Journal of Applied Mathematics and Physics, 2023, 11, 2838-2847
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

A Comparison of Four Methods of Estimating the Scale Parameter for the Exponential Distribution

Huda M. Alomari

Department of Mathematics, Al-Baha University, Al Baha, KSA
Email: haldrawshah@bu.edu.sa

How to cite this paper: Alomari, H.M. (2023) A Comparison of Four Methods of Estimating the Scale Parameter for the Exponential Distribution. *Journal of Applied Mathematics and Physics*, **11**, 2838-2847.
https://doi.org/10.4236/jamp.2023.1110186

Received: September 23, 2023
Accepted: October 16, 2023
Published: October 19, 2023

Abstract

In this paper, the estimators of the scale parameter of the exponential distribution obtained by applying four methods, using complete data, are critically examined and compared. These methods are the Maximum Likelihood Estimator (MLE), the Square-Error Loss Function (BSE), the Entropy Loss Function (BEN) and the Composite LINEX Loss Function (BCL). The performance of these four methods was compared based on three criteria: the Mean Square Error (MSE), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC). Using Monte Carlo simulation based on relevant samples, the comparisons in this study suggest that the Bayesian method is better than the maximum likelihood estimator with respect to the estimation of the parameter that offers the smallest values of MSE, AIC, and BIC. Confidence intervals were then assessed to test the performance of the methods by comparing the 95% CI and average lengths (AL) for all estimation methods, showing that the Bayesian methods still offer the best performance in terms of generating the smallest ALs.

Keywords

Bayes Estimator, Maximum Likelihood Estimator, Mean Squared Error (MSE), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC)

1. Introduction

The exponential distribution is one of the most commonly used continuous distributions applied in life data analysis. In particular, it is commonly used for systems exhibiting some form of constant failure rate. For a continuous random variable $X = (X_1, X_2, \cdots, X_n)$, the exponential distribution density function, the pdf, is thus given by

$$f(X, \lambda) = \lambda e^{-X\lambda}, \quad X \geq 0. \tag{1}$$

where λ is the constant failure rate parameter. For more details on exponential distribution and its applications, [1] and [2] offer a useful reference.

Empirical Bayes estimators of exponential distribution parameters have been introduced by multiple authors, such as [3] and [4] to support the generation of this function. [5] also investigated the use of the power function distribution as a conjugate prior to the estimation of the parameters required for the exponential distribution, while [6] used a prior drawn from the gamma to assess Bayesian estimation of the exponential distribution, applying the maximum likelihood estimator and three different loss functions to estimate the parameters of the exponential distribution.

This paper aimed to study the estimation of exponential distribution across a variety of loss functions, applying three different criteria-based methods to compare the resulting estimators. These were mean square error (MSE), the Akaike Information Criterion (AIC), and the Bayesian Information Criterion (BIC).

The remainder of this paper is thus organized as follows. In Section 2, the mathematical derivations of the estimation methods are present, while the final model selection is explained in Section 3. The Monte Carlo simulation study results are then used in Section 4 to perform a comparison of the estimation methods, based on applying the appropriate criteria for the mean square error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC), allowing a conclusion based on the results findings to be offered in Section 5.

2. Different Estimators of the λ Parameter

This section outlines the various estimates of the parameter of the exponential distribution, as given in (1).

2.1. Maximum Likelihood Estimator (MLE)

The maximum likelihood estimator (MLE) is a technique used for estimating the parameters of a given distribution as discussed in [7] [8] and [9]. Suppose that $X = (X_1, X_2, \cdots, X_n)$; based on this the exponential distribution is as shown in (1), and the likelihood of λ can be described as

$$L(X, \lambda) = \lambda^n e^{-\lambda \sum_{i=1}^{n} X_i}.$$

Taking the natural logarithm of both sides yields

$$\ln L(X, \lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} X_i$$

and the MLE estimator of λ can thus be obtained by solving the following equation:

$$\frac{\partial}{\partial\lambda}\ln L(X,\lambda)=0.$$

Hence,

$$\frac{n}{\lambda}-\sum_{i=1}^{n}X_i=0,$$

and the maximum likelihood estimator, $\hat{\lambda}_{\text{MLE}}$, is then given by

$$\hat{\lambda}_{\text{MLE}}=\frac{n}{\sum_{i=1}^{n}X_i}. \tag{2}$$

2.2. Bayesian Estimator

This section demonstrates the process of deriving Bayesian estimates of the scale parameter for the exponential distribution. Three different loss functions are used to achieve this, which are the squared error loss function, the entropy loss function, and the composite LINEX loss function.

The gamma (α,β) can be considered as a conjugate prior of λ with its density function written in the form

$$h(\lambda,\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda},\quad \lambda,\alpha,\beta>0,$$

where α and β are the shape parameter and scale parameter, respectively.

The posterior density function of λ for the given random sample $X=(X_1,X_2,\cdots,X_n)$ is thus obtained as

$$\begin{aligned}\pi(\lambda,X)&=\frac{\prod_{i=1}^{n}f(\lambda,X)h(\lambda,\alpha,\beta)}{\int_0^{\infty}\prod_{i=1}^{n}f(\lambda,X)h(\lambda,\alpha,\beta)\mathrm{d}\lambda}\\&=\frac{\lambda^{n}\mathrm{e}^{-\lambda\sum_{i=1}^{n}X_i}\beta^{\alpha}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}}{\int_0^{\infty}\lambda^{n}\mathrm{e}^{-\lambda\sum_{i=1}^{n}X_i}\beta^{\alpha}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}\mathrm{d}\lambda}\\&=\frac{\lambda^{n+\alpha-1}\mathrm{e}^{-\left(\sum_{i=1}^{n}X_i+\beta\right)\lambda}}{\int_0^{\infty}\lambda^{n+\alpha-1}\mathrm{e}^{-\left(\sum_{i=1}^{n}X_i+\beta\right)\lambda}\mathrm{d}\lambda}.\end{aligned}$$

This in turn implies that the posterior distribution can be written as

$$\pi(\lambda,X)=\frac{\lambda^{n+\alpha-1}\mathrm{e}^{-\left(\sum_{i=1}^{n}X_i+\beta\right)\lambda}\left(\sum_{i=1}^{n}X_i+\beta\right)^{n+\alpha}}{\Gamma(n+\alpha)},$$

which, as can be plainly observed, is a gamma distribution with parameters $(n+\alpha)$ and $\left(\sum_{i=1}^{n}X_i+\beta\right)$.

The three different loss functions used to develop a Bayes estimate for the parameter λ are discussed below.

2.2.1. Squared-Error Loss Function (SE)

The SE as discussed by [10] and [11], can be defined as

$$L(\hat{\lambda},\lambda)=(\hat{\lambda}-\lambda)^2.$$

The Bayesian estimator of λ under the squared error loss function is the mean of the posterior density function. The Bayes estimator of λ under the squared error loss function is then denoted as $\hat{\lambda}_{\text{SE}}$ which can be written as

$$\begin{aligned}
E(\lambda \mid X) &= \int_0^\infty \lambda\pi(\lambda,X)\,\mathrm{d}\lambda \\
&= \frac{\left(\sum_{i=1}^n X_i+\beta\right)^{n+\alpha}}{\Gamma(n+\alpha)}\int_0^\infty \lambda^{n+\alpha}\mathrm{e}^{-\left(\sum_{i=1}^n X_i+\beta\right)\lambda}\,\mathrm{d}\lambda \\
&= \frac{\Gamma(n+\alpha+1)}{\left(\sum_{i=1}^n X_i+\beta\right)^{n+\alpha+1}}\cdot\frac{\left(\sum_{i=1}^n X_i+\beta\right)^{n+\alpha}}{\Gamma(n+\alpha)} \\
&= \frac{n+\alpha}{\sum_{i=1}^n X_i+\beta}.
\end{aligned}$$

Such that,

$$\hat{\lambda}_{\text{SE}}=\frac{n+\alpha}{\sum_{i=1}^n X_i+\beta}, \tag{3}$$

where $E(.)$ indicates the posterior expectation.

2.2.2. Entropy Loss Function

The entropy loss function, as discussed by [12], can be obtained in the form

$$L(\hat{\lambda},\lambda)=K\left(\frac{\hat{\lambda}}{\lambda}-\ln\left(\frac{\hat{\lambda}}{\lambda}\right)-1\right).$$

The Bayes estimator of λ based on the entropy loss function, denoted by $\hat{\lambda}_{\text{BEN}}$, is then given as

$$\hat{\lambda}_{\text{BEN}}=\left(E\left(\lambda^{-1}\right)\right)^{-1}.$$

From this, it is possible to derive $E(\lambda^{-1})$ as follows:

$$\begin{aligned}
E\left(\lambda^{-1}\right) &= \int_0^\infty \lambda^{-1}\pi(\lambda,X)\,\mathrm{d}\lambda \\
&= \frac{\left(\sum_{i=1}^n X_i+\beta\right)^{n+\alpha}}{\Gamma(n+\alpha)}\int_0^\infty \lambda^{(n+\alpha-1)-1}\mathrm{e}^{-\left(\sum_{i=1}^n X_i+\beta\right)\lambda}\,\mathrm{d}\lambda \\
&= \frac{\left(\sum_{i=1}^n X_i+\beta\right)^{n+\alpha}}{\Gamma(n+\alpha)}\frac{\Gamma(n+\alpha-1)}{\left(\sum_{i=1}^n X_i+\beta\right)^{n+\alpha-1}}.
\end{aligned}$$

The Bayesian estimation of λ based on the entropy loss function is therefore written as

$$\hat{\lambda}_{\text{BEN}}=\frac{n+\alpha-1}{\sum_{i=1}^n X_i+\beta}. \tag{4}$$

2.2.3. Composite LINEX Loss Function

The composite LINEX loss function is defined by [13] and is given as

$$L\left(\hat{\lambda},\lambda\right)=\mathrm{e}^{-C\left(\hat{\lambda},\lambda\right)}+\mathrm{e}^{C\left(\hat{\lambda},\lambda\right)}-2,\quad C>0.$$

The Bayes estimator of λ based on composite LINEX loss function is denoted as $\hat{\lambda}_{\mathrm{BCL}}$ and can be expressed as

$$\hat{\lambda}_{\mathrm{BCL}}=\frac{1}{2C}\ln\left(\frac{E\mathrm{e}^{(C\lambda|X)}}{E\mathrm{e}^{(-C\lambda|X)}}\right).$$

To find $E\left(\mathrm{e}^{(C\lambda|X)}\right)$,

$$\begin{aligned}E\left(\mathrm{e}^{(C\lambda|X)}\right)&=\int_0^\infty \mathrm{e}^{(C\lambda,X)}\pi\left(\lambda,X\right)\mathrm{d}\lambda\\&=\frac{\sum_{i=1}^n X_i+\beta}{\Gamma(n+\alpha)}\int_0^\infty \mathrm{e}^{-\left(\sum_{i=1}^n X_i+\beta-C\right)}\lambda^{n+\alpha-1}\mathrm{d}\lambda\\&=\frac{\sum_{i=1}^n X_i+\beta}{\left(\sum_{i=1}^n X_i+\beta-C\right)^{n+\alpha}}.\end{aligned}$$

Here, in a similar manner, $E\left(\mathrm{e}^{(-C\lambda|X)}\right)$ where

$$E\left(\mathrm{e}^{(-C\lambda|X)}\right)=\frac{\sum_{i=1}^n X_i+\beta}{\left(\sum_{i=1}^n X_i+\beta+C\right)^{n+\alpha}}.$$

The Bayesian estimation of λ based on the composite LINEX loss function can thus be expressed as

$$\hat{\lambda}_{\mathrm{BCL}}=\frac{n+\alpha}{2C}\ln\left(\frac{\sum_{i=1}^n X_i+\beta+C}{\sum_{i=1}^n X_i+\beta-C}\right). \tag{5}$$

3. Model Selection Criterion

To compare the efficiency of MLE, SE, BEN, and BCL, estimators the mean square error, Akaike information criterion, and Bayesian information criterion methods were used to test their accuracy.

3.1. Mean Square Error (MSE)

The mean square error references the mean squared distance between observed and predicted values.

The MSE is thus calculated as

$$\mathrm{MSE}=\frac{\sum_{i=1}^n\left(\lambda_i-\hat{\lambda}_i\right)^2}{n}, \tag{6}$$

where $\hat{\lambda}$ is the estimator of the parameter λ on the i^{th} run and n is the sample size.

Estimates values with the lowest rates of MSE are preferred, as this means that $\hat{\lambda}$ is closer to the actual values of λ.

3.2. Akaike Information Criterion (AIC)

The Akaike information criterion (AIC) is defined by the equation

$$\mathrm{AIC} = 2K - 2\log L(\hat{\lambda}), \tag{7}$$

where K is the number of estimated parameters and $L(\hat{\lambda})$ is the maximum value of the likelihood estimate of the parameters.

Higher values for the likelihood function give a better fit, with the minimum AIC; however, the value of the AIC increases as more parameters are added to the first component.

In the case of small data set, $\frac{n}{K} < 40$, the second-order AIC, $\mathrm{AIC_c}$ can be used more effectively. The $\mathrm{AIC_c}$ takes the form

$$\mathrm{AIC_c} = -2\log(\hat{\lambda}) + 2K + \frac{2K+1}{n-K-1}, \tag{8}$$

where $\frac{2K+1}{n-K-1}$ is the bias-correction factor. As n increases, $\frac{2K+1}{n-K-1}$ tends to zero; at that point, the $\mathrm{AIC_c}$ gives results that more closely resemble the AIC.

3.3. Bayesian Information Criterion (BIC)

The Bayesian (or Schwarz) information criterion is expressed as

$$\mathrm{BIC} = -2\log L(\hat{\lambda}) + K\log(n). \tag{9}$$

Further information on AIC and BIC can be found in [14] and similar references.

The method with the smallest values of MSE, AIC, and BIC can, however, be assumed to be the most efficient method to estimate the parameter to estimate the exponential distribution, offering estimated values of λ close to its true value. The best method of estimating the parameter can also be determined by calculating which offers the highest log-likelihood value.

4. The Simulation Study

This section discusses the use of a Monte Carlo simulation to compare the MSE, AIC, and BIC criteria, as defined in (6), (7), and (9) respectively, to estimate the parameter of the exponential distribution based on the classical methods of comparison using MLE and Bayes estimators under the loss functions, including SE, BEN, and BCL, which can be computed as shown in Section 2 in equations 2 to 5. Sample sizes n = 10, 30, 150, 300, and 1,000 were used to achieve this, with data generated from the exponential distribution for the scale parameter $\lambda = 1$, with an arbitrary prior parameter (α, β, C) = (1, 2, 0.5). The number of replications used was 1,000,000 for each sample size.

The efficiencies of the estimation methods are compared in Table 1 and Table 2. The results of MSE, AIC, and BIC are presented in Table 1, while the 95% CIs, along with average confidence interval length. ALs, as computed using the estimators are displayed in Table 2. Table 1, which shows the estimated values

Table 1. Estimated value of the rate parameter, $\hat{\lambda}$, for the true parameter λ, mean square error (MSE), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), with log-likelihood (L) when (α, β, C) = (1, 2, 0.5).

	Estimators			
n = 10				
	MLE	BSE	BEN	BCL
$\hat{\lambda}$	1.033686	1.003013	0.9978051	1.160923
MSE	0.166474	0.07241902	0.07111619	0.072879
L	-6.350954	-4.9385	-13.81305	-10.41417
AIC	15.07691	12.25219	30.00111	23.20334
BIC	15.00449	12.17978	29.92869	23. 13093
n = 30				
$\hat{\lambda}$	0.9816438	0.991284	0.99964	1.005751
MSE	0.039308	0.030747	0.029941	0.030747
L	-13.1564	-15.2890	-24.64788	-13.21051
AIC	28.42004	32.68516	51.40291	28.52816
BIC	29.71409	33.9792	52.6969	29.8222
n = 150				
$\hat{\lambda}$	0.995469	0.998004	0.9965661	1.001133
MSE	0.0069195	0.006578301	0.006535639	0.0065790
L	-79.08826	-75.59041	-105.6675	-81.1737
AIC	160.1765	153.2011	213.3351	164.3474
BIC	163.1871	156.1915	216.3457	167.358
n = 300				
$\hat{\lambda}$	0.9949605	1.013197	0.9936484	0.9942531
MSE	0.003398617	0.0032999	0.003304556	0.003321389
L	-183.2983	-182.7589	-165.4214	-156.1926
AIC	368.5965	367.5179	332.8529	314.3853
BIC	372.3003	317.2217	336.5466	318.0891
n = 1000				
$\hat{\lambda}$	1.001877	1.000891	1.001926	0.9990892
MSE	0.00100733	0.000980377	0.000997429	0.000996724
L	-567.925	-604.1954	-555.8084	-541.6347
AIC	1137.85	1210.389	1113.617	1085.269
BIC	1142.758	1215.297	1118.524	1090.179

Table 2. The lower (L) the upper(U) and average length AL of 95% confidence intervals for λ of exponential distribution.

n	(L, U) AL estimators MLE	BSE	BEN	BCL
10	(0.7903452, 1.277027)	(0.83655, 1.169469)	(0.8462149, 1.149395)	(0.993913, 1.327934)
	0.4866818	0.332919	0.3031801	0.334021
30	(0.911776, 1.051511)	(0.928542, 1.054027)	(0.93897, 1.060325)	(0.9430085, 1.068498)
	0.139735	0.125485	0.121355	0.1254895
150	(0.9822032, 1.008736)	(0.9850244, 1.010984)	(0.9836738, 1.009458)	(0.988153, 1.014113)
	0.0265328	0.0259596	0.0257842	0.02596
300	(0.9878749, 1.001046)	(0.9999318, 1.012933)	(0.9871547, 1.000142)	(0.9877433, 1.000763)
	0.013197	0.0130012	0.0129873	0.0130197
1000	(0.9999109, 1.003843)	(0.9989327, 1.002849)	(0.9999695, 1.003882)	(0.9971325, 1.001046)
	0.0039321	0.0039163	0.0039125	0.0039135

and the values of MSE, AIC, and BIC with log-likelihoods for the selected sample size, also offers the MSE values obtained using all four estimators approached. For small sample sizes (n = 10, 30), the BEN method can thus be seen to perform better than other estimation methods, based on it offering the smallest value of MSE. With an increasing sample size, however, the BSE method provides the lowest MSE. The results further indicate that the lowest values of AIC and BIC were obtained by using BCL as n increases further, while the values of AIC and BIC are close to each other for all estimation methods. Overall, all methods offer good performance with respect to the estimation of the parameter, though the higher the log-likelihood value, the better the parameter estimation method, and the results also suggest that the BCL method gives a higher log-likelihood value than all other approaches. Comparing the results in this paper to that in [8], we see that both of these results show that the Bayesian method is better than the MLE. The results show that the MSE and L values of all methods decreased with increasing the sample size. The values of AIC and BIC increase as the sample size increases. For all estimation methods the 95% confidence intervals and average lengths for the parameter were computed. The narrow 95% band indicates that confidence levels are high. The results in Table 2 show that AL values decrease as the sample size increases, becoming more similar and closer to each other. The smallest values of ALs were found by applying the BEN method.

5. Conclusion

In this study, four methods of estimating the parameter of the exponential distribution were compared. Estimating the exponential parameter using classical

MLE was thus compared with the use of the Bayesian method assessed using three criteria, the MSE, the AIC, and the BIC. The results indicate that the Bayesian method performs better than the maximum likelihood estimator for the estimation of the parameter, based on it having the smallest values of MSE, AIC, and BIC, with narrow 95% CIs and the shortest ALs.

Acknowledgements

Sincere thanks to the members of JAMP for their professional performance.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Balakrishnan, K. (2019) Exponential Distribution: Theory, Methods, and Applications. Routledge, Milton Park, Abingdon-on-Thames. https://doi.org/10.1201/9780203756348

[2] Verdu, S. (1996) The Exponential Distribution in Information Theory. *Izdatel'stvo Nauka*, **32**, 100-111.

[3] Sarhan, A.M. (2003) Empirical Bayes Estimates in Exponential Reliability Model. *Applied Mathematics and Computation*, **135**, 319-332. https://doi.org/10.1016/S0096-3003(01)00334-4

[4] Yang, L., Zhou, H. and Yuan, S. (2013) Bayes Estimation of Parameter of Exponential Distribution under a Bounded Loss Function. *Research Journal of Mathematics and Statistics*, **5**, 28-31. https://doi.org/10.19026/rjms.5.5802

[5] Rahman, H., Roy, M.K. and Baizid, A.R. (2012) Bayes Estimation under Conjugate Prior for the Case of Power Function Distribution. *American Journal of Mathematics and Statistics*, **2**, 44-48. https://doi.org/10.5923/j.ajms.20120203.06

[6] Hasan, M.R. and Baizid, A.R. (2017) Bayesian Estimation under Different Loss Functions Using Gamma Prior for the Case of Exponential Distribution. *Journal of Scientific Research*, **9**, 67-78. https://doi.org/10.3329/jsr.v1i1.29308

[7] Cohen, A.C. (1965) Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and on Censored Samples. *Technometrics*, **7**, 579-588. https://doi.org/10.1080/00401706.1965.10490300

[8] Harter, H.L. and Moore, A.H. (1965) Maximum-Likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples. *Technometrics*, **7**, 639-643. https://doi.org/10.1080/00401706.1965.10490304

[9] Nwobi, F. and Ugomma, C. (2014) A Comparison of Methods for the Estimation of Weibull Distribution Parameters. *Advances in Methodology and Statistics*, **11**, 65-78. https://doi.org/10.51936/bddv8875

[10] Al-Bossly, A. (2021) E-Bayesian and Bayesian Estimation for the Lomax Distribution under Weighted Composite LINEX Loss Function. *Computational Intelligence and Neuroscience*, **2021**, Article ID 2101972. https://doi.org/10.1155/2021/2101972

[11] Ali, S., Aslam, M., Abbas, N. and Kazmi, S.M.A. (2012) Scale Parameter Estimation of the Laplace Model Using Different Asymmetric Loss Functions. *International*

Journal of Statistics and Probability, **1**, 105-127.
https://doi.org/10.5539/ijsp.v1n1p105

[12] Parsian, A. and Nematollahi, N. (1996) Estimation of Scale Parameter under Entropy Loss Function. *Journal of Statistical Planning and Inference*, **52**, 77-91.
https://doi.org/10.1016/0378-3758(95)00026-7

[13] Zhang, R. (2007) Estimating Parameters in the Composite Linex Loss of Symmetry. Master's Degree Paper of Dalian University of Technology, Dalian University of Technology, Dalian, China.

[14] Fabozzi, F.J., Focardi, S.M., Rachev, S.T. and Arshanapalli, B.G. (2014) The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications. John Wiley and Sons, Hoboken. https://doi.org/10.1080/14697688.2015.1080486