การเปรียบเทียบประสิทธิภาพของการประมาณช่วงความเชื่อมั่นของพารามิเตอร์ในการแจกแจงแกมมา

A COMPARISON OF EFFICIENCY OF CONFIDENCE INTERVAL ESTIMATION OF PARAMETER ON GAMMA DISTRIBUTION

ฉัตรวดี กิจแก้ว อัชฌา อระวีพร*

***Chatwadee Kitkeaw*,** Autcha Araveeporn*

ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang.

****Corresponding author, e-mail:** bewkawon_suju@hotmail.com*

***Received:** March 27, 2018; **Revised:** May 31, 2018; **Accepted:** June 26, 2018*

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยนี้ เพื่อเปรียบเทียบประสิทธิภาพของการประมาณช่วงความเชื่อมั่นของพารามิเตอร์ (λ) ของการแจกแจงแกมมา ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยพิจารณาจากค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น กล่าวคือ ถ้าวิธีการประมาณใดมีค่าสัมประสิทธิ์ความเชื่อมั่นอยู่ในช่วงที่กำหนด และมีค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นน้อยที่สุด จะถือว่า วิธีการประมาณนั้นให้ช่วงความเชื่อมั่นที่เหมาะสมที่สุดในสถานการณ์นั้น โดยทำการจำลองข้อมูลที่มีการแจกแจงแกมมา ซึ่งกำหนดค่าพารามิเตอร์รูปร่าง (α) เท่ากับ 2, 3, 4, 5, 6, 7 และ 8 กำหนดค่าพารามิเตอร์สเกล (λ) เท่ากับ 2 กำหนดขนาดตัวอย่างเท่ากับ 30, 50 และ 70 และกำหนดระดับความเชื่อมั่น เท่ากับ 95% และ 99% ผลการวิจัยพบว่า วิธีของเบส์เป็นวิธีที่มีประสิทธิภาพดีที่สุดในเกือบทุกกรณี ยกเว้นกรณีที่ $n=50$ เมื่อ $\alpha=5$ สำหรับวิธีมาร์คอฟ เชน มอนติคาร์โล มีประสิทธิภาพในการประมาณช่วงความเชื่อมั่นดีในบางกรณี คือกรณีที่ $n=50$ เมื่อ $\alpha=5$ และ $n=70$ เมื่อ $\alpha=8$ ที่ระดับความเชื่อมั่น 99% และวิธีภาวะน่าจะเป็นสูงสุด เป็นวิธีที่มีประสิทธิภาพต่ำที่สุดในการประมาณช่วงความเชื่อมั่น

คำสำคัญ: การแจกแจงแกมมา ภาวะน่าจะเป็นสูงสุด เบส์ มาร์คอฟ เชน มอนติคาร์โล

Abstract

The objective of this research was to estimate the confidence interval of parameter (λ) on gamma distribution by using Maximum Likelihood (ML), Bayes', and Markov Chain Monte Carlo (MCMC) methods. The efficient performance of these methods was considered by Confidence Coefficients (CC) and Average Width (AW). Therefore, the best estimation was the estimation, which having CC within the range of the fixed confidence interval, and having the lowest AW in each situation. The data was simulated from gamma distribution by setting the shape parameter (α) as 2, 3, 4, 5, 6, 7, and 8, the scale parameter or called true parameter (λ) as 2, sample sizes (n) as 30, 50, and 70, and the 95% and 99% confidence interval. The results revealed that Bayes' method showed the best performance in most all cases, except $n = 50$, $\alpha = 5$. MCMC method had the efficient performance to estimate the confidence interval in some cases such as $n = 50$, $\alpha = 5$ and $n = 70$, $\alpha = 8$ at the 99% confidence interval. Maximum likelihood method was the lowest efficient method to estimate the confidence interval.

Keywords: Gamma Distribution, Maximum Likelihood, Bayes', Markov Chain Monte Carlo

บทนำ

โดยทั่วไปการศึกษาข้อมูลของประชากร (Population) ทั้งหมดนั้นเป็นเรื่องที่ทำได้ยาก เนื่องจากประชากรมีขนาดใหญ่ ประกอบกับค่าใช้จ่ายที่ค่อนข้างสูง และต้องใช้เวลาค่อนข้างมากในการเก็บรวบรวมข้อมูล จึงได้ทำการสุ่มอย่าง (Sample) ขึ้นมาและศึกษาข้อมูลจากตัวอย่างแทน ซึ่งค่าที่คำนวณได้จากตัวอย่างที่สุ่มมานั้น เรียกว่า ตัวประมาณ (Estimator) หลังจากที่ทำการศึกษาข้อมูลจากตัวอย่างแล้ว จะนำผลที่ได้จากตัวอย่างอ้างอิงเพื่อหาข้อสรุปเกี่ยวกับประชากร เรียกว่า พารามิเตอร์ (Parameter) โดยวิธีการทางสถิติดังกล่าวนี้ เรียกว่า การอนุมานเชิงสถิติ (Statistical Inference) ซึ่งแบ่งออกเป็น 2 ลักษณะ ได้แก่ การประมาณค่าพารามิเตอร์ (Parameter Estimation) และการทดสอบสมมติฐาน (Hypothesis Testing) โดยการประมาณค่าพารามิเตอร์ จะแบ่งออกเป็น 2 วิธี คือ การประมาณค่าแบบจุด (Point Estimation) และการประมาณค่าแบบช่วง (Interval Estimation)

การประมาณค่าแบบจุด (Point Estimation) เป็นการประมาณค่าพารามิเตอร์ของประชากรด้วยค่าสถิติเพียงค่าเดียว ซึ่งเรียกว่า ตัวประมาณ โดยค่าที่ได้จะขึ้นอยู่กับตัวอย่างที่สุ่มมา จึงทำให้การประมาณค่าแบบจุดเกิดความคลาดเคลื่อนได้ ถ้าตัวอย่างที่สุ่มมาไม่เป็นตัวแทนที่ดีของประชากรควรเลือกใช้การประมาณค่าแบบช่วง (Interval Estimation) ซึ่งเป็นการประมาณค่าพารามิเตอร์ของประชากรในช่วงๆ หนึ่งที่เชื่อว่าจะครอบคลุมค่าพารามิเตอร์ด้วยระดับความเชื่อมั่นที่กำหนด ซึ่งการประมาณค่าแบบช่วงจะครอบคลุมค่าพารามิเตอร์มากกว่าการประมาณค่าแบบจุด โดยใช้ค่าประมาณแบบจุดและการแจกแจงความน่าจะเป็นของตัวสถิติในการสร้างขอบเขตบนและขอบเขตล่าง ซึ่งการประมาณค่าแบบช่วงสามารถบอกถึงโอกาสหรือความน่าจะเป็นที่ค่าพารามิเตอร์จะตกอยู่ในช่วงประมาณนั้น ในทางสถิติเรียกค่าความน่าจะเป็นนั้นว่า สัมประสิทธิ์ความเชื่อมั่น (Coefficient of Confidence) แทนด้วย $1-\alpha$ โดยที่ $0<\alpha<1$ โดย α คือ ค่าความน่าจะเป็นที่เกิดจากความผิดพลาดประเภทที่ 1 หรือความน่าจะเป็นที่จะปฏิเสธสมมติฐานหลักเมื่อสมมติฐานหลักเป็นจริง และมีช่วงความเชื่อมั่น $(1-\alpha)100\%$

วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method) เป็นวิธีการที่ใช้กันแพร่หลายสำหรับการประมาณค่าพารามิเตอร์ ซึ่งเป็นวิธีการหาตัวประมาณค่าที่ทำให้ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function) มีค่าสูงสุดจากฟังก์ชันความหนาแน่นความน่าจะเป็น (Probability Density Function : pdf) ซึ่งขึ้นอยู่กับค่าพารามิเตอร์และตัวประมาณค่าที่ได้จากวิธีนี้จะมีคุณสมบัติของตัวประมาณค่าที่ดี โดยคุณสมบัติดังกล่าว ได้แก่ ความไม่เอนเอียง ความคงเส้นคงวา ความพอเพียง ความสมบูรณ์

ซึ่งเรียกว่าตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุด (Maximum Variance Unbiased Estimator : MVUE) [1] แต่วิธีการนี้ก็ยังมีปัญหาในการหาตัวประมาณค่า กล่าวคือ ในบางครั้งการแก้สมการเพื่อหาตัวประมาณค่าอาจทำได้ยาก ถ้าสมการภาวะน่าจะเป็นเป็นสมการที่มีกำลังสูงๆ หรือเป็นเศษส่วนที่ซับซ้อน และในบางกรณีก็ไม่สามารถหาตัวประมาณค่าได้ด้วยวิธีอนุพันธ์ สำหรับวิธีของเบส์ (Bayes' Method) [2] จะนำความรู้เดิมที่เกี่ยวกับพารามิเตอร์มาใช้ให้เป็นประโยชน์ เพื่อให้การประมาณค่าพารามิเตอร์ได้ผลดียิ่งขึ้น วิธีของเบส์นั้นเป็นการพิจารณาฟังก์ชันภาวะน่าจะเป็นที่ได้จากวิธีภาวะน่าจะเป็นสูงสุดและการแจกแจงก่อน (Prior Distribution) ซึ่งในการแจกแจงก่อนจะทำให้ทราบข้อมูลเกี่ยวกับพารามิเตอร์ก่อนที่จะกำหนดข้อมูล จากฟังก์ชันภาวะน่าจะเป็นและการแจกแจงก่อนจะนำมาคำนวณหาการแจกแจงภายหลัง (Posterior Distribution) ซึ่งการแจกแจงภายหลังแสดงให้เห็นถึงข้อมูลทั้งหมดที่เกี่ยวกับพารามิเตอร์เมื่อกำหนดข้อมูลแล้ว หลังจากนั้นจะได้ตัวประมาณเบส์ภายหลังจากการคำนวณค่าเฉลี่ยของการแจกแจงภายหลัง โดยขึ้นอยู่กับพารามิเตอร์ของการแจกแจงก่อน แต่วิธีการนี้ก็ยังมีปัญหาในการประมาณค่าพารามิเตอร์ของการแจกแจงก่อน ซึ่งจะประมาณค่าพารามิเตอร์ได้ยาก จึงได้นำวิธีมาร์คอฟ เชน มอนติคาร์โล (Markov Chain Monte Carlo Method) มาใช้ในการแก้ปัญหาดังกล่าว โดยวิธีมาร์คอฟ เชน มอนติคาร์โล เป็นวิธีการที่จำลองค่าจากการแจกแจงภายหลังซึ่งพัฒนาขึ้นจากวิธีของเบส์ และวิธีการสุ่มตัวอย่างแบบกิบส์ (Gibb Sampling) [3] ถูกนำมาใช้ในวิธีมาร์คอฟ เชน มอนติคาร์โล ซึ่งวิธีการสุ่มตัวอย่างแบบกิบส์เป็นวิธีที่ได้รับความนิยมในการสร้างค่าจากการแจกแจงภายหลัง ซึ่งประมาณเป็นตัวประมาณค่ามาร์คอฟ เชน มอนติคาร์โล

จากการศึกษางานวิจัยที่เกี่ยวข้องกับวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล พบว่า Araveeporn [4] ได้เปรียบเทียบวิธีประมาณค่าพารามิเตอร์แบบช่วงของการแจกแจงปัวซง (Poisson Distribution) โดยวิธีภาวะน่าจะเป็นสูงสุด วิธีมาร์คอฟ เชน มอนติคาร์โล และวิธีของเบส์ โดยประมาณค่าพารามิเตอร์ของการแจกแจงก่อน (λ) พบว่า วิธีภาวะน่าจะเป็นสูงสุดเป็นวิธีการประมาณที่ดีที่สุด เมื่อพารามิเตอร์มีขนาดเล็ก ($\lambda = 0.5$) สำหรับทุกขนาดตัวอย่าง เมื่อขนาดตัวอย่างมีขนาดใหญ่ ($n = 100, 300, 500$) วิธีมาร์คอฟ เชน มอนติคาร์โล จะเป็นวิธีการประมาณที่ดีที่สุดเมื่อ $\lambda = 5$ ส่วนวิธีของเบส์เป็นวิธีการประมาณที่ดีที่สุดในทุกกรณี ศรสวรรค์ บุญเพ็ญ และคณะ [5] ได้ศึกษาและเปรียบเทียบวิธีการประมาณค่าแบบช่วงสำหรับพารามิเตอร์ขนาด (η) ของการแจกแจงไวบูลล์ (Weibull Distribution) แบบสองพารามิเตอร์ เมื่อทราบค่าพารามิเตอร์ (β) ด้วยวิธีการประมาณ 2 วิธี คือ วิธีบูตสแตรปที (Bootstrap-t Method) และวิธีแจ็คไนฟ์ (Jackknife Method) พบว่า วิธีบูตสแตรปทีเป็นวิธีที่ดีที่สุดสำหรับทุกๆ สถานการณ์ที่ทำการศึกษา Pradhan และ Kundu [6] ได้ศึกษาการประมาณเบส์และการทำนายของการแจกแจงแกมมา (Gamma Distribution) 2 พารามิเตอร์ โดยสมมติว่า พารามิเตอร์สเกล (λ) มีการแจกแจงแกมมา ส่วนพารามิเตอร์รูปร่าง (α) มีการแจกแจงล็อก-คอนเคฟ (Log-Concave Distribution) และทำการคำนวณช่วงความเชื่อมั่น 95% จากวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ที่ใช้การประมาณของลินลี่ (Lindley Estimation) และวิธีของเบส์ที่ใช้วิธีมาร์คอฟ เชน มอนติคาร์โล ในการประมาณ พบว่า เมื่อไม่ทราบการแจกแจงของข้อมูล วิธีของเบส์และวิธีภาวะน่าจะเป็นสูงสุดมีประสิทธิภาพในการประมาณช่วงความเชื่อมั่นใกล้เคียงกัน ในขณะที่ถ้าทราบการแจกแจงของ

ข้อมูล วิธีของเบส์จะมีประสิทธิภาพในการประมาณช่วงความเชื่อมั่นดีกว่าวิธีภาวะน่าจะเป็นสูงสุด

ในงานวิจัยนี้ ผู้วิจัยสนใจศึกษาข้อมูลที่มีการแจกแจงแกมมา โดยสนใจตัวแปรสุ่ม X แทนช่วงเวลาในการรอคอยที่จะเกิดความสำเร็จขึ้น α ครั้ง ด้วยจำนวนสิ่งที่สนใจโดยเฉลี่ยใน 1 หน่วยเวลา (λ) โดยที่ λ เป็นพารามิเตอร์ของการแจกแจงปัวซง นอกจากนี้การแจกแจงแกมมายังเกี่ยวข้องกับการแจกแจงแบบเลขชี้กำลัง (Exponential Distribution) ถ้าช่วงเวลาในการรอคอยที่จะเกิดความสำเร็จเกิดขึ้นเป็นครั้งแรก กล่าวคือ การแจกแจงแบบเลขชี้กำลังเป็นกรณีเฉพาะของการแจกแจงแกมมา สำหรับการแจกแจงแกมมาสามารถนำไปประยุกต์ใช้ในด้านการเงิน วิศวกรรมโยธา และด้านเศรษฐมิติ เป็นต้น

ดังนั้นผู้วิจัยจึงมีความสนใจที่จะศึกษาการประมาณช่วงความเชื่อมั่นของพารามิเตอร์ในการแจกแจงแกมมา โดยสนใจประมาณค่าพารามิเตอร์ (λ) ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล นอกจากนี้ ในด้านศึกษาต่อ ผู้วิจัยสามารถใช้วิธีการประมาณอื่นๆ ในการศึกษาเปรียบเทียบการประมาณช่วงความเชื่อมั่นของพารามิเตอร์ในการแจกแจงแกมมา เช่น วิธีสคอร์ (Score Method) และวิธีบูตสแตรป (Bootstrap Method) เป็นต้น

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาการประมาณช่วงความเชื่อมั่น สำหรับพารามิเตอร์ของการแจกแจงแกมมา ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล
2. เพื่อเปรียบเทียบประสิทธิภาพของการประมาณช่วงความเชื่อมั่น สำหรับพารามิเตอร์ของการแจกแจงแกมมา ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล

วิธีการประมาณค่า

ในงานวิจัยครั้งนี้ศึกษาวิธีการประมาณค่าพารามิเตอร์ λ ของการแจกแจงแกมมา 3 วิธี ดังนี้

1. วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method : ML)

กำหนดให้ $X_1,...,X_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงแกมมาที่มีพารามิเตอร์รูปร่าง (shape parameter) α และพารามิเตอร์สเกล (Scale Parameter) λ หรือเขียนได้ว่า $X_i \sim Gamma(\alpha,\lambda)$ โดยมีฟังก์ชันความหนาแน่นความน่าจะเป็น ดังนี้

$$f(x;\alpha,\lambda)=\begin{cases}\dfrac{1}{\Gamma(\alpha)}\lambda^{\alpha}x^{\alpha-1}e^{-\lambda x} & ,\ x\geq 0\\ 0 & ,\ x \text{ มีค่าอื่นๆ}\end{cases}$$

โดยมีค่าคาดหมาย $E(X)=\dfrac{\alpha}{\lambda}$ และค่าความแปรปรวน $Var(X)=\dfrac{\alpha}{\lambda^2}$

ตัวประมาณด้วยวิธีภาวะน่าจะเป็นสูงสุดของ λ พิจารณาจากฟังก์ชันภาวะน่าจะเป็น ดังนี้

$$
\begin{aligned}
L(\alpha,\lambda) &= \prod_{i=1}^{n} f(x_i;\alpha,\lambda) \\
&= \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^{n} \prod_{i=1}^{n} x_i^{\alpha-1} e^{-\lambda\sum_{i=1}^{n} x_i} \\
\ln L(\alpha,\lambda) &= \alpha n \ln(\lambda) - n\ln(\Gamma(\alpha)) + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \lambda\sum_{i=1}^{n} x_i \\
\frac{\partial \ln L(\alpha,\lambda)}{\partial\lambda} &= \frac{\partial}{\partial\lambda}\left(\alpha n \ln(\lambda) - n\ln(\Gamma(\alpha)) + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \lambda\sum_{i=1}^{n} x_i\right) \\
\frac{n\alpha}{\lambda} &= \sum_{i=1}^{n} x_i \\
\lambda &= \frac{\alpha}{\bar{x}} \qquad (1)
\end{aligned}
$$

และอนุพันธ์อันดับสองของ $\ln L(\alpha,\lambda)$ คือ $\frac{\partial^2 \ln L(\alpha,\lambda)}{\partial\lambda^2} = -\frac{n\alpha}{\lambda^2} < 0$ จะได้ว่า ค่าประมาณที่ได้มีค่าสูงสุด

ดังนั้น ตัวประมาณภาวะน่าจะเป็นสูงสุดของ λ คือ $\hat{\lambda}_{ML} = \frac{\alpha}{\bar{x}}$

จาก $\hat{\lambda}_{ML} = \frac{\alpha}{\bar{x}}$ ค่าคาดหมายของตัวประมาณภาวะน่าจะเป็นสูงสุด จาก (1) หาได้จาก

$$E\left(\hat{\lambda}_{ML}\right) = E\left(\frac{\alpha}{\bar{X}}\right) = n\alpha E\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) \qquad (2)$$

เนื่องจาก $X_i \sim Gamma(\alpha,\lambda)$ และ $X_1,...,X_n$ เป็นอิสระกัน จะได้ $\sum_{i=1}^{n} X_i \sim Gamma(n\alpha,\lambda)$

จาก (2) จะได้ว่า $$E\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) = \frac{\lambda^{n\alpha}}{\Gamma(n\alpha)}\left(\frac{1}{\lambda^{n\alpha-1}}\right)\int_0^{\infty}\left(\lambda\sum_{i=1}^{n} x_i\right)^{n\alpha-2} e^{-\lambda\sum_{i=1}^{n} x_i} d\left(\lambda\sum_{i=1}^{n} x_i\right) \qquad (3)$$

จากฟังก์ชันแกมมา $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1}e^{-x}dx$ และ $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

จาก (3) จะได้ $$E\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) = \frac{\lambda}{\Gamma(n\alpha)}\left(\Gamma(n\alpha-1)\right) = \frac{\lambda}{n\alpha-1}$$

จาก (2) สรุปได้ว่า $n\alpha E\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) = \frac{n\alpha\lambda}{n\alpha - 1} = \frac{n\alpha^2}{(n\alpha - 1)\overline{x}}$ (4)

ดังนั้น ค่าคาดหมายเท่ากับ $E\left(\hat{\lambda}_{ML}\right) = \frac{n\alpha^2}{(n\alpha - 1)\overline{x}}$

จาก $\hat{\lambda}_{ML} = \frac{\alpha}{\overline{x}}$ ค่าความแปรปรวนของตัวประมาณภาวะน่าจะเป็นสูงสุด หาได้จาก

$$Var\left(\hat{\lambda}_{ML}\right) = Var\left(\frac{\alpha}{\overline{X}}\right) = n^2\alpha^2 Var\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) \tag{5}$$

จาก (5) เนื่องจาก $Var\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) = E\left(\frac{1}{\left(\sum_{i=1}^{n} X_i\right)^2}\right) - \left(E\left(\frac{1}{\sum_{i=1}^{n} X_i}\right)\right)^2$ (6)

จาก (6) ค่าคาดหมาย เป็นดังนี้

$$E\left(\frac{1}{\left(\sum_{i=1}^{n} X_i\right)^2}\right) = \frac{\lambda^{n\alpha}}{\Gamma(n\alpha)}\left(\frac{1}{\lambda^{n\alpha-2}}\right)\int_0^\infty \left(\lambda\sum_{i=1}^{n} x_i\right) e^{-\lambda\sum_{i=1}^{n} x_i} d\left(\lambda\sum_{i=1}^{n} x_i\right)$$

$$= \frac{\lambda^{n\alpha}}{\Gamma(n\alpha)}\left(\frac{\Gamma(n\alpha - 2)}{\lambda^{n\alpha-2}}\right) = \frac{\lambda^2}{(n\alpha - 1)(n\alpha - 2)}$$

จาก (6) จะได้ค่าความแปรปรวน ดังนี้

$$Var\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) = \frac{\lambda^2}{(n\alpha - 1)(n\alpha - 2)} - \left(\frac{\lambda}{n\alpha - 1}\right)^2 = \frac{\lambda^2}{(n\alpha - 1)^2(n\alpha - 2)} \tag{7}$$

จาก (5) สรุปได้ว่า $Var\left(\hat{\lambda}_{ML}\right) = n^2\alpha^2 Var\left(\frac{1}{\sum_{i=1}^{n} X_i}\right) = \frac{n^2\alpha^4}{(n\alpha - 1)^2(n\alpha - 2)\overline{x}^2}$ (8)

ดังนั้น ค่าความแปรปรวนเท่ากับ

$$Var\left(\hat{\lambda}_{ML}\right) = \frac{n^2\alpha^4}{(n\alpha - 1)^2(n\alpha - 2)\overline{x}^2}$$

2. วิธีของเบส์ (Bayes' Method : Bayes')

กำหนดให้ตัวอย่างสุ่ม $X_1,...,X_n$ มีการแจกแจงแกมมาด้วยจำนวนครั้งที่เกิดเหตุการณ์โดยเฉลี่ยใน 1 หน่วยเวลา (λ) ซึ่งเป็นค่าพารามิเตอร์ที่ต้องการประมาณ สามารถเขียนได้ในรูปแบบ $X_i|\lambda \sim Gamma(\alpha,\lambda)$ ฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่ม X เมื่อกำหนด α และ λ คือ

$$\begin{aligned} f\left(x_1,...,x_n|\alpha,\lambda\right) &= \prod_{i=1}^{n} f\left(x_i;\alpha,\lambda\right) \\ &= \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^n \prod_{i=1}^{n} x_i^{\alpha-1} e^{-\lambda\sum_{i=1}^{n} x_i} \end{aligned}$$

โดยพิจารณาฟังก์ชันการแจกแจงก่อนให้มีการแจกแจงแกมมา ที่มีพารามิเตอร์ a และ b ดังนี้

$$f\left(\lambda|a,b\right) = \begin{cases} \dfrac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda} & ,\lambda > 0 \\ 0 & ,\lambda \text{ มีค่าอื่นๆ} \end{cases}$$

จากรูปแบบวงศ์คู่สังยุค (Conjugate Distribution) จะได้ฟังก์ชันการแจกแจงภายหลัง คือ

$$\begin{aligned} h\left(\lambda|x_1,...,x_n\right) &= L(\alpha,\lambda) f\left(\lambda|a,b\right) \\ h\left(\lambda|x_1,...,x_n\right) &= \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^n \prod_{i=1}^{n} x_i^{\alpha-1} e^{-\lambda\sum_{i=1}^{n} x_i} \left(\frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda}\right) \\ &= \frac{b^a}{\left(\Gamma(\alpha)\right)^n\left(\Gamma(a)\right)} \prod_{i=1}^{n} x_i^{\alpha-1} \lambda^{n\alpha+a-1} e^{-\lambda\left(n\bar{x}+b\right)} \end{aligned}$$

จากฟังก์ชันการแจกแจงภายหลังของ λ คือ $h(\lambda|x_1,...,x_n) \sim Gamma(n\alpha+a, n\bar{x}+b)$

ดังนั้น ตัวประมาณเบส์ภายหลังของ λ คือ $\hat{\lambda}_{Bayes'} = E\left(\lambda|x_1,...,x_n\right) = \dfrac{n\alpha+a}{n\bar{x}+b}$ และ $Var\left(\lambda|x_1,...,x_n\right) = \dfrac{n\alpha+a}{\left(n\bar{x}+b\right)^2}$

3. วิธีมาร์คอฟ เชน มอนติคาร์โล (Markov Chain Monte Carlo Method : MCMC)

วิธีมาร์คอฟ เชน มอนติคาร์โล (MCMC) [7] เป็นวิธีที่นิยมใช้เมื่อไม่ทราบฟังก์ชันการแจกแจงบางส่วน (Marginal Distribution) ของตัวแปรสุ่ม โดยประกอบด้วยการสุ่มตั วอย่างตัวแปรจากวิธีมาร์คอฟ เชน (Markov Chain) จากการแจกแจงก่อนของค่าพารามิเตอร์ที่สนใจ และวิธีการสุ่มตัวอย่างแบบกิบส์ (Gibb Sampling) ได้ถูกนำมาใช้ในวิธีของ MCMC สำหรับประมาณค่าพารามิเตอร์จากฟังก์ชันการแจกแจงภายหลัง

ขั้นตอนการสุ่มตัวอย่างจากวิธีมาร์คอฟ เชน มอนติคาร์โล ประกอบด้วย

(1) กำหนดค่าเริ่มต้น $a^{(t)}$ จากฟังก์ชันการแจกแจงเลขชี้กำลัง และ $b^{(t)}$ จากฟังก์ชันการแจกแจงแกมมา โดยที่ $a > 0,\ b > 0$

(2) สร้างค่าจาก (1) มา T ค่า เมื่อ $t = 1,2,...,T$

(3) สร้างค่า $\lambda^{(t)}$ จากฟังก์ชันการแจกแจงภายหลังของการแจกแจงแกมมาที่ค่าพารามิเตอร์ $a^{(t)}$ และ $b^{(t)}$ ที่ได้จาก (1) และคำนวณค่าคาดหมาย ค่าความแปรปรวน จากฟังก์ชันการแจกแจงภายหลัง

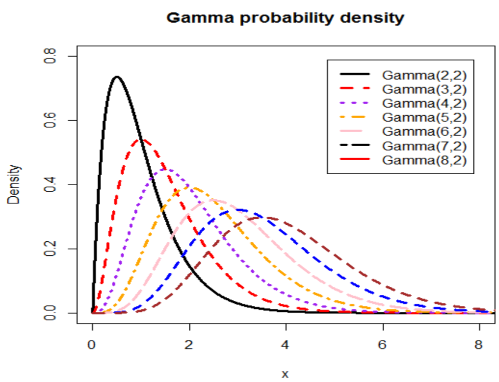
ดังนั้น ตัวประมาณมาร์คอฟ เชน มอนติคาร์โลของ λ คือ $\hat{\lambda}_{MCMC} = E\left(\hat{\lambda}_{MCMC}\right) = \frac{1}{T}\sum_{t=1}^{T}\lambda^{(t)}$ และ

$$Var\left(\hat{\lambda}_{MCMC}\right) = \frac{1}{T-1}\sum_{t=1}^{T}\left(\lambda^{(t)} - \bar{\lambda}\right)^2$$

วิธีดำเนินการวิจัย

ในการวิจัยครั้งนี้มีขั้นตอนดังนี้

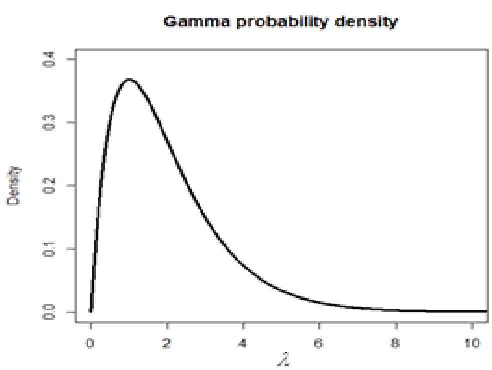
1. จำลองข้อมูลที่ใช้ในการวิจัยด้วยโปรแกรม R เวอร์ชัน 3.3.1 โดยกำหนดให้ข้อมูลมีการแจกแจงแกมมา ซึ่งกำหนดค่าพารามิเตอร์ (a) เท่ากับ 2, 3, 4, 5, 6, 7 และ 8 และกำหนดค่าพารามิเตอร์ (λ) เท่ากับ 2 ซึ่งมีกราฟแสดงฟังก์ชันความหนาแน่นความน่าจะเป็นของการแจกแจงแกมมา ดังแสดงในภาพที่ 1



ภาพที่ 1 การแจกแจงแกมมาที่มีพารามิเตอร์ (2,2) (3,2) (4,2) (5,2) (6,2) (7,2) และ (8,2)

2. กำหนดขนาดตัวอย่าง (n) เท่ากับ 30, 50 และ 70

3. กำหนดค่าพารามิเตอร์ของการแจกแจงก่อนเป็นการแจกแจงแกมมา โดยมีค่าพารามิเตอร์เท่ากับ $Gamma(2,1)$ หรือการแจกแจงแกมมา (2,1) ซึ่งการที่เลือกการแจกแจงแกมมา (2,1) ดังแสดงในภาพที่ 2 เป็นฟังก์ชันการแจกแจงก่อน เนื่องจากการแจกแจงแกมมา (2,1) มีประสิทธิภาพดีกว่าการแจกแจงอื่นๆ [8] เช่น การแจกแจงแบบเลขชี้กำลัง การแจกแจงไคกำลังสอง (Chi-Square Distribution)



ภาพที่ 2 การแจกแจงแกมมาที่มีพารามิเตอร์ (2,1)

4. กำหนดระดับความเชื่อมั่น 2 ระดับ คือ 95% และ 99%

5. คำนวณช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของพารามิเตอร์ (λ) จากวิธีการประมาณทั้ง 3 วิธี โดยใช้ทฤษฎีลิมิตเข้าสู่ส่วนกลาง (Central Limit Theorem) [9] เมื่อไม่ทราบรูปแบบการแจกแจง แต่ขนาดตัวอย่าง (n) มีจำนวนมาก พบว่า การแจกแจงจะลู่เข้าสู่การแจกแจงปกติมาตรฐาน (Standard Normal Distribution) ดังนั้น ช่วงความเชื่อมั่น $(1-\alpha)100\%$ หาได้จาก

$$P\left(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

5.1 วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method)

ตัวประมาณภาวะน่าจะเป็นสูงสุดของ λ คือ $\hat{\lambda}_{ML} = \frac{\alpha}{\bar{x}}$ เมื่อ $E\left(\hat{\lambda}_{ML}\right) = \frac{n\alpha^2}{(n\alpha-1)\bar{x}}$ และ

$Var\left(\hat{\lambda}_{ML}\right) = \frac{n^2\alpha^4}{(n\alpha-1)^2(n\alpha-2)\bar{x}^2}$ จะได้ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ λ คือ

$$\left(\frac{n\alpha-1}{n\bar{x}} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\alpha^2}{(n\alpha-2)\bar{x}^2}}\ ,\ \frac{n\alpha-1}{n\bar{x}} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\alpha^2}{(n\alpha-2)\bar{x}^2}}\right) \tag{9}$$

5.2 วิธีของเบส์ (Bayes' Method)

ตัวประมาณเบส์ภายหลังของ λ คือ $\hat{\lambda}_{Bayes'} = E\left(\lambda | x_1,...,x_n\right) = \frac{n\alpha+a}{n\bar{x}+b}$

และ $Var\left(\lambda | x_1,...,x_n\right) = \frac{n\alpha+a}{(n\bar{x}+b)^2}$ จะได้ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ λ คือ

$$\left(\left(\frac{n\alpha-a}{n\bar{x}+b}\right) - Z_{\frac{\alpha}{2}}\sqrt{\frac{n\alpha+a}{(n\bar{x}+b)^2}}\ ,\ \left(\frac{n\alpha-a}{n\bar{x}+b}\right) + Z_{\frac{\alpha}{2}}\sqrt{\frac{n\alpha+a}{(n\bar{x}+b)^2}}\right) \tag{10}$$

5.3 วิธีมาร์คอฟ เชน มอนติคาร์โล (Markov Chain Monte Carlo Method)

ตัวประมาณมาร์คอฟ เชน มอนติคาร์โลของ λ คือ $\hat{\lambda}_{MCMC} = E\left(\hat{\lambda}_{MCMC}\right) = \frac{1}{T}\sum_{t=1}^{T}\lambda^{(t)}$ และ

$Var\left(\hat{\lambda}_{MCMC}\right) = \frac{1}{T-1}\sum_{t=1}^{T}\left(\lambda^{(t)} - \bar{\lambda}\right)^2$ จะได้ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ λ คือ

$$\left(\frac{1}{T}\sum_{t=1}^{T}\lambda^{(t)} - Z_{\frac{\alpha}{2}}\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(\lambda^{(t)}-\bar{\lambda}\right)^2}\ ,\ \frac{1}{T}\sum_{t=1}^{T}\lambda^{(t)} + Z_{\frac{\alpha}{2}}\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(\lambda^{(t)}-\bar{\lambda}\right)^2}\right) \tag{11}$$

6. คำนวณค่าประมาณสัมประสิทธิ์ความเชื่อมั่น (Confidence Coefficient : CC)

โดยนำช่วงความเชื่อมั่นที่ได้จากการประมาณจากวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ และวิธีมาร์คอฟ เชน มอนติคาร์โล มาพิจารณาว่าครอบคลุมค่าพารามิเตอร์ λ หรือไม่ ถ้าช่วงความเชื่อมั่นใด ครอบคลุมค่าพารามิเตอร์ λ จะทำการนับจำนวนครั้ง หลังจากนั้นนำจำนวนครั้งทั้งหมดมาหารด้วยจำนวนรอบ (m)

โดยค่าที่ได้เรียกว่า ค่าประมาณสัมประสิทธิ์ความเชื่อมั่น $(1-\hat{\alpha})$ ซึ่งคำนวณได้ ดังนี้

$$1-\hat{\alpha}=\frac{\text{จำนวนครั้งที่ช่วงความเชื่อมั่นครอบคลุมค่าพารามิเตอร์}}{m}$$

7. เปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่นกับค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด

เป็นการเปรียบเทียบค่าประมาณสัมประสิทธิ์ความเชื่อมั่น $(1-\hat{\alpha})$ ว่ามีค่าอยู่ในช่วงของค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด $(1-\alpha_0, 1+\alpha_0)$ หรือไม่ ซึ่งพิจารณาจาก

$$P_0-Z_{\frac{\alpha}{2}}\sqrt{\frac{P_0(1-P_0)}{m}}=1-\alpha_0\leq 1-\hat{\alpha}\leq 1+\alpha_0=P_0+Z_{\frac{\alpha}{2}}\sqrt{\frac{P_0(1-P_0)}{m}}$$

โดยที่ P_0 แทน สัมประสิทธิ์ความเชื่อมั่นที่กำหนด หรือ $P_0=1-\alpha_0$ ในที่นี้กำหนดให้ $P_0=0.95$ และ 0.99

α แทน ระดับนัยสำคัญของการทดสอบ ในที่นี้กำหนดให้ $\alpha=0.05$

m แทน จำนวนรอบของการทดลอง ในที่นี้กำหนดให้ $m=1,000$

โดยทำการพิจารณาแต่ละกรณี ดังนี้

กรณีที่ 1 เมื่อ $P_0=0.95$ จะถือว่า วิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นอยู่ในช่วงค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อ $1-\alpha_0=0.95-1.96\sqrt{\frac{0.95(1-0.95)}{1,000}}=0.9365$ และ

$$1+\alpha_0=0.95+1.96\sqrt{\frac{0.95(1-0.95)}{1,000}}=0.9635$$

ดังนั้น ถ้าค่าประมาณสัมประสิทธิ์ความเชื่อมั่นอยู่ภายในช่วง $(0.9365, 0.9635)$ ก็จะถือว่า วิธีการประมาณนั้นผ่านเกณฑ์

กรณีที่ 2 เมื่อ $P_0=0.99$ จะถือว่า วิธีการประมาณที่ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นอยู่ในช่วงค่าสัมประสิทธิ์ความเชื่อมั่นที่กำหนด เมื่อ $1-\alpha_0=0.99-1.96\sqrt{\frac{0.99(1-0.99)}{1,000}}=0.9838$ และ

$$1+\alpha_0=0.99+1.96\sqrt{\frac{0.99(1-0.99)}{1,000}}=0.9962$$

ดังนั้น ถ้าค่าประมาณสัมประสิทธิ์ความเชื่อมั่นอยู่ภายในช่วง ก็จะถือว่า วิธีการประมาณนั้นผ่านเกณฑ์

8. คำนวณค่าเฉลี่ยความกว้างของช่วงความเชื่อมั่น (Average Width of Confidence Interval : AW)

โดยทำการหาผลต่างระหว่างขอบเขตบนและขอบเขตล่างของช่วงความเชื่อมั่น และนำค่าที่ได้มาหารด้วยจำนวนรอบของการทดลอง ซึ่งค่าที่ได้เรียกว่า ค่าประมาณของความกว้างเฉลี่ยของช่วงความเชื่อมั่น โดยคำนวณได้จาก

$$\text{ค่าประมาณของความกว้างเฉลี่ยของช่วงความเชื่อมั่น}=\frac{\sum_{j=1}^{m}(U_j-L_j)}{m}$$

โดยที่ U_j แทน ขอบเขตบนของช่วงความเชือมัน ในการทำซ้ำครั้งที่ j

L_j แทน ขอบเขตล่างของช่วงความเชื่อมั่น ในการทำซ้ำครั้งที่ j

m แทน จำนวนรอบของการทดลอง โดยในที่นี้เท่ากับ 1,000 รอบ

ผลการวิจัย

ผลการวิจัยที่ได้จากวิธีการประมาณช่วงความเชื่อมั่นทั้ง 3 วิธี ได้แก่ วิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงแกมมา (2,1) และวิธีมาร์คอฟ เชน มอนติคาร์โล แสดงไว้ในตารางที่ 1 และ 2 โดยจะพิจารณาจากค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ซึ่งผลที่ได้จะแยกตามระดับความเชื่อมั่นที่กำหนด โดยมีสัญลักษณ์แทนความหมายต่างๆ ดังนี้

ML แทน วิธีภาวะน่าจะเป็นสูงสุด

Bayes' แทน วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงแกมมา (2,1)

MCMC แทน วิธีมาร์คอฟ เชน มอนติคาร์โล

1. ระดับความเชื่อมั่น 95%

ผลที่ได้จากวิธีการประมาณทั้ง 3 วิธี แสดงดังตารางที่ 1

ตารางที่ 1 ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณที่ระดับความเชื่อมั่น 95%

ขนาดตัวอย่าง (n)	พารามิเตอร์รูปร่าง (α)	วิธีการประมาณค่า					
		ML		Bayes'		MCMC	
		CC	AW	CC	AW	CC	AW
30	2	0.996	-	0.946	**1.01271**	0.945	1.03031
	3	0.998	-	0.947	**0.82822**	0.941	0.83737
	4	0.999	-	0.953	**0.71660**	0.950	0.72305
	5	0.997	-	0.941	**0.64127**	0.942	0.64625
	6	1.000	-	0.945	**0.58457**	0.945	0.58848
	7	0.998	-	0.942	**0.54129**	0.942	0.54332
	8	1.000	-	0.945	**0.50639**	0.943	0.50824
50	2	0.996	-	0.943	**0.78367**	0.943	0.79172
	3	0.995	-	0.956	**0.64172**	0.955	0.64661
	4	0.999	-	0.942	**0.55495**	0.936	-
	5	0.998	-	0.934	-	0.935	-
	6	1.000	-	0.954	**0.45242**	0.952	0.45380
	7	0.998	-	0.942	**0.41906**	0.940	0.42108
	8	1.000	-	0.951	**0.39211**	0.948	0.39267
70	2	0.998	-	0.955	**0.66234**	0.950	0.66580
	3	1.000	-	0.951	**0.54196**	0.948	0.54418
	4	1.000	-	0.959	**0.46903**	0.954	0.47016
	5	0.997	-	0.954	**0.41937**	0.953	0.42081
	6	1.000	-	0.939	**0.38239**	0.943	0.38398
	7	0.998	-	0.941	**0.35418**	0.941	0.35475
	8	1.000	-	0.948	**0.33135**	0.947	0.33191

หมายเหตุ: ค่าที่ขีดเส้นใต้ หมายถึง ค่าสัมประสิทธิ์ความเชื่อมั่นที่ได้จากการประมาณไม่อยู่ในเกณฑ์ที่กำหนดค่าที่พิมพ์ด้วยตัวหนา หมายถึง ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่แคบที่สุดในแต่ละสถานการณ์

- หมายถึง ไม่พิจารณาค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น เนื่องจากให้ค่าสัมประสิทธิ์ความเชื่อมั่นไม่อยู่ในเกณฑ์ที่กำหนด

จากตารางที่ 1 พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงแกมมา (2,1) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่แคบที่สุดเกือบทุกสถานการณ์ ยกเว้นกรณีที่ $n = 50$ เมื่อ $\alpha = 5$ สำหรับวิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าประมาณสัมประสิทธิ์ความเชื่อมั่น (CC) อยู่ในเกณฑ์เกือบทุกสถานการณ์ ยกเว้นกรณีที่ $n = 50$ เมื่อ $\alpha = 4$ และ 5 แต่วิธีการประมาณนี้ไม่ได้ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นที่แคบที่สุดในทุกสถานการณ์ และวิธีภาวะน่าจะเป็นสูงสุด เป็นวิธีที่ค่าประมาณสัมประสิทธิ์ความเชื่อมั่นไม่อยู่ในเกณฑ์ที่กำหนดในทุกสถานการณ์

2. ระดับความเชื่อมั่น 99%

ผลที่ได้จากวิธีการประมาณทั้ง 3 วิธี แสดงดังตารางที่ 2

ตารางที่ 2 ค่าสัมประสิทธิ์ความเชื่อมั่น (CC) และค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่ได้จากการประมาณที่ระดับความเชื่อมั่น 99%

ขนาดตัวอย่าง (n)	พารามิเตอร์รูปร่าง (α)	วิธีการประมาณค่า					
		ML		Bayes’		MCMC	
		CC	AW	CC	AW	CC	AW
30	2	0.998	-	0.990	**1.33093**	0.990	1.35534
	3	0.999	-	0.992	**1.08846**	0.990	1.10053
	4	1.000	-	0.989	**0.94178**	0.988	0.94936
	5	1.000	-	0.986	**0.84277**	0.982	-
	6	1.000	-	0.988	**0.76826**	0.990	0.77271
	7	1.000	-	0.992	**0.71138**	0.992	0.71488
	8	1.000	-	0.992	**0.66551**	0.991	0.66780
50	2	1.000	-	0.991	**1.02992**	0.988	1.04050
	3	0.999	-	0.986	**0.84337**	0.987	0.84892
	4	1.000	-	0.986	**0.72933**	0.985	0.73226
	5	1.000	-	0.983	-	0.984	**0.65663**
	6	1.000	-	0.987	**0.59457**	0.986	0.59615
	7	0.999	-	0.991	**0.55074**	0.988	0.55173
	8	1.000	-	0.992	**0.51532**	0.990	0.51659
70	2	1.000	-	0.990	**0.87046**	0.988	0.87667
	3	1.000	-	0.988	**0.71226**	0.986	0.71556
	4	1.000	-	0.988	**0.61641**	0.986	0.61878
	5	1.000	-	0.986	**0.55114**	0.987	0.55311
	6	1.000	-	0.989	**0.50254**	0.989	0.50330
	7	1.000	-	0.981	-	0.981	-
	8	1.000	-	0.990	0.43546	0.992	**0.43543**

จากตารางที่ 2 พบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงแกมมา (2,1) ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่แคบที่สุดเกือบทุกสถานการณ์ ยกเว้นกรณีที่ $n = 50$ เมื่อ $\alpha = 5$ และ $n = 70$ เมื่อ $\alpha = 7$ และวิธีมาร์คอฟ เชน มอนติคาร์โล ให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ที่แคบที่สุดที่ $n = 50$ เมื่อ $\alpha = 5$ และ $n = 70$ เมื่อ $\alpha = 8$ และวิธีภาวะน่าจะเป็นสูงสุด เป็นวิธีที่ค่าประมาณสัมประสิทธิ์ความเชื่อมั่น (CC) ไม่อยู่ในเกณฑ์ที่กำหนดในทุกสถานการณ์

สรุปและอภิปรายผล

ในงานวิจัยนี้เป็นการเปรียบเทียบประสิทธิภาพของการประมาณช่วงความเชื่อมั่นของพารามิเตอร์ในการแจกแจงแกมมาที่ได้จากวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงแกมมา (2,1) และวิธีมาร์คอฟ เชน มอนติคาร์โล โดยพิจารณาวิธีการประมาณที่มีประสิทธิภาพดีที่สุดจากค่าสัมประสิทธิ์ความเชื่อมั่นและค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น ซึ่งแยกตามระดับความเชื่อมั่นที่กำหนดไว้ ผลการวิจัยพบว่า วิธีของเบส์ที่มีการแจกแจงก่อนเป็นการแจกแจงแกมมา (2,1) เป็นวิธีที่มีประสิทธิภาพดีที่สุดในการประมาณช่วงความเชื่อมั่น เพราะว่าเป็นวิธีที่นำข้อมูลที่เกี่ยวกับพารามิเตอร์มาใช้ให้เป็นประโยชน์ จึงทำให้การประมาณช่วงความเชื่อมั่นดีกว่าวิธีการอื่นๆ สำหรับวิธีมาร์คอฟ เชน มอนติคาร์โล ให้ผลการประมาณช่วงความเชื่อมั่นดีในบางกรณี ได้แก่ กรณีที่ $n=50$ เมื่อ $\alpha=5$ และ $n=70$ เมื่อ $\alpha=8$ นอกจากนี้เมื่อค่าพารามิเตอร์รูปร่าง (α) มีค่าเพิ่มขึ้นทำให้ความแปรปรวนเพิ่มขึ้น ซึ่งส่งผลให้ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่นลดลงในทุกสถานการณ์ เมื่อขนาดตัวอย่างเพิ่มมากขึ้นจะพบว่า ค่าความกว้างเฉลี่ยของช่วงความเชื่อมั่น (AW) ของวิธีของเบส์จะให้ค่าใกล้เคียงกับวิธีมาร์คอฟ เชน มอนติคาร์โล ดังนั้นถ้าผู้ใช้ไม่ทราบข้อมูลเกี่ยวกับพารามิเตอร์ ก็สามารถนำวิธีมาร์คอฟ เชน มอนติคาร์โล มาใช้ได้เช่นกัน เช่นเดียวกันกับงานวิจัยของ Kummaraka และ Araveeporn [10] ได้เปรียบเทียบช่วงความเชื่อมั่นของพารามิเตอร์ p ของการแจกแจงทวินามเชิงลบ (Negative Binomial Distribution) ด้วยวิธีภาวะน่าจะเป็นสูงสุด วิธีของเบส์และวิธีมาร์คอฟ เชน มอนติคาร์โล ซึ่งแสดงให้เห็นว่าวิธีมาร์คอฟ เชน มอนติคาร์โลมีประสิทธิภาพในการประมาณช่วงความเชื่อมั่นได้ดีเทียบเท่ากับวิธีของเบส์ ดังนั้นถ้าผู้วิจัยไม่ทราบค่าพารามิเตอร์ก็สามารถใช้วิธีมาร์คอฟ เชน มอนติคาร์โลได้เช่นกัน

เอกสารอ้างอิง

[1] Prachoom Suwattee. (2010). *Theory of statistical inference*. 3th ed. Bangkok: WVO officer of Printing Mill.

[2] Dagpunar S John. (2007). *Simulation and Monte Carlo : with applications in finance and MCMC*. Chichester: John Wiley & Sons Ltd.

[3] Gamerman Dani. (1997). *Markov Chain Monte Carlo : stochastic simulation for Bayesian inference.* London: Chapman & Hall.

[4] Araveeporn, A. (2015). The Interval Estimation of Poisson Distribution by Using Maximum Likelihood, Markov Chain Monte Carlo, and Bayes methods. In *International Conference on Applied Statistics 2015.* Pattaya: pp. 36-42.

[5] Sornsawan Boonpen, Boonorm Chomtee; and Apinya Hirunwong. (2015, October-December). A Comparison of Intervals Estimation Methods for Scale Parameter of the Two-Parameter Weibull Distribution. *Thai Science and Technology Journal Thammasat University.* 23(4): 579-587.

[6] Pradhan, Biswabrata; and Kundu, Debasis. (2011, September). Bayes estimation and prediction of the two-parameter gamma distribution. *Statistical Computation and Simulation*. 81(9): 1187-1198.

[7] Autcha Araveeporn. (2012). Bayesian Analysis from WinBUGS Program to R Program. *NU Science Journal.* 9(1): 30-44.

[8] Chatwadee Kitkeaw. (2017). *Efficiency Comparison of Parameter Estimation of Gamma Distribution by Maximum Likelihood, Bayes', and Markov Chain Monte Carlo Methods.* Master of Science (Applied Statistics). Bangkok: Department of Statistics Faculty of Science KMITL.

[9] Saichon Sinsomboonthong. (2011). *Mathematical Statistics 1.* 5th ed. Bangkok: Chamchuri Products.

[10] Kummaraka, Unyamanee; and Araveeporn, Autcha. (2018, January-June). A Comparison of Confidence Intervals of Negative Binomial Parameter p by Maximum Likelihood, Bayesian and Markov Chain Monte Carlo Methods. *Journal of Applied Statistics and Information Technology.* 3(1): 8-17.