ช่วงความเชื่อมั่นสำหรับผลต่างของค่ามัธยฐานของการแจกแจงล็อกปรกติ

Confidence Intervals for Difference of Median of the Lognormal Distribution

อุษณีย์ จันทะสุวรรณ์ (Usanee Janthasuwan)* ดร.วราฤทธิ์ พานิชกิจโกศลกุล (Dr.Wararit Panichkitkosolkul)**

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบช่วงความเชื่อมั่นสำหรับผลต่างของค่ามัธยฐานของการแจกแจงล็อกปรกติ โดยใช้วิธี Method of Variance Estimates Recovery (MOVER) ซึ่งช่วงความเชื่อมั่นที่ศึกษามี 3 วิธี ได้แก่ ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior การวิจัยครั้งนี้ใช้วิธีการจำลองด้วยเทคนิคมอนติคาร์โลโดยในการเปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่นดังกล่าว จะพิจารณาจากค่าความน่าจะเป็นคุ้มรวมและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น ผลการศึกษาสรุปได้ดังนี้ ช่วงความเชื่อมั่นแบบเบส์มีประสิทธิภาพสูงกว่าช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด กล่าวคือ เมื่อพิจารณาค่าความน่าจะเป็นคุ้มรวม ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุดและช่วงความเชื่อมั่นแบบเบส์เข้าใกล้ระดับความเชื่อมั่นที่กำหนด และค่าความน่าจะเป็นคุ้มรวมไม่แตกต่างกันมาก เมื่อพิจารณาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น ช่วงความเชื่อมั่นแบบเบส์มีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่สั้นกว่าช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด

ABSTRACT

The purpose of this research is to study and compare the confidence intervals for difference of median of the lognormal distribution based on the Method of Variance Estimates Recovery (MOVER). They are three methods of the confidence intervals studied: the confidence interval for maximum likelihood estimator, Bayes credible interval with right invariant prior and Bayes credible interval with left invariant prior. This research uses the Monte Carlo simulation technique by comparing the efficiency of the confidence interval. By considering the coverage probability and expected length for the differences of confidence intervals. The results of this study showed that Bayes credible interval was more effective than confidence interval for maximum likelihood estimator. Considering the coverage probability, the confidence interval for maximum likelihood estimator and Bayes credible interval were close to the determined confidence level and not too different. Considering the expected length confidence interval, Bayes credible interval had shorter expected length confidence interval than the confidence interval for maximum likelihood estimator.

คำสำคัญ: ช่วงความเชื่อมั่น ค่ามัธยฐาน การแจกแจงล็อกปรกติ

Keywords: Confidence interval, Median, Lognormal Distribution

*นักศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
**รองศาสตราจารย์ สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

บทนำ

ในการเชิงอนุมานแบ่งออกเป็น 2 ลักษณะคือ การทดสอบสมมติฐานและการประมาณค่า ซึ่งการประมาณค่าพารามิเตอร์นั้นแบ่งออกเป็น 2 แบบ คือ การประมาณค่าแบบจุดและการประมาณค่าแบบช่วง โดยที่การประมาณค่าแบบช่วงเป็นการประมาณค่าพารามิเตอร์ของประชากรโดยอาศัยตัวประมาณแบบจุดของพารามิเตอร์และการแจกแจงของตัวประมาณนั้น ซึ่งช่วงความเชื่อมั่นที่หามาได้นั้นจะต้องครอบคลุมค่าพารามิเตอร์ที่สนใจศึกษา และจะมีค่าอยู่ในช่วงของขอบเขตล่าง (Lower Bound: L) และขอบเขตบน (Upper Bond: U) โดยสามารถเขียนรูปแบบของช่วงความเชื่อมั่น ได้ดังนี้ $L<\theta<U$ เมื่อ θ แทนค่าพารามิเตอร์ให้อยู่ภายในช่วงนี้ด้วยความเชื่อมั่นระดับหนึ่ง (สายชล, 2555)

การแจกแจงล็อกปรกติ (Lognormal Distribution) เป็นการแจกแจงหนึ่งที่ข้อมูลประชากรที่สนใจไม่ได้มีการแจกแจงแบบปรกติ การแจกแจงแบบล็อกปรกติเป็นการแจกแจงความน่าจะเป็นแบบต่อเนื่องของตัวแปรสุ่ม ซึ่งนำไปประยุกต์ใช้กับงานด้านต่าง ๆ อย่างกว้างขวาง เช่น งานด้านการแพทย์นำไปใช้อธิบายเกี่ยวกับระยะเวลาในการแพร่กระจายของเชื้อโรคในผู้ป่วย งานด้านบริหารธุรกิจและเศรษฐศาสตร์นำไปใช้อธิบายเกี่ยวกับระยะเวลาของการจ้างงาน งานด้านสังคมศาสตร์นำไปใช้อธิบายเกี่ยวกับการวิเคราะห์ข้อมูลตลาดหุ้น งานทางด้านวิศวกรรมศาสตร์นำไปใช้อธิบายเกี่ยวกับระยะเวลาซ่อมของระบบไฟฟ้าจากเหตุไฟฟ้าขัดข้อง เป็นต้น และการแจกแจงล็อกปรกติเป็นการแจกแจงแบบล็อก – โลเคชั่น – สเกล (Log - Location - Scale) ซึ่งมีฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function; PDF) คือ

$$f(x;\mu,\sigma^2)=\frac{1}{x\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\left(\frac{\log(x)-\mu}{\sigma}\right)^2}$$

โดยที่ $x>0$, $-\infty<\mu<\infty$ และ $\sigma>0$ เมื่อ μ เป็นพารามิเตอร์แสดงตำแหน่ง (Location Parameter)) หรือพารามิเตอร์ค่าเฉลี่ยและ σ เป็นพารามิเตอร์กำหนดรูปร่าง (Shape Parameter) สำหรับค่ามัธยฐานหาได้จาก $\exp(\mu)$

ในการวิเคราะห์ข้อมูลนั้น เมื่อเราเก็บรวบรวมข้อมูลได้แล้ว วิธีการหนึ่งที่เราจะอธิบายลักษณะของข้อมูลที่เก็บรวบรวมมาได้ได้คือการอธิบายค่ากลางของข้อมูลซึ่งเป็นตัวแทนของข้อมูลทั้งหมด หรือที่เรียกว่าการวัดแนวโน้มเข้าสู่ส่วนกลาง โดยมีวิธีการหาได้หลายวิธี แต่ที่นิยมใช้กันทั่ว ๆ ไปคือการหาค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม แต่ละวิธีมีทั้งของดีและข้อเสีย และมีความเหมาะสมไม่เหมือนกัน ขึ้นอยู่กับลักษณะของข้อมูลที่รวบรวมมา ซึ่งถ้าข้อมูลที่ลักษณะสมมาตร การวัดแนวโน้นเข้าสู่ส่วนกลางสามารถใช้ค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม อย่างใดอย่างหนึ่งก็ได้ เนื่องจากค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยมมีค่าเท่ากัน แต่ถ้าข้อมูลมีลักษณะไม่สมมาตรหรือมีการเบ้ไปทางใดทางหนึ่ง ข้อมูลลักษณะนี้ควรใช้ค่ามัธยฐาน และค่ามัธยฐานสามารถคำนวณได้ง่าย เข้าใจง่าย อีกทั้งยังไม่ถูกกระทบกระเทือนเมื่อข้อมูลมีค่าผิดปรกติ

จากการศึกษางานวิจัยของ Zellner (1971) ได้เสนอตัวประมาณค่าของค่าเฉลี่ยและค่ามัธยฐานของการแจกแจงล็อกปรกติหลายตัวประมาณ และมีการพิจารณาถึงปัญหาของการอนุมานเกี่ยวกับพารามิเตอร์ของการแจกแจงล็อกปรกติ อีกทั้งยังมีการเปรียบเทียบค่าประมาณของเบส์ ต่อมา Padgett and Wei (1977) ได้มีการพัฒนาตัวประมาณของเบส์ของฟังก์ชันการแจกแจงล็อกปรกติ โดยมีการแจกแจงก่อน 2 การแจกแจง คือ normal prior สำหรับค่าเฉลี่ย และ gamma prior สำหรับการผกผันของพารามิเตอร์กำหนดรูปร่าง จากนั้น D’ Cunha and Rao (2014) ได้ศึกษาตัวประมาณแบบเบส์และช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยของการแจกแจงล็อกปรกติ ซึ่งมีการแจกแจงก่อนคือ Right invariant prior และ Left invariant prior ต่อมา D’ Cunha and Rao (2016) ได้เปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่นแบบเบส์

กับช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าจะเป็นสูงสุด สำหรับค่ามัธยฐานของการแจกแจงล็อกปรกติ โดยมีการแจกแจงก่อนคือ Uniform prior, Right invariant prior , Left invariant prior และ Jeffreys rule prior

เมื่อเราสนใจเกี่ยวกับประชากร 2 กลุ่ม ซึ่งได้ศึกษางานวิจัยของ Hettmansperger (1984) ได้เสนอวิธีการสร้างช่วงความเชื่อมั่นสำหรับผลต่างของค่ามัธยฐาน โดยที่สองประชากรมีความคล้ายกันและมีการแจกแจงเดียวกัน ต่อมา ปวริศร์ (2547) ได้ศึกษาช่วงความเชื่อมั่นสำหรับผลต่างของมัธยฐานของตัวอย่างโดยไม่จำกัดการแจกแจง โดยทำการเปรียบเทียบการประมาณค่าแบบช่วงสำหรับประชากร 2 กลุ่ม

Zou et al. (2009) เสนอวิธีการที่เรียกว่า " Method of Variance Estimates Recovery " (MOVER) ซึ่งวิธี MOVER ได้รับการออกแบบเพื่อนำไปประยุกต์ใช้กับสถานการณ์ทั่วไปและเพื่อให้มีอัตราความครอบคลุมที่เพียงพอในขั้นตอนการประมาณที่เกี่ยวข้องกับข้อมูลที่มีการแจกแจงล็อกปรกติ ข้อได้เปรียบของวิธี MOVER คือง่ายต่อการใช้งาน โดยมีความรู้เบื้องต้นเกี่ยวกับพื้นฐานทางสถิติเพียงเล็กน้อยเท่านั้น ต่อมา Harvey et al. (2012) ได้เปรียบเทียบประสิทธิภาพของการประมาณช่วงความเชื่อมั่นแบบเบส์สำหรับค่าเฉลี่ยของการแจกแจงล็อกปรกติ โดยวิธี MOVER ซึ่งมีการแจกแจงก่อนคือ Jeffreys prior และ Jeffreys rule prior จากนั้น Krishnamoorthy and Hasacn (2017) ได้ศึกษาปัญหาการประมาณอัตราส่วนของค่าสัมประสิทธิ์การแปรผันของประชากรสองกลุ่มที่เป็นอิสระของการแจกแจงล็อกปรกติ โดยใช้วิธี MOVER

จากงานวิจัยที่ผ่านมายังไม่มีนักวิจัยท่านใดศึกษาช่วงความเชื่อมั่นสำหรับผลต่างของค่ามัธยฐานของการแจกแจงล็อกปรกติ โดยใช้วิธี MOVER ดังนั้นในงานวิจัยนี้ ผู้วิจัยจึงมุ่งเน้นศึกษาและเปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด (Maximum Likelihood Estimator: MLE) กับช่วงความเชื่อมั่นแบบเบส์ (Bayes credible interval) โดยใช้วิธี MOVER ซึ่งในการประมาณค่าแบบเบส์กำหนดการแจกแจงก่อน 2 การแจกแจง คือ Right invariant prior และ Left invariant prior สำหรับประชากรสองกลุ่มให้ค่าความน่าจะเป็นคุ้มรวมเข้าใกล้ระดับความเชื่อมั่นและให้มีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่สั้น

วัตถุประสงค์การวิจัย

1. เพื่อเสนอช่วงความเชื่อมั่นสำหรับผลต่างของค่ามัธยฐานของการแจกแจงล็อกปรกติ 3 วิธี ได้แก่ ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และ Left invariant prior โดยใช้วิธี MOVER

2. เพื่อเปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่นโดยพิจารณาค่าความน่าจะเป็นคุ้มรวม (Confidence Probability) และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น (Expected Length) สำหรับผลต่างของค่ามัธยฐานของประชากรที่มีการแจกแจงล็อกปรกติ ซึ่งมีช่วงความเชื่อมั่นทั้งหมด 3 วิธี ได้แก่

   2.1 ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด

   2.2 ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

   2.3 ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

วิธีการวิจัย

งานวิจัยนี้สนใจศึกษาช่วงความเชื่อมั่นสำหรับผลต่างของค่ามัธยฐานของการแจกแจงล็อกปรกติ 3 วิธี ได้แก่ ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior โดยใช้วิธี MOVER โดยมีขั้นตอนการศึกษาดังนี้

1. สร้างตัวแปรสุ่มจากการแจกแจงล็อกปรกติ โดยจำลองข้อมูลในโปรแกรม R จากนั้นกำหนดขนาดตัวอย่างและค่าพารามิเตอร์

2. นำตัวแปรสุ่มที่ได้ มาคำนวณค่าเฉลี่ยและค่าความแปรปรวนประชากร

3. หาช่วงความเชื่อมั่นของค่ามัธยฐานสำหรับประชากรที่มีการแจกแจงล็อกปรกติ ทั้งช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุดและช่วงความเชื่อมั่นแบบเบส์

3.1 ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด กำหนดให้ $X_1,\dots,X_n$ เป็นตัวอย่างสุ่มขนาด n ที่สุ่มมาจากประชากรที่มีการแจกแจงปรกติ โดยมีพารามิเตอร์ ค่าเฉลี่ย μ และความแปรปรวน σ^2 กำหนดให้ $Z_i=\log X_i$, $i=1,2,\dots,n$ เมื่อ Z มีการแจกแจงล็อกปรกติที่มีพารามิเตอร์ μ และ σ^2 ตัวประมาณภาวะน่าจะเป็นสูงสุดแทน μ และ σ^2 ด้วย $\bar{Z}$ และ S_z^2 ตามลำดับ โดยที่ $\bar{Z}=\sum_{i=1}^{n}Z_i\Big/n$ และ $S_Z^2=\sum_{i=1}^{n}\left(Z_i-\bar{Z}\right)^2\Big/n$ สำหรับการประมาณภาวะน่าจะเป็นสูงสุดของค่ามัธยฐานของการแจกแจงล็อกปรกติแทน $\theta=e^{\mu}$ ด้วย $\hat{\theta}=e^{\bar{z}}$ ความแปรปรวนเชิงเส้นกำกับของ $\hat{\theta}$ สามารถคำนวณได้จากวิธีเดลตา (Delta Method) จะได้เป็น

$$\widehat{Var}\left(\hat{\theta}\right)=e^{2\bar{z}}\frac{S_z^2}{n}+\frac{S_z^2}{n} \tag{1}$$

จากทฤษฎีบทขีดจำกัดส่วนกลางจะได้ว่า $Z=\dfrac{\hat{\theta}-\theta}{\sqrt{\widehat{Var}\left(\hat{\theta}\right)}}\sim N\left(0,1\right)$ ดังนั้นช่วงความเชื่อมั่นเชิงเส้นกำกับ $(1-\alpha)100\%$ ของพารามิเตอร์ θ ของตัวประมาณแบบภาวะน่าจะเป็นสูงสุด คือ

$$\left(\hat{\theta}-Z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}\right)}\;,\;\hat{\theta}+Z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}\right)}\right) \tag{2}$$

โดยที่ $z_{\alpha/2}$ คือ เปอร์เซ็นไทล์ที่ $(\alpha/2)\times 100$ ของการแจกแจงปรกติมาตรฐาน

α คือ ระดับนัยสำคัญที่กำหนด ในที่นี้กำหนดให้ $\alpha=0.05$

3.2 ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior: $\pi\left(\mu,\sigma^2\right)=\dfrac{1}{\sigma}$ จะได้การแจกแจงหลัง (Posterior distribution) คือ

$$\pi\left(\mu\,,\,\sigma^2\,|\,z_1,\dots,z_n\right)=\frac{1}{\sqrt{2\pi}}\frac{\sqrt{n}}{\sigma}e^{-\frac{1}{2}\frac{\left(\bar{z}-\mu\right)}{\sigma^2/n}}\frac{\left(\dfrac{\left(n-1\right)S_z^2}{2}\right)^{\frac{n+2}{2}}}{\Gamma\left(\dfrac{n+2}{2}\right)}\left(\frac{1}{\sigma^2}\right)^{\left(\frac{n+2}{2}\right)-1}e^{-\frac{1}{2}\left(\frac{n-1}{2}\right)S_z^2}$$

เนื่องจากฟังก์ชันการแจกแจงไม่มีรูปแบบปิด (Closed-form solution) ในการจำลองมอนติคาร์โล (Monte Carlo Simulation) จึงดำเนินการโดยใช้วิธี Importance sampling จะได้เป็น $E\left(e^{\mu} \mid z_1,\dots,z_n\right)=\frac{1}{M}\sum_{i=1}^{M}e^{\mu_i}$ โดยที่ M เป็นจำนวนรอบของการทำซ้ำ และสำหรับช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior จะได้เป็น

$$\left(R_{\alpha/2},\, R_{1-\alpha/2}\right) \tag{3}$$

โดยที่ $R_{\alpha/2}$ คือ เปอร์เซ็นไทล์ที่ $(\alpha/2)\times 100$ ของกรแจกแจงหลังของ e^{μ} เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และ $R_{1-\alpha/2}$ คือ เปอร์เซ็นไทล์ที่ $(1-\alpha/2)\times 100$ ของกรแจกแจงหลังของ e^{μ} เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

3.3 ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior: $\pi\left(\mu,\sigma^2\right)=\frac{1}{\sigma^2}$ จะได้การแจกแจงหลัง (Posterior distribution) คือ

$$\pi\left(\mu\,,\,\sigma^2 \mid z_1,\dots,z_n\right)=\frac{1}{\sqrt{2\pi}}\frac{\sqrt{n}}{\sigma}e^{-\frac{1}{2}\frac{(\bar{z}-\mu)}{\sigma^2/n}}\frac{\left(\frac{(n-1)S_z^2}{2}\right)^{\frac{n+3}{2}}}{\Gamma\left(\frac{n+3}{2}\right)}\left(\frac{1}{\sigma^2}\right)^{\left(\frac{n+3}{2}\right)-1}e^{-\frac{1}{2}\left(\frac{n-1}{2}\right)S_z^2}$$

เนื่องจากฟังก์ชันการแจกแจงไม่มีรูปแบบปิด (Closed-form solution) ในการจำลองมอนติคาร์โล (Monte Carlo Simulation) จึงดำเนินการโดยใช้วิธี Importance sampling จะได้เป็น $E\left(e^{\mu} \mid z_1,\dots,z_n\right)=\frac{1}{M}\sum_{i=1}^{M}e^{\mu_i}$ โดยที่ M เป็นจำนวนรอบของการทำซ้ำ และสำหรับช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior จะได้เป็น

$$\left(L_{\alpha/2},\, L_{1-\alpha/2}\right) \tag{4}$$

โดยที่ $L_{\alpha/2}$ คือเปอร์เซ็นไทล์ที่ $(\alpha/2)\times 100$ ของกรแจกแจงหลังของ e^{μ} เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior $L_{1-\alpha/2}$ คือเปอร์เซ็นไทล์ที่ $(1-\alpha/2)\times 100$ ของกรแจกแจงหลังของ e^{μ} เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

4. คำนวณหาช่วงความเชื่อมั่นของผลต่างของค่ามัธยฐานสำหรับประชากรที่มีการแจกแจงล็อกปรกติ ระหว่างช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุดกับช่วงความเชื่อมั่นแบบเบส์ โดยใช้วิธี MOVER

ในเบื้องต้นจะอธิบายหลักการของวิธี MOVER เพื่อนำไปหาช่วงความเชื่อมั่นของผลต่างของมัธยฐานต่อไป

Donner and Zou (2008) ได้สร้างช่วงความเชื่อมั่นสำหรับฟังก์ชันของค่าเฉลี่ยประชากรปรกติ และจากทฤษฎีขีดจำกัดส่วนกลาง (The Central Limit Theorem: CLT) ทำให้ได้ช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ $\theta_1+\theta_2$ คือ

$$\left(\hat{\theta}_1+\hat{\theta}_2\right)\pm z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_1\right)+\widehat{Var}\left(\hat{\theta}_2\right)}$$

โดยที่ $z_{\alpha/2}$ คือ เปอร์เซ็นไทล์ที่ $100\times(\alpha/2)^{th}$ ของการแจกแจงปรกติมาตรฐาน

$\hat{\theta}_1$, $\hat{\theta}_2$ คือ ตัวประมาณที่ไม่เอนเอียงของพารามิเตอร์ θ_1 , θ_2

$\widehat{Var}\left(\hat{\theta}_1\right)$, $\widehat{Var}\left(\hat{\theta}_2\right)$ คือ ค่าความแปรปรวนของ $\hat{\theta}_1$, $\hat{\theta}_2$ ที่มี $\widehat{Var}\left(\hat{\theta}_1\right)$ และ $\widehat{Var}\left(\hat{\theta}_2\right)$ เป็นอิสระต่อกัน

กำหนดให้ (l_i, u_i) เป็นช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ θ_i และให้ $l_1 + l_2$ มีค่าใกล้เคียงกับ L มากที่สุดและ $u_1 + u_2$ มีค่าใกล้เคียงกับ U มากที่สุด จะได้ช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ $\theta_1 + \theta_2$ คือ (L, U) โดยที่

$$L = \left(\hat{\theta}_1 + \hat{\theta}_2\right) - z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_1\right) + \widehat{Var}\left(\hat{\theta}_2\right)}$$

$$= \left(\hat{\theta}_1 + \hat{\theta}_2\right) - z_{\alpha/2}\sqrt{\frac{\left(\hat{\theta}_1 - l_1\right)^2}{z_{\alpha/2}^2} + \frac{\left(\hat{\theta}_2 - l_2\right)^2}{z_{\alpha/2}^2}}$$

$$U = \left(\hat{\theta}_1 + \hat{\theta}_2\right) + z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_1\right) + \widehat{Var}\left(\hat{\theta}_2\right)}$$

$$= \left(\hat{\theta}_1 + \hat{\theta}_2\right) + z_{\alpha/2}\sqrt{\frac{\left(u_1 - \hat{\theta}_1\right)^2}{z_{\alpha/2}^2} + \frac{\left(u_2 - \hat{\theta}_2\right)^2}{z_{\alpha/2}^2}}$$

ดังนั้นช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับ $\theta_1 + \theta_2$ คือ

$$(L, U) = \left(\left(\hat{\theta}_1 + \hat{\theta}_2\right) - \sqrt{\left(\hat{\theta}_1 - l_1\right)^2 + \left(\hat{\theta}_2 - l_2\right)^2}\ ,\ \left(\hat{\theta}_1 + \hat{\theta}_2\right) + \sqrt{\left(u_1 - \hat{\theta}_1\right)^2 + \left(u_2 - \hat{\theta}_2\right)^2}\right) \quad (5)$$

ในการหาช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับผลต่างของ $\theta_1 - \theta_2$ คือ

$$(L, U) = \left(\left(\hat{\theta}_1 - \hat{\theta}_2\right) - \sqrt{\left(\hat{\theta}_1 - l_1\right)^2 + \left(u_2 - \hat{\theta}_2\right)^2}\ ,\ \left(\hat{\theta}_1 - \hat{\theta}_2\right) + \sqrt{\left(u_1 - \hat{\theta}_1\right)^2 + \left(\hat{\theta}_2 - l_2\right)^2}\right) \quad (6)$$

หลังจากทราบหลักการของวิธี MOVER แล้วจะนำหลักการดังกล่าวหาช่วงความเชื่อมั่นของผลต่างของมัธยฐาน ซึ่งจะอาศัยช่วงความเชื่อมั่นของมัธยฐานกลุ่มเดียว

4.1 จากช่วงความเชื่อมั่นสำหรับค่ามัธยฐานในสมการ (2) กำหนดให้ $\theta_1 = e^{\mu_1}$, $\theta_2 = e^{\mu_2}$ จะได้ว่า

$$\left(l_1^*, u_1^*\right) = \left(\hat{\theta}_1 - Z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_1\right)}\ ,\ \hat{\theta}_1 + Z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_1\right)}\right)$$

$$\left(l_2^*, u_2^*\right) = \left(\hat{\theta}_2 - Z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_2\right)}\ ,\ \hat{\theta}_2 + Z_{\alpha/2}\sqrt{\widehat{Var}\left(\hat{\theta}_2\right)}\right)$$

โดยที่ $\widehat{Var}\left(\hat{\theta}_1\right) = e^{2\bar{z}_1}\frac{S_{z_1}^2}{n_1} + \frac{S_{z_1}^2}{n_1}$ และ $\widehat{Var}\left(\hat{\theta}_2\right) = e^{2\bar{z}_2}\frac{S_{z_2}^2}{n_2} + \frac{S_{z_2}^2}{n_2}$

กำหนดให้ $\hat{\theta}_1 = e^{\bar{z}_1}$, $\hat{\theta}_2 = e^{\bar{z}_2}$, $\alpha = 0.05$, $l_1 = l_1^*$, $l_2 = l_2^*$, $u_1 = u_1^*$ และ $u_2 = u_2^*$ จากนั้นแทนค่าทั้งหมดลงในสมการ (6) ดังนั้น ช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับผลต่างของ $\theta_1 - \theta_2$ คือ

$$CI_{MLE} = \left(\left(e^{\bar{z}_1} - e^{\bar{z}_2}\right) - \sqrt{\left(e^{\bar{z}_1} - l_1^*\right)^2 + \left(u_2^* - e^{\bar{z}_2}\right)^2}\ ,\ \left(e^{\bar{z}_1} - e^{\bar{z}_2}\right) + \sqrt{\left(u_1^* - e^{\bar{z}_1}\right)^2 + \left(e^{\bar{z}_2} - l_2^*\right)^2}\right) \quad (7)$$

4.2 จากช่วงความเชื่อมั่นสำหรับค่ามัธยฐานในสมการ (3) กำหนดให้ $\theta_1 = e^{\mu_1}$, $\theta_2 = e^{\mu_2}$ จะได้ว่า

$$\left(l_{R1}^*, u_{R1}^*\right) = \left(R_{(\alpha/2)1},\ R_{(1-\alpha/2)1}\right)$$

$$\left(l_{R2}^*, u_{R2}^*\right) = \left(R_{(\alpha/2)2}, R_{(1-\alpha/2)2}\right)$$

กำหนดให้ $\hat{\theta}_1 = e^{\bar{z}_1}$, $\hat{\theta}_2 = e^{\bar{z}_2}$, $\alpha = 0.05$, $l_1 = l_{R1}^*$, $l_2 = l_{R2}^*$, $u_1 = u_{R1}^*$ และ $u_2 = u_{R2}^*$ จากนั้นแทนค่าทั้งหมดลงในสมการ (2.6) ดังนั้น ช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับผลต่างของ $\theta_1 - \theta_2$ คือ

$$CI_{Right} = \left(\left(e^{\bar{z}_1} - e^{\bar{z}_2}\right) - \sqrt{\left(e^{\bar{z}_1} - l_{R1}^*\right)^2 + \left(u_{R2}^* - e^{\bar{z}_2}\right)^2} \ , \left(e^{\bar{z}_1} - e^{\bar{z}_2}\right) + \sqrt{\left(u_{R1}^* - e^{\bar{z}_1}\right)^2 + \left(e^{\bar{z}_2} - l_{R2}^*\right)^2} \right) \quad (8)$$

4.3 จากช่วงความเชื่อมั่นสำหรับค่ามัธยฐานในสมการ (4) กำหนดให้ $\theta_1 = e^{\mu_1}$, $\theta_2 = e^{\mu_2}$ จะได้ว่า

$$\left(l_{L1}^*, u_{L1}^*\right) = \left(L_{(\alpha/2)1}, L_{(1-\alpha/2)1}\right)$$

$$\left(l_{L2}^*, u_{L2}^*\right) = \left(L_{(\alpha/2)2}, L_{(1-\alpha/2)2}\right)$$

กำหนดให้ $\hat{\theta}_1 = e^{\bar{z}_1}$, $\hat{\theta}_2 = e^{\bar{z}_2}$, $\alpha = 0.05$, $l_1 = l_{L1}^*$, $l_2 = l_{L2}^*$, $u_1 = u_{L1}^*$ และ $u_2 = u_{L2}^*$ จากนั้นแทนค่าทั้งหมดลงในสมการ (6) ดังนั้น ช่วงความเชื่อมั่น $100(1-\alpha)\%$ สำหรับผลต่างของ $\theta_1 - \theta_2$ คือ

$$CI_{Left} = \left(\left(e^{\bar{z}_1} - e^{\bar{z}_2}\right) - \sqrt{\left(e^{\bar{z}_1} - l_{L1}^*\right)^2 + \left(u_{L2}^* - e^{\bar{z}_2}\right)^2} \ , \left(e^{\bar{z}_1} - e^{\bar{z}_2}\right) + \sqrt{\left(u_{L1}^* - e^{\bar{z}_1}\right)^2 + \left(e^{\bar{z}_2} - l_{L2}^*\right)^2} \right) \quad (9)$$

5. ตรวจสอบว่าช่วงความเชื่อมั่นสำหรับผลต่างของมัธยฐานของการแจกแจงล็อกปรกติที่คำนวณได้ครอบคลุมผลต่างของมัธยฐานประชากรของการแจกแจงล็อกปรกติที่กำหนดไว้หรือไม่

6. คำนวณค่าความน่าจะเป็นคุ้มรวมของช่วงความเชื่อมั่นที่คำนวณได้ ซึ่งคำนวณได้จากสูตรคือจำนวนครั้งที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ ในงานวิจัยนี้มีการทำซ้ำทั้งหมด 10,000 ครั้ง

$$CP = \frac{\sum_{i=1}^{10,000} I_i\left(e^{\mu_1} - e^{\mu_2}\right)}{10,000}$$

เมื่อ $\sum_{i=1}^{10,000} I_i\left(e^{\mu_1} - e^{\mu_2}\right)$ คือจำนวนครั้งที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์ $e^{\mu_1} - e^{\mu_2}$

$I_i\left(e^{\mu_1} - e^{\mu_2}\right)$ มีค่าเท่ากับ 1 เมื่อ $e^{\mu_1} - e^{\mu_2}$ ตกอยู่ภายในช่วงความเชื่อมั่น (L_i, U_i) ที่ประมาณขึ้นในครั้งที่ i และ $I_i\left(e^{\mu_1} - e^{\mu_2}\right)$ มีค่าเท่ากับ 0 เมื่อ $e^{\mu_1} - e^{\mu_2}$ ตกอยู่นอกช่วงความเชื่อมั่น (L_i, U_i) ที่ประมาณขึ้นในครั้งที่ i

7. คำนวณค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่คำนวณได้ ซึ่งคำนวณได้จากสูตร

$$EL = \frac{\sum_{i=1}^{10,000} Len_i\left(e^{\mu_1} - e^{\mu_2}\right)}{10,000}$$

โดยที่ $Len_i\left(e^{\mu_1} - e^{\mu_2}\right) = U_i - L_i$ คือ ความยาวของช่วงความเชื่อมั่น เมื่อ L_i คือค่าขอบเขตล่างของช่วงความเชื่อมั่น และ U_i คือค่าขอบเขตบนของช่วงความเชื่อมั่นของพารามิเตอร์

8. เปรียบเทียบค่าความน่าจะเป็นคุ้มรวม และค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นของช่วงความเชื่อมั่นที่คำนวณได้ทั้งช่วงความเชื่อมั่นแบบเบส์และช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ถ้าช่วงความเชื่อมั่นใดให้

ค่าความน่าจะเป็นคุ้มรวมเข้าใกล้ระดับความเชื่อมั่น และค่าประมาณความยาวเฉลี่ยของช่วงความเชื่อมั่นที่สั้นที่สุด แสดงว่าช่วงความเชื่อมั่นนั้นมีประสิทธิภาพสูงที่สุด

ผลการวิจัย

ในการศึกษาความน่าจะเป็นคุ้มรวมและความยาวเฉลี่ยของช่วงความเชื่อมั่น โดยทำการจำลองมอนติคาร์โล ซึ่งในการศึกษาและเปรียบเทียบช่วงความเชื่อมั่นสำหรับผลต่างของมัธยฐานของการแจกแจงล็อกปรกติที่ใช้ในงานวิจัยนี้ ได้จำลองข้อมูลมาจากโปรแกรม R โดยมีการกำหนดค่าต่าง ๆ ไว้ดังนี้ กำหนดให้ข้อมูลมีการแจกแจงปรกติที่มีค่าเฉลี่ยประชากรสองกลุ่มเป็น log(1000) กำหนดขนาดตัวอย่างของประชากรสองกลุ่มเป็น n_1 =10, 20, 40, 60, 80, 100, 150, 200 และ n_2 =10, 20, 40, 60, 80, 100, 150, 200 ระดับความเชื่อมั่นที่กำหนดในงานวิจัยนี้ คือ 95% กำหนดค่าสัมประสิทธิ์การแปรผันของประชากรทั้งสองกลุ่มเป็น CV_1 = 0.1, 0.3, 0.5, 0.7, 1.5, 2, 2.5 และ CV_2 = 0.1, 0.3, 0.5, 0.7, 1.5, 2, 2.5 จากการแจกแจงล็อกปรกติ ค่าสัมประสิทธิ์การแปรผัน เท่ากับ $CV = \left(\exp\left(\sigma^2\right)-1\right)^{1/2}$ นั่นคือ $\sigma^2 = \ln\left(CV^2+1\right)$ ทำการจำลองทั้งหมด 10,000 ครั้ง ในแต่ละสถานการณ์ที่กำหนด

ตารางที่ 1 ความน่าจะเป็นคุ้มรวมและความยามเฉลี่ยของช่วงความเชื่อมั่นในแต่ละค่าสัมประสิทธิ์การแปรผัน เมื่อขนาดตัวอย่างของประชากร $n_1 = 40$ และ $n_2 = 40$

ช่วงความเชื่อมั่น		ความน่าจะเป็นคุ้มรวม(ความยาวเฉลี่ยของช่วง) เมื่อ CV เท่ากับ							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
MLE		0.944	0.948	0.948	0.952	0.950	0.958	0.955	0.953
		(87.14)	(256.51)	(415.04)	(555.55)	(737.30)	(970.21)	(1142.53)	(1276.31)
Bayes	Right	0.942	0.945	0.949	0.956	0.951	0.950	0.951	0.953
		(86.68)	(255.51)	(414.69)	(557.23)	(744.77)	(990.12)	(1177.00)	(1324.91)
	Left	0.943	0.946	0.950	0.951	0.950	0.951	0.954	0.955
		(85.06)	(250.74)	(406.86)	(546.67)	(730.18)	(971.00)	(1153.25)	(1298.47)

หมายเหตุ : Right = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

Left = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

ตารางที่ 2 ความน่าจะเป็นคุ้มรวมและความยามเฉลี่ยของช่วงความเชื่อมั่นในแต่ละค่าสัมประสิทธิ์การแปรผัน เมื่อขนาดตัวอย่างของประชากร $n_1 = 40$ และ $n_2 = 80$

ช่วงความเชื่อมั่น		ความน่าจะเป็นคุ้มรวม(ความยาวเฉลี่ยของช่วง) เมื่อ CV เท่ากับ							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
MLE		0.946	0.949	0.948	0.952	0.952	0.953	0.952	0.958
		(75.35)	(222.11)	(360.18)	(484.39)	(644.43)	(851.70)	(1011.26)	(1134.57)
Bayes	Right	0.944	0.945	0.945	0.947	0.950	0.952	0.950	0.955
		(74.71)	(220.24)	(357.17)	(480.22)	(638.75)	(844.27)	(1002.17)	(1123.65)
	Left	0.940	0.945	0.946	0.949	0.950	0.951	0.952	0.953
		(73.80)	(217.47)	(352.53)	(474.22)	(630.52)	(833.58)	(989.31)	(1109.18)

หมายเหตุ : Right = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

Left = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

ตารางที่ 3 ความน่าจะเป็นคุ้มรวมและความยาวเฉลี่ยของช่วงความเชื่อมั่นในแต่ละค่าสัมประสิทธิ์การแปรผัน เมื่อขนาดตัวอย่างของประชากร $n_1 = 80$ และ $n_2 = 80$

ช่วงความเชื่อมั่น		ความน่าจะเป็นคุ้มรวม(ความยาวเฉลี่ยของช่วง) เมื่อ CV เท่ากับ							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
MLE		0.952	0.947	0.947	0.951	0.948	0.953	0.956	0.955
		(61.69)	(181.70)	(293.31)	(393.27)	(521.67)	(687.16)	(809.61)	(906.04)
Bayes	Right	0.950	0.945	0.945	0.947	0.949	0.952	0.954	0.954
		(61.49)	(181.07)	(292.39)	(392.04)	(519.91)	(684.35)	(807.34)	(903.44)
	Left	0.949	0.944	0.946	0.950	0.949	0.952	0.955	0.951
		(60.89)	(179.34)	(289.54)	(388.34)	(515.16)	(678.49)	(799.41)	(894.14)

หมายเหตุ : Right = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

Left = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

ตารางที่ 4 ความน่าจะเป็นคุ้มรวมและความยาวเฉลี่ยของช่วงความเชื่อมั่นในแต่ละค่าสัมประสิทธิ์การแปรผัน เมื่อขนาดตัวอย่างของประชากร $n_1 = 80$ และ $n_2 = 100$

ช่วงความเชื่อมั่น		ความน่าจะเป็นคุ้มรวม(ความยาวเฉลี่ยของช่วง) เมื่อ CV เท่ากับ							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
MLE		0.951	0.946	0.951	0.953	0.950	0.955	0.956	0.953
		(58.62)	(172.52)	(277.93)	(371.93)	(491.28)	(644.38)	(753.80)	(840.10)
Bayes	Right	0.948	0.948	0.950	0.949	0.950	0.951	0.958	0.952
		(58.30)	(171.77)	(276.99)	(371.26)	(492.01)	(648.55)	(761.67)	(851.83)
	Left	0.946	0.946	0.949	0.947	0.951	0.950	0.954	0.950
		(57.85)	(170.37)	(274.92)	(368.53)	(487.99)	(643.24)	(755.69)	(845.26)

หมายเหตุ : Right = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

Left = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

ตารางที่ 5 ความน่าจะเป็นคุ้มรวมและความยามเฉลี่ยของช่วงความเชื่อมั่นในแต่ละค่าสัมประสิทธิ์การแปรผัน เมื่อขนาดตัวอย่างของประชากร $n_1=100$ และ $n_2=100$

ช่วงความเชื่อมั่น		ความน่าจะเป็นคุ้มรวม(ความยาวเฉลี่ยของช่วง) เมื่อ CV เท่ากับ							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
MLE		0.948	0.945	0.955	0.951	0.948	0.950	0.952	0.955
		(55.23)	(162.54)	(261.80)	(350.97)	(463.67)	(606.48)	(711.64)	(798.99)
Bayes	Right	0.947	0.944	0.953	0.951	0.952	0.950	0.951	0.950
		(54.98)	(161.91)	(261.05)	(350.63)	(464.42)	(610.13)	(718.48)	(802.26)
	Left	0.946	0.947	0.951	0.948	0.950	0.950	0.953	0.956
		(54.56)	(160.69)	(259.03)	(347.86)	(460.83)	(605.30)	(712.80)	(796.16)

หมายเหตุ : Right = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior
Left = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

ตารางที่ 6 ความน่าจะเป็นคุ้มรวมและความยามเฉลี่ยของช่วงความเชื่อมั่นในแต่ละค่าสัมประสิทธิ์การแปรผัน เมื่อขนาดตัวอย่างของประชากร $n_1=100$ และ $n_2=150$

ช่วงความเชื่อมั่น		ความน่าจะเป็นคุ้มรวม(ความยาวเฉลี่ยของช่วง) เมื่อ CV เท่ากับ							
		0.1	**0.3**	**0.5**	**0.7**	**1**	**1.5**	**2**	**2.5**
MLE		0.948	0.948	0.948	0.950	0.951	0.952	0.953	0.955
		(50.44)	(148.70)	(238.99)	(319.94)	(422.74)	(552.93)	(647.52)	(719.86)
Bayes	Right	0.946	0.945	0.946	0.948	0.949	0.949	0.951	0.950
		(50.18)	(147.95)	(238.04)	(319.24)	(422.57)	(554.60)	(651.89)	(726.98)
	Left	0.948	0.946	0.948	0.948	0.949	0.950	0.950	0.952
		(49.88)	(147.11)	(236.58)	(317.35)	(420.27)	(551.53)	(648.22)	(722.02)

หมายเหตุ : Right = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior
Left = ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior

จากตารางที่ 1 ถึงตารางที่ 6 ขนาดตัวอย่างของทั้งสองประชากรเท่ากันและขนาดตัวอย่างของประชากรไม่เท่ากัน ผลลัพธ์ที่ได้ไม่แตกต่างกันมาก เมื่อพิจารณาค่าความน่าจะเป็นคุ้มรวม แสดงให้เห็นว่าค่าความน่าจะเป็นคุ้มรวมของช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior เข้าใกล้ระดับความเชื่อมั่นที่กำหนด และค่าความน่าจะเป็นคุ้มรวมของทั้ง 3 วิธี มีค่าไม่แตกต่างกันมาก

เมื่อค่าสัมประสิทธิ์การแปรผันเพิ่มขึ้นจะทำให้ค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นมีแนวโน้มเพิ่มขึ้นเรื่อย ๆ และเมื่อพิจารณาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่น แสดงให้เห็นว่าช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior ในส่วนมากจะมีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่สั้นกว่าค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่ได้จากตัวประมาณ

ภาวะน่าเป็นสูงสุด ดังนั้นช่วงความเชื่อมั่นแบบเบส์มีประสิทธิภาพสูงกว่าช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด

หากพิจารณาเฉพาะช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และ Left invariant prior จากตารางแสดงให้เห็นว่า ค่าความน่าจะเป็นคุ้มรวมของช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior เข้าใกล้ระดับความเชื่อมั่นที่กำหนด และค่าความน่าจะเป็นคุ้มรวมทั้ง 2 วิธีนี้มีค่าไม่แตกต่างกันมาก แต่เมื่อพิจารณาค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นจะได้ว่า ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior มีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่สั้นกว่าค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior ดังนั้นช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior มีประสิทธิภาพสูงกว่าช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior

อภิปรายและสรุปผลการวิจัย

ในการวิจัยนี้ได้เปรียบเทียบประสิทธิภาพของช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด (Maximum Likelihood Estimator) กับช่วงความเชื่อมั่นแบบเบส์ (Bayes credible interval) สำหรับผลต่างของค่ามัธยฐานของการแจกแจงล็อกปรกติ โดยใช้วิธี MOVER ซึ่งมีช่วงความเชื่อมั่นทั้งหมด 3 วิธี ได้แก่ ช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Right invariant prior และช่วงความเชื่อมั่นแบบเบส์ เมื่อกำหนดการแจกแจงก่อนเป็น Left invariant prior โดยพิจารณาความน่าจะเป็นคุ้มรวมและค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นเป็นเกณฑ์ในการเปรียบเทียบ จากผลการวิจัยแสดงให้เห็นว่าค่าความน่าจะเป็นคุ้มรวมของช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุดกับช่วงความเชื่อมั่นแบบเบส์เข้าใกล้ระดับความเชื่อมั่นที่กำหนด และค่าความน่าจะเป็นคุ้มรวมไม่แตกต่างกันมาก แต่ช่วงความเชื่อมั่นแบบเบส์มีค่าความยาวเฉลี่ยของช่วงความเชื่อมั่นที่สั้นกว่าช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด ดังนั้นช่วงความเชื่อมั่นแบบเบส์มีประสิทธิภาพสูงกว่าช่วงความเชื่อมั่นที่ได้จากตัวประมาณภาวะน่าเป็นสูงสุด

กิตติกรรมประกาศ

ผู้วิจัยขอขอบพระคุณ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ที่ให้ทุนสนับสนุนการนำเสนอผลงานวิทยานิพนธ์ ในประเทศประจำปีงบประมาณ 2562

เอกสารอ้างอิง

ปวริศร์ โกสีย์รัตน์. ช่วงความเชื่อมั่นสำหรับผลต่างของมัธยฐานของตัวอย่างโดยไม่จำกัดการแจกแจง [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์]. ธรรมศาสตร์: มหาวิทยาลัยธรรมศาสตร์; 2546.

มานพ วราภักดิ์. ทฤษฎีความน่าจะเป็น. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์ มหาวิทยาลัย; 2548.

สายชล สินสมบูรณ์ทอง. สถิติเบื้องต้น. พิมพ์ครั้งที่ 10. กรุงเทพฯ: จามจุรีโปรดักท์; 2555.

D’Cunha, J. G., and Rao, K. A. Bayesian inference for mean of the lognormal distribution. International Journal of Scientific and Research Publications 2014; 4(10): 195-203.

D’Cunha, J. G., & Rao, K. A. Bayesian inference for Median of the lognormal distribution. International Journal of Scientific and Research Publications 2016; 2(15): 526-535.

Donner, A., and Zou, G.Y. Construction of confidence limits about effect measures: A general approach. Statistics on Medicine 2008; 27(10): 1693-1720.

Harvey, J., and Van der Merwe A. J. Bayesian confidence intervals for means and variances of lognormal and bivariate lognormal distributions. Journal of Statistical Planning and Inference 2012; 6(142): 1294-1309.

Hettmansperger, P., Thomas. Two-Sample Inference Based on One-Sample Sign Statistics. Journal of the Royal Statistical Society 1984; 1(33): 45-51.

Krishnamoorthy, K., and Sazib Hasan, Improved Confidence Intervals for the Ratio of Coefficients of Variation of Two Lognormal Distributions. Journal of Statistical Theory and Applications 2017; 3(16): 345-353.

Padgett, W. J., and Wei, L. J. Bayes estimation of reliability for the two-parameter lognormal distribution. Communications in Statistics—Theory and Methods 1977; 6(5): 443-457.

Zellner, A. Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression. Journal of the American Statistical Association 1971; 66(334): 327-330.

Zou, Y.Z., Taleban, J., Huo, C.Y. Confidence interval estimation for lognormal data with application to health economics. Computational Statistics and Data Analysis 2009; 53: 3755–3764.