณอัตราส่วนแบบเลขชี้กำลังภายใต้ปัญหาการไม่ตอบสนอง ในการสำรวจตัวอย่าง ซึ่งมีความสำคัญต่อการเพิ่มความแม่นยำและความน่าเชื่อถือของผล

การประมาณในการสำรวจเชิงสถิติ อีกทั้งการศึกษาเชิงลึกเกี่ยวกับตัวแปรช่วยช่วยให้การออกแบบแบบสำรวจยังมีความเหมาะสมและสามารถช่วยลดความคลาดเคลื่อนของข้อมูลในทุกขั้นตอนของกระบวนการสำรวจ

วิธีดำเนินการวิจัย : งานวิจัยนี้ทำการศึกษาคุณสมบัติที่สำคัญบางประการของตัวประมาณที่นำเสนอขึ้นมาใหม่ ที่ผู้วิจัยได้พัฒนามาจากตัวประมาณที่ถูกนำเสนอโดย Rao (1986) และ Singh *et al.* (2009) เช่น ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุด (Minimum Mean Squared Error: MMSE) และจะทำการเปรียบเทียบประสิทธิภาพของตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้องในเชิงทฤษฎี การประยุกต์ใช้กับข้อมูลจริง และการจำลองข้อมูล ดังรายละเอียดต่อไปนี้ 1) การเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้องในเชิงทฤษฎี ผู้วิจัยจะแสดงเงื่อนไขต่าง ๆ ที่ทำให้ตัวประมาณที่นำเสนอมีประสิทธิภาพดีกว่าตัวประมาณอื่น ๆ ที่เกี่ยวข้อง 2) การเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้องจากการประยุกต์ใช้กับข้อมูลจริง ผู้วิจัยจะใช้ข้อมูลจริงที่เก็บรวบรวมโดย Khare & Sinha (2010) และ Yadav *et al.* (2019) มาใช้ในการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ โดยจะใช้ค่าร้อยละประสิทธิภาพสัมพัทธ์ (Percentage Relative Efficiencies: PRE) เป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพของตัวประมาณต่าง ๆ และ 3) การเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้องจากการจำลองข้อมูล ผู้วิจัยจะทำการจำลองข้อมูลประชากร (X, Y) ที่มีการแจกแจงแบบปกติสองตัวแปร โดยใช้โปรแกรม R โดยที่ตัวแปรที่สนใจศึกษา Y มีค่าเฉลี่ยเท่ากับ 150 มีความแปรปรวนเท่ากับ 5 และตัวแปรช่วย X มีค่าเฉลี่ยเท่ากับ 180 มีความแปรปรวนเท่ากับ 10 (เมื่อ $Y = 114.10 + 0.2X$) โดยกำหนดให้ข้อมูลจำนวน 25% จากข้อมูลทั้งหมดถือเป็นข้อมูลที่ไม่ตอบสนอง กำหนดขนาดประชากรเท่ากับ 1,000 และกำหนดค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient: ρ_{yx}) คือ 0.2, 0.5 และ 0.8 ตามลำดับ จากนั้นจะทำการสุ่มตัวอย่างของตัวแปรช่วย X และตัวแปรที่สนใจศึกษา Y จากประชากรที่สร้างขึ้น โดยจะทำการเลือกตัวอย่างแบบสองเฟส ด้วยขนาดตัวอย่างเท่ากับ 20 และ 120 ตามลำดับ และจะเปรียบเทียบประสิทธิภาพของตัวประมาณต่าง ๆ ที่ศึกษา จากการพิจารณาจากค่าร้อยละประสิทธิภาพสัมพัทธ์ ซึ่งผลจากการจำลองข้อมูลนี้จะช่วยให้ผู้วิจัยเข้าใจถึงพฤติกรรมของตัวประมาณภายใต้สถานการณ์ต่าง ๆ อย่างละเอียดและช่วยให้ผลการวิจัยมีความน่าเชื่อถือมากขึ้น

ผลการวิจัย : จากการเปรียบเทียบประสิทธิภาพของตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้อง ทั้งในทางทฤษฎี การประยุกต์ใช้กับข้อมูลจริง และการจำลองข้อมูล จะพบว่า ตัวประมาณที่นำเสนอมีประสิทธิภาพที่ดีกว่าตัวประมาณอื่น ๆ เมื่อเงื่อนไขต่าง ๆ จาก Table 1 เป็นจริง อีกทั้งผลจากการประยุกต์ใช้กับข้อมูลจริง แสดงดัง Table 2 พบว่า ตัวประมาณที่นำเสนอมีค่า MSE ที่น้อยกว่า ค่า MSE ของตัวประมาณอื่น ๆ อีกทั้งเมื่อพิจารณาประสิทธิภาพของตัวประมาณต่าง ๆ จากค่า PRE จะพบว่า ตัวประมาณที่นำเสนอมีค่า PRE มากกว่าค่า PRE ของตัวประมาณอื่น ๆ ที่ทำการศึกษากภายใต้สถานการณ์เดียวกัน ผลจากการวิเคราะห์ในเชิงลึกชี้ให้เห็นว่าความแตกต่างของ PRE ระหว่างตัวประมาณที่นำเสนอและตัวประมาณอื่น ๆ มีแนวโน้มที่สูงขึ้น เมื่อความสัมพันธ์ระหว่างตัวแปรที่สนใจศึกษาและตัวแปรช่วยมีค่าสูงขึ้น และสำหรับผลการเปรียบเทียบประสิทธิภาพของตัวประมาณต่าง ๆ จากการจำลองข้อมูล จะพบว่า เมื่อค่าสัมประสิทธิ์สหสัมพันธ์และขนาดตัวอย่างมีค่าเท่ากัน

แต่ค่า m มีค่าที่เพิ่มขึ้น จะส่งผลให้ค่า PRE ของตัวประมาณต่าง ๆ มีค่าลดลง แต่ตัวประมาณที่น่าเสนอยังคงเป็นตัวประมาณที่มีค่า PRE ที่สูงที่สุดเช่นเดิม ซึ่งยืนยันได้ว่าตัวประมาณที่น่าเสนอยังคงมีประสิทธิภาพเหนือกว่าตัวประมาณตัวอื่น ๆ ที่เกี่ยวข้องในทุกสถานการณ์ที่ศึกษา

สรุปผลการวิจัย : จากการนำเสนอตัวประมาณอัตราส่วนแบบเลขชี้กำลังสำหรับค่าเฉลี่ยประชากร โดยใช้สารสนเทศจากตัวแปรช่วย ภายใต้การเลือกตัวอย่างสุ่มแบบสองเฟส เมื่อเกิดปัญหาการไม่ตอบสนองของข้อมูลในการสำรวจตัวอย่าง โดยพัฒนามาจากตัวประมาณที่ถูกนำเสนอโดย Rao (1986) และ Singh *et al.* (2009) พบว่า ผลจากการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่น่าเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้อง ทั้งในเชิงทฤษฎี การประยุกต์ใช้กับข้อมูลจริง และการจำลองข้อมูลให้ผลที่สอดคล้องกัน คือ ตัวประมาณอัตราส่วนแบบเลขชี้กำลังสำหรับประมาณค่าเฉลี่ยประชากรที่น่าเสนอขึ้นมาใหม่เป็นตัวประมาณที่มีประสิทธิภาพดีกว่าตัวประมาณอื่น ๆ ที่เกี่ยวข้องทุกตัว ภายใต้ชุดข้อมูลและสถานการณ์เดียวกันกับที่ศึกษาในงานวิจัยฉบับนี้ โดยสรุป ตัวประมาณที่น่าเสนอในงานวิจัยฉบับนี้เป็นเครื่องมือที่มีประโยชน์และเชื่อถือได้สำหรับนักวิจัยหรือนักสถิติในการสำรวจเชิงปริมาณ สามารถประยุกต์ใช้ได้ในหลายสาขา โดยเฉพาะในกรณีที่เกิดปัญหาการไม่ตอบสนอง เนื่องจากช่วยเพิ่มความแม่นยำในการประมาณค่าพารามิเตอร์ประชากร และได้รับการตรวจสอบประสิทธิภาพทั้งจากข้อมูลจริงและข้อมูลจำลอง

คำสำคัญ: ตัวประมาณอัตราส่วน ; ปัญหาการไม่ตอบสนอง ; การสำรวจตัวอย่าง ; ค่าเฉลี่ยประชากร

Abstract

Background and Objectives: In survey sampling, the estimation of population parameters, such as the mean (Mean), proportion (Proportion), variance (Variance), or total (Total), based on sampling methods often faces limitations in the precision of estimators. This is especially true when the sample size is limited due to various factors, such as budget, time, or the resources required for data collection. These factors directly affect the resulting estimators, leading to high variance and an inability to accurately reflect the true values of the population parameters. To mitigate these limitations, most researchers have developed approaches using auxiliary variables to construct or extend the estimators of interest. An auxiliary variable is information or a characteristic of a variable that researchers believe is correlated with the study variable and can be accessed easily or at a low cost. The integration of auxiliary variables into parameter estimation not only reduces bias and variance of the estimators but also increases the efficiency of the estimators under study compared to traditional estimation methods. Therefore, the study of the role of auxiliary variables in the development of estimators is important both theoretically and practically, including applications, as researchers can use the benefits of auxiliary variables in constructing estimators across various fields, including demography, economics, agriculture, industry, and social sciences, all of which may face limitations in resources and data completeness. Moreover, considering the relationship between auxiliary variables

and the study variable helps to understand the characteristics of the population in greater depth and can be used to improve survey planning and resource allocation efficiently. This aligns with research that aims to use information from auxiliary variables to develop exponential ratio-type estimators under the presence of non-response in sample surveys, which is important for increasing the accuracy and reliability of survey estimators. Additionally, an in-depth study of auxiliary variables helps to ensure that survey design is appropriate and can reduce measurement errors throughout all stages of the survey process.

Methodology: This research studies certain key properties of the newly proposed estimators, which were developed from the estimators introduced by Rao (1986) and Singh *et al.* (2009), such as mean squared error (MSE) and minimum mean squared error (MMSE). The efficiency of the proposed estimators will be compared with other related estimators theoretically, through applications to real data, and via simulation studies. The study comprises three components: 1) Theoretical comparison: The conditions under which the proposed estimators perform better than other related estimators will be presented. 2) Comparison using real data: Real data collected by Khare & Sinha (2010) and Yadav *et al.* (2019) will be used to compare the efficiency of the proposed estimators with other estimators, using percentage relative efficiency (PRE) as the criterion. 3) Comparison using simulation: A bivariate normal population (X, Y) will be generated using the R program, where the study variable Y has a mean of 150 and variance of 5, and the auxiliary variable X has a mean of 180 and variance of 10 (when $Y = 114.10 + 0.2X$) while 25% of the total observations are considered as non-respondents. The population size is set to 1,000, and correlation coefficients (ρ_{yx}) are set at 0.2, 0.5, and 0.8, respectively. Samples of the study variable Y and auxiliary variable X will be drawn from the simulated population using a two-phase sampling design, with sample sizes of 20 and 120, respectively. The efficiency of the studied estimators will be compared using percentage relative efficiency (PRE). This simulation helps researchers to understand the behavior of the estimators under different scenarios in detail and enhances the reliability of the study findings.

Main Results: From the comparison of the efficiency of the proposed estimators with other related estimators, in theory, applications to real data, and simulation studies, it was found that the proposed estimators perform better than other estimators when the conditions in Table 1 are satisfied. Moreover, the application to real data, as shown in Table 2, indicates that the proposed estimators have lower MSE than other estimators. When considering the efficiency of the estimators from PRE values, the proposed estimators have higher PRE than other studied estimators under the same conditions. In-depth analysis shows that the difference in PRE between the proposed estimators and other estimators tends to increase when the correlation between the study variable and the auxiliary variable is higher. For the comparison of the efficiency of estimators from the simulation study, it was found that when the

correlation coefficient and sample size are the same, an increase in the relevant parameter leads to a decrease in PRE for all estimators. However, the proposed estimators remain the estimators with the highest PRE, confirming that the proposed estimators continue to perform better than other related estimators in all scenarios studied.

Conclusions: The study proposed an exponential ratio-type estimator for the population mean using information from auxiliary variables under two-phase sampling with non-response, developed from the estimators introduced by Rao (1986) and Singh *et al.* (2009). Results from the efficiency comparison between the proposed estimators and other related estimators, using theoretical analysis, real data applications, and simulation studies, consistently show that the newly proposed exponential ratio-type estimator for the population mean outperforms all other related estimators under the same datasets and scenarios. In conclusion, the estimators proposed in this study provide a useful and reliable tool for researchers and statisticians in quantitative surveys. They can be applied across various fields, particularly in situations involving non-response, as they enhance the accuracy of population parameter estimation and have been validated using both real and simulated data.

Keywords : ratio estimator, problem of non-response, sample surveys, population mean

*Corresponding author. E-mail : napattchan.dan@kmutt.ac.th

Introduction

โดยทั่วไปสาเหตุหลักของการสำรวจในด้านต่าง ๆ เช่น การเกษตร การศึกษา อุตุนิยมวิทยา ชีวการแพทย์ วิศวกรรมศาสตร์ เป็นต้น มักเกิดจากการที่ผู้วิจัยไม่สามารถเก็บรวบรวมข้อมูลที่สนใจได้ครบถ้วนจากแหล่งข้อมูล โดยเฉพาะอย่างยิ่งแหล่งข้อมูลที่เป็นคน หรือบุคคล ทั้งนี้อาจเกิดจากการขาดความร่วมมือในการตอบแบบสอบถามหรือแบบสำรวจจากแหล่งข้อมูล ทั้งในแง่ของการไม่ตอบ (Non-response) หรือผู้ตอบไม่ให้ข้อมูลที่ถูกต้อง (Response Errors) โดยปัญหาดังที่กล่าวมาข้างต้นในทางทฤษฎีจะเรียกว่า เกิดปัญหาการไม่ตอบสนอง (Nonresponse Problem) ต่อข้อมูล ซึ่งการไม่ตอบสนองถือเป็นความท้าทายที่สำคัญในการสำรวจตัวอย่างเกือบทั้งหมด และอัตราการไม่ตอบสนองจะมีแนวโน้มที่จะเพิ่มขึ้น เมื่อกลุ่มตัวอย่างหรือแหล่งข้อมูลมีขนาดใหญ่ สำหรับงานทางด้านสถิติ โดยเฉพาะอย่างยิ่งหากผู้วิจัยสนใจที่จะสร้างตัวประมาณ (Estimator) เพื่อประมาณค่าเฉลี่ย (μ) ค่าสัดส่วน (P) หรือค่าความแปรปรวน (σ^2) ของประชากรที่สนใจศึกษา ผู้วิจัยจะใช้ค่าเฉลี่ยของตัวอย่าง ($\bar{x}$) ค่าสัดส่วนของตัวอย่าง ($\hat{p}$) และค่าความแปรปรวนของตัวอย่าง (s^2) เป็นต้น แทนการใช้ค่าที่แท้จริงที่คำนวณได้จากประชากร โดยค่าที่คำนวณได้จากตัวอย่างดังที่กล่าวมา หากเป็นค่าที่ได้จากข้อมูลที่เกิดปัญหาการไม่ตอบสนอง จะทำให้ความแม่นยำของตัวประมาณค่าเหล่านี้นับน้อยลง และเพื่อจัดการกับปัญหาการไม่ตอบสนองของข้อมูลสำหรับประมาณค่าเฉลี่ยประชากร ($\bar{Y}$) นักวิจัยส่วนใหญ่จึงนิยมใช้เทคนิคการเลือกตัวอย่างย่อย (Sub-sampling Technique) ภายใต้การเลือกตัวอย่างสุ่มแบบสองเฟส (Two-phase Sampling) ร่วมกับการประยุกต์ใช้ตัวประมาณสำหรับประมาณค่าเฉลี่ยประชากร (Population mean: $\bar{Y}$) ของตัวแปรที่สนใจศึกษา (Study Variable: Y) ที่นำเสนอโดย Hansen และ

Hurwitz (1946) มาใช้ในการเพิ่มประสิทธิภาพของตัวประมาณ (Estimator) ที่สร้างขึ้น โดยตัวประมาณดังกล่าวพบว่า มีคุณสมบัติเป็นตัวประมาณที่ไม่มีความเอนเอียง (Unbiased)

โดย Hansen & Hurwitz (1946) จะทำการเลือกตัวอย่างขนาด n โดยใช้การเลือกตัวอย่างแบบง่ายแบบไม่ใส่คืน (Simple Random Sampling Without Replacement: SRSWOR) จากประชากรขนาด N ทั้งนี้ก่อนการเลือกตัวอย่างจะแบ่งกลุ่มประชากรออกเป็น 2 กลุ่มคือ กลุ่มที่หนึ่งมีขนาด N_1 จะถูกกำหนดให้เป็นกลุ่มที่ข้อมูลมีการตอบสนอง และกลุ่มที่สองมีขนาด $N_2 = N - N_1$ คือกลุ่มที่ข้อมูลไม่มีการตอบสนอง ตามลำดับ ภายหลังจากการเลือกตัวอย่างขนาด n แล้วจะทำการแยกตัวอย่างขนาด n ออกเป็นสองกลุ่มได้แก่ ตัวอย่างกลุ่มที่หนึ่งมีขนาด n_1 จะถูกกำหนดให้เป็นกลุ่มที่พบว่าข้อมูลมีการตอบสนอง ในขณะที่ตัวอย่างกลุ่มที่สองซึ่งมีขนาด $n_2 = n - n_1$ จะเป็นกลุ่มที่ประกอบด้วยข้อมูลที่ไม่มีการตอบสนอง โดยตัวอย่างกลุ่มที่สองซึ่งมีขนาด n_2 นี้จะถูกเลือกตัวอย่างแบบง่ายแบบไม่ใส่คืนอีกครั้ง โดยกำหนดขนาดตัวอย่างที่ถูกเลือกในขั้นตอนนี้มีขนาด s , $s = n_2 / m$ (เมื่อ $m > 1$) ดังนั้นขนาดตัวอย่างของข้อมูลที่มีการตอบสนองจะมีค่าเท่ากับ $n_1 + s$ และตัวประมาณค่าเฉลี่ยของประชากร $\bar{Y}$ ที่นำเสนอโดย Hansen & Hurwitz (1946) มีรูปแบบดังนี้

$$\bar{y}^* = w_1 \bar{y}_1 + w_2 \bar{y}_{2(s)} \quad (1)$$

โดยที่ความแปรปรวนของ $\bar{y}^*$ ได้แก่

$$V(\bar{y}^*) = \bar{Y}^2 \left[\delta C_{y_1}^2 + \delta^* C_{y(s)}^2 \right] \quad (2)$$

$$\text{เมื่อ } w_1 = n_1 / n, w_2 = n_2 / n, \bar{y}_1 = \sum_{i=1}^{n_1} y_i / n_1, \bar{y}_{2(s)} = \sum_{i=1}^s y_i / s, \delta = (N - n) / Nn, \delta^* = w_2 (m - 1) / n, C_{y_1}^2 = S_{y_1}^2 / \bar{Y}^2, \\ C_{y(s)}^2 = S_{y(s)}^2 / \bar{Y}^2, S_{y_1}^2 = (N - 1)^{-1} \sum_{i=1}^{N_1} (y_i - \bar{Y})^2 \text{ และ } S_{y(s)}^2 = (N_2 - 1)^{-1} \sum_{i=1}^{N_2} (y_i - \bar{Y}_2)^2$$

นอกจากนี้นักวิจัยอีกหลาย ๆ คน ยังนิยมเก็บข้อมูลอื่น ๆ นอกเหนือจากการเก็บเฉพาะข้อมูลที่สนใจศึกษา เพื่อนำข้อมูลเหล่านี้มาสร้างเป็นตัวแปรช่วย (Auxiliary Variable: X) เพื่อใช้ในการช่วยเพิ่มประสิทธิภาพของตัวแปรที่สนใจศึกษา Y สำหรับการประมาณค่าเฉลี่ยประชากร ($\bar{Y}$) ดังเห็นได้จากงานวิจัยของ Rao (1986) Singh *et al.* (2009) Yadav & Adewara (2013) Yadav *et al.* (2014, 2015) Jerajuddin & Kishun (2016) Abid *et al.* (2016) Ganie *et al.* (2022) Singh *et al.* (2023) และ Ahmadini *et al.* (2025) เป็นต้น แม้ว่าตัวประมาณต่าง ๆ ที่ถูกนำเสนอมาก่อนหน้านี้จะใช้สารสนเทศจากตัวแปรช่วย มาช่วยเพิ่มประสิทธิภาพในการประมาณค่าเฉลี่ยประชากรภายใต้กรณีไม่ตอบสนอง แต่ประสิทธิภาพดังกล่าว ยังคงถูกจำกัดด้วยปัจจัยหลายประการ อาทิ การเลือกใช้แผนการสุ่มตัวอย่างและการเลือกตัวแปรช่วยที่เหมาะสม

จากตัวประมาณต่าง ๆ ดังที่กล่าวมาแล้วข้างต้น ในงานวิจัยฉบับนี้ผู้วิจัยจึงสนใจที่จะนำเสนอตัวประมาณอัตราส่วนแบบเลขชี้กำลังสำหรับค่าเฉลี่ยประชากร โดยใช้สารสนเทศจากตัวแปรช่วย ภายใต้อการเลือกตัวอย่างแบบสองเฟส เมื่อเกิดปัญหาการไม่ตอบสนองของข้อมูลในการสำรวจตัวอย่าง โดยเฉพาะในตัวแปรที่สนใจศึกษา Y โดยตัวประมาณที่นำเสนอจะ

พัฒนามาจากตัวประมาณที่นำเสนอโดย Rao (1986) และ Singh *et al.* (2009) นอกจากนี้คุณสมบัติที่สำคัญบางประการของตัวประมาณที่นำเสนอขึ้นมาใหม่ เช่น ความคลาดเคลื่อนกำลังสองเฉลี่ย และความคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุดจะถูกนำเสนอพร้อมทั้งจะทำการเปรียบเทียบประสิทธิภาพของตัวประมาณดังกล่าวกับตัวประมาณตัวอื่น ๆ ที่เกี่ยวข้องในเชิงทฤษฎีการประยุกต์ใช้กับข้อมูลจริง และการจำลองข้อมูล

Methodology

จากตัวประมาณที่นำเสนอโดย Rao (1986) และ Singh *et al.* (2009) ดังสมการที่ (3) และสมการที่ (4) ดังต่อไปนี้

$$t_1 = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}} \right) \quad (3)$$

$$t_2 = \bar{y}^* \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \quad (4)$$

ผู้วิจัยมีความสนใจที่จะพัฒนาตัวประมาณนำเสนอโดย Rao (1986) และ Singh *et al.* (2009) โดยการนำเสนอตัวประมาณอัตราส่วนแบบเลขชี้กำลังสำหรับค่าเฉลี่ยประชากร ภายใต้การเลือกตัวอย่างสุ่มแบบสองเฟส เมื่อเกิดปัญหาการไม่ตอบสนองของข้อมูลในการสำรวจตัวอย่าง โดยเฉพาะในตัวแปรที่สนใจศึกษา Y ดังรายละเอียดต่อไปนี้

ตัวประมาณ t_{NEW_1} เกิดจากการนำตัวประมาณ t_1 จากสมการที่ (3) แทนลงในค่า $\bar{y}^*$ ของตัวประมาณ t_2 ในสมการที่ (4) จะได้รูปแบบของตัวประมาณ t_{NEW_1} ดังสมการที่ (5)

$$t_{NEW_1} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}} \right) \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) \quad (5)$$

ในขณะที่ตัวประมาณ t_{NEW_2} เกิดจากการนำตัวประมาณ t_2 จากสมการที่ (4) แทนลงในค่า $\bar{y}^*$ ของตัวประมาณ t_2 ในสมการที่ (4) จะได้รูปแบบของตัวประมาณ t_{NEW_2} ดังสมการที่ (6)

$$t_{NEW_2} = \bar{y}^* \exp \left(\frac{2(\bar{X} - \bar{x})}{\bar{X} + \bar{x}} \right) \quad (6)$$

และตัวประมาณ t_{NEW_3} เกิดจากการนำตัวประมาณ t_{NEW_1} และตัวประมาณ t_{NEW_2} มารวมเข้าด้วยกัน โดยมีรูปแบบของตัวประมาณดังนี้

$$t_{NEW_3} = \alpha \bar{y}^* \left(\frac{\bar{X}}{\bar{x}} \right) \exp \left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}} \right) + (1 - \alpha) \bar{y}^* \exp \left(\frac{2(\bar{X} - \bar{x})}{\bar{X} + \bar{x}} \right) \quad (7)$$

เมื่อ α คือ ค่าคงที่ใด ๆ ที่ทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ t_{NEW_3} มีค่าที่ต่ำที่สุด

จากนั้นจะทำการหาค่าเคลื่อนกำลังสองเฉลี่ยของตัวประมาณที่นำเสนอโดยจะทำการจัดรูปแบบของตัวประมาณจากสมการที่ (5) ถึงสมการที่ (7) ให้อยู่ในเทอมความคลาดเคลื่อน (Error terms) โดยกำหนดให้ $\bar{y}^* = \bar{Y}(1 + e_0^*)$, $\bar{x} = \bar{X}(1 + e_1)$ เมื่อ e_0^* และ e_1 คือ ความคลาดเคลื่อนระหว่างค่าเฉลี่ยตัวอย่างกับค่าเฉลี่ยประชากรของตัวแปรที่สนใจศึกษา Y และตัวแปรช่วย X ตามลำดับ ดังนั้นจะได้

$$t_{NEW_1} = \bar{Y}(1 - e_1 + e_1^2 + e_0^* - e_0^*e_1) \exp \left(\frac{-e_1}{2} \left(1 + \frac{e_1}{2} \right)^{-1} \right) \quad (8)$$

$$t_{NEW_2} = \bar{Y}(1 + e_1^*) \exp \left(e_1 \left(1 + \frac{e_1}{2} \right)^{-1} \right) \quad (9)$$

$$t_{NEW_3} = \alpha \bar{Y} \left[-\frac{\alpha e_1}{2} + \frac{3\alpha e_1^2}{8} + \frac{\alpha e_1^2}{2} - \frac{\alpha e_0^* e_1}{2} + 1 - e_1 + e_1^2 + e_0^* - e_0^*e_1 \right] \quad (10)$$

จากสมการที่ (8) ถึงสมการที่ (10) ผู้วิจัยจะใช้ทฤษฎีบททวินาม (Binomial Theorem) และอนุกรมเทย์เลอร์ (Taylor's Series) ในการกระจายพจน์ที่มีเลขชี้กำลังติดลบ ดังนั้นจะได้

$$t_{NEW_1} = \bar{Y}(1 - e_1 + e_1^2 + e_0^* - e_0^*e_1) \exp \left(-\frac{e_1}{2} + \frac{e_1^2}{4} \right) \quad (11)$$

$$t_{NEW_2} = \bar{Y}(1 + e_1^*) \exp \left(-e_1 + \frac{e_1^2}{2} \right) \quad (12)$$

$$t_{NEW_3} = \alpha \bar{Y}(1 - e_1 + e_1^2 + e_0^* - e_0^*e_1) \exp \left(-\frac{e_1}{2} + \frac{e_1^2}{4} \right) + (1 - \alpha) \bar{Y}(1 + e_0^*) \exp \left(-e_1 + \frac{e_1^2}{2} \right) \quad (13)$$

และผู้วิจัยจะกระจายของพจน์ $\exp(x)$ โดยพจน์ที่ยกกำลัง 3 เป็นต้นไป จะถูกกำหนดให้เป็นค่าคลาดเคลื่อนจากการตัดปลาย (Truncation Error) ดังนั้นจะได้

$$t_{NEW_1} = \bar{Y} \left(1 - \frac{e_1}{2} + \frac{3e_1^2}{8} - e_1 + \frac{e_1^2}{2} + e_1^2 + e_0^* - \frac{e_0^* e_1}{2} - e_0^* e_1 \right) \quad (14)$$

$$t_{NEW_2} = \bar{Y} (1 - e_1 + e_1^2 + e_0^* - e_0^* e_1) \quad (15)$$

$$t_{NEW_3} = \alpha \bar{Y} \left[-\frac{\alpha e_1}{2} + \frac{3\alpha e_1^2}{8} + \frac{\alpha e_1^2}{2} - \frac{\alpha e_0^* e_1}{2} + 1 - e_1 + e_1^2 + e_0^* - e_0^* e_1 \right] \quad (16)$$

สำหรับการคำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ t_{NEW_1} , t_{NEW_2} และ t_{NEW_3} สามารถคำนวณได้จาก $MSE(t_{NEW_i}) = E[t_{NEW_i} - \bar{Y}]^2$; $i=1,2,3$ นั่นคือ

$$\begin{aligned} MSE(t_{NEW_1}) &= E[t_{NEW_1} - \bar{Y}]^2 \\ &= \bar{Y}^2 E \left[\frac{e_1^2}{4} + e_1^2 + e_0^{*2} + e_1^2 - e_0^* e_1 - 2e_0^* e_1 \right] \end{aligned} \quad (17)$$

แทนค่า $E(e_0^*) = E(e_1) = 0$, $E(e_0^{*2}) = \delta C_y^2 + \delta^* C_{y(s)}^2$, $E(e_1^2) = \delta C_x^2$, $E(e_0^* e_1) = \delta \rho_{yx} C_y C_x$ ลงในสมการที่ (17) และจัดรูปสมการใหม่อีกครั้ง จะได้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ t_{NEW_1} ดังนี้

$$MSE(t_{NEW_1}) = \bar{Y}^2 \left[(\delta C_y^2 + \delta^* C_{y(s)}^2) + \frac{1}{4} \delta C_x^2 (9 - 12K) \right] \quad (18)$$

เมื่อ $K = \rho_{yx} C_y / C_x$

ในทำนองเดียวกันจะได้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ t_{NEW_2} และ t_{NEW_3} ดังนี้

$$MSE(t_{NEW_2}) = \bar{Y}^2 \left[(\delta C_y^2 + \delta^* C_{y(2)}^2) + \delta C_x^2 (1 - 2K) \right] \quad (19)$$

$$MSE(t_{NEW_3}) = \bar{Y}^2 \left[\left(\delta C_y^2 + \delta^* C_{y(2)}^2 \right) + \delta C_x^2 \left[\left(\frac{\alpha^2}{4} + 1 + \alpha \right) - K(\alpha + 2) \right] \right] \quad (20)$$

ทั้งนี้เนื่องจากตัวประมาณ t_{NEW_3} เป็นตัวประมาณเดียวที่มีค่าคงที่ใด ๆ (α) ดังนั้นจึงต้องทำการหาค่า α ที่ทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ t_{NEW_3} มีค่าที่ต่ำที่สุด โดยการนำอนุพันธ์อันดับที่หนึ่งของ $MSE(t_{NEW_3})$ มาเทียบกับค่า α และให้เท่ากับ 0 จะได้

$$\frac{\partial MSE(t_{NEW_3})}{\partial \alpha} = \frac{2}{4} \alpha \delta C_x^2 + \delta C_x^2 - K \delta C_x^2 = 0 \quad (21)$$

แก้สมการหาค่า α จากสมการที่ (21) จะได้

$$\alpha_{(opt)} = \frac{2(K \delta C_x^2 - \delta C_x^2)}{\delta C_x^2} = 2(K - 1) \quad (22)$$

และจะหาอนุพันธ์อันดับที่สองของ $MSE(t_{NEW_3})$ เทียบกับค่า α จะได้

$$\frac{\partial^2 MSE(t_{NEW_3})}{\partial \alpha^2} = 2(K - 1) > 0 \quad (23)$$

ซึ่งพบว่ามีค่ามากกว่า 0 ทำให้ได้ $MSE(t_{NEW_3})$ มีค่าต่ำที่สุด และแทนค่า $\alpha_{(opt)}$ ลงในสมการที่ (20) จะได้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่ต่ำที่สุดของตัวประมาณ t_{NEW_3} ดังนี้

$$MMSE(t_{NEW_3}) = \bar{Y}^2 \left[\left(\delta C_y^2 + \delta^* C_{y(2)}^2 \right) - \delta K C_x^2 \right] \quad (24)$$

Results

1. ในการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้องในเชิงทฤษฎี จะพิจารณาจากเงื่อนไขจากสมการใน Table ต่อไปนี้

Table 1 Presents the conditions under which the proposed ratio estimator outperforms other related estimators in terms of theoretical efficiency

Related estimators	Conditions for Higher Efficiency of the Proposed Ratio Estimator		
	t_{NEW_1}	t_{NEW_2}	t_{NEW_3}
$\bar{y}^*$	$\rho_{yx} > \frac{3C_x}{4C_y}$	$\rho_{yx} > \frac{C_x}{2C_y}$	$\rho_{yx} > \frac{C_x \left(\frac{\alpha^2}{4} + 1 + \alpha \right)}{(\alpha + 2)C_y}$
t_1	$\rho_{yx} < \frac{-5C_x C_y + 12C_x^2}{C_x C_y + 8C_y^2}$	$\rho_{yx} > 0$	$\rho_{yx} > \frac{C_x \left(\frac{\alpha^2}{4} + \alpha \right)}{\alpha C_y}$
t_2	$\rho_{yx} < -\frac{5C_x C_y + 12C_x^2}{C_x C_y + 4C_y^2}$	$\rho_{yx} > \frac{3C_x}{4C_y}$	$\rho_{yx} > \frac{C_x \left(\alpha + \frac{\alpha^2}{4} + \frac{3}{4} \right)}{(\alpha + 1)C_y}$

ทั้งนี้สมการต่าง ๆ ภายใน Table 1 จะพิจารณาภายใต้เงื่อนไข $MSE(t_{NEW_i}, MMSE(t_{NEW_3})) < MSE(\bar{y}^*, t_1, t_2)$; $i = 1, 2$

2. ในการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณอัตราส่วนตัวใหม่กับตัวประมาณอื่น ๆ ที่เกี่ยวข้องจากการประยุกต์ใช้กับข้อมูลจริง ผู้วิจัยจะใช้ข้อมูลจริงจำนวน 2 ชุด ซึ่งเก็บรวบรวมโดย Khare & Sinha (2010) และ Yadav *et al.* (2019) โดยลักษณะประชากรของข้อมูลทั้ง 2 ชุด แสดงดังนี้

ข้อมูลชุดที่ 1 : $N = 95$, $n = 70$, $w_2 = 0.25$, $\bar{Y} = 19.497$, $\bar{X} = 57.217$, $C_y^2 = 0.036$, $C_x^2 = 0.005$,
 $C_{y(s)}^2 = 0.034$, $\rho_{yx} = 0.834$

ข้อมูลชุดที่ 2 : $N = 150$, $n = 40$, $w_2 = 0.25$, $\bar{Y} = 32.97$, $\bar{X} = 3.94$, $C_y^2 = 0.562$, $C_x^2 = 0.532$,
 $C_{y(s)}^2 = 0.518$, $\rho_{yx} = 0.860$

และจะใช้ค่าร้อยละประสิทธิภาพสัมพัทธ์ (PRE) เป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพของตัวประมาณ ซึ่งมีสูตรการคำนวณแสดงดังสมการที่ (25) ดังนี้

$$PRE = \frac{MSE(\bar{y}^*, t_1, t_2)}{MSE(t_{NEW_i})} \times 100; \quad i = 1, 2, 3 \quad (25)$$

ทั้งนี้ตัวประมาณ t_{NEW_i} จะเป็นตัวประมาณที่มีประสิทธิภาพสูงกว่าตัวประมาณ $\bar{y}^*, t_1, t_2$ ก็ต่อเมื่อ $MSE(t_{NEW_i}) < MSE(\bar{y}^*, t_1, t_2)$ หรืออาจกล่าวได้ว่าตัวประมาณที่มีค่า PRE ที่สูงจะถือว่าเป็นตัวประมาณที่มีประสิทธิภาพสูง โดยผลการคำนวณค่า MSE และค่า PRE ของตัวประมาณต่าง ๆ แสดงดัง Table 2

Table 2 PRE Values of the Estimators Based on Data Sets 1 and 2

m	Data 1: Khare & Sinha (2010)			Data 2: Yadav <i>et al.</i> (2019)		
	Estimators	MSE	PRE	Estimators	MSE	PRE
1	$\bar{y}^*$	0.097606	100.00	$\bar{y}^*$	14.731851	100.00
	t_1	0.072770	134.13	t_1	6.584991	223.72
	t_2	0.083402	117.03	t_2	8.003418	184.07
	t_{NEW_1}	0.065712	148.54	t_{NEW_1}	10.476572	140.62
	t_{NEW_2}	0.072770	134.13	t_{NEW_2}	6.584991	223.72
	t_{NEW_3}	0.061822	157.88	t_{NEW_3}	6.441006	228.72
2	$\bar{y}^*$	0.143765	100.00	$\bar{y}^*$	18.253798	100.00
	t_1	0.118929	120.88	t_1	10.106939	180.61
	t_2	0.129561	110.96	t_2	11.525366	158.38
	t_{NEW_1}	0.111871	128.51	t_{NEW_1}	13.998519	130.40
	t_{NEW_2}	0.118929	120.88	t_{NEW_2}	10.106939	180.61
	t_{NEW_3}	0.107981	133.14	t_{NEW_3}	9.962954	183.22
3	$\bar{y}^*$	0.189924	100.00	$\bar{y}^*$	21.775746	100.00
	t_1	0.165088	115.04	t_1	13.628887	159.78
	t_2	0.175720	108.08	t_2	15.047313	144.72
	t_{NEW_1}	0.158030	120.18	t_{NEW_1}	17.520467	124.29
	t_{NEW_2}	0.165088	115.04	t_{NEW_2}	13.628887	159.78
	t_{NEW_3}	0.154140	123.22	t_{NEW_3}	13.484902	161.48
4	$\bar{y}^*$	0.236083	100.00	$\bar{y}^*$	25.297694	100.00
	t_1	0.211247	111.76	t_1	17.150835	147.50
	t_2	0.221879	106.40	t_2	18.569261	136.23
	t_{NEW_1}	0.204189	115.62	t_{NEW_1}	21.042415	120.22
	t_{NEW_2}	0.211247	111.76	t_{NEW_2}	17.150835	147.50
	t_{NEW_3}	0.200299	117.87	t_{NEW_3}	17.006850	148.75

m	Data 1: Khare & Sinha (2010)			Data 2: Yadav <i>et al.</i> (2019)		
	Estimators	MSE	PRE	Estimators	MSE	PRE
5	$\bar{y}^*$	0.282242	100.00	$\bar{y}^*$	28.819642	100.00
	t_1	0.257406	109.65	t_1	20.672782	139.41
	t_2	0.268038	105.30	t_2	22.091209	130.46
	t_{NEW_1}	0.250348	112.74	t_{NEW_1}	24.564362	117.32
	t_{NEW_2}	0.257406	109.65	t_{NEW_2}	20.672782	139.41
	t_{NEW_3}	0.246458	114.52	t_{NEW_3}	20.528797	140.39

3. การเปรียบเทียบประสิทธิภาพของตัวประมาณโดยการจำลองข้อมูล

ผู้วิจัยจะทำการจำลองข้อมูลประชากร (X, Y) ที่มีการแจกแจงแบบปกติสองตัวแปร โดยใช้โปรแกรม R โดยที่ตัวแปรที่สนใจศึกษา Y มีค่าเฉลี่ยเท่ากับ 150 มีความแปรปรวนเท่ากับ 5 และตัวแปรช่วย X มีค่าเฉลี่ยเท่ากับ 180 มีความแปรปรวนเท่ากับ 10 (เมื่อ $Y = 114.10 + 0.2X$) โดยกำหนดให้ข้อมูลจำนวน 25% จากข้อมูลทั้งหมดถือเป็นข้อมูลที่ไม่ตอบสนอง กำหนดขนาดประชากรเท่ากับ 1,000 และกำหนดค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient: ρ_{yx}) คือ 0.2, 0.5 และ 0.8 ตามลำดับ จากนั้นจะทำการสุ่มตัวอย่างของตัวแปรช่วย X และตัวแปรที่สนใจศึกษา Y จากประชากรที่สร้างขึ้น โดยจะทำการเลือกตัวอย่างแบบสองเฟส ด้วยขนาดตัวอย่างเท่ากับ 20 และ 120 ตามลำดับ โดยผลการเปรียบเทียบประสิทธิภาพของตัวประมาณ แสดงดัง Table 3 ดังนี้

Table 3 Results of Efficiency Comparison Between the Proposed Estimator and Related Estimators Through Data Simulation Based on PRE Values

m	Estimators	$\rho_{yx} = 0.2$		$\rho_{yx} = 0.5$		$\rho_{yx} = 0.8$	
		n= 20	n= 120	n= 20	n= 120	n= 20	n= 120
1	$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000	100.000
	t_1	193.587	136.932	124.322	131.196	115.573	118.037
	t_2	103.477	111.703	116.060	109.393	110.064	102.771
	t_{NEW_1}	150.289	145.840	141.058	137.887	134.238	126.781
	t_{NEW_2}	193.587	136.932	124.322	131.196	115.573	118.037
	t_{NEW_3}	227.154	162.563	223.969	157.380	178.522	129.360
2	$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000	100.000
	t_1	135.616	146.906	133.796	135.990	124.429	116.550
	t_2	102.709	115.814	101.741	112.314	101.674	101.465

m	Estimators	$\rho_{yx} = 0.2$		$\rho_{yx} = 0.5$		$\rho_{yx} = 0.8$	
		n= 20	n= 120	n= 20	n= 120	n= 20	n= 120
3	t_{NEW_1}	161.141	169.390	140.218	141.255	139.242	128.638
	t_{NEW_2}	135.616	146.906	133.796	135.990	124.429	116.550
	t_{NEW_3}	178.360	183.770	158.553	155.360	154.501	127.454
	$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000	100.000
	t_1	135.483	140.426	126.715	132.258	111.053	121.944
	t_2	108.523	105.375	107.892	116.009	104.691	121.095
	t_{NEW_1}	154.418	155.051	144.134	135.478	120.798	130.579
	t_{NEW_2}	135.483	140.426	126.715	132.258	111.053	121.944
	t_{NEW_3}	166.262	163.772	157.038	144.126	151.213	134.475
4	$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000	100.000
	t_1	113.875	147.855	119.009	130.482	113.660	115.824
	t_2	107.244	128.797	103.496	118.130	103.048	115.241
	t_{NEW_1}	143.313	159.525	122.155	132.839	121.686	121.624
	t_{NEW_2}	128.935	147.855	119.009	130.482	113.660	115.824
	t_{NEW_3}	151.919	166.234	130.466	138.934	123.486	123.213
	$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000	100.000
5	t_1	112.080	121.114	111.680	114.149	103.560	112.373
	t_2	101.117	108.056	105.788	105.413	103.060	111.930
	t_{NEW_1}	127.879	128.810	121.361	115.787	117.002	116.725
	t_{NEW_2}	117.080	121.114	111.680	114.149	103.560	112.373
	t_{NEW_3}	134.229	133.136	128.064	120.032	125.836	117.902
	$\bar{y}^*$	100.000	100.000	100.000	100.000	100.000	100.000

Discussion

จากการเปรียบเทียบประสิทธิภาพของตัวประมาณที่นำเสนอกับตัวประมาณอื่นๆ ที่เกี่ยวข้อง ในทางทฤษฎี จะพบว่า ตัวประมาณที่นำเสนอจะมีประสิทธิภาพดีกว่าตัวประมาณอื่นๆ ถ้าเงื่อนไขต่าง ๆ จากสมการใน Table 1 เป็นจริง

ในขณะที่ผลการเปรียบเทียบประสิทธิภาพของตัวประมาณที่นำเสนอกับตัวประมาณอื่นๆ ที่เกี่ยวข้อง จากการประยุกต์ใช้กับข้อมูลจริง แสดงดัง Table 2 จะพบว่า เมื่อพิจารณาจากชุดข้อมูลและเงื่อนไขเดียวกัน ค่า MSE ของตัวประมาณที่นำเสนอขึ้นมาใหม่ t_{NEW_3} มีค่าน้อยกว่าค่า MSE ของตัวประมาณอื่นๆ ในขณะที่เมื่อพิจารณาจากค่า PRE ของตัวประมาณแต่ละตัว จะพบว่า ตัวประมาณ t_{NEW_3} มีค่า PRE มากกว่าค่า PRE ของตัวประมาณอื่นๆ ที่ทำการศึกษา

และสำหรับผลการเปรียบเทียบประสิทธิภาพของตัวประมาณต่าง ๆ จากการจำลองข้อมูล จะพบว่า เมื่อค่าสัมประสิทธิ์สหสัมพันธ์และขนาดตัวอย่างมีค่าเท่ากัน แต่ค่า m มีค่าที่เพิ่มขึ้น จะส่งผลให้ค่า PRE ของตัวประมาณต่าง ๆ มีค่าลดลง โดยตัวประมาณที่มีค่า PRE ที่สูงที่สุด ได้แก่ ตัวประมาณ t_{NEW_3} และเมื่อพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเพิ่มขึ้น คือ จาก $\rho_{yx} = 0.2$ ไปเป็น $\rho_{yx} = 0.5$ และ $\rho_{yx} = 0.8$ ในขณะที่ขนาดตัวอย่างและค่า m มีค่าเท่ากัน จะพบว่า ค่า PRE ของตัวประมาณ ต่าง ๆ มีค่าที่ลดลง แต่ตัวประมาณ t_{NEW_3} ยังคงเป็นตัวประมาณที่มีค่า PRE ที่สูงที่สุดเช่นเดิม โดยค่า ρ_{yx} ที่เพิ่มจะส่งผลให้ค่า PRE ลดลง ในขณะที่เมื่อ m มีค่าเพิ่มขึ้น ณ ขนาดตัวอย่างเดียวกัน จะพบว่าค่า PRE ของตัวประมาณจะมีค่าลดลง

เนื่องจากค่า m หรือ inverse sampling rate ซึ่งมีความสัมพันธ์กลับกับอัตราการสุ่มตัวอย่างที่แท้จริง (Sampling Fraction) เมื่อค่า m มีค่าเพิ่มขึ้น จะส่งผลให้อัตราการสุ่มตัวอย่างลดลง ทำให้จำนวนตัวอย่างที่ได้มีค่าน้อยลง และจะส่งผลให้ความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของตัวประมาณเพิ่มขึ้น เนื่องจากตัวแปรช่วยมีข้อมูลน้อยลงและความสามารถในการอธิบายความแปรปรวนของตัวแปรที่สนใจลดลง ทำให้ประสิทธิภาพสัมพัทธ์ (PRE) ของตัวประมาณลดลงตามไปด้วย อาจกล่าวอีกนัยหนึ่งว่าการเพิ่มค่า m จะทำให้ความแม่นยำของตัวประมาณลดลง เนื่องจากความแปรปรวนของตัวอย่างสูงขึ้น และตัวแปรช่วยไม่สามารถลดความคลาดเคลื่อนได้อย่างมีประสิทธิภาพ และเมื่อแทนค่าข้อมูลจากการประยุกต์ใช้กับข้อมูลจริงและค่าข้อมูลที่ได้จากการจำลองลงในเงื่อนไขต่าง ๆ ตามสมการในตารางที่ 1 จะพบว่าทุกเงื่อนไขเป็นจริง จึงกล่าวได้ว่าผลการเปรียบเทียบประสิทธิภาพในทางทฤษฎี การประยุกต์ใช้กับข้อมูลจริง และการจำลองข้อมูลให้ผลที่สอดคล้องกัน ดังนั้นจึงสามารถสรุปได้ว่า ตัวประมาณที่นำเสนอ t_{NEW_3} เป็นตัวประมาณที่มีประสิทธิภาพดีกว่าตัวประมาณอื่น ๆ จากการประยุกต์ใช้กับข้อมูลจริงและการจำลองข้อมูล

Conclusions

จากการนำเสนอตัวประมาณอัตราส่วนแบบเลขชี้กำลังสำหรับค่าเฉลี่ยประชากร โดยใช้สารสนเทศจากตัวแปรช่วยภายใต้การเลือกตัวอย่างสุ่มแบบสองเฟส เมื่อเกิดปัญหาการไม่ตอบสนองของข้อมูลในการสำรวจตัวอย่าง โดยพัฒนามาจากตัวประมาณที่แนะนำเสนอโดย Rao (1986) และ Singh *et al.* (2009) พบว่า ผลจากการเปรียบเทียบประสิทธิภาพระหว่างตัวประมาณที่นำเสนอกับตัวประมาณอื่น ๆ ที่เกี่ยวข้อง ทั้งในเชิงทฤษฎี การประยุกต์ใช้กับข้อมูลจริง และการจำลองข้อมูลให้ผลที่สอดคล้องกัน กล่าวคือ ภายใต้สถานการณ์ที่ความสัมพันธ์ระหว่างตัวแปรช่วย X และตัวแปรที่สนใจศึกษา Y มีค่าเป็นบวก ตัวประมาณที่นำเสนอจะยังคงมีประสิทธิภาพดีกว่าตัวประมาณอื่น ๆ ดังนั้นจึงสามารถกล่าวได้ว่า ตัวประมาณอัตราส่วนแบบเลขชี้กำลังสำหรับประมาณค่าเฉลี่ยประชากรที่นำเสนอขึ้นมานี้ใหม่เป็นตัวประมาณที่มีประสิทธิภาพดีกว่าตัวประมาณอื่น ๆ ที่เกี่ยวข้องทุกตัว ภายใต้ชุดข้อมูลและสถานการณ์เดียวกันกับที่ศึกษาในงานวิจัยฉบับนี้ งานวิจัยในอนาคตอาจขยายตัวประมาณดังกล่าวไปสู่การสุ่มตัวอย่างแบบแบ่งชั้นหรือแบบเป็นระบบ ภายใต้ปัญหาการไม่ตอบสนอง เพื่อประเมินและยืนยันประสิทธิภาพของตัวประมาณในสถานการณ์ที่หลากหลายและซับซ้อนยิ่งขึ้น นอกจากนี้ ตัวประมาณที่นำเสนอยังมีนัยสำคัญ

เชิงปฏิบัติ (practical implication) โดยสามารถเพิ่มประสิทธิภาพและความแม่นยำของการสำรวจข้อมูลทางเศรษฐกิจ สังคม และการเกษตรได้อย่างมีนัยสำคัญ โดยเฉพาะในสถานการณ์ที่ไม่ตอบแบบสอบถามเกิดขึ้นในอัตราสูง

Acknowledgments

งานวิจัยฉบับนี้สำเร็จได้อย่างสมบูรณ์ด้วยดี โดยได้รับความอนุเคราะห์จากสาขาวิชาสถิติและวิทยาการข้อมูล ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ที่เอื้อเฟื้อสถานที่และอำนวยความสะดวกในการทำงานวิจัย

References

- Abid, M., Abbas, N. Sherwani, R.A.K., & Nazir, H.Z. (2016). Improved ratio estimators for the population mean using non-conventional measures of dispersion. *Pakistan Journal of Statistics and Operations Research*, XII(2), 353-367.
- Ahmadini, A.A.H, Singh, R., Raghav, Y.S., & Kumari, A. (2025). Estimation of population mean using ranked set sampling in the presence of measurement errors. *Kuwait Journal of Science*, 51(3), 1-14.
- Ganie, Z.A., Tarray, T.A., & Rather, K.U.I. (2022). A New Data Gathering Exponential Type Ratio Estimator for the Population Mean. *Journal of Scientific Research & Reports*, 28(10), 148-153.
- Hansen, M.H., & Hurwitz, W.N. (1946). The problem of nonresponse in sample surveys. *Journal of the American Statistical Association*, 41, 517-529.
- Jerajuddin, M., & Kishun, J. (2016). Modified ratio estimators for population mean using size of the sample. selected from population, *IJSRSET*, 2(2), 10-16.
- Khare, B.B., & Sinha, R.R. (2010). On class of estimators for the product of two population means using auxiliary character in presence of non-response. *International Review of Applied Sciences and Engineering*, 2(4), 841-846.
- Rao, P.S.R.S. (1986). Ratio estimation with sub sampling the non-respondents. *Survey Methodology*, 12, 217-230.

- Singh, H.P., Gupta, A., & Tailor, R. (2023). An efficient approach for estimating population mean in simple random sampling using an auxiliary attribute. *Thailand Statistician*, 21(3), 631–659.
- Singh, R., Kumar, M., Chaudhary, M.K., & Smarandache, F. (2009). Estimation of mean in presence of non-response using exponential estimator. *arXiv preprint arXiv:0906.2462*.
- Yadav, S.K., & Adewara, A.A. (2013). On Improved estimation of population mean using qualitative auxiliary information. *Mathematical Theory and Modeling*, 3(11), 42-50.
- Yadav, S.K. Mishra, S.S., & Shukla, A.K. (2014). Improved ratio estimators for population mean based on median using linear combination of population mean and median of an auxiliary variable. *American Journal of Operational Research*, 4(2), 21-27.
- Yadav, S.K. Mishra, S.S., & Shukla, A.K. (2015). Estimation approach to ratio of two inventory population means in stratified random sampling. *American Journal of Operational Research*, 5(4), 96-101.
- Yadav, S.K., Dixit M.K., Dungana, H.N., & Mishra, S.S. (2019). Improved estimators for estimating average yield using supplementary variable. *The International Journal of Mathematical, Engineering and Management Sciences*, 4(5), 1228-38.