CONFIDENCE INTERVAL USING MAXIMUM LIKELIHOOD ESTIMATION FOR THE PARAMETERS OF POISSON TYPE LENGTH BIASED EXPONENTIAL CLASS MODEL

[1]Rajesh Singh, [2]Preeti A. Badge, [3]Pritee Singh

•

[1,2] Department of Statistics, S. G. B. Amravati University, Amravati, India
[3]Department of Statistics, Institute of Science, Nagpur, India.
rsinghamt@hotmail.com
preetibadge10@gmail.com
priteesingh25@gmail.com

Abstract

In this research paper, Confidence interval using Maximum likelihood estimation is obtained for Poisson type Length biased exponential class for the parameters. Failure intensity, mean time to failure and likelihood function for the parameter is obtained. Confidence interval has been derived for parameters using maximum likelihood estimation. To study the performance of confidence interval, average length and coverage probability are calculated using Monte Carlo simulation technique. From the obtained intervals it is concluded that Confidence interval for the parameter perform better for appropriate choice of execution time and certain values of parameters.

Keywords: Length biased exponential distribution, Software reliability growth model, Maximum likelihood estimation (MLE), Confidence interval using MLE, Average length and coverage probability.

I. Introduction

In this research paper Poisson type length biased exponential class model is considered according to Musa and Okumoto [9] classification scheme. Seenoi et al [12] proposed length biased exponentiated invented weibull distribution including some probability functions and moments of this distribution. Mir et al [6] introduced a length biased Beta distribution and also given a test for detection of length biasedness of beta distribution. The exponential exponentiated distribution proposed by Gupta and Kundu [4] which is special case of the exponentiated Weibull family. Mudholkar and Shrivastava [7] proposed the exponentiated weibull distribution as an extension of the weibull family obtained by adding the second shape parameter. Gupta and Keating [3] developed relationship between the survival function, the failure rate and mean residual life of exponential distribution and its length biased form.

Mudholkar et al [8] applied exponentiated weibull distribution to serve survival data and showed those hazard rates are increasing, decreasing bathtub shape and unimodal. Neppala et al [11] proposed Pareto type II based software reliability growth model with interval domain data using maximum likelihood estimation to estimate the parameter. Singh et al [13] proposed Bayes estimators for length biased distribution compared with ML estimators. Cunha and Rao [1] estimated credible interval and confidence interval through MLE for lognormal distribution also compared average length and coverage probability of the calculated interval. In the field of software reliability most of the work done on point estimation which give single guess value. Interval estimation with confidence interval gives more information than a point estimate. Confidence interval will be derived for both finite and infinite failure type models. Interval estimate indicates the error related to point estimate by the extent of its range and by probability of the true population parameter lying within that range. Thus, the purpose of this research paper is to study confidence interval using maximum likelihood estimation.

The association of the paper is such that section II presents length biased exponential model and derivation of failure intensity and expected number of failures using Length biased exponential distribution. Section III presents Likelihood function and derivation of maximum likelihood estimates of the parameters. Section IV contains derivation of confidence limits for the parameters θ_0 and θ_1 using maximum likelihood estimation. Results are discussed in the section V while concluding remarks are provided in section VI.

II. Model Formulation and Evaluation

Consider that software is tested for its performance and observed the time of failure occurs during software system performance. Let the number of failures present in software be θ_0,and t_e be the execution time i.e. time during which CPU is busy and m_e be the number of failures observed up to execution time t_e. Consider that time between the failures t_i (i=1,2,………m_e) follows the exponential distribution with parameter θ_1. The length biased exponential distribution is given as

$$f^*(t) = \begin{cases} t\theta_1^2 e^{-\theta_1 t} & , t > 0, \theta_1 > 0, E[t] \neq 0 \\ 0 & otherwise \end{cases} \tag{1}$$

Where $f^*(t)$ denotes the length biased exponential distribution.

The failure intensity function is obtained by using equation (1)

$$\lambda(t) = \theta_0 t\theta_1^2 e^{-\theta_1 t}, t > 0, \theta_0 > 0 \tag{2}$$

Where, θ_0 express the number of failures and θ_1 express the for failure rate.

The mean failure function i.e. expected number of failures at time t_e can be obtained by using equations (2) and given by:

$$\mu(t_e) = \theta_0\theta_1^2 I_1 \tag{3}$$

Where, $I_1 = \int_0^{t_e} t_i e^{-\theta_1 t_i}\, dt$ and by solving (see Gradshteyn and Ryzhik [2]) we get,

$$\mu(t_e) = \theta_0\left[1-(1+\theta_1 t_e)e^{-\theta_1 t_e}\right]\ , t>0, \theta_0>0, \theta_1>0 \tag{4}$$

The study of behavior of failure intensity and expected number of failure of length biased exponential class model has been done by Singh et al [13]. They have compared the MLE's and Bayesian estimators on the basis of risk efficiencies.

III. Maximum Likelihood Estimation

Maximum likelihood estimation is most preferable because of its easy computation, greater efficiency and better numerical stability. It requires likelihood function for estimation. The likelihood function of parameters $\theta_0$and θ_1 is obtained with the help of failure intensity (2) and expected number of failures (4) (see for details Musa et al [10]) given by:

$$L(\theta_0,\theta_1) = \theta_0^{m_e}\theta_1^{2m_e}\left[\prod_{i=1}^{m_e} t_i\right]e^{-T\theta_1}e^{-\theta_0\left[1-(1+\theta_1 t_e)e^{-\theta_1 t_e}\right]} \tag{5}$$

Where, $\sum_{i=1}^{m_e} t_i = T$

After taking the logarithm of both sides of above equation and applying the procedure of obtaining the MLE's for parameters $\theta_0$andθ_1, the MLE's are

$$\hat{\theta}_{m0} = \left[\frac{m_e}{\left(1-(1+\hat{\theta}_{m1}t_e)e^{-\hat{\theta}_{m1}t_e}\right)}\right] \tag{6}$$

and

$$\hat{\theta}_{m1} = \left[\frac{(2m_e - T\theta_{m1})e^{\hat{\theta}_{m1}t_e}}{\hat{\theta}_{m0}t_e^2}\right]^{1/2} \tag{7}$$

respectively.

The values of $\hat{\theta}_{m0}$ and $\hat{\theta}_{m1}$ can be obtained by solving simultaneous equations (6) and (7) using any available standard numerical method viz. Bisection Method, Newton Rapson method. Singh et al [13] obtained maximum likelihood estimates for parameters of length biased exponential model. They compared maximum likelihood estimates and Bayes estimates on the basis of risk efficiencies and concluded that Bayes estimates preferred over maximum likelihood estimates.

IV. Confidence Interval using maximum likelihood estimation

Now to obtain confidence interval for both the parameter, it requires variance-covariance matrix for Σ all the MLE. Variance-covariance matrix is derived using Fisher information matrix. For asymptotic variance we can calculate Fisher information matrix which is negative second partial derivative of log likelihood function (see for details Kale [5]). Second derivative of log likelihood function can be given as follows:

$$\text{Var}\ (\hat{\theta}_0) = \frac{\hat{\theta}_0^2}{me} \tag{8}$$

$$\text{Var}\ (\hat{\theta}_1) = [(1/2m_e + \theta_0\theta_1^2 e^{-\hat{\theta}_1 t_e}t_e^2 - \theta_0\theta_1^3 e^{-\hat{\theta}_1 t_e}t_e^3)] \tag{9}$$

Using equation (8) confidence limits for parameter θ_0 is given by:

$$\hat{\theta}_{0L} = \hat{\theta}_0 + Z_{\alpha/2}(\hat{\theta}_0/(m_e)^{1/2}) \tag{10}$$

$$\hat{\theta}_{0U} = \hat{\theta}_0 - Z_{\alpha/2}(\hat{\theta}_0/(m_e)^{1/2}) \tag{11}$$

For parameter θ_1 confidence limits using equation (9) are given by:

$$\hat{\theta}_{1L} = \hat{\theta}_1 - Z_{\alpha/2}[(1/2m_e + \hat{\theta}_0\hat{\theta}_1^2 e^{-\hat{\theta}_1 t_e} t_e^2 - \hat{\theta}_0\hat{\theta}_1^3 e^{-\hat{\theta}_1 t_e} t_e^3)]^{1/2} \tag{12}$$

$$\hat{\theta}_{1U} = \hat{\theta}_1 + Z_{\alpha/2}[(1/2m_e + \hat{\theta}_0\hat{\theta}_1^2 e^{-\hat{\theta}_1 t_e} t_e^2 - \hat{\theta}_0\hat{\theta}_1^3 e^{-\hat{\theta}_1 t_e} t_e^3)]^{1/2} \tag{13}$$

By substituting tabulated values of Z_α, 95% confidence interval can be obtained.

V. Discussion and Results

Here, 95% confidence interval using maximum likelihood estimation is obtained for the parameters θ_0 i.e. total number of failures and θ_1. Confidence interval for the proposed length biased exponential model is calculated as defined equation (1). To study the performance of confidence interval, sample size m_e was generated up to execution time t_e and it was repeated 1000 times from the length biased exponential distribution for distinct values of θ_0 and θ_1.Using Monte Carlo simulation technique 95% confidence interval has been obtained. The values of average length and coverage probability have been obtained by assuming execution time$t_e(= 15,20,25,30,35)$, and parameters$\theta_0(= 24(2)32)$, and$\theta_1(= 0.02(0.02)0.1)$. Average length and coverage probability obtained for confidence interval has been summarized in the tables 1to 10

Table 1 to 5 represents the 95% confidence interval for the parameter θ_0. From the table, it is seen that as values of parameter θ_0 increases the calculated average length decreases and there is slight increase in calculated average length as values of parameter θ_1 increases. From table it can be seen that values of coverage probability decreases as θ_0 increases and as θ_1 increases coverage probability increases. Similarly, it can also be observed that as execution time t_e increases average length increases. As average length decreases it effects on coverage probability. Coverage probability decreases as average length decreases and it increases as execution time increases.

The table 6 to 10 represents the 95% confidence interval for the parameter θ_0. From the table, it is seen that the values of average length increases as θ_0 increases and average length increases slightly as θ_1 increases. From table it can be seen that values of coverage probability increases as θ_0 increases and as θ_1 increases coverage probability also increases. Similarly, it can also be observed that execution time t_e increases average length decreases and coverage probability also decreases as execution time increases.

Table 1: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 15*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.16894 (0.995)	56.38039 (0.995)	55.55027 (0.994)	54.65604 (0.994)	53.83833 (0.992)
0.04	57.222518 (0.995)	56.39313 (0.995)	55.63409 (0.995)	54.70139 (0.994)	54.12473 (0.993)
0.06	57.233046 (0.995)	56.42863 (0.994)	55.65794 (0.995)	54.90752 (0.994)	54.16633 (0.993)
0.08	57.233615 (0.995)	56.45414 (0.995)	55.67368 (0.995)	54.92776 (0.994)	54.23347 (0.993)
0.1	57.243652 (0.995)	56.50778 (0.995)	55.71389 (0.995)	54.93576 (0.994)	54.27069 (0.993)

*The values in the parenthesis are coverage probability.

Table 2: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 20.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.605773 (0.995)	57.18747 (0.995)	56.78515 (0.994)	56.38402 (0.994)	55.94543 (0.992)
0.04	57.612651 (0.995)	57.21467 (0.995)	56.84414 (0.994)	56.41711 (0.994)	56.02504 (0.993)
0.06	57.61924 (0.995)	57.23135 (0.995)	56.85961 (0.994)	56.44942 (0.9944)	56.07487 (0.993)
0.08	57.62179 (0.995)	57.2444 (0.995)	56.89678 (0.994)	56.46993 (0.994)	56.09547 (0.993)
0.1	57.63312 (0.995)	57.26667 (0.995)	56.90905 (0.995)	56.53445 (0.995)	56.18467 (0.994)

*The values in the parenthesis are coverage probability.

Table 3: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 25.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.687913 (0.995)	57.35914 (0.995)	57.05328 (0.995)	56.72293 (0.994)	56.39379 (0.993)
0.04	57.694764 (0.995)	57.38959 (0.995)	57.05003 (0.995)	56.73697 (0.994)	56.42135 (0.994)
0.06	57.70429 (0.996)	57.38601 (0.995)	57.07483 (0.995)	56.77586 (0.994)	56.47019 (0.994)
0.08	57.725421 (0.996)	57.38709 (0.995)	57.09107 (0.995)	56.77883 (0.994)	56.48683 (0.994)
0.1	57.728669 (0.996)	57.40079 (0.995)	57.11992 (0.995)	56.78598 (0.995)	56.50296 (0.994)

*The values in the parenthesis are coverage probability.

Table 4: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 30.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.757765 (0.996)	57.47327 (0.995)	57.18907 (0.995)	56.92065 (0.994)	56.70381 (0.994)
0.04	57.760038 (0.996)	57.48198 (0.995)	57.19682 (0.995)	56.95059 (0.994)	56.72307 (0.994)
0.06	57.760261 (0.996)	57.48631 (0.995)	57.21524 (0.995)	56.96878 (0.995)	56.73889 (0.994)
0.08	57.765491 (0.996)	57.49554 (0.995)	57.23952 (0.995)	57.01431 (0.995)	56.75868 (0.994)
0.1	57.765513 (0.997)	57.53795 (0.996)	57.27481 (0.995)	57.04687 (0.995)	57.02576 (0.995)

*The values in the parenthesis are coverage probability.

Table 5: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M0}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e = 35.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	57.784104 (0.997)	57.53507 (0.996)	57.33528 (0.995)	57.04729 (0.995)	56.85084 (0.994)
0.04	57.786394 (0.997)	57.54886 (0.996)	57.34892 (0.995)	57.06216 (0.995)	56.87181 (0.994)
0.06	57.793615 (0.997)	57.55129 (0.996)	57.35886 (0.995)	57.07193 (0.995)	56.87427 (0.994)
0.08	57.797047 (0.997)	57.57595 (0.996)	57.35927 (0.995)	57.08127 (0.995)	56.88975 (0.994)
0.1	57.798991 (0.998)	57.58077 (0.996)	57.37618 (0.995)	57.13901 (0.995)	57.00542 (0.995)

*The values in the parenthesis are coverage probability.

Table 6: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e = 15.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.023096 (0.996)	0.023270 (0.996)	0.023425 (0.996)	0.023611 (0.997)	0.023803 (0.997)
0.04	0.023100 (0.996)	0.023273 (0.996)	0.023426 (0.996)	0.023648 (0.997)	0.023842 (0.997)
0.06	0.023103 (0.996)	0.023279 (0.996)	0.023448 (0.996)	0.023651 (0.997)	0.023856 (0.997)
0.08	0.023103 (0.996)	0.023297 (0.996)	0.023461 (0.996)	0.023671 (0.997)	0.023865 (0.997)
0.1	0.023107 (0.996)	0.023298 (0.996)	0.023488 (0.996)	0.023691 (0.997)	0.023871 (0.998)

*The values in the parenthesis are coverage probability.

Table 7: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 20.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.019317 (0.995)	0.019425 (0.995)	0.019531 (0.995)	0.019644 (0.996)	0.019751 (0.996)
0.04	0.019321 (0.995)	0.019431 (0.995)	0.019535 (0.995)	0.019657 (0.996)	0.019779 (0.996)
0.06	0.019321 (0.995)	0.019435 (0.995)	0.019546 (0.995)	0.019664 (0.996)	0.019786 (0.996)
0.08	0.019323 (0.995)	0.019440 (0.995)	0.019551 (0.995)	0.019679 (0.996)	0.019802 (0.996)
0.1	0.019325 (0.995)	0.019448 (0.995)	0.019568 (0.995)	0.019685 (0.996)	0.019826 (0.996)

*The values in the parenthesis are coverage probability.

Table 8: *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 25.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.016918 (0.994)	0.017003 (0.995)	0.017076 (0.995)	0.017164 (0.995)	0.017245 (0.995)
0.04	0.016919 (0.994)	0.017006 (0.995)	0.017083 (0.995)	0.017166 (0.995)	0.017262 (0.995)
0.06	0.016924 (0.994)	0.017008 (0.995)	0.017088 (0.995)	0.017167 (0.995)	0.017269 (0.995)
0.08	0.016927 (0.995)	0.017010 (0.995)	0.017094 (0.995)	0.017177 (0.995)	0.017276 (0.994)
0.1	0.016928 (0.995)	0.017013 (0.995)	0.017093 (0.995)	0.017181 (0.995)	0.017284 (0.996)

*The values in the parenthesis are coverage probability.

Table 9 : *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 30.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.015232 (0.994)	0.015288 (0.994)	0.015349 (0.994)	0.015427 (0.994)	0.015466 (0.994)
0.04	0.015235 (0.994)	0.015298 (0.994)	0.015363 (0.994)	0.015431 (0.994)	0.015473 (0.994)
0.06	0.015236 (0.994)	0.015301 (0.994)	0.015364 (0.994)	0.015434 (0.994)	0.015477 (0.994)
0.08	0.015237 (0.994)	0.015303 (0.994)	0.015368 (0.994)	0.015442 (0.994)	0.015481 (0.994)
0.1	0.015256 (0.994)	0.015328 (0.994)	0.015372 (0.994)	0.015467 (0.994)	0.015489 (0.994)

*The values in the parenthesis are coverage probability.

Table 10 : *Average length and coverage probability of 95% confidence interval $\hat{\theta}_{M1}$ calculated for the different values parameters θ_0 = (24:2:32) and θ_1 = (0.02:0.02:0.1) and t_e= 35.*

θ_1 \ θ_0	**24**	**26**	**28**	**30**	**32**
0.02	0.013966 (0.993)	0.01402 (0.993)	0.014061 (0.993)	0.014124 (0.993)	0.014167 (0.994)
0.04	0.013967 (0.993)	0.014014 (0.993)	0.014064 (0.993)	0.014129 (0.993)	0.014171 (0.994)
0.06	0.013968 (0.993)	0.014023 (0.993)	0.014067 (0.993)	0.014136 (0.993)	0.014176 (0.994)
0.08	0.013969 (0.993)	0.014029 (0.993)	0.014076 (0.993)	0.014141 (0.993)	0.014184 (0.994)
0.1	0.013972 (0.993)	0.014033 (0.993)	0.014086 (0.993)	0.014153 (0.994)	0.014186 (0.994)

*The values in the parenthesis are coverage probability.

VI. Conclusion

Confidence interval through MLE is derived in this paper, the confidence interval suggested for the parameters of Poisson type length biased exponential class as SRGM. From the computation and above discussion it is concluded that proposed confidence interval maintained high coverage probability for different values of parameters for fixed execution time. Confidence interval for parameter θ_0 can be preferred for small execution time whereas Confidence interval for parameter θ_1 can be preferred for large execution time.

References:

[1] D'Cunha, J. G. and Rao, K. A. (2016). Frequentist Comparison of the Bayesian Credible and Maximum Likelihood Confidence interval for the Median of the Lognormal Distribution for the Censored Data.*International Journal of Scientific and Research Publications*, 6:61-66.

[2] Gradshteyn, I. S and Ryzhik, I. M. Table of Integrals, Series, and Products, Alan Jeffrey (editor) 5th Ed., Academic Press, New York, (1994).

[3] Gupta, R. C. and Keating, J. P. (1986).Relations for reliability measures under length biased sampling. *Scand Journal of Statistics*, 13: 49-56.

[4] Gupta, R. D. and Kundu, D. (2009). A new class of weighted exponential distribution. Statistics, 43:621 - 634.

[5] Kale, B. K., A First Course on Parametric Inference (Second Edition). Narosa, (1999).

[6] Mir, K. A., Ahmad, A. Reshi, J. A. (2013). Structural properties of length biased Beta distribution of first kind. *American Journal of Engineering Research,* 2:1-6.

[7] Mudholkar, G.S. and Shrivastava, D. K. (1993). ExponentiatedWeibull Family for Analyzing Bathtub Failure Rate Data.*IEEE Transanctions on Reliability*, 42:299-302.

[8] Mudholkar, G.S., Shrivastava, D. K., Kollia G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data.*Journal of the American Statistical Association,* 91:1575-1585.

[9] Musa, J.D. and Okumoto, K.(1984). A logarithmic Poisson execution time model for software reliability measurement. *Proc. 7th International Conference on Software Engineering*, Orlando, Florida, 230–238.

[10] Musa, J. D., Lannino, A. and Okumoto, K. Software Reliability: Measurement, Prediction, Application, New York, McGraw-Hill, (1987).

[11] Neppala, G. Satya Prasad, R. Kantam, R.R.L. (2011).Software Reliability Growth Model using Interval Domain Data. *International Journal of Computer Applications*, 34: 5-9.

[12] Seenoi, P. Supapakorn, T., Bodhiswan, W. (2014). The length-biased exponetiated inverted Weibull distribution. *International Journal of Pure and Applied Mathematics*, 92:191-206.

[13] Singh, R. Singh, P. and Kale, K. (2016). Bayes Estimators for the parameters of Poisson type length biased exponential class model using non- informative priors. *Journal of Reliability and Statistical Studies*, 9: 21-28.