การเปรียบเทียบการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีของโมเมนต์สำหรับการแจกแจงแชงเกอร์

A Comparison of Parameter Estimation Using Maximum Likelihood Estimation and Method of Moment for Shanker Distribution

ศศิชา นามเสนา[1,*] และ สุธาสินี ไทยส่วย[1]
Sasicha Namsena[1,*] and Suthasinee Thaisuay[1]
[1]สาขาวิชาวิทยาศาสตร์และวิทยาการคอมพิวเตอร์, คณะวิทยาศาสตร์และเทคโนโลยี, มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี, ปทุมธานี, 12120, ประเทศไทย
[1]Branch of applied statistics, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Pathum Thani, 12120, Thailand
***Corresponding Author E-mail:** sasichaaaa@gmail.com

รับบทความ (Received) : May 25, 2025 /ปรับแก้ไข (Revised) : June 2, 2025 /ตอบรับบทความ (Accepted) : June 17, 2025

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์ เพื่อศึกษาและเปรียบเทียบประสิทธิภาพในการประมาณค่าพารามิเตอร์ของการแจกแจงแชงเกอร์ระหว่างวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation : MLE) และวิธีของโมเมนต์ (Method of Moment : MOM) การศึกษาโดยการจำลองข้อมูลเมื่อกำหนด θ = 0.3 0.5 1.0 1.5 และ 2.5 ขนาดตัวอย่างที่ใช้ในการศึกษาได้แก่ n= 20 40 60 80 และ 100 นอกจากนี้มีการนำเสนอการประมาณค่าพารามิเตอร์ทั้งสองวิธีกับข้อมูลจริงและการทดสอบภาวะสารูปดี ผลการศึกษาจากการจำลองข้อมูลพบว่า วิธีภาวะน่าจะเป็นสูงสุด (MLE) และ วิธีของโมเมนต์ (MOM) มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ของการแจกแจงแชงเกอร์เนื่องจากให้ค่าประมาณ $\hat{\theta}$ เข้าใกล้ค่าจริง θ เมื่อขนาดตัวอย่างเพิ่มขึ้น เมื่อพิจารณาจากค่าสัมบูรณ์ของ BIAS พบว่า วิธี MLE ให้ค่าสัมบูรณ์ของ BIAS น้อยกว่าวิธี MOM ในทุกกรณี ยกเว้นเมื่อ θ = 0.3 และ 0.5 (θ<1) ขนาดตัวอย่าง n= 80 และ 100 ซึ่งวิธี MOM ให้ค่าสัมบูรณ์ของ BIAS น้อยกว่าวิธี MLE อย่างไรก็ตามวิธี MOM เป็นวิธีที่มีค่าความแปรปรวนน้อยกว่าวิธี MLE จากการเปรียบเทียบการประมาณค่าพารามิเตอร์ทั้งสองวิธีกับข้อมูลจริง 3 ชุด พบว่า วิธี MLE ให้ค่าประมาณที่มีความสอดคล้องและเหมาะสมกับข้อมูลจริงมากกว่าวิธี MOM เนื่องจากวิธี MLE ให้ค่า K-S test ที่มีค่า p-value>0.05 และมีค่ามากกว่าวิธี MOM ส่วนวิธี MOM ให้จำนวนรอบในการลู่เข้าหาคำตอบได้เร็วกว่าวิธี MLE

คำสำคัญ : การแจกแจงแชงเกอร์, วิธีภาวะน่าจะเป็นสูงสุด, วิธีของโมเมนต์

Abtract

The research aims to study and compare the efficiency of estimating the parameters of the Shanker distribution between the maximum likelihood estimation (MLE) and the method of moment (MOM) methods by simulating the data when setting = 0.3, 0.5, 1.0, 1.5 and 2.5, respectively. The sample sizes used in the study were 20, 40, 60, 80 and 100. In addition, both parameter estimation methods are applied for real data with goodness-of-fit tests presented. The results of the simulation study showed that the maximum likelihood method (MLE) and the method of moments (MOM) were effective in estimating the parameters of the Shanker distribution because they gave an estimate that approached the true value as the sample size increased. Considering the absolute value of BIAS, it is found that the MLE method gives a lower absolute value of BIAS than the MOM method, except for = 0.3 and 0.5 (< 1) and n = 80 and 100, where the MOM method gives a lower absolute value of BIAS than the MLE method. However, the MOM method has a lower variance than the MLE method. From the comparison of the parameter estimation of both methods with the three real datasets, it is found that the MLE method gives estimates that are more consistent with the real datasets than the MOM method. On the other hand, the MOM method gives a faster number of rounds to converge to the answer than the MLE method.

Keywords : Shanker distribution, Maximum Likelihood Estimation, Method of Moments

บทนำ

การแจกแจงช่วงชีวิต (lifetime distribution) คือ การแจกแจงที่ใช้ในการอธิบายและวิเคราะห์ข้อมูล โดยเฉพาะข้อมูลที่เกี่ยวข้องกับช่วงอายุการใช้งาน เช่น อายุการใช้งานของอุปกรณ์ เครื่องจักร การรอดชีวิตในทางการแพทย์ ฯลฯ การแจกแจงช่วงชีวิตที่นิยมใช้ในงานประเภทนี้มีหลากหลายรูปแบบ เช่น การแจกแจงเลขชี้กำลัง (Exponential Distribution) และการแจกแจงไวบูลล์ (Weibull Distribution) แต่ในบางกรณีการแจกแจงเหล่านี้ไม่สามารถอธิบายข้อมูลช่วงชีวิตที่เก็บรวบรวมได้อย่างเหมาะสม เนื่องจากข้อจำกัดของตัวแบบและพารามิเตอร์ที่อาจไม่สอดคล้องกับข้อมูลจริงจึงจำเป็นต้องมีการพิจารณาการแจกแจงแบบใหม่ที่สามารถอธิบายลักษณะของข้อมูลได้ดียิ่งขึ้นการแจกแจงแชงเกอร์ (Shanker Distribution) เป็นหนึ่งในตัวเลือกที่น่าสนใจ เนื่องจากมีความยืดหยุ่นและสามารถใช้อธิบายข้อมูลที่มีการกระจายแบบเบ้หรือไม่สมมาตรได้ดี การแจกแจงนี้ยังมีรูปแบบที่ค่อนข้างง่ายต่อการคำนวณและการนำไปใช้งาน โดยเฉพาะในกรณีที่ข้อมูลมีขนาดตัวอย่างเล็ก หรือข้อมูลที่ไม่เป็นไปตามการแจกแจงแบบปกติ (non-normal data) ด้วยคุณสมบัตินี้ การแจกแจงแชงเกอร์จึงได้รับความสนใจในการนำมาใช้อธิบายข้อมูลช่วงชีวิตที่หลากหลาย [1] การประมาณค่าพารามิเตอร์เป็นขั้นตอนสำคัญในการประยุกต์ใช้การแจกแจงแชงเกอร์ สำหรับการสร้างตัวแบบอธิบายข้อมูลช่วงชีวิต โดยวิธีที่นิยมใช้ได้แก่ วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: MLE) และวิธีของโมเมนต์ (Method of Moments: MOM) ซึ่งทั้งสองวิธีมีข้อดีและข้อจำกัดที่แตกต่างกัน วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: MLE) มีข้อได้เปรียบในเรื่องของความแม่นยำและประสิทธิภาพเมื่อขนาดตัวอย่างมีขนาดใหญ่ ในขณะที่วิธีของโมเมนต์ (Method of Moments: MOM) มีขั้นตอนการคำนวณที่ง่ายกว่าและสามารถใช้ได้ดีในกรณีที่ขนาดตัวอย่างมีขนาดเล็ก [2] โดยงานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation: MLE) และวิธีของโมเมนต์ (Method of Moments: MOM) ในการประมาณค่าพารามิเตอร์ของการแจกแจงแชงเกอร์ โดยพิจารณาความแม่นยำในกรณีที่ขนาดตัวอย่างแตกต่างกัน ผลการศึกษานี้คาดว่าจะช่วยให้เกิดความเข้าใจที่ชัดเจนยิ่งขึ้นเกี่ยวกับการเลือกใช้วิธีการประมาณค่าพารามิเตอร์ที่เหมาะสมในสถานการณ์ต่าง ๆ และเป็นแนวทางสำหรับนักวิจัยในการนำไปประยุกต์ใช้ในงานวิจัยหรือการวิเคราะห์ข้อมูลต่อไป

1. วัตถุประสงค์การวิจัย

1.1 เพื่อศึกษาประสิทธิภาพในการประมาณค่าพารามิเตอร์ของการแจกแจงแชงเกอร์โดยใช้วิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation : MLE) และวิธีของโมเมนต์ (Method of Moment : MOM)

1.2 เพื่อเปรียบเทียบประสิทธิภาพการประมาณค่าระหว่างวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation : MLE) และวิธีของโมเมนต์ (Method of Moment : MOM)

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 การแจกแจงแชงเกอร์ (Shanker Distribution) เป็นการแจกแจงแบบต่อเนื่องหลายแบบสำหรับสร้างข้อมูลช่วงชีวิต

ฟังก์ชันความหนาแน่นของความน่าจะเป็น (Probability Density Function : PDF) คือ ฟังก์ชันที่อธิบายความหนาแน่นของความน่าจะเป็นของเหตุการณ์ที่ใช้ค่าที่เป็นตัวแปรสุ่มแบบต่อเนื่อง (Continuous Random Variable) PDF ของการแจกแจงแชงเกอร์เกิดจากการผสมกันระหว่างการแจกแจงเลขชี้กำลัง(Exponential Distribution) ด้วยค่าพารามิเตอร์ θ และการแจกแจงแกมมา (Gamma Distribution) ด้วยค่าพารามิเตอร์ $(2,\theta)$ [3]

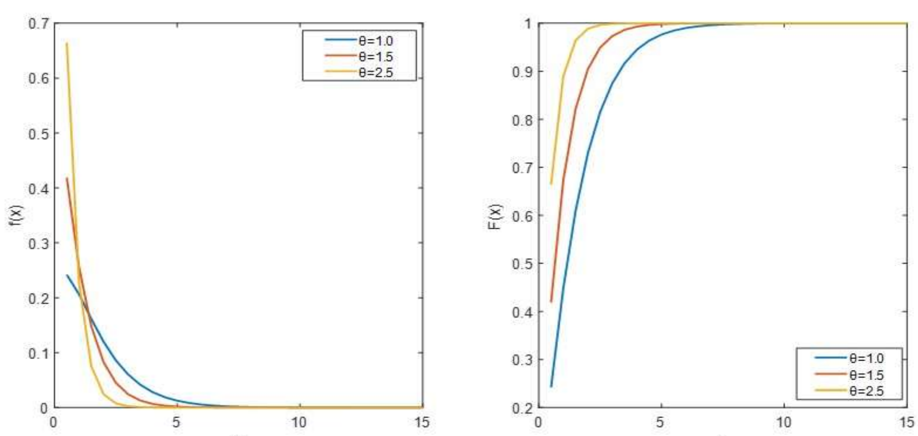
ฟังก์ชันการแจกแจงความน่าจะเป็น แสดงดังสมการที่ (1)

$$f(x;\theta)=\frac{\theta^2(\theta+x)}{(\theta^2+1)}e^{-\theta x} \quad ; \quad x>0,\theta>0 \tag{1}$$

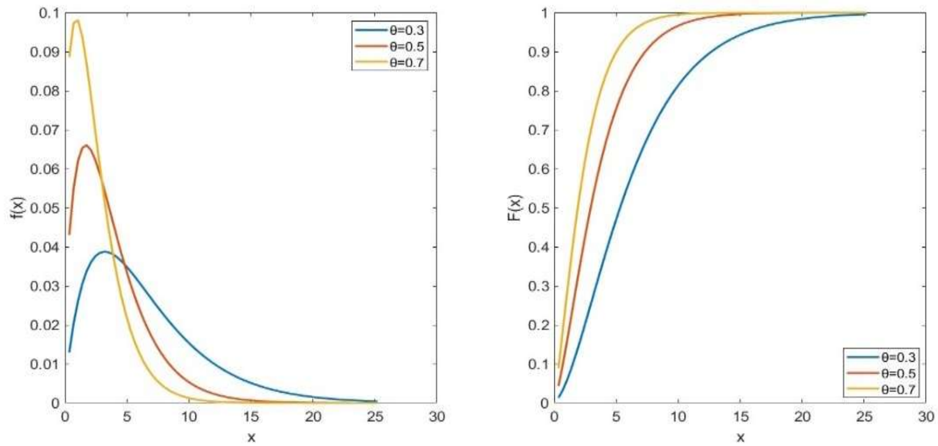
ฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function : CDF) คือ ฟังก์ชันของความน่าจะเป็น ที่ค่าของตัวแปรสุ่มจำนวนจริง x น้อยกว่าหรือเท่ากับ x สำหรับตัวแปรสุ่มแบบต่อเนื่อง [3] ฟังก์ชันการแจกแจงสะสม แสดงดังสมการที่ (2)

$$F(x;\theta)=1-\frac{(\theta^2+1)+\theta x}{\theta^2+1}e^{-\theta x} \quad ; \quad x>0,\theta>0 \tag{2}$$

กราฟ PDF และ CDF ของการแจกแจงแชงเกอร์แสดงดังรูปที่ 1 และ รูปที่ 2



รูปที่ 1 PDF และ CDF ของการแจกแจงแชงเกอร์เมื่อ $\theta \geq 1$



รูปที่ 2 PDF และ CDF ของการแจกแจงแชงเกอร์เมื่อ $\theta < 1$

3) โมเมนต์ (Moments of Associated Measures) โมเมนต์รอบจุดกำเนิดของการแจกแจงแชงเกอร์แสดงดังสมการที่ (3)

$$\mu_r' = \frac{r!(\theta^2 + r + 1)}{\theta^r(\theta^2 + 1)} \tag{3}$$

การคำนวณโมเมนต์เหล่านี้จะใช้ในการศึกษาคุณสมบัติทางสถิติของการแจกแจง เช่น ความคาดหวังและความแปรปรวนของข้อมูล ซึ่งโมเมนต์รอบที่ 1 แสดงดังสมการที่ (4)

$$\mu_1' = \frac{\theta^2 + 2}{\theta(\theta^2 + 1)} \tag{4}$$

2.2 การประมาณค่าพารามิเตอร์

2.2.1) วิธีภาวะน่าจะเป็นสูงสุด (MLE)

ให้ $\tilde{x}_i = x_1, x_2, \ldots, x_n$ เป็นตัวแปรสุ่มขนาด n ที่มีการแจกแจง SD การประมาณค่าพารามิเตอร์ทำได้ดังนี้ กำหนดให้ L คือ likelihood function ของ SD แสดงดังสมการ (5) [3]

$$\begin{aligned} L &= \prod_{i=1}^{n}\left(\left(\frac{\theta^2}{\theta^2+1}\right)(\theta + x_i)e^{-\theta X_i}\right) \\ &= \left(\frac{\theta^2}{\theta^2+1}\right)^n \prod_{i=1}^{n}(\theta + x_i)e^{-\theta X_i} \end{aligned} \tag{5}$$

log – likelihood function SD เขียนด้วย $\ell(\theta|\tilde{x})$

$$\ell(\theta|\tilde{x}) = n\log\left(\frac{\theta^2}{\theta^2+1}\right) + \sum_{i=1}^{n}\log(\theta + x_i)e^{-\theta X_i}$$

$$= n\log\left(\frac{\theta^2}{\theta^2+1}\right) + \sum_{i=1}^{n}\log\left(\theta + x_i\right) - \theta\sum_{i=1}^{n} x_i$$

$$= n\log\left(\frac{\theta^2}{\theta^2+1}\right) + \sum_{i=1}^{n}\log\left(\theta + x_i\right) - \theta\bar{X}n$$

ดังนั้นการประมาณค่าพารามิเตอร์ ($\hat{\theta}$) คำนวณได้จากการแก้สมการอนุพันธ์เทียบ θ

$$\frac{d\left(\ell\left(\theta|\tilde{x}\right)\right)}{d\theta} = \frac{2n}{\theta\left(\theta^2+1\right)} + \sum_{i=1}^{n}\left(\frac{1}{\theta + x_i}\right) - n\bar{X} = 0$$

เมื่อ $\bar{X}$ คือค่าเฉลี่ยของตัวอย่างการประมาณค่าพารามิเตอร์ θ ทำได้โดย $\frac{d\left(\ell\left(\theta|\tilde{x}\right)\right)}{d\theta} = 0$ สมการดังกล่าวไม่อยู่ในรูปแบบปิด (closed from) จึงแก้สมการโดยใช้ระเบียบวิธีนิวตัน-ราฟสัน

$$\hat{\theta}_{n+1} = \hat{\theta}_n - \frac{f(\hat{\theta}_n)}{f'(\hat{\theta}_n)}$$

2.2.2) วิธีของโมเมนต์ (MOM)

โมเมนต์รอบที่ 1 ของ SD แสดงดังนี้ [3]

$$M_1' = E(x) = \frac{\theta^2 + 2}{\theta\left(\theta^2+1\right)}$$

ดังนั้น

$$\frac{\theta^2+2}{\theta(\theta^2+1)} = \bar{X}$$

$$\theta^2 + 2 = \bar{X}\left(\theta\left(\theta^2+1\right)\right)$$

$$\theta^2 + 2 = \bar{X}\theta^3 + \bar{X}\theta$$

$$\bar{X}\theta^3 + \bar{X}\theta - \theta^2 - 2 = 0 \quad (6)$$

คำตอบของ $\hat{\theta}$ มาจากการแก้สมการที่ (6) ซึ่งสมการดังกล่าวไม่อยู่ในรูปแบบปิด (closed from) การแก้สมการจึงใช้ระเบียบวิธีนิวตัน-ราฟสัน เช่นเดียวกับวิธี MLE

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีของโมเมนต์สำหรับการแจกแจงแชงเกอร์ ซึ่งมีวิธีการดำเนินงานวิจัยออกเป็น 5 ขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับ SD, MLE และ MOM

ขั้นตอนที่ 2 สร้างข้อมูลให้มีการแจกแจงแชงเกอร์โดยใช้ควอนไทล์ฟังก์ชันของการแจกแจงแชงเกอร์ แสดงรายละเอียดดังนี้

การจำลองข้อมูลโดยควอนไทล์ฟังก์ชัน (Quantile Function ; Q(u))

กำหนดให้ $X \sim SD$ ดังนั้นควอนไทล์ฟังก์ชัน $Q(u) = F^{-1}(x)$

ให้ $u = 1 - \frac{(\theta^2+1)+\theta x}{\theta^2+1}e^{-\theta x}$

$[(\theta^2+1)+\theta x]e^{-\theta x} = (1-u)(\theta^2+1)$ [จาก Lambert W Function, $W(Z)\cdot e^{W(Z)} = Z$ [4]]

จะได้ว่า
$$x=\frac{-1}{\theta}[W_{-1}[(u-1)(\theta^2+1)e^{-(\theta^2+1)}]+(\theta^2+1)]$$

การสร้างตัวแปรสุ่มที่มีการแจกแจง SD

ให้ $x_i=F^{-1}(u_i)$ ค่าของ $x_i, i=1,2,...,n$ สามารถสร้างตัวแปรสุ่มดังขั้นตอนต่อไปนี้

ขั้นที่ 1 ให้ $u_i; i=1,2,...,n \sim uniform(0,1)$

ขั้นที่ 2 คำนวณค่า $W_i=W_{-1}[(u_i-1)(\theta^2+1)e^{-(\theta^2+1)}]$ โดยใช้ Lambert W Function

ขั้นที่ 3 คำนวณตัวแปรสุ่ม $x_i=\frac{-1}{\theta}[W_i+(\theta^2+1)]$

โดยในการศึกษาครั้งนี้จะกำหนดขนาดตัวอย่างในการศึกษา n= 20, 40, 60, 80 และ 100 ตามลำดับ และจะศึกษาโดยการกำหนดค่าพารามิเตอร์ $\theta<1$ และ $\theta\geq 1$ และประมาณค่าพารามิเตอร์ด้วยวิธี MLE และ MOM ในหัวข้อที่ 2.2.1 และ 2.2.2

ขั้นตอนที่ 4 เปรียบเทียบการประมาณ $\hat{\theta}$ ระหว่างวิธี MLE และวิธี MOM ด้วยค่า $|BIAS|$ MSE และ Variance ดังสมการที่ (7) และ (8)

$$|BIAS|=|E(\hat{\theta})-\theta| \tag{7}$$

$$MSE(\hat{\theta})=(BIAS(\hat{\theta},\theta))^2+Variance(\hat{\theta}) \tag{8}$$

ขั้นตอนที่ 5 ประยุกต์การประมาณค่า $\hat{\theta}$ ด้วยวิธี MLE และ MOM กับข้อมูลจริง พร้อมเปรียบเทียบประสิทธิภาพด้วยการทดสอบภาวะสารูปดี ด้วยค่าการทดสอบคอลโมโกรอฟ-สมีร์นอฟ หรือ K-S test

ผลการวิเคราะห์ข้อมูลและอภิปรายผล

ผลการวิเคราะห์ข้อมูลแบ่งออกเป็นสองส่วน ได้แก่ การแปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ด้วยวิธี MLE และวิธี MOM จากการจำลองข้อมูลและการประยุกต์การประมาณค่าพารามิเตอร์ทั้ง 2 วิธีกับข้อมูลจริง ซึ่งรายละเอียดแสดงดังนี้

1. **การเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ด้วยวิธี MLE และวิธี MOM จากการจำลองข้อมูล**

เมื่อกำหนดค่าพารามิเตอร์ $\theta<1$ $\theta=1$ และ $\theta>1$ โดยจะทำการจำลองข้อมูลขนาด $n=20, 40, 60, 80$ และ 100 และมีการทำซ้ำจำนวน 500 รอบ ผลการศึกษาพบว่า MLE และ MOM มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ของการแจกแจงแชงเกอร์เนื่องจากมีค่าประมาณ $\hat{\theta}$ ใกล้เคียงกับค่าพารามิเตอร์ θ เมื่อ $n\to\infty$ อย่างไรก็ตามการประมาณ $\hat{\theta}$ ของทั้ง 2 วิธีมีความแตกต่างกันเล็กน้อย ดังตารางที่ 1 ถึง 5 หากพิจารณาเป็นรายกรณีพบว่ากรณี $\theta=0.3$ และ $\theta=0.5(\theta<1)$ วิธี MLE จะให้ค่าประมาณ $\hat{\theta}$ ที่มีค่า $|BIAS|$ ต่ำกว่าวิธี MOM เมื่อ n=20 40 และ 60 ในทางกลับกัน เมื่อ n=80 และ 100 วิธี MOM จะมี

ค่า $|BIAS|$ ต่ำกว่า MLE ส่วนในกรณี $\theta = 1.0$ 1.5 และ 2.5 $(\theta \geq 1)$ พบว่าวิธี MLE ให้ $|BIAS|$ ต่ำกว่า MOM ทุกๆค่าขนาดตัวอย่าง อย่างไรก็ตามวิธี MOM นั้นจะให้ค่า MSE ต่ำกว่าวิธี MLE ในทุกกรณี ยกเว้นกรณี $\theta = 0.3$ และ n =20 40 และ 60 ซึ่งวิธี MOM จะมีค่า MSE สูงกว่าเล็กน้อย เมื่อพิจารณาค่าความแปรปรวน (Variance) ของค่าประมาณพารามิเตอร์ จะยกตัวอย่างการนำเสนอ ดังรูปที่ 3 กรณี $\theta = 0.3(\theta < 1)$ รูปที่ 4 กรณี $\theta = 1$ และรูปที่ 5 กรณี $\theta = 2.5(\theta > 1)$ พบว่าวิธี MOM เป็นตัวประมาณที่มีค่าความแปรปรวนต่ำกว่าวิธี MLE ซึ่งกราฟเปรียบเทียบการประมาณค่าพารามิเตอร์ระหว่างวิธี MLE และ MOM ด้วยค่าความแปรปรวน (Variance) แสดงได้ดังรูปที่ 3 ถึง รูปที่ 5

ตารางที่ 1 ค่าประมาณพารามิเตอร์ด้วยวิธี MLE และวิธี MOM เมื่อ $\theta = 0.3$ กรณี $\theta < 1$

n	MLE			MOM		
	$\hat{\theta}$	$\|BIAS\|$	MSE	$\hat{\theta}$	$\|BIAS\|$	MSE
20	0.4585	0.1585	0.0269	0.4670	0.1670	0.0284
40	0.2478	0.0521	0.0029	0.2451	0.0548	0.0030
60	0.3387	0.0387	0.0017	0.3436	0.0436	0.0019
80	0.2653	0.0346	0.0013	**0.2698**	**0.0301**	**0.0009**
100	0.2775	0.0224	0.0006	**0.2844**	**0.0155**	**0.0002**

ตารางที่ 2 ค่าประมาณพารามิเตอร์ด้วยวิธี MLE และวิธี MOM เมื่อ $\theta = 0.5$ กรณี $\theta < 1$

n	MLE			MOM		
	$\hat{\theta}$	$\|BIAS\|$	MSE	$\hat{\theta}$	$\|BIAS\|$	MSE
20	0.6265	0.1265	0.0204	0.6379	0.1379	0.0197
40	0.5647	0.0648	0.0064	0.5729	0.0729	0.0054
60	0.4685	0.0315	0.0023	0.4659	0.0346	0.0012
80	0.5074	0.0074	0.0018	**0.5060**	**0.0060**	**0.00005**
100	0.5020	0.0020	0.0015	**0.5007**	**0.0007**	**0.00001**

ตารางที่ 3 ค่าประมาณพารามิเตอร์ด้วยวิธี MLE และวิธี MOM เมื่อ $\theta = 1$

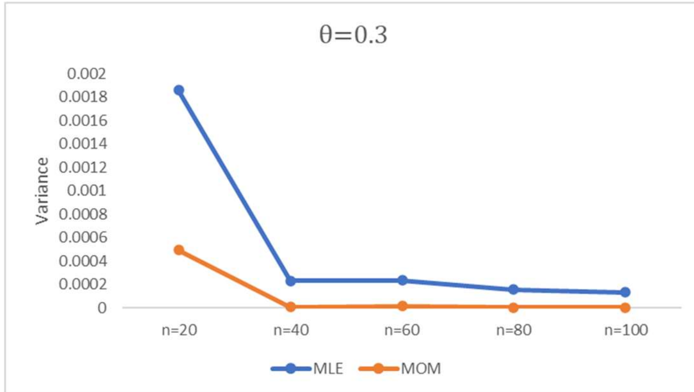
n	MLE			MOM		
	$\hat{\theta}$	$\lvert BIAS \rvert$	MSE	$\hat{\theta}$	$\lvert BIAS \rvert$	MSE
20	1.2319	0.2319	0.0790	1.2637	0.2637	0.0722
40	1.1200	0.1200	0.0301	1.1400	0.1400	0.0200
60	1.0642	0.0642	0.0163	1.0709	0.0709	0.0051
80	0.9828	0.0172	0.0098	0.9740	0.0260	0.00072
100	1.0137	0.0137	0.0072	1.0256	0.0256	0.00071

ตารางที่ 4 ค่าประมาณพารามิเตอร์ด้วยวิธี MLE และวิธี MOM เมื่อ $\theta = 1.5$ กรณี $\theta > 1$

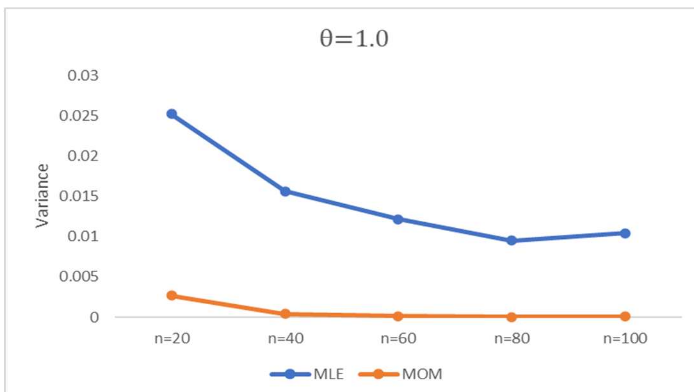
n	MLE			MOM		
	$\hat{\theta}$	$\lvert BIAS \rvert$	MSE	$\hat{\theta}$	$\lvert BIAS \rvert$	MSE
20	1.3210	0.1790	0.0522	1.3009	0.1990	0.0408
40	1.5438	0.0438	0.0307	1.5502	0.0502	0.0034
60	1.4815	0.0185	0.0225	1.4666	0.0334	0.0013
80	1.4960	0.0040	0.0211	1.4870	0.0130	0.0003
100	1.5021	0.0021	0.0184	1.5111	0.0111	0.0002

ตารางที่ 5 ค่าประมาณพารามิเตอร์ด้วยวิธี MLE และวิธี MOM เมื่อ $\theta = 2.5$ กรณี $\theta > 1$

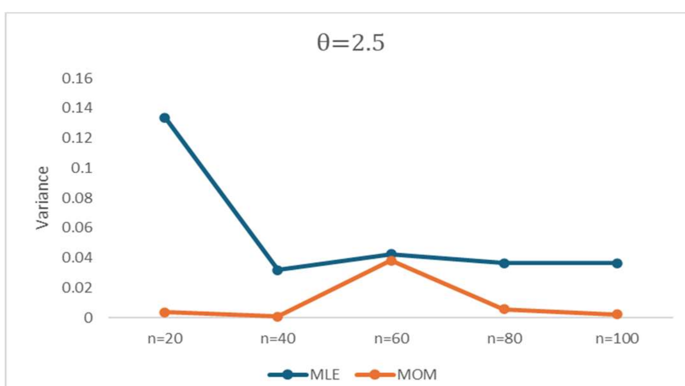
n	MLE			MOM		
	$\hat{\theta}$	$\lvert BIAS \rvert$	MSE	$\hat{\theta}$	$\lvert BIAS \rvert$	MSE
20	2.0415	0.4585	0.3441	2.0259	0.4741	0.2284
40	2.0902	0.4098	0.1997	2.0606	0.4394	0.1939
60	2.3968	0.1032	0.0531	2.8500	0.3500	0.1607
80	2.4338	0.0662	0.0409	2.6846	0.1846	0.0397
100	2.5084	0.0084	0.0366	2.6302	0.1302	0.0192



รูปที่ 3 ความแปรปรวนของ MLE และ MOM เมื่อ $\theta = 0.3$



รูปที่ 4 ความแปรปรวนของ MLE และ MOM เมื่อ $\theta = 1.0$



รูปที่ 5 ความแปรปรวนของ MLE และ MOM เมื่อ $\theta = 2.5$

2. การประยุกต์การประมาณค่าพารามิเตอร์ด้วยวิธี MLE และวิธี MOM กับข้อมูลจริง

ในส่วนนี้ได้ดำเนินการศึกษาโดยจะนำเสนอการเปรียบเทียบการประมาณค่าพารามิเตอร์ด้วย MLE กับ MOM กับข้อมูลจริง จำนวน 3 ชุดพร้อมการตรวจสอบภาวะสารูปดีด้วยค่า K-S test ดังสมการที่ (9)

$$K-S=\sup_{x}\left|F_n(x)-F_0(x)\right| \tag{9}$$

เมื่อ $F_0(x)$ คือ ฟังก์ชันการแจกแจงสะสมของตัวอย่าง

$F_n(x)$ คือ ฟังก์ชันการแจกแจงสะสมของประชากรที่คาดไว้

n คือ จำนวนตัวอย่าง

p คือ จำนวนพารามิเตอร์

ซึ่งหากค่า K-S test มีค่าต่ำและ p -value > 0.05 แสดงว่าข้อมูลจริงมีความสอดคล้องกับการแจกแจงที่ต้องการทดสอบ รายละเอียดของข้อมูลจริงทั้ง 3 ชุดที่ใช้ในการศึกษา ได้แก่

ข้อมูลชุดที่1 มีการศึกษาตลับลูกปืน (ball bearings) จำนวน 23 ชิ้น โดยวัดว่ามันหมุนได้กี่ล้านรอบก่อนจะเสีย (failure) ข้อมูลนี้ถูกนำมาใช้เป็นตัวอย่างเพื่ออธิบายแนวคิด GASIP รายละเอียดของข้อมูลดังนี้ [5] 17.88 28.92 33 41.52 42.12 45.60 48.40 51.84 51.96 54.12 55.56 67.80 68.64 68.64 68.88 84.12 93.12 98.64 105.12 105.84 127.92 128.04 173.40

ข้อมูลชุดที่2 เวลาที่เกิดความล้มเหลวของกระจกหน้าของเครื่องบิน 63 บาน ซึ่งข้อมูลถูกวัดเป็นหน่วยพันชั่วโมง เพื่อให้สามารถอ้างอิงได้ง่าย ข้อมูลมีดังนี้ [6] 0.046 1.436 2.592 0.140 1.492 2.600 0.150 1.580 2.670 0.248 1.719 2.717 0.280 1.794 2.819 0.313 1.915 2.820 0.389 1.920 2.878 0.487 1.963 2.950 0.622 1.978 3.003 0.900 2.053 3.102 0.952 2.065 3.304 0.996 2.117 3.483 1.003 2.137 3.500 1.010 2.141 3.622 1.085 2.163 3.665 1.092 2.183 3.695 1.152 2.240 4.015 1.183 2.341 4.628 1.244 2.435 4.806 1.249 2.464 4.881 1.262 2.543 5.140

ข้อมูลชุดที่3 ชุดข้อมูลจากการทดสอบอายุการใช้งานของไดโอดเปล่งแสง (LED) มีข้อมูลตัวอย่าง 58 ตัวอย่าง ที่มีระดับการเซ็นเซอร์ 0.25 เพื่อความเรียบง่าย เราจึงแยกตัวอย่างความล้มเหลวที่สังเกตได้ภายใต้เงื่อนไขการใช้งานหรือความเค้นเป็นชุดข้อมูล PALT ที่มีความเค้นคงที่ทั้งหมด ซึ่งแสดงข้อมูลดังนี้ [7] 0.18, 0.19, 0.19, 0.34, 0.36, 0.40 0.44, 0.44, 0.45, 0.46, 0.47, 0.53, 0.57, 0.57, 0.63, 0.65, 0.70, 0.71, 0.71, 0.75, 0.76, 0.76, 0.79, 0.80, 0.85, 0.98, 1.01, 1.07, 1.12, 1.14, 1.15, 1.17, 1.20, 1.23, 1.24, 1.25, 1.26, 1.32, 1.33, 1.33, 1.39, 1.42, 1.50, 1.55, 1.58, 1.59, 1.62, 1.68, 1.70, 1.79, 2.00, 2.01, 2.04, 2.54, 3.61, 3.76, 4.65, 8.97

ตารางที่ 6 ค่าประมาณพารามิเตอร์ด้วยวิธี MLE วิธี MOM และ K-S test ของข้อมูลจริง

ข้อมูลชุดที่	MLE			MOM		
	$\hat{\theta}$	จำนวนรอบในการลู่เข้าหาคำตอบ	K-S (p-value)	$\hat{\theta}$	จำนวนรอบในการลู่เข้าหาคำตอบ	K-S (p-value)
1 (n=23)	0.02766	5	0.1885 (0.3432)	0.02768	3	0.1888 (0.3414)
2 (n=63)	0.76662	7	0.1439 (0.1331)	0.77822	5	0.1529 (0.0945)
3 (n=58)	1.08561	8	0.1403 (0.1850)	1.09665	6	0.1441 (0.1628)

จากตารางที่ 6 เมื่อพิจารณาจากข้อมูลชุดที่ 1 ถึง ชุดที่ 3 พบว่าทั้งวิธี MLE และวิธี MOM มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ของ SD เนื่องจากทั้งสองวิธีให้ค่า K-S test ต่ำและให้ค่า p-value > 0.05 แสดงว่าข้อมูลจริงในแต่ละชุดดังกล่าวมีความสอดคล้องกับ SD แต่อย่างไรก็ตามวิธี MLE จัดว่าเป็นวิธีการประมาณค่าพามิเตอร์ที่มีค่า K-S test ที่ให้ค่า p-value $\geq$ 0.05 และสูงกว่าวิธี MOM ในทุกกรณี อย่างไรก็ตามวิธี MOM จะลู่เข้าหาคำตอบได้เร็วกว่าวิธี MLE

สรุปผลการวิจัย

การวิจัยนี้นำเสนอการศึกษาและการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีของโมเมนต์สำหรับการแจกแจงแชงเกอร์ โดยการจำลองข้อมูลให้มีการแจกแจงแชงเกอร์กำหนด θ = 0.3 0.5 1.0 1.5 และ 2.5 ทำการศึกษา กรณี n = 20 40 60 80 และ 100 ผลการศึกษาพบว่าวิธี MLE และวิธี MOM มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ของการแจกแจงแชงเกอร์เนื่องจากมีค่าประมาณ $\hat{\theta}$ ใกล้เคียงกับค่าพารามิเตอร์ θ เมื่อ $n \to \infty$ อย่างไรก็ตาม กรณี $\theta < 1$ วิธี MLE จะให้ค่าประมาณ $\hat{\theta}$ ที่มีค่า $|BIAS|$ ต่ำกว่าวิธี MOM เมื่อ n = 20,40 และ 60 ในทางกลับกันเมื่อ n = 80 และ 100 วิธี MOM จะมีค่า $|BIAS|$ ต่ำกว่าวิธี MLE กรณี $\theta \geq 1$ พบว่าวิธี MLE ให้ค่า $|BIAS|$ ต่ำกว่าวิธี MOM ทุกๆค่าขนาดตัวอย่าง อย่างไรก็ตามวิธี MOM นั้นจะให้ค่า Variance ต่ำกว่าวิธี MLE ในทุกกรณีและให้ค่า MSE ต่ำกว่าวิธี MLE ในเกือบทุกกรณี เมื่อพิจารณาการประยุกต์กับข้อมูลจริง 3 ชุด ผลการศึกษาพบว่าทั้งวิธี MLE และวิธี MOM มีประสิทธิภาพในการประมาณค่าพารามิเตอร์ของ SD เนื่องจากทั้งสองวิธี มีค่า K-S test ที่ให้ค่า p-value > 0.05 แต่อย่างไรก็ตาม MLE จัดว่าเป็นวิธีการประมาณค่าพารามิเตอร์ที่มีค่า K-S test ที่ให้ค่า p-value สูงกว่า วิธี MOM ในเกือบทุกกรณี แสดงว่า MLE ให้ค่าประมาณพารามิเตอร์ที่มีความสอดคล้องกับข้อมูลจริงมากกว่า MOM ส่วนวิธี MOM จะมีจำนวนรอบในการลู่เข้าหาคำตอบได้รวดเร็วกว่าวิธี MLE

ข้อเสนอแนะ

1) ควรมีการศึกษาเพิ่มเติมเกี่ยวกับผลกระทบของการแจกแจงข้อมูลที่แตกต่างกันต่อความแม่นยำของวิธีการประมาณค่าพารามิเตอร์

2) ในการศึกษาค้นคว้าต่อไปอาจมีการนำการประมาณค่าพารามิเตอร์ในรูปแบบอื่นๆ มาพิจารณาเปรียบเทียบประสิทธิภาพ เช่น วิธีการแปลงการลงจุดตำแหน่ง (Modified Plotting Position) และวิธีโมเมนต์ถ่วงน้ำหนักความน่าจะเป็น (Probability Weighted Moments)

เอกสารอ้างอิง

[1] Liang, J. and Fang, K. (2022). Mixture of Shanker Distributions: Estimation. Simulation and Application. Computational Statistics & Data Analysis, 12(3), 231.

[2] Tan, W. Y. and Chang, W. C. (1972). Some Comparisons of the Method of Moments and the Method of Maximum Likelihood in Estimating Parameters of a Mixture of Two Normal Densities. Journal of the American Statistical Association, 67(339), 702-708.

[3] Shanker, R. (2015). Shanker Distribution and Its Applications, International Journal of Statistics and Applications, 5(6), 338-348.

[4] Jorda, P. (2010). Computer generation of random variables with Lindley or Poisson–Lindley distribution via the Lambert W function. Mathematics and Computers in Simulation,81, 851-859.

[5] Lawless, J. F. (2003). Statistical Models and Methods for Lifetime Data, Wiley, New York.

[6] Tahir, M. H., Cordeiro, G. M., Mansoor, M. and Zubair, M. (2014). The Weibull-Lomax distribution: Properties and applications. Hacettepe Journal of Mathematics and Statistics, 44 (14), 1–474.

[7] Cheng, Y.F. and Wang, F.K. (2012). Estimating the Burr XII parameters in constant-stress partially accelerated life tests under multiple censored data. Communications in Statistics - Simulation and Computation, 41, 1711–1727.