Applied Mathematical Modelling 39 (2015) 515–530

Contents lists available at ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

Mixture of the inverse Rayleigh distribution: Properties and estimation in a Bayesian framework

Sajid Ali

Department of Decision Sciences, Bocconi University, Milan, Italy

ARTICLE INFO

Article history:
Received 8 May 2013
Received in revised form 19 March 2014
Accepted 23 May 2014
Available online 20 June 2014

Keywords:
Bayesian analysis
Censored data
Inverse Rayleigh distribution
Loss functions
Mixture model
Posterior risk

ABSTRACT

An engineering process is an output from a set of combined processes which may be homogeneous or heterogeneous. To study the lifetime of such processes, we need a model which can accommodate the nature of such processes. Single probability models are not capable of capturing the heterogeneity of nature. However, mixture models of some suitable lifetime distributions, have the potential to highlight such interesting feature. Due to time and cost constraint, in the most lifetime testing experiments, censoring is an unavoidable feature of most lifetime data sets. This article deals with the modeling of the heterogeneity existing in the lifetime processes using the mixture of the inverse Rayleigh distribution, and the spotlight is the Bayesian inference of the mixture model using non-informative (the Jeffreys and the uniform) and informative (gamma) priors. We are considering this particular distribution due to two reasons; the first one is due to its skewed behavior, i.e. in engineering processes, an engineer suspects that high failure rate in the beginning, but after continuous inspection, the failure goes down and the second reason is due to its vast application in many applied fields. A Gibbs sampling algorithm based on adaptive rejection sampling is designed for the posterior computation. A detailed simulation study is carried out to investigate the performance of the estimators based on different prior distributions. The posterior risks are evaluated under the squared error, the weighted, the quadratic, the entropy, the modified squared error and the precautionary loss functions. Posterior risks of the Bayes estimates are compared to explore the effect of prior information and loss functions.

1. Introduction

Mixture distributions occur in many contexts, especially when a statistical population contains two or more sub-populations. Mixture densities express complex situations in terms of simpler ones, and are used to provide good models for certain data sets. These densities can be used to model a statistical population with subpopulations, where the mixture components are the densities of the subpopulations and the weights represent the proportions of each subpopulation in the overall population. These densities may also be used to model experimental error or contamination, where we assume that most of the samples measure the desired phenomenon. Sultan et al. [1] investigated the properties of a two-component mixture of inverse Weibull distributions using a classical approach and the identifiably property of the mixture model was also discussed. Recently, Kazmi et al. [2] introduced a mixture of the generalized class of distributions where the

E-mail address: sajidali.qau@hotmail.com

http://dx.doi.org/10.1016/j.apm.2014.05.039

Maxwell distribution was a particular example of their defined class using censored data, and they studied different properties of the developed mixture under various loss functions.

The Rayleigh distribution is often used in physics-related fields to model processes such as sound and light radiation, wave heights, and wind speed, as well as in communication theory to describe the hourly median and instantaneous peak power of received radio signals. It has been used to model the frequency of different wind speeds over a year at wind turbine sites and the daily average wind speed. In probability and statistics, the Rayleigh distribution is used to study the wind speed, while it is used as a lifetime model in reliability theory. This distribution also plays an important role in land mobile radio because it can accurately describe the instantaneous amplitude and power of a multipath fading signal [3]. This distribution is an important distribution in statistics and operations research. It is applied in areas such as health, agriculture, biology, and other sciences. Different studies have used the inverse Rayleigh distribution for various purposes. For example, Soliman and Al-Aboud [4] used Bayesian and classical techniques for parameter estimation based on a set of upper record values from the Rayleigh distribution. Bayesian estimators have been developed under symmetric (squared error) and asymmetric (LINEX and general entropy (GE)) loss functions. Howlader et al. [5] used a Bayesian approach to predict the bounds for Rayleigh and inverse Rayleigh lifetime models. Aslam and Jun [6] designed an acceptance sampling plan from a truncated life test when the lifetime of an item followed either an inverse Rayleigh or a log–logistic distribution, where multiple items in a group could be tested simultaneously by a tester. Soliman et al. [7] discussed the problems of Bayesian and non-Bayesian estimation of an unknown parameter for an inverse Rayleigh distribution based on lower record values. They obtained a maximum likelihood (ML) estimator of the unknown parameter and Bayesian analysis was addressed using squared error and zero-one loss functions. The informative prior used to derive these estimators and the predictive intervals were also addressed with a real life dataset. Later, Rosaiah and Kantam [8] discussed an acceptance sampling plan based on an inverse Rayleigh model when the life test was truncated at a pre-assigned time. In the support of their results, they obtained and discussed various properties, such as confidence levels, the ratio of the fixed experimental time to the specified mean life, the minimum sample size, producer risk, and operating characteristics curves. Dey [9] obtained Bayesian estimates of an inverse Rayleigh distribution using squared error and LINEX loss functions.

Several types of data are encountered in everyday life, including simple data, grouped data, truncated data, censored data, and progressively censored data. In the present study, random observations taken from this population are assumed to be characterized by one of two distinct unknown members of the inverse Rayleigh distribution. Right censoring is applied and the observations greater than the fixed cut-off censor value, T, are taken as censored observations. The problem of estimating unknown parameters in statistical distributions, especially in the mixture distributions used to study certain phenomenon, is an important problem that is encountered constantly in applied statistics. In the present study, we consider the estimation of unknown parameters in a mixture of inverse Rayleigh distribution. The aim of this study is to specify a record by introducing a mixture of inverse Rayleigh distributions when data are censored using a Bayesian approach. A mixture of this type has not been reported previously, to the best of our knowledge.

The remainder of this study is organized as follows: The mixture model and the statistical properties of its different shapes, hazard rate of mixture, and ML functions are derived in Section 2. The posterior distributions that assume informative and noninformative priors are given in Section 3. For Bayesian computation, we need to specify the loss function, thus some loss functions used to derive Bayesian estimators and their respective posterior risks are discussed in Section 4. Limiting expressions are also given in the same section. A detailed simulation study based on Gibbs sampling is presented in Section 5. In Section 6, a real-life mixture of inverse Rayleigh distribution is considered. Finally, the last section provides some concluding remarks and proposals for future work.

2. Mixture model and its properties

A finite mixture distribution with k-component densities of specified parametric form and unknown mixing weights (p) is defined as:

$$f(x) = \sum_{i=1}^{k} p_i f_i(x); \quad 0 < p_i < 1, \quad \sum_{i=1}^{k} p_i = 1. \tag{1}$$

The following inverse Raleigh distribution is assumed for k components of the mixture:

$$f_i(x|\theta_i) = \frac{2\theta_i}{x^3} \exp\left(\frac{-\theta_i}{x^2}\right), \quad i = 1, 2, \ldots k. \tag{2}$$

Thus, the mixture model (1) takes the form:

$$f(x) = \sum_{i=1}^{k} p_1 \left\{ \frac{2\theta_i}{x^3} \exp\left(\frac{-\theta_i}{x^2}\right) \right\}, \quad 0 < p_i < 1, \quad \sum_{i=1}^{k} p_i = 1.$$

The corresponding mixture distribution function is:

$$F(x) = pF_1(x) + qF_2(x) = \sum_{i=1}^{k} p_i \left\{ \exp\left(\frac{-\theta_i}{x^2}\right) \right\}, \quad 0 < p_i < 1, \quad \sum_{i=1}^{k} p_i = 1.$$

Graphical representations of different selections of parameter's values for the mixture model are shown in Fig. 1. The mode and median of the two-component mixture of the inverse Rayleigh model are obtained by solving the following nonlinear equations with respect to x

$$2\theta_1^2 p \exp\left(-\frac{\theta_1}{x^2}\right) - 3x^2\theta_1 p \exp\left(-\frac{\theta_1}{x^2}\right) - 3x^2\theta_2(1-p)\exp\left(-\frac{\theta_2}{x^2}\right) + 2\theta_2^2(1-p)\exp\left(-\frac{\theta_2}{x^2}\right) = 0$$

and $\sum_{i=1}^{2} p_i\{\exp\left(\frac{-\theta_i}{x^2}\right)\} = 0.5$, where $p_2 = 1 - p_1$.

The parameter values for the two-component mixture were selected to demonstrate unimodal and bimodal cases. Table 1 shows clearly that the mode and median exhibit decreasing trends as the proportions of the mixed components increase.

2.1. Reliability and failure rate functions

The survival function (reliability function) of the considered mixture model is given as $R(x) = p\{1 - \exp\left(-\frac{\theta_1}{x^2}\right)\} + (1-p)\{1 - \exp\left(-\frac{\theta_2}{x^2}\right)\}$. The hazard function is given as:

$h(x) = \frac{\frac{2\theta_2}{x^3}\exp\left(\frac{-\theta_2}{x^2}\right) + \frac{2\theta_2}{x^3}\exp\left(\frac{-\theta_2}{x^2}\right)}{p\left\{1-\exp\left(-\frac{\theta_1}{x^2}\right)\right\} + (1-p)\left\{1-\exp\left(-\frac{\theta_2}{x^2}\right)\right\}}$, which can be written as (Sultan et al. [1] and references cited therein) $h(x) = h_1(x)r(x) + h_2(x)(1 - r(x))$ and it has the following form of derivative $h'(x) = h_1'(x)r(x) + h_2'(x)(1 - r(x)) - r'(x)(h_1(x) - h_2(x))^2$ where $i = 1,2$; $h_i(x) = \frac{f_i(x)}{R_i(x)}$, $r(x) = \left(1 + \frac{qR_2(x)}{pR_1(x)}\right)^{-1}$, and $R_i(x) = 1 - \exp\left(-\frac{\theta_i}{x^2}\right)$.

Lemma 1. *$\lim_{x\to 0} h(x) = 0$ and $\lim_{x\to\infty} h(x) = 0$*

Proof. Using a power series expansion, we can express $h_i(x)$ for $i = 1, 2$; which is given as $h_i(x) = \frac{2\theta_i}{x^3\left(\exp\left(\frac{\theta_i}{x^2}\right) - 1\right)} = \frac{2}{x + \frac{\theta_i}{2x} + \frac{\theta_i^2}{6x^3} + \dots}$.

Hence, $\lim_{x\to 0} h(x) = 0$ and $\lim_{x\to\infty} h(x) = 0$. Moreover, $\lim_{x\to 0} r(x) = p$. Note that $\frac{qR_2(x)}{pR_1(x)} \geqslant 0$, hence $\lim_{x\to\infty} \frac{qR_2(x)}{pR_1(x)} \neq -1$, thus it follows that $|r(x)| < \infty$.

Interpretation of the failure rate curves: Let us assume that $x_1 = \min(x_1^*, x_2^*)$ and $x_2 = \max(x_1^*, x_2^*)$, where for $i = 1, 2, x_i^*$ represents the mode of the density function $f_i(x)$. From $h_i(x) = \frac{f_i(x)}{R_i(x)}$, we can see that both $f_i(x)$ in the numerator of $h_i(x)$ increases on $(0, x_1)$, where the denominator decreases on the same interval. Therefore, $h(x)$ increases on $(0, x_1)$. In addition, at $x \to \infty, h(x) \to 0$. The following two cases occur in the interval (x_1, ∞).

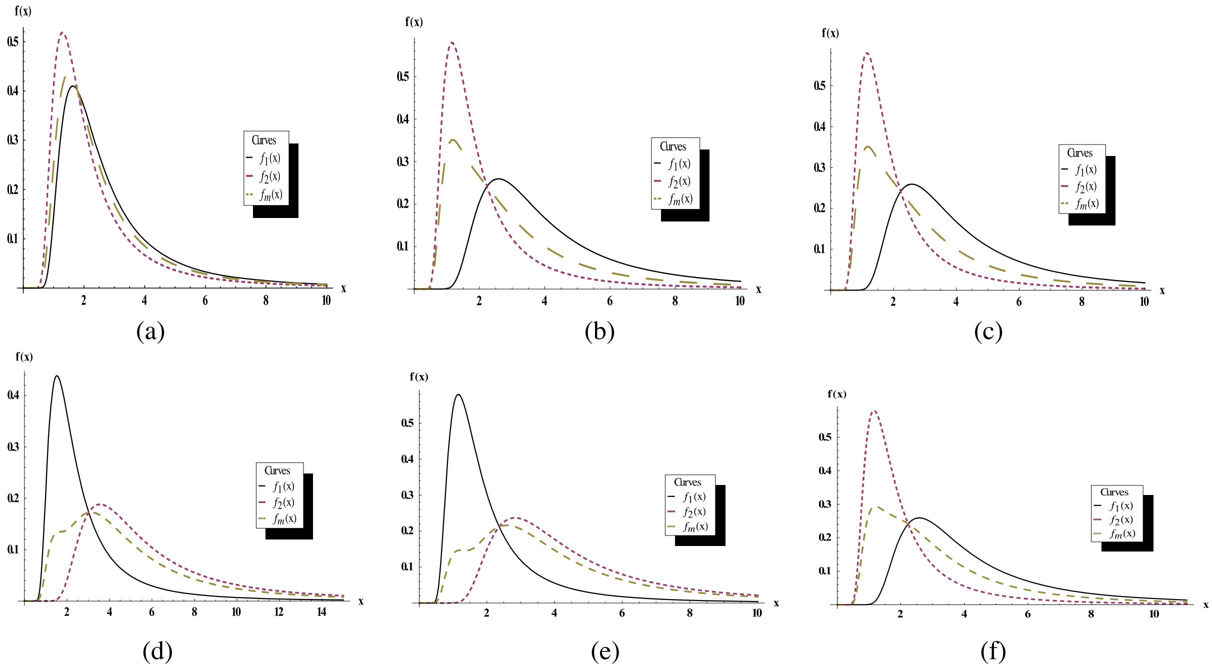

Fig. 1. Density function components and their mixtures: (a) (0.4, 3, 4), (b) (0.6, 4, 2.5), (c) (0.4, 10, 2), (d) (0.3, 3.5, 19), (e) (0.25, 1, 12), (f) (0.5, 10, 2).

Table 1
Modes and medians for the two-component mixture of the inverse Rayleigh distribution.

(p,θ_1,θ_2)	Mode	Median	(p,θ_1,θ_2)	Mode	Median	(p,θ_1,θ_2)	Mode	Median
(0, 3, 4)	1.63299	2.40224	**(0.6, 3, 4)**	1.49725	2.20684	**(0.4, 10, 2)**	1.17845	2.42475
(0.1, 3, 4)	1.61056	2.36947	**(0.7, 3, 4)**	1.47540	2.17481	**(0.3, 3.5, 19)**	1.69921, 1.82461	4.30867
(0.2, 3, 4)	1.58787	2.33673	**(0.8, 3, 4)**	1.45419	2.14309	**(0.25, 1, 12)**	2.67616, 2.82194	3.42954
(0.3, 3, 4)	1.56505	2.30407	**(0.9, 3, 4)**	1.43375	2.11155	**(0.5, 10, 2)**	1.19628	2.66691
(0.4, 3, 4)	1.54224	2.27151	**(1.0, 3, 4)**	1.41421	2.08041	**(0.35, 10, 2)**	1.17296	2.30763
(0.5, 3, 4)	1.51959	2.23909	**(0.6, 4, 2.5)**	1.48868	2.19614	**(0.3, 1, 2)**	1.04082	1.54461

- *Unimodal case:* Suppose that x^* is the maximum point of the failure rate mixture. When the difference Δ between $h_1(x)$ and $h_2(x)$ on the interval (x_1,x^*) is sufficiently small that the first two terms of $h'(x)$ dominate the third term, then $h'(x) > 0$ on (x_1, x^*). In summary, the failure rate of the mixture model increases on $(0,x^*)$ and decreases on (x^*,∞), as $x \to \infty, h(x) \to 0$ (see Fig. 2 (a-b)).
- *Bimodal case:* The failure rate of the mixture model increases on $(0,x^*)$, decreases on (x^*, x^{***}), increases on (x^{***},x^{**}) and decreases again on (x^{**},∞), and $x \to \infty, h(x) \to 0$ (see Fig. 2 (c-d)).

Identifiability: Using the approach adopted by Sultan et al. [1] based on the idea of [10], let ϕ be a transform associated with each $F_i \in \Phi$, which has the domain defined by D_{ϕ_i} with the linear map $M{:}F_i \to \phi_i$. If there exists a total ordering ($\leqslant$) of Φ such that:

- $F_1 \leqslant F_2$, $(F_i \in \Phi)$ implies $D_{\phi_1} \subseteq D_{\phi_2}$;
- For each $F_i \in \Phi$, $(i = 1,2)$, there exists some $s_1 \in D_{\phi_1}, \phi_1(s) \neq 0$ such that $\lim_{s \to s_1} \phi_2(s)/\phi_1(s) = 0$ for $F_1 < F_2$,

then the class Λ of all finite mixture distributions is identifiable relative to Φ.
Using Chandra's approach, we can prove the following proposition. □

Proposition. *The class of all finite mixture distributions is identifiable relative to the inverse Rayleigh distribution.*

Proof. It is immediate because the distribution that is considered belongs to one parameter family. □

2.2. Likelihood function

Assume that '*n*' units from the defined mixture model are used in a lifetime testing experiment with a fixed test termination time *T*. After the test has been conducted, it is observed that of *n* units, *r* units failed up to the test termination

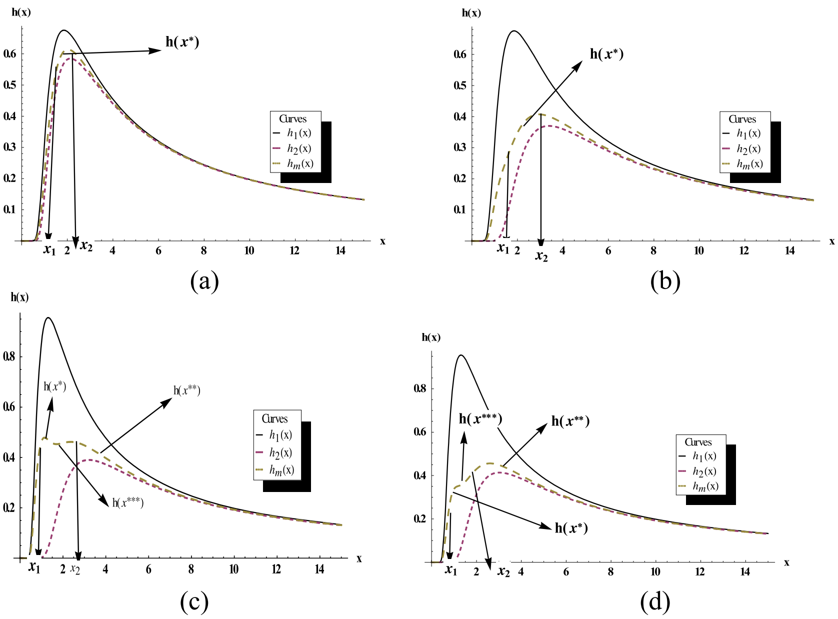

Fig. 2. Hazard function components and their mixtures with parameters: (a) (0.4,3,4), (b) (0.4,3,10), (c) (0.6,1.5,9), (d) (0.45,1.5,8).

time T whereas the remaining $n-r$ units are still functioning. As described by Mendenhall and Hader [11], in many real-life situations, only the failed items can be recognized easily as members of subpopulation 1 or subpopulation 2, etc. For example, an engineer may categorize a failed electronic object as a member of the first or the second subpopulation based on the reasons for its failure. Thus, depending on the cause of the failure, r_i failures may be in the ith sub-population. Obviously, the remaining $n-r$ censored objects provide no information about the subpopulation to which they belong and $r=\sum_{i=1}^{k} r_i$ is the number of uncensored observations. The information about the label, i.e., the cause of failure, can be obtained only after a failure has occurred. Let us define x_{ij} as the failure time of the jth unit from the ith subpopulation, where $j = 1,2,3,\ldots,r_i$, $i = 1,2$, $0 < x_{1j}, x_{2j} \leqslant T$. Hence, the likelihood function has the following form:

$$L(\boldsymbol{\theta},\mathbf{p},\mathbf{x}) \propto \left\{\prod_{j=1}^{r_1} p_1 f_1(x_{1j})\right\}\left\{\prod_{j=1}^{r_2} p_2 f_2(x_{2j})\right\}\cdots\left\{\prod_{j=1}^{r_k} p_k f_k(x_{kj})\right\}\{S(T)\}^{n-r},$$

where $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_k)$; $\mathbf{p}=(p_1,\ldots,p_k)$; $\mathbf{x}=(x_{11},x_{12}\ldots,x_{1r_1},x_{21},x_{22},\ldots,x_{2r_2},\ldots,x_{k1},x_{k2},\ldots,x_{kr_k})$ are the observed failure times for the non-censored observations. After simplification, the likelihood function becomes:

$$L(\boldsymbol{\theta},\mathbf{p},\mathbf{x}) \propto \sum_{k_1,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,\ldots,k_k}\prod_{l=1}^{k} p_l^{r_l+k_l}\theta_l^{r_l}\exp\left(-\theta_l\left\{\sum_{i=1}^{n_k}\frac{1}{x_{lj}^2}+\frac{k_l}{T^2}\right\}\right).$$

where H_{n-r}^k denotes the number of all the k-array sequences $(k_1,\ldots,k_k)$ of non-negative integers with $\sum_{l=1}^{k} k_i = n-r$ and $\binom{n-r}{k_1,\ldots,k_k}=\frac{(n-r)!}{k_1!,\ldots,k_k!}$ in the expansion of the multinomial $\{S(T)\}^{n-r}$, as discussed in Chuan-Chong and Mhee-Meng [12]. For a two-component mixture, the likelihood function can be written as:

$$L(\theta_1,\theta_2,p) \propto \sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k+m}q^{r_2+m}\theta_1^{r_1}\exp\left(-\theta_1\left\{\sum_{i=1}^{r_1}\frac{1}{x_{1j}^2}+\frac{k-m}{T^2}\right\}\right)\theta_2^{r_2}\exp\left(-\theta_2\left\{\sum_{i=1}^{r_2}\frac{1}{x_{2j}^2}+\frac{m}{T^2}\right\}\right). \quad (3)$$

Taking $\mathrm{A}=\left\{\sum_{j=1}^{r_1}\frac{1}{x_{1j}^2}+\frac{k-m}{T^2}\right\}$ and $\mathrm{B}=\left\{\sum_{j=1}^{r_2}\frac{1}{x_{2j}^2}+\frac{m}{T^2}\right\}$, Eq. (3) can be simplified to:

$$L(\theta_1,\theta_2,p) \propto \sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k+m}q^{r_2+m}\theta_1^{r_1}\exp(-\theta_1\mathrm{A})\theta_2^{r_2}\exp(-\theta_2\mathrm{B}).$$

2.3. ML estimators and variances

This method is very important for estimation in classical statistics and it is also used widely for estimating the parameters in mixture models. The ML estimates are obtained by differentiating the log of the likelihood with respect to the unknowns θ_1,θ_2,p and then solving the resulting nonlinear equations, as given in Appendix A. The equations reported in Appendix A require some iterative methods, such as Newton Raphson. Kazmi et al. [13] also showed that the ML equations of mixture of Maxwell distributions are not in closed form. The variances of the ML estimates are on the main diagonal of the inverted information matrix, which is the expectation of the negative Hessian matrix. However, we focus on the Bayesian analysis so we do not discuss ML estimation further. However, it will be interesting to study different properties of ML estimators for the proposed mixture.

3. Posterior distribution using the gamma prior (GP, a conjugate prior)

Our distribution belongs to the skewed family and it is also a member of the natural exponential family, thus the natural conjugate prior for the inverse Rayleigh distribution is a gamma distribution. The GP for the unknown parameter θ_i is defined as $\theta_i \sim Gamma(\mu_i,\delta_i)$ and $p_l \sim \mathrm{Dir}(a_l)$, where Dir stands for Dirichlet distribution. The joint prior distribution is $g(\boldsymbol{\theta},\mathbf{p}) \propto \prod_{l=1}^{k}\theta_l^{\mu_l-1}e^{-\theta_l\delta_l}p_l^{a_l-1}$ and the joint posterior distribution is:

$$g(\boldsymbol{\theta},\mathbf{p}|\mathbf{x}) \propto \sum_{k_1,\ldots,k_k}^{H_{n-r}^k}\binom{n-r}{k_1,\ldots,k_k}\prod_{l=1}^{k} p_l^{r_l+a_l+k_l-1}\theta_l^{r_l+\mu_l-1}\exp\left(-\theta_l\left\{\sum_{i=1}^{n_k}\frac{1}{x_{lj}^2}+\frac{k_l}{T^2}+\delta_l\right\}\right).$$

For a two-component mixture, the joint posterior distribution of θ_1,θ_2,p is

$$g(\theta_1,\theta_2,p|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k+m+a-1}q^{r_2+m+b-1}(\theta_1)^{r_1+\mu_1-1}e^{-\theta_1\{A+\delta_1\}}(\theta_2)^{r_2+\mu_2-1}e^{-\theta_2\{B+\delta_2\}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a,r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}. \quad (4)$$

The marginal posterior distribution of θ_1 is:

$$g(\theta_1|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)(\theta_1)^{r_1+\mu_1-1}e^{-\theta_1\{A+\delta_1\}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}} \quad (5)$$

The marginal posterior distribution of θ_2 is:

$$g(\theta_2|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)(\theta_2)^{r_2+\mu_2-1}e^{-\theta_2\{B_k+\delta_2\}}\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}. \quad (6)$$

Finally, for 'p'

$$g(p|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k-m+a-1}q^{r_2+m+b-1}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}\frac{\Gamma(r_2+\mu_2)}{(A+\delta_1)^{r_1+\mu_1}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}} \quad (7)$$

The predictive distribution contains information about independent future random observations given the known observations. Bansal [14] provided a detailed discussion of the predictive distribution. For details of the posterior predictive distribution and the credible interval derivation based on GP, see Appendix B.

3.1. Bayesian estimation of the mixture model assuming noninformative priors

Among the techniques that have been proposed for determining uninformative priors, Jeffreys [15] suggested a method based on the square-root of the Fisher information, which is the most widely used method. Box and Taio [16] defined a noninformative prior as a prior that provides little information related to the experiment. Later, Geisser [17] also proposed techniques for noninformative priors. Bernardo and Smith [18] used a similar definition, where they stated that noninformative priors have a minimal effect on the final inference relative to the data and they regarded noninformative priors as a mathematical tool. Bernardo [19] argued that a noninformative prior should be regarded as a reference prior, i.e., a prior that is convenient for use as a standard when analyzing statistical data. The most common examples of noninformative priors are uniform priors and the Jeffreys priors. Both priors are used only when no formal prior information is available.

3.1.1. Posterior distribution using uniform prior

The uniform prior for the unknown parameter θ_i can easily be written as $\theta_i \sim$ Uniform $(0,\infty)$, $i = 1,2$. We assume a priori that (θ_i,p), $i = 1,2$ are independent and we also assume that $p \sim$ Uniform (0,1). Thus, the joint prior distribution of θ_1, θ_2, p is $g(\theta_1,\theta_2, p) \propto 1$. By combining the likelihood function given in (3) and uniform prior information, we obtain the joint posterior distribution of θ_1, θ_2, p as:

$$g(\theta_1,\theta_2,p|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k+m}q^{r_2+m}(\theta_1)^{r_1}e^{-\theta_1 A}(\theta_2)^{r_2}e^{-\theta_2 B}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+1)}{(A)^{r_1+1}}\frac{\Gamma(r_2+1)}{(B)^{r_2+1}}}, \quad (8)$$

where $A = \left\{\sum_{j=1}^{r_1}\frac{1}{x_{1j}^2}+\frac{k-m}{T^2}\right\}$, $B = \left\{\sum_{j=1}^{r_2}\frac{1}{x_{2j}^2}+\frac{m}{T^2}\right\}$. The marginal posterior distribution of θ_1 is simply the probability distribution of θ_1 that ignores other irrelevant information about θ_2 and p, which is calculated by integrating the joint probability distribution with respect to other parameters as

$$g(\theta_1|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_2+1)}{(B)^{r_2+1}}(\theta_1)^{r_1}e^{-\theta_1 A}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_2+1)}{(A)^{r_1+1}}\frac{\Gamma(r_2+1)}{(B)^{r_2+1}}}. \quad (9)$$

Similarly, the marginal posterior distribution of θ_2 is obtained as

$$g(\theta_2|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+1)}{(A)^{r_1+1}}(\theta_2)^{r_2}e^{-\theta_2 B}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+1)}{(A)^{r_1+1}}\frac{\Gamma(r_2+1)}{(B)^{r_2+1}}}. \tag{10}$$

For the mixing proportion 'p', we have:

$$g(p|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k-m}q^{r_2+m}\frac{\Gamma(r_1+1)}{(A)^{r_1+1}}\frac{\Gamma(r_2+1)}{(B)^{r_2+1}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1+1)}{(A)^{r_1+1}}\frac{\Gamma(r_2+1)}{(B)^{r_2+1}}}. \tag{11}$$

3.1.2. Posterior distribution using the Jeffreys prior

Jeffreys [15] proposed a formal rule for obtaining a noninformative prior, as follows: if θ is a k-vector valued parameter, $g(\theta) \propto \sqrt{|\det I(\theta)|}$, where $I(\theta)$ is a $k \times k$ Fisher's (information) matrix, in which the (i,j)th element is $-E\left[\frac{\partial^2 \log L(\theta/x)}{\partial\theta_i \partial\theta_j}\right]$; $i,j = 1,2,\ldots,k$. Fisher's information matrix is not related directly to the notation of the lack of information. The relationship is derived from the role of Fisher's matrix in asymptotic theory. The Jeffreys prior is based on Fisher's information matrix, which often leads to a family of improper priors. The Jeffreys prior for the unknown parameter θ_i can easily be calculated as $g(\theta_i) \propto \frac{1}{\theta_i}$. We assume that $(\theta_i, p), i = 1,2$, are a priori independent and we also assume that $p \sim$ Uniform (0,1). Thus, the joint prior distribution of θ_1, θ_2, p is $g(\theta_1, \theta_2, p) \propto \frac{1}{\theta_1\theta_2}$. The joint posterior of θ_1, θ_2, p is as follows:

$$g(\theta_1, \theta_2, p|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k+m}q^{r_2+m}(\theta_1)^{r_1-1}e^{-\theta_1 A}(\theta_2)^{r_2-1}e^{-\theta_2 B}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1)}{(A)^{r_1}}\frac{\Gamma(r_1)}{(B)^{r_2}}}. \tag{12}$$

The marginal posterior distributions for the Jeffreys prior using (8) are:

$$g(\theta_1|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)(\theta_1)^{r_1-1}e^{-\theta_1 A}\frac{\Gamma(r_2)}{(B)^{r_2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1)}{(A)^{r_1}}\frac{\Gamma(r_2)}{(B)^{r_2}}}, \tag{13}$$

$$g(\theta_2|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)(\theta_2)^{r_2-1}e^{-\theta_2 B}\frac{\Gamma(r_1)}{(A)^{r_1}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1)}{(A)^{r_1}}\frac{\Gamma(r_2)}{(B)^{r_2}}} \tag{14}$$

and

$$g(p|\mathbf{x}) = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}p^{r_1+k-m}q^{r_2+m}\frac{\Gamma(r_2)}{(B)^{r_2}}\frac{\Gamma(r_1)}{(A)^{r_1}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+1, r_2+m+1)\frac{\Gamma(r_1)}{(A)^{r_1}}\frac{\Gamma(r_2)}{(B)^{r_2}}}. \tag{15}$$

4. Bayesian estimation under different loss functions

For the evaluation of Bayes estimators and their respective posterior risks, we need to specify a loss function. The selection of a loss function is a difficult task, and its choice is often made for the reasons of mathematical convenience without any particular decision problem of current interest except cost effect. Since in the risk analysis, both the potentiality of an undesired event and its consequences are investigated. This potentiality is usually measured by either a probability or a failure rate. The Bayesian approach is widely applied to estimate this failure rate. When dealing with disastrous consequences, it can be worse to underestimate the potentiality of an event rather than to overestimate it. This is important when the risk level is the basis of a risk-reducing initiative, either by reducing the potentiality or the consequences. An erroneously low estimate of the risk level can lead to the lack of necessary steps to reduce the risk level. Thus, it is unreasonable to use a loss function that allows the estimation of a failure probability of zero. A positive loss at the origin allows the estimation of zero and, in a risk analysis, estimating a zero failure probability simply means that no risk is anticipated [20]. This section presents the derivation of different loss functions for the posterior distributions. Six different loss functions are used to obtain the Bayesian estimators and their respective posterior risks, i.e., the squared error loss function (SELF), weighted squared error loss function (WSELF), quadratic loss function (QLF), entropy loss function (ELF), modified squared error loss function (MSELF), and the precautionary loss function (PLF).

The SELF is the most commonly used loss function and it was proposed to develop the least squares theory. It is defined as $L_1 = L(\theta, d) = (\theta - d)^2$. The WSELF is $L_2 = L(\theta, d) = \frac{(\theta-d)^2}{\theta}$ and the QLF is $L_3 = L(\theta, d) = (1 - d/\theta)^2$. In many practical applications, it appears to be more reasonable to express the loss in terms of the ratio $\widehat{\theta}/\theta$. Thus, the entropy loss allows the explicit estimation of the natural parameter, which is the canonical form of the exponential family (which makes it suitable because the inverse Rayleigh distribution belongs to this family). Calabria and Pulcini [21] defined the generalized ELF as $L_4 = L(\theta, d) = b[(d/\theta)^c - c \log(d/\theta) - 1], c \neq 0$, which is a valid alternative to the modified linex loss function. For a special case of entropy loss function where $c = 1$, we have $L_4 = L(\theta, d) = [d/\theta - \log(d/\theta) - 1]$. Norstrom [20] introduced an alternative asymmetric PLF and also presented a general class of precautionary loss functions as a special case, which is defined as $L_5 = L(\theta, d) = (\theta - d)^2/d$. According to Norstrom [20], this loss function approaches infinitely close to the origin to prevent underestimation, thereby yielding conservative estimators, especially when underestimation may lead to serious results. Thus, using nonsymmetric loss functions, we can handle cases where it is more damaging to miss the target on one side than the other. The MSELF was introduced by Degroot [22] and it is defined as: $L_6 = L(\theta, d) = \left(\frac{\theta-d}{d}\right)^2$. These loss functions have also been studied for different purposes. Recently, Ali et al. [23] used these loss functions for Lindley distribution parameter estimation. Fig. 3 shows graphical presentation of these loss functions.

The mathematical expectation of each parameter with respect to its marginal distribution gives the Bayesian estimator of that parameter (we can find the posterior risk of the respective parameter in a similar manner). Table 2 shows the Bayes estimators and their respective posterior risks under different loss functions.

4.1. Bayes estimators using GP for SELF

The Bayes estimator and the respective posterior variance under different loss functions can be obtained easily using Eqs. (5)–(7). These estimators are simply obtained by multiplying the respective parameter by its marginal density and integrating, i.e., $E(\theta_i|\mathbf{x}) = \int_0^\infty \theta_i g(\theta_i|\mathbf{x}) d\theta_i$, where $g(\theta_i|\mathbf{x})$ is the respective marginal distribution. The following is the Bayesian estimator of θ_1

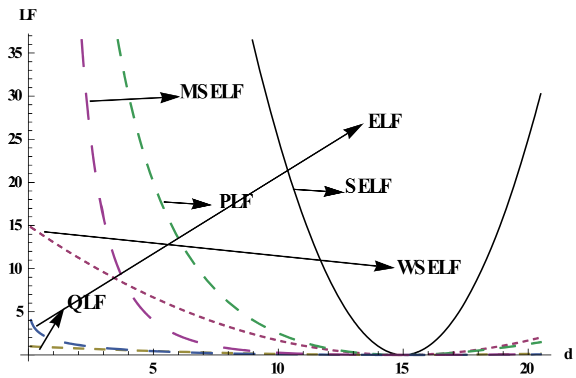

Fig. 3. Effects of the estimates of θ on decision d using $\theta = 15$.

Table 2
Bayesian estimators and posterior risks under different loss functions.

Loss function	Bayesian estimator	Posterior risk
SELF	$E(\theta\|\mathbf{x})$	$E(\theta^2\|\mathbf{x}) - (E(\theta\|\mathbf{x}))^2$
WSELF	$\frac{1}{E(\theta^{-1}\|\mathbf{x})}$	$E(\theta\|\mathbf{x}) - \frac{1}{E(\theta^{-1}\|\mathbf{x})}$
QLF	$\frac{E(\theta^{-1}\|\mathbf{x})}{E(\theta^{-2}\|\mathbf{x})}$	$1 - \frac{[E(\theta^{-1}\|\mathbf{x})]^2}{E(\theta^{-2}\|\mathbf{x})}$
ELF	$(E(\theta^{-1}\|\mathbf{x}))^{-1}$	$E(\log\theta\|\mathbf{x}) + \log(E(\theta^{-1}\|\mathbf{x}))$
MSELF	$\frac{E(\theta^2\|\mathbf{x})}{E(\theta\|\mathbf{x})}$	$\frac{\mathrm{Var}(\theta\|\mathbf{x})}{E(\theta^2\|\mathbf{x})}$
PLF	$\sqrt{E(\theta^2\|\mathbf{x})}$	$2\left(\sqrt{E(\theta^2\|\mathbf{x})} - E(\theta\|\mathbf{x})\right)$

$$d_1^* = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1+1)}{(A+\delta_1)^{r_1+\mu_1+1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}$$

and for θ_2, we have

$$d_2^* = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2+1)}{(B+\delta_2)^{r_2+\mu_2+1}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}$$

Finally, for the mixing proportion parameter, the Bayesian estimator is as follows.

$$d_3^* = \frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a+1, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}$$

4.2. Posterior risk/variances for SELF

To determine the posterior risk/variance, we need to find $E(\theta_i^2|\mathbf{x}) = \int_0^\infty \theta_i^2 g(\theta_i|\mathbf{x})d\theta_i$ with the addition of $E(\theta_i|\mathbf{x})$. The posterior risk under SELF is simply the variance. For θ_1, the risk estimator is

$$V(\theta_1|\mathbf{x}) = \left\{\frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1+2)}{(A+\delta_1)^{r_1+\mu_1+2}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}} - [d_1^*]^2\right\}$$

For θ_2, the risk estimator is

$$V(\theta_2|\mathbf{x}) = \left\{\frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2+2)}{(B+\delta_2)^{r_2+\mu_2+2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}} - [d_2^*]^2\right\}$$

Finally, for p we have:

$$V(p|\mathbf{x}) = \left\{\frac{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a+2, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}}{\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^k\binom{n-r}{k}\binom{k}{m}\beta(r_1+k-m+a, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{(A+\delta_1)^{r_1+\mu_1}}\frac{\Gamma(r_2+\mu_2)}{(B+\delta_2)^{r_2+\mu_2}}} - [d_3^*]^2\right\}$$

For reasons of space, we do not present the expressions for the Bayesian estimators and the respective posterior risks using other loss functions for noninformative and informative priors.

4.3. Limiting expressions for the complete dataset

Suppose that $T \to \infty$, where all of the observations included in our analysis are uncensored, therefore r tends to n, $n_1 + n_2$ tends to n, r_1 tends to unknown n_1, and r_2 tends to unknown n_2. As a result, the amount of information in the sample is increased, thereby reduces the risks of the estimates. Thus, we can easily obtain the limiting expressions for GP and ML. The expressions of the complete Bayesian estimates and the risks for the samples are simplified in Table 3.

Note that in Table 3, $\text{A} = \sum_{j=1}^{n_1}\frac{1}{x_{1j}^2}$ and $\text{B} = \sum_{j=1}^{n_2}\frac{1}{x_{2j}^2}$.

Table 3
Limiting expressions for the BE (GP) and ML estimators as $T \to \infty$.

Parameter	BE (GP)	Risk (GP)	ML	Variance
θ_1	$\frac{n_1+\mu_1}{A+\delta_1}$	$\frac{n_1+\mu_1}{(A+\delta_1)^2}$	$\frac{n_1}{A}$	$\frac{\hat{\theta}_1^2}{n_1}$
θ_2	$\frac{n_2+\mu_2}{B+\delta_2}$	$\frac{n_2+\mu_2}{(B+\delta_2)^2}$	$\frac{n_2}{B}$	$\frac{\hat{\theta}_2^2}{n_2}$
p	$\frac{n_1+a}{n+a+b}$	$\frac{(n_1+a)(n_2+b)}{(n+a+b)^2(n+a+b+1)}$	$\frac{n_1}{n}$	$\frac{n_1(1-p)^2+n_2p^2}{p^2(1-p)^2}$

5. Simulation study

A simulation study was conducted to determine the Bayesian estimates and the respective posterior risks, and to highlight the effects of small and large sample sizes, different censoring rates, loss functions, and prior distributions. Samples of $n = 25, 50, 100, 500$, and 1000 were generated from the two-component mixture of the inverse Rayleigh distribution with the parameters: $T = 5, \theta_1 = 3, \theta_2 = 4$, and $p \in (0.40, 0.60)$. Probabilistic mixing was used to generate the mixture data. To generate observations from the mixture model, 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 inverse Rayleigh distribution with parameter θ_1) and if '$u > p$', the observation was taken randomly from F_2 (the inverse Rayleigh distribution with parameter θ_2). The hyperparameters for the GP were $\mu_1 = 6, \delta_1 = 2, \mu_2 = 8, \delta_2 = 2$. Right censoring was performed using a fixed censoring time T. All observations that exceeded T were treated as censored. Different fixed censoring times T were used to evaluate the effect of the censoring time on the estimates. For each of the combinations of parameters, sample sizes, and censoring rates, we generated 10,000 samples using the statistical package R. In each case, only failures were recognized as members of subpopulation-1 or subpopulation-2 of the mixture. For each of the 10,000 samples, the Bayesian estimates and posterior risks were computed using the package R, and Tables 4–9 show the averages of the 10,000 estimates. A Gibbs sampling algorithm was designed to compute the posterior summary. The joint posterior distribution is

$$g(\boldsymbol{\theta}, \mathbf{p}|\mathbf{x}) \propto \left\{\prod_{j=1}^{r_1} p_1 f_1(x_{1j})\right\}\left\{\prod_{j=1}^{r_2} p_2 f_2(x_{2j})\right\}\{S(T)\}^{n-r}\prod_{l=1}^{2}\theta_l^{\mu_l-1}e^{-\theta_l\delta_l}p_l^{a_l-1}.$$

The conditional distribution for the mixing proportion component can be written as follows:

$g(p|\boldsymbol{\theta}, \mathbf{x}) \propto \left\{p_1^{r_1+a_1-1}p_2^{r_2+a_2-1}\right\}\{S(T)\}^{n-r}$. This conditional does not have a standard form, thus adaptive rejection sampling (ARS) is useful in this situation. However, ARS requires that we prove that it is a log-concave function with respect to the mixing proportion. Taking the second derivative of the logarithmic conditional distribution with respect to "p," we can show that it is strictly negative and hence the conditional distribution is log-concave in the mixing proportion p, so the parameter p can be generated using the ARS procedure.

Remark. As a special case when $r = n$ (complete data case), the conditional distribution of the mixing proportion is a standard beta distribution and random numbers from this distribution can be generated directly using a standard beta random number generator instead of the ARS technique.

Similarly, by ignoring the terms that are not related to θ_1, we can find the conditional density as follows.

$$g(\theta_1|\theta_2, p, \mathbf{x}) \propto (\theta_1)^{r_1+\mu_1-1} e^{-\theta_1\left\{\sum_{j=1}^{r_1}\frac{1}{x_{1j}^2}+\delta_1\right\}+(n-r)\ln S(T)}$$

Table 4
Bayesian estimates and posterior risk using SELF (UP: uniform prior; JP: Jeffreys prior; GP: gamma prior).

Prior	UP			JP			GP		
	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\ \|\mathbf{x})$	$E(p\|\mathbf{x})$
n	$p = 0.6$								
25	3.69887	4.09724	0.692305	3.60181	3.5674	0.695204	2.98707	4.44854	0.59544
	(0.769719)	(2.2489)	(0.009074)	(0.732527)	(1.98766)	(0.008998)	(0.445278)	(1.161750)	(0.007173)
50	3.03982	4.21958	0.588797	2.94333	4.01657	0.58938	2.95376	4.21059	0.591293
	(0.321328)	(0.906406)	(0.005378)	(0.311907)	(0.867389)	(0.005305)	(0.257657)	(0.603972)	(0.004381)
100	3.19252	4.04239	0.599075	3.14179	3.94063	0.599381	3.1128	4.13650	0.596302
	(0.176785)	(0.434158)	(0.002726)	(0.17417)	(0.428438)	(0.002727)	(0.155332)	(0.417463)	(0.002451)
500	3.02118	4.05135	0.599843	3.01144	4.0306	0.599905	3.01124	4.09237	0.599912
	(0.032189)	(0.088993)	(0.000561)	(0.0320928)	(0.0885852)	(0.000560)	(0.031409)	(0.084803)	(0.000547)
1000	3.02118	4.04097	0.600082	3.0115	4.03059	0.600113	3.0114	4.02996	0.600115
	(0.016074)	(0.044407)	(0.000281)	(0.016049)	(0.044305)	(0.000280)	(0.015877)	(0.043338)	(0.000278)
n	$p = 0.4$								
25	3.47587	5.36504	0.390375	3.12999	5.05368	0.38849	3.38765	4.14854	0.365443
	(1.27856)	(1.87148)	(0.010355)	(1.16533)	(1.76963)	(0.010324)	(0.454278)	(1.101705)	(0.009137)
50	3.43305	4.39402	0.399853	3.26188	4.25811	0.399111	3.25379	4.11509	0.391923
	(0.623137)	(0.658448)	(0.005388)	(0.595239)	(0.639389)	(0.005384)	(0.276557)	(0.639702)	(0.004918)
100	3.1008	4.32929	0.397987	3.02202	4.26089	0.397663	3.21182	4.10651	0.396320
	(0.257177)	(0.321842)	(0.002702)	(0.251204)	(0.317022)	(0.002700)	(0.175323)	(0.447613)	(0.002561)
500	3.02576	4.01902	0.400545	3.01026	4.00606	0.400481	3.15124	4.07273	0.399129
	(0.052807)	(0.056708)	(0.000556)	(0.049289)	(0.056535)	(0.000556)	(0.032405)	(0.084913)	(0.000616)
1000	3.01446	4.01621	0.399809	3.00671	4.00973	0.399777	3.08145	4.03986	0.400151
	(0.024661)	(0.028323)	(0.000278)	(0.024607)	(0.044305)	(0.000279)	(0.016807)	(0.043638)	(0.000287)

Table 5
Bayesian estimates and posterior risk using WSELF (UP: uniform prior; JP: Jeffreys prior; GP: gamma prior).

Prior	UP			JP			GP		
	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$
n	$p = 0.6$								
25	3.48972	3.5414	0.677977	3.30577	3.00174	0.681049	2.83755	4.18629	0.582648
	(0.209152)	(0.555842)	(0.014328)	(0.209576)	(0.565668)	(0.014155)	(0.149518)	(0.262224)	(0.012793)
50	2.93382	4.00358	0.579249	2.83704	3.7993	0.579835	2.86636	3.85956	0.58361
	(0.105999)	(0.215998)	(0.009549)	(0.106290)	(0.217271)	(0.009545)	(0.087400)	(0.151028)	(0.007682)
100	3.13707	3.93471	0.594419	3.08628	3.83266	0.594727	3.06285	4.23851	0.592108
	(0.055451)	(0.107679)	(0.004656)	(0.055515)	(0.107973)	(0.004654)	(0.049955)	(0.096484)	(0.004195)
500	3.01053	4.02937	0.598904	3.00078	4.00861	0.598966	3.00081	4.0081	0.598996
	(0.010657)	(0.021078)	(0.000939)	(0.010660)	(0.021991)	(0.000939)	(0.010433)	(0.021057)	(0.000916)
1000	3.01104	4.02997	0.599612	3.00617	4.01959	0.599643	3.00631	4.0192	0.599951
	(0.005329)	(0.010992)	(0.000469)	(0.005330)	(0.010995)	(0.000469)	(0.005273)	(0.010757)	(0.000464)
n	$p = 0.4$								
25	3.10487	5.01313	0.361514	2.75397	4.69993	0.359602	2.98575	4.28619	0.426548
	(0.370996)	(0.351911)	(0.028861)	(0.376024)	(0.353752)	(0.028881)	(0.159411)	(0.264245)	(0.013723)
50	3.25076	4.24361	0.385772	3.07855	4.10735	0.385016	2.98686	4.21965	0.418364
	(0.182281)	(0.150410)	(0.014081)	(0.183323)	(0.150762)	(0.014095)	(0.089409)	(0.151182)	(0.007729)
100	3.01772	4.25483	0.391049	2.93875	4.18636	0.390723	3.03259	4.18354	0.402918
	(0.083079)	(0.074460)	(0.006938)	(0.083271)	(0.074527)	(0.006940)	(0.051956)	(0.106844)	(0.004285)
500	3.00939	4.0049	0.399151	2.99388	3.99194	0.399087	3.02801	4.15018	0.401996
	(0.016372)	(0.014145)	(0.001394)	(0.016379)	(0.014117)	(0.001395)	(0.010443)	(0.021175)	(0.000967)
1000	3.00628	4.00916	0.399111	3.00617	4.01959	0.399079	3.03601	4.10129	0.409953
	(0.008182)	(0.007053)	(0.000696)	(0.008178)	(0.007050)	(0.000696)	(0.005372)	(0.010689)	(0.000478)

Table 6
Bayesian estimates and posterior risk using ELF (UP: uniform prior; JP: Jeffreys prior; GP: gamma prior).

Prior	UP			JP			GP		
	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$
n	$p = 0.6$								
25	3.27936	3.97763	0.662249	3.28782	3.92634	0.665506	3.23892	4.07763	0.652298
	(0.195322)	(0.543212)	(0.014310)	(0.189576)	(0.542686)	(0.014148)	(0.184342)	(0.534221)	(0.014130)
50	3.18751	3.98629	0.569258	3.17342	3.98059	0.569847	3.18719	4.06796	0.569485
	(0.103978)	(0.207854)	(0.009494)	(0.101200)	(0.207712)	(0.009445)	(0.101108)	(0.207654)	(0.009442)
100	3.08154	3.98674	0.589655	3.03068	3.97438	0.589965	3.11584	4.05764	0.589534
	(0.053249)	(0.106481)	(0.004566)	(0.053215)	(0.106173)	(0.004554)	(0.053209)	(0.106099)	(0.004546)
500	2.99987	4.00738	0.597961	2.99012	3.98660	0.598023	3.10978	4.03738	0.597899
	(0.010578)	(0.021061)	(0.000929)	(0.010570)	(0.021054)	(0.000925)	(0.010564)	(0.020986)	(0.000918)
1000	3.00571	4.01898	0.599141	3.00084	4.00860	0.599172	3.07571	4.01975	0.599138
	(0.005279)	(0.010952)	(0.000457)	(0.005280)	(0.010951)	(0.000449)	(0.005269)	(0.010950)	(0.000443)
n	$p = 0.4$								
25	2.73049	4.09259	0.360031	2.87396	4.14182	0.382119	3.07409	4.08256	0.390031
	(0.370869)	(0.351879)	(0.028798)	(0.365968)	(0.348962)	(0.027889)	(0.364879)	(0.347895)	(0.027798)
50	2.87968	4.12959	0.371046	2.89435	3.95593	0.370279	3.05968	4.06529	0.397146
	(0.172819)	(0.150407)	(0.013989)	(0.172189)	(0.150396)	(0.013979)	(0.172119)	(0.150279)	(0.013965)
100	2.93845	4.18024	0.383957	2.85533	4.11171	0.392297	3.03495	4.04644	0.398957
	(0.083065)	(0.074399)	(0.006929)	(0.082897)	(0.074378)	(0.006920)	(0.082765)	(0.074367)	(0.006919)
500	2.99301	3.99079	0.39775	2.07749	3.97782	0.397686	3.01381	4.01979	0.397541
	(0.016289)	(0.014138)	(0.001382)	(0.016194)	(0.014127)	(0.001379)	(0.016181)	(0.014118)	(0.001380)
1000	2.99810	4.00210	0.398411	2.99034	3.99562	0.398379	3.00981	4.00187	0.398741
	(0.008159)	(0.007047)	(0.000685)	(0.008149)	(0.007040)	(0.000686)	(0.008139)	(0.007038)	(0.000675)

The conditional distribution of θ_1 does not have a standard form, but its form suggests that a gamma distribution can be used as a source distribution, i.e., gamma $(r_1 + \mu_1,\phi)$, where ϕ is a parameter that can be obtained by solving the saddle point problem $\max_{\theta_1}\min_{\phi}\mathrm{ADR}(\phi,\theta_1)$, where ADR (ϕ,θ_1) is the acceptance rate, which is defined by the ratio of the conditional distribution (as given above) relative to the distribution source adopted (gamma in this case), and the logarithm of this ratio is given by:

$$\ln \mathrm{ADR}(\phi,\theta_1) = -\theta_1\left\{\sum_{j=1}^{r_1}\frac{1}{x_{1j}^2} + \delta_1 - \phi\right\} + (n-r)\ln S(T) - (r_1+\mu_1)\ln\phi$$

Table 7
Bayesian estimates and posterior risk using PLF (UP: uniform prior; JP: Jeffreys prior; GP: gamma prior).

Prior	UP			JP			GP		
	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$
n	$p = 0.6$								
25	3.801494	4.363058	0.698828	3.702102	3.835884	0.701646	3.060697	4.577254	0.601433
	(0.205248)	(0.531636)	(0.013045)	(0.200585)	(0.536967)	(0.012884)	(0.147254)	(0.257429)	(0.011986)
50	3.092221	4.325652	0.593346	2.995847	4.123133	0.593863	2.997058	4.085193	0.594986
	(0.104803)	(0.212143)	(0.009099)	(0.105034)	(0.213125)	(0.008966)	(0.086595)	(0.149207)	(0.007385)
100	3.220088	4.095739	0.601347	3.169387	3.994622	0.601652	3.137651	4.382886	0.598354
	(0.055137)	(0.106697)	(0.004543)	(0.055194)	(0.107983)	(0.004541)	(0.049703)	(0.095772)	(0.004104)
500	3.026503	4.062318	0.60031	3.016764	4.041574	0.600372	3.016451	4.039879	0.600368
	(0.010645)	(0.021936)	(0.000934)	(0.010648)	(0.021948)	(0.000934)	(0.010422)	(0.021019)	(0.000912)
1000	3.023839	4.046461	0.600316	3.014164	4.036082	0.600347	3.014035	4.035333	0.600346
	(0.005318)	(0.010982)	(0.000469)	(0.005327)	(0.010985)	(0.000469)	(0.005270)	(0.010747)	(0.000463)
n	$p = 0.4$								
25	3.655165	5.536708	0.40342	3.310916	5.225831	0.401557	3.166095	4.672574	0.401334
	(0.358590)	(0.343336)	(0.026091)	(0.361853)	(0.344302)	(0.026135)	(0.172454)	(0.274592)	(0.014968)
50	3.522637	4.468317	0.406535	3.35188	4.332538	0.4058	3.147506	4.385913	0.409968
	(0.179173)	(0.148595)	(0.013363)	(0.180000)	(0.148857)	(0.013378)	(0.126955)	(0.249270)	(0.011358)
100	3.141996	4.366302	0.401367	3.0633	4.29793	0.401044	3.137564	4.288816	0.398435
	(0.082392)	(0.074024)	(0.006759)	(0.082561)	(0.074081)	(0.006762)	(0.097430)	(0.205727)	(0.009140)
500	3.034474	4.026069	0.401239	3.018436	4.01311	0.401175	3.064751	4.209789	0.400386
	(0.017427)	(0.014097)	(0.001387)	(0.016352)	(0.014100)	(0.001387)	(0.070425)	(0.041109)	(0.001102)
1000	3.018548	4.019735	0.400157	3.010799	4.015251	0.400125	3.030345	4.135334	0.400364
	(0.008175)	(0.007049)	(0.000696)	(0.008178)	(0.011042)	(0.000696)	(0.008261)	(0.020774)	(0.000936)

Table 8
Bayesian estimates and posterior risk using QSELF (UP: uniform prior; JP: Jeffreys prior; GP: gamma prior).

Prior	UP			JP			GP		
	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$
n	$p = 0.6$								
25	3.27934	3.90774	0.662234	3.27782	3.90639	0.665511	3.18392	4.07673	0.642928
	(0.195222)	(0.543209)	(0.014130)	(0.189566)	(0.542678)	(0.014128)	(0.184332)	(0.534120)	(0.014130)
50	3.19875	3.97628	0.569238	3.19342	3.98050	0.569827	3.17191	4.06966	0.639845
	(0.103978)	(0.207754)	(0.009490)	(0.101200)	(0.207710)	(0.009420)	(0.101018)	(0.207653)	(0.009440)
100	3.16154	3.98762	0.589675	3.13068	3.97483	0.589865	3.15184	4.05746	0.629354
	(0.053249)	(0.106294)	(0.004560)	(0.053213)	(0.106170)	(0.004545)	(0.053204)	(0.106069)	(0.004543)
500	3.09987	4.00734	0.597861	3.10992	3.98657	0.598230	3.13987	4.04783	0.617989
	(0.010574)	(0.021060)	(0.000920)	(0.010569)	(0.021050)	(0.000923)	(0.010563)	(0.020986)	(0.000916)
1000	3.07571	4.01889	0.599101	3.08084	4.00852	0.599100	3.08571	4.02957	0.609183
	(0.005271)	(0.010950)	(0.000450)	(0.005263)	(0.010950)	(0.000445)	(0.005259)	(0.010941)	(0.000440)
n	$p = 0.4$								
25	3.07349	4.08295	0.350301	3.07396	4.10182	0.342191	3.05490	4.06165	0.390021
	(0.370861)	(0.351869)	(0.028788)	(0.365965)	(0.348892)	(0.027789)	(0.359879)	(0.347885)	(0.027788)
50	3.05968	4.07959	0.340746	3.04935	4.09553	0.330729	3.04868	4.05629	0.397132
	(0.172814)	(0.150400)	(0.013979)	(0.172180)	(0.150395)	(0.013971)	(0.172109)	(0.150269)	(0.013957)
100	3.03845	4.06024	0.333057	3.05353	4.08171	0.329297	3.03459	4.04638	0.398867
	(0.083062)	(0.074389)	(0.006911)	(0.082890)	(0.074362)	(0.006910)	(0.082756)	(0.074360)	(0.006914)
500	3.01993	4.04979	0.327175	3.04749	4.07782	0.317866	3.02318	4.03997	0.397540
	(0.016280)	(0.014136)	(0.001380)	(0.016184)	(0.014120)	(0.001369)	(0.016171)	(0.014108)	(0.001370)
1000	3.01998	4.03210	0.308141	3.02934	4.05629	0.308397	3.01981	4.01087	0.398701
	(0.008155)	(0.007045)	(0.000680)	(0.008146)	(0.007038)	(0.000676)	(0.008129)	(0.007028)	(0.000655)

Differentiating ln ADR(ϕ,θ_1) with respect to ϕ,θ_1 and by equating the results to zero, the optimal value of the parameter ϕ is $\frac{(r_1+\mu_1)}{\theta_1}$, where θ_1 can be obtained by solving the nonlinear equation:

$$-\left\{\sum_{j=1}^{r_1}\frac{1}{x_{1j}^2}+\delta_1\right\}+\frac{(r_1+\mu_1)}{\theta_1}+(n-r)\frac{p\exp\left(\frac{-\theta_1}{T^2}\right)}{S(T)T^2}=0$$

Note that for the complete sample, the conditional distribution is a gamma $\left(r_1+\mu_1, \sum_{j=1}^{n_1}\frac{1}{x_{1j}^2}+\delta_1\right)$. Using the same procedure, we can obtain the conditional distribution of θ_2. Hence, the final Gibbs sampling algorithm can be written as follows.

Table 9
Bayesian estimates and posterior risk using MSELF (UP: uniform prior; JP: Jeffreys prior; GP: gamma prior).

Prior	UP			JP			GP		
	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$
n	$p = 0.6$								
25	3.906965 (0.053263)	4.646122 (0.118138)	0.705412 (0.018581)	3.805187 (0.053447)	4.124574 (0.135086)	0.708148 (0.018277)	3.136139 (0.047532)	4.709692 (0.055450)	0.607486 (0.019830)
50	3.145525 (0.033605)	4.43439 (0.048442)	0.59793 (0.015276)	3.049301 (0.034752)	4.232523 (0.051022)	0.59838 (0.015042)	3.040991 (0.028685)	3.963530 (0.036190)	0.598702 (0.012375)
100	3.247894 (0.017049)	4.149792 (0.025881)	0.603628 (0.007538)	3.197226 (0.017339)	4.049354 (0.02685)	0.603932 (0.007533)	3.162700 (0.015778)	4.643948 (0.021732)	0.600413 (0.006846)
500	3.031835 (0.003514)	4.073316 (0.005393)	0.600777 (0.001557)	3.022097 (0.003526)	4.052578 (0.005423)	0.600839 (0.001554)	3.021671 (0.003452)	3.988061 (0.005196)	0.600824 (0.001518)
1000	3.026500 (0.001758)	4.051959 (0.002712)	0.600550 (0.000780)	3.016830 (0.001767)	4.041581 (0.002720)	0.600581 (0.000777)	3.016672 (0.001748)	4.040713 (0.002661)	0.600577 (0.000771)
n	$p = 0.4$								
25	3.843709 (0.095699)	5.713869 (0.061049)	0.416901 (0.063626)	3.502300 (0.106305)	5.403846 (0.064800)	0.415064 (0.064026)	2.959030 (0.045318)	5.262803 (0.050461)	0.440750 (0.056727)
50	3.614562 (0.050217)	4.543870 (0.032979)	0.413329 (0.032601)	3.444363 (0.052980)	4.408267 (0.034063)	0.412601 (0.032695)	3.044694 (0.027916)	4.674559 (0.033255)	0.428844 (0.029261)
100	3.183739 (0.026051)	4.403630 (0.016882)	0.404776 (0.016773)	3.105144 (0.026770)	4.335292 (0.017162)	0.404454 (0.016787)	3.065025 (0.017810)	4.479215 (0.024335)	0.400561 (0.016132)
500	3.043213 (0.005735)	4.033130 (0.003499)	0.401934 (0.003454)	3.026634 (0.005410)	4.020172 (0.003510)	0.401870 (0.003455)	2.980636 (0.003450)	4.351460 (0.004791)	0.401647 (0.003843)
1000	3.022642 (0.002707)	4.023263 (0.001753)	0.400505 (0.001736)	3.014894 (0.002715)	4.020780 (0.002748)	0.400473 (0.001743)	2.980088 (0.001830)	4.233064 (0.002552)	0.400577 (0.001790)

Table 10
Bayesian estimates and posterior risk for the real dataset (UP: uniform prior; JP: Jeffreys prior).

Prior	UP			JP		
LF	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$	$E(\theta_1\|\mathbf{x})$	$E(\theta_2\|\mathbf{x})$	$E(p\|\mathbf{x})$
SELF	0.183057 (0.00172821)	7.62493 (19.8644)	0.676655 (0.0412756)	0.127617 (0.00102926)	6.13288 (15.0845)	0.60305936 (0.0229137)
PLF	0.187718 (0.009322)	8.831985 (2.41411)	0.706497 (0.059683)	0.131588 (0.0079417)	7.25925 (2.252741)	0.621767 (0.037415)
MSELF	0.192498 (0.0014904)	10.23012 (0.254659)	0.737654 (0.042694)	0.135682 (0.0009442)	8.592491 (0.180251)	0.641055 (0.019272)
QLF	0.185073 (0.001531)	6.96423 (11.86104)	0.626505 (0.043272)	0.107414 (0.009918)	6.25828 (9.980451)	0.613603 (0.020524)
ELF	0.170573 (0.001613)	6.96423 (17.8464)	0.646545 (0.039267)	0.117471 (0.001018)	6.05882 (13.90845)	0.603610 (0.020942)
WSELF	0.180537 (0.0016228)	7.64923 (18.8446)	0.656565 (0.040276)	0.117627 (0.001022)	6.08828 (14.0854)	0.593060 (0.021924)

Initialize the starting points θ_1^0, θ_2^0 and p^0
For $j = 1,\ldots,N$
Generate θ_1^j from $g\left(\theta_1|\theta_2 = \theta_2^{j-1}, p^{j-1}\right)$.
Generate θ_2^j from $g\left(\theta_2|\theta_1 = \theta_1^j, p = p^{j-1}\right)$.
Generate p^j from $g\left(p|\theta_1 = \theta_1^j, \theta_2 = \theta_2^j\right)$.

Based on Tables 4–9, some points are very clear. The first important point that requires attention is that the posterior risk decreases as we increase the sample size. The second point is related to the comparison of loss functions. In the case of SELF, the Bayesian estimates are overestimated and the Bayesian estimates become very close to the true values of the parameters as we increase the sample size (this is not specific to the SELF because it was also verified for all other loss functions). With a large parameter value, the associated posterior risk has a high value. The second parameter value is overestimated when using a mixing component probability with a small value (Tables 4–9). There is a direct relationship between the mixing component probability and the posterior risk, where the posterior risk decreases as we increase the mixing probability. Kazmi et al. [2] reported the same relationship between the mixing component and the posterior risk for a mixture of the Maxwell distribution. For the Jeffreys prior, there is a smaller posterior risk compared with its counterpart: the uniform prior. Using an informative prior (the GP), the posterior risk is smaller than that with the noninformative priors (the uniform prior and the Jeffreys prior). The MSELF performs the best based on its lower posterior risk value.

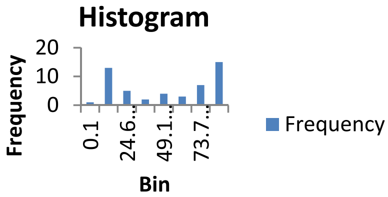

Fig. 4. Graphical representation of the original dataset.

6. Real data application

The following dataset was obtained from Aarset [24] and it represents the lifetimes of 50 devices.

1, 7, 18, 40, 45, 50, 55, 86, 85, 85, 85, 84, 84, 84, 79, 75, 72, 67, 67, 63, 60, 18, 11, 6, 3, 2, 1, 1, 0.2, 0.1
1, 18, 86, 85, 85, 83, 82, 82, 67, 67, 63, 47, 46, 36, 32, 21, 18, 12, 7, 1

Fig. 4 shows a graphical presentation of this dataset, which shows that its behavior approximates the mixture of inverse Rayleigh distribution.

We used the Kolmogorov Smirnov test and the difference obtained was 0.165421. Thus, this dataset followed the inverse Rayleigh distribution. Fig. 4 shows clearly that first 30 observations behaved differently compared with the other 20 observations. Thus, we took $n_1 = 30$ and $n_2 = 20$, and using this information we found $p = 0.60$. By taking the censoring time $T = 70$, the remainder of the information for the mixture setup is as follows.

$$n_1 = 30,\ n_2 = 20,\ r_1 = 20,\ r_2 = 14,\ \sum_{j=1}^{r_1} x_{1j}^{-2} = 128.4266,\ \sum_{j=1}^{r_2} x_{2j}^{-2} = 2.039164$$

We have no prior information about the process, thus we restrict our attention to the noninformative priors. Table 10 shows that MSELF had the best performance based on its minimum risk value, and the Jeffreys prior had lower posterior risks compared with the uniform prior.

7. Final remarks

In this study, we proposed a mixture of inverse Rayleigh model for lifetime study in engineering processes and we discussed its properties. Because of the effectiveness of Bayesian analysis, we performed a comprehensive study to address the problems of selecting priors (comparing informative and noninformative priors) and loss functions (based on the posterior risk using different asymmetric and symmetric loss functions) for the mixture of the inverse Rayleigh model. The simulation study obtained some interesting results related to the Bayesian estimates. The posterior risks of the estimates of the parameters appeared to be quite large with relatively larger values of the parameters, and vice versa. However, the posterior risk of parameters decreased as the sample size increased in each case. Another interesting finding related to the posterior risk of the estimates of the parameters is that increasing (decreasing) the proportion of the component in the mixture reduced (increased) the posterior risk of the estimate of the corresponding parameter. To address the problem of selecting priors and loss functions, we can categorize the posterior risk under different loss functions in the following order: MSELF $\leqslant$ ELF $\leqslant$ QSELF $\leqslant$ PLF < WSELF < SELF; and GP < JP < UP. The same pattern was observed for the real-life application. In the future, this method could be extended by comparing MLE with Bayesian estimates. Moreover, we also plan to use mixtures with record values and to obtain their predictions.

Acknowledgements

The author would like to thanks the anonymous reviewers and Editor for carefully reading the article and for their comments which greatly improved the manuscript. The author is also indebted to Professor Sonia Petrone, Department of Decision Sciences Bocconi University for providing constructive comments to reshape this idea.

Appendix A

The log-likelihood equation of the two-component mixture of the inverse Rayleigh distribution is given as

$$l = r_1 \ln p + r_1 \ln \theta_1 - \sum_{j=1}^{r_1} \frac{\theta_1}{x_{1j}^2} + r_2 \ln q + r_2 \ln \theta_2 - \sum_{j=1}^{r_2} \frac{\theta_2}{x_{2j}^2} + (n-r) \ln \left\{ 1 - \left(p \exp\left(-\frac{\theta_1}{T^2}\right) + q \exp\left(-\frac{\theta_2}{T^2}\right) \right) \right\}.$$

The normal equations used to obtain the ML estimates are as follows:

$$\frac{\partial l}{\partial \theta_1} = \frac{r_1}{\theta_1} - \sum_{j=1}^{r_1} \frac{1}{x_{1j}^2} + \frac{(n-r)(pT^{-2}\exp(-\theta_1 T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))} = 0,$$

$$\frac{\partial l}{\partial p} = \frac{r_1}{p} - \frac{r_2}{q} - \frac{(n-r)(\exp(-\theta_1 T^{-2}) - \exp(-\theta_2 T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))} = 0,$$

$$\frac{\partial l}{\partial \theta_2} = \frac{r_2}{\theta_2} - \sum_{j=1}^{r_2} \frac{1}{x_{2j}^2} + \frac{(n-r)(qT^{-2}\exp(-\theta_2 T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))} = 0.$$

To derive the variance of the likelihood estimates, the partial derivatives used in the Fisher information matrix are as follows:

$$\frac{\partial^2 l}{\partial \theta_1^2} = -\frac{r_1}{\theta_1^2} + \frac{(n-r)(pT^{-4}\exp(-\theta_1 T^{-2}) - pqT^{-4}\exp(-(\theta_1+\theta_2)T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))^2}$$

$$\frac{\partial^2 l}{\partial \theta_2^2} = -\frac{r_1}{\theta_2^2} + \frac{(n-r)(pqT^{-4}\exp(-(\theta_1+\theta_2)T^{-2}) - qT^{-4}\exp(-\theta_2 T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))^2}$$

$$\frac{\partial^2 l}{\partial p^2} = -\frac{r_1}{p^2} - \frac{r_2}{q^2} + \frac{(n-r)(\exp(-\theta_1 T^{-2}) - \exp(-\theta_2 T^{-2}))^2}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))^2}$$

$$\frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} = \frac{(n-r)(pqT^{-4}\exp(-(\theta_1+\theta_2)T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))^2}$$

$$\frac{\partial^2 l}{\partial p \partial \theta_1} = \frac{(n-r)(T^{-2}\exp(-\theta_1 T^{-2}) - T^{-2}\exp(-(\theta_1+\theta_2)T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))^2}$$

$$\frac{\partial^2 l}{\partial p \partial \theta_2} = \frac{(n-r)(T^{-2}\exp(-\theta_2 T^{-2}) - T^{-2}\exp(-(\theta_1+\theta_2)T^{-2}))}{(1-p\exp(-\theta_1 T^{-2}) - q\exp(-\theta_2 T^{-2}))^2}$$

Appendix B

Predictive distribution using GP

The posterior predictive distribution of a future observation y can be obtained as:

$p(y|\mathbf{x}) = \int_0^\infty \int_0^\infty \int_0^1 g(\theta_1, \theta_2, p|\mathbf{x}) p(y|\theta_1, \theta_2, p) dp d\theta_1 d\theta_2$, where $p(y|\theta_1, \theta_2, p) = p\left(\frac{2\theta_1}{y^3}\exp\left(-\frac{\theta_1}{y^2}\right)\right) + q\left(\frac{2\theta_2}{y^3}\exp\left(-\frac{\theta_2}{y^2}\right)\right)$, and $g(\theta_1, \theta_2, p|\mathbf{x})$ is given in (4). After simplification, we obtain the posterior predictive distribution of the future observation y as follows:

$$p(y|\mathbf{x}) = \frac{2}{C}\sum_k^{n-r}\sum_m^k (-1)^k \binom{n-r}{k}\binom{k}{m}\left[\begin{array}{l} \beta(r_1+k-m+a+1, r_2+m+b)\frac{\Gamma(r_1+\mu_1+1)}{y^3\left(C_1+\frac{1}{y^2}\right)^{(r_1+\mu_1+1)}}\frac{\Gamma(r_2+\mu_2)}{(C_2)^{(r_2+\mu_2)}} \\ +\beta(r_1+k-m+a, r_2+m+b+1)\frac{\Gamma(r_2+\mu_2+1)}{y^3\left(C_2+\frac{1}{y^2}\right)^{(r_2+\mu_2+1)}}\frac{\Gamma(r_1+\mu_1)}{(C_1)^{(r_1+\mu_1)}} \end{array}\right],$$

where $C_2 = B + \delta_2$ and $C_1 = A + \delta_1 C = \sum_k^{n-r}\sum_m^k (-1)^k \binom{n-r}{k}\binom{k}{m}\frac{\Gamma(r_2+\mu_2)}{(C_2)^{(r_2+\mu_2)}}\frac{\Gamma(r_1+\mu_1)}{(C_1)^{(r_1+\mu_1)}}\{\beta(r_1+k-m+a+1, r_2+m+b) + \beta(r_1+k-m+a, r_2+m+b+1)\}$. A $(1-\alpha)100\%$ Bayesian interval (L, U) can be obtained by solving the following two equations $\int_0^L p(y/x)dy = \frac{\alpha}{2} = \int_U^\infty p(y/x)dy$,

which can also be expressed after simplification as:

$$\frac{\alpha}{2} = \frac{1}{C}\sum_{k=0}^{n-r}\sum_{m=0}^k (-1)^k \binom{n-r}{k}\binom{k}{m}\left[\begin{array}{l} \beta(r_1+k-m+a+1, r_2+m+b)\frac{\Gamma(r_1+\mu_1)}{\left(C_1+\frac{1}{L^2}\right)^{(r_1+\mu_1)}}\frac{\Gamma(r_2+\mu_2)}{(C_2)^{(r_2+\mu_2)}} \\ +\beta(r_1+k-m+a, r_2+m+b+1)\frac{\Gamma(r_2+\mu_2)}{\left(C_2+\frac{1}{L^2}\right)^{(r_2+\mu_2)}}\frac{\Gamma(r_1+\mu_1)}{(C_1)^{(r_1+\mu_1)}} \end{array}\right]$$

and

$$\frac{\alpha}{2}=\frac{1}{C}\sum_{k=0}^{n-r}\sum_{m=0}^{k}(-1)^{k}\binom{n-r}{k}\binom{k}{m}\left[\begin{array}{l}\beta(r_1+k-m+a+1,r_2+m+b)\frac{\left(C_1U^2\left\{-1+\left(1+\frac{1}{C_1U^2}\right)^{(r_1+\mu_1+1)}\right\}-1\right)\Gamma(r_1+\mu_1)}{U^2\left(C_1+\frac{1}{U^2}\right)^{(r_1+\mu_1+1)}}\\ \frac{\Gamma(r_2+\mu_2)}{(C_2)^{(r_2+\mu_2)}}\\ +\beta(r_1+k-m+a,r_2+m+b+1)\frac{\left(C_2U^2\left\{-1+\left(1+\frac{1}{C_2U^2}\right)^{(r_2+\mu_2+1)}\right\}-1\right)\Gamma(r_2+\mu_2)}{U^2\left(C_2+\frac{1}{U^2}\right)^{(r_2+\mu_2+1)}}\\ \frac{\Gamma(r_1+\mu_1)}{(C_1)^{(r_1+\mu_1)}}\end{array}\right]$$

These predictive intervals can be used to evaluation the precision of the Bayesian estimates in terms of hyperparameters. If a trend is observed in terms of the hyper-parameters for the narrower predictive intervals, then a form of objectivity may be added based on prior information provided by a number of experts.

References

[1] K.S. Sultan, M.A. Ismail, A.S. Al-Moisheer, Mixture of two inverse Weibull distributions: properties and estimation, Comput. Stat. Data Anal. 51 (2007) 5377–5387.
[2] S.M.A. Kazmi, M. Aslam, S. Ali, N. Abbas, Selection of suitable prior for the Bayesian mixture of a class of lifetime distributions under type-I censored datasets, J. Appl. Stat. 40 (8) (2013) 1639–1658.
[3] E. Gomez-Déniz, L. Gomez-Déniz, A generalisation of the Rayleigh distribution with applications in wireless fading channels, Wireless Commun. Mob. Comput. 13 (2013) 85–94.
[4] A.A. Soliman, F.M. Al-Aboud, Bayesian inference using record values from Rayleigh model with application, Eur. J. Oper. Res. 185 (2) (2008) 659–672.
[5] H.A. Howlader, M.A. Hossain, O. Makhnin, Bayesian prediction bounds for Rayleigh and inverse Rayleigh lifetime models, J. Appl. Stat. Sci. 17 (1) (2008) 131–142.
[6] M. Aslam, Chi-Hyuck Jun, A group acceptance sampling plans for truncated life tests based on the inverse Rayleigh and log–logistic distributions, Pak. J. Stat. 25 (2) (2009) 107–119.
[7] A. Soliman, E.A. Amin, A.A. Abd-El Aziz, Estimation and prediction from Inverse Rayleigh distribution based on lower record values, Appl. Math. Sci. 4 (62) (2010) 3057–3066.
[8] K. Rosaiah, R.R.L. Kantam, Acceptance sampling based on the Inverse Rayleigh distribution, Econ. Qual. Control 20 (2) (2010) 277–286.
[9] S. Dey, Bayesian estimation of the parameter and reliability function of an inverse Rayleigh distribution, Malaysian J. Math. Sci. 6 (1) (2012) 113–124.
[10] S. Chandra, On the mixtures of probability distributions, Scand. J. Statist. 4 (1977) 105–112.
[11] W. Mendenhall, R.A. Hader, Estimation of parameter of mixed exponentially distribution failure time distribution from censored life test data, Biometrika 45 (3–4) (1958) 504–520.
[12] C. Chuan-Chong, K. Chee-Meng, Principles and Techniques in Combinatorics, World Scientific Publishing Co. Pvt. Ltd, Singapore, 1992.
[13] S.M.A. Kazmi, M. Aslam, S. Ali, A note on the maximum likelihood of mixture of Maxwell distribution using censored data, Open Stat. Probab. J. 3 (2011) 31–35.
[14] A.K. Bansal, Bayesian Parametric Inference, Narosa Publishing House Pvt. Ltd, New Delhi, 2007.
[15] H. Jeffreys, Theory of Probability, Clarendon Press, Oxford, UK, 1961.
[16] G.E.P. Box, G.C. Taio, Bayesian Inference in Statistics Analysis, Wiley & Sons, New York, 1973.
[17] S. Geisser, On prior distribution for binary trials, Am. Stat. 38 (4) (1984) 224–251.
[18] J.M. Bernardo, A.F.M. Smith, Bayesian Theory, John Wiley & Sons, Chichester, 1994.
[19] J.M. Bernardo, Reference posterior distribution for Bayesian inference (with discussion), J. R. Stat. Soc. Ser. B 41 (1979) 113–147.
[20] J.G. Norstrom, The use of precautionary loss functions in risk analysis, IEEE Trans. Reliab. 45 (1996) 400–403.
[21] R. Calabria, G. Pulcini, Point estimation under asymmetric loss functions for left truncated exponential samples, Commun. Stat. Theory Methods 25 (3) (1996) 585–600.
[22] M.H. DeGroot, Optimal Statistical Decisions, McGraw-Hill, New York, 1970.
[23] S. Ali, M. Aslam, S.M.A. Kazmi, A study of the effect of the loss function on Bayes estimate, posterior risk and hazard function for Lindley distribution, Appl. Math. Modell. 37 (8) (2013) 6068–6078.
[24] M.V. Aarset, How to identify bathtub hazard rate, IEEE Trans. Reliab. 36 (1) (1987) 106–108.