JAMBURA JOURNAL OF PROBABILITY AND STATISTICS
Volume 4 Nomor 2, November 2023
Journal Homepage: http://ejurnal.ung.ac.id/index.php/jps
DOI: https://doi.org/10.37905/jjps.v4i2.22549

BAYES ESTIMATOR OF EXPONENTIAL DISTRIBUTION PARAMETERS OF TYPE II CENSORED DATA WITH LINEAR EXPONENTIAL LOSS FUNCTION METHOD BASED ON JEFFREY PRIORS

Anggara Teguh Previan[1], Ardi Kurniawan[2*], M. Fariz Fadillah Mardianto[3], Sediono[4]
[1,2,3,4] Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

***e-mail**: *ardi-k@fst.unair.ac.id*

Abstract
Survival analysis is often used in the application of analyzing the survival of an object such as living things or objects. This analysis is identical to data censoring which is divided into three, namely: type I, II, and III censored data. Type II censored data is data censoring done by determining the number of objects to be analyzed (r) *from the total number of observation objects* (n)*. Type II censored data is used when the analysis is intended to maximize the results of the analysis. Bayesian Linear Exponential (LINEX) loss function is one method that can be used to estimate parameters in survival analysis by minimizing the expected value of LINEX. The purpose of this study is to determine the Bayesian LINEX loss function parameter estimation on type II censored data using exponential distribution. This method uses the concept of posterior distribution and prior distribution. The prior distribution used is the Jeffrey prior distribution which has objective properties and is based on Fisher information theory The application of the parameter estimation results is carried out on the survival data of lung cancer patients obtained from NCCTG. Based on the results of parameter estimation, it is concluded that the greater the value of the controller* (a) *can produce a smaller value of parameter estimation results* $(\hat{\theta})$*. The results of this study can be used as a reference in conducting survival tests using type II censored exponential distribution data using the LINEX loss function method based on Jeffrey priors.*
Keywords: *Bayesian Method, Exponential Distribution, LINEX loss function, Jeffey Prior*

1. INTRODUCTION

Survival analysis is a statistical analysis used to test the survival time of an object (Kurniawan, et al., 2023). Events that often use this analysis include: illness, death, recovery, or certain experiences that are interesting to observe in a particular object. This analysis has three types of censored data, namely censored type I, type II, and type III (Wang, et al., 2019). Type II censored data has high accuracy in event measurement, so that the analysis results can be maximized. The commonly used distribution is the exponential distribution because it is consistent with the survival graph of living things. The exponential distribution is one of the continuous distribution models with parameters θ (Mahdavi & Kundu, 2017). In survival analysis, the estimation process is also carried out. Estimation is a step used to obtain the value of a parameter in the distribution model. There are two approaches used to estimate parameters, namely the classical approach and the Bayesian approach (Saha, et al., 2019).

The Bayesian approach is used to obtain parameter estimates before knowing the sample data information. The estimation process is used by using sample data information (objective) and information from prior knowledge (subjective) about the distribution used and considering random variables that have a prior distribution. The Bayesian Linear Exponential (LINEX) loss function method is one of the methods used in the Bayesian

Copyright © 2023 Jambura Journal of Probability and Statistics
Received: 2023-10-23 | Accepted: 2023-11-25 | Available Online: 2020-11-30

approach (Al-Duais & Hmood, 2020). This method has instruments used in parameter estimation, namely the likelihood function, posterior distribution, and prior distribution. The prior distribution plays a role in estimating when the distribution in the data is still unknown, this distribution has two types, namely informative prior distribution and non-informative prior distribution (Ramos, et al., 2017). A non-informative prior distribution is used to estimate data for which the distribution is unknown (Banerjee & Seal, 2022). The Jeffrey prior is one example of a non-informative prior distribution (Yanuar, et al., 2019). There is previous research conducted by Afriani, et al. (2023) which discusses parameter estimation of the exponential distribution using the LINEX loss function method with Jeffrey priors on type I censored data. In addition, there is research that proves that the Bayesian LINEX loss function method is able to provide better parameter estimation results than using the Maximum Likelihood Estimator (MLE) method on survival data (Rizki, et al., 2017). Thus, a research update was conducted by estimating parameters using the Bayesian LINEX loss function method for type II censored data on the exponential distribution.

2. RESEARCH METHOD

Determination of the parameter estimation results of the exponential distribution on type II censored data using the Bayesian LINEX loss function method based on Jeffrey priors can be done with the following steps

1. Assume to observe n random samples $X_1, X_2, \dots, X_n$ and determine the amount of data to be observed (r), where $1 < r < n$.
2. Assume the survival data $T_1, T_2, \dots, T_{r-1}, T_r$ come from an exponential distribution with parameter θ.
3. Determining the PDF of the exponential distribution as in the following Equation
$$f(x) = \begin{cases} \theta e^{-\theta x}, & x \geq 0; \theta > 0 \\ 0, & x < 0 \end{cases} \tag{1}$$
(Sahoo, 2008).
4. Determining the survival function of the exponential distribution as in the following Equation
$$S(t) = 1 - F(t) = 1 - \left(1 - e^{-\theta x}\right) = e^{-\theta x} \tag{2}$$
5. Determine the form of the likelihood function $L(\theta)$ of the exponential distribution on type II censored data as in the following Equation
$$L(\theta) = \frac{n!}{(n-r)!} \cdot f(t_1) \cdot f(t_2) \cdot \dots \cdot f(t_r) \cdot \left(S(t_r)\right)^{n-r} \tag{3}$$
(Ren & Gui, 2021).
6. Determine the prior distribution based on the Jeffrey prior which is one type of non-informative prior distribution with the following procedure
   a. Calculating the value of $\ln\left(f(x;\theta)\right)$.
   b. Calculating the first derivative of $\ln\left(f(x;\theta)\right)$.
   c. Calculating the second derivative of $\ln\left(f(x;\theta)\right)$.
   d. Calculating Fisher information $I(\theta)$ as in Equation
$$I(\theta) = -E\left[\frac{\partial^2 \ln\left(f(x;\theta)\right)}{\partial^2 \theta}\right] \tag{4}$$

(He, et al., 2023).

e. Determining the prior distribution $P(\theta)$ as in the following Equation

$$P(\theta) \propto \sqrt{I(\theta)} \tag{5}$$

(He, et al., 2023).

7. Determining the posterior distribution $P(\theta|X)$ as in the following Equation

$$P(\theta|X) = \frac{L(\theta) \cdot P(\theta)}{\int_0^\infty L(\theta) \cdot P(\theta)\, d\theta} \tag{6}$$

8. Estimate parameters using the LINEX loss function method based on the posterior distribution $P(\theta|X)$ obtained. The estimation steps are as follows
   a. Calculating the LINEX loss function $L(\hat{\theta}, \theta)$.
   b. Determining the expected value of the posterior distribution of the LINEX loss function $E_\theta[L(\hat{\theta}, \theta)]$.
   c. Calculating the value of the Bayes estimator of the exponential distribution on type II censored data with the LINEX loss function method approach based on Jeffrey priors as in Equation

$$\hat{\theta}_L = -\frac{1}{a}\left(\ln\left(E_\theta\left[e^{-a\theta}\right]\right)\right) \tag{7}$$

where,

$$E_\theta\left[e^{-a\theta}\right] = \int_0^\infty e^{a\theta} P(\theta|X_i)\, d\theta \tag{8}$$

3. RESULTS AND DISCUSSION

3.1 Likelihood Function Censored Type II

Substituting Equations (1) and (2) into Equation (3) will give the form of the exponential distribution Likelihood function on type II censored data.

$$\begin{aligned} L(\theta) &= \frac{n!}{(n-r)!} f(t_1) \cdot f(t_2) \cdot \ldots \cdot f(t_r)\left(S(t_r)\right)^{n-r} \\ &= \frac{n!}{(n-r)!}\left(\prod_{i=1}^{r} \theta e^{-\theta t_i}\right)\left(e^{-\theta t_r}\right)^{n-r} \\ &= \frac{n!}{(n-r)!} \theta^r e^{-\theta[(\sum_{i=1}^{r} t_i) + t_r(n-r)]} \end{aligned}$$

$$L(\theta) = \frac{n!}{(n-r)!} \theta^r e^{-\theta T} \tag{9}$$

with $T = [(\sum_{i=1}^{r} t_i) + t_r(n-r)]$.

3.2 Jeffrey Prior Distribution

Determining the shape of the Jeffrey prior distribution begins with determining the value of Fisher information, which is the second derivative of the value of $\ln(f(x;\theta))$ where $f(x;\theta)$ is the CDF of the exponential distribution.

$$\ln\big(f(x;\theta)\big) = \ln\big(\theta e^{-\theta x}\big)$$
$$\ln\big(f(x;\theta)\big) = \ln(\theta) - \theta x \tag{10}$$

Next, the first derivative of Equation (10)

$$\frac{\partial\big(\ln\big(f(x;\theta)\big)\big)}{\partial\theta} = \frac{1}{\theta} - x \tag{11}$$

Thus, the second derivative can be determined by deriving Equation (11)

$$\frac{\partial^2\big(\ln\big(f(x;\theta)\big)\big)}{\partial\theta} = -\frac{1}{\theta^2} \tag{12}$$

Then, Fisher Information is obtained as follows

$$I(\theta) = \frac{1}{\theta^2} \tag{13}$$

After obtaining the Fisher Information value, the Jeffrey prior distribution can be obtained as follows

$$P(\theta) \propto \sqrt{I(\theta)}$$
$$P(\theta) \propto \frac{1}{\theta} \tag{14}$$

3.3 Posterior Distribution

The posterior distribution value can be obtained by substituting Equation (9) and (14) into Equation (6) as follows

$$P(\theta|X) = \frac{L(\theta)\cdot P(\theta)}{\int_0^\infty L(\theta)\cdot P(\theta)\,d\theta}$$
$$= \frac{\frac{n!}{(n-r)!}\theta^r e^{-\theta[(\sum_{i=1}^r t_i)+t_r(n-r)]}\cdot\left(\frac{1}{\theta}\right)}{\int_0^\infty \frac{n!}{(n-r)!}\theta^r e^{-\theta[(\sum_{i=1}^r t_i)+t_r(n-r)]}\cdot\left(\frac{1}{\theta}\right)d\theta}$$
$$P(\theta|X) = \frac{[(\sum_{i=1}^r t_i)+t_r(n-r)]^r\theta^{r-1}e^{-\theta[(\sum_{i=1}^r t_i)+t_r(n-r)]}}{\Gamma(r)} \tag{15}$$

3.4 Bayesian LINEX Loss Function Estimation Value

The Bayesian LINEX loss function parameter estimation is obtained by minimizing the Bayesian LINEX loss function expectation. The expectation can be obtained as follows

$$E_\theta\big[e^{-a\theta}\big] = \int_0^\infty e^{-a\theta}\cdot P(\theta|X_i)\,d\theta$$
$$= \int_0^\infty e^{-a\theta}\cdot\left(\frac{[(\sum_{i=1}^r t_i)+t_r(n-r)]^r\theta^{r-1}e^{-\theta[(\sum_{i=1}^r t_i)+t_r(n-r)]}}{\Gamma(r)}\right)d\theta$$
$$E_\theta\big[e^{-a\theta}\big] = \frac{[(\sum_{i=1}^r t_i)+t_r(n-r)]^r}{\big[(\sum_{i=1}^r t_i)+t_r(n-r)+a\big]^r} \tag{16}$$

Thus, the Bayesian LINEX loss function estimation value is obtained by substituting Equation (16) into Equation (7) as follows

$$\hat{\theta} = -\frac{1}{a}\left(\ln\left(E_\theta\left[e^{-a\theta}\right]\right)\right)$$

$$\hat{\theta} = -\frac{1}{a}\left(\ln\left(\frac{\left[\left(\sum_{i=1}^{r} t_i\right) + t_r(n-r)\right]}{\left[\left(\sum_{i=1}^{r} t_i\right) + t_r(n-r) + a\right]}\right)^r\right) \quad (17)$$

3.5 Implementing Parameter Estimation

Application of Bayesian LINEX loss function estimation on data obtained from the North Central Cancer Treatment Group (NCCTG) in 2020 regarding lung cancer survival data. Thus, the data used is data on lung cancer patients who died after treatment in patients using type II censored data with a total of 20 data from 22 observations made as in Table 1 below.

Table 1. Survival Time Data of Patient Lung Cancer

Patient	Survival Time (Days)	Patient	Survival Time (Days)
1	5	11	305
2	15	12	371
3	53	13	433
4	88	14	519
5	118	15	558
6	142	16	613
7	180	17	641
8	207	18	707
9	245	19	814
10	286	20	883

The survival data of lung cancer patients was tested for distribution using the Kolmogorov-Smirnov test to determine whether the survival data was exponentially distributed or not. The hypothesis used in the test is as follows

H_0 : Data is exponentially distributed

H_1 : Data is not exponentially distributed

with the critical region, reject H_0 when $P_{value} < 0.05$.

Based on the Kolmogorov-Smirnov test, it is concluded that the survival data of lung cancer patients is exponentially distributed because the P_{value} is 0.536 or greater than 0.05. After that, parameter estimation is carried out based on Equation (17) with the value of a in this test carried out in the range of $1 < a \leq 10$ to find out how the parameter estimation results on different measurement controllers.

Table 2. Parameter Estimation Results with Different Controllers

Controller (a)	Parameter Estimation $(\hat{\theta})$	Controller (a)	Parameter Estimation $(\hat{\theta})$
1	0.000970545	6	0.000161757
2	0.000485272	7	0.000138649
3	0.000323515	8	0.000121318
4	0.000242636	9	0.000107838
5	0.000194109	10	0.000097054

Based on Table 2, it can be seen that the greater the value of the controller (a) can produce a smaller value of parameter estimation results $(\hat{\theta})$ on the data used.

4. CONCLUSION

The formula for the parameter estimation equation of type II censored exponential distribution data using the LINEX loss function method based on Jeffrey priors is as follows

$$\hat{\theta} = -\frac{1}{a}\left(\ln\left(\frac{[(\sum_{i=1}^{r} t_i) + t_r(n-r)]}{[(\sum_{i=1}^{r} t_i) + t_r(n-r) + a]}\right)^r\right)$$

The application of the equation estimation results was carried out on secondary data obtained from NCCTG in 2020. Tests were carried out with different measurement controller values (a) and it was concluded that the the greater the value of the controller (a) can produce a smaller value of parameter estimation results $(\hat{\theta})$ on the data used.

REFERENCES

Afriani, Kurniawan, A. & Ana, E., 2023. Bayes Estimator Parameters Exponential Distribution of Type I Sensor Data Using Linear Exponential Loss Function Method Based on Prior Jeffrey. International Journal of Statistical Distributions and Applications, pp. 62-67. https://doi.org/10.11648/j.ijsd.20230902.12

Al-Duais, F. S. & Hmood, M. Y., 2020. Bayesian and non-Bayesian estimation of the Lomax model based on upper record values under weighted LINEX loss function. Periodicals of Engineering and Natural Sciences, 8(3), pp. 1786-1794. http://dx.doi.org/10.21533/pen.v8i3.1608

Banerjee, P. & Seal, B., 2022. Partial Bayes Estimation of Two Parameter Gamma Distribution Under Non-Informative Prior. Statistics, Optimization & Information Computing, 10(4), pp. 1110-1125. https://doi.org/10.19139/soic-2310-5070-1110

He, D., Sun, D. & Zhu, Q., 2023. Bayesian analysis for the Lomax model using noninformative priors. Statistical Theory and Related Fields, 7(1), pp. 61-68. https://doi.org/10.1080/24754269.2022.2133466

Kurniawan, A., Previan, A. T. & Nurrohmah, Z., 2023. SURVIVAL FUNCTION AND HAZARD FUNCTION ANALYSIS OF EXPONENTIAL DISTRIBUTION IN TYPE I CENSORED SURVIVAL DATA: A CASE STUDY OF BREAST CANCER PATIENTS. BAREKENG: Jurnal Ilmu Matematika dan Terapan, 17(3), pp. 1795-1802. https://doi.org/10.30598/barekengvol17iss3pp1795-1802

Mahdavi, A. & Kundu, D., 2017. A new method for generating distributions with an application to exponential distribution. *Communications in Statistics-Theory and Methods,* 46(13), pp. 6543-6557.https://doi.org/10.1080/03610926.2015.1130839

Ramos, P. L. et al., 2017. Bayesian analysis of the generalized gamma distribution using non-informative priors. Statistics, 51(4), pp. 824-843. https://doi.org/10.1080/02331888.2017.1327532

Ren, J. & Gui, W., 2021. Statistical analysis of adaptive type-II progressively censored competing risks for Weibull models. Applied Mathematical Modelling, Volume 98, pp. 323-342. https://doi.org/10.1016/j.apm.2021.05.008

Rizki, S. W., Mara, M. N. & Sulistianingsih, E., 2017. Survival Bayesian Estimation of Exponential-Gamma Under Linex Loss Function. s.l., IOP Publishing, p. 012036. https://doi.org/10.1088/1742-6596/855/1/012036

Saha, M., Dey, S., Yadav, A. S. & Kumar, S., 2019. Classical and Bayesian inference of C py for generalized Lindley distributed quality characteristic. Quality and Reliability Engineering International, 35(8), p. Quality and Reliability Engineering International. https://doi.org/10.1002/qre.2544

Sahoo, P., 2008. Probability and Mathematical Statistics. Louisville: Department of Mathematics University of Louisville.

Wang, P., Li, Y. & Reddy, C. K., 2019. Machine learning for survival analysis: A survey. ACM Computing Surveys (CSUR), 51(6), pp. 1-36. https://doi.org/10.1145/3214306

Yanuar, F., Yozza, H. & Rescha, R. V., 2019. Comparison of Two Priors in Bayesian Estimation for Parameter of Weibull Distribution. Science and Technology Indonesia, 4(3), pp. 82-87. https://doi.org/10.26554/sti.2019.4.3.82-87 .