Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v7n2p326

Bayesian analysis of censored Burr XII distribution

By Danish M. Y., Aslam M.

Published: 14 October 2014

Electronic Journal of Applied Statistical Analysis
Vol. 07, Issue 02, 2014, 326-342
DOI: 10.1285/i20705948v7n2p326

Bayesian analysis of censored Burr XII distribution

Muhammad Yameen Danish*[a] and Muhammad Aslam[b]

[a] *Department of Mathematics and Statistics, AIOU, Islamabad, Pakistan*
[b] *Department of Statistics, Quaid-i-Azam University, Islamabad, 45320, Pakistan*

Published: 14 October 2014

The paper deals with Bayesian estimation of unknown parameters of Burr type XII distribution under the Koziol-Green model of random censorship assuming both the informative and noninformative priors. We use different symmetric and asymmetric loss functions to obtain the Bayes estimates. It is seen that the closed-form expressions for the Bayes estimators cannot be obtained; we propose Gibbs sampling scheme to obtain the approximate Bayes estimates. Monte Carlo simulation is carried out to observe the behavior of the proposed estimators and also to compare with the maximum likelihood estimators. Based on the simulation study, we propose a set of estimators of the model parameters. One real data analysis is performed and it is seen that the proposed set of estimators fit the data best than the rest.

keywords: Random censoring, Bayes estimate, log-concave density, Gibbs sampling, Markov chain Monte Carlo.

1 Introduction

There are several types of censoring schemes. The most popular among these are the type I and type II censoring schemes. In these censoring schemes, objects on test are removed from the test at the final termination of the test. In situations where it is desirable to remove the objects from the test other than the final termination point, random censoring scheme provide a suitable plan. Random censoring scheme is an important type of right censoring in which the time of censoring is not fixed but taken as random. In a clinical trial, for example, patients often enter into the study after some medical operation,

*Corresponding author: yameendanish42sb@yahoo.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

therefore the enrolment time and hence the censoring time is random. In some medical studies and longitudinal designs, individuals enter into the study simultaneously but the censoring time depends on other random factors e.g., patients lost to follow-up, drop out of the study, etc. The type I censoring scheme can be considered as a special case of random censoring scheme in which censoring takes place at some fixed time point.

Let $X_1, X_2, ..., X_n$ be independent identically distributed random variables with distribution function F(x), density function f(x) and let $T_1, T_2, ..., T_n$ be also independent identically distributed random variables with distribution function G(t) and density function g(t). In the context of reliability and life testing experiments, X_i's are the true survival times of n individuals censored by T_i's from the right. The experiment thus results in independent identically distributed random pairs $(Y_1, D_1), ..., (Y_n, D_n)$ where $Y_i = \text{Min}(X_i, T_i)$ and $D_i = \text{I}(X_i \leq T_i)$ is indicator of noncensored observation, for $i = 1, 2, ..., n$. In random censorship model it is assumed that X_i and T_i, Y_i and D_i are independent. Now it is simple to show that the joint density function of Y and D is

$$f_{Y,D}(y, d) = \{f_X(y)(1 - G_T(y))\}^d \{g_T(y)(1 - F_X(y))\}^{1-d}; \qquad y \geq 0, d = 0, 1. \tag{1}$$

In Green (1976) of random censorship, variables X and T satisfy the relation for some $\beta > 0$

$$1 - G_T(y) = \{1 - F_X(y)\}^\beta \tag{2}$$

The relation (2) differentiates the Koziol-Green model from the general model of random censorship. From (1) and (2), we have

$$f_{Y,D}(y, d) = f_X(y)\{1 - F_X(y)\}^\beta \beta^{1-d}; \qquad y > 0, d = 0, 1. \tag{3}$$

The Burr type XII is the most important distribution among the twelve distributions introduced by Burr (1942). It covers a variety of curve shapes and provides a wide range of values of skewness and kurtosis that can be used to model any general lifetime data. The probability density and cumulative distribution functions of Burr type XII distribution are

$$f_X(x; \theta, \lambda) = \theta\lambda x^{\lambda-1}(1 + x^\lambda)^{-\theta-1}; \qquad x > 0, \lambda > 0, \tag{4}$$

$$F_X(x; \theta, \lambda) = 1 - (1 + x^\lambda)^\theta. \tag{5}$$

For the density function in (4) and the distribution function in (5), the expression in (3) takes the following form

$$f_{Y,D}(y, d; \theta, \beta) = \theta\lambda y^{\lambda-1}(1 + y^\lambda)^{-\theta(1+\beta)-1}\beta^{1-d}; \qquad y, \theta, \lambda, \beta > 0, d = 0, 1. \tag{6}$$

The marginal distributions of Y and D can be obtained from (6) as

$$f_Y(y; \theta, \lambda, \beta) = \theta\lambda(1 + \beta)y^{\lambda-1}(1 + y^\lambda)^{-\theta(1+\beta)-1}; \qquad y, \theta, \lambda, \beta > 0,$$

$$f_D(d;p) = p^d(1-p)^{1-d}; \qquad 0 \le p \le 1,$$

where $p = P(X_i \le T_i) = \frac{1}{1+\beta}$.

Some of the important references on the topic include the following. Liang (2004) considered the exponential distribution under Koziol-Green model of random censorship with known censoring parameter. Friesl (2007) dealt with Bayesian estimation in exponential distribution under random censorship and investigated the asymptotic properties of the estimators with particular stress on the Bayesian risk. Nigm (1988) discussed the prediction bounds for the Burr model. Rampone and Ruse (2012) obtained the approximate Bayes estimates of the Burr XII distribution using Lindleys approximation. Upadhyay (2004) considered Bayesian analysis of generalized four-parameter Burr distribution for complete and censored samples via Gibbs sampler. Li and Chai (2007) focused on the empirical Bayes estimators of reliability performances for progressively type II censored Burr XII distribution. Hussaini and Hussein (2011) provided a number of references on the applications of Burr model in different fields of applied statistics and derived the Bayes estimates using censored data from exponentiated Burr type XII population.

Although extensive work has been done on the statistical inferences of the unknown parameters of the Burr distribution, however, this type of work has not been addressed up to now. The aim of the paper is to obtain the Bayes estimators of the unknown parameters of the model derived in (6) under different loss functions assuming different sets of priors.

The rest of the paper is organized as follows. In Section 2, we derive the maximum likelihood (ML) estimators of the unknown parameters. Section 3 contains the prior distributions, loss functions, Bayes estimates using Gibbs sampling scheme. A simulation study is considered in Section 4. A real data set is analyzed in Section 5 and finally, we conclude the paper in Section 6.

2 Maximum Likelihood Estimation of Parameters

In this section, we derive the ML estimators $\hat{\theta}$,$\hat{\lambda}$ and $\hat{\beta}$ of θ, λ and β, respectively, assuming the model defined in (6) holds. For the observed sample $(y_1, d_1), (y_2, d_2), ..., (y_n, d_n) = (y, d)$ of size n from (6), the likelihood function can be written as

$$l(\theta, \lambda, \beta; (y,d)) = \theta^n \lambda^n \prod_{i=1}^{n} y_i^{\lambda-1} \prod_{i=1}^{n} (1+y_i^{\lambda})^{-\theta(1+\beta)-1} \beta^{n-\sum_{i=1}^{n} d_i}. \tag{7}$$

Thus the log-likelihood function is

$$L(\theta,\lambda,\beta;(y,d)) = \ln\left(l(\theta,\lambda,\beta;(y,d))\right) = n\ln\theta + n\ln\lambda + \left(n - \sum_{i=1}^{n} d_i\right)\ln\beta$$
$$- (1+\theta(1+\beta))\sum_{i=1}^{n}\ln(1+y_i^{\lambda}) + (\lambda-1)\sum_{i=1}^{n}\ln y_i. \tag{8}$$

Differentiating (6) with respect to θ, λ and β, the three normal equations thus obtained are

$$\frac{n}{\theta} - (1+\beta)\sum_{i=1}^{n}\ln(1+y_i^{\lambda}) = 0, \tag{9}$$

$$\frac{n}{\lambda} - \theta(1+\beta)\sum_{i=1}^{n}\frac{y_i^{\lambda}\ln y_i}{1+y_i^{\lambda}} + \sum_{i=1}^{n}\frac{\ln y_i}{1+y_i^{\lambda}} = 0, \tag{10}$$

$$\frac{n - \sum_{i=1}^{n} d_i}{\beta} - \theta\sum_{i=1}^{n}(1+y_i^{\lambda}) = 0. \tag{11}$$

Solving these equations simultaneously, we have

$$\hat{\beta} = \frac{n - \sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} d_i}, \tag{12}$$

$$\theta(\lambda) = \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n}\ln(1+y_i^{\lambda})}, \tag{13}$$

$$\lambda = h(\lambda), \tag{14}$$

where

$$h(\lambda) = \left[\frac{\sum_{i=1}^{n}\frac{y_i^{\lambda}\ln y_i}{1+y_i^{\lambda}}}{\sum_{i=1}^{n}\ln(1+y_i^{\lambda})} - \frac{1}{n}\sum_{i=1}^{n}\frac{\ln y_i}{1+y_i^{\lambda}}\right]^{-1}. \tag{15}$$

Some numerical procedures are required to solve (13) and (14). We suggest the following procedure: Start with a suitable initial value of λ, say $\lambda^{(0)}$, and obtain $\lambda^{(1)}$ from $\lambda^{(1)} = h(\lambda^{(0)}), \lambda^{(2)}$ from $\lambda^{(2)} = h(\lambda^{(1)})$ and finally $\lambda^{(n)}$ from $\lambda^{(n)} = h(\lambda^{(n-1)})$. Stop the process when $|\lambda^{(n)} - \lambda^{(n-1)}| < \epsilon$, where ϵ is a pre-assigned tolerance limit. Once the MLE of λ is obtained, the MLE of θ can be obtained from (13).

3 Bayesian Estimation of Parameters

In this section, we discuss prior distributions for unknown parameters, loss functions and Bayes estimates.

3.1 Prior distributions

The Bayesian analysis requires the choice of appropriate priors for the unknown parameters in addition to the experimental data. Arnold and Press (1983) correctly pointed out that there is no clear cut way in which one can say that one prior is better than the other. The important thing in this connection is the relationship between the prior distribution and the loss function. The model under consideration has two shape parameters and one censoring parameter, and continuous conjugate priors for these parameters do not exist. Nevertheless, we consider both the informative and noninformative priors and observe the results. First, we assume the following independent gamma priors for θ, λ and β

$$\left.\begin{aligned} \pi_1(\theta) &= \tfrac{b_1^{a_1}}{\Gamma(a_1)}\theta^{a_1-1}e^{-b_1\theta}; \quad a_1, b_1, \theta > 0 \\ \pi_2(\lambda) &= \tfrac{b_2^{a_2}}{\Gamma(a_2)}\lambda^{a_2-1}e^{-b_2\lambda}; \quad a_2, b_2, \lambda > 0 \\ \pi_3(\beta) &= \tfrac{b_3^{a_3}}{\Gamma(a_3)}\beta^{a_3-1}e^{-b_3\beta}; \quad a_3, b_3, \beta > 0. \end{aligned}\right\} \tag{16}$$

The assumption of independent gamma priors is not unreasonable. Many authors have used these priors on the scale and the shape parameters of the two-parameter lifetime distributions, see Gupta and Kundu (2006), Kundu (2008), Wahed (2006). It is to be noted that the noninformative priors on the scale and the shape parameters are the special cases of independent gamma priors. The joint prior density of the unknown parameters can be written from (16) as

$$\pi(\theta, \lambda, \beta) \propto \theta^{a_1-1}e^{-b_1\theta}\lambda^{a_2-1}e^{-b_2\lambda}\beta^{a_3-1}e^{-b_3\beta} \tag{17}$$

Second, we consider a conditional gamma prior for θ given λ and a gamma prior for λ, and an independent beta prior for p as

$$\left.\begin{aligned} \pi_4(\theta|\lambda) &= \tfrac{\lambda^{a_4}}{\Gamma(a_4)}\theta^{a_4-1}e^{-\lambda\theta}; \quad a_4 > 0 \\ \pi_5(\lambda) &= \tfrac{b_5^{a_5}}{\Gamma(a_5)}\lambda^{a_5-1}e^{-b_5\lambda}; \quad a_5, b_5 > 0 \\ \pi_6(p) &= \tfrac{\Gamma(a_6+b_6)}{\Gamma(a_6)\Gamma(b_6)}p^{a_6-1}(1-p)^{b_6-1}; \quad a_6, b_6 > 0 \end{aligned}\right\} \tag{18}$$

For the transformation $p = \dfrac{1}{1+\beta}$, the prior density for β is obtained from $\pi_6(p)$ as

$$\pi_7(\beta) = \frac{\Gamma(a_6+b_6)}{\Gamma(a_6)\Gamma(b_6)}\frac{\beta^{b_6-1}}{(1+\beta)^{a_6+b_6}}. \tag{19}$$

Thus the joint prior density for θ, λ and β in this case is

$$g(\theta, \lambda, \beta) \propto \frac{\theta^{a_4-1} e^{-\lambda\theta} \lambda^{a_4+a_5-1} e^{-b_5\lambda} \beta^{b_6-1}}{(1+\beta)^{a_6+b_6}}. \tag{20}$$

This prior density is formulated following Upadhyay (2004), Friesl (2007) and Hussaini and Hussein (2011).

3.2 Loss functions

In order to select a best decision in decision theory, an appropriate loss function must be specified. The most commonly used loss function is the squared error (SE) loss function defined by $l_1(\hat{\theta}_{SE}, \theta) = (\hat{\theta}_{SE} - \theta)^2$, where $\hat{\theta}_{SE}$ is a decision rule to estimate parameter θ. The Bayes estimator under SE loss function is

$$\hat{\theta}_{SE} = E(\theta), \tag{21}$$

where E denotes the expectation with respect to the posterior distribution of θ.

The SE loss function is used when the loss is symmetric with respect to over estimation and under estimation of equal magnitude. When the true loss is not symmetric with respect to over estimation and under estimation, then the asymmetric loss functions are used to represent the consequences of different errors (Zellner (1986), Norstrom (1996)). Since there is no specific way to identify a suitable loss function for a particular problem, we, therefore, consider symmetric as well as asymmetric loss functions in our Bayesian analysis.

The second loss function is the asymmetric precautionary (AP) loss function defined by $l_2(\hat{\theta}_{AP}, \theta) = \frac{(\hat{\theta}_{AP}-\theta)^2}{\hat{\theta}_{AP}}$. This loss function is a special case of the general class of precautionary loss functions introduced by Norstrom (1996). The Bayes estimator under AP loss function is

$$\hat{\theta}_{AP} = [E(\theta^2)]^{\frac{1}{2}}. \tag{22}$$

The third loss function is the quadratic loss function which is defined as $l_3(\hat{\theta}_Q, \theta) = \frac{(\hat{\theta}_Q-\theta)^2}{\theta}$. The Bayes estimator under quadratic loss function is

$$\hat{\theta}_Q = \frac{E(\theta^{-1})}{E(\theta^{-2})}. \tag{23}$$

The last one is the squared-log error (SLE) loss function defined by

$$l_4(\hat{\theta}_{SE}, \theta) = (\ln \hat{\theta}_{SE} - \ln \theta)^2 = \left(\ln \frac{\hat{\theta}_{SE}}{\theta}\right)^2.$$

The Bayes estimator under SLE loss function is

$$\hat{\theta}_{SE} = \exp[E(\ln \theta)]. \tag{24}$$

3.3 The Bayes estimators under $\pi(\theta, \lambda, \beta)$

In this section, we obtain the Bayes estimates of θ, λ and β under different loss functions assuming the prior density as in (17). Combining the likelihood function in (7) and the prior density in (17), the joint posterior density function of θ, λ and β given data is

$$\pi(\theta, \lambda, \beta|(y, d)) \propto \theta^{n+a_1-1} e^{-\theta\left(b_1+\sum_{i=1}^{n} \ln(1+y_i^\lambda)\right)} \lambda^{n+a_2-1} e^{-\lambda b_2}$$
$$\prod_i^n \frac{y_i^{\lambda-1}}{1+y_i^\lambda} \beta^{n-\sum_{i=1}^{n} d_i+a_3-1} e^{-\beta\times\left(b_3+\theta\sum_{i=1}^{n} \ln(1+y_i^\lambda)\right)}. \tag{25}$$

Thus the posterior expectation of any function of parameters, say $U(\theta, \lambda, \beta)$ can be written as

$$\hat{U}_B(\theta, \lambda, \beta) = E(U(\theta, \lambda, \beta)|(y, d)) = \frac{\int_0^\infty\int_0^\infty\int_0^\infty U(\theta, \lambda, \beta)\pi(\theta, \lambda, \beta|(y, d))d\theta d\lambda d\beta}{\int_0^\infty\int_0^\infty\int_0^\infty \pi(\theta, \lambda, \beta|(y, d))d\theta d\lambda d\beta}. \tag{26}$$

However, it is not possible to evaluate (26) in closed-form. We use Gibbs sampling scheme to obtain the Bayes estimates.

The full conditional forms of θ, λ and β can be obtained from (25) up to proportionality as

$$\pi_1(\theta|\lambda, \beta, (y, d)) \propto \theta^{n+a_1-1} e^{-\theta\left(b_1+(1+\beta)\sum_{i=1}^{n} \ln(1+y_i^\lambda)\right)}, \tag{27}$$

$$\pi_2(\lambda|\theta, \beta, (y, d)) \propto \lambda^{n+a_2-1} e^{-\lambda b_2} e^{-\theta(1+\beta)\sum_{i=1}^{n} \ln(1+y_i^\lambda)} \prod_i^n \frac{y_i^{\lambda-1}}{1+y_i^\lambda}, \tag{28}$$

$$\pi_3(\beta|\lambda, \theta, (y, d)) \propto \beta^{n-\sum_{i=1}^{n} d_i+a_3-1} e^{\beta(b_3+\theta\sum_{i=1}^{n} \ln(1+y_i^\lambda))}, \tag{29}$$

The full conditional forms (27) and (29) are the gamma densities, so the samples of θ and β can be easily generated using any of the gamma generating routines. The full conditional form (28) is log-concave since

$$\frac{\partial^2(\pi_2(\lambda|\theta, \beta, (y, d)))}{\partial\lambda^2} = \frac{-(n+a_2-1)}{\lambda^2} - (1+\theta(1+\beta))\sum_{i=1}^{n} \frac{y_i^\lambda(\ln y_i)^2}{(1+y_i^\lambda)^2} < 0.$$

Thus the samples of λ can be generated using the method suggested by Devroye (1984). Now following the idea of Geman (1984) and using (27), (28), (29), it is possible to generate samples of (θ, λ, β) from posterior distribution (25) and then to obtain the Bayes estimates. Starting with suitable choice of initial values, say $(\theta_0, \lambda_0, \beta_0)$, we suggest the following procedure to generate the posterior samples and then to obtain the Bayes estimates:

Step 1: Generate θ_1 from gamma $\left(n + a_1, b_1 + (1 + \beta_0)\sum_{i=1}^{n} \ln(1 + y_i^{\lambda_0})\right)$.

Step 2: Generate λ_1 from the log-concave density function (28) using the method suggested by Devroye (1984).

Step 3: Generate β_1 from gamma $\left(n - \sum\limits_{i=1}^{n} d_i + a_3, b_3 + \theta_0 \sum\limits_{i=1}^{n} \ln(1 + y_i^{\lambda_0})\right)$.

Step 4: Repeat Steps 1-3 M times to obtain $(\lambda_1, \theta_1, \beta_1), ..., (\lambda_M, \theta_M, \beta_M)$.

Now the approximate Bayes estimates of θ, λ and β under SE loss function can obtained from

$$\hat{\theta}_{SE} = \frac{1}{M}\sum_{j}^{M}\theta_j, \qquad \hat{\lambda}_{SE} = \frac{1}{M}\sum_{j}^{M}\lambda_j, \qquad \text{and} \qquad \hat{\beta}_{SE} = \frac{1}{M}\sum_{j}^{M}\beta_j.$$

The approximate Bayes estimates of θ, λ and β under AP loss function can be obtained from

$$\hat{\theta}_{AP} = \left(\frac{1}{M}\sum_{j}^{M}\theta_j\right)^{\frac{1}{2}}, \qquad \hat{\lambda}_{AP} = \left(\frac{1}{M}\sum_{j}^{M}\lambda_j\right)^{\frac{1}{2}}, \qquad \text{and} \qquad \hat{\beta}_{AP} = \left(\frac{1}{M}\sum_{j}^{M}\beta_j\right)^{\frac{1}{2}}.$$

The approximate Bayes estimates of θ, λ and β under quadratic loss function can be obtained from

$$\hat{\theta}_{Q} = \frac{\frac{1}{M}\sum\limits_{j}^{M}\frac{1}{\theta_j}}{\frac{1}{M}\sum\limits_{j}^{M}\frac{1}{\theta_j^2}}, \qquad \hat{\lambda}_{Q} = \frac{\frac{1}{M}\sum\limits_{j}^{M}\frac{1}{\lambda_j}}{\frac{1}{M}\sum\limits_{j}^{M}\frac{1}{\lambda_j^2}} \qquad \text{and} \qquad \hat{\beta}_{Q} = \frac{\frac{1}{M}\sum\limits_{j}^{M}\frac{1}{\beta_j}}{\frac{1}{M}\sum\limits_{j}^{M}\frac{1}{\beta_j^2}}.$$

The approximate Bayes estimates of θ, λ and β under SLE loss function can be obtained from

$$\hat{\theta}_{SLE} = \exp\left(\frac{1}{M}\sum_{j}^{M}\ln\theta_j\right), \qquad \hat{\lambda}_{SLE} = \exp\left(\frac{1}{M}\sum_{j}^{M}\ln\lambda_j\right),$$
$$\text{and} \qquad \hat{\beta}_{SLE} = \exp\left(\frac{1}{M}\sum_{j}^{M}\ln\beta_j\right).$$

3.4 The Bayes estimators under $g(\theta, \lambda, \beta)$

In this section, we obtain the Bayes estimates of θ, λ and β under different loss functions assuming the prior density as in (20). The joint posterior density function of θ, λ and β given data is obtained by combining the likelihood function in (7) and the joint prior density in (20) as

$$\pi(\theta, \lambda, \beta|(y, d)) \propto \theta^{n+a_4-1} e^{-\theta\left(\lambda + (1+\beta)\sum\limits_{i=1}^{n}\ln(1+y_1^{\lambda})\right)} \lambda^{n+a_4+a_5-1} e^{-\lambda b_5} \frac{\beta^{n-\sum\limits_{i=1}^{n} d_i + b_6 - 1}}{(1+\beta^{a_6+b_6})} \prod_{i=1}^{n}\frac{y_i^{\lambda-1}}{1+y_i^{\lambda}}. \tag{30}$$

Thus the Bayes estimate of any function of parameters, say $U(\theta,\lambda,\beta)$ can be written as

$$\hat{U}_B(\theta,\lambda,\beta) = E(U(\theta,\lambda,\beta)|(y,d)) = \frac{\int_0^\infty\int_0^\infty\int_0^\infty U(\theta,\lambda,\beta)\pi(\theta,\lambda,\beta|(y,d))d\theta d\lambda d\beta}{\int_0^\infty\int_0^\infty\int_0^\infty \pi(\theta,\lambda,\beta|(y,d))d\theta d\lambda d\beta}. \tag{31}$$

However, it is not possible to evaluate (31) in closed-form. We again use the Gibbs sampling procedure to approximate the Bayes estimates. The full conditional distributions of θ,λ and β can be obtained from (30) as

$$\pi_1(\theta|\lambda,\beta,(y,d)) \propto \theta^{n+a_4-1} e^{-\theta\left(\lambda+(1+\beta)\sum_{i=1}^n \ln(1+y_i^\lambda)\right)}, \tag{32}$$

$$\pi_2(\lambda|\theta,\beta,(y,d)) \propto \theta^{n+a_4+a_5-1} e^{-\lambda(b_5+\theta)} e^{-\theta(1+\beta)\sum_{i=1}^n \ln(1+y_i^\lambda)} \prod_i^n \frac{y_i^{\lambda-1}}{1+y_i^\lambda}, \tag{33}$$

$$\pi_3(\beta|\theta,\lambda,(y,d)) \propto \frac{\beta^{n-\sum_{i=1}^n d_i+a_6-1}}{(1+\beta)^{a_6+b_6}} e^{-\beta\theta\sum_{i=1}^n \ln(1+y_i^\lambda)}. \tag{34}$$

The full conditional form (32) is gamma density, so the samples of θ can be easily generated. The full conditional forms (33) and (34) are log-concave since

$$\frac{\partial^2(\pi_2(\lambda|\theta,\beta,(y,d)))}{\partial\lambda^2} = -\frac{(n+a_4+a_5-1)}{\lambda^2} - (1+\theta(1+\beta))\sum_{i=1}^n \frac{y_i^\lambda(\ln y_i)^2}{(1+y_i^\lambda)^2} < 0,$$

$$\frac{\partial^2(\pi_3(\beta|\theta,\lambda,(y,d)))}{\partial\beta^2} = -\frac{(n-\sum_{i=1}^n d_i+b_6-1)}{\beta^2} + \frac{a_6+b_6}{(1+\beta)^2} < 0, \forall \quad n-\sum_{i=1}^n d_i - 1 \geq b_6.$$

Thus the samples of λ and β can be generated using the method suggested by Devroye (1984). Starting with suitable choice of initial values, say $(\theta^0,\lambda^0,\beta^0)$, we suggest the following procedure to generate samples of (θ,λ,β) from posterior distribution (30) and then to obtain the Bayes estimates:

Step 1: Generate θ_1 from gamma $\left(n+a_4, \lambda^0+(1+\beta^0)\sum_{i=1}^n \ln(1+y_i^{\lambda^0})\right)$.

Step 2: Generate λ_1 and β_1 from the log-concave densities (33) and (34), respectively, using the method suggested by Devroye (1984).

Step 3: Repeat Steps 1 and 2 M times to obtain $(\lambda_1,\theta_1,\beta_1),...,(\lambda_M,\theta_M,\beta_M)$.

Now the Bayes estimates of θ,λ and β under different loss functions can be obtained following the procedure given in previous section.

4 Simulation

In this section, we perform a Monte Carlo simulation to observe the behavior of the proposed estimators of the parameters for different sample sizes, for different priors, for different loss functions and for different proportions of uncensored observations. We consider different sample sizes: $n = 20, 40, 60$; different proportions of uncensored observations: $p = 0.50, 0.80$; different sets of parameter values: $\theta = 2, \lambda = 1.5, \beta = 1, \theta = 2, \lambda = 1.5, \beta = 0.25$; and different combinations of hyperparameters: $a_1 = 0, b_1 = 0, a_2 = 0, b_2 = 0, a_3 = 0, b_3 = 0$ (prior-1), $a_1 = 4, b_1 = 2, a_2 = 3, b_2 = 2, a_3 = 2, b_3 = 2$ (prior-2), $a_4 = 3, a_5 = 3, b_5 = 2, a_6 = 4, b_6 = 3$ (prior-3) when $\theta = 2, \lambda = 1.5, \beta = 1$ and prior-1, $a_1 = 4, b_1 = 2, a_2 = 3, b_2 = 2, a_3 = 1, b_3 = 4$ (prior-2), $a_4 = 3, a_5 = 3, b_5 = 2, a_6 = 9, b_6 = 2$ (prior-3) when $\theta = 2, \lambda = 1.5, \beta = 0.25$. Where prior-1 denotes the noninformative priors for θ, λ and β when all the hyperparameters in (17) are zero, prior-2 and prior-3 are defined in (17) and (20). It is to be noted that in prior-2 and prior-3, the hyperparameters are taken so that the priors means are the same as the original means.

To generate random samples from (6), we suggest the following procedure:

1. Generate N uniform $(0, 1)$ random numbers $u_1, ..., u_N$ and compute the corresponding $x_1, ..., x_N$, where $x_i = ((1 - u_i)^{-\theta} - 1)^{\frac{1}{\lambda}}$.
2. Generate another set of N uniform $(0, 1)$ random numbers $v_1, ..., v_N$ and compute the corresponding $t_1, ..., t_N$, where $t_i = ((1 - v_i)^{-\theta\beta} - 1)^{\frac{1}{\lambda}}$.
3. Obtain $d_1, ..., d_N$, where $d_i = 1$ if $x_i \leq t_i$ and $d_i = 0$ otherwise.
4. Compute $y_1, ..., y_N$, where $y_i = x_i d_i + t_i(1 - d_i)$.
5. Obtain the required number of samples of required size from the generated pairs (y_i, d_i), for $i = 1, ..., N$.

For a particular combination, we generate 1000 randomly censored samples from (6) and for each sample we compute the Bayes estimates under different loss functions based on 20,000 MCMC samples with 10,000 samples as burn-in period. The average values of the MLEs and the Bayes estimates and the corresponding mean square errors (MSEs) are obtained from these 1000 replications. The results are reported in Tables 1-3. Some of the points are very clear from these results. It is observed that as the sample size increases the MSEs of the estimators decrease in all cases. This rate of decrease in MSEs is relatively higher for informative priors as compared to noninformative priors. This is true for both the cases of 50% and 80% proportions of noncensored observations. It is further observed that the performance of the Bayes estimators of θ, λ and β is different under different loss functions. That is for the shape parameter θ, the Bayes estimator of θ under squared log-error loss function performs better than the rest. For the shape parameter λ, the Bayes estimator of λ under quadratic loss function performs better than the rest. For the censoring parameter β, the ML estimator of β performs better than the rest. It is also noted that the noinformative prior based Bayes estimators of θ, λ under AP loss function perform worse than the corresponding ML estimators and perform better under the other loss functions. When comparing the Bayes estimators under prior2 and prior3, it is seen that the Bayes estimators of parameter θ under prior2 perform better than the corresponding Bayes estimators under prior3 and the Bayes estimators of λ, β under prior3 perform better than the corresponding Bayes estimators under prior2.

Table 1: Average values of the different estimators of θ and the corresponding MSEs (in parenthesis).

Prior	**p**	**n**	$\hat{\theta}_{SE}$	$\hat{\theta}_{Q}$	$\hat{\theta}_{AP}$	$\hat{\theta}_{SLE}$	$\hat{\theta}$
1	0.50	20	2.2451 (0.6132)	1.7189 (0.3304)	2.3815 (0.8102)	2.1101 (0.4698)	2.2899 (0.8171)
		40	2.1098 (0.1758)	1.8695 (0.1343)	2.1704 (0.2063)	2.0492 (0.1534)	2.1147 (0.1825)
		60	2.0603 (0.1015)	1.9054 (0.0884)	2.0992 (0.1128)	2.0213 (0.0935)	2.0686 (0.1067)
	0.80	20	2.1139 (0.3633)	1.8308 (0.2780)	2.1844 (0.4143)	2.0429 (0.3241)	2.1304 (0.3664)
		40	2.0456 (0.1449)	1.9120 (0.1311)	2.0789 (0.1543)	2.0122 (0.1379)	2.0567 (0.1473)
		60	2.0251 (0.0883)	1.9375 (0.0840)	2.0469 (0.0919)	2.0032 (0.0858)	2.0295 (0.0864)
2	0.50	20	2.1173 (0.1758)	1.8256 (0.1423)	2.1916 (0.2134)	2.0432 (0.1501)	2.2899 (0.8171)
		40	2.0763 (0.1016)	1.8991 (0.0861)	2.1211 (0.1159)	2.0316 (0.0915)	2.1147 (0.1825)
		60	2.0483 (0.0732)	1.9220 (0.0660)	2.0801 (0.0802)	2.0166 (0.0683)	2.0686 (0.1067)
	0.80	20	2.0652 (0.1715)	1.8600 (0.1515)	2.1164 (0.1906)	2.0138 (0.1581)	2.1304 (0.3664)
		40	2.0367 (0.1051)	1.9227 (0.0973)	2.0651 (0.1113)	2.0082 (0.1006)	2.0567 (0.1473)
		60	2.0214 (0.0718)	1.9427 (0.0689)	2.0410 (0.0745)	2.0017 (0.0699)	2.0295 (0.0864)
3	0.50	20	2.1729 (0.2250)	1.8508 (0.1517)	2.2544 (0.2793)	2.0915 (0.1852)	2.2899 (0.8171)
		40	2.1125 (0.1193)	1.9258 (0.0892)	2.1594 (0.1384)	2.0655 (0.1048)	2.1147 (0.1825)
		60	2.0763 (0.0832)	1.9446 (0.0680)	2.1093 (0.0926)	2.0432 (0.0759)	2.0686 (0.1067)
	0.80	20	2.0816 (0.1932)	1.8696 (0.1627)	2.1344 (0.2158)	2.0286 (0.1765)	2.1304 (0.3664)
		40	2.0426 (0.1069)	1.9295 (0.0973)	2.0707 (0.1135)	2.0143 (0.1020)	2.0567 (0.1473)
		60	2.0282 (0.0744)	1.9492 (0.0701)	2.0479 (0.0774)	2.0084 (0.0721)	2.0295 (0.0864)

Table 2: Average values of the different estimators of λ and the corresponding MSEs (in parenthesis).

Prior	**p**	**n**	$\hat{\lambda}_{SE}$	$\hat{\lambda}_{Q}$	$\hat{\lambda}_{AP}$	$\hat{\lambda}_{SLE}$	$\hat{\lambda}$
1	0.50	20	1.5862 (0.0857)	1.4997 (0.0702)	1.6072 (0.0918)	1.5648 (0.0804)	1.5967 (0.0903)
		40	1.5373 (0.0332)	1.4962 (0.0303)	1.5471 (0.0345)	1.5269 (0.0322)	1.5433 (0.0342)
		60	1.5176 (0.0216)	1.4909 (0.0207)	1.5242 (0.0220)	1.5109 (0.0212)	1.5244 (0.0218)
	0.80	20	1.5864 (0.0910)	1.4932 (0.0744)	1.6091 (0.0978)	1.5633 (0.0852)	1.5878 (0.0929)
		40	1.5376 (0.0365)	1.4933 (0.0332)	1.5486 (0.0379)	1.5266 (0.0353)	1.5396 (0.0363)
		60	1.5252 (0.0249)	1.4961 (0.0234)	1.5324 (0.0256)	1.5179 (0.0244)	1.5270 (0.0247)
2	0.50	20	1.5594 (0.0599)	1.4844 (0.0518)	1.5778 (0.0637)	1.5409 (0.0568)	1.5967 (0.0903)
		40	1.5304 (0.0285)	1.4921 (0.0264)	1.5399 (0.0295)	1.5209 (0.0277)	1.5433 (0.0342)
		60	1.5178 (0.0198)	1.4922 (0.0190)	1.5242 (0.0203)	1.5114 (0.0195)	1.5244 (0.0218)
	0.80	20	1.5663 (0.0685)	1.4823 (0.0578)	1.5868 (0.0733)	1.5455 (0.0645)	1.5878 (0.0929)
		40	1.5331 (0.0325)	1.4904 (0.0299)	1.5435 (0.0337)	1.5226 (0.0315)	1.5396 (0.0363)
		60	1.5231 (0.0232)	1.4949 (0.0219)	1.5301 (0.0237)	1.5160 (0.0227)	1.5270 (0.0247)
3	0.50	20	1.551 (0.0520)	1.4849 (0.0451)	1.5723 (0.0553)	1.5377 (0.0494)	1.5967 (0.0903)
		40	1.5306 (0.0267)	1.4937 (0.0247)	1.5397 (0.0277)	1.5215 (0.0260)	1.5433 (0.0342)
		60	1.5173 (0.0190)	1.4922 (0.0182)	1.5235 (0.0194)	1.5110 (0.0187)	1.5244 (0.0218)
	0.80	20	1.5595 (0.0572)	1.4823 (0.0487)	1.5784 (0.0611)	1.5403 (0.0539)	1.5878 (0.0929)
		40	1.5303 (0.0301)	1.4905 (0.0278)	1.5402 (0.0312)	1.5204 (0.0292)	1.5396 (0.0363)
		60	1.5224 (0.0220)	1.4949 (0.0507)	1.5292 (0.0225)	1.5155 (0.0215)	1.5270 (0.0247)

Table 3: Average values of the different estimators of β and the corresponding MSEs (in parenthesis).

Prior	**p**	**n**	$\hat{\beta}_{SE}$	$\hat{\beta}_{Q}$	$\hat{\beta}_{AP}$	$\hat{\beta}_{SLE}$	$\hat{\beta}$
1	0.50	20	1.1136 (0.0130)	0.7274 (0.0744)	1.2395 (0.0576)	1.0018 (0.0001)	1.0000 (0.0000)
		40	1.0540 (0.0030)	0.8585 (0.0202)	1.1102 (0.0121)	1.0011 (0.0003)	1.0000 (0.0000)
		60	1.0352 (0.0013)	0.9040 (0.0092)	1.0709 (0.0051)	1.0007 (0.0000)	1.0000 (0.0000)
	0.80	20	0.2917 (0.0203)	0.1363 (0.0232)	0.3359 (0.0291)	0.2495 (0.0160)	0.2500 (0.0000)
		40	0.2721 (0.0079)	0.1936 (0.0090)	0.2928 (0.0098)	0.2518 (0.0070)	0.2500 (0.0000)
		60	0.2664 (0.0049)	0.2141 (0.0052)	0.2800 (0.0058)	0.2530 (0.0045)	0.2500 (0.0000)
2	0.50	20	1.0600 (0.0080)	0.8093 (0.0391)	1.1282 (0.0213)	0.9945 (0.0041)	1.0000 (0.0000)
		40	1.0308 (0.0020)	0.8854 (0.0143)	1.0794 (0.0071)	0.9984 (0.0012)	1.0000 (0.0000)
		60	1.0271 (0.0011)	0.9172 (0.0072)	1.0561 (0.0032)	0.9994 (0.0000)	1.0000 (0.0000)
	0.80	20	0.2718 (0.0087)	0.1556 (0.0151)	0.3013 (0.0114)	0.2423 (0.0078)	0.2500 (0.0000)
		40	0.2655 (0.0053)	0.1993 (0.0069)	0.2825 (0.0064)	0.2486 (0.0049)	0.2500 (0.0000)
		60	0.2632 (0.0037)	0.2167 (0.0042)	0.2751 (0.0043)	0.2514 (0.0035)	0.2500 (0.0000)
3	0.50	20	1.0091 (0.0032)	0.7582 (0.0602)	1.0831 (0.0103)	0.9413 (0.0064)	1.0000 (0.0000)
		40	0.9993 (0.0013)	0.8471 (0.0243)	1.0403 (0.0023)	0.9583 (0.0023)	1.0000 (0.0000)
		60	1.0010 (0.0002)	0.8913 (0.0124)	1.0311 (0.0013)	0.9732 (0.0013)	1.0000 (0.0000)
	0.80	20	0.2620 (0.0062)	0.1653 (0.0117)	0.2878 (0.0080)	0.2367 (0.0058)	0.2500 (0.0000)
		40	0.2590 (0.0044)	0.2003 (0.0061)	0.2744 (0.0051)	0.2439 (0.0041)	0.2500 (0.0000)
		60	0.2586 (0.0032)	0.2154 (0.0039)	0.2698 (0.0036)	0.2476 (0.0030)	0.2500 (0.0000)

5 Data Analysis

To illustrate the proposed methods we analyze a real data set from Fleming (1991). The data belongs to Group IV of the Primary Biliary Cirrhosis (PBC) liver study conducted by Mayo Clinic. The event of interest is the time to death of PBC Patients. The data on the survival times (in days) of 36 patients who had the highest category of bilirubin are: $400, 77, 859, 71, 1037, 1427, 733, 334, 41, 51, 549, 1170, 890, 1413, 853, 216, 1882^+, 1067^+, 131, 223, 1827, 2540, 1297, 264, 797, 930, 1329^+, 264, 1350, 1191, 130, 943, 974, 790, 1765^+, 1320^+$. The observations with + indicate censored times. For computational ease, each data value is divided by 1000. Since we do not have any prior information about the unknown parameters, we use noninformative priors with all the hyperparameters equal to zero, that is $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$, for Bayes estimates. We compute the MLEs and the Bayes estimates of θ, λ and β under different loss functions.

Table 4: The estimates of unknown parameters under different loss functions and the corresponding p-values of the Kolomogorov-Smirnov test.

Method	θ	λ	β	**K-S D**	**p-value**
Bayes(SE)	1.4860	1.5070	0.1670	0.1262	0.7251
Bayes (Q)	1.3902	1.4483	0.0947	0.1594	0.3939
Bayes (AP)	1.5093	1.5212	0.1858	0.1287	0.6962
Bayes (SLE)	1.4617	1.4920	0.1482	0.1286	0.6973
SG	1.4617	1.4483	0.1613	0.1238	0.7533
ML	1.4904	1.5108	0.1613	0.1273	0.7123

Based on the simulation study we suggest the Bayes estimator of θ under SLE loss function, the Bayes estimator of λ under quadratic loss function and the ML estimator of β, and call these estimators as simulation guided (SG) estimators. To test the goodness of fit of the model to this data, we compute the Kolomogorov-Smirnov D statistics in each case. The results are reported in Table 4. Based on the Kolomogorov-Smirnov test, we can say that all the methods fit the data quite well. However, the SG estimates perform slightly better than the rest.

6 Conclusion

In this paper, we consider the Bayesian estimation in Burr type XII distribution under the Koziol-Green model of random censorship. We assume two different set of informative priors for the unknown parameters since no continuous conjugate priors exist. We use squared error, quadratic, precautionary and squared log-error loss functions to obtain the Bayes estimates. It is seen that the closed-form expressions for the Bayes estimators are not possible we suggest Gibbs sampling scheme to obtain the approximate Bayes estimates. To observe the behavior of the Bayes estimators and to compare them with the ML estimators some simulation study is carried out. It is observed that as the sample size increases the MSEs of the estimators decrease. This rate of decrease in MSEs is relatively higher for informative priors as compared to noninformative priors. It is further observed that the Bayes estimator of θ under SLE loss function performs best, the Bayes estimator of λ under quadratic loss function performs best and the ML estimator of β performs best in all the cases. A real data analysis is performed to illustrate the proposed methods and to confirm the simulation results. The performance of these different methods is judged by the Kolomogorov-Smirnov test. It is seen that all the methods fit the data quite well. However, the simulation guided best estimators fit the data best than the rest.

Acknowledgment

The authors would like to thank the anonymous referees and the corresponding editor for their constructive comments and suggestions which had greatly improved the earlier version of this manuscript.

References

Al-Hussaini, E.K., Jaheen, Z.F. (1992). Bayesian estimation of parameters, reliability and failure rate functions of the Burr type XII failure model. *J. Statist. Comput. Simulation*, 41, 31-40.

Al-Hussaini, E.K., Hussein, M. (2011). Estimation using censored data from exponentiated Burr type XII population. *Amer. Open J. Statist.*, 1, 33-45.

Arnold, B.C., Press, S.J. (1983). Bayesian inference for Pareto populations. *J. Econom.*, 21, 287-306.

Burr, I.W. (1942). Cumulative frequency functions. *Ann. Math. Statist.*, 13, 215-232.

Devroye, L. (1984). A simple Algorithm for generating random variates with a log-concave density. *Computing*, 33, 247-257.

Fleming, T.R., Harrington, D.P. (1991). Counting Processes and Survival Analysis. *New York: John Wiley.*

Friesl, M., Hurt, J. (2007). On Bayesian estimation in an exponential distribution under random censorship. *Kybernetika*, 43, 45-60.

Geman, S., Geman, A. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-740.

Gupta, R.D., Kundu, D. (2006). Comparison of the Fisher information matrices of the Weibul and generalized exponential distributions. *J. Statist. Plann. Inference*, 136, 3130-3144.

Koziol, J.A., Green, S.B. (1976). A Cramer-von Mises statistic for random censored data. *Biometrika*, 63, 465-474.

Kundu, D. (2008), Bayesian inference and life testing plan for Weibull distribution in presence of progressive censoring. *Technometrics*, 50, 144-154.

Liang, T. (2004). Empirical Bayes estimation with random right censoring. *Internat. J. Inform. Manag. Sci.*, 15, 1-12.

Li, X., Shi, Y., Wei, J., Chai, J. (2007). Empirical Bayes estimators of reliability performances using LINEX loss under progressively type II censored samples. *Math. Comp. Simul.*, 73, 320-326.

Lindley, D.V. (1980). Approximate Bayes methods. *Trabajos de Estadistica*, 31, 223-237.

Nigm, A.M. (1988). Prediction bounds for the Burr model. *Comm. Statist. A*, 17, 287-297.

Norstrom, J.G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Trans. Reliab.*, 45, 400-403.

Upadhyay, S.K., Javed, I.A., Peshwani, M. (2004). Bayesian analysis of generalized four-pamater Burr distribution via Gibbs sampler. *Metron-Internat. J. Statist.*, LXII, 115-135.

Wahed, A.S. (2006). Bayesian inference using Burr model under asymmetric loss function: An application to Carcinoma survival data. *J. Statist. Research*, 40, 45-47.

Zellner, A. (1986). A Bayesian estimation and prediction using asymmetric loss function. *J. Amer. Statist. Assoc.*, 81, 446-451.