



## Approximate Bayes estimators applied to the Bilal model



Ayman M. Abd-Elrahman<sup>a,\*</sup>, Sherif F. Niazi<sup>b</sup>

<sup>a</sup> Department of Mathematics, University of Assiut, 71516 Assiut, Egypt

<sup>b</sup> Mathematics Department, Al-Azhar University, Assuit Branch, 71524 Assiut, Egypt

### ARTICLE INFO

#### Article history:

Received 4 June 2015

Revised 22 May 2016

Accepted 24 May 2016

Available online 13 October 2016

#### MSC:

60E05

62B15

62F10

62N02

62N05

#### Keywords:

Jiffery's prior

Tierney and Kadanes' approximation form

Gibb's sampling approach

Maximum likelihood and EM algorithm

Estimated risks

### ABSTRACT

This paper develops approximate Bayes estimators of the parameter of the Bilal failure model by using the method of Tierney and Kadane [Accurate approximations for posterior moments and marginal densities, *J. Amer. Statist. Assoc.* 81 (1986) 82–86.] based on Type-2 censored sample and four different loss functions. Existence and uniqueness theorem for the maximum likelihood estimate are established. Based on Monte Carlo simulation study, comparisons are made between those estimators and their corresponding Bayes estimators obtained by using Gibb's sampling approach.

© 2016 Egyptian Mathematical Society. Production and hosting by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license.

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

### 1. Introduction

The Bilal distribution,  $Bilal(\theta)$ , is introduced by Abd-Elrahman [2] as a member of some families of distributions. He shows that, this distribution is a member of the class of new better than average renewal failure rates, and its density function is always unimodal and has less of skewness and kurtosis than the density of the exponential distribution by about 25% and 28%, respectively. Furthermore, however, the distribution function,  $q$ th quantile, failure rate function are in compact forms and the different moments are obtained in explicit forms in terms of the exponential function. The cumulative distribution (cdf), and the probability density (pdf) functions of  $Bilal(\theta)$  distribution are respectively as follows

$$F_X(x; \theta) = 1 - e^{-\frac{2x}{\theta}} (3 - 2e^{-\frac{x}{\theta}}), \quad x \geq 0, (\theta > 0), \quad (1)$$

$$f_X(x; \theta) = \frac{6}{\theta} e^{-\frac{2x}{\theta}} (1 - e^{-\frac{x}{\theta}}). \quad (2)$$

The  $q$ th quantile,  $x_q$ , is an important quantity, especially for generating random varieties using the inverse transformation method.

This quantity is obtained from (1) as [2]

$$x_q = -\theta \ln [\mathcal{U}(a_q)], \quad (3)$$

where  $a_q = \frac{1}{3} \arctan\left(\frac{2\sqrt{q(1-q)}}{2q-1}\right)$ , and

$$\mathcal{U}(a_q) = \begin{cases} 0.5 + \sin(a_q + \frac{\pi}{6}) & \text{if } 0 < q < 0.5, \\ 0.5 & \text{if } q = 0.5, \\ 0.5 - \cos(a_q + \frac{\pi}{3}) & \text{if } 0.5 < q < 1. \end{cases}$$

Although the Bilal model has only one parameter, this distribution is much better than some other recent distributions for fitting two different real data sets, namely, (i) Hinkley's [3] data; and (ii) the data for waiting times (in minutes) before service of 100 bank customers that is examined and analyzed, respectively, by Ghitany et al. [4] and Zakerzadeh and Mahmoudi [5] in fitting the Lindely and the compounding Lindely-Geometric distributions, see [2] for reference.

In this paper, based on Type-2 censored sample, both maximum likelihood (MLE) and Bayesian estimates are considered. For MLE,  $\hat{\theta}_{ML}$ , we established existence and uniqueness theorem. The Fisher information about the unknown parameter as well as the corresponding asymptotic confidence interval (ACI) are obtained with the use of asymptotic normality of the MLE and the missing information principle. An EM algorithm for estimating  $\theta$ , say  $\hat{\theta}_{EM}$ , is also presented. For Bayesian approach, we set Bayesian es-

\* Corresponding author.

E-mail address: [ayman275@aun.edu.eg](mailto:ayman275@aun.edu.eg) (A.M. Abd-Elrahman).

timates based on squared error, LINEX, Entropy and Precautionary loss functions. By using the method of Tierney and Kadane [1], approximate Bayesian estimators for the parameter  $\theta$  are developed. On the other hand, by using Gibbs's sampling approach, Bayesian estimators and the highest posterior density (HPD) credible intervals for the parameter  $\theta$  are derived. Simulation studies are conducted for comparing the resulting estimators, for various sample and censoring sizes. Finally, we draw some concluding remarks.

## 2. Maximum likelihood estimation

Suppose  $n$  items, which follow  $Bilal(\theta)$  distribution, are put on a life-testing experiment and we observe only the first  $r$  failure times, say  $x_{(1)} < x_{(2)} < \dots < x_{(r)}$ . It follows from (1) and (2) that, the likelihood function of  $\theta$  can be written as

$$\ell(\text{Data}|\theta) \propto \theta^{-r} e^{-\frac{2}{\theta}} \left( (n-r)x_{(r)} + \sum_{i=1}^r x_{(i)} \right) \times \left( 3 - 2e^{-\frac{x_{(r)}}{\theta}} \right)^{n-r} \prod_{i=1}^r \left( 1 - e^{-\frac{x_{(i)}}{\theta}} \right). \quad (4)$$

The corresponding likelihood equation is then given by

$$\frac{\partial \ln(\ell)}{\partial \eta} = \frac{r}{\eta} - 2(n-r)x_{(r)} \left\{ 1 - \frac{e^{-\eta x_{(r)}}}{3 - 2e^{-\eta x_{(r)}}} \right\} - \sum_{i=1}^r x_{(i)} \left\{ 3 - (1 - e^{-\eta x_{(i)}})^{-1} \right\} = 0, \quad \eta = \frac{1}{\theta}. \quad (5)$$

**Theorem 2.1.** The MLE,  $\hat{\theta}_{ML}$ , for  $\theta$  based on a Type-2 censored sample exists and it is unique.

**Proof.** See Appendix.

The solution of (5), say  $\eta^*$ , can be numerically obtained, and by using the invariance property of the MLE,  $\hat{\theta}_{ML}$  is then equal to  $\frac{1}{\eta^*}$ .  $\square$

### 2.1. Fisher information about the parameter $\theta$

The well known missing information principle will be used for computing the Fisher information about the unknown parameter  $\theta$  under Type-2 censoring data from  $Bilal(\theta)$  distribution. This technique have been used by, e. g., Ng et al. [6] and Abd-Elrahman [7]. In order to do this, first of all:

1. Let  $\mathbf{X} = (x_{(1)}, x_{(2)}, \dots, x_{(r)})'$  denote the ordered observed censored data.
2. Let  $\mathbf{Y} = (X_{(r+1)}, X_{(r+2)}, \dots, X_{(n)})'$  denote the unobserved ordered data.

The vector  $\mathbf{Y}$  can be thought of as the missing data. Combine  $\mathbf{X}$  and  $\mathbf{Y}$  to form  $\mathbf{W}$ , which is the complete data set.

As an especial case of the Theorem of Ng et al. [6], one can see that, for  $s = r+1, r+2, \dots, n$ , the conditional distribution of each  $X_{(s)} \in \mathbf{Y}$  given  $X_{(s)} > x_{(r)}$  follows the truncated underlying distribution with left truncation at  $x_{(r)}$ . Therefore, using (1) and (2), we have

$$f_{X_{(s)}|X_{(r)}}(x|X_{(s)} > x_{(r)}; \theta) = \frac{6e^{-\frac{2(x-x_{(r)})}{\theta}} (1 - e^{-\frac{x}{\theta}})}{\theta (3 - 2e^{-\frac{x_{(r)}}{\theta}})}, \quad x > x_{(r)}, (\theta > 0). \quad (6)$$

The Fisher information of the ordered observed censored and complete data are denoted by  $I_{\mathbf{X}}(\theta)$  and  $I_{\mathbf{W}}(\theta)$ , respectively. Denote

$I_{\mathbf{Y}|\mathbf{X}}(\theta)$  for the ordered unobserved (missing) information related to the vector  $\mathbf{Y}$ . Hence, based on the conditional distribution in (6), the expected Fisher information related to the vector  $\mathbf{Y}$  is given by

$$I_{\mathbf{Y}|\mathbf{X}}(\theta) = -(n-r) \mathbb{E} \left[ \frac{\partial^2 \ln[f_{X_{(s)}|\mathbf{X}}(x|X_{(s)} > x_{(r)}; \theta)]}{\partial \theta^2} \right] = (n-r) T_1(x_{(r)}, \theta), \quad (7)$$

where

$$T_1(x_{(r)}, \theta) = \frac{1}{\theta^2} \left\{ 1 - \left( \frac{6e^{-\frac{x_{(r)}}{\theta}}}{3 - 2e^{-\frac{x_{(r)}}{\theta}}} \right) \left[ \frac{\left( \frac{x_{(r)}}{\theta} \right)^2}{3 - 2e^{-\frac{x_{(r)}}{\theta}}} - \sum_{j=0}^{\infty} \frac{\left( 1 + \left( 1 + \frac{(3+j)x_{(r)}}{\theta} \right)^2 \right) e^{-\frac{jx_{(r)}}{\theta}}}{(3+j)^3} \right] \right\}, \quad (8)$$

which is the Fisher information related to each  $x_{(s)}$ ,  $s = r+1, r+2, \dots, n$ , where  $x_{(s)}$  is distributed as in (6). Therefore, it follows from (8) that, the Fisher information about the parameter  $\theta$  related to the complete data set  $\mathbf{W}$  is given by

$$I_{\mathbf{W}}(\theta) = n \lim_{t \rightarrow 0^+} T_1(t, \theta) = \frac{n(1 + 12 \sum_{j=0}^{\infty} (3+j)^{-3})}{\theta^2} = \frac{(24\zeta(3) - 25)n}{2\theta^2} = \frac{1.92468284n}{\theta^2}, \quad (9)$$

$$\zeta(3) = \sum_{i=1}^{\infty} i^{-3} = 1.202056903.$$

Note that, Eq. (9) goes in line with the result due to Abd-Elrahman [2]. Therefore, the Fisher information gains from a given Type-2 censored sample,  $x_{(1)}, x_{(2)}, \dots, x_{(r)}$ , from  $Bilal(\theta)$  distribution is then given by

$$I_{\mathbf{X}}(\theta) = \frac{1.92468284n}{\theta^2} - (n-r) T_1(x_{(r)}, \theta), \quad (10)$$

where  $T_1(x_{(r)}, \theta)$  is given by (8). Once  $I_{\mathbf{X}}(\theta)$  is calculated at  $\theta = \hat{\theta}_{ML}$ , the asymptotic variance of the MLE of the parameter  $\theta$  is then given by

$$\widehat{\text{Var}}(\hat{\theta}_{ML}) = \left\{ I_{\mathbf{X}}(\hat{\theta}_{ML}) \right\}^{-1}.$$

Consequently, the asymptotic 100(1- $\alpha$ )% confidence interval, ACI, of  $\hat{\theta}_{ML}$  is given by

$$\left[ \hat{\theta}_{ML} - Z_{\frac{\alpha}{2}} \text{SD}(\hat{\theta}_{ML}), \hat{\theta}_{ML} + Z_{\frac{\alpha}{2}} \text{SD}(\hat{\theta}_{ML}) \right], \quad (11)$$

where  $\text{SD}(\hat{\theta}_{ML}) = \sqrt{\widehat{\text{Var}}(\hat{\theta}_{ML})}$  and  $Z_{\frac{\alpha}{2}}$  is the percentile  $(1 - \frac{\alpha}{2})$  of the standard normal distribution.

### 2.2. EM algorithm

The EM algorithm was proposed by Dempster et al. [8] as a very powerful tool in handling the incomplete data problem. It is an iterative method by repeating to fill in the missing data with expected values and to update the parameter estimates. An EM algorithm for obtaining the MLE,  $\hat{\theta}_{EM}$ , when the data are Type-2 censored from  $Bilal(\theta)$  distribution, can be easily constricted as follows:

In the E-Step, the expectation of  $X_{(s)} [3 - (1 - e^{-\frac{x_{(s)}}{\theta}})^{-1}]$  given  $X_{(s)} > x_{(r)}$ ,  $s = r+1, r+2, \dots, n$ , is needed. By using (6), it is easy

to show that

$$\begin{aligned} E \left[ X_{(s)} \left( 3 - \left( 1 - e^{-\frac{x_{(s)}}{\theta}} \right)^{-1} \right) \middle| X_{(s)} > x_{(r)} \right] \\ = \theta + 3x_{(r)} \left( 1 - \frac{1}{3 - 2e^{-\frac{x_{(r)}}{\theta}}} \right). \end{aligned}$$

In the M-Step, on the  $(h + 1)$ th iteration of the EM algorithm, the value of  $\hat{\theta}_{(h+1)}$  may then be calculated as follows

$$\begin{aligned} \hat{\theta}^{(h+1)} = \frac{1}{n} \left\{ (n-r) \left[ \hat{\theta}_{(h)} + 3x_{(r)} \left( 1 - \frac{1}{3 - 2e^{-x_{(r)}/\hat{\theta}_{(h)}}} \right) \right] \right. \\ \left. + \sum_{j=1}^r x_{(j)} \left( 3 - \frac{1}{1 - e^{-x_{(j)}/\hat{\theta}_{(h)}}} \right) \right\}. \end{aligned} \tag{12}$$

Dempster et al. [8] have shown that, the EM algorithm ensures the convergence of this procedure to a local maximum, irrespective of the starting point. However, a reasonable starting value for  $\hat{\theta}_{(0)}$  is the estimate of the parameter  $\theta$  based on the observed vector of observations  $\mathbf{X} = (x_{(1)}, x_{(2)}, \dots, x_{(r)})'$  as a “pseudo-complete” sample of size  $r$ . For a given data  $\mathbf{X}$ ,  $\hat{\theta}_{EM}$  can be iteratively calculated using (12). These iterations will be repeated 1000 times. As a stopping rule, the iterations will be terminated after some value of  $h \leq 1000$  with a level of accuracy less than  $1.2 \times 10^{-7}$  of the absolute relative error, which is defined as

$$\epsilon = \left| \frac{\hat{\theta}^{(h+1)} - \hat{\theta}^{(h)}}{\hat{\theta}^{(h)}} \right|.$$

### 3. Bayes estimations

It is assumed that the prior for  $\eta$ ,  $\eta = \frac{1}{\theta}$ , has a gamma prior distribution with hyper parameters  $c$  and  $d$ ; and it has the pdf

$$\pi_1(\eta) = \frac{d^c}{\Gamma(c)} \eta^{c-1} e^{-d\eta}, \quad \eta > 0, \quad (c, d > 0). \tag{13}$$

This prior is frequently used when the range of a population parameter under estimation is from zero to infinity. The hyper parameters can be chosen to suit the prior belief of the experimenter in terms of location and variability of the prior distribution. It may be easy to show that, the non-informative prior, Jiffery’s prior, of  $\eta$  is an especial case of (13), which can be obtained by substituting  $c = d = 0$  in (13).

Combining (4) and (13), the posterior density function of  $\eta$ ,

$$\pi(\eta|\text{Data}) = \frac{\ell(\text{Data}|\eta)\pi_1(\eta)}{\int_0^\infty \ell(\text{Data}|\eta)\pi_1(\eta) d\eta}, \tag{14}$$

takes the form

$$\pi(\eta|\text{Data}) = C_0 \eta^{c+r-1} \exp[-a\eta + T(\eta)], \tag{15}$$

where

$$C_0^{-1} = \int_0^\infty \eta^{c+r-1} \exp[-a\eta + T(\eta)] d\eta,$$

$$a = d + 2[(n-r)x_r + \sum_{j=1}^r x_j] \text{ and } T(\eta) = (n-r) \ln(3 - 2e^{-x_{(r)}\eta}) + \sum_{i=1}^r \ln(1 - e^{-x_{(i)}\eta}).$$

In the Bayesian frame work, a loss function is needed. However, it is well known that, the squared error loss function (SLF) is well justified when the losses are symmetric in nature. But losses may not be symmetric. Therefore, in this article we consider, in addition to SLF, three different asymmetric loss functions which are found in literatures, namely, LINEX (LLF), Entropy (ELF), and Precautionary (PLF) loss functions, (see [9–11], ).

Under any of these loss functions, the Bayes estimators of a function  $g(\eta)$  requires the evaluation of a ratio of two integrals of the form

$$\mathcal{E}_G = \frac{\int_0^\infty G(\eta) Q(\eta) d\eta}{\int_0^\infty Q(\eta) d\eta}, \tag{16}$$

where  $Q(\eta)$  is the posterior density of  $\eta$  except for the normalizing constant and  $G(\eta)$  is a continuous function of  $\eta$ , which is related to the function  $g(\eta)$ . The integrals involved in (16) are usually not obtainable in closed form. However, an approximate procedure can be used for this propose. We will adopt the approximation form developed by Tierney and Kadane [1].

#### 3.1. Tierney and Kadanes’ approximation

Tierney and Kadane [1] gave an approximated expression, of order  $n^{-2}$ , where  $n$  is the sample size, for the evaluation of the ratio of integrals of the form (16) by writing the two expression

$$L(\eta) = \frac{1}{n} \ln(Q(\eta)), \quad L^*(\eta) = L(\eta) + \frac{1}{n} \ln(G(\eta)).$$

So that (16) takes the form

$$\mathcal{E}_G = \frac{\int_0^\infty \exp[n(L^*(\eta))] d\eta}{\int_0^\infty \exp[n(L(\eta))] d\eta}. \tag{17}$$

Following Tierney and Kadane [1], Eq. (17) can be written in the approximate form

$$\hat{\mathcal{E}}_G = \frac{\sigma^*}{\sigma} \exp[n(L^*(\hat{\eta}^*) - L(\hat{\eta}))], \tag{18}$$

where  $\hat{\eta}$  and  $\hat{\eta}^*$  are the modes of  $L(\eta)$  and  $L^*(\eta)$ , respectively,

$$\sigma = -1/L''(\hat{\eta}), \quad \sigma^* = -1/L''(\hat{\eta}^*) \quad \text{and} \quad ' \equiv \frac{\partial}{\partial \eta}.$$

This form requires that each of the two integrands in (17) is unimodal function of  $\eta$ .

In our case, it follows from (15) that, the function  $L(\eta)$  can be written as

$$L(\eta) = \frac{1}{n} [(c+r-1) \ln(\eta) - a\eta + T(\eta)], \tag{19}$$

which satisfies

$$\begin{aligned} L''(\eta) &= \frac{\partial^2 L(\eta)}{\partial \eta^2} \\ &= -\frac{1}{n} \left[ \frac{r+c-1}{\eta^2} + \frac{6(n-r)x_r^2 e^{-x_r\eta}}{(3-2e^{-x_r\eta})^2} + \sum_{j=1}^r \frac{x_j^2 e^{-x_j\eta}}{(1-e^{-x_j\eta})^2} \right] < 0. \end{aligned} \tag{20}$$

Using a similar technique which is used for proving Theorem 2.1, it is easy to show that, the following equation

$$L'(\eta) = \frac{1}{n} \left[ \frac{c+r-1}{\eta} - a + T'(\eta) \right] = 0, \tag{21}$$

where

$$T'(\eta) = \frac{2(n-r)x_r e^{-x_r\eta}}{3-2e^{-x_r\eta}} + \sum_{j=1}^r \frac{x_j e^{-x_j\eta}}{1-e^{-x_j\eta}},$$

has a unique root at  $\hat{\eta}$ , say, which can be obtained numerically.

Now, for the Bayes estimators of  $\theta = g(\eta) = \frac{1}{\eta}$  with respect to SLF, LLF, ELF and PLF, respectively, which are depicted in Table 1, we consider

$$\begin{aligned} L_{SLF}^*(\eta) &= L(\eta) - \ln(\eta)/n, & L_{LLF}^*(\eta) &= L(\eta) - a_1/(n\eta), \quad a_1 > 0, \\ L_{ELF}^*(\eta) &= L(\eta) + \ln(\eta)/n, & L_{PLF}^*(\eta) &= L(\eta) - 2 \ln(\eta)/n. \end{aligned}$$

It may be easy to show that, each of these functions is a unimodal function of  $\eta$ . Their modes, respectively, are at  $\hat{\eta}_{SLF}^*$ ,  $\hat{\eta}_{LLF}^*$ ,  $\hat{\eta}_{ELF}^*$  and

**Table 1**  
SLF and some asymmetric loss functions that are found in literature.

Loss function	Functional form	Bayesian estimator
SLF	$L_1(\Delta) \propto \Delta^2, \Delta = \hat{\eta} - \eta$	$IE_{\eta \text{Data}}[g(\eta)]$
LLF	$L_2(\Delta) \propto (e^{a_1 \Delta} - a_1 \Delta - 1), (a_1 \neq 0)$	$-\frac{1}{a_1} \ln \left( E_{\eta \text{Data}} \left[ e^{-a_1 g(\eta)} \right] \right)$
ELF	$L_3(\delta) \propto [\delta - \ln(\delta) - 1], \delta = \frac{\hat{\eta}}{\eta}$	$\left\{ E_{\eta \text{Data}} \left[ \frac{1}{g(\eta)} \right] \right\}^{-1}$
PLF	$L_4(\hat{\eta}, \eta) \propto \frac{(\hat{\eta} - \eta)^2}{\hat{\eta}}$	$\left\{ E_{\eta \text{Data}} [g^2(\eta)] \right\}^{\frac{1}{2}}$

$\hat{\eta}_{PLF}^*$ , which can be numerically obtained by solving the following equations, separately,

$$0 = L_{SLF}'(\eta) = L'(\eta) - 1/(n\eta), \quad 0 = L_{LLF}'(\eta) = L'(\eta) + a_1/(n\eta^2), \quad a_1 > 0,$$

$$0 = L_{ELF}'(\eta) = L'(\eta) + 1/(n\eta), \quad 0 = L_{PLF}'(\eta) = L'(\eta) - 2/(n\eta^2),$$

where  $L'(\eta)$  is as given by (21). Therefore,

$$L_{SLF}''(\eta) = L''(\eta) + 1/(n\eta^2), \quad L_{LLF}''(\eta) = L''(\eta) - 2a_1/(n\eta^3), \quad a_1 > 0,$$

$$L_{ELF}''(\eta) = L''(\eta) - 1/(n\eta^2), \quad L_{PLF}''(\eta) = L''(\eta) + 2/(n\eta^2),$$

where  $L''(\eta)$  is as given by (20). Note that, once  $\hat{\eta}$  and  $\hat{\eta}_{\xi}^*$  are calculated,  $\sigma = -1/L''(\hat{\eta})$ ,  $\sigma_{\xi}^* = -1/L''_{\xi}(\hat{\eta}_{\xi}^*)$ ,  $L(\hat{\eta})$  and  $L_{\xi}^*(\hat{\eta}_{\xi}^*)$  are readily obtained, where  $\xi$  stands for SLF, LLF, ELF or PLF.

**Remark.** During our simulation, given below, we observe that when  $a_1 < 0$ , the unimodality of  $L_{LLF}^*(\eta)$  is not valid for some cases. This may be expected since, for  $a_1 < 0$ , the function  $\exp(-a_1/\eta)$  is not unimodal or even pounded in the rang  $(0, \infty)$  of  $\eta$ , i.e.,  $\exp(-a_1/\eta)$  does not satisfy one of the regularity conditions of Tierney and Kadanes' approximation form.

In view of Table 1 and Eq. (18), Tierney and Kadanes' approximate Bayes estimators for the parameter  $\theta$ , with respect to SLF, LLF, ELL and PLF, respectively, can be written as

$$\begin{aligned} \hat{\theta}_s &= \frac{\sigma_{SLF}^*}{\sigma} \exp \{n [L_{SLF}^*(\hat{\eta}_{SLF}^*) - L(\hat{\eta})]\}, \\ \hat{\theta}_l &= -\frac{1}{a_1} \left\{ \ln \left( \frac{\sigma_{LLF}^*}{\sigma} \right) + n [L_{LLF}^*(\hat{\eta}_{LLF}^*) - L(\hat{\eta})] \right\}, \\ \hat{\theta}_e &= \frac{\sigma}{\sigma_{ELF}^*} \exp \{n [L(\hat{\eta}) - L_{ELF}^*(\hat{\eta}_{ELF}^*)]\}, \\ \hat{\theta}_p &= \sqrt{\frac{\sigma_{PLF}^*}{\sigma}} \exp \left\{ \frac{n}{2} [L_{PLF}^*(\hat{\eta}_{PLF}^*) - L(\hat{\eta})] \right\}. \end{aligned} \quad (22)$$

### 3.2. Approximate Bayesian estimation using MCMC technique

Following Pradhan and Kundu [12], it can be easy to construct an approximate Bayesian estimator for  $\theta$  using MCMC technique, with respect to SLF, LLF, ELF or PLF, and their corresponding credible intervals. Where, by using the algorithm of Devroye [13], it is possible to generate a Gibb's sample from the posterior density function of  $\eta$ . This algorithm requires to ensure that (15) has a log-concave density function property. The function (15) is already the case, compare (20). We use this algorithm for generating a sample  $\eta_1, \dots, \eta_M, M = 1000$  say, from (15). In view of Table 1, the approximate Bayes estimators for  $g(\eta)$ , with respect to SLF, LLF, ELF and PLF, respectively, can be written as

$$\begin{aligned} \widetilde{g(\eta)}_s &= \frac{1}{M} \sum_{i=1}^M g(\eta_i), & \widetilde{g(\eta)}_l &= -\frac{1}{a_1} \ln \left[ \frac{\sum_{i=1}^M e^{-a_1 g(\eta_i)}}{M} \right], \\ \widetilde{g(\eta)}_e &= M \left[ \sum_{i=1}^M \frac{1}{g(\eta_i)} \right]^{-1}, & \widetilde{g(\eta)}_p &= \left[ \frac{\sum_{i=1}^M (g(\eta_i))^2}{M} \right]^{1/2}. \end{aligned} \quad (23)$$

Denoting  $\lfloor x \rfloor$  the largest integer less than or equal to  $x$ , the  $100(1 - \alpha)\%$  HPD symmetric credible interval for  $g(\eta)$  becomes  $[g_{(\lfloor M\alpha/2 \rfloor + 1)}, g_{(\lfloor M(1 - \alpha)/2 \rfloor)}]$ . When  $M = 1000$ , 95% HPD symmetric credible interval for  $g(\eta)$  becomes  $[g_{(26)}, g_{(975)}]$ .

### 4. Numerical computations and comparisons

In order to compare the performance of the estimators obtained in the above sections, we design some simulation experiments. Note that, throughout our simulation, we find that  $\hat{\theta}_{ML}$  is, almost, as the same as  $\hat{\theta}_{EM}$ . This may be expected since the maximum likelihood estimate of  $\theta$  is unique. Therefore, to avoid repetitions, we will denote this unique MLE as  $\hat{\theta}_M$ .

The following steps were applied for the simulation:

- Step (1) Values of  $n, r$  and the hyper parameters  $c$  and  $d$  were selected as  $c=5$  and  $d=8$ . This gives  $E(\theta) = \frac{d}{c-1} = 2$ .
- Step (2) A value for the parameter of  $\theta$ , say  $\theta_0$ , is generated from gamma( $c, d$ ) distribution using the IMSL [14] routine DRNGAM.
- Step (3) Using (3) and the IMSL [14] routine DRNUN together with some sorting routine, ordered sample from  $Bilal(\theta_0)$  distribution can be generated.
- Step (4) By using (12) and (22),  $\hat{\theta}_M, \hat{\theta}_s, \hat{\theta}_l, \hat{\theta}_e$  and  $\hat{\theta}_p$ , can be computed and stored.
- Step (5) A Gibb's sample of size  $M=10,000$  can be generated from the posterior density (15), using Devroye's algorithm [13]. Substituting  $g(\eta) = \frac{1}{\eta} = \theta$  in (23), and using this generated Gibb's sample,  $\tilde{\theta}_s, \tilde{\theta}_l$  (with  $a_1 = 0.5$  and  $a_1 = 1$ ),  $\tilde{\theta}_e$  and  $\tilde{\theta}_p$  can be calculated and stored.
- Step (6) By using (11), lower and upper bounds of the asymptotic  $100(1 - \alpha)\%$  confidence interval, ACI, of  $\hat{\theta}_M$  as well as its length be calculated and stored.
- Step (7) As described in Section 3.2, lower and upper bounds of the 95% HPD HPD symmetric credible interval for the Bayes estimator of  $\theta$  as well as its length can be calculated and stored.
- Step (8) The generated value  $\theta_0$ , which is obtained from Step (2), may be used as the true value of the parameter  $\theta$ . The biases as well as the squared errors for the estimators can then be calculated and stored.
- Step (9) Steps (2)–(8) were repeated 10,000 times. The average biased (AB) and estimated risk (ER) for MLE are calculated, respectively, as

$$AB(\hat{\theta}_M) = \frac{1}{K} \sum_{k=1}^K \left| \hat{\theta}_{M,k} - \theta_{0,k} \right|, \quad K = 10,000,$$

$$ER(\hat{\theta}_M) = \sqrt{\frac{1}{K} \sum_{k=1}^K \left( \hat{\theta}_{M,k} - \theta_{0,k} \right)^2},$$

where  $\hat{\theta}_{M,k}$  and  $\theta_{0,k}$ , respectively, are the MLE and the true value for the parameter  $\theta$  in the  $k$ th repetition. Similarly, the biases and ERs for the other estimators can be calculated. The averages of the lower and upper bounds and their corresponding length of the asymptotic confidence interval for the MLE as well as the symmetric credible interval for the Bayes estimator, which are obtained from Step (6) and Step (7), respectively, over the 10,000 runs, can be calculated and stored.

On the other hand, we consider also the case when no prior information about the parameter  $\theta$  are available. This is done by setting  $c = d = 0$  in (13), corresponding to the non-informative prior. Similar steps like the above ones are carried out, but the samples were generated using  $\theta_0 = 2$ . The numerical results are presented in Tables 2–5.

**Table 2**  
Average Biased and ERs of the MLs and Bayes estimates for  $\theta$ .

$n$	$r$	$\hat{\theta}_M$ AB (ER)	$\hat{\theta}_S$ AB (ER)	$\hat{\theta}_L, a_1 = 0.5$ AB (ER)	$\hat{\theta}_L, a_1 = 1$ AB (ER)	$\hat{\theta}_E$ AB (ER)	$\hat{\theta}_p$ AB (ER)	
15	9	0.3530 (0.5114)	I	0.3227 (0.4726)	0.3241 (0.4806)	0.3276 (0.4994)	0.3268 (0.4891)	0.3255 (0.4701)
			II	0.3259 (0.4702)	0.3244 (0.4803)	0.3277 (0.4983)	0.3236 (0.4792)	0.3321 (0.4725)
	12	0.3167 (0.4590)	I	0.2952 (0.4301)	0.2963 (0.4372)	0.2996 (0.4536)	0.2988 (0.4435)	0.2970 (0.4280)
			II	0.2970 (0.4279)	0.2962 (0.4368)	0.2992 (0.4526)	0.2958 (0.4351)	0.3014 (0.4295)
	15	0.2953 (0.4263)	I	0.2774 (0.4027)	0.2783 (0.4089)	0.2815 (0.4238)	0.2805 (0.4142)	0.2788 (0.4008)
			II	0.2791 (0.4005)	0.2782 (0.4081)	0.2811 (0.4224)	0.2779 (0.4064)	0.2828 (0.4019)
30	18	0.2481 (0.3645)	I	0.2350 (0.3470)	0.2351 (0.3498)	0.2367 (0.3593)	0.2361 (0.3532)	0.2365 (0.3465)
			II	0.2367 (0.3471)	0.2352 (0.3501)	0.2367 (0.3590)	0.2351 (0.3498)	0.2397 (0.3486)
	24	0.2239 (0.3234)	I	0.2147 (0.3128)	0.2151 (0.3162)	0.2166 (0.3250)	0.2157 (0.3184)	0.2156 (0.3121)
			II	0.2157 (0.3120)	0.2151 (0.3160)	0.2166 (0.3245)	0.2147 (0.3148)	0.2178 (0.3127)
	30	0.2091 (0.3014)	I	0.2017 (0.2911)	0.2020 (0.2933)	0.2031 (0.3005)	0.2025 (0.2955)	0.2025 (0.2906)
			II	0.2027 (0.2906)	0.2021 (0.2932)	0.2031 (0.3000)	0.2018 (0.2926)	0.2044 (0.2913)
60	36	0.1731 (0.2504)	I	0.1695 (0.2470)	0.1701 (0.2498)	0.1714 (0.2558)	0.1702 (0.2502)	0.1699 (0.2464)
			II	0.1700 (0.2467)	0.1702 (0.2501)	0.1715 (0.2561)	0.1697 (0.2485)	0.1709 (0.2469)
	48	0.1572 (0.2279)	I	0.1541 (0.2251)	0.1543 (0.2272)	0.1551 (0.2319)	0.1543 (0.2274)	0.1545 (0.2247)
			II	0.1545 (0.2249)	0.1544 (0.2274)	0.1551 (0.2319)	0.1541 (0.2262)	0.1554 (0.2250)
	60	0.1459 (0.2120)	I	0.1436 (0.2102)	0.1439 (0.2122)	0.1446 (0.2164)	0.1440 (0.2122)	0.1439 (0.2098)
			II	0.1440 (0.2098)	0.1439 (0.2121)	0.1445 (0.2161)	0.1437 (0.2109)	0.1446 (0.2098)

( $c = 5, d = 8, 10,000$  repetitions, simulated mean for  $\theta = 1.9983$ ), I results related to Tierney and Kadanes' Bayes approximation form, II results related to MCMC technique.

**Table 3**  
Same as in Table 2, but based on Jiffery's prior and  $\theta_0 = 2$ .

$n$	$r$	$\hat{\theta}_M$ AB (ER)	$\hat{\theta}_S$ AB (ER)	$\hat{\theta}_L, a_1 = 0.5$ AB (ER)	$\hat{\theta}_L, a_1 = 1$ AB (ER)	$\hat{\theta}_E$ AB (ER)	$\hat{\theta}_p$ AB (ER)	
15	9	0.3531 (0.4447)	I	0.3621 (0.4599)	0.3514 (0.4452)	0.3356 (0.4217)	0.3493 (0.4369)	0.3761 (0.4814)
			II	0.3759 (0.4812)	0.3519 (0.4462)	0.3366 (0.4232)	0.3535 (0.4455)	0.3962 (0.5095)
	12	0.3176 (0.4010)	I	0.3232 (0.4107)	0.3158 (0.4007)	0.3044 (0.3833)	0.3148 (0.3943)	0.3329 (0.4259)
			II	0.3334 (0.4267)	0.3165 (0.4017)	0.3052 (0.3845)	0.3176 (0.4008)	0.3475 (0.4468)
	15	0.2957 (0.3731)	I	0.2995 (0.3801)	0.2938 (0.3723)	0.2849 (0.3585)	0.2934 (0.3673)	0.3072 (0.3919)
			II	0.3081 (0.3932)	0.2946 (0.3735)	0.2856 (0.3596)	0.2956 (0.3727)	0.3191 (0.4089)
30	18	0.2480 (0.3119)	I	0.2515 (0.3176)	0.2480 (0.3130)	0.2419 (0.3040)	0.2465 (0.3089)	0.2566 (0.3252)
			II	0.2567 (0.3252)	0.2482 (0.3133)	0.2422 (0.3045)	0.2483 (0.3123)	0.2636 (0.3350)
	24	0.2247 (0.2820)	I	0.2270 (0.2856)	0.2246 (0.2824)	0.2201 (0.2759)	0.2234 (0.2795)	0.2306 (0.2909)
			II	0.2309 (0.2912)	0.2248 (0.2827)	0.2203 (0.2763)	0.2247 (0.2820)	0.2359 (0.2981)
	30	0.2097 (0.2630)	I	0.2112 (0.2656)	0.2093 (0.2631)	0.2058 (0.2579)	0.2086 (0.2609)	0.2140 (0.2697)
			II	0.2143 (0.2701)	0.2095 (0.2634)	0.2060 (0.2582)	0.2096 (0.2628)	0.2181 (0.2756)
60	36	0.1748 (0.2211)	I	0.1761 (0.2231)	0.1750 (0.2216)	0.1727 (0.2184)	0.1743 (0.2200)	0.1779 (0.2257)
			II	0.1781 (0.2258)	0.1752 (0.2218)	0.1730 (0.2186)	0.1751 (0.2214)	0.1805 (0.2292)
	48	0.1578 (0.1990)	I	0.1586 (0.2003)	0.1578 (0.1992)	0.1561 (0.1969)	0.1573 (0.1981)	0.1599 (0.2021)
			II	0.1600 (0.2023)	0.1579 (0.1994)	0.1563 (0.1971)	0.1578 (0.1991)	0.1617 (0.2047)
	60	0.1464 (0.1850)	I	0.1471 (0.1859)	0.1464 (0.1851)	0.1451 (0.1832)	0.1460 (0.1842)	0.1481 (0.1874)
			II	0.1484 (0.1877)	0.1467 (0.1854)	0.1453 (0.1835)	0.1466 (0.1851)	0.1498 (0.1897)

(I) results related to Tierney and Kadanes' Bayes approximation form, (II) results related to MCMC technique.

**5. Concluding remarks**

- (1) All of the obtained results can be adapted to the case of complete sample by taking  $r = n$ .
- (2) It may be clear from Tables 2 and 3 that, for fixed  $r$ , as  $n$  is increased the ABs as well as ERs of the MLs and/or Bayes estimates for  $\theta$  are decreased. On the other hand, Tables 4 and 5 that, for fixed  $r$ , as  $n$  is increased the average lengths of the 95% ACI and/or 95% HPD for  $\theta$  are decreased. This may indicate that all of the estimators of the parameter  $\theta$  are simultaneously consistent.
- (3) Table 2 may show that the ABs and ERs related to Tierney and Kadanes' Bayes approximation form, with respect to any of the four loss functions, are quite closed to their corresponding ones related to the Bayes estimators obtained by using the MCMC technique. Moreover, in all cases, the ABs and ERs related to ML method are bigger than the corresponding ones obtained using any of the Bayesian techniques. Furthermore, Table 2 may also show that, in most cases, the approximate Bayes estimators by using MCMC technique with respect to the squared error loss function, SLF, has the

**Table 4**  
Averages of 95% credible intervals for  $\theta$  and their corresponding average length.

$n$	$r$	ACI			HBD		
		Lower	Upper	(Average length)	Lower	Upper	(Average length)
15	9	1.1240	2.8685	(1.7445)	1.3473	2.9604	(1.6131)
	12	1.2141	2.7866	(1.5725)	1.3927	2.8616	(1.4688)
	15	1.2712	2.7311	(1.4599)	1.4236	2.7963	(1.3727)
30	18	1.3842	2.6198	(1.2357)	1.4959	2.6793	(1.1833)
	24	1.4454	2.5571	(1.1117)	1.5341	2.6051	(1.0710)
	30	1.4853	2.5179	(1.0326)	1.5607	2.5588	(0.9982)
60	36	1.5631	2.4350	(0.8719)	1.6182	2.4695	(0.8513)
	48	1.6075	2.3928	(0.7854)	1.6514	2.4197	(0.7683)
	60	1.6351	2.3646	(0.7295)	1.6724	2.3877	(0.7153)

( $c = 5, d = 8, 10,000$  repetitions, simulated mean for  $\theta = 1.9983$ ).

smallest estimated risk. Therefore, the estimator  $\tilde{\theta}_S$  may be recommended to be used when prior information about the parameter  $\theta$  are available.

**Table 5**  
Same as in Table 4, but based on Jiffery's prior and  $\theta_0 = 2$ .

n	r	ACI			HBD		
		Lower	Upper	(Average length)	Lower	Upper	(Average length)
15	9	1.1262	2.8742	(1.7480)	1.3538	3.2740	(1.9202)
	12	1.2166	2.7925	(1.5759)	1.4034	3.0933	(1.6899)
	15	1.2737	2.7366	(1.4628)	1.4356	2.9831	(1.5476)
30	18	1.3856	2.6226	(1.2370)	1.5087	2.8005	(1.2918)
	24	1.4475	2.5608	(1.1133)	1.5475	2.6971	(1.1496)
	30	1.4870	2.5206	(1.0337)	1.5732	2.6336	(1.0604)
60	36	1.5653	2.4384	(0.8732)	1.6302	2.5204	(0.8902)
	48	1.6093	2.3956	(0.7863)	1.6622	2.4586	(0.7965)
	60	1.6372	2.3676	(0.7304)	1.6827	2.4206	(0.7380)

- (4) For a fixed loss function, Table 3 may show that, in all cases, the method of Tierney and Kadane results in approximate Bayes estimator which is much better than the corresponding ones obtained by using the MCMC technique or ML method, in terms of ABs and ERs point of view. Moreover, the use of ML method produces estimate risks smaller than the corresponding ones by using the MCMC technique under any of four loss functions except LLF with  $\alpha_1 = 1$ . Moreover, Table 3 may also show that, in all cases, the approximate Bayes estimators form from Tierney and Kadane with respect to the LINEX loss function, LLF, (with  $\alpha_1 = 1$ ) has the smallest estimated risk. Therefore, the estimator  $\hat{\theta}_L$  with  $\alpha_1 = 1$  may be recommended to be used when no prior information about the parameter  $\theta$  are available.
- (5) This article provides classical and Bayesian approaches for finding an estimated value for the parameter  $\theta$  of *Bilal*( $\theta$ ) distribution based on a given Type-2 right censoring sample. One of these estimators are based on our provided EM algorithm, which can be readily applied to the progressive Type-2 censoring scheme. When prior information about the parameter  $\theta$  are available, the use of the MCMC technique, for obtaining Bayes estimator for  $\theta$ , can be readily utilized for obtaining prediction intervals for unobserved lifetimes in the same sample, which follows *Bilal*( $\theta$ ) model; and in a future sample from the same population based on a Type-2 censored sample. Work is in proses with these further researches.

## Acknowledgment

The authors would like to express their sincere thanks to the editor and referees for various suggestions and helpful comments, which improved the current presentation of the paper.

## Appendix

*Proof of existence and uniqueness of the solution of (5):*

It follows from (5) that,

$$\frac{\partial^2 \ln(\ell)}{\partial \eta^2} = -\frac{r}{\eta^2} - \frac{6(n-r)x_{(r)}^2 e^{-x_{(r)}\eta}}{(3-2e^{-x_{(r)}\eta})^2} - \sum_{i=1}^r \frac{x_{(i)}^2 e^{-x_{(i)}\eta}}{(1-e^{-x_{(i)}\eta})^2} < 0.$$

This may imply that, the MLE,  $\hat{\theta}_{ML} = \frac{1}{\hat{\eta}_{ML}}$ , for  $\eta$  is unique. To insure that  $\hat{\eta}_{ML}$  exists, following Balakrishnan and Kateri [15] with a modification, we rewrite (5) as  $h_1(\eta) = h_2(\eta)$ , where  $h_1(\eta) = \frac{r}{\eta}$  and

$$h_2(\eta) = 2(n-r)x_{(r)} \left\{ 1 - \frac{e^{-\eta x_{(r)}}}{3-2e^{-\eta x_{(r)}}} \right\} + \sum_{i=1}^r x_{(i)} \left\{ 3 - (1 - e^{-\eta x_{(i)}})^{-1} \right\}.$$

Since  $\frac{\partial h_2(\eta)}{\partial \eta} = \sum_{i=1}^r \frac{x_{(i)}^2 e^{-x_{(i)}\eta}}{(1-e^{-x_{(i)}\eta})^2} + \frac{6(n-r)x_{(r)}^2 e^{-x_{(r)}\eta}}{(3-2e^{-x_{(r)}\eta})^2} > 0$ ,  $\lim_{\eta \rightarrow 0^+} h_2(\eta) = -\infty$  and  $\lim_{\eta \rightarrow \infty} h_2(\eta) = \sum_{i=1}^r 2x_{(i)} + 2x_{(r)}(n-r) > 0$ ,  $h_2(\eta)$  is then a pounded increasing function of  $\eta$ . But  $h_1(\eta)$  is a positive strictly decreasing function with right limit  $+\infty$  at 0. This insure that  $h_1(\eta) = h_2(\eta)$  holds exactly once at some value  $\eta = \eta^*$ .

## References

- [1] L. Tierney, J.B. Kadane, Accurate approximations for posterior moments and marginal densities, *J. Amer. Statist. Assoc.* 81 (1986) 82–86.
- [2] A.M. Abd-Elrahman, Utilizing ordered statistics in lifetime distributions production: a new lifetime distribution and applications, *J. Probability Stat. Sci.* 11 (2) (2013) 153–164.
- [3] D. Hinkley, On quick choice of power transformations, *Appl. Stat.* 26 (1977) 67–69.
- [4] M.E. Ghitany, B. Atieh, S. Nadarajah, Lindley distribution and its application, *Math. Comput. Simul.* 78 (4) (2008) 493–506.
- [5] H. Zakerzadeh, E. Mahmoudi, A new two parameter lifetime distribution: model and properties, 2012, 1204.4248v1 [stat.CO].
- [6] H.K.T. Ng, P.S. Chan, N. Balakrishnan, Estimation of parameters from progressively censored data using EM algorithm, *Comput. Stat. Data Anal.* 39 (2002) 371–386.
- [7] A.M. Abd-Elrahman, Asymptotic variance-covariance matrix of the ML estimates based on general progressive censored data using EM algorithm, *J. Probability Stat. Sci.* 5 (2) (2007) 171–185.
- [8] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, *J. Roy. Statist. Soc. B* 39 (1977) 1–38.
- [9] H.R. Varian, *A Bayesian Approach to Real Estate Assessment*, 1975, North Holland, Amsterdam.
- [10] R. Calabria, G. Pulcini, An engineering approach to bayes estimation for the weibull distribution, *Micro-electron. Reliab.* 34 (5) (1994) 789–802.
- [11] J.G. Norstrom, The use of precautionary loss functions in risk analysis, *IEEE Trans. Reliab.* 45 (3) (1996) 400–403.
- [12] B. Pradhan, P. Kundu, Bayes estimation and prediction of the two-parameter gamma distribution, *J. Stat. Comput. Simul.* 81 (9) (2011) 1187–1198.
- [13] L. Devroye, A simple algorithm for generating random variates with a log-concave density function, *Computing* 33 (1984) 247–257.
- [14] IMSL, *IMSL Reference Manual*, IMSL, Inc., Houston, TX., 1995.
- [15] N. Balakrishnan, M. Kateri, On the maximum likelihood estimation of parameters of weibull distribution based on complete and censored data, *Stat. Probability Lett.* 78 (2008) 2971–2975.