Preference of Prior for Bayesian Analysis of the Mixed Burr Type X Distribution under Type I Censored Samples

Tabassum Naz Sindhu
Department of Statistics
Quaid-i-Azam University 45320, Islamabad 44000, Pakistan
sindhuqau@gmail.com

Navid Feroze
Department of Mathematics and Statistics
Allama Iqbal Open University, Islamabad, Pakistan
navidferoz@hotmail.com

Muhammad Aslam
Department of Statistics
Quaid-i-Azam University 45320, Islamabad 44000, Pakistan
aslamsdqu@yahoo.com

Abstract

The paper is concerned with the preference of prior for the Bayesian analysis of the shape parameter of the mixture of Burr type X distribution using the censored data. We modeled the heterogeneous population using two components mixture of the Burr type X distribution. A comprehensive simulation scheme, through probabilistic mixing, has been followed to highlight the properties and behavior of the estimates in terms of sample size, corresponding risks and the proportion of the component of the mixture. The Bayes estimators of the parameters have been evaluated under the assumption of informative and non-informative priors using symmetric and asymmetric loss functions. The model selection criterion for the preference of the prior has been introduced. The hazard rate function of the mixture distribution has been discussed. The Bayes estimates under exponential prior and precautionary loss function exhibit the minimum posterior risks with some exceptions.

Keywords: Inverse Transformation Method, Loss Functions, Prior Predictive distributions, Credible Intervals.

1. Introduction

Burr (1942) introduced twelve different forms of cumulative distribution functions for modeling lifetime data. Among those twelve distribution functions, Burr-Type X and Burr-Type XII have received the maximum attention. Surles and Padgett (2001) observed that the Burr-Type X distribution can be used quite effectively in modeling strength data and general lifetime data. Several aspects of the one-parameter (Scale = 1) Burr-Type X distribution have been studied by Sartawi and Abu-Salih (1991), Jaheen (1996), Ahmad et al. (1997) and Raqab (1998). The distribution function and the density function of a Burr-Type X distribution have closed form. As a consequence of that, it can be used very conveniently even for censored data.

Mixture models play a vital role in many applications. The direct applications of finite mixture models are in the fields of physics, biology, geology, medicine, engineering and economics, and many others. Detailed applications and examples are given by Mclachlan and Peel (2000), Mcculloch and Searle (2001), Ismail and El Khodary (2001), Demidenko (2004), Sultan et al. (2007), Nair and Abdul (2010), Afify (2011) and Erisoglu et al. (2011). Gosh and Ebrahimi (2001) have made the Bayesian analysis of the mixing function in a mixture of two exponential distributions. Saleem and Aslam (2008) presented a comparison of the Maximum Likelihood (ML) estimates with the Bayes estimates assuming the uniform and the Jeffreys priors for the parameters of the Rayleigh mixture. Saleem et al. (2010) considered the Bayesian analysis of the mixture of Power function distribution using the complete and the censored sample.

The problem of censoring is more commonly encountered in life-time data because no experiment may remain sustained for an infinite time due to restrictions on the available time or cost for testing. There are different kinds of censoring schemes which include right, left and interval censoring, single or multiple censoring and type-I or type-II censoring. Type-I and type-II censoring schemes are most popular among them. Saleem et al. (2010) considered the Bayesian analysis of the power function mixture distribution using type-I censored data. Shi and Yan (2010) discussed the empirical Bayes estimates of two-parameter exponential distribution under type-I censoring.

We have focused on the selection of suitable prior for Bayesian analysis of the mixture of Burr type X distribution under type I censored samples. It worth mentioning here, that the mixture of this model under Bayesian approach has not been considered in the literature yet.

The article is outlined as follows. In the section 2, we defined the mixture model for Burr type X distribution, sampling and likelihood function for type I censored samples. In the section 3, the posterior distributions have been derived under different priors. The loss functions for the derivation of Bayes estimators and posterior risks have been introduced in the section 4. Method of elicitation of hyper-parameters for the mixture of Burr type X distribution via prior predictive approach has been discussed in the section 5. Credible intervals for the parameters of the model have been derived in the section 6. The posterior predictive distributions have been presented in the section 7. A simulation study along with graphical representation of the results has been performed in the section 8. A real life example has been included in the section 9. The section 10 contains the discussion regarding the hazard rate function of the mixture model. The model selection criterion has been introduced in the section 11. The section 12 presents the conclusion of the study.

2. The Population and the Model

A population is postulated to be composed to two subpopulations with specified parameters. The subpopulations are mixed in proportion $w,(1-w)$ where $0<w<1$. A finite mixture distribution function with the two component densities of specified parametric form (but with unknown parameters, λ_1 and λ_2) and with unknown mixing weights, w and $(1-w)$ is $F(x)=wF_1(x)+(1-w)F_2(x)$, $0<w<1$, with the two component distribution functions of specified parametric (Burr type X distribution) form

$F_1(x)=\left\{1-exp\left(-x^2\right)\right\}^{\lambda_1}$ and $F_2(x)=\left\{1-exp\left(-x^2\right)\right\}^{\lambda_2}$. The corresponding finite mixture density function has its probability density function (pdf) as:

$$p(x\,|\,\lambda_1,\lambda_2,w)=w2\lambda_1 x\exp\left(-x^2\right)\left\{1-\exp\left(-x^2\right)\right\}^{\lambda_1-1}+(1-w)2\lambda_2 x\exp\left(-x^2\right)\left\{1-\exp\left(-x^2\right)\right\}^{\lambda_2-1},$$
$$\lambda_i>0,\quad i=1,2;\quad 0<x<\infty \tag{1}$$

The graphs of the mixture model, given in (1), are presented in the following. The abbreviations used in the legends are: PR1: $\lambda_1 = 1.50$, $\lambda_2 = 0.75$; PR2: $\lambda_1 = 2.75$, $\lambda_2 = 2.00$; PR3: $\lambda_1 = 3.50$, $\lambda_2 = 3.00$; PR4: $\lambda_1 = 4.50$, $\lambda_2 = 4.00$; PR11: $\lambda_1 = 0.50$, $\lambda_2 = 0.75$; PR12: $\lambda_1 = 2.75$, $\lambda_2 = 4.50$; PR13: $\lambda_1 = 3.50$, $\lambda_2 = 5.00$; PR14: $\lambda_1 = 4.50$, $\lambda_2 = 6.00$.

From the graphs it can be seen that the mixture model shifts its origin to the right for larger values of the parameters. In addition, the larger values of the parameters decrease the spread and increase the height of the curves.

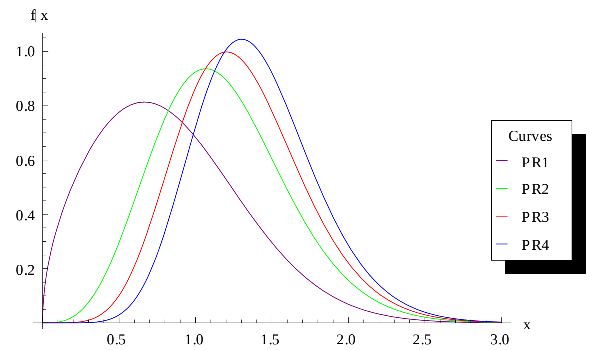

Fig. 1: Graph of the mixture model using $\pi = 0.25$

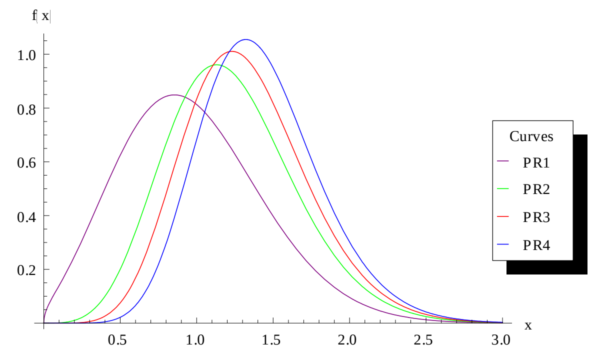

Fig. 2: Graph of the mixture model using $\pi = 0.75$

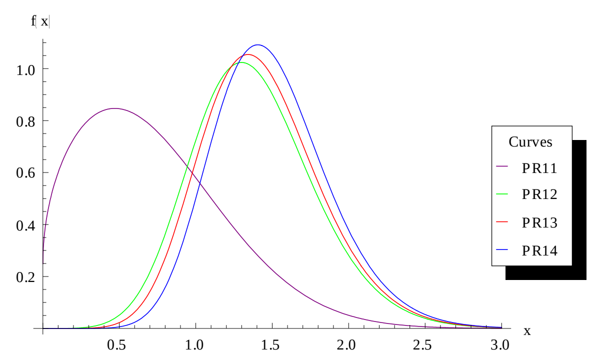

Fig. 3: Graph of the mixture model using $\pi = 0.25$

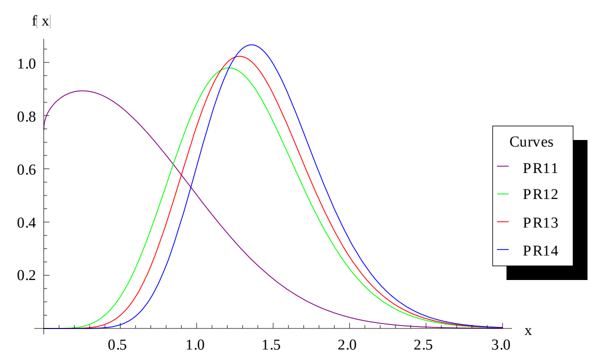

Fig. 4: Graph of the mixture model using $\pi = 0.75$

2.1. Sampling

A random sample of size n units from the above mixture model is operating to a life testing experiment. The test is terminated at a fixed time T. Let the test be conducted and it is observed that out of n, r units have lifetime in the interval $[0,T]$ and $(n-r)$ units are still functioning when the test termination time T is over. Hence $(n-r)$ units that have not failed by the time T are censored objects and yield no information. According to Mendenhall and Hader (1958), in many real life situations only the failed objects can easily be identified as member of either subpopulation 1 or subpopulation 2. So, depending upon the cause of failure it may be observed that r_1 and r_2 objects are identified as members of the first subpopulation and the second subpopulation respectively. Obviously $r=r_1+r_2$ and remaining $(n-r)$ units provide no information about the subpopulation to which they belong. Furthermore, let x_{ij} as the failure time of the j^{th} unit to the i^{th} subpopulation, where $j=1,2,...,r_i, i=1,2; 0\le x_{1j},x_{2j}\le T$.

2.2. The Maximum Likelihood Function

The likelihood function for a two-component mixture with n items under study, the probability that r_1 will fail due to cause 1, r_2 will fail due to cause 2 and remaining $(n-r_1-r_2)$ will survive at time T when test is terminated is given as:

$$L(\lambda_1,\lambda_2,w\,|\,\mathrm{x})\propto\prod_{j=1}^{r_1} wf_1\left(x_{1j}\right)\prod_{j=1}^{r_2}(1-w)\,f_2\left(x_{2j}\right)\left\{1-F\left(T\right)\right\}^{n-r}$$

$$L\left(\lambda_1,\lambda_2,w\,|\,\mathrm{x}\right)\propto\left\{\prod\nolimits_{j=1}^{r_1} w2\lambda_1 x_{1j}\exp\left(-x_{1j}{}^2\right)\left\{1-\exp\left(-x_{1j}{}^2\right)\right\}^{\lambda_1-1}\right\}\left\{\prod\nolimits_{j=1}^{r_2}(1-w)2\lambda_2 x_{2j}\exp\left(-x_{2j}{}^2\right)\left\{1-\exp\left(-x_{2j}{}^2\right)\right\}^{\lambda_2-1}\right\}$$
$$\times\left[1-\left\{w\exp\left\{-\lambda_1\ln\left(1-\exp\left[-T^2\right]\right)^{-1}\right\}+\left(1-w\right)\exp\left\{-\lambda_2\ln\left(1-\exp\left[-T^2\right]\right)^{-1}\right\}\right\}\right]^{n-r}$$

$$L\left(\lambda_1,\lambda_2,w\,|\,\mathrm{x}\right)\propto\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}\prod_{i=1}^{2}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}w^{r_1+k_1}\left(1-w\right)^{r_2+k_2}\lambda_i^{r_i}\exp\left[-\lambda_i\left\{\psi_{ij}\left(x_{ij}\right)\right\}\right]\qquad(2)$$

where $\mathbf{x}=(\mathbf{x_1},\mathbf{x_2})=(x_{11},x_{12},...,x_{1r_1},x_{21},x_{22},...,x_{2r_2})$ is data,

$$\psi_{1j}\left(x_{1j}\right)=\sum_{j=1}^{r_1}\ln\left\{1-\exp\left(-x_{1j}{}^2\right)\right\}^{-1}+k_1\ln\left(1-\exp\left[-T^2\right]\right)^{-1}$$

and $\psi_{2j}\left(x_{2j}\right)=\sum_{j=1}^{r_2}\ln\left\{1-\exp\left(-x_{2j}{}^2\right)\right\}^{-1}+k_2\ln\left(1-\exp\left[-T^2\right]\right)^{-1}$

3. Prior and Posterior Distributions

In case of an informative prior, the use of prior information is equivalent to add a number of observations to the given sample size and hence leads to a reduction of posterior risks of the Bayes estimates based on the said informative prior. Bolstad (2004) has discussed a method to evaluate the worth of a prior information in terms of the number of additional observations supposed to be added to the given sample size. While, when significant information is not avaiable regarding the parameters of the sampling distribution, a non-informative prior can be formed. We have assumed both informative and non-informative priors for the posterior estimation.

3. 1. Posterior Distribution under Uniform Prior

Let us assume a state of ignorance that λ_1, λ_2 and w are uniformly distributed over $(0,\infty)$. Hence

$$p_1(\lambda_1)=k_1,\, p_2(\lambda_2)=k_2, \lambda_i>0 \;\; p_3(w)\propto 1, \quad i=1,2,\; 0<w<1 \tag{3}$$

Assuming independence, we have an improper joint prior that is proportional to a constant. The joint prior is incorporated with the likelihood (2) to yield a proper joint posterior distribution of λ_1, λ_2 and *w.* The joint posterior distribution of λ_1, λ_2 and w is

$$p(\lambda_1,\lambda_2,w\,|\,\mathrm{x})=\frac{1}{\mathbb{R}_{1k}}\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}\prod_{i=1}^{2}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}w^{r_1+k_1}(1-w)^{r_2+k_2}\,\lambda_i^{r_i}\exp\left[-\lambda_i\left\{\psi_{ij}\left(x_{ij}\right)\right\}\right]$$

$$\lambda_i>0, \quad 0<w<1 \tag{4}$$

where $\Im_1=r_1+k_1+1,\;\; \Im_2=r_2+k_2+1$, $B\left(\Im_1,\Im_2\right)$ is a standard beta function and $\mathbb{R}_{1k}$ is defined as:

$$\mathbb{R}_{1k}=\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}B\left(\Im_1,\Im_2\right)\frac{\Gamma\left(r_1+1\right)\Gamma\left(r_2+1\right)}{\left\{\psi_{1j}\left(x_{1j}\right)\right\}^{(r_1+1)}\left\{\psi_{2j}\left(x_{2j}\right)\right\}^{(r_2+1)}}$$

3. 2. Posterior Distribution under Jeffreys Prior

Jeffreys prior is locally uniform and hence non-informative. An appealing property of Jeffreys prior is that it is invariant with respect to one-to-one transformations. For the Burr type x model given in Section 2, the Jeffreys priors are $p_1(\lambda_1)=\lambda_1^{-1}, 0<\lambda_1<\infty$, $p_2(\lambda_2)=\lambda_2^{-1}, 0<\lambda_2<\infty$ and $p_3(w)\propto 1, 0<w<1$. assuming independence, the joint prior $g(\lambda_1,\lambda_2,w)\propto\left(\lambda_1\lambda_2\right)^{-1}$ is incorporated with the likelihood (2) to yield a proper joint posterior distribution of λ_1, λ_2 and w. The joint posterior distribution under Jeffreys prior is:

$$p(\lambda_1,\lambda_2,w\,|\,\mathrm{x})=\frac{1}{\mathbb{R}_{2k}}\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}\prod_{i=1}^{2}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}w^{r_1+k_1}(1-w)^{r_2+k_2}\,\lambda_i^{r_i-1}\exp\left[-\lambda_i\left\{\psi_{ij}\left(x_{ij}\right)\right\}\right]$$

$$\lambda_i>0,\;\; 0<w<1 \tag{5}$$

where $\mathbb{R}_{2k}$ is defined as:

$$\mathbb{R}_{2k}=\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}B\left(\Im_1,\Im_2\right)\frac{\Gamma\left(r_1\right)\Gamma\left(r_2\right)}{\left\{\psi_{1j}\left(x_{1j}\right)\right\}^{(r_1)}\left\{\psi_{2j}\left(x_{2j}\right)\right\}^{(r_2)}}.$$

3.3. Posterior Distribution under Gamma Prior

In case of an informative prior, the use of prior information is equivalent to add a number of observations to the given sample size and hence leads to a reduction of posterior risks of the Bayes estimates based on the said informative prior.

Let $\lambda_1 \sim Gamma(a_1, b_1)$, $\lambda_2 \sim Gamma(a_2, b_2)$ are the gamma priors and $p_3(w) \propto 1$, $a_i, b_i > 0,\ i = 1, 2,\ 0 < w < 1$

Assuming independence, the joint prior is incorporated with the likelihood to give the joint posterior distribution, that is

$p(\lambda_1, \lambda_2, w) \propto \lambda_1^{r_1 + a_1 - 1} \lambda_2^{r_2 + a_2 - 1} \exp[-b_1 \lambda_1] \exp[-b_2 \lambda_2]$. The joint posterior distribution of λ_1, λ_2 and w is

$$p(\lambda_1, \lambda_2, w \mid \mathrm{x}) = \frac{1}{\mathbb{R}_{3k}} \sum_{k_1=0}^{n-r} \sum_{k_2=0}^{n-r-k_1} \prod_{i=1}^{2} (-1)^{k_1+k_2} \binom{n-r}{k_1} \binom{n-r-k_1}{k_2} w^{r_1+k_1} (1-w)^{r_2+k_2}\, \lambda_i^{r_i + a_i - 1} \exp\left[-\lambda_i \left\{ b_i + \psi_{ij}\left(x_{ij}\right) \right\}\right]$$

$$\lambda_i > 0,\ 0 < w < 1 \qquad (6)$$

where $\mathbb{R}_{3k}$ is defined as:

$$\mathbb{R}_{3k} = \sum_{k_1=0}^{n-r} \sum_{k_2=0}^{n-r-k_1} (-1)^{k_1+k_2} \binom{n-r}{k_1} \binom{n-r-k_1}{k_2} B\left(\Im_1, \Im_2\right) \frac{\Gamma\left(r_1 + a_1\right) \Gamma\left(r_2 + a_2\right)}{\left\{ b_1 + \psi_{1j}\left(x_{1j}\right) \right\}^{(r_1 + a_1)} \left\{ b_2 + \psi_{2j}\left(x_{2j}\right) \right\}^{(r_2 + a_2)}}$$

3.4. Posterior Distribution under Exponential Prior

Let $\lambda_1 \sim Exp(\nu_1)$, $\lambda_2 \sim Exp(\nu_2)$ are the exponential priors and $p_3(w) \propto 1$, $\nu_i > 0$, $i = 1, 2,\ 0 < w < 1$

Assuming independence, the joint prior is incorporated with the likelihood to give the joint posterior distribution, that is

$p(\lambda_1, \lambda_2, w) \propto \exp[-\nu_1 \lambda_1] \exp[-\nu_2 \lambda_2]$. The joint posterior distribution of λ_1, λ_2 and w is

$$p(\lambda_1, \lambda_2, w \mid \mathrm{x}) = \frac{1}{\mathbb{R}_{4k}} \sum_{k_1=0}^{n-r} \sum_{k_2=0}^{n-r-k_1} \prod_{i=1}^{2} (-1)^{k_1+k_2} \binom{n-r}{k_1} \binom{n-r-k_1}{k_2} w^{r_1+k_1} (1-w)^{r_2+k_2}\, \lambda_i^{r_i} \exp\left[-\lambda_i \left\{ \nu_i + \psi_{ij}\left(x_{ij}\right) \right\}\right]$$

$$\lambda_i > 0,\ 0 < w < 1 \qquad (7)$$

where $\mathbb{R}_{4k}$ is defined as:

$$\mathbb{R}_{4k} = \sum_{k_1=0}^{n-r} \sum_{k_2=0}^{n-r-k_1} (-1)^{k_1+k_2} \binom{n-r}{k_1} \binom{n-r-k_1}{k_2} B\left(\Im_1, \Im_2\right) \frac{\Gamma\left(r_1 + 1\right) \Gamma\left(r_2 + 1\right)}{\left\{ \nu_1 + \psi_{1j}\left(x_{1j}\right) \right\}^{(r_1 + 1)} \left\{ \nu_2 + \psi_{2j}\left(x_{2j}\right) \right\}^{(r_2 + 1)}}.$$

The graphs for the marginal posterior distributions for the parameters of the mixture density, given in (3), (5), (6) and (7) under different priors are presented in the following. The graphs are based on the simulated data from the mixture model using a sample of size 50. The legends in the graphs contain following abbreviations: UP: Uniform prior; JP: Jeffreys prior; GP: Improved gamma prior; EP: Exponential prior.

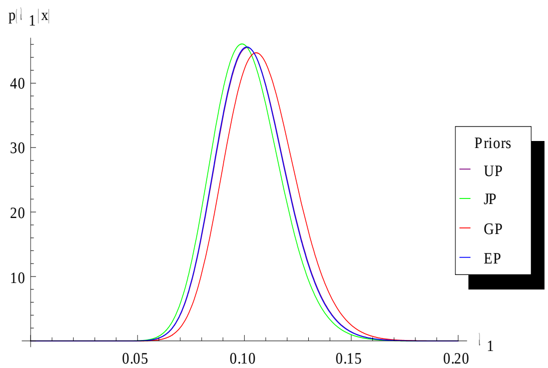

Fig. 5: Graph of the marginal posterior distribution of λ_1 under different priors

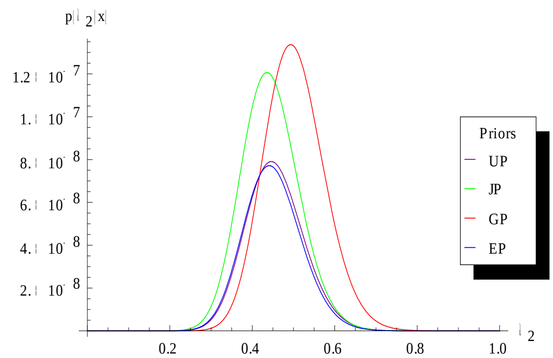

Fig. 6: Graph of the marginal posterior distribution of λ_2 under different priors

The graphs indicate that the marginal posterior distributions for the parameters λ_1 and λ_2 are slightly positively skewed. In case of first component, it can be observed that shape of posterior distribution under all the prior are close to each other with slightly different origins. In case of second component, the curves of the posterior distributions under uniform and exponential priors are similar, while the shapes of posterior distributions under Jeffreys and gamma priors are different.

4. Loss Function

The squared error loss function (SELF) is the commonly chosen loss function for the estimation of the parameter. The squared error loss function $L\left(\lambda,\lambda^*\right)=\left(\lambda-\lambda^*\right)^2$ was proposed by Legendre (1805) and Gauss (1810). This loss function is broadly used because it gives apparently sound Bayesian solution, i.e., those one would usually suggest as estimators for a non-decision theoretic inference based on the posterior distribution. A very useful and simple asymmetric precautionary loss function (PLF) is: $L(\lambda^*,\lambda)=\dfrac{(\lambda^*-\lambda)^2}{\lambda^*}$. The Bayes estimator and the posterior risk under PLF are $\lambda^*=\sqrt{E_{\lambda|\mathbf{x}}\left(\lambda^2|\mathbf{x}\right)}$ and $E_{\lambda|\mathbf{x}}(L(d,\lambda)=2\left\{\sqrt{E_{\lambda\mathbf{x}}\left(\lambda^2|\mathbf{x}\right)}-E_{\lambda|\mathbf{x}}\left(\lambda|\mathbf{x}\right)\right\}$ respectively. The Bayes estimators are also evaluated under weighted squared error loss function (WSELF). The Bayes estimator and the posterior risk under WSELF are $\lambda^*=\left\{E_{\lambda|\mathbf{x}}\left(\lambda^{-1}|\mathbf{x}\right)\right\}^{-1}$ and $E_{\lambda|\mathbf{x}}(L(\lambda^*,\lambda)=\left\{E_{\lambda|\mathbf{x}}\left(\lambda|\mathbf{x}\right)-\left\{E_{\lambda|\mathbf{x}}\left(\lambda^{-1}|\mathbf{x}\right)\right\}^{-1}\right\}$. Hence, we consider symmetric as well as asymmetric loss functions for getting better understanding in our Bayesian analysis.

5. Elicitation of Hyper-parameters of Informative Prior through Prior Predictive Probabilities

To elicit a prior density, Aslam (2003) forms some new methods base on the prior predictive distribution. For the elicitation of hyper-parameters, he considers prior predictive probabilities, predictive mode and confidence level. In this study, the method of prior predictive probabilities is used for obtaining the hyper-parameters of the considered informative prior. In fact, prior predictive removes the uncertainty in parameter (s) to reveal a distribution for the data point only. We suppose that prior predictive probabilities satisfy the laws of probability because this law ensure the expert would be consistent in eliciting the probabilities and some inconsistencies may arise which are not very serious.

A function $\xi(a,b)$ is defined in such a way that the hyper-parameters a and b are to be chosen by minimizing this $\xi(a,b)=\min\limits_{a,b}\sum\limits_{y}\left\{\frac{p(y)-p_0(y)}{p(y)}\right\}^2$, where $p(y)$ denote the prior predictive probabilities characterized by the hyper-parameters a and b and $p_0(y)$ denote the elicited prior predictive probabilities. The above equations solved simultaneously by applying 'PROC SYSLIN' of the SAS package for eliciting the required hyper-parameters.

5.1. Elicitation of hyper-parameters of Gamma Prior

The equation of prior predictive under the gamma prior is given as:

$$p(y)=\int_0^\infty\int_0^\infty\int_0^1 p(\lambda_1,\lambda_2,w)\,p(y\mid\lambda_1,\lambda_2,w)\,dw\,d\lambda_1\,d\lambda_2 \tag{8}$$

where $p(\lambda_1,\lambda_2,w)=\frac{b_1^{a_1}}{\Gamma(a_1)}\lambda_1^{a_1-1}e^{-\lambda_1 b_1}\,\frac{b_2^{a_2}}{\Gamma(a_2)}\lambda_2^{a_2-1}e^{-\lambda_2 b_2}$

$$p(y\mid\lambda_1,\lambda_2,w)=w2\lambda_1 y\exp\left(-y^2\right)\left\{1-\exp\left(-y^2\right)\right\}^{\lambda_1-1}+(1-w)2\lambda_2 y\exp\left(-y^2\right)\left\{1-\exp\left(-y^2\right)\right\}^{\lambda_2-1}$$

After some algebra

$$p(y)=y\exp\left(-y^2\right)\left\{1-\exp\left(-y^2\right)\right\}^{-1}\left[\frac{a_1 b_1^{a_1}}{\left\{b_1+\ln\left\{1-\exp\left(-y^2\right)\right\}^{-1}\right\}^{a_1+1}}+\frac{a_2 b_2^{a_2}}{\left\{b_2+\ln\left\{1-\exp\left(-y^2\right)\right\}^{-1}\right\}^{a_2+1}}\right],$$
$$0<y<\infty \tag{9}$$

By using the method of elicitation, discussed above, we get the following hyper-parametric values $a_1=2.69875$, $b_1=0.08456$, and $a_2=5.58269, b_2=0.089586$.

5.2. Elicitation of hyper-parameters of Exponential Prior

The equation of prior predictive under the exponential prior is given as:

$$p(y)=\int_0^\infty\int_0^\infty\int_0^1 p(\lambda_1,\lambda_2,w)\,p(y\,|\,\lambda_1,\lambda_2,w)\,dw\,d\lambda_1 d\lambda_2$$

where $p(\lambda_1,\lambda_2,w)=\nu_1 e^{-\lambda_1\nu_1}\nu_2 e^{-\lambda_2\nu_2}$

$$p(y\,|\,\lambda_1,\lambda_2,w)=w2\lambda_1 y\exp\left(-y^2\right)\left\{1-\exp\left(-y^2\right)\right\}^{\lambda_1-1}+(1-w)2\lambda_2 y\exp\left(-y^2\right)\left\{1-\exp\left(-y^2\right)\right\}^{\lambda_2-1}$$

After some simplifications the prior predictive distribution becomes:

$$p(y)=y\exp\left(-y^2\right)\left\{1-\exp\left(-y^2\right)\right\}^{-1}\left[\frac{\nu_1}{\left\{\nu_1+\ln\left\{1-\exp\left(-y^2\right)\right\}^{-1}\right\}^2}+\frac{\nu_2}{\left\{\nu_2+\ln\left\{1-\exp\left(-y^2\right)\right\}^{-1}\right\}^2}\right],\quad 0<y<\infty \tag{10}$$

By using the method of elicitation, mentioned above, we get the following hyper-parametric values $\nu_1=0.925806$, $\nu_2=1.023728$.

6. Credible Interval

The Bayesian counterpart of the confidence interval is named as credible interval. Unlike classical confidence interval, the 95% Bayesian credible interval contains the true parameter value with approximately 95% confidence. The credible interval is defined as: *Let* $p(\lambda|\mathrm{x})$ *be the posterior distribution; then a* $100(1-\alpha)\%$ *credible interval for parameter* λ*, in any set* C is *such that* $P_{p(\lambda|\mathrm{x})}(C)=1-\alpha$. According to Eberly and Casella (2003) the credible interval can also be defined as: $\int_0^L p(\lambda|\mathrm{x})d\lambda=\frac{\alpha}{2}$, $\int_U^\infty p(\lambda|\mathrm{x})d\lambda=\frac{\alpha}{2}$ where L and U are the lower and upper limits of the credible interval respectively and α is level of significance.

7. Posterior Predictive Distributions

The predictive distribution contains the information about the independent future random observation given preceding observations. In context of Bayesian inference the predictive distribution is referred as the posterior predictive distribution. Bolstad (2004) and Bansal (2007) have given a detailed discussion about the posterior predictive distribution. The posterior predictive distribution can be defined as:

$$g\left(y|\mathrm{x}\right)=\int_0^\infty\int_0^\infty\int_0^1 h(\lambda_1,\lambda_2,w|\mathrm{x})\,f\left(y;\lambda_1,\lambda_2,w\right)dw\,d\lambda_1 d\lambda_2 \tag{11}$$

where $h(\lambda_1, \lambda_2, w \mid \mathrm{x})$ is the posterior mixture distribution, $f\left(y; \lambda_1, \lambda_2, w\right)$ is mixture density for future observation and y = x_{n+1} is the future observation given the sample information $x = x_1, x_2, \ldots, x_n$, from of the model (1). The posterior predictive distribution using (1) and (12) can be obtained as:

The posterior predictive distribution under uniform prior is:

$$p(\lambda_1,\lambda_2,w\mid \mathrm{x})=\frac{1}{\mathbb{R}_{1k}}\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}\frac{e^{-y^2}}{1-e^{-y^2}}\left[\frac{B(\Im_1+1,\Im_2)\Gamma(r_1+2)\Gamma(r_2+1)}{\left\{\psi_{1j}(x_{1j})-\ln\left(1-e^{-y^2}\right)\right\}^{r_1+2}\left\{\psi_{2j}(x_{2j})\right\}^{r_2+1}}+\frac{B(\Im_1,\Im_2+1)\Gamma(r_1+1)\Gamma(r_2+2)}{\left\{\psi_{1j}(x_{1j})\right\}^{r_1+1}\left\{\psi_{2j}(x_{2j})-\ln\left(1-e^{-y^2}\right)\right\}^{r_2+2}}\right]$$

The posterior predictive distribution under Jeffreys prior is:

$$p(\lambda_1,\lambda_2,w\mid \mathrm{x})=\frac{1}{\mathbb{R}_{1k}}\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}\frac{e^{-y^2}}{1-e^{-y^2}}\left[\frac{B(\Im_1+1,\Im_2)\Gamma(r_1+1)\Gamma(r_2)}{\left\{\psi_{1j}(x_{1j})-\ln\left(1-e^{-y^2}\right)\right\}^{r_1+1}\left\{\psi_{2j}(x_{2j})\right\}^{r_2}}+\frac{B(\Im_1,\Im_2+1)\Gamma(r_1)\Gamma(r_2+1)}{\left\{\psi_{1j}(x_{1j})\right\}^{r_1}\left\{\psi_{2j}(x_{2j})-\ln\left(1-e^{-y^2}\right)\right\}^{r_2+1}}\right]$$

The posterior predictive distribution under gamma prior is:

$$p(\lambda_1,\lambda_2,w\mid \mathrm{x})=\frac{1}{\mathbb{R}_{1k}}\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}\frac{e^{-y^2}}{1-e^{-y^2}}\left[\frac{B(\Im_1+1,\Im_2)\Gamma(r_1+a_1+1)\Gamma(r_2+a_2)}{\left\{\psi_{1j}(x_{1j})+b_1-\ln\left(1-e^{-y^2}\right)\right\}^{r_1+a_1+1}\left\{\psi_{2j}(x_{2j})+b_2\right\}^{r_2+a_2}}+\frac{B(\Im_1,\Im_2+1)\Gamma(r_1+a_1)\Gamma(r_2+a_2+1)}{\left\{\psi_{1j}(x_{1j})+b_1\right\}^{r_1+a_1}\left\{\psi_{2j}(x_{2j})+b_2-\ln\left(1-e^{-y^2}\right)\right\}^{r_2+a_2+1}}\right]$$

The posterior predictive distribution under exponential prior is:

$$p(\lambda_1,\lambda_2,w\mid \mathrm{x})=\frac{1}{\mathbb{R}_{1k}}\sum_{k_1=0}^{n-r}\sum_{k_2=0}^{n-r-k_1}(-1)^{k_1+k_2}\binom{n-r}{k_1}\binom{n-r-k_1}{k_2}\frac{e^{-y^2}}{1-e^{-y^2}}\left[\frac{B(\Im_1+1,\Im_2)\Gamma(r_1+2)\Gamma(r_2+1)}{\left\{\psi_{1j}(x_{1j})+v_1-\ln\left(1-e^{-y^2}\right)\right\}^{r_1+2}\left\{\psi_{2j}(x_{2j})+v_2\right\}^{r_2+1}}+\frac{B(\Im_1,\Im_2+1)\Gamma(r_1+1)\Gamma(r_2+2)}{\left\{\psi_{1j}(x_{1j})+v_1\right\}^{r_1+1}\left\{\psi_{2j}(x_{2j})+v_2-\ln\left(1-e^{-y^2}\right)\right\}^{r_2+2}}\right]$$

8. Simulation Study

A simulation study is carried out in order to investigate the performance of Bayes estimators and impact of sample size and mixing proportion in the fit of model. We take random samples of sizes n = 25, 50, 100 and 300 from the two component mixture of Burr type x distribution with $(\lambda_1, \lambda_2) = \{(3,\ 7), (9,\ 11)\}$, $w = 0.35$. To generate a mixture data we make use of probabilistic mixing with probability w and $(1-w)$. A uniform number u is generated n times and if $u < w$ the observation is taken randomly from F_1 (Burr Type X distribution with parameter λ_1) otherwise from F_2 (from Burr Type X distribution with parameter λ_2). Hence the parameters to be estimated are known to be (λ_1, λ_2) and w. To implement censored sampling, all the observations greater than T are declared as censored ones while calculations are conducted. The choice of the censoring time is made in such a way that the censoring rate in the resultant sample to be approximately 20%. To avoid an extreme sample, we simulate 1000 data sets each of size n. The Bayes estimates and posterior risks (in parenthesis) are computed using Mathematica 8.0. The average of these estimates and corresponding risks are reported in Tables 1-20 and Figs 1-6. These Tables depict that the Bayes estimates with informative (Exponential) prior have smaller posterior risks, however, a few exceptions are observed. The quality of Bayes (Exponential) depends upon the quality of prior information. The hyper-parameters can be considered as outcomes of the prior information. The informative Bayes estimates may turn out to be the most efficient, provided that useful prior information and consequently, the appropriate hyper-parameters are available. The comparison observed is summarized in last section.

Table 1: Bayes Estimates (Uniform) of Burr Type X mixture parameters under SELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	3.84830 (1.70731)	7.90969 (3.94036)	0.368222 (0.008304)	11.29690 (8.549523)	12.5169 (9.86615)	0.376611 (0.008446)
50	3.36193 (0.62968)	7.44322 (1.73275)	0.364750 (0.004340)	10.17520 (5.79473)	11.69280 (4.27949)	0.369313 (0.004461)
100	3.16872 (0.287169)	7.22388 (0.801778)	0.352024 (0.002214)	9.67339 (2.681690)	11.37621 (1.992660)	0.360692 (0.002277)
300	3.06021 (0.089149)	7.04191 (0.254291)	0.347940 (0.000749)	9.13739 (0.794986)	11.09790 (0.644356)	0.358282 (0.000771)

Table 2: Bayes Estimates (Jeffreys) of Burr Type X mixture parameters under SELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	3.30176 (1.38407)	7.40462 (3.660540)	0.367358 (0.008280)	10.11900 (8.130706)	11.63690 (9.06919)	0.380166 (0.008483)
50	3.19966 (0.607675)	7.17565 (1.661650)	0.361885 (0.004347)	9.37646 (5.178210)	11.35580 (4.16771)	0.370783 (0.004458)
100	3.08849 (0.280949)	7.12302 (0.793535)	0.350966 (0.002198)	9.25758 (2.523130)	11.22860 (1.96647)	0.356951 (0.002262)
300	3.02764 (0.088078)	7.06139 (0.257066)	0.349584 (0.000747)	9.08948 (0.794586)	11.09280 (0.634268)	0.355318 (0.000771)

Table 3: Bayes Estimates (Gamma) of Burr Type X mixture parameters under SELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	4.27375 (1.17377)	9.52703 (3.444110)	0.369743 (0.008279)	11.239910 (7.30536)	14.96091 (8.066221)	0.377196 (0.008415)
50	3.57024 (0.552699)	8.29246 (1.38648)	0.363936 (0.004340)	0.48351 (4.612960)	12.85350 (3.52851)	0.370791 (0.004451)
100	3.30485 (0.258689)	7.61572 (0.734721)	0.351129 (0.002109)	9.70951 (1.87269)	11.90730 (1.738860)	0.357211 (0.002245)
300	3.09703 (0.086911)	7.202190 (0.249998)	0.348890 (0.000747)	9.34969 (0.719263)	11.32190 (0.602342)	0.360447 (0.000770)

Table 4: Bayes estimates (Exponential) of Burr Type X mixture parameters under SELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	2.72059 (0.792067)	5.28858 (1.696070)	0.368870 (0.008289)	5.34828 (2.947610)	6.93937 (2.898230)	0.375599 (0.008459)
50	2.86723 (0.451839)	6.028280 (1.125201)	0.363530 (0.004352)	6.65760 (2.390270)	8.54611 (2.249980)	0.375230 (0.004499)
100	2.93150 (0.244612)	6.48006 (0.644396)	0.350943 (0.002100)	7.58245 (1.62458)	9.68817 (1.438101)	0.359929 (0.002269)
300	2.947962 (0.084543)	6.82801 (0.239022)	0.348355 (0.000746)	8.47065 (0.682571)	10.45710 (0.560437)	0.357804 (0.000770)

Table 5: Bayes Estimates (Uniform) of Burr Type X mixture parameters under PLF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	3.63081 (0.372643)	7.81626 (0.466748)	0.379789 (0.021960)	10.76150 (1.10449)	12.36130 (0.738157)	0.391052 (0.022816)
50	3.22798 (0.172190)	7.30170 (0.222966)	0.370674 (0.011806)	9.89830 (0.528004)	11.58250 (0.353685)	0.378963 (0.012049)
100	3.14062 (0.087854)	7.12907 (0.108429)	0.354127 (0.006196)	9.41779 (0.265448)	11.31340 (0.172076)	0.361765 (0.006358)
300	3.04858 (0.029090)	7.06232 (0.054014)	0.349588 (0.002135)	9.17393 (0.086752)	11.096440 (0.056676)	0.357110 (0.002203)

Table 6: Bayes Estimates (Jeffreys) of Burr Type X mixture parameters under PLF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	3.97559 (0.370027)	8.14414 (0.458917)	0.380948 (0.022075)	11.6103 (1.08062)	12.9970 (0.73241)	0.386649 (0.022521)
50	3.46248 (0.175344)	7.56759 (0.224237)	0.369176 (0.011778)	10.2128 (0.517186)	11.90650 (0.352806)	0.378516 (0.012144)
100	3.25057 (0.088455)	7.22929 (0.108306)	0.353921 (0.006232)	9.73641 (0.264949)	11.43030 (0.171243)	0.366667 (0.006352)
300	3.06459 (0.028708)	7.10131 (0.036093)	0.351089 (0.002121)	9.23960 (0.086554)	11.13260 (0.056583)	0.361796 (0.002198)

Table 7: Bayes Estimates (Gamma) of Burr Type X mixture parameters under PLF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	4.39203 (0.352954)	9.76091 (0.437120)	0.378361 (0.022084)	12.19460 (0.979985)	15.15220 (0.678564)	0.386882 (0.022684)
50	3.69332 (0.172217)	8.34134 (0.217613)	0.368883 (0.011804)	10.63280 (0.495797)	13.24760 (0.345611)	0.380310 (0.012066)
100	3.36767 (0.087592)	7.68878 (0.107789)	0.354849 (0.006195)	9.88059 (0.256992)	12.04450 (0.168852)	0.364419 (0.006406)
300	3.11248 (0.028697)	7.25726 (0.036046)	0.349397 (0.002131)	9.30716 (0.085821)	11.32090 (0.056229)	0.356236 (0.002167)

Table 8: Bayes estimates (Exponential) of Burr Type X mixture parameters under PLF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	2.81632 (0.262129)	5.44615 (0.306887)	0.377792 (0.022090)	5.60420 (0.521610)	7.22505 (0.407127)	0.391388 (0.022505)
50	2.94209 (0.148991)	6.07860 (0.180117)	0.369035 (0.011772)	6.81321 (0.345029)	8.63802 (0.255956)	0.380553 (0.012140)
100	2.99418 (0.081478)	6.48661 (0.097179)	0.355187 (0.006215)	7.74382 (0.210726)	9.65180 (0.144598)	0.361397 (0.006318)
300	2.99668 (0.027978)	6.83629 (0.034746)	0.348832 (0.002128)	8.54831 (0.080078)	10.55130 (0.053628)	0.359814 (0.002168)

Table 9: Bayes Estimates (Uniform) of Burr Type X mixture parameters under WSELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	3.43773 (0.381970)	7.49636 (0.468523)	0.344168 (0.023988)	10.12221 (1.124690)	11.71480 (0.732172)	0.350418 (0.024484)
50	3.19497 (0.177499)	7.26792 (0.227122)	0.351243 (0.012408)	9.45622 (0.525346)	11.29910 (0.353094)	0.356714 (0.012628)
100	3.09010 (0.088289)	7.17604 (0.110401)	0.343517 (0.006387)	9.22026 (0.263436)	11.18990 (0.172153)	0.355046 (0.006505)
300	3.01432 (0.028708)	7.02753 (0.036039)	0.346242 (0.002142)	9.09046 (0.086576)	11.05680 (0.056701)	0.353429 (0.002193)

Table 10: Bayes Estimates (Jeffreys) of Burr Type X mixture parameters under WSELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	2.94121 (0.367651)	6.96661 (0.46641)	0.343578 (0.024165)	8.89329 (1.111660)	10.97050 (0.731364)	0.351393 (0.024907)
50	2.96825 (0.174603)	6.97154 (0.224888)	0.349698 (0.012346)	8.91495 (0.524409)	10.97320 (0.349895)	0.363096 (0.012690)
100	2.99877 (0.088199)	6.99661 (0.109322)	0.345632 (0.006359)	9.01686 (0.263027)	10.98961 (0.171265)	0.351578 (0.006541)
300	3.00289 (0.028310)	6.99948 (0.036006)	0.347717 (0.002153)	8.93985 (0.085960)	10.99760 (0.056689)	0.353594 (0.002181)

Table 11: Bayes estimates (Gamma) of Burr Type X mixture parameters under WELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	3.88854 (0.363457)	9.20844 (0.447387)	0.344651 (0.024142)	10.86930 (1.01594)	14.0801 (0.684076)	0.355318 (0.024890)
50	3.42251 (0.173743)	8.15975 (0.223049)	0.350421 (0.012402)	9.92928 (0.504056)	12.59300 (0.344234)	0.358373 (0.012587)
100	3.21551 (0.087619)	7.51155 (0.107951)	0.345565 (0.006369)	9.39564 (0.256021)	11.81140 (0.169747)	0.355374 (0.006594)
300	3.07914 (0.028858)	7.19696 (0.036060)	0.346770 (0.002149)	9.15469 (0.085799)	11.20710 (0.056153)	0.356592 (0.002192)

Table 12: Bayes estimates (Exponential) of Burr Type X mixture parameters under WELF

n	$\lambda_1 = 3$	$\lambda_2 = 7$	$w = 0.35$	$\lambda_1 = 9$	$\lambda_2 = 11$	$w = 0.35$
25	2.42438 (0.269376)	5.01342 (0.313339)	0.344712 (0.023980)	4.83713 (0.537459)	6.56521 (0.410325)	0.351672 (0.024566)
50	2.74146 (0.152303)	5.85477 (0.182962)	0.349687 (0.012335)	6.31033 (0.350574)	8.30897 (0.259655)	0.363016 (0.012676)
100	2.85994 (0.081713)	6.40238 (0.098498)	0.344886 (0.006382)	7.36190 (0.21034)	9.45204 (0.145416)	0.350197 (0.006543)
300	2.95728 (0.028165)	6.77115 (0.034724)	0.347872 (0.002152)	8.41971 (0.080188)	0.43219 (0.053498)	0.354439 (0.002182)

7.1. Graphical Representation of posterior Risks under Different Priors of Component Densities Graphs of Component density λ_1 Graphs of Component density λ_2

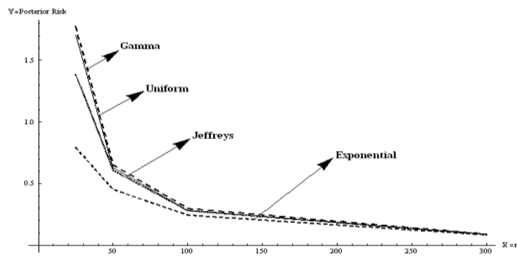

Fig. 7. Posterior Risks for $\lambda_1 = 3$ assuming different priors under SELF using $\lambda_2 = 7$ and $w = 0.35$.

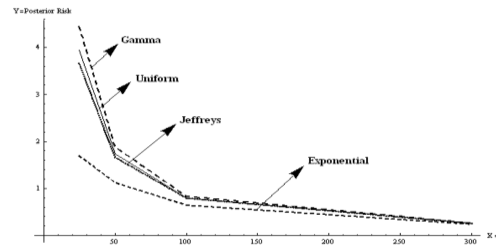

Fig. 10. Posterior Risks for $\lambda_2 = 7$ assuming different priors under SELF using $\lambda_1 = 3$ and $w = 0.35$.

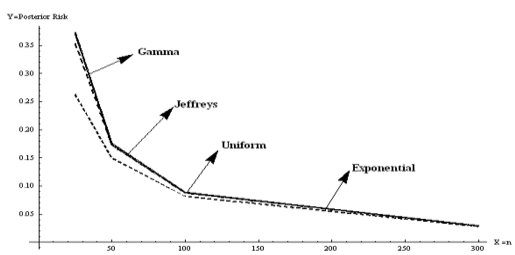

Fig. 8. Posterior Risks $\lambda_1 = 3$ for assuming different priors under PLF using $\lambda_2 = 7$ and $w = 0.35$.

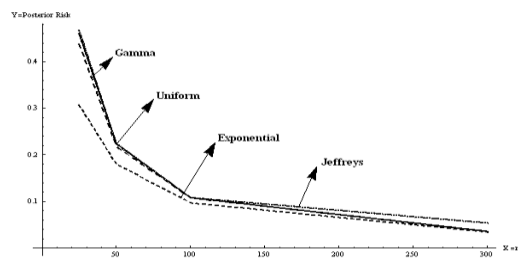

Fig. 11. Posterior Risks for $\lambda_2 = 7$ assuming different priors under PLF using $\lambda_1 = 3$ and $w = 0.35$.

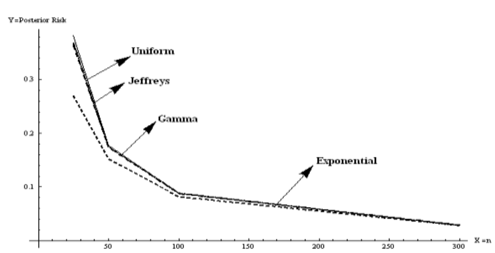

Fig. 9. Posterior Risks for $\lambda_1 = 3$ assuming different priors under WSELF using $\lambda_2 = 7$ and $w = 0.35$.

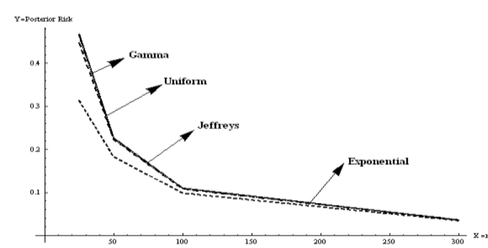

Fig. 12. Posterior Risks for $\lambda_2 = 7$ assuming different priors under WSELF using $\lambda_1 = 3$ and $w = 0.35$.

The simulation study has displayed some interesting properties of the Bayes point estimates. The posterior risks of the estimates of the lifetime parameters seem to be quite large (small) for the relatively larger (smaller) values of the parameters. However, in each case the posterior risks of estimates of lifetime parameters are reduced as the sample size increases.

The Bayes estimates under exponential prior are more precise with few exceptions than existing informative as well as non-informative counterparts as the averaged posterior risks of the mixture components are smaller than others. The Bayes (gamma) estimates of the first and second lifetime parameters are over-estimated whereas the Bayes (exponential) estimates of the lifetime parameters are under-estimated with having smaller posterior risks with few exceptions. Both the estimates of the lifetime parameters are overestimated under uniform and gamma priors but the tendency of over-estimation is greater in case of Bayes (gamma). On the other hand, estimates of the mixing proportion parameter are over-estimated/under-estimated. The trend of over-estimation is higher (lower) for smaller (larger) true values of the mixing parameter. In addition, in comparing the loss functions the posterior risks under PLF are less than SELF and WLF with few exceptions. It may be mentioned here that because of space restriction, results for all the variations in the parameters are not shown graphically. Only selected figures are included. The figure 7, 8 and 9 are the show the estimated posterior risks of the parameter of first component density under different priors. Similarly, the figures 10, 11, and 12 are the show the estimated posterior risks of the parameter of second component density under the different priors. The graphs for other combinations of the parametric values have the similar patterns. It is observed that as *n* increases, the risks of all the estimators decrease in all the considered cases. The posterior risks, under the exponential prior, are smaller for the parameter of both component densities as compared to other priors. Therefore, the Bayesian point estimation suggests the preference of the exponential prior under PLF for estimation.

Table 13: The lower limit (LL), the upper limit (UL) and the width of the 95% credible intervals under different priors

n	$\lambda_1 = 3$			$\lambda_2 = 7$			$w = 0.35$		
	LL	UL	Width	LL	UL	Width	LL	UL	Width
	Uniform Prior								
25	1.5835	5.6418	4.0582	4.3922	11.524	7.1252	0.2022	0.5566	0.3544
50	1.9006	4.7265	2.8259	4.9218	9.7879	4.8661	0.2412	0.4993	0.2580
10	2.1733	4.1957	2.0223	5.4999	8.9266	3.4267	0.2636	0.4477	0.1841
30	2.4929	3.6513	1.1583	6.0841	8.0528	1.9687	0.2982	0.4056	0.1073
	Jeffreys Prior								
25	1.4590	5.2054	3.8464	4.0562	10.972	6.9167	0.2010	0.5521	0.3510
50	1.7724	4.5223	2.7499	4.7424	9.5339	4.7914	0.2405	0.4964	0.2558
10	2.1013	4.0953	1.9939	5.4052	8.8058	3.4006	0.2631	0.4434	0.1803
30	2.4669	3.6198	1.1528	6.0507	8.0144	1.9637	0.2958	0.4030	0.1072
	Gamma Prior								
25	1.6191	5.6359	4.0168	4.1144	11.050	6.9362	0.2017	0.5535	0.3518
50	1.9192	4.7315	2.8123	4.7730	9.5715	4.7984	0.2409	0.4949	0.2540
10	2.1829	4.2003	2.0174	5.4210	8.8240	3.403	0.2613	0.4459	0.1846
30	2.4960	3.6535	1.1574	6.0561	8.0203	1.9641	0.2971	0.4028	0.1057
	Exponential Prior								
25	1.2127	4.3208	3.1080	4.3922	7.9255	3.5332	0.2019	0.5513	0.3494
50	1.6472	4.0964	2.4492	4.9218	8.0110	3.0892	0.2393	0.4961	0.2568
10	2.0127	3.8856	1.8729	5.4999	8.0398	2.5399	0.2629	0.4452	0.1823
30	2.4283	3.5567	1.1283	6.0841	7.7674	1.6833	0.2978	0.4020	0.1042

Table 13 gives the results for interval estimation. It is interesting to see that all the credible intervals contain the true value of the parameter. The credible intervals tend to be more specific under the assumption of the exponential prior. The width of credible interval is inversely proportional to sample size. The findings of the interval estimation also advocate that in order to estimate λ_1, λ_2 and w, the use of exponential prior can be preferred. It should be noted that the credible intervals for the other combination of the parametric values have not been presented as they follow the similar patterns.

9. Real Life Example

This section covers the analysis of real life data set regarding the breaking strengths of 64 single carbon fibers of length 10, presented Lawless (2003). The idea has been to see whether the results and properties of the Bayes estimators, explored by simulation study, are applicable to a real life situation. We have taken n = 64 and T = 3.501 in order to have censoring rate close to 20% (that has been used in simulation study). The results of the analysis have been reported in the following tables. The amounts of posterior risks associated with each estimate have been presented in the parenthesis in the tables.

Table 14: Bayes estimates and posterior risks under real life data using $w = 0.35$

Loss Function	λ_1	λ_2	w	λ_1	λ_2	w
	Uniform			Jeffreys		
SELF	586.64784 (5998.50714)	684.96707 (6139.78479)	0.36710 (0.03126)	581.87619 (5963.70792)	683.80672 (6103.83917)	0.36561 (0.03112)
PLF	590.09875 (6.90183)	689.01264 (8.09115)	0.37445 (0.01469)	585.49791 (7.24344)	688.33968 (9.06591)	0.37322 (0.01520)
WSELF	577.62248 (9.02535)	674.96094 (10.00612)	0.35092 (0.01618)	571.51824 (10.35795)	671.49608 (12.31064)	0.34893 (0.01668)
	Gamma			Exponential		
SELF	594.25344 (5954.74938)	709.85097 (6076.50709)	0.36769 (0.03069)	570.01915 (5645.80140)	666.24748 (6006.24078)	0.35989 (0.02945)
PLF	596.46348 (4.42009)	713.06151 (6.42107)	0.37289 (0.01039)	571.94323 (3.84816)	668.29623 (4.09751)	0.36344 (0.00909)
WSELF	587.48067 (6.77277)	700.92605 (8.92492)	0.35209 (0.01560)	563.57439 (6.44476)	660.81255 (5.43493)	0.34767 (0.01222)

Table 15: Bayes estimates and posterior risks under real life data using $w = 0.45$

Loss Function	λ_1	λ_2	w	λ_1	λ_2	w
	Uniform			Jeffreys		
SELF	578.77102 (5725.84773)	675.89204 (8807.06835)	0.44713 (0.03831)	574.04339 (5680.47114)	674.76413 (80343.45159)	0.44572 (0.03848)
PLF	582.13541 (6.72878)	680.07069 (8.35729)	0.45661 (0.01895)	577.43224 (6.67769)	680.12762 (7.72698)	0.45596 (0.02049)
WSELF	571.06410 (7.70692)	661.19582 (14.69623)	0.42753 (0.01960)	566.79832 (7.24507)	661.61529 (13.14884)	0.42410 (0.02162)
	Gamma			Exponential		
SELF	586.29208 (5670.51313)	700.42877 (9689.56536)	0.44837 (0.03779)	570.65717 (5280.48484)	667.87697 (7600.81798)	0.41940 (0.03613)
PLF	588.44681 (4.30948)	704.01181 (7.16607)	0.45426 (0.01178)	571.34555 (1.37675)	669.45127 (3.14860)	0.42444 (0.01007)
WSELF	579.21176 (7.08031)	688.99712 (11.43165)	0.43218 (0.01619)	564.21325 (6.44392)	656.67974 (11.19723)	0.40908 (0.01332)

The tables 14-15 contain the Bayes estimates and posterior risks for the mixture distribution using real life data. It can be observed from the analysis that the performance of the informative priors is better than the non-informative priors. In comparison of informative priors, the least amounts of the posterior risks have been observed under exponential prior for each loss function. On the other hand, the performance of the PLF seems better than SELF and WSELF for all priors. The increase in the value of the mixing parameter has a positive impact on the performance of the estimates for the parameter of the first component of the mixture. This is due to the reason that the increase in the values of the mixing parameter will add more values for the analysis of the first component of the mixture. Therefore, the findings from the analysis of real life data are in accordance with those of simulation study, suggesting the preference of exponential prior under PLF.

10. Hazard Rate for the Mixture of Burr Type X Distribution

The hazard rate is a useful way of describing the distribution of 'time to event' because it has a natural interpretation that relates to the aging of a population. The hazard function is the risk of failure in a small time interval, given survival at the beginning of the time interval. As a function of time, a hazard function may be increasing; meaning as time increases the rate for failure increases, for example, when a patient is untreated for a disease such as cancer or the medication do not work properly; may be decreasing, for example, as a person is recovering from severe trauma like a surgery; or may be constant, meaning the rate of failure is the same regardless of how much time has passed. The constant hazard rate is mostly unrealistic. The hazard rate for the mixture of Burr Type X distribution has been compared under a range of parametric values.

Hazard rate function for mixture of Burr Type X distribution is:

$$H(t)=\frac{\pi 2\lambda_1 x\exp\left(-x^2\right)\left\{1-\exp\left(-x^2\right)\right\}^{\lambda_1-1}+(1-\pi)2\lambda_2 x\exp\left(-x^2\right)\left\{1-\exp\left(-x^2\right)\right\}^{\lambda_2-1}}{1-\left\{\pi\left\{1-\exp\left(-x^2\right)\right\}^{\lambda_1}+(1-\pi)\left\{1-\exp\left(-x^2\right)\right\}^{\lambda_2}\right\}} \qquad (22)$$

The graphs for the hazard rate of the mixture model, for different parametric values and for the various ranges of the variable, are presented in the following. The abbreviations in the graphs are: HR: hazard rate; PR1: $\theta_1 = 0.50$, $\theta_2 = 0.75$; PR2: $\theta_1 = 0.75$, $\theta_2 = 2.00$; $\theta_1 = 1.50$, $\theta_2 = 3.00$; PR4: $\theta_1 = 2.50$, $\theta_2 = 4.00$.

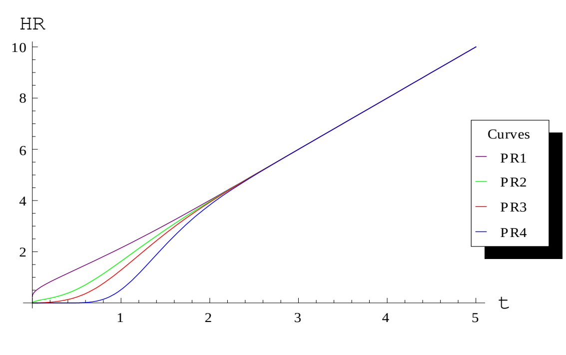

Fig. 13: Graph of hazard rates for mixture of model using $\pi = 0.25$ and $0 < t < 5$

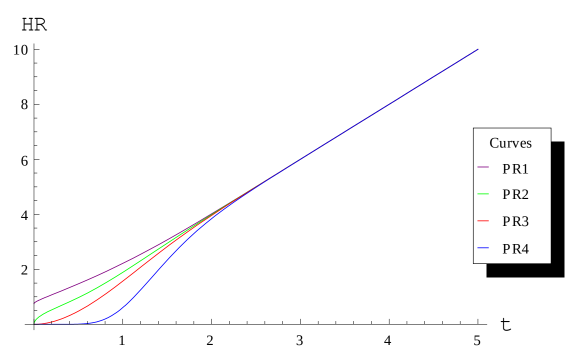

Fig. 15: Graph of hazard rates for mixture of model using $\pi = 0.75$ and $0 < t < 5$

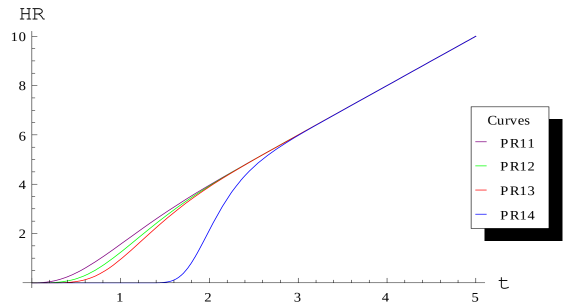

Fig. 14: Graph of hazard rates for mixture of model using $\pi = 0.25$ and $0 < t < 5$

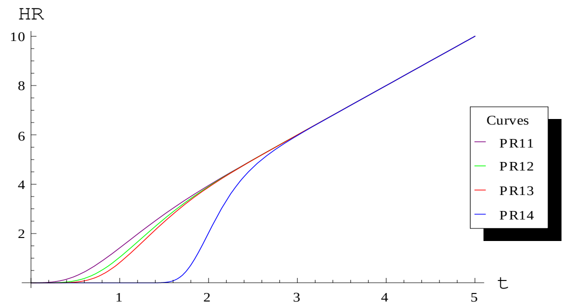

Fig. 16: Graph of hazard rates for mixture of model using $\pi = 0.75$ and $0 < t < 5$

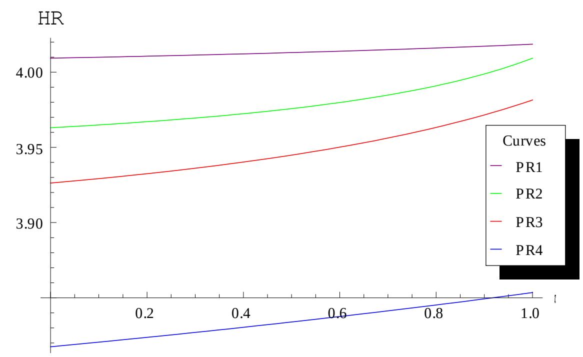

Fig. 17: Graph of hazard rates for mixture of model using $t = 2$ and $0 < \pi < 1$

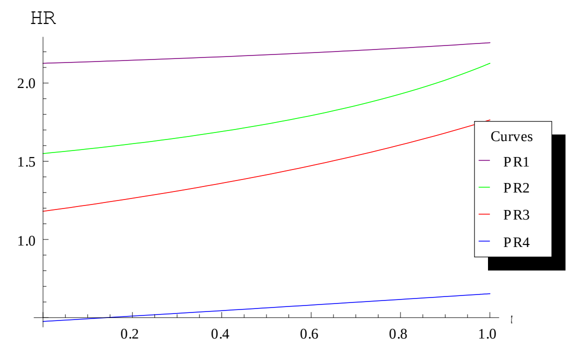

Fig. 18: Graph of hazard rates for mixture of model using $t = 1$ and $0 < \pi < 1$

The graphs suggest that the hazard rate for the mixture model tend to decrease for smaller 't' and larger parametric values. However, for $t > 3$ the choice of parametric values does not have a significant impact on the behavior of the hazard rate. The figures simply suggest that the hazard rate is an increasing function. The increase in the value of the mixing parameter (w) increases the hazard rate. The results are in accordance with the theory; because by increasing the mixing weight there are chances of more failures.

11. Model Comparison Criteria

Bernardo (1979) proposed that under Bayesian inference, the performance of a model is based on the posterior predictive distribution. The criterion used to compare them is based on the use of the logarithmic score as a utility function in a statistical decision. When the uncertainty is contained in the value of a future observation $y = x_{n+1}$ the logarithmic score $\log\left(g_k\left(y|\mathrm{x}\right)\right)$ is used, where $g_k\left(y|\mathrm{x}\right)$ denotes the posterior predictive density under model M_k. Then, the posterior predictive expected utility is given by: $\bar{U}_k = \int \log\left(g_k\left(y|\mathrm{x}\right)\right) g_k\left(y|\mathrm{x}\right) dy$. The optimal solution to the decision problem of choosing among the competing models M_0, M_1, ... , Mw is given by the model M_{k*}, such that: $\bar{U}_{k^*} = \max\limits_{k\in(0,1,\ldots,w)} \bar{U}_k$. Practically, the $\bar{U}_k$ can be estimated as: $\hat{\bar{U}}_k = \frac{1}{m}\sum_{i=1}^{m} \log\left(g_k\left(y_i|\mathrm{x}\right)\right)$.

where $y_1, y_2, \ldots, y_m$ are an independent and identically distributed random sample from $g_k\left(y|\mathrm{x}\right)$. This criterion can be used for selection of a suitable prior for the posterior analysis. The prior for which the posterior predictive distribution produces the maximum amount of posterior predictive expected utilities can be considered to be the best model. The more details can be seen from: Martin and Perez (2009). For numerical illustration, we have generated a random sample of size 30 from mixture of Burr type x distribution with for $\lambda_1 = 1.80$, $\lambda_2 = 2.00$ and w = 0.35. The values of the sample are:

Table 16: Simulated sample from mixture of Burr type x distribution for $\lambda_1 = 1.80$, $\lambda_2 = 2.00$ and w = 0.35

0.571361	0.926639	1.102480	1.247594	1.456517	1.689385
0.666807	0.967292	1.137095	1.270262	1.499276	1.715857
0.748350	1.005153	1.162367	1.316861	1.545267	1.848867
0.830025	1.034984	1.190743	1.368140	1.600780	1.940499
0.881668	1.058690	1.215792	1.411192	1.643376	2.127397

Now a sample of size 1000 for the parameters of the posterior distribution under each prior is generated. Based on these samples the estimated values for the posterior predictive expected utilities have been presented in the following table.

Table 17: The posterior predictive expected utilities under different priors

Uniform	Jeffreys	Gamma	Exponential
-1.77146	-1.75013	-1.70721	-1.63065

Now based on these posterior predictive expected utilities it can be assessed that the posterior distribution under exponential prior is the best among all the posterior densities. Hence, the most suitable prior is exponential prior. The findings from the model comparison criteria endorsed the preference of the exponential prior as suggested by the analysis of simulated and real life data sets.

12. Conclusion

The study has been planned to select a suitable prior for the Bayesian analysis of the parameters of the two-component mixture of the Burr Type X distribution. Four informative and non-informative priors have been assumed for posterior analysis under three loss functions. The model selection criterion for selection of a suitable prior has also been introduced. The findings of the study suggest that the exponential prior is the most suitable prior for the estimation of the parameters of the mixture density. The study is useful for the analysts looking to model the heterogeneous data through some lifetime distributions under censored samples. The study can further be extended by considering some other censoring scheme or mixing two different distributions from Burr family.

References

1. Afify, W. M. (2011). Classical estimation of mixed Rayleigh distribution in type i progressive censored, *Journal of Statistical Theory and Applications*, 10(4), 619-632.
2. Ahmad, K. E., Fakhry, M. E. and Jaheen, Z. F. (1997). Empirical Bayes estimation of P(Y < X) and characterization of Burr-type X model, *Journal of Statistical Planning and Inference*, 64, 297-308**.**
3. Aslam, M. (2003). An application of prior predictive distribution to elicit the prior density, *Journal of Statistical Theory and Applications*. 2(1), 70-83.
4. Bernardo, J. M. (1979). *Expected information as expected utility*. Ann. Statist. 7, 686-690.
5. Bolstad, W. M. (2004). Introduction to Bayesian Statistics. *John Wiley and Sons, Inc. New York*.
6. Burr, I. W. (1942). Cumulative frequency distribution, *Annals of Mathematical Statistics*, 13, 215-232.
7. Gauss, C. F. (1810). M´ethode des Moindres Carr´es. M´emoire sur la Combination des Observations. Transl. J. Bertrand (1955). Mallet-Bachelier, Paris.
8. Demidenko, E. (2004). Mixed Models: Theory and applications. *Wiley, New Jersey*.
9. Eberly, L. E. and Casella, G. (2003). Estimating Bayesian credible intervals. *Journal of Statistical Planning and Inference*, 112, 115-132.
10. Erisoglu, U., Erisoglu, M. and Erol, H. (2011). A mixture model of two different distributions approach to the analysis of heterogeneous survival data, *World Academy of Science, Engineering and Technology*, 78, 41-45.
11. Ghosh, S. K. Ebrahimi, N. (2001). Bayesian analysis of the mixing function in a mixture of two exponential distributions. Tech. Rep. 2531, Institute of Statistics *Mimeographs, North Carolina State University, North Carolina State University*.
12. Ismail, S. A. and El-Khodary, I. H. (2001). Characterization of mixtures of exponential family distributions through conditional expectation, *Annual Conference on Statistics and Computer Modeling in Human and Social Sciences*, 13, 64-73.
13. Jaheen. Z. F. (1996). Empirical Bayes estimation of the reliability and failure rate functions of the Burr type X failure model, *Journal of Applied Statistical Sciences*, 3, 281-288.
14. Legendre, A. (1805). Nouvelles M´ethodes pour la D´etermination des Orbitesdes Com.etes. Courcier, Paris.
15. McLachlan, G. J. and Peel, D. (2000). Finite Mixture Models, *John Wiley and Sons, New York*.
16. Martin, J. and Perez, C. J. (2009). Bayesian analysis of a generalized lognormal distribution, *Computational Statistics and Data Analysis*, 53, 1377-1387.
17. Mcculloch, C. E. and Searle, S. R. (2001). *Generalized, Linear and Mixed Models*, Wiley, New York.

18. Mendenhall, W. and Hadar, R. J. (1958). Estimation of parameters of mixed exponen-tially distributed failure time distributions from censored life test data. *Biometrika*, 45(3-4), 504-520.

19. Nair, M. T. and Abdul, E. S. (2010). Finite mixture of exponential model and its applications to renewal and reliability theory, *Journal of Statistical Theory and Practice*, 4(3): 367-373.

20. Raqab, M. Z. (1998). Order statistics from the Burr type X model, *Computers Mathematics and Applications*, 36, 111-120.

21. Saleem M. and Aslam M. (2008). Bayesian analysis of the two component mixture of the Rayleigh distribution assuming the Uniform and the Jeffreys prior from censored data, *J. App. Statist. Science*, 16(4), 105-113.

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

23. Saleem, M., Aslam, M. and Economou, P. (2010). On the Bayesian analysis of the mixture of power function distribution using the complete and the censored sample, *Journal of Applied Statistics*, 37(1), 25-40.

24. Sartawi, H. A. and Abu-Salih, M. S. (1991). Bayes prediction bounds for the Burr type X model, *Communications in Statistics - Theory and Methods*, 20, 2307-2330.

25. Shi, Y. and Yan, W. (2010). The EB estimation of scale parameter for two parameter exponential distribution under the type-I censoring life test, *J. Phys. Sci.* 4, 25-30.

26. Sultan, K. S., Ismail, M. A. and Al-Moisheer, A. S., 2007. Mixture of two inverse Weibull distributions: properties and estimation, *Computational Statistics & Data Analysis*, 51, 5377-5387.

27. Surles, J. G. and Padgett, W. J. (2001). Inference for reliability and stress-strength for a scaled Burr Type X distribution, *Lifetime Data Analysis*, 7, 187-200.