Research Article

On Random Censored Exponential Distribution

Chris Bambey Guure

Institute for Mathematical Research, University Putra Malaysia, 43400, Serdang, Selangor, Malaysia
Bolgatanga Senior High School, Mail Box 176, Bolgatanga, Ghana

Abstract: The Exponential distribution has attracted the attention of statisticians working on theory and methods as well as in various fields of lifetime data analysis. In this study, we employ gamma non-informative prior and generalised (data-dependent) non-informative prior proposed by Guure and Ibrahim (2012) using squared error loss function. The Bayesian estimate of the scale parameter of the Exponential distribution is obtained by making use of Lindley's approximation procedure and compared with the classical maximum likelihood estimator. Mean Squared Error (MSE) and the absolute bias of the estimators are determined via simulation study for the purpose of comparison. It has been observed from the simulation study that, Bayes estimator with the generalised non-informative prior outperformed the gamma non-informative prior and the classical maximum likelihood estimator.

Keywords: Bayesian, exponential distribution, gamma and generalised non-informative priors, random censoring, simulation study

INTRODUCTION

Exponential distribution is the most exploited distribution for life-time data analysis. However, its suitability is restricted to a constant hazard rate, which is difficult to justify in many practical problems. This led to the development of alternative models for life-time data. A number of distributions such as Weibull and gamma have been extensively used for analysing life-time data, particularly, in those situations where the hazard rate is monotonically increasing or decreasing.

Censoring is a feature that is recurrent in lifetime and reliability data analysis, it occurs when exact lifetimes or run-outs can only be collected for a portion of the inspection units.

According to Horst (2009), "A data sample is said to be censored when, either by accident or design, the value of the variables under investigation is unobserved for some of the items in the sample".

Guure et al. (2012a) studied Bayesian estimation of two-parameter Weibull distribution using extension of Jeffreys' prior information with three loss functions. Al-Kutubi and Ibrahim (2009), studied Jeffery prior information to get the modify Bayes estimator for exponential distribution and compared it with standard Bayes estimator and maximum likelihood estimator to find the best (less MSE and MPE). The extension of Jeffery prior information with new loss function for estimating the parameter exponential distribution of life time is presented by Alkutubi and Ibrahim (2009).

Elfessi and Reineke (2001) studied the relationship between Bayesian and classical estimation approach using exponential distribution. He showed how the classical and the Bayesian estimators could be obtained with regards to the exponential distribution parameter.

Al-Bayyati (2002), further studied the problem of estimating parameters of the Weibull distribution and the reliability function where there is lack or insufficient information about the parameters.

An important element, in the point estimation problem, is the specification of the loss function. The most popular loss function used in the estimation problem is the Squared Error Loss Function (SELF), which can be easily justified on the grounds of minimum variance-unbiased estimation.

However, the weakness of this loss function is that it is symmetrical and gives equal weight to the overestimation and underestimation of the same magnitude. A number of asymmetric loss functions are available in the statistical literature and one of the most widely used asymmetric loss function is the LINEX loss function, originally proposed by Varian (1974) and popularized by Zellner (1986). But it has been pointed out by various authors that LINEX loss function is not as appropriate for the estimation of the scale parameter as it is for the location parameter.

Exponential distribution is the most commonly used model in reliability and life-testing analysis. In this study we present and compare Bayes estimator using gamma and generalised non-informative priors for the estimation of the exponential scale parameter with the squared error loss function against the classical maximum likelihood estimator.

Maximum Likelihood Estimation: Let $(t_1, \ldots, t_n)$ be the set of n random lifetimes from exponential

distribution with parameter θ. The probability density function of exponential distribution is given by:

$$f(t,\theta) = \theta^{-1} \exp\left(-\frac{t}{\theta}\right)$$

(1)

The survival function is:

$$S(t,\theta) = \exp\left(-\frac{t}{\theta}\right)$$

(2)

The likelihood function with respect to right (random) censored data is:

$$L(t_i,\theta,\delta_i) = \prod_{i=1}^{n} \left[f(t_i,\theta)\right]^{\delta_i} \left[S(t_i,\theta)\right]^{1-\delta_i}$$

Implying that:

$$L(t_i,\theta,\delta_i) = \prod_{i=1}^{n} \left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta_i} \left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta_i}$$

(3)

where, $\delta = 1$ is for failure and $\delta = 0$ for censored observation and $S(.)$ is the survival function. To obtain the maximum likelihood estimator for the unknown parameter, we take the natural log of the likelihood function, differentiate with respect to the parameter θ and equate to zero as follows:

$$\frac{\partial \ell}{\partial \theta} = -\frac{\sum_{i=1}^{n}\delta_i}{\theta} + \frac{\sum_{i=1}^{n}t_i}{\theta^2} = 0$$

(4)

Hence,

$$\hat{\theta} = \left(\frac{\sum_{i=1}^{n}t_i}{\sum_{i=1}^{n}\delta_i}\right)$$

Bayesian Analysis: The Bayesian approach has received a lot of attention for analysing failure data and other time-to-event data and has often being proposed as a suitable alternative to the traditional statistical methods. The Bayesian approach to lifetime data analysis allows prior subjective knowledge on lifetime parameters and technical information on the failure mechanism, as well as experimental data, to be incorporated into the inferential procedure. Hence Bayesian methods usually require less sample data to achieve the same quality of inferences than methods

based on sampling theory, which becomes extremely important in case of expensive testing procedures. Often there exists in failure data analysis some prior information concerning the parameters of the model; such knowledge can be translated into a prior density.

The information summarized by this prior density may be either subjective (i.e., based on individual's experience, belief and preference) or objective (i.e., based on test data from a comparable experiment and technical information) or both. When reliable prior subjective knowledge or technical information concerning the parameter is not available, or an inference from the data at hand is desired, it may be suitable to resort to the use of a default (or non-informative) prior density, (Al-Aboud, 2009).

Gamma prior: Let the scale parameter of the Exponential distribution take a Gamma (a, b) prior such that a>0 and b>0 and $v(\theta)$ represent the prior:

$$Gamma(a,b) \propto \frac{b^a}{\Gamma a} \theta^{a-1} \exp(-\theta b)$$

The joint posterior density function is:

$$\pi(t;\theta) = v(\theta)\prod_{i=1}^{n}\left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta}\left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta}$$

(5)

and the marginal distribution function is:

$$\int v(\theta)\prod_{i=1}^{n}\left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta}\left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta} d\theta$$

(6)

Therefore, the posterior distribution under squared error loss function is given as:

$$\pi^*(t;\theta) \propto \frac{\int u(\theta)v(\theta)\prod_{i=1}^{n}\left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta}\left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta} d\theta}{\int v(\theta)\prod_{i=1}^{n}\left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta}\left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta} d\theta}$$

(7)

Generalised non-informative prior: According to Guure and Ibrahim (2012), the generalised non-informative prior is solely data dependent and is given with respect to the scale parameter of the exponential distribution as:

$$v_1(\theta) = \left(\frac{1}{\hat{\theta}}\right)^c, c > 0$$

(8)

The posterior distribution according to the generalised prior is:

$$\pi^*(t;\theta) \propto \frac{\int u(\theta)v_1(\theta)\prod_{i=1}^{n}\left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta}\left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta}d\theta}{\int v_1(\theta)\prod_{i=1}^{n}\left[\theta^{-1}\exp\left(-\frac{t}{\theta}\right)\right]^{\delta}\left[\exp\left(-\frac{t}{\theta}\right)\right]^{1-\delta}d\theta} \quad (9)$$

Since the posterior distributions with respect to the two priors contain some ration of integrals, we propose to use a numerical approximation approach known as Lindley approximation procedure to reduce the ratio of integrals.

With the one-parameter exponential distribution, the Lindley approach is given as:

$$\hat{\theta} = u + \frac{1}{2}(u_{11}\sigma_{11}) + u_1\rho_1\sigma_{11} + \frac{1}{2}\left(\ell_{30}u_1\sigma_{11}^2\right) \quad (10)$$

with l being the log-likelihood equation given in (4) Therefore,

$$\rho = \ln(v(\theta)) \text{ and } \rho_1 = -\left(\frac{1}{\theta}\right)$$

$$u = \theta, u_1 = 1, u_{11} = 0.$$

$$\ell_{20} = \frac{\sum_{i=1}^{n}\delta_i}{\theta^2} - 2\frac{\sum_{i=1}^{n}t_i}{\theta^3}, \quad \sigma_{11} = (-\ell_{20})^{-1}$$

$$\ell_{30} = -2\frac{\sum_{i=1}^{n}\delta_i}{\theta^3} + 6\frac{\sum_{i=1}^{n}t_i}{\theta^4}$$

Considering the generalised non-informative prior, $v(\theta)$ is substituted for $v\_1(\theta)$.

Simulation study: A simulation study is carried out to determine the best estimator for the scale parameter of the exponential distribution with 40% random censoring, taking into consideration a sample size of n = 25, 50 and 100. The coding and the analysis are performed using the R programming language which is freely available on line. The exponential parameter is estimated with maximum likelihood and Bayesian using gamma non-informative prior and generalised non-informative prior approach. To obtain gamma non-informative prior, we make an assumption such that the hyper-parameters of the gamma distribution equal zero, i.e., a = b = 0. The values for the generalised non-informative prior were chosen to be c = 0.4, 0.8 and 1.6 without loss of generality. The scale parameter θ = 0.5, 1.0 and 1.5. These were iterated (R) 1000 times. The mean squared error values and that of the absolute bias are determined and presented below for the purpose of comparison. The mean squared error and absolute bias are given respectively as:

$$MSE(\hat{\theta}) = \frac{\sum_{r=1}^{R}(\hat{\theta}^r - \theta)^2}{R-1} \text{ and } Abs(\hat{\theta}) = \frac{\sum_{r=1}^{R}|\hat{\theta}^r - \theta|}{R-1}$$

CONCLUSION

Bayes gamma non-informative prior and the generalised non-informative prior estimators of the exponential parameter are obtained using squared error loss function by employing Lindley approximation. Comparisons are made between the estimators using mean squared error and absolute bias based on simulation study. The performance of MLE and Bayes using the non-informative priors were examined and the following conclusions made.

Table 1, shows the mean squared error values of the exponential parameter. It's been observed that the generalised Bayes non-informative prior estimator has the smallest mean squared error values than the others. But at θ = 0.5 and 1.0 with n = 25 and θ = 1.0 with n = 50 that, maximum likelihood estimator gave the smallest mean squared error. We must state that the generalised non-informative prior performed better than the gamma non-informative prior.

Table 1: Mean squared error for $\hat{\theta}$ of the estimators

Size	θ	$\hat{\theta}_{MLE}$	$\hat{\theta}_{BG}$	$\hat{\theta}_{BGN}$ c = 0.4	c = 0.8	c = 1.6
25	0.5	0.3674	0.4318	0.3312	0.3805	0.4470
	1.0	0.0290	0.0379	0.0330	0.0380	0.0396
	1.5	0.0149	0.0133	0.0132	0.0130	0.0135
50	0.5	0.1623	0.1754	0.1549	0.1676	0.1761
	1.0	0.0107	0.0124	0.0119	0.0126	0.0128
	1.5	0.0071	0.0066	0.0064	0.0063	0.0062
100	0.5	0.0750	0.0779	0.0748	0.0779	0.0800
	1.0	0.0043	0.0046	0.0042	0.0046	0.0046
	0.5	0.0034	0.0032	0.0031	0.0032	0.0033

Table 2: Absolute bias for $\hat{\theta}$ of the estimators

Size	θ	$\hat{\theta}_{MLE}$	$\hat{\theta}_{BG}$	$\hat{\theta}_{BGN}$		
				c = 0.4	c = 0.8	c = 1.6
25	0.5	0.1126	0.1217	0.1089	0.1150	0.1235
	1.0	0.0268	0.0311	0.0297	0.0305	0.0313
	1.5	0.0221	0.0205	0.0208	0.0204	0.0208
50	0.5	0.0551	0.0572	0.0539	0.0559	0.0573
	1.0	0.0126	0.0137	0.0133	0.0137	0.0138
	1.5	0.0113	0.0108	0.0106	0.0106	0.0104
100	0.5	0.0269	0.0274	0.0268	0.0274	0.0278
	1.0	0.0060	0.0062	0.0061	0.0063	0.0063
	1.5	0.0056	0.0055	0.0054	0.0055	0.0055

The absolute biases of the estimated values are presented in Table 2 observe.

Bayes using the generalised non-informative prior outperformed Bayes using the gamma non-informative prior. This is followed by maximum likelihood estimator with respect to minimum bias. All the estimators had their mean squared error and absolute bias values decreasing with an analogous increase in sample space. It implies that the generalised non-informative prior is a good alternative to the gamma non-informative prior.

REFERENCES

Al-Aboud, F.M., 2009. Bayesian estimation for the extreme value distribution using progressive censored data and asymmetric loss. Int. Math. Forum, 4: 1603-1622.

Al-Bayyati, H.N., 2002. Comparing methods of estimating weibull failure model using simulation. Ph.D. Thesis, College of Administration and Economics, Baghdad, Iraq.

Al-Kutubi, H.S. and N.A. Ibrahim, 2009. Bayes estimator for exponential distribution with extension of Jeffery prior information. Malays. J. Math. Sci., 3(2): 297-313.

Elfessi, A. and D.M. Reineke, 2001. Bayesian look at classical estimation: The exponential distribution. J. Stat. Educ., 9(1).

Guure, C.B. and N.A. Ibrahim. 2012. Bayesian analysis of the survival function and the failure rate of weibull distribution with censored data. Math. Probl. Eng., Article ID: 329489, pp: 18, Doi:10.1155/2012/329489.

Guure, C.B., N.A. Ibrahim and A.M. Al Omari, 2012a. Bayesian estimation of two-parameter weibull distribution using extension of Jeffreys' prior information with three loss functions. Math. Probl. Eng., Article ID: 589640, pp: 13, Doi: 10.1155/2012/589640.

Horst, R., 2009. The Weibull Distribution Handbook. CRC Press, Taylor and Francis Group LLC, New York.

Varian, H.R., 1974. A Bayesian Approach to Real Estate Assessment. In: Fienberg, S.E. and A. Zellner (Eds.), Studies in Bayesian Econometrics and Statistics. North-Holland Press, Amsterdam, pp: 195-208.

Zellner, A., 1986. Bayesian estimation and prediction using asymmetric loss functions. J. Am. Stat. Assoc., 81: 446-451.