STATISTICA, anno LXXXI, n. 4, 2021

INFERENCE ON $P(Y < X)$ BASED ON RANKED SET SAMPLE FOR GENERALIZED PARETO DISTRIBUTION

Manoj Chacko
Department of Statistics, University of Kerala, Kerala, India
Shiny Mathew[1]
Department of Statistics, University of Kerala, Kerala, India

1. INTRODUCTION

The concept of ranked set sampling (RSS) was first introduced by McIntyre (1952), as a process of improving the precision of the sample mean as an estimator of the population mean. RSS as described in McIntyre is applicable whenever ranking of a set of sampling units can be done easily by judgment method without error and additional cost. The procedure involves randomly choosing n sets of size n each from a population and ranking the units in each set visually or using some method that has negligible cost. From the first set of n units, the unit ranked lowest is chosen for actual quantification. From the second set of n units, the unit ranked second lowest is chosen for actual quantification. The process is continued until the unit ranked highest is chosen from the nth set for actual quantification. Let $X_{r(r)}$ be the observation measured on the variable of interest, say X with cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$ from the unit chosen from the rth set. If the ranking made on each set is perfect, then clearly $X_{r(r)}$ is the rth order statistic arising from a random sample of size n. Moreover $X_{r(r)}, r = 1, 2, \dots, n$ are independent. For convenience through out the paper we denote $X_{r(r)}$ by X_r. Then, the joint density of $X_1, X_2, \dots, X_n$ is given by

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} \frac{n!}{(i-1)!(n-i)!} [F(x_i)]^{i-1} [1 - F(x_i)]^{n-i} f(x_i). \quad (1)$$

The RSS is seen applied in many areas such as forest, agriculture, animal sciences, medicine etc.Wang *et al.* (2009) used RSS in fisheries research. Tiwari and Pandey (2013) considered an application of RSS design in environmental studies. For more details on applications of RSS see Dong *et al.* (2012) and Chen *et al.* (2004). In reliability theory, the

[1] Corresponding Author. E-mail: mshiny15@gmail.com

stress-strength model describes the life of a component which has a random strength X and is subjected to random stress Y. The component fails when the stress applied to it exceeds the strength and the component will function whenever $Y < X$. Thus $R = P(Y < X)$ can be considered as a measure of component reliability. It has found applications in many life testing problems and engineering. For more details on applications of R in engineering see, Nadarajah and Kotz (2006). In the context of system reliability, Muttlak *et al.* (2010), Akgul and Senoglu (2017), Mahdisadeh and Zamanzade (2018), Safariyan *et al.* (2019) obtained the estimator of stress-strength reliability based on RSS using parametric or non-parametric methods. Chacko and Mathew (2019) considered inference on $P(X < Y)$ for bivariate normal distribution based on ranked set sample.

The problem of estimation of $R = P(Y < X)$ has been widely used in the statistical literature. Inference on R was carried out by several authors for the majority of common distribution families. The maximum likelihood estimator (MLE) of $P(Y < X)$, when X and Y are normally distributed, has been considered by Downtown (1973) and Govidarajulu (1967). Tong (1977) considered the estimation of $P(Y < X)$ when X and Y are independent exponential random variables. Kundu and Gupta (2005) considered the problem of estimation of $P(Y < X)$, when X and Y are independent generalized exponential distribution. Awad *et al.* (1981) considered the MLE of R, when X and Y are bivariate exponential distributions. Ahmad *et al.* (1997) and Surles and Padgett (1998) considered the problem of estimation of $P(Y < X)$, where X and Y are Burr type X random variables. Rezaei *et al.* (2010) considered the estimation of $P(Y < X)$ when X and Y are independent generalized Pareto distributions.

In this study, we are interested in making inference about R using RSS, when X and Y follow independent generalized Pareto (GP) distribution. A random variable X is said to follow GP distribution if its pdf is given by (see Rezaei *et al.*, 2010)

$$f_X(x) = \alpha\lambda(1+\lambda x)^{-(\alpha+1)};\ x > 0,\ \lambda > 0 \text{ and } \alpha > 0. \tag{2}$$

Here α and λ are the shape and scale parameters respectively. This distribution is also known as Pareto distribution of the second type or Lomax distribution. Shi *et al.* (1999) used generalized Pareto distribution to estimate the size of the maximum inclusion in clean steels. Wong (2012) considered the problem of interval estimation of $P(Y < X)$ for generalized Pareto distribution. Through out the paper we denote $GP(\alpha, \lambda)$ for a generalized Pareto distribution with shape parameter α and scale parameter λ. The cdf of the $GP(\alpha, \lambda)$ distribution with pdf defined in Eq. (2) is given by

$$F(x) = 1 - (1+\lambda x)^{-\alpha};\ \lambda > 0 \text{ and } \alpha > 0. \tag{3}$$

In this paper, we consider the problem of estimation of $R = P(Y < X)$ based on RSS data when X and Y follow generalized Pareto distributions. In Section 2, the maximum likelihood estimation of R when all the parameters of X and Y are unknown and different is considered. In Section 3, Bayes estimation of R is considered under both symmetric and asymmetric loss functions. In Section 4, we use a real data to illustrate the inferential procedures described in the previous Sections. Section 5 is devoted to some simulation studies. Finally, a conclusion is given in Section 6.

2. MAXIMUM LIKELIHOOD ESTIMATION OF R

In this Section, we consider the problem of estimation of R when all the parameters of X and Y are different. Let $X \sim GP(\alpha,\lambda_1)$ and $Y \sim GP(\beta,\lambda_2)$ where X and Y are independent random variables. Then, R is given by (see Surles and Padgett, 2001)

$$\begin{aligned} R &= \int_0^\infty \alpha\lambda_1(1+\lambda_1 x)^{-(\alpha+1)} 1-(1+\lambda_2 x)^{-\beta}\,dx \\ &= 1-\alpha\left(\frac{\lambda_2}{\lambda_1}\right)^{-\beta}\int_0^\infty (1+t)^{-(\alpha+1)}\left(\frac{\lambda_1}{\lambda_2}+t\right)^{-\beta} dt. \end{aligned} \tag{4}$$

Considering the integral of the form

$$\int_0^\infty x^{\nu-1}(\beta+x)^{-\mu}(x+\gamma)^{-Q}dx = \beta^{-\mu}\gamma^{\nu-Q}B(\nu,\mu-\nu+Q)_2F_1\left(\mu,\nu;\mu+Q;1-\frac{\gamma}{\beta}\right), \tag{5}$$

where $\nu>0$, $\mu>\nu-Q$ and ${}_2F_1(.)$ is Gauss' hypergeometric function given by

$${}_2F_1(a,b;c;\theta)=\sum_{j=0}^\infty \frac{(a)_j(b)_j}{(c)_j}\frac{\theta^j}{j!}, \tag{6}$$

where $(x)_k = x(x+1)(x+2)\cdots(x+k-1)$ for $k\geq 1$ with $(x)_0=1$. For more details of this function see Mathai and Haubold (2008). Then,

$$\begin{aligned} R &= 1-\alpha\left(\frac{\lambda_2}{\lambda_1}\right)^{-\beta}\left(\frac{\lambda_1}{\lambda_2}\right)^{1-\beta}B(1,\alpha+\beta)_2F_1\left(\alpha+1,1;\alpha+\beta+1;1-\frac{\lambda_1}{\lambda_2}\right) \\ &= 1-\frac{\alpha}{\alpha+\beta}\frac{\lambda_1}{\lambda_2}{}_2F_1\left(\alpha+1,1;\alpha+\beta+1;1-\frac{\lambda_1}{\lambda_2}\right). \end{aligned} \tag{7}$$

Let $X_i, i=1,2,\ldots,m$ be a ranked set sample from $GP(\alpha,\lambda_1)$ and $Y_j, j=1,2,\ldots,n$ be a ranked set sample from $GP(\beta,\lambda_2)$. If we denote $\delta=(\alpha,\beta,\lambda_1,\lambda_2)$, then the likelihood function is given by

$$\begin{aligned} L(\delta) &= K\alpha^m\beta^n\lambda_1^m\lambda_2^n\prod_{i=1}^m\left[1-(1+\lambda_1 x_i)^{-\alpha}\right]^{i-1}\left[1+\lambda_1 x_i\right]^{\alpha(i-m-1)-1} \\ &\quad \prod_{j=1}^n\left[1-(1+\lambda_2 y_j)^{-\beta}\right]^{j-1}\left[1+\lambda_2 y_j\right]^{\beta(j-n-1)-1}, \end{aligned} \tag{8}$$

where

$$K=\prod_{i=1}^{m}\frac{m!}{(i-1)!(m-i)!}\prod_{j=1}^{n}\frac{n!}{(j-1)!(n-j)!}. \tag{9}$$

The log-likelihood function is given by

$$\begin{aligned}\log L(\delta) &= \log c+\sum_{i=1}^{m}(i-1)\log\left(1-(1+\lambda_1 x_i)^{-\alpha}\right)\\ &+m\left[\log\alpha+\log\lambda_1\right]+\sum_{j=1}^{n}(j-1)\log\left(1-(1+\lambda_2 y_j)^{-\beta}\right)\\ &+n\left[\log\beta+\log\lambda_2\right]+\sum_{i=1}^{m}[\alpha(i-m-1)-1]\log(1+\lambda_1 x_i)\\ &+\sum_{j=1}^{n}[\alpha(j-n-1)-1]\log(1+\lambda_2 y_j).\end{aligned} \tag{10}$$

Thus

$$\begin{aligned}\frac{\partial\log L}{\partial\alpha} &= \frac{m}{\alpha}+\sum_{i=1}^{m}(i-m-1)\log(1+\lambda_1 x_i)\\ &+\sum_{i=1}^{m}(i-1)\frac{(1+\lambda_1 x_i)^{-\alpha}\log(1+\lambda_1 x_i)}{1-(1+\lambda_1 x_i)^{-\alpha}},\end{aligned} \tag{11}$$

$$\begin{aligned}\frac{\partial\log L}{\partial\beta} &= \frac{n}{\beta}+\sum_{j=1}^{n}(j-n-1)\log(1+\lambda_2 y_j)\\ &+\sum_{j=1}^{n}(j-1)\frac{(1+\lambda_2 y_j)^{-\beta}\log(1+\lambda_2 y_j)}{1-(1+\lambda_2 y_j)^{-\beta}},\end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial\log L}{\partial\lambda_1} &= \frac{m}{\lambda_1}+\sum_{i=1}^{m}(i-1)\frac{\alpha x_i(1+\lambda_1 x_i)^{-(\alpha+1)}}{1-(1+\lambda_1 x_i)^{-\alpha}}\\ &+\sum_{i=1}^{m}[\alpha(i-m-1)-1]\frac{x_i}{1+\lambda_1 x_i}\end{aligned} \tag{13}$$

and

$$\begin{aligned}\frac{\partial\log L}{\partial\lambda_2} &= \frac{n}{\lambda_2}+\sum_{j=1}^{n}(j-1)\frac{\beta y_j(1+\lambda_2 y_j)^{-(\beta+1)}}{1-(1+\lambda_2 y_j)^{-\beta}}\\ &+\sum_{j=1}^{n}[\beta(j-n-1)-1]\frac{y_j}{1+\lambda_2 y_j}.\end{aligned} \tag{14}$$

The MLEs of α,β,λ_1 and λ_2 say $\hat{\alpha},\hat{\beta},\hat{\lambda}_1$ and $\hat{\lambda}_2$ can be obtained by equating each of the Equations from (11) to (14) to 0 and by solving those equations.

Then, by the invariance property of the ML estimators, the MLE of R becomes

$$\hat{R} = 1 - \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} \frac{\hat{\lambda}_1}{\hat{\lambda}_2} {}_2F_1\left(\hat{\alpha}+1, 1; \hat{\alpha}+\hat{\beta}+1; 1-\frac{\hat{\lambda}_1}{\hat{\lambda}_2}\right). \tag{15}$$

2.1. *Bootstrap Confidence Interval*

Next we consider the percentile bootstrap confidence interval for R based on MLEs. For that first we compute the MLEs $\hat{\alpha}^{(0)}$, $\hat{\beta}^{(0)}$, $\hat{\lambda}_1^{(0)}$ and $\hat{\lambda}_2^{(0)}$ of α, β, λ_1 and λ_2 using original ranked set sample. Generate a bootstrap RSS of size m from $GP\left(\hat{\alpha}^{(0)}, \hat{\lambda}_1^{(0)}\right)$ and a bootstrap RSS of size n from $GP\left(\hat{\beta}^{(0)}, \hat{\lambda}_2^{(0)}\right)$. Obtain the MLEs $\hat{\alpha}^{(1)}$, $\hat{\beta}^{(1)}$, $\hat{\lambda}_1^{(1)}$ and $\hat{\lambda}_2^{(1)}$ using the bootstrap samples and find the MLE $\hat{R}_1$. Repeat the procedure B times to have $\hat{R}_k$ for $k = 1, 2, \ldots, B$. Arrange $\hat{R}_k$ for $k = 1, 2, \ldots, B$ in ascending order as $\hat{R}_{(1)} \leq \hat{R}_{(2)}, \ldots, \leq \hat{R}_{(B)}$. Then, the $100(1-\nu)$ percentile bootstrap CI for R is given by $\left(\hat{R}_{([B(\nu/2)])}, \hat{R}_{([B(1-\nu/2)])}\right)$.

REMARK 1. *If the scale parameters of X and Y are same that is, $\lambda_1 = \lambda_2$, then $R = P(Y < X)$ is given by*

$$R = \frac{\beta}{\alpha+\beta} \tag{16}$$

3. BAYESIAN ESTIMATION OF R

In this Section, we obtain the Bayes estimator of R when the scale parameters of both X and Y are different. Let X_i, $i = 1, 2, \ldots, m$ be a ranked set sample from $GP(\alpha, \lambda_1)$ and $Y_j, j = 1, 2, \ldots, n$ be a ranked set sample from $GP(\beta, \lambda_2)$. Then, the likelihood function is given by

$$\begin{aligned} L(\delta) &= K\alpha^m \beta^n \lambda_1^m \lambda_2^n \prod_{i=1}^{m} \left[1-(1+\lambda_1 x_i)^{-\alpha}\right]^{i-1} \left[1+\lambda_1 x_i\right]^{\alpha(i-m-1)-1} \\ & \prod_{j=1}^{n} \left[1-(1+\lambda_2 y_j)^{-\beta}\right]^{j-1} \left[1+\lambda_2 y_j\right]^{\beta(j-n-1)-1}. \end{aligned} \tag{17}$$

For the Bayes estimation we assume that the prior distributions of α, β, λ_1 and λ_2 follow independent gamma distributions with density functions given by

$$\pi_1(\alpha|a_1, b_1) = \frac{b_1^{a_1}}{\Gamma(a_1)} \alpha^{a_1-1} e^{-b_1\alpha}; \; a_1 > 0 \text{ and } b_1 > 0, \tag{18}$$

$$\pi_2(\beta|a_2,b_2)=\frac{b_2^{a_2}}{\Gamma(a_2)}\beta^{a_2-1}e^{-b_2\beta};\, a_2>0 \text{ and } b_2>0, \tag{19}$$

$$\pi_3(\lambda_1|a_3,b_3)=\frac{b_3^{a_3}}{\Gamma(a_3)}\lambda_1^{a_3-1}e^{-b_3\lambda_1};\, a_3>0 \text{ and } b_3>0 \tag{20}$$

and

$$\pi_4(\lambda_2|a_4,b_4)=\frac{b_4^{a_4}}{\Gamma(a_4)}\lambda_2^{a_4-1}e^{-b_4\lambda_2};\, a_4>0 \text{ and } b_4>0. \tag{21}$$

Then, the joint prior density of $\delta=(\alpha,\beta,\lambda_1,\lambda_2)$ is given by

$$\pi(\delta)=\frac{b_1^{a_1}b_2^{a_2}b_3^{a_3}b_4^{a_4}}{\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)\Gamma(a_4)}\alpha^{a_1-1}\beta^{a_2-1}\lambda_1^{a_3-1}\lambda_2^{a_4-1}e^{-b_1\alpha}e^{-b_2\beta}e^{-b_3\lambda_1}e^{-b_4\lambda_2}. \tag{22}$$

Then, the joint posterior density of α, β, λ_1 and λ_2 given the data is

$$\pi^*(\delta|\text{data})=\frac{L(\delta)\pi(\delta)}{\int L(\delta)\pi(\delta)d\delta}. \tag{23}$$

Since the posterior pdf in Eq. (23) cannot be reduced analytically to a closed form we propose MCMC method to find the Bayes estimates of R.

3.1. MCMC Method

In this Sub-Section, we consider MCMC method to find the Bayes estimate of R. The joint posterior density given in Eq. (23) can be written as

$$\begin{aligned}\pi^*(\delta|\text{data}) \;\propto\;& \alpha^{a_1+m-1}\beta^{a_2+n-1}\lambda_1^{a_3+m-1}\lambda_2^{a_4+n-1}\exp(-b_1\alpha)\exp(-b_2\beta)\\ & \exp(-b_3\lambda_1)\exp(-b_4\lambda_2)\prod_{i=1}^{m}\left(1-(1+\lambda_1x_i)^{-\alpha}\right)^{i-1}\\ & (1+\lambda_1x_i)^{\alpha(i-m-1)-1}\prod_{j=1}^{n}\left(1-(1+\lambda_2y_j)^{-\beta}\right)^{j-1}(1+\lambda_2y_j)^{\beta(j-n-1)-1}.\end{aligned} \tag{24}$$

From Eq. (24) the conditional posterior density of α given β, λ_1, λ_2 and data is given by

$$\pi_1^*(\alpha|\beta,\lambda_1,\lambda_2,data)\propto\alpha^{a_1+m-1}\exp\left(-(b_1-T_1)\alpha\right)\prod_{i=1}^{m}\left(1-(1+\lambda_1x_i)^{-\alpha}\right)^{i-1}, \tag{25}$$

where $T_1=\sum_{i=1}^{m}(i-m-1)\log(1+\lambda_1x_i)$. The conditional posterior density of β given α, λ_1, λ_2 and the data is given by

$$\pi_2^*(\beta|\alpha,\lambda_1,\lambda_2,\text{data})\propto\beta^{a_2+n-1}\exp\left(-(b_2-T_2)\beta\right)\prod_{j=1}^{n}\left(1-(1+\lambda_2y_j)^{-\beta}\right)^{j-1}, \tag{26}$$

where $T_2=\sum_{j=1}^{n}(j-n-1)\log(1+\lambda_2 y_j)$. The conditional posterior density of λ_1 given α, β, λ_2 and the data is given by

$$\pi_3^*(\lambda_1|\alpha,\beta,\lambda_2,\text{data}) \propto \lambda_1^{a_3+m-1}\exp(-b_3\lambda_1)\prod_{i=1}^{m}\left(1-(1+\lambda_1 x_i)^{-\alpha}\right)^{i-1} (1+\lambda_1 x_i)^{\alpha(i-m-1)-1}. \tag{27}$$

The conditional posterior density of λ_2 given α, β, λ_1 and the data is given by

$$\pi_4^*(\lambda_2|\alpha,\beta,\lambda_1,\text{data}) \propto \lambda_2^{a_4+n-1}\exp(-b_4\lambda_2)\prod_{j=1}^{n}\left(1-(1+\lambda_2 y_j)^{-\beta}\right)^{j-1} (1+\lambda y_j)^{\beta(j-n-1)-1}. \tag{28}$$

We use Metropolis-Hasting (M-H) algorithm with in the Gibbs sampling procedure to generate samples from conditional posterior distributions. By setting initial values $\alpha^{(0)}$, $\beta^{(0)}$, $\lambda_1^{(0)}$ and $\lambda_2^{(0)}$, let $\alpha^{(t)}$, $\beta^{(t)}$, $\lambda_1^{(t)}$ and $\lambda_2^{(t)}$, $t=1,2,\ldots,N$ be the observations generated from Equations (25),(26),(27) and (28), respectively. Then, the Bayes estimators of R under SEL, LL and EL by taking first M iterations as burn-in period, are respectively given by

$$\hat{R}_s=\frac{1}{N-M}\sum_{t=M+1}^{N}\hat{R}^{(t)}, \tag{29}$$

$$\hat{R}_L=\frac{-1}{h}\log\left[\frac{1}{N-M}\sum_{t=M+1}^{N}e^{-h\hat{R}^{(t)}}\right] \tag{30}$$

and

$$\hat{R}_E=\left[\frac{1}{N-M}\sum_{t=M+1}^{N}\left(\hat{R}^{(t)}\right)^{-q}\right]^{\frac{-1}{q}}, \tag{31}$$

where

$$\hat{R}^{(t)}=1-\frac{\hat{\alpha}^{(t)}}{\hat{\alpha}^{(t)}+\hat{\beta}^{(t)}}\frac{\hat{\lambda}_1^{(t)}}{\hat{\lambda}_2^{(t)}}\,{}_2F_1\left(\hat{\alpha}^{(t)}+1,1;\hat{\alpha}^{(t)}+\hat{\beta}^{(t)}+1;1-\frac{\hat{\lambda}_1^{(t)}}{\hat{\lambda}_2^{(t)}}\right). \tag{32}$$

4. Illustration using real data

In this Section, we illustrate the inferential procedures described in the previous Sections using real data sets given in Pandit and Joshi (2018), which were originally reported by Badar and Priest (1982). The data represents the strength measured in GPA for single carbon fibers and impregnated 1000-carbon fiber tows. The single fibers were tested under tension at gauge lengths of 20mm (data set I) and 10 mm (data set II). Pandit and Joshi (2018) fitted the GP distribution to both data sets. Let X and Y denote the strength for single carbon fiber under tension at gauge lengths 20mm and 10mm respectively. Then, the problem is to estimate the probability that the strength for single carbon fiber under tension at gauge length 10mm is less than the strength at 20mm,that is, $P(Y<X)$, using RSS. For that choose 49 ($m=7$) units randomly from data set I and arrange them into 8 sets of 8 units each. Similarly choose 49 ($n=7$) units randomly from data set II and arrange them into 7 sets of 7 units each. Then, the RSS observations from X and Y are given in Table 1.

TABLE 1
RSS observations from the strength for single carbon fiber under tension at gauge lengths 20mm (X) and 10mm (Y).

i	1	2	3	4	5	6	7
X	2.055	2.301	2.253	2.684	2.566	2.586	3.433
Y	2.006	1.479	2.575	2.624	2.880	2.566	3.585

We have obtained the MLEs and Bayes estimates of $P(Y<X)$ in the general case based on RSS (Table 2). For Bayes estimation we took non-informative priors for α, β, λ_1 and λ_2 and are obtained when $a_1=0$, $b_1=0$, $a_2=0$, $b_2=0$, $a_3=0$, $b_3=0$, $a_4=0$ and $b_4=0$. Bayes estimates are obtained using 50000 iterations in which we take first 5000 iterations as burn-in period. Results show that the estimated values of $P(Y<X)$ is more or less 0.5 for all the cases. Therefore we cannot claim that the strength for single fiber carbon under tension at gauge length 10mm is relatively smaller than the strength at 20mm.

TABLE 2
MLEs and Bayes estimates of $P(Y<X)$ from RSS observations in Table 1.

MLE			0.517
MCMC method	SEL		0.487
	LL	$h=-1$	0.515
		$h=1$	0.528
	EL	$q=-0.5$	0.491
		$q=0.5$	0.542

5. SIMULATION STUDY

In this Section, a Monte Carlo simulation is performed to study the efficiency of the estimators developed in the previous Sections. We have considered the case when the scale parameters are unknown and different. We have obtained the average bias and MSE of MLEs of R, bootstrap confidence intervals for R and are given in Table 3. For the Bayesian estimation we took the hyper parameters for the prior distributions of α, β, λ_1 and λ_2 as $a_1 = 2$, $b_1 = 2$, $a_2 = 2$, $b_2 = 2$, $a_3 = 2$ $b_3 = 2$, $a_4 = 2$ and $b_4 = 2$. We have obtained the bias and MSE of Bayes estimator of R under SEL, LL and EL functions (Table 4). The 95% HPD credible intervals of R are also obtained by using the method given in Chen and Shao (1999) (Table 3).

TABLE 3
The bias & MSE of MLEs for R and AIL & CP for CIs when $\lambda_1 = 0.5$ and $\lambda_2 = 1$.

m/n	α	β	R	MLE		Bootstrap		HPD	
				Bias	MSE	AIL	CP	AIL	CP
2	0.5	0.5	0.586	0.137	0.195	0.289	82	0.274	92
2	1.5	0.5	0.343	0.245	0.294	0.452	78	0.271	91
2	0.5	1.5	0.828	0.179	0.212	0.212	75	0.274	92
2	2	1.5	0.562	0.097	0.230	0.352	82	0.273	94
4	0.5	0.5	0.586	0.097	0.180	0.339	78	0.272	89
4	1.5	0.5	0.343	0.192	0.284	0.311	75	0.269	88
4	0.5	1.5	0.828	0.173	0.177	0.260	73	0.259	91
4	2	1.5	0.562	-0.112	0.215	0.377	75	0.170	93
6	0.5	0.5	0.586	-0.084	0.178	0.307	77	0.167	90
6	1.5	0.5	0.343	-0.172	0.237	0.226	72	0.161	91
6	0.5	1.5	0.828	-0.163	0.163	0.261	82	0.168	92
6	2	1.5	0.562	0.103	0.187	0.325	79	0.156	90
8	0.5	0.5	0.586	-0.076	0.144	0.230	78	0.152	88
8	1.5	0.5	0.343	0.152	0.175	0.233	76	0.142	92
8	0.5	1.5	0.828	-0.152	0.148	0.194	78	0.157	92
8	2	1.5	0.562	0.097	0.162	0.197	73	0.142	94
10	0.5	0.5	0.586	-0.060	0.013	0.190	76	0.142	89
10	1.5	0.5	0.343	-0.111	0.146	0.162	82	0.133	91
10	0.5	1.5	0.828	0.146	0.127	0.178	80	0.136	92
10	2	1.5	0.562	-0.075	0.089	0.176	79	0.136	91
15	0.5	0.5	0.586	-0.056	0.012	0.184	79	0.140	92
15	0.5	1	0.343	-0.105	0.116	0.158	78	0.124	93
15	1	1.5	0.828	0.125	0.118	0.138	81	0.116	92
15	2	1.5	0.562	-0.067	0.081	0.158	79	0.125	91

TABLE 4

The bias and MSE for Bayes estimators for R when $\lambda_1 = 0.5$ and $\lambda_2 = 1$.

m/n	α	β	R	SEL		LL				EL			
				Bias	MSE	$h=-1$		$h=1$		$q=-.05$		$q=0.5$	
						Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
2	0.5	0.5	0.585	-0.141	0.082	-0.097	0.070	-0.193	0.169	-0.165	0.095	-0.273	0.181
2	0.5	1	0.753	-0.313	0.184	-0.287	0.172	-0.411	0.184	-0.465	0.182	-0.442	0.382
2	1	1.5	0.712	-0.315	0.186	-0.323	0.154	-0.390	0.187	-0.479	0.219	-0.565	0.352
2	2	1.5	0.561	-0.088	0.048	-0.082	0.035	-0.186	0.087	-0.149	0.148	-0.192	0.082
4	0.5	0.5	0.585	-0.131	0.063	-0.092	0.068	-0.173	0.139	-0.158	0.085	-0.233	0.163
4	0.5	1	0.753	-0.264	0.178	-0.233	0.162	-0.341	0.161	-0.457	0.171	-0.425	0.376
4	1	1.5	0.712	-0.293	0.176	-0.230	0.151	-0.259	0.182	-0.453	0.196	-0.562	0.342
4	2	1.5	0.561	-0.079	0.031	-0.075	0.027	-0.171	0.043	-0.137	0.132	-0.171	0.079
6	0.5	0.5	0.585	-0.129	0.055	-0.086	0.063	-0.162	0.082	-0.149	0.074	-0.213	0.141
6	0.5	1	0.753	-0.240	0.173	-0.138	0.157	-0.321	0.160	-0.443	0.205	-0.387	0.351
6	1	1.5	0.712	-0.283	0.163	-0.273	0.149	-0.329	0.174	-0.439	0.166	-0.555	0.329
6	2	1.5	0.561	-0.069	0.025	-0.067	0.021	-0.168	0.039	-0.132	0.123	-0.154	0.063
8	0.5	0.5	0.585	-0.124	0.047	-0.072	0.061	-0.156	0.062	-0.128	0.068	-0.153	0.137
8	0.5	1	0.753	-0.238	0.156	-0.136	0.142	-0.281	0.153	-0.419	0.186	-0.332	0.293
8	1	1.5	0.712	-0.279	0.150	-0.235	0.133	-0.318	0.160	-0.422	0.185	-0.544	0.303
8	2	1.5	0.561	-0.063	0.021	-0.059	0.018	-0.145	0.037	-0.131	0.116	-0.127	0.059
10	0.5	0.5	0.585	-0.123	0.037	-0.062	0.059	-0.146	0.058	-0.127	0.067	-0.140	0.127
10	0.5	1	0.753	-0.232	0.118	-0.130	0.104	-0.265	0.143	-0.361	0.144	-0.531	0.280
10	1	1.5	0.712	-0.247	0.132	-0.232	0.115	-0.317	0.150	-0.393	0.166	-0.529	0.293
10	2	1.5	0.561	-0.052	0.020	-0.045	0.017	-0.121	0.034	-0.125	0.104	-0.112	0.048
15	0.5	0.5	0.585	-0.112	0.028	-0.057	0.048	-0.128	0.051	-0.113	0.052	-0.124	0.118
15	0.5	1	0.753	-0.215	0.073	-0.123	0.063	-0.197	0.096	-0.285	0.091	-0.245	0.220
15	1	1.5	0.712	-0.229	0.063	-0.203	0.051	-0.277	0.097	-0.272	0.087	-0.449	0.226
15	2	1.5	0.561	-0.033	0.016	-0.041	0.015	-0.112	0.026	-0.102	0.083	-0.103	0.043

The bias and MSE of all estimators decrease when the number of RSS units m and n increase. Bias and MSEs of Bayes estimators are smaller than the corresponding MLEs. Among the estimators, Bayes estimators under LINEX loss function when $h=-1$ have smaller bias and MSE. AILs of HPD intervals are smaller and the associated CPs are higher than that of bootstrap confidence intervals.

6. Conclusion

In this work, we considered the problem of estimation of $R=P(Y<X)$ for generalized Pareto distribution using RSS. The maximum likelihood and Bayesian estimators have been obtained for R. For obtaining the Bayes estimates, MCMC method has been applied. Based on the simulation study we have concluded that Bayes estimators perform better than the corresponding MLEs. Among the Bayes estimators, estimators under LINEX loss function perform better in terms of bias and MSE. AILs of HPD intervals are smaller and the associated CPs are higher than that of bootstrap confidence intervals.

Acknowledgements

The authors would like to thank the referees for the valuable comments.

References

K. Ahmad, M. Fakhry, Z. Jaheen (1997). *Empirical Bayes estimation of $P(Y<X)$ and characterizations of the Burr-type X model*. Journal of Statistical Planning and Inference, 64, pp. 297–308.

F. Akgul, B. Senoglu (2017). *Estimation of $P(X<Y)$ using ranked set sampling for the Weibull distribution*. Quality Technology and Quantity Management, 14(3), pp. 296–309.

A. Awad, M. Azzam, M. Hamdan (1981). *Some inference results on $P(Y<X)$ in the bivariate Exponential model*. Communications in Statistics - Theory and Methods, 10, pp. 2515–2525.

M. G. Badar, A. Priest (1982). *Statistical aspects of fibre and bundle strength in hybrid composites*. In T. Hayashi, K. Kawata, S. Umekawa (eds.), *Progress in Science and Engineering Composites*, ICCM-IV, Tokyo, pp. 1129–1136.

M. Chacko, S. Mathew (2019). *Inference on $P(X<Y)$ for bivariate Normal distribution based on ranked set sample*. Metron, 77, pp. 239–252.

M. Chen, Q. Shao (1999). *Monte Carlo estimation of Bayesian credible and HPD intervals*. Journal of Computational and Graphical Statistics, 8, pp. 69–92.

Z. Chen, Z. Bai, B. Sinha (2004). *Ranked Set Sampling, Theory and Applications.* Springer, New York.

X. Dong, L. Cui, F. Liu (2012). *A further study on reliable life estimation under ranked set sampling*. Communications in Statistics, Theory and Methods, 21, pp. 3888–3902.

F. Downtown (1973). *The estimation of $P(Y < X)$ in the Normal case*. Technometrics, 15, pp. 551–558.

Z. Govidarajulu (1967). *Two sided confidence limits for $P(X > Y)$ based on Normal samples of X and Y*. Sankhya, B29, pp. 35–40.

D. Kundu, R. Gupta (2005). *Estimation of $P(X > Y)$ for generalized Exponential distribution*. Metrika, 61, pp. 291–308.

M. Mahdisadeh, E. Zamanzade (2018). *A new reliability measure in ranked set sampling*. Statistical Papers, 59, no. 3, pp. 861–91.

A. M. Mathai, H. Haubold (2008). *Special Functions for Applied Scientists*. Springer, New York.

G. McIntyre (1952). *A method for unbiased selective sampling using ranked set*. Australian Journal of Agricultural Research, 3, pp. 385–390.

H. Muttlak, W. A. Abu-Dayyeh, M.F.Saleh, E. Al-Sawi (2010). *Estimating $P(Y < X)$ using ranked set sampling in case of the Exponential distribution*. Communications in Statistics, Theory Methods, 39, pp. 1855–1868.

S. Nadarajah, S. Kotz (2006). *Reliability for some bivariate Exponential distributions*. Mathematical Problems in Engineering, 2006, pp. 1–14.

P. Pandit, S. Joshi (2018). *Reliability estimation in multicomponent stress-strength model based on generalized Pareto distribution*. American Journal of Applied Mathematics and Statistics, 6, pp. 210–217.

S. Rezaei, R. Tahmasbi, M. Mahmoodi (2010). *Estimation of $P[Y < X]$ for generalized Pareto distribution*. Journal of Statistical Planning and Inference, 140, pp. 480–494.

A. Safariyan, M.Arashi, R.A.Melaghi (2019). *Improved point and interval estimation of the stress-strength reliability based on ranked set sampling*. Statistics, 53(1), pp. 101–25.

G. Shi, H. Atkinson, C. Sellaars, C. Anderson (1999). *Application of the generalized Pareto distribution to the estimation of the size of the maximum inclusion in clean steels.* Acta Materiala, 47, pp. 1455–1468.

J. SURLES, W. PADGETT (1998). *Inference for $P(Y < X)$ in the Burr-type X model.* Journal of Applied Statistical Sciences, 7, pp. 225–238.

J. SURLES, W. PADGETT (2001). *Inference for reliability and stress-strength for a scaled Burr type X distribution*. Metrics, 7, pp. 187–200.

N. TIWARI, G. PANDEY (2013). *Application of ranked set sampling design in environmental investigations for real data set*. Thailand Statistician, 11, no. 2, pp. 173–184.

H. TONG (1977). *On the estimation of $P(Y < X)$ for Exponential families*. IEEE Transactions on Reliability, 26, pp. 54–56.

Y. G. WANG, Y. YE, D. MILTON (2009). *Efficient designs for sampling and subsampling in fisheries research based on ranked sets*. Journal of Marine Science, 66, pp. 928–934.

A. WONG (2012). *Interval estimation of $P(Y < X)$ for generalized Pareto distribution*. Journal of Statistical Planning and Inference, 142, pp. 601–607.

SUMMARY

In this paper, the problem of estimation of $R = P(Y < X)$ based on ranked set sampling, when (X, Y) follows generalised Pareto distribution (GPD) is considered. The maximum likelihood (ML) estimators and Bayes estimators of R are obtained. A Monte Carlo simulation is also performed to study the behaviour of different estimators.

Keywords: Ranked set sample; Generalised Pareto distribution; Maximum likelihood estimator; Bayes estimation