Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 4, Issue 2 (2011): 57-82

IMPROVED INFORMATIVE PRIOR FOR THE MIXTURE OF LAPLACE DISTRIBUTION UNDER DIFFERENT LOSS FUNCTIONS

Sajid Ali[a]*, Muhammad Aslam[a] and Syed Mohsin Ali Kazmi[b]
[a*, a, b] Department of Statistics, Quaid-i-Azam University Islamabad, 45320, Pakistan.
E Mail: **[a]***sajidali.qau@hotmail.com

Abstract

In this study, a new informative prior is developed for the scale parameter of the mixture of Laplace distribution when data is censored and can be used to model various real world problems. The basic proposal is to merge both informative and non-informative priors for improvement of prior information. There are many real world problems in which investigator has different opinion than prior information e.g. one doctor provides information that the harmfulness of medicine is 20% and another chemist observes the chemical combination of medicine and thinks that medicine is harmful 30% due to one element, so if we combine both doctor and chemist opinion as a prior our analysis will improve. An inclusive simulation scheme including a large number of parameter is followed to highlight properties and behavior of the estimates in terms of sample size, censoring rate and proportion of the component of the mixture. A simulated mixture data with censored observations is generated by probabilistic mixing for computational purposes. Elegant closed form expressions for the Bayes estimators and their posterior risk are derived for the censored sample as well as for the complete sample. Some interesting comparison and properties of the estimates are observed and presented. The complete sample expressions for ML estimates and for their variances are derived and also the components of the information matrix are constructed as well. The Elicitation of hyper-parameters of mixture through prior predictive approach and a real-life mixture data example has also been discussed. The Bayes estimates are evaluated under squared error loss function and precautionary loss function.

Keywords: Censored Sampling; Inverse Transformation Method; Squared Error Loss Function (SELF); Precautionary Loss Function; Hyperparameters; Prior Elicitation; Fixed Test Termination Time; Mixture Distribution; Posterior Risk; Improved Informative Prior.

2010 Classification: 62N01, 62N02 ,62N05

1. Introduction

In the last few decades, there has been a growing interest in the construction of flexible parametric classes of probability distributions in Bayesian as compared to classical. The Laplace distribution is a popular topic in probability theory due to simplicity of its characteristics function and density, 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. The application areas of the Laplace distribution are wide ranging: business firm growth, chemical physics, civil engineering, decision sciences, distribution and velocity of money, dynamics of electricity prices, dynamics of manufacturing companies, earth's magnetic field, ecosystem respiration, electronics and communications, financial data, geographical information systems, grain size distribution, human heredity, information theory, management science, neuroscience, pattern recognition, quality technology, stock returns and exchange rate changes, vision, image and signal processing.

Censoring is an important feature of the lifetime data because most of times it is not possible to continue the experiment until last observation in order to obtain a complete data set, i.e. a data set with exact life times of all objects. There are three types of censored observations, the left, the interval and the right censored observations. A right censored observation may be of Type-I or Type-II. Censoring is said to be of Type-I if the censoring time is fixed and the number of failures is random. On the other hand in Type-II censoring the number of failures in the sample is predetermined and so the time to complete the test is random. Romeu (2004) have given an account of censoring.

Mixture distributions have been used in wide range of important practical situations because these provide a powerful way to extend common parametric families of distributions to fit datasets not adequately fit by single common parametric distributions. In our daily life we often try to fit a life time model that tells us about failure of system or fit a model on data that comes from failure of units (electronic device, system components etc.). Mixture models have been used in the physical, chemical, social science, biological and other fields. As examples, Kanji (1985) described wind shear data using mixture distributions. In human exposure and risk assessment, Burmaster (1998) used mixture of lognormal models to re-analyze data set collected by the U.S. EPA for the concentration of Radon-222 in drinking water.

It needs to be mentioned here that Mixture of Laplace distribution has been considered before in literature. For example, Aryal and Rao (2005) study different characteristics of truncated from the left at zero skew-Laplace distribution and reliability of this model is compared with a two parameter Gamma model. Balakrishnan and Chandramouleeswaran (1996) present an estimator for the reliability function based on the best linear unbiased estimators (BLUEs) for the location and scale parameters of Laplace distribution based on Type-II censored samples, Bhowmick et al. (2006) recognize that Laplace model as a long-tailed alternative to the normal distribution is taken for identifying differentially expressed genes in microarray experiments, and provide an extension to asymmetric over or under-expression. Childs and Balakrishnan (2000) consider the progressively type-II right censored sample for analysis of Laplace distribution.

Mitianoudis and Stathaki (2005) explore the use of Laplacian mixture models (LMMs) to address the overcomplete blind source separation problem in the case that the source signals are very sparse. Rabbani and Vafadust (2008) use mixture of Laplace distribution for image/video denoising new algorithm with local parameters in multidimensional complex wavelet domain. Sabarinath and Anilkumar (2008) use the concept of modeling for sunspot numbers by a modified binary mixture of Laplace distribution function. Inusah and Kozubowski (2006) define a discrete analogue of the Laplace distribution. Nadarajah (2009) uses the Laplace distribution random variables with application to price indices. Kozubowski and Nadarajah (2010) motivated by recent popularity of Laplace distribution, provide a comprehensive review of the known Laplace distributions along with their properties and applications. Choi and Nadarajah

(2009) derive the information matrix for a mixture of two Laplace distributions without censoring involvement. Ali and Nadarajah (2007) drive the information matrices for normal and Laplace mixtures. Nadarajah (2004) has a discussion about the reliability of Laplace and related heavy tailed distributions while Scallan (1992) focuses on the maximum likelihood estimation for the Normal / Laplace mixture for wind shear data.

In this article, a population of certain objects is assumed to be composed of two subgroups mixed together in an unknown proportion. The random observations taken from this population are supposed to be characterized by one of the two distinct unknown members of a Laplace distribution. So two-component mixture of the Laplace distribution is recommended to model this population. Right censoring is considered and the observations greater than the fixed cut off censor value, T, are taken as censored ones. The inverse transformation method of simulation, the probabilistic mixing and the computations involved are conducted in Minitab 12.0, Mathematica 6.0, and Maple 13.0.

The Laplace mixture model is defined in Section 2 and its likelihood is developed in Section 3. In Section 4 loss functions are defined. In Section 5 and 6, the expressions for the Bayes estimates, posterior risk and predictive intervals are derived. Method of Elicitation of hyper-parameter for the mixture of Laplace distribution via prior predictive approach is discussed in Section 7. Limiting expressions of these estimates and their posterior risk are derived in Section 8. Simulation study is performed in Section 9, and a real life data is used in Section 10 for the evaluation of Bayes estimates. Some concluding remarks and further research proposal are given in Section 11.

2. The Population and the Model

A finite mixture distribution with two component densities of specified parametric form and 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 considering location parameter zero:

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

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

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

where A= λ_1=10, λ_2=15,*p*=0.30, B= λ_1=15, λ_1=15,*p*=0.40 and C= λ_1=20, λ_2=40,p=0.60, μ=20

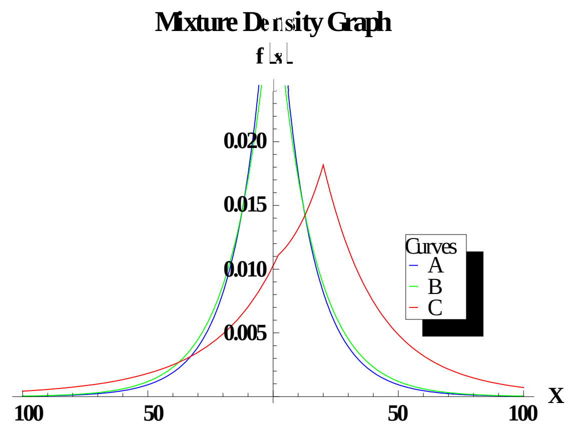

Fig. 1

When we consider the location parameter values, the graph is like bimodal distribution (Fig.1). The corresponding mixture distribution function is given as:

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

3. Sampling

Suppose n units from the above mixture model are employed to a life testing experiment with a fixed test termination 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 (1958), in many real life situations only the failed objects can easily be identified as member of either subpopulation 1 or subpopulation 2. An engineer, for example, may identify a failed electronic object as a member of the first or the second subpopulation based on the cause of its failure. So 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. Let we define, x_{ij} as the failure time of the j^{th} unit belonging to the i^{th} subpopulation, where $j = 1, 2, 3, \dots, ri,\ i = 1, 2,\ 0 < x_{1j},\ x_{2j} \le T$.

3.1 The Likelihood Function

The likelihood function for the above condition defined data is

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

where $\mathbf{x} = \left(|x_{11}|, |x_{12}|, \dots, |x_{1r_1}|, |x_{21}|, |x_{22}|, \dots, |x_{2r_2}|\right)$ are the observed failure times for the non-censored observations.

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

After a little sort of simplifications, the (2) can be represented as:

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

The generalized form of above equation can be written as:

$$L(\boldsymbol{\lambda},p|\mathbf{x})\propto\{\prod_{j=1}^{r_1}p_1\ f_1(x_{1j})\}\ \{\prod_{j=1}^{r_2}\mathrm{p}_2\ f_2(x_{2j})\}\ldots\{\prod_{j=1}^{r_k}\mathrm{p}_k\ f_k(x_{kj})\}\ \{(1-F(T))^{n-r}\}$$

$$L(\boldsymbol{\lambda},\mathbf{p}\,|\,\mathbf{x})\propto\sum_{i}^{H_{n-r}^{m}}\binom{n-r}{m_1,m_2,\ldots,m_m}\left\{p_1^{r_1+m_1}\ p_2^{r_2+m_2}\cdots p_m^{r_m+m_m}\right\}\left\{\left(\frac{1}{\lambda_1}\right)^{r_1}\left(\frac{1}{\lambda_2}\right)^{r_2}\cdots\left(\frac{1}{\lambda_m}\right)^{r_m}\right\} e^{-\frac{1}{\lambda_1}\left[\sum_{j=1}^{r_1}|x_{1j}|+(m_1)T\right]}\times e^{-\frac{1}{\lambda_2}\left[\sum_{j=1}^{r_2}|x_{2j}|+m_2T\right]}\ldots e^{-\frac{1}{\lambda_m}\left[\sum_{j=1}^{r_m}|x_{1j}|+(m_m)T\right]}$$

Here 'm' denotes the components i.e. m_1 means first component; m_2 2nd component and so on where H_{n-r}^{k} denotes the number of distinct terms in the expansion of the multinomial

$$\left(1-\left\{p_1\left\{1-0.5e^{-\frac{T}{\lambda_1}}\right\}+p_2\left\{1-0.5e^{-\frac{T}{\lambda_2}}\right\}+\ldots+p_m\left\{1-0.5e^{-\frac{T}{\lambda_m}}\right\}\right\}\right)^{n-r}$$ as

discussed in Chuan-Chong and Mhee-Meng (1992). Maximum Likelihood Estimates of $\boldsymbol{\lambda}$ and of $\mathbf{p}$ are obtained by solving the system of $2m-1$ nonlinear equations. But in this paper we will present only for two component mixture estimators and in similar way can be extended to '**m**' component.

4. Loss Functions

The Bayes estimators are evaluated under squared error loss function (SELF) and precautionary loss function (PLF). The squared error loss function (SELF) $L_1=L(\lambda,\lambda^*)=\left(\lambda-\lambda^*\right)^2$ was used by proposed by Legendre (1805) and Gauss (1810) to develop least square theory. This is 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 (1996) introduced an alternative asymmetric precautionary loss function, and also presented a general class of precautionary loss functions as a special case which is defined as $L_2 = L(\lambda, d) = \frac{(\lambda - d)^2}{d}$, Bayes estimator using this loss function is $d^* = \sqrt{E(\lambda^2 \mid \mathbf{x})}$ and $E_{\lambda|\mathbf{x}} L(\lambda, d) = 2\left(\sqrt{E(\lambda^2 \mid \mathbf{x})} - E(\lambda \mid \mathbf{x})\right)$ is posterior risk. According to him these loss functions approach infinitely near the origin to prevent underestimation, thus giving conservative estimators, especially when underestimation may lead to serious consequence. Taking expectation of each parameter with respect to its marginal distributions gives the Bayes estimator of the parameters.

5. Bayes Estimators assuming Jeffreys Inverse Gamma as a proposed 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 decline of the posterior risk of the Bayes estimates. Bansal (2007) discussed a method to evaluate the significance of a prior information in terms of the number of additional observations supposed to be added to a given sample size. There are many real world situations in which we have two types of information and we can call them prior information e.g. in medicine an experiment is conducted for its effectiveness but at the same time we know from expert that due to some specific ingredient, medicine will be effective this level. So now we have two types of information one is within data or experiment and other relevant information based on expert opinion. If we ignore any type of available information, our analysis will be distorted. Similar example can be found in other fields like reliability analysis of electric components etc. In view of above example we will use the new improved informative prior which is Jeffreys Inverse Gamma prior for analysis. Haq (2009) introduced this idea for normal distribution.

5.1 The Jeffreys Inverse Gamma prior

Since Jeffreys prior for $f(\lambda_1, \lambda_2) \propto 1/\lambda_1\lambda_2$, now suppose $\lambda_1 \sim JInvGamma(a_1, b_1)$, $\lambda_2 \sim JInvGamma\ (a_2, b_2)$ and $p \sim U(0,1)$.

Assuming independence, we have a joint prior $h(\lambda_1, \lambda_2, p) \propto \left(\frac{1}{\lambda_1}\right)^{(a_1+2)} e^{-\frac{b_1}{\lambda_1}} \left(\frac{1}{\lambda_2}\right)^{(a_2+2)} e^{-\frac{b_2}{\lambda_2}}$, which is combined with the likelihood function given in (3) to get a joint posterior distribution of λ_1, λ_2 and p. The marginal distribution of each parameter is obtained by integrating out the nuisance parameters.

5.2 Bayes Estimators using the Jeffreys Inverse Gamma Prior

The expectation of each parameter with respect to its marginal distributions gives the Bayes estimators of the parameter. The joint posterior distribution is

$$g(\lambda_1,\lambda_2,p\,|\,\mathbf{x}) \propto \sum_{m=0}^{n-r}\binom{n-r}{m}p^{(n-r_2-m)}q^{(r_2+m)}\left(\frac{1}{\lambda_1}\right)^{(D+2)}\left(\frac{1}{\lambda_2}\right)^{(E+2)}e^{-\frac{1}{\lambda_1}[B]}e^{-\frac{1}{\lambda_2}[C]} \quad (3)$$

The marginal posterior distribution of λ_1, λ_2 and p are follow

$$g(\lambda_1\,|\,\mathbf{x}) = H^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\left(\frac{1}{\lambda_1}\right)^{r_1+a_1+2}e^{-\frac{1}{\lambda_2}[B]}A\left\{\Gamma(E+1)/C^{E+1}\right\},\ 0<\lambda_1<\infty$$

$$g(\lambda_2\,|\,\mathbf{x}) = H^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}\left(\frac{1}{\lambda_2}\right)^{r_2+a_2+2}e^{-\frac{1}{\lambda_2}[C]}A\left\{\Gamma(D+1)/B^{(D+1)}\right\},0<\lambda_2<\infty$$

and

$$g(p\,|\,\mathbf{x}) = H^{-1}\sum_{m=0}^{n-r}\binom{n-r}{m}p^{(n-r_2-m)}q^{(r_2+m)}\left\{\left(\Gamma(D+1)\Gamma(E+1)\right)/\left(B^{(D+1)}C^{(E+1)}\right)\right\},0<p<1$$

where

$$A = \beta(n-r_2-m+1,r_2+m+1)\,,\ B=\sum_{j=1}^{r_1}\left|x_{1j}\right| + (n-r-m)T+b_1\,,$$

$$C=\sum_{j=1}^{r_2}\left|x_{2j}\right| + mT + b_2\,,\ D=r_1+a_1\,,\ E=r_2+a_2 \text{ and}$$

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

5.3 The Posterior Risk of the Bayes Estimators using the Jeffreys Inverse Gamma Prior

Under SELF posterior risk is equivalent to variance. The variances of the Bayes estimators of λ_1, λ_2 and p using the Jeffreys Inverse Gamma prior are given as

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

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

and

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

respectively. Similarly Bayes Estimators and Posterior Risk can be found under precautionary loss function using above equations.

5.4 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 preceding observations. Bolstad (2004) and Bansal (2007) have given a great detailed discussion about the posterior predictive distribution.

5.5 Predictive distribution intervals using the Jeffreys Inverse Gamma 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

$g(\lambda_1, \lambda_2, p \mid \mathbf{x}) \propto \sum_{m=0}^{n-r} \binom{n-r}{m} p^{(n-r_2-m)} q^{(r_2+m)} \left(\frac{1}{\lambda_1}\right)^{(D+2)} \left(\frac{1}{\lambda_2}\right)^{(E+2)} e^{-\frac{1}{\lambda_1}[B]} e^{-\frac{1}{\lambda_2}[C]}$, $-\infty < y < \infty$ is the joint posterior distribution obtained by incorporating the Inverse Gamma prior with the likelihood given by equation (3).

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

$$p(y \mid \mathbf{x}) = \frac{1}{L} \sum_{m=0}^{n-r} \binom{n-r}{m} \left[\begin{array}{l} \left\{ \beta(n-r_2-m+2, r_2+m+1) \left(\Gamma(D+2)\Gamma(E+1)/(B+|y|)^{D+2} C^{E+1} \right) \right\} + \\ \left\{ \beta(n-r_2-m+1, r_2+m+2) \left(\Gamma(D+1)\Gamma(E+2)/C^{E+2} B^{D+1} \right) \right\} \end{array} \right] \tag{4}$$

where $L = \frac{H}{2}$

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

$$\frac{\alpha}{2} = \frac{1}{L} \sum_{m=0}^{n-r} \binom{n-r}{m} \left[\begin{array}{l} \left\{ \beta(n-r_2-m+2, r_2+m+1) \frac{\Gamma(D+2)\Gamma(E+1)/C^{E+1}}{(D+1)(B+L)^{D+1}} \right\} \\ + \\ \left\{ \beta(n-r_2-m+1, r_2+m+2) \frac{\Gamma(D+1)\Gamma(E+2)}{(B)^{(D+1)} (C+L)^{(E+1)} (E+1)} \right\} \end{array} \right]$$

and

$$\frac{\alpha}{2}=\frac{1}{L}\sum_{m=0}^{n-r}\binom{n-r}{m}\left[\begin{array}{l}\left\{\beta(n-r_2-m+2,r_2+m+1)\dfrac{\Gamma(D+2)\Gamma(E+1)/C^{E+1}}{(D+1)(B+U)^{D+1}}\right\}\\ +\\ \left\{\beta(n-r_2-m+1,r_2+m+2)\dfrac{\Gamma(D+1)\Gamma(E+2)}{(B)^{(D+1)}(C+U)^{(E+1)}(E+1)}\right\}\end{array}\right]$$

respectively.

These posterior predictive intervals can be evaluated for a number of combinations of the hyper-parameters which help us to determine a range of hyper-parameters that may lead to informative Bayes estimates having smaller posterior risk than the non-informative Bayes estimates. Saleem and Aslam (2008) used predictive intervals for the Rayleigh mixture to discuss precision of Bayes estimates in terms of hyper-parameters. If a trend in terms of the hyper-parameters is observed for the narrower predictive intervals, then a sort of objectivity may be added to prior information provided by a number of experts. Since, we are using an elicitation method as defined in Section (7), so we need not to solve these intervals.

6. Bayesian Analysis using Inverse Gamma Informative Prior

Suppose $\lambda_1 \sim InvGamma(a_3,b_3)$, $\lambda_2 \sim InvGamma(a_4,b_4)$ *and* $p \sim U(0,1)$. Assuming independence, we have a joint prior $h(\lambda_1,\lambda_2,p)\propto\left(\frac{1}{\lambda_1}\right)^{(a_3+1)}e^{-\frac{b_3}{\lambda_1}}\left(\frac{1}{\lambda_2}\right)^{(a_4+1)}e^{-\frac{b_4}{\lambda_2}}$, which is combined with the likelihood function given in (3) to get a joint posterior distribution of λ_1, λ_2 and *p*. The marginal distribution of each parameter is obtained by integrating out the nuisance parameters.

6.1 Posterior Distribution using Inverse Gamma Prior

The joint posterior distribution of λ_1, λ_2 and '*p*' is as following

$$g(\lambda_1,\lambda_2,p\,|\,\mathbf{x})\propto\sum_{m=0}^{n-r}\binom{n-r}{m}p^{(n-r_2-m)}q^{(r_2+m)}\left(\frac{1}{\lambda_1}\right)^{(G+1)}\left(\frac{1}{\lambda_2}\right)^{(H+1)}e^{-\frac{1}{\lambda_1}[I]}e^{-\frac{1}{\lambda_2}[J]} \quad (5)$$

The marginal distribution of each parameter is obtained by integrating out the nuisance parameters

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

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

And

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

Where G, H, I, J, E and K are defined as

$$G = r_1 + a_3, H = r_2 + a_4, I = \sum_{j=1}^{r_1} |x_{1j}| + (n-r-m)T + b_3, J = \sum_{j=1}^{r_2} |x_{2j}| + mT + b_4,$$

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

6.2 Bayes Estimators using Inverse Gamma Prior

The expectation of each parameter with respect to its marginal distributions gives the Bayes estimators of the parameter. The Bayes estimators of λ_1, λ_2 and p assuming the Inverse Gamma prior are given as follow

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

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

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.

6.3 Posterior Risk of the Bayes Estimators using the Inverse Gamma Prior

The posterior risk of the Bayes estimators of λ_1, λ_2 and p using the Inverse Gamma prior are given as

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

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

And

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

respectively.

6.4 Predictive Distribution and Intervals using the Inverse Gamma 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

$$g(\lambda_1, \lambda_2, p \mid \mathbf{x}) \propto \sum_{m=0}^{n-r} \binom{n-r}{m} p^{(n-r_2-m)} q^{(r_2+m)} \left(\frac{1}{\lambda_1}\right)^{(G+1)} \left(\frac{1}{\lambda_2}\right)^{(H+1)} e^{-\frac{1}{\lambda_1}[I]} e^{-\frac{1}{\lambda_2}[J]}, \quad (6)$$

$-\infty < y < \infty$

is the joint posterior distribution obtained by incorporating the Inverse Gamma prior with the likelihood given by equation (3).

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

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

where $X = \frac{K}{2}$

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

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

and

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

respectively.

7. Elicitation

Elicitation is the process/technique of taking out the expert knowledge about some unknown quantity of interest, or the probability of some future event, which can then be used to supplement any numerical data that we may have. If the expert in question does not have a statistical background, as often happens, translating their beliefs into a statistical form suitable for use in our analyses can be a challenging task as described Dey (2007).

Prior elicitation is an important component of Bayesian statistics and yet to be invented. In any statistical analysis there will typically be some form of background knowledge available in addition to 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 Berger (1985), Kadane, et al. (1980), Garthwaite and Dickey (1992), Al-Awadhi and Garthwaite (1998), Oakley and O'Hagan (2005), Kass and Greenhouse (1989), Klir and Wierman (1999), Cooke (1991) and Rowe (1992), Ayyub (2001), León et al. (2003), Gaioni (2008) , Gajewsk et al. (2007), O'Hagan et al. (2006) and Jenkinson (2005)).

7.1 Hyperparameter Elicitation

Hyperparameter elicitation from the prior $g(\lambda)$ directly is conceptually difficult task because we first have to identify prior distribution and then its parameters. The harmony of opinion among researchers is now to elicit expert knowledge about hyperparameters from observable quantities 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 parameters $g(\lambda)$, yet is still a function of the unknown hyperparameters. The moments (mean, variance . . .) have 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 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. In principle, both tertiles would suffice when there are two hyperparameters to be determined, whereas the quartiles would be needed to determine three hyperparameters.

To determine (elicit) a prior density, Aslam (2003) develops some new methods base on the prior predictive distribution. In his paper, he uses prior predictive probabilities, predictive mode and confidence level for elicitating the hyperparameters. The method for elicitation of hyperparameters is developed by following the ideas of Chaloner and Duncan (1983); Kadane (1980); Kadane et al. (1980) and Winkler (1980). Using the beta-binomial as a predictive distribution and comparing it with the expert's assessment of this distribution, he selects those hyperparameters that make the difference between elicited and observed probabilities very closely. The following method of elicitation is used in this study for determining hyper-parameters of informative prior.

7.2 Method of Elicitation through Prior Predictive Probabilities

Infact prior predictive remove 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 $\xi(a_1,a_2)$ is defined in such a way that the hyperparameters $a_1 \text{ and } a_2$ are to be chosen by minimizing this function $\xi(a_1,a_2) = \min_{a_1,a_2} \sum_{n_{12}}^{r_{12}} \left\{ \frac{p(n_{12}) - p_0(n_{12})}{p(n_{12})} \right\}^2$, where $p(n_{12})$ denote the prior predictive probabilities characterized by the hyperparameters a_1 and a_2 and $p_0\left(n_{12}\right)$ denote the elicited prior predictive probabilities. If the prior predictive distribution is symmetrical then the hyperparameters a_1 and a_2 are equal (say =*c*), so above equation becomes $\xi(c) = \min_{c} \sum_{n_{12}=0}^{r_{12}} \left\{\frac{p(n_{12})-p_0(n_{12})}{p(n_{12})}\right\}^2$, where $p\left(n_{12}\right)$ and $p_0\left(n_{12}\right)$ are the symmetrical prior predictive probabilities characterized by the hyperparameters c and the elicited prior predictive probabilities respectively. Solving the above equations simultaneously by applying 'PROC SYSLIN' of the SAS package for eliciting the requires hyperparameters. We have taken prior probabilities '*0.01* and *0.09*' and initial values of hyperparameters $a_1=1$, $a_2=1$, $b_1=1$, $b_2=2$ and data points by looping concept from '*1 to 200*' because distribution is symmetric and continuous.

7.2.1 Elicitation through Prior Predictive Probabilities when prior is Jeffreys Inverse Gamma

The equation of prior predictive using the Jeffreys Inverse Gamma prior, where '*y*' be future observation,

$$p(y) = \int_0^\infty \int_0^\infty \int_0^1 p(\lambda_1,\lambda_2, p)\, p(y \mid \lambda_1, \lambda_2, p)\, d\lambda_1 d\lambda_2 dp \tag{7}$$

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

$$f(y) = \frac{1}{2}\frac{(b_1)^{a_1}(b_2)^{a_2}}{\Gamma(a_1)\,\Gamma(a_2)}\int_0^1\int_0^\infty\int_0^\infty \sum_{k=0}^{1}\binom{1}{k}\left(\frac{p}{\lambda_1}\exp\left(-\frac{|y|}{\lambda_1}\right)\right)^{1-k}\left(\frac{q}{\lambda_2}\exp\left(-\frac{|y|}{\lambda_2}\right)\right)^{k}\left(\frac{1}{\lambda_1}\right)^{(a_1+2)}\exp\left(-\frac{b_2}{\lambda_1}\right)\left(\frac{1}{\lambda_2}\right)^{(a_2+2)}\exp\left(-\frac{b_2}{\lambda_2}\right)d\lambda_1 d\lambda_2 dp$$

So

$$f(y) = \frac{1}{4}\left[\frac{a_1(b_1)^{a_1}}{\left(|y|+b_1\right)^{a_1+2}} + \frac{a_2(b_2)^{a_2}}{\left(|y|+b_2\right)^{a_2+2}}\right] \tag{8}$$

7.2.2 Elicitation of Hyper-parameters of Inverse Gamma prior

The equation of prior predictive using the Inverse Gamma prior is:

$$f(y) = \frac{1}{2}\frac{(b_3)^{a_3}}{\Gamma(a_3)}\frac{(b_4)^{a_4}}{\Gamma(a_4)}\int_0^1\int_0^\infty\int_0^\infty \sum_{k=0}^{1}\binom{1}{k}\left(\frac{p}{\lambda_1}\exp\left(-\frac{|y|}{\lambda_1}\right)\right)^{1-k}\left(\frac{q}{\lambda_2}\exp\left(-\frac{|y|}{\lambda_2}\right)\right)^{k}$$
$$\left(\frac{1}{\lambda_1}\right)^{(a_3+1)}\exp\left(-\frac{b_3}{\lambda_1}\right)\left(\frac{1}{\lambda_2}\right)^{(a_4+1)}\exp\left(-\frac{b_4}{\lambda_2}\right)d\lambda_1 d\lambda_2 dp$$

Which simplifies that

$$f(y)=\frac{1}{2}\left[0.5\times\frac{a_3(b_3)^{a_3}}{(|y|+b_3)^{a_3+1}}+0.5\times\frac{a_4(b_4)^{a_4}}{(|y|+b_4)^{a_4+1}}\right] \tag{9}$$

7.3 Elicitation of hyperparameters of Jeffreys Inverse Gamma Prior

By using the method of elicitation defined above we get the following hyperparameters values $a_1=0.269735$, $b_1=0.00003712$ and $a_2=1.015421$, $b_2=3.302469$.

8. Limiting Expressions for Complete Data Set

Assuming $T\to\infty$, the entire observations that are incorporated in our study are uncensored, and thus 'r' tends to 'n', r_1 tends to unknown n_1 and r_2 tends to unknown n_2. As a result, the quantity of information enclosed in sample is increasing, which accordingly results in the decline of the Posterior Risk of the estimates. The expressions for the complete sample ML and Bayes estimates and their Posterior Risk are simplified as given in Table 1-2. The off diagonal terms of the information matrix vanish as can be seen from the second-order derivatives of the log likelihood function given in Equations (8) - (13). This obviously results in a diagonal information matrix, which can contentedly be inverted by simply inverting the terms on the main diagonal. Further, this shows the linear independence of the ML estimates.

Parameters	BE (Inverse Gamma)	BE (Jeffreys Inverse Gamma)
λ_1	$\frac{\sum_{j=1}^{n_1}\lvert X_{1j}\rvert+(b_3)}{n_1+(a_3)-1}$	$\frac{\sum_{j=1}^{n_1}\lvert X_{1j}\rvert+(b_1)}{n_1+(a_1)}$
λ_2	$\frac{\sum_{j=1}^{n_2}\lvert X_{2j}\rvert+(b_4)}{n_2+(a_4)-1}$	$\frac{\sum_{j=1}^{n_2}\lvert X_{2j}\rvert+(b_2)}{n_2+(a_2)}$
P	$\frac{n_1+1}{n+2}$	$\frac{n_1+1}{n+2}$

Table 1: The limiting expressions for the Bayes Estimators as $T\to\infty$

Parameters	Variance of Bayes estimators (Inverse Gamma)	Variance of Bayes estimators (Jeffreys Inverse Gamma)
λ_1	$\dfrac{\left(\sum_{j=1}^{n_1}\left\|X_{1j}\right\|+b_3\right)^2}{\left(n_1+a_3-1\right)^2\left(n_1+a_3-2\right)}$	$\dfrac{\left(\sum_{j=1}^{n_1}\left\|X_{1j}\right\|+b_1\right)^2}{\left(n_1+a_1\right)^2\left(n_1+a_1-1\right)}$
λ_2	$\dfrac{\left(\sum_{j=1}^{n_2}\left\|X_{2j}\right\|+b_4\right)^2}{\left(n_2+a_4-1\right)^2\left(n_2+a_4-2\right)}$	$\dfrac{\left(\sum_{j=1}^{n_2}\left\|X_{2j}\right\|+b_2\right)^2}{\left(n_2+a_2\right)^2\left(n_2+a_2-1\right)}$
P	$\dfrac{n_1+1}{n+2}\left\{(n_1+2)-\dfrac{n_1+1}{n+2}\right\}$	$\dfrac{n_1+1}{n+2}\left\{(n_1+2)-\dfrac{n_1+1}{n+2}\right\}$

Table 2: The limiting expressions for the Variance (Posterior Risk) of the Bayes Estimators as $T \rightarrow \infty$

9. Simulation Study

A simulation study was carried out in order to scrutinize the performance of the Bayes estimators and the impact of small and large sample size and different censoring rate in the fit of the model. Samples of size *n=25, 50, 100, 200, 300, 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 $(\lambda_1,\lambda_2) \in \{(0.5,1),(2,2.5),(3,4)\}$ and $p \in \{0.30, 0.40, 0.60\}$.

Probabilistic mixing is 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 are greater than *T* were declared as censored ones. Different fixed censoring times *T* are chosen in order to evaluate the impact of censoring rate on estimates. 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* samples were generated using a routine Minitab. 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 using Mathematica 6.0 and the average of the *5000* estimate is presented in Tables 3-8.

Prior	**IGP**			**JIGP**		
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	0.594600 (0.144168)	1.125320 (0.144495)	0.348853 (0.011704)	0.497680 (0.084770)	1.060400 (0.114392)	0.337015 (0.010950)
100	0.566313 (0.046469)	1.030610 (0.035173)	0.318277 (0.004072)	0.521467 (0.038074)	1.022540 (0.032572)	0.311763 (0.003815)
500	0.515868 (0.008064)	1.002900 (0.006592)	0.305206 (0.000816)	0.506033 (0.007535)	1.002040 (0.006436)	0.303588 (0.000791)
1000	0.512492 (0.003881)	0.997859 (0.003220)	0.302700 (0.000400)	0.507508 (0.003741)	0.997463 (0.003177)	0.301875 (0.000393)
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	3.869540 (5.496760)	3.989808 (1.788972)	0.361430 (0.011695)	3.185710 (3.581152)	3.972995 (1.507573)	0.352324 (0.011488)
100	3.440090 (1.452681)	4.097440 (0.534585)	0.319976 (0.003697)	3.206120 (1.259895)	4.077720 (0.506900)	0.315034 (0.003598)
500	3.099652 (0.252159)	3.994830 (0.097285)	0.305399 (0.000746)	3.049834 (0.241551)	4.079429 (0.095572)	0.304181 (0.000735)
1000	3.071440 (0.124264)	4.031734 (0.050000)	0.302518 (0.000371)	3.045835 (0.121257)	4.031389 (0.049484)	0.301878 (0.000369)

Table 3: BE and PR using IGP, JIGP under SELF when p=0.30.

Prior	**IGP**			**JIGP**		
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	0.506015 (0.069183)	1.090880 (0.154711)	0.418679 (0.011578)	0.495236 (0.047288)	1.024150 (0.116697)	0.410525 (0.011146)
100	0.522184 (0.024946)	1.024400 (0.041975)	0.412723 (0.004194)	0.497127 (0.022160)	1.012680 (0.038679)	0.408487 (0.004070)
500	0.510100 (0.005489)	0.999371 (0.008549)	0.403654 (0.000943)	0.504201 (0.005294)	0.997907 (0.008359)	0.402498 (0.000930)
1000	0.505287 (0.002658)	1.002289 (0.004249)	0.401991 (0.000466)	0.502319 (0.002607)	1.001568 (0.004199)	0.401402 (0.000462)
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	3.209540 (2.428853)	3.990822 (1.662848)	0.427006 (0.011145)	2.988555 (1.876001)	4.130386 (1.285309)	0.425601 (0.011113)
100	3.129700 (0.731678)	3.981479 (0.548577)	0.412906 (0.003690)	3.038690 (0.696683)	4.073035 (0.519189)	0.412097 (0.003701)
500	3.066413 (0.166978)	3.996993 (0.129186)	0.403773 (0.000843)	3.045358 (0.155401)	4.039576 (0.127866)	0.403446 (0.000845)
1000	3.039902 (0.083620)	3.997885 (0.065208)	0.402313 (0.000426)	3.029078 (0.083187)	4.049799 (0.064866)	0.402129 (0.000426)

Table 4: BE and PR using IGP, JIGP under SELF when p=0.40.

Prior	IGP			JIGP		
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	0.519843 (0.039855)	1.720150 (0.524384)	0.598498 (0.011987)	0.496693 (0.033953)	1.303910 (0.321169)	0.598265 (0.012202)
100	0.501135 (0.011761)	1.152220 (0.087609)	0.595366 (0.004038)	0.490275 (0.011754)	1.080660 (0.074224)	0.596181 (0.004143)
500	0.499959 (0.002957)	1.026796 (0.017115)	0.600263 (0.001006)	0.498254 (0.002994)	1.013279 (0.016654)	0.600675 (0.001021)
1000	0.500676 (0.001559)	1.009212 (0.008614)	0.600032 (0.000507)	0.499852 (0.001570)	1.002541 (0.008511)	0.600255 (0.000530)
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	2.990113 (1.153275)	4.775470 (4.249685)	0.597983 (0.011083)	2.988830 (1.159995)	3.973453 (2.229191)	0.599758 (0.011156)
100	2.996903 (0.362921)	4.256700 (1.159805)	0.595338 (0.003651)	3.053910 (0.381154)	3.997765 (0.979646)	0.606027 (0.003632)
500	3.010219 (0.088868)	4.074771 (0.260261)	0.598862 (0.000871)	3.037340 (0.090126)	3.996125 (0.253647)	0.601903 (0.000875)
1000	3.003780 (0.045920)	4.025923 (0.131870)	0.598994 (0.000450)	3.018258 (0.046432)	3.996795 (0.130533)	0.600596 (0.000450)

Table 5: BE and PR using IGP, JIGP under SELF when p=0.60.

Prior	IGP			JIGP		
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	0.705491 (0.221781)	1.187788 (0.124936)	0.365243 (0.032779)	0.576589 (0.157819)	1.113032 (0.105264)	0.352887 (0.031744)
100	0.605953 (0.079281)	1.047535 (0.033850)	0.324611 (0.012668)	0.556778 (0.070622)	1.038345 (0.031609)	0.317822 (0.12119)
500	0.523626 (0.015515)	1.006181 (0.006562)	0.306540 (0.002668)	0.513424 (0.014782)	1.005246 (0.006413)	0.304888 (0.002599)
1000	0.516264 (0.007545)	0.999471 (0.003224)	0.303360 (0.001320)	0.511180 (0.007345)	0.999054 (0.003182)	0.302525 (0.001300)
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	4.524389 (1.309699)	4.208033 (0.436449)	0.377262 (0.031664)	3.705388 (1.039356)	4.158396 (0.370803)	0.368266 (0.031885)
100	3.645120 (0.410060)	4.162163 (0.129446)	0.325702 (0.011451)	3.396925 (0.381610)	4.139408 (0.123376)	0.320693 (0.011319)
500	3.140064 (0.080824)	4.006988 (0.024316)	0.306618 (0.002438)	3.089181 (0.078694)	4.091126 (0.023394)	0.305387 (0.002411)
1000	3.091603 (0.040325)	4.037930 (0.012392)	0.303130 (0.001225)	3.065676 (0.039681)	4.037522 (0.012265)	0.302488 (0.001221)

Table 6: BE and PR using IGP, JIGP under L_2 when p=0.30.

Prior	IGP			JIGP		
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	0.570293 (0.128556)	1.159625 (0.137490)	0.432285 (0.027211)	0.540876 (0.091279)	1.079620 (0.110941)	0.423883 (0.026716)
100	0.545547 (0.046727)	1.044687 (0.040573)	0.417773 (0.010099)	0.518937 (0.043619)	1.031600 (0.037841)	0.413439 (0.009903)
500	0.515452 (0.010704)	1.003639 (0.008536)	0.404820 (0.002333)	0.509424 (0.010446)	1.002086 (0.008359)	0.403652 (0.002307)
1000	0.507910 (0.005247)	1.004406 (0.004235)	0.402570 (0.001158)	0.504907 (0.005176)	1.003662 (0.004188)	0.401977 (0.001150)
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	3.567912 (0.716745)	4.193985 (0.406325)	0.439863 (0.025713)	3.287470 (0.597831)	4.283153 (0.305533)	0.438462 (0.025723)
100	3.244488 (0.229575)	4.049784 (0.136610)	0.417350 (0.008888)	3.151241 (0.225102)	4.136279 (0.126488)	0.416563 (0.008932)
500	3.093520 (0.054214)	4.013121 (0.032256)	0.404815 (0.002085)	3.070766 (0.050817)	4.055372 (0.031591)	0.404492 (0.002092)
1000	33.053625 (0.027446)	4.006032 (0.016294)	0.402842 (0.001058)	3.042778 (0.027401)	4.057799 (0.016001)	0.402658 (0.001059)

Table 7: BE and PR using IGP, JIGP under L_2 when p=0.40.

Prior	IGP			JIGP		
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	0.556859 (0.074032)	1.866360 (0.292420)	0.608430 (0.019864)	0.529771 (0.066155)	1.421742 (0.235664)	0.608377 (0.020225)
100	0.512735 (0.023200)	1.189630 (0.074820)	0.598747 (0.006763)	0.502119 (0.023688)	1.114473 (0.067626)	0.599645 (0.006929)
500	0.502907 (0.005897)	1.035097 (0.016601)	0.601100 (0.001675)	0.501249 (0.005991)	1.021464 (0.016369)	0.601524 (0.001698)
1000	0.502230 (0.003109)	1.013471 (0.008517)	0.600454 (0.000845)	0.501420 (0.003136)	1.006777 (0.008471)	0.600696 (0.000883)
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	3.177113 (0.374001)	5.201423 (1.889445)	0.598396 (0.000827)	3.176964 (0.376268)	4.244705 (0.542504)	0.608987 (0.018459)
100	3.056853 (0.119899)	4.390820 (0.268240)	0.598796 (0.006117)	3.115689 (0.123558)	4.118467 (0.241404)	0.609016 (0.005978)
500	3.024944 (0.029450)	4.106582 (0.063623)	0.599589 (0.001453)	3.052140 (0.029600)	4.027736 (0.063223)	0.602629 (0.001453)
1000	3.011414 (0.015268)	4.042267 (0.032689)	0.599369 (0.000751)	3.025940 (0.015364)	3.998428 (0003265)	0.600970 (0.000749)

Table 8: BE and PR using IGP, JIGP under L_2 when p=0.60.

It is instantaneous from Table 3-8, that the posterior risk of the estimates reduces as the sample size increases. As a result of censoring, the λ parameter and proportion parameter is over or under-estimated when $\lambda_1 < \lambda_2$ and proportion parameter in few cases is under-estimated. When we make comparison between informative priors, we see that λ_1 as well as proportion is under-estimated in some cases but talking in terms of posterior risk one can easily see that Jeffreys-Inverse Gamma informative prior has smaller posterior risk than Inverse Gamma informative prior. The quality of Bayes (Jeffreys Inverse Gamma and Inverse Gamma) 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-parameter value are available. With the increment of sample size our posterior risk are also reduced. For using large degree of censoring we can see that our posterior risk are reduced for large parameters value. It is to be noted here that '*T*' failure time may be in days, seconds, years or months depending upon situation under study.

10. Real Life Application

Solar activity prediction is nowadays a topic of great interest in the scientific community because the emission of solar particles and electromagnetic radiations affects not only telecommunication systems, electric power transmission lines, long-term climate variations, weather, and other ionosphere parameters but also space activities concerning operations of low-Earth-orbiting satellites. The yearly averaged sunspot numbers are taken from the data available at the NOAA Website (ftp://ftp.ngdc.noaa.gov/STP/SOLARDATA/SUNSPOTNUBERS/) from the year 1749 to 2008.

Data description as follows

Variable	N	Mean	Median	Tr Mean	StDev	SE Mean	Min	Max	Q1	Q3
Sunspot	260	53.14	43.95	49.76	43.45	2.69	0.00	260.20	16.30	77.65

The division of data using different mixing proportion and T=50 minutes as

$$p = 0.30,\ n_1 = 78,\ n_2 = 182,\ r_1 = 42,\quad r_2 = 107,\ \sum_{j=1}^{r_1}\left|X_{1j}\right| = 978.20,\ \sum_{j=1}^{r_2}\left|X_{2j}\right| = 2462.8$$

$$p=0.375,\ n_1=98,\ n_2=162,\ r_1=51,\quad r_2=98,\ \sum_{j=1}^{r_1}\left|X_{1j}\right|=1261.7,\ \sum_{j=1}^{r_2}\left|X_{2j}\right|=2179.3$$

$$p=0.40,\ n_1=104,\ n_2=156,\ r_1=58,\quad r_2=97,\ \sum_{j=1}^{r_1}\left|X_{1j}\right|=91208.9,\ \sum_{j=1}^{r_2}\left|X_{2j}\right|=2242.9$$

$$p=0.60,\ n_1=156,\ n_2=104,\ r_1=88,\quad r_2=61,\ \sum_{j=1}^{r_1}\left|X_{1j}\right|=2076.9,\ \sum_{j=1}^{r_2}\left|X_{2j}\right|=1364.1$$

BE	JIG			IG		
	E(λ_1|x)	E(λ_2|x)	E(p|x)	E(λ_1|x)	E(λ_2|x)	E(p|x)
0.30	66.328328 (547.8318)	57.075954 (94.56082)	0.303658 (0.0053)	71.288099 (581.3665)	56.317492 (96.1258)	0.314234 (0.0052)
0.375	79.965755 (416.6888)	50.748247 (108.3146)	0.403599 (0.0057)	79.480973 (438.5733)	50.760257 (106.7806)	0.407383 (0.0054)
0.40	47.670545 (224.0768)	60.207503 (102.2374)	0.384504 (0.0045)	50.327143 (243.0507)	59.572086 (105.0853)	0.373065 (0.0046)
0.60	66.140057 (136.7402)	50.522844 (215.8391)	0.627000 (0.0047)	65.947835 (138.2687)	52.695842 (224.4549)	0.620670 (0.0048)

Table 9: BEs and PR using JIG and IG under L_1

BE	JIG			IG		
	E(λ_1|x)	E(λ_2|x)	E(p|x)	E(λ_1|x)	E(λ_2|x)	E(p|x)
0.30	70.336896 (8.017135)	57.898405 (1.644903)	0.312232 (0.017149)	75.255296 (7.934393)	57.164550 (1.694115)	0.322433 (0.016398)
0.375	82.530060 (5.128609)	51.804431 (2.112368)	0.410592 (0.013981)	82.193664 (5.425381)	51.801392 (2.082271)	0.413943 (0.013120)
0.40	49.965565 (4.590040)	61.050641 (1.686276)	0.390354 (0.011700)	52.686545 (4.718803)	60.447652 (1.751133)	0.379151 (0.012172)
0.60	67.165819 (2.051525)	52.615557 (4.185426)	0.630790 (0.007580)	66.987952 (2.080233)	54.784183 (4.176682)	0.624476 (0.007612)

Table 10: BEs and PR using JIG and IG under L_2

From the abovesaid examples, one can infer that Jeffreys-Inverse Gamma prior results are better than Inverse Gamma prior, although there are very few cases in which Inverse Gamma is better like mixing proportion parameter against *p=0.30* and *0.375* but overall Jeffreys-Inverse Gamma results are stable. Using large mixing component parameter, results are more accurate than smaller mixing component parameter.

10.1 Graphical representation of Marginal Posterior distributions using various Priors

Figures 2-5 show the Graphical representation of marginal posterior of λ_1, λ_2 and p . The graphical representation is well-matched with our numerical values. Jeffreys prior and Uniform prior behave approximately in the same way with the minor difference. Similar conclusions can be drawn for Inverse Gamma and Jeffreys Inverse Gamma priors.

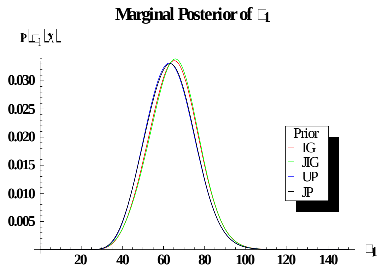

Fig.2

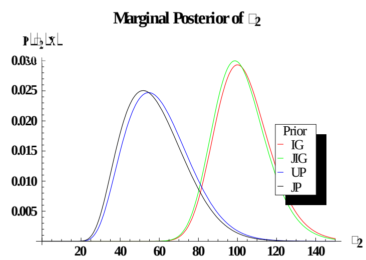

Fig.3

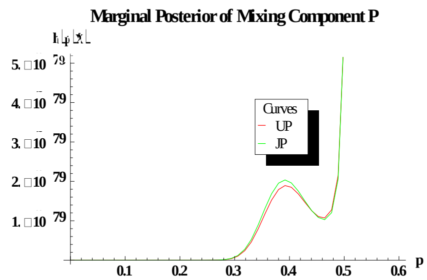

Fig. 4

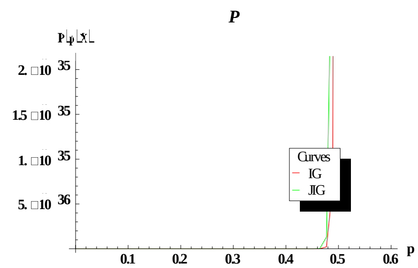

Fig. 5

10.2 Testing of Hypothesis using the marginal posterior of Mixture distribution

For Bayes Factor conclusion Jeffreys (1961) define the following rules.

i. BF>1, H_1 is supported.

ii. $10^{-0.5} \le BF < 1$, Minimal evidence against H_1.

iii. $10^{-1} \le BF < 10^{-0.5}$, Substantial evidence against H_1.

iv. $10^{-2} \le BF < 10^{-1}$, Strong evidence against H_1.

v. $BF < 10^{-2}$, Decisive evidence against H_1.

Using abovesaid rules, from the Table 11 we can easily see some interesting results. According to Jeffreys rule of Bayes Factor (BF), in case of $\lambda_1 \le 45.0$ using IG or JIG prior, we have strong evidence against H_1. Similarly in case of $\lambda_1 \le 75.0$ using IG and JIG prior, we support H_1. For $\lambda_2 \le 45.0$ we have decisive evidence and for $\lambda_2 \le 55.0$ we have substantial evidence against H_1. Similarly, in case of mixing component ‘p’, we have decisive evidence against H_1. So on the basis of testing, we prefer to use JIG as a prior information.

Table 11: Hypothesis Testing

Null Hypothesis H_1	Alternative Hypothesis H_2	Prior Distribution		Posterior Probability $P(H_1)$	Posterior Probability $P(H_2)$	Bayes Factor B
$\lambda_1 \le 45.0$	$\lambda_1 > 45.0$	IP	IG	0.0431903	0.9568097	0.0451399
			JIG	0.0431726	0.9568274	0.0451206
$\lambda_2 \le 45.0$	$\lambda_2 > 45.0$	IP	IG	0.000121313	0.9998787	0.000121328
			JIG	0.000164619	0.9998354	0.000164646
$p \le 0.50$	$p > 0.50$	IP	IG	0	1	0
			JIG	0	1	0
$\lambda_1 \le 75.0$	$\lambda_1 > 75.0$	IP	IG	0.800588	0.199412	4.014743
			JIG	0.797352	0.202648	3.934665
$\lambda_2 \le 55.0$	$\lambda_2 > 55.0$	IP	IG	0.0775849	0.9224151	0.0841106
			JIG	0.107972	0.892028	0.1210410
$p \le 0.90$	$p > 0.90$	IP	IG	0.0000000012	0.9999999	0.0000000012
			JIG	0.0000000012	0.9999999	0.0000000012

11. Conclusion and Suggestions

A new technique has been introduced to merge the prior information before we have data and with the information contained in data about the unknown parameter, in order to make a new informative prior which performs better than other simple informative and non-informative priors. Here failure time which is denoted by T depends upon the situation under study, so if we consider genes in microarray experiments data it can be in days etc. The simulation study has displayed some interesting properties of the Bayes estimates. The posterior risks of the parameters estimate seem to be quite large (small) for the relatively larger (smaller) values of the parameters in case of Jeffreys-Inverse-Gamma prior but when we used Inverse Gamma prior this theme is reverse. However, in each case the posterior risk of parameters are reduced as the sample size increases in both loss functions.

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. The effect of censoring on the estimates of parameters is in the form over-estimation. To be more specific, larger degree of censoring results in large size of over-estimation. However, as we increase the sample size the effect of censoring reduces. On the other hand, in some cases the proportion parameter is either under-estimated or over-estimated depending upon the values of the parameters or censoring degree.

Particularly, the proportion parameter is over-estimated (some degree under-estimated) whenever the parameter of the first subgroup is smaller (greater) than the parameter of the second subgroup. Also, the extent of over or under-estimation is more intensive for larger parameter values of the proportion parameter. Furthermore, increasing the sample size reduces the posterior risk of the estimate of the proportion parameter. The increase in proportion of a component in the mixture does not guaranty the reduction in variance of estimate of the proportion parameter. Contrary, the posterior risk of the estimate of population proportion is slightly increased for large values of the proportion parameter. As a cut off censor value tends to infinity, the complete sample expressions for the estimators and posterior risk are greatly simplified which will cause further reducing as there is no more effect of censoring.

Both of the estimates of the second parameter are over or under-estimated depending on the sample size but the size of under or over-estimation is greater in case of Bayes (IG) but again with JIG decreased posterior risk. The Bayes (JIG) estimates are much closer to the corresponding parameter value in case of second parameter. Also both estimates of the mixing proportion parameter are over-estimated but the degree of over-estimation is quite smaller in case of Bayes (JIG). The Bayes estimates with proposed informative (Jeffreys-Inverse Gamma) prior seem to be more efficient than existing noninformative (UP, JP which tables can be obtained from the corresponding authors) informative (Inverse Gamma) counterparts with a few exceptions in terms of under or over-estimation of Bayes estimate. The posterior risk of Jeffreys-Inverse Gamma informative prior reduce than the Inverse Gamma informative prior and both noninformative priors. One thing which can be commonly observed is that as we increase sample size the posterior risk decreases and increasing the degree of censoring does not guarantee that the risk will decrease. Also comparing noninformative priors, we can see that Jeffreys prior have smaller posterior risk in both loss function, however our risk increased using precautionary loss function but underestimation prevented.

In the real life examples, Jeffreys-Inverse Gamma prior posterior risk is less than the Inverse Gamma prior. Also this study suggests that at least *50* 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 Bayes estimates is also large.

In future this work can be extended using other informative priors based on above idea and using mixture of truncated Laplace distribution, considering location parameter and eliciting the hyperparameters of mixing component by taking Beta prior.

Acknowledgement

We are indebted to Dr Peter M. Lee, University of York, for his valuable suggestions in this paper. We are also thankful to the reviewers for their potential comments on the article that lead to a significant improvement therein.

References

1. Ali, M. M. and Nadarajah, S. (2007). Information matrices for normal and Laplace mixtures. Information Sciences, 177, p. 947–955.
2. Al-Awadhi, S.A. & Garthwaite, P.H. (1998). An elicitation method for multivariate normal distributions, Communications in Statistics: Theory and Methods, 27, p. 1123–1142.
3. Aryal, G. and Rao, V. N. A. (2005). Reliability model using truncated skew-Laplace distribution, Nonlinear Analysis, 63(5-7), p. 639-e646.
4. Aslam, M. (2003). An application of prior predictive distribution to elicit the prior density, Journal of Statistical Theory and Applications. 2(1), p. 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 Reliab. 36(3), p. 375-378.
7. Bansal, A. K., (2007). Bayesian Parametric Inference, Narosa Publishing House Pvt. Ltd., New Delhi.
8. Bhowmick, D., Davison, A. C., Goldstein R.D. and Ruffieux, Y. (2006). A Laplace mixture model for identification of differential expression in microarray experiments, Journal of Biostatistics, 4 (4), p. 630-641.
9. Berger O. James (1985). Statistical Decision Theory and Bayesian Analysis, 2nd edition, Springer Series in Statistics, ISBN-13: 978-0387-96098-2.
10. Bolstad, W. M. (2004). Introduction to Bayesian Statistics, John Willey & Sons, Inc., New Jersey.
11. Burmaster, D.E. and Thompson, K.M. (1998). Fitted second-order parametric distributions to data using maximum likelihood estimation, Human and Ecological Risk Assessment, 3(2), p. 235-255.
12. Chaloner, K. M. and Duncan, G. T. (1983). Assesment of Beta Prior Distribution: PM elicitation. The Statistician, 32, p. 174-180.
13. Childs, A. and Balakrishnan, N, (2000). Conditional inference procedures for the Laplace distribution when the observed samples are progressively censored.**,** Metrika, 52(3), p. 253-265.
14. Choi, D. and Nadarajah, S. (2009). Information matrix for a mixture of two Laplace distributions. Stat Papers, 50, p. 1–12, DOI 10.1007/s00362-007-0053-8.
15. Chuan-Chong, C. and Mhee-Meng, K. (1992). Principles and Techniques in Combinatorics, World Scientific Publishing Co. Pvt. Ltd. Singapore.
16. Cooke, R.M., (1991). Experts in Uncertainty, Oxford University Press, New York.
17. 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.
18. Gaioni, E., Dey, K. D. and Grigoriu, M. (2008). Semiparametric Functional Estimation Using Quantile-Based Prior Elicitation, samsi Technical Report #2008-6, June 23, 2008, Statistical and Applied Mathematical Sciences Institute PO Box 14006 ,Research Triangle Park, NC 27709-4006, www.samsi.info.

19. Gajewski, J. B., Simon D. S. and Carlson, E. S. (2007). Predicting accrual in clinical trials with Bayesian posterior predictive distributions. Statist. Med. Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/sim.3128
20. Garthwaite, H.P., and Dickey, M.J., (1992). Elicitation of Prior Distributions for Variable-Selection Problems in Regression, The Annals of Statistics, 20(4), p. 1697-1719.
21. Gauss, C. F. (1810), Method des Moindres Carres Memoire sur la Combination des Observations, Translated by J. Bertrand (1955). Mallet-Bachelier, Paris.
22. Haq, A. (2009), New Improved Informative Priors for Variance of Normal Distribution. Available at: http://interstat.statjournals.net/YEAR/2009/abstracts/0911003.php
23. Inusah, S. and Kozubowski, J. T. (2006). A discrete analogue of the Laplace distribution, Journal of Statistical Planning and Inference, 136, p.1090 – 1102.
24. Jeffreys, H. (1961). Theory of Probability (3rd edt.), Oxford University Press.
25. Jenkinson, D. (2005). The Elicitation of Probabilities: A Review of the Statistical Literature, Department of Probability and Statistics, University of Sheffield.
26. Jones, N. P. and Mclachlan, J. G. (1990). Laplace-normal mixtures fitted to wind shear data, Journal of Applied Statistics, 17(2), p.271-276.
27. Kadane, J. B. (1980). Predictive and Structural methods for eliciting prior distributions, In Bayesian Analysis in Econometrics and Statistics, Zellner, A. (Ed), Amesterdam: North- Holland, p. 89-93.
28. Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. and Peter, S. C. (1980). Interactive elicitation of opinion for a normal linear model, Journal of the American Statistical Association. 75, p. 845-854.
29. Kass, R.E., Greenhouse, J.B., (1989). Comment on Investigating therapies of potentially great benefit: ECMO, by Ware, J.H., (1989), Stat. Sci. 4, p. 310–317.
30. Kanji, K. G. (1985). A mixture model for wind shear data, Journal of Applied Statistics, 12(1), p. 49-58.
31. Klir, G.J. and Wierman, M.J., (1999). Uncertainty-based information: elements of generalized information theory, Studies in Fuzziness and Soft Computing, Physica-Verlag, New York.
32. Kozubowski, J. T. and Nadarajah, S. (2010). Multitude of Laplace distributions, Stat Papers, 51, p. 127–148, DOI 10.1007/s00362-008-0127-2.
33. Legendre, A. (1805). Nouvelles Methodes Pour la Determination des Orbites des Cometes, Courcier, Paris.
34. León, J. C., Vázquez -Polo ,J. F. and González, L.R. (2003). Environmental and Resource Economics Vol. 26 p.. 199–210.© 2003 Kluwer Academic Publishers. Printed in the Netherlands.
35. Mendenhall, W. and Hader, R. A. (1958). Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data, Biometrika, 45, p. 1207-1212.

36. Mitianoudis, N. and Stathaki, T., (2005). Overcomplete source separation using Laplacian mixture models. IEEE Signal Process. Letters, 12(4), p. 277-280. Digital Object Identifier 10.1109/LSP.2005.843759.
37. Nadarajah, S., (2009). Laplace random variables with application to price indices. AStA Adv Stat Anal. 93, p. 345–369. DOI 10.1007/s10182-009-0108-3
38. Nadarajah, S. (2004). Reliability for Laplace Distribution. Mathematical Problems in Engineering, 2, p. 169–183.
39. Norstrom, J. G. (1996). The use of precautionary loss functions in risk analysis, IEEE Trans. Reliab, 45(1), p. 400-403.
40. 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.
41. O' Hagan, A., Buck, E. C., Daneshkhah, A., Eiser, J. E., Garthwaite, H.P., Jenkinson, J. D., Jeremy E. Oakley, E. J., and Rakow, T. (2006). Uncertain Judgments: Eliciting Experts' Probabilities. John Wiley & Sons, Ltd. ISBN-13: 978-0-470-02999-2 (HB),ISBN-10: 0-470-02999-4 (HB).
42. Rabbani, H., and Vafadust, M. (2008). Image/ video denoising based on a mixture of Laplace distributions with local parameters in multidimensional complex wavelet domain, Signal Processing, 88(1), p. 158-173.
43. Romeu, L. J. (2004). Censored Data. START, 11(3), p. 1-8.
44. Rowe, G., (1992). Perspectives on expertise in aggregation of judgments, in Wright, G., and Bolger, F. (Eds.), Expertise and Decision Support, Plenum Press, New York, p. 155–180.
45. Sabarinath, A. and Anilkumar, A. K., (2008). Modeling of Sunspot numbers by a modified binary mixture of Laplace distribution functions. Solar Phys, 250 (1), p. 183-197.
46. Saleem, M. and Aslam, M. (2008). Bayesian analysis of the two component Mixture of the Rayleigh distribution assuming the Uniform and Jeffreys priors. Journal of Applied Statistical Science, 16 (4), p. 105-113.
47. Scallan, A. J. (1992). Maximum likelihood estimation for a Normal/Laplace mixture distribution. The Statistician, 41, p. 227-231.
48. Winkler, R. L. (1980). Prior information, predictive distributions, and Bayesian model-building, In Bayesian Analysis in Econometrics and Statistics, A. Zellner (Ed.). North Holland, Amsterdam.