BAYESIAN ANALYSIS OF MIXTURE DISTRIBUTIONS

By

Muhammad Saleem

DEPARTMENT OF STATISTICS
QUAID-I-AZAM UNIVERSITY
ISLAMABAD, PAKISTAN
2010

BAYESIAN ANALYSIS OF MIXTURE DISTRIBUTIONS

By

Muhammad Saleem

Supervised by

Prof. Dr. Muhammad Aslam

A Thesis Submitted in the Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
IN
STATISTICS

DEPARTMENT OF STATISTICS
QUAID-I-AZAM UNIVERSITY
ISLAMABAD, PAKISTAN

2010

To my parents

especially my sweet mother (Aappaa Jee) Maryam Bibi (late),

may Allah rest her soul in peace.

ACKNOWLEDGEMENTS

The completion of this dissertation is simply because of the grace of Allah Almighty who granted me with the spirit and enthusiasm required for this adventurous task.
I am highly indebted to the Higher Education Commission (HEC) for granting me Indigenous PhD Scholarship and the Education Department, Government of the Punjab, Pakistan, for the grant of study leave for this project.

This outcome owes a great deal to a number of great names. The supreme gratitude is due to Dr. Muhammad Aslam, Professor and Formar Chairman, Department of Statistics, Quaid-i-Azam University, Islamabad, whose glorious patronage, encouragement and guidance made this dream come true. I extend my profound thanks to Dr. Javed Shabbir (Chairman) for his cooperation and moral support throughout my stay at the campus.

Many thanks are due to Dr. Ioannis Ntzoufras for his invitation and for making me learn MCMC simulation techniques, using WinBUGS and R, during my six months stay at Athens University of Economics and Business, Athens, Greece financed by International Research Support Initiative Program (IRSIP) of Higher Education Commission (HEC) of Pakistan. I am thankful once again to Dr. Ioannis Ntzoufras for encouraging me to participate and present in the "First Athens-Pavia Meeting On Statistics" (held at Mati, Attica, Greece from June 4 to June 6, 2008) which was published afterwards in the Pakistan journal of Statistics and is a part of this thesis.

Dr. Polichronas Economou was another pleasant finding of the "Mati Conference" and a coauthor of my first impact factor publication in the Journal of Applied Statistics which is also a part of this thesis. Many thanks are due to Economou's girl friend Vasia for being my great well wisher.

Dr. Munir Ahmad, founding president of ISOSS, must be saluted for his lifetime struggle for promoting research culture and establishing a pioneer research platform in the field of Statistics in Pakistan.

Thanks are due to Muhammad Yaseen and Muhammad Arif of Agriculture University Faisalabad, Khurshid Mir of Srinagar, Ahmar Mahmood, Zaheer, Naser, Qamar, Madad Khan, Sarwar Saeed, Mazhar, Umar, Amanullah Dar and Tariq Javed of Mathematics Department, QAU, Islamabad for their support and encouragement during my stay at QAU.

I am highly grateful to Dr. M. Ahsanullah, Rider University, USA, and Syed Ejaz Ahmad Shah, University of Windsor, Canada who always encouraged me whenever I contacted them to seek their help.

Sincere gratitude is due to Syed Maqbool Hassan Shah, Manzoor Hussain, and Mukkaram Idrees who always supported me at many critical moments of my career and hence ensured my survival.

Anonymous refrees of various journals have been a source of learning with many respects and helped a lot to improve this thesis, in content and value. I would especially like to mention the Pakistan Journal of Statistics, Journal of applied Statistics and Journal of Statistical Computation and Simulation.

Finally, I owe a lot to my beloved wife, Muniba Baig, and our loving kids Sanabil, Abdul Raafay and Maaham who suffered a lot during this whole work but displayed patience and provided me with the best moral support. I highly value and acknowledge Sana's love which always had a soothing effect on my nerves during the accomplishment of this task.

Date: ____________ Signature________________

(Muhammad Saleem)

DECLARATION

I Muhammad Saleem S/O Muhammad Tufail, Registration No. 1448-ST/Ph.D-2004, a student of Doctor of Philosophy at Quaid-i-Azam University, Islamabad, Pakistan, do hereby solemnly declare that the thesis entitled “Bayesian Analysis of Mixture Distributions” submitted by me in partial fulfillment of the requirements for PhD degree in Statistics, is my original work and has not been submitted and shall not, in future, be submitted by me for obtaining any degree from this or any other University or Institution. And that to the best of my knowledge and belief, this thesis contains no material previously published except where due reference is made in the text of the thesis.

Date: ____________ Signature________________

(Muhammad Saleem)

APPROVAL SHEET

The thesis entitled "Bayesian Analysis of Mixture Distributions" proposed and submitted by Mr. Muhammad Saleem in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Statistics is hereby accepted. We accept this dissertation as it fulfills the conditions and standards required by the Quaid-i-Azam University, Islamabad.

____

Prof. Dr. Muhammad Aslam
(Supervisor)

_____________________ _____________________

External Examiner **External Examiner**

____

Dr. Javed Shabbir
(Chairman)

____

Dean Faculty of Sciences **Date:___________________**

ABSTRACT

In this thesis, we consider type-I mixtures of the members of a subclass of one parameter exponential family. This subclass includes Exponential, Rayleigh, Pareto, a Burr type XII and Power function distributions. Except the Exponential, mixtures of distributions of this subclass get either no or least attention in literature so far.

The elegant closed form expressions for the Bayes estimators of the parameters of each of these mixtures are presented along with their variances assuming uninformative and informative priors. The proposed informative Bayes estimators emerge advantageous in terms of their least standard errors. An extensive simulation study is conducted for each of these mixtures to highlight the properties and comparison of the proposed Bayes estimators in terms of sample sizes, censoring rates, mixing proportions and different combinations of the parameters of the component densities. A type-IV sample consisting of ordinary type-I, right censored observations is considered. Bayesian analysis of the real life mixture data sets is conducted as an application of each mixture and some interesting observations and comparisons have been observed.

The systems of non-linear equations to evaluate the classical maximum likelihood estimates, the components of the information matrices, complete sample expressions, the posterior predictive distributions and the equations for the evaluation of the Bayesian predictive intervals are derived for each of these mixtures as relevant algebra. The predictive intervals are evaluated in case of the Rayleigh mixture only for a number of combinations of the hyperparameters to look for a trend among the hyperparameters that may lead towards an efficient estimation.

CONTENTS

CHAPTER 1

INTRODUCTION AND REVIEW OF LITERATURE

There are many practical situations where an underlying population is believed to consist of a finite number of categories such that the individual object, on which the observation X is made, belongs to one of these categories. In many practical situations, the use of the finite mixture model becomes inevitable when data are not available for each component or conditional distribution separately rather for the overall mixture distribution. The use of mixture models in such situations is called the direct application of the mixture models. On the other hand, the indirect applications include, for instance, the kernel based density estimation, the identification of contaminants in a data, cluster analysis, modeling prior densities and random variable generation. In this thesis we have considered the direct application of mixture models.

A random variable (X) is said to follow a finite mixture distribution if it has a probability density function of the form

$$p(x) = \sum_{j=1}^{k} \pi_j f_j(x|\theta_j) \ ; \ \pi_j > 0,\ j = 1, 2, ..., k,\ \sum_{j=1}^{k} \pi_j = 1,\ f_j(.) \geq 0,\ \int f_j(x)\ dx = 1$$

The parameters, π_j 's, are mixing weights, $f_j(x|\theta_j)$ is the j^th component density of the mixture and θ_j is the parameter of the j^th component density. The conditional probability of X given that the observation actually belongs to the category j is summarized by $f_j(x|\theta_j)$ and π_j is the probability that the observation belongs to the category j .

The history of methodology for the finite mixture models dates back to 19^{th} century when Pearson (1894) made use of method of moments to estimate the five parameters of a two-component normal mixture. The literature of finite mixture models has grown rapidly since the advent of computers and the recent powerful computational techniques. Finite mixtures have applications in a large number of areas e.g., econometrics, sociology, engineering, reliability estimation, quality control, electrophoresis, switching regression, medical diagnosis and prognosis, multi target signal environment, remote sensing, clustering and discriminant analysis.

The first comprehensive monographs on finite mixtures include Everitt and Hand (1981) and Titterington (1985) and McLachlan and Basford (1988). There is an extensive bibliography in Titterington (1985). The Theory, geometry and applications of mixture models are presented by Lindsay (1995). Mi (1999) underlined the age-smooth properties of mixture models. Böhning (2000) presents computer assisted analysis of mixtures and applications. Finite mixture models are focused by Geoffrey (2000) while a nice account of finite mixtures is given by Schnatter (2006) with a touch of Markov switching models. Böhning et al. (1998), Böhning and Seidel (2003) and Böhning et al. (2007) highlights the advances in Mixture Models. A detailed discussion of medical applications of finite mixture models can be seen in Schlattmann (2009).

Deciding on the number of categories (components of the mixture), if it is not known, is a difficult problem and relatively a little work has been done to this effect so far. The work done so far can be divided into the informal graphical methods and the formal hypothesis testing. The informal graphical methods include the study of the sample histogram. There are two limitations of graphical methods. First, unimodality of the underlying data does not imply that it is not a mixture. Secondly, bimodality of the sample histogram is not necessarily an indication of a mixture data having two components. Titterington (1985) has given an account of bitangentiality which is a less dramatic departure from unimodality. Bimodality implies bitangentiality but bitangentiality does not imply bimodality. He has declared bimodality an extra hump and bitangentiality an extra bump. The formal hypothesis testing includes the generalized likelihood ratio test but, unfortunately, is not that straightforward. Solow (1994) analyze the Bayesian estimation of the number of species in a community. Dietz and Böhning (1997) analyze two-component mixture models with one completely or partly known component. Zheng and Frey (2004) discuss evaluation of sample size, mixing weights, and separation between components.

The identifiability problem for finite mixtures is extensively studied by a number of authors. Teicher (1960, 1961, 1963 and 1967) emphasizes that before attempting the parameter estimation in finite mixture distributions, one must make sure that the class of distributions considered is identifiable. He reviews the conditions of identifiability in several senses i.e., T-identifiability, P-identifiability and B-identifiability. Yakowitz and Spragins (1968) discuss that the mixtures of the members of the exponential family of distributions are identifiable. In this thesis, the mixtures of a subclass of the distributions of the exponential family are considered. Chandra (1977) is among the others who contributed to the problem of identifiability of mixture density functions.

As quoted by Soegiarso (1992), four types of mixture samples are confronted with in real life applications. Type-I sample consists of observations which are not labeled. In Type-II samples, the observations are completely labeled and no censoring. The Type-III samples comprise unlabeled observations which are completely labeled subsequently. The Type-IV samples consist of unlabeled observations, of which some are labeled afterwards but rest of them are labeled due to censoring. The real life illustration, we are dealing with in our thesis, considers a Type-IV sample. The Type-IV samples are most frequently encountered in real life applications.

Li (1983) and Li and Sedransk (1982 and 1988) focus various aspects of mixture distributions and quote two types of mixture models. A type-I mixture is defined as the mixture of probability density functions from the same family. While a mixture of density functions from several families is called a type-II mixture. Rachev and SenGupta (1993) and Scallan, A.J. (1992) are two examples of type-II mixtures. In some practical situations, a mixture population may have the known component densities and we need to infer about only the mixing weights. On the other hand in many real life applications, there are known functional form of the component densities with unknown parameters but the mixing weights may be considered known. But in most of the applications the functional form and the number of the component densities are known but with unknown parameters of the component densities as well as unknown mixing weights. However, the component densities may or may not belong to the same parametric family. In this thesis, type-I mixture models with unknown parameters of the known number of component densities belonging to the same parametric family and with unknown mixing weights are considered. Some practical applications are discussed and analyzed with a touch of novelty.

Censoring is an unavoidable feature of the most of the lifetime applications and is a form of missing data problem. An account of censoring can be seen in Sinha (1986), Leemis (1995), Dietz et al. (1996), Klein and Moeschberger (1997) and Kalbfleisch and Prentice (2002) and Smith (2002) which are valuable contributions on survival analysis techniques for censored and truncated data. Raqab (1992) discussed predictors of future order statistics from type-II censored samples. Jiang and Kececioglu (1992) deals with maximum likelihood estimates using censored data for mixed Weibull distributions while Wang and Li (2005) considers the estimators for survival function when censoring times are known. Saleem and Aslam (2009) has considers a Rayleigh distributed random censoring time. Censoring is divided into three kinds, i.e., left, right and interval censoring. Right censoring may be of type-I, type-II or random right censoring. Type-I censoring is further divided into ordinary type-I, progressive type-I and generalized type-I censoring while the type-II sampling is categorized as ordinary type-II and progressive type-II censoring. An observation on lifetime of an object is said to be censored one if the exact life length of that object is unknown. Censoring is called right (left) if the unknown life length lies on the right (left) of the end (start) of study. The random right censoring is said to be employed if lifetime of an object is greater than an independent random number. In type-I (type-II) right censoring the life-test termination time (the number of dead objects) is pre-specified. In ordinary (progressive) type-I right censoring the life-test termination time is the same (different) for all the objects. However, in generalized type-I right censoring objects enter a life test at different time points while the life-test termination time is fixed. The lifetime of an object is called interval censored if it is known to fall between a known time-interval. Interval censoring has applications in industrial life-time experiments where objects are inspected periodically. In ordinary (progressive) type-II right censoring the life-test terminates after a (series of) fixed number of deaths occurs in a single phase (a series of phases). In this thesis, an ordinary type-I, right censoring is considered with a fixed life-test termination time.

Posterior distribution is the basis of Bayesian Inference. Posterior distribution is obtained when prior information is combined with the likelihood. Therefore, the prior information is a necessary part of Bayesian inference. Prior information often represents purely the subjective assessment of an expert before any data have been observed. In Bayesian analysis, specification, elicitation and formulation of a prior distribution is often difficult, so the use of conventional priors reflecting little or no information is recommended. These priors are called uninformative, indifference, ignorance, vague or reference priors. The uninformative priors are supposed to yield the Bayes estimates that are approximately as precise as maximum likelihood estimates. An uninformative prior is usually improper i.e. it does not have a proper density function but the resulting posterior distribution is a proper density function. Uninformative priors are often formal and impersonal simply to enable the theory to begin with. Burger (1985) declares the Bayesian analysis assuming the uninformative prior is likely to yield a sensible answer for a given investment of effort. Uniform and Jeffreys are the most commonly used uninformative priors. A nice discussion on Uniform prior can be seen in Jeffreys (1961). Kass (1989) declares the Jeffreys prior a uniform measure in information metric. Bernado (1979) states that when there is no nuisance parameter then the Jeffreys prior is an appropriate reference prior.

An informative prior carry fairly precise, specific, definite and scientific information about the unknown parameter of interest. The conjugate priors being compatible with the likelihood are often used as informative priors. Ghosh et al. (2006) states an advantage of the conjugate prior in terms of the calculation of posterior quantities in closed form. Parameters of the informative prior distributions are called hyperparameters, to distinguish them from the parameters of the data generating model. As the formal elicitation of the hyperparameters, is not a focus of this thesis, so appropriate values of the hyperparameters are assumed to accomplish computations wherever required.

Titterington et al. (1985) compile the scope and frequency of applications compiled from a wide variety of fields and point out the overwhelming preponderance of normal mixtures. However, binomial, poisson, gamma, exponential, lognormal and von-mises mixtures are also encountered. Al-Hussaini et al. (1997) makes parametric and nonparametric estimation for finite mixtures of lognormal components. Wiper et al. (2001) work with mixture of Gamma distributions and its applications. Sultan et al. (2006) discusses the mixture of two inverse Weibull distributions in terms properties and estimation. They also indicate that many of these applications concern univariate mixtures and very often only two components are involved.

The maximum likelihood and moments methods are the mostly used estimation methods in the mixture distributions. Redner and Walker (1984) have given a note on maximum likelihood estimation and the EM Algorithm while Arcidiacono and Jones (2003) estimated finite mixture distributions using Sequential Likelihood and the EM Algorithm. However, the Bayesian estimation in connection with the finite mixture models is paid relatively little attention in literature so far. History of Bayesian analysis of mixture distributions is not that old. Sinha (1998) and Nobile (1998) are among the first who went for the Bayesian estimation of the finite mixture models.

The focus of this thesis is the mixtures of distributions of a subclass of one parameter exponential family. This subclass includes Exponential, Rayleigh, Pareto, Burr type-XII, and Power Function distributions. Except the Exponential, Bayesian analysis of type-I mixtures of distributions of this subclass get either no or least attention in literature so far. The contribution in this thesis to the Bayesian analysis of the Power function mixture is published in the Journal of applied statistics. The work on Rayleigh mixture is already published in the Journal of Applied Statistical Science and Pakistan Journal of Statistics. While the works on the Burr and Pareto mixture are under review by the International Statistical Review and Metrika respectively.

The elegant closed form Bayes estimators are derived for each of these mixtures assuming informative and uninformative priors. Also an extensive simulation study is conducted to highlight some interesting properties and comparison of the Bayes estimates using type-I, right censored mixture data. Also, some real life examples are presented in a novel way. Although the presentation of the maximum likelihood estimates of the mixture parameters is not a focus of this thesis, yet the systems of non linear equations which are required for the evaluation of maximum likelihood estimates and the elements of the negative Hessian matrix are derived as a related algebra.

Chapter 2 consists mainly of the comprehensive simulation study of the Exponential mixture to highlight properties and comparison of Bayes estimates, derived parallel to Sinha (1998) assuming uninformative and informative priors. A real life application is presented as well. Chapter 3 presents Bayes estimators assuming uninformative and informative priors for a two component mixture of the Rayleigh distribution. A simulation study is conducted to highlight the properties of the said Bayes estimates along with a real life application. The work done in this chapter partially appears in Saleem and Aslam (2008a) and Saleem and Aslam (2008b).

Chapter 4 deals with a comparison among a number of available conjugate priors based on predictive intervals. The contents of Chapter 4 are published in Saleem and Aslam

(2008b). Another important member of the Exponential family is one parameter Pareto distribution. The predictive intervals are derived and are evaluated extensively in Chapter 4. Chapter 5 includes an account of Pareto mixture in terms of derivations of Bayes estimators and their properties based on a comprehensive simulation study along with a real life application. The work done in this chapter is under review of Metrika.

A study has been conducted in Chapter 6 of a Burr Type-XII mixture and has been submitted for possible publication in the International Statistical Review. We are the first to contribute to the area of mixture of Power Function distribution. The contents of Chapter 7 include details and some nice features of Power Function mixtures are published in Saleem and Aslam (2010). Overall conclusion is summarized in Chapter 8 along with some recommendations for the possible future extensions of this work.

CHAPTER 2

PROPERTIES AND COMPARISON OF THE BAYES ESTIMATES OF THE EXPONENTIAL FINITE MIXTURE PAREMETERS

2.1 Introduction

Exponential distribution has applications in life testing of the objects which do not age with time and have constant hazard rate. If a population of certain objects is assumed to be composed of two subgroups mixed together in an unknown proportion and the observations taken from this population are supposed to be characterized by one of the two distinct unknown members of an Exponential distribution, then the two component mixture of the Exponential mixture is recommended to model such a population provided the data is not available on the individual components rather on the mixture only. Ahsanullah (1988) worked on record values of exponentially distributed random variables. McCullagh and Peter (1994) combine Exponential mixtures and quadratic Exponential families. Sinha (1998) has compared the ML estimates and the Bayes estimates assuming the Jeffreys prior for the Mendenhall and Hader (1958) mixture model. Ebrahimi (2001) focused on the mixing fraction of the Exponential mixture. Raqab and Ahsanullah (2001) estimated location and scale parameters of generalized Exponential distribution based on order statistic. Hebert and Scariano (2004) compare the location estimators for Exponential mixtures under Pitman's measure of closeness. Ali et al. (2005) discussed the Bayes estimators of the Exponential distribution. Gosh and Ahmed et al. (2005) have discussed robustness of the ML estimation of the Exponential parameters. Elsherpieny (2007) estimated parameters of mixed generalized exponentially distributions from censored type-I samples while Abu-Taleb et al. (2007) considered Bayesian estimation of exponential survival time using exponential censor time.

The Exponential distribution, because of its memory-less property, is used for the life-testing of the products that do not age with time. There are several electronic devices whose failure rate does not depend upon their age and, therefore, the Exponential distribution is considered. In this chapter, the said Exponential mixture is defined in Section 2.2. The likelihood specific expressions are developed in Section 2.3. The ML estimates and the components of the information matrix are derived in Section 2.3. In Section 2.4 and Section 2.5 the expressions for the Bayes estimators assuming uninformative (the Uniform and the Jeffreys) and the informative (Inverted Gamma) priors, are derived along with their variances. The complete sample expressions, as the test termination time tends to infinity, of the said estimators and variances are derived in Section 2.6. A comprehensive simulation study is conducted in Section 2.7 to compare and highlight the properties of the Bayes estimates in terms of sample sizes, censoring rates and parameter points. Type-IV samples are simulated with ordinary type-I, right censoring from a two component type-I, Exponential mixture. A real life application is discussed in Section 2.8. Some concluding remarks are given in Section 2.9.

2.2 The Exponential Finite Mixture Model

A finite mixture density function with a known number ($k > 1$), of component densities of specified parametric form (exponential) but with k unknown parameters, $\theta_1 < \theta_2 < \ldots < \theta_k$ depending upon their cause of death and k unknown mixing weights, $0 < \pi_i < 1,\ i = 1, 2, \ldots, k$ where $\pi_k = 1 - \sum_{i=1}^{k-1} \pi_i$ is defined as follows.

$$f(x) = \sum_{i=1}^{k} \pi_i \ f_i(x) \tag{2.1}$$

The following exponential distribution is assumed for the k components of the mixture.

$$f_i(x)=\frac{1}{\theta_i}\exp(-x/\theta_i),\ i=1,2,...,k\ ;\ j=1,2,3,...,r_i;\ 0<\theta_i<\infty\ ;\ 0\le x_{ij}<\infty$$

The corresponding mixture survival function is given by $S(x)=\sum_{i=1}^{k}\pi_i\ e^{-x/\theta_i}$.

2.3 The Maximum Likelihood Estimates for Censored Data

The sampling scheme includes a Type-IV sample of size n units from the mixture model described above under ordinary type-I, right censoring. Let a minor inspection of the dead objects shows that $r_i,\ i=1,2,\ldots,k$ of the r failed objects are failed because of the 1st, 2nd,…, kth cause of death respectively such that $r=\sum_{i=1}^{k}r_i$ and the remaining $n-r$ objects are still alive and hence cannot be labeled because of censoring. We define, x_{ij} as the failure time of the jth unit belonging to the ith subpopulation, where $j=1,2,3,\ldots,r_i$; $i=1,2,\ldots,k$ and $0<x_{1j},x_{2j},\ldots,x_{kj}\le T$. So the likelihood function for this censored sample is as under.

$$L(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})\propto\{\prod_{j=1}^{r_1}\pi_1 f_1(x_{1j})\}\ \{\prod_{j=1}^{r_2}\pi_2 f_2(x_{2j})\}\ldots\{\prod_{j=1}^{r_k}\pi_{\mathrm{k}} f_k(x_{kj})\}\ \{(S(T))^{n-r}\} \qquad (2.2)$$

$\mathbf{x}=(\mathbf{x_1},\mathbf{x_2},\ldots,\mathbf{x}_k)=(x_{11},x_{12},\ldots,x_{1r_1},x_{21},x_{22},\ldots,x_{2r_2},\ldots,x_{k1},x_{k2},\ldots,x_{kr_k})$ is data while the $2k-1$ parameters are $\boldsymbol{\theta}=(\theta_1,\theta_2,\ldots,\theta_k)$ and $\boldsymbol{\pi}=(\pi_1,\pi_2,\ldots,\pi_k),\ \pi_k\ =\ 1-\sum_{i=1}^{k-1}\pi_i$.

$$L(\boldsymbol{\theta},\boldsymbol{\pi}|\mathbf{x})\propto(\prod_{i=1}^{k}\pi_i^{r_i})(\prod_{i=1}^{k}\theta_i^{-r_i})\{\sum_{i=1}^{k}(\pi_i\ e^{-T\theta_i^{-1}})\}\exp[\sum_{i=1}^{k}\{\theta_i^{-1}(\sum_{j=1}^{r_i}x_{ij})\}] \qquad (2.3)$$

After some manipulation, the likelihood function in (2.3) takes the form as follows.

$$L(\boldsymbol{\theta},\boldsymbol{\pi}|\mathbf{x})\propto\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^{k}}\binom{n-r}{k_1,k_2,\ldots,k_k}(\prod_{i=1}^{k}\pi_i^{r_i+k_i})(\prod_{i=1}^{k}\theta_i^{-r_i})\exp\left[-\sum_{i=1}^{k}\{\theta_1^{-1}(\sum_{j=1}^{r_1}x_{1j}+k_1T)\}\right](2.4)$$

$H^k_{n-r} = \binom{n-r+k-1}{n-r}$ denotes the number of all *k-ary* sequences $(k_1, k_2, \ldots, k_k)$ of non-negative integers with $\sum_{i=1}^{k} k_i = n-r$ and $\binom{n-r}{k_1, k_2, \ldots, k_k} = \frac{(n-r)!}{k_1!k_2!\ldots k_k!}$ in the expansion of the multinomial $(\sum_{i=1}^{k} \pi_i T^{-\alpha_i})^{n\text{-}r}$ as discussed in Chuan-Chong and Mhee-Meng (1992). The ML estimates are obtained by simultaneously solving the k equations obtained by setting the first order derivatives of the natural log of the likelihood (2.3) with respect to $\theta_1,\ \theta_2, \ldots, \theta_k,\ \pi_1,\ \pi_2, \ldots,\ \pi_{k-1}$ to zero.

$$\frac{-r_i}{\theta_i} + \frac{\sum_{j=1}^{r_i} x_{ij}}{\theta_i^2} + \frac{(n-r)\pi_i T \exp(-T/\theta_i)}{\theta_i^2 \left\{ \sum_{i=1}^{k} \pi_i \exp(-T/\theta_i) \right\}} = r_i \ , \ i = 1, 2, \ldots, k \tag{2.5}$$

$$\frac{r_i}{\pi_i} - \frac{r_k}{\pi_k} + \frac{(n-r)\left\{ \exp(-T/\theta_i) - \exp(-T/\theta_k) \right\}}{\left\{ \sum_{i=1}^{k} \pi_i \exp(-T/\theta_i) \right\}} = r_k,\ i = 1, 2, \ldots, k-1 \tag{2.6}$$

It is not possible to solve analytically the above system of $2k-1$ non-linear equations. However, they can be solved by the numerical iterative procedure. Let $\boldsymbol{\theta} = (\theta_1,\ \theta_2, \ldots,\ \theta_k,\ \pi_1,\ \pi_2, \ldots,\ \pi_{k-1})$ and it is a well known result that $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta},\ \mathrm{I}^{-1}(\boldsymbol{\theta}))$ asymptotically. Here $\mathbf{I}(\boldsymbol{\theta})$ is a symmetric matrix of order $2k-1$ as given bellow.

$$\mathbf{I}(\boldsymbol{\theta}) = -E\left(\frac{\partial^2 l}{\partial \boldsymbol{\theta}\ \partial \boldsymbol{\theta}'}\right) \tag{2.7}$$

Matrix (2.7) is the information matrix of order $2k-1$, inverting it we can find the variances of ML estimates on the main diagonal. Following are the elements of the symmetric information matrix.

$$E(\frac{-\partial^2 l}{\partial\theta_i^2}) = \frac{r_i}{\theta_i^2} + \frac{(\prod_{i=1}^{k}\pi_i)e^{-(\sum_{i=1}^{k}T/\theta_i)}(T-2\theta_i) + 2\pi_i^{\,2}\theta_i e^{-2T/\theta_i}}{T^{-1}(n-r)^{-1}\theta_i^4(\sum_{i=1}^{k}\pi_i e^{-T/\theta i})^2},\ i = 1, 2, \ldots, k. \tag{2.8}$$

$$E(\frac{-\partial^2 l}{\partial\pi_i^2}) = \frac{r_i}{\pi_i^2} - \frac{r_k}{\pi_k^2} + \frac{(n-r)(e^{-T/\theta_i} - e^{-T/\theta_k})^2}{(\sum_{i=1}^{k}\pi_i e^{-T/\theta i})^2},\ i = 1, 2, \ldots, k-1. \tag{2.9}$$

$$E(\frac{-\partial^2 l}{\partial\theta_i\partial\theta_j}) = E(\frac{-\partial^2 l}{\partial\theta_j\partial\theta_i}) = \frac{-(n-r)\ T^2(\prod_{i=1}^{k}\pi_i)\ e^{-(\sum_{i=1}^{k}T/\theta_i)}}{\theta_i^2\ \ \theta_j^2\ \ (\sum_{i=1}^{k}\pi_i e^{-T/\theta i})^2},\ j > i = 1, 2, \ldots, k. \tag{2.10}$$

$$E(\frac{-\partial^2 l}{\partial\theta_i\partial\pi_j}) = E(\frac{-\partial^2 l}{\partial\theta_j\partial\pi_i}) = -(n-r)\ \theta_i^{-2}T\ e^{-(\sum_{i=1}^{k}T/\theta_i)}\ (\sum_{i=1}^{k}\pi_i e^{-T/\theta i})^{-2},$$
$$i = 1, 2, \ldots, k;\ \ j = 1, 2, \ldots, k-1. \tag{2.11}$$

$$E(\frac{-\partial^2 l}{\partial\pi_i\partial\pi_j}) = E(\frac{-\partial^2 l}{\partial\pi_j\partial\pi_i}) = -\frac{r_k}{\pi_k^2} - \frac{(n-r)\ (e^{-T/\theta_i} - e^{-T/\theta_k})}{(e^{-T/\theta_j} - e^{-T/\theta_k})^{-1}(\sum_{i=1}^{k}\pi_i e^{-T/\theta i})^2},$$
$$i = 1, 2, \ldots, k-2,\ j = i+1 = 2, \ldots, k-1. \tag{2.12}$$

The computation of the above elements of Information matrix can be conducted using Mathematica software after replacing the parameters by their estimates obtained by iterative numerical solution of equations (2.5)-(2.6) .

2.4 The Posterior Distributions assuming the Conjugate Prior

Having a look on the likelihood, Inverted Gamma prior is considered as an informative prior for the unknown component density parameters of the mixture.

2.4.1 The Posterior Distributions assuming the Inverted Gamma Prior

For the parameters of the exponential component densities given in Section 2, the inverted gamma priors are assumed as follows. $g_i\left(\theta_i\right) \propto (1/\theta_i)^{m_i+1} e^{-s_i/\theta_i}$, $\theta_i > 0,\ m_i, s_i > 0,\ i = 1, 2, \ldots, k$, and $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k) \sim Dirichlet(1,1,\ldots,1)$. Assuming independence, the joint prior is incorporated with the likelihood (2.3) to have the following joint posterior.

$$g_{IG}(\boldsymbol{\theta}, \boldsymbol{\pi}|\mathbf{x}) = \Omega_{IG}^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1, k_2, \ldots, k_k} \{\prod_{i=1}^{n} \pi_i^{r_i+k_i}\} \{\prod_{i=1}^{n} \theta_i^{-(r_i+m_i+1)}\} \exp\left\{-\sum_{i=1}^{k} \{\theta_i^{-1}(s_i + \sum_{j=1}^{r_i} x_{ij} + k_i T)\right\}$$

where $\Omega_{IG} = \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \dfrac{\prod_{i=1}^{k}\Gamma(r_i+m_i)}{\prod_{i=1}^{k}\left\{s_i + \sum_{j=1}^{r_i} x_{ij} + k_i T\right\}^{r_i+m_i}}$.

Marginal posterior distributions of $\theta_i,\ i = 1, 2, \ldots, k$ are obtained by integrating out the nuisance parameters.

$$\begin{aligned} g_{IGi}(\theta_i|\mathbf{x}) &= \Omega_{IG}^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \\ &\times \left\{\prod_{j\neq i}^{k} \frac{\Gamma(r_i+m_i)}{(A_i+s_i)^{r_i+m_i}}\right\} \theta_i^{-(r_i+m_i+1)} \exp\{-\theta_i^{-1}(A_i+s_i)\},\ 0<\theta_i<\infty,\ i=1,2,\ldots,k. \end{aligned}$$

The marginal posterior distributions of $\pi_i,\ i = 1, 2, \ldots, k-1$ are obtained on the same lines. The expectations of $\theta_i,\ i = 1, 2, \ldots, k$ and of $\pi_i,\ i = 1, 2, \ldots, k-1$ with respect to the respective marginal distributions give the Bayes estimators under the squared error loss function. The Bayes estimators assuming the inverted Gamma prior are derived as follows.

$$\hat{\theta}_i = \Omega_{IG} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots r_k+k_k+1) \frac{\Gamma(r_i+m_i-1)}{(A_i+s_i)^{r_i+m_i-1}} \prod_{j\neq i} \frac{\Gamma(r_j+m_j)}{(A_j+s_j)^{r_j+m_j}},\ i=1,2,\ldots,k$$

$$\hat{\pi}_i = \Omega_{IG} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_i+k_i+2,\ldots,r_k+k_k+1) \prod_{j=1}^{k} \frac{\Gamma(r_j+m_j)}{(A_j+s_j)^{r_j+m_j}},\ i=1,2,\ldots,k-1.$$

2.4.2 Variances of Bayes Estimators assuming Inverted Gamma Prior

The expressions for the variances of the Bayes estimators are derived as follows.

$$V(\hat{\theta}_i) = \Omega_{IG} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \frac{\Gamma(r_i+m_i-2)}{(A_i+s_i)^{r_i+m_i-2}} \prod_{j\neq i} \frac{\Gamma(r_j+m_j)}{(A_j+s_j)^{r_j+m_j}} - \\ -\left\{ \Omega_{IG} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \frac{\Gamma(r_i+m_i-1)}{(A_i+s_i)^{r_i+m_i-1}} \prod_{j\neq i} \frac{\Gamma(r_j+m_j)}{(A_j+s_j)^{r_j+m_j}} \right\}^2,\ i=1,2,\ldots k.$$

$$V(\hat{\pi}_i) = \Omega_{IG} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_i+k_i+3,\ldots,r_k+k_k+1) \prod_{j=1}^{k} \frac{\Gamma(r_1+m_1)}{(A_1+s_1)^{r_1+m_1}} - \\ -\left\{ \Omega_{IG} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_i+k_i+2,\ldots,r_k+k_k+1) \prod_{j=1}^{k} \frac{\Gamma(r_j+m_j)}{(A_j+s_j)^{r_j+m_j}} \right\}^2,\ i=1,2,\ldots,k-1.$$

where $A_i = \sum_{j=1}^{ri} x_{ij} + k_i T,\ i=1,2,\ldots k$; $\sum_{i=1}^{k} r_i = r$ and $\sum_{i=1}^{k} k_i = n-r$ **.**

2.5 Bayes Estimators assuming Uninformative Priors

The Uniform and the Jeffreys priors are two common examples of uninformative priors which materialize the use of Bayesian estimation methods when no prior information is available. The uninformative priors are supposed to produce estimates that are approximately as precise as ML estimates if the mode of the posterior distribution is used as a Bayes estimate.

2.5.1 Bayes Estimators assuming the Uniform Prior

Let us assume a state of ignorance i.e., $\theta_i,\ \ i=1,2,\ldots,k$ is distributed uniformly over $(0,\infty)$

and $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k) \sim Dirichlet(1,1,\ldots,1)$. Hence $f_i(\theta_i) \propto k_i$, $i = 1,2,\ldots,k$. Assuming independence we have an improper joint prior that is proportional to a constant and is incorporated with the likelihood (2.4) to yield a proper joint posterior distribution. The arithmetic means of the parameters with respect to the respective marginal distributions yield the Bayes estimators under the squared error loss function. These estimators can be had by replacing $m_i = -1,\ s_i = 0$ in equations of Section 2.4. The expressions for the variances are obtained on the same lines.

2.5.2 Bayes Estimators assuming the Jeffreys Prior

Let $g(\theta_i) \propto \sqrt{|I(\theta_i)|}$, $I(\theta_i) = -E[\frac{\partial^2 f_i(x|\theta_i)}{\partial \theta_i^2}]$ $\forall$ $f_i(x|\theta_i)$, $i = 1,2,\ldots,k$. For the parameters of the exponential component densities given in Section 2, the Jeffreys prior of θ_i is $g_i(\theta_i) \propto 1/\theta_i$, $i = 1,2,\ldots,k$ as used by Sinha (1998) while as in Section 2.5.1, we assume $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k) \sim Dirichlet(1,1,\ldots,1)$. Assuming independence, the joint prior is incorporated with the likelihood (2.4) to have the joint posterior. The expectations of the parameters with respect to the respective marginal distributions give the Bayes estimators under the squared error loss function. These Bayes estimators and their variances are obtained by replacing $m_i = 0,\ s_i = 0$ in the expressions given in Section 2.4.

2.6 The Complete Sample Expressions

When the sample is uncensored, T tends to ∞, r tends to n, r_i tends to n_i, $i = 1,2,\ldots,k$, and consequently all the observations are incorporated into our analysis. Therefore, the amount of information contained in the sample is increased. The expressions for the Bayes estimators and their variances are simplified as given in Tables 2.1-2.2 for $k = 2$. It is immediate that the efficiency of the estimates is increased as well because of inclusion of all the observations in our sample. This is clear from the second order derivatives of the log likelihood given in Section 2.3 that the off diagonal terms of the information matrix vanish and ensure the linear independence of the ML estimates. The information matrix becomes a diagonal matrix which can be inverted very comfortably by simply inverting the terms on the main diagonal.

Table 2.1 The complete sample expressions for the Bayes and ML estimators as $T \to \infty$

Parameters	Bayes Estimators (Uniform)	Bayes Estimators (Jeffreys)	ML Estimators
θ_1	$\frac{\sum x_{1j}}{n_1 - 2}$	$\frac{\sum x_{1j}}{n_1 - 1}$	$\frac{\sum x_{1j}}{n_1}$
θ_2	$\frac{\sum x_{2j}}{n_2 - 2}$	$\frac{\sum x_{2j}}{n_2 - 1}$	$\frac{\sum x_{2j}}{n_2}$
π_1	$\frac{n_1 + 1}{n + 2}$	$\frac{n_1 + 1}{n + 2}$	$\frac{n_1}{n_1 + n_2}$

Table 2.2 The complete sample expressions for the variances of the Bayes and ML estimators as $T \to \infty$

Parameters	Variances of Bayes Estimators (Uniform prior)	Variances of Bayes Estimators (Jeffreys prior)	Estimated Variances of ML Estimators
θ_1	$\frac{\left(\sum x_{1j}\right)^2}{(n_1 - 3)(n_1 - 2)^2}$	$\frac{\left(\sum x_{1j}\right)^2}{(n_1 - 1)^2 (n_1 - 2)}$	$\frac{(\sum x_{1j})^2}{n_1^2}$
θ_2	$\frac{\left(\sum x_{2j}\right)^2}{(n_2 - 3)(n_2 - 2)^2}$	$\frac{\left(\sum x_{2j}\right)^2}{(n_2 - 1)^2 (n_2 - 2)}$	$\frac{(\sum x_{2j})^2}{n_2^2}$
π_1	$\frac{(n_1 + 1)(n_2 + 1)}{(n + 2)^2 (n + 3)}$	$\frac{(n_1 + 1)(n_2 + 1)}{(n + 2)^2 (n + 3)}$	$\frac{n_1 n_2}{(n_1 + n_2)(n_2 - n_1)}$

2.7 A Simulation Study

A simulations study is carried out in order to investigate the performance of the Bayes estimators in terms of sample size, censoring rate and parameter size for $k = 2$. Samples of sizes $n = 50,\ 100,\ 200,\ 300$ were generated from the two component mixture of Exponential distribution with various combinations of the parameters θ_1, θ_2 and π_1 such that $(\theta_1, \theta_2) \in \{(15,20),(20,25),(24,30),(32,40)\}$ and $\pi_1 \in \{0.40, 0.60\}$. Probabilistic mixing was used to generate the mixture data. For each observation a random number u was generated from the uniform distribution on [0, 1]. If $u < \pi_1$, the observation was taken randomly from the Exponential distribution with parameter θ_1 otherwise it is from the exponential distribution with parameter θ_2. Ordinary, type-I, right-censoring is carried out using a fixed test termination time T. The choice of the censoring time is made in such a way that the censoring rate in the resulting sample to be approximately 15% and 30%. For each of the 64 combination of parameters, sample size, censoring rate, one thousand samples were generated using a routine in Minitab software. In each case failures are identified to be a member of either Subpopulation-1 or Subpopulation-2 of the mixture. For each of the 1000 samples, the Bayes estimates were computed using a routine in Mathematica software. The results are presented in Tables 2.3-2.7. Tables 2.3-2.6 display some interesting properties of the Bayes estimates in terms of sample sizes, censoring rates, size of life time parameters and mixing proportion parameters. Table 2.7 presents an interesting comparison between the three Bayes estimates. The properties and the comparison observed are summarized in Section 2.9.

2.8 A Real Life Example

Davis (1952) mixture data denoted by $\mathbf{x} = (x_{11}, x_{12}, \ldots, x_{1r_1}, x_{21}, x_{22}, \ldots, x_{2r_2})$ consists of hours to failure of an indicator valve and of a transmitter valve both used in the aircraft radar sets. Until failure occurs at or before the test termination time $T = 800$ hours is over, it is not known that which valve will fail and hence the category of failure is unknown. Inspection of the failed units allowed the engineers to allocate the failed units to two subpopulations. The test was conducted 1003 times till the test termination time was over. So this is a type-I, right censored sample censored at $T = 800$ hours.

The mixture failure data can be found on Page 76 of Everitt and Hand (1981). The mixture parameters $(\theta_1,\ \theta_2,\ \pi_1)$ can be evaluated using estimators derived in Sections 2.4-2.5. The data summary is, $n = 1003,\ r_1 = 891,\ r_2 = 92,\ r = r_1 + r_2 = 983$, $\sum_{j=1}^{r_1} x_{1j} = 151130$, $T = 800$, $\sum_{j=1}^{r_2} x_{2j} = 22550$ and $n - r = 20$. It is interesting to note that Bayes (Inverted Gamma) estimates have the least variances but are slightly under-estimated. Table 2.8 displays the Bayes estimates using the real life Davis mixture data.

Table 2.3 Bayes estimates (Jeffreys)* of Exponential mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 15,\ \theta_2 = 20,$ $\pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$.

$(\theta_1,\ \theta_2,\ \pi_1)$	n	**15% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(15, 20, 0.40)	50	17.9572(5.73418)	20.3207(5.03935)	0.423806(0.0695231)
	100	16.6404(3.99468)	19.9765(3.52293)	0.411601(0.0573622)
	200	15.9022(2.80208)	19.9752(2.61393)	0.405829(0.0404456)
	300	15.6225(2.3830)	19.8788(2.04341)	0.405732(0.0338122)
(15, 20, 0.60)	50	16.2107(3.96154)	22.2075(7.09718)	0.601036(0.0744419)
	100	15.5932(2.87992)	20.9877(4.93697)	0.600474(0.0555195)
	200	15.3002(2.05167)	20.5198(3.49403)	0.599715(0.0396292)
	300	15.1837(1.73016)	20.2205(2.88425)	0.600330(0.0337118)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_1$	$\hat{\pi}_1$
(15, 20, 0.40)	50	19.5237(7.27322)	20.3257(6.21582)	0.43813(0.0830511)
	100	17.5926(5.0248)	19.5375(4.09142)	0.424621(0.0660233)
	200	16.608(3.88785)	19.5167(3.35918)	0.4182(0.0533334)
	300	16.3683(3.27019)	19.4262(2.68632)	0.416343(0.0466534)
(15, 20, 0.60)	50	16.1496(4.59663)	23.6305(9.25259)	0.590764(0.0846607)
	100	15.5278(3.49805)	21.5041(6.11363)	0.592843(0.071388)
	200	15.3379(2.71106)	20.3556(4.68857)	0.60185(0.057036)
	300	15.2473(2.29145)	20.2701(3.94430)	0.600198(0.0487674)

*Bayes estimates (Jeffreys) means Bayes estimates assuming the Jeffreys prior

Table 2.4 Bayes estimates (Jeffreys) of Exponential mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 20,\ \theta_2 = 25,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$.

$(\theta_1,\ \theta_2,\ \pi_1)$	n	**15% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(20, 25, 0.40)	50	23.5442(7.51464)	25.6308(5.95837)	0.417219(0.0763365)
	100	21.917(5.04182)	24.912(4.28264)	0.410358(0.0558259)
	200	21.0258(3.72203)	24.8975(3.16793)	0.408054(0.0409154)
	300	20.8471(3.09791)	24.8001(2.67412)	0.406275(0.0339158)
(20, 25, 0.60)	50	21.1924(5.08998)	28.0574(8.61439)	0.594901(0.0735825)
	100	20.6477(4.00456)	26.4942(6.16346)	0.598145(0.0557135)
	200	20.3572(2.76600)	25.6145(4.66263)	0.601029(0.0429821)
	300	20.1368(2.29770)	25.2627(3.68680)	0.598905(0.0345350)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(20, 25, 0.40)	50	25.6965(9.17545)	25.0019(7.20281)	0.433689(0.0825697)
	100	23.5141(6.80246)	24.492(5.20641)	0.425992(0.0691188)
	200	22.4065(5.10022)	24.2453(4.15857)	0.42048(0.0542228)
	300	21.6077(4.20428)	24.3823(3.52196)	0.412942(0.0470203)
(20, 25, 0.60)	50	21.1954(6.04377)	29.8813(11.2816)	0.585169(0.0851966)
	100	20.6525(4.55708)	26.7465(8.2803)	0.59932(0.0687103)
	200	20.4043(3.42707)	25.6673(5.82455)	0.598988(0.0556101)
	300	20.2769(2.92470)	25.6005(5.11278)	0.599101(0.0481021)

*Bayes estimates (Jeffreys) means Bayes estimates assuming the Jeffreys prior

Table 2.5 Bayes estimates (Jeffreys) of Exponential mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 24,\ \theta_2 = 30,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$.

$(\theta_1,\ \theta_2,\ \pi_1)$	n	**15% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(24, 30, 0.40)	50	28.4525 (8.74296)	30.526 (7.33529)	0.41943 (0.0713214)
	100	26.429 (6.30768)	30.3471(5.26008)	0.408741 (0.0558356)
	200	25.3443 (4.63340)	29.9643 (3.91221)	0.407988 (0.0403168)
	300	25.1215 (3.67935)	29.8533 (3.15081)	0.406332 (0.0340150)
(24, 30, 0.60)	50	25.3886 (6.05612)	33.6163 (10.4633)	0.598055 (0.0750837)
	100	24.6994 (4.52877)	31.8179 (7.27465)	0.596326 (0.0558714)
	200	24.4408 (3.26834)	30.5856 (5.39760)	0.599955 (0.0394830)
	300	24.2564 (2.73516)	30.2074 (4.34678)	0.600264 (0.0349255)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(24, 30, 0.40)	50	30.7791 (11.0733)	30.113 (8.67572)	0.43502 (0.0827341)
	100	28.0551 (8.05774)	29.5693 (6.74718)	0.42164 (0.0685295)
	200	26.6445 (5.9591)	29.013 (5.03314)	0.42097 (0.0550208)
	300	25.8135 (5.01346)	29.4243 (4.20910)	0.413082 (0.0464281)
(24, 30, 0.60)	50	25.8844 (7.478)	35.5836 (13.2931)	0.587896 (0.0833194)
	100	24.8118 (5.53324)	32.2868 (9.39505)	0.595039 (0.068322)
	200	24.3729 (4.09697)	30.8291 (7.14997)	0.595966 (0.0566869)
	300	24.2964 (3.55407)	30.5110 (5.95118)	0.601142 (0.0485244)

Table 2.6 Bayes estimates (Jeffreys) of Exponential mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 32,\ \theta_2 = 40,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$

$(\theta_1,\ \theta_2,\ \pi_1)$	n	15% Censoring		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(32, 40, 0.40)	50	37.8677 (12.2342)	41.486 (10.1606)	0.415329 (0.0740837)
	100	35.6497 (8.75582)	40.1724 (7.50539)	0.413157 (0.0546148)
	200	33.8065 (6.08417)	40.0185 (5.12281)	0.407268 (0.0405597)
	300	32.8936 (4.77743)	40.0535 (4.23213)	0.404206 (0.033730)
(32, 40, 0.60)	50	34.0149 (8.31294)	44.6464 (13.8107)	0.595337 (0.0725503)
	100	32.9748 (6.01155)	42.0038 (9.83337)	0.601799 (0.0570538)
	200	32.4800 (6.89396)	40.6446 (7.44020)	0.596124 (0.0573039)
	300	32.2165 (3.69155)	40.9386 (6.14946)	0.598584 (0.0335595)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	30% Censoring		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(24, 30, 0.40)	50	41.1463 (15.1363)	40.4395 (11.5441)	0.432 (0.08033)
	100	37.338 (10.4582)	38.9936 (8.27658)	0.426787 (0.0690458)
	200	35.1042 (7.87611)	39.2213 (6.61285)	0.419857 (0.0533039)
	300	34.5057 (6.46954)	39.1094 (5.60577)	0.412626 (0.0464537)
(24, 30, 0.60)	50	34.628 (9.74306)	46.6247 (18.2667)	0.591098 (0.0856927)
	100	32.829 (7.09451)	42.2487 (12.4881)	0.600617 (0.07125)
	200	32.3489 (5.60999)	41.6362 (9.78215)	0.596277 (0.0557039)
	300	32.3612 (4.64040)	40.8479 (7.88144)	0.598940(0.0477019)

Table 2.7 A comparison of the Bayes (Uniform), Bayes (Jeffreys) and Bayes (Inverted Gamma)* estimates of Exponential mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 15$, $\theta_2 = 20$, $\pi_1 = 0.40$ and censoring rate, $C = 15\%$.

Prior	π_1	n	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
Uniform	0.40	50	19.4787(6.14801)	21.0817(5.22377)	0.426302(0.06735)
		100	17.3561(4.08402)	20.2799(3.563)	0.41343(0.05630)
		200	16.2309(2.80386)	20.1879(2.60380)	0.408579(0.04063)
		300	15.8535(2.39774)	19.9589(2.04979)	0.406552(0.03359)
	0.60	50	16.736(4.04247)	24.3933(7.95425)	0.598514(0.07171)
		100	15.8177(2.87994)	21.8991(5.05488)	0.59903(0.05413)
		200	15.4015(2.04131)	20.9437(3.51146)	0.598937(0.03899)
		300	15.2484(1.71999)	20.4960(2.88400)	0.599790(0.03330)
Jeffreys	0.40	50	17.9572(5.73418)	20.3207(5.03935)	0.423806(0.069523)
		100	16.6404(3.99468)	19.9765(3.52293)	0.411601(0.05736)
		200	15.9022(2.80208)	19.9752(2.61393)	0.405829(0.04045)
		300	15.6225(2.3830)	19.8788(2.04341)	0.405732(0.033812)
	0.60	50	16.2107(3.96154)	22.2075(7.09718)	0.601036(0.074442)
		100	15.5932(2.87992)	20.9877(4.93697)	0.600474(0.055520)
		200	15.3002(2.05167)	20.5198(3.49403)	0.599715(0.03963)
		300	15.1837(1.73016)	20.2205(2.88425)	0.600330(0.03372)
Inverted Gamma	0.40	50	17.9044(5.44333)	20.2078(4.87113)	0.423992(0.0693908)
		100	16.6479(3.88489)	19.9207(3.46098)	0.4179(0.0572292)
		200	15.9929(2.84958)	19.7225(2.52160)	0.409934(0.0404716)
		300	15.6360(2.35914)	19.8569(2.03032)	0.405855(0.0337667)
	0.60	50	16.2564(3.83635)	21.8731(6.65261)	0.601799(0.0743165)
		100	15.6306(2.82744)	20.8434(4.7899)	0.601009(0.0553806)
		200	15.3253(2.03060)	20.4459(3.43912)	0.600064(0.0395534)
		300	15.2019(1.71700)	20.1716(2.85266)	0.600583(0.0336511)

* Bayes (Inverted Gamma) means Bayes estimates with Inverted Gamma Prior.

Table 2.8 A comparison of Bayes (Uniform), Bayes (Jeffreys) and Bayes (Inverted Gamma) estimates of parameters of two component Exponential mixture using Davis mixture data.

Uniform		
$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
179.77 (6.610)	331.786 (44.223)	0.898 (0.010)
Jeffreys		
$\hat{\theta}_2$	$\hat{\theta}_2$	$\hat{\pi}_1$
179.75 (6.606)	326.399 (43.463)	0.899 (0.010)
Inverted Gamma		
$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
172.189 (6.247)	320.840 (42.117)	0.898 (0.010)

2.9 Conclusion

The simulation study has displayed that the estimates of all the parameters are over-estimated but the extent of this over-estimation is reduced as the sample size increases. The effect of increase (decrease) in censoring on the estimates of the life time parameters is in the form of the reduction (increase) in the extent of the over-estimation. The same is true for the estimate of the mixing proportion but reduction in the extent of over-estimation is sometimes turns into the slight under-estimation. The extent of over-estimation is observed to be proportional to the size of the corresponding lifetime or mixing proportion parameter and is inversely proportional to the sample size. The estimate of the mixing proportion is over-estimated if true mixing proportion of the first component of the mixture is less than the one-half. However, when the true mixing proportion of the first component of the mixture is greater than one-half, the estimate of the mixing proportion is slightly under-estimated or very close to the true parameter value.

The increase (decrease) in the proportion of some component of the mixture reduces (increases) the standard errors of the corresponding parameter estimate. The standard errors of all the life time and mixing proportion estimates are reduced (increased) as the sample size increases (decreases). The increase (decrease) in the censoring rate increases (decreases) the standard error of all the three estimates.

The comparison among the three Bayes estimators is as follows. The Bayes (Inverted Gamma) estimates of the life time parameters display the least over-estimation while the extent of over-estimation is more in case of Bayes (Uniform) than that of Bayes (Jeffreys). The Bayes (Inverted Gamma) estimates of the life time parameters have the least standard errors as compared to the Bayes (Uniform) and the Bayes (Jeffreys). However, Bayes (Jeffreys) have lesser standard errors than the Bayes (Uniform) estimates. On the other hand, the estimates of the mixing proportion parameter π have almost the same standard errors for all the three Bayes estimates. The informative Bayes (Inverted Gamma) estimates may be the most efficient ones subject to availability of useful prior information that results in appropriate hyper parameters. In the real life example, the proposed estimates are superior to ones presented in Everitt and Hand (1981) page 77 in terms of the use of Bayesian analysis, information on and size of the standard errors.

CHAPTER 3

PROPERTIES AND COMPARISON OF THE BAYES ESTIMATES OF THE RAYLEIGH FINITE MIXTURE PARAMETERS

3.1 Introduction

The Exponential distribution, because of its memory-less property, is used for the life-testing of the products that do not age with time. But there are several radio-wave and electro-vacuum devices whose failure rate depends upon their age and, therefore, the Rayleigh distribution is preferred. Rayleigh distribution has applications in communication engineering. It is used to model lifetimes of the objects which age with time and have increasing hazard rate. Mixture models received considerable attention in the area of survival analysis and reliability. The use of mixture model becomes inevitable when the data are not available for each component. Rather, in many real life situations the data are available only for the overall mixture distribution. 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 Rayleigh distribution. So the two component mixture of the Rayleigh distribution is recommended to model this population.

Hirai (1972) derives quadratic coefficients estimators for the two-parameter Rayleigh distribution. This is a highly efficient complete sample method, even for small samples. Hirai (1976) considers a Type-II censored sample to estimate scale-parameter of the Rayleigh distribution. Raqab (1992) also considered Type-II censored sample in prediction problems with Rayleigh distribution. Sinha and Howlader (1983) and Lalitha and Mishra (1996) have quoted useful references on the Rayleigh model. Fernandez (2000) made Bayesian inference using Type-II doubly censored Rayleigh data. Mostert et al. (1998) conducted Bayesian analysis of survival data using the Rayleigh model and Linex Loss. Saleem and Aslam (2008a, 2008b and 2009) focused Bayesian estimation of the Rayleigh mixture with fixed and Rayleigh distributed censor times.

In this chapter, the said Rayleigh mixture model is defined in Section 2. The likelihood specific expressions are developed in Section 3. In Sections 4-5 the expressions for the uninformative Bayes estimators are derived assuming the Uniform, the Jeffreys and the Square Root Inverted Gamma priors. The expressions for the variances of the said estimators are presented as well. The complete sample expressions of the said estimators and variances are derived in Section 6. To highlight and compare the properties of the said Bayes estimates, a comprehensive simulation study is conducted considering various sample sizes, different censoring rates and a number of combinations of the parameters of the mixture model. A comprehensive simulation study is conducted in Section 7. A real life example is discussed in Section 8 while the concluding remarks are given in Section 9. The squared error loss function is assumed, i.e., posterior means are used as the Bayes estimators.

3.2 The Rayleigh Mixture Model

Recall a k component, type-I mixture model of Section 2.2 with finite mixture density function given by $f(t)=\sum_{i=1}^{k}\pi_i\, f_i(t),\ i=1,2,...,k$; $0<\alpha_i<\infty$; $0\le t_{ij}<\infty$. The k component densities are Rayleigh ones given by $f_i(t)=2\ t\ \exp\left(-t^2/\theta_i^2\right)\big/\theta_i^2\ ,\theta_i>0,\ i=1,2,...,k$ and $t>0$.

3.3 The Maximum Likelihood Estimates for Censored Data

The sampling scheme used in Section 2.3 is adopted here. We define, t_{ij} as the failure time of the j^th unit belonging to the i^th subpopulation, where $j = 1, 2, 3, ..., r_i$, $i = 1, 2$ and $0 < t_{1j}, t_{2j} \le T$. The likelihood function $L(\theta_1, \theta_2, \pi_1 | \mathbf{t})$ for the above conditions takes the form as in equation (2.2), where $S(T) = \sum_{i=1}^{2} \pi_i \exp(-T^2 / \theta_i^2)$ $\mathbf{t} =$ $[\mathbf{t}_{1j}, \mathbf{t}_{2j}]$ is data with $\mathbf{t}_{ij} = [t_{11}, t_{12}, \ldots, t_{1r_i}]$, $i = 1, 2$.

For the functional forms defined in Section 3.2, the likelihood (2.2) takes the following form.

$$L(\theta_1, \theta_2, \pi_1 | \mathbf{t}) \propto \left\{ \prod_{i=1}^{2} \left(\frac{\pi_i}{\theta_i^2} \right)^{r_i} \right\} \exp\{-(\sum_{i=1}^{2} \frac{r_i \overline{t_i^2}}{\theta_i^2})\} \{\sum_{i=1}^{k} \pi_i \exp(-\frac{T^2}{\theta_i^2})\}^{(n-r)} \qquad (3.1)$$

$$\begin{aligned} L(\theta_1, \theta_2, \pi_1 | \mathbf{t}) &\propto \sum_{k=0}^{n-r} \binom{n-r}{k} \left\{ \prod_{i=1}^{2} \pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)} \right\} \\ &\times \left\{ \prod_{i=1}^{2} \theta_i^{-2r_i} \right\} \exp\left[-\sum_{i=1}^{2} \left\{ \frac{r_i \overline{t_i^2} + \{(2k+r-n)i + (2n-2r-3k)\} T^2}{\theta_i^2} \right\} \right] \end{aligned} \qquad (3.2)$$

Here $\overline{t_i^2} = \sum_{j=1}^{r_i} \frac{t_{ij}^2}{r_i}$, $i = 1, 2$. The ML estimates are obtained by simultaneously solving the three equations obtained by setting the first order derivatives of the natural log of the likelihood (3.1) with respect to θ_1, θ_2 and π_1 to zero.

$$\frac{r_i \overline{t_i^2}}{\theta_i^2} + \frac{T^2 (n-r) \pi_i \exp(-T^2 / \theta_i^2)}{\theta_i^2 \{\sum_{i=1}^{2} \pi_i \exp(-T^2 / \theta_i^2)\}} = r_i, \; i = 1, 2 \qquad (3.3)$$

$$(\frac{1-\pi_2}{\pi_1}) r_1 + \frac{\pi_2 (n-r)(\exp(-T^2 / \theta_1^2) - \exp(-T^2 / \theta_2^2))}{\sum_{i=1}^{2} \pi_i \exp(-T^2 / \theta_i^2)} = r_2 \qquad (3.4)$$

It is not possible to solve the above system of three non-linear equations analytically. However, they can be solved by any numerical iterative procedure. A well known asymptotic property of maximum likelihood estimates is that $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta},\ \mathrm{I}^{-1}(\boldsymbol{\theta}))$. The information matrix, $\mathbf{I}(\boldsymbol{\theta})$, is defined in equation (2.7) and $\mathrm{I}^{-1}(\boldsymbol{\theta})$ is a matrix of order 3×3 having the variances of ML estimates on the main diagonal and the covariance on the off diagonal positions. Following are the elements of the symmetric information matrix.

$$E(\frac{-\partial^2 l}{\partial\theta_i^2}) = \frac{4r_i}{\theta_i^2} - \frac{2\{\pi_{3-i}(2T^2-3\theta_i^2)e^{-(T^2/\theta_1^2+T^2/\theta_2^2)} - 3\,\pi_i\,\theta_i^2 e^{-2T^2/\theta_i^2}\}}{T^{-2}(n-r)^{-1}\pi_i^{-1}\theta_i^6\left\{\sum_{i=1}^{2}\pi_i\exp(-T^2/\theta_i^2)\right\}^2},\ i=1,2 \qquad (3.5)$$

$$E(\frac{-\partial^2 l}{\partial\pi_1{}^2}) = \frac{r_1}{\pi_1{}^2} + \frac{r_2}{\pi_2{}^2} + \frac{(n-r)(e^{-T^2/\theta_1^2} - e^{-T^2/\theta_2^2})^2}{\left\{\sum_{i=1}^{2}\pi_i\exp(-T^2/\theta_i^2)\right\}^2} \qquad (3.6)$$

$$E(\frac{-\partial^2 l}{\partial\theta_1\partial\theta_2}) = \frac{4T^4(n-r)\pi_1\pi_2 e^{-(\sum_{i=1}^{2}(T^2/\theta_i^2))}}{\theta_1^3\theta_2^3\left\{\sum_{i=1}^{2}\pi_i\exp(-T^2/\theta_i^2)\right\}^2} = E(\frac{-\partial^2 l}{\partial\theta_2\partial\theta_1}) \qquad (3.7)$$

$$E(\frac{-\partial^2 l}{\partial\theta_i\partial\pi_1}) = \frac{2T^2(n-r)e^{-\sum_{i=1}^{2}(T^2/\theta_i^2)}}{\theta_i^3\left\{\sum_{i=1}^{2}\pi_i\exp(-T^2/\theta_i^2)\right\}^2} = E(\frac{-\partial^2 l}{\partial\pi_1\partial\theta_i}),\ i=1,2 \qquad (3.8)$$

The approximate values of these components of information matrix can easily be obtained by replacing the parameters involved by their respective maximum likelihood estimates obtained by the iterative solution of equations (3.3)-(3.4).

3.4 The Posterior Distributions assuming the Conjugate Prior

Square Root Inverted Gamma prior, being compatible with the likelihood, is considered as a conjugated prior for the evaluation of posterior distribution and hence Bayes estimators.

3.4.1 The Posterior Distributions assuming the Square Root Inverted Gamma Prior

The Bayesian framework combines the prior information with the information contained in the sample data to formulate the posterior distribution. The posterior distribution is the basis for the Bayesian inference. Under the square error loss function, the mean of the posterior distribution is considered as the Bayes estimator.

Let $\theta_i \sim SRIG(m_i, s_i),\ i=1,2$ and $\pi_1 \sim U(0,1)$. Assuming independence, we have a joint prior $g(\theta_1,\theta_2,\pi_1) \propto \prod_{i=1}^{2} \theta_i^{-(2m_i+1)} \exp\{-\sum_{i=1}^{2}(s_i/\theta_i^2)\},\ \theta_i > 0,\ i=1,\ 2$ and $0<\pi_1<1$.

Here $m_i,\ s_i,\ i=1,2$ are the hyperparameters to be elicited. This joint prior is incorporated with the likelihood (3.3) to yield a joint posterior distribution of $\theta_i,\ i=1,2$ and π_1.

$$g_{SRIG}(\theta_1,\theta_2,\pi_1|\mathbf{t}) = \Omega_{SRIG} \sum_{k=0}^{n-r} \binom{n-r}{k} \left\{\prod_{i=1}^{2} \pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\} \left\{\prod_{i=1}^{2} \theta_i^{-(2r_i+2m_i+1)}\right\}$$
$$\times \exp\left\{-\sum_{i=1}^{2} \theta_i^{-2}(A_{ik}+s_i)\right\},\ 0<\theta_i<\infty,\ 0<\pi_i<1,\ i=1,2.$$

where $\Omega_{SRIG}^{-1} = \sum_{k=0}^{n-r}\binom{n-r}{k} \mathrm{B}(n-r_2-k+1, r_2+k+1) \prod_{i=1}^{2} \frac{\Gamma(r_i+m_i)}{(A_{ik}+\mathrm{s_i})^{r_i+m_i}}$ and

$\mathrm{B}(n-r_2-k+1, r_2+k+1)$ is Euler Beta function. The marginal posterior distribution of $\theta_i,\ i=1,2$ and π_1 is obtained by integrating out the nuisance parameters.

$$g_{SRIGi}(\theta_i|\mathbf{t}) = \Omega_{SRIG} \sum_{k=0}^{n-r}\binom{n-r}{k} \mathrm{B}(n-r_2-k+1, r_2+k+1) \frac{\Gamma(r_{3-i}+m_{3-i})\ \exp\{-\theta_i^{-2}(A_{ik}+s_i)\}}{\theta_i^{(2r_i+2m_i+1)}(A_{(3-i)k}+\mathrm{s}_{3-i})^{r_{3-i}+m_{3-i}}},$$
$$i=1,2\ ,0<\theta_i<\infty.$$

$$g_{SRIG3}(\pi_1|\mathbf{t}) = \Omega_{SRIG} \sum_{k=0}^{n-r}\binom{n-r}{k} \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i+m_i)}{(A_{ik}+\mathrm{s_i})^{r_i+m_i}}\right\} \left\{\prod_{i=1}^{2} \pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\},\ 0<\pi_1<1.$$

3.4.1.1 Bayes Estimators assuming the Square Root Inverted Gamma Prior

The Bayes estimators of $\theta_i,\ i=1,2$ and π_1 assuming the SRIG prior are obtained by taking expectations of $\theta_i,\ i=1,2$ and π_1 with respect to their respective marginal posterior distributions.

$$\hat{\theta}_i=\Omega\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}(n-r_2-k+1,r_2+k+1)\Gamma(r_1+m_1-(2-i)/2)\Gamma(r_2+m_2-(i-1)/2)}{(A_{1k}+\ s_1)^{r_1+m_1-1/2}(A_{2k}+\ s_2)^{r_2+m_2-(i-1)/2}},i=1,2\,,$$

$$\hat{\pi}_1=\Omega\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+2,r_2+k+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+m_i)}{(A_{ik}+\ s_i)^{r_i+m_i}}\right\},$$

where $\Omega^{-1}=\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+1,r_2+k+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+m_i)}{(A_{ik}+\ s_i)^{r_i+m_i}}\right\}.$

3.4.1.2 Variances of Bayes Estimators assuming Square Root Inverted Gamma Prior

The expressions for the variances of the Bayes estimators are as under.

$$V(\hat{\theta}_i)=\Omega\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+1,r_2+k+1)\frac{\Gamma(r_1+m_1-(2-i))\Gamma(r_2+m_2-(i-1))}{(A_{1k}+\ s_1)^{r_1+m_1-(2-i)}(A_{2k}+\ s_2)^{r_2+m_2-(i-1)}}-$$
$$-\left\{\Omega\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+1,r_2+k+1)\frac{\Gamma(r_1+m_1-(2-i)/2)\Gamma(r_2+m_2-(i-1)/2)}{(A_{1k}+\ s_1)^{r_1+m_1-1/2}(A_{2k}+\ s_2)^{r_2+m_2-(i-1)/2}}\right\}^2,$$
$$i=1,2$$

$$V(\hat{\pi}_1)=\Omega\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+2,r_2+k+1)\frac{\Gamma(r_1+m_1)\Gamma(r_2+m_2)}{(A_{1k}+\ s_1)^{r_1+m_1}(A_{2k}+\ s_2)^{r_2+m_2}}-$$
$$-\left\{\Omega\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+2,r_2+k+1)\frac{\Gamma(r_1+m_1)\Gamma(r_2+m_2)}{(A_{1k}+\ s_1)^{r_1+m_1}(A_{2k}+\ s_2)^{r_2+m_2}}\right\}^2$$

where $A_{ik}=\sum_{j=1}^{ri}t^2{}_{ij}+\{(2k+r-n)i+(2n-2r-3k)\}T^2,\ i=1,2\,.$

3.5 The Posterior Distributions assuming the Uninformative Priors

Uniform and the Jeffreys are the two most commonly used uninformative priors.

3.5.1 The Posterior Distributions assuming the Uniform Prior

Let $\theta_i \sim Uniform \;\; \forall\; \theta_i \in (0,\infty),\; i=1,2$ and $\pi_1 \sim U(0,1)$. Assuming independence we have a joint prior that is proportional to a constant. This joint prior is incorporated with the likelihood (3.2) to yield the following joint posterior distribution of $\theta_i,\; i=1,2$ and π_1.

$$g_U(\theta_1,\theta_2,\pi_1|\mathbf{t}) = \Omega_U \sum_{k=0}^{n-r}\binom{n-r}{k}\left\{\prod_{i=1}^{2}\pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\}\left\{\prod_{i=1}^{2}\theta_i^{-2r_i}\right\}$$
$$\times\exp\left[-\sum_{i=1}^{2}\left\{\frac{r_i\,\overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2}{\theta_i^2}\right\}\right],\; 0<\theta_1,\theta_2<\infty,\; 0<\pi_1<1.$$

where $\Omega_U^{-1} = \sum\binom{n-r}{k}\mathrm{B}(n-r_2-k+1, r_2+k+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i-1/2)}{A_{ik}^{\;r_i-1/2}}\right\}$. The marginal posterior distribution parameters, $\theta_i,\; i=1,2$ and π_1 is obtained by integrating out the nuisance parameters.

$$g_{Ui}(\theta_i|\mathbf{t}) = \Omega_U \sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+1, r_2+k+1)$$
$$\times\frac{\Gamma(r_{3-i}-1/2)\exp[-\theta_i^{-2}\{r_i\,\overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2\}]}{\theta_i^{2r_i}A_{(3-i)k}^{\;r_{3-i}-1/2}},\; 0<\theta_i<\infty,\; i=1,2.$$

$$g_{U3}(\pi_1|\mathbf{t}) = \Omega_U \sum_{k=0}^{n-r}\binom{n-r}{k}\left\{\prod_{i=1}^{2}\frac{\Gamma(r_1-1/2)}{A_{1k}^{\;r_1-1/2}}\right\}\left\{\prod_{i=1}^{2}\pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\},\; 0<\pi_1<1$$

3.5.1.1 Bayes Estimators assuming Uniform Prior

The Bayes estimators of $\theta_i,\; i=1,2$ and π_1 (under the squared error loss function) are the expected values of $\theta_i,\ i=1,2$ and π_1 with respect to their respective posterior distributions. It is observed that the said estimators and their variances can be obtained by choosing $m_i=-1/2$ and $s_i=0$ in the expressions given in Section 3.4.1.

3.5.2 The Posterior Distributions assuming the Jeffreys Prior

Using definition of Section 2.5.2 and assuming independence, we obtain a joint prior $g(\theta_1,\theta_2,\pi_1)\propto\dfrac{1}{\theta_1\theta_2}$, which is incorporated with the likelihood (3.2) to yield a joint posterior distribution of $\theta_i,\ i=1,2$ and π_1.

$$g_J(\theta_1,\theta_2,\pi_1|\mathbf{t})=\Omega_J\sum_{k=0}^{n-r}\binom{n-r}{k}\left\{\prod_{i=1}^{2}\pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\}\left\{\prod_{i=1}^{2}\theta_i^{-(2r_i+1)}\right\}$$
$$\times\exp[-\theta_i^{-2}\{r_i\ \overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2\}],\ 0<\theta_1,\theta_2<\infty,\ 0<\pi_1<1.$$

where $\Omega_J^{-1}=\sum\binom{n-r}{k}\mathrm{B}(n-r_2-k+1,r_2+k+1)\left\{\prod_{i=1}^{2}\dfrac{\Gamma(r_i)}{A_{ik}^{\ r_i}}\right\}$. The marginal posterior distribution of each parameter is obtained by integrating out the nuisance parameters.

$$g_{Ji}(\theta_i|\mathbf{t})=\Omega_J\sum_{k=0}^{n-r}\binom{n-r}{k}\mathrm{B}(n-r_2-k+1,r_2+k+1)\frac{\Gamma(r_{3-i})}{A_{(3-i)k}^{\ r_{3-i}}}\ \theta_i^{-(2r_i+1)}$$
$$\times\exp[-\theta_i^{-2}\{r_i\ \overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2\}],\ 0<\theta_i<\infty,\ i=1,2$$

$$g_{J3}(\pi_1|\mathbf{t})=\Omega_J\sum_{k=0}^{n-r}\binom{n-r}{k}\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i)}{A_{ik}^{\ r_i}}\right\}\left\{\prod_{i=1}^{2}\pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\},0<\pi_1<1$$

3.5.2.1 Bayes Estimators assuming the Jeffreys Prior

Under the squared error loss function, the expectations of $\theta_i,\ i=1,2$ and π_1 with respect to their respective posterior distributions are the Bayes estimators assuming the Jeffreys prior.

The said estimators and their variances are observed to be obtained by choosing $m_i = 0$ and $s_i = 0$ in the expressions given in Sections 3.4.1.

3.6 The Complete Sample Expressions

Under the conditions given in Section 2.6, the expressions for the Bayes estimators and their variances are simplified as given in Table 3.1 and Table 3.2. The comments regarding amount of information, computational ease and simplification quoted in Section 2.6 also applies here.

3.7 A Simulation Study

A simulations study is conducted in order to investigate the properties of the Bayes estimators in terms of sample sizes and censoring rates. Samples of size $n = 50,\ 100,\ 200,\ 300$ are generated from the two component mixture of Rayleigh distribution with a number of combinations of parameters such that $(\theta_1, \theta_2) \in \{(3,5),\ (8,12),\ (16,20),\ (25,36)\}$ and $\pi_1 \in \{0.40,\ 0.60\}$. Probabilistic mixing was used to generate the mixture data. For each observation a random number u was generated from the uniform on [0, 1] distribution. If $u < \pi_1$, the observation is taken randomly from F_1 (the Rayleigh distribution with parameter θ_1), and the observation is taken randomly from F_2 (the Rayleigh distribution with parameter θ_2) otherwise. Remaining details of the simulation scheme are the same as mentioned in Section 2.7. The results of the simulation study are presented in Tables 3.3-3.7.

Table 3.1 The complete sample expressions for the Bayes and ML estimators as $T \to \infty$

Parameters	Bayes Estimators (Uniform)	Bayes Estimators (Jeffreys)	ML Estimators
θ_1	$\dfrac{\Gamma(n_1-1)\sqrt{(\sum t_{1j}^2)}}{\Gamma(n_1-0.5)}$	$\dfrac{\Gamma(n_1-0.5)\sqrt{(\sum t_{1j}^2)}}{\Gamma(n_1)}$	$\sqrt{\dfrac{\sum t_{1j}^2}{n_1}}$
θ_2	$\dfrac{\Gamma(n_2-1)\sqrt{(\sum t_{2j}^2)}}{\Gamma(n_2-0.5)}$	$\dfrac{\Gamma(n_2-0.5)\sqrt{(\sum t_{2j}^2)}}{\Gamma(n_2)}$	$\sqrt{\dfrac{\sum t_{2j}^2}{n_2}}$
π_1	$\dfrac{n_1+1}{n+2}$	$\dfrac{n_1+1}{n+2}$	$\dfrac{n_1}{n}$

Table 3.2 The complete sample expressions for the variances of the Bayes and ML estimators as $T \to \infty$

Parameters	Variances of Bayes Estimators(Uniform)	Variances of ML Estimators
θ_1	$\dfrac{(\sum t_{1j}^2)}{\Gamma^2(n_1-0.5)}\left\{\prod_{i=1}^{2}\Gamma(n_1+i-2.5)-\Gamma^2(n_1-1)\right\}$	$\dfrac{(\sum t_{1j}^2)}{4n_1^2}$
θ_2	$\dfrac{(\sum t_{2j}^2)}{\Gamma^2(n_2-0.5)}\left\{\prod_{i=1}^{2}\Gamma(n_2+i-2.5)-\Gamma^2(n_2-1)\right\}$	$\dfrac{(\sum t_{2j}^2)}{4n_2^2}$
π_1	$\dfrac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\dfrac{n_1 n_2}{n^3}$

Parameters	Variances of Bayes Estimators(Jeffreys)
θ_1	$(\sum t_{1j}^2)\left\{\prod_{i=1}^{2}\Gamma(n_1+i-2)-\Gamma^2(n_1-0.5)\right\}\Big/\Gamma^2(n_1)$
θ_2	$(\sum t_{2j}^2)\left\{\prod_{i=1}^{2}\Gamma(n_2+i-2)-\Gamma^2(n_2-0.5)\right\}\Big/\Gamma^2(n_2)$
π_1	$(n_1+1)(n_2+1)\big/(n+2)^2(n+3)$

Table 3.3 Bayes estimates (Jeffreys)* of Rayleigh mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 3,\ \theta_2 = 5,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$

$(\theta_1,\ \theta_2,\ \pi_1)$	n	**15% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(3,5, 0.40)	50	3.23134(0.502282)	5.02938(0.625413)	0.417125(0.0698353)
	100	3.09397(0.330599)	5.03397(0.416239)	0.409646(0.0508097)
	200	3.046974(0.21901)	5.020636(0.282609)	0.402913(0.0367326)
	300	3.0329198(0.1743307)	5.00748(0.220199)	0.402404(0.028347)
(3, 5, 0.60)	50	3.12489(0.392458)	5.11615(0.877498)	0.610581(0.0722052)
	100	3.08993(0.290635)	5.03247(0.600118)	0.606719(0.0554734)
	200	3.04246(0.187932)	5.00586(0.398795)	0.605195(0.0376934)
	300	3.030238(0.160767)	5.01088(0.30495)	0.60319(0.030999)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(3,5, 0.40)	50	3.41044(0.641822)	4.79069(0.784077)	0.450247(0.089137)
	100	3.27045(0.4744)	4.87132(0.581756)	0.433759(0.0692577)
	200	3.15174(0.335481)	4.9296(0.38126)	0.418022(0.0481476)
	300	3.108805(0.291567)	4.95816(0.318616)	0.4114630(0.03858)
(3, 5, 0.60)	50	3.20048(0.455437)	5.01647(1.17027)	0.630236(0.0870769)
	100	3.11603(0.360813)	4.86894(0.83644)	0.622412(0.0699798)
	200	3.0948(0.285844)	4.86557(0.611822)	0.616464(0.0589762)
	300	3.069089(0.23587)	4.91305(0.490298)	0.613555(0.047567)

*Bayes estimates (Jeffreys) means the Bayes estimates assuming the Jeffreys prior

Table 3.4 Bayes estimates (Jeffreys) of Rayleigh mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 8,\ \theta_2 = 12,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$

$(\theta_1,\ \theta_2,\ \pi_1)$	n	**15% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(8,12, 0.40)	50	8.57369(1.34349)	12.0273(1.43818)	0.421886(0.0721387)
	100	8.30014(0.96195)	11.9924(1.03634)	0.413154(0.0546364)
	200	8.134013(0.669734)	12.03165(0.70057)	0.4064439(0.037505)
	300	8.116058(0.493910)	12.015357(0.545469)	0.404751(0.030085)
(8,12, 0.60)	50	8.35362(1.06451)	12.0887(2.09924)	0.609797(0.0722502)
	100	8.18552(0.807363)	12.0015(1.45755)	0.608234(0.0562353)
	200	8.147259(0.562267)	11.916232(1.012702)	0.6086179(0.0396566)
	300	8.1027563(0.440342)	11.942353(0.80511)	0.604298(0.0330128)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$
(8,12, 0.40)	50	9.024(1.76041)	11.489(1.86603)	9.024(1.76041)
	100	8.62143(1.29162)	11.6813(1.35801)	8.62143(1.29162)
	200	8.454500(0.953146)	11.749945(0.986117)	8.454500(0.953146)
	300	8.33747(0.798912)	11.80245(0.795017)	8.33747(0.798912)
(8,12, 0.60)	50	8.35329(1.2233)	12.0088(2.64088)	8.35329(1.2233)
	100	8.33408(0.958812)	11.6562(1.91376)	8.33408(0.958812)
	200	8.23932(0.724008)	11.66111(1.356878)	8.23932(0.724008)
	300	8.1442254(0.6374216)	11.790946(1.224265)	8.1442254(0.6374216)

Table 3.5 Bayes estimates (Jeffreys) of Rayleigh mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 16,\ \theta_2 = 20,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$

$(\theta_1,\ \theta_2,\ \pi_1)$	n	**15% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(16, 20, 0.40)	50	17.2856(2.77738)	19.7064(2.47503)	0.418106(0.0743458)
	100	16.7525(1.99085)	19.8082(1.74753)	0.413326(0.0533546)
	200	16.392669(1.443508)	19.91990(1.240641)	0.406413(0.03950327)
	300	16.27085(1.185301)	19.88942(1.03848)	0.4050460(0.03288)
(16, 20, 0.60)	50	16.4437(2.05832)	20.2439(3.25763)	0.603891(0.0732955)
	100	16.2059(1.53027)	19.8916(2.40581)	0.605693(0.054399)
	200	16.1263(1.115813)	19.93559(1.70904)	0.603247(0.039846)
	300	16.117628(0.941156)	19.846232(1.4213144)	0.602854(0.0349025)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(16, 20, 0.40)	50	17.6617(3.29426)	19.3868(2.96606)	0.437069(0.0867775)
	100	17.0776(2.50343)	19.4974(2.18107)	0.428409(0.0709385)
	200	16.8004(1.842101)	19.51368(1.66737)	0.421603(0.054826)
	300	16.547134(1.69240)	19.60145(1.40818)	0.4165624(0.0463104)
(16, 20, 0.60)	50	16.3197(2.46724)	20.4771(4.06554)	0.597395(0.0859806)
	100	16.2811(1.75689)	19.9413(3.03369)	0.607658(0.0704141)
	200	16.132797(1.415933)	19.92792(2.350997)	0.605024(0.055708)
	300	16.1678494(1.23173)	19.66430(2.057206)	0.6077401(0.049660)

Table 3.6 Bayes estimates (Jeffreys) of Rayleigh mixture parameters and their standard errors (in parenthesis) with $\theta_1 = 25,\ \theta_2 = 36,\ \pi_1 = 0.40,\ 0.60$ and censoring rates, $C = 15\%,\ 30\%$.

$(\theta_1,\ \theta_2,\ \pi_1)$	n	15% Censoring		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(25,36,0.40)	50	27.091(4.42658)	35.8883(4.50264)	0.424279(0.0698265)
	100	26.2495(3.12986)	35.7654(3.13701)	0.414077(0.0561989)
	200	25.62280(2.157442)	35.93667(2.21848)	0.407359(0.038810)
	300	25.34289(1.681530)	35.876649(1.71885)	0.4035629(0.031579)
(25,36,0.60)	50	25.9119(3.22669)	36.4048(5.85731)	0.607454(0.0710905)
	100	25.7317(2.41903)	35.6736(4.11674)	0.607249(0.0557861)
	200	25.366839(1.726969)	35.849775(2.96943)	0.60542(0.040997)
	300	25.27275(1.394435)	35.79714(2.45937)	0.604629(0.032762)
$(\theta_1,\ \theta_2,\ \pi_1)$	n	**30% Censoring**		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
(25,36,0.40)	50	28.4221(5.36772)	34.6144(5.70795)	0.451977(0.0879951)
	100	27.1634(4.05112)	34.849(4.12275)	0.436692(0.0693779)
	200	26.51757(3.04613)	35.0538(2.94486)	0.42363(0.05165)
	300	26.095194(2.552831)	35.1981398(2.454655)	0.418069(0.043926)
(25,36,0.60)	50	25.971(3.82141)	35.893(7.28253)	0.613014(0.0840109)
	100	25.7334(2.90073)	35.4097(5.45395)	0.614933(0.0691857)
	200	25.5885078(2.3552110)	35.0946921(4.293542)	0.6143718(0.057976)
	300	25.48305(2.01465)	35.18045(3.60990)	0.613207(0.0492847)

Table 3.7 A comparison of Bayes (Jeffreys), Bayes (Uniform) and Bayes (SRIG)* estimates and standard errors (in parenthesis) of Rayleigh mixture parameters $\theta_1 = 3,\ \theta_2 = 5,\ \pi_1 = 0.40,\ 0.60$ with censoring rate, $C = 15\%$.

Prior	π_1	n	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\pi}_1$
UNIFORM	0.40	50	3.23134(0.502282)	5.02938(0.625413)	0.417125(0.0698353)
		100	3.12326(0.33616)	5.0582(0.42017)	0.410248(0.050769)
		200	3.0607218(0.22106)	5.03261(0.28395)	0.4031739(0.03675)
		300	3.041816(0.17545)	5.015486(0.22094)	0.4025654(0.02836)
	0.60	50	3.15025(0.39047)	5.24344(0.89774)	0.609891(0.07088)
		100	3.10297(0.28877)	5.08651(0.60070)	0.606502(0.054863)
		200	3.04946(0.18748)	5.02996(0.39936)	0.605204(0.037509)
		300	3.03510(0.16060)	5.02623(0.30547)	0.603234(0.030917)
JEFFREYS	0.40	50	3.23134(0.50228)	5.02938(0.62541)	0.417125(0.0698353)
		100	3.09397(0.33059)	5.03397(0.41624)	0.409646(0.0508097)
		200	3.046974(0.21901)	5.020636(0.28261)	0.402913(0.0367326)
		300	3.0329198(0.17433)	5.00748(0.22019)	0.402404(0.028347)
	0.60	50	3.12489(0.39246)	5.11615(0.87749)	0.610581(0.0722052)
		100	3.08993(0.29064)	5.03247(0.60012)	0.606719(0.0554734)
		200	3.04246(0.18793)	5.00586(0.39879)	0.605195(0.0376934)
		300	3.030238(0.16077)	5.01088(0.30495)	0.60319(0.030999)
SRIG	0.40	50	3.41649(0.46943)	5.02906(0.60632)	3.41649(0.46943)
		100	3.19301(0.32525)	5.03024(0.41549)	3.19301(0.32525)
		200	3.095991(0.21871)	5.019696(0.28319)	3.095991(0.21871)
		300	3.065164(0.17442)	5.00724(0.22066)	3.065164(0.17442)
	0.60	50	3.24634(0.36121)	5.13588(0.79265)	3.24634(0.36121)
		100	3.15724(0.27792)	5.03233(0.57548)	3.15724(0.27792)
		200	3.07901(0.18526)	5.00091(0.39436)	3.07901(0.18526)
		300	3.05520(0.15972)	5.00671(0.30394)	3.05520(0.15972)

*Bayes (SRIG) means the Bayes estimates assuming the Square Root Inverted Gamma prior.

3.8 A Real Life Example

Mendenhall and Hader (1958) mixture data $\mathbf{t} = (t_{11}, t_{12}, \ldots, t_{1r_1}, t_{21}, t_{22}, \ldots, t_{2r_2})$ consists of hours to failure for ARC-1 VHF radio transmitter receivers of a single commercial airline. The radio transmitter receivers that seemed to be failed at or before 630 hours of operation were removed from the aeroplanes as a general policy of the airline giving Type-I right censored observations at $T = 630$ hours. On the other hand, inspection of the failed units allowed the engineers to allocate the failed units to any one of the two different subpopulations. The mixture failure data can be found on page 509 in Mendenhall and Hader (1958). Mendenhall and Hader fitted Exponential distribution to this data. The transformation $x = \sqrt{t}$ of an Exponential random variable $(\mathbf{t})$ yields a Rayleigh random variable $(\mathbf{x})$. This property allows us to use the transformed Mendenhall and Hader data set for our analysis. The transformed test termination time will be the square root of the termination time used by Mendenhall and Hader. It is interesting to note that despite the transformation almost no major computations are required to have the data summary required to evaluate the proposed estimates. For instance, $\sum_{j=1}^{r_1} x_{1j}^2 = \sum_{j=1}^{r_1} t_{1j} = 20458$ and $\sum_{j=1}^{r_1} x_{2j}^2 = \sum_{j=1}^{r_2} t_{2j} = 50056$. Other sample characteristics required are as follows. $n = 369,\ r_1 = 107,\ r_1 = 218,\ r = r_1 + r_2 = 325$. The Rayleigh mixture parameters $(\theta_1, \theta_2, \pi_1)$ can be evaluated using the estimators derived in Section 3.4. The Bayes (Jeffreys) estimates of Rayleigh mixture parameters are computed using equations of the form given in Section 4 as $\hat{\theta}_1 = 241.25951,\ \hat{\theta}_2 = 334.84301$ and $\hat{\pi}_1 = 0.31299$ (corrected to five decimal places) with their respective standard errors, $SE(\hat{\theta}_1) = 33.80173$, $SE(\hat{\theta}_2) = 25.666077$ and $SE(\hat{\pi}_1) = 0.46371$. The estimates are compatible with ones presented in Mendenhall and Hader (1958) and Sinha (1998). It can easily be shown that having an informative (SRIG) prior, the standard error of estimates can further be reduced. Also, it is encouraging to note that the proposed lifetime estimates are much greater than the corresponding sample average lifetimes of the two subgroups i.e., $\overline{t_1} = 191.2 << 241.25951$, $\overline{t_2} = 229.6 << 334.84301$ as is expected in the right censoring situations. The values of Bayes estimates assuming different priors are presented in Table 3.8.

Table 3.8 A comparison of Bayes (Uniform), Bayes (Jeffreys) and Bayes (Informative) estimates of the mixture parameters using Mendenhall and Hader mixture data.

Uniform		
θ_1	θ_2	π_1
245.080(34.607)	335.653 (25.881)	0.3137 (0.0265)
Jeffreys		
θ_1	θ_2	p
241.260(33.802)	334.843 (25.666)	0.3130 (0.4637)
Informative Prior		
θ_1	θ_2	π_1
225.369(30.392)	331.071(24.719)	0.310 (0.026)

3.9 Conclusion

Some interesting properties of the Bayes estimates are highlighted by the simulation study. The estimates of all the mixture parameters are over-estimated. The extent of over-estimation is higher in case of the first lifetime parameter than the second one. Increasing the sample size reduces the extent of over-estimation of all the estimates. The extent of over-estimation is higher for the estimates of parameters of larger size.

Another interesting remark concerning the variances of the estimates of the lifetime parameters is that increasing (decreasing) the proportion of a component in the mixture reduces (increases) the variance of the estimate of the corresponding lifetime parameter. The variances of estimates of lifetime and proportion parameters are reduced as the sample size increases. The effect of increase in censoring rate on the estimates of the first lifetime parameter is always observed in the form of an increase in the extent of over-estimation. The same is true for the estimates of the mixing proportion parameter. The effect of an increase in the censoring rate of the second life time parameter is a bit interesting, slight over-estimation is observed to change into a slight under-estimation in most of the cases. However, a slight fall in the extent of over-estimation is observed in some rare exceptions. Increasing the censoring rate decreases the variances of estimates of all the mixture parameters.

The Bayes (Jeffreys), Bayes (Uniform) and Bayes (SRIG) estimates of the lifetime parameters and those of the proportion parameters are over-estimated. The extent of over-estimation is slightly higher in case of Bayes (SRIG) but with lesser standard errors of all the estimates for lifetime parameters. The extent of over-estimation and variances of the estimates are slightly higher in case of Bayes (Uniform) than the Bayes (Jeffreys). So the Bayes lifetime estimates with informative (SRIG) prior seem to be more efficient than their uninformative counterparts. A better choice of hyperparameters may further improve the efficiency of Bayes (SRIG) estimates. In the real life example the results are presented using the three Bayes estimates with the help of a Rayleigh mixture. These estimates are comparable with those presented in Sinha (1998) with the help of an Exponential mixture. The Bayes (SRIG) estimates seem superior in terms of lesser standard errors.

CHAPTER 4

THE PRIOR SELECTION FOR THE MIXTURE OF RAYLEIGH DISTRIBUTION USING PREDICTIVE INTERVALS

4.1 Introduction

Rayleigh model is especially suitable for the life-testing of the products that age with time. In this chapter, the 95% Bayes predictive intervals are evaluated for the two component mixture of the Rayleigh distribution assuming three conjugate priors i.e., Inverted Chi, the Inverted Rayleigh and the Square Root Inverted Gamma priors. The conjugate priors have functional form compatible with the likelihood. The Bayesian predictive intervals are evaluated for different choices of the hyperparameters. The motivation is to explore the prior that produce the most precise estimates. The trends are also explored in terms of the hyperparameters of each prior distribution as to how do they affect the scatter of the respective predictive intervals. These trends serve as a sort of partial prior elicitation, reduce prior subjectivity and increase precision of the estimates. A type-IV mixture sample data is simulated and the type-I, right censoring is employed. Sloan and Sinha (1991) constructed Bayesian predictive intervals for a mixture of Exponential failure-time distributions. Dey and Das (2005) explores Bayesian predictive intervals for Rayleigh distribution. Saleem and Aslam (2008b) observe the behavior of Bayesian predictive intervals in terms of hyperparameters of three conjugate priors.

4.2 The Rayleigh Mixture Model

A finite type-I mixture distribution as described in Section 3.2 is considered with two Rayleigh component densities with unknown parameters and unknown mixing weights.

4.3 Sampling

A type-IV mixture sample as stated in Section 3.3 is simulated to conduct the computations involved with ordinary type-I, right censoring.

4.3.1 The Likelihood Function for Censored Data

For the said type-I mixture with two Rayleigh components and with a sample of type-IV, the likelihood function as developed in Section 3.3 is adopted.

4.4 The Posterior Distribution assuming the Inverted Chi Prior

We assume that θ_1 and θ_2 are independent a priory and follow Inverted Chi (IC) distributions with a_1 and a_2 degrees of freedom respectively. We further assume that π_1 is Uniform random variable with support [0, 1]. So the joint prior distribution of $\theta_1,\ \theta_2$ and π_1 is

$$g\left(\theta_1,\theta_2,\pi_1\right) \propto \left\{\prod_{i=1}^{2}\theta_i^{-\left(a_i+1\right)}\right\}\exp\{-\sum_{i=1}^{2}(1/2\theta_i^2)\},\ i=1,2;\ \theta_i>0,\ 0<\pi_1<1 \tag{4.1}$$

Combining likelihood and prior we get joint posterior distribution of $\theta_1,\ \theta_2$ and p as follows

$$\begin{aligned} g_{IC}\left(\theta_1,\theta_2,\pi_1\,|\,\mathrm{t}\right) \propto \sum_{k=0}^{n-r}\binom{n-r}{k}&\left\{\prod_{i=1}^{2}\pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\}\left\{\prod_{i=1}^{2}\theta_i^{-\left(2r_i+a_i+1\right)}\right\}\\ &\times\exp[-\theta_i^{-2}\{r_i\ \overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2+0.5\}],\\ &\theta_i>0,\ 0<\pi_1<1,\ i=1,2. \end{aligned} \tag{4.2}$$

Here a_1,a_2 are the hyperparameters to be elicited.

4.4.1 Bayes Estimators assuming the Inverted Chi Prior

The expectations of $\theta_1,\ \theta_2$ and π_1 with respect to their respective marginal posterior distributions are called the Bayes estimators of θ_1, θ_2 and π_1 (posterior means) under the square error loss function.

$$\hat{\theta}_i = \Omega\ \Gamma(\frac{a_i - 1}{2} + r_i)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k, b_k\right)}{A_k^{\frac{a_1-(2-i)}{2}+r_1} B_k^{\frac{a_2-(i-1)}{2}+r_2}},\ i = 1, 2$$

$$\hat{\pi}_1 = \Omega\ \Gamma(\frac{a_i}{2} + r_i)\ \sum_{k=0}^{n-r}\binom{n-r}{k}\frac{B\left(a_k + 1, b_k\right)}{A_k^{\frac{a_1}{2}+r_1} B_k^{\frac{a_2}{2}+r_2}}$$

where $\Omega^{-1} = \Gamma(\frac{a_i}{2} + r_i)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k, b_k\right)}{A_k^{\frac{a_1}{2}+r_1} B_k^{\frac{a_2}{2}+r_2}}$.

The algebraic expressions for the variances of the above estimates are given as under.

$$V\left(\hat{\theta}_{\mathrm{i}}\right) = \Omega\ \Gamma(\frac{a_i}{2} + r_i - 1)\ \sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k, b_k\right)}{A_k^{\frac{a_1}{2}+r_1-(2-i)} B_k^{\frac{a_2}{2}+r_2-(i-1)}}$$
$$-\left\{\Omega\ \Gamma(\frac{a_i - 1}{2} + r_i)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k, b_k\right)}{A_k^{\frac{a_1-(2-i)}{2}+r_1} B_k^{\frac{a_2-(i-1)}{2}+r_2}}\right\}^2,\ i = 1, 2$$

$$V\left(\hat{\pi}_1\right) = \Gamma(\frac{a_i}{2} + r_i)\ \sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k + 2,\ b_k\right)}{A_k^{\frac{a_1}{2}+r_1} B_k^{\frac{a_2}{2}+r_2}} - \left\{\Gamma(\frac{a_i}{2} + r_i)\ \sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k + 1,\ b_k\right)}{A_k^{\frac{a_1}{2}+r_1} B_k^{\frac{a_2}{2}+r_2}}\right\}^2$$

4.4.2 Bayesian Predictive Intervals assuming the Inverted Chi Prior

The predictive distribution of the future observation y is

$$p\left(y|\mathrm{t}\right) = \int_0^1\int_0^\infty\int_0^\infty g\left(\theta_1, \theta_2, \pi_1 \mid \mathrm{t}\right)\ p\left(y \mid \theta_1, \theta_2, \pi_1\right)\ d\theta_1\, d\theta_2\, d\pi_1 \qquad (4.3)$$

Where $g\left(\theta_1, \theta_2, \pi_1 \mid \mathrm{t}\right)$ is the posterior distribution given by equation (4.2) and $p(y \mid \theta_1, \theta_2, \pi_1)$ is the data model described in Section 4.2.

$$p\left(y\middle|\mathrm{t}\right)=C^{-1}\sum_{k=0}^{n-r}\binom{n-r}{k}y\;[B\left(a_k+1,\;b_k\right)\Gamma\left(r_1+1+\frac{a_1}{2}\right)\Gamma\left(r_2+\frac{a_2}{2}\right)/\left(A_k+y^2\right)^{r_1+1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}}$$

$$+B\left(a_k,b_k\right)\Gamma\left(r_1+\frac{a_1}{2}\right)\Gamma\left(r_2+1+\frac{a_2}{2}\right)/A_k^{\;r_1+\frac{a_1}{2}}\left(B_k+y^2\right)^{r_2+1+\frac{a_2}{2}}\;] \qquad (4.4)$$

where $C=\Gamma\left(r_1+\frac{a_1}{2}\right)\Gamma\left(r_2+\frac{a_2}{2}\right)\sum_{k=0}^{n-r}\binom{n-r}{k}y\left\{\mathrm{B}\left(a_k+1,b_k\right)+\mathrm{B}\left(a_k,b_k+1\right)\right\}/A_k^{\;r_1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}}$.

A $(1-\alpha)100\%$ Bayesian prediction interval (L,U) is obtained by solving the two equations $\int_0^L p\left(y\middle|\mathrm{t}\right)dy=\frac{\alpha}{2}=\int_U^\infty p\left(y\middle|\mathrm{t}\right)dy$. After manipulation these equations become

$$\frac{\alpha}{2}=\sum_{k=0}^{n-r}\binom{n-r}{k}G_k\{\frac{(n+2)}{A_k^{\;r_1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}}}-\frac{a_k}{\left(A_k+L^2\right)^{r_1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}}}-\frac{b_k}{A_k^{\;r_1+\frac{a_1}{2}}(B_k+L^2)^{r_2+\frac{a_2}{2}}}\} \qquad (4.5)$$
$$\times\;[1/\sum_{k=0}^{n-r}((n+2)\binom{n-r}{k}G_k/A_k^{\;r_1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}})]$$

$$\frac{\alpha}{2}=\sum_{k=0}^{n-r}\binom{n-r}{k}G_k\{\frac{a_k}{\left(A_k+U^2\right)^{r_1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}}}+\frac{b_k}{A_k^{\;r_1+\frac{a_1}{2}}\left(B_k+U^2\right)^{r_2+\frac{a_2}{2}}}\} \qquad (4.6)$$
$$\times\;[1/(\sum_{k=0}^{n-r}(n+2)\binom{n-r}{k}G_k/A_k^{\;r_1+\frac{a_1}{2}}B_k^{\;r_2+\frac{a_2}{2}})]$$

where $A_k=r_1\,\overline{t_1^2}+(n-r-k)\;T^2+\frac{1}{2},B_k=r_2\overline{t_2^2}+k\;T^2+\frac{1}{2},\;a_k=n-r_2-k+1$ and

$b_k=r_2+k+1,\;G_k=\Gamma(n-r_2-k+1)\;\Gamma(r_2+k+1)$.

4.5 The Posterior Distribution assuming the Inverted Rayleigh Prior

We assume that θ_1 and θ_2 are independent a priory and follow Inverted Rayleigh (IR) distribution with parameters a_1 and a_2 respectively. We further assume that π_1 is a Uniform random variable with support as [0, 1]. Here is the joint prior distribution of $\theta_1,\;\theta_2$ and π_1.

$$g(\theta_1,\theta_2,\pi_1) \propto \left\{\prod_{i=1}^{2}\theta_i^{-3}\right\}\exp\{-\sum_{i=1}^{2}(a_i/\theta_i^2)\},\ \theta_1,\ \theta_2>0,\ 0<\pi_1<1 \tag{4.7}$$

Combining likelihood and prior we get joint posterior distribution of θ_1,θ_2 and π as

$$\begin{aligned} g_{IR}\ (\theta_1,\theta_2,\pi_1\,|\,\mathrm{t}) &\propto \sum_{k=0}^{n-r}\binom{n-r}{k}\left\{\prod_{i=1}^{2}\pi_i^{(2r_2+2k-n)i+(2n-3k-3r_2)}\right\}\left\{\prod_{i=i}^{2}\theta_i^{-(2r_i+3)}\right\} \\ &\times\exp\left[-\theta_i^{-2}\left\{r_i\ \overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2+a_i\right\}\right], \\ &\qquad \theta_1,\ \theta_2>0,\ 0<\pi_1<1. \end{aligned} \tag{4.8}$$

Here a_1, a_2 are the hyperparameters to be elicited.

4.5.1 Bayes Estimators assuming the Inverted Rayleigh Prior

The expectations of θ_1,θ_2 and π_1 with respect to their respective marginal posterior distributions are called the Bayes estimators (posterior means) under the square error loss function.

$$\hat{\theta}_i=\Omega\ \Gamma\left(r_i+\frac{1}{2}\right)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}(a_k,b_k)}{A_{1k}^{\,r_1+(i/2)}A_{2k}^{\,r_2+(\frac{3-i}{2})}},\ i=1,2$$

$$\hat{\pi}_1=\Omega\ \Gamma(r_i+1)\ \sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}(a_k+1,b_k)}{A_{1k}^{\,r_1+1}A_{2k}^{\,r_2+1}}$$

where $\Omega^{-1}=\Gamma(r_i+1)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}(a_k,b_k)}{A_{1k}^{\,r_1+1}A_{2k}^{\,r_2+1}}$.

The algebraic expressions for the variances of the above estimates are as under.

$$V\left(\hat{\theta}_i\right)=\Omega\ \Gamma(r_i)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}(a_k,b_k)}{A_{1k}^{\,r_1+(i-1)}A_{2k}^{\,r_2+(2-i)}}-\left\{\Omega\ \Gamma\left(r_i+\frac{1}{2}\right)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}(a_k,b_k)}{A_{1k}^{\,r_1+(i/2)}A_{2k}^{\,r_2+(\frac{3-i}{2})}}\right\}^2,\quad i=1,2.$$

$$V\left(\hat{\pi}_1\right)=\Omega\,\Gamma\left(r_i+1\right)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k+2,b_k\right)}{A_{1k}^{\;r_1+1}A_{2k}^{\;r_2+1}}\}-\left\{\Omega\,\Gamma\left(r_i+1\right)\sum_{k=0}^{n-r}\binom{n-r}{k}\frac{\mathrm{B}\left(a_k+1,b_k\right)}{A_{1k}^{\;r_1+1}A_{2k}^{\;r_2+1}}\right\}^2$$

4.5.2 Bayesian Predictive Intervals assuming the Inverted Rayleigh Prior

The predictive distribution of the future observation y is again what is stated in (4.3) where $g\left(\theta_1,\theta_2,\pi_1\mid \mathrm{t}\right)$ is the posterior distribution given by equation (4.8) and $p\left(y\mid\theta_1,\theta_2,\pi_1\right)$ is the data model described in Section 4.2.

$$p\left(y\middle|\mathrm{t}\right)=D^{-1}\sum_{k=0}^{n-r}\binom{n-r}{k}y\left\{\frac{\mathrm{B}\left(a_k+1,b_k\right)\Gamma\left(r_1+2\right)\Gamma\left(r_2+1\right)}{\left(A_{1k}+y^2\right)^{r_1+2}A_{2k}^{\;r_2+1}}+\frac{\mathrm{B}\left(a_k,b_k+1\right)\Gamma\left(r_1+1\right)\Gamma\left(r_2+2\right)}{A_{1k}^{\;r_1+1}\left(A_{2k}+y^2\right)^{r_2+2}}\right\},\,y>0. \tag{4.10}$$

where $D=\Gamma\left(r_1+1\right)\Gamma\left(r_2+1\right)\sum_{k=0}^{n-r}\binom{n-r}{k}y\left\{\mathrm{B}\left(a_k+1,b_k\right)+\mathrm{B}\left(a_k,b_k+1\right)\right\}/\,A_{1k}^{\;r_1+1}A_{2k}^{\;r_2+1}$.

The $(1-\alpha)100\%$ Bayesian Prediction Interval (L,U) is obtained by solving the two equation as given in Section 4.4.2. On manipulation these equations become

$$\begin{aligned}\frac{\alpha}{2}=&\sum_{k=0}^{n-r}\binom{n-r}{k}G_k\{\frac{(n+2)}{A_{1k}^{\;r_1+1}A_{2k}^{\;r_2+1}}-\frac{a_k}{\left(A_{1k}+L^2\right)^{r_1+1}A_{2k}^{\;r_2+1}}-\frac{b_k}{A_{1k}^{\;r_1+1}\left(A_{2k}+L^2\right)^{r_2+1}}\}\\&\times\;[1/(\sum_{k=0}^{n-r}\binom{n-r}{k}G_k(n+2)/\,A_{1k}^{\;r_1+1}A_{2k}^{\;r_2+1})]\end{aligned} \tag{4.11}$$

$$\begin{aligned}\frac{\alpha}{2}=&\sum_{k=0}^{n-r}\binom{n-r}{k}G_k\{\frac{a_k}{\left(A_{1k}+U^2\right)^{r_1+1}A_{2k}^{\;r_2+1}}+\frac{b_k}{A_{1k}^{\;r_1+1}\left(A_{2k}+U^2\right)^{r_2+1}}\}\\&\times\;[1/(\sum_{k=0}^{n-r}\binom{n-r}{k}G_k(n+2)/\,A_{1k}^{\;r_1+1}A_{2k}^{\;r_2+1})]\end{aligned} \tag{4.12}$$

where $A_{ik}=r_i\overline{t_i^2}+\{(2k+r-n)i+(2n-2r-3k)\}T^2+a_i,\; i=1,2.$

4.6 The Posterior Distribution assuming the Square Root Inverted Gamma Prior

We assume that θ_1 and θ_2 are independent a priory and follow Square Root Inverted Gamma (SRIG) distribution with parameters (m_1, s_1) and (m_2, s_2) respectively. Here m_1, m_2 and s_1, s_2 are the hyperparameters to be elicited. We further assume that π_1 is a Uniform random variable with support on [0, 1]. So the joint prior distribution of θ_1, θ_2 and π_1 is the same as derived in Section 3.4.1. Combining likelihood and prior we get joint posterior distribution of θ_1, θ_2 and π_1 is the same as derived in Section 3.4.1.

4.6.1 Bayes Estimators assuming the Square Root Inverted Gamma Prior

The expectations of θ_1, θ_2 and π_1 with respect to their respective marginal posterior distributions are called the Bayes estimators (posterior means) under the square error loss function. The resulting Bayes estimates and their respective variances are given in Section 3.4.1 and Section 3.4.2 respectively.

4.6.2 Bayesian Predictive Intervals assuming the Square Root Inverted Gamma Prior

The predictive distribution of the future observation y is defined in (4.3). Where $g\left(\theta_1, \theta_2, \pi_1 \mid \mathrm{t}\right)$ is the posterior distribution given in Section 3.4.1 and $p\left(y \mid \theta_1, \theta_2, \pi_1\right)$ is the data model described in Section 4.2.

$$p\left(y|\mathrm{t}\right) = E^{-1}\sum_{k=0}^{n-r}\{\frac{\mathrm{B}\left(a_k+1, b_k\right)\Gamma\left(r_1+a_1+1\right)\Gamma\left(r_2+a_2\right)}{\left(B_{1k}+y^2\right)^{r_1+a_1+1}B_{2k}^{\;r_2+a_2}}+\frac{\mathrm{B}\left(a_k, b_k+1\right)\Gamma\left(r_1+a_1\right)\Gamma(r_2+a_2+1)}{B_{1k}^{\;r_1+a_1}\left(B_{2k}+y^2\right)^{r_2+a_2+1}}\},\ y>0 \tag{4.14}$$

$$E = \binom{n-r}{k}\ \Gamma\left(r_1+a_1\right)\ \Gamma\left(r_2+a_2\right)\sum_{k=0}^{n-r}\mathrm{B}\left(a_k+1, b_k\right)+\mathrm{B}\left(a_k, b_k+1\right)/\,B_{1k}^{\;r_1+a_1}B_{2k}^{\;r_2+a_2}\,.$$

The $(1-\alpha)100\%$ Bayesian Prediction Interval (L,U) is obtained by solving the two equations $\int_0^L p\left(y|\mathrm{t}\right)dy = \frac{\alpha}{2} = \int_U^\infty p\left(y|\mathrm{t}\right)dy$. On algebraic manipulation, these equations become

$$\begin{aligned}\frac{\alpha}{2} &= \sum_{k=0}^{n-r}\binom{n-r}{k}\Big\{\frac{(n+2)}{B_{1k}{}^{r_1+a_1}B_{2k}{}^{r_2+a_2}} - \frac{a_k}{\left(B_{1k}+L^2\right)^{r_1+a_1}B_{1k}{}^{r_2+a_2}} - \frac{b_k}{B_{1k}{}^{r_1+a_1}\left(B_{2k}+L^2\right)^{r_2+a_2}}\Big\} \\ &\times \Big[\,1/\Big(\sum_{k=0}^{n-r}\binom{n-r}{k}G_k(n+2)/B_{1k}{}^{r_1+a_1}B_{2k}{}^{r_2+a_2}\Big)\Big]\end{aligned} \tag{4.15}$$

$$\begin{aligned}\frac{\alpha}{2} &= \sum_{k=0}^{n-r}\binom{n-r}{k}\Big\{\frac{(n-r_2-k+1)}{\left(B_{1k}+U^2\right)^{r_1+a_1}B_{2k}{}^{r_2+a_2}} + \frac{(r_2+k+1)}{B_{1k}{}^{r_1+a_1}\left(B_{2k}+U^2\right)^{r_2+a_2}}\Big\} \\ &\times \Big[\,1/\Big(\sum_{k=0}^{n-r}\binom{n-r}{k}G_k\frac{(n+2)}{B_{1k}{}^{r_1+a_1}B_{2k}{}^{r_2+a_2}}\Big)\Big]\end{aligned} \tag{4.16}$$

where $B_{ik} = r_i\ \overline{t_i^2} + \left\{(2k+r-n)i + (2n-2r-3k)\right\}T^2 + b_i,\ i=1,\ 2.$

4.7 An Example based on Simulated Data

Consider a random sample of size $n=400$ from the mixture of two Rayleigh distributions with test termination time fixed at $T=15$. To generate a mixture data we make use of probabilistic mixing with probability π_1 and π_2 and take $\pi_1 = 0.375$. A uniform number u is generated 400 times and if $u < \pi_1$ the observation is taken randomly from F_1 (the Rayleigh distribution with parameter $\theta_1 = 8$) otherwise from F_2 (from the Rayleigh distribution with parameter $\theta_2 = 12$). Hence the parameters to be estimated are known to be $\theta_1 = 8$, $\theta_2 = 12$ and $\pi_1 = 0.375$. To avoid an extreme sample, we repeat this simulation 1000 times using a computer program and an averaged data summary is used to conduct computations. Actually, the observations that are greater than T are not observed during a real life test. Hence all the observations that are greater than T are considered as censored ones while calculations are conducted. In practical situations, elements generating observations are easily distinguished to be a member of either Subpopulation 1 or Subpopulation 2 of the mixture after minor inspection regarding their cause of death. The above data yields $n=400$, $r_1 \cong 146$, $r_2 \cong 197$, $\overline{t_1^2}=56.9146$, $\overline{t_2^2}=84.3234$, $r=343$ and $n-r=57$.

4.7.1 Bayesian Predictive Intervals assuming the Inverted Chi Prior

Bayesian Predictive Interval assuming the IC prior are evaluated using equations (4.5) and (4.6) for different combinations of the hyperparameters, a_1 and a_2. We used combinations of $a_1=10,\ 40,\ 70,\ 100$ and $a_2=10,\ 40, 70,\ 100$. We are only reporting here the values of L and U for those combinations of a_1 and a_2 that produce the shortest predictive intervals. Here the trend observed is "the smaller the a_1, the greater the a_2, the shorter will be the predictive intervals". Hence for $a_1=10$ and $a_2=10$, the lower and upper limits of the 95% predictive interval are found to be L = 1.54969 and U = 21.41852 respectively, while for $a_1=10$ and $a_2=100$, these are L = 1.50622 and U = 19.15067 respectively. Hence the width of the interval reduced from $\delta=19.86883$ to $\delta=17.64445$ for the said change in the hyperparameters as is clear from Table 4.1. On comparison of Tables 4.1-4.3, it can be observed that this reduction is greater than that of with IR prior and is lesser than that of with the SRIG prior.

4.7.2 Bayesian Predictive Intervals assuming the Inverted Rayleigh Prior

Bayesian Predictive Interval assuming the IR prior are evaluated using equations (4.11) and (4.12) for the said combinations of the hyperparameters. Although, there is a negligible effect on the spread of the predictive interval but a trend is again observed which is “the lower the a_2 , the higher the a_1, the narrower are the predictive intervals”. Hence for $a_1 = 10$ and $a_2 = 10$, the lower and upper limits of the 95% predictive interval are found to be L = 1.57204 and U = 21.59739 respectively, while for $a_1 = 100$ and $a_2 = 10$, they are L = 1.57798 and U = 21.57672. Hence the width of the interval reduced from $\delta = 20.02535$ to $\delta = 19.99874$ for the said change in the hyperparameters. Here, no further reduction is observed, whatever be the combination of the said hyperparameters be used as is clear from Table 4.2. A comparison of Tables 4.1-4.3 tell that this reduction is the smaller than those of the IR prior and the SRIG priors. This is an indication that the IR is the weakest prior to be used.

4.7.3 Bayesian Predictive Intervals assuming the Square Root Inverted Gamma Prior

Bayesian Predictive Interval assuming the SRIG prior are evaluated using equations (4.15) and (4.16) for the said combinations of the hyperparameters. “The lower are the values of $s_1,\ s_2$ and the higher are the values of $m_1,\ m_2$, the efficient are the predictive intervals” is the trend observed. Hence for $m_1 = 10,\ m_2 = 10,\ s_1 = 10,\ s_2 = 10$, the lower and upper limits of the 95% predictive interval were found to be L = 1.52469 and U = 21.20159 respectively, while for $m_1 = 100,\ m_2 = 100,\ s_1 = 10,\ s_2 = 10$, these were L = 1.21844 and U = 17.91917. Hence the width of the interval reduced from $\delta = 19.6769$ to $\delta = 16.70073$ for the said change in the hyperparameters as is immediate from Table 4.3. This reduction is the greater than those of the IR prior and the IC prior as depicted by Tables 4.1-4.3. This points out that the SRIG is the best among these three conjugate priors. It is interesting to note that as compared to scale parameter the change in the shape parameter of the SRIG prior has more control on the length of predictive intervals. Also, the change in upper limit is more rapid as compared to the change in the lower limit of the predictive interval.

4.7.4 The Objectivity, the Efficiency and the Partial Prior Elicitation

To chose among the prior distributions each having a functional form compatible with the likelihood, a trend can be looked for after studying the length of the predictive intervals for a number of combinations of hyperparameters. If a trend is observed on the pattern given in Table 4.3, we may proceed as follows to further filter the available prior information. Let the prior information from n experts be processed by some suitable method of prior elicitation to yield n sets of hyperparameters. Let $[(a_{1i},\ a_{2i},\ b_{1i},\ b_{2i}),\ i=1,2,3,\ldots,n]$ be the n sets of hyperparameters, one set for each expert, where n is the number of experts. Here the available prior information varies from expert to expert but the said trend observed may help us to reach a consensus by choosing a single set of hyperparameters as

$[a_1,\ a_2,\ b_1,\ b_2]=[a_1=\max(\ a_{1i}),\ a_2=\max(\ a_{2i}),\ b_1=\min(\ b_{1i}),\ b_2=\min(\ b_{2i})].$

Hence the subjective prior information turned into objective prior information with the help of the said trend observed in the possible values of the hyperparameters. Secondly, this choice of hyperparameters would obviously result in the efficient estimation and prediction. As the said trend observed narrows the possible range of the unknown hyperparameters, it can be called a sort of partial prior elicitation.

4.8 Conclusion

The Bayesian Predictive Intervals of the future observation assuming the Inverted Chi prior, the Inverted Rayleigh prior and the Square Root Inverted Gamma prior are constructed for various choices of the hyperparameters. The Square Root Inverted Gamma prior can produce more precise estimates and predictive intervals than its competitors, Inverted Chi prior and Inverted Rayleigh prior. Inverted Chi performs better than Inverted Rayleigh in terms of efficiency and precision. Inverted Rayleigh prior makes almost no improvement in the efficiency of the estimates.

If a trend can be established with the help of predictive intervals in terms of more favorable combinations of the hyperparameters, it is a sort of partial prior elicitation. The said trend adds objectivity to the subjective prior information and guaranties more precise estimation and prediction.

The role of the shape parameter of the square root inverted gamma prior is more influential as compared to its scale parameter on the breadth of predictive intervals. The changes in predictive intervals are mainly contributed by the changes in the upper limits and the changes in the lower limits are relatively slow.

Table 4.1 Bayesian Predictive Interval for different values of the hyperparameters, a_1 and a_2 of the Inverted Chi prior.

	$a_1 = 10$	$a_1 = 40$	$a_1 = 70$	$a_1 = 100$	$a_1 = 150$
$a_2 = 10$	$L = 1.585855$ $U = 22.673992$ $\delta = 21.088137$	$L = 1.527165$ $U = 22.826026$ $\delta = 21.298861$	$L = 1.477658$ $U = 22.909813$ $\delta = 21.432155$	$L = 1.434378$ $U = 22.958885$ $\delta = 21.524507$	$L = 1.371958$ $U = 23.002211$ $\delta = 21.6302$
$a_2 = 40$	$L = 1.567585$ $U = 21.781478$ $\delta = 20.213893$	$L = 1.508695$ $U = 21.958387$ $\delta = 20.449692$	$L = 1.45995$ $U = 22.052681$ $\delta = 20.592731$	$L = 1.417656$ $U = 22.106942$ $\delta = 20.689286$	$L = 1.356888$ $U = 22.154273$ $\delta = 20.797385$
$a_2 = 70$	$L = 1.551772$ $U = 20.962175$ $\delta = 19.410403$	$L = 1.491471,$ $U = 21.173564$ $\delta = 19.682093$	$L = 1.443135$ $U = 21.280562$ $\delta = 19.837427$	$L = 1.401654$ $U = 21.340768$ $\delta = 19.939114$	$L = 1.342373$ $U = 21.392575$ $\delta = 20.050202$
$a_2 = 100$	$L = 1.540310$ $U = 20.189093$ $\delta = 18.648783$	$L = 1.475632$ $U = 20.456404$ $\delta = 18.980772$	$L = 1.427213$ $U = 20.579458$ $\delta = 19.152245$	$L = 1.386349$ $U = 20.646516$ $\delta = 19.260167$	$L = 1.328387$ $U = 20.703284$ $\delta = 19.374897$
$a_2 = 150$	$L = 1.555455$ $U = 19.822219$ $\delta = 18.266764$	$L = 1.453654$ $U = 19.371259$ $\delta = 17.917605$	$L = 1.402766$ $U = 19.539050$ $\delta = 18.136284$	$L = 1.362360$ $U = 19.620528$ $\delta = 18.258168$	$L = 1.306194$ $U = 19.686750$ $\delta = 18.380556$

$L =$ the lower limit, $U =$ the upper limit, $\delta =$ the length of the predictive interval

Table 4.2 Bayesian Predictive Interval for different values of the hyperparameters, a_1 and a_2 of the Inverted Rayleigh prior.

	$a_1 = 10$	$a_1 = 40$	$a_1 = 70$	$a_1 = 100$	$a_1 = 150$
$a_2 = 10$	$L = 1.609639$ $U = 22.870927$ $\delta = 21.261288$	$L = 1.611813$ $U = 22.863517$ $\delta = 21.251704$	$L = 1.613980$ $U = 22.856029$ $\delta = 21.242049$	$L = 1.616140$ $U = 22.848464$ $\delta = 21.232324$	$L = 1.619725$ $U = 22.835686$ $\delta = 21.215961$
$a_2 = 40$	$L = 1.609883$ $U = 22.882947$ $\delta = 21.273064$	$L = 1.612057$ $U = 22.875558$ $\delta = 21.263501$	$L = 1.614224$ $U = 22.868092$ $\delta = 21.253868$	$L = 1.616385$ $U = 22.860549$ $\delta = 21.244164$	$L = 1.619970$ $U = 22.847808$ $\delta = 21.227838$
$a_2 = 70$	$L = 1.610127$ $U = 22.894957$ $\delta = 21.28483$	$L = 1.612301$ $U = 22.887589$ $\delta = 21.275288$	$L = 1.614469$ $U = 22.880144$ $\delta = 21.265675$	$L = 1.616629$ $U = 22.872622$ $\delta = 21.255993$	$L = 1.620215$ $U = 22.859919$ $\delta = 21.239704$
$a_2 = 100$	$L = 1.610370$ $U = 22.906956$ $\delta = 21.296586$	$L = 1.612545$ $U = 22.899609$ $\delta = 21.287064$	$L = 1.614713$ $U = 22.892185$ $\delta = 21.277472$	$L = 1.616874$ $U = 22.884685$ $\delta = 21.267811$	$L = 1.620460$ $U = 22.872018$ $\delta = 21.251558$
$a_2 = 150$	$L = 1.610776$ $U = 22.926931$ $\delta = 21.316155$	$L = 1.612951$ $U = 22.919619$ $\delta = 21.306668$	$L = 1.615119$ $U = 22.912230$ $\delta = 21.297111$	$L = 1.617281$ $U = 22.904766$ $\delta = 21.287485$	$L = 1.620869$ $U = 22.892159$ $\delta = 21.27129$

$L =$ the lower limit, $U =$ the upper limit, $\delta =$ the length of the predictive interval

Table 4.3 Bayesian Predictive Interval for different values of the hyper parameters m_1, m_2 and s_1, s_2 of the Square root inverted Gamma prior.

	$m_1 = 10$ $m_2 = 10$	$m_1 = 40$ $m_2 = 40$	$m_1 = 70$ $m_2 = 70$	$m_1 = 100$ $m_2 = 100$	$m_1 = 150$ $m_2 = 150$
$s_1 = 10$ $s_2 = 10$	$L = 1.559359$ $U = 22.433777$ $\delta = 20.874418$	$L = 1.4239907$ $U = 21.066623$ $\delta = 19.642632$	$L = 1.321602$ $U = 19.871959$ $\delta = 18.550357$	$L = 1.239922$ $U = 18.840033$ $\delta = 17.60011$	$L = 1.133306$ $U = 17.411904$ $\delta = 16.278598$
$s_1 = 40$ $s_2 = 40$	$L = 1.561678$ $U = 22.439364$ $\delta = 20.877686$	$L = 1.426038$ $U = 21.074081$ $\delta = 19.648043$	$L = 1.323459$ $U = 19.880053$ $\delta = 18.556594$	$L = 1.241641$ $U = 18.848243$ $\delta = 17.606602$	$L = 1.134858$ $U = 17.419886$ $\delta = 16.285028$
$s_1 = 70$ $s_2 = 70$	$L = 1.563991$ $U = 22.444904$ $\delta = 20.880913$	$L = 1.428080$ $U = 21.081505$ $\delta = 19.653425$	$L = 1.325312$ $U = 19.888125$ $\delta = 18.562813$	$L = 1.243356$ $U = 18.856438$ $\delta = 17.613082$	$L = 1.136409$ $U = 17.427860$ $\delta = 16.291451$
$s_1 = 100$ $s_2 = 100$	$L = 1.566300$ $U = 22.450397$ $\delta = 20.884097$	$L = 1.430119$ $U = 21.088896$ $\delta = 19.658777$	$L = 1.327161$ $U = 19.896173$ $\delta = 18.569012$	$L = 1.245067$ $U = 18.864617$ $\delta = 17.61955$	$L = 1.137956$ $U = 17.435824$ $\delta = 16.297868$
$s_1 = 150$ $s_2 = 150$	$L = 1.570135$ $U = 22.459449$ $\delta = 20.889314$	$L = 1.433508$ $U = 21.101137$ $\delta = 19.667629$	$L = 1.330234$ $U = 19.909537$ $\delta = 18.579303$	$L = 1.247911$ $U = 18.878214$ $\delta = 17.630303$	$L = 1.140527$ $U = 17.449078$ $\delta = 16.308551$

$L =$ the lower limit, $U =$ the upper limit, $\delta =$ the length of the predictive interval

CHAPTER 5

ESTIMATION AND APPLICATION OF THE PARETO MIXTURE

5.1 Introduction

Pareto model is often used for investigating the distribution of many empirical phenomena including personal incomes, city population sizes, the sizes of firms and the lifetimes. In this chapter, a lifetime population of certain objects is assumed to be composed of $1 < k < \infty$ subgroups mixed together in an unknown proportion. The random observations taken from this population are supposed to be characterized by one of the k distinct unknown members of a Pareto distribution. So the k-component mixture of the Pareto distribution is recommended to model this population provided the data is not available on the individual components rather on the mixture only.

Abdel-All et al. (2003) discussed geometrical properties of Pareto distribution. Ismail (2004) presented a simple estimator for the shape parameter of the Pareto distribution. Bhat (2005) focus Bayes estimation and reliability functions for a two component mixture of Pareto lifetime distributions. Sankaran and Nair (2005) discussed the properties of finite mixture of Pareto distributions in the context of income analysis. Nadarajah and Kotz (2005) focused the information matrix for a mixture of two Pareto distributions. A truncated Pareto distribution is studied by Ali and Nadarajah (2006).

The uninformative and informative Bayes estimators of $2k$-1 parameters of the k-components Pareto mixture are derived. In this chapter, the said Pareto mixture, the likelihood and the system of three non-linear equations, required to be solved iteratively for the computations of maximum likelihood estimates, are developed in Section 2. The components of the information matrix are constructed as well. In Section 3, the expressions for the Bayes estimators and their variances are presented along with the expression for the posterior predictive distribution and the equations required for finding the predictive intervals. The complete sample expressions for the ML and Bayes estimators and variances are derived in Section 4. In Section 5, a comprehensive simulation scheme consisting of a large number of parameter points is accomplished to highlight the properties and behavior of the estimates in terms of sample size, censoring rate, parameters size and the proportion of the components of the mixture. A real life data set is used to evaluate and compare the Bayes estimates in Section 6. Some interesting comparisons and properties of the proposed estimates are discussed in Section 7 as concluding remarks. The Bayes estimates are evaluated under the squared error loss function.

5.2 The Maximum Likelihood Estimates for Censored Data

A finite mixture density function with a known integer number ($k > 1$), of component densities of specified parametric form but with k unknown parameters, $\alpha_i,\ i = 1, 2, \ldots, k$ and with k unknown mixing weights, $\pi_i,\ i = 1, 2, \ldots, k$ where $\pi_k = 1 - \sum_{i=1}^{k-1} \pi_i$ is defined as in (2.1). The following Pareto distribution is assumed for the k components of the mixture.

$$f_i(x) = \alpha_i\ x_{ij}^{-(\alpha_i+1)},\ i = 1, 2, ..., k\ ;\ j = 1, 2, 3, ..., r_i;\ 0 < \alpha_i < \infty\ ;\ 1 \leq x_{ij} < \infty$$

And the corresponding Survivor functions are $S_i(x) = x_{ij}^{-\alpha_i}$, $i = 1, 2, ..., k$. The corresponding mixture distribution function is given by $F(x) = \sum_{i=1}^{k} \pi_i F_i(x)$. The sampling scheme and likelihood function (2.2) of Section 2.3 is considered with a Type-IV sample of size n units from the Type-I mixture model described above under ordinary type-I, right censoring. Here $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x}_k) = (x_{11}, x_{12}, \ldots, x_{1r_1}, x_{21}, x_{22}, \ldots, x_{2r_2}, \ldots, x_{k1}, x_{k2}, \ldots, x_{kr_k})$ is data while the $2k-1$ parameters are $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ and $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k)$, $\pi_k = 1 - \sum_{i=1}^{k-1} \pi_i$.

$$L(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x}) \propto \{\prod_{j=1}^{r_1} (\pi_1 \ \alpha_1 \ x_{1j}^{-(\alpha_1+1)})\}\{\prod_{j=1}^{r_2} (\pi_2 \ \alpha_2 \ x_{2j}^{-(\alpha_2+1)})\}\ldots\{\prod_{j=1}^{r_k} (\pi_k \ \alpha_k \ x_{kj}^{-(\alpha_k+1)})\}$$
$$\times\{(\pi_1 T^{-\alpha_1} + \pi_2 T^{-\alpha_2} + \ldots + \pi_k T^{-\alpha_k})^{n\text{-}r}\}$$

$$L(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x}) \propto \left\{ \prod_{i=1}^{k} \prod_{j=1}^{r_i} \pi_i \ \alpha_i \ x_{ij}^{-(\alpha_i+1)} \right\} \left\{ \prod_{i=1}^{k} \pi_i T^{-\alpha_i} \right\}^{n\text{-}r} \tag{5.2}$$

The likelihood function in (5.2) can take the following form

$$L(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x}) \propto \sum_{l}^{H_{n-r}^k} \binom{n-r}{k_1, k_2, \ldots, k_k} \left\{ \prod_{i=1}^{k} \pi_i^{r_i+k_i} \right\} \left\{ \prod_{i=1}^{k} \alpha_i^{r_i} \right\} \exp\left[-\sum_{i=1}^{k} \left\{ \alpha_i \sum_{j=1}^{r_i} \left(\ln x_{ij} + k_i \ln T \right) \right\} \right] \tag{5.3}$$

Here H_{n-r}^k denotes the number of distinct terms in the expansion of a multinomial as mentioned in Section 2.3. Maximum Likelihood Estimates of $\boldsymbol{\alpha}$ and of $\boldsymbol{\pi}$ are obtained by solving the system of $2k-1$ nonlinear equations (5.4) obtained by setting first order derivatives of the natural log of the likelihood (5.2) to zero.

$$\frac{\partial l}{\partial \alpha_i} = \frac{r_i}{\alpha_i} - \sum_{j=1}^{r_i} \ln(x_{ij}) - \frac{(n-r) p_i T^{-\alpha_i} \ln(T)}{\sum_{i=1}^{k} \pi_i T^{-\alpha_i}} = 0, \ i = 1, 2, \ldots, k \tag{5.4}$$

$$\frac{\partial l}{\partial \pi_i} = \frac{r_i}{\pi_i} - \frac{r_k}{\pi_k} + \frac{(n-r)(T^{-\alpha_i} - T^{-\alpha_k})}{\sum_{i=1}^{k} \pi_i T^{-\alpha_i}} = 0, \ i = 1, 2, \ldots, k-1 \tag{5.5}$$

Solving the system of nonlinear equations (5.4)-(5.5) with the help of an iterative numerical procedure, the ML estimates can be found. The information matrix with the following elements can help find variances as discussed in Section 2.3.

$$E\left(\frac{-\partial^2 l}{\partial \alpha_i^2}\right) = \frac{\left(\prod_{i=1}^{k} \pi_i\right)(n-r)T^{-\left(\sum_{i=1}^{k} \alpha_i\right)} (\ln T)^2}{\left(\sum_{i-1}^{k} \pi_i T^{-\alpha_i}\right)^2} - \frac{r_i}{\alpha_i^2}, \; i = 1, 2, \ldots, k \tag{5.6}$$

$$E\left(\frac{-\partial^2 l}{\partial \pi_i^2}\right) = -\frac{r_i}{\pi_i^2} - \frac{r_k}{\pi_k^2} - \frac{(n-r)(T^{-\alpha_i} - T^{-\alpha_k})^2}{\left(\sum_{i-1}^{k} \pi_i T^{-\alpha_i}\right)^2}, i = 1, 2, \ldots, k-1 \tag{5.7}$$

$$E\left(\frac{-\partial^2 l}{\partial \alpha_i \partial \alpha_j}\right) = E\left(\frac{-\partial^2 l}{\partial \alpha_j \partial \alpha_i}\right) = -\frac{\left(\prod_{i=1}^{k} \pi_i\right)(n-r)T^{-\left(\sum_{i=1}^{k} \alpha_i\right)}}{(\ln T)^{-2}\left(\sum_{i-1}^{k} \pi_i T^{-\alpha_i}\right)^2}, \; j > i = 1, 2, \ldots, k \tag{5.8}$$

$$E\left(\frac{-\partial^2 l}{\partial \alpha_i \partial \pi_j}\right) = E\left(\frac{-\partial^2 l}{\partial \pi_i \partial \alpha_j}\right) = -\frac{(n-r)T^{-\left(\sum_{i=1}^{k} \alpha_i\right)} (\ln T)}{\left(\sum_{i-1}^{k} \pi_i T^{-\alpha_i}\right)^2}, \quad j = 1, 2, \ldots, k-1; \; i = 1, 2, \ldots, k. \tag{5.9}$$

$$E\left(\frac{-\partial^2 l}{\partial \pi_i \partial \pi_j}\right) = \frac{r_k}{\pi_k^2} - \frac{(n-r)(T^{-\alpha_i} - T^{-\alpha_k})(T^{-\alpha_j} - T^{-\alpha_k})}{\left(\sum_{i-1}^{k} \pi_i T^{-\alpha_i}\right)^2}, \quad i = 1, 2, \ldots, k-2; \; j = i+1 = 2, \ldots, k-1 \tag{5.10}$$

5.3 Bayes Estimators assuming the Conjugate Prior

Here Gamma prior is used as a conjugate prior. Let $\alpha_i \sim Gamma(m_i, s_i) \;\forall\; i = 1, 2, ..., k$ and $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_k) \sim Dirichlet(1,1,\ldots,1)$, so $g_i(\alpha_i) \propto \alpha_i^{m_i - 1} e^{-\alpha_i s_i}$, $i = 1, 2, ..., k$. Assuming independence, the joint prior is incorporated with the Likelihood (5.3) to give the joint posterior and then marginal posterior densities are obtained. The expressions for the Bayes estimators under the squared error loss function are given by the respective expectations under the marginal posterior distributions. The following are the expressions for the k estimators of the k parameters of the k component densities of the mixture.

$$\hat{\alpha}_i = \Omega \sum_{k_1,k_2,\dots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\dots,k_k} \mathrm{B}(r_1+k_1+1, r_2+k_2+1,\dots,r_k+k_k+1)$$
$$\times \frac{\Gamma(r_i+m_i+1)}{\left\{s_i+\sum_{j=1}^{r_i}\ln(x_{ij})+k_i\ln(T)\right\}^{r_1+m_1+1}} \prod_{j\neq i} \frac{\Gamma(r_j+m_j)}{\left\{s_j+\sum_{j=1}^{r_j}\ln(x_{ij})+k_j\ln(T)\right\}^{r_j+m_j}},\ i=1,2,\dots,k$$

The following are the k estimators of the k mixing proportions of the finite mixture.

$$\hat{\pi}_i = \Omega \sum_{k_1,k_2,\dots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\dots,k_k} \mathrm{B}(r_1+k_1+1,\dots,r_i+k_i+2,\dots,r_k+k_k+1) \prod_{j=1}^{k} \frac{\Gamma(r_j+m_j)}{\left\{s_j+\sum_{j=1}^{r_i}\ln(x_{ij})+k_j\ln(T)\right\}^{r_j+m_j}},\ i=1,2,\dots,k$$

where $\Omega^{-1} = \sum_{k_1,k_2,\dots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\dots,k_k} \mathrm{B}(r_1+k_1+1,\dots,r_k+k_k+1) \prod_{i=1}^{k} \frac{\Gamma(r_i+m_i)}{\left\{s_i+\sum_{j=1}^{r_i}\ln(x_{ij})+k_i\ln(T)\right\}^{r_i+m_i}}$ and

$\mathrm{B}(\boldsymbol{\tau})$, $\boldsymbol{\tau}=(\tau_1,\tau_2,\dots,\tau_k)$ is the multinomial Beta function having k arguments and can be expressed in terms of Gamma functions $\Gamma(\tau_i),\ i=1,2,\dots,k$. The expressions for the variances of the Bayes estimators can be evaluated on the same lines as under.

$$V(\hat{\alpha}_i) = \Omega \sum_{k_1,k_2,\dots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\dots,k_k} \mathrm{B} \frac{\Gamma(r_i+m_i+2)}{\left\{s_i+\sum_{j=1}^{r_i}\ln(x_{ij})+k_i\ln(T)\right\}^{r_1+m_1+2}} \prod_{j\neq i} \frac{\Gamma(r_j+m_j)}{\left\{s_j+\sum_{j=1}^{r_j}\ln(x_{ij})+k_j\ln(T)\right\}^{r_j+m_j}}$$
$$-\left\{\Omega \sum_{k_1,k_2,\dots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\dots,k_k} \mathrm{B} \frac{\Gamma(r_i+m_i+1)}{\left\{s_i+\sum_{j=1}^{r_i}\ln(x_{ij})+k_i\ln(T)\right\}^{r_1+m_1+1}} \prod_{j\neq i} \frac{\Gamma(r_j+m_j)}{\left\{s_j+\sum_{j=1}^{r_j}\ln(x_{ij})+k_j\ln(T)\right\}^{r_j+m_j}}\right\}^2,$$
$$i=1,2,\dots,k.$$

Here B stands for $\mathrm{B}(r_1+k_1+1,\dots,r_k+k_k+1)$.

$$V(\hat{\pi}_i) = \Omega \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_i+k_i+3,\ldots,r_k+k_k+1) \prod_{j=1}^{k} \Gamma(r_j+m_j) \left\{ s_j + \sum_{j=1}^{r_i} \ln(x_{ij}) + k_j \ln(T) \right\}^{-(r_j+m_j)}$$

$$-\left\{ \Omega \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_i+k_i+2,\ldots,r_k+k_k+1) \prod_{j=1}^{k} \Gamma(r_j+m_j) \left\{ s_j + \sum_{j=1}^{r_i} \ln(x_{ij}) + k_j \ln(T) \right\}^{-(r_j+m_j)} \right\}^2,$$

$$i = 1,2,\ldots,k.$$

5.3.1 Bayes Estimators assuming the Uninformative Priors

The simplest and the oldest uninformative priors are the Uniform and the Jeffreys.

5.3.1.1 Bayes Estimators assuming the Uniform Prior

Let us assume a state of ignorance i.e., $\alpha_i,\ i=1,2,\ldots,k$ is uniformly distributed over $(0,\infty)$ and $\pi_i,\ i=1,2,\ldots,k$ are uniformly distributed over [0,1]. Hence $f_i(\alpha_i) \propto k_i,\ \ 0<\alpha_i<\infty$ $\forall\ i=1,2,...,k$ and $\boldsymbol{\pi} = (\pi_1,\pi_2,\ldots,\pi_k) \sim Dirichlet(1,1,\ldots,1).$ Assuming independence we have an improper joint prior that is proportional to a constant and is incorporated with the likelihood (5.3) to yield a proper joint posterior distribution. The respective marginal posterior distributions yield the Bayes estimators under the squared error loss function. The expressions for the estimators assuming Uniform prior are obtained by replacing $m_i = 2,\ i=1,2,\ldots,k$ and $s_i = 0,\ i=1,2,\ldots,k$ equations of Section 5.3.

5.3.1.2 Bayes Estimators assuming the Jeffreys Prior

For the components of the Pareto mixture model given in Section 2, based on the definition of Section 2.5.2, the Jeffreys priors are assumed as $g_i(\alpha_i) \propto 1/\alpha_i,\ \ 0<\alpha_i<\infty,\ \forall$ $i=1,2,...,\ k$ and $\boldsymbol{\pi} = (\pi_1,\pi_2,\ldots,\pi_k) \sim Dirichlet(1,1,\ldots,1).$ Assuming independence, the joint prior is incorporated with the likelihood (5.3) to have the joint posterior. Then the respective marginal posterior distributions give the Bayes estimators under the squared error loss function. The expressions for the Bayes estimators using the Jeffreys prior are obtained by replacing $m_i = 1,\ i = 1, 2, \ldots, k$ and $s_i = 0,\ i = 1, 2, \ldots, k$ in equations of Section 5.3.

5.3.2 The Posterior Predictive Distribution and Predictive Intervals

Equation (4.3) defines the posterior predictive distribution. Assuming $k = 2$, the predictive distribution of the future observation y given the data $\mathbf{x}$ is

$$p(y|\mathbf{x}) = \Omega \sum_{k=0}^{n-r} \binom{n-r}{k} \tfrac{1}{y} \{\mathrm{B}(n-r_2-k+2, r_2+k+1) \frac{\Gamma(r_1+m_1+1)\Gamma(r_2+m_2)}{(s_1+A_{1k})^{r_1+\mu_1+1}(s_2+A_{2k})^{r_2+m_2}} + \mathrm{B}(n-r_2-k+1, r_2+k+2) \frac{\Gamma(r_1+m_1)\Gamma(r_2+m_2+1)}{(s_1+A_{1k})^{r_1+\mu_1}(A_{2k}+s_2+\ln\ y)^{r_2+m_2+1}}\},\ y>0 \tag{5.10}$$

The $(1-\alpha)100\%$ Bayesian Prediction Interval (L, U) is obtained by solving the two equations $\int_1^L p(y|\mathbf{x})\ dy = \frac{\alpha}{2} = \int_U^\infty p(y|\mathbf{x})\ dy$. On necessary manipulation these equations become

$$\Omega \sum_{k=0}^{n-r} \binom{n-r}{k} \{\frac{\mathrm{B}(n-r_2-k+2, r_2+k+1)}{(s_2+A_{2k})^{r_2+m_2}} (\frac{1}{(s_1+A_{1k})^{r_1+m_1}} - \frac{1}{(s_1+A_{1k}+\ln L)^{r_1+m_1}}) + \frac{\mathrm{B}(n-r_2-k+1, r_2+k+2)}{(s_1+A_{1k})^{r_1+m_1}} (\frac{1}{(s_2+A_{2k})^{r_2+m_2}} - \frac{1}{(s_2+A_{2k}+\ln L)^{r_2+m_2}})\} - \frac{\alpha}{2} = 0 \tag{5.11}$$

$$\Omega \sum_{k=0}^{n-r} \binom{n-r}{k} \{\frac{\mathrm{B}(n-r_2-k+2, r_2+k+1)}{(s_2+A_{2k})^{r_2+m_2}(s_1+A_{1k}+\ln U)^{r_1+m_1}} + \frac{\mathrm{B}(n-r_2-k+1, r_2+k+2)}{(s_1+A_{1k})^{r_1+m_1}(s_2+A_{2k}+\ln U)^{r_2+m_2}}\} - \frac{\alpha}{2} = 0 \tag{5.12}$$

where $\Omega^{-1} = \sum_{k=0}^{n-r} \binom{n-r}{k} \left\{ \prod_{i=1}^{2} \frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}} \right\} \{\mathrm{B}(n-r_2-k+2, r_2+k+1) + \mathrm{B}(n-r_2-k+1, r_2+k+2)\}$.

These predictive intervals when evaluated for a number of combinations of the hyperparameters can help locate a range of hyperparameters that may lead to the Informative Bayes estimates having lesser variances than the uninformative Bayes estimates. Saleem and Aslam (2008a) used predictive intervals for the Rayleigh mixture to discuss precision of Bayes estimates in terms of hyperparameters. Also, a sort of objectivity can be added to the prior information provided by a number of experts provided a trend is observed for the narrower predictive intervals in terms of the hyperparameters.

5.4 The Complete Sample Expressions

Under the conditions given in Section 2.6, the expressions for the Bayes estimators and their variances are simplified as given in Tables 5.1-5.2. The comments regarding amount of information, computational ease and simplification quoted in Section 2.6 also applies here.

5.5 A Simulation Study

We take random samples of sizes $n = 50,\ 100,\ 200,\ 300$ from the two component mixture of Pareto distribution with $(\alpha_1, \alpha_2) = \{ (0.5,1.5), (1.0, 4.0), (2.5, 0.5), (4.0,1.0) \}$ and $\pi = 0.25$, 0.40. Censoring rates assumed are $C = 10\%,\ 20\%$. To generate a mixture data we make use of probabilistic mixing. A uniform number u is generated n times and if $u < \pi$ the observation is taken randomly from F_1 (the Pareto distribution with parameter α_1) otherwise from F_2 (from the Pareto distribution with parameter α_2). Remaining details of the simulation study is the same as mentioned in Section 2.7. Some interesting properties of the Bayes estimates are highlighted in Tables 5.3-5.4 while a comparison of the estimates is summarized in Table 5.5.

5.5.1 WinBUGS Code for computations using Gibbs Sampling

The full Bayesian model can be fit in WinBUGS. Here's an example with $n = 15,\ r = 10,\ 5$ observations come from component 1 and 5 come from component 2.

Here zi is the component indicator and xi are data. The code is given bellow which makes use of the Exponential-Pareto connection described in Section 5.6.

```
model {
        for (i in 1:n) {cens.log[i] <- log(cens[i]); z[i] ~ dcat(p[1:k])}
        for (i in 1:r) {y[i] <- log(x[i]); y[i] ~ dexp(alpha[z[i]])}
        p[1:k] ~ ddirch (ones[1:k])
        for (i in 1:k) {alpha[i] ~ dgamma(1, 1)}
        for (i in (r+1):n) {x[i] ~ dexp(alpha[z[i]]) I(cens.log[i],)}


list (ones = c(1, 1),
cens = c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 10, 10, 10, 10, 10),

x = c(2.0, 2.5, 3.0, 2.7, 1.5, 5.5, 8.0, 7.2, 9.4, 8.7, NA, NA, NA, NA, NA),
z = c(1, 1, 1, 1, 1, 2, 2, 2, 2, 2, NA, NA, NA, NA, NA))
```

5.6 A Real Life Example

Mendenhall and Hader mixture data $\mathbf{t} = (t_{11},\ t_{12}, ...,\ t_{1r_1},\ t_{21},\ t_{22}, ...,\ t_{2r_2})$ consists of hours to failure for ARC-1 VHF radio transmitter receivers. Radio transmitters that had not failed by 630 hours were removed from the aeroplanes anyway, so these are Type-I right censored data at 630 hours. Inspection of failed units allowed the engineers to allocate the failed units to two different subpopulations. Mendenhall and Hader fitted Exponential distributions to this data. The transformation $x = \exp(t)$ of an Exponential distribution yields a Pareto distribution. This transformation allowed us to use Mendenhall and Hader (MH) data set for our analysis with the obvious transformation of the data. It is interesting to note that despite the transformation

$x = \exp(t)$ almost no major computations are required to have the data summary required to evaluate the estimates. For instance, $\sum_{j=1}^{r_1} \ln(x_{1j}) = \sum_{j=1}^{r_1} t_{1j} = 20458$ and $\sum_{j=1}^{r_2} \ln(x_{2j}) = \sum_{j=1}^{r_2} t_{2j} = 50056$, $n = 369,\ r_1 = 107,\ r_2 = 218,\ r = r_1 + r_2 = 325,\ n - r = 44$. Other sample characteristics required are also made available easily. Pareto mixture parameters $(\alpha_1,\ \alpha_2,\ \pi_1)$ are estimated using estimators derived in Section 5.3. The Bayes (Jeffreys) estimates of the MH mixture lifetime parameters, after an obvious re-parameterization as evident from the functional form of the component densities of the mixture given in Section 5.2, are found to be $(\hat{\alpha}_{11},\ \hat{\alpha}_{22}) = (1/\hat{\alpha}_1,\ 1/\hat{\alpha}_2) = (237, 333)$ where $\hat{\alpha}_1 = 0.00422,\ \hat{\alpha}_2 = 0.00300$ (correct to five decimal places) are the Bayes (Jeffreys) estimates of Pareto mixture parameters with $SE(\hat{\alpha}_1) = 0.000574$, $SE(\hat{\alpha}_2) = 0.00023$ respectively. The variances of the lifetime parameters of the MH mixture are computed as $SE(\hat{\alpha}_{11}) = 33.8015,\ SE(\hat{\alpha}_{22}) = 25.6659$. The estimate of the proportion parameter of the MH mixture is $\hat{\pi} = 0.313$ with $SE(\hat{\pi}) = 0.02642$. This is encouraging to note that the estimates are much greater than the two respective subgroup sample means i.e., $\overline{t_1} = 191.2 << 237,\ \overline{t_2} = 229.6 << 333$ which happens in the right censoring situations. Also the proposed estimates presented here are slightly more precise than the estimates presented in Sinha (1998). A comparison of the various estimates based on the Mendenhall and Hader (1958) mixture data is displayed in Table 5.6.

5.7 Conclusion

The simulation study depicts some interesting properties of the Bayes estimates. The properties of the estimates are highlighted in terms of sample sizes, sizes of mixing proportion parameters, sizes of the component densities parameters and censoring rates.

The estimates of the parameters of the component densities are generally over-estimated with a few exceptions in case of the second component. The extent of over-estimation is higher in case of the estimates of the first component density parameter. On the other hand the estimates of the mixing proportion parameter are observed to be under-estimated with a few exceptions. It is interesting to note that the estimates seem to approach the true parameter values with the increase in sample size.

Another interesting remark concerning the variances of the estimates of the component densities' parameters is that increasing (decreasing) the proportion of a component in the mixture reduces (increases) the variance of the estimate of the corresponding component density parameter. The variances of estimates of all three mixture parameters reduce with an increase in sample size. However, the variances of the estimates of the component densities' parameters seem to be quite large in cases when the value of the parameters are large and quite small for relatively smaller values of the parameters.

Table 5.1 The complete sample expressions for the Bayes (Uniform)*, Bayes (Jeffreys) and ML estimators as $T \to \infty$.

Parameters	Bayes Estimators (Uniform)	Bayes Estimators (Jeffreys)	ML Estimators
α_1	$\frac{(n_1+1)}{(\sum \ln x_{1j})}$	$\frac{n_1}{(\sum \ln x_{1j})}$	$\frac{n_1}{(\sum \ln x_{1j})}$
α_2	$\frac{(n_2+1)}{(\sum \ln x_{2j})}$	$\frac{n_2}{(\sum \ln x_{2j})}$	$\frac{n_2}{(\sum \ln x_{2j})}$
π_1	$\frac{(n_1+1)}{(n+2)}$	$\frac{(n_1+1)}{(n+2)}$	$\frac{n_1}{n}$

*Bayes (Uniform) means the Bayes estimates assuming the Uniform prior.

Table 5.2 The complete sample expressions for the variances of the Bayes (Uniform), Bayes (Jeffreys) and ML estimators as $T \to \infty$.

Parameters	Variances of Bayes Estimators (Uniform prior)	Variances of Bayes Estimators (Jeffreys prior)	Variances of ML Estimators
α_1	$\frac{n_1+1}{(\sum \ln x_{1j})^2}$	$\frac{n_1}{(\sum \ln x_{1j})^2}$	$\frac{n_1}{(\sum \ln x_{1j})^2}$
α_2	$\frac{n_2+1}{(\sum \ln x_{2j})^2}$	$\frac{n_2}{(\sum \ln x_{2j})^2}$	$\frac{n_2}{(\sum \ln x_{2j})^2}$
π_1	$\frac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\frac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\frac{n_1 n_2}{n^3}$

Table 5.3 Bayes (Jeffreys) estimates of Pareto mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 0.5,\ \alpha_2 = 1.5,\ \pi_1 = 0.25,\ 0.40$ and censoring rates, $C = 10\%,\ 20\%$

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**10% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.5, 1.5, 0.25)	50	0.63926(0.36444)	1.51497(0.31810)	0.25334(0.06394)
	100	0.56961(0.21440)	1.50283(0.22687)	0.24911(0.04824)
	200	0.53881(0.13534)	1.49248(0.15310)	0.24785(0.03389)
	300	0.517657(0.0944522)	1.49677(0.129143)	0.249782(0.027997)
(0.5, 1.5, 0.40)	50	0.55870(0.22151)	1.50016(0.34748)	0.39219(0.06957)
	100	0.52027(0.11974)	1.50942(0.22711)	0.39877(0.05048)
	200	0.50815(0.07325)	1.50428(0.16261)	0.39890(0.03515)
	300	0.5054(0.0566698)	1.50235(0.13542)	0.398518(0.0289886)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.5, 1.5, 0.25)	50	0.75179(0.55476)	1.58108(0.41216)	0.25461(0.06546)
	100	0.67048(0.33805)	1.50765(0.26856)	0.24579(0.05274)
	200	0.60366(0.20462)	1.49444(0.20106)	0.24513(0.04295)
	300	0.573052(0.165515)	1.49599(0.175607)	0.244256(0.0375923)
(0.5, 1.5, 0.40)	50	0.65813(0.35527)	1.48759(0.38060)	0.38131(0.07868)
	100	0.57586(0.18428)	1.48787(0.30622)	0.38774(0.05822)
	200	0.53164(0.11201)	1.48527(0.20755)	0.39231(0.04185)
	300	0.519594(0.0837556)	1.49218(0.18015)	0.392906(0.0359372)

Table 5.4 Bayes estimates (Jeffreys) of Pareto mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 1.0,\ \alpha_2 = 4.0,\ \pi_1 = 0.25,\ 0.40$ and censoring rates, $C = 10\%,\ 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	10% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(1.0, 4.0, 0.25)	50	1.33812(1.06763)	4.06876(0.82902)	0.25284(0.06420)
	100	1.10941(0.43288)	4.03982(0.58555)	0.24878(0.04572)
	200	1.03321(0.22509)	3.991(0.40274)	0.25084(0.03235)
	300	1.02742(0.185037)	3.97965(0.327512)	0.2498889(0.0265039)
(1.0, 4.0, 0.40)	50	1.08311(0.45645)	4.05159(0.84810)	0.39995(0.07131)
	100	1.02705(0.23688)	4.01479(0.60296)	0.40027(0.05200)
	200	1.00512(0.14181)	4.01193(0.40115)	0.39855(0.03521)
	300	1.00049(0.108703)	3.99171(0.331273)	0.399337(0.0285543)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(1.0, 4.0, 0.25)	50	1.81191(2.62688)	4.09448(0.96668)	0.245927(0.06863)
	100	1.38455(0.97628)	3.98446(0.69090)	0.24081(0.05626)
	200	1.18649(0.45700)	3.96388(0.52966)	0.24374(0.04474)
	300	1.12527(0.322429)	3.98573(0.443237)	0.245717(0.0358915)
(1.0, 4.0, 0.40)	50	1.29493(0.66903)	3.94695(0.97085)	0.38350(0.07792)
	100	1.10029(0.35337)	3.97788(0.71211)	0.39131(0.05627)
	200	1.04657 (0.20215)	3.95922(0.50577)	0.39668(0.04054)
	300	1.02093(0.1474)	3.98558(0.423434)	0.39626(0.0326351)

Table 5.5. A comparison of the Bayes (Uniform), Bayes (Jeffreys) and Bayes (Gamma)* estimates of Pareto mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 0.5,\ \alpha_2 = 1.5,\ \pi_1 = 0.25$ and censoring rate, $C = 10\%$.

Uniform			
n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}$
50	0.79498(0.53338)	1.5329(0.33227)	0.24861(0.06751)
100	0.62560(0.25225)	1.51213(0.23418)	0.24692(0.04991)
200	0.56126(0.14727)	1.49772(0.15623)	0.24695(0.03457)
Jeffreys			
n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}$
50	0.63926(0.36444)	1.51497(0.31810)	0.25334(0.06394)
100	0.56961(0.21440)	1.50283(0.22687)	0.24911(0.04824)
200	0.53881(0.13534)	1.49248(0.15310)	0.24785(0.03389)
Gamma			
n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}$
50	0.56053(0.24039)	1.40222(0.26659)	0.25142(0.06208)
100	0.54739(0.17978)	1.44430(0.20676)	0.24765(0.04748)
200	0.53243(0.12688)	1.46210(0.14604)	0.24684(0.03364)

*Bayes (Gamma) means the Bayes estimates assuming the Gamma prior.

Table 5.6 A comparison of the various estimates based on the Mendenhall and Hader (1958) mixture data.

parameters	Mendenhall and Hader (1958)	Sinha (1998)	Saleem and Aslam (2010)	Saleem and Aslam (proposed)
α_1	234.234	241.227 (33.957)	200 (17.7513)	237 (33.8015)
α_2	335.664	334.845 (25.956)	250 (14.6994)	333 (25.6659)
π_1	0.3098	0.313 (0.0265)	0.331 (0.0260)	0.313 (0.02642)

The effect of an increase in censoring rate on the estimates of the component density parameters is observed in terms of an increase in the extent of over-estimation with some rare exceptions in case of estimates of the second component density parameter. But the effect of an increase in censoring rate on the estimates of the mixing proportion parameter is observed in terms of an increase in the extent of under-estimation except a few exceptional cases when its over-estimation turns into a slight under estimation with the increase in censoring rate. However, this is interesting to note that the variances of all the estimates of the component density and mixing proportion parameters are increased with an increase in the censoring rate. As the cut off sensor value gets infinitely large, the complete sample expressions for the estimators and variances are greatly simplified. Also variances of the complete sample estimates are expected to be reduced further as is clear from the effect of censoring rates.

An over-estimation is observed in Bayes (Uniform) and Bayes (Jeffreys) estimates of the component density parameters with some rare exceptions in case of the second component density estimates. On the other hand, the Bayes (Uniform) and the Bayes (Jeffreys) estimates of the mixing proportion parameter are generally under-estimated with some rare exceptions. The Bayes (Gamma) estimates of the first component density and mixing proportion parameters are over-estimated while under-estimated for the second component density parameter. The extent of over-estimation is higher in case of Bayes (Uniform) as compared to Bayes (Jeffreys) but the latter has relatively smaller variance. All the Bayes estimates get more precise with the increase in sample size. The informative Bayes (Gamma) estimates have least variances than the uninformative Bayes estimates. The efficiency of Bayes (Gamma) can further be improved with an improvement in the prior information.

In the real life example, the proposed estimates evaluated with the help of Pareto mixture appears to be slightly more precise than those based on Exponential mixture as presented in Sinha (1998).

CHAPTER 6

PROPERTIES AND COMPARISON OF THE BAYES ESTIMAES OF THE BURR MIXTURE PARAMETERS

6.1 Introduction

Burr (1942) has suggested a number of cumulative distribution functions yielding a wide range of values of skewness and kurtosis and hence can be used to fit almost any given set of unimodal data. Johnson et al. (1994) presented the twelve forms for the cumulative distribution function of Burr distribution.

I. $F(y)=y,\ 0<y<1$
II. $F(y)=(e^{-y}+1)^{k}$
III. $F(y)=(y^{-c}+1)^{-k},\ 0<y$
IV. $F(y)=[((c-y)/y)^{1/c}+1],\ 0<y<c$
V. $F(y)=(ce^{-\tan y}+1)^{-k},\ -\pi/2<y<\pi/2$
VI. $F(y)=(ce^{-k\sinh y}+1)^{-k}$
VII. $F(y)=2^{-k}(1+\tanh y)^{k}$
VIII. $F(y)=[(2/\pi)\tan^{-1}e^{y}]^{k}$
IX. $F(y)=1-2/\{c\{1+e^{y}\}^{k}-1\}+2\}$
X. $F(y)=(1+e^{y^{2}})^{k},\ 0\ <\ y$
XI. $F(y)=[y-(1/2\pi)\sin 2\pi y]^{k},\ 0<y<1$
XII. $F(y)=1-(1+y^{c})^{-k},\ 0<y$

A special case of Burr Type-XII distribution is discovered as a transformed version of the Pareto distribution. The proposed distribution has an advantage over the Pareto distribution in terms of its more realistic support. In life testing and reliability we confront with many applications where a population under study is supposed to comprise of a number of subpopulations mixed together in an unknown proportion. If the observations are supposed to be characterized by the proposed Burr distribution, the use of the finite mixture Burr distribution becomes inevitable provided the data is not available on the individual components rather on the mixture only. Burr (1942, 1968 and 1973), Burr and Cislak (1968), and Rodriguez (1977) devoted special attention to one of these forms, denoted by Type-XII whose distribution function $F(x)$ is given as below.

$$F(x) = 1-(1+x^{c})^{-k},\ x\ >\ 0,\ c\ >\ 0,\ k\ >\ 0$$

Both c and k are shape parameters. The probability density function is

$$f(x) = k\ c\ x^{c-1}(1+x^{c})^{-(k+1)},\ x\ \geq 0,\ k\ >\ 0,\ c\ > 0$$

The r^{th} moment about the origin of X can easily be shown to be

$$E(X^{r}) = \mu_{r}^{'}\ = k\ \Gamma(k-r/c)\ \Gamma(r/c+1)/\Gamma(k+1)\ ,\ ck\ >\ r$$

Tadikamalla (1980) presents a nice account of the Burr and related distribution. Mixtures of Burr distribution have not been paid much attention in literature so far. Economou and Caroni (2005) have considered a Burr distribution in terms of Graphical tests. Saleem and Aslam (2010) are the first to consider a two component mixture of one parameter Burr type-XII distribution. In this Chapter, a two component mixture of the proposed Burr distribution is considered to model a lifetime mixture data. The said mixture model is defined in section 2. The likelihood is discussed in section 3. In section 4, the expressions are derived for the Bayes estimators assuming uninformative priors, the uniform and the Jeffreys. Section 5 consists of the expressions of the Bayes estimators assuming informative Gamma prior. The complete sample expressions for the ML and Bayes estimates are given in Section 6. A simulation study is conducted in section 7 to highlight some interesting properties of the estimates of the proposed Burr mixture in terms of different sample sizes, component density and mixing proportion parameter values and sample sizes. A real life application of the proposed mixture is presented as well in Section 8 and the concluding remarks are given in Section 9.

6.2 A Burr Finite Mixture Model

A finite mixture density function with the k component densities of specified parametric form (but with unknown parameters, $\alpha_i,\ i=1,2,\dots,k$) and with unknown mixing weights ($\pi_i,\ i=1,2,\dots,k$) is defined as in (2.1). Consider the following Burr distribution (a special case of Burr Type-XII) to assume for k components of the mixture.

$$f_i(x)=\alpha_i\ (1+x\ )^{-(\alpha_i+1)},\ i=1,2,\dots,k\ ;\ 0<\alpha_i<\infty\ ;\ 0\le x<\infty$$

This Burr distribution is a practical transformed version of Pareto distribution having a life time specific support. It has an interesting relation with the Exponential distribution as well through Burr-Exponential link. The motivation for this Burr distribution is derived from the fact that it is a transformed version of the one parameter Pareto distribution given by

$$f_i(x)=\alpha_i\ x^{-(\alpha_i+1)},\ i=1,2,\dots,k\ ;\ 0<\alpha_i<\infty\ ;\ 1\le x<\infty$$

But, unlike the above Pareto distribution, the considered Burr distribution has support on the positive x-axis and hence seems more suitable to fit lifetime data. The mixture model defined in (2.1) takes the following form

$$f(x)=\sum_{i=1}^{k}\pi_i\ \alpha_i\ (1+x_{ij})^{-(\alpha_i+1)}\ ;\ \ 0\ <\ \pi_i<\ 1,\ i=1,2,\dots,k\,.$$

The corresponding mixture Survival function is given by

$$S(T)=\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i}$$

where T is the fixed test termination point used in the ordinary type-I, right censoring.

6.3 The Maximum Likelihood Estimates for Censored Data

The sampling scheme of Section 2.3 is considered with a Type-IV sample of size n units from the Type-I mixture model described above under ordinary type-I, right censoring. The likelihood function $L(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})$ for the censored data is considered as given in equation (2.2).

Where $\mathbf{x}=(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_k)$ is data where $\mathbf{x}_i=(x_{i1},\ x_{i2},\ldots,x_{ir_i}),\ i=1,2,\ldots,k$

$$L(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})\propto\{\sum_{i=1}^{k}\pi_i\ (1+T)^{-\alpha_i}\}^{n-r}\prod_{i=1}^{k}\{\prod_{j=1}^{r_i}\pi_i\ \alpha_i\ (1+x_{ij})^{-(\alpha_i+1)}\} \tag{6.1}$$

$$\begin{aligned} l=\ln L(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})\propto\ &\sum_{i=1}^{k}r_i\ \ln\ \pi_i\ +\ \sum_{i=1}^{k}r_i\ln\ \alpha_i\ -\ \sum_{i=1}^{k}\Big[(\alpha_i+1)\ \{\ \sum\ln\ (1+x_{ij})\}\Big] \\ &+(n-r)\ \ln\ \{\sum_{i=1}^{k}\pi_i\ (1+T)^{-\alpha_i}\} \end{aligned} \tag{6.2}$$

The likelihood function in (6.2) can take the following form as well

$$L(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})\propto\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^{k}}\binom{n-r}{k_1,k_2,\ldots,k_k}(\prod_{\mathrm{i}=1}^{\mathrm{k}}\pi_i^{r_i+k_i})(\prod_{i=1}^{k}\alpha_i^{r_i})\exp\left\{-\sum_{i=1}^{k}\alpha_i\left\{\sum\ln\ (1+x_{ij})+k_i\ln\ (1+T)\right\}\right\} \tag{6.3}$$

H_{n-r}^{k} denotes the number of all k-ary sequences $(k_1,k_2,\ldots,k_k)$ of non-negative integers as defined in Section 2.3. Maximum Likelihood Estimates of $\alpha_i,\ i=1,2,\ldots,k$ and of $\pi_i,\ i=1,2,\ldots,k$ are obtained by solving the system of nonlinear equations (6.4)-(6.5) obtained by setting first order derivatives of the log likelihood (6.3) to zero.

$$\frac{\partial l}{\partial\alpha_i}=\frac{r_i}{\alpha_i}-\sum_{j=1}^{r_i}\ln(1+x_{ij})-\frac{(n-r)\pi_i(1+T)^{-\alpha_i}\ \ln(1+T)}{\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i}}=0,\ i=1,2,\ldots,k \tag{6.4}$$

$$\frac{\partial l}{\partial\pi_i}=\frac{r_i}{\pi_i}-\frac{r_k}{\pi_k}+\frac{(n-r)((1+T)^{-\alpha_i}-(1+T)^{-\alpha_k})}{\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i}}=0,\ i=1,2,\ldots,k-1. \tag{6.5}$$

Variances of the maximum likelihood estimates are on the main diagonal of the inverted Information matrix. The elements of the Information matrix as given in equation (2.7) are as below.

$$E(\frac{-\partial^2 l}{\partial\alpha_i^2})=\frac{(\prod_{i=1}^{k}\pi_i)(n-r)(1+T)^{-(\sum_{i=1}^{k}\alpha_i)}}{(\ln(1+T))^{-2}\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i}}-\frac{r_i}{\alpha_i^2},\ i=1,2,\ldots,k. \tag{6.6}$$

$$E(\frac{-\partial^2 l}{\partial\pi_i^{\ 2}})=-\frac{r_i}{\pi_i^{\ 2}}-\frac{(n-r)((1+T)^{-\alpha_i}-(1+T)^{-\alpha_k})^2}{(\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i})^2},\ i=1,2,\ldots,k-1 \tag{6.7}$$

$$E(\frac{-\partial^2 l}{\partial\alpha_i\partial\alpha_j})=E(\frac{-\partial^2 l}{\partial\alpha_j\partial\alpha_i})=\frac{-(\prod_{i=1}^{k}\pi_i)(n-r)(\ln(1+T))^2}{(1+T)^{(\sum_{i=1}^{k}\alpha_i)}(\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i})^2},\ j>i=1,2,\ldots,k \tag{6.8}$$

$$E(\frac{-\partial^2 l}{\partial\alpha_i\partial\pi_j})=E(\frac{-\partial^2 l}{\partial\pi_j\partial\alpha_i})=\frac{-(1+T)^{-(\sum_{i=1}^{k}\alpha_i)}(\ln(1+T))}{(n-r)^{-1}(\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i})^2},\ i=1,2,\ldots,k;\ j=1,2,\ldots,k-1$$

(6.9)

$$E(\frac{-\partial^2 l}{\partial\pi_j\partial\pi_i})=-\frac{r_k}{\pi_k^2}+\frac{\{(1+T)^{-\alpha_i}-(1+T)^{-\alpha_k}\}\{(1+T)^{-\alpha_j}-(1+T)^{-\alpha_k}\}}{(n-r)^{-1}\left\{\sum_{i=1}^{k}\pi_i(1+T)^{-\alpha_i}\right\}^2}=0, \tag{6.10}$$

$$i=1,2,\ldots,k-2,\ j=i+1=2,\ldots,k-1$$

The equations (6.5)-(6.6) can be solved using an iterative numerical to reach ML estimates and the estimated variances can be evaluated by using equations (6.7)-(6.11) and by inverting the information matrix as mentioned above.

6.4 The Posterior Distributions assuming the Uninformative Priors

Uninformative priors works in the state of ignorance about the parameter of interest.

6.4.1 The Posterior Distributions assuming the Uniform Prior

Let us assume a state of ignorance, that is, are uniformly distributed over $(0,\infty)$. Hence $f_i(\alpha_i)=k_i,\ 0<\alpha_i<\infty$, $i=1,2,\ldots,k$ and $\boldsymbol{\pi}=(\pi_1,\pi_2,\ldots,\pi_k)$ □ $Dirichlet(1,1,\ldots,1)$. Assuming independence we have an improper joint prior that is proportional to a constant which is incorporated with the likelihood (6.4) to yield a proper joint posterior distribution as follows.

$$g_U(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})=\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\Big(\prod_{i=1}^k\pi_i^{r_i+k_i}\Big)\Big(\prod_{i=1}^k\alpha_i^{r_i}\Big)\exp\Big\{-\sum_{i=1}^k(\alpha_iA_{ik})\Big\},$$
$$0<\alpha_i<\infty,\ 0<\pi_i<1,\ i=1,2,\ldots,k.$$

where $\Omega_U=\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\left\{\prod_{i=1}^k\frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\},\ \sum_{i=1}^k r_i=r$,

$A_{ik}=\sum_{j=1}^{r_i}\ln\ (1+x_{ij})+k_i\ln\ (1+T)$, $i=1,2,\ldots,k$; and $\sum_{i=1}^k k_i=n-r$.

The following are the respective marginal posterior distributions of $\alpha_i,\ i=1,2,\ldots,k$.

$$g_{Ui}(\alpha_i|\mathbf{x})=\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\left\{\prod_{j\neq i}^k\frac{\Gamma(r_2+1)}{A_{2k}^{r_2+1}}\right\}\alpha_i^{r_i}\exp(-\alpha_iA_{ik}),$$
$$0<\alpha_i<\infty,\ i=1,2,\ldots,k.$$

Marginal distributions of $\pi_i,\ i=1,2,\ldots,k$ can be obtained on the same lines as well.

6.4.1.1 Bayes Estimators assuming the Uniform Prior

Under the squared error loss function, the posterior expectations of $\alpha_i,\ i=1,2,\ldots,k$ and $\pi_i,\ i=1,2,\ldots,k-1$ with respect to the above marginal posterior distributions are the Bayes estimators are as under.

$$\hat{\alpha}_i = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \frac{\Gamma(r_i+2)}{A_{ik}^{r_i+2}} \left\{ \prod_{i\neq j} \frac{\Gamma(r_j+1)}{A_{jk}^{r_j+1}} \right\}, \quad i=1,2,\ldots,k.$$

$$\hat{\pi}_i = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \left[\binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+2,\ldots,r_k+k_k+1) \left\{ \prod_{j=1}^{k} \frac{\Gamma(r_j+1)}{A_{jk}^{r_j+1}} \right\} \right], \quad i=1,2,\ldots,k.$$

The expressions for the variances of the Bayes estimators are

$$V(\hat{\alpha}_i) = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \frac{\Gamma(r_i+3)}{A_{ik}^{r_i+3}} \left\{ \prod_{i\neq j} \frac{\Gamma(r_j+1)}{A_{jk}^{r_j+1}} \right\}$$
$$- \left[\Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1) \frac{\Gamma(r_i+2)}{A_{ik}^{r_i+2}} \left\{ \prod_{i\neq j} \frac{\Gamma(r_j+1)}{A_{jk}^{r_j+1}} \right\} \right]^2, \quad i=1,2,\ldots,k.$$

$$V(\hat{\pi}_i) = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+3,\ldots,r_k+k_k+1) \left\{ \prod_{j=1}^{k} \frac{\Gamma(r_j+1)}{A_{jk}^{r_j+1}} \right\} -$$
$$- \left[\Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+2,\ldots,r_k+k_k+1) \left\{ \prod_{j=1}^{k} \frac{\Gamma(r_j+1)}{A_{jk}^{r_j+1}} \right\} \right]^2, \quad i=1,2,\ldots,k-1.$$

Here $A_{ik} = \sum_{j=1}^{r_i} \ln\ (1+x_{ij}) + k_i \ln\ (1+T),\ i=1,2,\ldots,k$; $\sum_{i=1}^{k} r_i = r$ and $\sum_{i=1}^{k} k_i = n-r$.

6.4.2 The Posterior Distributions assuming the Jeffreys Prior

For the Burr models given in Section 6.2, the Jeffreys priors, as defined in Section 2.5.2, are $g_i(\alpha_i) \propto 1/\alpha_i,\ i=1,2,\ldots,k$, $0<\alpha_i<\infty$ and $\boldsymbol{\pi} = (\pi_1,\pi_2,\ldots,\pi_k) \sim Dirichlet(1,1,\ldots,1)$. Assuming independence, the joint prior is incorporated with the likelihood (6.4) to have the joint posterior distribution as follows.

$$g_J(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x}) = \Omega_J^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \left\{\prod_{i=1}^{k} \pi_i^{r_i+k_i}\right\}\left\{\prod_{i=1}^{k} \alpha_1^{r_1-1}\right\} \exp\left\{-\sum_{i=1}^{k}(\alpha_i A_{ik})\right\},$$
$$0<\alpha_i<\infty,\ 0<\pi_i<1,\ i=1,2,\ldots,k.$$

where $\Omega_J = \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\left\{\prod_{i=1}^{k}\frac{\Gamma(r_i)}{A_{ik}^{r_i}}\right\},\ \sum_{i=1}^{k} r_i = r$,

$A_{ik} = \sum_{j=1}^{r_i} \ln\ (1+x_{ij}) + k_i \ln\ (1+T),\ i=1,2,\ldots,k$; and $\sum_{i=1}^{k} k_i = n-r$.

The following are the respective marginal posterior distributions of $\alpha_i,\ i=1,2,\ldots,k$.

$$g_{J_i}(\alpha_i|\mathbf{x}) = \Omega_J^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\left\{\prod_{j\neq i}^{k}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right\}\alpha_i^{r_i-1}\ \exp(-\alpha_i A_{ik}),$$
$$0<\alpha_i<\infty,\ i=1,2,\ldots,k.$$

Marginal distributions of $\pi_i,\ i=1,2,\ldots,k$ can be obtained similarly.

6.4.2.1 Bayes Estimators assuming the Jeffreys Prior

The expressions for the Bayes estimators of $\alpha_i,\ \pi_i,\ i=1,2,\ldots,k$, assuming the Jeffreys prior are obtained by taking expectations of $\alpha_i,\ \pi_i,\ i=1,2,\ldots,k$ with respect to their respective marginal posterior distributions.

$$\hat{\alpha}_i = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\left\{\frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\}\left\{\prod_{i\neq j}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right\},$$
$$i=1,2,\ldots,k.$$

$$\hat{\pi}_i = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+2,\ldots,r_k+k_k+1)\left\{\prod_{j=1}^{k}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right\},$$
$$i=1,2,\ldots,k.$$

The expressions for the variances of the Bayes estimators are

$$V(\hat{\alpha}_i)=\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ \ldots,r_k+k_k+1)\left\{\frac{\Gamma(r_i+2)}{A_{ik}^{r_i+2}}\right\}\left\{\prod_{i\neq j}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right\}$$
$$-\left[\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ \ldots,r_k+k_k+1)\left\{\frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\}\left\{\prod_{i\neq j}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right\}\right]^2,$$
$$i=1,2,\ldots,k.$$

$$V(\hat{\pi}_i)=\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+3,\ldots,r_k+k_k+1)\left\{\prod_{j=1}^{k}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right\}-$$
$$-\left[\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+2,\ldots,r_k+k_k+1)\prod_{j=1}^{k}\frac{\Gamma(r_j)}{A_{jk}^{r_j}}\right]^2,$$
$$i=1,2,\ldots,k-1.$$

Here $A_{ik}=\sum_{j=1}^{r_i}\ln\ (1+x_{ij})+k_i\ln\ (1+T),\ i=1,2,\ldots,k$ **;** $\sum_{i=1}^{k}r_i=r$ and $\sum_{i=1}^{k}k_i=n-r$.

6.5 The Posterior Distributions assuming the Conjugate Prior

The suitable informative conjugate prior to be used in this case is the Gamma prior.

6.5.1 The Posterior Distributions assuming the Gamma Prior

Let $\alpha_i\sim Gamma(m_i,s_i),\ i=1,2,\ldots,k$ and $\boldsymbol{\pi}=(\pi_1,\pi_2,\ldots,\pi_k)\sim Dirichlet(1,1,\ldots,1)$. That is, $g(\alpha_i)\propto\alpha_i^{m_i-1}\exp(-s_i\alpha_i),\ i=1,2,\ldots,k.$ Assuming independence, the joint prior is incorporated with the Likelihood to give the following joint posterior distribution of $\alpha_i,\ \pi_i\ (i=1,2,\ldots k)$ as follows.

$$g_G(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})=\Omega_G^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\left\{\prod_{i=1}^{k}\pi_i^{r_i+k_i}\right\}\left\{\prod_{i=1}^{k}\alpha_1^{r_1+m_1-1}\right\}\exp\left\{-\sum_{i=1}^{k}(\alpha_i(A_{ik}+s_i))\right\},$$
$$0<\alpha_i<\infty,\ 0<\pi_i<1,\ i=1,2,\ldots,k.$$

where $\Omega_G = \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\prod_{i=1}^{k}\frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}}$. The following marginal posterior densities are obtained by integrating out the nuisance parameters.

$$g_{iG}(\alpha_i|\mathbf{x}) = \Omega_G^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ r_2+k_2+1,\ldots,r_k+k_k+1)$$
$$\times\left\{\prod_{i\neq 1}\frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}}\right\}\ \alpha_i^{r_i+m_i-1}\exp\{-\alpha_i(A_{ik}+s_i)\},\ 0<\alpha_i<\infty,\ i=1,2,\ldots,k$$

Marginal distributions of $\pi_i,\ i=1,2,\ldots,k$ can be obtained in the same fashion as well.

6.5.1.1 Bayes Estimators assuming the Gamma Prior

To find the expressions for the Bayes estimators of $\alpha_i,\ \pi_i,\ i=1,2,\ldots,k$ under the squared error loss function are given by the posterior means of $\alpha_i,\ \pi_i,\ i=1,2,\ldots,k$ with respect to the respective marginal posterior distributions. The posterior means or the Bayes estimators of $\alpha_i,\ \pi_i,\ i=1,2,\ldots,k$ under the squared error loss function are given bellow.

$$\hat{\alpha}_i = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\frac{\Gamma(r_i+m_i+1)}{(s_i+A_{ik})^{r_i+m_i+1}}\prod_{i\neq j}\frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}},$$
$$i=1,2,\ldots,k$$

$$\hat{\pi}_i = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \left[\binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+2,\ldots,r_k+k_k+1)\prod_{j=1}^{k}\frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}}\right],$$
$$i=1,2,\ldots,k-1.$$

The expressions for the variances of the Bayes estimators are

$$V(\hat{\alpha}_i) = \Omega_U^{-1} \sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k} \binom{n-r}{k_1,k_2,\ldots,k_k} \mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\frac{\Gamma(r_i+m_i+2)}{(s_i+A_{ik})^{r_i+m_i+2}}\prod_{i\neq j}\frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}}$$

$$-\left[\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,r_k+k_k+1)\frac{\Gamma(r_i+m_i+1)}{(s_i+A_{ik})^{r_i+m_i+1}}\prod_{i\neq j}\frac{\Gamma(r_i+m_i)}{(s_i+A_{ik})^{r_i+m_i}}\right]^2,$$

$$i=1,2,\ldots,k.$$

$$V(\hat{\pi}_i)=\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+3,\ldots,r_k+k_k+1)\left\{\prod_{j=1}^{k}\frac{\Gamma(r_j+m_j)}{(s_j+A_{jk})^{r_j+m_j}}\right\}$$

$$-\Omega_U^{-1}\sum_{k_1,k_2,\ldots,k_k}^{H_{n-r}^k}\left[\binom{n-r}{k_1,k_2,\ldots,k_k}\mathrm{B}(r_1+k_1+1,\ldots,\ r_i+k_i+2,\ldots,r_k+k_k+1)\left\{\prod_{j=1}^{k}\frac{\Gamma(r_j+m_j)}{(s_j+A_{jk})^{r_j+m_j}}\right\}\right]^2,$$

$$i=1,2,\ldots,k-1.$$

Here $A_{ik}=\sum_{j=1}^{r_i}\ln\ (1+x_{ij})+k_i\ln\ (1+T),\ i=1,2,\ldots,k$ **;** $\sum_{i=1}^{k}r_i=r$ and $\sum_{i=1}^{k}k_i=n-r$.

6.6 The Complete Sample Expressions

Under the conditions given in Section 2.6, the expressions for the Bayes estimators and their variances are simplified as given in Tables 6.1-6.2. The comments regarding amount of information, computational ease and simplification quoted in Section 2.6 also applies here.

Table 6.1 The complete sample expressions for the Bayes and Maximum likelihood estimators as $T\to\infty$

Parameters	Bayes Estimators (Uniform)	Bayes Estimators (Jeffreys)	ML Estimators
α_1	$\dfrac{n_1+1}{\sum\ln(1+x_{1j})}$	$\dfrac{n_1}{\sum\ln(1+x_{1j})}$	$\dfrac{n_1}{\sum\ln(1+x_{1j})}$
α_2	$\dfrac{n_2+1}{\sum\ln(1+x_{2j})}$	$\dfrac{n_2}{\sum\ln(1+x_{2j})}$	$\dfrac{n_2}{\sum\ln(1+x_{2j})}$
π_1	$\dfrac{n_1+1}{n+2}$	$\dfrac{n_1+1}{n+2}$	$\dfrac{n_1}{n}$

Table 6.2 The complete sample expressions for the variances of the Bayes and Maximum likelihood estimators as $T \to \infty$

Parameters	Variances of Bayes Estimators (Uniform prior)	Variances of Bayes Estimators (Jeffreys prior)	Variances of ML Estimators
α_1	$\frac{n_1+1}{(\sum \ln(1+x_{1j}))^2}$	$\frac{n_1}{(\sum \ln(1+x_{1j}))^2}$	$\frac{n_1}{(\sum \ln(1+x_{1j}))^2}$
α_2	$\frac{n_2+1}{(\sum \ln(1+x_{2j}))^2}$	$\frac{n_2}{(\sum \ln(1+x_{2j}))^2}$	$\frac{n_2}{(\sum \ln(1+x_{2j}))^2}$
π_1	$\frac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\frac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\frac{n_1 n_2}{n^3}$

6.7 A Simulation Study

A simulations study is conducted to investigate the performance of the Bayes estimators in terms of sample size, censoring rate and various parameter points. Samples of sizes $n = 50,\ 100,\ 150,\ 250$ from the two component mixture of Burr distribution with parameters $\alpha_1,\ \alpha_2$ and p such that $(\alpha_1, \alpha_2) \in \{(0.5,1.5),(1,4),(1.5,4.5),(2,6),(3,9)\}$ and $p \in \{0.4, 0.6\}$. Probabilistic mixing was used to generate the mixture data. For each observation a random number u was generated from the Uniform distribution on [0, 1]. If $u < p$, the observation was taken randomly from F_1 (the Burr distribution with parameter α_1) and if $u > p$, the observation was taken randomly from F_2 (the Burr distribution with parameter α_2). The choice of the censoring time, in each case, was made in such a way that the censoring rate in the resulting sample to be approximately 10% and 20%. The remaining details of the simulation study are the same as mentioned in Section 2.7. The findings of the simulation study are 6.3-6.8. Some properties of Bayes estimates are depicted in Tables 6.3-6.7 in terms of sample sizes, censoring rates and parameter points while Table 6.8 displays a comparison among the three Bayes estimates.

6.8 A Real Life Example

Davis (1952) mixture data, $\mathbf{t} = (t_{11}, t_{12}, \ldots, t_{1r_1}, t_{21}, t_{22}, \ldots, t_{2r_2})$, consist of hours to failure for electronic valves, an indicator valve and for a transmitter valve, both used in aircraft radar sets. The category of the failure is not known until the failure occurs. Inspection of failed units allows the engineers to allocate the failed units to two different subpopulations. The total number of tests carried out was 1003. The data set can be seen in Everitt and Hand (1981) on page 76.

Davis (1952) fitted exponential distributions to this data. The transformation $x = \exp(-t) - 1$ of an exponential distribution yields the said Burr distribution. This transformation allowed us to use the Davis mixture data set for our analysis with the obvious transformation of the data.

It is interesting to note that despite the transformation $x = \exp(-t) - 1$ almost no major computations are required to have the data summary required to evaluate the estimates.

For instance, $\sum_{j=1}^{r_1} \ln(1 + x_{1j}) = \sum_{j=1}^{r_1} t_{1j} = 151130$ and $\sum_{j=1}^{r_2} \ln(1 + x_{2j}) = \sum_{j=1}^{r_2} t_{2j} = 22550$.

Other sample characteristics required are also made available easily.

$n = 1003,\ r_1 = 891,\ r_2 = 92,\ r = r_1 + r_2 = 983,\ n - r = 20$

Burr mixture parameters $(\alpha_1,\ \alpha_2,\ p)$ are evaluated using estimators derived in Sections 4.2. The Bayes (Jeffreys) estimates $(\hat{\alpha}_{11},\ \hat{\alpha}_{22})$ of the Davis mixture lifetime parameters, after an obvious re-parameterization as evident from the functional form of the component densities of the mixture given in section 2, are found to be

$(\hat{\alpha}_{11},\ \hat{\alpha}_{22}) = (1/\hat{\alpha}_1,\ 1/\hat{\alpha}_2) = (179.553, 320.513)$ where $\hat{\alpha}_1 = 0.00557,\ \hat{\alpha}_2 = 0.00312$ (correct to five decimal places) are the Bayes (Jeffreys) estimates of Burr mixture parameters. with $SE(\hat{\alpha}_1) = 0.000204599,\ SE(\hat{\alpha}_2) = 0.000414361$ respectively. The standard errors of the lifetime estimates of the Davis mixture are computed as $SE(\hat{\alpha}_{11}) = 6.60642,\ SE(\hat{\alpha}_{22}) = 34.6218$. The estimate of the proportion parameter of the Davis mixture is $\hat{p} = 0.901774$ with $SE(\hat{p}) = 0.00968426$. This is encouraging to note that the estimates are much greater than the two respective subgroup sample means i.e., $\overline{t}_1 = 169.618 << 179.533,\ \overline{t}_2 = 245.109 << 320.513$ which happens in the right censoring situations. Also the proposed estimates presented here are superior to those presented in Everitt and Hand (1981) in terms of Bayesian analysis and information on standard error of the estimates.

6.9 Conclusion

The simulation study highlights some interesting properties of the Bayes estimates. The estimates of the component density parameters are generally over-estimated with some rare exceptions in case of the second component. Also the extent of over-estimation is higher in case of the estimates of the first component density. The estimates of the mixing proportion parameter are under-estimated.

The variances of the estimates of the component density parameters seem to be quite large (small) for the relatively larger (smaller) values of the parameters. The variances of the estimates of all the mixture parameters are reduced as the sample size increases. Another remark concerning the variances of the estimates of the component density parameter is that

Table 6.3 Bayes estimates (Jeffreys) of Burr mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 0.5,\ \alpha_2 = 1.5,\ \pi_1 = 0.4,\ 0.6$ and censoring rates, $C = 10\%, 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	10% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.50,1.50,0.40)	50	0.563245(0.212829)	1.51724(0.350963)	0.394582(0.0705581)
	100	0.520936(0.120153)	1.51118(0.24209)	0.396616(0.0495987)
	150	0.512384(0.0894173)	1.4997(0.196023)	0.398852(0.0422593)
	250	0.507501(0.0652004)	1.49638(0.144842)	0.39834(0.0307418)
(0.50,1.50,0.60)	50	0.519267(0.12693)	1.5424 (0.425452)	0.586989(0.070666)
	100	0.504399(0.082749)	1.50667(0.271674)	0.596691(0.0507526)
	150	0.502445(0.0608871)	1.5077(0.214871)	0.598474(0.0377672)
	250	0.50212(0.0464165)	1.50679(0.168577)	0.597064(0.0296618)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.50,1.50,0.40)	50	0.6404 (0.34735)	1.51658(0.411874)	0.385982(0.0745899)
	100	0.586221(0.195429)	1.47665(0.296009)	0.386156(0.0580364)
	150	0.549889(0.145323)	1.47479(0.239981)	0.391429(0.0497621)
	250	0.532881(0.100711)	1.47944(0.190058)	0.392174(0.0385402)
(0.50,1.50,0.60)	50	0.571622(0.189876)	1.47721(0.467739)	0.573809(0.0750944)
	100	0.526989(0.112306)	1.50218(0.344042)	0.587839(0.0551717)
	150	0.51612(0.0783961)	1.48119(0.260268)	0.591603(0.0439547)
	250	0.508395(0.0575559)	1.49048(0.2059)	0.594128(0.0328895)

Table 6.4 Bayes estimates (Jeffreys) of Burr mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 1,\ \alpha_2 = 4,\ \pi_1 = 0.4,\ 0.6$ and censoring rates, $C = 10\%, 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**10% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(1,4,0.40)	50	1.05859 (0.432553)	4.0607 (0.880235)	0.402192(0.0704995)
	100	1.01501(0.214597)	4.01675(0.584037)	0.399608(0.0488987)
	150	1.01573 (0.164883)	4.01755(0.481591)	0.397903(0.0405221)
	250	1.0097 (0.125912)	4.00205(0.362668)	0.397923 (0.0301273)
(1,4,0.60)	50	1.01158(0.220562)	4.11042 (1.05833)	0.59385(0.0662204)
	100	1.00607(0.154689)	4.07697 (0.720826)	0.595179(0.049105)
	150	0.99606 (0.117763)	4.08422 (0.558807)	0.598152(0.0397743)
	250	1.00347(0.091507)	4.03637(0.419032)	0.600054(0.0315375)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(1,4,0.40)	50	1.26429 (0.665624)	4.01434(1.06246)	0.380388(0.0783151)
	100	1.11181(0.381456)	3.9762 (0.708017)	0.389651(0.0583055)
	150	1.07679 (0.264695)	3.97594 (0.595045)	0.392915 (0.0461595)
	250	1.03148(0.16494)	3.96117 (0.458245)	0.395455(0.0349119)
(1,4,0.60)	50	1.09388(0.375397)	4.09782 (1.26045)	0.583676(0.0720112)
	100	1.02327(0.19868)	4.01138(0.759477)	0.592166 (0.0502647)
	150	1.01544 (0.140035)	4.01229(0.642166)	0.596443(0.0408702)
	250	1.00772 (0.105068)	4.00914 (0.480287)	0.597176(0.0310371)

Table 6.5 Bayes estimates (Jeffreys) of Burr mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 1.5,\ \alpha_2 = 4.5,\ \pi_1 = 0.4,\ 0.6$ and censoring rates, $C = 10\%, 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**10% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(1.5, 4.5, 0.40)	50	1.65432 (0.663069)	4.53824 (1.02221)	0.395974 (0.0678431)
	100	1.56651 (0.330227)	4.5318 (0.70716)	0.399456(0.0512226)
	150	1.53337 (0.25748)	4.55327(0.57368)	0.397583(0.0417845)
	250	1.5147 (0.19115)	4.49578 (0.444445)	0.399321 (0.0315452)
(1.5, 4.5, 0.60)	50	1.55619 (0.396521)	4.5808 (1.32379)	0.593287(0.0684327)
	100	1.51279(0.228073)	4.56592 (0.840181)	0.597119(0.0492198)
	150	1.5024 (0.182506)	4.55109 (0.662697)	0.599052(0.0397222)
	250	1.50809 (0.139536)	4.51429(0.505096)	0.597286 (0.0309605)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(1.5, 4.5, 0.40)	50	2.09186 (4.34854)	4.53329(1.2932)	0.382362(0.0784554)
	100	1.76558(0.600229)	4.45613 (0.901008)	0.384719(0.0611575)
	150	1.64011 (0.396642)	4.46345 (0.734239)	0.389264 (0.0496912)
	250	1.58799(0.277652)	4.41593 (0.545297)	0.391451(0.0396456)
(1.5, 4.5, 0.60)	50	1.69817(0.552653)	4.42644 (1.45622)	0.574746(0.0764883)
	100	1.57887 (0.317611)	4.41687(1.00175)	0.585484 (0.0535618)
	150	1.54821 (0.242906)	4.44615(0.829013)	0.593493(0.0435132)
	250	1.528 (0.162178)	4.47749 (0.598865)	0.596234 (0.0337839)

Table 6.6 Bayes estimates (Jeffreys) of Burr mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 2,\ \alpha_2 = 6,\ \pi_1 = 0.4,\ 0.6$ and censoring rates, $C = 10\%, 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	10% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(2, 6, 0.40)	50	2.24802 (0.892374)	6.08347 (1.46152)	0.39222 (0.0735677)
	100	2.1047 (0.48607)	6.04929 (0.988346)	0.39982(0.0489828)
	150	2.05072 (0.341908)	5.98034(0.755177)	0.398359 (0.0410617)
	250	2.03708 (0.262091)	6.00147 (0.578209)	0.398681 (0.0316621)
(2, 6, 0.60)	50	2.06376 (0.508258)	6.11963 (1.65882)	0.59171(0.0699953)
	100	2.01328 (0.299106)	6.04311 (1.10052)	0.594093(0.0513299)
	150	2.01817 (0.252408)	6.00469 (0.866215)	0.597663(0.0394122)
	250	2.01162 (0.182618)	5.98786(0.636022)	0.59811 (0.0304095)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	20% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(2, 6, 0.40)	50	2.55501 (1.21309)	6.11759(1.72849)	0.383844(0.0760835)
	100	2.31269 (0.733036)	5.96783 (1.23128)	0.387768(0.0602393)
	150	2.18727 (0.544893)	5.98064 (0.975847)	0.393029 (0.0500159)
	250	2.11128(0.379879)	5.94117 (0.769935)	0.393224 (0.0380757)
(2, 6, 0.60)	50	2.29083(0.801201)	5.93851 (1.97904)	0.577608(0.0797272)
	100	2.12503 (0.449143)	5.8851(1.41371)	0.582634 (0.0550745)
	150	2.05912 (0.317596)	5.97591 (1.04254)	0.590399(0.0442136)
	250	2.03132 (0.225605)	5.95041 (0.85228)	0.597053 (0.033136)

Table 6.7 Bayes estimates (Jeffreys) of Burr mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 3,\ \alpha_2 = 9,\ \pi_1 = 0.4,\ 0.6$ and censoring rates, $C = 10\%, 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	10% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(3, 9, 0.40)	50	3.37042 (1.20231)	9.06943 (2.15165)	0.394986 (0.0691276)
	100	3.13938 (0.703413)	9.07597 (1.41536)	0.397902(0.0498643)
	150	3.09128 (0.559522)	8.917 (1.16236)	0.395241 (0.0414236)
	250	3.04876 (0.382345)	9.0191 (0.850516)	0.398479 (0.0318075)
(3, 9, 0.60)	50	3.1006 (0.774)	9.22706 (2.52174)	0.588959 (0.0720894)
	100	3.04615 (0.481875)	9.07567 (1.71137)	0.595205(0.0488415)
	150	3.00422 (0.362654)	9.04443 (1.31639)	0.595519(0.0397315)
	250	3.0096 (0.267927)	9.04052(1.01117)	0.598233 (0.0305004)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	20% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(3, 9, 0.40)	50	3.89071 (1.82021)	9.05081(2.49022)	0.378545 (0.0764212)
	100	3.49255 (1.17498)	8.91389 (1.8142)	0.386272(0.0618352)
	150	3.26047 (0.780437)	8.88824 (1.42102)	0.391153 (0.0488964)
	250	3.17515(0.564824)	8.92507 (1.08309)	0.39318 (0.0391872)
(3, 9, 0.60)	50	3.4299 (1.17686)	8.9686 (3.0807)	0.570009(0.075344)
	100	3.14887 (0.660037)	8.86383 (2.00151)	0.588208 (0.0554302)
	150	3.08956 (0.468894)	8.94374 (1.59991)	0.59083(0.0424624)
	250	3.04201 (0.331964)	9.01183 (1.17878)	0.597894 (0.0318113)

Table 6.8 A comparison of the Bayes (Uniform) and Bayes (Jeffreys) estimates of Burr mixture parameters and their standard errors (in parenthesis) with $\alpha_1 = 0.5$, $\alpha_2 = 1.5$, $\pi_1 = 0.25$ and censoring rate, $C = 10\%$.

Prior	π_1	n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
UNIFORM	0.40	50	0.611974(0.250889)	1.56884(0.36821)	0.393967(0.072079)
		100	0.53933(0.126465)	1.53816(0.247555)	0.396633(0.049935)
		150	0.5238(0.092298)	1.518(0.198886)	0.398937(0.042411)
		250	0.514029(0.066073)	1.50749(0.14596)	0.398427(0.030777)
	0.60	50	0.538668(0.13124)	1.63726(0.447465)	0.588103(0.070883)
		100	0.513777(0.083804)	1.55149(0.277023)	0.597118(0.050738)
		150	0.508742(0.061413)	1.537(0.217735)	0.598709(0.037749)
		250	0.505952(0.046685)	1.52397(0.169821)	0.59718(0.029653)
JEFFREYS	0.40	50	0.563245(0.212829)	1.51724(0.350963)	0.394582(0.070558)
		100	0.520936(0.120153)	1.51118(0.24209)	0.396616(0.049599)
		150	0.512384(0.089417)	1.4997(0.196023)	0.398852(0.042259)
		250	0.507501(0.0652)	1.49638(0.144842)	0.39834(0.030742)
	0.60	50	0.519267(0.12693)	1.5424(0.425452)	0.586989(0.070666)
		100	0.504399(0.082749)	1.50667(0.271674)	0.596691(0.050753)
		150	0.502445(0.060887)	1.5077(0.214871)	0.598474(0.037767)
		250	0.50212(0.046417)	1.50679(0.168578)	0.597064(0.029662)
GAMMA	0.40	50	0.549771(0.191977)	1.46817(0.327981)	0.549771(0.191977)
		100	0.51722(0.117482)	1.48681(0.235083)	0.51722(0.117482)
		150	0.510209(0.088367)	1.48359(0.192396)	0.510209(0.088367)
		250	0.506287(0.064957)	1.48678(0.143341)	0.506287(0.064957)
	0.60	50	0.515472(0.124858)	1.4655(0.387584)	0.515472(0.124858)
		100	0.502549(0.082407)	1.47172(0.261841)	0.502549(0.082407)
		150	0.501148(0.060726)	1.48489(0.209638)	0.501148(0.060726)
		250	0.549771(0.191977)	1.46817(0.327981)	0.549771(0.191977)

increasing (decreasing) the proportion of a component in the mixture reduces (increases) the variance of the estimate of the corresponding parameter.

The effect of censoring on the estimates of the first component density parameter is in the form of an increase in the extent of over-estimation. To be more specific, larger degree of censoring results in bigger size of over-estimation. The slight over-estimation in the estimates of the second component density parameter turns into a slight under-estimation and in the exceptional cases there is an increase in the extent of the under-estimation in response to an increase in the censoring rate. On the other hand the extent of under-estimation is further increased in case of estimates of the proportion parameter as a result of an increase in the censoring rates. It is interesting to note that the size of this over or under-estimation is directly proportional to the amount of the censoring rates and inversely proportional to the sample size. Also the extent of over-estimation is more intensive for larger parameter values. The increase in censoring rate increases the variances of estimates of all the mixture parameters.

Furthermore, increasing the sample size reduces the variance of all the estimates without any exception. The increase in proportion of a component in the mixture reduces the variance of the estimate of the corresponding parameter.

As the cut off sensor value tends to infinity, the complete sample expressions for the estimators and variances are greatly simplified. In addition, variances of the complete sample estimates are expected to be reduced further as there in no more effect of censoring.

The Bayes (Uniform), the Bayes (Jeffreys) and the Bayes (Gamma) estimates of the parameter of the first component density are over-estimated but the extent of over-estimation is higher in case of Uniform and the least in case of Gamma. On the other hand, the Bayes (Uniform) estimates of the parameter of the second component density are over-estimated as well, while the Bayes (Jeffreys) estimates are generally over- estimated but with some exceptions and the Bayes (Gamma) estimates are under-estimated. It is interesting to note that the all the three Bayes estimates of the mixing proportion are under-estimated.

The Bayes estimates of the parameters of the component densities with informative (Gamma) prior have smaller variances than their Uniform and Jeffreys counterparts while the Bayes (Jeffreys) show smaller variances than Bayes (Uniform). However, the variances of the Bayes (Gamma) estimates of the mixing proportion parameter may not be the least all the times. In other words, the Bayes estimates with informative (Gamma) prior seem to be more efficient than their uninformative counterparts with a few exceptions only in case of the mixing proportion estimates. Actually, the quality of Bayes (Gamma) depends upon the quality of prior information. The hyperparameters 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 hyperparameters are available. In the real life example, the proposed estimates are superior in terms of the Bayesian analysis, information on and size of standard error of the estimates.

CHAPTER 7

PROPERTIES AND APPLICATION OF A POWER FUNCTION MIXTURE

7.1 Introduction

The Power Function distribution is often used to study the electrical component reliability. A population of lifetimes of a certain electrical elements may be divided into a number of subpopulations depending upon the possible causes of failure. If the random observations taken from this population are supposed to be characterized by one of the two distinct members of a Power Function distribution, then the two components mixture of the Power Function distribution is recommended to model this population provided the data is not available on the individual components rather on the mixture only. Acheson and McElwee (1952) reported that an electronic tube may fail either due to gaseous defects or mechanical defects or normal deterioration of the cathode. Mixed failure populations can be found in Davis (1952), Epstein and Sobel (1953) and Mendenhall and Hader (1958). Ahsanullah and Lutful Kabir (1974) gave a brief characterization of the Power Function distribution**.** Meniconi and Barry (1996) discussed the electrical component reliability using the Power Function distribution. Ali et al. (2005) considers a characterization of Power Function distribution. Saleem and Aslam (2010) focused Bayesian analysis of the Power Function mixture. There are a few works available in literature on the Bayesian analysis of the Power Function distribution and its mixture.

In this chapter, the Power Function mixture model is defined in Section 2 and its likelihood is developed in Section 3. The system of three non-linear equations, required to be solved iteratively for the computations of maximum likelihood estimates, is derived. The components of the information matrix are constructed as well. In Sections 4-5, the elegant closed form expressions for the Bayes estimators and their variances are derived. In Section 5, the posterior predictive distribution with the informative prior is derived and the equations required to find the lower and upper limits of the predictive intervals are constructed. The complete sample expressions of these estimators and their variances are derived in Section 6. In Section 7, a comprehensive simulation scheme including a large number of parameter points is followed to highlight the properties and behavior of the estimates in terms of sample size, censoring rate, parameters size and the proportion of the components of the mixture. A simulated mixture data with censored observations is generated by probabilistic mixing for the computational purposes. A real life data is used in Section 8 for the evaluation of Bayes estimates. Some concluding remarks are given in Section 9. The Bayes estimates are evaluated under the squared error loss function.

7.2 A Power Function Mixture Model

A finite mixture density function with the two component densities of specified parametric form (but with unknown parameters, $\alpha_i,\ i=1,2,\ldots,k$) and with unknown mixing weights ($\pi_i,\ i=1,2,\ldots,k$) is defined as in (2.2). And the corresponding mixture survival function is as follows. $S(x)=1-F(x)=1-\sum_{i=1}^{k}\pi_i\ F_i(x)$. The Power Function distributions assumed for the two components of the mixture are as under.

$$f_i(x)=\alpha_i x^{\alpha_i-1},\ i=1,2,\ldots,k\ ;\ \alpha_i>0\ ;\ 0\le x\le 1$$

And the corresponding distribution functions are given as follow.

$$F_i(x)=1-x^{\alpha_i},\ i=1,2,\ldots,k\ ;\ \alpha_i>0\ ;\ 0\le x\le 1$$

7.3 The Maximum Likelihood Estimates for Censored Data

The sampling scheme of Section 2.3 is considered with a Type-IV sample of size n units from the Type-I mixture model described above under ordinary type-I, right censoring. We define, x_{ij} as the failure time of the jth unit belonging to the ith subpopulation, $j = 1, 2, 3, ..., r_i$, where $i = 1, 2, \ldots, k;\ 0 < x_{ij} \le T$. The likelihood function, $L(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x})$, under the above conditions takes the form as given in (2.2) where $\mathbf{x} = (x_{11}, x_{12}, \ldots, x_{1r_1}, x_{21}, x_{22}, \ldots, x_{2r_2})$ is data.

$$L(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x}) \propto \{\prod_{j=1}^{r_1} \pi_1\ \alpha_1\ x_{1j}^{\alpha_1 - 1}\}\{\ \prod_{j=1}^{r_2} \pi_2\ \alpha_2\ x_{2j}^{\alpha_2 - 1}\}\{(1 - (\pi_1 T^{\alpha_1} + \pi_2 T^{\alpha_2}))^{n-r}\} \qquad (7.1)$$

After some simplifications the above likelihood can be expressed as

$$L(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x}) \propto \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k} \binom{k}{m} \pi_1^{r_1 + k - m} \pi_2^{r_2 + m}$$
$$\times \alpha_1^{r_1} \exp[-\alpha_1 \{\sum_{j=1}^{r_1} \ln(\frac{1}{x_{1j}}) - (k-m)\ln T\}]\ \alpha_2^{r_2} \exp[-\alpha_2 \{\sum_{j=1}^{r_2} \ln(\frac{1}{x_{2j}}) - m \ln T\}] \qquad (7.2)$$

The ML estimates of $\alpha_1,\ \alpha_2$ and π_1 are obtained by solving the system of nonlinear equations (7.3)-(7.4), obtained by differentiating the likelihood (7.2) with respect to $\alpha_1,\ \alpha_2$ and π_1 respectively.

$$\frac{\partial l}{\partial \alpha_i} = \frac{r_i}{\alpha_i} + \sum_{j=1}^{r_i} \ln(x_{ij}) - \frac{\pi_i (n-r) T^{\alpha_i} (\ln T)}{(1 - \pi_1 T^{\alpha_1} - \pi_2 T^{\alpha_2})} = 0,\ i = 1, 2 \qquad (7.3)$$

$$\frac{\partial l}{\partial \pi_1} = \frac{r_1}{\pi_1} - \frac{r_2}{\pi_2} - \frac{(n-r)(T^{\alpha_1} - T^{\alpha_2})}{(1 - \pi_1 T^{\alpha_1} - \pi_2 T^{\alpha_2})} = 0 \qquad (7.4)$$

Variances of the ML estimates are on the main diagonal of the inverted Information matrix. The Information matrix is given by the expectation of the negative Hessian matrix of the form as given by equation (2.7), where elements of the Information matrix are:

$$E\left(\frac{-\partial^2 l}{\partial\alpha_i^2}\right)=\frac{-r_i}{\alpha_i^2}-\frac{\pi_i(n-r)(\ln T)^2 T^{\alpha_i}(1-\pi_{3-i}T^{\alpha_{3-i}})}{(1-\pi_1 T^{\alpha_1}-\pi_2 T^{\alpha_2})^2},\ i=1,2 \tag{7.5}$$

$$E\left(\frac{-\partial^2 l}{\partial\pi_1^{\,2}}\right)=-\frac{r_1}{\pi_1^{\,2}}-\frac{r_2}{\pi_2^{\,2}}-\frac{(n-r)(T^{\alpha_1}-T^{\alpha_2})^2}{(1-\pi_1 T^{\alpha_1}-\pi_2 T^{\alpha_2})^2} \tag{7.6}$$

$$E\left(\frac{-\partial^2 l}{\partial\alpha_1\partial\alpha_2}\right)=E\left(\frac{-\partial^2 l}{\partial\alpha_2\partial\alpha_1}\right)=\frac{-\pi_1\pi_2(n-r)(\ln T)^2 T^{\alpha_1+\alpha_2}}{(1-\pi_1 T^{\alpha_1}-\pi_2 T^{\alpha_2})^2} \tag{7.7}$$

$$E\left(\frac{-\partial^2 l}{\partial\alpha_i\partial\pi_1}\right)=E\left(\frac{-\partial^2 l}{\partial\pi_1\partial\alpha_i}\right)=\frac{-(n-r)(\ln T)\ T^{\alpha_i}(1-T^{\alpha_{3-i}})}{(1-\pi_1 T^{\alpha_1}-\pi_2 T^{\alpha_2})^2},\ i=1,2 \tag{7.8}$$

7.4 The Posterior Distributions assuming the Uninformative Priors

Uniform and Jeffreys are considered as uninformative priors to implement Bayesian inference.

7.4.1 The Posterior Distributions assuming the Uniform Prior

Let $\alpha_i \sim Uniform\ \forall\,\alpha_i \in (0,\infty),\ i=1,2$ and $\pi_1 \sim U(0,1)$. Assuming independence, we have a joint prior that is proportional to a constant. This joint prior is incorporated with the likelihood (7.3) to yield a joint posterior distribution of $\alpha_1,\ \alpha_2$ and π_1 as follows.

$$g_U(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})=\Omega_U^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\left\{\prod_{i=1}^{2}\pi_i^{(2m+r_2-r_1+k)i+(2r_1-r_2-3m+2k)}\right\}\left\{\prod_{i=1}^{2}\alpha_i^{r_i}\right\}$$
$$\times\exp\left[-\sum_{i=1}^{2}\alpha_i\left\{\sum_{j=1}^{r_i}\ln\left(\frac{1}{x_{ij}}\right)-\{(2m-k)i+(2k-3m)\}\ln T\right\}\right]$$

where $\Omega_U=\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1,r_2+m+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\}$. The marginal posterior distribution of each parameter is obtained by integrating out the nuisance parameters.

$$g_{iU}(\alpha_1|\mathbf{x}) = \Omega_U^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \frac{\Gamma(r_{3-i}+1)}{A_{3-i,k}^{r_{3-i}+1}}$$

$$\times \alpha_i^{r_i} \exp[-\alpha_i\{\sum_{j=1}^{r_i} \ln(\frac{1}{x_{ij}}) - (k-m)\ln T\}],\ 0<\alpha_i<\infty,\ i=1,2.$$

$$g_{3U}(\pi_1|\mathbf{x}) = \Omega_U^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\}\left\{\prod_{i=1}^{2} \pi_i^{(2m+r_2-r_1-k)i+(2r_1-r_2-3m+2k)}\right\},\ 0<\pi_1<1.$$

7.4.1.1 Bayes Estimators assuming the Uniform Prior

The expectation of each parameter with respect to its marginal posterior distribution gives the Bayes estimator of the parameter under the square error loss function. The Bayes estimators of α_1, α_2 and π_1 assuming the uniform prior, are given by

$$\hat{\alpha}_i = \Omega \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \frac{\Gamma(r_1+3-i)\Gamma(r_2+i)}{A_{1k}^{r_1+3-i}\ A_{2k}^{r_2+i}},\ i=1,2$$

$$\hat{\pi}_1 = \Omega \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+2, r_2+m+1) \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\}$$

where $\Omega^{-1} = \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i+1)}{A_{ik}^{r_i+1}}\right\}$ and

$A_{ik} = \sum_{j=1}^{r_i} \ln(\frac{1}{x_{ij}}) - \{(2m-k)i + (2k-3m)\}\ln T,\ i=1,2.$ $\mathrm{B}(x,y)$ is Beta function and $\Gamma(z)$ is Gamma function.

7.4.1.2 Variances of the Bayes Estimators assuming the Uniform Prior

The variances of the Bayes estimators of α_1, α_2 and π_1 assuming the uniform prior are:

$$V(\hat{\alpha}_i)=\Omega\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1,r_2+m+1)\frac{\Gamma(r_1+5-2i)\Gamma(r_2+2i-1)}{A_{1k}^{r_1+5-2i}\;A_{2k}^{r_2+2i-1}}-$$
$$-\left[\Omega\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1,r_2+m+1)\frac{\Gamma(r_1+3-i)\Gamma(r_2+i)}{A_{1k}^{r_1+3-i}\;A_{2k}^{r_2+i}}\right]^2,\ i=1,2$$

$$V(\hat{\pi}_1)=\Omega\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+3,r_2+m+1)\frac{\Gamma(r_1+1)\Gamma(r_2+1)}{A_{1k}^{r_1+1}\;A_{2k}^{r_2+1}}-$$
$$-\left[\Omega\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+2,r_2+m+1)\frac{\Gamma(r_1+1)\Gamma(r_2+1)}{A_{1k}^{r_1+1}\;A_{2k}^{r_2+1}}\right]^2$$

7.4.2 The Posterior Distributions assuming the Jeffreys Prior

According to the definition of Jeffreys prior given in Section 2.5.2, Let $g(\alpha_i)\propto\sqrt{|I(\alpha_i)|}$

where $I(\alpha_i)=-E\left[\frac{\partial^2\ln L(x\,|\,\alpha_i)}{\partial\alpha_i^2}\right]$, $i=1,2$ and $\pi_1\sim U(0,1)$, where $f_i(x\,|\,\alpha_i)$ are given in

Section 7.2. Assuming independence, the joint prior, $g(\alpha_1,\alpha_2,\pi_1)\propto\frac{1}{\alpha_1\alpha_2}$ is incorporated with

the likelihood (7.3) to yield a joint posterior distribution of α_1, α_2 and π_1 as follows**.**

$$g_J(\boldsymbol{\alpha},\boldsymbol{\pi}|\mathbf{x})=\Omega_U^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\left\{\prod_{i=1}^{2}\pi_i^{(2m+r_2-r_1+k)i+(2r_1-r_2-3m+2k)}\right\}\left\{\prod_{i=1}^{2}\alpha_i^{r_i-1}\right\}$$
$$\times\exp\left[-\sum_{i=1}^{2}\alpha_i\left\{\sum_{j=1}^{r_i}\ln(\frac{1}{x_{ij}})-\{(2m-k)i+(2k-3m)\}\ln T\right\}\right]$$

where $\Omega_J=\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1,r_2+m+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i)}{A_{ik}^{r_i}}\right\}$.

The marginal posterior distribution of each parameter is obtained by integrating out the nuisance parameters. The expectation of each parameter with respect to its marginal posterior distribution gives the Bayes estimator of the parameter under the square error loss function. Marginal posterior distributions follow as under.

$$g_{iJ}(\alpha_i|\mathbf{x}) = \Omega_J^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \frac{\Gamma(r_{3-i})}{A_{3-i,k}^{r_{3-i}}}$$
$$\times \alpha_i^{r_i-1} \exp[-\alpha_i\{\sum_{j=1}^{r_i} \ln(\frac{1}{x_{ij}}) - (k-m)\ln T\}],\ 0<\alpha_i<\infty,\ i=1,2$$

$$g_{3J}(\pi_1|\mathbf{x}) = \Omega_J^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i)}{A_{ik}^{r_i}}\right\} \left\{\prod_{i=1}^{2} \pi_i^{(2m+r_2-r_1+k)i+(2r_1-r_2-3m+2k)}\right\},\ 0<\pi_1<1.$$

7.4.2.1 Bayes Estimators assuming the Jeffreys Prior

The Bayes estimators of $\alpha_1,\ \alpha_2$ and π_1 assuming the Jeffreys prior are:

$$\hat{\alpha}_i = \Omega \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \frac{\Gamma(r_1+2-i)\Gamma(r_2+i-1)}{A_{1k}^{r_1+2-i}\ A_{2k}^{r_2+i-1}},\ i=1,2$$

$$\hat{\pi}_1 = \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+2, r_2+m+1) \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i)}{A_{ik}^{r_i}}\right\}$$

$$\Omega_J^{-1} = \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i)}{A_{ik}^{r_i}}\right\}$$

7.4.2.2 Variances of the Bayes Estimators assuming the Jeffreys Prior

The variances of the Bayes estimators of $\alpha_1,\ \alpha_2$ and π_1 using the Jeffreys prior are:

$$V(\hat{\alpha}_i) = \Omega_J^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \frac{\Gamma(r_1+4-2i)\Gamma(r_2+2i-2)}{A_{1k}^{r_1+4-2i}\ A_{2k}^{r_2+2i-2}} -$$
$$-\left[\Omega_J^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+1, r_2+m+1) \frac{\Gamma(r_1+2-i)\Gamma(r_2+i-1)}{A_{1k}^{r_1+2-i}\ A_{2k}^{r_2+i-1}}\right]^2,\ i=1,2$$

$$V(\hat{\pi}_1) = \Omega_J^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+3, r_2+m+1) \frac{\Gamma(r_1)\Gamma(r_2)}{A_{1k}^{r_1}\ A_{2k}^{r_2}} -$$
$$-\left[\Omega_J^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \mathrm{B}(r_1+k-m+2, r_2+m+1) \frac{\Gamma(r_1)\Gamma(r_2)}{A_{1k}^{r_1}\ A_{2k}^{r_2}}\right]^2$$

where $A_{ik} = \sum_{j=1}^{r_i} \ln(\frac{1}{x_{ij}}) - \{(2m-k)i + (2k-3m)\} \ln T,\ i = 1, 2.$

7.5 The Posterior Distributions assuming the Conjugate Prior

The appropriate informative prior, being compatible with the likelihood, is the Gamma distribution.

7.5.1 The Posterior Distribution assuming the Gamma Prior

Let $\alpha_i \sim Gamma(\mu_i, \sigma_i),\ i = 1, 2$ and $\pi_1 \sim U(0,1)$. Assuming independence, the joint prior, $g(\alpha_1, \alpha_2, \pi_1) \propto \alpha_1^{\mu_1 - 1}\alpha_2^{\mu_2 - 1} \exp\{-(\sigma_1\alpha_1 + \sigma_2\alpha_2)\}$, is incorporated with the likelihood (7.3) to yield a joint posterior distribution of α_1, α_2 and π_1**.** The marginal posterior distribution of each parameter is obtained by integrating out the nuisance parameters. Assuming independence, the joint prior is incorporated with the Likelihood (7.3) to give the joint posterior as follows.

$$g_G(\boldsymbol{\alpha}, \boldsymbol{\pi} | \mathbf{x}) = \Omega_G^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m} \left\{\prod_{i=1}^{2} \pi_i^{(2m+r_2-r_1+k)i+(2r_1-r_2-3m+2k)}\right\}\left\{\prod_{i=1}^{2} \alpha_i^{r_i-1}\right\}$$
$$\times \exp\left[-\sum_{i=1}^{2} \alpha_i \left\{\sum_{j=1}^{r_i} \ln(\frac{1}{x_{ij}}) - \{(2m-k)i + (2k-3m)\} \ln T\right\}\right],\ 0 < \alpha_1,\ \alpha_2 < \infty,\ 0 < \pi_1 < 1.$$

Here $\Omega_G = \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m}\ \mathrm{B}(r_1+k-m+1, r_2+m+1) \left\{\prod_{i=1}^{2} \frac{\Gamma(r_i+\mu_i)}{C_{ik}^{r_i+\mu_i}}\right\}$. Then the marginal posterior densities are obtained by integrating out the nuisance parameters as follows.

$$g_{iG}(\alpha_i|\mathbf{x}) = \Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_{3-i}+\mu_{3-i})}{C_{3-i,k}^{r_{3-i}+\mu_{3-i}}}$$

$$\times\alpha_i^{r_i+\mu_i-1}\exp[-\alpha_i\{\sum_{j=1}^{r_i}\ln(\frac{1}{x_{ij}})-(k-m)\ln T+\sigma_i\}],\ 0<\alpha_i<\infty,\ i=1,2$$

$$g_{3G}(\pi_1|\mathbf{x}) = \Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+\mu_i)}{C_{ik}^{r_i+\mu_i}}\pi_i^{(2m+r_2-r_1+k)i+(2r_1-r_2-3m+2k)}\right\},\ 0<\pi_1<1.$$

7.5.1.1 Bayes Estimators assuming the Gamma Prior

The Bayes estimators under the squared error loss function are obtained by takingexpectations of α_1, α_2 and π_1 with respect to their respective marginal posterior densities as under.

$$\hat{\alpha}_i = \Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+\mu_1+2-i)\Gamma(r_2+\mu_2+i-1)}{C_{1k}^{r_1+\mu_1+2-i}\ C_{2k}^{r_2+\mu_2+i-1}},\ i=1,2$$

$$\hat{\pi} = \Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+2, r_2+m+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+\mu_i)}{C_{ik}^{r_i+\mu_i}}\right\},$$

where $C_{ik} = \sigma_i + \sum_{j=1}^{r_i}\ln(\frac{1}{x_{ij}}) - \{(2m-k)i+(2k-3m)\}\ln T,\ i=1,2.$

7.5.1.2 Variances of the Bayes Estimators assuming the Gamma Prior

The expressions for the variances of the above estimators are as follows.

$$V(\hat{\alpha}_i) = \Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+\mu_1+4-2i)\Gamma(r_2+\mu_2+2i-2)}{C_{1k}^{r_1+\mu_1+4-2i}\ C_{2k}^{r_2+\mu_2+2i-2}} -$$

$$-\left[\Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+\mu_1+2-i)\Gamma(r_2+\mu_2+i-1)}{C_{1k}^{r_1+\mu_1+2-i}\ C_{2k}^{r_2+\mu_2+i-1}}\right]^2, i=1,2$$

$$V(\hat{\pi}_1)=\Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+3,r_2+m+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+\mu_i)}{C_{ik}^{r_i+\mu_i}}\right\}-$$
$$-\left[\Omega_G^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\mathrm{B}(r_1+k-m+2,r_2+m+1)\left\{\prod_{i=1}^{2}\frac{\Gamma(r_i+\mu_i)}{C_{ik}^{r_i+\mu_i}}\right\}\right]^2$$

7.5.2 Predictive Distribution

The predictive distribution contains the information about the independent future random observation given the observations, already accomplished.

7.5.2.1 Predictive Intervals assuming the Gamma Prior

The posterior predictive distribution of a future observation y is defined as follows

$$p(y|\mathbf{x})=\int_0^\infty\int_0^\infty\int_0^1 g(\alpha_1,\alpha_2,\pi_1|\mathbf{x})\,p(y|\alpha_1,\alpha_2,\pi_1)\,d\pi_1\,d\alpha_2\,d\alpha_1$$

Where $p(y|\alpha_1,\alpha_2,\pi_1)=\pi_1\alpha_1 y^{\alpha_1-1}+\pi_2\alpha_2 y^{\alpha_2-1}$ is the Pareto mixture model described in Section 7.2 and $g(\alpha_1,\alpha_2,\pi_1|\mathbf{x})$ is the joint posterior distribution given in Section 7.5.1.

The posterior predictive distribution of the future observation y becomes as under.

$$p(y|\mathbf{x})=\Omega^{-1}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\{\mathrm{B}(r_1+k-m+2,r_2+m+1)\frac{\Gamma(r_1+\mu_1+1)}{y(C_{1k}-\ln y)^{r_1+\mu_1+1}}$$
$$\times\frac{\Gamma(r_2+\mu_2)}{C_{2k}^{r_2+\mu_2}}+\mathrm{B}(r_1+k-m+1,r_2+m+2)\frac{\Gamma(r_1+\mu_1)}{C_{1k}^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2+1)}{y(C_{2k}-\ln y)^{r_2+\mu_2+1}}\},\ y>0$$

$$\Omega=\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\left\{\prod_{i=1}^{k}\frac{\Gamma(r_i+\mu_i)}{C_{ik}^{r_i+\mu_i}}\right\}\{\mathrm{B}(r_1+k-m+2,r_2+m+1)+\mathrm{B}(r_1+k-m+1,r_2+m+2)\}$$

A $(1-a)100\%$ Bayesian predictive interval (L,U) can be obtained by solving the two equations

$\int_0^L p(y|\mathbf{x})\, dy = \frac{\alpha}{2} = \int_U^1 p(y|\mathbf{x})\, dy$ which can be expressed as follows.

$$\frac{\alpha}{2} = \Omega^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m}\{ \mathrm{B}(r_1+k-m+2, r_2+m+1) \frac{\Gamma(r_1+\mu_1)}{(C_{1k}-\ln L)^{r_1+\mu_1}} \frac{\Gamma(r_2+\mu_2)}{C_{2k}{}^{r_2+\mu_2}}$$
$$+\mathrm{B}(r_1+k-m+1, r_2+m+2) \frac{\Gamma(r_1+\mu_1)}{C_{1k}{}^{r_1+\mu_1}} \frac{\Gamma(r_2+\mu_2)}{(C_{2k}-\ln L)^{r_2+\mu_2}}\}$$

$$\frac{\alpha}{2} = \Omega^{-1} \sum_{k=0}^{n-r} \sum_{m=0}^{k} (-1)^k \binom{n-r}{k}\binom{k}{m}[\frac{\mathrm{B}(r_1+k-m+2, r_2+m+1)\Gamma(r_1+\mu_1)\Gamma(r_2+\mu_2)}{C_{2k}{}^{r_2+\mu_2}}\{\frac{1}{C_{1k}{}^{r_1+\mu_1}}$$
$$-\frac{1}{(C_{1k}-\ln U)^{r_1+\mu_1}}\} + \frac{\mathrm{B}(r_1+k-m+1, r_2+m)\Gamma(r_1+\mu_1)\Gamma(r_2+\mu_2)}{C_{1k}{}^{r_1+\mu_1}}\{\frac{1}{C_{2k}{}^{r_2+\mu_2}} - \frac{1}{(C_{2k}-\ln U)^{r_2+\mu_2}}\}]$$

These predictive intervals when evaluated for a number of combinations of the hyperparameters can help locate a range of hyperparameters that may lead to the informative Bayes estimates having smaller variances than the uninformative Bayes estimates. Saleem and Aslam (2008b) used predictive intervals for the Rayleigh mixture to discuss precision of Bayes estimates in terms of hyperparameters. If a trend in terms of the hyperparameters is observed for the narrower predictive intervals, then a sort of objectivity may be added to the prior information provided by a number of experts.

7.6 The Complete Sample Expressions

Under the conditions given in Section 2.6 and letting $T \to 1$ the expressions for the Bayes estimators and their variances are simplified as given in Tables 7.1-7.2. The comments regarding amount of information, computational ease and simplification quoted in Section 2.6 also apply here.

7.7 A Simulation Study

A simulations study reflects the performance of Bayes estimates under the impact of sample size and censoring rate. Samples of sizes $n = 50, 100, 200$ were generated from the two component mixture of Power Function distribution with parameters α_1, α_2 and π_1 such that

Table 7.1 The complete sample expressions for the Bayes (Uniform), Bayes (Jeffreys) and ML estimators as $T \rightarrow 1$

Parameters	Bayes Estimators (Uniform)	Bayes Estimators (Jeffreys)	ML Estimators
α_1	$\dfrac{n_1+1}{\sum \ln(\frac{1}{x_{1j}})}$	$\dfrac{n_1}{\sum \ln(\frac{1}{x_{1j}})}$	$\dfrac{n_1}{\sum \ln(\frac{1}{x_{1j}})}$
α_2	$\dfrac{n_2+1}{\sum \ln(\frac{1}{x_{2j}})}$	$\dfrac{n_2}{\sum \ln(\frac{1}{x_{2j}})}$	$\dfrac{n_2}{\sum \ln(\frac{1}{x_{2j}})}$
π_1	$\dfrac{n_1+1}{n+2}$	$\dfrac{n_1+1}{n+2}$	$\dfrac{n_1}{n}$

Table 7.2 The complete sample expressions for the variances of the Bayes (Uniform), Bayes (Jeffreys) and ML estimators as $T \rightarrow 1$

Parameters	Variances of Bayes Estimators (Uniform prior)	Variances of Bayes Estimators (Jeffreys prior)	Variances of ML Estimators
α_1	$\dfrac{n_1+1}{(\sum \ln(\frac{1}{x_{1j}}))^2}$	$\dfrac{n_1}{(\sum \ln(\frac{1}{x_{1j}}))^2}$	$\dfrac{n_1}{(\sum \ln(\frac{1}{x_{1j}}))^2}$
α_2	$\dfrac{n_2+1}{(\sum \ln(\frac{1}{x_{2j}}))^2}$	$\dfrac{n_2}{(\sum \ln(\frac{1}{x_{2j}}))^2}$	$\dfrac{n_2}{(\sum \ln(\frac{1}{x_{2j}}))^2}$
π_1	$\dfrac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\dfrac{(n_1+1)(n_2+1)}{(n+2)^2(n+3)}$	$\dfrac{n_1 n_2}{n^3}$

$\pi_1 \in \{0.25, 0.40\}$ and $(\alpha_1, \alpha_2) \in \{(0.25, 0.8), (0.25, 4.0), (4.0, 1.2)\}$. Probabilistic mixing was used to generate the mixture data. For each observation a random number u was generated from the Uniform on [0, 1] distribution. If $u < \pi_1$, the observation was taken randomly from F_1 (the Power Function distribution with parameter α_1) and if $u > \pi_1$, the observation was taken randomly from F_2 (the Power Function distribution with parameter α_2). The choice of the censoring time, in each case, was made in such a way that the censoring rate in the resulting sample to be approximately 10% and 20%. The results of the simulation study are presented in Tables 7.3-7.5. A careful study of the Tables 7.3-7.4 depicts many interesting properties of the Bayes estimates while Table 7.5 presents an interesting among three Bayes estimates.

7.8 A Real Life Example

Mendenhall and Hader (1958) mixture data $\mathbf{t} = (t_{11}, t_{12}, \ldots, t_{1r_1}, t_{21}, t_{22}, \ldots, t_{2r_2})$ consisting of hours to failure for ARC-1 VHF radio transmitter receivers of a single commercial airline is considered. The radio transmitter receivers that seemed to be failed at or before 630 hours of operation were removed from the aeroplanes as a general policy of the airline giving Type-I right censored observations at $T = 630$ hours. On the other hand, inspection of the failed units allowed the engineers to allocate the failed units to any one of the two different subpopulations. The mixture failure data can be found on page 509 in Mendenhall and Hader (1958). Mendenhall and Hader fitted Exponential distribution to this data. The transformation $x = \exp(-t)$ of an Exponential random variable $(\mathbf{t})$ yields a Power Function random variable $(\mathbf{x})$. This property allows us to use the transformed Mendenhall and Hader data set for our analysis. It is interesting to note that despite the transformation almost no major computations are required to have the data summary required to evaluate the proposed estimates. For instance, $\sum_{j=1}^{r_1} \ln\left(\frac{1}{x_{1j}}\right) = \sum_{j=1}^{r_1} t_{1j} = 20458$ and $\sum_{j=1}^{r_2} \ln\left(\frac{1}{x_{2j}}\right) = \sum_{j=1}^{r_2} t_{2j} = 50056$. While other sample characteristics required are:

$n = 369, r_1 = 107, r_1 = 218, r = r_1 + r_2 = 325$, $n = 369,\ r_1 = 107,\ r_2 = 218,\ r = r_1 + r_2 = 325$.

The Power Function mixture parameters $(\alpha_1,\ \alpha_2,\ \pi_1)$ can be evaluated using the estimators derived in Sections 7.4-7.5. The Bayes (Jeffreys) estimates of Power Function mixture parameters are computed as $\hat{\alpha}_1 = 0.005,\ \hat{\alpha}_2 = 0.004$ and $\hat{\pi}_1 = 0.331$ (corrected to three decimal places) with their respective standard errors, $SE(\hat{\alpha}_1) = 0.0005$, $SE(\hat{\alpha}_2) = 0.000305$ and $SE(\hat{\pi}_1) = 0.0261$. The lifetime estimates of the average lifetimes of the two subgroups of the radio transmitter receivers require a subtle re-parameterization and turn out to be $\hat{\alpha}_{11} = \frac{1}{\hat{\alpha}_1} = \frac{1}{0.005} = 200$ and $\hat{\alpha}_{22} = \frac{1}{\hat{\alpha}_2} = \frac{1}{0.004} = 250$ while the proportion estimate remains the same as $\hat{\pi}_1 = 0.331$. The standard errors of these lifetime estimates are computed as $SE\left(\hat{\alpha}_{11}\right) = 17.7513$, $SE\left(\hat{\alpha}_{22}\right) = 14.6994$ while the standard error of proportion estimate is again unchanged as $SE(\hat{\pi}_1) = 0.0261$. In general, the standard errors of all the three proposed estimates are much lower than the respective standard errors of the estimates presented in Mendenhall and Hader (1958) and Sinha (1998). The proposed proportion estimate is almost the same while the proposed lifetime estimates seem to be under-estimated but on the other hand it is encouraging to note that the proposed lifetime estimates are greater than the corresponding sample average lifetimes of the two subgroups i.e., $\overline{t_1} = 191.2 < 200, \overline{t_2} = 229.6 < 250$ as is expected in the right censoring situations.

7.9 Conclusion

The simulation study displays some interesting properties of the Bayes estimates. The estimates of the first component density parameter seem to be under-estimated with some exceptions. It is interesting to note that when the component density parameters are very well separated, the estimates of the first component density parameter seem to be over-estimated with rare exception. The estimates of the second density parameter are under-estimated while the estimates of the proportion parameter are over-estimated.

The effect of an increase in censoring on the estimates of component density and proportion parameters is in terms of an increase in the intensity of over or under-estimation. However, with an increase in the censoring rate, the variances of the estimates of the component densities parameters are reduced while the variances of the estimates of mixing proportion parameter are increased.

The extent of over-estimation or under-estimation is more intensive for larger parameter values of the component density and proportion parameters. Also, the variances of the estimates of the component density parameters seem to be quite large (small) for the relatively larger (smaller) values of the parameters. The variance of the estimate of population proportion is slightly increased for larger values of proportion parameter. Furthermore, increasing the sample size reduces the variance of estimate of the component density and proportion parameters. The increase (decrease) in proportion of a component in the mixture reduces (increases) the variance of the estimate of the corresponding component density parameter.

As the cut off sensor value tends to unity, the complete sample expressions for the estimators and variances are greatly simplified. In addition, variances of the complete sample estimates are expected to be reduced further as there in no more effect of censoring.

The Bayes (Jeffreys) estimates of the first component density parameter are under-estimated while the Bayes (Uniform) estimate of the first lifetime parameter are over-estimated with rare exception. Both the uninformative estimates of the second component density parameter are under-estimated but the size of under-estimation is greater in case of Bayes (Jeffreys) estimates which have smaller variances. On the other hand, both Bayes (Uniform) and Bayes (Jeffreys) estimates of the mixing proportion parameter are over-estimated but the degree of over-estimation is relatively lower in case of Bayes (Uniform). The Bayes (Gamma) estimates under-estimate the component density parameters while over-estimate the mixing proportion parameter. The Bayes estimates with informative (Gamma) prior
seem to be more efficient than their uninformative counterparts with a few exceptions. A better choice of hyperparameters may further improve the efficiency of Bayes (Gamma) estimates.

In the real life example, the proposed estimates of lifetime parameters are under-estimated as compared with Mendenhall and Hader estimates but they are much greater than the respective sample mean lifetime hours as expected in the censoring situations. However, the proposed estimate of the mixing proportion parameter is very close to its Mendenhall and Hader counterpart. On the other hand the standard errors of all the three proposed estimates are much smaller than the respective standard errors of the estimates based on the Mendenhall and Hader data which are found in literature so far.

Table 7.3 Bayes (Jeffreys) estimates and standard errors (in parenthesis) of the Power Function mixture parameters $\alpha_1 = 0.25,\ \alpha_2 = 0.80,\ \pi_1 = 0.25,\ 0.40$ with censoring rates, $C = 10\%,\ 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	10% Censoring		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.25,0.80,0.25)	50	0.25699(0.07913)	0.69890(0.11625)	0.27988(0.06417)
	100	0.24843(0.05176)	0.70151(0.08360)	0.27580(0.04644)
	200	0.24476(0.03550)	0.70452(0.05513)	0.27178(0.03296)
(0.25,0.80,0.40)	50	0.24768(0.05574)	0.68351(0.12279)	0.43998(0.07646)
	100	0.24332(0.04001)	0.68707(0.09317)	0.42836(0.05211)
	200	0.24182(0.02811)	0.68538(0.06095)	0.42771(0.03745)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.25,0.80,0.25)	50	0.23887(0.06792)	0.59072(0.09852)	0.31059(0.07440)
	100	0.23515(0.04750)	0.59785(0.06784)	0.30317(0.05306)
	200	0.22895(0.03266)	0.59675(0.04842)	0.29708(0.03675)
(0.25,0.80,0.40)	50	0.22746(0.05258)	0.57930(0.10895)	0.46877(0.07859)
	100	0.22499(0.03453)	0.57602(0.07515)	0.46362(0.05484)
	200	0.22450(0.02422)	0.57533(0.04989)	0.46520(0.03873)

* Bayes (Jeffreys) estimates means the Bayes estimates assuming the Jeffreys prior.

Table7. 4 Bayes (Jeffreys) estimates and standard errors (in parenthesis) of Power Function mixture parameters $\alpha_1 = 0.25,\ \alpha_2 = 4.0,\ \pi_1 = 0.25,\ 0.40$ with censoring rates, $C = 10\%,\ 20\%$.

$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**10% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.25, 4.0,0.25)	50	0.27262(0.08456)	3.45464(0.58315)	0.29606(0.06569)
	100	0.25680(0.05241)	3.44268(0.40634)	0.28711(0.04848)
	200	0.25372(0.03697)	3.45236(0.27788)	0.28400(0.03504)
(0.25, 4.0,0.40)	50	0.25840(0.06170)	3.33087(0.63844)	0.45969(0.07703)
	100	0.25343(0.04250)	3.3663(0.44998)	0.44790(0.05318)
	200	0.24856(0.02779)	3.33515(0.29534)	0.44766(0.03752)
$(\alpha_1,\ \alpha_2,\ \pi_1)$	n	**20% Censoring**		
		$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
(0.25, 4.0,0.25)	50	0.26582(0.08309)	2.88349(0.47852)	0.26582(0.08309)
	100	0.25693(0.05164)	2.91572(0.33675)	0.25693(0.05164)
	200	0.24911(0.03504)	2.90049(0.22766)	0.24911(0.03504)
(0.25, 4.0,0.40)	50	0.25738(0.05936)	2.71112(0.50059)	0.25738(0.05936)
	100	0.24967(0.03942)	2.71219(0.36629)	0.24967(0.03942)
	200	0.24615(0.02740)	2.72858(0.25181)	0.24615(0.02740)

Table 7.5 A comparison of Bayes (Uniform), Bayes (Jeffreys) and Bayes (Gamma)* estimates and standard errors (in parenthesis) of Power Function mixture parameters $\alpha_1 = 0.25,\ \alpha_2 = 0.80,\ \pi_1 = 0.25,\ 0.40$ with censoring rate, $C = 10\%$.

Prior	π_1	n	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\pi}_1$
Uniform	0.25	50	0.28052(0.08780)	0.72078(0.11999)	0.27987(0.06421)
		100	0.25918(0.05433)	0.71233(0.08487)	0.27580(0.04645)
		200	0.24995(0.03629)	0.70991(0.05554)	0.27178(0.03296)
	0.40	50	0.261191(0.05930)	0.711567(0.12798)	0.440075(0.07652)
		100	0.249875(0.04112)	0.700587(0.09493)	0.428405(0.05212)
		200	0.24503(0.02850)	0.69206 (0.06151)	0.42773 (0.03746)
Jeffreys	0.25	50	0.23887(0.06792)	0.59072(0.09852)	0.31059(0.07440)
		100	0.23515(0.04750)	0.59785(0.06784)	0.30317(0.05306)
		200	0.22895(0.03266)	0.59675(0.04842)	0.29708(0.03675)
	0.40	50	0.24768(0.05574)	0.68351(0.12279)	0.43998(0.07646)
		100	0.24332(0.04001)	0.68707(0.09317)	0.42836(0.05211)
		200	0.24182(0.02811)	0.68538(0.06095)	0.42771(0.03745)
Gamma	0.25	50	0.22801(0.05992)	0.62844(0.09346)	0.27966(0.06407)
		100	0.23524(0.04584)	0.66504(0.07513)	0.275701 (0.04641)
		200	0.23844(0.03360)	0.68595(0.05231)	0.27173(0.03295)
	0.40	50	0.23131(0.04766)	0.59756(0.09364)	0.43952(0.07633)
		100	0.23542(0.03734)	0.64292(0.08174)	0.42815(0.05207)
		200	0.23795(0.02717)	0.66306(0.05717)	0.42761(0.03744)

* Bayes (Gamma) estimates means the Bayes estimates assuming the Gamma prior.

CHAPTER 8

CONCLUSION AND RECOMMENDATIONS

The type-I mixtures of distributions belonging to a subclass of single parameter exponential family are considered. All the members of this subclass are related with exponential distribution by one one link or the other as described in the WinBUGS code in case of Pareto mixture. Also the common feature of this subclass is the availability of the closed form Bayes estimates of the mixture parameters. The motivation behind the consideration of this subclass is to explore the closed form expressions for the Bayes estimators and their variances as well as the conduct of the real life applications in a number of ways. However, the MCMC methods e.g., Gibbs Sampling can be applied to deal with gerneralized versions of the distributions.

An extensive simulation study is conducted for Exponential, Rayleigh, Pareto, Burr and Power Function mixtures. The simulation study is conducted to explore some interesting properties of the Bayes estimates for various combinations of the component density parameters, various mixing proportions, different sample sizes and for different censoring rates. Although the findings for each mixture were reported in the respective chapters yet it is considered useful to present an overall conclusion in terms of parameters of the component densities, mixing proportion parameter of all the mixtures under study.

The first component density parameter is observed to be over-estimated in all the mixtures except the Power Function mixture where it is under-estimated. The second component density parameter is also under-estimated in Power Function mixture while in rest of the mixtures, its estimates have a mixed behavior, sometimes over-estimated and sometimes under-estimated. On the other hand, the mixing proportion parameter is over-estimated in all the mixtures under study except in Pareto and Burr mixtures where it is under-estimated.

The effect of censoring rates on estimates of parameters in different mixtures is as follows. The degree of over-estimation of the estimates of the first component density parameter increases with the increase in the censoring rate except in Power Function mixture where degree of under-estimation of under-estimated estimates increases. The response of the estimates of the second component density parameter is different. It is observed that the slight over-estimation of these estimates turns into a slight under-estimation while the degree of under-estimation increases with the increase in censoring rate for the under-estimated estimates. So far as the behavior of the estimates of the mixing proportion parameter is concerned, their degree of over-estimation increases in Exponential, Rayleigh and Power Function mixtures whereas the degree of under-estimation increases in Burr and Pareto mixtures.

The effect of censoring on the variance of estimates is another aspect of interest. It is interesting to note that the variances of the estimates of the parameter of the first component density increase with the increase in censoring rates in all the mixtures. The variances of the estimates of the parameters of the second component density parameter also increase except in Power Function mixture where the variance of estimates decreases with an increase in the censoring rates. The same is the response of the estimates of the mixing proportion parameter towards an increase in the censoring rate.

How does the magnitude of proportion of a component density in a mixture affect the estimates of parameter of that component density? This is interesting to note that the variances of estimates of the parameter of a component density decreases when its proportion to the mixture is increased. Also the degree of over-estimation of the estimates of parameter of a component density decreases provided the mixing proportion of that component in the mixture is increased. The situation is a bit different in case of the Power Function mixture where the degree of under-estimation is increased and hence the under-estimation becomes more severe. However, the reduction of variance in the estimates of the mixing proportion is not ensured. The variances of the estimates of the mixing proportion parameter decrease with some exceptions in Exponential and Burr mixtures while increase in Pareto and Power Function mixtures while increase in Rayleigh mixture with some rare exceptions and display a mixed behavior in case of Burr mixture. Recall that the estimates of the mixing proportion parameter are over-estimated in Exponential, Rayleigh and Power Function mixtures while under-estimated in Burr and Pareto mixtures. With an increase in the mixing proportion, the degree of over-estimation of the mixing proportion estimates decreases (increases) in Exponential and Rayleigh (Power Function) mixtures while trends are mixed in case of Burr and Pareto mixtures.

What is the effect of the magnitudes or sizes of the parameters of the component densities on the estimates in various mixtures? It is observed that the extent of over-estimation of the estimates of the parameter of the first component density increases in Exponential and Rayleigh mixtures while no definite trend could be observed in Pareto, Burr and Power Function mixtures. In case of the estimates of the parameter of the second component density, the extent of over-estimation is seen to be increased in Exponential and Rayleigh mixtures with rare exceptions, while no certain pattern could be established in case of Pareto, Burr and Power Function mixtures. Now we come to the effect of parameter sizes on the variances of the estimates. It is interesting to note that variances of the estimates of the first component density parameter tends to increase in Exponential, Rayleigh, Burr and Power Function mixtures while in case of Pareto mixture, there has been observed some rare exceptions. So far as the estimates of the parameter of the second component density is concerned, the variances are noticed to be increased in Exponential, Rayleigh, Pareto, Burr and Power Function mixtures in response to an increase in the parameter sizes.

The above mentioned properties and trends are analyzed with the help of only the Bayes estimates assuming the Jeffreys prior. In each simulation study for every mixture under consideration, the Bayes estimates with Uniform prior, the Bayes estimates with Jeffreys prior and the Bayes estimates with some informative conjugate prior have also been computed and compared. Here comes an overall comparison of these Bayes estimates in terms of various mixtures and parameters.

The Bayes (Uniform), the Bayes (Jeffreys) and the Bayes (Informative) estimates of the parameter of the first component density are all observed to be over-estimated in all the mixtures except the Power Function mixture where the Bayes (Jeffreys) and Bayes (Gamma) are under-estimated while the Bayes (Uniform) estimates are over-estimated with some exceptions. In Exponential mixture, the Bayes (Inverted Gamma) estimates are closer to the true parameter while in Rayleigh mixture, the Bayes (Jeffreys) are observed close to the parameter. In Pareto mixture, the Bayes (Gamma) estimates are found close to the parameter and the Bayes (Gamma) are close to the parameter in Burr mixture. However, in case of the Power Function mixture, the Bayes (Uniform) are very close to the parameter as compared to the other two Bayesian counter parts.

As is quoted in many classical sources, that precision is more appreciable property of an estimate as compared to the unbiasedness. It is observed in all the mixtures under consideration that the Bayes estimates of the parameters of the component densities of the mixtures which are based on suitable informative prior are more efficient than their Uniform and Jeffreys counterparts. However, there may be some rare exceptions which can easily be overcome by a better choice of the hyperparameters. As the hyperparameters of the informative prior are outcome of prior information, so the quality of the Bayes estimates based on informative prior can further be improved if some better prior information is available. So the quality the Bayes estimates with informative prior depends upon the quality of the prior information available.

In our study a real life mixture data has been analyzed by different mixtures using some suitable transformations of the data, so that the results may be compared with each other and with the results of Davis (1952), Mendenhall and Hader (1958) and Sinha (1998). The estimates of Exponential mixture parameters using real life data are superior to the previous results found in literature. The estimates obtained with the help of Pareto mixture are close to the previously available estimates but having lesser variances. Finally the estimates based on Power Function mixture are, although, under-estimated but have much lower standard errors than those given in literature so far.

All the mixtures discussed in this thesis are of type-I, samples of type-IV with ordinary type-I, right censoring. The elegant closed form Bayes estimates are provided. Obviously, the work may be extended in many possible directions. Some mixtures of the single (multiple) parameter distributions, discrete or continuous, may be considered which have not been explored so far. Such mixtures can be explored with a careful study of many real life situations where the observations coming from some heterogeneous populations which can be assumed to be characterized by the some particular component densities. So a possible extension is to consider mixtures of the generalized versions of the distributions where the use of Monte Carlo Markov Chains (MCMC) techniques would be inevitable. A WinBUGS code is presented in Chapter 5 to deal with such cases. A code in R could meet the same purpose.

Although algebraic expressions are derived for finite mixtues in most of the chapters, yet simulation study for each mixture are conducted for two component case. The simulation study can also be extended to larger number of components of the mixtures with various mixing weights. As the number of components of the mixture depends upon the number of causes of death of the objects under study, so the mixtures with more than two components can also be considered in the light of real life bio-medical and industrial applications.

REFERENCES

Abdel-All, N.H., Mahmoud, M.A.W. and Abd-Ellah H. N. (2003). Geometrical properties of Pareto distribution. Applied Mathematics and Computation, 145: 321–339.

Abu-Taleb, A.A., Smadi, M.M. and Alawneh, A.J. (2007). Bayes estimation of the lifetime parameters for the Exponential distribution. J. Math. Stat., 3(3): 106-108.

Acheson, M.A. and McElwee, E.M. (1952). Concerning the reliability of electron tubes. Proceedings of the IRE., 40(10): 1204-1206.

Ahmed, E.S., Volodin, A.I. and Hussein, A. (2005). Robust weighted likelihood estimation of Exponential parameters. IEEE Trans. Reliab., 54(3): 389-395.

Ahsanullah, M. (1988). Introduction to Record Statistics. Needham Heights, MA: Ginn Press.

Ahsanullah, M. and Lutful Kabir, A.B.M. (1974). A characterization of the Power Function distribution. The Canadian journal of Statistics, 2(1): 95-98.

AlHussaini, E.K., Mousa, M.A. and Sultan, K.S. (1997). Parametric and nonparameteric estimation of $P(Y<X)$ for finite mixtures of lognormal components. Communications in Statistics- Theory and Methods, 26(5): 1269-1289.

Ali, M.M. and Nadarajah, S. (2006). A truncated Pareto distribution. Computer Communications, 30: 1-4.

Ali, M.M., Woo, J. and Nadarajah, S. (2005). Bayes estimators of the Exponential distribution. Journal of statistics and management systems, 8(1): 53-58.

Ali, M.M., Woo, J. and Nadarajah, S. (2005). On the Ratio $X/(X + Y)$ for the Power Function Distribution. Pakistan Journal of Statistics, 21(2): 131–138

Arcidiacono, P. and Jones, J.B. (2003). Finite mixture distributions, sequential likelihood and the EM Algorithm. Econometrica, 71(3): 933-946.

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

Bernardo, J.M. (1979). Reference posterior distribution for Bayesian inference. Journal of Royal Statistical Society, B, 41: 113-147.

Bhat, B.A. (2005). On Bayes Estimation and Reliability Functions for Mixtures of Two Pareto Life Time Distributions from Right Censored Samples. International Workshop /Conference on Bayesian Statistics and its Applications, January 6-8, 2005 Banaras Hindu University, Varanasi, India.

Böhning, D. (2000). Computer Assisted Analysis of Mixtures and Applications: Meta-Analysis, Disease Mapping, and Others. Chapman & Hall/CRC, London.

Böhning, D., Dietz, E. & Schlattmann, P. (1998). Recent developments in C.A.MAN (Computer Assisted Analysis of Mixtures). Biometrics, 54: 525-536

Böhning, D. and Seidel, W. (2003). Editorial: Recent developments in mixture models. Computational Statistics and Data Analysis, 41: 349-357.

Böhning, D., Seidel, W., Alfó, M., Garel, B. Patilea, V. and Walther, G. (2007). Advances in Mixture Models. Computational Statistics and Data Analysis, 51: 5205-5210.

Bolstad, W.M. (2004). Introduction to Bayesian Statistics. John Wiley & Sons, Inc. New York.

Burger, J.O. (1985). Statistical decision theory and Bayesian analysis. 2^{nd} Ed. Springer-Verlag, New York.

Burr, I.W. (1942). Cummulative frequency functions, Annals of Mathematical Statistics, 13: 215-232.

Burr, I.W. (1968). On a general system of distributions III, The sample range. Journal of the American Statistical Association, 63: 636-643.

Burr, I.W. (1973). Parameters for a general system of distributions to match a grid of c3 and a4. Communications in Statistics, 2: 1-21.

Burr, I.W. and Cislak, P.J. (1968). On a general system of distributions I. Its curve-shape characteristics, II. The sample median. Journal of the American Statistical Association, 63: 627-635.

Chandra, S. (1977). On the mixtures of probability distributions. Scandinavian Journal of Statistics, 4: 105-112.

Chuan-Chong, C. and Mhee-Meng, K. (1992). Principles and Techniques in Combinatorics, World Scientific Publishing Co. Pvt. Ltd. Singapore.

Davis, D.J. (1952). An analysis of some failure data. Journal of American Statistical Association, 47: 113-50.

Dey, S. and Das, M.K. (2005). A note on Predictive Intervals for a Rayleigh distribution: Bayesian approach. International Workshop/Conference on Bayesian Statistics and its applications. January 6-8, 2005. Banaras Hindu University, Varanasi, India.

Dietz, E., and Böhning, D. (1997). The use of two-component mixture models with one completely or partly known component. Computational Statistics, 12:219-234.

Dietz, K., Gail, M., Krickeberg, K. and Singer, B. (1996). Introduction to Survival Analysis. Springer-Verlag, Inc. New York.

Economou, P.and Caroni, C. (2005). Graphical Tests for the Assumption of Gamma and Inverse Gaussian Frailty Distributions. Lifetime Data Analysis, 11(4): 565-582.

Elsherpieny, E.A. (2007). Estimation of mixed generalized exponentially distributions from censored type-I samples. Journal of Applied Sciences Research, 3(12): 1696-1700.

Epstein, B. and Sobel, M. (1953). Life testing. Journal of American Statistical Association, 48: 486-502.

Everitt, B. and Hand, D.J. (1981). Finite Mixture Distributions. Chapman and Hall. New York.

Fernandez, A.J. (2000). Bayesian inference from Type-II doubly censored Rayleigh data. Statistics & Probability Letters, 48(4): 393-399.

Geoffrey, J.M. and David, P. (2000). Finite Mixture Models. John Wiley and Sons.

Ghosh, J.K., Delampady, M. and Samanta, T. (2006). An Introduction to Bayesian Analysis: Theory and Methods. Springer Science+Business Media LLC, New York.

Ghosh, S.K. and Ebrahimi, N. (2001). Bayesian analysis of the mixing function in a mixture of two Exponential distributions. Institute of Statistics Mimeo Series No. 2531.

Hebert, J.L. and Scariano, S.M. (2004). Comparing location estimators for Exponential mixtures under Pitman's measure of closeness. Communications in Statistics- Theory and Methods, 33(1): 29-46.

Hirai, A.S. (1972). Estimation of scale and location parameters, using quadratic coefficients. Australian Journal of Statistics, 14: 222-226.

Hirai, A.S. (1976). Estimation of the scale parameter from the Type-II censored Rayleigh distribution. Punjab University Journal of Mathematics, 9: 61-67.

Ismail, S. (2004). A simple estimator for the shape parameter of the Pareto distribution with economics and medical applications. Journal of Applied Statistics, 31(1): 3–13.

Jeffreys, H. (1961). Theory of Probability, 3rd Ed. Clarendon Press. Oxford.

Jiang, S. and Kececioglu, D. (1992). Maximum likelihood estimates, from censored data, for mixed- Weibull distributions. IEEE Transactions on Reliability, 41(2): 241-247

Johnson, N., Kotz, S. and Balakrishnan, N. (1994). Continuous Univariate Distributions. John & Wiley Sons, Inc. New York.

Kalbfleisch, J.D. and Prentice, R.L. (2002). The Statistical Analysis of Failure Time Data. John Wiley & Sons, Inc. New York.
Kass, R.E. (1990). Data translated likelihood and Jeffreys Rule. Biometrika, 77: 117-134.

Klein, J. and Moeschberger, M. (1997). Survival Analysis Techniques for Censored and Truncated Data. Springer-Verlag, Inc. New York.

Lalitha, S. and Mishra, A. (1996). Modified maximum likelihood estimation for Rayleigh distribution. Communications in Statistics: Theory and Methods, 25: 389-401.

Lee, P.M. (1997). Bayesian Statistics: An Introduction. 2nd Ed. Arnold, London.

Leemis, L.M. (1995). Reliability: Probabilistic Models and Statistical Methods. Prentice-Hall, Englewood Cliffs. New Jersey.

Li, L.A. (1983). Decomposition theorems, conditional probability, and finite mixtures distributions. Thesis, State University, Albany. New York,

Li, L.A. and Sedransk, N. (1982). Inference about the presence of a mixture. Technical Report, State University, New York, Albany.

Li, L.A. and Sedransk, N. (1988). Mixtures of distributions: A topological approach. The annals of Statistics, 16(4): 1623-1634.

Lindsay, B.G. (1995). Mixture Models: Theory, Geometry, and Applications. IMS.

McCullagh, Peter. (1994). Exponential mixtures and quadratic Exponential families. Biometrika, 81(4): 721-9.

McLachlan, G.J. and Bastford, K.E. (1988). Mixture Models: Inference and Applications to Clustering. Marcel Dekker, New York.

Mendenhall, W. and Hader, R.A. (1958). Estimation of parameters of mixed Exponentially distributed failure time distributions from censored life test data. Biometrika, 45(3-4): 504-520.

Meniconi, M. and Barry, D.M. (1996). The Power Function distribution: a useful and simple distribution to assess electrical component reliability. Microelectronics Reliability, 36(9): 1207-1212.

Mi, J. (1999). Age-smooth properties of mixture models. Statistics and Probability Letters, 43: 225-236.

Mostert, P.J., Bekker, A. and Roux, J.J.J. (1998). Bayesian analysis of survival data using the Rayleigh model and linex loss, South Africa Statistical Journal, 32(1): 19-42.

Nadarajah, S. and Kotz, S. (2005). Information matrix for a mixture of two Pareto distributions. Iranian Journal of Science and Technology Trans A., 29(A3): 377-385.

Nobile, A. (1998). Bayesian Analysis of Finite Mixture Distribution. UMI.

Pearson, K. (1894). Contribution to the mathematical theory of evolution. Phil. Trans. Roy. Soc. A., 185: 71-110.

Rachev, S.T. and SenGupta, A. (1993). Laplace-Weibull mixtures for modeling price changes. Management Science, 39 (8): 1029-1038.

Raqab, M.Z. (1992). Predictors of Future Order Statistics from Type-II censored samples. Unpublished PhD Thesis, Columbus, Ohio State University.

Raqab, M.M. and Ahsanullah, M. (2001). Estimation of location and scale parameters of generalized Exponential distribution based on order statistic. Journal of Statistical Computation and Simulation, 22: 112-127.

Redner, R.A. and Walker, H.F. (1984). Mixture densities, maximum likelihood and the EM Algorithm. SIAM Review, 26(2): 195-239.

Rodriguez, R.N. (1977). A guide to the Burr Type-XII distributions. Biometrika, 64: 129-134.

Saleem, M. and Aslam, M. (2008a). 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.

Saleem, M. and Aslam, M. (2008b). On prior selection for the mixture of Rayleigh distribution using predictive Intervals. Pakistan Journal of Statistics, 24(1): 21-35.

Saleem, M. and Aslam, M. (2009). On Bayesian analysis of the Rayleigh survival time assuming the random censor time. Pakistan Journal of Statistics, 25(2): 71-82.

Saleem, M., Aslam, M. and Economou, P. (2010). On Bayesian analysis of the Power Function mixture using censored and complete sample. Journal of Applied Statistics, 37(1): 25-40.

Sankaran, P.G. Nair, M.T. (2005). On finite mixture of Pareto distributions. CSA Bulletin, 57: 225-226.

Scallan, A.J. (1992). Maximum likelihood estimation for a Normal/Laplace Mixture distribution. The Statistician, 41(2): 227-231.

Schlattmann, P. (2009). Medical Applications of Finite Mixture Models. Springer.

Schnatter, S.F. (2006). Finite Mixture and Markov switching models. Springer.

Sinha, S.K. (1986). Reliability and Life Testing. Wiley Eastern Limited, New Delhi.

Sinha, S.K. (1998). Bayesian Estimation, New Age International (P) Limited. Publishers, New Delhi.

Sinha, S.K. and Howlader, H.A. (1983). Credible and HPD intervals of the parameter and reliability of Rayleigh distribution. IEEE Trans. Reliability, 32: 217-220.

Sloan, J.A. and Sinha, S.K. (1991). Bayesian predictive intervals for a mixture of Exponential failure-time distributions with censored samples. Statistics and probability letters, 11(6): 537-545.

Smith, P.J. (2002). Analysis of Failure and Survival Data. Chapman & Hall/CRC. New York.

Soegiarso, R.W. (1992). A covariate model in finite mixture survival distributions. A PhD Thesis. Department of Epidemiology and Biostatistics, Case Western Reserve University.

Solow, A.R. (1994). On the Bayesian Estimation of the Number of Species in a community. Ecology, 75(7): 2139-2142.

Sultan, K.S., Ismail, M.A. and Al-Moisheer, A.S. (2006). Mixture of two inverse Weibull distributions: properties and estimation. Computational Statistics and Data Analysis, 51(11). 5377-5387.

Tadikamalla, P.R. (1980). A look at the Burr and Related Distributions. International Statistical Review, 48, 337-344.

Teicher, H. (1960). On mixture of distributions. Annals of Mathematical Statistics, 31: 55-73.

Teicher, H. (1961). Identifiability of mixtures. Annals of Mathematical Statistics, 32: 244-248.

Teicher, H. (1963). Identifiability of finite mixtures. Annals of Mathematical Statistics, 34: 1265-1269.

Teicher, H. (1967). Identifiability of mixtures of product measures. Annals of Mathematical Statistics, 38: 1300-1302.

Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985). Statistical Analysis of Finite Mixture Distributions. John Wiley and Sons Ltd. Great Britian.

Wang, J. and Li, Y. (2005). Estimators for survival function when censoring times are known. Communications in Statistics-Theory and Methods, 34: 449-459.

Wiper, M., Insua, D.R. and Ruggeri, F. (2001). Mixtures of Gamma distributions with applications. Journal of Computaional and Graphical Statistics, 10(3): 440-454.

Yakowitz, S.J. and Spragins, J.D. (1968). On the identifiability of finite mixtures. Annals of Mathematical Statistics, 39: 209-214.

Zheng, J and Frey, H.C. (2004). Quantification of Variability and Uncertainty using mixture distributions: Evaluation of Sample size, mixing weights, and separation between components. Risk Analysis, 24(3): 553-571.