ว.วิทย. มข. 42(3) 523-531 (2557) KKU Sci. J. 42(3) 523-531 (2014)

การประมาณค่าด้วยวิธีความควรจะเป็นสูงสุดสำหรับการแจกแจงเลขชี้กำลังเชิงเส้นด้วยข้อมูลแบ่งกลุ่ม

Maximum likelihood Estimation for the Linear Exponential Distribution Using Grouped Data

คณิศา โชติจันทึก[1]

บทคัดย่อ

การใช้ฟังก์ชันการแจกแจงทางสถิติ เพื่ออธิบายลักษณะข้อมูล เป็นกระบวนการทางวิทยาศาสตร์ที่ยอมรับกันทั่วโลก ซึ่งจะทำให้ทราบแนวโน้มและสารสนเทศของข้อมูล เพื่อการตัดสินใจเชิงประยุกต์ในด้านต่าง ๆ จึงจำเป็นที่จะต้องประมาณค่าพารามิเตอร์สำหรับฟังก์ชันการแจกแจงทางสถิตินั้น

อย่างไรก็ตามพบว่าการสังเกตลักษณะข้อมูลบางสถานการณ์ โดยเฉพาะสถานการณ์ที่สอดคล้องกับการแจกแจงเลขชี้กำลังเชิงเส้น ซึ่งขึ้นอยู่กับเวลา การสังเกตข้อมูลให้ครบทุกหน่วยนั้นทำได้ยาก เทคนิคหนึ่งที่นำมาแก้ปัญหานี้ ก็คือ การสังเกตข้อมูลในลักษณะข้อมูลแบ่งกลุ่ม

บทความนี้ได้นำเสนอ การประมาณค่าด้วยวิธีความควรจะเป็นสูงสุด สำหรับการแจกแจงเลขชี้กำลังเชิงเส้น ด้วยข้อมูลแบ่งกลุ่ม ซึ่งผลที่ได้จากการศึกษาจะเป็นแนวทางที่สำคัญเบื้องต้น ที่จะนำไปสู่วิธีการอธิบายลักษณะข้อมูลที่สอดคล้องกับการแจกแจงนี้ ด้วยการสังเกตข้อมูลแบ่งกลุ่ม

[1]ภาควิชาคณิตศาสตร์ สถิติและคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยอุบลราชธานี

E-mail: sckanich@ubu.ac.th

ABSTRACT

Using statistical distribution functions to explain data behavior is the scientific process accepted all around the world. They could show tendency and information of data used in decision making to apply in any area. Estimation of parameters for statistical distribution functions are seen as a crucial preliminary process.

However, it is found that data observing in some situations could not be completely collected especially in the situation rested on linear exponential distribution because they depended on time that is very difficult to observe. A technique used to solve this problem is collecting the grouped data.

This paper presents using maximum likelihood method to estimate parameters for linear exponential distribution by observing grouped data. The result from this study is a fundamental alternative technique to explain data corresponded with linear exponential distribution using grouped data.

คำสำคัญ: การประมาณค่า วิธีความควรจะเป็นสูงสุด การแจกแจงเลขชี้กำลังเชิงเส้น

Keywords: Estimation, Maximum likelihood method, Linear exponential distribution

บทนำ

ในยุคปัจจุบันข้อมูลทางสถิติมีประโยชน์ต่อการค้นคว้าวิจัยในหลายด้าน เช่น อายุการเก็บรักษาของผลิตภัณฑ์ ความพึงพอใจของลูกค้าที่มีต่อผลิตภัณฑ์เพื่อหาตลาดให้กับสินค้าชนิดใหม่ เป็นต้น อย่างไรก็ตามข้อมูลที่เก็บรวบรวมมาอาจไม่สมบูรณ์ เนื่องจากมีความผิดพลาดระหว่างเก็บข้อมูล ความผิดพลาดดังกล่าว อาจเกิดจากตัวบุคคลหรืออุปกรณ์เครื่องมือที่ใช้ในการตรวจสอบ และยังไม่มีเครื่องมือใดที่สามารถตรวจสอบข้อมูลได้อย่างละเอียดทุกหน่วยการทดสอบให้เป็นไปได้อย่างต่อเนื่อง ในทางสถิติศาสตร์เรียกข้อมูลลักษณะนี้ว่า ข้อมูลไม่สมบูรณ์ (censored data) จึงแก้ปัญหาโดยการแบ่งกลุ่มข้อมูล (grouped data) คือ แบ่งข้อมูลออกเป็นกลุ่ม ๆ ตามช่วงเวลาการเก็บข้อมูล โดยการนับค่าสังเกตที่เกิดขึ้นในช่วงนั้น ๆ ตลอดระยะเวลาที่ทำการตรวจสอบ นอกจากนี้วิธีการแบ่งกลุ่มข้อมูลช่วยลดอุปสรรคทางด้านปัจจัยต่าง ๆ เช่น ตัวบุคคล วิธีการ หรือเวลาที่ส่งผลต่อการเก็บข้อมูลและมีความสะดวกต่อผู้วิเคราะห์ข้อมูล เวลานั้น ๆ ตลอดระยะเวลาที่ทำการตรวจสอบ นอกจากนี้วิธีการเก็บข้อมูลแบ่งกลุ่มมักเป็นที่นิยมเนื่องจากช่วยลดอุปสรรคทางด้านปัจจัยต่าง ๆ ที่มีผลต่อการเก็บข้อมูล มีความสะดวกและรวดเร็วต่อผู้วิเคราะห์ข้อมูล (Ampai and Kanisa, 2011)

ข้อมูลที่มีความเกี่ยวข้องกับเวลาจะอยู่ในกลุ่มฟังก์ชันเลขชี้กำลัง ซึ่งการแจกแจงเลขชี้กำลังเชิงเส้น (linear exponential distribution; LED) เป็นรูปแบบทั่วไปของการแจกแจงเลขชี้กำลัง (exponential distribution; ED) เมื่อ $\beta = 0$ และการแจกแจงเรย์ลี (Rayleigh distribution; RD) เมื่อ $\alpha = 0$ ลักษณะข้อมูล

ดังกล่าวจะนำไปใช้ในการวิเคราะห์การอยู่รอด (survival analysis) ซึ่งพบในงานวิจัยด้านการแพทย์ ชีววิทยา วิศวกรรม เป็นต้น (Broadbent, 1958; Carbone et al., 1967; Lai et al., 2001)

โดยทั่วไปแล้วรูปแบบของฟังก์ชันการแจกแจงจะถูกกำหนดด้วยค่าพารามิเตอร์ แต่ค่าพารามิเตอร์ดังกล่าวจะไม่ทราบค่า ดังนั้นการประมาณค่าพารามิเตอร์จึงมีความสำคัญ การประมาณค่าด้วยวิธีความควรจะเป็นสูงสุด เป็นวิธีพื้นฐานที่นำไปใช้ในการประมาณค่าพารามิเตอร์ สามารถคำนวณหาค่าตัวประมาณโดยใช้วิธีการเชิงตัวเลข (numerical methods) ด้วยวิธีนิวตัน-ราฟสัน (Newton-Raphson method) ต่อมา Walter et al. (1971) และ Al-khedhairi (2008) ได้นำวิธีความควรจะเป็นสูงสุด มาประยุกต์ใช้กับข้อมูลที่ถูกจัดกลุ่ม

บทความนี้ได้นำเสนอประมาณค่าด้วยวิธีความควรจะเป็นสูงสุด ของการแจกแจงเลขชี้กำลังเชิงเส้น ด้วยข้อมูลแบ่งกลุ่ม ซึ่งผลที่ได้จากการศึกษาจะเป็นแนวทางที่สำคัญเบื้องต้น และนำไปสู่การอธิบายลักษณะข้อมูลที่สอดคล้องกับการแจกแจงดังกล่าว

การแจกแจงเลขชี้กำลังเชิงเส้น

ฟังก์ชันความหนาแน่นของความน่าจะเป็น (probability density function)

$$f(x:\alpha,\beta)=(\alpha+\beta x)\exp\left\{-\alpha x-\frac{\beta}{2}x^2\right\} \tag{1}$$

ฟังก์ชันการแจกแจงความน่าจะเป็นแบบสะสม (cumulative distribution function)

$$F(x;\alpha,\beta)=1-\exp\left\{-\alpha x-\frac{\beta}{2}x^2\right\} \quad , \; x>0,\, \alpha,\beta>0 \tag{2}$$

ฟังก์ชันอัตราภาวะภัย (hazard rate function)

$$h(x;\alpha,\beta)=\alpha+\beta x \tag{3}$$

ค่าเวลาเฉลี่ยก่อนการเสียหาย (mean time to failure; MTTF) คือ

$$\text{MTTF}=\frac{K\left(\alpha^2/2\beta\right)}{\sqrt{\beta/2}}-\frac{\alpha}{\beta} \tag{4}$$

โดยที่ $K(\upsilon)=\exp\{\upsilon\}\int_{\upsilon}^{\infty}x^{1/2}\exp\{-x\}\,dx$

ฟังก์ชันการอยู่รอด (survival function) คือ

$$S(t)=\exp\left\{-\alpha t-\frac{\beta}{2}t^2\right\} \tag{5}$$

ตัวประมาณค่าความควรจะเป็นสูงสุด (maximum likelihood method: MLE)

เมื่อ $t_1, t_2, t_3, \ldots, t_{k-1}$ เป็นข้อมูลที่สังเกต ณ เวลาหนึ่ง มีจำนวนข้อมูลทั้งหมด n จำนวน ถูกแบ่งเป็น k กลุ่ม ซึ่งจะประกอบด้วย $[0,t_1),[t_1,t_2),\ldots,[t_{k-1},\infty)$ แต่ละช่วง $\left[t_{i-1},t_i\right)$ มีจำนวนข้อมูล $n_i; i=1,2,3,\ldots,k-1$ ฟังก์ชันความควรจะเป็น (Likelihood Function) สำหรับค่าพารามิเตอร์ (θ) ใด ๆ

$$L(\theta)=\frac{n!}{\prod_{i=1}^{k} n_i}\prod_{i=1}^{k} P_i^{n_i} \tag{6}$$

สามารถพิจารณา ฟังก์ชันล็อก-ความควรจะเป็น (log-likelihood function) ของ $[0,t_1),[t_1,t_2),\ldots,\left[t_{k-1},\infty\right)$

$$\log L(\theta)=\log n!-\log\prod_{i=1}^{k} n_i+\sum_{i=1}^{k} n_i \log P_i \tag{7}$$

โดยที่ $P_i = P\left(t_{i-1} < X < t_i\right)$

สำหรับ LED (α,β)

$$\begin{aligned} P_i &= P\left(t_{i-1} < X < t_i\right) \\ &= \exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)-\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right) \end{aligned} \tag{8}$$

อนุพันธ์อันดับหนึ่งของ $\log L(t;\alpha,\beta)$ เทียบกับ α เป็นดังนี้

$$\begin{aligned} \frac{\partial}{\partial\alpha}\log L(t_i;\alpha,\beta) &= \sum_{i=1}^{k} n_i \frac{1}{P_i}\frac{d}{d\alpha}P_i \\ &= \sum_{i=1}^{k} n_i \frac{t_i \exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)-t_{i-1}\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)}{\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)-\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)} \end{aligned}$$

และกำหนดให้เท่ากับ 0 ดังสมการ (9)

$$\sum_{i=1}^{k} n_i \frac{t_i \exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)-t_{i-1}\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)}{\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)-\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)}=0 \tag{9}$$

ให้ $d=t_i-t_{i-1}$ จัดรูปสมการ (9) ใหม่ดังนี้

$$\sum_{i=1}^{k-1} n_i\left[\frac{d}{\exp\left(\alpha d+\frac{\beta}{2}d\left(t_i+t_{i+1}\right)\right)-1}-t_{i-1}\right]=0 \tag{10}$$

อนุพันธ์อันดับหนึ่งของ $\log L(t;\alpha,\beta)$ เทียบกับ β เป็นดังนี้

$$\frac{\partial}{\partial\beta}\log L(t;\alpha,\beta)=\sum_{i=1}^{k}n_i\frac{1}{P_i}\frac{d}{d\beta}P_i$$

$$=\sum_{i=1}^{k}n_i\frac{\frac{t_i^2}{2}\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)-\frac{t_{i-1}^2}{2}\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)}{\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)-\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)}$$

และกำหนดให้เท่ากับ 0 ดังสมการ (11)

$$\sum_{i=1}^{k}n_i\frac{\frac{t_i^2}{2}\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)-\frac{t_{i-1}^2}{2}\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)}{\exp\left(-\alpha t_{i-1}-\frac{\beta}{2}t_{i-1}^2\right)-\exp\left(-\alpha t_i-\frac{\beta}{2}t_i^2\right)}=0 \tag{11}$$

ค่าประมาณ α , β คือ $\hat{\alpha}$, $\hat{\beta}$ ตามลำดับ สามารถหาค่าที่เหมาะสมโดยใช้วิธีนิวตัน-ราฟสัน

1. กรณี $\beta=0$ จาก LED จะกลายเป็น ED มีพารามิเตอร์ คือ α ซึ่งสามารถหาค่าประมาณพารามิเตอร์ $(\hat{\alpha})$ โดยเขียนสมการ (10) ใหม่ดังนี้

$$\sum_{i=1}^{k-1}n_i\left[\frac{d}{\exp(\alpha d)-1}-t_{i-1}\right]=0 \tag{12}$$

$$\sum_{i=1}^{k-1}n_i\frac{d}{\exp(\alpha d)-1}-\sum_{i=1}^{k}n_i t_{i-1}=0 \tag{13}$$

เมื่อ $t_0=0$ ดังนั้น $d=t_1$ สามารถเขียนสมการ (13) ใหม่ ดังนี้

$$\sum_{i=1}^{k\text{-}1}\frac{n_i d}{\exp(\alpha d)-1}-\sum_{i=2}^{k}(i-1)n_i x_1=0 \tag{14}$$

จาก Gulati และ Neus (2003) ได้แสดง $\sum_{i=1}^{k\text{-}1}n_i d=(n-n_k)t_1$ สามารถจัดรูปสมการ (14) ใหม่ได้ดังนี้

$$\frac{(n-n_k)t_1}{\exp(\alpha t_1)}-\sum_{i=2}^{k}(i-1)n_i t_1=0 \tag{15}$$

จากสมการที่ (15) สามารถแก้สมการหาค่า α ได้ดังนี้

$$\hat{\alpha}=\frac{1}{t_1}\ln\left(1+\frac{n-n_k}{\sum_{i=2}^{k}(i-1)n_i}\right) \tag{16}$$

2. กรณี $\alpha=0$ จาก LED จะกลายเป็น RD มีพารามิเตอร์คือ β สามารถซึ่งสามารถหาค่าประมาณพารามิเตอร์ $(\hat{\beta})$ โดยเขียนสมการ (11) ใหม่ดังนี้

$$\sum_{i=1}^{k} n_i \frac{\frac{t_i^2}{2}\exp\left(-\frac{\beta}{2}t_i^2\right)-\frac{t_{i-1}^2}{2}\exp\left(-\frac{\beta}{2}t_{i-1}^2\right)}{\exp\left(-\frac{\beta}{2}t_{i-1}^2\right)-\exp\left(-\frac{\beta}{2}t_i^2\right)}=0 \tag{17}$$

ค่าประมาณ $\hat{\beta}$ ได้จากการแก้สมการ (17)

ขอบเขตความเชื่อมั่นเชิงเส้นกำกับ (asymptotic confidence bounds)

ค่าพารามิเตอร์$\left(\hat{\alpha},\hat{\beta}\right)$ที่ประมาณด้วย MLE ไม่สามารถเขียนให้อยู่ในรูปแบบปิดได้ อาจทำให้เกิดความคลาดเคลื่อน จึงทำการประมาณช่วงความเชื่อมั่นของค่าพารามิเตอร์ ภายใต้พื้นฐานการแจกแจงเชิงกำกับ (asymptotic distributions) ของ MLE โดย Miller (1981) ได้ศึกษาตัวประมาณของค่าพารามิเตอร์ $\hat{\theta}$ ทำให้อยู่ในรูปการแจกแจงเชิงกำกับ ดังนี้

$$\left(\theta-\hat{\theta}\right)\rightarrow N_2\left(0,I^{-1}\left(\theta\right)\right) \tag{18}$$

เมื่อ $I^{-1}\left(\theta\right)$ คือ เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม (variance-covariance matrix) ของพารามิเตอร์ที่ไม่ทราบค่า $\theta=\left(\alpha,\beta\right)$ มีขนาด 2×2 $I_{ij}\left(\theta\right),\, i,j=1,2,3$ สามารถประมาณด้วย $I_{ij}\left(\hat{\theta}\right)$ แล้ว

$$I_{ij}\left(\hat{\theta}\right)=-\left.\frac{\partial^2 L\left(\theta\right)}{\partial\theta_i\partial\theta_j}\right|_{\theta=\hat{\theta}} \tag{19}$$

จากสมการ (9) และ (11) หาอนุพันธ์อันดับสอง จะได้

$$\frac{\partial^2}{\partial\alpha^2}=\sum_{i=1}^{k} n_i \frac{P_i P_{i\alpha^2}-\left[P_{i,\alpha}\right]^2}{P_i^2}$$

$$\frac{\partial^2}{\partial\beta^2}=\sum_{i=1}^{k} n_i \frac{P_i P_{i\beta^2}-P_{i,\alpha}P_{i,\beta}}{P_i^2}$$

$$\frac{\partial^2}{\partial\alpha\partial\beta}=\sum_{i=1}^{k} n_i \frac{P_i P_{i\alpha\beta}-P_{i,\alpha}P_{i,\beta}}{P_i^2}$$

เมื่อ $P_i=S_{k-1}-S_i$, $i=1,2,3,\ldots,k$ ให้ $S_i=\exp\left(-\alpha t_i-\frac{\beta}{2}t_1^2\right)$

$$P_{i,\alpha}=-t_{i-1}S_{i-1}+t_iS_i$$

$$P_{i,\alpha^2}=t_{i-1}^2S_{i-1}+t_i^2S_i$$

$$P_{i,\beta}=-\frac{1}{2}t_{i-1}^2S_{i-1}+\frac{1}{2}t_i^2S_i$$

$$P_{i,\beta^2}=-\frac{1}{4}t_{i-1}^4S_{i-1}-\frac{1}{4}t_i^4S_i$$

$$P_{i,\alpha\beta} = \frac{1}{2}t_{i-1}^{3}S_{i-1} + \frac{1}{2}t_{i}^{3}S_{i}$$

ช่วงความเชื่อมั่น $100(1-\gamma)\%$ ของ α คือ $\hat{\alpha} \pm Z_{\frac{\gamma}{2}}\sqrt{I_{11}^{-1}\left(\hat{\alpha}\right)}$

และช่วงความเชื่อมั่น $100(1-\gamma)\%$ ของ β คือ $\hat{\beta} \pm Z_{\frac{\gamma}{2}}\sqrt{I_{22}^{-1}\left(\hat{\beta}\right)}$

เมื่อ $Z_{\frac{\gamma}{2}}$ คือ เปอร์เซ็นไทล์ที่ $100(1-\gamma)\%$ ของการแจกแจงปรกติมาตรฐาน โดยที่ γ คือ ระดับนัยสำคัญ

ตัวอย่างการประยุกต์ใช้

นำข้อมูลของ Conover (1971) มาประยุกต์ใช้กับวิธีการประมาณค่าที่ได้นำเสนอ โดยข้อมูลแสดงอายุการใช้งานของแบตเตอรี่ไฟฟ้า จำนวน 500 ก้อน มีการจดบันทึกข้อมูลดังนี้

ช่วงเวลา (ชม.)	จำนวน
[0, 50)	208
[50, 100)	112
[100, 150)	75
[150, 200)	40
[200, 250)	30
[250, 300)	18
[300, ∞)	17
รวม	500

กำหนดให้ข้อมูลมีลักษณะการแจกแจง ED ค่าพารามิเตอร์ที่ประมาณด้วยวิธี MLE $\hat{\alpha} = 0.010565$ ฟังก์ชันล็อก-ความควรจะเป็น คือ -355.320

กำหนดให้ข้อมูลมีการแจกแจง LED ค่าพารามิเตอร์ที่ประมาณด้วยวิธี MLE $\hat{\alpha} = 0.0101516$ และ $\hat{\beta} = 4.9243 \times 10^{-6}$ ฟังก์ชันล็อก-ความควรจะเป็น คือ -355.055

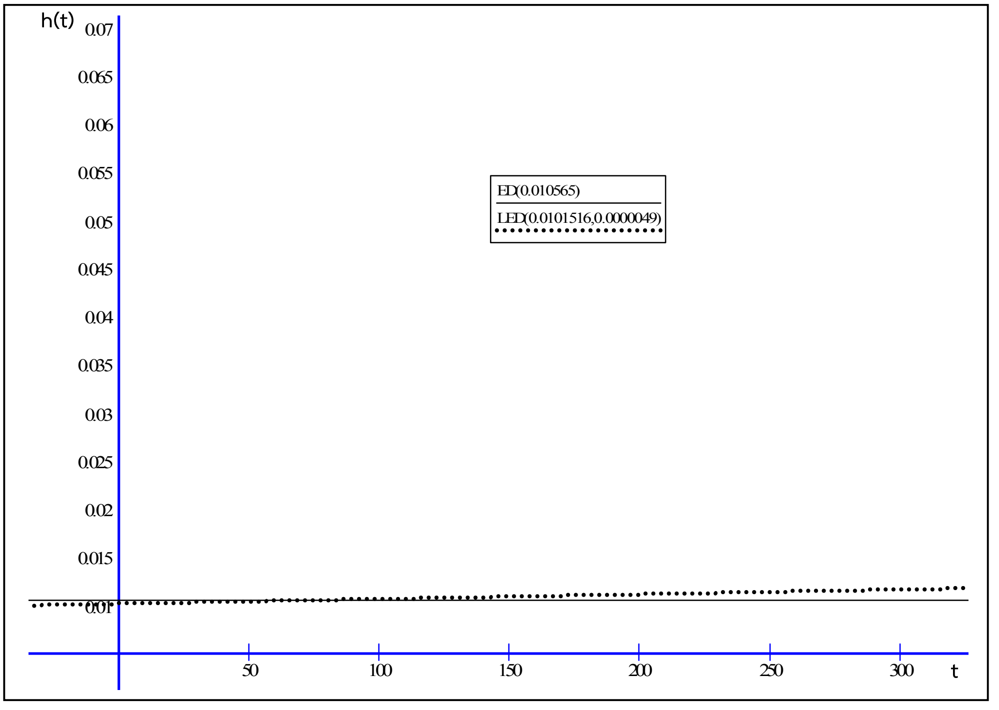
เมื่อพิจารณาจากค่าฟังก์ชันล็อก-ความควรจะเป็น* สรุปได้ว่า การแจกแจง LED มีความเหมาะสมกับข้อมูลมากกว่า การแจกแจง ED เล็กน้อย

เมทริกซ์ความแปรปรวน-ความแปรปรวนร่วม ของการแจกแจง LED(α, β) คือ

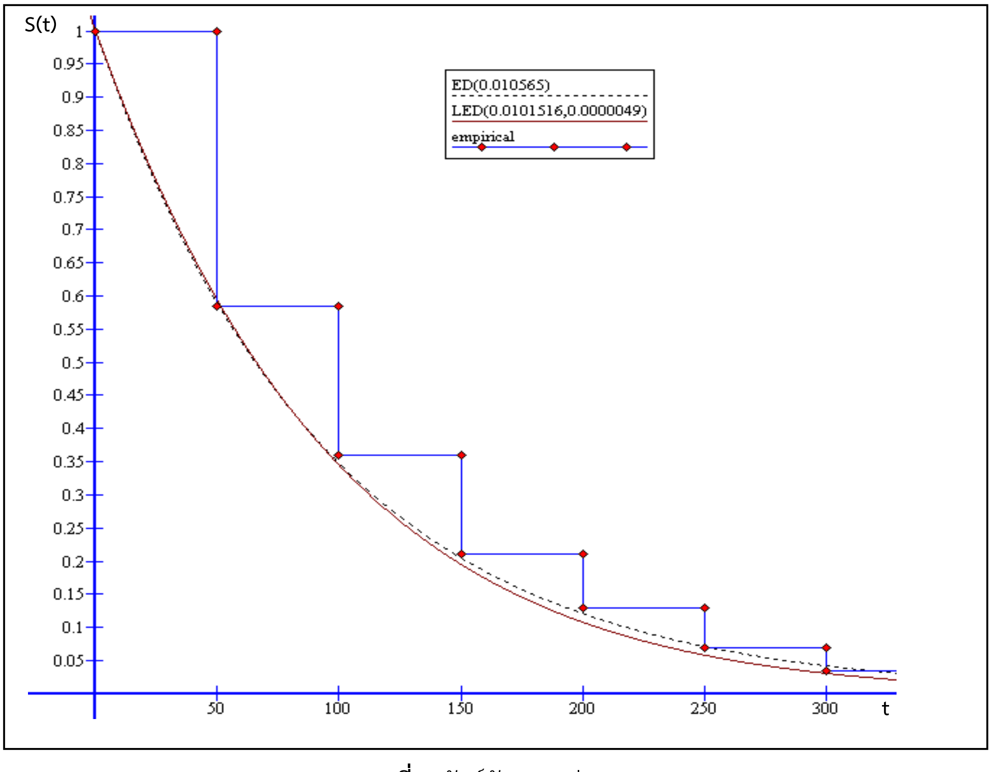
$$I^{-1} = \begin{bmatrix} 7.858\times10^{-7} & -6.7252\times10^{-9} \\ -6.7252\times10^{-9} & 8.23\times10^{-11} \end{bmatrix}$$

ดังนั้น ความแปรปรวนของ α และ β คือ $Var(\hat{\alpha}) = 7.858\times10^{-7}$ และ $Var(\hat{\beta}) = 8.23\times10^{-11}$ การประมาณค่าช่วงความเชื่อมั่น 95% ของ α อยู่ในช่วง $\left[8.41\times10^{-3}, 1.19\times10^{-2}\right]$ และ β อยู่ในช่วง $\left[0, 2.27\times10^{-5}\right]$ โดยที่ γ เท่ากับ 0.05

*ค่าฟังก์ชันล็อก-ความควรจะเป็น คือค่าที่ใช้พิจารณาความเหมาะสมของรูปแบบการแจกแจงของข้อมูล โดยรูปแบบการแจกแจงที่มีค่าฟังก์ชันล็อก-ความควรจะเป็นสูงกว่า จะมีความเหมาะสมกับข้อมูลนั้นได้ดีกว่ารูปแบบการแจกแจงที่มีค่าฟังก์ชันล็อก-ความควรจะเป็นต่ำกว่า



รูปที่ 1 ฟังก์ชันอัตราภาวะภัย



รูปที่ 2 ฟังก์ชันการอยู่รอด

รูปที่ 1 แสดงกราฟฟังก์ชันอัตราภาวะภัยของ ED และ LED เมื่อค่าพารามิเตอร์ใช้ MLE ในการประมาณค่า และรูปที่ 2 แสดงกราฟฟังก์ชันการอยู่รอดเมื่อค่าพารามิเตอร์ใช้ MLE ในการประมาณค่า เมื่อพิจารณารูปที่ 1 และรูปที่ 2 ลักษณะกราฟของ ED และ LED ใกล้เคียงกันมาก เนื่องจาก $\hat{\beta} = 4.9243 \times 10^{-6}$ ของ LED มีค่าใกล้เคียง 0 จึงเป็นไปตามกรณีที่ 1 ของ LED เมื่อ $\beta = 0$ รูปแบบ LED จะกลายเป็น ED

สรุป

บทความนี้ได้นำเสนอ การประมาณค่าพารามิเตอร์ของ LED คือ α, β กรณีที่ข้อมูลถูกจัดกลุ่มโดยวิธี MLE ทั้งประมาณค่าจุดและแบบช่วง และนำไปใช้ประยุกต์ใช้กับข้อมูลของ Conover (1971) ได้กำหนดให้ลักษณะข้อมูลเป็น ED และ LED โดยใช้ฟังก์ชันล็อก-ความควรจะเป็น พิจารณาความเหมาะสมของรูปแบบข้อมูล ผลลัพธ์ คือ LED มีความเหมาะสมมากกว่า ED และจากตัวอย่างแสดงให้เห็นว่า เมื่อค่าพารามิเตอร์ของ LED คือ β ใกล้ 0 ลักษณะกราฟมีความใกล้เคียง ED ดังรูปที่ 2

เอกสารอ้างอิง

Al-khedhairi, A. (2008). Parameters estimation for a linear exponential distribution based on grouped data. International Mathematical Forum 33.

Ampai Thongteeraparp and Kanisa Chodjuntug. (2011). A power of Comparison of goodness of fit tests for Exponential distribution with grouped data. Journal of The Thai Statistical Association 9(1).

Broadbent, S. (1958). Simple mortality rates. Journal of Applied Statistics 7.

Carbone, P., Kellerthouse, L. and Gehan, E. (1967). Plasmacytic myeloma: A study of the relationship of survival to various clinical mani festations and anomalous protein type in 112 patients. American Journal of Medicine 42.

Gulati, S. and J. Neus. (2003). Goodness of fit statistics for the Exponential distribution when the data are grouped. Communication in Statistics Theory and Methods 32.

Miller, Jr., R. G., (1981). Survival analysis. New York: John Wiley.

Lai, C.D., Xie, M., Murthy, D.N.P. (2001). Bathtub shaped failure rate distributions. In: Balakrishnan, N., Rao, C.R. (Eds.), Handbook in Reliability 20.

Conover, W. J. (1971). Practical nonparametric statistics, John Wiley and Sons.

Nelson, W. (1982). Applied life data analysis. New York: John Wiley.

Walter H. Carter, Jr., Jacob, Van Bowen, Jr. and Raymond, H. M. (1971). Maximum likelihood estimation from grouped Poisson data. Journal of the American Statistical Association 66.

❑❑❑❑❑