Electronic Journal of Applied Statistical Analysis
EJASA (2013), Electron. J. App. Stat. Anal., Vol. 6, Issue 1, 32 – 56
e-ISSN 2070-5948, DOI 10.1285/i20705948v6n1p32

CHOICE OF SUITABLE INFORMATIVE PRIOR FOR THE SCALE PARAMETER OF THE MIXTURE OF LAPLACE DISTRIBUTION

Sajid Ali[1,2]*, Muhammad Aslam[2], Syed Mohsin Ali Kazmi[3]

[1]*Department of Department of Decision Sciences, Bocconi University, Milan, Italy.*
[2]*Department of Statistics, Quaid-i-Azam University Islamabad, 45320, Pakistan.*
[3]*Sustainable Development Policy Institute Islamabad 44000, Pakistan.*

Received 07 December 2011; Accepted 28 July 2012
Available online 26 April 2013

***Abstract**: The major problem in Bayesian analysis is the choice of prior for the specified model. In the current study, the motivation is the comparsion of the informative priors for the mixture of Laplace distribution under different loss functions. A prior is selected based on the minimum posterior risks criteria where Bayes estimates and posterior risk are evaluated using the square error loss function, the precautionary loss function, the weighted squared error loss function and the modified (quadratic) squared error loss function. Bayes estimates and respective posterior risks are evaluated in terms of sample size, censoring rate and proportion of the component of the mixture using Levy and Gumbel Type-II informative priors. Limiting expressions for the complete sample are also derived. A real-life mixture data application has been discussed. The Elicitation of hyperparameters of mixture through prior predictive approach has also argued.*

***Keywords**: Censored sampling, Inverse transformation method, Hyperparameters; Elicitation, Fixed test termination time, Informative prior, Mixture distribution, Predictive intervals.*

1. Introduction

In our daily life it is very difficult to find homogenous life time data in real life so instead of simple model we will prefer a mixture of a family of life time distributions. A finite mixture of some appropriate probability distribution is recommended to study a population that is supposed to comprise a number of subpopulations mixed in an unknown proportion. A population of

* Email: sajidali.qau@hotmail.com

lifetimes of certain electrical elements may be divided into a number of subpopulations depending upon the possible cause of failure. Mixture models have been used in the physical, chemical, social science, biological and other fields. As examples, Harris [19] applied mixture distributions to modeling crime and justice data and Kanji [25] described wind shear data using mixture distributions.

For many years the Laplace distribution was a popular topic in probability theory due to simplicity of its characteristics function and density, the curious phenomenon that a random variable with only slight different characteristics function loses the simplicity of the density function and other numerous attractive probabilistic features enjoyed by this distribution. Censoring is imperative trait of the lifetime data because most of the times it is not possible to continue the experiment until the last observation in order to obtain a complete data set, i.e. a data set with exact life times of all the objects. Romeu [33] and Gijbels [17] have given an account of censoring.

In the last few decades with the advent of numerous computational methods, there has been a growing interest in the construction of flexible parametric classes of probability distributions in the Bayesian as compared to Classical approach. The aim of the present study is to investigate the heterogeneous population using the two-component Mixture of Laplace probability distribution using informative priors when data is censored and can be used to model various real world problems. Different aspects of Mixture of Laplace distribution have been considered before in literature by various authors like [1-3, 6, 10-14, 20, 22-26, 29-30, 34, 37].

The sequence of this paper is as follows. Section 2 introduces the Laplace mixture model, and its likelihood is developed in the Section 3. In Section 4, posterior distributions assuming informative priors are discussed. The Bayes estimators and their posterior risk under different loss functions are derived and discussed in Section 5. Posterior predictive distribution and predictive intervals are derived in the Section 6, while Section 7 deals with the method of Elicitation of hyperparameter for the mixture of Laplace distribution via prior predictive approach. Limiting expressions of these estimates and their posterior risk are presented in Section 8. A simulation study is performed in the Section 9 and a real life data for illustration purpose is discussed in the Section 10 for the evaluation of proposed method. The model comparison approach for the choice of prior is discussed in the Section 11. Some concluding remarks and future research proposal are given in the last Section 12.

2. The Population and the Model

A finite mixture distribution function with the two component densities of specified parametric form with unknown mixing weights (p, $q=1-p$) is defined as follows:

$$f(x) = pf_1(x) + (1-p)f_2(x), \qquad 0 < p < 1. \tag{1}$$

The following Laplace distribution is assumed for both components of the mixture with location parameter zero:

$$f_i(x) = \frac{1}{2\lambda_i}\exp\left(-\frac{|x|}{\lambda_i}\right), \qquad \lambda_i > 0,\ i = 1,2;\ -\infty < x < \quad .$$

So the mixture model (1) takes the following form:

$$f(x) = \frac{p}{2\lambda_1}\exp\left(-\frac{|x|}{\lambda_1}\right) + \frac{q}{2\lambda_2}\exp\left(-\frac{|x|}{\lambda_2}\right); \quad q = 1 - p,\ 0 < p < 1.$$

The graph of mixture density with different parameter values as follows:

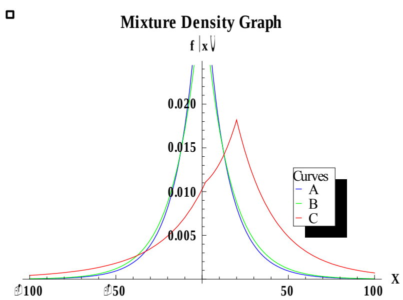

Figure 1. PDF of Mixture Density.

and the corresponding mixture cumulative distribution function is given by:

$$F(x) = pF_1(x) + qF_2(x) = p\left(1 - \frac{1}{2}e^{-\frac{x}{\lambda_1}}\right) + q\left(1 \ \frac{1}{2}e^{-\frac{x}{\lambda_2}}\right).$$

3. Sampling

Suppose n units from the above mixture model are engaged to a life testing experiment with a fixed test extinction time T. Let the test be conducted and it is observed that out of n, r units failed until the test termination time T is over and the remaining $n-r$ units are still functioning. As described in Mendenhall and Hader [28] and Ali et al. [1], in many real life situations only the failed objects can easily be identified as member of either subpopulation 1 or subpopulation 2. Hence, depending upon the cause of failure, it may be observed that r_1 and r_2 failures are from the first and the second subpopulation, respectively. Obviously the remaining $n-r$ censored objects provide no information about the subpopulation to which they belong, and $r = r_1 + r_2$ is the number of uncensored observations. For example, an engineer may categorize a failed electronic object as a member of the first or the second subpopulation based on the reason of its failure because he knows whether this component failed due to electricity shock or has some other manufacture problem based on his experience. Let we define, x_{ij} as the failure time of the j^{th} unit belonging to the i^{th} subpopulation, where $j = 1,2,3,...,r_i$, $i = 1,2$, $\left|x_{ij}\right| \le T$.

3.1 The Likelihood Function

The likelihood function for the above situation is:

$$L\left(\lambda_1,\lambda_2,p|\mathbf{x}\right) \propto \left\{\prod_{j=1}^{r_1} p\, f_1\left(x_{1j}\right)\right\}\left\{\prod_{j=1}^{r_2} q\, f_2\left(x_{2j}\right)\right\}\left\{(1-F(T))\right\}^{n\text{-}r}.$$

where $\mathbf{x} = \left(|x_{11}|, |x_{12}|,.....,|x_{1r_1}|, |x_{21}|, |x_{22}|,.....,|x_{2r_2}|\right)$ are the observed failure times for the non-censored observations. Since T should be positive, so further we are assuming positive side of Laplace model.

$$L(\lambda_1,\lambda_2,p\,|\,\mathbf{x}) \propto \left[\prod_{j=1}^{r_1}\frac{p}{2\lambda_1}\exp\left(-\frac{\left|x_{1j}\right|}{\lambda_1}\right)\right]\left[\prod_{j=1}^{r_2}\frac{q}{2\lambda_1}\exp\left(-\frac{\left|x_{2j}\right|}{\lambda_2}\right)\right]\left[1-\left\{p\left(1-\frac{1}{2}\mathrm{e}^{-\frac{T}{\lambda_1}}\right)+q\left(1\quad\frac{1}{2}\mathrm{e}^{-\frac{T}{\lambda_2}}\right)\right\}\right]^{(n-r)}$$

After simplifications, the above equation can be represented as:

$$L(\lambda_1,\lambda_2,p\,|\,\mathbf{x}) \propto \sum_{m=0}^{n-r}\binom{n-r}{m}p^{(n-r_2-m)}q^{(r_2+m)}\left(\frac{1}{\lambda_1}\right)^{r_1}\left(\frac{1}{\lambda_2}\right)^{r_2} e^{-\frac{1}{\lambda_1}\left[\sum_{j=1}^{r_1}\left|x_{1j}\right|+(n-r-m)T\right]} e^{-\frac{1}{\lambda_2}\left[\sum_{j=1}^{r_2}\left|x_{2j}\right|+mT\right]} \qquad (2)$$

4. Posterior Distributions assuming Informative Priors

In case of an informative prior, the use of prior information is equivalent to adding a number of observations to a given sample size, and therefore leads to a reduction of the variance/posterior risk of the Bayes estimates. We used the Levy and Gumbel Type-II informative priors for analysis.

4.1 Posterior Distribution using the Inverse Levy Prior (LP)

Suppose $\lambda_1 \sim Levy(b_1), \lambda_2 \sim Levy(b_2)$ and $p \sim U(0,1)$, also assuming independence, we have a joint prior distribution $h(\lambda_1,\lambda_2,p) \propto \left(\frac{1}{\lambda_1}\right)^{(1.5)} e^{-\frac{b_1}{2\lambda_1}}\left(\frac{1}{\lambda_2}\right)^{(1.5)} e^{-\frac{b_2}{2\lambda_2}}$, to get a joint posterior distribution of λ_1, λ_2 and 'p'. The joint posterior distribution of λ_1, λ_2 and 'p' is as following:

$$h(\lambda_1,\lambda_2,p\,|\,\mathbf{x}) \propto \sum_{m=0}^{n-r}\binom{n-r}{m}p^{(n-r_2-m)}q^{(r_2\quad m)}\left(\frac{1}{\lambda_1}\right)^{(A+1)}\left(\frac{1}{\lambda_2}\right)^{(B+1)} e^{-\frac{1}{\lambda_1}[C]} e^{-\frac{1}{\lambda_2}[D]}, \lambda_i > 0, i = 1,2, 0 < p < 1.$$

The marginal distributions are given in equations (3-5) below:

$$h(\lambda_1 \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\left(\frac{1}{\lambda_1}\right)^{(A+1)} e^{-\frac{1}{\lambda_1}[C]} E\left(\Gamma\left(B\right)/ D^{B}\right), \lambda_1 > 0 \tag{3}$$

$$h(\lambda_2 \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\left(\frac{1}{\lambda_2}\right)^{(B+1)} e^{-\frac{1}{\lambda_2}[D]} E\left(\Gamma\left(A\right)/ C^{A}\right), \lambda_2 > 0 \tag{4}$$

and

$$h(p \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m} p^{(n-r_2-m)} q^{(r_2+m)} \left(\left(\Gamma\left(A\right)\Gamma\left(B\right)\right)/ \left(C^{A}D^{B}\right)\right), 0 < p < 1 \tag{5}$$

where A, B, C, D, E and F are defined as:

$$A = r_1 + \frac{1}{2}, B = r_2 + \frac{1}{2}, C = \sum_{j=1}^{r_1}\left|x_{1j}\right| + (n-r-m)T + \frac{b_1}{2}, D = \sum_{j=1}^{r_2}\left|x_{2j}\right| + mT + \frac{b_2}{2},$$

$$E = \beta(n - r_2 - m + 1, r_2 + m + 1) \text{ and } F = \sum_{m=0}^{n-r}\binom{n-r}{m} E\left(AB/\left(C^{A}D^{B}\right)\right)$$

4.2 Posterior Distribution using the Gumbel Type-II Prior

Suppose $\lambda_1 \sim Gumbel\,Type - II(1, b_3)$, $\lambda_2 \sim Gumbel\,Type - II(1, b_4)$ and $p \sim U(0,1)$. Assuming independence, we have a joint prior $P(\lambda_1, \lambda_2, p) \propto \left(\frac{1}{\lambda_1}\right)^{(2)} e^{-\frac{b_3}{\lambda_1}} \left(\frac{1}{\lambda_2}\right)^{(2)} e^{-\frac{b_4}{\lambda_2}}$, which is combined with the likelihood function to get a joint posterior distribution of λ_1, λ_2 and *p*. The joint posterior distribution of λ_1, λ_2 and '*p*' is as following:

$$h(\lambda_1, \lambda_2, p \mid \mathbf{x}) \propto \sum_{m=0}^{n-r}\binom{n-r}{m} p^{(n-r_2-m)} q^{(r_2+m)} \left(\frac{1}{\lambda_1}\right)^{(G+1)} \left(\frac{1}{\lambda_2}\right)^{(H+1)} e^{-\frac{1}{\lambda_1}[I]} e^{-\frac{1}{\lambda_2}[J]}$$
$$0 < p \le 1, 0 < \lambda_i \le \infty, i = 1, 2$$

The marginal distribution of each parameter is obtained in equations (6-8) as below:

$$h(\lambda_1 \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\left(\frac{1}{\lambda_1}\right)^{(G+1)} e^{-\frac{1}{\lambda_1}[I]} E\left(\Gamma\left(H\right)/ J^{H}\right), \lambda_1 > 0 \tag{6}$$

$$h(\lambda_2 \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\left(\frac{1}{\lambda_2}\right)^{(H+1)} e^{-\frac{1}{\lambda_2}[J]} E\left(\Gamma\left(G\right)/ I^{G}\right), \lambda_2 > 0 \tag{7}$$

and

$$h(p \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m} p^{(n-r_2-m)} q^{(r_2+m)} \left(\left(\Gamma\left(G\right)\Gamma\left(H\right)\right)/\left(I^{G} J^{H}\right)\right), 0 < p < 1 \tag{8}$$

where G, H, I, J, E and K are defined as:

$$G = r_1 + 1, H = r_2 + 1, I = \sum_{j=1}^{r_1}\left|x_{1j}\right| + (n-r-m)T + b_3, J = \sum_{j=1}^{r_2}\left|x_{2j}\right| + mT + b_4,$$

$$E = \beta(n - r_2 - m + 1, r_2 + m + 1) \text{ and } K = \sum_{m=0}^{n-r}\binom{n-r}{m} E\left(GH/\left(I^{G} J^{H}\right)\right)$$

5. Bayes Estimators and Posterior Risks under Different Loss Functions

The Bayes estimators are evaluated under squared error loss function (SELF), weighted squared error loss function (WSELF) and modified (quadratic) squared error loss function (M/QSELF). Since this is a symmetrical loss function that assigns equal losses to overestimation and underestimation and it is often used because it does not lead to extensive numerical computation. Norstrom [31] introduced an alternative asymmetric precautionary loss function, which approaches infinity near the origin to prevent underestimation, thus giving conservative estimators especially when underestimation may lead to serious consequence. Following table 1, shows general form of Bayes estimators under different loss functions with their respective posterior risk.

Table 1. Bayes Estimators (BE) and Posterior Risk of different Loss Functions.

Loss Function Name	Mathematical Form	Bayes Estimator	Posterior Risk
SELF	$L_1 = L(\lambda, \lambda^*) = \left(\lambda - \lambda^*\right)^2$	$d^* = E(\lambda \mid \mathbf{x})$	$E_{\lambda\mid\mathbf{x}} L(\lambda, d) = E(\lambda^2 \mid \mathbf{x}) - E(\lambda \mid \mathbf{x})^2$
WSELF	$L_3 = L(\lambda, d) = \frac{(\lambda - d)^2}{\lambda}$	$d^* = \frac{1}{E(\lambda^{-1} \mid \mathbf{x})}$	$E_{\lambda\mid\mathbf{x}} L(\lambda, d) = E(\lambda \mid \mathbf{x}) - \frac{1}{E(\lambda^{-1} \mid \mathbf{x})}$
M/Q SELF	$L_2 = L(\lambda, d) = \left(1 \quad \frac{d}{\lambda}\right)^2$-	$d^* = \frac{E(\lambda^{-1} \mid \mathbf{x})}{E(\lambda^{-2} \mid \mathbf{x})}$	$E_{\lambda\mid\mathbf{x}} L(\lambda, d) = 1 - \frac{\left[E(\lambda^{-1} \mid \mathbf{x})\right]^2}{E(\lambda^{-2} \mid \mathbf{x})}$
PLF	$L_4 = L(\lambda, d) = \frac{(\lambda - d)^2}{d}$	$d^* = \sqrt{E(\lambda^2 \mid \mathbf{x})}$	$E_{\lambda\mid\mathbf{x}} L(\lambda, d) = 2\left(\sqrt{E(\lambda^2 \mid \mathbf{x})} - E(\lambda \mid \mathbf{x})\right)$

5.1 Bayes Estimators Using the LP

Taking expectation of each parameter with respect to its marginal distributions gives the Bayes estimator of the parameters. The Bayes estimators of λ_1, λ_2 and 'p' assuming the Levy prior are given as follows:

$$E(\lambda_1 \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma\left(A-1\right)\Gamma\left(B\right)/\,C^{(A-1)}D^{B}\right)$$

$$E(\lambda_2 \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma\left(A\right)\Gamma\left(B-1\right)/\,C^{A}D^{(B-1)}\right)$$

and

$$E(p \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m+2, r_2+m+1)\left(\Gamma\left(A\right)\Gamma\left(B\right)/\,C^{A}D^{B}\right)$$

respectively. Following are the Bayes Estimates of λ_1, λ_2 and 'p' assuming under MSELF:

$$E(\lambda_1^{-1} \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A+1)\Gamma(B)/\left(C^{A+1}D^{B}\right)\right) \tag{9}$$

$$E(\lambda_1^{-2} \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A+2)\Gamma(B)/\left(C^{A+2}D^{B}\right)\right) \tag{10}$$

dividing equation (9) by (10) we get Bayes Estimate of λ_1.

$$d_1^* = \frac{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A+1)\,\Gamma(B)/\left(C^{A+1}D^{B}\right)\right)}{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A+2)\,\Gamma(B)/\left(C^{A+2}D^{B}\right)\right)}$$

$$E(\lambda_2^{-1} \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A)\Gamma(B+1)/\left(C^{A}D^{B+1}\right)\right) \tag{11}$$

$$E(\lambda_2^{-2} \mid \mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A)\Gamma(B+2)/\left(C^{A}D^{B+2}\right)\right) \tag{12}$$

dividing equation (11) by (12) we get Bayes Estimate of λ_2.

$$d_2^* = \frac{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A)\Gamma(B+1)/\left(C^A D^{B+1}\right)\right)}{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A)\Gamma(B+2)/\left(C^A D^{B+2}\right)\right)}$$

and for '*p*' mixing component:

$$E(p^{-1}\,|\,\mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m, r_2+m+1)\left(\Gamma(A)\Gamma(B)/\left(C^A D^B\right)\right) \tag{13}$$

$$E(p^{-2}\,|\,\mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m-1, r_2+m+1)\left(\Gamma(A)\Gamma(B)/\left(C^A D^B\right)\right)$$

(14)

dividing equation (13) by (14) we get Bayes Estimate of '*p*'.

$$d_p^* = \frac{\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m, r_2+m+1)\left(\Gamma(A)\Gamma(B)/\left(C^A D^B\right)\right)}{\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m-1, r_2+m+1)\left(\Gamma(A)\Gamma(B)/\left(C^A D^B\right)\right)}$$

respectively, where A, B, C, D, E and F are defined above. The Bayes estimators under WSELF and precautionary loss function can be evaluated using above two loss function equations.

5.2 Posterior Risk (PR) using the Levy Prior

The posterior risk of the Bayes estimators of λ_1 , λ_2 and *p* using the Levy prior are given as:

$$Var(\lambda_1\,|\,\mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A-2)\Gamma(B)/C^{(A-2)}D^B\right) - \left\{E(\lambda_1|\mathbf{x})\right\}^2$$

$$Var(\lambda_2\,|\,\mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(A)\Gamma(B-2)/C^A D^{(B-2)}\right) - \left\{E(\lambda_2|\mathbf{x})\right\}^2$$

and

$Var(p\,|\,\mathbf{x}) = F^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m+3, r_2+m+1)\left(\Gamma(A)\Gamma(B)/C^A D^B\right) - \left\{E(p|\mathbf{x})\right\}^2$ respectively. The PR under WSELF and precautionary loss function can be evaluated using above two loss function equations.

5.3 Bayes Estimators using the Gumbel Type-II Prior

The Bayes estimators of λ_1, λ_2 and p assuming the Gumbel Type-II prior are given as follow:

$$E(\lambda_1 \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G-1)\Gamma(H)/I^{(G-1)}J^{H}\right)$$

$$E(\lambda_2 \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G)\Gamma(H-1)/I^{G}J^{(H-1)}\right)$$

and

$$E(p \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m+2, r_2+m+1)\left(\Gamma(G)\Gamma(H)/I^{G}J^{H}\right)$$

respectively. Following are the Bayes Estimates of λ_1, λ_2 and p assuming MSELF:

$$E(\lambda_1^{-1} \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G+1)\Gamma(H)/\left(I^{G+1}J^{H}\right)\right) \tag{15}$$

$$E(\lambda_1^{-2} \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G+2)\Gamma(H)/\left(I^{G+2}J^{H}\right)\right) \tag{16}$$

dividing equation (15) by (16) we get Bayes Estimate of λ_1.

$$d_1^* = \frac{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G+1)\Gamma(H)/\left(I^{G+1}J^{H}\right)\right)}{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G+2)\Gamma(H)/\left(I^{G+2}J^{H}\right)\right)}$$

$$E(\lambda_2^{-1} \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G)\Gamma(H+1)/\left(I^{G}J^{H+1}\right)\right) \tag{17}$$

$$E(\lambda_2^{-2} \mid \mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G)\Gamma(H+2)/\left(I^{G}J^{H+2}\right)\right) \tag{18}$$

dividing equation (17) by (18) we get Bayes Estimate of λ_2.

$$d_2^* = \frac{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G)\Gamma(H+1)/\left(I^G J^{H+1}\right)\right)}{\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G)\Gamma(H+2)/\left(I^G J^{H+2}\right)\right)}$$

and for 'p' mixing component:

$$E(p^{-1}\,|\,\mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m, r_2+m+1)\left(\Gamma(G)\Gamma(H)/\left(I^G J^H\right)\right) \tag{19}$$

$$E(p^{-2}\,|\,\mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m-1, r_2+m+1)\left(\Gamma(G)\Gamma(H)/\left(I^G J^H\right)\right) \tag{20}$$

dividing equation (19) by (20) we get Bayes Estimate of 'p'.

$$d_p^* = \frac{\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m, r_2+m+1)\left(\Gamma(G)\Gamma(H)/\left(I^G J^H\right)\right)}{\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m-1, r_2+m+1)\left(\Gamma(G)\Gamma(H)/\left(I^G J^H\right)\right)}$$

respectively, where A, B, C, D, E and G are defined above. Similarly the Bayes Estimators under precautionary and WSELF can be evaluated using above two loss function Bayes Estimators equations.

5.4 Posterior Risk using the Gumbel Type-II Prior

The posterior risk of the BEs under SELF of λ_1, λ_2 and p using the Gumbel Type-II prior are given as:

$$Var(\lambda_1\,|\,\mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G-2)\Gamma(H)/I^{(G-2)}J^H\right) - \left\{E\left(\lambda_1\,|\,\mathbf{x}\right)\right\}^2$$

$$Var(\lambda_2\,|\,\mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}E\left(\Gamma(G)\Gamma(H-2)/I^G J^{(H-2)}\right) - \left\{E\left(\lambda_2\,|\,\mathbf{x}\right)\right\}^2$$

and

$$Var(p\,|\,\mathbf{x}) = K^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\beta(n-r_2-m+3, r_2+m+1)\left(\Gamma(G)\Gamma(H)/I^G J^H\right) - \left\{E(p\,|\,\mathbf{x})\right\}^2$$

respectively.

6. Predictive Distribution

The posterior distributions of the parameters λ_1, λ_2 and p given the data, likelihood and prior are recapitulate to have Bayes estimates of the parameters. The predictive distribution contains the information about the independent future random observation given the preceding observations. Bansal [7] has given a great detailed discussion about the posterior predictive distribution.

6.1 Predictive Distribution Intervals using the Levy Prior

The posterior predictive distribution of the future observation y=$|x_{(n+1)}|$ is:

$$p(y\mid\mathbf{x})=\int_0^1\int_0^\infty\int_0^\infty g(\lambda_1,\lambda_2,p|\mathbf{x})\,f(y\mid\lambda_1,\lambda_2,p)\,d\lambda_1 d\lambda_2 dp$$

where $f(y\mid\lambda_1,\lambda_2,p)=\frac{p}{2\lambda_1}e^{-\frac{y}{\lambda_1}}+\frac{q}{2\lambda_2}e^{-\frac{y}{\lambda_2}}$ is the future observation density and:

$$h(\lambda_1,\lambda_2,p\mid\mathbf{x})\propto\sum_{m=0}^{n-r}\binom{n-r}{m}p^{(n-r_2-m)+}q^{(r_2\ \ m)}\left(\frac{1}{\lambda_1}\right)^{(A+1)}\left(\frac{1}{\lambda_2}\right)^{(B+1)}e^{-\frac{1}{\lambda_1}[C]}e^{-\frac{1}{\lambda_2}[D]}$$

is the joint posterior distribution obtained by incorporating the Levy prior with the likelihood given by equation (2). The posterior predictive distribution of the future observation "y" is:

$$p(y\mid\mathbf{x})=\frac{1}{Z}\sum_{m=0}^{n-r}\binom{n-r}{m}\Big[\Big\{\beta(n-r_2-m+2,r_2+m+1)\Big(\Gamma\big(A+1\big)\Gamma\big(B\big)/(C+y)^{(A+1)}D^B\Big)\Big\}+$$
$$+\Big\{E\Big(\Gamma\big(A\big)\Gamma\big(B+1\big)/C^A(D+y)^{(B+1)}\Big)\Big\}\Big] \tag{21}$$

where $Z=\frac{F}{2}$.

A (1-α) 100% Bayesian interval *(L, U)* can be obtained by solving the following two equations simultaneously:

$$\int_{-\infty}^{L}p(y|\mathbf{x})dy=\frac{\alpha}{2}=\int_{U}^{\infty}p(y|\mathbf{x})dy$$

which further can be expressed as:

$$\frac{\alpha}{2}=\frac{1}{Z}\sum_{m=0}^{n-r}\binom{n-r}{m}\Big[\Big\{\beta(n-r_2-m+2,r_2+m+1)\Big(\Gamma\big(A+1\big)\Gamma\big(B\big)/A(C+L)^{A}D^B\Big)\Big\}+$$
$$+\Big\{E\Big(\Gamma\big(A\big)\Gamma\big(B+1\big)/BC^A(D+L)^{B}\Big)\Big\}\Big] \tag{22}$$

and

$$\frac{\alpha}{2} = \frac{1}{Z}\sum_{m=0}^{n-r}\binom{n-r}{m}\Big[\Big\{\beta(n-r_2-m+2, r_2+m+1)\Big(\Gamma\big(A+1\big)\Gamma\big(B\big)/A(C+U)^A D^B\Big)\Big\}+$$
$$+\Big\{E\Big(\Gamma\big(A\big)\Gamma\big(B+1\big)/BC^A(D+U)^B\Big)\Big\}\Big] \tag{23}$$

respectively. These posterior predictive intervals can be evaluated for a number of combinations of the hyperparameters which help us to determine a range of hyper-parameters that may lead to informative Bayes estimates having smaller variances than the non-informative Bayes estimates. Saleem and Aslam [35] and Saleem et al. [36] used predictive intervals for the Rayleigh mixture to discuss precision of Bayes estimates in terms of hyper-parameters.

6.2 Predictive Distribution Intervals using the Gumbel Type-II Prior

The posterior predictive distribution of the future observation "y" is:

$$p(y\,|\,\mathbf{x}) = \frac{1}{X}\sum_{m=0}^{n-r}\binom{n-r}{m}\Big[\Big\{\beta(n-r_2-m+2, r_2+m+1)\Big(\Gamma\big(G+1\big)\Gamma\big(H\big)/(I+y)^{(G+1)}J^H\Big)\Big\}+$$
$$+\Big\{E\Big(\Gamma\big(G\big)\Gamma\big(H+1\big)/I^G(J+y)^{(H+1)}\Big)\Big\}\Big] \tag{24}$$

where $X = \frac{K}{2}$ and predictive intervals are:

$$\frac{\alpha}{2} = \frac{1}{X}\sum_{m=0}^{n-r}\binom{n-r}{m}\Big[\Big\{\beta(n-r_2-m+2, r_2+m+1)\Big(\Gamma\big(G+1\big)\Gamma\big(H\big)/G(I+L)^G J^H\Big)\Big\}+$$
$$+\Big\{E\Gamma\big(G\big)\Gamma\big(H+1\big)/HI^G(J+L)^{(H)}\Big\}\Big] \tag{25}$$

and

$$\frac{\alpha}{2} = \frac{1}{X}\sum_{m=0}^{n-r}\binom{n-r}{m}\Big[\Big\{\beta(n-r_2-m+2, r_2+m+1)\Big(\Gamma\big(G+1\big)\Gamma\big(H\big)/G(I+U)^G J^H\Big)\Big\}+$$
$$\Big\{E\Gamma\big(G\big)\Gamma\big(H+1\big)/HI^G(J+U)^{(H)}\Big\}\Big] \tag{26}$$

respectively.

7. Elicitation

Elicitation is a system of extracting professional knowledge about some unknown measure of interest, or the probability of some prospect event, which can then be used to enrichment any numerical data that we may have. If the expert in question does not have a statistical background as is often the case, translating their beliefs into a statistical form suitable for use in our analyses can be a challenging task (see Dey [15]).

Prior elicitation is an important and yet under researched element of Bayesian statistics because we first have to decide prior distribution type and then parameters value. In any statistical analysis there will typically be some form of background knowledge available in addition to the data at hand. For example, suppose we are investigating the average lifetime of a component. We can do tests on a sample of components to learn about their average lifetime, but the designer/engineer of the component may have their own expectations about its performance. There are various methods available in literature for detail see [5, 8, 16, 21, and 32] and references cited therein.

7.1 Hyperparameter(s) Elicitation

Hyperparameter elicitation from the prior $g(\lambda)$ directly is conceptually difficult. The consensus of opinion amongst researchers is now to elicit expert knowledge about hyperparameters from observable quantities only. In fact prior predictive distribution removes the uncertainty in parameter (s) and reveals a distribution for the data point only. This superior approach is achievable by specifying summary features of the prior predictive density (mass) function $f(x)=\int_{-\infty}^{\infty} f(x|\lambda)g(\lambda)d\lambda$, which describes the probability distribution of the random variables X without conditioning on the parameter(s) $g(\lambda)$, yet is still a function of the unknown hyperparameters. The moments (mean, variance . . .) are unreasonable summary features of $f(x)$, as they are based on the non-trivial concept of mathematical expectation. The mode (most likely value) is perhaps the obvious summary feature, though ambiguity arises if the maximum is at an endpoint. Furthermore, the mode's extensions to relative likelihoods are not usually amenable for analysis. Perhaps the best summary features are quantiles or cumulative probabilities.

To determine (elicit) a prior density, Aslam [4] extend some new methods base on the prior predictive distribution. In his paper, he uses prior predictive probabilities, predictive mode and confidence level for eliciting the hyperparameters. The following method of elicitation is used in this study for determining hyperparameters of informative prior.

7.2 Method of Elicitation through Prior Predictive Probabilities Approach

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 $\psi\left(a_1,a_2\right)$ is defined in such a way that the hyperparameters $a_1\, and\, a_2$ are to be chosen by minimizing this function $\psi\left(a_1,a_2\right)=\min\limits_{a_1,a_2}\sum\limits_y\left\{\frac{p(y)-p_0(y)}{p(y)}\right\}^2$, where $p\left(y\right)$ denote the prior predictive probabilities characterized by the hyperparameters a_1 and a_2 and $p_0\left(y\right)$ denote the elicited prior predictive probabilities. The above equations solved simultaneously by applying 'PROC SYSLIN' of the SAS package for eliciting the required hyperparameters.

7.2.1 Elicitation through Prior Predictive Approach when Prior is Levy

The equation of prior predictive using the Inverse Chi-Squared prior is:

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

$$f(y)=\frac{p}{2\lambda_1}\exp\left(-|y|/\lambda_1\right)+\frac{q}{2\lambda_1}\exp\left(-|y|/\lambda_2\right);\quad q=1-p,\ 0<p<1.$$

$$p(y)=\frac{1}{2}\sqrt{\frac{b_1}{2\pi}}\sqrt{\frac{b_2}{2\pi}}\int_0^1\int_0^\infty\int_0^\infty\sum_{k=0}^{1}\binom{1}{k}\left(\frac{p}{\lambda_1}\exp\left(-\frac{|y|}{\lambda_1}\right)\right)^{1-k}\left(\frac{q}{\lambda_2}\exp\left(-\frac{|y|}{\lambda_2}\right)\right)^{k}$$
$$\left(\frac{1}{\lambda_1}\right)^{(1.5)}\exp\left(-\frac{b_1}{2\lambda_1}\right)\left(\frac{1}{\lambda_2}\right)^{(1.5)}\exp\left(-\frac{b_2}{2\lambda_2}\right)d\lambda_1 d\lambda_2 dp$$

So simplified form of above equation is:

$$f(y)=\frac{1}{8}\left[\frac{\left(0.5b_1\right)^{0.5}}{\left(|y|+0.5b_1\right)^{1.5}}+\frac{\left(0.5b_2\right)^{0.5}}{\left(|y|+0.5b_2\right)^{1.5}}\right]\tag{27}$$

we get the following hyper-parameters values $b_1=0.105369$ and $b_2=3.296676$.

7.2.2 Elicitation through Prior Predictive Approach when Prior is GTII

The equation of prior predictive using the Gumbel Type-II prior is:

$$f(y)=\frac{b_3b_4}{2}\int_0^1\int_0^\infty\int_0^\infty\sum_{k=0}^{1}\binom{1}{k}\left(\frac{p}{\lambda_1}\exp\left(-\frac{|y|}{\lambda_1}\right)\right)^{1-k}\left(\frac{q}{\lambda_2}\exp\left(-\frac{|y|}{\lambda_2}\right)\right)^{k}$$
$$\left(\frac{1}{\lambda_1}\right)^{(2)}\exp\left(-\frac{b_3}{\lambda_1}\right)\left(\frac{1}{\lambda_2}\right)^{(2)}\exp\left(-\frac{b_4}{\lambda_2}\right)d\lambda_1 d\lambda_2 dp$$

Which simplifies that:

$$f(y)=\frac{1}{4}\left[\frac{b_3}{\left(|y|+b_3\right)^2}+\frac{b_4}{\left(|y|+b_4\right)^2}\right] \tag{28}$$

we get the following hyper-parameters values $b_3 = 0.170230$ and $b_4 = 2.066578$.

8. Limiting Expressions for Complete Data Set

Suppose $T \to \infty$ i.e. all observations that are incorporated in our analysis are uncensored, and therefore r tends to n, r_1 tends to unknown n_1 and r_2 tends to unknown n_2. As a result, the amount of information contained in the sample is increasing, which consequently results in the reduction of the variances of the estimates. Following table 2, contains the limiting expressions for complete data set.

Table 2. The limiting expressions for the BE (LP, GTP) and Variance (LP, GTP) as $T \to \infty$.

Parameters	BE (LP)	BE (GTP)	Variance of BE (LP)	Variance of BE (GTP)
λ_1	$\frac{\sum_{j=1}^{n_1}\lvert x_{1j}\rvert+(0.5b_1)}{n_1-0.5}$	$\frac{\sum_{j=1}^{n_1}\lvert x_{1j}\rvert+(b_3)}{n_1}$	$\frac{\left(\sum_{j=1}^{n_1}\lvert x_{1j}\rvert+b_1\right)^2}{(n_1-0.5)^2(n_1-1.5)}$	$\frac{\left(\sum_{j=1}^{n_1}\lvert x_{1j}\rvert+b_3\right)^2}{(n_1)^2(n_1-1)}$
λ_2	$\frac{\sum_{j=1}^{n_2}\lvert x_{2j}\rvert+(0.5b_2)}{n_2-0.5}$	$\frac{\sum_{j=1}^{n_2}\lvert x_{2j}\rvert+(b_4)}{n_2}$	$\frac{\left(\sum_{j=1}^{n_2}\lvert x_{2j}\rvert+b_2\right)^2}{(n_2-0.5)^2(n_2-1.5)}$	$\frac{\left(\sum_{j=1}^{n_2}\lvert x_{2j}\rvert+b_4\right)^2}{(n_2)^2(n_2-1)}$
p	$\frac{n_1+1}{n+2}$	$\frac{n_2+1}{n+2}$	$\frac{n_1+1}{n+2}\left\{(n_1+2)-\frac{n_1+1}{n+2}\right\}$	$\frac{n_2+1}{n+2}\left\{(n_2+2)-\frac{n_2+1}{n+2}\right\}$

9. Simulation Study

A simulation study was carried out in order to scrutinize the performance of the Bayes estimators and the impact of sample size and censoring rate in the fit of the model. Samples of sizes *n=25, 100, 500* and *1000* were generated from the two component mixture of the Laplace distribution (location parameter considering zero) with parameters, λ_1, λ_2 and p such that $(\lambda_1,\lambda_2)\in\{(0.5,1),(3,4)\}$ and $p\in\{0.30, 0.40, 0.60\}$.

Probabilistic mixing was used here to generate the mixture data. For each observation a random number 'u' was generated from the uniform on *(0, 1)* distribution. If '$u<p$', the observation was taken randomly from F_1 (the Laplace distribution with parameter λ_1) and if '$u>p$', the observation was taken randomly from F_2 (the Laplace distribution with parameter λ_2).

Right censoring was carried out using a fixed censoring time *T*. All observations that were greater than *T* were declared as censored ones. The choice of the censoring time, in each case, was made in such a way that the censoring rate in resulting sample was to be approximately *15%* to *30%*. For each of the combinations of parameters, sample size, censoring rate, *5000* sample were generated. In each case, only failures were identified to be a member of either subpopulation-1 or subpopulation-2 of the mixture. For each of the *5000* samples, the Bayes estimates were computed and the average of the *5000* estimate is presented in Tables *3-7*.

Extensive tables are available upon request from the corresponding author for above mentioned mixing component values. Here we will present only tables for mixing component *p=0.60*. The general findings are: (i) the posterior risk of the estimates reduces as the sample size increases; (ii) as a result of censoring, the λ parameter and proportion parameter is over-estimated when $\lambda_1 < \lambda_2$, but when $\lambda_1 = \lambda_2$ the parameters are either over or under -estimated and proportion parameter in few cases is also under-estimated.

Table 3. BEs using LP, GTIIP and PRs in parentheses when p=0.60 under L_1.

Prior	**LP**			**GTIIP**		
λ_1, λ_2	**E(λ_1|x)**	**E(λ_2|x)**	**E(p|x)**	**E(λ_1|x)**	**E(λ_2|x)**	**E(p|x)**
n	**T=1, λ_1=0.5, λ_2=1.0**					
25	0.508386 (0.041251)	1.537960 (0.534699)	0.595883 (0.012235)	0.493542 (0.037047)	1.488890 (0.530467)	0.598389 (0.012133)
50	0.498789 (0.020989)	1.139700 (0.150044)	0.591047 (0.007016)	0.498976 (0.019663)	1.135890 (0.149454)	0.598086 (0.006942)
100	0.501151 (0.012313)	1.114850 (0.086039)	0.597290 (0.004153)	0.495988 (0.011835)	1.114700 (0.082754)	0.595962 (0.004108)
500	0.500435 (0.003019)	1.018933 (0.017143)	0.600899 (0.001021)	0.499159 (0.002982)	1.019790 (0.016957)	0.600502 (0.001015)
1000	0.500949 (0.001576)	1.005264 (0.008631)	0.600368 (0.000513)	0.500286 (0.001566)	1.005764 (0.008583)	0.600158 (0.000508)
λ_1, λ_2	**T=6, λ_1=3.0, λ_2=4.0**					
n	**E(λ_1|x)**	**E(λ_2|x)**	**E(p|x)**	**E(λ_1|x)**	**E(λ_2|x)**	**E(p|x)**
25	3.025460 (1.269692)	4.316770 (4.263897)	0.598738 (0.011155)	2.981300 (1.152665)	4.098980 (3.634163)	0.587400 (0.011154)
50	3.021990 (0.664566)	4.295840 (1.891769)	0.594805 (0.006383)	2.997084 (0.637340)	4.085688 (1.761677)	0.595008 (0.006309)
100	3.021140 (0.386853)	4.114360 (1.140642)	0.599160 (0.003680)	2.996570 (0.370708)	4.057780 (1.103921)	0.599349 (0.003686)
500	3.024388 (0.089910)	4.040885 (0.259025)	0.599961 (0.000875)	3.019689 (0.089782)	4.029402 (0.257821)	0.600020 (0.000877)
1000	3.011235 (0.046229)	4.008492 (0.131658)	0.599578 (0.000451)	3.008902 (0.046203)	4.002816 (0.131389)	0.599608 (0.000451)

Table 4. BEs using LP, GTIIP and PRs in parentheses when p=0.60 and both parameters of mixture are assumed same under L_1 .

Prior	LP			GTIIP		
λ_1, λ_2	E(λ_1|x)	E(λ_2|x)	E(p|x)	E(λ_1|x)	E(λ_2|x)	E(p|x)
n	T=1, λ_1=0.5, λ_2=0.5					
25	0.494288 (0.023977)	0.772573 (0.062169)	0.587776 (0.010124)	0.483523 (0.021553)	0.782578 (0.059785)	0.585625 (0.010078)
50	0.496126 (0.014175)	0.712744 (0.036190)	0.597338 (0.005861)	0.495255 (0.013290)	0.722999 (0.034852)	0.597117 (0.005823)
100	0.496814 (0.007438)	0.588154 (0.015296)	0.598067 (0.003131)	0.496244 (0.007169)	0.596850 (0.015110)	0.597890 (0.003118)
500	0.527938 (0.001883)	0.528540 (0.002925)	0.602188 (0.000683)	0.498213 (0.001882)	0.525737 (0.002556)	0.611344 (0.000657)
1000	0.494743 (0.000950)	0.512639 (0.001837)	0.597077 (0.000368)	0.493797 (0.000946)	0.514151 (0.001808)	0.596717 (0.000367)
λ_1, λ_2	E(λ_1|x)	E(λ_2|x)	E(p|x)	E(λ_1|x)	E(λ_2|x)	E(p|x)
n	T=1, λ_1=1.0, λ_2=1.0					
25	0.992102 (0.232703)	1.900970 (1.001083)	0.592008 (0.019009)	0.986245 (0.196151)	1.842960 (0.926908)	0.605186 (0.018754)
50	0.998467 (0.101707)	1.381690 (0.232813)	0.591908 (0.011833)	0.985518 (0.093067)	1.382730 (0.282348)	0.596406 (0.011668)
100	0.998722 (0.055439)	1.312620 (0.178549)	0.594207 (0.008170)	0.998519 (0.062438)	1.319560 (0.172021)	0.597454 (0.008088)
500	0.994885 (0.021164)	1.084057 (0.046237)	0.598047 (0.002803)	0.992617 (0.020970)	1.089610 (0.045931)	0.597881 (0.002799)
1000	0.996734 (0.012170)	1.052970 (0.026342)	0.598207 (0.001628)	0.996389 (0.012121)	1.056363 (0.026271)	0.598724 (0.001628)
λ_1, λ_2	E(λ_1|x)	E(λ_2|x)	E(p|x)	E(λ_1|x)	E(λ_2|x)	E(p|x)
n	T=6, λ_1=3.0, λ_2=3.0					
25	2.990792 (1.034421)	3.293300 (2.120675)	0.597720 (0.010246)	2.980434 (0.933457)	3.157930 (1.849078)	0.587714 (0.010210)
50	2.996737 (0.590415)	3.146315 (1.338092)	0.590049 (0.006001)	2.991655 (0.564636)	3.383260 (1.251652)	0.590193 (0.006003)
100	2.993837 (0.289712)	3.097310 (0.553571)	0.591819 (0.003152)	2.991393 (0.283712)	3.106270 (0.537189)	0.591951 (0.003147)
500	2.997033 (0.066291)	3.039416 (0.126781)	0.598213 (0.000715)	2.996579 (0.066066)	3.031977 (0.126101)	0.598275 (0.000706)
1000	2.993823 (0.034431)	3.014454 (0.064948)	0.599192 (0.000363)	2.991587 (0.034374)	3.010674 (0.064778)	0.599229 (0.000357)
λ_1, λ_2	E(λ_1|x)	E(λ_2|x)	E(p|x)	E(λ_1|x)	E(λ_2|x)	E(p|x)
n	T=6, λ_1=4.0, λ_2=4.0					
25	3.962822 (2.079020)	4.602730 (5.303277)	0.597444 (0.012291)	3.988300 (1.876063)	4.365150 (4.536665)	0.587454 (0.011469)
50	3.990070 (1.398641)	4.280470 (2.680277)	0.597333 (0.007707)	3.991566 (1.337707)	4.162480 (2.506060)	0.597865 (0.007665)
100	3.998394 (0.716908)	4.233242 (1.560137)	0.598864 (0.004528)	3.998073 (0.705243)	4.266590 (1.513910)	0.598121 (0.004525)
500	3.996185 (0.184342)	4.121362 (0.375143)	0.595869 (0.001159)	3.995674 (0.184108)	4.106422 (0.373610)	0.596093 (0.001162)
1000	3.997717 (0.095801)	4.019819 (0.190346)	0.598833 (0.000605)	3.995430 (0.095747)	4.012078 (0.189967)	0.598969 (0.000604)

Table 5. BEs using LP, GTIIP and PRs in parentheses when p=0.60 under L_2.

Prior	LP			GTIIP		
λ_1, λ_2	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$
n	**T=1, λ_1=0.5, λ_2=1.0**					
25	0.490132 (0.118653)	1.012540 (0.183820)	0.593267 (0.041993)	0.483382 (0.113511)	1.018950 (0.168256)	0.573446 (0.042000)
50	0.492714 (0.072074)	0.998844 (0.113217)	0.595690 (0.022027)	0.492162 (0.069915)	0.990951 (0.106117)	0.584231 (0.021918)
100	0.495707 (0.044268)	0.995205 (0.071402)	0.593128 (0.012067)	0.495317 (0.043395)	0.997107 (0.068291)	0.591824 (0.011987)
500	0.498817 (0.011594)	0.994444 (0.017366)	0.597510 (0.002825)	0.497651 (0.011513)	0.995728 (0.017127)	0.597129 (0.002813)
1000	0.494792 (0.006137)	0.997787 (0.008815)	0.598606 (0.001400)	0.494162 (0.006113)	0.998414 (0.008750)	0.598400 (0.001401)
λ_1, λ_2	**T=6, λ_1=3.0, λ_2=4.0**					
n	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$
25	3.364490 (0.113412)	4.192420 (0.173471)	0.594373 (0.039215)	3.291370 (0.111486)	4.582030 (0.167498)	0.584341 (0.039154)
50	3.162681 (0.067374)	4.141400 (0.109150)	0.597171 (0.018265)	3.258430 (0.067052)	4.307221 (0.107522)	0.587127 (0.019036)
100	3.078474 (0.039117)	4.058852 (0.066935)	0.596413 (0.010680)	3.176188 (0.039159)	4.154107 (0.066689)	0.596547 (0.010492)
500	2.995574 (0.009708)	3.991209 (0.016114)	0.597022 (0.002360)	3.096865 (0.009717)	4.009822 (0.016133)	0.597075 (0.002425)
1000	2.990711 (0.005018)	3.992464 (0.008308)	0.598068 (0.001226)	2.997836 (0.005019)	3.996826 (0.008304)	0.598096 (0.001236)

Table 6. BEs using LP, GTIIP and PRs in parentheses when p=0.60 under L_3.

Prior	LP			GTIIP		
λ_1, λ_2	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$
n	**T=1, λ_1=0.5, λ_2=1.0**					
25	0.464254 (0.065731)	1.240590 (0.297376)	0.596344 (0.022489)	0.472443 (0.061069	1.225070 (0.263822)	0.576153 (0.022367)
50	0.491151 (0.037638)	1.013600 (0.126104)	0.597873 (0.012208)	0.495386 (0.035900)	1.017840 (0.118404)	0.587674 (0.012211)
100	0.497825 (0.022901)	1.039420 (0.075425)	0.590250 (0.007040)	0.493755 (0.022233)	1.042260 (0.072440)	0.598884 (0.006978)
500	0.494551 (0.005884)	1.001843 (0.017091)	0.599202 (0.001696)	0.493331 (0.005828)	1.002906 (0.016884)	0.598814 (0.001688)
1000	0.497847 (0.003101)	0.996582 (0.008682)	0.599486 (0.000882)	0.497201 (0.003085)	0.997139 (0.008625)	0.599278 (0.000880)
λ_1, λ_2	**T=6, λ_1=3.0, λ_2=4.0**					
n	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$
25	2.966660 (0.358498)	3.975330 (0.778844)	0.596648 (0.020338)	2.978880 (0.339422)	4.387740 (0.711244)	0.595662 (0.020408)
50	2.981658 (0.205415)	4.052629 (0.432112)	0.593536 (0.010269)	2.997700 (0.200795)	4.234430 (0.414546)	0.598368 (0.011111)
100	2.990052 (0.120628)	4.084595 (0.268417)	0.592864 (0.006197)	2.998768 (0.119727)	4.179409 (0.263692)	0.593025 (0.006123)
500	2.994864 (0.029524)	3.997656 (0.064320)	0.598495 (0.001366)	2.990160 (0.029429)	4.019659 (0.064210)	0.598551 (0.001369)
1000	2.995942 (0.015293)	3.997553 (0.031960)	0.598824 (0.000754)	2.993603 (0.015228)	3.996875 (0.032041)	0.598852 (0.000754)

Table 7. BEs using LP, GTIIP and PRs in parentheses when p=0.60 under L_4 .

Prior	LP			GTIIP		
λ_1, λ_2	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$
n	T=1, λ_1=0.5, λ_2=1.0					
25	0.547455 (0.078139)	1.702944 (0.329969)	0.606062 (0.020359)	0.529746 (0.072408)	1.657486 (0.337192)	0.608442 (0.020107)
100	0.513289 (0.024275)	1.152792 (0.075884)	0.600756 (0.006933)	0.507778 (0.023581)	1.151221 (0.073042)	0.599399 (0.006873)
500	0.503442 (0.006015)	1.027311 (0.016755)	0.601748 (0.001698)	0.502137 (0.005956)	1.028070 (0.016561)	0.601346 (0.001689)
1000	0.502519 (0.003141)	1.009548 (0.008567)	0.600795 (0.000854)	0.501849 (0.003125)	1.010022 (0.008516)	0.600581 (0.000846)
λ_1, λ_2	T=6, λ_1=3.0, λ_2=4.0					
n	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$	$E(\lambda_1\|x)$	$E(\lambda_2\|x)$	$E(p\|x)$
25	3.228483 (0.406045)	4.785227 (0.936914)	0.607982 (0.018488)	3.168724 (0.374849)	4.520597 (0.843235)	0.596819 (0.018838)
100	3.084499 (0.126720)	4.250718 (0.272715)	0.602223 (0.006126)	3.057799 (0.122460)	4.191599 (0.267637)	0.602416 (0.006134)
500	3.039216 (0.029656)	4.072809 (0.063849)	0.600699 (0.001416)	3.034519 (0.029659)	4.061268 (0.063733)	0.600750 (0.001461)
1000	3.018901 (0.015333)	4.024881 (0.032778)	0.599954 (0.000752)	3.016569 (0.015336)	4.019194 (0.032757)	0.599984 (0.000752)

By making judgment between priors, one can see that λ is over-estimated for small sample size, but in terms of posterior risk we can observe that using the Gumbel Type-II, the posterior risk is smaller than the posterior risk using the Levy informative prior. Also note λ_1 is under-estimated in some case and also proportion. The Levy prior has smaller value of posterior as compared to the Gumbel type-II prior which is in fact due to the hyper parameters values; because the quality of Bayes (Levy and Gumbel Type-II) depends upon the quality of prior information. For using large degree of censoring we can see our posterior risk reduced for large parameters values. Based on minimum posterior risk value, L_2 is the best.

10. Real Life Application

Kanji [25] considered that wind shear is an important factor affecting the safety of aircraft during take-off and landing period. Measurements of the distribution of wind shears of particular form and magnitude encountered during the approach to lending of aircraft are needed to provide a rational choice of wind shear for: (a) assessing the effectiveness and safety of aircraft control systems, and (b) training pilots to recognize and react correctly when they encounter a wind shear. Considered under a wide variety of conditions atmospheric disturbance show both ‘order’ and ‘disorder’ trend.

There are 24 cases and each case identified with band 1 to 4. He fitted the mixture distribution and concluded that cases against Band 1 and Band 2 are appropriate for Laplace or exponential type distribution and Band 3, 4 cases are suitable for normal type distribution(s). Latter Jones and McLachlan [22] used same data set for the mixture of Laplace and normal distribution. For our purpose, we divide the data by taking mixing weight *0.375* and *T=400*. Other result summery as follow:

Data description as follows:

Variable	*N*	*Mean*	*Median*	*Tr Mean*	*StDev*	*SE Mean*	*Min*	*Max*	*Q1*	*Q3*
Case-1	*23*	*343*	*24*	*269*	*659*	*137*	*0*	*2237*	*1*	*433*
combine	*276*	*220.2*	*2.0*	*138.9*	*485.2*	*29.2*	*0.0*	*2237.0*	*0.0*	*118.0*

$$n_1 = 14,\ n_2 = 9,\ r_1 = 7,\quad r_2 = 10,\ \sum_{j=1}^{r_1} |x_{1j}| = 41,\ \ \sum_{j=1}^{r_2} |x_{2j}| = 550$$

$$n_1 = 104,\ n_2 = 172,\ r_1 = 88,\quad r_2 = 148,\ \sum_{j=1}^{r_1} |x_{1j}| = 4528,\ \ \sum_{j=1}^{r_2} |x_{2j}| = 5113$$

From table 8, it is clear that Gumbel type-II is the most suitable prior for the mixture of Laplace distribution on the basis of minimum posterior risk. However, there are some Bayes estimates which have lower posterior risk than Gumbel Type-II prior which is due to the quality of hyperparameters. Also, about the choice of loss function as concerned, the $L_2(\lambda, d) = \left(1 - \frac{d}{\lambda}\right)^2$ is the most appropriate los function.

Table 8. BEs and PRs for real data set under different loss functions.

P=0.375	**LP**			**GTIIP**		
BE	**E(λ_1|x)**	**E(λ_2|x)**	**E(p|x)**	**E(λ_1|x)**	**E(λ_2|x)**	**E(p|x)**
Loss Function	L_1					
Case-1	6.3160 (2.7083)	310.7000 (106.5690)	0.3200 (0.0915)	5.8815 (2.4055)	295.2070 (98.4023)	0.3200 (0.0915)
12 cases	234.5893 (25.2248)	34.6857 (2.8705)	0.4640 (0.0298)	233.2582 (25.0094)	34.5710 (2.8560)	0.4640 (0.0299)
Loss Function	L_2					
Case-1	4.829730 (0.117678)	256.6620 (0.086965)	0.260870 (0.105590)	4.574470 (0.111111)	246.0050 (0.083336)	0.260870 (0.105590)
12 cases	229.346438 (0.011175)	34.220243 (0.006708)	0.460131 (0.004218)	228.074055 (0.011112)	34.108736 (0.006686)	0.460132 (0.004218)
Loss Function	L_3					
Case-1	5.473700 (0.842307)	281.109000 (29.590900)	0.291667 (0.028333)	5.146280 (0.735239)	268.370000 (26.837100)	0.291667 (0.028333)
12 cases	231.938265 (2.651062)	34.451358 (0.234302)	0.462080 (0.001935)	230.637016 (2.621204)	34.338309 (0.232727)	0.462081 (0.001935)
Loss Function	L_4					
Case-1	6.526879 (0.421759)	310.871451 (0.342902)	0.440341 (0.240681)	6.082560 (0.402121)	295.373620 (0.333239)	0.440341 (0.240681)
12 cases	234.683058 (0.107515)	34.727054 (0.082708)	0.495072 (0.062143)	233.311803 (0.107205)	34.612282 (0.082563)	0.495173 (0.062345)

10.1 Graphs of Marginal Posterior distributions using Levy and GTII Prior when p=0.60.

Following Figures 2-4 show the graphical presentation of marginal posterior of λ_1, λ_2 and p .

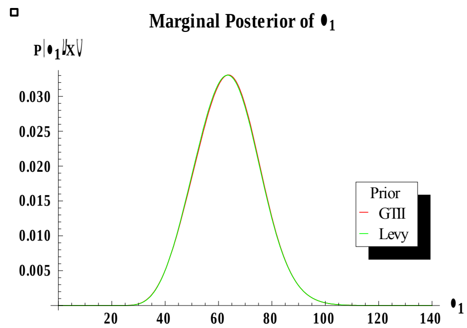

Figure 2. Marginal Posterior Density of λ_1.

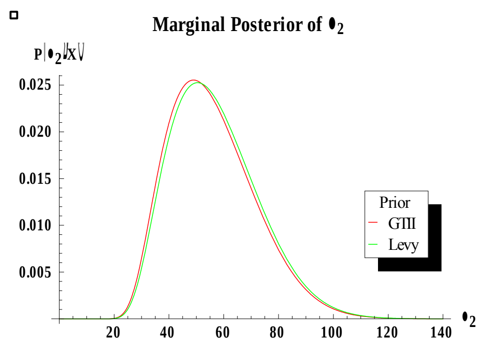

Figure 3. Marginal Posterior Density of λ_2.

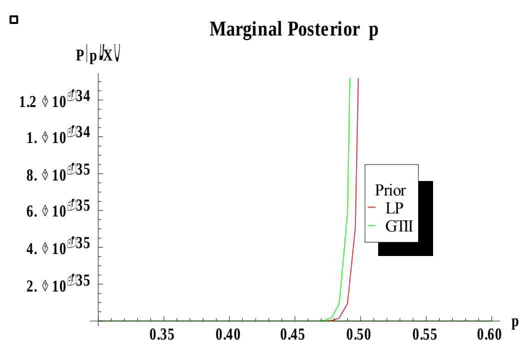

Figure 4. Marginal Posterior Density of p

From above figures 2-4, one can easily observe that Gumbel Type-II prior has heavier tail as compared to the Levy Prior, and since Laplace distribution is also heavy tailed distribution so Gumbel Type-II distribution is the most suitable prior for mixture of Laplace distribution.

11. Model Comparison

The comparison of model performances is proposed to be based on the generated posterior predictive distributions. The criterion used to compare them is based on the use of the logarithmic score as a utility function in a statistical decision framework. This was proposed by Bernardo [9] and used, for example, by Gutiérrez-Peña and Walker [18], and Martín and Pérez [27] in a similar context. In situations where the uncertainty is contained in the value of a future observation $y = x_{n+1}$, the logarithmic score $\log\left(p_k(y\,|\,\mathbf{x})\right)$ is used, where $p_k(y\,|\,\mathbf{x})$ denotes the posterior predictive density under model M_k. Then, the posterior predictive expected utility is given by: $\bar{U}_k = \int \log\left(p_k(y\,|\,\mathbf{x})\right) p_k(y\,|\,\mathbf{x})dy$. The optimal solution to the decision problem of choosing among the competing models M_0; M_1, ... , M_l is given by the model M_{k*} , such that: $\bar{U}_{k*} = \max_{k\in\{0,1,\ldots,l\}} \bar{U}_k$. From a practical viewpoint, $\bar{U}_k$ can be estimated as: $\hat{\bar{U}}_k = \frac{1}{m}\sum_{i=1}^{m} \log\left(p_k(y_i\,|\,\mathbf{x})\right)$, where y_1; y_2, ... , y_m are an independent and identically distributed random sample from $p_k(y\,|\,\mathbf{x})$. In order to illustrate this method and its applicability in the context of our proposed approaches, we considered this approach as a prior selection criterion. A random sample of sizes *20* were generated from mixture of Laplace distribution with mean *0*, mixing proportion *0.40* and scale parameter equal to *3* and *4* using Minitab v *12* (*0.60438, 0.08135, 0.53826, 0.36013, 2.38849, 0.54444, 1.12126, 3.84595, 3.04917, 3.05758, 2.18903, 1.01963, 0.30160, 4.58365, 4.30493, 5.77560, 0.58089, 0.29435, 0.63610, 1.86025*). The posterior information is used of *n=100* and we have the results: U_{Levy}=-94.88978 and U_{GTII}=-93.56026. Hence, the Gumbel Type-II prior is the best based on predictive utility criteria.

12. Conclusion and Suggestions

The simulation study has displayed various attractive properties of the Bayes estimates. The posterior risk of the estimates seems to be quite large (small) for the relatively larger (smaller) values of the parameters. However, in each case the posterior risk of parameters and effect of censoring reduces as the sample size increases. Another interesting remark concerning the posterior risk of the estimates is that increasing (decreasing) the proportion of the component in the mixture reduces (increases) the posterior risk of the estimate of the corresponding parameter. In some cases the proportion parameter is either under-estimated or over-estimated depending upon the values of the parameters or censoring degree.

Levy prior has greater posterior risk value than the Gumbel Type-II prior. The posterior risk of second component parameter is less as compared to first component parameter value, and the Gumbel Type-II prior results are the best based on its minimum posterior risks values. Also, this study suggests that at least *100* or above sample size is required for this type of mixture because for small sample size we can easily see that degree of over-estimation is large and posterior risk (variances) of the Bayes estimates is also larger. Based on above evidence (Posterior Risk, sample size effect, mixing proportion parameter, prior, the Bayes estimates, graphical presentation and model comparison method) we prefer the Gumbel Type-II prior as the most suitable prior for the mixture of Laplace distribution, and the loss function $L_2(\lambda,d)=\left(1-\frac{d}{\lambda}\right)^2$ performance is the best.

In future, this work can be extended using mixture of truncated Laplace distribution and eliciting the hyper-parameters of mixing component by taking Beta prior. Here we analyzed the two component mixture of Laplace distribution and only focus on the scale parameter but in future, a possible extension of this work is by considering mixture of more than two component mixture with location parameter.

Acknowledgement

The authors are thankful to the anonymous referees for their valuable suggestions and comments, which have helped in improving the present investigation.

References

[1]. Ali, S., Aslam, M. and Kazmi, S.M.A. (2012), On the Bayesian Analysis of the Mixture of Laplace Distribution using the Complete and the Censored Sample under different Loss Functions. *Model Assisted Statistics and Applications*, 7 (3), 209-227.

[2]. Ali, M. M. and Nadarajah, S. (2007) Information matrices for normal and Laplace mixtures. *Information Sciences*, 177, 947–955.

[3]. Aryal, G. and Rao, V. N. A., (2005), Reliability model using truncated skew-Laplace distribution, *Nonlinear Analysis*, 63 (5-7), e639-e646.

[4]. Aslam, M. (2003). An application of prior predictive distribution to elicit the prior density. *Journal of Statistical Theory and Applications*. 2 (1), 70-83.

[5]. Ayyub, B. M., (2001), *Elicitation of Expert Opinions for Uncertainty and Risks*, CRC Press, ISBN 0-8493-1087-3.

[6]. Balakrishnan, N. and Chandramouleeswaran, P. M., (1996), Reliability estimation and tolerance limits for Laplace distribution based on censored samples. *Microelectron Reliability,* 36 (3), 375-378.

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

[8]. Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, 2nd edition, Springer Series in Statistics, ISBN-10: 0-387-96098-8 and -13: 978-0387-96098-2.

[9]. Bernardo, J.M., (1979). Expected information as expected utility. *Ann. Statist*. 7, 686-690.

[10]. Bhowmick, D., Davison, A. C., Goldstein, R.D., Ruffieux, Y. (2006), A Laplace mixture model for identification of differential expression in microarray experiments. *Journal of Biostatistics*, 4 (4), 630-641.

[11]. Childs, A. and Balakrishnan, N, (2000), Conditional inference procedures for the Laplace distribution when the observed samples are progressively censored. *Metrika*, 52 (3), 253-265.

[12]. Childs, A. and Balakrishnan, N. (1997), Maximum likelihood estimation of Laplace parameters based on general Type-II censored samples. *Statistical Papers*, 38 (3), 343-349.

[13]. Childs, A. and Balakrishnan, N., (1996), Conditional inference procedures for the Laplace distribution based on Type-II right censored samples. *Statistics & Probability Letters*, 31 (1), 31-39.

[14]. Choi, D. and Nadarajah, S. (2009) Information matrix for a mixture of two Laplace distributions. *Statistical Papers*, 50, 1–12, DOI 10.1007/s00362-007-0053-8.

[15]. Dey, K. D., (2007), *Prior Elicitation from Expert Opinion*, (Lecture Notes), University of Connecticut and Some parts joint with: Junfeng Liu Case Western Reserve University.

[16]. Gajewski, J. B., Simon D. S. and Carlson, E. S., (2008), Predicting accrual in clinical trials with Bayesian posterior predictive distributions. *Statistical Methodology,* 27 (13), 2328-2340.

[17]. Gijbels, I. (2010), Censored Data, *WIREs Computational Statistics*, 2, March/April 2010, 178-188.

[18]. Gutiérrez-Peña, E. and Walker, S. G., (2001), A Bayesian predictive approach to model selection, *Journal of Statistical Planning and Inference*, 93 259-276.

[19]. Harris, C.M., (1983), On finite mixtures of geometric and negative binomial distributions. *Communication in Statistics-Theory and Method,* 12, 987-1007.

[20]. Inusah, S. and Kozubowski, J. T., (2006) A discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference*, 136, 1090 – 1102.

[21]. Jenkinson, D. (2005), *The Elicitation of Probabilities: A Review of the Statistical Literature*. Department of Probability and Statistics, University of Sheffield.

[22]. Jones, P. N. and McLachlan, J. G., (1990), Laplace-Normal mixtures fitted to wind shear data. *Journal of Applied Statistics*, 17 (2), 271-276.

[23]. Kappenman, Russell F. (1975), Conditional confidence intervals for Double Exponential distribution parameters. *Technometrics*, 17 (2), 233-235.

[24]. Kappenman, Russell F. (1977), Tolerance intervals for the Double Exponential distribution. *Journal of the American Statistical Association*, 72 (360), 908-909.

[25]. Kanji, K. G. (1985), A mixture model for wind shear data. *Journal of Applied Statistics*, 12 (1), 49-58.

[26]. Kozubowski, J. T. and Nadarajah, S., (2010), Multitude of Laplace distributions. *Statistical Papers*, 51, 127–148,DOI 10.1007/s00362-008-0127-2

[27]. Martín, J. and Pérez. C. J. (2009), Bayesian analysis of a generalized lognormal distribution. *Computational Statistics & Data Analysis*, 2009: 1377-1387.

[28]. Mendenhall, W. and Hader, R. A. (1958), Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data, *Biometrika*, 45, 1207-1212.

[29]. Nadarajah, S., (2009) Laplace random variables with application to price indices. *AStA. Adv. Stat. Anal.,* 93, 345–369, DOI 10.1007/s10182-009-0108-3

[30]. Nadarajah, S. (2004), Reliability for Laplace Distribution. *Mathematical Problems in Engineering*, 2, 169–183.

[31]. Norstrom, J. G. (1996), The use of precautionary loss functions in risk analysis, *IEEE Trans. Reliab.* 45 (1), 400-403.

[32]. Oakley, J., and O'Hagan, A. (2005). *Uncertainty in prior elicitations: a non-parametric approach*. Revised version of research report No. 521/02 Department of Probability and Statistics, University of Sheffield.
[33]. Romeu, L. J. (2004), Censored Data. *START*, 11 (3), 1-8.
[34]. Sabarinath, A. and Anilkumar, A. K., (2008), Modeling of Sunspot numbers by a modified binary mixture of Laplace distribution functions. *Solar Physics*, 250 (1), 183-197.
[35]. Saleem, M. and Aslam, M. (2008), Bayesian Analysis of the Two Component Mixture of the Rayleigh Distribution assuming the Uniform and the Jeffreys Priors. *Journal of Applied Statistical Science*, 16 (4), 105-113
[36]. Saleem, M. and Aslam, M. and Economus, P. (2010), On the Bayesian analysis of the mixture of the power distribution using the complete and censored sample. *Journal of Applied Statistics*, 37 (1), 25-40.
[37]. Scallan, A. J. (1992), Maximum likelihood estimation for a Normal/Laplace mixture distribution. *The Statistician*, 41, 227-231.