Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/040203

Bayesian Estimation and Prediction for Pareto Distribution based on Ranked set Sampling

M. M. Mohie El-Din[1], *M. S. Kotb*[1] *and H. A. Newer*[2,*]

[1] Department of Mathematics, Faculty of Science, Al-Azhar University, Nasr City, Cairo 11884, Egypt
[2] Department of Mathematics, Faculty of Education, Ain-Shams University, Nasr City, Cairo 11757, Egypt

Received: 22 Nov. 2014, Revised: 3 Apr. 2015, Accepted: 4 Apr. 2015
Published online: 1 Jul. 2015

Abstract: In this paper, we provide Bayesian estimation for the parameters of the Pareto distribution based on simple random sample (SRS) and ranked set sampling (RSS) in two cases, one cycle and m-cycle. Posterior risk of the derived estimators are also obtained by using squared error loss (SEL). Two-sample Bayesian prediction for future observations are obtained by using SRS and RSS in two cases, one cycle and m-cycle. A simulation data for SRS and RSS for one cycle and two cycle are used to illustrate the results.

Keywords: RSS; SRS; Bayesian estimation; SEL; Bayesian prediction; Posterior risk; ordinary order statistic.

1 Introduction

Sampling methods play an important role in all kinds of disciplines, such as medical sciences, engineering, education, and industrial processes. RSS has been suggested by Mclntyre [12] as a useful technique for improving estimates of the mean and variance of population and he found that the estimator based on RSS is more efficient than SRS. The RSS method can be described as follows: randomly select n^2 units from the target population and put in n sets each of them are a SRS of size n and then the n units of each set are ranked visually in ascending order with respect to the variable of interest as the following: $X_{11}, X_{21}, \ldots, X_{n1}, X_{12}, X_{22}, \ldots, X_{n2}, \ldots, X_{1m}, X_{2m}, \ldots, X_{nm}$. From the first set of n units, the smallest ranked unit is measured, the second smallest ranked unit is measured from the second set of n units. The process is continued until the n-th largest ranked unit is measured from the last set. If we repeat this method m times, we get a RSS of size mn.
Assume that the variable of interest X has Pareto distribution which was suggested by Pareto [17] with a probability density function (pdf) is given by

$$f(x;\alpha,\beta) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1}, \quad x > \beta, \quad \beta, \alpha > 0, \tag{1}$$

and cumulative distribution function (cdf) is given by

$$F(x;\alpha,\beta) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}. \tag{2}$$

Al-Hadhrami and Al-Omari [1] introduced Bayesian inference of the variance of the normal distribution by using moving extremes ranked set sampling. Al-Omari and Jaber[3] used double RSS method for estimating the population mean. Al-Omari et al. [4] used extreme RSS method to find estimators of the population mean. Al-Saleh and Samuh [2] suggested multistage RSS. Chacko and Thomas [6] derived different estimators of Morgenstern type bivariate logistic distribution by using RSS. Efron and Morris [7] used risk Bayesian problem of estimating the mean of a normal distribution when the mean itself has a normal prior. Ghafoori et al. [8] discussed two-sample Bayesian prediction by using progressively Type-II censored data. Ibrahim and Syam [9] applied stratified median RSS method for estimating the population mean. Islam et al. [10] described the modified maximum likelihood estimator of location and scale parameters based on selected RSS for normal, uniform and two-parameter exponential distributions. Mohammadi and

* Corresponding author e-mail: haidynewer@edu.asu.edu.eg

Pazira. [13] showed Bayesian estimations of the parameters of the generalized exponential distribution by using censored data. Mohie El-Din et al. [14] studied two-sample Bayesian prediction intervals for order statistics based on the class of the inverse exponential-type distributions using a right censored sample. Mohie El-Din et al. [15] used multiply type-II censored data to find two-sample Bayesian prediction intervals. Panaitescu et al. [16] used Bayesian and non-Bayesian estimators using record statistics of the modified-inverse Weibull distribution. Sadek et al. [19] used the asymmetric loss function to derive the Bayesian estimate of the parameter of the exponential distribution based on RSS. Soliman et al. [20] made a comparison of estimates using record statistics from Weibull model by using Bayesian and non-Bayesian approaches. Zellner [21] introduced Bayesian estimation by using asymmetric loss function.
In the current investigation, Bayesian estimators under SEL function for the parameters of Pareto distribution are obtained based on SRS and RSS in two cases, one cycle RSS and m-cycle RSS in Section 2. Two-sample Bayesian prediction scheme by using SRS and RSS when both parameters are unknown is presented in Section 3. Simulation result is presented in Section 4. Finally, we make some concluding remarks in Section 5.

2 Bayes Estimation

In this section, Bayesian estimators under SEL function for the parameters of Pareto distribution are derived based on SRS and RSS.

2.1 Bayes estimation based on SRS

Suppose that $X_{11}, X_{21}, \ldots, X_{n1}, X_{12}, X_{22}, \ldots, X_{n2}, \ldots, X_{1m}, X_{2m}, \ldots, X_{nm}$ be m sets of order statistics each of size n, then the joint density in the case of dependence of $x_{i\ell}$, $i = 1, 2, \ldots, n$, $\ell = 1, 2, \ldots, m$ is given by

$$\begin{aligned} f(\underline{\mathbf{x}}|\alpha,\beta) &= (n!)^m \prod_{\ell=1}^{m}\prod_{i=1}^{n} f(x_{i\ell},\alpha,\beta) \\ &\propto \alpha^{nm} \exp\left(-\alpha \sum_{\ell=1}^{m}\sum_{i=1}^{n} \ln\left(\frac{x_{i\ell}}{\beta}\right)\right). \end{aligned} \tag{3}$$

To obtain the joint posterior density of α and β, we will use the prior density, suggested by Lwin [11] and generalized by Arnold and Press [18] of α and β which is given by

$$\pi(\alpha,\beta;\delta) = \pi_1(\alpha)\pi_2(\beta|\alpha) \propto \frac{\alpha^{\eta}}{\beta} \exp\left(-\alpha\left(\rho + a\ln\left(\frac{b}{\beta}\right)\right)\right), \tag{4}$$

where $\delta = (a,b,\eta,\rho)$, $\pi_1(\alpha)$ is a gamma distribution with parameters η and ρ and $\pi_2(\beta|\alpha)$ is a power function distribution with parameters αa and b.
From Eq.(3) and (4), the posterior density function based on SRS can be written as

$$\pi^*(\alpha,\beta|\underline{\mathbf{x}}) = A^{-1}\frac{\alpha^c}{\beta} \exp\left(-\alpha\left(\rho + a\ln\left(\frac{b}{\beta}\right) + \sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{\beta}\right)\right)\right), \quad 0<\beta<L \tag{5}$$

where $\underline{\mathbf{x}} = (x_{11}, x_{21}, \ldots, x_{n1}, x_{12}, x_{22}, \ldots, x_{n2}, \ldots, x_{1m}, x_{2m}, \ldots, x_{nm})$, $c = nm+\eta$,
$A = \frac{\Gamma(c)}{nm+a}\left(\rho + a\ln\left(\frac{b}{L}\right) + \sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{L}\right)\right)^{-c}$ and $L = min(x_{1\ell}, b)$.

Hence, Bayesian estimators of α and β under a SEL function are

$$\begin{aligned} \widehat{\alpha}_{BS} &= E(\alpha|\underline{\mathbf{x}}) \\ &= \int_0^L\int_0^\infty \alpha\pi^*(\alpha,\beta|\underline{\mathbf{x}})\mathrm{d}\alpha\mathrm{d}\beta \\ &= c\left(\rho + a\ln\left(\frac{b}{L}\right) + \sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{L}\right)\right)^{-1}, \end{aligned} \tag{6}$$

$$E(\alpha^2|\underline{\mathbf{x}}) = c(c+1)\left(\rho + a\ln\left(\frac{b}{L}\right) + \sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{L}\right)\right)^{-2}, \tag{7}$$

$$\widehat{\beta}_{BS} = A^{-1}\Gamma(c+1)\int_0^L \left(\rho + a\ln\left(\frac{b}{\beta}\right) + \sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{\beta}\right)\right)^{-(c+1)} \mathrm{d}\beta$$
$$= A^{-1}\Gamma(c+1)\frac{e^{\frac{\kappa}{\upsilon}}}{\upsilon}(\kappa - \upsilon\ln(L))^{-c}\Psi\left(1+c, \frac{\kappa}{\upsilon} - \ln(L)\right), \tag{8}$$

and

$$E(\beta^2|\underline{\mathbf{x}}) = A^{-1}\Gamma(c+1)\frac{e^{\frac{2\kappa}{\upsilon}}}{\upsilon}(\kappa - \upsilon\ln(L))^{-c}\Psi\left(1+c, \frac{2\kappa}{\upsilon} - 2\ln(L)\right), \tag{9}$$

where $\Psi(\cdot)$ be Exp. Integral function, $\kappa = \rho + a\ln(b) + \sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln(x_{i\ell})$ and $\upsilon = nm + a$.

The posterior risk (minimum posterior expected loss (MPEL)) of θ is the posterior variance. The posterior model is essentially an updated version of our prior knowledge about θ in light of knowledge of the sample data. So the risk function of $\theta = (\alpha, \beta)$ under a SEL function is given by

$$\widehat{\theta}_{Risk} = E(\theta^2|\underline{\mathbf{x}}) - (E(\theta|\underline{\mathbf{x}}))^2. \tag{10}$$

2.2 Bayes estimation based on RSS

In this subsection, we introduce Bayesian estimators under SEL function for the parameters based on RSS in two cases, one cycle RSS and m-cycle RSS.
Consider $Y_1, Y_2, \ldots, Y_n$ be a one cycle RSS from a Pareto distribution obtained from a complete sample of SRS, then the joint pdf of the independent $Y_1, Y_2, \ldots, Y_n$ is given by

$$f(y_1, y_2, \ldots, y_n) = \prod_{j=1}^{n} g(y_j), \tag{11}$$

where

$$g(y_j) = j\binom{n}{j}(F(y_j))^{j-1}(\overline{F}(y_j))^{n-j}f(y_j), \tag{12}$$

is the pdf for the j-th order statistic for a SRS of size n (see Arnold et al. [5]).

2.2.1 Bayes estimates based on one cycle RSS

From Eq.(12), the density function of the j-th order statistic Y_j can be written as

$$g(y_j|\alpha,\beta) = j\binom{n}{j}\frac{\alpha}{y_j}\sum_{k=0}^{j-1}\binom{j-1}{k}(-1)^k\exp\left(-\alpha(n-j+k+1)\ln\left(\frac{y_j}{\beta}\right)\right). \tag{13}$$

Then from Eq.(11) the joint density function of the RSS in the case of the independence of y_j is given by

$$f(\mathbf{y}|\alpha,\beta) = \sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\left(\prod_{j=1}^{n}\frac{1}{y_j}\right)\alpha^n$$
$$\times\exp\left(-\alpha\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{\beta}\right)\right), \tag{14}$$

where $c_{i_j}(j) = j\binom{j-1}{i_j}\binom{n}{j}(-1)^{i_j}$.
Using Eq.(4) and (14), the posterior density function can be derived as

$$\pi^*(\alpha,\beta|\mathbf{y}) = R^{-1}\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\frac{\alpha^{c_1}}{\beta}$$
$$\times\exp\left(-\alpha\left(\rho + a\ln\left(\frac{b}{\beta}\right) + \sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{\beta}\right)\right)\right), \tag{15}$$

where

$$R=\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\frac{\Gamma(c_1)}{(a+\sum_{j=1}^{n}(n-j+i_j+1))}$$
$$\times\left(\rho+a\ln\left(\frac{b}{L}\right)+\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{L}\right)\right)^{-c_1}. \quad (16)$$

Hence, Bayesian estimation of α and β under a SEL function is

$$\widehat{\alpha}_{BS}=R^{-1}\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\frac{\Gamma(c_1+1)}{(a+\sum_{j=1}^{n}(n-j+i_j+1))}$$
$$\times\left(\rho+a\ln\left(\frac{b}{L}\right)+\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{L}\right)\right)^{-(c_1+1)}, \quad (17)$$

$$E(\alpha^2|\mathbf{y})=R^{-1}\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\frac{\Gamma(c_1+2)}{(\sum_{j=1}^{n}(n-j+i_j+1)+a)}$$
$$\times\left(\rho+a\ln\left(\frac{b}{L}\right)+\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{L}\right)\right)^{-(c_1+2)}, \quad (18)$$

$$\widehat{\beta}_{BS}=R^{-1}\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\Gamma(c_1+1)\frac{e^{\frac{\kappa_1}{\upsilon_1}}}{\upsilon_1}(\kappa_1-\upsilon_1\ln(L))^{-c_1}$$
$$\times\Psi\left(1+c_1,\frac{\kappa_1}{\upsilon_1}-\ln(L)\right), \quad (19)$$

and

$$E(\beta^2|\mathbf{y})=R^{-1}\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]\Gamma(c_1+1)\frac{e^{\frac{2\kappa_1}{\upsilon_1}}}{\upsilon_1}(\kappa_1-\upsilon_1\ln(L))^{-c_1}$$
$$\times\Psi\left(1+c_1,\frac{2\kappa_1}{\upsilon_1}-2\ln(L)\right), \quad (20)$$

where $\kappa_1=\rho+a\ln(b)+\sum_{j=1}^{n}(n-j+i_j+1)\ln(y_j)$, $\upsilon_1=a+\sum_{j=1}^{n}(n-j+i_j+1)$ and $c_1=n+\eta$.

2.2.2 Bayes estimates based on m-cycle RSS

Let y_{jl}, $j=1,2,\ldots,n$, $\ell=1,2,\ldots,m$ be m-cycle RSS from Pareto distribution. The joint density function in this case is given by

$$f(\underline{\mathbf{y}}|\alpha,\beta)=\prod_{\ell=1}^{m}\prod_{j=1}^{n}f(y_{j\ell}|\alpha,\beta)$$
$$=\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\left(\prod_{\ell=1}^{m}\prod_{j=1}^{n}\frac{1}{y_{j\ell}}\right)\alpha^{nm}$$
$$\times\exp\left(-\alpha\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{\beta}\right)\right), \quad (21)$$

where $K_{i_j}^{\ell}=\left[\prod_{\ell=1}^{m}\prod_{j=1}^{n}c_{i_j^\ell}(j)\right]$.

Using Eq.(4) and (21), the posterior density function can be derived as

$$\pi^*(\alpha,\beta|\underline{\mathbf{y}}) = D^{-1}\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\frac{\alpha^{c_2}}{\beta} \times \exp\left(-\alpha\left(\rho+a\ln\left(\frac{b}{\beta}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{\beta}\right)\right)\right), \tag{22}$$

where

$$D = \prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\frac{\Gamma(c_2)}{(a+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1))} \times\left(\rho+a\ln\left(\frac{b}{L}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{L}\right)\right)^{-c_2}. \tag{23}$$

Hence, Bayesian estimation of α and β under a SEL function is

$$\widehat{\alpha}_{BS} = D^{-1}\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\frac{\Gamma(c_2+1)}{(a+\sum_{l=1}^{m}\sum_{j=1}^{n}(n-j+i_j+1))} \times\left(\rho+a\ln\left(\frac{b}{L}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{L}\right)\right)^{-(c_2+1)}, \tag{24}$$

$$E(\alpha^2|\underline{\mathbf{y}}) = D^{-1}\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\frac{\Gamma(c_2+2)}{(a+\sum_{l=1}^{m}\sum_{j=1}^{n}(n-j+i_j+1))} \times\left(\rho+a\ln\left(\frac{b}{L}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{L}\right)\right)^{-(c_2+2)}, \tag{25}$$

$$\widehat{\beta}_{BS} = D^{-1}\Gamma(c_2+1)\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\frac{e^{\frac{\kappa_2}{\upsilon_2}}}{\upsilon_2}(\kappa_2-\upsilon_2\ln(L))^{-c_2} \times\Psi\left(1+c_2,\frac{\kappa_2}{\upsilon_2}-\ln(L)\right), \tag{26}$$

and

$$E(\beta^2|\underline{\mathbf{y}}) = D^{-1}\Gamma(c_2+1)\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}\frac{e^{\frac{2\kappa_2}{\upsilon_2}}}{\upsilon_2}(\kappa_2-\upsilon_2\ln(L))^{-c_2} \times\Psi\left(1+c_2,\frac{2\kappa_2}{\upsilon_2}-2\ln(L)\right), \tag{27}$$

where $\kappa_2=\rho+a\ln(b)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln(y_{j\ell})$, $\upsilon_2=a+\sum_{j=1}^{n}(n-j+i_j+1)$ and $c_2=nm+\eta$.

3 Bayes Prediction

This section provides two-sample Bayesian prediction scheme by using SRS and RSS when both parameters are unknown. Suppose that $w_1,w_2,\ldots,w_{n_1}$ be a second independent sample of size n_1. To predict the future sample w_s, $s=1,2,...,n_1$

based on a complete sample of SRS and RSS, then the density function of w_s is given by

$$\begin{aligned} g(w_s|\alpha,\beta) &= s\binom{n_1}{s}(F(w_s))^{s-1}(\overline{F}(w_s))^{n_1-s}f(w_s) \\ &= s\binom{n_1}{s}\frac{\alpha}{w_s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k \\ &\times \exp\left(-\alpha(n_1-s+k+1)\ln\left(\frac{w_s}{\beta}\right)\right). \end{aligned} \tag{28}$$

3.1 Two sample Bayesian prediction intervals based on SRS

By using Eq.(5) and (28), the predictive density function of w_s is given by

$$\begin{aligned} f(w_s|\underline{\mathbf{x}}) &= \int_0^L\int_0^\infty g(w_s|\alpha,\beta)\pi^*(\alpha,\beta|\underline{\mathbf{x}})\mathrm{d}\alpha\mathrm{d}\beta \\ &= A^{-1}s\binom{n_1}{s}\frac{1}{w_s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k\frac{\Gamma(c+1)}{(nm+(n_1-s+k+1)+a)} \\ &\times\left(\rho+a\ln\left(\frac{b}{L}\right)+(n_1-s+k+1)\ln\left(\frac{w_s}{L}\right)+\sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{L}\right)\right)^{-(c+1)}. \end{aligned} \tag{29}$$

Hence, the predictive survival function is given by

$$\begin{aligned} P[w_s>\nu|\underline{\mathbf{x}}] &= \int_\nu^\infty f(w_s|\underline{\mathbf{x}})\mathrm{d}w_s \\ &= A^{-1}s\binom{n_1}{s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k \\ &\times\frac{\Gamma(c)}{(nm+(n_1-s+k+1)+a)(n_1-s+k+1)} \\ &\times\left(\rho+a\ln\left(\frac{b}{L}\right)+(n_1-s+k+1)\ln\left(\frac{\nu}{L}\right)+\sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{L}\right)\right)^{-c}. \end{aligned} \tag{30}$$

So the lower and upper $100\tau\%$ prediction bounds $[L(x),U(x)]$ for w_s are obtained by equating Eq.(30) to $(1+\tau)/2$ and $(1-\tau)/2$, respectively.
From Eq.(29), the predictive estimator of w_s, $s=1,2,...,n_1$ can be obtained as

$$\begin{aligned} \widehat{w}_s = E(w_s|\underline{\mathbf{x}}) &= \int_\beta^\infty w_s f(w_s|\underline{\mathbf{x}})\mathrm{d}w_s \\ &= A^{-1}s\binom{n_1}{s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k\frac{\Gamma(c+1)}{(nm+(n_1-s+k+1)+a)} \\ &\times\int_\beta^\infty(\rho+a\ln\left(\frac{b}{L}\right)+(n_1-s+k+1)\ln\left(\frac{w_s}{L}\right) \\ &+\sum_{\ell=1}^{m}\sum_{i=1}^{n}\ln\left(\frac{x_{i\ell}}{L}\right))^{-(c+1)}\mathrm{d}w_s. \end{aligned} \tag{31}$$

3.2 Two sample Bayesian prediction intervals based on one cycle RSS

By using Eq.(15) and (28), the predictive density function of w_s is given by

$$f(w_s|\mathbf{y}) = R^{-1}s\binom{n_1}{s}\frac{1}{w_s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]$$
$$\times\frac{\Gamma(c_1+1)}{\left(\sum_{j=1}^{n}(n-j+i_j+1)+(n_1-s+k+1)+a\right)}$$
$$\times(\rho+(n_1-s+k+1)\ln\left(\frac{w_s}{L}\right)$$
$$+a\ln\left(\frac{b}{L}\right)+\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{L}\right))^{-(c_1+1)}). \tag{32}$$

Hence, the predictive survival function is given by

$$P[w_s>v|\mathbf{y}] = R^{-1}s\binom{n_1}{s}\sum_{k=0}^{s-1}\binom{s-1}{k}\frac{(-1)^k}{(n_1-s+k+1)}\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]$$
$$\times\frac{\Gamma(c_1)}{\left(\sum_{j=1}^{n}(n-j+i_j+1)+(n_1-s+k+1)+a\right)}$$
$$\times(\rho+(n_1-s+k+1)\ln\left(\frac{v}{L}\right)$$
$$+a\ln\left(\frac{b}{L}\right)+\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{L}\right))^{-c_1}). \tag{33}$$

So the lower and upper $100\tau\%$ prediction bounds $[L(x),U(x)]$ for w_s are obtained by equating Eq.(33) to $(1+\tau)/2$ and $(1-\tau)/2$, respectively.
From Eq.(32), the predictive estimator of w_s, $s=1,2,...,n_1$ can be obtained as

$$\widehat{w}_s = R^{-1}s\binom{n_1}{s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k\sum_{i_1=0}^{0}\sum_{i_2=0}^{1}\cdots\sum_{i_n=0}^{n-1}\left[\prod_{j=1}^{n}c_{i_j}(j)\right]$$
$$\times\frac{\Gamma(c_1+1)}{\left(\sum_{j=1}^{n}(n-j+i_j+1)+(n_1-s+k+1)+a\right)}$$
$$\times\int_{\beta}^{\infty}(\rho+(n_1-s+k+1)\ln\left(\frac{w_s}{L}\right)$$
$$+a\ln\left(\frac{b}{L}\right)+\sum_{j=1}^{n}(n-j+i_j+1)\ln\left(\frac{y_j}{L}\right))^{-(c_1+1)})dw_s. \tag{34}$$

3.3 Two sample Bayesian prediction intervals based on m-cycle RSS

By using Eq.(22) and (28), the predictive density function of w_s is given by

$$f(w_s|\underline{y}) = D^{-1}s\binom{n_1}{s}\frac{1}{w_s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}$$
$$\times\frac{\Gamma(c_2+1)}{\left(\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)+1)+(n_1-s+k+1)+a\right)}$$
$$\times(\rho+(n_1-s+k+1)\ln\left(\frac{w_s}{L}\right)$$
$$+a\ln\left(\frac{b}{L}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{L}\right))^{-(c_2+1)}). \tag{35}$$

Hence, the predictive survival function is given by

$$P[w_s > v|\underline{\mathbf{y}}] = D^{-1}s\binom{n_1}{s}\sum_{k=0}^{s-1}\binom{s-1}{k}\frac{(-1)^k}{(n_1-s+k+1)}\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}$$
$$\times\frac{\Gamma(c_2)}{\left(\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)+(n_1-s+k+1)+a\right)}$$
$$\times\ (\rho+(n_1-s+k+1)\ln\left(\frac{v}{L}\right)$$
$$+\,a\ln\left(\frac{b}{L}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{L}\right))^{-c_2}). \tag{36}$$

So the lower and upper $100\tau\%$ prediction bounds $[L(x),U(x)]$ for w_s are obtained by equating Eq.(36) to $(1+\tau)/2$ and $(1-\tau)/2$, respectively.
From Eq.(35), the predictive estimator of w_s, $s=1,2,\ldots,n_1$ can be obtained as

$$\widehat{w}_s = D^{-1}s\binom{n_1}{s}\sum_{k=0}^{s-1}\binom{s-1}{k}(-1)^k\prod_{\ell=1}^{m}\sum_{i_1^\ell=0}^{0}\sum_{i_2^\ell=0}^{\ell}\cdots\sum_{i_n^\ell=0}^{n-1}K_{i_j}^{\ell}$$
$$\times\frac{\Gamma(c_2+1)}{\left(\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)+1)+(n_1-s+k+1)+a\right)}$$
$$\times\int_{\beta}^{\infty}(\rho+(n_1-s+k+1)\ln\left(\frac{w_s}{L}\right)$$
$$+\,a\ln\left(\frac{b}{L}\right)+\sum_{\ell=1}^{m}\sum_{j=1}^{n}(n-j+i_j^\ell+1)\ln\left(\frac{y_{j\ell}}{L}\right))^{-(c_2+1)})\mathrm{d}w_s. \tag{37}$$

4 Illustrative example

In this section, we present a simulation study to illustrate our previous theoretical results for Pareto distribution when both parameters are unknown.

4.1 Simulation Study

To illustrate Bayesian estimation and two-sample Bayesian prediction intervals results for the Pareto distribution based on SRS and RSS, we perform a simulation study using different sample sizes according to the following steps:

1. To compute Bayesian estimation, we choose the parameter values $(a,b,\eta,\rho)=(1,1,2,3)$ and then generate $\alpha=0.7497$ from $\pi_1(\alpha)$ and $\beta=0.6660$ from $\pi_2(\beta|\alpha)$.
2. By using the transformation $X_i=\beta(1-U_i)^{\frac{-1}{\alpha}}$ where U_i from $U(0,1)$, the generated sample of size $n=4,6,8$ from Pareto distribution can be obtained for RSS of one cycle $(m=1)$ and two cycle $(m=2)$ and SRS when $m=1$ and $m=2$.
3. Compute Bayesian estimates which derived in the previous sections by using the generated samples of SRS of size $n=4,6,8$ when $(m=1)$ and $(m=2)$ and RSS of size $n=4,6,8$ for one cycle and two cycle.
4. By using Eq.(10), posterior risk can be obtained and then we replicate the steps $2-3$ for 1000 times to compute the average of posterior risk. The results are displayed in Table 1.

Table 1: Posterior risk of the Bayesian estimates based on SRS and RSS in two cases $(m=1)$ and $(m=2)$.

		$m=1$		$m=2$	
n	*Par.*	SRS	RSS	SRS	RSS
4	α	1.9412	0.1903	0.3065	0.1163
	β	0.0171	0.0099	0.0081	0.0060
6	α	0.2742	0.0565	0.1262	0.0454
	β	0.0120	0.0051	0.0052	0.0035
8	α	0.1392	0.0261	0.0760	0.0254
	β	0.0087	0.0030	0.0038	0.0022

5.To compute two-sample Baysian prediction of w_s, $s = 1,2,3$, we choose the parameter values $(a,b,\eta,\rho) = (1,1,2,1)$ to generate $\alpha = 2.0002$ from $\pi_1(\alpha)$ and $\beta = 0.2499$ from $\pi_2(\beta|\alpha)$ and then by using the transformation $X_i = \beta(1-U_i)^{\frac{-1}{\alpha}}$ where U_i from $U(0,1)$, we generate SRS and RSS of size $n = 4,6,8$ in two cases when $(m = 1)$ and $(m = 2)$. The results are displayed in Table 2 and 3, respectively.

Table 2: The generate SRS of size $n = 4,6,8$.

m	n								
1	4	0.4075	0.5028	1.0191	1.8870				
	6	0.25017	0.2578	0.2672	0.3784	0.5256	1.2568		
	8	0.2510	0.2716	0.2834	0.3035	0.3854	0.4091	0.6121	0.6360
2	4	0.3445	0.4346	0.7466	0.7716	0.2959	0.8054	0.8512	1.2061
	6	0.2778	0.2866	0.3143	0.3745	0.5006	1.5104	0.2666	0.3044
		0.3632	0.4121	0.4260	0.4765				
	8	0.2634	0.2993	0.3013	0.3430	0.3758	0.9270	0.1833	1.2689
		0.2569	0.2785	0.2794	0.2830	0.2953	0.3433	0.4485	0.5120

Table 3: The generate RSS of size $n = 4,6,8$.

m	n								
1	4	0.4075	0.4302	0.4691	0.5194				
	6	0.2598	0.3269	0.3287	0.4293	0.6185	0.7071		
	8	0.2522	0.2697	0.2864	0.2986	0.3289	0.3728	0.3825	0.3957
2	4	0.2506	0.3065	0.3513	0.3102	0.3065	0.3102	0.7563	0.3581
	6	0.2536	0.2620	0.3109	0.2601	0.4774	0.2799	0.2621	0.2601
		0.2799	0.3293	0.8306	0.5316				
	8	0.2512	0.2573	0.2711	0.2664	0.3509	0.2575	0.5182	0.2859
		0.2572	0.2664	0.2575	0.2859	0.2805	0.4007	0.5025	0.9134

6.By using the generated SRS of size $n = 4,6,8$ when $(m = 1)$ and $(m = 2)$, a 95% two-sample Baysian prediction intervals of w_s, $s = 1,2,3$ are obtained from Eq.(30) and then we can find the predictive estimator of w_s, $s = 1,2,3$ by using Eq.(31). The results are displayed in Tables 4 and 5.

7.By using the generated RSS samples of size $n = 4,6,8$ for one cycle and two cycle, a 95% two-sample Baysian prediction intervals of w_s, $s = 1,2,3$ are obtained from Eq.(33) and Eq.(36), respectively. In this case we can find the predictive estimator of w_s, $s = 1,2,3$ by using Eq.(34) and Eq.(37), respectively. The results are displayed in Tables 4 and 5.

Table 4: The Bayesian prediction bounds for w_s and $s = 1,2,3$ based on SRS and RSS for $m = 1$.

		SRS				RSS			
n	s	Lower	Upper	Width	$\widehat{w}_s$	Lower	Upper	Width	$\widehat{w}_s$
4	1	0.3807	0.7484	0.3677	0.4473	0.2697	0.4464	0.1767	0.3056
	2	0.3829	1.5584	1.1755	0.6553	0.2701	0.7478	0.4777	0.3869
	3	0.4116	3.4369	3.0253	2.0389	0.2988	2.2376	1.9388	0.7648
6	1	0.2399	0.4062	0.1663	0.2830	0.2472	0.3801	0.1329	0.2773
	2	0.2411	0.7032	0.4621	0.3496	0.2499	0.5790	0.3291	0.3359
	3	0.2627	2.2549	1.9922	0.6655	0.2738	1.4175	1.1437	0.5265
8	1	0.2438	0.3812	0.1374	0.2804	0.2500	0.3436	0.0936	0.2724
	2	0.2400	0.6004	0.3594	0.3328	0.2620	0.4674	0.2054	0.3126
	3	0.2702	1.5728	1.3026	0.5583	0.2879	0.8973	0.6094	0.4228

Table 5: The Bayesian prediction bounds for w_s and $s = 1,2,3$ based on SRS and RSS for $m = 2$.

		SRS				RSS			
n	s	Lower	Upper	Width	$\widehat{w}_s$	Lower	Upper	Width	$\widehat{w}_s$
4	1	0.2815	0.6202	0.3387	0.3497	0.2446	0.3565	0.1119	0.2690
	2	0.3100	1.3824	1.0724	0.5320	0.3000	0.5159	0.2159	0.3188
	3	0.3295	2.9199	2.5904	1.8839	0.2944	1.1370	0.8426	0.4677
6	1	0.2627	0.3596	0.0969	0.2749	0.2453	0.3391	0.0938	0.2599
	2	0.3031	0.4910	0.1879	0.3257	0.3116	0.4757	0.1641	0.3038
	3	0.2818	1.0000	0.7182	0.4396	0.3315	1.0131	0.6816	0.4314
8	1	0.2540	0.3497	0.0957	0.2700	0.2447	0.3304	0.0857	0.2551
	2	0.3000	0.4775	0.1775	0.3169	0.3011	0.4427	0.1416	0.2948
	3	0.2758	0.9242	0.6484	0.4294	0.3255	0.8489	0.5234	0.3975

5 Conclusion

We present Bayesian estimation and two-sample Bayesian prediction scheme based on SRS and RSS. Pareto distribution is used as application example to illustrate our results. We compute posterior risk of the derived Bayesian estimates and then make a comparison between SRS and RSS. Our observations about the results are stated in the following points:

1. From Table 1, posterior risk of the Bayes estimates under SEL function decrease with increasing n based on SRS and RSS in two cases ($m = 1$) and ($m = 2$).
2. Posterior risk of two cycle RSS is better than posterior risk of one cycle RSS and we notice that posterior risk of SRS in case of $m = 2$ is better than posterior risk of SRS in case of $m = 1$.
3. Posterior risk of the Bayes estimates based on RSS for one cycle and two cycle are better than posterior risk based on SRS in two cases ($m = 1$) and ($m = 2$), respectively.
4. From Tables 4 and 5, we notice that the lengths of the prediction intervals are increasing with increasing s and decrease with increasing n for SRS and RSS in two cases ($m = 1$) and ($m = 2$).
5. It is evident from Tables 4 and 5 that the predictive estimator are increasing with increasing s and decrease with increasing n for SRS and RSS in two cases ($m = 1$) and ($m = 2$).
6. It is clear that the lengths of the prediction intervals based on two cycle RSS are better than the lengths of the prediction intervals based on one cycle RSS and lengths of the prediction intervals based on SRS for $m = 1$ are better than the lengths of the prediction intervals based on SRS for $m = 2$.
7. We obtain better results of lengths of the prediction intervals based on RSS than the lengths of the prediction intervals based on SRS.

References

[1] Al-Hadhrami, S. A. and Al-Omari, A. I. (2009). Bayesian inference on the variance of normal distribution using moving extremes ranked set sampling, J. of Mod. Appl. Stat. Meth.; 8(1), 273 − 281.
[2] Al-Saleh,M. F. and Samuh, m. h. (2008). On multistage ranked set sampling for distribution and median estimation, Computat. Stat. and Data Analy.; 52, 2066 − 2078.
[3] Al-Omari, A. I. and Jaber, K. H. (2008). Percentile double ranked set sampling, J. of Math. and Stat.; 4(1), 60 − 64.
[4] Al-Omari, A. I., Jaber, K. and Al-Omari, A. (2008). Modified ratio-type estimators of the mean using extreme ranked set sampling, J. of Math. and Stat.; 4(3), 150 − 155.
[5] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992). A First course in order statistics, John Wiley and Sons, New York.
[6] Chacko, M. and Thomas, P. Y. (2009). Estimation of parameters of Morgenstern type bivariate logistic distribution by ranked set sampling, J. Ind. Soc. Agril. Statist.; 63(1), 77 − 83.
[7] Efron, B. and Morris, C. (1971). Limiting the risk of Bayes and empirical Bayes estimatorspart I: the Bayes case, J. of the Amer. Stat.l Assoc.; 66, 807 − 815.
[8] Ghafoori, S., Habibi Rad, A. and Doostparast, M. (2011). Bayesian two-sample prediction with progressively Type-II censored data for some lifetime models, JIRSS; 10(1), 63 − 86.
[9] Ibrahim, K. and Syam, M. (2010). Estimating the population mean using stratified median ranked set sampling, Appl. Math. Sc.; 4(47), 2341 − 2354.

[10] Islam, T., Shaibur, M. R., and Hossain, S.S. (2009). Effectivity of modified maximum likelihood estimators using selected ranked set sampling data, Aust. J. of Stat.; 38(2), 109 – 120.
[11] Lwin, T. (1972). Estimating the tail of the Paretian law. Skandinavian Aktuarietiaskr; 55, 170 – 178.
[12] McIntyre, G.A. (1952). A method for unbiased selective sampling using ranked sets, Aust. J. Agri. Res.; 3, 385 – 390.
[13] Mohammadi, M. Y. and Pazira, H. (2010). Classical and Bayesian estimations on the generalized exponential distribution using censored data, Int. J. Math. Analy.; 4(29), 1417 – 1431.
[14] Mohie El-Din, M.M., Abdel-Aty, Y. and Shafay, A.R. (2011). Two sample Bayesian prediction intervals for order statistics based on the inverse exponential-type distributions using right censored sample, J. of the Egy. Math. Societ.; 19, 102 – 105.
[15] Mohie El-Din, M.M., Abdel-Aty, Y. and Shafay, A.R. (2012). Two sample Bayesian prediction intervals of generalized order statistics based on multiply type-II censored data, Comm. in Statist. Theor. and Method.; 41, 381 – 392.
[16] Panaitescu, E., Popesc, P. G., Cozma, P. and Popa, M. (2006). Bayesian and non-Bayesian estimators using record statistics of the modified-inverse Weibull distribution, Proceed. of the Roman. Acad., Series A; 11(3), 224 – 231.
[17] Pareto, V. (1897). Cours d'Economie Politique, Paris: Rouge et Cie.
[18] Arnold, B. C. and Press, S. J. (1989). Bayesian estimation and prediction for Pareto data, J. of the Amer. Statist. Associat.; 84, 1079 – 1084.
[19] Sadek, A., Sultan, K.S. and Balakrishnan, N. (2009). Bayesian estimation based on ranked set sampling using asymmetric loss function, Bull. Malays. Math. Sci. Soc.
[20] Soliman, A. A., Abd Ellah, A. H. and Sultan, K. S. (2006). Comparison of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches, Comput. Stat. Data Anal.; 51, 2065 – 2077.
[21] Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss function, J. Amer. Statist. Assoc.; 81, 446 – 451.