

การเปรียบเทียบวิธีการประมาณการแจกแจงของข้อมูลสมมาตรและไม่สมมาตร ด้วยวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป

A Comparison of Estimation Methods for Symmetric and Asymmetric Distributed Data by Jackknifing Method and Bootstrapping Method

กัญญ์พิชญา พุทธะไชยทัศน^{1*}

Kanpichaya Puttachaiyatsad^{1*}

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีประมาณการแจกแจงระหว่างวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป โดยจะประมาณพารามิเตอร์ค่าเฉลี่ย ความแปรปรวน ความเบ้ และความโด่ง ทั้งการประมาณแบบจุด และการประมาณแบบช่วง ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองโดยเทคนิคมอนติคาร์โล สำหรับการแจกแจงปกติ ปโลมปน มีพารามิเตอร์กำหนดความแปรปรวน (σ^2) เท่ากับ 25 เปอร์เซนต์การปโลมปน(p) เท่ากับ 10% และ 30% และสเกลแฟกเตอร์(c) เท่ากับ 5 และ 10 สำหรับการแจกแจงซีกก้าง มีค่าพารามิเตอร์ β เท่ากับ 0.1, 0.5, 1 และ 1.5 สำหรับการแจกแจงแกมมา มีค่าพารามิเตอร์ α เท่ากับ 3, 4, 6 และ 8 ค่าพารามิเตอร์ β เท่ากับ 0.1 โดยกำหนดขนาดตัวอย่าง เท่ากับ 100, 200, 300, 400, 500, 600, 700, 800, 900 และ 1,000 การจำลองข้อมูลกระทำซ้ำ 500 ครั้ง ในแต่ละสถานการณ์ ค่าเฉลี่ยความคลาดเคลื่อนค่าเฉลี่ยสอง และความเอนเอียง เป็นเกณฑ์ที่ใช้ในการเปรียบเทียบการประมาณค่าแบบจุด และสัมประสิทธิ์ความเชื่อมั่น เป็นเกณฑ์ที่ใช้ในการเปรียบเทียบการประมาณค่าแบบช่วง ผลการวิจัยสรุปได้ดังนี้

กรณีการประมาณค่าแบบจุด : พบว่า โดยส่วนใหญ่ประมาณ 71.43% วิธีบูตสเตรป จะมีประสิทธิภาพดีกว่าวิธีแจ๊คไนฟ์

กรณีการประมาณค่าแบบช่วง : พบว่า ทุกค่าพารามิเตอร์ ในทุกๆลักษณะการแจกแจง วิธีบูตสเตรป จะมีประสิทธิภาพดีกว่าวิธีแจ๊คไนฟ์

คำสำคัญ : วิธีแจ๊คไนฟ์ วิธีบูตสเตรป การประมาณค่าแบบจุด การประมาณค่าแบบช่วง

¹ นิสิตสาขาวิชาสถิติ หลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

* Corresponding author: โทรศัพท์ 083-7733456 Email-nu_naung@hotmail.com

Abstract

The objective of this study is to compare the efficiency between Jackknifing method and Bootstrapping method. The parameter to be estimated are mean, variance, skewness and kurtosis. Both point estimation and interval estimation are calculated. Data are generated by using Monte Carlo Simulation technique based on several distribution those are scale contaminated normal distribution, exponential distribution and gamma distribution. For scale contaminated normal distribution, the data are generated specified σ^2 to be 25, percent of contamination to be 10% and 30% and scale factor to be 5 and 10. For exponential distribution, specified to be β 0.1, 0.5, 1 and 1.5. For gamma distribution, specified α to be 3, 4, 6 and 8 and to be β 0.1. The sample sizes are 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1,000. The generated data were repeated 500 times under each situations. The mean square error and biasedness are criteria for comparing the point estimations. The confidence coefficient is criteria for comparing the interval estimations. The results of this study are summarized as follows:

It is found that about 71.43% Bootstrapping method is more efficient than Jackknifing method in point estimations.

It is found that every specified parameter in every generated distribution, Bootstrapping method is more efficient than Jackknifing method in interval estimations.

Keywords : Jackknife, Bootstrap, Point Estimation, Interval Estimation

คำนำ

การศึกษาวิจัยดำเนินงานต่างๆ เพื่อหาข้อสรุปเกี่ยวกับประชากร มักจะใช้ระเบียบวิธีทางสถิติ ที่เรียกว่า การอนุมานเชิงสถิติ ซึ่งประกอบด้วย การประมาณค่าพารามิเตอร์ และการทดสอบสมมติฐาน

การศึกษาค่าที่บอกถึงลักษณะของประชากร ซึ่งเรียกว่าค่าพารามิเตอร์นั้น จะทราบได้ก็ต่อเมื่อ จะต้องทำการแจกแจงทุกๆ หน่วยในประชากร ในกรณีที่ประชากรมีขนาดใหญ่ การแจกแจงทุกหน่วยในประชากร จะต้องเสียเวลา และค่าใช้จ่ายจำนวนมาก ดังนั้นจึงใช้การเลือกตัวแทนเพียงบางส่วนของประชากรมาศึกษา แล้วใช้ค่าที่ได้จากตัวอย่าง ซึ่งเรียกว่า ค่าสถิติ ไปประมาณค่าพารามิเตอร์ของประชากร ซึ่งการประมาณค่าพารามิเตอร์จะสามารถแบ่งออกได้เป็นสองประเภท ได้แก่ การประมาณแบบจุด และการประมาณแบบช่วง ดังนั้นจึงควรเลือกใช้วิธีการประมาณที่มีความเหมาะสมกับเรื่องที่สนใจศึกษา และสอดคล้องกับลักษณะของประชากร เพื่อให้ค่าประมาณที่คำนวณได้มีความถูกต้อง และน่าเชื่อถือมากยิ่งขึ้น

ในกรณีที่ตัวอย่างถูกสุ่มมาจากประชากร ด้วยการเลือกตัวอย่างสุ่มแบบง่าย เมื่อเราต้องการทราบลักษณะการแจกแจงของข้อมูล เราสามารถประมาณการแจกแจงโดยใช้การประมาณค่าพารามิเตอร์ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ และค่าความโด่ง ได้จากข้อมูลตัวอย่างแล้วนำไปวาดกราฟ ซึ่งการหาค่าพารามิเตอร์ค่าเฉลี่ย ความแปรปรวน ความเบ้ และ ความโด่ง จะสามารถบอกถึงลักษณะการแจกแจงของประชากรได้ แต่ในกรณีที่ไมทราบการแจกแจงของประชากรมาก่อน หรือข้อมูลตัวอย่างมีจำนวนน้อย หรือตัวสถิติที่ใช้ประมาณค่าพารามิเตอร์อยู่ในรูปที่ยุ่งยาก การประมาณค่าดังกล่าว ไม่อาจจะหาได้ด้วยวิธีง่าย ๆ

ในปี ค.ศ.1956 เควโนอิล (Quenouille) ได้เสนอวิธีแจ๊คไนฟ์ และได้มีการรวบรวมและพัฒนาต่อโดย มิลเลอร์ (Miller) ในปี ค.ศ.1974 ต่อมาปี ค.ศ.1979 แบรดลีย์ เอฟรอน (Bradley Efron) ได้เสนอวิธีบูตสเตรปเป็นคนแรก ซึ่งวิธีแจ๊คไนฟ์และวิธีบูตสเตรป เป็นวิธีการหาตัวประมาณของพารามิเตอร์ ด้วยวิธีการทางนอนพารามตริก (Nonparametric) ที่ไม่มีข้อจำกัดเกี่ยวกับลักษณะการแจกแจงของข้อมูล ซึ่งอาศัยหลักการสุ่มซ้ำ (Resampling) จากตัวอย่างสุ่มเพียงชุดเดียว เพื่อสร้างความผันแปรให้กับค่าของตัวสถิติ โดยถือเสมือนว่า ตัวอย่างที่ได้จากการสุ่มซ้ำจากตัวอย่างชุดเดี่ยวนั้น เป็นตัวอย่างสุ่มซ้ำๆ จากประชากรจริงๆ อย่างไรก็ตาม สองวิธีนี้มีข้อจำกัดคือ การสุ่มตัวอย่างซ้ำขึ้นอยู่กับตัวอย่างเพียงชุดเดียว ที่สุ่มได้จากประชากรคั้งนั้นอาจทำให้ค่าบางค่า ไม่มีโอกาสที่จะเกิดขึ้นในการสุ่มตัวอย่างซ้ำได้เลย

ผู้วิจัยสนใจที่จะทำการศึกษาและเปรียบเทียบประสิทธิภาพ ของวิธีแจ๊คไนฟ์ และวิธีบูตสเตรป ในการประมาณค่าแบบจุด และการประมาณแบบช่วง สำหรับพารามิเตอร์ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ และค่าความโด่ง โดยสนใจที่จะศึกษาข้อมูลที่มีการแจกแจงสมมาตรและไม่สมมาตร เพื่อที่จะหาข้อสรุป ไว้เป็นตัวเลือกในการประมาณการแจกแจงของประชากรได้

อุปกรณ์และวิธีการ

ข้อมูลที่ใช้ในการศึกษาสร้างจากตัวอย่างสุ่มที่มีการแจกแจงแบบต่างๆ

สร้างข้อมูลที่มีการแจกแจงตามลักษณะที่ต้องการศึกษา สำหรับการประมาณค่า ภายใต้ขอบเขตที่ต้องการศึกษาดังนี้ [1]

1. การแจกแจงปกติปลอมปน (Scale Contaminated Normal Distribution)

ฟังก์ชันความหนาแน่นของ X อยู่ในรูปของ

$$f(x) = (1 - p)N(\mu, \sigma^2) + pN(\mu, c^2\sigma^2)$$

โดยที่ $c > 0$, $0 \leq p \leq 1$

ค่าคาดหวัง $E(X) = \mu$

ความแปรปรวน $Var(X) = (1 - p)\sigma^2 + pc^2\sigma^2$

ความเบ้ $\gamma_1 = 0$

ความโด่ง $\gamma_2 = 0$

ขอบเขตของการวิจัย กำหนดให้ข้อมูลมีค่าเฉลี่ย (μ) เท่ากับ 0 พารามิเตอร์ที่กำหนด

ความแปรปรวน (σ^2) เท่ากับ 25 เปอร์เซนต์การปลอมปน (p) เท่ากับ 10%, 30% และ

สเกลแฟคเตอร์ (c) เท่ากับ 5, 10

2. การแจกแจงชี้กำลัง (Exponential Distribution)

ฟังก์ชันความหนาแน่นของ X อยู่ในรูปของ

$$f(x) = \beta e^{-\beta x} \quad , x > 0$$

โดยที่ $\beta > 0$

$$\text{ค่าคาดหวัง } E(X) = \frac{1}{\beta}$$

$$\text{ความแปรปรวน } Var(X) = \frac{1}{\beta^2}$$

$$\text{ความเบ้ } \gamma_1 = 2$$

$$\text{ความโด่ง } \gamma_2 = 6$$

ขอบเขตของการวิจัย กำหนดให้ข้อมูลมีค่าพารามิเตอร์ β เท่ากับ 0.1, 0.5, 1 และ 1.5

3. การแจกแจงแกมมา (Gamma Distribution)

ฟังก์ชันความหนาแน่นของ X อยู่ในรูปของ

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad , x > 0$$

โดยที่ $\alpha > 0$, $\beta > 0$

$$\text{ค่าคาดหวัง } E(X) = \frac{\alpha}{\beta}$$

$$\text{ความแปรปรวน } Var(X) = \frac{\alpha}{\beta^2}$$

$$\text{ความเบ้ } \gamma_1 = \frac{2}{\sqrt{\alpha}}$$

$$\text{ความโด่ง } \gamma_2 = \frac{6}{\alpha}$$

ขอบเขตของการวิจัย กำหนดให้ข้อมูลมีค่าพารามิเตอร์ α เท่ากับ 3, 4, 6 และ 8

ค่าพารามิเตอร์ β เท่ากับ 0.1

การประมาณค่าพารามิเตอร์

วิธีการประมาณค่าของพารามิเตอร์ค่าเฉลี่ย(μ) ค่าความแปรปรวน(σ^2) ค่าความเบ้(γ_1) และค่าความโด่ง(γ_2) สำหรับการแจกแจงดังกล่าวด้วยวิธีแฉีกไนฟ กับ วิธีบูตสแตรป

กำหนดให้ θ เป็นค่าจริงของพารามิเตอร์

$\hat{\theta}_j$ เป็นค่าประมาณของพารามิเตอร์ด้วยวิธีแฉีกไนฟ

$\hat{\theta}_B$ เป็นค่าประมาณของพารามิเตอร์ด้วยวิธีบูตสแตรป

โดยที่ $\theta = \{\mu, \sigma^2, \gamma_1, \gamma_2\}$, $\hat{\theta}_j = \{\hat{\mu}_j, \hat{\sigma}_j^2, \hat{\gamma}_{1j}, \hat{\gamma}_{2j}\}$, $\hat{\theta}_B = \{\hat{\mu}_B, \hat{\sigma}_B^2, \hat{\gamma}_{1B}, \hat{\gamma}_{2B}\}$

จำนวนตัวอย่างสุ่ม(n) ที่ใช้ในการศึกษาครั้งนี้เป็น 100, 200, 300, 400, 500, 600, 700, 800, 900 และ 1,000

การศึกษาครั้งนี้ ทำการจำลองข้อมูลโดยใช้เทคนิคมอนติคาร์โล (Monte Carlo Simulation : M) โดยการจำลองในแต่ละสถานการณ์จะกระทำซ้ำ 500 รอบ

1. การประมาณค่าพารามิเตอร์ด้วยวิธีแจ็กไนฟ์ (Jackknifing Method)

วิธีแจ็กไนฟ์ถูกเสนอโดย Quenouille [2] เพื่อลดความเอนเอียงของตัวประมาณ วิธีการดังกล่าวนี้ สามารถหาค่าประมาณของตัวพารามิเตอร์ และค่าประมาณความแปรปรวนของตัวประมาณต่างๆได้ ซึ่งวิธีแจ็กไนฟ์ได้มีการรวบรวมและพัฒนาต่อโดย Miller [3] วิธีแจ็กไนฟ์จัดเป็นวิธีที่ใช้การเลือกตัวอย่างซ้ำวิธีหนึ่ง ซึ่งเป็นการเลือกตัวอย่างใหม่จากตัวอย่างสุ่มเพียงชุดเดียวโดยมีวิธีดำเนินการดังนี้

สุ่มตัวอย่างมา n ตัว คือ $x_1, x_2, \dots, x_n$ จากประชากรที่มีการแจกแจงแบบต่างๆ ให้ θ เป็นพารามิเตอร์ที่ต้องการประมาณในประชากรดังกล่าวนี้ และให้ $\hat{\theta}_j$ เป็นค่าประมาณของพารามิเตอร์ θ ด้วยวิธีแจ็กไนฟ์ ที่คำนวณจากข้อมูลตัวอย่างขนาด n ในการหาค่าประมาณด้วยวิธีแจ็กไนฟ์ จะเริ่มจาก

ครั้งที่ 1 ตัดค่า x_1 ออกจากตัวอย่าง แล้วคำนวณค่าประมาณของ θ จาก $x_2, x_3, \dots, x_n$ จะได้ค่าประมาณคือ $\hat{\theta}_{(1)}$

ครั้งที่ 2 ตัดค่า x_2 ออกจากตัวอย่าง แล้วคำนวณค่าประมาณของ θ จาก $x_1, x_3, \dots, x_n$ จะได้ค่าประมาณคือ $\hat{\theta}_{(2)}$

⋮

ครั้งที่ n ตัดค่า x_n ออกจากตัวอย่าง แล้วคำนวณค่าประมาณของ θ จาก $x_1, x_2, \dots, x_{n-1}$ จะได้ค่าประมาณคือ $\hat{\theta}_{(n)}$

ด้วยการทำซ้ำ ดังที่กล่าวมาแล้วจำนวน n ครั้ง จะได้ค่าประมาณของ θ จำนวน n ตัวคือ $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \dots, \hat{\theta}_{(n)}$ ให้ $\hat{\theta}_j$ เป็นตัวประมาณของพารามิเตอร์ θ ด้วยวิธีแจ็กไนฟ์ ซึ่งการหาค่าประมาณแบบจุดจะถูกกำหนดโดย

$$\hat{\theta}_j = \frac{\sum_{i=1}^n \hat{\theta}_{(i)}}{n}$$

การหาค่าประมาณแบบช่วงของพารามิเตอร์ θ ด้วยวิธีแจ็กไนฟ์ ที่ระดับนัยสำคัญ α จะได้ว่า

$$P(\hat{\theta}_{jL} < \theta < \hat{\theta}_{jU}) = 1 - \alpha$$

ซึ่งจะหาค่าประมาณของ $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \dots, \hat{\theta}_{(n)}$ ที่ได้นำมาจัดเรียงจากค่าน้อยไปหามาก จากนั้นคำนวณค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(\alpha/2)$ กำหนดให้เป็น $\hat{\theta}_{jL}$ และค่าที่ตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(1 - \alpha/2)$ กำหนดให้เป็น $\hat{\theta}_{jU}$ ดังนั้น จะได้ช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ด้วยวิธีแจ็กไนฟ์ คือ $[\hat{\theta}_{jL}, \hat{\theta}_{jU}]$

2. การประมาณค่าพารามิเตอร์ด้วยวิธีบูตสแตรป (Bootstrapping method)

วิธีบูตสแตรปจัดเป็นวิธีการประมาณค่าพารามิเตอร์โดยใช้การสุ่มตัวอย่างซ้ำ เช่นเดียวกับวิธีของแจ็กไนฟ์ แต่วิธีการสุ่มตัวอย่างซ้ำจะใช้การสร้างตัวอย่างชุดใหม่จากตัวอย่างสุ่มที่มีเพียงชุดเดียว โดยการสุ่มตัวอย่างแบบคืนที่ (Resampling with Replacement) วิธีการนี้ถูกเสนอโดย Efron [4] และได้พัฒนาต่อมาโดย Efron [5]

Efron [4] เสนอให้ใช้วิธีการสุ่มตัวอย่างแบบคืนที่ขนาด n จากตัวอย่างสุ่มชุดเดียวที่มี เพื่อสร้างชุดตัวอย่าง

ขนาด n ที่เป็นไปได้ นั่นคือ แทนที่จะสุ่มตัวอย่างซ้ำ ๆ จากประชากรที่มีฟังก์ชันการแจกแจง F โดยตรง จะใช้การสุ่มตัวอย่างจาก Empirical distribution function (F_n) ของข้อมูลตัวอย่างโดยมีวิธีการดำเนินการดังนี้

สุ่มตัวอย่างมา n ตัว คือ $x_1, x_2, \dots, x_n$ ที่เป็นอิสระกันมาจากการแจกแจงแบบต่างๆ ให้ θ เป็นพารามิเตอร์ที่ต้องการประมาณในประชากรดังกล่าวนี้ และให้ $\hat{\theta}_B$ เป็นค่าประมาณของพารามิเตอร์ θ ด้วยวิธีบูตสเตรปที่คำนวณจากข้อมูลตัวอย่างขนาด n สร้างฟังก์ชันการแจกแจงโดยให้ความน่าจะเป็นของ $x_i, i = 1, 2, \dots, n$ เป็น $\frac{1}{n}$ ซึ่งเรียกฟังก์ชันการแจกแจงแบบนี้ว่า Empirical distribution function

การสุ่มตัวอย่างจะทำการสุ่มตัวอย่างทีละ 1 ค่าจำนวน n ครั้ง จากชุดของตัวอย่าง $x_1, x_2, \dots, x_n$ โดยค่าที่ได้จะคืนกลับไปในชุดตัวอย่างก่อนที่จะมีการสุ่มตัวอย่างครั้งต่อไป ให้ $x_1^*, x_2^*, \dots, x_n^*$ เป็นชุดของตัวอย่างขนาด n ที่สุ่มได้ ซึ่งจะเรียกชุดของตัวอย่างดังกล่าวนี้ว่า ตัวอย่างบูตสเตรป (Bootstrap sample) ซึ่งการหาค่าประมาณด้วย วิธีบูตสเตรป จะเริ่มจาก

ครั้งที่ 1 ทำการสุ่มตัวอย่างทีละ 1 ค่าแบบคืนที่จำนวน n ครั้ง จากชุดของตัวอย่าง จะได้ $x_1^*, x_2^*, \dots, x_n^*$ แล้วคำนวณค่าประมาณของ θ จะได้ค่าประมาณ คือ $\hat{\theta}_1^*$

ครั้งที่ 2 ทำการสุ่มตัวอย่างทีละ 1 ค่าแบบคืนที่จำนวน n ครั้ง จากชุดของตัวอย่าง จะได้ $x_1^*, x_2^*, \dots, x_n^*$ แล้วคำนวณค่าประมาณของ θ จะได้ค่าประมาณ คือ $\hat{\theta}_2^*$

⋮

ครั้งที่ B ทำการสุ่มตัวอย่างทีละ 1 ค่าแบบคืนที่จำนวน n ครั้ง จากชุดของตัวอย่าง จะได้ $x_1^*, x_2^*, \dots, x_n^*$ แล้วคำนวณค่าประมาณของ θ จะได้ค่าประมาณ คือ $\hat{\theta}_B^*$

ด้วยการทำซ้ำ ดังที่กล่าวมาแล้วจำนวน B ครั้ง จะได้ค่าประมาณของ θ จำนวน B ตัวคือ $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$ นำมาสร้างฮิสโตแกรม (histogram) โดยกำหนดให้แต่ละตัวมีความน่าจะเป็นเท่ากัน เท่ากับ $\frac{1}{B}$ จะได้การแจกแจงของตัวสถิติตัวอย่างบูตสเตรป (the bootstrap sampling distribution)

ให้ $\hat{\theta}_B$ เป็นตัวประมาณของพารามิเตอร์ θ ด้วยวิธีบูตสเตรป ซึ่งการหาค่าประมาณแบบจุดจะถูกกำหนดโดย

$$\hat{\theta}_B = \frac{\sum_{i=1}^B \hat{\theta}_i^*}{B}$$

การหาค่าประมาณแบบช่วงของพารามิเตอร์ θ ด้วยวิธีบูตสเตรป ที่ระดับนัยสำคัญ α จะได้ว่า

$$P(\hat{\theta}_{BL} < \theta < \hat{\theta}_{BU}) = 1 - \alpha$$

ซึ่งจะหาจาก การแจกแจงตัวสถิติตัวอย่างบูตสเตรป $\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*$ ที่ได้ นำมาจัดเรียงจากค่าน้อยไปหามาก จากนั้นคำนวณหาตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(\alpha/2)$ กำหนดให้เป็น $\hat{\theta}_{BL}$ และหาตำแหน่งเปอร์เซ็นต์ไทล์ที่ $100(1 - \alpha/2)$ กำหนดให้เป็น $\hat{\theta}_{BU}$ ดังนั้น จะได้ช่วงความเชื่อมั่น $(1 - \alpha)100\%$ ด้วยวิธีบูตสเตรป คือ $[\hat{\theta}_{BL}, \hat{\theta}_{BU}]$

เกณฑ์ในการตัดสินใจ

การประมาณค่าแบบจุด ใช้การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error หรือ MSE) และค่าความเอนเอียง (Biasedness)

1. ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error หรือ MSE) ซึ่งหาจากค่าเฉลี่ยของกำลังสองของผลต่างระหว่างค่าประมาณของพารามิเตอร์ กับ ค่าจริงของพารามิเตอร์ ดังนี้

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองวิธีเจ็คนไฟ

$$MSE(\hat{\theta}_J) = \frac{\sum_{i=1}^M (\hat{\theta}_{Ji} - \theta)^2}{M}$$

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองวิธีบูตสเตรป

$$MSE(\hat{\theta}_B) = \frac{\sum_{i=1}^M (\hat{\theta}_{Bi} - \theta)^2}{M}$$

โดย M คือจำนวนรอบที่กระทำซ้ำในแต่ละสถานการณ์ ในงานวิจัยนี้กำหนดให้ M มีค่าเท่ากับ 500

ถ้าวิธีการประมาณค่าแบบใดที่ทำให้ได้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ ต่ำกว่าจะถือว่าการประมาณค่าจากวิธีนั้นเป็นวิธีที่มีประสิทธิภาพมากกว่า

2. ความเอนเอียง (Biasedness) เป็นค่าที่ใช้วัดว่าค่าเฉลี่ยของตัวสถิติที่ได้ห่างจากฟังก์ชันพารามิเตอร์ θ มากน้อยเพียงใด นอกจากนี้ยังบอกทิศทางได้ด้วยว่า ตัวสถิติที่ได้ให้ค่าสูงหรือต่ำกว่า พารามิเตอร์ ซึ่งค่าความเอนเอียง หาจาก

ค่าความเอนเอียง ของวิธีเจ็คนไฟ

$$Bias(\hat{\theta}_J) = E[\hat{\theta}_J] - \theta$$

ค่าความเอนเอียง ของวิธีบูตสเตรป

$$Bias(\hat{\theta}_B) = E[\hat{\theta}_B] - \theta$$

โดยถ้าวิธีการประมาณค่าแบบใดที่ทำให้ได้ค่าความเอนเอียงของตัวประมาณ ต่ำกว่าจะถือว่าการประมาณค่าจากวิธีนั้นเป็นวิธีที่มีประสิทธิภาพมากกว่า และเราเรียก $\hat{\theta}$ ว่าเป็นตัวประมาณที่ไม่เอนเอียงของพารามิเตอร์ θ ถ้า

$$E[\hat{\theta}] = \theta$$

การประมาณค่าแบบช่วง ใช้การเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) จากการตรวจสอบว่าช่วงความเชื่อมั่นที่คำนวณจากแต่ละวิธีการประมาณครอบคลุมค่าพารามิเตอร์ θ หรือไม่ในแต่ละรอบ หากช่วงความเชื่อมั่นที่คำนวณได้ครอบคลุมค่าพารามิเตอร์ θ จะทำการนับจำนวนครั้งและบวกสะสมค่าไว้ นำมาหาค่าสัมประสิทธิ์ความเชื่อมั่น ดังนี้

$$\text{ค่าสัมประสิทธิ์ความเชื่อมั่น} = \frac{\text{จำนวนครั้งที่ทั้งหมดที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ } \theta}{M}$$

โดย M คือจำนวนรอบที่กระทำซ้ำในแต่ละสถานการณ์ ในงานวิจัยนี้กำหนดให้ M มีค่าเท่ากับ 500
ถ้าวิธีการประมาณค่าแบบใดที่ทำให้ค่าสัมประสิทธิ์ความเชื่อมั่น สูงกว่าจะถือว่าค่าประมาณค่าจากวิธีนั้นเป็นวิธี
ที่เหมาะสมมากกว่าในการประมาณค่าแบบช่วงสำหรับสถานการณ์นั้นๆ

ผลการวิจัยและอภิปรายผล

จากการวิจัยการเปรียบเทียบประสิทธิภาพวิธีประมาณการแจกแจงระหว่างวิธีแฉกไนฟ์ กับ วิธีบูตสเตรป
โดยจะประมาณพารามิเตอร์ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ และค่าความโด่ง ซึ่งแบ่งเป็น 2 กรณีคือ
การประมาณค่าแบบจุด และ การประมาณค่าแบบช่วง เมื่อข้อมูลมีการแจกแจงปกติปลอมปน การแจกแจงซีกกำลัง
และการแจกแจงแกมมา ได้ผลการวิจัยดังนี้

กรณีที่ 1 การเปรียบเทียบประสิทธิภาพวิธีการประมาณค่าแบบจุด ระหว่างวิธีแฉกไนฟ์ กับ วิธีบูตสเตรป โดยจะ
พิจารณาจากการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error) และค่าความเอนเอียง (Biasedness)
รวมทั้งการทดสอบสมมติฐาน ที่ระดับนัยสำคัญ 0.05 ดังนี้

สมมติฐานสำหรับการทดสอบค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง

$$H_0: MSE_J = MSE_B$$

$$H_1: MSE_J \neq MSE_B$$

สมมติฐานสำหรับการทดสอบความเอนเอียง

$$H_0: |BIAS_J| = |BIAS_B|$$

$$H_1: |BIAS_J| \neq |BIAS_B|$$

ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 1-4

ตารางที่ 1 แสดงค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error) ระหว่างวิธีแจกไนฟี กับ วิธีบูตสเตรป
เมื่อข้อมูลมีการแจกแจงแบบปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$

n	ค่าเฉลี่ย		ความแปรปรวน		ความเบ้		ความโด่ง	
	วิธีแจกไนฟี	วิธีบูตสเตรป	วิธีแจกไนฟี	วิธีบูตสเตรป	วิธีแจกไนฟี	วิธีบูตสเตรป	วิธีแจกไนฟี	วิธีบูตสเตรป
100	2.0614	2.1053	3020.3950	2986.5550	0.7041	0.5601	5.3722	4.3311
200	1.0507	1.0620	1499.5740	1513.8200	0.4182	0.3676	2.9127	2.4132
300	0.7116	0.7148	978.4244	977.2894	0.2748	0.2508	2.0367	1.8355
400	0.4957	0.4991	821.9920	824.1560	0.1967	0.1819	1.6910	1.5353
500	0.4052	0.4067	634.2891	629.6037	0.1699	0.1589	1.2775	1.1674
600	0.3420	0.3457	516.0232	514.1922	0.1347	0.1274	1.2844	1.1678
700	0.3048	0.3064	407.6393	408.3515	0.1244	0.1195	1.0409	0.9741
800	0.2447	0.2444	400.0039	401.5295	0.1007	0.0963	0.9706	0.8794
900	0.2410	0.2411	339.0845	338.3073	0.0951	0.0915	0.7917	0.7291
1000	0.1862	0.1869	295.1873	294.4032	0.0791	0.0763	0.6001	0.5767

จากตารางที่ 1 การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ระหว่างวิธีแจกไนฟี กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$ สามารถอธิบายได้ดังนี้

สำหรับการประมาณค่าเฉลี่ย พบว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจกไนฟี ต่ำกว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป ทุกขนาดตัวอย่าง ยกเว้นที่ $n=800$

สำหรับการประมาณความแปรปรวน พบว่า ส่วนใหญ่ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป ต่ำกว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจกไนฟี ยกเว้นที่ $n=200,400,700,800$

สำหรับการประมาณความเบ้ พบว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป ต่ำกว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจกไนฟี ทุกขนาดตัวอย่าง

สำหรับการประมาณความโด่ง พบว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป ต่ำกว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจกไนฟี ทุกขนาดตัวอย่าง

ตารางที่ 2 ตารางการเปรียบเทียบการทดสอบสมมติฐานของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error) ระหว่างวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$

ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง				
พารามิเตอร์	วิธีประมาณ	MSE เฉลี่ย	t^1	p-value
ค่าเฉลี่ย	วิธีแจ๊คไนฟ์	0.306	-	0.052
	วิธีบูตสเตรป	0.307		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	-0.001		
ความแปรปรวน	วิธีแจ๊คไนฟ์	235.214	0.941	0.372
	วิธีบูตสเตรป	234.319		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	0.895		
ความเบ้	วิธีแจ๊คไนฟ์	0.142	4.343	0.002
	วิธีบูตสเตรป	0.123		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	0.019		
ความโด่ง	วิธีแจ๊คไนฟ์	10.860	7.106	0.000
	วิธีบูตสเตรป	8.158		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	2.702		

¹ สมมติฐานของการทดสอบ คือ $H_0: MSE_J = MSE_B$, $H_1: MSE_J \neq MSE_B$

จากตารางที่ 2 การเปรียบเทียบการทดสอบสมมติฐานของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ระหว่างวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$ สามารถอธิบายได้ดังนี้

สำหรับการประมาณค่าเฉลี่ย และความแปรปรวน พบว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจ๊คไนฟ์ และ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป ไม่มีนัยสำคัญของความแตกต่าง

สำหรับการประมาณความเบ้ และความโด่ง พบว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจ๊คไนฟ์ และ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป มีนัยสำคัญแสดงความแตกต่าง โดยที่ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีบูตสเตรป มีค่าน้อยกว่า ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของวิธีแจ๊คไนฟ์

ตารางที่ 3 แสดงค่าความเอนเอียง(Biasedness) ระหว่างวิธีแจกไนฟี กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงแบบปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$

N	ค่าเฉลี่ย		ความแปรปรวน		ความเบ้		ความโค้ง	
	วิธีแจกไนฟี	วิธีบูตสเตรป	วิธีแจกไนฟี	วิธีบูตสเตรป	วิธีแจกไนฟี	วิธีบูตสเตรป	วิธีแจกไนฟี	วิธีบูตสเตรป
100	0.0323	0.0346	-0.0471	-1.6738	0.0172	0.0144	-0.6133	-0.9935
200	-0.0301	-0.0319	-0.8295	-1.8191	0.0541	0.0490	-0.2366	-0.4649
300	-0.0076	-0.0081	-2.3398	-2.9965	0.0049	0.0039	-0.2853	-0.4416
400	0.0235	0.0222	-0.0680	-0.5617	-0.0079	-0.0071	-0.2051	-0.3365
500	-0.0001	0.0009	-0.4907	-0.9141	0.0320	0.0299	-0.1355	-0.2428
600	0.0062	0.0058	-1.7778	-2.1193	0.0121	0.0126	-0.1154	-0.2079
700	-0.0298	-0.0293	0.0786	-0.1751	-0.0149	-0.0142	-0.1884	-0.2657
800	-0.0100	-0.0102	0.2609	-0.0071	-0.0009	-0.0013	-0.0809	-0.1542
900	0.0146	0.0133	-0.2336	-0.5186	0.0175	0.0168	0.0027	-0.0660
1000	0.0075	0.0072	0.9250	0.7025	-0.0109	-0.0112	-0.1261	-0.1836

จากตารางที่ 3 การเปรียบเทียบค่าความเอนเอียงระหว่างวิธีแจกไนฟี กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$ สามารถอธิบายได้ดังนี้

สำหรับการประมาณค่าเฉลี่ย พบว่า ค่าความเอนเอียงของวิธีแจกไนฟี ต่ำกว่า ค่าความเอนเอียงของวิธีบูตสเตรป ยกเว้นที่ $n=400,600,700,900,1000$

สำหรับการประมาณความแปรปรวน พบว่า ส่วนใหญ่ค่าความเอนเอียงของวิธีแจกไนฟี ต่ำกว่า ค่าความเอนเอียงของวิธีบูตสเตรป ยกเว้นที่ $n=800,1000$

สำหรับการประมาณความเบ้ พบว่า ส่วนใหญ่ค่าความเอนเอียงของวิธีบูตสเตรป ต่ำกว่า ค่าความเอนเอียงของวิธีแจกไนฟี ยกเว้นที่ $n=600,800,1000$

สำหรับการประมาณความโค้ง พบว่า ค่าความเอนเอียงของวิธีแจกไนฟี ต่ำกว่า ค่าความเอนเอียงของวิธีบูตสเตรป ทุกขนาดตัวอย่าง

ตารางที่ 4 ตารางการเปรียบเทียบการทดสอบสมมติฐานของค่าความเอนเอียง(Biasedness) ระหว่างวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$

ความเอนเอียง				
พารามิเตอร์	วิธีประมาณ	BIAS เฉลี่ย	t ¹	p-value
ค่าเฉลี่ย	วิธีแจ๊คไนฟ์	0.016	-0.516	0.618
	วิธีบูตสเตรป	0.016		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	0.000		
ความแปรปรวน	วิธีแจ๊คไนฟ์	0.705	-2.501	0.034
	วิธีบูตสเตรป	1.149		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	-0.444		
ความเบ้	วิธีแจ๊คไนฟ์	0.017	2.215	0.054
	วิธีบูตสเตรป	0.016		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	0.001		
ความโด่ง	วิธีแจ๊คไนฟ์	0.199	-4.322	0.002
	วิธีบูตสเตรป	0.336		
	วิธีแจ๊คไนฟ์ - วิธีบูตสเตรป	-0.137		

¹ สมมติฐานของการทดสอบ คือ $H_0: |BIAS_J| = |BIAS_B|$, $H_1: |BIAS_J| \neq |BIAS_B|$

จากตารางที่ 4 การเปรียบเทียบการทดสอบสมมติฐานของค่าความเอนเอียง ระหว่างวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$ สามารถอธิบายได้ดังนี้

สำหรับการประมาณค่าเฉลี่ย พบว่า ค่าความเอนเอียงของวิธีแจ๊คไนฟ์ และ ค่าความเอนเอียงของวิธีบูตสเตรป ไม่มีนัยสำคัญของความแตกต่าง

สำหรับการประมาณความแปรปรวน พบว่า ค่าความเอนเอียงของวิธีแจ๊คไนฟ์ และ ค่าความเอนเอียงของวิธีบูตสเตรป มีนัยสำคัญแสดงความแตกต่าง โดยที่ ค่าความเอนเอียงของวิธีแจ๊คไนฟ์ มีค่าน้อยกว่า ค่าความเอนเอียงของวิธีบูตสเตรป

กรณีที่ 2 การเปรียบเทียบประสิทธิภาพวิธีการประมาณค่าแบบช่วง ระหว่างวิธีแจ๊คไนฟ์ กับ วิธีบูตสเตรป โดยจะพิจารณาจากการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) ซึ่งผลการวิจัยได้นำเสนอในตารางที่ 5

ตารางที่ 5 แสดงค่าสัมประสิทธิ์ความเชื่อมั่น ระหว่างวิธีแจกไนฟ์ กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงแบบปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$

n	ค่าเฉลี่ย		ความแปรปรวน		ความเบ้		ความโด่ง	
	วิธีแจกไนฟ์	วิธีบูตสเตรป	วิธีแจกไนฟ์	วิธีบูตสเตรป	วิธีแจกไนฟ์	วิธีบูตสเตรป	วิธีแจกไนฟ์	วิธีบูตสเตรป
100	0.1920	0.9340	0.1060	0.8980	0.0860	0.8960	0.0940	0.8180
200	0.1320	0.9360	0.0940	0.9140	0.0340	0.8920	0.0780	0.8600
300	0.1120	0.9400	0.0780	0.9240	0.0280	0.9180	0.0660	0.8340
400	0.1020	0.9440	0.0700	0.9220	0.0340	0.9240	0.0360	0.8500
500	0.0820	0.9620	0.0540	0.9380	0.0160	0.9340	0.0440	0.8820
600	0.0800	0.9440	0.0340	0.9300	0.0200	0.9400	0.0360	0.8860
700	0.0620	0.9360	0.0560	0.9580	0.0220	0.9280	0.0340	0.8700
800	0.0920	0.9380	0.0560	0.9360	0.0120	0.9300	0.0320	0.8800
900	0.0660	0.9380	0.0580	0.9480	0.0180	0.9340	0.0380	0.9220
1000	0.0600	0.9540	0.0440	0.9420	0.0180	0.9380	0.0380	0.9140

จากตารางที่ 5 การเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่น ระหว่างวิธีแจกไนฟ์ กับ วิธีบูตสเตรป เมื่อข้อมูลมีการแจกแจงแบบปกติปลอมปน ซึ่งมีค่า $p=30\%$, $c=5$ สามารถอธิบายได้ดังนี้

สำหรับการประมาณค่าเฉลี่ย ความแปรปรวน ความเบ้ และความโด่ง พบว่า ค่าสัมประสิทธิ์ความเชื่อมั่นของวิธีบูตสเตรป สูงกว่า ค่าสัมประสิทธิ์ความเชื่อมั่นของวิธีแจกไนฟ์ ทุกขนาดตัวอย่างในทุกๆค่าประมาณพารามิเตอร์

สรุปผลการวิจัย

กรณีที่ 1 การเปรียบเทียบประสิทธิภาพวิธีการประมาณค่าพารามิเตอร์แบบจุด ระหว่างวิธีแจกไนฟ์ กับ วิธีบูตสเตรป

โดยพิจารณาจากการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error) ของทั้ง 2 วิธีก่อน โดยถ้าวิธีการประมาณค่าแบบใดที่ทำให้ได้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ ต่ำกว่าจะสรุปได้ว่าการประมาณค่าจากวิธีนั้นเป็นวิธีที่มีประสิทธิภาพมากกว่า แต่ถ้าวิธีการประมาณค่าทั้ง 2 วิธี ให้ค่าเฉลี่ยความคลาดเคลื่อนกำลังสองของตัวประมาณ เท่ากัน จะไม่สามารถสรุปได้จึงพิจารณาต่อจากการเปรียบเทียบค่าความเอนเอียง (Biasedness) โดยถ้าวิธีการประมาณค่าแบบใดที่ทำให้ได้ค่าความเอนเอียงของตัวประมาณ ต่ำกว่าจะถือว่าวิธีการประมาณค่าจากวิธีนั้นเป็นวิธีที่มีประสิทธิภาพมากกว่าในการประมาณค่าแบบจุด โดยจะประมาณพารามิเตอร์ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ และค่าความโด่งของข้อมูลที่มีการแจกแจงแบบปกติปลอมปน การแจกแจงชี้กำลัง และการแจกแจงแกมมา ซึ่งผลการวิจัยสรุปได้ว่า

1. สำหรับการประมาณค่าเฉลี่ย พบว่า วิธีแจ็คไนฟ์ มีประสิทธิภาพดีเท่ากับวิธีบูตสเตรป ทุกค่าพารามิเตอร์
 2. สำหรับการประมาณความแปรปรวน พบว่า โดยส่วนใหญ่วิธีแจ็คไนฟ์ มีประสิทธิภาพดีเท่ากับ วิธีบูตสเตรป ยกเว้นบางกรณี วิธีแจ็คไนฟ์ มีประสิทธิภาพดีกว่า วิธีบูตสเตรป
 3. สำหรับการประมาณความเบ้ พบว่า โดยส่วนใหญ่วิธีบูตสเตรปมีประสิทธิภาพดีกว่า วิธีแจ็คไนฟ์ ยกเว้นบางกรณี วิธีแจ็คไนฟ์ จะมีประสิทธิภาพดีเท่ากับ วิธีบูตสเตรป
 4. สำหรับการประมาณความโค้ง พบว่า วิธีบูตสเตรปมีประสิทธิภาพดีกว่า วิธีแจ็คไนฟ์ ทุกค่าพารามิเตอร์
- ดังนั้น การเปรียบเทียบประสิทธิภาพวิธีการประมาณค่าพารามิเตอร์แบบจุด ระหว่างวิธีแจ็คไนฟ์ กับ วิธีบูตสเตรป สามารถสรุปได้ว่า โดยส่วนใหญ่ประมาณ 71.43% วิธีบูตสเตรป จะมีประสิทธิภาพดีกว่า วิธีแจ็คไนฟ์

กรณีที่ 2 การเปรียบเทียบประสิทธิภาพวิธีการประมาณค่าพารามิเตอร์แบบช่วง ระหว่างวิธีแจ็คไนฟ์ กับ วิธีบูตสเตรป

โดยพิจารณาจากการเปรียบเทียบค่าสัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient) ของทั้ง 2 วิธี โดยถ้าวิธีการประมาณค่าแบบใดที่ทำให้ค่าสัมประสิทธิ์ความเชื่อมั่น สูงกว่าจะถือว่าการประมาณค่าจากวิธีนั้นเป็นวิธีที่มีประสิทธิภาพมากกว่าในการประมาณค่าแบบช่วง โดยจะประมาณพารามิเตอร์ค่าเฉลี่ย ค่าความแปรปรวน ค่าความเบ้ และค่าความโค้งของข้อมูลที่มีการแจกแจงแบบปกติปลอมปน การแจกแจงชี้กำลัง และการแจกแจงแกมมา ซึ่งผลการวิจัยสรุปได้ว่า

การประมาณพารามิเตอร์แบบช่วงด้วยวิธีบูตสเตรปจะมีประสิทธิภาพดีกว่า วิธีแจ็คไนฟ์ ทุกขนาดตัวอย่าง ทุกค่าพารามิเตอร์ในทุกๆการแจกแจง

ข้อสังเกต :

เนื่องจากค่าเฉลี่ยความแปรปรวนของตัวประมาณด้วยวิธีบูตสเตรป จะมีค่ามากกว่าค่าเฉลี่ยความแปรปรวนของตัวประมาณด้วยวิธีแจ็คไนฟ์ ทุกกรณี ทำให้การประมาณช่วงความเชื่อมั่นด้วยวิธีบูตสเตรป จะกว้างกว่า การประมาณช่วงความเชื่อมั่นด้วยวิธีแจ็คไนฟ์ ทุกกรณี ดังนั้น การประมาณค่าแบบช่วงของพารามิเตอร์ด้วยวิธีบูตสเตรป จะสามารถครอบคลุมค่าพารามิเตอร์ ได้มากกว่า การประมาณค่าแบบช่วงของพารามิเตอร์ด้วยวิธีแจ็คไนฟ์ ทุกกรณี

เอกสารอ้างอิง

- [1] ชีระพร วีระถาวร.(2539). **ความน่าจะเป็น กับ การประยุกต์**. กรุงเทพฯ : วิทยพัฒน์
- [2] Quenouille, M. H. (1956). Notes on Bias in Estimation. **Biometrika** 43, 353-360.
- [3] Miller, R. G. (1974). The Jackknife - a Review. **Biometrika** 61, 1-15.
- [4] Efron, B.(1979). Bootstrapping methods : Another Look at the Jackknife. **The Annals of Statistics** 7, 1-26.
- [5] Efron, B.(1982). **The Jackknife, the Bootstrap and Other Resampling Plans**. Philadelphia : SIAM