ON BAYESIAN ANALYSIS OF THE EXPONENTIAL SURVIVAL TIME ASSUMING THE EXPONENTIAL CENSOR TIME

M. Saleem and A. Raza

Department of Statistics, Government College University, Faisalabad 38000, Pakistan

ABSTRACT: Random censoring is a type of right censoring in time-to-event studies. The exponential survival time with an exponential censor time is focused to derive classical and the Bayes estimators. The uniform and the Inverted Gamma priors are assumed to carry out the Bayesian analysis. The posterior predictive distribution is derived and the equations required for the construction of predictive intervals are developed. The construction of the credible intervals and that of the Highest Posterior Density (HPD) intervals is elaborated theoretically as well as conducted numerically. A comprehensive simulation study assuming various parameter points and sample sizes is conducted to highlight the properties and comparison of the estimates.

Key words: Information matrix, predictive interval, credible intervals, highest posterior density intervals, inverse transform method

INTRODUCTION

Censoring is an unavoidable feature of the lifetime data. The lifetimes with constant hazard rates are modeled using exponential distribution which is known for its memory less property. Ahmed *et al*. (2005) and Ali *et al*. (2005) presented robust weighted likelihood and Bayes estimation of the Exponential parameters respectively. Raqab and Ahsanullah (2001) focused the ordered generalized exponential situation. Saleem and Aslam (2008a and b) use ordinary type I right censored data for Bayesian analysis of Rayleigh mixture. Saleem and Aslam (2009) considered Rayleigh survival time assuming random censor time. Saleem *et al*. (2010) used ordinary type I right censored data for Bayesian analysis of Power function mixture. This paper is an extension of Abu-Taleb *et al*. (2007) with some added features in terms of algebraic expressions along with numerical results and simulation study to explore and compare the properties of the estimators. The posterior predictive distribution is derived and the equations required for the construction of predictive intervals are presented. The construction of the credible intervals and that of the Highest Posterior Density (HPD) intervals is elaborated theoretically as well as conducted numerically. Computations are performed in Mathematica and Minitab.

MATERIALS AND METHODS

Maximum Likelihood Estimators: Let the survival time (x) and the censor time (T) independently follow exponential distributions with unknown parameters θ and λ respectively. The densities are

$$f(x|\theta) = \theta^{-1}e^{-\theta^{-1}x}\ ;\ x, \theta > 0 \quad \text{and}$$

$$f(t|\lambda) = \lambda^{-1}e^{-\lambda^{-1}t}\ ;\ t, \lambda > 0 \qquad (1)$$

A sample as described in Saleem and Aslam (2009) is considered and the discrete probability distribution of D_i and the joint density of (Y_i, D_i) are given in Abu-Taleb (2007) respectively along with $E(Y_i) = \frac{\theta\lambda}{\lambda+\theta}$ and $E(D) = \frac{\lambda}{\lambda+\theta}$. The natural log of the likelihood can be described as under.

$$l = \ln L(\theta,\lambda|_{\mathbf{z}}) \propto (-\sum d_i)\ln\theta - (n-\sum d_i)\ln\lambda - \frac{1}{\theta}\sum y_i - \frac{1}{\lambda}\sum y_i \qquad (2)$$

The ML estimators $\hat{\theta} = \frac{\sum y_i}{\sum d_i}$ and $\hat{\lambda} = \frac{\sum y_i}{n-\sum d_i}$ are the Bayes estimators under the Zero-One Loss and the uniform prior with variances $V(\hat{\theta}) = \frac{\theta^2(\theta+\lambda)}{n\lambda}$ and $V(\hat{\lambda}) = \frac{\lambda^2(\theta+\lambda)}{n\theta}$. Information matrix is derived as under.

$$I(\boldsymbol{\theta}) = \begin{bmatrix} \frac{n\beta}{\alpha^2(\alpha+\beta)} & 0 \\ 0 & \frac{n\alpha}{\beta^2(\alpha+\beta)} \end{bmatrix} \qquad (3)$$

where $\boldsymbol{\theta} = (\theta, \lambda)$ is the vector of unknown parameters. The expressions for the estimated variances are $v(\hat{\theta}) = \frac{(\sum y_i)^2}{(\sum d)^3}$ and $v(\hat{\lambda}) = \frac{(\sum y_i)^2}{(n - \sum d)^3}$.

Bayes Estimates Assuming Informative Prior: In case of a formal informative prior is available, the use of prior information is just like adding a fixed number of observations to the given sample size and consequently leads to the reduction of variance of the Bayes estimates based on the said informative prior. Bolstad (2004) has given an account of a method to evaluate the worth of a formal prior information in terms of the number of additional observations supposed to be added to the given sample size. It is interesting to note that the use of Uniform prior contributes zero number of observations to the given sample. Let θ and λ follow the Inverted gamma distribution with hyper-parameters $(a_1,\ b_1)$ and $(a_2,\ b_2)$ respectively. Assuming independence, we have a joint prior that is incorporated with the likelihood to yield the following joint posterior.

$$p(\theta, \lambda|\mathbf{z}) = \frac{(\sum y_i + b_1)^{\Sigma d_i + a_1 - 2} (\sum y_i + b_2)^{n - \Sigma d_i + a_2 - 2}}{\Gamma(\sum d_i + a_1 - 2)\ \Gamma(n - \sum d_i + a_2 - 2)}\ \theta^{-(\Sigma d_i + a_1 - 1)}$$

(4)

The marginal posterior distributions of θ and λ are

$$p_1(\theta|\mathbf{z}) = \frac{(\sum y_i + b_1)^{\Sigma d_i + a_1 - 2}}{\Gamma(\sum d_i + a_1 - 2)} \theta^{-(\sum d_i + a_1 - 1)} e^{-\frac{1}{\theta}(\Sigma y_i + b_1)}; \theta > 0 \tag{5}$$

and

$$p_2(\lambda|\mathbf{z}) = \frac{(\Sigma y_i + b_2)^{n - \Sigma d_i + a_2 - 2}}{\Gamma(n - \Sigma d_i + a_2 - 2)}\ \lambda^{-(n - \Sigma d_i + a_2 - 1)} e^{-\frac{1}{\lambda}(\Sigma y_i + b_2)}; \tag{6}$$

Using the squared error loss function, the Bayes estimators of θ and λ are found to be

$\hat{\theta} = \frac{\sum y_i + b_1}{\sum d_i + a_1 - 3}$ and $\hat{\lambda} = \frac{\sum y_i + b_2}{n - \sum d_i + a_2 - 3}$.

Each of these Bayes estimators is a linear combination of its ML as well as its Bayesian (Uniform prior) counterpart. The expressions for the Variances of the Bayes estimators with the Inverted Gamma prior are derived as

$$V(\hat{\theta}|\mathbf{z}) = \frac{(\sum y_i + b_1)^2}{(\sum d_i + a_1 - 3)^2 (\sum d_i + a_1 - 4)}$$ and

$$V(\hat{\lambda}|\mathbf{z}) = \frac{(\sum y_i + b_2)^2}{(n - \sum d_i + a_2 - 3)^2\ (n - \sum d_i + a_2 - 4)}.$$

Each of these variances is a linear combination of its ML as well as its Bayesian (Uniform prior) counterpart.

Bayes Estimates Assuming Uninformative Priors: The uniform prior is the most famous example of uninformative prior which materialize the use of Bayesian estimation methods when no formal prior information is available. $f_1(\theta) = k_1;\ 0 < \theta < \infty$ and $f_2(\lambda) = k_2;\ 0 < \lambda < \infty$. Assuming independence we have an improper joint prior that is proportional to a constant and is incorporated with the likelihood to yield a proper joint posterior distribution. Using the squared error loss function, the Bayes estimators of θ and λ assuming prior can be obtained on setting $b_1 = b_2 = 0$ in the Bayes estimators in case of informative prior.

The Posterior Predictive Distribution: The predictive distribution contains the information about the independent future random observation given the already accomplished observations. Bolstad (2004) and Bansal (2007) have given a detailed account of the predictive distribution. The posterior predictive distribution of the future observation, y is defined as

$$p(y|\mathbf{z}) = \int_0^\infty \int_0^\infty p(\ \theta, \lambda|\mathbf{z})\, f(y|\theta, \lambda)\ d\theta\ d\lambda$$

On simplification it reduces to

$$p(y|z) = [\frac{(\sum d + a_1 - 2)(\sum y + b_1)^{\Sigma d + a_1 - 2} (\sum y + b_2)^{n - \Sigma d + a_2 - 2}}{(\sum y + b_1 + y)^{\Sigma d + a_1 - 1} (\sum y + b_2 + y)^{n - \Sigma d + a_2 - 2}} + \frac{(n - \sum d + a_2 - 2)(\sum y + b_1)^{\Sigma d + a_1 - 2} (\sum y + b_2)^{n - \Sigma d + a_2 - 2}}{(\sum y + b_1 + y)^{\Sigma d + a_1 - 2} (\sum y + b_2 + y)^{n - \Sigma d + a_2 - 1}}];\ y > 0$$

(7)

with $a_1 = a_2 = a$ and $b_1 = b_2 = b$, it further reduces to

$$p(y|\mathbf{z}) = \frac{(n + 2a - 4)(\sum y + b)^{n + 2a - 4}}{(\sum y + b + y)^{n + 2a - 3}};\ y > 0 \tag{8}$$

The Bayesian Predictive Intervals: The $(1-\alpha)100\%$ Bayesian predictive interval (L, U) can be obtained by solving the two equations $\int_0^L p(y|\mathbf{z})\ dy = \frac{\alpha}{2} = \int_U^\infty p(y|\mathbf{z})\ dy$, where

$$[\frac{(n + 2a - 3)(n + 2a - 4)(\sum y + b)^{n + 2a - 4}}{(\sum y + b)^{n + 2a - 2}} - \frac{(n + 2a - 3)(n + 2a - 4)(\sum y + b)^{n + 2a - 4}}{(\sum y + b + L)^{n + 2a - 2}}] = 0$$

$p(y|\mathbf{z})$ given by (8). These equations, on simplification become

$$\frac{(n+2a-3)(n+2a-4)(\sum y+b)^{n+2a-4}}{(\sum \mathrm{y+b+U})^{n+2a-2}}=0$$

and

Evaluating these predictive intervals for various combinations of the hyper-parameters, a trend in hyper-parameters can be determined which leads to enhance the efficiency of the Bayes estimates.

The Bayesian Credible Intervals: If $p_1(\theta|\mathbf{z})$ is the posterior distribution given the sample and prior of the parameter of interest θ, we may be interested in finding an interval (θ_1,θ_2) such that

$$P(\theta\in(\theta_1,\theta_2)|\mathbf{z})=\int_{\theta_1}^{\theta_2} p_1(\theta|\mathbf{z})\,d\theta=1-\alpha$$

. The credible interval may be declared as a Bayesian counterpart of the classical confidence interval. However credible interval may not be unique even for a unimodal posterior density. The $(1-\alpha)100\%$ Bayesian credible intervals are obtained using the marginal distribution of the respective parameter of interest. The marginal posterior densities of θ and λ (assuming uniform prior) give the following credible intervals.

$$\frac{2\sum \mathrm{y}}{\chi^2_{\frac{\alpha}{2},2(\sum d-1)}}\le\theta\le\frac{2\sum \mathrm{y}}{\chi^2_{1-\frac{\alpha}{2},2(\sum d-1)}}$$

and

$$\frac{2\sum \mathrm{y}}{\chi^2_{\frac{\alpha}{2},2(n-\sum d-1)}}\le\lambda\le\frac{2\sum \mathrm{y}}{\chi^2_{1-\frac{\alpha}{2},2(n-\sum d-1)}}.$$

And the marginal posterior densities of θ and λ (assuming Inverted Gamma prior) result in the following credible intervals.

$$\frac{2(\sum \mathrm{y+b_1})}{\chi^2_{\frac{\alpha}{2},2(\sum d+a_1-1)}}\le\theta\le\frac{2(\sum \mathrm{y+b_1})}{\chi^2_{1-\frac{\alpha}{2},2(\sum d+a_1-1)}}$$

and

$$\frac{2(\sum \mathrm{y+b_2})}{\chi^2_{\frac{\alpha}{2},2(n-\sum d+a_2-1)}}\le\lambda\le\frac{2(\sum \mathrm{y+b_2})}{\chi^2_{1-\frac{\alpha}{2},2(n-\sum d+a_2-1)}}.$$

Where $\chi^2_{\frac{\alpha}{2},m}$, for example, has an area $\frac{\alpha}{2}$ on its right and the area beyond $\chi^2_{1-\frac{\alpha}{2},m}$ is $1-\frac{\alpha}{2}$.

The Highest Posterior Density (HPD) Intervals: The Highest Posterior Density (HPD) Interval may be declared as an advanced version of the Credible Interval. In case of a unimodal density, an additional condition can be imposed on a Credible Interval to make it unique by tilting it towards the most concentrated part of the posterior density. This unique Credible Interval is called an HPD Interval. The HPD interval is defined on the posterior density such that the posterior density at every point inside the HPD interval is greater than the posterior density at every point outside the HPD interval. An interval (θ_1,θ_2) would be a $(1-\alpha)100\%$ HPD interval for θ if it satisfy the following two conditions simultaneously as given in Sinha (1998).

$$\int_{\theta_1}^{\theta_2} p_1(\theta|_{\mathbf{z}})\,d\theta=1-\alpha$$ and $$p_1(\theta_1|\mathbf{z})=p_1(\theta_2|\mathbf{z})$$

where $p_1(\theta|\mathbf{z})$ is given by (5). On simplification, these two conditions reduce to

$$\Gamma(\sum d \quad {}_1a \quad \frac{\sum y_i+ {}_2b}{\theta_2}, \quad) \quad (d \quad {}_1a \quad \alpha- 2\Gamma, \sum \frac{y_i}{1} {}^b \quad \sum \quad) \quad {}_\theta($$

and

$$(\sum d+a_1-1)\ \ln(\frac{\theta_2}{\theta_1})-(\frac{1}{\theta_1}-\frac{1}{\theta_2})(\sum y+b_1)=0$$

Solving these two equations simultaneously gives the HPD interval (θ_1,θ_2) for θ. Here

$$\Gamma(\sum d+a_1-2,\ \frac{\sum y_i+b_1}{\theta_1})=\int_{\frac{\sum y_i+b_1}{\theta_1}}^{\infty}\theta^{\sum d+a_1-3}e^{-\theta}d\theta$$

is the incomplete Gamma function. While

$$\Gamma(\sum d+a_1-2)=\int_0^{\infty}\theta^{\sum d+a_1-3}e^{-\theta}\,d\theta$$

is the Gamma function. Similarly a $(1-\alpha)100\%$ HPD interval (λ_1,λ_2) for λ is obtained by solving the following two equations.

$$\int_{\lambda_1}^{\lambda_2} p_2(\lambda|_{\mathbf{z}})\,d\lambda=1-\alpha$$ and $$p_2(\lambda_1|_{\mathbf{z}})=p_2(\lambda_2|_{\mathbf{z}})$$

where $p_2(\lambda|_{\mathbf{z}})$ is given by (8).

$$(n-\sum d+a_2-1)\ \ln(\frac{\lambda_2}{\lambda_1})-(\frac{1}{\lambda_1}-\frac{1}{\lambda_2})(\sum y+b_2)=0$$

and

$$\Gamma(n \sum -d_2 \quad \frac{(\sum y+ {}_2b \quad)}{a \quad \theta_2} 2, + \quad - \quad) \quad (n \quad \alpha \quad d_2 \quad a \quad \frac{(y \quad {}_2b \quad)}{1} 2, \quad -$$

RESULTS AND DISCUSSION

We simulate random samples of sizes $n = 50, 100, 250$ of exponential survival and termination times with parameters $(\theta,\ \lambda) = (75,50)$ and $(175,200)$. The results of the simulation study are presented in Tables 1-2.

Table1 The ML estimates and the Bayes estimates (with standard errors in the parenthesis) assuming the Uniform and the SRIG priors of parameters $\theta = 75,\ \lambda = 50$ Hyper-parameters assumed are $a_1 = a_2 = 4.5$, $b_1 = b_2 = 0$.

Sample Size	ML estimates		Methods Bayes estimates (Uniform)		Bayes estimates (SRIG)	
n	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$
50	77.792 (17.9267)	50.838 (9.3925)	87.1963 (22.067)	54.5816 (10.6576)	72.01 (16.3495)	48.355 (8.8546)
100	75.6471 (12.0794)	50.4732 (6.5632)	79.7318 (13.2612)	52.2431 (6.9742)	72.8519 (11.5556)	49.223 (6.3731)
250	75.9146 (7.6485)	50.1305 (4.0986)	77.4906 (7.9302)	50.8103 (4.1967)	74.7742 (7.5141)	49.6325 (4.0511)

Table2 The ML estimates and the Bayes estimates (with standard errors in the parenthesis) assuming the Uniform and the SRIG priors of parameters $\theta = 175,\ \lambda = 200$. Hyper-parameters assumed are $a_1 = a_2 = 4.5$, $b_1 = b_2 = 0$.

Sample Size	**ML estimates**		**Methods Bayes estimates (Uniform)**		**Bayes estimates (SRIG)**	
n	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$
50	179.357 (35.3257)	206.047 (43.5989)	194.559 (40.8574)	226.53 (51.7014)	169.457 (33.0329)	193.019 (40.349)
100	176.087 (24.319)	201.777 (29.8625)	183.092 (26.0545)	211.058 (32.3381)	171.178 (23.525)	195.341 (28.7456)
250	175.937 (15.2653)	202.125 (18.8039)	178.629 (15.6778)	205.689 (19.3905)	173.971 (15.066)	199.532 (18.522)

Table3. The 95% Bayesian Credible Intervals and the HPD Intervals assuming the Uniform and the Inverted Gamma prior. Hyper-parameters assumed are $a_1 = 20, a_2 = 25$, $b_1 = 3, b_2 = 1$.

Parameters	**Credible Intervals (Uniform prior)**	**HPD Intervals (Uniform prior)**	**Credible Intervals (Inverted Gamma)**	**HPD Intervals (Inverted Gamma)**
$\theta = 150$	(133.391,171.948)	(132.63,171.03)	(123.678,157.846)	(123.0,157.1)
$\lambda = 100$	(93.2583,115.044)	(92.9,114.63)	(87.3198,106.957)	(87.0,106.59)

It is immediate from the Table 1 that increasing the sample size reduces the standard errors of the estimates. ML estimates and Bayes (Uniform) estimates are overestimated but the extent of this overestimation is higher in the latter. The Bayes (SRIG) estimates are under estimated but are more précised than the rest of estimates. The extent of over or under-estimation is reduced with the increase in sample size. Also the censor time parameter has low standard errors than that of the life time parameter. It is evident from Table 2 that increasing the sample size reduces the standard errors of the estimates. ML estimates and Bayes (Uniform) estimates are overestimated but the extent of this overestimation is higher in the latter. The Bayes (SRIG) estimates are under estimated but are more précised than the rest of estimates. The extent of over or under-estimation is reduced with the increase in sample size.

Also the life time estimates have lesser standard errors than that of censor time estimates.

Table 3 shows that both the credible intervals and the HPD intervals assuming the Informative prior (The Inverted Gamma) are pretty shorter than those obtained by assuming the state of ignorance (uniform prior).

This again highlights the advantage of incorporating the informative conjugate prior information into the analysis rather than an uninformative one that has either no proper density or involves no hyper-parameters. The HPD intervals are not only narrower than their corresponding credible intervals but are also slightly left aligned to capture the most denser parts of the concerned marginal posterior distributions. This is supported by the fact that the marginal distributions of the unknown parameters are positively skewed. A more extensive analysis can be managed by considering a number of different parameter points and a variety of sample sizes.

The simulation study shows that increasing the sample size reduces the standard errors of the estimates. ML estimates and Bayes (Uniform) estimates are overestimated but the extent of this overestimation is higher in the latter. The Bayes (SRIG) estimates are under estimated but are more précised than the rest of estimates. This shows that the use of an informative prior is more paying than an uninformative prior. The extent of over or under-estimation is reduced with the increase in sample size. Also the life time estimates have lesser (greater) standard errors than that of censor time estimates if the lifetime parameters are smaller (larger) than the censor time parameters. The Credible intervals assuming the Inverted Gamma prior are much narrower than the credible intervals assuming the Uniform prior. The HPD intervals assuming the Inverted Gamma prior are more precise than the HPD intervals assuming the state of ignorance. It is the use of prior information that makes a difference everywhere. As the marginal posterior densities are positively skewed, so the HPD intervals are slightly left aligned as compared to the corresponding credible intervals. Also, the lengths of the HPD intervals are shorter than the lengths of the corresponding credible intervals. The predictive intervals can be used to discover a trend and a range of the hyper-parameters that ensure the more precise estimates. This can also be used to further refine the prior pieces of information already provided by a number of experts.

REFERENCES

Abu-Taleb A. A, M. M. Smadi, and A. J. Alawneh. Bayes estimation of the lifetime parameters for the exponential distribution. J. Mathe. and Stat., 3 (3): 106-108. (2007)

Ahmed, E. S., A. I. Volodin, and A. Hussein. Robust weighted likelihood estimation of exponential parameters, IEEE Transactions on Reliability, 54(3): 389-395. (2005)

Ali, M. M, J. Woo and S. Nadarajah. Bayes estimators of the exponential distribution, J. Statist. Mgmt. Sys., 8(1): 53-58. (2005)

Bansal, A. K. Bayesian Parametric Inference. Narosa Publishing House Pvt. Ltd. New Delhi. (2007)

Bolstad, W. M.. Introduction to Bayesian Statistics. John Wiley & Sons, Inc. New Jersey. (2004)

Raqab, M. M, M. Ahsanullah. Estimation of location and scale parameters of generalized exponential distribution based on order statistic. Journal of Statistical Computation and Simulation, 22: 112-127. (2001)

Saleem, M. and M. Aslam,. On prior selection for the mixture of Rayleigh distribution using predictive Intervals, Pakistan J. Stat. 24(1): 21-35. (2008a)

Saleem, M. and M. Aslam, Bayesian Analysis of the two component mixture of the Rayleigh Distribution with uniform and Jeffreys priors, J. Applied Stat. Sci., 16(4): 105-113. (2008b)

Saleem, M. and M. Aslam,. On bayesian analysis of the rayleigh survival time Assuming the random censor time, Pakistan. J. Stat. 25(2): 71-82. (2009)

Saleem, M., M. Aslam, and P. Economou.. On the Bayesian analysis of the mixture of Power function distribution Using the complete and the censored sample, J. Applied Stat., 37(1): 25-40. (2010)

Sinha, S. K.. Bayesian Estimation. New Age International (P) Limited, Publishers, New Delhi. (1998)