ISSN 1684-8403
Journal of Statistics
Volume 24, 2017. pp. 92-111

Parametric Survival Analysis Assuming the Proportional Hazard Functions of Survival and Censoring Times Distributions

Muhammad Yameen Danish[1] and Irshad Ahmad Arshad[2]

Abstract

The article deals with frequentist and Bayesian Survival analysis in Proportional Hazards model of random censorship using Burr type XII distribution. The Joint Conjugate Prior distribution of the model parameters does not exists while computing the Bayes estimates; we suggest pairwise independent gamma priors for the shape and scale parameters. The closed-form expressions for the Bayes estimates are not possible; we consider two different methods of Bayesian computation, namely, importance sampling and Lindley's approximation to obtain the Bayes estimates. The Maximum Likelihood estimation is presented in a novel way. Monte Carlo simulation study is carried out to observe the behavior of the Maximum Likelihood estimators and Bayes estimators for different combinations of the quantities involved. One real data analysis is performed for illustration.

Keywords

Random censorship, Maximum likelihood estimation, Bayesian estimation, Importance sampling, Lindley's approximation.

1. Introduction

Reliability and life testing experiments are usually expensive and time consuming. Several censoring mechanisms are used in order to reduce the experimental cost and time. The most popular among these are the right censoring schemes because of their crucial importance in life testing experiments.

[1] Department of Statistics, Government College Fateh Jang. Attock, Pakistan.
Email: yameendanish42sb@yahoo.com
[2] Department of Statistics, Allama Iqbal Open University, Islamabad, Pakistan.
Email: irshadahmad@aiou.edu.pk

The unified feature of Type I and Type II right censoring schemes is that the exact observation window is known for each unit in the sample. The third type of right censoring is random censoring in which censoring time is not fixed but depends upon other random factors which are modeled by an independent random variable. Consider the situation where the patients with cancer enter simultaneously into the study and we want to observe their lifetimes but censoring occurs in the following forms: loss to follow-up (e.g. the patient may decide to move elsewhere), drop out (e.g. due to bad side effects or refusal to participate), death from other diseases or termination of the study. Clearly, these random factors are beyond the control of an investigator and are modeled by a censoring time variable.

The Proportional Hazards (PH) model has been studied by several authors including Koziol and Green (1976), Csorgo and Horvath (1983), Hollander and Pena (1989), and Csorgo and Faraway (1998). It was Cox (1972) who first introduced the PH model to add covariates in Regression type models. This idea can also be used to add an additional shape parameter to base distribution. It is observed that the addition of new parameter(s) to the distribution make it richer and more flexible. Marshall and Olkin (1997) added a positive parameter to a general Survival function. Al-Hussaini and Ghitany (2005) added two parameters (r and p) to a Survival function by considering a countable mixture of positive integer powers of general Survival functions where the mixing proportions are the Pascal (r, p). Similarly, Al-Hussaini and Gharib (2009) obtained a new family of distributions as a countable mixture with Poisson added parameter.

A number of lifetime distributions such as Exponential, Generalized Exponential, Gamma, Weibull, Lognormal, etc., are commonly used for modeling failure time data. In this article, we consider Burr Type XII distribution introduced by Burr (1942). The motivation behind this distribution is that much of the region covered by Gamma and Lognormal distributions in skewness-kurtosis plane is also covered by the Burr Type XII distribution (Rodriguez, 1977). It is somewhat surprising to observe that a limited attention has been paid for modeling the randomly censored lifetimes when the lifetimes are not Exponential. The reason may be that the analysis becomes too difficult and may not be tractable. However, the density function of Burr distribution can take different curve shapes and provides a variety of curve shapes. For a nice account of some randomly censored lifetime distributions, see Danish and Aslam (2015, 2014, and 2013).

Suppose *n* identical units are put in life testing experiment and their lifetimes are recorded. Let, $X_1,...,X_n$ be their actual survival times which are independent and identically distributed (i.i.d.) random variables with distribution function $F_X(t)$ and density function $f_X(t)$. Further, suppose that $T_1,...,T_n$ are their censoring times which are also i.i.d. random variables with distribution function $F_T(t)$ and density function $f_T(t)$. Since, only one of the X_i's and T_i's is actually observed and it is not known in advance which one first. Let, the actual observed time be $Y_i = \min(X_i, T_i)$ with indicator variable $W_i = \mathrm{I}(X_i \le T_i)$. The indicator variable indicates whether the observation is censored or non-censored for $1 \le i \le n$. Now, it is simple to show that the observed Y_i constitute a random sample from $F_Y(t)$, where, $1 - F_Y(t) = (1 - F_X(t))(1 - F_T(t))$. This is the usual model of random censorship studied by several authors including Kaplan and Meier (1958), Efron (1967), Breslow and Crowley (1974) in nonparametric context. Under this model, Kaplan and Meier introduced their historic product limit estimator of the Survival function given by $S(y) = \prod_{i:Y_i \le y} \left(\frac{n - R_i}{n - R_i + 1} \right)^{w_i}$, where R_i is the rank of *ith* unit in the observed sample. Under PH model, the variables X and T are connected as $1 - F_T(y) = \{1 - F_X(y)\}^{\delta}$ for some positive constant δ.

It is simple to show that the joint density function of Y and W is,

$$f_{Y,\mathrm{W}}(y, \mathrm{w}) = f_X(y)\{1 - F_X(y)\}^{\delta} \delta^{1-w}; \quad y \ge 0, \ w = 0,1. \tag{1.1}$$

The density and distribution functions of the Burr Type XII distribution are,

$$f_X(x; \alpha, \beta) = \alpha\beta x^{\beta-1}\left(1 + x^{\beta}\right)^{-\alpha-1}; \ \ x > 0, \ \alpha, \beta > 0, \tag{1.2}$$

$$F_X(x; \alpha, \beta) = 1 - \left(1 + x^{\beta}\right)^{-\alpha}, \tag{1.3}$$

Using eq. (1.2) and eq. (1.3) in expression (1.1), we have

$$f_{Y,\mathrm{W}}(y, \mathrm{w}; \alpha, \beta, \delta) = \alpha\beta y^{\beta-1}\left(1 + y^{\beta}\right)^{-\alpha(1+\delta)-1} \delta^{1-w}; \quad y > 0, \mathrm{w} = 0,1. \tag{1.4}$$

The marginal distribution of Y can be obtained from eq. (1.4) as,

$$f_Y(y; \alpha, \beta, \delta) = \alpha\beta(1 + \delta) y^{\beta-1}\left(1 + y^{\beta}\right)^{-\alpha(1+\delta)-1}; \quad y, \alpha, \beta, \delta > 0.$$

The rest of the paper is organized as follows.

In the next section, we provide the Maximum Likelihood (ML) estimation. Section 4 consists of Bayesian estimation using importance sampling and Lindley's approximation. A simulation study is carried out in Section 5. A real data analysis is performed in Section 6 and finally, we conclude the paper in Section 7.

2. Maximum Likelihood estimation

In this section, we derive the ML estimators $\hat{\alpha},\hat{\beta}$ and $\hat{\delta}$ of the parameters α,β and δ assuming the model defined in eq. (1.4) holds. For an observed random sample $(y_1, \mathrm{w}_1),...,(y_n, \mathrm{w}_n) = (y, \mathrm{w})$ of size *n* from (4), the Likelihood function is,

$$l(\alpha,\beta,\delta; y, w) = \alpha^n \beta^n \prod_{i=1}^{n} y_i^{\beta-1} \prod_{i=1}^{n} \left(1+y_i^{\beta}\right)^{-\alpha(1+\delta)-1} \delta^{n-\sum_{i=1}^{n} w_i}. \tag{2.1}$$

The Log-Likelihood function can be written as,

$$L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w}) = \ln\left(l(\boldsymbol{\theta}; y, \mathrm{w})\right) = n\ln\alpha + n\ln\beta + (n - S_1)\ln\delta$$
$$+(\beta-1)S_2 - \left(1+\alpha(1+\delta)\right)\sum_{i=1}^{n} \ln\left(1+y_i^{\beta}\right), \tag{2.2}$$

where,

$$\boldsymbol{\theta} = (\alpha,\beta,\delta),\ \mathrm{S}_1 = \sum_{i=1}^{n} w_i \ \text{ and } \ S_2 = \sum_{i=1}^{n} \ln y_i .$$

The Likelihood equations are obtained by differentiating eq. (2.1) as,

$$\frac{n}{\alpha} - (1+\delta)\sum_{i=1}^{n} \ln\left(1+y_i^{\beta}\right) = 0,\ \ \frac{n-S_1}{\delta} - \alpha\sum_{i=1}^{n} \ln\left(1+y_i^{\beta}\right) = 0,$$

$$\frac{n}{\beta} - \alpha(1+\delta)\sum_{i=1}^{n} \frac{y_i^{\beta} \ln y_i}{1+y_i^{\beta}} + \sum_{i=1}^{n} \frac{\ln y_i}{1+y_i^{\beta}} = 0.$$

Solving these equations simultaneously, we have

$$\hat{\delta} = \frac{(n-S_1)}{S_1}, \qquad \hat{\alpha}(\beta) = \frac{S_1}{\sum_{i=1}^{n} \ln\left(1+y_i^{\beta}\right)}, \tag{2.3}$$

$$\sum_{i=1}^{n} \frac{y_i^{\beta} \ln y_i}{1+y_i^{\beta}} \left(\frac{1}{\beta} + \frac{1}{n}\sum_{i=1}^{n} \frac{\ln y_i}{1+y_i^{\beta}} \right) = \sum_{i=1}^{n} \ln\left(1+y_i^{\beta}\right). \tag{2.4}$$

Any suitable iterative procedure can be used to solve eq. (2.4) for β. Once the ML estimate of β is obtained from eq. (2.4), the ML estimate of α can be obtained from eq. (2.3). We need the following results for further development.

The elements of observed Fisher information matrix $\Theta(\boldsymbol{\theta})$ are obtained as,

$$\Theta_{11}(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\alpha^2} = -\frac{n}{\alpha^2} \ ,$$

$$\Theta_{12}(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\alpha\partial\beta} = -(1+\delta)\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}} \ ,$$

$$\Theta_{13}(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\alpha\,\partial\delta} = -\sum_{i=1}^{n}\ln\left(1+y_i^{\beta}\right)$$

$$\Theta_{22}(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\beta^2} = -\frac{n}{\beta^2} - \left(1+\alpha(1+\delta)\right)\sum_{i=1}^{n}\frac{y_i^{\beta}\left(\ln y_i\right)^2}{\left(1+y_i^{\beta}\right)^2},$$

$$\Theta_{23}(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\beta\,\partial\delta} = -\alpha\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}},\ \Theta_{33}(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\delta^2} = -\frac{1}{\delta^2}\left(n-S_1\right).$$

The elements of expected Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$ are derived as,

$$\mathrm{I}_{11}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\alpha^2}\right) = \frac{n}{\alpha^2},$$

$$\mathrm{I}_{12}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\alpha\partial\beta}\right) = \frac{n(1+\delta)}{\beta(1+\gamma)}k,$$

$$\mathrm{I}_{13}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\alpha\partial\delta}\right) = \frac{n}{\gamma},$$

$$\mathrm{I}_{33}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\delta^2}\right) = \frac{n}{\delta(1+\delta)},$$

$$\mathrm{I}_{22}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 L(\boldsymbol{\theta}; \mathrm{y}, \mathrm{w})}{\partial\beta^2}\right) = \frac{n}{\beta^2} + \frac{n\gamma\left(\left(k-\frac{1}{\gamma}\right)^2 + \Psi'(2) + \Psi'(1+\gamma)\right)}{\lambda^2(2+\gamma)},$$

$$I_{23}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 L}{\partial\beta\partial\delta}\right) = \frac{n\alpha}{\beta(1+\gamma)}k.$$

Where,

$\gamma = \alpha(1+\delta)$, $k = \Psi(2) - \Psi(\gamma)$, $\Psi(.)$ is Digamma function and $\Psi'(.)$ is its derivative.

Theorem 1: The matrix $\mathbf{I}(\boldsymbol{\theta}) = \left[I_{rs}(\boldsymbol{\theta})\right]$ is positive-definite and its determinant $\left|\mathbf{I}(\boldsymbol{\theta})\right|$ is finite for all $\boldsymbol{\theta} = (0 < \theta, \lambda, \beta < \infty)$ and $r, s = 1, 2, 3.$

Proof: The determinant of the expected information matrix $\mathbf{I}(\boldsymbol{\theta})$ is,

$$\left|\mathbf{I}(\boldsymbol{\theta})\right| = \frac{n^3}{\alpha^2\beta^2}\left\{\frac{1}{\delta(1+\delta)} + \frac{\gamma\left(k-\gamma^{-1}\right)^2 + \Psi'(2) + \Psi'(\gamma+1)}{\delta(1+\delta)(\gamma+2)} - \frac{\alpha^2 k^2}{(\gamma+1)^2}\right\}$$

$$-\frac{n^3 k^2}{\delta\beta^2(\gamma+1)^2} + \frac{n^3}{\gamma\beta^2}\left\{\frac{(1+\delta)\alpha k^2}{(\gamma+1)^2} - \frac{1}{\gamma} - \frac{\left(k-\gamma^{-1}\right)^2 + \Psi'(2) + \Psi'(\gamma+1)}{\gamma+2}\right\},$$

Careful simplification gives

$$\left|\mathbf{I}(\boldsymbol{\theta})\right| = \frac{n^3}{\delta\gamma\beta^2(\gamma+2)}\left[\frac{\left\{\gamma k - (\gamma+1)^2\right\}^2}{\gamma^2(\gamma+1)^2} + \Psi'(2) + \Psi'(\gamma+1)\right] > 0.$$

Thus the matrix $\mathbf{I}(\boldsymbol{\theta})$ is positive-definite and its determinant $\left|\mathbf{I}(\boldsymbol{\theta})\right|$ is finite for all $\boldsymbol{\theta} = (0 < \alpha, \beta, \delta < \infty)$.

It follows from Theorem 1 that there exists at least one solution of the Likelihood equations which is consistent estimate of the true parameter vector $\boldsymbol{\theta}$, see Chanda (1954) for more detail.

Lemma 1: For $y_i \geq 0$, suppose $g(\theta) = \sum_{i=1}^{n} y_i^{\theta}$, then $g(\theta)g''(\theta) - \left\{g'(\theta)\right\}^2 \geq 0.$

Proof: Since $g'(\theta) = \sum_{i=1}^{n} y_i^{\theta} \ln y_i$ and $g''(\theta) = \sum_{i=1}^{n} y_i^{\theta} (\ln y_i)^2$,

therefore, $g(\theta)g''(\theta) - \left\{g'(\theta)\right\}^2 = \left(\sum_{i=1}^{n} y_i^{\theta}\right)\left(\sum_{i=1}^{n} y_i^{\theta} (\ln y_i)^2\right) - \left(\sum_{i=1}^{n} y_i^{\theta} \ln y_i\right)^2$

$$= \sum_{1\le i\le j}^{n} y_i^{\theta} y_j^{\theta} \left(\ln y_i - \ln y_j \right)^2 \ge 0.$$

Theorem 2: The observed information matrix $\Theta(\boldsymbol{\theta}) = \left[\Theta_{rs}(\boldsymbol{\theta})\right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ for $r, s = 1, 2, 3$, where $\hat{\boldsymbol{\theta}}$ is the consistent root of the Likelihood equations, is negative-definite with probability tending to unity.

Proof: The determinant $\left|\Theta(\boldsymbol{\theta})\right|$ of the observed information matrix is,

$$\left|\Theta(\boldsymbol{\theta})\right| = -\frac{n}{\alpha^2}\left[\frac{(n-S_1)}{\delta^2}\left\{\frac{n}{\beta^2} + (1+\gamma)\sum_{i=1}^{n}\frac{y_i^{\beta}(\ln y_i)^2}{\left(1+y_i^{\beta}\right)^2}\right\} - \alpha^2\left(\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}}\right)^2\right] +$$

$$(1+\delta)\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}}\left[\frac{(1+\delta)(n-S_1)}{\delta^2}\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}} - \alpha\sum_{i=1}^{n}\ln\left(1+y_i^{\beta}\right)\left(\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}}\right)\right]$$

$$-\sum_{i=1}^{n}\ln\left(1+y_i^{\beta}\right)\left[\gamma\left(\sum_{i=1}^{n}\frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}}\right)^2 - \sum_{i=1}^{n}\ln\left(1+y_i^{\beta}\right)\left\{\frac{n}{\beta^2} + (1+\gamma)\sum_{i=1}^{n}\frac{y_i^{\beta}(\ln y_i)^2}{\left(1+y_i^{\beta}\right)^2}\right\}\right].$$

Substituting ML estimates and simplifying, we have

$$\left|\Theta(\boldsymbol{\theta})\right| = -\frac{n}{\hat{\beta}^2\hat{\delta}^2}\sum_{i=1}^{n}\ln\left(1+y_i^{\hat{\beta}}\right)$$

$$-\frac{n}{\hat{\delta}}\left\{\left(\sum_{i=1}^{n}\ln\left(1+y_i^{\hat{\beta}}\right)\right)\left(\sum_{i=1}^{n}\frac{y_i^{\hat{\beta}}(\ln y_i)^2}{\left(1+y_i^{\hat{\beta}}\right)^2}\right) - \left(\sum_{i=1}^{n}\frac{y_i^{\hat{\beta}}\ln y_i}{1+y_i^{\hat{\beta}}}\right)^2\right\}.$$

Since,

$$\frac{y_i^{\hat{\beta}}}{1+y_i^{\hat{\beta}}} < \ln\left(1+y_i^{\hat{\beta}}\right) \text{ for all } y_i^{\hat{\beta}} > 0,$$

it follows from Lemma 1 that the 2nd term on the right is positive and so $\left|\Theta(\boldsymbol{\theta})\right| < 0$. Thus, the matrix $\Theta(\boldsymbol{\theta})$ is negative-definite.

Therefore, the Likelihood function has a relative maximum at the consistent roots of the Likelihood equations. Now the asymptotic normality result of the ML estimators can be stated as follows.

The vector $\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)$ has asymptotically Multivariate Normal distribution with mean vector **0** and the variance-covariance matrix $\mathbf{V}=\mathbf{I}^{-1}\left(\boldsymbol{\theta}\right)$.

3. Bayesian estimation

For the Bayesian estimation of unknown parameters, one needs prior distributions of these parameters. We assume the following independent Gamma priors of α,β and δ

$$\left.\begin{array}{ll}\pi_1\left(\alpha\right)=\dfrac{b_1^{a_1}}{\Gamma\left(a_1\right)}\alpha^{a_1-1}e^{-b_1\alpha}; & a_1,b_1,\alpha>0\\ \pi_2\left(\beta\right)=\dfrac{b_2^{a_2}}{\Gamma\left(a_2\right)}\beta^{a_2-1}e^{-b_2\beta}; & a_2,b_2,\beta>0\\ \pi_3\left(\delta\right)=\dfrac{b_3^{a_3}}{\Gamma\left(a_3\right)}\delta^{a_3-1}e^{-b_3\delta}; & a_3,b_3,\delta>0\end{array}\right\}. \tag{3.1}$$

It may be noted that as the hyper-parameters in the Gamma density approach zero, it becomes inversely proportional to its argument. This density is often used as non-informative Gamma prior for the parameters in the range 0 to ∞. The joint prior density of the unknown parameters can be written as,

$$\pi\left(\alpha,\beta,\delta\right)\propto\alpha^{a_1-1}e^{-b_1\alpha}\beta^{a_2-1}e^{-b_2\beta}\delta^{a_3-1}e^{-b_3\delta}. \tag{3.2}$$

The posterior distribution is obtained by combining the Likelihood function in eq. (2.1) and joint prior in eq. (3.2) as,

$$\pi\left(\alpha,\beta,\delta\,|\,y,\mathrm{w}\right)\propto\alpha^{a_1+n-1}e^{-\alpha\left(b_1+\sum_{i=1}^{n}\ln\left(1+y_i^{\beta}\right)\right)}\beta^{a_2+n-1}e^{-\beta\left(b_2-S_2\right)}\prod_{i}^{n}\frac{1}{1+y_i^{\beta}}\times$$

$$\delta^{a_3+n-S_1-1}e^{-\delta\left(b_3+\alpha\sum_{i=1}^{n}\ln\left(1+y_i^{\beta}\right)\right)}=\pi_C\left(\alpha,\beta,\delta\,|\,y,\mathrm{w}\right). \tag{3.3}$$

The Bayes estimators involve the posterior expectation of a parameter or a function of parameters. In general, the posterior expectation of any function of parameters, say $U\left(\alpha,\beta,\delta\right)$, can be written as,

$$E\left(U\left(\alpha,\beta,\delta|y,\mathrm{w}\right)\right)=\frac{\int_0^\infty\int_0^\infty\int_0^\infty U\left(\alpha,\beta,\delta\right)\pi_C\left(\alpha,\beta,\delta|y,\mathrm{w}\right)}{\int_0^\infty\int_0^\infty\int_0^\infty \pi_C\left(\alpha,\beta,\delta|y,\mathrm{w}\right)}. \tag{3.4}$$

However, it is not possible to evaluate eq. (3.4) in closed-form. We use two different methods, namely, importance sampling and Lindley's approximation to evaluate it.

3.1 Importance sampling: Monte Carlo importance sampling is the most commonly used method of computing posterior expectations and provides reliable accuracy of computation. Here, we use it to obtain the Bayes estimates of α,β and δ. The posterior distribution eq. (3.3) can be written as,

$$\pi\left(\alpha,\beta,\delta|y,w\right)\propto g_\alpha\left(a_1+\mathrm{n},b_1+\sum_{i=1}^n\ln\left(1+y_i^\beta\right)\right)g_\beta\left(a_2+\mathrm{n},b_2-S_2\right)\times$$
$$g_\delta\left(a_3+\mathrm{n}-S_1,b_3+\alpha\sum_{i=1}^n\ln\left(1+y_i^\beta\right)\right)\xi\left(\alpha,\beta\right), \tag{3.5}$$

where,

g_α, g_β and g_δ are gamma densities and

$$\xi\left(\alpha,\beta\right)=\frac{\prod_i^n\left(1+y_i^\beta\right)^{-1}}{\left(b_1+\sum_{i=1}^n\ln\left(1+y_i^\beta\right)\right)^{n+a_1}\left(b_3+\alpha\sum_{i=1}^n\ln\left(1+y_i^\beta\right)\right)^{n-S_1+a_3}}.$$

Now, we suggest the following procedure to obtain the posterior samples and in turn to obtain the Bayes estimates and corresponding highest posterior density (HPD) credible intervals:

Step 1: Generate $\beta_1\sim\mathrm{gamma}\left(a_2+\mathrm{n},b_2-S_2\right)$.

Step 2: Generate $\alpha_1|\beta_1\sim\mathrm{gamma}\left(a_1+\mathrm{n},b_1+\sum_{i=1}^n\ln\left(1+y_i^{\beta_1}\right)\right)$.

Step 3: Generate $\delta_1|\left(\beta_1,\alpha_1\right)\sim\mathrm{gamma}\left(a_3+\mathrm{n}-S_1,b_3+\alpha_1\sum_{i=1}^n\ln\left(1+y_i^{\beta_1}\right)\right)$.

Step 4: Repeat Steps 1, 2 and 3 M times to obtain $\left(\alpha_1,\beta_1,\delta_1\right),\ldots,\left(\alpha_M,\beta_M,\delta_M\right)$.

The Bayes estimates of α, β and δ under Squared Error (SE) Loss function can be obtained from

$$\hat{\alpha}_{BI} = \frac{\sum_{j=1}^{M} \alpha_j \xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)}, \quad \hat{\beta}_{BI} = \frac{\sum_{j=1}^{M} \beta_j \xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)} \quad , \hat{\delta}_{BI} = \frac{\sum_{j=1}^{M} \delta_j \xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)}.$$

The HPD credible intervals can be constructed following the procedure described in Kundu and Pradhan (2009) as follows:

Let $\upsilon_j = \frac{\xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)}; \; j = 1, \ldots, M.$

Arrange the pairs $(\alpha_1, \upsilon_1), \ldots, (\alpha_M, \upsilon_M)$ as $(\alpha_{(1)}, \upsilon_{(1)}), \ldots, (\alpha_{(M)}, \upsilon_{(M)})$,

where $\alpha_{(1)} < \ldots < \alpha_{(M)}$.

The Bayes estimate of ϑ_p is $\hat{\vartheta}_p = \alpha_{(M_p)}$,

where M_p is the integer satisfying $\sum_{j=1}^{M_p} \upsilon_{(i)} \le p < \sum_{j=1}^{M_p+1} \upsilon_{(i)}$.

Now construct all the $100(1-\lambda)\%$ credible intervals for α as $(\hat{\vartheta}_\gamma, \hat{\vartheta}_{\gamma+1-\lambda})$, for

$\gamma = \upsilon_{(1)}, \upsilon_{(1)} + \upsilon_{(2)}, \ldots, \sum_{j=1}^{M_{1-\lambda}} \upsilon_{(j)}$.

The HPD credible interval for α is the interval that has the shortest length.

The Bayes estimates of α, β and δ under Linear Exponential (LE) Loss function with Loss function parameter *c* can be obtained from

$$\hat{\alpha}_{BI} = -\frac{1}{c} \ln \left(\frac{\sum_{j=1}^{M} e^{-c\alpha_j} \xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)} \right), \quad \hat{\beta}_{BI} = -\frac{1}{c} \ln \left(\frac{\sum_{j=1}^{M} e^{-c\beta_j} \xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)} \right),$$

$$\hat{\delta}_{BI} = -\frac{1}{c} \ln \left(\frac{\sum_{j=1}^{M} e^{-c\delta_j} \xi(\alpha_j, \beta_j)}{\sum_{j=1}^{M} \xi(\alpha_j, \beta_j)} \right).$$

The Bayes estimates of α, β and δ under General Entropy (GE) Loss function with Loss function parameter q can be obtained from

$$\hat{\alpha}_{BI} = \left[\frac{\sum_{j=1}^{M} \alpha_j^{-q} \xi\left(\alpha_j, \beta_j\right)}{\sum_{j=1}^{M} \xi\left(\alpha_j, \beta_j\right)} \right]^{\frac{-1}{q}}, \ \hat{\beta}_{BI} = \left[\frac{\sum_{j=1}^{M} \beta_j^{-q} \xi\left(\alpha_j, \beta_j\right)}{\sum_{j=1}^{M} \xi\left(\alpha_j, \beta_j\right)} \right]^{\frac{-1}{q}},$$

$$\hat{\delta}_{BI} = \left[\frac{\sum_{j=1}^{M} \delta_j^{-q} \xi\left(\alpha_j, \beta_j\right)}{\sum_{j=1}^{M} \xi\left(\alpha_j, \beta_j\right)} \right]^{\frac{-1}{q}}.$$

3.2 Lindley's approximation: Lindley (1980) proposed a procedure to approximate the ratio of two integrals such as eq. (3.4). The Linley's approximation plays an important role in Bayesian analysis. It can be used quite effectively to obtain the Bayes estimates that are more accurate than the usual normal approximation and not computationally as intensive as numerical methods. The procedure is explained in appendix.

The Bayes estimates of α, β and δ under SE Loss function using the Lindley's approximation are,

$$\hat{\alpha}_{BL} = \hat{\alpha} + \rho_1\sigma_{11} + \rho_2\sigma_{21} + \rho_3\sigma_{31} + \frac{1}{2}\left(A_1\sigma_{11} + A_2\sigma_{21} + A_3\sigma_{31}\right), \tag{3.6}$$

$$\hat{\beta}_{BL} = \hat{\beta} + \rho_1\sigma_{12} + \rho_2\sigma_{22} + \rho_3\sigma_{32} + \frac{1}{2}\left(A_1\sigma_{12} + A_2\sigma_{22} + A_3\sigma_{32}\right), \tag{3.7}$$

$$\hat{\delta}_{BL} = \hat{\delta} + \rho_1\sigma_{13} + \rho_2\sigma_{23} + \rho_3\sigma_{33} + \frac{1}{2}\left(A_1\sigma_{13} + A_2\sigma_{23} + A_3\sigma_{33}\right). \tag{3.8}$$

Similar expressions for the Bayes estimates of α, β and δ under LE and GE loss functions can be obtained using the Lindley's approximation. The closed-form expressions for $\rho_1,\ \rho_2,\ \rho_3,\ A_1,\ A_2$ and A_3 are provided in Appendix.

4. Simulation

In this section, we perform a Monte Carlo simulation to observe the behavior of the ML estimators and Bayes estimators for different sample sizes, different priors, different Loss functions and for different proportions of non-censored observations. We consider different sample sizes: n = 20, 40, 60; different proportions of non-censored observations: p = 0.50, 0.80; different values of loss function parameters c and q: –0.5, –0.9, –0.3, 0.3, 0.9, 1.5; different sets of parameter values and different combinations of hyper-parameters as given in Table 1. It may be noted that NGP represents the non-informative gamma priors when all the hyper-parameters in eq. (3.2) are zero and IGP represents the informative Gamma priors with prior means equal to the corresponding parameter values. For a particular case, 1000 randomly censored samples are generated from the model in eq. (1.4) and for each sample we compute the ML estimates and the corresponding 95% confidence intervals based on observed information matrix, the Bayes estimates and corresponding 95% credible intervals based on 1000 importance samples. The average values of the ML estimates, Bayes estimates and MSEs are reported in Tables 2 and 3. Some of the foregoing points are very clear from these results. It is observed that as the sample size increases, the biases and MSEs decrease reasonably. However, it is seen that the rate of decrease in biases and MSEs is higher for small to medium sample sizes as compared with medium to large sample sizes. It is further observed that the Bayes estimators under NGP based on importance sampling perform slightly better than the ML estimators for small sample sizes and their performance is very similar for large sample sizes. The Bayes estimators under NGP based on the Lindley's approximation perform relatively better than both the ML estimators and the Bayes estimators under NGP based on importance sampling. A similar behavior is observed under the informative priors. However, the Bayes estimators under IGP perform quite better than both the ML estimators and the Bayes estimators under NGP. When comparing the Bayes estimators under different loss functions, it is seen that to estimate α and δ under LE loss function with minimum MSEs, the appropriate range should be $0.3 < c < 1.5$. Similarly to estimate α under GE loss function, the corresponding loss function parameter range should be $0.3 < q < 1.5$. The range of loss function parameter for the estimation of shape β parameter should be round about 1.5 under LE and GE loss functions. The appropriate range of GE loss function parameter for the Bayes estimator of δ is $0.3 < q < 1.5$ in case of 50% non-censoring rate and it changed to $-0.9 < q < 0.3$ in case of 80% non-censoring rate. In their appropriate ranges, the Bayes estimators under LE loss function perform slightly better than the Bayes estimators under GE loss function and both perform slightly better than the Bayes estimators under SE loss function.

5. Data analysis

To illustrate the proposed methods we analyze a real data set from Fleming and Harrington (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 non-informative Gamma priors with all hyper-parameters in (10) equal to zero, that is $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$ for Bayes estimates. We compute the ML estimates and Bayes estimates of parameters under different Loss functions. The results are reported in Table 4. To test the Goodness-of-Fit of the model to the data at hand, we compute *p*-values of the Kolomogorov-Smirnov test. Based on this test we can say that all the methods fit the data quite well with slightly better fit for the Bayes estimates under GE loss function based on importance sampling. Figure 1 shows the survival function of Burr type XII distribution fitted to the K-M survival curve of the data using different methods of estimation. The fitted survival functions provide a closed, but smoothed, summary of the K-M survival curve.

6. Conclusion

In this paper, we consider the survival analysis in Proportional Hazards model of random censorship using Burr Type XII distribution. We use independent Gamma priors for the unknown model parameters for Bayes estimates. The Bayes estimates under different Loss functions are obtained using importance sampling and Lindley's approximation. A simulation study is performed to observe the behavior of the Maximum Likelihood and Bayes estimators. It is observed that the Bayes estimators under non-informative Gamma priors based importance sampling perform slightly better than the Maximum Likelihood estimators for small sample sizes and their performance is very similar for large sample sizes. The Bayes estimators under non-informative Gamma priors based on the Lindley's approximation perform relatively better than both the Maximum Likelihood estimators and the Bayes estimators under non-informative priors based on importance sampling. Similarly, the Bayes estimators under informative

Gamma priors perform quite better than both the Maximum Likelihood estimators and the non-informative Bayes estimators. The Bayes estimators under Linear Exponential Loss function perform slightly better than the Bayes estimators under General Entropy Loss function and both perform relatively better than the Bayes estimators under Squared Error Loss function. However, the ML estimate of censoring parameter outperforms than the rest. A real data analysis is performed to illustrate the proposed methodology. The Goodness-of-Fit of the model is checked by the Kolomogorov-Smirnov test of fit. It is observed that all the methods fit the data well with slightly better results for the Bayes estimates under GE Loss function based on importance sampling.

References

1. Al-Hussaini, E. K. and Gharib, M. (2009). A new family of distributions as a countable mixture with Poison added parameter. *Journal of Statistical Theory and Applications*, **8**, 169-190.
2. Al-Hussaini, E. K. and Gitany, M. (2005). On certain countable mixtures of absolutely continuous distributions. *Metron*, **LXIII (1)**, 39-53. 3.
4. Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Annuals of Statistics*, **2**, 437-453.
5. Burr, I. W. (1942). Cumulative frequency functions. *Annuals of Mathematical Statistics*, **13**, 215-232.
6. Chanda, K. C. (1954). A note on the consistency and maxima of the roots of Likelihood equations. *Biometrika*, **46**, 56-61.
7. Cox, D. (1972). Regression models and life tables. *Journal of the Royal Statistical Society, Series-B*, **34**, 187-200.
8. Csorgo, S. and Faraway, J. (1998). The paradoxical nature of the Proportional Hazards model. *Statistics*, **31**, 61-78.
9. Csorgo, S. and Horváth, L. (1983). On the Koziol-Green model for random censorship II. *Studia Scientiarum Mathematicarum Hungarica*, **18**, 195-203.
10. Danish, M. Y., & Aslam, M. (2013). Bayesian estimation for randomly censored generalized exponential distribution under asymmetric loss functions. *Journal of Applied Statistics*, **40** (5), 1106-1119.
11. Danish, M. Y. and Aslam, M. (2014). Bayesian inference for the randomly censored Weibull distribution. *Journal of Statistical Computation and Simulation*, **84(1)**, 215-230.

12. Danish, M. Y. and Aslam, M. (2015). *Bayesian Analysis of some Randomly Censored Lifetime Distribution, Ph.D. thesis*. Faculty of Science, Allama Iqbal Open University, Islamabad, Pakistan.
13. Effron, B. (1967). The two-sample problem with censored data. *Proceeding of 5th Berkeley Symposium,* **4**, 831-853.
14. Fleming, T. R. and Harrington, D. P. (1991). *Counting processes and survival analysis*. Wiley, New York.
15. Hollander, M. and Peña, E. (1989). Families of confidence bands for the survival function under the general random censorship model and the Koziol-Green model. *The Canadian Journal of Statistics*, **17**, 59-74.16.
17. Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association*, **53**, 457-481.
18. Koziol, J. A. and Green, S. B. (1976). A Cramer-von Mises statistic for random censored data. *Biometrika*, **63**, 465-474.
19. Kundu, D. and Pradhan, B. (2009). Bayesian inference and life testing plans for generalized exponential distribution. *Science in China, Series-A,* **52**, 1373-1388.
20. Lindley, D. V. (1980). Approximate Bayes methods. *Trabajos de Estadistica*, **31**, 223-237.
21. Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families. *Biometrika*, **84**, 641-652.
22. Rodriguez, R. N. (1977). A guide to the Burr type XII distributions. *Biometrika*, **64**, 129-134.

Appendix

The posterior expectation eq. (3.4), using the notations $(\alpha,\beta,\delta)=(\theta_1,\theta_2,\theta_3)$ and $\rho(\theta_1,\theta_2,\theta_3)=\ln\pi(\theta_1,\theta_2,\theta_3)$, can be written as,

$$E\left(U(\theta_1,\theta_2,\theta_3)\middle|y,d\right)=\frac{\int_{(\theta_1,\theta_2,\theta_3)} U(\theta_1,\theta_2,\theta_3)e^{L(\theta_1,\theta_2,\theta_3)+\rho(\theta_1,\theta_2,\theta_3)}d(\theta_1,\theta_2,\theta_3)}{\int_{(\theta_1,\theta_2,\theta_3)} e^{L(\theta_1,\theta_2,\theta_3)+\rho(\theta_1,\theta_2,\theta_3)}d(\theta_1,\theta_2,\theta_3)}. \qquad (A1)$$

For large *n*, the expression in (*A1*) is evaluated by the Lindley's method as,

$$\hat{U}_B(\theta_1,\theta_2,\theta_3)=U(\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3)+(U_1d_1+U_2d_2+U_3d_3+d_4+d_5)$$
$$+\tfrac{1}{2}[A_1(U_1\sigma_{11}+U_2\sigma_{12}+U_3\sigma_{13})+A_2(U_1\sigma_{21}+U_2\sigma_{22}+U_3\sigma_{23})$$
$$+A_3(U_1\sigma_{31}+U_2\sigma_{32}+U_3\sigma_{33})], \qquad (A2)$$

where,

$$d_i=\rho_1\sigma_{i1}+\rho_2\sigma_{i2}+\rho_3\sigma_{i3}, \qquad U_i=\frac{\partial U(\theta_1,\theta_2,\theta_3)}{\partial\theta_i}, \qquad \rho_i=\frac{\partial\rho(\theta_1,\theta_2,\theta_3)}{\partial\theta_i},$$

$$L_{ijk}=\frac{\partial^3 L(\theta_1,\theta_2,\theta_3)}{\partial\theta_i\partial\theta_j\partial\theta_k}, \quad i,j,k=1,2,3, \qquad d_4=U_{12}\sigma_{12}+U_{13}\sigma_{13}+U_{23}\sigma_{23},$$

$$d_5=\frac{1}{2}(U_{11}\sigma_{11}+U_{22}\sigma_{22}+U_{33}\sigma_{33}),$$
$$A_1=\sigma_{11}L_{111}+2\sigma_{12}L_{121}+2\sigma_{13}L_{131}+2\sigma_{23}L_{231}+\sigma_{22}L_{221}+\sigma_{33}L_{331},$$
$$A_2=\sigma_{11}L_{112}+2\sigma_{12}L_{122}+2\sigma_{13}L_{132}+2\sigma_{23}L_{232}+\sigma_{22}L_{222}+\sigma_{33}L_{332},$$
$$A_3=\sigma_{11}L_{113}+2\sigma_{12}L_{123}+2\sigma_{13}L_{133}+2\sigma_{23}L_{233}+\sigma_{22}L_{223}+\sigma_{33}L_{333}.$$

Moreover, σ_{ij} is ijth element of minus the inverse of observed information matrix and all the quantities are evaluated at the ML estimates $(\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3)$.

In our case,

$$\rho_1=\frac{a_1-1}{\hat{\alpha}}-b_1, \qquad \rho_2=\frac{a_2-1}{\hat{\beta}}-b_2, \qquad \rho_3=\frac{a_3-1}{\hat{\delta}}-b_3,$$

$$A_1=\frac{2n\sigma_{11}}{\hat{\alpha}^3}-2\sigma_{23}\sum_{i=1}^{n}\eta_i'-(1+\hat{\delta})\sigma_{22}\sum_{i=1}^{n}\eta_i'',$$

$$A_2=\frac{2n\sigma_{22}}{\hat{\beta}^3}-2\sigma_{13}\sum_{i=1}^{n}\eta_i'-2\left((1+\hat{\delta})\sigma_{12}+\hat{\alpha}\sigma_{23}\right)\sum_{i=1}^{n}\eta_i''-\left(1+\hat{\alpha}(1+\hat{\delta})\right)\sigma_{22}\sum_{i=1}^{n}\eta_i''',$$

$$A_3=\frac{2\sigma_{33}}{\hat{\delta}^3}(n-S_1)-2\sigma_{12}\sum_{i=1}^{n}\eta_i'-\alpha\sigma_{22}\sum_{i=1}^{n}\eta_i'',$$

where,

$$\eta_i = \ln\left(1+y_i^{\beta}\right), \eta_i' = \frac{y_i^{\beta}\ln y_i}{1+y_i^{\beta}}, \eta_i'' = \frac{y_i^{\beta}\left(\ln y_i\right)^2}{\left(1+y_i^{\beta}\right)^2}, \eta_i''' = \frac{y_i^{\lambda}\left(1+y_i^{\beta}\right)\left(\ln y_i\right)^3}{\left(1+y_i^{\beta}\right)^3}.$$

To derive expression eq. (3.6) for the Bayes estimate of α, take $U\left(\alpha,\beta,\delta\right)=\alpha$ in (*A2*) so $U_1=1$ and all other U-terms in (*A2*) are zero. Now expression eq. (3.6) follows from these substitutions in (*A2*). The eq. (3.7) and eq. (3.8) can be obtained similarly.

Table 1: The values of parameters and hyper-parameters used in simulation

Notation	α	β	δ	a_1	b_1	a_2	b_2	a_3	b_3
NGP	2	1.5	1	0	0	0	0	0	0
IGP	2	1.5	1	4	2	3	2	4	2
NGP	2	1.5	0.25	0	0	0	0	0	0
IGP	2	1.5	0.25	4	2	3	2	2	4

Table 2: The average values of ML estimators and Bayes estimators under NGP and the corresponding MSEs (in parenthesis) when (*a*) p = 0.50 and (*b*) p = 0.80

(*a*)

$U(\theta)$	***n***	**ML**	**Bayes Estimates (Importance Sampling)**					**Lindley**
		Estimates	SE	LE (0.3)	LE (0.9)	GE (0.3)	GE (0.9)	SE
α	20	2.2899 (0.8171)	2.2451 (0.6132)	2.1445 (0.4412)	1.9855 (0.2805)	2.0698 (0.4369)	1.9895 (0.3842)	2.8603 (0.2248)
	40	2.1147 (0.1825)	2.1098 (0.1758)	2.0701 (0.1535)	1.9975 (0.1247)	2.0310 (0.1483)	1.9947 (0.1402)	2.0324 (0.1112)
	60	2.0686 (0.1067)	2.0603 (0.1015)	2.0357 (0.0937)	1.9892 (0.0829)	2.0097 (0.0917)	1.9863 (0.0890)	1.9958 (0.0776)
β	20	1.5967 (0.0903)	1.5862 (0.0857)	1.5758 (0.0820)	1.5556 (0.0754)	1.5584 (0.0790)	1.5454 (0.0765)	1.5360 (0.0700)
	40	1.5433 (0.0342)	1.5371 (0.0332)	1.5324 (0.0325)	1.5230 (0.0313)	1.5239 (0.0319)	1.15177 (0.0314)	1.5138 (0.0302)
	60	1.5244 (0.0218)	1.5170 (0.0216)	1.5145 (0.0213)	1.5085 (0.0208)	1.5089 (0.0211)	1.5049 (0.0210)	1.5025 (0.0205)
δ	20	1.0000 (0.0000)	1.1136 (0.0130)	1.0728 (0.0054)	1.0070 (0.0001)	0.9706 (0.0009)	0.9108 (0.0080)	0.9546 (0.0021)
	40	1.0000 (0.0000)	1.0540 (0.0030)	1.0366 (0.0014)	1.0050 (0.0001)	0.9860 (0.0002)	0.9561 (0.0020)	0.9769 (0.0006)

	60	1.0000 (0.0000)	1.0352 (0.0013)	1.0242 (0.0006)	1.0035 (0.0000)	0.9906 (0.0001)	0.9707 (0.0009)	0.9844 (0.0003)

(*b*)

U(θ)	***n***	**ML**	**Bayes Estimates (Importance Sampling)**					**Lindley**
		Estimates	SE	LE (0.3)	LE (0.9)	GE (0.3)	GE (0.9)	SE
α	20	2.1304 (0.3664)	2.1139 (0.3633)	2.0658 (0.3164)	1.9796 (0.2559)	2.0642 (0.3347)	2.0216 (0.3147)	2.1038 (0.2244)
	40	2.0567 (0.1473)	2.0456 (0.1449)	2.0246 (0.1369)	1.9843 (0.1250)	2.0223 (0.1398)	2.0022 (0.1363)	1.9863 (0.1177)
	60	2.0295 (0.0864)	2.0251 (0.0883)	2.0116 (0.0854)	1.9853 (0.0810)	2.0098 (0.0864)	1.9966 (0.0852)	1.9600 (0.0783)
β	20	1.5878 (0.0929)	1.5864 (0.0910)	1.5752 (0.0868)	1.5534 (0.0794)	1.5704 (0.0869)	1.5564 (0.0837)	1.5323 (0.0735)
	40	1.5396 (0.0363)	1.5376 (0.0365)	1.5325 (0.0356)	1.5223 (0.0342)	1.5299 (0.0356)	1.5233 (0.0350)	1.5124 (0.0330)
	60	2.0295 (0.0864)	1.5252 (0.0249)	2.0116 (0.0854)	1.9853 (0.0810)	2.0098 (0.0864)	1.9966 (0.0852)	1.9680 (0.0783)
δ	20	0.2712 (0.0138)	0.2917 (0.0203)	0.2872 (0.0191)	0.2789 (0.0170)	0.2620 (0.0169)	0.2371 (0.0155)	0.2715 (0.0134)
	40	0.2618 (0.0059)	0.2721 (0.0079)	0.2703 (0.0077)	0.2668 (0.0073)	0.2578 (0.0072)	0.2457 (0.0069)	0.2634 (0.0052)
	60	0.2608 (0.0037)	0.2664 (0.0049)	0.2653 (0.0048)	0.2630 (0.0046)	0.2570 (0.0046)	0.2490 (0.0044)	0.2549 (0.0025)

Table 3: The average values of Bayes estimators under IGP and the corresponding MSEs (in parenthesis) when (*a*) p = 0.50 and (*b*) p = 0.80

(*a*)

U(θ)	***n***	**Bayes Estimates (Importance Sampling)**					**Lindley**
		SE	LE (0.3)	LE (0.9)	GE (0.3)	GE (0.9)	SE
α	20	2.1173 (0.1758)	2.1696 (0.2106)	2.0689 (0.1503)	2.0211 (0.1447)	1.9769 (0.1370)	2.9820 (0.1203)
	40	2.0763 (0.1016)	2.1060 (0.1137)	2.0479 (0.0920)	2.0182 (0.0893)	1.9915 (0.0860)	2.0947 (0.0795)
	60	2.0483 (0.0732)	2.0688 (0.0791)	2.0285 (0.0685)	2.0070 (0.0672)	1.9880 (0.0656)	1.9906 (0.0620)
β	20	1.5594 (0.0599)	1.5683 (0.0623)	1.5507 (0.0577)	1.5353 (0.0560)	1.5240 (0.0546)	1.5334 (0.0539)
	40	1.5304 (0.0285)	1.5349 (0.0291)	1.5260 (0.0279)	1.5180 (0.0275)	1.5123 (0.0271)	1.5173 (0.0270)

	60	1.5178 (0.0198)	1.5207 (0.0201)	1.5149 (0.0196)	1.5095 (0.0194)	1.5057 (0.0193)	1.5091 (0.0192)
δ	20	1.0600 (0.0080)	1.0833 (0.0113)	1.0383 (0.0052)	0.9744 (0.0041)	0.9359 (0.073)	0.9992 (0.0032)
	40	1.0380 (0.0020)	1.0512 (0.0034)	1.0252 (0.0014)	0.9864 (0.0009)	0.9632 (0.0020)	1.0015 (0.0007)
	60	1.0270 (0.0010)	1.0365 (0.0016)	1.0183 (0.0006)	0.9904 (0.0004)	0.9738 (0.0009)	1.0013 (0.0003)

(*b*)

U(θ)	***n***	**Bayes Estimates (Importance Sampling)**					**Lindley**
		SE	LE (0.3)	LE (0.9)	GE (0.3)	GE (0.9)	SE
α	20	2.0652 (0.1715)	2.0325 (0.1568)	1.9713 (0.1372)	2.0293 (0.1615)	1.9984 (0.1551)	2.9154 (0.1279)
	40	2.0367 (0.1051)	2.0189 (0.1002)	1.9847 (0.0929)	2.0168 (0.1018)	1.9996 (0.0996)	2.0522 (0.0886)
	60	2.0214 (0.0718)	2.0093 (0.0696)	1.9858 (0.0665)	2.0076 (0.0704)	1.9958 (0.0694)	1.9630 (0.0646)
β	20	1.5663 (0.0685)	1.5564 (0.0656)	1.5369 (0.0607)	1.5518 (0.0656)	1.5392 (0.0635)	1.5181 (0.0601)
	40	1.5331 (0.0325)	1.5282 (0.0318)	1.5186 (0.0306)	1.5256 (0.0318)	1.5194 (0.0313)	1.5091 (0.0296)
	60	1.5231 (0.0232)	1.5198 (0.0228)	1.5134 (0.0223)	1.5182 (0.0228)	1.5139 (0.0226)	1.5071 (0.0218)
δ	20	0.2718 (0.0087)	0.2692 (0.0084)	0.2642 (0.0078)	0.2511 (0.0078)	0.2334 (0.0078)	0.2595 (0.0074)
	40	0.2655 (0.0053)	0.2641 (0.0052)	0.2613 (0.0050)	0.2537 (0.0049)	0.2436 (0.0048)	0.2586 (0.0048)
	60	0.2632 (0.0037)	0.2622 (0.0037)	0.2603 (0.0036)	0.2549 (0.0035)	0.2479 (0.0034)	0.2585 (0.0036)

Table 4: The ML estimates and Bayes estimates and the associated *p*-values of Kolomogorov-Smirnov test.

Method	LF	α	β	δ	*p-value*
ML	–	1.4904	1.5108	0.1613	0.7123
Importance	SE	1.4860	1.5070	0.1670	0.7251
Sampling	LE	1.4344	1.4745	0.1621	0.7274
	GE	1.4978	1.5141	0.1764	0.7403
Lindley	SE	1.4867	1.5112	0.1665	0.7146
	LE	1.5390	1.5438	0.1710	0.6848

GE	1.4988	1.5185	0.1755	0.7286

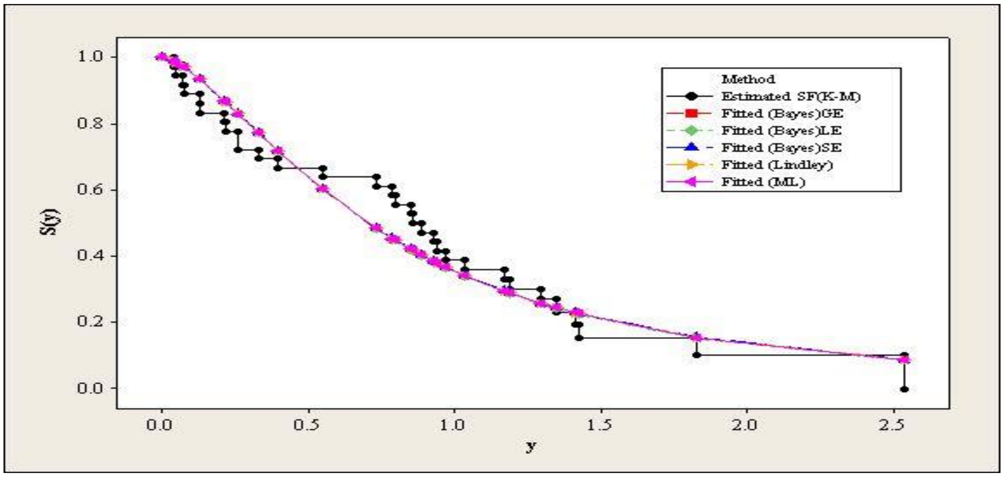

Figure 1: The Burr XII survival function fitted to K-M survival curve using different methods