Minimax Estimation of the Parameter of ЭРланга Distribution Under Different Loss Functions

Lanping Li

Department of Basic Subjects, Hunan University of Finance and Economics, Changsha, China

Email address: Lilanping1981@163.com

To cite this article:

Lanping Li. Minimax Estimation of the Parameter of ЭРланга Distribution Under Different Loss Functions. Science Journal of Applied Mathematics and Statistics. Vol. 4, No. 5, 2016, pp. 229-235. doi: 10.11648/j.sjams.20160405.16

Received: August 31, 2016; Accepted: September 12, 2016; Published: October 8, 2016

Abstract: The aim of this article is to study the estimation of the parameter of ЭРланга distribution based on complete samples. The Bayes estimators of the parameter of ЭРланга distribution are obtained under three different loss functions, namely, weighted square error loss, squared log error loss and entropy loss functions by using conjugate prior inverse Gamma distribution. Then the minimax estimators of the parameter are derived by using Lehmann’s theorem. Finally, performances of these estimators are compared in terms of risks which obtained under squared error loss function.

Keywords: Bayes Estimator, Minimax Estimator, Squared Log Error Loss Function, Entropy Loss Function

1. Introduction

In reliability and supportability data analysis field, the most commonly used distribution are the exponential distribution, normal distribution and Weibull distribution, etc. But in some practical application, such as the repair time, guarantee the distribution delay time, the above several distributions does not just as one wish. At this time ЭРланга distribution was proposed as a suitable alternative distribution [1].

Suppose that the repair time $T$ obeys the ЭРланга distribution with the following probability density function (pdf) and distribution function respectively:

$$f(t; \theta) = 4t\theta^2 e^{-2\theta t}, \quad t \geq 0, \theta > 0$$  \hspace{1cm} (1)

$$F(t; \theta) = 1 - (1 + 2\theta t) e^{-2\theta t}, \quad t \geq 0, \theta > 0$$  \hspace{1cm} (2)

Here, $\theta$ is the unknown parameter. It is easily to see that $ET = \theta^{-1}$, and then the parameter $\theta^{-1}$ is also often referred to as the mean time to repair equipment.

Lv et al. [1] studied the characteristic parameters, such as mean, variance and median of the maximum likelihood estimation of ЭРланга distribution was also derived. Pan et al. [2] studied the interval estimation and hypothesis test of ЭРланга distribution based on small sample, and the difference of exponential distribution with 3 ЭРланга distribution was also discussed. Long [3] studied the estimation of the parameter of ЭРланга distribution based on missing data. Yu et al. [4] used the ЭРланга distribution to fit the battlefield injury degree, and established the simulating model, then proposed a new method to solve the problem in the production and distribution of battlefield injury in campaign macrocosm. Long [5] studied the Bayes estimation of ЭРланга distribution under type-II censored samples on the basis conjugate prior, Jeffreys prior and no information prior distributions.

The purpose of this paper is to study maximum likelihood estimation (MLE) and Bayes estimation of the parameter of ЭРланга distribution. Further, by using Lehmann’s theorem we derive minimax estimators under three loss functions.

2. Maximum Likelihood Estimation

Let $X = (X_1, X_2, \ldots, X_n)$ be a sample drawn from $\text{Erlang}$ distribution with pdf $\text{(1)}$, and $x = (x_1, x_2, \ldots, x_n)$ is the observation of $X$. For given sample observation, we can get the likelihood function of the parameter $\theta$ as follows:

$$l(\theta; x) = \prod_{i=1}^{n} f(x; \theta) = \prod_{i=1}^{n} 4\theta^2 x_i e^{-2\theta x_i} = \theta^{2n} \cdot (\prod_{i=1}^{n} x_i) \cdot e^{-2n\theta \sum_{i=1}^{n} x_i}$$ (3)

That is

$$l(\theta; x) = (\prod_{i=1}^{n} 4x_i)^{\cdot} \cdot \theta^{2n} e^{-2\theta \sum_{i=1}^{n} x_i} = (\prod_{i=1}^{n} 4x_i)^{\cdot} \cdot \theta^{2n} e^{-\theta \sum_{i=1}^{n} x_i}$$ (4)

Here $t = \sum_{i=1}^{n} x_i$ is the observation of $T = 2\sum_{i=1}^{n} X_i$.

Then the log-likelihood function is

$$\text{Ln}[l(\theta; x)] = \sum_{i=1}^{n} \ln(4x_i) + 2n \ln \theta - \theta t$$

By solving log-likelihood equation

$$\frac{d\text{Ln}[l(\theta; x)]}{d\theta} = 0,$$

the maximum likelihood estimator of $\theta$ can be easily derived as follows:

$$\hat{\theta}_M = \frac{n}{t}$$ (5)

And by Eq. (1), we can easily show that $T$ is a random variable distributed with the Gamma distribution $\Gamma(2n, \theta)$, which has the following probability density function:

$$f_T(t; \theta) = \frac{\theta^t 
\Gamma^{-1} e^{-\theta \theta}}{t}, \quad t > 0, \theta > 0$$ (6)

3. Bayesian Estimation

In Bayesian statistical analysis, loss function plays an important role in the Bayes estimation and Bayes test problems. Many loss function are proposed in Bayesian analysis, and squared error loss function is the most common loss function, which is a symmetric loss function. In many practical problems, especially in the estimation of reliability and failure rates, symmetric loss may be not suitable, because it is to be thought the overestimation will bring more loss than underestimation [17]. Then some asymmetric loss functions are developed. For example, Zellner [18] proposed the LINEX loss in Bayes estimation, Brown [19] put forward the squared log error loss function for estimating unknown parameter, Dey et al. [20] proposed the entropy loss function in the Bayesian analysis.

In this paper, we will discuss the Bayes estimation of the unknown parameter $\theta$ of $\text{Erlang}$ distribution under the following loss functions:

(i) Weighted squared error loss function

$$L_1(\theta, \delta) = \frac{(\delta - \theta)^2}{\theta^2}$$ (7)

Under the weighted squared error loss function (7), the Bayes estimator of $\theta$ is

$$\hat{\delta}_{BE} = \frac{E[\theta^{-1} | X]}{E[\theta^{-2} | X]}$$ (8)

(ii) Squared log error loss function

Squared log error loss function is a asymmetric loss function, which first proposed by Brown for estimating scale parameter. This loss function can also be found in Kiapour and Nematollahi [21] with the following form:

$$L_2(\theta, \delta) = (\ln \delta - \ln \theta)^2$$ (9)

Obviously, $L_2(\theta, \delta) \to \infty$ as $\delta \to 0$ or $\infty$. The loss function (9) is not always convex, and it is convex for $\frac{\delta}{\theta} \leq e$ and concave otherwise. But the risk function of this function has minimum value, which we also call it the Bayes estimator $\hat{\delta}_{SL}$ under squared log error loss function. That is

$$\hat{\delta}_{SL} = \exp[E(\ln \theta | X)]$$ (10)

(iii) Entropy loss function

In many practical situations, it appears to be more realistic to express the loss in terms of the ratio $\frac{\delta}{\theta}$. In this case, Dey et al. [20] pointed out a useful asymmetric loss function named entropy loss function:

$$L_3(\theta, \delta) = \frac{\delta}{\theta} - \ln \frac{\delta}{\theta} - 1$$ (11)

Whose minimum occurs at $\delta = \theta$. Also, this loss function has been used in Singh et al. [22], Nematollahi and Motamed-Shariati [23]. The Bayes estimator under the entropy loss (11) is denoted by $\hat{\delta}_{BE}$, obtained by

$$\hat{\delta}_{BE} = [E(\theta^{-1} | X)]^{-1}$$ (12)

In this section, we will estimate the unknown parameter $\theta$ on the basis of the above three mentioned loss functions. We further assume that some prior knowledge about the parameter $\theta$ is available to the investigation from past experience with the $\text{Erlang}$ model. The prior knowledge can often be summarized in terms of the so-called prior densities on parameter space of $\theta$. In the following
discussion, we assume the following Jeffrey’s non-informative quasi-prior density defined as,
\[
\pi(\theta) \propto \frac{1}{\theta^d}, \quad \theta > 0 \tag{13}
\]

Hence, \(d = 0\) leads to a diffuse prior and \(d = 1\) to a non-informative prior.

Let \(X = (X_1, X_2, \ldots, X_n)\) be a sample drawn from ЭРланга distribution with pdf (1), and \(x = (x_1, x_2, \ldots, x_n)\) is the observation of \(X\). Combining the likelihood function (3) with the prior density (13), the posterior probability density of \(\theta\) can be derived using Bayes Theorem as follows

\[
h(\theta | x) \propto l(\theta; x) \cdot \pi(\theta) \propto \theta^x e^{-\theta d} \propto \theta^{x-d} e^{-\theta} \tag{14}
\]

Theorem 1. Let \(X = (X_1, X_2, \ldots, X_n)\) be a sample of ЭРланга distribution with probability density function (1), and \(x = (x_1, x_2, \ldots, x_n)\) is the observation of \(X\). Then if the Bayes risk function (9) can be derived using Bayes Theorem as follows

\[
E(\ln \theta | X) = \frac{T^{2n-d+1}}{\Gamma(2n-d+1)} \int_0^{\infty} \ln \theta \cdot \theta^{2n-d+1-1} e^{-\theta} d\theta
\]

Then the Bayes estimator under the squared log error loss function (9) is come out to be

\[
\hat{\delta}_{BE} = \exp[E(\ln \theta | X)] = \frac{\Psi(n)}{T}
\]

(iii) By Eqs. (12) and (17), the Bayes estimator under the entropy loss function (11) is given by

\[
\hat{\delta}_{BE} = [E(\theta^{-1} | X)]^{-1} = \frac{2n-d}{T}
\]

4. Minimax Estimation of ЭРланга Distribution

This section will derive the minimax estimators of ЭРланга Distribution by using Lehmann’s Theorem, which depends on specific prior distribution and loss functions of a Bayesian method. The Lehmann’s Theorem is stated as follows:

Lemma 1 Let \(\tau = \{F_{\theta}; \theta \in \Theta\}\) be a class distribution functions and \(D\) be the estimators of \(\theta\). Suppose that \(\theta^* \in D\) is a Bayes estimator, which derived on the basis of a prior distribution \(\pi^*(\theta)\) on \(\Theta\). Then if the Bayes risk function \(R(\theta, \theta)\) equals constant on \(\Theta\), then \(\theta^*\) is a minimax estimator of \(\theta\).

Theorem 2 Let \(X = (X_1, X_2, \ldots, X_n)\) be a sample drawn from ЭРланга distribution with pdf (1), and \(x = (x_1, x_2, \ldots, x_n)\) is the observation of \(X = (X_1, X_2, \ldots, X_n)\). Suppose that \(t = 2\sum_{i=1}^{n} x_i\) is the observation of the statistics \(T = 2\sum_{i=1}^{n} X_i\). Then

(i) Under the weighted square error loss function (7),

\[
\hat{\delta}_{WS} = \frac{2n-d-1}{T} \quad \text{is the minimax estimator of parameter } \theta
\]
(ii) Under the squared log error loss function,
\[ \hat{\delta}_{SL} = \frac{\theta^{n-d+1}}{T} \] is the minimax estimator of parameter \( \theta \).

(iii) Under the entropy loss function, \( \hat{\delta}_{SL} = \frac{3n-2d}{T} \) is the minimax estimator of parameter \( \theta \).

Proof. To use Lehmann’s Theorem for the proof of the results. We need calculate the risk function of Bayes estimators and prove these risk functions are constants.

For the case (i), we can derive the risk function of the Bayes estimator \( \hat{\delta}_{BS} \) under the weighted square error loss function (7) as follows:

\[ R(\theta) = E\left[ L\left(\theta, \hat{\delta}_{BS}\right)\right] = E\left[ L\left(\theta, \hat{\delta}_{BS}\right)\right] = \frac{1}{\theta^2} \left[ (2n-d-1)^2 E(T^{-1}) - 2\theta(2n-d-1)E(T^{-1}) + \theta^2 \right] \]

From equation (6), we have \( T \sim \Gamma(2n, \theta) \), then we have

\[ E(T^{-1}) = \frac{\theta^2}{(2n-1)(2n-2)} \]

Consequently,

\[ R(\theta) = \frac{1}{\theta^2} \left[ (2n-d-1)^2 - 2\theta(2n-d-1) + \theta^2 \right] = 1 - \frac{2n-d-1}{2n-1} + \frac{(2n-d-1)^2}{(2n-1)(2n-2)} \]

Then, for Bayes estimator \( \hat{\delta}_{BS} \), the risk function \( R(\theta) \) is a constant on the parameter \( \theta \). So, According to Lemma 1, \( \hat{\delta}_{BS} \) is the minimax estimator of parameter \( \theta \) under the weighted square error loss function (7).

For the case (ii). The risk function of the Bayes estimator \( \hat{\delta}_{SL} \) is

\[ R(\theta) = E\left[ L\left(\theta, \hat{\delta}_{SL}\right)\right] = E\left[ \left( \ln \hat{\delta}_{SL} - \ln \theta \right)^2 \right] = E[\ln \hat{\delta}_{SL}]^2 - 2 \ln \theta \cdot E[\ln \hat{\delta}_{SL}] + (\ln \theta)^2 \]

By \( T \sim \Gamma(2n, \theta) \), we can easily get the result

\[ E(\ln T) = \Psi(2n) - \ln \theta \]

Then

\[ E[\ln \hat{\delta}_{SL}] = E[\Psi(2n-d+1) - \ln T] = \Psi(2n-d+1) - [\Psi(2n) - \ln \theta] = \Psi(2n-d+1) - \Psi(2n) + \ln \theta \]

\[ E[\ln \hat{\delta}_{SL}]^2 = E[\Psi(2n-d+1) - \ln T]^2 = \Psi^2(2n-d+1) - 2\Psi(2n-d+1)E[\ln T] + E[\ln T]^2 \]

Let \( Y \sim \Gamma(n, 1) \), then we can prove that \( Y = T \theta \sim \Gamma(2n, 1) \).

The derivative of \( \Psi(n) \) is

\[ \Psi'(n) = \int_0^{\infty} (\ln y)^2 y^{n-1} e^{-y} dy - \int_0^{\infty} (\ln y)^2 y^{n-1} e^{-y} dy = E[(\ln Y)^2] - \Psi^2(n) \]

Then

\[ \Psi^2(2n) + \Psi'(2n) = E[(\ln Y)^2] = E[(\ln T)^2] = E[(\ln T)^2] \]

\[ + 2 \ln \theta \cdot E[\ln(T)] + (\ln \theta)^2 = E[(\ln T)^2] + 2 \ln \theta \cdot (\Psi(2n) - \ln \theta) + (\ln \theta)^2 \]

From above results, we can get the fact

\[ E[(\ln T)^2] = \Psi^2(2n) + \Psi'(2n) - 2 \ln \theta \cdot \Psi(2n) + (\ln \theta)^2 \]

Further, we have

\[ E[\ln \hat{\delta}_{SL}]^2 = \Psi^2(2n-d+1) - 2\Psi(2n-d+1)E[\ln T] + E[\ln T]^2 \]
\[
R(\theta) = E[\ln \hat{\delta}_{LS}]^2 - 2 \ln \theta \cdot E[\ln \hat{\delta}_{LS}] + (\ln \theta)^2 = \Psi^2(2n-d+1) - 2\Psi(2n-d+1)\ln \theta + \Psi^2(2n) - 2\ln \theta \cdot \Psi(2n)
\]
\[
+ (\ln \theta)^2 - 2 \ln \theta \cdot \Psi(2n-d+1) + \Psi(2n) + \ln \theta + (\ln \theta)^2 = \Psi^2(2n-d+1) - 2\Psi(2n)\Psi(2n-d+1) + \Psi^2(2n) + \Psi^2(2n)
\]

Then \(R(\theta)\) is also a constant about the parameter \(\theta\). So, according to Lemma 1, we know that, \(\hat{\delta}_{LS}\) is a minimax estimator for the parameter \(\theta\) under the squared log error loss function.

For the case (iii). The risk function of the Bayes estimator \(\hat{\delta}_{BE}\) can be obtained as follows:

\[
R(\theta) = E\left[ L(\theta, \hat{\delta}_{BE}) \right] = E\left[ \frac{2n-d}{T^2 \theta} \cdot \ln \frac{2n-d}{T \theta} + 1 \right] = \frac{2n-d}{\theta} E(T) - \ln(2n-d) + \ln \theta + E(\ln T) + 1
\]

Then \(R(\theta)\) is also a constant about the parameter \(\theta\). So, according to lemma 1, we know that, \(\hat{\delta}_{BE}\) is a minimax estimator for the parameter \(\theta\) under the entropy loss function.

5. Performances of Bayes Estimators

To illustrate the performance of these Bayes estimators, squared error loss function \(L(\hat{\theta}, \theta) = (\delta - \theta)^2\) is used as a loss function to compare them. We note \(R(\hat{\delta}_{LS}), R(\hat{\delta}_{BE})\) and \(R(\hat{\delta}_{BL})\) are the risk functions of estimators \(\hat{\delta}_{LS}, \hat{\delta}_{BE}\) and \(\hat{\delta}_{BL}\) relative to the squared error loss, respectively. They can be easily derived as follows:

\[
R(\hat{\delta}_{LS}) = E[L(\hat{\delta}_{LS}, \theta)] = E[(\hat{\delta}_{LS} - \theta)^2] = E\left[ \frac{2n-d}{T} - \theta^2 \right]
\]

\[
= \int_0^\infty \left[ \frac{(2n-d-1)Y}{t^2} \right] \cdot \frac{2(2n-d-1)\theta^2}{t} \cdot f_Y(t) dt = \theta^2 \left[ \frac{(2n-d-1)Y}{(2n-1)(2n-2)} - \frac{2(2n-d-1)}{2n-1} + 1 \right],
\]

\[
R(\hat{\delta}_{BE}) = E[L(\hat{\delta}_{BE}, \theta)] = E[(\hat{\delta}_{BE} - \theta)^2] = E\left[ \frac{e^{\Psi(2n-d+1)}}{T} - \theta^2 \right] = \theta^2 \left[ \frac{e^{\Psi(2n-d+1)}}{(2n-1)(2n-2)} - \frac{2e^{\Psi(2n-d+1)}}{2n-1} + 1 \right]
\]

\[
R(\hat{\delta}_{BL}) = E[L(\hat{\delta}_{BL}, \theta)] = E[(\hat{\delta}_{BL} - \theta)^2] = E\left[ \frac{2n-d}{T} - \theta^2 \right] = \theta^2 \left[ \frac{2n-d}{(2n-1)(2n-2)} - \frac{2(2n-d)}{2n-1} + 1 \right]
\]

Let \(L_1, L_2\) and \(L_3\) are the ratio of the risk functions to \(\theta^2\), which are plotted in Figs. 1-4 with different sample sizes, \((n=10, 20, 30, 50)\)

![Figure 1. Performance of estimators with n=10.](image1.png)

![Figure 2. Performance of estimators with n=20.](image2.png)
From Figure 1 to Figure 4, we know that no of these estimators is uniformly better than other estimators. Then in practice, we recommend to select the estimator according to the prior parameter value \( d \) when assuming the quasi-prior as the prior distribution.

6. Conclusion

This paper derived Bayes estimators of the parameter of \( \text{Erlanda} \) distribution under weighted squared error loss, squared log error loss and entropy loss functions. Monte Carlo simulations show that the risk functions of these estimators, defined under squared error loss function, are all decrease as sample size \( n \) increases. The risk functions more and more close to each other when the sample size \( n \) is large, such as \( n>50 \).

Acknowledgement

This study is partially supported by Natural Science Foundation of Hunan Province (No. 2015JJ3030) and Foundation of Hunan Educational Committee (No. 15C0228). The author also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

[1] Lv H. Q., Gao L. H. and Chen C. L., 2002. \( \text{Erlanda} \) distribution and its application in supportability data analysis. Journal of Academy of Armored Force Engineering, 16 (3): 48-52.

[2] Pan G. T., Wang B. H., Chen C. L., Huang Y. B. and Dang M. T., 2009. The research of interval estimation and hypothetical test of small sample of \( \text{Erlanda} \) distribution. Application of Statistics and Management, 28 (3): 468-472.

[3] Long B., 2013. The estimations of parameter from \( \text{Erlanda} \) distribution under missing data samples. Journal of Jiangxi Normal University (Natural Science), 37 (1): 16-19.

[4] Yu C. M., Chi Y. H., Zhao Z. W., and Song J. F., 2008. Maintenance-decision-oriented modeling and emulating of battlefield injury in campaign macrocosm. Journal of System Simulation, 20 (20): 5669-5671.

[5] Long B., 2015. Bayesian estimation of parameter on \( \text{Erlanda} \) distribution under different prior distribution. Mathematics in Practice & Theory, 4: 186-192.

[6] Jiao J., Venkat K., Han Y., Weissman T., 2015. Minimax estimation of functionals of discrete distributions. Information Theory IEEE Transactions on, 61 (5): 2835-2885.

[7] Gao C., Ma Z., Ren Z., Zhou H. H. 2014. Minimax estimation in sparse canonical correlation analysis. Annals of Statistics, 43 (5): 905-912.

[8] Kogan M. M., 2014. LMI-based minimax estimation and filtering under unknown covariances. International Journal of Control, 87 (6): 1216-1226.

[9] Tchakian T. T., Zhuk S., 2015. A macroscopic traffic data-assimilation framework based on the Fourier–Galerkin method and Minimax estimation [J]. IEEE Transactions on Intelligent Transportation Systems, 16 (1): 452-464.

[10] Roy, M. K., Podder C. K. and Bhuiyan K. J., 2002. Minimax estimation of the scale parameter of the Weibull distribution for quadratic and MLINEX loss functions, Jahangirnagar University Journal of Science, 25: 277-285.

[11] Podder, C. K., Roy M. K., Bhuiyan K. J. and Karim A., 2004. Minimax estimation of the parameter of the Pareto distribution for quadratic and MLINEX loss functions, Pak. J. Statist., 20 (1): 137-149.

[12] Dey, S., 2008. Minimax estimation of the parameter of the Rayleigh distribution under quadratic loss function, Data Science Journal, 7 (1): 23-30.

[13] Shadrokh, A. and Pazira H., 2010. Minimax estimation on the Minimax distribution, International Journal of Statistics and Systems, 5 (2): 99-118.

[14] Rasheed H. A., Al-Shareefi E. F., 2015. Minimax estimation of the scale parameter of Laplace distribution under squared-log error loss function. Mathematical Theory & Modeling, 5 (1): 183-193.
[15] Li L. P., 2014. Minimax estimation of the parameter of exponential distribution based on record values. International Journal of Information Technology & Computer Science, 6 (6): 47-53.

[16] Li L. P., 2016, Minimax estimation of the parameter of Maxwell distribution under different loss functions, American Journal of Theoretical and Applied Statistics, 5 (4): 202-207.

[17] Li X., Shi Y., Wei J., Chai J., 2007. Empirical Bayes estimators of reliability performances using LINEX loss under progressively Type-II censored samples [J]. Mathematics & Computers in Simulation, 2007, 73 (5): 320-326.

[18] Zellner, A., 1986. Bayesian estimation and prediction using asymmetric loss function. Journal of American statistical Association, 81 (394): 446-451.

[19] Brown L., 1968. Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. Annals of Mathematical Statistics, 39 (1): 29-48.

[20] Dey, D. K., Ghosh M. and Srinivasan C., 1987. Simultaneous estimation of parameters under entropy loss, J. Statist. Plan. and Infer., 15 (3): 347-363.

[21] Kiapour A. and Nematollahib N., 2011. Robust Bayesian prediction and estimation under a squared log error loss function. Statistics & Probability Letters, 81 (11): 1717-1724.

[22] Singh S. K., Singh U. and Kumar D., 2011. Bayesian estimation of the exponentiated Gamma parameter and reliability function under asymmetric loss function. REVSTAT, 9 (3): 247-260.

[23] Nematollahi N., Motamed-Shariati F., 2009. Estimation of the scale parameter of the selected Gamma population under the entropy loss function. Communication in Statistics- Theory and Methods, 38 (7): 208-221.