STATISTICA, anno LXXVIII, n. 4, 2018

ESTIMATION OF STRESS-STRENGTH RELIABILITY FOR THE PARETO DISTRIBUTION BASED ON UPPER RECORD VALUES

Rahmath Manzil Juvairiyya [1]
Department of Statistics, Farook College, Calicut, India
Parameshwaranpillai Anilkumar
Department of Statistics, Farook College, Calicut, India

1. INTRODUCTION

The problem of estimating stress-strength reliability has many applications in a variety of fields. In the stress-strength model with X as stress applied and Y as the strength of the system, $R = P(X > Y)$ is called the stress-strength reliability and it measures the chance that the system fails. We want to estimate the stress-strength reliability $R = P(X > Y)$ using record data. This probability arises in life testing experiments when X and Y represent the lifetimes of two devices and it gives the probability that the device with lifetime Y fails before the other. If X and Y represent the life lengths of a product with same guarantee period produced by two companies, then $P(X > Y)$ represents the probability that one is better than the other. This probability can be considered as a general measure of difference between populations. Thus the probability $P(X > Y)$, even though it is called stress-strength reliability has applications (Kotz *et al.*, 2003) beyond evaluation of the actual stress-strength reliability.

Records were unexplored until Chandler (1952) introduced and studied some properties of record values. Since then abundant literature was devoted to the study of records (Arnold *et al.*, 1998). Record values and associated statistics have an important role in many real life applications involving data relating to meteorology, hydrology, sports and life tests. In industry and reliability many products may fail under stress. For example, a wooden beam breaks when sufficient perpendicular force is applied to it, a battery dies under stress of time, an electronic component ceases to function in an environment of too high temperature. In such experiments for getting the precise failure point, measurements may be made sequentially and only values larger (or smaller) than all previous ones are recorded. Data of this type are called record data. Let $X_1, X_2, \ldots, X_n$

[1] Corresponding Author. E-mail: juvairiyya@farookcollege.ac.in

be a sequence of i.i.d random variables with an absolutely continuous cumulative distribution function (cdf) F(x) and probability density function (pdf) f(x). An observation X_j is called an upper record if its value exceeds all previous observations. That is X_j is an upper record if $X_j > X_i$ for every $i < j$.

Estimation of stress-strength reliability based on record data was considered by Baklizi (2008) and Asgharzadeh *et al.* (2017) for the one and two parameter exponential distribution, and for the generalized exponential distribution respectively. Baklizi (2012) also studied inference on stress-strength reliability in the two-parameter Weibull model based on records. Estimation of stress-strength parameter using record values from proportional hazard model was considered by Basirat *et al.* (2016) and for the two-parameter bathtub shaped life time distribution based on upper record values was presented by Tarvirdizade and Ahmadpour (2016). Rezaei *et al.* (2010) studied the estimation of $R = P(X < Y)$ when X and Y are two independent generalized Pareto random variables with common scale parameters and different shape parameters. In this paper, we consider the estimation of $R = P(X > Y)$ based on upper records when X and Y are independent random variables having Pareto distributions with the same scale parameters and different shape parameters.

The organization of this paper is as follows. In Section 2, we discuss the likelihood inference of the stress-strength parameter. Section 3 describes the Bayesian inference. In Section 4, a simulation study is conducted to investigate and compare the performance of point estimators presented in this paper. Section 5 presents a real data analysis for the illustration of the proposed estimation methods . Finally some conclusions are given in Section 6.

2. LIKELIHOOD INFERENCE

The Pareto distribution is quite popular in describing the distribution of wealth in a given population. Recently, the generalized Pareto distribution was considered by Rezaei *et al.* (2010) for estimation of stress-strength reliability. The pdf and the cdf of the Pareto distribution with parameters α and β (both positive) are given by

$$\begin{aligned} f(x) &= \frac{\beta\alpha^{\beta}}{x^{\beta+1}}, \quad x \geq \alpha, \alpha > 0, \beta > 0, \\ F(x) &= 1-(\frac{\alpha}{x})^{\beta}. \end{aligned} \tag{1}$$

Let X and Y be two independent random variables from the Pareto distribution with parameters (α, β_1) and (α, β_2) respectively. Then using (1)

$$R = P(X > Y) = \int_{\alpha}^{\infty}\int_{y}^{\infty} \frac{\beta_1\alpha^{\beta_1}}{x^{\beta_1+1}}\frac{\beta_2\alpha^{\beta_2}}{y^{\beta_2+1}} dx dy = \frac{\beta_2}{\beta_1+\beta_2} .$$

We are interested in estimating R based on upper record values on both variables. Let $\underset{\sim}{r} = (r_1, r_2, ..., r_n)$ be a set of upper records from distribution of X with pdf f and cdf F and let $\underset{\sim}{s} = (s_1, s_2, ..., s_m)$ be an independent set of upper records from distribution of Y with pdf g and cdf G. The likelihood functions are given by

$$L(\alpha, \beta_1|\underset{\sim}{r}) = f(r_n)\prod_{i=1}^{n-1}\left(\frac{f(r_i)}{1-F(r_i)}\right), 0 < r_1 < r_2 < ..... < r_n < \infty, \quad (2)$$

$$L(\alpha, \beta_2|\underset{\sim}{s}) = g(s_m)\prod_{i=1}^{m-1}\left(\frac{g(s_i)}{1-G(s_i)}\right), 0 < s_1 < s_2 < ..... < s_m < \infty. \quad (3)$$

Substituting f, F, g and G, the joint likelihood and the joint log-likelihood are respectively given by

$$\begin{aligned} L(\alpha, \beta_1, \beta_2, \underset{\sim}{r}, \underset{\sim}{s}) &= \frac{\beta_1^n \alpha^{\beta_1}}{r_n^{\beta_1+1}}\prod_{i=1}^{n-1} r_i^{-1}\frac{\beta_2^m \alpha^{\beta_2}}{s_m^{\beta_2+1}}\prod_{i=1}^{m-1} s_i^{-1}, \\ l(\alpha, \beta_1, \beta_2, \underset{\sim}{r}, \underset{\sim}{s}) &= n\log\beta_1 + \beta_1\log\alpha - \sum_{i=1}^{n-1}\log r_i - (\beta_1+1)\log r_n \\ &+ m\log\beta_2 + \beta_2\log\alpha - \sum_{i=1}^{m-1}\log s_i - (\beta_2+1)\log s_m. \end{aligned} \quad (4)$$

$$\frac{\delta l}{\delta\beta_1} = 0 \Rightarrow \frac{n}{\beta_1} = \log\left(\frac{r_n}{\alpha}\right). \quad (5)$$

$$\frac{\delta l}{\delta\beta_2} = 0 \Rightarrow \frac{m}{\beta_2} = \log\left(\frac{s_m}{\alpha}\right). \quad (6)$$

When α is known,

$$\hat{\beta}_1 = \frac{n}{\log(r_n/\alpha)}, \quad \hat{\beta}_2 = \frac{m}{\log(s_m/\alpha)}. \quad (7)$$

Then the MLE of R is given by $\hat{R} = \frac{\hat{\beta}_2}{\hat{\beta}_1+\hat{\beta}_2}$.

REMARK 1. *While estimating shape parameters using record values it is assumed that the scale parameters are common and known. The procedure is same if we assume that the scale parameters α_1 and α_2 are assumed to be different but known. On the other hand if α_1 and α_2 are assumed to be unknown we have to estimate them from the upper record data. But due to the nature of the scale parameter ($x \geq \alpha$) in the model, α cannot be estimated consistently from upper record values alone. Hence the assumption that the scale parameters are known cannot be relaxed. In the data set analysed we are using a data with same scale parameter. Intuitively this is an appropriate situation when we are comparing the performance of two products with same warranty period.*

Now we shall study the distribution of $\hat{\beta}_1$ and $\hat{\beta}_2$. Consider first $\hat{\beta}_1 = z_1 = \frac{n}{\log(r_n/\alpha)}$. We know that the pdf of R_n is given by

$$\begin{aligned} f_{R_n}(r_n) &= \frac{1}{(n-1)!} f(r_n)[-\log(1-F(r_n))]^{n-1} \\ &= \frac{1}{(n-1)!} \frac{\beta_1^n \alpha^{\beta_1}}{r_n^{\beta_1+1}} [\log(r_n/\alpha)]^{n-1} \quad r_n > \alpha. \end{aligned} \tag{8}$$

Therefore the pdf of $z_1 = \hat{\beta}_1$ is given by

$$f_{Z_1}(z_1) = \frac{(n\beta_1)^n}{(n-1)!} \frac{\exp^{\frac{-n\beta_1}{z_1}}}{z_1^{n+1}}, \quad z_1 > 0. \tag{9}$$

Here

$$Z_1 \sim \text{Inv-Gamma}(n, n\beta_1).$$

Similarly

$$Z_2 \sim \text{Inv-Gamma}(m, m\beta_2).$$

Therefore we can find the pdf of

$$\begin{aligned} \hat{R} &= \frac{\hat{\beta}_2}{\hat{\beta}_1 + \hat{\beta}_2} \\ &= \frac{Z_2}{Z_1 + Z_2} = \frac{1}{1 + Z_1/Z_2}. \end{aligned} \tag{10}$$

Consider $\frac{Z_1}{Z_2}$. By the properties of the inverse-gamma distribution and its relation with the gamma distribution, we have $\frac{n\beta_1}{z_1} \sim \text{Gamma}(n, 1)$ and $\frac{m\beta_2}{z_2} \sim \text{Gamma}(m, 1)$. Hence $\frac{2n\beta_1}{z_1} \sim \chi^2_{2n}$ and $\frac{2m\beta_2}{z_2} \sim \chi^2_{2m}$. Since the two random variables are independent we have

$$\frac{\left(\frac{2m\beta_2}{2mZ_2}\right)}{\left(\frac{2n\beta_1}{2nZ_1}\right)} = \frac{\beta_2 Z_1}{\beta_1 Z_2} = \frac{R}{1-R} \frac{\hat{\beta}_1}{\hat{\beta}_2} \sim F(2m, 2n).$$

This can be used to construct the following $100(1-\alpha)\%$ confidence interval for R

$$\left(\left(1 + \frac{\hat{\beta}_1}{\hat{\beta}_2 F_{\frac{\alpha}{2},2m,2n}}\right)^{-1}, \left(1 + \frac{\hat{\beta}_1}{\hat{\beta}_2 F_{1-\frac{\alpha}{2},2m,2n}}\right)^{-1} \right). \tag{11}$$

3. Bayesian inference

It is also interesting how supplementary information other than upper record values available can be incorporated. A convenient vehicle is the method of Bayesian inference. We have considered both informative and non-informative priors. The sampling distribution of β_1 and β_2 and hence R will instigate the use of an appropriate conjugate prior. We assume the conjugate family of prior distribution to be gamma family of distributions. So,

$$\pi(\beta_1) = \frac{1}{\Gamma(\gamma_1)}\theta_1^{\gamma_1}\beta_1^{\gamma_1-1}\exp^{-\theta_1\beta_1} \quad \beta_1>0;\theta_1,\gamma_1>0, \tag{12}$$

$$\pi(\beta_2) = \frac{1}{\Gamma(\gamma_2)}\theta_2^{\gamma_2}\beta_2^{\gamma_2-1}\exp^{-\theta_2\beta_2} \quad \beta_2>0;\theta_2,\gamma_2>0. \tag{13}$$

Using the priors and the likelihood functions, the posterior distributions of β_1 and β_2 are obtained as

$$\begin{aligned}(\beta_1|\underset{\sim}{r}) &\sim \text{Gamma}\Big(n+\gamma_1\,,\,\theta_1+\log\Big(\frac{r_n}{\alpha}\Big)\Big),\\ (\beta_2|\underset{\sim}{s}) &\sim \text{Gamma}\Big(m+\gamma_2\,,\,\theta_2+\log\Big(\frac{s_m}{\alpha}\Big)\Big).\end{aligned} \tag{14}$$

Since β_1 and β_2 are independent, using standard transformation techniques and after some manipulations the posterior pdf of R is given by

$$f_R(r)=C\frac{(1-r)^{n+\gamma_1-1}r^{m+\gamma_2-1}}{\left[(1-r)\left(\theta_1+\log\frac{r_n}{\alpha}\right)+r\left(\theta_2+\log\frac{s_m}{\alpha}\right)\right]^{n+m+\gamma_1+\gamma_2}}\ ,\quad 0<r<1$$

where

$$C=\frac{\Gamma(n+m+\gamma_1+\gamma_2)}{\Gamma(n+\gamma_1)\Gamma(m+\gamma_2)}\left(\theta_1+\log\left(\frac{r_n}{\alpha}\right)\right)^{n+\gamma_1}\left(\theta_2+\log\left(\frac{s_m}{\alpha}\right)\right)^{m+\gamma_2}.$$

Under squared error loss function, the Bayes estimator of R is the expected value of R. This expected value contains an integral which is not obtainable in a simple closed form. Therefore using the approximate method of Lindley (1980), we can find the approximate Bayes estimator R_B relative to square error loss function. By the approximate method of Lindley, the Bayes estimator for $u(\theta)$ for a prior $v(\theta)$ is given by

$$\begin{aligned}\frac{\int_\theta u(\theta)v(\theta)\exp^{L(\theta)}d\theta}{\int_\theta v(\theta)\exp^{L(\theta)}d\theta} &= [u^*+\frac{1}{2}(u_{11}\sigma_{11}+u_{22}\sigma_{22})+\rho_1u_1\sigma_{11}+\rho_2u_2\sigma_{22}\\ &+ \frac{1}{2}[\sigma_{11}\sigma_{22}(u_1L_{12}+u_2L_{21})+u_1\sigma_{11}^2L_{30}\\ &+ u_2\sigma_{22}^2L_{03}]+O(\frac{1}{n^2})]_{at\,\hat{\theta}}\,.\end{aligned} \tag{15}$$

where $u(\theta) = \frac{\beta_2}{\beta_1+\beta_2}$; ρ= logarithm of joint priors=$\log C + (\gamma_1 - 1)\log\beta_1 - \theta_1\beta_1 + (\gamma_2 - 1)\log\beta_2 - \theta_2\beta_2$; u^* is the MLE of $u(\theta)$and $L(\theta)$ is the logarithm of likelihood function. C is independent of β_1 and β_2. Further

$$u_1 = \frac{\delta u}{\delta \beta_1} = \frac{-\beta_2}{(\beta_1+\beta_2)^2};\ u_2 = \frac{\delta u}{\delta \beta_2} = \frac{\beta_1}{(\beta_1+\beta_2)^2}.$$

$$u_{11} = \frac{\delta^2 u}{\delta \beta_1^2} = \frac{2\beta_2}{(\beta_1+\beta_2)^3};\ u_{22} = \frac{\delta^2 u}{\delta \beta_2^2} = \frac{-2\beta_1}{(\beta_1+\beta_2)^3}.$$

$$\rho_1 = \frac{\gamma_1 - 1}{\beta_1} - \theta_1;\ \rho_2 = \frac{\gamma_2 - 1}{\beta_2} - \theta_2.$$

$$\sigma = \left[-L_{ij}\right]^{-1} \quad \text{where} \quad L_{ij} = \left[\frac{\delta^2 L}{\delta \beta_i \delta \beta_j}\right].$$

$$\sigma = \begin{bmatrix} \frac{\beta_1^2}{n} & 0 \\ 0 & \frac{\beta_2^2}{m} \end{bmatrix};\ L_{30} = \frac{\delta^3 L}{\delta \beta_1^3} = \frac{2n}{\beta_1^3};\ L_{03} = \frac{\delta^3 L}{\delta \beta_2^3} = \frac{2m}{\beta_2^3}.$$

Substituting in (15) we get the Bayes estimator as

$$\bar{R_B} = \hat{R} + \hat{R}(1-\hat{R}) \left[\frac{\left(1 - \hat{R} - \gamma_1 + \hat{\beta}_1\theta_1\right)}{n} + \frac{\left(\gamma_2 - \hat{R} - \hat{\beta}_2\theta_2\right)}{m} \right]. \tag{16}$$

Furthermore, it follows from (14) that $2(\theta_1 + \log\frac{r_n}{\alpha})(\beta_1 | \underset{\sim}{r}) \sim \chi^2_{2(n+\gamma_1)}$ and $2(\theta_2 + \log\frac{s_m}{\alpha})(\beta_2 | \underset{\sim}{s}) \sim \chi^2_{2(m+\gamma_2)}$. It follows that $\pi(R | \underset{\sim}{r}, \underset{\sim}{s})$, the posterior distribution of R is equal to that of $(1 + AW)^{-1}$ where $W \sim F_{2(n+\gamma_1),2(m+\gamma_2)}$ and $A = \frac{(n+\gamma_1)(\theta_2+\log\frac{s_m}{\alpha})}{(m+\gamma_2)(\theta_1+\log\frac{r_n}{\alpha})}$. Therefore a Bayesian $(1-\alpha)$% confidence interval for R is given by

$$\left(\left(AF_{1-\frac{\alpha}{2},2(n+\gamma_1),2(m+\gamma_2)} + 1\right)^{-1}, \left(AF_{\frac{\alpha}{2},2(n+\gamma_1),2(m+\gamma_2)} + 1\right)^{-1}\right). \tag{17}$$

When we are ignorant about the parameter we use a non-informative prior. Here we take Jeffreys' non informative prior. We assume

$$\pi(\beta_1) \propto \frac{1}{\beta_1} \quad \beta_1 > 0, \tag{18}$$

$$\pi(\beta_2) \propto \frac{1}{\beta_2} \quad \beta_2 > 0. \tag{19}$$

Using the priors and the likelihood functions, the posterior distributions of β_1 and β_2 are obtained as

$$(\beta_1|\underset{\sim}{r}) \sim \text{Gamma}\left(n\ ,\ \log\left(\frac{r_n}{\alpha}\right)\right).$$
$$(\beta_2|\underset{\sim}{s}) \sim \text{Gamma}\left(m\ ,\ \log\left(\frac{s_m}{\alpha}\right)\right). \tag{20}$$

Since β_1 and β_2 are independent, then, using standard transformation techniques and after some manipulations the posterior pdf of R is given by

$$f_R(r) = C\frac{(1-r)^{n-1} r^{m-1}}{\left[(1-r)\log\left(\frac{r_n}{\alpha}\right) + r\log\left(\frac{s_m}{\alpha}\right)\right]^{n+m}} \quad 0 < r < 1$$

where

$$C = \frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m)}\left(\log\left(\frac{r_n}{\alpha}\right)\right)^n \left(\log\left(\frac{s_m}{\alpha}\right)\right)^m.$$

Under squared error loss function, the Bayes estimator of R is the expected value of R. This expected value contains an integral which is not obtainable in a simple closed form. Therefore using the approximate method of Lindley (1980), we can find the approximate Bayes estimator R_B relative to square error loss function. By the approximate method of Lindley, the Bayes estimator for $u(\theta)$ for a prior $v(\theta)$ and a likelihood function is given by

$$\begin{aligned}\frac{\int_\theta u(\theta)v(\theta)\exp^{L(\theta)}\,d\theta}{\int_\theta v(\theta)\exp^{L(\theta)}\,d\theta} &= [u^* + \frac{1}{2}(u_{11}\sigma_{11} + u_{22}\sigma_{22}) + \rho_1 u_1\sigma_{11} + \rho_2 u_2\sigma_{22} \\ &+ \frac{1}{2}[\sigma_{11}\sigma_{22}(u_1 L_{12} + u_2 L_{21}) + u_1\sigma_{11}^2 L_{30} \\ &+ u_2\sigma_{22}^2 L_{03}] + O(\frac{1}{n^2})]_{at\hat{\theta}}\,,\end{aligned} \tag{21}$$

where

$$u(\theta) = \frac{\beta_2}{\beta_1+\beta_2};\ \ v(\theta) = \frac{1}{\beta_1\beta_2};\ \ \rho = \log v(\theta) = \log\left(\frac{1}{\beta_1\beta_2}\right).$$

The MLE of $u(\theta)$ is u^* and the logarithm of the likelihood function is $L(\theta)$.
Further

$$u_1 = \frac{\delta u}{\delta\beta_1} = \frac{-\beta_2}{(\beta_1+\beta_2)^2};\ \ u_2 = \frac{\delta u}{\delta\beta_2} = \frac{\beta_1}{(\beta_1+\beta_2)^2}.$$
$$u_{11} = \frac{\delta^2 u}{\delta\beta_1^2} = \frac{2\beta_2}{(\beta_1+\beta_2)^3};\ \ u_{22} = \frac{\delta^2 u}{\delta\beta_2^2} = \frac{-2\beta_1}{(\beta_1+\beta_2)^3}.$$
$$\rho_1 = \frac{\delta\rho}{\delta\beta_1} = \frac{-1}{\beta_1};\ \ \rho_2 = \frac{\delta\rho}{\delta\beta_2} = \frac{-1}{\beta_2}.$$

$$\sigma = \left[-L_{ij}\right]^{-1} \quad \text{where} \quad L_{ij} = \left[\frac{\delta^2 L}{\delta\beta_i \delta\beta_j}\right].$$

$$\sigma = \begin{bmatrix} \frac{\beta_1^2}{n} & 0 \\ 0 & \frac{\beta_2^2}{m} \end{bmatrix}; \; L_{30} = \frac{\delta^3 L}{\delta\beta_1^3} = \frac{2n}{\beta_1^3}; \; L_{03} = \frac{\delta^3 L}{\delta\beta_2^3} = \frac{2m}{\beta_2^3}.$$

Substituting in (21) we get the Bayes estimator as

$$\bar{R_B} = \hat{R} + \hat{R}(1-\hat{R})\left[\frac{\left(1-\hat{R}\right)}{n} - \frac{\hat{R}}{m}\right]. \tag{22}$$

Furthermore, it follows from (20) that $2\log(\frac{r_n}{\alpha})(\beta_1|\underset{\sim}{r}) \sim \chi^2_{2n}$ and $2\log(\frac{s_m}{\alpha})(\beta_2|\underset{\sim}{s}) \sim \chi^2_{2m}$. It follows that $\pi(R|\underset{\sim}{r},\underset{\sim}{s})$, the posterior distribution of R is equal to that of $(1+AW)^{-1}$ where $W \sim F_{2n,2m}$ and $A = \frac{n\log(\frac{s_m}{\alpha})}{m\log(\frac{r_n}{\alpha})}$. Therefore a Bayesian $(1-\alpha)\%$ credible interval for R is given by

$$\left(\left(AF_{1-\frac{\alpha}{2},2n,2m}+1\right)^{-1}, \left(AF_{\frac{\alpha}{2},2n,2m}+1\right)^{-1}\right), \tag{23}$$

which happens to be the same as the confidence interval based on MLE.

4. A SIMULATION STUDY

In this section, a Monte Carlo simulation study is conducted to investigate and compare the performance of point estimators and confidence intervals presented in this paper. The performance of MLE and Bayes estimators is compared in terms of their biases and mean squared errors (MSE). We consider only one case, when the scale parameter α is known. We use the parameter values $(\beta_1, \beta_2) = (4,1),(2,2),(1,3),(1,9)$. Therefore $R_{\text{exact}} = 0.2, 0.5, 0.75, 0.9$. To compute the Bayes estimators we consider two methods: one with conjugate prior and the other with Jeffreys' invariant prior. We report all the results based on 2000 replications.

When the scale parameter is known we obtain the average estimates, biases and MSEs of the MLE and the approximate Bayes estimator of R. We also compute the expected length for the confidence intervals obtained by using the ML method, the Jeffreys' prior Bayes method and the conjugate prior Bayes method. The results are reported in Table 1 and Table 2.

From the simulation results, it is observed that as the sample size (n,m) increases the biases and the MSEs decrease. Thus the consistency properties of all the methods are verified. It is observed that the bias of the estimators become negative for values of R larger than 0.5. It is also observed that the intervals based on all methods are maximized when R=0.5 and they becomes shorter and shorter as we move away to smaller and larger values. Increasing the sample size on either variable also results in shorter

TABLE 1
Average estimates (AVR), bias and MSE of the estimators of R.

(n,m)	R	MLE			Bayes non-inform			Bayes conjugate		
		AVR	Bias	MSE	AVR	Bias	MSE	AVR	Bias	MSE
(4,4)	0.2	0.225	0.025	0.016	0.244	0.044	0.017	0.303	0.103	0.021
(4,7)		0.212	0.012	0.011	0.236	0.036	0.013	0.299	0.099	0.019
(4,10)		0.208	0.008	0.009	0.233	0.033	0.011	0.297	0.097	0.017
(7,4)		0.229	0.029	0.013	0.236	0.036	0.013	0.264	0.064	0.013
(7,7)		0.216	0.016	0.008	0.228	0.028	0.009	0.262	0.062	0.011
(7,10)		0.210	0.010	0.007	0.224	0.024	0.007	0.259	0.059	0.009
(10,4)		0.232	0.032	0.013	0.234	0.034	0.012	0.249	0.049	0.009
(10,7)		0.218	0.018	0.008	0.225	0.025	0.008	0.247	0.047	0.008
(10,10)		0.213	0.013	0.006	0.221	0.021	0.006	0.245	0.045	0.007
(4,4)	0.5	0.502	0.002	0.028	0.502	0.002	0.023	0.502	0.002	0.011
(4,7)		0.491	-0.009	0.023	0.504	0.004	0.019	0.526	0.026	0.012
(4,10)		0.486	-0.014	0.021	0.504	0.004	0.018	0.534	0.034	0.011
(7,4)		0.515	0.015	0.023	0.501	0.001	0.019	0.478	-0.022	0.012
(7,7)		0.503	0.003	0.017	0.503	0.003	0.015	0.502	0.002	0.011
(7,10)		0.498	-0.002	0.014	0.503	0.003	0.013	0.510	0.010	0.009
(10,4)		0.521	0.021	0.021	0.502	0.002	0.018	0.472	-0.028	0.011
(10,7)		0.509	0.009	0.014	0.503	0.003	0.013	0.495	-0.005	0.009
(10,10)		0.505	0.005	0.012	0.504	0.004	0.011	0.504	0.004	0.009
(4,4)	0.75	0.729	-0.020	0.019	0.712	-0.038	0.018	0.668	-0.083	0.018
(4,7)		0.725	-0.026	0.016	0.720	-0.030	0.015	0.701	-0.049	0.012
(4,10)		0.723	-0.028	0.015	0.724	-0.026	0.013	0.714	-0.036	0.010
(7,4)		0.743	-0.007	0.014	0.719	-0.030	0.015	0.669	-0.082	0.017
(7,7)		0.739	-0.011	0.010	0.728	-0.023	0.011	0.701	-0.049	0.011
(7,10)		0.737	-0.013	0.009	0.731	-0.019	0.009	0.714	-0.036	0.009
(10,4)		0.751	0.001	0.012	0.724	-0.026	0.013	0.672	-0.078	0.015
(10,7)		0.746	-0.004	0.008	0.732	-0.018	0.009	0.704	-0.046	0.009
(10,10)		0.744	-0.006	0.007	0.736	-0.014	0.007	0.717	-0.033	0.007
(4,4)	0.9	0.882	-0.018	0.006	0.865	-0.035	0.008	0.766	-0.134	0.027
(4,7)		0.880	-0.020	0.005	0.872	-0.028	0.006	0.819	-0.080	0.012
(4,10)		0.879	-0.020	0.005	0.875	-0.024	0.005	0.841	-0.059	0.008
(7,4)		0.891	-0.001	0.004	0.873	-0.027	0.006	0.775	-0.125	0.022
(7,7)		0.889	-0.010	0.003	0.879	-0.020	0.004	0.826	-0.074	0.009
(7,10)		0.889	-0.010	0.003	0.883	-0.017	0.003	0.846	-0.054	0.006
(10,4)		0.895	-0.005	0.003	0.877	-0.023	0.005	0.781	-0.119	0.019
(10,7)		0.895	-0.005	0.002	0.884	-0.016	0.003	0.830	-0.070	0.008
(10,10)		0.894	-0.006	0.002	0.887	-0.013	0.002	0.849	-0.050	0.005

TABLE 2
Expected length (EL) for the confidence intervals.

(n,m)	R	MLE		Bayes non-inform		Bayes conjugate	
		95% EL	90% EL	95% EL	90% EL	95% EL	90% EL
(4,4)	0.2	0.457	0.384	0.457	0.384	0.489	0.414
(4,7)		0.418	0.348	0.418	0.348	0.451	0.379
(4,10)		0.400	0.332	0.400	0.332	0.433	0.363
(7,4)		0.398	0.336	0.398	0.336	0.412	0.349
(7,7)		0.346	0.290	0.346	0.290	0.365	0.307
(7,10)		0.323	0.269	0.323	0.269	0.342	0.287
(10,4)		0.373	0.316	0.373	0.316	0.379	0.321
(10,7)		0.317	0.266	0.317	0.266	0.328	0.276
(10,10)		0.290	0.243	0.290	0.243	0.303	0.254
(4,4)	0.5	0.585	0.503	0.585	0.503	0.574	0.494
(4,7)		0.535	0.458	0.535	0.458	0.523	0.448
(4,10)		0.512	0.438	0.512	0.438	0.499	0.427
(7,4)		0.535	0.458	0.535	0.458	0.523	0.449
(7,7)		0.470	0.401	0.470	0.401	0.462	0.394
(7,10)		0.441	0.375	0.441	0.375	0.433	0.369
(10,4)		0.513	0.439	0.513	0.439	0.500	0.428
(10,7)		0.441	0.376	0.441	0.376	0.433	0.369
(10,10)		0.406	0.345	0.406	0.345	0.399	0.339
(4,4)	0.75	0.499	0.423	0.499	0.423	0.513	0.436
(4,7)		0.439	0.373	0.439	0.373	0.443	0.376
(4,10)		0.413	0.352	0.413	0.352	0.411	0.349
(7,4)		0.457	0.383	0.457	0.383	0.473	0.399
(7,7)		0.385	0.324	0.385	0.324	0.393	0.332
(7,10)		0.353	0.298	0.353	0.298	0.358	0.302
(10,4)		0.438	0.366	0.438	0.366	0.453	0.382
(10,7)		0.359	0.301	0.359	0.301	0.369	0.310
(10,10)		0.324	0.272	0.324	0.272	0.329	0.277
(4,4)	0.9	0.314	0.256	0.314	0.256	0.416	0.347
(4,7)		0.259	0.215	0.259	0.215	0.317	0.265
(4,10)		0.238	0.198	0.238	0.198	0.275	0.229
(7,4)		0.279	0.226	0.279	0.226	0.381	0.315
(7,7)		0.219	0.179	0.219	0.179	0.278	0.230
(7,10)		0.195	0.161	0.195	0.161	0.234	0.195
(10,4)		0.263	0.212	0.263	0.212	0.363	0.299
(10,7)		0.200	0.164	0.201	0.164	0.258	0.213
(10,10)		0.175	0.144	0.175	0.144	0.213	0.177

intervals. Except for $R = 0.5$ expected length for non-informative prior is shorter than that for conjugate prior.

5. Real data analysis

In this section we analyze a real dataset to illustrate the use of our proposed estimation methods. The data from Crowder (2000) give the lifetimes of the steel specimens tested at two different stress levels.
Dataset 1. (38.5 stress level): $60, 51, 83, 140, 109, 106, 119, 76, 68, 67$.
Dataset 2. (38 stress level): $100, 90, 59, 80, 128, 117, 177, 98, 158, 107$.
We fit a Pareto distribution to the two datasets separately. The estimated parameters (based on ML methods), Kolmogorov-Smirinov (K-S) distances between the fitted and the empirical distribution functions and the corresponding *p*-values are presented in Table 3.

TABLE 3
Results of the real data analysis.

Data set	Scale parameter	Shape parameter	K-S distance	P-value
1	51	2.0173	0.2233	0.6247
2	51	1.3592	0.3577	0.119

From the table it is clear that the Pareto distribution with common scale parameter fits quite well to both data. For the above data we observe the upper record values as follows

$$\underset{\sim}{r} : 60, 83, 140 \quad \underset{\sim}{s} : 100, 128, 177.$$

Based on these record values, we take $\alpha = 51$, We obtain the MLE of β_1 and β_2 from (7) as 3.9611 and 3.2146 respectively. Therefore the MLE of R becomes $\hat{R} = 0.4479$. The corresponding 95% confidence interval based on (11) is equal to $(0.1224, 0.8253)$. In the Bayesian inference, for the first estimator we take the values of the hyperparameters as $\gamma_1 = \gamma_2 = \theta_1 = \theta_2 = 0.5$. Then we obtain the Bayes estimator $\bar{R}_1 = 0.4775$. Also the 95% credible interval from (17) is (0.1295, 0.8121). Using a non-informative prior, the Bayes estimator $\bar{R}_2 = 0.4544$. The corresponding 95% credible interval from (23) becomes (0.1224, 0.8253).

6. Conclusion

This paper considers the estimation of stress-strength reliability $R = P(X > Y)$ based on upper record values where X and Y are independent random variables from the Pareto distribution with same scale parameter but different shape parameters. The results for estimation of R by maximum likelihood estimation and the Bayesian approach are reported when the scale parameter is known. From the simulation results, it is observed that as the sample size (n,m) increases the biases and the MSEs decrease. Thus the consistency properties of all the methods are verified. It is observed that the bias of the estimators become negative for values of R larger than 0.5. It is also observed that the interval based on MLE is maximized when $R = 0.5$ and it becomes shorter and shorter as we move away to smaller and larger values.Increasing the sample size on either variable also results in shorter intervals. Except for $R = 0.5$, the expected length for non-informative prior is shorter than that for conjugate prior.

ACKNOWLEDGEMENTS

The authors express their sincere thanks to the Editor and the Referees for their constructive criticism and excellent suggestions which led to a considerable improvement in the presentation of the paper.

REFERENCES

B. C. ARNOLD, N. BALAKRISHNAN, H. N. NAGARAJA (1998). *Records*. John Wiley & Sons, Inc., New York.

A. ASGHARZADEH, R. VALIOLLAHI, M. Z. RAQAB (2017). *Estimation of Pr(Y < X) for the two-parameter generalized exponential records*. Communications in Statistics - Simulation and Computation, 46, no. 1, pp. 379–394.

A. BAKLIZI (2008). *Estimation of Pr(X < Y) using record values in the one and two parameter exponential distributions*. Communications in Statistics - Theory and Methods, 37, no. 5, pp. 692–698.

A. BAKLIZI (2012). *Inference on Pr(X<Y) in the two-parameter Weibull model based on records*. ISRN Probability and Statistics, 2012, pp. 1–11.

M. BASIRAT, S. BARATPOUR, J. AHMADI (2016). *On estimation of stress–strength parameter using record values from proportional hazard rate models*. Communications in Statistics - Theory and Methods, 45, no. 19, pp. 5787–5801.

K. N. CHANDLER (1952). *The distribution and frequency of record values*. Journal of the Royal Statistical Society, Series B, 14, no. 2, pp. 220–228.

M. CROWDER (2000). *Tests for a family of survival models based on extremes*. In N. LIMNIOS, M. NIKULIN (eds.), *Recent Advances in Reliability Theory Methodology, Practice, and Inference*, Birkhäuser Boston, Boston, MA, pp. 307–321.

S. KOTZ, I. A. P. LUMELSKII, M. PENSKY (2003). *The Stress-Strength Model and its Generalizations: Theory and Applications*. World Scientific, Singapore.

D. V. LINDLEY (1980). *Approximate Bayesian methods*. Trabajos de Estadistica Y de Investigacion Operativa, 31, no. 1, pp. 223–245.

S. Rezaei, R. Tahmasbi, M. Mahmoodi (2010). *Estimation of $P(Y<X)$ for generalized Pareto distribution*. Journal of Statistical Planning and Inference, 140, no. 2, pp. 480–494.

B. Tarvirdizade, M. Ahmadpour (2016). *Estimation of the stress-strength reliability for the two-parameter bathtub-shaped lifetime distribution based on upper record values*. Statistical Methodology, 31, pp. 58–72.

Summary

In this paper, the estimation of stress-strength reliability based on upper record values is considered when X and Y are independent random variables having a Pareto distribution with the same scale parameter and with different shape parameters. The maximum likelihood estimator (MLE), the approximate Bayes estimators and the exact confidence interval of the stress-strength reliability are obtained. A Monte Carlo simulation study is conducted to investigate the merits of the proposed methods. A real data analysis is presented for illustrative purpose.

Keywords: Stress-strength reliability; Record values; Pareto distribution; Maximum likelihood estimator; Bayes estimator.