South African Statistical Journal
2018, Vol. 52, No. 1, 57–71

BAYESIAN ESTIMATION OF $P(Y > X)$ IN THE TWO-PARAMETER EXPONENTIAL DISTRIBUTION UTILIZING AN INITIAL GUESS

Mohammed Obeidat [1]
Department of Statistics, Yarmouk University, Irbid, Jordan
e-mail: *mohammad.obidat@yu.edu.jo*

Ayman Rawshdeh
Department of Statistics, Yarmouk University, Irbid, Jordan
e-mail: *ayman.r@yu.edu.jo*

This paper develops a Bayesian approach to estimate the stress-strength reliability, the probability that one random variable exceeds another. The proposed methodology utilizes an initial guess of this reliability through an informative prior, which constitutes the cornerstone of the model. Emphasis lies on exponentially distributed data, but the proposed method is applicable in a wider range of models with similar form of stress-strength reliability. A Monte Carlo simulation study is conducted to compare the performance of the new estimators with both the Maximum Likelihood and the Shrinkage estimators. The comparison is conducted with respect to the Mean Squared Error (MSE) for different values of the rate parameters of the exponential distribution. The proposed method outperforms the two aforementioned alternative methods. A demonstration is conducted through analyzing a real data set.

Key words: Bayes estimator, Exponential distribution, MCMC, Prior knowledge, Shrinkage estimator, Stress strength reliability.

1. Introduction

Birnbaum (1956) introduced the idea of estimating $R = P(X > Y)$, the probability that one random variable exceeds another, whereafter the idea attracted the attention of many authors in literature. R has many applications in a variety of different fields such as reliability analysis, in which R is known as the stress-strength model reliability. Furthermore, if Y models the strength of the device and X models the stress subjected on it, then the device fails any time its strength is exceeded by the stress applied on it. So the importance of R arises because its uses as a measure of system performance. Another interpretation of the parameter R is the effectiveness of a treatment or a drug when X and Y are the response variables for treatment and control groups, respectively. See Ventura and Racugno (2011).

Inference on R has received a great amount of attention and it has been studied extensively in various contexts, including parametric and non-parametric estimation using Bayesian and frequentist methods based on different data structures. Enis and Geisser (1971), Awad, Azzam and Hamdan

[1] Corresponding author.
MSC2010 subject classifications. 62F10, 62F15, 62N05.

(1981), and Tong (1974) discussed the problem of estimating R when X and Y are either independent or bivarate exponential random variables. When X and Y are independent normal random variables, estimation of R is considered by Govindarazulu (1967) and Woodward and Kelley (1977). Estimation of R is also considered by McCool (1991), Kundu and Gupta (2006), and Davarzani, Haghighi and Parsian (2009) for the Weibull case, and by Nadarajah (2004b), Nadarajah (2004a), Nadarajah (2005a), and Nadarajah (2005b) for logistic, Laplace, beta and gamma distributions respectively. Most recently, Samawi, Helu, Rochani, Yin and Linder (2016) considered the case when X and Y are dependent random variables with a bivariate underlying distribution.

Inference on R has been investigated given different contexts of the structure of the underlying data. Muttlak, Abu-Dayyeh, Saleh and Al-Sawi (2010) used ranked set sampling. Elfattah and Marwa (2008) considered the case based on censored samples. Baklizi (2008a), Baklizi (2008b) and Baklizi (2014) discussed point and interval estimation of R in the exponential case based on record values. Khamnei (2013) and Abdel-Hady (2014) studied the inference on R in the presence of outliers.

In some fields, an expert often possesses some prior information of R based on either past experience or from the technical structure of the system. Given a prior estimate R_0 of R, we are looking for an estimator that incorporates this information. In the frequentist paradigm, estimators that take advantage of the given prior guess are introduced by Thompson (1968) and termed as shrinkage estimators. Baklizi and Abu Dayyeh (2003) and Baklizi and El-Masri (2004) discussed different shrinkage estimators of R when X and Y are exponential. Chaturvedi and Nandchahal (2016) studied the characteristics of the shrinkage estimators of the reliability of a family of lifetime distributions. Montazer Haghighi and Shayib (2009) considered the shrinkage estimation of R for the Weibull type of distributions with common shape parameter. The use of shrinkage estimators in the context of linear models has been extensively studied by Saleh (2006).

The availability of the prior information of R is an ideal setting to be implemented within the Bayesian paradigm. Enis and Geisser (1971) considered a Bayesian approach for estimating R in the presence of prior guess R_0 for the one parameter exponential case. However, complicated procedures and expressions are given in their works and there were more decisions on the hyper-parameters of the prior distribution.

In this paper we discuss Bayesian estimation of the stress-strength reliability R, in the case of two-parameter exponential distribution with common location parameter. We assume that there is an expert prior guess, R_0 of R. The main original component of our proposed method is the elicitation of a joint prior density that is considered as jointly informative and marginally non-informative, and only uses the available prior guess R_0. Although we focus on data coming from an exponential distribution, our proposed method can be used in other models where R takes a similar form and/or in which the shrinkage estimation method may not be feasible. We then compare our proposed Bayesian method with the shrinkage estimator proposed by Baklizi and El-Masri (2004).

The structure of the remainder of the paper is as follows. The elicitation of a prior reflecting the previous knowledge of R together with the construction of the Bayesian model are presented in Section 2. The effectiveness of the proposed methodology is demonstrated using Monte Carlo simulation in Section 3. In Section 4 an illustrative example is considered through analyzing a real-world data set consisting of survival times of head and neck cancer patients. Section 5 concludes the paper.

2. Estimation methods

A positive random variable X follows a two-parameter exponential distribution with mean $1/\theta$, denoted by $\mathrm{Exp}(\theta, \mu)$, with probability density function $f_X(x; \theta, \mu) = \theta e^{-\theta(x-\mu)}$, for $x > \mu$ and $\theta, \mu > 0$. In this study, let X and Y be two independent random variables following the two-parameter exponential distribution with means $1/\theta$ and $1/\lambda$, respectively. The location parameter μ is assumed to be unknown but common. The probability that Y exceeds X is

$$\begin{aligned} R &= P(Y > X) \\ &= \int\int_{(Y>X)} f(x, y)\, dxdy = \int_{\mu}^{\infty} f_X(x; \theta, \mu)\left(\int_{x}^{\infty} f_Y(y; \theta, \mu)\, dy\right)dx = \frac{\theta}{\theta+\lambda}. \end{aligned} \tag{1}$$

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample from $\mathrm{Exp}(\theta, \mu)$ and $\mathbf{Y} = (Y_1, \ldots, Y_m)$ be a random sample from $\mathrm{Exp}(\lambda, \mu)$, with $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_m)$ being their realizations. The likelihood function of the parameter vector (θ, λ, μ) is

$$L(\theta, \lambda, \mu; \mathbf{x},\mathbf{y}) = \theta^n \lambda^m e^{-\left(\theta \sum\limits_{i=1}^{n}(x_i-\mu)+\lambda \sum\limits_{j=1}^{m}(y_j-\mu)\right)}. \tag{2}$$

The MLE of R can be shown to be $\hat{R}_{MLE} = \tilde{Y}/(\tilde{X} + \tilde{Y})$, where $\tilde{X} = \sum_{i=1}^{n}(x_i - u)/n$, $\tilde{Y} = \sum_{i=1}^{m}(y_i - u)/m$, and $u = \min(x_1, \cdots, x_n, y_1, \cdots, y_m)$.

2.1 Bayesian estimation

This section presents the building blocks of the Bayesian model, i.e. the prior, likelihood and posterior. The Bayesian estimators are developed using the squared error loss function, so the posterior mean is used to estimate the parameter of interest.

2.1.1 Prior intuition

In this section we discuss the translation of the available prior information about R, within the Bayesian paradigm, to a bivariate prior density $\pi = \pi(\theta, \lambda)$ defined on $\mathbb{R}_+^2$, where $\mathbb{R}_+^2 = \{(\theta, \lambda); \theta > 0, \lambda > 0\}$. The elicitation of the bivariate prior density $\pi(\theta, \lambda)$ is based on the prior information that the true value of R is believed to be close to R_0.

Equivalently to (1), we can write $\theta = m_R\lambda$, where $m_R = R/(1 - R)$, so each R value uniquely determines a line $L_{m_R} \subseteq \mathbb{R}_+^2$ that passes through the origin with a slope m_R, where $L_{m_R} = \{(\theta, \lambda) \in \mathbb{R}_+^2; \theta = m_R\lambda\}$. Set $\boldsymbol{\ell}$ to be the class of all possible lines, i.e. $\boldsymbol{\ell} = \{L_{m_R} \subseteq \mathbb{R}_+^2;$ where $0 < R < 1\}$, which constitute a partition of $\mathbb{R}_+^2$ in the sense that each $(\theta, \lambda) \in \mathbb{R}_+^2$ belongs to one and only one of the members of $\boldsymbol{\ell}$, which is L_{m_R}. Notice that the initial guess R_0 determines a line L_{m_0} in $\boldsymbol{\ell}$ with slope, $m_0 = R_0/(1 - R_0)$.

The bivariate prior distribution π should assign density values to points in $\mathbb{R}_+^2$ only through the class $\boldsymbol{\ell}$, i.e. π assigns density values to the lines L_R in $\boldsymbol{\ell}$, hypothetically speaking, instead of the points $(\theta, \lambda) \in \mathbb{R}_+^2$. Moreover, all points (θ, λ) that belong to the same line L_{m_R}, for some $0 < R < 1$, should be assigned with the same density value, while points belonging to different lines in $\boldsymbol{\ell}$ should be assigned with different density values (see Figure 1 for a clearer manifestation of this idea). The bivariate prior density $\pi(\theta, \lambda)$ is considered as both jointly informative and marginally non-informative in the following sense:

- **Jointly** informative: π prefers lines in $\boldsymbol{\ell}$ with slope m_R close to m_0. Particularly, $\pi(\theta, \lambda)$ assigns the highest density values for points (θ, λ) in the line L_{m_0}, while density values decreases for points in lines $L_{m_R} \in \boldsymbol{\ell}$, as m_R/m_0 deviates away from 1. Notice that only the relationship between θ and λ is imposed through the prior $\pi(\theta, \lambda)$.
- **Marginally** non-informative: Both marginal densities $\pi(\lambda) = \int_0^\infty \pi(\theta, \lambda) d\theta$ and $\pi(\theta) = \int_0^\infty \pi(\theta, \lambda) d\lambda$ should be non-informative. This reflects our lack of prior knowledge about the actual values of θ and λ.

2.1.2 Prior construction

Given the value of m_0, the joint prior $\pi(\theta, \lambda)$ is determined below. Let $\pi(\theta, \lambda)$ be expressed as $\pi(\lambda)\pi(\theta \mid \lambda, m_0)$, where $\pi(\lambda)$ and $\pi(\theta \mid \lambda, m_0)$ are as follows:

1. The gamma density is assigned as a prior to the parameter λ. Therefore, $\pi(\lambda) = \pi(\lambda \mid \alpha_0, \beta_0) \propto \lambda^{\alpha_0 - 1} e^{-\beta_0 \lambda}$. The hyper-parameters α_0 and β_0 are determined such that $\pi(\lambda)$ is either non-informative or weakly informative about the value of λ, which could be accomplished by the following, respectively:
    (a) Non-informative priors are widely used through the Jeffreys' prior (Jeffreys, 1961). Jeffreys' prior $\pi_J(\lambda)$ is proportional to $\sqrt{I(\lambda)}$, where $I(\lambda) = -E_X[\partial^2 \log L / \partial \lambda^2]$ and L is the likelihood function per observation given by $L = \log(\lambda) - (y - \mu)\lambda$. Differentiating L twice with respect to λ, implies $\partial^2 \log L / \partial \lambda^2 = -4/\lambda^2$. Therefore $\pi_J(\lambda) \propto \sqrt{I(\lambda)} = 1/\lambda$. Jeffreys' prior, $\pi_J(\lambda)$, is obtained by assigning the hyper-parameters $\alpha_0 = \beta_0 \to 0$.
    (b) Weakly-informative priors are obtained by assigning the hyper-parameters α_0 and β_0 small and equal values such as 0.01 or 0.001. One advantage of preferring weakly-informative prior is due to the fact that the resulting marginal posterior distribution of λ is a proper density. In our simulations we used a weakly informative prior for λ with $\alpha_0 = \beta_0 = 0.001$
2. The conditional prior distribution $\pi(\theta \mid \lambda, m_0)$ is constructed such that $\pi(\theta, \lambda)$ is jointly informative.
    (a) Given λ and m_0, the density value assigned to θ depends on the ratio of the slope of $L_{\theta/\lambda}$ to m_0. Specifically, $\pi(\theta \mid \lambda, m_0)$ is inversely proportional to deviations of the ratio $(\theta/\lambda)/m_0$ from 1. Therefore, set
    $$\pi(\theta \mid \lambda, m_0) \propto e^{-\frac{1}{2\sigma_0^2}\left(\frac{\theta}{m_0\lambda} - 1\right)^2} = e^{-\frac{1}{2\sigma_0^2}\left(\frac{\theta - m_0\lambda}{m_0\lambda}\right)^2},$$
    a truncated normal distribution with mean $m_0\lambda$ and standard deviation $\sigma_0 m_0 \lambda$. The hyperparameter σ_0^2 is carefully determined below.
    (b) One of the unique advantages of the proposed method is its ability to still incorporate the prior knowledge even when sample sizes n and m are large. This advantage is contrary to all shrinkage estimators in the literature. Set $\sigma_0 = \sqrt{1/(2n) + 1/(2m)}$. As the sample sizes n and m increase, the term $\sqrt{1/(2n) + 1/(2m)}$ decreases the variability of the bivariate distribution $\pi(\theta, \lambda)$. This enhances the impact of the prior information on the resulting

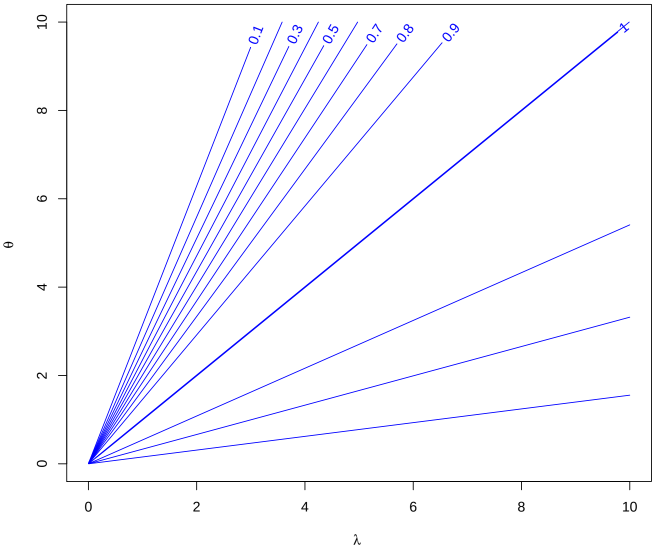

Figure 1. Contour plot of $\pi(\theta, \lambda)$ when $R_0 = 0.5$.

posterior mean of R, when n and m are large. Notice that the factor $\sqrt{1/(2n) + 1/(2m)}$ is motivated by the pooled standard deviation obtained from two independent samples.

Therefore the resulting joint prior $\pi(\theta, \lambda)$ is

$$\pi(\theta, \lambda) \propto \lambda^{\alpha_0 - 2} e^{-\beta_0 \lambda} e^{-\frac{1}{2}\left(\frac{1}{2n} + \frac{1}{2m}\right)^{-1}\left(\frac{\theta - m_0\lambda}{m_0\lambda}\right)^2}. \tag{3}$$

3. The parameter μ is assigned a flat prior over its domain, that is,

$$\pi(\mu) \propto 1, \quad 0 < \mu < \min(x_1, \cdots, x_n, y_1, \cdots, y_n). \tag{4}$$

2.1.3 Full conditional posterior densities

Combining the likelihood in (2) with the priors in (3) and (4), the joint posterior distribution of (θ, λ, μ) is as follows:

$$\pi(\theta, \lambda, \mu | \mathbf{x}, \mathbf{y}) \propto \theta^n \lambda^{\alpha_0 + m - 2} \times \exp\left\{ -\left(\theta \sum_{i=1}^{n} (x_i - \mu) + \lambda\left(\beta_0 + \sum_{j=1}^{m} (y_j - \mu)\right) + \frac{1}{2}\left(\frac{1}{2n} + \frac{1}{2m}\right)^{-1} \left[\frac{\theta - m_0\lambda}{m_0\lambda}\right]^2\right)\right\}.$$

Assuming squared error loss function, the Bayes estimator of R is the posterior mean, that is

$$\hat{R}_{BAY} = \frac{\int\int\int R \times \pi(\theta, \lambda, \mu | \mathbf{x}, \mathbf{x}) \, d\theta \, d\lambda \, d\mu}{\int\int\int \pi(\theta, \lambda, \mu | \mathbf{x}, \mathbf{y}) \, d\theta \, d\lambda \, d\mu}.$$

Clearly, the above integral is analytically intractable, so the Gibbs sampling, together with MCMC methods, are employed to obtain the posterior samples from (θ, λ, μ). Straightforward algebra shows that the conditional posterior distributions for one parameter given others are shown below

1. $\pi(\theta\,|\lambda, \mu, \cdot) \propto \theta^n \exp\{-(\theta - \mu_\theta)^2/2\sigma_\theta^2\}$, where $\mu_\theta = m_0\lambda(1 - m_0\lambda\sigma_\lambda^2\sum_{i=1}^n(x_i - \mu))$ and $\sigma_\theta = \sqrt{1/(2n) + 1/(2m)}\,(m_0\lambda)$. Notice that $\pi\,(\theta\,|\lambda, \mu\,\cdot)$ is a uni-modal log-concave function. Therefore, we can use the adaptive rejection sampling algorithm (ARS) proposed by Gilks and Wild (1992) to sample directly from its full conditional distribution.

2. $\pi(\lambda\,|\theta, \mu, \cdot) \propto \lambda^{\alpha_0+m-2} \exp\{-(\beta_0+\sum_{j=1}^m(y_j-\mu))\lambda\} \exp\{-\frac{1}{2}(1/(2n)+1/(2m))^{-1}(\theta/(m_0\lambda)-1)^2\}$. Although it is not easy to show the log-concavity of $\pi\,(\lambda\,|\theta, \mu,\ \cdot)$, our simulation studies showed that the second derivative of $\log\pi\,(\lambda\,|\theta, \mu,\ \cdot)$ was always negative in all cases considered. Therefore, ARS is used to obtain samples from $\pi\,(\lambda\,|\theta, \mu, \cdot)$.

3. $\pi(\mu|\theta, \lambda, \cdot) \propto e^{\mu(n\theta+m\lambda)}$ for $0 < \mu < \min(x_1, \ldots, x_n, y_1, \ldots, y_m)$. Sampling from this posterior can be easily accomplished using the inverse transform sampling.

The obtained posterior samples are used to approximate $\hat{R}_{BAY}$. The Gibbs sampler algorithm is summarized as follows:

- Start with an initial guess $\lambda^{(0)}$, and $\theta^{(0)}$,
- for $i \geq 1$,
  - generate $\mu^{(i)}$ from $\pi(\mu|\theta^{(i-1)}, \lambda^{(i-1)}, \cdot)$ using the inverse transform method,
  - generate $\theta^{(i)}$ from $\pi\left(\theta\,|\lambda^{(i-1)}, \mu^{(i)}, \cdot\right)$ using ARS,
  - generate $\lambda^{(i)}$ from $\pi\left(\lambda\,|\theta^{(i)}, \mu^{(i)}, \cdot\right)$ using ARS,
  - calculate $R^{(i)} = \theta^{(i)}/(\theta^{(i)} + \lambda^{(i)})$,
- repeat the previous step M times, and
- calculate $\hat{R}_{BAY} = \sum R^{(i)}/M$.

In addition to the point estimator, $\hat{R}_{BAY}$, we can obtain a Bayesian credible interval from the posterior samples, $R^{(i)}$, $i = 1, \ldots, M$. One popular credible interval is the Highest Posterior Density (HPD) credible interval. The HPD interval can be constructed from the empirical cumulative distribution function (cdf) of the posterior samples as the shortest interval for which the difference in the empirical cdf values of the endpoints is the desired nominal probability.

2.2 Shrinkage estimation

The shrinkage estimator of R is a linear combination between the MLE of R, $\hat{R}_{MLE}$, and the prior guess R_0, that is $\hat{R}_{SHR} = wR_0 + (1-w)\hat{R}_{MLE}$, where $0 \leq w \leq 1$ is a weight to be chosen. Baklizi and El-Masri (2004) proposed several methods for selecting w. The choice of w that led to the highest relative efficiency of the shrinkage estimator of R compared to the MLE was the one that minimized the mean square error of the shrinkage estimator $\hat{R}_{SHR} = wR_0 + (1-w)\hat{R}_{MLE}$. Baklizi and El-Masri (2004) found w to be

$$w = \frac{(\hat{R}_{MLE} - R_0)(E(\hat{R}_{MLE}) - R_0)}{E(\hat{R}_{MLE}^2) - 2R_0E(\hat{R}_{MLE}) + R_0^2},$$

where $E(\hat{R}_{MLE})$ and $E(\hat{R}^2_{MLE})$ are the first and second moment of $\hat{R}_{MLE}$ respectively, and are defined as follows:

$$E(\hat{R}_{MLE}) = \frac{n^{1-m}m^{1-n}}{n\theta + m\lambda}\frac{\theta^n\lambda^m}{n+m-1}\left(\frac{\theta}{m}\right)^{1-n-m} \times \left((m-1)\mathbf{2F1}\left[m, n+m-1, n+m, 1-\frac{m\lambda}{n\theta}\right] + m\mathbf{2F1}\left[1+m, n+m-1, n+m, 1-\frac{m\lambda}{n\theta}\right]\right),$$

and

$$E(\hat{R}^2_{MLE}) = \frac{n^{1-m}m^{1-n}}{n\theta + m\lambda}\frac{\theta^{n+1}\lambda^m}{(n+m)(n+m-1)}\left(\frac{\theta}{m}\right)^{-n-m} \times \left((m-1)\mathbf{2F1}\left[1+m, n+m-1, n+m+1, 1-\frac{m\lambda}{n\theta}\right] + (m+1)\mathbf{2F1}\left[2+m, n+m-1, n+m+1, 1-\frac{m\lambda}{n\theta}\right]\right),$$

where $\mathbf{2F1}[a, b, c, d]$ is the Gaussian hypergeometric function. The expectations above depend on the unknown parameters θ and λ and they may be replaced with their maximum likelihood estimates to estimate $E(\hat{R}_{MLE})$ and $E(\hat{R}^2_{MLE})$. See Baklizi and El-Masri (2004) for more details.

3. Monte Carlo simulation study

To gain insight into the efficiency of the three estimators of R, we perform a Monte Carlo simulation study. Different combinations of the sample sizes were considered. For each combination of n and m, 2000 samples were generated from X and Y. Three values of the reliability were considered: $R = 0.2, 0.5, 0.8$. For each value of R, three values of θ are studied: $\theta = 0.2, 1, 5$. The values of λ are obtained by $\lambda = \theta R/(1-R)$. The initial guess of R was taken to be $R \pm 0.15$ with a step of 0.05. In all cases μ is set to zero.

The three estimators are calculated for each data set. The Bayesian estimator $\hat{R}_{BAY}$ is obtained from each simulated dataset with an MCMC of length 1000 iterations such that 400 simulations are for burn-in, while the last 600 draws of the posterior samples are used in the calculation. The mean square error of all estimators is calculated using the formula

$$\text{MSE}(\hat{R}) = \frac{\sum(\hat{R} - R)^2}{2000}.$$

The relative efficiency of $\hat{R}_{BAY}$ and $\hat{R}_{SHR}$ with respect to the MLE are then calculated using the formula

$$\text{RE}(\hat{R}) = \frac{\text{MSE}(\hat{R}_{MLE})}{\text{MSE}(\hat{R})}.$$

Tables 1-6 show the relative efficiency of the Bayes and Shrinkage estimators with respect to the MLE. It can be seen that the Bayes estimator is successful in taking advantage of the prior guess R_0 over the MLE and the shrinkage estimators, especially when R_0 lies within 5–15% of the true value of R. The performance of the Bayes estimator improves as the sample sizes get small or when the true value of R is around 0.5.

Table 1. MSE of the MLE and RE of the Bayes and shrinkage estimators of R when $\theta = 0.2$ and $(n, m) = \{(5, 5), (5, 10), (10, 5)\}$.

		$R = 0.2$				$R = 0.5$				$R = 0.8$			
n	m	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$
5	5	0.05	0.0123	0.7027	1.0033	0.35	0.0266	1.8397	1.2122	0.65	0.0118	0.8596	1.0420
5	5	0.10	0.0123	1.8516	1.2946	0.40	0.0278	3.5386	1.4545	0.70	0.0128	1.7569	1.3527
5	5	0.15	0.0122	5.8771	1.7108	0.45	0.0276	6.1448	1.6667	0.75	0.0130	3.6537	1.7028
5	5	0.20	0.0130	11.1301	1.9124	0.50	0.0281	7.9829	1.7307	0.80	0.0132	6.1810	1.8718
5	5	0.25	0.0120	4.4018	1.7078	0.55	0.0275	5.9299	1.6735	0.85	0.0129	4.6357	1.7357
5	5	0.30	0.0135	2.1386	1.3383	0.60	0.0289	3.5399	1.4505	0.90	0.0133	2.3053	1.3115
5	5	0.35	0.0127	1.1584	1.0594	0.65	0.0270	2.0606	1.2304	0.95	0.0122	1.1760	1.0015
5	10	0.05	0.0102	0.5693	1.0424	0.35	0.0201	1.4338	1.1875	0.65	0.0084	0.6282	0.9843
5	10	0.10	0.0104	1.5886	1.2793	0.40	0.0207	3.0190	1.4339	0.70	0.0085	1.3258	1.2361
5	10	0.15	0.0096	5.5912	1.7836	0.45	0.0198	6.2317	1.7421	0.75	0.0083	3.2714	1.6918
5	10	0.20	0.0102	12.3504	2.1083	0.50	0.0208	8.9566	1.8303	0.80	0.0086	7.1957	1.9605
5	10	0.25	0.0094	3.9493	1.7213	0.55	0.0198	5.5304	1.6921	0.85	0.0084	3.8854	1.6133
5	10	0.30	0.0095	1.6533	1.2181	0.60	0.0193	2.7647	1.4013	0.90	0.0081	1.5734	1.0991
5	10	0.35	0.0102	1.0050	0.9690	0.65	0.0200	1.6845	1.1283	0.95	0.0080	0.9103	0.9092
10	5	0.05	0.0085	0.4960	0.9194	0.35	0.0206	1.4298	1.1197	0.65	0.0106	0.7907	0.9675
10	5	0.10	0.0082	1.2650	1.1021	0.40	0.0197	2.7234	1.3858	0.70	0.0095	1.4110	1.2207
10	5	0.15	0.0083	4.3356	1.5768	0.45	0.0197	5.3831	1.6760	0.75	0.0097	3.2627	1.6794
10	5	0.20	0.0084	8.7734	1.9393	0.50	0.0206	7.5626	1.8190	0.80	0.0103	6.8312	2.0886
10	5	0.25	0.0086	3.7075	1.6621	0.55	0.0209	5.4384	1.7126	0.85	0.0106	4.6108	1.7785
10	5	0.30	0.0082	1.5373	1.2478	0.60	0.0193	2.9813	1.4862	0.90	0.0091	1.8991	1.2786
10	5	0.35	0.0083	0.8740	0.9779	0.65	0.0189	1.6779	1.1669	0.95	0.0092	0.9891	1.0266

In addition to point estimation we constructed a 95% HPD credible interval of the reliability for different combinations of the parameter values. We noticed that the coverage probability is higher than the nominal probability when R_0 within 10–15% of R for small samples and when it is within 5–10% of R for larger sample sizes. The coverage probability gets less than the nominal rate as R_0 moves further away from R. This conclusion is true regardless of the parameter values. It was also noticed that the length of the credible interval was not affected by the initial guess. The coverage probability and half length of the HPD interval are summarized in Table 7.

4. Real data analysis

4.1 Data description

To illustrate our method on real data, we consider a data presented in Singh, Singh, Yadav and Viswkarma (2015), which consists of the survival times of two treatment groups of head and neck cancer patients. The first group represents the survival times of 58 head and neck cancer patients treated with radiotherapy, while the other group represents the survival times of 45 head and neck cancer patients treated with combined radiotherapy and chemotherapy. The survival times are provided in Table 8. The histograms of these data are given in Figure 2. Histograms reveals that the exponential distribution is a good fit for the data.

4.2 Data analysis results

In this section we analyzed 1 000 bootstrap samples that were randomly selected with replacement from the above two groups such that $(n, m) = \{(5, 5), (10, 10), (20, 20)\}$. By analyzing the complete data as to be our populations, the true value of R is found to be 0.508. Table 9 presents the mean

Table 2. MSE of the MLE and RE of the Bayes and shrinkage estimators of R when $\theta = 0.2$ and $(n, m) = \{(10, 10), (10, 20), (20, 10), (20, 20)\}$.

		$R = 0.2$				$R = 0.5$				$R = 0.8$			
n	m	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$
10	10	0.05	0.0056	0.3228	0.9451	0.35	0.0129	0.9668	1.0253	0.65	0.0057	0.4576	0.8421
10	10	0.10	0.0060	0.9332	1.0564	0.40	0.0136	2.1141	1.3248	0.70	0.0057	0.9684	1.0835
10	10	0.15	0.0056	3.3870	1.5517	0.45	0.0131	4.8098	1.6896	0.75	0.0059	2.6768	1.5775
10	10	0.20	0.0056	9.6949	2.0642	0.50	0.0129	8.2085	1.9154	0.80	0.0056	7.2100	2.0579
10	10	0.25	0.0059	3.0559	1.5722	0.55	0.0133	4.8288	1.6890	0.85	0.0056	3.3754	1.5715
10	10	0.30	0.0059	1.2339	1.1057	0.60	0.0136	2.3076	1.3175	0.90	0.0060	1.3093	1.0637
10	10	0.35	0.0057	0.6750	0.8403	0.65	0.0134	1.3086	1.0393	0.95	0.0059	0.7586	0.9596
10	20	0.05	0.0044	0.2464	0.9791	0.35	0.0097	0.7264	0.9640	0.65	0.0040	0.3184	0.7995
10	20	0.10	0.0046	0.7041	1.0533	0.40	0.0099	1.6678	1.2306	0.70	0.0040	0.7189	0.9803
10	20	0.15	0.0043	2.8795	1.5292	0.45	0.0095	4.5111	1.6813	0.75	0.0038	2.2029	1.4769
10	20	0.20	0.0041	10.0323	2.1646	0.50	0.0094	8.4771	1.9862	0.80	0.0039	7.4552	2.0657
10	20	0.25	0.0045	2.6063	1.4839	0.55	0.0098	4.2599	1.6402	0.85	0.0040	2.7059	1.4132
10	20	0.30	0.0046	1.0432	0.9922	0.60	0.0099	1.8746	1.2024	0.90	0.0040	0.9667	0.9355
10	20	0.35	0.0046	0.5961	0.7892	0.65	0.0101	1.0680	0.9393	0.95	0.0042	0.6433	0.9322
20	10	0.05	0.0041	0.2453	0.9041	0.35	0.0099	0.7650	0.9246	0.65	0.0045	0.3768	0.7418
20	10	0.10	0.0041	0.6594	0.9533	0.40	0.0096	1.6314	1.1901	0.70	0.0043	0.7794	0.9485
20	10	0.15	0.0041	2.5935	1.4462	0.45	0.0100	4.2562	1.6655	0.75	0.0044	2.3021	1.4936
20	10	0.20	0.0040	8.5118	1.9979	0.50	0.0101	7.7600	1.9403	0.80	0.0047	7.2850	2.1216
20	10	0.25	0.0041	2.5722	1.4975	0.55	0.0098	4.5005	1.7157	0.85	0.0044	3.1850	1.5805
20	10	0.30	0.0040	0.9194	0.9897	0.60	0.0095	1.8442	1.2343	0.90	0.0042	1.0033	1.0205
20	10	0.35	0.0040	0.5166	0.8045	0.65	0.0100	1.0662	0.9762	0.95	0.0047	0.6373	1.0177
20	20	0.05	0.0027	0.1544	0.9391	0.35	0.0064	0.4980	0.8501	0.65	0.0027	0.2292	0.7313
20	20	0.10	0.0027	0.4220	0.9219	0.40	0.0063	1.1279	1.0820	0.70	0.0026	0.4961	0.8457
20	20	0.15	0.0026	1.7611	1.2777	0.45	0.0060	3.2922	1.5802	0.75	0.0025	1.5940	1.2738
20	20	0.20	0.0026	9.2600	2.1426	0.50	0.0062	8.4922	2.0649	0.80	0.0026	7.8816	2.1909
20	20	0.25	0.0027	1.8924	1.3096	0.55	0.0065	3.5190	1.5662	0.85	0.0027	2.1345	1.2909
20	20	0.30	0.0026	0.6659	0.8512	0.60	0.0061	1.3151	1.0739	0.90	0.0025	0.6450	0.8938
20	20	0.35	0.0025	0.3639	0.7359	0.65	0.0061	0.7088	0.8546	0.95	0.0026	0.4384	0.9553

Table 3. MSE of the MLE and RE of the Bayes and shrinkage estimators of R when $\theta = 1$ and $(n, m) = \{(5, 5), (5, 10), (10, 5)\}$.

		$R = 0.2$				$R = 0.5$				$R = 0.8$			
n	m	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$
5	5	0.05	0.0123	0.7028	1.0033	0.35	0.0266	1.8402	1.2122	0.65	0.0118	0.8620	1.0420
5	5	0.10	0.0123	1.8533	1.2946	0.40	0.0278	3.5471	1.4545	0.70	0.0128	1.7582	1.3527
5	5	0.15	0.0122	5.8660	1.7108	0.45	0.0276	6.1597	1.6667	0.75	0.0130	3.6450	1.7028
5	5	0.20	0.0130	11.1611	1.9124	0.50	0.0281	8.0159	1.7307	0.80	0.0132	6.1556	1.8718
5	5	0.25	0.0120	4.4073	1.7078	0.55	0.0275	5.9251	1.6735	0.85	0.0129	4.6091	1.7357
5	5	0.30	0.0135	2.1342	1.3383	0.60	0.0289	3.5443	1.4505	0.90	0.0133	2.2934	1.3115
5	5	0.35	0.0127	1.1555	1.0594	0.65	0.0270	2.0596	1.2304	0.95	0.0122	1.1777	1.0015
5	10	0.05	0.0102	0.5694	1.0424	0.35	0.0201	1.4338	1.1875	0.65	0.0084	0.6286	0.9843
5	10	0.10	0.0104	1.5893	1.2793	0.40	0.0207	3.0186	1.4339	0.70	0.0085	1.3253	1.2361
5	10	0.15	0.0096	5.5884	1.7836	0.45	0.0198	6.2345	1.7421	0.75	0.0083	3.2775	1.6918
5	10	0.20	0.0102	12.3948	2.1083	0.50	0.0208	8.9431	1.8303	0.80	0.0086	7.2281	1.9605
5	10	0.25	0.0094	3.9503	1.7213	0.55	0.0198	5.5174	1.6921	0.85	0.0084	3.8972	1.6133
5	10	0.30	0.0095	1.6504	1.2181	0.60	0.0193	2.7602	1.4013	0.90	0.0081	1.5735	1.0991
5	10	0.35	0.0102	1.0046	0.9690	0.65	0.0200	1.6854	1.1283	0.95	0.0080	0.9096	0.9092
10	5	0.05	0.0085	0.4962	0.9194	0.35	0.0206	1.4307	1.1197	0.65	0.0106	0.7904	0.9675
10	5	0.10	0.0082	1.2655	1.1021	0.40	0.0197	2.7211	1.3858	0.70	0.0095	1.4137	1.2207
10	5	0.15	0.0083	4.3485	1.5768	0.45	0.0197	5.4065	1.6760	0.75	0.0097	3.2619	1.6794
10	5	0.20	0.0084	8.7827	1.9393	0.50	0.0206	7.5475	1.8190	0.80	0.0103	6.8428	2.0886
10	5	0.25	0.0086	3.7085	1.6621	0.55	0.0209	5.4260	1.7126	0.85	0.0106	4.5918	1.7785
10	5	0.30	0.0082	1.5388	1.2478	0.60	0.0193	2.9784	1.4862	0.90	0.0091	1.9026	1.2786
10	5	0.35	0.0083	0.8735	0.9779	0.65	0.0189	1.6752	1.1669	0.95	0.0092	0.9916	1.0266

Table 4. MSE of the MLE and RE of the Bayes and shrinkage estimators of R when $\theta = 1$ and (n, m)={(10,10),(10,20),(20,10),(20,20)}.

		$R = 0.2$				$R = 0.5$				$R = 0.8$			
n	m	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$
10	10	0.05	0.0056	0.3228	0.9451	0.35	0.0129	0.9672	1.0253	0.65	0.0057	0.4574	0.8421
10	10	0.10	0.0060	0.9341	1.0564	0.40	0.0136	2.1150	1.3248	0.70	0.0057	0.9673	1.0835
10	10	0.15	0.0056	3.3879	1.5517	0.45	0.0131	4.8066	1.6896	0.75	0.0059	2.6642	1.5775
10	10	0.20	0.0056	9.6850	2.0642	0.50	0.0129	8.2199	1.9154	0.80	0.0056	7.1716	2.0579
10	10	0.25	0.0059	3.0536	1.5722	0.55	0.0133	4.8317	1.6890	0.85	0.0056	3.3651	1.5715
10	10	0.30	0.0059	1.2299	1.1057	0.60	0.0136	2.3047	1.3175	0.90	0.0060	1.3104	1.0637
10	10	0.35	0.0057	0.6746	0.8403	0.65	0.0134	1.3109	1.0393	0.95	0.0059	0.7580	0.9596
10	20	0.05	0.0044	0.2465	0.9791	0.35	0.0097	0.7263	0.9640	0.65	0.0040	0.3186	0.7995
10	20	0.10	0.0046	0.7044	1.0533	0.40	0.0099	1.6666	1.2306	0.70	0.0040	0.7190	0.9803
10	20	0.15	0.0043	2.8834	1.5292	0.45	0.0095	4.5119	1.6813	0.75	0.0038	2.1959	1.4769
10	20	0.20	0.0041	10.0606	2.1646	0.50	0.0094	8.4617	1.9862	0.80	0.0039	7.4804	2.0657
10	20	0.25	0.0045	2.6029	1.4839	0.55	0.0098	4.2535	1.6402	0.85	0.0040	2.7028	1.4132
10	20	0.30	0.0046	1.0401	0.9922	0.60	0.0099	1.8695	1.2024	0.90	0.0040	0.9675	0.9355
10	20	0.35	0.0046	0.5951	0.7892	0.65	0.0101	1.0664	0.9393	0.95	0.0042	0.6439	0.9322
20	10	0.05	0.0041	0.2454	0.9041	0.35	0.0099	0.7652	0.9246	0.65	0.0045	0.3768	0.7418
20	10	0.10	0.0041	0.6596	0.9533	0.40	0.0096	1.6303	1.1901	0.70	0.0043	0.7794	0.9485
20	10	0.15	0.0041	2.5960	1.4462	0.45	0.0100	4.2635	1.6655	0.75	0.0044	2.3057	1.4936
20	10	0.20	0.0040	8.5878	1.9979	0.50	0.0101	7.7903	1.9403	0.80	0.0047	7.3005	2.1216
20	10	0.25	0.0041	2.5775	1.4975	0.55	0.0098	4.4965	1.7157	0.85	0.0044	3.1904	1.5805
20	10	0.30	0.0040	0.9184	0.9897	0.60	0.0095	1.8447	1.2343	0.90	0.0042	1.0038	1.0205
20	10	0.35	0.0040	0.5160	0.8045	0.65	0.0100	1.0654	0.9762	0.95	0.0047	0.6373	1.0177
20	20	0.05	0.0027	0.1544	0.9391	0.35	0.0064	0.4981	0.8501	0.65	0.0027	0.2290	0.7313
20	20	0.10	0.0027	0.4221	0.9219	0.40	0.0063	1.1280	1.0820	0.70	0.0026	0.4964	0.8457
20	20	0.15	0.0026	1.7592	1.2777	0.45	0.0060	3.2950	1.5802	0.75	0.0025	1.5936	1.2738
20	20	0.20	0.0026	9.2723	2.1426	0.50	0.0062	8.4808	2.0649	0.80	0.0026	7.8568	2.1909
20	20	0.25	0.0027	1.8914	1.3096	0.55	0.0065	3.5210	1.5662	0.85	0.0027	2.1379	1.2909
20	20	0.30	0.0026	0.6661	0.8512	0.60	0.0061	1.3166	1.0739	0.90	0.0025	0.6452	0.8938
20	20	0.35	0.0025	0.3637	0.7359	0.65	0.0061	0.7086	0.8546	0.95	0.0026	0.4382	0.9553

Table 5. MSE of the MLE and RE of the Bayes and shrinkage estimators of R when $\theta = 5$ and $(n, m) = \{(5, 5), (5, 10), (10, 5)\}$.

		$R = 0.2$				$R = 0.5$				$R = 0.8$			
n	m	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$
5	5	0.05	0.0123	0.7033	1.0033	0.35	0.0266	1.8426	1.2122	0.65	0.0118	0.8614	1.0420
5	5	0.10	0.0123	1.8601	1.2946	0.40	0.0278	3.5514	1.4545	0.70	0.0128	1.7626	1.3527
5	5	0.15	0.0122	5.9325	1.7108	0.45	0.0276	6.2038	1.6667	0.75	0.0130	3.6691	1.7028
5	5	0.20	0.0130	11.2945	1.9124	0.50	0.0281	8.0494	1.7307	0.80	0.0132	6.1885	1.8718
5	5	0.25	0.0120	4.3946	1.7078	0.55	0.0275	5.9109	1.6735	0.85	0.0129	4.5919	1.7357
5	5	0.30	0.0135	2.1294	1.3383	0.60	0.0289	3.5328	1.4505	0.90	0.0133	2.2893	1.3115
5	5	0.35	0.0127	1.1501	1.0594	0.65	0.0270	2.0576	1.2304	0.95	0.0122	1.1743	1.0015
5	10	0.05	0.0102	0.5696	1.0424	0.35	0.0201	1.4353	1.1875	0.65	0.0084	0.6290	0.9843
5	10	0.10	0.0104	1.5902	1.2793	0.40	0.0207	3.0223	1.4339	0.70	0.0085	1.3264	1.2361
5	10	0.15	0.0096	5.6100	1.7836	0.45	0.0198	6.2521	1.7421	0.75	0.0083	3.2879	1.6918
5	10	0.20	0.0102	12.4816	2.1083	0.50	0.0208	8.9418	1.8303	0.80	0.0086	7.2642	1.9605
5	10	0.25	0.0094	3.9439	1.7213	0.55	0.0198	5.5230	1.6921	0.85	0.0084	3.8945	1.6133
5	10	0.30	0.0095	1.6435	1.2181	0.60	0.0193	2.7600	1.4013	0.90	0.0081	1.5688	1.0991
5	10	0.35	0.0102	1.0010	0.9690	0.65	0.0200	1.6814	1.1283	0.95	0.0080	0.9067	0.9092
10	5	0.05	0.0085	0.4965	0.9194	0.35	0.0206	1.4331	1.1197	0.65	0.0106	0.7909	0.9675
10	5	0.10	0.0082	1.2683	1.1021	0.40	0.0197	2.7255	1.3858	0.70	0.0095	1.4124	1.2207
10	5	0.15	0.0083	4.3807	1.5768	0.45	0.0197	5.4178	1.6760	0.75	0.0097	3.2769	1.6794
10	5	0.20	0.0084	8.9229	1.9393	0.50	0.0206	7.6298	1.8190	0.80	0.0103	6.8087	2.0886
10	5	0.25	0.0086	3.7024	1.6621	0.55	0.0209	5.4448	1.7126	0.85	0.0106	4.6234	1.7785
10	5	0.30	0.0082	1.5325	1.2478	0.60	0.0193	2.9739	1.4862	0.90	0.0091	1.9044	1.2786
10	5	0.35	0.0083	0.8696	0.9779	0.65	0.0189	1.6746	1.1669	0.95	0.0092	0.9889	1.0266

Table 6. MSE of the MLE and RE of the Bayes and Shrinkage Estimators of R When $\theta = 5$, and $(n, m) = \{(10, 10), (10, 20), (20, 10), (20, 20)\}$.

		$R = 0.2$				$R = 0.5$				$R = 0.8$			
n	m	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$	R_0	MLE	$\hat{R}_B$	$\hat{R}_S$
10	10	0.05	0.0056	0.3229	0.9451	0.35	0.0129	0.9685	1.0253	0.65	0.0057	0.4581	0.8421
10	10	0.10	0.0060	0.9350	1.0564	0.40	0.0136	2.1150	1.3248	0.70	0.0057	0.9674	1.0835
10	10	0.15	0.0056	3.3944	1.5517	0.45	0.0131	4.8113	1.6896	0.75	0.0059	2.6746	1.5775
10	10	0.20	0.0056	9.7046	2.0642	0.50	0.0129	8.2198	1.9154	0.80	0.0056	7.2033	2.0579
10	10	0.25	0.0059	3.0484	1.5722	0.55	0.0133	4.8248	1.6890	0.85	0.0056	3.3709	1.5715
10	10	0.30	0.0059	1.2298	1.1057	0.60	0.0136	2.3032	1.3175	0.90	0.0060	1.3095	1.0637
10	10	0.35	0.0057	0.6724	0.8403	0.65	0.0134	1.3088	1.0393	0.95	0.0059	0.7568	0.9596
10	20	0.05	0.0044	0.2465	0.9791	0.35	0.0097	0.7267	0.9640	0.65	0.0040	0.3184	0.7995
10	20	0.10	0.0046	0.7046	1.0533	0.40	0.0099	1.6653	1.2306	0.70	0.0040	0.7185	0.9803
10	20	0.15	0.0043	2.8858	1.5292	0.45	0.0095	4.5128	1.6813	0.75	0.0038	2.1954	1.4769
10	20	0.20	0.0041	10.0658	2.1646	0.50	0.0094	8.4391	1.9862	0.80	0.0039	7.4347	2.0657
10	20	0.25	0.0045	2.6022	1.4839	0.55	0.0098	4.2592	1.6402	0.85	0.0040	2.7036	1.4132
10	20	0.30	0.0046	1.0406	0.9922	0.60	0.0099	1.8716	1.2024	0.90	0.0040	0.9672	0.9355
10	20	0.35	0.0046	0.5944	0.7892	0.65	0.0101	1.0660	0.9393	0.95	0.0042	0.6422	0.9322
20	10	0.05	0.0041	0.2454	0.9041	0.35	0.0099	0.7653	0.9246	0.65	0.0045	0.3767	0.7418
20	10	0.10	0.0041	0.6604	0.9533	0.40	0.0096	1.6329	1.1901	0.70	0.0043	0.7799	0.9485
20	10	0.15	0.0041	2.6061	1.4462	0.45	0.0100	4.2815	1.6655	0.75	0.0044	2.3079	1.4936
20	10	0.20	0.0040	8.5478	1.9979	0.50	0.0101	7.7699	1.9403	0.80	0.0047	7.2853	2.1216
20	10	0.25	0.0041	2.5705	1.4975	0.55	0.0098	4.5056	1.7157	0.85	0.0044	3.1875	1.5805
20	10	0.30	0.0040	0.9163	0.9897	0.60	0.0095	1.8437	1.2343	0.90	0.0042	1.0030	1.0205
20	10	0.35	0.0040	0.5140	0.8045	0.65	0.0100	1.0644	0.9762	0.95	0.0047	0.6367	1.0177
20	20	0.05	0.0027	0.1544	0.9391	0.35	0.0064	0.4978	0.8501	0.65	0.0027	0.2291	0.7313
20	20	0.10	0.0027	0.4222	0.9219	0.40	0.0063	1.1281	1.0820	0.70	0.0026	0.4966	0.8457
20	20	0.15	0.0026	1.7651	1.2777	0.45	0.0060	3.2944	1.5802	0.75	0.0025	1.5943	1.2738
20	20	0.20	0.0026	9.2626	2.1426	0.50	0.0062	8.5030	2.0649	0.80	0.0026	7.8190	2.1909
20	20	0.25	0.0027	1.8889	1.3096	0.55	0.0065	3.5249	1.5662	0.85	0.0027	2.1327	1.2909
20	20	0.30	0.0026	0.6643	0.8512	0.60	0.0061	1.3147	1.0739	0.90	0.0025	0.6443	0.8938
20	20	0.35	0.0025	0.3631	0.7359	0.65	0.0061	0.7088	0.8546	0.95	0.0026	0.4385	0.9553

Table 7. Coverage probability (CP) and half length (HL) of a 95% HPD interval of R when $\theta = 1$, $R = (0.5, 0.8)$ and $(n, m) = \{(5, 5), (5, 10), (10, 5), (10, 20), (20, 10), (20, 20)\}$.

		$R = 0.5$			$R = 0.8$					$R = 0.5$			$R = 0.8$		
n	m	R_0	CP	HL	R_0	CP	HL	n	m	R_0	CP	HL	R_0	CP	HL
5	5	0.35	0.74	0.14	0.65	0.19	0.13	10	20	0.35	0.01	0.09	0.65	0.00	0.08
		0.40	0.98	0.15	0.70	0.91	0.13			0.40	0.82	0.10	0.70	0.22	0.08
		0.45	1.00	0.16	0.75	1.00	0.12			0.45	0.99	0.10	0.75	0.98	0.08
		0.50	1.00	0.17	0.80	1.00	0.12			0.50	1.00	0.11	0.80	1.00	0.07
		0.55	1.00	0.18	0.85	0.98	0.12			0.55	0.99	0.12	0.85	0.94	0.07
		0.60	0.98	0.19	0.90	0.84	0.12			0.60	0.92	0.12	0.90	0.71	0.08
		0.65	0.92	0.20	0.95	0.72	0.13			0.65	0.8	0.13	0.95	0.67	0.09
5	10	0.35	0.53	0.12	0.65	0.00	0.12	20	10	0.35	0.02	0.09	0.65	0.00	0.08
		0.40	0.96	0.13	0.70	0.82	0.11			0.40	0.80	0.09	0.70	0.26	0.07
		0.45	0.99	0.14	0.75	0.99	0.11			0.45	0.99	0.10	0.75	0.96	0.07
		0.50	1.00	0.15	0.80	1.00	0.11			0.50	1.00	0.11	0.80	1.00	0.07
		0.55	1.00	0.16	0.85	0.98	0.11			0.55	0.99	0.12	0.85	0.94	0.08
		0.60	0.97	0.17	0.90	0.82	0.11			0.60	0.92	0.12	0.90	0.72	0.08
		0.65	0.88	0.17	0.95	0.69	0.11			0.65	0.83	0.13	0.95	0.67	0.10
10	5	0.35	0.51	0.12	0.65	0.04	0.11	20	20	0.35	0.00	0.07	0.65	0.00	0.06
		0.40	0.95	0.13	0.70	0.79	0.11			0.40	0.54	0.08	0.70	0.02	0.06
		0.45	1.00	0.14	0.75	0.99	0.11			0.45	0.98	0.08	0.75	0.93	0.06
		0.50	1.00	0.15	0.80	0.001	0.11			0.50	1.00	0.09	0.80	1.00	0.06
		0.55	0.99	0.16	0.85	0.97	0.11			0.55	0.98	0.10	0.85	0.89	0.06
		0.60	0.97	0.17	0.90	0.84	0.11			0.60	0.87	0.10	0.90	0.62	0.06
		0.65	0.90	0.18	0.95	0.72	0.12			0.65	0.71	0.11	0.95	0.62	0.08

Table 8. Survival times of two treatment groups of head and neck cancer patients.

Radio (X)	6.53, 7, 10.42, 14.48, 16.1, 22.7, 34, 41.55, 42, 45.28, 49.4, 53.62, 63, 64, 83, 84, 91, 108, 112, 129, 133, 133, 139, 140, 140, 146, 149, 154, 157, 160, 160, 165, 146, 149, 154, 157, 160, 160, 165, 173, 1146, 1417, 176, 218, 225, 241, 248, 273, 277, 297, 405, 417, 420, 440, 523, 583, 594, 1101
Radio & Chemo (Y)	12.2, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81, 43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776

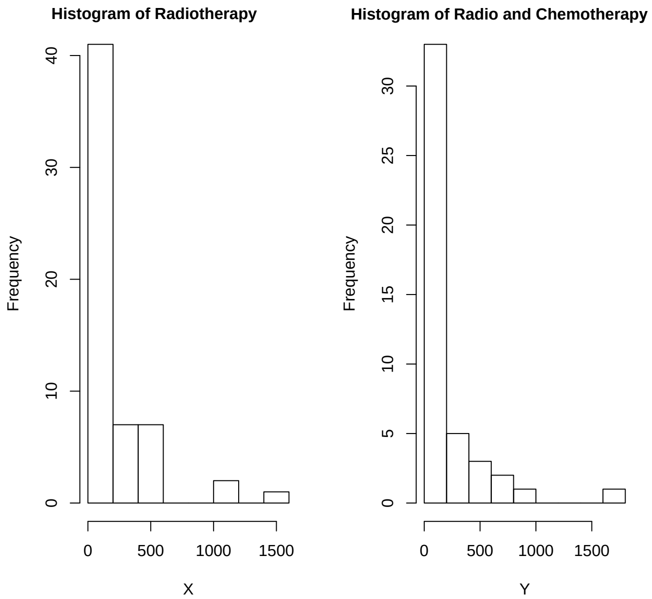

Figure 2. Histogram of radiotherapy (left) and radio and chemotherapy (right) survival data.

square error of the MLE and the relative efficiency of the Bayes and shrinkage estimators. Four different initial guess values are used $R_0 = \{0.3, 0.4, 0.5, 0.6\}$. Clearly, $\hat{R}_{BAY}$ outperforms both $\hat{R}_{MLE}$ and $\hat{R}_{SHR}$ especially when the initial guess is within 0.15 of the true value. The performance of the Bayes estimator improves for smaller samples sizes. This result agrees with findings of the simulation results.

5. Conclusion

The present study proposes a new Bayesian approach to estimate the stress strength reliability in the exponential case. The proposed method utilizes an initial guess of the reliability through an informative prior. The results of the simulation study demonstrate that the Bayesian estimator outperforms the existing shrinkage estimators as long as the initial guess is within 5–15% of the true value, regardless of the underlying distribution parameter values. The use of the Bayesian estimator is worth considering especially if available sample size is small. As generally expected, as the sample size increases, the precision of MLE estimator increases, while both Bayesian and shrinkage estimators are still affected by the prior guess.

Table 9. MSE of $\hat{R}_{MLE}$ and RE of $\hat{R}_{BAY}$ and $\hat{R}_{SHR}$ for the Real Data Example.

			Initial guess			
n	m		0.3	0.4	0.5	0.6
5	5	MSE($\hat{R}_{MLE}$)	0.037	0.037	0.038	0.037
		RE($\hat{R}_{BAY}$) = MSE($\hat{R}_{MLE}$)/MSE($\hat{R}_{BAY}$)	1.218	3.220	5.988	3.956
		RE($\hat{R}_{SHR}$) = MSE($\hat{R}_{MLE}$)/MSE($\hat{R}_{SHR}$)	1.050	1.321	1.538	1.335
10	10	MSE($\hat{R}_{MLE}$)	0.020	0.020	0.020	0.020
		RE($\hat{R}_{BAY}$)	0.714	2.162	5.945	3.392
		RE($\hat{R}_{SHR}$)	0.891	1.200	1.406	1.255
20	20	MSE($\hat{R}_{MLE}$)	0.007	0.007	0.007	0.007
		RE($\hat{R}_{BAY}$)	0.262	0.939	6.359	1.966
		RE($\hat{R}_{SHR}$)	0.801	1.019	1.702	1.158

Although the new approach is proposed for exponentially distributed data, it is, however, useful outside the current framework. It is also easy to apply in settings where the shrinkage estimation methods are not feasible. This is true because we can use our proposed prior distribution as long as the reliability R takes the same form, no matter what is the assumed distribution of the data. However, deriving a suitable weight w for the shrinkage estimator has to be done for each distribution separately, which in turn may not be easy to obtain.

Acknowledgments. We express our sincere thanks to the referees and editor Sarah Radloff for their constructive and valuable suggestions, which led to an improvement of this article.

References

Abdel-Hady, D. H. (2014). Estimation of $P(X < Y)$ for Lomax distribution in the presence of k outliers. *Int. J. Acad. Res.*, **6**, 260–267.

Awad, A. M., Azzam, M. M., and Hamdan, M. A. (1981). Some inference results on $\Pr(X < Y)$ in the bivariate exponential model. *Comm. Statist. A – Theor.*, **10**, 2515–2525.

Baklizi, A. (2008a). Estimation of $\Pr(Y < X)$ using record values in the one and two parameter exponential distributions. *Comm. Statist. A – Theor.*, **37**, 692–698.

Baklizi, A. (2008b). Likelihood and Bayesian estimation of $\Pr(Y < X)$ using lower record values from the generalized exponential distribution. *Comput. Stat. Data Anal.*, **52**, 3468–3473.

Baklizi, A. (2014). Interval estimation of the stress-strength reliability in the two-parameter exponential distribution based on records. *J. Stat. Comput. Sim.*, **84**, 2670–2679.

Baklizi, A. and Abu Dayyeh, W. (2003). Shrinkage estimation of $P(Y < X)$ in the exponential case. *Commun. Stat. – Simulat.*, **32**, 31–42.

Baklizi, A. and El-Masri, A. E. Q. (2004). Shrinkage estimation of $P(X < Y)$ in the exponential case with common location parameter. *Metrika*, **59**, 163–171.

Birnbaum, Z. W. (1956). On a use of the Mann-Whitney statistic. *In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*. University of California Press, Berkeley and Los Angeles, 933–945.

Chaturvedi, A. and Nandchahal, S. (2016). Shrinkage estimators of the reliability characteristics of a family of lifetime distributions. *Statistica*, **76**, 57.

Davarzani, N., Haghighi, F., and Parsian, A. (2009). Estimation of $P(X \leq Y)$ for a bivariate Weibull distribution. *J. Appl. Probab. Stat.*, **4** (2), 227–238.

Elfattah, A. and Marwa, O. M. (2008). Estimating of $P(Y < X)$ in the exponential case based on censored samples. *arXiv preprint arXiv:0801.0922*.

Enis, P. and Geisser, S. (1971). Estimation of the probability that $Y < X$. *J. Amer. Statist. Assoc.*, **66**, 162–168.

Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Appl. Stat.*, **41**, 337–348.

Govindarazulu, Z. (1967). Two-sided confidence limits for $P(X < Y)$ based on normal samples of X and Y. *Sankhyā. Ser. B*, **29**, 35–40.

Jeffreys, H. (1961). *Theory of Probability*. Third edition. Oxford University Press, Oxford.

Khamnei, H. J. (2013). Reliability for Lindley distribution with an outlier. *Bull. Math. Sci. Appl.*, **2**, 23–28.

Kundu, D. and Gupta, R. D. (2006). Estimation of $P[Y < X]$ for Weibull distributions. *IEEE Trans. Rel.*, **55**, 270–280.

McCool, J. I. (1991). Inference on $P\{Y < X\}$ in the Weibull case. *Comm. Statist. Simulation Comput.*, **20** (1), 129–148.

Montazer Haghighi, A. and Shayib, M. A. (2009). Shrinkage estimators for calculating reliability, Weibull case. *J. Appl. Statist. Sci.*, **17**, 219–233.

Muttlak, H. A., Abu-Dayyeh, W. A., Saleh, M. F., and Al-Sawi, E. (2010). Estimating $P(Y < X)$ using ranked set sampling in case of the exponential distribution. *Comm. Statist. A – Theor.*, **39**, 1855–1868.

Nadarajah, S. (2004a). Reliability for Laplace distributions. *Math. Probl. Eng.*, **2004**, 169–183.

Nadarajah, S. (2004b). Reliability for logistic distributions. *Elektronnoe Modelirovanie*, **26**, 65–82.

Nadarajah, S. (2005a). Reliability for some bivariate beta distributions. *Math. Probl. Eng.*, **2005**, 101–111.

Nadarajah, S. (2005b). Reliability for some bivariate gamma distributions. *Math. Probl. Eng.*, **2005**, 151–163.

Saleh, A. M. E. (2006). *Theory of Preliminary Test and Stein-Type Estimation with Applications*. John Wiley & Sons, New York.

Samawi, H. M., Helu, A., Rochani, H. D., Yin, J., and Linder, D. (2016). Estimation of $P(X > Y)$ when X and Y are dependent random variables using different bivariate sampling schemes. *Commun. Stat. Appl. Methods*, **23**, 385–397.

Singh, S. K., Singh, U., Yadav, A., and Viswkarma, P. K. (2015). On the estimation of stress strength reliability parameter of inverted exponential distribution. *Int. J. Sci. World*, **3**, 98–112.

Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean. *J. Amer. Statist. Assoc.*, **63**, 113–122.

Tong, H. (1974). A note on the estimation of Pr $\{Y < X\}$ in the exponential case. *Technometrics*,

16, 625.

Ventura, L. and Racugno, W. (2011). Recent advances on Bayesian inference for $P(X < Y)$. *Bayesian Anal.*, **6**, 411–428.

Woodward, W. A. and Kelley, G. D. (1977). Minimum variance unbiased estimation of $P[Y < X]$ in the normal case. *Technometrics*, **19** (1), 95–98.

Manuscript received 2017-07-19, revised 2018-01-11, accepted 2018-02-02.